



SMART 2025

The Fourteenth International Conference on Smart Systems, Devices and
Technologies

ISBN: 978-1-68558-251-7

April 6 - 10, 2025

Valencia, Spain

SMART 2025 Editors

Lasse Berntzen, University of South-Eastern Norway, Norway

SMART 2025

Forward

The Fourteenth International Conference on Smart Cities, Systems, Devices and Technologies (SMART 2025), held between April 6th, 2025, and April 10th, 2025, in Valencia, Spain, continued a series of co-located events covering tendencies towards future smart cities, specialized technologies and devices, environmental sensing, energy optimization, pollution control and socio-cultural aspects.

Digital societies take rapid developments toward smart environments. More and more social services are digitally available to citizens. The concept of 'smart cities' including all devices, services, technologies and applications associated with the concept sees a large adoption. Ubiquity and mobility added new dimensions to smart environments. Adoption of smartphones and digital finder maps, as well as increasing budgets for technical support of services to citizens, settled a new behavioral paradigm of city inhabitants.

We take here the opportunity to warmly thank all the members of the SMART 2025 technical program committee, as well as all the reviewers. The creation of such a high-quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to SMART 2025. We truly believe that, thanks to all these efforts, the final conference program consisted of top-quality contributions. We also thank the members of the SMART 2025 organizing committee for their help in handling the logistics of this event.

We hope that SMART 2025 was a successful international forum for the exchange of ideas and results between academia and industry for the promotion of progress in the area of smart cities, systems, devices, and technologies.

SMART 2025 Chairs

SMART 2025 Steering Committee

Lasse Berntzen, University of South-Eastern Norway, Norway
Young-Joo Suh, POSTECH, Korea

SMART 2025 Publicity Chairs

Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politècnica de València, Spain
José Miguel Jiménez, Universitat Politècnica de València, Spain

SMART 2025 Committee

SMART 2025 Steering Committee

Lasse Berntzen, University of South-Eastern Norway, Norway
Young-Joo Suh, POSTECH, Korea

SMART 2025 Publicity Chairs

Francisco Javier Díaz Blasco, Universitat Politècnica de València, Spain
Ali Ahmad, Universitat Politècnica de València, Spain
Sandra Viciano Tudela, Universitat Politècnica de València, Spain
José Miguel Jiménez, Universitat Politècnica de València, Spain

SMART 2025 Technical Program Committee

Lounis Adouane, Université de Technologie de Compiègne / Heudisayc, France
Ramakalyan Ayyagari, National Institute of Technology Tiruchirappalli, India
Karolina Baras, University of Madeira, Portugal
Lasse Berntzen, University of South-Eastern Norway, Norway
DeJiu Chen, KTH Royal Institute of Technology, Sweden
Patrizio Dazzi, University of Pisa, Italy
María del Mar Gómez Zamora, Universidad Rey Juan Carlos, Spain
Daniel Delahaye, ENAC LAB, Toulouse, France
Rachid El Bachtiri, USMBA University, Morocco
Abdelali Elmoufidi, University of Hassan II Casablanca, Morocco
Xiang Fang, Nanyang Technological University, Singapore
Steffen Fries, Siemens AG, Germany
Marco Furini, University of Modena and Reggio Emilia, Italy
Angelo Furno, LICIT laboratory - University of Gustave Eiffel / University of Lyon, France
Ivan Ganchev, University of Limerick, Ireland / University of Plovdiv "Paisii Hilendarski", Bulgaria
Maria Helena Gomes de Almeida Gonçalves Nadaís, University of Aveiro, Portugal
Jorge J. Gomez-Sanz, Universidad Complutense de Madrid, Spain
Javier Gozálvez, Universidad Miguel Hernández de Elche, Spain
Damas Gruska, Comenius University, Slovakia
Fakhrul Hazman Yusoff, Sohar University, Oman
Gerold Hölzl, University of Passau, Germany
Tzung-Pei Hong, National University of Kaohsiung, Taiwan
Javier Ibanez-Guzman, Renault, France
Sergio Ilarri, University of Zaragoza, Spain
Abdelmajid Khelil, Landshut University of Applied Sciences, Germany
Pinar Kirci, Bursa Uludag University, Turkey
Hongzhi Kuai, Maebashi Institute of Technology, Japan
Tshepo Godfrey Kukuni, Central University of Technology, South Africa
Wolfgang Leister, Norsk Regnesentral, Norway
Giovanni Livraga, Università degli Studi di Milano, Italy

Giuseppe Mangioni, University of Catania, Italy
Manuel Mastrofini, Università degli studi di Roma Tor Vergata & LAST Horizon, Italy
Kevin Meehan, Atlantic Technological University, Ireland
Natarajan Meghanathan, Jackson State University, USA
Georges Mykoniatis, Ecole Nationale de l'Aviation Civile (ENAC), France
Dmitry Namiot, Lomonosov Moscow State University, Russia
Fawzi Nashashibi, INRIA - Paris Research Centre, France
Ouail Ouchetto, University of Hassan II - Casablanca, Morocco
Wilma Penzo, DISI - University of Bologna, Italy
Cathryn Peoples, Ulster University, UK
Christine Perakslis, Johnson & Wales University, USA
Ermanno Pietrosemoli, International Centre for Theoretical Physics, Italy
Marco Polignano, University of Bari Aldo Moro, Italy
Sherif Rashad, Morehead State University, USA
Marius Rohde Johannessen, University of South-Eastern Norway - School of Business, Norway
Enrique Romero-Cadaval, University of Extremadura, Spain
Demetrios Sampson, University of Piraeus, Greece
Lucie Schmidt, University of Applied Sciences Jena, Germany
Farhan Siddiqui, Dickinson College, USA
Steffen Späthe, Friedrich Schiller University Jena, Germany
Young-Joo Suh, POSTECH, Korea
Konstantinos Votis, Information Technologies Institute | Centre for Research and Technology Hellas, Thessaloniki, Greece
Qinggang Wu, Second Research Institute of Civil Aviation Administration of China, China
Mudasser F. Wyne, National University, USA
Wuyi Yue, Konan University, Japan
Sherali Zeadally, University of Kentucky, USA
Sotirios Ziavras, New Jersey Institute of Technology, USA

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission to reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Smart City Road Maintenance: A LiDAR and AI-Driven Approach for Detecting and Mapping Road Defects <i>Giovanni Nardini, Roberto Nucera, Alessandro Ulleri, Stefano Cordiner, Eugenio Martinelli, Arianna Mencattini, and Iulian Gabriel Coltea</i>	1
IoT System for Indoor Air Quality Monitoring Using ThingSpeak: Promoting Healthy Work Environments <i>Ilber Adonayt Ruge Ruge, Ingrid Carolina Ortiz Alvarez, and Fabian Rolando Jimenez Lopez</i>	5
Deep Learning-based Failure Detection for Safety Diagnostics of Hydrogen Storage Vessels <i>Dongju Kim, Da-Hyun Kim, Hyo-Jin Kim, and Young-Joo Suh</i>	10
Disaster Detection Framework for Smart Cities: An AI YOLOv8 and IoT Approach <i>Hossam Kamel</i>	16
On-the-Edge Inference Enabled Vision System for Smart Cities <i>Carmelo Scribano, Ignacio Sanudo Olmedo, Micaela Verucchi, Danda Pani Paudel, Marko Bertogna, and Luc Van Gool</i>	24
Enhancing Smart City Sustainability: Anaerobic Treatment of Semiconductor Industry Wastewater for Improved Efficiency and Environmental Impact <i>Bruno Escobar, Ana Santos, Flavio Silva, and Helena Nadais</i>	29
Using Artificial Intelligence to Support Emergency Management Training - Creating Efficient and Realistic Scenarios with ChatGPT <i>Lasse Berntzen, Jarle Lowe Sorensen, Marius Rohde Johannessen, Lauritz Rauer Nilsen, Lene Sandberg, Eric Carlstrom, and Amir Khorram-Manesh</i>	34

Smart City Road Maintenance: A LiDAR and AI-Driven Approach for Detecting and Mapping Road Defects

Giovanni Nardini*, Roberto Nucera*, Alessandro Ulleri*, Stefano Cordiner[†],
Eugenio Martinelli[‡], Arianna Mencattini[‡], Iulian Gabriel Coltea*

*Key to business s.r.l., Department of [†]Industrial and [‡]Electronic Engineering, University of Rome Tor Vergata
Rome, Italy

*e-mail: info@key2.it

Abstract—This work is focused on the development of an integrated system designed to detect, map, and analyze road surface defects, contributing to Smart City infrastructure maintenance. The system is installed on vehicles and leverages a multi-sensor approach, combining Light Detection and Range (LiDAR) point clouds, visual information from Red-Green-Blue (RGB) cameras, inertial data and Global Navigation Satellite Systems (GNSS) coordinates. Road defects such as potholes and alligator cracks are detected in RGB images by a custom deep learning model based on instance segmentation. The scene understanding is committed to a second Artificial Intelligence (AI) model based on semantic segmentation in order to perceive objects locations and the overall structure of the road. Afterward, all results are processed together and translated into the 3D domain of LiDAR data. This can be done through a proper camera calibration procedure and LiDAR-Camera data alignment with the estimation of intrinsic and extrinsic parameters. Then, AI segmentation results are projected to 3D point clouds in order to isolate the detected items from the rest of the point cloud and obtain three-dimensional models of each of them, enabling measurements like the affected surface extension, depth and volumes. GNSS and inertial data are fused together to obtain the correct orientation and location of the system, enabling geographic positioning of all detected items on the map. Results are displayed on a map-based portal, enabling easy access to near real-time defect data. This approach advances road monitoring by automating the mapping and analysis of surface conditions, enhancing urban infrastructure management. In addition, the strengths of this approach are the possibility of deploying the pipeline in edge devices enabling real-time computation, the use of pre-existing training datasets based on RGB images alone, and good accuracy on the geographical localization and estimation of defect measurements.

Keywords—cities; road maintenance; LiDAR; AI; computer vision.

I. INTRODUCTION

As urban areas grow the need to monitor road conditions efficiently becomes crucial for keeping infrastructure intact and promoting road safety. The conventional methods of inspecting roads are laborious, time consuming and frequently fall short of providing the accuracy required for repairs. However, recent progress in sensor technology, artificial intelligence and data integration present fresh opportunities for monitoring road conditions. Over the past few years, many approaches have been explored. Sometimes using inertial data [1], pure machine learning and computer vision methods [2][3], sometimes exploiting more sophisticated deep learning models [4], and other times combining vision and depth sensing together with spatial AI [5][6]. The technologies that have been tested for

depth estimation are based on stereoscopy, Red-Green-Blue-Depth (RGB-D) cameras and LiDAR. However, each has its own disadvantages: stereoscopy generally does not work with feature-poor surfaces, RGB-D cameras based on Time of Flight (ToF) technology, while achieving good accuracy, drop their performance in outdoor environments and are limited to a range of few meters, while LiDAR provides the most long-range and accurate measurements but at the expense of lower point density and the need for an additional imaging system to obtain the scene picture. Furthermore, approaches using RGB-D images as input for AI detection models, while achieving good performance due to depth information, are strongly affected by the context, sensor position and framing of the training data, and therefore require the acquisition of huge amounts of images from every possible angle and distance, in order to replicate all possible setups. Our approach, on the other hand, bases AI inference solely on RGB images and transports the detection information to the LiDAR domain, via camera-LiDAR registration, as shown in Figure 1.

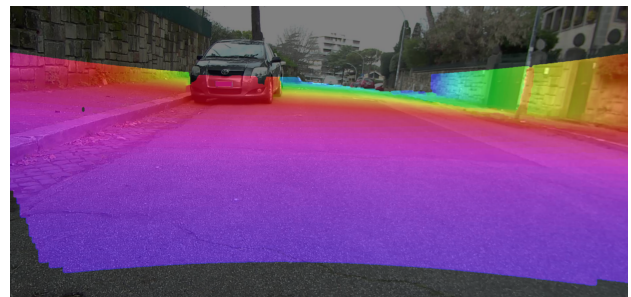


Figure 1. Camera-LiDAR Registration.

This allows the use of pre-existing datasets without having to create a custom dataset and re-labelling all images. This work suggests a setup (see Figure 2) that utilizes LiDAR technology along with RGB imaging, inertial and GNSS data within a framework based on Robot Operating System (ROS), as shown in Figure 3, in order to identify and pinpoint road surface issues efficiently.

From an economic standpoint, the system's adaptability to city vehicles, including public transport, could potentially transform routine operations into continuous, cost-effective road monitoring. Combining this distributed sensing with on-the-ground human supervision, such as cleaning personnel, creates

a hybrid model that optimizes resource use and enhances data accuracy, leading to efficient urban road maintenance.

In Section 2, the methods employed are detailed, including the hardware components, the software architecture, the design and training of the AI models. In Section 3, the results of the system's validation are presented, focusing on the performance metrics of the AI models and the accuracy of defect measurements and positioning. In Section 4, the paper concludes with a discussion of the system's contributions and potential future developments.

II. METHODS

The proposed system integrates the following hardware components: an Hybrid Solid-State LiDAR with 128-channels of resolution, a global shutter camera sensor with 4k resolution at 30 fps, a navigation system with 9-axis accelerometer INS and dual antenna GNSS and a Nvidia Jetson AGX module where the software runs. The LiDAR, camera and navigation system are mounted on the vehicle's roof, while the Jetson unit is installed inside the cabin and connected to the vehicle's power supply.

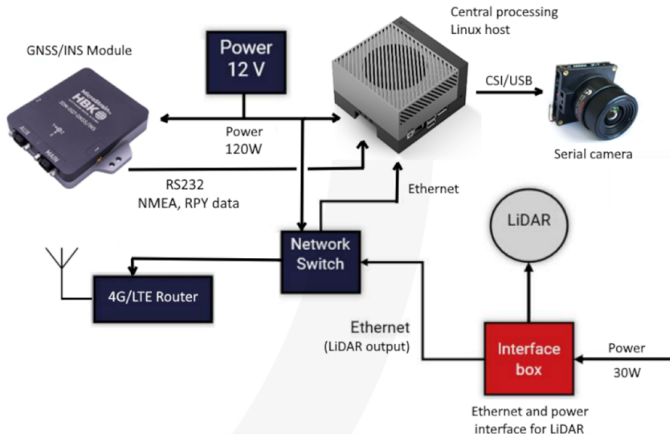


Figure 2. Hardware setup.

The software architecture is based on ROS and is made up of the following nodes: driver nodes to collect data from each sensor and publish to topics, data processing nodes to apply AI model inference on images and get results, projection nodes to map defects from RGB domain into 3D domain and make measurements and navigation nodes to estimate precise latitudes and longitudes of each defects. All collected results are then submitted to the visualization platform.

There are two types of custom trained AI models: an instance segmentation model based on the You Only Look Once (YOLOV8-Seg) small architecture [7] and a semantic segmentation model based on the SegFormerB1 architecture [8]. The former was trained on the RDD22 dataset [9]: since it is an object detection dataset, it was necessary to re-label the annotations with the addition of pothole and alligator crack segmentation mask. In order to speed up the process, the Segment Anything Model (SAM) [10] was adopted, enabling a quick annotation of the images from the bounding boxes using a

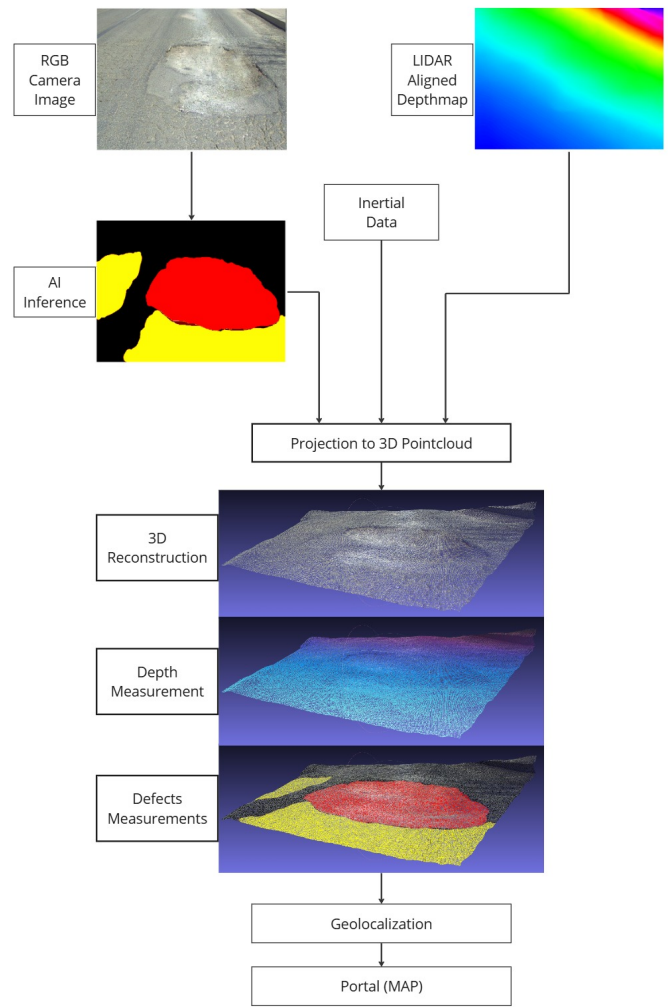


Figure 3. Software architecture components in ROS framework.

dedicated tool. The latter was trained on the Cityscapes dataset [11], which provides over 5000 densely annotated images with 30 segmentation classes. The purpose of the two models is to identify potholes and alligator cracks in road images captured by the camera, along with their segmentation masks, and to verify their placement within the "road" class of the semantic segmentation model, in order to limit false positives. Qualitative results for both models are shown in Figure 4.

III. RESULTS & DISCUSSION

The system was rigorously tested on various urban road sections, demonstrating strong performance in identifying potholes and alligator cracks. The validation of the AI models on dedicated test sets yielded key computer vision metrics that underline their effectiveness. For the YOLO model, a mean Average Precision (mAP) of 0.56 at thresholds of Intersection Over Union (IoU) ranging from 0.5 to 0.95 reflects its robustness in detecting and segmenting defects across different scales and conditions. The mAP is calculated with

the following equation:

$$\text{mAP} = \frac{1}{N_c} \sum_{c=1}^{N_c} \frac{1}{N_{\text{IoU}}} \sum_{i=1}^{N_{\text{IoU}}} \text{AP}_c^{(i)} \quad (1)$$

where N_c is the total number of classes, N_{IoU} is the number of IoU thresholds and $\text{AP}_c^{(i)}$ is the average precision for class c at IoU threshold i .

Additionally, an F1-score of 0.57 indicates a balanced performance in terms of precision (reducing false positives) and recall (capturing true positives). The formula of F1-score is the following:

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R} \quad (2)$$

where P is the Precision and R is the Recall value. For the SegFormerB1 model used in road segmentation, a mean Intersection over Union (meanIoU) score of 0.43 demonstrates its capacity to accurately delineate the "road" class, while an exceptional F1-score of 0.98 highlights its precision and reliability in avoiding misclassifications. The following formulas provide the way for calculating meanIoU for semantic segmentation:

$$\text{IoU}_c = \frac{|A_c \cap B_c|}{|A_c \cup B_c|} \quad (3)$$

$$\text{meanIoU} = \frac{1}{N_c} \sum_{c=1}^{N_c} \text{IoU}_c \quad (4)$$

where A_c is the predicted segmentation for class c , B_c is the ground truth segmentation for class c and N_c is the number of classes. These metrics are considered strong, given the complexity of urban environments and variability in road textures.



Figure 4. YOLOV8s-Seg results for pothole and crack segmentation (left), SegFormerB1 road segmentation results (right).

Moreover, the integration of segmentation results with LiDAR data allowed for accurate 3D reconstruction and spatial measurements, achieving an error margin of less than 10% for defect dimensions (surface area and depth). The system's navigation module further enhanced functionality, delivering geolocation with Global Positioning System (GPS) accuracy suitable for effective road management applications at the city scale. Finally, deployment on the Nvidia Jetson AGX 64 GB Edge device and model optimization using TensorRT enabled real-time processing. The YOLO model achieved a remarkable throughput of 312 Frames-Per-Second (FPS), while the SegFormer model delivered 18 FPS, ensuring a processing rate exceeding 10 Hz—well-aligned with the LiDAR's sampling

rate. This ensures that the system can operate seamlessly in real-time, offering both efficiency and scalability.

IV. CONCLUSION

Our work presents an advanced, integrated system for detecting and mapping road surface defects, marking a significant step forward in Smart City infrastructure maintenance. By leveraging a multi-sensor approach, including LiDAR, RGB cameras, inertial data, and GNSS, the system achieves precise localization and accurate measurements of defects like potholes and alligator cracks. The innovative application of RGB-based AI models combined with LiDAR domain projection enables the use of existing datasets, minimizing the need for extensive retraining. Deployment on an edge device ensures real-time processing, while the ROS-based framework facilitates seamless data integration and visualization. The achieved accuracy in defect detection, spatial measurement, and geolocation demonstrates the system's potential for scalable implementation in urban road management. Future developments could further enhance adaptability to diverse environments, driving even greater efficiency in urban infrastructure maintenance.

ACKNOWLEDGEMENT

This project has been co-financed by the European Union through the PR FESR 2021–2027 RSI program of Regione Lazio, managed by LazioInnova. The authors would like to thank the European Union and Regione Lazio for their support in enabling this research. Additionally, we extend our gratitude to all partners and collaborators who contributed to the successful implementation and validation of the proposed system.

REFERENCES

- [1] S. Girisan, T. V. Sreelakshmi, N. V. Swetha, V. Suresh, and K. M. Vipin, "Pothole Detection Based on Accelerometer Method", *International Journal of Research in Engineering, Science and Management*, vol. 3, issue 5, May 2020.
- [2] A. Ahmadi, S. Khalesi, and M. R. Bagheri, "Automatic road crack detection and classification using image processing techniques, machine learning and integrated models in urban areas: A novel image binarization technique", *Journal of Industrial and Systems Engineering*, vol. 11, pp. 85-97, 14th International Industrial Engineering Conference, 2018.
- [3] K. L. Qi, "Fast and Accurate Road Crack Detection Based on Adaptive Cost-Sensitive Loss Function", *Institute of Electrical and Electronics Engineers (IEEE)*, vol. 53, pp. 1051–1062, *IEEE Transactions on Cybernetics*, February 2023.
- [4] Y. Li, C. Yin, Y. Lei, J. Zhang, and Y. Yan, "RDD-YOLO: Road Damage Detection Algorithm Based on Improved You Only Look Once Version 8", *Applied Sciences*, vol. 14, April 2024.
- [5] E. Tran et al., "SHREC 2022: Pothole and crack detection in the road pavement using images and RGB-D data", *SHREC2022 3D Shape Retrieval Challenge*, 2022.
- [6] S. A. Talha, "A LiDAR-camera fusion approach for automated detection and assessment of potholes using an autonomous vehicle platform", *Innovative Infrastructure Solutions*, vol. 8, September 2023.
- [7] J. Zeng, H. Ouyang, M. Liu, L. U. Leng, and X. Fu, "Multi-scale YOLACT for instance segmentation", *Journal of King Saud University - Computer and Information Sciences*, vol. 34, October 2022.
- [8] E. Xie et al., "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers", *Neural Information Processing Systems (NeurIPS)*, 2021.

- [9] D. Arya, H. Maeda, S. K. Ghosh, D. Toshniwal, and Y. Sekimoto, "RDD2022: A multi-national image dataset for automatic Road Damage Detection", Crowdsensing-based Road Damage Detection Challenge (CRDDC), 2022.
- [10] A. Kirillov et al., "Segment Anything", Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 4015-4026, 2023.
- [11] M. Cordts et al., "The Cityscapes Dataset for Semantic Urban Scene Understanding", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3213-3223, 2016.

