Optimization of Gas Sensor Array in Electronic Nose Using Whale Optimization Algorithm for Tuberculosis Detection

2025 International Conference on Computer Sciences, Engineering, and Technology Innovation (ICoCSETI) | 979-8-3315-0861-6/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICoCSETI63724.2025.11020012

Dava Aulia
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
davaaulia2000@gmail.com

Riyanarto Sarno
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
riyanarto@its.ac.id

Kelly Rossa Sungkono
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
kelly@its.ac.id

Muhammad Rivai
Department of Electrical Engineering
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
muhammad_rivai@its.ac.id

Sheva Aulia
Department of Information System
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
shevaaulia2004@gmail.com

Dwi Sunaryono
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
dwi@its.ac.id

***Abstract*—A lung infected with tuberculosis contains volatile organic compounds (VOCs) caused by Mycobacterium tuberculosis. An exhaled breath analysis approach can detect biomarkers of tuberculosis. An electronic nose can be proposed as a device for the early diagnosis of tuberculosis. The primary problem with this device is to specify the optimal number of gas sensors while maintaining high performance. The whale optimization algorithm (WOA) is an optimization method that uses the hunting mechanism of humpback whales, including bubble-net feeding and spiral movement. This algorithm is straightforward to implement, uncomplicated, and converges quickly. This study aimed to determine the optimal number of gas sensors in an electronic nose for differentiating between patients with tuberculosis and healthy controls through exhaled breath using WOA. The implementation was performed using Python programming and Jupyter Notebook for data processing, optimization, and classification. The experimental results demonstrate that the WOA efficiently selected the optimal gas sensor combinations in a short optimization time. Furthermore, two gas sensors were selected, and the support vector machine (SVM) successfully distinguished between tuberculosis and healthy controls with accuracy, precision, recall, and f1-score above 98%.**

Keywords*—*diseases, electronic nose, tuberculosis, whale optimization algorithm.

I. Introduction

Tuberculosis is a respiratory condition caused by Mycobacterium tuberculosis. It can be dispatched through droplets and causes persistent cough [1]. The world health organization (WHO) reported that 10 million people develop tuberculosis annually, with 845,000 cases occurring in Indonesia [2]. Tuberculosis can be diagnosed through chest imaging, blood tests, and sputum smear [3]. However, most of these methods are impractical, expensive, and uncomfortable. In addition, exhaled breath analysis is a safe, inexpensive, and rapid method for detecting biomarkers of tuberculosis [4]. A lung infected with tuberculosis contains volatile organic compounds (VOCs) derived from the bacteria [5]. Furthermore, gas chromatography is a costly gas analysis technique that requires qualified operators. Thus, an electronic nose system, which combines gas sensors and machine learning, can address these issues and be used to diagnose tuberculosis [4], chronic obstructive pulmonary disease (COPD) [6], and asthma [7] through exhaled breath.

An electronic nose with an appropriate number of gas sensors can provide effective and accurate breath gas measurements [8]. Therefore, a gas sensor selection approach is required to eliminate gas sensors that do not perform significantly at the classification stage. There are two popular gas sensor selection methods, namely the firefly algorithm (FA) [8] and genetic algorithm (GA) [7]. The FA utilizes the behavior of fireflies, where brightness represents the quality of the solution. This method is effective and efficient because its theory is simple and has fewer parameters, which makes it easier to adjust. However, its computational architecture makes the process more time-consuming [9]. In addition, GA uses a genetic evolution approach through selection, crossover, and mutation, in which better chromosomes tend to survive and produce the best solutions [7]. This method can solve complex problems; however, it involves more parameters and requires careful adjustment.

The whale optimization algorithm (WOA) is an optimization algorithm inspired by the hunting mechanism of humpback whales, which utilizes bubble-net feeding and spiral movement [10], [11]. The whale generates bubbles around its prey, representing the optimal solution, to narrow the search area and focus on its efforts. Thereafter, the whale moves in a spiral to expand the search area, enabling broader exploration and avoiding getting stuck in a local optimum, thus helping find the optimal solution in the optimization process. The advantages of this method include its fast convergence, simplicity, and ease of implementation, which make it ideal for addressing feature selection problems [12].

