Multi-modal Learning in Sensing Systems and Machine Learning for IOT: Enhancing Gas Leakage Detection in Industrial and Residential Environments

Julio Contreras
Rutgers University-New Brunswick
Perth Amboy, NJ
JRC397@Scarletmail.Rutgers.edu

Jake Giannotto
Rutgers University-New Brunswick
Manalapan, NJ
JG1631@Scarletmail.Rutgers.edu

***Abstract*—This study presents an approach, to gas detection using a combination of imaging and semiconductor gas sensor data. The unique aspect of this research lies in its network design, which integrates a Convolutional Neural Network (CNN) a Long Short Term Memory (LSTM) network and a Multi-Layer Perceptron (MLP). The CNN is structured with layers consisting of 32 filters sized at 3x3 along with MaxPooling layers that effectively extract spatial features from the thermal images. The LSTM, composed of 32 units, handles the time series sensor data; capturing temporal dynamics. Acting as a classifier, the MLP consists of layers with 128 and 64 neurons respectively. This neural network harnesses the spatial feature extraction capabilities of the CNN, the temporal data handling abilities of the LSTM, and the classification prowess of the MLP to create a model for gas detection.**

I. Introduction

Multimodal learning in sensing systems and machine learning for the Internet of Things (IoT) represents an innovative approach that integrates various sensory inputs, such as audio, video, text, and other forms of data, to enable more comprehensive and accurate decision-making processes. In essence, it leverages multiple sources of information to enhance the capabilities of IoT devices and systems. These systems consist of machine learning techniques to analyze and interpret the diverse data streams, allowing them to perceive, comprehend, and respond to complex real-world situations. For instance, in smart environments, such as smart homes or cities, multimodal sensors gather data from different sources like cameras, microphones, temperature sensors, and more, while machine learning algorithms process this collection of data to recognize patterns, detect anomalies, and make predictions. This fusion of modalities not only enhances the information gathered but also improves the system's ability to adapt and make informed decisions in uncertain environments, advancing the potential for more efficient and responsive IoT applications across various systems. Understanding multimodal learning in sensing systems is crucial for future IoT innovations, offering tremendous opportunities to revolutionize industries, enhance user experiences, and solve complex real-world problems. While technology has significantly transformed various aspects of the chemical industry, the persistent threat of industrial hazards, especially related to gas leakages causing environmental harm, remains a pressing concern. These leaks pose serious risks such as explosions, emissions, fires, and environmental pollution, highlighting the urgent need for reliable sensing systems capable of quick and accurate detection [1], [2]. An article presented in the media revealed that burning wood, biomass, and dung led to 326,000 of the estimated 645,000 premature deaths from outdoor air pollution, which constitutes about 50% of the total deaths due to outdoor pollution [3]. Additionally, activities such as improper residential waste disposal increase the dangers of hazardous gas emissions, leading to air pollution and potential health risks. Despite rigorous pre-installation tests, gas leakage incidents continue due to the limitations of current sensors that struggle to effectively identify gas mixtures. Manual identification using chemical tools is not feasible and unsafe, particularly in situations with poor visibility, like during smoke leakages. Thus, there's a crucial need for an innovative approach: an automatic, intelligent gas identification system that can efficiently recognize various gas compositions during leakages. A system like this not only reduces risks to human life but also protects industrial assets. This project aims to explore multimodal learning in sensing systems and machine learning for the Internet of Things (IoT), specifically focusing on improving gas leakage detection in industrial and residential settings. By enhancing advancements in sensing technologies and machine learning algorithms, this research intends to overcome the limitations of existing gas detection mechanisms. The primary research question we aim to answer is: How can multimodal learning techniques integrated with machine learning algorithms enable accurate and rapid gas identification in scenarios involving complex gas mixtures? The ability to efficiently and accurately identify gas compositions in real-time not only saves lives but also prevents environmental disasters and protects critical industrial infrastructures. The following sections of this report will outline the project's scope, methodologies, related work, experimental setups, results, conclusions, and future work. Through a deep exploration of multimodal learning techniques, machine learning applications, and IoT advancements, this paper aims to contribute to a fundamental shift in gas leakage detection systems, promoting safer environments and more efficient industrial operations.

