DOI: https://doi.org/10.62985/j.huit_ojs.vol26.no2E.420

A LIGHTWEIGHT MACHINE LEARNING APPROACH FOR ALCOHOL CONCENTRATION ESTIMATION USING AN ESP32 AND MEMS GAS SENSOR

Minh-Hai Le[1,*], Duc-Toan Nguyen[2], Huu-Phuc Dang[1], Song-Toan Tran[1], Phuong-Thao Cao[1]

[1]*College of Engineering and Technology, Tra Vinh University, Vietnam*

[2]*Faculty of Agriculture and Aquaculture, Tra Vinh University, Vietnam*

*Email: *lmhai@tvu.edu.vn*

Received: 12 July 2025; Revised: 12 March 2026; Accepted: 17 April 2026

ABSTRACT

This paper presents a low-cost alcohol concentration measurement system based on an ESP32 microcontroller and a GM-302B MEMS alcohol gas sensor. The sensor output is sampled in analog mode to capture the time-domain response to ethanol vapor. From this response, a set of six features is extracted, including peak value, peak time, response slope, area under the curve, ambient temperature, and humidity. Alcohol concentration is estimated using a lightweight regression model implemented as a small fully connected neural network with two hidden layers. The trained model is converted to TensorFlow Lite for Microcontrollers and embedded directly into the ESP32 firmware using a TinyML framework, enabling on-device real-time inference without external computation. Experimental results indicate that the proposed system provides a practical and accessible solution for fermentation monitoring and educational research, offering an effective alternative to conventional alcohol measurement instruments.

Keywords: Lightweight machine learning, MEMS, TinyML, alcohol concentration estimation.

1. INTRODUCTION

Accurate measurement of alcohol concentration plays a crucial role in various applications, including fermentation monitoring, beverage quality control, safety regulation, and educational experimentation [1]. Traditional methods for determining alcohol content, such as distillation followed by densitometry, refractometry, and gas chromatography, provide high accuracy but typically require bulky equipment, laboratory infrastructure, trained personnel, and relatively long processing times [2]. These limitations restrict their applicability in field monitoring, real-time process control, and low-cost educational environments.

Recent advances in micro-electro-mechanical systems (MEMS) gas sensors and low-power microcontrollers have enabled the development of compact and inexpensive sensing platforms [3]. In particular, alcohol-sensitive semiconductor sensors, such as the GM-302B, offer the ability to detect ethanol vapor through changes in electrical resistance. These sensors are low-cost and easy to integrate; however, their practical use is challenged by nonlinear response characteristics, sensitivity to temperature and humidity, sensor drift, and cross-sensitivity to other volatile organic compounds commonly present in fermented beverages [4], [5].

Conventional signal processing approaches typically rely on steady-state sensor readings or simple threshold-based detection. Such approaches often fail to capture the rich dynamic behavior of the sensor response, which contains valuable information related to adsorption, reaction, and desorption processes occurring at the sensor surface [6]. Moreover, static measurements are particularly vulnerable to environmental variations and interfering compounds, leading to reduced accuracy and poor generalization across different operating conditions.

In parallel, the emergence of Tiny Machine Learning (TinyML) has made it possible to deploy data-driven models directly on resource-constrained embedded platforms [7]. By combining feature extraction, lightweight neural networks, and on-device inference, TinyML enables intelligent sensing systems that can adapt to nonlinear sensor behavior while maintaining low computational and energy overhead [8]. This paradigm offers a promising pathway for enhancing the performance of low-cost chemical sensing systems without requiring cloud connectivity or external computation.

