Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 13, No. 2, June 2025, pp. 294~308
ISSN: 2089-3272, DOI: 10.52549/ijeei.v13i2.6146

Trainer Kit for Aroma Classification Using Artificial Intelligence

Jazi Eko Istiyanto, Danang Lelono, Oskar Natan,
Shafa Khamila, Hafizha Adhiyant, Ikhlasul Amal Abda'i, Ilyaz Raukhillah Adzaqi
Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences,
Universitas Gadjah Mada, Indonesia

Article Info

Article history:

Received Dec 12, 2024
Revised Mar 12, 2025
Accepted Mar 27, 2025

Keywords:

Aroma Classification
Machine Learning
Neural Network
Electronic Nose

ABSTRACT

This research focused on the development and evaluation of machine learning algorithms for aroma classification using sensor data, implemented within the e-Trainose system. Various algorithms, including Neural Network, Support Vector Machines, and Random Forest, were tested to determine their effectiveness in distinguishing between different aroma samples, namely alcohol, coffee, and tea. The study utilized an array of metal oxide semiconductor sensors to capture the volatile organic compounds associated with each aroma. The features tested included sensor responses such as resistance changes and Gaussian smoothing of sensor data. Among the algorithms tested, Neural Network demonstrated the highest accuracy (98.89%), precision (99.10%), recall (99.10%), and F1 score (99.10%), making it the most reliable model for this task. These results highlight the potential of using machine learning with e-Trainose for real-time aroma detection and classification. The research paves the way for future advancements in the field, including the development of hybrid models and further optimization of sensor-based classification systems.

Corresponding Author:

Jazi Eko Istiyanto
Department of Computer Science and Electronics,
Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada,
Bulaksumur BLS 21, Yogjakarta 55281, Indonesia.
Email: jazi@ugm.ac.id

1. INTRODUCTION

To remain competitive, cover current markets, explore new areas, and satisfy important consumer demands, food companies worldwide must continuously produce new products. This continuous advancement imposes significant requirements for excellence across all dimensions, especially for food safety, integrity, quality, nutrition, and other health impacts [1]. Artificial systems that detect sensory criteria such as aroma, taste, and external appearance would facilitate objective assessment of food quality, presenting a significant opportunity for the food business [2]. Every food possesses a distinct aroma determined by its composition, influenced by origin, processing, storage, and level of deterioration [3]. In the food business, the electronic nose has been used to assess meat quality, fruit quality, and spices, and to analyze coffee and wine, as well as to detect fragrance changes during dough fermentation or pastry aging [4], [5], [6].

In 1982, Persaud and Dodd published the first electronic nose (E-nose) design based on chemical sensors and pattern recognition. An e-nose is a device that consists of a series of reversible but only semi-selective gas sensors connected to a pattern recognition algorithm [7]. The instrument's selectivity is obtained by applying pattern recognition algorithms to the sensor array's responses [8]. Despite significant advancements in electronic nose technology, its promotion and use remain constrained by the development of gas-sensitive materials, sensor fabrication procedures, data processing, and pattern recognition methodologies. Conventional gas sensors are bulky and cumbersome, exhibit high power consumption, lack integration, and are limited to single detection capabilities [9].

Previous research by Tayagornkul and Numsek [10] successfully proves the efficacy of e-nose technology in detecting chicken flesh deterioration. Utilizing seven sensors improves detection capabilities, with PCA validating different alterations in odor patterns across time. Even though the study shows encouraging findings, future research should concentrate on fixing sensor drift difficulties, comparing the system to existing detection techniques, and scaling it for industrial usage. Zhang and Tai [11], on the other hand, concentrated on the theoretical aspects of data preprocessing techniques for electronic nose systems and emphasized their significance in enhancing gas categorization accuracy. The analysis is enlightening; nonetheless, subsequent research should prioritize experimental validation, sensor performance assessment, and hybrid algorithm methodologies to improve the dependability of e-nose systems in practical applications. Liu and Tang's study [12] successfully introduces the MDIP (Minimum Distance Inliers Probability) feature selection method, which improves gas classification in e-nose systems by removing unneeded features and increasing accuracy. The method is computationally efficient and enhances pattern separation. However, more study is needed to assess its usefulness in real-world scenarios, compare it to other feature selection strategies, and broaden its applicability to various gas classifications. These previous studies highlight the potential of e-nose technology in gas classification but also underscore key challenges such as sensor drift, data processing limitations, and the need for more robust AI-driven methodologies. This research aims to bridge these gaps through the development of e-Trainose.

Unlike conventional e-nose systems, which can be bulky and complex, e-Trainose is designed to be a high-performance, AI-driven training tool for educational and research applications. By integrating AI-driven aroma analysis with an intuitive interface, e-Trainose enhances accessibility for students and researchers exploring advanced gas classification techniques. While e-nose technology has been widely applied in food quality assessment, its use in coffee aroma classification remains challenging due to variations in humidity, temperature sensitivity, and the difficulty of distinguishing subtle differences between similar coffee varieties. Addressing these challenges requires a robust AI-driven approach capable of accurately identifying and classifying complex aroma profiles.

According to the International Coffee Organization (ICO), Indonesia is one of the world's largest coffee exporters, ranking fourth globally between 2022 and 2024 [13]. Coffee consumption has steadily increased, with annual growth rates of 2% in consumption and 17% in coffee bean production in recent years [14]. Global coffee production is dominated by two species: Robusta coffee (*Coffea canephora* var. Robusta), which accounts for 30%, and Arabica coffee (Coffea arabica), which makes up the remaining 70% [15]. Robusta is generally more heat-resistant but more susceptible to low temperatures than Arabica.

Electronic nose technology has become increasingly important for aroma analysis across various industries, such as food quality control, beverage authentication, and environmental monitoring. The key advantages of e-nose systems include fast response times, low power consumption, and exceptional sensitivity to odors [16]. The signals generated by the e-nose system correspond to distinct patterns, which can be classified based on their recognized characteristics [17].

This research addresses traditional e-nose systems' limitations by developing e-Trainose, a user-friendly and cost-effective trainer kit. The novelties of this research include:

- Development of e-Trainose, a high-performance trainer kit that simplifies the integration of e-nose technology with AI for educational purposes.
- Interactive and user-friendly design that bridges the gap between theoretical learning and practical application.
- Incorporation of gas sensors, AI algorithms, and an intuitive interface, making it accessible to both students and researchers.
- The integration of AI-driven aroma analysis to provide hands-on learning and research opportunities in the rapidly advancing field of AI-based sensory technologies.

An overview of the e-Trainose system design and its functionality can be seen in Figure 1. Section 2 provides a literature review on existing sensing systems and artificial intelligence models (such as SVM, NN, and Random Forest) in aroma classification. Section 3 outlines the research methodology, including the proposed model, system block diagram, sensors used in this research, dataset details, data preprocessing, feature extraction, and the evaluation metrics used for performance assessment. Section 4 presents the results and discussion, focusing on the application of the models and the evaluation of performance based on the metrics results.

