SASYR Symposium of Applied Science for Young Researchers

4th Symposium of Applied Science for Young Researchers

PROCEEDINGS 2024

July 3 , 2024

4th Symposium
of
Applied Science for Young Researchers

Proceedings

SASYR 2024

3 July 2024

Editors

Florbela P. Fernandes
Research Centre in Digitalization and Intelligent Robotics (CeDRI)
Instituto Politécnico de Bragança

Helena Torres
Applied Artificial Intelligence Laboratory (2Ai)
Instituto Politécnico do Cávado e do Ave

Pedro Pinto
Applied Digital Transformation Laboratory (ADiT-LAB)
Instituto Politécnico de Viana do Castelo

Instituto Politécnico de Bragança — 2024
Campus de Santa Apolónia
5300-253 Bragança – Portugal
ISBN: 978-972-745-341-2

Book cover: Natália Santos, Instituto Politécnico do Cávado e do Ave

Welcome

This document presents the proceedings of the 4th Symposium of Applied Science for Young Researchers - SASYR 2024. This scientific event welcomed works by junior researchers on any research topic covered by the following three research centers: ADiT-lab (from IPVC, Instituto Politécnico de Viana do Castelo), 2Ai (from IPCA, Instituto Politécnico do Cávado e do Ave) and CeDRI (from IPB, Instituto Politécnico de Bragança).

The main objective of SASYR is to provide a friendly and relaxed environment for young researchers to present their work, discuss recent results, and develop new ideas. In this way, this event offered an opportunity for the ADiT-lab, 2Ai, and CeDRI research communities to gather synergies and promote collaborations, thus improving the quality of their research. The SASYR 2024 took place at Instituto Politécnico de Bragança, Bragança, Portugal, on the 3rd of July, 2024.

The SASYR 2024 Organizing Committee,
Florbela P. Fernandes
Helena Torres
Pedro Pinto

Committees

Organizing Committee

Florbela P. Fernandes, CeDRI, Instituto Politécnico de Bragança
Helena Torres, 2Ai, Instituto Politécnico do Cávado e do Ave
Pedro Pinto, ADiT-lab, Instituto Politécnico de Viana do Castelo

Advisory Committee

Ana Pereira, CeDRI, Instituto Politécnico de Bragança
José Lima, CeDRI, Instituto Politécnico de Bragança
Paulo Leitão, CeDRI, Instituto Politécnico de Bragança
João L. Vilaça, 2Ai, Instituto Politécnico do Cávado e do Ave
António Miguel Cruz, ADiT-lab, Instituto Politécnico de Viana do Castelo
Jorge Garcia, ADiT-lab, Instituto Politécnico de Viana do Castelo

Scientific Committee

Ana Pereira, Instituto Politécnico de Bragança
André Mendes, Instituto Politécnico de Bragança
Andreia Teixeira, Institituto Politécnico de Viana do Castelo
Ângela Ferreira, Instituto Politécnico de Bragança
Ângela Silva, Institituto Politécnico de Viana do Castelo
António Miguel Rosado da Cruz, Institituto Politécnico de Viana do Castelo
Carla A. S. Geraldes, Instituto Politécnico de Bragança
Carlos Balsa, Instituto Politécnico de Bragança
Carlos R. Cunha, Instituto Politécnico de Bragança
Cátia Alves, Instituto Politécnico do Cávado e do Ave
Clara Bento Vaz, Instituto Politécnico de Bragança
Daniel Miranda, Instituto Politécnico do Cávado e do Ave
Duarte Duque, Instituto Politécnico do Cávado e do Ave
Estela Vilhena, Instituto Politécnico do Cávado e do Ave
Estrela Ferreira Cruz, Institituto Politécnico de Viana do Castelo
Fernando Monteiro, Instituto Politécnico de Bragança
Fernando Veloso, Instituto Politécnico do Cávado e do Ave
Florbela P. Fernandes, Instituto Politécnico de Bragança
Helena Torres, Instituto Politécnico do Cávado e do Ave
João Borges, Instituto Politécnico do Cávado e do Ave
João Carlos Silva, Instituto Politécnico do Cávado e do Ave
Joao Coelho, Instituto Politécnico de Bragança
João Paulo Teixeira, Instituto Politécnico de Bragança
Joaquim Gonçalves, Instituto Politécnico do Cávado e do Ave

Jorge Ribeiro, Instititution Politécnico de Viana do Castelo
José Henrique Brito, Instituto Politécnico do Cávado e do Ave
José Lima, Instituto Politécnico de Bragança
José Rufino, Instituto Politécnico de Bragança
Luís Barreto, Instititution Politécnico de Viana do Castelo
Luís Teófilo, Instititution Politécnico de Viana do Castelo
Maria F Pacheco, Instituto Politécnico de Bragança
Maria João Varanda Pereira, Instituto Politécnico de Bragança
Maria Mourão, Instititution Politécnico de Viana do Castelo
Natália Maria De Bessa Rego, Instituto Politécnico do Cávado e do Ave
Nuno Lopes, Instituto Politécnico do Cávado e do Ave
Óscar R. Ribeiro, Instituto Politécnico do Cávado e do Ave
Patricia Leite, Instituto Politécnico do Cávado e do Ave
Paula Alexandra Rego, Instititution Politécnico de Viana do Castelo
Paulo Alves, Instituto Politécnico de Bragança
Paulo Caldas, Instititution Politécnico de Viana do Castelo
Paulo Costa, Instititution Politécnico de Viana do Castelo
Paulo Leitao, Instituto Politécnico de Bragança
Paulo Matos, Instituto Politécnico de Bragança
Pedro Coutinho, Instititution Politécnico de Viana do Castelo
Pedro Faria, Instititution Politécnico de Viana do Castelo
Pedro Filipe Oliveira, Instituto Politécnico de Bragança
Pedro Morais, Instituto Politécnico do Cávado e do Ave
Pedro Pinto, Instititution Politécnico de Viana do Castelo
Pedro Rodrigues, Instituto Politécnico de Bragança
Rogério Tavares, Instituto Politécnico de Bragança
Rui Pedro Lopes, Instituto Politécnico de Bragança
Sara Paiva, Instititution Politécnico de Viana do Castelo
Sérgio Lopes, Instititution Politécnico de Viana do Castelo
Teresa Abreu, Instituto Politécnico do Cávado e do Ave
Tiago Pedrosa, Instituto Politécnico de Bragança
Vítor Carvalho, Instituto Politécnico do Cávado e do Ave

Local Organizing Committee

Fátima Pacheco, CeDRI, Instituto Politécnico de Bragança
André Mendes, CeDRI, Instituto Politécnico de Bragança

Secretariat / Technical Support

Carla Fontes, Instituto Politécnico de Bragança
Beatriz Cabrera, Instituto Politécnico de Bragança
Rebeca Kalbermatter, Instituto Politécnico de Bragança
Fernanda Roman, Instituto Politécnico de Bragança

Gabriel Henrique Silva Cangussu, Instituto Politécnico de Bragança
Clarisse Pais, Instituto Politécnico de Bragança
Evandro Alves, Instituto Politécnico de Bragança
Natália Santos, Instituto Politécnico do Cávado e do Ave

Table of Contents

Optimization of RNA classification using the Resonant Recognition Model

Felipe Bueno de Souza[1,2], Matheus Henrique Pimenta-Zanon[1], Dora Henriques[3], M. Alice Pinto[3], Carlos Balsa[2], José Rufino[2], and Fabrício Martins Lopes[1]

[1] Universidade Tecnológica Federal do Paraná (UTFPR), Campus Cornélio Procópio, Brasil
felipebuenosouza@alunos.utfpr.edu.br, matheus.pimenta@outlook.com, fabricio@utfpr.edu.br

[2] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Laboratório para a Sustentabilidade e Tecnologia em Regiões de Montanha (SUSTEC), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
balsa@ipb.pt, rufino@ipb.pt

[3] Centro de Investigação de Montanha (CIMO), Laboratório para a Sustentabilidade e Tecnologia em Regiões de Montanha (SUSTEC), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
dorasmh@ipb.pt, apinto@ipb.pt

Abstract. The development of high throughput sequencing technologies, such as RNA-Seq, has enabled the generation of large volumes of biological data. Thus, it is necessary to develop computational methods to interpret this massive volume of data and contribute to knowledge discovery. RNA sequences are products of the transcription of genomic DNA sequences, and represent the gene expression process that organisms use to synthesize protein or RNA molecules. These RNA sequences can be compared between organisms of the same or different species to demonstrate similar functional proteins. The RNA sequences have different biological functions, such as mRNAs, rRNAs, tRNAs and ncRNA, to name a few. The correct identification of each class of RNA sequences is important because of the huge volume of unlabeled data available. In this context, this study proposes an approach based on the Resonant Recognition Model for feature extraction and classification regarding the ncRNA and mRNA classes. To assess the proposed approach, it was adopted the dataset from the PLEK method. Despite the reduction of the input data size achieved using the RMM method, the results show high accuracy, indicating the potential of the proposed approach.

Keywords: RNA · RRM · RNAs classification · Feature Extraction · Bioinformatics · Pattern Recognition.

1 Introduction

One of the current scientific challenges is the interpretation and discovery of knowledge from the large volumes of data that are nowadays generated in the most diverse fields of science. Therefore, it is important to develop efficient mathematical and computational methods to deal with such challenge.

Regarding biological data, there has recently been a significant advance in the development of high throughput sequencing technologies "Next-Generation Sequencing", making it possible to popularize the sequencing of organisms [6]. Adequate Bioinformatics tools and algortihms are thus becoming essential to analyze the huge amounts of biological data generated by these technologies, which requires efficiency and scalability to extract useful information from it [4].

In particular, DNA sequences are transcribed to RNA sequences, providing different classes of RNAs (transcriptome), which are important because each class of RNA can perform different biological functions in the organisms, from regulation of cells and dosage compensation, their relationship with genetic diseases and autoimmune disorders [2, 5].

This study proposes a method to extract features from RNA sequences and classify two different classes of RNA: non-coding RNA (ncRNA) and messenger RNA (mRNA). The proposed approach starts by transforming DNA sequences into numerical series and analyzing their frequency spectra provided by Discrete Fourier Transform (DFT) as features, which are the input feature vector for classification. The proposed approach reduces the dimensionality of the inputted numerical series by applying the Resonant Recognition Model (RMM) [3], selecting only the common frequencies of each RNA class. This approach proved to be efficient, achieving the expected results with lower dimensional data input.

This paper is structured in four main sections, in addition to the Introduction: Section 2 covers the method applied in this study; Section 3 describes the dataset used, the experimental methodology applied and the results obtained; finally, Section 4 lays out the conclusions and delineates future work.

2 Resonant Recognition Model

The Resonant Recognition Model (RRM) method [3] is a digital signal processing method that uses numerical series that represent amino acids sequences, in order to extract the most discriminative information about their biological functionalities. These discrete sequences are transformed to the frequency domain by the Discrete Fourier Transformation (DFT), using the Fast Fourier Transformation (FFT) algorithm.

To get the numerical series from the DNA nucleotides string sequences, triads of nucleotides can be translated as amino acids, where each one has its EIIP value, as shown in Table 1. The values are the average energy states of the amino acid's valence electrons. In this way it is possible to convert strings sequences into numerical series, to be analyzed by digital signal processing methods. The data transformation follows the steps presented at Figure 1.

RMM extracts common frequencies between spectres through a cross-spectrum function, for two frequency spectres of different sequences. The cross-spectrum function can be described as the multiplication of the DFT coefficient X_n from a $x(m)$ series, by the conjugate complex Y_n^* of the DFT coefficients of another series $y(m)$, as shown in Equation 1:

$$S_n = X_n Y_n^* \quad n = 1, 2, 3, ..., N/2 \tag{1}$$

In this study, DNA sequences of different sizes were used. To define common frequencies among protein sequences, it is calculated the absolute value M_n for each coefficient of a multiple cross-spectral function, as defined by Equation 2:

$$|M_n| = |X1_n|.|X2_n|.|X3_n|...|XM_n| \quad n = 1, 2, 3, ..., N/2 \tag{2}$$

Table 1: Electron-Ion Interaction Potential (EIIP) Values for Amino Acids [3]

Name	Amino Acid	Letter	EIIP Value
Leucine	Leu	L	0.0000
Isoleucine	Ile	I	0.0000
Asparagine	Asn	N	0.0036
Glycine	Gly	G	0.0050
Valine	Val	V	0.0057
Glutamic Acid	Glu	E	0.0058
Proline	Pro	P	0.0198
Histidine	His	H	0.0242
Lysine	Lys	K	0.0371
Alanine	Ala	A	0.0373
Tyrosine	Tyr	Y	0.0516
Tryptophan	Trp	W	0.0548
Glutamine	Gln	Q	0.0761
Methionine	Met	M	0.0823
Serine	Ser	S	0.0829
Cysteine	Cys	C	0.0829
Thereonine	Thr	T	0.0941
Phenyalanine	Phe	F	0.0946
Arginine	Arg	R	0.0959
Aspartic Acid	Asp	D	0.1263

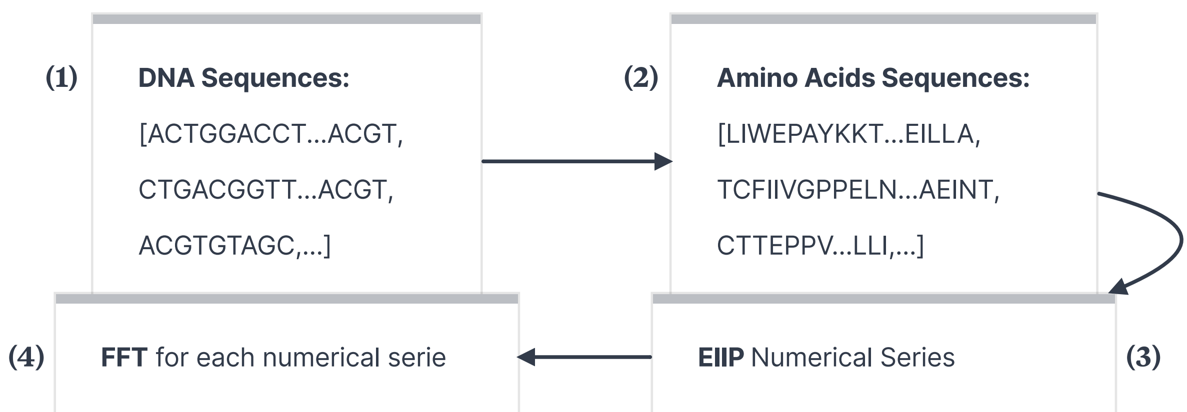


Fig. 1: From DNA sequences to Frequency Spectres: (1) Represent DNA sequences as array of strings. (2) Translate DNA strings into amino acids strings. (3) Transform each amino acid string in a numerical series using the EIIP value for each amino acid letter. (4) Use the FFT method to create a frequency spectre for each numerical series.

This multiple cross-spectrum function is called *Consensus Spectrum* for a large group of protein sequences with the same biological function [3].

3 Experiments

This sections is organized as follows: first, it is provided a characterization of the dataset used (Subsection 3.1); then, it is presented the methodology applied to classify the RNA classes (Subsection 3.2); finally, a comparison is made between the results obtained in this work and a previous related one (Subsection 3.3).

3.1 Dataset

To evaluate the RMM method, a dataset from the PLEK site (predictor of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme) [5] was used. This dataset includes different numbers of sequences of two RNA classes: ncRNA and mRNA. The dataset comprises data from 9 different organisms species (see Table 2).

Table 2: PLEK dataset [5]

Species	RNA Class	Number of Sequences	Sequences Mean Size	σ
Gorilla gorilla	mRNA	33025	2775.4	2080.6
(Western gorilla)	ncRNA	367	291.5	88.4
Macaca mulatta	mRNA	5709	2044.9	1388.9
(Rhesus macaque)	ncRNA	359	292.8	87.9
Bos taurus	mRNA	13190	2302.3	1507.4
(Cow)	ncRNA	182	296.8	116.9
Danio rerio	mRNA	14493	2088.8	1257.1
(Zebrafish)	ncRNA	419	593.1	471.6
Mus musculus	mRNA	35765	2659.8	2269.2
(House mouse)	ncRNA	8032	530.5	929.6
Pan troglodytes	mRNA	1906	1922.7	1204
(Chimpanzee)	ncRNA	1164	289.7	50.4
Pongo abelii	mRNA	3401	2836.9	1195.6
(Sumatran orangutan)	ncRNA	392	290.4	86.1
Sus scrofa	mRNA	3978	1823.7	1412.8
(Boar)	ncRNA	241	381.2	247.9
Xenopus tropicalis	mRNA	8874	2294.3	1350.1
(Clawed frog)	ncRNA	279	205.2	110.5

Because of the different amount of sequences, to create an unbiased prediction model, for each species, the number of sequences was limited to the minimum number of sequences between the two classes, slicing the sequence's array of the class with greater amount from index 0 to index of the minimum value, producing a balanced number of sequences in each class of RNA.

3.2 Methodology

The aim of this study was to develop a binary classification to classify sequences between ncRNAs and mRNAs, focusing on reducing the dimensionality of the input sequences.

For that, common frequencies from mRNA and ncRNA were extracted using miscellaneous sequences from the same class. This process was carried out individually for each one of the nine target species listed in Table 2.

To perform the spectre analysis, the frequencies were first distributed in a histogram of $N = 512$ bins, where frequencies range from 0.0 to 0.5. Note that 0.5 is the maximum frequency value of the spectre, because the mean distance of amino acids in a peptide chain is considered equidistant; Therefore, the distance between points in a numerical series is set arbitrarily with a value of $d = 1$, making the maximum frequency to be $F = d/2 = 0.5$ [3].

The distribution of values in a histogram ensures that sequences of different sizes are analyzed equivalently. First, two histograms were assembled, one for mRNA and one for ncRNA. These histograms have their coefficient values multiplied to form a spectrum of common frequencies by considering normalized values between 0 and 1, extracting which are the discriminatory frequencies for each class. The frequencies with a magnitude lower than 0.1 were filtered and considered noise. Figure 2 shows the process by which the histogram is built.

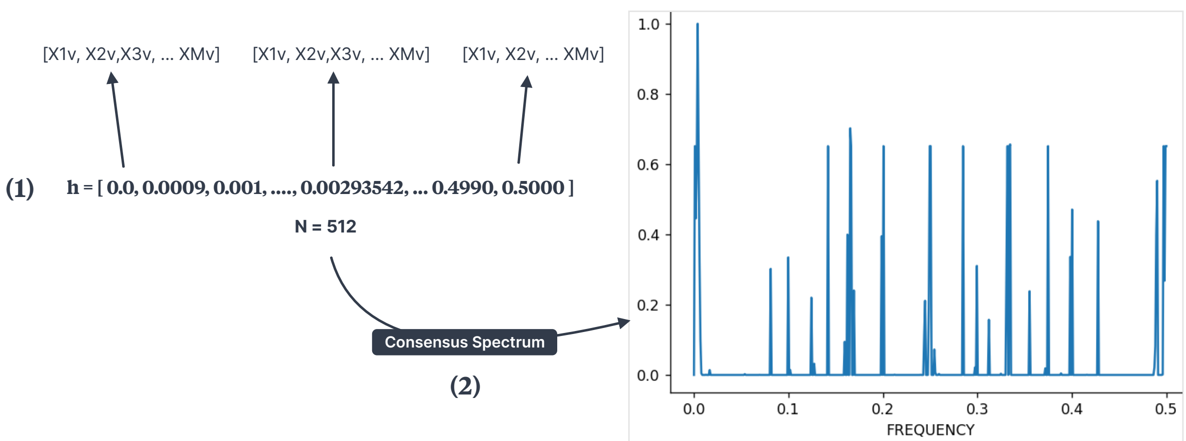


Fig. 2: Building the frequencies histogram based on the multiple cross-spectrum function *Consensus Spectrum*: (1) For each spectrum in the class, DFT sequences were iterated to assign each value to the respective frequency in the histogram; Because sequences have different dimensions, this process created a proportional histogram for the cross-spectral function. (2) With the histogram assembled, Equation 2 is applied, to create the spectrum signal with the peaks of common frequencies.

Then, to validate the selected frequencies as discriminatory parts, a classification was performed by considering histograms representing the DFTs as feature vectors, ensuring that sequences of different sizes were analyzed equally since larger sequences have more frequency points than smaller ones. Thus, different frequency points were allocated to the frequencies closest to the histogram, and a modulus of these values was performed when several values are assigned to a single frequency. This process can be visualized in Figure 3. The result is a collection of histograms that can be interpreted as discrete sequences of values.

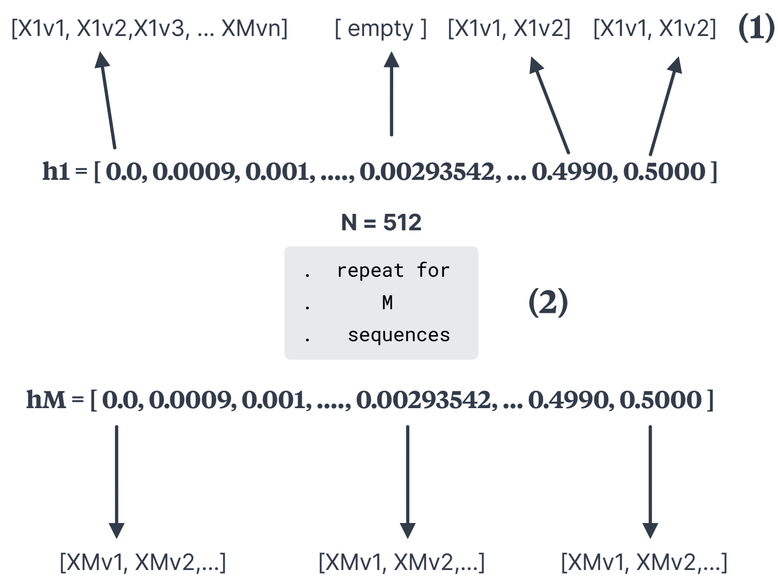


Fig. 3: Construction of the frequency histograms to analyze all the sequences spectra as a same dimension signal of $N = 512$ frequency intervals from 0.0 to 0.5: (1) The DFT sequence was iterated to assign each value to the respective frequency in the frequency histogram. (2) The first step was repeated for each one of the DFT sequences, creating M histograms that will be analyzed as a spectrum signal.

The discrete sequences were classified with their full size, and then, for each sequence, only the most discriminating frequency indices extracted through RRM in the first stage were selected, thus reducing the size of the input sequences and analyzing the performance of the two classifications.

The adopted classification algorithm was a Decision Tree, using the optimized version of the CART algorithm [1], with 10-fold cross-validation to assess the training performance of the classification model.

3.3 Results

The results of the classifications for each species can be seen in Table 3.

Table 3 shows similar results to those obtained with the PLEK tool [5], some better than others, but overall with similar accuracies. Comparing the results between sequences with size $N = 512$, and sequences with reduced sizes, it is possible to see that the peak values from common frequency can be interpreted as characteristics coefficients to be analyzed when the approach is to classify sequences. Therefore, by reducing the size of the input data, it is possible to achieve better performance and computational efficiency.

4 Conclusions

Biological sequences may be analysed in many different ways. Moreover, an increasingly amount of biological data is becoming available. Thus, computational solutions to increase the performance of the analysis of such are very important.

Table 3: Classification results for mRNA and ncRNA classes for each specie from PLEK dataset [5]

Species	RNA class	PLEK accuracy	N-512 DFT histogram accuracy	N-size sequences frequency peaks accuracy	N
Gorilla gorilla	mRNA	83.8%	91%	94%	41
	ncRNA	99.7%	93%	96%	
Macaca mulatta	mRNA	85%	92%	98%	38
	ncRNA	100%	93%	94%	
Bos taurus	mRNA	94.8%	98%	100%	38
	ncRNA	99.5%	98%	100%	
Danio rerio	mRNA	91.3%	78%	76%	3
	ncRNA	90.9%	82%	77%	
Mus musculus	mRNA	88.1%	79%	81%	2
	ncRNA	89.9%	80%	79%	
Pan troglodytes	mRNA	87.1%	94%	99%	42
	ncRNA	99.9%	97%	95%	
Pongo abelii	mRNA	98%	98%	98%	40
	ncRNA	100%	97%	96%	
Sus scrofa	mRNA	85.1%	88%	85%	8
	ncRNA	98.3%	88%	86%	
Xenopus tropicalis	mRNA	94.5%	96%	98%	90
	ncRNA	100%	97%	97%	

This study presents an approach based on RRM as a functional method for dimensionality reduction for analyzing the frequency spectra of DNA sequences. Here, using a simple decision tree as a classifying algorithm, high accuracies were achieved, even when reducing the input data dimensionality.

For future work, it will be considered the development of a method to map which parts of the original sequences generate the common frequency peaks that were used for dimensionality reduction in this work.

Acknowledgments:

This work was supported by national funds through the Fundação Araucária (Grant number 035/2019, 138/2021 and NAPI - Bioinformática), CNPq 440412/2022-6 and 408312/2023-8), FCT/MCTES (PIDDAC): CeDRI, UIDB/05757/2020 (DOI: 10.54499/UIDB/05757/2020) and UIDP/05757/2020 (DOI: 10.54499/UIDB/05757/2020); CIMO, UIDB/00690/2020 (DOI: 10.54499/UIDB/00690/2020) and UIDP/00690/2020 (DOI: 10.54499/UIDP/00690/2020); and SusTEC, LA/P/0007/2020 (DOI: 10.54499/ LA/P/0007/2020).

References

1. Breiman, L., Friedman, J., Stone, C., Olshen, R.: Classification and Regression Trees. Taylor & Francis (1984), `https://books.google.pt/books?id=JwQx-WOmSyQC`
2. Breve, M.M., Pimenta-Zanon, M.H., Lopes, F.M.: Basinetentropy: an alignment-free method for classification of biological sequences through complex networks and entropy maximization (2022)
3. Cosic, I.: Macromolecular bioactivity: Is it resonant interaction between macromolecules? - theory and applications. IEEE transactions on bio-medical engineering **41**, 1101–14 (01 1995). https://doi.org/10.1109/10.335859
4. Iqbal, N., Kumar, P.: From data science to bioscience: emerging era of bioinformatics applications, tools and challenges. Procedia Computer Science **218**, 1516–1528 (2023)

5. Li, A., Zhang, J., Zhou, Z.: Plek: a tool for predicting long non-coding rnas and messenger rnas based on an improved k-mer scheme. BMC Bioinformatics **15**(1), 311 (2014). https://doi.org/10.1186/1471-2105-15-311
6. Villaseñor-Altamirano, A.B., Balderas-Martínez, Y.I., Medina-Rivera, A.: Review of gene expression using microarray and rna-seq. In: Rigor and Reproducibility in Genetics and Genomics, pp. 159–187. Elsevier (2024)

A survey on the applicability of Artificial Intelligence, Virtual and Augmented Reality in Training in the Industrial Environment

Pedro Saldanha[1], Duarte Duque[1], and Jõao L. Vilaça[1]

Instituto Politécnico do Cávado e do Ave, IPCA, Portugal
psaldanha@ipca.pt, dduque@ipca.pt, jvilaca@ipca.pt

Abstract. Artificial Intelligence (AI) has been a powerful tool in education and training, offering personalized and adaptive solutions to improve learning performance. In industry, especially the footwear industry, its potential is evident. Training becomes different and dynamic, thus creating a way of learning. Augmented Reality (AR) and Virtual Reality (VR) can be used to complement immersive training and provide more realistic and interactive learning experiences. However, rapid technological evolution has created a growing gap between the level of technological advancement and the skills needed to operate it. This leads to an urgent need for highly qualified professionals, both in terms of advanced technical skills and interpersonal skills, to cope with the demands of new technologies. This phenomenon is exacerbated by the lack of adequate investment in training and education programs that keep pace with technological developments.With this in mind, this article aims to present a systematic literature review on the applications of AI in the training of professionals in the industry.

Keywords: Artificial Intelligence · Augmented reality · Virtual Reality · Industry · Training · Education Support .

1 Introduction

In recent years, the industry has witnessed an unprecedented convergence between Artificial Intelligence (AI) and Virtual Reality (VR) and Augmented Reality (AR) technologies [1]. This synergy has not only boosted the industrial sector but has also introduced new dimensions to training and skills development.

Traditional training, in certain respects, is still limited in its ability to provide trainees with knowledge and skills, due to lengthy processes and significant requirements in terms of resources and time [2], centered on direct learning on machines and heavy equipment. With the arrival of AI, this reality has begun to change significantly. There are now training systems that not only virtually replicate the working environment, but also adapt to the pace and learning style of users [3]. However, the implementation of these technologies faces substantial challenges, such as which advanced hardware and software technology is best, and the learning curve to effectively integrate these technologies into existing processes.

The use of VR and AR provides deep immersion through simulations that replicate environments, allowing workers to practice skills in risk-free scenarios. However, the widespread implementation of these technologies faces barriers, such as resistance to change on the part of workers [4] used to traditional methods and dependence on a robust technological infrastructure. Today, AI, VR and AR are redefining training standards in the manufacturing industry [2], increasing the safety and effectiveness of training programs and offering a platform for continuous skills development and adaptation to rapid technological change. The integration of these technologies promises to overcome traditional training challenges, offering safer and more efficient solutions that are essential for the advancement of the industry.

However, the widespread implementation of these technologies faces major obstacles. The high initial costs of hardware and software, as well as the development of specific or exclusive content [5]. In addition, there is a variation in the effectiveness of Virtual Reality-based training, which is highly dependent on the type of task and the quality of the software and hardware used [6]. Resistance to change on the part of workers is still present, however, head-mounted displays(HMDs) are an innovation, but the adherence rate is still low compared to adherence with mobile AR, which is significantly higher [6].

Despite these challenges, Artificial Intelligence (AI), Virtual Reality (VR) and Augmented Reality are creating new standards in the manufacturing industry, increasing safety and thus evolving the effectiveness of training programs. The integration of these technologies with personalized and efficient solutions is not only an industrial breakthrough. It also improves precision, reduces errors, and increases knowledge retention. Additionally, it allows for adaptable training and the possibility of rapid changes in the face of evolving technology [5].

In the face of these rapid technological changes, conducting a systematic review becomes crucial. This review aims to analyze existing studies to identify effective patterns, highlight critical points and uncover gaps in current knowledge, providing a solid foundation for future research and practical implementation. By deepening our understanding of these technologies and their application in industrial training, we hope to promote continuous improvements and facilitate more efficient adaptation to new technologies, thus boosting advances in the sector.

2 Related Work

This section discusses contributions to the fields of research associated with Virtual Reality (VR), Augmented Reality (AR) and Artificial Intelligence (AI). The existing literature reveals trends, challenges, and opportunities for improving the technologies applied in various sectors.

The study [7] presents the integration of AI and VR with support for future development, focusing on training simulation for energy systems, highlighting the ability of glasses to replicate complex working environments, in which they are essential for training advanced techniques. The research identifies some important gaps in the interaction between AI and VR. Firstly, there is a notable lack of adaptability in AI and VR systems to dynamically adjust to the different rhythms and learning styles of each worker. This adaptability is essential for creating personalized and effective learning environments. Just as current technologies struggle to create sufficiently complex training scenarios, they often fail to simulate real-world scenarios with the necessary accuracy, which is crucial for practical training applications in complex industrial environments. In addition, the depth of interaction between AI-generated content and VR environments requires significant improvement. Better interaction is needed to provide more immersive and engaging training experiences, so the use of haptic technologies [7] is essential for the user to "feel" or perceive what is happening in the virtual world.

In [3] the authors discussed the use of VR to train operators in the era of Industry 4.0, pointing to the critical need to adapt user interfaces to improve the learning experience and maximize operational efficiency. The research highlights the application of VR in processes such as design, assembly and prototyping, in which VR is used to train employees through 3D simulations, using Unity as the software to create the virtual simulations. Unity, as well as Unreal Engine, are game design development software; however, these applications allow detailed modeling and programming of 2D, 3D, VR and AR environments, making them the main software for the development of these applications.

The analysis of AR and VR applications carried out by [4] highlights the development of applications in industry for training, design and testing. This study emphasizes the importance of improving the use of haptic devices and artificial intelligence to improve the effectiveness of these simulations, providing more natural and immediate iterations between the physical and virtual environments. The need for more robust technological integration between VR and manufacturing systems exists, as does the need for wider acceptance of these technologies. The study highlights that although VR offers effective and immersive simulation, AR still faces significant challenges in integration and acceptance. Thus, the article proposes focusing on training methods that integrate AI more deeply with VR and AR, however, not only creating simulations that instruct, but also responding dynamically to inputs.

The study [1] details how to integrate AR with artificial intelligence, highlighting the combination of this technology as a promising vector for the future in industries. This study shows that AR and AI are integrated through the use of machine learning techniques, which are applied to improve the digital information on which they are superimposed on the real world, i.e., the study seeks to develop real-time guidance and training software using specific frameworks such as ARKit and ARCore, together with AI platforms such as TensorFlow, allowing the creation of an application that can not only display information in real time but also learn to adapt to the environment and user interactions. However, the study highlights that the need to overcome technical challenges and end-user acceptance is still a problem, and according to [1] the effectiveness of AR solutions is still limited by the maturity of the surrounding technologies and their interaction with existing production systems, especially on a large scale. The article therefore proposes improved methods of interaction with the user and the AR system, as well as a more adapted AI to optimize interface dynamics.

The literature analyzed provides a wealth of knowledge on the methods, implications and limitations of both VR and AI. By identifying and highlighting gaps, it is vital, at this early stage of the project, to draw on best practice in the use of AI, as well as in the application of VR/AR technologies in training and learning. Thus, this state of the art provides a more comprehensive view of current technologies and their applications in the context of industrial education, highlighting the technologies used, opportunities and challenges, as well as investigating emerging technologies and their impact on education and training with a focus on industry, especially the footwear industry.

In the article [8] a Vocational Immersive Storytelling (vIS) system is presented, which integrates storytelling with vocational training in Virtual Reality (VR). The vIS system focuses on teaching users how to operate a mechanical micrometer through an engaging narrative set in a virtual factory environment. According to the authors, storytelling can significantly improve long-term recall and user engagement with the system. The study involved a two-phase user experiment with 30 participants, comparing the vIS system's effectiveness against traditional 2D video and text-based training methods.

Participants using the vIS system showed higher retention rates and long-term recall of the training material, as evidenced by their ability to accurately use the micrometer a week after the training session. Additionally, the vIS system received a high system usability score (SUS) of 76, indicating greater user satisfaction compared to the 2D video method, which scored 72, and the text-based method, which scored 62. The immersive nature of VR, combined with the narrative-driven approach, provided a more engaging and effective training experience, demonstrating the potential of integrating storytelling into VR-based vocational training.

3 Methodology

The methodology for this review was structured to ensure an in-depth and systematic analysis of existing publications on Artificial Intelligence (AI) applications with AR/VR in education and training in the manufacturing industry. Initially, we defined the objectives of the review, clearly specifying the research questions.

Q1: How is Artificial Intelligence being used to support teaching, particularly in the context of task-based training within the industrial sector?