II. Related Work

The research on gas leakage detection covers a wide range of methods and technologies aimed at accurately identifying gas compositions. Recently, there have been advancements in using Internet of Things (IoT) systems with inexpensive sensors to detect gases in different places. However, these systems are not very effective because the sensors they use are not very sensitive, especially when dealing with mixtures of gases. Some methods like Colorimetric Tape [4] and Gas Chromatography [5] have been important in finding specific gas amounts in mixed environments. Also, different machine learning methods, like least square-based approaches and Deep Neural Networks, have been studied a lot for gas detection [6]. But these methods mostly depend on data from many gas-detecting sensors that are sensitive to different gases, which makes it hard to detect complex gas mixtures accurately. However, relying only on gas sensor-based methods has limitations. Sometimes, gas levels in the air are so low that standard sensor arrays can't detect certain gases. This makes it hard to get accurate results. Also, using cheap sensors might make measurements less precise. To deal with these issues, thermal imaging cameras have been used to find temperature changes caused by gas leaks, which helps identify leaks from far away. Having accurate training data is really important as artificial intelligence and data analytics are used more. But current datasets for gas detection mainly rely on sensor array data and lack important multimodal information needed to fully identify and classify gases. Detecting gas leaks, especially in places like industries and homes, is very important because these leaks can be dangerous. They can cause accidents like explosions, fires, and air pollution. Problems with sensor-based detection systems make these risks worse, especially in situations where it's unsafe for humans to intervene due to things like smoke or safety concerns. Some methods like Colorimetric Tape and Gas Chromatography work well for specific gas detection. Artificial intelligence methods like machine learning and neural networks have shown promise in classifying gases and estimating their amounts. But there are still challenges in accurately identifying complex gas mixtures and reducing false results from using single-mode sensing methods. A possible solution to these issues is to use a new approach that combines different kinds of data using deep learning frameworks such as CNNs, LSTMs, MLPs, SVMs, etc. [7], [8]. By bringing together data from various gas sensors and thermal imaging cameras, this method aims to improve the accuracy and reliability of gas detection and classification, especially for finding specific gases in mixed environments. Using different data types helps to get a more complete analysis and reduces false results that might come from individual sensor limitations.

TABLE I
Sensitive Gas(s) for Each Sensor

Sensor	Sensitive Gas
MQ2	LPG, Butane, Methane, Smoke
MQ3	Smoke, Ethanol, Alcohol
MQ5	LPG, Natural Gas
MQ6	LPG, Butane
MQ7	Carbon Monoxide
MQ8	Hydrogen
MQ135	Air Quality (Smoke, Benzene)

III. Dataset Description

The dataset collection involves two primary modalities: gas sensors and a thermal camera.

A. *Gas Sensors:*

Gas sensors used are Metal-Oxide Semiconductor (MQ) technology-based. These sensors identify the presence of specific gasses by producing electrical signals. Each sensor responds to various gasses (Table I). The gas sources used for data collection were Park Avenue Deodorants (containing 95% alcohol) as perfume and incense sticks to produce smoke, generating a combination of gasses including Carbon Monoxide, Carbon Dioxide, Nitrogen Dioxide, Sulfur Dioxide, and others in small quantities.

Biases/Considerations:
Each gas sensor is designed to detect specific gasses, which can lead to sensor-specific biases towards those gasses and missing others not within their sensitivity range. Additionally, the process of gas-generation, the manner in which the gasses were released, whether it be spraying perfume or burning incense sticks, it might affect their composition and detection by sensors. There were efforts made to hold the gas discharge consistent throughout the experiment to avoid conflicts and ensure consistent data collection.

