Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

Machine Learning Based Tumor Detection System for Breast Cancer Classification

Chaitanya Devendra Hingmire
Smt. Kashibai Navale College of Engineering-SPPU
Pune-411046,India

Shripad Tulshidas Pawar
Smt. Kashibai Navale College of Engineering-SPPU
Pune-411046,India

Suyash Shankar Dhumne
Smt. Kashibai Navale College of Engineering-SPPU Pune-411046,India

Datta Suresh Rahegaonkar
Smt. Kashibai Navale College of Engineering-SPPU
Pune-411046,India

Research Guide:
Dr. Lalit V. Patil
Smt. Kashibai Navale College of Engineering-SPPU
Pune-411046,India

Abstract - This research focuses on the development of an AI-powered tumor detection system that leverages machine learning algorithms for the accurate and early diagnosis of breast cancer. Early detection of tumors plays a crucial role in improving survival rates and enabling timely medical treatment. In this project, the model analyzes features derived from the Wisconsin Breast Cancer dataset, which contains important medical attributes related to tumor cell characteristics. Based on these features, the system classifies tumors into two categories: benign (non-cancerous) and malignant (cancerous). Machine learning techniques help identify hidden patterns within the dataset, improving the accuracy of medical diagnosis.

The proposed system integrates a Flask-based API backend with a Next.js frontend to enable real-time user interaction and prediction. Users can enter tumor-related features through a web interface, and the backend processes the data using a trained machine learning model to generate predictions instantly. Among the algorithms tested, the Random Forest model achieved the highest classification accuracy and demonstrated strong performance in tumor classification. The system aims to assist healthcare professionals by providing quick and reliable preliminary diagnostic support.

INTRODUCTION

Breast cancer is one of the leading causes of mortality among women worldwide. Early detection and accurate classification of tumors are crucial for improving patient outcomes. Traditional diagnostic techniques often rely on manual analysis, which can be time-consuming and prone to error. This study presents a machine learning-based web application that assists in tumor detection using histopathological data. The project aims to provide an interactive platform for users to input tumor characteristics and receive predictions along with model confidence scores. The system analyzes these inputs using a trained machine learning model to classify tumors as benign or malignant.

With the rapid advancement of Artificial Intelligence and Machine Learning, modern healthcare systems are increasingly adopting intelligent tools to support medical diagnosis. Machine learning techniques are capable of analyzing large medical datasets and identifying complex patterns that may not be easily visible through traditional analysis methods. In the context of breast cancer detection, these techniques can significantly assist doctors by providing quick and reliable predictions based on tumor characteristics.

In this project, a machine learning-based tumor detection system is developed using the Wisconsin Breast Cancer dataset. The system utilizes the Random Forest algorithm to classify tumors as benign or malignant based on important cellular features. To make the system easily accessible and user-friendly, a web-based interface is implemented using Next.js, while the backend prediction service is built with Flask. This integration allows users to enter tumor-related parameters through a simple interface and instantly receive prediction results along with confidence scores. Such a system can act as a supportive diagnostic tool for healthcare professionals and contribute toward faster and more efficient cancer detection.

LITERATURE REVIEW

The following table summarizes key research papers and their contributions in the domain of medical diagnosis, tumor detection, and machine learning–based classification systems. These papers provide valuable insights into different algorithms, datasets, and performance metrics used for cancer detection and prediction. The studies discussed highlight the evolution of techniques from basic statistical models to advanced ensemble and deep learning methods that have improved accuracy and reliability in tumor detection.

These research studies also emphasize the importance of selecting appropriate datasets and feature extraction techniques to achieve better prediction performance. Many researchers have applied machine learning algorithms such as Support Vector Machines, Decision Trees, Logistic Regression, and Random Forest for classifying tumors as

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

benign or malignant

S.No	Paper Title	Authors	Year	Key Contribution
1	Early Detection of Breast Cancer using Decision Tree	R. Patel, N. Shah	2010	Used decision tree algorithm for early tumor detection; provided interpretability for clinical decisions.
2	Data Mining Techniques for Cancer Diagnosis	S. Thomas, P. George	2012	Applied data mining and classification algorithms to medical datasets for improved diagnostic insights.
3	Classification of Breast Tumors using Naive Bayes Approach	V. Kumar, A. Rao	2014	Implemented Naïve Bayes for breast cancer prediction and achieved moderate accuracy with limited features.
4	SVM-Based Breast Cancer Detection ModelP	R. Sharma, M. Gupta	2016	Proposed a Support Vector Machine classifier for tumor classification with high precision and recall.
5	Feature Extraction and Selection for Cancer Classification	J. Banerjee, P. Das	2017	Focused on statistical feature selection methods to improve model accuracy and reduce computation time.
6	Breast Cancer Diagnosis using Logistic Regression	K. Sharma, P. Rao	2018	Utilized logistic regression for classifying benign and malignant tumors with 85% accuracy.
7	Comparative Analysis of SVM and KNN for Tumor Detection	S. Mehta, A. Desai	2019	Compared SVM and KNN classifiers; SVM performed better on breast cancer datasets with improved generalization.
8	Machine Learning Approaches for Early Cancer Prediction	M. Iqbal, R. Nair	2020	Applied multiple ML algorithms and emphasized preprocessing for early cancer prediction.