Despite this potential, few studies have explored the integration of time-domain feature extraction with TinyML for alcohol concentration estimation in embedded systems. In particular, there remains a lack of work focusing on the exploitation of sensor dynamic response patterns, rather than static values, to improve robustness against sensor drift and cross-sensitivity in real-world beverage monitoring scenarios. This paper addresses this gap by proposing a low-cost alcohol concentration measurement system based on an ESP32 microcontroller and a GM-302B MEMS alcohol gas sensor. Instead of relying on steady-state measurements, the system captures the temporal response of the sensor and extracts a compact set of descriptive features, including peak response, peak time, response slope, area under the response curve, temperature, and humidity. These features are mapped to alcohol concentration using a lightweight neural network regression model, which is trained offline and then embedded directly into the microcontroller using TensorFlow Lite for Microcontrollers [9]. The main contributions of this work are as follows:

- A time-domain feature-based framework for alcohol concentration estimation using a low-cost MEMS gas sensor.
- The design and deployment of a lightweight regression neural network suitable for real-time inference on an ESP32 microcontroller.
- An embedded TinyML implementation enabling fully standalone operation without cloud or external computation.
- An experimental validation demonstrating improved robustness compared with conventional single-point measurement approaches.

The proposed system provides a compact, low-cost, and intelligent alternative to traditional alcohol measurement instruments and is well suited for fermentation monitoring, portable analysis, and educational research.

2. METHODS

2.1. Hardware Architecture

The hardware architecture of the proposed system is shown in Fig. 1. The system consists of a GM-302B MEMS alcohol gas sensor, an ESP32-C6 microcontroller, AMS1117 2.5V, and auxiliary temperature and humidity sensors (SHT30).

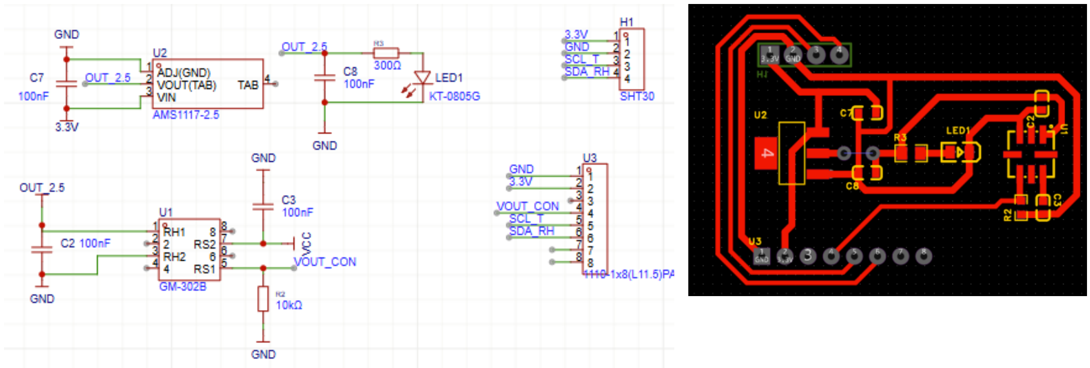

Fig. 1. Our proposed schematic and layout

The GM-302B sensor requires two supply voltages: a heater voltage (VH = 2.5 V) to maintain the sensing material at its operating temperature, and a circuit voltage (VC = 3.3 V) for signal measurement. The heater voltage is generated using a low-dropout linear regulator, while the measurement circuit is powered directly from the ESP32 3.3 V rail.

The sensing element of the GM-302B is modeled as a variable resistance Rs, whose value decreases with increasing ethanol concentration. Rs is connected in series with an external load resistor RL to form a voltage divider. The output voltage across RL is given by:

$$V_{out} = V_C \cdot \frac{R_L}{R_L + R_s} \tag{1}$$

This voltage is low-pass filtered using an RC filter to suppress high-frequency noise and is then fed into the ESP32 ADC input. An optional buffer amplifier may be used to provide impedance isolation, although the system can operate without it if the ADC input impedance is sufficiently high.

Temperature and humidity sensors are integrated into the system to capture environmental conditions, which are known to affect the sensor response. These values are used as additional inputs to the machine learning model to improve robustness and generalization.

