Frontiers in Artificial Intelligence and Applications

NOVELTIES IN INTELLIGENT DIGITAL SYSTEMS

Proceedings of the 1st International Conference (NIDS 2021), Athens, Greece, September 30 - October 1, 2021

Edited by
Claude Frasson
Katerina Kabassi
Athanasios Voulodimos

IOS Press

NOVELTIES IN INTELLIGENT DIGITAL SYSTEMS

Artificial intelligence and intelligent digital systems have become indispensible to many areas of modern life.

This book presents the proceedings of the 1st International Conference on Novelties in Intelligent Digital Systems (NIDS2021), held in Athens, Greece, from 30 September to 1 October 2021. The conference took place as a virtual event due to COVID-19 restrictions. The NIDS conference lays special emphasis on the novelties of intelligent systems and on the interdisciplinary research which enables, supports, and enhances Artificial Intelligence (AI) in software development. It promotes high-quality research, creating a forum for the exploration of challenges and new advances in AI, and addresses experts, researchers and scholars in the fields of artificial and computational intelligence in systems and in computer sciences in general, enabling them to learn more about pertinent, strongly related and mutually complementary fields. The conference promotes an exchange of ideas, reinforcing and expanding the network of researchers, academics, and market representatives.

The 30 accepted papers included here have each been reviewed rigorously by two or three reviewers through a double-blind process which reflects the commitment of the IIS academic community to make NIDS a top-flight, selective and high-quality conference. They are grouped in 6 sections, and cover the topics of learning; extended reality; data mining and machine learning; health and environment; brain assessment and reasoning; and computer vision

Describing some very significant research and reflecting many interesting new ideas, the book will be of interest to all those working in the field.

9 781643 682044

ISBN 978-1-64368-204-4 (print)
ISBN 978-1-64368-205-1 (online)
ISSN 0922-6389 (print)
ISSN 1879-8314 (online)

NOVELTIES IN INTELLIGENT DIGITAL SYSTEMS

Frontiers in Artificial Intelligence and Applications

The book series Frontiers in Artificial Intelligence and Applications (FAIA) covers all aspects of theoretical and applied Artificial Intelligence research in the form of monographs, selected doctoral dissertations, handbooks and proceedings volumes. The FAIA series contains several sub-series, including 'Information Modelling and Knowledge Bases' and 'Knowledge-Based Intelligent Engineering Systems'. It also includes the biennial European Conference on Artificial Intelligence (ECAI) proceedings volumes, and other EurAI (European Association for Artificial Intelligence, formerly ECCAI) sponsored publications. The series has become a highly visible platform for the publication and dissemination of original research in this field. Volumes are selected for inclusion by an international editorial board of well-known scholars in the field of AI. All contributions to the volumes in the series have been peer reviewed.
The FAIA series is indexed in ACM Digital Library; DBLP; EI Compendex; Google Scholar; Scopus; Web of Science: Conference Proceedings Citation Index – Science (CPCI-S) and Book Citation Index – Science (BKCI-S); Zentralblatt MATH.

Volume 338

Recently published in this series

Vol. 337. H. Fujita and H. Perez-Meana (Eds.), New Trends in Intelligent Software Methodologies, Tools and Techniques – Proceedings of the 20th International Conference on New Trends in Intelligent Software Methodologies, Tools and Techniques (SoMeT_21)
Vol. 336. A. Biere, M. Heule, H. van Maaren and T. Walsh (Eds.), Handbook of Satisfiability – Second Edition
Vol. 335. M. Nørskov, J. Seibt and O.S. Quick (Eds.), Culturally Sustainable Social Robotics – Proceedings of Robophilosophy 2020 – August 18–21, 2020, Aarhus University and online
Vol. 334. S. Villata, J. Harašta and P. Křemen (Eds.), Legal Knowledge and Information Systems – JURIX 2020: The Thirty-third Annual Conference, Brno, Czech Republic, December 9–11, 2020
Vol. 333. M. Tropmann-Frick, B. Thalheim, H. Jaakkola, Y. Kiyokir and N. Yoshida (Eds.), Information Modelling and Knowledge Bases XXXII
Vol. 332. A.J. Tallón-Ballesteros and C.-H. Chen (Eds.), Machine Learning and Artificial Intelligence – Proceedings of MLIS 2020

ISSN 0922-6389 (print)
ISSN 1879-8314 (online)

Novelties in Intelligent Digital Systems

Proceedings of the 1st International Conference (NIDS 2021),
Athens, Greece, September 30 – October 1, 2021

Edited by

Claude Frasson

University of Montreal, Canada

Katerina Kabassi

Ionian University, Greece

and

Athanasios Voulodimos

University of West Attica, Greece

IOS Press

Amsterdam • Berlin • Washington, DC

ISBN 978-1-64368-204-4 (print)
ISBN 978-1-64368-205-1 (online)
Library of Congress Control Number: 2021946231
doi: 10.3233/FAIA338

Publisher
IOS Press BV
Nieuwe Hemweg 6B
1013 BG Amsterdam
Netherlands
fax: +31 20 687 0019
e-mail: order@iospress.nl

For book sales in the USA and Canada:
IOS Press, Inc.
6751 Tepper Drive
Clifton, VA 20124
USA
Tel.: +1 703 830 6300
Fax: +1 703 830 2300
sales@iospress.com

Preface

The 1st International Conference on **Novelties in Intelligent Digital Systems** (**NIDS2021**) was held in Athens, Greece, under the auspices of the Institute of Intelligent Systems (IIS). The Hosting Institution of the conference was the University of West Attica (Greece). NIDS2021 was implemented virtually due to COVID-19 restrictions, on the scheduled dates, that is from September 30 to October 1, 2021.

NIDS lays special emphasis on the novelties of intelligent systems and on interdisciplinary research which enables, supports, and enhances Artificial Intelligence (AI) in software development. It promotes high-quality research, creating a forum for the exploration of challenges and novel advancements in AI.

NIDS addresses experts, researchers and scholars in the fields of artificial and computational intelligence in systems and in computer sciences in general, enabling them to learn more about pertinent, strongly related and mutually complementary fields. It triggers an exchange of ideas, reinforcing and expanding the network of researchers, academics, and market representatives.

Topics within the scope of NIDS series include, but are not limited to:

Adaptive Systems
Affective Computing
Augmented Reality
Big Data
Bioinformatics
Cloud Computing
Cognitive Systems
Collaborative Learning
Cybersecurity
Data Analytics
Data Mining and Knowledge Extraction
Decision Making Systems
Deep Learning
Digital Marketing
Digital Technology
Distance Learning
E-Commerce
Educational Data Mining
E-Learning
Environmental Informatics
Expert Systems
Fuzzy Systems
Genetic Algorithm Applications
Human-Machine Interaction
Information Retrieval
Intelligent Information Systems
Intelligent Modeling
Machine Learning
Medical Informatics
Mobile Computing
Multi-Agent Systems
Natural Language Processing
Neural Networks
Pattern Recognition
Personalized Systems and Services
Pervasive Multimedia Systems
Recommender Systems
Reinforcement Learning
Semantic Web Applications
Sentiment Analysis
Serious Gaming
Smart Cities
Smart Grid
Social Media Applications
Social Network Analytics
Text Mining
Ubiquitous Computing
User Modelling
Virtual Reality
Web Intelligence

The Call for Scientific Papers solicited work presenting substantive new research results in using advanced computer technologies and interdisciplinary research for enabling, supporting and enhancing intelligent systems. The Posters Track provided an interactive forum for authors to present research prototypes and work in progress to the conference participants.

The international Program Committee consisted of 50 leading members of the Intelligent Systems community, as well as of highly promising younger researchers. The Conference (General) Chairs were Cleo Sgouropoulou and Ioannis Voyatzis from the University of West Attica (Greece), whereas the Program Committee Chairs were Claude Frasson from the University of Montreal (Canada), Katerina Kabassi from Ionian University (Greece), and Athanasios Voulodimos from the University of West Attica (Greece).

Scientific papers were reviewed rigorously by two and in some cases by three reviewers (one of which was senior) through a double-blind process, thus reflecting the commitment of the IIS academic community to make NIDS a top-flight, selective, high-quality conference. We believe that the chosen full papers describe some very significant research, the short papers reflect some very interesting new ideas, while the posters present research in progress that deserves close attention. In the review process, the reviewers' evaluations were generally respected. The management of the review process and the preparation of the proceedings were handled through EasyChair.

We would like to thank all those who have contributed to the conference: the authors, the Program Committee members and the Organization Committee with its chair, Kitty Panourgia, as well as the Institute of Intelligent Systems. Special thanks to Christos Troussas and Akrivi Krouska to have launched and carefully followed this conference.

We are also thankful to our conference sponsor, the Entropy journal (MDPI) for their financial contribution to the Best Paper Award.

Claude Frasson
Katerina Kabassi
Athanasios Voulodimos

Organization

Conference Committee

General Conference Chairs

Cleo Sgouropoulou	University of West Attica, Greece
Ioannis Voyatzis	University of West Attica, Greece

Honorary Chair

Riichiro Mizoguchi	JAIST, Japan

Program Committee Chairs

Claude Frasson	University of Montreal, Canada
Aikaterini Kabassi	Ionian University, Greece
Athanasios Voulodimos	University of West Attica, Greece

Program Advising Chairs

Jaime Caro	University of the Philippines
George Magoulas	University of London, UK

Workshop and Tutorial Chair

Andreas Floros	Ionian University, Greece

Posters and Demos Chair

Kyparisia Papanikolaou	ASPETE, Greece

Doctoral Consortium Chair

Lanndon A. Ocampo	CTU, Philippines

Organization Chair

Kitty Panourgia	Neoanalysis Ltd, Greece

Publicity Chair

Christos Troussas	University of West Attica, Greece

Sponsorship & Exhibition Chair

Akrivi Krouska	University of West Attica, Greece

Program Committee

Ali Abd Almisreb	IUS, Bosnia and Herzegovina
Shahzad Ashraf	Hohai University, China
Abul K.M. Azad	Northern Illinois University, USA
Costin Badica	University of Craiova, Romania
Maumita Bhattacharya	Charles Sturt University, Australia
Siddhartha Bhattacharyya	Christ University, India
Karima Boussaha	University of Oum El Bouaghi, Algeria
Ivo Bukovsky	CTU, Czech Republic
George Caridakis	University of the Aegean, Greece

Jaime Caro	University of the Philippines, Philippines
Adriana Coroiu	Babeș-Bolyai University, Romania
Athanasios Daradoumis	University of the Aegean, Greece
Samia Drissi	University of Souk Ahras, Algeria
Eduard Edelhauser	University of Petrosani, Romania
Kurt Junshean Espinosa	University of Manchester, UK
Andreas Floros	Ionian University, Greece
Foteini Fotopoulou	University of Patras, Greece
Claude Frasson	University of Montreal, Canada
Peter Hajek	University of Pardubice, Czech Republic
Nantia Iakovidou	King’s College London, UK
Katerina Kabassi	Ionian University, Greece
Athanasios Kakarountas	University of Thessaly, Greece
Yasushi Kambayashi	Nippon Institute of Technology, Japan
Achilles Kameas	Hellenic Open University, Greece
George Kolezas	NTUA, Greece
Petia Koprinkova-Hristova	Bulgarian Academy of Sciences, Bulgaria
Akrivi Krouska	University of West Attica, Greece
Florin Leon	Technical University of Iasi, Romania
George Magoulas	Birkbeck College, University of London, UK
Phivos Mylonas	Ionian University, Greece
Vasileios Nittas	University of Zurich, Switzerland
Vera Novikova	Tomsk Polytechnic University, Russia
Stavros Ntalampiras	University of Milan, Italy
Lanndon A. Ocampo	Cebu Technological University, Philippines
Kyparisia Papanikolaou	ASPETE, Greece
Nikolaos Polatidis	University of Brighton, UK
Spyros Polykalas	Ionian University, Greece
Theodosios Sapounidis	International Hellenic University, Greece
Filippo Sciarrone	Roma Tre University, Italy
Cleo Sgouropoulou	University of West Attica, Greece
Allan Sioson	Ateneo de Naga University, Philippines
Dimitris Sotiros	WUST, Poland
Antonio Staiano	University of Naples Parthenope, Italy
George Styliaras	University of Patras, Greece
Christos Troussas	University of West Attica, Greece
Panagiotis Vlamos	Ionian University, Greece
Athanasios Voulodimos	University of West Attica, Greece
Ioannis Voyiatzis	University of West Attica, Greece
Davide Zambrano	EPFL, Switzerland

Contents

Preface v
Claude Frasson, Katerina Kabassi and Athanasios Voulodimos

Organization vii

Learning

An Alternative Educational Tool Through Interactive Software over Facebook in the Era of COVID-19 3
Christos Troussas, Akrivi Krouska, Filippos Giannakas, Cleo Sgouropoulou and Ioannis Voyiatzis

Forced ε-Greedy, an Expansion to the ε-Greedy Action Selection Method 12
George Angelopoulos and Dimitris Metafas

A Methodology for Assessing the Impact of Error Components in Gait Analysis Using Closed-Loop Testing on a Biomimetic Rig 21
Christos Kampouris, Philip Azariadis and Vasilis Moulianitis

Development of an Innovative Learning Methodology Aiming to Optimise Learners' Spatial Conception in an Online Mechanical CAD Module During COVID-19 Pandemic 31
Zoe Kanetaki, Constantinos Stergiou, Christos Troussas and Cleo Sgouropoulou

Enhancing the Effectiveness of Intelligent Tutoring Systems Using Adaptation and Cognitive Diagnosis Modeling 40
Akrivi Krouska, Christos Troussas, Filippos Giannakas, Cleo Sgouropoulou and Ioannis Voyiatzis

Ambient Intelligence and Smart Environments: A Preliminary Overview 46
Maria Poli

Extended Reality

Virtual Reality Tool for Rehabilitation of Patients with Parkinson's Disease: A Conceptual Design Review 55
Josiah Cyrus Boque, Jaime Caro, Richelle Anne Juayong and Roland Dominic Jamora

Toward Personalizing Alzheimer's Disease Therapy Using an Intelligent Cognitive Control System 63
Hamdi Ben Abdessalem and Claude Frasson

On the Development of a Personalized Augmented Reality Spatial Ability Training Mobile Application 75
Christos Papakostas, Christos Troussas, Akrivi Krouska and Cleo Sgouropoulou

Zoo Therapy for Alzheimer's Disease with Real-Time Speech Instruction and Neurofeedback System 84
Yan Ai, Hamdi Ben Abdessalem and Claude Frasson

A Framework for Personalized Fully Immersive Virtual Reality Learning Environments with Gamified Design in Education 95
Andreas Marougkas, Christos Troussas, Akrivi Krouska and Cleo Sgouropoulou

A Novel Framework Incorporating Augmented Reality and Pedagogy for Improving Reading Comprehension in Special Education 105
Andrianthi Kapetanaki, Akrivi Krouska, Christos Troussas and Cleo Sgouropoulou

GunitaHu: A VR Serious Game with Montessori Approach for Dementia Patients During COVID-19 111
Aira Jazel Y. Ang, Roberto D. Principio Jr., Richelle Ann B. Juayong and Jaime D.L. Caro

Intelligence and Intelligent Simulation 117
V.Yu. Meytus

Data Mining and Machine Learning

Creating a Metamodel for Predicting Learners' Satisfaction by Utilizing an Educational Information System During COVID-19 Pandemic 127
Zoe Kanetaki, Constantinos Stergiou, Georgios Bekas, Christos Troussas and Cleo Sgouropoulou

Evaluating the Usefulness of Unsupervised Monitoring in Cultural Heritage Monuments 137
Charalampos Zafeiropoulos, Ioannis N. Tzortzis, Ioannis Rallis, Eftychios Protopapadakis, Nikolaos Doulamis and Anastasios Doulamis

Implementing Hierarchical Indoor Semantic Location Identity Classification: A Case Study for COVID-19 Proximity Tracking in the Philippines 147
Irvin Kean Paulus Paderes, Ligayah Leah Figueroa and Rommel Feria

Data Mining for Improving Online Higher Education Amidst COVID-19 Pandemic: A Case Study in the Assessment of Engineering Students 157
Zoe Kanetaki, Constantinos Stergiou, Georgios Bekas, Christos Troussas and Cleo Sgouropoulou

Visual Analytics in Process Mining for Supporting Business Process Improvement 166
Antonia Kaouni, Georgia Theodoropoulou, Alexandros Bousdekis, Athanasios Voulodimos and Georgios Miaoulis

ActiveCrowds: A Human-in-the-Loop Machine Learning Framework 176
Lazaros Toumanidis, Panagiotis Kasnesis, Christos Chatzigeorgiou, Michail Feidakis and Charalampos Patrikakis

A Fuzzy C-Means-Based Algorithm for the Surveillance of Dengue Cases Distribution in Local Communities 185
Jozelle C. Addawe, Jaime D.L. Caro and Richelle Ann B. Juayong

Application Programming Interface for a Customer Experience Analysis Tool 191
George Kopsiaftis, Ioannis Georgoulas, Ioannis Rallis, Ioannis Markoulidakis, Kostis Tzanettis, Michael Sfakianos and Nikolaos Doulamis

Health and Environment

An Introduction to the euPOLIS Project 197
Charalampos Zafeiropoulos, Matthaios Bimpas, Eftychios Protopapadakis, Emmanuel Sardis, Nikolaos Doulamis, Anastasios Doulamis, Cedo Maksimovic, Stanislava Boskovic, Ranko Bozovic and Maja Lalic

Assessing the Lockdown Effects on Air Quality During COVID-19 Era 207
Ioannis Kavouras, Eftychios Protopapadakis, Maria Kaselimi, Emmanuel Sardis and Nikolaos Doulamis

Smart Technologies and the Case of People with Disabilities: A Preliminary Overview 217
Maria Poli

Designing Personalisation in a Environmental Recommendation System: Differences of AHP and Fuzzy AHP 223
Katerina Kabassi

Brain Assessment and Reasoning

Prediction of Amusement Intensity Based on Brain Activity 229
Gabrielle Toupin, Mohamed S. Benlamine and Claude Frasson

Ontology Reasoning for Explanatory Feedback Generation to Teach How Algorithms Work 239
Anton Anikin, Oleg Sychev and Mikhail Denisov

Computer Vision

Real-Time Face Mask Detector Using Convolutional Neural Networks Amidst COVID-19 Pandemic 247
Efstratios Kontellis, Christos Troussas, Akrivi Krouska and Cleo Sgouropoulou

Unsupervised Man Overboard Detection Using Thermal Imagery and Spatiotemporal Autoencoders 256
Nikolaos Bakalos, Iason Katsamenis, Eleni Eirini Karolou and Nikolaos Doulamis

Subject Index 265

Author Index 267

This page intentionally left blank

Learning

This page intentionally left blank

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210069

An Alternative Educational Tool Through Interactive Software over Facebook in the Era of COVID-19

Christos TROUSSAS[1], Akrivi KROUSKA, Filippos GIANNAKAS, Cleo SGOUROPOULOU, Ioannis VOYIATZIS
Department of Informatics and Computer Engineering
University of West Attica, Greece

Abstract. During the COVID-19 period, education has been required to reform and adapt to a modern, imposed paradigm in which all of its activities must be delivered in a completely digital format. As a result, in light of this pressing need, educational institutions should either repurpose existing facilities, such as learning management systems, or develop new online education options in a timely manner. In view of the above, this paper proposes an alternative educational tool over Facebook by exploiting the technology of social networking. This Facebook-based application was used as the sole platform for supporting asynchronous learning of high school students in the tutoring of mathematics during the COVID-19 lockdown. This modern educational approach through Facebook encompasses interactive software which involves an e-class environment with social characteristics, multimedia-based learning material delivery and learning activities and assessment of different types. The presented Facebook-based interactive educational software has been evaluated by school students with very encouraging results.

Keywords. Coronavirus, COVID-19, education, Facebook, online learning, technology, social networks.

1. Introduction

In the era of COVID-19, education has been forced to reform and adapt to a new, imposed reality where all its activity must be conveyed in a fully digital way. As such, in view of this compelling need, education institutes should either use their already existing equipment, namely platforms, or to offer new possibilities of online education within a reasonable time. Until the time the pandemic appeared, education institutes worldwide have mainly incorporated Learning Management Systems (LMSs) that can support online learning either in the form of solely electronic or blended learning. Blended learning combines the traditional classroom learning with online learning, in which students can, in part, control the time, pace, and place of their learning [1].

Indeed, educators report[2], in a percentage of over 90%, that they use one LMS, while in a percentage of over 85% they believe that through LMSs, education can be further enhanced. From students' perspective, they use only the basic features and functionalities

[1] Corresponding Author; E-mail: ctrouss@uniwa.gr.
[2] https://www.educause.edu/ecar

of the LMSs, and dislike the others that require more engagement [2]. The e-lesson is the core component of the LMSs. Each lesson is an autonomous entity on the platform that incorporates a number of subsystems [3]. Essentially, the course is a modular structure that is organized and managed by the responsible instructor, depending on the material and the e-learning model s/he adopts (from a simple informative educational website to a fully dynamic learning environment).

At the same time, social networking sites have evolved thanks to the second generation of the internet (web 2.0) [4,5] and their users can, on the one hand, publish the content they want and, above all, interact with each other. One representative example is Facebook, which has been also used in the educational settings [6]. According to previous studies [7], Facebook has been used for self-directed learning, with the potential for cooperative teaching. This is to be anticipated, given that Facebook's integrated module for offline and instant messaging can facilitate contact between learners and their educators. Researchers [8-10] emphasize the use of Facebook in education because it offers a pleasant interface for promoting learner-computer communication, teamwork, and the exchange of knowledge and materials. As a result, it is crucial that tutors all over the world concentrate on the integration of such technology into education, as well as the aspects that can help with the adoption of educational platforms that result from them. Furthermore, according to the authors of [11], educational institutes should encourage learners to incorporate Facebook into their learning scenario in order to develop their knowledge and capacities. Indeed, learners are increasingly requesting the use of social media practices in their studies.

Concerning the incorporation of Facebook in instructional settings, research efforts [12-20] have mostly highlighted the interaction of learners and instructors in its social media environment; however, so far, instructors have not used it as the sole tool for supporting education, but they believe that it can offer learners a pleasant and effective learning experience.

In view of the above, this paper presents an alternative educational tool through adaptive software over Facebook in the era of COVID-19. More specifically, the presented application has been used as an alternative of LMSs during the imperatively imposed online education as a result of the pandemic. This paper exploits the technology of social networking and proposes an alternative educational approach through interactive software over Facebook. This application was used as the sole platform for asynchronous learning by high school students in the tutoring of mathematics during the COVID-19 lockdown. This modern educational approach through Facebook encompasses interactive software which involves an e-class environment with social characteristics, multimedia-based learning material delivery and learning activities and assessment of different types.

2. Description of modules

This section provides an overview of the system’s main modules, focusing on its user interface, learning material delivery and e-assessment.

2.1. E-class environment

For the environment of the e-class, the Facebook groups are utilized. Facebook groups provide a closed space for small groups of users who can communicate in the contexts of their educational interests. A group can be created to resemble an e-class for teaching a specific domain. Group members can chat, exchange messages, post text, link to websites, videos, photos and photo galleries, create and collaborate on shared text files, add files of various types as well as create events. Also, they can receive a notification when a member makes posts in the group. Such posts can be the upload of learning material or assessment. In terms of privacy, a group can belong to one of the following categories: i) open, ii) closed and iii) secret. In closed and secret groups, on the one hand the new members are accepted or added by the pre-existing ones and on the other hand the posts are visible only to the group members. The difference between closed groups and secret groups is that closed groups are visible to both members and non-members of the group, while secret groups are visible only to group members. Accessing Facebook's open and closed groups is similar to accessing the open and closed courses of e-learning systems.

The collaborative teaching method can be supported by the Facebook social networking program, through the groups that can be created. Students can communicate with their classmates and their teacher and then discuss the relevant information in class. Teachers can post files, activities, presentations, organize events and remind students of various events. All team members can see the posts but also interact. The interaction can be done in two ways:

- with a comment, where they can express their views in text or post relevant links and related photos and videos and
- by adding a reaction to express their dis/agreement with the post. Also, instead of a “Like”, they can react by choosing one of the following: “Love”, “Haha”, “Wow”, “Sad”, “Angry”.

The role of teachers is mainly coordinating-guiding. In web 2.0 applications, the focus is placed on the student and for this reason it is necessary to adjust and differentiate the role of the teacher.

Figure 1. Environment of the e-class.

2.2. Learning material delivery

The learning material is delivered to learners in the form of text, images, videos and audio and external website links (Figure 2). As such, the learners have the opportunity to study better according to the way of learning they prefer. For instance, if they are visual learners, then they can choose to see an image or watch a video. The domain to be taught included mathematical concepts, such as natural, decimal and positive/negative numbers, percentages, fractions and equations.

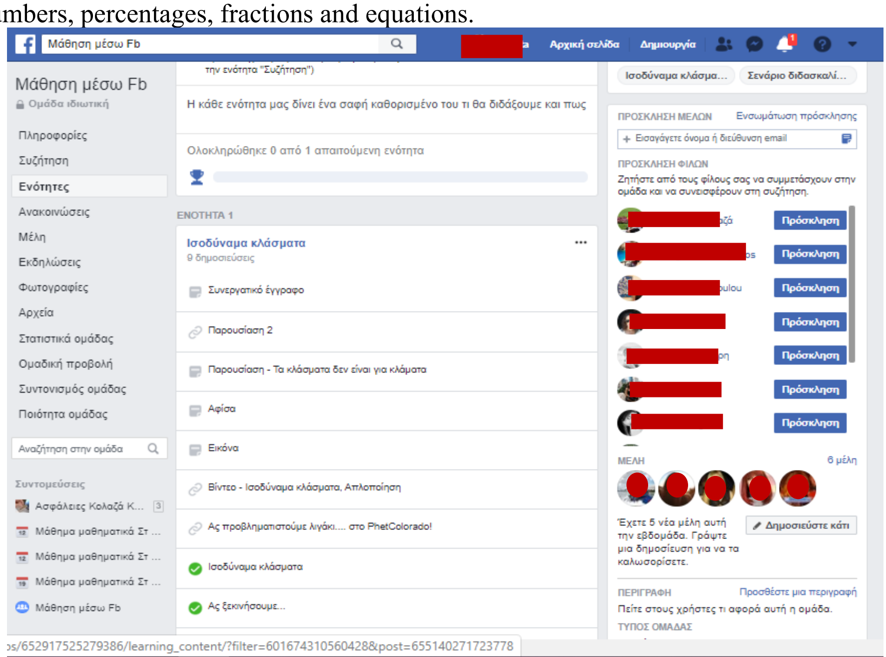


Figure 2. Learning material delivery.

2.3. Learning activities and assessment

Through the platform, the students have the opportunity to assess their knowledge level, which is of utter importance in online learning environments [21,22]. Towards this direction, the Facebook application had multiple quizzes for each one of the chapters being taught as well as revision exercises (Figure 3). The exercises were of several different types, such as multiple choice and file upload. Especially in the multiple-choice questions, the system provided automated feedback to students in order to motivate them towards advancing their knowledge.

The progress of learners is also monitored; with the option of group statistics, the instructor can have an overview about the degree of completion and monitoring of the submitted activities by students. With the option of getting results from the quiz, the instructor has a complete picture of the students' answers, so that the evaluation can be carried out both qualitatively and quantitatively.

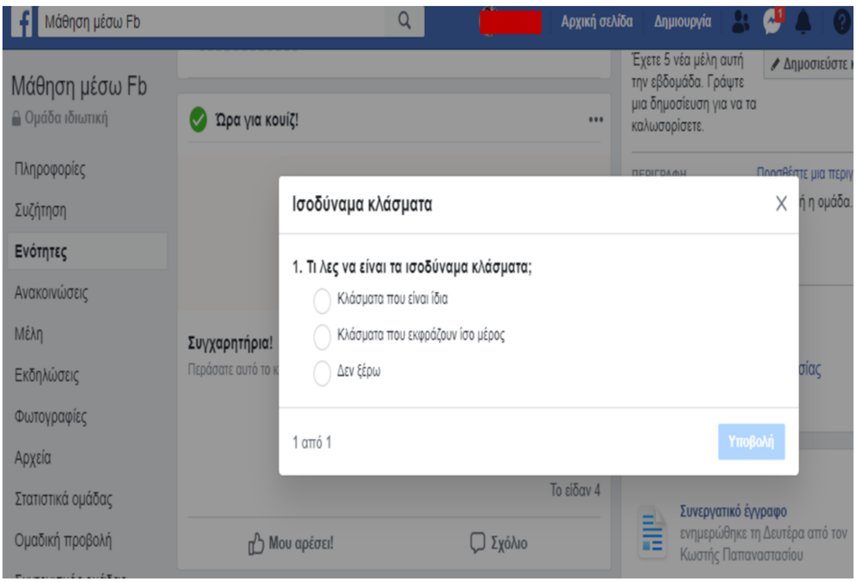


Figure 3. Learning activity.

3. Experimental results and Discussion

This section presents an evaluation of the presented Facebook-based application. The first part of the evaluation involves the perception of learners, while the second part describes the assessment of their knowledge level advancement.

The population of the experiment includes 60 students in the first years of a public high school in the capital city of the country (population A). Also, at the time of the experiment, the students had reached the minimum legal age to have a Facebook account[3].The experiment took place during the COVID-19 lockdown for a period of 4 months. The students have used the presented Facebook application, as the only tool of asynchronous learning, for the following reasons: a) download the learning material, b) communicate with peers/instructors, c) take quizzes, d) ask and resolve possible questions, e) interact in the e-learning environment. Apart from the asynchronous learning offered by the application, the students also had synchronous meeting with the instructors via corresponding tools. The evaluators described the user interface and functionalities of the application to students. After the period of the four months, the participants were asked to answer the following questions (5-Point Likert Scale: 1-Strongly disagree; 2-Disagree; 3-Neither agree nor disagree; 4-Agree; 5-Strongly agree):

- Question 1: Did the software assist you in advancing your knowledge?
- Question 2: Was your stance toward e-learning technology changed as a result of this experience?

The results are presented in Fig. 4. Analyzing the results, it can be inferred that a) students have reached the learning objectives and met the desired outcomes, namely the knowledge acquisition has been achieved in a high rate, b) the type and number of learning activities were appropriate to the knowledge level of students, c) the system is easy to use and its interface is friendly, d) students are very engaged and motivated throughout their interaction with the system.

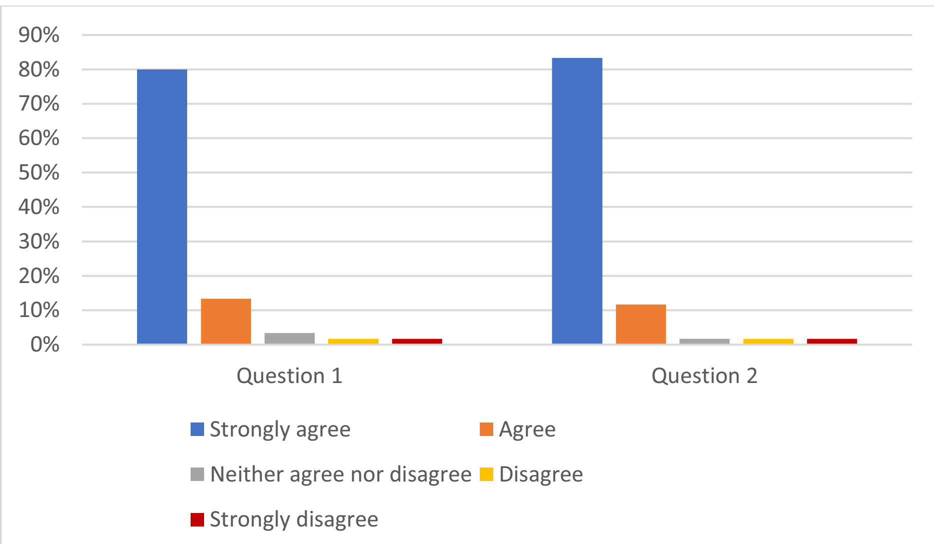


Figure 4. Results.

The second part of the evaluation includes the assessment of the improvement of the students' knowledge level. Using students' t-test, we compared the grades of the 60 students (population A) that used the presented Facebook application with other 60 students (population B) who are also in the first years of the same public high school.

[3] https://www.facebook.com/help/157793540954833

Population B used Edmodo[4], which is an educational platform that refines the concepts of social networking and adapts them for use in a classroom, during the same period of time and for the same reasons, i.e. asynchronous learning. The grades of students were between 0 (lowest) to 20 (highest).

Table 1. Results of t-test.

	Population A	*Population B*
Mean	15,93333333	14,81666667
Variance	7,114124294	7,982768362
Observations	60	60
Hypothesized Mean Difference	0	
df	118	
t Stat	2,226154971	
P(T<=t) one-tail	0,013951352	
t Critical one-tail	1,657869522	
P(T<=t) two-tail	0,027902703	
t Critical two-tail	1,980272249	

The grade averages of the two populations were compared. Table 1 summarizes the findings. According to the findings, the performance of Population A (15.93) outperformed the performance of Population B (14.81). We also note that P(T<=t) two-tail value is less than 0.05. This indicates that the difference in performance between the two means is statistically significant. Based on the t-test results, it can be inferred that the provided scheme, which is a social e-learning application over Facebook, helps the students to enhance their performance on the domain taught.

4. Conclusions

This paper presents a modern way of online education beyond the established LMSs. Students can perceive an alternative way to use Facebook, apart from its social dimension. As presented in this paper, this modern educational approach through Facebook encompasses interactive software which involves an e-class environment with social characteristics, multimedia-based learning material delivery and learning activities and assessment of different types.

Students can interact and communicate with peers, exchange ideas, collaborate and express themselves in a variety of ways. Comments become arguments as students think about the style and way they can comment. This cultivates creativity and the expression of emotions and thoughts. Submitting comments allows students to express their views, whether they agree or disagree, which helps to cultivate the self-awareness of team members. Consequently, smooth socialization is promoted, since student-centered discussions and collaborative learning in groups cultivate responsibility and positive relationships between groups.

4 https://new.edmodo.com/

Through the use of the presented Facebook-based educational tool, students can be more engaged during the educational process. On top of that, their technological skills can be strengthened and expanded. The presented application has been evaluated using questionnaire and t-test with very promising results, showing the acceptance by its users and knowledge advancement of students. Future steps include the exploration of different social media platforms as educational tools. Also, a more extensive evaluation is in our future plans.

References

[1] I. Cabrera, J. Villalon and J. Chavez, Blending Communities and Team-Based Learning in a Programming Course, *IEEE Transactions on Education*, **60**(4) (2017), 288-295.

[2] J. D. Galanek, D. C. Gierdowski, and D. C. Brooks. *ECAR Study of Undergraduate Students and Information Technology*, Research report, Louisville, CO: ECAR, 2018.

[3] S. Dlalisa, Acceptance and usage of learning management system amongst academics, *2017 Conference on Information Communication Technology and Society* (2017), 1-7.

[4] M. Virvou, C. Troussas and E. Alepis, Machine learning for user modeling in a multilingual learning system, *International Conference on Information Society (i-Society 2012)* (2012) 292-297.

[5] M. Virvou, C. Troussas, CAMELL - Towards a Ubiquitous Multilingual e-Learning System, *International Conference on Computer Supported Education (CSEDU)* (2011), 509-513.

[6] C. Troussas, A. Krouska, E. Alepis and M. Virvou. Intelligent and adaptive tutoring through a social network for higher education, *New Review of Hypermedia and Multimedia* (2021). DOI: 10.1080/13614568.2021.1908436

[7] J. Prescott, M. Stodart, G. Becket, and S. Wilson, S. The Experience of using Facebook as an Educational Tool, *HSCE*, **0**(0) (2015), 1-5.

[8] D. Dogan, Y. Gulbahar, Using Facebook as Social Learning Environment, *Informatics in education*, **17**(2), (2018), 207-228.

[9] A. Krouska, C. Troussas and M. Virvou, Social networks as a learning environment: Developed applications and comparative analysis, *2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA)* (2017), 1-6. doi: 10.1109/IISA.2017.8316430.

[10] C. Troussas, M. Virvou, K. Junshean Espinosa, Using Visualization Algorithms for Discovering Patterns in Groups of Users for Tutoring Multiple Languages through Social Networking, *Journal of Networks* **10**(12) (2015), 668-674.

[11] C. Saini, J. Abraham, Implementing Facebook-based instructional approach in pre-service teacher education: An empirical investigation, *Computers & Education* **128** (2019), 243-255.

[12] I. Awidi, M. Paynter, and T. Vujosevic, Facebook group in the learning design of a higher education course: An analysis of factors influencing positive learning experience for students. *Computers & Education* **129** (2019), 106-121.

[13] D. Dogan, and Y. Gulbahar, Using Facebook as Social Learning Environment, *Informatics in Education* **17**(2) (2018), 207-228.

[14] T. Teo, T. Doleck, and P. Bazelais, The role of attachment in Facebook usage: a study of Canadian college students, *Interactive Learning Environments* **26**(2) (2017), 256-272.

[15] J. Wu, Y. Hsiao, and M. Nian, Using supervised machine learning on large-scale online forums to classify course-related Facebook messages in predicting learning achievement within the personal learning environment, *Interactive Learning Environments* **28**(1) (2018), 65-80.

[16] E. Gallardo-Echenique, M. Bullen and A. Castillo, Peruvian undergraduate students' usage of digital technology in academic context, *2020 IEEE World Conference on Engineering Education (EDUNINE)* (2020), 1-5.

[17] N. N. Zulkifli, N. D. Abd Halim, N. Yahaya and H. Van Der Meijden, Patterns of Critical Thinking Processing in Online Reciprocal Peer Tutoring Through Facebook Discussion, *IEEE Access* **8** (2020), 24269-24283.

[18] S. Ramadanty, Y. Safitri and H. Suhendra, Social Media Content Strategy for Higher Education, Indonesia, *2020 IEEE International Conference on Information Management and Technology* (2020), 336-341.

[19] M. Y. Alghamdi, A Study of Social Network Applications in University Education, *2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress* (2020), 138-145.

[20] L. -K. Lee, Y. -C. Fung, Y. -W. Pun, K. -K. Wong, M. T. -Y. Yu and N. -I. Wu, Using a Multiplatform Chatbot as an Online Tutor in a University Course, *2020 IEEE International Symposium on Educational Technology* (2020), 53-56, doi: 10.1109/ISET49818.2020.00021.

[21] C. Troussas, A. Krouska, and M. Virvou, Using a Multi Module Model for Learning Analytics to Predict Learners' Cognitive States and Provide Tailored Learning Pathways and Assessment, *Virvou M., Alepis E., Tsihrintzis G., Jain L. (eds) Machine Learning Paradigms, Intelligent Systems Reference Library* **158** (2020), Springer, Cham. https://doi.org/10.1007/978-3-030-13743-4_2.

[22] K. Chrysafiadi, C. Troussas and M. Virvou, A Framework for Creating Automated Online Adaptive Tests Using Multiple-Criteria Decision Analysis, *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (2018), 226-231. doi: 10.1109/SMC.2018.00049.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210070

Forced ε-Greedy, an Expansion to the ε-Greedy Action Selection Method

George ANGELOPOULOS[a] and Dimitris METAFAS[a]
[a] *Department of Electrical and Electronic Engineering, University of West Attica, Athens Greece*

Abstract. Reinforcement Learning methods such as Q Learning, make use of action selection methods, in order to train an agent to perform a task. As the complexity of the task grows, so does the time required to train the agent. In this paper Q Learning is applied onto the board game Dominion, and Forced ε-greedy, an expansion to the ε-greedy action selection method is introduced. As shown in this paper the Forced ε-greedy method achieves to accelerate the training process and optimize its results, especially as the complexity of the task grows.

Keywords. Q Learning, action selection, ε-greedy, forced ε-greedy

1. Introduction

Reinforcement Learning (RL) methods such as Q Learning, Monte Carlo Trees and deep Q Learning have become popular over the past few years. The goal of RL is to train an agent to make the best decisions in order to complete a specific task, which could be as simple as getting out of a maze in the fastest way possible, or as complex as finding the best strategy in order to win a board game. Many researchers use board games to test new methods of RL. Backgammon [1], the game of go [2], [3], Othello [4], [5], the settlers of Catan [6], [7] and other classic board games, were used to create AI agents that evolved strategies through playing the actual game, instead of their choices being controlled by a handwritten deterministic code. Strategies developed by RL agents can be so successful, that they are being adopted worldwide by human players, as demonstrated by G. Tesauro [1].

Not only must the agent be successful in completing the desired task, but the training process must also be as fast as possible. The choice of the action selection method during the training can be crucial in achieving the best possible results, as well as how fast those results are being achieved. The most popular action selection methods are ε-greedy and softmax [8]. Quite a few attempts have been made in order to improve those methods. Some of them are adaptive ε-greedy [9], [10], temporally-extended ε-greedy [11], and ε-BCM [12], which is a Bayesian Ensemble Approach to ε-greedy.

Some different approaches of RL have been tested on the board game Dominion. Researchers used Monte Carlo trees [13], [14], [15], others combinations of artificial neural networks [16], and even balancing cards sets [17] was researched. In this paper Q Learning is applied on the card game Dominion. Our goal is to train an agent able to oppose and dominate over AI opponents, who use popular strategies developed by human players. We also test a new expansion to ε-greedy method, defined as the Forced ε-greedy, in order to accelerate the training. The agent trained with this new method not only achieves better wining rate, than the one trained with the classic ε-greedy method, but the training process is much faster.

2. Reinforcement Learning

The purpose of reinforcement learning (RL), is to train an agent to perform a task, through interaction with the environment and by using the trial and error method. Firstly we create a virtual episode of the problem, and define the environment as everything else, except for the agent in training. The environment's current state (s) is described by several variables, which are provided to the agent. Then the agent has to make a choice between all possible actions $a(s)$ given the current state of the environment. The choice is made based on a policy $\pi(s, a)$. After the action is performed, the agent receives a numerical reward (r), as well as the environment's new state (s'). The task of the agent is to find an optimal policy $\pi'(s, a)$, in order to maximize the cumulative rewards. To do so the agent, at the beginning of the training, assigns a random value $Q(s, a)$ to every state-action pair, which equals to the expected reward if he takes action a, when the state is s. Every time the agent finds itself to the same state s, taking the same action a, he updates the value of $Q(s, a)$, for that state-action pair. After many iterations, and after passing all possible states, the agent will have an optimal approximation of the true values of the $Q(s, a)$ function. The best policy $\pi'(s, a)$ then is proven [8] to be the greedy policy, in which the agent always selects the action with the highest $Q(s, a)$ value, among all possible actions, in order to maximize the cumulative reward.

2.1. Exploration versus Exploitation

When an agent follows the greedy policy, it is considered an exploitation of current knowledge, in order to maximize the reward. This policy works well when the values of $Q(s, a)$ function are the correct ones. But at the start of the training we assign random values to the $Q(s, a)$ function. If the agent follows the greedy policy during the training, he will never explore some actions, which have lower Q values than others, therefore he may not find the optimal solution to the problem.

During the training process it is better to use a policy that allows the agent to explore alternative actions, even if they have low Q values. It is not wise though to let the agent make random choices, because then the training time would be huge. It would be better for the agent to focus on the actions with higher Q values, in order to get a better approximation of them, but occasionally to explore, in order to find better actions that have been assigned low Q values by chance.

Such action selection methods, which match those criteria, are ε-greedy and softmax. The ε-greedy method states that the agent chooses the greedy action with probability $1 - \varepsilon$, and a random action with probability ε, where ε is a number between zero and one. When ε equals zero the agent becomes greedy, and when ε equals one the agent always chooses random actions. According to the softmax method, the agent chooses an action α with probability $p(\alpha)$ shown in Eq. (1), where τ is a positive parameter called temperature and n the number of possible actions in the current state of the environment. In the limit where $\tau \rightarrow 0$ the agent becomes greedy, while higher temperatures make all the actions equiprobable.

$$p(\alpha) = \frac{e^{Q(a)/\tau}}{\sum_{b=1}^{n} e^{Q(b)/\tau}} \tag{1}$$

2.2. Temporal Difference Learning

Temporal Difference learning (TD) is a method of updating the values of $Q(s, a)$ function. Temporal difference learning methods are divided into on policy (S.A.R.S.A.) and off policy (Q Learning). The difference between them is that in S.A.R.S.A. we estimate the Q values for the policy the agent uses to make choices, while in Q Learning we estimate the Q values for the greedy policy, regardless which policy the agent uses. In both methods, a virtual episode of the process is created, and random values $Q(s, a)$ are set, for every state-action pair. The agent finds themselves at the initial state s_0, and selects the action to take between all possible actions, either by the ε-greedy or by the softmax method.

In S.A.R.S.A., at every time step of the episode the agent finds the environment in a state s and selects an action a, based on the $Q(s, a)$ value. Then the agent receives a reward r as well as the environment's new state s'. The agents then selects another action a' with $Q(s', a')$ value and updates $Q(s, a)$ as shown in Eq. (2), where e is the learning rate and γ is the discount factor. This process is repeated until the episodes terminate. After the simulation of a large number of episodes the estimates of the Q values will converge to the true values.

$$Q(s, a) = Q(s, a) + e \cdot [r + \gamma \cdot Q(s', a') - Q(s, a)] \tag{2}$$

During Q Learning the update of $Q(s, a)$ is quite different. Instead of waiting until the next action a', the agent updates $Q(s, a)$ immediately after taking an action, using the maximum Q value of all the possible actions a' at the new state s', as shown in Eq. (3).

$$Q(s, a) = Q(s, a) + e \cdot [r + \gamma \cdot \max_{a'} Q(s', a') - Q(s, a)] \tag{3}$$

3. Dominion Board Game

Dominion is a card game, consisting of three types of cards, treasure cards, victory cards and kingdom cards. Treasure cards are the currency of the game providing coins when played. There are three kinds of treasure cards, copper that provide one coin, silver that provide two coins and gold that provide three coins. Victory cards cannot be played, but they are worth points, and acquiring as many as possible of them, is the way to win. Estates worth one point, duchies worth two points, provinces worth three points, while curses worth minus one point. There are many different kingdom cards, with each one having a different effect when played.

Each player starts the game with a deck of ten cards, seven coppers and three estates. The remaining treasure and victory cards are placed in piles, and ten piles of kingdom cards are created, each consisting of ten cards. The players draw five cards from their decks, and take turns, starting with the first player. Each turn has three phases, action phase, buy phase and clean-up. At the action phase the active player can play one kingdom card from his hand. Kingdom cards have various effects, as card drawing, providing additional actions or buys, etc. At the buy phase, the active player can buy one card, or more if he has gained additional buys at the action phase, paying the cost of the bought cards by playing treasure cards from his hand. Finally at the clean-up phase, the active player discards his hand, as well as all the cards he played and bought, and draws five new cards from his deck. If the deck runs out of cards, he shuffles the discarded cards to create a new deck. The players keep taking turns until the provinces pile is empty, or three or more other piles are empty, whichever comes first, and the game ends. The player with the most points is the winner.

3.1. JDominion

A training and testing environment was created, written entirely in Java, called JDominion. The application allows the creation of two, three or four player games, and the option to train using the Q Learning method or test the agent. The ten kingdom cards, which were integrated in the application, as well as a small description and their prices are shown in Table 1**.** An AI opponent was created in order to train and test the agent, Money. The Money policy buys only silver and gold treasure cards, until two gold cards are present in the deck, from that point on it is also possible to buy victory cards. The money policy was also used by Winder [16], to train and test his agent.

Table 1. Kingdom Cards

Kingdom Card	Description	Price
Festival	+2 actions, +1 buy, +2 coins	5
Chapel	Thrash up to four cards from your hand	2
Bazaar	+2 actions, +1 card, +1 coin	5
Adventurer	Reveal cards from your deck, until you draw two treasure cards	6
Conspirator	+2 coins, if you played three or more actions +1 card, +1 action	4
Smithy	+3 cards	4
Moneylender	Thrash a copper from your hand, if you do +3 coins	4
Village	+1 card, +2 actions	3
Woodcutter	+1 buy, +2 coins	3
Workers village	+1 card, +2 actions, +1 buy	4

The JDominion application, not only trains the agent, but also tests the efficiency of the trained agent, alongside with the training. After every 10.000 games, the training process stops temporarily, and 10.000 testing games are conducted, using the same number of players as the training games. During the testing games the agent uses the greedy policy, according to how the values of the Q function have shaped so far. The results of the training and the testing games are stored, as well as various other statistics, like the average number of each card bought during the game, the maximum and the minimum number of each card bought etc. The application also stores the current values of the Q function, the values of the Q function for which the testing games had the best results, and the number of times the agent has found himself to every specific state-action pair during the training. The JDominion application can be found online at https://sourceforge.net/projects/jdominion/files/.

4. Results

To define the state of the environment several variables were used as shown in Table 2. For every variable we used an upper bound, in order to diminish the total number of different state-actions pairs. Two sets of training games were conducted, the first with the agent against one AI opponent (two players game) and the second with the agent against three AI opponents (four players games).That is because in four player games the provinces pile, tend to exhaust faster, causing the game to end in a lower number of turns than the two players games. So the strategy the agent evolved in the slow two player games, isn't proved to be successful in the fast four player games, and vice versa. Also we chose different upper bound for four players games and for two players games .The upper bound in each case, was chosen as the rounded up average of every variable, which we found during preliminary training games [18].

Table 2 Variables used to describe the environment

	Four players games		**Two players games**	
Variables	**Upper bound**	**Number of states**	**Upper bound**	**Number of states**
Number of turns	*12*	*13*	*19*	*20*
Copper cards in deck	*8*	*2*	*8*	*2*
Silver cards in deck	*4*	*5*	*6*	*7*
Gold cards in deck	*2*	*3*	*4*	*5*
Festivals in deck	*1*	*2*	*1*	*2*
Chapels in deck	*1*	*2*	*1*	*2*
Bazaars in deck	*1*	*2*	*1*	*2*
Adventurers in deck	*1*	*2*	*1*	*2*
Conspirators in deck	*1*	*2*	*1*	*2*
Smithies in deck	*1*	*2*	*2*	*3*
Moneylenders in deck	*1*	*2*	*1*	*2*
Villages in deck	*1*	*2*	*1*	*2*
Woodcutters in deck	*1*	*2*	*1*	*2*
Workers villages in deck	*1*	*2*	*1*	*2*
Possible buys		*18*		*18*
Total number of states		*7.188.480*		*38.707.200*

The purpose of the training was for the agent to learn the best card to buy in every round, in order to win the game. So with 18 different choices in every round, the total number of state-action pairs was 7,188,480 for four player games, and 38,707,200 for two player games.

4.1. Forced Exploration

The action selection method used during both sets of training was ε-greedy, with $\varepsilon = 0.2$, the learning rate $e = 0.2$ and the discount factor $\gamma = 0.95$, which are the recommended values by R. Sutton [3]. Training games with different values for these parameters were conducted, but the results were worst, so we fixed their values as above. Although the total number of different state-action pairs was quite large in both cases, we found that the agent only explored only a small portion of them. In the case of four player games, after 10,000,000 training games, the agent explored 27.2% of the state-action pairs (1,958,822 of 7,188,480) and only in 11.1% (800,087 of 7,188,480) he chose the action, given the state, more than 10 times. In the case of two player games, after 15,000,000 training games, the corresponding numbers are 24.1% (9,358,686 of 38,707,200) and 8.4% (3,265,873 of 38,707,200).

In order to force the agent to explore more, we created an expansion to the ε-greedy method, that we called Forced ε-greedy. To apply this method, we kept track of how many times during the training the agent visited every state-action pair, and we named every visit a "pass". Before the agent select the actions, he checks the passes of every possible action he can take. If there is an action with less than 10 passes, he selects that action. If there are more than one actions with 10 or less passes, he selects the one with the least passes, and if two or more have the least passes he selects one at random. Finally if all the actions have 10 passes or more he selects his next action with the ε-greedy action selection method. The training for both two and four player games was repeated, using the forced ε-greedy method. We slightly modified the Forced ε-greedy method to prefer actions with less than 5 passes instead of 10 in the case of four player games, due to the smaller total number of states. As we can see in Table 3, we forced the agent to explore significantly more in both cases.

Table 3. State-actions pairs explored for classic and forced ε-greedy

	Two players games		**Four players games**	
	Classic ε-greedy	**Forced ε-greedy**	**Classic ε-greedy**	**Forced ε-greedy**
Total Number of states	7,188,480	7,188,480	38,707,200	38,707,200
Number of states explored	1,958,822	2,830,574	9,358,686	15,009,555
Number of states explored (%)	27.2	39.3	24.1	38.7
Number of states with more than 10 passes	800,087	1,213,700	3,265,873	5,702,073
Number of states with more than 10 passes (%)	11.1	16.8	8.4	14.7

In order to verify our results the training for both methods was repeated, for both two and four player games, ten times. In Figure 1 we can see the average winning rate of the agent (5 point moving average) for the four player games, for both classic and forced ε-greedy. The forced ε-greedy method has slightly better results, as the agent reaches the average winning rate, earlier than the agent trained with the classic ε-greedy method.

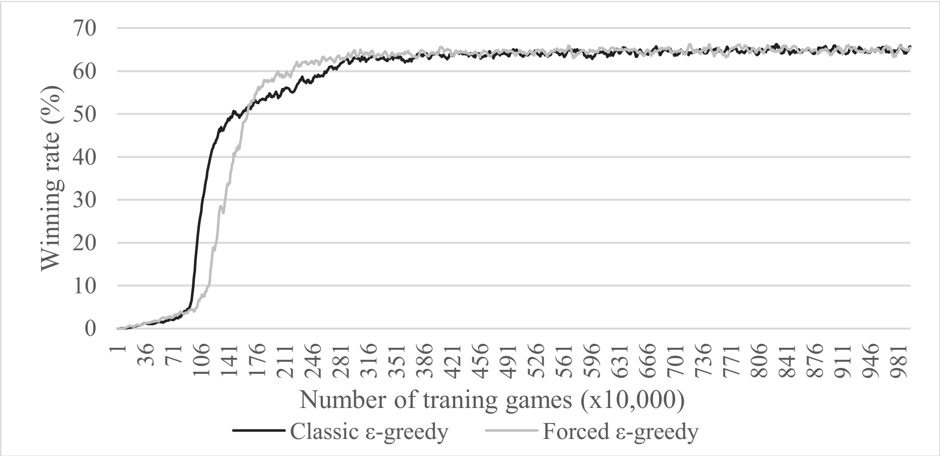


Figure 1. Average winning of Classic and Forced ε-greedy (4 player games)

In Figure 2 we can see the average winning rate of the agent (5 point moving average) for the two player games this time, for both classic and forced ε-greedy. The agent trained with the forced ε-greedy method, not only is faster, but his average winning rate is better, which suggest that he found a better strategy, than the agent trained with the classic ε-greedy method. The large number of total states, clearly favors the forced ε-greedy method.

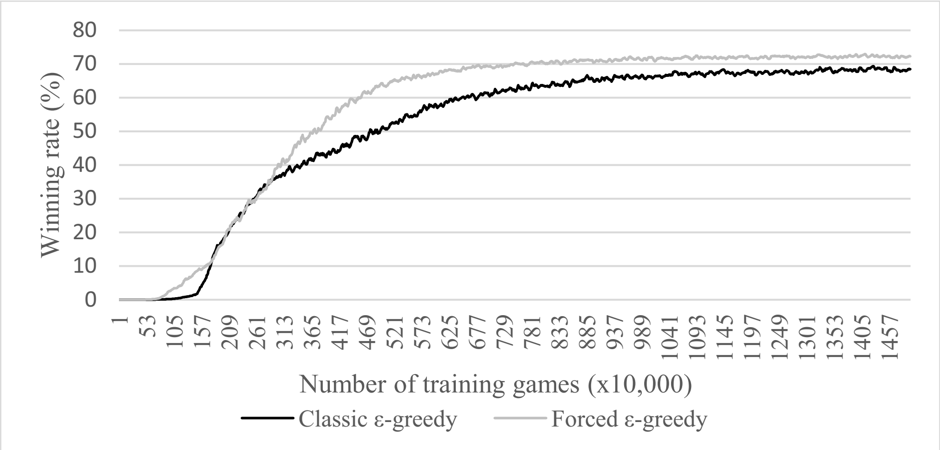


Figure 2. Average winning of Classic and Forced ε-greedy (2 player games)

In the case of the two player games, the success of the trained agent is slightly better than the one trained by Winder [16], who also tested his agent against the money policy, and found that the agent wining rate was 74.7%. The agent trained with the Forced ε-greedy method had average wining rate of 76.9% (the best result found during training).

5. Discussion and Future Work

In Q Learning, it is crucial to give the agent the option to explore new states, in order to find the optimal solution for the task at hand. Given the update algorithm of the Q values for every state-action pair, it is clear that with every update we get a better approximation of the true Q value. With the classic ε-greedy method, the agent only by chance selects low Q value actions. If the number of the total state-action pair is small, ant the number of training games large, we can be quite sure, that most of the possible actions, will get selected, and the agent will find the optimal solution. But in more complex environments, where the number of states is huge, the ε-greedy method can overlook some actions, causing the training process to be slower and possibly miss the optimal solution. In such environments, forcing the agent to explore actions with initial low Q values, which were assigned at random, will lead to better approximations for those values, causing the higher of them to be chosen again and the training process to present faster and better results.

The Forced exploration is not a new action selection method, but merely an expansion to the ε-greedy method. Other actions selection methods, as softmax, can also be expanded the in same way. It is also our opinion that the forced exploration could find an application in Deep Q Learning methods, involving neural networks.

References

[1] Gerald Tesauro. 1994. TD-Gammon, a self-teaching backgammon program, achieves master-level play. Neural computation 6, 5 (March 1994), 215-219. https://doi.org/10.1162/neco.1994.6.2.215

[2] Nicol Schraudolph, Peter Dayan, and Terrence Sejnowski. 1994. Temporal difference learning of position evaluation in the game of Go. In Proceedings of the Advances in Neural Information Processing Systems. 6. 817-824

[3] David Silver, Richard Sutton, and Martin Müller. 2007. Reinforcement Learning of Local Shape in the Game of Go. In Proceedings of the 20th international joint conference on Artificial intelligence. 7. 1053-1058

[4] Michiel Van Der Ree, and Marco Wiering. 2013. Reinforcement learning in the game of Othello: Learning against a fixed opponent and learning from self-play. In Proceedings of the IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning. 108-115

[5] Nees Jan van Eck, and Michiel van Wezel. 2008. Application of reinforcement learning to the game of Othello. Computers \& Operations Research 35, 6 (June 2008), 1999-2017. https://doi.org/10.1016/j.cor.2006.10.004

[6] Stván Szita, Guillaume Chaslot, and Pieter Spronck. 2009. Monte-carlo tree search in settlers of catan. In Proceedings of the Advances in Computer Games. 21-32

[7] Michael Pfeiffer. Reinforcement learning of strategies for Settlers of Catan. 2004. in Proceedings of the International Conference on Computer Games: Artificial Intelligence, Design and Education

[8] Richard Sutton, and Andrew Barto. 1998. Reinforcement Learning: An Introduction. Cambridge, UK: MIT Press

[9] Tokic M. (2010) Adaptive ε-Greedy Exploration in Reinforcement Learning Based on Value Differences. In: Dillmann R., Beyerer J., Hanebeck U.D., Schultz T. (eds) KI 2010: Advances in Artificial Intelligence. KI 2010. Lecture Notes in Computer Science, vol 6359. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-16111-7_23

[10] Dos Santos Mignon, Alexandre, and Ricardo Luis de Azevedo da Rocha. "An Adaptive Implementation of ε-Greedy in Reinforcement Learning." Procedia Computer Science 109 (2017): 1146-1151.
[11] Dabney, Will, Georg Ostrovski, and André Barreto. "Temporally-extended {\epsilon}-greedy exploration." arXiv preprint arXiv: 2006.01782 (2020).
[12] Gimelfarb, Michael, Scott Sanner, and Chi-Guhn Lee. "{\epsilon}-BMC: A Bayesian Ensemble Approach to Epsilon-Greedy Exploration in Model-Free Reinforcement Learning." arXiv preprint arXiv: 2007.00869 (2020).
[13] Rasmus Bille Fynbo, and Christian Sinding Nellemann. 2010. Developing an agent for dominion using modern ai-approaches. PhD Thesis, M. Sc. IT, Media Technology and Games (MTG-T) Center for Computer Games Research
[14] Robin Tollisen, Jon Vegard Jansen, Morten Goodwin, and Sondre Glimsdal. 2015. AIs for Dominion Using Monte-Carlo Tree Search. In Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. 43-52
[15] Sondre Glimsdal. 2015. AIs for Dominion Using Monte-Carlo Tree Search. In Proceedings of the Current Approaches in Applied Artificial Intelligence: 28th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. 43
[16] Ransom Winder. 2014. Methods for approximating value functions for the Dominion card game. Evolutionary Intelligence 6, 4 (March 2014). 195–204. https://doi.org/10.1007/s12065-013-0096-9
[17] Tobias Mahlmann, Julian Togelius, and Georgios N. Yannakakis. 2012. Evolving card sets towards balancing dominion. In Proceedings of the 2012 IEEE Congress on Evolutionary Computation. 1-8
[18] George Angelopoulos and Dimitris Metafas. 2020. Q Learning applied on the Board Game Dominion. In Proceedings of the 2020 24th Pan-Hellenic Conference on Informatics (PCI 2020). 34–37. https://doi.org/10.1145/3437120.3437269

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210071

A Methodology for Assessing the Impact of Error Components in Gait Analysis Using Closed-Loop Testing on a Biomimetic Rig

Christos KAMPOURIS [a,b,1], Philip AZARIADIS [a] and Vasilis MOULIANITIS [a]
[a] *University of the Aegean*
[b] *University of West Attica*

Abstract. Scientific gait analysis methods aim to offer objective measurements, to assist physicians towards an accurate diagnosis or pre-diagnosis of ailments before they actually manifest through noticeable symptoms. This paper reviews selected gait analysis system technologies, trends, applications and discusses errors and precision in spatial and angular readings. Furthermore, we propose a novel test and calibration method using a biomimetic rig. To illustrate this, we conduct three tests on an optical single-camera gait analysis system based on a mobile android smart-phone with specially developed software.

Keywords. gait analysis, biomimetic rig, exoskeleton systems

1. Introduction

The multiple applications of human gait analysis can be categorized in eight main fields: health, sports, entertainment, education, ergonomics, exoskeletons, robotics and security, with health being the most important. The global ageing tendency and the associated ailments (e.g., Parkinson's and osteoarthritis) amplify motorial problems dramatically. In addition, increased traffic accidents, obesity, mass sporting activities, orthopaedic surgeries, urge for solutions in compensating incorrect gait cycle. Exoskeleton mechanisms for the human limbs can help alleviate such problems. The design procedure of safe exoskeleton mechanisms (including prototyping, programming, debugging, calibrating, testing, foolproofing) necessitates the study of a significant number of gait cycles derived from numerous individuals both patients and healthy volunteers.

Each human being has its own distinctive gait, or, rather, set of gaits, adapting to speed, terrain, carried load, fatigue, etc. Individual gait patterns change due to age, occupation, working conditions, sport activities, hobbies, life style, health issues. Various abnormalities, asymmetries of the gait cycle, or significant deviations from the standards, are indicative of possible underlying pathological conditions that can help doctors focus on specific ailments or conduct prediagnosis [1]. Benchmarking, i.e., comparison with deviation analysis, against a previously executed gait analysis at a

[1]Kampouris Christos, University of the Aegean, School of Engineering, Department of Product and Systems Design Engineering, Konstantinoupoleos 1, Ermoupolis, 84100, Syros, Cyclades, Greece; E-mail: cab@aegean.gr

younger age, or against standard (typical) healthy population values, may reveal heart diseases [2], imminent cardioarrest, Parkinson's, and other syndromes including myopathies [3]. Gait analysis is the systematic study of human locomotion objectively (using instruments to record physical measurements), subjectively (based on the observation of experienced physicians), or semi-subjectively. Numerous protocols cover both objective and semi-subjective gait measuring techniques, by recording and analyzing measurable gait parameters such as: step length, walking speed, swing and stance durations, joint angles, muscles force, left and right asymmetries. For example, Extra Laboratory Gait Assessment Method evaluates gait parameters such as step length, speed and head turning ability while walking, in order to identify risk factors for falls among the elderly at home [4]. Miniaturized sensors and mechatronic devices for gait study can be categorized in three main groups:

- Contact Wearables, including in-shoe and in-sole systems, body sensors such as accelerometers, gyroscopes, magnetometers, extensometers, goniometers, EMG electromyography, EEG encephalography, etc., allow the subject to walk freely in an uncontrolled environment [5]. Combined study of multiple sensor data is a common method [6].
- Contact Non-Wearables, e.g., pressure measuring platforms such as forceplates and walkways with capacitive sensors, piezoelectric, piezoresistive and ultrasonic sensors, can be significantly more precise and accurate than wearables, but also more expensive, typically designed for lab and clinical experiments, constraining motion to a few square meters or less.
- Non Contact, image recording and processing systems with analog or digital cameras, laser range scanners, infrared sensors, active or passive markers, time-of-flight cameras and usual or structured lighting.

In order to identify subtle gait differences, the measurement system should be of known accuracy. Preferably, the error components should be an order of magnitude smaller than the characteristics to be detected. In general, gait measurement errors are induced by instruments, software, physical phenomena, protocols, practitioners, subjects. Numerous papers deal with the precision of gait analysis systems and the validity of the data obtained, aiming at improving overall accuracy in diagnosis.

In the following section (No2), we discuss about error components, accepted limits, and comment on objectives and research questions of the current paper. Additionally an alternative methodology for evaluating gait analysis systems, employing a biomimetic rig, without resulting to Golden standard systems, is proposed in section 3. The use of the rig (subsection 3.2) offers two major benefits, firstly, the exclusion of human volunteers from the experiments, and secondly, the exclusion of human error factors in gait repetition. Illustrated test examples of the methodology are presented in section 4. They can check hypotheses such as "the particular gait analysis system has the required knee angle measurement precision for the diagnosis of that specific ailment". The hypothesis is discussed in last paragraph of section 4. Discussion section (No 5), reviews presented methodology and the contributions of the current paper.

2. Related Work

Modern multi-camera systems can offer 3D markers' coordinate accuracies better than 0.03mm. Consequently, other error factors become prevalent in lab setups. The most crucial factor affecting inferred joint coordinates and infusing errors - often beyond the acceptability of the gait analysis community - is the misplacement of markers on the body of the subject. Misplacement (e.g., due to different protocols / practices among physicians, obesity, subject's morphology / wounds / cooperation spirit) may account for an angle error exceeding the 5° acceptance limit which is what a trained eye is likely to differentiate [1]. The study of 10 children with cerebral palsy against 10 aged-matched typical developing children, showed that anterior-posterior misplacement of the lateral epicondyle marker, led to hip internal-external rotation angle offsets of 5.3° per 10mm of displacement [7]. Determination of the ankle's internal-external rotation angle, demonstrated a sensitivity of 4.4° per 10mm offset. Naturally, measurements tend to worsen according to the ratio of the misplacement to leg length: the smaller the length, the bigger the angle error. It was concluded that in order to achieve an error below the limit of 5° on all joints, a physician needs a repeatability precision below 1.2% of the leg length when placing the markers [7].

In some cases, the disparity of angle measurements between labs reached 34° [8]. It can be attributed mainly to different testers and the plurality of marker placement protocols.

Based on the reported experiments conducted in swimming pool, with calm water [9], a two action cameras setup, with extensive calibration and non-linear optical distortion model, yielded reconstruction accuracy of 1.5mm at highres mode 1920x1080p and 2.5mm at lowres 1280x720p. The work volume was 1x4x1.5m and cameras were placed 1m away. Linear camera model increased the error up to 10mm.

A single RGB-D camera (Microsoft Kinect V2, 2.5D color +infrared camera) was used for the gait analysis of 20 subjects [10]. Machine learning algorithms processed the data and achieved a step-length mean absolute error of 42mm (with standard deviation of 42mm) when walking towards the camera, compared against the gold standard Qualisys system with 12 cameras laboratory setup.

Two versions of Kinect (V1 & V2) were compared, against a laboratory motion capture system [11]. Joints coordinate errors ranged from 50 to 100mm and varied according to distance from center of camera.

The MO^2CA single iPhone camera system was compared to the golden standard Qualisys with 8 cameras setup [12]. Although spatial MO^2CA measurements had an error up to 10mm, a non-inferiority statistical test showed that regarding stride length, stride time, stride length variability, stride time variability, MO^2CA was not inferior to multi-camera Qualisys.

According to the DMS method [13], a single camera setup with multiple markers on joints and head could reconstruct 3D coordinates with maximum linear displacement error of 77mm, with averages from 4 to 33 mm. Inferred joint angles maximum error is 38°, with averages from 2.4 to 11.6°.

In more complex environments, physical phenomena such as water turbulence and air bubbles during swimming [11], or snow spraying in skiing activities [14], can blind optical equipment leaving large gaps of unmeasured track.

Other factors such as subject's speed do induce errors well above standard stance, that in addition to soft tissue/suit artifacts, could result in a total measurement of up to 8.3 +/- 7.1mm, with absolute maximum values being several times higher [14].

Synopsizing, although Golden standard systems, for clinical and lab use, achieve sub-millimeter accuracy, yet, reports from labs with such systems still contain significant measurement errors, mainly due to personnel, protocols and less-than-ideal conditions.

Furthermore, there is increased interest in new low-cost and portable gait analysis systems [15] which are now forming a new market trend. These systems are typically benchmarked with the aid of the golden standard systems [16]. There is also increased interest in new methods that could assess the accuracy of such portable gait systems, at a lower cost, or in a portable manner, especially for systems located in distant areas, away from the ease of reach of Golden systems. Also, another missing ingredient towards the evaluation, is a way to benchmark and calibrate such systems by excluding the human factor noise from the human gait repetition. Such a method is to be presented in the following section.

3. Proposed Methodology

In order to assess any gait analysis system, we have it record and analyze the gait of a biomimetic rig. Afterwards, we compare the results against the known properties of the emulated gait and determine the intra-equipment variability. The rig manages to exclude the inherent human intra-subject variability. This section presents examples of open (A,B) and closed-loop testing (C). The scope of the paper is not to assess the particular optical gait analysis system under test, but to demonstrate that induced errors are easily visualized and identified and secondly present the utility of the proposed methodology. Such tests, summarized in figure 1, could be conducted after a gait system's calibration at the initial setup, and periodically later on. They form a procedure that could assess/enhance the manufacturers' calibration and the users' efficiency.

	Test phases	Test A	Test B	Test C
1	Select location (plain background, without moving objects)	√	√	√
2	Setup treadmill (for easy access and good side-view)			√
3	Fix phone / tablet on tripod / base for a perpendicular side-view	√	√	√
4	Inspect scene through camera and focus frame on the desired area	√	√	√
5	Setup lights. Ensure uniform lighting, without shadows/bright spots.	√	√	√
6	Brief subject (volunteer) on the experiment to be conducted			√
7	Get subject's written approval / consent			√
8	Attach markers on human subject			√
9	Start recording			√
10	Conduct experiment (stand / walk / jog / run)			√
11	Stop recording (pause / resume / stop)			√
12	Extract data from device for further analysis			√
13	Remove targets from subject			√
14	Attach markers on rig	√	√	√
15	Download data to rig	√	√	√
16	Start recording	√	√	√
17	Start rig motion / set to desired pose	√	√	√
18	Stop recording	√	√	√
19	Extract data from device for further analysis	√	√	√
20	Remove targets from rig	√	√	√
21	End session	√	√	√

Figure 1. Work flowchart for Tests A, B and C

It is apparent that not all tests include the same phases. For example, in test A and B there is no need for a human volunteer, so human related phases are omitted.

Test A, defines the absolute minimum joint angle error. Note that, for optical systems, this error varies volumetrically, i.e. it is not the same in all points of the observed volume. The rig can assist in precise volumetric error mapping, thus help calibrate the measurements inside the working envelope. Furthermore, the same procedure can reveal length estimation errors. Both tests B & C are dynamic. If the errors observed are significantly higher than test A, then further investigation could reveal a problematic algorithm in the gait capture system, e.g., look-ahead sampling. Besides positional and angular errors, dynamic tests can reveal timing errors, e.g., in stride time and the efficiency of critical software algorithms e.g., gait cycle detection. Given the assessed system's precision, one can decide if specific clinical gait measurements can be reliably performed. The following subsections present basic information for two of the main components utilized in the tests.

3.1. Gait Analysis System Under Test

For illustration purposes, the Device Under Test is an optical, single-camera gait analysis system based on a smart-phone with specially developed software. The software identifies colored markers placed on hip, knee, ankle, captures their coordinates, and records them in a tabular file. Note that the camera is not calibrated, to demonstrate error factors. Our gait-capture application runs on android platform using OpenCV4.0, and records plane coordinates and video to file. New marker colors and max-min marker sizes can be taught-in at any time. Each frame is time-stamped using the system clock. Through scaling, the captured marker coordinates are transformed from pixel units to millimeters. Offset detection algorithms transform human gait cycle to rig data. For example, a human subject walking on a steady-speed treadmill is bound to infuse periodic horizontal axis displacement at his joints coordinates, as he can't achieve an absolute steady pace. Similarly, a treadmill with cushioning, induces vertical axis periodic displacements. Such "noise" is detected and excluded from the data downloaded to the rig.

3.2. Biomimetic Rig

The rig bears one or two independent limbs, with four motorized orthogonal linear axes, two for each limb, controlled by a microprocessor, within a volume of 160x120x75cm. It is part of an ongoing study for the design and evaluation of various knee exoskeletical mechanisms. The limbs are interchangeable and length adjustable in order to match different physiologies. Each has three rotating joints: hip, knee, angle, and can be mounted on the rig at various orientations. The limbs are pathetic without motors or actuators, as they are designed to accept exoskeleton mechanisms for study. At this set of experiments, the specific limbs' knee joints are rotational, although, other sets of limbs could be used depending on the experiments' specifications, e.g. typical four ~ six bar mechanisms. The current rig setup focuses on knee flexion/extension angle (between femur and tibia) measurements. The rig emulates walking cycles or other sequences e.g., stand & walk, jog & run, squats, sitting & standing repetitive cycles, offering major advantages versus human subjects: memory, repeatability, stamina, adaptability, controlled variability. It can perform the same squat thousands of times, so that a gait analysis system can record it from various angles, distances, lighting conditions.

4. Evaluation

4.1. Test A

It is designed to reveal static divergence, suitable for optical but not for inertial measurements. The setup uses:

1. A test limb that is adjusted to specific, known measurements (ankle-knee distance, hip-knee distance).
2. The biomimetic rig to drive the test limb to a specific, known angle (90^o ankle-knee-hip angle).
3. The optical gait measurement system to capture the rig's stance for a few seconds. Three points are recorded, yielding one triangle per frame.

The scope is to determine the measured error against a trusted 90^o angle. The same experiment could be conducted at various degrees, and various positions within the observed volume. In our case, we used a calibrated laser cross system to confirm the rig's correct placement at 90^o (Figure 2 left). Note that the tablet camera is not fixed on a tripod, to demonstrate hand jitter which is vividly presented on Figure 2, centre-right, as offset multiple triangles. The test algorithm is: *Capture* an object of known spatial properties (width, height, angle) with the camera. *Compare* the measurement results to object's already known properties. Finally, extract differences and *categorize* possible errors, such as axes discrepancies, errors of perspective, angle errors, etc. Figure 2 offers an example of the above. The mean measured angle from the captured data (76 frames) is 88.5^o, instead of 90^o, revealing an error of ~1.5^o, well below the 5^o acceptable limit. For a thorough system characterization, the same experiment should be repeated for various angles, at different positions within the observation volume, which is beyond the scope of this paper.

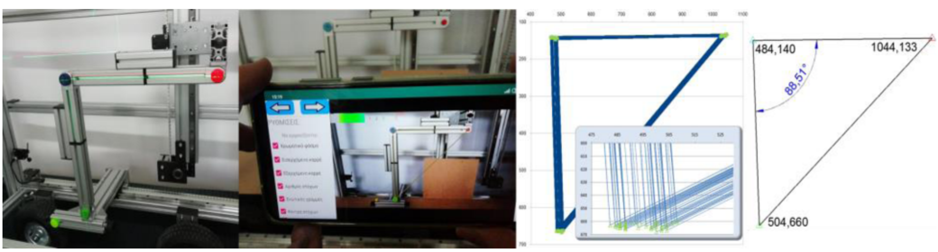


Figure 2. Test A. Left: partial photo of the rig posing. Centre-left: pose captured with our marker tracking software. Centre-right: visualization of 76 frames coordinates. Right: angle measurement on a single frame.

4.2. Test B

It is dynamic, suitable for optical and / or inertial measurements. It is designed to reveal measurement divergence due to motion, and the accuracy of the gaps-filling estimation procedures. Setup:

1. Recorded, open source, human gait joint coordinates data (e.g. [17]) of known quality.
2. The biomimetic rig to emulate the recorded human gait.
3. The gait measurement system to capture the rig's gait. The camera's tripod is placed at a distance of 130cm away from the rig.

If the quality of the data which will drive the rig is not known, simple tests can illustrate the precision/variability of the data. In this example, the hip-knee distance variation is approximately 30.6mm, as shown in Figure 3 (left) which displays three pairs of coordinates per frame (of the hip, knee and ankle). The limb(s) of the biomimetic rig are continuously driven with the gait data. The gait is then captured with the optical system and the results are compared to initial data, in order to assess differences. By plotting the initial and captured frames per gait cycle, in different colors (e.g., grey versus red), with the aid of a design platform, the errors "pop-up", i.e. become apparent even to the untrained eye (Figure 3, right). Note the increasing distortion / noise towards the far ends of the gait cycle, as represented by the "misplacement" of captured ankles (red dots) against the original data. Plotting also facilitates fast, indicative measurements, with the ability to isolate frames of particular interest. In this example, measurements at selected frames reveal spatial errors around 24mm and angular deviations around 2° which are typical for non-calibrated cameras. Normally, to assess the optical-instrument-induced-variability, statistical analysis and point-by-point comparisons can be used as illustrated in [18].

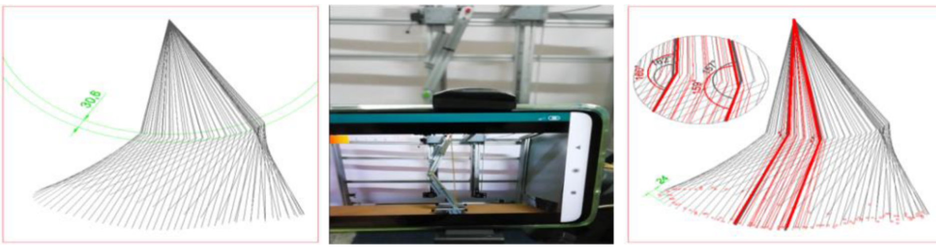


Figure 3. Test B. Left: initial data visualized. Centre: biomimetic rig emulating walk. Right: captured data (red) superimposed on initial data (grey). Measurements on selected frames.

4.3. Test C

It is a dynamic test, suitable for optical and inertial measurements, performed with a human subject, designed to reveal divergence by magnifying error components. Setup:

1. A human volunteer with attached markers.
2. A treadmill for the volunteer, so as to keep the capture camera steady.
3. The optical gait measurement system to capture the human's gait.
4. The biomimetic rig to emulate the recorded human gait.
5. The same gait measurement system to capture the rig's gait.

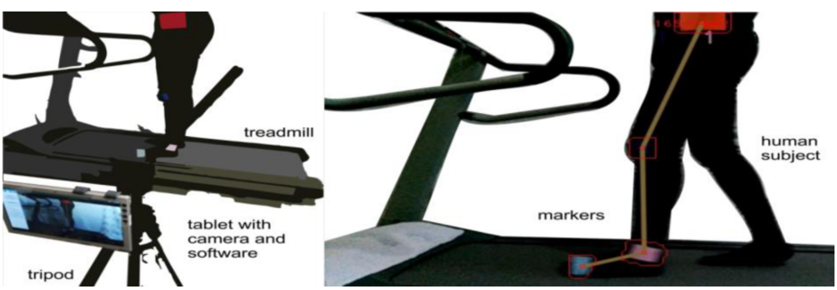


Figure 4. Left: Subject on treadmill, tablet on tripod. Right: Close-up of tablet monitor with markers and connecting lines. Background has been obscured afterwards.

An example is exhibited in Figures 4~6. The camera of the measuring system captures the sagittal plane at a distance of approximately 130cm from the treadmill. It is noted that the setup includes controlled light sources to help reduce shadows on markers. The male volunteer with height 182cm, 75kg weight, 45cm angle-knee distance, and 53cm knee-hip distance, walks for a few minutes on the treadmill, enough for the gait system to capture several complete gait cycles. Tabular joint data are analyzed, transformed and downloaded to the biomimetic rig to drive the limbs.

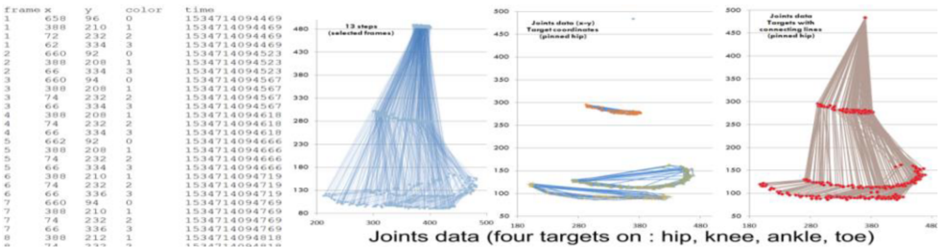


Figure 5. Test C. Data from human subject. Left: walking data sample as captured using 4 targets on left foot. Right: selected frames from the swing stage, visualized after processing.

Figure 6. Test C. Left: selected angle measurements from human captured swing stage. Centre: Rig walks. Right: visualization and measurements on selected frames.

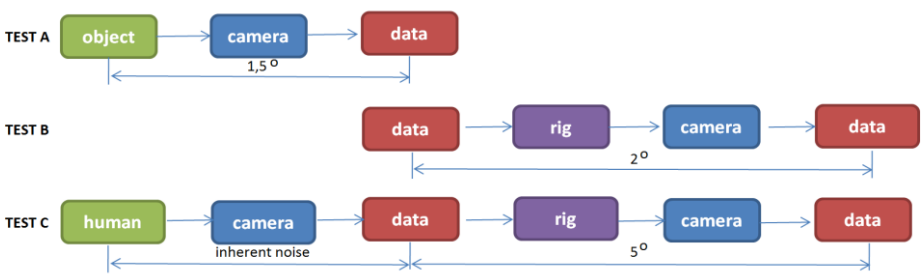


Figure 7. Error components

In Figure 6 left, note that the hip-knee distance from the captured human gait varies approximately 20.4mm. The rig's gait is then captured with the same optical system. The rig's gait data are plotted as in test B. Compared against the human gait, they reveal satisfactory fit: knee flexion/extension angular error reaches, but does not exceed, the 5° limit. In this case the initial hypothesis would be confirmed, if 5° errors were satisfactory for diagnosing that specific ailment, or rejected otherwise. In our particular example, the difference in angular errors (2° versus 5°), between test B & C, reveals the effect of inherent noise in the "camera plus markers plus software" system. In test C, this "noise" is added twice (Figure 7), given that the data used in test B are from significantly smaller variability.

5. Discussion

In this paper, we discussed the error components in gait analysis systems and presented a methodology for assessing the intra-equipment variability. The main contribution is that we managed to exclude the human Intra-Subject variability, which up to now has been largely uncontrolled, unknown, and unmeasured, thus affecting the overall accuracy of the gait analysis [18]. The remaining variability is due to the rest factors (equipment + therapist), that can now be revealed, thus easier to control.

Furthermore, with the aid of a biomimetic rig, it is now feasible to duplicate human gait patterns from past analysis or from distant labs data, and study them repeatedly, using various software / hardware configurations.

Similar methodologies can find interesting uses. Experiments like Test B, based on biomimetic rigs, could assist in training A.I. neural networks for gait capture. For example, by repeating a known pathological gait on the rig multiple times, the A.I. system can receive new training data (coordinate streams or video files) [19], from various camera positions, along the sagittal, frontal and transversal planes, at different lighting conditions. The above training data are automatically tagged, since they derive from the same pathological gait and thus, supervised machine learning training can be highly facilitated. The efficiency and maturity of a trained A.I. system can be assessed by presenting data from new perspectives. Again, using the biomimetic rig, this process can be automated and run extensively. An experienced A.I. could later, identify specific pathological gaits, in public human environments.

The same methodology can be adapted to other needs. E.g., gait labs/clinics, that will focus on the application of exoskeletical mechanisms, could be assisted by the biomimetic rigs, in the customization and tuning phases of such mechanisms, as much as possible, prior to the patients' visit (thus minimizing initial discomfort and fatigue). I.e., the first series of customization can be performed on the test rig, with the aim of reducing the differences (error components) between the pathological and the standard gait.

6. Conclusion

For assessing gait capture systems and labs, the human gait cycle variability (intra-subject) can be minimized by substituting the human subject with a robotic device. With the presented methodology, even low-cost, single camera systems can easily be benchmarked, without resulting to Golden standard systems, and thus may be used for certain gait analysis tasks within their specifications.

Identification of errors and variability in posture measurements can be accomplished with static tests, such as test A. Errors and variability in gait cycle characteristics estimation can be studied with dynamic tests, such as test B and C. Intra-Equipment variability can be amplified by closing the loop, i.e., by repeating the motion capture analysis twice within the same set of experiments, as in test C. The presented tests and measurements are just indicative and not limiting the possible tests and measurements that could be produced by the same or other setups.