Q2: What are the applications of Artificial Intelligence in supporting teaching and training tasks within the footwear industry?

Q3: In what ways is Virtual Reality utilized in the teaching and training of tasks across the broader industry and specifically within the footwear industry?

To explore the application of emerging technologies in teaching and training within industrial contexts, specific keywords were established to guide the search in academic databases. The keywords used for the first question were Teaching or Training or education support or e-learning or Education and Industry or Manufacturing and "artificial intelligence" or AI. The second question focuses on the footwear industry, so the keywords used were artificial intelligence or AI and footwear industry or "shoes industry. The keywords used for the third question, which focuses on the use of Augmented Reality in teaching, were Augmented reality or AR and Teaching or Training or education support, or e-learning or Education and Industry or Manufacturing.

To ensure the accuracy of our analysis, we defined exclusion criteria applied specifically to each question investigated. These criteria were designed to ensure focus on the study's objective. The choice of the time frame (2015-2024) is justified by the rapid evolution of technology, ensuring that the selected studies attempt to reflect the most efficient practices and innovations. Only articles in English were included in order to maintain linguistic consistency. In addition, studies in the context of informal learning, school, medical or other areas outside the manufacturing industry were excluded because the focus of the study is on applications in industrial contexts and mainly in the footwear manufacturing industry. Finally, studies that do not address AI for human development are also excluded. These criteria can be found in Table 1. After applying these criteria, the studies listed in Table 3 were selected.

Table 1: Table of inclusion and exclusion criteria .

Research	Exclusion criteria	Inclusion Criteria
Use of Artificial Intelligence to support teaching, especially applied to the teaching/training of task in industry	-Studies outside this time (1/2015- 1/2024) -Language not English. -Studies carried out in informal learning contexts, school or medical contexts (e.g. K12 teaching). -Use of AI for non-human training	-Period (1/2015- 1/2024) -Language English -Studies carried out in apprenticeship contexts in the manufacturing industry -Using AI to help and support humans in training
Use of Artificial Intelligence to support teaching, especially applied to the teaching/training of tasks in the footwear industry	-Studies outside this time (1/2015- 1/2024) -Language not English. -Studies carried out in any area other than the footwear industry. -Use of AI for non-human training.	-Period (1/2015- 1/2024) -Language English -Studies carried out in apprenticeship contexts, especially in the shoe industry. -Using AI to help and support humans in training
Use of Virtual Reality and Augmented Reality in the teaching/training of tasks in industry and the footwear industry.	-Studies outside this time (1/2015- 1/2024) -Language not English. -Using mobile for training. -Studies carried out in school, medical or aviation learning contexts.	-Period (1/2015- 1/2024) -Language English -Studies carried out in apprenticeship contexts in the manufacturing industry and in the footwear industry. -Use of head-mounted displays for training.

Table 2: Table of results from the literature analyzed.

Research	Databases (# of registries)	Duplicated	Excluded	Assessed for Eligibility	Included
Use of Artificial Intelligence to support teaching, especially applied to the teaching/training of tasks in industry	Web of Science (N=2634)	N= 22	Language Not English. (N=66) Use of AI for non-human training. (N=572) Studies carried out in apprenticeship contexts. (N = 1946)	Studies carried out contexts in the manufacturing industry. (N=5) Don't have access. (N = 8)	Studies included in review. (N=15)
Use of Artificial Intelligence to support teaching, especially applied to the teaching/training of tasks in the footwear industry.	Web of Science (N=5) Science Direct (N=1)	N=0	Studies carried out in other areas. (N=6)	-	-
Use of Virtual Reality and Augmented Reality in the teaching/training of tasks in industry and the footwear industry	Web of Science (N=891)	N=11	Using mobile for training. (N=57) Studies carried out in school, medical or aviation learning contexts or any other non-manufacturing industry context. (N= 678) Don't use of head-mounted Displays. (N=28) Language Not English. (N=16)	Study focused only on construction. (N=67)	Studies included in review. (N=34)

4 Conclusion

Based on the analysis of the articles reviewed so far, we can conclude that there is a great need for continuous training in Industry 5.0. Emerging technologies such as VR, AR and AI offer immersive and safe learning environments, making it essential for training to evolve. In other words, although the monetary costs are high compared to traditional methods [9] , the results have been positive [10]. However, it is still too early to say that these technologies are completely flexible or moldable, thus showing the continued need for adaptation and development.

5 Acknowledgements

This work was financed by the financed by national funds, through FCT - Foundation for Science and Technology and FCT / MCTES under the project UIDB/05549/2020, UIDP/05549/2020 and LASI-LA/P/0104/2020. This project was also funded by the FAIST - Fábrica Ágilente Inteligente Sustentável e Tecnológica, co-funded from the "Mobilizing Agendas for Business Innovation" of the "Next Generation EU" program of Component 5 of the Recovery and Resilience Plan (RRP), concerning "Capitalization and Business Innovation", under the Regulation of the Incentive System "Agendas for Business Innovation".

References

1. Jeevan S Devagiri, Sidike Paheding, Quamar Niyaz, Xiaoli Yang, and Samantha Smith. Augmented reality and artificial intelligence in industry: Trends, tools, and future challenges. *Expert Systems with Applications*, 207:118002, 2022.
2. Feng-Kuang Chiang, Xiaojing Shang, and Lu Qiao. Augmented reality in vocational training: A systematic review of research and applications. *Computers in Human Behavior*, 129:107125, 2022.
3. Mohamad Fallaha, Orhan Korhan, and Qasim Zeeshan. Virtual reality: A possibility for training operator 4.0. In *Industrial Engineering in the Internet-of-Things World: Selected Papers from the Virtual Global Joint Conference on Industrial Engineering and Its Application Areas, GJCIE 2020, August 14–15, 2020*, pages 335–347. Springer, 2022.
4. Sylvain Fleury, Fabien Bernard, Raphaël Paquin, Philippe Blanchard, and Simon Richir. Augmented and virtual reality simulation in industry. *Computer*, 56(12):55–64, 2023.
5. Lea M Daling and Sabine J Schlittmeier. Effects of augmented reality-, virtual reality-, and mixed reality–based training on objective performance measures and subjective evaluations in manual assembly tasks: a scoping review. *Human factors*, 66(2):589–626, 2024.
6. Sanika Doolani, Callen Wessels, Varun Kanal, Christos Sevastopoulos, Ashish Jaiswal, Harish Nambiappan, and Fillia Makedon. A review of extended reality (xr) technologies for manufacturing training. *Technologies*, 8(4):77, 2020.
7. Miao Li, Lijuan Li, Risheng Jiao, and Hongguang Xiao. Virtrul reality and artificial intelligence support future training development. In *2017 Chinese Automation Congress (CAC)*, pages 416–419. IEEE, 2017.
8. Sanika Doolani, Luke Owens, Callen Wessels, and Fillia Makedon. vis: An immersive virtual storytelling system for vocational training. *Applied sciences*, 10(22):8143, 2020.
9. Juan José Ortega Gras, Maria Victoria Gomez Gomez, María Delgado, Josefina Garrido-Lova, and Gregorio Cañavate. Designing a technological pathway to empower vocational education and training in the circular wood and furniture sector through extended reality. *Electronics*, 12:2328, 05 2023.
10. Andreas Triantafyllidis, Rosa Carretón, Maria Loeck, Isidoros Michalellis, George Michalakis, Georgios Chantziaras, Sofia Segkouli, Dimitrios Giakoumis, Konstantinos Moustakas, Konstantinos Votis, and Dimitrios Tzovaras. *A User Evaluation Study of Augmented and Virtual Reality Tools for Training and Knowledge Transfer*, pages 291–304. 06 2022.

Intrinsic Wastes in Digital Era Industries and Their Influence on Technological Innovation

Arthur Henrique Gomes Rossi[1], Joseane Pontes[1], and Paulo Leitão[2]

[1] Federal Technological University of Parana (UTFPR), Ponta Grossa, 84016-210 Parana, Brazil
arthurrossi@alunos.utfpr.edu.br, joseane@utfpr.edu.br
[2] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal
pleitao@ipb.pt

Abstract. Waste is still generated even in the technological era and emerges as a consequence of inefficient use of digital technologies, its implementation and lack of culture to avoid waste aiming at making processes leaner and with lower investments. As such emerges the Lean philosophy as a solution for the new called digital waste, or digital Muda in the context of digital transformation. Dig- ital waste is present in some forms, from data usage and storage wastes to techno invasion, digital interference in the physical realm on human habits which hin- ders companies' technological innovation. Therefore, this study aims to explore the literature for digital waste and innovation to comprehend the literature topics and convergence. As such it was caried an exploratory review using Scopus data base to analyze 10 articles until 2022. Main results are the presentation for some of Digital waste terminologies and typology mapping found in portfolio and the discussion on how Lean can improve the innovation by avoiding digital waste.

Keywords: Digital Waste · Industry 4.0 · Technological Innovation.

1 Introduction

Technological innovation plays a pivotal role in the evolution of industrial processes, products, and services, bringing along tangible benefits such as operational efficiency, cost reduction, and quality enhancement. At the heart of this process, workers play a crucial role, being the primary end-users of implemented technologies. Their active involvement is essential for the success of innovation initiatives, providing valuable insights into process needs and challenges, as well as contributing to the identification and elimination of potential waste. The transition from Industry 4.0 to Industry 5.0 highlights a significant shift in the approach to automation and connectivity towards closer collaboration between humans and machines. This new paradigm, referred to as the Digital Era, seeks not only operational efficiency but also the improvement of the worker experience and the creation of more humanized work environments. However, the rapid evolution of digital technologies often outpaces companies' ability to assimilate them, making it challenging to understand how workers cope with these changes.

In this context, there arises a need to explore the intrinsic waste within the Digital Era, where the contexts of Industry 4.0 and 5.0 are contextualized, which can hinder technological innovation. The accumulation of digital waste can negatively impact companies' ability to add value and lead to loss of control over digitization processes.

Therefore, this study aims to present an analysis of intrinsic waste in Digital Era Industries, which may impair technological innovation, as evidenced by relevant literature. We will explore how adopting an exploratory literature approach. This approach is the first step of literature review involving a broad and initial search for information on a specific topic. It seeks to identify key issues, concepts, and research gaps in the area, without necessarily seeking in-depth analysis or a complete synthesis of existing studies. It is useful for gaining an overview of the field of study and for better defining the focus of more detailed research later on.

Thus, articles addressing the term "Digital Waste" were obtained, with 10 articles retrieved from the Scopus database and read for analysis. The Vosviewer software was used to conduct bibliometric analysis and word spectrum. This will enable addressing digital waste and its effective waste management, aiming to identify clusters and groupings of themes that may drive technological innovation and promote a more resilient, sustainable, and human-centered work environment.

2 Results & Discussion

In this section is presented the bibliometric and content analysis results along with the discussion. The most cited keyword is digital waste directly by 4 articles and Lean is cited in other 3 articles. Other highlights show emerging of Lean 4.0 and Industry 4.0 with 3 citations.

The clustering of keywords indicates four main clusters in the articles: one about Industry 4.0 and tools in yellow, the digital waste focused on the worker in blue, lean management for agile projects and manufacturing in green, and the red cluster about assembling performance indicators to avoid waste in the context of Industry 4.0. When analyzing over time with Fig 2, it becomes apparent that the search for understanding lost talents and a focus on the worker begins to stand out in the newer articles. Meanwhile, older clusters focus on comprehending the Lean model and theorizing it in the context of Industry 4.0 and its waste.

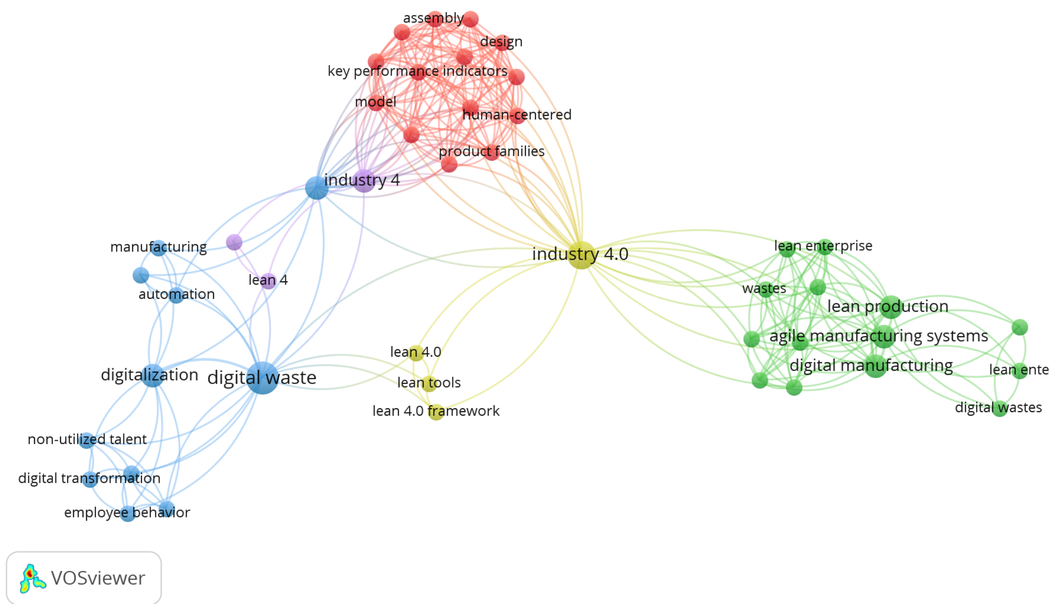


Fig. 1: Keywords clustering

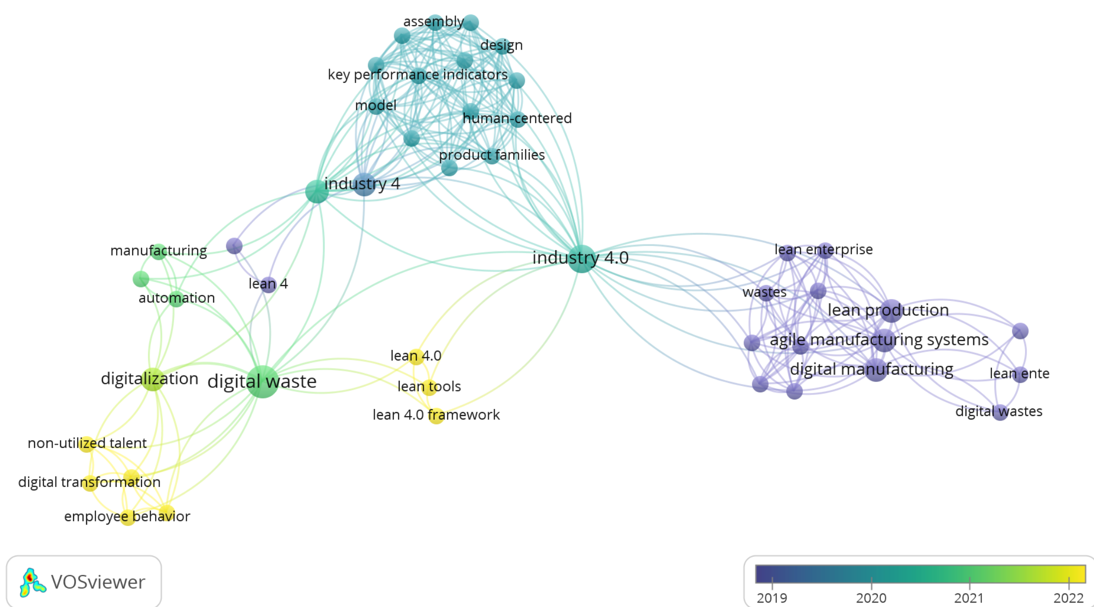


Fig. 2: Keywords in time distribution

Following the article analysis, Table 1 presents main concept and terminologies for digital waste and its implication in technological innovation.

Table 1: Digital Waste Terminologies.

Terminologies	Concept	Technological Innovation implications	References
Digital Waste	Waste generated by the use of digital technologies affecting both the physical and virtual worlds.	It hinders digitization processes, increases costs, alienates workers, and impacts the excess generation of data and processing.	[9], [10], [5], [6], [1], [3], [2], [7], [8], [4]
Reduction Mechanism	Mechanisms for reducing digital waste include technologies and practices that enhance visibility, accuracy, speed, feedback, employee engagement, and prevention of issues in digital processes.	By adopting these mechanisms, organizations can minimize waste associated with inefficiencies and errors in digital and operational systems.	[7], [8], [4]
Avoid Digital Waste	Addressing digital waste is essential to prevent disruptions in processes, improve efficiency, and avoid organizational problems.	Failure to address digital waste can lead to data instability, procedural insecurity, and incomplete decisions, ultimately resulting in organizational problems.	[9], [10], [1], [3], [2]
Digital MUDA	"MUDA," in the digital environment. It includes issues like data overload, inefficient data collection, misallocation of human resources, and underutilization of talents in Industry 4.0	Innovation Delay, Digital impacts the physical world, information instability	[1], [7], [8]

Portfolio analysis present inefficiencies in data acquisition, transmission, processing, unbalanced production workflows, and misalignment between digital technologies and organizational goals. Addressing these inefficiencies is crucial for optimizing operations in the digital era. Literature highlights the emergence of new forms of waste within Lean 4.0, such as data overload and inefficient human resource allocation, emphasizing the need for streamlined processes and improved training to mitigate digital waste and enhance operational efficiency. Main digital wastes detected in the portfolio was iden- tified. Waste of Data: encompasses issues related to inefficient data acquisition, trans- mission, and processing. Problems such as poor sensor quality, unnecessary data col- lection, and presenting data in complex formats contribute to this type of waste. Addi- tionally, inefficient processing and storage of data can lead to energy waste, hardware deterioration, and poor decision-making, further exacerbating digital waste. Waste of production: in digital contexts encompasses challenges like inefficient automation, un- balanced workflows, and digitized processes that don't add value. These inefficiencies lead to delays, errors, and increased costs. Lean management principles are vital for addressing this waste by optimizing processes, reducing unnecessary activities, and im- proving resource utilization. By embracing lean methodologies, organizations can streamline workflows, identify and eliminate inefficiencies, and enhance overall efficiency, thus mitigating digital waste and fostering a culture of continuous improvement. Waste of technology: The waste of technology refers to inefficiencies stemming from the misalignment between digital technologies and organizational goals, leading to suboptimal resource utilization and productivity hindrances. It encompasses issues like inadequate employee training, technology misuse, and allocating resources to tasks that could be automated. To address this, organizations must prioritize comprehensive train- ing programs and strategic planning to ensure effective utilization of digital tools in line with their objectives, thus optimizing productivity and minimizing waste.

3 Conclusion

The digitalization of a large amount of data has an overwhelming impact on operators, with augmented reality assistance, complex human-machine interfaces, confusing dashboards, and exaggerated safety systems, leading to alarm fatigue. This affects creativity, innovation, and can exacerbate turnover and alienation. Additionally, excessive training to deal with technologies overburdens operators, potentially creating a technostress environment via techno-invasion and techno-overload. Lean principles can play a crucial role in addressing this challenge by focusing on streamlining processes, eliminating the digitalization excess and preparing workers to deal with technologies. Additionally, advanced data analytics tools can sift through vast datasets to identify patterns and anomalies, enabling proactive decision-making and fostering innovation by uncovering new insights and opportunities. By embracing a combination of digital lean philosophy and Lean 4.0 tools updated with digital technologies, organizations can effectively tackle passive digital waste and optimize their digital operations for enhanced efficiency, productivity, and technological innovation. This is achieved by identifying and reducing inefficiencies at every stage of the digital management process, making Lean practices a critical success factor in driving production improvements and fostering technological innovation while avoiding the generation of data and human wastes caused by technologies. Future work can aim to analyze the relationship between digital waste and energy efficiency, as well as the reduction of energy consumption through the reduction of data waste. This works also suggest to new research to investigate the digital waste in the emerging context of Industry 5.0 and their relationship with human workers.

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

1. Alieva, J., Haartman, R.: Digital muda-the new form of waste by industry 4.0. OPERATIONS AND SUPPLY CHAIN MANAGEMENT **13**(3), 269–278 (2020)
2. Alieva, J., Powell, D.: The automation of lean practices: Digitalized or digitally wasted? In: Dolgui, A., Bernard, A., Lemoine, D., von Cieminski, G., Romero, D.(eds) Advances in Production

Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems. APMS 2021. IFIP Advances in Information and Communi-cation Technology, 633. Springer, Cham. (2021). https://doi.org/https://doi.org/10.1007/978-3-030-85910-7_69.
3. Alieva, J., Powell, D.: The significance of employee behaviours and soft management prac-tices to avoid digital waste during a digital transformation. International Journal of Lean Six Sigma **14**(1), 1–32 (2023). https://doi.org/I10.1108/IJLSS-07-2021-0127
4. Cifone, F., Hoberg, K., Holweg, M., Staudacher, A.: 'lean 4.0': How can digital technologies support lean practices? International Journal of Production Economics **241** (2021). https://doi.org/https://doi.org/10.1016/j.ijpe.2021.108258
5. Ejsmont, K., Gładysz, B.: Lean industry 4.0—wastes versus technology frame-work. Lec-ture Notes in Mechanical Engineering. In: Panuwatwanich, K., Ko, CH. (eds) The 10th In-ternational Conference on Engineering, Project, and Production Management. Springer, Singapore (2020). https://doi.org/https://doi.org/10.1007/978-981-15-1910-9_44.
6. Miqueo, A., Torralba, M., Yagüe-Fabra, J.A.: Lean manual assembly 4.0: A systematic review. Applied Sciences **10**(23), 8555 (2020). https://doi.org/http://dx.doi.org/10.1002/andp.19053221004
7. Romero, D., Gaiardelli, P., Powell, D., Wuest, T., Thürer, M.: Digital lean cyber-physical production systems: The emergence of digital lean manufacturing and the signif-icance of digital waste. In: Moon, I., Lee, G., Park, J., Kiritsis, D., von Cieminski, G. (eds) Advances in Production Management Systems. Production Management for Data-Driven, Intelligent, Collaborative, and Sustainable Manufacturing. APMS 2018. IFIP Advances in Information and Communication Technology, 535. Springer, Cham. (2018). https://doi.org/https://doi.org/10.1007/978-3-319-99704-9_2
8. Romero, D., Gaiardelli, P., Thürer, M., Powell, D., Wuest, T.: Cyber-physical waste identification and elimination strategies in the digital lean manufacturing world. In: Ameri, F., Stecke, K., von Cieminski, G., Kiritsis, D. (eds) Advances in Production Man-agement Systems. Production Management for the Factory of the Future. APMS 2019. IFIP Advances in Information and Communication Technology, 566. Springer, Cham. (2019). https://doi.org/https://doi.org/10.1007/978-3-030-30000-5_5.
9. Rossi, A., Marcondes, G., Pontes, J., Leitão, P., Treinta, F., De Resende, L., Mosconi, E., Yoshino, R.: Lean tools in the context of industry 4.0: Literature review, implementation and trends. Sustainability **14**, 12295 (2022). https://doi.org/https://doi.org/10.3390/su141912295
10. Rossi, A.H.G., Marcondes, G.B., Paulo Leitão, J.P.: Digital waste in the industry 4.0 context: Characterization, typology and solutions. PROCEEDINGS OF THE 32nd CONFERENCE ON MANAGEMENT OF TECHNOLOGY - IAMOT (2023). https://doi.org/(ISBN: 978-65-00-80827-8).

Smart Small Ruminant Device: An IoT-Based System for Real-Time Monitoring and Management of Sheep and Goat Mobility

Mateus Araujo[1], Paulo Leitão[1,3], Marina Castro[2,3], José Castro[2,3], and Miguel Bernuy[4]

[1] Research Center in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
a57010@alunos.ipb.pt, pleitao@ipb.pt
[2] Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
marina.castro@ipb.pt, mzecast@ipb.pt
[3] Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
[4] Universidade Tecnológica Federal do Paraná, Campus Cornélio Procópio, Brasil
angel@utfpr.edu.br

Abstract. The traditional agricultural practice of silvopastoralism, combining cultural heritage and ecological management, plays a vital role in food production and landscape management. This article explores the integration of precision livestock farming (PLF) into silvopastoral landscapes, using Internet of Things (IoT) technologies for monitoring and data transmission. Commercially available tracking devices lack interesting ergonomics, price and autonomy, and most companies do not focus on goats and sheep. A customized IoT device was developed and tested on goats, employing temperature/humidity sensors, relative position, geolocation and a long-range communication module. The system architecture enabled real-time data flow through the sensing, service and interface layers, resulting in data visualization. The results demonstrate the high reliability, low cost, interesting autonomy and appropriate ergonomics of the device, with the animals exhibiting normal behavior during the tests. The study highlights the effectiveness of IoT-enabled PLF for understanding grazing dynamics and animal behavior in silvopastoral systems, highlighting its potential for sustainable landscape management and the welfare of goats and sheep.

Keywords: IoT · Animal Tracking · Animal Monitoring · Precision Livestock Farming · Silvopasture.

1 Introduction

The practice of Silvopasture holds significant cultural and ecological importance in many societies, serving as a cornerstone of food production and landscape management [5, 13]. Beyond its agricultural contributions, Silvopasture fosters social cohesion and identity, embodying a nexus of human interaction with natural systems and providing ecosystem services [1].

The rise of the Internet of Things (IoT) has revolutionized our interaction with the environment, enabling real-time data collection and informed decision-making across diverse sectors. From smart homes to industrial operations and environmental monitoring, IoT has ushered in a new era of connectivity and innovation. Central to this advancement is the development of efficient IoT nodes capable of robust data collection and reliable communication [7].

In recent decades, the need to evaluate and compensate ecosystem services has underscored the importance of integrating technology and data acquisition systems within Silvopastoral landscapes. IoT and PLF methodologies have emerged as a key approach to accurately monitor and quantify grazing activities, essential for assessing and remunerating the ecological benefits of extensive grazing in forested and natural areas.

PLF encompasses a suite of technologies tailored to individual animal monitoring, particularly valuable in extensive grazing systems. By integrating PLF applications, herders can enhance pasture utilization, animal health monitoring, and overall livestock management. Data acquisition forms the foundation of PLF, with parameters ranging from temperature and weight to GPS-based position tracking and grazing behavior analysis [1].

Utilizing technologies like Global Navigation Satellite Systems (GNSS) and Geographic Information Systems (GIS), farmers can efficiently monitor grazing dynamics and pasture utilization. Combining GPS with accelerometers offers deeper insights into grazing behavior and rest patterns, enhancing small ruminant management and theft prevention. The integration of RFID, virtual fencing, UAVs (Unmanned Aerial Vehicles), and other emerging technologies further expands the toolkit for optimizing small ruminant husbandry within silvopastoral systems [12].

The application of IoT and PLF techniques enables precise recording of grazing patterns, contributing critical data for developing localized systems to reward ecosystem services, mitigate fire risks, recycle nutrients, and protect fragile mountain ecosystems [4]. Global Positioning Systems (GPS) have been pivotal in this regard, providing insights into animal behavior and resource utilization that benefit both livestock welfare and ecosystem management [1, 3, 11].

To respond to the need to track small ruminants, companies such as Domodis and Pastoral have developed commercial solutions. Domodis is suitable for farm animals, including horses, sheep, cows, goats and wild animals, although with concerns about the ergonomics of the device for goats and sheep. Pastoral provides a back-mounted tracking device that incorporates a solar panel, but this poses potential risks such as entanglement in vegetation for goats. Despite offering low-cost entry points, starting at 95 dollars, Pastoral's devices generally have a shorter battery life compared to Domodis, which starts at 397.00 euros per unit, and both also require monthly subscriptions for use [2, 9].

This work involves the development of a data acquisition system for goats and sheep based on IoT technologies and architecture. This system must guarantee twenty days of autonomy with acquisitions every five minutes, not exceed 200 euros, not cause any harm to the welfare of the animals, and have a communication range of at least two kilometers. In addition, the data collected is latitude, longitude, speed, temperature and humidity of the environment and the relative position of the animal's neck.

This article is organized into an introduction, which covers the context and importance of the application, materials and methods, which shows which technologies were used and how the system works in general, results, which show the performance of the entire system, followed by a discussion of the benefits of the system developed. Finally, the conclusion provides a summary of the work and a direction for future work.

2 Materials and Methods

This research aims to develop a data acquisition device for sheep and goats and its performance on a collar worn by the animal. The device is part of an IoT architecture and includes a series of sensors to acquire and transmit data.

2.1 Architecture

Although this work is focused on the device, in order for the whole system to work, an IoT architecture has been established that contains all the necessary layers. The layers are divided into an sensing layer, a network layer, a service layer and an interface layer. These layers ensure the flow of acquired data from its collection on the device to its presentation to the end user. Figure 1 gives an overview of the architecture adopted. Fig. 1 gives an overview of the adopted architecture.

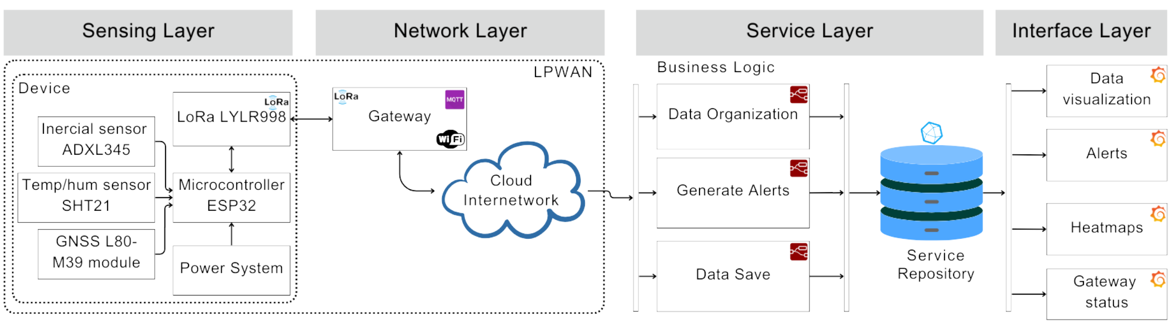


Fig. 1: IoT architecture overview. You can see the data sensing, network, service and user interface layers. Each block indicates a part of the system and some show the technology used.

2.2 Sensing Layer

The sensing layer is responsible for collecting and sending data. Belonging to the this layer, the device developed in this workis equipped with a high-precision GNSS module (L80-M39) that accurately records the animal's spatial movements, operating over a wide range of temperatures and voltages. In addition, temperature and humidity sensors (SHT21) have been incorporated for precise measurements of environmental conditions, as well as an inertia sensor (ADXL345) to detect changes in the movement of the animal's neck. The devices use Long Range (LoRa) communication (LYLR998 module) for long-distance data transmission and low power consumption.

The system's power solution combined a lithium polymer (LiPo) battery with a high-efficiency solar panel, ensuring continuous operation by harnessing solar energy to recharge the battery. All these functions are controlled and integrated by a robust microcontroller (ESP32), which offers various interfaces and peripherals for operation in harsh environments.

2.3 Network Layer

The network layer is responsible for receiving data from the device and forwarding it to the internet. This is done through a gateway that was connected to a local network and configured to forward received via LoRa data packets to a cloud server via the Message Queuing Telemetry Transport (MQTT) protocol, according to the architecture seen in 1. MQTT was chosen because it is extremely efficient in terms of bandwidth and resources, ideal for IoT devices with power and connectivity constraints. It uses a lightweight publish/subscribe model, enabling efficient asynchronous communication between devices and MQTT servers. In addition, MQTT supports distributed network architectures and is highly scalable, making for IoT systems.

LoRa was chosen because of its ability to transmit data over long distances with low power consumption, making it ideal for applications in extensive grazing areas where traditional connectivity may be limited. Its modulation technology allows it to reach greater distances without the need for complex network infrastructures, which makes it suitable for rural environments with varied topography, according to the case study [6].

2.4 Service Layer

The service layer is responsible for executing all the business logic and storing the data, for which purpose a Node-red application was developed to receive, filter, analyze, generate alerts and forward the data to the database. As Node-red is a platform for building task flows, it is easy to integrate and has little code, making it the ideal choice for this application.

The database used in this work is InfluxDB, an open source, highly available and scalable time series database designed to handle large volumes of time series data. It is especially suitable for storing, querying and visualizing data that changes over time, perfect for IoT applications.

2.5 Interface Layer

The interface layer is responsible for presenting all the data, averages, alerts and other relevant information to the end user. To do this, an application was developed in Grafana for data visualization due to its flexible dashboard creation tools and native support for integration with InfluxDB, which is the database chosen for the application. InfluxDB was selected for its efficiency in storing and querying time series data, offering high performance and scalability for handling large volumes of data. In addition, its ability to process queries in real time facilitates the analysis and visualization of the data collected.

With Grafana application, customized dashboards were developed to present historical and real-time data on grazing patterns, environmental conditions and device status.

3 Results

The tests carried out in this experiment provided various analyses of the system and its operation. Analyses ranged from connection reliability to consistent data presentation.

The tests were carried out on sunny pastoral days and the collars were fitted to random adult animals. Four tests were carried out on different days, i.e. the animals only wore the collars with the device while grazing, they were installed at the beginning of the grazing and removed at the end. Fig. 2 shows the collar fitted to the animal before it goes out to graze.

Fig. 2: Device fitted to the animal. In the image you can see the arrangement of the device on the animal, where the top supports all the modules and sensors and the bottom the feeding system. This distribution ensures that the bottom part counterbalances the device so that the top part is always on top of the animal.

3.1 Device Performance and Animal Behavior

During the four grazing tests, the device was programmed to remain idle for 10s. Each data acquisition process took an average of 20s (exiting idle mode, waiting for data from the GNSS and other modules and sending data). With these settings, the device averaged two data collections every minute. There were a few moments when the data was not collected correctly, due to unexpected module failures, but the device still showed a reliability of 98% during the tests. This is due to the reliability of the LoRa technology and the way the data rules are executed on the device.

In bench tests, the device had a consumption equivalent to 37 days of battery life, assuming a linear discharge with data being acquired and sent every five minutes. In addition, with these configurations it is possible to make approximately twelve thousand acquisitions of animal data. In the grazing tests, the solar panel showed a significant difference in the results. In one of the tests carried out, the device showed a gain in charge after grazing, demonstrating the contribution of the solar panel.

Animal behavior in relation to the collar was also analyzed. During all the tests, the animals that received the collar did not show any discomfort, as they grazed normally. The collar did not prevent them from running, eating, walking, climbing trees or any other activity. This guarantees that the shape and weight of the collar are efficient for its purpose.

3.2 Data Visualization

Application made in provides a rich presentation of all the data collected by the device and the status of the gateway. As the main focus of this work is on data relating to the animal's position, the panel that show the heat maps generated by application made in Grafana are more relevant. Fig. 3 shows two heat maps generated in two of the four tests carried out, on the left the heat map generated during a three-hour grazing session and on the right a seven-hour grazing session.

