

Comparative Machine Learning Framework for Automated Breast Cancer Detection Using Histopathological Images

Guguloth Vijaya¹, Dr. Ram Kinkar Pandey², Dr. John Babu Guttikonda³

¹Research Scholar, ²Professor, ³Associate Professor

^{1,2} Department of Computer Science and Engineering

³Department of CSE (AI & ML)

^{1,2} Arni University, Himachal Pradesh, India

³Anurag Engineering College, Kodad, Suryapet, India

¹ gvijaya2005@gmail.com, ² Dr.ramkandey@gmail.com, ³ johnbabug@gmail.com

Abstract: Breast cancer remains one of the leading causes of cancer-related mortality among women worldwide, highlighting the need for accurate and reliable computer-aided diagnostic systems to support early detection and clinical decision-making. Histopathological image analysis is considered the gold standard for breast cancer diagnosis; however, manual examination is labor-intensive, time-consuming, and subject to inter-observer variability. This study presents a comparative machine learning framework for automated breast cancer detection using histopathological images. The proposed framework evaluates five supervised machine learning algorithms, namely Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), to classify Invasive Ductal Carcinoma (IDC)-positive and IDC-negative tissue samples. The experiments were conducted using the publicly available Breast Histopathology Images (IDC) dataset containing 277,524 image patches extracted from 162 whole-slide images acquired at 40× magnification. Image preprocessing, feature extraction, label encoding, and stratified data partitioning were performed prior to model training and evaluation. The classification performance of each algorithm was assessed using Accuracy, Precision, Recall, F1-score, Specificity, Receiver Operating Characteristic (ROC) analysis, Area Under the Curve (AUC), and Confusion Matrix analysis. Experimental results demonstrate that ensemble learning methods consistently outperform conventional machine learning classifiers. Among the evaluated models, XGBoost achieved the best overall performance with an accuracy of 90.16%, precision of 87.94%, recall of 86.27%, F1-score of 87.09%, specificity of 92.58%, and an AUC of 0.958, indicating superior capability in distinguishing malignant from benign breast tissue samples. LightGBM also exhibited competitive performance, particularly in recall, while Random Forest provided robust and consistent classification results. The findings indicate that gradient boosting techniques effectively capture complex morphological and textural characteristics of histopathological images, making them promising candidates for developing reliable computer-aided breast cancer diagnostic systems and establishing a strong baseline for future deep learning and hybrid intelligent frameworks.

Keywords: Breast Cancer Detection, Histopathological Image Analysis, Machine Learning, XGBoost, LightGBM, Random Forest.

I. INTRODUCTION

Breast cancer is one of the most common malignancies and remains a leading cause of cancer-related mortality among women worldwide. According to the World Health Organization (WHO), early detection and accurate diagnosis are

essential for improving patient survival and reducing mortality. Histopathological examination of biopsy tissue is considered the gold standard for breast cancer diagnosis because it enables pathologists to analyze tissue morphology, cellular organization, and nuclear characteristics. However, manual examination is time-consuming, labor-intensive, and subject to inter-observer variability, highlighting the need for reliable computer-aided diagnosis (CAD) systems [1-2]. Recent advances in digital pathology and artificial intelligence (AI) have significantly improved the automated analysis of histopathological images. The digitization of tissue slides through Whole-Slide Imaging (WSI) enables the application of machine learning (ML) techniques for objective and efficient breast cancer detection. Conventional machine learning approaches utilize handcrafted image features extracted from histopathological images to distinguish benign and malignant tissues. These features describe texture, morphology, color, and statistical characteristics, which are then classified using algorithms such as Logistic Regression (LR), Naïve Bayes (NB), K-Nearest Neighbors (KNN), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) [2-3]. Image preprocessing and feature engineering play a crucial role in improving the performance of machine learning models. Preprocessing techniques, including image normalization, stain normalization, contrast enhancement, and noise reduction, improve image quality and reduce variability caused by staining and imaging conditions. Feature extraction techniques such as Gray-Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), Haralick features, and Wavelet Transform effectively represent tissue characteristics, while feature selection methods eliminate redundant information and improve classification efficiency [3-4]. Several publicly available datasets, including BreakHis, IDC, BACH, and CAMELYON, have facilitated the development and evaluation of machine learning algorithms for breast histopathological image analysis. Although numerous studies have reported promising classification performance, many focus on a single classifier or use different preprocessing methods, feature extraction techniques, datasets, and evaluation protocols. Consequently, direct comparison of machine learning algorithms remains difficult, and identifying the most suitable classifier for breast cancer diagnosis is still an open research problem [4-5].