The novelty of this study was the application of the WOA method to identify the optimal number of gas sensors in an electronic nose system to differentiate tuberculosis from healthy controls through exhaled breath. The comparative evaluations were performed using two other optimization methods, namely FA and GA. The paper is organized as follows: Section 1 presents the introduction, Section 2 explains the methodology, Section 3 analyzes the results, and Section 4 offers the conclusions and suggests future work.

II. Materials and Methods

A. Subjects and Proposed Research Design

Exhaled breath samples were collected at Dr. Soetomo General Hospital in Surabaya, East Java, Indonesia. This

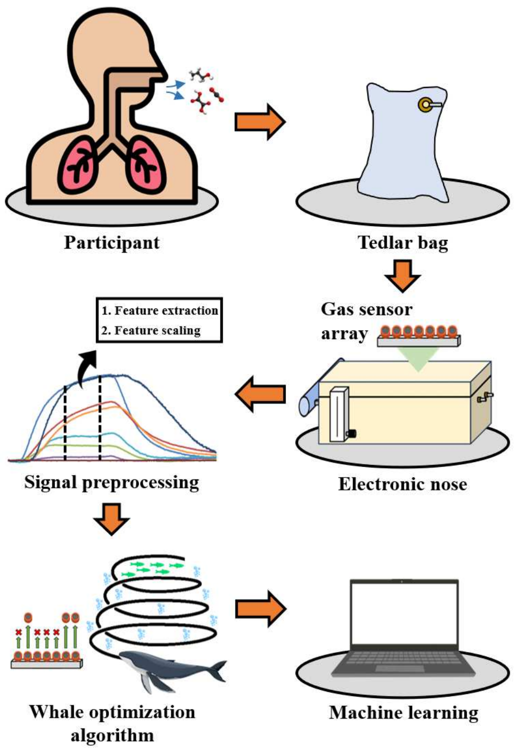

Fig. 1. Research design of the study.

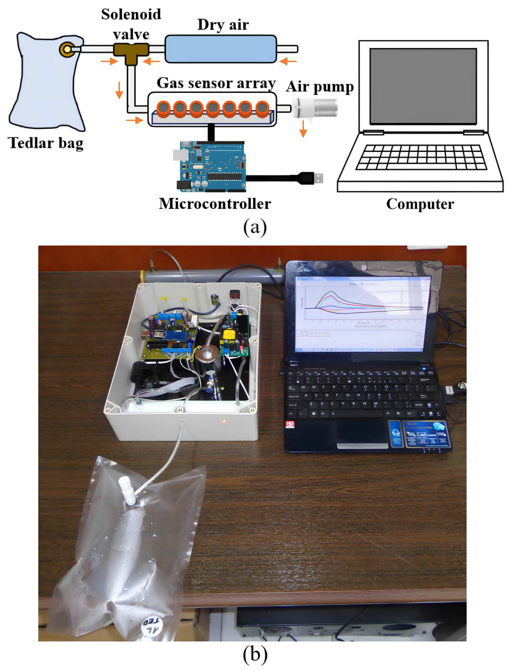

Fig. 2. Electronic nose configuration: (a) schematic and (b) actual implementation.

study involved six patients with tuberculosis and six healthy controls. In addition, all research criteria, with participant categorization, were determined by a pulmonary doctor.

Fig. 1 presents the research design of the study. The participants provided their exhaled breath samples in a 1-L Tedlar bag. An electronic nose was used as the instrument for measurement. The response of each gas sensor generates different curves. Signal preprocessing involves feature extraction and feature scaling techniques. A support vector machine (SVM) was used as both a machine learning model for classification and a WOA fitness function for gas sensor array optimization. The fewest number of gas sensors with the highest performance were selected. The machine learning evaluation was performed using standard metrics.

TABLE I. GAS SENSORS IN AN ELECTRONIC NOSE

Sensor	Gas
MQ-7 (S00)	Hydrogen and carbon monoxide
MQ-8 (S01)	Hydrogen and alcohol
MQ-131 (S02)	Ozone
MQ-136 (S03)	Hydrogen sulfide
MQ-137 (S04)	Ammonia
MQ-138 (S05)	VOCs
TGS4161 (S06)	Carbon dioxide

B. Electronic Nose

Fig. 2 illustrates the electronic nose configuration. The electronic nose was composed of a gas sensor array, a microcontroller, and a computer, as shown in Fig. 2a. The electronic circuitry was mounted on a printed circuit board. The high sensitivity of metal oxide semiconductor gas sensors, which is one of their key advantages [13], [14], was selected in this study. Table I lists the gas sensors used in the experiments, including S00, S01, S02, S03, S04, S05 and S06. In addition, these gas sensors are often used to detect lung disease [4], [6]. The gas sensor array was placed in a 240 mL chamber to prevent interference. The realization of the electronic nose is demonstrated in Fig. 2b.

