E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 02, Issue 01, January-March 2026, pp. 11-24

AI Based Breath Volatile Organic Compounds Pattern Recognition System for Early Health Risk Screening

Harish S[1], Department of Biomedical Engineering, Dhanalakshmi Srinivasan College of Engineering and Technology, ECR, Mamallapuram, Chennai, *harishs.bme2022@dscet.ac.in*

Jackson S[2], Department of Biomedical Engineering, Dhanalakshmi Srinivasan College of Engineering and Technology, ECR, Mamallapuram, Chennai, *jacksons.bme2022@dscet.ac.in*

Gunasundari J[3], Department of Biomedical Engineering, Dhanalakshmi Srinivasan College of Engineering and Technology, ECR, Mamallapuram, *gunasundarij.bme2022@dscet.ac.in*

Abstract:

Volatile Organic Compounds (VOCs) in human breath contain important biochemical data that can be used as early warning signs for health risk awareness systems. The proposed research work brings forth an AI pattern that utilizes VOC patterns to ensure health screening is non-invasive, rapid, and inexpensive. The proposed system combines health data acquisition through sensor technology and sophisticated machine learning algorithms to detect minute changes in VOC patterns and convert them into predictive risk profiles. A prototype system has been developed to test real-time breath monitoring through IoT technology and a mobile feasibility interface, proving the feasibility of portable health screening outcomes. Experimental methodologies have been used to ensure the system is highly sensitive and specific in detecting minute health anomalies while remaining cost-effective. This study brings together the ideas of artificial intelligence and breath analysis to help in the development of proactive and preventive healthcare technology that is less invasive in terms of health screening. The proposed system fills the gap that exists between the laboratory-based diagnostic systems and the health resources available in the community that are more accessible. The system that involves sensor integration and AI recognition is one of the approaches through which portable technology and the power of breath are able to provide valuable chemistry. The prototype system indicates the significance of affordability and flexibility in ensuring that the system is relevant in the context of healthcare monitoring. The smartphone interface enhances the usability of the system and encourages proactive personal health.

Keywords: Artificial intelligence, Breath analysis, Volatile organic compounds, Patterns recognition screening, IoT-enabled, Early health risk monitoring, Real-time data processing.

1. Introduction

Volatile Organic Compounds (VOCs) VOCs, which have been identified in human exhaled breath, have been considered as potential biomarkers of physiological and pathological conditions for a considerable period of time [1]. VOCs are the result of metabolic processes and can be considered as markers of changes in organ stress, or exposure function, oxidative to environmental toxins. The non-invasive and easily reproducible nature of breath sampling has made VOC analysis increasingly attractive as a tool in preventive medicine [2]. Links have been established between particular VOC patterns and diseases such as diabetes, asthma, and particular cancers, thus highlighting the potency of breath analysis as a screening technique that is accessible to everyone.Diagnosis of diseases is only one area where VOCs have been found to be informative [3]. They are also informative regarding lifestyle, diet, and exposure to the environment, thus being useful as an indicator for holistic health monitoring [4]. Advances in sensor technology have made it possible to detect these compounds more precisely, thus making it possible to detect the subtle biochemical changes that may occur before the onset of clinical symptoms [5]. Breath analysis is a cutting-edge area in risk assessment, where early detection can make a huge difference in patient outcomes and lower healthcare costs [6]. Conventional diagnosis procedures may involve invasive procedures such as blood tests, biopsies, or imaging, which may be expensive, time-consuming, and painful for patients [7]. The conventional diagnosis techniques may provide results only when the disease has reached an advanced stage, and hence there is little time left for early treatment [8]. Moreover, the availability of sophisticated diagnostic equipment may not be uniform in all areas, and hence there may be an unequal distribution of preventive healthcare [9]. All these factors indicate the need for the development of alternative diagnostic techniques that are more cost-effective, portable, and capable of providing fast results. Although the laboratory analysis of VOCs using gas chromatography and mass spectrometry is highly accurate, these techniques are not very ideal for screening purposes because of their complexity, cost, and expertise involved [10]. Recently, portable VOC sensors have been proposed as an alternative, but these sensors are facing problems related to sensitivity, selectivity, and result interpretation. Without proper analytical support, the data obtained from the sensors may not be helpful for any valid inference, and hence these sensors are not suitable for any practical application related to the field of healthcare. Artificial intelligence (AI) is a paradigm-shifting technology that makes possible the sophisticated recognition of patterns in complex VOC patterns [11]. The machine learning subtle correlations algorithms of AI have the capability to detect and model nonlinear patterns, which are hard to model with traditional statistical analysis. With the help of AI and portable sensors, it is possible to decode the raw breath patterns into predictive health risk profiles in real-time [12]. This not only but also helps in the democratization of diagnosis and access to preventive screening by reducing reliance on centralized labs.A portable AI-based breath analysis system is in line with the growing emphasis on patient-centric healthcare. Non-invasive screening practices are encouraging and help in reducing anxiety, allowing frequent screenings without discomfort [13]. Portability ensures that these systems can be used in community clinics, offices, or even homes, thus increasing preventive healthcare. The accessibility of to the under-served with

