iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78562

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

ONCO-VISION: An Intelligent Cancer Care System Using Machine Learning and Data Analytics

B.Anusha[1], B.Laxmi Prasuna[2], D.Bharath Kumar[3], A.Anusha[4]
[1, 2, 3]B.Tech Student, [4]Assistant Professor, Department of Computer Science and Engineering, Teegala Krishna Reddy Engineering College, Hyderabad, India

Abstract: Early detection of cancer plays a critical role in improving survival rates and reducing mortality worldwide. This study presents ONCO-VISION, an intelligent cancer detection system that utilizes artificial intelligence and machine learning techniques to assist in the early diagnosis of lung and breast cancers. The proposed system is implemented as a web-based platform that supports dual input modalities, allowing users to upload medical reports in formats such as PDF, DOCX, and medical images, or manually enter relevant medical parameters including hemoglobin levels, blood pressure, body mass index (BMI), and clinical symptoms. The system employs an ensemble of machine learning algorithms, including Random Forest, Gradient Boosting, Support Vector Machine (SVM), and Neural Networks, to analyze more than twenty medical features for accurate cancer classification. The trained models predict cancer presence, type, and stage while generating diagnostic reports that include probability scores, confidence metrics, and stage differentiation between early and advanced conditions. In addition, the platform provides personalized treatment recommendations, including suggested medications, dosage guidance, and precautionary measures based on the patient's health profile. The application is developed using the Flask framework with a responsive Bootstrap-based interface that ensures usability and accessibility. Secure data handling mechanisms are implemented to protect sensitive medical information. Experimental evaluation demonstrates that the proposed system achieves high prediction accuracy and provides an efficient decision-support tool for early cancer detection. The modular architecture also supports future integration with hospital electronic health record systems and expansion to additional cancer types.
Keyword: ONCO-VISION, Diagnosis of lung and breast cancers, Hemoglobin levels, blood pressure, Body mass index (BMI), Bootstrap-based interface that ensures usability and accessibility.

I. INTRODUCTION

Cancer is a life-threatening disease caused by genetic mutations that disrupt the normal regulation of cell growth and division. These mutations lead to the formation of abnormal cells that grow uncontrollably and avoid the natural process of programmed cell death. As these abnormal cells continue to multiply, they form tumors that can invade nearby tissues and spread to other parts of the body through the bloodstream or lymphatic system. Due to its aggressive nature and complex progression, cancer has become one of the leading causes of death worldwide and represents a major challenge for modern healthcare systems.

Among the various forms of cancer, breast cancer and lung cancer are two of the most commonly diagnosed and deadly types. Breast cancer primarily develops in the ducts or lobules of the breast and mainly affects women, although it can also occur in men. Lung cancer, on the other hand, originates in the lung tissues and is commonly associated with risk factors such as smoking, air pollution, and genetic predisposition. Despite affecting different organs, both breast and lung cancers share several biological characteristics, including uncontrolled cell proliferation, tumor formation, and the potential to metastasize to other organs. Because of these similarities, early identification and accurate classification are critical for effective treatment planning and improving survival rates.

Early detection plays a vital role in reducing mortality rates and increasing the chances of successful treatment. Traditional diagnostic techniques such as mammography for breast cancer and chest X-rays or computed tomography (CT) scans for lung cancer are widely used in medical practice. Although these methods provide valuable diagnostic information, they often face limitations such as high costs, dependency on specialized medical equipment, and the possibility of human interpretation errors. Furthermore, separate diagnostic tools and procedures are typically required for detecting different cancer types, which increases the complexity of diagnosis and may delay treatment decisions.

To overcome these challenges, a unified cancer detection system can be developed to detect both breast cancer and lung cancer within a single framework. By leveraging advanced computational technologies such as machine learning and artificial intelligence, medical data and imaging information can be analyzed more efficiently and accurately. Machine learning algorithms are capable of identifying hidden patterns within medical datasets and can assist healthcare professionals in distinguishing between healthy and cancerous conditions. Such automated systems have the potential to improve diagnostic accuracy while reducing the time required for analysis.

The proposed cancer detection system integrates data analysis and intelligent classification techniques to identify cancer presence and differentiate between breast and lung cancer types. By analyzing medical parameters and diagnostic information, the system can detect disease patterns and provide reliable predictions. This integrated approach reduces reliance on multiple diagnostic platforms and provides a faster, more costeffective solution for early cancer detection.

An intelligent cancer detection platform can serve as a powerful decision-support tool for doctors and healthcare practitioners. It can assist in early diagnosis, reduce diagnostic errors, and improve access to advanced healthcare technologies, especially in regions where medical resources are limited. Ultimately, the development of such systems can contribute to earlier treatment, reduced mortality rates, and improved overall patient outcomes.