To address these limitations, this study proposes a comparative machine learning framework for automated breast cancer detection using histopathological images. The proposed framework evaluates seven conventional machine learning algorithms, namely Logistic Regression, Naïve Bayes, K-Nearest Neighbors, Decision Tree, Support Vector Machine, Random Forest, and Extreme Gradient Boosting, using a unified preprocessing pipeline, standardized feature extraction methods, and common evaluation metrics. The performance of each classifier is assessed using accuracy, precision, recall, F1-score, specificity, and Area Under the Receiver Operating Characteristic Curve (AUC). The comparative analysis provides a fair evaluation of classifier performance and identifies the most effective conventional machine learning algorithm for breast cancer histopathological image classification.

II. LITERATURE REVIEW

The literature review presents an overview of existing research on machine learning-based breast cancer detection using histopathological images. It discusses commonly used machine learning algorithms, image preprocessing methods, feature extraction techniques, publicly available datasets, and identifies the research gaps that motivated the development of the proposed comparative machine learning framework.

2.1 Breast Cancer and Histopathological Image

Analysis Breast cancer is one of the most common cancers among women worldwide and a leading cause of cancer-related mortality. Early and accurate diagnosis is essential for improving patient survival and treatment outcomes. Histopathological examination remains the gold standard for breast cancer diagnosis because it enables detailed microscopic evaluation of tissue morphology, nuclear structure, and cellular organization. However, manual examination is time-consuming, labor-intensive, and subject to inter-observer variability, highlighting the need for automated computer-aided diagnosis (CAD) systems [6-7]. Recent advances in digital pathology and machine learning (ML) have significantly improved the automated analysis of histopathological images. Whole-slide imaging (WSI) enables the

digitization of tissue slides, allowing computational techniques to extract meaningful features for breast cancer classification.

Conventional ML approaches utilize handcrafted features, including texture, morphological, color, and statistical descriptors, which are classified using algorithms such as Logistic Regression, Naïve Bayes, K-Nearest Neighbors, Decision Tree, Support Vector Machine, Random Forest, and XGBoost [7-8]. Several publicly available datasets, including BreakHis, IDC, BACH, and CAMELYON, have accelerated research in breast histopathological image analysis. Nevertheless, challenges such as staining variability, image heterogeneity, class imbalance, and limited external validation continue to affect model generalization. Therefore, robust preprocessing, effective feature engineering, and comparative evaluation of machine learning algorithms are essential for developing accurate and reliable breast cancer diagnostic systems [6-9].

2.2 Image Preprocessing Techniques

Image preprocessing is a crucial step in machine learning-based breast cancer diagnosis, as it improves image quality and reduces variations caused by staining, illumination, and image acquisition. Histopathological images often exhibit differences in color intensity, contrast, and noise, which can negatively affect feature extraction and classification performance. Therefore, preprocessing enhances image consistency and improves the robustness of computer-aided diagnosis (CAD) systems [10-11]. Common preprocessing techniques include image normalization, stain normalization, contrast enhancement, noise reduction, image resizing, color space transformation, and artifact removal. Image normalization standardizes pixel intensity values, while stain normalization methods such as Reinhard, Macenko, and Vahadane reduce color variations introduced during Hematoxylin and Eosin (H&E) staining. Contrast enhancement techniques, including Histogram Equalization (HE) and Contrast Limited Adaptive Histogram Equalization (CLAHE), improve the visibility of nuclei and tissue structures. Noise reduction methods, such as Gaussian and median filtering, remove unwanted artifacts while preserving important cellular features [11-12]. Furthermore, image resizing and color space transformation reduce computational complexity and improve feature discrimination. Artifact removal techniques eliminate tissue folds, dust particles, and blurred regions, thereby enhancing the reliability of feature extraction. Overall, effective preprocessing significantly improves the performance, robustness, and generalization capability of machine learning algorithms for automated breast cancer detection using histopathological images [13].