7. Acknowledgments

The authors would like to thank stakam s.a., for the support and funding. Part of this research has been done under the framework of the "SCILED" project, 601137-EPP-1-2018-1-RO-EPPKA2-KA, co-funded by the E.U. Erasmus+ programme.

References

[1] Perry, J., Burnfield, J., "Gait Analysis, Normal and Pathological Function",J Sports Sci Med.9(2), (2010), 353.

[2] Pulignano, G., Del Sindaco, D., Di Lenarda, A., Alunni, G., Senni, M., TArantini, L., Cioffi, G., Tinti M. D., Barbati, G., Minardi, G., Uguccioni, M., "Incremental Value of Gait Speed in Predicting Prognosis of Older Adults With Heart Failure: Insights From the IMAGE-HF Study", JACC: Heart Failure, vol 4, issue 4, (2016), 289-298.

[3] Muro-de-la-Herran, A., Garcia-Zapirain, B., Mendez-Zorrilla, A., "Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications", Sensors, 14, (2014), 3362-3394.

[4] Fried, A.V., Cwikel, J., Ring, H., Galinsky, D., "ELGAM, Extra laboratory gait assessment method: Identification of risk factors for falls among the elderly at home", Int. Disabil. Stud.,12,(1990), 161-164.

[5] Anderson, B., Mingqian, S., Vincent, T., Ye, W. "Mobile Gait Analysis Using Foot-Mounted UWB Sensors", (2019), Article No 73.

[6] Bovi, G., Rabuffetti, M., Mazzoleni, P., Ferrarin, M., "A multiple-task gait analysis approach: Kinematic, Kinetic, and EMG reference data for healthy young and adult subjects", Gait & Posture, Vol 33(1), (2011), 6-13

[7] Fonseca, M., Gasparutto, X., Leboeuf, F., Dumas, R., Armand, S., "Impact of knee marker misplacement on gait kinematics of children with cerebral palsy using the Conventional Gait Model-A sensitivity study", PLoS ONE 15(4), (2020).

[8] Gorton G. E., Hebert D. A., Gannotti M. E., "Assessment of the kinematic variability among 12 motion analysis laboratories" Gait posture, vol. 29, no 3, (2009), pp. 398-402,

[9] Bernardina, G.R.D., Cerveri, P., Barros, R.M.L., Marins, J.C.B., Silvatti, A.P., "Action Sport Cameras as an Instrument to Perform a 3D Underwater Motion Analysis" PLoS ONE 11(8), (2016).

[10] Rocha, A.P., Choupina, H.M.P., Vilas-Boas, M.C., Fernandes, J.M., Cunha, J.P.S, "System for automatic gait analysis based on a single RGB-D camera", PLoS ONE 13(8), (2018).

[11] Wang, Q.,Kurillo, G., Ofli, F., Bajcsy, R., "Evaluation of Pose Tracking Accuracy in the First and Second Generations of Microsoft Kinect", IEEE International Conference on Healthcare Informatics (ICHI), (20156), 380-389.

[12] Parks, M., Chien J.H., Siu K.C., "Development of a Mobile Motion Capture (MO^2CA) System for Future Military Application", Military Medicine. vol. 184, (2019), 3/4:65.

[13] Cohen, E.J., Bravi, R., Minciacchi, D., "3D reconstruction of human movement in a single projection by dynamic marker scalling", PLoS ONE 12(10), (2017).

[14] Sporri, J., Schiefermuller, C., Muller, E., "Collecting Kinematic Data on a Ski Track with Optoelectronic Stereophotogrammetry:A Methodological Study Assessing the Feasibility of Bringing the Biomechanics Lab to the Field", PLoS ONE 11(8), (2016).

[15] Dawe, R.J., Yu, L., Leurgans, S.E., Truty, T., Curran, T., Hausdorff, J.M., et al., "Expanding instrumented gait testing in the community setting: A portable, depth-sensing camera captures joint motion in older adults", PLoS ONE 14(5), (2019).

[16] Muller, B.,Ilg, W., Giese, M.A., Ludolph, N., "Validation of enhanced kinectsensor based motion capturing for gait assessment", PLoS ONE 12(4), (2017).

[17] Fukuchi, C., Fukuchi, R., Duarte, M., "A public dataset of overground and treadmill walking kinematics and kinetics in healthy individuals", PeerJ6, (2018), e4640.

[18] Schwartz, Michael H., Trost, Joyce P., Wervey, Roy A., "Measurement and Management of Errors in Quantitative Gait Data" Gait and Posture 20, (2004), pp. 196-203

[19] KidzinskiL., Yang B., Hicks J.L., Rajagopal A., Delp S.L, Schwartz M.H., "Deep neural networks enable quantitative movement analysis using single-camera videos", Nature Communications, 11, (2020), 4054.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210072

Development of an Innovative Learning Methodology Aiming to Optimise Learners' Spatial Conception in an Online Mechanical CAD Module During COVID-19 Pandemic

Zoe KANETAKI, Constantinos STERGIOU, Christos TROUSSAS, Cleo SGOUROPOULOU
University of West Attica, Athens, Greece

Abstract. The COVID-19 pandemic struck humanity in February 2020. Closures of educational institutions, worldwide, resulted to the creation of emergency remote teaching environments as a substitute to face to face learning. The disruption caused in the academic community has stimulated innovative learning methods within all levels of the educational sector. New parameters affecting knowledge transmission are getting involved while students follow courses apart on a common virtual learning environment. This research is based on a first-semester Mechanical Engineering CAD module in tertiary education. A learning strategy has been applied by reforming the traditional face-to-face leaning mode to a fully remote learning environment. The methods applied have been tested using statistical analysis and have shown to contribute significantly in students' spatial perception in 2-Dimentional Drawings. The outcomes of this research reveal a novel teaching strategy that improved students' academic achievements in CAD during the lockdown. Specific aspects can be considered sustainable on their return back to normality.

Keywords. CAD, Learning strategy, Online Learning, Engineering Education, COVID-19.

1. Introduction

In February 2020, the impact of the COVID-19 pandemic struck the academic community by resulting to the closure of educational institutions around the world. With 94% of the world's student population affected, the pandemic has caused the biggest disruption in education systems in history[1]. Emergency remote teaching environments were created as a substitute for face-to-face learning. During the second quarter of 2020, schools and universities around the world responded to the situation by offering synchronous and asynchronous online learning. Engineering needs, on the other hand, have sparked new learning approaches at all levels of the educational system. Advanced technology tools are introduced into the instructional process [1-6], such as digital platforms and social media channels, serving as the principal mean of knowledge transmission.

[1] United Nations Policy Brief: Education during COVID-19 and beyond AUGUST 2020

The adoption of online-based education in higher education settings around the world offered a unique opportunity to collect electronic data and track students' academic progress, which motivated this study. Specific aspects and learning goals achievement level of the presented modules could be extract out of platform's activity reports, including assessments' results and online surveys.
The novelty of this study consists on the fact that pandemic restrictions have just recently been imposed on educational systems, fully virtual learning environments that provide students' data under these enforced limitations have yet to be evaluated.

1.1. Related work

In order to create a successful learning environment, goals have been determined in each phase of the online module's content flow based on previous studies in the field of Engineering Education. In [7], students' engagement in a learning framework can be differentiated into four behavioral modes during lectures: passive, active, constructive and interactive. Those four stages have been applied in the online module of Mechanical Design with CAD module, under the extenuating learning circumstances of COVID-19 pandemic, when all face-to-face classes have been cancelled and new knowledge sharing solutions should be implemented in a fully remote learning environment. [8]. Previous research has demonstrated a variety of strategies to operationalize the "taking notes" as an index of activity engagement [7]. Asynchronous support can be provided especially in a remote teaching environment, using student friendly modes, including social media channels [9]

In [10], Fuchs made an important argument of using cases in engineering education, aiming to "bring the outside reality into the classroom". Students increase their awareness in real world tasks which motivates them to learn the concepts they need, in order to maximize their performance in engineering disciplines [11]. By introducing tasks, with which engineers will have to deal in their future work, helps developing basic concepts and problem-solving skills, while providing professional experience to the students [12].

2. Module content and description

During the face-to-face learning environment, after attending the ten laboratory lectures, students should be able to represent in 2d views a simple (basic) object referring to a mechanical part, shown in a 3-Dimentional isometric view, by applying the rules of mechanical drawings (orthographic projections hidden edges, visible edges, view placement, top views, side views, sectional views etc.) in a CAD environment. While students are able to represent the visible edges of an object from the isometric given view, their difficulties have been detected on conceiving the bottom views, back views, and sectional views determining the inner form of an object.

One of the basic rules of mechanical drawings is defining an object throughout a minimum number of views, which makes spatial conception an essential aspect of mechanical design. Back, bottom and sectional views are not directly visible in a single isometric view, therefore, have to be interpreted through the hidden edges symbolized by a dashed line. In physical laboratory classrooms, the course flow can be described in the following diagram (figure 1):

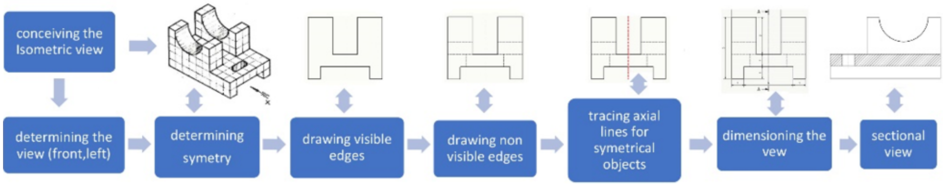


Figure 1. Teaching strategy applied during the face-to-face learning environment

3. The reform of the Mechanical Design CAD module

The reform of the curriculum, due to the COVID-19 imposed distance learning environment, has been based on previous experience of the Emergency Remote Teaching Environment [12] that took place during the spring semester 2019-2020 [13] for the module Mechanical Design CAD II (3-Dimentional modeling module). The outcomes of this research and the student's feedback have been used as guidance for a complete restructure on the Mechanical Design CAD I (2-Dimentional representation module) online module's context.

The goals of this previous experience have been centered on the unexpected intrusion of the COVID-19 pandemic and its social impact to the world's academic population.

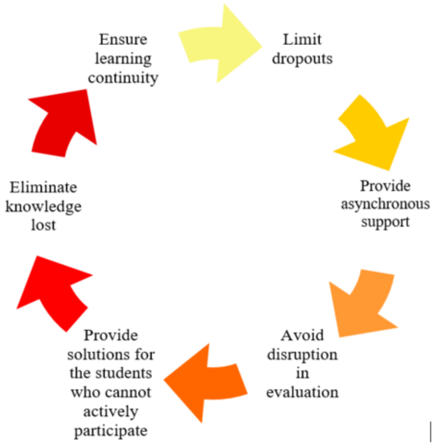
The learning strategy has been entirely reformed, while focusing on achieving the learning goals, in a fully remote teaching environment (figure 2). The research has been dwelt on first year students, who had previous experience in online learning, through an ERTE[2] [8] performed on their last high school year, during the first lockdown period in Greece.

The E-course flow is divided in 3 stages:

- The first part consists on passive learning, a lecture during which the theoretical part is presented (figure 3). During this part, learners are engaged in a passive mode [7] where their behavior is limited on listening and paying attention without doing anything else (no need to take notes, a pdf is uploaded in Microsoft Teams files, in the section of each lecture content). Since online synchronous lectures include a real time constructive phase, previous engagement of a passive learning stage should be "pen-free" in order to help students concentrate in the presented topic, relieving them from the anxiety of the note-taking action.
- The second part is followed by a quiz, a self-evaluating activity, aiming to determine the level of understanding of the lecture content, as a cognitive and recall new knowledge stage (figure 3). The Quiz has been created with MS Forms, with inserted images of geometrical shapes corresponding to multiple choice questions. Responders where able to see their results and correct answers immediately after submitting the quiz.
- The third part consists on the first minds on activity: Students have to represent on a 2-Dimentional sketch (scaled freehand drawing) the 3 views of a mechanical part. Methodology is assisted by a video in order to help students apply the new knowledge, after being student passive recipients (figure 3). At

[2] Emergency Remote Teaching Environments

this point, students cannot actively manipulate the video since the instructor holds control while constructively generates the presented methodology by explaining steps and refers to prior knowledge (figure 7). The object is viewed in a 3-Dimentional video of the model while rotated in order to carefully inspect it and fully conceive it [14].



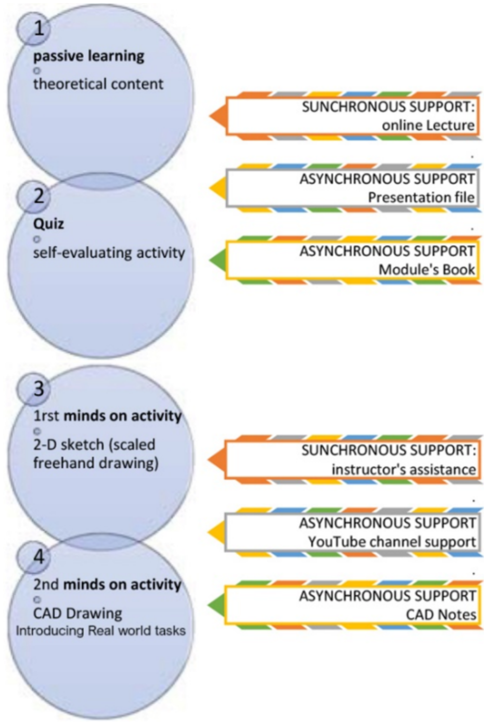


Figure 2. basic goals of the module's reform

Figure 3. learning's reform strategy

- The fourth part consists on the CAD drawing of the sketched part. Students watch real time instructions via screen sharing, followed by an interactive dialog with their instructor and debating about different ways to achieve the task. Students are also provided with asynchronous instructional support via MCAD I YouTube channel videos. During this phase, a specific task has been assessed to help students on understanding the relation of the tasks assigned through the online CAD I module with real-world tasks, and with their future work in Mechanical Engineering [15] (figure 5).

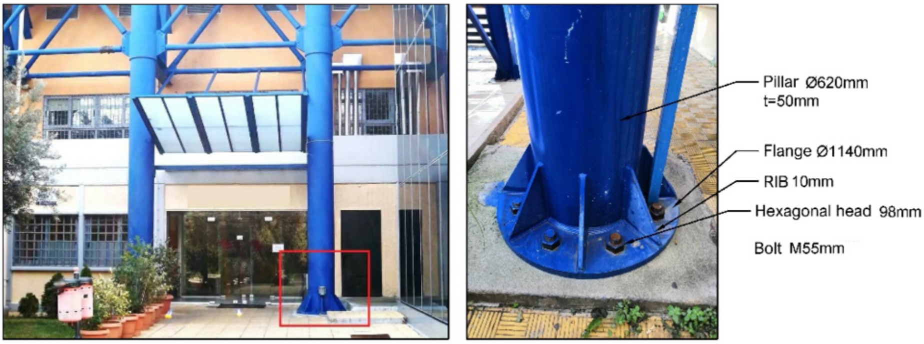


Figure 5. Assignment related to real-world tasks in Mechanical Engineering, uploaded on You Tube

The specific project consists on a CAD drawing of an existing large scale metallic construction and aims to promote the transfer of new knowledge and develop students' co-creating skills. Students should be able to represent in a 2-Dimensional CAD drawing; the existent metallic structure of the library building located in the Ancient Olive campus of the University of West Attica (figure 5). Since most students didn't have the opportunity to visit the campus due to the lockdown, the presentation of the assignment needed to be precise and integrate the most relevant characteristics to the students' future work, which includes most aspects of metallic assemblies like beams, flanges, ribs, bolts, nuts, washers. A video representation of the metallic structure has been uploaded as an added link in the resource option of MS Teams assignment[3] (figure 6), as well as a screen and audio recording of the drawing methodology of basic geometric shapes[4] and AutoCAD commands[5] (polygon, write block, arc, polar array).

Figure 6. Instances from video representation of the 15th assignment on You Tube channel

It should be noted at this point those students who had already developed their special conception had quit sketching. Therefore, they have managed to skip the third part of the online module and did not need to create a scaled freehand drawing in order to proceed on a Computer Aided Design.

All activities where announced, uploaded and graded on MS Teams platform to avoid disruption with several LMS platforms.

[3] https://youtu.be/RWY5EuLB9nQ
[4] https://youtu.be/HT-oJB2JZ9k
[5] Autodesk Autocad 2017 Educational Version

4. Developing students' spatial conception

One of the most critical challenges of this research has been to apply technology features and empirical techniques that could favorize the development of students' spatial perception. The reform strategy induced the surface planes as an exploded view of the given object. Each isometric view has been analyzed in surface planes parallel to one another, perpendicular or even not related to each other. The surface planes where tangent to each edge set, determining a single view. The methodology resulted on helping students conceiving the representation of continuous line segments for visible edges, as well as the tracing of dashed line segments for the non-visible edges, belonging to parallel back planes, while passing from isometric views to 2-Dimentional Drawings. When sectional views have been announced in the virtual classroom, the presence of the cutting plane determining the point on the plane where the sectional view should be drawn, has been highlighted in the isometric view, instead of the typical chain line with thick ends and two arrows in the direction of the view. Each object that has been assessed as task has been previously modeled in Autodesk Inventor 2020[6] in order to be presented in a photoreal view. In the reformed virtual laboratory classrooms, the course flow can be described in the following diagram:

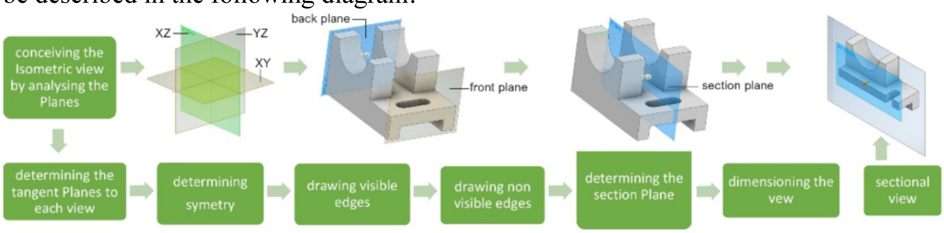


Figure 7. Teaching strategy applied during the virtual learning environment

5. Evaluation of the learning strategy

The online reformed module has been applied to 216 (N=216) students which were divided in 11 online platform teams, assisted by a group of 5 instructors in MS Teams. 190 (n=190) students replied to the surveys out of which 165 where first semester students. Out of the 190 participants, 88% where male and 12% female students.

A univariate analysis has been performed throughout a survey containing 101 questions (variables). Variables that have proven to have a significant relation with the students' final grade in the specific module (Table 1) have been filtered out through a correlation analysis performed in SPSSv2020 using data mined from four sources:

- *1st online survey conducted during the 1rst online lecture:*

The questions provided information about student's previous educational information (i.e. university entrance exams grades in four courses)

- *2nd online questionnaire conducted during the ninth (out of ten) online lecture,*

An online questionnaire including various constructs of the learning strategy has been shared to the students during the eighth out of nine lectures of the academic semester. Constructs' categories included variables indicating the students' emotional

---

[6] Autodesk Inventor Professional 2020 Educational Version.

state during the lockdown, insecurity, social distancing, evaluation factors in a Likert scale, spatial conception and module content convenient to real world tasks in mechanical engineering.

- *General characteristics of the students, the class croups and the instructors*

The above include gender, age, high school type, class time (morning or afternoon class) number of instructors, instructors id.

- Insights from MS Teams Platform, assignment Grades, Final's exam grade.

6. Qualitative Analysis

The learning strategy has been tested and proved to have a general acceptance from the population that participated.

High ratings have been shown in specific aspects as overall evaluation of the learning strategy, organization, enjoyability of the online module, preference to other laboratory and theoretical courses. The new introduced features aiming to help students improve their special conception have been positively received and highly evaluated by the majority of participants. The sustainability of the learning strategy has been evaluated as well and significant percentages have been localized in the way weekly tasks are assessed 62% (via MS Teams platform), parallel asynchronous support from the MCAD I UNIWA YouTube channel 66.8% and 85% of the students rated the videos between 4-5. Full time lecture recording has not been widely suggested, as the percentage is limited in 45%. Parallel online transmission of each 3hour lecture has been suggested by 51% of the students, which implies that students prefer short videos describing the solving methodology than 3-hour lecture recordings (figure 8).

Table 1: Significant variables (survey questions) corelated to student's positive evaluation

	Significant variables (survey)	Evaluation /5	Evaluation range 4-5
1	Enjoyable vs other Labs	4.33	89%
2	CAD I helped familiarising to MS Teams	4.16	78%
3	CAD I vs other Modules	4.27	83%
4	Confident for the finals CAD I	3.47	53%
5	Confident for the finals all modules	3.12	31%
6	Organisation	4.25	84%
7	Theory contributes on tasks	3.59	55%
8	Planes well-conceived	4.16	81%
9	Importance of section plane	4.47	90%
10	Quality of videos	4.27	85%
11	Clearness of videos	4.17	84%
12	Pleasant experience	4.38	90%
13	Interactivity	3.60	61%
14	Questions resolved	4.32	83%
15	Assignments related to future tasks	3.84	62%
16	15th assignment video	4.01	76%

17	15th assignment related to real world tasks	4.08	77%
18	Evaluation vs other labs	4.19	82%
19	Overall evaluation	4.21	85%

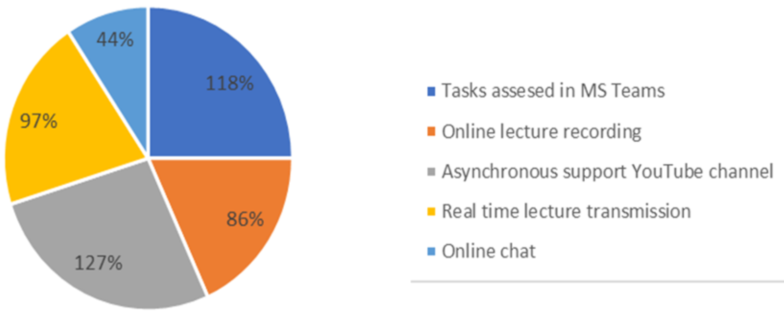


Figure 8. Percentage of preference in Sustainable learning modes and support

7. Conclusions and Future work

Learning materials, online accessibility, student involvement, and communication have always been essential to the e-learning philosophy [16]. During the transition from face-to-face to the COVID-19 emerging learning environments, most University online modules' lectures and Laboratories in the University of West Attica have been using simultaneously MS Teams (Communication Platform) for the synchronous lectures and Moodle or E-class (Learning Management Systems) for asynchronous support. The last included tasks' assessment, gradings, modules' notes and final semester exams. The procedure previously mentioned aimed to facilitate the transition mostly for educators who have been using LMSs before the pandemic, due to the lack of time for preparing the material but also for getting more familiarized to MS Teams. During the first period of the pandemic, the use of both platforms has caused confusion in most students with inferior computer skills.

One of the most important outcomes of this study is that using a single online platform for transmitting synchronous lectures and at the same time uploading tasks' assessment, module notes, and grading and asynchronous support facilitates task's organization in both students and instructors. In this case where asynchronous support has been provided by social media channels (in this case YouTube MCAD I UNIWA channel), integrating assignments and resources links uploaded on social media channels has contributed on organizing functionalities for educators and at the same time provided direct access and control to students, for managing their homework schedule.

By determining educational patterns and learning modes that have been issued while transforming an existent face-to-face laboratory course to a fully remote teaching environment, the question is whether some of those learning modes can be applied in the future, on the return to real classroom environment, by the means of a synchronous or asynchronous support, or even as a point of reference of each lecture content.

Future work consists of applying the single platform educational mode to first-semester students who will be attending the same module during the academic year 2021-2022, in a similar, blended or virtual learning environment in order to determine the sustainability of the methodology suggested by this study, in future learning environments.

References

[1] C. Troussas, A. Krouska, C. Sgouropoulou, I. Voyiatzis, Ensemble Learning Using Fuzzy Weights to Improve Learning Style Identification for Adapted Instructional Routines, *Entropy* **22**(7) (2020), 735. https://doi.org/10.3390/e22070735.

[2] A. Krouska, C. Troussas, C. Sgouropoulou, Fuzzy Logic for Refining the Evaluation of Learners' Performance in Online Engineering Education, *European Journal of Engineering and Technology Research* **4**(6) (2019), 50-56. DOI: https://doi.org/10.24018/ejers.2019.4.6.1369.

[3] C. Troussas, F. Giannakas, C. Sgouropoulou, I. Voyiatzis, Collaborative activities recommendation based on students' collaborative learning styles using ANN and WSM, *Interactive Learning Environments* (2020). DOI: 10.1080/10494820.2020.1761835.

[4] C. Troussas, A. Krouska and C. Sgouropoulou, A Novel Teaching Strategy Through Adaptive Learning Activities for Computer Programming, *IEEE Transactions on Education* **64**(2) (2021), 103-109. DOI: 10.1109/TE.2020.3012744.

[5] C. Troussas, A. Krouska, C. Sgouropoulou, Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education, *Computers and Education* **144** (2019), 103698. https://doi.org/10.1016/j.compedu.2020.103698.

[6] C. Troussas, A. Krouska and C. Sgouropoulou, Improving Learner-Computer Interaction through Intelligent Learning Material Delivery Using Instructional Design Modeling, *Entropy* **23**(6) (2021), 668. https://doi.org/10.3390/e23060668

[7] M. Chi, R. Wylie, The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. Educational Psychologist. 49 (2014). 10.1080/00461520.2014.965823.

[8] C. Hodges, S. Moore, B. Lockee, T. Trust, A. Bond, The Difference Between Emergency Remote Teaching and Online Learning, Educausereview March 27, (2020).

[9] W. Al-Rahmi, M. Othman, M. Musa, The Improvement of Students' Academic Performance by Using SocialMedia through Collaborative Learning in Malaysian Higher Education. Asian Social Science. (2014). 10. 10.5539/ass. v10n8p210

[10] H.O. Fuchs, Outside reality inside the classroom. Journal of Engineering Education 60 (7) (1970).: 745–47.

[11] A. Yadav, G.M. Shaver, P. Meckl, Lessons Learned: Implementing the Case Teaching Method in a Mechanical Engineering Course. Journal of Engineering Education. (2010), 99: 55-69. https://doi.org/10.1002/j.2168-9830.2010.tb01042.x

[12] C. Whittle, S. Tiwari, S. Yan, J. Williams, "Emergency remote teaching environment: a conceptual framework for responsive online teaching in crises", in Information and Learning Sciences, (2020) Vol. 121 No. 5/6, pp. 311-319, published on 29/06/2020, https://doi.org/10.1108/ILS-04-2020-0099

[13] Z. Kanetaki, C. Stergiou, G. Bekas and E. Kanetaki, "Machine Learning and Statistical Analysis applied on Mechanical Engineering CAD course: A Case Study During ERTE Pahse in the Context of Higher Education," 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), (2020) pp. 1-13, doi: 10.1109/ISMSIT50672.2020.9254924.

[14] K.H. James, G.K. Humphrey, T. Vilis, B. Corrie, R. Baddour, M.A. Goodale, "Active" and "passive" learning of three-dimensional object structure within an immersive virtual reality environment. Behavior research methods, instruments, & computers: a journal of the Psychonomic Society. (2002). Inc. 34. 383-90. 10.3758/BF03195466

[15] M. Binkley, O. Erstad, J. Herman, S. Raizen, M. Ripley, M. Miller-Ricci, M. Rumble, Defining Twenty-First Century Skills. (2011). 10.1007/978-94-007-2324-5_2

[16] M. Krašna, I. Pesek, "Influence of Moodle and MS Teams on teaching-learning-studying (TLS) processes," *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, (2020), pp. 612-616, doi: 10.23919/MIPRO48935.2020.9245356.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210073

Enhancing the Effectiveness of Intelligent Tutoring Systems Using Adaptation and Cognitive Diagnosis Modeling

Akrivi KROUSKA [a,1], Christos TROUSSAS [a], Filippos GIANNAKAS [a], Cleo SGOUROPOULOU [a] and Ioannis VOYIATZIS [a]
[a] *Department of Informatics and Computer Engineering University of West Attica, Greece*

Abstract. This paper presents a novel framework for developing educational hypermedia systems incorporating adaptation techniques and tailored feedback. In particular, the adaptation techniques refers to the content presentation; where the system hides/displays information according to students' knowledge level and learning goals, and to the navigation design; where the system proposes the learning path that is better to be followed based on their profile. Finally, the framework embodies a diagnostic model that analyzes the students' misconceptions and provides tailored feedback and advices on bridging students' knowledge gap. This framework aims to enhance the effectiveness of learning process, increasing student engagement through the adaptive content and navigation and improving student performance through the tailored feedback.

Keywords. Adaptive content, Adaptive navigation, Buggy model, Personalized learning, Tailored feedback

1. Introduction

Adaptive hypermedia increases the functionality of hypermedia by tailoring their presentation to each individual, and thus, providing an alternative to the traditional "one-size-fits-all" approach applied to the development of hypermedia systems [1]. Adaptive educational hypermedia systems construct a dynamic students' model based on their goals, preferences and knowledge, which is updated according to students' interactions with the system [1, 2]. Hence, the systems adapt the presentation of their content and the navigation interface to each student's needs [1, 3].

An educational system focuses on a certain piece of knowledge, namely a course. The students' goal in an educational system is to learn the whole content provided or a significant part of it. The use of hypermedia helps students learn better. In particular, adapting the hypermedia to students' characteristics can improve students' performance and boost their engagement in learning process [4]. The main student characteristic that adaptive hypermedia systems use is student knowledge; since, the level of knowledge varies from one student to other, as well as each student has his/her own pace of learning [5]. Student knowledge emerged from the mistakes made during the assessment process. However, identifying why the student makes these mistakes can further help the student

[1] Corresponding Author, E-mail: akrouska@uniwa.gr.

improve his/her cognitive state [6, 7]. The diagnosis of student misconception is used for providing tailored feedback and advice with the intention of bridging the knowledge gap [8].

To this direction, this paper introduces a novel framework for developing effective educational hypermedia systems through the incorporation of adaptation techniques and cognitive diagnosis model. This framework can provide a new learning experience to students increasing student engagement and improving learning outcomes.

2. Adapting content presentation and system navigation based on student knowledge level

Adapting learning environments is crucial for confronting the differences between the students in their needs and preferences [9]. Users with different goals and knowledge may be interested in different pieces of information presented on a hypermedia page and may use different links for navigation [10]. Adaptive hypermedia systems include two basic areas for customizing their operations according to each student's individual needs: adaptive presentation and adaptive navigation.

Adaptive presentation refers to the function of adapting the content of a page based on student characteristics, such as knowledge level, learning goals and others [11]. In particular, it determines what information should be presented and how it should be organized and displayed. For example, the system should provide more detailed and advanced material to expert students, while it should provide more explanations and examples to novice students. The techniques used include the conditional text, stretch text, page variants, fragment variants and frame variants.

Adaptive navigation supports students in the learning environment by adapting the appearance and structure of user interface [12]. Hence, systems can support users in their navigation by limiting browsing space, suggesting most relevant links to follow, or providing adaptive comments to visible links, according to their profile. With this adaptation, students' learning experience and outcomes are expected to be improved. The adaptive navigation techniques used include the direct guidance, adaptive ordering, hiding and adaptive annotation.

3. Providing tailored feedback based on student misconceptions

Most of the tutoring systems consider only the number of mistakes a student makes in order to evaluate his/her knowledge. However, understanding why the student makes these mistakes is of major importance, since incomplete learning and forgetting content can be detected and faced using corresponding teaching strategies [10, 6]. Thus, an enriched model of student cognitive state, including his/her misconception, is a prerequisite for guiding him/her properly in learning/training process [13].

To accomplish this goal, a diagnostic model for student misconception that also takes into consideration pedagogical aspects, needs to be developed. A solution includes the adoption of the Repair Theory introduced by [14]. According to this theory, every bug is related to a series of operations that generate an incomplete procedure and to a series of operations that represent the repair of the procedure so that it can proceed [15]. These two parts are independent; thus, the kind of repair attempted depends only on the procedure and its current impasse, not on how the incomplete procedure was derived.

In view of the above, the diagnostic module needs to build a large repository of predefined bugs that is used for inserting the corresponding misconception to student model. This repository stores for each incorrect answer of test items, the kind of misconception concerned and the material referred to. Hence, every time students make mistakes, the system updates their profile and advices them which actions to be followed, e.g. the section of the lesson that needs to be studied more carefully, in order to "repair" their misconception[2].

4. A framework of adaptive tutoring systems incorporating a diagnostic model of student misconceptions

This paper proposes a framework for developing an adaptive educational hypermedia system that facilitates the learning process and improves learning outcomes. The architecture of the system is based on the elements of intelligent tutoring systems. As such, it includes four modules, namely student, domain, tutor and interface module.

Regarding the student module, every time students give a test, the system records their mistaking behavior and updates properly their student profile, namely their knowledge level and misconceptions detected. The domain module stores data about the course content, assessment items, teaching strategies and buggy rules, as well as the relationship between them. The domain module feeds with data both the student module for building student profile and the tutor module for the adaptation process. The tutor module includes the adaptation techniques incorporated into the system for tailoring user interface and navigation, and the diagnostic model used for providing personalized feedback. To accomplish this, the tutor module combines the information emerged from student and domain module. Finally, the interface module delivers to students adaptive content regarding its presentation and navigation design. Furthermore, it provides tailored feedback to students according to their misconceptions in order to bridge the knowledge gap and improve learning outcomes.

The novelty of this framework is the adaptation techniques used and the diagnostic model applied. In particular, the proposed approach for the adaptive presentation is this of stretchtext, since it covers better the students' needs and system's goals for learning. Using the stretchtext approach, the system hides the expert knowledge from novice students and provides them to advanced students who need expertise. On the other hand, it provides additional examples and further explanations to novice students for enhancing the acquisition of knowledge. Regarding the adaptive navigation, the system incorporates the direct guidance and adaptive annotation. Using the direct guidance, when students log in, the system leads them to the lesson they need to study based on their progress in their previous visit. According to adaptive annotation approach, the system uses explanation icons next to lessons' links indicating the students' progress and thus, suggesting them an optional learning path to follow. It should be noted that students have the capability to study any lesson they want without the system restricting it to them. Finally, the system provides tailored feedback to students every time they give a test. The diagnostic model used for this purpose provides a mechanism for explaining the reason why a student makes a mistake, and not only identifies the mistakes. To this direction, a repository of buggy rules is constructed where each incorrect answer of a test item corresponds to a certain misconception of a lesson content. The buggy rules are

[2] http://tecfaetu.unige.ch/staf/staf-d/krige/staf11/buggy.html

combined with the teaching strategies defined based on the learning goals that need to be achieved, providing useful advices to students on the learning path that can lead them to improve their knowledge.

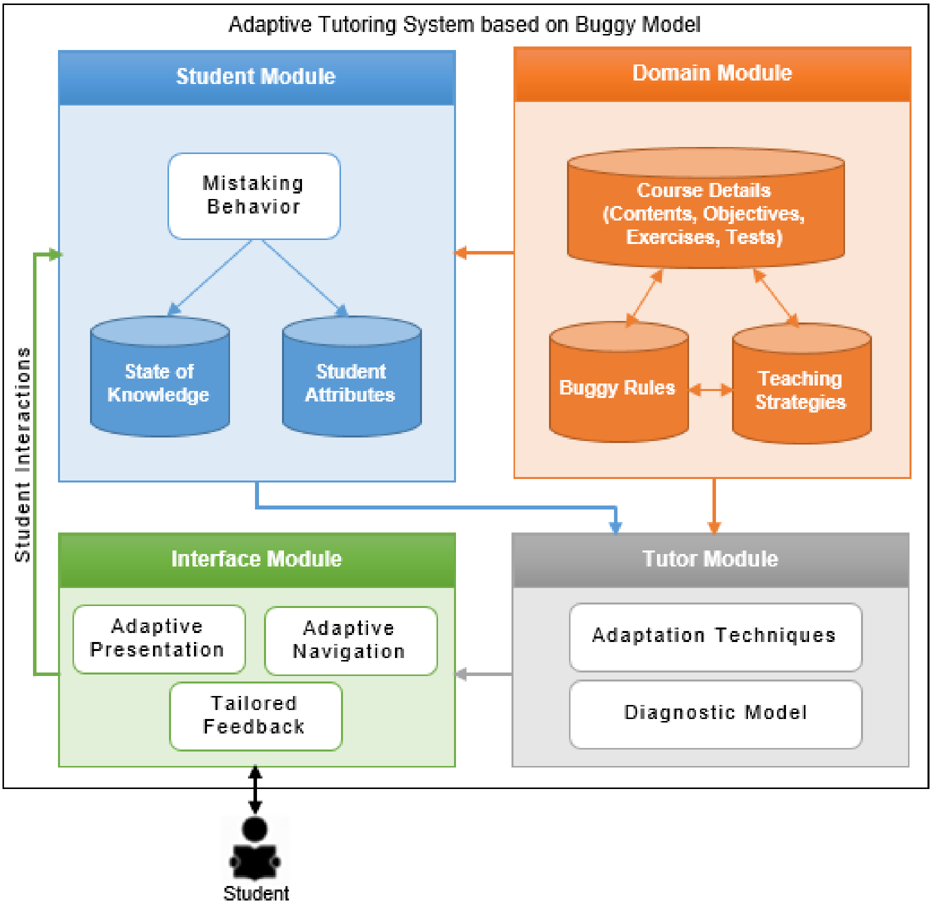


Figure 1. Architecture of proposed framework.

5. Conclusions

Adaptive educational hypermedia systems aims to guide students in their learning process in a tailored way, providing individualized learning instead of the traditional "one-size-fits-all" approach. To achieve their objective, they incorporate adaptation techniques and teaching strategies in order to deliver material to students based on their profile, needs, preferences and learning goals. The majority of tutoring systems consider students' knowledge level as the main characteristic for personalization identifying only the mistakes the students make. However, the diagnosis of the reason why a student make a mistake is of major importance for guiding properly in order to bridge knowledge gap.

To this direction, this paper introduces a novel framework for enhancing the effectiveness of tutoring systems. In particular, it incorporates adaptation techniques, namely adaptive presentation and navigation, and a cognitive diagnosis model based on student misconceptions. This framework is expected to boost student engagement and improve learning outcomes.

References

[1] P. Brusilovsky, Adaptive educational hypermedia, *In International PEG Conference* **10** (2001), 8-12.

[2] A. Krouska, C. Troussas and C. Sgouropoulou, A Personalized Brain-Based Quiz Game for Improving Students' Cognitive Functions, in *Brain Function Assessment in Learning. BFAL 2020. Lecture Notes in Computer Science* **12462** (2020), Cham, Springer.

[3] A. Krouska and M. Virvou, An Enhanced Genetic Algorithm for Heterogeneous Group Formation based on Multi-Characteristics in Social Networking-based Learning, *IEEE Transactions on Learning Technologies* **13(3)** (2020), 465-476.

[4] C. Troussas, A. Krouska and C. Sgouropoulou, Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education, *Computers & Education* **144** (2020), 103698.

[5] A. Krouska, C. Troussas and M. Virvou, Computerized Adaptive Assessment Using Accumulative Learning Activities Based on Revised Bloom's Taxonomy, *In Joint Conference on Knowledge-Based Software Engineering* (2018), 252-258.

[6] A. Krouska, C. Troussas and C. Sgouropoulou, Fuzzy logic for refining the evaluation of learners' performance in online engineering education, *European Journal of Engineering Research and Science* **4(6)** (2019), 50-56.

[7] N. T. Heffernan, K. R. Koedinger and L. Razzaq, Expanding the model-tracing architecture: A 3rd generation intelligent tutor for algebra symbolization, *International Journal of Artificial Intelligence in Education* **18(2)** (2008), 153-178.

[8] C. Troussas, A. Krouska and C. Sgouropoulou, Improving Learner-Computer Interaction through Intelligent Learning Material Delivery Using Instructional Design Modeling, *Entropy* **23(6)** (2021), 668.

[9] C. Troussas, A. Krouska, C. Sgouropoulou and I. Voyiatzis, Ensemble Learning Using Fuzzy Weights to Improve Learning Style Identification for Adapted Instructional Routines, *Entropy* **22(7)** (2020), 735.

[10] C. Troussas, A. Krouska, M. Virvou and E. Sougela, Using Hierarchical Modeling of Thinking Skills to Lead Students to Higher Order Cognition and Enhance Social E-Learning, *In 2018 9th International Conference on Information, Intelligence, Systems and Applications (IISA)* (2018), 1-5.

[11] A. Elmabaredy, E. Elkholy and A. Tolba, Web-based adaptive presentation techniques to enhance learning outcomes in higher education, *Research and Practice in Technology Enhanced Learning* **15(1)** (2020), 1-18.

[12] A. Hamzah and K. Irianto, Adaptive Navigation Support: A Literature Review of Its Application and Implication in Learning Process, *International Journal of Emerging Technologies in Learning (iJET)* **14(16)** (2019), 181-187.

[13] C. Troussas, A. Krouska and M. Virvou, "sing a Multi Module Model for Learning Analytics to Predict Learners' Cognitive States and Provide Tailored Learning Pathways and Assessment, in *Machine Learning Paradigms. Intelligent Systems Reference Library* **158** (2020), Cham, Springer.

[14] J. Brown and K. VanLehn, Repair Theory: A Generative Theory of Bugs in Procedural Skills, *Cogn. Sci.* **4** (1980), 379-426.

[15] J. S. Brown and R. R. Burton, Diagnostic models for procedural bugs in basic mathematical skills, *Cogn. Sci.* **2(2)** (1978), 155-192.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210074

Ambient Intelligence and Smart Environments: A Preliminary Overview

Maria POLI[a]
[a]*Lecturer, Department of Interior Architecture, University of West Attica, Aigaleo, Greece*

Abstract. Nowadays human activities are incoming at a digitalization stage. The introduction of information technology along with new forms of communication, influence a variety of forms of human action and focus mainly on the integration and the convergence of the digital and physical worlds. The use of more intelligent – electronic solutions, improves the lives of people around the world, according to studies carried out on the ingress of new smart technologies. Artificial and Ambient intelligence nowadays getting more and more attention about the development of smart, digital environments. The Smart Cities designed for All must aim to arrange the disparity in cities through smart technology, making cities both smart and accessible to a range of users regardless of their abilities or disabilities. The birth of "Artificial Intelligence" (AI) has facilitated the complex computations for reality simulation the new communication era of wireless 5G, all combined have given the hope for a new and better future, to reverse disability to empower the humans with more capabilities, to be faster than they can ever be, stronger than they can ever dream. This paper provides an overview of Ambient Intelligence and smart environments, as well as how technological advancements will benefit everyday usage by devices in common spaces such as homes or offices, and how they will interact and serve as a part of an intelligent ecosystem by bringing together resources such as networks, sensors, human-computer interfaces, pervasive computing, and so on.

Keywords. Ambient Intelligence, Smart Environments, Smart Cities, Augmented Reality Accessibility, Disabled.

1. Introduction

Researchers focus on Computational Intelligence (CI), Machine Learning and developing new techniques of intelligent environments. The spread of new technologies and Internet of Things (IoT) provide opportunities creating new web-based smart systems for all levels of education and related educational and learning tasks. [1]
This study aims to explore the prospects of: Section 1) Ambient Intelligence and Smart Environments, Section 2) Smart Cities, Section 3) With Augmented Reality and Internet of Things Improve Accessibility, a Paradigm of a Smart Store (RFID) technology, Section 3) Smart Store and RFID technology, paradigm of a super market, a definition what is smart, intelligent a packaging paradigm. From the conceptual model to the Smart and to the Intelligent. Smart - intelligent designing system a basic for design for All. Section 4) Supporting people with disabilities in their Daily Life in a Smart City. Section 5) Conclusion and Future.

Furthermore, the authors' perspective is extended by a glimpse into the future on how technological advancements such as Augmented Reality and the Internet of Things will be applied in a smart shop using RFID technology. Intelligently designed goods for everybody, as well as the inclusion of persons with disabilities in the city of the future as a smart city, are investigated.

2. Ambient Intelligence and Smart Environments

Computer devices become part of our life. This technology can be networked and used with the coordination of highly intelligent software to understand the events and relevant context of a specific environment and to take sensible decisions in real-time or a posteriori [2].
Appliances with the advancement of science have become smaller in size with more technology included, easy to be used providing to consumers benefits that years ago we had not imagined. Devices have become part of our body movements extending of our hand. Computing devices by minimizing their size to small chips can embodied to our society in all the environments public places and private places. Our inside home machines have been enhanced with the possibility of increased options making autonomous decisions.
Our home our personal in home or outside devices as coffee, laundry machines, refrigerator, cars have so many sensors and activators to anticipate or activate situations. The existing smart systems easily provide us the opportunity to turn on and off lights in our house scheduling the time or using our voice, order shopping from super markets, be moved easily to another place. We can easily realize how computing will affect all environments in the near future [3].
These computing devices will have to be coordinated by intelligent systems that integrate the resources available to provide an “intelligent environment”. This confluence of topics has led to the introduction of the area of “Ambient Intelligence” (AmI) [4]: It is AmI which brings together networks, sensors, human-computer interfaces, pervasive computing, artificial intelligence (including robotics and multi-agent systems) and many other areas to provide flexible and intelligent services to users acting in their environments [5]. It is necessary the existence of sensible-intelligent system. Being sensible demands recognizing the user, learning or knowing her/his preferences, and the capability to exhibit empathy with or react to the user’s mood and the prevailing situation, i.e, it implicitly requires for the system to be sensitive [6]. With the use of term Smart Environment, we make clear the existence and the necessity to support the post system with (sensors, actuators and networks).

3. Smart Cities

Nowadays, cities are getting bigger and more numerous, with a lot of urban problems like air pollution, traffic congestions, lack of resources [7]. Societies unfold all their problems and citizens all their peculiarities and social values become more sensitive and seeks solutions more than ever. The need for societies to respond to both the changes of nature and the demands of citizens and socio-economic change in general leads cities become Smarter. Several working definitions have been put forward and adopted in both

practical and academic use [8]. The development of eight factors as technology, policy, built a nature environment etc., identify the term Smart; it depends from local governments how will be envisioned the term and which factor will develop firstly in order to become a Smarter City [9]. Giffinger et al. (2007) suggest a smart city framework consisting of six main components (smart economy, smart people, smart governance, smart mobility, smart environment, and smart living) [10]. Local and international accessibility are important aspects of Smart Mobility as well as the availability of information and communication technologies and modern and sustainable transport systems [10].
In a Smart City the implementation of information technology includes all the devices, networks, procedures that are utilized in the information and telecommunication and technology (ICT) fields to promote interaction amongst different stakeholders [11]. ICT infrastructure includes wireless infrastructure (fiber optic channels, Wi-Fi networks, wireless hotspots, kiosks) [12]. Smart object networks play a crucial role in making smart cities a reality [13]. According H. Chourabi etal., (2012) in Initiatives framework all factors have an impact to Smart City, Due to the fact that many smart city initiatives are intensively using technology, it could be seen as a factor that in some way influences all other success factors in this framework [14].

4. Smart Cities Augmented Reality and Internet of Things Improve Accessibility Paradigm of a smart store-RFID technology

In a Smart City the living conditions must serve all citizens equally [15]. Even if there are important accessing problems or insignificant issues, technology enables us to overcome obstacles to solve difficulties and to provide a society of equality and socially acceptable to all [16].
For people as wheelchair users who have limited independence in their everyday life and are not able to do shopping because of a limited ability to reach upper surfaces as the super market shelves Augmented Reality (AR) and Radio Frequency Identification (RFID) technology give the solution. The resulting experience is close to being able to browse a shelf, clicking on it and obtaining information about the items it contains, allowing wheelchair users to shop independently, and providing autonomy in their everyday activities [17]. The Smart Shelf enabled with the RFID into a super market area can easily provide the information to user when the labeled with RFID tags items change location. The RFID system is composed of electronic tags (attached to objects), a reader or interrogator and an Information System (IS) managing the system's operations [18]. It is a low cost, item level identification for products, being the best for IoT technology. On the shelf of a store every item can attached with RFID tags. In that way the interactive interface gives all the necessary information for the location of products into the market area, firstly indicating the products into the shop and secondly with another indicating the products on the selves.
According Z. Rashid. et al. (2017) paradigm a database within the IS stores information about each item including EPC (i.e. ID code), an image (i.e. cover) and all available information on the package. An inventory list, consisting of all objects' EPCs, together with their approximate locations is periodically uploaded to the database from the RFID system [19]. With the help of a mobile device, a touch screen which is connected with the IS and the real conditions of shelfs and their products all the groups of users disabled or not can easily find the required product inside the store. Web interfaces present

information about location and existence of the products into the shop and on the shelves thanks to RFID update. A system including Augmented Reality using touch screen interfaces and with on time realization of Internet of Things (IoT) technology can give solutions [20]. Technology helps to avoid stigma between able-bodied and wheelchair users. The combination of both AR and RFID is a unique step towards a practical solution to wheelchair users to shop independently and autonomously in the context of Smart Cities [21]. Shopping independence help not only people with disabilities but all shoppers.

In Smart Cities people with disabilities as motor disabled people have solutions to their problems through the Internet of Things (IoT) which connects the physical objects with the people, the Internet, information systems and among themselves [22]. There are researchers showing the assisting technology providing solutions to everyday life of disabled people as people with motor disability that help them to escape of their problem. There are a big variety of people who interact only with head or mouth, with robotic arms, with smart-phone [23].

5. Smart - intelligent designing system a basic for design for All

The Internet-based Systems aid to design applications that could easily help designers to design for All. At the level of design and modelling, recent advances in areas, such as adaptive user interfaces and software agents, provide solutions to conceptual and engineering issues related to design-for-all [24]. The need to redesign products is because of the rapidly change of our society, the increasing changes of population needs the diversity of activities, the required needs of the planet and the expansion of technological platforms [25]. Designing products using technology in a way to make the life easier for people with a disability we succeed to have applications widely accessible and to incorporate people requirements. For that reason, we have to evaluate the existing design environment use the modelling theories and methods include all the required recommendations for disabled users and develop the areas of electronic virtual use and the redesign of the existing practical everyday issues.

6. What is smart, intelligent a package paradigm. From the conceptual model to the smart and to the intelligent

New technologies familiar from the Smart City and Home as intelligent products and packaging, radiofrequency identification (RFID), we find them in Smart Products and in their packaging replacing the traditional packaging. The new intelligent packaging incorporates microchips, antennas and new materials such as thermochromic inks and various indicators that track changes and provide real-time responses and wireless communication capabilities [26]. From the literature packaging is defined as the mean to protect the product from environmental changes and distribution, with some graphical information in a way to communicate with the consumer and provide product information. The conventional packaging is the packaging that can contain material product, to protect it, to communicate with the consumer, but also to have compatibility capabilities of the construction material (container), with the content.

Smart packaging has simple changes; offers better graphics, bar code, better protection. Smart packaging monitors changes in a product or its environment (smart), but also acts after these changes in active behavior. It uses chemical sensors or biosensors to monitor the quality and safety of products (food-drugs, etc.) from the producer to users-customers-consumers. But along with smart packaging, it uses a variety of sensors to monitor the quality and safety of sensitive products. This is done through the detection of chemicals, such as pathogens, leaks, carbon dioxide, oxygen, pH level, time or temperature, which means freshness in a food products e.tc. Intelligent packaging has a communicative character with printed electronics and the existence of microchips, antennas and batteries. RFID tags provide a product description and allow content identification [26].

Intelligent packaging allows a product to be located (tracked) and detected throughout its life cycle (time) and the analysis and control of the environment inside or outside the package, to inform the manufacturer, retailer or the condition of the product at all times. It also gives information about the place and time consumed, mainly for the smooth operation of the supply chain. We have capabilities for managing metadata, but also in recording any counterfeit products (smuggled).

7. Supporting people with disabilities in their Daily Life in a Smart City

With the evolution of technology smart cities is the future of urban development. To define what smart city is becomes difficult because the ideal city has to be perfect in all the aspects and that is difficult. From the citizen point of view the smart city has to be accessible and friendly to users in a natural way without feeling like technology is actual. Innovations technology became the tool for independence [27]. It would be very promising and comforting for this technology to promote growth and development. A completely hands-free city with stores, schools, hospitals accessible through mobile applications, automatic doors and voice control [28].

Imagine a city where a person in a wheelchair or pushing a stroller can chart a route to the local park using curb cuts and avoiding barriers; wirelessly log onto the park itself and receive notifications of upcoming park events, and perhaps even participate in an interactive lesson on the trees and flowers currently in bloom; where refrigerators will provide alerts of any diminishing essentials so that caregivers can adjust their grocery list before they visit their parents' home [29].

Toronto is slated to be the first smart city in Canada thanks to Alphabet's (Google's owner) Sidewalk Labs; The project's mission is to "blend people-centered urban design with cutting-edge technology to achieve new standards of sustainability, affordability, mobility, and economic opportunity [30]." When the U.S. Department of Transportation held its Smart City Challenge in 2016, one of the application requirements was to increase inclusivity, including for people with disabilities [31].

Kansas City has launched a network of interactive digital kiosks that can collect and share information. According (Bowman E. 2017) the kiosks have an audio jack so that visually impaired users won't feel self-conscious about listening to information available in the text-to-voice feature. They simply plug in their own headphones and listen privately [31]. Microsoft has created the Smart Cities for All proposal with G3ict and World Enabled and try to affect investors to become reality defining accessibility.

8. Conclusions

Several task-based scenarios for people with mild cognitive disabilities have been developed like for example preparing a meal, medication or handling daily activities like morning routines. Our longterm vision and our main purpose must be focused on promoting an accessible intelligent environment based on design for All, using architecture and innovative solutions, positioning citizens play the leading role in the intelligent society. Very significant is the role of Smart technology in many sciences. Artificial Intelligence and Bionics can offer to the disabled humans to experience things that they would have not otherwise been able to if surrender to the disability state [32]. By evaluating present technologies and needs, the author tackles themes that will convert modern living into the life of the future, simplifying the daily lives of individuals with different requirements and impairments.

References

[1] A. B. Salem, A. Nikitaeva, Knowledge Engineering Paradigms for Smart Education and Learning Systems, Conference: 2019, *42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO),* 2019 DOI: 10.23919/MIPRO.2019.8756685
[2] J. C. Augusto, H. Nakashima, H. Aghajan, Ambient Intelligence and Smart Environments: A State of the Art, 2010 *Handbook of Ambient Intelligence and Smart Environments* ,Springer, (2010),p2
[3] C. Papakostas, C. Troussas, A. Krouska, & C. Sgouropoulou, User acceptance of augmented reality welding simulator in engineering training, Educ Inf Technol (2021), https://doi.org/10.1007/s10639-020-10418-7
[4] J. C. Augusto, H. Nakashima, H. Aghajan, Ambient Intelligence and Smart Environments: A State of the Art, *Handbook of Ambient Intelligence and Smart Environments* ,Springer, (2010), p3
[5] C. Troussas, A. Krouska, C. Sgouropoulou: Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education. *Computers & Education*, (2020), p144
[6] J. C. Augusto, H. Nakashima, H. Aghajan, Ambient Intelligence and Smart Environments: A State of the Art, *Handbook of Ambient Intelligence and Smart Environments* p4, Springer, 2010
[7] A. Krouska, C.Troussas, C. Sgouropoulou: Usability and Educational Affordance of Web 2.0 tools from Teachers' Perspectives. *PCI 2020-24th Pan-Hellenic Conference on Informatics*, (2020), Pages 107-110, November 2020, https://doi.org/10.1145/3437120.3437286
[8] H. Chourabi, T. Nam, S.Walker, J. R. Gil-Garcia, S. Mellouli, K. Nahon, T. A. Pardo, H. J. Scholl, Understanding Smart Cities: An Integrative Framework, *45th Hawaii International Conference on System Sciences*, 2012
[9] A. Krouska, C. Troussas, C. Sgouropoulou: A Personalized Brain-Based Quiz Game for Improving Students' Cognitive Functions, *BFAL 2020*, pp 102-106
[10] R. Giffinger, C. Fertner, H. Kramar, R. Kalasek, N. Pichler-Milanović & E. Meijers, Smart Cities: Ranking of European Medium-Sized Cities. Vienna, Austria: Centre of Regional Science (SRF), Vienna University of Technology, 2007
[11] C.Troussas, A. Krouska, C. Sgouropoulou, Towards a Reference Model to Ensure the Quality of Massive Open Online Courses and E-Learning. In: Frasson C., Bamidis P., Vlamos P. (eds) *Brain Function Assessment in Learning.* BFAL 2020, Lecture Notes in Computer Science, vol 12462. Springer, 2020 https://doi.org/10.1007/978-3-030-60735-7_18
M. Al-Hader, A. & Rodzi, (2009). The smart city infrastructure development & monitoring. Theoretical and Empirical Researches in Urban Management, Research Centre in Public Administration and Public Services, Bucharest, Romania, vol. 4(2(11)), (2009), pp 87-94,
[12] J. Vasseur, Smart cities and urban networks. In Vasseur, J. & Dunkels, A. (Eds.), *Interconnecting Smart Objects with IP: The Next Internet* (pp. 360-377). Burlington, MA: Morgan Kaufmann, 2010
[13] H. Chourabi, T. Nam, S. Walker, J. R. Gil-Garcia, S. Mellouli, K. Nahon, T. A. Pardo, H. J. Scholl, Understanding Smart Cities: An Integrative Framework, 45th Hawaii International Conference on System Sciences, 2012

[14] M. Poli, Greece, Tourism and Disability, Strategic Innovative Marketing and Tourism, International Journal of Strategic Innovative Marketing 8th ICSIMAT, Springer, 2019 https://doi.org/10.1007/978-3-030-36126-6_74

[15] M. Poli, Covid-19 Effects in People's Daily Life, Social Media's Role and the Power of Emotions and State of Mind, International Journal of Strategic Innovative Marketing 9th ICSIMAT, Athens, Greece, Springer, 2020 https://doi.org/10.1007/978-3-030-66154-0

[16] Z. Rashid, J. Melis-Segui, R. Pous, E. Peig, Using Augmented Reality and Internet of Things to Improve Accessibility of People with Motor Disabilities in the Context of Smart Cities, (2017), p. 248-250 doi: 10.1016/j.future.2016.11.030

[17] Z. Rashid, J. Melià-Seguí, R. Pous, E. Peig, Using Augmented Reality and Internet of Things to improve accessibility of people with motor disabilities in the context of Smart Cities, *Future Generation Computer Systems*, University of Dundee, (2017), pp 250-255 https://doi.org/10.1016/j.future.2016.11.030

[18] Z. Rashid, J. Melià-Seguí, R. Pous, E. Peig, Using Augmented Reality and Internet of Things to improve accessibility of people with motor disabilities in the context of Smart Cities, *Future Generation Computer Systems*, University of Dundee, (2017), p 255-259 https://doi.org/10.1016/j.future.2016.11.030

[19] C. Troussas, C. Sgouropoulou: Innovative Trends in Personalized Software Engineering and Information Systems - The Case of Intelligent and Adaptive E-learning Systems. *Frontiers in Artificial Intelligence and Applications 324*, IOS Press (2020), pp1-96, ISBN 978-1-64368-096-5

[20] Z. Rashid, J. Melià-Seguí, R. Pous, E. Peig, Using Augmented Reality and Internet of Things to improve accessibility of people with motor disabilities in the context of Smart Cities, Future Generation Computer Systems, University of Dundee, (2017) pp 258-261 https://doi.org/10.1016/j.future.2016.11.030

[21] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, Exploration of Augmented Reality in Spatial Abilities Training: A Systematic Literature Review for the Last Decade, Informatics in Education, no.1, (2020-2021), 107-130, DOI 10.15388/infedu.2021.06

[22] C. Troussas, F. Giannakas, C. Sgouropoulou & I. Voyiatzis , Collaborative activities recommendation based on students' collaborative learning styles using ANN and WSM, *Interactive Learning Environments*, 2020 DOI: 10.1080/10494820.2020.1761835

[23] P. Koutsabasis, J. S. Darzentas, J. Abascal, T. Spyrou, J. Darzentas, Designing Internet-based Systems and Services for All: Problems and Solutions, the 9th International Conference on Human-Computer Interaction), USA, 2001

[24] C. Troussas, A. Krouska, C. Sgouropoulou, I. Voyiatzis: Ensemble Learning Using Fuzzy Weights to Improve Learning Style Identification for Adapted Instructional Routines, Entropy 22(7): 735, 2020

[25] S. Nomikos, Smart packaging: new communications concepts and models, 33rd International Research Conference of Iarigai Germany, 2006

[26] F. Giannakas, C. Troussas, I. Voyiatzis, C. Sgouropoulou, A deep learning classification framework for early prediction of team-based academic performance, *Applied Soft Computing*, Volume 106, 2021, 107355 https://doi.org/10.1016/j.asoc.2021.107355

[27] A. Krouska, C. Troussas, C. Sgouropoulou, Applying Genetic Algorithms for Recommending Adequate Competitors in Mobile Game-Based Learning Environments. In: Kumar V., Troussas C. (eds) *Intelligent Tutoring Systems*. ITS 2020. Lecture Notes in Computer Science, vol 12149. 2020, Springer, Cham https://doi.org/10.1007/978-3-030-49663-0_23

[28] Smart Cities for All: A Vision for an Inclusive, Accessible Urban Future, AT&T https://smartcities4all.org/wp-content/uploads/2017/06/Smart-Cities-for-All-A-Vision-for-an-Inclusive-Accessible-Urban-Futur...-min.pdf

[29] M. Stangl, The Future of Smart Cities for People with Disabilities, 2017 website https://medium.com/gettecla/the-future-of-smart-cities-for-people-with-disabilities-8a54e1cfc2a0

[30] E. L. Bowman, Smart cities work for people with disabilities, ShareAmerica, 2017, website https://share.america.gov/smart-cities-work-for-people-with-disabilities/

[31] K. Elbehiery, H. Elbehoery. Artificial Intelligent Bionics; Disability to Exceptional Ability, The International Journal of Engineering and Science (IJES), (2020), vol 9, pp23-33, ISSN (p): 20-24-1805

Extended Reality

This page intentionally left blank

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210076

Virtual Reality Tool for Rehabilitation of Patients with Parkinson's Disease: A Conceptual Design Review

Josiah Cyrus BOQUE, Jaime CARO, Richelle Anne JUAYONG and Roland Dominic JAMORA
Department of Computer Science,
College of Engineering,
University of the Philippines – Diliman

Abstract. Parkinson's Disease is a progressive, irreversible disease that is only slowed down with the use of medications and therapy. These are the only way to slow down the progression of the disease to more severe situation. Currently, virtual reality is being use on games and entertainment. Some doctors prefer the use of virtual reality like Wii to help supplement the therapy done on rehabilitation. However, devices to cater virtual reality for mid-class to lower-class patients cannot afford such devices. An alternative device, usually accessible to everyone is proposed to cater these virtual reality applications that can help in therapy of Parkinson's patients. With that, the capabilities of VR can be accessible to more patients who cannot avail expense medication and devices for virtual reality.

Keywords. Virtual Reality, Parkinsons Disease, VR Rehabilitation

1. Introduction

In recent years, mental health and disorder has become a concern all over the world and is increasing. Instances of imbalances and involuntary seizures may show different symptoms enough for someone to be diagnosed with this disorder. One of these is Parkinson's Disease (PD). Up to this date, it is irreversible. However, the progression of symptoms can be slow down, and let the disease stop prevailing at the early stages through medication and continuous therapy[1].

There hasn't been a focus of research to this disease on slow-developing country like in the Philippines. Moreover, ratio of doctors and scientists that studies for treatment and core medication per population is relatively small. Hence, patients with PD endure on having the disease as they age, then eventually die bed-ridden [2]. Moreover, PD can be acquired at any age, but as they age, incidence of PD symptoms are also imminent. Risk of a person to have PD can be high due to different factors like race, demography, environmental noise and pollution, and others [1].

On this paper, problems on available treatment will be addressed, then a design of a proposed alternative solution using other means of Virtual Reality (VR) technology will be created to supplement and help catalyze the therapy treatment process done on PD patients. Hence, to help slow down the progression of symptoms in patients.

2. **Review of Related Literature**

2.1. Parkinson's Disease

Parkinson's disease (PD) is a nervous system disorder where the nerve cells that produces dopamine are being affected. Hence, signs of this disease can be early depicted by the following symptoms: muscle rigidity, tremors, and changes in speech and gait (walking posture). A part of your brain called the substantia nigra makes our dopamine, a chemical gives way for signals from neurons to our brain, especially when we think and do something. This is also the chemical that manages how smooth and evenly we move. But having this disorder, we are tend to become disoriented on how we navigate, do our activities or even on the way we talk [3].

PD severity can be measured via the Hoehn and Yahr staging of PD based on the study of Hoehn and Yahr that focused on factors within and around a person that may have a significant effect which triggers the disease's severity, or the symptoms patients may show. Some of these factors include age, physical activities, hobbies and etc. And when these factors are identified, the patient undergoes a test that involves certain movement routine [4]. Using this scale, they found out that progression for each patient to jump to a higher stage takes 2-5 years even if patient is on rehabilitation [5]

Balance and posture are usually lost in PD patients. Tests such as the Berg Balance Scale (BBS) are used to assess patient. The test includes scoring of 14 movement items, with 4 being the highest score for each item. Each item is scored based on assessment if the patient cannot do the task or has a struggle, or pain doing it. Once the scores have been summarized, the accumulated points will determine the condition of the patient [6].

In the present, drug medication are also still used as a part of the treatment. Usual prescription of doctors may cost patients too much. Computation of each prescription of doctors to the patients can show that each patient have to spend of at least $6.00 just for the medicines, or at an average of $6,000 or PhP 300,000 annually. Dosage of medicines may increase depending on how symptoms persist with the patient. [2].

2.2. Virtual Reality and its use to Parkinson's Therapy

Virtual Reality(VR) is a computer-based technology that immerses its users to replicated environments, natural or fictional, as if the user is experiencing it for real. Interaction within the virtual environment can be experience depending on what material is being used. This technology brings us to places or let us do activities that we cannot do in real world. Virtual Reality nowadays is becoming prominent among people, since most mobile devices are now equip with gyroscope or gyro sensors which stimulates the phone to feel heights and direction does the phone is facing towards. Some equipments that are used in this technology are: the computer that renders the visual interface of the environment it want to portrait; the output monitor through a projector, a VR goggles, or digital display (Simulated Television, etc.) [7].

VR as a clinical therapy tool has been found to be effective in improving mobility of patients undergoing such therapy like for cerebral palsy patients. With their 59 participants, improvements with their performance is noticeable as seen in Figure 1.

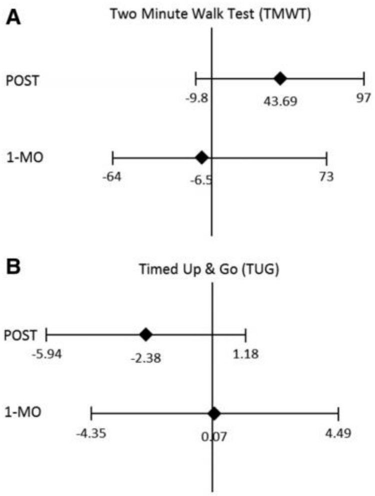


Figure 1. Scores for the Post assessment (POST) and after-a-month assessment (1-MO) for TMWT and TUG

Ninety-five percent confidence intervals and effect size (black diamonds) for the difference in improvements immediately after the final training session (POST) and 1 month after the cessation of training (1 MO) are shown. For the Two-Minute Walk Test (TMWT; A), the effect size to the right of the zero line indicates an improvement, whereas for the Timed Up and Go test (TUG; B), the effect size to the left is indicative of improvement in favor of the treatment group. Moreover, they even did another intervention a month after their post intervention with the patient. They found increase of performance from both of their experimental and controlled group. This only shows that VR therapy has no significant difference with the therapy done on rehabilitation centers. But with the advantage that it can be done at home and by their own means, a lot of patients can now have therapy at their own reach [8].

Also, [9] cited the use of Wii Fit, a VR gaming application on Wii, is being use for therapy sessions of stroke patients. Activities included during the intervention involves arms and shoulder movements, flexion and extension of elbow, wrist and finger stretches. With their 22 randomly selected patients, their performance where assessed using the Wolf Motor Function Test. They say that most of their patients have improved significantly in terms of muscle and joint flexes of arms and shoulder. This tells us that VR therapy can really boost patients improvement while enjoying therapy through this kind of technology.

2.3. Wii Technology

Wii, a VR technologies, has been used as a tool for Parkinson's Disease therapy with the application called Wii Fit created back 2012. This application is a group of games that are made to train couple of patients' body parts including balance and standing posture test when moving the arms, legs, hips, bending sideways and other extremities.Some of this games include Ski Slalom, Tightrope walk. This is tested with 50 patients, where a study group, 27 of them were able to rehab with the application, while the other 23, the control group, undergo the conventional therapy. For their preliminary experiment, each patient undergone pre-assessment and post-assessment using BBS and the Dizziness Handicap Inventory (DHI). DHI is a 25 questionnaire-type assessment tool for balance where patients are scored between 4 (affirmative), 2 (sometimes), 0 (negative).

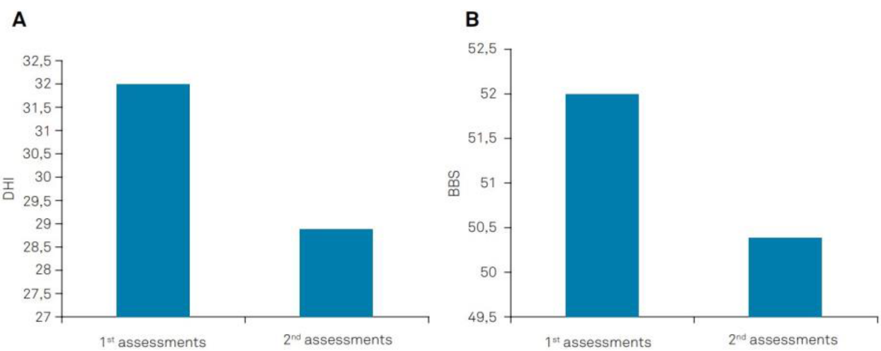


Figure 2. Results of 1st and 2nd Assessment for (A) DHI and (B) BBS.

Based on Figure 2, patients that undergone therapy with Wii Fit shows lowers scores on both DHI and BBS. This indicates that Wii Fit indeed helps to lessen and tremors that patients experience. Moreover, it help the patients improve their body balance and walking posture with reduced fall-risks and increase self-confidence and independence. Test results show that the performance of the study group when performing the games on the Wii Fit were better than that of the control group. This tells us that such use of technology, like the Wii Fit can be a reliable and valid measuring tool to train and assess patients' posture and balance improvements. Moreover, patients experiencing it for the first time, found the technology motivating and exciting to use every time they will do rehabilitation with the technology [10].

Also another paper [11] showed great improvements on patient when they used Wii Fit as part of their therapy session. They created the study as a pilot study to check the significant effect of training when using such gadget. According to the authors, for 10 weeks, 8 subjects or patients with ages between 30-60 years old underwent their 30min sessions per week which includes yoga, balance, aerobic and strength Wii Fit™ activities. Each activity are also encourage to be modified with increasing difficulty to increase their performance and improvement.

Table 1. Effect of Wii Fit™ intervention

Measurement	**Pre-intervention**	**Post-intervention**
6-min walk test (m)	650.8 (±70.9)	683.6 ±85.8)
Step test (number)		
Right	21.75 (±3.4)	22.38 (±3.38)
Left	22.5 (±3.34)	23.13 (±3.36)
Timed Up and Go test(s)	4.93 (±0.76)	5.00 (±0.73)
Timed Up and Go cognitive test*	5.39 (±0.97)	4.94 (±0.78)

As seen in Table 1, it shows us the scored they gathered from different test before and after the intervention. They used the 6-mins walk test, Timed Up and Go, Step test and other test as metrics to measure improvement for the patients. Based on the scored, the use of Wii Fit™ help in improving each patient's strength and balance after training with it. They have seen big improvement on their functional balance measures, flexibility, reaction time or somatosensation.

With these results, the authors are not confident to say that Wii Fit™ can effectively help patient to have better conditions so they suggest to increase the number of subjects which includes healthy subjects so they can compare how well this technology can go further to improve the performance of the patients as same as with those that are healthy.

2.4. Motor Imagery

Moreover, patient assessment using Motor Imagery (MI) together with VR was also studied to improve patients' walking gait, balance and mobility. They collated various articles and papers that indicates improvement on patients with freezing of gait (FOG). These collectively concluded that MI can really help improve patients cognitive and visual learn. And with the help of immersion to VR, patients greatly improve their locomotor movements even more [12].

2.5. Visuo-motor Learning

Also, a study showed what with the use of visuo-motor learning through immersive 3D VR can help improve their conditions. They immersed patients from different age group. First they noticed that elders and young subjects have different pace movement. However they are able to observe same result that as they go on with the trials, both age group manage to improve their performance, and lessen movement errors. Second, they found out that PD patients, as compared with their healthy proponents, shows slower comprehension with training using biaxial visuo-motor discordance. Also, they have more errors with the proper movement corrections. However, trends of improvement from all subjects shows decreasing amount of movement errors as number of trials are being done.

In general, these applications created for patient rehabilitation are found efficient since these can really help the patients improve their walking gait, balance to gain independence, and can somehow do their daily living activities independently. However, to warrant all claims of improvement when using VR for PD patients' rehabilitation, further studies are still needed to be done. Since there are some proven concepts of improvement, other capacities of VR on brain rehabilitation can now be explored, not just for PD patients' use but also for other neuro-diseases. [13,14].

3. **Proposed Design**

From the previous sections, there are proven concept presented when using VR as therapy for PD patients. However, they are only limited to be used in rehabilitation centers. So to cater patients from places far from these centers, we can create VR applications that can be placed on smartphones and use it as daily therapy for the patients even if they are at staying at their homes.

With Google, they extended the accessibility of VR for making android phones capable of running these applications and programs for a Head-mounted Device (HMD) . With the use of gyroscopes, the phones can manage to sense device orientation, proximity and elevation [15].

Considering the accessibility of android phones to the public that can be a HMD, we can use this as an advantage to make applications for therapy . With that, an application can be created to let its users do the activities similar with what they do in rehabilitation center. If in therapy sessions with Wii Fit, they let the patients bend their body sideways, there can be activities in VR Android phones that will let the users or patients bend their bodies to perform the activities indicated. Also, this application can include games that will test patients' balance, stature strength, and their ability to walk. These are the aspects of the body we can target for their therapy.

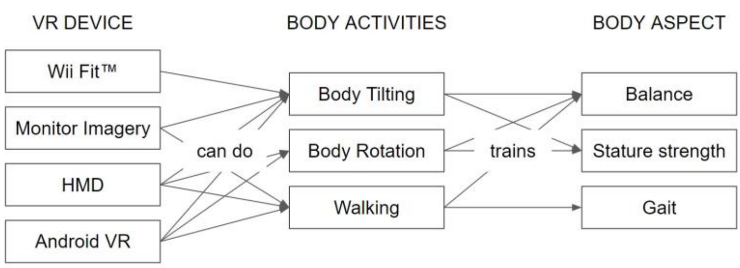


Figure 3. Mapping of activities that can be done with each VR Device. Also the mapping to the body aspects that each activity targets to train

As seen in Figure 3, there are similarities on activities that all VR device can do like from body tilting to actual walking. For Parkinson's Disease, it is very helpful to keep patients moving through walking and limbs movements. These decrease the prevalence of having frozen joints and lessen their gait. So forcing the patients to have fun with VR activities will really be helpful in keeping them moving. Since we target those body parts that other VR devices do too, we can expect that it can also improve the balance, stature strength and gait in training them to walk.

3.1. Patient Specification

For the research, it is recommended to have at least 16 subjects with Parkinson's disease for a pilot test. However there are some limitation about who we will be needed. These patients must :

1. have Parkinson's Disease of stage 2, 2.5, or 3 (according to Hoehn and Yahr Scale);
2. be able to sit and/or stand even if with or without assistance;
3. be able to walk of at least 4 to 6 inches per pace;
4. not have difficulty(no pain) in moving upper(arms, hands, shoulders, etc.) andlower extremities (limbs, feet, legs, etc.).

3.2. Testing Process

The testing phase will have two parts. First, after the application is created, a dummy testing will be done to test it for any bugs that a healthy individual can detect or notice while using the current application. After assessing it through a Usability test and Reliability test, the application will then be modified to cater tester's issues with the application, then another testing will be done using the same assessing tests to gather other hanging issues about the application.

Second part of the testing phase involves the participation of random select patients as subjects prior to our set limitations. A pre-assessment will be done to all of them using the application. After which, half of the subjects will undergo rehabilitation with the use of the application for 10 mins. in a span of 4 weeks schedule. Same through with the other other subjects, but they will not be using the application during the testing schedule. After 4 weeks prior to the start of this testing phase, a post-assessment will be done in terms of their performances according to doctor's perspective.

3.3. Risk and Safety Analysis

Each of our subjects are assigned a personal assistant(PA) since from the start of their rehabilitation. To be able to successfully apply this testing, coordination with their PA on how to use and guide the subjects with the use of the device and application should be observed. Careful handling of the patient while the testing is on going will dictate safety in using this technology.

4. **Conclusion**

VR has a lot of potential use with its wide variety of activities that can be done and enjoyed according to the literature. Because of the fluidity that the activities that can be ran across different devices, it is surely that application for android phones that will target similar body aspects is also promising. With the proposed use of HMD on smartphones, use of VR for therapy can open accessibility to PD patients living far from rehabilitation centers.

To really test out its capabilities, clinical trials can be done to test the significance of using VR as a supplementary therapy for our patients. Also, Augmented reality (AR) is noticeable nowadays. It it because of its capabilities to place the virtual objects into the real world with the use of mobile devices. This connects the bridge for innovations with AR and VR for health and other life aspects.

Acknowledgement

The researchers would like to acknowledge the Engineering Research and Development for Technology under the Department of Science and Technology of the Philippines for funding this research and submission for this conference.

References

[1] L.C. Tan, Epidemiology of Parkinson's disease, Neurology Asia 2013 (2013), 231–238.
[2] R.D.G. Jamora and J.M. Miyasaki, Treatment gaps in Parkinson's disease care in the Philippines, Neurodegenerative Disease Management 7(4) (2017), 245–251. doi:10.2217/nmt-2017-0014.
[3] Parkinson's Disease: Symptoms, Causes, Diagnosis, Treatment, WebMD, 2017. https://www.webmd.com/parkinsons-disease/parkinsons-disease-overview1.
[4] E. Modestino, Hoehn and Yahr staging of Parkinson's disease in relation to neuropsychological measures, Frontiers In Bioscience, Landmark 23 (2018), 1370–1379. doi:10.2741/4649.
[5] M.M. Hoehn and M.D. Yahr, Parkinsonism: onset, progression, and mortality, Neurology 17(5) (1967), 427–427–. doi:10.1212/wnl.17.5.427.
[6] F.B. Horak, D.M. Wrisley and J. Frank, The Balance Evaluation Systems Test (BESTest) to Differentiate Balance Deficits, Physical Therapy 89(5) (2009), 484–498–. doi:10.2522/ptj.20080071.
[7] Step into a new world of Virtual Reality(VR). https://www.completegate.com/2016070154/blog/virtualreality-explained.
[8] D. Mcewen, A. Taillon-Hobson, M. Bilodeau, H. Sveistrup and H. Finestone, Virtual Reality Exercise Improves Mobility After Stroke, Stroke 45(6) (2014), 1853–1855–. doi:10.1161/strokeaha.114.005362.
[9] V. Fung, A. Ho, J. Shaffer, E. Chung and M. Gomez, Use of Nintendo Wii Fit in the rehabilitation of outpatients following total knee replacement: a preliminary randomized controlled trial, Physiotherapy 98 (2012), 183–188.
[10] J.C. Nitz, S. Kuys, R. Isles and S. Fu, Is the Wii Fit™ a new-generation tool for improving balance, health and well-being? A pilot study, Climacteric 13(5) (2010), 487–491. doi:10.3109/13697130903395193.
[11] A. Mirelman, I. Maidan and J.E. Deutsch, Virtual reality and motor imagery: Promising tools for assessment and therapy in Parkinsons disease, Movement Disorders 28(11) (2013), 1597–1608–. doi:10.1002/mds.25670.
[12] K. Dockx, V.V.D. Bergh, E.M. Bekkers, P. Ginis, L. Rochester, J.M. Hausdorff, A. Mirelman and A. Nieuwboer, Virtual reality for rehabilitation in Parkinsons disease, Cochrane Database of Systematic Reviews (2013). doi:10.1002/14651858.cd010760.
[13] N.-Y. Lee, D.-K. Lee and H.-S. Song, Effect of virtual reality dance exercise on the balance, activities of daily living, and depressive disorder status of Parkinson's disease patients, Journal of Physical Therapy Science 27(1) (2015), 145–147–. doi:10.1589/jpts.27.145.
[14] A. Faisal, Computer science: Visionary of virtual reality, Nature 551(7680) (2017), 298–299–. doi:10.1038/551298a.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210077

Toward Personalizing Alzheimer's Disease Therapy Using an Intelligent Cognitive Control System

Hamdi BEN ABDESSALEM[1] and Claude FRASSON[2]
Département d'Informatique et de Recherche Opérationnelle
Université de Montréal, Canada

Abstract. Subjective cognitive decline is an early state of Alzheimer's Disease which affects almost 10 million people every year. It results from negative emotions such as frustration which are more present than healthy adults. For this reason, our work focuses on relaxing subjective cognitive decline patients using virtual reality environments to improve their memory performance. We proposed in our previous work a neurofeedback approach which adapts the virtual environment to each patient according to their emotions using a Neural Agent. We found that the Neural Agent can adapt the environment to each participant but have limitations. This work is a continuation of our approach in which we propose a *Limbic Agent* able to monitor the interactions between the Neural Agent and patients' emotional reactions, learn from these interactions, and modify the Neural Agent in order to enhance the adaptation to each patient with an Intelligent Cognitive Control System. Our goal is to create a system able to support the Limbic System which is the main area in charge of controlling emotions and creating memory in the human brain. We used data collected form our previous work to train the Limbic Agent and results showed that the agent is capable of modifying the weight of existing rules, generating new intervention rules, and predicting if they will work or not.

Keywords. EEG, Virtual Reality, Limbic System, Intelligent Agent, Neurofeedback, Alzheimer's Disease, Intelligent Cognitive Control System.

1. Introduction

Alzheimer's disease (AD) is one of the most crucial diseases of our century affecting millions every year. AD has a direct negative effect on memory and cognitive functions which trigger a negative effect on the emotions. Negative emotions such as anxiety, frustration, and apathy are common in AD patients which reduce their wellbeing significantly [1]. This disease affects more commonly people aged 65 and older with a vulnerability increasing with age [2]. Alzheimer's disease is progressive, and its earliest phase is characterized by a state of subjective cognitive decline (SCD).

Which makes this disease important is that no cure has yet been found, although, some pharmacological interventions may reduce the symptoms. Considering that AD patients experience negative emotions, non-pharmacological treatments which aim to

[1] hamdi.ben.abdessalem@umontreal.ca
[2] frasson@iro.umontreal.ca

reduce these negative emotions could be a good alternative to reduce AD symptoms and promotes cognitive performance [3,4].

Virtual Reality (VR) has proven to be efficient in treating certain phobias, for instance, spider's phobia [5], and can be used to calm negative emotions. The immersion provided by VR gives a feeling of safety and encourages imagination. In addition, the patient is isolated from external distraction which increases the effect of the environment presented in VR and reduces the possible negative reactions caused by external factors, electroencephalography (EEG) devices help track the emotional states of AD patients interacting with VR environment.

In a previous work [6], we presented a novel approach which uses VR to relax SCD patients, reduce their negative emotions and improve their cognitive performance. We also proposed to use a Neural Agent which adapts the relaxing VR environment and optimize its effect. We conducted experiments and results showed that the approach effectively reduces negative emotions and improve memory performance. In addition, results showed that the Neural Agent is able to optimize the relaxing environment in order to personalize it for each participant by triggering intervention rules on the environment according to patients' emotions.

We also found that the modification performed by the Neural Agent does not always work as intended, some intervention rules leading to another emotional state not targeted by the rule.

In this work, we aim to verify the following hypotheses: H1: **it is possible to learn from the interactions between the neural agent and the patient's emotional reactions in order to generate new intervention rules** and H2: **it is possible to predict the impact of the new generated rules on the patient's emotions.**

The rest of this paper is organized as follows. In section 2, we present an outline of the related works. Section 3 presents our approach. In section 4 we detail the experiments. Finally, in section 5 we present and discuss the results.

2. Related Works

2.1. Brain Assessment

Due to its low cost and non-invasiveness, EEG represents a good equipment for brain assessment. In the field of emotion recognition, Liu et al. [7], proposed a multi-level features guided capsule network for multi-channel emotion recognition based on EEG signals. Their framework can identify intrinsic relationship among various EEG channels effectively. Moreover, Zhu et al. [8], analyzed EEG signals in order to explore social emotion perception and emotion classification. They designed a classifier based on convolutional neural network for social emotion classification using EEG. While in the field of psychology, Li et al. [9], proposed a method for quantitative evaluation of people's Big Five personality using EEG. They developed machine learning models in order to predict the user's personality based on their EEG signals while presented to emotional video clips. In the AD and mild cognitive impairment (MCI) detection, Oltu et al., [10] developed a machine learning approach that discriminate AD, MCI and healthy control individuals using EEG signals.

2.2. Intelligent Agents and Adaptation

Adaptative systems are often used to personalize experiences for each user. In the field of video games, researchers proposed to adapt a video game in real-time [11]. Using intelligent agents, they track in real-time the player's cognitive and emotional states, then adapt the video game's parameters according to these states. Furthermore, Spencer Rugaber et al. [12] created a model-based technique for self-adaptation in game-playing agents using intelligent agents. Doctor et al. [13] applied intelligent agents in the field of neuroscience. They used adaptation rules to facilitate learned behaviors' tuning and behavioral changes' tracking in the monitoring of dementia patients living in their own homes. Peña et al. [14] used intelligent agents in the field of learning. They adapted their system to learners by using an adaptation technique which focuses on the navigation tools, navigation strategies and the didactic contents.