IoT System for Indoor Air Quality Monitoring Using ThingSpeak: Promoting Healthy Work Environments

Ilber Adonayt Ruge Ruge
 Universidad Pedagógica y Tecnológica de Colombia
 Facultad de Ingeniería
 Tunja, Colombia
 email: ilber.ruge@uptc.edu.co

Ingrid Carolina Ortiz Álvarez
 Universidad Pedagógica y Tecnológica de Colombia
 Facultad de Ingeniería
 Tunja, Colombia
 email: ingrid.ortiz05@uptc.edu.co

Fabian Rolando Jiménez López
 Universidad Pedagógica y Tecnológica de Colombia
 Facultad de Ingeniería
 Tunja, Colombia
 email: fabian.jimenez02@uptc.edu.co

Abstract— This article presents the development of an embedded system for air quality monitoring in enclosed environments, utilizing the ESP8266 microcontroller and the Arduino® IDE. The system integrates DHT11 and MQ-4 sensors to measure temperature, humidity, and CO₂ concentration, respectively. The collected data is displayed on an LCD screen and stored on a 4GB microSD card in Comma-Separated Values (CSV) format, ensuring easy access for further analysis. A DS3231 real-time clock module timestamps each measurement, enhancing data organization. Additionally, an alert system is implemented to log events when CO₂ levels exceed a predefined threshold. To enable remote monitoring, the system transmits data in real-time to the ThingSpeak platform using the ESP8266's WiFi connectivity, allowing for interactive visualization and analysis. This implementation provides a low-cost, scalable solution for air quality assessment, facilitating decision-making in indoor environmental management.

Keywords—Air quality monitoring; ESP8266-embedded system; Internet of Things; ThingSpeak.

I. INTRODUCTION

Monitoring indoor air quality is essential for protecting health and well-being, as prolonged exposure to high carbon dioxide levels (CO₂) and other pollutants can cause headaches, fatigue, and cognitive impairment. While elevated CO₂ levels in classrooms do not pose direct health risks, they indicate insufficient ventilation for the number of occupants, as human exhalation increases these levels. Therefore, ensuring proper ventilation is crucial to maintaining air quality [1], [2].

To address this issue, a monitoring system based on the ESP8266 microcontroller was designed, incorporating CO₂, temperature, and humidity sensors. Measuring temperature and humidity helps establish correlations between CO₂ concentration and the presence and number of people in the monitored space using artificial intelligence techniques, although this analysis falls outside the scope of this article.

The electronic device displays the measured variables on an LCD screen, records critical CO₂ levels in a local storage module, and transmits the data in real-time to MathWorks® ThingSpeak, an Internet of Things (IoT) platform. The system is programmed in C/C++, using the Arduino® IDE, which allows efficient control of the microcontroller's resources and facilitates the integration of essential functionalities to ensure optimal system performance.

This paper is organized as follows: Section 2 – Related Work reviews prior research and recent developments in air quality monitoring systems and IoT platforms. Section 3 – Materials and Methods describes the hardware selection process and the algorithmic logic implemented for data acquisition and processing. Section 4 – Results and Discussion presents the collected IoT data along with a detailed cost analysis. Finally, Section 5 – Conclusion and Future Work summarizes the key findings and outlines potential improvements for future implementations.

II. RELATED WORK

Several studies have explored microcontroller- and sensor-based platforms for environmental monitoring. Chouhan et al. [3] found that the MQ-4 sensor effectively detects variations in CO₂ concentration. Kadir et al. [4] demonstrated that the ThingSpeak platform is well-documented for its capability to store and display real-time data, facilitating remote analysis and decision-making.

Srivatsa et al. [5] propose an IoT system comprising three main components: a network of wireless sensors to monitor CO₂ levels, a wireless access point that transmits data via Wi-Fi, and a server that processes and stores the information while issuing alerts when CO₂ levels rise. Salamone et al. [6] present a wireless system for indoor air quality control using an Arduino UNO, a K30 CO₂ sensor, an XBee S2 communication module, and a DS1307 RTC module.

Bhattacharya et al. [7] describe a wireless solution for monitoring indoor air quality, measuring temperature, humidity, gaseous pollutants, and particulate matter. The system utilizes the Air Quality Index to regulate ventilation and air conditioning in smart buildings. Similarly, Wang et al. [8] investigate the impact of air quality on students' concentration during classes by measuring temperature, relative humidity, and CO₂ levels using a wireless system with DHT11 and MG-811 sensors. The system was installed in two classrooms, and the results revealed a significant correlation between environmental parameters and students' academic performance. Beyond indoor applications, Liu et al. [9] developed a monitoring system for urban air pollution, consisting of a sensor node, a gateway, and a LabVIEW™-controlled platform. This system was deployed on major streets in Taipei to measure carbon monoxide (CO) levels from vehicle emissions.

Finally, Kang et al. [10] designed an advanced air quality monitoring system integrating multiple communication technologies. The system employs Wi-Fi for Internet connectivity, Bluetooth for smart device setup, and a Radio Frequency (RF) module (IEEE 802.15.4g) for home network integration. Based on the TI MSP430 processor, it includes sensors for particulate matter, volatile organic compounds, carbon monoxide, temperature, and humidity, while supporting additional sensors via Universal Asynchronous Receiver-Transmitter (UART), Serial Peripheral Interface (SPI), and Inter-Integrated Circuit (I²C) interfaces. The results demonstrated that the system enables real-time air quality monitoring with high resolution.

III. MATERIALS AND METHODS

Poor Indoor Air Quality (IAQ) poses a significant threat to public health, as individuals spend over 90% of their time indoors, where pollutants such as tobacco smoke, carbon, CO, CO₂, NO₂, and microorganisms can negatively impact health [11]. Although temperature and humidity monitoring is common, real-time air quality monitoring is rarely implemented in most buildings. To address this gap, an IoT-based monitoring system was developed to provide a low-cost, easy-to-install solution capable of measuring key air pollutants in real-time and generating alerts when excessive concentrations are detected.

The embedded system consists of an ESP8266 microcontroller module for Wi-Fi communication, a sensor unit integrating an MQ-4 sensor for CO₂ measurement and a DHT11 sensor for temperature and humidity monitoring, a DS3231 Real-Time Clock (RTC) module for precise timestamping of data, and a 4GB microSD module for local data storage. The recorded data is stored in CSV format, ensuring compatibility with text editors and advanced analysis software such as Microsoft Excel and IBM SPSS Statistics, among others. This electronic system was developed as a final project for the Microcontrollers Course (Semester II-2024) at the School of Electronic Engineering of the Pedagogical and Technological

University of Colombia, Tunja Campus. The project was conducted following the Project-Based Learning (PBL) methodology.

While the current implementation measures a single gas, the system can be expanded by integrating additional MQ-series sensors to enhance its monitoring capabilities. The collected data is stored on the ThingSpeak platform (See Fig. 1), allowing remote access via a website or mobile application. The system's software is structured following a flowchart-based design, ensuring continuous measurement of environmental variables, real-time visualization, and data transmission to the cloud at 15-second intervals.

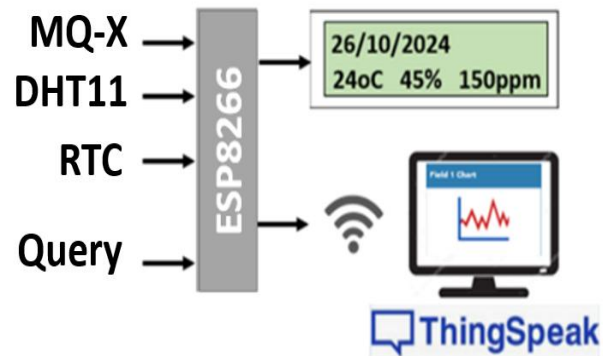


Figure 1. Embedded system block diagram.

Additionally, an alert mechanism was implemented to notify users when CO₂ levels exceed 700 ppm for more than 10 seconds, with this information being stored in Electrical Erasable Programmable Read-Only Memory (EEPROM) memory for future reference.

Once the block diagram was structured, the flowchart was developed, incorporating the main program logic and interrupt routines. This design ensures the orderly execution of the system, optimizes process integration, and facilitates interaction between the sensors, real-time clock module, LCD, and communication with the ThingSpeak IoT platform.

A critical component of the system is the CONSULT interrupt routine, which allows the user to display on the LCD screen the latest alert events related to exposure to excessive gas levels. According to NTC 6199 [12], the permissible CO₂ concentration is set at 700 ppm for an exposure time ranging from 1 to 8 hours. In real-time applications, as shown in Fig. 2, the use of interrupts to manage high-priority events is essential, as it enables immediate system response without disrupting the execution of the main algorithm.

To validate the system's functionality, preliminary testing was conducted using ISIS Proteus 8.17 SP2 simulation software [13], as shown in Fig. 3. The simulation included the Tensilica L106 microcontroller integrated into the ESP8266 module, which provides Wi-Fi connectivity for transmitting data to the ThingSpeak platform.

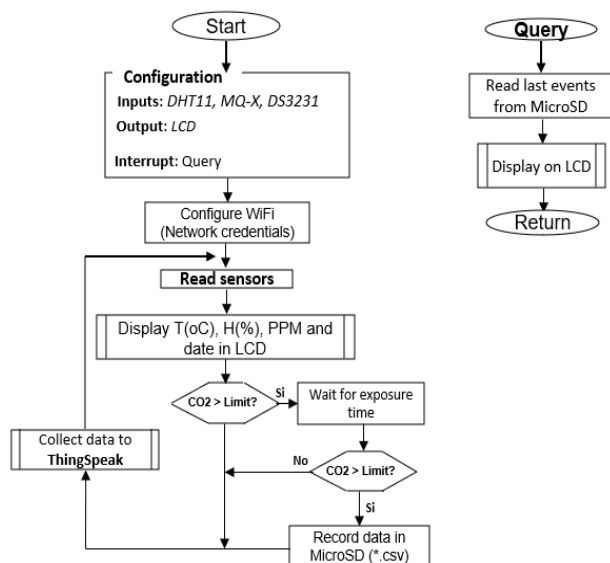


Figure 2. Flowchart and system interrupt routine for indoor air quality monitoring.

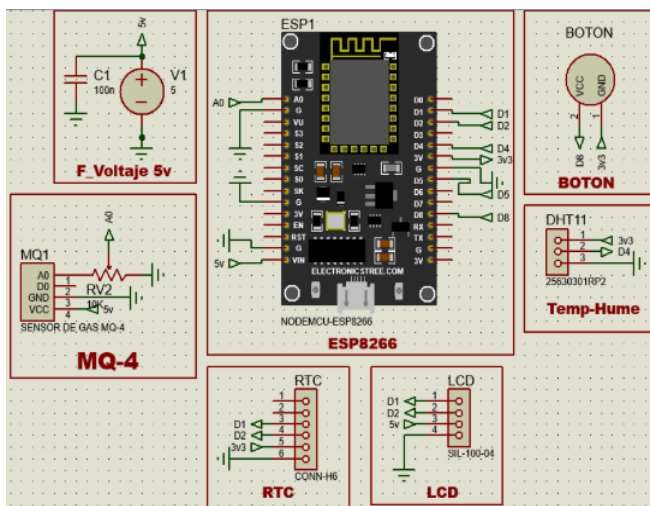


Figure 3. Simulation of monitoring system in ISIS Proteus.

The simulation integrates the DHT11 temperature and humidity sensor, the MQ-4 sensor for CO₂ concentration detection, the DS3231 RTC digital calendar module to label the data with date and time, and a 16x2 LCD with I²C adapter. In addition, there is a button to display on the screen the latest events recorded when CO₂ levels exceed the established limit.

The Schematic Capture and Printed Circuit Board (PCB) Layout functionalities of the ISIS Proteus software were used to design and generate the system's PCB (See Fig. 4).

These tools enable seamless integration between the schematic and the physical layout, facilitating the organization of components and the routing of connections essential for the device's proper operation.

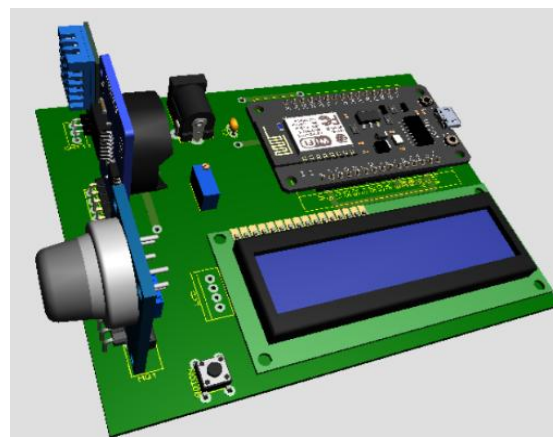


Figure 4. PCB design of the monitoring device in ISIS Proteus.

IV. RESULTS AND DISCUSSION

The developed system enables data retrieval in multiple formats: as numerical values displayed on the LCD screen, as graphical representations via the ThingSpeak IoT platform, and optionally through a smartphone interface using ThingView – ThingSpeak Viewer, available on the Google Play Store for Android devices. Sample data collected by the system are presented in Fig. 5 and Fig. 6, demonstrating its capability to effectively monitor and visualize environmental conditions.

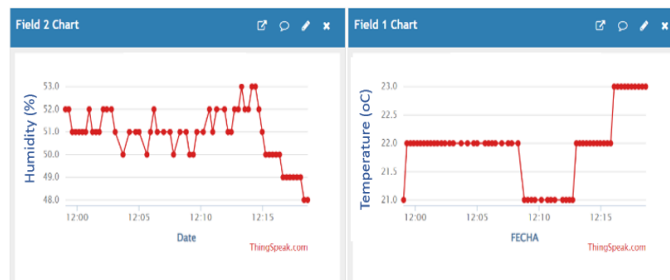


Figure 5. Humidity and temperature were recorded during testing.

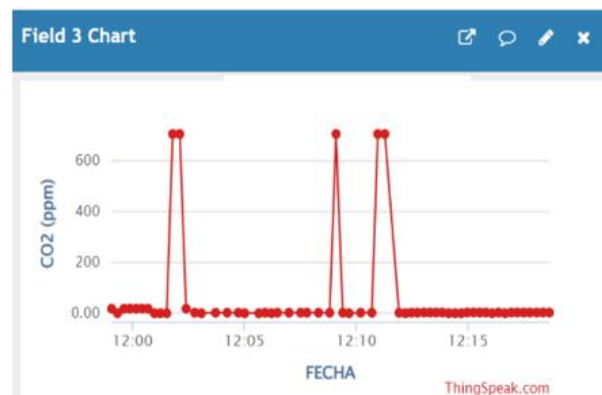


Figure 6. Concentration of Carbon Dioxide (CO₂) obtained in the tests.

To conduct the tests, environmental conditions in the Digital Electronics Laboratory of the School of Electronic Engineering at the Pedagogical and Technological University of Colombia, Tunja campus, were monitored. To modify CO₂ levels in a controlled manner, the device was exposed to a lighter or briquette, which contains gases such as butane (C₄H₁₀), propane (C₃H₈), and methane (CH₄). When burned, these gases generate carbon dioxide (CO₂) and small amounts of Volatile Organic Compounds (VOCs).

For a rapid evaluation of the alert system, a critical exposure time of 10 seconds was configured, differing from the NTC 6199 standard [14], which establishes maximum exposure times of 1 hour and 8 hours based on permissible levels. The exposure time parameter can be adjusted in the algorithm to ensure compliance with regulatory requirements in real-world applications. Fig. 7 illustrates some of the alert events recorded during testing.

Alarm events are logged on a 4GB microSD card in *.txt format, structured for compatibility with CSV files to facilitate further analysis. A sample of the recorded alarm logs is shown in Fig. 7, illustrating the comparison between these records and the data transmitted to ThingSpeak, resulting in an accuracy of 87%. The observed discrepancy in timestamp synchronization is attributed to the 30-second latency inherent to the free version of ThingSpeak.

Date	Temperature (oC)	Humidity (%)	CO2 (ppm)
28/11/2024 9:12	20,6	62	693,3
28/11/2024 9:12	20,6	62	686,2
28/11/2024 9:12	20,6	62	650,9
28/11/2024 9:12	20,6	62	533,1

Figure 7. Alarm log stored in CSV-compliant TXT file.

The system systematically stores alert events on the microSD card, building a database for future applications involving Artificial Intelligence IA-based analysis. Although email or SMS notifications were not implemented, these functionalities are available through ThingSpeak, and their configuration can be referenced in the MathWorks® Help Center.

Finally, the IoT-based air quality monitoring system is a cost-effective solution, with a total development cost of 26.64 USD. Fig. 8 and Table I provide detailed cost breakdowns. The component prices were obtained from the Mouser Electronics website [18].

The proposed system improves upon previous air quality monitoring approaches by integrating real-time IoT capabilities and efficient data storage. Chouhan et al. [3] employed the MQ-4 sensor for CO₂ detection; however, their system lacked cloud connectivity, limiting remote monitoring capabilities. Similarly, Bhattacharya et al. [7] implemented a wireless sensor network for comprehensive indoor air quality assessment, but its complexity and cost reduce its accessibility.

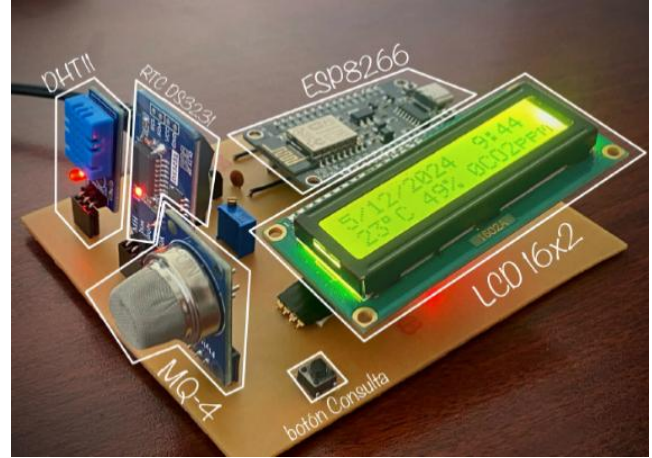


Figure 8. Implementation of the IoT air quality monitoring system.

TABLE I. COST OF AIR QUALITY MONITORING IOT SYSTEM.

Component	Cost
ESP32	8.0 USD
MQ-4	2.17 USD
DHT11	1.83 USD
DS3231	4.35 USD
MicroSD Module 4GB	4.57 USD
Cables and box	5.72 USD
Total	26.64 USD

In contrast, the system presented in this work employs the MQ-135 sensor for multi-gas detection, systematically logs alarm events on a microSD card, and supports future AI-based analysis. The total development cost of the proposed system was 26.64 USD, substantially lower than the USD 122 required for the Klein Tools ET120 [15] gas detector and the USD 485 for the AR8900 Smart Sensor [16], underscoring its economic feasibility compared to commercial alternatives. Although the current implementation focuses on CO₂ monitoring, the system is designed to be scalable, enabling the integration of additional MQ-X series sensors for multi-gas detection and broader environmental monitoring capabilities.

V. CONCLUSION AND FUTURE WORK

Systems for monitoring and controlling environmental variables—such as temperature, humidity, and gas concentrations—are essential in critical environments such as laboratories and offices, where sudden fluctuations can pose health and safety risks. These systems enable early warning mechanisms, enhance occupational health and safety, and ensure compliance with regulations such as [14] and ASHRAE 62-2001 [17], which define critical gas concentration limits in enclosed spaces. By adhering to air quality standards, these systems contribute to maintaining a healthy and safe indoor environment.

The ESP8266 board is well-suited for IoT applications due to its versatility, integrated Wi-Fi™ module, and compatibility with serial communication protocols. Its seamless integration

with the Arduino IDE and extensive library support simplifies its implementation in complex projects, offering efficient and flexible connectivity. Additionally, for more demanding applications, the ESP32 board presents a superior alternative, featuring increased RAM capacity (from 64 KB to 512 KB), higher processing speed (from 80 MHz to 240 MHz), and improved Analog-Digital Converter (ADC) resolution from 10 bits to 12 bits, among other enhancements.

However, a key limitation of the current prototype is its dependence on the free tier of ThingSpeak, which introduces a 30-second latency in data updates. Additionally, the absence of real-time notifications via email or SMS may delay critical alerts in emergencies. Future research could address these limitations by integrating IoT platforms with lower latency, implementing edge computing for local decision-making, and incorporating artificial intelligence algorithms for predictive air quality analysis. These enhancements would significantly improve system performance, responsiveness, and reliability in real-world applications.

REFERENCES

- [1] National Collaborating Centre for Environmental Health. *Carbon dioxide in indoor air*, 2023. [Online]. Available from: <https://nccch.ca/content/carbon-dioxide-indoor-air>. Accessed: 10-Jan-2025.
- [2] Health Canada. *Exposure to indoor air pollutants, 2009*. [Online]. Available from: https://www.hc-sc.gc.ca/ewh-semt/alt_formats/hecs-sesc/pdf/pubs/air/exposure-exposition/exposure-exposition-eng.pdf. Accessed: 10-Jan-2025.
- [3] S. Chouhan, R. Singh, and P. Sharma, "Development of an Air Quality Monitoring System for Urban Environment Using IoT," *International Journal of Computer Applications*, vol. 160, no. 7, pp. 24-30, 2017.
- [4] A. Kadir, M. Alias, D. Dzaki, N. Din, S. Deros and M. Haron, "Cloud-Based IoT Air Quality Monitoring System," *Proc. In 26th IEEE Asia-Pacific Conference on Communications (APCC)*, Kuala Lumpur, Malaysia, 2021, pp. 121-127, doi: 10.1109/APCC49754.2021.9609897.
- [5] P. Srivatsa and A. Pandhare, "Indoor Air Quality: IoT Solution," *International Journal of Research in Advent Technology, Proceedings of the National Conference NCPCL-2016*, pp. 218-220, [Online]. Available from: https://ijrat.org/downloads/Conference_Proceedings/ncpci2016/ncpci-46.pdf. Accessed: 10-Jan-2025.
- [6] F. Salamone, L. Belussi, L. Danza, T. Galanos, M. Ghellere, and I. Meroni, "Design and Development of a Nearable Wireless System to Control Indoor Air Quality and Indoor Lighting Quality," *in Sensors*, vol. 17, no. 5, pp. 1021, 2017, doi: 10.3390/s17051021.
- [7] S. Bhattacharya, S. Sridevi, and R. Pitchiah, "Indoor air quality monitoring using wireless sensor network," *Proc. IEEE 2012 Sixth Int. Conf. Sensing Technol. (ICST)*, Kolkata, India, 18–21 Dec. 2012, pp. 422–427.
- [8] S. Wang, S. Chew, M. Jusoh, A. Khairunissa, K. Leong, and A. Azid, "WSN based indoor air quality monitoring in classrooms," *in AIP Conf. Proc.*, AIP Publishing, Melville, NY, USA, 2017.
- [9] J. Liu et al., "Developed urban air quality monitoring system based on wireless sensor networks," *Proc. of the 2011 Fifth International Conference on Sensing Technology (ICST)*, Palmerston North, New Zealand, 28 Nov.–1 Dec. 2011, IEEE, Piscataway, NJ, USA, 2011, pp. 549–554.
- [10] J. Kang and K.-I. Hwang, "A Comprehensive Real-Time Indoor Air-Quality Level Indicator," *in Sustainability*, vol. 8, no. 881, 2016.
- [11] J. Miller, S. Semple, and S. Turner, "High carbon dioxide concentrations in the classroom: the need for research on the effects of children's exposure to poor indoor air quality at school," *in Occup. Environ. Med.*, vol. 67, no. 11, p. 799, Nov. 2010, doi: 10.1136/oem.2010.057471.
- [12] Ministry of National Education, *Guide for the Design of Healthy Educational Spaces*, 2017. [Online]. Available from: https://www.mineducacion.gov.co/1780/articles-355996_recurso_1.pdf. Accessed: 10-Jan-2025.
- [13] Labcenter Electronics, *Proteus Design Suite: PCB Design & Simulation Made Easy*. [Online]. Available from: <https://www.labcenter.com/>. Accessed: Mar. 30, 2025.
- [14] Colombian Institute of Technical Standards and Certification (ICONTEC), *Colombian Technical Standard NTC 6199: Planning and Design of Environments for Early Childhood Education within the Framework of Comprehensive Care*, 2016. [Online]. Available: https://www.mineducacion.gov.co/1780/articles-355996_recurso_1.pdf. Accessed: Mar. 30, 2025.
- [15] Klein Tools, *Combustible Gas Leak Detector - ET120*, [Online] Available: <https://www.kleintools.com.mx/catalog/detectores-de-gas/detector-de-fuga-de-gas-combustible>. Accessed: Mar. 30, 2025.
- [16] Smart Sensor, *AS8900 Multi-Gas Monitor*, [Online] Available: <https://en.smartsensor.cn/product/38.html>. Accessed: Mar. 30, 2025.
- [17] American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), *Addendum n to ANSI/ASHRAE Standard 62-2001: Ventilation for Acceptable Indoor Air Quality*, 2004. [Online]. Available: <https://www.ashrae.org/File%20Library/Technical%20Resources/Standards%20and%20Guidelines/Standards%20Addenda/62-2001/62-2001-Addendum-n.pdf>. Accessed: Mar. 30, 2025.
- [18] Mouser Electronics, [Online] Available: <https://co.mouser.com/>, Accessed: Mar., 2025.