B. *Thermal Camera:*

A Seek Compact Thermal Imaging Camera was used to capture thermal images. It detects temperature variations based on infrared light. The camera creates RGB images based on temperature configurations, and each pixel serves as an infrared temperature sensor.

Biases/Considerations: The camera's temperature measurement ranges from -40 to 330 °C, potentially limiting its ability to capture extreme temperatures beyond this range. Additionally, it operates at a frame rate of less than 9 Hz, which will make it difficult to efficiently capture rapid changes in temperature.

TABLE II
STATISTICAL ANALYSIS OF SENSOR DATA

	MQ2	MQ3	MQ5	MQ6	MQ7	MQ8	MQ135
Mean	677.6	462	404.6	399.8	566	542.5	416.7
Std	92.91	70.28	55.67	45.09	83.13	151.02	76.68
Min	502	337	291	311	361	220	275
$25^{th}\%$	591	405	366	366	524	447	354
$50^{th}\%$	701	486	400	393	576	576	437
$75^{th}\%$	756	529	443	426	629	642	473
Max	824	543	596	524	796	794	589

C. Dataset Description, Collection, and Biases:

The dataset comprises 6400 samples divided into four classes: perfume, smoke, a mixture of perfume and smoke, and a neutral environment (no gas). Each class contains 1600 samples. The data of the sensors were collected manually by the authors of the MDPI paper [9]. They used Park Avenue Deodorant as the source of perfume which contains 95% alcohol and incense sticks for the source of smoke. During the experiment, the gas sensors were spread out 1mm apart. The data was logged at a frequency of 2 seconds continuously for 90 minutes. The gas was sprayed into the environment at different intervals over the 90 minutes: 15s for the first 30mins, 30s for the next 30mins, and 45s for the last 30mins. Those who conducted the experiment took careful consideration into spraying the gasses to maintain uniformity and to spray a constant amount of gas every time. The only data preprocessing that needed to be done was re-scaling and resizing for the thermal images. There were also no challenges or limitations while using this dataset because of the preprocessing that was already performed by the owner of the dataset. Statistical analysis of the sensor readings are presented in Table II to showcase data variations.

Biases/Considerations:
The challenges in this study stem from various factors. Firstly, the issue of class imbalance is notable, as an equal distribution of samples across classes do not reflect real-world scenarios where specific gasses or combinations may occur more frequently. Additionally, the experimental environment, although controlled for gas dispersal, might not precisely emulate real-world conditions where gases disperse differently based on varying environmental factors. Moreover, the dataset's focus primarily on perfume and smoke raises concerns about its broader applicability, potentially limiting its relevance to other gas types or environmental conditions not explicitly represented within this study's scope. These limitations collectively highlight the need for further exploration and consideration when extrapolating findings beyond the specific confines of this dataset.The multimodal dataset, comprising both gas sensor readings and thermal imaging data, provides a richer context for addressing various questions related to environmental monitoring, gas detection, and pattern recognition. Some potential questions that this multimodal dataset could help answer are discussed next.

1) **Gas Identification and Classification:**
   Q: Can we accurately identify and classify different gasses or gas combinations?

   Multimodal Data Context: By combining gas sensor readings from different sensors (each responsive to specific gasses) and thermal images capturing temperature variations caused by gas emissions, a comprehensive dataset can help in creating models to distinguish between gasses. The combination of sensor data and thermal signatures might enhance the accuracy of gas identification compared to using either modality alone.

2) **Environmental Monitoring and Anomaly Detection:**
   Q: How can we monitor and detect anomalies in an environment concerning gas emissions or thermal changes?

   Multimodal Data Context: Integrating gas sensor data with thermal imaging allows for a more holistic view of environmental changes. Patterns in gas concentrations detected by sensors combined with corresponding temperature alterations in thermal images might enable the detection of abnormal events or deviations from expected environmental conditions.