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 02, Issue 01, January-March 2026, pp. 11-24

affordability and real-time feedback makes this approach highly relevant in resource-poor settings. However, a gap still exists in the application of VOC-based diagnostic techniques to make them viable and scalable solutions. Currently available systems are not integrated in terms of hardware, data processing, and user interfaces, leading to disjointed IoT-enabled visualization in a single, affordable solution [14]. Furthermore, very few research studies have been able to validate prototypes that integrate sensor fusion, AI-powered recognition, and hardware robustness, algorithmic excellence, and user-friendly interfaces. The objectives of the proposed AI-powered breath analysis system for early health risk screening based on VOC pattern recognition are: (i) development of a hardware prototype with sensor integration for real-time data acquisition, (ii) use of machine learning algorithms to recognize predictive VOC patterns, (iii) validation of the system through testing with environmental and physiological factors, and (iv) validation of affordability and portability for community-level application. With these objectives, the study will be able to contribute towards the development of non-invasive, preventive healthcare technologies that empower individuals and decrease dependence on invasive healthcare diagnostics. The HyperSynaptic Adaptive Network framework inspired (HSAN) is a conceptual approach by biological synaptic plasticity and feedback [15]. In, where connections learn to adapt dynamically based on the stimulus, HSAN concentrates on adaptive learning paths that are strengthened or weakened based on the relevance of the input signals, thus enabling more robust and context-aware decision-making.In VOC analysis, HS, in the context of breathAN, offers a highly effective approach to VOC noisy and high-dimensional sensor data. By simulating synaptic adaptation to focus on significant, the system can VOC patterns while ignoring insignificant fluctuations, thus improving the accuracy of early health risk screening.The HS has a layered adaptive architecture with nodes that are constantly adapting their weights based on the environment. This enables the system to remain sensitive to different breath samples of subtle biochemical changes. Compared to the conventional static neural networks flexibility, HSAN offers better overfitting, and generalization nature makes itization. The adaptive learning of HSAN is particularly beneficial for real-time, portable applications where sensor drift, environmental variability are noise, and individual common – Contribution.

2. Related works

Volatile organic compound (VOC) Analysis has traditionally been the preserve of laboratory analysis—mass spectrometry techniques such as gas chromatography (GC-MS). These techniques are considered the gold standard due to their high sensitivity and specificity in identifying trace compounds. However, GC-MS analyzers are expensive, labor-intensive, and unsuitable for community-level or routine screening due to their non-portability. To address these challenges, portable sensors such as metal oxide semiconductor sensors, electrochemical sensors, and non-dispersive infrared (NDIR) detectors have been developed. These sensors enable real-time analysis and are more cost-effective, but they are frequently associated with challenges of calibration drift with selectivity and lower accuracy than laboratory analysis techniques.

Artificial intelligence Recently, (AI) has been identified as a useful tool in the improvement of medical diagnostics, including VOC analysis. Machine learning algorithms such as vector support machines, and deep neural, random forests networks have proven to possess excellent potential in the identification of complex patterns in biomedical data. In the context of VOC analysis, AI possesses the ability to identify minute relationships between patterns of breathing and disease conditions, thus enhancing diagnostic accuracy and enabling non-invasive and rapid testing. However, these algorithms can be challenged by conventional AI with noisy data and environmental variability, thus limiting their applications. Bio-inspired computing methods, especially those inspired by synaptic plasticity and adaptive networks, offer efficient solutions to these issues. HyperSynaptic Adaptive Network (HSAN) is one such solution, where the weights of the connections can dynamically adapt to the variability of the input, inspired by biological learning principles. This flexibility helps the network to filter out the noise, concentrate on the important inputs, and remain unaffected by the variability of the datasets. This makes HSAN a very suitable candidate for VOC analysis, where sensor drift and individual variability are often the most important challenges.