2.3. Virtual Reality and Relaxing Environments

Recently, VR started gaining interest in many fields due to its remarkable advantages. Its main characteristics are immersion, sense of presence, and interactivity [15]. In fact, VR tricks the users' mind and makes them believe they are in a real world and thus promotes their performance [16]. This technology is applied in the field of psychology. Dana et al. [5] used it to as treatments for certain types of phobias. They created virtual environments in order to simulate situations and thus help patients face their fear. Furthermore, Lorenzo G. et al. [17] used VR in order to enhance the emotional skills of children diagnosed with autism spectrum disorders. Moreover, Coyle et al. [18], used it in the cognitive training. In fact, they assessed the efficacy of computerized cognitive training and virtual reality cognitive training programs for individuals living with MCI or dementia.

3. Our Approach: Intelligent Cognitive Control System

We propose an Intelligent Cognitive Control System (ICCS), the goal of an ICCS is to detect emotional and cognitive state of the user, alleviate negative emotions and increase positive emotions, using strategies to adapt the virtual environment in order to create the best favorable conditions in order to improve cognitive process. The ICCS is composed of a Savannah VR environment, Neural Agent and Limbic agent. In this section, we detail Savannah VR, Neural Agent, and Limbic Agent which compose our ICCS.

3.1. Savannah VR

We developed a Savannah VR environment using Unity 3D software in order to relax SCD patients. This environment provides is a therapeutic experience aiming to improve the negative state of SCD patients. The patients follow an avatar walking through a savannah and speaking in a soft and relaxing voice. Combined with VR techniques, this environment offers an immersive experience to the participants, so they feel like in a real savannah. The environment was developed [19] taking into consideration several parameters able to calm the patients, for instance the dominant color is warm, the animals are calm, and their movement is slow. A soothing piano tune is played as music background. Figure 1 illustrates a screen capture of Savannah VR.

Figure 1. Screenshots of Savannah VR [19]

Participants automatically follow a gazelle that moves along a precise path with breaking points. They are free to look around them by rotating their head in 360 degrees. To avoid nausea caused by movement in virtual reality, users follow the gazelle at low speed. The animal is in front of them to imitate a third-person view that is less likely to cause motion sickness [20].

Our therapeutic approach proposes that the environment's parameters have to be modifiable in real-time. Thus, functions have been implemented that triggers the modification of the environment's parameters.

One of the important parameters to be modifiable is the volume, we must choose the volume carefully; too high volume can cause noise pollution [21]. Another adjustable parameter is the color and intensity of the light because light influence perception and decision-making process [22]. The color can also relieve stress more quickly [23].

As an environment with more trees can also relieve stress more quickly and effectively [24], it is therefore important to control dynamically the number of trees in the environment. The number of animals can also be decreased, and the sky colors can be changed to have a soothing sunset.

3.2. Neural Agent

The goal of our work is to optimize and adapt the therapeutic environment to each participant, we propose to use the Neural Agent developed in [11]. This agent is able to modify the parameters of Savannah VR detailed in the previous section based on intervention rules triggering a neurofeedback approach.

This agent possesses a rules base containing multiple intervention rules. The Neural Agent operates by tracking the emotions of the patients and modifies Savannah VR accordingly using intervention rules. Giving the main emotion of the patient, the agent selects a rule from this rules base and modifies the VR environment to adapt it to each patient. The structure of one intervention rule is as following: Rule id, Triggering Emotion, Target Emotion, Action on VR, Weight.

After each intervention, the agent observes the emotional reactions of the patient to check whether or not it does modify the emotional state as desired and updates the weight of the intervention rule.

The Neural Agent is composed of two modules: "Analysis, decision and action" and "Rules base manager" as well as two databases: "Decisions base" and "Rules base".

Figure 2 illustrates the architecture of the Neural Agent and its interactions with the Emotions module and Savannah VR.

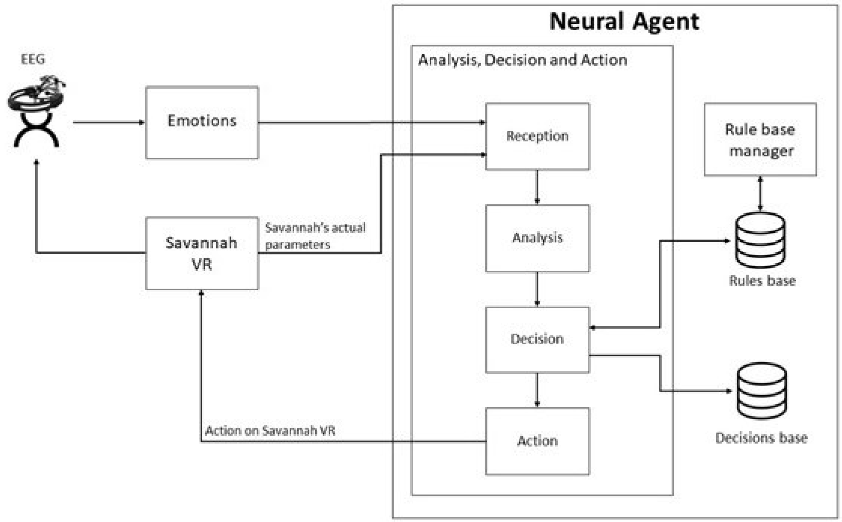


Figure 2. Architecture of the Neural Agent [6].

The "Analysis, Decision and Action" handles the data reception, the data analysis, the decisions to make and the actions to perform on Savannah VR. This module stores the agent's decisions in addition to the user's emotional reaction in the "Decisions base". The structure of a decision in the base is as follows: "Participant Id, Rule Id, Triggering Emotion, Target Emotion, Action on VR, Emotion_1 before Intervention, ..., Emotion_n before Intervention, Emotion_1 after Intervention, ..., Emotion_n after Intervention, Worked or Not?"

The "Rules base manager" handles the creation, modification and deletion of the intervention rules and stores them in the "Rules base".

3.3. Limbic Agent

The results of our previous work were satisfying but its cognitive capabilities were limited. In order to improve this system, we are interested on how the human brain works and specially the part of the brain which handles the emotions and the memory. We found that the Limbic System is the one which handles the emotions in the brain [25]. Sometimes, the Limbic system cannot regulate the emotions and access to memory. Thus, we propose to create a Limbic Agent which allows our system to assist the human Limbic System by modifying the emotions.

The goal of the Limbic Agent in our system is to observe and analyze the interactions between the Neural Agent and the emotional reactions of the patients, then tries to optimize it and improve it. Thus, it should learn from these interactions, create new intervention rules if needed, and predict its impact on the patient.

The Limbic Agent is composed of two modules "Observation and Data preparation" and "Learning and prediction" in addition to one database "Knowledge Base". Figure 3 illustrates the architecture of the Limbic Agent and its interactions with the system.

The role of "Observation and Data preparation" is to collect information from all system components through the Neural Agent's Decisions base. This module manages the reception of data and its preparation in the correct format, and sends these data to the "Learning and prediction" module.

The "Learning and prediction" module is composed of four components 'Analysis', 'Learning', 'Rule creation' and 'Prediction'. The role of 'Analysis' component is to receive structured data, analyze which rule works and which rule did not work according to multiple executions of the same rule, analyze the patient's emotional reaction, and change the weight of rules if needed. Then it transmits the data to the 'Learning' and 'Rule Creation' component.

The role of 'Learning' component is to learn from the execution of the intervention rules and the resulting emotional reaction. It creates a prediction model and transmit it to the 'Prediction' component. It stores the information in the 'Knowledge base' so more information it gets, more the knowledge base growth and more its model will be precise.

The 'Rule creation' component role is to analyze the different emotional impacts of each rule. It does not check if the rule worked as expected or not (this is the role of 'Analysis component'), however, it verifies if the multiple executions of one rule generates a different emotional reaction. Then, a new intervention rule is generated and sent to the 'Prediction' component.

The role of the 'Prediction' component is to predict the emotional impact of a new intervention rule just created by the 'Rule Creation' component using the learning model created by the 'Learning' component. Then, it inserts the new intervention rule in the Neural Agent's rules base with the predicted weight so it can be used in the next iteration.

The *Limbic Agent* is a cognitive agent since it could perform complex tasks autonomously by analyzing the interactions between the Neural Agent and the emotional reactions of the patients, learning from these interactions, modifying existing intervention rules, creating new ones, and finally predicting their impact on a patient.

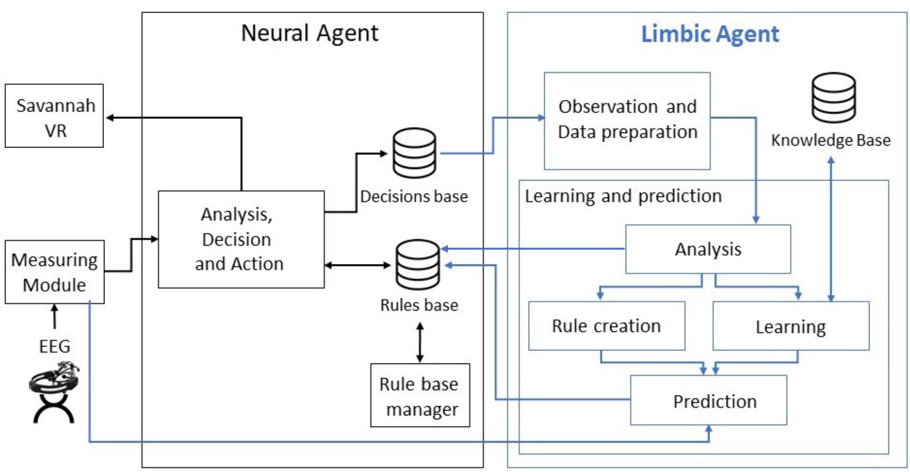


Figure 3. Architecture of the Limbic Agent and its interaction with the system.

4. Experiments

In our previous work [6] we aimed to analyze the impact of the Savannah VR on the memory and attention performances, and for that we created 6 attention and memory exercises. Exercises 1, 2 and 3 are attention tests and exercises 4, 5 and 6 are memory tests.

- Exercise 1: a series of numbers are presented vocally to the participants and they are asked to repeat them in the order of presentation using a numerical pad; then, a second series of numbers are presented vocally and the participants are now asked to report them in the backward order.
- Exercise 2: participants hear a list of letters at a rate of one letter per second and are invited to click the space bar every time they hear the letter "A".
- Exercise 3: images of different objects are shown for a short period of time. Each image is then replaced by four letters and the participant is asked to select the first letter of the object's name.
- Exercise 4: participants are asked to memorize a series of objects presented visually or verbally. Participants are then presented to a series of objects and are asked to determine whether the object was seen visually, auditorily or never presented.
- Exercise 5: several circles are presented to the participant. A series of these circles is highlighted one by one in order to create a sequence. The participants are asked to memorize and reproduce the same sequence.
- Exercise 6: participants are asked to memorize a set of three pictures for a short period of time. Then, we present four sets and the participants are asked to choose the set which corresponds to the one they saw.

4.1. Process of the experiments

We conducted experiments on 19 participants (12 females) with SCD and a mean age = 71 (SD=8.39).

The participants took part in two sessions: in the first one, we make sure that they met eligibility criteria to perform the experiments during a pre-experiment session (one hour). Our eligibility criteria were the following:

- Older than aged 60 of age,
- Francophone,
- Normal or correct-to-normal vision,
- Normal hearing,
- Met the Consortium for the Early Identification of Alzheimer's Disease – Quebec (CIMA-Q) criteria for SCD:
  - o Presence of a complaint defined as a positive answer to the following statements: "My memory is not as good as it used to be" "and it worries me."
  - o MoCA 20-30
  - o No impairment on the logical memory scale based on the education-adjusted CIMA-Q cut-off scores.

During this first session, participants were provided with oral and written description of the study and were invited to sign a consent form. The session the included the clinical tests that are necessary to confirm diagnosis and characterize participants. If the participants were eligible, they were invited to the second session.

In the experimental session, the participants start by filling a pre-session form. Then, we equip them with the Emotiv Epoc+ EEG headset, and they begin to perform attention and memory exercises. When the exercises were completed, we equip them with the Fove VR headset, and they started the savannah relaxing neurofeedback system.

Following the savannah, participants completed other examples of the attention and memory exercises and finished by filling a post-experiment form. Figure 4 details the process of the experiment.

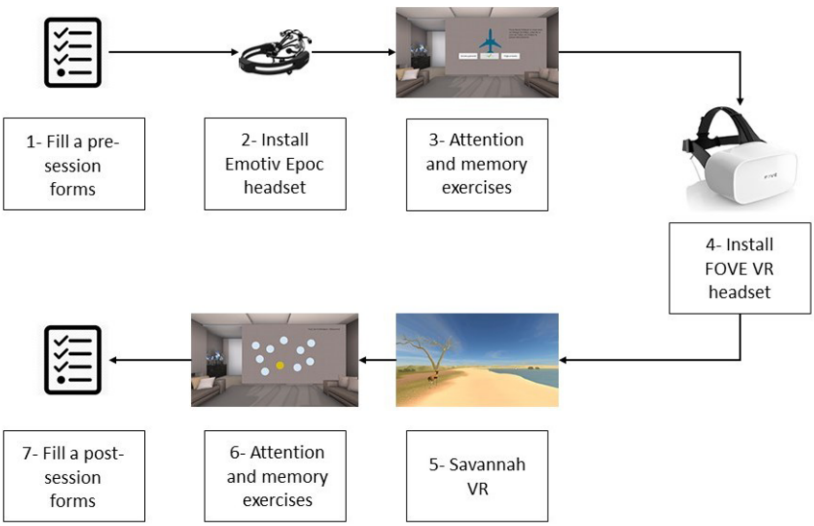


Figure 4. Process of the experiment.

As mentioned in section 3.2, the Neural Agent saves the decisions taken during its interactions with Savannah VR environment in addition to the participants' emotional reactions. We will use the data that the Neural Agent stored during these experiments in order to train the Limbic Agent.

5. Results and Discussion

We started by analyzing the interaction between the Neural Agent and the emotional reactions, we noticed that the agent's interventions worked expected 244 on a total of 414 (59%), however that the interventions did not work 170 times (41%). Figure 5 illustrates a histogram of the number of worked vs not worked intervention rules.

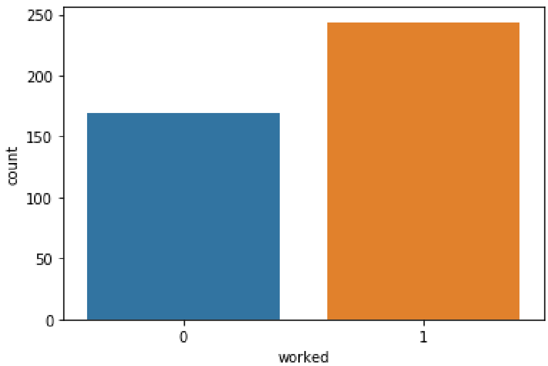


Figure 5. Histogram of the number of worked vs not worked intervention rules.

As we mentioned in section 3.3, the 'Analysis' component of the Limbic Agent analyses the interactions between the Neural Agent and the patients' emotional reaction to each rule and modifies the weight of the existent intervention rules that does not work. For instance, the Limbic Agent found that for participant 4, the intervention rule number 1 was not efficient 75%, so it changes it weight to minimum in order to inhibit it for this particular user. The same rule worked 100% for participant 18, so it increases its weight to maximum.

One of the Limbic Agent's objective is to generate new intervention rules considering the interactions between the Neural Agent and the emotional reactions that each rule generates after the intervention. The 'Rule creation' component of the Limbic Agent generated this new rule after analyzing the Neural Agents' Decisions base:

rule_id=14 | triggering _emotion=negative_engagement |
target_emotion=positive_excitement | action= light-intensity-incr

This indicates that most of the time, when the patients' engagement was negative and we increase the light intensity, their excitement went up which results to the creation of this intervention rule.

In order to predict the impact of the new intervention rules, we split the Decisions base into 70% for training and 30% for testing. We used the 'triggering_Emotion', 'targetting_Emotion', 'action', 'engagement_trend_before', 'frustration_tren_before', 'excitement_trend_before', 'valence_trend_before', 'meditation_trend_before' columns as input features and the model should predict the 'worked' column (labeled 0 or 1). We note that the 'Observation and Data preparation' component already normalized the data as follows: it gives a unique integer for each 'trigerring_emotion', 'target_emotion' and action. And for each emotion trend before the intervention rule, it put **1** if the emotion goes up, **-1** if it goes down and **0** if the trend did not change.

We trained the model that predicts if the intervention rules works or not, and we compared the predicted results of testing data with the already-known labels to analyze the efficiency of the models. Four (4) supervised learning algorithms have been tested in our study, namely: Decision Tree, Random Forest, K-Nearest Neighbors (KNN) and Support Vector Machine (SVM). For Decision tree, we used the Classification and Regression Trees (CART) version [26]. The random forest was set up with 200 estimators. For KNN, the number of k-nearest neighbors was set to 4. Finally. The SVM was set with C equal to 1.0 and gamma as default.

In order to evaluate the model of each algorithm, we use the precision, recall and F1-score indicators. **Precision** is the number of true positives divided by the number of true positives plus the number of false positives; it reflects the ability of a classification model to identify only the relevant data points. **Recall** is the number of true positives divided by the number of true positives plus the number of false negatives; it reflects the ability of a model to find all relevant cases within a dataset. **F1-Score** is a harmonic mean of precision and recall, it reflects how good a model is making a good balance between precision and recall. Table 1 details the results of each algorithm.

Table 1. Classification reports of tested algorithms

		Precision	Recall	F1-score	Support
	0	0.51	0.32	0.39	60
Decision Tree	1	0.53	0.72	0.61	65
	Avg/total	0.52	0.53	0.51	125
	0	0.55	0.28	0.37	60
Random Forest	1	0.54	0.78	0.64	65
	Avg/total	0.54	0.54	0.51	125
	0	**0.58**	**0.62**	**0.60**	60
KNN	1	**0.62**	**0.58**	**0.60**	65
	Avg/total	**0.60**	**0.60**	**0.60**	125
	0	0.62	0.22	0.32	60
SVM	1	0.55	0.88	0.67	65
	Avg/total	0.58	0.56	0.50	125

Average precision ranged from 0.52 to 0.6, average recall ranged from 0.53 to 0.6, and F1-score ranged from 0.5 to 0.6. However, we note that for all algorithms except KNN, the models miss-predicted too much class 0 which justifies the low recall and F1-score. KNN model is the only one which predicted both classes (0 and 1) with an accuracy of 60% and has an acceptable Recall and F1-score for both classes. Table 2 details the resulting confusion matrix for each tested algorithm.

Table 2. Confusion matrix of tested algorithms

		Predicted 0	Predicted 1
Decision Tree	Actual 0	19	41
	Actual 1	18	47
Random Forest	Actual 0	17	43
	Actual 1	14	51
KNN	Actual 0	37	23
	Actual 1	27	38
SVM	Actual 0	13	47
	Actual 1	8	57

These results can be improved by performing more experiments to increase the Neural Agents' Decisions base and thus increase the size of training data. The more data the 'Learning' component can have, the more knowledge it can get and the more precision the 'Prediction' component can produce.

6. Conclusion and Future Work

We presented in this work an approach aiming to reduce the negative emotions of patients with subjective cognitive decline. We used a Neural Agent which intervenes on a relaxing environment aiming to optimize it and better reduce negative emotions. Experiments were performed and results showed that we can optimize the reduction of negative emotions by controlling the relaxing environment, we proposed an approach consisting of an Intelligent Cognitive Control System (ICCS) able to interact with patient through its brain limbic system. This ICCS is composed of the Neural Agent in addition to a Limbic Agent.

The Limbic Agent was created to make the interactions between the system and the human's Limbic System works better. It analyzes the Neural Agent's intervention and the patients' emotional reaction to this intervention, learn from these interactions and makes corrections to the rules base. In addition, it generates new intervention rules that could work better and predicts their impact. Results showed that the Limbic Agent is capable to analyze the interaction and modify the existing rules by increasing and decreasing their weight. Furthermore, results showed that the agent can generate new rules and predict their efficiency using KNN model with a precision of 60%.

Future work consists of making more experiments to collect more data so that we can improve the trained models and get better results. With more data we could use deep learning techniques which could produce better results. Also, we aim to perform multiple sessions experiments with the same persons so that we can see the evolution of the personalization for each person.

Acknowledgments

We acknowledge NSERC-CRD (National Science and Engineering Research Council Cooperative Research Development), Prompt, and BMU (Beam Me Up) for funding this work.

References

[1] M. Benoit, I. Dygai, O. Migneco, P. Robert, C. Bertogliati, J. Darcourt, J. Benoliel, V. Aubin-Brunet, D. Pringuey, Behavioral and psychological symptoms in Alzheimer's disease, Dementia and Geriatric Cognitive Disorders. 10 (1999) 511–517.

[2] P.A. Boyle, L. Yu, S.E. Leurgans, R.S. Wilson, R. Brookmeyer, J.A. Schneider, D.A. Bennett, Attributable risk of Alzheimer's dementia attributed to age-related neuropathologies, Annals of Neurology. 85 (2019) 114–124.

[3] H.B. Abdessalem, A. Byrns, M. Cuesta, V. Manera, P. Robert, M.-A. Bruneau, S. Belleville, C. Frasson, Application of Virtual Travel for Alzheimer's Disease, 9th International Conference on Sensor Networks. (2020) 52–60.

[4] A. Byrns, H.B. Abdessalem, M. Cuesta, M.-A. Bruneau, S. Belleville, C. Frasson, Adaptive Music Therapy for Alzheimer's Disease Using Virtual Reality, in: Springer, 2020: pp. 214–219.

[5] D. Horváthová, V. Siládi, Creating virtual environments for phobia treatment, Open Computer Science. 6 (2016) 138–147. https://doi.org/10.1515/comp-2016-0012.

[6] Ben Abdessalem, H., Byrns, A. et Frasson, C. Optimizing Alzheimer's Disease Therapy Using a Neural Intelligent Agent-Based Platform. International Journal of Intelligence Science. (2021). 11(02), 70-96. https://doi.org/10.4236/ijis.2021.112006

[7] Y. Liu, Y. Ding, C. Li, J. Cheng, R. Song, F. Wan, X. Chen, Multi-channel EEG-based emotion recognition via a multi-level features guided capsule network, Computers in Biology and Medicine. 123 (2020) 103927. https://doi.org/10.1016/j.compbiomed.2020.103927.

[8] L. Zhu, C. Su, J. Zhang, G. Cui, A. Cichocki, C. Zhou, J. Li, EEG-based approach for recognizing human social emotion perception, Advanced Engineering Informatics. 46 (2020) 101191. https://doi.org/10.1016/j.aei.2020.101191.

[9] W. Li, X. Hu, X. Long, L. Tang, J. Chen, F. Wang, D. Zhang, EEG responses to emotional videos can quantitatively predict big-five personality traits, Neurocomputing. 415 (2020) 368–381. https://doi.org/10.1016/j.neucom.2020.07.123.

[10] B. Oltu, M.F. Akşahin, S. Kibaroğlu, A novel electroencephalography based approach for Alzheimer's disease and mild cognitive impairment detection, Biomedical Signal Processing and Control. 63 (2021) 102223. https://doi.org/10.1016/j.bspc.2020.102223.
[11] Ben Abdessalem, H. et Frasson, C. Real-time Brain Assessment for Adaptive Virtual Reality Game : A Neurofeedback Approach. Dans C. Frasson et G. Kostopoulos (dir.), Brain Function Assessment in Learning (2017) 133-143. https://doi.org/10.1007/978-3-319-67615-9_12.
[12] S. Rugaber, A.K. Goel, L. Martie, GAIA: A CAD Environment for Model-Based Adaptation of Game-Playing Software Agents, Procedia Computer Science. 16 (2013) 29–38. https://doi.org/10.1016/j.procs.2013.01.004.
[13] F. Doctor, R. Iqbal, R.N.G. Naguib, A fuzzy ambient intelligent agents approach for monitoring disease progression of dementia patients, J Ambient Intell Human Comput. 5 (2014) 147–158. https://doi.org/10.1007/s12652-012-0135-x.
[14] C. Peña, J.-L. Marzo, J.D.L. Rosa, Intelligent Agents in a Teaching and Learning Environment on the Web, in: 2002.
[15] K.R. Walsh, S.D. Pawlowski, Virtual Reality: A Technology in Need of IS Research, CAIS. 8 (2002) 296–314. https://doi.org/10.17705/1CAIS.00820.
[16] F. Biocca, M.R. Levy, Communication in the age of virtual reality, 1995.
[17] G. Lorenzo, A. Lledó, J. Pomares, R. Roig, Design and application of an immersive virtual reality system to enhance emotional skills for children with autism spectrum disorders, Computers & Education. 98 (2016) 192–205. https://doi.org/10.1016/j.compedu.2016.03.018.
[18] H. Coyle, V. Traynor, N. Solowij, Computerized and Virtual Reality Cognitive Training for Individuals at High Risk of Cognitive Decline: Systematic Review of the Literature, The American Journal of Geriatric Psychiatry. 23 (2015) 335–359. https://doi.org/10.1016/j.jagp.2014.04.009.
[19] Dakoure, C., Ben Abdessalem, H., Boukadida, M., Cuesta, M., Bruneau, M.-A., Belleville, S. et Frasson, C. Virtual Savannah: An Effective Therapeutic and Relaxing Treatment for People with Subjective Cognitive Decline. Dans P. Vlamos, P. Bamidis et C. Frasson (dir.), Brain Function Assessment in Learning (2020) 107-112.
[20] A.M. Durko, J.F. Petrick, Travel as Relationship Therapy: Examining the Effect of Vacation Satisfaction Applied to the Investment Model, Journal of Travel Research. 55 (2016) 904–918. https://doi.org/10.1177/0047287515592970.
[21] A. Biamonti, S. Gramegna, B. Imamogullari, A design experience for the enhancement of the quality of life in people with alzheimer's, in: PRT, 2014: pp. 273–284.
[22] K. Schaler, Travel Therapy: Where Do You Need to Go?, Seal Press, 2009.
[23] G.E. Lancioni, N.N. Singh, M.F. O'Reilly, J. Sigafoos, C. Renna, M. Ventrella, K. Pinto, M.G. Minervini, D. Oliva, J. Groeneweg, Supporting daily activities and indoor travel of persons with moderate Alzheimer's disease through standard technology resources, Research in Developmental Disabilities. 34 (2013) 2351–2359.
[24] L.M. Heinrich, E. Gullone, The clinical significance of loneliness: A literature review, Clinical Psychology Review. 26 (2006) 695–718.
[25] J.W. Papez, A PROPOSED MECHANISM OF EMOTION, Arch NeurPsych. 38 (1937) 725. https://doi.org/10.1001/archneurpsyc.1937.02260220069003.
[26] W. Loh, Classification and regression trees, WIREs Data Mining Knowl Discov. 1 (2011) 14–23. https://doi.org/10.1002/widm.8.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210078

On the Development of a Personalized Augmented Reality Spatial Ability Training Mobile Application

Christos PAPAKOSTAS[1], Christos TROUSSAS, Akrivi KROUSKA, Cleo SGOUROPOULOU
Department of Informatics and Computer Engineering
University of West Attica, Greece

Abstract. Augmented Reality has been integrated in educational settings in the field of engineering. Prior research has examined the learning outcomes and the pedagogical affordances of this technology. However, training undergraduate engineers, from diverse knowledge level, requires customized training approach, tailored to the individual learning pace. In this paper, we present PARSAT (Personalized Augmented Reality Spatial Ability Training), which is a mobile Augmented Reality application for the enhancement of students' spatial visualization skills. The application takes into account the theoretical contents of engineering design, deployed through video tutorials, and student-computer interaction with 3D objects. Students interpret different views of a 3D object, which are represented on their mobile screen. PARSAT efficaciously strengthens students' recognition of spatial structures and views, adjusted to the fulfillment of their personal needs. In terms of personalization, PARSAT consists of different levels, which do not follow a linear flow, as each student takes part in a different sequence of activities, according to their time spent in the 3D object manipulation, and their assessment scores at the end of each level. Furthermore, an agent is used to analyze students' knowledge level, and send them feedback. The system reduces unnecessary cognitive load and, at the same time, improves students learning experience in learning engineering drawing.

Keywords. Spatial ability, personalized augmented reality, mobile application, customized training, machine learning

1. Introduction

Spatial ability, as a factor of human intelligence, was initially recognized and studied by Thorndike [1]. Thorndike's proposed model consisted of three components, namely abstract, mechanical and social intelligence, and served as the early stage research for later studies on spatial ability. Thurstone [2] suggested that the intelligence consisted of seven primary mental abilities, one of which is spatial visualization, as involved in visualizing and manipulating objects. Thurstone [3] defined three core components of

[1] Corresponding Author; E-mail: cpapakostas@uniwa.gr.

spatial ability, namely mental rotation, spatial visualization, and spatial perception. Mental rotation is defined as the ability to recognize an object being moved in different directions or angles; spatial visualization is defined as the ability to recognize the parts of an object when it is moved or displaced from its original position; and spatial perception is defined as the ability to use the own body orientation to interact with the environment and, therefore, with spatial orientation.

Gardner [4] introduced the concept of multiple intelligences, rather than defining intelligence as a single general ability. Gardner's argument stated that there is a wide range of cognitive abilities, not necessarily correlated between them, namely music-rhythmic and harmonic, visual-spatial, linguistic-verbal, logical-mathematical, bodily-kinesthetic, interpersonal, intrapersonal, and naturalistic. Many modern theories proposed additional factors or cognitive processes, in their effort to better understand spatial ability [5,6]. Each research has added significantly to the definition of spatial ability, as a form of intelligence, where a person demonstrates the capacity to mentally generate, transform, and rotate a visual image and thus, understand and recall spatial relationships between real and imagined objects.

Smith [7] reported a list of 84 different occupations requiring top-level spatial ability, 26 of which were related to engineering, while 14 of them were related to graphics. There is significant correlation between spatial ability and many scientific fields, such as geometry, physics and technical drawing. Field [8] designed a 52-hours course aiming to develop skills in representing spatially visualized objects through projections. Martín - Dorta et al. [9] also launched a fast remedial course based on 3-dimensional modelling for improving spatial abilities of engineering students. Designing 3-dimensional objects is highly considered to be a crucial factor in the development of spatial skills [10,11].

However, many students find difficulties in visualizing abstract concepts or the geometry of 3-dimensional objects [12]. The training of spatial skills is either based on the manipulation of physical models used for exercises, or modelling in computer software [13]. Zaretsky & Bar [14] positively introduced virtual reality (VR) for object rotation training, and respectively with the increased duration that they observed in students' concentration, they finally enhanced their academic achievements. Overall, the presentation of an object in 3-dimensional view using a smartphone or a tablet, can help students having difficulties to understand its geometry, to watch it from different angles. Hidden details of the object are revealed due to the software's availability for various views orientations, and the technical drawing can be implemented faster.

Apart from VR, Augmented Reality (AR) is another alternative technology for spatial skills training, considered by many recent studies as the best alternative teaching approach [15–18]. AR superimposes a computer-generated image on a user's view of the real world, thus providing a composite view. Azuma et al. [19] defined AR as a system which a) combines real and virtual objects in a real environment; b) runs interactively in real time; and c) aligns real and virtual objects with each other. The evaluation of AR in educational settings is gaining more place, mainly due to the emergence of smartphones and tablets. AR has a great potential in enhancing students' spatial ability and learning experience.

A few researchers have focused on the integration of mobile AR in spatial ability training. Tumkor [20] tested mixed reality (MR) technologies for more than 3 years in engineering design courses, representing two-dimensional sketches in front of students, while also rotating the virtual objects as they wished. The results reported an overall improvement of students' visualization skills. Figueiredo et al. [21] presented EducHolo, a mobile AR learning tool for the visualization and interaction of 3D models, providing students with a better perception of models and improving theirs sketching at 2D orthographic views. De Ravé et al. [22] designed DiedricAR, a mobile AR exercise workbook, aimed at the learning of descriptive geometry and explored the application's benefits for students' spatial ability. Kaur et al. [23] designed and developed GeoSolvAR, an AR-based solution for visualizing 3D solids. Omar et al. [24] emphasized on the development of visualization skills and concept understanding, by teaching and learning orthographic projections using an AR engineering drawing application, namely AREDApps. Papakostas et al. [25] explored the advantages and trends of AR in spatial ability training based on the revision of articles over the decade 2010-2019. The trend, in the development of AR applications, is to use three-dimensional objects towards improving spatial ability. The considerable findings identified the lack of personalized applications adapting their content to the learner's level of visualization skills. That is an area where there is still vast potential development.

To address this problem, we designed a mobile personalized AR spatial ability training application, named PARSAT, which is tailored to meet the user's spatial skills level. The diversity of the characteristics that each student presents, such as the level of knowledge, the gender, the learning style, etc, define the process of learning. The novelty of our research lies on the introduction of the concept of adaptive AR, as the next-generation advanced educational application, designed to improve students' visualization skills. Adaptive AR is among the most promising technologies in teaching and learning processes, enabling the students to benefit from the relations with the 3D objects in the space, and thus, keeping pace with their own abilities and interest.

2. Overview of the system

The main difference, between PARSAT and the existing AR applications, is that the educational process is adaptive [26–28]. The basic idea of the application is to deliver relevant content based, at first, on student's background knowledge and secondly, on student's assessment score [29–31].

The application consists of three different stages, each one covering terms of the Orthographic Projection chapter of engineering design graphics course. The most fundamental technique is called graphical projection, by which a three dimensional object is represented on a planar surface. Orthographic projection uses a series of two-dimensional views, arranged in a standard manner, in order to fully document the object's geometry [32].

Students can receive a short, medium or long tutorial at the beginning of the training, depending on their skill level and past knowledge. PARSAT then assigns each student the alternative learning path that maximizes his/her target key performance indicator (KPI).

In order to achieve the research goal, a framework has been developed reflecting the typical spatial abilities training, comprising of three modules which are: i) the User Interface Model (UIM); ii) the Domain Knowledge Model (DKM); and iii) the Personalization Model (PM), analyzed at the next subsections. Figure 1 illustrates the logical architecture of PARSAT.

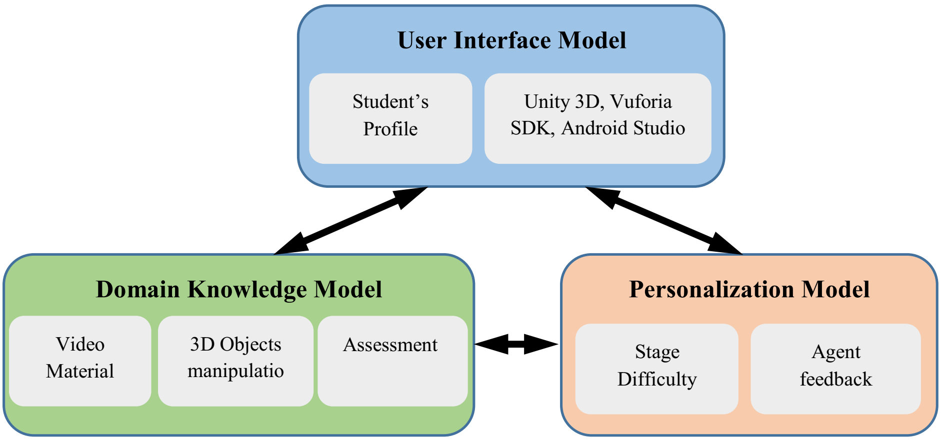


Figure 1. Logical architecture of PARSAT

2.1. User interface model

The user interface model includes the context and the function of the AR application, making the student experience as clean and simple as possible. As such, students create their profile, based on the demographic characteristics of age and gender, and cognitive characteristics of prior knowledge of the topic, before training starts.

One of the key difficulties in developing AR applications is the tracking of the user's viewpoint. The application has to know all of the time where the user is looking in the real world, using the mobile's rear-view camera. To this direction, we used Vuforia[2] Software Development Kit (SDK), one of the best-known AR tool sets, adding advanced computer vision functionality to both Android and iOS applications, and creating AR experiences that realistically interact with the 3D geometrical objects displayed at each stage. We took advantage of Vuforia's marker-based recognition and imported 3D objects with virtual buttons embedded, so that the students can move, rotate and zoom them.

Regarding the development framework, we used Unity3D[3] cross-platform game engine. In 2016, Vuforia announced a partnership with Unity to integrate the Vuforia platform into Unity's popular game engine and development platform. One of Unity's most powerful features is the extensible editor it has. We used editor scripting in C-Sharp

[2] https://developer.vuforia.com/

[3] https://unity.com/

(C#) programming language to extend the functionalities of the application. Finally, Android Studio Integrated Development Environment (IDE) is interfaced with Unity 3D to build the application for Android environment.

2.2. Domain knowledge model

The domain knowledge model includes the basic, intermediate and advanced topics from the domain of engineering graphics in an undergraduate level of the School of Engineering. Each one of the three levels of complexity consists of 5 sections (Table 1).

Table 1. Domain knowledge of PARSAT

Difficulty Level	Sections
Basic	1. Orthographic projection introduction
	2. The six principal views
	3. The glass box method
	4. Standard views – view alignment
	5. Line type – line weight
Intermediate	6. Creating an orthographic projection
	7. Drawing an orthographic projection
	8. Basic dimensioning rules
	9. Parts of dimensions
	10. Scales
Advanced	11. Cutting plane
	12. Cutting plane line
	13. Section lining
	14. Full sections
	15. Half sections

The aforementioned sections, covered by PARSAT, are deployed firstly through recorded video tutorials, and secondly through interaction with augmented 3D objects, which were made in 3D software. Among a variety of 3D designing tools, we selected Autodesk 3D Studio Max[4], as it is a product more suitable for modeling, architecture designs, engineering, and construction. All 3D models, created in 3D Studio Max, were exported to Unity using an appropriate generic file format.

The video tutorials have different details and length providing the alternatives of short, medium, and long tutorial, according to the student's background knowledge level.

2.3. Personalization model

The adaptation, or else personalization, model is modeling the users and customizes the components of the learning content and the flow of the learning. Each student follows a different sequence of activities that allows them to acquire knowledge and reach the topic's objectives.

[4] https://www.autodesk.com/products/3ds-max

Each student's default starting stage is the second one, and a Difficulty Personalization module (DPM) defines whether the difficulty should be easier or harder for the student. Therefore, each student is either less frustrated, in case of a hard level, or less bored, in case the current level is too easy. DPM takes a set of inputs, such as firstly the time spent by the student on manipulating augmented 3D objects, and secondly the scores in each level's assessment, and derives the order of the stages delivered to the students.

As far as the evaluation is concerned, we used an agent to analyze the assessment score and send feedback to the user. Agent could help the student identify potential gaps in the knowledge domain, and proposes specific sections that should be repeated and/or book pages that should be thoroughly studied. Moreover, software agents could also enhance students' support level, by encouraging them to keep up the good effort.

3. Evaluation

The first task in evaluating the prototype, is to ask the target users for initial feedback on design, usability, and user experience. Towards this direction, 35 postgraduate students were recruited to test the application on their own. The participants were all students of a Master Degree program on Computer Science and Engineering of the Department of Informatics and Computer Engineering. The participants were asked to operate the PARSAT by themselves, to watch the video tutorials, to rotate the 3D objects in order to visualize and understand their structures, and finally took the spatial skills test to evaluate the training effect.

After interacting with the AR application, the participants answered a questionnaire, in order to determine their personal feelings about the training. All the items of the questionnaire (Table 2) were measured using a 7-point Likert scale ranging from (1) strongly disagree to (7) strongly agree.

Table 2. Questions asked about PARSAT

Constructs	Question
Design	1. PARSAT has a simple and understandable menu
	2. PARSAT has good functionality
	3. PARSAT presents high quality videos and 3D objects
Usability	4. Using PARSAT enhances my spatial skills
	5. Using PARSAT increases my educational performance
	6. I find PARSAT useful for my course
User experience	7. I would like to use PARSAT in more knowledge sections
	8. I will recommend others to use PARSAT
	9. AR has become one of my favorite technologies

Descriptive analysis of mean and standard deviation was used to analyze the feedback towards using PARSAT. Other than that, short interview sessions were used, discussing the participants' experience with the system. Usability and system performance were measured.

In the following paragraph we briefly report our findings regarding the framework application. Based on the results gathered, all participants agreed that PARSAT had a positive training effect (mean level of agreement = 6.1; SD = 0.62; score from 1 to 7). The participants strongly agreed that the AR integration allowed them to manipulate the 3D objects, improving their understanding of their geometry, thus the adaptive behavior of PARSAT significantly reduced the cognitive load.

In the rest of this section, we present some qualitative results, collected through opinion surveys and interviews with the participants (Figure 2). Three group interviews were carried out, and they described the application as challenging, the 3D objects very attractive and appropriate, and they craved for more teaching material. In the individual interviews, participants showed satisfaction with the use of these 3D tools during the course.

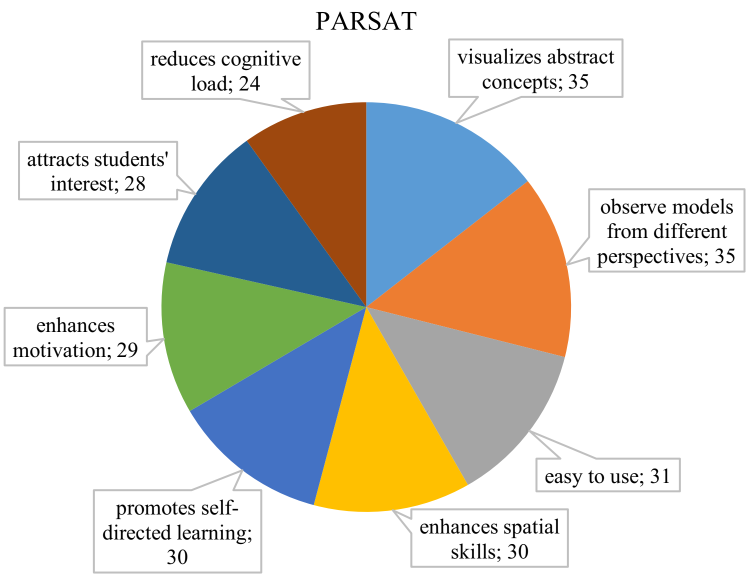


Figure 2. Learners' outcome of PARSAT

4. Conclusion

AR can currently help towards the digitalization of education, and the enhancement of learning achievements. To this direction, this paper focused on the development of PARSAT, which is a mobile AR application incorporating personalization for the enhancement of learners' visualization skills.

The findings of this research show that there is a significant difference in learners' spatial skills after using PARSAT application. PARSAT's optimizations are personalized and dynamic for each student, based on their state and progression. The adaptive process is critical for students' retention, and depends on the tuning of key parameters, such as the difficulty of the stages and the tutorial duration. Once there is enough collected data, PARSAT can select alternative for the student by a machine learning model.

5. Future work

Our 3D object controller has only a rotate mode, while next version of PARSAT will embed zoom and move mode. Furthermore, the three stages of difficulty will be extended to five, so that each level will have less learning goals and students will not feel stressed about the cognitive load. Lastly, personalization, based on machine learning algorithm implemented in C# script, is a long process that requires input data from a lot of participants, in order to train the model. So, we will incorporate the improvements in PARSAT and distribute it to engineering students' classrooms for training.

References

[1] E.L. Thorndike, On the Organization of Intellect., *Psychol. Rev.* **28** (1921), 141–151. https://doi.org/10.1037/h0070821.

[2] L.L. Thurstone, *Primary mental abilities*, Psychometric Monograph 1, Chicago, IL: Univer, 1938.

[3] L. Thurstone, Some Primary Abilities in Visual Thinking, *Proc. Am. Philos. Soc.* **94** (1950), 517–521.

[4] H. Gardner, *Frames of mind : the theory of multiple intelligences*, Basic Books, New York, 1983.

[5] J.W. Pellegrino, E.B. Hunt, Cognitive models for understanding and assessing spatial abilities, *Intell. Reconceptualization Meas.* (1991), 203–225.

[6] J.B. Carroll, *Human cognitive abilities: A survey of factor-analytic studies*, Cambridge University Press, 1993.

[7] I.M. Smith, *Spatial ability : its educational and social significance*, R.R. Knapp, San Diego, Calif., 1964.

[8] B.W. Field, A course in spatial visualization. *Proceedings of the 6th International Conference on Engineering Design Graphics and Descriptive Geometry* (1994), 257–261.

[9] N. Martín-Dorta, J.L. Saorín, M. Contero, Development of a fast remedial course to improve the spatial abilities of engineering students, *J. Eng. Educ.* **97** (2008), 505–513.

[10] A. Šafhalter, S. Glodež, A. Šorgo, M. Ploj Virtič, Development of spatial thinking abilities in engineering 3D modeling course aimed at lower secondary students, *Int. J. Technol. Des. Educ.* (2020). https://doi.org/10.1007/s10798-020-09597-8.

[11] R.H. McKim, *Experiences in visual thinking*, Brooks/Cole Pub. Co., Monterey, Calif., 1980.

[12] A.R. Arslan, S. Dazkir, Technical Drafting and Mental Visualization in Interior Architecture Education, *Int. J. Scholarsh. Teach. Learn.* **11** (2017). https://doi.org/10.20429/ijsotl.2017.110215.

[13] H.B.P. Gerson, S.A. Sorby, A. Wysocki, B.J. Baartmans, The development and assessment of multimedia software for improving 3-D spatial visualization skills, *Comput. Appl. Eng. Educ.* **9** (2001), 105–113.

[14] E. Zaretsky, V. Bar, Intelligent virtual reality and its impact on spatial skills and academic achievements, 10th Int. Conf. Inf. Syst. Anal. Synth. (2004), 107–113.

[15] O. Huerta, A. Kus, E. Unver, R. Arslan, M. Dawood, M. Kofoğlu, V. Ivanov, A design-based approach to enhancing technical drawing skills in design and engineering education using VR and AR tools, *VISIGRAPP 2019 - Proc. 14th Int. Jt. Conf. Comput. Vision, Imaging Comput. Graph. Theory Appl.* **3** (2019), 306–313. https://doi.org/10.5220/0007566003060313.

[16] R.M. Yilmaz, Augmented Reality Trends in Education between 2016 and 2017 Years, *State Art Virtual Real. Augment. Real. Knowhow*. (2018). https://doi.org/10.5772/intechopen.74943.

[17] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, User acceptance of augmented reality welding simulator in engineering training, *Educ. Inf. Technol.* (2021). https://doi.org/10.1007/s10639-020-10418-7.

[18] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, Measuring User Experience, Usability and Interactivity of a Personalized Mobile Augmented Reality Training System, *Sensors.* **21** (2021). https://doi.org/10.3390/s21113888.

[19] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, B. MacIntyre, Recent advances in augmented reality, *IEEE Comput. Graph. Appl.* **21** (2001), 34–47. https://doi.org/10.1109/38.963459.

[20] S. Tumkor, Personalization of engineering education with the mixed reality mobile applications, *Comput. Appl. Eng. Educ.* **26** (2018), 1734–1741. https://doi.org/10.1002/cae.21942.

[21] M. Figueiredo, P.J.S. Cardoso, J.M.F. Rodrigues, R. Alves, Learning Technical Drawing with Augmented Reality and Holograms, *Recent Adv. Educ. Technol. Methodol*. (2014), 1–20.

[22] E.G. de Ravé, F.J. Jiménez-Hornero, A.B. Ariza-Villaverde, J. Taguas-Ruiz, DiedricAR: a mobile augmented reality system designed for the ubiquitous descriptive geometry learning, *Multimed. Tools Appl.* **75** (2016), 9641–9663. https://doi.org/10.1007/s11042-016-3384-4.

[23] N. Kaur, R. Pathan, U. Khwaja, S. Murthy, GeoSolvAR: Augmented reality based solution for visualizing 3D Solids, *Proc. - IEEE 18th Int. Conf. Adv. Learn. Technol. ICALT 2018* (2018), 372–376. https://doi.org/10.1109/ICALT.2018.00093.

[24] M. Omar, D.F. Ali, A.N. Nasir, M.S. Sunar, AREDApps: Integrating mobile augmented reality in orthographic projection teaching and learning, *Int. J. Recent Technol. Eng.* **8** (2019), 821–825.

[25] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, Exploration of Augmented Reality in Spatial Abilities Training: A Systematic Literature Review for the Last Decade, *Informatics Educ.* **20** (2021), 107–130. https://doi.org/10.15388/infedu.2021.06.

[26] A. Krouska, C. Troussas, M. Virvou, K. Espinosa, C. Troussas, J. Caro, M. Virvou, Language Learning Assisted by Group Profiling in Social Networks*, Int. J. Emerg. Technol. Learn.* **8** (2013), 35-38. http://dx.doi.org/10.3991/ijet.v8i3.2684.

[27] C. Troussas, A. Krouska, and M. Virvou, Using a Multi Module Model for Learning Analytics to Predict Learners' Cognitive States and Provide Tailored Learning Pathways and Assessment, *Virvou M., Alepis E., Tsihrintzis G., Jain L. (eds) Machine Learning Paradigms, Intelligent Systems Reference Library* **158** (2020), Springer, Cham. https://doi.org/10.1007/978-3-030-13743-4_2.

[28] A. Krouska, C. Troussas, M. Virvou, Computerized Adaptive Assessment Using Accumulative Learning Activities Based on Revised Bloom's Taxonomy, *Virvou M., Kumeno F., Oikonomou K. (eds) Knowledge-Based Software Engineering: 2018. JCKBSE 2018. Smart Innovation, Systems and Technologies* **108** (2019), Springer, Cham. https://doi.org/10.1007/978-3-319-97679-2_26

[29] C. Troussas, A. Krouska, C. Sgouropoulou, Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education, *Comput. Educ.* **144** (2019), 103698. https://doi.org/10.1016/j.compedu.2020.103698.

[30] A. Krouska, C. Troussas, M. Virvou, Social networks as a learning environment: Developed applications and comparative analysis, *8th International Conference on Information, Intelligence, Systems & Applications (IISA)* (2017), 1-6. https://doi.org/10.1109/IISA.2017.8316430.

[31] C. Troussas, A. Krouska, C. Sgouropoulou, I. Voyiatzis, Ensemble Learning Using Fuzzy Weights to Improve Learning Style Identification for Adapted Instructional Routines, *Entropy* **22** (2020), 735. https://doi.org/10.3390/e22070735.

[32] J.D. Camba, J. Otey, M. Contero, M. Alcaniz, *Visualization and Engineering Design Graphics with Augmented Reality*, SDC Publications, Seiten, 2013.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210079

Zoo Therapy for Alzheimer's Disease with Real-Time Speech Instruction and Neurofeedback System

Yan AI[1], Hamdi BEN ABDESSALEM[2] and Claude FRASSON[3]
Département d'Informatique et de Recherche Opérationnelle,
Université de Montréal, Canada

Abstract. There is an increasing number of people with Alzheimer's disease. Negative emotions are not only one of the symptoms of AD, but also the accelerator of the disease. Animal therapy can have a positive impact on the negative emotions of patients, but it has strict requirements for both environments and animals. In this study, we aim to explore the effectiveness of using virtual animals and their impact on the reduction of patients' negative emotions to improve the user's cognitive functions. This approach has been implemented in the Zoo Therapy project, which presents an immersive 3D virtual reality animal environment, where the impact on the patient's emotion is measured in real-time by using electroencephalography (EEG). In addition to creating highly realistic virtual animals, the innovation of Zoo Therapy is also in its communication mechanism as it implements bidirectional human-computer interaction supported by 3 interaction methods: 3D buttons, speech instruction, and Neurofeedback. Patients can actively interact with virtual animals through 3D buttons or speech instructions. The Neurofeedback system will guide the animal to actively interact with the patients according to their real-time emotional changes to reduce their negative emotions. Experiments and preliminary results show that it is possible to interact with virtual animals in Zoo Therapy, and the Neurofeedback system can intervene in Zoo VR environment when the emotional value goes down and might reduce patients' negative emotions.

Keywords. Alzheimer's Disease, Virtual Reality, EEG, Intelligent Agent, Human-Computer Interaction, Speech Recognition, Zoo Therapy, Emotions, Neurofeedback system, Reinforcement Learning, Auto-encoder

1. Introduction

Alzheimer's disease (AD) is a chronic progressive neurodegenerative disease. In the progression of AD, when the gradual deterioration of cognitive function appears, emotional changes include emotional agitation, anxiety, irritability, depression, and social withdrawal also follow the entire process [1, 2]. Such emotional fluctuations, with negative emotions, can cause distress to AD patients, their relatives and friends, and even medical staff. This leads to more negative emotions, a vicious circle. These negative emotions will in turn affect and destroy the cognitive function area of the patient's brain and push the further deterioration of the disease [3,4,5]. Due to the side effects of

[1] yan.ai@umontreal.ca
[2] hamdi.ben.abdessalem@umontreal.ca
[3] frasson@iro.umontreal.ca

pharmacological treatment [6,7,8], non-pharmacological therapy has attracted attention. Among them, some stimulation-oriented treatments are designed to reduce the negative emotions of users have shown positive effects in studies [9,10,11].

Animal-assisted therapy (AAT) is also a stimulation-oriented treatment that uses animals as the treatment medium is different from therapies such as objects, music, and sports. Animals are living organisms. They have emotions, can perceive human behavior and emotions, and can interact with people. This human-animal interaction constitutes a strong emotional background. Non-organisms cannot give this kind of emotional feedback. Many studies [13,14] have shown the AAT can have a positive impact on the social behavior, the negative emotions, mental state, social interaction ability and quality of life of the AD patients. However, the various limitations such as maintenance AAT team [14, 15], cost, and some ethical concerns [16,17] of AAT prevent this therapy from being widely promoted. Using virtual animals instead of real animals to participate in the treatment is a solution.

Although some studies [11,18] have tried to use virtual animals to complete AAT, there are still many shortcomings, especially how to perform the most important interactive part of AAT. Existing virtual animal treatments usually use cartoon animals, and their interaction is done by clicking on the interactive patterns on the user interface. It is not a direct interaction between the user and the animal. It is more like that a user interacts with a computer, rather than an animal. And these virtual animals cannot give emotional feedback because it cannot perceive users. The connection between participants and virtual animals is in one direction. All the emotional changes of the participants do not affect the actions and behaviors of the virtual animals.

As an overall improvement, also our contributions, we propose a brand-new virtual animal interaction therapy environment called Zoo Therapy. 1) It uses 360 degrees VR technology to allow users to interact with virtual animals immersivity. 2) Under the combined action of the three interactive methods 3D buttons, discourse (speech recognition system), and the user's emotions (Neurofeedback system), Zoo Therapy realizes the bidirectional interaction, to maximize the emotional feedback between the user and the virtual animal to ensure timely reduce the negative emotions of users. 3) we are the first to use the reinforcement learning-based method to autonomously learn the reaction rules of humans' EEG emotions in specific environments.

Our research questions are: (1) Is it possible to interact by the user's voice in a virtual reality environment? (2) The Neurofeedback system can send corresponding instructions to the virtual reality environment in real-time to intervene on the user's emotions. (3) Is it possible to reduce the user's negative emotions by interacting with virtual animals?

The rest of this paper is organized as follows. In section 2, we present our animal therapy and detail the different modules of the therapy that we developed. In section 3 we detail the experimental procedure. In section 4 we discuss the obtained results. Finally, in section 5, conclusion.

2. Our Approach: Zoo VR Therapy System

Our approach consists of a Zoo VR Therapy System which has realistic virtual horses and virtual dogs to participate in the interaction and supports three different interaction methods. The three interaction methods can be divided into two parts: subjective interaction and non-subjective interaction. Subjective interaction refers to the interactive instructions actively issued by the user and includes 3D Buttons and a Speech Instruction system. Non-subjective interaction is the interaction that the Neurofeedback system that guides the animal to actively interact with the user based on the user's real-time emotions. Neurofeedback system's EEG measuring module measures users' emotional reactions to the environment and the intelligent agent receives users' emotions and Zoo environment's information in real-time and intervenes in the Zoo environment by commanding animals depending on the emotions of the users and information of the environment. As shown in Figure 1.

2.1. Zoo VR Environment

"VR Zoo" creates a safe and economical virtual interactive environment including virtual animals. In this environment, the user can call the animal to come, eat, walk, stand up or ask it to leave at any time. Animals respond immediately when they receive instructions. Interactive instructions can come from three interactive methods. Our environment can be divided into five Functional Modules, namely, scene module, animal module, sound effect module, map module and human-computer interaction module. Following is a description for each module.

Scene Module - The overall appearance of the Zoo environment. In the scene module, we created a 3D treatment room (shown as Figure 2).

Animal models - Animals in the environment. The most common forms of AAT are dogs and horses. Therefore, we created a 3D treatment horse and a 3D treatment dog. (Figure 3 is the 3D horse in the treatment room). To complete the interaction between the virtual animals and the users, we also created an animation system when creating the animals to realize the animals' walking, eating, standing and other actions.

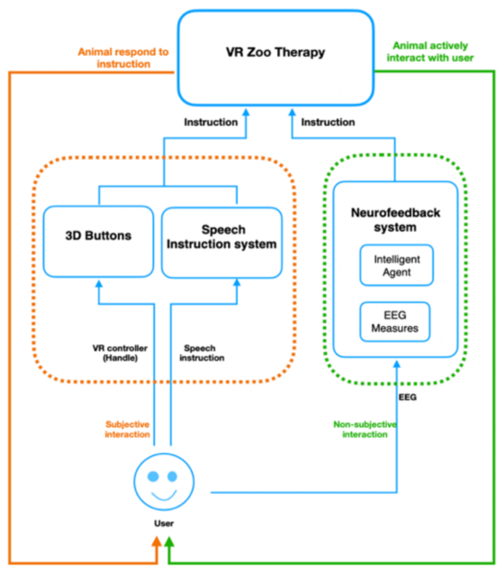


Figure 1. Architecture of Zoo Therapy.

Figure 2. 3D treatment room

Figure 3. The 3D horse in the treatment room

Sound effect module - Environment and animal sound effects. We do not only play soothing background music in the entire 3D environment. We also add different animal sound effects. For example, horse and dog can make several different calls in response to different commands from users, and the sounds of horse walking and eating, the dog's panting. The sound effect will accompany the animal's movement and change in real-time according to the different actions of the animals.

Map Module - The trajectory of animal movement and the generation of interactive routes. The function of the map module is to calculate and generate a feasible path according to the animal's real-time position and state, so that the animal will update its new path depending on the user's new instructions and approach or go away from the user.

Human-computer interaction module - User interaction with the environment. First, the user can interact with the animals by selecting the 3D button in the environment. Then, after completing the model training of the speech recognition system, we will directly add the speech recognition system. Users can talk to the virtual animals to make them come, walk, eat, or leave. In addition, a Neurofeedback system has been also added to the environment to influence animal behavior by identifying emotional changes in the user's EEG. When the three instruction modules work together, we give priority to the 3D button and speech recognition system. We will introduce the speech instruction system in section 2.2 and the Neurofeedback system in 2.3.

Finally, we integrate the above five modules to form the final VR Zoo Therapy.

2.2. Speech instruction system

We added the speech recognition system that allows users to interact with animals by using their voices. We directly selected an End-to-End ASR (Transformer [21]) model [22] which accepts a spectrogram parsed from the user's voices as the input and the output is text. This system's goal is to input the user's voice, recognize the text content of the voice through our trained speech recognition model, and determine whether the user has issued an instruction to the animals based on the content of the text. If there are instructions, the model will send them to the environment, and the animals in the environment will complete the task according to the instructions.

The horse in Zoo therapy can receive five commands: come, eat, walk, stand up, and leave. Therefore, the speech recognition model is particularly sensitive to these words. During the use of the Zoo environment, which is synchronized to 16 kHz sampling rate recorded audio. A waveform audio file is generated every three seconds, and then passed to the trained ASR model in turn for further processing. Once the ASR model receives audio file, it will parse it to a spectrogram, and then predict the audio content. The model converts it into text information. Finally, the predicted audio content will be judged whether it contains the above five instructions. If so, the model will send instructions to the Zoo environment and execute them.

2.3. Neurofeedback System

The Neurofeedback System monitors the user's emotional state in real-time and allows the animals to respond to the user's emotions. Our Neurofeedback system includes EEG measures and an RL-based agent, between which EEG measures are responsible for monitoring the user's EEG and passing it to the intelligent agent. The agent will automatically determine whether to allow the animal to actively interact with the user without the user giving an instruction to the animal based on the user's EEG information, thereby reducing the user's negative emotions.

2.3.1. EEG Measures

In this research we use the Emotiv Epoc EEG headset to track emotions. The headset contains 14 electrodes spatially organized according to the International 10-20 system, moist with a saline solution. The detailed position of the measured regions is shown in Figure 4.

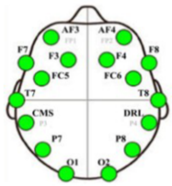


Figure 4. Emotiv headset sensors placement

The Emotiv system generates raw EEG data with 128 Hz sampling rate as well as the five frequency bands. The system uses internal algorithms to measure the following mental states: frustration, engagement, excitement, and valence. Although we don't have access to the system's proprietary algorithms, studies have provided evidence showing the reliability of its output [19]. We design our Neurofeedback system of Zoo Therapy based on these four emotional values.

2.3.2. RL-based Agent

In the past, our agent used the rule-based method. In this method, the agent will issue instructions to the animal based on the human-made feedback rules of the emotion that changes fastest per unit time among the four emotion values. However, this approach does not account the correlation between multiple emotions and cannot simultaneously show an increase or decrease in their value. The rules are artificially provided in advance, it is difficult to cover all changes in the emotional value of all users. Hence, we develop a reinforcement learning-based Neurofeedback system. The benefit of RL-based method is that it can learn the rules from the user's emotion in a flexible way, especially when the artificial rules are not accurate and hard to extract. Reinforcement learning emphasizes how to act based on the environment to maximize the expected benefits. It is inspired by behaviorism in psychology, that is, how an organism, stimulated by environmental rewards or punishments, gradually forms the expectation of the stimulus, and produces the habitual behavior that can obtain the maximum benefit. An agent must exploit existing experience to gain benefits, as well as explore unobserved status so that it can make better choices.

Our RL model has three components: agent, Zoo environment, and reward function. The agent is the decision-maker who will select an action from the action space given by Zoo Therapy and send it to the Zoo environment. The Zoo environment is the virtual animals and users which is the medium of agent interaction. The user's emotional value, as the agent's goal to maximize, is the reward. To fully reflect the user's emotions, we designed and trained a deep autoencoder model to encode 4 emotional values into one value to the latent representation of the current state of user emotions, called comprehensive emotional value. This value is used as an important component of the reward function. The reward function of this approach is defined such that if the user's comprehensive emotional value increases positively, then the reward is positive and vice-versa. We adopted the open-source codebase of the Proximal Policy Optimization (PPO) algorithm [24] to train the model.

Since the EEG changes of different users may change at any time and are unpredictable, our RL-based agent can perceive the user's emotion fluctuations, and when it finds that the user's comprehensive emotional value has decreased, it will immediately take corresponding actions to try to improve. Instead of giving instructions based on rules like the rule-based method at a fixed time.

3. Experiments

Our Zoo Therapy will be used to stimulate positive emotions and reduce negative emotions, thereby achieving the effect of slowing the progression of Alzheimer's Disease.

To analyze the effectiveness of our approach, we aimed to experiment the entire Zoo Therapy System with participants, unfortunately, due to the Covid-19 pandemic, experiments could not be performed with VR Zoo Therapy, but we tested the three research questions in this article separately.

(1) Is it possible to interact by the user's voice in a virtual reality environment? We tested the Zoo Therapy in a non-VR setting with the speech recognition system on five real participants of different ages, genders, and accents. This test is mainly to verify whether the speech recognition system works in Zoo Therapy and how effective it is. We let users experience using 3D buttons or the speech instruction system to interact with animals, and then tell the researchers their overall experience and rating Zoo Therapy after use.

(2) The Neurofeedback system can send corresponding instructions to the virtual reality environment in real-time to intervene on the user's emotions? We used the synthetic AD EEG data to train the agent and simulate the experiment of the RL-based Neurofeedback system. We generated 12 distributions of four different EEG emotions and the EEG data of AD patients comes from a similar project [20]. We used these 12 distributions to randomly generate a hundred sets of emotional values to simulate users and test our system. This set of experiments test whether the RL - based Neurofeedback system can learn rules from simulated user's EEG data and can give feedback when the emotional value changes. Simulated users can be roughly divided into three categories: like animals, don't like animals, and neutral.

(3) Is it possible to reduce the user's negative emotions by interacting with virtual animals? We tested the Zoo Therapy environment in a follow-up project [23], it combined the VR Zoo Therapy with gesture recognition and rule-based neurofeedback system and proved the effectiveness by evaluating one participant's EEG emotion value. We started by equipping the participant with an EEG headset and then added the Fove VR headset in which we installed the Leap Motion devise (used for gesture prediction), and the participant started the immersive experiment. In the whole process, the user's emotion value detected by the EEG helmet can reflect in real-time whether the user's negative emotions have decreased.

4. Results

The following results correspond to the experiments of each research question.

(1) Is it possible to interact by the user's voice in a virtual reality environment? We test Zoo Therapy with speech recognition system with five participants. All 5

participants said that the speech interaction method brought surprises to people. Although the virtual horse is in the computer, compared with a mouse click in 3D buttons on menu panels, the animal's direct response to the participant's voice is still delightful. The participants expressed that they prefer such animals and want to continue to interact with them. Participants rated the Zoo Therapy with speech command system from 0-10 subjectively in the "Evaluation" based on their own experience, the average evaluation of the five users was 7. Figure 5 shows the model's learning curve. In the test, the prediction accuracy of the speech recognition model for the user's voice instruction is 98%-99%.

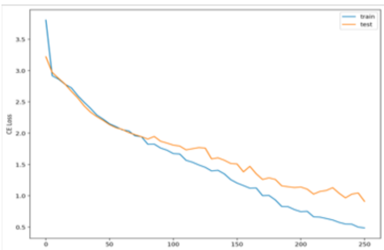


Figure 5. Cross Entropy Loss Curve

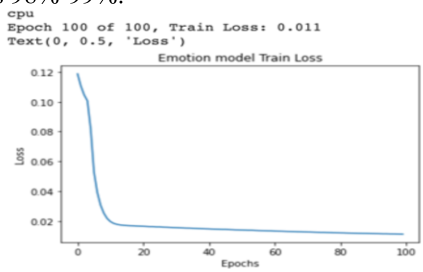


Figure 6. The train loss curve of autoencoder model

(2) The Neurofeedback system can send corresponding instructions to the virtual reality environment in real-time to intervene on the user's emotions? This part of the experiment consists of two parts, the autoencoder model, and the RL model.

The deep autoencoder model to reflect the user's four-dimensional emotions (Mentioned in 2.3.1) in one-dimension comprehensive emotional value (red line in Figure 7) and serve as an important part of the reward function of the reinforcement learning model. Autoencoding is first mapping the 4 emotions as a four-dimension vector into a one-dimension latent representation, and then reconstruct this four-dimension vector from the latent representation. It can be viewed as feature extraction and dimension reduction method, i.e., reduce the 4 emotions dimension to the one comprehensive emotional value, while still maintaining the reconstruction ability. The model is trained by minimizing the reconstruction error between the input four-dimension vector and the reconstructed four-dimension vector. Through 100 epochs of learning, it can be seen from Figure 6 that the model can quickly converge, and the MSE error of reconstructing 4 emotion values is 0.011.

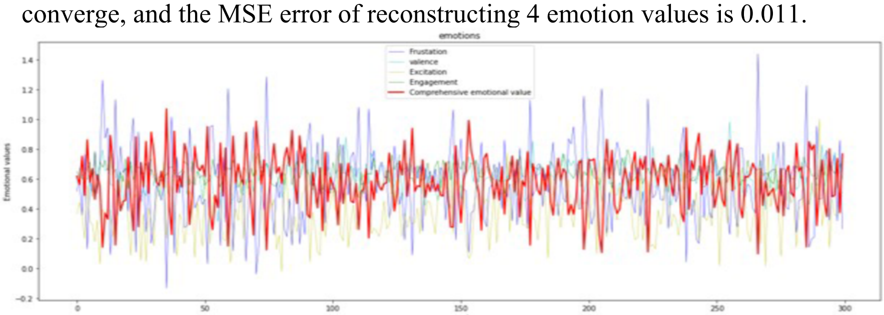


Figure 7. The effect of the autoencoder model is shown at a continuous point in time.

We used this model to train and test the RL-base agent. The behavior of RL agents can be divided into 3 parts. ①: For users who like animals, the RL-based agent gives proximity instructions, which account for about 63% of all instructions for this kind of user and 60-70% of the approach instructions occur when the user's emotion drops. For simulated users who don't like animals, the approach instructions given by the RL-based agent account for 59% of all instructions for this kind of user. The natural user command accounted for 54% of all instructions. ②: The instruction to leave is issued more times to users who don't like animals. But the gap between the three types of people is not large. ③: For the instruction of "stand up", the number of times sent to users who don't like animals is 120%-130% of the number of times sent to users who like animals.

①and ②show the system will bring animals closer to users who like animals, and less proactively approach users who do not like animals. Through the analysis, ③ is the product of the combined effect of multiple emotions. For users who don't like animals, when the animal is very close to the user, the user's frustration value may rise. The other three emotional values may also rise. When the system chooses to stand up, the target location of this behavior is a certain distance away from the user. It will not be too close to the user to cause negative emotions but try to get closer to increase other emotional values. Through learning, the agent finds a strategy to balance the relationship between distance and multiple emotions to maximize the user's positive emotions. The agent has learned a potential feedback rule.

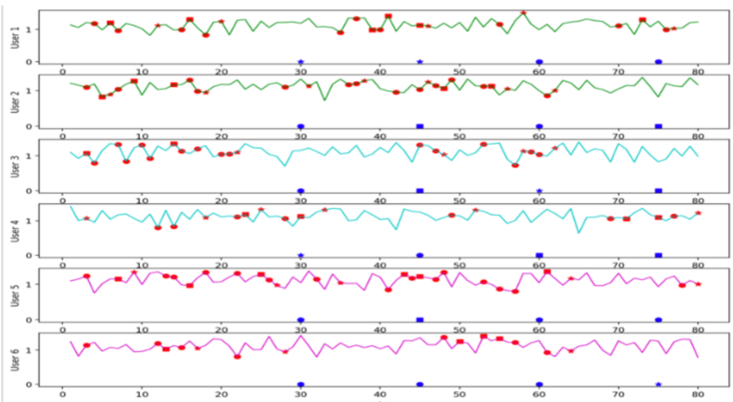


Figure 8. 6 of 100 experimental results. The green line is the EEG curve of simulated users who like animals, cyan corresponds to users who don't like animals, and purple corresponds to natural users. RL-based agent: red markers. Rule-based agent: blue markers. The circles represent "come" or "eat", the squares represent "stand up", and the stars represent "walk" or "leave".

In addition, it can be seen from Figure 8 (blue markers) that the rule-based method sends instruction at a fixed time, and it is unable to give feedback on the emotional changes between the two instructions. Its characteristic is that it will give instructions in strict accordance with our rules. The RL-based method will give different strategies for different simulated users. From the red markers in Figure 8, the timing of the RL-based agent's instructions is more flexible. It doesn't need to wait for a fixed time but immediately gives feedback based on real-time emotional changes. It is more sensitive to emotional changes and can be changed at any time after the order is issued. The most important thing is that it can be generalized to a more comprehensive data set and learn new potential rules what we didn't know.

(3) Is it possible to reduce the user's negative emotions by interacting with virtual animals?

According to the emotion value recorded in the cooperative project, we analyzed the mean frustration of the participant before, during, and after Zoo Therapy. Results show that, before the therapy, the mean frustration was 0.524, during Zoo Therapy, the mean frustration was 0.429, and after the mean frustration was 0.486. Figure 9 shows the difference between the mean frustration before, during, and after Zoo Therapy.

Although this result is carried out under the gesture interaction method, in view of the positive feedback from users under the speech instruction system, we believe that the use of speech interaction and the enhanced intervention of the Neurofeedback system can also achieve the same positive effect of reducing the user's negative emotions.

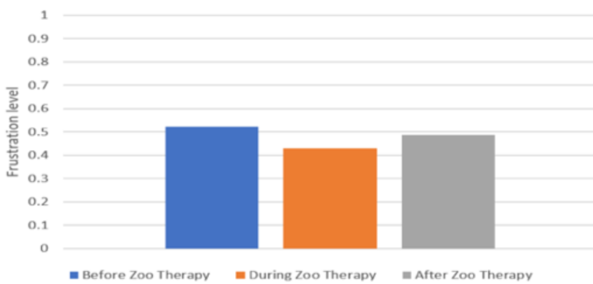


Figure 9. Histogram of general mean frustration

5. Conclusion and future work

In this research, we aim to design and develop a virtual reality animal treatment that can replace real AAT for Alzheimer's disease, called Zoo Therapy, which stimulates the user's positive emotion feedback through the interaction with the virtual animal dog or virtual animal horse to reduce the user's negative emotions and relieve AD symptoms. The salient feature of Zoo Therapy is that in the 360-degree VR HD scenario, users can interact with the virtual animals through multi-interaction methods.

We created a VR environment in which users can interact with virtual animals. We designed three ways of interactions in the Zoo Therapy: 3D button, speech instruction, and the Neurofeedback system. Using the 3D button, the interaction is as simple as clicking on the required action for the animal. The speech recognition system using an automatic speech recognition (ASR) model that can accept trained voice instructions and the animals can follow them as if they can hear the users in real-life. The Neurofeedback system can instruct the animals to react as per the user's emotions.

Due to the Covid-19 epidemic, it was not possible to perform the experiments with real subjects. Hence, we performed preliminary experiments with some lab students and reported the results in this study. We get a good result for the one real AD patient. We received positive feedback regarding the design of the Zoo Therapy. The users were very excited with the horse's response and the overall experience of the Zoo Therapy elicited positive emotions in them. The Neurofeedback system was tested on a hundred simulated users, and it was able to instruct the animals by considering the user's current emotional state. All that show our Zoo Therapy will lead a positive emotion and might be used to reduce AD symptoms.

In the future, Zoo Therapy plans to invite patients to experience. In particular, the Neurofeedback system will use real patient EEG data for training and adjustment, to improve the accuracy of real-time intervention of the model in users' emotions. We will also try to extend it to relevant projects that can perceive users' emotions and give feedback, to investigate whether the model can autonomously learn new feedback rules and intervene with patients in different specific environments.

Acknowledgments

We acknowledge NSERC-CRD (National Science and Engineering Research Council Cooperative Research Development), Prompt, and BMU (Beam Me Up) for funding this work.

References

[1] ALZHEIMER'S ASSOCIATION REPORT: 2020 alzheimer's disease facts and figures. https://doi.org/10.1002/alz.12068, 2020.
[2] Guy M McKhann, David S Knopman, Howard Chertkow, Bradley T Hyman, Cliord R Jack Jr, Claudia H Kawas, William E Klunk, Walter J Koroshetz, Jennifer J Manly, Richard Mayeux et al.: The diagnosis of dementia due to alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. Alzheimer's & dementia, 7(3):263–269, 2011.
[3] Guy G Potter et David C Steffens: Contribution of depression to cognitive impairment and dementia in older adults. The neurologist, 13(3):105–117, 2007.
[4] Xiaoyan Sun, David C Steffens, Rhoda Au, Marshal Folstein, Paul Summergrad, Jacqueline Yee, Irwin Rosenberg, D Mkaya Mwamburi et Wei Qiao Qiu: Amyloid-associated depression: a prodromal depression of alzheimer disease? Archives of general psychiatry, 65(5):542–550, 2008.
[5] DB Miller et JP O'callaghan : Aging, stress and the hippocampus. Ageing research reviews, 4(2): 123–140, 2005.
[6] Brian K. Alldredge.: Applied therapeutics: the clinical use of drugs 10th. baltimore: Wolters kluwer health/lippincott williams & wilkins. ISBN:978-1609137137, 2013.
[7] Forest Pharmaceuticals: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5329133//, 2008.
[8] Clive Ballard, Maria Luisa Hanney, Megan Theodoulou, Simon Douglas, Rupert McShane, Katja Kossakowski, Randeep Gill, Edmund Juszczak, Ly-Mee Yu, Robin Jacoby et al.: The de- mentia antipsychotic withdrawal trial (dart-ad): long-term follow-up of a randomised placebo-controlled trial. The Lancet Neurology, 8(2):151–157, 2009.
[9] Hamdi Ben Abdessalem, Alexie Byrns, Marc Cuesta, Valeria Manera, Philippe Robert, Marie- Andrée Bruneau, Sylvie Belleville et Claude Frasson : Application of virtual travel for alzheimer's disease. In SENSORNETS, pages 52–60, 2020.
[10] Alexie Byrns, Hamdi Ben Abdessalem, Marc Cuesta, Marie-Andrée Bruneau, Sylvie Belle- ville et Claude Frasson: Adaptive music therapy for alzheimer's disease using virtual reality. In International Conference on Intelligent Tutoring Systems, pages 214–219. Springer, 2020.
[11] Caroline Dakoure, Hamdi Ben Abdessalem, Marwa Boukadida, Marc Cuesta, Marie-Andrée Bru- neau, Sylvie Belleville et Claude Frasson : Virtual savannah: An eective therapeutic and relaxing treatment for people with subjective cognitive decline. In International Conference on Brain Function Assessment in Learning, pages 107–112. Springer, 2020.
[12] Sibel Cevizci, Halil Murat Sen, Fahri Güne et Elif Karaahme : Animal assisted therapy and activities in alzheimer disease. InTech Publishing: Rijeka, Croatia, 2013.
[13] Sandra Wesenberg, Christoph Mueller, Frank Nestmann et Vjera Holthoff-Detto: Eects of an animal-assisted intervention on social behaviour, emotions, and behavioural and psychological symptoms in nursing home residents with dementia. Psychogeriatrics, 19(3):219–227, 2019.
[14] Antonio Santaniello, Susanne Garzillo, Alessia Amato, Mario Sansone, Annalisa Di Palma, Annamaria Di Maggio, Alessandro Fioretti et Lucia Francesca Menna : Animal-assisted therapy as a non-pharmacological approach in alzheimer's disease: A retrospective study. Animals, 10(7):1142, 2020.
[15] Antonio Santaniello, Mario Sansone, Alessandro Fioretti et Lucia Francesca Menna : Systematic review and meta-analysis of the occurrence of eskape bacteria group in dogs, and the related zoonotic risk in animal-assisted therapy, and in animal-assisted activity in the health context. International journal of environmental research and public health, 17(9):3278, 2020.

[16] Lisa Maria Glenk : Current perspectives on therapy dog welfare in animal-assisted interventions. Animals, 7(2):7, 2017.
[17] Marta De Santis, Laura Contalbrigo, Marta Borgi, Francesca Cirulli, Fabio Luzi, Veronica Redaelli, Annalisa Stefani, Marica Toson, Rosangela Odore, Cristina Vercelli et al. : Equine assisted interventions (eais): Methodological considerations for stress assessment in horses. Veterinary sciences, 4(3):44, 2017.
[18] Henrique Dantas Ferreira: Animal-assisted therapy for alzheimer patients using virtual reality. Master Dissertation, Universidade da Madeira, April 2019.
[19] Aspinall, P.,Mavros,P.,Coyne,R.,&Roe,J.(2015).Theurbanbrain: Analysingoutdoor physical activity withmobile EEG. British Journal of Sports Medicine, 49(4), 27
[20] Manish Kumar Jha : A 3d virtual reality orientation game with real-time guidance system for cognitive training. Master Dissertation,University of Montreal, 2020.
[21] AshishVaswani, NoamShazeer, NikiParmar, JakobUszkoreit, LlionJones, Aidan NGomez,ŁukaszKaiseretIlliaPolosukhin:Attention is all you need.In Advances in neural information processing systems, pages5998–6008, 2017
[22] Genta IndraWinata: End-to-end speech recognition on pytorch. https://github.com/gentaiscool?tab=repositories,2019.
[23] Hamdi Ben Abdessalem Yan Ai Marulasidda Swamy KS et Claude Frasson : Virtual reality zoo therapy for alzheimer's disease using real-time gesture recognition. Genetics, Geriatrics and Neurodegenerative diseases 2020, FORTH Foundation The Foundation for Research and Technology - Hellas (FORTH) Heraklion, Crete, Greece, October 8-11, 2020
[24] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford et Oleg Klimov : Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347, 2017.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210080

A Framework for Personalized Fully Immersive Virtual Reality Learning Environments with Gamified Design in Education

Andreas MAROUGKAS[1], Christos TROUSSAS, Akrivi KROUSKA and Cleo SGOUROPOULOU
Department of Informatics and Computer Engineering, University of West Attica, Greece

Abstract. Traditional learning methods frequently fail to provoke students' interest, stimulate their enjoyment and encourage them to participate in learning activities, resulting in discomfort, distractions, and disengagement, if not quitting. Education's goal is to improve the quality and effectiveness of teaching and learning methods. This paper aims to present a framework based on Virtual Reality (VR) technology and contemporary Head Mounted Displays, that incorporates game-based techniques and adaptive design according to the student's profile. As a result, this paper analyzes the relevant literature, the VR apparatus, the importance of VR, as well as gamification, personalization and adaptive design in education, which are the learning foundations on which the framework is based. Finally, the framework's modules and structure are presented, taking into account all of the previously mentioned parameters. This novel framework aspires to serve as a basis for educational applications that use immersive Virtual Reality technologies to transform learning procedures into entertaining, engaging, enjoyable, and effective experiences.

Keywords. Virtual Reality, framework, Virtual Reality Learning Environments , personalization, gamification

1. Introduction

Education is becoming a more demanding and complex field over the years with educators trying to deliver more efficient ways of transferring knowledge to the class. Traditional teaching methods have many disadvantages towards students. Educators have tested, researched and employed a wide variety of methods and platforms to facilitate educational approaches such as social networks [1,2], desktop computers [3], smartphones and tablets [4], Augmented Reality (AR) [5], mobile applications and online platforms [6].

The technological advancements of the last decade led researchers and educators to explore alternative methods of teaching students by taking advantage of new upcoming technologies. Virtual Reality (VR), Mixed Reality (MR) [7] and Augmented Reality

[1] Corresponding Author; E-mail: amarougkas@uniwa.gr.

(AR) [8] are very promising approaches for an effective way of teaching and learning trying to solve various problems. Disengagement, lack of interest and enjoyment, reduced participation, discomforts and distractions are only a few of the problems that lead to students' poor learning outcomes and false perceptions about the subject of learning. VR made a giant leap from passive learning methods to more interactive experiences that students find useful, exciting, approachable, enjoyable, interesting and engaging [9].

After decades of failed attempts to be established in tech community, the hardware limitations related to severe oculomotor malfunctions, such as nausea and vertigo, have finally be solved. Smartphones can be incorporated with cardboard headsets delivering an affordable alternative access to VR applications. But the most noteworthy hardware is the Head Mounted Displays that offer fully immersive experiences accompanied by hardware controllers, hand tracking and gesture recognition. They also offer high visual fidelity, option to use it as a stand-alone wireless device or to connect it with a computer for more complex graphics. Contemporary HMDs offer a new path to education where the student is offered active participation in the learning process and a hands-on experience.

VR provides a new fresh, exciting and effective method of learning [10] for students, but sometimes that is not enough for students to capture their interest across the learning process. As an innovative approach to adaptive learning, this paper presents a framework that incorporates gamification design and personalization techniques into VR environments. Although gamified VR applications have been successfully employed, there is a lack of incorporating adaptive learning design in gamified Virtual Reality Learning Environments (VRLEs). It is essential to take into consideration students' unique characteristics and abilities to diversify the learning process creating a dynamic system serving students' academic requirements [2].

2. Relevant work

The improvements in VR hardware throughout the last decade have led many researchers to reformulate Virtual Environments for educational purposes and to further explore the potential of VR in the field of education. In certain cases, VR has been used passive learning [11]; while in other cases, the student is offered the opportunity of integrating with the virtual world as part of the learning [12,13].

Cochrane et al. [14] proposed a design-based research (DBR) framework for implementing VRLEs to mobile devices. Furthermore, they presented two ongoing projects for higher education use as examples.

The findings of a thorough review in relevant studies led us to the conclusion that the most notable approaches to VR regarding education incorporate game-based logic and personalization techniques. Even if a VRLE is generally effective in terms of achieving the desired educational outcomes, this does not indicate that students perceive it to be an appealing learning method [15]. As a result, the willingness to use it and acknowledge it as a learning method could be negatively impacted, potentially leading to disengagement during the learning process.

Mapping gamification techniques to educational applications is a way to pique students' interest, engage and motivate them in the learning process. Hung et al. [16] developed a Digital Game-Based Learning (DGbL) environment for elementary school mathematics based on e-books. Although the application was not supported by a VR approach, it demonstrates the effectiveness of implementing gamification strategies in Learning Environments with immersive technologies.

Rychkova et al. [13] presented “Orbital Battleship”, a game-based VR application that combined educational content in the field of chemistry with multiplayer gameplay. The application produced beneficial results for students, exceeding the authors' expectations. Despite the gamification techniques and a fully immersive Head Mounted Display used, as well as the positive learning outcomes emerged, the application was not designed to provide students with an adaptive experience that could be dynamically adjusted to their academic needs.

Wilson et al. [17] evaluated a VR ophthalmoscopy training application using gamification techniques, such as in-app, virtual awards to capture students' attention and keep them engaged throughout the procedure. Students felt a sense of accomplishment, which is one of the intrinsic factors that can lead to student retention and engagement. Despite the fact that the application provided methods for personalizing the learning process, the system lacked adaptivity in accordance with the academic requirements of the students.