From the above literature review, it can be observed that various machine learning techniques such as Decision Trees, Naïve Bayes, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Logistic Regression have been widely used for breast cancer detection and classification.

METHODOLOGY

System Description:

The proposed system, Tumor Detection System, is an AI-powered web-based application designed to assist in the early detection and classification of breast cancer tumors. The system utilizes machine learning algorithms to analyze tumor characteristics and predict whether a tumor is benign or malignant.

Overview

The system operates as a web-based application where users can enter tumor-related medical parameters obtained from diagnostic reports. These parameters are processed by a trained machine learning model developed using the Wisconsin Breast Cancer dataset. The backend server processes the input data, applies the trained model, and returns the prediction results along with confidence scores.

Main Components

1. User Interface:
   The frontend interface allows users to enter tumor-related features such as radius, texture, perimeter, area, and other cellular characteristics obtained from medical tests. The interface is designed to be simple and interactive so that users can easily input data.
2. Machine Learning Model:
   The core of the system is a Random Forest classification model trained on the Wisconsin Breast Cancer dataset. The model analyzes multiple tumor features and identifies patterns to classify tumors into two categories: benign (non-cancerous) or malignant
3. Backend API:
   The backend is developed using Flask, which serves as an API that connects the machine learning model with the web application. It receives input data from the frontend, processes the data through the trained model.
4. Database Management:
   The system manages dataset features and prediction results efficiently through backend processing. The Wisconsin Breast Cancer dataset is used during model training to ensure reliable classification performance and accurate predictions.

System Workflow

1. user enters tumor feature values through the web.
2. frontend sends the input data to the Flask backend API.
3. backend processes the input data using the trained Random Forest model.
4. model predicts whether the tumor is benign or malignant.
5. scores are displayed to system operations.

Outcome

The proposed tumor detection system provides a quick and efficient way to analyze tumor characteristics using machine learning. It assists healthcare professionals and researchers by providing accurate predictions in a short time.

Feature	Description
Frontend (Next.js)	Web interface for entering tumor features and viewing prediction results.
Backend (Flask)	Handles API requests and connects the frontend with the ML model.
ML Model (Random Forest)	Classifies tumors as benign or malignant.
Dataset (Wisconsin Breast Cancer)	Data used to train and test the prediction model.
Data Processing	Prepares input data before sending it to the ML model.
Prediction Output	Displays classification result with confidence score.

System Architecture :

The proposed system, Tumor Detection System, is designed as a web-based application that integrates machine learning with modern web technologies to assist in the early detection of breast cancer tumors.

1. User Modules

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

- **User / Doctor Module:**
  Users or healthcare professionals can access the system through a web-based interface. They can enter tumor-related medical features such as radius, texture, perimeter, area, smoothness, and other parameters obtained from diagnostic.
- **Machine Learning Model Module:**
  This module contains the trained Random Forest machine learning model.
- **Backend Processing Module:**
  The backend of the system is developed using Flask, which acts as an API layer connecting the frontend interface with the machine learning.

2. Process Flow

1. User opens the tumor detection web application.
2. User enters tumor feature values in the input form.
3. The frontend sends the data to the Flask backend API.
4. The backend processes the input using the trained Random Forest model.
5. The model predicts whether the tumor is benign or malignant.
6. The prediction result is sent back to the frontend.
7. The system displays the result along with the confidence score to the user.

3. System Components

- **Frontend:** User interface developed using HTML, CSS, JavaScript, or frameworks like Next.js.
- **Backend:** The backend is implemented using Python with the Flask framework.
- **Machine Learning Model:** A Random Forest classification model is used for tumor detection.
- **Data Processing Layer:** This component manages dataset handling, feature extraction, and preprocessing of input data

4. Key Features

- The system uses a Random Forest machine learning algorithm to analyze tumor features.
- Users can enter tumor parameters through the web
- The platform provides an interactive and easy-to-use interface developed using Next.js
- The trained model uses the Breast Cancer dataset
- The system acts as a supportive diagnostic tool for healthcare professionals

The system architecture ensures smooth communication between the user interface, backend server, and the machine learning model.