3) **Fusion of Modalities [10] for Improved Detection:**
   Q: Can combining gas sensor data with thermal imaging enhance the accuracy of gas leak detection or environmental monitoring?

   Multimodal Data Context: The fusion of gas sensor data and thermal images allows for a more comprehensive analysis. Certain gasses might not be easily detected by sensors alone, but their thermal signatures could provide additional clues. By integrating both modalities, the dataset enables the development of models that leverage the strengths of each modality, potentially leading to more accurate and reliable detection systems.

4) **Real-time Monitoring and Response Systems:**
   Q: How can this multimodal data assist in creating real-time monitoring systems for immediate responses to gas-related incidents?

   Multimodal Data Context: The dataset collected at various time intervals and with different gas dispersion rates can aid in building predictive models. These models, incorporating both sensor readings and thermal imaging data, could be utilized in real-time systems to predict and respond promptly to changes in gas concentrations or temperature anomalies.

5) **Environmental Impact Assessment:**
   Q: How do different gasses or gas combinations impact the thermal environment?

Multimodal Data Context: By analyzing the correlation between gas concentrations detected by sensors and the corresponding thermal changes captured by the camera, it becomes possible to understand the thermal effects of different gasses. This understanding can be crucial in assessing the environmental impact of specific gas emissions.

IV. Methodology

Data Preparation

Our dataset, comprising 6400 samples, is categorized into four distinct classes: No Gas, Gas, Mixture, and Perfume. Each sample in the dataset encapsulates dual modalities of data: thermal imagery and readings from an array of seven semiconductor gas sensors. The thermal images provide a visual representation of the gases, while the sensor readings offer quantitative chemical analysis. Before feeding into the model, these images are resized to a uniform dimension of 240x320 pixels and normalized to ensure pixel values lie within a standardized range. The sensor data, on the other hand, undergoes a standardization process using the Standard-Scaler from scikit-learn, scaling each feature to have zero mean and unit variance, crucial for maintaining consistency and improving model performance. It also undergoes one-hot encoding to derive labels.

Model Architecture

Convolutional Neural Network (CNN)

The CNN is the cornerstone of our architecture for processing thermal images. It consists of layers designed to automatically and adaptively learn spatial hierarchies of features from the images. The initial layer is a Conv2D layer with 32 filters of a 3x3 kernel size, capturing essential features such as edges and textures. Following this, a 2x2 MaxPooling layer is employed to reduce the spatial dimensions of the output, thus diminishing the computational load and potential overfitting. The activation function throughout the CNN is the ReLU (Rectified Linear Unit), instrumental in introducing non-linearity to the network, enabling it to learn more complex patterns in the data. The final output of the CNN is flattened, transforming the two-dimensional feature maps into a one-dimensional feature vector, facilitating the integration with subsequent network components.

Long Short Term Memory (LSTM) Network

The LSTM network is tailored to handle the time-series data from the gas sensors. Its architecture allows it to capture temporal dependencies and patterns that might be indicative of specific gas types or concentrations. Our LSTM layer comprises 32 units and processes the sequential sensor data, which is crucial for understanding the temporal dynamics of gas detection. The ability of LSTM to remember long-term dependencies makes it exceptionally suited for our application, where sensor readings over time provide vital clues for accurate gas classification.

Multi Layer Perceptron (MLP)

The MLP acts as the final classification stage of our model, integrating the outputs from both the CNN and LSTM. It consists of two dense layers: the first with 128 neurons and the second with 64 neurons, both employing the ReLU activation function for non-linear transformation. The output layer uses a softmax activation function, which is ideal for multi-class classification problems like ours. This layer outputs a probability distribution over the four gas categories, providing both the predicted class and a confidence level for the prediction. The MLP's role is pivotal in combining and interpreting the features extracted from both image and sensor data to make accurate and reliable gas predictions.