Zhang et al. [18] created a VR simulation for fire safety education. Through their exploration of the virtual environment, students were given the ability to interact with virtual elements in order to construct their knowledge. The authors applied gamification techniques to captivate students' interest, such as reward strategies, that made the procedure appear both challenging and appealing. Through interactions, students received personalized feedback on learning errors, allowing them to learn more effectively. However, due to the subject's nature, it did lack adaptive techniques that could target students' learning needs and personal preferences.

3. Virtual Reality as a medium

Virtual Reality is built around a computer-generated space that enables the user to participate in a three-dimensional simulation while feeling immersed and experiencing realism. The element of interaction between the user and the virtual environment is a fundamental aspect of VR. Over the years, researchers and developers have used the term "Virtual Reality" to refer to a wide range of hardware devices such as CAVE VR [19], Desktop VR [3], Mobile VR [20], Head Mounted Device VR [13] glasses using stereoscopic imaging technology [21] and more. During the last decade, VR finally laid the groundwork for mass acceptance. After years of research and development, companies, such as HTC and Oculus, made Head Mounted Display devices accessible to the public through retail versions, which are now intertwined with the definition of VR. HMDs are now a well-established type of hardware that is widely available on the market.

Head-Mounted Displays efficiently transport the user from the physical to the virtual world by combining visual simulations, sounds and real-time interactions with virtual objects. Users currently employ two types of VR headsets depending on their requirements and preferences: hardware systems, such as the Oculus Quest and HTC Vive, and Cardboards, which use smartphone devices.

Each specific device is chosen based on what users evaluate that it is appropriate in each unique circumstance, with each having its own set of key advantages and disadvantages, as shown in Table 1.

Table 1. An apparatus comparative chart

	Hardware HMD's	Cardboards
Advantages	Higher visual fidelity Six Degrees of Freedom (6-DoF) provide movement flexibility Better interaction experience with virtual environment with to two advanced controllers	Easy access with any smartphone device Cost effective for group use applications
Limitations	Requires acquisition of new device High costs for group use applications	Three Degrees of Freedom (3-DoF) provide limited movement flexibility Lower visual fidelity

4. Virtual Reality in education

4.1. Virtual Reality as a Virtual Learning Environment

Throughout the last decade, Virtual Reality has gradually attracted more educators and researchers to work on projects demonstrating that it is a promising teaching and learning tool. Despite the fact that it is still under-researched technological approach in terms of its potential, it has already been implemented in various fields of education such as architecture [22], biology [20] chemistry [23], engineering [10], health sciences [17], physics [24], religion [25] and video game development [26].

When compared to traditional Virtual Learning Environments, Virtual Reality Learning Environments (VRLE) are becoming a more effective and appealing alternative. VRLEs have been used mostly in desktop computers for decades as part of the learning process, being limited in terms of interaction and field of view. They make use of a keyboard, mouse, and monitor, but VRLEs make use of advanced controllers or even hand gestures. VRLEs also provide a broader field of view with no framed boundaries, resulting in a more intuitive sense of freedom, interaction and perspective within the artificial environment, promoting a "Learning-by-doing" experience [27].

4.2. Benefits

In accordance with the "Flow Theory", VR immerses users in the experience, while their attention is completely absorbed within the virtual simulation [28]. With contemporary VR headsets, users can achieve high levels of immersion and embodiment. Using controllers, they interact in real-time with the environment, transforming the overall experience into an effective aid to students' learning and developing a sense of being there. VR allows users view realistic portraits of real-world environments, situations and objects, enhancing their spatial presence [11].

VR applications in education, taking into account all of the aforementioned properties, can have a positive impact on a variety of aspects of students' perceptions, like perceived learning motivation, which encourages them to participate in related activities [9]. Furthermore, the intention to use it increases enjoyability because students find it appealing [29].

Moreover, students lose interest in traditional in-class learning events, resulting in a circumstance in which their educational results are harmed. Students' interest can be captivated when VR opens up a previously uncharted territory, encouraging them to not only participate more than they normally would, but also to be more involved in the learning process, which can lead to improved learning outcomes [11].

VR has a high didactic and practical value, and when used in conjunction with traditional methods, it has the potential to improve students' understanding [29], assist them to master knowledge [30] and produce better learning outcomes [31]. It may also enable students to carry out curriculum-related experiments and procedures that would otherwise be prohibitively expensive, dangerous, or impossible [32].

4.3. Limitations

Despite the benefits and potential of using VR as a teaching or learning method, some significant limitations must be considered. For example, despite the limited amount of time the user operated the system, some oculomotor-related symptoms such as nausea have been reported in some cases [12]. Dizziness, vertigo, eye strain, and headaches are also symptoms of users' sensory conflict. Therefore, it is critical to check users' backgrounds before involving them in any VR-related process, due to cybersickness symptoms [33].

5. Gamification

5.1. A gamified classroom

When compared to commercial video games, in-class learning activities reveal similar behavioral patterns among students and gamers [34]. For example, both allow students to prioritize their tasks, as in turn-based role-playing games, design and develop strategies in the same way as in real-time strategy games. Moreover, they can develop new skills and improve old ones through practice, such as in sports games; learn to be adaptable and creative in new situations by playing adventure games; develop critical thinking skills, as in puzzle games; learn to be persistent, patient and goal-oriented, just like in action role-playing games and in multiplayer and online games, where players can change their gaming style to adjust to group activities.

Gamification manages to combine commercial games and education by employing game design, game mechanics, and techniques in non-gaming contexts of the curriculum to improve students' learning experiences by making them quite enjoyable and engaging. This can help to solve the problem of students' lack of interest in getting involved in school activities by increasing their motivation and interest in the relevant curricula.

5.2. Gamification and Virtual Reality in education

The logic, techniques, and mechanics of video games can be successfully implemented in the frameworks of Virtual Reality Learning Environments to create simulations based on educational content.

Students can interact with the artificial environment and interfere in assigned tasks designed according to their educational needs using VR controllers, which are individually created for each head-mounted display system, just as they would do in video game quests of commercial games using ordinary video gaming controllers.

When combined with gamified features, VR educational content that focuses on student motivation and engagement can lead to better learning outcomes [18], improve memory retention [33] and enhance content knowledge [35]. Gamification techniques of various types can be incorporated into Virtual Reality Learning Environments for educational purposes. Live frequent updates on information about students' scores, ongoing state, progress status, and milestones achieved in each stage are among the features that can assist students throughout the VR learning process. Furthermore, rewarding students with badges for game-like achievements, and leaderboards as a projection of ranking based on their scores [15] are both effective methods of improving student retention and enhancing their excitement. Students can also collaborate or compete in a multiplayer mode [13]. A help support repository in the menu, different levels of difficulty, objectives, puzzles, and item inventories are also gamification techniques.

6. Personalized Virtual Learning Environments

An important way to support a virtual learning environment efficiently is to use personalization techniques with the goal of delivering a tailored experience based on student profiling. This can be accomplished by taking into account students' preferences, interests, needs, capabilities, skills, prior knowledge of the subject, learning styles, and even special needs, thereby addressing the challenges of the heterogeneous features of twenty-first-century education. A personalized Virtual Learning Environment can benefit students by assisting them in improving their content knowledge, learning outcomes, and academic achievements. It also eliminates any perceptual discomforts that may arise during the learning process, such as cognitive load, frustration, or a lack of interest. These discomforts may act as impediments to learning, resulting in ineffective or distorted learning outcomes, or even quitting.

To clear up any confusion, a personalized learning environment is an adaptive automated system based on individualization, as opposed to a customizable environment, which requires manual adjustments before it can be applied to any educational intervention.

Various techniques and theories can be used to compose a model design with flexible modules, dynamic data filtering, data logging for registering behavioral parameters and preferences, and thus creating an application based on the cognitive diversity and mentality of each student [36].

Troussas et al. [6] presented a model of learning analytics regarding learning environments which included multiple modules such as cognitive state and behavior prediction module, identification of targeted educational material module, curriculum improvement module and personalization module. The findings of this study, which included students' evaluations of a web-based prototype application based on the aforementioned modules, were very promising, as students perceived the system as an efficient way to participate in a personalized experience.

Regarding learning analytics in VR, Srimadhaven et al. [37] proposed an architecture of a VR mobile game for undergraduate students related to the subject of programming languages. In this project, students had to complete three stages of the game by performing coding tasks. Each level provided a different level of difficulty, allowing students to earn points based on question correctness. The authors designed the system's architecture to assess students' cognitive behavior, self-regulation, affective behavior, and self-efficacy in order to improve students' academic performance, with learning analytics assisting the entire process by providing a personalized experience based on student results individualization.

7. A proposed framework

After the thorough literature review presented above, we came up with the conclusion that the integration of VR Learning Environments with personalized gamified mechanisms that take into account students' cognitive states and behavioral patterns while also balancing gameplay and learning objectives has never been approached with fully immersive VR Head Mounted Displays as medium hardware. It is critical to identify students' learning gaps in order to provide a tailored solution that not only meets his/her needs and competencies [38], but also to pique his/her interest and increase excitement, resulting in an engaged and overall enjoyable experience.

The following proposed framework (as shown in Fig. 1) aspires to be a reference point for future research and educational designs, enabling the inclusion of all of the above-mentioned approaches. The framework is comprised of the pedagogical design, which includes the formulation of educational objectives, learning theories, and strategies as part of the student-centered approach. The Technological Design includes all of the technological mechanisms, software, and hardware required to connect the theoretical design to the artificial environment. The Data Collection Layer is responsible for the generation, storage, and retrieval of data produced by the user's interactions with the system.

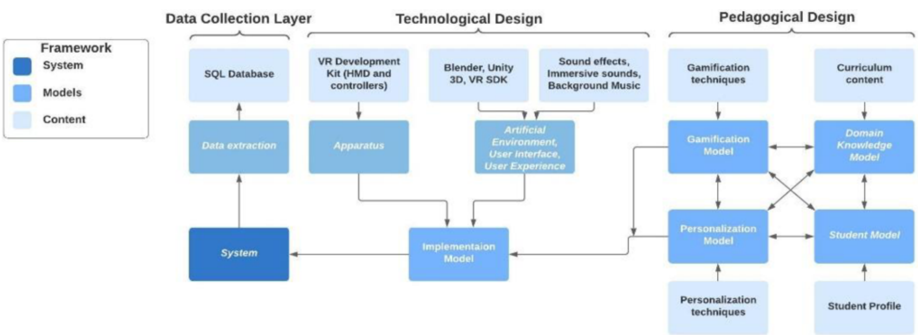


Figure 1. A framework for adaptive Virtual Reality applications using gamification and personalization techniques

7.1. Pedagogical Design

The Domain Knowledge Model contains the educational material needed to construct the specific knowledge that defines the curriculum's educational goals.

The Student Model constructs the student's profile, which is in charge of assisting the learning process. It is provided by a system that is designed to take into account student-related factors such as cognitive states, behavioral patterns, preferences, and prior knowledge related to the subject and provide relevant feedback and recommendations as the procedure progresses. This could be accomplished through data collection via questionnaires and surveys, or through an in-app procedure similar to those used in commercial games in the form of mini-games when the application first launches as the first task to complete.

The Gamification Model consists of the mechanisms that the system employs to motivate students to use the application and remain engaged throughout the learning process. It aims to increase student retention by providing an enjoyable experience through mechanisms similar to those found in commercial video games.

The Personalization Model works in tandem with the Gamification Model and Student Model to create a tailored learning experience that incorporates data input from student assessments during the Student Model phase, as well as gamification techniques that transform the learning environment into a more exciting and fun learning process that is tailored to students' competencies, skills, and preferences.

7.2. Technological Design

The Artificial Environment is where the student participates in the active learning. It is a three-dimensional virtual environment in which students can interact with virtual objects that have been created and configured to allow them to experience life-like situations. The environment consisted of 3D visual representations, interactions, User Interface, and User experience using the game development platform Unity3D or Unreal Engine that incorporate the Virtual Reality SDK provided by the Head Mounted Display Hardware Development Kit such as Oculus SDK or VIVE Wave SDK. Blender is a software that allows users to create 3D models and animations and offers the option to import them in the game development platform. Furthermore, in order to achieve high levels of immersion and provide students with a flow experience, the systems must produce immersive sound effects and background music.

A Head Mounted Display (HMD) hardware is required for this framework to allow the user to experience three-dimensional visualizations and to serve as a mediator for interactions between the user and the virtual environment and its virtual objects. Since contemporary HMDs use advanced controllers and hand gesture recognition systems, the HTC Vive and Oculus HMDs series are sufficient to meet the requirements of this framework. They also provide high-fidelity visualizations because advanced displays are used.

7.3. Data Collection Layer

Data extraction will occur after the required data has been collected as a result of the user's interactions with any virtual elements within the virtual environment. A database will gather all of the required grouped data for further analysis.

8. Conclusion

This paper describes a novel framework for designing adaptive, fully immersive Virtual Reality Learning Environments with gamified techniques and aims to become the blueprint for future researches and applications. Taking into account students' cognitive diversity and characteristics, the design of this framework aims to optimize the learning experience by transforming it into an effective, appealing, and enjoyable educational approach, as well as creating a tailored educational interactive environment that can be automatically adjusted to the students' profile.

References

[1] A. Krouska, C. Troussas and M. Virvou, Social networks as a learning environment: Developed applications and comparative analysis, *2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA)* (2017), 1-6. doi: 10.1109/IISA.2017.8316430.

[2] C. Troussas, M. Virvou, J. Caro, K. J. Espinosa, Language Learning Assisted by Group Profiling in Social Networks, *International Journal of Emerging Technologies in Learning* (IJET) **8**(3) (2013), 35.

[3] G. Makransky, R. E. Mayer, N. Veitch, M. Hood, K. B. Christensen, H. Gadegaard, Equivalence of using a desktop virtual reality science simulation at home and in class, *Plos One* **14**(4) (2019), e0214944.

[4] B. Haßler, L. Major, S. Hennessy, Tablet use in schools: a critical review of the evidence for learning outcomes, *Journal of Computer Assisted Learning* **32**(2) (2015), 139-156. doi: 10.1111/jcal.12123.

[5] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, User acceptance of augmented reality welding simulator in engineering training, *Education and Information Technologies* (2021).

[6] C. Troussas, Krouska A., Virvou M., Using a Multi Module Model for Learning Analytics to Predict Learners' Cognitive States and Provide Tailored Learning Pathways and Assessment, *In: Virvou M., Alepis E., Tsihrintzis G., Jain L. (eds) Machine Learning Paradigms. Intelligent Systems Reference Library, vol 158. Springer* (2020). Cham. https://doi.org/10.1007/978-3-030-13743-4_2

[7] R. Lindgren, M. Tscholl, S. Wang, E. Johnson, Enhancing learning and engagement through embodied interaction within a mixed reality simulation, *Computers & Education* ***95 (2016)***, 174–187.

[8] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, Exploration of Augmented Reality in Spatial Abilities Training: A Systematic Literature Review for the Last Decade, *Informatics in Education* **20** (2021), 107–130. https://doi.org/10.15388/infedu.2021.06

[9] F. A. Monita, J. Ikhsan, Development Virtual Reality IPA (VR-IPA) learning media for science learning, *Journal of Physics: Conference Series* **1440** (2020), 012103.

[10] J. R. Brown, I. Kuznetcova, E. K. Andersen, N. H. Abbott, D. M. Grzybowski, C. D. Porter, Implementing Classroom-Scale Virtual Reality into a Freshman Engineering Visuospatial Skills Course, *2019 FYEE Conference* (2019).

[11] K. Cheng, C. Tsai, A case study of immersive virtual field trips in an elementary classroom: Students' learning experience and teacher-student interaction behaviors, *Computers & Education* **140** (2019), 103600. doi: 10.1016/j.compedu.2019.103600

[12] J. Pirker, I. Lesjak, C. Gütl, An Educational Physics Laboratory in Mobile Versus Room Scale Virtual Reality – A Comparative Study, *International Journal of Online Engineering (IJOE*) **13**(08) (2017), 106.

[13] A. Rychkova, A. Korotkikh, A. Mironov, A. Smolin, N. Maksimenko, M. Kurushkin, Orbital Battleship: A Multiplayer Guessing Game in Immersive Virtual Reality, *Journal of Chemical Education* **97**(11) (2020), 4184–4188. https://doi.org/10.1021/acs.jchemed.0c00866

[14] T. Cochrane, S. Cook, S. Aiello, D. Christie, D. Sinfield, M. Steagall, C. Aguayo, A DBR framework for designing mobile virtual reality learning environments, *Australasian Journal Of Educational Technology* **33**(6) (2017), 54-68. doi: 10.14742/ajet.3613

[15] P.Chen, The Design of Applying Gamification in an Immersive Virtual Reality Virtual Laboratory for Powder-Bed Binder Jetting 3DP Training, *Education Sciences* **10**(7) (2020), 172.

[16] C. Hung, I. Huang, G. Hwang, Effects of digital game-based learning on students' self-efficacy, motivation, anxiety, and achievements in learning mathematics, *Journal of Computers In Education* **1**(2-3) (2014), 151-166. doi: 10.1007/s40692-014-0008-8

[17] A. S. Wilson, J. O'Connor, L. Taylor, D. Carruthers, A 3D virtual reality ophthalmoscopy trainer, *The Clinical Teacher* **14**(6) (2017), 427–431. https://doi.org/10.1111/tct.12646

[18] K. Zhang, J. Suo, J. Chen, X. Liu, L., Gao Design and Implementation of Fire Safety Education System on Campus based on Virtual Reality Technology, *Federated Conference on Computer Science and Information Systems (FedCSIS)* (2017), 1297-1300. 10.15439/2017F376.

[19] P. Häfner, V. Häfner, J. Ovtcharova, Teaching Methodology for Virtual Reality Practical Course in Engineering Education, *Procedia Computer Science* **25** (2013), 251-260. doi: 10.1016/j.procs.2013.11.031

[20] A. P. Johnston, J. Rae, N. Ariotti, B. Bailey, A. Lilja, R. Webb, C. Ferguson, S. Maher, T. P. Davis, R. I. Webb, J. McGhee, R. G. Parton, Journey to the centre of the cell: Virtual reality immersion into scientific data, *Traffic* **19**(2) (2017), 105–110. https://doi.org/10.1111/tra.12538
[21] T. Barrett, M. Hegarty, Effects of interface and spatial ability on manipulation of virtual models in a STEM domain, *Computers in Human Behavior* **65** (2016), 220-231. doi: 10.1016/j.chb.2016.06.026
[22] C. Rosa-Jimenez, F. Ventura-Blanch, A transversal methodology for the implementation of virtual reality in Architecture learning, *Headache* (2020), 1347-1354. 10.4995/HEAd20.2020.11274.
[23] M. A. Z. Megat Zakaria, H. Abuhassna, Virtual Reality Acceptance in Classrooms: A Case Study in Teaching Science, *International Journal of Advanced Trends in Computer Science and Engineering* **9**(2) (2020), 1280–1294. https://doi.org/10.30534/ijatcse/2020/58922020
[24] D. Becerra, J. Herrera Quispe, R. Aceituno, G. Vargas, F. Zamora, J. Mango, G. Anccasi Figueroa, A. Vizcarra, J. Chana, Evaluation of a Gamified 3D Virtual Reality System to Enhance the Understanding of Movement in Physics*, Proceedings of the 9th International Conference on Computer Supported Education - Volume 1: CSEDU* (2017), 395-401. 10.5220/0006328003950401.
[25] W. Yahaya, A. Ahmad, The effectiveness of signaling principle in virtual reality courseware towards achievement of transfer learning among students with different spatial ability, *AIP Conference Proceedings* **1891** (2017), 020144. 10.1063/1.5005477
[26] E. Surer, M. Erkayaoğlu, Z.N. Öztürk, et al. , Developing a scenario-based video game generation framework for computer and virtual reality environments: a comparative usability study, *J Multimodal User Interfaces* (2020). https://doi.org/10.1007/s12193-020-00348-6
[27] O. A. Meyer, M. K. Omdahl, G. Makransky, Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment, *Computers & Education* **140** (2019), 103603. https://doi.org/10.1016/j.compedu.2019.103603
[28] S.-C. Wu, W.-H. Hsiang, The study of the communication effect on emotional design in virtual reality technology using VR application - Facebook spaces as example, *Towards Sustainable Technologies and Innovation - Proceedings of the 27th Annual Conference of the International Association for Management of Technology IAMOT* 2018. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85085666893&partnerID=40&md5=3d2636a012138e1dd890786055197e74
[29] W. Wizaka, G. Suharjanto, W. Wangidjaja, The new teaching method using virtual reality technology in building technology subject, *IOP Conference Series: Earth and Environmental Science* **426** (2020), 012080. https://doi.org/10.1088/1755-1315/426/1/012080
[30] J. Calvert, R. Abadia, Impact of immersing university and high school students in educational linear narratives using virtual reality technology, *Computers & Education* ***159** (2020)*, 104005.
[31] J. Ikhsan, K. H. Sugiyarto, T. N. Astuti, Fostering Student's Critical Thinking through a Virtual Reality Laboratory, *International Journal of Interactive Mobile Technologies (IJIM)* **14**(08) (2020), 183.
[32] S. Al Kork, T. Beyrouthy, Interactive Virtual Reality Educational Application, *Advances in Science, Technology and Engineering Systems Journal* **3** (2018), 72-82. 10.25046/aj030409
[33] J. Varela-Aldás, G. Palacios-Navarro, R. Amariglio, I. García-Magariño, Head-Mounted Display-Based Application for Cognitive Training, *Sensors* **20**(22) (2020), 6552. https://doi.org/10.3390/s20226552
[34] A.Domínguez, J. Saenz-de-Navarrete, L. de-Marcos, L. Fernández-Sanz, C. Pagés, J. Martínez-Herráiz, Gamifying learning experiences: Practical implications and outcomes, *Computers & Education* **63**(2013), 380-392. doi: 10.1016/j.compedu.2012.12.020
[35] J. Bendeck Soto, D. Toro Ocampo, L. Beltrán Colon, A. Valencia Oropesa, Perceptions of ImmerseMe Virtual Reality Platform to Improve English Communicative Skills in Higher Education, *International Journal Of Interactive Mobile Technologies (Ijim)* **14**(07) (2020), 4. doi: 10.3991/ijim.v14i07.12181
[36] A. Klašnja-Milićević, B. Vesin, M. Ivanović, Z. Budimac, E-Learning personalization based on hybrid recommendation strategy and learning style identification, *Computers & Education* **56**(3) (2011), 885–899. https://doi.org/10.1016/j.compedu.2010.11.001
[37] T. Srimadhaven & AV. Chris, N. Harshith, S. Jessenth, S. Shabari, M. Priyaadharshini, Learning Analytics: Virtual Reality for Programming Course in Higher Education, *Procedia Computer Science* **172** (2020), 433-437. 10.1016/j.procs.2020.05.095.
[38] A. Krouska, C., Troussas, M. Virvou, Computerized Adaptive Assessment Using Accumulative Learning Activities Based on Revised Bloom's Taxonomy, *In: Virvou M., Kumeno F., Oikonomou K. (eds) Knowledge-Based Software Engineering: 2018. JCKBSE 2018. Smart Innovation, Systems and Technologies, vol 108. Springer* (2019). Cham. https://doi.org/10.1007/978-3-319-97679-2_26

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210081

A Novel Framework Incorporating Augmented Reality and Pedagogy for Improving Reading Comprehension in Special Education

Andrianthi KAPETANAKI [a], Akrivi KROUSKA [a,1], Christos TROUSSAS [a] and Cleo SGOUROPOULOU [a]

[a] *Department of Informatics and Computer Engineering University of West Attica, Greece*

Abstract. Augmented Reality (AR) is an emerging technology thriving in recent years. The implementation of AR in education offers great opportunities to enhance educational environments achieving better learning outcomes. As students with learning disabilities struggle with reading comprehension, an AR learning environment provides them support to better understand texts they actually read. Even though few studies have tried to explore the impact of AR technology to reading comprehension for students with learning disabilities in Secondary Education, there is a lack of research grounded in the incorporation of learning theories and personalization technologies. The goal of this paper is to present an AR educational environment capable of supporting meaningful learning outcomes by taking into consideration each student special educational needs and learning style. The novelty of this study lies in the student-centered and personalized design, which leads to improved understanding, student interaction and self-learning.

Keywords. Augmented Reality, Educational Technology, Learning Disabilities, Personalized Learning, Reading Comprehension, Special Education.

1. Introduction

One of the main goals of education is to help students develop knowledge, skills and strategies, in order to become proficient and independent readers who read with meaning. The activity of reading includes seeing, perceiving, understanding, vocalizing and mentally constructing through the combined use of eyes, ears and brain [1]. However, students with learning disabilities struggle with reading comprehension which may be due to not only word recognition skills that are not automatic, but also to problems with working memory or to the fact that they may have difficulty in applying metacognitive strategies while reading a text [2].

Reading comprehension intervention can be provided by using a gamified approach and Augmented Reality (AR). Augmented Reality is a rapidly emerging technology that supplements the real world with virtual objects that appear to coexist in the same space as the real world [3]. AR books link text, image, sound and movement

[1] Corresponding Author, E-mail: akrouska@uniwa.gr.

as AR application on the mobile phone tracks images in the printed book and overlays virtual content onto them [4]. AR provides a multisensory learning environment with active students' involvement during the reading session [5]. Studies indicate that multisensory approach is the most effective teaching method for students with learning disabilities as it incorporates several learning pathways in the brain at the same time so that opportunities for memory and learning are increased [6].

The purpose of this paper is to investigate the exploitation of augmented reality in order to help students with learning disabilities to better understand texts in a meaningful student-centered learning environment.

2. Literature Review

According to the literature review, Augmented Reality is mainly applied in the field of Science [7, 8]. This may be related to the fact that AR learning environments facilitate comprehension of science concepts with the help of 3D models, improving students' laboratory skills [9, 10]. In the field of Social Sciences and Humanities, AR is mainly used in second language learning [11].

Although many studies have been conducted on AR use in education, only few of them are related to special education [12]. However, there is a growing interest in designing AR learning environments to support students with special educational needs. Researchers studied its affordance in skill acquisition of individuals with intellectual disabilities [13], autism spectrum disorder (ASD) [14], learning disabilities [7], hearing [15] and visual impairments [16]. It is remarkable that studies have shown that AR enables students with special educational needs to gain independent life skills, reduce behavioral problems and increase their level of academic achievements, enthusiasm and readiness by bringing them in real-life experiences [17].

In view of the above, the design and development of an AR system incorporating learning theories and personalization techniques targeting to enhancement of students' with learning disabilities reading comprehension skills is crucial. Student-centered and personalized design is expected to promote student interaction and self-learning and achieve better learning outcomes.

3. Research Methology

AR application is mainly focused on STEM education, while there are only few studies in Social Sciences and Humanities. Furthermore, research in AR targeting in reading comprehension for students with learning disabilities in Secondary Education is limited, while there is a lack of research grounded in learning theories, which could provide evidence how AR learning environments should be designed in order to improve learning outcomes. Additionally, AR systems designed are not personalized and not adapted in special educational needs of each particular student. Hence, within this context the research questions addressed by this study are the following:
RQ1: Does AR technology leads students with learning disabilities to the enhancement of their reading comprehension?
AR technology offers great opportunities to enrich learning procedure. Students who use AR material have a better performance in reading comprehension and learning

permanency compared to students who read with traditional methods [18]. Therefore, it would be quite interesting to investigate the effect of an AR environment on students with learning disabilities in Secondary Education.

RQ2: How could the pedagogical affordance of an AR system in special education be enhanced?

Many studies have been conducted in order to explore the impact of AR in education. However, only few of them implemented pedagogical theories. To optimize learning outcomes the incorporation of learning theories is needed in order to design a meaningful educational AR environment [19, 20].

RQ3: How could a student-centered AR learning environment be provided in special education?

A student-centered learning environment is focused on students' interests, abilities and learning styles [21, 22]. It is quite challenging to design a student-centered framework by using artificial intelligence techniques, in order to adjust the AR system to specific needs of students in special education. In that way, personalized learning is provided, satisfying the demand for inclusion and enabling students with special educational needs to take over an active role in learning process.

In view of the above, this paper proposes the design and development of an AR system with the aforementioned features, namely the incorporation of learning theories and personalization techniques, for supporting students with learning disabilities in reading comprehension.

In order to examine whether the use of AR may have a positive impact on this field, quasi-experimental design and a mixed method, combining the collection of both qualitative and quantitative data, will be used for evaluating the developed AR system. Qualitative data will be collected by interviews, while quantitative data will be collected by Likert-scale questionnaire. The sample will be composed of students with learning disabilities in Secondary Education. The experimental group will participate in reading activities of textbooks using AR application, while traditional methods will be used for control group.

Fig. 1 illustrates the steps of the research methology of this study.

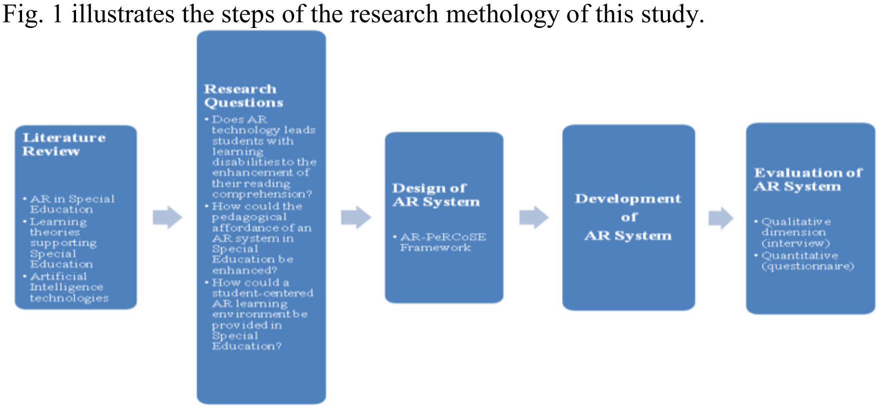


Figure 1. Research design of the study

4. The AR-PeRCoSE Framework : Augmented Reality and Pedagogy in Reading Comprehension in Special Education

This study introduces a novel framework, called AR-PeRCoSE, which uses AR technology in conjunction with pedagogical theories and AI techniques for promoting reading comprehension in Special Education. To this direction, school books will be used and enriched with AR content by providing real-time visual and audio feedback. Personalized AR content, consisted of image, audio, video and text, will be prepared, in order to allow multiple entry point to process learning and demonstrate understanding for each particular student. AR technology will be based on markers and a tablet will scan book pages and detect markers via an AR application. Markers and multimedia objects will be designed and saved in system's database in the cloud, so as to enable tablet's camera tracking. AR content will be designed taking into consideration domain model, pedagogical model by integrating learning and special education theories, student model, Augmented Reality (AR) and Artificial Intelligence (AI) techniques. As tablet's camera tracks the markers, the display will show supplemental material from which every student can choose what to use according to its needs and learning style. Interactivity between students and the device is expected to stimulate their interest and encourage them to study efficiently texts at their own pace.

Fig. 2 presents the proposed AR-PeRCoSE framework.

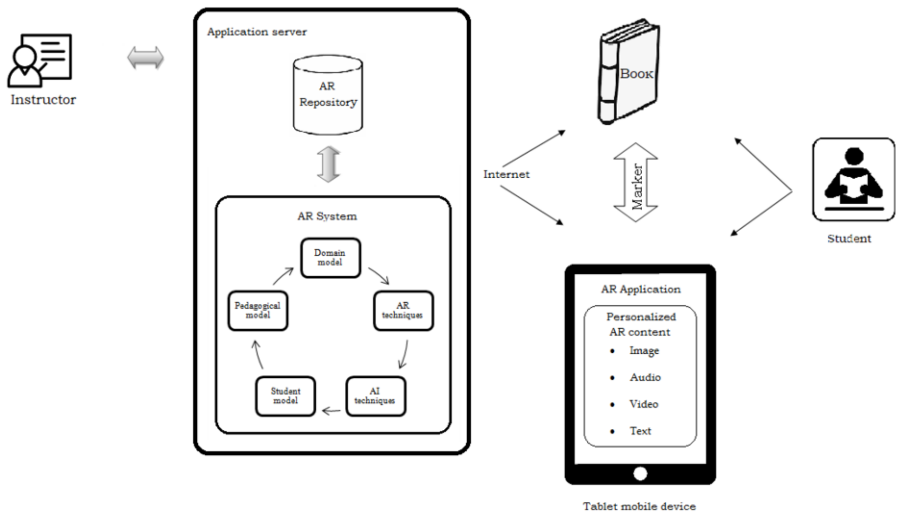


Figure 2. Augmented Reality and Pedagogy in Special Education Framework

5. Conclusion and Future Work

This paper introduces a novel framework which uses AR technology incorporating pedagogical theories and personalization techniques in order to enhance students' with learning disabilities reading comprehension skills. The main aim of this framework is to foster personalized student-centered learning environment and achieve better learning outcomes.

In addition, future studies need to use this framework so as to compare its feasibility to diverse groups of special educational needs, examine its usability and affordance to each particular group and detect potential gaps for its improvement. Furthermore, AR use could be further expanded for other skills acquisition for students in Special Education.

References

[1] Palani, K. (2012). Promoting Reading Habits and Creating Literate Society. *Researchers World, 3*, 90.

[2] Jitendra, A. & Gajria, M. (2011). Reading Comprehension Instruction for Students with Learning Disabilities. *Focus on Exceptional Children*. 43 (8). 1-16. https://doi.org/10.17161/foec.v43i8.6690

[3] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. Computer Graphics and Applications IEEE, 21(6), 34-47.

[4] Billinghurst, M. & Duenser, A. (2012). Augmented Reality in the Classroom. *Computer, 45(7), 56–63.* https://doi.org/10.1109/MC.2012.111

[5] Abas, H., & Zaman, H.B. (2011). Scaffolding models for remedial students in using augmented reality storybook. *Proceedings of the 2011 International Conference on Electrical Engineering and Informatics*, 1-5.

[6] Algrni N.S. (2020). The Effectiveness of Using Multisensory Approach in Enhancing Achievement and Retention of English Vocabulary Amongst Intermediate Female Students with EFL Learning Disabilities. *Journal of Education and Practice*. 11 (9).

[7] Bacca, J., Baldiris, S., Fabregat, R., Graf, S., & Kinshuk. (2014). Augmented Reality Trends in Education: A systematic review of research and Applications. *Educational Technology & Society*, *17* (4), 133–149.

[8] Papakostas, C., Troussas, C., Krouska, A., & Sgouropoulou, C. (2021). Exploration of Augmented Reality in Spatial Abilities Training: A Systematic Literature Review for the Last Decade. *Informatics in Education*, 20(1), 107-130.

[9] Papakostas, C., Troussas, C., Krouska, A., & Sgouropoulou, C. (2021). User acceptance of augmented reality welding simulator in engineering training. *Education and Information Technologies*, 1-27.

[10] Akçayır, M., Akçayır, G., Pektaş, H. M., &Ocak, M. A. (2016). Augmented reality in science laboratories: The effects of augmented reality on university students' laboratory skills and attitudes toward science laboratories. *Computers in Human Behavior*, 57, 334–342.

[11] Liu, Y., Holden, D., & Zheng, D. (2016). Analyzing students' language learning experience in an augmented reality mobile game: an exploration of an emergent learning environment. *Procedia-Social and Behavioral Sciences,* 228, 369-374. https://doi.org/10.1016/j.sbspro.2016.07.055

[12] Sirakaya, M., & Alsancak Sirakaya, D. (2018). Trends in Educational Augmented Reality Studies: A Systematic Review. *Malaysian Online Journal of Educational Technology*, *6*(2), 60-74.

[13] McMahon, D. D., Smith, C. C., Cihak, D. F., Wright, R., & Gibbons, M. M. (2015). Effects of digital navigation aids on adults with intellectual disabilities:

Comparison of paper map, Google maps, and augmented reality. *Journal of Special Education Technology*, 30(3), 157-165.

[14] Bhatt, S., De Leon, N., & Al-Jumaily, A. (2014). Augmented reality game therapy for children with autism spectrum disorder. *International Journal on Smart Sensing and Intelligent Systems*, 7 (2), 519-536.

[15] Mirzaei, M. R., Ghorshi, S., & Mortazavi, M. (2014). Audio-visual speech recognition techniques in AR environments. *The Visual Computer,* 30, 245–257.

[16] Albouys-Perrois, J., Laviole, J., Briant, J., & Brock, A. (2018). Towards a multisensory augmented reality map for blind and low vision people: A participatory design approach. In *CHI 2018, Conference on Human Factors in Computing Systems*, 21 - 26 April 2018, Montréal, Canada: ACM.

[17] Cakir, R., & Korkmaz, O. (2019). The Effectiveness of Augmented Reality Environments on Individuals with Special Education Needs. *Education and Information Technologies,* 24(2), 1631–1659. https://doi.org/10.1007/s10639-018-9848-6

[18] Bursali, H., & Yilmaz, R., (2019). Effect of augmented reality applications on secondary school students' reading comprehension and learning permanency, *Computers in Human Behavior*, 95, 126-135. https://doi.org/10.1016/j.chb.2019.01.035

[19] Krouska A., Troussas C., & Virvou M. (2019). Computerized Adaptive Assessment Using Accumulative Learning Activities Based on Revised Bloom's Taxonomy. In: Virvou M., Kumeno F., Oikonomou K. (eds) Knowledge-Based Software Engineering: 2018. JCKBSE 2018. Smart Innovation, Systems and Technologies, vol 108. Springer, Cham. https://doi.org/10.1007/978-3-319-97679-2_26

[20] Krouska, A., Troussas, C., & Virvou, M. (2017, August). Social networks as a learning environment: Developed applications and comparative analysis. In *2017 8th International Conference on Information, Intelligence, Systems & Applications* (IISA) (pp. 1-6). IEEE.

[21] Troussas, C., Krouska, A., & Virvou, M. (2020). Using a multi module model for learning analytics to predict learners' cognitive states and provide tailored learning pathways and assessment. In *Machine Learning Paradigms* (pp. 9-22). Springer, Cham.

[22] Troussas, C., Krouska, A., & Sgouropoulou, C. (2021). Impact of social networking for advancing learners' knowledge in E-learning environments. *Education and Information Technologies*, 1-21.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210082

GunitaHu: A VR Serious Game with Montessori Approach for Dementia Patients During COVID-19

Aira Jazel Y. ANG [a], Roberto D. PRINCIPIO Jr. [a], Richelle Ann B. JUAYONG [a] and Jaime D.L. CARO [a]

[a] *Service Science and Software Engineering Laboratory, Department of Computer Science, University of the Philippines Diliman*

Abstract.
The study aims to explore VR Serious Games as a form of therapy for people with dementia. It seeks to establish the utility of VR-based interventions with the application of Montessori Method. This study also serves as a basis for researchers, healthcare professionals, and developers who plan to incorporate VR therapy with other therapeutic approaches and to create a system that may be replicated for other illnesses via telemedicine to address the most vulnerable sectors. The main beneficiaries of this study are people with dementia and those who directly interact with them such as their doctors, caregivers, and family members of the patient.

Keywords. dementia, virtual reality, telemedicine, montessori method

1. Introduction

According to the World Health Organization [1], dementia is a syndrome in which there is deterioration in memory, thinking, behaviour and the ability to perform everyday activities. There are currently around 50 million people affected with dementia and almost 10 million new cases each year. With a projection of 82 million by 2030 and 152 million by 2050, researchers are continuously trying to find a cure for dementia.

There are pharmacological and non-pharmacological treatments for dementia. For pharmacological treatments, the commonly prescribed medicine to the patients are mostly to regulate the hormones related to learning, but they have short effectiveness and may cause side effects in the long run, example of these are cholinesterase inhibitors and memantine [2]. Non-pharmacological treatments include multiple types of therapies that are used to alleviate the symptoms of dementia. One study explores reminiscence therapy on dementia [3]. It involves discussion of memories and past experiences using photographs and music as mementos for memories. The study claims that the therapy has some positive effects on quality of life, cognition, communication, and mood. Game therapy is another kind of therapy that improves cognitive functions [4] along with social and emotional functions [5].

Virtual Reality is a promising technology for therapy for its real-life scenario simulation without risk. Interventions using VR can be useful for people with mild cognitive impairment or dementia [6]. Content, immersion, and controls must be taken into account for persons with dementia [7]. In this study, VR technology is implemented as a serious game. This meta-analysis on VR-based interventions states that VR is cost-effective, flexible, comprehensive, and potentially useful for patient-centered care. VR therapy has a positive effect on physical fitness, cognition, and emotion. Dementia patients would benefit greater accessibility in their own homes or residential aged care since each experience can be tailored to the interest of the patient [7].

Due to the nature of COVID-19, older adults with dementia are more vulnerable in contracting COVID-19. Hence, local authorities have banned visits to nursing homes and long-term care facilities [8]. Anxiety and depression are the most common psychological effects caused by the pandemic. It also raises the issue of high rates of pre-existing depressive symptoms in the elderly and lack of access to mental health care [9]. Technology-based interventions that can be remotely accessed by people with dementia can address this problem [10]. Due to COVID-19, progress of telemedicine has hastened [11]. This study is a response to the need for more technology-based interventions for people with dementia.

The Montessori Method dates back to the 1900s when Dr. Maria Montessori was looking for a scientific approach to teach children. The philosophy revolves around bringing the person's potential by giving him creative freedom in an environment. Later on, researchers and doctors start to apply the concept to dementia patients where he reminisces his interests or hobby then work the therapy from there [12]. According to [13]van Rijn's team, the Montessori principles applied should include the following: (1) cueing, (2) building on existing skills, (3) providing clear specific tasks related to an activity, and (4) multisensory activities. The Montessori method is a very helpful way to help the patient set up the environment he is most comfortable with as the main goal is to make the patient feel more independent and at the same time boost his sense of purpose, so he will not feel like a dead weight [14]. For this study, we are going to focus on individual engagement rather than group activity due to the pandemic.

2. Objectives and Considerations

The study has four primary objectives: (1) Develop A VR Serious Game with Montessori Approach, (2) Add a real-time system for the application, (3) preserve VR application features for dementia and (4) test for usability, and user feedback with performance metrics and analysis of the application.

The first objective consists of two features: (1) Contain a variety of Montessori-based activities, with a touch of (2) personalization [15]. The rooms are divided into two where each room has its own set of Montessori activities

1. Living Room - Activities are (1) Turning on the TV and (2) the radio. Personalized features are (1) Television show, (2) Radio music, (3) Book, movie, poster covers and (4) Personal photographs.
2. Lanay or Garden - Activities are (1) Flower arrangement and (2) watering plants.

The second objective is to develop a real-time system for the VR application. The patient and caregiver can be in the same virtual room in real-time. This objective addresses the specific constraints and additional conditions brought about by COVID-19.

The third objective consists of the following features: (1) have easy-to-understand game mechanics, (2) have intuitive interface and controls, and (3) provide an ambient environment. The study uses the Serious Game design assessment framework [16] to further expand upon these features.

The fourth objective is to do various tests on the application. Namely, these are System Usability Test, user feedback, and performance metrics and analysis on the application.

3. Application Design

The conceptual framework used for this study is divided into 4 parts corresponding to each of the main objectives. The first part is about developing the game with Montessori approach. To ensure that the game follows the Montessori approach, the activities are designed following the Montessori Principles that [13] used in designing their own game. Rather than introducing new skills, it is important to build on the existing skills of the patients to prevent frustration. Clear and specific tasks must be provided to make sure that the patients understand such as visible cues. Finally, activities must be multi-sensory and dynamic. Adhering to these principles support the Montessori approach of the application.

The second part is supported by [17], stating that there is indeed an existing need for a remote connection between patients and caregivers. COVID-19's restriction on social distancing can be alleviated through the use of telemedicine, therefore a real-time system for the application is a crucial component.

The third part uses the Serious Game design assessment framework (SGDA) provided by [16] to preserve VR application features. There are 6 key design elements: (1) purpose, (2) content & information, (3) game mechanics, (4) fiction & narrative, (5) aesthetics & graphics and (6) framing.

- Purpose - The main purpose of the application is to develop a VR serious game with the Montessori approach, provide a comfortable, relaxing, and entertaining experience, accommodate various hobbies and interests by providing a multitude of Montessori-based activities, and improve overall quality of life of the patient
- Content & Information - The Montessori-based activities that were chosen to be included in the application are from literature and professional opinion. Simple and straightforward instructions are given to the user for each activity.
- Game Mechanics - The application offers different types of activities. To keep the game simple and easy to use, the application only features one game mechanic: point-and-click.
- Fiction & Narrative - The application is a collection of different Montessori-based activities. The user does not need to keep track of a linear story as there is none. This ensures that the user focuses on the gameplay itself.
- Aesthetics & Graphics - The application's interface is intuitive in order to prevent overwhelming the user with information and cause confusion. The application has simple yet recognizable graphics. There is an ambient environment with background music for relaxation. Auditory and visual cues are present to aid the patient.

- Framing - The target audience for the application is people with dementia. The game design is specifically developed for their usage, from the game mechanics to the user interface. The caregiver personalizes the virtual world for the patient.

The final part is testing the application using the Software Usability Scale [18] and performing user interviews. Performance metrics and analysis are also done.

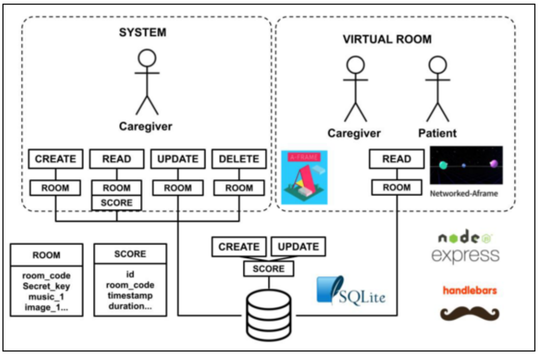


Figure 1. Application Architecture

The application is a web application using NodeJS and Express and Handlebars as its view engine. The database used is SQLite as it is good for small-scale prototypes. There are two tables in the database. The first table is the room table that contains information about a specific room such as patient name. The second table is the points table that contains post-session data for the caregiver to analyze later on.

The application is separated into two parts: The system and the virtual room. In the system, the caregiver can perform CRUD (Create, Read, Update, Delete) on the room table, where uploading of personal media is performed. The caregiver may also view the post-session data of the patient.

Figure 2. GunitaHu's Virtual Room

The next part is the Virtual Room which is developed using A-Frame, a WebVR framework and Networked-Aframe built on top of A-Frame used for the multiplayer aspect. The caregiver and patient are able to interact with each other through voice chat, and may interact with the virtual room. After the session, the system automatically creates a post-session data report of the patient for the caregiver to view later. The pointing system is simple: completing a task gives you a point.

In order to give a nostalgic feeling to the patient, the environment is comprised of a simple home with a living room and a garden.

4. Setup and Implementation Specifics

- Participants - healthy people, preferably those between 50 and 70 years old.
- Equipment - The recommended internet speed is be at least 8mpbs for the videos to load properly with maximum latency of 40ms. The computer on the caretaker side should have an operating system of Windows 10 with a minimum of i3-6100, GPU of NVIDIA GTX 1050 Ti or greater, and 8GB RAM. The Oculus Go is standalone, so there is no need to connect the device to the computer. It has a Qualcomm Snapdragon 835 with GPU of Qualcomm Adreno 540 GPU and 3GB of memory.

The user flow of the patient-caregiver are as follows: Caregiver A is in location A. Caregiver B and Patient are in location B. Caregiver A, using a PC or a HMD connects to the application and enters the virtual room. On the other hand, caregiver B assists Patient such as setting up the HMD, chair, etc. The patient connects to the application thru a HMD. While the session is ongoing, caregiver B monitors the state of Patient and may also view their screen through a PC through screen casting. After the session, the patient exits the virtual world and the caregiver exits and views the post-session data.

5. Testing Plan

The end users of the application are people with mild to moderate dementia. Due to COVID-19, actual testers are healthy participants in the age range of 50 - 70. Clinical testing is outside the scope of this study due to time constraint and the pandemic. As such, the effectiveness of the application is based on usability tests, interviews, not clinical assessments.

The researcher conducts the experiment after setup and profiling the patient then records the patient doing the tasks in any order. The application would give a point for every task done. The session lasts for 30 - 40 minutes with a 5 minute break from the VR experience in between. At the end of the session, the patient answers a questionnaire about usability and user feedback (System Usability Scale)[18]. The source code is publicly available on Github, and the live demo is deployed on Heroku.

6. Conclusion

Exploring this area of Montessori Method integrated in VR is promising. An additional data point such as duration of the session may be helpful for caregivers. This study serves as a case study for future researchers, healthcare professionals, and developers. The COVID-19 pandemic and its effects may persist for a few more years, hence expanding telemedicine is encouraged.

Testing on a larger test sample, as well as on actual dementia patients would be preferable. More scenarios, rooms, and activities would greatly increase the application's features. Enriching some tasks by adding multiple states, such as adding channels on the TV, would also be good. Other game mechanics such as dragging and dropping can be explored by using other VR HMDs with better performance than the Oculus Go. Using a full-fledged game engine like Unity, Unreal, or Oculus SDK allows for more control in the performance side. Healthcare professionals and caregivers can add clinical testing and since the application is open-source, the system may be replicated for other illnesses and ultimately contribute in expanding telemedicine.

References

[1] Dementia. World Health Organization; 2020.

[2] Dementia: What medicines treat it?. WebMD; 2020.

[3] Woods B, O'Philbin L, Farrell EM, Spector AE, Orrell M. Reminiscence therapy for dementia. Cochrane Database of Systematic Reviews. 2018 Mar.

[4] Kyriazis M, Kiourti E. Video games and other online activities may improve health in ageing. Frontiers in Medicine. 2018 Jan;5:8.

[5] Manera V, Petit PD, Derreumaux A, Orvieto I, Romagnoli M, Lyttle G, et al. "Kitchen and cooking", a serious game for mild cognitive impairment and Alzheimer's disease: a pilot study. Frontiers in Aging Neuroscience. 2015 Mar;7.

[6] Kim O, Pang Y, Kim JH. The effectiveness of virtual reality for people with mild cognitive impairment or dementia: a meta-analysis. BMC Psychiatry. 2019 Dec;19(1):219.

[7] D'Cunha N, Nguyen D, Naumovski N, McKune A, Kellett J, Georgousopoulou E, et al. A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment. Gerontology. 2019;65(4):430–440.

[8] Wang H, Li T, Barbarino P, Gauthier S, Brodaty H, Molinuevo JL, et al. Dementia care during COVID-19. The Lancet. 2020 Apr;395(10231):1190–1191.

[9] Rajkumar RP. COVID-19 and mental health: A review of the existing literature. Asian Journal of Psychiatry. 2020 Aug;52:102066.

[10] Cheung G, Peri K. Challenges to dementia care during COVID-19: Innovations in remote delivery of group Cognitive Stimulation Therapy. Aging & Mental Health. 2020 Jul:1–3.

[11] Cuffaro L, Di Lorenzo F, Bonavita S, Tedeschi G, Leocani L, Lavorgna L. Dementia care and COVID-19 pandemic: a necessary digital revolution. Neurological Sciences. 2020 Aug;41(8):1977–1979.

[12] Intriago J. The montessori approach: a breakthrough in dementia care? SeniorsMattercom. 2021 May.

[13] van Rijn H, van Hoof J, Stappers PJ. Designing leisure products for people with dementia: developing "the chitchatters" game. American Journal of Alzheimer's Disease & Other Dementias. 2010 Feb;25(1):74–89.

[14] Judge KS, Camp CJ, Orsulic-Jeras S. Use of Montessori-based activities for clients with dementia in adult day care: Effects on engagement. American Journal of Alzheimer's Disease. 2000 Jan;15(1):42–46.

[15] Laird EA, Ryan A, McCauley C, Bond RB, Mulvenna MD, Curran KJ, et al. Using mobile technology to provide personalized reminiscence for people living with dementia and their carers: appraisal of outcomes from a quasi-experimental study. JMIR Mental Health. 2018 Sep;5(3):e57.

[16] Mitgutsch K, Alvarado N. Purposeful by design?: a serious game design assessment framework. In: Proceedings of the International Conference on the Foundations of Digital Games - FDG '12. Raleigh, North Carolina: ACM Press; 2012. p. 121.

[17] Mantovani E, Zucchella C, Bottiroli S, Federico A, Giugno R, Sandrini G, et al. Telemedicine and virtual reality for cognitive rehabilitation: a roadmap for the covid-19 pandemic. Frontiers in Neurology. 2020 Sep;11:926.

[18] Lewis J, Sauro J. Revisiting the factor structure of the system usability scale. Journal of Usability Studies. 2017 Aug;12:183–192.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210083

Intelligence and Intelligent Simulation

V.Yu. MEYTUS
International Research and Training Center for Information Technologies and Systems of the National Academy of Sciences of Ukraine

Abstract. The work presents a new approach to the study of problems associated with the initial definition of the concept of "intelligence". In work, intelligence is a property that ensures, at some level, the successful interaction of a subject or system with their environment, which is specified in the form of a subject area. For this, the subject area is modeled (smart modeling). Modeling is the process of representing a domain in which the subject must solve the corresponding problems.

The domain model is presented as a body of knowledge covering the constituent elements of this domain. Knowledge is a means of describing this model and is determined using a language that includes the logic of describing the corresponding entities of the subject area. Language properties are used to model some abstract formal structure in the form of a mathematical representation. The modeling process consists in a sequential transition from the description of the subject area to an interconnected structure of knowledge, associated with the constituent elements of this area, which are highlighted in the modeling process.

Keywords. Intelligence, intelligent system, subject area, mathematical structure, logic, modeling, intelligent modeling, knowledge, knowledge representation, ontology.

1. Introduction

There are many science fiction novels and films, the heroes of which are machines created by man, but possessing reason and intellect. And therefore, they are capable of solving the most difficult problems on which the existence of mankind in the future depends, which way will its development go, how it will be able to protect itself from internal and external threats.

The real path to the creation of such machines began when computer systems emerged as the basis for collecting, accumulating, storing and processing data about the world around a person. With the help of these systems, first of all, they made an attempt to solve problems that a person successfully solves, relying on his knowledge and his intelligence, on his feelings and emotions. These were the tasks of recognizing signs, texts, images, competing with a person in various games - checkers, chess, poker, go - understanding what a person is talking about, translating from one language into another, writing poetry and creating musical works. In other words, we are talking about solving problems that are considered creative and have always been the prerogative of a person.

These tasks have historically been called artificial intelligence (AI) [1 - 3]. But, over time, this concept has expanded. The tools that simulated creative functions expanded their capabilities to new areas of application. Now AI is widely used in developments related to entire areas of human activity: industry, robotics, transport, medicine, defense, and education. AI has found applications in finance, trade, security systems, and data mining. Algorithms and methods related to AI are used wherever it is necessary to process large amounts of various data, identify existing dependencies and use them to optimize decisions.

At the same time, a new approach has been outlined for constructing systems that are supposed to have intelligence, and in their direction of research differ from classical AI. The basis of this direction is the definition of intelligence as a certain property, the presence of which in a subject or system means that such a subject or system can be called intellectual. The definition of intelligence simultaneously shows that there are different representations for intelligence, although they all have some common properties.

This work is devoted to the study of the main properties and possibilities associated with this new approach. Its purpose is to highlight and detail the basic elements that collectively define intelligence. In particular, this element is smart modeling, discussed below.

2. Basic concepts

2.1. Intelligence.

Consider the basic concepts related to intelligence. Firstly, any intelligence is related to and manifests itself in a certain subject area. The subject area is a part of the real, imaginary or in some other way a given world, environment, within a certain context..

It is assumed that any subject area can be described by a certain set of concepts, rules and restrictions that are specific to this area. Usually a person works in a certain field and, if you need to characterize him, then say a financier, mathematician, physicist, composer, programmer or athlete, thereby in a sense defining what a person does and where he manifests himself. Although in some cases a person can express himself in different subject areas. For example, the physicist A. Einstein played the violin well, and the politician W. Churchill painted beautiful pictures.

Note that along with the natural material world, considered as a set of subject areas - place of residence, home, study, work, entertainment (theaters, clubs, concerts) - one can consider the worlds of ideas, representations, fantasies, beliefs, hopes, expectations, virtual structures. Subject areas include the fields of science, art, management, and production.

In order to carry out any actions in the subject area, you first need to set or describe these actions for yourself. For this, there is some representation of the subject area, called its model. In this model, actions are formally performed. And intelligence is used in the context of this model. Therefore, the second necessary property of intelligence is the ability and ability to adequately model the subject area, based on their perception of the individual components of this area. The process of modeling itself, since it involves a connection with intelligence, is called intelligent modeling.

Finally, the last third property of intelligence, in which it manifests itself most vividly, is the solution of problems in the subject area. First, the problem is solved in the constructed model, and then, for implementation in reality, it is transferred to the subject area, in which the elements of the model are interpreted. Success in solving problems depends on how adequate the model of the domain was built, and on whether the intellectual subject will be able to find a solution to problems using the constituent elements of the model.

Let's combine these properties in a single concept. Then the ***intellect of the subject is understood as a property that allows the subject to model the subject area, which he perceives and with which he interacts, so that at the level of the constructed model to solve the problems that can be posed to the subject in this area***. [4]

A system with intelligence, that is, capable of simulating its environment and solving problems on the basis of this model, is called an intelligent system. Note that it is possible to consider systems for which a domain model is specified, and the system only solves problems based on this model. For example, this is how the teaching of mathematics at school is structured: textbooks contain information and rules, the use of which allows you to solve problems associated with these rules.

Subject area modeling. Let the subject area $\mathcal{A}$ be given, for which an intelligent system (IS) $\mathfrak{I}$ is being built. This system is created to organize interaction with $\mathcal{A}$. The interaction itself is determined by the tasks coming from $\mathcal{A}$ to system $\mathfrak{I}$, and the solutions of these problems by $\mathfrak{I}$. It is assumed that the system $\mathfrak{I}$ forms for itself some idea of the subject area - the model of area $\mathcal{A}$ - on the basis of information flows coming from $\mathcal{A}$ to $\mathfrak{I}$. These streams are formed using the means of perception that the IC has, or other means used by the IC to characterize the area $\mathcal{A}$.

Model $\mathcal{A}$ is a part of a certain space that includes a partially named set U of objects and structures of the area $\mathcal{A}$, together with the connections and relations existing in this area, perceived by the IS and reflected by it in the model.

The set U is called the universe $\mathcal{A}$. Partial naming implies that individual elements of $\mathcal{A}$ or their constructions from these elements can have their own name. Concepts are associated with them as a description of a set of properties characteristic of this part. Behind the concepts are the classes of objects that exist in $\mathcal{A}$.

Concepts (N), individual representatives of the class of objects (Id) defined by concepts, attributes of concepts (M), relations R in the aggregate make up the universe U, and hereinafter are called concepts of model $\mathcal{A}$. A variant of perception of the area $\mathcal{A}$ is possible not immediately, but gradually. For example, with the help of a training sequence, the elements of which form separate concepts that define the characteristics $\mathcal{A}$. In this case, the IS does not even directly perceive the subject area. All the necessary information is contained in the training sequence.

In addition to the formation of a set of concepts at the stage of modeling, the IS chooses the type of spaces (mathematical structures) in which the model will be built. These can be algebraic structures, topological spaces, manifolds, categories, functional spaces, and neural networks of various shapes, probabilistic models and networks, fuzzy spaces, automata. The features of the selected space are used in the construction of the model, being a constituent part of it. Let's call the selected structure the bearing mathematical structure of the domain model.

The second component of the understanding of intelligence is the logic $\mathcal{L}$, which underlies both the construction of a domain model and the solution of problems that are posed to the IS. In the general case, logic, as a means of formalizing reasoning, shows how, in the condition of certain assumptions, from some statements - the conditions of the problem - it is possible to obtain other statements, considered as a result of solving this problem. But classical logic is not well suited for this. Therefore, in theory of modelling, an attempt was made to expand the possibilities of logical inference, improving it and bringing it closer to the methods with which a person solves problems.

To build IS, one can use, for example, predicate logic, modal and fuzzy logics, evolutionary variants of logic, default logic, probabilistic logic, and descriptive logic. In each case, the system receives its own version of intelligence with its own problem-solving capabilities.

2.2. *Knowledge*.

In order to build a model of a subject area, IS collects information about the structure, constituent elements, connections between the elements of this area, IS collects information about the structure, constituent elements, connections between the elements of this area, the possibility of transforming some elements into others.. Moreover, this information comes in separate portions and is stored in the system, forming in it an idea of the subject area in the form of a body of knowledge. [4, 5]

Each knowledge is generated by an objective reality in which an IP (subject) exists. But this reality is transformed through the perception of the IP (subject). Or it is set ready-made, as a result of the perception of other subjects. The process of knowledge formation can be empirical: the collected information is brought together and transformed into the form of knowledge. Or theoretical, passing first through the creation of a theory that is valid in a certain area, and only then including theoretical hypotheses and conclusions in the form of knowledge about the environment. Although between these forms of obtaining knowledge, in reality, transitions and interconnections are possible and exist.

Based on the collected knowledge, the IS performs modeling. The process of creating a model can be one-time: "{knowledge} $\rightarrow$ model", or sequential and multiple, in which knowledge is gradually built up and a new amount of knowledge allows you to form a new approximation of the domain model.

Naturally, the collected knowledge should contain enough information to serve as the basis for a domain model, and be consistent with the capabilities of IP. It is proposed to consider the multicomponent task of knowledge, which from different sides define the object to which the knowledge is compared. In the future, this knowledge can be used to build a model.

Let's choose an arbitrary object X belonging to the area $\mathcal{A}$. Perceiving this object, the IS forms a certain set of features $\alpha^X = (\alpha_1^X, \dots, \alpha_n^X)$, built on the basis of this perception. Signs can be as simple as estimating the size or color of an object. And they can be complex if the object is associated with some concepts of the subject area, for example, a tree, a garden, a stream, and a ravine. Suppose that the set α^X is the value of some operator Λ on the object X. And this value Λ(X) is one of the components of knowledge ς_X about the object X. Λ(X) is called the natural interpretation of the object X.

The second component of knowledge is built on the basis of logic $\mathcal{L}$. This component consists of a logical description of the object X in the form of a formula in the language of logic $\mathcal{L}$. This formula will be denoted as $\mathcal{L}$(X). In area $\mathcal{A}$, object X is linked to other objects from $\mathcal{A}$ by different relationships and dependencies. It is assumed that the set of such objects associated with X is finite. Such a set of pairs (object Y, relation R), for which object Y is connected with object X by relation R in $\mathcal{A}$, constitutes the third component of knowledge about object X. Finally, the last fourth component of knowledge is the ontological description of object X. The use of ontology is a prerequisite for human understanding of knowledge representation. What cannot be described at the level of ontology is not knowledge for a person either.

Note that, in the general case, ontology can include various means of description for different subject areas: the language of mathematical theory and chemical formulas, mythological representations, metaphors, fantastic images and descriptions. Interpretation of an object from $\mathcal{A}$ in the form of an ontological description is the basis on which the interaction of the IS with a person is further built.

2.3. Intelligent Modeling Algorithm.

The concepts discussed above are combined in a basic algorithm for constructing a domain model. A variant of such an algorithm was proposed in [6].

The intelligent modeling process is structured as follows.

1. The supporting mathematical structure [6], considered as the skeleton of the model, is selected as the basic basis for the created model. The choice of the supporting structure largely determines the methods that can be used in the model to find solutions to problems. The components of this structure in the subject area are compared to entities that the IS (subject) perceives, separates from the area and finds for them a possible set of properties (features, attributes).

2. Entities in the subject area include:

- objects perceived at this level as some kind of integrity;
- instances of objects;
- classes of objects;
- structures built from various objects, classes;
- connections and relationships between objects, classes, structures;
- a set of objects considered as a single entity;
- processes considered as a sequential change in a set of objects.

It is assumed that each of these entities can be perceived and separated from the subject area for modeling. Moreover, some entities that have the same or similar properties can be combined into a common class. Finally, an entity can be formed during the modeling process when other entities are selected, or it can arise based on the choice of an ontology associated with the subject area.

Note that other approaches to considering entities in the subject area are also possible. For example, when considering different variants of neural networks as a carrier structure, sets of weights of individual neuron configurations can be considered as entities.

To accumulate experience of its work, the IS contains a knowledge base in which various variants of subject areas and their components are presented and accumulated in the course of work. The intelligent modeling algorithm assumes the possibility of accessing this database. The request is based on the information that the subject endows with the entities selected from the subject area. At the same time, the possible environment of the entity in question can be selected from the base.

3. After building a domain model, intelligence appears as the basis for solving problems in this model. The supporting mathematical structure associated with the subject area sets the core of a formal approach to solving future problems, to the choice of the necessary methods and an abstract approach to the presentation of information that needs to be used in the solution. In turn, the logic associated with intelligence sets the rules on the basis of which the IS will seek a solution to the tasks assigned to it.

4. Logic is the formal basis for solving future problems. It coincides with the logic that is chosen to build the domain model. The logical description of knowledge is one of the components of its presentation. And logical view processing is an integral part of the problem-solving path. Above, we have considered possible options for specifying logic that can be used to describe knowledge and the subsequent conclusion of consequences when solving a problem. In fact, there are much more possible options for specifying logic. And research in this direction is constantly ongoing. In many respects, the efficiency of the intellect in solving the problem, as well as the complexity of the problems being solved, depends on the task of logic.

5. The intelligent modeling process usually includes several stages. In other words, the domain model is not created immediately as a one-time sequential execution of the first four steps of the overall building process. As practice shows, first, based on the selected set of entities, the first approximation of the model is built. It is then analyzed and viewed as a representation of the domain for comparison and validation of the model and the domain. It is possible that the constructed model reflects only some characteristics of the region and requires clarification of the elements presented in the model. Then the modeling process is reapplied to the pair (subject area, model). The constructed model is expanded and refined; new elements and entities are introduced into it so that it more adequately reflects the subject area.

6. The finished model is immersed in the IP knowledge base. In the future, many of these models determine the experience of an intelligent subject or system, making it possible to simplify the modeling process. Therefore, the process of sequential building of the model is transformed into the process of its training.

7. Along with the process of intelligent modeling and consideration of its result - the model of the subject area - the knowledge base includes information related to the choice of the supporting mathematical structure and the choice of the logic of description and construction of the model. Such a choice can be determined by analogy with ready-made solutions or randomly. But determining the nature of the supporting structure and the logic of solving problems in a specific subject area is one of the complex unsolved problems associated with the development of the theory of intelligent systems.

3. Conclusion.

The work gives a definition of intelligence as a property that provides, at some level, the successful interaction of a subject or system with their environment in the form of a subject area. The system (subject) possessing this property is then called intellectual. The property of intelligence is expressed in the fact that such a system models a subject area, and then uses this model in order to solve problems in the model that determine the interaction “system ↔ area”.

The intelligent modeling algorithm selects for building a model some supporting mathematical structure, which is associated by the system with the perceived structure of the subject area. And to solve problems in the model, it uses one of the possible variants of logic. Naturally, for the same subject area, different versions of intelligences can be built, which have different capabilities, determined by the classes of problems to be solved.

How for a specific area to choose the most suitable type of intelligence corresponding to it remains an open problem, which can be solved only after building different intelligences and comparing them in solving problems.

References

[1] Russell, S.J. Norvig, P. Artificial Intelligence: A Modern Approach (4nd ed.), Prentice Hall, 2020.- 1115 p.
[2] Luger, G., Stubblefield, W. Artificial Intelligence: Structures and Strategies for Complex Problem Solving (5th ed.), Benjamin/Cummings Publishing Company, Inc., 2004.- 720 p.
[3] Hlybovets MM, Oletsky OV Artificial intelligence: a textbook. - K .: "KM Academy", 2002. 366 p.
[4] Meytus V. Problems of building intelligent systems. Knowledge representation. *Cybernetics and Systems Analysis*. 2019, №4. p. 3-14.
[5] Meytus V. Introduction to the theory of intelligent systems. Basic views. Saarbrücken: Palmarium academic publishing, 2015. 189 p.
[6] Meytus V. Problems of building intelligent systems. Intellectual simulation. *Cybernetics and Systems Analysis.* 2021, №4. p.3-19.

This page intentionally left blank

Data Mining and Machine Learning

This page intentionally left blank

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210085

Creating a Metamodel for Predicting Learners' Satisfaction by Utilizing an Educational Information System During COVID-19 Pandemic

Zoe KANETAKI, Constantinos STERGIOU, Georgios BEKAS, Christos TROUSSAS, Cleo SGOUROPOULOU
University of West Attica, Athens, Greece

Abstract. Faced with the disruption generated by the COVID-19 pandemic, the advent of enforced and exclusive online learning presented a challenging opportunity for researchers worldwide, to quickly adapt curricula to this new reality and gather electronic data by tracking students' satisfaction after attending online modules. Many researchers have looked into the subject of student satisfaction to discover if there is a link between personal satisfaction and academic achievement. Using a set of data, filtered out of a statistical analysis applied on an online survey, with 129 variables, this study investigates students' satisfaction prediction in a first-semester Mechanical Engineering CAD module combined with the evaluation and the effectiveness of specific curriculum reforms. A hybrid machine learning model that has been created, initially consists of a Generalized Linear Model (GLAR), based on critical variables that have been filtered out after a correlation analysis. Its fitting errors are utilized as an extra predictor, that is used as an input to an artificial neural network. The model has been trained using as a basis the 70% of the population (consisting of 165 observations) to predict the satisfaction of the remaining 30%. After several trials and gradual improvement, the metamodel's architecture is produced. The trained hybrid model's final form had a coefficient of determination equal to 1 ($R = 1$). This indicates that the data fitting method was successful in linking the independent variables with the dependent variable 100 percent of the time (satisfaction prediction).

Keywords. machine learning, students' satisfaction, CAD, COVID-19, online learning, hybrid model, Engineering education.

1. Introduction

As recognizing the customers' needs is critical to business success, in the educational field, recognizing the learners' needs and taking measures to satisfy them can be the key to enhance students' academic achievements. Many scholars have explored the topic of students' satisfaction in order to discover whether a connection can be established between personal satisfaction and academic performance. It has been described as a comparison of expectations and perceived service quality [1]. However, student satisfaction can be considered as a good indicator of retaining existing students [2], and especially during their first year in higher education. Due to the global COVID-19 pandemic outbreak, traditional teaching methodologies needed to be reformed through online platforms. Technology features and social media channels have been applied in the field of education in order to create challenging virtual learning environments, especially for the first semester students that have not even visited the University Campus due to lockdown measures.

In view of the above, this paper presents a model that predicts the students' satisfaction in a first-semester mechanical engineering course. The model consists of 24 critical variables of an extended online survey. The novelty of this paper lies in the forecasting of students' satisfaction in learning environments controlled fully by pandemic restrictions.

1.1. Related work

It is important to investigate significant factors in online learning that indicate the success of the method applied. Those success factors can be measured in terms of network education platforms corresponding to the needs of instructors and learners, remote teaching on completing learning tasks efficiently and whether online education can become an efficient tool for specific periods. Outcomes can reveal recommendations based on the research findings, in order to support the sustainability of online education strategies [3]. In [4], the online portion of a blended-learning degree program for pharmacists has been evaluated, using a formal self-assessment and peer review. An instrument systematically devised according to Moore's principles of transactional distance [5] has been applied and the research pointed out that a number of course elements for adjustment could enhance the structure, dialog, and autonomy of the student learning experience. In [6] a virtual reality tour-guiding network has been constructed, and 391 students from a Taiwanese technical university took part in the experiment. The findings of this research revealed their learning efficacy and acceptance of technology in the educational system. During the first stage of the outbreak, in most educational establishments online education took the form of class-based instruction and is an expansion of the original online education.

Previous research on the satisfaction of online education platforms did not consider the new variables introduced by the epidemic, such as ease of use and interaction quality [3]. Nevertheless, new constructs categories have been revealed in [7] such as Students transition from Face-to-Face learning to an Emergency remote Teaching environment [8] and virtual classroom fatigue, that can affect learners' attitude towards a specific online curriculum.

The present study evaluates the learning strategy of an online first-year engineering curriculum from the viewpoint of students, in the context of public health emergency. Factors referenced in previous studies [9] have been optimized, and shown to contribute on student's perception of assignments relating to real world tasks.

2. Research methodology

This research has been conducted on students (population) completing their first semester during the academic year 2020-2021 at the University of West Attica, School of Engineering, Department of Mechanical Engineering (Athens, Greece). The curriculum selected is laboratory course, named "Mechanical design, Computer Aided Design CAD I". The online module's learning strategy has been applied to 216 students which were divided in 11 online groups in MS Teams platform, assisted by a group of 5 instructors (N=216), with a valid number of 165 participants (n=165).

A large amount of information regarding engineering students and their interactions with their virtual learning environment has been obtained. Data were mined out of two web-based surveys[1], and additional students' related data. Quizzes, freehand drawings of object views (sketches) and CAD drawings of object views including sectional views were among the weekly assignments. The purpose of the mind-on assignments was to improve students' spatial perception. Specific activities were centred on an existing metallic structure on campus, aiming to establish tasks' relevance to real-world mechanical engineering cases, as well as their importance for students' future employment. [10, 11].

The measuring methods of students' satisfaction of the module can be schematized in the following organogram:

Figure 1. Data mined from different sources.

Following the analysis of the collected data, a matrix with the dimensions 129X165 was created (Figure 1,2), where 165 is the number of students and 129 is the number of variables to be evaluated. A statistical analysis was conducted, which included a correlation analysis (Spearman's rank correlation coefficient) in SPSS v20 to highlight the most significant correlations between all of the ordinal variables. The aforementioned analysis filtered out 24 variables that affect students' evaluation of the module, related to their satisfaction during COVID-19 pandemic circumstances.

The methodology performed is illustrated in Figure 2.

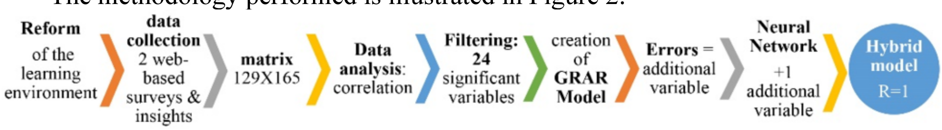


Figure 2. Methodology of the study.

3. Metamodel of interpreting students' satisfaction

The presented method aims to combine the benefits of the GLAR method in fitting transformed predictors with linear logic, with the effectiveness of neural networks in non-linear data fitting applications. Neural networks are considered adequate for handling non-linear applications. The machine learning technique known as an artificial neural network (ANN) is used to define a function that connects a set of inputs to a set of outputs. As a result, a generalized linear model generates a linear combination of transformed predictors (through a link function's response).

A separate file containing the most important variables is established for the aim of building a function to associate the students' satisfaction from the module with the database's most influential variables for this pursuit.

[1] 1. pre-course, 2. post-course

The function created will have the form presented below:

$$y = f(x_1, x_2, x_3, x_4 \ldots x_n)$$

where y represents the students' satisfaction and $x_1, x_2, \ldots, x_n$ represent the predictors (table 1), which in this case are the variables (survey questions) that have a statistically significant correlation with the student's satisfaction.

Artificial neural networks are known to be excellent at handling non-linear issues. Therefore, artificial neural networks (ANN) can be conceived as a machine learning methodology for defining the function (metamodel) that connects a set of inputs to a set of outputs. In the relevant literature [12, 13], ANNs have been used in a number of studies.

The procedure mentioned above is performed by connecting the input layer or a set of hidden layers to an intermediate hidden layer (or a set of hidden layers) that is connected to the output layer. The size of the input layer is relative to the number of variables that are being taken into account. Each layer's nodes (also called neurons) pass information on to the nodes that come after them [14-16].

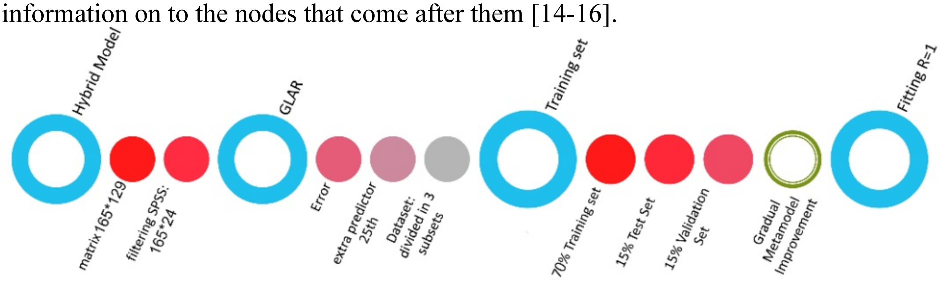


Figure 3. Satisfaction Metamodel neural Network's Architecture.

The suggested method in this study aims to combine the benefits of the GLAR in fitting transformed predictors with linear logic with the efficacy of neural networks in non-linear data fitting applications. The generalized linear model integrates a variety of statistical models, including linear regression, logistic regression, and Poisson regression. It is essentially an iteratively repeatable least squares method that can maximize the likelihood of model parameters.

The survey answers and related data of the 165 students in 24 statistically significant variables (with p-values less than 5% and Spearman rho coefficients equal to absolute 0.20 or more) are isolated in a matrix with the following dimensions: 165x25 (table 1). The students' satisfaction rate is listed in the table's last column (25^{th}). In multivariate regressions, the absence of tools dealing with ordinal and nominal variables is a problem that will be handled by the technique applied in the present study. In the process of modelling continuous variables, similar issues do not occur. The errors generated by the Generalized Linear Model will be used in the hybrid model as the 25^{th} variable, apart from the 24 variables previously stated. A neural network for predicting student satisfaction from the Mechanical Design CAD I module, is finally trained. All simulations of this research have been performed in MATLAB R2020b.

The process can be described in the following scheme:

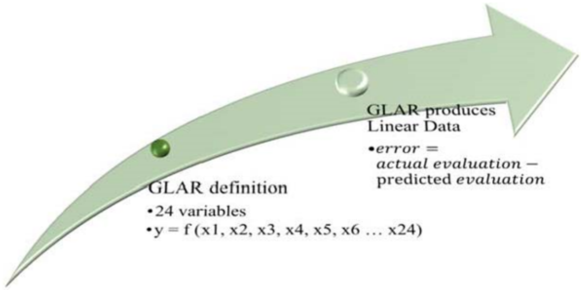
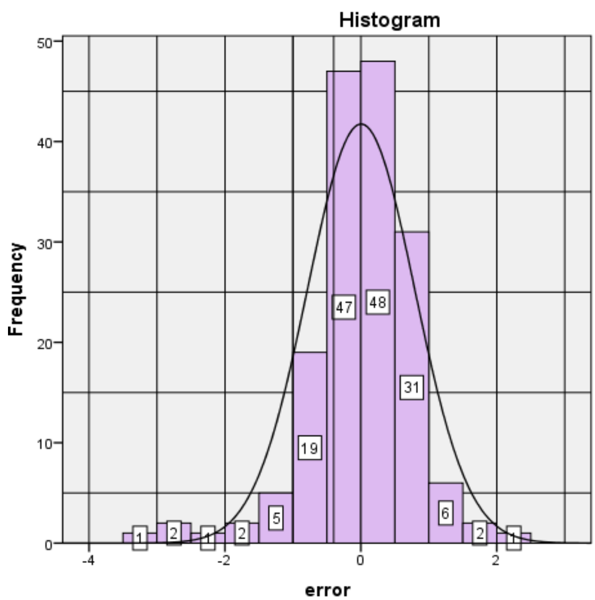


Figure 4. Process scheme.

As a result, the errors originating from the GLAR model are displayed in the histogram (Figure 5).

A normal distribution curve is also included in the histogram in order to compare the generated errors with the ones that would occur if the data were normally distributed (Figure 5). The majority of data are located in the center of the histogram, that represents a range of values between -1 and +1. A general symmetry can be observed with a slight negative left-skewness (-1,024). The left tail of the distribution graph is longer, but since the mean is centrally located on the distribution peak (0,00), it can be concluded that the students' satisfaction prediction variable is normally distributed. As a result, more errors can be found in the region where underestimation of the model's prediction occurs (and not overestimation).

The histogram's bin size was chosen to be equal to 0.5 (Figure 5). It can be seen that the forecasting errors are generally minor.



Statistics

error

N	Valid	165
	Missing	0
Mean		,00
Median		,05
Mode		-0.59[a]
Std. Deviation		,782
Variance		,611
Skewness		-1,024
Std. Error of Skewness		,189
Kurtosis		3,587
Std. Error of Kurtosis		,376
Range		6
Minimum		-3.49
Maximum		2.07

a. Multiple modes exist. The smallest value is shown

Figure 5. Histogram of Errors and mean error.

The table 1 indicates the values of the coefficient estimates from the GLAR model, which are presented with the following form, which correspond to the equation below:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + intercept$$

where: β_0, β_1, β_2 …, are the coefficient estimates from the GLAR model.

Table 1. variables' description, Spearman's RCC and GLAR coefficients.

Variables	Description	Spearman's rank correlation coefficient	coefficients of the GLAR
(Intercept)			0.16
x1	Enjoyable vs other labs	0.527	-0.06
x2	Familiarised to MS Teams vs other modules	0.430	0.06
x3	Satisfied vs other modules	0.673	0.23
x4	Able to complete assignments (difficulties)	0.329	0.00
x5	Satisfied with assignments' grades	0.314	-0.10
x6	Well organised	0.588	0.14
x7	How well tasks are assessed	0.366	-0.32
x8	Theory contributes on accomplishing assignments	0.427	0.14
x9	Evaluate class notes	0.391	0.03
x10	Quizzes contribute on understanding the theory	0.388	-0.03
x11	Evaluate assignments variety (Quizzes, sketches, CAD)	0.451	0.16
x12	Quality of videos in assignments' methodology	0.443	0.06
x13	Enjoyability CAD II lab vs other theoretical modules	0.470	0.28
x14	Classroom fatigue	0.469	0.04
x15	Evaluate interactivity	0.608	0.25
x16	Opportunity to express out loud questions during online lectures	0.418	-0.08
x17	Questions being answered during online lectures	0.485	0.08
x18	Resent late assignment gradings	0.309	0.21
x19	Instructor's comments helped understand mistakes	0.325	-0.05
x20	Assignments related to future work	0.340	0.06
x21	Presentation and clarity of instructions in the 15th assignment	0.376	0.09
x22	Assignments related to Real World Tasks	0.343	0.05
x23	Sustainability of the learning strategy in Face-to-Face learning	0.284	-0.03
x24	Evaluate CAD I vs other modules	0.630	0.31

The majority of errors in 145 instances (out of the total 165 instances) fall into a range between -1 and +1 (Figure 5). Therefore in 87,88% of cases, the errors fall into the aforementioned range.

The following diagrams are obtained: R (coefficient of determination/ prediction success rate; in this case equal to 1), error histogram (with classification of errors per subset (there are three distinct subsets in the present study: training subset (representing 70% of the observations), validation subset (representing 15% of the observations), and test subset (representing 15% of the observations). The overall artificial neural network performance is also displayed in Figure 6.

It's worth noting that the process of breaking down observations into sets (sets) contributes to the neural network's statistical independence, whereas a model derived from a large set of observations (training set) is used to generate predictions in other subgroups of observations.

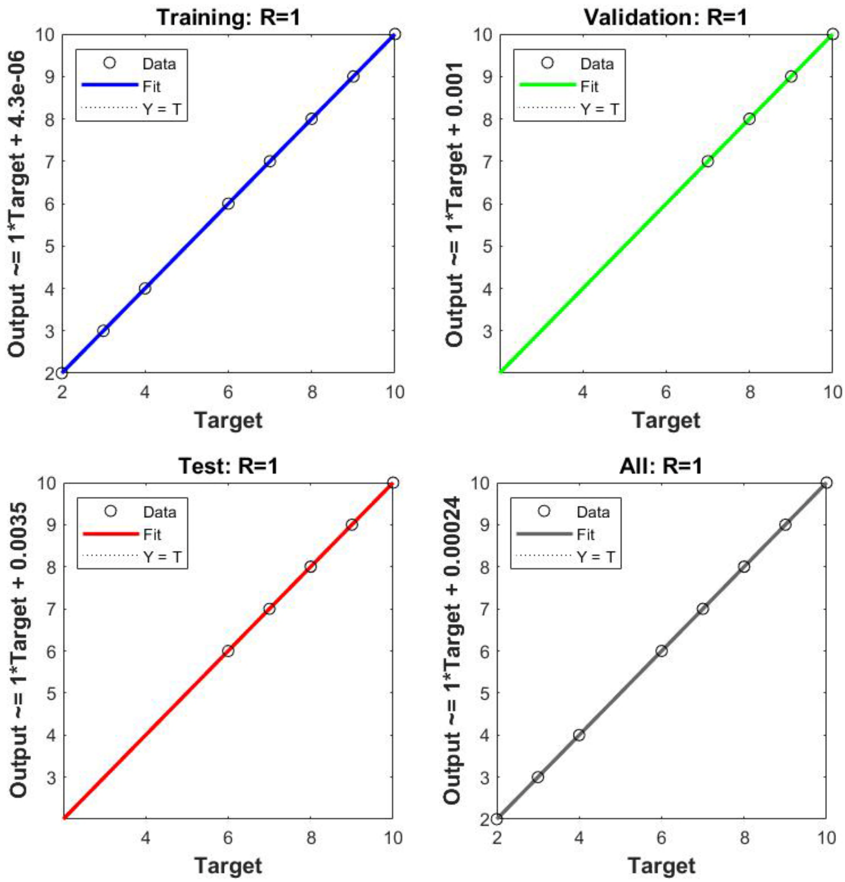


Figure 6. Training set, validation set.