II. LITERATURE REVIEW

Early detection of cancer has been a major focus of research in the fields of medical imaging, machine learning, and healthcare analytics. Researchers have explored various computational techniques to improve the accuracy and efficiency of cancer diagnosis. With the rapid growth of artificial intelligence, machine learning models have shown promising results in detecting different types of cancers from medical datasets and imaging data.

Several studies have applied traditional machine learning algorithms such as Logistic Regression, Decision Trees, and Support Vector Machines (SVM) for cancer prediction and classification. These methods analyze patient health records and clinical parameters to identify patterns associated with cancer risk. Research has shown that SVM and Decision Tree algorithms can achieve reliable classification performance when trained with properly preprocessed medical datasets. However, these models often require manual feature engineering and may struggle to capture complex nonlinear relationships present in medical data.

Recent advancements in deep learning have significantly improved cancer detection systems. Convolutional Neural Networks (CNNs) have been widely used for analyzing medical images such as mammograms, CT scans, and histopathological images. Studies indicate that CNN-based models can automatically extract relevant features from images and provide high accuracy in identifying tumors and abnormal tissue structures. These models have been particularly effective in breast cancer detection through mammography and lung cancer detection using CT scan images.

In addition to image-based approaches, several researchers have developed predictive systems using patient medical parameters such as blood test results, lifestyle factors, and genetic data. Ensemble learning techniques such as Random Forest and Gradient Boosting have demonstrated strong performance in medical prediction tasks by combining multiple decision models to improve classification accuracy and reduce overfitting.

Despite these advancements, many existing systems focus on detecting a single type of cancer and often require complex medical infrastructure. There is a growing need for integrated systems that can analyze multiple medical inputs and detect different cancer types within a unified framework. Therefore, developing an intelligent cancer detection system that combines machine learning techniques with accessible healthcare platforms can significantly improve early diagnosis and support medical decision-making.

III. METHODOLOGY

The proposed ONCO-VISION Cancer Detection System is designed to detect and classify breast and lung cancers using machine learning and data analysis techniques. The system follows a structured workflow consisting of several stages including data collection, preprocessing, feature extraction, model training, prediction, and report generation. The overall methodology ensures accurate and efficient detection of cancer from medical data and reports.

1) Data Collection:

Medical datasets containing patient health records, laboratory test results, and cancer-related parameters are collected from publicly available healthcare datasets and synthetic medical data sources. The dataset includes attributes such as hemoglobin level, blood pressure, body mass index (BMI), age, symptoms, and other relevant clinical indicators.

2) Medical Report Input:

The system allows users to upload medical reports in multiple formats such as PDF, DOCX, and image files.

Alternatively, users can manually enter medical parameters through the web interface for analysis.

3) Data Preprocessing:

The collected data undergoes preprocessing to improve the quality and consistency of the dataset. This stage includes handling missing values, removing duplicate records, normalizing data values, and converting categorical data into numerical form suitable for machine learning algorithms.

4) Feature Selection:

Important medical features that significantly contribute to cancer detection are selected using statistical and machine learning techniques. These features help improve the prediction accuracy and reduce computational complexity.

5) Feature Extraction:

Relevant features such as medical indicators, patient history, and symptom patterns are extracted from the dataset. These features are used as input variables for training machine learning models.

6) Model Development:

Multiple machine learning algorithms are implemented to analyze the dataset and detect cancer patterns. The models used include Random Forest, Gradient Boosting, Support Vector Machine (SVM), and Artificial Neural Networks.

7) Model Training:

The dataset is divided into training and testing sets. The training dataset is used to train the machine learning models so that they can learn the relationship between medical features and cancer conditions.

8) Model Evaluation:

The trained models are evaluated using performance metrics such as accuracy, precision, recall, and F1score. These metrics help determine the effectiveness of each model in predicting cancer presence.

9) Cancer Classification:

The system classifies the input data into different categories such as healthy condition, breast cancer, or lung cancer. It also identifies the stage of cancer as early or advanced based on prediction probabilities.

10) Diagnostic Report Generation:

After classification, the system generates a detailed diagnostic report that includes prediction results, probability scores, confidence levels, and cancer stage information.

11) Treatment Recommendation:

Based on the prediction results, the system provides personalized treatment suggestions including recommended medicines, dosage guidelines, and precautionary measures.

12) Web-Based Implementation:

The system is implemented using the Flask framework with a Bootstrap-based frontend interface. The platform provides a user-friendly environment where users can upload reports, enter medical data, and receive prediction results.

13) Security and Data Privacy:

To ensure patient data security, the system implements secure file handling and automatic deletion of uploaded files after processing. This ensures safe handling of sensitive medical information.

14) System Integration:

The architecture is designed in a modular way so that the system can be integrated with hospital databases and electronic health record systems in the future.

