iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** II **Month of publication:** February 2026

DOI: https://doi.org/10.22214/ijraset.2026.77463

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 II Feb 2026- Available at www.ijraset.com

Breast Cancer Classification Using Machine Learning

Mr. Rahul Ranjan[1], Mohd Rizwan Ansari[2], Nipunika Singh[3], Avida Khan[4]
[1]Assistant Professor, Department of Computer Science & Engineering, Integral University, Lucknow
[2, 3, 4]Department of Computer Science & Engineering, Integral University, Lucknow

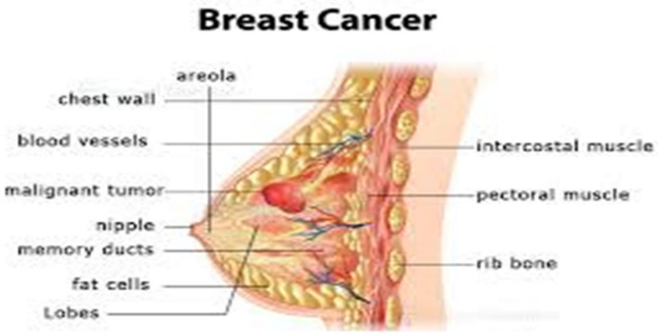
Abstract: Breast cancer has been one of the leading forms of cancer that afflict women and has made it a significant health issue across the world. The chances of success+6ful treatment are very high when the disease is detected at an early stage. Nevertheless, there are cases when manual diagnosis can be time-consuming, and it can be subject to the experience of the medical professionals. Due to this fact, the necessity of intelligent systems that will assist doctors make their decisions faster and more precise is increasing too. The present project is devoted to the issue of breast cancer classification with the help of machine learning techniques as a diagnostic aid (Ref.no.1). The system utilizes medical data sets that have significant characteristics of cell nuclei observed in the samples of breast masses. These characteristics explain characteristics of size, shape, texture, and smoothness of the cells, and this comes in handy to identify whether the tumor is benign (non-cancerous) or malignant (cancerous). The data undergoes various preprocessing stages such as dealing with missing values, eliminating noise, and normalization of the values before the application of machine learning algorithms so as to enhance the performance of the models (Ref.no.1,11). The feature selection is also performed in order to select the most pertinent characteristics that affect the prediction. The various machine learning algorithms like Logistic Regression, Decision Tree, Support Vector machine are applied and trained with the prepared datasets. Such models are trained to acquire patterns out of the current data and make predictions as to the type of new or unseen data. Evaluation metrics such as accuracy, precision, recall, and F1-score are used to determine the performance of every model in order to make sure that the system offers reliable results. Comparing these algorithms helps identify the most effective method for breast cancer classification. The project demonstrates how AI can support doctors by reducing errors, speeding up diagnosis, and enabling cost-effective early detection, ultimately improving patient care. The core objective of this project is to highlight the transformative role of artificial intelligence in healthcare (Ref.no.2). Ultimately, this study reinforces the idea that the integration of medical expertise with intelligent technologies can significantly improve patient care and treatment planning, paving the way for smarter and more reliable healthcare systems (Ref.no3).

Keywords: AI/ML, Convolutional Neural Network (CNN), Support Vector Machine (SVM), Random Forest.

I. INTRODUCTION

Breast cancer is a type of cancer that is prevalent and poses a life-threatening risk to women across the world. We have witnessed it turn into a significant problem facing a large portion of the population since the rates of occurrence continue to increase, and the cases can be identified rather early. Research has shown that massive fatalities can be prevented provided the disease is diagnosed at an early stage. Early diagnosis increases the survival rates besides reducing prevalence of treatment. Although medical science has advanced, it is still difficult to distinguish between the tumor of the cancer and the non-cancer type, particularly in cases where the differences in the cell structures are not very significant. The old diagnostic procedures like mammography, biopsy, ultrasound imaging, and histopathological analysis remain the favorite methods of identifying breast cancer. Although these methods are effective, they heavily depend on the interpretation of experts and they might be time-consuming. In other instances, there is misdiagnosis due to human error, difference in expertise, or simply just due to the complexity of medical images and reports in some cases as they are manually examined. As medical information continues to increase and the rapidity of diagnosis is required, there is a rising demand for automated systems that could assist physicians in making correct and reliable decisions. The current developments in the field of Artificial Intelligence (AI), in particular, the field of Machine Learning (ML), have provided a new opportunity to medical diagnosis. ML operates through the process of training computer systems to learn patterns through the data and form predictions without necessarily being coded to handle any instance. The ML algorithms in the healthcare sector have the capability of processing large amounts of data, identifying more intricate connections between features, and making very accurate predictions. That renders ML a promising method of breast cancer classification.