Deep Learning-based Failure Detection for Safety Diagnostics of Hydrogen Storage Vessels

Dong-Ju Kim, Da-Hyun Kim, Hyo-Jin Kim, and Young-Joo Suh
Institute of Artificial Intelligence, POSTECH, Pohang, South of Korea
e-mail: {kkb0320, kdhyun8011, hyojinkim, yjsuh}@postech.ac.kr

Abstract - Hydrogen is a clean energy source that is essential for responding to climate change and ensuring energy security. Typically, hydrogen storage vessels are exposed to high pressure environments, which can pose an immediate risk of explosion in the event of failure. Therefore, technologies are needed to detect and resolve failures early through diagnostics of hydrogen storage vessels. In this paper, we propose a deep learning-based multimodal failure detection technique to ensure the safety of hydrogen storage vessels. To develop the failure detection technique, we first performed tensile tests on the storage vessel material to collect Acoustic Emission (AE) signals, and also collected failure and normal data based on tensile load graphs. The Synthetic Minority Over-sampling Technique (SMOTE) method was applied to solve the data imbalance. Finally, we developed a multimodal deep learning model using time-domain waveforms and frequency spectra for failure detection, and the proposed method achieved an accuracy of 99.19% and an F1 score of 0.9733, demonstrating excellent failure detection performance. Furthermore, we confirmed that the proposed method shows better performance than using only time-domain waveforms or frequency spectra, and we expect that this research will contribute to the safety diagnosis and maintenance of hydrogen storage vessels.

Keywords - Hydrogen Storage Vessels; Acoustic Emission; Multimodal; Deep-learning.

I. INTRODUCTION

Hydrogen is a clean energy source that emits no greenhouse gases when burned, and is essential for responding to climate change and ensuring energy security. However, hydrogen exists as a gas at room temperature and is highly flammable and bulky, requiring advanced storage and transport technologies. Therefore, high-pressure vessels for compressed hydrogen storage are essential for hydrogen mobility and infrastructure development [1].

In general, hydrogen storage vessels are classified into types 1 to 4 according to their materials and structures, as shown in Figure 1. Type 1 vessels, made entirely of metals such as aluminum or steel, are cost-effective and ideal for transporting hydrogen at low pressures of 200 bar [2]. Type 2 vessels consist of a metal liner reinforced with an outer layer of Glass Fiber Reinforced Plastic (GFRP), allowing a maximum pressure limit of 300 bar. Type 3 vessels enhance this design by using Carbon Fiber Reinforced Plastic (CFRP) as the outer layer, significantly increasing the pressure limit to 700 bar. Unlike Type 2, which does not fully wrap the liner with fibers, Type 3 uses a fully wrapped

structure, providing superior reinforcement. Type 4 vessels, on the other hand, use a resin liner as the inner layer and CFRP as the outer layer, achieving a lightweight design while maintaining the same pressure rating as Type 3.

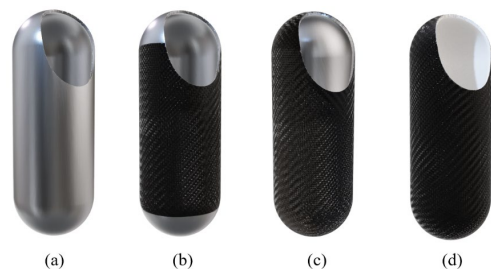


Figure 1. Types of hydrogen gas storages: (a) Type 1, (b) Type 2, (c) Type 3, (d) Type 4.

The metallic composition of Type 1 hydrogen storage vessels makes them susceptible to fatigue, corrosion and cracking, increasing the risk of hydrogen leakage or explosion. Periodic inspection and failure detection are therefore essential for safety. Traditional diagnostics often require disassembly of the vessel, which is not feasible during operation. Failure of a high-pressure vessel poses an immediate risk of explosion, emphasizing the need for in-service diagnostics to detect and resolve failures early, while maintaining reliability and efficiency. Non-Destructive Testing (NDT) technologies are therefore essential. NDT technologies, such as ultrasonic, radiographic and Acoustic Emission Testing (AET) provide real-time in-service safety diagnostics. AET is particularly effective because it analyses acoustic signals generated during failure, making it ideal for high-pressure vessel diagnostics. These methods can ensure safety while preventing catastrophic events such as explosions. However, previous research has focused on specific materials or single failure modes, and in particular on Type 2 and Type 3 vessels.

Therefore, this paper proposes a deep learning based multimodal failure detection technique to ensure the safety of Type 1 hydrogen storage vessels. To do so, we first perform tensile tests on specimens made of aluminum, stainless steel and steel to collect AE signals during failures, and construct a dataset of AE signals using tensile load plots. A multimodal deep learning model using time-domain waveform and frequency spectrum data is developed to improve detection accuracy and reliability. As a result, the

multimodal model achieved an accuracy of 99.19% and an F1 score of 0.9733, demonstrating excellent performance.

II. BACKGROUNDS

A. Failure Modes of Hydrogen Storage Vessels

Failures in hydrogen storage vessels refer to structural deformations caused by external impacts, exceeding allowable pressure limits or material durability issues. The high temperature and high pressure conditions resulting from repeated loading and unloading cycles lead to fatigue-related failures, and Type 1 vessels are particularly susceptible due to their lower allowable pressure limits.

As Type 1 vessels are made entirely of metal, any potential failures are also limited to metal failures. Metals have a single molecule structure and failures occur sequentially depending on the fatigue level of the material. Failures in metals are classified into elasticity, plasticity and fracture. Elasticity occurs when the stress exceeds the yield strength, i.e., a deformation that is reversible when the stress is removed. Plasticity, on the other hand, refers to the permanent deformation that occurs even after the stress is removed [3]. Fracture refers to cracks and ultimate fracture caused by excessive stress. Figure 2 shows examples of the three failure modes in Type 1 hydrogen storage vessels.

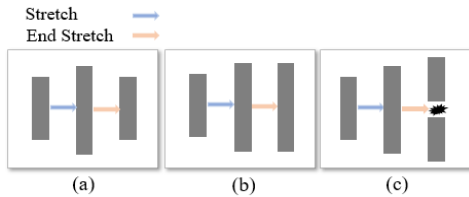


Figure 2. Failure modes of Type 1 vessel: (a) Elasticity, (b) Plasticity, (c) Fracture.

B. AET-Based Non-Destructive Testing

NDT inspects objects without damage, enabling real-time failure detection in hydrogen storage tanks. AET evaluates material failure by analyzing the elastic waves generated during deformation. AET systems consist of AE sensors for signal detection, Data Acquisition (DAQ) systems for digital signal conversion, and analysis for interpretation. Figure 3 shows an example of an AET system.

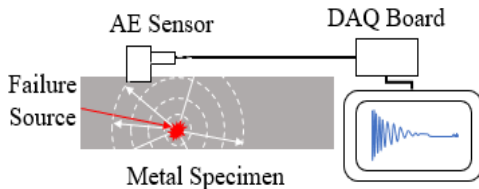


Figure 3. An Example of an AET system.

Accurate AE data acquisition requires appropriate sampling rates. Hits are defined using parameters such as preamplifier gain, threshold and Hit Definition Time (HDT).

Noise is filtered with High Pass Filters (HPF) and Low Pass Filters (LPF), and signal features are extracted in the time domain (e.g. maximum amplitude, rise time) and frequency domain (e.g. peak frequency, average frequency). Figure 4 illustrates the AE waveform and the DAQ parameters used to define hits. To collect accurate data from AE sensors, an appropriate sampling rate must be set. Event occurrences, or hits, are defined using parameters such as preamplifier gain, threshold, Peak Definition Time (PDT), Hit Definition Time (HDT), Maximum Hit Duration (MHD) and Hit Lockout Time (HLT).

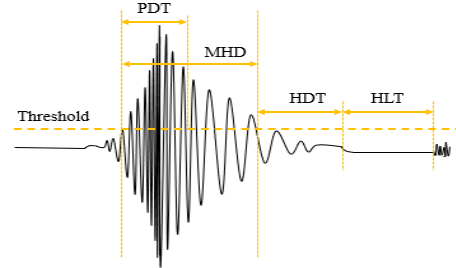


Figure 4. Example of DAQ parameters for defining hits.

C. Tensile Testing

Tensile testing involves pulling customized specimens using a Universal Testing Machine (UTM) to apply stress until failure occurs. Figure 5 shows an example of an AET-based tensile test and a tensile stress graph. When a specimen reaches failure, its properties change, resulting in variations in the applied stress. Generally, failures appear as inflection points on the tensile load graph. By analyzing these inflection points, changes in the specimen properties can be identified.

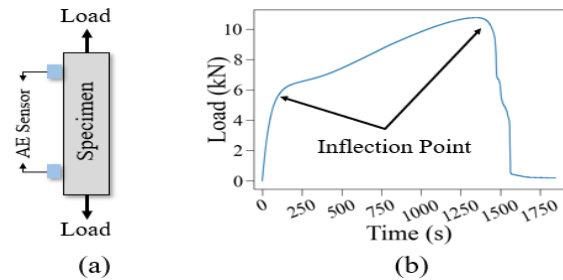


Figure 5. (a) Tensile testing, (b) Tensile-load graph and inflection point.

III. RELATED WORKS

Research on hydrogen storage failure is divided into Finite Element Model (FEM)-based approaches and AE signal analysis methods. FEM simulates stress and fatigue under operational conditions and analyses potential failures experimentally [4]. However, FEM studies primarily focus on correlations between failure phenomena and fatigue levels, limiting their application for real-time detection during operation. In contrast, AET-based research is simpler as it avoids detailed numerical analysis and relies on acoustic signals generated during failures. This allows for real-time

failure detection. However, AET studies have mainly focused on composite materials, such as CFRP in Type 2, 3 and 4 vessels, and there is a lack of research on Type 1 vessels, despite their widespread use and advantages.

Type 1 vessels are made from a variety of metals such as steel, stainless steel and aluminum and require extensive failure analysis. Recent AET and deep learning studies have analyzed failures in metallic vessels, but have been limited to 4130X steel [5]. Therefore, this study addresses this gap by collecting AE signals from steel, stainless steel, and aluminum through tensile testing and constructing an accurate dataset. A multimodal classification model was also developed using time domain waveforms and frequency spectrum data.

IV. DATA PROCESSING

For safety reasons, it is impractical to directly charge and discharge Type 1 hydrogen storage vessels or to apply destructive pressure. Instead, AE failure signals have been obtained by performing tensile tests on specimens of container materials under predetermined parameters.

A. Data Acquisition

In this study, the specimens included stainless steel (SUS304), steel (SS400) and aluminum (AL6106-T6), all of which are widely used in hydrogen storage vessels [6][7]. These specimens were fabricated in accordance with Korean standard KS B 0801 No. 5, and Table I shows example images of each specimen.

TABLE I. EXAMPLE IMAGES OF SPECIMENS

Material	Standard	Example Images
Stainless steel	SUS304	
Steel	SS400	
Aluminum	AL6106-T6	

The specimens were subjected to tensile testing to induce material specific failures. Tensile loads were applied using the Sintech 30/G model (MTS system) and AE signals were recorded using the IDK-AES-H150 resonant sensor at 1 MHz. Failures typically occur below 500 kHz, while signals below 10 kHz are often noise or equipment vibration, so a digital filter was applied to remove noise. Hit detection parameters were set to accurately capture peak values and event intervals. Table II summarizes the DAQ settings used in the experiment.

Tensile tests were performed on three specimen types to collect AE waveform data. Specimens were loaded to failure and only event waveform data was collected based on the sensor settings. Normal data was also collected by attaching AE sensors to Type 1 vessels operated within allowable pressure limits. Figure 6 shows the test environment and failed specimens, while Table III lists the number of samples collected.

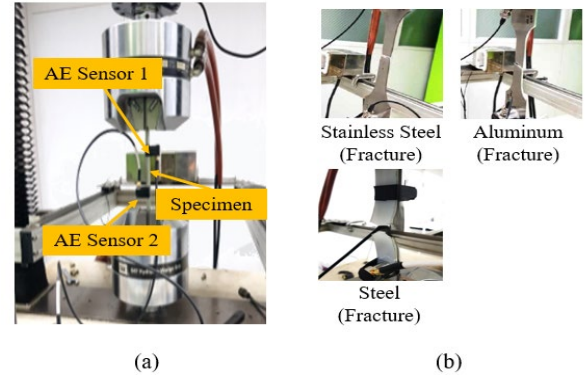


Figure 6. (a) Tensile testing environment, (b) Specimens after testing.

TABLE II. DAQ PARAMETERS CONFIGURED TO DEFINE HITS

Parameter Type	Parameters	Value	Unit
Sensor	Sampling Rate	1	MHz
	Pre-Amp Gain	40	dB _{ac}
	Threshold	30	dB
Digital Filter	High Pass Filter (HPF)	10	kHz
	Low Pass Filter (LPF)	500	kHz
Hit Detection	Peak Definition Time (PDT)	200	μs
	Hit Definition Time (HDT)	400	μs
	Maximum Hit Duration (MHD)	1	ms
	Hit Lock-out Time (HLT)	10	ms

TABLE III. NUMBER OF COLLECTED DATA

Specimen	Number of data
Stainless steel	333
Aluminum	2,056
Steel	44,792
Type 1 Storage (Normal)	69,243

B. Data Labeling

Metal failure occurs when the tensile load exceeds the yield strength, changing atomic arrangements and material properties. These changes vary with the rate of load increase, allowing failure regions to be identified using a time dependent load curve. The elastic region occurs when stresses remain below the yield point, causing minimal deformation and a continuous increase in load. Plasticity begins when the stress exceeds the yield point, resulting in significant deformation and slower load increase. An inflection point marks the transition from elasticity to plasticity. Fracture occurs when the material can no longer support the stress, causing the load to drop rapidly to zero,

creating another inflection point. These points divide failure regions and matching their times to the event waveforms allows failure labelling. Figure 7 illustrates tensile load graphs and failure region subdivisions, summarized in Table IV.

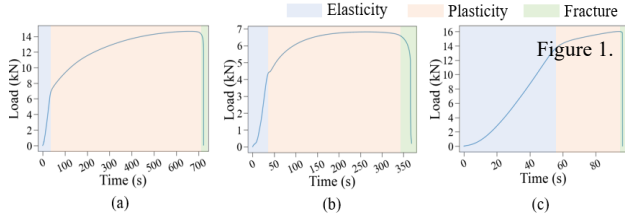


Figure 7. Tensile-load graph and failure region of each material specimens: (a) Stainless steel, (b) Steel, (c) Aluminum.

C. Data Preprocessing

The dataset was pre-processed for deep learning training, incorporating frequency domain information to improve performance.

1) Unify Waveform Lengths and Min-Max Scaling

All waveform lengths were unified to 1024 samples. AE waveforms vary in length depending on event duration, but consistent input sizes are required for deep learning. This study determined the optimal length to minimize information loss and computational load. Waveform lengths were statistically analyzed and outlier information segments were identified to set the unified length to 1024, as shown in the histogram in Figure 8.

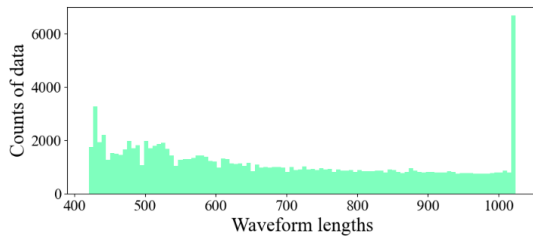


Figure 8. Histogram of waveform lengths.

Second, the amplitude of the waveform was scaled to [-1, 1] using min-max scaling. Without scaling, data values could

vary widely, causing instability and inefficiency during the weight update process. Scaling creates a uniform distribution, reducing variability and stabilizing training. Since waveform data includes negative values, the scaling range was set to [-1, 1], as expressed below:

$$x' = 2 \left(\frac{x_i - \min(x)}{\max(x) - \min(x)} \right) - 1 \quad (1)$$

where $\max(x)$ and $\min(x)$ is:

$$\max(x) = \begin{cases} |\max(x)|, & |\max(x)| \geq |\min(x)| \\ |\min(x)|, & |\max(x)| < |\min(x)| \end{cases} \quad (2)$$

$$\min(x) = \begin{cases} -|\max(x)|, & |\max(x)| \geq |\min(x)| \\ -|\min(x)|, & |\max(x)| < |\min(x)| \end{cases} \quad (3)$$

For x , the baseline is 0, but x' maps the maximum and minimum of x to 1 and -1 respectively. If their absolute values differ, the centers of the waveforms can vary, increasing the variance of the data. To overcome this, the larger absolute value is mapped to 1 and the smaller to -1, centering the waveform at 0. Figure 9 shows the original waveforms after min-max scaling and alignment to 1024 length, comparing standard scaling and the adapted method.

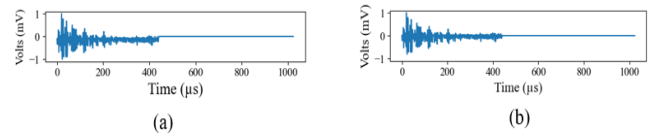


Figure 9. Plot a waveform after preprocessing: (a) General min-max, scaling (b) Ours.

2) Construction of Frequency Spectrum Dataset

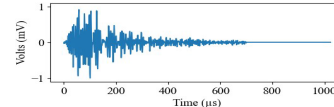
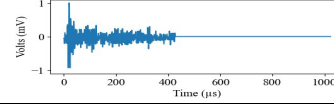
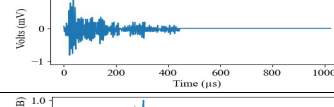
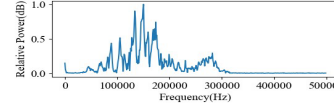
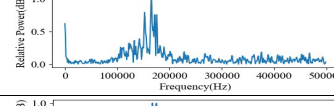
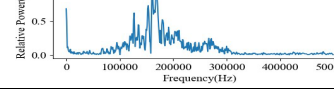
Understanding both intrinsic and frequency characteristics is essential in waveform analysis. Frequency characteristics minimize the effect of sensor type and placement, facilitating generalized classification methods. To incorporate this, the frequency domain data was constructed using the Fourier Transform (FT). As the signals were digital and discrete, the Discrete Fourier Transform (DFT) was applied using the Fast Fourier Transform (FFT) algorithm for computational efficiency.

TABLE IV. TIME AND NUMBER OF ACQUIRED WAVEFORMS ACCORDING TO FAILURE REGIONS IN EACH SPECIMEN

Specimen	Time by Region (Seconds)				Number of Waveforms		
	Total Experiment	Elasticity	Plasticity	Fracture	Elasticity	Plasticity	Fracture
Aluminum	100.54	0 ~ 57.21	57.2 ~ 100.4	100.4 ~ 100.5	1,583	467	6
Steel	379.85	0 ~ 37.47	37.4 ~ 365.9	366.7 ~ 379.8	4,093	40,621	78
Stainless steel	728.94	0 ~ 33.94	34.6 ~ 706.7	706.5 ~ 728.9	184	107	42
Total					5,860	41,195	126

The FFT transformation produced discrete frequency spectrum graphs as line plots of the frequency distribution. Unlike the waveform data, the frequency spectrum lacks negative values and was scaled to $[0, 1]$. Each FFT result contained 1024 samples, corresponding to the length of the waveform. Due to y-axis symmetry, only positive frequencies were retained, reducing the length to 512. Table V shows examples of the original waveform, the pre-processed waveform and the scaled spectrum data after FFT transformation.

TABLE V. EXAMPLE OF PREPROCESSED WAVEFORM AND FREQUENCY SPECTRUM

Data Type	Specimen	Example
Waveform	Stainless steel	
	Aluminum	
	Steel	
Frequency Spectrum	Stainless steel	
	Aluminum	
	Steel	

V. CLASSIFICATION MODEL FOR FAILURE DETECTION

A deep learning model was developed using the constructed dataset to classify type 1 storage failures. Fracture, which indicates material rupture and explosion, was excluded as it is irrelevant to safety diagnostics. Only elasticity, plasticity and normal data were used to distinguish these states. The data set was divided into training, validation and test sets in a ratio of 60:20:20. However, there is a significant imbalance between elasticity and plasticity data. To address this, the SMOTE method was used to balance the training data [8]. Table VI shows the number of training, validation and test data augmented by SMOTE.

A binary failure and normal classification model was trained and tested using the collected dataset. The model, designed as a one-dimensional convolutional neural

network (1D-CNN), extracted features from both waveforms and frequency spectra for classification. To better capture temporal characteristics, an extended causal 1D-CNN architecture was used. Figure 10 shows the structure of the diluted causal 1D CNN.

TABLE VI. NUMBER OF TRAIN/VALID/TEST DATASET

Failure Mode	Train		Validation	Test
	Before Augmentation	After Augmentation		
Elasticity	3,516	41,545	1,172	1,172
Plasticity	24,717	41,545	8,239	8,239
Normal	41,545	41,545	13,849	13,849
Total	69,778	124,635	23,260	23,260

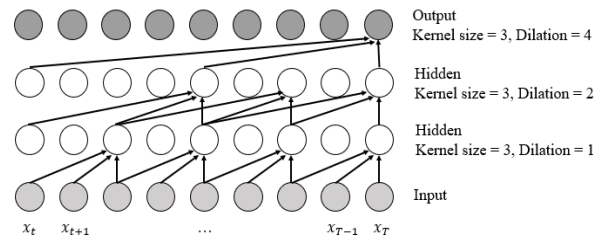


Figure 10. Dilated Causal 1D-CNN

Convolution, batch normalization and max-pooling (size: 2) layers were used to process the input data to extract features, which were then transformed into a 1-dimensional vector with 3 outputs using global average pooling, bypassing the need for a fully connected layer. The outputs were passed through the softmax activation function to compute the final probabilities for each failure type. The Nadam optimizer was used to train the model. This study evaluated the classification performance under three input scenarios: waveform data only, frequency spectrum data only, and combined features of both. In addition, the ResNet-50 architecture was used as the feature extractor to evaluate the performance improvements from a deeper network design.

Table VII shows the experimental results for models using different inputs and architectures. The result shows that the 12-layer multimodal model achieved the best performance with 99.19% and an F1 score of 0.9733. The superior results of the flatter model compared to ResNet suggest that a more complex architecture is not necessary for this classification task. Furthermore, the higher performance of the multimodal model compared to single input models (waveform or frequency spectrum) indicates that the two types of data are complementary for this classification problem. Figure 11 shows the structure of the best performing model, i.e. the 12-layer 1D CNN multimodal model.

TABLE VII. CLASSIFICATION PERFORMANCE

<i>Input Data Type</i>	<i>Model Structure</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Waveform Only	12-Layer 1D-CNN	98.87%	0.9603	0.9622	0.9613
	ResNet-50 1D-CNN	98.83%	0.9602	0.9655	0.9628
Frequency Spectrum Only	12-Layer 1D-CNN	98.94%	0.9599	0.9712	0.9654
	ResNet-50 1D-CNN	98.93%	0.9604	0.9624	0.9614
Multimodal	12-Layer 1D-CNN	99.19%	0.9723	0.9743	0.9733
	ResNet 1D-CNN	98.89%	0.9590	0.9686	0.9637

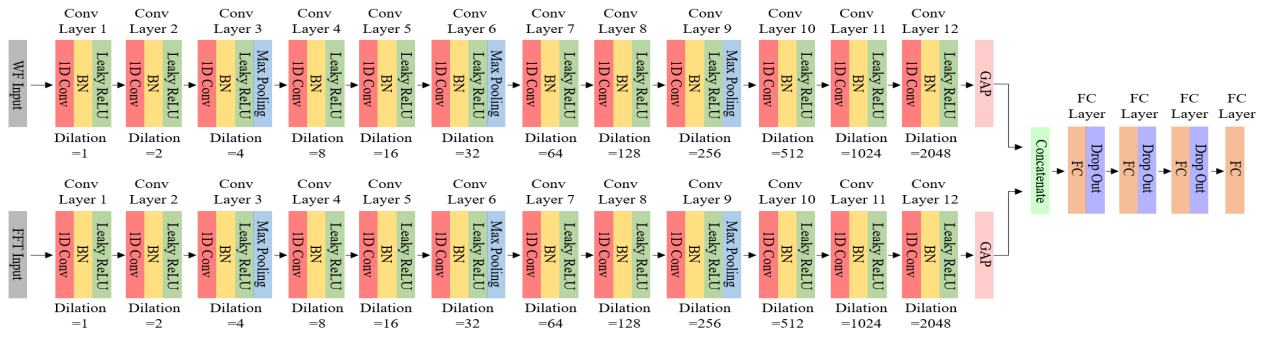


Figure 11. Proposed 1D-CNN multimodal model.