2. LITERATURE REVIEW

2.1. Sensing System

Aroma sensing technology, particularly in electronic nose (e-nose) systems, often relies on semiconducting gas sensors. Gas sensors are designed to detect specific gases or groups of gases. There are several types, including electrochemical, semiconductor (metal oxide), and infrared (IR) sensors.

Electrochemical sensors detect gases through chemical reactions that generate an electrical signal proportional to gas concentration. These sensors are widely used in environmental monitoring, industrial safety, and medical diagnostics, detecting compounds such as carbon monoxide (CO), hydrogen sulfide (H_2S), nitrogen dioxide (NO_2), and oxygen (O_2) [18]. MOS sensors operate based on changes in electrical conductivity when exposed to gases. They are cost-effective, highly responsive, and capable of detecting a broad range of volatile organic compounds (VOCs), including methane, propane, and ethanol [19]. However, they suffer from cross-reactivity, meaning they respond to multiple gases simultaneously, reducing selectivity.

To address these limitations, researchers have explored hybrid sensor arrays, combining MOS and electrochemical sensors to enhance accuracy. For instance, Chowdhury et al. [20] demonstrated that a hybrid e-nose system for black tea aroma classification significantly improved accuracy, highlighting the benefits of multi-sensor integration. Despite these advances, calibration complexity, environmental sensitivity, and sensor aging remain challenges that need further research.

Electronic nose technology has been widely applied in food quality assessment, environmental monitoring, and industrial safety. Various studies have explored different sensor configurations, ranging from 6 to 15 gas sensors, to detect and classify aroma profiles in applications such as liquor identification, gas detection, and meat spoilage monitoring [10], [12], [21], [22].

Feature selection techniques like Minimum Distance Inliers Probability (MDIP) have been used to refine the most relevant sensor inputs, improving model efficiency and accuracy [12]. Additionally, Principal Component Analysis (PCA) has been employed for dimensionality reduction, with key components capturing over 95% of data variance, helping improve classification performance [10]. Table 1 compares these studies, summarizing the sensor configurations, dataset size, and performance results.

Table 1. Comparison of Previous Studies

Title	Sensor Used	Number of Data Samples	Key Results
"A Novel Olfactory Neural Network for Classification of Chinese Liquors Using Electronic Nose"	6 gas sensors	50 measurements (10 per liquor type)	Classification accuracy: - 94% (olfactory neural network) - 84% (traditional methods)
"High-Performance Sensing Application of FPGA-based Winograd CNN Accelerator in Electronic Nose Systems"	15 metal oxide gas sensors	150 measurements (10 per gas type, 15 times each)	Classification accuracy: - 99.78% (full precision) - 99.5% (quantized model)
"Analysis of Chicken Meat Degradation Using Electronic Nose Technology"	7 gas sensors (including TGS825, TGS826, and TGS2602)	Measurements taken every 6 hours over 48 hours	PCA variance: 96.98% (PC1: 87.24%, PC2: 9.74%)
"A Minimum Distance Inliers Probability (MDIP) Feature Selection Method to Enhance Gas Classification for an Electronic Nose System"	8 gas sensors (TGS2611, TGS2612, TGS2610, TGS2602)	5 replicates per gas type (ethanol, methane, ethylene, CO)	Classification accuracy: - Without MDIP: 68%-99% - With MDIP: 79%-100%

While earlier studies have shown that e-nose systems are feasible for classifying aromas, a number of issues still need to be addressed. First, many studies rely on small or controlled datasets, restricting their applicability in real-world conditions. Models based in Convolutional Neural Network (CNN), for example, require large labeled datasets for optimal performance, but collecting and labeling gas data is often time-consuming and costly [23].

Second, MOS sensors suffer from sensor drift and cross-reactivity, reducing long-term reliability. Sensor drift occurs when a sensor's response gradually changes over time, leading to inconsistent readings, while cross-reactivity makes it difficult to distinguish similar compounds [24]. These challenges highlight the need for AI models that can dynamically compensate for sensor instability.

Lastly, hybrid sensor arrays have been proposed to enhance selectivity and sensitivity, but their real-time application remains underexplored. While laboratory studies have confirmed their effectiveness, industrial and environmental applications still require further research. This underscores the importance of developing advanced calibration techniques and robust machine-learning frameworks to ensure stability and accuracy in dynamic conditions [25].

Despite the advancements in sensor technology, accurately classifying aroma profiles remains challenging due to sensor drift, cross-reactivity, and variations in environmental conditions. To address these limitations, machine learning models play a crucial role in enhancing the reliability and precision of e-nose systems.

2.2. Artificial Intelligence Model

The roast level of coffee beans significantly influences their flavor profile, chemical composition, and consumer appeal. The e-nose, capable of identifying variations in aroma, provides an accurate chemical profile of the coffee roast. Its application provides a fascinating, efficient, and less resource-intensive analysis procedure than standard methods. However, the complex data collected by the e-nose requires the application of machine learning methods. Machine learning techniques find patterns within this data to maximize its value. Combining e-nose and machine learning simplifies and enhances the analysis of coffee roast fragrance [26], [27].

By applying machine learning techniques, researchers can enhance data interpretation, pattern recognition, and predictive accuracy in aroma classification tasks. Integrating machine learning with e-nose systems not only improves aroma analysis efficiency but also allows for more precise differentiation between similar aroma profiles. Among the various machine learning techniques available, Support Vector Machines (SVM), Neural Network (NN), and Random Forest are widely used for their robustness and effectiveness in aroma classification tasks. These models were chosen over alternatives such as K-Nearest Neighbors (KNN) and deep learning architectures for their adaptability to small and medium-sized datasets and computational efficiency, and they demonstrated success in previous research.

Support Vector Machine (SVM) is a supervised learning algorithm that finds the optimal hyperplane to separate different data classes while maximizing the margin between them. This makes SVM highly effective in high-dimensional spaces, essential for tasks like text classification and image recognition [28]. Distante et al. [29] have demonstrated that SVMs are suitable for distinguishing subtle differences in volatile compounds, as evidenced by their widespread use in odor recognition. This is due to their capacity to manage complex, non-linear patterns in olfactory signal data. In another scent identification research by Kim et al. [30], SVM models classified four distinct fragrance samples across three concentration levels, proving their resilience to changes in concentration and aroma.

Neural Network, particularly deep learning models, are increasingly popular in aroma classification due to their ability to model non-linear relationships and capture complex patterns in data. NNs consist of an input layer, multiple hidden layers, and an output layer, with each layer learning from the weighted connections between them. NN have proven highly effective in aroma classification, especially when integrated with semiconductor gas sensors, enabling precise aroma recognition [31]. Furthermore, artificial neural network have been used to forecast scent quality and threshold values of pyrazine derivatives, with a classification accuracy of 93.7% for three separate olfactory impressions [32]. This illustrates NN's efficacy in modelling the complex connections between chemical composition and perceived scent.

Conversely, Random Forest has been utilized in numerous aroma categorization investigations due to its resilience to noise and overfitting. Researchers utilized Random Forest models to categorize 98 pyrazine derivatives into three aroma classifications—Green, Nutty, and Bell-Pepper—demonstrating Random Forest's capacity to manage intricate classification challenges associated with volatile organic molecules [33]. Furthermore, another study utilized Random Forest for the classification of odor emission sources based on molecular descriptors, demonstrating its efficacy in distinguishing between diverse odor-emitting compounds [34].

