International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

EMPOWERING EARLY BREAST CANCER DETECTION: HARNESSING THE RELIEF ALGORITHM AND DIVERSE AI MODELS FOR IMPROVED DIAGNOSTICS

[1] S. Meenakshi Phd Scholar, [2] Dr. A. S. Arunachalam,

(Vels University, VISTAS, Chennai – 600117.)

Email: -whitecross2013meena@gmail.com.

Corresponding

MCA.,MPhil.,Ph.D., Professor,

Department of Computer Science, Vels Institute of Science,

Technology & Advance Studies (VISTAS),

Pallavaram, Chennai – 600117.)

Email: arunachalam.1976@gmail.com

Abstract— Early detection of breast cancer significantly improves patient outcomes and survival rates, underscoring its importance in healthcare. Early-stage breast cancer allows for less aggressive treatments and generally leads to better prognoses, as such cancers are smaller and localized, making them amenable to treatment with radiation or surgery. This study introduces a novel approach to early detection using the Relief Algorithm and advanced AI models (Multilayer Perceptron, Convolutional Neural Network, and Support Vector Machine), targeting the enhancement of diagnostic precision. Given breast cancer's role in cancer-related mortality among women worldwide, our research, utilizing the Kaggle Breast Cancer Dataset, focuses on improving early detection. We embarked on a comprehensive data preprocessing and feature selection process, employing the Relief Algorithm to identify key features for accurate classification. Our models demonstrated high accuracy in detecting breast cancer: MLP (94.68%), CNN (92.76%), and SVM (91.42%), highlighting the benefits of combining the Relief Algorithm with AI technologies in improving early detection accuracies.

Keywords: Breast Cancer Detection, Artificial Intelligence (AI), Feature Selection, Relief Algorithm, Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Support Vector Machine (SVM)

Introduction

Breast cancer, a leading concern globally, arises from the uncontrolled growth of abnormal cells in the breast tissue, affecting individuals of all ages and genders, though predominantly diagnosed in women. This disease encompasses a range of subtypes, each with distinct characteristics and behaviors, highlighting the necessity of grasping its complexities for effective prevention, detection, and treatment [1]. The global burden of breast cancer is significant, with the World Health Organization reporting approximately 2.3 million new cases annually, positioning it as the most common cancer among women. Moreover, it accounts for

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

over 685,000 deaths each year, ranking as the seventh deadliest cancer. These figures emphasize the critical need for holistic approaches to mitigate breast cancer's profound impact on global health.

Early detection is crucial in the fight against breast cancer, greatly improving treatment success and patient survival. It enables the discovery of cancer in its early stages, when treatments are most effective and survival rates are higher [2]. Healthcare professionals can then quickly and effectively treat patients, providing them with the best chance for a positive outcome. Screening programs for people without symptoms play a key role in early detection, allowing for early intervention and potentially life-saving treatments. Mammography, a non-invasive imaging technique, is a common method used to find early signs of breast cancer in specific age groups, with screening guidelines varying worldwide [3].

Regular mammographic screenings help healthcare providers spot early signs of breast tissue changes that may need further evaluation. Follow-up diagnostics such as diagnostic mammography, ultrasound, or MRI are employed to closely examine detected abnormalities like masses or calcifications. These detailed examinations are crucial for accurately diagnosing and staging breast cancer, enabling personalized treatment plans tailored to the patient's specific needs [4] [5]. The importance of early detection in influencing disease outcome and treatment efficacy cannot be overstated. Early-stage identification allows for less invasive treatments, minimizing side effects and maintaining quality of life. It also expands treatment options to include breast-conserving surgeries, targeted therapies, and hormone treatments, which might not be options at more advanced stages.

Early detection significantly impacts healthcare systems and society by reducing the costs and emotional burdens of advanced breast cancer [6]. Screening programs catching cancer early often mean less intensive treatments and shorter hospital stays, leading to lower healthcare costs and better resource use. Early diagnosis empowers individuals to make informed decisions about their care and treatment options. It offers hope and opportunity in the fight against breast cancer, marking a path towards earlier, more effective detection and potentially avoiding the disease altogether [7] [8]. Through comprehensive screening and timely action, maximizing the benefits of early detection can save lives, promote health, and envision a better future for all.