VI. CONCLUSION

This paper proposes a deep learning-based multimodal fault detection technique to ensure the safety of hydrogen storage vessels. To this end, we first collected AE signals from Type 1 hydrogen storage vessels by tensile testing and constructed data sets for elastic, plastic and normal regions. We then developed a multimodal deep learning fault detection model using waveform and frequency spectral data. From the experimental results, we confirmed that the proposed multimodal model achieved an accuracy of 99.19% and an F1 score of 0.9733, demonstrating excellent failure detection performance. In the future, the proposed method is expected to enable real-time fault detection of Type 1 vessels, contributing to efficient and reliable safety diagnostics.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No.2022R1A6A1A03052954)

REFERENCES

- [1] D. Cecere, E. Giacomazzi, and A. Ingenito, "A review on hydrogen industrial aerospace applications," *Int. J. Hydrogen Energy*, vol. 39, no. 20, pp. 10731–10747, Jul. 2014.
- [2] H. Barthélémy, "Hydrogen storage – Industrial perspectives," *Int. J. Hydrogen Energy*, vol. 37, no. 22, pp. 17364–17372, 2012.
- [3] R. Hill, *The Mathematical Theory of Plasticity*. Oxford, U.K.: Oxford University Press, 1950.
- [4] A. El-Abbassi, A. Mekki, and A. Medles, "A comparative study of hydrogen storage in solid-state materials," *Int. J. Hydrogen Energy*, vol. 45, no. 9, pp. 5367–5375, 2020.
- [5] R. M. Gaikwad, A. Kulkarni, and S. V. Joshi, "Experimental investigation of hydrogen storage in composite vessels," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 810, no. 1, p. 012016, 2020.
- [6] S. Kumar and S. K. Dwivedi, "Composite Pressure Vessels: Design, Analysis, and Manufacturing," *J. Pressure Vessel Technol.*, vol. 140, no. 6, 2018.
- [7] O. Comond, D. Perreux, F. Thiebaud, and M. Weber, "Methodology to Improve the Lifetime of Type III HP Tank with a Steel Liner," *Int. J. Hydrogen Energy*, vol. 34, no. 2, pp. 576–586, 2009.
- [8] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

Disaster Detection Framework for Smart Cities: An AI YOLOv8 and IoT Approach

Hossam Kamel

Swiss School of Business and Management Geneva

SSBM Geneva

Geneva, Switzerland

email: dr.hossamkamel.a@gmail.com

Abstract— Disaster detection is vital for smart city resilience and public safety. This paper presents a framework for detecting fire and flood incidents using the You Only Look Once version 8 (YOLOv8) algorithm on a Raspberry Pi Internet of Things (IoT) device, which transmits data to IoT operation center. An initial experiment using a laptop and mobile phone demonstrated the effectiveness of machine learning in fire detection.

Keywords—Smart Cities; Disaster Detection; Fire Detection; Flood Detection; IoT; YOLOv8; Artificial Intelligence (AI); Machine Learning (ML); Raspberry Pi.

I. INTRODUCTION

The rapid expansion of urban populations has placed increasing pressure on city infrastructure, requiring innovative solutions to enhance resilience and disaster preparedness. Smart cities leverage technology to mitigate risks posed by natural and man-made disasters, integrating AI, IoT devices, and real-time data processing to improve urban safety. The goal of smart cities is to leverage technology and data analytics to improve the quality of life in urban areas [1]. There are several ongoing research efforts focused on using technology to monitor and manage in-city disasters, either at the macro level or for specific types of disasters [2].

This paper presents a comprehensive disaster detection framework that integrates IoT-based environmental sensing, AI-driven image processing, and real-time data transmission to an IoT operations center. The framework is designed to detect and respond to disasters such as fires and floods using a combination of Closed-Circuit Television (CCTV) surveillance, edge AI processing on Raspberry Pi devices, and automated alerts to emergency responders.

The core contribution of this work is the implementation and validation of an AI-based disaster detection system within the proposed smart city framework. Specifically, we evaluate the effectiveness of YOLOv8, a state-of-the-art object detection algorithm, in identifying fire hazards using real-time image analysis. Additionally, we discuss the potential for extending the system to flood detection.

The remainder of the document is organised as follows: Section II discusses the challenges in traditional disaster detection systems, emphasizing the need for real-time AI-based solutions. Section III introduces the proposed disaster detection framework, outlining its key components and role in smart city resilience. Section IV details the system implementation, explaining how YOLOv8 and IoT

components work together for fire and flood detection. Section V presents experimental results, highlighting the effectiveness of the proposed system in identifying fire incidents. Finally, Section VI discusses the broader implications, including potential improvements for flood detection and integration into smart city infrastructure.

II. CHALLENGES IN TRADITIONAL DISASTER DETECTION SYSTEMS

Conventional disaster detection systems deploy various types of sensing devices, which can be categorized into different groups. Static sensing devices are permanently located at a specific geographical site, cumulating data over time. Examples include seismometers and weather sensors. On the other hand, mobile sensing devices are portable and can be strategically deployed at various locations or moved over time. Such devices include smart phones and Unmanned Aerial Vehicles UAVs [3].

Despite the technological advancements in disaster sensing and detection devices, the currently deployed disaster detection systems face several challenges in effectively detecting disasters [4]. These challenges can be categorized into three groups: technological challenges, operational challenges, and situational challenges [5].

The first technological challenge is slow disaster response time. The disaster response time consists of three time delays: sensing time, processing time, and communication time. The traditional disaster detection systems take considerable time to detect disasters due to their reliance on fixed and static sensors. These sensors must first sense the arrival of a disaster. Then, the cumulative data are sent to a centralized location for processing, which results in significant communication delays. The second technological challenge is a lack of integration. Currently, the disaster detection systems in developed countries do not integrate with one another, even though multiple disasters are detected by using different devices on numerous occasions. As a result, significant time delays occur in detecting and preventing disasters. It is vital that systems be developed that can share disaster information in real time. The third technological challenge is a lack of high-resolution data. Currently, most of the data used in the detection systems are remotely gathered data, which restrict the detection systems from generating high-resolution data [6]. This limitation results in difficulties in precisely locating the site of a disaster.

The first operational challenge is the lack of trained human resources [7]. Most of the developing countries'

governments have limited resources to employ. The second operational challenge is maintenance costs [8]. The structural deterioration of the sensing devices is a key reason for inoperable disaster detection systems. The third operational challenge is community engagement [9]. A lack of community awareness and participation can lead to ignored emergency alerts, delaying evacuation and reducing disaster response effectiveness. While many disaster detection systems have been successfully implemented in developing countries, the local community remains largely uninterested. As a result, several systems are rendered useless because the community does not provide adequate resources for operating the systems.

III. KEY COMPONENTS OF SMART CITIES FRAMEWORKS

In recent years, numerous cities have adopted smart city frameworks, which outline principles, policies, and goals for smart city development [10]. The foundation for the frameworks is a definition of the essential components of smart city systems, describing what needs to be integrated and how this integration works [11].

The comprehensive framework for establishing advanced smart city systems is fundamentally centered around six key and crucial Smart City Pillars: Smart Governance & Education, Smart Living, Smart Healthcare, Smart Transportation, Smart Economy, and Smart Environment [12]. These vital pillars serve as the essential focus areas, meticulously aimed at significantly enhancing urban life, fostering community engagement, and ensuring a sustainable future for all residents. Each of these pillars plays a unique and impactful role in fostering innovation while improving overall quality of life within metropolitan areas. By integrating these pillars effectively, cities can promote technological advancement and create a conducive environ-

ment for growth and development. Furthermore, the interconnectedness of these pillars increases the potential for synergies, enabling cities to tackle complex urban challenges more efficiently and sustainably. The holistic approach of the framework ensures that every aspect of city living is considered, providing a comprehensive strategy for modern urban management and planning, thereby encouraging a well-rounded development that benefits everyone.

The proposed framework consists of three core layers, each playing a distinct role in disaster detection:

1. **IoT Sensor and Camera Layer** – This includes Raspberry Pi-based edge AI devices, CCTV cameras, and environmental sensors deployed across the city. These devices capture real-time visual and environmental data.

2. **AI Processing and Detection Layer** – The captured data is analyzed using YOLOv8 running on Raspberry Pi to detect fire or flood incidents. This edge computing approach ensures faster detection and reduces reliance on cloud processing.

3. **IoT Operations Center and Response Layer** – Detected events are transmitted via wireless or wired networks to a central operations center, where emergency services are notified. Alerts can also be sent to residents via mobile applications or warning systems.

By structuring the solution within this framework, we ensure that the proposed system is scalable, adaptable, and aligned with existing smart city initiatives.

IV. SMART CITY FRAMEWORK

A plethora of smart city frameworks can be unearthed through a wide-ranging investigation of publicly accessible smart city features, models, methodologies, scaffolds, architectures, and pilot schemes [13].

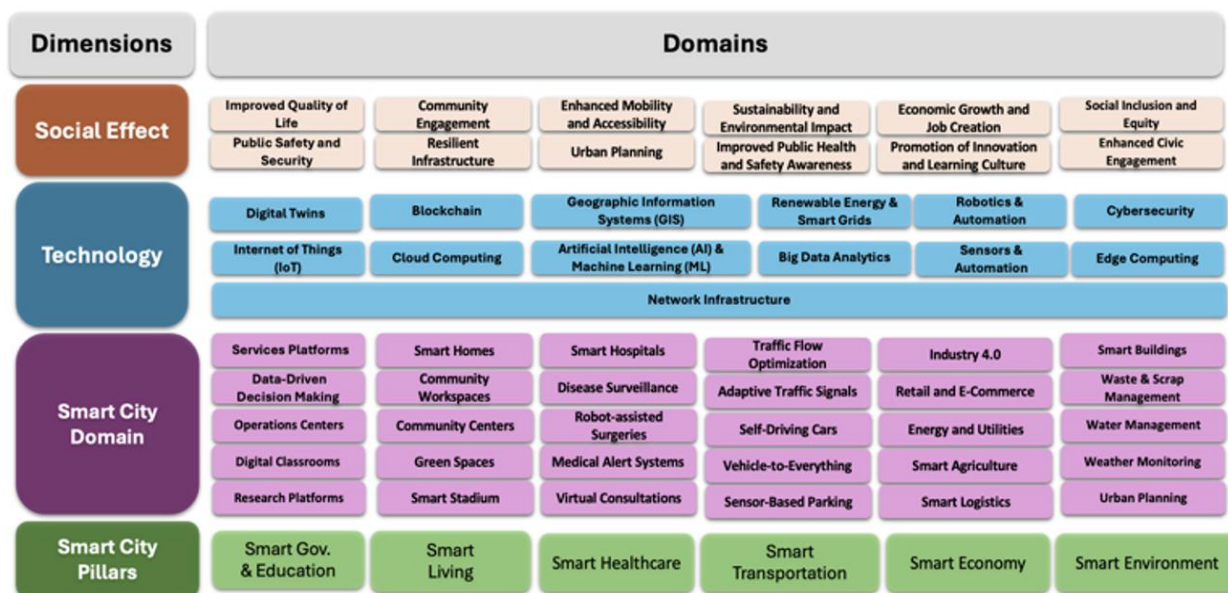


Figure 1. Smart City Framework [14].

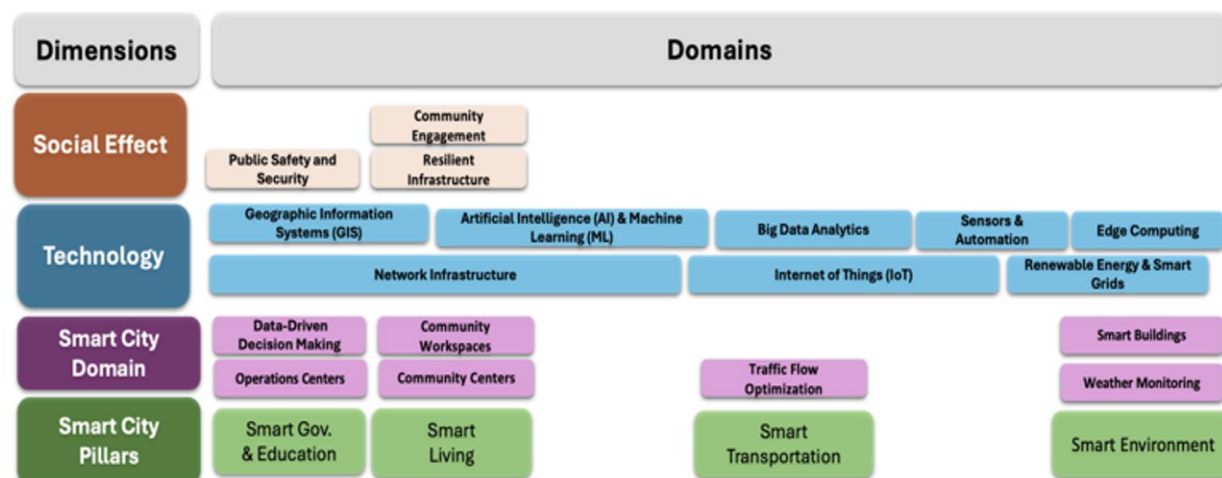


Figure 2. Proposed Disaster Detection Smart City Framework.

The smart city framework in Figure 1 is an adaptive model that integrates key components of smart city development, combining core pillars, specialized domains, advanced technologies, and social impacts. When focusing on a specific activity—such as waste management or smart healthcare—only relevant elements are retained, streamlining efforts and emphasizing expected social benefits like public health, sustainability, and economic growth.

Structured around six central pillars, the framework aligns targeted domains such as smart homes, traffic optimization, and waste management with data-driven decisions and robust infrastructure. It highlights cutting-edge technologies, including IoT, AI, blockchain, and renewable energy, ensuring seamless operation through a strong network infrastructure.

By linking domains and technologies to tangible outcomes, the framework offers a focused, efficient, and socially impactful roadmap for smart city initiatives. It is important to address data security, particularly in data transmission. Ensuring the authenticity and integrity of transmitted data can prevent cyber threats, misinformation, and unauthorized access. Implementing encryption, secure communication protocols, and blockchain-based verification could strengthen the system against tampering or data manipulation, enhancing trust and reliability in disaster response operations.

V. PROPOSED DISASTER DETECTION SMART CITY FRAMEWORK

A Disaster Detection Framework (DDF) for smart cities, as shown in Figure 2, integrates advanced technologies and smart city components to enhance fire prevention, detection, and response while promoting sustainability. IoT sensors, Geographic Information System (GIS), and weather monitoring systems are deployed to track fire risks and detect incidents in real time, while AI and edge computing analyze data for early detection and predictive modeling. Operations centers coordinate responses, leveraging adaptive traffic systems for evacuation routes and notifying

communities through automated alerts. Smart buildings and resilient infrastructure are equipped with automated safety measures, and waste management systems handle post-fire debris sustainably.

Community engagement and public safety initiatives educate citizens on fire prevention, while renewable energy sources power detection systems, ensuring sustainability. This holistic approach combines technology, smart city domains, and proactive strategies to minimize fire risks and enhance safety in urban environments.

VI. AI & ML ROLE IN DDF

The advancements of AI and ML have made urban centers smarter and more self-sufficient [15]. However, the sustainable development of smart cities is still an ongoing challenge, especially in disaster-prone areas. These technologies can monitor and identify a disaster, as well as estimate the resources needed to handle it.

AI and ML offer data analysis tools that can enhance performance beyond traditional methods, fostering innovation in diverse fields. Natural and man-made disasters significantly affect societal development, underscoring the importance of early detection to reduce risks, economic losses, and casualties [16].

Despite great interest in using artificial intelligence and machine learning for disaster detection, several obstacles must be overcome to ensure successful implementation. Primarily, the quality and availability of data pose significant challenges. While many cities provide access to public data, such datasets are often not collected with the same parameters or standards, limiting their comparability. Moreover, the implementation of smart city technologies is frequently hindered by privacy concerns, particularly regarding the use of personal data.

In modern smart cities, disaster detection systems are crucial for safety. Urbanization has increased flood and fire vulnerabilities. The object detection technology has evolved significantly in recent years, driven by its successful applications in various domains. The development of these detectors follows a “model zoo” approach, where different

models trained using varying methodologies are made publicly available [17].

The YOLO series is one of the most well-known object detector families. In particular, YOLO version 8 (YOLOv8) is a complete object detection and instance segmentation model that overtakes its predecessors [18]. It uses the framework for model implementation, training, evaluation, and inference. The YOLO family of models has played a pioneering role in advancing real-time object detection, owing to their unique architecture that integrates model training and inference on a single neural network. YOLOv8 excels as a real-time object detection model, quickly identifying and classifying objects within diverse classes in images and videos. It utilizes a single convolutional neural network to simultaneously predict bounding boxes, class probabilities, and object counts for detected classes [19]. YOLOv8 architecture consists of five key stages: image preprocessing and augmentation, backbone, neck, detector, and postprocessing [20].

VII. FIRE AND FLOOD DETECTION IN SMART CITIES USING YOLOv8

Fire is one of the disasters that poses a significant threat to human life in urbanized areas. This is compounded by other potential hindrances based on the infrastructure. Fire detection in cities is particularly difficult as they are typically crowded spaces, leading to obstructions in the view field of the cameras. Furthermore, flames in general spread rapidly, leading to the idea of having an early detection mechanism [21].

Based on the above, integrating the algorithm with preexisting CCTV cameras on the roads would be an efficient alternative. Currently, most detection systems rely on either thermal cameras or a combination of both thermal and visual cameras. This necessitates the need for a separate camera system installed in addition to the standard CCTV cameras on roadways. Consequently, a new detection framework that makes use of road surveillance CCTV cameras for fire detection is presented. Recent advancements in the YOLO family, namely YOLOv8, are utilized to train a model that can detect fires. There are various approaches to integrating this model, either with an already preexisting detection system based on image processing techniques or outside detection systems based on just monitoring the images.

The urban environment poses an additional challenge for fire detection since fires are anticipated to be detected at a greater distance as opposed to other environments like industrial complexes. Hence, it is critical to have cameras that can cover a wider area. Empirical results with real-world implementation to monitor and detect fire in the surroundings of a highway are provided. Detection systems of this nature are necessary, particularly in high-speed roadways, as the response time for vehicles approaching an accident is crucial. Since the detection system is based on image processing techniques, the data can be processed in real-time on the edge to ensure rapid detection and a timely response. A thorough discussion of the framework is provided, along with case studies and examples where

YOLOv8 has been useful in detecting fires [22]. This aims to provide effective implementations of such technologies and inspire the future and current endeavors in this field.

Flooding is one of the most serious calamities in urban settings that arise due to sudden and massive downpour events with gradual drainage of the platform. A flood is a complex catastrophe that involves many crucial and complicated occurrences, which happen concurrently.

There are many cataclysmic events that create flooding in a city, such as storms, tsunamis, dam breaking, heavy rain, melting snow, landslides, etc. Though there are many elements accountable for a flood catastrophe, urbanization is found to be the most evident one [23].

The real-time image processing-based flood detection is designed to detect flooding swiftly using the video stream taken from the camera placed on the roadside. The framework for flood detection in smart cities is implemented using YOLOv8. Flood can be detected using infrared, visible, or depth images taken from the camera installed in public places like traffic signals, malls, parking areas, etc. The video stream from the camera is processed using YOLOv8 to detect the flood situation. When flood is detected, it generates an alert which can be sent to the control room or concerned authority. The framework can be integrated with other sensors like rainfall, water depth, temperature, humidity etc. to take precautionary measures. The framework can also be used with the GIS system to view the flood affected area on the city map. The flood detection using YOLOv8 is tested with various videos taken from real urban flood scenarios during heavy downpour. The framework is successfully able to detect the flood condition.

Fire detection can benefit from CCTV cameras, even if they are not infrared, due to their widespread installation in urban areas, reducing the need for additional infrastructure. Leveraging existing CCTV networks allows for cost-effective fire monitoring, real-time surveillance, and integration with AI-based detection systems like YOLOv8, enabling early detection and response without requiring specialized thermal imaging cameras.

VIII. IOT ROLE IN THE DETECTION

Raspberry, as shown in Figure 3, can serve as a powerful IoT edge device capable of running advanced AI models such as YOLOv8 to detect fire and flood incidents in real time. Equipped with camera modules and environmental sensors, the Raspberry Pi can analyze visual and sensor data locally, leveraging its processing power to identify potential hazards with high accuracy. Once a threat is detected, the device can connect to the country's Internet



Figure 3. Raspberry with Camera.

network via Wi-Fi, Ethernet, or cellular modules to transmit critical data, including alerts and images, to a centralized IoT operation center. This seamless integration enables authorities to respond swiftly to emergencies, enhancing disaster management efforts with a cost-effective and scalable solution.

Local processing with YOLO could enhance disaster detection by reducing latency and reliance on network connectivity, enabling faster responses. However, back-end processing allows for centralized analysis, resource optimization, and integration with larger datasets, making it more scalable. A hybrid approach, combining local edge processing for real-time detection with back-end verification, could improve efficiency and reliability.

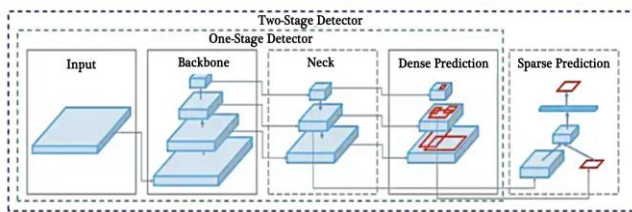


Figure 4. YOLO Model.

Figure 4 represents the architecture of object detection models. The one-stage detector processes input through a backbone for feature extraction, a neck to refine features, and a dense prediction layer to detect objects directly. The two-stage detector refines detection by using a sparse prediction layer, improving accuracy by first generating region proposals before classification.

IX. THE PROPOSED MODEL ALGORITHM

Figure 5 presents the solutions architecture for an IoT-enabled fire detection and response system, integrating UAVs and cameras. The system utilizes the YOLOv8 algorithm for real-time fire detection by analyzing visual data from UAVs and surveillance cameras. An IoT network facilitates communication, with the IoT operation center managing data processing and response coordination. Continuous model training and feedback loops enhance detection accuracy and system performance. The network infrastructure ensures reliable connectivity, enabling rapid UAV deployment for fire suppression. This architecture demonstrates the integration of AI, IoT, and UAV technologies to improve fire safety in smart city environments.

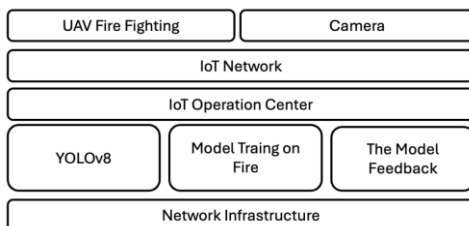


Figure 5. The Proposed Solutions Architecture.

Although YOLOv8 is used in different works, the authors only focus on one detection, either fire or flood [23] [24] [25]. Detecting fire and flood events in real time is a critical task for mitigating potential disasters and protecting people and property. This algorithm is proposed as a solution to the problem of automated fire and flood detection using the YOLO deep learning framework. By leveraging a single multi-class model, we can efficiently identify both threats within the same scene, simplifying the deployment process and reducing the computational overhead.

High-quality data is the foundation of any successful detection model. We begin by gathering a wide range of images showing fire and flood scenarios under various conditions (different lighting, angles, scales). We include some negative examples (images without fire or flood) to help the model learn what backgrounds typically look like. We use a labeling tool—such as Labellmg, Roboflow, or CVAT—to draw bounding boxes around the areas containing fire or flood. Each bounding box should be labeled with the appropriate class name: fire or flood. Once labeled, we split the data into training and validation sets, maintaining a similar distribution of classes in both sets. The directory structure typically follows YOLO's expected format, and we will need a data.yaml file that specifies paths to images, the number of classes, and their names.