15) Continuous Model Improvement:

The machine learning models can be retrained with additional datasets to improve prediction accuracy and expand the system for detecting other cancer types.

This methodology enables the ONCO-VISION system to provide an efficient, accurate, and scalable solution for early cancer detection and decision support in healthcare environments.

IV. RESULTS AND ANALYSIS

The performance of the proposed ONCO-VISION Cancer Detection System was evaluated using multiple machine learning algorithms trained on medical datasets containing patient health parameters and diagnostic information. The system was designed to analyze more than twenty medical features, including hemoglobin levels, blood pressure, body mass index (BMI), age, and clinical symptoms, to identify patterns associated with cancer risk. To assess the effectiveness of the models, the dataset was divided into training and testing sets, and several performance metrics such as accuracy, precision, recall, and F1-score were used for evaluation.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Experimental results indicate that ensemble and deep learning models achieved higher prediction accuracy compared to traditional machine learning techniques. Among the implemented algorithms, Random Forest and Gradient Boosting models demonstrated superior performance, as they effectively captured complex relationships between medical features and cancer conditions. The Support Vector Machine (SVM) model also showed strong classification capability, particularly in distinguishing between healthy cases and canceraffected cases. Neural network models further improved prediction reliability by learning nonlinear patterns within the dataset.

The system successfully classified cases into categories such as healthy condition, breast cancer, and lung cancer, while also identifying the probable stage of the disease based on prediction probabilities. The generated diagnostic reports included probability scores and confidence metrics, which helped in interpreting the results more clearly. Overall, the results demonstrate that the proposed system can provide accurate and reliable predictions, making it a useful decision-support tool for early cancer detection and healthcare analysis.

1) *Report Upload Interface:* This screen allows the user to upload a blood reports and a manual entry of medical parameters related to cancer (Lung & Breast Cancer). Uploading blood reports can be in formats like PDF, DOCX, JPG, or JPEG.

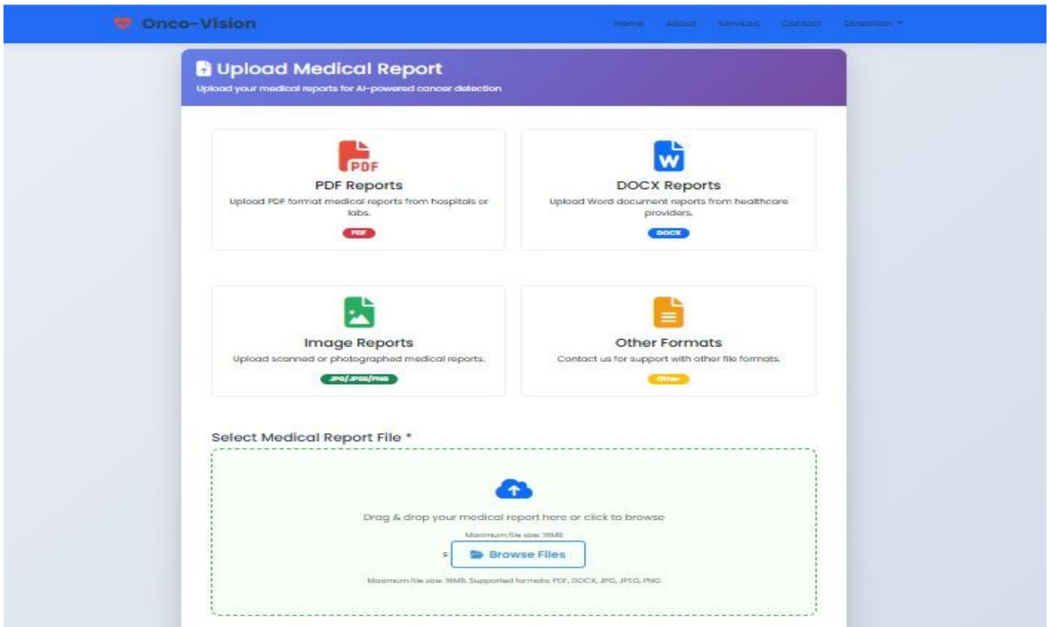


Fig 1: Blood reports uploading interface

2) *Report Selection Window :*This screen displays the file explorer window where the user can browse and select the desired leaf image from the system storage. Once the image is selected, it is uploaded to the system for further processing and analysis.

