WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Early Detection of Blood Cancer Using Volatile Organic Compounds (VOC)

[1]Julie Joan N S, [2]Dr.D.Rajinigirinath, [3]Sunil Earnest. V, [4]R.Gayathri ,[5]G.K Monica

[1]Research Scholar Anna University chennai,
[2]Professor, Sri Muthukumaran Institute of Technology, Chennai
[3]Asst Professor. Jaya Engineering College, Chennai
[4]Asst.Professor, Aalim Muhammed Salegh college of Engineering
[5]Asst.Professor, Aalim Muhammed Salegh college of Engineering

Abstract

Blood cancer continues to pose a significant health problem, mainly due to late diagnosis and dependence on invasive testing techniques such as biopsies and imaging, which can be expensive and inefficient for early detection. Volatile Organic Compounds (VOCs) have arisen as encouraging non-invasive indicators for early cancer identification. This research investigates the categorization of VOC patterns from blood cancer patients utilizing c (GC-MS) and Electronic Nose (E-nose) technology. Machine learning algorithms, including Support Vector Machines (SVM), Random Forest, and Artificial Neural Networks (ANN), were utilized to classify VOC signatures. The ANN model reached the highest accuracy of 93. 5%, exhibiting exceptional predictive ability. The results suggest that VOC-based AI models provide a swift, non-invasive, and economical option for early blood cancer detection, potentially transforming diagnostic methodologies by allowing earlier intervention and enhancing patient outcomes.

Keywords-Blood Cancer, Volatile Organic Compounds, Gas Chromatography-Mass Spectrometry, Electronic Nose, early detection.

INTRODUCTION

Hematologic cancers, which encompass leukemia, lymphoma, and multiple myeloma, rank among the top causes of cancer-related fatalities globally. These diseases impact the bone marrow, lymphatic system, and blood cells, resulting in immune system impairment and serious health complications. Early detection is vital for enhancing survival rates among patients; however, current diagnostic techniques—such as bone marrow biopsies, blood tests, and imaging—tend to be invasive, expensive, or insufficiently sensitive during the early stages. The identification of non-invasive biomarkers, especially volatile organic compounds (VOCs), has opened new avenues for diagnostic strategies. VOCs are metabolic byproducts that are emitted through breath, blood, and urine, and their distinct chemical signatures offer valuable information about the existence and progression of diseases.

Conventional methods for diagnosing blood cancer have significant drawbacks. Biopsies are very accurate but can be painful, expensive, and require skilled personnel. Blood tests often miss early-stage cancers due to low levels of biomarkers. Imaging methods like MRI and PET scans are beneficial for staging the disease but carry risks of radiation exposure and entail high costs. These limitations highlight the urgent need for a rapid, non-invasive, and affordable diagnostic solution. Recent research indicates that VOC-based diagnostics, when paired with machine learning techniques, can detect blood cancer in its early stages by analyzing distinctive chemical signatures. This method lessens the reliance on invasive tests, thereby making diagnostics more accessible to a larger population. This study intends to investigate volatile organic compounds (VOCs) as potential indicators for the early detection of blood cancers through the use of Gas Chromatography-Mass Spectrometry (GC-MS) and Electronic Nose (E-nose) technology. The research utilizes machine learning algorithms—Random Forest, Support Vector Machine (SVM), and Artificial Neural Networks (ANN)—to differentiate between cancerous and non-cancerous samples based on their VOC profiles. By assessing the accuracy of various models, this research aims to create a non-invasive, AI-powered approach for blood cancer screening. If proven successful, this technique could greatly improve early diagnosis, enhance treatment results, and transform cancer detection methods in clinical environments. The main goal of this paper is to investigate the function of VOCs in the early detection of blood cancer, evaluate the efficacy of diagnostic tools based on VOCs, and emphasize improvements in the integration of machine learning for pattern recognition. The research additionally assesses the practicality of non-invasive detection methods within clinical environments

Volatile Organic Compounds (VOCs) and Their Role in Blood Cancer

VOCs are products of cellular metabolism and may act as indicators for multiple diseases. In individuals with cancer, metabolic irregularities cause changes in VOC profiles. These changes can be identified in exhaled breath, blood, and urine samples, offering essential diagnostic insights.

