SASYR Symposium of Applied Science for Young Researchers

5th Symposium of Applied Science for Young Researchers

BOOK OF ABSTRACTS 2025

July 2 , 2025

5th Symposium
of
Applied Science for Young Researchers

Book of Abstracts

SASYR 2025

2 July 2025

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
Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development (GECAD)
Instituto Politécnico do Porto

Silvestre Malta
Applied Digital Transformation Laboratory (ADiT-LAB)
Instituto Politécnico de Viana do Castelo

Instituto Politécnico de Bragança – 2025
Campus de Santa Apolónia
5300-253 Bragança - Portugal
ISBN: 978-972-745-361-0

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

Welcome

These are the abstracts of the 5th Symposium of Applied Science for Young Researchers – SASYR. This scientific event welcomed scientific works on the topics covered by the following four research centers:

– CeDRI (from IPB, Instituto Politécnico de Bragança)
– 2Ai (from IPCA, Instituto Politécnico do Cávado e do Ave)
– GECAD (from IPP, Instituto Politécnico do Porto)
– ADiT-lab (from IPVC, Instituto Politécnico de Viana do Castelo)

The primary objective of SASYR 2025 is to create a welcoming and relaxed environment for young researchers to present their work, discuss recent findings, and explore new ideas. In this way, this event offers an opportunity for the CeDRI, 2Ai, GECAD, and ADiT-lab research communities to leverage synergies and promote collaborations, thereby enhancing the quality of their research.

The SASYR 2025 took place at Instituto Politécnico de Viana do Castelo, Viana do Castelo, Portugal, on 2 July 2025.

The SASYR 2025 Organizing Committee,
Florbela P. Fernandes, Helena Torres, Pedro Pinto, and Silvestre Malta

Committees

Organizing Committee

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

Advisory Committee

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

Scientific Committee

Ana Almeida, Instituto Politécnico do Porto
Ana Faria, Instituto Politécnico do Porto
Ana I. Pereira, Instituto Politécnico de Bragança
Ana Madureira, Instituto Politécnico do Porto
Ana Teixeira, Instituto Politécnico do Porto
Ângela Ferreira, Instituto Politécnico de Bragança
António Miguel Rosado da Cruz, Instituto Politécnico de Viana do Castelo
Beatriz Flamia Azevedo, Instituto Politécnico de Bragança
Bertil P. Marques, Instituto Politécnico do Porto
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
Constantino Martins, Instituto Politécnico do Porto
Daniel Miranda, Instituto Politécnico do Cávado e do Ave
Diogo Martinho, Instituto Politécnico do Porto
Duarte Duque, Instituto Politécnico do Cávado e do Ave
Estela Vilhena, Instituto Politécnico do Cávado e do Ave
Eva Maia, Instituto Politécnico do Porto
Eva Oliveira, Instituto Politécnico do Cávado e do Ave

Fernando Lezama, Instituto Politécnico do Porto
Fernando Monteiro, Instituto Politécnico de Bragança
Fernando Veloso, Instituto Politécnico do Cávado e do Ave
Flávia Pires, Instituto Politécnico de Bragança
Florbela P. Fernandes, Instituto Politécnico de Bragança
Francisco Pereira, Instituto Politécnico do Porto
Gil Lopes, Instituto Politécnico do Porto
Helena Sofia Rodrigues, Instituto Politécnico de Viana do Castelo
Helena Torres, Instituto Politécnico do Cávado e do Ave
Inês Sena, Instituto Politécnico de Bragança
Isabel S. Jesus, Instituto Politécnico do Porto
Ivone Amorim, Instituto Politécnico do Porto
João Borges, Instituto Politécnico do Cávado e do Ave
João Braun, Instituto Politécnico de Bragança
João Carlos Silva, Instituto Politécnico do Cávado e do Ave
João Coelho, Instituto Politécnico de Bragança
João Mendes, Instituto Politécnico de Bragança
João Paulo Teixeira, Instituto Politécnico de Bragança
Joaquim Gonçalves, Instituto Politécnico do Cávado e do Ave
Joaquim Santos, Instituto Politécnico do Porto
Jorge Esparteiro Garcia, Instituto Politécnico de Viana do Castelo
Jorge Ribeiro, Instituto Politécnico de Viana do Castelo
José Henrique Brito, Instituto Politécnico do Cávado e do Ave
José Lima, Instituto Politécnico de Bragança
Luís Conceição, Instituto Politécnico do Porto
Luís Ferreira, Instituto Politécnico do Cávado e do Ave
Luis Gomes, Instituto Politécnico do Porto
Luiz Faria, Instituto Politécnico do Porto
Maria F. Pacheco, Instituto Politécnico de Bragança
Maria João Varanda Pereira, Instituto Politécnico de Bragança
Marílio Cardoso, Instituto Politécnico do Porto
Martinha Pereira, Instituto Politécnico do Cávado e do Ave
Murillo Ferreira Dos Santos, Instituto Politécnico de Bragança
Natália Maria De Bessa Rego, Instituto Politécnico do Cávado e do Ave
Nuno Pereira, Instituto Politécnico do Porto
Orlando Sousa, Instituto Politécnico do Porto
Óscar R. Ribeiro, Instituto Politécnico do Cávado e do Ave
Patricia Leite, Instituto Politécnico do Cávado e do Ave
Paula Alexandra Rego, Instituto Politécnico de Viana do Castelo
Paulo Alves, Instituto Politécnico de Bragança
Paulo Costa, Instituto Politécnico de Viana do Castelo
Paulo Matos, Instituto Politécnico de Bragança
Paulo S. Matos, Instituto Politécnico do Porto
Pedro Coutinho, Instituto Politécnico de Viana do Castelo

Pedro Faria, Instituto Politécnico de Viana do Castelo
Pedro Filipe Oliveira, Instituto Politécnico de Bragança
Pedro Morais, Instituto Politécnico do Cávado e do Ave
Pedro Pinto, Instituto Politécnico do Porto
Pedro Rodrigues, Instituto Politécnico de Bragança
Ramiro Barbosa, Instituto Politécnico do Porto
Ricardo Almeida, Instituto Politécnico do Porto
Ricardo Severino, Instituto Politécnico do Porto
Rogério Tavares, Instituto Politécnico de Bragança
Rui Pedro Lopes, Instituto Politécnico de Bragança
Sara Baltazar, Instituto Politécnico de Viana do Castelo
Sara Cruz, Instituto Politécnico do Cávado e do Ave
Sara Paiva, Instituto Politécnico de Viana do Castelo
Silvestre Malta, Instituto Politécnico de Viana do Castelo
Teresa Abreu, Instituto Politécnico do Cávado e do Ave
Tiago Franco, Instituto Politécnico de Bragança
Tiago Pedrosa, Instituto Politécnico de Bragança
Vítor Carvalho, Instituto Politécnico do Cávado e do Ave

Local Organizing Committee

Helena Sofia Rodrigues, Instituto Politécnico de Viana do Castelo
Jorge Garcia, Instituto Politécnico de Viana do Castelo
Sara Baltazar, Instituto Politécnico de Viana do Castelo
Silvestre Malta, Instituto Politécnico de Viana do Castelo

Secretariat / Technical Support

Beatriz Cabrera, Instituto Politécnico de Bragança
Carla Fontes, Instituto Politécnico de Bragança
Clarisse Pais, Instituto Politécnico de Bragança
Evandro Alves, Instituto Politécnico de Bragança
Francisco Dias, Instituto Politécnico de Viana do Castelo
Gabriel Brandão, Instituto Politécnico de Viana do Castelo
Ivo Ramoa, Instituto Politécnico de Viana do Castelo
Luís Filipe Dias, Instituto Politécnico de Viana do Castelo
Luís Pinto, Instituto Politécnico de Viana do Castelo
Mario Gomes, Instituto Politécnico de Viana do Castelo
Natália Santos, Instituto Politécnico do Cávado e do Ave
Rodrigo Machado, Instituto Politécnico de Viana do Castelo
Rui Filipe Silva, Instituto Politécnico de Viana do Castelo
Tiago Baptista, Instituto Politécnico de Viana do Castelo

Table of Contents

Personalized Feedback Through a Digital Application: Supporting the Recovery of Unacquired Learning in Basic Education

José Mário Baptista[1] and Duarte Duque[1,2]

[1] Instituto Politécnico do Cávado e do Ave, IPCA, Portugal
[2] Lab. Associado de Sistemas Inteligentes, Portugal
{a33773,dduque}@ipca.pt

Abstract. This paper presents an online application that identifies and records learning gaps, providing automated and personalized feedback to support learning recovery in basic education. Grounded in Design Science Research (DSR), the tool translates assessment data into actionable insights for teachers and personalized guidance for students. The system architecture integrates curriculum-aligned objectives, minimum performance thresholds, and feedback generation modules. Evaluation will involve qualitative (Think Aloud) and quantitative (System Usability Scale) methods. Results will inform future enhancements, including AI-based recommendation of open educational resources. This initiative promotes inclusive, adaptive, and data-informed education for all.

Keywords: Personalized Feedback · Learning Recovery · Educational Technology · Basic Education · Learning Analytics

1 Introduction

The integration of learning analytics into education is increasingly seen as a key factor in improving pedagogical practices. By enabling a detailed analysis of students' progress, these technologies facilitate early identification of learning gaps and support timely intervention. This regard, personalized feedback has become central to such strategies, where digital tools can bridge gaps in equity and access.

Education faces serious challenges regarding learning loss, commonly understood as the decline or absence of expected knowledge and skills, often identified through assessment data, and the demand for differentiated learning. Personalized feedback, when adapted to individual needs, has been shown to significantly improve learning outcomes [5]. Automated systems for feedback and learning analysis can improve intervention efficiency by up to 80% [1].

The project responds to these needs through a application that analyzes students' assessment data and provides tailored feedback, aligning with Portugal's national strategy "Plano 23–24 Escola+" [2], which prioritizes innovative and inclusive practices to recover unacquired learning.

2 Objectives and Methodology

The application is designed to streamline data processing for teachers and provide timely support to students, aimings to minimize the time between assessment and feedback, ensuring that interventions are faster and more impactful [1]. This approach aligns with pedagogical best practices, where prompt, specific feedback contributes to deeper understanding and improved learner autonomy.

To minimize the gap between assessment and intervention, the platform will streamline data processing and feedback delivery by providing individualized, actionable feedback based on written assessments, along with structured data to support instructional adjustmentswhile, at the same time reducing teacher workload through automation.

2.1 Methodological Framework

The project uses the Design Science Research (DSR) methodology [3, 4], particularly the Educational Design Research (EDR) framework. This methodology follows three iterative phases: analysis and exploration to define pedagogical and system requirements; design and construction of a modular system for data input, learning gap detection, and feedback generation; and evaluation and reflection through Think Aloud and SUS protocols to validate usability and relevance.

Iterative Development - Each development cycle includes formative feedback loops, enabling the system to evolve in response to pilot testing and classroom realities.

Functional Architecture - The workflow begins with inputting class/student data and curriculum-based Learning Objectives (LOs), linked via a matrix to test items. After grading, the system compares scores to a Minimum Performance Level (MPL) and LOs below MPL trigger feedback generation and summary reports. Outputs for teachers include detailed statistics that inform instruction and support decision-making. To ensure ethical compliance and protect privacy, the system operates without using real student data. The system is built on a Python/Django stack, which supports both frontend and backend development within an integrated environment. Django's MVT pattern and built-in ORM ensure seamless interaction with a PostgreSQL database, allowing efficient data flow between assessment input, learning gap detection, and personalized feedback generation.

Evaluation and Reflection - Validation using Think Aloud and SUS ensure usability and relevance while Think Aloud protocol will involve 6 to 8 participants (teachers and educational technology experts). All sessions will occur in a controlled school setting, where users will verbalize their thoughts while interacting with the prototype. Recordings will be transcribed and thematically analysed to identify usability issues and guide iterative improvements. Participation will be voluntary and anonymized, with informed consent to ensure confidentiality. Following this, the System Usability Scale (SUS) will be applied according to the standardized protocol, providing quantitative usability metrics to complement the qualitative findings.

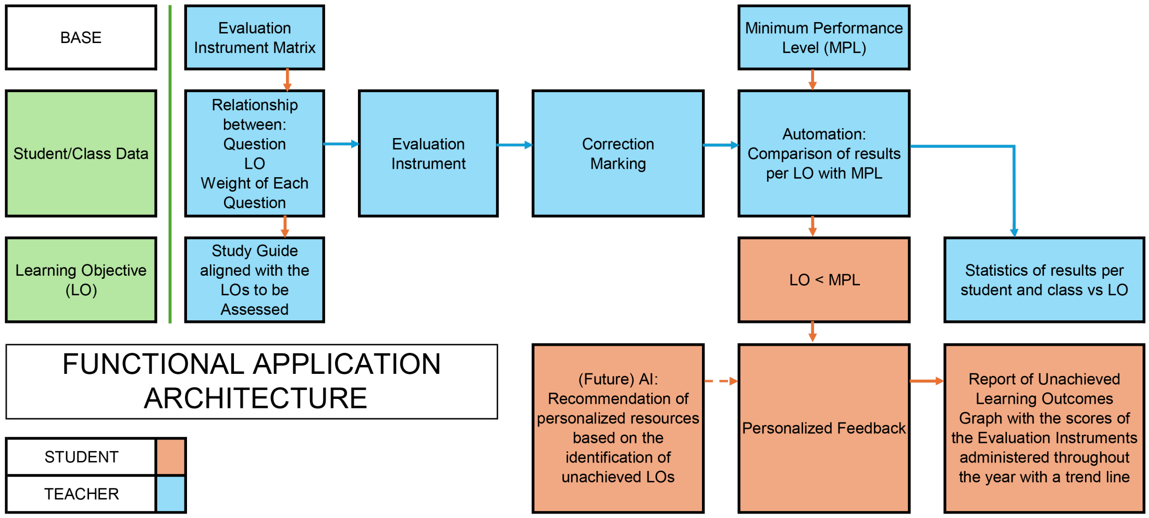


Fig. 1: Functional architecture of the personalized feedback application.

3 Results and Future Work

A functional proof of concept will be tested with target users, and expected outcomes include high usability and positive teacher feedback, particularly regarding the system's ability to identify unacquired learning.

Subsequent phases will integrate AI modules for resource recommendation, enhancing learner autonomy and promoting adaptive instruction.

References

1. Bernacki, M.L., Greene, M.J., Lobczowski, N.G.: A systematic review of research on personalized learning. Educational Psychology Review **33**(4), 1675–1715 (2021). https://doi.org/10.1007/s10648-021-09615-8
2. Conselho de Ministros: Resolução de conselho de ministros n.º 80-b/2023, de 18 de julho (2023), diário da República n.º 136/2023, 2.º Suplemento, Série I
3. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Quarterly **28**(1), 75–105 (2004)
4. McKenney, S., Reeves, T.: Educational design research. In: Handbook of Research on Educational Communications and Technology, pp. 131–140. Springer (2013). https://doi.org/10.1007/978-1-4614-3185-5_11
5. Rodrigues, L.e.a.: Mathaide: A qualitative study of teachers' perceptions of an its unplugged for underserved regions. International Journal of Artificial Intelligence in Education (2024). https://doi.org/10.1007/s40593-024-00397-y

A FIWARE-Based Digital Twin Architecture for Expedite IIoT Integration: the Adalberto Textile Solutions Use Case

Gonçalo Nuno Esperança Quesado[1], João Pedroso[3], and Sérgio Ivan Lopes[1,2]

[1] ADiT-Lab, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
[2] CiTin - Centro de Interface Tecnológico Industrial, Inovarcos, Arcos de Valdevez, 4970-786, Portugal
[3] Adalberto Textile Solutions, Rebordões, 4795-177

Abstract. This paper presents the transition of Adalberto Textile Solutions's monitoring system from a Home Assistant-based approach to a modular and scalable architecture built on the FIWARE framework. The proposed solution integrates Industrial IoT (IIoT) data using a custom-developed MQTT IoT Agent, which translates machine sensor data into NGSI format for processing in the Orion Context Broker. This enables real-time monitoring, improves scalability, and lays the base for future Digital Twin capabilities. Although features such as predictive analytics and ERP integration are planned, the current implementation already supports contextual data management, flexible sensor integration, and real-time visualization using Grafana. The work demonstrates how FIWARE components may be efficiently customized to support IIoT data pipelines in industrial environments

Keywords: FIWARE · Digital Twin · Industrial IoT · Smart Manufacturing · Textile Industry

1 Introduction

The textile and clothing sector is facing a digital revolution driven by the adoption of Industrial Internet of Things (IIoT) technologies [1]. The change aims to improve productivity, quality control, and responsiveness in dynamic production environments. Digital Twin (DT) architectures are becoming enablers of the change, with real-time representations of industrial systems [2]. FIWARE, an open-source platform, has gained traction to create scalable and interoperable DTs by enabling contextual data management and integration [3, 4].

Adalberto Textile Solutions, a Portuguese company in the textile and clothing sector, currently employs Home Assistant to monitor processes and transmit data to its ERP system and Grafana dashboard via MQTT. While this setup is functional, it lacks the scalability and flexibility required for system expansion and detailed monitoring. In particular, the inability to track work-in-progress or seamlessly integrate new sensors constrains operational visibility and responsiveness.

2 Objectives and Motivation

The main focus of this project is to overcome the limitation of the present system and implement a scalable Digital Twin model based on FIWARE. Specific goals are as follows:

- Replace Home Assistant with a FIWARE-compatible architecture;

– Develop a custom IoT Agent to subscribe to MQTT and convert data into NGSI format;
– Integrate sensor data with Orion Context Broker for centralized contextual management;
– Enable real-time visualization in Grafana and integration with Adalberto's ERP system.

3 Methodology and Implementation

The proposed architecture shown in Fig. 1 follows a modular and incremental development approach, inspired by the FIWARE-based TEXP@CT reference model. It integrates multiple components to enable the collection, processing, and contextualization of IIoT data from textile machinery:

– **Custom IoT Agent:** Developed in Python, this agent subscribes to existing MQTT topics and parses machine data in the NGSI v2 format.
– **Orion Context Broker:** Acts as the central aggregator of contextual data, supporting real-time updates and historical tracking.
– **MongoDB:** Serves as the internal storage backend for the Orion Context Broker.
– **Custom Agent:** A second agent was developed to subscribe to relevant context changes from Orion, extract selected attributes from entities, and structure them into time-series format for storage in InfluxDB.
– **InfluxDB + Grafana:** Through a custom agent, data is streamed into InfluxDB for visualization on customized Grafana dashboards for operational needs.

The first deployment imports data from three types of sensors (energy, gas, and distance), which are each represented by different entities in FIWARE. The topics are structured as `bp/sensorType/sensorID/.` so that the sensor identity can be automatically parsed and attributes mapped. The scheme ensures that a new sensor with the same topic structure can automatically be identified, mapped, and integrated into the architecture without additional configuration, allowing for scalability and flexibility when introducing more components into the system.

4 Preliminary Results and Discussion

Initial results show clear improvements in modularity, scalability, and visibility of machine performance. The custom IoT Agent reliably processes and routes MQTT data to Orion, with Grafana providing near real-time dashboards to production teams.

The FIWARE-based system runs in parallel with the existing Home Assistant infrastructure, facilitating direct comparison under live conditions. Data stored in InfluxDB is being used by external IT teams to develop strategic dashboards, supporting broader decision-making. Adalberto has already confirmed gains in system adaptability and production tracking.

Although the current deployment does not yet support behavioral simulation or predictive logic, it delivers a key layer of contextual data integration and real-time monitoring – foundational steps toward a full Digital Twin. Future work will evolve from this baseline.

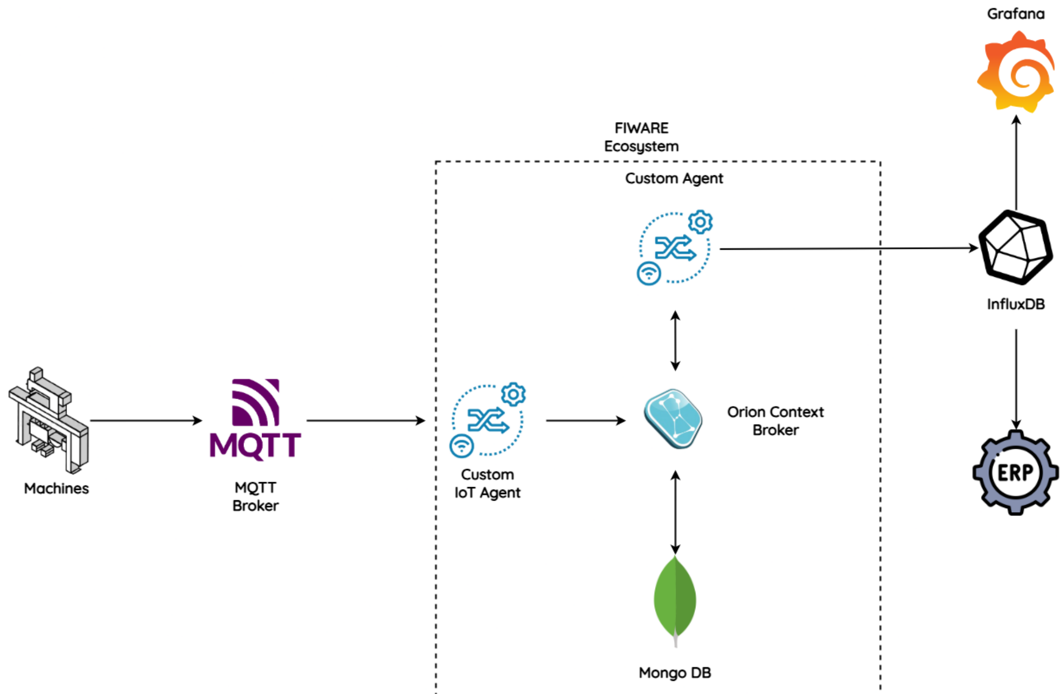


Fig. 1: Proposed architecture integrating MQTT data with FIWARE components and visualization systems.

5 Conclusion and Future Work

The ongoing use of this FIWARE-based architecture shows the industrial benefits of open and interoperable systems in practice. Performing in parallel with the current system, it has been established to be capable of centralizing IIoT data and expanding monitoring capabilities on a scalable and flexible basis.

Future activity includes expanding the number of machines connected, improving the Digital Twin logic to support more advanced behavioral modeling, and deploying predictive analytics mechanisms. A second project will be undertaken to develop a visual Digital Twin of the factory floor, with each machine represented graphically in a plant layout. In this deployment, operators will also be able to interact with the factory model, the individual machines for real-time information, historical behavior, and state of health, so that they have better situational awareness and can inform their decision-making on the shop floor.

ACKNOWLEDGMENTS: This work was carried out within the project "WP 2.2 (PPS#8): Digital Twin Platform textile Manufacturing", as part of the "Agenda: TEXP@CT: Pacto de Inovação para a Digitalização do STV", funded under Recovery and Resilience Plan - 02-C05i01.01-2022.PC644915249-00000025. This work has been partially supported through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project: UID/06121/2023.

References

1. Dal Forno, A. J., Bataglini, W. V., Steffens, F., & Ulson de Souza, A. A.: Industry 4.0 in textile and apparel sector: A systematic literature review. *Research Journal of Textile and Apparel*, 27(1), 95-117 (2023). https://doi.org/10.1108/RJTA-08-2021-0106
2. Deloitte: Industry 4.0 and the digital twin technology. *Deloitte Insights* (2021). https://www2.deloitte.com/us/en/insights/focus/industry-4-0/digital-twin-technology-smart-factory.html
3. FIWARE Foundation: FIWARE: Open Source platform for our smart digital future. https://www.fiware.org
4. Zyrianoff, I. D., Heideker, A., Silva, D., & Kamienski, C.: A FIWARE-based IoT platform for enabling digital twins in a greenfield smart factory. *Sensors*, (2020). https://www.researchgate.net/publication/354510938
5. Kamienski, C., Soininen, J. P., Taumberger, M., Dantas, R., & Toscano, A.: Towards efficient industrial IoT data management using FIWARE. *International Journal of Distributed Sensor Networks*, 15(6) (2019)

Adoption of a Fiware-based architecture to estimate material consumption by production order using Machine Learning

Vítor Amorim[1] and António Miguel Cruz[1,2]

[1] ADiT-Lab, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
vitoram@ipvc.pt
[2] Algoritmi Research Centre, Universidade do Minho, Guimarães, Portugal
miguel.cruz@estg.ipvc.pt

Abstract. The adoption of data-driven architectures in industry has seen significant advances in the efficiency and accuracy of production processes. In the Textile and Clothing sector, the implementation of solutions that integrate technologies is obviously very relevant. In this part of the industry, there is a major problem with the traceability of each company's manufacturing orders, which hinders some processes. Traceability in the textile supply chain is a recognized challenge, with direct implications for sustainability and compliance with regulations.

Keywords: Fiware · Machine Learning · Industry · Predictive analytics

1 Introduction and Objectives

Despite the relevance of traceability to ensure ethical and sustainable practices, the textile chain faces significant challenges, especially related to the collection, integration and analysis of large volumes of data. Accurately estimating material consumption per manufacturing order is essential for companies looking to minimize waste and optimize production.

This study is part of this context and is being carried out in partnership with the TextP@ct project, together with a company dedicated to the production of socks. The company faces difficulties in the initial phase of the production process, related to managing the quantities of each component and material used in each Manufacturing Order. To solve this problem, a Machine Learning (ML) model will be implemented to predict the precise consumption of each material in each Manufacturing Order, allowing for a more efficient allocation of resources.

The use of technologies such as FIWARE, an open-source platform that facilitates the development of digital applications, combined with ML, offers a promising approach to addressing these gaps. FIWARE can be integrated into industrial architectures to improve real-time data collection and analysis, while ML enables the development of robust predictive models. With this, the work seeks to contribute to greater efficiency, sustainability and competitiveness in the sector [1].

The main objectives are:

- To design and implement an architecture that supports real-time data ingestion, processing, and prediction using FIWARE.
- To develop and evaluate regression models for estimating material consumption based on historical production data.
- To benchmark model performance against traditional baseline methods and demonstrate improvements in forecasting accuracy.

2 Methodology

For the description of the methodology in this article, the Fig. 1 shows a diagram of the design science research methodology.

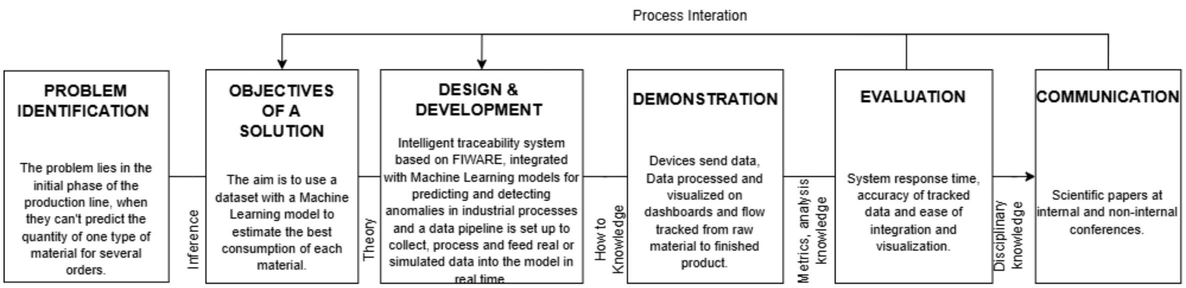


Fig. 1: Design science research

3 Discussion and Results

This project is currently under development. At an early stage, some studies are being carried out on various types of Machine Learning models, in order to understand and test

We will assess performance using some standard regression metrics:

- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)
- Coefficient of Determination (R^2)

This theoretical evaluation will later be complemented with actual results, enabling us to validate forecasting accuracy in real scenarios, retrain models periodically as new data becomes available and explore more complex architectures once sufficient historical data exists.

Below is a representation of how the dataset is structured, both for training, represented in Table 1, and prediction represented in Table 2.

Table 1: Solution — Training Dataset.

Material	Consumed Quantity	Unit	Order Manufacture	Produced Quantity	Production Unit	Type of Final Product
MaterialType_1	1000	Kg	XPTO1234	750	Kg	ThreadType_1
MaterialType_1	2700	Kg	XPTO12345,	2400	Kg	ThreadType_1
ThreadType_1	500	Kg	ZX1234	250	pairs	SockModel X

Table 2: Solution — Prediction.

Material	Quantity	Order	N pairs Output	Type of final Product	Output - Prediction
MaterialType_1	Quantity_X	Order_1, Order_2, Order_3 ...	580 750 1000	SockModel X SockModel X SockModel X	MaterialQuantity x1 MaterialQuantity x2 MaterialQuantity x3

4 Conclusion

The application of a FIWARE-based architecture combined with Machine Learning models has proved to be a promising approach to tackling the challenges of traceability and forecasting material consumption in the textile industry. This integration makes it possible to optimize the resources used in each order, promoting greater efficiency and sustainability. Although the project is still in the development phase, the first results point to a high potential for applicability in real industrial contexts.

References

1. Agrawal, T.K., Pal, R.: Traceability in Textile and Clothing Supply Chains: Classifying Implementation Factors and Information Sets via Delphi Study. Sustainability **11**(6), 1698 (2019). https://doi.org/https://doi.org/10.3390/su11061698

Design and Development of a FIWARE-based Digital Twin for Predictive Maintenance in the Textile and Clothing Sector

Tiago Lemos[1], Ricardo Severino[2], and Sérgio Ivan Lopes[1,3,4]

[1] ADiT-Lab, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
tiagolemos@ipvc.pt
[2] ISEP, Instituto Politécnico do Porto & INESC TEC Porto, 4200-465 Porto, Portugal
rar@isep.ipp.pt
[3] CiTin - Centro de Inteface Tecnológico Industrial, Inovarcos, 4970-786 Arcos de Valdevez, Portugal
sil@estg.ipvc.pt
[4] IT- Instituto de Telecomunicações, Aveiro, Portugal

Abstract. Competitiveness in the Textile & Clothing sector remains hindered by unplanned downtime and fragmented data flows. An open-standard, FIWARE-based digital twin has therefore been developed, streaming telemetry from legacy and new MTEX NS machines into an Orion Context Broker via a customised IoT-Agent. Early shop-floor tests have demonstrated sub-millisecond latency, stable throughput and more intelligible alarm codes, enabling faster operator reaction. The three-phase roadmap – connectivity, security, and predictive maintenance – is currently in Phase I and is projected to reduce downtime by 30% and maintenance costs by 20%. These results confirm FIWARE's potential for vendor-neutral predictive maintenance.

Keywords: Digital Twin · Predictive Maintenance · FIWARE · IIoT

1 Introduction

The Textile & Clothing (T&C) sector faces volatile demand, stricter sustainability goals, and costly unplanned downtime, all of which erode competitiveness. Data-driven Business-Intelligence systems can mitigate some of these pressures [1], yet open research questions remain around the enabling technologies for Digital Twins (DTs) in manufacturing [3].

FIWARE, an open-source Next Generation Service Interfaces (NGSI) ecosystem, offers vendor-neutral IoT Agents, the Orion Context Broker and modular security that harmonizes heterogeneous shop-floor assets [2]; it has already proven effective for predictive maintenance in flexible manufacturing pilots [5].

FIWARE is applied to MTEX NS through a three-phase architecture: Phase I connectivity has been prototyped on the TPPPS4 conductive-ink printer, and latency as well as alarm readability have been benchmarked. The architecture, implementation and preliminary results are detailed in accordance with the six-activity Design-Science Research Process (DSRP) [4], thereby preparing upcoming phases focused on role-based security and fault prediction.

2 Methodology

The project pursues six goals: (i) integrate heterogeneous shop-floor assets via FIWARE; (ii) develop an IoT Agent that maps Message Queuing Telemetry

Transport (MQTT) topics to NGSI-Linked Data (NGSI-LD) context; (iii) define a Smart Data Model for textile equipment; (iv) implement a DT that renders real-time state and anticipate failures; (v) deploy a predictive-maintenance pipeline and quantify benefits; and (vi) validate the solution in the MTEX NS environment.

Following the DSRP, six activities were carried out: (i) **Problem identification & motivation:** a review of MTEX NS alarm logs, maintenance strategy, and system architecture revealed non-intuitive numeric codes that must be looked up by operators, prolonging line stops. **(ii) Solution objectives:** legacy and new equipment are harmonised using vendor-neutral FIWARE components, shop-floor context exposed as NGSI-LD entities, and alarm intelligibility improved to enable predictive maintenance; **(iii) Design & development:** a three-phase FIWARE architecture stack (see Fig. 1) - Phase I connectivity (custom IoT-Agent + Orion), Phase II role-based security (Keyrock/Policy Enforcement Point Proxy (PEP)/AuthZForce) and Phase III predictive analytics (AI/ML forecaster) - containerised with Docker-Compose for staging. **(iv) Demonstration:** the connectivity prototype streams machine states and human-readable alarms from the TPPPS4 printer and simulators in real time. **(v) Evaluation:** upcoming shop-floor trials will benchmark downtime, repair time and operator-alert comprehension against historical baselines. **(vi) Communication:** updates are shared continuously with MTEX NS personnel and validated on the shop floor.