While K-Nearest Neighbors (KNN) is computationally simple, it struggles with high-dimensional sensor data and is highly sensitive to noise. CNN, although powerful, requires large labeled datasets and significant computational resources, making them quite impractical for real-time e-nose applications. In contrast, SVM, NN, and Random Forest balance between accuracy, efficiency, and scalability, making them ideal choices for aroma classification.

3. RESEARCH METHOD

3.1. Proposed Model

The main goal of this project is to create a real-time system for classifying aromas by combining machine learning algorithms with gas sensors. This research introduces a portable e-nose prototype that utilizes Jetson Orin Nano for edge computing and Gaussian smoothing to enhance signal stability, in contrast to prior studies. This system is designed to categorize three specific aromas: alcohol, coffee, and tea, rendering it suitable for food quality assessment, environmental monitoring, and research training.

The research is divided into two main phases: training and evaluation. In the training phase, aroma samples are collected, preprocessed, and used to train three machine learning models: Support Vector Machine (SVM), Neural Network (NN), and Random Forest. These models are optimized to identify the chemical patterns associated with each aroma. In the evaluation phase, the models are tested on unseen samples to assess their accuracy, precision, and robustness in real-world applications. The best-performing model is identified for practical deployment in aroma classification tasks.

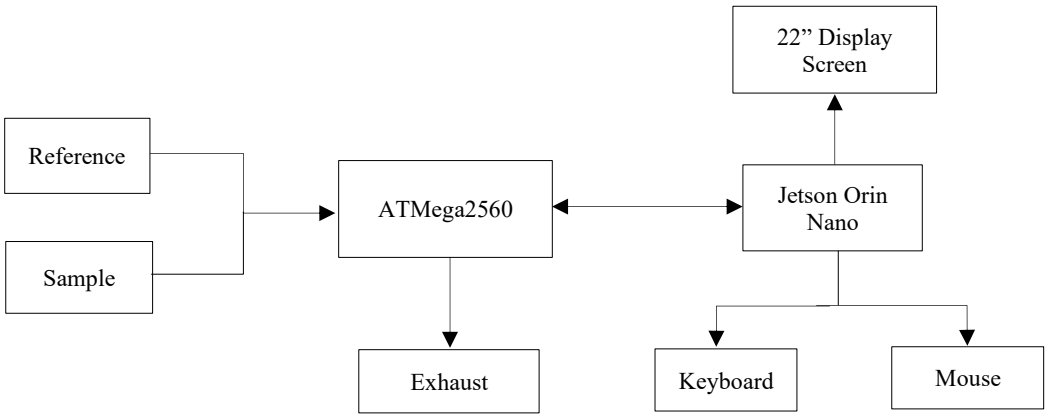

Figure 1. Block Diagram System

Figure 1 shows that the system prototype incorporates MOS sensors, a Jetson Orin Nano, and an ATMega2560 microprocessor for real-time aroma classification. The Jetson Orin Nano was selected for its superior GPU processing capabilities, facilitating on-device machine learning inference independent of cloud computing. The ATMega2560 was chosen for its proficiency in effectively managing multiple analog gas sensor inputs, guaranteeing stable communication between the sensor array and the AI processor.

Before the operation, the e-Trainose undergoes a 10-minute pre-calibration phase. This step eliminates residual odors, ensuring that sensor measurements remain unaffected by previous samples, thereby maintaining data integrity. In the classification process, a pump mechanism distributes sample aroma and clean air through a 5/2 valve into the ATMega2560 via interconnected pipes. The system continuously exhausts air through a separate pipe to maintain a stable atmosphere, and the clean airflow keeps samples from becoming contaminated.

The ATMega2560 microcontroller transmits real-time sensor data to the Jetson Orin Nano through serial connection. The Jetson Orin Nano processes this data utilizing the trained machine learning models (SVM, NN, Random Forest) and classifies the aroma in real time. A dongle facilitates the connection of a keyboard and mouse, enabling user interaction with the system. The categorization outcomes are presented on a 22-inch display, offering real-time feedback and enhancing the system's usability in educational and industrial contexts.

3.2. Sensors

In this research, various sensors have been selected to detect different gases associated with specific aroma profiles. These sensors can detect various gases useful for applications in environmental monitoring, safety systems, and consumer electronics. By utilizing these sensors, we can capture the diverse chemical signatures present in various substances, facilitating more accurate aroma classification.

Table 2. Sensor Array Specifications List

No.	Sensor	Description
1	TGS2600	Ammonia (NH_3)
2	TGS2602	Hydrogen (H_2) and other combustible gases
3	TGS816	Carbon Monoxide (CO)
4	TGS813	Methane (CH_4)
5	MQ8	Methane (CH_4)
6	TGS2611	Methane (CH_4) and Liquefied Petroleum Gas (LPG)
7	TGS2620	Alcohol ($Ethanol$) and Carbon Monoxide (CO)
8	TGS822	Alcohol ($Ethanol$) and other organic solvents
9	MQ135	Ammonia (NH_3), Benzene (C_3H_6), Alcohol (C_2H_5OH), and smoke
10	MQ3	Alcohol ($Ethanol$)

Table 2 summarizes the sensors used in this study and the gases each sensor is designed to detect. These sensors form the core of the e-Trainose system, allowing it to recognize and classify various aromas based on the volatile organic compounds (VOCs) they emit. By utilizing these sensors, the system can

efficiently capture the chemical signatures of alcohol, coffee, and tea aromas, contributing to accurate aroma classification.

3.3. Dataset

The dataset used in this study consists of sensor readings collected during a series of inhaling and flushing cycles. As Figure 2 shows, in one completed cycle of inhale and flush, the system records 156 sensor value readings. Each line represents the sensor's response to the air during either the inhaling or flushing phase. During the inhaling phase, the system draws in the sample aroma to the sensing chamber, allowing the sensors to register the VOCs associated with the specific aroma. During the flushing phase, the system draws in neutral air to neutralize the system and ensure that sensor readings are not affected by residual aromas from the previous cycle. This ensures that the sensing chamber only measures the new sample aroma when it is inhaled.