By developing and training a unified multi-class YOLO model with meticulously labeled datasets encompassing both fire and flood scenarios, the proposed algorithm effectively facilitates simultaneous real-time detection of these two critical hazards. The process begins with the collection and annotation of diverse images representing fire, flood, and non-threatening environments, which are subsequently organized into training and validation subsets adhering to YOLO's standardized format. Utilizing the Ultralytics YOLO framework, the model undergoes extensive training to learn distinguishing features of each class, resulting in a robust best.pt weight file capable of identifying both fire and flood instances with high accuracy.

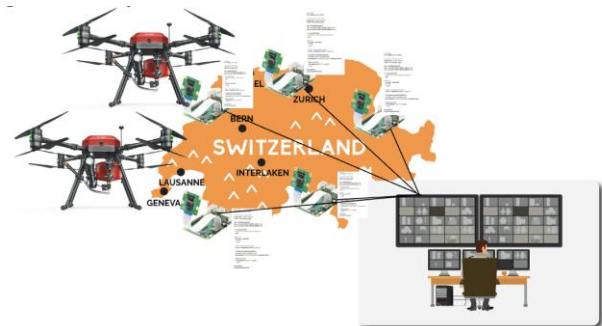


Figure 6. Overview of The Proposed Solutions.

The main Python script integrates OpenCV to capture live video streams, wherein each frame is processed by the trained YOLO model to perform detections based on

predefined confidence and Intersection over Union (IoU) thresholds. Detected objects are annotated with bounding boxes and class labels directly on the video feed, enabling immediate visualization and potential activation of alert mechanisms. This streamlined approach not only enhances computational efficiency by employing a single model for dual-class detection but also ensures prompt and reliable identification of fire and flood events, thereby contributing significantly to automated disaster monitoring and mitigation systems.

Figure 6 presents the proposed solution, which integrates a Raspberry Pi running the YOLOv8 algorithm to detect fire and flood, transmitting real-time data to an IoT operation center. Additionally, UAVs are deployed to respond rapidly and assist in firefighting efforts. It also highlights the role of geographic distribution in detecting and tracking spreading disasters. By using IoT sensors and AI, the system enables real-time monitoring across urban areas, improving early detection and response efficiency to ensure that the framework is not just theoretical, but directly supports the proposed system, Table I clarifies how each framework component integrates with the disaster detection model.

TABLE I. PROPOSED FRAMEWORK COMPONENTS AND THEIR ROLES IN DISASTER DETECTION

Framework Component	Role in the Proposed Disaster Detection System
IoT Sensors & Cameras	Capture visual and environmental data (fire, smoke, water levels)
Raspberry Pi (Edge AI)	Runs the YOLOv8 model to detect disasters locally
YOLOv8 AI Model	Identifies fire and flood incidents with real-time object detection
Communication Networks	Sends detection alerts to the IoT Operations Center
IoT Operations Center	Processes incoming disaster data and triggers emergency responses
Public Alert Systems	Notifies authorities and affected populations via smart notifications

X. EXAMINATION OF THE ALGORITHM

In our experiment, we utilized the YOLOv8 model - and algorithm as in Figure 7 - with a picture from [24] (see Figure 8) to detect fire as a potential hazard within a given environment. Initially, the model was employed without incorporating fire-specific learning, resulting in the detection of various objects present in the scene, but failing to identify fire accurately, as shown in Figures 9 and 10. However, after integrating fire learning detection into the model, it successfully recognized fire occurrences with improved precision. The experiment was conducted using a standard laptop for processing and presenting the fire images situations while a mobile phone is used as a real-time camera to capture live video streams, demonstrating the feasibility of implementing fire detection in practical scenarios, as shown in Figures 11 and 12.

```
import cv2
from ultralytics import YOLO

# Load the multi-class model
model_path = '/path/to/best.pt'
model = YOLO(model_path)

cap = cv2.VideoCapture(0)
cap.set(cv2.CAP_PROP_FRAME_WIDTH, 640)
cap.set(cv2.CAP_PROP_FRAME_HEIGHT, 480)

if not cap.isOpened():
    print("Error: Could not open the camera")
    exit()

while True:
    ret, frame = cap.read()
    if not ret:
        break

    # Perform fire & flood detection
    results = model(frame, conf=0.7, iou=0.4)

    # Annotate and display detections
    annotated_frame = results[0].plot()
    cv2.imshow('Fire & Flood Detection', annotated_frame)

    # Press 'q' to quit
    if cv2.waitKey(1) & 0xFF == ord('q'):
        break

cap.release()
cv2.destroyAllWindows()
```

Figure 7. The Proposed Python Script.



Figure 8. Test Picture from [24].

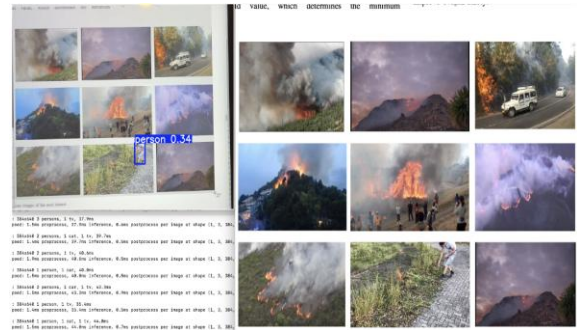


Figure 9. Screenshot from the Laptop - No Fire Detection.

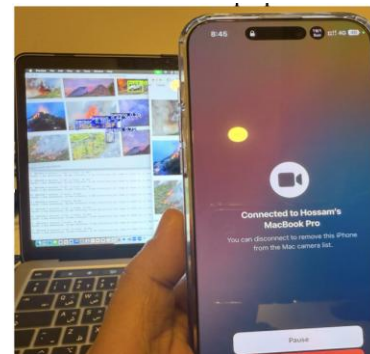


Figure 10. Experiment on the Laptop - No Fire Detection.

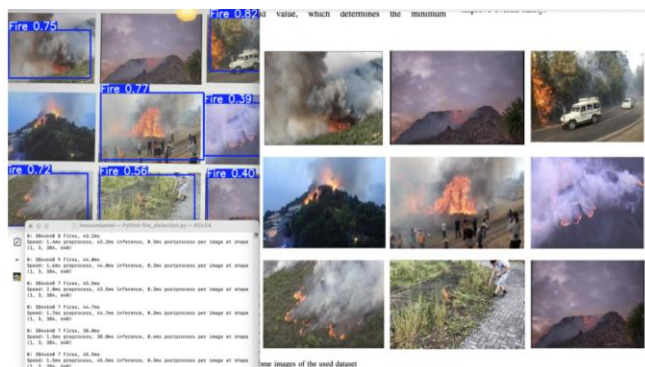


Figure 11. Screenshot from the Laptop - Fire Detection.



Figure 12. Experiment on the Laptop - Fire Detection.

XI. CONCLUSION AND FUTURE WORK

Smart city disaster detection is a critical topic that requires innovative solutions to enhance urban resilience and safety. In this work, we presented a comprehensive framework for disaster detection, focusing on fire and flood scenarios. We proposed a solution leveraging the YOLOv8 algorithm for real-time fire and flood detection, demonstrating its potential for effective disaster response. To validate our approach, we conducted an experiment using a laptop and a mobile phone, which successfully proved the effectiveness of machine learning in detecting fire incidents. As part of our future work, we aim to extend the validation to flood detection and conduct further tests under various flood scenarios to ensure the robustness and reliability of our solution.

In this study, we tested the fire detection capabilities of the YOLOv8 algorithm using laptop connected to mobile camera as a simulation of camera in a rural environment. The choice of a rural test setting was intentional, as wildfires are a major threat to smart cities. In many cases, early detection of wildfires in forests and suburban areas can prevent fires from spreading into urban zones, which lack sufficient open-space fire barriers.

While urban areas typically use CCTV-based fire detection, the proposed system is also designed for deployment in forests and highways where traditional fire detection is limited. Our initial tests in rural environments demonstrate

the feasibility of detecting fire hazards before they reach populated areas.

Planned Urban Testing: Future work will integrate the same YOLOv8-based detection system with CCTV feeds from city surveillance cameras, allowing detection of street fires, car fires, and industrial fires in real-time.

The system is designed to complement existing fire detection methods by focusing on early wildfire detection in rural and peri-urban areas, where traditional fire alarms are not available. While urban buildings have smoke detectors and suppression systems, wildfires pose a greater risk to city outskirts, requiring AI-powered monitoring. The proposed approach leverages image-based detection using Raspberry Pi and UAVs, ensuring early intervention before fires spread to cities. Given recent events like the LA wildfires, this use case is both timely and necessary.

Flood detection is planned as future work, with efforts focused on training YOLOv8 on flood datasets and integrating IoT water level sensors for real-world validation.

REFERENCES

- [1] H. M. Kim, S. Sabri, and A. Kent, "Smart cities as a platform for technological and social innovation in productivity, sustainability, and livability: A conceptual framework," *Smart Cities Technol. Soc. Innov.*, pp. 9–28, 2021, doi: 10.1016/B978-0-12-818886-6.00002-2.
- [2] A. Sharifi, A. R. Khavarian-Garmsir, and R. K. R. Kummitha, "Contributions of Smart City Solutions and Technologies to Resilience against the COVID-19 Pandemic: A Literature Review," *Sustainability*, vol. 13, no. 14, Art. no. 14, Jan. 2021, doi: 10.3390/su13148018.
- [3] R. Gaire *et al.*, "Internet of Things (IoT) and Cloud Computing Enabled Disaster Management," in *Handbook of Integration of Cloud Computing, Cyber Physical Systems and Internet of Things*, R. Ranjan, K. Mitra, P. Prakash Jayaraman, L. Wang, and A. Y. Zomaya, Eds., Cham: Springer International Publishing, 2020, pp. 273–298. doi: 10.1007/978-3-030-43795-4_12.
- [4] A. Khan, S. Gupta, and S. K. Gupta, "Emerging UAV technology for disaster detection, mitigation, response, and preparedness," *J. Field Robot.*, vol. 39, no. 6, pp. 905–955, 2022, doi: 10.1002/rob.22075.
- [5] W. He, Z. (Justin) Zhang, and W. Li, "Information technology solutions, challenges, and suggestions for tackling the COVID-19 pandemic," *Int. J. Inf. Manag.*, vol. 57, p. 102287, Apr. 2021, doi: 10.1016/j.ijinfomgt.2020.102287.
- [6] H. Jiang, M. Peng, Y. Zhong, H. Xie, Z. Hao, J. Lin, X. Ma, and X. Hu, "A Survey on Deep Learning-Based Change Detection from High-Resolution Remote Sensing Images." Accessed: Jan. 15, 2025. [Online]. Available: <https://www.mdpi.com/2072-4292/14/7/1552>
- [7] H.-T. Nguyen, N. Ngoc, and D. Anh, "The situation of high-quality human resource in FDI enterprises in Vietnam: Exploitation and development solutions," Jan. 2020.
- [8] Z. Ren, A. S. Verma, Y. Li, J. J. E. Teuwen, and Z. Jiang, "Offshore wind turbine operations and maintenance: A state-of-the-art review," *Renew. Sustain. Energy Rev.*, vol. 144, no. C, 2021, Accessed: Jan. 15, 2025. [Online]. Available: <https://ideas.repec.org/a/eee/rensus/v144y2021ics1364032121001805.html>

- [9] N. T. Nwosu, "Reducing operational costs in healthcare through advanced BI tools and data integration," *World J. Adv. Res. Rev.*, vol. 22, no. 3, pp. 1144–1156, Jun. 2024, doi: 10.30574/wjarr.2024.22.3.1774.
- [10] E. Ismagilova, L. Hughes, Y. K. Dwivedi, and K. R. Raman, "Smart cities: Advances in research—An information systems perspective," *Int. J. Inf. Manag.*, vol. 47, pp. 88–100, 2019.
- [11] E. Ismagilova, L. Hughes, N. P. Rana, and Y. K. Dwivedi, "Security, Privacy and Risks Within Smart Cities: Literature Review and Development of a Smart City Interaction Framework | Information Systems Frontiers." Accessed: Jan. 12, 2025. [Online]. Available: <https://link.springer.com/article/10.1007/s10796-020-10044-1>
- [12] H. M. Alshuwaikhat, Y. Adenle, and T. Almuhaideb, "A Lifecycle-Based Smart Sustainable City Strategic Framework for Realizing Smart and Sustainability Initiatives in Riyadh City." Accessed: Jan. 13, 2025. [Online]. Available: <https://www.mdpi.com/2071-1050/14/14/8240>
- [13] P. Bellini, P. Nesi, and G. Pantaleo, "IoT-Enabled Smart Cities: A Review of Concepts, Frameworks and Key Technologies," *Appl. Sci.*, vol. 12, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/app12031607.
- [14] H. Kamel, "Smart Scrap Management for the Smart City: Enhancing Urban Sustainability through Technology-Driven Smart Circular Economy Model Framework," *Urban. Sustain. Soc.*, 2025, Under Review.
- [15] F. Cugurullo, "Frankenstein Urbanism: Eco, Smart and Autonomous Cities, Artificial Intelligence and the End of the City," Routledge & CRC Press. Accessed: Jan. 18, 2021. [Online]. Available: <https://www.routledge.com/Frankenstein-Urbanism-Eco-Smart-and-Autonomous-Cities-Artificial-Intelligence-and-the-End-of-the-City/Cugurullo/p/book/9781138101784>
- [16] M. M. Kuglitsch, I. Pelivan, S. Ceola, M. Menon, and E. Xoplaki, "Facilitating adoption of AI in natural disaster management through collaboration," *Nat. Commun.*, vol. 13, p. 1579, Mar. 2022, doi: 10.1038/s41467-022-29285-6.
- [17] W. Ouyang *et al.*, "BioImage Model Zoo: A Community-Driven Resource for Accessible Deep Learning in BioImage Analysis," Jun. 08, 2022, *bioRxiv*. doi: 10.1101/2022.06.07.495102.
- [18] R. Sapkota, D. Ahmed, and M. Karkee, "Comparing YOLOv8 and Mask R-CNN for instance segmentation in complex orchard environments," *Artif. Intell. Agric.*, vol. 13, pp. 84–99, Sep. 2024, doi: 10.1016/j.aiia.2024.07.001.
- [19] P. M. Harikrishnan, A. Thomas, V. P. Gopi, P. Palanisamy, and K. A. Wahid, "Inception single shot multi-box detector with affinity propagation clustering and their application in multi-class vehicle counting," *Appl. Intell.*, vol. 51, no. 7, pp. 4714–4729, Jul. 2021, doi: 10.1007/s10489-020-02127-y.
- [20] J. Chappidi and D. M. Sundaram, "Novel Animal Detection System: Cascaded YOLOv8 With Adaptive Preprocessing and Feature Extraction," *IEEE Access*, vol. 12, pp. 110575–110587, 2024, doi: 10.1109/ACCESS.2024.3439230.
- [21] O. Zell, J. Pålsson, K. Hernandez-Diaz, F. Alonso-Fernandez, and F. Nilsson, "Image-Based Fire Detection in Industrial Environments with YOLOv4," in *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods*, Lisbon, Portugal: SCITEPRESS - Science and Technology Publications, 2023, pp. 379–386. doi: 10.5220/0011689400003411.
- [22] O. M. Khare, S. Gandhi, A. M. Rahalkar, and S. Mane, "YOLOv8-Based Visual Detection of Road Hazards: Potholes, Sewer Covers, and Manholes," Oct. 31, 2023, *arXiv*: arXiv:2311.00073. doi: 10.48550/arXiv.2311.00073.
- [23] P. R. Kumbam, and K. M. Vejre., "FloodLense: A Framework for ChatGPT-based Real-time Flood Detection." Accessed: Jan. 18, 2025. [Online]. Available: <https://arxiv.org/html/2401.15501v1>
- [24] F. M. Talaat and H. ZainEldin, "An improved fire detection approach based on YOLO-v8 for smart cities," *Neural Comput. Appl.*, vol. 35, no. 28, pp. 20939–20954, Oct. 2023, doi: 10.1007/s00521-023-08809-1.
- [25] S. Norkobil Saydirasulovich, A. Abdusalomov, M. K. Jamil, R. Nasimov, D. Kozhamzharova, and Y.-I. Cho, "A YOLOv6-Based Improved Fire Detection Approach for Smart City Environments," *Sensors*, vol. 23, no. 6, Art. no. 6, Jan. 2023, doi: 10.3390/s23063161.

On-the-Edge Inference Enabled Vision System for Smart Cities

Carmelo Scribano^{1,2,3}, Ignacio Sanudo Olmedo², Micaela Verucchi², Danda Pani Paudel¹,
Marko Bertogna², Luc Van Gool¹

¹ INSAIT, Sofia University “St. Kliment Ohridski”, Bulgaria ² HIPERT, University of Modena and Reggio Emilia, Italy

³ Institute of Informatics and Telematics, National Research Council, Italy

e-mail: carmelo.scribano@unimore.it, ignacio.sanudo@hipert.it, micaela.verucchi@hipert.it,
danda.paudel@insait.ai, marko.bertogna@unimore.it, luc.vangool@insait.ai

Abstract—This work introduces a next-generation smart city platform using a novel embedded vision system. i.e., HAura. The HAura system integrates a dual camera and other sensors with a powerful embedded computing unit. The powerful perception stack, based on robust deep learning and computer vision techniques, provides a perfect baseline for implementing a variety of security, traffic management and urban planning policies. Choosing to process images directly on the device and transmit only metadata ensures compliance with privacy and security, as well as bandwidth efficiency. The next evolution of the vision stack will finally improve capabilities by introducing a new multi-task perception model.

Keywords—smart-city; edge inference; safety; privacy-preserving AI.

I. INTRODUCTION

Increasing urbanization brings several challenges regarding public safety, congestion control, and the search for more efficient infrastructure. More urban centers are turning to camera-based smart-city solutions, which use computer vision, machine learning, and built-in data anonymization techniques to monitor public areas in real time, detecting and recognizing vehicles and pedestrians. These technologies are potentially the way to enable a new level of situational awareness and decision-making. In this paper, we present a comprehensive solution for smart-city infrastructure implementation, within the mentioned context. The designed system includes cameras and other sensors, in addition to communications systems. A powerful but efficient embedded computing unit performs local image analysis, leveraging a modern computer vision stack based on deep learning techniques. Running the vision stack on the edge device, instead of on a central compute server, provides all the flexibility and scalability of a fully decentralized architecture. In addition, the images are never transmitted, which makes the system effective in protecting the privacy of citizens. The next chapter discusses the technical hardware and software details of the developed system. In the following one, the details of the current vision stacks are presented. In the end, the development of a next-generation vision stack is introduced, which will largely improve the abilities of the system without overburdening the computational capabilities. The remainder of this paper is structured as follows:

- In Section II, we discuss the smart city application scenario in which we operate the proposed HAura system.

- In Section III, we detail the hardware and software technical characteristics of the HAura system, including the execution stack and data representation format.
- In Section IV, we delve in the details of the proposed preception stack, based on powerful and efficient deep models. We also anticipate ongoing work on the development of a next-generation perception model.
- Section V summarizes this overview and provides additional insights.

II. OUTLINE

A. Motivation

In accordance with projections, it is anticipated that 68% of the global population is expected to live in urban areas by 2050 [1]. Consequently, there is an imperative need for improved city management, particularly in terms of security and safety measures. An urban monitoring system is implemented using a connected camera infrastructure, with several technological and non-technological challenges involved. As pointed out by [2], human monitoring operators are easily overwhelmed by simultaneous monitoring of multiple screens. Therefore, there is a prevailing need for automated and accurate monitoring systems. Current computer vision systems are already used to implement sophisticated systems for traffic monitoring [3][4], road safety, emergency detection [4][5] and urban planning [6]. However, the algorithmic scenario is constrained by the ability to comply with privacy regulations and technological limitations dictated by available economic resources.

The HAura system processes footage on the local edge computing unit, sending only the resulting metadata over a dedicated low-latency network (bypassing the public Internet), achieving an end-to-end latency below 150 ms. In contrast, conventional IP camera setups rely on the Internet, introducing hundreds of milliseconds—or even seconds—of delay. Low-latency is essential to enable the interaction between the smart city and connected vehicles, enhancing vehicle perception by providing critical information.

B. HAura embedded system

The HAura embedded kit, shown in Figure 1, is composed of a dual camera, computing board, and software, and enables real-time identification, geolocalization, and tracking of vehicles, pedestrians, and various road users. HAura computes all the data onboard. The metadata produced by the device is



Figure 1. HAura hardware installed at the Modena Automotive Smart Area (MASA) [7]. Modena, Italy.

seamlessly transmitted to a server. Depending on the municipality or the private entity, the server, leveraging the metadata produced by the HAura infrastructure, implements different applications to monitor road users and execute smart urban strategies. At the time of writing, HAura is being implemented in several Italian cities to improve public safety and optimize traffic flow, including Modena, Reggio Emilia, and Torino.

III. TECHNICAL DESCRIPTION

A. HAura Technology stack

The proposed device, named HAura, is a smart road side unit designed for safety management and data analysis in smart cities and industrial contexts. Specifically, the system processes data from two cameras continuously, with an image transmission frequency of 10Hz. The metadata produced is sent to a server that can implement any urban monitoring policy utilizing the produced data.

a) Hardware Description: Enclosed in a rugged waterproof case, the HAura's computing heart is based on the Nvidia Orin Nano embedded platform. This choice is popular in the embedded computer vision domain because of the performance of the Nvidia Graphics Processing Unit (GPU) included in the Orin System on Chip (SoC).

- **Computing:** Specifically, the Orin Nano SoC is based on a 6-core Arm Cortex A78A Central Processing Unit (CPU), an Ampere-based Nvidia GPU with 1024 Cuda cores and 32 Tensor cores. It is also equipped with 8Gb of unified Low-Power Double Data Rate 5 (LPDDR5) memory.
- **Sensors:** The sensor set comprises two Red-Green-Blue (RGB) cameras. These cameras offer a wide 120° field of view, ensuring comprehensive coverage of the surveillance area. The system supports a resolution from 640x480 to 1920x1080
- **Connectivity:** Mainly the device is designed for low-latency 5G connectivity. The system also supports Wi-Fi (2.4GHz and 5GHz) and Ethernet. To complement this, it is equipped with a Global Positioning System

(GPS) antenna, which is useful for automating the post-installation operations, ensuring accurate localization.

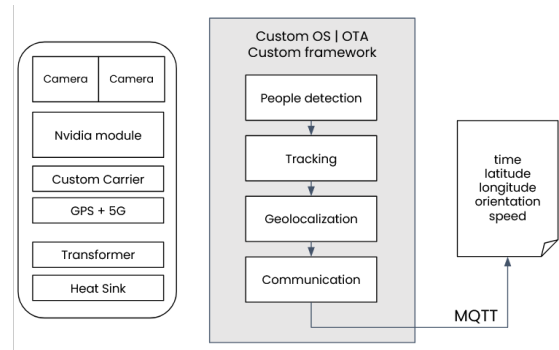


Figure 2. Diagram of the HAura's hardware, processing pipeline, and data produced.

b) Software Description: The full software suite is ready to power a full-fledged smart city infrastructure. At the base of the software stack, we have a custom Linux-based operating system (OS), boasting essential capabilities of remote performance assessment and over-the-air (OTA) updates. A key component, the HAura's perception stack depicted in Figure 2, facilitates real-time object detection and tracking over time. The upcoming frames from both cameras are processed in parallel using the powerful computer vision infrastructure detailed at Section IV-A. The obtained detections include pedestrians and different vehicles (cars, bicycles, vans, buses and motorcycles). The output is processed to recover GPS coordinates of detected objects and perform tracking of detections over time. The resulting JSON, exemplified in Figure 3, includes categorized information represented by numeric IDs (e.g., 0 for a person), latitude, longitude, tracking ID, device ID, and detection timestamp.

```

1 {
2   "camIdx": 0,
3   "nObjects": 1,
4   "objects": [
5     {
6       "latitude": 45.06582260131836,
7       "longitude": 7.662070274353027,
8       "speed": 0.0,
9       "orientation": 0,
10      "id": 1089,
11      "cl": 2
12    }
13  ]
14 }

```

Figure 3. Example json snippet with object data.

Only this metadata is transmitted, leveraging the Message Queuing Telemetry Transport (MQTT) protocol. No images of any kind are included (faces, license plates etc.). This choice is an important factor in preserving citizens' privacy and assuring compatibility with the strictest regulations.