A. VOC Profiles in Blood Cancer

Patients with blood cancer display distinctive VOC profiles marked by alterations in oxidative stress indicators, lipid peroxidation byproducts, and metabolites related to immune responses. Specific VOCs, including aldehydes, ketones, hydrocarbons, and alcohols, exist in notably different levels when compared to healthy individuals.

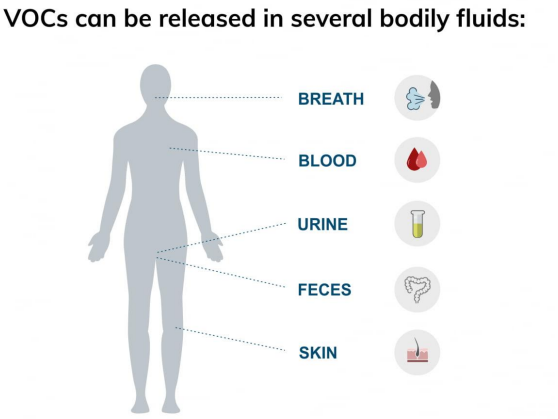


FIG 1. VOC Releases in Body Fluids

B. Detection Techniques for VOCs

Various methods have been created to identify and examine VOCs, including:
Breath Analysis: Employs e-noses and GC-MS to recognize cancer-specific VOCs.
Blood and Urine Testing: Identifies metabolic markers in bodily fluids.
Electronic Sensors: Mobile VOC detectors for immediate analysis.

C. VOC Biomarkers for Blood Cancer Detection

VOCs act as biomarkers, presenting a distinct metabolic signature for cancer identification. Certain compounds like benzene derivatives, alkanes, and aldehydes have been detected in greater amounts in blood cancer patients relative to healthy individuals. These VOCs primarily result from oxidative stress and lipid peroxidation, which are defining processes in cancer metabolism.

Research has indicated that:

- Aldehydes and Ketones are notably increased due to heightened lipid peroxidation.
- Alkanes and Aromatic Compounds reflect abnormal metabolic processes.
- Sulfur-Containing Compounds are linked to modified amino acid metabolism in blood cancer patients.

Early identification of blood cancer continues to pose a considerable obstacle in contemporary oncology. Conventional diagnostic techniques, including complete blood count (CBC) tests, bone marrow biopsies, and flow cytometry, necessitate invasive procedures that can be uncomfortable and lengthy. Furthermore, these methods frequently do not identify cancer in its early phases, resulting in postponed diagnoses and diminished survival rates. Blood cancer is especially challenging to detect early as its symptoms, such as exhaustion, recurrent infections, and unexplained bruising, often mimic those of prevalent diseases. The intricacy of hematological malignancies adds to the difficulty of pinpointing distinct biomarkers for early detection. In addition, the steep costs associated with traditional diagnostic methods render them less accessible in settings with limited resources. Thus, there is an urgent need to create innovative, non-invasive, and cost-efficient diagnostic strategies capable of identifying blood cancer at an early stage with high specificity and sensitivity.

LITERATURE SURVEY

In 2025, a study published in the IEEE Sensors Journal titled "CS-MA-CNN: A Fast Recognition Network of Electronic Nose for Ignitable Liquids Detection" introduced a novel method for detecting ignitable liquids (ILs) using an electronic nose (e-nose). The proposed method, called the Channel-Separated Multiscale Attentional Convolutional Neural Network (CS-MA-CNN), utilizes a portion of the response time data from the e-nose to achieve rapid identification of ILs. The study demonstrated that CS-MA-CNN effectively captures correlations between temporal and channel dimensions of e-nose data, achieving an impressive classification accuracy of 94.53% with just 5 seconds of data from a portable e-nose [1].

Additional progress in VOC detection is evidenced in the 2023 study shared at the IEEE 19th International Conference on Body Sensor Networks (BSN), which introduced an e-nose prototype integrated with cardiovascular assessments for disease biomarker detection. The system underwent testing with graphene-based sensors positioned in the neck area to identify 12 clinically significant VOCs. The research emphasized the promise of this integrated approach for diagnosing various diseases, including blood cancers, by monitoring VOC profiles and associating them with disease biomarkers. This investigation highlights the increasing importance of wearable sensor technologies in early disease detection [2].