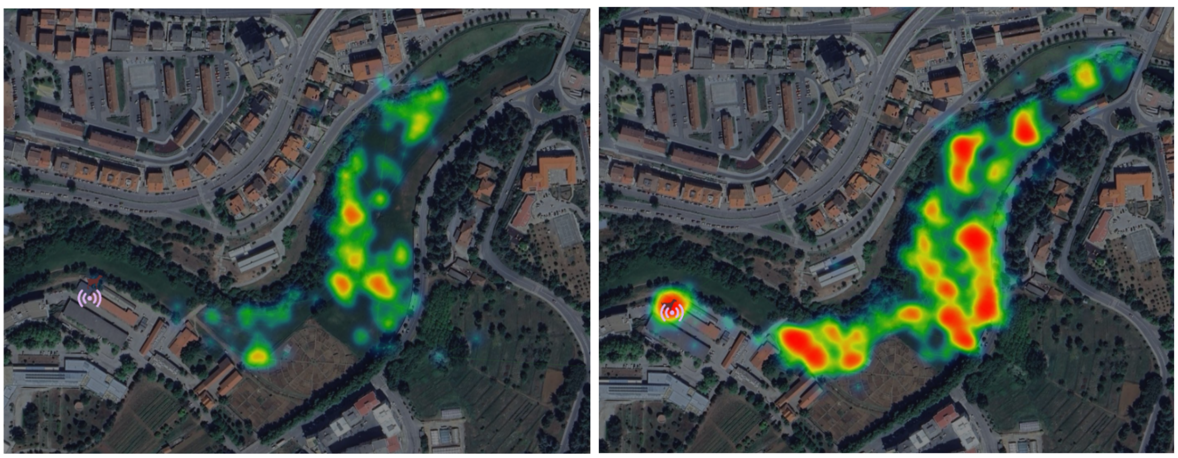


Fig. 3: Heat maps generated by Grafana. These heat maps were generated based on two days of grazing recorded by the system. In both cases, there is a distribution of green and red spots which indicate the areas preferred by the animals (the red ones indicate longer stays by the animals).

4 Discussion

The device developed is a breakthrough in the development of tracking devices for small ruminants. The sensors chosen allow for a series of data analyses and pattern detection that can be put to better use in future work. The price of this device, between 100-140 euros, despite being more expensive than the one developed by Pastoral, is cheaper than Domodis and other commercial devices [8].

Its format, where the power supply system is below the neck and the modules above, ensured that data acquisition and transmission were maximized, mainly by ensuring that the GNSS sensor and LoRa module remained facing upwards. In addition, the device's autonomy proved to be excellent, although further tests to assess the contribution of the solar panel need to be carried out.

The application developed in Grafana allows all the data to be presented in a rich and intuitive way. Heat maps are essential for analyzing animal patterns and their contribution to the silvopastoral system [10]. In addition, the alerts that this application generates quickly show possible risks or dangers that the animals may be facing.

5 Conclusion

The commercial devices available for tracking small ruminants have a number of problems in terms of ergonomics, high cost and autonomy. In this work, an intelligent device integrated with an IoT architecture was developed in order to solve the problems arising from commercial devices.

Due to the choice of three batteries and LoRa, the system had excellent autonomy and range, the shape of the device ensured non-invasive ergonomics for the animals and the choice of low-cost sensors ensured that the device did not exceed the value initially stipulated.

For future work, it is recommended to study the need to use temperature/humidity and relative neck position sensors, the removal of which would ensure a smaller device. In addition, the construction of a waterproof device will ensure that real-life tests can be carried out.

6 Acknowledgment

This work was supported by the Foundation for Science and Technology (FCT, Portugal) through national funds FCT/MCTES (PIDDAC) to CeDRI (UIDB/05757/2020 and UIDP/05757/2020) and SusTEC (LA/P/0007/2021).

References

1. Aquilani, C., Confessore, A., Bozzi, R., Sirtori, F., Pugliese, C.: Review: Precision Livestock Farming technologies in pasture-based livestock systems. Animal **16**(1), 100429 (2022)
2. Domodis: Collares GPS. https://www.domodis.com (2024)
3. Gordon, I.: Animal-based techniques for grazing ecology research. Small Ruminant Research **16**(3), 203–214 (1995)
4. Hamidi, D., Komainda, M., Tonn, B., Harbers, J., Grinnell, N.A., Isselstein, J.: The Effect of Grazing Intensity and Sward Heterogeneity on the Movement Behavior of Suckler Cows on Semi-natural Grassland. Frontiers in Veterinary Science **8** (2021)
5. Köhler-Rollefson, I.: Hoofprints on the Land. Chelsea Green Publishing UK, Somerset House, London, UK, 9st edn. (01 2023)
6. Lauridsen, M., Nguyen, H., Vejlgaard, B., Kovacs, I.Z., Mogensen, P., Sorensen, M.: Coverage comparison of GPRS, NB-IoT, LoRa, and SigFox in a 7800 km^2 area. In: 2017 IEEE 85th Vehicular Technology Conference (VTC Spring). pp. 1–5. IEEE (2017)
7. Madakam, S., Ramaswamy, R., Tripathi, S.: Internet of things (iot): A literature review. Journal of Computer and Communications **3**, 164–173 (2015)
8. Maroto-Molina, F., Navarro-García, J., Príncipe-Aguirre, K., Gómez-Maqueda, I., Guerrero-Ginel, J.E., Garrido-Varo, A., Pérez-Marín, D.C.: A low-cost iot-based system to monitor the location of a whole herd. Sensors **19**(10) (2019)
9. Pastoral: Healthier Livestock Reduced Carbon Emissions Profitable Outcomes. https://www.pastoral.ai (2024)

10. Torres-Manso, F., Marta-Costa, A.A., Castro, M., Tibério, L.: Silvopastoral Systems as a Tool for Territorial Sustainability and Biodiversity, pp. 317–333. Springer Singapore, Singapore (2017)
11. Trieu, L.L., Bailey, D.W., Cao, H., Son, T.C., Scobie, D.R., Trotter, M.G., Hume, D.E., Sutherland, B.L., Tobin, C.T.: Potential of Accelerometers and GPS Tracking to Remotely Detect Perennial Ryegrass Staggers in Sheep. Smart Agricultural Technology **2**, 100040 (2022)
12. Tzanidakis, C., Tzamaloukas, O., Simitzis, P., Panagakis, P.: Precision Livestock Farming Applications (PLF) for Grazing Animals. Agriculture **13**(2) (2023)
13. Zhao, J., Cao, J., Che, Z., Guo, Y., Ma, C., Zhang, Q.: Contribution of Sheep Grazing to Plant Diversity in Natural Grasslands. Diversity **14**(6) (2022)

Integration of AI and Embedded Systems for Enhanced Intrusion Detection System: A Case Study with NVIDIA Jetson Nano

Rui Fernandes[1], Nuno Lopes[1,2], and Joaquim Gonçalves[1,3]

[1] 2AI - School of Technology - IPCA, Barcelos, Portugal
rpfernandes@ipca.pt
nlopes@ipca.pt
jgoncalves@ipca.pt
[2] LASI – Associate Laboratory of Intelligent Systems, Portugal
[3] LIACC, Laboratório de Inteligência Artificial e Ciência de Computadores da Universidade do Porto, Portugal

Abstract. Intrusion Detection Systems are essential at all levels of companies, from startups to big companies, and they are still being optimized. Machine learning-based systems require resources - money and energy - and they can be reduced by using embedded systems. Edge computing and low-power devices are being introduced to run ML models due to their reduced cost and low energy consumption. This use contributes to increasing safety in networks without the need for a bigger investment.
This research leverages an ML-based IDS, using an embedded system (NVidia Jetson Nano) to behave as an IDS, proposing a system architecture to deploy an ML model into an embedded system and showing a guide on how to start from a dataset to implement the model integrated with a well-known software - CICFlowMeter.

Keywords: AI-based IDS · Cybersecurity · Intrusion Detection System · Machine Learning · Nvidia Jetson Nano.

1 Introduction

Cybersecurity is responsible for producing and processing a huge quantity of data, coming from firewalls, Intrusion Detection Systems (IDS), or Intrusion Prevention Systems (IPS) in the form of network traffic metadata and systems logs, to keep network systems operational and safe from Internet attacks. IDS have an important role in the network infrastructure because they represent a layer of security that can detect a potential attack going on through the analysis of network traffic, in real-time, and raise a warning about a specific threat.

The high cost and energy needed to run ML models raise concerns that create space for innovation and research around low-budget and low-energy consuming devices.

Embedded computing systems are fast proliferating in every aspect of human endeavour today, finding useful applications in areas such as wearable systems for health monitoring, wireless systems for military surveillance, networked systems as found in the Internet of Things (IoT), smart appliances for home automation, and antilock braking systems in automobiles, among others [1].

This research focuses on solving an actual problem - the classification of intrusion in networks - using embedded systems and the end device that will run those systems to behave as an IDS.

2 State of Art

Due to the subject of Cybersecurity being a sensitive area, we can assume that every IDS is outdated because there will always be a new exploit or attack. Thus, those systems must be adaptive and ready to learn more. This requirement is the reason why Machine Learning (ML) is mandatory. A Machine Learning-Based Intrusion Detection System, testing different ML algorithms in different datasets, [10] has been developed achieving great results in classifying network traffic. ML algorithms used to classify the attacks are in the areas of supervised and unsupervised learning as well as Reinforcement Learning (RL) [13].

For this research, it was decided to use Nvidia Jetson Nano to evaluate the models trained and if possible, to train and predict in real-time. This device has been used for some studies [2,3,5,11] but for the simple task of evaluating the time needed to predict. This research will use the existing research and leverage an IDS to work in real time.

3 Embedded AI-based IDS

The process of this research was split into three parts: the use of an embedded system - Nvidia Jetson Nano (subsection 3.1); the ML algorithms trained with IDS datasets (subsection 3.2); and a last part where we combine both the embedded system and the ML model trained to build a fully-working IDS in real-time for packet classification (subsection 3.3). The overall system architecture is displayed in Figure 1 where the three parts are together in the workflow.

As input the system receives the network packets in real-time or a file with packets pre-recorded and ML model as ".pkl" file. As output the program returns some identifiable features and the traffic category predicted. All this data is then loaded to a web server for monitoring.

3.1 Nvidia Jetson Nano

NVIDIA Jetson Nano Developer Kit is described by NVIDIA as a small, powerful computer designed to run different Neural Networks in parallel for different use cases related to computer vision, speech processing robotics, and others [14]. Although the Jetson Nano is seen as an end-device where a model is used to infer something and predict a certain output it can also be employed for parallel retraining while doing its normal job (predicting) [12].

3.2 AI-based IDS

The chosen IDS dataset for training the ML algorithm that is further used for deployment was the HIKARI-2021 [9]. Our previous studies [6–8] where we applied this dataset allows us to say that it is a reference and the most recent published IDS dataset. Regarding the data cleaning process in this dataset, four features were removed [8] and the label encoder from scikit-learn was used to encode the categorical features.

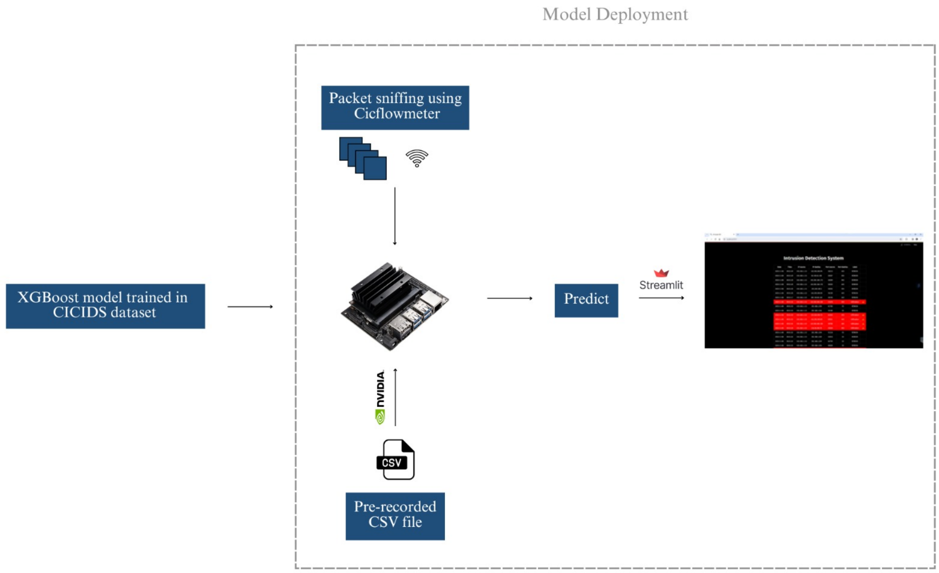


Fig. 1: System Architecture

The algorithm trained was the XGBoost, a well-known ML algorithm that has proven its value in this matter [4]. The parameters for training were the default and the method of splitting the dataset was the hold-out method with 70% for train and 30% for test.

Although for this problem different metrics can be used, in our case and due to the sensitivity of this problem the F1-score was the reference achieving a total of 76.3% after the training process which was done 10 times to prevent any problem that could interfere negatively with the training.

3.3 Integration and Development

To integrate the pre-trained model with the embedded system (NVidia Jetson Nano), the CICFlowMeter - a network traffic flow generator and analyzer - was used. This tool was the one that was used to build the CICIDS dataset. Developed in 2017 and open source published in GitHub [15], it was built in Java Code and the most recent developments are a Python version developed in 2021.

With the CICFlowMeter tool, a new integration step was carried out to make a call to the IDS system in order to utilize the data recorded by CICFlowMeter and employ it in our trained model mentioned above. To build a complete ecosystem, a web interface in Python harnessing the streamlit library was made, where it is possible to see the flow of data as well as the traffic category after the packet is predicted.

Figure 2 shows the webpage where a table is displayed with the information relative to the source IP, destiny IP, source Port, destiny Port and the label which corresponds to benign traffic or other type of attack. In case of an attack, the line correspondent of the row will be filled in red in the background and a new column in the right will have a warning sign.

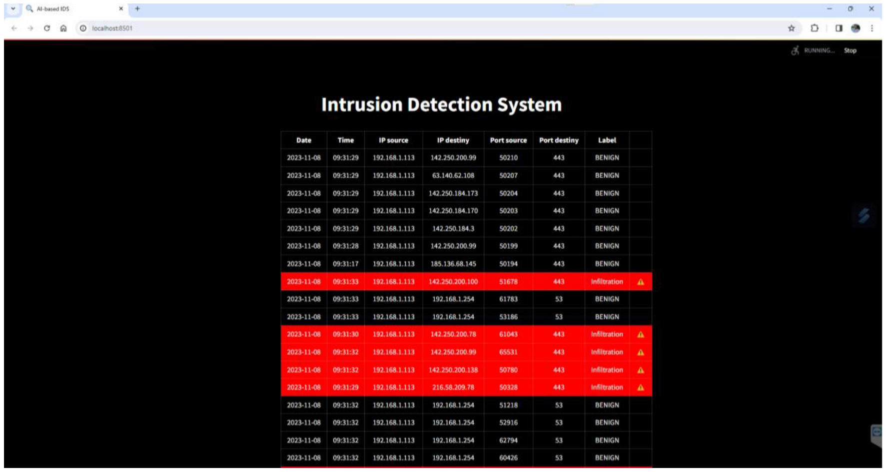


Fig. 2: IDS Web interface

4 Conclusions and Future Work

This research represents a step further in the development of AI-based IDS as well as in the deployment of AI-based systems in embedded systems.

The results from this research and the discussion of the built solution propose the deployment of an ML model into an embedded system and demonstrate a guide on how to start from the training step to the deployment, to build a ready-to-use system that behaves as an IDS.

Further research must be done to evaluate that system by comparing it with others used in the industry and analyzing the time of prediction, also future research is required to evaluate if the system can train and predict at the same time to improve itself over time.

References

1. Ajani, T.S., Imoize, A.L., Atayero, A.A.: An overview of machine learning within embedded and mobile devices–optimizations and applications. Sensors **21**(13) (2021). https://doi.org/10.3390/s21134412, `https://www.mdpi.com/1424-8220/21/13/4412`
2. Alladi, T., Kohli, V., Chamola, V., Yu, F.: A deep learning based misbehavior classification scheme for intrusion detection in cooperative intelligent transportation systems. Digital Communications and Networks (07 2022). https://doi.org/10.1016/j.dcan.2022.06.018
3. Bierbrauer, D.A., De Lucia, M.J., Reddy, K., Maxwell, P., Bastian, N.D.: Transfer learning for raw network traffic detection. Expert Systems with Applications **211**, 118641 (2023). https://doi.org/https://doi.org/10.1016/j.eswa.2022.118641, `https://www.sciencedirect.com/science/article/pii/S0957417422016840`
4. Chen, Z., Jiang, F., Cheng, Y., Gu, X., Liu, W., Peng, J.: Xgboost classifier for ddos attack detection and analysis in sdn-based cloud. In: 2018 IEEE International Conference on Big Data and Smart Computing (BigComp). pp. 251–256 (01 2018). https://doi.org/10.1109/BigComp.2018.00044
5. Chougule, A., Kohli, V., Chamola, V., Yu, F.R.: Multibranch reconstruction error (mbre) intrusion detection architecture for intelligent edge-based policing in vehicular ad-hoc networks. IEEE Transactions on Intelligent Transportation Systems **24**(11), 13068–13077 (2023). https://doi.org/10.1109/TITS.2022.3201548
6. Fernandes, R., Lopes, N.: Network intrusion detection packet classification with the hikari-2021 dataset: a study on ml algorithms. In: 2022 10th International Symposium on Digital Forensics and Security (ISDFS). pp. 1–5 (2022). https://doi.org/10.1109/ISDFS55398.2022.9800807

7. Fernandes, R., Lopes, N.: A simple distributed approach for running machine learning based simulations in intrusion detection systems. In: Pereira, T., Impagliazzo, J., Santos, H., Chen, J. (eds.) Internet of Everything. pp. 66–75. Springer Nature Switzerland, Cham (2024)
8. Fernandes, R., Silva, J., Ribeiro, O., Portela, I., Lopes, N.: The impact of identifiable features in ml classification algorithms with the hikari-2021 dataset. In: 2023 11th International Symposium on Digital Forensics and Security (ISDFS). pp. 1–5 (2023). https://doi.org/10.1109/ISDFS58141.2023.10131864
9. Ferriyan, A., Thamrin, A.H., Takeda, K., Murai, J.: Generating network intrusion detection dataset based on real and encrypted synthetic attack traffic. Applied Sciences **11**(17) (2021). https://doi.org/10.3390/app11177868
10. Halimaa A., A., Sundarakantham, K.: Machine learning based intrusion detection system. In: 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI). pp. 916–920 (2019). https://doi.org/10.1109/ICOEI.2019.8862784
11. Hong, W.C., Huang, D.R., Chen, C.L., Lee, J.S.: Towards accurate and efficient classification of power system contingencies and cyber-attacks using recurrent neural networks. IEEE Access **PP**, 1–1 (07 2020). https://doi.org/10.1109/ACCESS.2020.3007609
12. Mirzai, A., Coban, A.Z., Almgren, M., Aoudi, W., Bertilsson, T.: Scheduling to the rescue; improving ml-based intrusion detection for iot. In: Proceedings of the 16th European Workshop on System Security. p. 44–50. EUROSEC '23, Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3578357.3589460, `https://doi.org/10.1145/3578357.3589460`
13. Nguyen, T.T., Reddi, V.J.: Deep reinforcement learning for cyber security. IEEE Transactions on Neural Networks and Learning Systems **34**(8), 3779–3795 (2023). https://doi.org/10.1109/TNNLS.2021.3121870
14. NVIDIA Corporation: Jetson nano developer kit. `https://developer.nvidia.com/embedded/jetson-nano-developer-kit` (2023), accessed: 2024-04-19
15. University of New Brunswick, Canadian Institute for Cybersecurity: Research applications. `https://www.unb.ca/cic/research/applications.html` (2023), accessed: 2024-04-19

On the Path to Develop a Recovery Cost-Utility Function for Ransomware Attacks

Luís Pinto[1], Pedro Pinto[1,2], and António Pinto[2,3]

[1] Instituto Politécnico de Viana do Castelo, IPVC, Portugal
lfilipepinto@ipvc.pt, pedropinto@estg.ipvc.pt
[2] INESC TEC, Porto, Portugal
[3] ESTG, Instituto Politécnico do Porto, 4610-156 Felgueiras, Portugal
apinto@inesctec.pt

Abstract. Ransomware stands out for its disruptive impact on organizations regarding cybersecurity threats. These attacks not only encrypt data, demanding ransoms for access, but also lead to significant operational, financial, and reputational damages. The complexity and frequency of ransomware incidents require robust risk management and recovery planning strategies.
Although there is worldwide recognition of the risks ransomware creates, organizations have difficulty in estimating and managing the costs associated with recovery. The variability in recovery times across different IT assets further complicates this issue. Current strategies are not enough to quantify the costs, leading to potential problems in crucial cybersecurity evaluation.
This paper introduces the development of an assessment that will be the base for the creation of a utility function for estimating recovery costs from ransomware attacks. Focusing on the varied recovery times across different Information & Technologies (IT) assets within organizations, this study establishes a methodological framework to quantify the factors that influence these times. Through an assessment involving nineteen questions, the research captures qualitative and quantitative data regarding the presence and efficacy of Disaster Recovery (DR) policies, redundancy systems, and the specific recovery times for diverse IT components. The implications of this research can offer insights into cybersecurity policy and the strategic planning of resources, contributing to cybersecurity risk management.

Keywords: Ransomware · Function · Costs · Recovery · Calculate.

1 Introduction

When considering cybersecurity threats, ransomware attacks have emerged as an enormous challenge that confronts organizations across the globe [2]. Defined as the unauthorized encryption of data and the demand for payment in exchange for the decryption keys [6], these attacks usually have relevant operational, financial, or reputational impact [1]. Beyond the immediate impact to an entity blocked from accessing its information, the broader impact must include the recovery effort, the financial loss due to downtime, and potential legal liabilities [5]. The complexity and frequency of ransomware incidents have underscored the urgent need for sophisticated risk management strategies and recovery planning.

This paper addresses a gap in cybersecurity risk management by focusing on the estimation of recovery times from ransomware attacks. The primary problem being tackled is the variability and lack of specificity in existing methodologies for quantifying the time and resources required for effective recovery across different IT assets within organizations. The study aims to introduce an assessment that will allow for the creation of a utility function that will improve the accuracy of cost and time assessments related to ransomware recovery.

The proposed utility function will allow companies to understand the impacts associated with ransomware attacks. By quantifying the time and resources required for recovery, this tool will aim to be a close representation of reality.

This work is an initial step towards effective cybersecurity risk management and decision-making actions regarding ransomware attacks. With our approach, by using data that is directly related to the time needed to recover from these cyberattacks, the intention is to have a utility function that will be capable of helping organizations understand the importance of preemptive measures to minimize the time needed for recovery.

Through an evaluation of ransomware recovery times, the paper proposes a framework for enhancing organizational preparation and mitigating the impacts of cyberattacks. By assessing the average duration of each factor and evaluating their relative contributions to the ransomware recovery process, the authors aim to develop a formula to calculate the total recovery time for a ransomware attack.

The paper is organized as follows: Section 1 provides an overview of the ransomware threats and highlights the evolving nature of these cyberattacks. Section 2 reviews the existing literature on ransomware impacts and recovery cost estimation, identifying gaps the current research aims to fill. Section 3 details the assessment's methodology, including the questions that will allow for data collection. Finally, Section 4 discusses the implications of the research for cybersecurity policy and practice, and makes some conclusions about the objectives of the assessment.

2 Related Work

The proliferation of ransomware attacks and the subsequent need for effective recovery strategies have allowed for some literature dedicated to the economic quantification of these incidents to arise. Often, the recovery cost exceeds the initial ransom demand [7] and some efforts in this domain aim to create a utility function that assists in forecasting the financial aftermath of ransomware attacks on organizations. This section surveys the milestones in the development of such models. The development of a recovery cost-utility function for ransomware attacks is an emerging field of study within cybersecurity economics. Researchers in [4] provide a critical analysis using the Global Cost of Cyber Risk Calculator, forecasting potential costs in Lithuania, Latvia, and Estonia, and positing that sectors like public business and defense are at heightened risk.

Sophos research indicates that the cost of ransomware attacks has been steadily increasing over the years [7]. The average total cost of remediation from a ransomware attack more than doubled between 2019 and 2020, from $761,106 to $1.85 million [7]. Interestingly, the average ransom paid was comparatively much lower, at just $170,404 in 2020 [7]. The high cost of recovery is attributed to the complexity of advanced ransomware attacks, which combine automation with hands-on human hacking [7]. These complex attacks require more complex recovery processes, contributing to an overall increase in ransomware remediation costs [7].

Recovery costs differ from sector to sector and are influenced by a range of factors [7]. For instance, outdated and fragmented IT systems contribute to high remediation costs [7]. Organizations often have to completely rebuild their systems after a ransomware attack, costing both time and money [7]. Investing upfront in defences can cut down remediation costs [7]. For instance, upgrading systems before they come under siege from cyber-criminals is one way to minimize costs [7].

Despite the increasing threat of ransomware, many organizations believe they are prepared to thwart a breach [3]. However, half of the organizations surveyed still fell victim to ransomware last year [3]. This highlights the need for more effective strategies and measures to mitigate the impact of ransomware attacks. Security leaders are prioritizing the implementation of advanced technologies such as artificial intelligence (AI) and machine learning (ML) for faster threat detection [3]. Internet-of-Things (IoT) security and next-generation firewalls (NGFWs) are among the areas where leaders plan to invest [3].

Finally, existing literature on ransomware recovery has some limitations like the lack of detailed temporal data. The studies do not specify the recovery times associated with different types of IT assets nor do they specify step-by-step breakdowns of recovery processes. Also, existing papers frequently do not account for the variability in recovery procedures across different organizational structures and technological environments, which can lead to underestimations or overestimations of recovery costs and times.

3 Assessment

To obtain ransomware recovery times, it is important to collect the time organizations need to recover their assets after a ransomware attack. This time is of major importance to minimize operational downtime and ensure business continuity. Once there was a great lack of information and detail about recovery times, there was a need to create an assessment of recovery practices within various organizations. The purpose of this assessment was to gather data on current recovery capabilities and to analyze the effectiveness of these capabilities in diverse IT environments and scenarios.

The assessment consists of nineteen questions and is designed to extract information regarding the presence and efficacy of DR policies, the extent of redundancy systems, and the recovery time for various IT components. It is divided into two sections, in the first section the questions are about the existence of DR and the geographical distribution of the resources, and the second is aimed at the restoration capabilities for IT assets. The questions are structured to elicit both qualitative and quantitative data, with some emphasis on the difference in recovery times when usable backups are available compared to scenarios with no backups. The questions were selected to cover a broad range of IT assets, from workstations and servers to user applications and network peripherals. This approach will allow the assessment to evaluate transversely the recovery operations across different technology and infrastructure components. Also, by distinguishing between recovery times with and without backups, the assessment provides insight into the real-world impact of backup strategies on DR processes.

The following are the questions that will be made:

- (Q01) Does your organization have a DR policy?
- (Q02) Does your organization have multiple locations?
- (Q03) Does your organization have a redundancy system (or fail-over) for essential services to its operation?
- (Q04) If you answered Yes to the previous question, how much time (in hours) is needed to get your redundancy or fail-over operational?
- (Q05) Can you restore a group of workstations automatically and simultaneously?
- (Q06) If you answered Yes to the previous question, how many can you restore simultaneously?
- How much time do you need to restore:
  - (Q07) a workstation.
  - (Q08) user applications used by the institution.
  - (Q09) a server.
  - (Q10) an enterprise application.
  - (Q11) a database management system.
  - (Q12) an email system.
  - (Q13) a backup system.
  - (Q14) a cloud service or a virtual machine.
  - (Q15) a storage, NAS, or equivalent.
  - (Q16) a security system (for example, a firewall).
  - (Q17) network equipment (routers, switches, wireless access points, etc.).
  - (Q18) a mobile device or tablet.
  - (Q19) a printer or another peripheral used in your institution.

Questions Q07 to Q19 will be asked in two different scenarios: one in which the company has usable backups and one in which there are no usable backups.

4 Conclusions

As ransomware attacks continue to evolve in complexity and frequency, organizations need to enhance their cybersecurity measures. This paper responds to the need to create a recovery-cost utility function that will allow for the existence of a tool capable of calculating the total time needed for recovery of an attack.

The authors present a study on the ransomware recovery practices of several organizations, with focus on the variability of recovery times across different types of IT assets.

With this assessment, the authors pretend to offer a view of the current recovery practices from ransomware attacks. The presence of DR policies and redundancy will allow for the verification of the importance placed on preparedness. One of the findings that the authors intend to evaluate is the great variability in recovery time-frames, especially when there are no backups. The research will evaluate how many of the inquired organizations have disaster recovery policies and redundant systems. It will also verify if there is a significant variability between organizations' recovery time. This assessment can provide critical data that can be used for the development of a recovery cost-utility. The variability in recovery times, if any, can reinforce the need for an approach to estimating recovery times and costs. It can also confirm that the existence of backups can reduce recovery times, making them a key variable in the utility.

In conclusion, the research done so far demonstrates the complexity and difficulty of evaluating ransomware recovery times but also shows that the development of a recovery cost-utility function can represent an innovative approach that allows organizations to test, forecast, and plan for the implications of a successful ransomware attack.

References

1. Brewer, R.: Ransomware attacks: detection, prevention and cure. Netw. Secur. **2016**, 5–9 (2016). https://doi.org/10.1016/S1353-4858(16)30086-1
2. Cristea, L.: Current security threats in the national and international context. JOURNAL OF ACCOUNTING AND MANAGEMENT INFORMATION SYSTEMS (2020). https://doi.org/10.24818/jamis.2020.02007
3. Fortinet: The 2023 global ransomware report (2023), `https://www.fortinet.com/content/dam/fortinet/assets/reports/report-2023-ransomware-global-research.pdf`, [Online; accessed 9-April-2024]
4. Gavėnaitė-Sirvydienė, J., Miečinskienė, A.: Forecasting costs of cyber attacks using the global cost of cyber risk calculator v 1.2. Contemporary Issues in Business, Management and Economics Engineering **2021** (2021). https://doi.org/10.3846/cibmee.2021.618
5. Pattnaik, N., Nurse, J.R.C., Turner, S., Mott, G., MacColl, J., Huesch, P., Sullivan, J.: It's more than just money: The real-world harms from ransomware attacks. ArXiv **abs/2307.02855** (2023). https://doi.org/10.48550/arXiv.2307.02855
6. Sharma, S.K., Singh, T.: Ransomware : A severe cyber security threat in the digital era. International Journal of Scientific Research in Science and Technology **6**, 211–218 (2019). https://doi.org/10.32628/IJSRST196120
7. Sophos: Ransomware recovery costs vs. the ransom demands (2024), `https://www.ft.com/partnercontent/sophos/ransomware-recovery-costs-vs-the-ransom-demands.html`, [Online; accessed 9-April-2024]

Performance analysis of annotated paraconsistent logic applied to the development of artificial neural networks

João H. R. L. Sousa[1], Edson A. Batista[1], José L. S. de M. Lima[2], and Marcos R. Araújo[1]

[1] UFMS, Universidade Federal do Mato Grosso do Sul, Brasil
[2] CeDRI, Instituto Politécnico de Bragança, Portugal
j.henrique@ufms.br, edson.batista@ufms.br, jllima@ipb.pt, marcos.r.araujo@ufms.br

Abstract. Neural networks often employ a variety of activation functions, such as the Log-sigmoid and hyperbolic tangent. However, in environments with limited computational resources, the pursuit of computational efficiency is crucial. In this context, this study compares an alternative neural network method, the Paraconsistent Neural Network (PNN), with traditional neural networks. The PNN uses the annotated two-value paraconsistent logic algorithm as its activation function, significantly simplifying calculations throughout the process. The results of the comparisons, both with two and three neurons in the hidden layer, were promising. The PNN achieved a mean squared error (MSE) of 0.083 when employing two neurons in the hidden layer, while the neural network with the sigmoid function reached an MSE of 0.2835, and the network with the *tanh* function did not converge, registering an MSE of 1.1841. In the test with three neurons in the hidden layer, the PNN obtained an MSE of 0.0574, close to the MSE of 0.0515 of the hyperbolic tangent function, and lower than the MSE of 0.1647 of the sigmoid function. These results demonstrate that the PNN not only presented the best performance in the first scenario but also satisfactory performance in the second. The findings are in line with the study's objective, which aims to achieve results comparable to or superior with a lower computational load compared to conventional approaches.

Keywords: Neural networks · Computational efficiency · Paraconsistent neural network.

1 Introduction

Paraconsistent logic, situated within the realm of non-classical logics, exhibits the capability to address uncertainties, inconsistencies, and ambiguities in data. Studies indicate that managing contradictory situations, as observed in the real world, finds its optimal application through paraconsistent logics [4]. Within this domain, emerges the two-valued annotated paraconsistent logic (PAL2v), a branch enabling the treatment of contradictions by manipulating degrees of certainty and contradiction [3, 4, 11].

The PAL2v approach employs the para-analyzer algorithm to calculate the real resulting degree of evidence based on the inputs of favorable evidence degree (μ) and opposing evidence degree (λ) [1]. These degrees vary independently from 0 to 1. At the end of the process, the resulting degree of evidence is determined through mathematical operations that involve measuring the Euclidean distance of the certainty and contradiction degrees, derived from mi and lambda through specific equations, in relation to the point of absolute truth (1, 0).

In this context, the increasing deployment of intelligent systems grounded in paraconsistent logic is witnessed [9], where, for instance, neural networks are designed to translate the burgeoning volume of generated data into useful information [10].

Addressing this theme, there is an observable uptick in the utilization of paraconsistent neural cells in various arrangements, as documented in [2, 4, 6, 11], adapting to each proposed problem.

The objective of this work consists of discussing the efficacy of a shallow paraconsistent neural network, in comparison to conventional networks. To this end, three distinct neural networks were developed, employing: the Real Analytical Paraconsistent Artificial Neural Cell (RaPANC), the Log-sigmoid function, and the hyperbolic tangent function, respectively, as activation mechanisms for the neurons in the hidden layer.

Additionally, the error backpropagation technique, as elucidated in [7], was employed to adjust the synaptic weights of the networks, necessitating adaptations in the Paraconsistent Neural Network (PNN) due to the PAL2v algorithm requiring two inputs. In the proposed model, the bias of each neuron was fixed at 0.5 and remained unchanged, as it represents the point of indefiniteness in paraconsistent logic.

Finally, the mean squared errors of each network were calculated as described in [8], serving as one of the fundamental metrics adopted to evaluate the efficiency of a neural network.