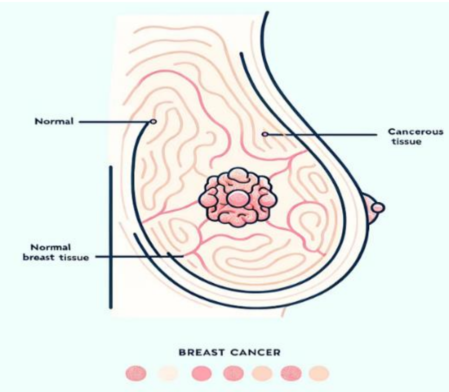


Figure 1. Breast Cancer Image

Received: August 08, 2025

Detection methods are crucial in identifying breast cancer early, with mammography being the primary tool. This X-ray imaging technique can spot lesions too small for physical detection [9]. Clinical breast exams by professionals further support mammography by identifying palpable changes. Advancements like digital breast tomosynthesis and MRI enhance detection capabilities, improving diagnosis accuracy and treatment outcomes by offering superior imaging that increases sensitivity and specificity.

Understanding breast cancer signs and symptoms is crucial for early detection and taking prompt action. While screenings help identify many cases, personal symptom awareness is also key. Common signs include changes in breast size or shape, nipple appearance or texture, skin changes like redness or dimpling, and new lumps in the breast or underarm. Recognizing these symptoms enables individuals to proactively seek medical advice, playing an active role in their health management [10]. A deep grasp of breast cancer statistics, detection techniques, early detection strategies, and symptoms is crucial for combating the disease effectively. Elevating awareness, investing in research and innovation, and emphasizing early detection can significantly lessen breast cancer's impact and enhance outcomes for those it affects.

Related Study

Cancer remains a significant global health issue, with breast cancer among the leading causes of death. The advent of high-throughput sequencing and deep learning has transformed prognosis and treatment strategies, focusing on gene expression for informed healthcare decisions. In diagnostic imaging, breast density and tissue texture are key indicators of cancer risk, guiding targeted screening and prevention. Deep learning has emerged as a pivotal tool for classification and prediction in breast imaging, leading to the development of models like InceptionResNetV2 based on transfer learning [11]. This model, tested on breast cancer patient data, showed promising results with a 91% accuracy rate, outperforming existing methods by incorporating risk factors into breast cancer risk assessments. Deep learning models enhance accuracy by including risk flags, highlighting their growing importance in improving medical imaging procedures and potentially automating them.

Breast cancer is a leading cause of death among women globally, with late detection significantly worsening outcomes. Early detection could drastically reduce mortality and associated risks. A study introduced a hybrid machine learning model aiming to improve early detection, using the Kaggle dataset to predict tumor development and size with methods like logistic regression, XBoost, MLPs, and random tree classification, implemented in Python [12]. The model achieved high accuracy (99.65%) in both training and testing, showing potential for enhancing detection, diagnosis, and treatment outcomes. It could also assist individuals in timely managing their health and lifestyle choices, crucial for overall well-being.

In combating breast cancer, leveraging deep learning (DL) and machine learning (ML) has significantly enhanced detection and diagnosis, promising better patient outcomes. A novel hybrid approach, combining DL's analytical strength with ML's precision via the pre-trained ResNet50V2 model, has shown exceptional performance on the Invasive Ductal Carcinoma (IDC) dataset, achieving notable metrics: 94.57% reliability, 94.86% accuracy, 94.32% recall, and an F1 score of 94.57%. The Light Boosting Classifier (LGB) proved to be the best ML match for ResNet50V2, enhancing both interpretability and generalizability [13]. This breakthrough in breast cancer detection underscores the potential for improved physician decision-making and patient care.

Breast cancer (BC) is a leading cause of female cancer mortality globally, highlighting the importance of early detection. This study introduces a novel framework combining numerical and visual data analysis with explainable Artificial Intelligence (XAI) for precise BC detection. It utilizes the U-NET transfer learning model for image-based predictions and an ensemble model incorporating SVM, random forest (RF), and a custom CNN for analyzing complex features [14]. Tested using the Wisconsin dataset, the approach demonstrated remarkable accuracy (99.99%) in BC diagnosis, indicating significant potential for enhancing early detection and patient outcomes. This research contributes to BC diagnostic advancements by offering a highly accurate machine learning-based method and showcasing how XAI can aid in interpreting the influencing factors.