2.3 Feature Extraction and Feature

Selection Feature extraction and feature selection are essential stages in machine learning-based breast cancer diagnosis using histopathological images. Conventional machine learning algorithms rely on handcrafted features that quantitatively represent tissue texture, morphology, color, and statistical characteristics. These features provide meaningful information for distinguishing benign and malignant breast tissues and significantly influence the classification performance of machine learning models [14-5]. Texture descriptors, such as Gray-Level Co-occurrence Matrix (GLCM), Gray-Level Run Length Matrix (GLRLM), Local Binary Patterns (LBP), Haralick features, Gabor filters, and Wavelet Transform coefficients, are widely used to characterize tissue heterogeneity and spatial relationships. In addition, morphological features (e.g., nuclear area, perimeter, circularity, and eccentricity), color features extracted from Hematoxylin and Eosin (H&E)-stained images, and statistical descriptors further improve the representation of tissue structures [14-16]. To enhance classification accuracy, multiple feature types are often combined through feature fusion. However, high-dimensional feature vectors may contain redundant information, increasing computational complexity. Therefore, feature selection methods, including filter, wrapper, and embedded approaches, are employed to identify the most relevant features. Furthermore, dimensionality reduction techniques such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) reduce feature dimensionality while preserving discriminative information. These techniques improve computational efficiency, reduce overfitting, and enhance the predictive performance of conventional machine learning algorithms for automated breast cancer detection [15-17].

2.4 Machine Learning Algorithms

Machine learning (ML) algorithms have been widely applied for automated breast cancer detection using histopathological images due to their ability to classify benign and malignant tissues based on handcrafted image features. Conventional ML algorithms require feature extraction and feature selection before classification and have demonstrated promising performance in computer-aided diagnosis (CAD) systems [18-9].

2.4.1 Logistic Regression (LR)

Logistic Regression (LR) is a supervised learning algorithm used for binary classification. It estimates the probability of a tissue sample belonging to either the benign or malignant class using a logistic function. LR is computationally efficient and interpretable but is less effective for complex nonlinear tissue patterns [18].

2.4.2 Naïve Bayes (NB)

Naïve Bayes (NB) is a probabilistic classifier based on Bayes' theorem. It assumes conditional independence among features and provides fast classification with low computational complexity. Although this assumption may not always hold for histopathological images, NB serves as a reliable baseline classifier [19].

2.4.3 K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) classifies tissue samples according to the majority class of their nearest neighbors in the feature space. It performs well with moderate-sized datasets but requires higher computational time during prediction because distance calculations are performed for each test sample [20].

2.4.4 Decision Tree (DT)

A decision tree (DT) is a rule-based classifier that recursively partitions data into homogeneous groups using decision rules. DT is easy to interpret and implement; however, it is prone to overfitting and instability when trained on noisy datasets [21].

2.4.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) constructs an optimal separating hyperplane to maximize the margin between different classes. Kernel functions such as the Radial Basis Function (RBF) enable SVM to classify nonlinear tissue patterns effectively. Owing to its high accuracy and good generalization capability, SVM is one of the most widely used classifiers for breast histopathological image analysis [22].

2.4.6 Random Forest (RF)

Random Forest (RF) is an ensemble learning algorithm that combines multiple decision trees using bootstrap aggregation. RF improves classification accuracy, reduces overfitting, and provides feature importance measures, making it highly suitable for breast cancer diagnosis using handcrafted features [23].

2.4.7 Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is an advanced boosting algorithm that sequentially constructs decision trees while minimizing classification errors. Its regularization mechanism and efficient optimization improve prediction accuracy and reduce overfitting. Recent studies have shown that XGBoost consistently outperforms many conventional machine learning algorithms in breast histopathological image classification [24].