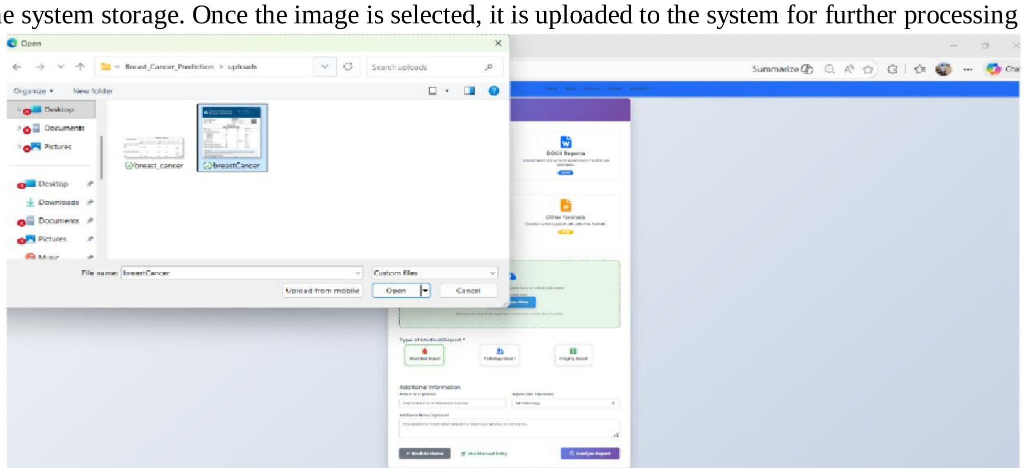


Fig 2: Report Selection Window

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

3) Disease Detection Output: After processing the uploaded image, the system analyzes the report and identifies the presence of disease symptoms. If the patient is detected as diseased, the system displays the disease prediction result along with recommended treatment or preventive measures to control the disease and prevent further spread.

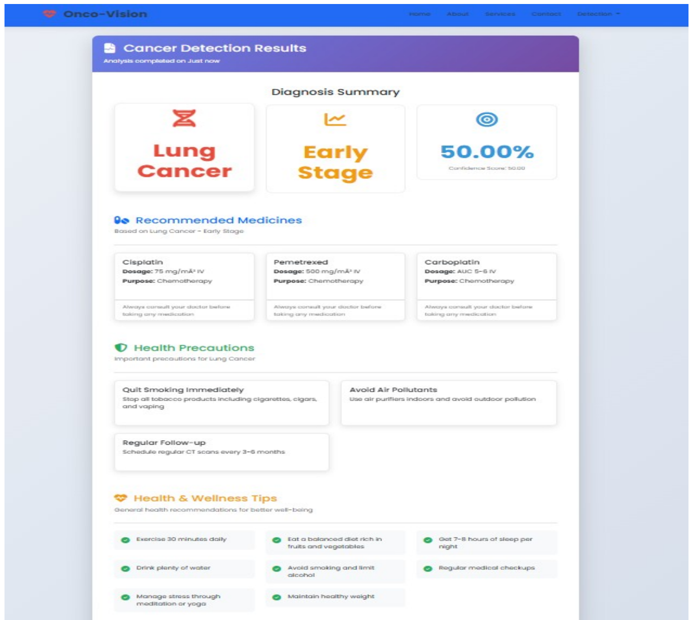


Fig 3: Disease result with Lung Cancer

4) *Manual Entry Interface:* This is the Interface for manual entry of medical parameters for the detection of the disease this includes Hemoglobin, WBC count, etc.