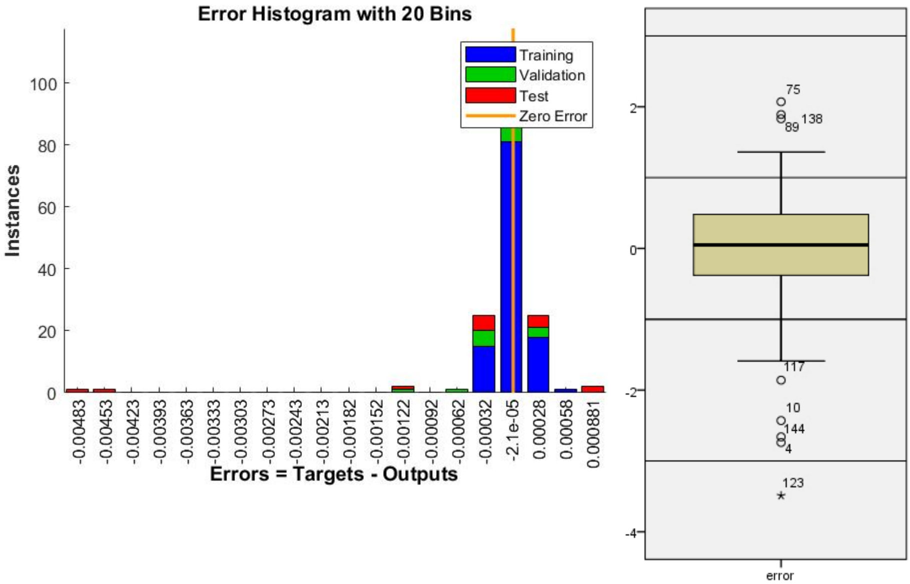


Figure 7. Error Histogram and boxplot.

As shown in the boxplot above (Figure 7), the median is equal to zero. The upper and lower whisker display the position of the first (Q1) and the third (Q3) quartile and predicted values (interquartile range) are mostly gathered between the whiskers, while outliers are displayed above and below the whiskers.

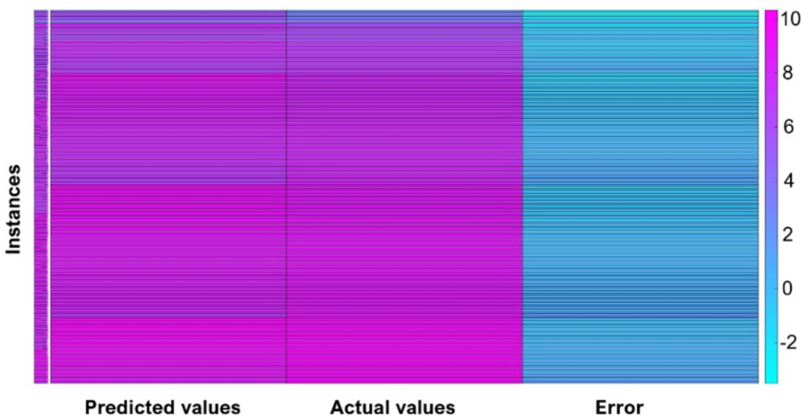


Figure 8. Heatmap of errors

In order to visualize the relation between the actual values and the predicted values a heatmap has been created (Figure 8), where numerical values have been replaced by colors, with three data columns: Predicted values, Actual values and Error. In Figure 8, the color of the cell represents the values. The color gradient in the predicted values is uniform and almost identical to the actual values color gradient. Values in between are shown faded and centered above the average.

4. Conclusions and Future work

It can be concluded that students' satisfaction prediction from a university module attendance during COVID-19 pandemic has not been widely researched until today, whereas satisfaction prediction of online attendance can be a useful component of academic experience [17-21], or even a recommendation tool for selecting specific subjects and university curricula. Students can improve their happiness by their own choices, which represents a great challenge when online learning becomes exclusive and imposed [22, 23]. Recognizing learners' needs and taking steps to meet them might help students achieve greater academic success.

For this purpose, a forecasting tool has been created, by filtering out the critical variables of an extended online survey, applied to students attending a first-semester Mechanical Engineering module, Mechanical Design CAD I, under pandemic circumstances. Therefore, a hybrid model has been created, using 24 critical variables as predictors, focusing on forecasting students' satisfaction. Since there has been no previously similar learning environment controlled by pandemic restrictions, the actual research has been centered on forecasting students' satisfaction.

The final hybrid model was separated into three subsets (Figure 3), with a training set of 70% of the data predicting the test set (15%) and the validation set (15%), resulting in a R=1 fit.

Future work consists of testing the performance of those 24 variables to the academic year 2021-2022 first-semester students, attending the specific module in a similar or blended learning environment in order to determine the overall success of the methodology suggested by this study.

References

[1] R.L. Oliver, A cognitive model of the antecedents and consequences of satisfaction decisions, *Journal of Marketing Research* 17 (1980), 460-469.

[2] I. Čavar, L. Bulian, N. Dubreta, Student Satisfaction in Higher Education: Factors Affecting Engineering Students' Satisfaction, ICERI2019 Proceedings (2019), 6782-6792.

[3] T. Chen, L. Peng, X. Yin, J. Rong, J. Yang, G. Cong. Analysis of User Satisfaction with Online Education Platforms in China during the COVID-19 Pandemic. Healthcare, (2020).

[4] K. Wilbur, Evaluating the online platform of a blended-learning pharmacist continuing education degree program. Med. Educ. Online (2016).

[5] M. Moore, The Theory of Transactional Distance (2013).

[6] H.M. Chiao, Y.L. Chen, W.H. Huang, Examining the usability of an online virtual tour-guiding platform for cultural tourism education. J. Hosp. Leis. Sport Tour. Educ. (2018) 32, 29–38.

[7] Z. Kanetaki, C. Stergiou, G. Bekas and E. Kanetaki, "Machine Learning and Statistical Analysis applied on Mechanical Engineering CAD course: A Case Study During ERTE Pahse in the Context of Higher Education," 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2020, pp. 1-13, doi: 10.1109/ISMSIT50672.2020.9254924.

[8] C. Hodges, S. Moore, B. Lockee, T. Trust, A. Bond, The difference between emergency remote teaching and online learning. Educause Review, (2020).

[9] M. Binkley, O. Erstad, J. Herman, S. Raizen, M. Ripley, M. Miller-Ricci, M. Rumble, Defining Twenty-First Century Skills. (2011)

[10] U. Dominguez, J. Magdaleno, Active Learning in Mechanical Engineering Education in Spain. Conference: WEE 2011At: Lisbon, Portugal (2011).

[11] Z. Tordai, I. Holik, Student's Characteristics as a Basis for Competency Development in Engineering Informatics Education. International Journal of Engineering Pedagogy, (2018) 8: 32-42.

[12] C. Troussas, F. Giannakas, C. Sgouropoulou, I. Voyiatzis, Collaborative activities recommendation based on students' collaborative learning styles using ANN and WSM, *Interactive Learning Environments* (2020). DOI: 10.1080/10494820.2020.1761835.

[13] F. Giannakas, C. Troussas, I. Voyiatzis, C. Sgouropoulou, A deep learning classification framework for early prediction of team-based academic performance, *Applied Soft Computing* **106** (2021), 107355. https://doi.org/10.1016/j.asoc.2021.107355
[14] K.G. Sheela, S.N. Deepa, Review on Methods to Fix Number of Hidden Neurons in Neural Networks. Mathematical Problems in Engineering, Hindawi Publishing Corporation (2013).
[15] P. Yao, Integrating Generalized Linear Auto-Regression and Artificial Neural Networks for Coal Demand Forecasting. Advances in Neural Networks - ISNN 2009 (2009) pp. 993-1001.2009.
[16] M.W. Trosset, An Introduction to Statistical Inference and Its Applications with R; CRC Press, Taylor & Francis Group: Boca Raton, FL, USA (2009).
[17] Z. Mihanović, A. Batinić, J. Pavičić, The link between students' satisfaction with faculty, overall students' satisfaction with student life and student performances. Review of Innovation and Competitiveness. (2016) 2. 37-60. 10.32728/ric.2016.21/3.
[18] I. Duzevic, Student satisfaction vs student achievements – should quality management system in higher education aim at student satisfaction or student achievements? Poslovna izvrsnost - Business excellence. (2020) 14. 51-67. 10.22598/pi-be/2020.14.2.51.2020
[19] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, User acceptance of augmented reality welding simulator in engineering training. *Education and Information Technologies* (2021). https://doi.org/10.1007/s10639-020-10418-7
[20] C. Troussas, A. Krouska, C. Sgouropoulou, Impact of social networking for advancing learners' knowledge in E-learning environments. *Education and Information Technologies* (2021). https://doi.org/10.1007/s10639-021-10483-6
[21] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou Measuring User Experience, Usability and Interactivity of a Personalized Mobile Augmented Reality Training System. *Sensors* 21(11) (2021), 3888. https://doi.org/10.3390/s21113888.
[22] M. Omar, Influence of Wi-Fi Performance on Students' Satisfaction. International Journal of Modern Education. 3. 226-237. 10.35631/IJMOE.380018.2021
[23] Kéri, Anita. Online Teaching Methods and Student Satisfaction during a Pandemic. Language Learning (2021)

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210086

Evaluating the Usefulness of Unsupervised Monitoring in Cultural Heritage Monuments

Charalampos Zafeiropoulos [a] Ioannis N. Tzortzis [a] Ioannis Rallis [a]
Eftychios Protopapadakis [a] Nikolaos Doulamis [a] Anastasios Doulamis [a]
[a] *National Technological University of Athens, {mpampiszafeiropoulos, itzortzis, eftprot }@mail.ntua.gr, irallis@central.ntua.gr, {adoulam, ndoulam}@cs.ntua.gr*

Abstract. In this paper, we scrutinize the effectiveness of various clustering techniques, investigating their applicability in Cultural Heritage monitoring applications. In the context of this paper, we detect the level of decomposition and corrosion on the walls of Saint Nicholas fort in Rhodes utilizing hyperspectral images. A total of 6 different clustering approaches have been evaluated over a set of 14 different orthorectified hyperspectral images. Experimental setup in this study involves K-means, Spectral, Meanshift, DBSCAN, Birch and Optics algorithms. For each of these techniques we evaluate its performance by the use of performance metrics such as Calinski-Harabasz, Davies-Bouldin indexes and Silhouette value. In this approach, we evaluate the outcomes of the clustering methods by comparing them with a set of annotated images which denotes the ground truth regarding the decomposition and/or corrosion area of the original images. The results depict that a few clustering techniques applied on the given dataset succeeded decent accuracy, precision, recall and f1 scores. Eventually, it was observed that the deterioration was detected quite accurately.

Keywords. Cultural Heritage, Hyperspectral data, Clustering, Unsupervised Monitoring

1. Introduction

Cultural Heritage assets (monuments, artefacts and sites) suffer from on-going deterioration through natural disasters, climate change and human negligence or interventions. Monuments are defined as structures created by a person or event and they symbolize a historic period of the corresponding place due to its artistic [1], historical, political, technical or architectural importance [2]. UNESCO considers as a first priority the preservation and valorisation of the tangible/intangible Cultural Heritage and applies innovative techniques for the capturing, digitizing, documenting and preserving prestigious monuments [3], [4].

Early detection of decay and deterioration is essential to preserve monuments. Material degradation leads to the failure of the the buildings components. Non-destructive techniques utilized for detection of monument decay. Most of these techniques come from the scientific fields of computer vision, whose great goal is the extraction of information regarding regions of interest (ROIs) from images or sequences of images [5].

The state-of-the-art for 3D/4D documentation and modelling of complex sites utilizes multiple sensors and technologies (e.g., LIDAR, photogrammetry, in-situ surveying, hyperspectral sensors) in order to define the preservation status of the monument[6]. The efficient use of these tools can give significantly better material detection and object recognition, and thus to identify even the smallest differences in spectral signatures of various objects. This continuous development of new sensors, data capturing methodologies, computer vision algorithms, multi-resolution 3D/4D representations and the improvements of existing ones are contributing significantly to the growth of the interdisciplinary cultural heritage domain.

Hyperspectral images are more suitable than RGB ones since they provide a large amount of information (high-spatial and high spectral resolution), allowing identifying the screened materials based on their chemical composition rather than only their size, shape, and visible colour [7]. Moreover, the recent advancement of sensors technologies has led to the development of hyperspectral imaging sensors with higher spectral and spatial resolution on-board various satellite, aerial, UAV and ground acquisition platforms.

In our study, we exploit image clustering techniques on hyperspectral images to detect and evaluate the corrosion of the stones on cultural heritage assets. In more details, an automated mechanism is proposed for the detection of the ROIs in an unsupervised way. In other words, the evaluation of specific ROIs identifiability, using unsupervised clustering techniques, is being attempted.

2. Related work

A plethora of methods for assessing and detecting the deterioration of stone monuments are available to researchers [8]. Those methods are distinguished into (a) destructive and (b) non-destructive techniques. The main drawback of the destructive approach is that a valuable piece of monument structure is taken [9]. On the other hand, the non-destructive approaches utilize methods that extract features of the examined surface in order to detect cracks, defects in the architectural surface and material degradation.

In [2], the authors proposed a method to identifying the exterior and interior surface flaws. This approach provides elastic features of the architectural structure materials in order to detect the crack and inclusion in the building taking into consideration the affected layer inside the material. In [10], the authors introduced a fuzzy clustering approach for extracting the local variance feature from an image. This method applied to define the transitional features implementing hybrid segmentation. In [11], the authors exploited infrared thermography to diagnose materials decay taking into account different historical periods.This approach is used as a tool in the diagnostic level, for the detection of invisible superficial cracks or/and disparities, as well as the revelation of moisture presence within structures.

In the case of cultural heritage, techniques such as clustering can be applied from archaeological artifacts to the entire archaeological site. In [12] the authors used multispectral images from different geomatic sensors, trying to define different construction materials and the main pathologies of Cultural Heritage elements by combining active and passive sensors recording data in different range. During this study, the unsupervised clustering method K-means is proposed. The results shows that an ideal sensors calibra-

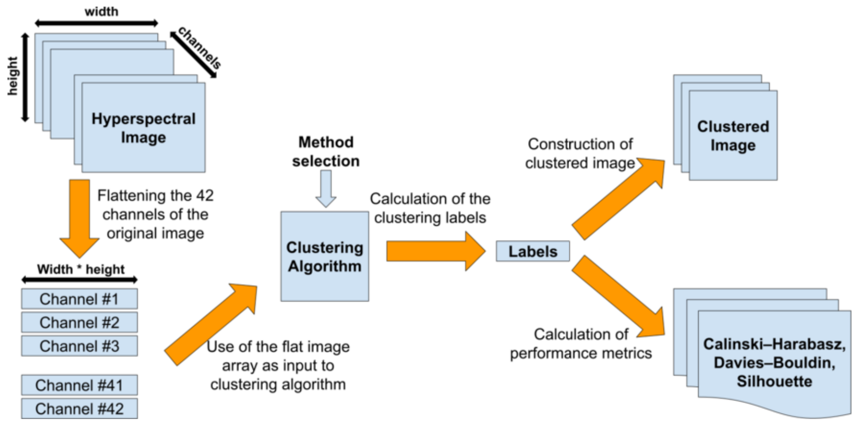


Figure 1. The proposed methodology workflow. Firstly, the original images are being converted to flat image arrays. At a second step, these flat images are used as input to clustering methods. Then, the clustering labels are produced. Finally, the clustered images are being constructed by the clustering labels and the performance metrics are being calculated for each clustering method.

tion can provide more accurate clustering. In [13] the authors used 3D models and data mapping on 3D surfaces in the context of the restoration documentation of Neptune's Fountain in Bologna. In [14] the authors used machine learning classifiers, support vector machines and classification trees for the masonry classification. In [15] the authors developed a correlation pipeline for the integration of semantic, spatial and morphological dimension of a built heritage.

3. Methodology Overview

Cluster based machine learning approaches have been used for the detection of the wall corrosion, regarding the aforementioned historic monument. As described in [16], clustering is an unsupervised learning technique that is being applied to data in order to group them into clusters according to some common characteristics. Several well known clustering algorithms were applied to the images of this study like K-means, Meanshift, Spectral, Birch, DBSCAN and Optics [17].

The data set, used for the purposes of this study, contains hyperspectral images with 42 channels, which represent the several frequencies of the electromagnetic spectrum. The first step of the proposed approach is the conversion of the initial images to flat image arrays, as shown in Figure 1. At this stage, the image content is in the appropriate form defined by the clustering methods. The application of the clustering algorithms bring as result an array with labels that correspond to the produced clusters. Finally, these arrays of labels are used to build the so-called clustered images and to calculate the performance metrics of each clustering method.

3.1. Our contribution

The major outcome of this study is the development of an automated mechanism for the detection of several deterioration types on historical walls. To achieve this, a pipelined approach was followed for the decomposition of the initial images, the clustering fitting procedure, the construction of the clustered images and their comparison with the annotated ones. A pixel based processing was applied to the images, offering a more detailed analysis. Another aspect of this study is the hyperspectral images selection as part of the dataset, since a correlation between the wall deterioration and the additional information from across the electromagnetic spectrum of the image was attempted.

4. The experimental Setup

Our proposed dataset consists of 14 final hyperspectral images of the Fort of Saint Nicholas, with 42 channels for each image. These measurements carried out using the HyperView [18] multi sensor hyperspectral sensing platform by 3D-one. This HyperView system is a dual head system combining one Visual (VIS) snap-shot camera and one Near Infrared (NIR) snap-shot camera, which are connected on a EP-12 board. These cameras acquire only one band per pixel (instead of acquiring all spectral bands for every pixel) while they acquire all the bands in small windows, 4x4 for the VIS head and 5x5 for the NIR head.

Then, the raw images turn into a low resolution hyperspectral image (1/4th or 1/5th of the initial resolution for VIS and the NIR camera respectively), an intermediate hyperspectral image and the final pansharepened hyperspectral image.

(a) The original image (b) The annotated image

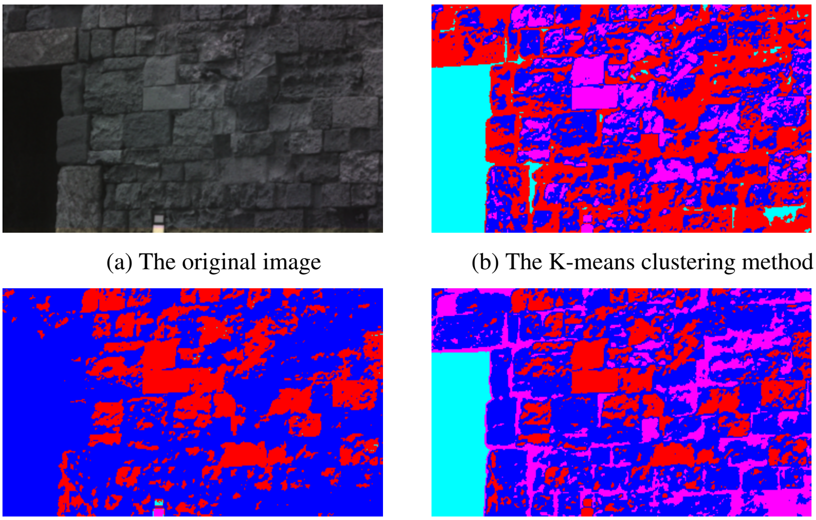
Figure 2. Less than 30% of the wall is affected by the erosion. In the left image a small part of the wall looks brighter from the sunrays so is not considered as damaged.

4.1. The annotation process

First step of the annotation process was the selection of the hyperspectral image channels that correspond to the natural colours. The appropriate channel combination is the following triplet: 15th channel for Red, 6th channel for Green, and 3rd channel for Blue. Next step was the highlighting of the ROIs, using white colour. The rest of the image was coloured black in order for the damaged areas to be more clearly distinctive. In Figure 2, the annotation process that implemented to one of the hyperspectral images, is presented. In the left photo (Figure 2a) the image with the natural colours is shown, while in the right one (Figure 2b) the corresponding annotated image is depicted.

4.2. Clustering areas depiction

The corrosion of the stones could be characterised by the different colour and the roughness of the relevant area surface on the corresponding images of the dataset. Thus, different clustering techniques could offer alternative views of the same data. In Figure 3a, the original image of the monument wall is depicted. Figure 3b presents the outcome of the K-means clustering algorithm, applied to that specific image, while Figures 3c, 3d depict the Birch and the Spectral partitioning accordingly.



(a) The original image

(b) The K-means clustering method

(c) The Birch clustering method

(d) The Spectral clustering method

Figure 3. In Figure (a), the original hyperspectral image is presented, while in Figures (b), (c), (d) the images produced by methods K-means, Birch and Spectral are presented accordingly.

5. Evaluation of Clustering methods

5.1. Clustering algorithms characterization

Essential part of this study is the characterization of the several clustering techniques regarding their performance. Metrics such as Calinski–Harabasz [19], Davies–Bouldin [20] indexes and Silhouette value [21] were calculated for the initial evaluation of the clustering methods.

For the calculation of these metrics, some definitions and assumptions should be provided. Let K denotes the number of clusters $\{C_k\}, k = 0, 1, 2, ..., K$. Let $X = \{x_1, x_2, ..., x_N\}$ be a vector containing N objects, where x_{ij} denotes the jth element of x_i. The grouping of all objects $x_i, i = 1, 2, ..., N$ in K clusters can be defined as follows:

$$w_{ki} = \begin{cases} 1, & iff x_i \in C_k \\ 0, & otherwise. \end{cases} \tag{1}$$

Eq. 1 ensures the uniqueness of the object to cluster association, which is a valid case for both hierarchical and partitioning cluster analysis

The Calinski–Harabasz index (CHI) is described by the Eq. 2:

$$CHI(k) = \frac{T_B/(K-1)}{T_W/(K-1)} \tag{2}$$

$$Where \;\; T_B = \sum_{k=1}^{K} |\bar{C_K}| \|C_K - \bar{x}\|, \qquad T_W = \sum_{k=1}^{K}\sum_{i=1}^{N} w_{ki}\|x_i - \bar{C_K}\|^2$$

According to [21], the maximum CHI value is associated with the optimal partitioning of the given data. By using constant number of clusters for all clustering methods, the most fitting one gives the maximum CHI value. The Davies–Bouldin index (DBI) is an internal evaluation scheme, where the quality of the clustering is being examined according to information extracted directly from the given dataset. The DBI is defined by the Eq. 3:

$$DB(k) = \frac{1}{K}\sum_{k=1}^{K} R_K \tag{3}$$

$$where \;\; R_K = max\left(\frac{S_k + S_j}{d_{kj}}\right), j = 1,2,\ldots,K, j \neq K,$$

$$and \;\; d_{kj} = \|\bar{x_k} - \bar{x_j}\|, \qquad S_K = \frac{1}{\sum_{i=1}^{N} w_{ki}} \sum_{i=1}^{N} w_{ki}\|x_i - \bar{x_k}\|$$

The minimum DBI value is related to the best partitioning solution. Thus, by using constant number of clusters for all clustering methods, the most fitting one gives the minimum DBI value. The silhouette value shows the similarity of an object regarding the cluster it belongs, compared to other clusters. The silhouette value is described by the Eq.4

$$s(x_i) = \frac{b(x_i) - a(x_i)}{max(b(x_i), a(x_i))} \tag{4}$$

where $a(x_i)$ represents the average dissimilarity of the object with all the other data in the same cluster and $b(x_i)$ represents the lowest average dissimilarity of the object to any other cluster. Since, Silhouette value ranges from -1 to 1, a value close to 1 ensures that the object is well matched to its own cluster.

5.2. Ground truth verification

To evaluate the outcome of the clustering methods, a set of annotated images was used that denote the ground truth regarding the corrosion area of the initial images. So, a comparison between the annotated and the clustered image was performed, using accuracy, precision, recall and f1 scores [22]. The detailed procedure is depicted in Figure 4 and is distinguished into (a) conversion of annotated and clustered images to flat image arrays, (b) assignation of colour triplets (RGB) to specific identifiers that represent the clustering labels and the corresponding clustering colour, (c) accuracy, precision, recall and f1-scores calculation and (d) design of the corresponding graphs.

In Figure 5, a simple example case is shown which offers a more descriptive view of the evaluation process. The initial images are being converted to RGB arrays. Each distinct RGB triplet is being assigned to a unique identifier that represents a specific cluster label (1-6) or one of the two distinct areas of the annotated images (0, 10). The two single-dimensional arrays are being adapted to the current clustering label which is under examination. Each identifier with number 10 of the annotated image single-dimensional array is being replaced by the identifier of the current clustering label. Each position of the clustered image single-dimensional array is being set to zero except from these which contain the same identifier of the current clustering label. The final single-dimensional arrays are used for the calculation of accuracy, precision, recall and f1 scores. This procedure is being repeated for each of the clustering labels (1-6). The most matching clustering label was extracted.

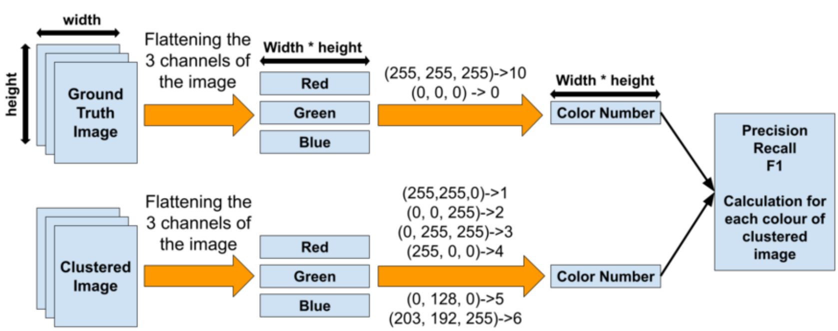


Figure 4. The clustered images evaluation. The annotated and the clustered images are being converted to flat RGB arrays. Each distinct RGB triplet gets associated to a unique identifier, representing the cluster label. The single dimensional arrays with the unique identifiers are combined for the calculation of accuracy, precision, recall and f1 scores

6. Experimental results

As it was mentioned above, the initial evaluation of the clustring techniques was performed using the cluster indexes (Calinski–Harabasz, Davies-Bouldin, Silhouette), which are characterized as internal metrics. According to Davies-Bouldin metric, Meanshift clustering technique presents the best partitioning quality, since its value is the closest to 0. At the same time, Meanshift seems to achieve better similarity among the objects of a common cluster because the Silhouette value is closer to 1 than in any other case (Figure 6).

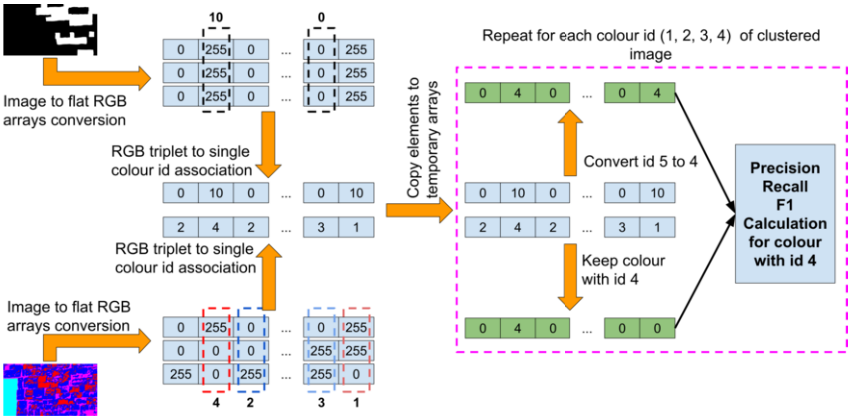


Figure 5. Results evaluation process. The annotated and the clustered images are being converted to flat RGB arrays. The annotated image RGB triplets get associated with identifiers 0 and 10, while the clustered image RGB triplets get associated with identifiers 1-6. These two single dimensional arrays are being adapted to the current clustering label identifier and the process is being repeated for each identifier.

A secondary, more practical approach was used for the evaluation of the clustering methods by comparing the clustered images with the annotated ones and calculating the performance scores accuracy, precision, recall and f1 (Figures 7, 8). From this evaluation, it arises that DBSCAN was the most fitting technique, since it achieved the best scores. Consequently, despite the better internal performance metrics of Meanshift, DBSCAN proved to be the technique that identified more sufficiently the ROIs of the given images.

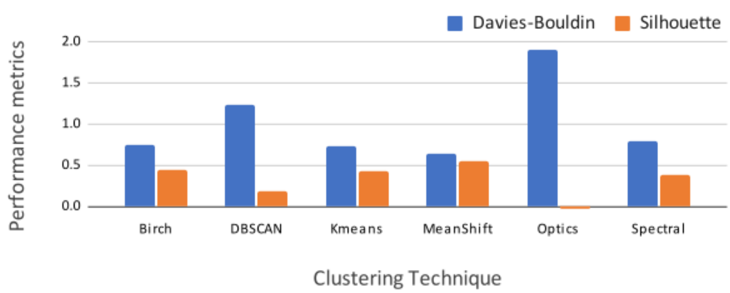


Figure 6. In this graph, the average cluster indexes (Calinski–Harabasz, Davies-Bouldin, Silhouette) for each clustering technique is presented. As shown, Meanshift method achieved the best performance, since its DBI is closest to 0 and Silhouette value is the closest to 1.

7. Conclusion

In our approach, we investigated whether the spectral signatures suffice to distinct various ROIs using trivial unsupervised machine learning techniques. Therefore, we investigate various clustering approaches to identify the feasibility of such methods. According to the results, unsupervised techniques can provide an early, still appropriate mechanism regarding the identification of certain regions of an image (monitoring, defect recognition).

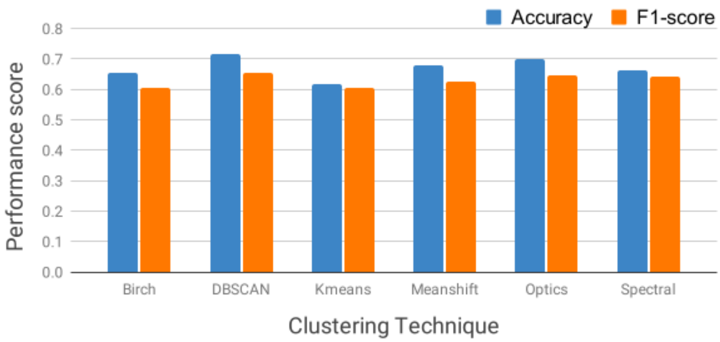


Figure 7. In this graph, the average accuracy and f1 performance scores are presented. As shown, DBSCAN achieved the best results.

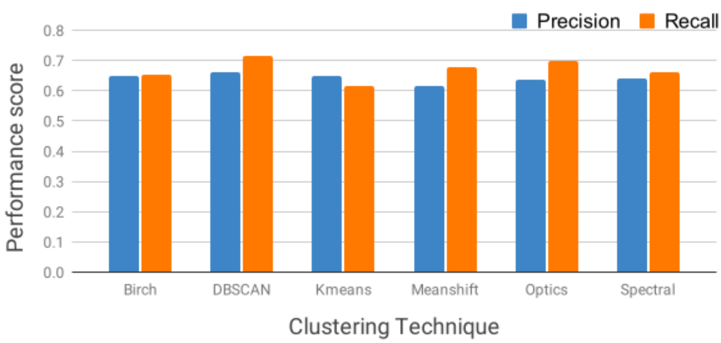


Figure 8. In this graph, the average precision and recall performance scores are presented. As shown, DBSCAN achieved the best results.

Acknowledgement

This paper is supported by the European Union Funded project Hyperion ”Development of a Decision Support System for Improved Resilience & Sustainable Reconstruction of historic areas to cope with Climate Change & Extreme Events based on Novel Sensors and Modelling Tools” under the Horizon 2020 program H2020-EU.3.5.6., grant agreement No 821054.

References

[1] Ioannis Rallis, Ioannis Georgoulas, Nikolaos Doulamis, Athanasios Voulodimos, and Panagiotis Terzopoulos. Extraction of key postures from 3d human motion data for choreography summarization. In *2017 9th international conference on virtual worlds and games for serious applications (VS-Games)*, pages 94–101. IEEE, 2017.

[2] Antonia Moropoulou, Kyriakos C Labropoulos, Ekaterini T Delegou, Maria Karoglou, and Asterios Bakolas. Non-destructive techniques as a tool for the protection of built cultural heritage. *Construction and Building Materials*, 48:1222–1239, 2013.

[3] Henrietta Marrie. The unesco convention for the safeguarding of the intangible cultural heritage and the protection and maintenance of the intangible cultural heritage of indigenous peoples. In *Intangible heritage*, pages 183–206. Routledge, 2008.

[4] E Adamopoulos, E Tsilimantou, V Keramidas, M Apostolopoulou, M Karoglou, S Tapinaki, C Ioannidis, A Georgopoulos, and A Moropoulou. Multi-sensor documentation of metric and qualitative information of historic stone structures. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 4, 2017.

[5] David A Forsyth and Jean Ponce. *Computer vision: a modern approach*. Pearson, 2012.

[6] Evangelos Maltezos, Anastasios Doulamis, Nikolaos Doulamis, and Charalabos Ioannidis. Building extraction from lidar data applying deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 16(1):155–159, 2018.

[7] Konstantinos Makantasis, Konstantinos Karantzalos, Anastasios Doulamis, and Nikolaos Doulamis. Deep supervised learning for hyperspectral data classification through convolutional neural networks. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 4959–4962. IEEE, 2015.

[8] Eleonora Grilli and Fabio Remondino. Classification of 3d digital heritage. *Remote Sensing*, 11(7):847, 2019.

[9] Bernd Fitzner. Damage diagnosis on stone monuments-in situ investigation and laboratory studies. In *Proceedings of the International Symposium of the Conservation of the Bangudae Petroglyph*, volume 7, pages 29–71, 2002.

[10] Priyadarsan Parida and Nilamani Bhoi. Fuzzy clustering based transition region extraction for image segmentation. *Engineering Science and Technology, an International Journal*, 21(4):547–563, 2018.

[11] Antonia Moropoulou, Nicolas P Avdelidis, Maria Karoglou, Ekaterini T Delegou, Emmanouil Alexakis, and Vasileios Keramidas. Multispectral applications of infrared thermography in the diagnosis and protection of built cultural heritage. *Applied Sciences*, 8(2):284, 2018.

[12] Susana Del Pozo, Pablo Rodríguez-Gonzálvez, Luis Sánchez-Aparicio, A. Muñoz-Nieto, David Hernandez, Beatriz Felipe, and Diego González-Aguilera. Multispectral imaging in cultural heritage conservation. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W5:155–162, 08 2017.

[13] Fabrizio Ivan Apollonio, Vilma Basilissi, Marco Callieri, Matteo Dellepiane, Marco Gaiani, Federico Ponchio, Francesca Rizzo, Angelo Raffaele Rubino, Roberto Scopigno, and Giorgio Sobra'. A 3d-centered information system for the documentation of a complex restoration intervention. *Journal of Cultural Heritage*, 29:89–99, 2018.

[14] Noelia Oses, Fadi Dornaika, and Abdelmalik Moujahid. Image-based delineation and classification of built heritage masonry. *Remote Sensing*, 6(3):1863–1889, 2014.

[15] Tommy Messaoudi, Philippe Véron, Gilles Halin, and Livio De Luca. An ontological model for the reality-based 3d annotation of heritage building conservation state. *Journal of Cultural Heritage*, 29:100–112, 2018.

[16] Xiaohui Huang, Yunming Ye, Huifeng Guo, Yi Cai, Haijun Zhang, and Yan Li. Dskmeans: A new kmeans-type approach to discriminative subspace clustering. *Knowledge-Based Systems*, 70:293–300, 2014.

[17] Mieczyslaw Klopotek Slawomir Wierzchon. *Modern Algorithms of Cluster Analysis*. Springer International Publishing, 2018.

[18] Alberto S Garea, Álvaro Ordóñez, Dora B Heras, and Francisco Argüello. Hyperview: an open source desktop application for hyperspectral remote-sensing data processing. *International Journal of Remote Sensing*, 37(23):5533–5550, 2016.

[19] Tadeusz Caliński and Harabasz JA. A dendrite method for cluster analysis. *Communications in Statistics - Theory and Methods*, 3:1–27, 01 1974.

[20] David L. Davies and Donald W. Bouldin. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2):224–227, 1979.

[21] Papakostas George A. Stacked autoencoders for outlier detection in over-the-horizon radar signals. *Computational Intelligence and Neuroscience*, 322, 2017.

[22] Ioannis Rallis, Nikolaos Doulamis, Anastasios Doulamis, Athanasios Voulodimos, and Vassilios Vescoukis. Spatio-temporal summarization of dance choreographies. *Computers & Graphics*, 73:88–101, 2018.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210087

Implementing Hierarchical Indoor Semantic Location Identity Classification: A Case Study for COVID-19 Proximity Tracking in the Philippines

Irvin Kean Paulus PADERES [a], Ligayah Leah FIGUEROA [a] and Rommel FERIA [a]
[a] *University of the Philippines - Diliman, Quezon City, Philippines*

Abstract. Efforts toward COVID-19 proximity tracking in closed environments focus on efficient proximity identification by combining it with indoor localization theory for location activity monitoring and proximity detection. But these are met with concerns based on existing considerations of the localization theory like costly infrastructure, multi-story support, and over-reliance on sensor networks. *Semantic location identities (SLI)*, or location data stored with additional meaningful context, has become a feasible localizing factor especially in locations that have multiple spaces with different usage from each other. There is also a novel method of classification framework, called *hierarchical classification*, that leverages the hierarchical structure of the labels to reduce model complexity. The research aims to provide a solution to proximity analysis and location activity monitoring considering guidelines released in a Philippine context that addresses concerns of indoor localization and handling of geospatial data by implementing a hybrid hierarchical indoor semantic location identity classification that focuses on observable events within context-unique locations.

Keywords. indoor localization, fingerprint localization, machine learning, classification problem, hierarchical classification, semantic location identities

1. Introduction

In the interest of public health due to the worldwide COVID-19 disease outbreak, the easing of community quarantines has been slow and steady for most countries. Governments enforced guidelines maintaining public health and social measures (PHSM) to continue mitigating the spread of the disease. Commonly, these countries have been collecting appropriate surveillance data in order to track public health indicators to allow proper authorities to reassure stakeholders that it is now safe to go back to the physical premises. Location-based solutions (LBS) [4] are one of the most commonly implemented with focus on contact tracing and proximity tracking, with most of the solutions implemented focus on contact tracing. Proximity tracking, on the other hand, is much harder to implement mainly because of the implementation barrier where it is tied to licensed software solutions. In the hopes of easing this barrier, the use of machine learning has been a good way to optimize *received signal strength indicator* (RSSI) from sensor nodes measurements for fingerprint-based localization [14]. The normal use-case of indoor localization is based on the estimation of precise location of the user over an indoor space, but in the context of proximity tracking, there is only a need to track the location usage within the sensing area. Machine learning (ML) classifiers do the job of estimating the user's location based on fingerprint data, but most research tends to ignore a latent aspect of localization, which is location data context. This research aims to introduce the use of *semantic location identities* (SLI) in order to utilize *semantic ontology* [10] and group fingerprint data into meaningful areas such as bathrooms, office rooms, hallways, etc., in order to reduce the localization model's learning complexity. On the other hand, most work on localization systems using ML algorithms [2] only focuses on using a traditional flat classification framework where the predicted labels have the same level with each other. This works for fingerprint-based localization, but as shown by SLIs, the context of any location has a natural hierarchy governing it in the form of buildings, floors, and the user's location. There is another type of framework that has the capability to utilize this hierarchical design of location data called hierarchical classification [12]. This classification framework accepts and utilizes the hierarchical relationships of each output label in order to utilize this inherent design by utilizing different local classifiers for each hierarchical class. Moving forward with this concept, one could leverage the use of cluster analysis on fingerprint data to identify similarities and extract groupings out of it that can be utilized as another hierarchical class. Zhang et al. [17] explored the use of cluster algorithms in providing another layer of hierarchy to the location data. Their work was able to show that introducing partitions that allow overlapping membership was able to improve the localization performance of hierarchical classifiers by allowing better stability and generalization of the input data. With this in mind, exploring other clustering algorithms that allow overlapping memberships can also be explored to improve the classification framework. This research aims to explore the performance of these hierarchically-structured local classifiers, in which each label type node has its own local classifier and the nodes have a parent-child relationship, in providing the final prediction of the system.

In order for proximity tracking to be an attractive COVID-19 mitigation solution to consider, there is a need to improve its performance within indoor spaces. Current solutions do not reliably maintain public health measures unobtrusively within indoor facilities without the use of proprietary software and equipment. The purpose of this research is to leverage the hierarchical structure of location data in improving fingerprint localization by the use of ***hierarchical classifier framework*** and integrate both ***SLI***, identities focusing on observable events within context-unique locations, and area partitions, focused on clustering fingerprint data as part of the SLIs, as integral parts of the hierarchical classification. The goal of this hybrid method is to decrease the complexity of the estimation models and bring improvement to the performance of proximity tracking systems in indoor locations. In this regard, the objectives of the research will be focused on the following:

1. Improve the area partitioning performance of the framework by exploring the use of clustering algorithms that allow overlapping membership,
2. Determine an optimal configuration of local classifiers by exploring different machine learning algorithms for each hierarchical class, i.e floor classification, cluster classification, and SLI classification

3. Leverage the hierarchical structure of location data with the innate cluster capabilities of semantic location identities into a hybrid hierarchical SLI classifier

2. Literature Review

Early work towards the use of hierarchy in order to improve fingerprint localization exists, albeit in different forms. The work of Kim et al. [7] was able to utilize deep neural networks (DNN) in order to integrate hierarchy into localization. They were able to reduce learning complexity and have favorable results although the composition of DNN brings forth the consideration of the added time and technological complexity required in order to train and predict in DNNs. On the other hand, Seçkin and Coşkun [11] shows a different approach to tackling hierarchical improvement to fingerprint localization by iteratively training and predicting over different hierarchies of location data (building, floor, then honing in to the longitudinal and latitudinal position). The study showed favorable improvements in classifying between the parent and sub-parent classes (e.g. building to floor), but its performance on estimating the latitudinal and longitudinal values of the user's location failed to present significant improvements compared to traditional algorithms. The work of Zhang et al. [17] tackled the use of hierarchical classification framework for localization, with the authors exploring area partitioning by grouping fingerprint data with the use of clustering algorithms. They introduced overlapping membership into the cluster algorithm, with them utilizing a modified K-Means Clustering algorithm that allows membership overlap, and found that it doesn't affect the predictive performance of the model. In reality, it improved its generalization capabilities of the framework by having more RSSI samples to match zones and allow more statistical information to be utilized. Though the authors were still able to show the improvement of overlapping clusters, the use of other clustering algorithms that allow overlapping data can still be explored to bring significant improvement to the performance of hierarchical classification.

This research hopes to explore two methods to create a hybrid hierarchical classification framework, one is the exploration of different clustering algorithms that have overlapping membership built into it, and the other focuses on utilizing SLI as the final hierarchical class to be classified. Zhang et al. [17] showed great performance with their improved k-means algorithm, but it failed to have a better performance compared to Fuzzy C-Means (FCM) which natively allows membership overlap. For this research, the authors aim to explore a version of DBSCAN called HDBSCAN [3]. The algorithm has yet to be explored in its performance for localization, but as shown by the study of Malzer and Baum [9], the hierarchical tree structure was able to show data points that has the same membership to a parent tree and it gave a high cluster validation measure of perfect Adjusted Rand Index (ARI) score and zero false noise classification rate in clustering GPS datasets. In this research, the author aims to utilize the innate use of hierarchy of HDBSCAN in zone cluster selection. Another substantial change that will be implemented to hierarchical classification is the use of *semantic location identities* as a classification factor of the final hierarchical level, the position classifier. Based on the study of [16], the fusion of SLI information in a sensing layer can not only provide a great context model engine for LBS applications but it can also provide enrichment of location identification of its system. The work of Abdelnasser et al. [1] showed its effectiveness in improving localization models in an indoor context. The study used landmarks based on sensor fusion and localization techniques to improve semantic localization and mapping (SLAM) model performance. Even if the study used dead-reckoning instead of fingerprinting, the process of capturing historical values to formulate semantic landmarks can be used to improve localization. Van Woensel et al. [15] implemented the use of SLI-improved fingerprinting as a verification model for the estimated position computed by a ML model based on C.50 Decision Trees that uses fingerprint data within the sensed area. The performance of the localization model with semantic location verification was tested compared to the utilization of a ML model alone and results from the evaluation show an increase in accuracy (92.8%) of the semantic location model in a flat classification framework.

For the localization models, the composition of the algorithms we compare as possible classifiers to the hybrid framework came from their previous performances as flat classifiers. In the study of [17], they utilized a simple 1-Nearest Neighbor algorithm, a special case of K-NN that classifies objects using a voting mechanism and selects the most voted class among the 1- nearest neighbors closest to the object. This relatively simple algorithm is appropriate enough for floor and area partition classification, but there are other algorithms that can be utilized that don't rely on explicit computations but have the possibility to perform better. Locality Sensitive Hashing (LSH) [13] is an approximate algorithm, given that it is able to provide a good approximation of the final estimate which can be used to identify neighbor membership. From the findings of Koga et al. [8], LSH's performance shows more than 0.94 cophenetic correlation coefficient (CCC) which means it has a high performance in detecting closely linked clusters while maintaining low complexity, which can be approximated to detect which cluster an input belongs to. The work of Seçkin and Coşkun [11] explored the use of Random Forest algorithm (RF) which is an ensemble of decision trees that has a random selection of the total features of the training dataset and the mode value from all the trees is used as output. In floor identification alone, this algorithm was able to acquire 95% classification accuracy in a building that was reinforced by prior training on a fingerprint classifier which can be utilized for floor and area partition classification, as well as SLI classification due to its reliance on training. The algorithms chosen for SLI classifiers should be more complex than the previous two hierarchical classes as it signifies the final position to be localized. Zhang et al. [17] found significant performance gains in location classification accuracy with the use of Support Vector Machines (SVM). It is an algorithm originally designed for binary classification problems, but a one-versus-one heuristic method can be used that splits the dataset into binary problems with the use of $N(N-1)/2$ binary SVM classifiers and the class from the multi-label setup with the most predictions or votes is selected to be the final output. Zhang et al. showed a 64.97% accuracy of the SVM+KNN hierarchical classifier in position classification, but this experiment was not utilizing SLI as the last hierarchical class and this will be explored in this research. This research will also use Random Forest and C5.0 Decision Trees as discussed previously. The last three models chosen were part of the algorithms discussed in the survey of [18], but performance measurements in terms of real-life localization classification were not disclosed. Extreme learning machines and radial basis function neural networks, which are feedforward neural networks that have the characteristics of close network structure and rapid learning that eliminates the problem of excessive training time. And the last algorithm, Adaptive Boosting (AdaBoost), is another ensemble algorithm that combines

Table 1. Initial Results of the Algorithms to be Explored for the Hybrid Classifier Framework Based on Existing Literature

Algorithm	Performance	Notes
Area Partitioning Algorithms		
Modified K-Means Algorithm with Overlap [17]	Matched actual reference points numbers	Dependent on p threshold value to determine number of clusters
Fuzzy C-Means (FCM) Clustering [17]	Natively allows overlap on clusters due to arbitrary p values,	
HDBSCAN [3]	Good AUC showing promising clustering in terms of data features	Also allow native class cluster overlap
Zone Classifier Algorithms		
1-Nearest Neighbor [17]	2.25 % zone classification rate	Provides fast and easy zone classification
Locality Sensitive Hashing [8]	94 % CCC	CCC measures performance in detecting closely-linked structures, meaning better determination of natural data overlap into other classes
Random Forest [11]	95% accuracy	High performance in general classification
SLI Classifier Algorithms		
Support Vector Machines (SVM) [17]	64.97 % accuracy	Tests here classify for actual position that is why it has lower accuracy compared to other algorithms
Random Forest [11]	95 % accuracy	Classifies for floor position which is more general than position and SLI values
C5.0 Algorithm [15]	92.8 % accuracy	The only test that utilizes SLIs for indoor classification
Extreme Learning Machines (ELM) [18]	<No test results >	These algorithms were recommended by the survey of Zhu et al. [18] based on their performance.
Radial Basis Function Neural Network (RBFNN) [18]	<No test results >	
Adaptive Boosting (AdaBoost) [18]	<No test results >	

several weak classifiers with their outputs put into a voting mechanism in order to provide a simple and fast improvement to the base classification capabilities of the weak model. A brief summary of the algorithms to be explored is found in Table 1. These algorithms were chosen to implement a hybrid localization framework that integrates hierarchical classification and semantic location identity together. The research also aims to achieve localization without provisioning costly infrastructure with the efficient use of sensor nodes.

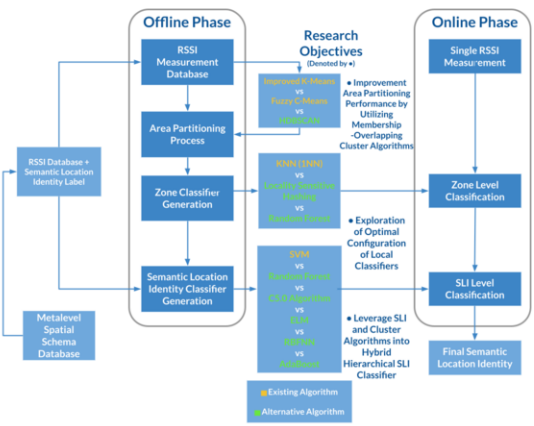


Figure 1. Hybrid Localization Process and Research Workflow Diagram

3. Methodology

Based on the defined objectives in Section 1, the proposed research focuses on providing an improved localization framework that will extend IPS by processing sensed location from various sensor data in a hierarchical classifier that systematically estimates the position within the sensed area from a general class to a specific SLI location. The evaluation method for the objectives will also be discussed in this section.

3.1. Proposed Experimental Design

Figure 1 visualizes the proposed localization framework and the direction of the research as a whole. The main localization process follows the principle of *divide-and-conquer*, separating the focus between classifying the general classes (i.e. floors and area clusters), and classifying the SLI as the current location of the user. Significant focus will also be given to the area clustering algorithm as related literature shows the correlation of the clustering algorithm to the performance of the cluster classifiers. In training the classifiers, a 6040 split will be conducted in order to prevent the machine learning algorithms from overfitting into the fingerprint dataset. In this split, 60% of the dataset will be for training and testing, while the latter 40% is for evaluation. The data belonging to the 60% split will also be subjected to a k-fold cross-validation procedure which will rotate the data between training and testing dataset. This procedure divides the observations into k groups of approximately equal size where the observations on the model will come from the mean of the k group scores and their variance.

3.2. Improvement of the Framework's Area Partitioning Performance

For this objective the research proposes the comparison of previously utilized cluster algorithms to other algorithms that allows membership overlap between area partitions. The final output of this objective will be the final clusters to be used by the proposed framework. One thing to note is that the creation of the area clusters with each of the area clusters is done to the whole fingerprint dataset before the initial 6040 split and k-fold cross validation as the area clusters are required by the proposed framework as a hierarchical class. The evaluation of the area partition classifiers will also act as the evaluation of this objective due to the direct effect of these cluster algorithms to this hierarchical class.

3.3. Determining the Optimal Configuration of Local Classifiers

The total localization performance of the hybrid framework will really depend on the classification models used for each hierarchical class, and because of this a sufficient amount of effort will be conducted in finding the final set of classifiers. The hierarchical classification framework starts with the classification of the floor where the RSSI was captured. In terms of the hierarchical levels, the classification difficulty at this first level is relatively low. The algorithms to be evaluated will be the same for both hierarchical classes. One thing to consider for area partition classification is that the presence of overlap over zones reduces the need for classification models for pre-training. SLI classification, on the other hand, carries the bulk of the complexity of the localization process. Even with the reduced scope due to floor and cluster classification, classifiers in this hierarchical class require a more robust training. In order to verify the performance of the local classifier algorithms without worrying about overfitting to the split 60% data, a cross-validation process will be conducted to get the training and testing data to be used for evaluating the local classifier algorithms. The validation process loops over different "folds" of the dataset with different training and testing subsets for each loop. This process allows the study to measure the real-world performance of the algorithms when classifying data by simulating scenarios where the model classifies data that it is not trained on.

For each loop within the cross-validation process, these measurements are taken:

- Accuracy - the most intuitive performance measure, ratio of correct predictions over the total number of observations
- Precision - ratio of correct positive predictions over total predicted positive observations, which are both the true and false positives
- Recall - checks the sensitivity of the prediction, the ratio of correct positive predictions over the total observations in the class to be predicted, which are the true positive and false negatives
- F1 Score - the weighted average of the Precision and Recall

After all of the loop is finished, the mean of these metrics will be taken as the final value of the classifier configuration. In order to measure the spread between the measurements in order to determine if the performance of the models for each loop varies, the Standard Deviation is taken. It measures the dispersion of the score values from the mean of the scores from all of the loop.

3.4. *Hybrid Hierarchical SLI Classifier*

Hierarchical classification framework follows the well-known *divide-and-conquer* (DAC) principle, with the aim to reduce the input space of different general subspaces in order to improve the framework's generalization capability as well as its learning performance. The use of location data hierarchy will reduce the complexity of the proximity tracing problem by having a smaller number of SLI to consider in classification. This change will still be applicable in the context of proximity tracing problem, by defining the area of the identity as metadata, the system can determine if the number of users to localize the same SLI reaches the approximate population density threshold that will break the enforced social distancing rule of **1 meter** as defined by the guidelines set by Philippine government institutions such as CHE and CSC [6, 5]. To help define these identities and their metadata, a metalevel spatial schema showing the relationship of static and dynamic elements of the sensed area will be formulated. The output framework will utilize fingerprint values based on all RSSI values captured from the Bluetooth sensor network. For evaluating the classification performance of the hierarchy-improved algorithms in each class of the data hierarchy, their performance will be compared to base algorithms in order to show if hierarchical classifiers provide significant improvement against flat classifiers. The comparison of the total running time of the different classifier configuration will be taken in order to check how the training and prediction time of the hierarchical classification framework compares to the flat classification framework. After determining the optimal configuration for the hierarchical classification framework, its performance will be tested on the remaining 40% data to add another verification on the performance of the model in classifying new data. The confusion matrix for each hierarchical class of this configuration will also be used to determine its performance in classifying the class labels for each hierarchy.

4. Results and Discussion

This research is an ongoing study that just passed its approval phase. But as defined in the previous section, most of the framework has been clearly defined based on existing results of other studies related to it. For this section, the work towards the proposed experimental design will be discussed, specifically on the contemporary algorithms that the research will explore in order to define what this study aims to bring into hierarchical classification implementation that differentiates it from existing studies.

4.1. Improvement of the Framework's Area Partitioning Performance

This research aims to explore the performance of HDBSCAN, an extended version of DBSCAN which is a density-based clustering algorithm where it now integrates hierarchy in order to get clusters, in improving the zone classification capabilities of the whole framework. It was chosen based on its ability to define a *distance threshold parameter* that defines the radius size and minimum population of points in order to control membership between clusters, allowing for overlapping membership. HDBSCAN is generally a transductive method where it uses new data points to alter the underlying clustering within its trained dataset. But for all intents and purposes of this research there is not a need to redefine the generated clusters as future RSSI captures are, by design, recaptured data within the localized area. This algorithm will be compared against previously utilized cluster algorithms such as modified K-Means and Fuzzy C-Means. K-means will be implemented here to utilize centroids of the RSSI measurements to assign them into clusters and group them based on a threshold value. Fuzzy C-Means will be implemented in this research as is due to it natively computing the degree of membership of data to a cluster. In previous works, the RSSI measurements are collected by the reference point (RP) where RPs are the one to be computed as to which area partition they belong to. But in order to fit this study, in the purposes of experimentation, there will be a switch from the use of RPs into semantic location identities. This change will still fit the algorithm, as each SLI will be a vector of RSSI measurements at different RP.

4.2. Determining the Optimal Configuration of Local Classifiers

For floor and area partition classifiers, the research introduces the following algorithms, Locality Sensitive Hashing and Random Forest, to be compared to the 1-Nearest Neighbor algorithm. The main idea behind the use of LSH is its low complexity just like 1NN. Given that it is an approximate algorithm, the research aims to leverage the algorithm's composition and see if it will bring significant improvement for classification accuracy, if not, the training time itself. Random forest, on the other hand, was chosen based on its previous performance on floor classification. Given that both floor and area partition classes have more generality, the RF algorithm can be sufficiently used for this use-case.

The algorithms chosen for SLI classifiers have more complexity than the previous ones for floor and area partitions. This study will use Support Vector Machines as a baseline SLI classifier for comparison to the following algorithms to be explored:

- Random Forest - the structure of the Random Forest to be utilized here is similar to the one used for area cluster and floor classification. Given that SLI is the last class to be classified, the study aims to leverage the use of all the fingerprint data features for prediction

- C5.0 Algorithm - C5.0 is an offshoot of the decision tree algorithm that uses less memory and a smaller rule set but it is able to achieve high accuracy. The available features used for this study are compact enough for the C5.0 algorithm to work well
- Extreme Learning Machines - an algorithm based on feedforward neural networks, the difference is that the hidden nodes are randomly assigned and it doesn't need to be tuned. The great thing behind this algorithm is its ability to be trained in a single step, essentially amounting to learning a linear model
- Radial Basis Function NN - also based on neural networks, the difference being that it uses radial basis function as the activation function. RBF allows neural networks to utilize the rate of closeness of a data point to its label for prediction
- AdaBoost - for this research, the AdaBoost algorithm will use Random Forest as the weak classifier. This study aims to utilize the performance of RF in an ensemble form to improve its performance

5. Conclusion

Based on Table 1, initial results from related literature shows the feasibility of these changes in providing improvements in the hybrid classification framework. The main purpose of this framework is to use it for a proximity tracking solution to provide localization capabilities to a system that is not too complicated but can bring comparative performance with lower sensor topology. The research also aims to show the feasibility of this mitigation solution as a response to COVID-19. Prior solutions show as well how LBS can utilize proximity tracking in maintaining health guidelines for COVID-19. This improvement is aimed for educational institutions and office agencies to provide reliable solutions for maintaining public health and social measures and help provide solutions should the need to go back to commercial and educational spaces arise. Most of the limitations of the study were due to the current quarantine situation affecting how the research will conduct the experiments. This study will not delve in exploring the performance of the localization model against device heterogeneity due to lack of sufficient sample devices. Also, this study will also not explore security concerns of crowdedness detection systems and the possible vulnerabilities that sensor-dependent systems. Future work for this study is the main output, a novel hierarchical localization framework revolving around comparison between the use of relative positioning with SLI and estimation of exact coordinates in a given space and the assembly of the hierarchical classification system based on the fusion of the optimal local learning models acquired through experimentation.

Acknowledgements

This research was funded by the University of the Philippines Diliman through the Office of the Vice Chancellor for Research and Development.

References

[1] Heba Abdelnasser, Reham Mohamed, Ahmed Elgohary, Moustafa Farid Alzantot, He Wang, Souvik Sen, Romit Roy Choudhury, and Moustafa Youssef. Semanticslam: Using environment landmarks for unsupervised indoor localization. *IEEE Transactions on Mobile Computing*, 15(7):1770–1782, 2015.

[2] Osamah Ali Abdullah and Ikhlas Abdel-Qader. Machine learning algorithm for wireless indoor localization. In *Machine Learning-Advanced Techniques and Emerging Applications*. IntechOpen, 2018.

[3] Ricardo JGB Campello, Davoud Moulavi, Arthur Zimek, and Jörg Sander. Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 10(1):1–51, 2015.

[4] Mohammad Jabed Morshed Chowdhury, Md Sadek Ferdous, Kamanashis Biswas, Niaz Chowdhury, and Vallipuram Muthukkumarasamy. Covid-19 contact tracing: Challenges and future directions. 2020.

[5] CSC. Revised interim guidelines for alternative work arrangements and support mechanisms for workers in the government during the period of state of national emergency due to covid-19 pandemic. `http://www.csc.gov.ph/phocadownload/MC2020/MC\%20No.\%2010,\%20s.\%202020.pdf`, 2020. Accessed: 2020-11-06.

[6] DepEd. Guidelines on the required health standards in basic education offices and schools. `https://www.deped.gov.ph/wp-content/uploads/2020/06/DO_s2020_014.pdf`, 2020. Accessed: 2020-11-06.

[7] Kyeong Soo Kim, Sanghyuk Lee, and Kaizhu Huang. A scalable deep neural network architecture for multi-building and multi-floor indoor localization based on wi-fi fingerprinting. *Big Data Analytics*, 3(1): 1–17, 2018.

[8] Hisashi Koga, Tetsuo Ishibashi, and Toshinori Watanabe. Fast hierarchical clustering algorithm using locality-sensitive hashing. In *International Conference on Discovery Science*, pages 114–128. Springer, 2004.

[9] Claudia Malzer and Marcus Baum. A hybrid approach to hierarchical density-based cluster selection. In *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 223–228. IEEE, 2020.

[10] R. Rocha. Middleware for location-based services. 2004.

[11] Ahmet Çağdaş Seçkin and Aysun Coşkun. Hierarchical fusion of machine learning algorithms in indoor positioning and localization. *Applied Sciences*, 9(18):3665, 2019.

[12] Carlos Silla and Alex Freitas. A survey of hierarchical classification across different application domains. *Data Mining and Knowledge Discovery*, 22:31–72, 01 2011. doi: 10.1007/s10618-010-0175-9.

[13] Malcolm Slaney and Michael Casey. Locality-sensitive hashing for finding nearest neighbors [lecture notes]. *IEEE Signal processing magazine*, 25(2):128–131, 2008.

[14] Santosh Subedi and Jae-Young Pyun. Practical fingerprinting localization for indoor positioning system by using beacons. *Journal of Sensors*, 2017, 2017.

[15] William Van Woensel, Patrice C Roy, Syed Sibte Raza Abidi, and Samina Raza Abidi. Indoor location identification of patients for directing virtual care: An ai approach using machine learning and knowledge-based methods. *Artificial Intelligence in Medicine*, 108:101931, 2020.

[16] Xin Wang, Jianga Shang, Fangwen Yu, and Jinjin Yan. Indoor semantic location models for location-based services. *Int. J. Smart Home*, 7(4):127–136, 2013.

[17] Chenbin Zhang, Ningning Qin, Yanbo Xue, and Le Yang. Received signal strength-based indoor localization using hierarchical classification. *Sensors*, 20(4):1067, 2020.

[18] Xiaoqiang Zhu, Wenyu Qu, Tie Qiu, Laiping Zhao, Mohammed Atiquzzaman, and Dapeng Oliver Wu. Indoor intelligent fingerprint-based localization: Principles, approaches and challenges. *IEEE Communications Surveys & Tutorials*, 2020.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210088

Data Mining for Improving Online Higher Education Amidst COVID-19 Pandemic: A Case Study in the Assessment of Engineering Students

Zoe KANETAKI, Constantinos STERGIOU, Georgios BEKAS, Christos TROUSSAS, Cleo SGOUROPOULOU
University of West Attica, Athens, Greece

Abstract. Instructional materials, internet accessibility, student involvement and communication have always been integral characteristics of e-learning. During the transition from face-to-face to COVID-19 new online learning environments, the lectures and laboratories at universities have taken place either synchronously (using platforms, like MS Teams) or asynchronously (using platforms, like Moodle). In this study, a case study of a Greek university on the online assessment of learners is presented. As a testbed of this research, MS Teams was employed and tested as being a Learning Management System for evaluating a single platform use in order to avoid disruption of the educational procedure with concurrent LMS operations during the pandemic. A statistical analysis including a correlation analysis and a reliability analysis has been used to mine and filter data from online questionnaires. 37 variables were found to have a significant impact on the testing of tasks' assignment into a single platform that was used at the same time for synchronous lectures. The calculation of Cronbach's Alpha coefficient indicated that 89% of the survey questions have been found to be internally consistent and reliable variables and sampling adequacy measure (Bartlett's test) was determined to be good at 0.816. Two clusters of students have been differentiated based on the parameters of their diligence, communication abilities and level of knowledge embedding. A hierarchical cluster analysis has been performed extracting a dendrogram indicating 2 large clusters in the upper branch, three clusters in the lower branch and an ensuing lower branch containing five clusters.

Keywords. CAD, Data analysis, Data mining, Online Learning, Engineering Education, COVID-19.

1. Introduction

In the absence of face-to-face classrooms, online learning is based on the use of information technology platforms to ensure that all students have access to learning resources. Asynchronous or synchronous communication is used, with asynchronous system tools such as e-mail, discussion boards, and learning newsgroups being used [1]. Because this epidemic was unexpected, many educational institutions were unprepared for the events that followed [2]. It is necessary to arrive into vital features of internet learning that indicate the method's effectiveness. These characteristics can be interpreted in terms of network education platforms that meet the needs of instructors and learners, remote instruction on effectively completing learning activities, and whether online education can become a useful tool for time periods where online learning is performed in a fully remote environment [3]. On the other hand, engineering needs have spawned novel learning methodologies at all levels of the educational system. Advanced technology tools, such as digital platforms and social media channels, are brought into the instructional process and serve as the primary means of information transmission [4-9]. Previous research found that a number of course features for adjustment could improve the structure, dialog, and autonomy of the student learning experience, using functionalities carefully created according to Moore's principles of transactional distance [10]. New factors have been revealed to affect online learning in [11] Emergency Remote Teaching Environments [12].

This study was inspired by the expanding use of online-based education in universities around the world, which provided a unique opportunity to collect electronic data and track students' academic progress.

The present study's novelty derives from the fact that pandemic restrictions were just recently imposed on educational institutions, and fully virtual learning environments have yet to be tested.

In the setting of a public health emergency, the current study assesses the learning approach of an online first-year engineering course from the perspective of students, evaluating tasks' assignment via a single learning platform. A statistical analysis has been performed, using data minded from online surveys, that revealed new factors having a significant correlation to the way students perceive tasks assignments during pandemic.

2. Module content and description

The present study has been performed during the first semester of the academic year 2020-2021 in the University of West Attica, School of Engineering, Department of Mechanical Engineering. The related module is called Mechanical Design CAD[1] I. It is a first-semester module and the objectives after attending the course where to be able to create 2-Dimensional sketches and CAD drawings of objects presented in 3-Dimensional views by applying the rules of Mechanical drawing (tracing and reading orthographic projections, views layout, sectional views, dimensioning, symbolic lines). The laboratory lectures had a 3-month duration, and had been performed exclusively online, in MS Teams platform due to pandemic restrictions.

2.1. Assigning tasks

All tasks were assigned in each lecture, including Quizzes, Practice tasks (non-existent 3-Dimensional objects) that had to be represent in 2-Dimensional drawings, as well as assignments related to real world tasks in Mechanical Engineering [13-15] (figure 1).

[1] Computer Aided Design

Figure 1. Assigning tasks

3. Data mining and filtering

Data have been collected by two online surveys in MS Forms, one pre-course[2] and one post-course[3]. Additional data have been mined out of MS Teams insights, and students' registration information. The total number of participants that attended CAD I during the semester where 216 (N=216) out of which 190 (subpopulation) participated in the study. 165 scores have been considered valid for the purpose of this study, representing first semester students (n=165). 87% of the participants where between 18-23 years old, 11% between 22-25 and the remaining 2% over 25 years old. 88% where male and 12% female.

A matrix of 129*165 has been created including several constructs [11] of variables that needed to be tested during research.

3.1. Correlation Analysis and discussion

A correlation analysis has been performed in SPSS 2020, among the 37 variables (survey questions) of the construct "assignments". (Table 3) The most significant relationships among all nominal variables are identified using a correlation analysis. The statistical significance is determined by identifying variables with a Spearman's rho coefficient greater than 0.20, indicating a relatively significant connection between two variables. In the following passage, correlations from the Inter-Item Correlation Matrix with a rho coefficient >0.400 are presented, and the conclusions have been schematized in the organogram shown in Figure 2:

The first variable, (Students find the CAD I module more enjoyable than other laboratory modules) is correlated with the synchronous platform familiarization (+0.432), the quality of videos on the supporting YouTube channel (+0.421). Students who find the module enjoyable have expressed lower classroom fatigue (+0.475) and consider that they are very likely to succeed in similar future tasks (+0.518). The enjoyability of the module's attendance has been proven to be highly correlated with their overall evaluation of the learning methodology, on a 10 Likert scale (+0.518). Students that highly evaluated the module find it very well organized (+0.532). Online module's task organization is also correlated with the evaluation of the class notes (+0.448). Tasks' assignment is an indicator of organization and is proven to be correlated with the 12th

[2] https://forms.office.com/Pages/ResponsePage.aspx?id=7kOJDHDDs0u6UTIfQG8y7KI65q2QN2ZEuqSfk2d0hHFUNjNBMEJCS1ZFVUg4WVNTOTEwVk4yWE1TSy4u

[3] https://forms.office.com/Pages/ResponsePage.aspx?id=7kOJDHDDs0u6UTIfQG8y7KI65q2QN2ZEuqSfk2d0hHFUMlEySFBTVVRNOVVKRERTQ0UwTzYxRVM5Ry4u

variable, of how well assigned tasks are assessed (+0.416). Assignments variety (14^{th}) has proven to contribute on understanding the theoretical part (13^{th}) and facilitated accomplishing the assignments (+0.431) and is also correlated with the quality of videos (20^{th}) presenting the methodology of tasks (+0.426). The 20^{th} variable can be considered as an enjoyability factor (+0.421) as well as an indicator of organization (+0.444) also correlated with the variety of tasks in the 14^{th} variable (+0.426). Highly significant correlation can be noticed when testing the 15^{th} assignment, which is dealing with an existing metallic structure located in the Ancient Olive Campus (of the University of West Attica) and has been shown to be relevant to future professional tasks that students consider very likely to encounter in their careers (+0.513). The variable testing the conception of 3-Dimensional planes, is correlated with the likelihood of succeeding in similar future tasks (+0.474) as well as with the overall evaluation of the module (+0.466).

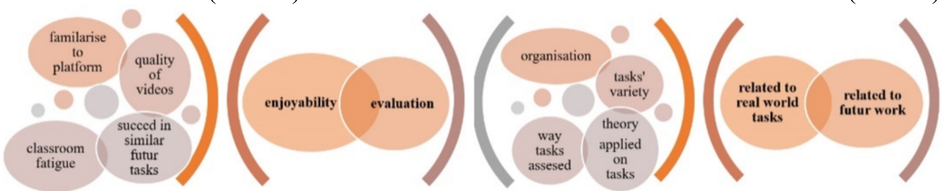


Figure 2. Organogram of correlations

3.2. Reliability analysis and discussion

A reliability analysis has been conducted and at the same time Cronbach's alpha has been calculated, as an estimate of the internal consistency associated with the values derived from the analysis. Cronbach's Alpha coefficient is indicative of the validity of the ordinal variables filtered out of the correlation analysis.

Table 1. Case Processing Summary

		N	%
Cases	Valid	164	99.4
	Excluded[a]	1	0.6
	Total	165	100.0

a. Listwise deletion based on all variables in the procedure.

Table 2. Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.891	0.894	37

As seen in Table 2, it can be estimated that 89% of the variances (37 survey questions) can be considered that concern internally reliable variables, therefore indicating a good internal consistency. The number of students participating in the study accounts to 165 (Table 1) and the valid answers are 164 (N=164). It can be noticed that one student has been excluded from the calculation, since missing values have been detected. In the column that contains Cronbach's Alpha, if the item of the 15^{th} question would be deleted, a coefficient of 0.894 would derive which combined with the result of 0.528 of the corrected item total correlation shows that if this variable would be entirely excluded, the overall coefficient (Table 2) will drop down by 0.004.

Table 3. Item-Total Statistics

	Variables	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
1	Enjoyable vs other Labs	292.45	0.577	0.886
2	CAD I helped familiarising to MS Teams VS other modules	292.59	0.484	0.887
3	Satisfied CAD I vs other modules	275.48	0.646	0.883
4	Insecure about following	293.20	0.392	0.889
5	Are you comfortable passing the finals	298.15	0.267	0.891
6	Affected by technical issues	296.62	0.320	0.890
7	Data consumption from downloading You Tube support channel videos	288.99	0.413	0.888
8	Assignments' load	299.96	0.279	0.890
9	Hours of studying during the week	295.33	0.343	0.889
10	Are you satisfied with your assignment grades	304.09	0.177	0.891
11	Is the module well organised	298.14	0.349	0.889
12	How well tasks are assessed during online lectures	280.07	0.554	0.885
13	Did the theory contribute on the assignments	300.46	0.452	0.889
14	Assignments variety (quizzes, sketches, CAD drawings)	293.02	0.487	0.887
15	Class notes, download and evaluate	288.15	0.367	0.890
16	Did quizzes help on understanding the theoretical part	295.50	0.484	0.888
17	Have you conceived the meaning of planes	293.97	0.573	0.887
18	cutting planes highlighted in 3d views	297.00	0.383	0.889
19	Have you fully perceived the object	300.67	0.319	0.890
20	Quality of videos concerning the solving methodology	302.54	0.234	0.891
21	Clearness of video, image and sound	284.40	0.508	0.886
22	Social Media applications skills	300.44	0.346	0.889
23	Classroom fatigue	296.36	0.275	0.891
24	Express out loud your questions	281.03	0.572	0.885
25	Questions being solved during synchronous lectures	295.49	0.494	0.888
26	Resent when instructor does not return graded tasks on time	291.37	0.574	0.886
27	Comments by the instruction help understand mistakes	294.57	0.365	0.889
28	Instructor helped meeting new people during synchronous lectures	293.87	0.337	0.890
29	Knowledge Weaknesses	297.81	0.291	0.890
30	Computer skills	298.81	0.308	0.890
31	Tasks assigned relevant to future work	305.38	0.126	0.892
32	Presentation and clarity of the 15th assignment	282.90	0.387	0.891
33	15th assignment related to real world tasks	292.65	0.466	0.887
34	All assignments related to real world tasks	298.70	0.293	0.890
35	Overall evaluation of the module	296.50	0.411	0.888

36	Dealing with weaknesses and knowledge lacks	293.77	0.496	0.887
37	Is it likely to succeed in a similar future task	271.75	0.738	0.881

3.3. Sample adequacy

The Kaiser-Meyer-Olkin sampling adequacy measure and the Bartlett's test were used, and the Kaiser-Meyer-Olkin measure was determined to be good at 0.816, while the Bartlett's test of sphericity indicated a significance level of 0.000 (Table 4), which is the highest significance level that can be obtained.

Table 4. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.816
Bartlett's Test of Sphericity	Approx. Chi-Square	2166.018
	df	666
	Sig.	0.00

3.4. Cluster Analysis and Anova analysis performed on the clusters

The k-means algorithm is used to perform a cluster analysis, which divides the sample into two groups of students (Table 5). Cluster 1 seems to enjoy more and feel more satisfied from the online module than cluster 2. Students belonging to cluster 1, feel less insecure and feel more comfortable about the final exam, are less affected by technical issues and do not resent data consumption from downloading the support You Tube channel videos. Cluster 1 includes more hard-working students who do not feel a heavy load on the tasks assigned, and find the module more organized than cluster 2. A two-score difference is noticed in the way tasks are assessed, with cluster 1 evaluating higher than cluster 2, expressing questions out loud and likely to succeed in similar future tasks.

From the cluster analysis, it can be deducted that dividing the participating students in two clusters, has been based on the diligence of students, their communication skills as well as their level of knowledge embedding, permitting them to believe that they can succeed in similar future tasks.

Table 5 shows the most important criteria with Sig.0.05 for the aforementioned clustering as well as an Anova analysis. Because the Anova analysis evaluates differences between individual clusters and within individual clusters (between groups/within groups comparison), the conclusion achieved is that out of 37 variables, only two variables are considered less distinctive, students' satisfaction with their assignment grades and tasks related to future work.

Table 5. Cluster analysis and Anova applied on the clusters

		Cluster		**Anova**
	Variables	**1**	**2**	**Sig.**
1	Enjoyable vs other Labs	5	4	0.00
2	CAD I helped familiarising to MS Teams VS other modules	4	4	0.00
3	Satisfied CAD I vs other modules	9	7	0.00
4	Insecure about following	4	3	0.00
5	Are you comfortable passing the finals	4	3	0.02
6	Affected by technical issues	4	3	0.00

7	Data consumption from downloading You Tube support channel videos	5	3	0.00
8	Assignments' load	4	3	0.00
9	Hours of studying during the week	6	5	0.00
10	Are you satisfied with your assignment grades	4	4	0.28
11	Is the module well organised	4	3	0.00
12	How well tasks are assessed during online lectures	9	7	0.00
13	Did the theory contribute on the assignments	4	4	0.00
14	Assignments variety (quizzes, sketches, CAD drawings)	4	3	0.00
15	Class notes, download and evaluate	3	2	0.00
16	Did quizzes help on understanding the theoretical part	4	4	0.00
17	Have you conceived the meaning of planes	4	4	0.00
18	cutting planes highlighted in 3d views	4	4	0.00
19	Have you fully perceived the object	5	4	0.02
20	Quality of videos concerning the solving methodology	4	4	0.01
21	Clearness of video, image and sound	9	8	0.00
22	Social Media applications skills	4	4	0.00
23	Classroom fatigue	5	4	0.01
24	Express out loud your questions	4	2	0.00
25	Questions being solved during synchronous lectures	5	4	0.00
26	Resent when instructor does not return graded tasks on time	5	4	0.00
27	Comments by the instruction help understand mistakes	4	3	0.00
28	Instructor helped meeting new people during synchronous lectures	4	3	0.00
29	Knowledge Weaknesses	3	2	0.00
30	Computer skills	4	4	0.00
31	Tasks assigned relevant to future work	5	4	0.23
32	Presentation and clarity of the 15th assignment	8	7	0.00
33	15th assignment related to real world tasks	4	3	0.00
34	All assignments related to real world tasks	4	4	0.00
35	Overall evaluation of the module	4	3	0.00
36	Dealing with weaknesses and knowledge lacks	4	3	0.00
37	Is it likely to succeed in a similar future task	9	7	0.00

3.5. hierarchical cluster analysis

A hierarchical cluster analysis is performed in SPSS to detect the number of clusters into which, the students can be divided, according to the proximity of their responses. This clustering method uses variance analysis techniques to calculate the distances between clusters. At each stage of the method, the clusters are joined, with the smallest sum of squared errors (SSE), and the sum of squares acts as a criterion of loss.

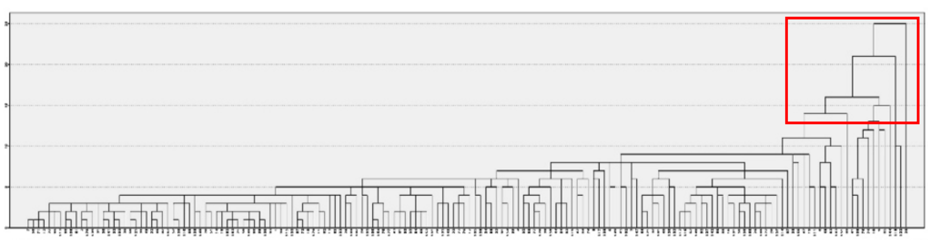


Figure 3. Dendrogram using average linkage (between groups)

The resulting tree diagram (dendrogram) is given below. In the upper branch of the hierarchical clustering tree, it is observed that there are two large clusters and this information is used to implement the k-means algorithm. The next lower branch contains three clusters, and after that there is a lower branch that contains five clusters.

4. Conclusions and Future work

The e-learning paradigm has always included instructional materials, internet accessibility, student interaction, and communication [16]. Most University online courses' lectures and laboratories in universities worldwide and specifically in the presented case of the University of West Attica have been using MS Teams (Communication Platform) for synchronous lectures and Moodle or e-Class (Learning Management Systems) for asynchronous support during the transition from face-to-face to the COVID-19 emerging learning environments.

In this study, the use of MS Teams as a Learning Management System has been used and tested in order to evaluate a single platform use [17] aiming to avoid disruption for students with parallel LMSs functions during the pandemic. Data from online surveys have been mined and filtered through a statistical analysis which included a correlation analysis and a reliability analysis. 37 variables have been considered to have an important significance on testing the performance of tasks' assignment into a single platform, used at the same time for synchronous lectures. The Cronbach's Alpha coefficient has been calculated and resulted to 89% of the survey questions being internally consistent reliable variables. A cluster analysis has been performed dividing students into two clusters, based on the criteria of their diligence, their communication skills, and their level of knowledge embedding.

One of the most noteworthy findings of this research is that having a single online platform for transmitting synchronous lectures while also uploading tasks' assignments, students' evaluation, module notes, and grading, as well as asynchronous support by integrating assignments and resources links uploaded on social media channels makes task organization easier for both students and instructors.

Future work consists on performing a statistical analysis on other constructs of the survey questions, aiming to reveal variables with important significance correlated to students' academic achievements (performance on the final exam), as well as their level of satisfaction, as an indicator of whether online education can become an efficient tool for online learning, or even virtual and blended environments [18].

References

[1] A. Chaudhary, Teaching During Covid-19. esculapio. (2021). 16. 10.51273/esc20.25164-guesteditorial
[2] J. Luo, J. Rajendra, M. Roberts, N. Rapallo, F. Khan, C. Mardon, (2020). Education During COVID-19.
[3] C. Whittle, S. Tiwari, S. Yan, J. Williams, "Emergency remote teaching environment: a conceptual framework for responsive online teaching in crises", in Information and Learning Sciences, (2020) Vol. 121 No. 5/6, pp. 311-319, published on 29/06/2020, https://doi.org/10.1108/ILS-04-2020-0099
[4] C. Troussas, A. Krouska, C. Sgouropoulou, I. Voyiatzis, Ensemble Learning Using Fuzzy Weights to Improve Learning Style Identification for Adapted Instructional Routines, *Entropy* **22**(7) (2020), 735. https://doi.org/10.3390/e22070735.
[5] A. Krouska, C. Troussas, C. Sgouropoulou, Fuzzy Logic for Refining the Evaluation of Learners' Performance in Online Engineering Education, *European Journal of Engineering and Technology Research* **4**(6) (2019), 50-56. DOI: https://doi.org/10.24018/ejers.2019.4.6.1369.

[6] C. Troussas, F. Giannakas, C. Sgouropoulou, I. Voyiatzis, Collaborative activities recommendation based on students' collaborative learning styles using ANN and WSM, *Interactive Learning Environments* (2020). DOI: 10.1080/10494820.2020.1761835.
[7] C. Troussas, A. Krouska and C. Sgouropoulou, A Novel Teaching Strategy Through Adaptive Learning Activities for Computer Programming, *IEEE Transactions on Education* **64**(2) (2021), 103-109. DOI: 10.1109/TE.2020.3012744.
[8] C. Troussas, A. Krouska, C. Sgouropoulou, Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education, *Computers and Education* **144** (2019), 103698. https://doi.org/10.1016/j.compedu.2020.103698.
[9] C. Troussas, A. Krouska and C. Sgouropoulou, Improving Learner-Computer Interaction through Intelligent Learning Material Delivery Using Instructional Design Modeling, *Entropy* **23**(6) (2021), 668. https://doi.org/10.3390/e23060668
[10] M. Moore, *Theory of transactional distance*, Keegan, D., ed. "Theoretical Principles of Distance Education, Routledge, UK, 1997.
[11] Z. Kanetaki, C. Stergiou, G. Bekas and E. Kanetaki, "Machine Learning and Statistical Analysis applied on Mechanical Engineering CAD course: A Case Study During ERTE Pahse in the Context of Higher Education," 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), (2020) pp. 1-13, doi: 10.1109/ISMSIT50672.2020.9254924.
[12] C. Hodges, S. Moore, B. Lockee, T. Trust, A. Bond, The Difference Between Emergency Remote Teaching and Online Learning, Educausereview March 27, (2020).
[13] H.O. Fuchs, Outside reality inside the classroom. Journal of Engineering Education 60 (7) (1970).: 745–47.
[14] A. Yadav, G.M. Shaver, P. Meckl, Lessons Learned: Implementing the Case Teaching Method in a Mechanical Engineering Course. Journal of Engineering Education. (2010), 99: 55-69. https://doi.org/10.1002/j.2168-9830.2010.tb01042.x
[15] M. Binkley, O. Erstad, J. Herman, S. Raizen, M. Ripley, M. Miller-Ricci, M. Rumble, Defining Twenty-First Century Skills. (2011). 10.1007/978-94-007-2324-5_2.
[16] M. Krašna, I. Pesek, "Influence of Moodle and MS Teams on teaching-learning-studying (TLS) processes," 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), (2020), pp. 612-616, doi: 10.23919/MIPRO48935.2020.9245356.
[17] K. Wilbur, Evaluating the online platform of a blended-learning pharmacist continuing education degree program. Med. Educ. Online (2016).
[18] K. Kabassi, I. Dragonas, A. Ntouzevits et al. Evaluating a learning management system for blended learning in Greek higher education. SpringerPlus 5, 101 (2016). https://doi.org/10.1186/s40064-016-1705-8

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210089

Visual Analytics in Process Mining for Supporting Business Process Improvement

Antonia KAOUNI[a], Georgia THEODOROPOULOU[a,1], Alexandros BOUSDEKIS[a], Athanasios VOULODIMOS[a], and Georgios MIAOULIS[a]
[a] Department of Informatics and Computer Engineering, University of West Attica, Athens, Greece

Abstract. The increasing amounts of data have affected conceptual modeling as a research field. In this context, process mining involves a set of techniques aimed at extracting a process schema from an event log generated during process execution. While automatic algorithms for process mining and analysis are needed to filter out irrelevant data and to produce preliminary results, visual inspection, domain knowledge, human judgment and creativity are needed for proper interpretation of the results. Moreover, a process discovery on an event log usually results in complicated process models not easily comprehensible by the business user. To this end, visual analytics has the potential to enhance process mining towards the direction of explainability, interpretability and trustworthiness in order to better support human decisions. In this paper we propose an approach for identifying bottlenecks in business processes by analyzing event logs and visualizing the results. In this way, we exploit visual analytics in the process mining context in order to provide explainable and interpretable analytics results for business processes without exposing to the user complex process models that are not easily comprehensible. The proposed approach was applied to a manufacturing business process and the results show that visual analytics in the context of process mining is capable of identifying bottlenecks and other performance-related issues and exposing them to the business user in an intuitive and non-intrusive way.