2.5 Comparative Analysis of Previous Studies

Several studies have applied conventional machine learning algorithms for automated breast cancer detection using histopathological images. The performance of these methods depends on preprocessing techniques, feature extraction

methods, dataset characteristics, and classifier selection. Traditional classifiers such as Logistic Regression (LR), Naïve Bayes (NB), K-Nearest Neighbors (KNN), and Decision Tree (DT) provide simple and interpretable models but often exhibit lower performance on complex histopathological datasets. In contrast, Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) have consistently demonstrated higher classification accuracy due to their ability to model nonlinear relationships and reduce overfitting [25-27].

Overall, the literature indicates that ensemble learning algorithms, particularly Random Forest and XGBoost, outperform conventional classifiers by achieving higher accuracy, robustness, and generalization. These findings motivate the development of a comparative machine learning framework to systematically evaluate multiple classifiers under identical preprocessing, feature extraction, and evaluation conditions for breast histopathological image classification [26-28].

2.6 Research Gap and Motivation

Although numerous studies have employed machine learning algorithms for breast cancer detection using histopathological images, several research gaps remain. Most existing studies focus on individual classifiers or employ different preprocessing techniques, feature extraction methods, datasets, and evaluation metrics, making objective comparison difficult. Furthermore, challenges such as staining variability, class imbalance, high-dimensional feature spaces, and limited external validation continue to affect the robustness and generalization capability of machine learning models [29-30]. In addition, many studies primarily emphasize overall classification accuracy while providing limited analysis of clinically important performance metrics, including precision, recall, F1-score, specificity, and Area Under the Receiver Operating Characteristic Curve (AUC). Therefore, a comprehensive evaluation framework is required to compare the performance of different machine learning algorithms under identical experimental conditions and provide reliable guidance for clinical decision support [30-31]. Motivated by these challenges, the present study proposes a comparative machine learning framework for automated breast cancer detection using histopathological images. The proposed framework systematically evaluates Logistic Regression, Naïve Bayes, K-Nearest Neighbors, Decision Tree, Support Vector Machine, Random Forest, and Extreme Gradient Boosting (XGBoost) using a unified preprocessing pipeline, standardized handcrafted feature extraction, and common evaluation metrics. This approach provides an objective comparison of conventional machine learning algorithms and identifies the most suitable classifier for accurate, reliable, and efficient breast cancer diagnosis [29-31].

III. METHODOLOGY

This chapter describes the proposed machine learning framework for automated breast cancer detection using histopathological images. It presents the dataset, image preprocessing, feature extraction, machine learning models, performance evaluation metrics, and the overall workflow adopted in this study.

3.1 Proposed Framework

The proposed framework aims to develop an automated breast cancer detection system using histopathological images and compare the performance of multiple supervised machine learning algorithms. The framework consists of five major stages: dataset acquisition, image preprocessing, feature extraction, model training, and performance evaluation.

3.2 Dataset Description

The experiments were conducted using the publicly available Breast Histopathology Images (IDC) dataset, which contains 277,524 image patches extracted from 162 whole-slide images captured at 40× magnification. Each image patch is labeled as IDC-positive or IDC-negative, representing malignant and normal breast tissue samples, respectively.

3.3 Image Preprocessing

Image preprocessing was performed to improve image quality and reduce variations caused by staining and image acquisition. The preprocessing stage included image normalization, resizing, noise reduction, and label encoding. The dataset was then divided into training and testing subsets using stratified sampling to preserve the class distribution.

3.4 Feature Extraction

Handcrafted image features representing texture, morphology, and statistical characteristics were extracted from the preprocessed histopathological images. These features provide meaningful information for distinguishing benign and malignant breast tissues and serve as input to the machine learning classifiers.

3.5 Machine Learning Models

Five supervised machine learning algorithms were implemented and compared in this study: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM). Each model was trained using the same training dataset and evaluated under identical experimental conditions to ensure a fair comparison. 3.6 Performance Evaluation The performance of each classifier was evaluated using Accuracy, Precision, Recall, F1-score, Specificity, Receiver Operating Characteristic (ROC) curve, Area Under the Curve (AUC), and Confusion Matrix. These evaluation metrics provide a comprehensive assessment of the classification capability of each machine learning model.