2 Methods

As previously mentioned, the schematic of the PNN was developed, employing a representative model composed of two neurons in the input layer, two neurons in the hidden layer, and one neuron in the output layer, as outlined in Fig. 1. In this neural network

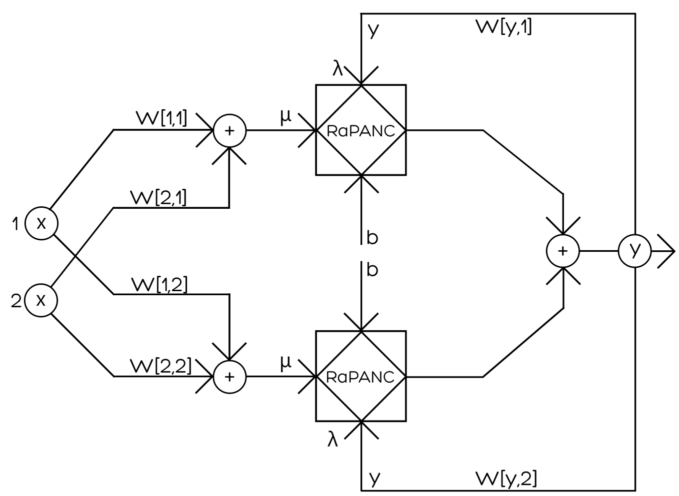


Fig. 1: Paraconsistent Neural Network model

representation, the absence of weights between the hidden layer and the output layer is evident. However, feedback loops from the output of the neural network connected to the input of the contradiction degree lambda of each cell in the hidden layer are observed, where such connections are established through synaptic weights. This approach deviates from the traditional neural network model, as illustrated in Fig. 2. The nomenclature adopted for the weights aims to facilitate the understanding and comparison of the methods presented here, where w[y,1] is numerically equivalent to w[1,y], and so forth. Continuing with the PNN, the algorithm of the real analytical artificial neural

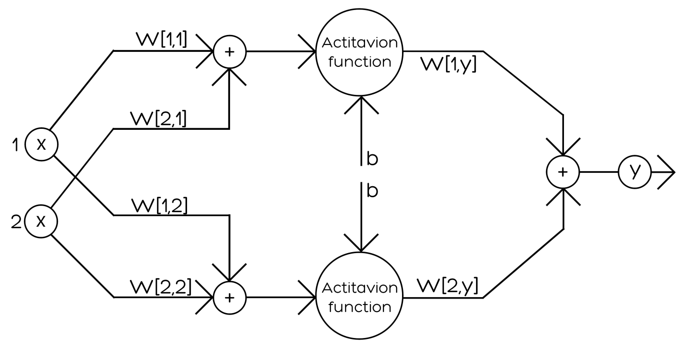


Fig. 2: Traditional Neural Network model

cell is presented in [4], along with biases fixed at 0.5 for all cells in the hidden layer and partial derivatives for weight adjustment, as detailed in [2]. Thus, Algorithm 1 contains the pseudocode designed for training the paraconsistent neural network elucidated in Fig. 1.

Algorithm 1 Paraconsistent neural network training algorithm.

Initialize the learning rate lr, the weights W and set the output $y = 0$.
for each epoch **do**
  for each input data **do**
    for each neuron **do**
      $\mu \leftarrow X_1 \cdot W[1, neuron] + X_2 \cdot W[2, neuron] + b$
      $\lambda \leftarrow y \cdot W[neuron, y] + b$
      $\mu_{ER} \leftarrow RaPANC(\mu, \lambda)$
    end for
    $y \leftarrow \sum \mu_{ER} - 0.5$
    $error \leftarrow y_{target} - y$
    for each neuron **do**
      $\delta_\mu, \delta_\lambda \leftarrow PartialDerivative(\mu_{ER})$
      $W[1, neuron] \leftarrow W[1, neuron] + \delta_\mu \cdot X_1 \cdot error \cdot lr$
      $W[2, neuron] \leftarrow W[2, neuron] + \delta_\mu \cdot X_2 \cdot error \cdot lr$
      $W[y, neuron] \leftarrow W[y, neuron] + \delta_\lambda \cdot y \cdot error \cdot lr$
    end for
  end for
end for

Similarly, Algorithm 2 presents the pseudocode for training a conventional neural network, where the derivatives of the log-sigmoid and hyperbolic tangent functions are expressed in equations (1) and (2), respectively.

$$\frac{d\sigma}{dx} = \sigma(x)(1 - \sigma(x)) \tag{1}$$

$$\frac{d\tanh}{dx} = 1 - \tanh^2(x) \tag{2}$$

Algorithm 2 Traditional neural network training algorithm.

Initialize the learning rate lr and the weights W.
for each epoch **do**
  for each input data **do**
    for each neuron **do**
      $u_{neuron} \leftarrow X_1 \cdot W[1, neuron] + X_2 \cdot W[2, neuron] + b$
      $y_{neuron} \leftarrow ActivationFunction(u_{neuron})$
    end for
    $y \leftarrow \sum y_{neuron} \cdot W[neuron, y]$
    $error \leftarrow y_{target} - y$
    $W_{old} \leftarrow W$
    for each neuron **do**
      $\delta \leftarrow Derivative(ActivationFunction)$
      $W[1, neuron] \leftarrow W[1, neuron] + \delta \cdot X_1 \cdot W_{old}[neuron, y] \cdot error \cdot lr$
      $W[2, neuron] \leftarrow W[2, neuron] + \delta \cdot X_2 \cdot W_{old}[neuron, y] \cdot error \cdot lr$
      $W[neuron, y] \leftarrow W[neuron, y] + \delta \cdot y_{neuron} \cdot error \cdot lr$
    end for
  end for
end for

The *tanh* and log-sigmoid functions were chosen because of their ability to introduce non-linearity into the network [5], a characteristic present in the data used in this study for training neural networks. Additionally, these functions are effective in capturing smooth and continuous variations, making them even more suitable for the specific data chosen.

Furthermore, as outlined in [8], it is possible to compute the mean squared error (MSE) using equation (3).

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{3}$$

3 Results and Discussion

Unconventional approaches in various research domains are being investigated to enhance performance in specific contexts or comprehensively. In this context, it remains to present a comparison of the training adequacy of each neural network, considering a set of two neurons in the hidden layer, 200 epochs, and a learning rate of 10%, as depicted in Fig. 3.

At first glance, a notable disparity is evident among the three neural networks, where the signals overlap, creating a turbulent scenario. Therefore, Fig. 4 illustrates an enlargement of the training for the last two epochs.

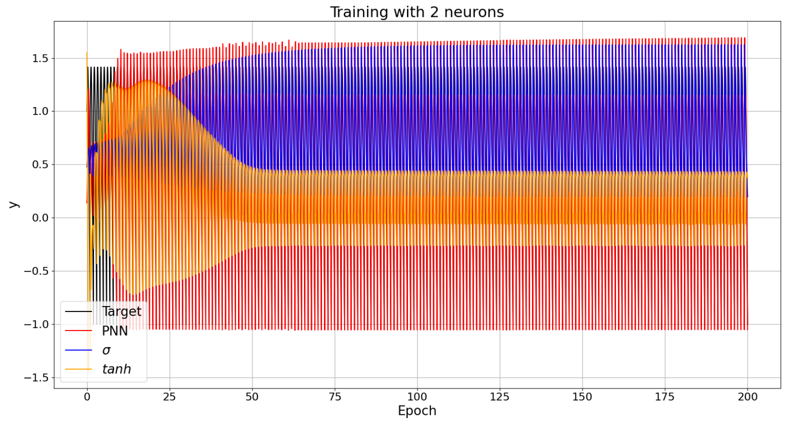


Fig. 3: Training neural network with 2 neurons

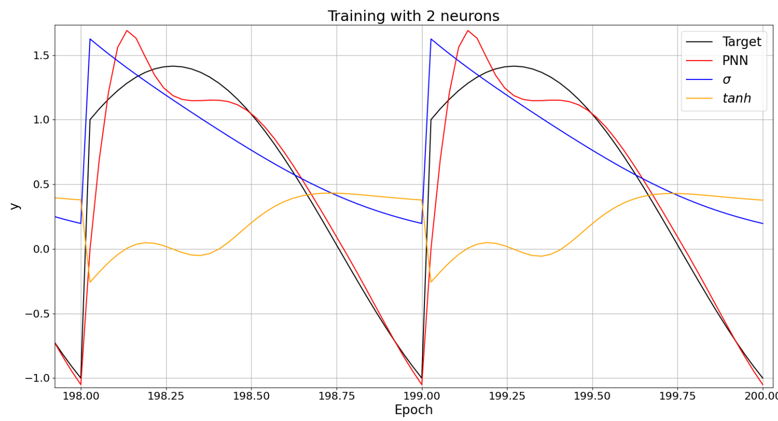


Fig. 4: Training neural network with 2 neurons - last 2 epochs

In this context, a more precise adjustment of the PNN is observed, where at the end of the 200 epochs, the model not only demonstrated a more suitable amplitude compared to the reference, but also was able to satisfactorily follow the reference compared to other activation functions. Thus, in seeking to measure the effectiveness of a neural network's adjustment, the equation for mean squared error was employed at the end of each epoch, as depicted in Fig. 5.

In this context, the MSE highlights a significant discrepancy among the neural networks for this specific application. The paraconsistent neural network, in this instance, demonstrated a more efficient ability to adjust the weights, resulting in a rapid reduction of the MSE and outperforming the neural network that employed the sigmoid function as the activation function. Additionally, it is observed that the neural network utilizing the hyperbolic tangent as the activation function failed to converge, exhibiting a significantly higher MSE compared to the other networks, considering the randomly chosen weights.

Similarly, it was decided to conduct an additional comparative analysis, this time employing three neurons in the hidden layer. The training results are depicted in Fig. 6 and 7, while the mean squared error is presented in Fig. 8.

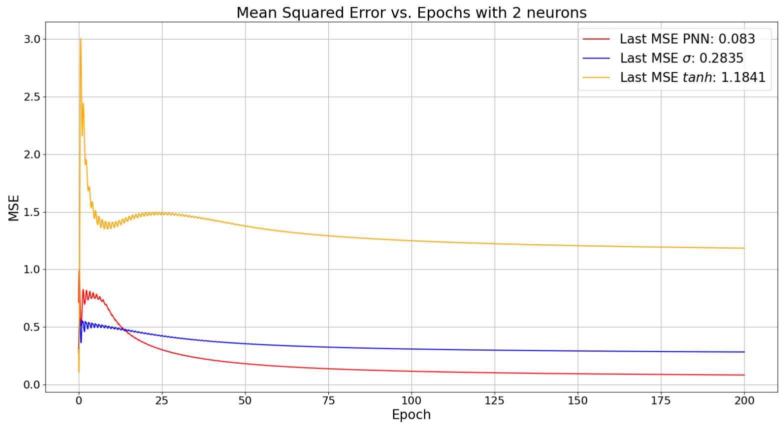


Fig. 5: Mean Squared Error of training a neural network with 2 neurons

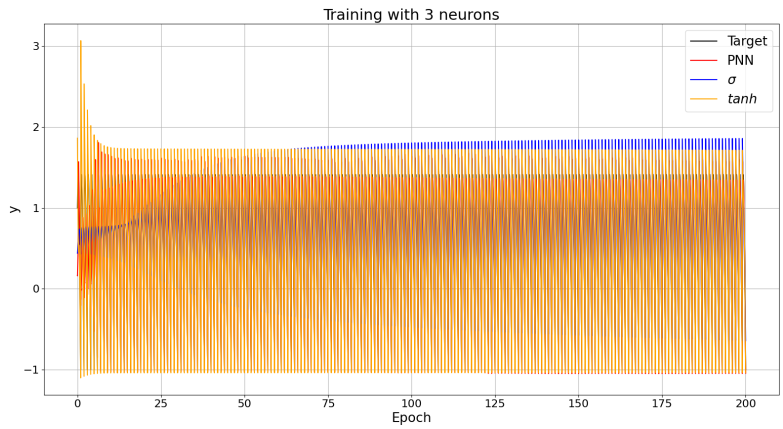


Fig. 6: Training neural network with 3 neurons

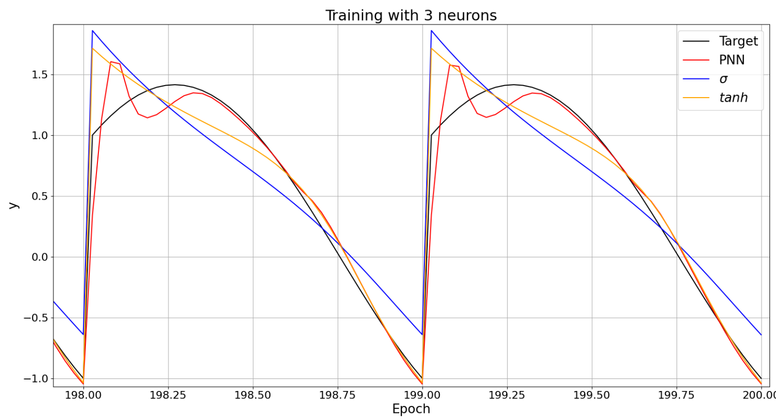


Fig. 7: Training neural network with 3 neurons - last 2 epochs

In Table 1, the comparative mean squared errors between the two tests for different activation functions are presented. When employing three neurons in the hidden layer, it is noteworthy that the neural network utilizing the *tanh* activation function

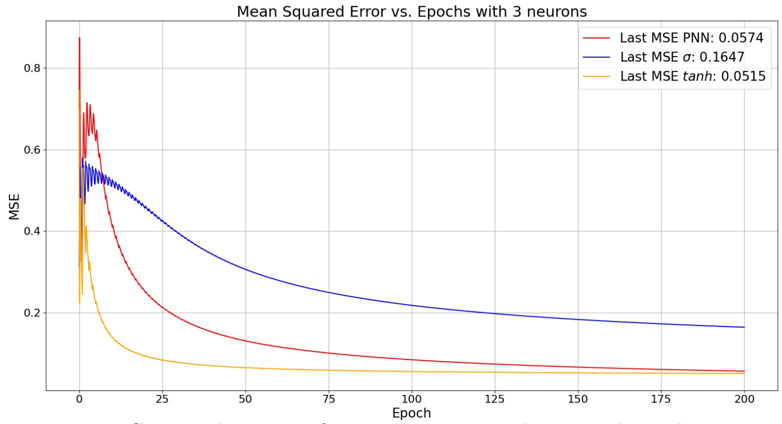


Fig. 8: Mean Squared Error of training a neural network with 3 neurons

Table 1: Comparative analysis of mean squared error using different activation functions and varying numbers of neurons in the hidden layer

Function	MSE with 2 neurons	MSE with 3 neurons
RaPANC	0.083	0.0574
Sigmoid	0.2835	0.1647
Hyperbolic tangent	1.1841	0.0515

demonstrated the greatest improvement compared to the previous analysis, followed by the neural network with the log-sigmoid activation function, and lastly, the PNN. This outcome is attributed to the satisfactory previous performance of the PNN and the non-convergence of the neural network with $tanh$ in the prior scenario. Additionally, it is observed that in this application, the neural network employing the hyperbolic tangent function achieved the lowest mean squared error. However, throughout the training process, the PNN approached this result considerably, resulting in similar final errors. Furthermore, upon examining the training graph, a convergence of outputs is perceived at a specific stage of each epoch between the outputs of the neural network with the hyperbolic tangent activation function and the PNN.

4 Conclusions and Future Works

The present study provided a detailed algorithm for constructing a paraconsistent neural network with a hidden layer, accompanied by a comparative analysis with conventional neural networks. This approach aimed to explore various techniques for solving problems using neural networks, with the goal of reducing computational cost through more efficient implementations that minimize the number of calculations required to achieve the desired results. Additionally, the simpler formulations of paraconsistent logic may make this method accessible in environments with limited computational resources, such as embedded systems, as discussed in [2, 4].

The PNN demonstrated versatility in the proposed scenarios, achieving a mean squared error of 0.083 when employing two neurons in the hidden layer. In contrast, the neural network with the sigmoid activation function reached an MSE of 0.2835, and the network with the *tanh* function failed to converge to the global minimum, registering an MSE of 1.1841. Furthermore, in the test with three neurons in the hidden layer, the PNN showed performance comparable to that of the hyperbolic tangent function, with MSEs of 0.0574 and 0.0515, respectively. The sigmoid function, although it presented a better MSE than in the previous scenario, still resulted in a relatively high value of 0.1647.

The results obtained demonstrated the effectiveness of this approach in the studied context, offering a viable alternative for solving similar problems. Future work could focus on exploring deep paraconsistent neural networks, improving optimization algorithms for weight updates and learning rate adaptation, or investigating the development of paraconsistent recurrent neural networks with long-term memory.

Acknowledgements The authors are grateful to the Foundation for Science and Technology (FCT, Portugal) for financial support through national funds FCT/MCTES (PIDDAC) to CeDRI, the Federal University of Mato Grosso do Sul, the National Council for Scientific and Technological Development (CNPq) and the Coordination for the Improvement of Higher Education Personnel (CAPES) UIDB/05757/2020 (DOI: 10.54499/UIDB/05757/2020) and UIDP/05757/2020 (DOI: 10.54499/UIDP/05757/2020) and SusTEC, LA/P/0007/2020 (DOI: 10.54499/LA/P/0007/2020).

References

1. Abe, J.M., Akama, S., Nakamatsu, K.: Introduction to annotated logics: foundations for paracomplete and paraconsistent reasoning, vol. 88. Springer (2015). https://doi.org/10.1007/978-3-319-17912-4
2. Carvalho Junior, A.d.: Identificação e controle de sistemas dinâmicos com rede neural paraconsistente. Ph.D. thesis, Universidade de São Paulo (2021). https://doi.org/10.11606/T.3.2021.tde-08102021-100149
3. Da Silva Filho, J.I., Abe, J.M., Marreiro, A.d.L., Martinez, A.A.G., Torres, C.R., Rocco, A., Côrtes, H.M., Mario, M.C., Pacheco, M.T.T., Garcia, D.V., et al.: Paraconsistent annotated logic algorithms applied in management and control of communication network routes. Sensors (2021). https://doi.org/10.3390/s21124219
4. Da Silva Filho, J.I., Torres, G.L., Abe, J.M.: Uncertainty treatment using paraconsistent logic: Introducing paraconsistent artificial neural networks. Ios Press (2010), ISBN: 978-1-60750-557-0
5. Goodfellow, I., Bengio, Y., Courville, A.: Deep learning. MIT press (2016), `http://www.deeplearningbook.org`
6. Kuntz, J.L.L.F.O., Mário, M.C.: Redes neurais artificiais paraconsistentes aplicadas no monitoramento óptico de produtos. Revista Brasileira de Computação Aplicada **13**(3), 62–76 (2021). https://doi.org/10.5335/rbca.v13i3.12427
7. Lamsal, R.: A step by step forward pass and backpropagation example (2021), `https://theneuralblog.com/forward-pass-backpropagation-example/`
8. Lamsal, R.: Understanding Gradient Descent Algorithm: The simplest way (2022), `https://theneuralblog.com/gradient-descent-algorithm/`
9. de Lima, A.W.B., de Lima, L.A., Abe, J.M., Gonçalves, R.F., Alves, D., Nakamatsu, K.: Paraconsistent annotated logic artificial intelligence study in support of manager decision-making. pp. 154–157 (2018). https://doi.org/10.1145/3278252.3278269
10. de Lima, L.A., Abe, J.M., Caique, J.S., Amado, N.: Um estudo em rede de análise paraconsistente (2018), `https://www.researchgate.net/publication/331384645`
11. Martinez, A.A., Abe, J.M., de Lima, L.A.: Deep learning paraconsistente: Perspectivas para a tomada de decisão. In: WICS0219 - Workshop on Intelligent Computer Science. pp. 17–43. Universidade Paulista (2021), anais Aspectos de Sistemas Inteligentes Baseados em Lógicas Anotadas, ISBN: 978-65-994728-0-0

Development of a Chatbot to support for IPB institutional website

Miguel Correia[1], Telmo Sampaio[1], and Pedro Filipe Oliveira[2]

[1] Instituto Politécnico de Bragança, Portugal
[2] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança, Campus de Santa Apolónia 5300-253 Bragança, Portugal
a48548@alunos.ipb.pt,a48561@alunos.ipb.pt,poliveira@ipb.pt

Abstract. This paper presents the development of a chatbot to support the institutional website of IPB, designed to enhance user interaction and information accessibility. Using techniques in natural language processing and machine learning, the chatbot extracts and processes data from the IPB website to provide accurate responses to user queries. Evaluation results demonstrate high accuracy in answering questions about IPB courses, showcasing the effectiveness of the developed system. Overall, the study underscores the potential of chatbots in improving user experience and information dissemination on educational websites.

Keywords: chatbot · machine-learning · html-scrapping.

Acronyms: IPB - Instituto Politécnico de Bragança, LLM - Large language model

1 Introduction

Currently, there is a significant increase in the world of online services, where everything is now being linked to websites [3]. As new generations, especially university students, show more and more interest in social networks than traditional websites for interaction and obtaining information, there is a preference for direct contact through the instant messaging mechanisms available on these platforms, rather than the use of indirect contact through non-instant messaging such as e-mail. These students prefer to use instant messaging because it is an effective method of communication that allows users to instantly share digital information such as text, audio and video [2]. Against this backdrop, there has been considerable investment in strengthening the presence on social networks, as well as the availability of interlocutors ready to answer questions posed via these platforms.

These interlocutors are also known by many other names, such as chatBot, dialog system, conversational agent, conversational interface, virtual assistant,and personal assistant. In higher education, chatbots facilitate more efficient interaction, promote sociability and provide access to information, which positively influences the educational flow to make it more interactive and dynamic. [1] [4]

This work aims to explain the state-of-the-art technique used and the purposes associated with building a chatbot. With the aim of providing comprehensive information for a complete understanding of how this system works. In addition to the theoretical approach, the main objective is to develop a chatbot to support the IPB website.

2 Materials and Methods

Figure 1, shows the process of obtaining the information from the IPB site to feed the chatbot.

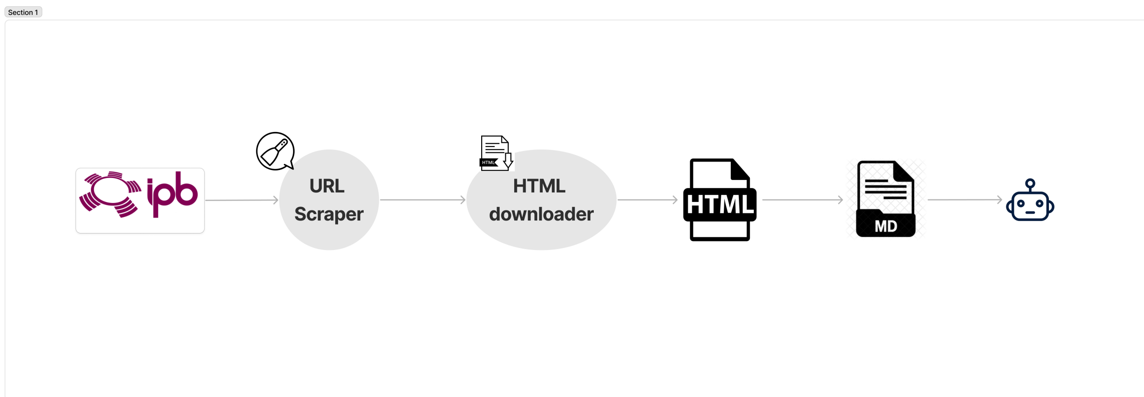


Fig. 1: Web scraping

To obtain the information to feed the chatbot, a script was used to obtain all the urls present in the ipb domain, after which the scraper was used to iterate through all the urls present and download the html content.

After the scraper, we used a separate script to transform the HTML content into Markdown format. This approach was essential in feeding the information used by the chatbot. The script was designed to filter out data we deemed unnecessary, such as the cookies from the HTML, and the menu on the right side of the IPB portal.

In the initial stages, we also removed any links from the site to optimize the chatbot's performance. However, recognizing the potential utility of links for user engagement and information accessibility, we are now working to incorporate them back into the chatbot's dataset. This enhancement will further enrich the user experience by providing relevant and direct sources of information during interactions with the chatbot.

Figure 2 shows how the information was fed to the chatbot is divided, stored, and then searched to give the user a correct answer based on the information present in the files.

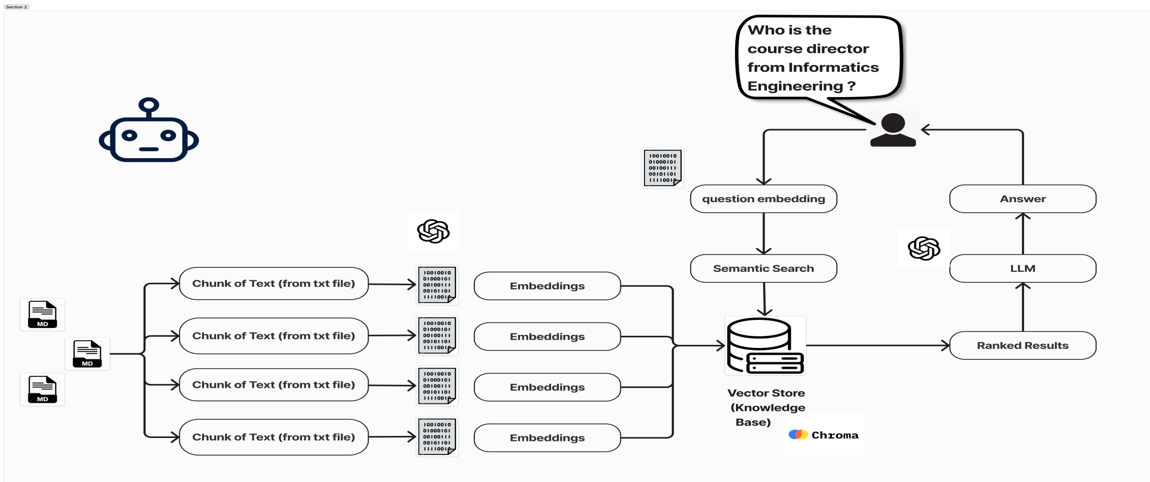


Fig. 2: Chatbot functioning

The md files are converted into smaller text chunks which, in turn, are transformed into embedding, which are numerical representations of the text that can be used for machine learning tasks. The embedding are now stored in a Vector Store where they can be accessed later. When the language model receives a question, a query is made in the Vector Store, which looks for relevant information based on the semantic similarity between the embedded question and the stored text embedding. The results of the semantic search are then sorted, returning the best possible answer to the question.

3 Results

To assess the results, the chatbot was asked 10 different questions about IPB courses,the accuracy of the results can be seen in the following table:

Table 1: Accuracy of the answers

Question	Accuracy
1	100%
2	50%
3	100%
4	100%
5	100%
6	0%
7	100%
8	100%
9	100%
10	100%

The answer where it is only 50% correct was an ambiguity problem, where there is a bachelor's and a master's course with the same name, and when asked a question about that course the chatbot always answers with the master's information. To mitigate such issues, we plan to implement follow-up questions into the chatbot's design. This enhancement will allow the chatbot to request additional information when a user's question lacks sufficient context, thereby ensuring a more accurate response. In the incorrect answer the chatbot was asked about the access requirements for a certain course, but instead of answering with the specific IPB requirements it was giving general information, meaning that the bot isn't reading the information of the files correctly. Overall, the results are very promising, with most answers being 100% correct and only one answer being completely wrong.

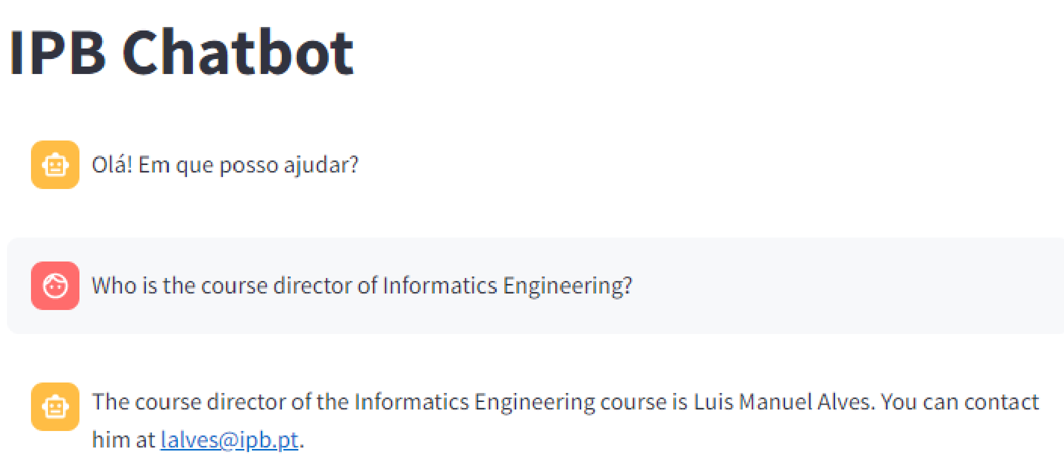


Fig. 3: Result of the first question

4 Conclusions

The development and assessment of the chatbot for the institutional website of the IPB have yielded promising results. By leveraging techniques such as web scraping, text processing, and semantic search, the chatbot demonstrated high accuracy in providing information about IPB courses. Despite minor inaccuracies, the majority of the answers were correct, indicating the efficacy of the implemented approach.

In the future, we will try to deploy the chatbot on the IPB server.This exposure will not only increase its usefulness of the chatbot but also provide us with valuable user feedback. With this we will continue to refine and enhance the chatbot, making it an even more effective and efficient tool for engaging with users.

References

1. Altinok, D.: An ontology-based dialogue management system for banking and finance dialogue systems. CoRR **abs/1804.04838** (2018), http://arxiv.org/abs/1804.04838
2. Barry, B., Tom, F.M.: Instant messaging: Standards, protocols, applications, and research directions (2010)
3. Ranoliya, B.R., Raghuwanshi, N., Singh, S.: Chatbot for university related faqs. In: 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). pp. 1525–1530. IEEE (2017). https://doi.org/10.1109/ICACCI.2017.8126057, https://ieeexplore.ieee.org/abstract/document/8126057/authors#authors
4. Sjöström, J., Aghaee, N., Dahlin, M., Ågerfalk, P.J.: Designing chatbots for higher education practice (2018), https://aisel.aisnet.org/siged2018/4/

Prototyping and application of an attendance control solution in an academic context, using Beacons

Sily Soares Simas[1] and Pedro Filipe Oliveira[2]

[1] Instituto Politécnico de Bragança, Portugal
a43759@alunos.ipb.pt
[2] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
poliveira@ipb.pt

Abstract. This project addresses the prototyping and application of an attendance control solution in an academic context, using Beacons. The introduction highlights the importance of effectively managing student attendance to ensure active participation in educational activities and provide insights into student engagement. BLE Beacons technology is compared to other RF signalling technologies, highlighting its advantages in terms of range, power consumption, cost, frequency, and data rate. BLE Beacon is chosen as the preferred technology due to its combination of adequate range, low power consumption, affordable cost, compatibility, security, and flexibility. The results section presents Estimote Beacons as devices used, highlighting their advanced features and techniques used, such as proximity and triangulation, to determine student presence. A visual summary of the project is presented in a flowchart that illustrates the operating processes of each stage. This solution promises to significantly improve the efficiency, safety, and quality of the academic environment, while simplifying the attendance recording process and providing valuable insights into student behavior.

Keywords: attendance-control · beacons · BLE.

1 Introduction

In recent years, technology has played an increasingly significant role in optimizing processes in various sectors, and education is no exception. In a dynamic and constantly evolving academic environment, effective management of student attendance is essential not only to ensure active participation in educational activities, but also to provide valuable insights into student engagement and teaching effectiveness.

In this context, the prototyping and application of an attendance control solution using Beacons emerges as a promising innovation. Beacons are small, low-cost devices that transmit short-range Bluetooth signals, allowing accurate identification of the location of nearby mobile devices. By integrating this technology with academic management systems, it is possible to create an efficient and automated solution to monitor student presence in the classroom, laboratories, libraries, and other areas of the campus.

This approach not only simplifies the attendance recording process, eliminating the need for error prone manual methods, but also offers a range of additional benefits. By collecting real-time data on student attendance, educators can identify patterns of behavior, detect absence trends, and proactively intervene to offer personalized support to students at risk of disengagement. Therefore, this project not only seeks to explore the possibilities offered by Beacons technology, but also aims to contribute to the continuous improvement of educational processes, promoting efficiency, safety, and quality in the academic environment.

2 Analysis and Methods

2.1 Bluetooth Low Energy Beacons (BLE Beacons) VS others Radio Frequency Beacons (RF)

BLE Beacons technology has numerous advantages compared to other radio frequency beacons, and figure 1 highlights each of these technologies and compares them to each other.

Fig. 1: Analysis and comparisons between technologies and radio beacons.

RF signaling technologies vary widely in terms of range/range. While Bluetooth Low Energy (BLE), Wi-Fi and Zigbee have a similar range of 10 to 100 meters, RFID can reach up to 100 meters. On the other hand, LoRa is capable of a much greater range, reaching up to 15 kilometers, while LTE can reach even longer distances, up to 30 kilometers. NFC has a very short range, less than 20 centimeters.

Regarding latency, BLE Beacon, WI-FI, Zigbee, RFID, NFC, have a low latency with values in milliseconds (ms), compared to other technologies such as LTE Beacon, which is medium and LoRa which has a high latency with values presented in seconds (s).

The frequencies vary from each other, with Bluetooth Low Energy operating at 2.4GHz, while Wi-Fi can operate on both 2.4GHz and 5GHz frequencies. Zigbee operates on 2.4GHz, 868MHz, and 915MHz, while RFID covers a range of frequencies including 125kHz, 13.56MHz, 433MHz and 900MHz. LoRa operates on 433 MHz, 868

MHz and 915 MHz, while NFC operates exclusively on 13.56 MHz. Finally, LTE operates on multiple frequencies. The data rate of BLE Beacons is higher than Zigbee Beacons, NFC Beacons, RFID, and LoRa making it ideal for applications that require faster data transfer.

BLE technology uses 2.4 GHz frequency band, and has a greater range than NFC Beacons and RFID Beacons, and consumes less energy than WI-FI Beacons and LTE Beacon.

The deployment cost of BLE Beacons is lower than WI-FI and LoRa Beacons, making them an economical solution for many applications [2]. Therefore, the choice of which RF beacon to use depends on the specific application requirements. For example, BLE Beacons can be suitable for indoor location-based services that require low power consumption, while RFID beacons can be suitable for inventory tracking that requires long-range reading [2].

2.2 Why BLE Beacons?

There are many factors that led us to choose this technology over another. BLE Beacons stand out as an ideal choice due to their combination of adequate range, low power consumption, affordable cost, compatibility and ease of integration, security, and flexibility.

- Moderate Range and Adequate Accuracy: BLE Beacons have a range of up to 100 meters, which is suitable for academic environments, where accuracy in detecting student presence is essential. While the range is moderate, it is enough to cover classrooms, hallways and common areas;
- Low Power Consumption: BLE Beacons devices consume considerably less power than other technologies, which is crucial in an academic environment where student mobile devices need to operate for long periods without recharging;
- Affordable: Deploying BLE Beacons is generally more cost-effective than other options such as Wi-Fi and LoRa, making it an affordable solution for academic institutions with limited budgets;
- Security and Privacy: BLE technology offers robust security and privacy, ensuring that student attendance data is transmitted securely and protected against unauthorized access;
- Flexibility and Scalability: BLE Beacons are highly flexible and scalable, allowing it to be used in a variety of academic environments and to easily increase the number of devices as needed.