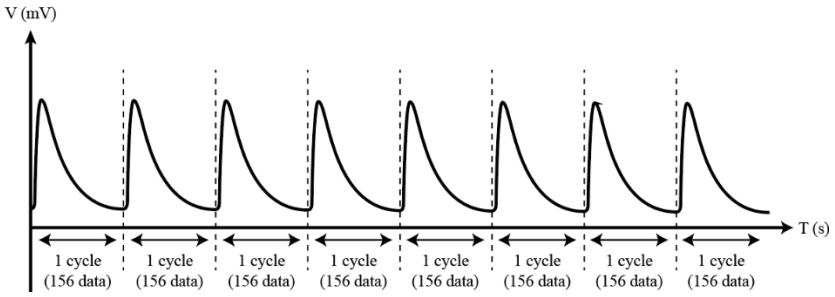

Figure 2. Sensor Response

Each aroma is further divided into specific types of categories, with each type associated with different brands. These brands were selected because they are quite diverse and readily available on the market. For the coffee category, there are three different types of coffee: Kapal Api, Good Day, and Nescafé. Each type of coffee was sampled with a total of 900 cycles for the coffee category. Similarly, for the tea category, there are three different brands: Teh Sosro, Teh Tong Tji, and Teh Sari Wangi, each sampled with a total of 900 cycles for the tea category. For the alcohol category, two types of alcohol were chosen: Alkohol 70% Ika and Alkohol 70% Dixol. Each type of alcohol was sampled with a total of 900 cycles for the alcohol category. In total, the dataset consists of 2,700 cycles, with 900 cycles for each of the three aroma categories. Each cycle includes 156 sensor value readings collected during both the inhaling and flushing phases, resulting in a total of 421,200 data.

3.4. Data Preprocessing

Preprocessing is the first step in analyzing sensor response data. This stage involves reducing factors like noise and irrelevant readings that can interfere with sensor data, making it easier to extract meaningful features [35]. The method applied in data preprocessing involves transforming the sensor response $X_{ij}(t)$, which is relative to the baseline or minimum value, denoted as $X_{ij}(0)$, where i represents the sensor and j represents the aroma sample.

$$Y_{ij}(t) = X_{ij}(t) - X_{ij}(0) \quad (1)$$

The sensor response $Y_{ij}(t)$, after preprocessing using the difference method as shown in equation (1), helps minimize errors such as additive noise, leading to a clearer signal in both the reference and the sensor response signals [36].

This study applies Gaussian smoothing as part of the data preprocessing to reduce noise and fluctuations in the sensor readings, ensuring a clearer and more consistent signal for analysis. The raw sensor data often contains high-frequency noise caused by environmental factors or sensor variability, which can obscure the underlying patterns essential for aroma classification. Gaussian smoothing mitigates this issue by applying a Gaussian function to the data, emphasizing the central values while gradually reducing the weight of surrounding data points. The Gaussian function is defined as:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (2)$$

Where $G(x)$ represents the value of the Gaussian function at point x, and σ is the standard deviation, which controls the spread of the Gaussian curve. In the smoothing process, the raw data is convolved with the Gaussian function to computer the smoothed signal:

$$S_{smoothed}(x) = \sum_{i=-n}^{n} S(x+1)G(i) \quad (3)$$

Here, $S(x)$ represents the original sensor data at point x and $S_{smoothed}(x)$ is the resulting smoothed value. The parameter n defines the range of the window applied during smoothing, which is determined based on the data characteristics.

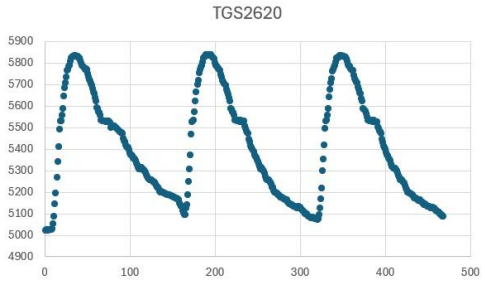

Figure 3. Raw Data of TGS2620 Sensor

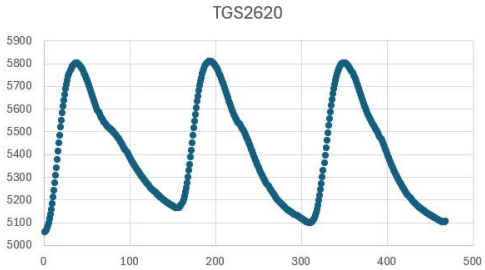

Figure 4. Preprocessed Data of TGS2620 using Gaussian Smoothing

The application of Gaussian smoothing significantly reduces noise while preserving the important trends in the sensor responses. Figure 3 illustrates the raw sensor data for a single aroma sample from the TGS2602 sensor during an inhaling/flushing cycle, showing significant fluctuations and noise. In contrast, Figure 4 shows the data after it has been preprocessed with Gaussian smoothing, revealing a much smoother and more consistent signal. The smoothed data makes it easier to extract meaningful features and enhances the classification accuracy by reducing the impact of irrelevant fluctuations in the sensor readings.

3.5. Feature Extraction

Feature extraction in an electronic nose is a technique used to identify the key characteristics of sensor responses by eliminating irrelevant information from the pattern recognition system [36]. During this process, the sensor response, once its features are extracted, becomes independent of time, as represented by equation (4).

$$X_{ij}(t) \rightarrow X_{ij} \quad (4)$$

Table 3. Feature Extraction Method Based on Statistical Analysis

No.	Method	Description	Formula
1	Mean	Average of all values in a dataset.	$Mean = \frac{(\sum x)}{n}$
2	Skewness	Inclination or the slope of a curve as seen from the difference in the location of the mean, median and mode.	$Skewness = \frac{3(\bar{x} - \tilde{x})}{s}$
3	Kurtosis	Sharpness or steepness of the peak of the data distribution.	$Kurtosis = \frac{\sum(x_i - \bar{x})^4}{N.\sigma^4}$
4	Maximum Value	Highest value in the data.	-
5	Minimum Value	Lowest value in the data.	-

Table 3 summarizes the feature extraction methods used in this research, which are based on statistical analysis of the sensor data. These methods are crucial for transforming raw sensor readings into meaningful features that can be used for machine-learning models in aroma classification. Each method focuses on extracting specific characteristics from the data, such as mean, skewness, and kurtosis, which provide insights into the sensor response data's distribution, variability, and shape.

3.6. Model

With the SVM model, the data is first cleaned by handling any missing values, ensuring that all features are properly prepared for training. The data is then divided into two parts: one for training the model and the other for testing it. When the SVM classifier is used, it learns to separate the different categories in the data by finding the best boundary between them (hyperplane). The model uses a linear kernel to create this optimal separating hyperplane in high-dimensional space, maximizing the margin between classes. Through the pipeline, the data is automatically standardized (scaled to mean = 0, variance = 1) before training, which is crucial for SVM's performance. The model is trained on the training data and evaluated through 5-fold cross-validation to estimate its overall performance, with probability estimates enabled for confidence scoring.

Using the Neural Network model, the data is first loaded and processed by handling any missing values. The dataset is then split into features and labels, with the labels being encoded into numerical values. The data is further divided into training and test sets. After that, the features and labels are converted into PyTorch tensors, which are required for training the neural network. A neural network model is defined as three connected hidden layers (with 256, 128, and 64 neurons for each respective layer), using ReLU activation functions for introducing non-linearity. A dropout rate of 30% is applied after each hidden layer to prevent overfitting by randomly disabling neurons during training. The model uses an Adam optimizer with a learning rate of 0.0005 for stable gradient updates and is trained for 50 epochs with a batch size of 64, balancing computational efficiency and gradient estimation accuracy. During training, gradient clipping is applied to avoid issues with exploding gradients. After each epoch, the model is evaluated on the test set to check its accuracy, and the best model (with the highest accuracy) is saved using early stopping.