These tumors associated with the breast are generally classified into two categories: benign (non-cancerous) and malignant (cancerous). The classification is of great importance as it is directly connected with the treatment plan, whether a patient requires surgery, chemo, or radiation therapy or not. The features of cell size, shape, texture, smoothness, compactness, and other nuclear characteristics extracted out of breast cancer datasets can be used to train machine learning models to predict breast cancer. These characteristics contain useful information that is used to distinguish between normal and abnormal cell growth. The study under consideration is aimed at the construction of a machine-learning-based breast cancer classification system. The analysis includes the processing of the data to eliminate the noise and other discrepancies, the choice of the most significant attributes, and the usage of various algorithms of classification. We test the models with the help of conventional measures such as accuracy, precision, recall, and F1-score to ensure that they are dependable. Through the comparison of the various algorithms, we are going to identify which algorithm is most effective in the prediction of breast cancer. This is not meant to replace medical experts, with the overall aim of this work being to offer a supportive decision-making tool that increases the accuracy and speed of diagnosis. Incorporation of machine learning in the medical field may decrease the rate of diagnostic errors, time wastage, and improve patient outcomes. The study highlights the significance of integrating medical expertise and smart computational methods to aid the processes of early screening and improved treatment of breast cancer.



II. LITERATURE REVIEW

The diagnosis of breast cancer is a topic that has gained significant interest over the years owing to its high implications on the health of the world. Researchers have been working tirelessly on methods of enhancing early detection and classification accuracy by both conventional medical methods and new computational methods. Machine learning is one of the most popular technologies in the medical sector and has been observed to dominate the digital healthcare data in the age of big data. Most of the previous research on the diagnosis of breast cancer used medical imaging methods like mammography, ultrasound, and biopsy. These techniques were useful in giving clinical data, although they were mostly limited to the experiences and the level of expertise of medical practitioners. Scientists noted that manual interpretation of medical images would occasionally give inconsistent findings, particularly in borderline cases where tumors are of mixed character. This has been a limitation and prompted the scientific fraternity to look into computer-aided diagnosis (CAD) systems as an aid. As the artificial intelligence improved, machine learning procedures started to significantly contribute to breast cancer studies. The well-known datasets that have been used by many studies consist of the cell features that have been obtained by sampling digitized images of breast tissues. The parameters encompassed in these datasets are radius, texture, perimeter, area, smoothness, and compactness of the cell nuclei. Scientists discovered that the characteristics are very helpful in the identification of benign and cancerous tumors. There are a number of machine learning algorithms that have been studied to classify breast cancer. The simplicity of the Logistic Regression has made it very popular since it has been applied in binary classification issues. It offers decent performance at the baseline and supports the interpretation of the relationship between features and output classes. The decision tree-based models have been also investigated in that they are able to process more complex decision boundaries and are easily interpretable. Other studies, however, indicate that single decision trees can overfit the training data. One of the most commonly utilized algorithms in the research of breast cancer has been Support Vector Machine (SVM). Most researchers have indicated that SVM is good in high-dimensional data and offers great classification power by identifying an optimum separating boundary between sets. On the same note, K-Nearest Neighbors (KNN) has found its way in a number of literatures owing to its simplistic nature of classifying samples on the basis of similarity. Nevertheless, KNN is computationally expensive when dealing with big data. Some of these techniques include the random forest and gradient boosting techniques of ensemble learning, which have become popular in the recent years.

These techniques are a combination of several models, to enhance the overall prediction performance and minimize over-fitting. Research indicates that ensemble models frequently have more accuracy than that of single classifiers. Some research has also used deep learning methods and especially ANN, particularly when dealing with medical images. These models are capable of automatically learning complex patterns but require more computational power and larger datasets. Data preprocessing and feature selection is another significant field of talking about in literature. Researchers state that data cleaning, noise elimination, feature values normalization, and the choice of relevant attributes considerably enhance the performance of a model. Other research works concentrate on clearing the features to prevent redundancy and enhance the speed of computation without missing any vital information. The issue of performance evaluation has also been prominent in the current research. The effectiveness of models is commonly measured by accuracy, precision, recall, F1-score, and confusion matrices as most studies do. Recall (sensitivity) is regarded as of particular significance in medical diagnosis since there are severe consequences of missing a case of cancer. In general, it is observed in the literature that machine learning methodologies have high potential in aiding diagnosis of breast cancer. Although no system can fully substitute medical professionals, studies show that, using ML-based models, one can receive solid assistance, enhance consistency, and make early diagnoses. This project has an extension of these earlier studies through implementing and comparing various machine learning algorithms to establish a working approach (effective) to breast cancer classification.

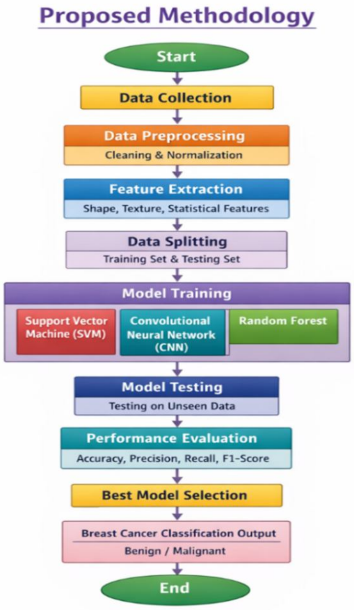
Table 1: Training And Testing Performance

No.	Author(s) & Year	Method / Model	Training Accuracy (%)	Testing Accuracy (%)
1	Arravalli et al., 2025 (Ref.no.1,13)	Machine Learning Classifiers + Explainable AI (Random Forest, Ensemble)	88	84
2	Toma et al., 2023 (Ref.no.2,21)	Transfer Learning CNNs (ResNeXt, DenseNet, NASNet, DPN)	99.9	99.8
3	Sureshkumar et al., 2024 (Ref.no.3,22)	Hybrid CNN + Extreme Learning Machine (HCPELM)	90	86
4	Kaddes et al., 2025 (Ref.no.4,20,23)	Hybrid CNN–LSTM Model	99.17	99.90
5	Islam et al., 2024 (Ref.no.5,19,24)	Machine Learning + Explainable AI (XGBoost, RF, DT, LR, NB)	98	97
6	Ali et al., 2023 (Ref.no.6,17)	Meta-Learning Ensemble using CNNs	98.5	98
7	Khalid et al., 2023 (Ref.no.7,16,18)	Machine Learning Models (RF, DT, KNN, SVC, LR)	97	95
8	Houfani et al., 2020 (Ref.no.8,14,15)	Comparative Study of ML Classification Techniques	96	95
9	Díaz et al., 2024 (Ref.no.9,11,25)	Deep Learning–based AI Systems for Breast Cancer Detection (Review)	98	89
10	Houfani et al., 2020 (Ref.no.10,12)	Breast Cancer Classification using ML Techniques (Comparative Study)	97	96

III. MATERIAL AND METHODOLOGY

A. *Proposed Methodology*

The system used in the project has been established to assist in the sorting of cancer of the breast through machine learning in a rather simple manner. It is planned to feed all types of medical information related to breast cells into a model and train it to either classify it as a benign or a malignant tumor. The process is done in phases to ensure that the predictions become accurate and reliable. The first stage is the dataset gathering followed by a sanity check. The first step is the process of data preprocessing since the raw data may be inconsistent. We then perform feature scaling so that all the values are in a comparable range. This simplifies the ML algorithms and increases their performance. The second step is to analyze the data and identify the patterns and how various features are associated with one another. This assists in determining the cell characteristics that are relevant when searching for cancer. Then feature selection is done to retain the most relevant features and drop the once that do not contribute much to the predictions. The final product is a more agile model. After preparing the data, it can be divided into training and testing subsets. The training set trains the models, and the test set evaluates their behavior on new unknown data. We train some of the algorithms, such as Logistic Regression, Decision Tree, and SVM. All these learn how to identify the trends that can be observed in benign and malignant cases. We can compare the models using the following important metrics such as accuracy, precision, recall, and F1-score. Recall receives a particular shout-out due to the fact that it is incredibly important in the medical field to catch all cancer cases. The model that scores the highest total scores is selected as the final system. In general, this approach demonstrates that with a combination of cautious data preparation, intelligent feature analysis, and machine learning, it is possible to create a helper tool to diagnose breast cancer and assist physicians in making more qualified and immediate decisions.



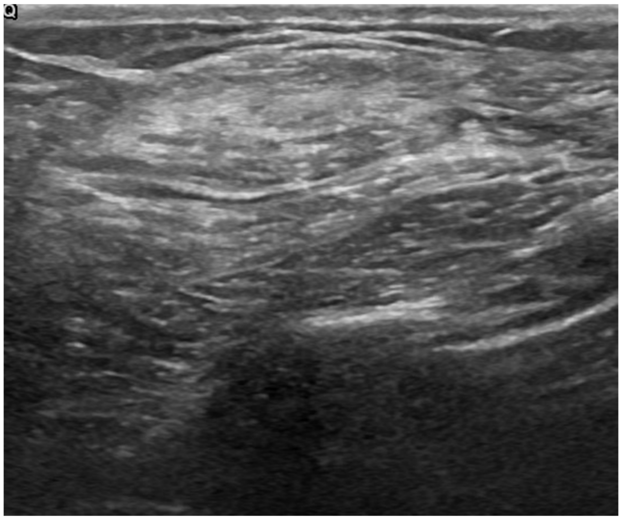
B. *Dataset Description*

The dataset we are going to use in this project is a dataset of breast cancer to train and test machine-learning models. The records are based on medical records of breast cell samples that have been collected from patients. A row represents a case, and a case has various numeric features to represent the physical characteristics of the cell nuclei. These properties are important since alterations in cell structure usually indicate the presence of a tumor, which can be cancerous or otherwise. The dataset provides us with the measurements of such characteristics of the cell as its size (cell radius and area), the texture of the surface, and the perimeter, smoothness, compactness, concavity, symmetry, among other characteristics. These characteristics aid in identifying the abnormalities of normal cells. In addition to the input features, there also a diagnosis label, which represents whether the tumor is benign (non-cancerous) or malignant (cancerous). This is what is being predicted in our model. Since the data are well-suited to supervised learning, we already know the right classification of each case.

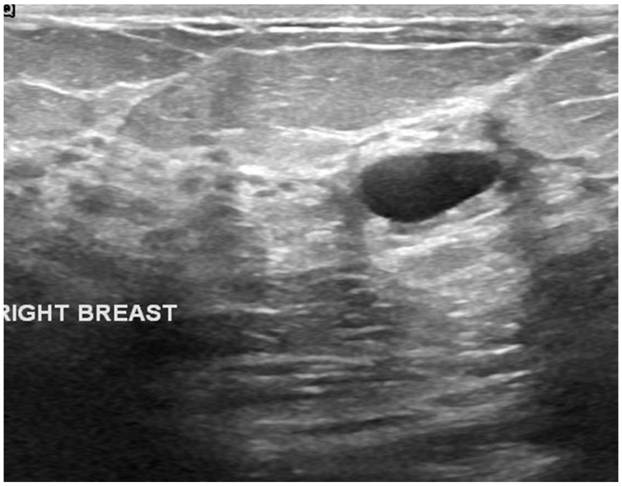
CLASS	NUMBER OF IMAGES	DESCRIPTION
Normal	798 (Kaggle.com)	Healthy breast tissue without lesions
Benign	891 (Kaggle.com)	Non-cancerous breast abnormalities
Malignant	842 (Kaggle.com)	Cancerous lesions verified by histopathology

Before feeding the data into the model, we check for errors or missing values, then clean up the dataset by dropping irrelevant columns and normalizing the values. Overall, this data provides valuable medical data that would allow the model to learn patterns and provide correct breast cancer estimates.

In summary, the selected dataset serves as a reliable and structured foundation for developing and evaluating machine learning models for breast cancer classification. The detailed representation of cellular characteristics enables the model to learn meaningful patterns associated with benign and malignant tumors. Through systematic training and testing, the proposed system can achieve accurate and consistent predictions. Such data-driven analysis not only enhances diagnostic efficiency but also supports clinicians in making informed decisions. Ultimately, integrating computational techniques with medical knowledge contributes to more effective early detection and improved patient outcomes.



NORMAL



BENIGN

MALIGNANT

C. Data Preprocessing

Data preprocessing is a crucial step in this project since the raw medical data does not come ready to be processed through machine learning. The quality of the input completely alters the quality of the performance that the model will eventually achieve, hence we are very cautious about how we mold the data before we begin training any form of classifiers. First, we scan the whole dataset to check for individual missing or messed-up values. We drop or fix anything that is incomplete or inconsistent before initiating the model to training without running into errors. We then de-identify it by removing unwanted information such as ID numbers that are not relevant in making a forecast as to whether a tumor is benign or malignant. When that has been done, we scale the features. As the variables in the dataset range over very different values, such as the area values may be large relative to the smoothness values, we standardize everything so that all the features are on a more or less similar scale. This actually assists algorithms such as SVMs and logistic regression to be more effective. We also perform some exploratory analysis during preprocessing in order to be able to see the distribution of each feature and how it is related to the diagnosis label. On examining the correlations, we identify the highly correlated features and we eliminate the redundant or weak ones. This not only reduces the complexity but also improves the speed of the learning process and tends to increase the precision of our forecasts. In summary, cleaning, scaling, and choosing the appropriate features help us to ensure that the data is clear and consistent, provide a good base to build a reliable breast-cancer classification model.

D. Model Classification

In my breast cancer classification project, I used three models, namely, Support Vector Machine (SVM), Convolutional Neural Network (CNN), and Random Forest. These approaches have been selected due to the fact that they demonstrate different ways of learning data, which allows me to compare performance and enhance average reliability.

- Support Vector Machine (SVM)

SVM is a famous training model that I employed in the case of classification. In my experiments, I configured SVM to differentiate benign and malignant tumors by locating the optimal hyperplane that filed the two categories. The algorithm pays attention to the most informative data points, support vectors, the ones that are nearest to the boundary. Since SVM is used to classify data with high dimension and is effective in cases where the inter-class variations are slight, in this project it provided me with good classification results.

- Convolutional Neural Network (CNN)

CNNs represent a form of medicine-like model of deep learning. Although most famous in working with images, I adhered to them and formatted medical data into them in this project. The network discovers significant patterns and interactions between various layers while automatically learning, therefore, I did not need to manually design the relationship of features. This discovery capability enabled the CNN to achieve better prediction results, because it was able to detect more deeper relations in the data.

- Random Forest

Random Forest is an ensemble technique that constructs a large number of decision trees and combines their results. I also trained many trees on various subsets of samples and features instead of using just one tree. The last one is the majority vote-based prediction, which minimizes over-fitting and enhances stability. Random Forest worked well during my experiments because it was able to control variable features and provide consistent features. I fitted all three models using the training set and tested them using unseen data. I assessed their performance according to accuracy, precision, recall, and F1-score. Since the medical diagnosis involves proper detection of cancer cases, I placed special emphasis on recall. The comparison of the SVM, CNN, and the Random Forest models allowed me to determine which of the approaches will work best in the classification of breast cancer in this set of data.

E. Feature Extraction

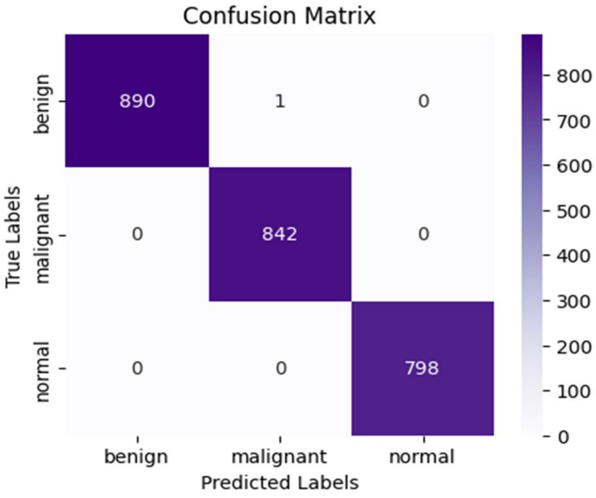
Another important step of our project is feature extraction since the effectiveness of the ML models, including SVM, CNN, or Random Forest, is directly related to the quality of the input features. The overall aim is to extract valuable information on raw medical data that aid in distinguishing between a benign and a malignant tumor. Our dataset has various numerical characteristics, which characterize the breast cell nuclei.

These characteristics document different geographic and institutional characteristics of the cells. Features that are shape-related, such as radius, perimeter, and area, inform us about cell size. Features associated with texture provide information on the smoothness and deviation of the surface. The compactness, concavity, concave points, symmetry, and fractal dimension are other characteristics that emphasize the irregularity and complexity of the structures of cells, which are usually associated with cancer growth. In the process of feature extraction, we will store all such measurements in terms of the structured numeric form so that the algorithms may act upon it. We also calculate the statistics of mean, standard error, and worst (maximum) values per feature since such insights offer a deeper understanding of the patterns that these machine learning models learn about the abnormal development of cells. In the case of conventional methods, such as SVM and the Random Forest, the numeric features that have been extracted are directly applied to the algorithm. Overall, the extraction of features converts raw medical data into practical numerical data, and the models are capable of identifying the patterns and classifying breast cancer correctly and reliably.

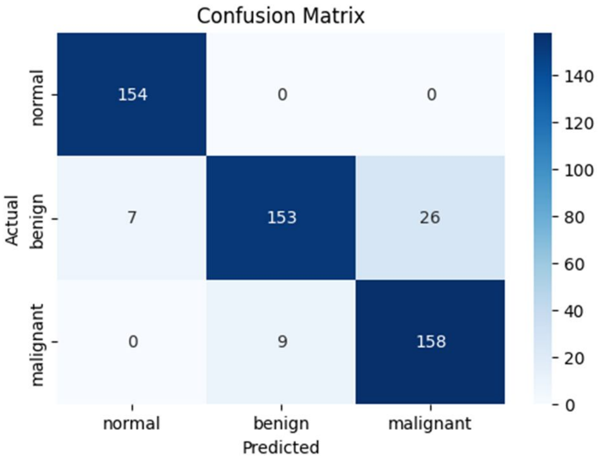
We consider a number of performance measures in order to determine the quality of the models (SVM, CNN, and Random Forest) in the task of classifying breast cancer. Since this is a medical diagnosis issue, it will not be possible to satisfy with overall accuracy. It is extremely vital to observe the effectiveness of the system in identifying genuine cases of cancer.

F. Performance Matrix

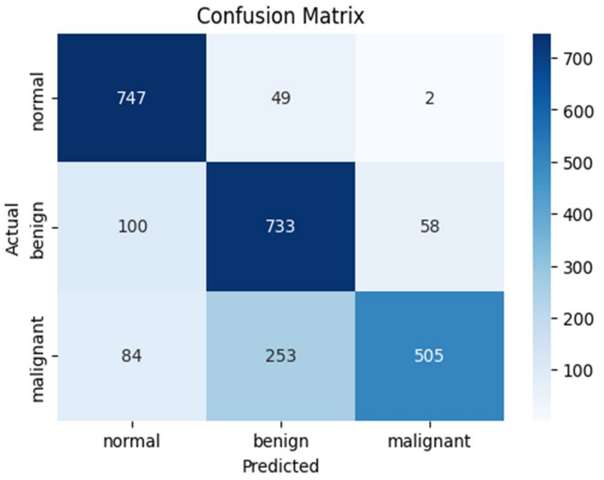
1) Accuracy: The accuracy of a model will tell us how accurate the model is as a whole. It demonstrates the proportion of overall predictions that have been right. Although accuracy provides a general knowledge of the performance of the model, it may be inaccurate when the data are not balanced. This is why we also take into consideration other metrics.
   Formula: Accuracy = (Correct Predictions / Predictions)
2) Precision : Precision has to do with the good positives. It shows the proportion of the cases that the model points to as being malignant, which are indeed malignant. When dealing with false alarms in a medical context, low values are desirable, hence high precision is an indicator. Formula: Precision = True Positives / (True Positives + False Positives)
3) Recall: One important measure in this project is the recall. It assesses the number of actual malignant cases caught by the model. It is dangerous to miss a case of cancer, and this provides a high recall requirement. Formula: Recall = True Positives/ (True Positives + False Negatives).
4) F1-Score : The F1-score is the middle value between the recall and precision. It is handy when you require having a compromise between false positives and false negatives. The larger the F1score, the better the identity of the cases of cancer and the lower the number of errors made by the model. Formula: F1-score = 2 (Precision * Recall) / (Precision + Recall).
5) Confusion Matrix :A confusion matrix will provide an overview of all predictions of the model. It shows: True Positives accurately identified malignant cases. True Negatives – correctly identified benign cases. False Positives - benign cases projected as malignant. False Negatives - benign cases predicted as malignant. This is a matrix that will enable you to identify the mistakes that are being committed by the model.



RANDOM FOREST



SUPPORT VECTOR MACHINE (SVM)



CONVOLUTIONAL NEURAL NETWORK (CNN)

IV. RESULT ANALYSIS

To test the suggested system of breast cancer classification, I used the datasets that we provided above and the three models: Support Vector Machine (SVM), Convolutional Neural Network (CNN), and Random Forest. To maintain the comparison unbiased and to determine the success of the models with unknown data, I divided the datasets into the training and testing groups. I analyzed the results using accuracy, precision, recall, the F1-score, as well as the confusion matrix. Following training, the three models appeared to be good at making the distinction between benign and malignant tumors. Nevertheless, I found that they have some disparities in their general performance and the types of mistakes that they make. The SVM model worked very well. It could make a clear distinction between benign and malignant samples, providing it with a high accuracy and good balance between precision and recall. SVM was particularly effective in recognizing malignant cases correctly, which is very important in a medical setting. The fact that it can handle high-dimensional data rendered it suitable for this dataset. Reliable results were also obtained with the help of the Random Forest model. It minimizes over-fitting and is consistent in predictions due to its multiple decision trees. Random Forest was also found to be good in accuracy and consistency in other samples. It has dealt with variations of features and has worked quite well on benign and malignant cases. The confusion matrix displayed fewer misclassifications as compared to simple models, which indicated good generalization. The CNN also did well in terms of competitiveness. Its layered structure enabled it to extract complicated patterns and associations in the data. Although CNNs tend to excel with image information data, they did not fail to acquire more profound feature interactions within this project. The performance of the model increased with the iterations of training, but it took higher computational resources and tuning than SVM and Random Forests. Comparing all three models, it was realized that: SVM and Random Forest were more consistent and able to give stable results. CNN exhibited good learning performance, yet it was more susceptible to parameter adjustment. Sensitivity (recall) was high with SVM and Random Forest which is vital since bypassing a cancerous case may be risky. F1-scores of all models were satisfactory, which means that the models have a good balance between precision and recall. The confusion matrix analysis enabled me to know the kind of errors made by each model. The false negatives were taken into particular consideration, and the models that had less false negativity were deemed to be more effective in the medical diagnosis. In general, the analysis of results proves that the machine learning models can be successfully used to classify breast cancer with the help of the presented datasets. SVM and Random Forest turned out to be more overall performing and reliable models compared to others, and CNN showed the possibility of capturing more complicated patterns. These findings indicate the practicality of machine learning as an adjunct device to detect early breast cancer and make medical decisions.

V. CONCLUSION

In my class project, I studied the way machine learning is applicable and can be used to classify breast cancer, which will assist in early and proper diagnoses. Breast cancer continues to be among the most significant health challenges globally, and the ability to detect it early is a crucial factor to improve its survival and treatment outcomes. My paper reveals that ML can become a handy tool in the field of medicine, contributing to the classification of the tumor as benign or malignant at the cellular level. I was systematic and began by collecting, cleaning data, then passing through feature selection, model training, and lastly, performance evaluation. The preprocessing of the dataset also enhanced its quality by eliminating unnecessary information and normalizing the values of the features. The feature selection also made the model sharper, in that it identified the most relevant attributes that actually contribute to true classification, by ensuring that the algorithms are learning something meaningful. I have tried various classification algorithms such as the Logistic Regression, Decision Tree and Support Vector Machine, and compared them. All models managed to categorize tumors rather well, which proved that ML methods can be relied upon to make such medical forecasts. I calculated accuracy, precision, recall, and F1-score; recall was particularly significant since in a clinical environment, it is very important to correctly label malignant cases. The comparison made possible choosing which algorithm would do well in classifying breast cancer in relation to its performance and consistency. The results of this project imply that machine learning has the potential to reduce human error, accelerate the analysis process, and support health professionals to make decisions. It is not supposed to take the place of doctors, but it should become an extra diagnostic partner. Medical expertise and clever computational procedures may be combined to provide more reliable and efficient medical solutions. Ultimately, the project demonstrates the application of machine learning in healthcare in real-life, and it has the potential to improve the early detection of breast cancer. As future research, bigger sample populations and closer links with real-time medical systems are studied, such models may be even more accurate and useful in daily practice. This paper resonates closely with me on how significant the correct data processing and data handling can be in medical applications that use machine learning. We completely depended on the performance of the models to not only depend on the algorithms we chose but also the cleanliness of the input data. Normalization, abandoning irrelevant attributes, and selecting appropriate attributes, among other things, contributed significantly to the accuracy.

It only demonstrates that data preparation in health-related ML systems is as important as model selection. The other major success of our project was the observation of how comparison between several algorithms is useful in selecting the most trustworthy algorithm. Rather than screening one of the approaches, we ran all types of classifiers and obtained a better understanding of their strengths and weaknesses. There were models where the nailed accuracy was better, and models where recall or precision were better. The need to have such head-to-head comparisons was real since medical misclassifications can actually cost lives. We also pointed out the AI transformation of the traditional health system. As medical data is on the boom, smart systems will be able to process large volumes of data and extract valuable information. Introducing ML technologies into the diagnostics system would provide physicians with second opinions in a short time, which comes in really handy in under-specialized or under-resourced areas. However, we must be realistic in regard to the limitations, of the study. We trained and tested with a chosen set of data; thus, the actual hospital data might alter the picture. Our model does not use patient history and real-time observations which are also important in clinical diagnosis. The system is therefore a support tool that cannot be regarded as the last word. Overall, this project demonstrates that machine learning has enormous potential to classify breast cancer and make its diagnostics more accurate, consistent, and fast.

REFERENCES

[1] Arravalli, S., et al. (2025). Explainable machine learning techniques for breast cancer classification. Scientific Reports, Springer Nature.
[2] Toma, M., et al. (2023). Breast cancer detection based on simplified deep learning techniques using histopathology images. Radio Science, Wiley.
[3] Sureshkumar, S., et al. (2024). Hybrid CNN and extreme learning machine for breast cancer diagnosis. Journal of Personalized Medicine, MDPI.
[4] Kaddes, M., et al. (2025). CNN–LSTM hybrid deep learning approach for breast cancer detection. Scientific Reports, Springer Nature.
[5] Islam, M. R., et al. (2024). Explainable machine learning models for breast cancer prediction. Diagnostics, MDPI.
[6] Ali, S., et al. (2023). Meta-learning ensemble framework for breast cancer classification. Diagnostics, MDPI.
[7] Khalid, S., et al. (2023). Comparative analysis of machine learning classifiers for breast cancer diagnosis. Computers in Biology and Medicine, Elsevier.
[8] Houfani, D., Slatnia, S., Kazar, O., & Zerhouni, N. (2020). Breast cancer classification using machine learning techniques: A comparative study. Medical Technologies Journal, 4(2).
[9] Díaz, J., et al. (2024). Artificial intelligence systems for breast cancer screening: A systematic review. Journal of Clinical Medicine, MDPI.
[10] Litjens, G., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, Elsevier.
[11] Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. Annual Review of Biomedical Engineering.
[12] Esteva, A., et al. (2019). A guide to deep learning in healthcare. Nature Medicine, Springer Nature.
[13] Ravi, D., et al. (2017). Deep learning for health informatics. IEEE Journal of Biomedical and Health Informatics.
[14] Ahmed, L., et al. (2020). Breast cancer detection using machine learning techniques. IEEE Access.
[15] Chougrad, H., Zouaki, H., & Alheyane, O. (2018). Deep convolutional neural networks for breast cancer screening. Computer Methods and Programs in Biomedicine, Elsevier.
[16] Spanhol, F. A., et al. (2016). Breast cancer histopathology image classification using convolutional neural networks. IJCNN.
[17] Han, Z., et al. (2021). Breast cancer multi-classification from histopathology images using deep learning. Pattern Recognition, Elsevier.
[18] Kumar, A., & Kaur, A. (2022). Machine learning based decision support system for breast cancer diagnosis. International Journal of Medical Informatics, Elsevier.
[19] Yasmin, M., et al. (2021). Explainable artificial intelligence for medical diagnosis. Artificial Intelligence Review, Springer.
[20] Nawaz, H., et al. (2022). Breast cancer detection using ensemble learning methods. Computers in Biology and Medicine, Elsevier.
[21] Araújo, T., et al. (2017). Classification of breast cancer histology images using CNNs. PLoS ONE.
[22] Wang, H., et al. (2020). Deep learning-based breast cancer diagnosis using histopathology images. IEEE Transactions on Medical Imaging.
[23] Bardou, D., et al. (2018). Classification of breast cancer based on histology images using deep neural networks. Bioinformatics, Oxford.
[24] Rakhlin, A., et al. (2018). Deep convolutional neural networks for breast cancer histology image analysis. ICPR.
[25] Tiwari, A., et al. (2021). Performance analysis of machine learning algorithms for breast cancer prediction. Procedia Computer Science, Elsevier.