And for these factors we chose to implement this technology to control attendance in the campus.

3 Results

This section presents the technologies used as well as the device and technique used to develop the project. Figure 2 [3] illustrates a Beacon device used in the project,

Fig. 2: Beacon Estimote

a great Beacon, with numerous features, and easy to implement, especially in indoor environments.

They are Bluetooth signaling devices manufactured by the company Estimote Inc. These devices are widely recognized for their quality, reliability, and advanced features.

They feature an elegant and compact design, usually in the form of small rectangular devices and are designed to be discreet so that they can be placed in different locations, such as walls, shelves, or even personal items.

They offer a wide range of advanced functionality, including:

- Motion detection;
- Temperature and humidity monitoring;
- Remote configuration;
- Advanced customization.

Estimotes Beacons are compatible with a variety of application development platforms, such as IOS and Android, making it easy to integrate with existing applications and develop new location-based solutions. Their reliability, quality, and advanced features make them a popular choice among businesses and developers looking to make the most of the potential of Beacon technology.

3.1 Techniques Used

Proximity and Triangulation:
There are several different methods that are used as position detection techniques, but in this project only two of these techniques were chosen, proximity and triangulation.

Proximity:
This technique determines the location of a moving object in relation to a known fixed position or reference point. When a device is sufficiently close to a Beacon, this information is used to estimate its position in relation to that Beacon, as the distance between them cannot be measured, due to the estimated positions.

The only information obtained from this detection is whether or not the mobile device is within the radar range of the transmitting device. If one of the devices detects a moving object, the location close to it is considered. If the signal is detected by two or more devices, the position is assigned to the closest Beacon device with the highest signal strength.

To better understand this position detection technique, figure 3 illustrates a small scene with two transmitting devices (beacons 1 and 2) and several mobile devices (A, B, C, D, and E) located in different positions.

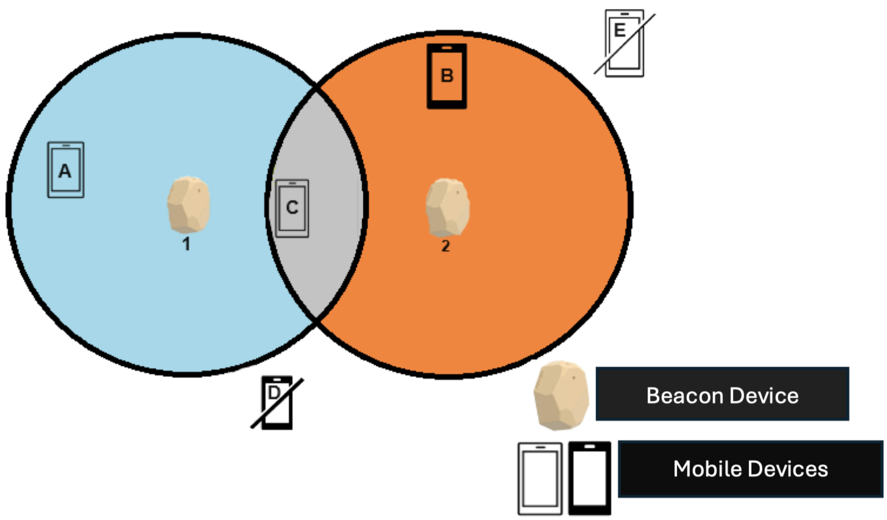


Fig. 3: Practical example of the Proximity technique.

Triangulation:
This technique is based on the geometric properties of triangles and can be based on a technique based on distance (lateration) or angular difference (angulation) between devices. By measuring the propagation time of the signal between the transmitter and several receivers, and knowing its propagation speed in advance, it is possible to estimate the distance of the locatable device to various reference points.

Likewise, measurements of the received power level, RSSI, make it possible to calculate the distance traveled by the signal by relating the emission and reception power with a propagation loss model. By obtaining at least three distances it is possible to estimate the position of the device. For a better understanding, figure 4 [4] illustrates an explanatory scenario with 3 Access Points (A, B, C).

3.2 Examples of Obtained Results

To illustrate the effectiveness of Estimote Beacons in attendance control, we present some examples of results obtained during the project tests.

Example 1: Classroom Presence Monitoring During a two-week period, Beacons were installed in a classroom to monitor student presence. The results were as follows:

- Number of students monitored: 30
- Presence detection accuracy: 95
- Average detection failures per day: 1.5 (due to sporadic interferences)
- User feedback: 85% of the students found the technology non-invasive and easy to use.

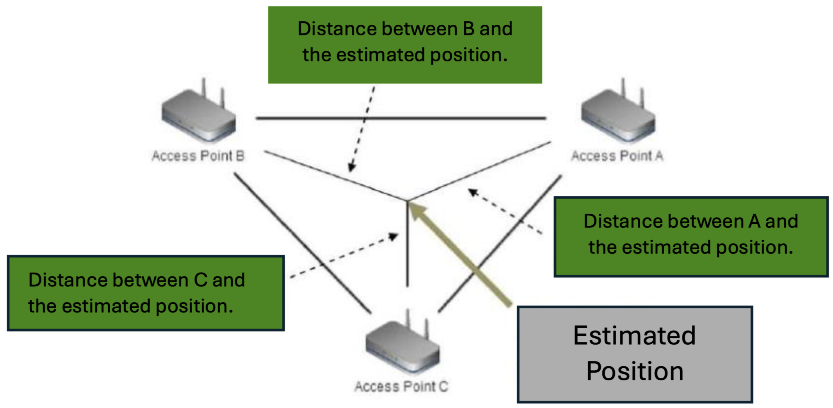


Fig. 4: Practical example of the Triangulation technique [1].

Example 2: Movement Detection in a Laboratory Beacons were used in a laboratory to monitor student movement between different workstations. The results showed:

- Number of workstations monitored: 10
- Average movements detected per session: 50
- Movement detection accuracy rate: 92
- Observations: There were some detection failures in areas with many metallic equipment pieces that interfered with the Beacon signals.

Example 3: Common Areas Monitoring on Campus To evaluate the coverage and effectiveness of Beacons in common areas of the campus, such as libraries and corridors, the devices were strategically positioned. The results were:

- Number of Beacons installed: 15
- Total monitored area: 1500 m^2
- Average location accuracy: 90%
- Number of entry/exit events detected per day: 200
- Challenges encountered: Interferences caused by large gatherings of people during peak hours.

These examples demonstrate the ability of Estimote Beacons to monitor the presence and movement of individuals in various academic environments, providing real-time data that can be used to improve campus management and security.

Figure 5 [1] illustrates a flowchart of the project, explaining the operating processes of each stage.

4 Discussion/Limitations

Despite the numerous advantages of using Beacons for attendance control in an academic environment, some limitations should be considered:

- Physical Obstacles: Walls, furniture, and other obstacles can affect the accuracy of the Beacons' signals, reducing the effectiveness of presence detection.

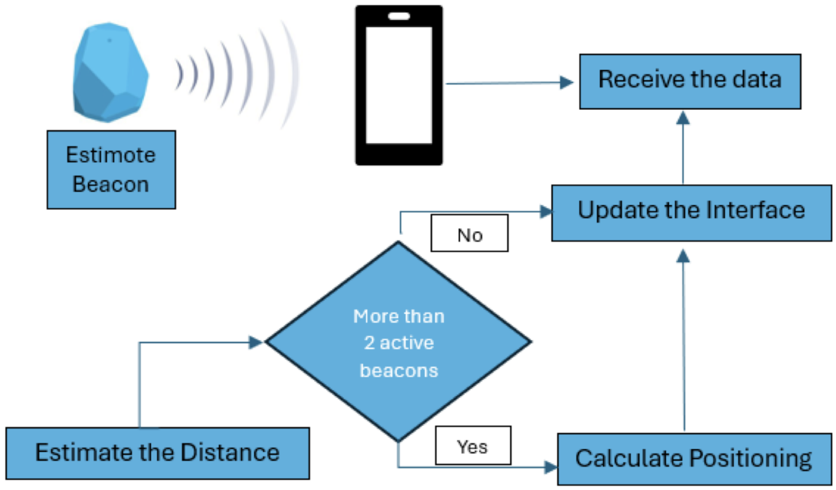


Fig. 5: Interface Flowchart.

- Electromagnetic Interference: Electronic devices and Wi-Fi networks can interfere with Beacon signals, leading to potential data collection failures.
- Battery Life: Despite low energy consumption, Beacons still require regular maintenance to replace batteries, which can be challenging on a large campus.
- Regular Maintenance: The need for periodic maintenance can increase operational costs and require additional effort for monitoring and logistics.
- Initial Investment: Installing Beacons on a campus may require a significant initial investment in terms of hardware and software acquisition, as well as integration with existing systems.
- Necessary Infrastructure: Effective implementation of Beacons may require a robust network infrastructure and specialized technical support to handle potential issues and updates.
- Continuous Movement: In areas with high foot traffic, such as corridors and common areas, the accuracy of presence detection may be compromised.
- Environmental Changes: Frequent changes in room layouts and spaces may require adjustments to the Beacons to maintain system effectiveness.

These limitations provide a critical view of the challenges that may arise with the implementation of Beacon technology, allowing for the identification of areas for improvement and future development.

5 Conclusions

The implementation of an attendance control solution in an academic environment using Beacons represents a significant advance in the optimization of educational processes.

Throughout this project, we highlight the advantages of Beacons, especially Estimote Beacons, compared to other RF signaling technologies. Beacons' ability to provide adequate range, low power consumption, affordability, compatibility, security, and flexibility makes them an ideal choice for monitoring student presence in classrooms, laboratories, and other areas on campus.

Furthermore, the proximity and triangulation techniques used in this project allowed accurate detection of student presence, providing real-time data that can be used to identify behavior patterns, detect absence trends, and offer personalized support to students at risk of disengagement.

The successful integration of Beacon technology in this academic context not only simplifies the attendance recording process, eliminating error-prone manual methods, but also provides valuable insights for educators and administrators, promoting continuous improvement in the efficiency, safety, and quality of the environment academic In the future, this solution is expected to evolve further, taking full advantage of the potential of Beacons and other emerging technologies to further improve educational processes and promote student engagement.

References

1. Bouchard, K., Ramezani, R., Naeim, A., et al.: Evaluation of bluetooth beacons behavior. In: 2016 IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). pp. 1–3. IEEE (2016)
2. Herrera Vargas, M.: Indoor navigation using bluetooth low energy (ble) beacons (2016)
3. Inaya, M., Meli, M., Sikeridis, D., Devetsikiotis, M.: A real-subject evaluation trial for location-aware smart buildings. In: 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). pp. 301–306. IEEE (2017)
4. Mohsin, N., Payandeh, S., Ho, D., Gelinas, J.P., et al.: Study of activity tracking through bluetooth low energy-based network. Journal of Sensors **2019** (2019)

Understanding Consumer Adoption of Electric Vehicles: A Comprehensive Review

Miguel Francisco P. Gonçalves[1] and Clara B. Vaz[1]

Research Center in Digitalization and Intelligent Robotics (CeDRI), Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
miguefpimenta@hotmail.com, clvaz@ipb.pt

Abstract. The transportation sector plays a significant role in greenhouse gas emissions, requiring the adoption of low-emission vehicles to mitigate climate change and address air quality concerns. Sustainable transportation policies worldwide emphasize strategies aimed at reducing emissions and promoting the adoption of electric vehicles (EVs) and plug-in hybrid electric vehicles (PHEVs). Consumer adoption is the main focus of this work, and in regards to EVs, it seems that consumer adoption is influenced by a multitude of factors, including cost considerations, range anxiety, charging infrastructure accessibility, limited model options, and technological evolution. Understanding these factors within the specific context of consumer behavior and preferences is essential for promoting widespread EV adoption. Through an extensive review of the literature, this study provides insights into the potential determinants of EV adoption, addressing the challenges and opportunities inherent in the transition to electric mobility. Furthermore, the study explores the current state of electric vehicles mainly in Portugal, highlighting a trend of rapid growth and adoption driven by maturing markets and advancing technology. Analyzing sales data from 2010 to 2023 provides valuable insights into regional dynamics and the progress of EV adoption in Portugal. The findings underscore the importance of understanding consumer preferences and behaviors in shaping the future of electric mobility, emphasizing the need for targeted interventions and strategies to overcome barriers and accelerate the transition to sustainable transportation.

Keywords: Electric vehicle · Battery electric vehicle · Plug-in hybrid electric vehicle · Hybrid electric vehicle

1 Introduction

Transportation is a major contributor to GHG emissions, accounting for a significant portion of total emissions in the United States [4]. Passenger cars and light and heavy trucks, in particular, contribute significantly to these emissions [7]. As a result, effective policies are necessary to mitigate GHG emissions and address climate change through sustainable transportation practices.

The adoption and advancement of electric vehicles (EVs) have become imperative in addressing the pressing need for sustainable transportation solutions, particularly in light of rising greenhouse gas (GHG) emissions from the transportation sector. This study focuses on understanding the key factors that can influence consumer adoption of electric vehicles and aims to shed light on the drivers behind their decision-making process.

[6] highlight the importance of sustainable transportation policies in reducing GHG emissions. Strategies employed globally, including those in European countries, Asia,

Africa, Australia, and the United States, can be broadly classified into three groups: Reduce, Avoid, and Replace, aiming to reduce GHG emissions per passenger per kilometer, avoid unnecessary energy consumption, and replace fossil fuels with low-emission alternative fuels [5, 6]

To improve sustainability in road transportation, the widespread adoption of plug-in electric vehicles (PEVs), including battery electric vehicles (BEVs) or Electric vehicles (EVs) and plug-in hybrid electric vehicles (PHEVs), is considered a key strategy [8]. PEVs offer several advantages, including reduced vehicle operating costs, decreased reliance on fossil fuels, and notably lower GHG emissions [2]. Therefore, understanding the factors that influence PEV adoption is crucial for assessing their environmental and economic impacts [9].

2 Literature Review

By closely examining the factors that contribute to the adoption of EVs, several studies [8–10] provide insights into a landscape of rapid growth and exciting opportunities on a global scale. The potential benefits of EVs, including reduced dependence on fossil fuels, decreased emissions, and improved energy security, position them as crucial elements in the future of sustainable mobility, not only in developed countries but also in developing nations. As the market continues to evolve and technology advances further, electric vehicles are expected to play an increasingly significant role in shaping the future of transportation worldwide. The global electric vehicle market is expanding at an accelerated pace, with an increasing number of EVs on the road. In 2021, although EVs represented a small fraction of total vehicle sales, their market share exponentially continues to grow, driven by the maturing market and advancing technology as shown is Fig. 1.

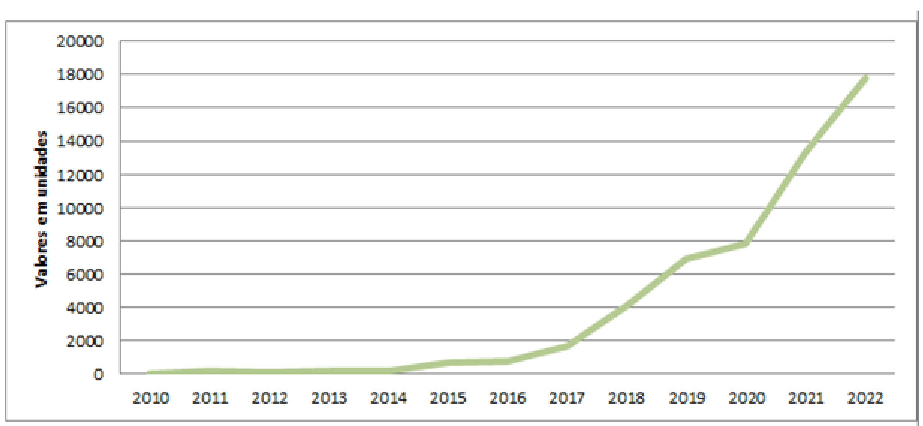


Fig. 1: EV's sales in Portugal 2010-2023 Source: [1]

Also, technological advancements, particularly in battery technology, have played a pivotal role in shaping the state of EVs. [10] discuss recent improvements in battery technology that have led to significant enhancements in the performance and cost-effectiveness of EVs. These advancements include the development of more advanced and longer-lasting batteries, as well as the establishment of faster and more convenient charging infrastructure. These breakthroughs have increased the accessibility and attractiveness of EVs to a broader range of consumers, while also reducing the key barriers to widespread adoption.

Despite the impressive growth and technological advancements, several challenges persist in the EV market. [3] highlight the need for a comprehensive charging infrastructure that is universally accessible and convenient. Meeting this challenge requires significant investments in infrastructure and technology, as well as collaboration among governments, industry stakeholders, and other relevant parties. Developing a robust charging infrastructure capable of meeting the needs of an expanding population of EV owners is essential for sustaining the growth and development of the EV market.

3 Conclusion

Synthesizing, mainly, the findings of the literature review we start to understand the complex dynamics influencing consumer adoption of EVs. Policymakers, industry stakeholders, and researchers must collaborate to address consumer concerns and promote the widespread adoption of EVs. Investments in charging infrastructure, technological advancements, and supportive policies are essential for facilitating the transition to sustainable transportation. In conclusion, understanding the factors influencing consumer adoption of EVs is essential for promoting sustainable transportation practices. By addressing consumer concerns and investing in infrastructure and policy support, stakeholders can hopefully accelerate the transition to a low-emission transportation system as well as better overall infrastructure. Future research however, should continue to explore consumer preferences and motivations to inform effective strategies for promoting EV adoption.

References

1. ACEA: Tech. rep. (2023), https://apve.pt/estatisticas/
2. Bradley, T.H., Quinn, C.W.: Analysis of plug-in hybrid electric vehicle utility factors. Journal of Power Sources **195**(16), 5399–5408 (2010). https://doi.org/https://doi.org/10.1016/j.jpowsour.2010.02.082, https://www.sciencedirect.com/science/article/pii/S0378775310003915
3. Duvall, T., Englander, F., Englander, V., Hodson, T.J., Marpet, M.: Ethical and economic issues in the use of zero-emission vehicles as a component of an air-pollution mitigation strategy. Science and Engineering Ethics **8**(4), 561–578 (2002). https://doi.org/10.1007/s11948-002-0009-2
4. IEA: World energy outlook 2022. Tech. rep., International Energy Agency (2022), www.iea.org/t&c/
5. J. Javid, R., Nejat, A., Hayhoe, K.: Selection of co2 mitigation strategies for road transportation in the united states using a multi-criteria approach. Renewable and Sustainable Energy Reviews **38**, 960–972 (2014). https://doi.org/https://doi.org/10.1016/j.rser.2014.07.005, https://www.sciencedirect.com/science/article/pii/S1364032114004572
6. Javid, R.J., Nejat, A.: A comprehensive model of regional electric vehicle adoption and penetration. Transport Policy **54**, 30–42 (2017). https://doi.org/https://doi.org/10.1016/j.tranpol.2016.11.003
7. Sarabi, S., Davigny, A., Courtecuisse, V., Riffonneau, Y., Robyns, B.: Potential of vehicle-to-grid ancillary services considering the uncertainties in plug-in electric vehicle availability and service/localization limitations in distribution grids. Applied Energy **171**, 523–540 (6 2016). https://doi.org/10.1016/j.apenergy.2016.03.064

8. Tal, G., Nicholas, M.A., Davies, J., Woodjack, J.: Charging behavior impacts on electric vehicle miles traveled. https://doi.org/10.3141/2454-07 **2454**, 53–60 (1 2014). https://doi.org/10.3141/2454-07, https://journals.sagepub.com/doi/10.3141/2454-07
9. Tamor, M.A., Gearhart, C., Soto, C.: A statistical approach to estimating acceptance of electric vehicles and electrification of personal transportation. Transportation Research Part C: Emerging Technologies **26**, 125–134 (2013). https://doi.org/https://doi.org/10.1016/j.trc.2012.07.007, https://www.sciencedirect.com/science/article/pii/S0968090X12000988
10. Uppaluri, M., Mishra, L., Subramaniam, A., Jang, T., Subramanian, V.R.: Modeling cathode/separator/lithium metal sandwich including porous electrode effects and plating/stripping - proper formulation, transformation, and efficient numerical simulation. ECS Meeting Abstracts **MA2022-02** (2022). https://doi.org/10.1149/ma2022-023277mtgabs

Tailored Training Solutions for Industrial Professionals: A Personalized Learning Platform

Helena A. Correia, Duarte Duque and Jõao L. Vilaça

2AI - Applied Artificial Intelligence Laboratory - Polytechnic Institute of Cávado and Ave, Barcelos, Portugal
hacorreia@ipca.pt, dduque@ipca.pt, jvilaca@ipca.pt

Abstract. According to the Future of Jobs 2023 report, skills gaps are the main obstacle to Industry 5.0 transformation. As industries transition towards sustainable and human-centric practices, the demand for training in complex industrial systems intensifies. Several authors have developed structured training programs, but concerns exist about oversimplification and generalization. Conventional methods fail to equip professionals with the necessary skills, resulting in a gap between training and practical tasks. This study seeks to bridge this gap by developing a digital platform for industrial training and customized learning. The project aims to deliver an immersive learning experience tailored to the specific requirements of Industry 5.0 and the individual skills of professionals, utilizing gamification techniques, artificial intelligence algorithms, and virtual/augmented reality technology. By crafting personalized training modules and integrating immersive technology, this study strives to cultivate a highly skilled and versatile workforce proficiently equipped to navigate the challenges inherent in a rapidly evolving industrial landscape.

Keywords: eLearning Platform · Immersive Technology · Artificial Intelligence · Gamification Techniques

1 Introduction

According to the Future of Jobs 2023 report, skills gaps in the local labor market pose a significant barrier to achieving Industry 5.0 transformation [1]. This shift emphasizes human-centric approaches and highlights the need for a highly skilled workforce to navigate modern challenges and technological advancements [2]. As Industry 5.0 reshapes professional roles, personalized training becomes increasingly crucial to prepare professionals for this new era, fostering continuous development and empowering them to adapt to evolving industrial demands [3].

Despite the recognition gained by structured training programs, concerns persist about oversimplification and generalization in their methods [4], [5], [6]. Current approaches fail to equip professionals with the skills needed to thrive in today's industrial environments, leading to a gap between training outcomes and practical tasks.

Recognizing this gap, the Industrial Metaverse emerges as a solution, offering immersive and interactive experiences tailored to professionals' specific needs. In this dynamic environment, professionals can engage in real-world scenarios and tasks, enhancing their adaptability and proficiency in response to the challenges presented by Industry 5.0. This study aims to bridge this gap by developing a digital platform for industrial training and personalized learning using immersive technologies, gamification techniques, and artificial intelligence. This platform enables the aggregation and use of relevant information to generate new digital content specifically designed to address the skills gap in professionals using technologies in the industry. By incorporating immersive training, adaptive algorithms, and gamification elements, effective, efficient, and personalized training can be achieved, leading to the continuous development of professionals, improvement of industrial operations, and adaptation to the ever-evolving demands of the sector.

2 Related Work

The training paradigm in Industry 4.0 is centered on technology; however, with the transition to Industry 5.0, the focus is on creating synergy between machines, robots, and humans [7]. Recognizing the inherent fragility of operators and the need for robust collaboration between humans and machines, a transformation in industrial training is urgent to prepare a new generation of 5.0 operators.

One of the most significant changes was the introduction of the Industrial Metaverse in professional training. This technology has revolutionized how professionals acquire knowledge and skills, offering highly interactive and engaging learning experiences [8]. Immersive applications, such as Virtual Reality (VR) and Augmented Reality (AR), allow professionals to simulate realistic work environments and practice complex tasks in a safe and controlled environment. Additionally, a VR review study conducted by Kumar et al. [9] reported that VR has completed a phase of technological adaptation and shows good potential as a training tool in a professional environment.

In this sense, several researchers have introduced metaverse applications in some industrial scenarios. For instance, Fracaro et al. [10] demonstrate that proper immersion can enhance professionals' training sessions, increasing motivation, and reducing learning time. This article incorporates design principles for a VR training environment, including features that enhance training effectiveness. Other studies have showcased the efficacy of VR in operator training, particularly concerning safety protocols [11], [12], [13]. These investigations leverage visual and graphical components within the virtual environment to aid users in comprehending assembly sequences, workspace configurations, and safety system procedures. Other applications focus on developing systems for production process operators. For example, [14] developed a VR training system to address the labor shortage in welding, resulting in faster acquisition of skills and better preparation for real work environments. [15] presented a training system combined with an assessment system for flexible assembly tasks. This study allows for adjustable learning speed tailored to the pace and dynamics of operators. Additional research uses extended reality (XR) technologies, which combine virtual and realistic interactions to offer an expanded environment for professional engagement [16], [17]. Maffei et al. [18] crafted an application guiding operators through assembly, resulting in heightened confidence and safety during task execution. Recent studies have revealed that introducing gamification into immersive experiences positively influences professionals' engagement, creativity, skill enhancement, and active participation [19], [20], [21].

Artificial intelligence (AI) advancements have opened new possibilities, especially in personalized and tailored education. Pratama et al. [22] explored how AI is revolutionizing education in the industry, concluding that AI holds significant potential for identifying each professional's learning styles, preferences, and unique talents. Thus, it facilitates the creation of personalized learning sessions, ensuring that professionals receive content and skills aligned with their specific needs. Building upon this, [23] and [24] developed intelligent AI-based learning systems providing real-time personalized assistance through guidance, feedback, and additional resources based on individual professional performance. Despite the promising opportunities that AI systems offer to personalize learning, automate routine tasks, and provide adaptive assessments, there is limited research on the impact of AI systems on skill acquisition.

3 Architecture of the Proposed System and Development Approach

This project involves the creation of a digital platform to provide industry professionals, both trainers, and trainees, with an innovative, flexible, and interactive means of personalized training. Figure 1 provides an overview of the proposed system's architecture, consisting mainly of four integrated modules: a platform for personalized training paths and information presentation, AI algorithms for performance analysis and optimization, immersive technology integration for virtual training, and advanced gamification techniques.

The first module develops a web application to streamline industry professionals' training. It grants access to pertinent data for crafting personalized training paths tailored to each professional's unique needs. The second module aims to develop AI algorithms for ongoing monitoring of professional performance, adjusting training difficulty dynamically, and providing personalized feedback. These algorithms will also assist in creating customized learning paths based on individual abilities and needs and incorporate non-playable characters (NPCs) to enhance training interactions. The third module innovatively integrates immersive technology into the training platform, allowing professionals to practice skills in VR and AR environments, fostering knowledge retention and practical skill development. Additionally, integrating gamification techniques will boost professionals' engagement and motivation, involving setting gamified objectives, designing game elements, incorporating challenges and missions, and establishing a reward system.

By combining these steps, it's possible to create an innovative and adaptable platform to tackle skill gaps among industry professionals and meet the ever-changing demands of the industrial sector. In addition to resolving present challenges, it fosters a dynamic and flexible learning environment that promotes professional excellence and ongoing advancement within the industry.

3.1 Platform for personalized training paths

The platform for personalized training paths revolutionizes the way industry professionals enhance their skills and knowledge. Designed as an online application, it offers an integrated training experience tailored to the unique needs of each professional. Utilizing cutting-edge AI algorithms, the platform analyzes relevant data to create personalized training paths that address specific competencies, objectives, and learning styles. With two distinct user roles, trainers and trainees, the platform satisfies the diverse needs

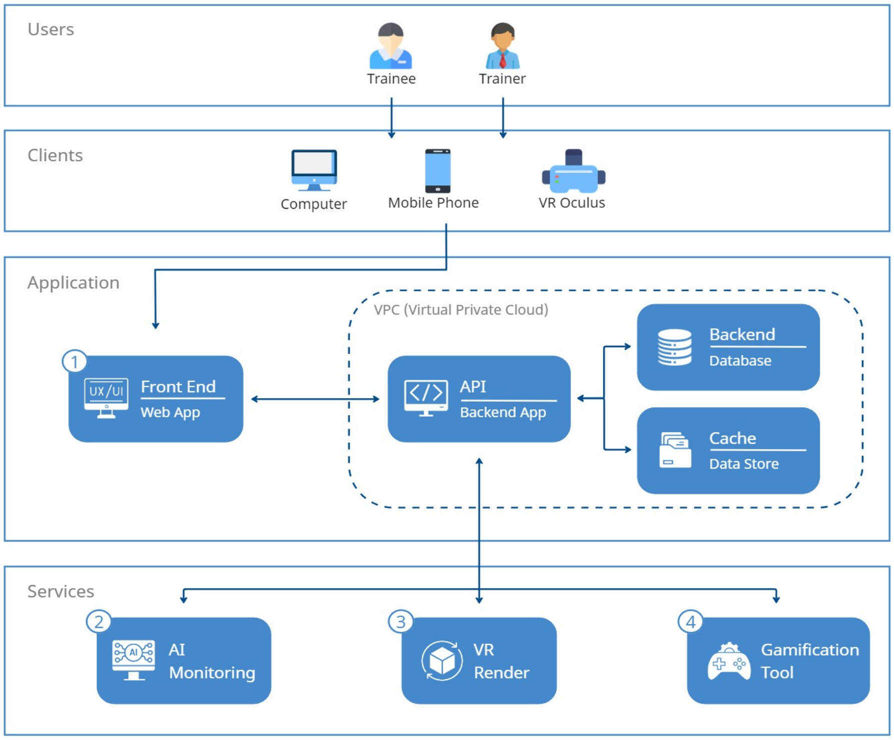


Fig. 1: Overview of our digital platform for industrial training and personalized learning.

of both educators and learners. Trainers have the power to create customized training experiences for VR Oculus devices, tailoring content, scenarios, and challenges based on the requirements of individual trainees.

The trainees benefit from a comprehensive range of features that allow them to track and analyze their training performance. Detailed insights into progress, completion status, and scores achieved across various modules empower trainees to monitor their development and identify areas for improvement. The platform allows access to training sessions through VR Oculus devices, enabling trainers to immerse themselves in the immersive environment. Real-time guidance and assessment enhance the interactive nature of training sessions, fostering a more engaging and effective learning environment.

To ensure that the platform has access to the necessary data to develop personalized training paths, it is essential that it can interact with information stored both in the database and in system files. Firstly, the platform will make requests to the database, which stores information about users. The platform will also need to access information stored in the file system, such as multimedia resources, training models, and other files necessary to customize the training experience.

3.2 AI algorithms for performance analysis and optimization

The development and integration of AI algorithms within the platform are pivotal for tailoring and adjusting training programs to suit the unique needs of industry professionals. These algorithms have various purposes, each aimed at enhancing the effectiveness and personalization of training programs. First, AI algorithms will play a crucial role in defining each professional's initial level of expertise. By analyzing their training history, competencies, and skill levels, the AI algorithm will discover a personalized starting point that optimizes the training process, ensuring that it aligns precisely with the individual's capabilities and requirements.

Moreover, the platform will utilize AI-driven feedback systems to monitor the performance of professionals during training sessions. These systems will dynamically adjust the difficulty level of exercises based on real-time performance metrics, ensuring that the learning experience continues appropriately and engaging for skill development. In addition, intelligent algorithms will help to create personalized learning paths adapted to each professional's skills and learning styles.

3.3 Immersive technology integration for virtual training

The integration of immersive technology into the learning platform is an essential step of the project aimed at providing professionals with a practical and highly applicable learning experience. The use of VR and AR provides the possibility to practice skills in a safe and controlled environment, promoting knowledge consolidation and the development of practical skills. The objective is to leverage these technologies to create immersive and personalized environments tailored to the training sessions outlined by the trainer on the platform. Initially, the instructional content and scenarios provided by the trainer will be analyzed and then converted into interactive experiences suitable for VR or AR.

The integration of Oculus VR with the platform enables the trainees to access virtual environments directly through VR devices. To enrich the learning experience and the sense of realism, trainees will be able to actively engage with objects and elements using their own hands, enabling them to perform specific actions and execute tasks in the virtual environment. Spatial audio will also be integrated, if enabled by the trainee, to provide an immersive sound experience, allowing users to locate the origin of sounds in the virtual environment.

3.4 Advanced gamification techniques

The training sessions will integrate advanced gamification techniques to boost engagement and motivation among professionals. These techniques are designed to create an immersive and motivating learning environment that encourages active participation and skill development. One such method involves implementing a points and levels system, where trainees earn points by completing tasks or achieving objectives. As they accumulate points, they progress to higher levels, unlocking new challenges and rewards along the way. Additionally, challenging missions will be designed to allow professionals to apply their skills in practical scenarios, with clearly defined goals and objectives providing a sense of purpose and achievement.

Virtual rewards, such as badges or trophies, will be offered to acknowledge significant achievements and milestones, displayed on individual profiles to showcase progress and accomplishments. By incorporating these advanced gamification techniques into training sessions, trainers can create dynamic and engaging learning environments that motivate trainees to actively participate in their skill development journey.

Conclusion

In conclusion, this project presents an innovative digital platform designed to provide innovative, flexible and interactive personalized training for industry professionals, covering both trainers and trainees. The proposed solution is based on four integrated modules. The first module focuses on the development of a web application that facilitates training for industry professionals. The second module involves the creation of AI algorithms for continuous monitoring and dynamic adjustment of training difficulty. The third module integrates immersive technologies into the training platform so that professionals practice skills in safe and controlled environments. The final module incorporates advanced gamification techniques, including defining gamified objectives, designing engaging game elements and incorporating challenges and missions. The study does not yet present results or conclusions due to the very early stage of the project's development. Further progress will therefore be needed to assess the effectiveness of the system and its impact on the learning of industry professionals.

Acknowledgement

This work was financed by the financed by national funds, through FCT - Foundation for Science and Technology and FCT / MCTES under the project UIDB/05549/2020, UIDP/05549/2020 and LASI-LA/P/0104/2020. This project was also funded by the FAIST - Fábrica Ágilente Inteligente Sustentável e Tecnológica, co-funded from the "Mobilizing Agendas for Business Innovation" of the "Next Generation EU" program of Component 5 of the Recovery and Resilience Plan (RRP), concerning "Capitalization and Business Innovation", under the Regulation of the Incentive System "Agendas for Business Innovation".

References

1. Attilio Di Battista, Sam Grayling, Else Hasselaar, T Leopold, Ricky Li, Mark Rayner, and Saadia Zahidi. Future of jobs report 2023. In *World Economic Forum, Geneva, Switzerland. https://www. weforum. org/reports/the-future-of-jobs-report-2023*, 2023.
2. Amr Adel. Future of industry 5.0 in society: human-centric solutions, challenges and prospective research areas. *Journal of Cloud Computing*, 11(1):40, 2022.
3. Didem Gürdür Broo, Okyay Kaynak, and Sadiq M Sait. Rethinking engineering education at the age of industry 5.0. *Journal of Industrial Information Integration*, 25:100311, 2022.
4. SI Wahidi, TW Pribadi, WS Rajasa, and MS Arif. Virtual reality based application for safety training at shipyards. iop conference series: Earth and environmental science, 972 (1), 2022.
5. Jessica Ulmer, Sebastian Braun, Chi-Tsun Cheng, Steve Dowey, and Jörg Wollert. Gamification of virtual reality assembly training: Effects of a combined point and level system on motivation and training results. *International Journal of Human-Computer Studies*, 165:102854, 2022.
6. Helen Crompton, Mildred V Jones, Yaser Sendi, Maram Aizaz, Katherina Nako, Ricardo Randall, and Eric Weisel. Examining technology use within the addie framework to develop professional training. *European Journal of Training and Development*, 2023.