Keywords. Visual analytics, process analytics, process intelligence, visualization, data analytics, business process management

1. Introduction

The amount of data recorded in various domains has been growing exponentially [1]. This offers opportunities for algorithmic techniques, but also creates new challenges. The availability of data has extended conceptual modeling as a research field of manually created models with automatic techniques for generating models from data [2]. Process mining is one of these recent extensions. Process mining involves a set of techniques aimed at extracting a process schema from an event log generated during process execution [3][4]

Combining automatic analysis of event logs and visualization methods face several challenges [3]. While automatic algorithms for process mining and analysis are certainly needed to filter out irrelevant data and to produce preliminary results, visual inspection, domain knowledge, human judgment and creativity are needed for proper

[1] Corresponding Author, Georgia Theodoropoulou, University of West Attica, 12243 Ag. Spyridonos Athens, Greece; E-mail: gtheodoropoulou@uniwa.gr.

interpretation of the results [1]. Moreover, a process discovery on an event log usually results in complicated and/or huge process models. Most of the current visualizations display the conformance results in their entirety. From a user perspective, displaying the conformance results on a big and complicated process model with many nodes does not provide significant added value [5]. Such techniques turned out to have several problems when addressing processes with little or no structure, where the flow of the activities mostly depends on the decisions taken by the involved people, thus leading to the so-called "spaghetti-like" models, almost incomprehensible for a human analyst [6].

On the other hand, visual analytics has the potential to provide an answer to these problems by proposing a tight integration between automatic techniques and visualization. In the past ten years, a variety of visual analytics methods have been proposed to make machine learning more explainable, trustworthy, and reliable [7][8][9][10]. Similarly to the way that machine learning has coupled with visualization methods, process mining needs to follow the same direction. In this paper we propose an approach for identifying bottlenecks in business processes by analyzing event logs and visualizing the results aiming at supporting human decisions. In this way, we exploit visual analytics in the process mining context in order to provide explainable and interpretable analytics results for business processes without exposing to the user complex process models that are not easily comprehensible.

The rest of the paper is organized as follows: Section 2 describes the theoretical background upon which the proposed approach is based. Section 3 presents the proposed approach for identifying bottlenecks in business processes with the use of visual analytics. Section 4 outlines the implementation of the proposed approach and demonstrates the results from its application to a manufacturing business process. Section 5 concludes the paper and outlines our plans for future work.

2. Theoretical Background

The term visual analytics was coined by Jim Thomas to mean "the science of analytical reasoning facilitated by visual interactive interfaces" [11]. Over time, the scope of visual analytics broadened. Now the term refers to a multidisciplinary field that combines elements from human-computer interaction, geo-spatial and temporal data processing, data analysis and statistics [1] [12].

During the last years, machine learning approaches in process mining have gathered an increased research and business research by automating data analytics algorithms for business processes deriving detection and prediction outcomes. However, such approaches require the business analyst to configure and train the models in order to make the most out of them. On the other hand, visualization approaches utilize the power of human perception to simultaneously process large amounts of data [13], thus allowing computer and human working in optimal complementarity with non-intrusive decision support.

In this context, visual analytics exploits both the computational power of the computer and the human's perception system to facilitate insights and enable knowledge discovery in large and complex bodies of data [13]. Visual analytics in process mining has been identified as a domain with high potential [1] and a key factor for the wide adoption of process mining methods, techniques and software tools [15], still facing several challenges [13].

Traditionally, visual analytics combined visualization with data mining techniques. However, given the maturity of process mining techniques, visual analytics based on process mining has been considered as a significant research direction [3]. Within process mining, performance analysis is used to explore issues such as identifying any bottlenecks in the process and the steps needed to optimize the process [5]. It is typically focused on performance indicators of the time dimension, such as the lead-, service- and waiting time and, as the name implies, is based on a process model [14]. Many commercial and free process mining tools allow doing such analysis; however, the commonly used model notations are not designed to project the time dimension on the model, i.e. changes over time cannot be represented in a comprehensible way [14]. On the other hand, although conformance results enable compliance and performance analysis, there are still some open challenges which could make the analysis difficult [5].

Since visual analytics aims at supporting human reasoning and gaining new insights, the quality of visual analytics approaches is hard to quantify [13]. However, process models extracted from the process mining algorithms aim at showing individual event sequences, conformance of individual event sequences, and variations of individual cases [13]. Therefore, they usually result in "spaghetti-like" process models that are not easily comprehensible by the business user, while they do not usually extract results related to time.

3. The Proposed Approach

The proposed approach aims at identifying bottlenecks in business processes by taking advantage of the large amounts of event logs becoming available through various information systems. In order to tackle the complexity of the process models extracted by the event logs, we propose a visual analytics approach that facilitates interpretability and explainability of the results in order to support business decisions. In this way, the human expert is in the loop having the capability of taking informed decisions. The proposed approach incorporates an incremental and iterative way and provides interactive visualization capabilities in order to facilitate the interaction with the human. The proposed approach consists of the following steps:

- **(A) Event Log Extraction**: The starting point is to extract the event logs. An event log is a file including the activities along with a timestamp. Every time a process is executed, a process instance is created. Each process instance is called a case (or trace). Each trace describes the lifecycle of a particular case (i.e., a process instance) in terms of the activities executed [15]. Every case has several ordered activities. The event log required for the proposed approach needs to incorporate the case number, the activity name, the timestamp (end time) and the start time.
- **(B) Filtering event logs**: We need to filter activities, which can be concluded to have faults from event log inspection. Also we can find outliers and filter them out. Filtering is done for two main reasons: cleaning the data or narrowing down the analysis. Filtering may concern removing a process instance (case), adding events, removing events or modifying events [1].
- **(C) Start and End Activities**: In this step, the proposed approach provides the identified start and end activities. In this way, the human may detect variations

in the process execution. For example, in case there are more than one end activities, the business process may suffer from incomplete instances.

(D) Start dates of traces and activities: In this step, the proposed approach provides visualization results about the duration of the various activities. From these results, the human may detect the most long-lasting activities and how their duration changed throughout the time period under examination.

(E) Duration of the activities: In this step, the visualization results provide information about the traces of the shortest and the longest duration in order to enable focusing on the improvement of the potentially problematic traces.

(F) Long-lasting traces: In this step, the visualization results detect the total time of completion of the traces in order to identify the most long-lasting traces as well as their patterns throughout the time period under examination. In addition to that, the human is exposed to a more detailed graph combining the start dates and the activities of each trace.

(G) Human cognition: In the final step, the human aggregates the aforementioned visualization results in order to make decisions about the efficiency of the business processes as well as about actions for improvement. The human decisions are supported by the results of visual analytics making capable of taking advantage of the large amounts of event logs generated by information systems without requiring advanced data analytics and machine learning skills.

4. Implementation of a Visual Analytics Web Application

The proposed approach was implemented as a web application with the Python programming language. For this implementation, we used 3 Python libraries:

- **Pandas** (https://pandas.pydata.org/pandas-docs/stable/index.html) for data structuring. In particular, we used Pandas DataFrame which is a two-dimensional data structure, i.e., data is aligned in a tabular fashion in rows and columns. Pandas DataFrame consists of three principal components, the data, rows, and columns.
- **PM4py** (https://pm4py.fit.fraunhofer.de/) for process mining. PM4py provides functionalities for handling and filtering event data, performing process discovery, managing Petri nets and conducting conformance checking. It also performs evaluation and simulation, and provides process statistics capabilities.
- **Plotly** (https://plotly.com/) for delivering machine learning and data science to business users. Plotly enables building and deploying analytic web apps using Python, R, and Julia without the use of JavaScript or DevOps.

5. Application to a Manufacturing Business Process

The proposed approach was applied to a manufacturing business process dealing with the control of a plant to increase its overall performance. The plant produces parts made of metal such as spurs, fastener, ball nuts, discs, tubes, wheel shafts, or clamps. To build these parts there are 28 machines for lapping, milling, turning, sinking, wire cutting, turning and milling, laser marking, and round and flat grinding [17] [18]. The wide adoption of IoT devices, sensors and actuators in manufacturing environments has fostered an increasing research interest on real-time data analytics; however, event logs contain information regarding the whole factory cycle, either they have sensors installed or not, thus, having the credentials to move towards providing an all-around view of manufacturing operations on the shopfloor [19]. In this way, they may provide valuable information for planning and resource allocation [18].

First, according to the steps (A) and (B) of the proposed approach, the event log is extracted and filtered in order to be structured in a way that can feed into the subsequent visual analysis. The dataset contains process data from a production process, including data on cases, activities, resources, timestamps, etc. Table 1 presents a data sample from the extracted event log. As already mentioned, the most important attributes for the proposed approach are: the *Case ID*, the *Start Timestamp*, the *Complete Timestamp*, and the *Activity*. The *Resource* provides additional information related to the Activity and can be used for further analysis.

Table 1. Data sample from the event log of the manufacturing business process.

Case ID	Start Timestamp	Complete Timestamp	Activity	Resource
Case 1	1/29/2012 23:24	1/30/2012 5:43	Turning & Milling - Machine 4	Machine 4 - Turning & Milling
Case 1	1/30/2012 5:44	1/30/2012 6:42	Turning & Milling - Machine 4	Machine 4 - Turning & Milling
Case 1	1/30/2012 6:59	1/30/2012 7:21	Turning & Milling - Machine 4	Machine 4 - Turning & Milling
Case 1	1/30/2012 7:21	1/30/2012 10:58	Turning & Milling - Machine 4	Machine 4 - Turning & Milling
Case 1	1/31/2012 13:20	1/31/2012 14:50	Turning & Milling Q.C.	Quality Check 1
Case 1	2/1/2012 8:18	2/1/2012 8:27	Laser Marking - Machine 7	Machine 7- Laser Marking
Case 1	2/14/2012 0:00	2/14/2012 1:15	Lapping - Machine 1	Machine 1 - Lapping
Case 1	2/14/2012 0:00	2/14/2012 1:15	Lapping - Machine 1	Machine 1 - Lapping
Case 1	2/14/2012 9:05	2/14/2012 10:20	Lapping - Machine 1	Machine 1 - Lapping
Case 1	2/14/2012 9:05	2/14/2012 9:38	Lapping - Machine 1	Machine 1 - Lapping

Figure 1 depicts part of the process model derived from the aforementioned event log. As it can be noticed, it suffers from a high complexity making it impossible for the business user to understand the process. However, explainability and transparency in manufacturing is essential in order to inspire confidence that the system is secure, robust, and controllable [20]. This can be achieved with a visual analytics approach.

According to the step (C), the proposed approach presents to the human the start and the end activities. In the scenario under examination, there are several different start and end activities, as shown in Figure 2 and in Figure 3. Four start activities have a high frequency (i.e. *Turning & Milling – Machine 4, Turning & Milling – Machine 8, Turning & Milling – Machine 5, Turning & Milling – Machine 6*). Their context indicates that there are parallel production lines. However, there are also several other start activities. On the other hand, there are mainly two end activities (*Packing, Final Inspection Q.C.*), although there is a wide distribution on other activities. These facts imply a complex business process with high variety, e.g. depending on the products, the maintenance operations, the remanufacturing, etc. On the other hand, this relatively unstructured process may indicate that the process either is executed in an inefficient way or it should be redesigned.

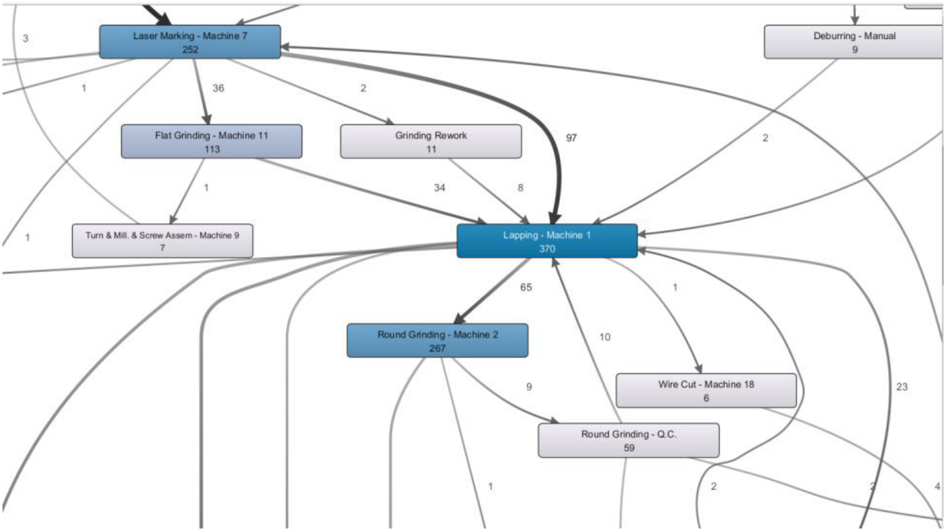


Figure 1. The process model of the manufacturing use case.

According to step (D), the proposed approach provides a Cleveland dot plot in order to result in an intuitive outcome regarding the activity durations through time. Cleveland dot plots are an alternative to bar graphs that reduce visual clutter and can be easier to read. Figure 4 depicts the duration of the various activities with respect to the start timestamp. Each data point has a different colored coding according to the activity it represents. In addition, Figure 5 depicts the mean duration of each activity enabling the detection of the most long-lasting activities.

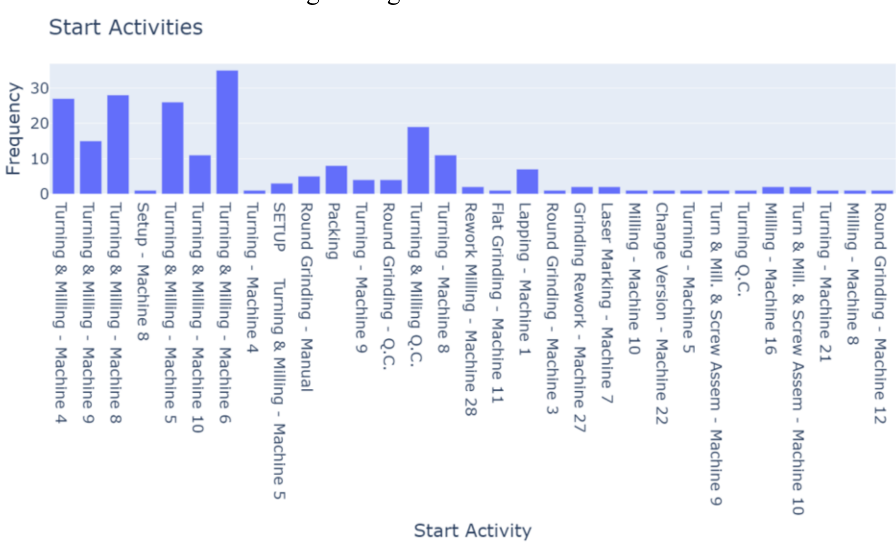


Figure 2. The start activities.

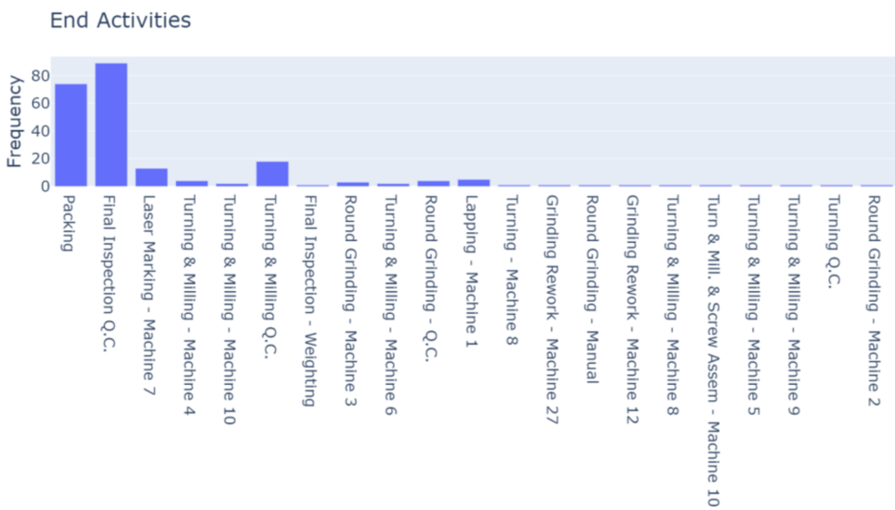


Figure 3. The end activities.

According to steps (E) and (F), Figure 6 depicts the total time for completion of the traces so that the business user is able to see the traces that required more time. For example, in the scenario under examination, one may notice that the traces tend to require much more time to be finished throughout time. Moreover, Figure 7 depicts a combination between the start timestamps of each trace and its activities.

According to step (G), the human decisions are supported by the aforementioned visual analytics on the manufacturing business process. These decisions may deal with both strategic and operational aspects of the business process. Therefore, for the scenario under examination, the business user may draw the following conclusions:

- The process under examination is subject to a high variety, since there are several different start and end activities. To some extent, this is reasonable due to the nature of the production network. However, the event log includes some start and end activities that seem to be outliers. Therefore, there are incomplete processes due to either the process execution or the process design itself.
- There are some variations with respect to the mean duration of similar activities. For example, the activities *Milling - Machine 8* and *Milling - Machine 16* have a significantly longer mean duration comparing to the activity *Milling - Machine 10*. Similarly, the activity *Grinding Rework – Machine 2* has a significantly longer mean duration comparing to the activity *Grinding Rework – Machine 12*.
- Some activities have a significantly longer duration comparing to others causing bottlenecks and leading to delays of the whole process.
- Some traces require more time to be completed comparing to others. Throughout time, the traces tend to have a longer duration revealing an increasing inefficiency and responsiveness in the process execution.

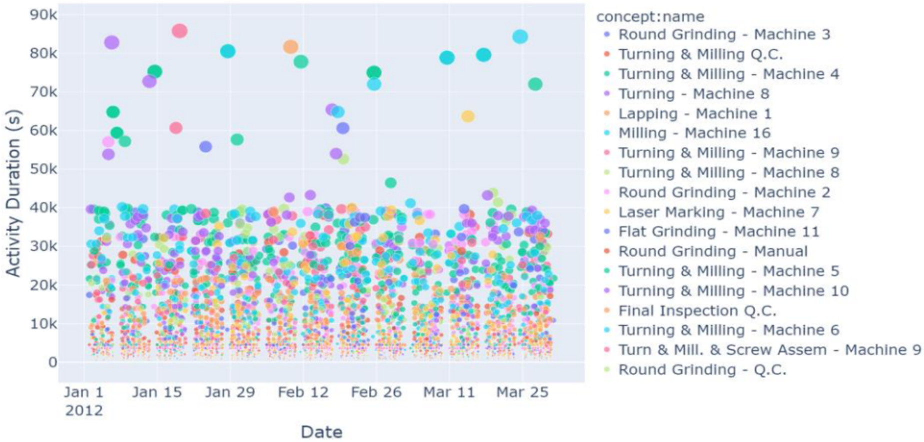


Figure 4. Activity duration through time (step D).

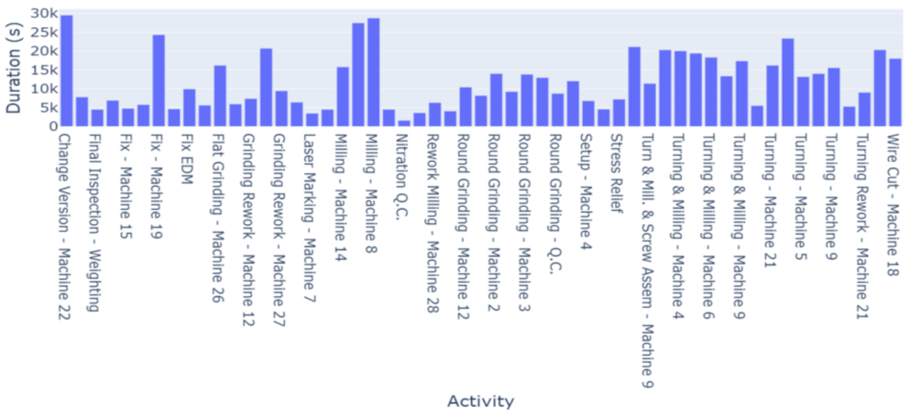


Figure 5. Mean duration of each activity (step D).

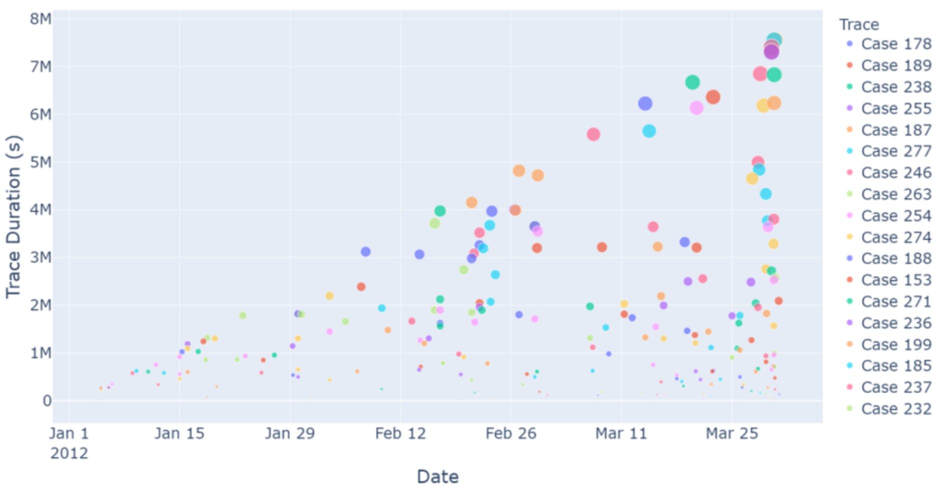


Figure 6. Trace durations through time (step E and F).

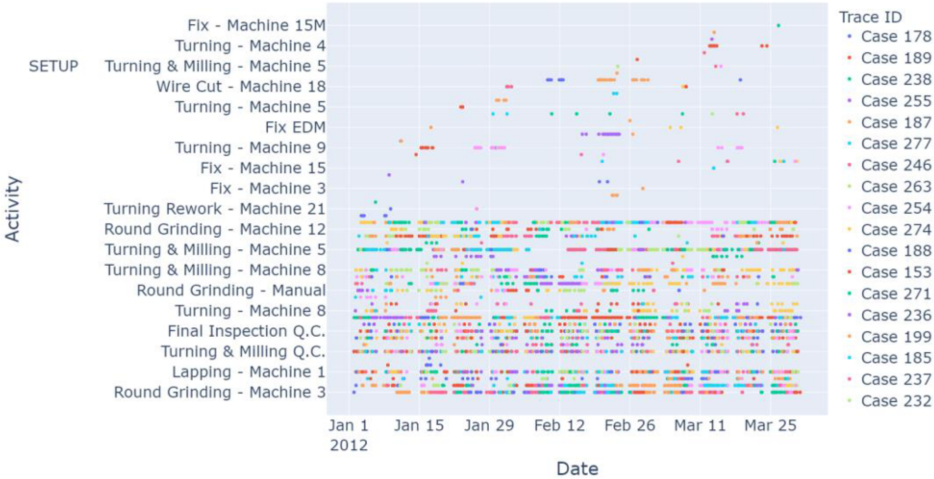


Figure 7. Activities and traces through time (step E and F).

6. Conclusions and Future Work

The increasing amounts of data have affected conceptual modeling as a research field by utilizing data analytics for generating process models. In this context, process mining involves a set of techniques aimed at extracting a process schema from an event log generated during process execution. While automatic algorithms for process mining and analysis are needed to filter out irrelevant data and to produce preliminary results, visual inspection, domain knowledge, human judgment and creativity are needed for proper interpretation of the results. Moreover, a process discovery on an event log usually results in complicated and/or huge process models not easily comprehensible by the business user. To this end, visual analytics has the potential to enhance process mining towards the direction of explainability, interpretability and trustworthiness in order to support human decisions.

In this paper we proposed an approach for identifying bottlenecks in business processes by analyzing event logs and visualizing the results. In this way, we exploit visual analytics in the process mining context in order to provide explainable and interpretable analytics results for business processes without exposing to the user complex process models that are not easily comprehensible. The proposed approach was applied to a manufacturing business process and the results show that visual analytics in the context of process mining is capable of identifying bottlenecks and other performance-related issues and exposing them to the business user in an intuitive and non-intrusive way.

Regarding our future work, we plan to implement advanced data analytics algorithms performing predictions on event logs. To do this, we will implement (deep) machine learning algorithms and we will focus on coupling them with visualization methods in order to achieve explainability and interpretability of complex algorithms and results.

References

[1] Van der Aalst, W. M., de Leoni, M., & ter Hofstede, A. H. (2011). Process mining and visual analytics: Breathing life into business process models. *BPM Center Report BPM-11-15, BPMcenter. org*, **17**, 699-730.

[2] Yeshchenko, Anton & Di Ciccio, Claudio & Mendling, Jan & Polyvyanyy, Artem (2019), Comprehensive Process Drift Detection with Visual Analytics.

[3] Van Der Aalst, W., Adriansyah, A., De Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., ... & Wynn, M. (2011b, August). Process mining manifesto. In International Conference on Business Process Management (pp. 169-194). Springer, Berlin, Heidelberg.

[4] Van Der Aalst, W. (2016). *Data science in action. In Process mining* (pp. 3-23). Springer, Berlin, Heidelberg.

[5] Buijs, J. C. A. M., & van der Aalst, W. M. P. (2017). Enabling interactive process analysis with process mining and visual analytics. *BIOSTEC* 2017, 573.

[6] Diamantini, C., Genga, L., Potena, D., & van der Aalst, W. M. (2015). *Towards Process Instances Building for Spaghetti Processes*. In SEBD (pp. 256-263).

[7] Cui, W. (2019). Visual analytics: A comprehensive overview. *IEEE Access*, **7**, 81555-81573.

[8] Yuan, J., Chen, C., Yang, W., Liu, M., Xia, J., & Liu, S. (2021). A survey of visual analytics techniques for machine learning. *Computational Visual Media*, 1-34.

[9] Ahn, Y., & Lin, Y. R. (2019). Fairsight: Visual analytics for fairness in decision making. IEEE transactions on visualization and computer graphics, 26(1), 1086-1095.

[10] Krak, I., Barmak, O., & Manziuk, E. (2020). Using visual analytics to develop human and machine - centric models: A review of approaches and proposed information technology. *Computational Intelligence*.

[11] Thomas, J. J., & Cook, K. A. (2005). Illuminating the path. Los Alamitos, Calif, *IEEE Computer Society*.

[12] Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008). Visual analytics: Scope and challenges, *Visual data mining* (pp. 76-90). Springer, Berlin, Heidelberg.

[13] Gschwandtner, T. (2015, December). Visual Analytics Meets Process Mining: Challenges and Opportunities, *International Symposium on Data-Driven Process Discovery and Analysis* (pp. 142-154). Springer, Cham.

[14] Denisov, V., Belkina, E., Fahland, D., & van der Aalst, W. M. (2018, September). The Performance Spectrum Miner: Visual Analytics for Fine-Grained Performance Analysis of Processes. *In BPM (Dissertation/Demos/Industry)* (pp. 96-100).

[15] Wong, P. C., Shen, H. W., Johnson, C. R., Chen, C., & Ross, R. B. (2012). The top 10 challenges in extreme-scale visual analytics. *IEEE computer graphics and applications*, **32** (4), 63-67.

[16] Wil M. P. Aalst, 2015. "Extracting Event Data from Databases to Unleash Process Mining," Management for Professionals, in: Jan vom Brocke & Theresa Schmiedel (ed.), *BPM - Driving Innovation in a Digital World, edition* **127**, pages 105-128, Springer.

[17] D. Levy, "Production Analysis with Process Mining Technology. Dataset," 2014. [Online]. Available: https://data.4tu.nl/repository/collection:event_logs

[18] Tello-Leal, E., Roa, J., Rubiolo, M., & Ramirez-Alcocer, U. M. (2018, November). Predicting activities in business processes with LSTM recurrent neural networks. In 2018 *ITU Kaleidoscope: Machine Learning for a 5G Future* (ITU K) (pp. 1-7). IEEE.

[19] Lepenioti, Katerina & Pertselakis, Minas & Bousdekis, Alexandros & Louca, Andreas & Lampathaki, Fenareti & Apostolou, Dimitris & Mentzas, Gregoris & Anastasiou, Stathis. (2020). Machine Learning for Predictive and Prescriptive Analytics of Operational Data in Smart Manufacturing. 10.1007/978-3-030-49165-9_1.

[20] Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). A human cyber physical system framework for operator 4.0–artificial intelligence symbiosis. *Manufacturing Letters*, **25**, 10-15.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210090

ActiveCrowds: A Human-in-the-Loop Machine Learning Framework

Lazaros TOUMANIDIS [a,1], Panagiotis KASNESIS [a], Christos CHATZIGEORGIOU [a], Michail FEIDAKIS [a] and Charalampos PATRIKAKIS [a]

[a] *University of West Attica, dept. of Electrical & Electronics Engineering, Egaleo, Greece*

Abstract. A widespread practice in machine learning solutions is the continuous use of human intelligence to increase their quality and efficiency. A common problem in such solutions is the requirement of a large amount of labeled data. In this paper, we present a practical implementation of the human-in-the-loop computing practice, which includes the combination of active and transfer learning for sophisticated data sampling and weight initialization respectively, and a cross-platform mobile application for crowdsourcing data annotation tasks. We study the use of the proposed framework to a post-event building reconnaissance scenario, where we utilized the implementation of an existing pre-trained computer vision model, an image binary classification solution built on top of it, and max entropy and random sampling as uncertainty sampling methods for the active learning step. Multiple annotations with majority voting as quality assurance are required for new human-annotated images to be added on the train set and retrain the model. We provide the results and discuss our next steps.

Keywords. Active Learning, Transfer Learning, Crowdsourcing, Mobile Computing

1. Introduction

There is no doubt that Machine Learning (ML), as a field of Artificial Intelligence (AI) and consequently as part of the 4th Industrial Revolution [1], has changed our work, our social interactions, our life in general. One of the major bottlenecks and open research topic in ML is the massive data collection and labeling requirement, which is essential for training supervised learning models. Towards, Transfer Learning (TL) constitutes a technique that reduces the amount of data that need to be used during the training and testing process. The objective of TL is to take advantage of data from an existing setting to extract information that may be useful when learning or even when making predictions in a different setting [2].

[1]Corresponding Author: Toumanidis Lazaros, University of West Attica, dept. of Electrical & Electronics Engineering, Egaleo, Greece; E-mail: laztoum@uniwa.gr.

In the case of supervised learning, further reduction in the data labeling process is achieved using Active Learning (AL) that reduces data labeling effort by actively selecting the most informative instances to be queried for labeling. In contrast with Passive Learning, where all labeled samples are obtained once and without a reference to the learning algorithm, new samples are interactively chosen from a pool of unlabeled data that may improve the learning process resulting on the reduction of the total amount of samples needed.

Crowdsourcing is a common technique for both data acquisition and data labeling, where a large group of people, not necessarily domain experts, are asked to either provide or label existing samples. The term crowdsourcing appeared in [3] as a paradigm, in which a specific service, information or task is offloaded to a crowd of individuals, often connected through a common goal or interest, as in an online community.

Using the above three techniques, i.e. TL, AL, and Crowdsourcing, we have implemented an end-to-end solution that can be applied in several domains in supervised ML, including computer vision, text and audio processing. Also, it can be easily used on-field to feed any ML pipeline by exploiting a user-friendly mobile application. Tasks are initialized with existing models using TL, and new samples are appended to the training data set using AL. Crowdsourcing is used for the labeling step of each task, by deploying a mobile application as an interface. In this paper, we apply our solution to post-event building reconnaissance use case, demonstrating its capabilities on reducing the amount of annotated data needed. In particular, we selected a binary image classification task using ResNet [4] as the initial model, a smaller image dataset with images of buildings that are either collapsed or not [5], and uncertainty sampling for querying new samples.

2. Related Work

TL techniques have been successfully applied in real-world applications [6] in several domains, including Natural Language Processing [7,8], Computer Vision [9,10], and Audio Processing [11,12]. In [13], Yang et al. present a study on the combination of TL and AL, focusing on the improvement of the learning accuracy, concerning the size of the required labels. Zhou et al. [14] present active fine-tuning (AFT), a new algorithm which naturally integrates AL and TL. Their results include a significant cut in the annotation cost compared to the state-of-the-art method [15]. High detection accuracy is achieved in [16], where Feng et al. use a deep residual network for defect detection and classification in infrastructure surface images, and apply AL strategy, asking experts to label the most informative subset of new images to retrain the network. Tong and Koller in [17] propose the use of Support Vector Machines (SVM) for samples selection in AL. By applying the latter to text classification, they achieve a significant reduction in the need for labeled instances.

There have been several studies on measuring and ensuring the quality of crowdsourced data. Wang et al. in [18] propose ARTSense, a framework to solve the problem of “trust without identity” in mobile sensing. The solution consists of a privacy-preserving provenance model, a data trust assessment scheme, and an anonymous reputation management protocol. Tian et al. have presented Max-Margin Majority voting as a new approach for improving the discriminative ability of majority voting in crowdsourcing [19].

Regarding the combination of AL and Crowdsourcing for data labeling, Fang et al. focus on the optimization of the sampling techniques in [20]. Zhang et al. propose a deep computational model with crowdsourcing for industrial IoT Big Data, trying to prevent overfitting and aggregate adequately labeled samples to train a model's parameters [21]. Song et al. propose a confidence-based Crowdsourcing approach for data labeling, in which the confidence of the crowd workers has been considered for aggregating the results [22]. A combination of the learning algorithm uncertainty, and the uncertainty derived from the crowd-sourced answers, is used to define a score function, for new instances selection. Zhao et al. study the use of AL and Crowdsourcing for activity recognition [23]. In their work, they present three methods to choose the most informative data points based on low confidence for the most probable activity class, the minimum difference between the confidence of the most and second most probable class, and high entropy among the probability of classes. A combination of ontological knowledge and AL is presented in Civitarese et al. [24], in which users' feedback is deployed to refine the correlations among sensor events and activity types initially extracted from a high-level ontology, in order to mine temporal patterns of sensor events that are frequently generated by the execution of specific activities.

In the domain of sound processing, MoodSwings [25] and TagAtune [26] are used for data labeling with mobile applications. The former is focusing on audio mood labeling, and the latter, on tagging music albums. A more generic approach that includes several tasks is provided in 'Crowdsource' [2] for Android and in 'Unbiased WorkForce' for Android[3] and iOS[4].

In this paper, we present an implementation that includes the three techniques mentioned (TL, AL, Crowdsourcing) above into one extendable platform, providing both the means for training new models, and a user interface for crowd-sourcing the data labeling requirement.

3. ActiveCrowds

In this section, we present 'ActiveCrowds', a platform that can be used for managing ML related tasks, utilizing the aforementioned techniques (TL, AL, Crowd-sourcing). It allows ML tasks to be carried out, utilizing the advantages of each technique. It aims at providing a flexible tool that could be used in several domains and extended with new implementations of either ML or crowdsourcing techniques.

3.1. Overview

A high-level overview of a task is displayed in Figure 1. First, a model initialization is required, using an existing pre-trained model. After collecting enough

[2]https://play.google.com/store/apps/details?id=com.google.android.apps.village.boond
[3]https://play.google.com/store/apps/details?id=io.datax.workstation
[4]https://apps.apple.com/tt/app/unbiased-workforce/id1502361378

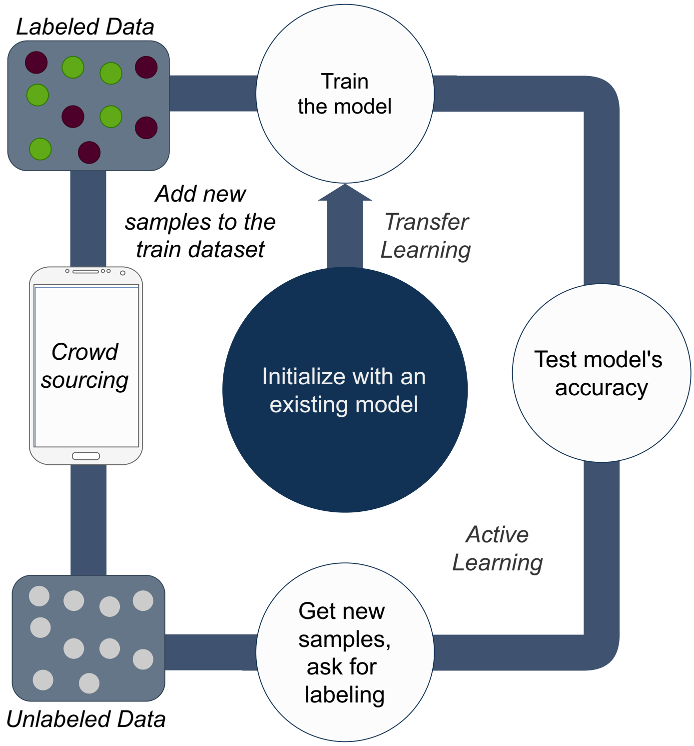


Figure 1. Overview

samples, we train and validate the model; we compute its accuracy on a test data set, we get new samples using one of the available sampling strategies and ask humans for labeling. The new samples that are labeled using the mobile application are added to the training set, and another loop of training–validation–testing sampling begins.

The platform is split into three main modules: (a) The core module related with the training, evaluation and testing of the available models, as well as the process of sampling entries to be labeled. (b) A web service acting as an interface to the core module. (c) A mobile application that is used from the users to complete the labeling tasks. For the core module, we have implemented a Python package that includes an abstract base class for the train, evaluation, testing of a model, and the sampling process. An image binary classification class implements the above methods on top of Torchvision and PyTorch [27]. We have also developed a storage class, responsible for handling the relevant datasets, and a scenario class that is used to hold information about the train–evaluation–test sample loop parameters.

Read and write access to the machine learning tasks is available through a REST API implemented using the Flask framework. The task parameters and their current state are stored in a PostgreSQL database. Background tasks that are triggered using the API, are performed using the Celery library with Redis being used as the message broker and the results' back-end storage. The mobile

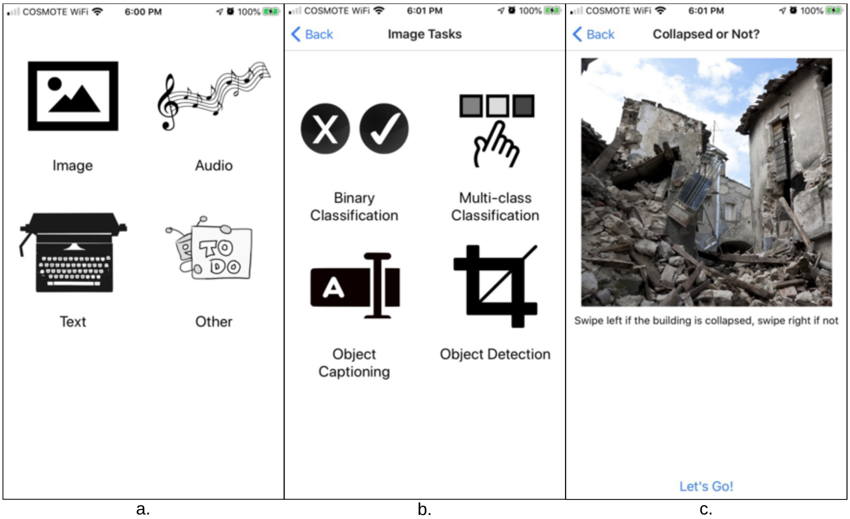


Figure 2. Mobile application screens: a) Task categories, b) Computer Vision tasks and c) Binary Classification task

application is implemented using the React Native framework, making it available to run on both Android and iOS based devices. A task initialization requires:

- The pre-trained model to be used,
- The required dataset sizes for the train, evaluation, and testing steps,
- The query strategy to be used for new samples,
- The minimum number of peer answers and the minimum required accuracy for the labeling decision

3.2. Implementation

An instance of the Scenario class consists of:

- The required length of samples to be used for the active learning phases, i.e., training, evaluation, testing, sampling, sampling strategy (currently one of: *Minimum Margin*, *Least Confidence* and *Max Entropy*),
- The category of the task that is image, audio, text or other,
- The task's subcategory, which the case of an image category may be a binary or multi-class classification, object detection or object captioning.

The REST API currently includes endpoints for the available tasks listing, new task creation, getting task details by its id or updating or deleting one, and a task's labeling answers submission endpoint. On each submission, the task's status is updated, the criterion about deciding for a new label is checked, and if applicable a new active learning loop is started. The mobile application, as depicted in Figure 2 consists of two main screens, (a) the lists of the task categories and subcategories respectively are presented after being fetched from the server, and (b) the tasks' specific screens. In the case of image binary classification, a cover image and a small description are presented, and a batch of the selected samples that need to be labeled are sequentially presented. After completing the batch labeling, the user answers are sent to the server, and the task is marked as locked. When new samples are available, the task is again available for a new submission.

4. Use Case Scenario

The proposed framework has been applied in a post-event building reconnaissance use case scenario. In such an event, a tremendous amount of perishable visual data can be generated in just a few days. As a result, data annotation from trained professional engineers about whether a building has collapsed or not, is time-consuming; it might take days to be completed, so there is a need of an autonomous classification tool [5]. Additionally, to train a deep Convolutional Neural Network (CNN) efficiently, a huge amount of annotated data is required. Therefore, we study the use of our tool in a simulated scenario, in which an earthquake has taken place damaging a lot of buildings.

In more technical terms, we study a binary classification problem, which deals with the prediction on whether a depicted building is collapsed or not. The dataset we used consists of 1850 images of collapsed buildings and 3420 of non-collapsed, for a total of 5270 images. Non-collapsed images mainly consist of undamaged buildings, damaged buildings, and irrelevant pictures, which represents a typical data set collected during an earthquake reconnaissance mission. It has been successfully used on a post-event building reconnaissance study in [5] with a collapse classification accuracy of about 91.5%, trained using the AlexNet CNN [28]. 1000 of them were used for validation and testing of the trained model, and the rest of them, as a pool of new samples to select from, for the AL sampling step (Figure 3, samples of the training dataset).

The task is created using the above dataset, while the initial parameters are set as follows:

- A team of 20 trained professional engineers investigates the area where the earthquake has taken place
- We have an initial training set of 30 images
- A fixed number of 30 images to be appended on this set for a train-validation-test loop of the model, and
- A peer accuracy of 80% with a minimum number of required peers (end-users) of 10, meaning that to decide for a class of an unlabeled image, input from at least 10 peers is required, and the 80% of the answers have to be the same for this image.

The server and the core modules are hosted on a computer workstation equipped with an NVIDIA GTX Titan X GPU, featuring 12 gigabytes RAM, 3072 CUDA cores, and bandwidth of 336.5 GB/s. We used Python as the programming language, and specifically the PyTorch library. To accelerate the tensor multiplications, we used CUDA Toolkit also supported by the cuDNN, which is the NVIDIA GPU-accelerated library for deep neural networks. The software is installed on a Linux operating system.

Figure 3. Train Preview

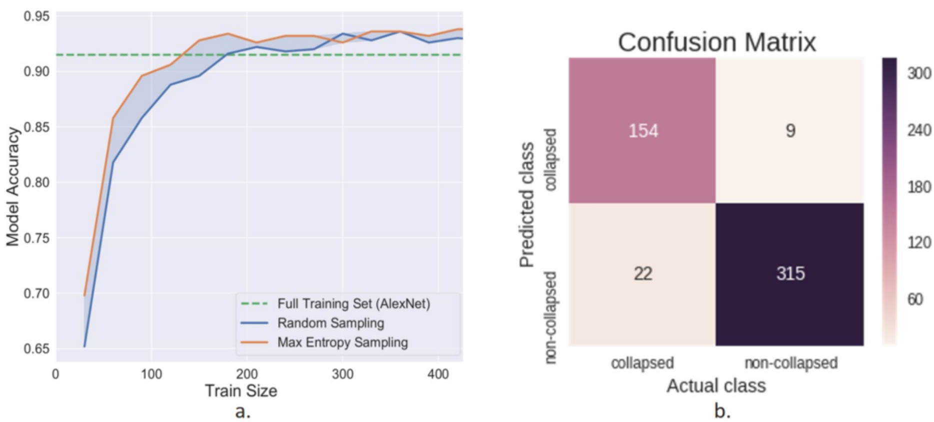


Figure 4. Obtained results: a) model accuracy and b) confusion matrix

Margin sampling and random sampling were deployed for the labelling of new images. As expected, testing the least confidence margin and entropy sampling methods gave precisely the same results with margin sampling, since this is a binary classification problem, and these three methods provide the same results in this case. We achieved an accuracy of about 93.5% with a training size of fewer than 420 images (Figure 4a). Moreover, it is worth mentioning that by exploiting the AL feature of our framework, we achieved almost 1% higher accuracy (92.47%) than the one reported in [5] (91.5%) using only 150 labeled images, which is around 0.09% of the labeled used in [5]. Testing the model on the 500 samples of the test set, gave us a result of 315 true-negative and 154 true-positive predictions presented in the form of a confusion matrix (Figure 4b).

5. Conclusions

This paper presents a novel framework that can be used in ML related tasks, utilizing the methods of Transfer Learning, Active Learning and Crowdsourcing. We successfully deployed and evaluated the platform using a post-earthquake building reconnaissance use case and by setting an image binary classification problem of collapsed building identification, achieving State-of-the-Art results, while using only few labeled instances. We expect that the proposed framework will be a breakthrough in domains where data labelling is both vital and time-consuming. Our next steps include the evaluation of more tasks both in the computer vision domain and in other domains like audio and text processing. Moreover, we intend to explore more active learning query strategies, as well as more methods for the quality assurance of the crowdsourced labels.

Acknowledgment

We would like to gratefully acknowledge Chul Min Yeum Shirley J. Dyke and Julio Ramirez from the Lyles School of Civil Engineering Purdue University for sharing the dataset which was used in our experiments, and the support of NVIDIA for a Titan X GPU card donation that has been used in this research.
This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH-CREATE-INNOVATE (project code: T1EDK-03487).

References

[1] K. Schwab, *The fourth industrial revolution.* Currency, 2017.
[2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning (Adaptive Computation and Machine Learning series).* MIT Press, 2016.
[3] J. Howe, "The rise of crowdsourcing," *Wired magazine*, vol. 14, pp. 1–4, June 2006.
[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 6 2016.
[5] C. M. Yeum, S. J. Dyke, and J. Ramirez, "Visual data classification in post-event building reconnaissance," *Engineering Structures*, vol. 155, pp. 16–24, 2018.
[6] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
[7] W. Dai, G.-R. Xue, Q. Yang, and Y. Yu, "Transferring naive bayes classifiers for text classification," in *Proceedings of the* 22^{nd} *National Conference on Artificial Intelligence - Volume 1*, AAAI'07, pp. 540–545, AAAI Press, 2007.
[8] X. Ling, G.-R. Xue, W. Dai, Y. Jiang, Q. Yang, and Y. Yu, "Can chinese web pages be classified with english data source?," in *Proceedings of the* 17^{th} *International Conference on World Wide Web*, WWW '08, (New York, NY, USA), pp. 969–978, Association for Computing Machinery, 2008.
[9] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection," *Construction and Building Materials*, vol. 157, pp. 322–330, Dec. 2017.
[10] A. Kensert, P. J. Harrison, and O. Spjuth, "Transfer learning with deep convolutional neural networks for classifying cellular morphological changes," *SLAS DISCOVERY: Advancing the Science of Drug Discovery*, vol. 24, no. 4, pp. 466–475, 2019.
[11] W.-N. Hsu, D. Harwath, and J. Glass, "Transfer learning from audio-visual grounding to speech recognition," *arXiv preprint arXiv:1907.04355*, 2019.
[12] M. E. A. Elshaer, S. Wisdom, and T. Mishra, "Transfer learning from sound representations for anger detection in speech," *arXiv preprint arXiv:1902.02120*, 2019.
[13] L. Yang, S. Hanneke, and J. Carbonell, "A theory of transfer learning with applications to active learning," *Machine Learning*, vol. 90, pp. 161–189, Feb. 2013.

[14] Z. Zhou, J. Shin, R. Feng, R. T. Hurst, C. B. Kendall, and J. Liang, "Integrating active learning and transfer learning for carotid intima-media thickness video interpretation," *Journal of Digital Imaging*, vol. 32, pp. 290–299, Apr. 2019.

[15] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.

[16] C. Feng, M. Y. Liu, C. C. Kao, and T. Y. Lee, "Deep active learning for civil infrastructure defect detection and classification," in *Computing in Civil Engineering 2017* (N. El-Gohary, P. Tang, N. El-Gohary, K.-Y. Lin, K.-Y. Lin, and P. Tang, eds.), pp. 298–306, American Society of Civil Engineers (ASCE), Jan. 2017.

[17] S. Tong and E. Chang, "Support vector machine active learning for image retrieval," in *Proceedings of the ninth ACM international conference on Multimedia*, pp. 107–118, ACM, 2001.

[18] X. Wang, W. Cheng, P. Mohapatra, and T. Abdelzaher, "Enabling reputation and trust in privacy-preserving mobile sensing," *IEEE Transactions on Mobile Computing*, vol. 13, pp. 2777–2790, Dec. 2014.

[19] T. Tian, J. Zhu, and Y. Qiaoben, "Max-margin majority voting for learning from crowds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, pp. 2480–2494, Oct. 2019.

[20] M. Fang, J. Yin, and D. Tao, "Active learning for crowdsourcing using knowledge transfer," in *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, AAAI'14, pp. 1809–1815, AAAI Press, 2014.

[21] Q. Zhang, L. T. Yang, Z. Chen, P. Li, and F. Bu, "An adaptive dropout deep computation model for industrial iot big data learning with crowdsourcing to cloud computing," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2330–2337, 2019.

[22] J. Song, H. Wang, Y. Gao, and B. An, "Active learning with confidence-based answers for crowdsourcing labeling tasks," *Knowledge-Based Systems*, vol. 159, pp. 244–258, 2018.

[23] L. Zhao, G. Sukthankar, and R. Sukthankar, "Robust active learning using crowdsourced annotations for activity recognition," in *Proceedings of the 11th AAAI Conference on Human Computation*, AAAIWS'11-11, pp. 74–79, AAAI Press, 2011.

[24] G. Civitarese, C. Bettini, T. Sztyler, D. Riboni, and H. Stuckenschmidt, "newnectar: Collaborative active learning for knowledge-based probabilistic activity recognition," *Pervasive and Mobile Computing*, vol. 56, pp. 88–105, 2019.

[25] Y. Kim, E. Schmidt, and L. Emelle, "Moodswings: A collaborative game for music mood label collection.," in *ISMIR*, pp. 231–236, 01 2008.

[26] E. L. M. Law, L. V. Ahn, R. B. Dannenberg, and M. Crawford, "Tagatune: A game for music and sound annotation," 2007.

[27] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, eds.), pp. 8024–8035, Curran Associates, Inc., 2019.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, (Red Hook, NY, USA), pp. 1097–1105, Curran Associates Inc., 2012.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210091

A Fuzzy C-Means-Based Algorithm for the Surveillance of Dengue Cases Distribution in Local Communities

Jozelle C. ADDAWE[a,1], Jaime D.L. CARO[a] and Richelle Ann B. JUAYONG[a]
[a]*Service Science and Software Engineering Laboratory, University of the Philippines Diliman , Quezon City Philippines*

Abstract. The analysis of disease occurrence over the smallest unit of a region is critical in designing data-driven and targeted intervention plans to reduce health impacts in the population and prevent spread of disease. This study aims to characterize groups of local communities that exhibit the same temporal patterns in dengue occurrence using the Fuzzy C-means (FCM) algorithm for clustering spatiotemporal data and investigate its performance in clustering data on dengue cases aggregated yearly, monthly and weekly. In particular, this study investigates similar patterns of Dengue cases in 129 barangays of Baguio City, Philippines recorded over a period of 9 years. Results have shown that the FCM has promising results in grouping together time series data of *barangays* when using data that is aggregated weekly.

Keywords. data mining, dengue, fuzzy c-means algorithm, time series clustering, disease surveillance

1. Introduction

Health systems around the world are continuously challenged with the emergence and re-emergence of diseases that continue to threaten the lives of many people. In the Philippines, emerging and re-emerging infectious disease (EREIDs) such as leptospirosis, dengue, meningococcemia, tuberculosis among others, continues to threaten the health system.

While the emergence of any disease is unpredictable, its health impact can be managed through planning and execution of proper interventions and proper surveillance [1]. In 2019, the Philippines suffered public health crises due to measles, dengue and polio outbreaks. Dengue cases during this time was noted to be the worst in the decade after recording more than 450,000 cases with about 1,500 deaths [2].

In particular, it is important to consider strengthening health surveillance systems even at the most local communities, commonly called *barangays* in the Philippines, in order to prevent major health crises or disease outbreaks at the national level. This is because health intervention planning is unique to an area or population due to regional differences in terms of available resources, health personnel, environmental conditions, and accessibility [3].

[1] Corresponding Author: Jozelle C. Addawe, Service Science and Software Engineering Laboratory, University of the Philippines Diliman, Quezon City Philippines E-mail: jcaddawe@up.edu.ph

The unpredictability of disease outbreaks, the sporadic growth of dengue, the increase in the number of deaths and strengthening surveillance in localized areas are the primary motivations of this research.

This study explores disease pattern at the barangay level by grouping together barangays that exhibit similar time series patterns and identify shared characteristics of the barangays belonging to the same group using a data clustering based approach. Furthermore, this research investigates how data granularity may affect the findings. This approach in analyzing disease data discovers patterns in data while considering the geographic location of the barangays which might affect disease distribution.

Particularly, this paper investigates patterns of dengue cases recorded in 129 barangays of Baguio City collected over a 9-year period. Baguio City is a highly urbanized city and a tourist spot located in the northern part of the Philippines which was under close monitoring after exceeding the dengue alert threshold following the declaration of the dengue alert in the country in 2019. The city government intensified its dengue preventive drives after observing that more than half of the barangays have recorded cases of mosquito-borne diseases. According to the City Heath Office, as of the end of July 2019, there has been a 29.77% increase in the dengue cases in the city as compared to 2018 with one dengue related death. Despite not yet being declared for an outbreak, government officials are alarmed with the increase in the deaths due to dengue fever [4].

The next sections of the paper presents previous works, the fuzzy C-means (FCM) algorithm used in this paper, the methodology and the results and discussion of the simulations. The conclusions and recommendations of the research are presented thereafter.

2. Previous Work

Dengue is a fast spreading vector-borne disease transmitted by Aedes female mosquitoes. Dengue fever thrives and spreads faster in sub-tropical and tropical regions such as the Philippines due to the mosquitoes' dependence on the ecological conditions present in these areas that support their lifecycle. Survival of mosquitoes highly depend on water, temperature, precipitation, human habitation, vegetation cover. Moreover, dengue fever often occurs with rapid urbanization [5].

Data on number of disease cases collected over time and space is a spatiotemporal type of data wherein each sample in the data is composed of a spatial and the temporal component. In epidemiology, analysis of such data is useful in the exploration of the distribution of disease over a particular region over time. Various researches in diseases use spatial analysis to study disease distribution by means of disease mapping or hotspot analysis [6, 7, 8], or disease modelling to forecast disease cases. However, most studies recommend that other factors be considered to further explain disease distribution, such as socioeconomic factors or environmental factors including cleanliness or terrain [9].

Izakian [10, 11] presented a fuzzy clustering method to group together subsequences of time series data within time windows to reveal patterns in historical data. The Fuzzy C-means is used to cluster spatiotemporal data based on the following objective function:

$$O = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m} d^2(v_i, x_k) \tag{1}$$

where $U = U_{ik} \in [0,1]$ is a fuzzy partition matrix containing the degree of membership of the k^{th} data point to an i^{th} cluster and $m > 1$ is the fuzzification coefficient. The distance between the i^{th} cluster, v_i and the k^{th} data point, x_k, d is computed by the Euclidean distance between these points [10].

$$d_\lambda^2 = (v_i, x_k) = ||v_i(s) - x_k(s)||^2 + \lambda||v_i(t) - x_k(t)||^2 \tag{2}$$

In this equation, the first term and second term is the distance between the spatial components and the temporal components, respectively and $\lambda \geq 0$ is a factor that controls the effect of the temporal component to the overall distance, *d*. In this paper, the spatial component of the data is the location of the barangay, the temporal component is the dengue cases over time and $\lambda = 0.1$ is used in all experiments. Data is assigned to the cluster with the highest degree of membership. The FCM uses the same randomly generated partition matrix for partitioning n barangays into c clusters, regardless of the data granularity. Reconstruction error is used to evaluate the results. This error is the difference between the original data x_k and the reconstructed data, $\hat{x}_k$ obtained from the FCM clustering result computed as [10]:

$$\hat{x}_k = \frac{\sum_{i=1}^{c} u_{ik}^m v_i}{\sum_{i=1}^{c} u_{ik}^m} \tag{3}$$

3. Results and Discussion

The fuzzy c-means clustering was implemented on time series data of dengue case records from 129 barangays of Baguio City obtained from the City Health Office. Results from initial implementations of the FCM using all barangays had low clustering performance which could have been affected by outliers in data, such as those barangays that have very low or zero dengue cases. It was used to cluster all 129 barangays (Case A), top 14 barangays having high dengue cases yearly (Case B) and top 19 barangays with high dengue incidence rates yearly (Case C) in order to extract relevant and information specific to a set of barangays. Dengue incidence rate is the number of dengue cases by the average population of a barangay.

The FCM was performed on time series data of dengue cases aggregated yearly, monthly or weekly for each barangay for arbitrarily chosen number of clusters $c = 3, 4, 5$ and 6.

The reconstruction errors of the FCM on yearly data aggregated weekly for Case B and Case C are summarized in Table 1 where shaded cells correspond to the least reconstruction errors. Observe that the optimal number of clusters of barangays is different each year with very small differences. For simplicity, $c=5$ clusters is chosen for Case B while $c=3$ clusters is used for Case C. The reconstruction errors for Case A is not shown but the optimal number of clusters is $c = 5$ for all years.

From these reconstruction results, the FCM was used to group the barangays in Case A, Case B and Case C using $c=5$, $c=5$ and $c=3$ clusters, respectively. Table 2 summarizes the reconstruction errors of the FCM in clustering the 9-year time series data of dengue cases aggregated weekly, monthly and weekly. Results show that using weekly aggregated data result to a lower reconstruction error although this is not the case for Case B.

Table 1. Reconstruction Errors of the FCM using weekly aggregated data

	Barangays with Top Cases				Barangays with Top Incidence			
YEAR	3 Clusters	4 Clusters	5 Clusters	6 Clusters	3 Clusters	4 Clusters	5 Clusters	6 Clusters
2010	10.030	9.818	9.727	9.936	10.118	10.118	10.118	10.118
2011	11.056	11.044	10.972	10.922	10.381	10.382	10.381	10.381
2012	9.968	9.968	9.926	9.876	12.086	12.086	12.086	12.086
2013	12.850	12.741	12.381	12.261	14.504	14.555	14.594	14.543
2014	11.278	11.192	11.050	11.182	10.527	10.542	10.542	10.542
2015	12.595	12.467	12.054	12.482	13.332	13.329	13.367	13.316
2016	12.670	12.510	12.434	12.516	16.833	16.923	16.989	16.917
2017	12.306	12.251	12.191	12.186	9.254	9.254	9.254	9.254
2018	11.091	11.086	10.976	11.037	9.704	9.705	9.706	9.706

Table 2. Reconstruction Errors using data aggregated yearly, monthly and weekly

	Yearly Aggregated	Monthly Aggregated	Weekly Aggregated
All Barangays (Case A)	136.160	132.72	43.868
Top Barangays (Case B)	11.580	11.612	12.270
Top Incidence Barangays (Case C)	19.959	20.013	16.791

Table 3. Hamming Distance between cluster assignments of barangays

	Case A	Case B	Case C
Yearly vs. Monthly	$D = 20$	(33333232221122) vs. (33333232221122) $D = 0$	(0011021010100011110) vs. (0011001010100011110) $D = 1$
Yearly vs. Weekly	$D = 25$	(33333232221122)vs. (33333232221122) $D = 0$	(0011021010100011110) vs. (0212002010100011010) $D = 5$
Weekly vs. Monthly	$D = 25$	$D = 0$	$D = 5$

Apart from the reconstruction error, a comparison of the cluster assignments of barangays obtained from the FCM using different data aggregations to tell more about the performance of the clustering approach. This was done by comparing the Hamming distances (D) between the cluster assignment of the barangays using different data aggregations. Results are summarized in Table 3, where the sequence of numbers enclosed in parentheses are the cluster number assignments of each barangay considered. Results show that the FCM results using yearly and monthly aggregated data generally partitions the barangays in the same clusters with very few differences. Furthermore, the FCM assigns Case B barangays to the same clusters when using different data aggregations but generally, a lower reconstruction error is obtained when using weekly aggregated data.

The FCM was tested in clustering barangays in Case B and Case C for different number of clusters and lambda values. Lower reconstruction errors were obtained with input c = 6 clusters and λ= 0.04 for Case B and c = 3 and λ = 0.04 for Case C.

The visualizations of the time series plots show similarities in the pattern of dengue cases. Figure 1 shows that barangays in Cluster 3 have relatively higher cases. Figure 2 shows that high incidence barangays belong to Cluster 0 while those with relatively the lowest dengue incidence belong to Cluster 1. Both results however show that there are a few barangays that are outliers of a cluster.

The barangays per cluster are mapped in Figure 3 to show the locations of these barangays within the city. Observe that barangays belonging to the same cluster

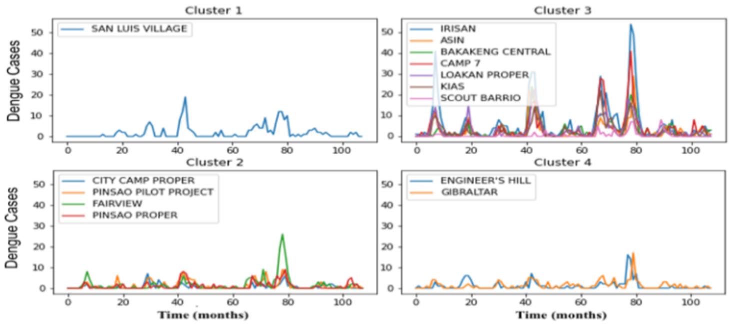


Figure 1. Cluster Assignments of Barangays with Top Dengue Cases

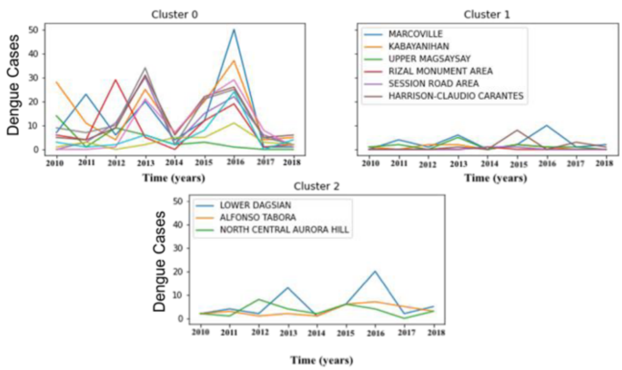


Figure 2. Cluster Assignments of Barangays with Top Dengue Incidence Rates

generally share the same boundaries. The FCM suggests that the barangays with top cases belonging to Cluster 3 are found in the western and southeastern part of the region. Barangays with top cases yearly relatively have high land area while those with high incidence areas are clustered closer to the city center with relatively smaller land areas. Moreover, top incidence barangays belonging to Cluster 1 are nearer to the central district while those in cluster 0 seem to be just around the center as shown in the map.

4. Conclusions and Future Works

This paper has shown the performance of an existing fuzzy c-means based approach in grouping together barangays that exhibit similar patterns of dengue cases. Results have shown that weekly aggregated data distinguishes similarities in time series data with lower reconstruction error. Further studies following the current findings may be to discover how cluster memberships of barangays change over the years or per quarter to observe significant disease pattern changes during a certain period. Necessary adjustments such as data transformation and proper scaling of the spatial and temporal components of the data is recommended.

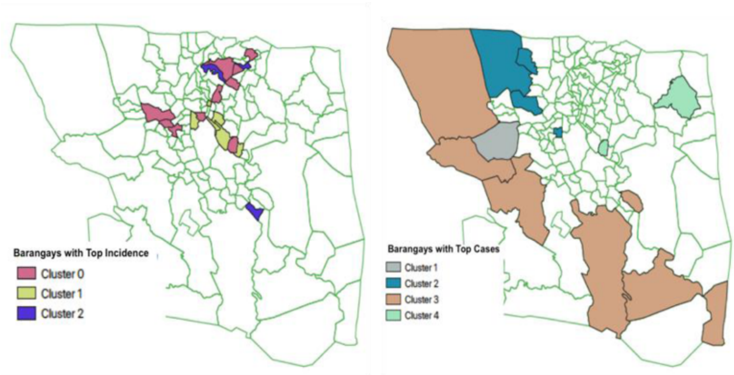


Figure 3. Cluster Assignments of Barangays

Acknowledgements

The authors would like to acknowledge the DOST - ERDT for funding the publication of this paper. The authors also thank the Baguio City Epidemiology and Surveillance Unit for the raw data on dengue cases.

References

[1] (WHO) WHO. Public Health Surveillance;. Available from: https://www.who.int/immunization/monitoring_surveillance/burden/vpd/en/.
[2] Santos AP. Philippines struggling to cope with back-to-back disease outbreaks; 2019. Available from: https://www.dw.com/en/philippines-struggling-to-cope-with-back-to-back-disease-outbreaks/a-51084152.
[3] Esparagoza C. Challenges in the Philippine Healthcare System: Social Determinants to health, Health system strengthening and Health engagement in development towards issues in equity. 2020 05.
[4] Rizavel C Addawe KSC Louie Ville A Balino. Spatio-temporal pattern distribution of dengue infections in baguio city. In: A Preprint; 2019. .
[5] Walsh M. Dengue Part 2: The Mosquito and its ecology'; 2011. Available from: http://www.infectionlandscapes.org/2011/01/dengue-part-2-mosquito-and-its-ecology.html.
[6] Pangilinan MAP, Gonzales DPG, Leong RNF, Co F. Spatial analysis of the distribution of reported dengue incidence in the National Capital Region, Philippines. Acta medica Philippina. 2017;51(2):126-32.
[7] Mutheneni SR, Mopuri R, Naish S, Gunti D, Upadhyayula SM. Spatial distribution and cluster analysis of dengue using self-organizing maps in Andhra Pradesh, India, 2011–2013. Parasite epidemiology and control. 2018;3(1):52-61.
[8] Ansari MY, Ahmad A, Khan SS, Bhushan G, et al. Spatiotemporal clustering: a review. Artificial Intelligence Review. 2019:1-43.
[9] Datoc HI, Caparas R, Caro J. Forecasting and data visualization of dengue spread in the Philippine Visayas island group. In: 2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA). IEEE; 2016. p. 1-4.
[10] Izakian H, Pedrycz W, Jamal I. Clustering spatiotemporal data: An augmented fuzzy c-means. IEEE transactions on fuzzy systems. 2012;21(5):855-68.
[11] Izakian H, Pedrycz W. Anomaly detection and characterization in spatial time series data: A cluster-centric approach. IEEE Transactions on Fuzzy Systems. 2014;22(6):1612-24.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210092

Application Programming Interface for a Customer Experience Analysis Tool

George KOPSIAFTIS [a,1], Ioannis GEORGOULAS [a], Ioannis RALLIS [a]
Ioannis MARKOULIDAKIS [a], Kostis TZANETTIS [b], Michael SFAKIANOS [b] and Nikolaos DOULAMIS [a]

[a] *School of Rural and Surveying Engineering, National Technical University of Athens*
[b] *AppArt S.A.*

Abstract. This paper analyzes the architecture of an application programming interface (API) developed for a novel customer experience tool. The CX tool aims to monitor the customer satisfaction, based on several experience attributes and metrics, such as the Net Promoter Score. The API aims to create an efficient and user-friendly environment, which allow users to utilize all the available features of the customer experience system, including the exploitation of state-of-the-art machine learning algorithms, the analysis of the data and the graphical representation of the results.

Keywords. Customer Experience, Net Promoter Score, Machine Learning

1. Introduction

Customer Experience (CX) has been recognized over the years as a key factor for the commercial success of many sectors, including telecommunication operators, banks, tourist enterprises, retail stores, etc. The cornerstones of a CX based strategy are the CX metrics, such as the Net Promoter Score (NPS) which expresses the customer's likelihood to recommend the company's product and/or services [1]. Based on an extended data set from the telecommunication sector, several state-of-the-art Machine Learning (ML) where employed in order to clarify the relation between the NPS and the most significant CX attributes [2], [3]. The ML techniques, along with several graphical representation features were incorporated in an integrated application dedicated to the analysis of the NPS index.

2. The architecture of the customer experience API tool

The CX API tool has been developed in Python, using the Flask [4] and FastAPI [5] frameworks. Figure 1 is a schematic representation of the CX system architecture. The end-users are able to communicate with the API using widely used web service protocols and standards. Specifically, a json-based [6] request/respond process was adopted,

[1] Corresponding Author; E-mail: gkopsiaf@survey.ntua.gr

according to the following steps: 1) the end-user selects the CX features he wants to analyze and the corresponding parameters through the user interface (UI) of the application, 2) an input.json file is created in accordance with the user preferences, 3) the CX tool performs the calculations and creates an output.json file, which contains the results and the metadata of the analysis and 4) the results are displayed using the user-friendly environment of the applications' front-end.

The current version of the CX API mainly focuses on the telecommunication industry, providing capabilities for detailed analytics for several crucial features of the sector, such as product attributes, touchpoints, website, call center, etc. Figures 2 and 3 present indicative screenshots of the application's UI. During the following steps of its development, the API will be further extended to the touristic sector, and specifically the analysis of the hotel customer satisfaction. It should be noted that the API conforms to the design principles of the Representational State Transfer (REST) architectural style and, therefore, it is flexible enough to adjust to a plethora of topics with minor alterations.

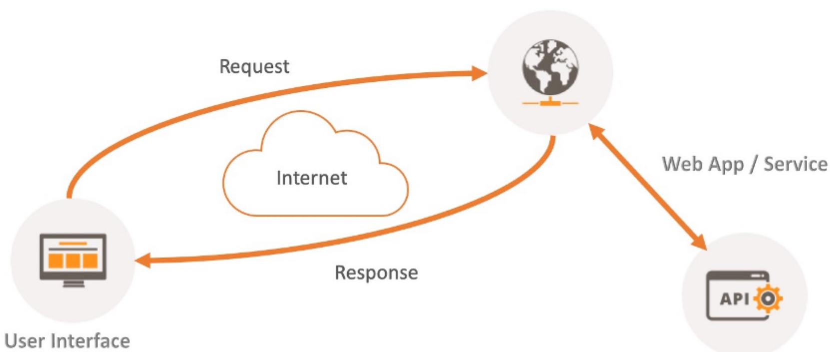


Figure 1. Schematic representation of the CX API tool architecture.

3. Machine Learning Module

One of the main objectives of the companies that use CX metrics to form and optimize their commercial strategies is to maximize the NPS through the improvement of the most crucial CX attributes. A major obstacle towards this goal is the lack of an accurate association between the NPS and the CX attribute scores, which, in turns, could be partly attributed to the lack of sufficient data. In order to tackle this deficiency, during an extended research stage, which preceded the development phase of the API, an effort was made to associate the NPS and the CX attribute scores using stat-of-the-art Machine Learning (ML) techniques and specifically the following [2]: 1) Artificial Neural Networks, 2) Convolutional Neural Networks, 3) Support Vector Machines, 3) Linear Regression, 4) Random Forest, 5) Decision Trees, 6) Naive Bayes and 7) K-Nearest Neighbors. All these techniques were calibrated using a relatively limited initial dataset, which

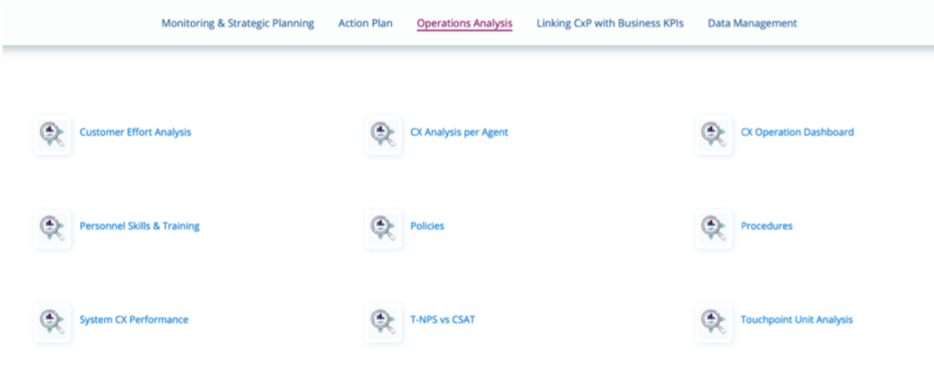


Figure 2. Screenshot of the CX API, presenting the features which are available for analysis for the telecommunication sector.

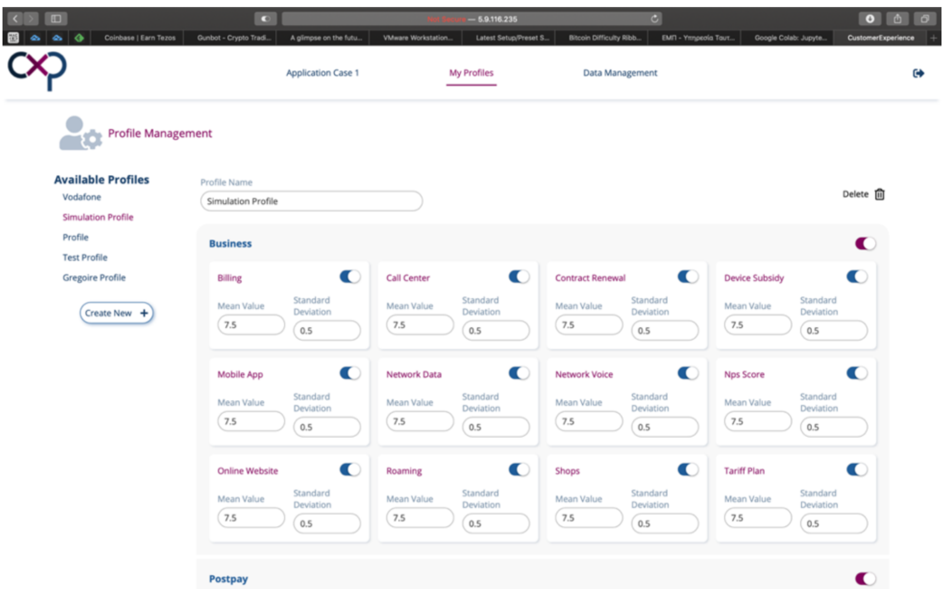


Figure 3. Indicative screenshots of the CX UI, depicting the available options in the Operations Analysis section of the API.

was based on true reviews from telecommunication companies. During the preparatory phase of the API the aforementioned ML techniques were incorporated in a ML module, which was integrated in the final version of the API. The ML module proved to be quite important for the API, as it allowed the formulation of several Use Case Scenarios and the full testing of the API's functionalities. The novelty of the overall CX system could be established on the utilization of advanced ML algorithms to tackle the major obstacles in the evaluation of the CX metrics: 1) the lack of sufficient data and 2) the fact that the CX metric scores depend not only on the CX attributes, but also on other unidentified attributes, such as perception or emotional related attributes. As shown in previous studies [2] the use of the ML module significantly improved most of the examined statistical metrics. Further investigation will be performed in the following steps of the CX system development to test the ability of the proposed methods to capture the relation between the CX attributes and the NPS indicator.

4. Conclusions

This paper briefly describes a CX application, which has the form of an integrated system, including a novel CX analysis tool, as well as a user-friendly UI. The application is currently at the last stages of its development and a full functional version will be soon available for both research and commercial purposes.

5. Acknowledgment

This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code: T1EDK-05063).

References

[1] F.F. Reichheld, The one number you need to grow, Harvard business review, 81(12), 46–55, 2003.
[2] I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis and N. Doulamis, A Machine Learning Based Classification Method for Customer Experience Survey Analysis, *Technologies* 8, 76, https://doi.org/10.3390/technologies8040076, 2020.
[3] I. Rallis, I. Markoulidakis, I. Georgoulas, and G. Kopsiaftis, G., A novel classification method for customer experience survey analysis, In Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments (pp. 1-9), 2020.
[4] Flask, 2018, [Online]. Available: https://flask.palletsprojects.com/en/1.1.x/
[5] FastAPI, 2021, [Online], Available: https://fastapi.tiangolo.com/
[6] JSON, 2013, [Online], Available: https://www.json.org/json-en.html

Health and Environment

This page intentionally left blank

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210094

An Introduction to the euPOLIS Project

Charalampos Zafeiropoulos [a] Matthaios Bimpas [a] Eftychios Protopapadakis [a] Emmanuel Sardis [a] Nikolaos Doulamis [a] Anastasios Doulamis [a] Cedo Maksimovic [b] Stanislava Boskovic [b] Ranko Bozovic [d] and Maja Lalic [e]

[a] *National Technological University of Athens, {mpampiszafeiropoulos, eftprot,sardism}@mail.ntua.gr, manthos.bibas@iccs.gr, {adoulam, ndoulam}@cs.ntua.gr*
[b] *Imperial College London,{c.maksimovic, s.boskovic }@imperial.ac.uk*
[c] *Faculty of Civil Engineering, anja@hikom.grf.bg.ac.yu*
[d] *EnPlus, ranko@enplustech.com*
[e] *Mikser Udruzenje, maja@mikser.rs*

Abstract. This paper introduces euPOLIS; an EU funded project, which emphasize on the appropriate development of urban ecosystems in a way that enhance Public Health (PH) and Well-Being (WB) without significant Life-Cycle costs. Such an approach has has the potential to regenerate urban ecosystems addressing multiple challenges, such as low environmental quality, fragmentation and low biodiversity in public spaces, water-stressed resources, undervalued use of space in deprived areas resulting in an improved urban livability. The proposed methodology is expected to improve people's quality of life, providing them with pleasant socializing open areas that stimulate social exchange while monitoring the impact of all those interventions to PH and WB of citizens. The euPOLIS suggested solutions will be demonstrated in 4 European cities: Belgrade, Lodz, Piraeus and Gladsaxe.

Keywords. Well-being, Public Health, Nature Based Solutions, Sustainability

1. Introduction

Urbanization, i.e. a population shift from rural to urban areas, dates back to 1800s. There are many reasons motivating such movements, including better education, housing, access to work, and less time and expense of commuting and transportation. Nevertheless, the continuously increasing trend of urbanization resulted in harmful social phenomena such as alienation, increased cost of living, and mass marginalization. A direct adverse effect is the deterioration of peoples' health, mental and physical.

On the one hand, Urbanization affects mental health through the influence of increased stressors and factors such as overcrowded and polluted environment, high levels of violence, and reduced social support [1]. On the other hand, urbanization is associated with increased risk of asthma due to the exposure to air pollutants such as nitrogen dioxide (NO2), carbon monoxide (CO) [2]. Both cases require for actions/solutions that are inspired and supported by nature, which are cost-effective, simultaneously provide environmental, social and economic benefits and help build resilience [3].

Nature-based solutions (NBS) is exactly what is needed to tackle most of the problems raised by increasing urbanization trends. One example of NBS is by protecting and/or restoring forests and wetlands. This way, water supplies, protection from floods, soil erosion and landslides can be secured and succeeded. Another example of NBS, is the protection of coastal habitants which will result a mitigation of climate change.

The most important NBS is succeeded by bringing nature into cities by planting trees in our gardens-balconies. The NBS will limit the danger of heatwaves and air pollution while, simultaneously, will have positive contribution on our mental and physical health. Generally, being close to greenery make us feel more ease and reduces our stress levels. Plants and their leaves tend to absorb and reflect noise, which makes the environment more comfortable to visitors, residents or workers. Moreover, trees and plants can also be used as a decor for squares and most abandoned buildings by adding color, ambiance, personality and life to a bland space.

2. Related Work

Protection of the environment, sustainable management of natural resources, water, biodiversity and ecosystems are research areas of great interest; Their interactions with social systems will play a significant role in sustaining the economy and human well-being. The development of integrated approaches to address water-related challenges and the transition to sustainable management and use of water resources and services is currently addressed by multiple organizations and joint actions.

IN-HABIT project [4] proposes a systemic urban planning framework based on innovative gender and diversity approaches to improve Inclusive Health and Well-being. This system has been applied in four peripheral small and medium size cities – Cordoba (Spain), Riga (Latvia), Lucca (Italy) and Nitra (Slovakia). Project VARCITIES [5] aims to implement innovated and visionary ideas by establishing sustainable models for increasing Health and Well-being of citizens. The purpose of VARCITIES is to create an operation model that will convert the cities of the future into human-centered cities. The GO GREEN project [6] proposes urban design based on NBS with the goal of fostering a positive human-nature relationship, flourishing nature connectedness and promoting citizen engagement through digital, educational and behavioural innovation. The purpose of GO GREEN is to position European cities as world ambassadors of urban sustainability.

Another project, similar to the above, is NICE [7]. The overall objective of NICE is to enhance existing NBS to provide circular urban water solutions. NICE will propose solutions which will produce reusable water for different purposes, in addition to mitigating pollution and runoff and composing an attractive and integral part of the urban landscape. Another ongoing project proGIreg [8], which proposes NBS on three cities: Dortmund (Germany), Turin (Italy), Zagreb (Hungary). The proposed NBS will be co-designed, co-created and co-implemented with local communities and on the economic level with market-ready business models. .

3. The euPOLIS Project

EuPOLIS' approach supports the development of urban ecosystems in order to enhance Public Health (PH) and Well-Being (WB) without adding significant Life-Cycle costs. Such an approach has the potential to activate the hidden possibilities and services of existing Natural and Engineered urban systems. During the process of integration, their joint environmental, social, cultural and economic effects will be defined, as a main vehicle for investments in advanced ESS and regenerate urban ecosystems.

All the above can be achieved, while facing challenges such as low environmental quality, fragmentation and low biodiversity in public spaces, water-stressed resources, undervalued use of space in deprived areas and therefore urban livability is improved. EuPOLIS will improve urban resilience via interventions designed using a set of proper urban planning matrices with the stakeholders' participation and with attention to every gender, age and disability perspectives within the process.

Another objective is the creation of inclusive and accessible urban spaces by systematically implementing gender mainstreaming strategies and novel participatory tools into all phases and processes of project development to ensure that the needs of diverse groups are considered. Therefore, an improvement of people's lifestyles is expected, providing them with pleasant socializing open areas that stimulate social exchange. Lastly, the project includes the monitoring and assessment of the impact of all BG \NBS interventions mainly regarding PH and WB of citizens, but also other factors, such as environmental and socioeconomical impacts, resilience-related (with emphasis on CC) results and long-term policy changes, etc.

3.1. *Strategic Objectives*

There are seven Scientific and Technical Objectives (STOs) in this project. The first STO refers to the systematic implementation of an NBS-based urban planning methodology, which will be enriched with cultural, economic, and societal aspects, based on the Blue Green Solutions (BGS). For instance, consider the creation of spatial and functional conditions that will enhance not only the PH and WB of citizens, but also the urban metabolism, the social cohesion and the resilience of cities to Climate Change (CC) and natural disasters.

The second STO proposes an intervention-aimed livability model, rooted in community needs and engaging community diverse potentials. This can be achieved through a three-pronged approach: 1) ensuring the relevance of interventions to specific communities' needs and preferences; 2) reflecting residents/users' voice in the planning process; 3) consolidating the sustainability of project outcomes by fostering a sense of ownership among citizens/users.

The third STO refers to the design and implementation of an online platform, to support and enhance participatory processes, through active socio-cultural hubs. This will re-design public spaces, while promoting new governance and new financing models. The fourth STO refers to the design and implement of customized spatial solutions for each case study (Front Runner -FR- cities of Belgrade, Lodz, Piraeus and Gladsaxe) and then monitor and assess their impact regarding PH and WB, as well as social and environmental results. A mixed-method approach will be developed via traditional questionnaires, ethnographies and interviews with quantitative data collected through wearable devices, behavioral games and mobile questionnaires. Climate hazards such as microclimate, biodiversity, pollution, Urban Heat Island (UHI) will be monitored via remote sensing and sensor networks.