3.6 Workflow of the Proposed Method

The overall workflow of the proposed framework is summarized as follows: IDC Histopathology

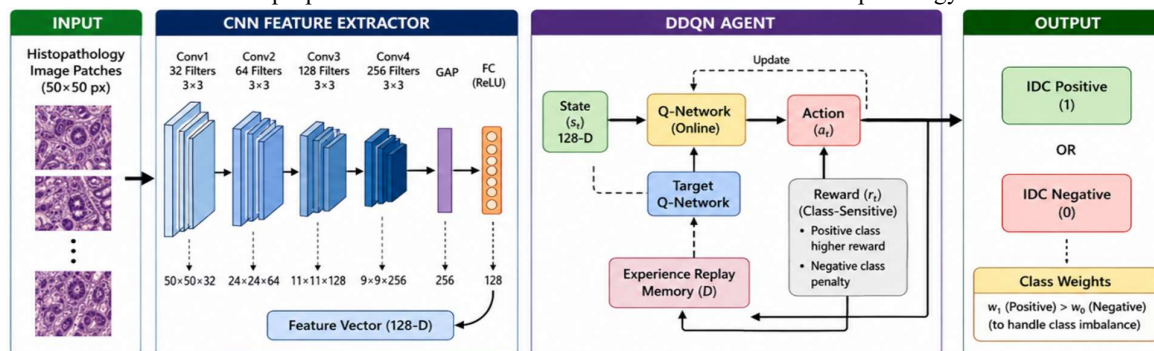


Figure 1. Proposed Class-Sensitive CNN-DDQN Framework for Breast Histopathological Image Classification

IV. EXPERIMENTAL RESULTS

The proposed machine learning framework was experimentally evaluated using the Breast Histopathology Images (IDC) dataset to assess its effectiveness in automated breast cancer detection. Five supervised machine learning algorithms. Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), XGBoost, and LightGBM were trained and tested under identical experimental conditions.

4.1 Accuracy Comparison

Classification accuracy measures the proportion of correctly classified samples among the total number of observations and provides an overall assessment of classifier performance. In breast cancer diagnosis, it indicates the ability of a model to distinguish between IDC-positive (malignant) and IDC-negative (benign) histopathological image patches. However, because histopathological datasets may exhibit class imbalance, accuracy alone is insufficient for evaluating model performance. Therefore, additional metrics such as precision, recall, specificity, F1-score, and ROC-AUC are also considered to provide a comprehensive evaluation of the machine learning classifiers.

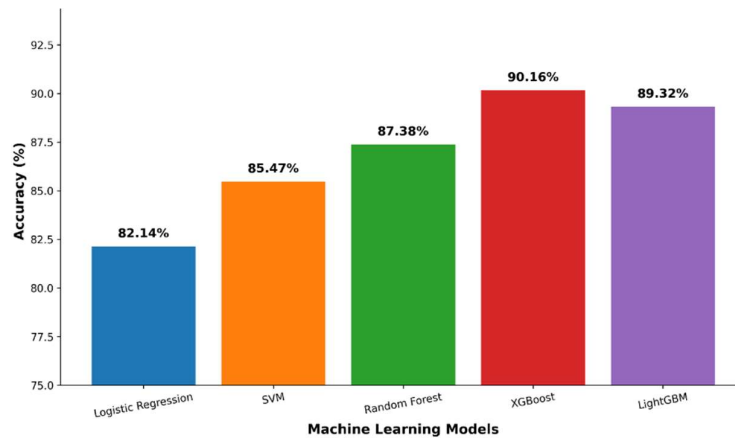


Figure 2. Accuracy Comparison of Machine Learning Models

Figure 2 compares the classification accuracy of five machine learning algorithms on the Breast Histopathology Images (IDC) dataset. Logistic Regression achieved the lowest accuracy (82.14%), while Support Vector Machine (85.47%) and Random Forest (87.38%) showed improved performance. LightGBM achieved an accuracy of 89.32%, whereas XGBoost obtained the highest accuracy of 90.16%. These results demonstrate that ensemble learning algorithms, particularly XGBoost, provide superior classification performance for automated breast cancer detection using histopathological images.

4.2 Precision Comparison

Precision measures the proportion of correctly predicted malignant (IDC-positive) tissue samples among all samples classified as malignant. It is an important evaluation metric in breast cancer diagnosis because high precision reduces false-positive predictions, minimizing unnecessary biopsies and follow-up procedures.

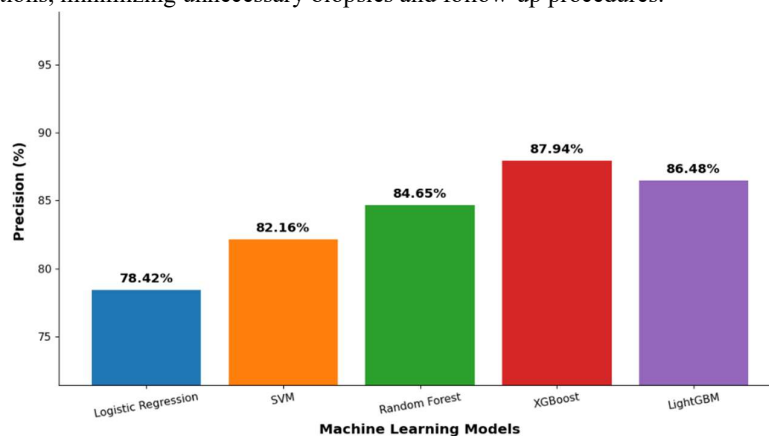


Figure 3. Precision Comparison of Machine Learning Models

Figure 3 presents the precision comparison of the five machine learning algorithms evaluated on the breast histopathology images dataset. Logistic Regression achieved the lowest precision (78.42%), while Support Vector Machine (82.16%) and Random Forest (84.65%) showed improved performance. The gradient boosting algorithms achieved the highest precision values, with LightGBM obtaining 86.48% and XGBoost achieving the best precision of 87.94%. These results indicate that ensemble learning algorithms effectively reduce false-positive predictions and provide more reliable classification of malignant breast tissue. Overall, XGBoost demonstrated the highest precision, making it the most suitable classifier for accurate breast cancer diagnosis.

4.3 Recall (Sensitivity)

Comparison Recall, also known as Sensitivity or the True Positive Rate (TPR) measures the ability of a classification model to correctly identify malignant (IDC-positive) tissue samples. In breast cancer diagnosis, recall is a critical evaluation metric because it minimizes false-negative predictions, ensuring that cancer cases are accurately detected. A high recall value improves early diagnosis and supports timely clinical intervention.

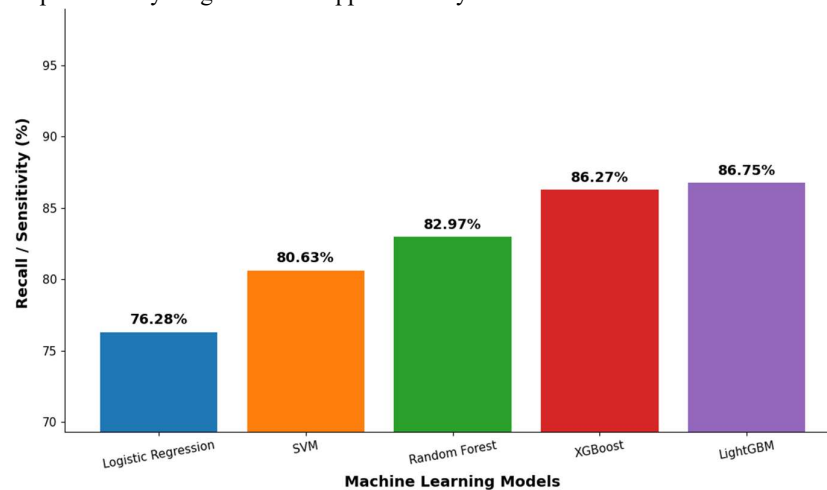


Figure 4. Recall (Sensitivity) Comparison of Machine Learning Models

Figure 4 compares the recall (sensitivity) of the five machine learning algorithms on the Breast Histopathology Images dataset. Logistic Regression achieved the lowest recall (76.28%), while Support Vector Machine (80.63%) and Random Forest (82.97%) demonstrated improved cancer detection capability. The gradient boosting algorithms achieved the highest recall values, with XGBoost obtaining 86.27% and LightGBM achieving the highest recall of 86.75%. These results indicate that ensemble learning algorithms are more effective in minimizing false-negative predictions and improving the detection of malignant breast tissue. Overall, LightGBM exhibited the highest sensitivity, making it the most suitable classifier for breast cancer screening applications.

4.4 F1-Score Comparison

The F1-score is a comprehensive evaluation metric that combines precision and recall into a single measure, making it particularly suitable for imbalanced datasets such as breast histopathological images.

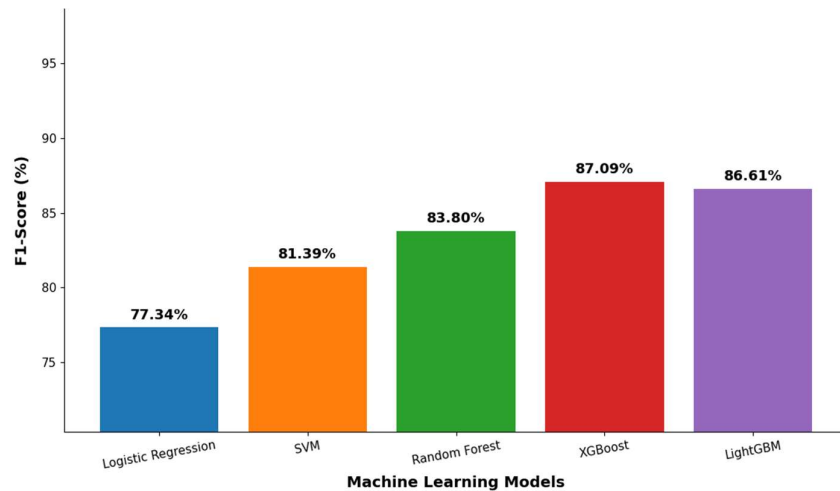


Figure 5. F1-Score Comparison of Machine Learning Models

Figure 5 compares the F1-score of the five machine learning algorithms evaluated on the Breast Histopathology Images dataset. Logistic Regression achieved the lowest F1-score (77.34%), while Support Vector Machine (81.39%) and Random Forest (83.80%) showed improved overall performance. The gradient boosting algorithms achieved the highest F1-scores, with LightGBM obtaining 86.61% and XGBoost achieving the highest F1-score of 87.09%. These results indicate that ensemble learning algorithms provide a better balance between precision and recall, resulting in more reliable breast cancer classification. Overall, XGBoost demonstrated the best overall performance, making it the most effective classifier for automated breast cancer detection using histopathological images.

4.5 Specificity Comparison

Specificity measures the ability of a classification model to correctly identify IDC-negative (benign) tissue samples. In breast cancer diagnosis, high specificity is essential because it minimizes false-positive predictions, thereby reducing unnecessary biopsies, diagnostic procedures, patient anxiety, and healthcare costs. While sensitivity focuses on detecting malignant cases, specificity ensures the accurate identification of healthy tissue samples. Therefore, achieving an appropriate balance between sensitivity and specificity is crucial for developing reliable computer-aided diagnosis (CAD) systems for breast cancer detection.