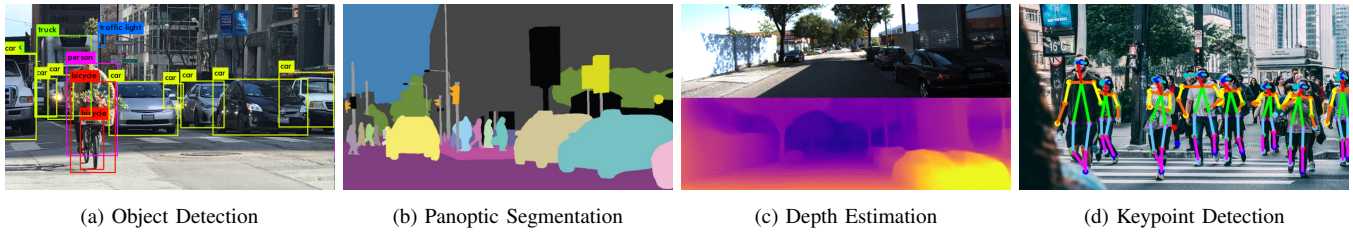


Figure 4. Visual representation of the different tasks implemented.

IV. COMPUTER VISION STACK

A. HAura perception stack

The current version of the HAura perception stack follows a conservative approach based on known techniques that have been proven to stand the test of time. The core of the vision system is based on the object detection task, which involves identifying and localizing objects of interest.

Overall, the current perception stack is structured as follows:

- **Self Diagnostic:** A small Resnet-18 [8] model, trained on a specialized proprietary dataset, is capable of classifying incoming images to detect abnormal conditions such as intense dirt or occlusion sources. This model is run sporadically (every several minutes) and is therefore not relevant to the overall latency.
- **Object Detection:** The core of the vision system is based on the YoLo-V4 [9] object detector, trained on the 80-classes MS COCO dataset [10]. YoLo-V4 is preferred to newer models because of its good balance between performance and low inference cost. Of the 80 classes, we select 6 of interest (person, car, bike, bicycle, truck, bus).
- **Multi-object tracking:** We use an extremely efficient tracker based on ByteTrack [11]. This tracker works by associating the detections of successive frames and does not require additional deep models, ensuring excellent execution performance.
- **Mapping to GPS:** Using calibrated camera extrinsic, the object detections are mapped to GPS coordinates using the inverse perspective mapping technique.

The vision stack runs entirely on NVidia embedded hardware. Model inference is accelerated using the proprietary TensorRT framework, currently version 8.6.1.

TABLE I. PERFORMANCE METRICS FOR DIFFERENT NUMBERS OF CAMERAS. THE REPORTED LATENCY (MILLISECONDS) OBTAINED BY AVERAGING OVER 1000 FRAMES.

Stage	1 Camera	2 Cameras	4 Cameras
Detection	19.98	32.41	64.85
GeoTracking	0.77	0.78	0.85
End2End	21.15	33.63	66.96

In Table I, we report an analysis of inference performance. Detection only replays the inference time of YoLo-V4 network. GeoTracking refers to the combination of ByteTrack tracker and Inverse Perspective Mapping (IPM) in GPS coordinates. End2End latency in the end includes the complete

execution cycle, including decoding the image and processing of the results in the desired format. From a performance standpoint, for precise identification, the system guarantees the following recognition ranges: 40 meters for pedestrians, 45 meters for cyclists, and 50 meters for cars.

B. Future Multi-Task perception model

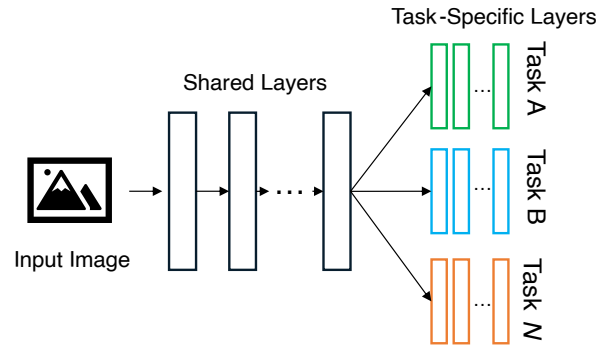


Figure 5. Outline of a multi-task learning architecture.

The current development effort is focused on the next generation of the perception stack. The underlying deep learning model is based on a multi-task learning paradigm [12]. A multi-task approach involves a single model being able to produce detection for multiple distinct tasks in a single forward pass. Compared with a classical approach, in which a specific and separate model is trained for each task, the multi-task approach has several advantages. Typically, as exemplified in Figure 5, a single backbone is used for feature extraction from the image. Only a shallow decoder is added for each task. This implies that the computational cost introduced by each additional task is marginal to the overall computational cost. In addition, in a positive-transfer effect, simultaneous learning of related tasks introduces a regularization effect that can potentially boost performance and generalization ability compared with single-task learning.

We are currently developing the model to include the following tasks, depicted in Figure 4, that presents an immediate application for numerous downstream applications.

- **Object Detection:** This is the same task underlying the current stack. Accurate prediction of bounding boxes remains a crucial element. The new enhanced model will lead to improved detection performance.

- **Panoptic Segmentation:** This task extends beyond object detection by assigning a semantic class label to each pixel while simultaneously assigning a unique label to each object instance (e.g., individual vehicles or pedestrians). Panoptic segmentation will enable a more comprehensive understanding of the urban scene.
- **Depth Estimation:** This task involves predicting per-pixel depth values to reconstruct the three-dimensional structure of the scene. This task is essential for estimating distances to objects and understanding their spatial relationship in the real world. Accurate depth data will enhance functionalities such as collision detection, and more precise GPS localization.
- **Keypoint estimation:** This task focuses on identifying and localizing critical points on objects, such as human body joints for pedestrians. This task enables fine-grained analysis of movement patterns of pedestrian, essential for advanced techniques of for behavior prediction.

The tasks of object detection, segmentation and keypoint estimation are all trainable on the MS COCO dataset, which provides the required annotations. The depth estimation ground-truth is not included, although there are dedicated datasets such as NYU Depth V2 [13] or KITTI [14], training a multi-task dataset on heterogeneous datasets is non-trivial. To overcome this limitation, we are considering leveraging pseudo labels for COCO images obtained using a powerful foundation model like DepthAnything [15].

C. Foundation Backbone

A second innovation, in addition to the multi-task paradigm, is to base the feature extractor of the new model on a powerful foundation model. In vision, a backbone foundation is obtained by pre-training the model with special techniques on a large scale, millions or even billions of images. A prominent example is Dino-V2 [16]. This backbone has been trained on a large dataset of 142 million images using a self-supervised learning approach derived from [17]. With the large-scale pre-training, the foundation models learn strong feature extraction ability, when fine-tuning on downstream tasks therefore the final model will show exceptional performances and strong generalization ability. The main disadvantage is that it is not possible to replicate pretraining on a large scale because of the huge costs and lack of proprietary training data. Therefore, we must start from the pretrained models released by the authors and keep the same model architecture. In particular, Dino-V2 is based on the Vision Transformer (ViT) family of models [18], which are generally considered expensive in terms of computational resources. For this reason, a crucial phase of the work is focused on reducing the computational cost of ViT models while maintaining compatibility with the pretrained weights of Dino-V2.

D. Computing cost reduction

Reducing computational cost, hence inference time, without degrading performance is a key goal for inference on edge devices. The TensorRT inference framework provided by

NVIDIA already implements a large set of generic techniques to accelerate inference: the proprietary TensorRT compiler is capable of optimizing the inference graph, performing complex fusion of operations and carefully selecting inference kernels to maximize performance. In addition, different tools are provided to implement techniques such as quantization and pruning [19]. In addition, there is extensive scientific literature of techniques to mitigate the inefficiencies of specific categories of models. Our current work includes developing a specific novel technique to further accelerate the inference of the ViT models on which DinoV2 is based that we use as the backbone for the multi-task model.

V. CONCLUSION

In this paper, we introduced the HAura hardware and software stack, a generic platform for smart city infrastructure. Our design leverages on-edge inference, ensuring privacy protection by transmitting only data that complies with existing regulations, thereby reducing the risk of exposing sensitive information. The platform allows adopters to develop front-end applications that utilize aggregated metadata for a variety of purposes, from real-time traffic monitoring to long-term urban planning. While our initial evaluations are promising, we recognize that further work is needed to thoroughly assess privacy guarantees and regulatory compliance in diverse settings. Future efforts will focus on developing a comprehensive front-end platform with the most requested functionalities and on evaluating the use of Large Language Models to simplify aggregate data querying. Although specialized systems are available for individual applications, to the best of our knowledge our platform represents a unique step toward a universal, upgradeable, and reconfigurable solution for smart city infrastructure.

ACKNOWLEDGEMENT

This research was partially funded by the dAIedge project (HORIZON-CL4-2022-HUMAN-02-02, Grant Agreement Number: 101120726) and the Ministry of Education and Science of Bulgaria (support for INSAIT, part of the Bulgarian National Roadmap for Research Infrastructure). C. Scribano work was partly funded by the Partenariato Esteso PE00000013 - "FAIR", funded by the European Commission under the NextGeneration EU program, PNRR - M4C2 - Investimento 1.3.

REFERENCES

- [1] *World Urbanization Prospects: The 2018 Revision*. UN, Aug. 2019, ISBN: 9789210043144. DOI: 10.18356/b9e995fe-en.
- [2] J. Aldridge, *CCTV operational requirements manual: Who will be the first to test your CCTV security or safety system?* 1994.
- [3] G. M. Lingani, D. B. Rawat, and M. Garuba, "Smart traffic management system using deep learning for smart city applications," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, 2019, pp. 0101–0106. DOI: 10.1109/CCWC.2019.8666539.
- [4] N. T. Ha *et al.*, "Leveraging deep learning model for emergency situations detection on urban road using images from cctv cameras," in *2022 International Conference on Engineering and Emerging Technologies (ICEET)*, IEEE, 2022, pp. 1–5.

- [5] R. Nouisser, S. K. Jarraya, and M. Hammami, "A review of vision-based abnormal human activity analysis for elderly emergency detection," in *2023 International Conference on Innovations in Intelligent Systems and Applications (INISTA)*, IEEE, 2023, pp. 1–6.
- [6] N. Naik, S. D. Kominers, R. Raskar, E. L. Glaeser, and C. A. Hidalgo, "Computer vision uncovers predictors of physical urban change," *Proceedings of the National Academy of Sciences*, vol. 114, no. 29, pp. 7571–7576, 2017, [Accessed 16-02-2025]. DOI: 10.1073/pnas.1619003114. eprint: <https://www.pnas.org/doi/pdf/10.1073/pnas.1619003114>.
- [7] MASA – a living lab for automated driving – Modena Automotive Smart Area — [automotivesmartarea.it](https://www.automotivesmartarea.it/), <https://www.automotivesmartarea.it/>, [Accessed 16-02-2025].
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [9] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, *Yolov4: Optimal speed and accuracy of object detection*, 2020. arXiv: 2004.10934 [cs.CV].
- [10] T.-Y. Lin *et al.*, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, Springer, 2014, pp. 740–755.
- [11] Y. Zhang *et al.*, "Bytetrack: Multi-object tracking by associating every detection box," in *European conference on computer vision*, Springer, 2022, pp. 1–21.
- [12] S. Vandenhende *et al.*, "Multi-task learning for dense prediction tasks: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 7, pp. 3614–3633, 2021.
- [13] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from rgb-d images," in *Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part V 12*, Springer, 2012, pp. 746–760.
- [14] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *International Journal of Robotics Research (IJRR)*, 2013.
- [15] L. Yang *et al.*, "Depth anything: Unleashing the power of large-scale unlabeled data," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 10 371–10 381.
- [16] M. Oquab *et al.*, "Dinov2: Learning robust visual features without supervision," *arXiv preprint arXiv:2304.07193*, 2023.
- [17] J.-B. Grill *et al.*, "Bootstrap your own latent-a new approach to self-supervised learning," *Advances in neural information processing systems*, vol. 33, pp. 21 271–21 284, 2020.
- [18] D. Alexey, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv: 2010.11929*, 2020.
- [19] T. Liang, J. Glossner, L. Wang, S. Shi, and X. Zhang, "Pruning and quantization for deep neural network acceleration: A survey," *Neurocomputing*, vol. 461, pp. 370–403, 2021.

Enhancing Smart City Sustainability: Anaerobic Treatment of Semiconductor Industry Wastewater for Improved Efficiency and Environmental Impact

Bruno Escobar¹, Ana Santos¹, Flávio Silva^{1,2}, Helena Nadais^{1,2}

¹Department of Environment and Planning

²CESAM, Centre for Environmental and Marine Studies

University of Aveiro, Portugal

e-mails: {bruno.p.escobar, arnds, flavio.silva, nadais}@ua.pt

Abstract- As cities transition into smart urban environments, managing industrial waste, especially from semiconductor manufacturing essential for smart technologies, becomes paramount. This study assesses the anaerobic treatment of semiconductor industry wastewater, which is laden with heavy metals and organic solvents. Our research focuses on the acclimation of anaerobic biomass and its effectiveness in treating these complex wastewaters over a 132-day period using semi-continuous reactors. We employed a phased approach starting with an initial stimulation using sucrose to boost microbial activity, followed by gradual increases in effluent concentration. The process culminated in stabilization phases where effluent mixtures were managed to evaluate adaptation and efficiency. Throughout these phases, we monitored methanization and Chemical Oxygen Demand (COD) removal, achieving average efficiencies of 61% and 79% respectively. Our findings underscore the complex dynamics between microbial communities and the unique constituents of semiconductor wastewater. While the presence of inhibitory substances challenged methanogenic activity, particularly in the latter stages with higher contaminant loads, the treatment system demonstrated significant resilience. This suggests that while the core anaerobic processes are effective, supplementary pre- and post-treatments could be necessary to handle the high concentrations of contaminants typical in semiconductor wastewater. The study confirms the feasibility of using anaerobic processes to manage the demanding effluents of semiconductor manufacturing, a critical component in smart cities. By enhancing wastewater treatment strategies, this research contributes to the sustainability of smart urban environments, reducing environmental impacts and supporting the continued development of essential smart technologies.

Keywords- smart cities; Sustainable urban development; Semiconductor wastewater; Anaerobic digestion; Biomass acclimation; Methane production; Chemical oxygen demand.

I. INTRODUCTION

In today's urban landscapes, where the concept of smart cities is becoming increasingly integral to sustainable development, semiconductors play a pivotal role. With a robust market value of \$543 billion [1], semiconductors are fundamental to the evolution and functionality of smart city technologies. These materials, known for their specific conductivity properties that can be tailored through doping processes [2], are crucial in determining the performance of electronic circuits and systems.

Silicon, the predominant material used in semiconductor manufacturing, along with other metalloids, forms the backbone of countless smart city applications. From traffic management systems that rely on sensors to public safety solutions empowered by smart surveillance technologies, semiconductors are at the heart of these innovations. They enable the development and efficient operation of LED lighting systems, advanced public transport networks, and integrated communication systems—all essential components of the smart city infrastructure. The production of semiconductors thus underpins not only traditional electronics like smartphones and LED TVs but also the sophisticated microelectronics that drive the smart cities of the future. This interconnection highlights the indispensable role of semiconductor technology in building urban environments that are more connected, sustainable, and responsive to the needs of their inhabitants.

In the evolving landscape of smart cities, the semiconductor manufacturing process plays a pivotal role, requiring high levels of purity to ensure the optimal performance of smart technologies. The complex manufacturing process involves multiple stages including wafer fabrication, oxidation, photolithography, etching, ion deposition and implantation, metallization, and electrical matrix sorting and packaging [3]. These processes use a broad spectrum of chemicals such as metals, solvents, and acids, necessitating extensive cleaning with ultrapure water. In 2022, the consumption of ultrapure water by the semiconductor industries globally was estimated at approximately $5.51 \times 10^8 \text{ m}^3$ [4].

The usage of such significant amounts of ultrapure water results in the production of complex WasteWater (WW), characterized by various pollutants like TetraMethylAmmonium Hydroxide (TMAH), phosphoric acid, ammonia, surfactants, organic solvents, and heavy metals [5]. These contaminants, often recalcitrant and environmentally harmful, predominantly exhibit organic characteristics, offering opportunities for biological treatment processes. Anaerobic digestion emerges as a promising method for treating semiconductor WW. Previous researchs validated the effectiveness of anaerobic processes in breaking down chemicals like DiMethyl SulfOxide (DMSO) and TMAH commonly found in these effluents [6][7].

For smart cities, the advantages of anaerobic treatment, such as minimal sludge production, energy recovery via methane production, and the feasibility of compact design, align well with the sustainability goals of reducing operational costs and enhancing energy efficiency [8]. However, the efficiency of anaerobic digestion is contingent upon the concentration of pollutants, as there is a threshold to the degradation capabilities of anaerobic microorganisms.

This study aims to conduct a preliminary analysis to assess the potential of anaerobic biomass for treating recalcitrant effluents produced in the semiconductor industry, underscoring its relevance in supporting sustainable urban development within smart cities. The Materials and Methods (Section II) details the origin and characteristics of the anaerobic inoculum and semiconductor WW (A. Materials), followed by the B. Experimental Setup and Operation, which outlines system configuration, operational parameters, maintenance, and assay duration. The Analytical Methods subsection C, specifies the primary analyses and methodologies employed. In the Results and Discussion (Section III), the complexity and variability of semiconductor WW are examined, highlighting its treatment potential through anaerobic digestion. Tables summarize WW characteristics before and after treatment, with comparative analysis against existing studies. Future research directions and treatment optimizations are also proposed. The Conclusion (Section VI) synthesizes key findings, emphasizing anaerobic digestion's effectiveness and prospects for further study. The paper concludes with Acknowledgments recognizing key contributors and a References section listing all cited sources.

II. MATERIALS AND METHODS

This section describes the materials used and the methods employed in this study.

A. Materials

Two types of effluent were obtained from a semiconductor industry located in northern Portugal, one with chemicals and diluted acid mixture (E1) and the other with a mixture of E1 and strong acids (E2). More specifically, it was also confirmed through the processes carried out by the industry that TMAH, isopropanol, a non-ionic surfactant, sodium persulfate, copper sulfate, citric acid, acetic acid, sulfuric acid, hydrofluoric acid and phosphoric acid were utilized in varying concentrations depending on the needs of production. The tanks from where the effluent was collected contained a mixture resulting from different processes, mainly lithography, packaging, plating, dicing, grinding and laser grooving and the many steps of washing and cleaning. To avoid any setbacks with the strong acidic content in the effluent E2, it was collected after the pH control step. Anaerobic inoculum was obtained from a local MWTP treating both domestic and industrial WW.

B. Experimental setup and operation

The experiments were performed on a laboratory scale, in four separate phases, lasting for a total of 132 days, as follows. In the Stimulation phase (36 days) the biomass was only fed with sucrose to enhance the metabolic activity and establish a baseline. The Acclimation phase (30 days) the biomass was fed with a continuous step-increase (10%) in effluent concentration, summing up 10 moments (10%, 13.3%, 17.7%, 23.6%, 31.4%, 41.8%, 55.6%, 74%, 98.5% and 100%). Stabilization 1 phase (15 days) consisted in feeding on 100% effluent. Finally, in Stabilization 2 (51 days) the biomass was fed with a mixture of effluents collected from different periods of the industrial operation, to evaluate a broader and more complex effluent. This phase was also incremented with a Simulated Wastewater (SW) solution made with sodium acetate (representing dissolved acetic acid, a common acid heavily used by this industry) and TMAH with a COD of 145 g L⁻¹.

The anaerobic assays were carried out in four glass reactors, two with a working volume of 5 L and two with working volume of 2 L. A fifth 2 L reactor was used as a control for growing inoculum fed only with sucrose. All assays started with a biomass concentration of 7 g VSS L⁻¹. Neutralized effluent was fed, and samples were collected every three days, with reactors maintained at 35°C. At each phase, nutrients and sodium bicarbonate were added to support digestion [9]. Biomass sludge samples were taken every six days to assess biomass concentration.

C. Analytical methods

Effluent characterization before and after anaerobic treatment followed Standard Methods [10], assessing COD, Biochemical Oxygen Demand (BOD, 5 days, Oxitop®), pH, Electrical Conductivity (EC), Alkalinity, and Total Volatile Acids (TVA). Anaerobic biomass concentration and activity were evaluated via Volatile Suspended Solids (VSS) and Specific Methanogenic Activity (SMA). Methane was purified by NaOH (20% w/w) gas washing [11] and quantified using a syringe [12].

III. RESULTS AND DISCUSSION

The management of wastewater from semiconductor manufacturing is a significant challenge, directly impacting urban sustainability and smart infrastructure. The complexity and variability of semiconductor WW are attributed to the diverse production techniques employed, the mixing of effluents from various processes, and particularly the dilution effects from cleaning operations, as previously highlighted [13]. These factors contribute to the broad range of characteristics observed in semiconductor WW, making it difficult to standardize treatment approaches. The WW samples used in this study are not different, and although they are from the same tank, the difference on collection day is enough to demonstrate high variability in all parameters, as

can be observed in Table I. Despite the variations, the characteristics are still within the WW profile of this type of industry [13].

TABLE I. AVERAGE PHYSICAL-CHEMICAL VALUES FOR THE SEMICONDUCTOR WASTEWATER USED IN THIS WORK

Parameters	Diluted Acids tank (E1)	Lowest - Highest values	Effluent Mixture tank (E2)	Lowest - Highest values
pH	6.06	4.74 - 9.10	9.47	5.48 - 11.50
EC (mS cm ⁻¹)	3.93	0.27 - 11.6	4.43	3.49 - 6.20
Alkalinity (mg L ⁻¹ of CaCO ₃)	98.18	68.75 - 125	283.33	150 - 425
TVA (mg L ⁻¹)	245.84	100 - 365.63	233.34	159.38 - 375
Kjeldahl Nitrogen (mg L ⁻¹)	13.16	11.20 - 14.84	13.72	10.36 - 17.08
Total Phosphorus (mg L ⁻¹)	1.03	0.27 - 1.64	0.42	0.29 - 0.50
COD (mgO ₂ L ⁻¹)	749.91	270.4 - 1,245.4	757.19	504.1 - 1,114.1
BOD (mg L ⁻¹)	259.94	162 - 411.1	281.43	184.89 - 353.6
Biodegradable COD fraction (%)	34.66	19.83 - 82.90	37.17	25.39 - 66.3
Total Solids (g L ⁻¹)	5.63	1.85 - 11.76	3.55	1.93 - 4.39
Dissolved Solids (g L ⁻¹)	5.35	1.76 - 11.32	3.29	1.63 - 4.32
Suspended Solids (g L ⁻¹)	0.21	0.03 - 0.54	0.26	0.06 - 0.34
Volatile Solids (g L ⁻¹)	0.11	0.01 - 0.18	0.13	0.05 - 0.18

The overarching goal of this study was to conduct a preliminary analysis and demonstrate the effectiveness of anaerobic processes for treatment of this WW with high variability in its composition, as has been proposed in previous studies [7][14]. By analyzing the characteristics of the WW, it is possible to determine that it has a considerable biodegradable content, with the presence of macronutrients nitrogen and phosphorus which are essential for microbial metabolism. In addition, most of the solids, and the organics, are dissolved. It is important to note that these favorable conditions for anaerobic processes are not always found for this type of WW [8]. Therefore, it is essential to emphasize the advantages of the application of a cost-effective method capable of WW treatment that enables energy and water recovery. Despite the various favorable conditions, it is important to highlight the presence of highly recalcitrant and inhibitory compounds for microorganisms, such as fluoride [15], copper [16], surfactants [17] and TMAH itself [14], potentially leading to metabolism disruption, but adaptation of the anaerobic microbiota is expected. Table II depicts the results of the treated effluent after 132 days of operation.