7. Praveen Kumar Reddy Maddikunta, Quoc-Viet Pham, B Prabadevi, Natarajan Deepa, Kapal Dev, Thippa Reddy Gadekallu, Rukhsana Ruby, and Madhusanka Liyanage. Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration*, 26:100257, 2022.
8. Sofia Garcia Fracaro, Jarka Glassey, Kristel Bernaerts, and Michael Wilk. Immersive technologies for the training of operators in the process industry: A systematic literature review. *Computers & Chemical Engineering*, 160:107691, 2022.
9. Vinod Vijay Kumar, Deborah Carberry, Christian Beenfeldt, Martin Peter Andersson, Seyed Soheil Mansouri, and Fausto Gallucci. Virtual reality in chemical and biochemical engineering education and training. *Education for Chemical Engineers*, 36:143–153, 2021.
10. Sofia Garcia Fracaro, Philippe Chan, Timothy Gallagher, Yusra Tehreem, Ryo Toyoda, Kristel Bernaerts, Jarka Glassey, Thies Pfeiffer, Bert Slof, Sven Wachsmuth, et al. Towards design guidelines for virtual reality training for the chemical industry. *Education for Chemical Engineers*, 36:12–23, 2021.
11. Morteza Dianatfar, Jyrki Latokartano, and Minna Lanz. Concept for virtual safety training system for human-robot collaboration. *Procedia Manufacturing*, 51:54–60, 2020.
12. Saeed Rokooei, Alireza Shojaei, Amin Alvanchi, Reza Azad, and Nasim Didehvar. Virtual reality application for construction safety training. *Safety science*, 157:105925, 2023.
13. SI Wahidi, TW Pribadi, WS Rajasa, and MS Arif. Virtual reality based application for safety training at shipyards. In *IOP Conference Series: Earth and Environmental Science*, volume 972, page 012025. IOP Publishing, 2022.
14. Ananya Ipsita, Levi Erickson, Yangzi Dong, Joey Huang, Alexa K Bushinski, Sraven Saradhi, Ana M Villanueva, Kylie A Peppler, Thomas S Redick, and Karthik Ramani. Towards modeling of virtual reality welding simulators to promote accessible and scalable training. In *Proceedings of the 2022 CHI conference on human factors in computing systems*, pages 1–21, 2022.
15. Kwang-Phil Park, Seung-Ho Ham, Woo-Yong Lee, and Byeong-Woo Yoo. Development of transporter training simulator based on virtual reality and vehicle dynamics model. *International Journal of Naval Architecture and Ocean Engineering*, 15:100547, 2023.
16. Sebastian Büttner, Michael Prilla, and Carsten Röcker. Augmented reality training for industrial assembly work-are projection-based ar assistive systems an appropriate tool for assembly training? In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pages 1–12, 2020.
17. Luke Beveridge, Keshav Dahal, and Marco Gilardi. Swagexr: Designing gamified ar applications for industrial training. In *European Conference on Games Based Learning*, volume 17, pages 765–773, 2023.
18. Antonio Maffei, Michela Dalle Mura, Fabio Marco Monetti, and Eleonora Boffa. Dynamic mixed reality assembly guidance using optical recognition methods. *Applied Sciences*, 13(3):1760, 2023.
19. Zhao Zhiqiang, Toh Da Jun, Ding Xiaoming, Ng Keng Chong, Sin See Choon, and Wong Yuen Choe. Gamification platform for manufacturing shopfloor training-a case study. In *The 15th International CDIO Conference*, 2019.
20. Aggelos Marinakis, Anastasia Katsamaki, Ioanna Pateraki, Evangelos Nikolidakis, and Aristomenis Antoniadis. A novel serious game for education and training of computer-aided manufacturing (cam) programming. In *2021 30th Annual Conference of the European Association for Education in Electrical and Information Engineering (EAEEIE)*, pages 1–6. IEEE, 2021.
21. Thomas Bohné, Ina Heine, Felix Mueller, Paul-David Joshua Zuercher, and Vera Maria Eger. Gamification intensity in web-based virtual training environments and its effect on learning. *IEEE Transactions on Learning Technologies*, 2022.
22. Muh Putra Pratama, Rigel Sampelolo, and Hans Lura. Revolutionizing education: harnessing the power of artificial intelligence for personalized learning. *Klasikal: Journal of Education, Language Teaching and Science*, 5(2):350–357, 2023.
23. Paul Rad, Mehdi Roopaei, Nicole Beebe, Mehdi Shadaram, and Yoris Au. Ai thinking for cloud education platform with personalized learning. 2018.
24. Zhisheng Chen. Artificial intelligence-virtual trainer: Innovative didactics aimed at personalized training needs. *Journal of the Knowledge Economy*, 14(2):2007–2025, 2023.

Intelligent Control and Monitoring of Indoor Microgreens Production System

Deborah Ishikawa da Silva[1], Lucas Santos Pinow[1], Cláudio Leones Bazzi[3], José Barbosa[1,2], and Paulo Leitão[1,2]

[1] Research Center in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
a58875@alunos.ipb.pt, {lucaspinow, pleitao, jbarbosa}@ipb.pt
[2] Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
[3] Universidade Tecnológica Federal do Paraná, Campus Medianeira, 85.884-000 Medianeira, Brasil
bazzi@utfpr.edu.br

Abstract. Due to the population growth in urban areas, the integration of urban agriculture has become increasingly common. Typically, microgreens cultures are known for their small size and faster growth cycle, which are suitable for the structures of these locations. However, conventional agriculture faces instabilities throughout the cultivation process, losses, and harvest failures due to environmental variations that directly affect its development. Urban agriculture's suffer even more from these environmental issues because they are embedded in environments without adequate infrastructure and climate adaptation. In this sense, controlled environments emerge as an alternative to maintain crop stability regardless of external conditions by simulating the necessary environmental conditions its expectations. Having this in mind, this paper describes the use of Internet of Things (IoT) technologies to assist conventional and urban agriculture with predictable, controllable and optimized processes. In the studied context, the implementation of these technologies also allows the adaptation of more dynamic, sustainable and efficient greenhouses in the field of agriculture, establishing parameters desired by the producer. For this work, it was possible to verify the effectiveness of an approach for controlling the internal microclimate of the greenhouse, using hysteresis and operating ranges of controlled variables in conjunction with low-cost sensors and actuators. Additionally, the practicality of a remote control and monitoring system using IoT technologies was demonstrated, enabling monitoring, forecasting, and data analysis.

Keywords: Microgreens · Environmental Indoor · Control · Monitoring · IoT.

1 Introduction

Environmental conditions directly affect the conventional agriculture and compromise the development of the productive cycle of its vegetation. Instability in these rural areas promotes damage to the crop development, crop failures and even income losses. Urban agriculture, in turn, has become increasingly common with the increase in demand for products in large centers. Generally, greenhouse cultivation accompanies microgreens cultivation due to their small size, rapid growth cycle and ease of cultivation to adapt to these environments. However, as this practice occurs in small spaces and in environments where external conditions can vary considerably, significant and efficient management is necessary to minimize negative effects. The integration of technologies in greenhouses allows you to optimize cultivation, provide the necessary resources efficiently and reduce damage caused by unstable environmental events. In this sense, the development of a controlled environment makes it possible to emulate the environmental conditions required according to the needs of each culture. The implementation of the Internet of Things (IoT) system involves circuits that are interconnected and monitored in real time to verify their proper functioning. Therefore, actuators were considered to control irrigation related to soil humidity, an atomizer to induce air humidity, fans and heaters that control the internal temperature of the environment and light-emitting diodes (LEDs) for artificial lighting. The control and monitoring of these variables makes it possible to establish limits according to the producer's expectations, since the agreement of practices establishes predictable parameters. Crops grown under these conditions present more economical and sustainable practices as well as reducing waste during all processes.

This paper is structured as follows: section 2 describes the research to be understood, the foundations and related work and section 3 the research proposal in accordance with the objective of this work. Finally, section 4 with the conclusion and notes on future work.

2 Related Works

The integration of the IoT across different industries has become one of the most widely adopted innovations due to its ability to promote connectivity and data exchange. Specifically, in the case of automated greenhouses, IoT has positioned itself as a technological innovation with high potential for optimizing agricultural production, including microgreens. Haris et al. [1] provide insight into IoT-based indoor agricultural automated systems under three scopes of operation: cloud, fog/edge and sensor/actuator. Soni et al. [2] emphasize the use of the Message Queuing Telemetry Transport (MQTT) protocol for IoT applications due to its efficient data transmission. Wardihani et al. [3] developed a greenhouse control and monitoring system with IoT, focusing on controlling air temperature and humidity. Similarly, Raj et al. [4] demonstrated the practical benefits of greenhouse automation through the implementation of a Netduino 3-based system automating tasks such as irrigation and ventilation, notable improvements in crop yields and in production efficiency. On the other hand, Stanghellini et al. [5] took a different approach, analyzing plant transpiration instead of directly monitoring temperature, humidity and solar intensity. Although positive results were achieved, it was concluded that an environment of precise control of environmental variables is necessary, otherwise there will be numerous issues. Teng et al. [6] carried out a summary of recent studies on multiple aspects of microgreens, which are young and very attractive plants. In this study, the authors emphasized technical challenges for widespread production and data gaps for growing microgreens and highlighted the use of IoT technologies for the development of the emerging microgreen industry.

3 Research Proposal

3.1 System Proposal

This paper addresses the development of an IoT-based automated modular greenhouse system designed for growing microgreens and a crop monitoring system capable of predicting when it will be ready for harvest. Based on existing literature, automation technologies and IoT applications, this system focuses on controlling and monitoring greenhouse temperature, humidity, soil moisture and illuminance. The monitoring interface is based on a node-red application, due to its intuitive implementation, flexibility, scalability, visualization capacity and data analysis. As a communication protocol, MQTT was adopted due to its data transmission efficiency, compatibility with IoT devices and the publish/subscribe model that simplifies data management and distribution. The edge device adopted was the ESP32 microcontroller due to its hardware versatility, Wi-Fi connectivity, MQTT compatibility, low cost and low power consumption.

The prediction model is based on thermal accumulation of the crop, but there are not enough studies that show exactly the time needed for microgreens to grow, only for conventional plants.

3.2 System Architecture

The system architecture is exemplified by Fig. 1, where it is possible to verify the data flow. This system is based on the acquisition of data from environmental variables through sensors connected to the microcontroller. After that, it creates a file based on JavaScript Object Notation (JSON) containing information about the analyzed variables and the current timestamp. Using the MQTT protocol, the JSON file is published to a specific topic on a cloud broker. The Node-RED application subscribes to this topic, performs the analysis and temporal graphical printing of all variables data, calculates the thermal accumulation of the crop and estimates the time until harvested.

All greenhouse control is based on pre-defined recipes for each specific type of microgreens, selected at the beginning of each cultivation in the Node-RED application. These recipes are setpoints for temperature, humidity, soil humidity and time of exposure to light, for the day and another for the night. The Node-RED application generates a JSON file and publishes it to a specific topic using MQTT. The microcontroller then receives this file through the topic subscription and makes the internal setpoint definitions according to the recipe.

Based on the defined setpoints and the developed control algorithm, the actuators are controlled to ensure compliance with the recipe. The actuators used in the greenhouse are fans and PTC heater for temperature control, water pump for soil irrigation, atomizer for water vaporization, contributing to increased air humidity and LED strips for illumination.

The system also includes a database based on MongoDB technology due to its flexible data model, performance and integration with the Node-RED application. In this database, the temporal values of all system variables are saved, allowing a possible more detailed analysis of the entire crop cultivation period.

3.3 System Prototype

Fig. 2 illustrates the automated greenhouse prototype developed to demonstrate the viability of the system. This greenhouse prototype consists of a small modular structure made with acrylic, equipped with sensors, actuators and IoT technology with the aim of creating an ideal environment for the growth of microgreens.

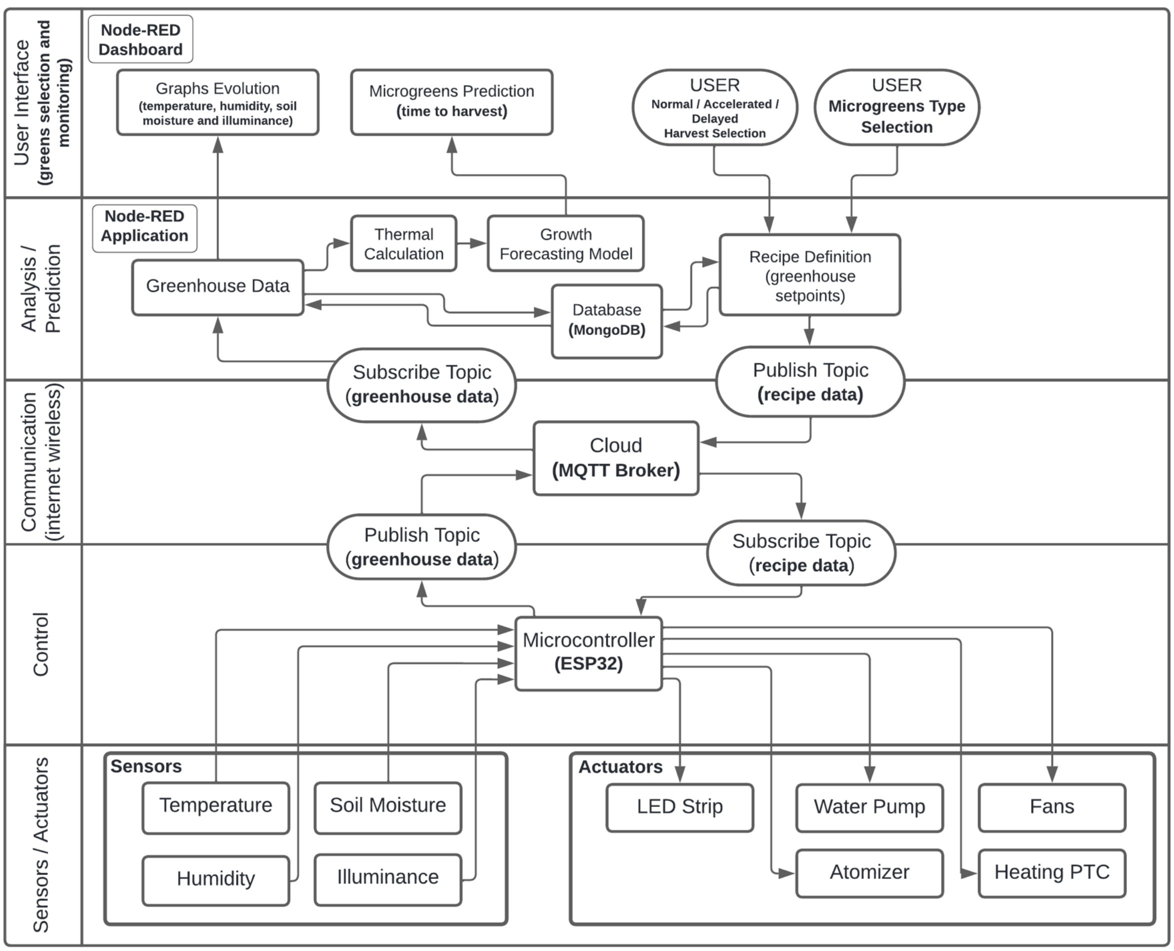


Fig. 1: System Architecture.

Regarding the module dimensions, it measures 28 cm in width, 56 cm in length, and 21 cm in height, excluding the base. Including the base, the system's dimensions are 60 cm in width, 60 cm in length, and 33 cm in height.

The prototype was built following the system architecture specifications described in subsection 3.2. Stress tests were carried out to evaluate the system's responsiveness to different recipes, to maintain adequate control of the environment.

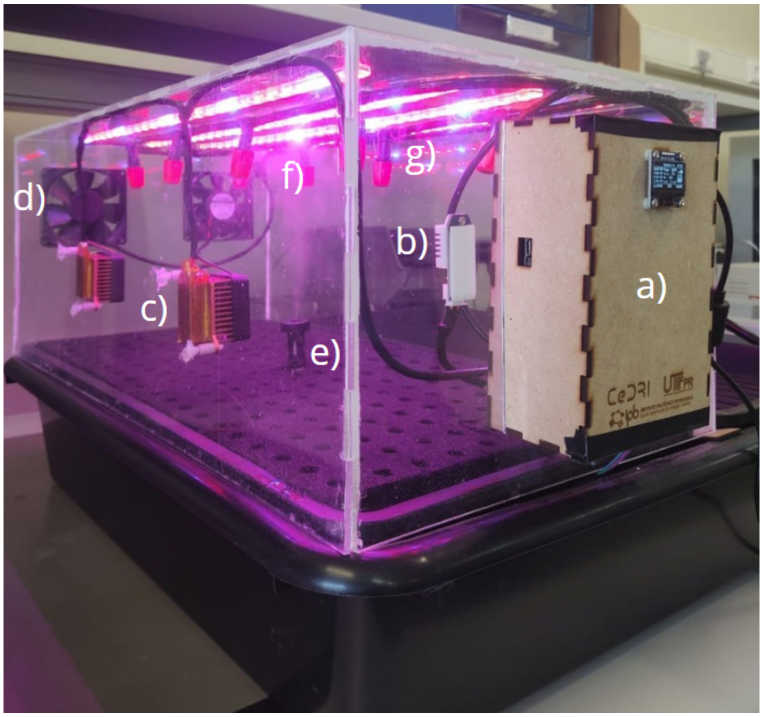


Fig. 2: System Prototype. a) Circuit box with display interface b) Sensor Temperature and Humidity c) Heaters d) Fans e) Atomizer f) Irrigation Circuit g) LED Strip

Similarly, Fig. 3 shows the monitoring system based on the node-red application. It is possible to check the current status and evolution of system variables. In addition to recipe selection, as indicated by the system architecture in subsection 3.2.

Based on the recipes defined by the crop monitoring system, the control system operates the actuators to maintain the environmental variables within an appropriate operating range, utilizing hysteresis for each variable. When the temperature is below the setpoint, the heater is turned on until it exceeds its hysteresis threshold, after which it is turned off. Ventilation is activated when the system's temperature surpasses this range and is turned off when it reaches the setpoint temperature. The approach for humidity is slightly different, the defined setpoint is considered a minimum value for both air and soil humidity. When the air humidity falls slightly below the setpoint, the atomizer is activated and deactivated once it reaches its operating range. For irrigation, a timed activation is used due to the high dead time between initiating irrigation and the system sensing a change in soil moisture. Thus, when the soil moisture is below the ideal level, the water pump is activated for a few seconds, followed by a waiting period, and the soil moisture is checked again. If it remains below the required level, re-activation occurs until the soil moisture is within acceptable levels.

Fig. 3: Node-RED Dashboard.

4 Conclusions

Vegetative growth in agriculture has been limited to seasons and oriented according to local climatic conditions. The implementation of IoT systems creates a controlled environment with graphical and virtual monitoring that allows the producer to monitor all variables and identify failures to apply immediate adjustments that optimize conditions during production. Internal system integrations control microgreen production and can be managed remotely with all the details about the cultivation. Both scenarios present the same environmental characteristics and functionalities with factors necessary for the development of microgreens. However, the implementation of a controlled environment in greenhouses presents the inexhaustible availability of resources according to the needs of the crop when compared to conventional agriculture. In this sense, microgreens developed in greenhouses result in high-quality, more nutritious and aesthetically more attractive products. Furthermore, the agreement of these factors establishes more sustainable, economical and efficient practices in the field of agriculture.

In the stress tests of the prototype, the effectiveness of the control system in maintaining environmental conditions close to those defined by the setpoint of the variables in the recipes was verified. In the monitoring system, it was possible to verify and analyze the system's behavior and reaction to the control action. To date, no microgreen production tests have been carried out, only empty tests to verify the functioning and behavior of the system. Finally, we identified that some improvements can be implemented to improve the ergonomics and practicality of the greenhouse. Initially, adjustments were made to the water reservoir to maintain adequate pressure, pumping again in a closed circuit to avoid waste and implementing level sensors at different levels for monitoring. It would be interesting to implement other sensors to improve internal control, such as CO_2 sensors, photosynthesis verification, identification of the presence of pests, verification of cultivar nutrition, among others. Furthermore, despite establishing local and remote access through different devices, it would also be interesting to develop a personalized application for restricted and exclusive access to the greenhouse.

5 Acknowledgments

This work was supported by the Foundation for Science and Technology (FCT, Portugal) through national funds FCT/MCTES (PIDDAC) to CeDRI (UIDB/05757/2020 and UIDP/05757/2020) and SusTEC (LA/P/0007/2021).

References

1. Isakovic Haris, Alexander Fasching, Lukas Punzenberger, and Radu Grosu. CPS/IoT ecosystem: Indoor vertical farming system. In *2019 IEEE 23rd International Symposium on Consumer Technologies (ISCT)*, pages 47–52. IEEE, 2019.
2. Dipa Soni and Ashwin Makwana. A survey on MQTT: a protocol of internet of things (IoT). In *International conference on telecommunication, power analysis and computing techniques (ICTPACT-2017)*, volume 20, pages 173–177, 2017.
3. Eni Dwi Wardihani, Eka Ulia Sari, Ari Sriyanto Nugroho Helmy, Yusnan Badruzzaman, Arif Nursyahid, Thomas Agung Setyawan, et al. Monitoring and controlling of IoT-based greenhouse parameters with the MQTT protocol. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 13(1), 2024.
4. Jennifer S Raj and J Vijitha Ananthi. Automation using IoT in greenhouse environment. *Journal of Information Technology*, 1(01):38–47, 2019.
5. C. Stanghellini and W.Th.m. van Meurs. Environmental control of greenhouse crop transpiration. *Journal of Agricultural Engineering Research*, 51:297–311, 1992.
6. Zi Teng, Yaguang Luo, Daniel J Pearlstein, Raymond M Wheeler, Christina M Johnson, Qin Wang, and Jorge M Fonseca. Microgreens for home, commercial, and space farming: a comprehensive update of the most recent developments. *Annual Review of Food Science and Technology*, 14:539–562, 2023.

Autonomous Navigation of Unmanned Aerial Vehicle in Software in the Loop

Gabriel O. Pimentel[1,2], José Lima[1], and Murillo F. dos Santos[2]

[1] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal
jllima@ipb.pt
[2] Federal Center for Technological Education of Minas Gerais, Brazil
gabriel.pimentel@aluno.cefetmg.br, murillo.ferreira@cefetmg.br

Abstract. Unmanned Aerial Vehicles (UAVs) have expanded with the development of innovative computing technologies in areas such as power line inspection, environmental monitoring, and precision agriculture. The latter, highlighted in this study, uses UAVs equipped with advanced sensors and cameras to provide detailed information on crop conditions and detect pests and diseases. However, these applications face challenges such as efficient battery management and adverse weather conditions. To overcome these obstacles, simulation in virtual environments is proving to be an effective approach for efficient flight planning. Through simulation in Software In The Loop (SITL), this study presents an algorithm developed for the autonomous navigation of a quadcopter in a vineyard field, allowing its effectiveness and accuracy to be analyzed and validated. The results demonstrate the quadcopter's ability to navigate autonomously, following the waypoints defined during the flight. These results are promising and highlight the potential of UAVs in precision agriculture, helping to increase the productivity and sustainability of agricultural activities.

Keywords: Software In The Loop · Autonomous navigation · Precision agriculture

1 Introduction

Unmanned Aerial Vehicles (UAVs) are characterized by a lack of human presence on board. The use of UAVs is currently expanding, driven by the emergence of innovative computer technologies, such as artificial intelligence and computer vision, mainly in the areas of reconnaissance and terrestrial mapping [7].

Previously, UAV applications were primarily for military purposes, such as geographic reconnaissance and unmanned inspection. However, UAVs have expanded into other applications such as aerial photography [3], power line inspection [1, 6], pollution monitoring [2], forest fire monitoring [9], rescue operations [12], disaster management [10], healthcare [5], and agricultural applications such as field inspection [4] and detection of crop stress and diseases [14].

In the field of agriculture, one well-known area is precision agriculture. Precision farming has emerged as a promising solution for improving agricultural productivity and reducing waste. Precision farming uses advanced technologies such as the Global Positioning System (GPS), sensors, and data analysis to enable farmers to make informed decisions about crop management and effectively monitor plantations [13].

Using UAVs equipped with a variety of sensors and cameras, detailed images can be captured that provide a variety of useful information over a large area, such as crop condition, soil analysis, vegetation analysis, pest and drought detection, and plant health monitoring. Thus, UAVs can deliver results quickly and accurately, reducing the need for manual monitoring and analysis [11].

With the expansion of precision agriculture, the use of UAVs is increasing due to factors such as their relatively low cost, availability in the market, and ease of use.

To ensure that flights are performed correctly, identify and correct possible flaws in control, navigation, and obstacle detection systems, and develop efficient and robust techniques for UAVs in these scenarios, simulations in virtual environments are an excellent option. Simulations provide a safe environment and are an effective way to address these issues.

In this context, this work presents a simulation of the autonomous navigation of a quadcopter using a developed algorithm to monitor operations in a vineyard field. The structure of this work is as follows: Section 1 covers the introduction described here; Section 2 describes the methodology used; Section 3 discusses the analysis of the results. Finally, Section 4 presents the conclusions drawn from this paper.

2 Methodology

The proposed methodology is detailed in this section. For this work, simulations were conducted using ArduPilot, an open-source platform that allows the development and deployment of unmanned autonomous systems, such as drones and ground vehicles.

The simulation was implemented using Software in the Loop (SITL), a methodology used to develop algorithms or even control strategies. It was validated using a simulation environment with vehicles such as fixed-wing aircraft, multi-rotor aircraft, and ground and underwater vehicles. Simulations are performed using a Flight Dynamics Model of the vehicle, which mimics the physics of how the vehicle moves. Inputs from SITL, which runs the ArduPilot firmware, control the behavior of the vehicle, and outputs such as vehicle status, position, and speed are fed back to the firmware simulation, as shown in Figure 1.

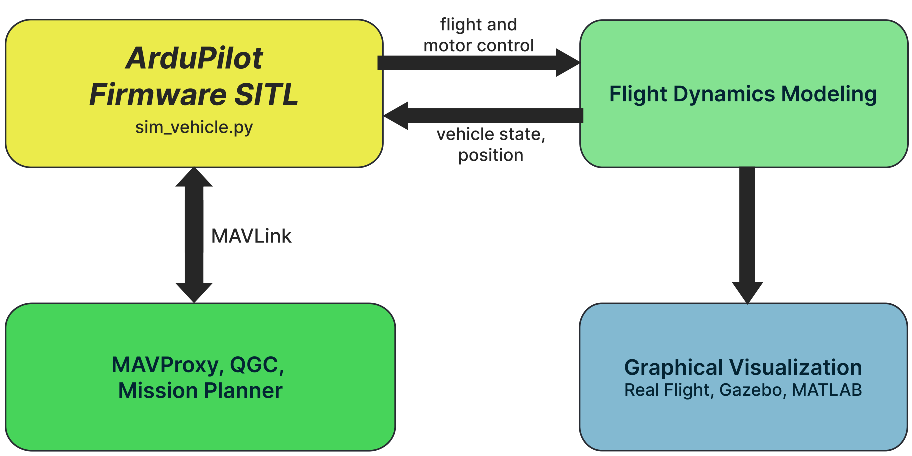


Fig. 1: SITL architecture

The exchange of information between these environments is done using the MAVlink protocol messages. MAVLink is a very lightweight messaging protocol for communication. This protocol follows a modern hybrid publish-subscribe and point-to-point design pattern, where data flows are sent, and published as topics, while the configuration subprotocols, such as the mission or parameter protocol, are point-to-point with retransmission [8].

This allows developed algorithms to use SITL. In this study, as a proof of concept, a powerful robot simulator, Gazebo, was used, which offers the ability to efficiently simulate robots in complex environments with real environmental variables, such as weather conditions. And as a simulation vehicle, using the model vehicle, the IRIS quadcopter described in Figure 2, was chosen. An algorithm was developed in Python to perform functions using the `dronekit` and `pymavlink` libraries to retrieve information from the vehicle and send commands via the MAVLink protocol.

Fig. 2: Quadcopter used in Gazebo

The exchange of information between these environments is done using MAVLink protocol messages. MAVLink is a very lightweight messaging protocol for communication. This protocol follows a modern hybrid publish-subscribe and point-to-point design pattern, where data flows are sent and published as topics, while the configuration subprotocols, such as the mission or parameter protocol, are point-to-point with retransmission [8].

Using Mission Planner, a tool for creating routes with waypoints (WPs) by "point and click" on maps such as Google Maps and OpenStreetMap, a mission with 38 waypoints was created for this case study to cover a field of vines located at the Polytechnic Institute of Bragança, as shown in Figure 3.

Fig. 3: Mission with waypoints

Once the mission has been defined, Mission Planner allows the mission to be downloaded in `txt` format with the real coordinates of the waypoints (WPs). From this file, the developed algorithm reads and interprets the coordinates, performs the necessary processing, and stores the coordinates in the format expected by MAVLink.

Each function of the algorithm performs a critical step in the mission. Through sophisticated functions, commands are sent to the quadcopter via MAVLink. In this case study, the adopted procedures to autonomously navigate the vehicle through the waypoints were to perform the functions of arming the engines and taking off to an altitude of 2 meters.

The quadcopter system in ArduPilot has 25 built-in flight modes, 10 of which are used regularly. These modes support different levels/types of flight stabilization and a sophisticated autopilot. Since the UAV was already set up and the waypoints loaded, it was decided to use the AUTO mode.

In AUTO mode, the quadcopter follows a pre-programmed mission script stored in the autopilot, consisting of navigation commands and "do" commands. This mode includes altitude and position control. Since the focus of this work was not on control, the original settings were kept.

After completing the mission, the quadcopter uses Return to Launch (RTL) mode to navigate from its current position to the home position, ensuring that the quadcopter returns after the mission completes. Without a final RTL command, the vehicle will simply stop at the final waypoint.

Finally, for this simulation, it was assumed that during the scanning of the plantation, the quadcopter detected an infestation on one of the crops, in this case represented by Waypoint 6. Therefore, the developed algorithm established a routine for the quadcopter to return to this waypoint after completing the navigation of the vine plantation and perform a scan to ensure a more accurate image to draw better conclusions about the possible infestation.

3 Results and Discussion

This section presents the results obtained from the execution of the planned mission and the implementation of the algorithm developed for the autonomous navigation of the quadcopter.

To validate the established algorithm, a flight simulation was conducted using the generated waypoints, with a return to Waypoint 6 to hypothetically detect an infestation on one of the crops, as described in the previous section. Figure 4 shows the graph of the path taken by the quadcopter.

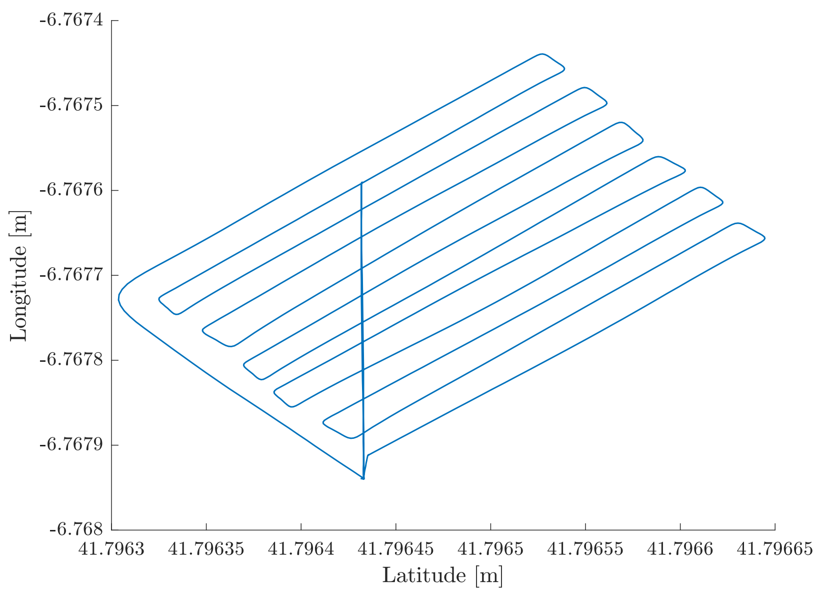


Fig. 4: Quadcopter path

This visualization provides a detailed analysis of the vehicle's navigation relative to the planned waypoints. For enhanced visualization, Figure 5 adds an extra dimension to display the path taken in the 3D plane.

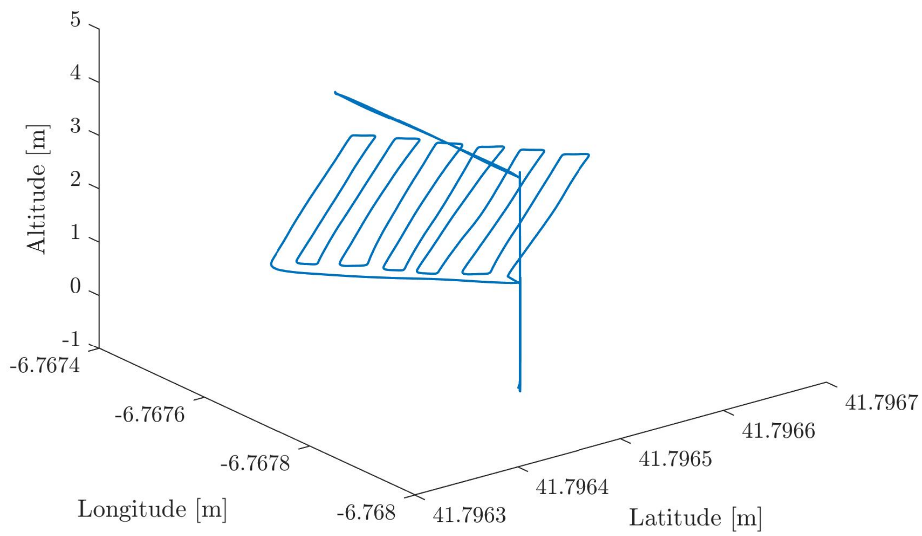


Fig. 5: Quadcopter 3D path

This allows observation of the vehicle's position in the plane alongside the variation in altitude over time. Such visualization is crucial for analyzing the vehicle's ability to maintain the desired altitude during the mission and for identifying potential variations in altitude control errors.

From the presented figures, it is evident that the quadcopter performed its tasks accurately and as expected. The trajectory followed by the vehicle consistently matched the planned waypoints, demonstrating the effectiveness of the developed algorithm for autonomous UAV navigation. Moreover, the variation in altitude throughout the flight remained within the desired parameters, indicating stable and consistent control throughout the mission.