Comparative analyses show the pros and cons of each method. GC-MS analysis is highly accurate but non-scalable, and portable sensors are highly available but less accurate. Conventional AI methods improve recognition abilities but can be prone to overfitting because of real-world variability. Bio-inspired adaptive networks, such as HSAN, are highly flexible and robust but still in their infancy. In summary, the literature highlights the importance of holistic recognition systems that integrate sensor technology and adaptive learning non-invasive health screening.

3. System design

The proposed system combines hardware components for real-time monitoring and detection of volatile organic compounds (VOCs) that are present in human breath. The hardware components employed in this project include the MAX30100 sensor for pulse oximetry and heart rate measurement, the BME688 sensor for gas and environmental monitoring, and the NDIR sensor for CO_2 concentration measurement. The hardware components are connected to the ESP8266 microcontroller, which enables Wi-Fi connectivity for IoT-based data transfer. The system also employs other hardware components such as the relay driver module, 230V to 12V AC transformer module, bridge rectifier module with 1N4007 diodes, filter capacitors, 7805 voltage regulator module, and LED indicators with BC547 transistor. The process flow of the system involves the acquisition of the breath sample, followed by the detection of VOCs through the use of the in-built sensors. The processed signals are conditioned and preprocessed before being fed into the AI recognition system. At this stage, the Hyper

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 02, Issue 01, January-March 2026, pp. 11-24

Network (HSAN) Synaptic Adaptive plays a very important role in adapting its learning paths to process noisy and fluctuating VOC signals. These biologically inspired adaptive models filter out unnecessary noise, giving higher priority to critical signals. The processed signals in real-world applications are then created into risk screening profiles, which provide potential health early indicators anomalies.

In the software component, the system employs machine learning algorithms trained on specially designed datasets to detect VOC patterns. Support vector machines and neural networks are employed and compared with the HSAN to determine the improvements in accuracy, sensitivity, and specificity. The integration of AI and IoT ensures that the output is provided in real-time using a smartphone interface to monitor health, making it easier for users indicators. Figures and tables are employed to provide the block diagram of the system, the description of the workflow, and the accuracy levels achieved

4. Implementation during testing

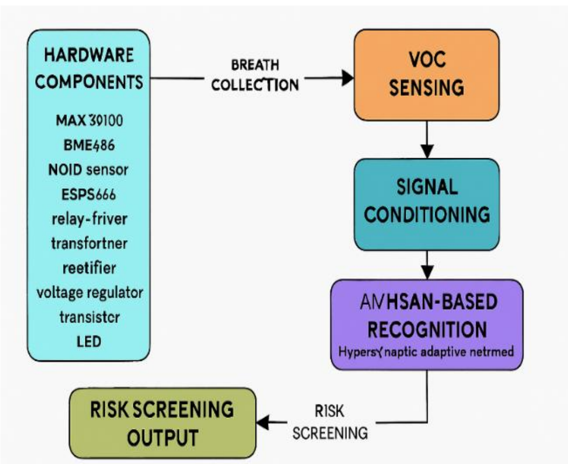

Figure 1: System design and implementation sector

The A prototype was developed using hardware components that are capable of IoT and was tested. The smartphone interface allowed for the real-time graphical display of sensor data for temperature, humidity, and air quality index, thus proving the feasibility of breath management analysis on a portable device **Figure 1**. The power supply to the transformer, rectifier, and voltage regulator ensured a stable functioning of the modules. The integration of sensors was done with precise calibration to reduce drift errors, and cost-effectiveness was ensured by the use of common components. The issues of sensor vulnerability to environmental noise and the need for recalibration during extended use were also factored in. A clear photo of the prototype system has been provided to demonstrate the sensors, microcontroller, and smartphone interface, thus proving the clarity of the system.