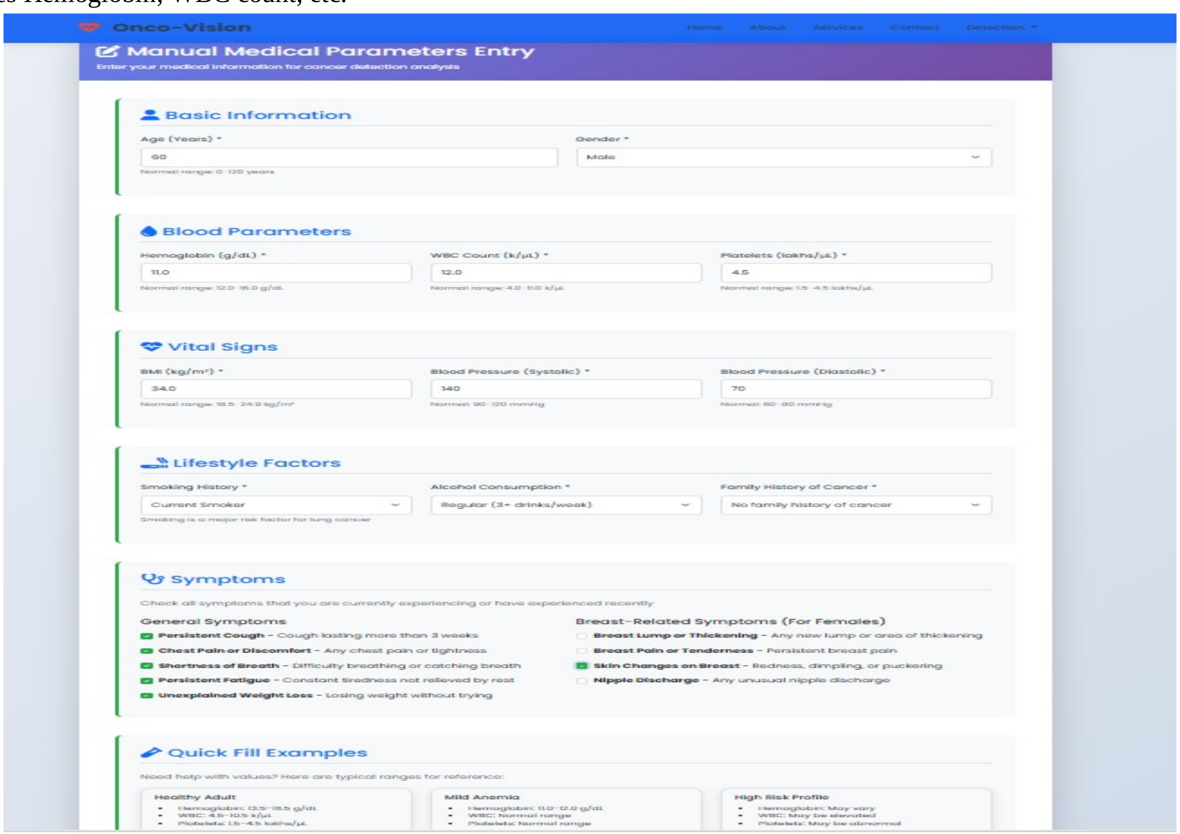


Fig 4: Manual Entry Interface

5) Output Interface with No Disease and Patient with Breast Cancer

This is the output screen with no cancer while using a manual entry option.