The sample measurement began by purging the gas sensor array with dry air for 15 seconds. Subsequently, the gas from the bag was analyzed for 255 seconds, followed by a 200-second cleaning phase to prepare the following measurements. The gas sensor signal output voltage was converted to digital data, transferred to a computer, and saved as an Excel file. The dataset included various response curves, which were subsequently analyzed on a computer.

C. Signal Preprocessing

This step implements feature extraction and feature scaling techniques. The feature extraction includes five features, namely mean $(\bar{x}_n)$, maximum $(x_n^{\max})$, median (x_n^{median}), minimum $(x_n^{\min})$, and standard deviation (σ), as expressed in (1) to (5), respectively.

$$\bar{x}_n = \frac{\sum_{T=1}^{k} x_n^T}{k} \tag{1}$$

$$x_n^{\max} = \max_{T=1,\ldots,k} x_n^T \tag{2}$$

$$x_n^{\text{median}} = \text{median}\,\{x_n^T\}_{T=1}^{k} \tag{3}$$

$$x_n^{\min} = \min_{T=1,\ldots,k} x_n^T \tag{4}$$

$$\sigma = \sqrt{\frac{1}{k}\sum_{T=1}^{k}(x_n^T - \bar{x}_n)^2} \tag{5}$$

where x_n^T is the response value of the n-th gas sensor at time T, k is the number of sample data, and n is the number of gas sensors. The five feature scaling techniques are the standard scaler, robust scaler, min-max scaler, max abs scaler, and l2 normalization, as defined in (6) to (10), each.

$$\text{Standard scaler} = \frac{X-\mu}{\sigma} \tag{6}$$

$$\text{Robust scaler} = \frac{X-\text{median}(X)}{\text{IQR}(X)} \tag{7}$$

$$\text{Min} - \text{max scaler} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (8)$$

$$\text{Max abs scaler} = \frac{X}{|X_{\max}|} \quad (9)$$

$$\text{L2 normalization} = \frac{X}{\|X\|_2} \quad (10)$$

where X is the data matrix of each feature calculated for each gas sensor.

D. Support Vector Machine (SVM)

SVM uses a decision boundary called a hyperplane to separate data into separate classes [15]. This machine-learning model has a kernel that can affect the hyperplane characteristics [16]. In this study, SVM was applied using the scikit-learn library. The parameters used were a cost function value of 1.0, gamma (γ) with a value of scale, and the radial basis function (RBF) kernel, as depicted in (11).

$$K(x, x') = exp(-\gamma\|x - x'\|^2) \quad (11)$$

where $\|x - x'\|$ is the Euclidean squared distance.

E. Whale Optimization Algorithm (WOA)

The gas sensor array optimization steps using WOA are shown in Fig. 3. The mathematical explanation is as follows:

1) Fitness function: The SVM evaluates the quality of each whale position as a candidate solution.

a) Position of the whale: Each whale position is a continuous vector arranged in an array, as described in (12).

$$\text{position} = [p_1, p_2, \ldots, p_d] \quad (12)$$

where p_j are the elements of the continuous position vector, with $p_j \in [0,1]$, and d is the number of dimensions in the search space.

b) Selected features condition: The selected features are those with $p_j > 0.5$, which is expressed in (13).

$$\text{selected\_features} = \{j | p_j > 0.5\} \quad (13)$$

If no features are selected ($\text{selected\_features} = \emptyset$), the fitness is assigned a high penalty value of 1.

c) Accuracy evaluation: The selected features were evaluated using stratified k-fold cross-validation with SVM, and the average accuracy was calculated, as defined in (14).

$$\text{average\_accuracy} = \frac{1}{k}\sum_{i=1}^{k} \text{accuracy\_score}_i \quad (14)$$

where k is the number of folds.

d) Fitness value: The fitness value is shown in (15).