The global challenge of breast cancer prompted the use of machine learning techniques such as XGBoost, logistic regression, K-nearest neighbor, and random forest for early detection. These models were evaluated based on recall, precision, accuracy, and F1-score, with recall being especially important for identifying malignant cells. Performance improvements were achieved by removing fifteen attributes identified through the Pearson correlation test [15]. The K- nearest neighbor approach applied cross-validation to determine the best k-value and used hierarchical sampling to manage sample imbalance. The XGBoost model emerged as the most effective, demonstrating superior performance with an F1-score of 0.980, perfect recall, precision of 0.960, and accuracy of 0.974, under an 8:2 division of training and testing sets.

Methodology

Our breast cancer detection method utilizes the Kaggle dataset, employing thorough preprocessing including missing value handling, outlier removal (IQR), data normalization (Robust Scaler), and categorical data conversion (Label Encoding). Key feature selection employed the Relief Algorithm, followed by MLP, CNN, and SVM models for pattern recognition and classification. Assessment metrics included accuracy, precision, specificity, and sensitivity, advancing early detection. Figure 2 depicted the model workflow.

Dataset Collection

The Kaggle Breast Cancer Dataset, comprising 569 cases and 32 attributes, categorizes tumors as "Malignant" or "Benign" and offers insights into breast cancer [16]. It encompasses detailed characteristics analyzed for mean values, standard error, and worst conditions, facilitating predictive modeling for tumor malignancy assessment. Exploratory Data Analysis (EDA) reveals essential patterns, enhancing detection accuracy, and advancing diagnostics, benefiting data scientists and medical professionals alike.

Data Preprocessing

Data preparation in data science and machine learning is vital, enhancing accuracy and efficiency. It involves cleaning data, filling missing values, normalizing features, and

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

converting categorical variables to numerical ones. Preprocessing includes exploratory analysis for pattern identification, aiding model and feature selection. It ensures compatibility between datasets by standardizing formats, laying the groundwork for reliable analysis and robust predictions in machine learning.

Handling Missing Values:

The efficacy of machine learning models in cancer diagnosis relies on handling missing values in the Breast Cancer Dataset. Imputation methods, like substituting missing values with statistical measures or common values, maintain data integrity. Advanced techniques, such as predictive modeling, estimate missing values based on dataset patterns, potentially improving predictive accuracy.

Replace missing value $x_{missing}$ with the mean (μ) of the available data points.

$$x_{missing} = \mu = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (1)$$

Removing columns or rows with extensive missing data simplifies datasets but risks losing crucial insights, notably in sensitive cases like the Breast Cancer Dataset. Algorithms like decision trees and random forests accommodate missing data without deletion or imputation, preserving the dataset integrity, vital for uncovering detailed cancer patterns, especially in medical datasets.

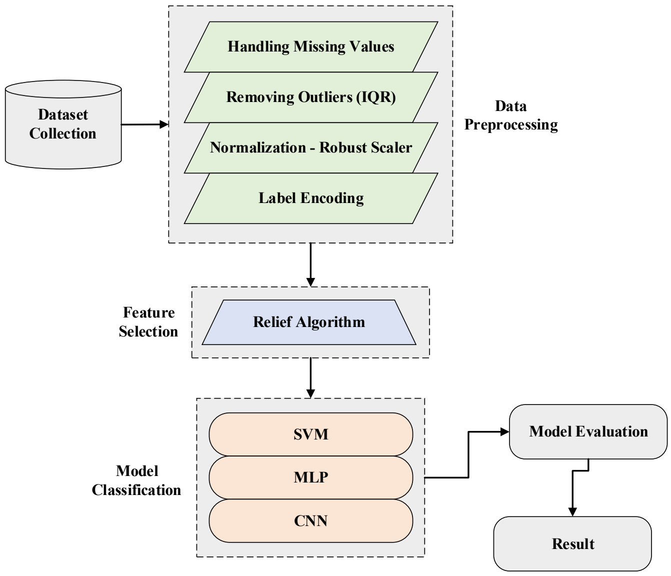


Figure 2. Workflow of Proposed Model

Removing Outliers: Interquartile Range (IQR):

Identifying and removing outliers in the Breast Cancer Dataset is vital for data integrity and accuracy. The Interquartile Range (IQR) method is commonly used to detect outliers by accessing data distribution, effectively excluding extreme values without their influence. The

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

IQR is calculated as the difference between the 75th percentile (Q_3) and the 25th percentile (Q_1) of the data, providing a measure of the middle 50% of the dataset's variability and spread.