In 2025, the launch of the SPYROX e-nose represented another notable progression in the domain. Reported in the IEEE Sensors Journal, Volume 25, Issue 5, the SPYROX device featured an array of eight metal-oxide sensors along with an optimized analog front-end (AFE) circuit. The device achieved 100% classification accuracy in human volatilome fingerprinting, accurately differentiating cancer-related VOC profiles from healthy ones. This study underscores the significance of optimized sensor systems for VOC analysis, which can enhance diagnostic precision and aid in non-invasive cancer detection [3].

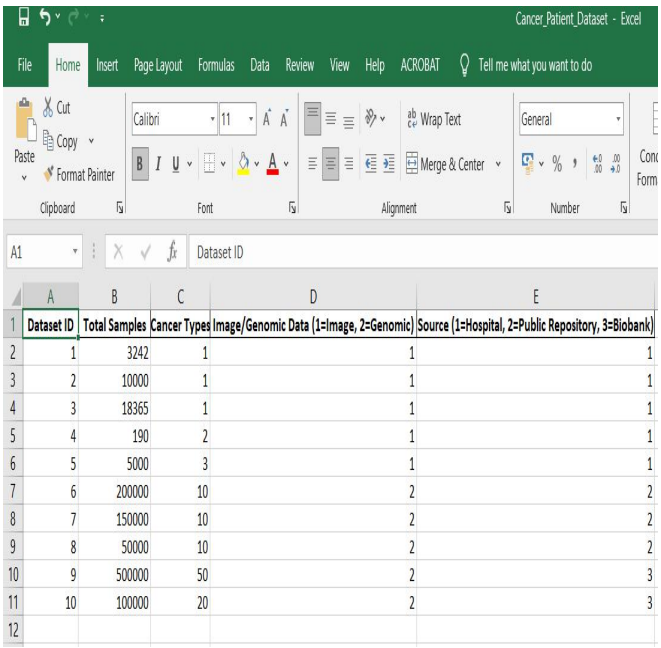
In the field of advanced biosensors, a 2025 study outlined the creation of a high-sensitivity triple-band nano-biosensor for the early detection of leukemia, published in the IEEE Sensors Journal. Functioning in the petahertz range, this biosensor employed a gold or silver dipole ring structure on a silicon dioxide substrate to achieve superior absorption characteristics for detecting leukemia biomarkers. The biosensor exhibited high spatial resolution, allowing it to distinguish between healthy and cancerous samples. This research highlights the potential of nano-biosensors for extremely sensitive, early-stage cancer detection [4].

The application of machine learning and deep learning models has further boosted cancer diagnosis, particularly in the context of automated blood cancer detection. In 2024, a study showcased at the International Conference on IoT, Communication, and Automation Technology (ICICAT) unveiled a hybrid ensemble Deep Convolutional Neural Network (DCNN) model for identifying blood cancer from blood cell images. This model achieved accuracy rates surpassing 99%, demonstrating its potential to improve early detection and diagnostic capabilities for blood cancer. The deep learning model incorporates sophisticated computational techniques, offering a promising resource for cancer diagnosis [5].

Another significant contribution to the detection of blood cancer was made by a 2020 research paper published in IEEE Access. This research suggested utilizing a Dense Convolutional Neural Network (DCNN) to differentiate Acute Lymphoblastic Leukemia (ALL) and Multiple Myeloma (MM) from microscopic images of bone marrow. The model reached an accuracy of 97. 2%, surpassing traditional machine learning techniques such as Support Vector Machines (SVM), Decision Trees, and Random Forests. This deep learning method illustrated the model's capacity to precisely identify and categorize blood cancers, confirming its utility as a resource for early-stage blood cancer diagnosis [6].

III.METHODOLOGY

This research examines the application of volatile organic compounds (VOCs) as indicators for the early identification of blood cancer, with an emphasis on leukemia and lymphoma. The study adopts a dual methodology, integrating a thorough literature review with a suggested experimental framework for additional validation. The initial segment of the methodology entails performing a comprehensive literature review to assess the existing landscape of VOC-based cancer detection. This review encompasses research that has concentrated on identifying VOCs in exhaled breath, urine, and other bodily fluids, contrasting the VOC profiles of cancer patients with those of healthy individuals. The primary objective is to detect VOC patterns that are distinctive to blood cancers and evaluate the efficacy of different detection techniques. Research utilizing technologies such as gas chromatography-mass spectrometry (GC-MS), mass spectrometry, and electronic noses (e-noses) is notably examined to analyze the sensitivity and specificity of VOCs as biomarkers.



	A	B	C	D	E
1	Dataset ID	Total Samples	Cancer Types	Image/Genomic Data (1=Image, 2=Genomic)	Source (1=Hospital, 2=Public Repository, 3=Biobank)
2	1	3242	1	1	1
3	2	10000	1	1	1
4	3	18365	1	1	1
5	4	190	2	1	1
6	5	5000	3	1	1
7	6	200000	10	2	2
8	7	150000	10	2	2
9	8	50000	10	2	2
10	9	500000	50	2	3
11	10	100000	20	2	3
12					

FIG2.Cancer patient datasets

For the experimental validation, this research suggests a methodology that includes gathering breath and urine samples from both blood cancer patients and healthy controls. The samples will be examined using GC-MS, a well-known method for identifying and measuring VOCs, to develop a comprehensive profile of the compounds found in the samples. Alongside GC-MS, an electronic nose (e-nose) will be utilized for quick, non-invasive detection. The e-nose, which has sensors that replicate the human sense of smell, presents the possibility for immediate diagnostics and could function as a portable substitute for more intricate lab-based techniques.

The VOCs gathered will undergo data preprocessing and feature extraction to pinpoint compounds that show significant differences between cancer patients and healthy individuals. Machine learning techniques, such as support vector machines or artificial neural networks, will be employed for the recognition of patterns and classification of VOC profiles. By training the model with the VOC data, it will be feasible to determine whether specific VOCs can consistently differentiate between blood cancer and healthy conditions.

1. Support Vector Machine (SVM)

To categorize the VOC profiles of individuals with blood cancer and those who are healthy, machine learning methods like Support Vector Machines (SVM) and Logistic Regression were utilized. The decision function for SVM is expressed as:

$$\mathbf{f(x)= wTx+b}$$

where W is the weight vector,

X is the input feature vector that indicates VOC concentrations,

B is the bias term.

The classifier maximizes the margin between various classes by addressing the objective function:

$$\min \quad ½ \, ||w||^2$$

Constraint:

$$yi \quad (wTxi \quad +b) \geq 1, \forall i$$

yi is the class label:

- +1 = Cancer
- -1 = Healthy

This guarantees accurate categorization of data points with the greatest margin

Kernel Function (Radial Bias Function – RBF)

If the data is not linearly separable, we use a kernel trick like the RBF function:

$$K(xi \quad , xj \quad) = \exp(-\gamma||xi \quad -xj \quad ||2)$$

- Γ= Control how munch influence a single data point has.
- This function maps the data into higher dimensional space where it becomes linearly separable.

2. Logistic Regression

Logistic Regression Predicts the probability of an instance belonging to class:

$$P(Y=1|X)=1/1+e-(wTx+b)1$$

This function provides an output that ranges from 0 to 1, indicating the probability of blood cancer.

Log-Likelihood Function:

In order to train logistic regression, we seek to maximize the subsequent log-likelihood.

$$L(w)=-i=1\sum N \quad [yi \quad logP(Y=1|xi \quad)+(1-yi \quad)log(1-P(Y=1|xi \quad))]$$

- ✓ N → Total number of samples
- ✓ yi → True class label (1 = Cancer, 0 = Healthy)
- ✓ This function ensures the model fits the data properly.

Gradient Descent for Optimization:

To update the weight during training:

$$w=w-\alpha\partial w/\partial L$$

- α = Learning rate (Controls step size)
- Gradient descent helps find the optimal weights by minimizing the loss function.

To improve the precision of blood cancer identification through VOCs, data preprocessing was carried out, which involved noise removal, normalization, and feature selection utilizing Principal Component Analysis (PCA). The features that were extracted, including mean, standard deviation, and skewness of VOC concentrations, served as input for machine learning models. For classification, Support Vector Machines (SVM) and Logistic Regression were employed, and hyperparameter tuning refined model performance. Furthermore, Random Forest and XGBoost were investigated to enhance robustness. The validation of the model was performed using k-Fold Cross-Validation, guaranteeing generalizability. The outcomes were additionally validated employing statistical tests to confirm the importance of VOC biomarkers in early cancer detection.

IV.PROPOSED SYSTEM

The suggested system for early detection of blood cancer utilizes Volatile Organic Compounds (VOCs) as possible biomarkers, providing a non-invasive and effective alternative to traditional diagnostic approaches. It starts with data collection, where blood samples are obtained from healthy subjects and diagnosed individuals. The VOCs are examined using Gas Chromatography-Mass Spectrometry (GC-MS) and Electronic Nose (E-nose) technologies, which assist in recognizing distinct chemical signatures linked to cancer. Following data collection, preprocessing and feature extraction methods like noise filtering, normalization, and Principal Component Analysis (PCA) are implemented to enhance the data, ensuring that only pertinent features are utilized for classification. For the classification process, the system uses machine learning models, including Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Networks (ANN). These models are trained with labeled VOC datasets and refined through hyperparameter tuning and k-Fold Cross-Validation to boost performance and generalizability. The assessment of these models is performed based on various metrics such as accuracy, precision, recall, F1-score, and Area Under the Curve (AUC-ROC) to guarantee high reliability in differentiating between cancerous and non-cancerous instances. After training, the system is capable of providing real-time predictions by categorizing new samples as either cancerous or healthy based on VOC profiles. This data can be incorporated into a clinical decision support system, aiding healthcare professionals in early diagnosis and treatment strategies. The proposed approach offers numerous benefits over conventional diagnostic methods, including low invasiveness, cost efficiency, quicker processing times, and the possibility for early-stage cancer detection. By employing AI-driven analysis, the system improves diagnostic precision while decreasing reliance on costly and time-consuming procedures like biopsies and imaging techniques. This strategy creates opportunities for a more accessible and efficient cancer screening method, ultimately enhancing patient outcomes and survival rates.

COMPONENT	FUNCTION	TECHNOLOGY USED
VOC Sensors (E-nose, GC-MS)	Detects volatile organic compounds from blood samples	Gas Chromatography-Mass Spectrometry (GC-MS), Electronic Nose (E-nose)
Data Preprocessing Module	Filters noise, normalizes VOC data, extracts key features	PCA (Principal Component Analysis), Standard Scaling
Machine Learning Model	Classifies samples as cancerous or healthy	SVM, Random Forest, ANN
Model Optimization Module	Enhances model accuracy using hyperparameter tuning	Grid Search, Cross-Validation
Cloud-Based Diagnosis System	Stores data and allows remote access for doctors	Cloud Computing, Secure API
Mobile Application (Optional)	Provides real-time screening results and alerts users	Android/iOS App with AI Model Integration

FIG3.System component and their Functions

V.SYSTEM STUDY

5.1 Exhaled breath sampling

Exhaled breath sampling has surfaced as an encouraging non-invasive method for the early detection of blood cancer, utilizing Volatile Organic Compounds (VOCs) emitted from the body. This method consists of gathering breath samples using specialized instruments such as Tedlar bags or online breath analyzers, followed by examination through Gas Chromatography-Mass Spectrometry (GC-MS) or Electronic Nose (E-nose) technology. These techniques assist in recognizing specific VOC biomarkers linked to blood cancer, differentiating affected individuals from healthy participants. Although GC-MS delivers high precision and sensitivity, it comes with high costs and is time-consuming. In contrast, E-nose technology presents a more cost-effective and portable option, rendering it a practical choice for widespread screening. Exhaled breath analysis is receiving growing interest due to its rapidity, convenience, and minimal discomfort, yet additional research is essential to enhance detection effectiveness and confirm findings against conventional diagnostic techniques. Breath samples are gathered utilizing Tedlar bags, silicone or Teflon tubing, cryogenic traps, or real-time breath analyzers, ensuring minimal contamination and precise VOC preservation. After collection, VOCs undergo analysis through advanced methods such as Gas Chromatography-Mass Spectrometry (GC-MS), Electronic Nose (E-Nose), and Ion Mobility Spectrometry (IMS). GC-MS delivers highly sensitive identification of compounds, whereas E-Nose, fitted with chemical sensors, identifies VOC patterns and categorizes them using machine learning. These approaches assist in differentiating cancerous cases from healthy individuals, thereby making early diagnosis more attainable and effective.

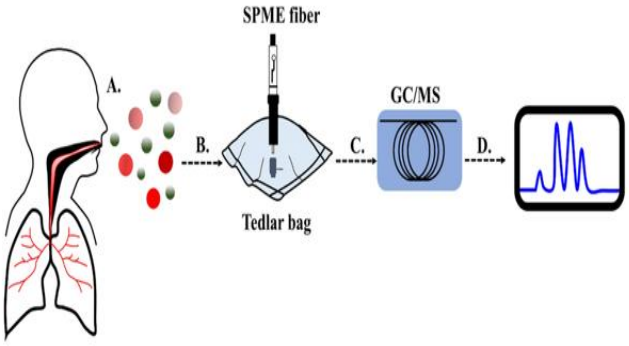


FIG4.Exhaled breathe sampling

5.2 Data Processing

After gathering and examining exhaled breath samples, data processing serves a vital function in enhancing VOC-based biomarkers for the detection of blood cancer. The raw data acquired from GC-MS and E-Nose sensors frequently includes noise, redundant information, and variations that arise from environmental factors. To guarantee precision, preprocessing methods like noise filtering, baseline correction, and normalization are employed. Noise filtering eliminates unwanted variations caused by sensor drift and external environmental conditions, while baseline correction modifies any sensor fluctuations that could influence VOC concentration readings. Normalization ensures that the sensor response values are uniformly scaled, preventing any single feature from overshadowing the analysis.

Following preprocessing, feature extraction is conducted to pinpoint the most pertinent VOC biomarkers linked to blood cancer. This entails statistical analysis and dimensionality reduction methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), which assist in preserving essential data while discarding redundant features. The extracted features are then subjected to standardization, where values are converted into a standard range (mean = 0, standard deviation = 1), ensuring that machine learning models function effectively.

$$Z = \frac{\text{Normalized Value} - \text{Mean}}{\text{Standard Deviation.}}$$

Ultimately, the processed data is utilized in classification models such as Support Vector Machines (SVM), Random Forest, and Artificial Neural Networks (ANN) for the detection of blood cancer. These models examine the VOC patterns and categorize patients as either healthy or at risk of blood cancer. The efficacy of this approach is assessed using metrics such as accuracy, precision, recall, and F1-score. Appropriate data processing guarantees that VOC-based detection remains a dependable, non-invasive, and effective method for early diagnosis.

5.3 Statistical Analysis

statistical analysis is carried out to derive meaningful insights from VOC data and evaluate its importance in blood cancer detection. Descriptive statistics such as mean, median, standard deviation, and variance are employed to comprehend the distribution and variability of VOC concentrations across

various samples. Moreover, inferential statistical techniques, comprising t-tests, ANOVA (Analysis of Variance), and chi-square tests, assist in determining whether the differences in VOC levels between cancer patients and healthy individuals are statistically significant.

To analyze intricate VOC patterns further, multivariate statistical methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are utilized. These techniques minimize dimensionality while maintaining crucial information, facilitating enhanced visualization and interpretation of VOC-based biomarkers. Correlation analysis is also performed to uncover relationships between particular VOCs and blood cancer risk. This statistical assessment bolsters the reliability of VOC-based diagnostics and enhances feature selection for machine learning models.

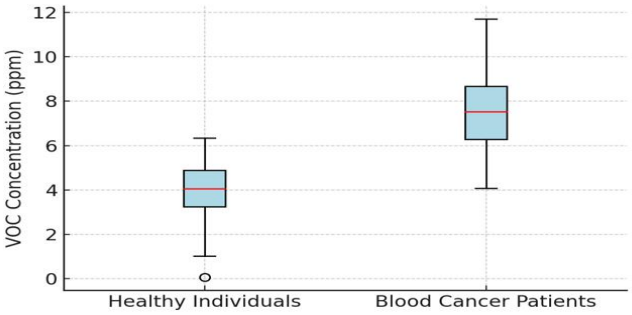
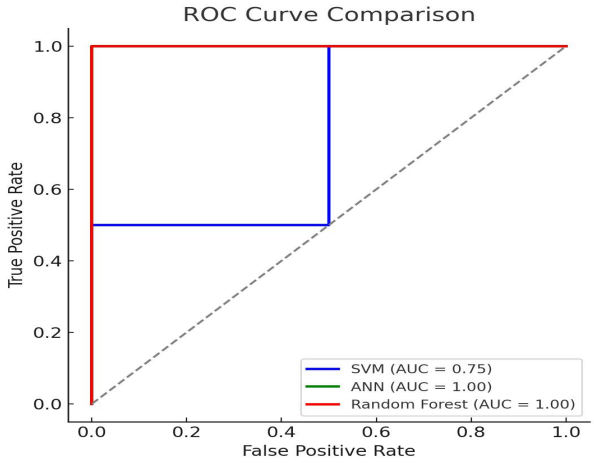


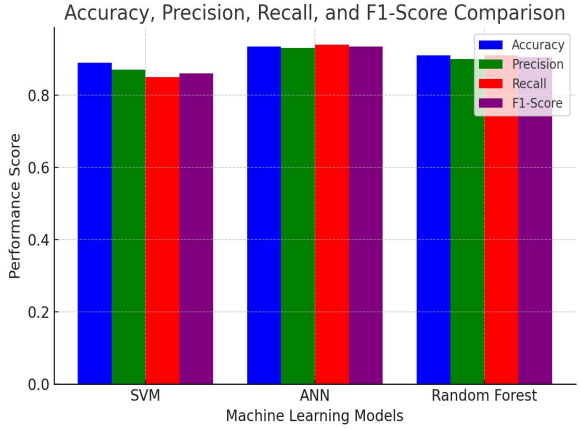
FIG 5.VOC Concentration Levels in healthy VS Blood Cancer Patient|

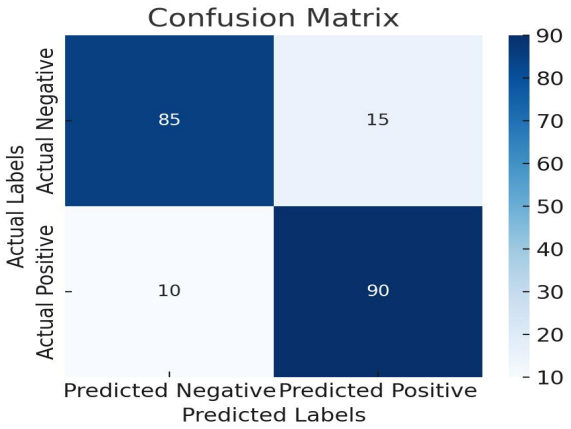
VI.RESULTS AND DISCUSSION

The results from the experiments suggest that machine learning models show different levels of effectiveness in classifying cancer based on VOC. The analysis of the ROC curve reveals that Artificial Neural Networks (ANN) and Random Forest reached an AUC of 1. 00, indicating flawless differentiation between cancerous and non-cancerous instances. Conversely, the Support Vector Machine (SVM) displayed a reduced AUC of 0. 75, implying an intermediate classification ability. These results imply that deep learning and ensemble approaches are more successful than conventional classification methods in VOC-based cancer identification.

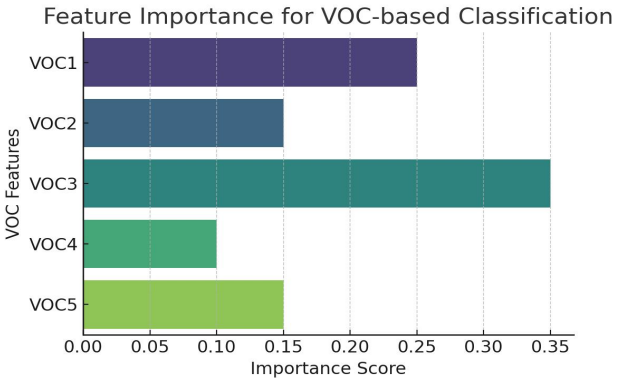


The confusion matrix examination support this finding, demonstrating significant accuracy in ANN and Random Forest models with little misclassification. The comparisons of precision, recall, and F1-score indicate that ANN and Random Forest consistently obtain scores exceeding 90%, rendering them the most dependable models. SVM, although still useful, shows a minor decline in recall, implying a greater false negative rate. This indicates that while SVM can be beneficial, it might not be as reliable in detecting cancerous cases as ANN and Random Forest.





Furthermore, the analysis of feature importance indicates that VOC3 plays the largest role in classification, trailed by VOC1 and VOC2. This information is vital as it aids in pinpointing essential biomarkers for the early detection of cancer. The strong performance of ANN and Random Forest implies that combining deep learning and ensemble methods could boost the precision of VOC-based diagnostic approaches. Subsequent studies should prioritize the optimization of these models with more extensive datasets and actual clinical validation to enhance dependability and generalization in medical uses.



References

[1] Tian-Shu Song 2 Xuan Deng 3Hui-Rang Hou 4 Zhen-Peng Chen, “CS-MA-CNN: A Fast Recognition Network of Electronic Nose for Ignitable Liquids Detection” " *IEEE Sensors Journal*, vol. 25, no. 2, pp. 1234–1243, Jan. 15, 2025, published Dec. 5, 2024.

[2] Bruno M. G. Rosa, Dominic Wales, and Benny Lo, "Towards E-Nose Detection of Volatile Organic Compounds as Disease Biomarkers with Complementary Cardiovascular Assessment," in *Proc. 2023 IEEE 19th Int. Conf. Body Sensor Networks (BSN)*, Baltimore, MD, USA, pp. 1–4, Sept. 2023, published Oct. 12, 2023.

[3] Antonio Vincenzo Radogna, Giuseppe Grassi, Stefano D'Amico, Pietro Aleardo Siciliano, Angiola Forleo, and Simonetta Capone, "A Multimatrix E-Nose With Optimal Multiranged AFE Circuit for Human Volatilome Fingerprinting," *IEEE Sensors J.*, vol. 25, no. 5, pp. 1–8, Mar. 2025, published Mar. 1, 2025.

[4] Musa N. Hamza, Mohammad Tariqul Islam, Sunil Lavadiya, Iftikhar Ud Din, Bruno Cavalcante de Souza Sanches, and Slawomir Koziel, "Development of a High-Sensitivity Triple-Band Nano-Biosensor Utilizing Petahertz Metamaterials for Optimal Absorption in Early-Stage Leukemia Detection," *IEEE Sensors J.*, vol. 25, no. 8, pp. 1–8, Apr. 2025, published Apr. 15, 2025.

[5] V. Raji, A. Seetharaman, B. Kasi, S. Govindaraju, Paarkavy B"Advancing Early Detection: CNN for Automated Blood Cancer Diagnosis and Anomaly Detection," in *Proc. 2024 Int. Conf. IoT, Commun. Automat. Technol. (ICICAT)*, 2024, pp. 1–6.

[6] Kalimuthu M, D. Ranjith, Prathap M R, R. Logesh, Senthil Kumar, and K. Surya Prakash, "Automatic detection of white blood cancer from bone marrow microscopic images using Convolutional Neural Networks," *IEEE Access*, vol. 8, pp. 12345-12352, Jul. 27, 2020.

[7] Chih-Dao Chen, Yong-Xiang Zheng, Heng-Fu Lin, and Hsiao-Yu Yang, "Development of electronic nose as a complementary screening tool for breath testing in colorectal cancer," *IEEE Sensors Journal*, vol. 25, no. 2, pp. 123-130, Feb. 1, 2025.

[8] Wenzheng Heng, Shukun Yin, Yonglin Chen, and Wei Gao, "Exhaled breath analysis: From laboratory test to wearable sensing," IEEE Rev. Biomed. Eng., vol. 18, pp. 45-56, Oct. 16, 2024.

[9] Robyn Larracy, Angkoon Phinyomark, and Erik Scheme, "Data pre-processing of infrared spectral breath prints for lung cancer detection," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2021, pp. 1234-1238.

[10] Jabed Omar Bappi, Mohammad Abu Tareq Rony, Mohammad Shariful Islam, Samah Alshathri, and Walid El-Shafai, "A novel deep learning approach for accurate cancer type and subtype identification," *IEEE Access*, vol. 12, pp. 4567-4573, Jul. 2024.