$$\text{fitness} = 1 - \text{average\_accuracy} \quad (15)$$

2) Bubble-net feeding and spiral movement: Two whale hunting behaviors used to solve optimization problems.

a) Bubble-net feeding (exploitation): When $|A| < 1$, the whale moves to the best whale position, given by (16).

$$X_i(t+1) = X_{best}(t) - A \times |C \times X_{best}(t) - X_i(t)| \quad (16)$$

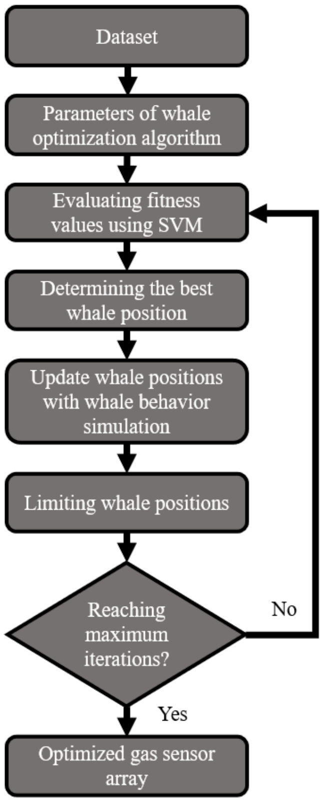

Fig. 3. WOA-based gas sensor array optimization.

TABLE II. WOA PARAMETERS

Parameter	Explanation
Number of whales	30
b	1.0
p	0.5
Upper	10
Lower	-10
Number of iterations	300

where $X_{best}(t)$ is the best whale position at iteration t, A is a coefficient calculated as $A = 2ar - a$, with a decreasing with iterations, and r is a random number between 0 and 1, C is a random coefficient between 0 and 2, $X_i(t)$ is the position of the whale i at iteration t. If $|A| \geq 1$, the whale moves toward a random position, as stated in (17).

$$X_i(t+1) = X_{random}(t) - A \times |C \times X_{random}(t) - X_i(t)| \quad (17)$$

where X_{random} is the randomly selected whale position.

b) Spiral movement (exploration): The whale moves in a spiral to find better solutions around the best position, as outlined in (18).

$$X_i(t+1) = D(t) \times e^{bp} \times \cos(p \times 2\pi) + X_{best}(t) \quad (18)$$

where $D(t)$ is the distance between whale i and the best whale at iteration t, namely $D(t) = |X_{best}(t) - X_i(t)|$. b is a coefficient that controls the shape of the spiral, p is a control parameter for the spiral.

3) Position constraints: After updating the position, the position is constrained within the range of [lower, upper], as presented in (19).

$$X_i(t+1) = \text{clip}(X_i(t+1), \text{lower}, \text{upper}) \quad (19)$$

TABLE III. FA PARAMETERS

Parameter	Explanation
Number of fireflies	30
Initial alpha	0.5
Beta0	1.0
Gamma	0.01
Theta	0.97
Number of iterations	300

TABLE IV. GA PARAMETERS

Parameter	Explanation
Number of chromosomes	30
Tournament size	5
Mutation rate	0.25
Crossover rate	0.25
Number of iterations	300

Table II shows the WOA parameters employed in this study, while Table III and Table IV provide the FA and GA parameters for comparative analysis, respectively.

F. Evaluation of the Machine Learning Model

The classification model was evaluated using four metrics. The key metrics were accuracy, precision, recall, and f1-score. These metrics were derived from true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values, defined in (20) to (23), respectively.

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+TN+FP} \tag{20}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{21}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{22}$$

$$\text{F1}-\text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{23}$$

G. Principal Component Analysis (PCA)

PCA can be used to visualize the data distribution for each class [17]. The total variance in PCA is used to measure the total information or spread of data that can be explained by the PC. This variance is calculated as the sum of the eigenvalues (λ) of the data covariance matrix, stated in (24).

$$\text{Total variance} = \sum_{i=1}^{d} \lambda_i \tag{24}$$

III. EXPERIMENTAL RESULTS

A. Experiment Tools and Setup

The research computer had the specifications of 12th gen Intel Core i5-12400F 2.50 GHz, 32 GB DDR4 RAM, NVIDIA GeForce RTX 4060 8 GB GDDR6, and Windows 11 Pro 64 Bit. All experimental programs used stratified k-fold cross-validation with 2 folds, shuffle was set to true, and random state values ranging from 0 to 100, resulting in 101 randomizations of the dataset to produce reliable results. The model's performance was assessed using the average (AVG) of ten trials across four evaluation metrics.