The IQR is calculated as:

$$IQR = Q_3 - Q_1 \qquad (2)$$

Normalization - Robust Scaler:

In datasets prone to outliers, like the Breast Cancer Dataset, normalization with the Robust Scaler is crucial to maintain prediction model accuracy. Traditional methods like Min-Max scaling or Z-score standardization may be skewed by outliers. The Robust Scaler addresses this by using the IQR and median, centering data by the median—less affected by outliers than the mean—and scaling based on the IQR, the range between the data's first and third quartiles.

The Robust Scaler approach protects the dataset's integrity by scaling data without being skewed by extreme values, crucial for analyses like breast cancer where accurate tumor measurement is key. It prevents predictive models from being biased by atypical data points, enhancing model reliability and accuracy. This makes the Robust Scaler an essential preprocessing tool for outlier-prone datasets.

Feature Encoding - Label Encoding:

Feature encoding, like Label Encoding, converts categorical data for machine learning, assigning unique numerical labels to categories, simplifying analysis without inflating dataset size. Unlike One-Hot Encoding, which expands features, Label Encoding's simplicity may introduce artificial ordinality, misleading algorithms, especially linear regression sensitive to variable magnitude and order.

Label Encoding assigns a unique integer to each category. If a categorical variable C has n unique values, then each unique value c_i is assigned a unique integer l_i such that:

$$l: c_i \rightarrow \{0,1,2,\ldots,n-1\} \qquad (3)$$

Feature Selection

Feature selection is essential in ML model prep, reducing overfitting and enhancing interpretability by identifying relevant characteristics, thus improving generalization and training speed. The versatile Relief algorithm evaluates feature relevance by distinguishing classes, assigning importance scores to optimize model input, ensuring accurate and interpretable models.

Relief Algorithm:

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Feature selection is crucial in machine learning for identifying the most relevant characteristics for model training. The Relief algorithm ranks features based on their ability to differentiate values among nearby examples. It randomly samples an instance and assesses its nearest hit and miss neighbors within and outside the same class, adjusting weights accordingly. This process highlights attributes that effectively distinguish between neighboring instances of different classes.

Let's α denotes a specific attribute of the dataset, $W[a]$ represents the weight of attribute a, initialized to 0, x stands randomly selected instance from the dataset, x_{hit} represents the nearest neighbor of x belonging to the same class, x_{miss} represents the nearest neighbor of x belonging to a different class.

The algorithm updates attribute weights iteratively based on the difference between the selected instance x and its nearest hit and miss.

$$W[a] = W[a] - \frac{diff(a, x, x_{hit})}{N} + \frac{diff(a, x, x_{miss})}{N} \qquad (4)$$

Where N denotes the total number of instances sampled, $diff(a, x, y)$ represents a function that measures the difference between the values of attribute a for instances x and y. For nominal attributes, $diff$ might be 0 if $x_a = y_a$ and 1 otherwise. For numerical attributes, $diff$ could be the absolute difference normalized by the range of attributes.

Steps in Application:

During initialization, features are set to zero weight, later updated to reflect their significance in distinguishing classes. Random sampling aids in grasping cancer indicators within the Breast Cancer Dataset. Identifying nearest hits and misses reveals local structure and diagnosis-relevant patterns. Feature weights adapt based on differentiation capability, emphasizing reliable indicators like tumor size or cell irregularity. Recurring sampling and updating ensure dataset representation. Finally, features with the highest weights are pivotal for classification, possibly reducing the dataset while improving predictive accuracy.

Classification Models

The classification phase is pivotal in machine learning pipelines, particularly in cancer diagnosis, aiming to differentiate between benign and malignant tumors accurately. Selected features, often chosen using methods like the Relief algorithm, train models for predicting classes of new instances. Notably, SVM, MLP, and CNN stand out as prominent models in this process.