Training Process

The training of the model is conducted on an 80/20 split of the dataset, ensuring a robust evaluation on unseen data. The loss function employed is categorical cross-entropy, which is particularly effective for multi-class classification problems. We use the Adam optimizer, known for its efficiency in computing adaptive learning rates for each parameter. Over 10 epochs, the model learns to refine its weights and biases to minimize the loss function, thereby enhancing its predictive accuracy.

Evaluation Metrics

The primary metric for evaluating our model is its classification accuracy on the test set. The combination of algorithims also increases accuracy [11]. This metric provides a direct measure of the model's ability to correctly classify the gases. Alongside accuracy, we delve into the model's interpretability by visualizing the activation maps from the CNN. These maps offer insights into what features the model is focusing on when making predictions. Additionally, we generate pairplot visualizations of the sensor data. These plots elucidate the relationships and distributions within the sensor readings, providing a deeper understanding of the data's structure and the model's decision-making process. Such visualizations not only affirm the model's performance but also enhance our understanding of its internal mechanics.

V. Results

The results section presents a visual narrative of our model's capabilities, starting from data distribution to detailed predictions on test samples. The following visualizations capture the essence of our model's analytical journey, providing insights into its predictive prowess and the underlying data patterns.

Gas Distribution Visualization

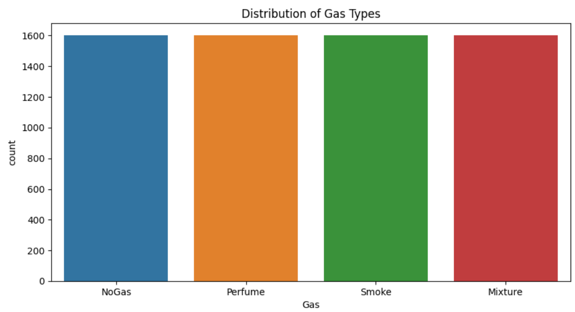

Fig. 1. Distribution of gas types present in the data sample

Figure 1 illustrates the frequency of each gas type within our dataset. It provides a fundamental understanding of the dataset's composition, which is crucial for evaluating the model's performance against varied gas occurrences.

Sensor Sensitivity Analysis

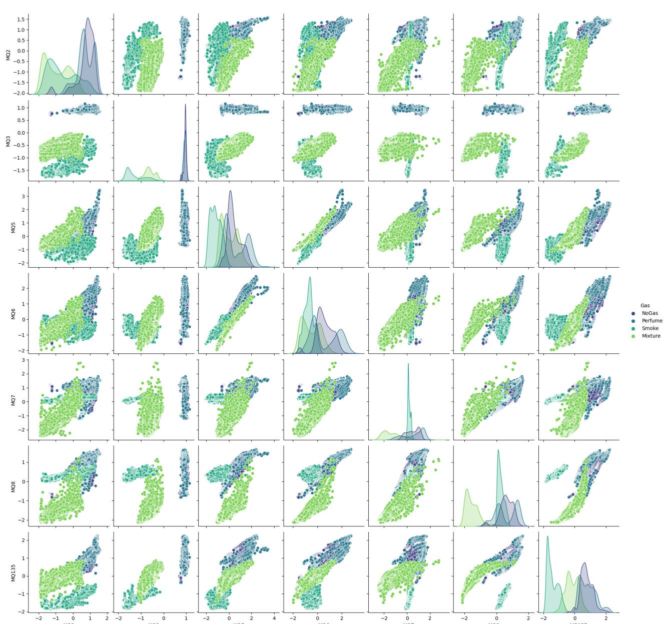

Fig. 2. Pairplots comparing the sensitivity of 2 different sensors to each gas category; the diagonal line places the same sensor against itself to visualize its specific sensitivity

Figure 2 delineates the relationship between different sensor readings, contrasting their responses to various gas types. The diagonal histograms offer a glimpse into each sensor's sensitivity, while the off-diagonal scatter plots reveal the interplay and correlation between sensor pairs.