B. Dataset Preparation

Table V presents the number of subjects, comprising six samples from patients with tuberculosis and six from healthy controls. Next, these samples formed a binary dataset with a total of twelve exhaled breath gas samples. Fig. 4 illustrates the response curves of the gas sensors for each category, showcasing distinct response patterns that facilitate analysis. The dataset used in this study was derived from sensor responses calibrated using a dynamic baseline, which was applied to minimize the effect of sensor drift. Fig. 5 shows the region determination for feature extraction. Four regions with a duration of 15 seconds were involved: Region 1 was from 190 to 204, Region 2 was from 205 to 219, Region 3 was from 220 to 234, and Region 4 was from 235 to 249.

TABLE V. NUMBER OF SUBJECTS IN EACH CATEGORY

Category	Total Subject
Tuberculosis	6
Healthy control	6
All	12

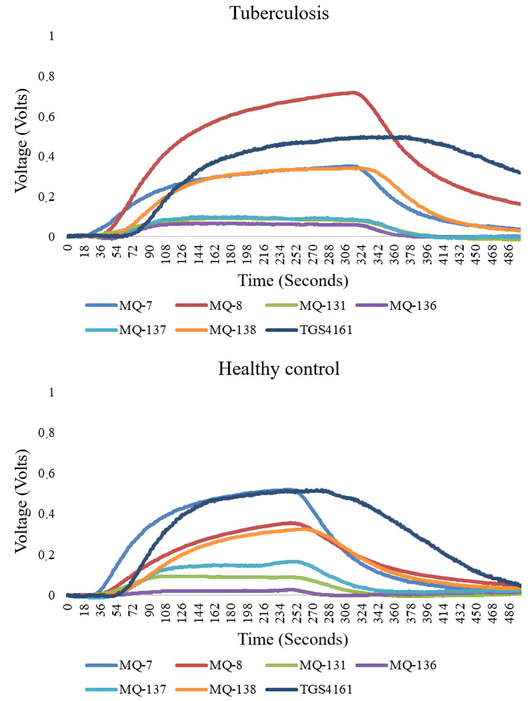

Fig. 4. Gas sensor response curves of exhaled breath from patients with tuberculosis and healthy controls.

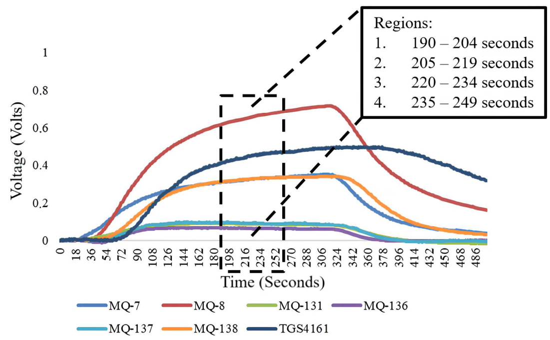

Fig. 5. Region determination strategy for feature extraction.

Table VI demonstrates the classification results for each gas sensor response region and feature variation. Each region provided the average performance calculated from each feature, and Region 1 was selected because it ranked first with a performance of 81.41%. In addition, the standard deviation feature was preferred because it achieved the

TABLE VI. SVM CLASSIFICATION OF GAS SENSOR RESPONSES IN EACH REGION AND FEATURE VARIATIONS.

Percentage (%)					
Region 1					
Metric	***Mean***	***Maximum***	***Median***	***Minimum***	***Standard deviation***
Accuracy	79.37	79.37	79.37	79.37	85.31
Precision	84.31	84.31	84.31	84.31	89.65
Recall	79.37	79.37	79.37	79.37	85.31
F1-score	77.81	77.81	77.81	77.81	84.40
Region 2					
Metric	***Mean***	***Maximum***	***Median***	***Minimum***	***Standard deviation***
Accuracy	78.55	78.38	70.28	79.04	61.30
Precision	83.46	83.32	68.63	84.00	62.62
Recall	78.55	78.38	67.86	79.04	61.30
F1-score	76.96	76.78	62.77	77.47	58.15
Region 3					
Metric	***Mean***	***Maximum***	***Median***	***Minimum***	***Standard deviation***
Accuracy	78.30	78.22	78.30	78.30	78.30
Precision	83.15	83.09	83.15	83.15	81.40
Recall	78.30	78.22	78.30	78.30	78.30
F1-score	76.73	76.64	76.73	76.73	77.04
Region 4					
Metric	***Mean***	***Maximum***	***Median***	***Minimum***	***Standard deviation***
Accuracy	77.56	77.48	77.64	77.15	73.51
Precision	82.00	81.96	82.15	81.57	73.82
Recall	77.56	77.48	77.64	77.15	73.51
F1-score	75.91	75.81	75.98	75.48	69.70