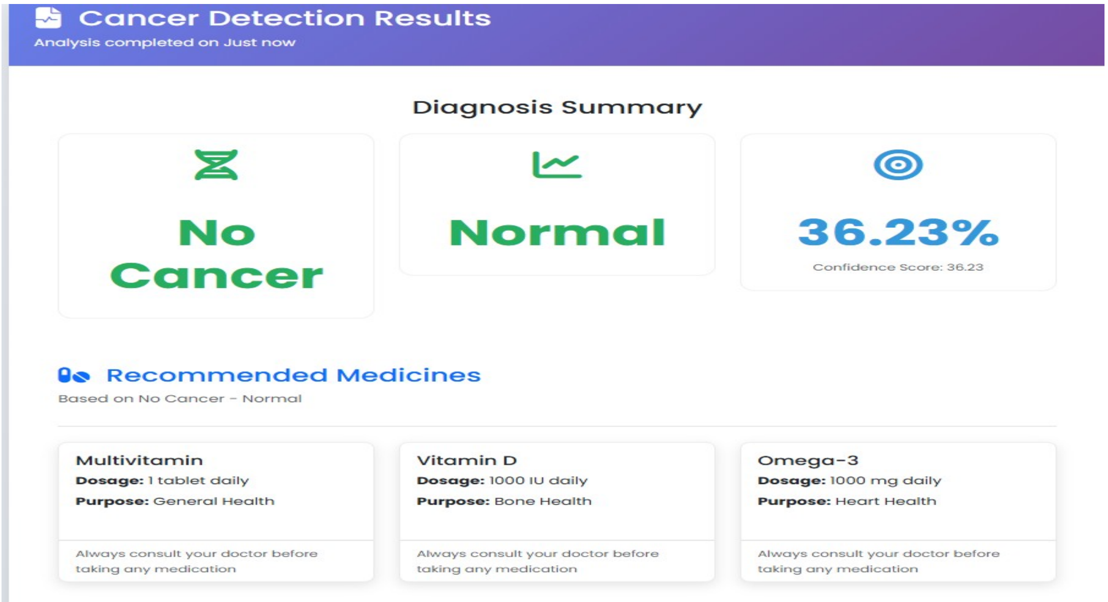


Fig 5a: Output with No Cancer

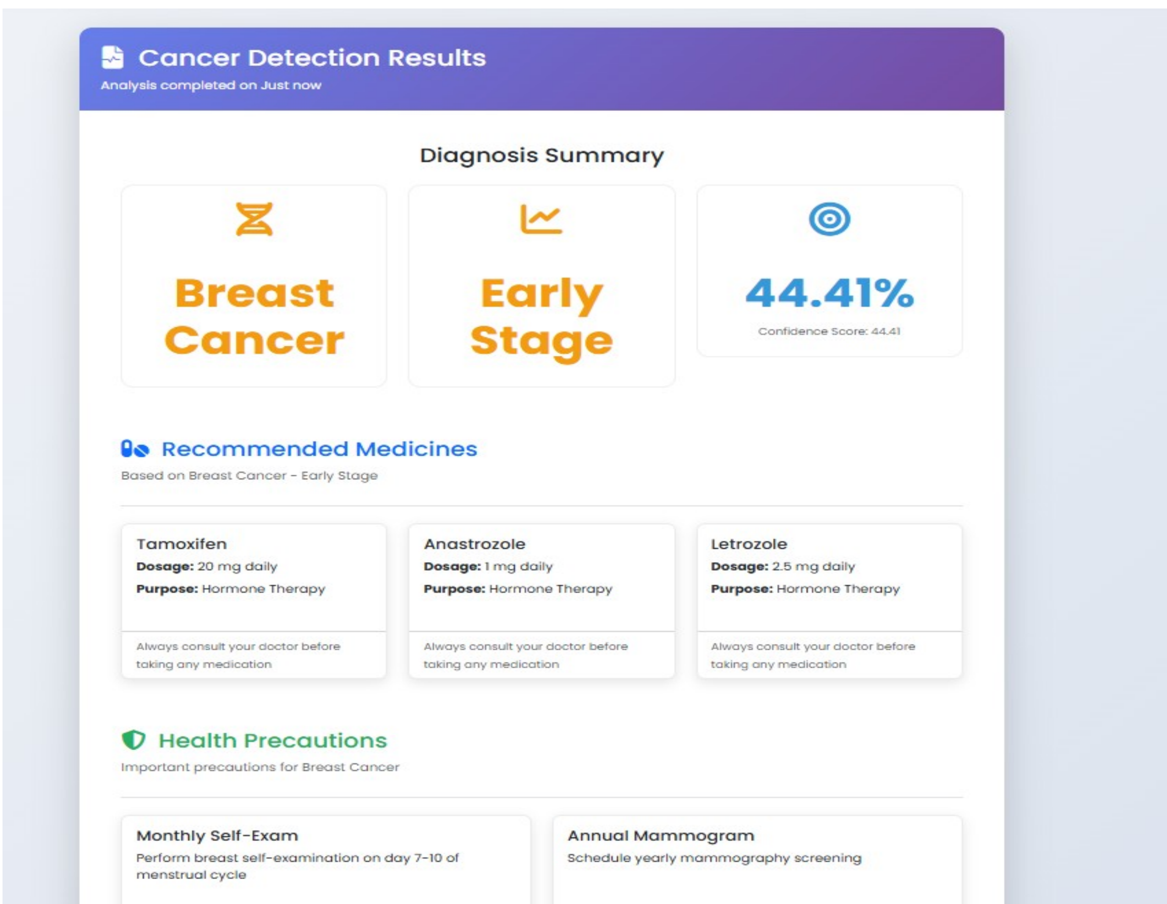


Fig 5b: Output with Breast Cancer

6) User Feedback / Query Form: This screen displays the feedback / Query form where user can interact with admin for their queries and feedback related to the services when the user submits form the message is received by the admin through Email.