However, it is important to note that real environmental variables, if introduced into the simulator, can impact these results. Although this work primarily focuses on simulating the algorithm for autonomous navigation, real-world factors such as weather can significantly influence UAV performance. Therefore, these real environmental variables are not negligible; they play a critical role in determining the reliability and robustness of UAV operations in practical scenarios.

4 Conclusion

The results obtained in this study demonstrate the viability and effectiveness of the algorithm developed for the autonomous navigation of quadcopters in complex environments such as vineyard plantations. Through the conducted simulation, the behavior of the vehicle during the execution of the planned mission was validated, including the hypothetical detection of an infestation in one of the crops.

The results indicate that the algorithm, as described above with its original control configurations, successfully guides the UAV accurately and consistently. It follows the defined waypoints and maintains a stable altitude throughout the flight. The functions developed contribute significantly to the efficiency and reliability of the operation, ensuring consistent results for future real-world applications.

Several future works can build upon the achievements of this study. Experimental tests are anticipated to validate the simulation results presented, incorporating sensors and cameras for infestation detection and monitoring. Furthermore, to assess system performance under varying conditions, different flight paths and real environmental variables should be introduced. Lastly, future work may explore implementing new control methodologies and control allocation strategies.

Acknowledgements The authors are grateful to the Foundation for Science and Technology (FCT, Portugal) for financial support through national funds FCT/MCTES (PIDDAC) to CeDRI, UIDB/05757/2020
(DOI: 10.54499/UIDB/05757/2020) and UIDP/05757/2020
(DOI: 10.54499/UIDP/05757/2020) and SusTEC, LA/P/0007/2020
(DOI: 10.54499/LA/P/0007/2020). The authors also thank the National Council for Scientific and Technological Development – CNPq, related to project 442696/2023-0.

References

1. Berger, G.S., Júnior, A.O., Braun, J., Lima, J., Pinto, M.F., Valente, A., Pereira, A.I., Cantieri, Á.R., Wehrmeister, M.A.: Sensorial testbed for high-voltage tower inspection with uavs. In: Tardioli, D., Matellán, V., Heredia, G., Silva, M.F., Marques, L. (eds.) ROBOT2022: Fifth Iberian Robotics Conference. pp. 353–364. Springer International Publishing, Cham (2023)
2. Bolla, G.M., Casagrande, M., Comazzetto, A., Dal Moro, R., Destro, M., Fantin, E., Colombatti, G., Aboudan, A., Lorenzini, E.C.: Aria: Air pollutants monitoring using uavs. In: 2018 5th IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace). pp. 225–229 (2018). https://doi.org/10.1109/MetroAeroSpace.2018.8453584
3. Evers, R., Masters, P.: The application of low-altitude near-infrared aerial photography for detecting clandestine burials using a uav and low-cost unmodified digital camera. Forensic science international **289**, 408–418 (2018)
4. Freeman, P.K., Freeland, R.S.: Agricultural uavs in the us: potential, policy, and hype. Remote Sensing Applications: Society and Environment **2**, 35–43 (2015)
5. Hiebert, B., Nouvet, E., Jeyabalan, V., Donelle, L.: The application of drones in healthcare and health-related services in north america: A scoping review. Drones **4**(3), 30 (2020)
6. Leal Lopes, V.M., Honório, L.M., Santos, M.F., Pancoti, A.A.N., Silva, M.F., Diniz, L.F., Mercorelli, P.: Design of an over-actuated hexacopter tilt-rotor for landing and coupling in power transmission lines. vol. 7 (2023)
7. Leme, F.P.: Sistema embarcado para controle de plataforma de dois eixos utilizada em veiculos aereos nao tripulados. UNICAMP (2017)
8. MAVLink: `https://mavlink.io/en/`
9. Nihei, K., Kai, N., Maruyama, Y., Yamashita, T., Kanetomo, D., Kitahara, T., Maruyama, M., Ohki, T., Kusin, K., Segah, H.: Forest fire surveillance using live video streaming from uav via multiple lte networks. In: 2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC). pp. 465–468 (2022). https://doi.org/10.1109/CCNC49033.2022.9700621
10. Noguchi, T., Komiya, Y.: Persistent cooperative monitoring system of disaster areas using uav networks. In: 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). pp. 1595–1600 (2019). https://doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00285

11. Phang, S.K., Chiang, T.H.A., Happonen, A., Chang, M.M.L.: From satellite to uav-based remote sensing: A review on precision agriculture. IEEE Access **11**, 127057–127076 (2023). https://doi.org/10.1109/ACCESS.2023.3330886
12. Rabta, B., Wankmüller, C., Reiner, G.: A drone fleet model for last-mile distribution in disaster relief operations. International Journal of Disaster Risk Reduction **28**, 107–112 (2018)
13. Radoglou-Grammatikis, P., Sarigiannidis, P., Lagkas, T., Moscholios, I.: A compilation of uav applications for precision agriculture. Computer Networks **172**, 107148 (2020). https://doi.org/10.1016/j.comnet.2020.107148
14. Yang, K.H., Lee, M.H., Yoe, H.: Designing an open field precision agriculture system using drones. In: 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE). pp. 1691–1694 (2023). https://doi.org/10.1109/CSCE60160.2023.00277

Decision Support Systems within Digital Twin: Bibliometric Study

Flávia Pires[1], António Paulo Moreira[3], and Paulo Leitão[1,2]

[1] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
fpires@ipb.pt, pleitao@ipb.pt
[2] Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
[3] INESC TEC - INESC Technology and Science, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, Porto, 4200-465, Portugal
amoreira@fe.up.pt

Abstract. The traditional decision-making process is facing a rising challenge due to the growing amounts of data generated by the digitalisation of manufacturing systems. The Digital Twin (DT) technology supports digitalisation and decision support and is a critical tool for decision support. This paper presents an introductory bibliometric study to assess the dynamic between DT technology and decision support systems (DSS) within the manufacturing domain, with the intent to understand how the research field is evolving, what are the main technologies being used and the main application domains. The study identified a growing interest in combining these topics among the scientific community.

Keywords: Digital Twin · Decision Support Systems · Industry 4.0. · Bibliometric Study

1 Introduction

Germany introduced Industry 4.0 in 2011 to stimulate the economy and transform the manufacturing industry [6], [9]. This was achieved through the digitalisation of systems using technologies like the Internet-of-Things (IoT) and Big Data, which generated vast amounts of data that posed a challenge to traditional decision-making processes [1], [7], [17], [11]. The analysis of large amounts of data can be very time-consuming, and many of the problems that require decision support have real-time constraints. Hence, there is a need for intelligent decision support systems (DSS) capable of handling large amounts of data in real-time [11]. The scientific community has explored using Digital Twin (DT) technology as a decision support platform as an effective solution. The DT is a virtual replica of a physical asset, process or entity capable of performing monitoring, data analysis, simulation and decision-support [13].

Traditional manufacturing systems rely on manual decision-making, which is often limited by decision-makers' experience and empirical knowledge [8]. This can be challenging for new decision-makers or when the manufacturing system requires flexibility and faster decision-making [4]. The DT supports digital transformation and enables decision support; once this is fully integrated with the manufacturing system, it is a central tool for decision support [8], [16].

This paper presents an introductory bibliometric study to assess how decision support evolves within the DT, determining the research field's evolution, how the research is being divided geographically, what are the main and more important contributions until now, the main technologies being used, and the application domains. This work was carried out as part of my PhD, which includes the development of a recommendation approach to integrate into a DT architecture as a DSS.

The paper is organised as follows: Section 2 presents the proposed study, illustrating the followed methodology and the main results from the conducted study. Lastly, in Section 3, are presented the main conclusion and future work.

2 Proposed Study

Bibliometric analysis is a technique usually used to explore vast scientific data and identify emerging areas or technologies [2].

2.1 Methodology

In the context of the development of a bibliometric study on decision support and DT, a well-established methodology was defined (based on [12]) and followed, which is illustrated in Fig. 1.

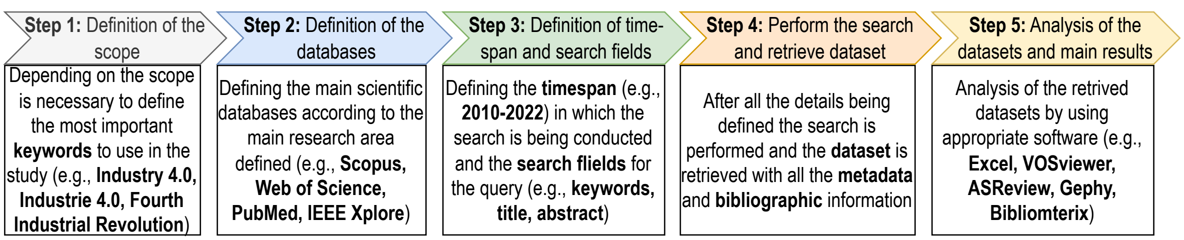


Fig. 1: General bibliometric methodology (based on [12]).

In this case, the methodology is divided into five steps: **Step 1**, the scope definition, DT and DSS; **Step 2**, the definition of the databases is conducted; in this study, it was considered the Scopus database, one of the most comprehensive scientific databases with over 90 million records; **Step 3** refers to the definition of the timespan (2010 to 2023) and the search fields to build the query (title, abstract and keywords); followed by **Step 4**, where the dataset was retrieved; and **Step 5** in which it was conducted the analysis using the VOSViewer, since it allows the generation, visualisation and analysis of bibliometric networks. and Excel. The search query was the following,
TITLE-ABS-KEY('digital twin') AND (('decision support' OR 'decision-support') AND ('decision-making' OR 'decision making') AND PUBYEAR > 2009 AND PUBYEAR < 2024.

This resulted in 1.733 papers, however only the English-written and in the final publication stage publications were analysed, resulting in a dataset of 1.630 publications. The range date considered in this query relates to the first appearance of the term DT in publications around 2011.

2.2 Results Analysis

The results from the bibliometric study were divided into two main categories, the *Performance Analysis* and the *Science Mapping*. The results focus on analysing the total number of publications and their evolution throughout time, the geographical distribution of the researchers working on these topics, the most influential publications on the topic, and the correlation between the keywords and the identification of possible research trends in this field.

Performance Analysis In this study were performed two techniques that belong to the performance analysis category, which is a type of study that examines the contributions made by various researchers to a particular field. This type of study intends to review and present the performance of different research contributors as authors, institutions, countries, and journals [3]. In this study, it was performed two publication-related metrics analysis, analysing the productivity per active year per publication of the field and the geographical distribution of the researchers. Fig. 2 presents the publication count and the evolution.

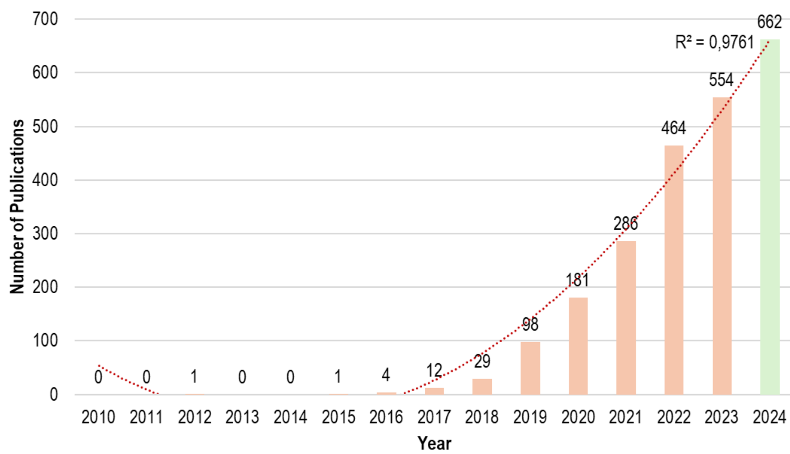


Fig. 2: Evolution of the number of publications of DT and DSS.

The number of publications that mention the two topics has grown in the last few years, representing a growing interest in applying DT to implement decision-support strategies or offer optimised decision-making. A trend line to show the possible evolution of the data has been added as a projection of the number of publications on these two topics in 2024, being around 662 publications. In 2024, until the month of May, 308 papers were published regarding the established search query, being expected to surpass the established projection. Fig. 3 presents the geographical distribution of the research being conducted on this topic.

Considering the results, the country that is most focused on the confluence between DT and decision support is China (16,1%), followed by the USA (14%) and the UK

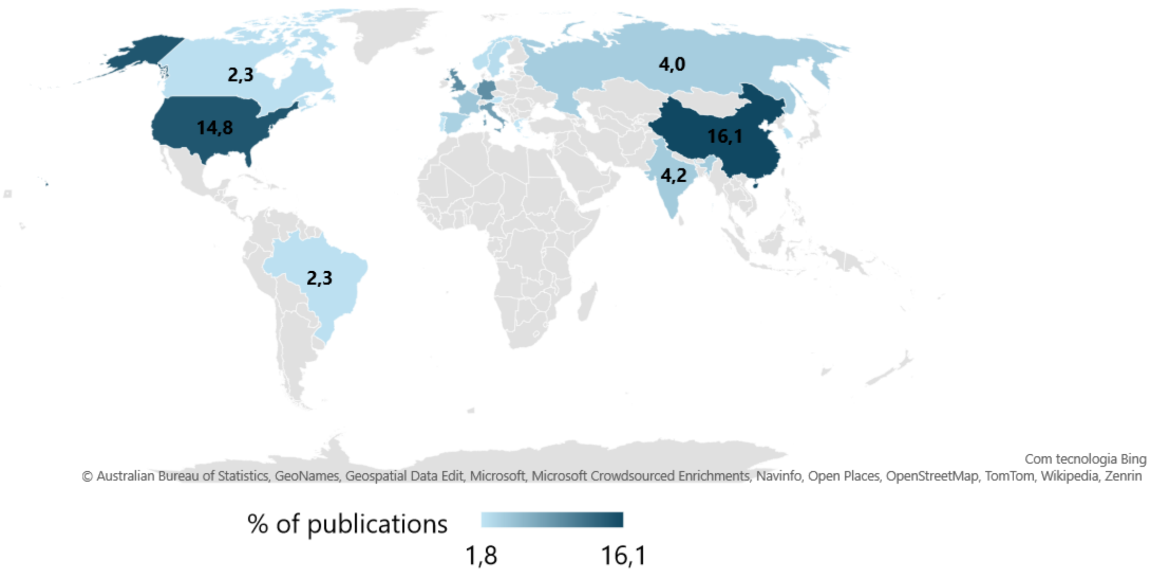


Fig. 3: Top 20 of the countries in terms of geographical distribution (% publications).

(9,9%). In fourth place is Germany (9,8%), and although the Industry 4.0 movement has started there, the focus is not on DT research as in the USA, but more on the Asset Administration Shell (AAS). In the seventeenth place is Portugal (1,9%), where there is already some expression on this topic despite the small size of the research community. Of these top 20 countries, more than half are European countries, demonstrating the efforts made in research into applying technologies and concepts linked to Industry 4.0.

Science Mapping The study performed two techniques of the science mapping category, which pretends to examine relationships between researchers. In this case, it was performed the citation analysis by identifying the most influential publications in the research field, and the co-word analysis intends to explore the existing or future relationships among topics in the research field, in this case focusing on the keywords [3]. Table 1 presents the publications with the most citations in the analysed dataset.

Based on the number of citations, it was possible to identify the top 5 most influential publications regarding DT and decision support. The identified publications focus on digital twin applications in various domains, such as production systems, epidemic outbreak analysis, and optimisation, showing the versatility and significance of DT technology across different domains. Additionally, the high citation could signify that these papers have made significant contributions to the topic of DT and decision support. Recurring to the VOSViewer, it is possible to assess how the decision support is linked to the DT, the research trends, and the main application domains. Fig. 4 presents the authors' keywords co-occurrence network.

The author's keywords co-occurrence network was divided into five clusters according to the parameters established in the VOSViewer. It was possible to identify the following keywords in each cluster, in Cluster 1 (Red), *'cloud computing'*, *'deci-*

Table 1: Highly cited publications.

Ref	Year	Title	Citations
[5]	2020	Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case	1237
[10]	2017	A Review of the Roles of Digital Twin in CPS-based Production Systems	965
[14]	2020	Digital twin: Values, challenges and enablers from a modeling perspective	662
[18]	2017	A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line	346
[15]	2019	Make more digital twins	322

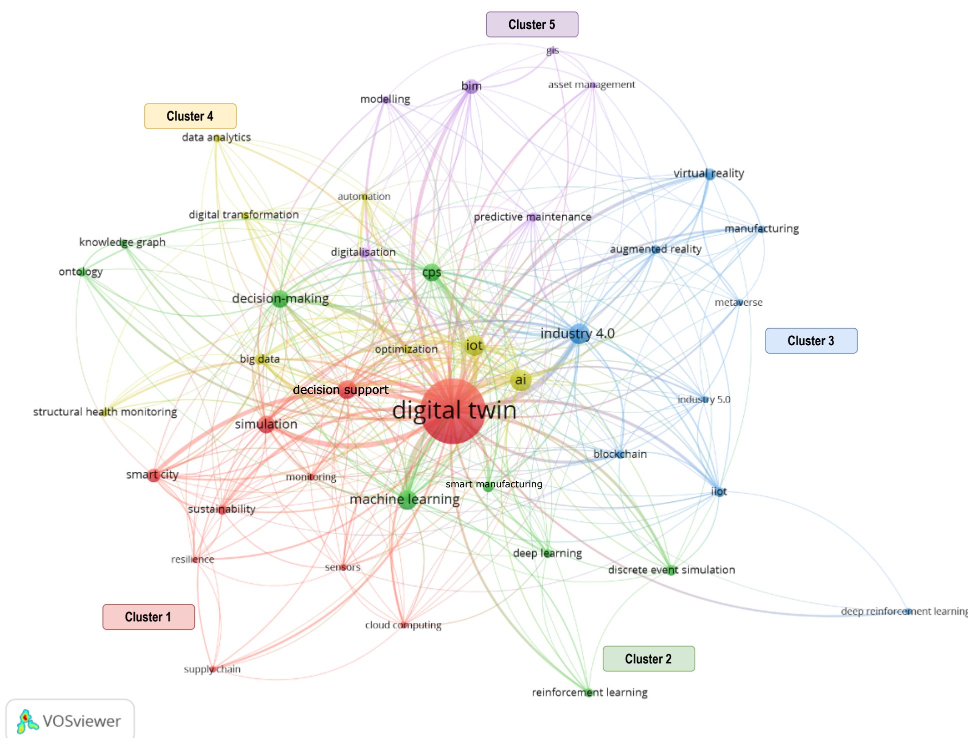


Fig. 4: Author keywords co-occurrence network map of the publications on DT and decision support (Time-frame 2010-2023; n=3880 keywords; threshold of 10 occurrences per keyword, display 42 keywords), with five clusters.

sion support', 'digital twin', 'monitoring', 'resilience', 'sensors', 'simulation', 'smart city', 'supply chain', 'sustainability'; Cluster 2 (Green) *'cps', 'decision-making', 'deep learning', 'discrete event simulation', 'knowledge graph', 'machine learning', 'ontology', 'reinforcement learning', "smar manufacturing*; Cluster 3 (Blue) *'augmented reality', 'blockchain', 'deep reinforcement learning', 'iiot', 'industry 4.0', 'industry 5.0', 'manufacturing', 'metaverse', 'virtual reality'*; Cluster 4 (Yellow) *'ai', 'automation', 'big data', 'data analytics', 'digital transformation', 'iot', 'optimization', 'structural health monitoring'*; and Cluster 5 (Purple) *'asset management', 'bim', 'digitalization', 'gris', 'modelling', 'predictive maintenance'*.

The cluster analysis, in general, revealed that the DT is strongly connected to decision support topics, alongside simulation, IoT, AI, Industry 4.0, machine learning, decision-making, and smart manufacturing. The connection between these topics makes sense as DT uses most of these technologies in the domain of Industry 4.0, focusing on smart manufacturing.

Since the study identified five clusters, its analysis will be further detailed hereafter. **Cluster 1** encompasses diverse digital transformation and smart systems topics, focusing on leveraging technologies like DT, sensors, and simulation for decision support and optimisation. The presence of keywords such as *'supply chain'* and *'sustainability'* suggests a particular emphasis on applying these technologies to enhance supply chain resilience and sustainability efforts. This cluster represents research focusing on DT technology and DSS.

Cluster 2 focuses on advanced computational methods and intelligent systems for decision-making and optimisation in manufacturing contexts. Keywords such as *'deep learning'*, *'machine learning'*, and *'reinforcement learning'* indicate a strong emphasis on leveraging AI and machine learning techniques within cyber-physical systems (CPS) and smart manufacturing environments. This cluster represents research to develop autonomous decision-making systems and optimise manufacturing processes by integrating AI and CPS technologies.

Cluster 3 explores emerging technologies and paradigms such as Industry 4.0, augmented reality (AR), and blockchain in the context of manufacturing and industrial systems. Keywords like *'industry 5.0'* and *'metaverse'* suggest a forward-looking perspective on the future of manufacturing and human-machine interactions. This cluster represents research on digital innovation, immersive technologies, and industrial automation, focusing on enhancing manufacturing processes and workforce capabilities through digital transformation.

Cluster 4 encompasses a broad range of topics related to data-driven decision-making, automation, and optimisation in various domains. Keywords such as *'big data'*, *'IoT'*, and *'data analytics'* highlight the importance of data-driven approaches for optimising processes and systems. This cluster represents research to leverage advanced analytics, automation technologies, and IoT devices to enhance operational efficiency, monitor structural health, and drive digital transformation initiatives across different sectors.

Lastly, **Cluster 5** focuses on asset management, digitalisation, and predictive maintenance strategies within the context of infrastructure and asset-intensive industries.

Keywords like *'BIM'* (Building Information Modelling) and *'predictive maintenance'* indicate a particular emphasis on leveraging digital technologies and modelling techniques to optimise asset performance and maintenance practices. This cluster represents research to improve asset lifecycle management, reduce downtime, and enhance operational efficiency through data-driven maintenance strategies and digital transformation initiatives.

3 Conclusions

The results from the bibliometric study conducted on DT and DSS made it possible to identify the growing interest in conjugating DT and decision support, given the increase in publications on the topic in recent years. This suggests a recognition within the research community of the potential of DT technology to enhance decision-making processes in manufacturing systems. The geographical distribution analysis showed a diverse and widespread interest in the topic, with contributions from countries such as China, the USA, and the UK. This global perspective highlights the importance of collaboration and knowledge sharing in advancing research and innovation in this field. The analysis of highly cited publications provided insights into influential research contributions in this field. These publications cover various topics, from epidemic outbreak analysis to production systems optimisation, underscoring the versatility and significance of DT technology across different domains. The cluster analysis revealed that DT is closely connected to various technologies and concepts, such as IoT, AI, machine learning, and Industry 4.0. This indicates a trend towards integrating DT into broader digital transformation initiatives within the manufacturing sector, leveraging advanced technologies to drive efficiency and innovation. The identified clusters enabled the identification of application domains and research trends related to DT and decision support. These include supply chain management, sustainability, smart manufacturing, asset management, and predictive maintenance. In terms of future research, a systematic literature review will be conducted to deepen and clarify the state-of-the-art research on the topic, identifying the main contributions and the research challenges of the field.

Acknowledgements This work was supported by national funds through FCT/ MCTES (PIDDAC): CeDRI, UIDB/05757/2020 (DOI:10.54499/UIDB/05757/2020) and UIDP/05757/2020 (DOI: 10.54499/UIDP/05757/2020); and SusTEC, LA/P/0007/2020 (DOI: 10.54499/LA/P/0007/2020). Flavia Pires thanks FCT Portugal for the PhD Grant SFRH/BD/143243/2019.

References

1. Ahuett-Garza, H., Kurfess, T.: A brief discussion on the trends of habilitating technologies for industry 4.0 and smart manufacturing. Manufacturing Letters pp. 60–63 (2018)
2. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W.M.: How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research **133**, 285–296 (2021)
3. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W.M.: How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research **133**, 285–296 (2021)

4. Franke, F., Franke, S., Riedel, R.: AI-based Improvement of Decision-makers' Knowledge in Production Planning and Control. IFAC-PapersOnLine **55**(10), 2240–2245 (2022)
5. Ivanov, D.: Predicting the impacts of epidemic outbreaks on global supply chains: A simulation- based analysis on the coronavirus outbreak (covid-19 / sars-cov-2) case. Transportation Research Part E Logistics and Transportation Review (2020)
6. Kagerman, H., Wahlster, W., Helbig, J.: Securing the future of German manufacturing industry Recommendations for implementing the strategic initiative Industrie 4.0 Final Report of the Industrie 4.0 Working Group. Tech. rep., ACATECH (2013)
7. Khan, M., Wu, X., Xu, X., Dou, W.: Big data challenges and opportunities in the hype of Industry 4.0. IEEE International Conference on Communications pp. 1–6 (2017)
8. Kunath, M., Winkler, H.: Integrating the Digital Twin of the manufacturing system into a decision support system for improving the order management process. In: 51st CIRP Conference on Manufacturing Systems. vol. 72, pp. 225–231 (2018)
9. Leitao, P., Pires, F., Karnouskos, S., Colombo, A.W.: Quo Vadis Industry 4.0? Position, Trends, and Challenges. IEEE OJIES **1**, 298–310 (2020)
10. Negri, E., Fumagalli, L., Macchi, M.: A review of the roles of digital twin in cps-based production systems. Procedia Manufacturing pp. 939–948 (2017)
11. Nouinou, H., Asadollahi-Yazdi, E., Baret, I., Nguyen, N.Q., Terzi, M., Ouazene, Y., Yalaoui, F., Kelly, R.: Decision-making in the context of Industry 4.0: Evidence from the textile and clothing industry. Journal of Cleaner Production **391** (2023)
12. Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., et al.: The prisma 2020 statement: An updated guideline for reporting systematic reviews. International Journal of Surgery **88** (2021)
13. Pires, F., Cachada, A., Barbosa, J., Moreira, A.P., Leitão, P.: Digital twin in industry 4.0: Technologies, applications and challenges. In: 2019 IEEE 17th INDIN. pp. 721–726 (2019)
14. Rasheed, A., San, O., Kvamsdal, T.: Digital twin: Values, challenges and enablers from a modeling perspective. IEEE Access (2020)
15. Tao, F., Qi, Q.: Make more digital twins. Nature **573**, 490–491 (2019)
16. VanDerHorn, E., Mahadevan, S.: Digital Twin: Generalization, characterization and implementation. Decision Support Systems **145** (2021)
17. Yinggang, L., Xiangrong, T.: Social recommendation system based on multi-agent deep reinforcement learning. In: 2022 IEEE 8th International Conference on Cloud Computing and Intelligent Systems (CCIS). pp. 371–377 (2022)
18. Zhang, H., Liu, Q., Chen, X., Zhang, D., Leng, J.: A digital twin-based approach for designing and multiobjective optimization of hollow glass production line. IEEE Access pp. 26901–26911 (2017)

Prediction of Hand Movements Using the k-Means Method

Gregory Marcelo Vergilino[1], José Lima[1], Uilian José Dreyer[2], Ana I. Pereira[1], and Péricles Valera Rialto Junior

[1] Research Center in Digitalization and Intelligent Robotics, Instituto Politécnico de Bragança, Portugal
a61451@alunos.ipb.pt {jllima,apereira}@ipb.pt
[2] Universidade Tecnológica Federal do Paraná - UTFPR Curitiba, Brazil uiliandreyer@utfpr.edu.br

Abstract. The major problem when using optical sensors to collect data for an automated prosthesis is predicting which finger is moving just by analyzing forearm data. This can be a significant challenge when designing an intelligent prosthesis, because it needs to be as efficient as possible. Nevertheless, it's all about improving health quality for those with motion disabilities. So, a crucial first step is to correctly analyze all this data. The proposal of this work is to use the k-Means method to solve this problem. k-Means is a machine learning technique renowned for its ability to group data into distinct clusters, based on similarity measures. By employing this method, the vast amount of data collected from the forearm can be efficiently organized into meaningful groups, or clusters, each representing different patterns of movement. This clustering process not only facilitates a more structured analysis but also enables the system to discern subtle variations in movement patterns, ultimately enhancing the accuracy and effectiveness of the automated prosthesis.

Keywords: Clustering · Data Science · k-Means.

1 Introduction

This research is based on a previous study [3], that developed an automated prosthesis using optical sensors on patients' forearms. However, the major problem was to recognize and identify which finger was moving analyzing the obtained data. If the prosthesis is not able to automatically discern this, it won't be an efficient solution. This is the principal objective that will be discussed further: how the automated prosthesis should be able to intelligently select which prosthetic finger is moving just by using the forearm data analysis.

This document is organized as follows. Section 2 presents the study methodology. The results are discussed in section 3 and in the last section the conclusions are presented.

2 Methodology

This work has the following methodology. In the first phase it was necessary to collect all data.The previous work is based on using Fiber Bragg Gratings (FBG) optical sensors. The FBG strain sensor information is encoded in wavelength, meaning the strain sensed by the FBG produces a wavelength shift proportionally [3]. The second phase, with all movements captured, it will be analysed all time-series by using k-Means method (dividing the series in clusters).

2.1 Data

The sensor analysis and data capture process involved instructing patients to move one finger at a time. Patients were asked to perform movements with four fingers individually: the index, middle, ring, and little finger. Due to the large volume of data, detailed data from only one patient is presented here, with summary statistics provided for all ten patients.

Then, it was obtained two important contents. For example Table 1 presents some data obtained from patient 1.

Table 1: Database from patient 1.

Time [s]	Fiber Bragg Grating 1 [nm]
0	1543.466
0.02	1543.466
0.04	1543.466
0.06	1543.465
0.08	1543.466
0.18	1543.467
...	...
184.92	1543.525

It can be observed that the obtained data was acquired from 0 s to 184.92 s. In Fiber Bragg Grating 1 column, it has the response in range from 1543.413 to 1543.594. Even these numbers migth be too large, the scale is still small (in nanometers order). So, the first patient has an database with 9949 lines. As presented in Table 2, it has number of lines that database has (count), its mean value (mean), deviation(std), its range(by min and max). And its distribution including median (by 25%,50%,75% columns). For the other patients the data is similar.

Table 2: Describe from patient 1.

	count	mean	std	min	25%	50%	75%	max
Time [s]	9247	92.460	53.390	0.000	46.230	92.460	138.690	184.920
Fiber Bragg Grating 1 [nm]	9247	1543.509	0.031	1543.413	1543.479	1543.513	1543.535	1543.594

2.2 Methods

Clustering strategies was used in this work. k-Means method is an unsupervised learning method used for clustering data points. It gathers points and divides the data into groups, known as clusters. In each cluster will be described the main characteristics (such as their minimum, mean and maximum points).

Considering that the data do not have information about how many fingers movements were done during tests, it was necessary to determine the optimal clusters number. Using Python programming language, it was made the elbow method [1] (see Fig. 1).

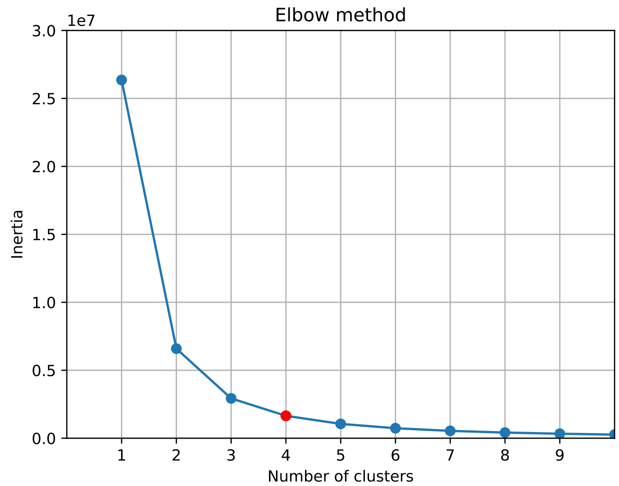


Fig. 1: Elbow method for selecting number of clusters

By the graph, it can be noticed that the optimal cluster number is four clusters. Such as the number of fingers that were used during the test.

3 Results

It is possible to obtain the first time and response graph, by Table 1 - Database from patient 1. Using scatter diagram, the signal is like Fig. 2.

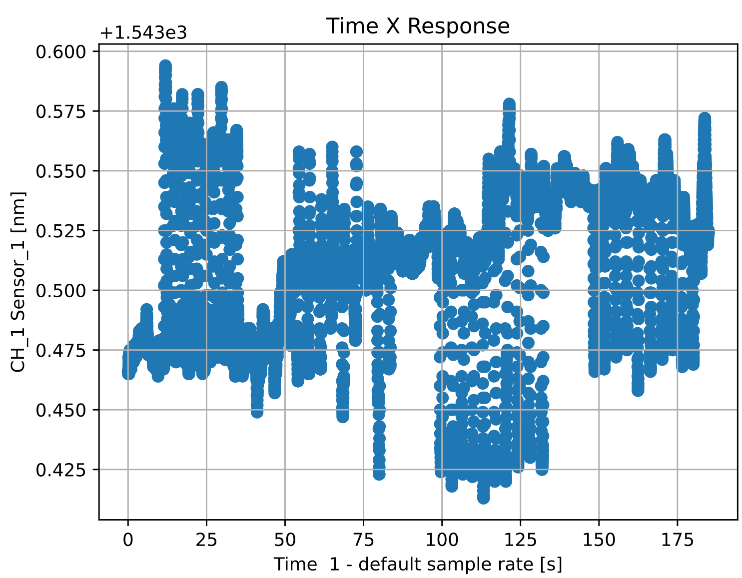


Fig. 2: Normal Time and Response graph.

Notice that it has 4 different regions in this chart, but it is difficult to distinguish where each one starts and ends. So by using k-Means method, with four clusters, it is obtained a new chart (Fig. 3).

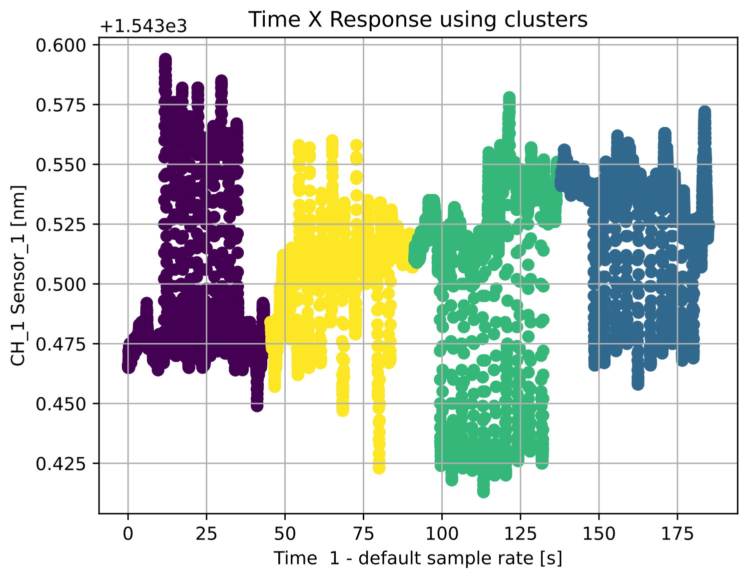


Fig. 3: Time and response graph using clusters.