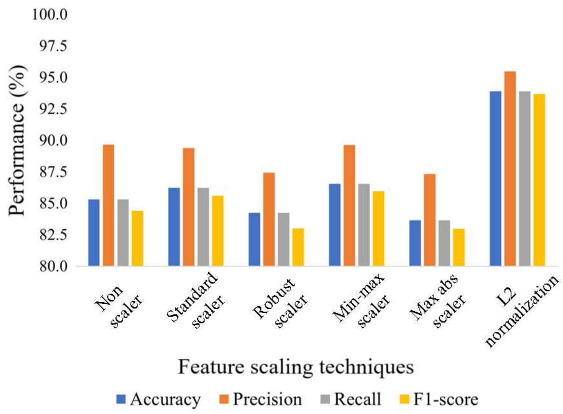

Fig. 6. Comparative results of all feature scaling techniques.

highest performance of 86.17%. Therefore, Region 1 and the standard deviation were selected for feature extraction due to their significant contribution to the classification phase.

The subsequent step was to implement all feature scaling techniques, as portrayed in Fig. 6. The standard scaler, robust scaler, min-max scaler, and max abs scaler did not significantly improve the performance. In addition, l2 normalization improved the performance from 86.17% to 94.24%. Therefore, this feature scaling technique was selected because it involves l2 calculations, which can normalize the dataset pleasingly. Fig. 7 visualizes the data distribution with l2 normalization using PCA. Each category can be separated nicely, with a total variance value of 90%.

The following step used different fold numbers in stratified k-fold cross-validation in the classification phase, with 3 to 5 folds. The evaluation results for classification with different numbers of folds are depicted in Fig. 8. It can be seen that increasing the number of folds did not improve performance. Therefore, two folds were maintained, and the

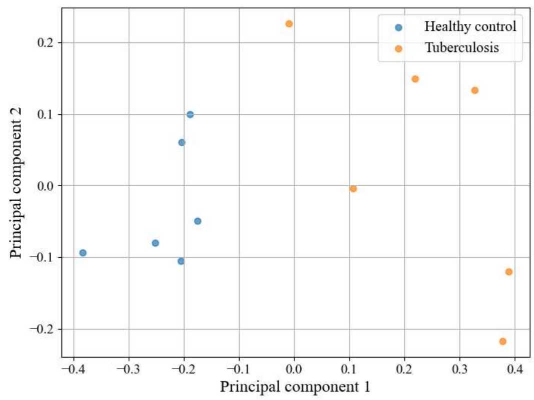

Fig. 7. Data distribution with l2 normalization using PCA visualization.

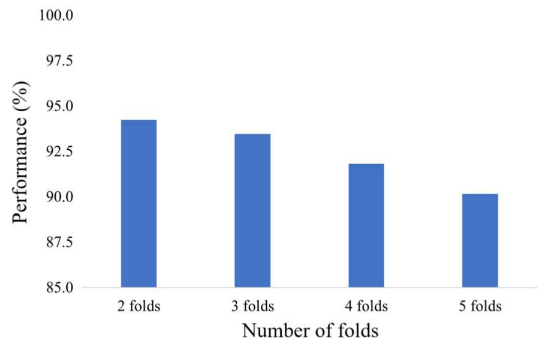

Fig. 8. Performance results for varying folds in stratified k-fold validation.

TABLE VII. COMPARATIVE RESULTS OF WOA, FA AND GA IN DETERMINING GAS SENSOR GROUPS.