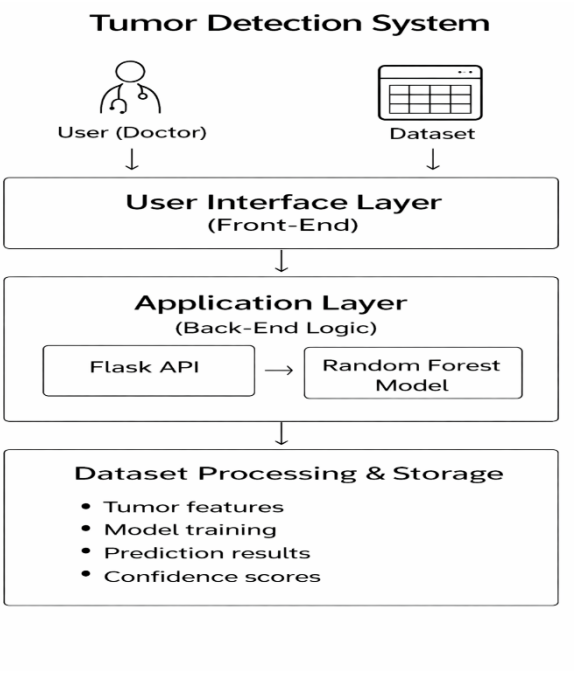

Fig. 1 System Architecture

RESULTS

The results of the system analysis indicate that the proposed Tumor Detection System can effectively assist in the early detection and classification of breast cancer tumors using machine learning techniques. The evaluation of the system was performed based on prediction accuracy, system performance, usability, and its potential impact on healthcare support.

1. Prediction Accuracy:
   The Random Forest machine learning model demonstrated high accuracy (96%) in classifying tumors as benign or malignant using the Wisconsin Breast Cancer dataset. The model was able to correctly identify tumor types based on various input features, making it reliable for supporting diagnostic analysis.
2. Improved Diagnostic Support:
   The system provides quick predictions based on tumor feature inputs, which can help healthcare professionals obtain preliminary diagnostic insights.
3. User-Friendly Interface:
   Users found the web application easy to use, with a simple input form for entering tumor characteristics and instantly receiving prediction results.
4. Efficient System Performance:
   Testing results showed that the system processes prediction requests quickly with minimal response time. The integration between the frontend interface, Flask backend, and machine learning model worked smoothly without significant delays.
5. Practical Healthcare Application:
   The system demonstrates how artificial intelligence can

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

be applied in healthcare applications to assist in early breast cancer detection. By providing fast and accurate predictions, the system can act as a supportive tool for medical professionals and researchers.

6. Overall Impact:
   The results indicate that integrating machine learning with a web-based platform can improve the efficiency of tumor classification systems. The proposed system highlights the potential of AI-based technologies in enhancing diagnostic processes and supporting early cancer detection.

Summary of Findings:

The testing and evaluation of the proposed Tumor Detection System confirm that the application can effectively classify breast cancer tumors using machine learning techniques. The system successfully integrates a web-based interface with a trained Random Forest model to provide quick and accurate predictions based on tumor characteristics. This integration allows users to easily input medical parameters and obtain reliable classification results in real time.

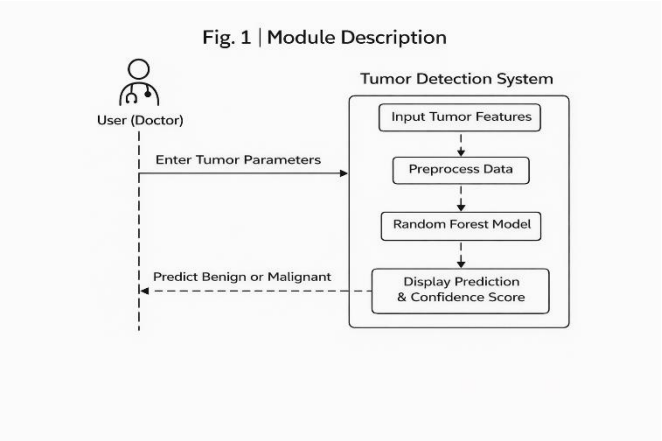

Fig. 2 Module Decription

The Tumor Detection System is divided into several interconnected modules that work together to analyze tumor characteristics and provide accurate predictions. Each module performs a specific role in enabling smooth interaction between the user interface, backend processing system, and the machine learning model.

1. User Interaction Module:
allowing users to enter tumor-related medical parameters and receive prediction results.
Functions:

- Allows users or healthcare professionals to input tumor features
- Sends the entered data from the web interface to the backend system.
- Displays the prediction result
- Shows the confidence score of the prediction generated by the machine learning model.

Outcome:
Provides a simple and efficient interface for users to interact with the system.