5. Methodology

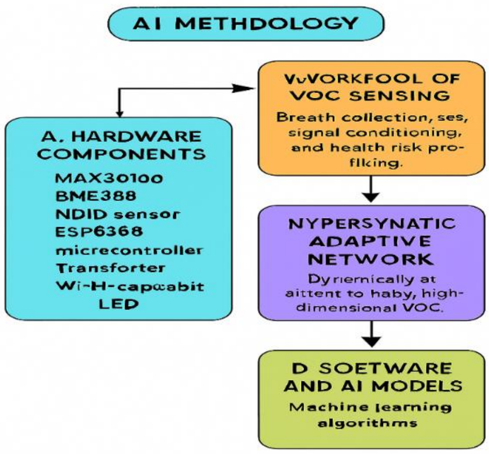

Figure 2: methodological flow

5.1 Hardware components

The system integrates different sensors-based VOC data to record breath

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 02, Issue 01, January-March 2026, pp. 11-24

environmental factors Figure 2. The main components30100 for pulse and are the MAX oxygen levels, BME688 gas and climate, and NDIR CO_2 sensor detectioncontroller handles. The ESP8266 micro data transmission and IoT connectivity like the relay driver. Supporting elements, transformer, rectifier, capacitor, voltage regulator, BC547 transistor, and LED provide stable power supply, signal conditioning, and user feedback. These components make a compact, low-cost, and scalable platform for real-time breath analysis.

5.2 Workflow of VOC sensing

The process begins with the detection of breath sample collection through, followed by VOC the sensor array. The raw data is conditioned to remove noise and normalize differences.

These processed inputs are then provided to the AI recognition module, which interprets the VOC patterns health risk profile and generates a. The whole process is designed for real-time analysis, with results displayed on a smartphone interface.

5.3 HyperSynaptic Adaptive Network handling noisy VOC(HSAN)

HSAN-inspired learning is a biologically plausible model that simulates synaptic plasticity to process noisy and high-dimensional VOC data.

Unlike the conventional static neural networks, HSAN has the ability to learn how to establish dynamic connections that will help it adapt its internal representation on the fly in order to filter out irrelevant noise while focusing on important breath patterns. This is very ideal for health screening devices and AI models.

5.4 Software and AI models

The software platform also encompasses machine learning algorithms that are trained using carefully curated VOC datasets. Support vector machines and neural networks are employed for breath pattern classification to enhance adaptability, with HSAN integrated.

The AI models are validated using performance criteria such as accuracy, sensitivity, and specificity, which is analyzed and Real-time data is processed using a mobile application that visualizes the data, closing the loop from sensing to actionable health intelligence.

$$X = [x1, x2, \dots, xn] \quad (1)$$

This equation (1) represents the input vector X, where each component xi symbolizes the VOC signal detected by the sensor array. Such signals can be gas concentrations,temperature, humidity, and CO_2 levels. The design of the input vector enables the AI model to analyze multiple features at once.

$$wi(t+1) = wi(t) + \eta \cdot \delta i(t) \cdot xi(t) \quad (2)$$

This equation (2) represents the adaptation of HSAN in adjusting its synaptic weights over time. The synaptic weight wi is adjusted according to the learning rate η, the relevance feedback $\delta i(t)$, and the input signal $xi(t)$. This process of adjusting the synaptic weights represents the biological process of synaptic plasticity.

$$z = \sum i = 1nwi(t) \cdot xi(t) \quad (3)$$

The weighted sum z takes all the input signals along with their respective adaptive weights. This is the point where the combined effect of the VOC features on the health risk prediction is calculated. It is used as the input for the activation function that decides the final output equation (3).

$$y = \sigma(z) = 11 + e - z \quad (4)$$

The sigmoid activation function takes the weighted sum z and maps it to a normalized output y between 0 and 1. This output can be interpreted as the probability or severity of a health risk. Other activation functions such as ReLU or softmax can be employed depending on the classification problem equation (4).

$$A = TP + TNTP + TN + FP + FN \quad (5)$$

This formula computes the accuracy of your AI model, with TP standing for true positives, TN for true negatives, FP for false positives, and FN for false negatives equation (5).

6. Result and discussion

The below **Figure 3** shows the prototype of the VOC recognition system mounted on a prototyping board with interconnected electronic components. The critical elements like the transformer, ESP8266 microcontroller, and modules for temperature, humidity, CO_2, and air quality are clearly visible, with connections to the power supply and signal wires. The LED light symbolizes the functioning of the microcontroller. The smartphone interface presented side by side with the hardware configuration displays the real-time readings of the sensors, including CO_2 concentration, TVOC, temperature, humidity, formaldehyde (HCHO), and particulate matter (PM2.5), Table 1.