Convolutional Neural Network Feature Extraction

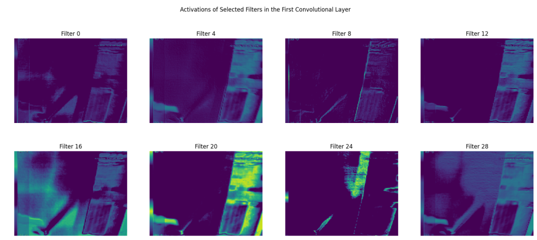

Fig. 3. Visualization of the CNN filtering process; rotating sensitivity to different thermal ranges for the same image

In Figure 3, we visualize the feature maps activated within the CNN when processing a thermal image. This provides an interpretative view of how the network discerns various thermal features, which contribute to its gas classification decisions.

Model Performance Metrics

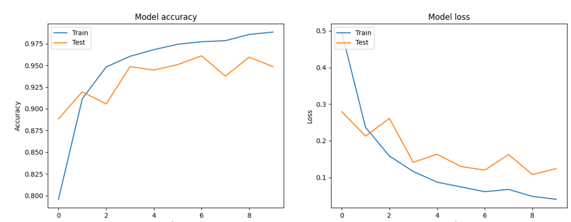

Fig. 4. Model loss across training epochs, reflecting the learning progress and convergence

Figure 4 tracks the model's learning curve, displaying the loss metric across training epochs. A descending curve indicates the model's improving accuracy and its ability to generalize from the training data. Our model ended up with an accuracy of 96.48% compared to the previous model's 96%.

Predictive Outcomes on Test Samples

The subsequent figures represent the model's predictions on randomly selected test images, showcasing the efficacy of the multi-modal neural network in classifying different gas types.

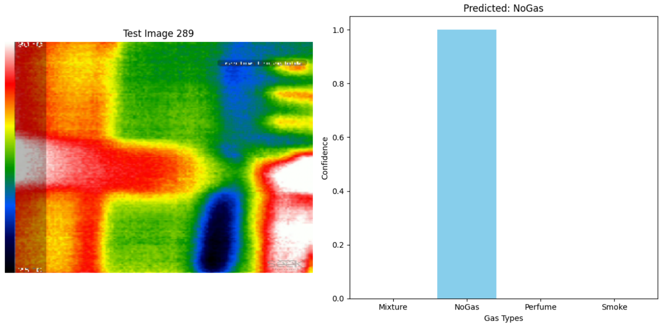

Fig. 5. Random Sample: 289; prediction: no gas

Figure 5 depicts the model's high confidence in predicting the absence of gas, aligning with the thermal image's lack of significant temperature variations typically associated with gas presence.

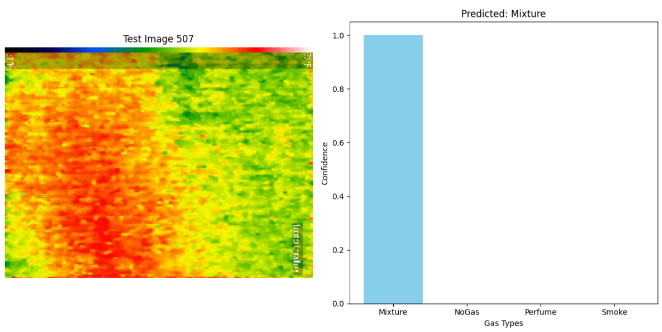

Fig. 6. Random Sample:507; prediction: mixture

In Figure 6, the model identifies a mixture of gases, reflecting the varied thermal signatures and corroborating with sensor data suggesting a blend of chemical constituents.