The fifth STO includes the replication and demonstration for corresponding advantages of euPOLIS innovations via mentoring and coaching of the follower cities of the

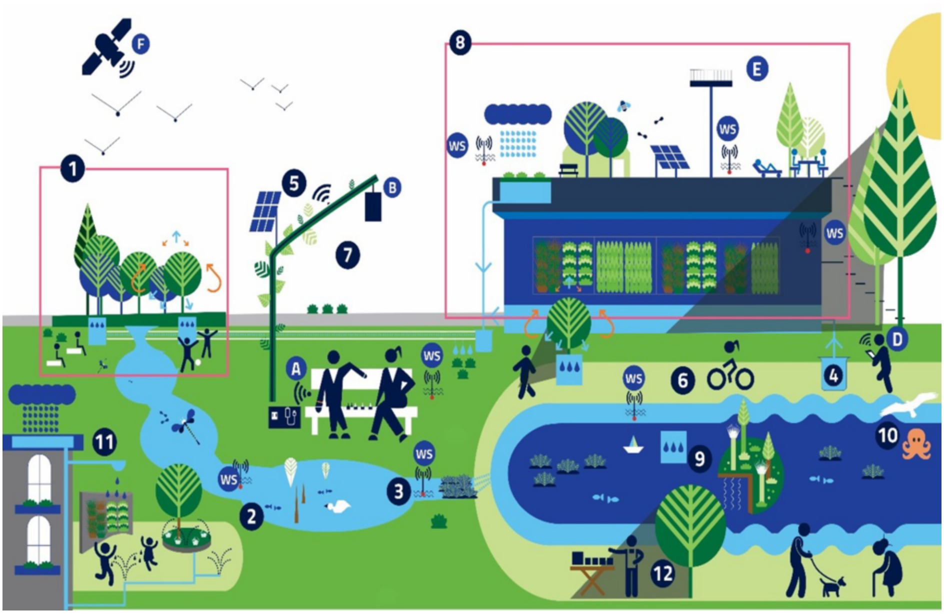


Figure 1. euPOLIS overall concept. 1. NBS-based MF pocket parks accessed by NBS locally conditioned pathways and shared spaces (1.1-1.9), 2. Waterway with mini biotope nodes, aquatic biodiversity – fead from groundwater aquifer or purified surface runoff, 3. NBS for surface runoff quality and pluvial flood management, 4. Groundwater abstraction for water, energy, greenery nexus, 5. MF NBS canopy for socializing, "recharging electronics", or "green bus stop" etc., 6. MF Live vegetation shaded waterfront promenade, 7. Air pollution abatement shrubs, trees and vertical green curtains, 8. Metabolic hub with MF ecotechnology demonstration/promotion, roof garden and art/culutural performance, 9. MF floating island, river water purification, 10. Coastal sea bottom marine aquatic biotope with euPOLIS-NBS, 11. MF euPOLIS Urban square/streetscape and other NBS (biotopes, sensory garden, waterfall, biodiversity kitchen garden for socialising, recreation), 12. Space for NBS business activation and promotion Monitoring- ICT System: A. Wearable devices for monitoring PH WB, B. Visualisation equpment, C. Renewable energy sources, D. Citizens observatories, E. Sensor network, F. Remote sensing, WS. Micro climate / wireless weather station.

project, i.e. Palermo, Limassol, Trebinje and Bogotá. The sixth STO includes the creation of long-term data platforms securing open, consistent data points about the impacts of the deployed approaches, mainly regarding PH and WB, and ensure interoperability with other relevant data infrastructures for effective public consultation, exchange and sharing of practices/experiences.

The seventh STO focus on the identification of sound business models, replicable to other markets and develop new resources activation techniques. This will support a gradual transition to a leading market position for the euPOLIS paradigm. One of the euPOLIS innovative planning tools is Business Activation Matrix. Communication and dissemination of project results, to various audiences through targeted activities, clustering with related projects and initiatives, aiming to spread the euPOLIS concept, is the last STO.

4. Action Plan

The euPOLIS is strongly depended on the deployment of NBS. The process is based on a two-fold approach. The first approach is the systematic activation of Ecosystem Services (ESS) and the use of their Multifunctional (MF) components to interact with the project's urban elements. EuPOLIS will assess ESS and their services, for each site, with the use of an analytical framework developed by the Working Group. Since ESS Provisioning functions (provision of clean air, crops, food, raw materials, etc.) have been dramatically degenerated, first action will be to identify present state of the Ecosystem and then introduce specific regenerative measures as primary planning criteria, in order to revitalize basic function - provisioning.

As for ESS Regulating functions such as microclimate, air quality, etc, euPOLIS will introduce mapping of existing regulating potentials and enhance systematically its impact with adequate urban components as greenery, water, etc. Lastly the Socio-Cultural ESS such as interaction and recreation facilities for mental and physical health or slow tourism, represent an area that is not included in the standard planning criteria and will be crucial for all planned interventions. Supporting services (habitat supporting, maintenance of genetic diversity, photosynthesis, nutrient cycling, etc.) will be treated as part of the underlying structures, processes and functions of urban Ecosystem. EuPOLIS will introduce a methodology for assessment and will use a matrix system to identify existing and potential NBS to enhance important functions, as shown in fig. 2.

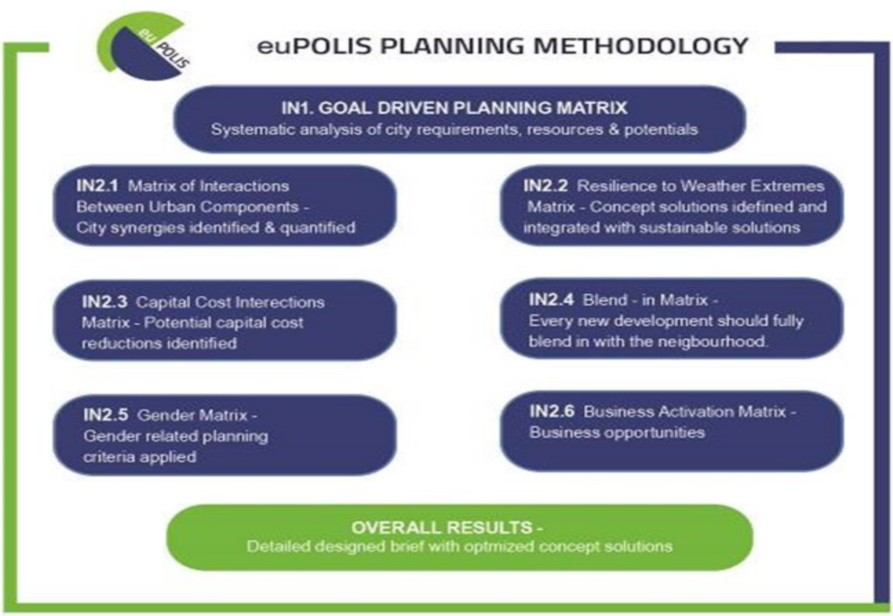


Figure 2. euPOLIS planning methodology

EuPOLIS social goal is to improve the relationship between citizens and the nearby places. Measuring the feelings triggered over citizens by a place, can be used as an indicator of Well Being, since many people tend to stay in a place that provides them safety and feeling of rootedness. Different dimensions of sense of place play a different role in establishing positive relation with place and will be therefore included in our research approach. One of the biggest challenges will be the lack of participation from the citizens. Synergies among euPOLIS' solutions, driven by citizens' initiatives, could generate a vast panorama of excellent practices, crystallized into our proposed methodology.

In case that technological tools are not attractive enough, some traditional methods will be implemented, such as public consultations and community workshops and introducing community card game to enable decision making as well as drama and live role-playing forms for releasing creativity and empathy among potential users. EuPOLIS will, also, employ social- cultural platform to engage people and enhance their knowledge, awareness and willingness for the visionary urban design. Connection with external digital libraries will also be supported, for further growth of its virtual repository. In addition, the project includes the development of a new, or support existing active, socio-cultural hubs in of the four cities. These nodes will communicate euPOLIS messages, carry the project's results and secure ownership and long-term commitment of local and creative communities, citizens including sensitive groups and general public to the euPOLIS participative process.

Data will be obtained through sensor networks and remote sensing. During days the most relevant weather conditions, ad-hoc Intense Observation Periods (IOP's) will be applied while permanent monitoring with continuous data acquisition, processing and storage will be performed all the other days. Monitoring data will be used for knowledge generation, model calibration, NBS evaluation, interactive modelling result visualisation and web-based dissemination/communication. Next, from the knowledge coming from employed platforms, euPOLIS expect to use a people-centred observation web to allow citizens to become the 'eyes' of the city authorities. The goal is to complement the monitoring of the carried-out interventions (e.g., satellite and remote sensing that are less cost-efficient and less dynamic) by using users' smartphones and tablets. Citizens, city authorities, policy makers, psychologists, sociologists and communication experts will be engaged during the adaptation of the euPOLIS toolbox to ensure it is practical and user-friendly, while respecting all relevant privacy issues.

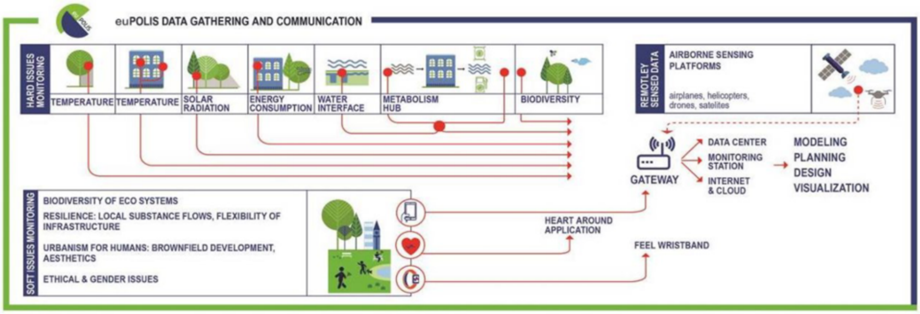


Figure 3. euPOLIS data gathering

Last but not least, data collection, as shown in fig 3, will involve a group of voluntary participants. Firstly, euPOLIS will provide them with feel wristband, which will be used for monitoring of wearer's bio-signals and the monitoring and assessment of the emotional status and stress/anxiety levels. Specific biological indicators will be sought for, and will be integrated with the urban planning methodology, while the different algorithmic components will be calibrated and validated for the use cases.The impact and value of the NBS interventions on people's stress/anxiety levels will be assessed both in the long run, by quantifying the pre- and post-implementation anxiety/stress levels of the citizens as well as in near-real time. Activity monitoring, video communication, analysis using effective computing services and integration of Serious Games will support provisioning, regulation and sociocultural ESS. Through the euPOLIS mobile apps, citizens will be informed about the effectiveness and the impact of NBS on Public Health and Well Being.

4.1. Design and implementations in the four cities

This section includes various insights regarding the upcoming actions on the four front-runner cities. This section will also focus on all health/environmental/social issues and the demonstration activities, of each city.

4.1.1. City of Belgrade

The city of Belgrade has (approximately) 200 km of riverfront length and population 1,374,000 people. Only the 25% of the population have access to water, greenery, and public spaces. EuPOLIS' interventions will be applied on "Linear park" and "Ušće", which is the biggest urban park and is located in new Belgrade. Currently, the lack of necessary infrastructures result in air/water/soil pollution, noise, and lowers PH and WB. Unemployment and an aggressive behaviour, related to stress, are also reported.

Eupolis will address the lack of proper NBS, which jeopardize PH. Also, in "Linear Park", project plans involve the creation of a visual and functional path between city main body and river. That way the ecological corridors will be connected with neighbouring greenery spots, including nodes related to NBS-based environmental education and local business activation. In Ušće park, an integrated constructed wetland will be proposed, with several bio filters of different types for storm water treatment. The project will also recommend NB connections with Sava river and multifunctional floating island with practical demonstration of natural river water purification, also used by visitors to be entertained and experience additional views on Belgrade panorama and learn of new technologies.

4.1.2. City of Lodz

The city of Lodz, with 696,708 inhabitants, includes a historical city center, inhabited by 152,292 people. In Lodz there are many buildings in poor technical condition, and a few green and attractive public spaces. Lodz is considered as low environmental quality city due to high air pollution in the urbanized center. In addition to low air quality, it also lacks a systemic approach to water management to mitigate pluvial floods, resulting in risk for raw sewerage overflows. There are no NBS to address drought and heat wave problems. Therefore, multiple MF-NBS interventions will be considered. Those NBS will include two multifunctional parks with green spaces, an air pollution abatement/mitigating greenery through shrubs and vertical green curtains and environmentally friendly corridors to create quality access to the historical city center.

4.1.3. City of Piraeus

The city of Piraeus, in which you largest port in Greece is located, is one of the most densely populated cities in Europe. EuPOLIS will intervene in three mutually interlinked neighboring sites at the main harbor area (Mikrolimano). These sites are Seaside Promenade Mikrolimano area, the riverine inland area in Akti Dilaveri,and the Ralio Complex Pilot School (RCPS). Different characteristics such as visibility, stage of development and need for interventions, apply for each site.

Piraeus, also, lacks open green spaces, parks and there is not enough shading for protection against solar radiation and other UHI effects. Moreover, the city has old high-density buildings and narrow streets with limited pedestrian free space and high levels of air pollution. In addition, the riverine inland area in Akti Dilaveri is connected to the promenade, which is underdeveloped with polluted water. The social profile of Piraeus has been affected from high unemployment rate, increased economically non-active population lack of SME business opportunities and substantial "people friendly" welfare structures.

During this project, the restoration of the coastal pedestrian promenade of Mikrolimano will include implement of acceptable Mediterranean types of pollution mitigating vegetation in open green spaces for the improvement of micro climate and reduction of air pollution. The project will also propose NBS public places conducive to small business for new job creation opportunities. Eupolis will also make two more interventions. First intervention refers to the extension of the coastal promenade to the riverine area closer to the Sailing Boat Marina and Athletic Area in the Akti Dilaveri which will be used as a demo site for training of local planners with full support of euPOLIS team. Live vegetation covered multifunctional-NBS canopies encompassing info-points will be established in the area that "merges" the running Promenade with Akti Dilaveri. This will include places, where people can be protected from direct solar exposure, socialize, sit, wait, relax, recharge batteries (electronic), socializing, enjoy the freshness of scent gardens, etc. Second intervention refers to coastal blue green sea bottom marine aquatic biotope with euPOLIS -NBS. This way city-ports have succeeded to make many former heavily polluted/ex harbor underwater areas attractive, clean and eco-friendly. The site in Piraeus will be the first international case, where the synergy between terrestrial and aquatic meet and interact. As for the area of RCPS, a Pocket Park in the wide sidewalk will be introduced and also an educational-eco center for promotion of NBS will be established.

4.1.4. Municipality of Gladsaxe

Last city is the Municipality of Gladsaxe, which is located in the north western suburb of Copenhagen. Previous analyses showed that the Gladsaxe Municipality represents a city district with a negative social development. Our attention is focused on the housing development Pilenparken which hosts a total of 1700 inhabitants.Main issues of the selected site are the poor environmental quality, the high groundwater level which exposes the local citizens to health risk during heavy rainfall, the backflow from the sewage system that occurs with high risk of infection caused by raw sewerage flooding. Generally, euPOLIS will propose NBS at this site that will rely on creative development of new solutions for urban revitalization based on facilitation of intensified evaporation to reuse locally harvested rainwater which in parallel support the other NBS. This will secure the unfolding of the recreational potential of the site and reduce the load on the sewer system thus prevent its harmful effects on public health during heavy rainfall events. In particular, NBS include natural cleaning of storm water to be used for other purposes. Moreover, euPOLIS will support the creation of evaporation playing field (underground water storage moistening grass and evaporating storm water), green sponge hill (a rolling landscape of artificially grown media to act as a water buffer magazine), an evaporation reactor, a water filtration unit, and pavement irrigation (sprinkling irrigation of paved areas during hot weather for speeding evaporation process, surface cooling and kids playing grounds).

These NBS will access pathways to make stronger unity in the community by making better connectivity between the facilities of the area. As for environmental awareness, water square playground (grove and water mist, art and culture interface, scent gardens) will create a framework for play and mediation for people in all ages, to create awareness of water and environment through play.

5. Measuring the Impact

There are 10 Key Performance Indicators (KPI) which will measure the impact of the proposed solution. First KPI is the psychological and physiological responses, psycho-emotional states. There will be optimization of relevant psychophysiological parameters among users of re-designed public space, including the reduction of stress, depression and anxiety levels. Second KPI includes health indicators related to physical activities such as walking, running, cycling, skateboarding, or activities related to an intervention such as running in the new park, strolling along the new pedestrian street. Third KPI includes health indicators related to improvements of local conditions. Forth KPI refers to the enhancement of social cohesion and cultural particularity through ensuring sense of security and inclusion for all (focusing on gender and age equality) allowing for the strengthening of exploratory and socializing/culture behaviors among users. Fifth KPI defines the sense of place and place attachment among users. There will be data from studies showing an increased positive emotional attachment to the neighborhood as well as re-designed public space. Sixth KPI will measure the density and strength of local community ties. Seventh KPI will include the number of planned natural systems. In particular, the improvements of local conditions will be quantified by implemented NBS. Eighth KPI will measure the improvement of habitat, biodiversity, resilience, EcoSystems (ES) in case studies. Ninth KPI will present a list of implemented business models. Last KPI will present the deployed communication activities. This includes all cities involved through technology adoption, and people involved in participatory processes.

6. Conclusion

This work provided a short introduction to euPOLIS; an EU funded project, which emphasizes on the appropriate development of urban ecosystems in a way that enhance PH and WB, without imposing significant Life-Cycle costs. Such an approach has has the potential to regenerate urban ecosystems addressing multiple challenges, such as low environmental quality, fragmentation and low biodiversity in public spaces, water-stressed resources, undervalued use of space in deprived areas and therefore urban livability is improved. The proposed methodology is expected to improve people's quality of life, providing them with pleasant socializing open areas that stimulate social exchange while monitoring the impact of all those interventions to PH and WB of citizens. The euPOLIS suggested solutions will be demonstrated in 4 European cities: Belgrade, Lodz, Piraeus and Gladsaxe.

Acknowledgement

This paper is supported by the European Union Funded project euPOLIS "Integrated NBS-based Urban Planning Methodology for Enhancing the Health and Well-being of Citizens: the euPOLIS Approach" under the Horizon 2020 program H2020-EU.3.5.2., grant agreement No 869448.

References

[1] Szabo, Christopher Paul: "Urbanization and mental health: a developing world perspective" *Current Opinion in Psychiatry*, May 2018 - Volume 31 - Issue 3 - p 256-257
[2] A.Alimissis K.PhilippopoulosC.G.TzanisD.Deligiorgi: "Spatial estimation of urban air pollution with the use of artificial neural network models" *Atmospheric Environment*, Volume 191, October 2018, Pages 205-213
[3] The EU and nature-based solutions *https://ec.europa.eu/info/research-and-innovation/research-area/environment/nature-based-solutions$_e$n*, $(May 2021)$
[4] INclusive Health And wellBeing In small and medium size ciTies(IN-HABIT), *https://www.inhabit-h2020.eu*, (May 2021)
[5] Visionary Nature Based Actions For Heath, Wellbeing Resilience In Cities (VARCITIES), *https://www.varcities.eu*, (May 2021)
[6] GO GREEN: Resilient Optimal Urban natural, Technological and Environmental Solutions, *http://gogreenproject.eu*, (May 2021)
[7] Innovative And Enhanced Nature-Based Solutions For Sustainable Urban Water Cycle (NICE),
[8] productive Green Infrastructure for post-industrial urban regeneration (proGIreg), *https://progireg.eu/*, (May 2021)

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210095

Assessing the Lockdown Effects on Air Quality During COVID-19 Era

Ioannis Kavouras [a] Eftychios Protopapadakis [a] Maria Kaselimi [a] Emmanuel Sardis [a] and Nikolaos Doulamis [a]

[a] *National Technical University of Athens, Zografou Campus 9, Iroon Polytechniou str, Athens, 15780, Greece, ikavouras@mail.ntua.gr, eftprot@mail.ntua.gr, mkaselimi@mail.ntua.gr, ndoulam@cs.ntua.gr*

Abstract.
In this work we investigate the short-term variations in air quality emissions, attributed to the prevention measures, applied in different cities, to mitigate the COVID-19 spread. In particular, we emphasize on the concentration effects regarding specific pollutant gases, such as carbon monoxide (CO), ozone (O_3), nitrogen dioxide (NO_2) and sulphur dioxide (SO_2). The assessment of the impact of lockdown on air quality focused on four European Cities (Athens, Gladsaxe, Lodz and Rome). Available data on pollutant factors were obtained using global satellite observations. The level of the employed prevention measures is employed using the Oxford COVID-19 Government Response Tracker. The second part of the analysis employed a variety of machine learning tools, utilized for estimating the concentration of each pollutant, two days ahead. The results showed that a weak to moderate correlation exists between the corresponding measures and the pollutant factors and that it is possible to create models which can predict the behaviour of the pollutant gases under daily human activities.

Keywords. environmental quality, air pollution, air quality levels, neural networks, COVID-19, machine learning, dynamic time warping

1. Introduction

By the end of January 2020, authorities, worldwide, introduced a series of measures, in order to reduce the transmissibility of the COVID-19 virus epidemic [2]. These measures include international travel controls (e.g. travel prohibition between countries), restrictions in internal movements (e.g. restrictions in transportation between municipalities), public event cancellations (e.g. closure of live concerts), restriction gatherings (e.g. limiting the amount of people over an area), closure of public transportation, closure of schools, stay home requirements (e.g. move outside the house only for work, emergency, etc) and closure of workplaces (e.g. closure of industries / commercial sector) [3].

The response measures appeared to have an impact on the Air Quality Levels (AQL) in cities. Indicative examples involve countries such as China [4], India [5], Brazil [6] and Spain [7]. However, there are no clear indication regarding the correlation between each major pollutant factor, i.e. Carbon Monoxide (CO), Ozone (O_3), Nitrogen Dioxide (NO_2) and Sulphur Dioxide (SO_2), and the applied measures. The general trends indicate that applied measures improved the AQL. There are various environmental models, based on simulations [8,9], indicating the causality between specific actions and AQL. Nevertheless, the models have not been sufficiently tested in real life. Thus, the lockdown, worldwide, created a unique opportunity for researching the significance of traffic reduction in AQL changes.

In this work we will investigate the correlation between the applied lockdown measures and the four major pollutant factors (CO, O_3, NO_2 and SO_2) for a total of four European Cities - Athens (Greece), Gladsaxe (Denmark), Lodz (Poland) and Rome (Italy) - during COVID-19 period. In particular, we focus on the identification of each measure and what scale affects each of the pollutant factors. This will be achieved by a variety of statistical correlation methods, where we will test the relation between each measure, at a time, and the density of each pollutant factor, for a yearly period divided by two days average values. The pollutant factor densities will be calculated using the Sentinel-5 TROPOMI satellite images of a two day period, clipped in a squared area of radius 0.25^o around the center of each city.

Also, a variety of machine learning methods will be trained using the most correlated COVID-19 measures and the densities of each one of the most dangerous pollutant gases (CO, O_3, NO_2, SO_2). A different model will be created, for each of the pollutant factors. In total, four models per machine learning methods will be trained. As input we will use a two days period COVID-19 measures and each pollutant's mean density. The output will be each pollutant's mean density, two days ahead.

2. Related Work

The multiple and synchronized applied measures (lockdown) for the minimization of COVID-19 transmissibility, provided the opportunity for an in depth research of the impact of industrial activities closure and traffic volume reduction, in cities, and AQL. During and after the lockdown, in China, researches proposed a reductive trend in CO, O_3, NO_2, SO_2, PM_{10} and $PM_{2.5}$ densities [10,11,12]. Indicative example, consists the Hubei region, in which during the lockdown period in January and February, 2020, the Angstrom Exponent (AE) and the Aerosol Optical Depth (AOD) increased and decreased respectively by 29.4% and 39.2% (only in Wuhan they increased and decreased by 45.3% and 31.0%) [13].

Another, significance example, includes India, in which has been observed reduction, during and after the lockdown, in critical pollutant factors [14,15,16]. A comparison in air pollution changes between India and China, before and after the lockdown, described in the work of [17]. Real-time measured densities of $PM_{2.5}$ and NO_2 are used for a total of 12 megacities in India and China (6 megacities from each country). An Air Quality Index (AQI) is calculated, for each density, before, during and after the lockdown, which resulted to a reduction of $45.25\%(AQI_{PM_{2.5}})$ and $64.65\%(AQI_{NO_2})$ in China and $37.42\%(AQI_{PM_{2.5}})$ and $65.80\%(AQI_{NO_2})$ in India after the lockdown.

Other related works, include, the [18], in which the densities of CO, O_3, NO_2 and SO_2 were predicted one week ahead, while, also, an AQI was calculated, which estimated that in countries, such as India, where the lockdown measures were stricter, the air quality improved, while on the other hand in countries like Australia where the lockdown measures, weren't so strict the air quality followed an incremented trend. In Madrid and

Barcelona (Spain) [19] the concentration of NO_2 reduced 62% and 50%, respectively, after a 75% in traffic volume. In Baghdad (Iraq) [20] the daily densities of NO_2, O_3, $PM_{2.5}$ and PM_{10} used before and during the lockdown, in addition to the AQI for the same time period.

3. Proposed methodology

In this work we focus on two research topics: a) how each of the specific prevention measures is related on each of the pollutant factor levels, and b) to what extent we could estimate (forecast) the trend of each pollutant level, when using as input the current level, plus the prevention measures. On the one hand, the estimation of the correlation between COVID-19 measures and pollutant factors can be evaluated using various case-specific techniques [22]. On the other hand, the estimation of pollutant factors' levels few days ahead, entails to traditional regression problem [21].

Let us denote as $\mathbf{x} = [x_1, ..., x_n]$, a set of predictive factors, involving prevention measures magnitude and current pollution levels, and $\mathbf{y}$ as the investigated pollutant factor. We will try to estimate a predictive function $f(\mathbf{x}) \longrightarrow \hat{\mathbf{y}}$, where $\hat{y}$ stands for the specific pollutant factor's level, two days ahead. Practically, we would like to have values as close as possible to actual ones, i.e. $\hat{\mathbf{y}} \simeq \mathbf{y}$ [23].

The input data, is a combination of: a) the densities of the most harmful pollutant gases (CO, O_3, NO_2 and SO_2), day-by-day period, and b) the magnitude of the applied prevention measure ranging from 0 (no measure applied) to 1 (the strictest form of the measure). Experiments rung in a total of 4 European Cities (Athens, Gladsaxe, Lodz, Rome). The adopted COVID-19 mitigation strategies, considered as predictive factors, are explained in the following lines.

Restrictions of Internal Movement (RE_IN_MOV) refers to the transportation between municipalities. **International Travel Controls** (IN_TR_CON) include any measure which restricts/prohibits the transportation between countries. **Cancellation of Public Events** (CA_PUB_EV) involves the cancellation of all public events, locally (e.g. football matches, concerts) or worldwide (e.g. Olympic Games in Japan, Euro-vision). **Restriction in Gatherings** (RE_GAT) is a general description for all measures that restricts more than a number of people be together, in the same place. **Close Public Transportation** (C_PUB_TRAN) includes the measures taken about the closure of public transportation (e.g. busses, metro). **School Closures** (C_SCHOOL) describe any measure involving the education system (e.g. school closure, distance learning, etc). **Stay at Home Requirement**(STAY_HOME_R) is considered the most intense measure, since it penalized any movement outside home, without a specified reason (e.g. visit to the doctor, pharmacy, provide assistance to relative, etc). **Workplace Closures** (C_WORKPLACE) refer to both the industrial activity and the commercial sector.

The COVID-19 dataset has been downloaded from Our World in Data (`https://ourworldindata.org/policy-responses-covid`) [3], which been updated daily. The densities of the four pollutant factors have been downloaded from the Sentinel-5/TROPOMI dataset of Google Earth Engine [24], in a day-by-day period, with 50m spatial analysis.

3.1. Investigating patterns similarity

As a first step, two statistical methods have been used. The core idea was to identify whether there is high correlation between the investigated pollutant factor and any of the applied prevention measures.These statistical measures are known as a) Pearson correlation and b) Dynamic Time Wrapping (DTW).

Pearson correlation estimates how related two signals are, using the best fit line approach. This method calculates three values, which refers to correlation (r), coefficient of determination (R^2) and hypothesis (p). The r can be any value between -1 and 1 (the sign shows the direction) and absolute value can be translated as:

- 0.00-0.09 = No Correlation
- 0.10-0.39 = Weak Correlation
- 0.40-0.69 = Moderate Correlation
- 0.70-0.89 = Strong Correlation
- 0.90-1.00 = Very Strong Correlation

The R^2 can be any value between [0, 1] and can be multiplied by 100.0 to express the percent of affection between correlated objects. The p can be any value between range [0, 1] representing the probability that this data would have arisen if the null hypothesis were true (**Null hypothesis**: There is no correlation between A and B in the overall population $r = 0$, **Alternative hypothesis**: There is a correlation between A and B in the overall population $r \neq 0$). [25,26]

DTW is a technique used for calculating the alignment between two given (time-dependent) signals, under certain restrictions. The sequences are warped in a non-linear fashion to match each other. DTW, originally, has been used for the comparison of different speech patterns in automatic speech recognition. [27]

3.2. Machine Learning Tools for Extracting Environmental Impact

Various machine approaches have been evaluated. These approaches range from traditional shallow learning techniques like k-nearest neighbors to more complex approaches as deep neural networks. The outcomes indicate that, for the problem at hand, and due to the low feature space dimensional, traditional approaches are sufficient for robust pollutant factors level estimations.

Deep Neural Network (DNN): A DNN method is a feed forward network that consists of multiple hidden layers. For the purpose of this research we used for input a set of predictive values $\mathbf{x} = [x_1, ..., x_n]$, and as output $\mathbf{y} = [y_1, y_2]$. The architecture, also, consists of three hidden layers (L_1, L_2, L_3), with sizes 20, 10 and 20 accordingly. As activation function between the hidden layers, the SELU (Scaled Exponential Linear Units) has been chosen, and for the output layer the sigmoid.

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) capable of learning order dependence in sequence prediction problems. The two technical problems, of RNN, overcome by LSTMs are vanishing gradients and exploding gradients, both related to how the network is trained. Hyperparameter optimization is feasible through Bayesian optimization schemes. [36]

Decision Tree Regressor (DTR): DTR method's core idea lies in sub-dividing the space into smaller regions and then fit simple models to them. Practically, every internal node tests an attribute and according to which cell the path, on the branches, ends, an average value over the available observations is calculated [29]. The main parameter a DTR need for running appropriately is the maximum depth of the tree. In our case we used a $maximumDepth = 5$.

Random Forest Regression(RFR) RFR is an alternative of the DTR (in case of regression can be called regression tree - RTR - as well) methodology. In this case the concept of the DTR(RTR) can be extended using the power of contemporary computers to generate hundreds of DTRs(RTRs), simultaneously, known as random forests. The final prediction can be estimated by the smallest MSE calculated from the average of the various DTRs(RTRs)[33,34]. RFR takes as parameters the maximum depth and a random state. We used $maximumDepth = 5$ and $randomState = 2$.

K-Nearest Neighbor (K-NN): K-NN is a non parametric supervised machine learning method, which makes no assumptions for the underlying data distribution. In this method, for every instance, the distances (usually Euclidean) between a x_i feature and all features of the training set, are calculated. Then the k-nearest neighbors are selected and the x_i feature is classified with the most frequent class among the k-nearest neighbors [30].

Linear Regression (LReg) can be expressed as the statistical method applied to a dataset to quantify and define the relation between the considered features. LReg can be used for forecasting. We have, also, consider other types of linear models, fitted by minimizing a regularized empirical loss with **Stochastic Gradient Descent (Multi O/P GBR - MGBR)**, or by fitting a regressor on the original dataset and then fits additional copies of the regressor on the same dataset but where the weights of instances are adjusted according to the error of the current prediction. As such, subsequent regressors focus more on difficult cases (**Multi O/P AdaB - MAdaB**) [32]. In both training methods (MGBR and MAdab) we used $estimators = 5$.

Lasso is a modification of linear regression, where the model is penalized for the sum of absolute values of the weights. Thus, the absolute values of weight will be (in general) reduced, and many will tend to be zeros. Lasso algorithm can be explained further by the LARS algorithm [31]. **Ridge** regression is another method for modeling the connection between a depended scalar variable with one or more explanatory variables. It can be seen as a step further to Lasso. Ridge regression penalizes the model for the sum of squared value of the weights. Thus, the weights not only tend to have smaller absolute values, but also really tend to penalize the extremes of the weights, resulting in a group of weights that are more evenly distributed [35].

4. Experimental Setup

4.1. Dataset description

The dataset is composed by 4 European Cities (Athens, Gladsaxe, Lodz, Rome). For each city has been calculated the measures per day-by-day period from the 1st of January, 2020, as the first day and 31st of December, 2020, as the last. This methodology splits the year into 183 periods (leap year) of two days. For each period has been calculated the mean values of the COVID-19 measures described in Section 3. The raw COVID-19 measures dataset can take integer values between 0 (no measures) to 4 (strict restriction).

These values have been normalized in range [0-no measure, 1-strict measure], over a day-by-day period calculation (average).

The pollutant concentrations of the most harmful gases (CO, O_3, NO_2 and SO_2) has been downloaded from Sentinel-5P/TROPOMI satellite from Google Earth Engine API [24]. To be more specific only the density band, which represent the density of pollutant per pixel in mol/m^2 unit has been downloaded. As clipped geometry has been chosen a square of $0.50deg$ with center coordinates the longitude and latitude of the city center, with a pixel size of $50m \times 50m$. The final pixel density has been calculated as the mean density per period. Furthermore for the linear regression to run the image needed to be transformed as a scalar value. The transformation achieved by calculated the mean and standard deviation per city for each of the four pollutants.

Finally, the two datasets combined and for each country the train/test input compromised from the 8 COVID-19 measures and one pollutant, out of four in total, (mean and standard deviation for the current week), while the train/test output compromised from the mean and standard deviation for the next week. In the end, experiments were based on a dataset of size 10 feature values $\times$ 182 periods $\times$ 4 cities. The same dataset has been, also, used to run the correlation tests.

4.2. *Experimental results*

Figure 1 indicates the coefficient of determination (CoD). The first observation is that all measures have $CoD < 20\%(0.20)$, which indicates weak to moderate correlation for most cases. In particular, internal travel restrictions had a moderate correlation to NO2 levels. Yet, the same restrictions appear to have no correlation to CO levels. Restrictions of Internal Movement seems to have no correlation with any of the pollutant factors' levels.

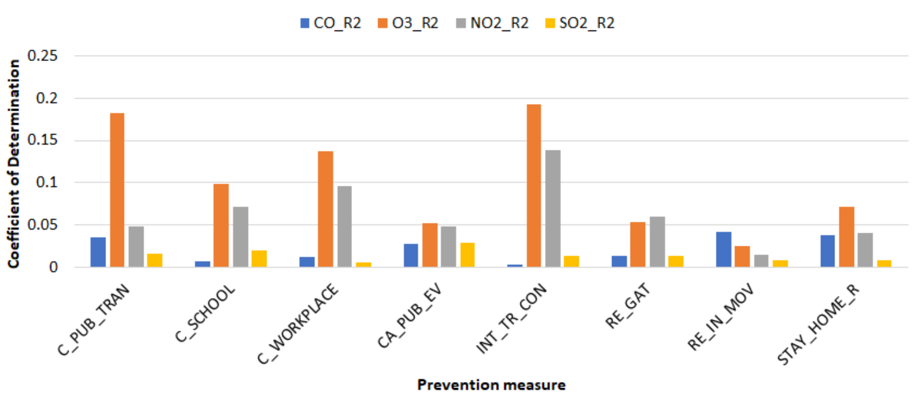


Figure 1. Coefficient of Determination (R^2) from Pearson Correlation

Figure 2 presents the alignment distance between the investigated pollutant factor and the applied measure. A small distance suggest that the behavioural patterns are more related. As such, closing the public transportation or imposing restrictions in movements signal patterns are closer to the ones of the pollutant factors. In this case the alignment distance was at least two times lesser compared to the one of international travel controls. The same is observed for the restriction in gatherings.

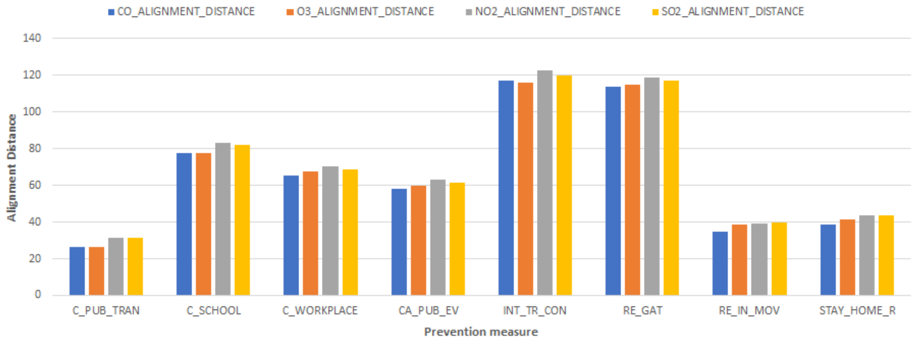


Figure 2. Alignment Distance from Dynamic Time Warping

Taking into account the previous analysis, we decided to use all measures for the training/testing the machine/deep learning techniques. From Figure 3 we can safely assume that RFR performs slightly better than the other techniques for both Carbon Monoxide(CO) and Ozone(O_3). Especially the DNN has the worst results for predicting the Ozone. In case of Nitrogen Dioxide and Sulphur Dioxide prediction (Figure 4), the best results observed in MAdaB technique. In this point we need to specify that the errors maybe seems low, however the difference with the input values (pollutant gases concentration) is in a 10% scale (e.g. if the original value was 0.000030 the output value could be 0.000028, in case of NO_2).

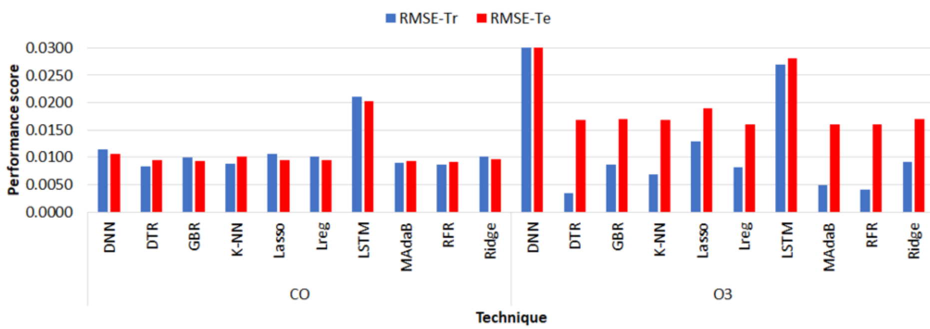


Figure 3. Root Mean Square Error (RMSE) for each machine and deep learning techniques for CO and O_3.

5. Conclusions

In this work we indicated that a weak correlation between the COVID-19 and the behaviour of the four most dangerous pollutant gases exists. To prove this correlation we performed a series of tests using Pearson and Dynamic Time Warping techniques for each of the 4 European Cities, using the each day-by-day period COVID-19 measure values and each pollutant gas density. Figures 1 and 2, which depict the coefficient of determination and the alignment distance, respectively, show that there is no measure with strong correlation, thus we used all the measures for the machine and deed learning models.

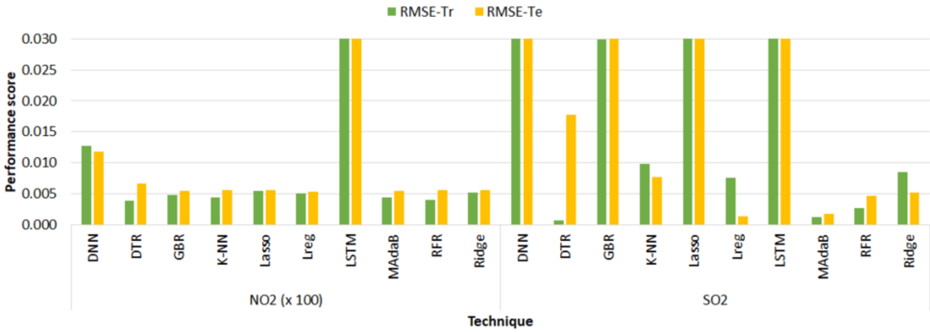


Figure 4. Root Mean Square Error (RMSE) for each machine and deep learning techniques for NO_2 and SO_2.

The machine learning outputs were sufficient enough for future predictions of the densities for each pollutant gas in a two days period. However, using bigger dataset the results can be further improved. Models like these, can be used, in the near future, for the estimation of the benefits from the replacement of petrol and oil vehicles, with other environmental friendly vehicles.

Acknowledgement

This paper is supported by the European Union Funded project euPOLIS ”Integrated NBS-based Urban Planning Methodology for Enhancing the Health and Well-being of Citizens: the euPOLIS Approach” under the Horizon 2020 program H2020-EU.3.5.2., grant agreement No 869448.

References

[1] World Health Organization. 2020. Events as they happen, https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen

[2] Ramadhan Tosepu, Joko Gunawan, Devi Savitri Effendy, La Ode Ali Imran Ahmad,Hariati Lestari, Hartati Bahar, and Pitrah Asfian, (2020) *Correlation betweenweather and Covid-19 pandemic in Jakarta, Science of The Total Environment* **725**, Indonesia. (2020), 138436. https://doi.org/10.1016/j.scitotenv.2020.138436

[3] Hannah Ritchie, Esteban Ortiz-Ospina, Diana Beltekian, Edouard Mathieu, JoeHasell, Bobbie Macdonald, Charlie Giattino, and Max Roser, (2020), Policy Responses to the Coronavirus Pandemic. https://ourworldindata.org/policy-responses-covid

[4] Shubham Sharma, Mengyuan Zhang, Anshika, Jingsi Gao, Hongliang Zhang, and Sri Harsha Kota, (2020) *Effect of restricted emissions during COVID-19on air quality in India. Science of The Total Environment* **728** (2020), 138878. https://doi.org/10.1016/j.scitotenv.2020.138878

[5] Sneha Gautam and Luc Hens, (2020) *SARS-CoV-2 pandemic in India: what might we expect? Environment, Development and Sustainability* **22** (2020), 3867–3869. https://doi.org/10.1007/s10668-020-00739-5

[6] Liane Yuri, Kondo Nakada and Rodrigo Custodio Urban, (2020) *COVID-19 pan-demic: Impacts on the air quality during the partial lockdown in São Paulostate, Brazil. Science of The Total Environment* **730** (2020), 139087. https://doi.org/10.1016/j.scitotenv.2020.139087

[7] Aurelio Tobías, Cristina Carnerero, Cristina Reche, Jordi Massagué, Marta Via, María Cruz Minguillón, Andrés Alastuey, and Xavier Querol, (2020) *Changesin air quality during the lockdown in Barcelona (Spain) one month into the SARS-CoV-2 epidemic. Science of The Total Environment* **726** (2020), 138540. https://doi.org/10.1016/j.scitotenv.2020.138540

[8] Tengku Azman, Tengku Mohd, Mohd Khair Hassan, and Wmk Aziz, (2015) *Mathematical Modeling and Simulation of an Electric Vehicle. International Journal of Mechanical Sciences* **8** (2015, July), 1312–1321. `http://dx.doi.org/10.15282/jmes.8.2015.6.0128`

[9] Enrico Ferrero, Stefano Alessandrini, and Alessia Balanzino, (2016) *Impact of the electric vehicles on the air pollution from a highway. Applied Energy* **169** (2016), 450–459. `https://doi.org/10.1016/j.apenergy.2016.01.098`

[10] Guojun He, Yuhang Pan, and Takanao Tanaka, (2020) *The short-term impacts of COVID-19 lockdown on urban air pollution in China. Nature Sustainability* **3** (122020), 1005—-101. `https://doi.org/10.1038/s41893-020-0581-y`

[11] Guojun He, Yuhang Pan, and Takanao Tanaka, (2020) *COVID-19, City Lockdowns, and Air Pollution: Evidence from China* **medRxiv** (2020), 39. `https://doi.org/10.1101/2020.03.29.20046649`

[12] Chanchan Gao, Shuhui Li, Min Liu, Fengying Zhang, V. Achal, Yue Tu, ShiqingZhang, and Chaolin Cai, (2021) *Impact of the COVID-19 pandemic on air pollution in Chinese megacities from the perspective of traffic volume and meteorological factors* **773** (2021), 145545. `https://doi.org/10.1016/j.scitotenv.2021.145545`

[13] Lijuan Shen, Tianliang Zhao, Honglei Wang, Jane Liu, Yongqing Bai, ShaofeiKong, Huang Zheng, Yan Zhu, and Zhuozhi Shu, (2021) *Importance of meteorology in air pollution events during the city lockdown for COVID-19 in Hubei Province Central China. Science of The Total Environment* **754** (2021), 142227. `https://doi.org/10.1016/j.scitotenv.2020.142227`

[14] Vikas Singh, Shweta Singh, Akash Biswal, Amit P. Kesarkar, Suman Mor, andKhaiwal Ravindra, (2020) *Diurnal and temporal changes in air pollution during COVID-19 strict lockdown over different regions of India. Environmental Pollution* **266** (2020), 115368. `https://doi.org/10.1016/j.envpol.2020.115368`

[15] Khurram Shehzad, Muddassar Sarfraz, and Syed Ghulam Meran Shah, (2020) *The impact of COVID-19 as a necessary evil on air pollution in India during the lockdown. Environmental Pollution* **266** (2020), 115080. `https://doi.org/10.1016/j.envpol.2020.115080`

[16] Susanta Mahato, Swades Pal, and Krishna Gopal Ghosh, (2020) *Effect of lockdown amid COVID-19 pandemic on air quality of the megacity Delhi, India. Science of The Total Environment* **730** (2020), 139086. `https://doi.org/10.1016/j.scitotenv.2020.139086`

[17] Aviral Agarwal, Aman Kaushik, Sankalp Kumar, and Rajeev Kumar Mishra, (2020) *Comparative study on air quality status in Indian and Chinese cities before and during the COVID-19 lockdown period. Air Quality, Atmosphere & Health* **13, 10** (Oct. 2020), 1167–1178. `https://doi.org/10.1007/s11869-020-00881-z`

[18] Ioannis Kavouras, Maria Kaselimi, Eftychios Protopapadakis, Nikolaos DOulamis, (2021) *Machine Learning Tools to Assess the Impact of COVID-19 CivilMeasures in Atmospheric Pollution* (2021). `https://doi.org/10.1145/3453892.3461327`

[19] José M. Baldasano, (2020) *COVID-19 lockdown effects on air quality by NO2 in the cities of Barcelona and Madrid (Spain). Science of The Total Environment* **741** (2020), 140353. `https://doi.org/10.1016/j.scitotenv.2020.140353`

[20] Bassim Mohammed Hashim, Saadi K. Al-Naseri, Ali Al-Maliki, and Nadhir Al-Ansari, (2021) *Impact of COVID-19 lockdown on NO_2, O_3, $PM_{2.5}$ and PM_{10} concentrations and assessing air quality changes in Baghdad, Iraq. Science of The Total Environment* **754** (2021), 14197. `https://doi.org/10.1016/j.scitotenv.2020.141978`

[21] Ioannis Kavouras, Eftychios Protopapadakis, Anastasios Doulamis, and Nikolaos Doulamis, (2018) *Skeleton extraction of dance sequences from 3D points using convolutional neural networks based on a new developed C3D visualization interface. In International Conference on Interactive Collaborative Learning* **Springer**, 267–279. `https://doi.org/10.1007/978-3-030-11935-5_26`

[22] Eftychios Protopapadakis, Athanasios Voulodimos, and Nikolaos Doulamis, (2018) *Multidimensional Trajectory Similarity Estimation via Spatial-Temporal Keyframe Selection and Signal Correlation Analysis. In Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference (PETRA '18). Association for Computing Machinery*, New York, NY, USA, 91–97. `https://doi.org/10.1145/3197768.3201533`

[23] George Kopsiaftis, Eftychios Protopapadakis, Athanasios Voulodimos, Nikolaos Doulamis, and Aristotelis Mantoglou, (2019) *Gaussian process regression tuned by bayesian optimization for seawater intrusion prediction. Computational intelligence and neuroscience* **2019** (2019). `https://doi.org/10.1155/2019/2859429`

[24] Earth Engine Data Catalog, n.d. Sentinel-5P. https://developers.google.com/earth-engine/datasets/catalog/sentinel-5p

[25] Zhong Shaobo,Yu Zhichen, Zhu Wei, (2019) *"Study of the Effects of Air Pollutants on Human Health Based on Baidu Indices of Disease Symptoms and Air Quality Monitoring Data in Beijing, China" Int. J. Environ. Res. Public Health* **16, no. 6**, 1014. https://doi.org/10.3390/ijerph16061014

[26] Xia Zhi, Song Yuexin, Ma Jin, Zhou Lujie, and Dong Zijian, (2017) *Research on the Pearson correlation coefficient evaluation method of analog signal in the process of unit peak load regulation. In 2017 13th IEEE International Conference on Electronic Measurement Instruments (ICEMI)*, 522–527. https://doi.org/10.1109/ICEMI.2017.8265997

[27] Müller, Meinard, (2007), *Dynamic Time Warping. Springer Berlin Heidelberg*, Berlin, Heidelberg, 69–84. https://doi.org/10.1007/978-3-540-74048-3_4

[28] L. Deng, G. Hinton, and B. Kingsbury, (2013) *New types of deep neural network learning for speech recognition and related applications: an overview. In2013IEEE International Conference on Acoustics, Speech and Signal Processing* (2013), 8599–8603. https://doi.org/10.1109/ICASSP.2013.6639344

[29] Pratap Sen, Mahimarnab Hajra, and Mitadru Ghosh, (2020) *Supervised Classification Algorithms in Machine Learning: A Survey and Review*, 99–111. https://doi.org/10.1007/978-981-13-7403-6_11

[30] Beckmann M., Ebecken N., and Pires de Lima B, (2015) *A KNN Undersampling Approach for Data Balancing. Journal of Intelligent Learning Systems and Applications* **7**, 104–116. https://doi.org/10.4236/jilsa.2015.74010

[31] Ryan J., Tibshirani, (2012) *The Lasso Problem and Uniqueness. arXiv:1206.0313* [math.ST]

[32] (2020) *Adaptive boosting of weak regressors for forecasting of crop production considering climatic variability: An empirical assessment. Journal of King Saud University - Computer and Information Sciences* **32, 8** (2020), 949–964. https://doi.org/10.1016/j.jksuci.2017.12.004

[33] Paul F. Smith, Siva Ganesh, and Ping Liu, (2013) *A comparison of random forest regression and multiple linear regression for prediction in neuroscience. Journal of Neuroscience Methods* **220, 1** (2013), 85–91. https://doi.org/10.1016/j.jneumeth.2013.08.024

[34] Yunjie Wang, Yifan Wen, Yue Wang, Shaojun Zhang, K. Max Zhang, HaotianZheng, Jia Xing, Ye Wu, and Jiming Hao, (2020) *Four-Month Changes in Air Quality during and after the COVID-19 Lockdown in Six Megacities in China. Environmental Science & Technology Letters* **7, 11** (Nov. 2020), 802–808. https://doi.org/10.1021/acs.estlett.0c00605Publisher:AmericanChemicalSociety.

[35] S. Hu, Q. Wang, J. Wang, S. S. M. Chow, and Q. Zou, (2016) *Securing Fast Learning!Ridge Regression over Encrypted Big Data. In 2016 IEEE Trustcom/BigDataSE/ISPA*, 19–26. https://doi.org/10.1109/TrustCom.2016.0041

[36] Kaselimi, M., Doulamis, N., Voulodimos, A., Protopapadakis, E. and Doulamis, A., (2020) *Context aware energy disaggregation using adaptive bidirectional LSTM models. IEEE Transactions on Smart Grid*, 11(4), pp.3054-3067. https://doi.org/10.1109/TSG.2020.2974347

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210096

Smart Technologies and the Case of People with Disabilities: A Preliminary Overview

Maria POLI[a]
[a] *Lecturer, Department of Interior Architecture, University of West Attica, Aigaleo, Greece*

Abstract. Recently, smart technologies as well as the digitization has entered dynamically into our lives. Cities become mega-cities because of the over population, human health conditions are gradually degraded. The existence of disabled people and the lack of their socialization the diversity of stakeholders as human beings and the difficulty of integrating them into society are few of some problems that Smart City and Smart Technology come to give manageable solutions. Solutions that could find into suggesting researches and intelligence analytics. It is important to refer to this relationship of intelligent capabilities and human resources. This study presents an overview of digital technology especially for people with disabilities. It highlights the contribution of technology to simple everyday habits of disabled and the ability to access the immediate environment. In conclusion, this article is based on the individual requirements, human rights, and preferences of people with disabilities and gives an intriguing perspective to a subject that will be in the limelight and provide effective solutions in the next years driven by technological developments.

Keywords. Smart Technology, Smart Cities, Accessibility, Disabled

1. Introduction

Nowadays human activities are entering a computerization phase. New tools have emerged, such as e-commerce sites, mobile systems and social media platforms facilitating the communication and development of our standards of living.

This study aims to explore the prospects of: Section 1) Smart Technology and what it consists of. Section 2) The definition of the disability the Disabled people, and their involvement with technology through everyday home devices. Section 3) Disabled people -Technology and Smart Applications. Section 4) Conclusion and Future

The characteristic of the 21st century society is the rapidly development of technologies information and communication. The technological revolution has an impact to everyday living conditions to our jobs even to our personal behavior. The social and economic changes are based to the rapid increase of the smart technology. For people with disabilities and especially people who needs assistance of another person all the time, technology effects their independence, privacy and autonomy. Providing an accessible life in their familiar surroundings through technology, their autonomy is increasing, their security is enhancing, and isolation is preventing by staying socially connected.

The study investigates the aspects of Smart Applications in relation to persons with disabilities, such as cognitive impairments or kinetic disabilities. Using modern digital technological tools, the author provides an overview of a sensitive subject that can be escalated to an established obligation, as well as a glimpse into the future of Smart Cities, which will become a domain of an all-inclusive society regardless of the diversity of people's disabilities.

2. About Smart Technology

What is smart technology? Is it the transfer of big data through internet? The connectivity with the Internet? The interconnection between machines and people through sensors and the ability that objects have acquired guide our lives? The existence of smartphones which give the opportunity of sharing personal and professional data? The artificial intelligence? or a combination of all these [1]?

Smart technology includes smart computers / devices with hardware, software and network technologies to be able to provide real-time real-world perception to help people make smarter decisions and provide new business solutions procedures and performance [2]. Artificial Intelligence (Al) is a machine's ability to imitate human cognitive abilities like problem solving and learning [3].

It also involves language, speech and strategic thinking, using basis of algorithmic machine learnings to build systems and find applications close to human factors creating intelligent machines and substituting human activities and functions. The application fields of Al have an enormous impact to important fields of science [4].

Through Internet we have the ability to connect the virtual with the real world. It is the magic tool that connect people with other people, with their work, with companies with public services and keeps all this together [5]. According the concept Internet of Things (IoT) all the objects or devices could be connected with others and identify them based on smart technology [6]. This connection becomes successfully realized through the assimilation of devices sensors and actuators. The intelligent systems through these get information and achieve reactions to real world [7]. The Smart devices are products and can easily programmed through an interface, a network connectivity, an application and give the characteristics of smart homes and environments [8]. The evolution of technology, the interconnection, synchronization and coordinated use of different technologies, can help to create an environment that is considered intelligent [9].

3. Disabled people – Definitions

According Pirkl J., (1994) disability, restriction or lack (resulting impairment) or ability to perform an activity in the manner or within the range considered normal for a human being. Three factors determine the presence of a disability. Each is a potential barrier limiting a person's freedom and independence: the objective, the ongoing condition results of medical, or injury conditions and the third the discrimination of our environment [10].

Disabled people are a minority in our population. The term disabled includes all the variety of population, people who are at a disadvantage due to an accident due to certain conditions such as pregnancy or even the elderly who become helpless in their old routine, to walk to climb a ladder to be flexible in their social life. While the percentage of mobile disabled people is much higher than what is officially shaken by the world health organization. About 15% of the global population, i.e. about 650 million people, have disabilities [11], [12].

For consistency reasons, we define disability according to the International Classification of Functioning, Disability and Health (ICF) of the World Health Organization, attributing to disability as a complex concept with multiple dimensions [13].

It is “the umbrella term for impairments, activity limitations and participation restrictions, referring to the negative aspects of the interaction between an individual (with a health condition) and that individual’s contextual factors (environmental and personal factors)” [14].

The term handicap refers to the inability of a person to access a set of barriers, coming from the structured environment and limit his ability to be active in social areas such as information, communication, education, e.tc. With the term (handicap) therefore emphasizes the inadequacies of the environment in which the individual lives and that prevent him from participating equally in the life of society [15]. According to the WHO, the cause is the damage (impairment) that has been caused, which can be physical, mental or otherwise, the result is disability in relation to the ability-normal in action and performance. There may be disability (physical, mental), to create disability always in relation to what is socially acceptable, but what leads to disability is the design of object function spaces in a way that gives the definition and percentage of social disability. A more careful approach, the WHO concluded that disability is the result of the relationship between an individual's abilities and the requirements of the environment. If the performance of the individual in the seven basic categories: (sight, hearing, communication, mental functions, movement, dexterity and stretching / reaching), fall short of his abilities, then the structured environment renders the individual helpless and prevents him from integrate equally into society. With this perception we lead to social disability created by social and environmental factors, various conditions, or mechanisms, deficiencies or imperfections of the environment [16]. A disability exists only when the degree and type of impairment cause difficulties in a certain environment with particular activity or capacity demands, such as in the work environment [17].

4. Disabled -Technology and Smart Applications

We try to improve the quality of life of All and especially people with cognitive kinetic or visual disabilities through technology [18]. To succeed in that we have to ensure proper living conditions providing their autonomy security and supporting their socially connection. As everybody personal social life is diverted in outside and insight of its house [19]. The wheelchair is one of the most commonly used assistive devices for improving personal mobility, which is not a tool is their legs is the condition for living in dignity. For people with kinetic problems wheelchair is the connection between the insight of their house and outside of it [20]. It is a shame for our society for the governments for people who work in research and support human rights to know that in our time 10% of disabled require a wheelchair. In 2003, it was estimated that more than 20 million of those requiring a wheelchair for mobility did not have one. Several systems with the support of technology, of smart applications, present that successful integration in many sectors of their daily life [21]. From the perspective of the Information Technology (IT) professional, the process of designing and developing for an inclusive information society requires awareness and to-the-point guidance with respect to these design-for-all tools [22]. For people with mild cognitive disabilities ensuring their daily pharmaceutical dosage or the ability to follow a daily schedule is very important for their personal independence. According (J. Haupert et al.2011) is provided the technical infrastructure such as the universal control hub or the digital product memory. Our longterm vision is to promote an accessible intelligent environment based on open standards and architectures and innovative solutions where everyone can continue to play a role in society [23]. In order the disabled people to be accepted in our society it is more important to be adapted in the word with the help of smart technology. They must be integrated in smart technology [24].

Although the access is usually addressed to people with sensory disabilities such as visual or listening impairment, or kinetic disabilities, there are also projects financial supported from the European Union that improve the lives of people with Intellectual or Development Disabilities (IDD). The Information Communication Technology (ICT) tools as mobile apps can improve their daily duties helping them to interact with the information society [25]. People with severe cerebral palsy and brain injury are often quadriplegic and computer and assistive technology is a great help for them [26]. Smart environments and interfaces systems help children with physical or learning disabilities, long-term care facilities for people with advanced neurological diseases, hospitals, and private homes [27]. There are devices which are important for everyday safety for people with motion impairments as reported by Javagopi et al (2009). Some examples of what we know and encounter in their daily lives are: A binary switch, a pull string, a mechanical touch, such as a hit plate, a tremble stick. It can also be an electrical switch, such as a voice-activated switch if the person can control making a sound, a tongue-movement detector, or a motion-sensitive photocell or infrared switch. These switches allow users with some motor control to operate wheelchairs or activate selection commands in assistive software [28].

One of the characteristics of the world-wide-web world is how has qualified the ordinary people to live in that digital revolution in a new way of live providing innovations to education entertainment and shopping [29]. Ordinary people utilize appliances with sensors and actuators during their daily life [30]. By coordinating the actions of networked devices or services, it is possible for the environment to behave in a holistic and reactive manner to satisfy the occupants needs; creating an intelligent environment [31].

Disrupting the used home appliances into sets supported by network accessible services it is possible to disrupt the appliance of our everyday life or to create new one combining network services and finally to create a new one called virtual appliance [32]. This principle can be extended to decompose and re-compose software applications allowing users to create their own bespoke applications [31]. Collectively, such user created entities are referred to as Meta – appliances or – applications, more generally abbreviated to Maps [33]. Deconstruction and user customized MAps raise exciting possibilities for occupants of future intelligent environments, and sets significant research challenges [34].

5. Conclusion and Future

It is necessary to believe that a Smart City has to be anthropocentric and to require from its citizens to participate to the design prosses, to realize the vision, to participate to take an active part in the realization of all plans. It is important to understand the severe problems of our planet the importance of the unique human being and to use technology in a proper way friendly to the environment and the people.

Smart Cities need to be designed to allow the inclusion of all kinds of citizens. Motor disabled people like wheelchair users, blind people who may have problems to interact with the city. Internet of Things (IoT), technologies provide the tools to include all citizens in the Smart City context. Using Augmented Reality and Internet of Things we can improve accessibility of people with motor disabilities in the context of Smart Cities. To improve processes using technology we connect people with physical objects, information systems and among themselves. We improve the relation between the citizens and their context. We promote independence, equal opportunities and human rights.

References

[1] C. Troussas, A. Krouska, & C. Sgouropoulou, *Impact of social networking for advancing learners' knowledge in E-learning environments*, Educ Inf Technol 2021 https://doi.org/10.1007/s10639-021-10483-6]

[2] D. Washburn, U. Sindhu, *Helping CIOs Understand Smart City Initiatives*. Forrester Research, 2010

[3] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou. *User acceptance of augmented reality welding simulator in engineering training.* Educ Inf Technol, 2021 https://doi.org/10.1007/s10639-020-10418-7

[4] C.Troussas, C. Sgouropoulou: Innovative Trends in Personalized Software Engineering and Information Systems - The Case of Intelligent and Adaptive E-learning Systems. *Frontiers in Artificial Intelligence and Applications 324,* IOS Press 2020, ISBN 978-1-64368-096-5, pp. 1-96

[5] C. Troussas, A. Krouska, F. Giannakas, C. Sgouropoulou, I. Voyiatzis: *Redesigning teaching strategies through an information filtering system.* PCI 2020-24th Pan-Hellenic Conference on Informatics, November 2020, Pages 111-114. https://doi.org/10.1145/3437120.3437287

[6] C. Troussas, A. Krouska, C. Sgouropoulou, I. Voyiatzis: *Ensemble Learning Using Fuzzy Weights to Improve Learning Style Identification for Adapted Instructional Routines*. Entropy 22(7): 735, 2020

[7] C. Troussas, A Krouska, C. Sgouropoulou: Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education. *Computers & Education*, (2020), p144

[8] A. Krouska, C. Troussas, C. Sgouropoulou, Applying Genetic Algorithms for Recommending Adequate Competitors in Mobile Game-Based Learning Environments. In: Kumar V., Troussas C. (eds*) Intelligent Tutoring Systems.* ITS 2020. Lecture Notes in Computer Science, vol 12149. Springer, Cham, 2020 https://doi.org/10.1007/978-3-030-49663-0_23

[9] U. Gretzel, L. Zhong, C. Koo, Application of smart tourism to cities. *International Journal of Tourism*, 2016

[10] J. Pirkl, Trangenerational design. *Van Nostrand Reinhold*, New York, (1994), Pp73

[11] World Helath Organization Fact sheet on wheelchairs, October 2010 https://apps.who.int/iris/bitstream/handle/10665/205041/B4616.pdf?sequence=1&isAllowed=y]

[12] World Health Organization. (2011). World Report on Disability: Summary, 2011. Geneva, Switzerland, 2011 https://apps.who.int/iris/bitstream/handle/10665/205041/B4616.pdf?sequence=1&isAllowed=y]

[13] World Health Organization. (2001). International Classification of Functioning, Disability and Health (ICF). Geneva, Switzerland, 2001

[14] World Health Organization, World Report on Disability: Summary, 2011. Geneva, Switzerland, 2011

[15] Accessibility http://www.disabled.gr/category/amea-disability-library/access/

[16] A. Vergou, D. Paganopoulou, P. Pelekanou, *Degree of user satisfaction from the level of accessibility of public services in the Municipality of Patras*, school of health professions and social work department, 2013

[17] K. Vornholt, P. Villotti, B. Muschalla, J. Bauer, A. Colella, F. Zijlstra, G. Van Ruitenbeek, S. Uitdewilligen & M. Corbière, Disability and employment – overview and highlights, *European Journal of Work and Organizational Psychology, 2017* doi: 10.1080/1359432X.2017.1387536

[18] M. Poli, Greece, Tourism and Disability, *Strategic Innovative Marketing and Tourism, International Journal of Strategic Innovative Marketing 8th ICSIMAT*, Springer, 2019 https://doi.org/10.1007/978-3-030-36126-6_74].

[19] A. Krouska, C. Troussas, C. Sgouropoulou: *A Personalized Brain-Based Quiz Game for Improving Students' Cognitive Functions*. BFAL 2020, p102-106

[20] World Health Organization Fact sheet on wheelchairs, Wheelchairs assist people with disabilities to become productive members of their communities, https://apps.who.int/iris/bitstream/handle/10665/205041/B4616.pdf, October 2010

[21] C. Troussas, F. Giannakas, C. Sgouropoulou & I. Voyiatzis (2020) Collaborative activities recommendation based on students' collaborative learning styles using ANN and WSM, *Interactive Learning Environments,* 2020 doi:10.1080/10494820.2020.1761835

[22] P. Koutsabasis1, J. S. Darzentas, T. Spyrou, A. Velasco, Y. Mohamad, J. Darzentas, *Towards an Environment that Supports Internet Design for All*, Conference paper, 2001

[23] R. Neßelrath, J. Haupert, J. Frey, B. Brandherm, *Supporting Persons with Special Needs in their Daily Life in a Smart Home, IEEE*, 2011 doi: 10.1109/IE.2011.75

[24] C. Troussas, A. Krouska, C. Sgouropoulou, (2020) *Dynamic Detection of Learning Modalities Using Fuzzy Logic in Students' Interaction Activities.* In: Kumar V., Troussas C. (eds) Intelligent Tutoring Systems. ITS 2020. Lecture Notes in Computer Science, vol 12149. Springer, Cham, 2020 https://doi.org/10.1007/978-3-030-49663-0_24

[25] J. Dekelver, J. Daems, S. Solberg, N. Bosch, L. Van de Perre, A. De Vliegher, T. More, A digital travel assistant for people with intellectual disabilities: *Modeling and design using contemporary intelligent technologies as a support for independent traveling of people with intellectual disabilities*, 2015 doi: 10.1109/IISA.2015.7388014

[26] M.Poli, Covid-19 Effects in People's Daily Life, Social Media's Role and the Power of Emotions and State of Mind, *International Journal of Strategic Innovative Marketing 9th ICSIMAT*, Athens, Greece, Springer, 2020 https://doi.org/10.1007/978-3-030-66154-0

[27] B. Margrit, Intelligent Interfaces to Empower People with Disabilities, *Handbook of Ambient Intelligence and Smart Environments, Springer,* 2010 pp 409-432

[28] D Jayagopi, H. Hung, C. Yeo, D. Gatica-Perez, (2009) Modeling dominance in group conversations using nonverbal activity cues. IEEE *Trans. on Audio, Speech, and Language Processing, Special Issue on Multimodal Processing for Speech-based Interactions, March (2009)* Vol. 17, No. 3, pp. 501-513

[29] C. Troussas, A. Krouska, C. Sgouropoulou, Towards a Reference Model to Ensure the Quality of Massive Open Online Courses and E-Learning. In: Frasson C., Bamidis P., Vlamos P. (eds) *Brain Function Assessment in Learning. BFAL 2020. Lecture Notes in Computer Science*, vol 12462. Springer, Cham., 2020 https://doi.org/10.1007/978-3-030-60735-7_18

[30] A. Krouska, C.Troussas, C. Sgouropoulou: Usability and Educational Affordance of Web 2.0 tools from Teachers' Perspectives. *PCI 2020-24th Pan-Hellenic Conference on Informatics*, November (2020), Pp 107-110. https://doi.org/10.1145/3437120.3437286

[31] J. Chin, V. Callaghan and G. Clarke, End-user Customisation of Intelligent Environments, *Handbook of Ambient Intelligence and Smart Environments*, 2010, p 370, Springer

[32] C. Papakostas, C. Troussas, A. Krouska, C. Sgouropoulou, Exploration of Augmented Reality in Spatial Abilities Training: A Systematic *Literature Review for the Last Decade, Informatics in Education*, 20(2021), no. 1, 107-130, doi 10.15388/infedu.2021.06

[33] J. Chin, V. Callaghan and G. Clarke, End-user Customisation of Intelligent Environments, *Handbook of Ambient Intelligence and Smart Environments,* Springer, (2010), p 372

[34] J. Chin, V. Callaghan, G. Clarke, Soft-appliances: A vision for user created networked appliances in digital homes, *Journal of Ambient Intelligence and Smart Environments* 1, IOS Press, (2009), 65–71, doi 10.3233/AIS-2009-0010,

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210097

Designing Personalisation in a Environmental Recommendation System: Differences of AHP and Fuzzy AHP

Katerina KABASSI
Department of Environment, Ionian University

Abstract. Different methods have been proposed for designing the personalization process in a recommendation system. In the past, multi-criteria decision making theories have been proposed for the design of stereotypes in a recommendation system for environmental awareness. The main objective of this paper is on presenting the main differences when applying the fuzzy AHP and AHP for designing the weights of criteria in a recommendation system that its personalization process is based on multi-criteria decision making theories.

Keywords. Multi-Criteria Decision Making, Personalisation, Recommendation Systems, Stereotypes

1. Introduction

The research discipline of recommender systems arose to address the problem of information or choice over-abundance, i.e., to help users find information or items that are most likely to be interesting to them or to be relevant to their needs [1]. Recommendation systems have been expanded for use in mobile phones since these devices have been widely spread. Especially, Smartphone crowdsourcing enables real-time data gathering and gives devices access to a wealth of information about each device's user. As a result, recommendation systems for mobile phones have been developed. The main advantage of those systems is that they can make use of the large number of modalities for interaction between them and their environment [2]. Implicit information coming from Smartphone crowdsourcing can be combined with location and time specific data can further utilize software adaptation [8].

Using all these information, we developed a mobile recommendation system for providing personalized information about Centuries-Old Olive Groves. The system is called OldGrove and is aimed to be used in the Ionian Islands. In order to provide personalized interaction, OldGrove employs user modeling techniques. More specifically, the system uses multi-criteria decision making techniques (MCDM) in combination with individual user modeling and stereotypes for providing personalized interaction. Stereotypes are used for providing the weights of the criteria and individual user modeling the values of the criteria. This paper aims at presenting the design of stereotypes using Analytic Hierarchy Process (AHP) [10]. The main objective of this research is to present a comparative analysis AHP and Fuzzy AHP (FAHP) [3] for designing stereotypes in a recommendation system. Comparison of AHP and FAHP has been implemented before but in completely different context [3, 6, 7].

2. Designing the Personalisation in a Recommendation system

Personalised interaction may be provided when the system maintains user models. A user model is a collection of information on the knowledge, interests and other characteristics of each user. Stereotypes [9] are used in user modeling in order to provide default assumptions about individual users belonging to the same category according to a generic classification of users that has previously taken place. This method has the advantage of providing personalized recommendations from the first interaction of the user with the system.

In this approach, the system uses a combination of stereotypes and individual user modeling. For the personalization of the recommendations, personalization uses multi-criteria decision making techniques for combining the data about the user that are stored in the user model. Taking into account the methods of personalization as well the multi-criteria decision making techniques used, the system uses stereotypes that provides the weights of the criteria and individual user model that provides the values of the criteria.The individual user model and the activated stereotype are loaded during the first phase of the recommendation mechanism, which is called cold start mechanism. Figure 1 presents all the steps the system implements for providing personalized recommendations.

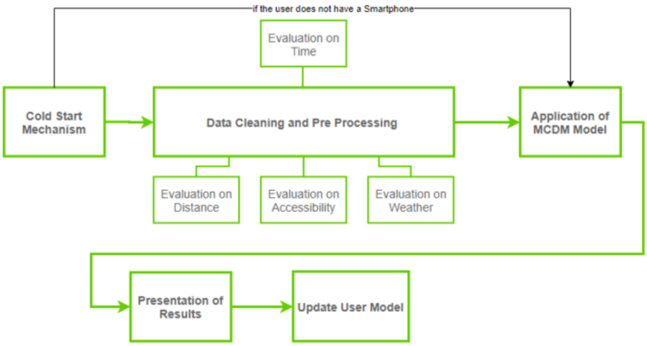


Figure 1. The mechanism of personalized recommendation of OldGrove

In the cold start mechanism, all possible recommendations are generated. If the system tracks a Smartphone, then the module ‘Data Cleaning and Pre-Processing’ runs. Otherwise, the system evaluates all recommendation that have been generated using the MCDM model. The ‘Data Cleaning and Pre-Processing’ module aims at reducing the possible recommendations before these are evaluated by the application of the MCDM model. The ‘Data Cleaning and Pre-Processing’ module uses criteria such as Distance, Accessibility, Weather and Time.

The MCDM model uses other four criteria for evaluating recommendations:

- Knowledge: Compatibility with the user’s knowledge is also considered very important while evaluating the different pieces of information.
- Interests: Compatibility with user’s interests is considered important in order to select the most interesting pieces of information for each user
- Skills: Compatibility with user’s skills is considered important so that the users could comprehend the information provided to them.
- Learning Needs: Compatibility with the learning needs of a user could also be use for evaluating different pieces of information.

These criteria are not considered as equally important by the system. The values of the weights of importance are acquired by the stereotype that has been activated during the cold start mechanism.

3. Stereotype Design: Fuzzy AHP vs AHP

In the early stages of the software's life cycle in order to find out the possible categories of users, we conducted an expert-based study [4], which revealed that the stereotypes that could be used for categorizing the possible users of the system are: Foresters, Agriculturalists, Historians, Farmers, Tourists, Residents.
AHP is a very well known MCDM theory. Its main advantage is that it presents a formal way of quantifying the qualitative criteria. A main reason for selecting AHP over other MCDM theories is that it presents a well defined way for calculating the weights of the criteria while most of MCDM theories have only defined the way criteria and weights are combined to rank alternatives but consider the weights of the criteria as already known.

Table 1. The weights of the criteria in the different stereotypes using AHP

	Knowledge	Interests	Skills	Learning Needs
Foresters	0.566	0.153	0.233	0.049
Agriculturalists	0.541	0.091	0.227	0.141
Historians	0.574	0.235	0.235	0.055
Farmers	0.473	0.081	0.302	0.144
Tourists	0.363	0.387	0.196	0.054
Residents	0.418	0.143	0.346	0.093

Table 2. The weights of the criteria in the different stereotypes using FAHP

	Knowledge	Interests	Skills	Learning Needs
Foresters	0.556	0.166	0.229	0.049
Agriculturalists	0.542	0.118	0.212	0.128
Historians	0.569	0.153	0.224	0.054
Farmers	0.417	0.136	0.308	0.139
Tourists	0.355	0.421	0.175	0.049
Residents	0.374	0.147	0.334	0.146

We have used AHP to execute three steps: (1) structuring the hierarchy of criteria; (2) assessing the expert evaluations by pair-wise comparison of the criteria; and (3) using the eigenvector method to yield weights for the four criteria. Through AHP, the importance of several criteria is obtained from a paired comparison process, in which the relevance of attributes or categories of drivers of intangible assets are matched two-by-two in a hierarchical structure [3]. The application of AHP revealed the weights of criteria for each stereotype of users (Table 1).

Except for AHP, fuzzy AHP has also been extensively used. Fuzzy AHP is based on the fuzzy interval arithmetic with triangular fuzzy numbers and confidence index α with an interval mean approach has been applied to weight evaluative element [3]. The main difference of AHP with FAHP is that the latter uses linguistic terms that are associated with triangular fuzzy numbers. A detailed application of FAHP for the estimation of weights for the different categories of users are presented in detail in [4] and are presented in summary in Table 2.

One can easily observe that the two models provide similar results but not identical. Very similar were the results for foresters, whereas main differences are presented in all the other stereotypes. In the Agriculturalists, Historians, Farmers, and Tourists the main difference is observed in the weights of the criterion Interests. For example, Interests are considered to be more important for agriculturalists when weights are calculated with AHP than when are calculated with FAHP. As far as the stereotype Residents is concerned differences are observed in the weights of criteria Knowledge and Learning Needs. Testing 5 scenarios where the values of the criteria were known, revealed that the differences were minor and did not affect the results of the recommendation system.

4. Conclusions

In this paper we have described the design of stereotypes in a recommendation system for environmental information. For the design of the stereotypes, we have used in turn AHP and FAHP. The main focus of the paper is on presenting the differences that occur when applying both models. Small differences occurred in the weights of the criteria but the testing of the both sets by using 5 specific scenarios revealed that the differences were not important. FAHP proved to be easier for expert users to implement. Similarly to [5], it is also demonstrated the advantage of being able to capture the vagueness of human thinking and to aid in solving the research problem through a structured manner and a simple process. It is among our future plans to implement the system with both sets of weights for each stereotype and test its effectiveness with real users.

References

[1] G. Adomavicius, Y. Kwon, Multi-criteria recommender systems, *Recommender Systems Handbook*, (2015), 847–880. http://dx.doi.org/10.1007/978-1-4899-7637-6_25.

[2] P. Bhargava, A. K. Agrawala, Modeling Users' Behavior from Large Scale Smartphone Data Collection, *EAI Endorsed Trans. Context-aware Systems & Applications*, **3(10)** (2016) https://doi.org/10.4108/eai.12-9-2016.151677

[3] Y. C. Chou, H. Y. Yen, C. C. Sun, J. S. Hon, Comparison of AHP and fuzzy AHP methods for human resources in science technology (HRST) performance index selection, *2013 IEEE International Conference on Industrial Engineering and Engineering Management*, 2013, 792-796, doi: 10.1109/IEEM.2013.6962520.

[4] K. Kabassi, A. Martinis, Studies for Designing Stereotypes in a Mobile Recommendation System for Environmental Awareness. *15th IEEE International Workshop on Semantic and Social Media Adaptation & Personalization* (2020).

[5] G. Kabir, M.A.A. Hasin, Comparative Analysis of AHP and Fuzzy AHP Models for Multicriteria Inventory Classification. *International Journal of Fuzzy Logic Systems (IJFLS),* **1**(1), 1-16.

[6] R. Mosadeghi, J. Warnken, R. Tomlinson, H. Mirfenderesk (2015) Comparison of Fuzzy-AHP and AHP in a spatial multi-criteria decision making model for urban land-use planning, *Computers, Environment and Urban Systems*, **49** (2015), 54-65,https://doi.org/10.1016/j.compenvurbsys.2014.10.001.

[7] M. Naderzadeh, H. Arabalibeik, M. R. Monazza, I. Ghasemi, Comparative Analysis of AHP-TOPSIS and Fuzzy AHP Models in Selecting Appropriate Nanocomposites for Environmental Noise Barrier Applications, *Fluctuation and Noise Letters*, **16(4)** (2017) 1750038

[8] A. M. Otebolaku, M. T. Andrade, User context recognition using smartphone sensors and classification models, *Journal of Network and Computer Applications*, **66** (2016), 33-51https://doi.org/10.1016/j.jnca.2016.03.013G.

[9] E. Rich, Users are individuals: individualizing user models, International Journal of Human–Computer Studies, **51** (1999), 323–338.

[10] T. Saaty, G. Hu, Ranking by Eigenvector Versus Other Methods in the Analytic Hierarchy Process. *Applied Mathematics Letters*, **11(4)** (1998), 121-125.

Brain Assessment and Reasoning

This page intentionally left blank

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210099

Prediction of Amusement Intensity Based on Brain Activity

Gabrielle TOUPIN [a], Mohamed S. BENLAMINE[a] and Claude FRASSON [a]
[a] *University of Montréal*

Abstract. Amusement can help modulate psychological disorders and cognitive functions. Unfortunately, algorithms classifying emotions still combine multiple positive emotions into a unique emotion, namely joy, making it hard to use amusement in a real-life setting. Here we train a Long-Short-Term-Memory (LSTM) on electroencephalography (EEG) to predict amusement on a categorical scale. Participants (n=10) watched and rated 120 videos with various funniness levels while their brain activity was recorded with an Emotiv Headset. Participant's ratings were divided into four bins of amusement (low, medium, high & very high) based on the participant's ranking's percentile. Nested cross-validation was used to validate the models. We first left out one video from each participant for the final model's validation and a leave-one-group-out technique was used to test the model on an unseen participant during the training phase. The nested cross-validation was tested on sixteen different videos. We created an LSTM model with five hidden layers, vatch size of 256 and an input layer of 14 x 128 (number of electrodes x 1 sec of recording) and four nodes representing the different levels of amusement. The best model obtained during the training phase was tested on the unseen video. While the level of accuracy between the validation videos varies slightly (mean=57.3%, std=13.7%), our best model obtained an accuracy of 82,4%. This high accuracy supports the use of brain activity to predict amusement. Moreover, the validation process we design conveys that models using this technique are transferable across participants and videos.

Keywords. amusement, LSTM, EEG, emotiv, emotions, humour

1. Introduction

1.1. Context & Motivation

Humour is a social behaviour that allows people to break the ice, relax the atmosphere, or gently pass a criticism [1]. It is a complex cognitive process that can result in an emotional state of amusement and can trigger laughter [2]. Research in positive psychology induces amusement to modulate psychological disorders, such as schizophrenia and depression [3-5]. This positive emotion can also benefit cognitive functions such as memory [5,6]. In addition to having different research uses, amusement differs from joy in terms of facial expressions, physiological signals, and feelings [7,8]. Nevertheless, predicting emotions still widely combines these positive emotions together [9]. Only a handful of studies can predict different positive emotions [8,10]. Thus, a better understanding and prediction of amusement would benefit research using amusement as a regulator, as well as new technologies.

The development of new models to predict emotions in artificial intelligence is on the rise. Those algorithms are trained to predict emotions based on facial expressions, electroencephalography (EEG) or physiological signals, such as electrocardiography (ECG). Algorithms using facials expressions to predict emotions can be complicated when used in real-world applications and experiments. First, using a filming process requires specific settings where the participant always faces a camera, making it notably difficult for moving subjects and situations where faces are hidden (e.g., virtual reality headset, wearing a mask, etc.). So far, algorithms based on artificial intelligence, like Emotient, are better than humans at classifying basic emotions when they are typical, exaggerated and static. However, the accuracy drastically drops when used on spontaneous, dynamic and mixed emotions [11-13]. While there is more work to be done in this area, using brain activity and physiology might be a better choice to train algorithms to predict emotion since it does not require the participant to be static in front of the camera and physiological signals cannot be intentionally controlled. With the use of new technologies like Emotiv (https://www.emotiv.com/), where the headset is affordable, requires minimum setup and connects via Bluetooth, new setup experiments and real-life applications are conceivable.

Researchers use different estimators and features to train algorithms to detect emotion based on EEG signals. If we look at the machine learning side, studies use estimators such as support vector machines (SVM), Naive Bayes (NB) and K-nearest neighbours (KNN) to classify emotions [8,10,12]. When looking at deep learning, there is no consensus on which algorithm is best for emotion classification [13]. In their study, Alhagry [14] reaches an accuracy score over 85% with a Long-Short Term Memory algorithm (LSTM) to predict the intensity of arousal and valence of the emotion base on EEG. LSTM is a recurrent neural network (RNN) architecture used in the field of deep learning. This algorithm is promising since it can learn from complex data and predict both on a continuous and categorical scale. Feeding raw EEG data allows us to create algorithms that do not require transformed data, which takes time to compute. Furthermore, LSTM can take more information into account than classical machine learning techniques, meaning that even some artifacts or movements detected by the EEG headset could help define the amusement intensity. Therefore, we propose to train an LSTM algorithm to predict amusement intensity with EEG data acquired with the Emotiv headset. This study brings new insights into the prediction of emotion intensity and amusement.

1.2. Objective

This paper’s objective is to develop a deep-learning model that can predict the amusement’s intensity of the participant based on its brain activity. We trained an LSTM to predict the categorical score of amusement (low, medium, high, very high) based on one second of brain activity from 14 electrodes. When developing our model, we took special care to ensure that our model was transferable to new participants and new visual content.

2. Methods

2.1. Participants

Ten participants (7 women, 3 men) were recruited for this experiment. They were approached on social media and were offered monetary compensation in exchange for their participation. Recruited participants were between the age of 18 and 30 and had similar education, standard or corrected-to-normal vision and no neurological or psychological disorders. One participant had to be excluded since he never completed the active task (*i.e.* pressing the space bar) during the experiment.

2.2. Material

One hundred twenty video clips were used as humorous stimuli. These video clips were selected in two steps. First, undergrad volunteers selected short portions of humorous and neutral videos from movies, short clips and video compilations. A total of 50 neutral videos and 100 humorous videos were selected. The videos were cropped to have a length between 8 to 12 seconds (mean of 10 seconds). Furthermore, black outlines and the sound, when they were present, were removed from the clip. Second, a preliminary study was conducted in order to validate the selected stimuli. Forty participants watched and rated every video on the following scales: arousal, valence and funniness. Those ratings were used to confirm that there is enough variability in the selected videos. We performed a K-Means clustering with three clusters on arousal, pleasantness and funniness. For the three clusters of funniness, corresponding to neutral, funny and very funny, we selected the best 40 videos of each as stimuli for this study.

2.3. Procedure

Participants arrived at the Functional Neuroimaging Unit and were inquired to read, understand, and sign the consent form. Participants were seated comfortably in a Faraday Cage. There was a screen, a mouse, and a keyboard in front of them to perform the experiment. The task used in this study was created with Psychopy 3 [15] and consisted of four blocks with 30 trials each. Each bloc was designed with a pseudo-randomized order and included ten neutral videos, ten funny videos and ten very funny videos. We made sure that there were at most three videos of the same type in a row. A single trial consists of a fixation cross (2-3 seconds), followed by a video (8-12 seconds), another fixation cross (3 seconds) and a single question ("how funny was this video"). The question was on a scale of 1 (not funny) to 100 (very funny) and was answered with the mouse. An active task was used while watching the video to keep the participant engaged in the task. The participant was asked to press the spacebar on the keyboard when he thinks the video was funny.