TABLE II. AVERAGE PHYSICAL-CHEMICAL VALUES OF FEED AND EFFLUENT FROM ANAEROBIC DIGESTION TREATMENT

Parameters	Diluted Acids tank (E1)	Lowest - Highest values	Effluent Mixture tank (E2)	Lowest - Highest values
Feed COD (mgO ₂ L ⁻¹) ^a	1,113.9	837.6 - 1,390.7	1,211.7	901.3 - 1,433.2
Final COD (mgO ₂ L ⁻¹) ^a	230.4	130.1 - 393.8	238.4	163.7 - 370.6
COD Removed (%)	79	71 - 84	80	74 - 81
Final pH ^a	8.04	7.45 - 8.54	8.07	7.53 - 8.66
Final Alkalinity ^a	1,262.90	543.75 - 2,262.50	1,356.78	612.50 - 2,337.50
Final EC (mS cm ⁻¹) ^a	3.97	2.77 - 4.60	4.42	3.24 - 5.36
Final Total Volatile Acids (mg L ⁻¹) ^a	58.77	27.50 - 140.62	62.53	33.75 - 126.06
CH ₄ Produced (mL) ^a	131.68	12 - 412	111.50	2 - 442
Methanisation Efficiency (%) ^a	66	46 - 83	53	27 - 69
SMA (g CH ₄ -COD g ⁻¹ VSS d ⁻¹) ^a	0.009	0.002 - 0.028	0.011	0.003 - 0.038

a. Global average for each feeding run

It is worth mentioning that methane production remained constant despite the different stages. COD removal was on average 79% for both types of WW, and although the values are lower when compared with other investigations, such amount of degradation is in accordance with other studies that reported COD removal ranging 70-90% with a influent COD of 1,800 mg/L, but when the organic load was increased to 8,000 mg/L the microbial culture was inhibited and COD removal dropped to values below 70% [6]. Another investigation reported a COD removal rate of 50% for a WW containing 1500 mg/L of COD before the acclimation of biomass to TMAH degradation and achieving COD removals of 90% on average after acclimation despite being operated in psychrophilic temperatures [18]. Following studies in such conditions, reported a removal of 96% of COD [7]. It is important to emphasize that in the studies cited, synthetic or diluted WW were used, thus reducing the impact of other contaminants in a real WW.

The methanisation efficiency obtained values ranged between 66% and 56% based on the input gCOD and produced gCOD.CH₄. However, SMA had low values when compared to other works, resembling SMA from reactors with a notable presence of sulfidogenic microbiota [19]. A low SMA and high conversion of COD to CH₄ may indicate that there is little substrate available for all microorganisms, indicating that there is the possibility of generating even more methane when at higher loads. Even considering the potential presence of sulfidogenic microbiota competing for resources, there was no significant impact on methane production, as a high consumption of TVA confirms that there was no

accumulation of volatile acids, demonstrating stability in the system. The constant production of volatile acids from acidogenesis was confirmed through the alkalinity reduction throughout the experiment. It is also possible to verify from Figure 1 that there was no accumulation of organic compounds over the different phases, confirming its degradation even when higher loads were added with different mixtures of WW.

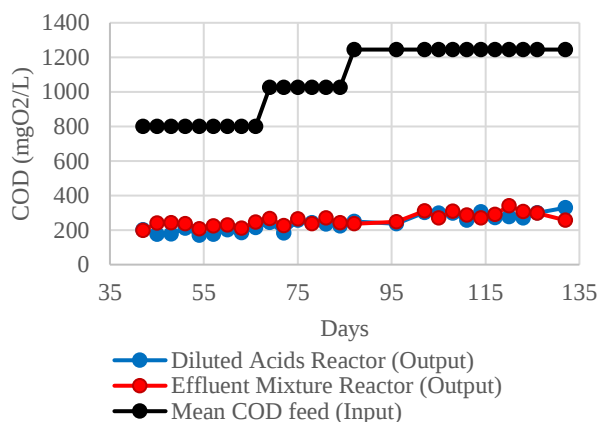


Figure 1. Mean COD concentration for feed and treated effluents, from Acclimation to Stabilization 2 phases.

Stabilization 1 achieved the highest COD removal and methanization, with SMA increasing from 0.0061 to 0.0137, indicating microbial adaptation to the effluent's composition and organic load. In contrast, Stabilization 2, with a mixed effluent containing higher contaminant concentrations, caused system destabilization, confirmed by the SMA decrease to 0.0070, suggesting inhibition. While complex effluents with harmful compounds like heavy metals or TMAH can be treated, their effectiveness is reduced. Other works also experienced a drop in overall treatment efficiency, by presence of heavy metals [20], especially copper [16] and high concentrations of TMAH [13]. Not only the compounds in the effluent can cause inhibitions, but also the degraded products of the digestion can also be inhibitory, such as the case with the degradation of TMAH, where its final product is ammonia which in high concentrations can destabilize the anaerobic system [21].

One study observed that not only the methanogens are responsible for the degradation of TMAH in semiconductor WW, corroborating that a reactor with greater diversity of microorganisms is more capable of degrading complex compounds [22]. Although sulfidogenic microorganisms can destabilize the anaerobic system due to competition for resources with methanogenic, other studies demonstrated that they are capable of mineralizing copper ions, thus being an alternative for reducing this metal in the final effluent and as a form to reduce the impacts generated by toxicity [23]. The presence of sulfidogenic microorganisms can also be effective in degrading other compounds that are not suitable for methanogenic archaea such as surfactants [24].

The experimental design and execution, characterized by semi-continuous feeding, inherently promotes the accumulation of inorganic compounds and other substances that are non-biodegradable by anaerobic microorganisms, such as copper, fluoride, and ammonia, within the reactor. Over time, as these compounds progressively accumulate, the efficiency of the anaerobic treatment is expected to decline, despite the system's adaptation to the WW. This context is possible to occur even in other types of reactors and in continuous systems due to a greater flow of effluent to the reactors, therefore, it is necessary to consider strategies such as pre and post treatments for this type of WW. Numerous studies have already sought the combination of treatments to increase the treatment efficiency of these effluents. Considering the increase of biodegradation to improve the degradation of organics, different authors proposed the use of oxidative systems as a pre-treatment to achieve this, either by using ozone [25], Fenton [26] or anodic oxidation [27] with varying degrees of efficiency. As a post-treatment for the removal of anaerobic degradation products, such as ammonia, authors have studied the use of aerobic and anoxic systems [28] reaching a maximum removal of nitrogen of 63% along with a TMAH reduction between 70 and 100%. One study combined a crystallization reactor filled with quartz salt and a sulfate reducing bioreactor to remove copper, reaching 99% and 70% removal of copper and COD, respectively [23]. Alternatively, electrochemical processes can complement anaerobic treatment by facilitating the removal of solids, including heavy metals by electro flotation, or promoting the degradation of complex compounds by electrooxidation [29].

IV. CONCLUSIONS

This study demonstrated that semicontinuous bioreactors inoculated with anaerobic microorganisms from a MWTP can effectively treat WW from the semiconductor industry. On average, methanization and COD removal efficiency were 61% and 79%, respectively, with methane production of 0.1 mL of CH₄ per mg of COD. The continuous production of methane, TVA consumption, alkaline decline and the absence of COD accumulation confirmed system stability. Low SMA and high methanization suggest limited carbon sources for microorganisms. The potential presence of sulfidogenic biota may help remove inhibitory compounds like copper and surfactants. The increased organic load and effluent complexity reduced treatment efficiency, highlighting the need for pre-treatment to improve biodegradability and post-treatment to remove inhibitory byproducts. Further research at larger scales is needed to validate these findings in industrial settings.

ACKNOWLEDGMENT

The authors express their gratitude to the University of Aveiro and Instituto Superior Técnico for providing the resources and support necessary for this research. Thanks, are also due to Portuguese Foundation for Science and

Technology (FCT)/Ministry of Science, Technology and Higher Education (MCTES), Portugal, for the financial support to UID Centro de Estudos do Ambiente e Mar (CESAM) + LA/P/0094/2020, through national funds. We acknowledge the funding support provided by the Microelectronics Agenda under the Recovery and Resilience Plan, financed by the European Union – NextGenerationEU.

REFERENCES

- [1] Semiconductor Industry Association, "Worldwide market billings (in thousands of U.S. Dollars)," Semiconductor Industry Association, 2022. Available from: <https://www.semiconductors.org/>, [retrieved: October, 2024].
- [2] K. Seeger, *Semiconductor Physics*, Springer Science & Business Media, Verlag, Vienna, 1973.
- [3] W. Tsai, H. Chen, and W. Hsien, "A review of uses, environmental hazards and recovery/recycle technologies of perfluorocarbons (PFCs) emissions from the semiconductor manufacturing processes," *Journal of Loss Prevention in the Process Industries*, vol. 15, no. 2, pp. 65–75, 2002.
- [4] United Microelectronics Corporation, "UMC sustainability report," 2022. Available from: <https://www.umc.com/>, [retrieved: October, 2024].
- [5] J. Sim, J. Lee, H. Rho, K. Park, Y. Choi, D. Kim, H. Kim, and Y. C. Woo, "A review of semiconductor wastewater treatment processes: Current status, challenges, and future trends," *Journal of Cleaner Production*, vol. 429, p. 139570, 2023.
- [6] H. Cheng et al., "Biological treatment of DMSO-containing wastewater from semiconductor industry under aerobic and methanogenic conditions," *Chemosphere*, vol. 236, p. 124291, 2019.
- [7] K. Syutsubo, T. Danshita, H. Sumino, A. Iguchi, and Y. Takemura, "Microbial properties of the granular sludge in a psychrophilic UASB reactor fed with electronics industry wastewater containing organic chemicals," *Journal of Environmental Science and Health, Part A*, vol. 56, no. 5, pp. 516–524, 2021.
- [8] M. Zieliński, J. Kazimierowicz, and M. Dębowski, "Advantages and limitations of anaerobic wastewater treatment—Technological basics, development directions, and technological innovations," *Energies*, vol. 16, no. 1, p. 83, 2022.
- [9] S. Rebac et al., "High-rate anaerobic treatment of malting wastewater in a pilot-scale EGSB system under psychrophilic conditions," *Journal of Chemical Technology and Biotechnology*, vol. 68, no. 2, pp. 135–146, 1997.
- [10] American Public Health Association (APHA), *Standard Methods for the Examination of Water and Wastewater*, 20th ed., APHA, Washington, DC, 1999.
- [11] M. S. Tahir, K. Shahzad, Z. Shahid, M. Sagir, M. Rehan, and A. Nizami, "Producing methane-enriched biogas using solvent absorption method," *Chemical Engineering Transactions*, vol. 45, pp. 1309–1314, 2015.
- [12] P. D. Nganyira, D. J. Mahushi, J. G. Balengayabo, G. N. Shao, and J. K. Emmanuel, "Quality of biogas generated through co-digestion of brewer's spent grain and cattle dung," *Energy Reports*, vol. 10, pp. 2330–2336, 2023.
- [13] E. A. Noman et al., "An insight into microelectronics industry wastewater treatment, current challenges, and future perspectives: A critical review," *Applied Water Science*, vol. 14, no. 4, 2024.
- [14] J. Lv, Q. Ni, J. Dong, C. Ou, Y. Cao, J. Gan, and F. Xiao, "Effects of influent organic load on TMAH (tetramethylammonium hydroxide) wastewater treatment by anaerobic process," *Journal of Environmental Chemical Engineering*, vol. 11, no. 2, p. 109315, 2023.
- [15] V. Ochoa-Herrera et al., "Toxicity of fluoride to microorganisms in biological wastewater treatment systems," *Water Research*, vol. 43, no. 13, pp. 3177–3186, 2009.
- [16] L. Otero-González, J. A. Field, and R. Sierra-Alvarez, "Inhibition of anaerobic wastewater treatment after long-term exposure to low levels of CuO nanoparticles," *Water Research*, vol. 58, pp. 160–168, 2014.
- [17] T. Hu et al., "Biological treatment of TMAH (tetramethylammonium hydroxide) in a full-scale TFT-LCD wastewater treatment plant," *Bioresource Technology*, vol. 113, pp. 303–310, 2012.
- [18] T. Danshita et al., "Influence of tetramethylammonium hydroxide (TMAH) on the microbial properties of anaerobic granular sludge acclimated to isopropyl alcohol (IPA) wastewater under psychrophilic conditions," *Journal of Environmental Science and Health, Part A*, vol. 53, no. 11, pp. 1015–1021, 2018.
- [19] V. M. Gámez, R. Sierra-Alvarez, R. J. Waltz, and J. A. Field, "Anaerobic degradation of citrate under sulfate-reducing and methanogenic conditions," *Biodegradation*, vol. 20, no. 4, pp. 499–510, 2008.
- [20] Y. Yang, C. Zhang, and Z. Hu, "Impact of metallic and metal oxide nanoparticles on wastewater treatment and anaerobic digestion," *Environmental Science: Processes & Impacts*, vol. 15, no. 1, pp. 39–48, 2012.
- [21] L. Whang et al., "Molecular analysis of methanogens involved in methanogenic degradation of tetramethylammonium hydroxide in full-scale bioreactors," *Applied Microbiology and Biotechnology*, vol. 99, no. 3, pp. 1485–1497, 2014.
- [22] Z. Li et al., "Anaerobic hydrolysis of recalcitrant tetramethylammonium from semiconductor wastewater: Performance and mechanisms," *Journal of Hazardous Materials*, vol. 459, p. 132239, 2023.
- [23] R. Sierra-Alvarez, J. Hollingsworth, and M. S. Zhou, "Removal of copper in an integrated sulfate-reducing bioreactor–crystallization reactor system," *Environmental Science & Technology*, vol. 41, no. 4, pp. 1426–1431, 2007.
- [24] J. Hollingsworth et al., "Anaerobic biodegradability and methanogenic toxicity of key constituents in copper chemical mechanical planarization effluents of the semiconductor industry," *Chemosphere*, vol. 59, no. 9, pp. 1219–1228, 2005.
- [25] S. Tripathi and T. Hussain, "Water and wastewater treatment through ozone-based technologies," in *Elsevier eBooks*, pp. 139–172, 2021.
- [26] S. Park, T. Yoon, J. Bae, H. Seo, and H. Park, "Biological treatment of wastewater containing dimethyl sulphoxide from the semiconductor industry," *Process Biochemistry*, vol. 36, no. 6, pp. 579–589, 2001.
- [27] E. Mousset, Z. Wang, H. Olvera-Vargas, and O. Lefebvre, "Advanced electrocatalytic pre-treatment to improve the biodegradability of real wastewater from the electronics industry—A detailed investigation study," *Journal of Hazardous Materials*, vol. 360, pp. 552–559, 2018.
- [28] H. Lin et al., "A combined upflow anaerobic sludge bed, aerobic, and anoxic fixed-bed reactor system for removing tetramethylammonium hydroxide and nitrogen from LED wastewater," *Environmental Technology*, vol. 37, no. 11, pp. 1325–1333, 2015.
- [29] H. G. Yang, H. Y. Chun, and D. Pak, "Improvement of sludge anaerobic degradability by combined electro-flotation and electro-oxidation treatment," *Biochemical Engineering Journal*, vol. 90, pp. 44–48, 2014.

Using Artificial Intelligence to Support Emergency Management Training

Creating Efficient and Realistic Scenarios with ChatGPT

Lasse Berntzen

Jarle Løwe Sørensen

Marius Rohde Johannessen

Center for Security, Crisis Management
and Emergency Preparedness

University of South-Eastern Norway

Borre, Norway

e-mail: lasse.berntzen@usn.no

jarle.sorensen@usn.no

marius.johannessen@usn.no

Laurits Rauer Nielsen

Lene Sandberg

Faculty of Social Education, Social
Work and Administration

University College, Copenhagen

Copenhagen N, Denmark

e-mail: laun@kp.dk

lesa@kp.dk

Eric Carlström

Amir Khorram-Manesh

Center for Disaster Medicine and
Gothenburg Emergency Medicine

Research Group (GEMREG),

Sahlgrenska Academy

Gothenburg, Sweden

e-mail: eric.carlstrom@gu.se

amir.khorram-manesh@surgey.gu.se

Abstract— The Erasmus+ project *TeleTraining in Crisis Management* is a collaborative initiative among the University of Gothenburg, the University of South-Eastern Norway, and the University College Copenhagen, aiming to enhance emergency management training through digital solutions. This paper explores the potential of Artificial Intelligence (AI) to improve training programs by generating realistic crisis scenarios and situational updates. Using a simulated emergency water landing of a passenger flight in international waters, the study demonstrates how AI-driven tools, such as ChatGPT, can create dynamic training environments for cross-border emergency response teams. The scenario emphasizes the need for standardized protocols, digital communication platforms, and AI-assisted decision-making in multinational crisis management. The paper further discusses the integration of AI within customized Learning Management Systems (LMS) to facilitate real-time exercises, improve coordination among emergency responders, and enhance crisis preparedness through adaptive and data-driven simulations.

Keywords—artificial intelligence; emergency management training; crisis management training; proof-of-concept; scenarios; ChatGPT.

I. INTRODUCTION

The Erasmus+ project *TeleTraining in Crisis Management* is a joint effort between the University of Gothenburg, the University of South-Eastern Norway, and the University College Copenhagen to explore new methods for crisis and emergency management training, with a focus on collaboration between different emergency services and management levels, both on national and transnational levels.

This paper explores the opportunities to use Artificial Intelligence (AI) to enhance training programs by effectively creating realistic training scenarios. We use ChatGPT to establish an initial scenario and then create situation reports as the scenario unfolds. The aim is to show that ChatGPT and similar tools can add value to emergency training exercises.

Eide, Lund-Kordahl, and Bakken [5] discussed how AI changes crisis management training and exercises. Their discussion is conceptual. Karinshak [3] promotes the idea of using simulation without going into details. A couple of papers discuss the use of ChatGPT in simulations in nursing training [1][8].

For this paper, we will use the following scenario:

A passenger flight from Oslo to Copenhagen performed an emergency landing in water 100 kilometers northwest of Gothenburg, Sweden. The incident occurred in international waters and required immediate cooperation between Danish, Norwegian, and Swedish emergency authorities and response units. The situation escalated quickly, with passengers and crew in distress, necessitating rapid and well-coordinated intervention. Sweden was closest to the scene of the incident, but as the flight originated in Norway and was on its way to Denmark, all three countries got involved. Given the international scope of the crisis, a tri-nation emergency response coordination team was established. This team used digital communication tools and real-time situational awareness platforms to ensure seamless collaboration.

This scenario highlights the critical need for cross-border emergency management training. Effective response requires standardized protocols, digital training platforms, and simulations that enable responders to practice coordinated decision-making under pressure. The operation's complexity demonstrates the importance of continuous, realistic training programs that bridge national and organizational differences in crisis management.

The Erasmus+ project *TeleTraining* aims to develop a learning management system for providing these kinds of training exercises. The project will also introduce new ideas for handling the specific problems of emergency management training. This paper shows an experiment to highlight the feasibility of using AI in crisis management training.

The following section discusses the various forms of digital training and learning management support. Section III discusses the particularities of emergency management

training. Section VI proposes some extensions to existing learning management systems to handle these particularities. Section V elaborates on extending the learning management system with AI to create better training scenarios. Section VI discusses our findings, and Section VII concludes our paper.

II. ONLINE LEARNING

A traditional learning management system operates in an asynchronous manner where students get access to training material and can study at their own pace.

Online training can be asynchronous, synchronous, or blended [6]. A Learning Management System (LMS) supports these three types of learning.

A. Asynchronous learning

Asynchronous learning allows learners to access materials at their convenience without real-time interaction. It includes pre-recorded lectures, discussion forums, and LMS like Moodle and Coursera. Table I analyzes the Strengths, Weaknesses, Opportunities, and Threats (SWOT) of asynchronous learning [4].

TABLE I. SWOT ANALYSIS OF ASYNCHRONOUS LEARNING

Strengths High flexibility Scalable for large audiences Cost-effective Allows learners to progress at their own pace.	Weaknesses: Lack of immediate feedback Potential for learner isolation Requires strong self-motivation and discipline.
Opportunities: Integration of AI for personalized learning paths Expansion into underserved regions, Use of mobile learning applications.	Threats: Reduced engagement compared to live sessions Digital literacy gaps Risk of outdated content if not regularly updated.

B. Synchronous learning

Synchronous training involves real-time instruction through virtual platforms such as Zoom, Microsoft Teams, and Google Meet. It also includes live webinars, virtual classrooms, and Instructor-Led Training (ILT). Table II analyzes the SWOT of synchronous learning [4].

TABLE II. SWOT ANALYSIS OF SYNCHRONOUS LEARNING

Strengths Immediate feedback Interactive learning Increased engagement Real-time discussions	Weaknesses Requires stable internet connection Scheduling difficulties Potential technical issues Dependency on instructor availability.
Opportunities Expansion of global learning communities Integration of AI-driven engagement tools Real-time collaboration across geographies	Threats Cybersecurity risks Digital fatigue Potential technological barriers for learners with limited access.

C. Blended learning

Blended learning [6] combines online and in-person instruction to provide flexible and personalized learning experiences. This approach can boost student engagement and offer access to a wide range of resources. However, it also comes with challenges, including technological hurdles, increased workload for teachers, reduced direct social interaction, and the need for students to have strong self-discipline to succeed in the online parts.

III. EMERGENCY MANAGEMENT TRAINING

Teams are multidisciplinary, coming from the police, fire and rescue, health, and military. The participants are not first responders; they are not out in the field. They are coordinators and managers who need to allocate resources and make decisions.

Taber [7] proposed *active critical learning* as an approach to learning how to act and make decisions in dynamic, unpredictable situations where incidents without precedence escalate rapidly. Active critical learning based on realistic scenarios entails engaging in the learning process, a progression of difficulty, access to peers, experts, and resources, a multimodal design, multiple routing, feedback, working through errors, and debriefing.

The TeleTraining program focuses on incidents involving more than one nation, which introduces some additional challenges due to differences in work practices and culture.

In most cases, emergency management is hierarchical, with a coordinator and team members from different sectors coordinating their contributions. Figure 1 shows the different contributors coordinated by the Joint Rescue Coordination Centre (an example from Norway).



Figure 1. Coordination and Responders.

Emergency management differs from traditional management in several ways. First, the situation is often elusive, and decisions frequently need to be made based on uncertain information. Data is also high volume, high velocity, and in different formats (audio, video). Spatial data is of particular interest, visualized in a Geographic Information System (GIS).

Virtual teams may coordinate emergency management operations. Ad hoc teams will be formed for cross-border incidents. Team members will be in their parent organization, but teamwork will be done through online collaboration.

Berlin and Carlström [2] investigated collaboration among organizations at the accident scene. They found that collaboration is avoided due to uncertainty and lack of

incentives. Training together on scenarios is a way to reduce uncertainty and see the advantages of close collaborations.

IV. CUSTOMIZED LEARNING MANAGEMENT SYSTEM

The TeleTraining program will offer an emergency training course with participants from different groups of stakeholders.

The learning management system will provide study material and assignments for asynchronous studies. Study material includes video lectures, podcasts, textbooks, and research papers. Since synchronous training sessions are an essential part of online emergency training where collaboration is in focus, the learning management system needs to be extended by the possibility for one-to-one and group communication.

We also need a shared workspace where announcements, incoming messages, and tasks are displayed. Tasks should be triaged based on priority. This shared workspace will be the topic of another paper.

Finally, we propose extending the learning management system with an AI-based plug-in to provide operational support and create incoming alerts for the emergency management team. This is the focus of the rest of this paper. Figure 2 shows the customized learning management system.

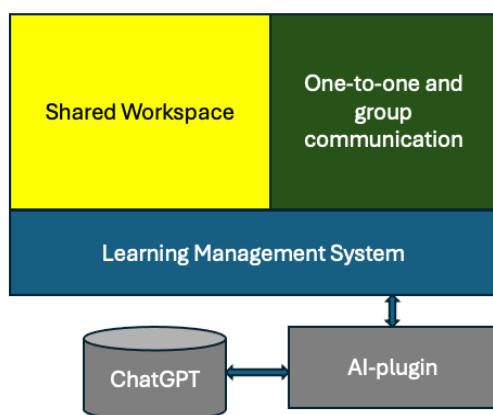


Figure 2. Customized Learning Management System.

A customized Learning Management System (LMS) is a platform uniquely designed to meet an organization's specific needs, differing from standard, pre-built LMS solutions. It offers functionalities, user experiences, branding, and scalability that match the organization's workflows and learning objectives. This customization often includes adding, modifying, or removing features, personalizing interfaces and learning paths, incorporating specific branding elements, ensuring seamless integration with other software, and addressing specialized training requirements. Ultimately, it creates a flexible and adaptable learning environment tailored to precise organizational needs.

V. INTEGRATION WITH AI TECHNOLOGY

Part of the training project involves conducting online synchronous training exercises. However, the participants

receive incomplete information about the incident. AI can help exercise planners improve the quality of simulations.