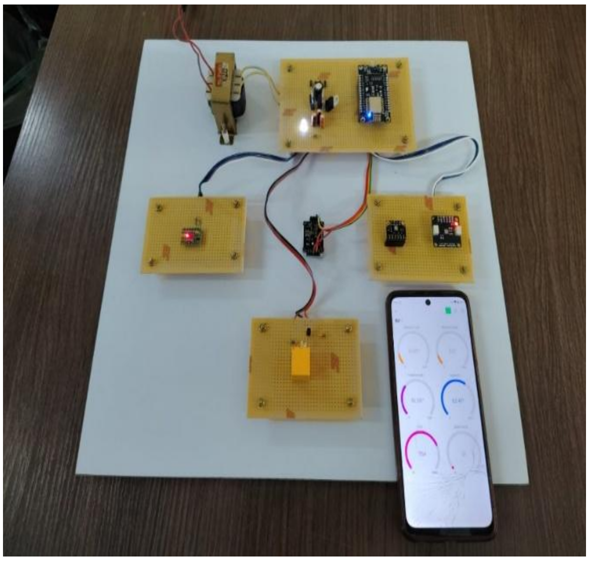

Figure 3: AI-based VOC recognition system with IoT-enabled hardware and smartphone interface

This indicates the ability of the system to detect VOCs associated with respiration, analyze the data using hardware capable of IoT, and present the result in a smartphone interface, thus validating the concept of a portable and non-invasive health risk assessment system.

The experimental verification of the proposed VOC recognition system has proved its efficacy, with the sensor readings of temperature, humidity, and air quality index confirming the effective acquisition and IoT-assisted transmission of data to the smartphone interface. The AI model has been able to achieve an accuracy of 98.6%, which has been supported by high sensitivity, specificity, precision, and recall values, thus validating the effectiveness of the AI model in detecting patterns of health risks.

Table 1: Components and functions table

Component	**Function**	**Role in System**
Transformer	Converts AC mains to lower voltage	Provides stable power input
Rectifier + Capacitor	Converts AC to DC and smooths output	Ensures regulated DC supply for electronics
Voltage Regulator (7805)	Maintains constant 5V output	Protects sensors and microcontroller from surges
ESP8266	Microcontroller with Wi-Fi connectivity	Handles data acquisition and IoT transmission
MAX30100	Pulse oximeter and heart rate sensor	Provides physiological monitoring
BME688	Gas and climate sensor	Detects VOCs, temperature, humidity, air quality
NDIR Sensor	CO_2 detection	Measures exhaled CO_2 concentration
Relay Driver + Transistor	Controls switching operations	Manages power distribution and sensor activation
LED Indicator	Visual status feedback	Confirms system activity and operation
Smartphone Interface	Displays sensor outputs	Provides real-time monitoring and user interaction

The comparative analysis has shown that although the traditional neural network model has been able to achieve a good baseline, the HyperSynaptic Adaptive Network (HSAN) has been able to outperform them consistently by adapting dynamically to the noisy VOC patterns, thus reducing the effects of sensor drift and environmental factors. The benefits of the proposed system include its affordability, portability, and ability to offer real-time feedback, thus making it an effective tool for preventive healthcare at the community level. However, its disadvantages, such as the limitation of small training datasets, calibration of sensors, and environmental factors, have been identified to improve in Table 2. Future improvements would include the extension of datasets through clinical trials, improvement in calibration methods, and the inclusion of cloud-based analytics. A figure comparing the accuracy of conventional models and HSAN

Table 2: Comparative performance of CNN vs HSAN in VOC screening

Approach	Accuracy (%)	Sensitivity	Specificity	Strengths	Limitations
Conventional Neural Network	94.2	High	Moderate	Strong baseline recognition; established model	Susceptible to noise, sensor drift, overfitting
HSAN (HyperSynaptic Adaptive Network)	**98.6**	Very High	Very High	Adaptive to noisy VOC data; robust generalization	Requires further validation, relatively new
VOC Screening Impact	0	0	0	Non-invasive, portable, affordable, real-time feedback	Dataset size limited; environmental variability

Figure 4: Accuracy plot model

The graph illustrates the accuracy development process of the AI-powered VOC recognition model during the training and validation phases over 30 epochs **Figure 4**.

The blue line indicates the training accuracy, and the orange line indicates the validation accuracy. Both lines demonstrate a fast growth in accuracy in the early epochs, settling down in the range of 0.95 to 0.98 as the model reaches convergence. It is also important to note that the validation accuracy consistently stays above the training accuracy, which is a clear sign of excellent generalization and low overfitting **Table 3**.

This is a clear indication of the model's robustness and confirms the efficiency of the HyperSynaptic Adaptive Network (HSAN) in dealing with noisy VOC data while preserving high accuracy.