The first part of the experiment consisted of a practice block where the participant got familiar with the trial design in the experimenter's presence. The participant was free to ask any question about the trial and the experimenter made sure that the task was understood. The practice block was followed by an emotional questionnaire where the participant evaluated the presence of 20 emotions on a Likert scale of 1-7. Then, a resting state (6 minutes) was measured. During the resting state, the participant was asked to look at the cross in the screen's center and stay neutral. The participant was then ready to start the four blocks of the experiment. Once the participant was ready, he could press the keyboard's space button to start the block. The experimenter went inside the room between each block and ensured the participant was still in good shape to proceed with the task. It was recommended to take a couple of minutes to relax between each block. After all the blocks were completed, the participant was presented with another resting state and the same emotional questionnaire.

2.4. EEG recording

The Emotiv Epoc headset was used to collect electrical activity during the task. EEGs were recorded from 14 electrodes (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4) with two reference nodes located behind the ears. The generated data are in µVolt with a sampling frequency of 128 samples per second. Electrodes were moisturized with a saline solution to maintain electrode impedance under the software's required level. Impedance was checked during the initial installation, followed by a rechecked before the start of each block.

2.5. Data Preparation

2.5.1. Participant Evaluation of Amusement

Only humorous videos were used to develop the model. Since the rating interval differs between the participants, we scaled the rating between 0 and 1 for each participant. The participant's lowest value was converted to 1, his highest value to 100 and every value in between was scaled proportionately. We computed the user's amusement rating level by dividing the rating scale into three intervals: [0..0.25[for low amusement, [0.25..0.50[for medium amusement, [0.50..0.75[for high amusement and [0.75..1[for very high amusement.

2.5.2. Time of Interest

Funniness appears mostly at the videos' end [16], leading us to choose the video's end as the time of interest for humorous videos (Figure 1). More precisely, if the participant pressed the spacebar to indicate that it is indeed a funny clip, we only used EEG data between the button press and the end of the video and assigned the user's reported amusement rating. On the other hand, if no button was pressed, particularly for less funny videos, we assumed that the reported funniness was stable across the video and used EEG data associated with the full video's length.

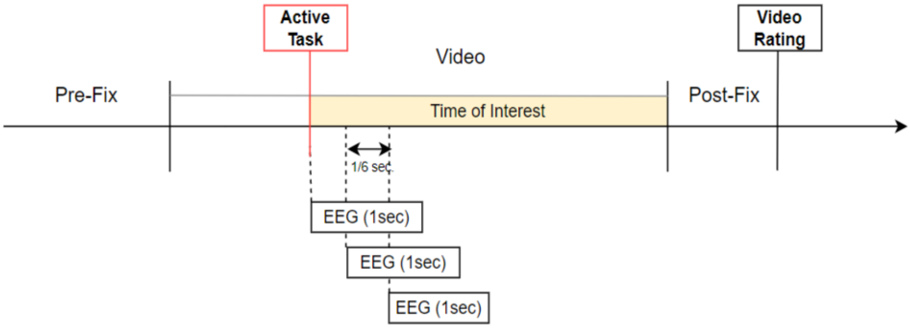


Figure 1. Task performed by the participants.

2.5.3. EEG Data Cleaning

EEG data collected via Emotiv were cleaned via the python's MNE library. Emotiv has a sampling frequency of 128hz per second. The first cleaning part is done automatically by the Emotiv Software, where it uses a 5th-order digital Sinc to filter between 0.2 Hz and 45 Hz. Plus, it uses a Notch filter at 60Hz since it is the frequency band for North America's electricity. It also removes most of the eye's blinks and heartbeat from the signal. Additionally, we complemented the cleaning from Emotiv Software with an additional process done with python's library MNE. To validate that all eyes and cardiac artifacts were well removed, we decomposed the EEG signal using an Independent Component Analysis (ICA). ICAs that were strongly correlated with either eye blinks or heartbeat were removed from the signal. Finally, we manually observed the signal of each participant and annotated then noisy parts of the signal. Epochs with those annotated parts were not used in further analysis.

2.5.4. Data Selection

Our model will attempt to predict the user's amusement rating from a 1 second EEG data from all electrodes. We used a data matrix of shape 128x14 which holds 1 second of recording for each of the 14 electrodes (Figure 1). This second of recording was associated with the rating of the participant for this specific video. For the length of the trial's time of interest, we move the data matrix 1/6 second in time and assign the participant's rating to the data matrix.

2.6. Model Training and validation

Deep-Learning models learn data representations within their hidden layer at multiple levels of abstraction [17]. We have constructed an LSTM model with 5 hidden layers (four layers of 14* 128 neurons and the fifth layer of 128 neurons), with a batch size of 256 and an input layer of 14*128 (see Figure 2). The Network output layer has four nodes representing the amusement level.

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 14, 128)	131584
lstm_1 (LSTM)	(None, 14, 128)	131584
lstm_2 (LSTM)	(None, 14, 128)	131584
lstm_3 (LSTM)	(None, 14, 128)	131584
lstm_4 (LSTM)	(None, 14, 128)	131584
lstm_5 (LSTM)	(None, 128)	131584
dropout (Dropout)	(None, 128)	0
dense (Dense)	(None, 4)	516

Figure 2. Neural network summary

To make sure our model can generalize on new content, we extracted the data of one video from all participants as our final test of the generated model. To ensure our validation accuracy is unbiased by the chosen video, we repeated our model training with 16 different videos and discussed the results below. Furthermore, to ensure the model is usable on an unseen participant, we used a leave-one-group-out (LOGO) technique during the training and testing phase (Figure 3). More precisely, we trained the algorithm on all 8 participants and tested it on the last one not previously seen by the algorithm. We repeat this procedure so that each participant is used as the test set once. The mean accuracy of all algorithms can describe the algorithms' performance.

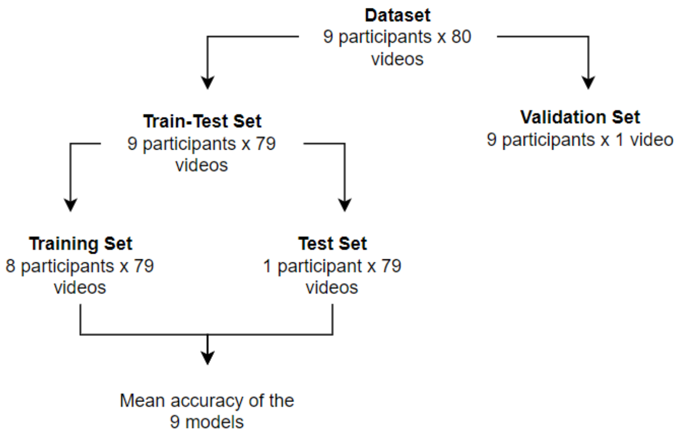


Figure 3. Leave-One-Group-Out cross-validation method

Initially, we set the number of training epochs to 100 by cross-validation. We have used early stopping techniques to prevent model overfitting (Figure 4). We set model monitoring on validation loss metric during training. Early stopping was used to evaluate different learning rate values for the model. The weights of the best model were recorded with minimum validation loss.

```
    earlyStopping = EarlyStopping(monitor='val_loss', patience=10,
verbose=0, mode='min')

    mcp_save = ModelCheckpoint('SavedModels/mdl_clf_wts.hdf5',
save_best_only=True, monitor='val_loss', mode='min')

    reduce_lr_loss = ReduceLROnPlateau(monitor='val_loss', factor=0.1,
patience=7, verbose=1, epsilon=1e-4, mode='min')
```

Figure 4. Early stopping code snapshot

3. Results and Discussion

3.1. Base Model

3.1.1. Generalization of the model

The validation accuracy of the base model, when tested on an unseen video can be found in Table 1 under *validation accuracy*. Taken together, our models predict an unseen video with an average of 64.2% (std=14.7%) with a maximum accuracy of 88.9% (model #1) and a minimum of 32.9% (model #6). Since there is high variability in the validation accuracy, we cannot conclude that this specific algorithm can yet be transferable to other content. On the other hand, when the algorithm is tested on an unseen participant during the training phase, accuracy is more stable. The column Mean Accuracy Training (STD) of Table 1 shows the model's mean accuracy when tested on each of the unseen participants (n=9). We obtain a mean accuracy of 74.9% (std=3.8%) with an accuracy as high as 87.5% (model #1) and as low as 71.5% (model #9). This high accuracy and low standard deviation show that our model can predict the amusement level based on a participant's brain activity that it has never seen before.

While looking at each model's confusion matrix, we saw that the fourth class, namely *very-high amusement*, is well represented in none of the models (see example in Figure 5), which may cause the high variability observed in the validation accuracy. Furthermore, the good results obtained might be due to overfitting on those unbalanced categories.

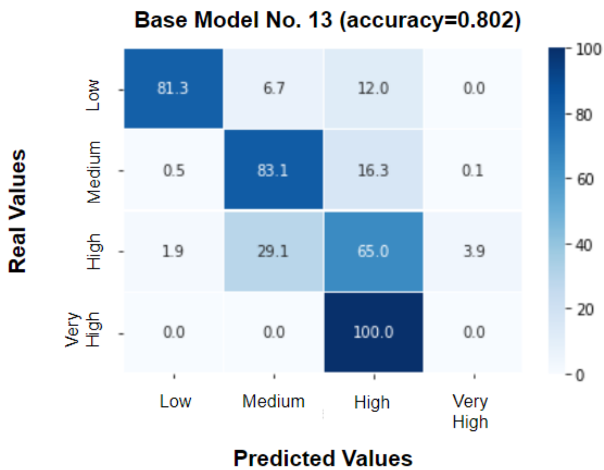


Figure 5. Example of a confusion matrix during validation phase where very high amusement is unwell represented

3.2. Model with Class Weight

3.2.1. Weight Classes

To overcome the fact that our classes are unbalanced, we assign each class a weight during the training phase. We used an automatic function that looks at the distribution of labels and produces weights to equally penalize under or over-represented classes in the training set. While each training model had different weights, the weight was very similar between models. A mean weight of 1.339 (std=0.014) was assign to low amusement, (std=0.009) to medium amusement, 0.591 (std=0.004) to high amusement and 6.284 (std=0.298) to very high amusement.

3.2.2. Generalization of the model

Out of the 16 models we trained, 10 of them saw their validation accuracy drop after the model was adjusted with weights (Table 1 under Validation Accuracy Gain/Loss). This confirms that our base model was overfitting on most of the models. We obtained a mean validation accuracy of 57.3% (std=13.7%). Like our base model, this high variability across the accuracy confirms that our model is not yet able to transfer perfectly to unseen videos. Our best model obtains an accuracy of 82.4% and our lowest is at 31.1%.

Training and testing accuracy with the adjusted weights also dropped for all models. During the training and testing phase. We obtained a mean accuracy of 63.1% (std=3.0%) where our best model has an accuracy of 72.6% (std=13.3%) and our worst model has an accuracy of 59.6% (std=15.5%). The mean accuracy is still above the theoretical chance level for four classes (chance=25%) and constant across our model. This low accuracy variability supports that our LSTM model can transfer across unseen participant's brain activity.

Table 1. Model Generalization results

	Base Model		Models with Weight		
Validation Video	**Validation Accuracy**	**Mean Accuracy Training**	**Validation Accuracy**	**Validation Accuracy Gain/Loss**	**Mean Accuracy Training**
1	0.3602	0.7336 (0.143)	0.3115	-0.0487	0.634 (0.152)
2	0.5584	0.7489 (0.140)	0.5121	-0.0463	0.611 (0.157)
3	0.5339	**0.7912 (0.135)**	0.4899	-0.0440	0.602 (0.154)
4	0.5588	0.7420 (0.133)	0.6110	0.0522	0.616 (0.154)
5	**0.7443**	0.7264 (0.133)	0.6018	-0.1425	**0.726 (0.133)**
6	0.3289	0.7328 (0.139)	0.3537	0.0248	0.629 (0.159)
7	0.5169	0.7466 (0.132)	0.5560	0.0391	0.622 (0.149)
8	0.5086	0.7229 (0.126)	0.5328	0.0242	0.620 (0.152)
9	0.4874	0.7155 (0.136)	0.3923	-0.0951	0.596 (0.155)
10	0.6858	0.7286 (0.131)	0.6344	-0.0514	0.643 (0.137)
11	**0.7596**	0.7372 (0.135)	0.6795	-0.0801	0.606 (0.155)
12	**0.7401**	0.7339 (0.134)	0.6450	-0.0951	0.645 (0.150)
13	**0.8015**	0.7290 (0.133)	**0.8241**	0.0226	0.632 (0.140)
14	0.6978	**0.7577 (0.137)**	**0.7326**	0.0348	0.634 (0.144)
15	**0.7840**	**0.7714 (0.149)**	0.6228	-0.1612	0.616 (0.153)
16	0.6786	0.7251 (0.134)	0.6687	-0.0099	0.667 (0.161)

3.2.3. Best model

Across our models, only one seems to both transfer across unseen videos and unseen participants. Our best model (model #13) can accurately predict the amusement level (low, medium, high, high amusement) of an unseen video based on the participant's brain activity with 82.41% accuracy. This high accuracy suggests that brain activity collected with a commercial headset can be used to predict amusement.

While our model reaches a high accuracy level, we can see from the confusion matrix (Figure 6) that our model still has difficulty distinguishing between high and very high amusement. Our model can accurately predict the low and medium levels of amusement, but high and very high amusement are still inadequately predicted. It is possible that the model cannot classify between high and very high because the brain activity is more alike in those two categories than in low and medium amusement.

Inspired by Liu [8], we believe that this problem could be resolved by first creating a model that classifies the data between three types of funniness: low, medium and high (where high is a combination of high and very high amusement). This would be followed by a second model trained to classify especially between high and very high amusement, thus increasing our model prediction.

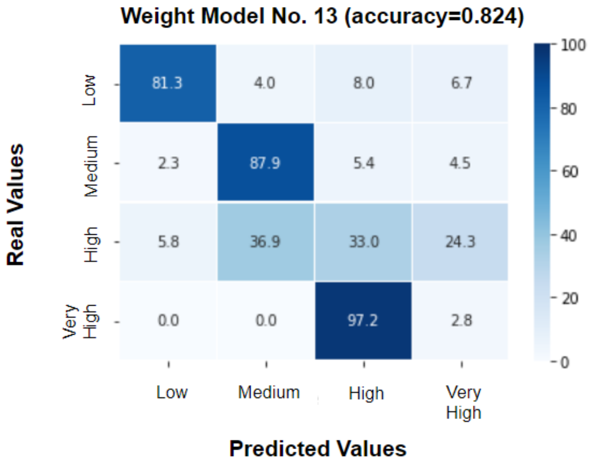


Figure 6. Confusion matrix of best weighted model during validation phase

4. Conclusion

In this study, we aimed to develop an algorithm that predicts amusement based on EEG data from a commercial headset. The objective of this paper was to develop a model that can predict amusement with high accuracy while ensuring that it is transferable across both new participants and new contents. Using an LSTM algorithm, we were able to obtain a model that can predict amusement with an accuracy of 82.4%. This high accuracy confirms that brain activity can accurately predict amusement experienced by the subject. While our model had, on average, a low variability when testing on unseen participants, models tested on unseen videos were more variable. This lets us believe that we can still improve our model. Classification of amusement in four-level (low, medium, high and very high) is our first step into creating a deep learning model that can predict amusement. In this study, we support both the use of EEG data and LSTM to predict amusement.

In our futures research, we want to improve our classification model by first creating a model that classifies the data between three types of funniness: low, medium, and high (where high is a combination of high and very high amusement). A second model would then be trained to classify between high and very high amusement. Furthermore, using the same nested-cross-validation, we will train an LSTM algorithm to predict a value between 0 (not funny) and 1 (very funny) on a continuous scale.

5. Acknowledgments

We acknowledge NSERC-CRD (National Science and Engineering Research Council Cooperative Research Development), Prompt, and BMU (Beam Me Up) for funding this work.

References

[1] Martin, R.A. (2007). The Social Psychology of Humor. In The Psychology of Humor, (Elsevier), pp. 113–152.
[2] Vrticka, P., Black, J. M., & Reiss, A. L. (2013). The neural basis of humour processing. Nature Reviews Neuroscience, 14(12), 860.
[3] Gelkopf, M., Kreitler, S., & Sigal, M. (1993). Laughter in a psychiatric ward. Somatic, emotional, social and clinical influences on schizophrenic patients. Journal of Nervous and Mental Disease, 181(5), 283–289.
[4] Hirosaki, M., Ohira, T., Kajiura, M., Kiyama, M., Kitamura, A., Sato, S., & Iso, H. (2013). Effects of a laughter and exercise program on physiological and psychological health among community‐dwelling elderly in Japan: Randomized controlled trial. Geriatrics & gerontology international, 13(1), 152-160.
[5] Zhao, J., Yin, H., Wang, X., Zhang, G., Jia, Y., Shang, B., ... & Chen, L. (2020). Effect of humour intervention programme on depression, anxiety, subjective well‐being, cognitive function and sleep quality in Chinese nursing home residents. Journal of Advanced Nursing, 76(10), 2709-2718.
[6] Savage, B. M., Lujan, H. L., Thipparthi, R. R., & DiCarlo, S. E. (2017). Humor, laughter, learning, and health! A brief review. Advances in Physiology Education.
[7] Herring, D. R., Burleson, M. H., Roberts, N. A., & Devine, M. J. (2011). Coherent with laughter: Subjective experience, behavior, and physiological responses during amusement and joy. International Journal of Psychophysiology, 79(2), 211-218.
[8] Liu, Y. J., Yu, M., Zhao, G., Song, J., Ge, Y., & Shi, Y. (2017). Real-time movie-induced discrete emotion recognition from EEG signals. IEEE Transactions on Affective Computing, 9(4), 550-562.
[9] Kim, J. H., Kim, B. G., Roy, P. P., & Jeong, D. M. (2019). Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. IEEE Access, 7, 41273-41285.
[10] Aydın, S. (2019). Deep Learning Classification of Neuro-Emotional Phase Domain Complexity Levels Induced by Affective Video Film Clips. IEEE Journal of Biomedical and Health Informatics, 24(6), 1695-1702.
[11] Stöckli, S.; Schulte-Mecklenbeck, M.; Borer, S. & Samson, A.C. (2018) Facial expression analysis with AFFDEX and FACET: A validation study. Behavior Research Methods, 50 (4), 1446-1460.
[12] Duan, R. N., Zhu, J. Y., & Lu, B. L. (2013, November). Differential entropy feature for EEG-based emotion classification. In 2013 6th International IEEE/EMBS Conference on Neural Engineering (NER) (pp. 81-84). IEEE.
[13] Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: a review. Journal of neural engineering, 16(3), 031001.
[14] Alhagry, S., Fahmy, A. A., & El-Khoribi, R. A. (2017). Emotion recognition based on EEG using LSTM recurrent neural network. Emotion, 8(10), 355-358.
[15] Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. Behavior Research Methods. 10.3758/s13428-018-01193
[16] Barral, O., Kosunen, I., & Jacucci, G. (2017). No need to laugh out loud: Predicting humor appraisal of comic strips based on physiological signals in a realistic environment. ACM Transactions on Computer-Human Interaction (TOCHI), 24(6), 1-29.
[17] LeCun, Y., Y. Bengio, and G. Hinton, Deep learning. Nature, 2015. 521(7553): p. 436-444.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210100

Ontology Reasoning for Explanatory Feedback Generation to Teach How Algorithms Work

Anton ANIKIN [a,b], Oleg SYCHEV [a], and Mikhail DENISOV [a]
[a] *Volgograd State Technical university, Russia*
[b] *Software Engineering School, Volgograd, Russia*

Abstract. Developing algorithms using control structures and understanding their building blocks are essential skills in mastering programming. Ontologies and software reasoning is a promising method for developing intelligent tutoring systems in well-defined domains (like programming languages and algorithms); it can be used for many kinds of teaching tasks. In this work, we used a formal model consisting of production rules for Apache Jena reasoner as a basis for developing a constraint-based tutor for introductory programming domain. The tutor can determine fault reasons for any incorrect answer that a student can enter. The problem the student should solve is building an execution trace for the given algorithm. The problem is a closed-ended question that requires arranging given actions in the (unique) correct order; some actions can be used several times, while others can be omitted. Using formal reasoning to check domain constraints allowed us to provide explanatory feedback for all kinds of errors students can make.

Keywords. Ontology reasoning, Constraint-based tutoring systems, Program execution trace, Error detection

1. Introduction

Ontology models and formal logic reasoning are used for knowledge representation and processing in different domains for a wide range of tasks. E.g., the Ontology Driven Software Engineering (ODSE) approach [1] implies using ontology models for various aspects of software engineering: modeling different parts of software systems, products, modules, and algorithms. In [2] the ontology model is used for declarative program analysis in software development. Most of these aspects are important in introductory programming courses as well, where ontologies are widely used for domain modeling [3,4].

One of the efficient approaches to introducing new learners to algorithms analysis and synthesis is the trace-based teaching approach that allows to decrease the dropout and grade failures by 25.49% and 8.51% respectively [5,6,7]. According to the structured programming approach, any algorithm can be represented as a tree of control structures. In the introductory programming teaching on the Problem Formulation step [6] the algorithmic reasoning skills ("a pool of abilities that are connected to constructing and understanding algorithms: to analyze given problems; to specify a problem precisely; to find basic actions that are adequate to the given problem; to construct a correct algorithm to a given problem using the basic actions" [8]) improvement is important. On the Solution Expression step, when the problem is formulated, and students should express a solution through programming structures, selecting the appropriate structures for solving the task is the main difficulty [9]. Finally, on the Solution Execution and Evaluation step, students should test and analyse the code to identify and correct problems, and code tracing activities are appropriate tasks on this step [5,10]. Explanatory feedback (e.g., error-flagging feedback [11] as well as other forms of explanations of student errors) has a significant effect on learning efficiency and results.

So, automated algorithm trace generation and analysis with explanatory feedback is an important task in introductory programming learning that can be solved using ontology domain models and formal logic reasoning. The reasoning rules allow to set the domain constraints and use these rules not only for the execution trace check for correctness but for the particular errors detection and corresponding explanation providing as well at the same time.

2. Intelligent Application to Teach Algorithms

We developed an online tutoring tool `How It Works: Algorithms`[1] using ontology reasoning to grade students' answers and generate explanatory feedback about their errors. Its input consists of an algorithm, represented as a tree of basic control structures – sequences, alternatives, and loops (see Fig. 1) – and the values of control conditions. The reasoner also receives the student-built trace as a sequence of control-statement execution acts. For complex control structures, the beginning and the end of their execution constitute separate execution acts to represent the nesting of control statements in the trace.

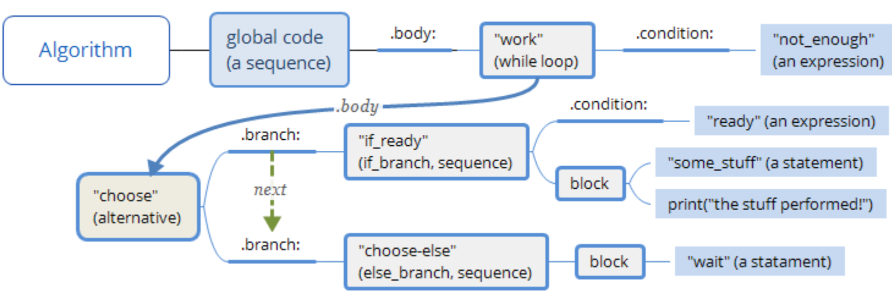


Figure 1. Fragment of algorithm represented as an abstract syntax tree

To generate explanatory feedback, we classified all the possible errors in execution-trace building creating 33 concepts to represent them (Fig. 2). The reasoning engine determines the error class and the additional information about the individuals related to the error for feedback generation.

[1] https://howitworks.app/algorithms

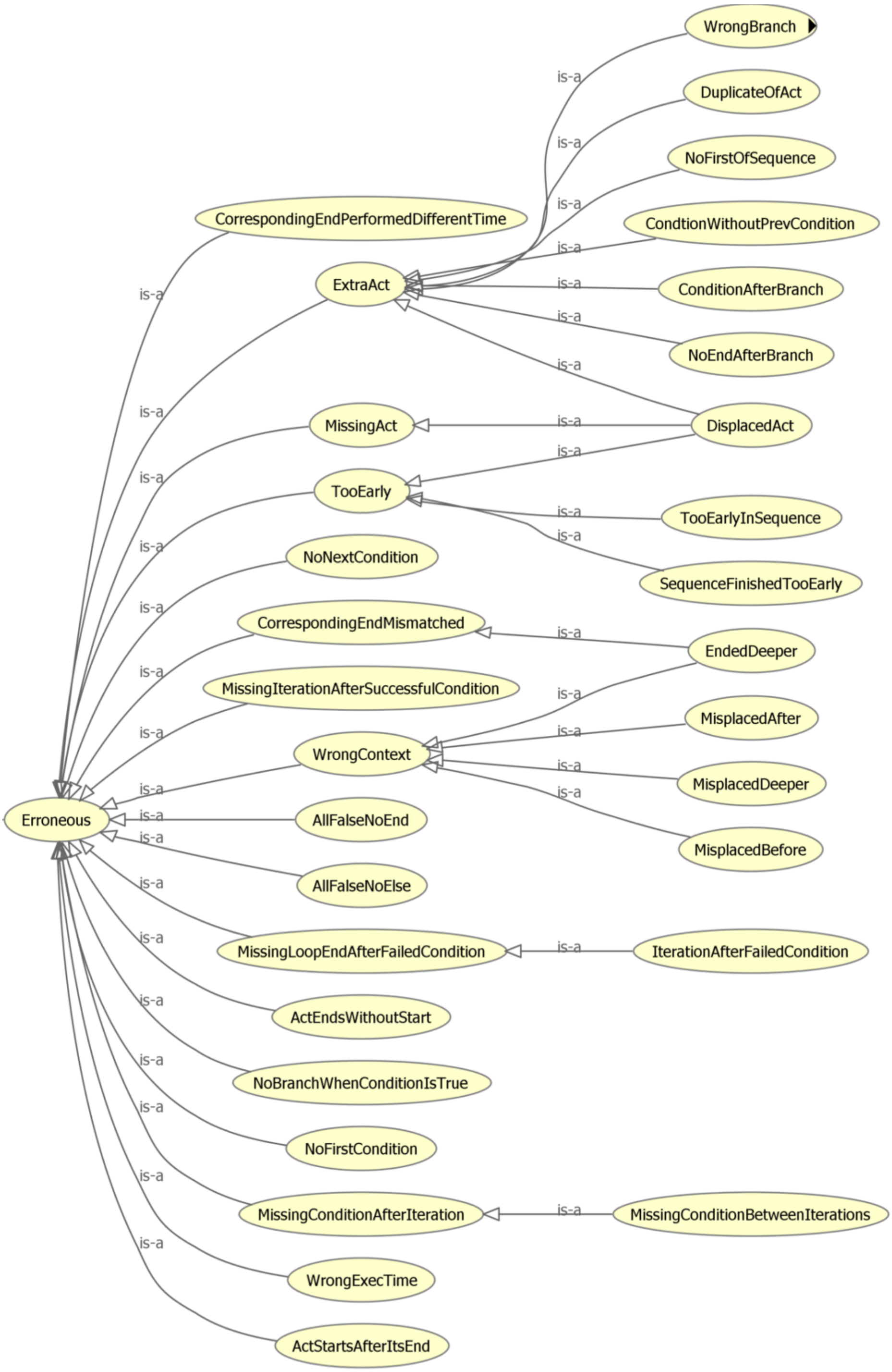


Figure 2. Concepts for classifying errors

To select a language for the reasoning rules description and appropriate reasoning engine, we performed a study of software reasoners to find the best one for our domain, comparing Pellet, Apache Jena, Apache Jena SPARQL query processor, SWI-Prolog with semweb package, and ASP (Answer Set Programming) solvers Clingo and DLV. The results show that Apache Jena performs inference quicker than other reasoners on most of the domain-specific tasks.

In particular, Apache Jena infers the correct trace and student's errors 2.4–2.9 times quicker than SWRL Pellet reasoner. Jena rules are also more expressive than SWRL, having full CRUD operations support (e.g., creation of concepts and individuals), negation support, and relation retraction.

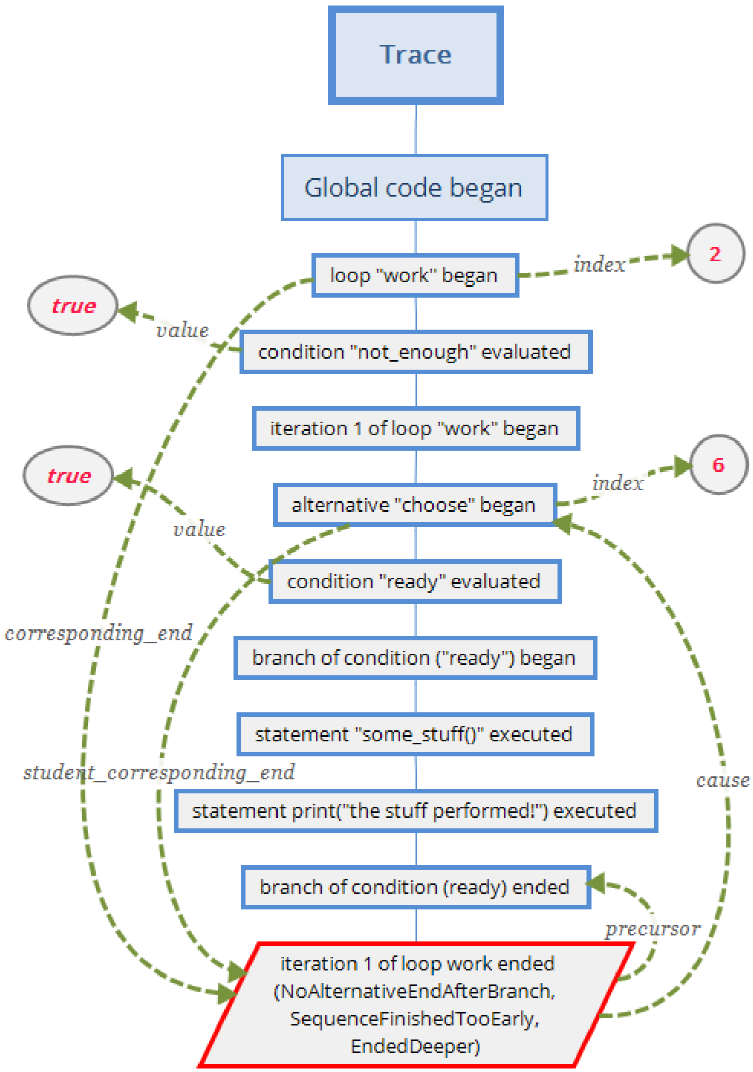


Figure 3. Execution trace processing example

So, the developed ontology contains about 30 concepts for algorithm elements, 7 trace acts, 29 kinds of errors, and 27 explanations for correct acts; 30 roles, and about 100 positive, negative, and helper rules. Using these rules, the execution trace can be generated based on the given algorithm, determine the domain constraints violated by the student, and provide all the necessary information about errors to generate explanatory feedback (Fig. 3).

The implemented software tool allows students to input the user execution trace for the teacher-defined algorithm (stored and available by URL for easy access) by clicking buttons inside the algorithm and show detailed explanatory feedback for errors made using the ontology reasoning described above (Fig. 4).

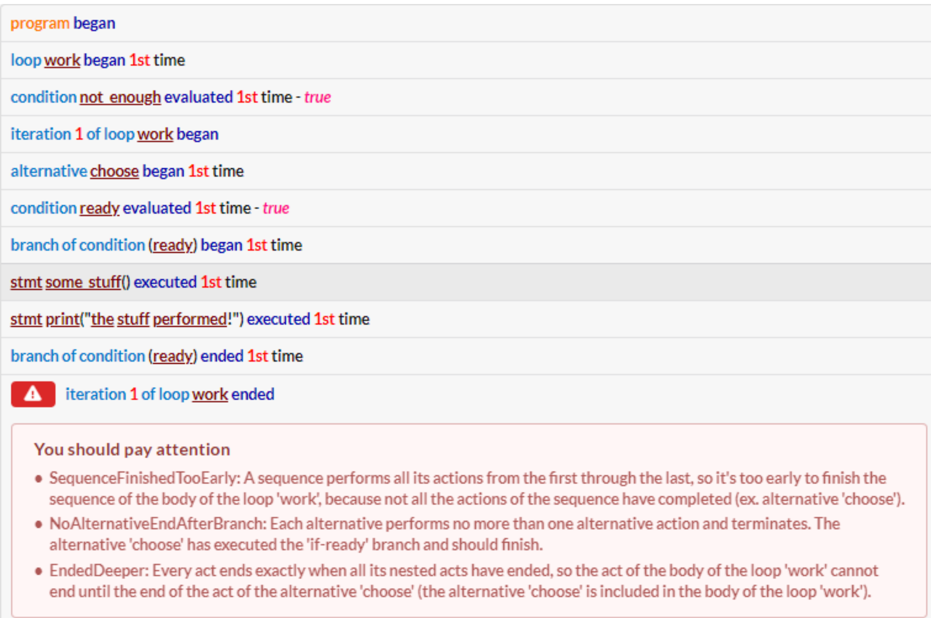


Figure 4. Explanatory feedback provided by the tutoring tool for an error in the trace

3. Conclusion and Future Work

In this study, we present an ontology with a set of reasoning rules that is able to build execution traces for a given algorithm, find errors in students' traces, and provide the necessary information to generate explanatory feedback about the violated constraints representing subject domain laws. The approach was implemented in a software tool, using Apache Jena inference engine for ontology reasoning. The usage of forward chaining RETE algorithm and Jena rules and reasoning allowed us to implement domain-specific rules with adequate performance to grade students' traces in real time step-by-step, showing feedback messages right after adding an erroneous line.

The software tool can be used as a basis for developing intelligent tutoring systems for improvement of algorithmic reasoning skills and developing understanding of program execution during introductory programming courses. The future work includes expanding the set of supported programming languages (C++, Python and Java are implemented at this moment), supporting recursive functions in the algorithms, and developing a constraint-based intelligent tutoring system based on the proposed approach for complex exercises implementation by adding learner's model and intelligent exercise selection.

Acknowledgements

The reported study was funded by RFBR, project number 20-07-00764 "Conceptual modeling of the knowledge domain on the comprehension level for intelligent decision-making systems in the learning".

References

[1] D. Strmečki, I. Magdalenić and D. Kermek, An Overview on the use of Ontologies in Software Engineering, *Journal of Computer Science* **12**(12) (2016), 597–610. doi:10.3844/jcssp.2016.597.610.

[2] Y. Zhao, G. Chen, C. Liao and X. Shen, Towards Ontology-Based Program Analysis, in: *30th European Conference on Object-Oriented Programming (ECOOP 2016)*, S. Krishnamurthi and B.S. Lerner, eds, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 56, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2016, pp. 26:1–26:25. doi:10.4230/LIPIcs.ECOOP.2016.26.

[3] M. Kultsova, A. Anikin, I. Zhukova and A. Dvoryankin, Ontology-based learning content management system in programming languages domain, *Communications in Computer and Information Science* **535** (2015), 767–777. doi:10.1007/978-3-319-23766-4_61.

[4] X.-T. Pham, T.-V. Tran, V.-T. Nguyen-Le, V.T. Pham and H.D. Nguyen, Build a search engine for the knowledge of the course about Introduction to Programming based on ontology Rela-model, in: *2020 12th International Conference on Knowledge and Systems Engineering (KSE)*, IEEE, 2020. doi:10.1109/kse50997.2020.9287775.

[5] M. Hertz and M. Jump, Trace-based teaching in early programming courses, in: *Proceeding of the 44th ACM technical symposium on Computer science education - SIGCSE '13*, ACM Press, 2013. doi:10.1145/2445196.2445364.

[6] R.P. Medeiros, G.L. Ramalho and T.P. Falcao, A Systematic Literature Review on Teaching and Learning Introductory Programming in Higher Education, *IEEE Transactions on Education* **62**(2) (2019), 77–90. doi:10.1109/te.2018.2864133.

[7] O. Sychev, M. Denisov and A. Anikin, Ways To Write Algorithms And Their Execution Traces For Teaching Programming, in: *EpSBS - Volume 102 - NININS 2020*, European Publisher, 2021, pp. 1029–1039. doi:10.15405/epsbs.2021.02.02.127.

[8] T. Wang, X. Su, P. Ma, Y. Wang and K. Wang, Ability-training-oriented automated assessment in introductory programming course, *Computers & Education* **56**(1) (2011), 220–226. doi:10.1016/j.compedu.2010.08.003.

[9] S. Xinogalos, Designing and deploying programming courses: Strategies, tools, difficulties and pedagogy, *Education and Information Technologies* **21**(3) (2014), 559–588. doi:10.1007/s10639-014-9341-9.

[10] A.N. Kumar, Solving Code-tracing Problems and its Effect on Code-writing Skills Pertaining to Program Semantics, in: *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, ACM, 2015. doi:10.1145/2729094.2742587.

[11] A.N. Kumar, Allowing Revisions While Providing Error-Flagging Support: Is More Better?, in: *Lecture Notes in Computer Science*, Springer International Publishing, 2020, pp. 147–151. doi:10.1007/978-3-030-52240-7_27.

Computer Vision

This page intentionally left blank

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210102

Real-Time Face Mask Detector Using Convolutional Neural Networks Amidst COVID-19 Pandemic

Efstratios KONTELLIS[1], Christos TROUSSAS, Akrivi KROUSKA, Cleo SGOUROPOULOU
Department of Informatics and Computer Engineering
University of West Attica
Greece

Abstract. The COVID-19 pandemic provoked many changes in our everyday life. For instance, wearing protective face masks has become a new norm and is an essential measure, having been imposed by countries worldwide. As such, during these times, people must wear masks to enter buildings. In view of this compelling need, the objective of this paper is to create a real-time face mask detector that uses image recognition technology to identify: (i) if it can detect a human face in a video stream and (ii) if the human face, which was detected, was wearing an object that it looked like a face mask and if it was properly worn. Our face mask detection model is using OpenCV Deep Neural Network (DNN), TensorFlow and MobileNetV2 architecture as an image classifier and after training, achieved 99.64% of accuracy.

Keywords. Virus Protection, COVID-19, Face Detection, Deep Learning, Convolutional Neural Network, Computer Vision, MobileNetV2

1. Introduction

The COVID-19 pandemic has forced many countries around the world to initiate new rules for face mask-wearing and social distancing. Also, governments have started working on new strategies [1] to stop the spreading of the virus by forcing people to wear face masks. The goal of wearing face masks is to lessen the transmission and spreading rate. As COVID-19 is transmitted [2] through air drops and close contact, more and more people are concerned about their health and public health is considered as top priority for governments.

According to World Health Organization (WHO) directions [3], the wearing of face masks appears as a solution for limiting the spread of COVID-19. That is the reason why face mask detection has become a trending application [4] due to the pandemic, to monitor that people are following this basic safety principle. Artificial Intelligence (AI) [5] based on Machine Learning and Deep Learning can help humanity fight against COVID-19 in many ways. For example, AI can be used to track and to predict how the disease will spread over time and space. Face mask detection means to identify whether a person is wearing a mask or not. There are two steps that should be implemented to achieve this. The first step is to detect human faces in a video stream and the second step

[1] Corresponding Author; E-mail: stratis96@outlook.com.

is to identify if the human face that was detected was wearing an object that it looked like a face mask and was properly worn. Face detection is one of the most popular applications of object detection and can be used in many areas like security, face emotions, biometrics, law enforcement and more.

In view of the above, the main objective of the research is to detect whether someone is wearing his face mask correctly to protect himself and the public health from COVID-19 spreading. To perform this task, a large dataset of masked and unmasked faces is necessary for training deep learning models towards detecting people wearing masks and those not wearing masks. However, at the moment, there are not large datasets available of masked faces and because of that, we applied a method called "Data Augmentation" in our training data. Our face mask detection model is using OpenCV Deep Neural Network (DNN), TensorFlow and MobileNetV2 [6] architecture as an image classifier. MobileNetV2 consists of Convolutional Neural Networks [7-11], which are ideal for image classification. The model is trained for 15 epochs which maintains a trade-off between accuracy and chances of overfitting. Finally, thanks to our accurate dataset, our trained model reaches 99,64% of accuracy.

2. Related Work

Object detection from an image is probably the deepest aspect of computer vision due to widely used in many cases such as, face recognition, neurobiology, robotic navigation etc. There has been supervised and unsupervised based learn in the field of computer vision to outfit the work of object detection in an image. This section conducts the recent academic papers for applying representative works related to object detection based on deep learning for the medical face mask. Most of face mask detection methods focus on face construction and face detection based on traditional machine learning techniques. In this paper, our focus is on detecting and finding the human who is wearing a face mask or not, to help in lessening the spreading of COVID-19.

In face detection method, a face is detected from an image that has several attributes in it. According to [12], research into face detection requires expression recognition, face tracking and pose estimation. Given a solidary 2D image, the challenge is to identify the face from the picture. Face detection is a difficult errand because faces change in color, size, shape, etc. and they are not immutable. In [13], the authors present occlusive face detection coming with two major challenges: (1) unavailability of sizably voluminous datasets containing both masked and unmasked faces, and (2) exclusion of facial expression in the covered area. According to the work reported in [14], convolutional neural networks (CNNs) in computer vision comes with a strict constraint regarding the size of the input image. The prevalent practice reconfigures the images before fitting them into the network to surmount the inhibition.

In [15], the authors presented a system for face mask detection that attains accuracy of 95.77%. They used TensorFlow, OpenCV and Data augmentation to reach this result. Another research [16] presented a DPM-based face mask detector using around 30,000 faces divided into masked and unmasked. This approach achieved an exceptional accuracy of 97.14 %. In [17], the authors presented a system for detecting the presence of a compulsory medical mask in public places. The proposed system achieved 97.1% accuracy.

Analyzing the related literature and towards advancing the literature, this paper presents a face mask detection model which has been developed using deep neural network modules from OpenCV and TensorFlow, containing a Single Shot Multibox Detector object detection model. Fine-tuned MobileNetV2 classifier has been used, which has been an improvement over MobileNetV1 architecture classifier. In conclusion, MobileNetV2 architecture (Table 1) is made up of 17, 3 x 3 convolutional layers in a row accompanied by a 1 x 1 convolution, an average layer of max pooling and a layer of classification. Also, thanks to the accurate dataset of masked and unmasked faces, our proposed method attains accuracy up to 99.64%.

Table 1. MobileNetV2 architecture

Input	Operator	t	c	n	s
224^2 x 3	conv2d	-	32	1	2
112^2 x 32	bottleneck	1	16	1	1
112^2 x 16	bottleneck	6	24	2	2
56^2 x 24	bottleneck	6	32	3	2
28^2 x 32	bottleneck	6	64	4	2
14^2 x 64	bottleneck	6	96	3	1
14^2 x 96	bottleneck	6	160	3	2
7^2 x 160	bottleneck	6	320	1	1
7^2 x 320	conv2d 1x1	-	1280	1	1
7^2 x 1280	avgpool 7x7	-	-	1	-
1 x 1 x 1280	conv2d 1x1	-	k	-	

where t: expansion factor, c: number of output channels, n: repeating number, s: stride. 3×3 kernels are used for spatial convolution.

3. Proposed Method

To predict whether a person has worn a face mask correctly, there are two phases as mentioned above. Firstly, an accurate face detection model is required to detect faces in real-time. Details about this model have been discussed above in Section 3.4. Secondly, a proper database can help us to train our model to reach high accuracy. Furthermore, to raise the accuracy of mask detection without being too resource-heavy, our classifier uses a pre-trained model MobileNetV2 to recognize whether a person is wearing a mask or not. Devices like smartphones and the Raspberry Pi are benefited from this approach. The methodology used in this paper is depicted in Figure 1.

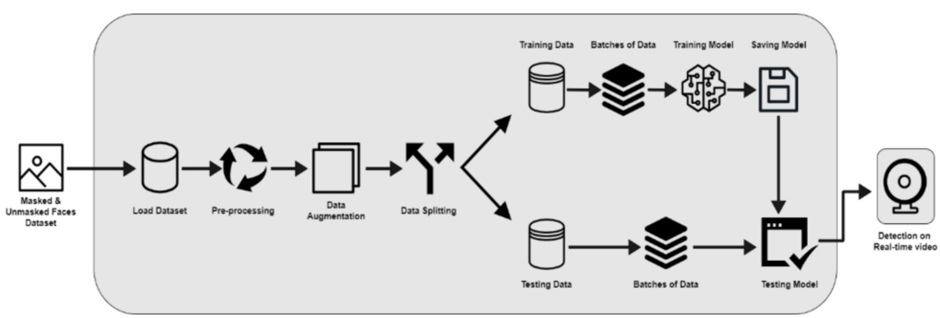


Figure 1. Flow Diagram of our model.

3.1. Dataset used

Until today, there are only a few images with masked people available for the detection of face masks [18]. Some of them are artificially created, which do not represent the real world accurately, while others are blurred with low resolution. By choosing the right dataset that contains the right images require a little effort [19]. The dataset used for training the model consists of 4000 total images out of which 2000 are of masked faces having the label "with_mask" and 2000 are of unmasked faces having the label "without_mask" [20]. All the images are actual images extracted from Kaggle datasets and cover diverse races i.e., Asian, Southern Europeans etc. As a result, the distribution between the two categories is visualized in Figure 2.

In our approach, we have dedicated 80% of the dataset as the training data and the remaining 20% as the testing data, which makes the split ratio as 0.8:0.2 of train to test set.

Figure 2. Samples of our dataset.

3.2. Pre-processing

There are four steps in the pre-processing phase which are resizing image size, converting the image into an array, pre-processing input using MobileNetV2 and the last is performing hot encoding on labels. But, before that, we should evaluate our dataset. Since a good dataset dictates the accuracy of the trained model, it is a vital part to manually remove corrupt images and repetitions. Cleaning, identifying, and correcting errors in a dataset removes adverse effects from any predictive model.

This phase explains the procedure of pre-processing the data and then training on data. The resizing image is a critical step in computer vision due to effectiveness of training models. The smaller the image size is, the better the model will perform. In this research, all the images are resized into 224x224 pixels.

The next step is to process all the images from the dataset into a NumPy array for faster calculation. After this, the process of data augmentation is applied to increase the training data as we have been discussed above in Section 3.3. Additionally, the images are used to pre-process input using MobileNetV2.

The last step in this phase is performing hot encoding on labels because a lot of machine learning algorithms cannot operate on data labeling directly. They require all the input and output labels to be numeric.

3.3. Data augmentation

For the training of our model, an enormous quantity of data is required to perform training effectively. Due to the non-availability of an adequate amount of data, the method of data augmentation [21] is used to solve this issue because neural networks are only as good as the data, we feed them. In this technique, methods like flipping, rotation, scaling, cropping, and translate the picture are used for generating numerous versions of a similar picture.

3.4. Face detection using OpenCV

Before detecting whether a person is wearing his face mask or not, we have to use a model called "Caffemodel" [22] to detect human faces in real-time using OpenCV. This model is a deep learning framework developed as an even faster and efficient alternative as compared to other object detection methods, i.e., Haar Cascades, and is created and managed by Berkeley AI Research (BAIR) and community contributors. Wearing a facial mask means that an object that it looks like a mask, is hiding some facial characteristics, i.e., nose, mouth etc. This means that some algorithms and pre-trained models may not be able to detect human faces with facial masks. For this reason, because the Caffemodel relies on deep learning neural networks, it makes it an ideal choice for masked or unmasked face detection. To implement this, the Caffemodel and prototxt files were loaded using *cv2.dnn.readNet* ("*path/to/prototxtfile*", "*path/to/caffemodel-weights*"). After applying the face detection model, we get the number of faces detected and the location of their bounding boxes. These video frame outputs are then used as input for the face mask classifier. Using this approach to detect masked or unmasked faces allows for real-time detection without much resource usage and it can also detect human faces in different orientations i.e., top, bottom, right and left with good accuracy.

3.5. Classification of images using MobileNetV2

MobileNetV2 is a Deep Neural Network based on Convolutional Neural Network architecture that has been deployed for the classification problem. Also, it aims to run very efficiently on mobile devices and it can be used as a basic image classifier or as a feature extractor. In our case, pretrained weights of ImageNet were loaded from TensorFlow and then, the base layers of our network are frozen to avoid impairment of already learned features. Subsequently new trainable layers are added and these are trained on collected dataset to classify features of a masked face from an unmasked face. Then, the model is fine-tuned, and then the weights are saved. By using pre-trained models helps to avoid needless computational costs and helps in taking advantage of already biased weights without losing any learned features (Figure 3).

4. Training

After the pre-processing phase, the model needs to be trained using our dataset and then to be tested against a different dataset. To do this, the data is split into two batches, which 80% of them is training data and the rest 20% is testing data. Each batch is containing both of masked and unmasked faces. The final step before training the model, is to set the hyperparameters. In machine learning, a hyperparameter is a parameter whose value is used to control the learning process. By contrast, the values of other parameters are derived via training. Initially, we trained with different values of hyperparameters by changing one and keeping the others constant and noted down the results in each case. By this way, we could select the hyperparameters that produced better performance though evaluation metrics. We have chosen the hyperparameters as follows: initial learning rate is taken as 0.0001, batch size is taken to be 32 and number of epochs as 15. The model is trained for 15 epochs which maintains a trade-off between accuracy and chances of overfitting.

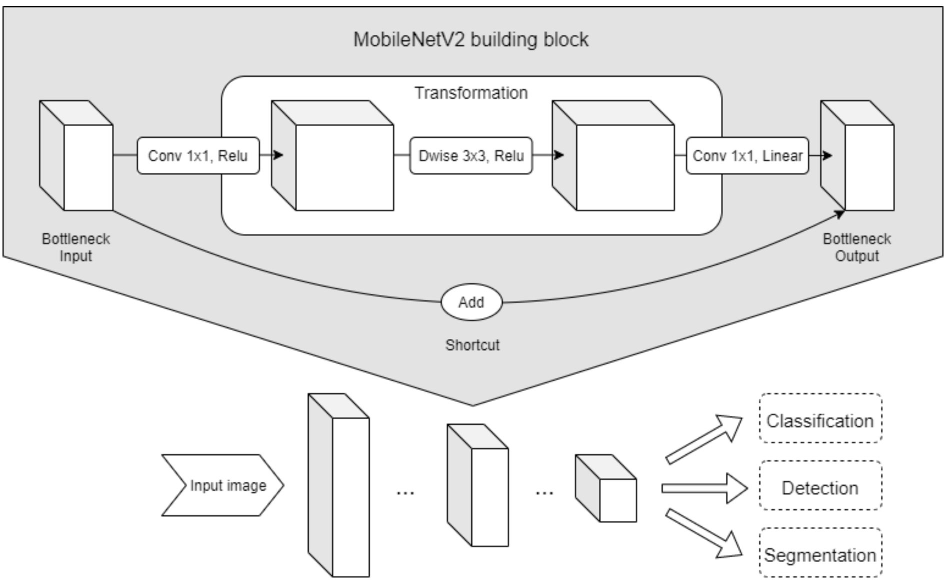


Figure 3. MobileNetV2 Convolutional Blocks.

5. Experimental Results

Our proposed model has simulated using Python and TensorFlow. The metrics selected for evaluation of our model are explained below.

$$\text{Accuracy} = \frac{(\text{Tp} + \text{Tn})}{(\text{Tp} + \text{Fp} + \text{Fn} + \text{Tn})}$$

$$\text{Precision} = \frac{\text{Tp}}{(\text{Tp} + \text{Fp})}$$

$$\text{Recall} = \frac{\text{Tp}}{(\text{Tp} + \text{Fn})}$$

$$\text{f1 score} = 2 * \frac{(\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})}$$

where Tp is the count of True positive samples, Tn is the count of True negative samples, Fp is the count of False positive samples and Fn is the count of False negative samples. Also, “Precision” is the number of correct predictions over how many occurrences of that class were in the test dataset and “Recall” is the actual true positives over how many times the classifier predicted that class. The “f1 score” can be interpreted as a weighted average of the precision and recall, where an f1 score reaches its best value at 1 and worst score at 0.

Corresponding to our dataset, the method attains accuracy up to 99,64% (Figure 4).

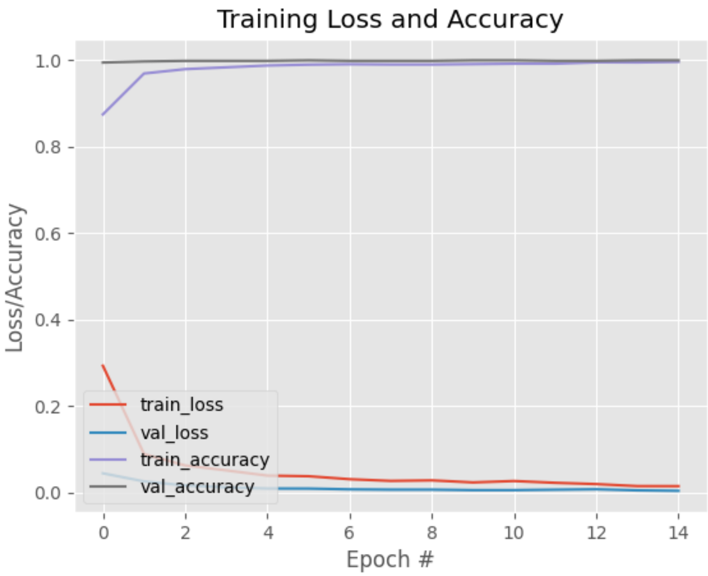


Figure 4. Training Results.

Table 2. Classification report

	Precision	Recall	f1 score	Support
with_mask	1.00	1.00	1.00	388
without_mask	1.00	1.00	1.00	414
accuracy			1.00	802
Macro average	1.00	1.00	1.00	802
Weighted average	1.00	1.00	1.00	802

Our real-time face mask detection model follows the procedure below. First of all, the video stream is scanning each frame to detect human faces. If a face is detected, it proceeds to the next process. From detected frames containing human faces, reprocessing will be carried out including resizing the image size, converting into an array and pre-processing input using MobileNetV2. The next step is predicting input data from the trained model. Afterward, the video frame will be labeled that the person is wearing or a mask or not. A testing example is depicted in Figure 5.

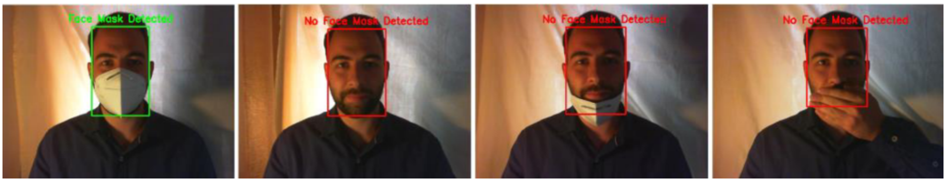


Figure 5. Real-time face mask detection results.

6. Conclusion

Due to the outbreak of COVID-19 pandemic, all the countries worldwide have imposed people to wear a face mask as a compulsory measure to prevent the spread of coronavirus. Manual observation of the face mask in crowded places is a critical task. To address this compelling need, this paper presents a real-time face mask detector. In this approach, we trained our model using MobileNetV2 with masked and unmasked faces. Furthermore, thanks to our accurate dataset, our model understands whether a face mask is properly worn. Later, we also applied this model into a real-time video to check our model's fps performance and accuracy. As future work, different types of face masks should be taken into consideration to expand the types of detecting objects and a hybrid model using deep learning techniques will be created to further investigate its accuracy in face masks detection.

References

[1] M. Loey, G. Manogaran, M. H. N. Taha, N. E. M. Khalifa, Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection, *Sustainable Cities and Society* **65** (2021).

[2] I. B. Venkateswarlu, J. Kakarla and S. Prakash, Face mask detection using MobileNet and Global Pooling Block, 2020 IEEE 4th Conference on Information & Communication Technology (2020), 1-5.

[3] Y. Chen et al., Face Mask Assistant: Detection of Face Mask Service Stage Based on Mobile Phone, *IEEE Sensors Journal* **21** (2021), 11084-11093.

[4] S. A. Sanjaya and S. Adi Rakhmawan, Face Mask Detection Using MobileNetV2 in The Era of COVID-19 Pandemic, *2020 IEEE International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy* (2020), 1-5.

[5] M. Loey, G. Manogaran, M. H. N. Taha, N. E. M. Khalifa, A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic, *Measurement* **167** (2021).

[6] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, MobileNetV2: Inverted Residuals and Linear Bottlenecks, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018) 4510-4520.

[7] F. Giannakas, C. Troussas, I. Voyiatzis, C. Sgouropoulou, A deep learning classification framework for early prediction of team-based academic performance, *Applied Soft Computing* **106** (2021).
[8] C. Troussas, F. Giannakas, C. Sgouropoulou, I. Voyiatzis, Collaborative activities recommendation based on students' collaborative learning styles using ANN and WSM, *Interactive Learning Environments* (2020). DOI: 10.1080/10494820.2020.1761835.
[9] C. Troussas, A. Krouska, M. Virvou, A multilayer inference engine for individualized tutoring model: adapting learning material and its granularity, Neural Computing and Applications (2021). https://doi.org/10.1007/s00521-021-05740-1.
[10] A. Krouska, C. Troussas, M. Virvou, *Deep Learning for Twitter Sentiment Analysis: The Effect of Pre-trained Word Embedding*, Tsihrintzis G., Jain L. (eds) Machine Learning Paradigms, Learning and Analytics in Intelligent Systems 18, Springer, Cham, 2020. https://doi.org/10.1007/978-3-030-49724-8_5.
[11] C. Troussas, C. Sgouropoulou, *Innovative Trends in Personalized Software Engineering and Information Systems - The Case of Intelligent and Adaptive E-learning Systems",* Frontiers in Artificial Intelligence and Applications, IOS Press, 324, Netherlands, 2020.
[12] D. Meena, R. Sharan, An approach to face detection and recognition, *2016 IEEE International Conference on Recent Advances and Innovations in Engineering* (2016), 1-6.
[13] S. Ge, J. Li, Q. Ye, Z. Luo, Detecting Masked Faces in the Wild with LLE-CNNs, *2017 IEEE Conference on Computer Vision and Pattern Recognition* (2017), 426-434.
[14] S. Ghosh, N. Das, M. Nasipuri, Reshaping inputs for convolutional neural network: Some common and uncommon methods, *Pattern Recognition* **93** (2019).
[15] A. Das, M. Wasif Ansari, R. Basak, Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV, *2020 IEEE 17th India Council International Conference* (2020), 1-5.
[16] K. Zhang, Z. Zhang, Z. Li, Y. Qiao, Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks, *IEEE Signal Processing Letters* **23** (2016), 1499-1503.
[17] A. Oumina, N. El Makhfi and M. Hamdi, Control The COVID-19 Pandemic: Face Mask Detection Using Transfer Learning, *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science* (2020), 1-5.
[18] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, J. Hemanth, SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2, *Sustainable Cities and Society* **66** (2021).
[19] M. Jiang, X. Fan, RetinaMask: A Face Mask detector, Hong Kong, 2020. https://arxiv.org/abs/2005.03950.
[20] M. R. Bhuiyan, S. A. Khushbu, M. S. Islam, A Deep Learning Based Assistive System to Classify COVID-19 Face Mask for Human Safety with YOLOv3, *2020 11th IEEE International Conference on Computing, Communication and Networking Technologies* (2020), 1-5.
[21] N. C. Ristea, R. T. Ionescu, Are you wearing a mask? Improving mask detection from speech using augmentation by cycle-consistent GANs, Romania, 2020. https://arxiv.org/abs/2006.10147.
[22] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, Caffe: Convolutional Architecture for Fast Feature Embedding, UC Berkeley EECS, 2014. https://arxiv.org/abs/1408.5093.

Novelties in Intelligent Digital Systems
C. Frasson et al. (Eds.)

doi:10.3233/FAIA210103

Unsupervised Man Overboard Detection Using Thermal Imagery and Spatiotemporal Autoencoders

Nikolaos Bakalos[a], Iason Katsamenis[a], Eleni Eirini Karolou[a] and Nikolaos Doulamis[a]
[a]*National Technical University of Athens*

Abstract. Man overboard incidents in a maritime vessel are serious accidents where the rapid detection of the even is crucial for the safe retrieval of the person. To this end, the use of deep learning models as automatic detectors of these scenarios has been tested and proven efficient, however, the use of correct capturing methods is imperative in order for the learning framework to operate well. Thermal data can be a suitable method of monitoring, as they are not affected by illumination changes and are able to operate in rough conditions, such as open sea travel. We investigate the use of a convolutional autoencoder trained over thermal data, as a mechanism for the automatic detection of man overboard scenarios. Morever, we present a dataset that was created to emulate such events and was used for training and testing the algorithm.

Keywords. Man overboard; Human detection; Deep learning Computer, thermal image processing

1. Introduction

Man overboard refers to an emergency scenario where a ship passenger or crew member has fallen off the vessel and into the sea. With a mortality rate of 79%, an estimate of 22 people annually lose their lives due to such incidents[1], with the majority of them being untrained passengers. The high mortality is caused by the low speed of detection and retrieval, combined by the usual low temperature and rough conditions of the waters that can quickly result in drowning or hypothermia. Thus, the use of intelligent systems is imperative, in order to continuously monitor for such incidents and raise timely alerts. To this end, models based on deep learning paradigms used for the analysis of video streams have displayed great performance.

However, even these approaches have some drawbacks, as they rely on the use of RGB video streams, i.e. data streams monitoring over the visible spectrum. While the use of such data is popular, due to the cost efficiency of installing normal video surveillance systems, and the high performance of algorithms for object detection and classification over such data, these streams are greatly affected by illumination changes, and poor visual conditions. This indicates that the use of additional or alternative data modalities is needed. A valid alternative is video streams using thermal capturing devices. These devices monitor the infrared spectrum and are not affected by the change of lighting.

1.1. Previous Work

In a universal maritime surveillance system, human detection is a key issue and must be completely independent of the environment as well as light and weather conditions. Several human detection methods have been presented in the literate and have emphasized the importance of real-time home surveillance systems ([2], [3]) that focus on fall detection through visual sensors, deep learning and computer vision applications (e.g [[4][5][6]), however, little work has been presented in the literature on the man overboard situation.

In essence though the incident can be modelled as an abnormal behavior detection problem, where the normal situation consists of a normal capturing a seafaring vessel, while the abnormality would be the capturing of a fall. To this end, the main approaches for abnormal event recognition involve either the use of supervised deep learning techniques to learn a dictionary of abnormal sub-events or unsupervised outlier detection techniques. in many applications [7]-[9]. Examples include surveillance in industrial environments [7] or critical infrastructures [9] for safety/security and quality assurance, traffic flow management [10] and intelligent monitoring of public places [11]

Regarding outlier detection, the works of [12], [[13], [14] learn dictionary of sub-events, through a training process, and then those events that do not lie in the partitioned sub-space are marked as abnormal ones.

Regarding deep learning, the work of [15] employs convolutional auto-encoders (ConvAE) to learn temporal regularity in videos, while auto-encoders are exploited in [16] to learn feature and reconstruct the input images. Then, one-class Support Vector Machines (SVMs) are used for detecting the abnormal events. The work of [17] introduces a hybrid scheme which aggregates ConvAE with Long Short-Term Memory (LSTM) encoder-decoder. Recently, deep generative models have been applied [15]-[17]. These models are trained to produce normal events while the abnormal ones are given as the difference between the original frames and the generated ones.

Computer-vision tools that operate outside of the visible spectrum (i.e., thermal sensors) are also gaining traction in this context, because they are not significantly affected by illumination changes [18]. However, such approaches do not capture texture or color information. Vision techniques focus on background and target modeling [8], object tracking [19], activity recognition [20], crowd dynamics, and identification of unusual and suspicious behavior [21]. These approaches seek to detect abnormalities in crowded environments by analyzing actions on both the spatial and temporal scales. Detailed surveys about video-based abnormal activity recognition have been published [23], [24].

Recently, unsupervised learning models are utilized for abnormal event detection. In [25], the anomalies in videos are scored independently of temporal ordering and without any training by simply discriminating between abnormal frames and the normal ones. Other approaches exploit on-line incremental coding [26], deep cascading neural networks, and unmasking (a technique previously used for authorship verification in text documents) [27]. Recently, the works of [28] and [29] incorporate autoencoders and supervised learning for abnormal event detection. Other approaches employ tracking algorithms to extract salient motion information which is then classified either as normal or abnormal [30], [31]. However, tracking fails in complex visual scenes where multiple humans are present.

In this paper we present the use of an unsupervised fall detection method for man overboard scenarios. Our approach is based on the use of convolutional spatiotemporal autoencoders trained using a thermal imagery dataset that simulates man overboard incidents.

2. Proposed System Architecture

The presented system using only thermal. video streams to identify overboard falls. Each image property was fed into the spatiotemporal autoencoder. Autoencoders are a type of Neural Network that manage to learn efficient data encodings by training the network to ignore signal noise. Their usefulness comes from the fact that they are trained in an unsupervised manner. They are essentially composed from two main components that are trained in parallel. The dimensionality reduction component aims at extracting an efficient encoding of the input signal, while the reconstruction side tries to generate from the reduced encoding a representation as close as possible to the original input. To identify the abnormalities, the reconstruction error of each autoencoder was monitored, and when the error was bigger than a predefined threshold, an alert was raised. The selection of the threshold took place during the training, to identify the exact value that maximized detection performance.

The autoencoders used for each image property had the structure presented in **Figure 1**. Each thermal frame was reduced to a grayscale image with a resolution of 227x227x1. A 10 frame batch was used for the analysis.

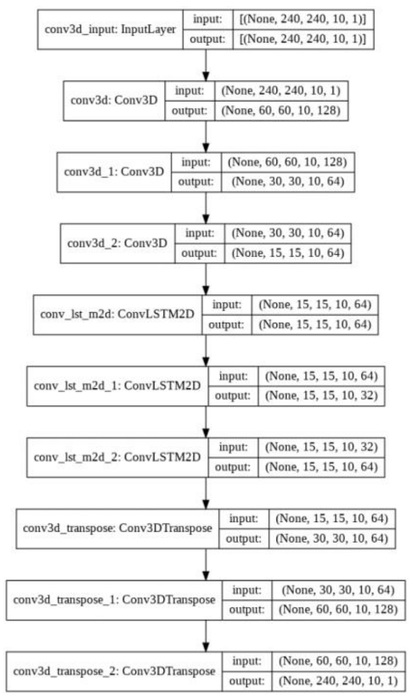


Figure 1. Proposed Model Structure

3. Dataset Description

In order to train and evaluate the proposed methodology, a mock man-overboard event was conducted that concerned the fall of a human-sized dummy from the balcony of a high-rise building. In particular, the human dummy, weighting 30 Kg, was thrown from an approximate height of 20 meters, which is roughly equivalent to two seconds of free-falling. For the needs of the experiment, we made 320 test throws of the dummy, to simulate a man-overboard event (see Figure 2(a)-(d)). Additionally, we recorded several videos without dropping the dummy as well as numerous throws of various objects, such as plastic bags and bottles (see Figure 2(e)). This way we can implement deep learning models that are not prone to false-positive alarms, triggered by non-human-related events.

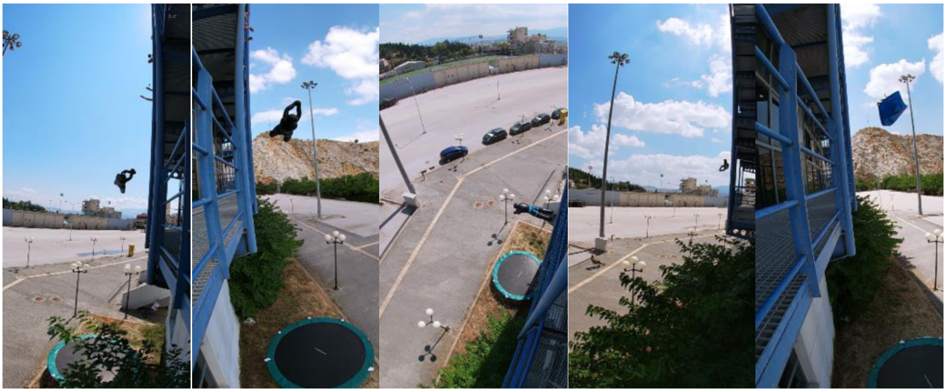

(a) (b) (c) (d) (e)

Figure 2. Test throws during the data collection experiments. The free fall (a)-(d) of the human dummy from different shooting angles (positive event), and (e) of a plastic bag (negative event).

The experiments took place in the surrounding area of Nikaia Olympic Weightlifting Hall, and lasted five days. Due to the fact that the test throws were carried out throughout the whole day, from 9:00 AM to 5:00 PM, the acquired videos vary in terms of illumination conditions (e.g. underexposure, overexposure). Additionally, we shot under various weather conditions (e.g. sunny, cloudy, rainy, windy, hot, cold), thus providing further variations in the background of the event.

In this paper, we are using a dataset consisted of RGB videos featuring the free falls of the dummy (see Figure 2(a)-(d)). For the dataset collection, which contains video sequences with a resolution of 1080 × 1920 pixels, we used a GoPro Hero 7 Silver. The camera was set to shoot at a high frame rate, at 50 frames per second, to ensure sufficient acquisition of data that concerns the critical event.

It is underlined that to avoid training bias and guarantee replicability of the results to other datasets, we placed the sensor in four different locations of the building, in order to obtain data that vary in terms of background, illumination, shooting angle, and distance (see Figure 2(a)-(d)). In particular, as depicted in Figure 3, we placed the RGB camera (i) on the left of the fall at a close distance of 7m (see Figure 2((a)), (ii) on the right of the fall at a close distance of 5m (see Figure 2((b)), (iii) on the top left of the fall at an angle of roughly 45° (see Figure 2((c)), and (iv) to the left of the fall at a long distance of 13m (see Figure 2((d)). It is emphasized that to further generalize the learning procedure, we augmented the training data by horizontally flipping the corresponding videos.

Figure 3. The four locations of the building that the optical sensor was placed, during the data acquisition experiments.

4. Experimental Evaluation

The proposed method was implemented in Python, using the Tensorflow and Keras libraries. The implementation used Python 3.6 and the Keras (1.08) and Tensorflow (2.1.0) machine learning libraries, in combination with a number of other scientific and data management libraries. The model was trained using the Google Collab Platform.
The Area Under Curve (AUC) metric was employed in assessing the performance of the proposed method and the compared ones. The AUC is computed with regard to ground-truth annotations at the frame-level and it is a common metric for many abnormal event detection methods. It measures the ability of the learning algorithm to correctly distinguish normal from abnormal events and summarises the ROC curve of the system, i.e. the probability curve that plots the raising a true alert (true positive rate) and a false alarm (false positive rate) at various thresholds. Our algorithm achieves an AUC score of 88.. Due to the fact that there are no similar publications for fall detection in man overboard scenarios, at least to the authors knowledge, a comparative analysis of the performance is hard to achieve. The performance of our system using these metrics can be viewed in **Figure 4**.

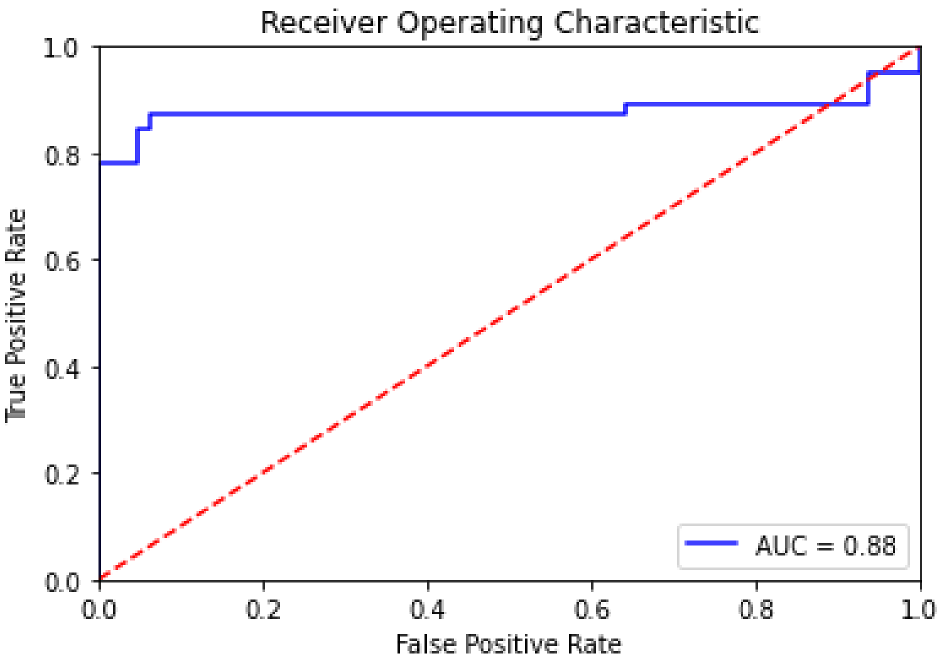


Figure 4: ROC Curve of spatiotemporal autoencoder

5. Conclusions

In this paper, we presented and evaluated a learning algorithm for man overboard detection over thermal data frames. The employed techniques use a deep machine learning framework, modelling a man overboard incident as an abnormal action recognition one. We then proceed in identifying falls by the autoencoders' success or failure to reconstruct a scene due to the presence of abnormal events.

Future work should include the fusion with additional imaging modules, such as normal RGB frames, and the studying of additional ways for inter and intra property encoding of all the available modalities to maximize the detection capabilities.

Acknowledgments

This paper is supported by the Greek Funded Project MHTIS No. 01169.

References

[1] E. Örtlund and M. Larsson, "Man Overboard detecting systems based on wireless technology," 2018.

[2] A. S. Voulodimos, D. I. Kosmopoulos, N. D. Doulamis, and T. A. Varvarigou, “A top-down event-driven approach for concurrent activity recognition,” Multimed. Tools Appl., vol. 69, no. 2, pp. 293–311, Mar. 2014, doi: 10.1007/s11042-012-0993-4.

[3] N. D. Doulamis, A. S. Voulodimos, D. I. Kosmopoulos, and T. A. Varvarigou, “Enhanced Human Behavior Recognition Using HMM and Evaluative Rectification,” in Proceedings of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams, New York, NY, USA, 2010, pp. 39–44, doi: 10.1145/1877868.1877880.

[4] K. Makantasis, E. Protopapadakis, A. Doulamis, N. Doulamis, and N. Matsatsinis, “3D measures exploitation for a monocular semi-supervised fall detection system,” Multimed. Tools Appl., vol. 75, no. 22, pp. 15017–15049, Nov. 2016, doi: 10.1007/s11042-015-2513-9.

[5] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau, “Robust video surveillance for fall detection based on human shape deformation,” IEEE Trans. Circuits Syst. Video Technol., vol. 21, no. 5, pp. 611–622, 2011.

[6] M. Yu, A. Rhuma, S. M. Naqvi, L. Wang, and J. Chambers, “A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment,” IEEE Trans. Inf. Technol. Biomed., vol. 16, no. 6, pp. 1274–1286, 2012.

[7] S. Lee, H. G. Kim and Y. M. Ro, "BMAN: Bidirectional Multi-Scale Aggregation Networks for Abnormal Event Detection," IEEE Trans. on Image Proc., vol. 29, pp. 2395-2408, 2020.

[8] A.S. Voulodimos, N.D. Doulamis, D.I. Kosmopoulos, and T.A. Varvarigou, “Improving multi-camera activity recognition by employing neural network based readjustment,” Applied Artificial Intelligence, 26(1-2), 97-118, 2012.

[9] N. Bakalos, et al. "Protecting water infrastructure from cyber and physical threats: Using multimodal data fusion and adaptive deep learning to monitor critical systems." IEEE Signal Processing Magazine, 36.2, pp. 36-48, 2019.

[10] S. Wan, X. Xu, T. Wang and Z. Gu, "An Intelligent Video Analysis Method for Abnormal Event Detection in Intelligent Transportation Systems," IEEE Trans. on Intell. Transportation Systems, (to be published)

[11] R. Leyva, V. Sanchez and C. Li, "Fast Detection of Abnormal Events in Videos with Binary Features," IEEE ICASSP, Calgary, AB, pp. 1318-1322, 2018.

[12] K.-W. Cheng, Y.-T. Chen, and W.-H. Fang, “Video anomaly detection and localization using hierarchical feature repre- sentation and Gaussian process regression,” IEEE CVPR, pp. 2909–2917, 2015.

[13] C. Lu, J. Shi, and J. Jia, “Abnormal Event Detection at 150 FPS in MATLAB,” IEEE ICCV, pages 2720–2727, 2013.

[14] H. Ren, W. Liu, S. I. Olsen, S. Escalera, and T. B. Moes- lund, “Unsupervised Behavior-Specific Dictionary Learning for Abnormal Event Detection,” Proc. of BMVC, pp. 28.1–28.13, 2015.

[15] M. Hasan, J. Choi, J. Neumann, A. K. Roy-Chowdhury, and L. S. Davis, “Learning temporal regularity in video sequences,” IEEE CVPR, pages 733–742, 2016.

[16] Xu Dan, Ricci Elisa, Yan Yan, Song Jingkuan and Sebe Nicu, "Learning Deep Representations of Appearance and Motion for Anomalous Event Detection", BMVC, 2015.

[17] L. Wang, F. Zhou, Z. Li, W. Zuo and H. Tan, "Abnormal Event Detection in Videos Using Hybrid Spatio-Temporal Autoencoder," 25th IEEE International Conference on Image Processing (ICIP), Athens, 2018, pp. 2276-2280, 2018.

[18] K. Makantasis, A. Nikitakis, A. Doulamis, N. Doulamis, and Y. Papaefstathiou, “Data-driven background subtraction algorithm for in-camera acceleration in thermal imagery,” IEEE Trans. Circuits Syst. Video Technol., vol. 28, no. 9, pp. 2090–2104, Sept. 2018.

[19] S. Herrero and J. Bescs, “Background subtraction techniques: Systematic evaluation and comparative analysis,” in Advanced Concepts for Intelligent Vision Systems (Lecture Notes in Computer Science, vol. 5807), J. Blanc-Talon, W. Philips, D. Popescu, and P. Scheunders, Eds. Berlin: Springer-Verlag, 2009.

[20] D. S. Yeo, “Superpixel-based tracking-by-segmentation using Markov chains,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2017, pp: 511–520.

[21] D. Kosmopoulos, A. Voulodimos, and A. Doulamis, “A system for multicamera task recognition and summarization for structured environments,” IEEE Trans. Ind. Informat., vol. 9, no. 1, pp. 161–171, 2013.

[22] H. M. Mousavi, “Analyzing tracklets for the detection of abnormal crowd behavior,” in Proc. IEEE Winter Conf. Applications Computer Vision, 2015, pp. 148–155.

[23] S. A. Ahmed, D. P. Dogra, S. Kar, and P. P. Roy, “Trajectory-based surveillance analysis: A survey,” IEEE Trans. Circuits Syst. Video Technol., 2018. doi: 10.1109/TCSVT.2018.2857489

[24] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, “Deep learning for computer vision: A brief review,” Computational Intell. Neurosci.,2018, article ID 7068349. [Online]. Available: https://doi.org/10.1155/2018/7068349

[25] A. Del Giorno, J. Bagnell, and M. Hebert, “A Discrimina- tive Framework for Anomaly Detection in Large Videos,” Proc. of ECCV, pp. 334–349, 2016.
[26] J. K. Dutta and B. Banerjee, “Online Detection of Abnormal Events Using Incremental Coding Length,” In Proceedings of AAAI, pages 3755–3761, 2015.
[27] R.T. Ionescu, S. Smeureanu, B. Alexe, and M. Popescu, “Un-masking the abnormal events in video,” IEEE ICCV, pp. 2895–2903, 2017.
[28] S. Smeureanu, R. T. Ionescu, M. Popescu, and B. Alexe, “Deep Appearance Features for Abnormal Behavior Detection in Video,” In Proceedings of ICIAP, Volume 10485, pages 779–789, 2017.
[29] Y. Liu, C.-L. Li, and B. Poczos, “Classifier Two-Sample Test for Video Anomaly Detections,” In Proceedings of BMVC, 2018.
[30] X. Mo, V. Monga, R. Bala, and Z. Fan, “Adaptive sparse representations for video anomaly detection,” IEEE Trans. Circuits Syst. Video Technol., vol. 24, no. 4, pp. 631–645, Apr. 2014.
[31] F. Jiang, Y. Wu, and A. K. Katsaggelos, “A dynamic hierarchical clustering method for trajectory-based unusual video event detection,” IEEE Trans. Image Process., vol. 18, no. 4, pp. 907–913, Apr. 2009.

Subject Index

accessibility 217
action selection 12
active learning 176
adaptive content 40
adaptive navigation 40
air pollution 207
air quality levels 207
Alzheimer's disease 63, 84
ambient intelligence 46
amusement 229
augmented reality 105
augmented reality accessibility 46
auto-encoder 84
biomimetic rig 21
buggy model 40
business process management 166
CAD 31, 127, 157
classification problem 147
clustering 137
computer vision 247
constraint-based tutoring systems 239
convolutional neural network 247
coronavirus 3
COVID-19 3, 31, 127, 157, 207, 247
crowdsourcing 176
cultural heritage 137
customer experience 191
customized training 75
data analysis 157
data analytics 166
data mining 157, 185
deep learning 247
deep learning computer 256
dementia 111
dengue 185
disabled 46, 217
disease surveillance 185
dynamic time warping 207
education 3
educational technology 105
EEG 63, 84, 229
ε-greedy 12
emotions 84, 229
emotiv 229
engineering education 31, 127, 157
environmental quality 207
error detection 239
exoskeleton systems 21
face detection 247
Facebook 3
fingerprint localization 147
forced ε-greedy 12
framework 95
fuzzy c-means algorithm 185
gait analysis 21
gamification 95
hierarchical classification 147
human detection 256
human-computer interaction 84
humour 229
hybrid model 127
hyperspectral data 137
indoor localization 147
intelligence 117
intelligent agent 63, 84
intelligent cognitive control system 63
intelligent modeling 117
intelligent system 117
knowledge 117
knowledge representation 117
learning disabilities 105
learning strategy 31
limbic system 63
logic 117
LSTM 229
machine learning 75, 127, 147, 191, 207
man overboard 256
mathematical structure 117
mobile application 75
mobile computing 176
MobileNetV2 247
modeling 117
Montessori method 111
multi-criteria decision making 223
nature based solutions 197

net promoter score 191
neural networks 207
neurofeedback 63
neurofeedback system 84
online learning 3, 31, 127, 157
ontology 117
ontology reasoning 239
Parkinson's disease 55
personalisation 95, 223
personalized augmented reality 75
personalized learning 40, 105
process analytics 166
process intelligence 166
program execution trace 239
public health 197
Q learning 12
reading comprehension 105
recommendation systems 223
reinforcement learning 84
semantic location identities 147
smart cities 46, 217
smart environments 46
smart technology 217
social networks 3
spatial ability 75
special education 105
speech recognition 84
stereotypes 223
students' satisfaction 127
subject area 117
sustainability 197
tailored feedback 40
technology 3
telemedicine 111
thermal image processing 256
time series clustering 185
transfer learning 176
unsupervised monitoring 137
virtual reality 55, 63, 84, 95, 111
virtual reality learning
environments 95
virtual reality rehabilitation 55
virus protection 247
visual analytics 166
visualization 166
well-being 197
zoo therapy 84

Author Index

Addawe, J.C. 185
Ai, Y. 84
Ang, A.J.Y. 111
Angelopoulos, G. 12
Anikin, A. 239
Azariadis, P. 21
Bakalos, N. 256
Bekas, G. 127, 157
Ben Abdessalem, H. 63, 84
Benlamine, M.S. 229
Bimpas, M. 197
Boque, J.C. 55
Boskovic, S. 197
Bousdekis, A. 166
Bozovic, R. 197
Caro, J.D.L. 55, 111, 185
Chatzigeorgiou, C. 176
Denisov, M. 239
Doulamis, A. 137, 197
Doulamis, N. 137, 191, 197, 207, 256
Feidakis, M. 176
Feria, R. 147
Figueroa, L.L. 147
Frasson, C. v, 63, 84, 229
Georgoulas, I. 191
Giannakas, F. 3, 40
Jamora, R.D. 55
Juayong, R.A.B. 55, 111, 185
Kabassi, K. v, 223
Kampouris, C. 21
Kanetaki, Z. 31, 127, 157
Kaouni, A. 166
Kapetanaki, A. 105
Karolou, E.E. 256
Kaselimi, M. 207
Kasnesis, P. 176
Katsamenis, I. 256
Kavouras, I. 207
Kontellis, E. 247
Kopsiaftis, G. 191
Krouska, A. 3, 40, 75, 95, 105, 247
Lalic, M. 197
Maksimovic, C. 197
Markoulidakis, I. 191
Marougkas, A. 95
Metafas, D. 12
Meytus, V.Yu. 117
Miaoulis, G. 166
Moulianitis, V. 21
Paderes, I.K.P. 147
Papakostas, C. 75
Patrikakis, C. 176
Poli, M. 46, 217
Principio Jr., R.D. 111
Protopapadakis, E. 137, 197, 207
Rallis, I. 137, 191
Sardis, E. 197, 207
Sfakianos, M. 191
Sgouropoulou, C. 3, 31, 40, 75, 95, 105, 127, 157, 247
Stergiou, C. 31, 127, 157
Sychev, O. 239
Theodoropoulou, G. 166
Toumanidis, L. 176
Toupin, G. 229
Troussas, C. 3, 31, 40, 75, 95, 105, 127, 157, 247
Tzanettis, K. 191
Tzortzis, I.N. 137
Voulodimos, A. v, 166
Voyiatzis, I. 3, 40
Zafeiropoulos, C. 137, 197

This page intentionally left blank