Method	Group	Gas Sensors	Performance (%)
WOA	6	S00, S02, S03, S04, S05, S06	97.66
	5	S02, S03, S04, S05, S06	98.84
	4	S02, S04, S05, S06	99.15
	3	S02, S05, S06	99.84
	2	S05, S06	98.92
	AVG		98.88
FA	6	S00, S02, S03, S04, S05, S06	97.66
	5	S00, S02, S03, S05, S06	98.21
	4	S02, S04, S05, S06	99.15
	3	S02, S05, S06	99.84
	2	S05, S06	98.92
	AVG		98.76
GA	6	S00, S02, S03, S04, S05, S06	97.66
	5	S00, S02, S03, S05, S06	98.21
	4	S00, S03, S05, S06	98.83
	3	S02, S05, S06	99.84
	2	S05, S06	98.92
	AVG		98.69

final recapitulation results before gas sensor selection were Region 1, standard deviation as feature extraction, and two folds in stratified k-fold cross-validation.

The final stage involved applying the gas sensor selection method using WOA, with FA and GA used for comparison. With 101 randomizations in the dataset, each optimization method generated 505 sensor groups. Table VII presents the best sensor groups based on the performance values for each method, with WOA exhibiting the highest performance of 98.88%. In addition, to confirm the WOA's superiority, its speed in finding the optimal solution was assessed, defined in Table VIII. WOA completed the optimization in just 4.7 seconds, which was significantly faster than the FA and GA,

TABLE VIII. DURATION OF EACH OPTIMIZATION PROCESS.

Method	Time (Seconds)
WOA	4.7
FA	435.3
GA	37.2

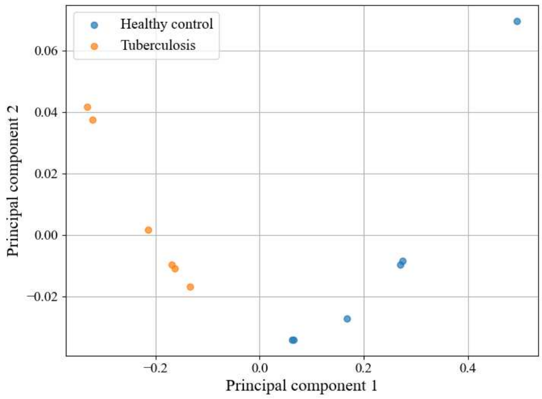

Fig. 9. Visualization of data distribution from two sensors using PCA.

with 435.3 and 37.2 seconds, respectively. These results highlighted the WOA's superior efficiency in selecting optimal gas sensors. In addition, MQ-138 and TG4161 were selected, achieving 98.92% performance, with MQ-138 capable of detecting VOCs for tuberculosis detection. Fig. 9 illustrates the PCA data distributions of these sensors, with well-separated classes and 100% total variance. Furthermore, the accuracy, precision, recall, and f1-score of these sensors were 98.84%, 99.31%, 98.84%, and 98.70%, respectively.

IV. CONCLUSION

This study optimized a gas sensor array in an electronic nose to detect tuberculosis using WOA. The electronic nose was integrated with seven metal oxide semiconductor gas sensors. The WOA identified the optimal number of gas sensors through the bubble-net feeding and spiral movement processes, which are the hunting mechanisms of humpback whales, while utilizing the SVM as the fitness function to evaluate the performance of the selected gas sensor combination. The main advantages of the WOA are straightforwardness, simplicity, and it achieves rapid convergence. The experimental results show that the WOA determined the optimal gas sensor combinations with the highest performance of 98.88%. The MQ-138 and TGS4161 gas sensors were selected as the optimal gas sensor array, delivering all metrics exceeding 98%. In addition, the WOA required only 4.7 seconds to complete the process, making it suitable for feature selection optimization problems.

Future research will aim to design a more compact electronic nose by minimizing the gas sensor chamber and control unit, collecting more samples, and applying different machine learning algorithms. Moreover, all methods will be deployed on a single-board device to enhance practicality.

ACKNOWLEDGMENT

This research was funded by Institut Teknologi Sepuluh Nopember (ITS) under Penelitian Kolaborasi Pusat and under the Project Scheme of the Publication Writing and Intellectual Property Rights (IPR) Incentive (Penulisan Publikasi dan Hak Kekayaan Intelektual/ PPHKI) Program 2024.

REFERENCES

[1] C. M. Gill, L. Dolan, L. M. Piggott, and A. M. McLaughlin, "New developments in tuberculosis diagnosis and treatment," *Breathe*, vol. 18, no. 1, pp. 1–15, 2022.