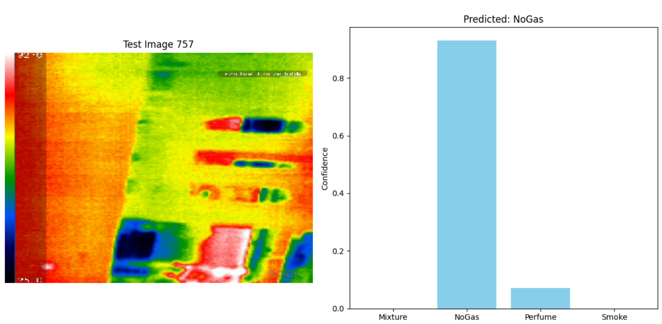

Fig. 7. Random Sample: 757; prediction: predominantly no gas, with a little perfume

Figure 7 reveals a subtler prediction where the model detects a trace of perfume amidst a largely gas-free environment, a testament to the model's sensitivity and nuanced detection ability.

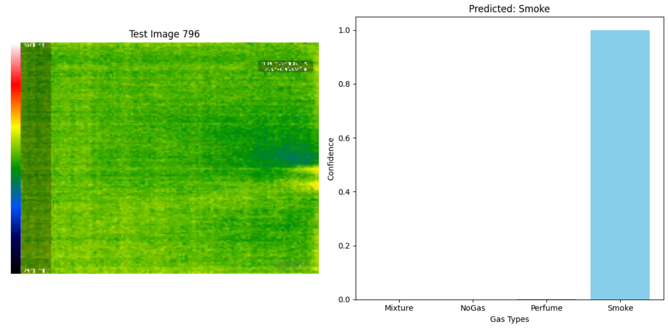

Fig. 8. Random Sample: 796; prediction: smoke

As shown in Figure 8, the model's prediction of smoke is corroborated by the distinct thermal anomalies associated with smoke's thermal footprint.

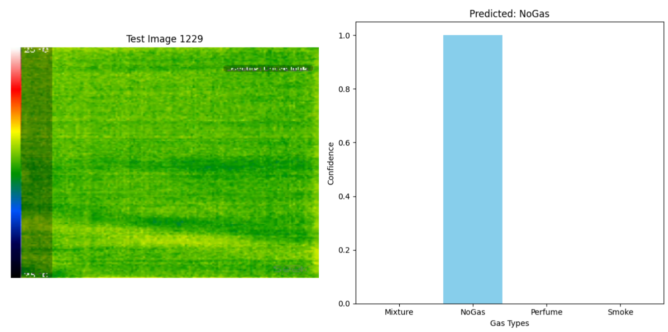

Fig. 9. Random Sample: 1229; prediction: no gas

Lastly, Figure 9 indicates the model's consistent performance in correctly identifying gas-free scenarios, further affirming its reliability. These visualizations collectively underscore the model's adeptness at navigating through the complexity of gas detection and classification, highlighting its potential as a robust tool for environmental monitoring and safety applications.

VI. Conclusion and Future Work

In conclusion, this study showcases the effectiveness of combining CNN and LSTM networks with an MLP classifier for gas detection and classification. The remarkable accuracy and nuanced comprehension of data indicate an advancement compared to traditional gas detection methods.

Regarding future work there are many directions to consider:

1) Increasing Dataset Diversity: Including a wider range of real-world data to enhance the model's robustness.
2) Real-time Processing: Adapting the model for analyzing data in real-time for applications.
3) Exploring Modalities: Investigating integrating different types of sensors to enhance detection accuracy further.

This research serves as a foundation, for developing dependable gas detection systems, which have the potential to revolutionize safety protocols across various industrial and environmental settings.

Poster and Video Demo Link:

Poster: Final IOT Poster
Video Demo: Demo Link

Acknowledgment

The authors would like to express their profound gratitude to the researchers who meticulously compiled and made available the dataset upon which this study is predicated. The article "Multimodal Machine Learning for Automated IOT Gas Sensor Drift Compensation" published in Sensors by MDPI has been an invaluable resource, providing both a robust foundation and inspiration for our research. We extend our sincere appreciation to the authors: Parag Narkhede, Rahee Walambe, Shruti Mandaokar, Shruti Mandaokar, Ketan Kotecha, George Ghinea; for their pioneering work that has significantly contributed to the advancement of gas detection technologies.