Lastly, in the Random Forest model, the data is loaded and processed by separating the features and labels. The labels are then encoded into numerical values using a label encoder. A Random Forest classifier is defined with 100 decision trees, where each tree is built using a random subset of the training data and considers only a random subset of features at each split. This ensemble approach reduces variance and improves generalization compared to a single decision tree. The trees are grown to their maximum depth (unless nodes become pure or contain fewer than two samples) and their predictions are combined through majority voting for classification. The model is trained on the full dataset and evaluated through 5-fold cross-validation to assess its robustness. After training and evaluation, the generated label encoder from each model is saved for usage in predicting aroma samples, along with the trained Random Forest model which maintains its ability to provide feature importance scores for interpretability.

3.7. Metrics

Metrics are quantitative measures used to evaluate the performance of a machine learning model. In classification tasks, metrics provide insight into how well a model predicts and classifies data. Four key classification metrics are employed to assess the performance of the machine learning models in this study: accuracy, precision, recall, and F1-score. These metrics comprehensively evaluate the classification models by measuring their ability to correctly classify aroma samples while balancing both correct predictions and errors. The choice of these metrics is supported by Sokolova and Lapalme [37], who emphasized that combining these evaluation measures offers a well-rounded assessment, particularly in multi-class classification tasks such as aroma recognition.

Accuracy is one of the simplest and most used metrics, representing the proportion of correct predictions made by the model out of all predictions. While accuracy provides a general performance overview, it can be misleading in imbalanced datasets, where one class dominates the predictions [38]. It is calculated as:

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Instances} \quad (5)$$

Precision measures the proportion of true positive predictions out of all the instances predicted as positive. High precision indicates that when the model predicts a sample to be positive, will likely to be correct. This metric is handy when false positives must be minimized, such as in fraud detection or medical diagnostics [37]. It is computed as:

$$Precision = \frac{True\ Positives}{True\ Positives+False\ Positives} \quad (6)$$

Recall, also known as sensitivity or True Positive Rate, measures the proportion of actual positives that were correctly identified by the model. This metric is essential in applications where missing a positive case is costly, such as anomaly detection or medical diagnosis [39], where failing to detect an issue could have serious consequences. The formula for recall is:

$$Recall = \frac{True\ Positives}{True\ Positives+False\ Negatives} \quad (7)$$

Since precision and recall often exhibit a trade-off, the F1-score balances the two. The F1-score is the harmonic mean of precision and recall, making it particularly effective when dealing with datasets where class distributions are slightly imbalanced [40]. A higher F1-score indicates better classification performance, ensuring that neither false positives nor false negatives disproportionately affect the results. It is defined as:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (8)$$

While other evaluation metrics, such as ROC-AUC (Receiver Operating Characteristic - Area Under the Curve) and Matthews Correlation Coefficient (MCC), could provide additional insights, they were not selected for this study. ROC-AUC is more suitable for binary classification problems, as it evaluates the trade-off between true positive and false positive rates across different classification thresholds, making it highly effective for imbalanced binary tasks [41]. However, since this research involves a multi-class classification problem, ROC-AUC was deemed less applicable. Similarly, MCC is often preferred for highly imbalanced datasets because it accounts for all four confusion matrix elements (true positives, false positives, true negatives, and false negatives), providing a single-score performance measure [42][43]. MCC was not considered necessary since the dataset used in this study maintains a relatively balanced number of samples across the three aroma categories.

By integrating accuracy, precision, recall, and F1-score, this research ensures a rigorous and well-rounded evaluation of aroma classification models. Accuracy provides a general correctness measure, precision ensures classification reliability, recall accounts for missed detections, and F1-score balances precision and recall. This selection allows for a comprehensive analysis of classification performance, ensuring that the system is effective and robust for real-world aroma classification applications.

4. RESULTS AND DISCUSSION

The development of e-Trainose was aimed at creating a high-performance trainer kit that simplifies the integration of e-nose technology with AI for educational purposes. Our evaluation of Neural Network (NN), Support Vector Machines (SVM), and Random Forest (RF) directly supports this goal by assessing their effectiveness in AI-driven aroma analysis.

The performance of the Neural Network (NN), Support Vector Machine (SVM), and Random Forest models was evaluated using accuracy, precision, recall, and F1 score. Overall, all three models performed well, with some notable differences, as shown in Table 4.

Table 4. Comparison Table

	Accuracy	Precision	Recall	F1-Score
Neural Network	98.89%	99.10%	99.10%	99.10%
SVM	94.78%	95.03%	94.78%	94.77%
Random Forest	97.67%	97.67%	97.67%	97.67%
Threshold-based	80.12%	81.34%	79.89%	80.10%

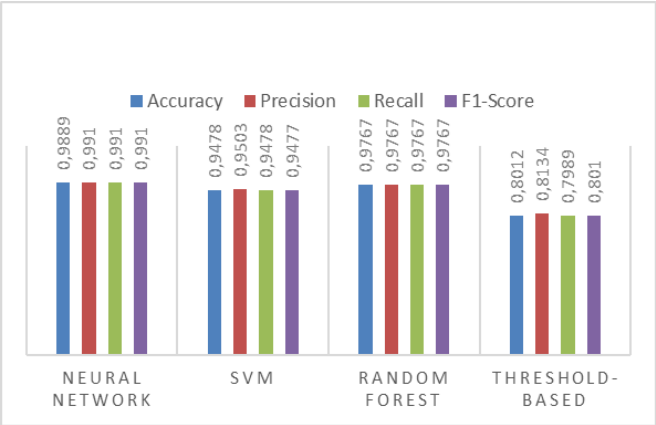

Figure 5. Comparison Chart

Neural Network achieved the highest accuracy of nearly 99%, with both precision and recall reaching about 99%. This excellent performance is likely due to the NN's ability to model complex, non-linear relationships in the data. The NN model achieved optimal classification performance by learning these patterns during training. Its balanced precision and recall indicate that the model successfully avoided both false positives and false negatives, providing a reliable classification for each aroma sample.

Random Forest performed closely behind NN, with an accuracy of approximately 97% and similarly high precision and recall values. The ensemble nature of Random Forest, which combines multiple decision trees, allows it to reduce overfitting and increase robustness. Although its accuracy was slightly lower than NN's, Random Forest is particularly effective at handling high-dimensional feature spaces and robustly dealing with variability in the data, making it a highly reliable model for aroma classification.

Support Vector Machine had the lowest performance among the three models, with an accuracy of approximately 94%. SVM relies on finding a hyperplane that separates classes. However, its linear decision boundary might not have been optimal for the inherent complexity of aroma data, resulting in higher misclassification rates and a slightly reduced balance between false positives and negatives.

The confusion matrix is presented to provide a more detailed evaluation of the models' performance. By examining the confusion matrix, it is possible to identify the overall accuracy and the specific misclassifications made by each model, such as false positives and false negatives. The following confusion matrices illustrate the performance of each model in terms of true positives, false positives, true negatives, and false negatives for the aroma classification task.