[2] D. Iskandar, A. A. Suwantika, I. S. Pradipta, M. J. Postma, and J. F. M. van Boven, "Clinical and economic burden of drug-susceptible tuberculosis in Indonesia: national trends 2017-19," *Lancet Global Health.*, vol. 11, no. 1, pp. e117–e125, 2023.

[3] T. N. Susilawati and R. Larasati, "A recent update of the diagnostic methods for tuberculosis and their applicability in Indonesia: a narrative review," *Medical Journal of Indonesia.*, vol. 28, no. 3, pp. 284–291, 2019.

[4] H. Hendrick, R. Hidayat, G.-J. Horng, and Z.-H. Wang, "Non-Invasive Method for Tuberculosis Exhaled Breath Classification Using Electronic Nose," *IEEE Sensors Journal.*, vol. 21, no. 9, pp. 11184–11191, 2021.

[5] M. Badola, A. Agrawal, D. Roy, R. Sinha, A. Goyal, and N. Jeet, "Volatile Organic Compound Identification-Based Tuberculosis Screening among TB Suspects: A Diagnostic Accuracy Study," *Advances in Respiratory Medicine.*, vol. 91, no. 4, pp. 301–309, 2023.

[6] D. Aulia, R. Sarno, S. C. Hidayati, A. N. Rosyid, and M. Rivai, "Identification of chronic obstructive pulmonary disease using graph convolutional network in electronic nose," *Indonesian Journal of Electrical Engineering and Computer Science.*, vol. 34, no. 1, pp. 264–275, 2024.

[7] D. Aulia, R. Sarno, S. C. Hidayati, and M. Rivai, "Optimization of the Electronic Nose Sensor Array for Asthma Detection Based on Genetic Algorithm," *IEEE Access*, vol. 11, pp. 74924–74935, 2023.

[8] M. Rivai, D. Aulia, and S. Aulia, "Reducing the Electronic Nose Sensor Array for Asthma Detection Using Firefly Algorithm," *International Journal of Intelligent Engineering and Systems.*, vol. 17, no. 2, pp. 700–714, 2024.

[9] C. Wang and K. Liu, "A Randomly Guided Firefly Algorithm Based on Elitist Strategy and its Applications," *IEEE Access*, vol. 7, pp. 130373–130387, 2019.

[10] M. Seyedali and L. Andrew, "The Whale Optimization Algorithm," *Advances in Engineering Software.*, vol. 95, pp. 51–67, 2016.

[11] S. Mahmood, N. Z. Bawany, and M. R. Tanweer, "A comprehensive survey of whale optimization algorithm: modifications and classification," *Indonesian Journal of Electrical Engineering and Computer Science.*, vol. 29, no. 2, pp. 899–910, 2023.

[12] M. Sharawi, H. M. Zawbaa, and E. Emary, "Feature selection approach based on whale optimization algorithm," in *2017 Ninth International Conference on Advanced Computational Intelligence (ICACI)*, Doha, Qatar, 2017, pp. 163–168.

[13] S. Wakhid, R. Sarno, and S. I. Sabilla, "The effect of gas concentration on detection and classification of beef and pork mixtures using E-nose," *Computers and Electronics in Agriculture.*, vol. 195, pp. 1–9, 2022.

[14] Misbah, M. Rivai, F. Kurniawan, Z. Muchidin, and D. Aulia, "Identification of diabetes through urine using gas sensor and convolutional neural network," *International Journal of Intelligent Engineering and Systems.*, vol. 15, no. 1, pp. 520–529, 2022.

[15] R. A. Putri, S. I. Sabilla, and R. Sarno, "Optimal Feature Selection Algorithm (FSA) for Electronic Nose Signal," in *2023 1st International Conference on Advanced Engineering and Technologies (ICONNIC)*, Kediri, Indonesia, 2023, pp. 310–315.

[16] M. Rivai and D. Aulia, "Use of Electronic Nose to Identify Levels of Cooking Cookies," *IEEE Access*, vol. 12, pp. 97235–97247, 2024.

[17] R. Sarno, S. I. Sabilla, D. R. Wijaya, and Hariyanto, "Electronic Nose for Detecting Multilevel Diabetes using Optimized Deep Neural Network," *Engineering Letters.*, vol. 28, no. 1, pp. 31–42, 2020.