It is more visible than the last one, and easier to analyze. Table 3 will present the standard deviation of the dataset response for each cluster (std), the minimum value (min), the mean value (mean), the maximum value(max), and the distribution including median (by 25%,50%,75% columns). All this procedure was repeat for each patient.

Table 3: Clusters database.

clusters	count	mean	std	min	25%	50%	75%	max
0	2300	1543.532	0.020384	1543.458	1543.527	1543.537	1543.543	1543.572
1	2320	1543.505	0.016987	1543.423	1543.502	1543.508	1543.514	1543.56
2	2316	1543.514	0.037107	1543.413	1543.512	1543.524	1543.539	1543.578
3	2311	1543.485	0.027189	1543.449	1543.472	1543.476	1543.482	1543.594

As can be seen, each cluster has different minimum, mean and maximum values. But, it is too slight. So with this database, it is possible to automatically identify each movement based on the centroid values [2]. After determining the values of each centroid, these values will contribute to analyzing the area covered by each finger and also determining the noise regions in the graph.

4 Conclusions and Future Work

This paper addressed an automated prosthesis that should be able to select the finger using data analysis. In this work K-Means method was used to identify four movements of ten patients. Although each cluster has is too slight, the results were encouraging. As future work, considering the information from each cluster, we want to identify what finger do the movement.

References

1. Patel, P., Sivaiah, B., Patel, R.: Approaches for finding optimal number of clusters using k-means and agglomerative hierarchical clustering techniques. In: 2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSP). pp. 1–6 (2022). https://doi.org/10.1109/ICICCSP53532.2022.9862439
2. Premkumar, M.S., Ganesh, S.H.: A median based external initial centroid selection method for k-means clustering. In: 2017 World Congress on Computing and Communication Technologies (WCCCT). pp. 143–146 (2017). https://doi.org/10.1109/WCCCT.2016.42
3. Valera Rialto Júnior, P., Henrique Dureck, E., Kalinowski, A., Gomes Fernandes, D., Carlos Cardozo da Silva, J., José Dreyer, U.: Automated prosthesis control by using fiber bragg grating forearm sensor. In: 2021 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC). pp. 1–3 (2021). https://doi.org/10.1109/IMOC53012.2021.9624845

Development of a wearable device for monitoring the activity of elderly people with dementia: first insights

Inês C. Rocha[1], Marcelo Arantes[1], Mariana Carvalho[1], António Moreira[1], João L. Vilaça[1], Pedro Morais[1], Demétrio Matos[2], and Vitor Carvalho[1]

[1] 2Ai, School of Technology, Polytechnic University of Cávado and Ave (IPCA), Campus of IPCA, 4750-810, Barcelos, Portugal
{irocha, jarantes, mccarvalho, amoreira, jvilaca, pmorais, vcarvalho}@ipca.pt
[2] ID+, Research Institute in Design, Media and Culture, School of Design, Polytechnic of Cávado and Ave, 4750-810 Barcelos, Portugal
dmatos@ipca.pt

Abstract. The world has been witnessing a demographic shift with an ageing population, leading to an increase in the prevalence of neurodegenerative diseases. Among all, dementias represent one of the major causes of dependency and morbidity among the elderly, causing changes in the behavior, memory, and the ability to perform daily tasks. The provision of care for the elderly faces significant challenges, with caregivers needing to develop physical and psychological capacities to perform their duties. This work presents the first steps in the development of a wearable device (wristband) for real-time activity monitoring, promoting adequate and minimally invasive care for elderly people with dementia. The wristband methodology to be developed is divided into two parts, the functionalities methodology and the design methodology, which are responsible for the functional and non-functional requirements.

Keywords: Wearable Device · Elderly People · Dementia · Design

1 Introduction

The United Nations notes that the global population is rapidly ageing, with all countries seeing a rise in elderly populations. This demographic shift has led to a sharp increase in dementia cases, affecting over 55 million people [1]. Caregivers play a fundamental role in the care of elderly with dementia, helping them to carry out daily activities, and their role is associated with great responsibility and physical and psychological demands [2]. Therefore, recognizing and understanding the specific needs faced by this target audience is essential for designing tailored interventions and systems that contribute to a better quality of life both for elderly and their caregivers [3]. The project were this work in included "Platform for Routine Measurement and Monitoring of Activity in Elderly People with Dementia" aims to develop and validate a new system for location and activity monitoring, promoting active, healthy, and inclusive ageing. This new system will feature two wearable data acquisition devices, the HowMi [4] (in-ear) and a wristband – which is the aim of the work presented in this paper. These devices will acquire biometric data such as body temperature, heart rate, blood oxygen saturation level, but also daily activity data (activity index, number of steps and distance traveled, fall detection, sleep rhythm analysis) and indoor and outdoor location. The data will be sent to a server, where it will be managed and monitored. The information will be made available on a web portal for institutional professionals, caregivers, and family members, helping them to decide on the most appropriate type of care for the elderly person's situation. This paper is organized in 4 sections, Section 2 "Related Work" describes the current literature framework in the area of wearable devices in the context of monitoring the activity parameters of elderly people with dementia, Section 3 "Methodology" describes the approach used for achieve the setting goals, and Section 4 "Future Work" describes the next steps that will be taken to continue the work according to the methodologies defined.

2 Related work

The evolution of wearable devices for monitoring daily activities has garnered significant importance in recent years for promoting safe, non-invasive, and cost-effective monitoring through intuitive use. Several studies have focused on the incorporation of wearable devices to monitor health metrics, behavioral and sleep patterns, activity levels and real-time tracking in vulnerable population. This section provides an overview of recent and notable contributions in this field. One of the studies in this domain explored the use of the WHOOP biosenser device for monitoring sleep and heart rate variability in individuals at risk for Alzheimer's disease [5]. Their findings demonstrated that wearable devices can allow caregivers or clinician researchers, an early detection and continuous monitoring, thereby facilitating timely and individualized interventions to manage and potentially delay the progression of cognitive decline. In addition, another study used an accelerometer wearable device for examine whether Parkinson's disease (PD) intensity, influences the relationship between functional connectivity and cognitive performance [6]. Their work revealed that performing moderate physical activity can improve cognitive functionality in individuals with PD, such as visuospatial function, memory, and executive function. Recent advancements have also been made in the field of artificial intelligence and the development of multi-modal sensors integrated in wearable devices for measuring a range of physiological and environmental parameters. Given this, a recent study introduced a multi-sensor wearable device to detect behavioral symptoms of dementia, including motor agitation, verbal, and physical aggression [7]. Their study revealed that using a device with personalized sensors can, not only, enhance the performance of generic models in classifying agitated behaviors but understanding the complex and progression of the pathology. Overall, the existing literature underscores the promising potential of wearable devices in improving the care of elderly individuals with dementia or prevent who are heading towards it. Despite the innovative advancements, there are still substantial challenges in terms of long-term usability since certain technologies need improvement once is given the specific needs of elderly individuals with dementia [3]. Another factor in consideration is that caregivers play a fundamental paper on the daily lives of these individuals. It is therefore of the utmost importance the adequately prepare and integration in the assessment and use of these wearable devices, enhancing the relationship between the caregiver and the patient [8]. The continuous collaboration among researches, healthcare providers, technology and design developers can ensure the development of cost-effective wearable devices by reducing the economic burden of dementia care.

3 Methodology

The structure of this work, particularly the design of the wristband, comprises two main sections: the functionalities methodology and the design methodology, as shown in Figure 1.

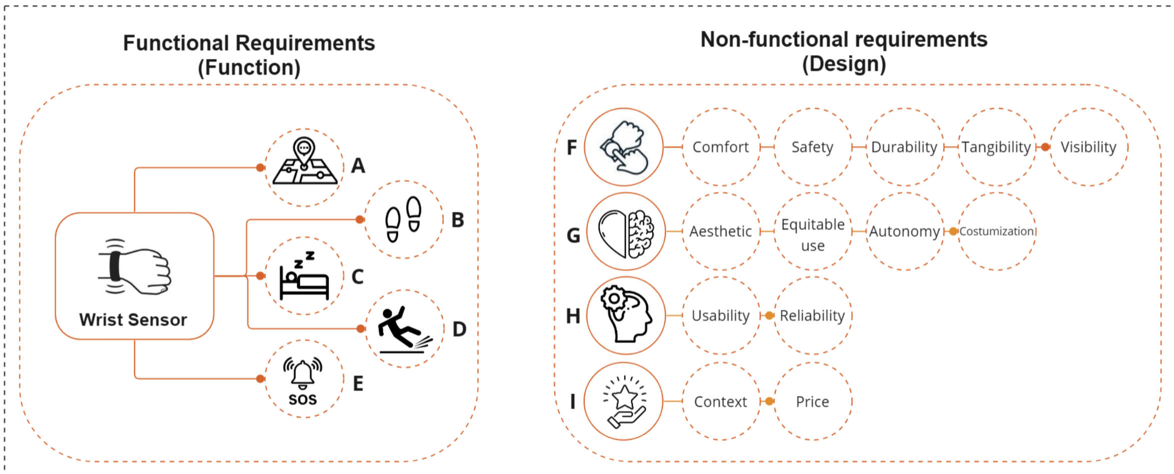


Fig. 1: Functional and non-functional requirements of the wristband.

These are outlined to establish the fundamental functional and non-functional requirements for the design of the wristband wearable device. In terms of functionality, the device must be able to communicate via ISM (Industrial, Scientific and Medical) Radio Frequency with the system's components. This communication will be done using the BLE (Bluetooth Low Energy) protocol in indoor environments and using the LoRa (Long Range) protocol in outdoor environments. Regarding the location requirement (A), indoor location will be given by connecting the wristband with BLE beacon devices and outdoor location will be given by a GPS/GNSS (Global Positioning System/Global Navigation Satellite Systems) module. This system will provide information to family members and caregivers about the patient's real-time location in cases of spatial memory loss episodes.It will have inertial sensors for sensory fusion and consequent analysis of daily activity patterns such as activity index, number of steps and sleep rhythm analysis (B,C).It will also include a fall detection system (D) and an emergency alert system (E), which will inform caregivers for a quick response to the necessary care, as this is a highly vulnerable group in terms of physical and cognitive abilities. On the other hand, non-functional requirements provide guidelines that facilitate the adoption of the wristband. In view of this, four domains of requirements have been defined: the physical domain (F) focuses on the direct relationship between the user and the product. The emotional domain (G) establishes considers ethical and visual issues, and the cognitive domain (H) ensures the complete usability of the wristband according to the abilities of the elderly. Finally, the requirements of the implementation success domain (I) focus on assisting in the integration of the wristband for the problem at hand.

4 Future Work

Further steps will follow the methodology being applied, starting by selecting the best sensors required for the wristband, including inertial sensors and GPS. We will then analyze and test the communication protocols (Bluetooth Low Energy and LoRa (Long Range)) and the integration with the HowMI device [4], internal location devices and the data server. Additionally, we will develop AI algorithms to analyze the collected data, enabling advanced functionalities such as localization tracking, routine monitoring, sleep pattern analysis, and anomaly detection. Finally, in terms of design, we aim to start ergonomic studies of the wrists and prepare a questionnaire for caregivers in order to obtain feedback on the idealized design of the wristband.

Acknowledgments

This project was funded through the Foundation for Science and Technology (FCT) under the projects UIDB/05549/2020 (DOI: 10.54499/UIDB/05549/2020), UIDP/05549/2020 (DOI: 10.54499/UIDP/05549/2020), LASI-LA/P/0104/2020, and CEECINST/00039/2021. This work was also funded by the Innovation Pact HfFP – Health From Portugal, co-funded from the "Mobilizing Agendas for Business Innovation" of the "Next Generation EU" program of Component 5 of the Recovery and Resilience Plan (RRP), concerning "Capitalization and Business Innovation", under the Regulation of the Incentive System "Agendas for Business Innovation".

References

1. World Health Organization. Dementia, 2024.
2. Márcia Maria Pires Camargo Novelli, Ricardo Nitrini, and Paulo Caramelli. Cuidadores de idosos com demência: perfil sociodemográfico e impacto diário. *Revista de Terapia Ocupacional da Universidade de São Paulo*, 21(2):139–147, 2010.
3. Júlia D. Rodrigues, Pedro Morais, and Vítor Carvalho. Routine measurement and monitoring system for the activity of elderly people with dementia: A systematic review. *9th International Conference on Sensors and Electronic Instrumentation Advances (SEIA' 2023)*, 20-22 September:139–144, 2023.
4. Claudino Costa, João M. Faria, Diana Guimarães, Demétrio Matos, António H.J. Moreira, Pedro Morais, João L. Vilaça, and Vítor Carvalho. A wearable monitoring device for covid-19 biometric symptoms detection. *IRBM*, 44(6):100810, 2023.
5. K. Niotis O. Scheyer A. Rahman M. Berkowitz R. Krikorian H. Hristov G. Sadek S. Bellara Richard S. Isaacson N. Saif P. Yan. Feasibility of using a wearable biosensor device in patients at risk for alzheimer's disease dementia. *The Journal of Prevention of Alzheimer's Disease*, 7:104–111, 2020.
6. Erin K. Donahue, Siva Venkadesh, Vy Bui, Angelie Cabrera Tuazon, Ryan K. Wang, Danielle Haase, Ryan P. Foreman, Jared J. Duran, Andrew Petkus, David Wing, Michael Higgins, Daniel P. Holschneider, Ece Bayram, Irene Litvan, Michael W. Jakowec, John Darrell Van Horn, Dawn M. Schiehser, and Giselle M. Petzinger. Physical activity intensity is associated with cognition and functional connectivity in parkinson's disease. *Parkinsonism & Related Disorders*, 104:7–14, 2022.
7. Andrea Iaboni, Sofija Spasojevic, Kristine Newman, Lori Schindel Martin, Angel Wang, Bing Ye, Alex Mihailidis, and Shehroz S. Khan. Wearable multimodal sensors for the detection of behavioral and psychological symptoms of dementia using personalized machine learning models. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 14(1):e12305, 2022.
8. Jesus Favela, Dagoberto Cruz-Sandoval, Arturo Morales-Tellez, and Irvin Hussein Lopez-Nava. Monitoring behavioral symptoms of dementia using activity trackers. *Journal of Biomedical Informatics*, 109:103520, 2020.

Empowering Santiago Island: Renewable Energy Integration in Cape Verde

Cintia Garcia[1], Luís Gabriel Teixeira Lopes[1], Gabriel Fonseca Oliveira Roma[1,2], and Ângela Paula Ferreira[1]

[1] CeDRI, SusTEC, Instituto Politécnico de Bragança, Portugal
[2] Department of Electrical Engineering (DAELE), Federal Technological University of Paraná (UTFPR), Brazil
a49427@alunos.ipb.pt, a48805@alunos.ipb.pt, gabrielroma@ipb.pt, apf@ipb.pt

Abstract. Cape Verde, an archipelago nation in the Atlantic Ocean, faces significant challenges in meeting its energy demands due to its dependence on imported fossil fuels. Through extensive data collection, this study evaluates the future state of renewable energy deployment, including the penetration levels of photovoltaic and wind renewable sources and their contribution to the overall energy supply on the island. This study aims primarily at critically assessing the scenarios outlined in the energy plan proposed by the government of Cape Verde, with the objective of identifying alternative solutions that offer substantial improvements.

Keywords: Renewable Energy Source · Energy Planning · Photovoltaic Power · Wind Power · Sustainable Energy.

1 Introduction

Given the increasing global concern and the increasing efforts to mitigate the impacts of climate change, the transition towards a more sustainable electricity grid emerges as a critical priority for many nations in their long-term strategic vision [1]. Island developing states, such as Cape Verde, face unique challenges in their energy transition characterized by geographical isolation, heavy reliance on fuel imports, and limited integration into global energy markets [2].

Over time, the islands of Cape Verde have witnessed a steady rise in both energy consumption and pricing, attributed to factors such as population expansion, industrial advancement, and reliance on the global market [3]. This trend has intensified the demand placed on power generation facilities, leading to a significant reliance on non-renewable energy sources, primarily due to insufficient investments in renewable energy-based power generation facilities.

In response to global trends, the Cape Verdean government has initiated a plan to expand the island's power generation capacity, with a focus on increasing the share of renewable energy-based power generation. This plan is structured into three phases [4]:

- 2025 scenario - improving the electrical grid to 30% of renewable energy-based power generation;
- 2030 scenario - improving the participation of renewable energy-based power generation to 50%.
- 2040 scenario - having nearly 100% participation of renewable energy-based power generation in the energy grid.

However, the integration of renewable energy sources (RES) presents a multifaceted challenge. It involves not only the physical integration of various energy sources into the grid but also addressing the dynamic nature of renewable energy generation. Managing fluctuations in generation, which can arise from sources such as photovoltaic and wind power, is essential for maintaining grid stability and reliability [5].

In addition, it is crucial to ensure that renewable energy sources (RES) are compatible with existing infrastructure and that grid operations are managed to accommodate intermittent generation. Addressing these challenges requires careful attention to fully harness the benefits of RES [6]. Therefore, it is essential to plan and simulate the behaviour of energy systems, which is fundamental to the successful integration of RES and the transition to a more sustainable energy system.

Advanced energy systems computer models play a vital role in predicting system variables, simulating scenarios, and informing strategies for a more resilient future grid. These models enable stakeholders to analyze parameters, plan for contingencies, and make informed choices, ultimately enhancing the resilience and performance of energy grids.

Recent research has been dedicated to energy-environmental planning and the shift towards renewable energy-based systems, particularly in the context of Cape Verde. Tavares et al. [7] investigated Cape Verde's energy landscape, finding a heavy reliance on fossil fuels despite the ample renewable potential, such as wind and solar photovoltaic. Their study, proposing an Autoregressive Integrated Moving Average (ARIMA) model to assess energy system behaviour, revealed significant under-utilization of renewables and resulting greenhouse gas emissions. Electrical losses, especially on Santiago Island, contribute to market imbalances.

Segurado et al. [8] examine ways to increase renewable energy integration on São Vicente Island from the Cape Verde archipelago, focusing on wind power. Using the Hydrogen to Renewable Energy Source model (H2RES), they explore scenarios for maximizing renewable energy utilization. Despite São Vicente's reliance on fossil fuels and intermittent wind power, the study suggests strategies, including pumped hydro storage, to enhance renewable energy penetration. Results indicate the potential to achieve over 30% yearly penetration of RES in the electricity supply system, mainly through wind power.

Pombo et al. [9] aimed to decarbonize Santiago Island's energy system in Cape Verde by proposing a flexibility-enabling planning approach. The study analyzed the island's energy system over a 20-year horizon, considering renewable energy targets set by the government. The research explored different scenarios and conducted a sensitivity analysis to understand Santiago's energy future. Results showed that flexibility exploitation could achieve up to 85% savings and enable significant decarbonization through the electrification of other sectors.

This study utilizes the EnergyPLAN advanced energy systems computer model [10] to forecast potential scenarios of renewable energy source (RES) facilities for the year 2030 in Santiago Island. Given its pivotal role in energy consumption and diverse economic activities, the focus of this study will be exclusively on this island of the Cape Verde archipelago. EnergyPLAN has been selected due to its user-friendly interface and its capability to simulate electrical systems on an hourly basis. Furthermore, the tool allows for the modelling and simulation of complex energy systems, incorporating factors such as energy demand, supply, conversion, and storage [11].

The paper is structured as follows: Section 2 is an exploration of the energy systems of Cape Verde and the island of Santiago, covering a brief history of the renewable energy sources' participation in the electrical grid. Section 3 introduces the methodology devised for a case study, displaying the data collection and utilized tools. Following that, Section 4 describes the simulation-generated dataset and the adjustments made aimed at the optimization of the RES model implemented. Section 5 summarizes the main conclusions drawn and outlines possible solutions in future research.

2 Energy Landscape of Cape Verde and Santiago

2.1 The archipelago of Cape Verde

Cape Verde, an archipelago in West Africa consisting of 10 islands, located approximately 600 km off the coast of continental Africa in the Atlantic Ocean. Despite its abundant solar and wind resources, the country is heavily dependent on imported refined oil for energy, with its 540,000 inhabitants dispersed across 9 islands facing extreme external energy dependence [9].

In alignment with global energy transition goals, the local government set targets in 2011 aiming for 50% and 100% renewable energy sources (RES) utilization. However, progress has been slow due to various factors, such as the impact of the COVID-19 pandemic and economic downturns on the islands [12]. The country initially made significant progress in its energy transition, reaching an 18% share of RES as early as 2012. However, progress stalled due to the general low reliability of the grid. In 2020, RES shares were still at 21% [13].

Despite the government's efforts, Cape Verde has faced challenges in achieving full energy access, with 10% of its population currently lacking access to electricity and up to 30% having very limited access. These challenges and future objectives are common to other developing island nations [2].

2.2 Santiago Island

Santiago Island, the largest and most populous island in Cape Verde, holds a pivotal position in the country's energy consumption, representing 49.8% of the total in 2022 [14]. However, despite its considerable potential for renewable energy, Santiago's reliance on fossil fuels is even more pronounced than the Cape Verdean average, with renewables accounting for only 15% of its electricity production. Specifically, wind and photovoltaic energy contribute a mere 12% and 3%, respectively, highlighting the island's significant dependence on non-renewable sources. In 2021, Santiago's energy infrastructure included fossil-fuelled generators, a wind farm, and a solar farm [15].

Wind and solar resources on Santiago Island exhibit peak availability from January to June, with a notable decline during the summer months, particularly for solar energy due to cloudiness during the wet season [9].

Another significant challenge faced by the island is the high rate of electricity losses (technical and commercial), especially on Santiago Island, where losses amount to approximately 36% [16]. Although there has been some improvement in the quality of service provided to customers since 2017, energy losses, particularly non-technical ones, continue to impact the distribution system operator (DSO) processes. These losses result in financial constraints for the operators and contribute significantly to the high cost of electricity in the archipelago [7].

3 Methodology

The primary aim of this research is to give insights on the deployment of the renewable energy infrastructure on Santiago Island, with the overarching goals of diminishing reliance on thermal power generation, mitigating CO_2 emissions, curtailing the necessity for fossil fuel imports, and striving towards a 50% integration of RES into the island's power grid.

To simulate the scenarios, the potential locations for the deployment of the renewable energy sources (RES) are the ones provided in the document published by the Cape-Verdean Ministry of Tourism, Industry and Energy [17], as shown in the Figure 1.

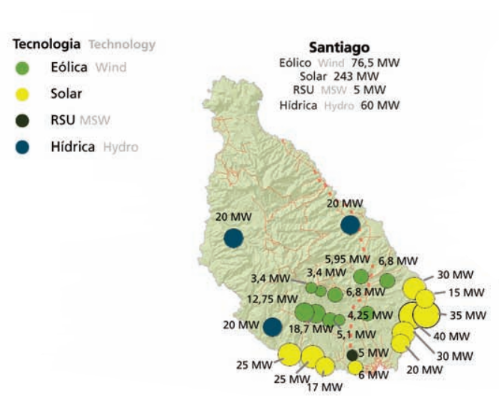


Fig. 1: Possible implementation locations of the RES.

The meteorological data with hourly distribution has been obtained from the platform Renewables.ninja, which provides high-resolution renewable energy data worldwide, including crucial parameters such as wind speed, solar irradiance and capacity factors [18]. Then, the load profile data of Santiago island was used from the Cape Verde reference system [19].

The sizing of both thermal and RES power plants was based on the Cape Verdean government bulletin (business-as-usual scenario) [4]. The bulletin provides the values for the whole country, so the values to dimension the energy sources in Santiago island were estimated to be 40% of the overall installed power, corresponding to the share of Santiago's energy consumption in relation to Cape Verde's total consumption.

4 Results

By introducing and applying all the aforementioned input data in the computational model provided by the EnergyPLAN tool, from the generated dataset, it is possible to depict the main energy sources Figure 2 describing the mix of energy production on Santiago island for the year 2030.

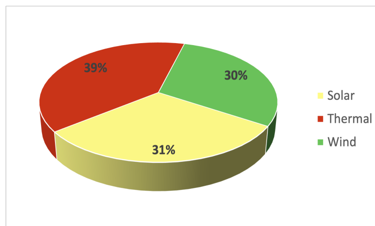
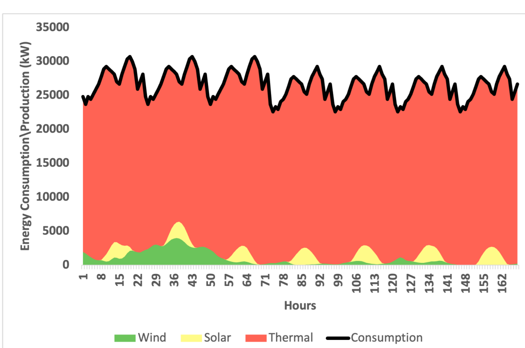


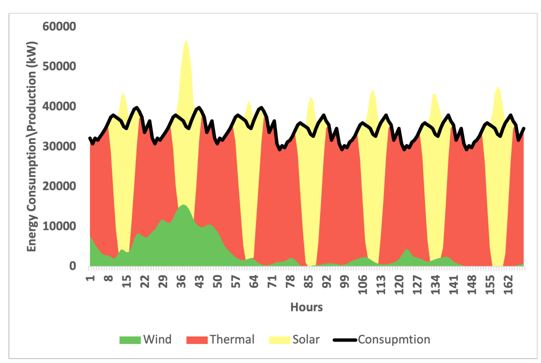
Fig. 2: Forecast of the primary energy mix for Santiago island in 2030.

The dynamics of energy generation and consumption on Santiago Island during typical summer and winter weeks for the year 2030 are illustrated in Figures 3 and 4. Figure 3a presents software-generated outputs of energy generation and consumption for a typical summer week using real data from the Electra report of 2022 [14]. Figure 3b displays the projected energy generation and consumption for a summer week in 2030, based on data from the Cape Verdean government bulletin [4].

Similarly, Figure 4a shows the energy generation and consumption for a typical winter week using data from the 2022 Electra report, while Figure 4b illustrates the same parameters for a winter week in 2030, using projections from the Cape Verdean government bulletin.

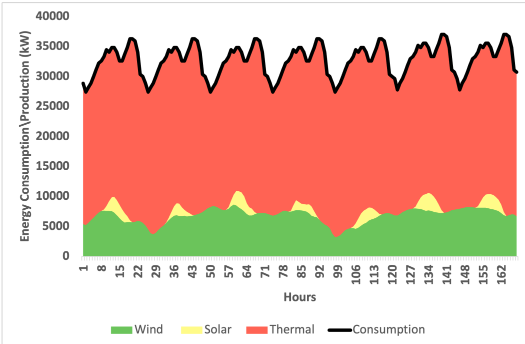


(a)

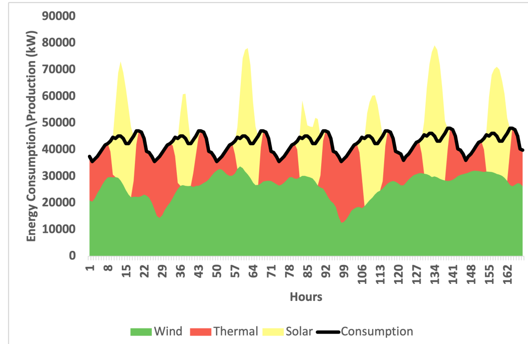


(b)

Fig. 3: Comparison of software generated energy production and consumption on Santiago Island during a) a summer week in 2022 and b) a summer week for 2030.



(a)



(b)

Fig. 4: Comparison of software generated energy production and consumption on Santiago Island during a) a winter week of 2022 and b) a winter week for 2030.

Throughout the year under analysis, the maximum production capacity of the thermal power plant would be 47 MW. This means that the originally used 72 MW thermal power [4] could be reduced up to 35%, resulting in a reduction in the need for fossil fuel imports and ownership costs.

An excess of production of the RES is observed, which could be caused by the oversizing of both thermal and RES power plants in the data extracted from the bulletin [4]. This output may suggest an excess of the installed power capacity of power plants based on RES. Avoiding excess energy production is crucial for ensuring the best use of energy resources, maximizing return on investment, effectively integrating renewable sources, protecting grid security, and preventing overvoltages and congestion, contributing to a more efficient, sustainable, and reliable energy system.

Tables 1 and 2 present a sensitivity analysis of the installed capacity of RES on the critical excess of electrical energy and the penetration of renewable energy in the system over the year. Table 1 details the impact of varying the photovoltaic capacity, while Table 2 focuses on wind power capacity. The analysis involves successive 5% reductions in the original installed capacity of each energy source.

Table 1: Sensitivity analysis of photovoltaic power installed capacity on the excess of electrical energy over the year.

Wind(MW)	Photovoltaic(MW)	Critical Excess(MWh/year)	Penetration of RES(%)
36.5	64	31.3	56.80
36.5	60.8	27.19	56.30
36.5	57.6	23.3	55.80
36.5	54.4	19.649	55.18
36.5	51.2	16.7	54.59

Table 2: Sensitivity analysis of wind power installed capacity on the excess of electrical energy over the year.

Wind(MW)	Photovoltaic(MW)	Critical Excess(MWh/year)	Penetration of RES(%)
36.50	64	31.30	56.80
34.68	64	29.59	55.64
32.85	64	27.93	54.50
31.03	64	26.32	53.36
29.20	64	24.75	52.19

It is observed that reducing the installed capacity of photovoltaic power has a higher impact on reducing overproduction with minimal impact on the RES penetration percentage. On the other hand, reducing wind power appears to have a less significant impact on overproduction.

Therefore, in terms of minimizing the impact on the percentage of renewable energy penetration, reducing photovoltaic power appears to be a more advantageous option than reducing wind power. This suggests that wind energy may have a more resilient or stable contribution to renewable penetration than photovoltaic energy, at least under the specific conditions of this study.

5 Conclusion

The analysis of the electrical energy sector on Santiago island, using the data foreseen for the year 2030, unveils significant challenges, notably the lack of complementarity among renewable energies and the excess production of photovoltaic power.

Owing to the absence of complementary resources, there are occasions when reliance on thermal power persists, underscoring the necessity for diversification in RES and the necessity of having a thermal energy source. Exploring storage solutions could greatly impact solving the lack of complementary resources and further reducing the dependence on thermal power.

Overdimensioning photovoltaic energy infrastructure can lead to unnecessary costs and resource wastage. To optimize efficiency and reliability, balanced integration strategies such as hybrid systems or energy storage are essential for a stable and sustainable energy supply. Cape Verde's abundant solar and wind resources provide significant potential for a sustainable energy future. A diversified, storage-integrated approach can enhance energy security, reduce costs, and contribute to a cleaner environment.

Excess photovoltaic energy could support storage solutions like green hydrogen production and battery banks (e.g., lithium-ion, lead-acid). Surplus energy can also be used for seawater desalination, addressing a critical need for the country. These alternative solutions will be explored in future research.

Acknowledgement.

The authors are grateful to the Foundation for Science and Technology (FCT, Portugal) for financial support through national funds FCT/MCTES (PIDDAC) to CeDRI, UIDB/05757/2020 (DOI: 10.54499/UIDB/05757/2020) and UIDP/05757/2020 (DOI: 10.54499/UIDP/05757/2020) and SusTEC, LA/P/0007/2020 (DOI: 10.54499/LA/P/0007/2020).

References

1. F. Riva, A. Tognollo, F. Gardumi, and E. Colombo, “Long-term energy planning and demand forecast in remote areas of developing countries: Classification of case studies and insights from a modelling perspective,” *Energy Strategy Reviews*, vol. 20, pp. 71–89, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2211467X18300099
2. D. V. Pombo, J. Martinez-Rico, and H. M. Marczinkowski, “Towards 100
3. A. Costa, “Relatório de base para cabo verde.” [Online]. Available: www.aler-renovaveis.org
4. imprensa nacional de Cabo Verde, “Boletim oficial i série número 40.” [Online]. Available: https://kiosk.incv.cv/V/2019/4/8/1.1.40.2752/
5. I. Westphal, “The effects of reducing renewable power intermittency through portfolio diversification,” *Renewable and Sustainable Energy Reviews*, vol. 197, p. 114415, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032124001382
6. P. Donk, S. Sterl, W. Thiery, and P. Willems, “Climate-combined energy modelling approach for power system planning towards optimized integration of renewables under potential climate change - the small island developing state perspective,” *Energy Policy*, vol. 177, p. 113526, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301421523001118
7. J. Tavares, M. Lopes, and F. Neto, “Climate and fundamentals of the energy offer in cape verde,” *Energy Reports*, vol. 6, pp. 370–377, 2020, the 6th International Conference on Energy and Environment Research - Energy and environment: challenges towards circular economy. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352484719306067
8. R. Segurado, G. Krajačić, N. Duić, and L. Alves, “Increasing the penetration of renewable energy resources in s. vicente, cape verde,” *Applied Energy*, vol. 88, no. 2, pp. 466–472, 2011, the 5th Dubrovnik Conference on Sustainable Development of Energy, Water and Environment Systems, held in Dubrovnik September/October 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261910002588
9. D. V. Pombo, J. Martinez-Rico, S. V. Spataru, H. W. Bindner, and P. E. Sørensen, “Decarbonizing energy islands with flexibility-enabling planning: The case of santiago, cape verde,” *Renewable and Sustainable Energy Reviews*, vol. 176, p. 113151, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032123000072
10. Lund, “EnergyPLAN - Advanced Energy Systems Analysis Computer Model,” jun 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6602938
11. M. G. Prina, M. Lionetti, G. Manzolini, W. Sparber, and D. Moser, “Transition pathways optimization methodology through energyplan software for long-term energy planning,” *Applied Energy*, vol. 235, pp. 356–368, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261918316672
12. D. V. Pombo, “Impact of the covid-19 pandemic on energy islands: The case of cape verde,” *Electric Power Systems Research*, vol. 213, p. 108611, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779622006885
13. Electra, “Relatório e contas 2019,” 2021, in Portuguese. [Online]. Available: https://www.electra.cv/index.php/2014-05-20-16-30-16/relatorios-nort
14. ELECTRA, “Relatorio e contas 2022.” [Online]. Available: https://www.electra.cv/frontend/web/index.php?r=site%2Frelatorio&type=2&inst=1
15. D. Electra, “Relatório e contas 2019 (in portuguese),” 2021.
16. Electra, “Management report,” Accessed on April 7, 2024, from http://www.electra.cv/, 2011–2017, cape Verde.
17. I. e. E. Ministério do Turismo, “Cape verde 50% renewable a roadmap to 2020.” [Online]. Available: https://gestoenergy.com/wp-content/uploads/2018/04/MASTER-PLAN-CAPE-VERDE-50-RENEWBALE.pdf
18. “Renewables.ninja,” https://www.renewables.ninja/, accessed: 2024-05-11.
19. D. Pombo, “Cape Verde Reference System v2.0,” 2 2022. [Online]. Available: https://figshare.com/articles/dataset/Cape_Verde_Reference_System_v2_0/17430413

Sponsors

– Fundação para a Ciência e a Tecnologia, FCT

Fundação
para a Ciência
e a Tecnologia

– Projeto STEP (project n.: 101078933, has been funded by the European Commission)

STEM Research
Equality, Diversity and Inclusion