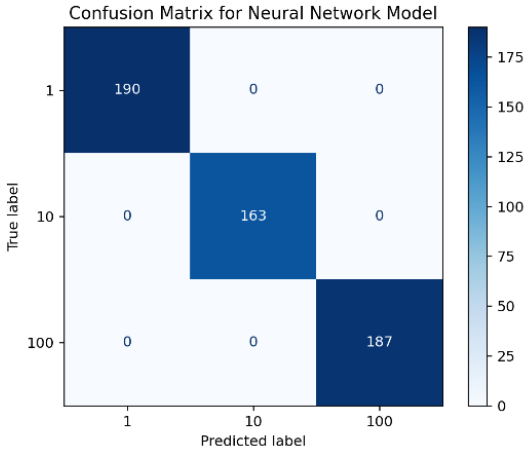

Figure 6. Confusion Matrix of Neural Network

Figure 6 presents the confusion matrix of the model or algorithm with the highest accuracy, the NN model with accuracy of 98.89%. To make AI and e-nose learning more accessible, we developed a trainer kit, e-Trainose, which simplifies AI training and sensor analysis. Figure 7 presents the physical design of the e-Trainose trainer kit, showcasing its compact form, integrated gas sensors, and essential components for aroma identification. This trainer kit provides a hands-on approach for users to interact with gas sensors and AI models, making it suitable for both educational and research purposes.

Figure 7. The e-Trainose Trainer Kit

In addition to its hardware design, e-Trainose is equipped with an interactive user interface (UI) developed using PyQt and Python. The UI provides an intuitive and user-friendly experience for researchers and students, allowing seamless interaction with the AI model and real-time sensor data. Figure 8 and Figure 9 illustrate the UI layout and its interactive components, enabling users to navigate through different functionalities with ease.

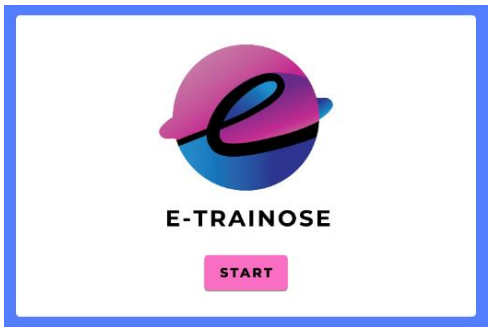

Figure 8. Startup Screen of The e-Trainose UI

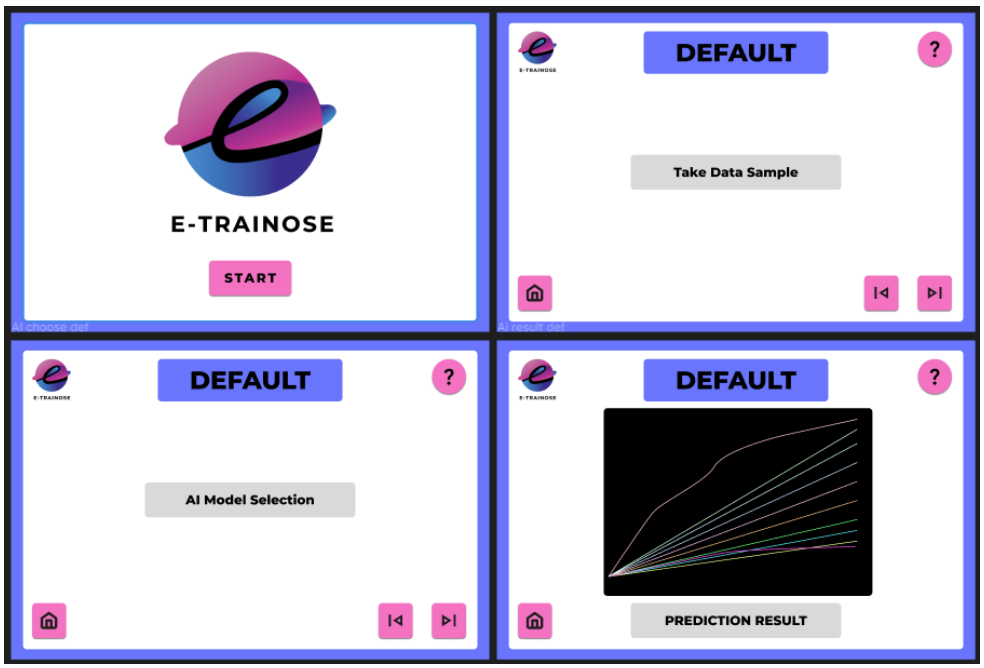

Figure 9. UI Layout During Operation.

In this study, the performance of each algorithm can be analyzed through evaluation metrics, allowing researchers to gain deeper insights into how each algorithm operates. Additionally, the findings from this research may provide strategies for improving the algorithms, which can be applied in future studies.

5. CONCLUSION

This study addressed the challenge of accurately classifying aroma samples using sensor-based data, proposing e-Trainose as an AI-driven solution. The system was developed to enhance aroma classification by integrating machine learning algorithms, namely Neural Network (NN), Support Vector Machine (SVM), and Random Forest. It was compared against a simpler threshold-based method.

Experimental results demonstrated that the Neural Network model achieved the highest classification performance, with an accuracy of 98.89%, precision of 99.10%, recall of 99.10%, and F1-score of 99.10%. The Random Forest model was followed closely, with an accuracy of 97.67%, indicating its strong generalization capability. Meanwhile, the SVM model exhibited lower performance with an accuracy of 94.78%, highlighting its limitations in handling complex aroma data patterns. In contrast, despite its simplicity, the threshold-based method achieved only 80.12% accuracy, confirming the significant advantage of AI-based approaches in aroma classification.

These findings underscore the effectiveness of AI-driven classification, particularly Neural Network, in improving the accuracy and reliability of aroma detection systems. Future research may focus on optimizing algorithmic performance, developing hybrid models, and addressing environmental factors such as temperature and humidity variations to enhance the robustness of e-Trainose further.

ACKNOWLEDGMENTS

This research would not have been possible without the generous support and funding provided by the Asian Development Bank under contract number 2397/UN1/DPU/HK.08.00/2024. Their financial assistance played a crucial role in enabling the successful completion of this project.

We would also like to express our sincere gratitude to PT Stechoq Robotika Indonesia, our partner in this endeavor, for their collaboration, expertise, and continuous support throughout the research process. Their contributions have been invaluable in advancing the objectives of this study.

Furthermore, we extend our deepest appreciation to Rifda Hakima Sari, Muhammad Hasani, Zaidan Hakim, Diandra Rizqi Yodatama, and Joshua Tito Amael for their guidance, insights, and unwavering support throughout this project. Their expertise and dedication have been instrumental in refining our methodologies and enhancing the overall quality of this research.

Lastly, we extend our appreciation to everyone involved in this research, whose hard work and contributions have been essential in bringing this project to fruition.

REFERENCES

[1] A. Zompanti, S. Grasso, M. Santonico, and G. Pennazza, "A Sensor System for Non-Destructive Monitoring of Food Ripening Processes," IEEE, 2020.

[2] J. Chilo, J. Pelegri-Sebastia, M. Cupane, and T. Sogorb, "E-nose application to food industry production," Feb. 2016.