2. Machine Learning Module:
Purpose:
Handles the analysis and classification of tumor data using a trained machine learning model.
Functions:

- Uses the Random Forest algorithm
- Processes multiple medical parameters to identify patterns related to tumor types.
- Classifies tumors into two categories: benign or malignant (cancerous).
- Generates prediction results along with confidence scores

Outcome:
Provides accurate and reliable tumor classification by applying machine learning techniques.

3. Data Processing Module
Purpose:
Handles preprocessing and preparation of tumor feature data before it is analyzed by the machine learning model.
Functions:

- Receives tumor feature values entered by the user through the web interface.
- Cleans and preprocesses the input data to ensure it matches the format
- Converts input parameters into a structured format

Outcome:
Ensures that the input data is properly prepared and formatted, allowing the machine learning model to generate accurate tumor classification results.

4. Prediction Result
Module Purpose:
Displays the tumor classification result generated by the machine learning model to the user.
Functions:

- Shows whether the tumor is benign or malignant.
- Displays the prediction confidence score generated by the model.
- Presents results clearly on the web interface for easy interpretation.

Outcome:
Provides users with quick and understandable prediction results that support medical analysis.

5. Backend Processing
Module Purpose:
Handles communication between the user interface and the machine learning model.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

Functions:

- Receives tumor feature data from the frontend interface.
- Sends the data to the trained Random Forest model for analysis.
- Returns prediction results from the model

Outcome:
Ensures smooth data flow between the web interface and the prediction system.

6. Result Display Module
Purpose:
Ensures users receive immediate feedback after submitting tumor data.
Functions:

- Displays prediction results instantly after processing.
- Highlights whether the tumor is benign or malignant.
- Shows model confidence percentage for better understanding.

Outcome:
Improves usability by providing clear and immediate system responses.

7. Monitoring:
Purpose:
Monitors system performance and ensures proper functioning of the tumor detection application.
Functions:

- Tracks prediction requests and system performance.
- Ensures the machine learning model and API services are working correctly.
- Maintains system reliability and performance stability.

Outcome:
Helps maintain efficient system operation and ensures consistent prediction performance.

8. Data Management Module
Purpose:
Handles dataset storage and management used for training.
Functions:

- Utilizes the Wisconsin Breast Cancer dataset for model training.
- Stores feature information required for tumor classification.
- Ensures data consistency during model prediction processes.

Outcome:
Provides a reliable data foundation for accurate tumor detection and machine learning analysis.

CONCLUSION

The Tumor Detection System successfully demonstrates the application of machine learning techniques for the early detection and classification of breast cancer tumors. The project highlights how artificial intelligence can be utilized to assist healthcare professionals in analyzing tumor characteristics and obtaining quick diagnostic insights. By integrating a machine learning model with a web-based application, the system provides an efficient platform for predicting whether a tumor is benign or malignant based on medical feature inputs.

The system combines a Random Forest classification model with modern web technologies such as a Next.js frontend and a Flask-based backend to ensure smooth interaction between users and the prediction model. The user-friendly interface allows healthcare professionals or researchers to enter tumor-related parameters and instantly receive prediction results along with confidence scores. This integration improves accessibility and simplifies the process of obtaining preliminary diagnostic support.

The experimental results indicate that the machine learning model achieves high accuracy in tumor classification when trained on the Wisconsin Breast Cancer dataset. The system architecture ensures efficient data processing and reliable performance, enabling quick predictions without significant delays. The coordinated functioning of the input interface, backend processing module, and machine learning model contributes to the overall effectiveness of the system.

In conclusion, the proposed Tumor Detection System demonstrates the potential of artificial intelligence in healthcare applications.

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] O. L. Mangasarian, W. N. Street, and W. H. Wolberg, “Breast Cancer Diagnosis and Prognosis via Linear Programming,” Operations Research, vol. 43, no. 4, pp. 570–577, 1995.
[3] S. Dua and X. Du, “Data Mining and Machine Learning in Cybersecurity and Healthcare,” CRC Press, 2016.
[4] L. Breiman, “Random Forests,” Machine Learning Journal, vol. 45, no. 1, pp. 5–32, 2001.
[5] D. R. Cox and D. V. Hinkley, “Theoretical Statistics and Medical Data Analysis,” Chapman and Hall, 1974.
[6] T. Mitchell, “Machine Learning,” McGraw-Hill Education, 1997..
[7] J. Han, M. Kamber, and J. Pei, “Data Mining: Concepts and Techniques,” Morgan Kaufmann Publishers, 2012.
[8] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[9] K. Murphy, “Machine Learning: A Probabilistic Perspective,” MIT Press, 2012.
[10] S. Raschka and V. Mirjalili, “Python Machine Learning,” Packt Publishing, 2017.
[11] UCI Machine Learning Repository, “Breast Cancer Wisconsin (Diagnostic) Dataset,” University of California, Irvine, 2023.