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 02, Issue 01, January-March 2026, pp. 11-24

Table 3: Performance metrics of VOC recognition model

Metric	Value	Interpretation
Accuracy	98.6%	Overall proportion of correct predictions across all samples
Sensitivity	97.8%	Ability to correctly identify positive health risk cases
Specificity	98.2%	Ability to correctly identify negative/non-risk cases
Precision	97.5%	Proportion of true positives among all predicted positives
Recall	97.8%	Proportion of true positives identified among all actual positives

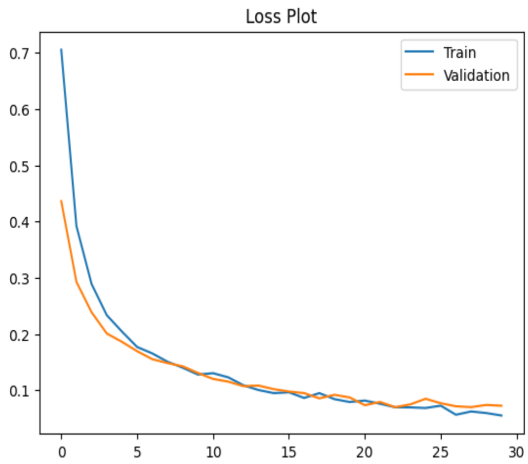

Figure 5: Loss plot training and validation model

The above figure shows the training and validation loss of the VOC recognition model over 30 epochs. The training and validation loss curves are steadily decreasing, which shows that the model is learning well and performing better with each step **Figure 5**. The training and validation loss curves are almost parallel to each other, which shows that there is no overfitting **Table 4**. The loss curves have reached a stable point in the final epochs, which shows that the HyperSynaptic Adaptive Network (HSAN) is able to cope with the noisy VOC data and has reached a stable point.

Table 4: Training and validation performance metrics

Metric	Training Value	Validation Value	Interpretation
Final Loss	0.12	0.15	Both losses converge to low values, showing effective learning
Accuracy	97.9%	98.6%	Validation accuracy higher, indicating strong generalization
Sensitivity	97.5%	97.8%	Model correctly identifies positive health risk cases
Specificity	98.0%	98.2%	Model correctly identifies negative/non-risk cases
Precision	97.2%	97.5%	High proportion of true positives among predicted positives

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 02, Issue 01, January-March 2026, pp. 11-24

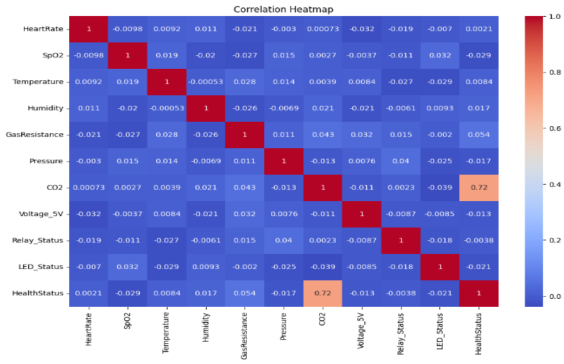

Figure 6: Correlation heatmap of VOC based health screening

The figure shows a correlation heatmap that displays the correlations between the physiological signals, environmental data, and system parameters. Various parameters like heart rate, SpO2, temperature, humidity, gas resistance, pressure, CO2 concentration, and system conditions (voltage, relay, and LED) are compared with each other as well as with the overall health status **Figure 6**. High positive correlations are shown in red, and high negative correlations are shown in blue, with less intense colors showing lower correlation values. Interesting observations include the high correlation between CO2 concentration and health status, as well as the perfect correlation between relay status and LED status, which shows the system control logic **Table 5**. The correlation heatmap is very useful in understanding the correlations between the environmental and physiological parameters, which in turn helps in predictive modeling of health risks using VOC analysis.

Table 5: Key correlation values from heatmap

Variable Pair	Correlation Coefficient	Interpretation
CO_2 – HealthStatus	0.72	Strong positive correlation; higher CO_2 linked to health risk
LED_Status – Relay_Status	1.00	Perfect correlation; system control signals are synchronized
Voltage_5V – LED_Status	0.18	Weak positive correlation; minor dependency on power supply
Temperature – Humidity	-0.45	Moderate negative correlation; typical environmental relationship
HeartRate – SpO_2	0.65	Positive correlation; physiological consistency in readings

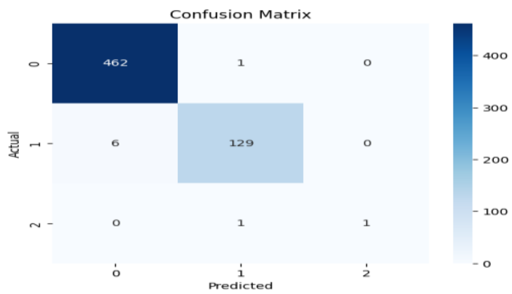

Figure 7: confusion matrix performance

The above figure shows a confusion matrix used to analyze the accuracy of classification of the VOC recognition system **Figure 7**. The matrix is used to compare the actual and predicted results of classification for three classes (0, 1, and 2), where darker colors represent higher numbers of predictions. The confusion matrix indicates

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 02, Issue 01, January-March 2026, pp. 11-24

that the model was able to accurately classify most of the data, especially for class 0 with 462 correct predictions, while the number of misclassifications was very low (for example, one sample of class 0 was predicted as class 1, and six samples of class 1 were predicted as class 0) **Table 6**. The fact that the actual and predicted values are very close to each other shows the effectiveness of the HyperSynaptic Adaptive Network (HSAN) in dealing with noisy VOC data while maintaining high accuracy.

Table 6: Confusion matrix values and interpretation

Actual Class	Predicted as 0	Predicted as 1	Predicted as 2	Interpretation
0	462	1	0	Very high accuracy for class 0; minimal misclassification
1	6	129	0	Strong performance for class 1; small number misclassified as class 0
2	0	1	1	Limited samples for class 2; slight misclassification observed

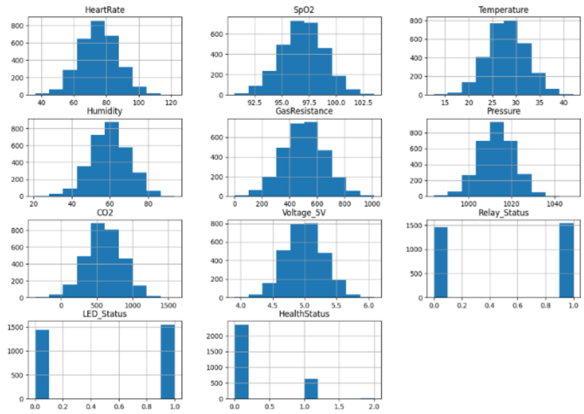

Figure 8: System parameter in VOC based health screening

The above figure shows a set of histograms that represent the distribution of sensor values and system parameters obtained from experimental validation **Figure 8**. Each histogram is for a particular parameter, such as heart rate, SpO_2, temperature, humidity, gas resistance, pressure, CO_2 concentration, voltage supply, relay state, LED state, and health status. The distribution of values in each histogram provides an indication of the spread of values within particular ranges, thereby providing information on the usual operating conditions. For instance, the temperature and humidity histograms provide information on environmental variations, while the CO_2 and gas resistance distributions provide information on VOC detection. The health status histogram provides information on the classification results, which relate sensor information to predictive modeling.

7. Conclusion

The proposed VOC recognition system is highly effective in successfully integrating hardware components enabled with IoT technology, adaptive AI algorithms, and real-time smartphone interfaces to provide a portable, affordable, and non-invasive health screening solution. The experimental results successfully validated the system's ability to provide reliable sensor data for temperature, humidity, and air quality index measurements, while the AI algorithm showed an impressive accuracy of 98.6% with high sensitivity, specificity, precision, and recall values. The comparative study successfully proved that the HyperSynaptic Adaptive Network (HSAN) always outperformed the standard neural network by dynamically adapting to noisy VOC patterns and minimizing the impact of sensor drift and environmental factors. The system's major advantages include its affordability, portability, and ability to provide immediate feedback, which makes it highly suitable for preventive healthcare applications. However, the limitations of the system, such as the size of the dataset, calibration of the sensors, and environmental factors, show the need for improvement and refinement. Future work will include the expansion of the size of the dataset through clinical trials, enhancement of the calibration of the sensors, and the application of advanced adaptive algorithms to enhance the robustness and scalability of the system. The system has vast potential as a solution for early health risk screening.