[3] M. Adamek, M. Buran, M. Reznicek, A. Adamkova, M. Buckova, and J. Matyas, "The E-nose for Orientation Screening in the Food Industry," in *Proceedings of the International Spring Seminar on Electronics Technology*, IEEE Computer Society, May 2021. doi: 10.1109/ISSE51996.2021.9467518.

[4] M. L. MENDÉZ and V. PREEDY, *Electronic noses and tongues in food science*. London: Academic Press, 2016.

[5] M. Adamek *et al.*, "Sensor Systems for Detecting Dough Properties Fortified with Grape Pomace and Mealworm Powders," *Sensors*, vol. 20, no. 12, p. 3569, Jun. 2020, doi: 10.3390/s20123569.

[6] M. Peris and L. Escuder-Gilabert, "Electronic noses and tongues to assess food authenticity and adulteration," *Trends Food Sci Technol*, vol. 58, pp. 40–54, 2016, doi: https://doi.org/10.1016/j.tifs.2016.10.014.

[7] S. H. Saeed, Z. Abbas, and B. Gopal, "Experimental Use of Electronic Nose for Analysis of Volatile Organic Compound (VOC)," Aligarh, Mar. 2009. doi: 10.1109/MSPCT.2009.5164187.

[8] D. James, S. M. Scott, Z. Ali, and W. T. O'Hare, "Chemical Sensors for Electronic Nose Systems," *Microchimica Acta*, vol. 149, no. 1, pp. 1–17, 2005, doi: 10.1007/s00604-004-0291-6.

[9] L. Li *et al.*, "Construction and Simple Application of a Primary Electronic Nose System," Fukuoka: IEEE, May 2019. doi: 10.1109/ISOEN.2019.8823340.

[10] N. Tayagornkul and N. Nimsuk, "Analysis of Chicken Meat Degradation using Electronic Nose Technology," Chiang Rai: IEEE, Jul. 2018. doi: 10.1109/ECTICon.2018.8619920.

[11] H. Zhang and H. Tai, "Several Data Analysis and Processing of Electronic Nose Data Preprocessing Subsystem," in *IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2021, pp. 1097–1100. doi: 10.1109/IAEAC50856.2021.9390785.

[12] Y.-T. Liu and K.-T. Tang, "A Minimum Distance Inliers Probablity (MDIP) Feature Selection Method To Enhance Gas Classification For An Electronic Nose System," Fukuoka: IEEE, May 2019. doi: 10.1109/ISOEN.2019.8823336.
[13] "Export of All Form of Coffee by Exporting Countries to All Destinations." Accessed: Mar. 23, 2025. [Online]. Available: https://ico.org/documents/prices/MTS-0324_T1.pdf
[14] "Home | ICEA." Accessed: Mar. 23, 2025. [Online]. Available: http://gaeki.or.id/en/home/
[15] C. Bunn, P. Läderach, O. Ovalle Rivera, and D. Kirschke, "A bitter cup: climate change profile of global production of Arabica and Robusta coffee," *Clim Change*, vol. 129, no. 1–2, pp. 89–101, Dec. 2014, doi: 10.1007/s10584-014-1306-x.
[16] H. K. Patel, *The Electronic Nose: Artificial Olfaction Technology*. New Delhi, 2014. [Online]. Available: http://www.springer.com/series/3740
[17] S. Baskara, D. Lelono, and T. W. Widodo, "Pengembangan Hidung Elektronik untuk Klasifikasi Mutu Minyak Goreng dengan Metode Principal Component Analysis," *IJEIS*, vol. 6, no. 2, pp. 221–230, Oct. 2016.
[18] Y. Raut and S. N. Chaudhri, "Comprehensive Overview of Gas and Bio-Sensors Enabling Artificial Intelligence in Healthcare," in *2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry, IDICAIEI 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/IDICAIEI61867.2024.10842798.
[19] G. Jasinski, P. Kalinowski, L. Wozniak, P. Koscinski, and P. Jasinski, "An electronic nose based on the semiconducting and electrochemical gas sensors," Warsaw, Sep. 2017. doi: 10.23919/EMPC.2017.8346895.
[20] S. S. Chowdhury, B. Tudu, R. Bandyopadhyay, and N. Bhattacharyya, "Portable Electronic Nose System for Aroma Classification of Black Tea," Kharagpur: IEEE, Dec. 2008. doi: 10.1109/ICIINFS.2008.4798403.
[21] Y.-Q. Jing, Q.-H. Meng, P.-F. Qi, and M. Zeng, *A novel olfactory neural network for classification of Chinese liquors using electronic nose*. Busan: IEEE, 2015. doi: 10.1109/ICSENS.2015.7370675.
[22] A. Tan, S. Duan, M. Chen, K. Ding, C. Li, and L. Wang, "High-Performance Sensing: Application of FPGA-based Winograd CNN Accelerator in Electronic Nose Systems," in *2024 5th International Conference on Artificial Intelligence and Computer Engineering, ICAICE 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 712–715. doi: 10.1109/ICAICE63571.2024.10863924.
[23] P. Hensman and D. Masko, "The Impact of Imbalanced Training Data for Convolutional Neural Networks," 2015.
[24] C. Bax, S. Prudenza, G. Gaspari, L. Capelli, F. Grizzi, and G. Taverna, "Drift compensation on electronic nose data for non-invasive diagnosis of prostate cancer by urine analysis," *iScience*, vol. 25, no. 1, Jan. 2022, doi: 10.1016/j.isci.2021.103622.
[25] J. Han, H. Li, J. Cheng, X. Ma, and Y. Fu, "Advances in metal oxide semiconductor gas sensor arrays based on machine learning algorithms," *J. Mater. Chem. C*, vol. 13, no. 9, pp. 4285–4303, 2025, doi: 10.1039/D4TC05220J.
[26] V. Somaudon and T. Kerdcharoen, "Unlocking the Aroma Profiles of Coffee Roasting Levels with an Electronic Nose Coupled with Machine Learning," in *Proceedings - 21st International Joint Conference on Computer Science and Software Engineering, JCSSE 2024*, Institute of Electrical and Electronics Engineers Inc., Jun. 2024, pp. 678–681. doi: 10.1109/JCSSE61278.2024.10613715.
[27] E. Bona and R. S. dos Santos Ferreira da Silva, "Chapter 4 - Coffee and the Electronic Nose," in *Electronic Noses and Tongues in Food Science*, M. L. Rodríguez Méndez, Ed., San Diego: Academic Press, 2016, pp. 31–38. doi: https://doi.org/10.1016/B978-0-12-800243-8.00004-4.
[28] C. Cortes, V. Vapnik, and L. Saitta, "Support-Vector Networks Editor," Kluwer Academic Publishers, 1995.
[29] C. Distante, N. Ancona, and P. Siciliano, "Support vector machines for olfactory signals recognition."
[30] S.-T. Kim, I.-H. Choi, and H. Li, "Identification of multi-concentration aromatic fragrances with electronic nose technology using a support vector machine," *Anal. Methods*, vol. 13, no. 40, pp. 4710–4717, Oct. 2021, doi: 10.1039/d1ay00788b.
[31] J. Brezmes, B. Ferreras, E. Llobet, X. Vilanova, and X. Correig, "Neural network based electronic nose for the classification of aromatic species," *Anal Chim Acta*, vol. 348, no. 1, pp. 503–509, 1997, doi: https://doi.org/10.1016/S0003-2670(97)00202-X.
[32] B. Wailzer, J. Klocker, G. Buchbauer, G. Ecker, and P. Wolschann, "Prediction of the aroma quality and the threshold values of some pyrazines using artificial neural networks," *J. Med. Chem.*, vol. 44, no. 17, pp. 2805–2813, Aug. 2001.
[33] K. Saadi, M. Korichi, V. Gerbaud, T. Talou, and P. Floquet, "Chapter 92 - Classification of the Aroma Quality of Pyrazine Derivatives using Random Forest Tree Technique," in *Flavour Science*, V. Ferreira and R. Lopez, Eds., San Diego: Academic Press, 2014, pp. 499–502. doi: https://doi.org/10.1016/B978-0-12-398549-1.00092-1.
[34] V. Distefano *et al.*, "Random forest for classifying odor emission sources * Random forest per la classificazione delle sorgenti emissive odorigene," Mar. 2024.
[35] F. C. Cynthiarani, Lelono, and D. U. K. Putri, "Performance of feature extraction method combination in arabica coffee roasting classification," in *IOP Conference Series: Earth and Environmental Science*, Institute of Physics, 2024. doi: 10.1088/1755-1315/1356/1/012029.
[36] J. Yan *et al.*, "Electronic nose feature extraction methods: A review," Nov. 02, 2015, *MDPI AG*. doi: 10.3390/s151127804.
[37] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf Process Manag*, vol. 45, no. 4, pp. 427–437, Jul. 2009, doi: 10.1016/j.ipm.2009.03.002.