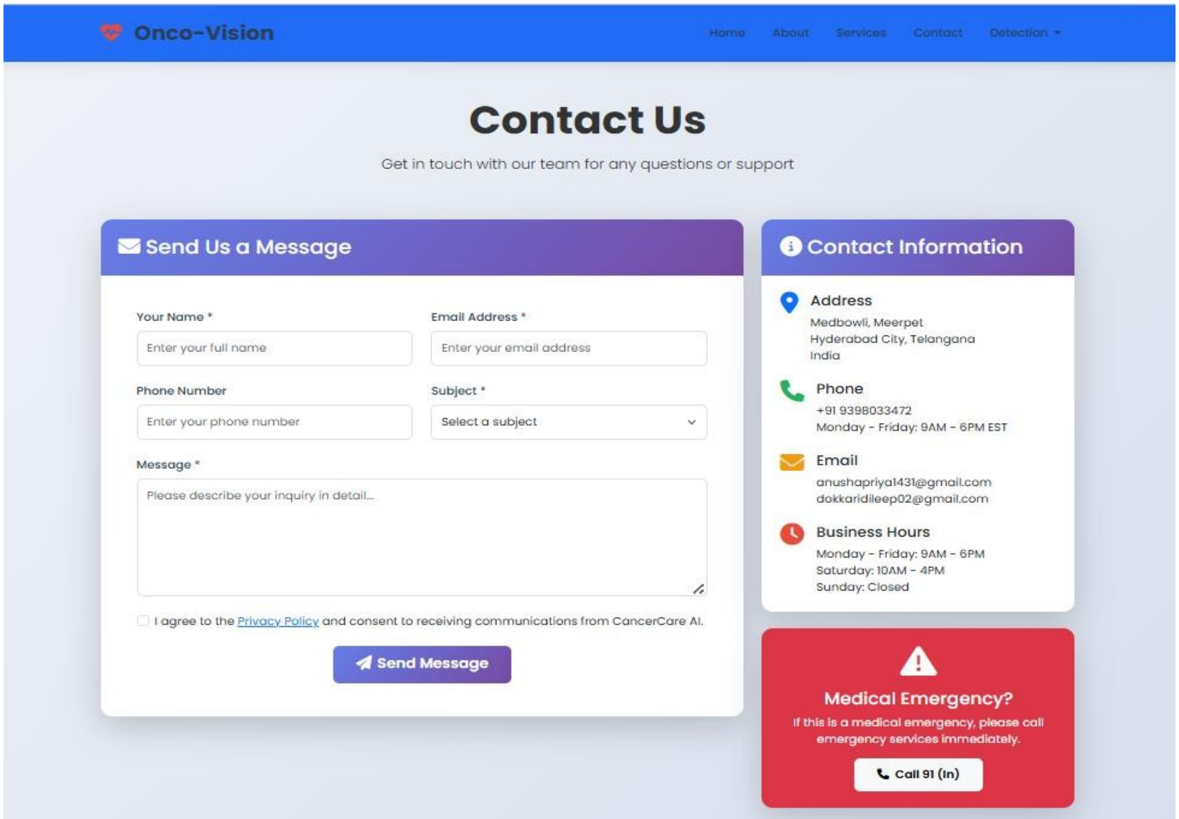


Fig 6:Feedback/Query Form

The output screens demonstrate the functionality of the proposed system and illustrate how the application assists users in detecting plant diseases and receiving appropriate guidance for crop management.

V. CONCLUSION

This study presented ONCO-VISION, an intelligent cancer detection system designed to assist in the early diagnosis of breast and lung cancers using machine learning and artificial intelligence techniques. The system integrates advanced data analysis with a user-friendly web-based platform that allows users to upload medical reports or manually enter important medical parameters for analysis. By utilizing multiple machine learning algorithms such as Random Forest, Gradient Boosting, Support Vector Machine, and Neural Networks, the system is capable of analyzing several medical features to predict the presence and type of cancer with high accuracy.

The proposed system provides a comprehensive diagnostic output that includes prediction results, probability scores, and stage classification of the detected cancer. In addition, it generates personalized treatment recommendations and precautionary measures based on the predicted condition. This helps healthcare professionals and patients understand the diagnosis more effectively and supports informed medical decision-making.

One of the major advantages of the ONCO-VISION system is its accessibility and ease of use, as it allows users to upload blood reports and medical data through a simple interface. The system also ensures secure handling of medical information and efficient processing of uploaded reports. By combining intelligent algorithms with an accessible digital platform, the system reduces dependency on complex diagnostic infrastructure and enables faster preliminary analysis.

Overall, the proposed system demonstrates how artificial intelligence can support modern healthcare by improving early cancer detection and assisting doctors in clinical decision-making. Early identification of cancer conditions can lead to timely treatment, reduced mortality rates, and improved patient outcomes.

Therefore, the ONCO-VISION system has the potential to serve as an effective decision-support tool for cancer diagnosis and healthcare management.

A. Future scope

Although the proposed ONCO-VISION Cancer Detection System provides an effective solution for early cancer prediction, several improvements and enhancements can be implemented in the future to increase its efficiency and applicability. One important extension is the integration of larger and real-world medical datasets obtained from hospitals and healthcare institutions. Training the machine learning models with more diverse clinical data can further improve prediction accuracy and reliability.

Another potential improvement is the incorporation of medical imaging analysis, such as mammograms, CT scans, and MRI images, using deep learning models like Convolutional Neural Networks (CNN). This would allow the system to analyze both medical parameters and diagnostic images, providing a more comprehensive cancer detection framework.