AI can be used to:

- Develop scenarios
- Provide situation reports and updates
- Create evaluation questions

ChatGPT was asked to create an initial scenario:

Please help create a scenario for emergency management training. A flight from Oslo to Copenhagen was forced to land on the water around 100 km northwest of Gothenburg. There are fresh breezes and rain, and the sea temperature is around 15 degrees Celsius. The extent of the damage is unknown. A Mayday signal was sent at 14:15. There is no more information at the present time.

Output from ChatGPT is shown in the frames with a grey background.

Emergency Management Training Scenario: Flight Emergency Water Landing

Scenario Overview:

At **14:15**, a **Mayday signal** was received from a commercial flight en route from **Oslo (OSL) to Copenhagen (CPH)**, indicating an emergency situation requiring an **immediate water landing**. The aircraft went down **approximately 100 km northwest of Gothenburg**, in open waters.

Weather conditions at the time of the incident:

- **Wind:** Fresh breezes (approx. 30–40 km/h)
- **Rain:** Moderate to heavy rainfall
- **Sea Temperature:** 15°C
- **Visibility:** Limited due to weather conditions

Initial Status:

- The extent of damage to the aircraft is **unknown**.
- The number of survivors is **unknown**.
- Communication with the flight **ceased** after the Mayday call.
- Search and Rescue (SAR) operations need to be **initiated immediately**.

ChatGPT provides the following emergency response activities:

Emergency Response Objectives:

Activate Search and Rescue (SAR) Operations:

- Dispatch **Coast Guard, SAR helicopters, and maritime rescue units** to the last known coordinates.
- Mobilize available **NATO or Swedish Air Force resources** if necessary.
- Deploy **civilian and military vessels** for search and recovery.

Assess and Prioritize Survival Factors:

- **Time-sensitive rescue:** The sea temperature (15°C) presents a **high risk of hypothermia** within **1-2 hours**.

- Strong winds and rainfall may **complicate rescue efforts**.
- **Life rafts & floating debris** may assist survivors—coordinate aerial surveillance to locate them.

Establish a Unified Command & Communication Strategy:

- Coordinate efforts between **Swedish, Norwegian, and Danish emergency services**.
- Use **emergency radio frequencies, satellite data, and air traffic control logs** to track the plane's last movements.
- Deploy **marine and aerial units with thermal imaging and sonar** to locate wreckage and survivors.

Medical & Logistical Preparedness:

- Hospitals in **Gothenburg, Oslo, and Copenhagen** placed on **high alert** for mass casualty incidents.
- **Evacuation procedures** for survivors—triage on ships and helicopters before hospital transport.
- Contingency plans for **handling fatalities and missing passengers**.

Media and Public Communication Strategy:

- **Prevent misinformation**—designate a spokesperson for coordinated press releases.
- Inform **families of passengers** and establish **emergency helplines**.
- Monitor **social media and aviation forums** for additional insights or passenger communications.

The training exercise will consist of four phases:

- Initial response
- Search and Recovery
- Operational Expansion
- Post-incident Management

A. Initial response

The initial response occurs within the first hour. The Mayday signal is received, the joint rescue coordination centers are alerted, search-and-rescue assets are mobilized in all three countries, cross-border communication between national coordinators is established, and the first reconnaissance flight is dispatched.

A participant asks: Where should the first reconnaissance aircraft be dispatched from? (Only show primary option)

Primary Dispatch Location:

Save Air Base (Säve Flygplats, ESGP, Gothenburg, Sweden)

Reason: Closest operational base with Swedish Coast Guard SAR aircraft, ensuring the fastest response time (~12–15 minutes to the crash site).

Another participant asked about the estimated time to get the aircraft airborne:

Immediate Standby (Best Case): 5–10 minutes

- If a Swedish Coast Guard Dash-8 Q300 is on standby and ready for rapid deployment.

Standard Readiness (Likely Case): 15–20 minutes

- Time for crew briefing, pre-flight checks, and taxiing.
- Assumes aircraft is fueled and prepared for emergency deployment.

Delayed Readiness (Worst Case): 30+ minutes

- If the aircraft is not immediately available or requires maintenance checks before takeoff.

B. Search and recovery

The search and recovery phase begins 1 to 4 hours after the incident. This phase includes identifying debris and establishing potential survivor locations. Search and rescue helicopters and vessels then move in to pick up victims. Survivors receive stabilizing treatment before being forwarded to the appropriate medical facility.

SITUATION REPORT – 14:40 CET

From: Swedish Coast Guard SAR Aircraft (Dash-8 Q300)

Callsign: SE-SAR 1

To: Joint Rescue Coordination Center (JRCC) Gothenburg

Current Status:

The aircraft reached the search area at 14:38 CET (~100 km NW of Gothenburg).

Weather conditions:

- Wind: Fresh breeze (~30–40 km/h)
- Rain: Moderate, reducing visibility
- Sea state: Rough, waves approx. 1.5–2 meters
- Sea temperature: 15°C

Observations:

- Debris field spotted: Approximately 2 NM south of last known coordinates.
- No sign of the main fuselage yet.
- Possible life rafts sighted: Infrared imaging detects multiple heat signatures in the water.
- No visible fire or smoke.

Next Actions:

- Conducting low-altitude pass for detailed assessment.
- Deploying emergency survival pods to survivors.
- Relaying coordinates to SAR helicopters & naval units.

Urgent: Survivors in water—immediate extraction required. Risk of hypothermia within 1-2 hours.

The accident with the location was confirmed. Additional search and rescue resources need to be activated.

A participant asked for additional search and rescue resources with a situation update.

SITUATION UPDATE – REQUEST FOR ADDITIONAL SAR RESOURCES – 14:45 CET

From: Joint Rescue Coordination Center (JRCC) Gothenburg

To: All SAR Units & Authorities Involved

Request for Additional SAR Assets:

1. Air Support (Urgent)

Swedish and Norwegian SAR Helicopters

Request immediate dispatch of 2–4 SAR helicopters from:

- Swedish Air Force (F7 Sätenäs) – Nearest base (~20 min ETA).
- Norwegian 330 Squadron (Sola Air Base) – (~40 min ETA).
- Danish EH101 SAR Helicopters (Skrydstrup) – If needed (~50 min ETA).

Task: Immediate extraction of survivors from water.

2. Maritime Response

- Rescue Vessels & Coast Guard Ships
- Deploy high-speed rescue boats from Gothenburg & nearby coastal stations.
- Request support from Swedish Navy (Karlskrona Base) for larger assets if needed.

3. Fixed-Wing Support

Additional Reconnaissance & Support Aircraft

- Request Swedish Air Force Gulfstream IV (SIGINT) to assist in debris tracking.
- Request Norwegian P-8 Poseidon maritime patrol aircraft (if available).

Immediate Actions:

- Directing first SAR helicopter to survivor location (~ETA 14:55).
- Marking survivor locations with emergency beacon drops.
- Relaying coordinates for coast guard and navy vessels.
- Critical: Survivors are exposed to 15°C water—risk of hypothermia within 1–2 hours. Immediate extraction is top priority.

C. Operational Expansion

After the search and recovery phase, a full-scale maritime and aerial search continues to identify additional survivors. Survivors are brought to appropriate medical facilities, and new resources arriving are used in search and rescue operations.

A participant asked for a situation report.

SITUATION REPORT – 16:15 CET

From: Joint Rescue Coordination Center (JRCC) Gothenburg
To: All SAR Units & Authorities Involved

1. Current Status of the Search and Rescue Operation

Incident: Commercial aircraft forced water landing ~100 km NW of Gothenburg.

Initial Mayday received: 14:15 CET.

First SAR aircraft on scene: 14:38 CET (Swedish Coast Guard Dash-8 Q300).

Weather conditions:

Winds: 35-40 km/h (fresh breeze)

Sea state: Rough, waves 2–2.5 meters

Rain: Intermittent, reducing visibility

Sea temperature: 15°C (risk of hypothermia within 1–2 hours)

2. Assets Deployed and Actions Taken

Airborne SAR Assets:

- Swedish Coast Guard Dash-8 Q300 (SE-SAR 1) – Overwatch & Coordination

- Swedish Air Force SAR Helicopters (2 UH-60 Black Hawks) – Ongoing Rescue
- Norwegian EH101 SAR Helicopter (Sola Air Base) – En Route (ETA 16:25)
- Danish EH101 SAR Helicopter (Skrydstrup) – En Route (ETA 16:40)
- Norwegian P-8 Poseidon (Maritime Patrol Aircraft) – Deployed for Wreckage Location

Maritime SAR Assets:

- Swedish Coast Guard Vessel (KBV 001 Poseidon) – En Route, ETA 17:00
- High-Speed Rescue Boats from Gothenburg – Arriving at Search Area
- Civilian Ships (Diverted Commercial Vessels in Vicinity) – Assisting SAR

3. Current Findings and Rescue Efforts

Debris Field Confirmed:

Located approx. 2–3 NM from last known aircraft position.

Floating wreckage includes aircraft parts, luggage, and life jackets.

Survivors Located & Rescued:

Confirmed Survivors: 23 individuals recovered (some in life rafts, others in open water).

Condition: Several cases of moderate hypothermia, no critical injuries reported yet.

First rescues by Swedish SAR helicopters – transported to Gothenburg hospitals.

Missing Persons:

Estimated ~70+ passengers and crew still unaccounted for.

Search extended to 10 NM radius from initial crash site.

No Sign of Main Wreckage (Fuselage):

Wreckage may have submerged.

Norwegian P-8 Poseidon conducting sonar scans for underwater debris.

4. Next Steps & Immediate Priorities

Continue Helicopter Rescue Operations

Priority: Recover survivors from life rafts and open water before hypothermia sets in.

Deploy More Naval Assets for Survivor Recovery & Wreckage Location

Swedish and Norwegian navy divers on standby for underwater search.

Civilian ships instructed to assist in search & survivor recovery. Expand Search Area

Given wind & current drift, survivors and debris may have moved further away.

Reconnaissance flights expanding search to a 15 NM radius.

Establish Emergency Response Centers in Gothenburg & Oslo

Coordination of survivor medical treatment & family notifications.

Swedish & Norwegian authorities managing public communication.

5. Key Challenges

Weather conditions deteriorating – rough sea impacting rescue ops. Limited daylight – sunset approaching (~17:30 CET), night SAR ops needed.

Time-sensitive – Risk of hypothermia increasing for those still in the water.

6. Summary

23 survivors rescued, 70+ still missing
Helicopter & naval rescues ongoing
No sign of main fuselage – P-8 Poseidon scanning for underwater wreckage
Search area expanding as wind & currents move debris and survivors
SAR operations remain in full emergency mode – Additional resources may be requested.

Later, a second participant asked for a situation update.

SITUATION REPORT – 18:15 CET

From: Joint Rescue Coordination Center (JRCC) Gothenburg
To: All SAR Units & Authorities Involved

1. Current Status of the Search and Rescue Operation

Incident: Commercial aircraft forced water landing ~100 km NW of Gothenburg.
Initial Mayday received: 14:15 CET.
First SAR aircraft on scene: 14:38 CET.
Weather conditions worsening:
Winds: 40–45 km/h (strong breeze)
Sea state: Rough, waves reaching 3 meters
Rain: Heavy, further reducing visibility
Sea temperature: 15°C (survival time in open water estimated at 1–2 hours)
**Nightfall approaching (~17:30 CET) – SAR operations now in low-visibility conditions.

2. SAR Assets Deployed and Actions Taken

Airborne SAR Assets:

- Swedish Coast Guard Dash-8 Q300 (SE-SAR 1) – Aerial Coordination (Returning to Base for Refuel at 18:30)
- Swedish Air Force UH-60 Black Hawks (2) – Rescuing Survivors
- Norwegian EH101 SAR Helicopter (Sola) – Recovering Survivors, Low on Fuel (RTB at 18:45)
- Danish EH101 SAR Helicopter (Skrydstrup) – Ongoing Search for Survivors
- Norwegian P-8 Poseidon (Maritime Patrol Aircraft) – Conducting Sonar Scan for Wreckage

Maritime SAR Assets:

- Swedish Coast Guard Vessel (KBV 001 Poseidon) – Arrived at Scene, Recovering Survivors
- High-Speed Rescue Boats (Gothenburg) – Multiple Survivors Recovered
- Civilian Ships (Diverted Vessels) – Assisting SAR, 2 Ships Actively Rescuing Survivors
- Swedish Navy Corvette (HMS Nyköping) – En Route, ETA 19:00 for Night Operations

3. Current Findings and Rescue Efforts

Debris Field Widened:
Floating wreckage spread over 5–6 NM radius, increasing search difficulty.
No large fuselage sections found.
Survivors Rescued (Confirmed):
41 individuals recovered (some in life rafts, others in open water).
Condition:
13 survivors suffering from severe hypothermia, airlifted to Gothenburg hospitals.

28 others in moderate or mild condition, receiving medical treatment on rescue vessels.
Survivor statements confirm aircraft broke apart upon impact with the water.
Missing Persons:
Estimated ~50+ passengers and crew still unaccounted for.
Increased risk of fatalities due to cold exposure.
No Sign of Main Wreckage (Fuselage):
Norwegian P-8 Poseidon continuing sonar scan – no submerged wreckage located yet.
Swedish Navy divers preparing for underwater search upon daylight.

4. Next Steps & Immediate Priorities

Continue Helicopter Rescue Operations Until Fuel Limits Require Rotation
Focus on recovering survivors from drifting rafts and open water.
Danish EH101 to take over aerial overwatch as Swedish/Norwegian helicopters rotate out.
Expand Maritime Search Efforts
Additional vessels arriving for overnight search & recovery.
Request possible assistance from Norwegian or Danish naval assets.
Nighttime Search Operations Initiated
Thermal imaging cameras being deployed.
Civilian ships equipped with spotlights assisting SAR.
Family Assistance & Crisis Response Centers Activated
Survivor medical triage ongoing in Gothenburg.
Swedish & Norwegian authorities coordinating passenger family notifications.

5. Key Challenges

- Weather conditions worsening – strong wind & high waves impacting SAR efforts.
- Nightfall reducing visibility – thermal imaging and sonar required.
- Cold exposure risk rising – immediate survivor recovery still top priority.
- SAR helicopters reaching fuel limits – rotations required to maintain coverage.

6. Summary

- 41 survivors rescued, 50+ still missing
 - Night SAR operations initiated
 - No sign of main fuselage – Norwegian P-8 continuing sonar scans
 - Search area expanded to ~10 NM as wind & currents push debris and survivors further
- SAR operations remain in full emergency mode – Additional assets may be requested.

D. Post-incident Management

During the first three phases, the focus is on life-saving efforts. When the life-saving efforts are finished, there are still things to be done. Survivors and non-survivors need to be recorded together with missing persons.

Aircraft recovery operations need to start (e.g., recovery of black boxes containing vital information about what happened. An incident investigation needs to be launched.

A participant asks for a post-incident report summary.

POST-INCIDENT REPORT SUMMARY

Incident: Emergency Water Landing – Commercial Flight from Oslo to Copenhagen

Date: [Insert Date]

Time of Mayday: 14:15 CET

Location: ~100 km NW of Gothenburg, Sweden (North Sea)

Weather Conditions:

Winds: 40–45 km/h (strong breeze)

Sea state: Rough (waves 2.5–3 meters)

Rain: Heavy at times, reducing visibility

Sea temperature: 15°C

1. Incident Overview

At 14:15 CET, air traffic control received a Mayday distress signal from a commercial aircraft en route from Oslo (OSL) to Copenhagen (CPH), reporting an emergency requiring an immediate water landing. Contact was lost shortly thereafter.

At 14:38 CET, a Swedish Coast Guard reconnaissance aircraft arrived at the scene and confirmed a debris field in the water approximately 2 NM from the last known coordinates. Multiple life rafts and individuals in the water were detected using thermal imaging.

2. Search and Rescue (SAR) Response

2.1 Deployed SAR Assets

- Airborne Assets:
Swedish Coast Guard Dash-8 Q300 – Aerial reconnaissance and coordination
- Swedish Air Force UH-60 Black Hawks (2) – Survivor extraction
- Norwegian EH101 SAR Helicopter (Sola Air Base) – Survivor extraction
- Danish EH101 SAR Helicopter (Skrydstrup Air Base) – Night SAR operations
- Norwegian P-8 Poseidon (Maritime Patrol Aircraft) – Sonar search for submerged wreckage

Maritime Assets:

- Swedish Coast Guard Vessel (KBV 001 Poseidon) – Survivor recovery and medical assistance
- Multiple high-speed rescue boats from Gothenburg and nearby stations
- Civilian vessels diverted from nearby shipping lanes
- Swedish Navy Corvette (HMS Nyköping) – Arrived for night SAR and debris recovery

3. Incident Outcome

3.1 Survivor Recovery

Total survivors recovered: 47

Medical conditions:

18 cases of severe hypothermia, requiring intensive care

29 cases of moderate to mild injuries, treated on scene or at local hospitals

Survivors reported that the aircraft broke apart on impact with the water.

3.2 Fatalities and Missing Persons

Confirmed fatalities: 19 (bodies recovered from the water)

Missing persons: 23 (presumed deceased, search efforts ended)

3.3 Wreckage & Investigation

No large fuselage sections located during initial SAR operations. Norwegian P-8 Poseidon sonar scans detected submerged wreckage at approximately 60 meters depth.

Swedish and Norwegian Navy dive teams scheduled for a detailed wreckage recovery operation.

Investigation agencies (Swedish SHK & European Aviation Authorities) coordinating recovery for black box retrieval and forensic examination.

4. Challenges Encountered

- Severe weather conditions (strong wind, rain, and high waves) delayed rescue efforts.
- Nightfall significantly reduced visibility, requiring thermal imaging and sonar assistance.
- SAR helicopters had to rotate frequently due to fuel limitations and poor weather conditions.
- Cold exposure risk was a critical factor—survival time in open water was estimated at 1–2 hours.

5. Recommendations for Future Response

- Enhanced Rapid Deployment of SAR Helicopters: Faster coordination with multiple agencies for immediate launch.
- Prepositioning of Maritime SAR Assets: Increase readiness of naval and coast guard vessels in high-risk areas.
- Cross-Border SAR Coordination Improvements: Improved protocols between Swedish, Norwegian, and Danish authorities.
- Weather-Adapted Search Techniques: Expand use of drift analysis modeling to predict survivor locations more accurately.
- Underwater Recovery Capabilities: Strengthen rapid deployment of navy dive teams and remotely operated vehicles (ROVs) for submerged wreckage.

6. Conclusion

The combined efforts of Swedish, Norwegian, and Danish SAR teams resulted in the successful rescue of 47 survivors.

However, 42 lives were lost due to the severity of the crash, weather conditions, and time required for rescue. Aviation and maritime authorities continue investigations to determine the cause of the incident and ensure improved emergency preparedness in the future.

Status: SAR operations concluded – Incident investigation ongoing.

VI. DISCUSSION

One objective of this paper was to investigate whether ChatGPT could deliver realistic content for scenarios and situation reports. We have presented one scenario but have also used it for other transnational incidents, including a cross-border wildfire, a solar storm, a pandemic, and a fire aboard a passenger ferry. In these cases, ChatGPT has helped create more realistic scenarios.

Experts have verified the answers from ChatGPT. The airborne and maritime SAR resources are correctly identified and located. The Norwegian got the P8 Poseidon naval patrol aircraft, and the Swedish used the Dash 8, as shown. The Swedish also use Black Hawk helicopters, and the KV

Poseidon is a real resource. Flight times and times to be airborne seem realistic. Estimations of cold exposure risk have been verified.

In some cases, ChatGPT was asked to elaborate on answers. For example, *why did you suggest Sola Air Base instead of Rygge Air Base* (which is closer)?

While **Rygge Air Base (ENRY)** is geographically closer to the search area than **Sola Air Base (ENZV)**, there are a few key reasons why Sola is prioritized for SAR helicopter dispatch:

1. Dedicated Search and Rescue (SAR) Squadron at Sola

- 330 Squadron of the Royal Norwegian Air Force operates Westland AW101 (EH101) SAR helicopters from Sola Air Base.
- These helicopters are specialized in long-range maritime rescue and are equipped with winches, medical capabilities, and thermal imaging systems, making them the best choice for this type of operation.

2. Rygge's Primary Focus is Military Operations

- Rygge (ENRY) is mainly a military base with F-35 fighter operations and limited SAR assets.
- There may be some SAR availability, but Sola remains the primary Norwegian SAR hub for maritime operations.

3. Faster Mobilization and Response Time

- Sola is already on high readiness for North Sea SAR operations.
- Sola's EH101s are better suited for rough weather conditions and long-range maritime missions.
- Despite a slightly longer flight time (~40 min from Sola vs. ~30 min from Rygge), Sola's helicopters are likely to be mission-ready faster.

Conclusion

Sola Air Base is the best choice for Norwegian SAR helicopter support due to its dedicated 330 Squadron, specialized aircraft, and high readiness for maritime rescues.

If additional assets are required, Rygge or Danish EH101s from Skrydstrup can be requested as backup.

VII. CONCLUSION

This paper has explored the potential of AI, specifically ChatGPT, in enhancing emergency management training through realistic scenario generation and adaptive situational updates. By integrating AI-driven simulations into a customized learning management system, training exercises can become more dynamic, data-driven, and reflective of real-world challenges. Our case study on a transnational air emergency demonstrated the capability of AI to generate coherent, contextually relevant crisis scenarios and support decision-making processes.

The AI tool can function both as a resource for the players and as an adversary controlled by the exercise leader. The player can provide adequate suggestions regarding status, risk assessments, and appropriate actions if the participants formulate the right questions. An AI opponent

can generate new challenges during an exercise, such as escalating crises, the need for reinforcement measures, and collaboration. A trained AI can also create an evaluation of the players' abilities, such as analysis, decision-making skills, and redundancy.

The results suggest that AI can significantly improve emergency training by facilitating cross-border coordination, refining decision-making processes, and enhancing real-time situational awareness. However, human oversight remains critical to validate AI-generated content and ensure alignment with established emergency protocols. Future work will focus on refining AI-assisted scenario development, integrating real-time data feeds, and further adapting AI tools for emergency response training across diverse crisis types.

ACKNOWLEDGMENTS

This research project is partially funded by the Erasmus+ Program of the European Union. Project 2024-1-NO01-KA220-HED-000253838 — TeleTraining in Crisis Management - The TTCM Exercise Model.

REFERENCES

- [1] M. Benfatah et al., "Assessing the efficacy of ChatGPT as a virtual patient in nursing simulation training: A study on nursing students' experience," *Teaching and Learning in Nursing*, vol. 19(3), 2024, pp. e486-e493, <https://doi.org/10.1016/j.teln.2024.02.005>.
- [2] J. M. Berlin and E. D. Carlström, "Why is collaboration minimised at the accident scene? A critical study of a hidden phenomenon," *Disaster Prevention and Management: An International Journal*, vol. 20(2), 2011, pp. 159-171.
- [3] E. Karinshak, "Simulations in human-AI crisis management systems," *Journal of Contingencies and Crisis Management*, vol. 32, 2023, e12525, <https://doi.org/10.1111/1468-5973.12525>.
- [4] C. M. Lee, "Descriptive SWOT Analysis about Online Learning," SSRN, [Online]. Available from <https://ssrn.com/abstract=4855003> 2025.03.01
- [5] K. Lie Eide, I. Lund-Kordahl, and B. Tallak Bakken, "Perspective Chapter: How Artificial Intelligence (AI) Fundamentally Changes Crisis Management Training and Exercises," in *Crisis Management Dynamics - Strategies, Challenges, and Best Practices*, M. Sarfraz, Ed. IntechOpen, 2025. DOI: 10.5772/intechopen.1008950
- [6] B. L. Moorhouse and K. M. Wong, "Blending asynchronous and synchronous digital technologies and instructional approaches to facilitate remote learning," *Journal of Computer Education*, vol. 9, 2022, pp. 51-70, <https://doi.org/10.1007/s40692-021-00195-8>
- [7] N. Taber, "Emergency response: Elearning for paramedics and firefighters," *Simulation & Gaming*, vol. 39(4), 2008, pp. 515-527.
- [8] J. Vaughn, S. H. Ford, M. Scott, C. Jones, and A. Lewinski, "Enhancing Healthcare Education: Leveraging ChatGPT for Innovative Simulation Scenarios," *Clinical Simulation in Nursing*, vol. 87, 2024, 101487, <https://doi.org/10.1016/j.ecns.2023.101487>.