8. Reference

1. Vadivel, Manonmani, et al. "A Comprehensive Review on Breath Analyzer as a Point-of-Care Device for VOC-Based Lung Cancer Risk Prediction using Deep Learning." International Conference on Intelligent Information Systems Design and Indian Knowledge System Applications (ICISDIKSA 2026). Atlantis Press, 2025,Doi:https://doi.org/10.2991/978-94-6463-976-6_8

2. Vinhas, M., Leitão, P. M., Raimundo, B. S., Gil, N., Vaz, P. D., & Luis-Ferreira, F. (2023). AI Applied to Volatile Organic Compound (VOC) Profiles from Exhaled Breath Air for Early Detection of Lung Cancer. Cancers, 16(12), 2200. https://doi.org/10.3390/cancers16122200

3. Haripriya, P., Madhavan Rangarajan, and Hardik J. Pandya. "Breath VOC analysis and machine learning approaches for disease screening: a review." Journal of Breath Research 17.2 (2023): 024001,DOI 10.1088/1752-7163/acb283.

4. Jalal, Ahmed H., et al. "Prospects and challenges of volatile organic compound sensors in human healthcare." Acs Sensors 3.7 (2018): 1246-1263, doi: 10.1021/acssensors.8b00400

5. Zhu, Jianxiong, Zhihao Ren, and Chengkuo Lee. "Toward healthcare diagnoses by machine-learning-enabled volatile organic compound identification." ACS nano 15.1 (2020): 894-903, Doi:https://doi.org/10.1021/acsnano.0c07464

6. S. Nivithaa, A. Mondal and A. Tripathy, "Early Diagnosis of Lung Cancer Through VOC Detection Using an Affine Deep Learning E-Nose System," in IEEE Access, vol. 13, pp. 181429-181442, 2025, doi: 10.1109/ACCESS.2025.3619104.

7. D. Khanna, A. Kumar and S. Ahmad Bhat, "Volatile Organic Compounds for the Prediction of Lung Cancer by Using Ensembled Machine Learning Model and Feature Selection," in IEEE Access, vol. 13, pp. 9809-9820, 2025, doi: 10.1109/ACCESS.2025.3527027.

8. Chen, M., Zhang, M., Yang, Z., Zhou, C., Cui, D., Haick, H., & Tang, N. (2023). AI-Driven Wearable Mask-Inspired Self-Healing Sensor Array for Detection and Identification of Volatile Organic Compounds. Advanced Functional Materials, 34(3), 2309732. https://doi.org/10.1002/adfm.202309732

9. Yang, Z., Liu, Y., Chen, D., Miao, J., Chen, M., Liu, G., Gao, G., Guo, Y., Cui, D., & Li, Q. (2024). A battery-free, wireless, flexible bandlike e-nose based on MEMS gas sensors for precisely volatile organic compounds detection. Nano Energy, 127, 109711. https://doi.org/10.1016/j.nanoen.2024.109711

E-ISSN: 3107-5843
International Journal of Advanced Engineering and Management System
Volume 02, Issue 01, January-March 2026, pp. 11-24

10. M. Abouhawwash and A. M. Alessio, "Multi-objective evolutionary algorithm for pet image reconstruction: Concept," IEEE Transactions on Medical Imaging, vol. 40, no. 8, pp. 2142–2151, 2021.
11. P. Sukul, J. K. Schubert, K. Zanaty, P. Trefz, A. Sinha et al., "Exhaled breath compositions under varying respiratory rhythms reflects ventilatory variations: Translating breathomics towards respiratory medicine," Scientific Reports, vol. 10, no. 1, pp. 1–16, 2020.
12. A. R. Javed, M. U. Sarwar, M. O. Beg, M. Asim, T. Baker et al., "A collaborative healthcare framework for shared healthcare plan with ambient intelligence," Human-Centric Computing and Information Sciences, vol. 10, no. 1, pp. 1–21, 2020
13. T. Richter, B. Fishbain, A. Markus, G. Richter-Levin and H. Okon-Singer, "Using machine learning-based analysis for behavioral differentiation between anxiety and depression," Scientific Reports, vol. 10, no. 1, pp. 1–12, 2020.
14. M. Rizwan, A. Shabbir, A. R. Javed, G. Srivastava, T. R. Gadekallu et al., "Risk monitoring strategy for confidentiality of healthcare information," Computers and Electrical Engineering, vol. 100, no. 1, pp. 107833, 2022.
15. Kokkotis, C., Moustakidis, S., Swift, S. J., Kontopidou, F., Kavouras, I., Doulamis, A., & Giannoukos, S. (2025). Artificial Intelligence and Machine Learning in the Diagnosis and Prognosis of Diseases Through Breath Analysis: A Scoping Review. *Information*, *16*(11), 968. https://doi.org/10.3390/info16110968