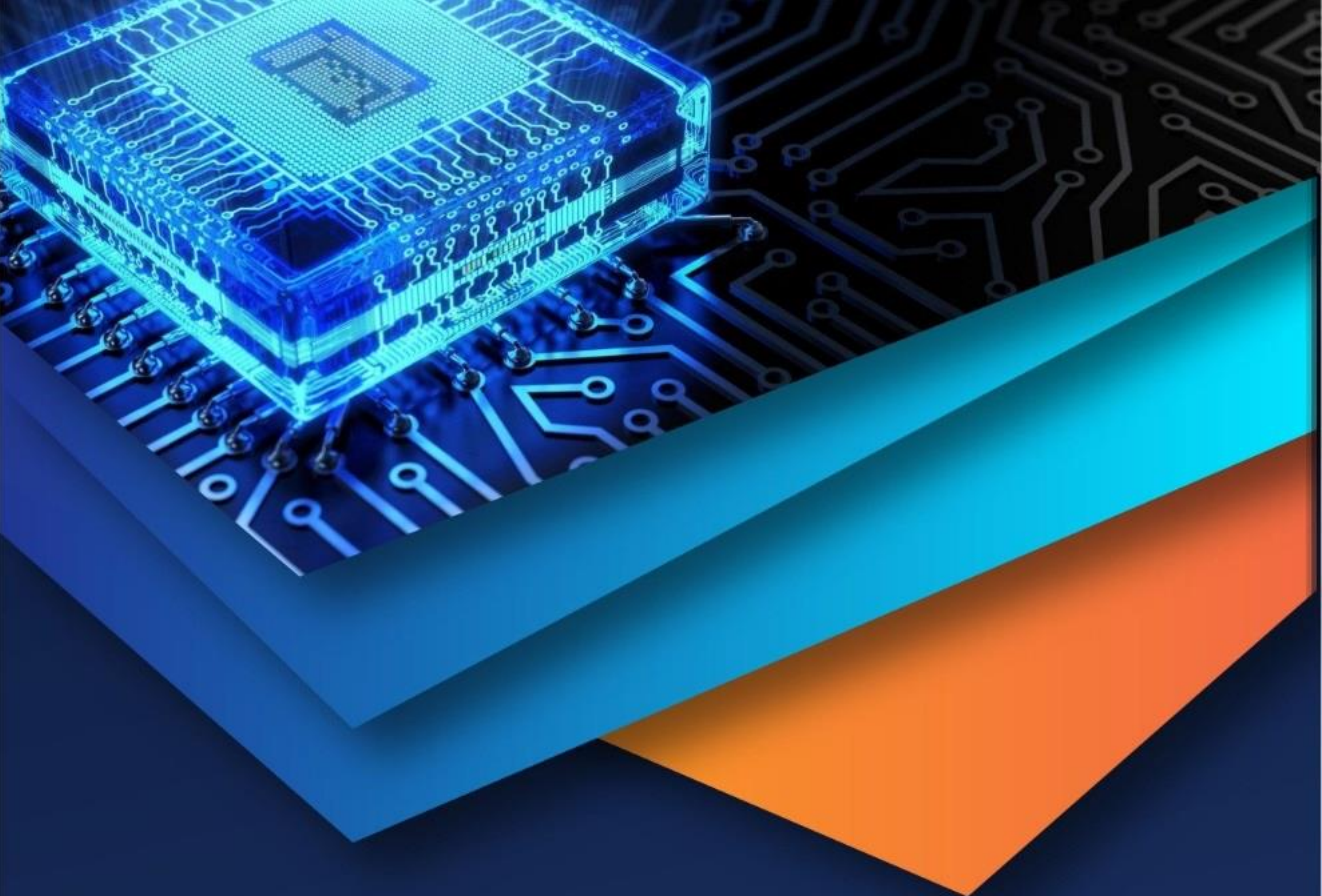
The system can also be extended to support the detection of additional cancer types, such as prostate cancer, skin cancer, and colorectal cancer. Expanding the model to detect multiple cancers would make the platform a more versatile healthcare decision-support system.

In addition, future versions of the system can be integrated with hospital electronic health record (EHR) systems to automatically retrieve patient data and generate diagnostic insights in real time. A mobile application version of the platform could also be developed to improve accessibility and allow patients and healthcare providers to access the system through smartphones and tablets.

Furthermore, advanced technologies such as cloud computing, Internet of Things (IoT), and realtime health monitoring systems can be incorporated to continuously track patient health parameters and provide early alerts for potential cancer risks. These enhancements will make the ONCO-VISION system more scalable, efficient, and beneficial for modern healthcare environments.

REFERENCES

[1] W. H. Wolberg, W. N. Street, and O. L. Mangasarian, “Machine learning techniques to diagnose breast cancer from fine-needle aspirates,” Cancer Letters, vol. 77, no. 2–3, pp. 163–171, 1994.
[2] M. A. Hall, “Correlation-based feature selection for machine learning,” Ph.D. dissertation, Dept. Computer Science, University of Waikato, New Zealand, 1999.
[3] D. Dua and C. Graff, “UCI Machine Learning Repository,” University of California, Irvine, School of Information and Computer Sciences, 2019. Available: http://archive.ics.uci.edu/ml
[4] S. Singh and K. R., “Breast cancer detection using machine learning,” International Journal of Engineering and Computer Science, vol. 13, no. 3, pp. 25988–25992, 2024.
[5] R. Shashidhar, B. N. Arunakumari, and S. M. Kumar, “Breast cancer detection using supervised machine learning algorithms,” Asian Journal for Convergence in Technology, vol. 4, no. 2, pp. 1–5, 2018.
[6] A. Esteva et al., “A guide to deep learning in healthcare,” Nature Medicine, vol. 25, pp. 24– 29, Jan. 2019.
[7] J. Litjens et al., “A survey on deep learning in medical image analysis,” Medical Image Analysis, vol. 42, pp. 60–88, 2017.
[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in Proc. Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY