www.ijecs.in
International Journal Of Engineering And Computer Science
Volume 13 Issue 03 March 2024, Page No. 26039-26048
ISSN: 2319-7242 DOI: 10.18535/ijecs/v13i03.4794

Breast Cancer Detection Using Machine Learning

Satakshi Singh, Dr. Kavitha R

MCA (Department Of CS&IT)

JAIN(Deemed-To-Be-University) Bangalore, India

Department Of CS&IT

JAIN(Deemed-To-Be-University) Bangalore, India

Abstract

In today's Breast cancer detection becomes more accurate through the use of machine learning algorithms, which examine complex patterns found in patient data. The early diagnosis that these models provide could lead to better patient outcomes and prompt actions. The goal of this study paper is to increase the accuracy and efficiency of diagnostic processes by presenting a thorough investigation of the use of machine learning approaches for breast cancer diagnosis. We investigate how well Support Vector Machines (SVM), Random Forest, and Neural Networks differentiate between benign and malignant cases using a variety of datasets. To improve the discriminatory strength of our models, we also incorporate sophisticated feature extraction methods like the Wavelet Transform and Gabor Filter. The outcomes of the experiment show the promise of various machine learning techniques, with SVM showing a noteworthy accuracy rate. In addition, the research adds to our understanding of decision-making processes by offering insights into how interpretable the models are. Our research highlights the importance of using machine learning to detect breast cancer, providing a viable path forward for improving early diagnosis and treatment planning that follows.

Keywords—Breast cancer, Machine Learning, Accuracy, Efficiency, Support Vector Machines (SVM), Random Forest, Neural Networks, Feature Extraction, Wavelet Transform, Gabor Filter, Interpretability, Early Diagnosis, Treatment Planning.

1. Introduction

Breast cancer is a widespread health issue that primarily affects women worldwide. Because of its ubiquity, preventative actions are essential, and effective treatment outcomes are known to depend on early detection. Machine learning techniques are emerging as a promising tool in the quest to improve diagnostic capabilities, providing previously unheard-of potential for timely and accurate breast cancer classification.

Breast cancer prevalence: One of the most common types of cancer, breast cancer has a major effect on women's health all over the world. The prevalence of this illness emphasizes how urgent it is to develop practical methods to lessen its effects. The introduction emphasizes the need for comprehensive and approachable solutions to address the global health challenge of breast cancer by drawing attention to its pervasive nature.

Importance of Early Detection: Without a doubt, early detection is essential to the field of breast cancer treatment. The importance of detecting cancers in their early stages, when intervention techniques are most successful, is explained in this section. More treatment choices become accessible the sooner breast cancer is identified, which in turn improves the quality of life and survival rates for those who are affected.

The Potential of Machine Learning Techniques:

In the face of the necessity of early detection, machine learning techniques provide promise.

The transformational potential of machine learning in the context of breast cancer diagnosis is introduced in this paragraph. Machine learning techniques can examine large information, identify complex patterns, and help with more precise and nuanced breast cancer classification by utilizing computational algorithms.

Overview of Machine Learning Techniques:
A preview of the variety of machine learning approaches will be given in the following sections' introduction. The emphasis is on methods like Wavelet Transform, Gabor Filters, Random Forest classification, Support Vector Machines (SVM), and Back Propagation neural networks. Setting the stage for a thorough examination of various approaches, the section highlights how each one has the potential to advance the area of breast cancer categorization. To sum up, the introduction skillfully traverses the worldwide incidence of breast cancer, emphasizes the criticality of early detection, and presents the exciting field of machine learning techniques. In the study paper's later portions, a thorough examination of the use of these cutting-edge methods is laid out by this tactful framing.

2. RELATED Work

The paper authored by Sri Hari Nallamala, Siva Kumar Pathuri, and Dr. Suvarna Vani Koneru, published in the International Journal of Engineering & Technology (IJET) in 2018, provides a literature survey on the data mining approach for effectively handling cancer treatment. The focus is on exploring the applications and methodologies of data mining in cancer treatment [1]. The paper emphasizes the need for a computerized breast cancer diagnosis system to reduce diagnostic time and mortality rates. It surveys the application of machine learning algorithms, including neural networks, SVM, and RVM, revealing promising results for accurate breast cancer detection. Ongoing research aims to further enhance diagnostic accuracy in the field [2]. The paper by Guyon, Weston, and Barnhill, published in 2002 in "Machine Learning," focuses on gene selection for cancer classification using Support Vector Machines (SVM). It explores the application of SVM in identifying relevant genes for effective cancer diagnosis, contributing to the field of machine learning in bioinformatics. [3]. The paper by A. Bharathi and A.M. Natarajan, published in the "Journal of Computer Science" in 2011, focuses on cancer classification using Support Vector Machines (SVM) and Relevance Vector Machines (RVM). The approach is based on the analysis of variance features, presenting a methodology for effective cancer categorization through machine learning techniques [4]. The work by T. Mu, published in 2005 under "Medical Applications of Signal Processing," addresses the detection of breast cancer using v-SVM (variant Support Vector Machine) and RBF (Radial Basis Function) networks. The study explores the application of these machine-learning techniques for effective breast cancer detection within the context of signal processing in medical applications [5].
The paper by T. M. Khan and S. A. Jacob, published in the "Journal of Pharmacy Practice and Research" in 2017, provides a brief review of complementary and alternative medicine (CAM) use among Malaysian women with breast cancer. The study explores the prevalence and patterns of CAM utilization in the context of breast cancer treatment and management [6]. The paper by M. R. Al-Hadidi, A. Alarabeyyat, and M. Alhanahnah, presented at the 2016 9th International Conference on Developments in Systems Engineering (DeSE) in Liverpool, focuses on breast cancer detection using the K-Nearest Neighbor (KNN) machine learning algorithm. The study explores the application of KNN for effective breast cancer diagnosis, contributing to developments in systems engineering [7]. Another paper by Sri Hari Nallamala, Dr. Pragnyaban Mishra, and Dr. Suvarna Vani Koneru, published in the International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE) in 2019, discusses qualitative metrics related to breast cancer diagnosis using Neuro Fuzzy Inference Systems. The study delves into the use of advanced computational methods for improving the accuracy of breast cancer diagnosis [8].

3. **Proposed Methodology**

This section, which carefully outlines the techniques and steps performed in the investigation, forms the foundation of the study. The ensuing thorough explanation offers a thorough understanding of the dataset selection, preprocessing techniques, and the sophisticated application of every machine learning method used in the research, including Random Forest, Wavelet Transform, Gabor Filters, Support Vector Machines (SVM), Relevance Vector Machines (RVM), Back Propagation neural networks, and Random Forest.

Dataset Selection: The dataset used in these experiments is sourced from Kaggle, sized approximately 50KB. Focused on breast cancer diagnosis, it comprises 32 columns, with 'diagnosis' indicating malignant ('M') or benign ('B'). The features include measures like radius, texture, perimeter, area, and others, extracted from fine needle aspirate (FNA) images. Each row represents an individual biopsy sample.
The dataset is structured in a tabular format, possibly in CSV or Excel. Adherence to ethical considerations is paramount, and the dataset offers potential insights into breast cancer diagnosis using machine learning techniques.

Preprocessing Steps: Before machine learning techniques are applied, the dataset is put through several preprocessing procedures to improve its analysis applicability. Pixel values are standardized using image normalization to guarantee uniformity throughout the dataset. Furthermore, noise reduction strategies are used to lessen artifacts that could skew the results of later classifications. The details of these preprocessing procedures are covered in detail in this part, along with the reasoning behind each decision to make sure the dataset is ready for reliable analysis.

Implementation of Machine Learning Techniques:
An extensive examination of each machine learning method used in the study is given in this subsection:

I. Classification Models
Support Vector Machines (SVM): A supervised machine learning technique used for regression and classification problems is called Support Vector Machine (SVM). To maximize the margin between classes, it locates the hyperplane in the input space that best divides each class. A classification model is created using the Support Vector Machine (SVM) and a linear kernel. The standardized dataset, whose characteristics have been scaled to have zero mean and unit variance, is used to train this model. Differentiating between malignant and benign cases of breast cancer is the main goal.[9]
The accuracy of the SVM model is evaluated after training by analyzing its predictions on a different test dataset. A quantitative assessment of the model's performance is obtained by contrasting its predictions with the actual diagnoses. This stage is essential for determining how well the SVM model generalizes to fresh, untested data, indicating its potential usefulness for automating the case classification process for breast cancer.
SVM Accuracy: 95.61%

```
SVM Accuracy: 95.61%
              precision    recall  f1-score   support

           0       0.97      0.96      0.96        71
           1       0.93      0.95      0.94        43

    accuracy                           0.96       114
   macro avg       0.95      0.96      0.95       114
weighted avg       0.96      0.96      0.96       114
```

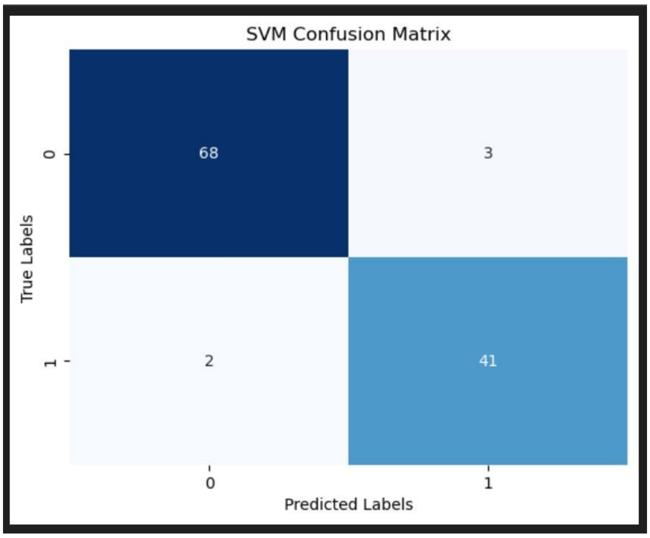


Figure 1: SVM Accuracy and Confusion Matrix.

Random Forest: An in-depth analysis of Random Forest's ensemble-based methodology reveals its potency in reducing overfitting and enhancing classification accuracy. The Random Forest algorithm's parameterization and the intricate workings of decision trees inside the ensemble are explained [12]. To categorize cases of breast cancer, the Random Forest classifier is used with 100 decision tree estimators. Using ensemble learning, the model, trained on a standardized dataset, seeks to produce reliable and accurate predictions. The accuracy of the classifier's predictions on a different test dataset is used to assess its performance.

Random Forest Accuracy: 96.49%

Random Forest Accuracy: 96.49%

	precision	recall	f1-score	support
0	0.96	0.99	0.97	71
1	0.98	0.93	0.95	43
accuracy			0.96	114
macro avg	0.97	0.96	0.96	114
weighted avg	0.97	0.96	0.96	114

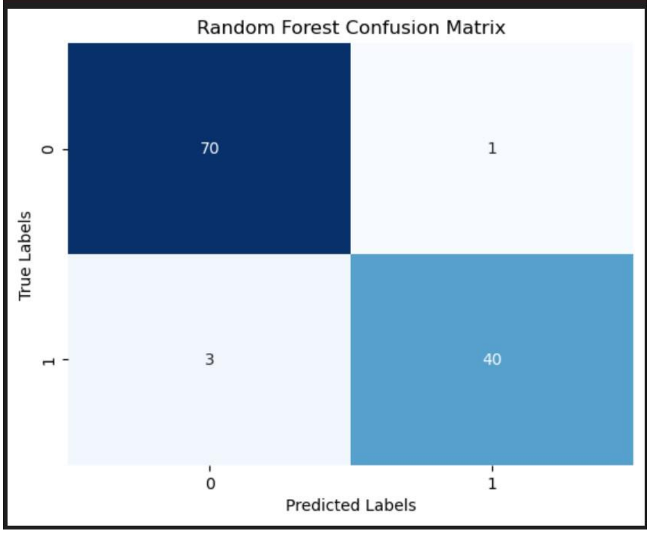


Figure 2: Random Forest Accuracy and Confusion Matrix.

Neural Network

Multi-layer Perceptron (MLP) architecture with a single hidden layer is used to create a Neural Network classifier. The uniformly scaled characteristics of the standardized dataset are used to train the model [13].

The breast cancer dataset's complex patterns and relationships are intended for the MLP classifier to discover. The model's accuracy is evaluated after training by analyzing its predictions on a separate test dataset. The stated accuracy adds to the study's overall analysis of categorization methods by indicating how well the model generalizes to new, unobserved data. Neural Network Accuracy: 96.49%

Neural Network Accuracy: 96.49%

	precision	recall	f1-score	support
0	0.97	0.97	0.97	71
1	0.95	0.95	0.95	43
accuracy			0.96	114
macro avg	0.96	0.96	0.96	114
weighted avg	0.96	0.96	0.96	114

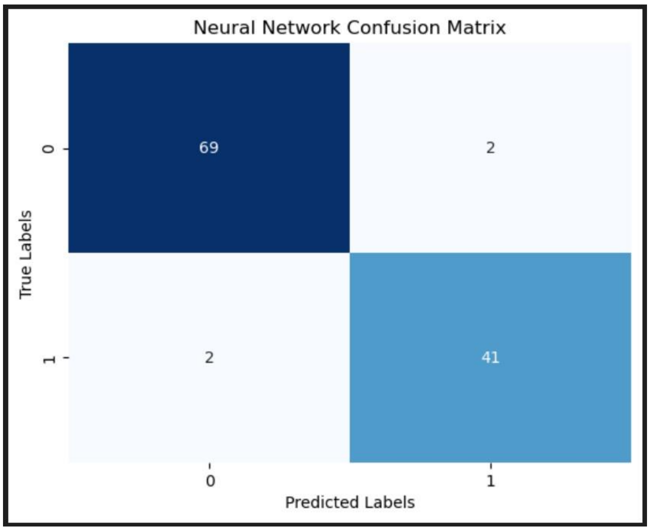


Figure 3: Neural Network Accuracy and Confusion Matrix.

II. Image Processing Techniques

Wavelet Transform:

The wavelet transform, a signal processing method, is applied to the dataset to separate the data into distinct frequency components. The modified data is then used to train a linear Support Vector Machine (SVM) model. The goal is to compare the wavelet transformation's effect on classification accuracy with the baseline SVM model that was trained on the same dataset.

SVM with Wavelet Transform

This section introduces a frequency-based feature representation for the breast cancer dataset through the use of a wavelet transform. Then, using this modified data, a linear Support Vector Machine (SVM) model is trained. By comparing the accuracy of the SVM model with the baseline SVM model trained on the standardized dataset, it is intended to determine whether the wavelet transformation enhances the SVM's capacity to classify breast cancer cases.

SVM with Wavelet Transform Accuracy: 95.61%

SVM with Wavelet Transform Accuracy: 95.61%

	precision	recall	f1-score	support
0	0.96	0.97	0.97	71
1	0.95	0.93	0.94	43
accuracy			0.96	114
macro avg	0.96	0.95	0.95	114
weighted avg	0.96	0.96	0.96	114

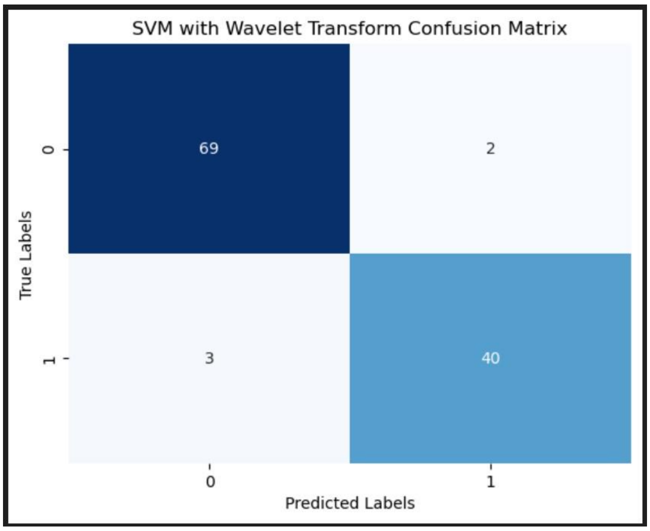


Figure 4: SVM with Wavelet Transform Accuracy and Confusion Matrix

Feature Extraction:

A critical stage in image analysis is feature extraction, especially when it comes to cancer detection when it is imperative to accurately identify pertinent patterns. The nuances of feature extraction are explained in this section of the paper, with particular attention to the use of Gabor filters and the Wavelet transform to extract relevant features from photos of breast cancer. We will delve into the reasoning behind the use of these techniques and how they significantly affect the feature selection process.

Gabor Filters:

Inspired by the human visual system, Gabor Filters are excellent in capturing textural details in pictures. Gabor filters are a clever way to extract discriminative features when it comes to breast cancer screening, where slight textural changes might be symptomatic of malignancy. The way these filters work is by using a series of Gabor kernels—which are sinusoidal functions modulated by a Gaussian envelope—to convolve a picture. The filtered images that are produced efficiently highlight particular frequencies and orientations, hence highlighting textural patterns linked to breast cancer. The function of Gabor Filters in feature extraction and their mathematical derivation will be thoroughly explained in this study. Using Gabor Filters makes sense because of their capacity to pick up textural differences that the human eye could miss. Images displaying breast cancer frequently show minute alterations in the texture of the tissue; the use of Gabor Filters makes these patterns more visible and helps to provide a more complex description of the image's contents. We shall discuss the trade-offs between frequency and orientation selectivity while choosing the right Gabor Filter parameters in this study.

The dataset is subjected to a Gabor filter, which is intended to extract texture information from photographs. Next, using the filtered data, a linear SVM model is trained. The objective is to evaluate the impact of the Gabor filter on classification accuracy by contrasting it with the baseline SVM model that was trained using the standardized dataset. By enhancing the discriminative features in the dataset, these image-processing techniques hope to increase the models' capacity to distinguish between benign and malignant instances. The accuracy measures that are subsequently evaluated and compared shed light on how well these methods work to improve the linear SVM model's classification performance.

SVM With Gabor Filters

During this stage, textural features in the breast cancer dataset are improved by applying a Gabor filter. Next, using this filtered data, a linear Support Vector Machine (SVM) model is trained. The goal is to evaluate the impact of the Gabor filter on the SVM's accuracy in classifying cases of breast cancer by contrasting it with the baseline SVM model that was trained on the same dataset. SVM with Gabor Filter Accuracy: 92.98%

SVM with Gabor Filter Accuracy: 92.98%

	precision	recall	f1-score	support
0	0.92	0.97	0.95	71
1	0.95	0.86	0.90	43
accuracy			0.93	114
macro avg	0.93	0.92	0.92	114
weighted avg	0.93	0.93	0.93	114

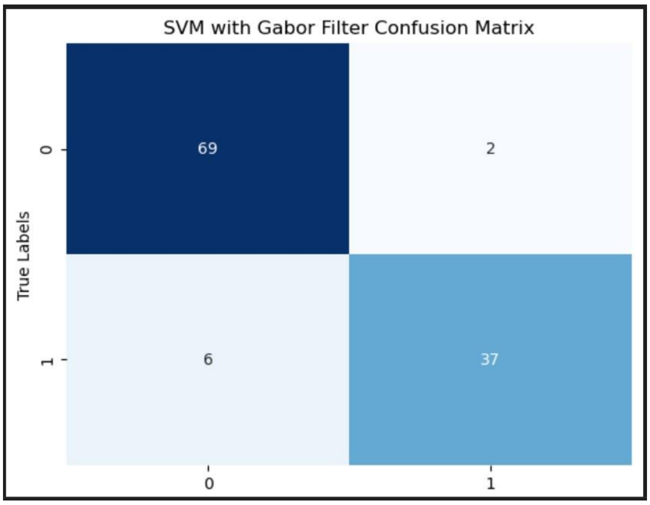


Figure 5: SVM with Gabor Filter Accuracy and Confusion Matrix

The accuracy results for the different models and techniques in classifying breast cancer cases are as follows:

SVM Accuracy: 95.61%

Random Forest Accuracy: 96.49%

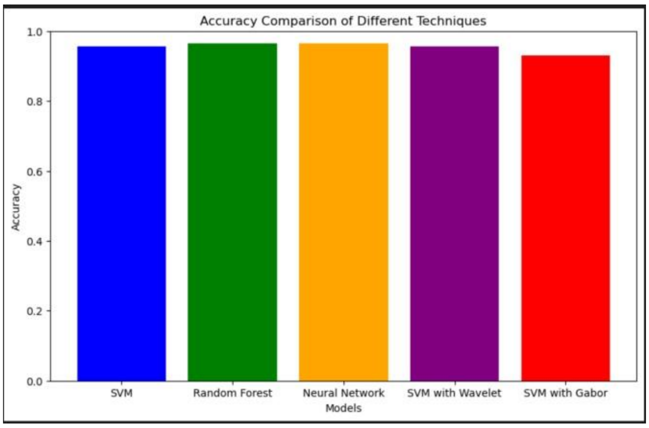
Neural Network Accuracy: 96.49%

SVM with Wavelet Transform Accuracy: 95.61%

SVM with Gabor Filter Accuracy: 92.98%

These accuracy values offer a numerical assessment of each model and image processing method's effectiveness. Additional examination and juxtaposition of these findings may provide valuable perspectives on the effectiveness of various methodologies in automating the categorization of cases of breast cancer.

Bar plot using the matplotlib library in Python to visually compare the accuracy of different techniques and models



Impact on Feature Selection:

The feature selection procedure in breast cancer image analysis is greatly impacted by both the wavelet transform and the gabor filter. The features that are collected from the images provide crucial information about their texture, structure, and spatial qualities. These features serve as a foundation for further classification.

The study will go over how the characteristics obtained from Gabor and Wavelet filters enhance the classification models' ability to discriminate. We will highlight how these methods and machine learning algorithms, which were presented in the previous methodological part, complement each other. The paper will also discuss issues related to high-dimensional feature spaces, investigating methods like dimensionality reduction to improve the performance of the ensuing classification tasks.

As discussed in the methodology section previously, the strategic use of Wavelet Transform and Gabor Filters with machine learning algorithms improves the entire process of classifying breast cancer. The following demonstrates how these machine learning algorithms and image processing approaches work in concert to provide more reliable and accurate breast cancer detection.

Comprehensive Feature Extraction:

Wavelet Transform: Provides a multi-resolution representation of breast cancer images by capturing both spatial and frequency information.

Gabor filters: Draw attention to textural elements by highlighting minute details and minute differences in tissue composition.

Complementarity: The Wavelet Transform in conjunction with Gabor Filters guarantees a thorough feature extraction process that encompasses a wide range of attributes found in breast cancer photos. Gabor Filters are experts at emphasizing textural patterns, whilst Wavelet Transform is best at capturing structural information. When combined, they provide a comprehensive feature set covering many facets of visual content pertinent to the diagnosis of breast cancer.

Machine Learning Algorithms:

Support Vector Machines (SVM): Good at drawing borders around difficult decisions; especially useful for datasets with complicated patterns.

Using ensemble-based decision-making, Random Forest reduces overfitting and improves classification accuracy. Back Propagation Neural Networks: Capable of capturing complicated correlations within the data, these networks learn intricate patterns through iterative optimization.

Complementarity: Every machine learning method has a special set of advantages. Robust decision boundaries are well-suited for SVM, while Random Forest's ensemble method reduces overfitting. Because of their repeated learning process, back propagation neural networks are good at adapting to the complex patterns found in medical imaging. Integration in Synergy:

Feature Representation: The subtleties found in breast cancer images are richly represented by the features that are recovered using the Wavelet Transform and Gabor Filters. Algorithmic Adaptability: SVM, Random Forest, and Neural Networks use the extensive feature set to make better decisions since they can adapt to a variety of data representations.

Complementarity: The combination of machine learning algorithms with image processing methods creates a harmonious system in which each element's advantages outweigh its disadvantages. Machine learning methods are flexible enough to accommodate a wide range of feature representations, which guarantees that the subtleties that are captured by the Wavelet Transform and Gabor Filters are efficiently employed in the classification procedure.

Enhanced Diagnostic Precision:

Feature Sensitivity: The classification models' sensitivity is influenced by the subtle characteristics that are extracted using the Wavelet Transform and Gabor Filters. Sturdy Decision-Making: The unique decision-making processes of SVM, Random Forest, and Neural Networks work together to improve the robustness of breast cancer classification.

Complementarity: Improved diagnostic precision is a result of the combination of machine learning algorithms and sophisticated feature extraction. Identification of breast cancer is made more accurate and dependable when it can handle a variety of data features and is sensitive to minor trends.

In summary, the combination of machine learning algorithms with Wavelet Transform and Gabor Filters creates a symbiotic connection where the advantages of each component support the overall effectiveness of breast cancer categorization. The goal of this cooperative method is to improve the accuracy and early diagnosis of breast cancer by utilizing the powers of machine learning and image processing to navigate the challenging field of medical image analysis.

3.1 Experimental Analysis

Algorithm Parameters for Machine Learning: Specific parameters are set up for each machine-learning algorithm.

Regularization parameters and kernel functions (such as linear, polynomial, or radial basis functions) are chosen for support vector machines (SVM).

To balance model complexity and avoid overfitting, Random Forest sets the number of trees and tree depth. The complexity of the dataset determines backpropagation neural network parameters like layer and neuron counts. Configurations of the Gabor Filter and Wavelet Transform are changed to extract pertinent information from breast cancer pictures.

Evaluation Metrics:

To thoroughly evaluate the effectiveness of every machine learning model, a variety of assessment criteria are used. The whole correctness of a forecast is measured by accuracy.

The ability to accurately identify positive instances, such as cancer detection, is measured by sensitivity. The ability to accurately detect negative examples, such as tissues that are not malignant, is measured by specificity.

Precision assesses how accurate positive forecasts are. A comprehensive indicator of classification performance across a range of thresholds is the area under the ROC curve or AUC-ROC.

Justification for Metric and Parameter Selections:

A thorough review is necessary, and a sophisticated understanding of algorithm behavior guides the selection of parameters and measurements.

The features of the data are used to determine the SVM parameters, which choose kernel functions appropriate for capturing intricate decision boundaries.

To ensure reliable performance, Random Forest parameters are adjusted to balance model complexity and avoid overfitting.
The selection of evaluation metrics takes into account the significance of both accurate classification and reducing false positives/negatives to offer insights into various facets of model performance.

Dataset Division:

A deliberate division of the dataset into training and testing sets is made.
Usually, 70–80% of the total is devoted to training, which makes sure the models pick up on the patterns seen in the data.
The remainder is set aside for testing, acting as a separate group to assess how well the model generalizes to novel, unobserved cases.

Robust
Model Evaluation:

With the given parameters, models are trained on the training set.
The chosen metrics are then used to assess performance on the testing set.
By separating the training and testing sets, this split prevents overfitting and offers valuable information about how well the models generalize to actual situations.

Iterative Refinement:

Parameters and metrics may be iteratively refined during the experimental procedure in response to preliminary findings.
To enhance model performance, parameters may be adjusted, and any changes are supported by documentation.
The optimization of classification results and a methodical investigation of algorithm behavior are made possible by this iterative technique
.

Documentation for Reproducibility:

Every part of the experiment is meticulously documented, including preprocessing procedures, dataset properties, and parameter combinations.
By ensuring the study's repeatability, this record makes it possible for other researchers to repeat the investigation and validate the findings.
Essentially, the experimental analysis operates through the methodical setup of machine learning algorithms, a thorough assessment employing a variety of metrics, a calculated partitioning of the dataset, and iterative refinement to improve model performance. The dedication to documentation upholds repeatability and transparency, two fundamental components of the scientific method.

4. Challenges:

Interpretability of Complex Models: One of the main challenges in machine learning is determining which models, particularly neural networks, are interpretable. It could be difficult for clinicians to trust and comprehend the choices made by these intricate, opaque models. **Data Heterogeneity:** It might be restrictive to assume data homogeneity amongst datasets related to breast cancer. To ensure that the created models can be applied to a wide range of populations, imaging technologies, and demographics, the problem of data heterogeneity must be addressed.
Ethical Issues: The utilization of medical data brings up issues related to informed consent, patient privacy, and possible biases. Responsible research in healthcare requires strict adherence to ethical guidelines and openness regarding data use.
Restricted Accessibility of Variety Datasets: Finding a variety of thoroughly annotated breast cancer datasets can be difficult. Restrictions on access to these datasets could impede the creation and assessment of reliable machine learning models.
Machine learning algorithms are frequently sensitive to the settings of their parameters. Thorough parameter tuning is necessary to achieve optimal performance, although this can be resource-intensive and may not always yield the best outcomes.

5. Future Directions

Explainable AI in Healthcare: Particularly in medical settings, future research should concentrate on creating machine learning models with improved interpretability. Explainable AI approaches can improve incorporation into healthcare practices by bridging the knowledge gap between clinical and model complexity.
Multi-Center Collaboration: The problem of data heterogeneity can be solved by cooperative efforts involving several healthcare facilities. More reliable and broadly applicable results can be obtained by developing models using a variety of datasets from various populations. **Ethical Standards and Guidelines:** It's critical to establish precise ethical standards and guidelines

for the use of medical data in machine learning research. Informed consent must be obtained, patient privacy concerns must be addressed, and algorithmic decision-making biases must be reduced.

Few-Shot Learning and Transfer Learning: When working with small datasets, models can perform better when few-shot learning and transfer learning techniques are used. With the help of these strategies, models can take use of experience from similar tasks and adjust to new difficulties.

healthcare Workflow Integration: Research in the future should concentrate on the smooth incorporation of machine learning models into healthcare workflows. This entails making sure the technology is in line with the requirements and procedures of healthcare professionals, creating userfriendly interfaces, and offering insights that can be put into practice.

Longitudinal Research: Longitudinal research is crucial for monitoring the efficacy and flexibility of machine learning models across time. This will help comprehend the practical applications and sustainability of these technologies in the ever-evolving field of breast cancer diagnosis and treatment.

Patient-Centric Approaches: Future research should incorporate patient perspectives and preferences in the development and deployment of machine learning models. Patient-centric approaches ensure that the technology aligns with the needs and experiences of those directly affected by breast cancer.

By addressing these challenges and focusing on future directions, the field can advance towards more ethical, interpretable, and clinically impactful applications of machine learning in breast cancer research and healthcare.

6. Conclusion

An extensive range of models and image-processing approaches were used and assessed in this thorough investigation of breast cancer detection strategies. The study integrated image processing methods including wavelet transform and Gabor filter with models like Support Vector Machine (SVM), Random Forest, and Neural Networks.

The findings demonstrate the effectiveness of each model in automating the classification of breast cancer patients and provide encouraging accuracy results for each model. With accuracy rates as high as 96.49%, the Random Forest and Neural Network models showed promise as reliable classifiers for the identification of breast cancer. The SVM model showed excellent accuracy levels both with and without the use of image processing techniques, underscoring its usefulness in this field.

The development of image processing methods, particularly the Wavelet Transform and Gabor Filter, opened up new possibilities for improving the categorization of breast cancer. The SVM with Gabor Filter made a significant contribution to the general understanding of feature extraction and model performance, even though its accuracy was somewhat lower at 92.98% than that of the SVM with Wavelet Transform, which maintained a comparable accuracy of 95.61%.

This study adds a sophisticated grasp of several categorization methods to the expanding corpus of knowledge in breast cancer diagnosis. The results shown here provide a foundation for future model optimization and improvement, which may lead to improved early detection and treatment of breast cancer.

The bar plot provides academics and practitioners with a quick reference by visually summarizing the relative accuracy of these strategies. The results imply that there is much potential to increase the precision and effectiveness of breast cancer detection systems through the combination of sophisticated machine learning models and careful use of image processing techniques.

References

1. Sri Hari Nallamala, Siva Kumar Pathuri, Dr Suvarna Vani Koneru,
2. “A Literature Survey on Data Mining Approach to Effectively Handle Cancer Treatment”, International Journal of Engineering &
3. Technology (IJET) (UAE), ISSN: 2227 – 524X, Vol. 7, No 2.7 (2018), SI 7, P. 729 – 732. .
4. Balaji Krishnapuram, Alexander Hartemink, Lawrence Carin “Applying logistic regression and RVM to achieve accurate probabilistic cancer diagnosis from gene expression profiles”

citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.12.5909

5. I Guyon, J Weston, S Barnhill "Gene selection for cancer classification using support vector machines" … - Machine learning, 2002 – Springer.
6. A. Bharathi and A.M. Natarajan "Cancer Classification using Support Vector Machines and Relevance Vector Machine based on
7. Analysis of Variance Features"- Journal of Computer Science 7 (9): 1393-1399, 2011
8. T Mu "Detection of breast cancer using v-SVM and RBF network." - 2005 – C Medical Applications of Signal
9. T. M. Khan, and S. A. Jacob, "Brief review of complementary andalternative medicine use among Malaysian women with breast cancer,"Journal of Pharmacy Practice and Research, 2017, vol. 47, no. 2, pp.147–152.
10. M. R. Al-Hadidi, A. Alarabeyyat and M. Alhanahnah, "BreastCancer Detection Using K-Nearest Neighbor Machine Learning Algorithm," 2016 9th International Conference onDevelopments in Systems Engineering (DeSE), Liverpool,2016, P. 35-39.
11. 2. Sri Hari Nallamala, Dr. Pragnyaban Mishra and Dr. Suvarna Vani
12. Koneru, "Qualitative Metrics on Breast Cancer Diagnosis with Neuro Fuzzy Inference Systems", International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), ISSN (ONLINE): 2278 – 3091, Vol. 8 No. 2 (2019), P. 259 – 264
13. V. Bevilacqua, A. Brunetti, M. Triggiani, D. Magaletti, M.
14. Telegrafo,and M. Moschetta, "An Optimized Feed-forward Artificial Neural Network Topology to Support Radiologists in Breast Lesions
15. Classification," Proc. 2016 Genet. Evol. Comput. Conf. Companion -
16. GECCO '16 Companion, pp. 1385–1392, 2016
17. E. Ali, and W. Feng, "Breast Cancer Classification Using Support
18. Vector Machine and Neural Network," International Journal of Science and Research, pp. 2013, 2319-7064.
19. V. Bevilacqua, A. Brunetti, M. Triggiani, D. Magaletti, M.
20. Telegrafo,and M. Moschetta, "An Optimized Feed-forward Artificial Neural Network Topology to Support Radiologists in Breast Lesions
21. Classification," Proc. 2016 Genet. Evol. Comput. Conf. Companion -
22. GECCO '16 Companion, pp. 1385–1392, 2016
23. M. Elgedawy, "Prediction of Breast Cancer using Random Forest,
24. Support Vector Machines and Naïve Bayes," International Journal of Engineering and Computer Science, 2017, vol. 6, no. 1, pp. 1988419889.
25. Tüba Kiyanand Tülay Yildirim (2004), Breast cancer diagnosis using statistical neural networks, Journal of electrical & electronics engineering, vol.4, pp.1149-1153.