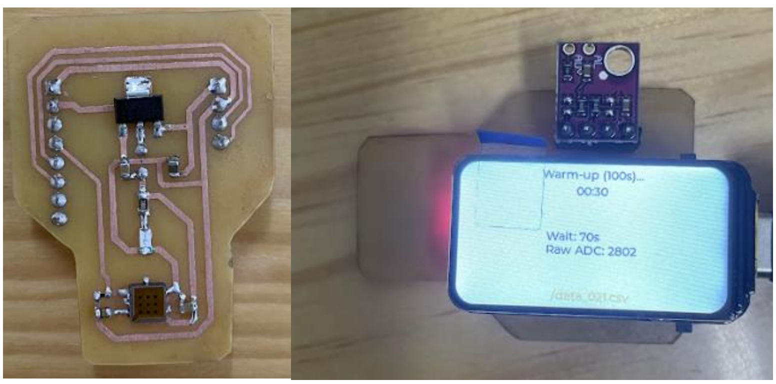

Fig. 2. Our proposed prototype

2.2. Sensor Characteristics and Motivation

The GM-302B sensor exhibits a nonlinear and gas-dependent sensitivity profile, typically represented by the ratio R_S/R_0 as a function of gas concentration (ppm). Fig. 3 illustrates the sensitivity curves of the sensor for several volatile compounds, including ethanol, acetone, toluene, and formaldehyde.

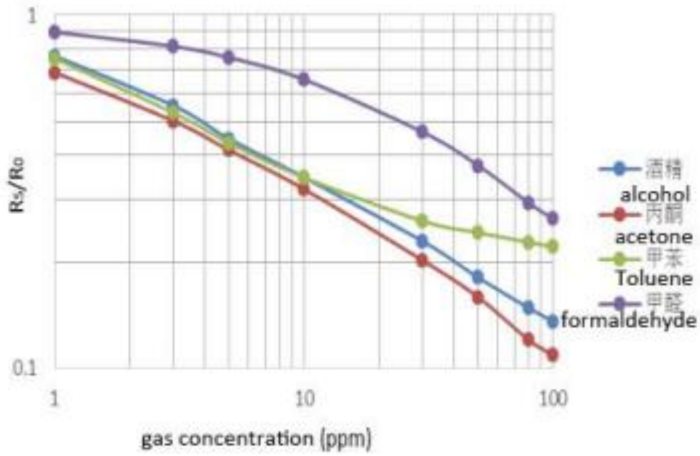

Fig. 3. Typical Sensitivity Curve from GM-302B datasheet

Two key observations can be made from these curves such as the sensitivity curves for different gases are not identical, indicating that the sensor has some degree of selectivity. However, the curves overlap significantly in the low-concentration region, and the curves for ethanol and acetone are particularly close, making it difficult to reliably distinguish ethanol from other volatile organic compounds using a single static measurement.

In addition, the sensor response is strongly affected by ambient temperature and humidity, which alter the adsorption and reaction kinetics on the sensing surface. As a result, a simple static mapping from output voltage to ethanol concentration is insufficient for accurate and robust measurement in realistic environments.

Instead, the dynamic response of the sensor contains additional information related to gas diffusion, adsorption rate, reaction speed, and desorption behavior. These transient characteristics differ across gases and environmental conditions, providing an opportunity for improved discrimination and estimation through time-domain analysis...

2.3. Software Workflow and Data Processing Pipeline

The proposed system is organized into two distinct processing pipelines: an offline data acquisition and training pipeline, and an online prediction pipeline embedded on the microcontroller.

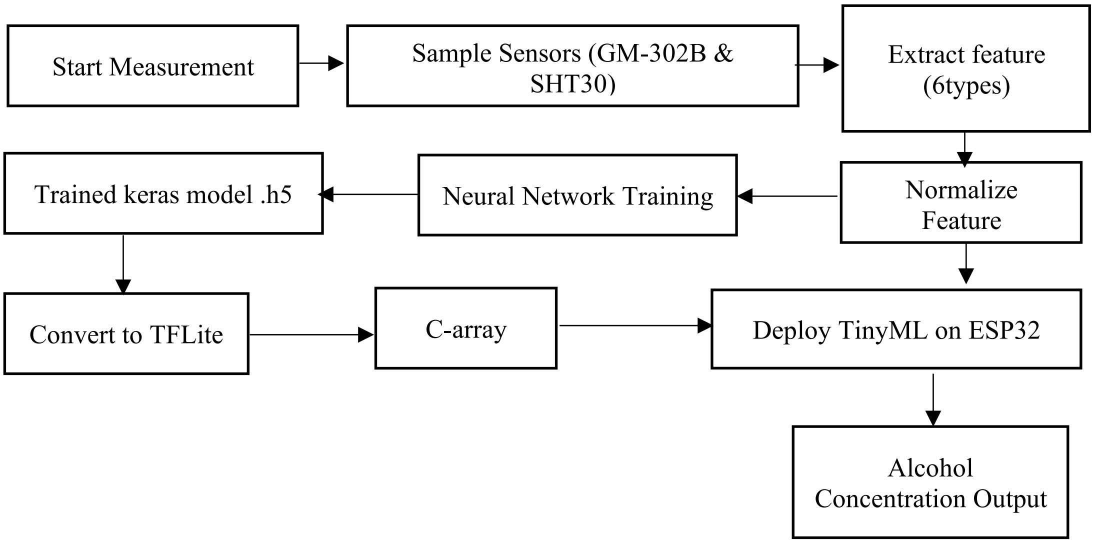

Fig. 4. Illustrations of our proposed workflow

2.3.1. Data Acquisition and Training Pipeline

The first pipeline is responsible for collecting labeled sensor data and training the machine learning model. This process is illustrated in Fig. 4.

The ESP32 samples the analog output of the GM-302B sensor together with ambient temperature and humidity at a fixed sampling interval. Each measurement session consists of a sensor exposure phase and a recovery phase, during which the temporal sensor response is recorded as a time series. The system integrates two primary sensing units: A GM-302B semiconductor alcohol sensor (chemical sensing), interfaced via a 12-bit Analog-to-Digital Converter (ADC) on GPIO1. To ensure high-fidelity signal acquisition, the ADC attenuation was set to 11dB, allowing a full-scale voltage range of approximately 3.1V. An Adafruit SHT40 sensor (environmental sensing), communicating via the I2C protocol (SDA: GP2, SCL: GP3), provides real-time ambient temperature and relative humidity data to compensate for environmental interference.

To ensure the stability of the metal-oxide sensor (MOS), a 100-second stabilization (warm-up) period was implemented before any measurement. Sensors signals were sampled at a frequency of 0.1 Hz (one sample every 10 seconds) throughout the entire measurement window. The raw ADC values were calibrated by subtracting the baseline value recorded at the end of the stabilization period to isolate the response triggered by alcohol exposure. During data acquisition, the alcohol vapor samples were collected in a controlled and repeatable manner. The sensor was placed at a fixed distance of approximately 7 cm from the vapor source to ensure consistent exposure conditions. Each measurement cycle started from ambient air and lasted for 460s, including an initial baseline period, a response phase during vapor exposure.

Instead of relying on a single steady-state sensor value, multiple time-domain features were extracted from the dynamic sensor response. Specifically, the following features were computed for each measurement cycle:

- Ambient Temperature (T): Average temperature during the sampling period.
- Relative Humidity (H): Average humidity during the sampling period.
- Maximum ADC Response (MaxADC): The peak calibrated ADC value recorded.
- Time to Peak (T_{peak}): The duration from the start of exposure to the achievement of MaxADC.

- Sensor Sensitivity Slope (S): Calculated as $S = MaxADC/T_{peak}$, representing the reaction rate.

- Area Under Curve (AUC): The integral of the ADC response over time, representing the total alcohol-sensor interaction as eq 2.

$$AUC = \sum_{i=1}^{n} (ADC_i \cdot \Delta t) \quad (2)$$

These six features form the input vector to the lightweight regression model and collectively capture both the magnitude and temporal characteristics of the sensor response. These features are stored on an SD card together with the corresponding reference alcohol concentration label. The collected dataset is then transferred to a personal computer, where the regression model is trained offline using standard machine learning tools.

Feature normalization is applied using a standard scaling operation defined as eq 3.

$$X_{norm} = \frac{X - \mu}{\sigma} \quad (3)$$

where μ and σ are the mean and standard deviation computed from the training dataset.

The dataset was collected at four alcohol concentration levels: 4%, 10%, 14%, and 19.5%. For each concentration level, 25 independent measurement samples were recorded under controlled experimental conditions, resulting in a total of 100 samples.

2.3.2. Model Architecture and TinyML Deployment

A Multilayer Perceptron (MLP) regression model was developed using the TensorFlow/Keras framework. The architecture consists of:

Input Layer: 6 neurons (one for each feature).

Hidden Layer 1: 16 neurons with Rectified Linear Unit (ReLU) activation.

Hidden Layer 2: 8 neurons with ReLU activation.

Output Layer: 1 neuron with linear activation to predict the continuous alcohol concentration.

The model was trained for 300 epochs using the Adam optimizer and Mean Squared Error (MSE) loss function. Post-training, the model was quantized and converted into a TensorFlow Lite (TFLite) flatbuffer, then embedded into the ESP32-C6 firmware as a C-style byte array for local inference. The dataset was split into training and testing subsets on a per-class basis, with 20 samples per concentration level used for training and the remaining 5 samples used for testing. This stratified split ensures that all concentration levels are represented in both training and evaluation sets while preventing data leakage across classes.

After training, the model is converted into TensorFlow Lite format and then transformed into a C array representation. The model parameters, together with the normalization constants μ and σ, are embedded into the ESP32-C6 firmware.

2.3.3. Prediction Pipeline

The second pipeline performs real-time alcohol concentration estimation directly on the ESP32 and is illustrated in Fig. 4.

During operation, the ESP32 acquires a new sensor response, extracts the same set of time-domain and environmental features, and normalizes them using the previously stored scaling parameters. The normalized feature vector is then passed to the embedded lightweight neural network model, which performs inference and outputs the estimated alcohol concentration.

The prediction result can be displayed locally, logged to the SD card, or transmitted wirelessly depending on the application requirements. This pipeline operates entirely on-device and does not require external computation or cloud connectivity.

Separating the system into an offline training pipeline and an online prediction pipeline offers several advantages. Offline training allows the use of more powerful computing resources and larger datasets to optimize model performance, while the embedded prediction pipeline ensures fast, low-latency, and energy-efficient operation in real-world deployments.

This architecture enables continuous improvement of the model through retraining while maintaining a lightweight and robust embedded implementation.

3. RESULTS & DISCUSSION

3.1. Performance Metrics

Model performance was evaluated using the mean absolute error (MAE) and root mean squared error (RMSE) between the predicted and reference alcohol concentrations:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{5}$$

where y_i and $\hat{y}_i$ denote the reference and predicted alcohol concentrations, respectively.

3.2. Overall Predictive Performance

The model was evaluated using a hold-out test dataset containing concentrations of 4%, 10%, 14%, and 19%. The quantitative performance metrics achieved are as Table 1.

Table 1. Model training results

Metrics	MAE	RMSE
Results	0.1842	0.2513

A MAE of 0.1842 indicates that the system's predictions deviate by less than 0.2% on average from the actual alcohol concentration. The proximity of the RMSE to the MAE suggests a low frequency of large outliers, indicating high model stability across different samples.

3.3. Class-Specific Error Analysis

An analysis of the MAE across specific concentration levels reveals insights into sensor behavior as Table 2

Table 2. The experimental results on mae

Actual Concentration (%)	MAE
4%	0.1214
10%	0.3528
14%	0.1512
19%	0.1114

The experimental results indicate that the model achieves its highest precision at the extreme ends of the tested range (4% and 19%). For the 4% concentration, the sensor operates near its baseline, where the response kinetics are highly predictable and less prone to turbulent fluctuations, resulting in a low MAE of 0.1214. Conversely, at the 19% concentration, the system achieved its most accurate performance (MAE = 0.1114). This is primarily due to the "saturation-like" yet distinct magnitude of the Area Under Curve (AUC), which reaches values near 58,900. This large numerical gap between the 19% class and the lower concentrations creates a high degree of linear separability in the feature space, allowing the Artificial Neural Network (ANN) to identify the 19% samples with near-perfect accuracy. In contrast, the 10% concentration class exhibited the highest variance, with an MAE of 0.3528. Observation of the raw data suggests that at this intermediate level, the GM-302B sensor experiences a "transition phase" in its surface-adsorption kinetics. This is evidenced by the wide spread of AUC values (spanning approximately 7,780 units) for the same 10% concentration. Such volatility in the input features introduces "noise" into the regression model, making it the most challenging range for the TinyML algorithm to map accurately. Notably, the 14% concentration showed a return to stability (MAE = 0.1512). Despite the higher ethanol content compared to the 10% samples, the MaxADC values remained remarkably consistent around 109 units. The ANN successfully utilized this consistency, along with the Slope feature, to differentiate 14% samples from the 10% and 19% classes effectively. This analysis confirms that while semiconductor sensors inherently possess non-linear characteristics, the integration of multi-feature TinyML models can successfully compensate for these fluctuations to maintain an overall system MAE of less than 0.19%.

The integration of T and H from the SHT40 sensor allowed the ANN to learn environmental dependencies. MOS sensors like the GM-302B are traditionally sensitive to humidity; however, the MLP model effectively utilized the humidity feature to normalize the MaxADC and AUC responses. The Slope feature was particularly useful in distinguishing the 10% samples, which exhibited a uniquely high reaction rate ($S \sim 1.04$) compared to the 19% samples ($S \sim 0.34$).

3.4. Computational Latency and Resource Utilization

The successful deployment of the TinyML model on the ESP32-C6 platform demonstrates the high feasibility of utilizing Edge AI for localized chemical sensing applications. Experimental measurements indicated that the system achieves exceptional computational efficiency, with a single inference cycle requiring approximately 5 ms. This rapid execution is supported by a minimal memory footprint; the TFLite model necessitated a Tensor Arena size of only 10 KB, which occupies a negligible fraction of the ESP32-C6's 512 KB total SRAM. Such resource efficiency ensures that the microcontroller can simultaneously manage sensor data acquisition, complex AI computations, and real-time UI rendering on the ST7789 display without any performance degradation. Crucially, this localized processing approach enables instantaneous feedback immediately after the sensing peak is detected, effectively eliminating the latency and connectivity dependencies typically associated with cloud-based processing.

3.5. Limitations of the study

Despite the promising predictive performance of the proposed system, several limitations must be acknowledged. First, due to experimental time constraints, the study focused on four discrete ethanol concentrations (4%, 10%, 14%, and 19.5%), which limits the validation of the model's interpolation and extrapolation capabilities for intermediate levels or high-proof spirits exceeding 20% ABV. Furthermore, the current evaluations were primarily conducted in controlled laboratory environments, leaving the system's robustness against extreme weather conditions or interfering volatile organic compounds (VOCs) that affect sensor selectivity not yet fully verified. Finally, the long-term stability of the trained TinyML model in relation to natural sensor drift and aging was not assessed, necessitating future research into on-device adaptive learning algorithms to maintain consistent measurement accuracy over time.

4. CONCLUSIONS AND FUTURE WORK

This study successfully demonstrated the development and deployment of a real-time alcohol concentration prediction system using a TinyML approach on the ESP32-C6 microcontroller. By integrating environmental data from an SHT40 sensor with dynamic kinetic features-specifically AUC, MaxADC, and Slope-from a GM-302B semiconductor sensor, the Artificial Neural Network (ANN) achieved high predictive performance. The model reached a Mean Absolute Error (MAE) of 0.1842 and a Root Mean Squared Error (RMSE) of 0.2513, proving its reliability across a concentration range of 4% to 19.5% ABV. Furthermore, the system exhibited exceptional efficiency for edge computing, with an inference latency of only 5 ms and a minimal memory footprint of 10 KB, allowing for instantaneous, localized feedback without cloud dependency.

Energy efficiency is an important consideration for the proposed embedded sensing system. The GM-302B sensor exhibits a heater power consumption about 50 mW, which dominates the sensor-side energy budget, while the SHT40 temperature and humidity sensor contributes negligible power due to its microampere-level current consumption. The ESP32-C6 platform integrates a RISC-V microcontroller and a color LCD display; although the exact power consumption of the complete development board is not explicitly specified, typical ESP32-class microcontrollers consume tens of milliamps in active mode without wireless communication and can enter microampere-level deep-sleep states when idle. The deployed TinyML model is deliberately lightweight, consisting of only two hidden layers and a small number of parameters, and is executed using TensorFlow Lite for Microcontrollers with fixed-point arithmetic. As a result, the computational overhead introduced by on-device inference is minimal compared to the overall system consumption, making the proposed architecture suitable for portable and battery-powered alcohol monitoring applications.

Despite the success in predicting relatively pure ethanol-water mixtures, future work is necessary to expand the system's applicability to more complex beverages, such as commercial wines. Wine presents a significant analytical challenge because it contains a myriad of interfering compounds, including sugars, organic acids, polyphenols, and various volatile organic compounds (VOCs). These constituents create complex cross-sensitivity patterns on the sensor surface, making concentration prediction significantly more difficult than in controlled solutions.

To address these complexities, future research will transition from a static MLP architecture to Long Short-Term Memory (LSTM) networks. LSTMs are specifically designed to process sequential data and capture temporal dependencies within the sensor's response curve over the entire sampling duration. By analyzing the time-series signature of the gas sensor rather than just fixed peaks or areas, an LSTM-based model is expected to more effectively decouple the ethanol signal from the background noise of wine's varied chemical matrix. This advancement will be a crucial step toward creating a universal, portable digital sommelier for high-precision beverage analysis.

Acknowledgement: We acknowledge the support of time and facilities from Tra Vinh University (TVU) for this study.

REFERENCES

[1] S. Kamiloglu, "Authenticity and traceability in beverages," *Food Chem.*, vol. 277, pp. 12-24, 2019, doi: https://doi.org/10.1016/j.foodchem.2018.10.091.

[2] T. Yarita, R. Nakajima, S. Otsuka, T. Ihara, A. Takatsu, and M. Shibukawa, "Determination of ethanol in alcoholic beverages by high-performance liquid chromatography–flame ionization detection using pure water as mobile phase," *J. Chromatogr. A*, vol. 976, no. 1-2, pp. 387–391, 2002, doi: https://doi.org/10.1016/S0021-9673(02)00942-1.

[3] G. H. Jain, “MOS Gas Sensors: What Determines Our Choice?,” in *Proc. 5th Int. Conf. Sens. Technol. (ICST)*, Palmerston North, New Zealand, 2011, pp. 66–72, doi: https://doi.org/10.1109/ICSensT.2011.6137067.

[4] N. Barsan and U. Weimar, “Conduction Model of Metal Oxide Gas Sensors,” *J. Electroceram.*, vol. 7, no. 3, pp. 143–167, 2001, doi: https://doi.org/10.1023/A:1014405811371

[5] P. Xu, K. Song, X. Xia, Y. Chen, Q. Wang, and G. Wei, “Temperature and Humidity Compensation for MOS Gas Sensor Based on Random Forests,” in *Intelligent Computing, Networked Control, and Their Engineering Applications*, D. Yue *et al*., Eds., vol. 762, *Communications in Computer and Information Science*. Singapore: Springer, 2017, pp. 135–145, doi: https://doi.org/10.1007/978-981-10-6373-2_14.

[6] R. Gutierrez-Osuna, "Pattern analysis for machine olfaction: A review," *IEEE Sens. J.*, vol. 2, no. 3, pp. 189–202, 2002, doi: https://doi.org/10.1109/JSEN.2002.800688

[7] P. Warden and D. Situnayake, *TinyML: Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers.* O'Reilly Media, 2019

[8] P. P. Ray, “A review on TinyML: State-of-the-art and prospects,” *J. King Saud Univ. Comput. Inf. Sci.,* vol. 34, no. 4, pp. 1595-1623, 2022, doi: https://doi.org/10.1016/j.jksuci.2021.11.019.

[9] R. David, J. Duke, T. Jain, V. Janapa Reddi, N. Jeffries, J. Li, N. Kreeger, I. Nappier, M. Natraj, T. Wang, P. Warden, and L. Rocky, "TensorFlow Lite Micro: Embedded Machine Learning on TinyML Systems," in *Proc. 3rd Conf. Machine Learning and Systems (MLSys)*, Austin, TX, USA, 2020.