[38] O. Natan, A. I. Gunawan, and B. S. B. Dewantara, "A New Feature Extraction Algorithm to Extract Differentiate Information and Improve KNN-based Model Accuracy on Aquaculture Dataset," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 9, no. 3, pp. 999-1007, May. 2019.
[39] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proceedings of the 23rd International Conference on Machine Learning*, in ICML '06. New York, NY, USA: Association for Computing Machinery, 2006, pp. 233–240. doi: 10.1145/1143844.1143874.
[40] A. A. G. Y. Pramana, H. M. Zidan, M. F. Maulana, and O. Natan, "ESDS: AI-Powered Early Stunting Detection and Monitoring System using Edited Radius-SMOTE Algorithm," in *Proceedings of the ACM International Conference on Computational Intelligence and Intelligent Systems (CIIS)*, Nagoya, Japan, Nov. 2024, pp. 111-118.
[41] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit Lett*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.
[42] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, Jan. 2020, doi: 10.1186/s12864-019-6413-7.
[43] O. Natan, A. I. Gunawan, B. S. B. Dewantara, and J. Ispianto, "Microorganism Estimation in a Shrimp Pond using Gaussian Process Regressor and Gradient Tree Boosting," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 3, pp. 1-10, Jun. 2020.

BIOGRAPHY OF AUTHORS

Jazi Eko Istiyanto received his Bachelor's degree in Physics from Gadjah Mada University, Indonesia, in 1986, his Master's degree in Computer Science from the University of Essex, UK, in 1988, and his Doctoral degree in Electronic Systems Engineering from the University of Essex in 1995. He has been a lecturer at Gadjah Mada University since 1988, and served as Chairman of the Indonesia Nuclear Energy Regulatory Agency (BAPETEN) from 2014 to 2021. His research focuses on Cyber-Physical Systems, their security, and FPGA design automation. Dr. Istiyanto has been awarded the ASEAN Engineer title and is an active member of various professional societies, including IEEE and APTIKOM.

Danang Lelono received his Master's degree in Electro/Technics (2000) and his undergraduate degree in Electronics & Instrumentation (1994) from Gadjah Mada University, Indonesia. He earned his Ph.D. in Physics from the same university in 2017. Since 1998, he has been a lecturer in the Department of Computer Science and Electronics at Universitas Gadjah Mada. His research interests include Instrumentation & Control, Smart Sensors, Intelligent Systems, Embedded Systems, and Emergency Nursing. He is actively engaged in teaching and research, contributing to advancements in computer science, electronics, and control systems.

Oskar Natan received the BASc degree in electronics engineering and the MAEng degree in electrical engineering from Politeknik Elektronika Negeri Surabaya, Indonesia, in 2017 and 2019, respectively, and the PhD(Eng) degree in computer science and engineering from Toyohashi University of Technology, Japan, in 2023. Since January 2020, he has been affiliated with the Department of Computer Science and Electronics, Universitas Gadjah Mada, Indonesia, first as a Lecturer and currently an Assistant Professor. He has been a reviewer/TPC for some reputable journals and conferences. His research interests include sensor fusion, deep learning, hardware acceleration, and end-to-end systems.

Shafa Khamila is currently pursuing a Bachelor's degree in Electronics and Instrumentation at Universitas Gadjah Mada, expected to graduate in 2026. She currently serves as an intern project manager in PT Stechoq Robotika Indonesia. Shafa is also the Manager of the Traditional Dancing Robot Team Alfan, where she leads technical coordination and team management. Shafa is actively learning and has strongly interested in hardware design, sensor integration, and machine learning algorithms. She is advancing her skills in data science, focusing on machine learning, Python programming, and data analysis.

Hafizha Adhiyant is currently pursuing a Bachelor's degree in Electronics and Instrumentation at Universitas Gadjah Mada, expected to graduate in 2026. He currently serves as an intern hardware engineer in PT Stechoq Robotika Indonesia. He is also involved in several technical roles, including as supervisor at Bimasakti Racing Team UGM, responsible for researching new electrical systems for FSAE cars. Hafizha has earned notable achievements, including 8th place in Formula SAE Italy 2023 and 4th place in KRBAI Regional 2023. He is also a recipient of the Pertamina Foundation Sobat Bumi Scholarship. Hafizha is quite skilled in areas such as programming, 3D design, PCB design, and electronics control systems.

Ikhlasul Amal Abda'i, a fifth-semester Electronics and Instrumentation student at Universitas Gadjah Mada, is an ambitious and versatile individual with expertise in control systems, programming, robotics, 3D design, and AI. He has demonstrated excellence in academic and competitive projects, including drone development and Chem-E-Car competitions. Known for his adaptability, creativity, and dedication, Ikhlas aspires to contribute innovative solutions in technology and engineering while continually expanding his knowledge and skills.

Ilyaz Raukhillah Adzaqi is a creative and adaptable individual who is passionate about learning and problem-solving and has expertise in software programming, hardware programming, and PCB design. Currently pursuing a degree in Electronics and Instrumentation at Universitas Gadjah Mada, Ilyaz has a strong background in robotics and programming. He is involved with various organizations such as ORCA GMRT UGM, contributing to mechanical and programming divisions, and the ELINS Research Club, focusing on robotics research.