We are especially indebted to Professor Jorge Ortiz for his invaluable counsel and steadfast support throughout this project; and issue him a heartfelt thank you. His profound expertise in the realm of machine learning applications within the Internet of Things (IoT) has been fundamental to the fruition of our research. His recommendation to integrate a Multi-Layer Perceptron (MLP) into our neural network framework was a transformative suggestion that significantly bolstered our model's efficacy. We are deeply grateful for the privilege of engaging in this formidable challenge under his tutelage and for the rich educational experience afforded by his machine learning IoT course.

We are also indebted to the dedicated team members who invested countless hours in data collection, experimentation, and the drafting of this paper. Their commitment to the project's goals and their meticulous attention to detail have been the cornerstone of this paper's realization. It is through their collective effort and collaborative spirit that this research was brought to fruition.

Lastly, our thanks extend to Rutgers University for providing the resources and environment conducive to such scholarly pursuits. The unwavering support of the academic and technical staff has been a constant source of encouragement and has greatly facilitated our research endeavors.

References

[1] Trivedi, P.; Purohit, D.; Soju, A.; Tiwari, R.R. Major industrial disasters in India An ofcial newsletter of ENVIS-NIOH, Oct-Dec 2014; Volume 9, No 4. Available online: http://niohenvis.nic.in/newsletters/vol9no4Indian%20Industrial%20Disasters.pdf

[2] Zhou, Yi, et al. "Research on fire and explosion accidents of oil depots." 3rd International Conference on Applied Engineering, April 22-25, 2016, Wuhan, China. Vol. 51. AIDIC-associazione italiana di ingegneria chimica, 2016.

[3] Mudur, G.S. Lakhs of early deaths tied to home emissions. Telegraph India Online 2015. Published online on 17 September 2015. Available online: https://www.telegraphindia.com/india/lakhs-of-early-deaths-tied-to-home-emissions/cid/1513045

[4] MDC Systems Inc. Detection Methods. Online Resource. Available online: https://mdcsystemsinc.com/detection-methods/

[5] Wang, T.; Wang, X.; Hong, M. Gas leak location detection based on data fusion with time difference of arrival and energy decay using an ultrasonic sensor array. Sensors 2018, 18, 2985.

[6] Yin, X.; Zhang, L.; Tian, F.; Zhang, D. Temperature modulated gas sensing E-nose system for low-cost and fast detection. IEEE Sens. J. 2015, 16, 464–474.

[7] Pan, Xiaofang, et al. "A fast and robust gas recognition algorithm based on hybrid convolutional and recurrent neural network." Ieee Access 7 (2019): 100954-100963.

[8] Liu, Q.; Hu, X.; Ye, M.; Cheng, X.; Li, F. Gas recognition under sensor drift by using deep learning. Int. J. Intell. Syst. 2015, 30, 907–922.

[9] Narkhede, P.; Walambe, R.; Mandaokar, S.; Chandel, P.; Kotecha, K.; Ghinea, G. Gas Detection and Identification Using Multimodal Artificial Intelligence Based Sensor Fusion. Appl. Syst. Innov. 2021, 4, 3. https://doi.org/10.3390/asi4010003

[10] Elmenreich, W. A review on system architectures for sensor fusion applications. In IFIP International Workshop on Software Technolgies for Embedded and Ubiquitous Systems; Springer: Berlin/Heidelberg, Germany, 2007; pp. 547–559.

[11] Luo, Yuan, et al. "Classification of data from electronic nose using gradient tree boosting algorithm." Sensors 17.10 (2017): 2376.