Support Vector Machine (SVM)

SVMs aim to find the hyperplane that optimally separates data points in an N-dimensional space based on their categories. Kernel functions play a vital role by facilitating operations in a modified feature space without directly computing its coordinates. This simplifies handling linearly separable data through higher-dimensional mapping.

Multilayer Perceptron (MLP)

Received: August 08, 2025

MLP neural networks have interconnected nodes across layers. Inputs to each neuron are weighted and pass through a nonlinear activation function. Backpropagation is used for training, adjusting weights to minimize the difference between expected and actual outputs.

Introducing regularization and a more detailed description of the backpropagation process for MLP, the weight update rule can be represented as:

$$w_{ij}^{(k+1)} = w_{ij}^{(k)} - \eta \frac{\partial \mathcal{E}}{\partial w_{ij}} - \lambda . sign\left(w_{ij}^{(k)}\right) \quad (5)$$

Where $w_{ij}^{(k+1)}$ and $w_{ij}^{(k)}$ are the weights connecting neurons i and j at iteration $k+1$ and k, respectively, η is the learning rate, ε is the loss function, λ is the regularization coefficient, $sign\left(w_{ij}^{(k)}\right)$ introduces $L1$ regularization for sparsity.

Convolutional Neural Network (CNN)

CNNs, a neural network subset, excel with grid-like data like images. Their architecture includes convolutional, pooling, and fully connected layers. Convolutional layers use filters to capture spatial and temporal data relationships through convolution operations on the input grid.

Results And Discussions

Deep learning on Google Colab with Intel® Core™ i7-1355U Processor and 8GB RAM streamlines computational tasks, offering free access to GPUs and TPUs for cost-effective model training. Its high performance manages large datasets efficiently, democratizing computing access for researchers. Google Colab's collaborative platform aids project sharing. In our analysis of the Kaggle Breast Cancer Dataset, we detailed its composition, preprocessing, and feature selection, revealing insights into diagnostics via tumor characteristics. Figure 3 showcased non-missing dataset values.

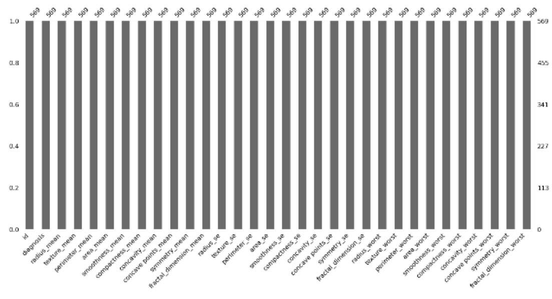


Figure 3. Non - Missing Value in Dataset

In our examination of the Kaggle Breast Cancer Dataset with 569 cases, we found 359 benign and 210 malignant cases, highlighting the need for accurate classification models. Figure 4 showed the distribution. Precise classification is crucial to minimize unnecessary interventions for benign conditions and ensure prompt treatment for malignancies. This understanding

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

guided our preprocessing and feature selection to enhance model accuracy in distinguishing between these crucial categories.

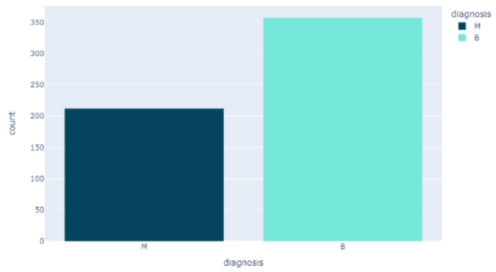


Figure 4. Target Distribution in Dataset

We addressed missing values through imputation, maintained data integrity. Outliers were removed using the IQR method, refining the dataset. Normalization with the Robust Scaler ensured balanced scaling. Feature Encoding with Label Encoding prepared categorical variables for analysis. Figure 5 showcased insights from our preprocessing of the breast cancer dataset.

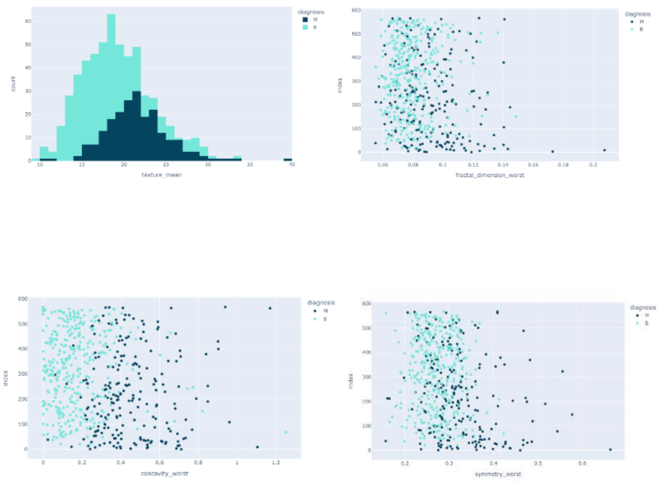


Figure 5. Exploratory Analysis of Breast Cancer Dataset

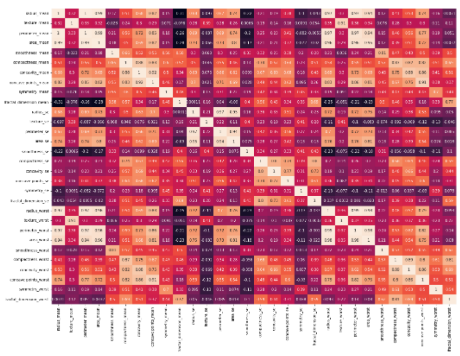


Figure 6. Correlation Matrix

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Understanding feature interactions before applying the Relief algorithm for selection is crucial. The correlation matrix reveals relationships, aiding in identifying positive or negative correlations. Visualizing these in a heatmap offers intuitive insights into interconnections and dependencies, informing informed decisions during selection. While the Relief algorithm isn't directly influenced by feature scale, understanding their interplay ensures a comprehensive approach. Figure 6, showing the Correlation Matrix, guides subsequent analysis with deep dataset understanding.

Table 1 and Figure 7 compare feature selection methods' accuracy in a dataset. PCA yields 88.50% accuracy by generating uncorrelated variables for dimensionality reduction. Genetic Algorithms achieve 93.77% accuracy, showing promise in trait selection. The Chi-Squared Test attains 90.65% accuracy for categorical data independence. Fisher Score achieves 94.20% accuracy in supervised learning. ANOVA F-value reaches 91.00% accuracy in comparing sample means. Notably, the Relief algorithm outperforms others with 96.34% accuracy, focusing on feature relevance to neighboring instances.

Table 1. Accuracy Comparison for Various Feature Selection

Feature Selection	**Accuracy**
PCA	88.50
GA	93.77
Chi-Squared Test	90.65
Fisher Score	94.20
ANOVA F-value	91.00
Relief	96.34

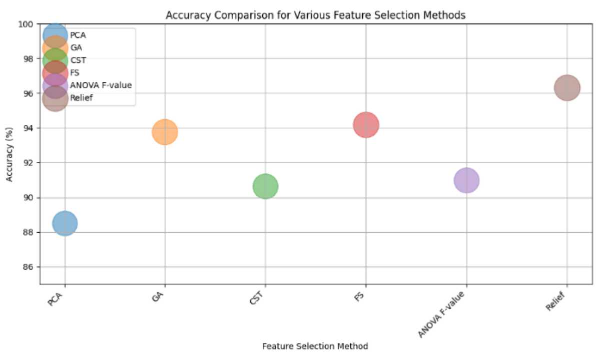


Figure 7. Accuracy Comparison

After thorough feature selection, we split the dataset into training and testing sets (80:20 ratio) for precise model evaluation. Utilizing MLP, CNN, and SVM diversifies our approach to breast

International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

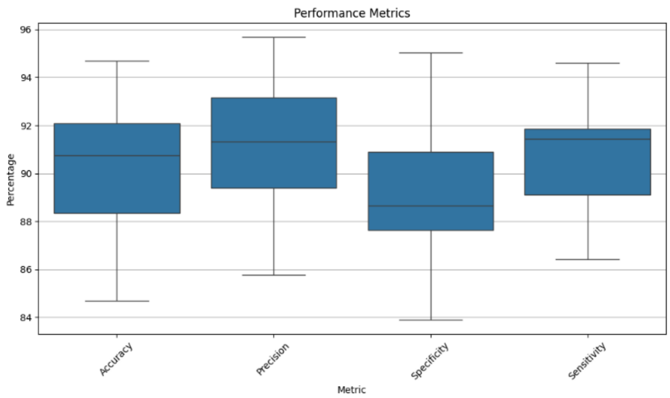
cancer classification, enhancing analysis robustness. Post-model application, we compare metrics like accuracy, sensitivity, and specificity to identify the most promising detection method.

Table 2. Model Evaluation Metrics

Model	Accuracy	Precision	Specificity	Sensitivity
Decision Tree	87.50	88.24	86.90	88.30
Gradient Boosting	90.76	91.32	89.45	92.10
K-Nearest Neighbors (KNN)	89.22	90.55	88.67	89.90
Logistic Regression (LR)	84.68	85.77	83.90	86.42
SVM	91.42	93.54	88.38	91.42
MLP	94.68	95.68	95.04	94.62
CNN	92.76	92.76	92.36	91.64

Table 2 and Figure 8 assesses machine learning models using Accuracy, Precision, Specificity, and Sensitivity, providing insights into their classification performance and areas for improvement in prediction accuracy and error reduction.

The Decision Tree model shows balanced performance with 87.50% accuracy, 88.24% precision, 86.90% specificity, and 88.30% sensitivity, indicating its versatility but with room for improvement. Gradient Boosting stands out with 90.76% accuracy and 91.32% precision, showcasing high efficacy in complex tasks with its 89.45% specificity and 92.10% sensitivity. The KNN model achieves 89.22% accuracy, demonstrating balanced identification and detection capabilities with 90.55% precision, 88.67% specificity, and 89.90% sensitivity.



International Journal of Applied Mathematics
Volume 38 No. 7s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Figure 8. Performance Evaluation Metrics

Logistic Regression provides solid results with 84.68% accuracy, 85.77% precision, and balanced identification metrics. SVM offers high precision at 93.54% and 91.42% accuracy, excelling in scenarios with high costs. MLP leads with 94.68% accuracy, excelling in precision and robust classification across various tasks. CNN showcases strong accuracy at 92.76%, with matched precision and high specificity and sensitivity, suitable for critical applications like medical imaging.

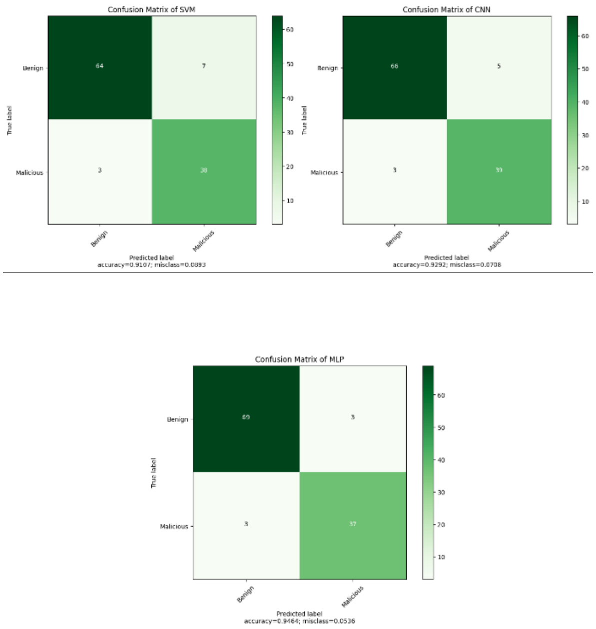


Figure 9. Confusion Matrix of (a) SVM (b) CNN (c) MLP

The confusion matrix (Figure 9) is crucial for evaluating SVM, CNN, and MLP in breast cancer diagnosis, comparing model predictions with ground truth labels (TP, TN, FP, FN). Analyzing these matrices reveals models' efficacy in distinguishing benign from malignant instances, crucial for clinical decisions and advancing early diagnosis methods for breast cancer.

Conclusion And Future Scope

Our research suggests that combining the Relief Algorithm with advanced AI models like MLP, CNN, and SVM can significantly enhance early breast cancer detection. By employing multiple machine learning algorithms, meticulous data preprocessing, and careful feature selection, we notably improved diagnostic accuracy, precision, specificity, and sensitivity. MLP achieved 94.68% accuracy, CNN 92.76%, and SVM 91.42%, potentially improving patient outcomes and survival rates. This study contributes to the growing body of literature supporting AI in medical diagnostics, paving the way for further research into diverse algorithms, feature selection methods, and larger datasets. Early diagnosis remains crucial in breast cancer therapy, and with AI and ML advancements, more accurate and accessible diagnostic tools are on the horizon, promising a brighter future for breast cancer management and treatment. Additionally, combining the Relief Algorithm with other AI techniques like decision trees and neural networks holds promise for even more accurate diagnosis, potentially unveiling insights into cancer biology and enabling earlier interventions for improved patient outcomes.

References

[1] Timothy B. Fisher, et al., (2024), "Digital image analysis and machine learning-assisted prediction of neoadjuvant chemotherapy response in triple-negative breast cancer", Breast Cancer Res 26, 12, DOI: 10.1186/s13058-023-01752-y

[2] Akhil Kumar Das, et al., (2024), "Machine Learning based Intelligent System for Breast Cancer Prediction (MLISBCP)", ESWA, Volume 242, 122673, ISSN 0957-4174, DOI: 10.1016/j.eswa.2023.122673

[3] Luyang Luo, et al., (2024), "Deep Learning in Breast Cancer Imaging: A Decade of Progress and Future Directions", RBME, DOI: 10.1109/RBME.2024.3357877

[4] Arslan Khalid, et al., (2023), "Breast Cancer Detection and Prevention Using Machine Learning", Diagnostics 13, no. 19: 3113, DOI: 10.3390/diagnostics13193113

[5] Ramdas Kapila, et al., (2023), "An efficient ensemble-based Machine Learning for breast cancer detection", BSPC, Volume 86, Part B, 105269, ISSN 1746-8094, DOI: 10.1016/j.bspc.2023.105269

[6] Mohamed Ebrahim, et al., (2023), "Accuracy Assessment of Machine Learning Algorithms Used to Predict Breast Cancer", Data 8, no. 2: 35, DOI: 10.3390/data8020035

[7] Elham Nazari et al., (2023), "Breast cancer prediction using different machine learning methods applying multi factors", JCRCO 149, 17133–17146, DOI: 10.1007/s00432-023-05388-5

Adyasha Sahu, et al., (2024), "An efficient deep learning scheme to detect breast cancer using mammogram and ultrasound breast images", BSPC, Volume 87, Part A, 105377, ISSN 1746-8094, DOI 10.1016/j.bspc.2023.105377.

[9] Chao Cong, et al., (2024), "MRI-Based Breast Cancer Classification and Localization by Multiparametric Feature Extraction and Combination Using Deep Learning", JMRI, 59: 148-161, DOI: 10.1002/jmri.28713

[10] Shallu Sharma, et al., (2024), "An ensemble of deep CNNs for automatic grading of breast cancer in digital pathology images", NCA, DOI: 10.1007/s00521-023-09368-1

[11] Humayun Mamoona, et al., (2023), "Framework for Detecting Breast Cancer Risk Presence Using Deep Learning", Electronics 12, no. 2: 403, DOI: 10.3390/electronics12020403

[12] Surjeet Dalal, et al., (2023), "A hybrid machine learning model for timely prediction of breast cancer", IJM, SSC, Vol. 14, No. 04, 2341023, Special Issue Section, DOI: 10.1142/S1793962323410234

[13] S. Sharmin, et al., (2023), "A Hybrid Dependable Deep Feature Extraction and Ensemble-Based Machine Learning Approach for Breast Cancer Detection", IEEE Access, vol. 11, pp. 87694-87708, 2023, DOI: 10.1109/ACCESS.2023.3304628.

[14] Raafat M. Munshi, et al., (2024), "A novel approach for breast cancer detection using optimized ensemble learning framework and XAI", IMAVC, Volume 142, 04910, ISSN 0262-8856, DOI: 10.1016/j.imavis.2024.104910

[15] Hua Chen, et al., (2023), "Classification Prediction of Breast Cancer Based on Machine Learning", CIN, vol. 2023, Article ID 6530719, 9 pages, DOI: 10.1155/2023/6530719

[16] https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset/data Accessed on 19th December 2023