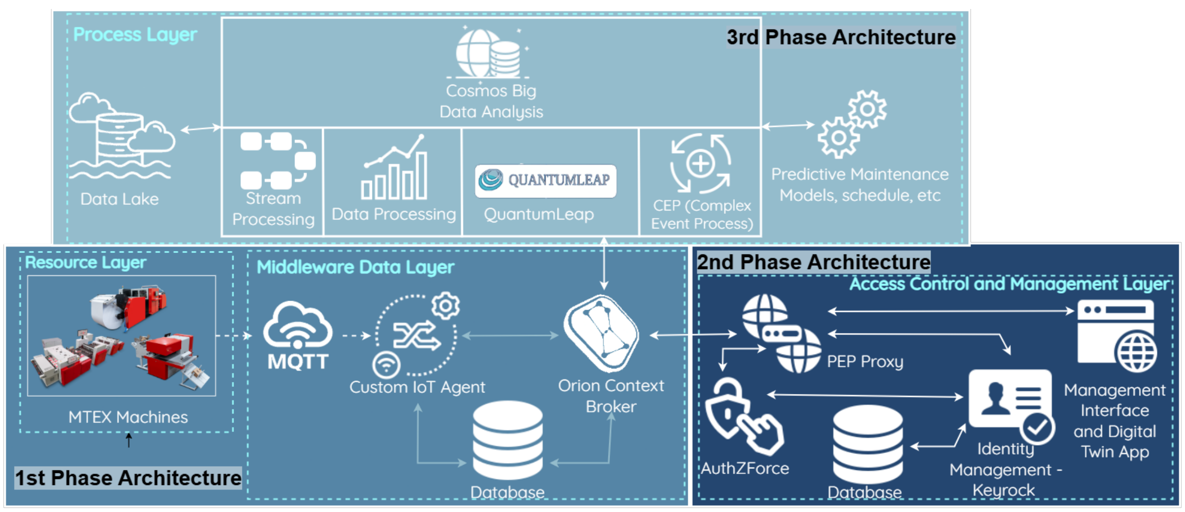


Fig. 1: Proposed FIWARE-based Digital Twin Architecture for MTEX NS

3 Ongoing Implementation

Phase I is operational: a custom IoT-Agent ingests TPPPS4 and simulated devices telemetry, translating the original *deviceId/state/attribute* MQTT topics to ≈ 30 NGSI-v2 entities stored in Orion and refreshed every few seconds. Although NGSI-v2 is still used for simplicity, the agent already accepts NGSI-LD payloads, easing future semantic enrichment.

Early shop-floor tests have exhibited steady ingestion of thousands of messages with end-to-end latency below 1 ms and no loss; mapping raw numeric status codes to descriptive NGSI attributes (e.g., “Warning: Uninitialised”) has reduced operator look-up time. On the staging server, role-based tokens are issued by the security stack, whose login features are in final testing for Phase II security, while an AutoRegressive Integrated Moving Average (ARIMA) model is being trialed as the analytics core of Phase III predictive maintenance.

4 Conclusions and Future Work

The project already demonstrates a viable FIWARE Digital Twin: Phase I delivers millisecond-level telemetry and real-time alarms; Phase II login trials confirm secure role-based access; and early tests with an ARIMA model show that Phase III fault prediction is technically feasible. Ongoing work will finish NGSI-LD migration, harden permissions, extend histories, and benchmark end-to-end KPIs toward 30% downtime and 20% maintenance-cost reductions.

ACKNOWLEDGMENTS: This work was carried out within the project "WP 2.2 (PPS#8): Digital Twin Platform for Textile Manufacturing", as part of the "Agenda: TEXP@CT: Pacto de Inovação para a Digitalização do STV", funded under Recovery and Resilience Plan - 02-C05i01.01-2022.PC644915249-00000025. This work has been partially supported through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project: UID/06121/2023.

References

1. Ahmad, S., Miskon, S., Alabdan, R., Tlili, I.: Towards sustainable textile and apparel industry: Exploring the role of business intelligence systems in the era of industry 4.0. Sustainability **12** (2020). https://doi.org/10.3390/su12072632
2. De Panfilis, S., Gusmeroli, S., Rodriguez, J., Benedicto, J.: FIWARE for Industry: A Data-driven Reference Architecture, chap. 21, pp. 171–178. John Wiley & Sons, Ltd (2018). https://doi.org/https://doi.org/10.1002/9781119564034.ch21, `https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119564034.ch21`
3. Fuller, A., Fan, Z., Day, C., Barlow, C.: Digital twin: Enabling technologies, challenges and open research. IEEE Access **8** (2020). https://doi.org/10.1109/ACCESS.2020.2998358
4. Peffers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S.: A design science research methodology for information systems research. Journal of Management Information Systems **24**, 45–77 (01 2007)
5. Sang, G.M., Xu, L., de Vrieze, P.T., Bai, Y.: Towards predictive maintenance for flexible manufacturing using fiware. Lecture Notes in Business Information Processing (CAiSE 2020 Workshops) pp. 17–28 (2020). https://doi.org/10.1007/978-3-030-49165-9_2

Learning-Based Autonomous Handling of Multi-Articulated Conveyor Trolleys

Joaquin Dillen[1], João L. Vilaça[1], João Borges[1], John Cosgrove[2], and António H. J. Moreira[1]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
{jdillen, jvilaca, jpbsilva, amoreira}@ipca.pt
[2] Technological University of the Shannon: Midlands Midwest, TUS, Ireland
john.cosgrove@tus.ie

Abstract. This project aims to develop a robotic framework for the autonomous transportation of non-rigid object convoys, a task commonly encountered in logistics and service environments. These convoys, composed of stacked or loosely coupled units, present substantial challenges due to variable traction, non-rigid articulation, and the absence of unified dynamic models. The proposed solution will address the limitations of current human-operated methods by designing a scalable robotic system capable of managing such convoys safely and efficiently. Development will rely on high-fidelity simulation platforms, including NVIDIA Isaac Sim, and virtualization tools to iteratively train and evaluate control strategies. Imitation learning will be used to replicate expert human behavior, enabling adaptive and context-aware motion planning. The project further aims to generalize the learned behavior across similar scenarios by developing a robust, transferable foundation model. The expected contribution is a validated robotic solution for convoy handling with real-world applicability in complex, unstructured environments.

Keywords: Autonomous Robotics, Non-Rigid Object Manipulation, Imitation Learning, Multi-Robot Systems, Simulation-Based Training, Convoy Transportation.

1 Introduction

Autonomous transportation of loosely connected object convoys such as multi-articulated conveyor trolleys (MACTs) remains a complex challenge in robotics. Unlike rigid systems, MACTs present nonholonomic constraints, variable structural dynamics, and unaligned motion vectors, complicating perception, planning, and control [1, 9].

This work proposes a unified, learning-based framework for autonomous MACT handling in both simulated and real environments. The system leverages multi-modal perception, imitation learning, and adaptive behavior generation to enable robust operation of loosely coupled convoys across structured and unstructured contexts. Prior research on robotic trolley collection has focused on constrained environments like airports and retail stores, often relying on predefined trajectories and handcrafted control policies [7, 8]. These solutions struggle with generalization.

Advanced control strategies using mobile manipulators and decentralized planning have improved performance in dynamic tasks but generally assume rigid couplings that do not reflect MACT dynamics [1, 4].

Learning-based approaches particularly imitation and reinforcement learning offer a model-free alternative for capturing expert behavior in complex systems [3].

Complementary advances in 3D perception, such as partial-to-full registration and sensor fusion, support localization in cluttered environments [5].

However, sim-to-real transfer remains an open problem due to sensory noise, actuation inaccuracies, and domain gaps [2, 6]. While progress has been made in individual subdomains, a comprehensive framework that integrates perception, learning, and control for MACT systems is yet to be established.

2 Methodology

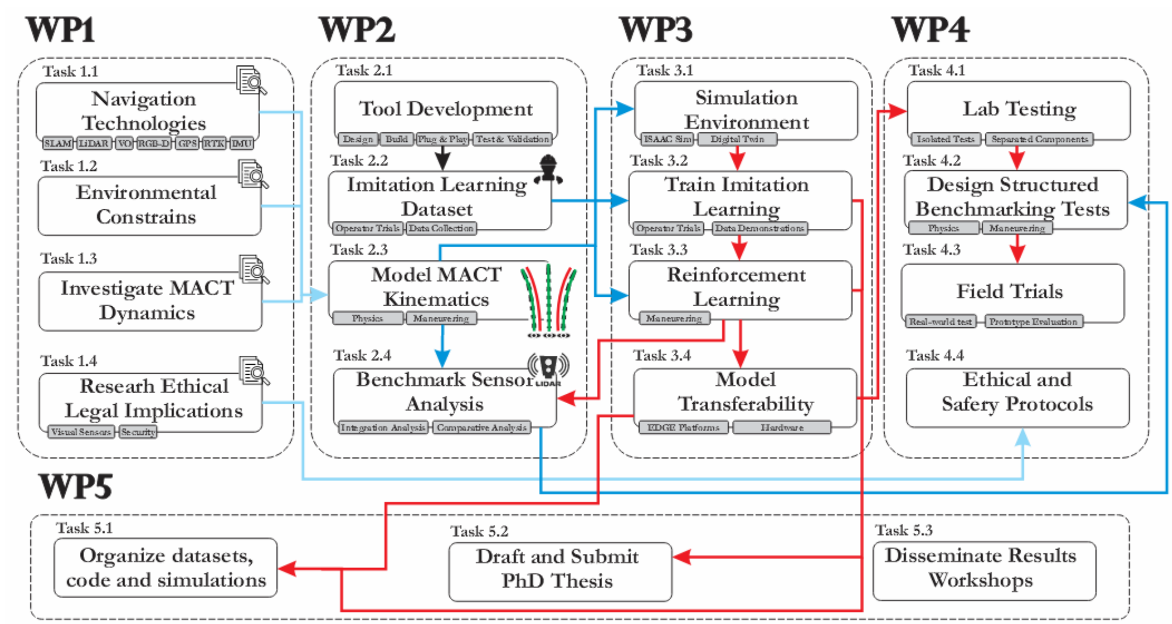


Fig. 1: Overall Perspective of Research Methodology

The proposed system integrates sensorized demonstration, simulation-driven learning, and real-world deployment to enable the autonomous handling of Multi-Articulated Conveyor Trolleys (MACTs). To collect expert demonstrations, a custom-designed handlebar equipped with inertial measurement units (IMUs) and RGB-D or LiDAR sensors will be developed, allowing the capture of full six degrees-of-freedom (6-DOF) motion data during human-guided operation. This rich dataset will form the basis for behavior modeling, motion strategy extraction, and kinematic analysis of non-rigid convoys. Fig. 1 illustrates the overall methodology, beginning with a review of navigation technologies and progressing through tool development, dataset creation, simulation experiments, and real-world field trials. The diagram also highlights how individual tasks interconnect across the research workflow, showing how each stage informs and supports subsequent phases of the project.

Simulated digital twin environments will be created using NVIDIA Isaac Sim, incorporating procedural randomness to improve generalization. Imitation learning (e.g., behavior cloning, DAgger) will be used to learn from expert data, followed by reinforcement learning (e.g., PPO, SAC) to refine policies under realistic dynamics in Isaac Gym. Trained models will be deployed on edge-AI hardware (e.g., NVIDIA Jetson), and evaluated in lab and field trials with variable MACT configurations and surfaces, focusing on policy transferability and operational robustness.

The proposed system will be evaluated in both simulated and real-world settings using key performance indicators. Task success rate will measure completion of MACT transport tasks without intervention, while trajectory similarity, assessed using Dynamic Time Warping (DTW) and path deviation, will quantify alignment with expert demonstrations. Generalization will be evaluated by testing across diverse layouts, load conditions, and convoy configurations. Robustness and safety will be measured through failure rates and the minimum distance maintained from obstacles during operation. Sim-to-real transfer performance will be assessed by comparing results between simulated and physical environments and by tracking the number of fine-tuning episodes needed to recover baseline functionality. These metrics collectively provide a comprehensive assessment of behavior quality, adaptability, and deployment feasibility.

3 Conclusion and Future Work

This work outlines a learning-based framework for the autonomous handling of Multi-Articulated Conveyor Trolleys (MACTs), targeting the automation of non-rigid convoys in unstructured environments. By combining sensorized human demonstrations, simulation-based learning, and adaptive policy generation, the approach aims to develop transferable motion strategies suitable for edge-AI deployment. Leveraging imitation and reinforcement learning within digital twin environments, the project explores scalable solutions for dynamic logistics scenarios.

Future work will focus on finalizing the sensor platform, collecting expert demonstrations, and initiating policy training in simulation. Further exploration will address complementary elements such as online adaptation, human-in-the-loop supervision, and transfer learning to improve generalization, safety, and real-world deployment.

Acknowledgements

This research was funded by the "Agenda Mobilizadora para a inovação empresarial do setor das Duas Rodas (AM2R)" 01/C05-i01/2021 and LASI-LA/P/0104 /2020, co-funded from the "Mobilising Agendas/Alliances for Business Innovation" of the "Next Generation EU" program of Component 5 of the Portuguese Recovery and Resilience Plan (RRP), concerning "Investment and innovation", under the Regulation of the Incentive System for "Mobilising Agendas/Alliances for Business Innovation".

References

1. Aguilera, S., Hutchinson, S.: Control of cart-like nonholonomic systems using a mobile manipulator. In: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 6801–6808 (2023). https://doi.org/10.1109/IROS55552.2023.10342088
2. Aguilera, S., Murtaza, M.A., Rogers, J., Hutchinson, S.: Modeling and inertial parameter estimation of cart-like nonholonomic systems using a mobile manipulator. In: 2023 IEEE International Conference on Robotics and Automation (ICRA). pp. 3073–3079 (2023). https://doi.org/10.1109/ICRA48891.2023.10161076
3. Dadiotis, I., Mittal, M., Tsagarakis, N., Hutter, M.: Dynamic object goal pushing with mobile manipulators through model-free constrained reinforcement learning (2025), https://arxiv.org/abs/2502.01546
4. Gao, X., Luan, H., Xia, B., Zhao, Z., Wang, J., Meng, M.Q.H.: A divide-and-conquer control strategy with decentralized control barrier function for luggage trolley transportation by collaborative robots. Robotica **41**(11), 3333–3348 (2023). https://doi.org/10.1017/S0263574723001005
5. Pan, J., Mai, X., Wang, C., Min, Z., Wang, J., Cheng, H., Li, T., Lyu, E., Liu, L., Meng, M.Q.H.: A searching space constrained partial to full registration approach with applications in airport trolley deployment robot. IEEE Sensors Journal **21**(10), 11946–11960 (2021). https://doi.org/10.1109/JSEN.2020.3042665
6. Scholz, J., Chitta, S., Marthi, B., Likhachev, M.: Cart pushing with a mobile manipulation system: Towards navigation with moveable objects. In: 2011 IEEE International Conference on Robotics and Automation. pp. 6115–6120 (2011). https://doi.org/10.1109/ICRA.2011.5980288
7. Xia, B., Luan, H., Zhao, Z., Gao, X., Xie, P., Xiao, A., Wang, J., Meng, M.Q.H.: Collaborative trolley transportation system with autonomous nonholonomic robots. In: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 8046–8053 (2023). https://doi.org/10.1109/IROS55552.2023.10341508
8. Xiao, A., Luan, H., Zhao, Z., Hong, Y., Zhao, J., Chen, W., Wang, J., Meng, M.Q.H.: Robotic autonomous trolley collection with progressive perception and nonlinear model predictive control. In: 2022 International Conference on Robotics and Automation (ICRA). pp. 4480–4486 (2022). https://doi.org/10.1109/ICRA46639.2022.9812455
9. Xie, P., Xia, B., Hu, A., Zhao, Z., Meng, L., Sun, Z., Gao, X., Wang, J., Meng, M.Q.H.: Autonomous multiple-trolley collection system with nonholonomic robots: Design, control, and implementation (2024), https://arxiv.org/abs/2401.08433

Deep Learning Methods for Left Atrial Appendage Segmentation in 3D TEE Images – A Comparative Study

Luís C. N. Barbosa[1], Siobhán Moane[2], João L. Vilaça[1], and Pedro Morais[1]

[1] Polytechnic of Cávado and Ave, IPCA, Portugal
{lbarbosa, jvilaca, pmorais}@ipca.pt
[2] Technological University of the Shannon - Midlands Midwest, TUS, Ireland
siobhan.moane@tus.ie

Abstract. Atrial fibrillation (AF) is the most prevalent form of cardiac arrhythmia, affecting around 2.5% of the global population. In cases of nonvalvular AF, roughly 90% of thrombi are found in the left atrial appendage (LAA), making LAA closure (LAAC) a clinically relevant intervention. This procedure, typically performed via percutaneous implantation of an occlusion device under fluoroscopic guidance, relies heavily on operator experience and involves significant radiation exposure. Transesophageal echocardiography (TEE) has emerged as a promising imaging modality for guiding LAAC with reduced radiation risk.

In this work, we explore the application of deep learning (DL) methods for segmenting the LAA in 3D TEE images. We compare four DL architectures: UNET, UNETR, DynUNET, and ViTAutoEnc using a dataset of 188 annotated 3D TEE volumes. Ground truth segmentations were generated via a semi-automated method. Among the models, DynUNET achieved the best performance, with an Average Surface Distance of 1.80±0.83 mm, a Dice score of 82.21±5.89%, and a 95th percentile Hausdorff Distance of 2.40±0.92 mm. These findings reinforce the potential of DL-based segmentation to support LAA assessment and contribute to image-guided LAAC procedures.

Keywords: left atrial appendage · artificial intelligence · 3D ultrasound imaging · transesophageal echocardiography · ultrasound segmentation.

1 Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia, affecting more than 30 million people worldwide and one in four middle-aged adults in Europe [1]. It is closely related to an increased risk of thromboembolic stroke, with approximately 90% of thrombi in patients with nonvalvular AF forming in the left atrial appendage (LAA). The LAA is a remnant of the embryonic left atrium and presents significant anatomical variability, including complex multilobular morphologies [2]. LAA closure (LAAC) has become an effective intervention for patients with high risk of stroke, helping prevent long-term bleeding complications [1]. During the procedure, an accurate 3D assessment of the LAA is essential to guide device selection and placement. In this context, LAA segmentation plays a key role by enabling full characterization of the anatomical structure for clinical planning and intervention. Ultrasound (US), particularly transesophageal echocardiography (TEE), is widely used for LAA imaging due to its real-time capability, high resolution, and lack of radiation exposure. Various methods have been proposed for automatic segmentation in US images, ranging from classical image processing techniques [3] to recent advances using deep learning (DL) and neural networks. These DL approaches have shown strong performance in segmenting complex anatomical structures. For example, the Transformer-based TransFusion network [4] improves segmentation accuracy in 3D TEE by integrating multiscale feature fusion strategies.

2 Methods

This study evaluated four deep learning (DL) architectures, UNET, DynUNET, UNETR, and ViTAutoEnc, for the segmentation of the left atrial appendage (LAA) in 3D transesophageal echocardiography (TEE) images. A total of 188 anonymized TEE volumes from LAAC procedures were retrospectively collected, pre-processed to a standardized voxel spacing of 0.4 mm, resized to 128×128×128, and normalized. Ground truth labels were generated via a previously validated semi-automatic approach. The dataset was randomly split into training (70%), validation (20%), and test (10%) sets. All models were implemented using the MONAI framework in Python and trained for up to 1000 epochs with Dice Loss optimization. Evaluation was conducted using Dice Coefficient (DC), 95th percentile Hausdorff Distance (HD), and Average Surface Distance Error (Averssd), reporting average, median, and standard deviation values. Fig.1 provides an overview of the full pipeline, from dataset acquisition to performance evaluation.

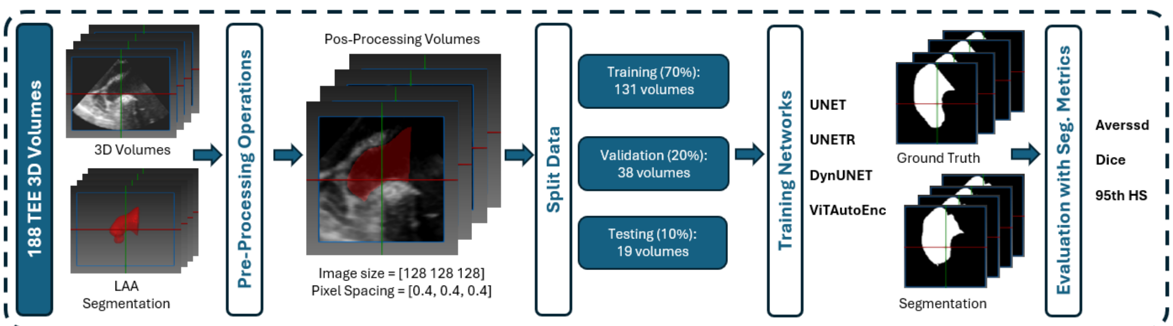


Fig. 1: Graphical summary of the methodology study.

3 Results

The test images were inserted into the trained models to obtain segmentation predictions. From the results obtained, the performance of each tested method is compared. An Averssd was obtained of 2.04±0.72mm, 2.65±1.98mm, 1.80±0.83mm, 2.70±1.80mm, of DC of 76.98±6.95%, 78.72±8.10%, 82.21±2.40%, 68.41±12.70%, and 95th HS of 4.29±1.69mm, 2.77±1.01mm, 2.40±0.92mm, 4.47±2.63mm, for UNET, UNETR, DynUNET and ViTAutoEnc, respectively. The state-of-the-art method proposed in [4] was also trained following the entire methodology detailed previously to measure its performance for the dataset used in this study. An Averssd of 3.15±1.43mm, a DC of 70.08±10.51%, and a 95th HS of 4.43±1.76mm were obtained. These findings highlight DynUNET as the most promising architecture in this study, combining high segmentation accuracy with real-time inference capability, and setting a strong baseline for future improvements and clinical integration.

Acknowledgment

This project was funded by the projects "NORTE-01-0145-FEDER- 000045" and "NORTE-01-0145-FEDER-000059", supported by Northern Portugal Regional Operational Programme (Norte2020), under the Portugal 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). It was also financed by national funds, through FCT - Foundation for Science and Technology and FCT/MCTES under the project UIDB/05549/2020, UIDP/05549/2020, CEECINST/00039/2021 and LASI-LA/
P/0104/2020. It was also funded by the Innovation Pact Health From Portugal, co-funded from the "Mobilizing Agendas for Business Innovation" of the "Next Generation EU" program 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. Jin, C., Feng, J., Wang, L., Yu, H., Liu, J., Lu, J., Zhou, J.: Left Atrial Appendage Segmentation Using Fully Convolutional Neural Networks and Modified Three-Dimensional Conditional Random Fields. IEEE Journal of Biomedical and Health Informatics **22**(6), 1906–1916 (nov 2018). https://doi.org/10.1109/JBHI.2018.2794552
2. Menke, J., Lüthje, L., Kastrup, A., Larsen, J.: Thromboembolism in atrial fibrillation. The American journal of cardiology **105**(4), 502–510 (feb 2010). https://doi.org/10.1016/J.AMJCARD.2009.10.018, `https://pubmed.ncbi.nlm.nih.gov/20152245/`
3. Morais, P., Queirós, S., De Meester, P., Budts, W., Vilaça, J.L., Tavares, J.M.R., D'hooge, J.: Fast segmentation of the left atrial appendage in 3-D transesophageal echocardiographic images. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **65**(12), 2332–2342 (dec 2018). https://doi.org/10.1109/TUFFC.2018.2872816
4. Wu, M., Zhang, D., Hua, Y., Si, M., Liu, P., Wang, Q.: TransFusion: Efficient Vision Transformer based on 3D transesophageal echocardiography images for the left atrial appendage segmentation. Expert Systems with Applications **255**, 124727 (dec 2024). https://doi.org/10.1016/J.ESWA.2024.124727

Cross domain I2I translation for improved Surgical Action Detection

Nuno S. Rodrigues[1,2,4], Helena Torres[1,3], Pedro Morais[1,3], Estevão Lima[2], and João L. Vilaça[1,3]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
[2] Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal
[3] LASI - Associate LAboratory of Intelligent Systems, Guimarães, Portugal
[4] ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal
ndrodrigues@ipca.pt

Abstract. The integration of artificial intelligence (AI) in surgical applications can introduce significant advantages. One of the limitations of deep learning (DL) applications in this context is the scarcity of extensive labeled datasets. To address this constraint, cross-domain knowledge approaches have been explored, with methods such as leveraging videos from surgical training phantoms. However, the lack of realistic traits in phantoms is also a limitation. We propose the implementation of an image-to-image (I2I) translation method to provide realism to surgical training phantoms, potentially improving the performance of deep learning tasks, such as action detection. Initial outcomes suggest that incorporating hyperrealistic synthetic data may contribute to improve the performance of action detection tasks in the context of surgical applications.

Keywords: Image-to-image translation · Surgical action detection · Generative adversarial networks

1 Introduction

Clinical Artificial Intelligence (AI) applications may help provide decision support tools to clinicians and contribute for better and more efficient therapies. However, such AI applications are data dependent [1]. The use of synthetic data to compensate for the lack of labelled clinical datasets has been increasing significantly [5]. Synthetic data can be created with different methods, such as using physical simulators, statistical models or hybrid models (consisting of a fusion of the two previous types) [4]. In this work we propose a multi domain approach (i.e. using two different surgical domains: real surgeries and realistic surgical training phantoms) as an attempt to boost model performance [2]. To address the lifeless and unrealistic appearance of the phantoms (i.e. lack of blood, bleeding, and conjunctive and adipose tissue), we use a GAN-based image-to-image (I2I) translation method, that maps patterns from the intraoperative domain onto the surgical simulation. The goal is to confer realistic traits to surgical training phantoms in order increase the performance of a Deep Learning (DL) surgical action detection model.

2 Methods

Images and labels from a publicly available dataset were used. Two sub-datasets are provided (real surgeries and phantom surgeries) containing the bounding box

coordinates and the respective action classes. In total the dataset contains 23 action classes from which 11 are common to both subsets. Only the common action classes were included and the datasets were re-arranged in the following manner:
Training:

- Set 1: 15739 images from two real surgeries.
- Set 2: 8358 images from surgical sessions performed on two surgical train-ing phantoms.
- Set 3: 8358 hyperrealistic phantom images, where the characteristics of the real surgeries were translated onto the phantom images from set 2.

Validation:

- 5789 images from real surgery 3.

Test:

- 4185 images from real surgery 1.

CycleGAN was used for the I2I task [6]. For the action detection task a Feature Pyramid Network (FPN) based architecture was used, combining a Convolutional Neural Network (CNN) with ResNet50 pre-trained weights [3].

Two separate trainings were performed:

1. Using Sets 1 and 2 (real surgeries and raw phantoms) as input images.
2. Using sets 1 and 3 (real surgeries and realistic phantoms) as input images.

3 Results

Average precision at IoU threshold values of 0.10, 0.30 and 0.50 are reported as AP10, AP30 and AP50, respectively, and meanAP reports the mean average precision computed by calculating the mean of three AP values. Results for training setup 1 are: AP10 - 26.9; AP30 - 16.0; AP50 - 6.9; MeanAP - 16.6. Results for training setup 2 are: AP10 - 27.4; AP30 - 16.7; AP50 - 6.9; MeanAP - 17.0.

4 Discussion and Conclusion

The achieved results indicate comparable performance between training setups 1 and 2, with slight improvements in the Real+Realistic configuration, supporting the hypothesis that conferring hyperrealism to surgical training phantoms through I2I translation has the potential to enhance the generalization capacity of the action detection model. In future developments we intend to implement concomitant AI methods in the I2I task to improve regularization and preservation of desired image features and ongoing efforts are directed towards refining the action detection model to further enhance performance.

Acknowledgements

The authors acknowledge the company KARL STORZ SE & Co. KG for their support of this research, and the Portuguese National funds, through the Innovation Pact HfPT – 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". The work was also funded by Foundation for Science and Technology (FCT) - project UIDB/50026/2020 (DOI 10.54499/UIDB/50026/2020), UIDP/50026/2020 (DOI 10.54
499/UIDP/50026/ 2020), LA/P/0050/2020 (DOI 10.54499/LA/P/0050/2020), UIDB/ 05549/2020 (DOI: 10.54499/UIDB/05549/2020), UIDP/05549/2020 (DOI: 10.54499/UI
DP/ 05549/2020), LASI/LA/P/0104/2020, and CEECINST/00039/2021.

References

1. Chen, R.J., Lu, M.Y., Chen, T.Y., Williamson, D.F., Mahmood, F.: Synthetic data in machine learning for medicine and healthcare. Nature Biomedical Engineering **5**, 493–497 (6 2021). https://doi.org/10.1038/s41551-021-00751-8
2. Li, X., Luo, M., Ji, S., Zhang, L., Lu, M.: Evaluating generative adversarial networks based image-level domain transfer for multi-source remote sensing image segmentation and object detection. International Journal of Remote Sensing **41**, 7327–7351 (2020). https://doi.org/10.1080/01431161.2020.1757782, `https://doi.org/10.1080/01431161.2020.1757782`
3. Lin, T., Dollár, P., Girshick, R.B., He, K., Hariharan, B., Belongie, S.J.: Feature pyramid networks for object detection. CoRR **abs/1612.03144** (2016), `http://arxiv.org/abs/1612.03144`
4. McDuff, D., Curran, T., Kadambi, A.: Synthetic data in healthcare. arXiv preprint arXiv:2304.03243 (2023)
5. Rajotte, J.F., Bergen, R., Buckeridge, D.L., El Emam, K., Ng, R., Strome, E.: Synthetic data as an enabler for machine learning applications in medicine. iScience **25**(11), 105331 (2022). https://doi.org/https://doi.org/10.1016/j.isci.2022.105331, `https://www.sciencedirect.com/science/article/pii/S2589004222016030`
6. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE International Conference on Computer Vision **2017-Octob**, 2242–2251 (2017). https://doi.org/10.1109/ICCV.2017.244

The Journey to Achieve a True Personalized Health Assistant

António Faria[1,2], Patrick Murray[2], Pedro Morais[1], and João L. Vilaça[1]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
afaria@ipca.pt,pmorais@ipca.pt,jvilaca@ipca.pt
[2] LIFE Research Institute, TUS, Limerick, Ireland
patrick.murray@tus.ie

Abstract. Recent developments in the artificial intelligence world more concrete the large language model open the possibility for the development of better personalized digital health assistants. But these systems need to have some key characteristics, trustworthiness, so the user is not afraid of using their answers, context-aware, so the answerers be personalized to a problem, and verifiable so we can understand why it gave that answer. This paper outlines our development journey, from an assistant for non-communicable disease prevention to OpAssist, an assistant customized for laparoscopic prostatectomy surgery. We analyze the methodologies used and compare them with recent advances to identify new paths to improvement and ensure relevance.

Keywords: Personalized Health Assistant · Digital Health · Surgical Assistance · Retrieval-Augmented Generation · Large Language Models.

1 Introduction

Although the potential of digital health assistants in personalized healthcare is noteworthy, accomplishing it is complex because the solution needs to be responsive, reliable, and adaptable to diverse scenarios, from disease prevention to disease treatment in all stages of life. We start by conceptualizing the system and its key components. This system was developed for the prevention of non-communicable disease (NCD) [4]. In this system, we focus on information reliability, a rudimentary version of Retrieval-Augmented Generation (RAG) was implemented [8], where its generating engine was the Large Language Model (LLM), LLAMA-3 Building on that, we developed OpAssist [5], a digital health assistant designed for laparoscopic prostatectomy surgery. In this system, we updated the LLM used, Gemma-2, and incorporated an updated version of the RAG technique where a custom-weighted ranking mechanism was used. This mechanism uses a combination of semantic keywords and metadata scores to rank the best context to the model. Finally, due to the required sterility of the environment, a voice interaction system was employed using Whisper [9] and SpeechT5 [2].

2 Discussion

The methodology used in the developed assistants reflects the AI practices at the time of the system development. Now, new and relevant techniques have been developed and have been established as a standard since then. In this chapter, we highlight the critical areas for refinement where rapid advancements in AI have been more prominent.

Retrieval-Augmented Generation: It's a key component for building answers with verifiable information and is used by the NCD assistant and OpAssist. Although this is a good technique when dealing with one document but it may face some limitations when it's required to do multi-hop reasoning [12]. To overcome these limitations, new architectures have been designed. Some of them are graph-based retrieval [13], iterative retrieval generation [11] or even through the use of a model that plans the needed retrievals, so all the information is available to the model when it's in the generating phase [7].

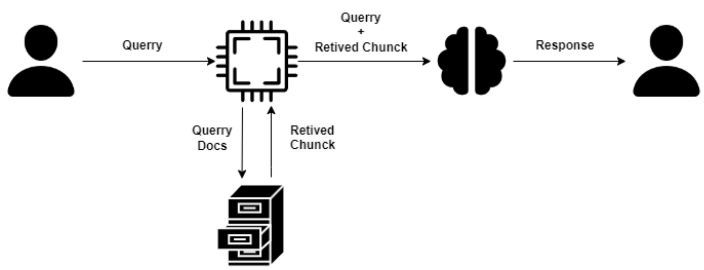


Fig. 1: An illustration of the Retrieval-Augmented Generation technique. Reproduced from [4].

Large Language Models: These models are responsible for generating the answer when a prompt is sent to the assistant. The LLMs used lack some of the newer capabilities, such as long-context, sophisticated reasoning, agentic behavior, and vision understanding. Anthropic's Claude 3 [1], or Alibaba's Qwen2.5-1M [10] models, are equipped with larger context windows, allowing the model to analyze the complete medical history of the patient. The vision understanding capability of the new models (Google's PaliGemma [6], and Alibaba's Qwen-VL [3]) can be used to analyze medical exams with visual representations.

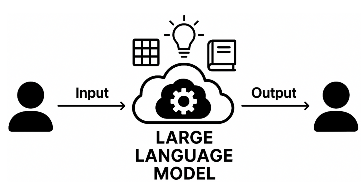


Fig. 2: An illustration of the information flow to a Large Language Model.

3 Conclusion

This paper presents our research path to a personalized health assistant that started with the development of an NCD prevention assistant and later evolved into a highly specialized surgical assistant. We detail the methodology and the techniques employed to provide personalized health assistance. Finally, this work sheds some light on how to overcome the challenges inherent in this domain and achieve the full potential of truly personalized health assistants.

Acknowledgement

This project was funded by NEXT GENERATION EU through Mobilizing Agendas for Business Innovation, Recovery and Resilience Plan (PRR), within the scope of the HfPT - PPS3 - Digital Thermometer - 41 project.

References

1. Anthropic: Introducing the claude 3 family (2024), https://www.anthropic.com/news/claude-3-family
2. Ao, J., et al.: Speecht5: Unified-modal encoder-decoder pre-training for spoken language processing (2021), https://arxiv.org/abs/2110.07205, arXiv:2110.07205
3. Bai, J., et al.: Qwen-vl: A versatile vision-language model for understanding, localization, text reading, and beyond (2023), https://arxiv.org/abs/2308.12966, arXiv:2308.12966
4. Faria, A., et al.: A personalized digital health assistant for lifelong self-care. In: Proceedings of the 4th Symposium of Applied Science for Young Researchers (SASYR 2024). pp. 52–54. Instituto Politécnico de Bragança (2024)
5. Faria, A., et al.: Opassist: An intelligent virtual assistant for laparoscopic prostatectomy. In: Medical Imaging 2025: Image-Guided Procedures, Robotic Interventions, and Modeling. Proc. SPIE, vol. 13408, p. 64 (2025), https://doi.org/10.1117/12.3047077
6. Google: Paligemma: A lightweight vision-language model (2024), https://ai.google.dev/gemma/docs/paligemma, google AI for Developers
7. Lee, M., et al.: Planrag: A plan-then-retrieval augmented generation for generative large language models as decision makers (2024), https://arxiv.org/abs/2406.12430, arXiv:2406.12430
8. Lewis, P., et al.: Retrieval-augmented generation for knowledge-intensive nlp tasks (2021), https://arxiv.org/abs/2005.11401, arXiv:2005.11401
9. Radford, A., et al.: Robust speech recognition via large-scale weak supervision (2022), https://arxiv.org/abs/2212.04356, arXiv:2212.04356
10. Yang, A., et al.: Qwen2.5-1m technical report (2025), https://arxiv.org/abs/2501.15383, arXiv:2501.15383
11. Yu, T., et al.: Auto-rag: Autonomous retrieval-augmented generation for large language models (2024), https://arxiv.org/abs/2411.19443, arXiv:2411.19443
12. Zhao, P., et al.: Retrieval-augmented generation for ai-generated content: A survey (2024), https://arxiv.org/abs/2402.19473, arXiv:2402.19473
13. Zhou, Y., et al.: In-depth analysis of graph-based rag in a unified framework (2025), https://arxiv.org/abs/2503.04338, arXiv:2503.04338

Assembly worker Digital Twin for workspace optimization

Ricardo N. C. Rodrigues[1,2,3], Jaime Fonseca[2,3], and António H. J. Moreira[1,2]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
rnrodrigues@ipca.pt
[2] LASI – Associate Laboratory of Intelligent Systems, Guimarães, Portugal
sec@lasi-research.pt
[3] Algoritmi Center, University of Minho, Guimarães, Portugal
secretaria@algoritmi.uminho.pt

Abstract. Industry 5.0 emphasizes human-centricity, aiming to integrate advanced digital technologies with human capabilities to create more adaptive and resilient industrial environments. This work then proposes the development of a Human Digital Twin (HDT) focused on optimizing and personalizing workspaces in manual assembly tasks. The proposed system aims to create a Cyber-Physical System (CPS) workbench equipped with an augmented reality interface and real-time computer vision, coupled with an upper body HDT modeled through imitation learning techniques, and a reinforcement learning-based workspace optimizer. During development data is to be collected from controlled assembly tasks involving varied object configurations and user behaviors. The expected outcome is a robust HDT that accurately models user-specific physical traits, enabling dynamic workspace adjustments that enhance efficiency, comfort, and safety. This project serves as a proof of concept for HDT-driven workspace personalization in Industry 5.0.

Keywords: Human Digital Twin · Cyber-Physical System · Imitation Learning · Reinforcement Learning

1 Introduction

Industry 5.0 marks a paradigm shift in industrial transformation, focusing on human-centricity, sustainability, and resilience beyond the automation-driven goals of Industry 4.0. At the core of this evolution lies the integration of advanced technologies with human capabilities, which fosters synergy between humans and intelligent systems [1, 2]. Assembly Human Digital Twins (HDTs) virtual representations of individuals enriched with real-time data and behavioral models emerge as an enabler in this context. By bridging the physical and digital realms, HDTs facilitate personalized system interactions, adaptive decision-making, and improved occupational safety, underscoring their critical role in shaping the future of work within Industry 5.0 frameworks [3, 4, 6].

Following this perspective, the authors propose a development pipeline to a HDT focused on the optimization and personalization of the workspace for assembly task workers.

2 Project Concept

This project consists in the creation of an upper body digital twin with the main objective of optimizing the workspace to the necessities and limitations of the users.

The proposed system will be divided in three different elements, see Fig. 1, a Cyber-Physical System (CPS) workbench equipped with augmented reality interface and real-time computer vision data acquisition system, an Upper body HDT of the user representing his physical peculiarities, a workspace optimizer that forecast the best configuration for the user.

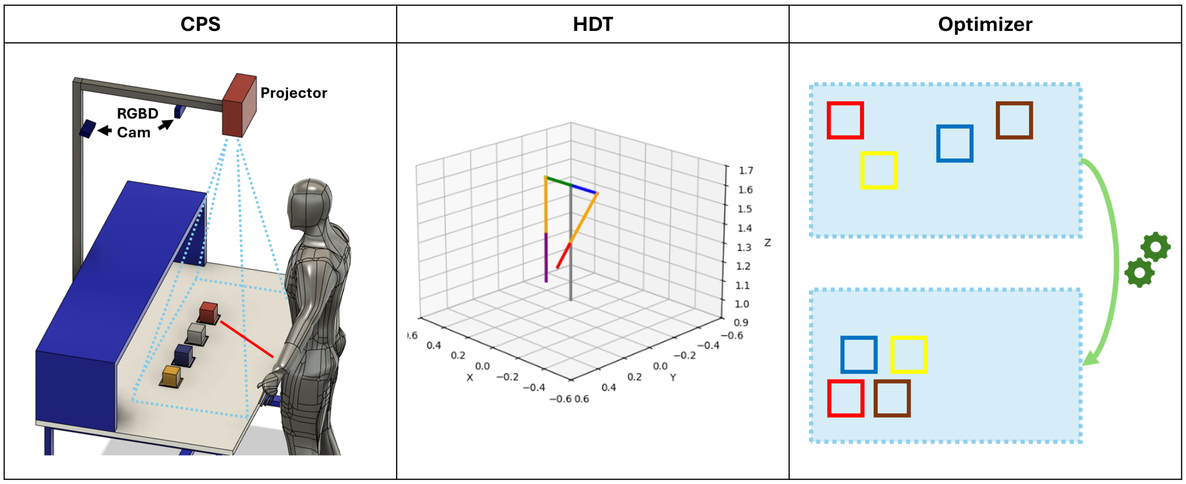


Fig. 1: Proposed system core elements, left- Representation of the CPS, center- Simplified representations of the HDT, right- Representation of the optimizer output

The CPS Workbench consists of an 1.6 by 1 meter long workspace coupled with a projector creating an augmented reality interface with the user, two RGBD cameras (realsense D435i) responsible for environment (workbench and human) data acquisition linked with a model from mmPose Toolbox for body joint tracking.

The user HDT will consist in deep learning model trained with the data acquired by the CPS, using imitation learning technics such Behavioral Cloning (BC), Density-Based Reward Modeling or Adversarial Inverse Reinforcement Learning (AIRL).

The Workspace Optimizer will consist in a deep learning model acquired using reinforcement learning algorithms, for instance Proximal Policy Optimization (PPO), trained with the HDT as an interactive element.

2.1 Dataset formation and training

The data (timestamps, worker body joints and picking/placing coordinates) for the HDT will be acquired in a laboratorial environment and will consist in the assembly of four colored (blue, brown, yellow and transparent) cubes in a predetermined configuration.

On an initial approach, one hundred assemblies will be recorded, where the subjects will be subjected to random initial picking positions on each assemble but to a constant pattern of assembly.

After validation of the first approach, the subjects will perform one hundred more repetitions, however with random picking positions and assembly patterns, to further generalize the HDT.

As the optimizer consists of a reinforcement learning model, a dataset is not needed [5] nonetheless the HDT will be included on the training environment introducing the variability and features that needed to be taken in consideration.

3 Expected Results and Conclusions

At the end of this project, is expected to have a framework that enable the development of a HDT capable of representing assembly worker movements tendencies and limitations, further capable of distinguish the configuration that best suits the user and task. Extended by an engaging workspace provided by a CPS. This project is intended to be a proof-of-concept that, it is possible to develop a reliable method to improve currently assembly processes where the users are not only taken in consideration but are the key element in it.

References

1. Adel, A.: Future of industry 5.0 in society: human-centric solutions, challenges and prospective research areas. Adel Journal of Cloud Computing **11**, 40 (2022). https://doi.org/10.1186/s13677-022-00314-5
2. Mourtzis, D., Angelopoulos, J., Panopoulos, N.: The future of the human–machine interface (hmi) in society 5.0. Future Internet 2023, Vol. 15, Page 162 **15**, 162 (4 2023). https://doi.org/10.3390/FI15050162
3. Qiu, C., Zhou, S., Liu, Z., Gao, Q., Tan, J.: Digital assembly technology based on augmented reality and digital twins: a review. Virtual Reality & Intelligent Hardware **1**, 597–610 (12 2019). https://doi.org/10.1016/J.VRIH.2019.10.002
4. Subramanian, K., Thomas, L., Sahin, M., Sahin, F.: Supporting human–robot interaction in manufacturing with augmented reality and effective human–computer interaction: A review and framework. Machines 2024, Vol. 12, Page 706 **12**, 706 (10 2024). https://doi.org/10.3390/MACHINES12100706
5. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. The MIT Press, second edn. (2018), `http://incompleteideas.net/book/the-book-2nd.html`
6. Zhang, X., Yang, Y., Zhang, X., Hu, Y., Wu, H., Li, M., Handroos, H., Wang, H., Wu, B.: A multi-level digital twin construction method of assembly line based on hybrid worker digital twin models. Advanced Engineering Informatics **62**, 102597 (10 2024). https://doi.org/10.1016/J.AEI.2024.102597

Imitation Learning for Adaptive Grasping and Precise Placement of Unstructured Metal Parts: A Case Study

Nuno Simões[1], Rafael Fernandes[1], Bruno Silva[1], Joaquin Dillen[1], João M. Faria[1], Luis Vilas-Boas[1], Ines Caetano[2], Luis Cardoso[3], Paulo Silva[4], João Borges[1], and António Moreira[1]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
nsimoes@ipca.pt
[2] Sistrade, Porto, Portugal
[3] Miranda, Águeda, Portugal
[4] SRAM, Coimbra, Portugal

Abstract. In this work, the use of imitation learning was investigated as a means to address the challenges associated with robotic grasping and placement of metallic parts in unstructured industrial environments. To this end, a cyclic, data-centric framework is proposed, integrating demonstration collection, data augmentation, iterative training, and a closed-loop re-picking mechanism to enable adaptability to variations in object position and orientation. Although the framework has not yet been implemented, its design has been outlined within a high-fidelity simulation context. Additionally, a set of evaluation metrics has been defined to support future empirical validation. It is expected that, once developed and deployed, the approach will facilitate robust and flexible manipulation, contributing to the automation of high-mix, low-volume manufacturing processes.

Keywords: Imitation Learning · Robotic Grasping · Repicking · Computer Vision

1 Introduction

Modern high-mix, low-volume manufacturing requires robotic systems capable of adapting swiftly to new parts and layouts—needs that often surpass model-based planners in unstructured environments [3]. Imitation learning offers an alternative by enabling control policies from demonstrations, reducing reliance on object models and synthetic datasets [1]. Yet naive methods like behavioral cloning suffer from covariate shift, motivating online strategies such as DAgger [6].

Recent work enables one-shot adaptation and hybrid assembly strategies: warping-based interaction learning generalizes SE(3) manipulation from a single demo [2], and combined imitation plus reinforcement pipelines enable precise assembly with minimal feedback [7]. In parallel, GG-CNN delivers real-time pixelwise grasping at 50 Hz [5], Dex-Net demonstrates robust sim-to-real generalization [4], and Transporter Networks support model-free 6-DOF pick-and-place [8].

Nonetheless, manipulating metallic parts remains challenging due to reflectivity and tight tolerances [9]. To address this, we propose a cyclic, data-centric pipeline integrating demonstration collection, augmentation, iterative training, and closed-loop re-picking. The paper details architecture, simulation setup, evaluation metrics, and future deployment plans.

2 Development

2.1 Research Questions

Despite substantial progress in adaptive robotic systems, the precise and flexible handling of metal parts in unstructured settings remains a critical challenge. Variability in part positioning and orientation greatly complicates both grasping and the subsequent placement into structured transport trays. To address these difficulties, this research is guided by the following two central questions.

RQ1: How can vision-based deep learning models be used to enhance the adaptability and compatibility between robotic grasping tools and parts?

RQ2: What is the impact of part positioning and orientation variability on the overall success rate of robotic manipulation, and how can these factors be accounted for in trajectory planning?

2.2 System Architecture and Data Acquisition Pipeline

The proposed system is composed of three tightly integrated modules: vision and data acquisition, planning and control, and continuous feedback for iterative refinement. A calibrated RGB-D camera is used to acquire a 3D representation of the workpieces, which is subsequently processed by an imitation learning framework to produce motion commands for grasping and placement. Instances of unsuccessful grasping or placement are automatically recorded, and the corresponding sensory action data are reintegrated into the training cycle, thereby closing the data-centric feedback loop.

At the core of this approach lies a cyclic, data-driven pipeline, illustrated in Fig. 1, which unfolds as follows: first, demonstration trajectories are collected in a simulated, unstructured environment with systematically varied part positions and orientations. Each demonstration is annotated with subtask labels, and the dataset is augmented to generate new variations from the original demonstrations. This dataset is then used to train the imitation learning model. During deployment, failures automatically trigger additional rounds of data collection and retraining, ensuring continuous improvement through feedback.

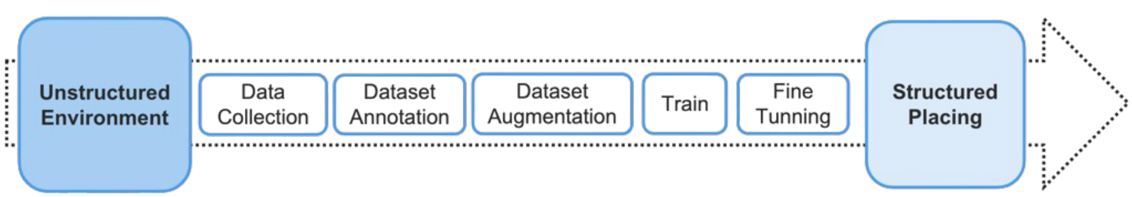


Fig. 1: Imitation learning framework used for demonstration collection, augmentation, training, and refinement in robotic grasping and placement tasks.

2.3 Simulation Environment and Evaluation Metrics

All experiments will be in NVIDIA Isaac Sim, using the Isaac Lab framework, which provides high fidelity physics simulation and ready to use task templates for imitation learning workflows. This setup allows for safe and repeatable generation of data prior to any real world deployment. System performance is evaluated using five key performance indicators: the success rate of first attempt grasps, the average number of repicks required for a successful grasp, spatial and angular placement error, total cycle time, and the generalization score on previously unseen part geometries.

3 Conclusion and Future Work

This paper outlines a system in which imitation learning and deep vision models are employed to tackle the challenges of grasping and placing parts in unstructured industrial environments. Through the use of a cyclic, data-driven pipeline comprising demonstration collection, augmentation, and iterative retraining, this work aims to enable generalization across varying part positions and orientations. The incorporation of a re-picking mechanism has further improved the system's reliability by facilitating closed-loop recovery from grasp failures. Although the system remains at the concept stage and lacks empirical validation, its design aligns well with prevailing trends in robotic learning research. Nevertheless, limitations such as reliance on simulated data, sim-to-real transfer constraints, and the computational demands of deep learning remain nontrivial.

Future work is expected to focus on deployment within real robotic platforms, where trained models will be transferred and validated under real world conditions. Specific challenges to be addressed include sensor noise, actuation latency, and discrepancies between simulated and physical domains. The inclusion of real world demonstrations is anticipated to enhance the robustness of the learning pipeline and further bridge the sim-to-real gap. Successful application of the system in an industrial context is viewed as essential for confirming its practical viability and uncovering new avenues for optimization.

In summary, this work demonstrates that imitation learning, when integrated with advanced computer vision techniques and developed under a data-centric methodology, holds significant potential for achieving flexible, robust, and adaptive robotic manipulation in high mix manufacturing environments.

Acknowledgment

This research was funded by the "Agenda Mobilizadora para a inovação empresarial do setor das Duas Rodas (AM2R)", co-funded from the "Mobilising Agendas/Alliances for Business Innovation" of the "Next Generation EU" program of Component 5 of the Recovery and Resilience Plan (RRP), concerning "Investment and innovation", under the Regulation of the Incentive System for "Mobilising Agendas/Alliances for Business Innovation".

References

1. Argall, B.D., Chernova, S., Veloso, M., Browning, B.: A survey of robot learning from demonstration. Robotics and Autonomous Systems **57** (2009). https://doi.org/10.1016/j.robot.2008.10.024
2. Biza, O., Thompson, S., Pagidi, K.R., Kumar, A., van der Pol, E., Walters, R., Kipf, T., van de Meent, J.W., Wong, L.L.S., Platt, R.: One-shot imitation learning via interaction warping (2023), `https://arxiv.org/abs/2306.12392`
3. Gan, Z.L., Musa, S.N., Yap, H.J.: A review of the high-mix, low-volume manufacturing industry (2023). https://doi.org/10.3390/app13031687
4. Mahler, J., Liang, J., Niyaz, S., Laskey, M., Doan, R., Liu, X., Ojea, J.A., Goldberg, K.: Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics. In: Proceedings of Robotics: Science and Systems (RSS). vol. 13 (2017). https://doi.org/10.15607/rss.2017.xiii.058
5. Morrison, D., Corke, P., Leitner, J.: Closing the loop for robotic grasping: A real-time, generative grasp synthesis approach. In: Robotics: Science and Systems (2018). https://doi.org/10.15607/RSS.2018.XIV.021
6. Ross, S., Gordon, G.J., Bagnell, J.A.: A reduction of imitation learning and structured prediction to no-regret online learning. In: Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS). pp. 627–635 (Apr 2011). https://doi.org/10.48550/arXiv.1011.0686
7. Wang, C., Su, C., Sun, B., Chen, G., Xie, L.: Extended residual learning with one-shot imitation learning for robotic assembly in semi-structured environments. Frontiers in Neurorobotics **18**, 1355170 (2024). https://doi.org/10.3389/fnbot.2024.1355170
8. Zeng, A., Florence, P., Tompson, J., Welker, S., Chien, J., Attarian, M., Armstrong, T., Krasin, I., Duong, D., Sindhwani, V., Lee, J.: Transporter networks: Rearranging the visual world for robotic manipulation. In: Proceedings of Machine Learning Research. vol. 155 (2020)
9. Zhang, J., Liu, H., Li, D., Yu, X., Geng, H., Ding, Y., Chen, J., Wang, H.: Dexgraspnet 2.0: Learning generative dexterous grasping in large-scale synthetic cluttered scenes (10 2024), `http://arxiv.org/abs/2410.23004`

Trocar-Point for robotic surgery

Sérgio G. Pereira[1], Pedro Morais[2], Estevão Lima[1], and João L. Vilaça[1]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
{sfpereira, pmorais, jvilaca}@ipca.pt
[2] ICVS/3B's - Laboratório Associado, Instituto de Ciências da Vida e da Saúde
estevaolima@med.uminho.pt

Abstract. Minimally invasive surgery was developed to replace traditional open surgery, aiming to reduce blood loss and enhance patient recovery and comfort. This surgical approach uses small incisions through which the surgeon inserts instruments and a camera to visualize the abdominal cavity. Today, several robotic systems exist to hold and position the camera, but these systems are manually controlled by the surgeon. There is an increasing need to automate this process to improve efficiency and precision. Many studies have explored image processing techniques to estimate the optimal camera position, but translating these methods into practical, functional devices remains a challenge. This work presents a proof-of-concept for a robotic system designed to automatically track a specific point during surgery. An experiment was conducted in which a robotic arm holding the camera followed an auxiliary robot performing movements along different axes. The system achieved satisfactory results, demonstrating moderate latency in rotational movements (133 ms for γ rotations and 233 ms for β rotations) and high accuracy in translational movements, with an average error of 1.76 mm over displacements of 30 mm. These findings demonstrate the effectiveness of the system for dynamic trocar point tracking.

Keywords: Dynamic trocar-point · laparoscopic camera control · minimally invasive surgery.

1 Introduction

Minimally invasive surgery (MIS) has become a significant alternative to traditional open surgery. By utilizing small incisions in the patient's abdomen, MIS offers multiple advantages, including reduced discomfort, lower blood loss, shorter hospital stays, and decreased surgical costs [2]. To visualize the surgical site, MIS employs a laparoscopic camera in conjunction with specialized instruments. Traditionally, a second surgeon manually maneuvers the camera under the direction of the lead surgeon [1]. In recent years, electromechanical systems have been introduced to hold and position the camera, offering greater stability and precision compared to manual handling [3]. Building on this progress, the present work introduces a proof-of-concept for dynamic trocar-point (TP) control, enabling robotic manipulation of the laparoscopic camera. To achieve this, a KUKA LBR iiwa 7R800 robot (Augsburg, Germany) was used to control the camera in real time, programmed to follow the movements generated by a second robot.

2 Methods

Fig. 1 shows an overview of the proposed workflow for the laparoscopic camera guidance system. First, we freely move the robot with the camera to the TP (A). Next, we establish communication between the different validation interfaces (robot to guide the camera and robot to transmit the target position) and start exchanging information via TCP socket (B). Then, when the tracking system gives a signal, the robot will start moving. To do this, this work was divided into three parts: designing the tool to hold the camera, developing the robotic control application and study of the dynamic TP.

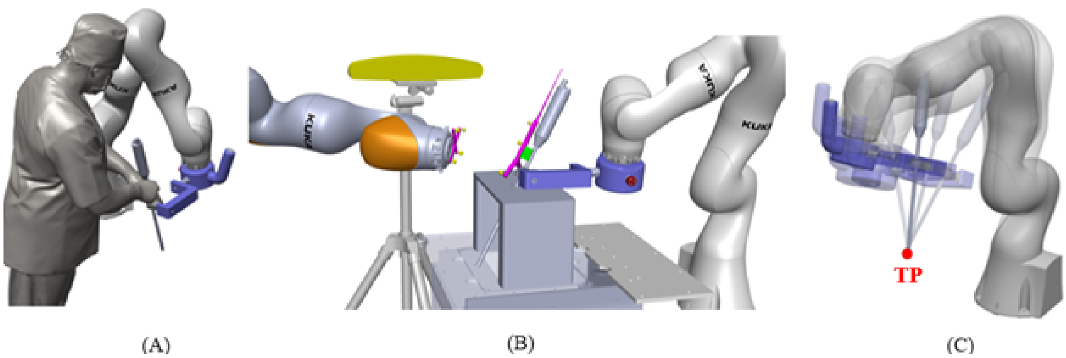


Fig. 1: System workflow: (A) hand-guiding moving to the TP; (B) connection between devices and definition of the TP; (C) start of the movement over the TP.

3 Experiments and results

To validate our system, we define sinusoidal movements with an amplitude of 30mm (each movement was made around an axis (x,y,z)). At the same time that robot 1 (R1) performed the movement, it sent its current position to the robot with the camera (C1) to track it. The first two movements were performed to validate the ability of our robot to rotate in relation to the TP. To do this, the robots are placed face-to-face and R1 only moves around one axis at a time (x or y). The value of this displacement is sent in real time to C1 and converts the displacement to degrees. Ideally, the TP should not shift as the camera moves, and, in these two trials, we also observed the movement made by the isolated marker on the tool at the TP, and it obtained an average displacement of 3.29mm for the gamma and 0.34mm for the beta, which are good results because this value should be as close to 0 as possible. Using data from the Polaris system, which captured markers positions at a sampling rate of 30Hz, we analyzed the peak values to determine the latency between system components. We observed delays of 133 ms during gamma movements and 233 ms during beta movements, and 28 ms in the X-translation.

References

1. Kremer: Minimally invasive abdominal surgery. Thieme (2001)
2. Pokorny: Minimally invasive versus open surgery for degenerative lumbar pathologies:a systematic review and meta-analysis. European Spine Journal **31**(10), 2502–2526 (2022). https://doi.org/http://dx.doi.org/10.1007/s00586-022-07327-3
3. Wijsman: Efficiency in image-guided robotic and conventional camera steering: a prospective randomized controlled trial. Surgical Endoscopy Journal **36**(4), 2334–2340 (2022). https://doi.org/http://dx.doi.org/10.1007/s00464-021-08508-9

Robotic-Assisted 3D Scanning for Assessing the Arrangement of Goods on Pallets in Logistics Warehouse

Rafael Fernandes[1], Bruno Silva[1], Nuno Simões[1], Joaquin Dillen[1], João M. Faria[1], Luís Vilas Boas[1], Inês Caetano[2], Luís Cardoso[3], Paulo Silva[4], João Borges[1], and António H. J. Moreira[1]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
rmfernandes@ipca.pt
[2] Sistrade, Porto, Portugal
[3] Miranda, Águeda, Portugal
[4] SRAM, Coimbra, Portugal

Abstract. Ensure accurate tracking of pallets during shipment is crucial to detecting fraud or tampering during transport. Currently, static scanning systems have limitations regarding the pose and dimensions of the pallet to be scanned, while manual systems suffer from inconsistency and low repeatability due to human intervention in the scanning process, giving these limitations, there is a need to make the digitalization process more robust and flexible. To address this problem, I propose the design and construction of a robotic-assisted 3D scanning using a mobile platform capable of autonomous navigation and detecting pallets in warehouse environments using LIDAR sensors and with the use of depth cameras, it enables dynamic and flexible 3D scanning, allowing for complete volumetric reconstruction and visual analysis of pallets. The data collected by the robotic system is processed on a server using artificial intelligence algorithms to further detect inconsistencies and anomalies. This proof of concept sets the foundation for a more intelligent, autonomous, and reliable pallet inspection process aligned with the principles of Industry 4.0.

Keywords: 3D scanning · Pallet expedition · Mobile Robot.

1 Introduction

In the context of Industry 4.0, one of the main concerns within the logistics and distribution sector is freight fraud, along with the challenges related to traceability and shipment alterations during transportation. To address these issues, there is a growing need to improve the tracking processes for pallets that contain packaged products. By improving the ability to track shipments inbound and outbound logistics,this tracking process can significantly mitigate the risk of fraud and quickly identify discrepancies or unauthorized changes during transit.

During transportation, pallets can be damaged due to incorrect handling or due to poor packaging practices, as well as the risk of intentional tampering such as theft of goods by the carrier or even the alteration of products in the orders for others of lower quality or the placement of illicit products. This research investigates the possibility of developing a robotic-assisted 3D scanning system to improve pallet scanning. Currently, scanning systems used in logistics, such as static systems, have significant limitations, being sensitive to the position and height of the pallet relative to the sensors. Manual scanning systems, on the other hand, have high variability and low repeatability, due to the human factor [1, 2, 5].

This research investigates the possibility of a mobile scanning system improving pallet scanning with the use of a mobile autonomous platform [3, 4].

2 Research & Development Methodology

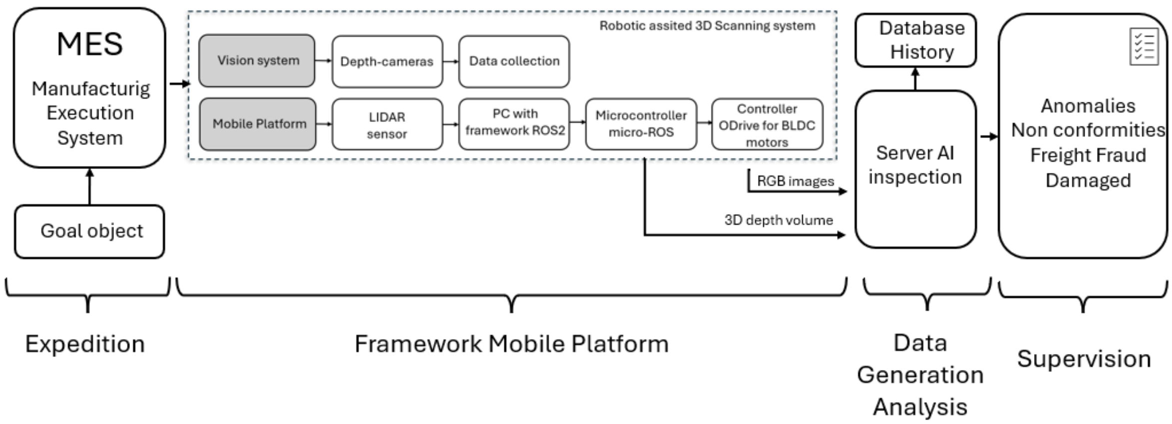


Fig. 1: Illustrate the overall architecture and operational flow of the proposed robotic-assisted 3D scanning system for pallet inspection in logistics

As illustrated in the Fig. 1 this study aims to overcome the limitations of static and manual scanning systems currently used in logistics. The proposed solution involves the design and development of a robotic-assisted 3D scanning system using a mobile robot capable of autonomously navigate and detect pallets in warehouse environments through the use of a LIDAR sensor (Livox MID-360). The robot is equipped with two Intel Realsense D435I depth cameras, which will be used to scan the pallets by collecting data such as volume measurements and capturing relevant information through RGB images, including components such as strapping tapes, protective plastic wrap, and informational labels.

2.1 Mobile Platform architecture

The mobile platform proposed for development consists of the implementation, development and analysis of the sensor structure and performance of the mobile system with a differential configuration and equipped with sensors and controllers that will allow it to perform navigation and data acquisition in logistics environments. The infrastructure will be supported by the ROS2 (Robot Operating System 2) framework running in intel N97 Mini PC, which ensures a distributed and efficient communication between the various system modules. The control sub-system will be build with a micro-controller ESP32 running micro-ROS. The perception of the surrounding environment will be guaranteed by a LIDAR sensor (Livox MID-360), which will allow the detection and location of pallets, while the movement of the platform will be ensured by BLDC motors, actuated by an ODrive controller, offering high precision and responsiveness in locomotion. Three-dimensional information for volumetric reconstruction and precise segmentation of the pallet will be ensured through the use of depth cameras.

The future platform aims to autonomously move around a Euro-pallet (800mm wide, 1200mm long and with maximum height of 2000mm) ensuring the complete scan with an accuracy better than 100mm, allowing for precise and controlled data acquisition from 360 perspectives.

2.2 Vision system and data generation analysis

The vision system mounted on the mobile platform will consist of two depth cameras that will allow essential three-dimensional information to be collected for volumetric reconstruction and precise segmentation of the Euro-pallet in a dynamic and continuous manner. Through the use of depth cameras, it will be possible to collect essential three-dimensional information for volumetric reconstruction and precise segmentation of the pallet, and also, through the RGB images with the use of AI algorithms like YOLOv11, it is possible to know the arrangement of elements and relevant visual characteristics, such as the detection of labels, anchoring tapes, and other additional information. The Ai algorithms will be processed on a central server, with the aim of detecting anomalies, non-conformities, estimating the volume, identifying labels, and extracting other information relevant to the shipping process.

3 Conclusion and Future Work

This study proposes a new innovative approach to pallet scanning in logistics environments through the development and construction of a robotic assisted 3D scanning system. The solution aims to overcome the static and manual systems currently used in order to improve the dynamic, achieving an accuracy above of 95% ,and autonomous scanning of the pallet, capable of collecting visual and volumetric data around the object at 360 degrees. The integration of LIDAR sensors, depth cameras and the use of the ROS2 framework give the platform high capacity for perception, navigation and data acquisition in real time. This study of the application of the robotic system to pallet scanning paves the way for more robust automated traceability of non-conformities or potential fraud in the transportation process. Although the system is in its conceptual phase, its objective demonstrates a high potential for application in shipping in the current context of Industry 4.0.

Acknowledgment

This research was funded by the “Agenda Mobilizadora para a inovação empresarial do setor das Duas Rodas (AM2R)" 01/C05-i01/2021 and LASI-LA/P/0104 /2020, co-funded from the “Mobilising Agendas/Alliances for Business Innovation" of the “Next Generation EU" program of Component 5 of the Portuguese Recovery and Resilience Plan (RRP), concerning “Investment and innovation", under the Regulation of the Incentive System for “Mobilising Agendas/Alliances for Business Innovation".

References

1. Artaso, P., López-Nicolás, G.: Volume estimation of merchandise using multiple range cameras. Measurement **89**, 223–238 (2016). https://doi.org/https://doi.org/10.1016/j.measurement.2016.04.005, https://www.sciencedirect.com/science/article/pii/S0263224116300616
2. Brylka, R., Bierwirth, B., Schwanecke, U.: Ai-based recognition of dangerous goods labels and metric package features pp. 245–272 (2021). https://doi.org/10.15480/882.3959, https://hdl.handle.net/10419/249618, urn:nbn:de:gbv:830-882.0161713; hdl:10419/249608; https://econpapers.repec.org/RePEc:zbw:hiclpr:31
3. Mohamed, I.S., Capitanelli, A., Mastrogiovanni, F., Rovetta, S., Zaccaria, R.: Detection, localisation and tracking of pallets using machine learning techniques and 2d range data. Neural Computing and Applications **32**, 8811–8828 (7 2020). https://doi.org/10.1007/S00521-019-04352-0/FIGURES/7, https://link.springer.com/article/10.1007/s00521-019-04352-0
4. Sasaki, M., Tsuda, Y., Matsushita, K.: Development of autonomous mobile robot with 3dlidar self-localization function using layout map. Electronics **13**(6) (2024). https://doi.org/10.3390/electronics13061082, https://www.mdpi.com/2079-9292/13/6/1082
5. Xiao, J., Lu, H., Zhang, L., Zhang, J.: Pallet recognition and localization using an rgb-d camera. International Journal of Advanced Robotic Systems **14**(6), 1729881417737799 (2017). https://doi.org/10.1177/1729881417737799, https://doi.org/10.1177/1729881417737799

Design and Construction of an Ultrasound-Compatible Kidney Phantom with Dynamic Breathing Simulation

Simão Valente[1,2,3,4], Nuno Rodrigues[1,2,3,4], Pedro Morais[1,4], Estevão Lima[2,3], and João L. Vilaça[1,4]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
spvalente@ipca.pt,
[2] Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal
[3] ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal
[4] LASI – Associate Laboratory of Intelligent Systems, Guimarães, Portugal

Abstract. Kidney imaging is crucial for the diagnosis and monitoring of kidney disease, with ultrasound (US) being a widely used medical imaging modality. Accurate interpretation of US images is a challenging task. Realistic US-compatible phantoms can be useful tools for medical training and testing new technologies. This paper describes the design and construction of a phantom with a mimic-breathing mechanism suitable for US and computed tomography (CT) imaging. To accurately replicate respiratory motion, the phantom uses a balloon-based bi-level positive airway pressure (Bi-PAP) system. An anatomically accurate kidney model is created using tissue-mimicking materials with acoustic properties similar to those of human renal tissue, ensuring fidelity in US and CT imaging. Validation experiments confirm the phantom's ability to simulate respiratory-induced motion artifacts in US imaging. Quantitative analysis demonstrates a strong correlation between the phantom image results and those from clinical datasets, affirming its suitability for research and training purposes. In summary, the developed kidney phantom provides a versatile platform for evaluating and refining US imaging protocols, assessing image quality, and training medical professionals. Its compatibility with US and CT imaging, as well as the ability to simulate realistic respiratory motion with a balloon-based BiPAP system, contribute to improving renal imaging research and education.

Keywords: Kidney imaging · Phantom · Ultrasound.

1 Introduction

Medical imaging plays an important role in the diagnosis, treatment planning, and procedures for a variety of renal diseases. Ultrasound (US) imaging is especially popular because it is safe, inexpensive, and radiation-free. However, because of US characteristics, accurate interpretation of imaging data is heavily dependent on the operator's skills and the quality of the imaging equipment, which can result in an incorrect diagnosis, complications throughout US-based procedures, and a long learning curve that limits US-based applications [4]. Phantoms are idealized tissue models that are used to evaluate clinical imaging, therapeutic device performance, and medical procedures in a controlled test environment without endangering human or animal subjects, making them essential for medical research, planning, testing, and practicing surgical procedures before they are used on real patients. Renal US-based systems, in particular, are frequently developed and tested using phantoms that simulate the in-vivo environment of the human renal anatomy or clinical scenarios relevant to renal imaging and intervention [3]. In this paper, we describe the design and construction of a novel kidney phantom that is compatible with both US and computed tomography (CT) and can simulate human respiratory breathing. Our phantom, which incorporates realistic anatomical features and simulates respiratory movement, provides an invaluable platform for advancing research in medical imaging, image-guided interventions, treatment planning, and medical education.

2 Materials and Methods

The kidney phantom was designed using SolidWorks software to ensure precise replication of renal anatomy. The kidney molds were designed to comprise four parts: the upper and lower models, and two calyces. Additionally, molds for the ureters were constructed using a three-part assembly. This assembly includes the upper, lower, and internal molds, collectively forming a tubular structure that simulates the ureter. An external mold was designed for the construction of the phantom, and another external case was designed to accommodate the phantom and fix the balloon. All the kidney and calyx molds and the water reservoir were fabricated using polylactic acid (PLA) material and produced with a Bambulab 3D printer. The external mold and external case were separately crafted from acrylic material using a laser cutting machine (GN1080; Gbos Laser Technology Company, Dongguan City, Guangdong Province, China).

Several criteria were used to design the kidney phantom, ensuring its efficacy, anatomical realism, and functionality. These criteria included anatomic kidney capsule shape, realistic kidney breathing movement, realistic US image appearance, CT kidney contrast, durability, ease of fabrication, and storage stability at room temperature. Inspired by LIDIA AL-ZOGBI, a mixture composed of 30% ballistic gel 70% Hummimic Gel was used to mimic the characteristics of the human body [1]. Similar to Caldas et al., 5% of the total volume of Glycerin was incorporated into the mixture to enhance realism in US imaging [2]. For the fabrication of the kidney, a silicone mixture comprised of additive HB FLEX 5513 A + B transparent (HBQuimica, Porto, Portugal) was selected. 50% of the total silicone mixture of plasticizer was added to improve the US imaging quality, the similarity with real kidney US imaging, and the flexibility and elasticity of the kidney. In contrast, the ureters were constructed using only the A and B parts to maintain silicone strength. To simulate the kidney's physiological breathing motion, a common 50 cm balloon was utilized in conjunction with a Bi-Level Positive Airway Pressure (BiPAP) GoodKnight 425ST.

The construction of the kidney phantom was carried out as follows: (i) The silicone mixture was prepared by combining 150g of part A, 150g of part B, 300g of plasticizer, and pigment. That was poured into the kidney mold and left to cure for approximately 24 hours; (ii) The ureters were filled with a silicone mixture composed of 50g of part A, 50g of part B, and pigment for approximately 12 hours; (iii The cured ureters were integrated into the kidney by pouring the silicone mixture; (iv) The kidney was fixed to the external mold using PLA fixators in the ureters format. The ballistic gel mixture, comprising 1200g of ballistic gel, 2800g of Humimic gel, and 0.2g of glycerin, was made in a container at 80ºC. After cooling to room temperature, the entire kidney phantom was removed from the external mold; (v) The kidney phantom was assembled into the external casing, along with the balloon.

Acknowledgement

The authors acknowledge the company KARL STORZ SE & Co. KG. 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". This work was funded by the Innovation Pact HfPT – 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". The work was also funded by Foundation for Science and Technology (FCT) - project UIDB/50026/2020 (DOI 10.54499/UIDB/50026/2020), UIDP/50026/2020 (DOI 10.54499/UIDP/50026/2020), LA/P/0050/2020 (DOI 10.54499/LA/P/0050/2020), 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 for their support of this research.

References

1. Al-Zogbi, L., Bock, B., Schaffer, S., Fleiter, T., Krieger, A.: A 3-d-printed patient-specific ultrasound phantom for fast scan. Ultrasound in Medicine & Biology **47**, 820–832 (3 2021). https://doi.org/10.1016/j.ultrasmedbio.2020.11.004, `https://linkinghub.elsevier.com/retrieve/pii/S0301562920304932`
2. Caldas, A., Valente, S., Rodrigues, N.S., Araujo, A.R.D., Storzs, R., Real, A., Ribeiro, R.R., Ferreira, M.R., Morais, P., Matos, D., Vilaca, J.L.: Development of a breast ultrasound phantom for medical training. Proceedings - IEEE Symposium on Computer-Based Medical Systems **2023-June**, 398–403 (2023). https://doi.org/10.1109/CBMS58004.2023.00251
3. McGarry, C.K., Grattan, L.J., Ivory, A.M., Leek, F., Liney, G.P., Liu, Y., Miloro, P., Rai, R., Robinson, A., Shih, A.J., Zeqiri, B., Clark, C.H.: Tissue mimicking materials for imaging and therapy phantoms: a review. Physics in Medicine & Biology **65** (9 2020). https://doi.org/10.1088/1361-6560/abbd17, `https://doi.org/10.1088/1361-6560/abbd17https://iopscience.iop.org/article/10.1088/1361-6560/abbd17`
4. Valente, S., Real, A., Gomes-Fonseca, J., Torres, H.R., Lima, E., Morais, P., Vilaca, J.L.: A new handheld ultrasound probe simulator for medical training. SeGAH 2021 - 2021 IEEE 9th International Conference on Serious Games and Applications for Health (8 2021). https://doi.org/10.1109/SEGAH52098.2021.9551859

Imitation Learning Strategy for enhanced pick and place handling of flexible/rigid objects

Bruno Silva[1], Rafael Fernandes[1], Joaquin Dillen[1], Nuno Simões[1], João M. Faria[1], Luís Vilas Boas[1], Ines Caetano[2], Luís Cardoso[3], Paulo Silva[4], João Borges[1], and António H. J. Moreira[1]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
brsilva@ipca.pt
[2] Sistrade, Porto, Portugal
[3] Miranda, Águeda, Portugal
[4] SRAM, Coimbra, Portugal

Abstract. This paper presents a proposed solution to improve the efficient pick and place of rigid and flexible products by an autonomous mobile manipulator robot that moves between different workstations. This is an area that has been little explored and continues to be a challenge, especially in relation to flexible products, due to their unpredictable and deformable nature. The aim is therefore to explore an imitation learning strategy, thus enabling the robot to be taught how to handle products efficiently based on demonstrations carried out by humans.
In addition to exploring an imitation learning strategy, the use of a mixed tool, consisting of robotic gripper and a vacuum system, will also be explored, assessing whether the vacuum system brings benefits when handling products. As the handling of rigid products is different from the handling of flexible products, a solution will also be explored to distinguish products by their type, thus allowing different forms of handling to be applied to each type of product.
In the end, it is hoped to obtain a solution that allows an autonomous mobile manipulator robot to handle rigid and flexible products with a success rate of over 90%. It is also hoped that the solution will allow the mobile base to be positioned more flexibly at each workstation, being able to achieve an error of up to 20 cm without compromising the success of the pick and place of the products.

Keywords: robot mobile manipulator · pick and place · flexible/rigid objects · imitation learning

1 Introduction

The growing demand for automated and flexible warehouses driven by Industry 5.0, is to foster the exploration of advanced robotic solutions, to create environments where human and robotic systems coexist efficiently [1]. Autonomous mobile manipulators have emerged as a promising technology, offering the ability to navigate between multiple product picking zones, eliminating the need for static layouts that direct all products to a single location. Combining mobile platforms with pick-and-place capability will enable efficient pick-and-place operations at different workstations [2].

In recent years, research on autonomous mobile manipulators has focused mainly on improving navigation systems [4], optimizing pick-and-place strategies [5], and achieving perfect coordination between navigation and object handling [3]. However, the effective manipulation of different types of products or packages, in particular rigid and flexible, with a single system remains an underexplored area. Flexible products and packages pose unique challenges due to their deformable and unpredictable nature, requiring specific, advanced and adaptable handling techniques/strategies when it is fundamental to control its final positioning. While current methods perform well with rigid products, they often fall short when it comes to the variability and complexity associated with flexible products [6].

This project aims to explore a solution for the effective manipulation of different types of product through the use of imitation learning, an emerging and promising approach that allows robots to be taught to perform complex tasks by imitating demonstrations performed by humans.

2 Methodology

Fig. 1 briefly illustrates the logic that will be used to develop this project and how it will be inserted into a real context. The right side of the figure shows the actions that the robot will have to perform in order to achieve the expected result, the efficient pick and place of rigid or flexible products; the left side of the figure shows how the robot is to be integrated into a logistics warehouse system.

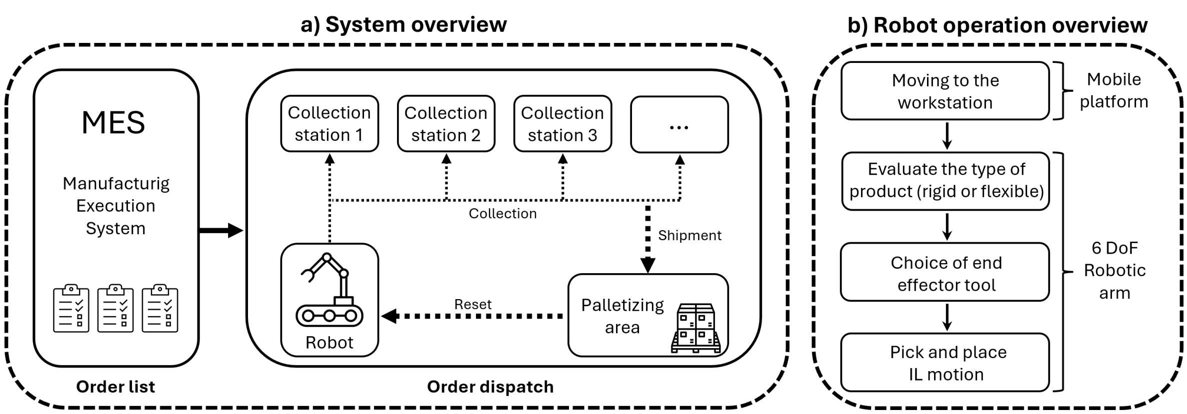


Fig. 1: a) Overview of the robot's operation in a real logistics warehouse context; b) Represents the different phases of the robot's operation to achieve efficient pick and place of rigid or flexible products, where IL mean imitation learning

The project methodology will follow an incremental strategy, divided into four development phases: 1) navigation and control of the mobile platform, 2) data acquisition for product type classification, 3) exploring and applying the imitation learning strategy, 4) development of the mixed tool and simulation in a real context. Each phase of the project consists of development, simulation and validation.

In the first development phase of the project, as the focus of the project is the handling of different products and not the navigation or control of the mobile platform, simple programming based on ROS 2 (Robot Operating System 2) will be used. With the help of the Mid-360 LiDAR, the robot will be able to navigate between workstations efficiently and make fine positioning adjustments when necessary to optimize picking and placing operations. The mobile platform that will be used in this project will have a square shape (1 x 1 m), with a load capacity of 100 kg. Four omnidirectional wheels will be used to move it. The use of this type of wheel provides high maneuverability, allowing movement in any direction without the need to turn the platform, which enables efficient and precise movements in dynamic environments.

In the second phase of the project, data acquisition for product type classification, a synthetic dataset will be created using the Nvidia Isaac Sim platform, complemented with real images of rigid (i.e. boxes) and flexible (i.e. bags) products. Once the dataset has been created, it will be used to train an AI classification model (YOLOv11), thus enabling products to be classified by their type.

In the third phase, which consists of exploring and applying the imitation learning strategy, a robotic arm with 6 degrees of freedom (6DoF) from Universal Robots (UR20) will be used. This phase will also use the Nvidia Isaac Sim platform to explore and apply the imitation learning strategy to the robot. Initially, a demonstration dataset will be created of how the different products should be handled, demonstrations will be carried out using only the robotic gripper and using the mixed tool, to assess which of the two options obtains the best result when the products must be placed in a specific position. Once the set of demonstration data has been obtained, it will be used by imitation learning models to teach the robot how to handle the different products and then evaluate the quality of the robot's learning. At this stage, it may be necessary to apply reinforcement learning alongside imitation learning in order to achieve more efficient learning.

Finally, in the fourth and last phase, when the entire system is functional in the simulation environment, the mixed tool (robotic gripper + vacuum system) will be developed to be attached to the UR20 robotic arm. During this phase, simulations will also be carried out in a real context in order to validate the solution developed with this project.

3 Conclusion

The aim of this project is to find a solution that allows an autonomous mobile manipulator robot to pick and place rigid and flexible products with a success rate of over 90%.

This solution also aims to ensure that an error of up to 20cm in the positioning of the mobile base at each workstation does not affect the success of the pick and place, which is also one of the reasons for using imitation learning, since this strategy allows the mobile platform to be positioned with greater flexibility.

At the end of this project, we hope to obtain an innovative solution that will allow a mobile manipulator robot to pick and place efficiently, especially flexible products, which are the biggest challenge and the main reason for developing this solution.

4 Acknowledgements

This research was funded by the "Agenda Mobilizadora para a inovação empresarial do setor das Duas Rodas (AM2R)" 01/C05-i01/2021 and LASI-LA/P/0104 /2020, co-funded from the “Mobilising Agendas/Alliances for Business Innovation" of the “Next Generation EU" program of Component 5 of the Portuguese Recovery and Resilience Plan (RRP), concerning “Investment and innovation", under the Regulation of the Incentive System for “Mobilising Agendas/Alliances for Business Innovation".

References

1. Dhanda, M., Rogers, B.A., Hall, S., Dekoninck, E., Dhokia, V.: Reviewing human-robot collaboration in manufacturing: Opportunities and challenges in the context of industry 5.0. Robotics and Computer-Integrated Manufacturing **93**, 102937 (2025). https://doi.org/https://doi.org/10.1016/j.rcim.2024.102937, https://www.sciencedirect.com/science/article/pii/S0736584524002242
2. Engemann, H., Du, S., Kallweit, S., Cönen, P., Dawar, H.: Omnivil—an autonomous mobile manipulator for flexible production. Sensors **20**(24) (2020). https://doi.org/10.3390/s20247249, https://www.mdpi.com/1424-8220/20/24/7249
3. Haviland, J., Sünderhauf, N., Corke, P.: A holistic approach to reactive mobile manipulation. IEEE Robotics and Automation Letters **7**(2), 3122–3129 (2022). https://doi.org/10.1109/LRA.2022.3146554
4. Lonklang, A., Botzheim, J.: Mobile robot path planning for unknown static obstacle avoidance by improved rrt* algorithm. In: 2024 10th International Conference on Automation, Robotics and Applications (ICARA). pp. 155–159 (2024). https://doi.org/10.1109/ICARA60736.2024.10553042
5. SEBBATA, W., KENK, M.A., BRETHÉ, J.F.: An adaptive robotic grasping with a 2-finger gripper based on deep learning network. In: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). vol. 1, pp. 620–627 (2020). https://doi.org/10.1109/ETFA46521.2020.9212163
6. Zhu, J., Cherubini, A., Dune, C., Navarro-Alarcon, D., Alambeigi, F., Berenson, D., Ficuciello, F., Harada, K., Kober, J., Li, X., Pan, J., Yuan, W., Gienger, M.: Challenges and outlook in robotic manipulation of deformable objects. IEEE Robotics and Automation Magazine **29**, 67–77 (5 2021). https://doi.org/10.1109/MRA.2022.3147415, https://arxiv.org/pdf/2105.01767

Enhanced Bladder and Prostate Segmentation in CT via Hybrid Contour-Distance Loss

Raul Ferrete Ribeiro[1,2,3], Siobhan Moane[2], Pedro Morais[1,3], and João L. Vilaça[1,3]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
rmribeiro@ipca.pt, pmorais@ipca.pt, jvilaca@ipca.pt
[2] LIFE Research Institute, Department of Applied Science, TUS, Ireland
[3] LASI – Associate Laboratory of Intelligent Systems, Guimarães, Portugal

Abstract. Segmentation of bladder and prostate in Computed Tomography (CT) scans is an important task in clinical practice for aiding in diagnosis, monitoring conditions, and identifying cancerous lesions, which affect around 2.1 million patients suffering from lower urinary tract symptoms. Traditionally, radiologists manually segment these structures on CT scans to identify and locate unusual tissues, shapes, and worrying areas. Despite its importance, it has been considered a backbreaking task due to: i) difficulty, ii) time-consuming, and iii) subject to inter- and intra-observer variability. This research proposes a novel approach for delineating the bladder and prostate in CT scans. The proposed approach consists of: (i) The implementation of a hybrid loss function to enhance bladder and prostate segmentation; and (ii) Comparative analyses of DL architectures to assess the improvement of the segmentation detail. The proposed method showed strong performance on the AMOS22 and PAEC datasets. For the AMOS22 dataset, mean average surface distance (aSSD) for the bladder and prostate were 1.5mm and 1.65mm, respectively, resulting in reductions of 17.13% and 9.34%, corresponding to mean Dice scores of 88.76% and 85.90%. The method on the PAEC dataset yielded mean aSSDs of 1.73mm and 3.26mm with reductions of 66.47% and 36.08% for the bladder and prostate, respectively, with average Dice scores of 90.80% and 80.57%. The obtained results support the suggested method's clinical practice potential and assist doctors in performing screenings of the bladder and prostate.

Keywords: Deep Learning · Segmentation · Distance Map Discrepancy · Bladder · Prostate · AMOS · PAEC

1 Introduction

Lower Urinary Tract Symptoms (LUTS) are estimated to afflict 45.8% of the total population and have increased by 18.4% since 2008 [5]. Nearly 2.1 million LUTS individuals received a new diagnosis of bladder or prostate cancer in 2022, with 618,025 of these cases resulting in a fatal outcome [1]. Therefore, effective treatment and a lower mortality rate depend on early disease detection. Studies emphasized that Computed Tomography (CT) segmentation is an effective and accurate tool for diagnosis and monitoring bladder and prostate [3]. Nonetheless, manual evaluations of CT screenings are often chosen to identify and locate concerning regions. Despite being critical, this process is difficult, time-consuming, and prone to observer variability [2]. In the last decades, Deep Learning (DL) techniques, such as Convolutional Neural Networks (CNNs), superseded the traditional segmentation approaches and gained popularity in the field of computer vision [4]. Several DL networks were adapted for the task of bladder and prostate segmentation in CT images. This research presents a novel approach for segmenting the prostate and bladder in CT images. It explores how segmentation accuracy can be improved by using a loss function that combining Dice coefficients with contours and distance maps. By adding distance maps, our method investigates the spatial link between predicted and real contours, allowing for a more comprehensive assessment of segmentation quality. The suggested method can be used as a useful tool in clinical practice.

2 Methodology

Our approach used DL architectures to generate a segmentation, and contour of the bladder and prostate. From the outputs of the network, a distance map is computed from a binary map containing the predicted contour of the bladder and prostate. The proposed loss function is applied to the networks' outputs and to the distance map to enhance their learning process. (section 2.1).

2.1 Loss Function

A hybrid loss function is proposed in this work to enhance bladder and prostate segmentation, as spatial inconsistencies and imprecise boundaries are common challenges in medical image segmentation. The hybrid loss function comprises three distinct components, each serving a specific purpose: i) Dice Loss, which measures the similarity between predicted segmentation and ground truths; ii) Contour Loss, which focuses on the contours of the structures; and iii) Distance Map Discrepancy Loss (DMD Loss), which aims to improve the segmentation in boundary regions, also having the potential to capture contextual information about the spatial relationships between the structures.

3 Results

Table 1 demonstrates how the proposed loss function performs better in terms of DSC, aSSD, and HD95 measures, when compared to the standard strategy Dice Loss, and the combination of Dice Loss and Contour Loss. DynUnet stood out from the other networks, achieving an aSSD values of 1.57mm and 2.14mm, it specifically achieves 89.38% DSC and 84.29% DSC for the bladder and prostate, respectively. Although it does not vary significantly in relation to the traditional Dice loss function, the proposed approach achieves a remarkable 66.21% reduction in aSSD. These outcomes demonstrate how well the suggested strategy works to improve segmentation accuracy.

4 Conclusion

This research presents a novel loss function for DL-based CT image-based bladder and prostate segmentation. The loss function worked well and was precise in improving the detail of the segmentations. Overall, the results supported the suggested method's clinical practice potential and assisted doctors to perform screenings of the bladder and prostate.

Table 1: Comparative analysis of Loss Functions and DL Methods for AMOS, PAEC, and Combined Datasets

Model	Loss	AMOS (30 volumes)						PAEC (13 volumes)						Total (43 volumes)					
		Bladder			Prostate			Bladder			Prostate			Bladder			Prostate		
		Dice	aSSD	HD95	Dice	aSSD	HD95	Dice	aSSD	HD95	Dice	aSSD	HD95	Dice	aSSD	HD95	Dice	aSSD	HD95
DynUnet	Dice Loss	88.72 ± 10.65	1.81 ± 1.53	8.11 ± 9.95	85.43 ± 11.36	1.82 ± 1.37	5.59 ± 3.36	90.72 ± 7.84	5.12 ± 19.94	12.48 ± 73.92	79.87 ± 9.14	5.10 ± 2.16	10.56 ± 4.59	89.33 ± 9.92	2.81 ± 11.58	9.43 ± 44.65	83.75 ± 11.03	2.81 ± 1.81	7.09 ± 4.00
	Dice Loss + Contour Loss	**89.22 ± 8.82**	**1.32 ± 1.26**	**6.05 ± 8.78**	85.74 ± 12.64	1.70 ± 1.67	5.10 ± 3.46	87.84 ± 14.39	2.15 ± 2.47	7.86 ± 8.65	79.07 ± 10.70	3.60 ± 2.49	9.84 ± 8.34	88.8 ± 10.83	1.57 ± 1.76	**6.60 ± 8.78**	83.72 ± 12.47	2.27 ± 2.14	6.53 ± 5.84
	Dice Loss + Contour Loss + DMD Loss	88.76 ± 9.02	1.50 ± 1.54	7.31 ± 11.34	**85.90 ± 12.93**	**1.65 ± 1.78**	**5.04 ± 3.42**	**90.80 ± 7.91**	**1.73 ± 1.35**	**7.82 ± 7.80**	**80.57 ± 7.07**	**3.26 ± 1.67**	**8.67 ± 3.18**	**89.38 ± 8.75**	**1.57 ± 1.49**	7.47 ± 10.4	**84.29 ± 11.71**	**2.14 ± 1.84**	**6.14 ± 3.48**
SegResNet	Dice Loss	81.67 ± 17.16	2.57 ± 2.24	**8.15 ± 8.44**	81.64 ± 11.52	2.21 ± 1.49	6.73 ± 4.58	73.72 ± 14.73	4.83 ± 2.26	12.23 ± 6.40	68.59 ± 7.45	7.01 ± 1.11	15.04 ± 4.56	79.27 ± 16.63	3.25 ± 2.28	**9.38 ± 8.17**	77.69 ± 10.95	3.66 ± 1.49	9.25 ± 4.68
	Dice Loss + Contour Loss	79.31 ± 17.21	3.27 ± 4.33	11.03 ± 16.13	80.07 ± 12.39	2.67 ± 2.54	7.35 ± 7.00	78.25 ± 17.42	4.63 ± 4.98	**10.78 ± 4.67**	71.84 ± 13.30	4.62 ± 3.30	8.47 ± 3.69	78.99 ± 17.27	3.68 ± 4.54	10.95 ± 14.11	77.58 ± 13.08	3.26 ± 2.86	7.70 ± 6.40
	Dice Loss + Contour Loss + DMD Loss	**82.54 ± 15.71**	**2.49 ± 2.56**	10.16 ± 10.83	**81.73 ± 13.04**	**2.15 ± 1.70**	**6.27 ± 3.84**	**79.59 ± 16.98**	**4.03 ± 4.54**	11.28 ± 6.43	**75.08 ± 15.53**	**3.51 ± 2.30**	**7.57 ± 3.85**	**81.65 ± 16.05**	**2.96 ± 3.23**	10.50 ± 9.93	**79.72 ± 14.26**	**2.56 ± 2.01**	**6.66 ± 3.87**
MedNext	Dice Loss	83.11 ± 12.09	3.11 ± 2.24	**8.43 ± 9.70**	82.70 ± 8.28	2.23 ± 1.29	6.55 ± 4.01	89.75 ± 7.94	1.86 ± 1.04	7.09 ± 3.15	76.89 ± 13.86	3.67 ± 5.12	8.45 ± 14.06	85.12 ± 11.29	2.73 ± 1.99	**8.02 ± 8.45**	80.94 ± 11.67	2.67 ± 3.43	7.13 ± 9.11
	Dice Loss + Contour Loss	84.48 ± 15.95	2.34 ± 3.68	9.69 ± 14.57	**84.88 ± 13.26**	**1.86 ± 1.95**	**5.73 ± 4.90**	**90.84 ± 7.79**	**1.63 ± 1.19**	6.85 ± 5.35	77.56 ± 13.26	4.23 ± 4.33	10.37 ± 12.87	86.40 ± 14.30	2.12 ± 3.16	8.83 ± 12.59	82.67 ± 13.98	2.57 ± 3.09	7.13 ± 8.45
	Dice Loss + Contour Loss + DMD Loss	**86.72 ± 10.26**	**1.88 ± 1.69**	8.86 ± 10.07	84.11 ± 13.42	1.94 ± 1.98	6.00 ± 5.12	89.44 ± 15.21	1.91 ± 2.46	**6.60 ± 5.89**	**80.00 ± 16.71**	**3.15 ± 3.95**	**8.06 ± 3.98**	**87.54 ± 12.02**	**1.87 ± 1.98**	8.18 ± 9.00	**82.87 ± 14.90**	**2.31 ± 2.85**	**6.62 ± 4.95**

Acknowledgment

This work was funded by the Innovation Pact HfPT – 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". The work was also funded by Foundation for Science and Technology (FCT) - projects UIDB/05549/2020 (DOI: 10.54499/UIDB/05549/2020), UIDP/05549/2020 (DOI: 10.54499/UIDP/05549/2020), LASI/LA/P/01
04/2020, and CEECINST/00039/2021 for their support of this research.

References

1. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., Jemal, A.: Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries **74** (2024). https://doi.org/10.3322/CAAC.21834
2. Celik, Y., Talo, M., Yildirim, O., Karabatak, M., Acharya, U.R.: Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images **133** (2020). https://doi.org/10.1016/J.PATREC.2020.03.011
3. Shahait, M., Abu-Hijlih, R., Farkouh, A., Obeidat, S., Salah, S., Abdlkadir, A.S., Al-Ibraheem, A.: Fluorodeoxyglucose positron emission tomography (18f-fdg pet)-computed tomography (ct) in the initial staging of bladder cancer: a single institution experience **35** (2023). https://doi.org/10.1186/S43046-023-00180-5/TABLES/3
4. Xu, X., Fu, L., Chen, Y., Larsson, R., Zhang, D., Suo, S., Hua, J., Zhao, J.: Breast region segmentation being convolutional neural network in dynamic contrast enhanced mri **2018-July** (2018). https://doi.org/10.1109/EMBC.2018.8512422
5. Zhang, A.Y., Xu, X.: Prevalence, burden, and treatment of lower urinary tract symptoms in men aged 50 and older: A systematic review of the literature (2018). https://doi.org/10.1177/2377960818811773

Collaborative Fabric Design with AI and 3D Simulation

João Santos[1], Nuno Dinis[2], Luís Romero[1], and Pedro Miguel Faria[1]

[1] ADiT-Lab, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
jhenriquesantos@ipvc.pt, {romero, pfaria}@estg.ipvc.pt
[2] RIOPELE, Pousada de Saramagos, Vila Nova de Famalicão, Portugal
nuno.dinis@riopele.pt

Abstract. This short paper incorporates the use of artificial intelligence and collaborative visualisation tools to provide an alternative approach to fabric/simulating 3D environments. The main aim is to create a machine learning model that follows the approach by classifying textile parameter combinations (i.e. composition, weight, elasticity) according to their suitability or compliance level to simplify the validation process. Approved combinations are exported to 3D modeling software for virtual visualization and then integrated into Unity, enabling real-time, collaborative interaction with the simulated fabrics.

Keywords: 3D Modeling · Real-Time Collaboration · Textile Simulation · Machine Learning

1 Introduction

Traditionally, the characterization of textile behavior is reliant on physical samples, which is a slow and inefficient process in today's design practice. New technologies and AI offer the opportunity to redesign the process. This short paper presents a system that combines machine learning and 3D simulation to support the validation of fabrics and enhance collaborative practices. Based on fabric properties classified using XGBoost and simulated in CLO 3D with real-time interaction built in Unity, our method enables quicker decision-making and less reliance on physical prototypes.

2 Related Work

A number of research studies have explored alternative approaches to the topic incorporating AI into textile simulation, virtual environments and collaborative environments. Different studies utilize deep learning models, such as convolutional neural networks (CNNs), gated recurrent units (GRUs) and extreme gradient boosting (XGBoost), for analyzing fabric, as well as analyzing wrinkle models [1], fabrics with defects [2] and fabric quality assessment [3]. Other studies investigated customization of avatars in a virtual environment with traditional clothing [4] and simulated tactile sensations with electro-vibration and AI [5].

In numerous experiments participants went for Unity and VR as a substitute for learning and designing, thus showing how client-server architectures and immersive simulations can boost study environments [6] [7].XR-based collaboration with Unity Mirror revealed limitations in scalability and device compatibility [8].

Generally speaking, there is a huge potential for AI in the domain of 3D modeling and collaborative technical that is able to change the textile and fashion industry via enhanced workflows and also more interesting and productive design and training environments.

3 System Architecture

A structured workflow is proposed to enable evaluation and visualization within a 3D collaborative environment.

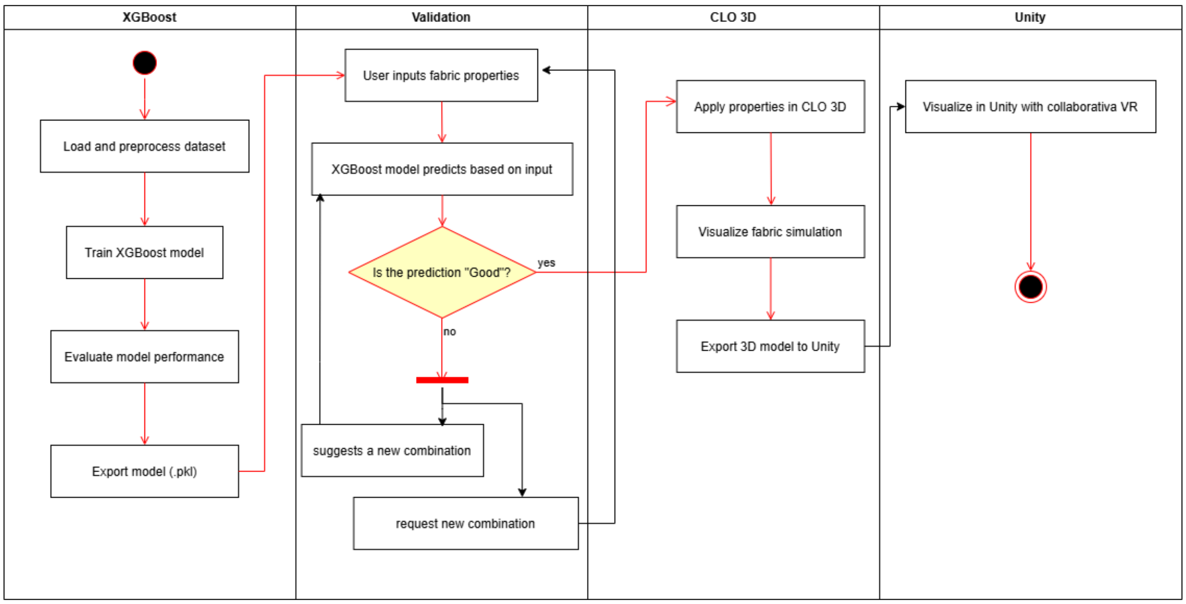


Fig. 1: Proposed workflow for the automatic validation of fabric property combinations using XGBoost, realistic simulation in CLO 3D, and collaborative visualization in Unity.

This architecture allows for the automated validation and realistic simulation of fabric properties in a 3D collaborative environment. The architecture starts with XGBoost, which classifies input by machine learning to find out if the specific textile parameters provide fair combinations of suitability. The verified suitable combinations are then simulated using the software CLO 3D to have the most realistic physical properties for fabrics while being able to utilize incorporated AI and other external data. The items generated in CLO 3D and their properties are then imported into Unity so that users can engage with the simulated fabrics on-location and in real-time using VR/AR technologies and synchronized adjacent environments at their own pace.

4 Conclusion

In conclusion, this short paper details an innovative use case that integrates machine learning and 3D simulation for textiles design and evaluation. Working with XGBoost for classification, CLO 3D for realistic simulated fabric, and Unity for collaborative VR environments, the proposed solution may support product development and potentially reduce reliance on physical samples when applied in practice.

Acknowledgements

This work received financial support from the integrated project Pacto Mobilizador TEXP@CT – Pacto de Inovação para a Digitalização dos Têxteis e do Vestuário (TC-C12-i01, Bioeconomia Sustentável n.º 02/C12-i01/202), promoted by the Recovery and Resilience Plan (PRR), Next Generation EU, for the period 2021-2026.

References

1. H. Yang. 3d clothing wrinkle modeling method based on gaussian curvature and deep learning. In *Proceedings - 2024 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC 2024)*, pages 136–140. Institute of Electrical and Electronics Engineers Inc., 2024.
2. N. Sajitha and S. P. Priya. Artificial intelligence based optimization with extreme gradient boosting for fabric defect detection and classification model. In *2nd International Conference on Sustainable Computing and Data Communication Systems ICSCDS 2023*, pages 509–515. IEEE, 2023.
3. S. A. Shifani, M. D. Suresh, M. Paramaiyappan, M. S. F. T. Selvi, and Giri. Intelligent deep learning model for real-time fabric evaluation using efficientnet, gru and xgboost. In *Proceedings of [conference name not provided]*, 2024.
4. A. Jalori, A. Prajod, A. Pandey, A. Pandey, and P. Kanwal. Enhancing 3d avatar realism with indian clothing and motion capture data. In *Proceedings of CONECCT 2024 - 10th IEEE International Conference on Electronics, Computing and Communication Technologies*. Institute of Electrical and Electronics Engineers Inc., 2024.
5. Y. Chen, W. Qiu, M. Zhang, and X. Wang. 2019 ieee the 2nd international conference on micro/nano sensors for ai, healthcare, and robotics (nsens). IEEE, 2019.
6. X. Luyao, D. Honghai, L. Jianfeng, and Z. Hao. Iicspi: 2018 ieee international conference of safety produce informatization: 10–12 december 2018, chongqing, china. page 475. Institute of Electrical and Electronics Engineers, 2019.
7. H. Huang, S. Wang, Z. Tang, S. Yue, Y. Zheng, and L. Tian. Immersive virtual simulation system for principles of communications based on unity 3d. In *Proceedings - 2024 39th Youth Academic Annual Conference of Chinese Association of Automation (YAC 2024)*, pages 2377–2380. Institute of Electrical and Electronics Engineers Inc., 2024.
8. B. Hall, J. Kessler, O. Edo-Ohanba, J. Collins, H. Zhang, N. Allegreti, Y. Duan, S. Wang, K. Palaniappan, and P. Calyam. Networked and multimodal 3d modeling of cities for collaborative virtual environments. In *Proceedings - 2022 IEEE/ACM 9th International Conference on Big Data Computing, Applications and Technologies (BDCAT 2022)*, pages 204–212. Institute of Electrical and Electronics Engineers Inc., 2022.

Acoustic Voice Analysis: Building a Reference Database for Clinical Applications

Ana Beatriz Oliveira Machado[1], Joana Filipa Teixeira Fernandes[2,3], and João Paulo Teixeira[1,2]

[1] Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal
a52260@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 (IPB), Bragança 5300, Portugal;
{joana.fernandes, joaopt}@ipb.pt
[3] Faculty of Engineering of University of Porto (FEUP), Porto, 4200-465, Portugal

Abstract. Acoustic voice analysis is a central technique in objectively assessing vocal alterations, enabling the non-invasive identification of laryngeal dysfunctions. This paper describes the development of a structured database containing voice samples collected in clinical and academic settings, aimed at supporting computational diagnoses and future research. The recordings are processed using advanced signal analysis algorithms, extracting parameters such as fundamental frequency, jitter, shimmer, and harmonic-to-noise ratio (HNR). This database, associated with physiological metadata (age, sex, height, and weight), represents a significant step toward creating clinical decision support systems, with potential applications in telemedicine and preventive screening.

Keywords: Voice Analysis · Computational Diagnosis · Database.

1 Introduction

The voice is a fundamental tool for communication and expression, and changes in its production can reflect pathological dysfunctions of the vocal tract. The clinical diagnosis of dysphonias and other laryngeal pathologies is traditionally carried out using invasive methods such as laryngoscopy, which, although effective, is expensive, uncomfortable, and not always accessible [1]. In this context, acoustic voice analysis emerges as a promising alternative by enabling functional voice assessment through quantitative parameters extracted from recordings. This non-invasive approach is especially relevant in population screening, monitoring of professional voice users, and therapeutic follow-up [1]. This paper aims to create a voice database, collected in clinical and academic environments, with the following objectives: to build a database, to provide support for future research in clinical phonetics, biomedical engineering, and voice technologies, and to lay the groundwork for implementing computational solutions for vocal diagnosis.

2 Related Work

Several pathological voice databases have been made available. The Saarbrücken Voice Database (SVD) [2] provides sustained vowel recordings from German speakers, but does not include detailed clinical metadata. Fraile et al. [3] developed an automatic classification system in uncontrolled environments but did not include laryngoscopic diagnoses. Bonilha et al. [1] proposed a voice evaluation protocol; however, they did not release a structured, publicly available database. This work, in turn, integrates high-resolution recordings, comprehensive physiological metadata, and includes evaluations by a voice specialist, who reviewed the recordings to confirm the results obtained by the proposed application, offering a more robust resource for research and clinical applications.

3 Methodology

3.1 Voice Collection

Voice recordings were performed following standardised vocal emission protocols. High-sensitivity microphones were used to ensure the integrity of the captured signal. The same equipment was used in all recordings to standardise data collection: an Edirol UA-101 audio interface connected to a TEF 04 microphone. The recordings were made using an initial prototype implemented in MATLAB 2021a (The MathWorks, Inc., Natick, MA, USA), with a sampling frequency of 50 kHz and 16-bit resolution. All participants gave informed consent and were instructed to maintain consistent posture and intensity during the recording. Ethical considerations were carefully observed, with approvals obtained from the Ethics Committees of both Hospital de São João and the Instituto Politécnico de Bragança.

At the time of recording, we also collected four key physiological descriptors: age (years), sex (male/female), height (cm), and weight (kg). These metadata enable demographic stratification (e.g., by age group or body mass index) and support subsequent correlation analyses between voice features and patient characteristics.

The vocal stimuli were adapted from the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) protocol [4], including sustained vowels “a”, “i”, and “u” produced at low, normal, and high pitches, and followed the Saarbrücken Voice Database (SVD) methodology [2]. A CAPE-V sentence (“A Marta e o avô vivem naquele casarão rosa velho”) and a sustained “a” at comfortable pitch for maximum duration were also recorded. The chosen vowels and pitch range reflect key articulatory differences in vocal tract configuration.

3.2 Experimental Procedures

The participants proceeded to the recording phase. Due to the institutional nature of the recording sites, such as hospitals and clinical offices, ideal soundproofing conditions could not always be guaranteed. However, every effort was made to minimize background noise and maintain recording consistency. Participants were instructed to adopt a balanced and upright posture throughout the recording. The study was briefly introduced, and written informed consent was obtained from each participant. The microphone was mounted on a fixed stand, positioned 20 centimeters from the participant’s lips and aligned to ensure a balance between adequate signal-to-noise ratio and avoidance of plosive artifacts, following the recommendations of the European Laryngological Society [5]. The recording tasks were then conducted according to the predefined vocal protocol. The resulting audio data were segmented into separate files, each corresponding to a specific vocal task and recording condition.

3.3 Acoustic Parameter Extraction

After the voice samples were collected, the signals were automatically processed by a custom-developed acoustic analysis system, which extracts a set of descriptive parameters of vocal production. The main parameters analyzed include fundamental frequency (F0), jitter (cycle-to-cycle variation of F0), shimmer (amplitude variation), and harmonic-to-noise ratio (HNR), which are widely used in the literature for the objective assessment of voice quality [6].

At this stage, the system does not automatically classify vocal pathologies. Instead, the extracted parameter values are made available to clinical voice specialists, who interpret the data in conjunction with auditory-perceptual analysis of the recordings to establish a diagnosis. This approach ensures diagnostic reliability while evaluating the potential of the system as a supportive tool in clinical practice.

4 Results

Table 1: Classification of the participants according to their diagnosed vocal pathology.

Group	Female	Male	Total	Average Age	Standard Deviation
IPB	58	89	147	30.69	12.38
FEUP	7	2	9	40.0	12.47
ULS São João					
Clinical Control (ULS São João)	3	1	4	43.8	16.6
Temporomandibular Disorder (TMD)	11	-	11	59.9	15.14
Unilateral Vocal Fold Paralysis (UVFP)	5	3	8	56.0	12.94
Nodular Keratosis (NK)	2	-	2	40.0	7.1
Cordectomy	-	1	1	62.0	-
Chronic Laryngitis	1	-	1	53.0	-
Vocal Polyp	-	1	1	71.0	-
Presbyphonia	-	1	1	45.0	-
Supraglottic Cyst	1	-	1	51.0	-
Total	88	98	186	50.2	12.77

As shown in Table 1, a total of 186 participants were included in the study, distributed across three main groups. Within the ULS São João, 4 participants were classified as clinical controls, while the remaining 26 were identified as presenting a vocal pathology based on otorhinolaryngology specialist evaluation of their recordings.

5 Conclusion

This paper represents a significant advancement in consolidating computational tools for voice analysis. The created database, enriched with comprehensive physiological metadata (age, sex, height, weight) and expert-validated pathology labels, can support healthcare professionals in monitoring vocal quality, detecting early functional alterations, and personalizing vocal treatments.

References

1. H.S. Bonilha, M. Sampaio, G. Madazio, and M. Behlau. The voice evaluation protocol for laryngoscopy: A proposal from a multiprofessional group. *Otolaryngology–Head and Neck Surgery*, 158(3):498–505, 2018.
2. M. Pützer, W.J. Barry, and J.R. Moringlane. The saarbrücken voice database. `http://www.stimmdatenbank.coli.uni-saarland.de`. Accessed: 2024-06-16.
3. R. Fraile, J.I. Godino-Llorente, M. Blanco-Velasco, and G. Cristobal. Automatic acoustic classification of voice disorders: Toward robust detection of pathological voices in uncontrolled environments. *Scientific Reports*, 7(1):16173, 2017.
4. Gail B. Kempster, Bruce R. Gerratt, Katherine Verdolini Abbott, Julie Barkmeier-Kraemer, and Robert E. Hillman. Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*, 18(2):124–132, 2009.
5. P.H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. Van de Heyning, M. Remacle, and V. Woisard. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *European Archives of Oto-Rhino-Laryngology*, 258(2):77–82, 2001.
6. Ingo R. Titze. *Workshop on Acoustic Voice Analysis: Summary Statement.* National Center for Voice and Speech, Denver, CO, 1995.

Development of a Scalable Digital Sample Archive for the Textile Sector

Válter Gomes[1], Pedro Machado[2], Gualter Barros[2], Sérgio Conceição[3] Luís Romero[1], and Pedro Miguel Faria[1]

[1] ADiT-Lab, Instituto Politécnico de Viana do Castelo, IPVC, Portugal
valtergomes@ipvc.pt, {romero, pfaria}@estg.ipvc.pt
[2] CITEVE, Centro Tecnológico das Indústrias Têxtil e do Vestuário de Portugal, Vila Nova de Famalicão, Portugal
{pmachado, gbarros}@citeve.pt
[3] Samplesss, Soluções Digitais para a Indústria da Moda, Porto, Portugal
sergio@sampless.io

Abstract. The textile and fashion industry is transitioning to digital, demanding solutions that improve efficiency, reduce waste, and support sustainability. This paper presents the design and initial deployment of a Minimum Viable Product (MVP) for a Digital Sample Archive platform, developed under the TEXP@CT initiative. This platform has taken the alternative of traditional sampling through a decentralized system that views digital fabric samples in a way that allows for management and mutual use. Thus, it tries to overcome challenges such as digital standards, digital certifications, and the skill gap while integrating with databases of physical properties and blockchain-registered ownership to establish traceability and sustainability across the textile value chain.

Keywords: Digital Samples · Textile Industry · Blockchain · Product Life Cycle Management · Sustainability.

1 Introduction

A great digital transformation is occurring in the fashion and textile industry. Conventional product development operations are time-consuming, drain resources, and dependent upon physical samples that generate waste and inefficiencies in logistics. This set of researches points out the need to leverage digital technologies toward smarter and more sustainable ways in the sector [2]. A few such breakthrough technologies include 3D designing software for prototyping in the early stages and blockchain technologies for the verification and traceability of digital assets [1]. This has opened new dimensions in designing, communicating, and verifying fashion products [5]. In this context, the paper presents the design of an MVP for a Digital Sample Archive that tackles key issues such as lack of standardization, limited digital skills, and absence of reliable certification. The scalable, interoperable platform centralizes sample management, supports 3D software integration, and uses blockchain for secure authentication—paving the way for a more efficient and sustainable digital textile industry.

2 Platform Architecture

This platform architecture emphasizes the chief components central to organizing, storing, and certification of textile samples, including storage modules, database, REST API, Sample Library, and Certification. In creating the system, security, scalability, and usability have always been held the highest priority. Equipped with a JWT-based authentication mechanism backed by secure cookies, the platform carries role-based access control and tokenized file-sharing mechanisms to preserve data privacy and enhance traceable end-user access. Interoperability considerations grant the structuring of metadata to conform with textile data standards, facilitating integration with tools like CLO 3D and Browzwear, with provision for future blockchain warehousing for digital certification.This architecture also ensures modularity and future scalability, allowing seamless updates and integration of new functionalities without disrupting existing services.

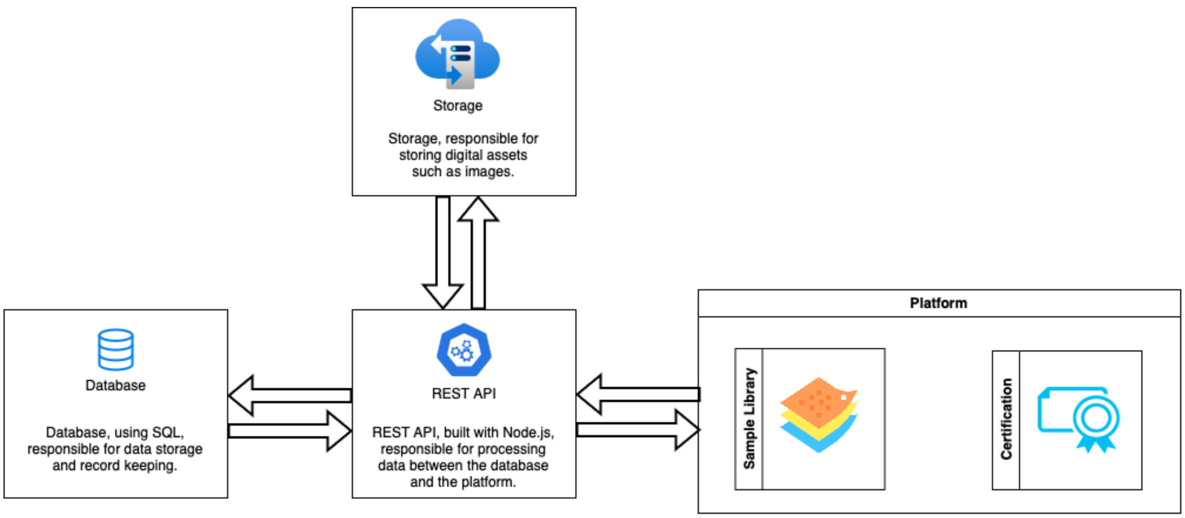


Fig. 1: Platform architecture for the Digital Sample Archive.

Component Descriptions

- **Storage:** Handles secure and efficient storage of digital assets, ensuring data integrity and easy access.
- **Database:** Powered by PostgreSQL [6], it stores all records—samples, users, and certifications—with scalability and security.
- **REST API:** Built with Node.js [4], it enables secure, efficient communication between the database and front-end.
- **Sample Library:** React-based [3] interface for users to manage, edit, and organize textile samples.
- **Certification:** Enables regulatory bodies to validate and certify samples against required standards securely.

3 Conclusion

The paper describes the design and development of a Minimum Viable Product (MVP) for the Digital Sample Archive platform that aims to foster digitization within the textile industry. Nowadays, textile sample management has become increasingly digital to reduce dependency on physical prototypes and encourage sustainable production. Hence the Digital Sample Archive provides an end-to-end solution for the management, archiving, and certification of digital textile samples: integrating an SQL database, scalable storage, RESTful API, with custom modules for sample handling and certification, thereby resolving major issues faced by digital product lifecycle management. The proposed architecture and functionalities have been validated through internal reviews and stakeholder evaluations from the textile domain, though a formal validation and user testing phase is still to come. Valuable feedback was gathered during the initial demos, especially related to the certification workflow, which has been considered during development. The next stages will be very structured pilots with industry partners to evaluate usability, interoperability and scalability in practical, operational contexts. This experience will inform further modifications and help to align with the operational requirements of the textile value chain to prepare the platform for large scale implementation. Enhancements will include linking physical property measurements databases, better interoperability with 3D design software, and blockchain-based certification, and the work will establish a solid technical base to promote the adoption of digital workflows as part of a more integrated, effective, and sustainable textile industry.

Acknowledgement

This work received financial support from the integrated project Pacto Mobilizador TEXP@CT – Pacto de Inovação para a Digitalização dos Têxteis e do Vestuário (TC-C12-i01, Bioeconomia Sustentável n.º 02/C12-i01/202), promoted by the Recovery and Resilience Plan (PRR), Next Generation EU, for the period 2021-2026.

References

1. Dai, N., Li, L., Xu, K., Lu, Z., Hu, X., Yuan, Y.: Development of a standardized data collection and intelligent fabric quality prediction system for the weaving department. Journal of Engineered Fibers and Fabrics **20**, 15589250241312778 (2025). https://doi.org/10.1177/15589250241312778, `https://doi.org/10.1177/15589250241312778`
2. Matos, J., Cunha, A., Faria, A., Vale, B., Caldas, L., Monteiro, M., Cunha, D., Meneses, G., Pinto, R., Barros, G., Paixão, G., Duarte, O.: Digital product: A textile use case. In: 2024 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC). pp. 1–8 (2024). https://doi.org/10.1109/ICE/ITMC61926.2024.10794226
3. Meta Platforms, Inc.: React – A JavaScript library for building user interfaces. `https://reactjs.org/` (2024), version 18.x
4. OpenJS Foundation: Node.js. `https://nodejs.org/` (2024), version 20.x
5. Petreca, B., Baurley, S., Bianchi-Berthouze, N., Tajadura-Jiménez, A.: Investigating nuanced sensory experiences in textiles selection. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. p. 989–994. UbiComp '16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2968219.2968264, `https://doi.org/10.1145/2968219.2968264`
6. PostgreSQL Global Development Group: PostgreSQL. `https://www.postgresql.org/` (2024), version 16.x

Purchase Behavior at Starbucks in Malaysia

Cristiana Silva[1], Alison Gonçalves[1], Inês Fernandes[1], and Elizabeth Fernandes[1,4]
Helena Sofia Rodrigues[2,3]

[1] Escola Superior de Ciências Empresariais, Instituto Politécnico de Viana do Castelo, IPVC, Portugal
{crafaelasilva,alisongoncalves,inesfernandes}@ipvc.pt, elisabeth.ferna@gmail.com
[2] ADiT-Lab, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
[3] CIDMA - University of Aveiro, Portugal
sofiarodrigues@esce.ipvc.pt
[4] Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal

Abstract. This study examines the consumer behaviour at Starbucks in Malaysia, a market that presents cultural diversity and high competition. By performing a quantitative analysis, interesting consumer insights were revealed to inform business data-driven decisions. The findings show that most customers are young adults with low income, visit Starbucks occasionally, and spend modestly. Preferences are split between dine-in and take-away, with short visits being common. Coffee is the item that is the most purchased, and promotions are found mainly on social media. Overall satisfaction is positive, especially with the service and atmosphere. Correlation analyses reveal distinct customer profiles and key factors influencing loyalty.

Keywords: Starbucks · Consumer Behavior · Data Analysis · Data-driven · Marketing Analytics.

1 Introduction

Consumer buying behavior is a key topic in marketing, especially for global brands operating in culturally diverse markets. The continued expansion of Starbucks in the country raises important questions about local consumer motivations, perceptions, and satisfaction, particularly in a competitive market where personalised experiences and innovative promotions are increasingly valued.

According to a Columbia Business School study [1], the opening of Starbucks stores in underserved areas increased local start-up activity by 11.8%, highlighting the brand's role beyond coffee, as a place to work, study and social interaction. Understanding how Malaysian consumers interact with Starbucks is essential for developing effective positioning and adaptation strategies. A different study by revealed that service quality has a significant impact on customer satisfaction at Starbucks in Malaysia, and most consumers prefer high-quality service, which in turn enhances brand loyalty.

Furthermore, another study demonstrated that brand reputation positively influences customer satisfaction at Starbucks in Malaysia [2], indicating that a strong and reliable brand image is correlated with higher levels of customer satisfaction. This study aims to analyze Starbucks consumer behavior in Malaysia by (i) profiling customer demographics; (ii) identifying patterns of visit frequency, duration, and spending; (iii) assessing satisfaction with different aspects of the experience; and (iv) examining the influence of promotions, ambiance, and pricing on purchase decisions.

2 Methodology

This study adopts a quantitative research approach to analyze consumer buying behavior at Starbucks in Malaysia. The data used for this research was obtained from a publicly available Kaggle dataset [3], which included 122 responses. The dataset provides a relevant and structured overview of consumer behavior, allowing for statistical analysis of trends and correlations. Descriptive statistics were used to profile consumers and summarize key behavioral indicators, this included frequencies, means, and standard deviations. Although the Likert scale is ordinal, it is common to use the mean and standard deviation in studies with reasonable sample sizes ($n > 100$), assuming an interval approximation. In addition, inferential analyzes, including correlation heatmaps, were conducted to explore the relationships between satisfaction drivers and behavioral patterns. A significance level of $p < 0.05$ was adopted for correlation analyses. All data analysis was performed using IBM SPSS Statistics version 29 software to ensure accuracy and reliability.

3 Main results

The sample consists of a total of 122 respondents, of which 65 are female (53.3%) and 57 are male (46.7%). Regarding age, the majority of respondents are between 20 and 29 years old. In addition, 50% of the respondents are employed, followed by 34.4% who are students. A smaller percentage, are self-employed or housewives. When asked about membership card ownership, the responses were almost evenly split: 50.8% of the respondents stated that they do not have a Starbucks card.

To assess consumer satisfaction with the perceived experience at Starbucks Malaysia, a 5-point Likert scale was used, where 1 represented the most negative rating and 5 the most positive. Statistical analysis was performed using SPSS software version 29.

Table 1: Customer Satisfaction Ratings at Starbucks Malaysia

Variable	**Mean**	**St. Dev.**
Rating of Starbucks compared to other brands	3.66	0.941
Rating of Starbucks' price range	2.89	1.082
Importance of discounts and promotions in the purchase decision	3.80	1.090
Evaluation of the environment (lighting, music, etc.)	3.75	0.930
Evaluation of Wi-Fi quality	3.25	0.958
Evaluation of service (speed, friendliness, etc.)	3.75	0.829
Likelihood of choosing Starbucks for work meetings or social gatherings	3.52	1.030

Table 1 presents the mean satisfaction ratings and standard deviations for various aspects of the Starbucks experience in Malaysia. Consumers reported the highest satisfaction with promotions and the general environment (means of 3.80 and 3.75, respectively), while the price range received the lowest rating (mean of 2.89). These results highlight the importance of ambiance and promotional strategies in shaping customer satisfaction, as evidenced by the positive reception of Starbucks Malaysia's store atmosphere and marketing efforts; however, the low price rating suggests that pricing may still represent a barrier for some consumers, indicating the need for more attractive pricing strategies or value-based promotions. Furthermore, the high service rating (mean score 3.75) reflects the appreciation of customers for the friendliness and efficiency of staff.

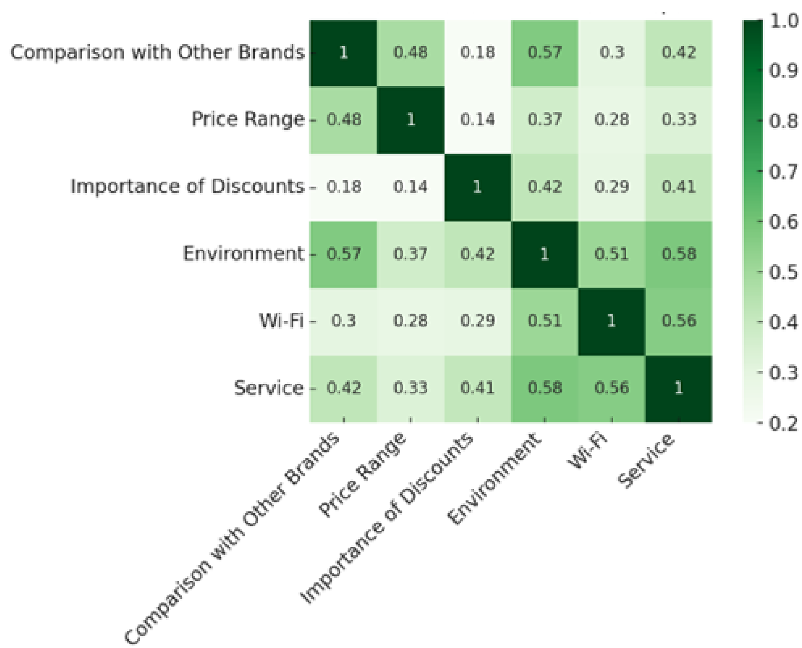


Fig. 1: Heatmap of consumers' perception of Starbucks

Fig. 1 shows the consumers' perception of Starbucks based on Spearman's Correlation. The strongest correlation was identified between 'Environment' and 'Service' provided (r = 0.58), suggesting that consumers who rate the environment positively also tend to perceive the service in a similarly positive way. This association indicates that strategies focused on improving the environment, such as music, lighting, and decor, can have a direct impact on perceived quality of service, thereby enhancing the overall in-store experience. Likewise, a significant correlation was found between the evaluation of the 'Environment' and the 'Comparison with other brands' (r = 0.57), suggesting that consumers who give high ratings to Starbucks' environment also tend to view the brand more favorably compared to competitors. On the other hand, the importance attributed to 'Discounts' shows weaker correlations with the other variables, especially with 'Comparison with other brands' (r = 0.18) and 'Price Range' (r = 0.14), revealing that this factor acts more independently in the purchasing decision.

That said, Spearman's heatmap analysis reveals patterns of association that are important for understanding consumer preferences, allowing us to identify which aspects of the experience have the greatest joint impact on the perception of Starbucks.

4 Conclusion

The results show that Starbucks consumers in Malaysia are predominantly young, low-income consumers who visit Starbucks occasionally, stay for short periods, and spend low amounts. While price may be a limitation, customers highly value promotions, service, and atmosphere. Strategies should prioritize service quality, ambience enhancement, and tailored promotions. Future research should include broader and more varied samples and consider tracking consumer behavior over time to understand changing trends. The insights provided can help Starbucks refine its positioning and customer engagement strategies within the Malaysian market.

References

1. Aryani, D. N., Singh, P., Liew, Y. W., Kee, D. M. H., Li, Y., Li, J., Lim, C. Y., & Arif, W. (2021). How brand reputation influences customer satisfaction: A case study of Starbucks, Malaysia. International Journal of Tourism & Hospitality in Asia Pacific, 4(3), 60-72.
2. Choi, J., Guzman, J., & Small, M. L. (2024). Third Places and Neighborhood Entrepreneurship: Evidence from Starbucks Cafés. NBER Working Paper No. 32604, Revised October 2024. National Bureau of Economic Research.
3. `https://www.kaggle.com/datasets/mahirahmzh/starbucks-customer-retention-malaysia-survey`, acessed on April, 2025

YOLO-Based CNN Model for Detecting Olive Knot Disease from UAV Images

Maurício Morais[1,2], João Mendes[1], José Lima[1], and Murillo Ferreira dos Santos[2]

[1] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança (IPB), Bragança, Portugal
mauricio.fofano@gmail.com, {joao.cmendes, jllima}@ipb.pt
[2] Department of Electroelectronics, Federal Center for Technological Education of Minas Gerais, Brasil
murillo.ferreira@cefetmg.br

Abstract. This paper proposes a practical methodology for detecting and localizing olive knot disease in olive groves using UAV imagery and Computer Vision. By leveraging the YOLO Deep Learning Model, the approach aims to quickly and accurately identify infected trees from UAV collected images. Integrating affordable UAV technology with Machine Learning enables early detection and more efficient disease management in olive groves, all the while improving crop management and complementing human labor.

Keywords: Plant Disease Detection · YOLO · UAV.

1 Introduction

Precision Agriculture (PA) leverages Unmanned Aerial Vehicles (UAVs) and Computer Vision (CV) to optimize crop management, enabling data-driven decisions that maximize yields while minimizing resource waste and environmental impact [1,5]. In recent years, UAVs equipped with high-resolution cameras have become powerful tools for monitoring plant health and detecting diseases, offering rapid and scalable coverage of agricultural fields [2].

In Portugal, olive cultivation is of significant economic and cultural importance, with the country ranking among the world's leading olive oil producers [4]. However, the sector faces major challenges from diseases such as the olive knot, a bacterial infection that can cause severe yield losses and diminish olive oil quality [3]. The symptoms of this disease are often concealed within dense tree canopies, making early detection particularly complex.

This work aims to address these challenges by developing a YOLO-based deep Convolutional Neural Network (CNN) model to detect the olive knot disease by using UAV imagery, while also providing GPS localization for rapid response. Combining affordable and low-cost UAV technology with state-of-the art CV, this approach seeks to enable early detection of infections and support a more effective management of the disease in olive groves.

2 Methodology

The main objective of this work is to develop an artificial intelligence system capable of detecting olive knot disease in olive groves using UAV imagery. Fig. 1 presents the system architecture. To achieve this, it was first necessary to assemble a comprehensive dataset of olive knot disease instances. This process involved several field visits to olive groves, where videos were captured at full HD resolution (1920×1080) during manual UAV flights over the olive trees. The collection was carried out with the assistance of PhD students, ensuring that the dataset reflected the variability found in real-world conditions, such as complex background, fluctuating lighting, and the frequent presence of leaves partially obscuring the disease symptoms. Images were subsequently extracted from these aerial videos for annotation, resulting in a final dataset of 2050 annotated images.

Once the images were gathered, the next critical step was annotation. Given the nature of olive knot disease, where knots often appear clustered together on trunks and branches, special care was taken during annotation to label each knot individually. This approach was adopted to prevent the model from interpreting multiple adjacent knots as a single instance, which would negatively impact detection accuracy and evaluation metrics. The annotation process was conducted by careful visual inspection, as illustrated in Fig. 1, where two closely positioned knots could easily be mistaken for one if not properly annotated.

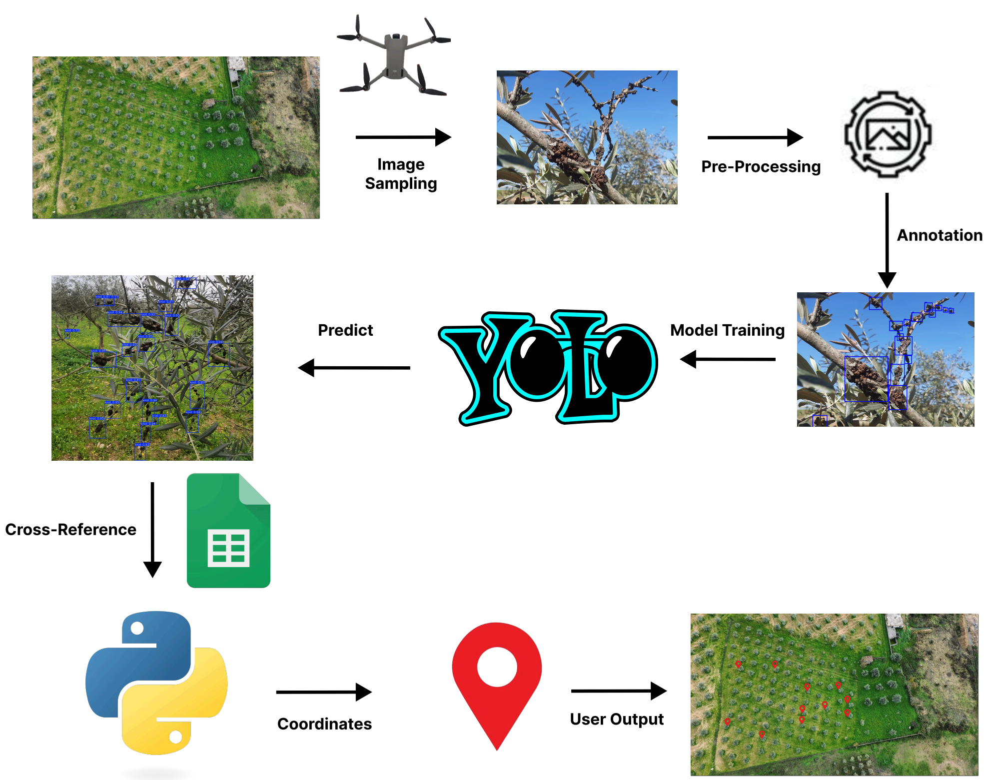


Fig. 1: System architecture.

Following dataset annotation, YOLO model training will proceed.This model was selected for its lightweight architecture, which makes it well-suited for deployment on resource-constrained devices. The training process will involve dividing the dataset into training, validation and test subsets, enabling evaluation and refinement of the model's performance. Once the model is trained, offline inference will be performed on a local computer to detect instances of olive knot disease. The detections will then be cross-referenced with the UAVs flight log data to determine the GPS coordinates of the UAV in the moment it detected a infected tree, giving a approximate location of the tree. This approach is designed to facilitate rapid and precise localization of disease outbreaks, ultimately providing producers with actionable information to help contain and manage the disease in their groves.

3 Conclusion and Future Work

This paper presents a methodology for developing an artificial intelligence system to detect olive knot disease in olive groves using UAV imagery and deep learning. Leveraging affordable UAV technology and lightweight CNN architectures like YOLO, this work has the potential to make advanced plant health monitoring more widely available to olive producers.

As hardware continues to improve and become more affordable, particularly with advances in GPUs and the increasing accessibility of UAVs, artificial intelligence methods are poised to deliver even greater impact across agriculture. Building on the methodology detailed in this paper, future work will focus on expanding the dataset with new images of the disease and testing the developed model under both real-time and real-world conditions.

References

1. Benos, L., Tagarakis, A.C., Dolias, G., Berruto, R., Kateris, D., Bochtis, D.: Machine learning in agriculture: A comprehensive updated review. Sensors **21**(11), 3758 (2021)
2. Bouguettaya, A., Zarzour, H., Kechida, A., Taberkit, A.M.: Recent advances on UAV and deep learning for early crop diseases identification: A short review. In: 2021 International conference on information technology (ICIT). pp. 334–339. IEEE (2021)
3. European Innovation Partnership for Agricultural Productivity and Sustainability (EIP-AGRI): Pests and diseases of the olive tree (2019), `https://ec.europa.eu/eip/agriculture/sites/default/files/eip-agri_fg33_pests_and_diseases_olive_tree_starting_paper_2019_en.pdf`
4. Food and Agriculture Organization of the United Nations (FAO): FAOSTAT Statistical Database (2023), `https://www.fao.org/faostat/en/#data/QCL`, accessed: 2025-01-16
5. Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D.: Machine Learning in Agriculture: A Review. Sensors **18**(8) (2018). https://doi.org/10.3390/s18082674, `https://www.mdpi.com/1424-8220/18/8/2674`

Surface Classification to Optimize Path Planning in Autonomous Mobile Robots

João M. Faria[1−3], João Borges[1,2], Luis Cardoso[4], John Cosgrove[3], and António H. J. Moreira[1,2]

[1] 2Ai – School of Technology, IPCA, Barcelos, Portugal
{jpfaria, jpbsilva, amoreira}@ipca.pt
[2] LASI – Associate Laboratory of Intelligent Systems, Guimarães, Portugal
[3] IDEAM Cluster, TUS, Limerick, Ireland
john.cosgrove@tus.ie
[4] Miranda & Irmão LDA, Águeda, Portugal
lcardoso@miranda.pt

Abstract. The integration of mobile robots into dynamic environments such as industrial facilities, healthcare institutions, and agricultural fields has grown significantly in recent years. For these systems to operate safely and efficiently alongside humans, it is essential that they are capable of adapting to changing environmental conditions. One of the critical challenges lies in perceiving and responding to different surface types — such as wet, dry, soft, or textured terrains — which directly impact locomotion stability and navigation safety. This paper investigates the application of artificial intelligence techniques for real-time surface classification to enhance the adaptive behavior of autonomous mobile robots (AMRs). Drawing inspiration from previous work that using vision-based convolutional neural networks and inertial sensor data for indoor surface recognition, this study aims to broaden the scope to include both indoor and outdoor terrains. By developing a robust classification model capable of identifying diverse surface conditions, the proposed approach enables dynamic adjustment of navigation parameters, including speed and trajectory. The outcomes of this work are expected to contribute to the advancement of intelligent navigation strategies in unstructured environments, increasing the operational resilience and safety of AMRs across multiple sectors. This capability represents a step forward in the digital transformation of mobile robotics, promoting more autonomous, context-aware systems for sustainable applications.

Keywords: Surface Classification · Optimized Route · Autonomous Robot Mobile (AMR).

1 Introduction

In recent years, mobile robotic systems have been increasingly integrated into multiples domains such as industry, agriculture, and healthcare, often working collaboratively alongside humans [2, 7]. As these robots become more present in dynamic and human-centered environments, their ability to adapt to changing conditions becomes critical to ensuring safety and operational efficiency [3, 4]. A key aspect of this adaptability lies in their capacity to perceive and respond to the physical properties of the surfaces they navigate — distinguishing between dry or wet, hard or soft, or textured terrain [5]. Asiye Demirtaş et al. [1] present a study for classification of indoor surfaces (carpet, tiles, and wood) for autonomous mobile robots (AMR). To collect the data, they used an RGB camera and convolution neural networks (CNNs) to classify the surface. David Tick et al. [6] classified the indoor surface using data collected from an inertial measurement unit (IMU) attached to the robot, extracting velocity and acceleration. In this study they analysed 5 types of surfaces (carpet, tile, linoleum, ceramic tiles-A, ceramic tiles-B) and used linear Bayes to classify the surfaces. This surface detection enables autonomous mobile robots (AMRs) to adjust their behavior in real time, such as modifying their speed, acceleration, or trajectory, thereby improving their stability and ensuring the safe transport of goods and interaction with humans [8, 9].

Therefore, this work explores the use of Artificial Intelligence (AI) to enhance the perception and adaptive capabilities of AMRs operating in unstructured environments. The goal is to develop intelligent navigation strategies that enable mobile robots to operate more effectively and safely across different indoor and outdoor surfaces. By focusing on real-time surface classification and response, this research aims to improve the performance of AMRs in sectors where flexibility, resilience, and human collaboration are essential for sustainable progress, such as agriculture, healthcare, and industry.

2 Expected Outcomes

With this research we hope to develop an algorithm capable of classifying different environmental conditions, namely surface characteristics, allowing the identification of different topologies (wet, dry, smooth, gravel) in multiple sectors, such as industry, health and agriculture. This capability will allow the autonomous platform to adjust parameters such as speed and movement trajectory according to the environment. This advancement is crucial for digital transformation, enhancing autonomous navigation with greater safety and reliability, Fig. 1.

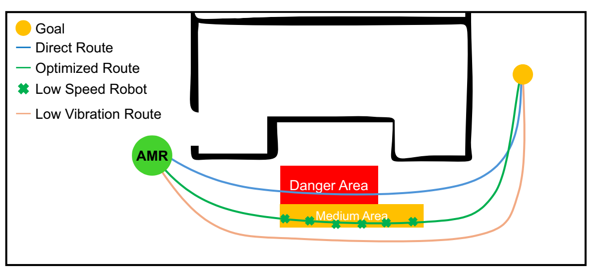


Fig. 1: Example of a route with a risk area, possible route optimization with different behavior along the path (simplified representation of the project objective).

Acknowledgment

This research was funded by the "Agenda Mobilizadora para a inovação empresarial do setor das Duas Rodas (AM2R)" 01/C05-i01/2021, co-funded from the "Mobilising Agendas/Alliances for Business Innovation" of the "Next Generation EU" program of Component 5 of thePortuguese Recovery and Resilience Plan (RRP), concerning "Investment and innovation", under the Regulation of the Incentive System for "Mobilising Agendas/Alliances for Business Innovation". This work also was financed by the financed by national funds, through FCT - Foundation for Science and Technology and FCT / MCTES under the project UIDB/05549:2Ai, UIDP/05549:2Ai and LASILA/P/0104/2020.

References

1. Erdemir, G., Demirtaşdemirtaş, A., Bayram, H.: Indoor surface classification for mobile robots. PeerJ Computer Science **10**, e1730 (2024). https://doi.org/10.7717/peerj-cs.1730, https://doi.org/10.7717/peerj-cs.1730
2. Markis, A., Papa, M., Kaselautzke, D., Rathmair, M., Sattinger, V., Brandstötter, M.: Safety of mobile robot systems in industrial applications. In: ARW OAGM Workshop. pp. 26–31 (2019). https://doi.org/10.3217/978-3-85125-663-5-04
3. Robotnik: The rise of machine learning robots: Explore machine learning in robotics. https://robotnik.eu/the-rise-of-machine-learning-robots-explore-machine-learning-in-robotics/ (2024), accessed: 2024-04-14
4. Robotnik: What is adaptive robotics? https://robotnik.eu/what-is-adaptive-robotics/ (2024), accessed: 2024-04-14
5. Rubio, F., Valero, F., Llopis-Albert, C.: A review of mobile robots: Concepts, methods, theoretical framework, and applications. Journal of Intelligent and Robotic Systems **16**(2) (2019). https://doi.org/10.1177/1729881419839596, https://doi.org/10.1177/1729881419839596
6. Tick, D., Rahman, T., Busso, C., Gans, N.: Indoor robotic terrain classification via angular velocity based hierarchical classifier selection. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). pp. 3594–3600 (2012). https://doi.org/10.1109/ICRA.2012.6225128, https://doi.org/10.1109/ICRA.2012.6225128
7. World Robotics 2022: Market presentation world robotics 2022 (2022)
8. Xue, F., Hu, L., Yao, C., Liu, Z., Zhu, Z., Jia, Z.: Sound-based terrain classification for multi-modal wheel-leg robots. In: 2022 7th IEEE International Conference on Advanced Robotics and Mechatronics (ICARM). pp. 174–179 (2022). https://doi.org/10.1109/ICARM54641.2022.9959511
9. Yang, M., Yang, E.: Two-stage multi-sensor fusion positioning system with seamless switching for cooperative mobile robot and manipulator system. International Journal of Intelligent Robotics and Applications **7**(2), 275–290 (2023). https://doi.org/10.1007/s41315-023-00276-0, https://doi.org/10.1007/S41315-023-00276-0

Intelligent Agent to enable Effectiveness in Industry 5.0

Eduardo Gonçalves[1] and Luis Ferreira[1]

2Ai – School of Technology, IPCA, Barcelos, Portugal
{emgoncalves, lufer}@ipca.pt

Abstract. In this paper, the question is explored: "What if our machines could genuinely understand and adapt to the humans who run them?". This question is especially important in industrial environments that are starting to transition from a state where machines were used mainly with the objective of maximizing efficiency and automation (in Industry 4.0), to a more sustainable, resilient and human-centric state, introduced by Industry 5.0. To support this transition, a technology agnostic architecture is proposed, that builds upon reference architectures developed within the Industry 4.0 context (e.g. Reference Architectural Model Industrie 4.0) and extends them by integrating Industry 5.0 standards and embedding Artificial Intelligence systems capable of dealing with ubiquitous-like human insight. This architecture introduces a new layer to maintain the human in the loop and stimulate meaningful human-machine collaboration.

Keywords: Ubiquitous human insight · Human in the loop · Industry 5.0 · Technology agnostic architecture · Human-machine collaboration.

1 Introduction

In modern industrial environments, a huge variety of heterogeneous systems is found, ranging from newly Internet of Things-enabled machinery to decades-old legacy Programmable Logic Controllers [5]. Each of these systems was designed to optimize or resolve a specific problem, making them critical to operations and difficult to replace. As a result, even though they were not originally designed for it, a strong need has been recognized in the industry for these different systems to be integrated so that communication among them can be achieved. This system interoperability, despite being essential, is also considered incredibly difficult, as different data formats or incompatible protocols are known to increase integration time, cost, and risk [1]. With this in mind, architectures based on reference models developed during the Industry 4.0 era (e.g., the Reference Architectural Model Industrie 4.0 or the Industrial Internet Reference Architecture) are typically relied upon by these industries [6]. These models are considered helpful up to a certain point, as they define data and semantic layers, standard communication protocols, and process cycles [4].

Yet the linchpin is missed: by assuming that the unidirectional flow from sensors → analytics → actuators is sufficient, the inclusion of a human feedback mechanism is neglected. So, while these architectures are regarded as well structured, "true effectiveness" is not achieved (that is, the enhancement of expected human-to-machine services [3] with pragmatic services, essential for human-to-human understanding, is absent). As a result, adaptability, transparency, and resilience, qualities only made possible through human interaction, are not fully realized [2].

Known problems of this:

- **Human Input Middleware Absence** - There's no standard endpoint for operators to share their observations;
- **Rigid automations** - In the absence of a feedback loop, systems cannot adapt to operator observations;
- **Fragmented situational awareness** - Human feedback lives outside of the core architecture affecting decision making.

So, if true effectiveness is to be achieved in these architectures, they must be evolved to embed human-in-the-loop capabilities.

2 Proposed Architecture

In Fig. 1 below, the proposed architecture is presented, structured by layers with the intent of achieving effectiveness, scalability and security if properly implemented. It is divided in 5 layers, including an interoperability middleware and an intelligent agent designed for assistance. The architecture is suitable for deployment in modern industrial environments (e.g. open-mining industry).

The layers are:

- **Data Collection Layer**: sensors or systems (e.g. smart tractors/machines) capturing and transmitting raw data;
- **Interoperability Layer**: matches the interoperability middleware;
- **Data Management Layer**: consists of multiple databases that organize the data collected via the data collection layer;
- **Service Layer**: composed of microservices capable of exposing analytics and APIs;
- **Presentation & Interaction Layer**: contains dashboards and collaboration channels enhanced via the intelligent assistant.

The **Middleware** is used to anchor the architecture by allowing interaction with heterogeneous systems via protocol and data standardization. In addition to this, it also guarantees security to the entire ecosystem.

The **Intelligent Agent**, is trained on the dataset assembled via the middleware and its function is designed as supporting human in the loop interaction by offering real time recommendations. The insight sharing becomes easier, the Artificial Intelligence models can be corrected and machine failures can be predicted.

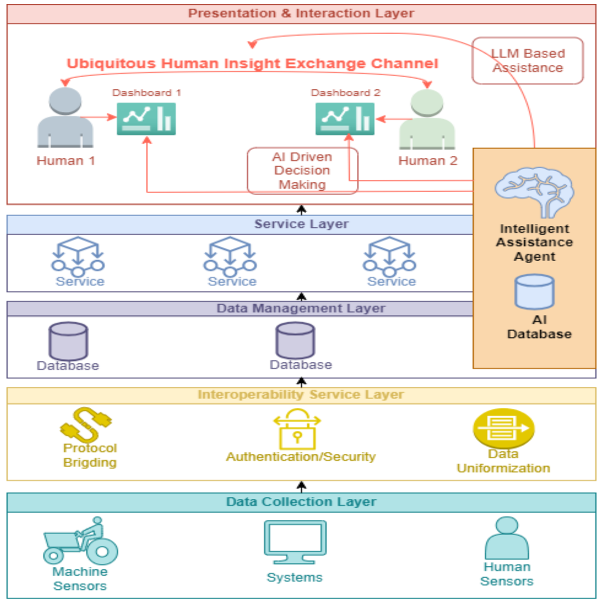


Fig. 1: Layered architecture.

3 Conclusion

While Industry 4.0 architectures such as RAMI 4.0 and IIRA have enabled machine connectivity and automation, the human operator has often been positioned outside the core decision-making loop. In the proposed architecture, a presentation and interaction layer, along with an intelligent assistance agent, has been introduced to embed human feedback directly into operations. This allows continuous insight sharing and real-time adaptation of AI-driven decisions, supporting the shift toward the human-centric principles of Industry 5.0.

References

1. Boyes, H., Hallaq, B., Cunningham, J., Watson, T.: The industrial internet of things (iiot): An analysis framework (2018)
2. Ferreira, L., Putnik, G.D., Cruz-Cunha, M.M., Putnik, Z., Castro, H., Alves, H., Shah, V., Varela, L.: A cloud-based architecture with embedded pragmatics renderer for ubiquitous and cloud manufacturing (2017)
3. Gorecky, D., Schmitt, M., Loskyll, M., Zühlke, D.: Human-machine-interaction in the industry 4.0 era (2014)
4. Jaloudi, S.: Communication protocols of an industrial internet of things environment: A comparative study (2019)
5. Korodi, A., Nițulescu, I.V., Fülöp, A.A., Vesa: Integration of legacy industrial equipment in a building-management system industry 5.0 scenario (2024)
6. Menghi, C., Kolagari, R.T., Trujillo, L., Ghezzi, C.: Industry 4.0 reference architectures: State of the art and future trends (2021)

Textile Pattern Design - Creative Platform

Diogo Araújo[1], Helena Silva[2] Luís Romero [1], and Pedro Faria [1]

[1] ADiT-Lab, Instituto Politécnico de Viana do Castelo, IPVC, Portugal
{diao, romero, pfaria}@ipvc.pt
[2] TMG - Têxtil Manuel Gonçalves SA, Portugal
Helena.Silva@tmg.pt

Abstract. As technologies emerge and industries evolve, new standards and solutions can be implemented to meet the increasing needs of the evolving market. The textile design industry is starting to take advantage of deep learning models, powered by artificial intelligence, where generative capabilities have been rapidly growing. In this short paper, we discuss the usage of generative models to facilitate textile designers' inspiration processes, the fine-tuning of a stable diffusion model, and the description of designs using vision models. The insights presented in this study aim to contribute to the broader application of generative tools in the textile design industry.

Keywords: Textile Design · Artificial Intelligence · Image Generation.

1 Introduction

In the fast-changing realm of digital design, the development of textile patterns continues to be a domain steeped in tradition, but also ready for innovation, as applications are already being researched and applied in the textile industry [1]. Designers often navigate through a complicated mix of aesthetics, technical limitations, and production needs, responsibilities that require creativity, tools, time, and accuracy. With the surge of artificial intelligence and deep learning models with strong generative capabilities, different applications are possible with the objective of facilitating the creative process of textile designers. The use of deep learning generative models, growing in presence, powered by artificial intelligence, such as diffusion models, is showing evolving capabilities [4] that could present relevant benefits for the textile design industry.

In the context of textile design, from different design approaches, this short paper focuses solely on woven designs. Feedback was obtained from three teams of textile designers involved in the project, which we observed through the following questions and Table **??**, directed at generative functionality. Q1 - Uses generative tools? Q2 - What tool mainly? Q3 - Has interest in having a tool that outputs based on the organization past work? Q4 - Interest in output based on the designer's preferred style? Q5 - The same prompt should output results with perceptible variation? Q6 - Tiling and rapport are essential?

This short paper presents some generative developments of a textile pattern design platform (in development), with the aim of facilitating and reducing the time invested in the creative process of designers, taking generative outputs designed to facilitate the inspiration aspect. This paper also reflects on the limitations faced through the development.

Table 1: Answers of 3 textile design teams.

Teams	1	2	TMG
Q1	Yes	Yes	No
Q2	midjourney	midjourney	-
Q3	Yes	Yes	Yes
Q4	Yes	Yes	Yes
Q5	Yes	Yes	Yes
Q6	Yes	Yes	Yes

2 Developments and Results

When starting the research on image generative models, we aimed for powerful tools that could be handled with our limited resources, and so, we were able to land on Stable Diffusion 1.5, which is a solid model with limitations such as only English, 75/77 tokens, and 512x512 pixels images. As training from scratch was not an option, we selected the weights based on the paper [2] as a base and fine-tuned the model with the data we acquired.

The data acquired was mainly scrapped from Google Images, existing datasets such as [5], and "real" data from one of the designer teams. When managing the data, we faced limitations of lack of image descriptions, the majority having vague or none at all. In the textile context, there are many specific terms (e.g. textile styles, textile applications, textile history, and composition) that need to be embedded if we want the output to be realistic to the specifications. To resolve this, we researched vision models that could handle this level of specification to a certain extent. The OpenAI Image Caption Generator seems promising, but as we lacked the paid licence to fully test it, we settled with Minicpm-v2.6 [3], which mainly described the designs in a satisfactory manner, as represented in Fig. 1, for the fine-tuning testing.

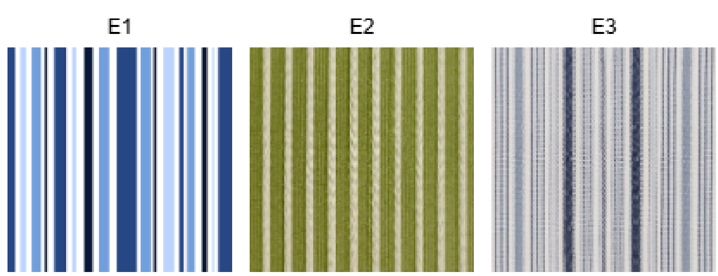


Fig. 1: 3 examples of designs

Following the examples in Fig. 1, the resumed generated descriptions obtained are: E1 – Pattern with vertical stripes in varying shades, including navy blue, white, light blue and black. Each stripe is meticulously aligned to form a harmonious pattern across the fabric's surface. This type of striped motif can be found in applications such as home decor, upholstery, or even fashion textiles. E2 – Pattern with vertical stripes in shades of olive green and white or cream. Each stripe has been meticulously aligned to form a harmonious pattern across the fabric's surface. This type of striped motif can be found in applications such as home decor, upholstery, or even fashion textiles. E3 – Pattern with vertical stripes that create a sense of depth and dimension featuring alternating shades of navy blue, white and light gray. Each stripe is outlined by thin stitching details, This type of pattern can be found in applications such as home decor, upholstery, or even fashion textiles.

For the case, represented in Fig. 2 and Fig. 3, we aimed to fine-tune the model to better fit realistic woven designs, training on simple vertical and horizontal striped compositions while trying to maintain the tilling property of the designs.

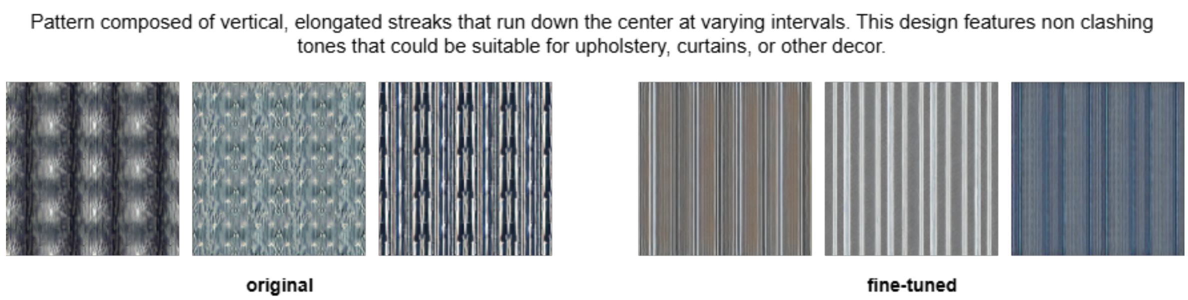


Fig. 2: 3 examples with the same prompt pre and after fine-tune

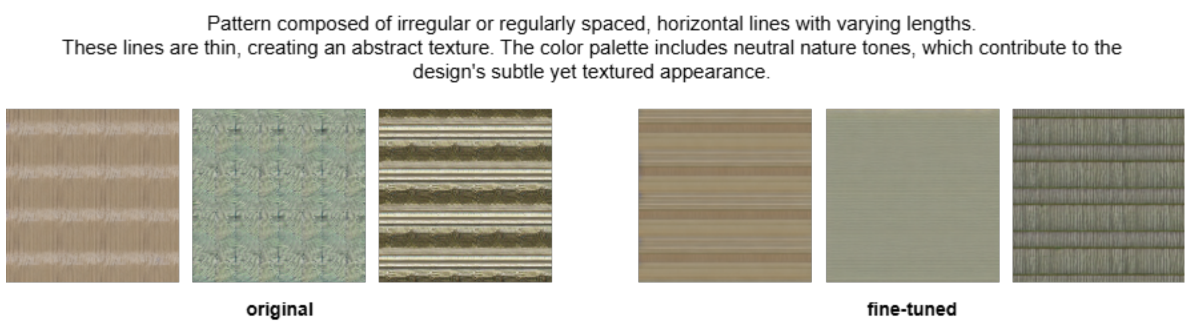


Fig. 3: 3 examples with the same prompt pre and after fine-tune

3 Conclusion

In terms of generative precision for the woven textile application, the results seem promising, with noticeable quality improvement from a simple fine-tuning test. Although using a more restrictive model such as Stable Diffusion 1.5, the output quality is not as bad as expected. Compared with other tools, such as Fermat and Midjourney, there is still a long way to go to achieve desired results. In the near future, we are expecting to test the capabilities of the model by the three professional designer teams involved. We are also striving for a more proper and efficient solution to describe designs to ensure an understanding of textile terms, and possibly train a vision model to automatically describe the data produced by the designers.

Acknowledgement

This work received financial support from the integrated project Pacto Mobilizador TEXP@CT – Pacto de Inovação para a Digitalização dos Têxteis e do Vestuário (TC-C12-i01, Bioeconomia Sustentável n.º 02/C12-i01/202), promoted by the Recovery and Resilience Plan (PRR), Next Generation EU, for the period 2021-2026.

References

1. Ingle, N., Jasper, W.J.: A review of the evolution and concepts of deep learning and ai in the textile industry. Textile Research Journal p. 00405175241310632 (2025)
2. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 10684–10695 (June 2022)
3. Wang, S., Lin, H., Luo, Z., Ye, Z., Chen, G., Ma, J.: Mfc-bench: Benchmarking multimodal fact-checking with large vision-language models. arXiv preprint arXiv:2406.11288 (2024)
4. Xie, E., Chen, J., Zhao, Y., Yu, J., Zhu, L., Wu, C., Lin, Y., Zhang, Z., Li, M., Chen, J., et al.: Sana 1.5: Efficient scaling of training-time and inference-time compute in linear diffusion transformer. arXiv preprint arXiv:2501.18427 (2025)
5. Yar, G.N.A.H., Taha, M., Afzal, Z., Zafar, F., Shahid, I.U.R., Noor-Ul-Hassan, A.: Texgan: Textile pattern generation using deep convolutional generative adversarial network (dcgan). In: 2023 IEEE International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEST). pp. 1–6 (2023). https://doi.org/10.1109/ICEST56843.2023.10138848

Development of a 2-DoF Antenna Radiation Measurement System

Mauro da Silva Leme[1], J. A. B. Coelho[2], and João Paulo Coelho[2]

[1] Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 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 (IPB), Bragança, Portugal
a52965@alunos.ipb.pt,{joaocoelho,jpcoelho}@ipb.pt

Abstract. Understanding antenna radiation patterns is crucial for optimizing wireless communication systems. Although simulation tools predict radiation patterns during development, environmental conditions in real-world settings substantially influence outcomes. Commercial measurement systems, though precise, are often prohibitively expensive for education. This work presents a low-cost, two-axis computerised system for measuring antenna radiation patterns at 868 MHz, integrating Arduino-based control, stepper motors, and power sensors. Tested with a Yagi-Uda antenna, it yields results consistent with simulations, serving as a practical tool for telecommunications classes. By enabling real-time visualization, this mechatronic platform enhances student understanding of antenna behaviour. Designs will be shared openly to support replication.

Keywords: Antenna Pattern Analysis, Wireless Systems, Mechatronic Education, Telecommunications Training

1 Introduction and System Overview

Antenna radiation patterns are essential in wireless communication engineering, defining signal coverage, interference, and system performance [1]. Simulation software predicts these patterns, yet real-world conditions, such as environmental interference and manufacturing variations, often yield different results [2]. Accurate measurements are thus critical for validating designs. Commercial measurement systems, while precise, are prohibitively expensive and complex for educational settings, limiting students to theoretical studies and simulations [3]. This restricts hands-on learning, making electromagnetic concepts harder to grasp and reducing practical problem-solving opportunities.

To address this challenge, as shown in Fig. 1, this work introduces a cost-effective, automated system for measuring antenna radiation patterns. The mechatronic platform uses Arduino-driven NEMA 17 stepper motors(1.8° resolution) and an RF power detector(based on the AD8313 integrated circuit) for dual-axis antenna rotation, with MATLAB[1] visualisation. Designed in SolidWorks[2], for telecommunications education, it enables students to analyse real-world patterns, fostering practical skills. This approach bridges theoretical knowledge with experimental practice, enhancing classroom learning.

[1] A campus version of this software is licensed to the Polytechnic Institute of Bragança.
[2] Educational version licensed to the Polytechnic Institute of Bragança.

(a) CAD rendering (b) Physical prototype

Fig. 1: System for antenna pattern measurement: SolidWorks (left) and assembled (right).

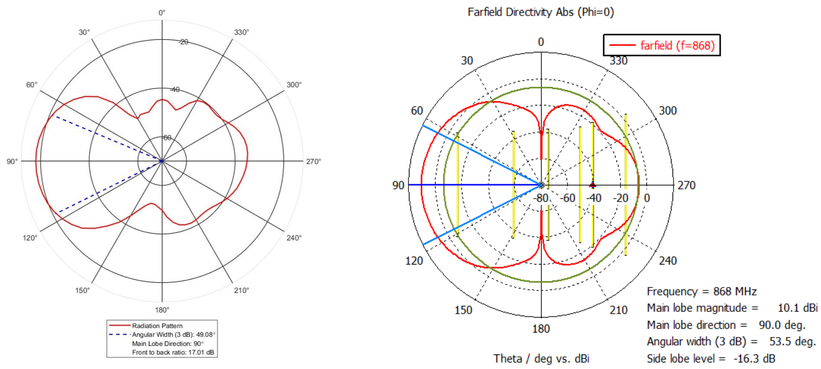
2 Results and Conclusion

The system was evaluated using a Yagi-Uda antenna at 868 MHz, with its measurements compared to CST Studio Suite[3] simulations. Fig. 2 illustrates the system's ability to capture radiation patterns, closely aligning with simulated outcomes. The primary lobe was detected at 90°, with a 3 dB beamwidth of 49.08°, near the simulated 53.5°, and front-to-back ratio reduced by 17.01 dB, compared to 16.3 dB theoretically. The noise limits of the RF detector were considered.

Small variations, due to real-world conditions like reflections and stray signals, appeared between measured and simulated patterns, underscoring difficulties in non-ideal test settings. Though external noise impacted precision, the system consistently captured radiation patterns, proving its value for classroom learning, even with its limitation. This reliability supports its role as an effective tool for teaching antenna concepts in educational environments.

Planned enhancements include filters to block unwanted signals, a balun to stabilize transmission, and optimized impedance alignment. Digital tutorials will be created to spark student interest in antenna principles through interactive visualizations. Sharing system designs openly will encourage collaborative improvements, strengthening its utility as an affordable platform for hands-on telecommunications education and fostering innovation.

[3] The free educational version of this software was used in the simulations.



(a) Measured data (b) Simulated outcome

Fig. 2: Measured and simulated radiation patterns.

Acknowledgment

This work was supported by national funds: UID/05757 - Research Centre in Digitalization and Intelligent Robotics (CeDRI); and SusTEC, LA/P/0007/2020 (DOI: 10.54499/LA/P/0007/2020).

References

1. Constantine A Balanis. *Antenna theory: analysis and design*. John wiley & sons, 2016.
2. Vittorio Picco and Keith Martin. An automated antenna measurement system utilizing wi-fi hardware. *IEEE Antennas and Propagation Magazine*, 53(6):173–184, 2011.
3. Asep Barnas Simanjuntak, Budi Mulyanti, Enjang Akhmad Juanda, Ade Gaffar Abdullah, Muhammad Naufal, and Raksa Reza Pratama. Preliminary design of automatic antenna radiation pattern measurement system for antenna and propagation laboratory course. In *Proceedings of the 5th UPI International Conference on Technical and Vocational Education and Training (ICTVET 2018)*, pages 84–87. Atlantis Press, 2019.

Flex ICT – Flexible and intelligent robotic platform for laboratory ICT

Márcio Gama[1], Emanuel Lomba[1,2], and Sérgio Lopes[1,2]

[1] ADiT-Lab, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
[2] CiTin - Centro de Interface Tecnológico Industrial, Inovarcos, 4970-786 Arcos de Valdevez, Portugal
marciogama@ipvc.pt

Abstract. The automotive industry demands high-quality and reliable electronic products, requiring thorough electrical testing. To address this, a flexible, automated In-Circuit Testing (ICT) station using a collaborative robot is proposed for testing High Voltage Coolant Heater (HVCH) PCBAs (Printed Circuit Board Assembly). This system will use Flying Probe Testing (FPT), eliminating the need for custom fixtures, reducing costs, and enhancing employee skills. Integrated machine vision will identify boards and guide the robot accordingly. The robot will also handle microcontroller programming via SWD and LIN protocols. This safe, automated approach improves flexibility, efficiency, and safety, filling a testing gap at BorgWarner Turbo and Thermal Technology.

Keywords: In-Circuit Test (ICT), Flying Probe Test (FPT), HVCH, industrial automation, collaborative robotics, computer vision, electronic measurement, Printed Circuit Board Assembly (PCBA).

1 Introduction

Due to the growing complexity and reliability demands of automotive electronics, there's an increasing need for advanced testing solutions during PCBA development and validation. Traditional manual methods and dedicated fixtures lack flexibility, limiting efficiency and raising costs. This project proposes a flexible, intelligent in-circuit test (ICT) station for BorgWarner Turbo and Thermal Technology. The solution integrates a collaborative robot to automate measurements and microcontroller programming (via SWD and LIN), along with a computer vision system to identify different PCBAs and adapt test procedures automatically – enhancing efficiency, accuracy, and test coverage.

Recent advancements in automated testing have demonstrated the advantages of integrating robotics and machine vision in electronic diagnostics. According to Muhammad Umar Anjum et. ALL the "pick and place" task is one of the most critical in industrial automation, especially in assembly lines and manufacturing processes where precision and efficiency are important. Also, the same authors state that collaborative robotics allows safe interaction between humans and robotics, and that it is a promising solution that allows increasing productivity without compromising the safety of people and the facilities in question [1]. Flying Probe Testing (FPT), for instance, has emerged as a viable alternative to fixture-based ICT by enabling high test coverage without requiring physical interfaces tailored to each board layout [2].

Studies also indicate that the use of collaborative robots in testing environments enhances operational safety and flexibility, especially when combined with intelligent vision systems and adaptive programming techniques. Building on these concepts, this project aims to design a cost-effective and scalable solution that bridges current gaps in laboratory testing, while promoting innovation and technical excellence within the automotive electronics sector [3].

2 Problem Identified

This project was initiated to address BorgWarner Turbo and Thermal Technology's need for an automated system for testing and analyzing Printed Circuit Board Assemblies (PCBA), designed to be easily adaptable for laboratory use.

Currently, testing is constrained by the design of a dedicated fixture for each PCBA, which limits the number of test points. Additionally, the measurement process is entirely manual, relying on an operator. This setup hampers quick responses in urgent situations, as it depends on the fixture's design to position the PCBA and requires external suppliers to develop components such as a bed of nails, leading to an average turnaround time of around three months. This approach suffers from multiple drawbacks, including low efficiency, reduced productivity, increased susceptibility to human error, lack of flexibility, and high costs.

3 Methodology

The development of the proposed In-Circuit Testing (ICT) station is structured in several key phases to ensure technical robustness and adaptability. The initial phase involves a comprehensive identification of system requirements and technical specifications, which serves as the foundation for all subsequent stages. A thorough analysis of the communication interfaces required for integrating the various hardware and software components is also conducted, ensuring system interoperability.

Subsequently, test points and the corresponding electrical quantities to be measured on the printed circuit board assemblies (PCBAs) are defined, followed by manual bench testing to validate the feasibility of the measurement procedures. To ensure consistency and precision during testing, a dedicated fixture is designed and implemented for the correct positioning of the PCBAs.

It is intended to develop a personalized measurement system to be operated by a collaborative robot, enabling automated, precise, and repeatable data acquisition. The integration of the collaborative robot into the testing workflow will allow for the automation of tasks traditionally performed manually, significantly increasing efficiency. Additionally, a vision system will be incorporated to enable automatic identification and differentiation of PCBAs, allowing the ICT station to dynamically adapt to various board configurations without the need for manual intervention or hardware modifications. Identification by the vision system will be carried out either by reading the board's identification code (Matrix code/QR code) or by using a standard image by components colors of the boards [4, 5] since this project is initially intended for laboratory testing rather than mass production.

The proposed conceptual architechture is presented in Fig. 1, and its core blocks can be described as follows:

- **Collaborative Robot** – Equipped with a three-function Gripper, it is responsible for measuring electrical parameters during tests.
- **Modular Fixture** – An adaptable fixturing system where PCBAs are positioned for testing.
- **Interface Machine** – Can be used to input test parameters, select functions, and view results.
- **Vision System** – Responsible for distinguishing different PCBAs by reading a code.
- **Electrical Measurement System** – Performs electrical measurements and sends the results to the central PC.
- **Programming Interface** – Standard communication interfaces for programming or diagnosing microcontrollers.
- **Central PC** – Acts as the system's brain; coordinates the robot, collects data from devices, and performs decision-making logic. Communicates with all other systems to synchronize operations.

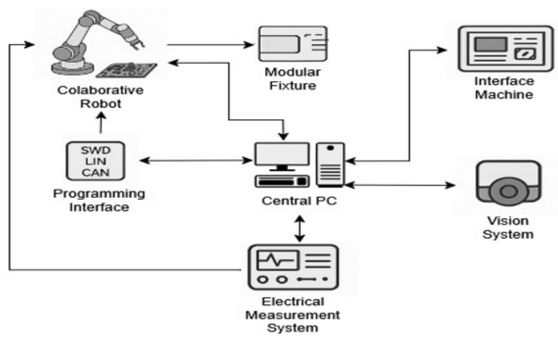


Fig. 1: Proposed architecture to development.

4 Ongoing implementation

In order to carry out the proof of concept, a Top 3 ranking was established based on historical laboratory data, identifying the most common defects found during the analysis of issues in Mercedes PCBAs (Table 1), which currently serve as the main focus for project development. Following this selection, a series of bench tests were conducted,

Table 1: Top 3 PCBA defects

Ranking	**Defect**
1	IGBT
2	Driver IGBT
3	Microcontroller Program

particularly resistance measurements between terminals of IGBT transistors to assess their operational status. The mechanical component of the project is currently under development (Fig. 2), including the design of a flexible base to allow adaptation to different PCBA formats, as well as a gripper equipped with three modules to perform electrical measurements and microcontroller programming.

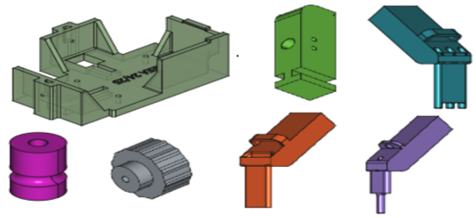


Fig. 2: Some drawings in development.

5 Conclusions

The development of a flexible and automated In-Circuit Testing (ICT) station represents a significant advancement in the testing and validation of PCBAs within the automotive industry. By integrating a collaborative robot, a vision system, and microcontroller programming capabilities, the proposed solution addresses the limitations of traditional testing methods, such as the dependency on manual measurements and the need for customized fixtures. The system improves test coverage, precision, and adaptabilitywhile simultaneously promoting operator safety and process efficiency.

In addition, the modular and scalable architecture of the ICT station enables its application to a wide range of PCBA designs, providing a cost-effective and technically robust solution for laboratory environments. This project not only meets the immediate needs of BorgWarner Turbo and Thermal Technology but also sets a precedent for future innovation in automated electronic testing within industrial contexts.

Acknowledgment

This work has been done in the context of Agenda Drivolution - Transition to the factory of the future, Ref. 02/C05-i01.02/2022.PC644913740-00000022, funded by the Portuguese Resilience and Recovery Plan (PRR) and has been partially funded by Missão Interface, an operation that offers public base funding for Technology and Innovation Centers (CTI) with project code 03/C05-i02/2022 as part of the Portuguese Plano de Recuperacão e Resiliência(PRR). This work has been partially supported through the Portuguese funding agency, FCT - Fundacão para a Ciência e a Tecnologia within project: UID/06121/2023.

References

1. Anjum, M.U., Khan, U.S., Qureshi, W.S., Hamza, A., Khan. W.A.: Vision-Based Hybrid Detection For Pick And Place Application (2023)
2. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern classification and scene analysis, vol.3. Wiley New York (1973)
3. Puttero, S., Verna, E., Genta, G., Galetto, M.: Collaborative robots for quality control. Journal of Intelligence Manufacturing (2025)
4. Sufyan, M.: Flying Probe Test: An Extensive Guide to the Technology and Applications (2023)

5. Vitalli, R.A.P., Clemente, A.O., Rosário, J.M.: Robot inspection and visual location using vision guided robot (VGR). Journal of Engineering Research **2**(5) (2022). Campinas State University (UNICAMP), Brazil

Economic Freedom in the European Union: insights from the 2021 Index

Ana Maria Silva[1], Ana Figueiredo[1], Ana Barbosa[1], Elizabeth Fernandes[1,4], and Helena Sofia Rodrigues[2,3]

[1] Escola Superior de Ciências Empresariais, Instituto Politécnico de Viana do Castelo, IPVC, Portugal
{anamariasilva,figueiredoana,ana.b}@ipvc.pt, elisabeth.ferna@gmail.com
[2] ADiT-Lab, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
[3] CIDMA - University of Aveiro, Portugal
sofiarodrigues@esce.ipvc.pt
[4] Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal

Abstract. This paper analyzes economic freedom in the European Union based on the 2021 Index of Economic Freedom. Despite high institutional integration, significant disparities exist among member states, particularly between Northern and Southern, and Western and Eastern Europe. Using descriptive statistics and correlation analysis, the study identifies key institutional and fiscal factors linked to higher economic freedom, such as strong rule of law, efficient regulation, and moderate public spending. The findings highlight persistent asymmetries and suggest the need for targeted reforms to foster convergence and resilience in the post-pandemic EU economic landscape.

Keywords: Economic Freedom · European Union · Data Analysis.

1 Introduction

Economic freedom has become a central concept in evaluating the institutional quality and policy effectiveness of modern economies. It reflects the extent to which individuals and businesses can operate with minimal state interference, and it is widely associated with economic growth, innovation, and social well-being [1]. Tools like the Index of Economic Freedom, published by the Heritage Foundation, provide a comparative framework to assess key dimensions such as rule of law, fiscal policy, regulatory efficiency, and market openness.

In the European Union (EU), composed of 27 member states with diverse economic structures and institutional capacities, economic freedom varies significantly [2]. Despite shared policy frameworks and a high degree of integration, pronounced disparities persist – particularly between Northern and Southern, and Western and Eastern Europe. These differences raise important questions about convergence, cohesion, and the overall functioning of the EU's economic model.

This study addresses the following research question: How is economic freedom distributed among EU member states in 2021, and what patterns or asymmetries can be identified through statistical analysis? By analyzing the 2021 index data, the paper identifies regional trends and explores potential explanatory factors, offering insights for both academic debate and policymaking in a post-pandemic European context.

2 Methodology

This study adopts a quantitative and descriptive approach based on secondary data from the 2021 Index of Economic Freedom, published by the Heritage Foundation [3]. The database was constructed using 16 indicators, grouped into three dimensions: legal (property rights, judicial effectiveness, government integrity), fiscal (tax burden, government spending, fiscal health, tax rates, public debt), and economic (business, labor, monetary, financial, and investment freedoms). The techniques applied include descriptive statistics, Pearson correlation analysis, cluster analysis, and principal component analysis. The analyses were conducted using SPSS, focusing on measures of central tendency, dispersion to identify patterns and institutional asymmetries among EU member states. Additionally, data visualization techniques were applied using Looker Studio to create interactive maps and dashboards that illustrate regional variations and facilitate interpretation of complex relationships across the dataset.

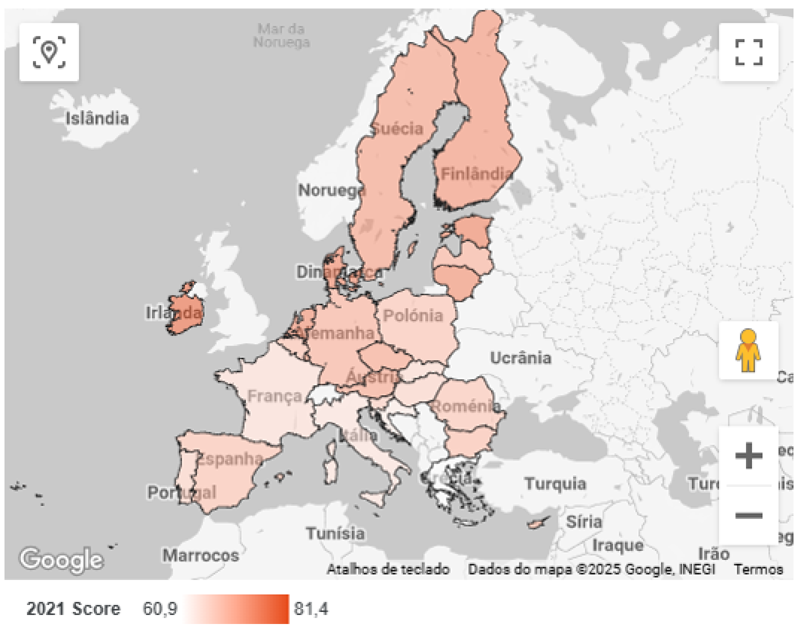


Fig. 1: Economic Freedom Index in the European Union by Country (2021)

.

3 Main results

According to Fig. 1, the analysis reveals disparities between member states. Scores range from 60.9 to 81.4, indicating a wide variation in performance throughout the continent. Northern European countries such as Sweden, Finland, Denmark, and Ireland rank highest, likely reflecting stronger institutional efficiency, higher development levels, or favorable policy environments.

In contrast, Southern and Eastern European countries like Portugal, Spain, Romania, and Poland display lower scores, suggesting structural challenges or weaker performance in key areas. These disparities highlight ongoing economic inequalities within the EU, underscoring the need for more targeted cohesion and development policies.

Table 1: Descriptive Statistics – Economic Freedom Indicators in the EU (2021)

Category	Indicator	Mean	Std. Error	Min	Max
Legal	Property Rights	76.99	8.93	55.5	91.9
	Judicial Effectiveness	64.47	12.83	41.1	86.9
	Government Integrity	70.67	16.25	46.8	97.2
Fiscal	Tax Burden	67.46	15.04	43.6	94.3
	Government Spending	43.52	17.58	6.3	81.1
	Fiscal Health	89.11	9.57	69.0	99.6
	Income Tax Rate (%)	36.97	14.03	10.00	57.00
	Corporate Tax Rate (%)	20.12	5.86	9.0	35.0
	Tax Burden / GDP (%)	34.54	8.37	14.59	46.09
	Gov. Expenditure / GDP (%)	42.84	7.02	25.08	55.90
	Public Debt / GDP (%)	63.55	38.76	8.4	179.2
Economic	Business Freedom	72.47	9.26	55.6	88.8
	Labor Freedom	59.84	10.29	43.6	77.1
	Monetary Freedom	80.76	2.95	74.8	85.2
	Trade Freedom	84.00	0.00	84.0	84.0
	Investment Freedom	79.63	9.19	55.0	95.0
	Financial Freedom	67.04	9.93	50.0	80.0
	Tariff Rate	3.00	0.00	3.0	3.0

Analysis of the Index indicators reveals notable disparities in the legal, fiscal, and economic dimensions (Table 1). Property Rights are consistently strong, while Judicial Effectiveness and Government Integrity vary more widely, pointing to institutional gaps. Fiscal indicators show high divergence – especially in Public Debt and Government Spending – despite generally strong Fiscal Health. Differences in tax burden and public expenditure reflect diverse fiscal strategies. Economically, the EU shows full alignment in Trade Freedom and Tariff Rates, but other areas like Labor, Investment, and Financial Freedom display significant variation, suggesting uneven regulatory environments.

References

1. De Haan, J., Lundström, S., Sturm, J.-E. (2022). Market-Oriented Institutions and Policies and Economic Growth: A Critical Survey. Journal of Economic Surveys, 36(2), 352-380.
2. Hall, J., Lawson, R. (2023). Economic Freedom and Well-Being: Evidence from the 21st Century. World Development, 160, 106075
3. Heritage Foundation. (2021). 2021 Index of Economic Freedom. The Heritage Foundation. Retrieved from https://www.heritage.org/index/

Application of Artificial Intelligence in Coffee Cultivation: Automated Diagnosis of Diseases and Nutritional Deficiencies Using Computer Vision

Gabriel Felippe Baptista[1], João Mendes[2], Eduardo Filgueiras Damasceno[3], José Lima[2], and Ana I. Pereira[2]

[1] Instituto Politécnico de Bragança, 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 (IPB), Bragança, Portugal
a65453@alunos.ipb.pt, {joao.cmendes,jllima,apereira}@ipb.pt
[3] Federal University of Technology – Paraná (UTFPR), Brazil
damasceno@utfpr.edu.br

Abstract. The application of artificial intelligence in agriculture has led to significant advances in coffee production. This work aims to develop a computer vision system based on convolutional neural networks to identify diseases and nutritional deficiencies in coffee plants. The proposed solution is designed to operate in real time, using a dataset of leaf images, and to be deployable on mobile devices or autonomous robotic platforms. By enabling early detection and intervention in the field, the system is expected to promote more efficient, scalable, and accessible agricultural management.

Keywords: Artificial Intelligence · Coffee Cultivation · Machine Learning · Computer Vision · Precision Agriculture.

1 Introduction

Agriculture has been a cornerstone of human civilization since the Neolithic era, around 10,000 years ago [5]. Over time, it has undergone significant technical, social, and environmental changes that shaped global production systems. [1]

Coffee cultivation, especially in Brazil, the world's leading producer and exporter since the 19th century, stands out for its economic and sociocultural importance [3, 4].

Currently, agriculture is central to global challenges such as food security and sustainability, with emerging technologies like artificial intelligence (AI), machine learning (ML), and computer vision driving innovation [2, 7]. For example, Maruthai et al. [6] achieved 93.66% accuracy detecting coffee berry borer pests using Hybrid Vision Graph Neural Networks, while Senanayake and Maduranga [8] applied convolutional neural networks to identify early-stage diseases with 89% accuracy via a mobile app.

This study explores the use of AI in precision coffee farming, emphasizing early detection of pests and diseases and yield estimation to promote more efficient and sustainable crop management.

The paper is organized as follows. After the paper introduction, Section 2 presents the study methodology, and the last section describes the discussion.

2 Methodology

This study proposes the development of an artificial intelligence system to identify diseases and nutritional deficiencies in coffee plants, using leaf images as the primary data source. The images were collected by the authors themselves from farms located in the state of São Paulo, Brazil, covering different seasons and environmental conditions.

The dataset currently comprises around 800 manually labeled coffee leaf images with a resolution of 640×640 pixels, categorized into healthy and unhealthy classes. In future stages, a more detailed dataset will be developed to include local images labeled with specific diseases and nutritional deficiencies.

Fig. 1 illustrates the overall methodological flow of the proposed system.

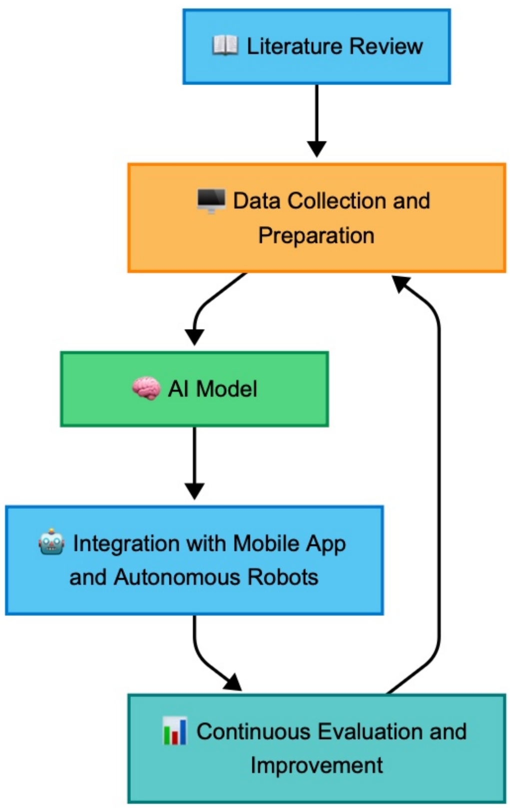


Fig. 1: Methodological flow of the AI-based system for coffee plants.

The system pipeline is currently under active implementation, involving dataset assembly, model training, and initial prototype development. However, the work remains conceptual from an academic perspective, with validations and field tests planned for future stages.

Initially, the system will be trained to classify whether a plant is healthy or not. Subsequently, additional databases containing labeled images of specific diseases and nutritional deficiencies will be incorporated, although this phase is still conceptual.

The chosen AI model is YOLOv11, which balances performance and lightweight design, making it suitable for field use. Training will be conducted using the PyTorch and Ultralytics YOLO libraries, and the final application is expected to integrate a mobile app for real-time diagnosis, as well as an autonomous robot for continuous crop monitoring, even in areas with limited digital infrastructure.

3 Discussion

Technology is rapidly transforming crop management, and the use of AI in coffee production is becoming a valuable tool for producers. Several studies have explored the use of leaf images for disease detection in coffee plants. However, the methodology proposed in this work differs by integrating mobile applications and autonomous soil robots, aiming to improve monitoring capabilities, especially in areas with limited infrastructure.

While previous research has made significant progress using image analysis for disease diagnosis, our approach aims to enhance those results by expanding the dataset and refining the models. By incorporating more diverse data and improving model performance, we anticipate increasing the reliability and accuracy of disease detection, thus contributing to more sustainable and productive agricultural practices.

Acknowledgment

This work was supported by national funds: UID/05757 - Research Centre in Digitalization and Intelligent Robotics (CeDRI); and SusTEC, LA/P/0007/2020 (DOI: 10.54499/LA/P/0007/2020).

References

1. Castanho, R.B., Teixeira, M.E.S.: A evolução da agricultura no mundo: da gênese até os dias atuais. Brazilian Geographical Journal: Geosciences and Humanities research medium **8**(1), 136–146 (2017)
2. Elgendy, A.: Computer vision in agriculture. `https://www.ultralytics.com` (2024), accessed on May 14, 2025
3. EMBRAPA: Coffee production system. `https://www.embrapa.br` (2018), accessed on May 14, 2025
4. Government of Brazil: Brazilian coffee: production and export. `https://www.gov.br` (2023), accessed on May 14, 2025
5. Guitarrara, M.: History of agriculture. Revista AgroHistória (2025)
6. Maruthai, R., Kumar, S., Gupta, A.: Detection of coffee pests using hybrid vision graph neural networks. Computers and Electronics in Agriculture **204**, 107439 (2025). https://doi.org/10.1016/j.compag.2025.107439, accessed on May 10, 2025
7. Santos, B.: Computer vision in precision agriculture. `https://blog.macnicadhw.com.br` (2025), accessed on May 14, 2025
8. Senanayake, L., Maduranga, K.: Early detection of coffee plant diseases using mobile-based cnn models. In: Proceedings of the 18th International Conference on Artificial Intelligence and Soft Computing. Lecture Notes in Computer Science, vol. 13984, pp. 451–460. Springer (2023). https://doi.org/10.1007/978-3-031-31392-2_42, accessed on May 10, 2025

VOC Pattern Recognition in Waste Using AI-Enabled BME688 Sensors

Felipe Merenda Izidorio[1], Mariam Jvarsheishvili[1], Getúlio Igrejas[1], Pedro João Rodrigues[1], and Rui Pedro Lopes[1]

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança (IPB), Bragança, Portugal
{felipeizidorio, mariam, igrejas, pjsr, rlopes}@ipb.pt

Abstract. Urban waste management faces significant challenges that demand innovative technological solutions. This paper presents an ongoing development of a comprehensive smart waste management system that combines sensor technology, artificial intelligence, and optimization algorithms. The paper focus on the initial phase of this project: evaluating the suitability of the BME688 environmental sensor for monitoring volatile organic compounds (VOCs) in organic waste containers. Using coffee grounds and ash as test materials, with normal air readings between samples, we collect and analyze sensor data to determine the device's ability to differentiate waste types through chemical signatures. Early findings indicate that the BME688 can detect distinctive VOC patterns from different waste types, supporting the development of a more sustainable and efficient urban waste management infrastructure.

Keywords: Smart waste management · VOC monitoring · Environmental sensor.

1 Introduction

Urban waste management constitutes a critical infrastructure challenge for modern cities. Traditional waste collection methodologies operate on fixed schedules regardless of actual container fill levels, resulting in inefficient resource allocation, unnecessary carbon emissions, and occasional overflow incidents. Additionally, the inability to detect hazardous materials in waste streams creates environmental and safety risks.

A transformative approach is being developed through an integrated system of smart sensors, artificial intelligence, and optimization algorithms. This system aims to monitor waste conditions in real-time, detect potentially dangerous materials, optimize collection routes dynamically, and provide actionable insights to waste management stakeholders. The environmental and economic benefits include reduced collection frequency, lower fuel consumption, decreased carbon emissions, and enhanced worker safety.

This paper focuses on evaluating the BME688 environmental sensor for detecting chemical signatures from organic waste materials, specifically examining its ability to differentiate between coffee grounds and ash residues, with ash detection being particularly important for safety, as improperly disposed hot ash can cause fires in waste containers [2].

2 Technology Overview and Methodology

The Bosch BME688 environmental sensor is used to detect volatile organic compounds and monitor environmental conditions in waste samples [1]. The sensor was chosen because of its ability to simultaneously measure gas concentration, temperature, humidity, and pressure in a compact and energy-efficient format.

The experimental setup connects the BME688 development kit to a microcontroller for continuous data acquisition. Measurements were carried out on controlled samples of coffee grounds and ash, interspersed with readings in ambient air. This approach establishes a reliable baseline and minimizes the risk of cross-contamination, making it possible to clearly distinguish between environmental variations and the specific signatures of residues (see Table 1 for a sample of the acquired sensor data).

Table 1: Sample of raw data from BME688 during coffee grounds testing

Data ID	**Gas Resistance**	**Temp. (°C)**	**Pressure (hPa)**	**Humidity (%)**	**Time On (s)**	**RTC**	**Error**	**Step ID**	**Cycle ID**
1	8.25e+05	29.3	933.0	25.0	0	1688772009	0	0	1
2	1.02e+08	29.3	933.0	25.0	279	1688772009	0	1	1
3	1.02e+08	29.5	933.0	24.9	1680	1688772011	0	2	1
4	1.02e+08	29.8	933.0	24.2	5879	1688772015	0	3	1
5	1.47e+07	30.3	933.0	24.2	6579	1688772016	0	4	1

The sensor was first calibrated in clean air, then each material was exposed in sealed conditions for 30 minutes, resulting in approximately 1000 cycles per sample, with a clean air baseline wash before the first exposure and after each sample. Throughout each exposure, gas resistance, temperature, humidity, and pressure were recorded at regular intervals. This cycle of sample exposure, baseline wash, and subsequent sample ensures clear, interference-free signatures for each material.

3 Preliminary Results

Fig. 1 presents the initial exploratory findings. The correlation heatmaps show that under ash all three readings vary independently, while coffee induces a strong negative link between temperature and humidity. Normal air exhibits moderate negative correlations between resistance and temperature and between temperature and humidity. The boxplots indicate that resistance is lowest and most uniform with ash, more variable with coffee (including several extreme spikes), and highest with the widest spread in clean air. These simple visuals confirm the BME688's capacity to tell ash, coffee, and air apart, paving the way for a lightweight classifier in the proposed smart waste system.

4 Conclusion

This paper presents the initial research phase of an intelligent waste management system. The findings indicate that the BME688 environmental sensor can reliably differentiate between waste materials and normal air through its VOC signatures, supporting its use in intelligent waste monitoring applications. This feature forms the

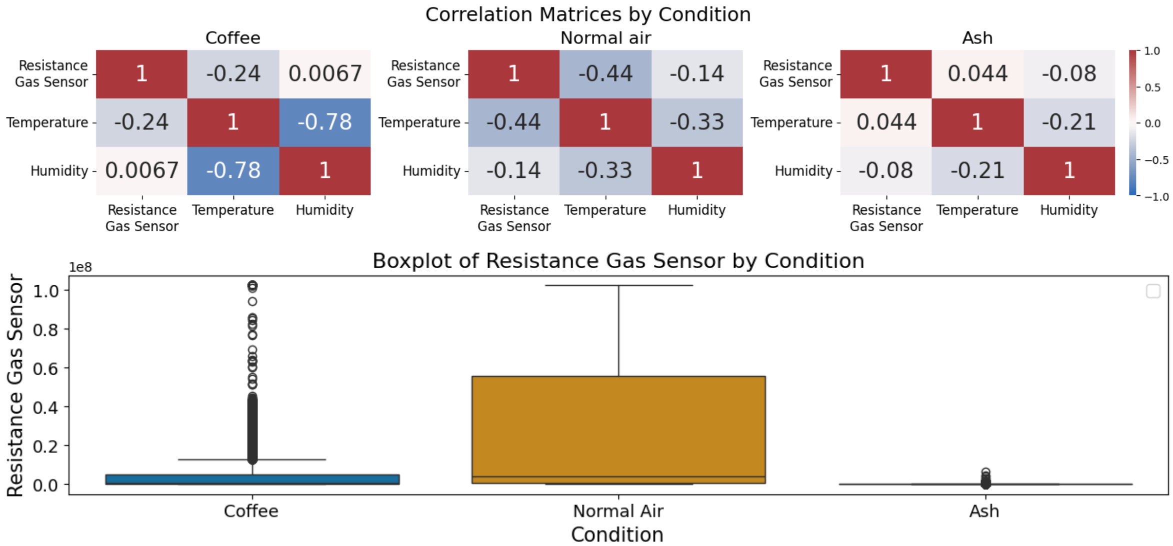


Fig. 1: (a) Correlation matrices between Resistance Gassensor, Temperature, and Relative Humidity under three environmental conditions: Ash, Coffee, and Normal Air; (b) Boxplot of Resistance Gassensor readings under the same conditions.

basis of a system that can detect hazardous materials, optimize collection routes and provide valuable insights into the data. One key limitation of the sensor is its baseline drift over time: even in a constant, gas-free environment, the sensor's resistance 'baseline' gradually changes due to aging of the material, surface contamination, and uncompensated environmental fluctuations, which can lead to significant errors in gas concentration estimates if not regularly recalibrated. Future research will focus on field-testing the sensor units in real waste containers, developing more sophisticated classification algorithms and integrating the sensor subsystem with route optimization and data analysis platforms.

Acknowledgment

This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the Project Scope: 2024.07316.IACDC/2024.

References

1. Gas sensor BME688, `https://www.bosch-sensortec.com/products/environmental-sensors/gas-sensors/bme688/`
2. Kaza, S., Yao, L., Bhada-Tata, P., Van Woerden, F.: What a waste 2.0: a global snapshot of solid waste management to 2050. World Bank Publications (2018)

Deep Dive into Data Poisoning: Proposal and Methodology

Hugo Costa[1], Silvestre Malta[1,2], and Ricardo Santos[3]

[1] Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal,
huca@ipvc.pt
[2] ADiT-Lab, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
smalta@estg.ipvc.pt
[3] CIICESI, ESTG, Politécnico do Porto, Portugal
rjs@estg.ipp.pt

Abstract. With the growing interest in artificial intelligence, now more than ever is necessary to understand the rising threats that accompany artificial intelligence. One of those threats is data poisoning. A threat that targets the training phase of a model. Machine learning algorithms relay on data to train models, and in most cases, the necessary data is collected from the Internet. Unfortunately, in many cases this data is collected from all types of unreliable sources, leaving space for possible issues regarding the models expected performance and outcomes. This paper aims to establish the methodology that is going to be applied to replicate some of the attacks under this threat in hopes of better understanding them.

Keywords: Artificial Intelligence · Data Poisoning · Machine Learning

1 Introduction

Within the subject of data poisoning multiple attacks have been established. Each one with its own goals, attack vectors, impacts, and necessary requirements. In this paper, a plan was established to define a clear exploration path and methodology for two of categories of attacks under this scope of data poisoning. This exploration is going to be directed at Availability Poisoning, as the as the starting point, following it with Target Poisoning. With Availability Poisoning the objective of the attacker is to compromise the availability of a model by targeting the data used in training [1]. Unlike the Availability Poisoning Attack, Target Poisoning is more directed to influence the model without impacting the accuracy of the model, or its functionality. In Targeted Poisoning the attacker poisons the training data to cause misclassifications on only a few target samples [1]. Within the scope of data poisoning, these represent only two of the many attacks established. And for the purpose of this paper these are the only two types of poisoning that are going to be explored.

2 Related work

When it comes to exploring the subject of data poisoning, one great starting point is the work made by NIST in a report [3] that emphasizes the creation of a taxonomy and the clarification of terms within adversarial machine learning. In this document, regarding data poisoning, four distinct categories of attacks are described, in terms of concept, available literature, and compilation of possible defenses available for each one. The described categories are availability poisoning, targeted poisoning, backdoor poisoning and model poisoning. A significant consideration is that the classification adopted in the report [3] is based on a framework established in an earlier study [1]. That specific paper [1] provides a systematization of poisoning attacks and defenses in the field of Machine Learning (ML), based on the review of an extensive list of papers published in the field. A fact that is different between both papers is the categories established under the concept of data poisoning. In [1] model poisoning [3] is not identified as a category under the scope of data poisoning.

3 Methodology

Section dedicated to explaining the methodology that is going to be used for the development of the attacks under the scope of data poisoning.

3.1 Settings and Baseline

To archive the objective of understanding the attacks, it's necessary establish the environment and define the settings in which these attacks are going to tested. For this purpose, the plan is to use Google Colab to train and run the models. The reason being the computational requirements associated with training a model. In terms of programming language, python was selected due to all the frameworks that ready exist and documentation available for this proposes. For the development/training of the models the plan is to follow a similar approach as [2] by using common datasets used for classification benchmarks. More specifically the BreastCancer and the Spambase datasets. Using these datasets clean models can be developed to serve as baselines when comparing with models compromised via poisoning.

3.2 First Scenario

For the first scenario the plan consists of exploring Data Poisoning via availability poisoning also reference as Indiscriminate Poisoning by the authors in [1], for a combination of factors. Within the scope of data poising, Availability Poisoning has one characteristic that make it possible to detect. While some attacks avoid impacting the accuracy of a model, Availability Poisoning is focused in targeting it. Meaning that it can be identified via metric analysis of the model, being the accuracy the main metric to consider, other metrics will be applied based on the type of model in question. It also proves useful as a starting point to test some of the attacks technics that can later be used for the following attacks. Technics like Label Flipping and Clear Label Attack. In this attack three attempts are going to be performed for each model by means of each technic, in the first attempt 25 percent of the dataset is going to be compromised both via label flipping and clean label attack, with the objective a simulating two different scenarios where the attack only has the ability to change the labels or the features but never both. After this stage is complete the metrics should be compared against the baseline models to look for indications of poisoning. The second and third attempt should follow the same approach, but this the dataset compromised by 35 percent and 45 percent respectably. Once all attempts are completed, mitigation's solutions are going to be applied and the attempts replicated to check the effectiveness of current strategies. These solutions consist of dataset sanitation based on the data analysis and robust training.

3.3 Second Scenario

For the second scenario the plan is to explore target poising, an attack that can be proven difficult to detect since it can occur without much change in terms of metrics. The setting for this scenario is more complex, the attempts are going to be performed on both whitebox and blackbox settings. In the first one a set of cases in the dataset are going to be flipped, then the models should be trained and compared with the original metrics of the baseline. After that step the model's prediction will be tested by presenting a similar sample as the poison one. Only when all tests in the whitebox context are concluded, do the blackbox tests begin where the experiment is the same but without having context in terms of what samples have been flipped. This will be achieved by means of a script intended to flip the labels of samples based on the wanted percentage of poisoning for the test. And the objective is to identify then via data analyses of the dataset, before the training and metrics analyses after the fact. Like mentioned in [3] this type of attack is challenging to defense against, meaning that the focus is investigating possible indicators of compromised in the datasets before and after training.

4 Conclusion

With this methodology in place the objective is to conduct both scenarios to better understand the attacks and analyze them in a more practical context. Testing current reported technics of identifying, exploring and mitigating them. While also evaluating the current solutions in terms of effectiveness. Keeping an eye out for the possibility of developing new ways to defend against this issue based on the results of the experiments and the study of both attacks in a similar context.

References

1. Cinà, A.E., Grosse, K., Demontis, A., Vascon, S., Zellinger, W., Moser, B.A., Oprea, A., Biggio, B., Pelillo, M., Roli, F.: Wild patterns reloaded: A survey of machine learning security against training data poisoning. ACM Computing Surveys **55**(13s), 1–39 (2023)
2. Paudice, A., Muñoz-González, L., Lupu, E.C.: Label sanitization against label flipping poisoning attacks (2018), `https://arxiv.org/abs/1803.00992`
3. Vassilev, A., Oprea, A., Fondyce, A., Anderson, H.: Adversarial machine learning: A taxonomy and terminology of attacks and mitigations. Nist (2023)

An Artificial Intelligence Method for Large Forest Areas Analysis

Raphael D. Britto[1,2], João Mendes[1], Murillo Ferreira dos Santos[2], João P. Castro[3], José Lima[1], and Ana I. Pereira[1]

[1] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança (IPB), Bragança, Portugal
{raphael, joao.cmendes, jllima, apereira}@ipb.pt

[2] Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), 36700-000, Minas Gerais, Brasil
murillo.ferreira@cefetmg.br

[3] CIMO, SusTEC, Instituto Politécnico de Bragança, Campus de Santa Apolónia, Bragança, 5300-253, Portugal
jpmc@ipb.pt

Abstract. The need to manage vegetation cover, which is often dispersed over extensive areas, is fundamental for environmental problems. Technological tools and resources can be applied to identify essential and relevant information. This study presents a methodology to predict carbon sequestration using UAV-based LiDAR scans, providing an innovative solution for the inspection of large forest areas. The methodology developed requires RGB images from forest areas. This paper also points out the results and conclusions already reached, as well as future possibilities.

Keywords: CNN · Forest areas · LiDAR · QGIS · RGB images.

1 Context and Motivation

Some countries are establishing governmental measures to normalize human actions and prevent an increase in environmental impacts. The regulation of toxic gas emission rates is a context of interest. In the case of carbon dioxide, its neutralization can be achieved through the existence of vegetation cover [3]. Forests play a key role in this scenario, considering typical high carbon storage indices.

The large extension of forest areas underscores the importance of intelligent and technological applications. Remote monitoring of these areas can be conducted with a reasonable degree of reliability. Thus, providing an efficient alternative in terms of the time and resources expended.

The use of uav presents a promising technique for facilitating the acquisition and analysis of forest data. This is due to their capacity to transport acquisition devices, such as multispectral cameras or Light Detection and Ranging (LiDAR) sensors [6].

Employing LiDAR scans, in which height and color values are assigned to pixels in a point cloud, facilitates the manipulation of the data for effective inspection [4]. This analysis can be integrated into artificial intelligence (AI) algorithms to achieve carbon stock amounts. The application of learning models is subject to certain requirements, including the availability of substantial data sets that are typically accompanied by labeled values.

The LAStools extension in Quantum Geographic Information System (QGIS) software offers a variety of tools for the processing of LiDAR data. Raster models facilitate the extraction of both RGB images and height averages from raw data [1]. In this regard, QGIS establishes a viable connection between LiDAR scans and learning algorithms.

This paper aims to summarize a research in progress, started in [2]. The final objective is to reach a robust AI approach to access carbon sequestration in forest areas. Achieved results and future work possibilities will be highlighted.

2 Methodology

Data will be acquired through LiDAR scans by Unmanned Aerial Vehicle (UAV) overflights. The raw scan data (point clouds) will then be processed in virtual environments, such as QGIS. Using LAStools, this data will be converted into rasters and Digital Elevation Models (DEMs). Fig. 1 shows an overview of the methodology in development.

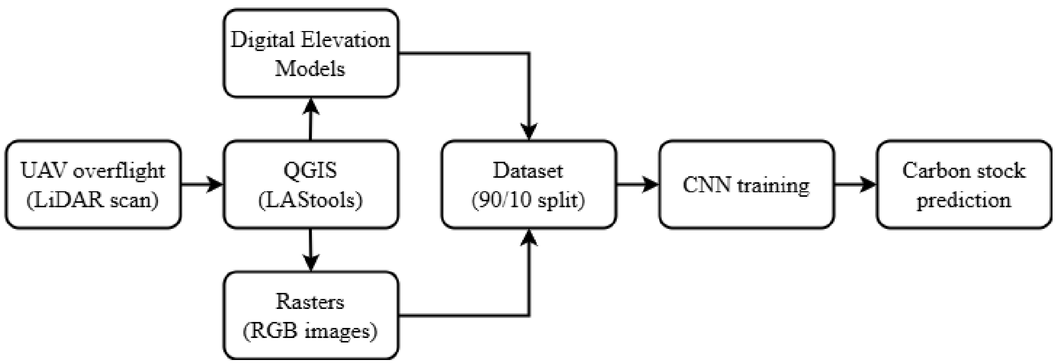


Fig. 1: Comprehensive methodology schematic.

After data processing, vegetation height values and RGB images will be organized into datasets. Implementing Deep Learning (DL) models, specifically Convolutional Neural Networks (CNN]s), relevant forest information will be estimated, such as carbon sequestration taxes. RGB images with 64×64 pixels dimension will be used as input data, applying a 90/10 split between the training and test sets. This division rate was defined considering a small dataset size.

3 Discussion

The work [2] presents the implementation of the proposed methodology, predicting heights in *Quercus pyrenaica* Willd. colonies. The stages of LiDAR data acquisition were outlined, along with its processing in QGIS software and its implementation in the VGG16 CNN] model [5]. The prediction results indicated a mean absolute error (MAE) of 0.7938 related to the QGIS ground truth values.

To explore the performance of the methodology, a more extensive study is being conducted using *Pinus pinaster* Ait. colonies. In addition to height predictions, the potential for accessing phytovolume, above ground biomass (AGB), and carbon sequestration is being examined. These variables are directly linked to the forest area coverage. To define a more reliable model, different CNN] architectures are being explored and compared using different evaluation metrics.

In future works, the performance of prediction models fitting data from diverse species will be researched. Favorable results allow for the investigation of scenarios with mixed species, encompassing a more common forest scenario. In addition, the use of other RGB images resources as dataset will also be addressed. For example, UAV imagery (conventional cameras) or free access digital orthophotos.

Acknowledgment

This study was funded by iCarbono project Fundação La Caixa (PL23-00038) and LIFE SILFORE project (LIFE21-CCA-ES-LIFE). The authors are also grateful to CeDRI (UID/05757), SusTEC (LA/P/0007/2021), CIMO (UIDP/00690/2020), CEFET-MG and the National Council for Scientific and Technological Development – CNPq, related to project 442696/2023-0.

References

1. Ballesteros, R., Ortega, J.F., Hernandez, D., Moreno, M.A.: Onion biomass monitoring using uav-based rgb imaging. Precision agriculture **19**, 840–857 (2018)
2. Britto, R.D., Mendes, J., Grilo, V., Castro, J.P., dos Santos, M.F., Castro, M., Pereira, A.I., Lima, J.: A deep learning approach for average height estimation in oak colony using rgb images. OL2A (2025)
3. Griscom, B.W., Adams, J., Ellis, P.W., Houghton, R.A., Lomax, G., Miteva, D.A., Schlesinger, W.H., Shoch, D., Siikamäki, J.V., Smith, P., et al.: Natural climate solutions. Proceedings of the National Academy of Sciences **114**(44), 11645–11650 (2017)
4. Lu, N., Zhou, J., Han, Z., Li, D., Cao, Q., Yao, X., Tian, Y., Zhu, Y., Cao, W., Cheng, T.: Improved estimation of aboveground biomass in wheat from rgb imagery and point cloud data acquired with a low-cost unmanned aerial vehicle system. Plant Methods **15**, 1–16 (2019)
5. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
6. Wallace, L., Lucieer, A., Watson, C., Turner, D.: Development of a uav-lidar system with application to forest inventory. Remote sensing **4**(6), 1519–1543 (2012)

Preliminary anatomical studies on the reliability of electromyography signal collection methods

Caio Oliveira[1], Luiz E. Luiz[2], and João P. Teixeira[2]

[1] Instituto Politécnico de Bragança, 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 (IPB), Bragança, Portugal
a52963@alunos.ipb.pt, {luiz.luiz, joaopt}@ipb.pt

Abstract. This work presents preliminary anatomical studies on the reliability of surface electromyogram (sEMG) acquisition during finger movements. An anatomical analysis was used to identify the muscles responsible for finger contraction and for compensatory wrist movement. The results showcase that this systematic approach brings reliability in research contexts with information regarding the objective movement and the most common compensatory interference movement.

Keywords: sEMG · Forearm muscles · Signal acquisition

1 Introduction

Surface electromyography (sEMG) offers key insights into neuromuscular function [2]. However, the forearm's complex anatomy challenges the acquisition of reliable signals, causing interference and inconsistent electrode placement [1, 2]. The reliability of these signals is critical when driven to areas such as medical diagnosis, muscle rehabilitation and the control of electric prostheses [3]. To overcome the limitations of standardised protocols like SENIAM [2] and ISEK [4], this study proposes and validates a systematic acquisition method grounded in a preliminary anatomical analysis to enhance data reproducibility.

2 Anatomical Study and Validation

From the 20 forearm muscles [5], this study used the accessible superficial and intermediate layers, the flexor carpi radialis (FCR), flexor digitorum superficialis (FDS), and brachioradialis (BR) [6] during palmar grip.

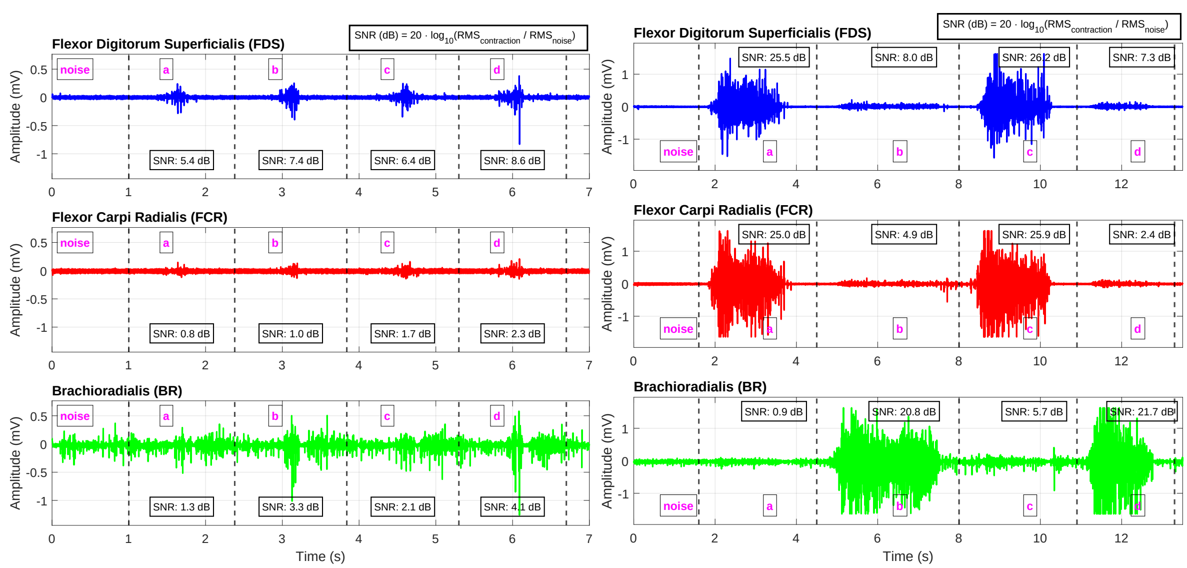
Based on anatomical considerations such as fibre orientation and proximity to bony landmarks (e.g., medial epicondyle and ulnar wrist), the FDS was identified as the primary candidate for isolating finger flexion. Despite its intermediate-layer classification, the FDS follows an oblique path that creates a concentrated and accessible region of muscle fibres [6]. The acquisition protocol required a stable seated position to mitigate known challenges like muscle-skin displacement and spatial distribution artefacts [1], thereby reducing such artefacts and suppressing crosstalk from synergistic muscles.

To quantitatively assess muscle selection, a Signal-to-Noise Ratio (SNR) contraction-relative analysis was performed. The SNR for each contraction event was calculated using equation 1,

$$SNR = 20 \cdot \log_{10} \frac{RMS_{contraction}}{RMS_{noise}} \; [dB] \tag{1}$$

where $RMS_{contraction}$ and RMS_{noise} are the root mean square values calculated from the sEMG signal within the respective contraction and baseline noise windows. The results of this quantitative analysis are presented in Fig. 1.

Fig. 1 (a) displays the sEMG signals from the FDS, FCR, and BR during finger flexion contractions. Across four contractions, the FDS yields superior SNRs, ranging from 5.4 dB to 8.6 dB. In contrast, both the FCR and BR muscles produce significantly lower values, with the FCR's SNR between 0.8 dB and 2.3 dB, and the BR's between 1.3 dB and 4.1 dB. Although the FCR exhibits low baseline noise, its amplitude during finger flexion is minimal, resulting in a poor SNR. The BR signal presents considerable baseline noise, which reduces its effectiveness, confirming that FDS is the optimal choice for this task.



(a) (b)

Fig. 1: sEMG data acquired during (a) 4 finger flexion movements labelled a to d and (b) during wrist flexion (a and c) and extension (b and d).

To validate FDS selectivity against compensatory movements, dedicated wrist flexion and extension were analysed, as shown in Fig. 1 (b). This task highlights the specific roles of the FCR and BR, which present strong activation during wrist flexion (SNR up to 25.9 dB) and extension (SNR up to 21.7 dB), respectively. Critically, the FDS also exhibited significant activation during wrist flexion (SNR of 25.5 dB), indicating considerable crosstalk for this compensatory motion. However, its signal was markedly lower during wrist extension (SNR not exceeding 8.0 dB). Despite the observed crosstalk, the FDS is confirmed as the optimal choice, as its signal-to-noise ratio during the primary task of finger flexion is substantially higher than that of the FCR. This quantitative evidence thus reinforces the selection of the FDS for its superior performance in isolating the target movement, while the FCR is validated as an effective reference channel for monitoring compensatory wrist flexion.

3 Conclusion

This work demonstrated that a preliminary anatomical investigation provides a systematic basis for muscle selection in forearm sEMG studies, validating the selection of the FDS which not only yielded a superior Signal-to-Noise Ratio (SNR) during the primary finger flexion task but also exhibited minimal activation during compensatory wrist movements, confirming its high selectivity. Furthermore, the identification of the FCR as an effective indicator of unintended wrist motion enhances the protocol's robustness. Overall, this anatomically-informed validation framework provides a validated methodology for enhancing the reliability of forearm sEMG acquisition in clinical and research applications.

Acknowledgement

This work was supported by national funds: UID/05757 - Research Centre in Digitalization and Intelligent Robotics (CeDRI); and SusTEC, LA/P/0007/2020 (DOI: 10.54499/LA/P/0007/2020).

References

1. Gazzoni, M., Celadon, N., Mastrapasqua, D., Paleari, M., Margaria, V., Ariano, P.: Quantifying forearm muscle activity during wrist and finger movements by means of multi-channel electromyography **9**(10), e109943 (2014). https://doi.org/10.1371/journal.pone.0109943
2. Hermens, H.J., Freriks, B., Disselhorst-Klug, C., Rau, G.: Development of recommendations for SEMG sensors and sensor placement procedures **10**(5), 361–374 (2000). https://doi.org/10.1016/s1050-6411(00)00027-4
3. Luiz, L.E., Da Silva, W.J., Soares, S., Leitão, P., Teixeira, J.P.: Portable system and user interface for ECG and EMG acquisition, conditioning, and parameters extraction **256**, 1216–1223. https://doi.org/10.1016/j.procs.2025.02.231
4. Merletti, R.: Standards for reporting emg data. Journal of Electromyography and Kinesiology **9**(1), III–IV (1999), https://isek.org/wp-content/uploads/2015/05/Standards-for-Reporting-EMG-Data.pdf
5. Mitchell, B., Whited, L.: Anatomy, shoulder and upper limb, forearm muscles. In: StatPearls. StatPearls Publishing (2025), http://www.ncbi.nlm.nih.gov/books/NBK536975/
6. NYSORA: Antebraço (2022), https://www.nysora.com/pt/m%C3%BAsculo-esquel%C3%A9tico/antebra%C3%A7o/, accessed: 07 May 2025

Hybrid Heuristic-Metaheuristic Strategies for Solving CVRPTW with OR-Tools

Arthur Brocco[1], Adriano S. Silva[2], Ana I. Pereira[2], and Florbela P. Fernandes[2]

[1] Instituto Politécnico de Bragança, 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 (IPB), Bragança, Portugal
a57652@alunos.ipb.pt,{adriano.santossilva, apereira, fflor}@ipb.pt

Abstract. This work explores hybrid heuristic–metaheuristic strategies using OR-Tools to solve the CVRPTW on Solomon benchmark instances. Combining initial solutions from classical heuristics with Tabu Search or Guided Local Search significantly improves solution quality. Results show that no single method is best for all cases, highlighting the potential of adaptive strategies based on instance characteristics.

Keywords: Optimization · Logistics · Transportation.

1 Introduction

The Capacitated Vehicle Routing Problem with Time Window (CVRPTW) is one of the most challenging and studied variants in the logistics field, as it adds temporal constraints to routes, reflecting real-world scenarios of goods distribution [2]. Developing algorithms for route optimization (minimization problem) is complex due to the strong influence of dataset characteristics on results, which is why standard benchmark datasets are commonly used for comparison (*e.g.* Solomon instances [3]). To solve the CVRPTW, methods such as savings-based algorithms, insertion heuristics, and metaheuristics are often applied. Google OR-Tools [1] is particularly effective in this context, offering a flexible and scalable library with built-in support for time windows, capacity constraints, and diverse search strategies.

2 Methodology

Solomon's benchmark instances, categorized as 'R' (random), 'C' (clustered), and 'RC' (mixed), each include 100 customers, one depot, capacity constraints, and time windows. Selected instances (C101–C109, C201–C208, R101–R112, R201–R211, RC101–108, RC201–208) were solved using Google OR-Tools with the CP-SAT solver. Initial solutions were generated using heuristics: Global Cheapest Arc, Local Cheapest Arc and Path Cheapest Arc (GCA, LCA and PCA, respectively), then refined with metaheuristics Guided Local Search and Tabu Search (GLS and TS, respectively). Hybrid approaches combining both were also tested. Each run was limited to 10 seconds, and results are reported as average distances for valid solutions across the C1 (9 instances), C2 (8 instances), R1 (12 instances), R2 (11 instances), RC1 (8 instances), and RC2 (8 instances) datasets.

3 Results and Discussion

The results obtained for the single-run heuristics and metaheuristics, along with the combination between heuristics and metaheuristics, are shown in Fig. 1.

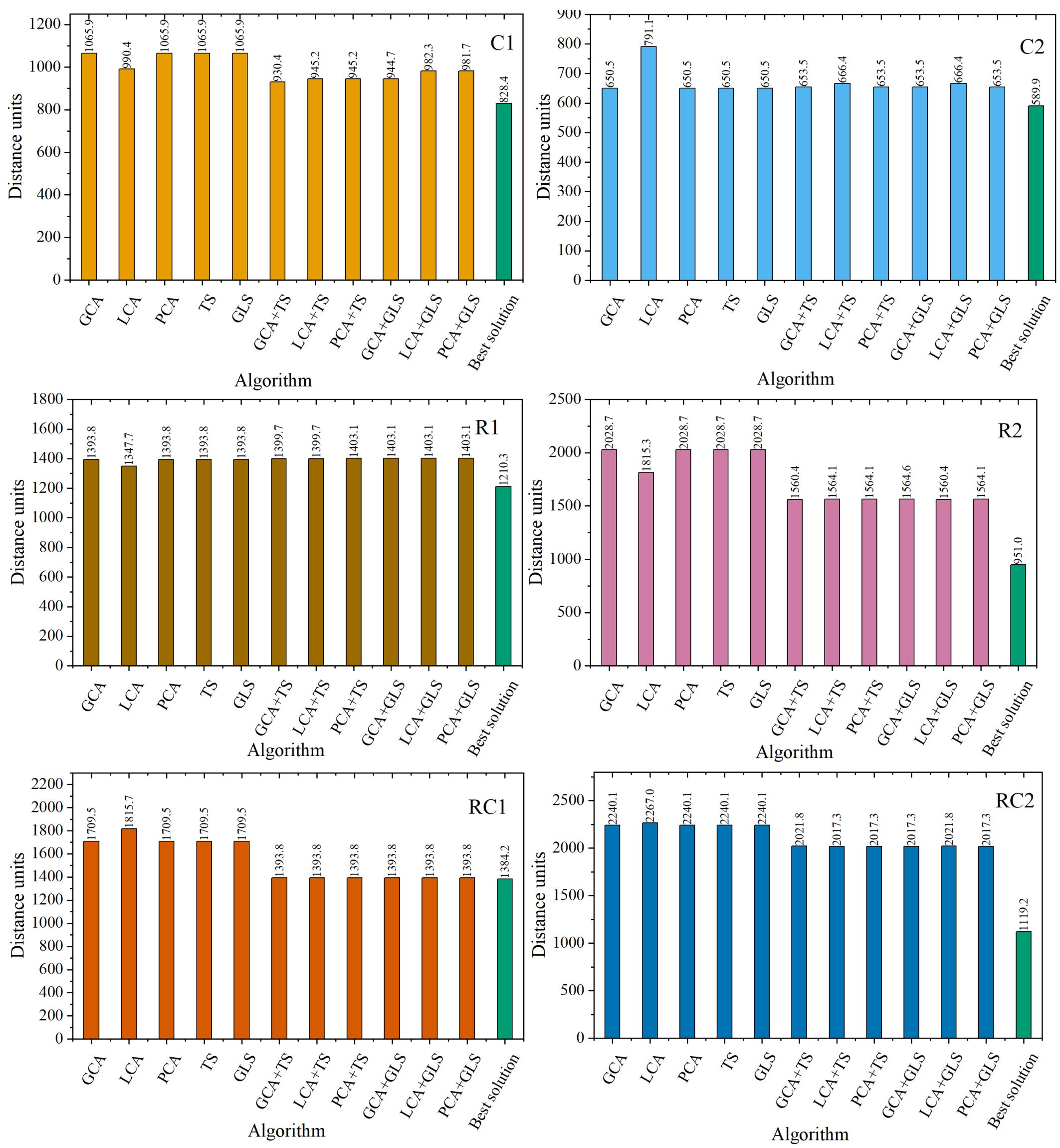


Fig. 1: Results obtained using single heuristics (GCA, LCA, PCA) and metaheuristics (TS and GLS), and the combination between heuristics and metaheuristics.

The results demonstrate clear differences in solution quality depending on the choice of heuristic and whether a metaheuristic is applied. Among the standalone heuristics, GCA and PCA had identical average results across all instances, which suggests that the search space explored by these methods is either very limited or leads to the same local optima. In contrast, LCA alone produces slightly lower average distances in clustered instances such as C1 and C2, indicating that this heuristic may be better suited for handling customer clusters. However, its performance deteriorates in more complex or randomly distributed instances such as R2 and RC2, where it is outperformed by the other methods. When combined with the TS metaheuristic, all heuristics show consistent improvement in solution quality. The improvement is particularly noticeable in instances like C1, where the average cost drops significantly from the baseline values, and also in RC1 and RC2, which are typically more challenging due to their mixed customer distribution and time window constraints.

Combinations with GLS also lead to better solutions compared to standalone heuristics. However, gains from GLS are generally slightly less pronounced than those from TS, especially in clustered instances. For example, while GCA+TS achieves a lower average cost in C1 compared to GCA+GLS, in other cases such as R1 and RC1, both metaheuristics produce similar improvements. The final column representing the best solution obtained among all strategies for each instance shows substantial reductions in cost, especially in R2 and RC2.

4 Conclusion

These results indicate that while combinations of heuristics and metaheuristics enhance performance, no single approach is dominant across all instance types. This variation suggests that an adaptive approach that selects the most appropriate strategy based on instance characteristics may be more effective than relying on a fixed combination. Overall, integrating metaheuristics into the OR-Tools framework leads to more robust and lower-cost solutions, particularly for the more complex Solomon instances.

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/UIDB/05757/2020) and SusTEC, LA/P/0007/2020 (DOI: 10.54499/LA/P/0007/2020).

References

1. Google: Or-tools: Operations research tools. `https://developers.google.com/optimization` (2023), available at `https://developers.google.com/optimization`
2. Silva, A.S., Alves, F., Diaz de Tuesta, J.L., Rocha, A.M.A.C., Pereira, A.I., Silva, A.M.T., Gomes, H.T.: Capacitated waste collection problem solution using an open-source tool. Computers **12**(1) (2023). https://doi.org/10.3390/computers12010015, `https://www.mdpi.com/2073-431X/12/1/15`
3. Solomon, M.M., Desrosiers, J., Dumas, Y.: Time window constrained routing and scheduling problems. Operations Research **35**(2), 254–265 (1987)

Grapevine Health Diagnosis with Deep Learning

Vinícius Sasaki[1], João Mendes[2], José Lima[2], and Ana I. Pereira[2]

[1] Instituto Politécnico de Bragança, 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 (IPB), Bragança, Portugal
a65415@alunos.ipb.pt, {joao.cmendes, jllima, apereira}@ipb.pt

Abstract. The increasing adoption of mobile robots and drones in precision agriculture has enabled continuous, high-resolution crop monitoring, opening up new opportunities for the early detection of plant health issues. This work proposes an integrated system for the acquisition and analysis of vine leaf and fruit images, aimed at identifying real-time signs of diseases and nutritional deficiencies. The proposed pipeline enables vineyard health monitoring through the segmentation and classification of grapevine leaves using computer vision techniques, namely convolutional neural networks (CNNs). By isolating the leaves from the surrounding structures and classifying their condition, the system aims to support rapid and localized diagnosis of plant health issues, enhancing decision-making in vineyard management.

Keywords: Grapevine · Deep learning · Precision viticulture · Computer vision

1 Introduction

Plant diseases and pests are important factors that affect crop yield and quality, and their identification can be effectively performed through digital image processing. Although digital imaging has long supported general crop-health monitoring, the complex structure of grapevine canopies and the seasonal variability in vineyards demand specialized approaches for reliable leaf and fruit analysis. The use of deep learning has shown significant improvements over traditional methods in this area, leading to growing interest in its application for plant health monitoring [8, 9].

The convergence of advanced technologies – particularly unmanned aerial vehicles and ground-based mobile platforms – alongside computer vision and artificial intelligence is transforming precision viticulture, enabling automated data capture and in-field analysis that were previously impractical [5, 15].

Recent studies demonstrate substantial progress in the use of deep learning for detecting diseases in grapevine leaves, a key factor in maintaining vineyard health and ensuring successful harvests. Komala et al. [7] reviewed several notable contributions, including the Grape Leaf Disease Detection Network (GLDDN) by Dwivedi et al. [4], which achieved 99.93 % accuracy using ResNet-18. Advances in convolutional architectures, such as those highlighted by Simonyan and Zisserman [13] and the application of Faster R-CNN with VGG16 under field conditions [12, 14] further illustrate this trend.

Beyond algorithmic advances, hardware platforms play a crucial role. RGB, thermal, and spectral cameras mounted on drones and ground vehicles have been employed to monitor large vineyard areas, detecting diseases, water stress, plant vigor, missing vines, and estimating yield [3, 6, 10]. Meanwhile, the democratization of digital sensors and storage has driven rapid growth in computer vision, and advances in AI and machine learning have enabled the processing of complex scenes and the automation of essential tasks, such as bunch counting for yield prediction. Both classical techniques (segmentation, shape recognition, feature extraction) and CNNs that learn representations directly from raw images are employed [11]. Notably, instance-segmentation architectures such as Mask R-CNN — enhanced with Soft-NMS post-processing — have achieved a +4.3 % gain in mAP and an F1-score of 96.13 % for grape bunch counting in high-resolution vineyard imagery [1]. Similarly, semantic-segmentation networks like U-Net outperform traditional classifiers (e.g. SVM, k-means), reaching Intersection over Union (IoU) scores up to 0.76 on fruit segmentation under natural backgrounds – about 17 pp above classical methods, underscoring the robustness of encoder–decoder CNNs in field conditions [2].

This study proposes an integrated drone and mobile-robot system featuring simultaneous leaf and fruit segmentation, optimized for real-time operation, and evaluates its robustness across diverse seasonal conditions.

The remainder of this paper is as follows. Following the Introduction, Section 2 outlines the methodological framework, detailing the system architecture and underscoring the article's central contributions. Section 3 then summarizes the findings and proposes avenues for future research.

2 Methodology

This work proposes the development of an intelligent system for the analysis of grapevine leaves and fruits, aiming to detect, in real time, signs of diseases, nutritional deficiencies, and other physiological anomalies. To achieve this, an architecture was designed using mobile sensors embedded in drones and ground robots, which will navigate through the vineyard to systematically capture images, as illustrated in Fig. 1.

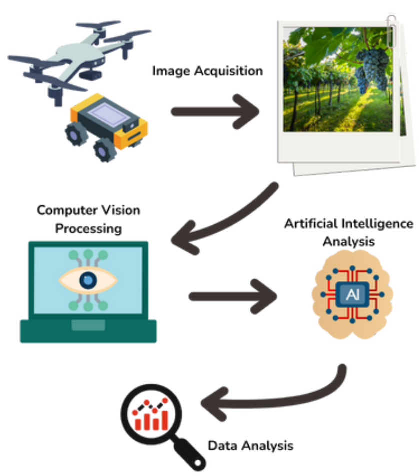


Fig. 1: Diagram illustrating the system architecture

One of the system's core components is the construction of a representative dataset composed of images collected under different environmental and operational conditions.

Data collection will be carried out throughout the different seasons of the year, allowing the observation of how seasonality affects crop health.

The captured images will undergo a preprocessing phase, including automatic cropping of regions of interest, resizing to a standardized format, and color normalization. The system uses the YOLO (You Only Look Once) architecture to detect vine leaves and fruits in real time, leveraging its efficiency in object detection under diverse conditions. The use of both aerial (drone) and ground (robot) platforms allows for image acquisition from multiple angles and heights, improving robustness and coverage.

After detection, a semantic segmentation model is applied to isolate the detected elements from the background. The segmented regions are then classified using convolutional neural networks trained on annotated datasets to distinguish healthy from diseased or nutritionally deficient tissues. The entire pipeline – detection, segmentation and classification – is designed to operate in real time. As a result, the system can immediately generate reports and alerts that support responsive decision-making in vineyard management.

Compared to existing approaches, the proposed system stands out for its simultaneous use of aerial and terrestrial imagery integrated into a unified analysis framework. The combination of YOLO-based detection with deep learning classifiers allows accurate, localized diagnostics across various environmental conditions. Furthermore, the incorporation of seasonal data enhances model generalization by accounting for phenological variability in grapevine development.

To support real-time inference in field conditions, the system relies on embedded hardware platforms with accelerated computing capabilities, such as the NVIDIA Jetson series. These devices offer a balance between computational power and energy efficiency, allowing onboard execution of detection and classification models without depending on cloud services.

The performance of the detection, segmentation, and classification models will be evaluated using standard metrics. These include accuracy, precision, recall, F1-score, mean Average Precision (mAP), and Intersection over Union (IoU), providing a comprehensive understanding of both classification quality and segmentation precision.

Despite its advantages, the system faces several limitations. Image quality is highly dependent on lighting conditions, which can vary significantly in outdoor environments. Furthermore, occlusion caused by dense foliage or overlapping plant structures may hinder detection and segmentation accuracy. Finally, the current model may require retraining or fine-tuning to adapt to different grapevine cultivars or geographical regions.

3 Conclusion and future work

In summary, the proposed system aims to support precision viticulture by integrating mobile robotics, aerial imagery, and deep learning. This approach is expected to contribute to early anomaly detection and informed decision-making in vineyard management by enabling the segmentation and classification of vine leaves and fruits under varying field conditions.

As a possible direction for future work, the integration of visual data with environmental variables, such as temperature, humidity, wind, and soil conditions, obtained via IoT sensors, could enhance the system's predictive capabilities. Hybrid AI models that combine supervised and unsupervised learning may also be explored to increase robustness and reduce false positives under complex field conditions.

Acknowledgment

This work was supported by national funds: UID/05757 - Research Centre in Digitalization and Intelligent Robotics (CeDRI); and SusTEC, LA/P/0007/2020 (DOI: 10.54499/LA/P/0007/2020).

References

1. Chauhan, A., Singh, M., et al.: Computer vision and machine learning based grape fruit cluster detection and yield estimation robot. Journal of Scientific & Industrial Research **81**(08), 866–872 (2022)
2. Chen, Y., Li, X., Jia, M., Li, J., Hu, T., Luo, J.: Instance segmentation and number counting of grape berry images based on deep learning. Applied Sciences **13**(11), 6751 (2023)
3. Di Gennaro, S.F., Toscano, P., Cinat, P., Berton, A., Matese, A.: A low-cost and unsupervised image recognition methodology for yield estimation in a vineyard. Frontiers in plant science **10**, 559 (2019)
4. Dwivedi, R., Dey, S., Chakraborty, C., Tiwari, S.: Grape disease detection network based on multi-task learning and attention features. IEEE Sensors Journal **21**(16), 17573–17580 (2021)
5. Figueiredo, N., Neto, A., Cunha, A., Sousa, J.J., Sousa, A.: Deep learning approach for terrace vineyards detection from google earth satellite imagery. In: IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium. pp. 5824–5827. IEEE (2022)
6. Kerkech, M., Hafiane, A., Canals, R.: Vine disease detection in uav multispectral images using optimized image registration and deep learning segmentation approach. Computers and Electronics in Agriculture **174**, 105446 (2020)
7. Komala, K., Lata, B., Venugopal, K.: Deep learning for accurate vineyard pathology detection. In: 2023 IEEE 5th PhD Colloquium on Emerging Domain Innovation and Technology for Society (PhD EDITS). pp. 1–2. IEEE (2023)
8. Lee, S.H., Chan, C.S., Mayo, S.J., Remagnino, P.: How deep learning extracts and learns leaf features for plant classification. Pattern recognition **71**, 1–13 (2017)
9. Liu, J., Wang, X.: Plant diseases and pests detection based on deep learning: a review. Plant methods **17**, 1–18 (2021)
10. Matese, A., Di Gennaro, S.F.: Practical applications of a multisensor uav platform based on multispectral, thermal and rgb high resolution images in precision viticulture. Agriculture **8**(7), 116 (2018)
11. Mohimont, L., Alin, F., Rondeau, M., Gaveau, N., Steffenel, L.A.: Computer vision and deep learning for precision viticulture. Agronomy **12**(10), 2463 (2022)
12. Rehana, H., Ibrahim, M., Ali, M.H.: Plant disease detection using region-based convolutional neural network. arXiv preprint arXiv:2303.09063 (2023)
13. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
14. Song, Z., Fu, L., Wu, J., Liu, Z., Li, R., Cui, Y.: Kiwifruit detection in field images using faster r-cnn with vgg16. IFAC-PapersOnLine **52**(30), 76–81 (2019)
15. Valente, J., Sanz, D., Barrientos, A., Del Cerro, J., Ribeiro, Á., Rossi, C.: An air-ground wireless sensor network for crop monitoring. Sensors **11**(6), 6088–6108 (2011)

Sponsors

- Fundação para a Ciência e a Tecnologia, FCT

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