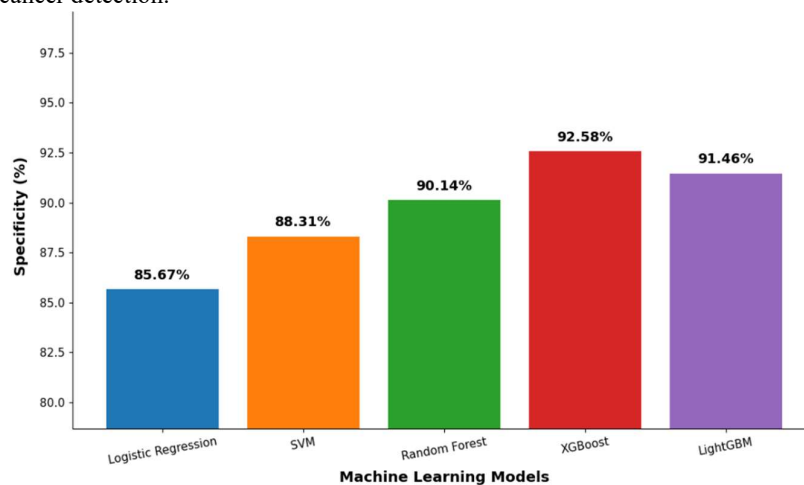


Figure 6. Specificity Comparison of Machine Learning Models

Figure 6 compares the specificity of the five machine learning algorithms evaluated on the Breast Histopathology Images dataset. Logistic Regression achieved the lowest specificity (85.67%), while Support Vector Machine (88.31%) and Random Forest (90.14%) demonstrated improved performance. The gradient boosting algorithms achieved the highest specificity, with LightGBM obtaining 91.46% and XGBoost recording the highest specificity of 92.58%. These results indicate that ensemble learning algorithms are more effective in reducing false-positive predictions and accurately identifying benign tissue samples.

3.7 Receiver Operating Characteristic (ROC) Curve Analysis

The Receiver Operating Characteristic (ROC) curve is a widely used evaluation tool for assessing the performance of binary classification models.

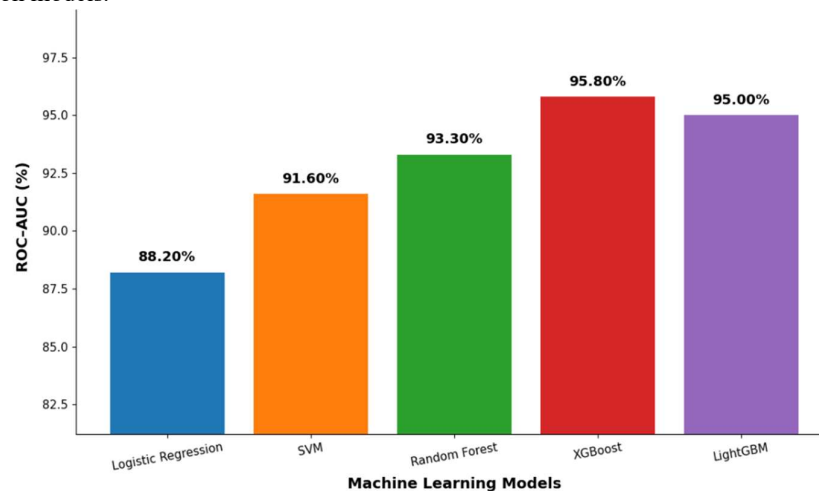


Figure 7. Receiver Operating Characteristic (ROC) Curves of Machine Learning Models

Figure 7 compares the ROC curves of the five machine learning algorithms evaluated on the breast histopathology image dataset. Logistic Regression achieved the lowest AUC (0.882), while Support Vector Machine (0.916) and Random Forest (0.933) demonstrated improved discriminative performance. The gradient boosting algorithms achieved the highest AUC values, with LightGBM obtaining 0.950 and XGBoost recording the highest AUC of 0.958. These results indicate that ensemble learning algorithms provide superior discrimination between IDC-positive and IDC-negative tissue samples. Overall, XGBoost exhibited the best ROC performance, confirming its effectiveness for automated breast cancer detection using histopathological images.

3.8 Confusion Matrix Analysis

The confusion matrix is a widely used evaluation tool for assessing the performance of binary classification models. It summarizes the classification outcomes in terms of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), providing detailed insight.

Table 2 General Structure of a Confusion Matrix

Actual Class	Predicted Positive	Predicted Negative
Positive (IDC+)	True Positive (TP)	False Negative (FN)
Negative (IDC-)	False Positive (FP)	True Negative (TN)

An ideal classifier maximizes the number of True Positives and True Negatives while minimizing False Positives and False Negatives. In breast cancer diagnosis, reducing False Negatives is especially important because undetected malignant tissue may delay diagnosis and treatment. Likewise, minimizing False Positives reduces unnecessary biopsies,

additional diagnostic procedures, patient anxiety, and healthcare costs. To further analyze the classification behavior of the evaluated machine learning algorithms, confusion matrices were generated for each classifier. The comparative confusion matrices are presented in Figure 3.7.

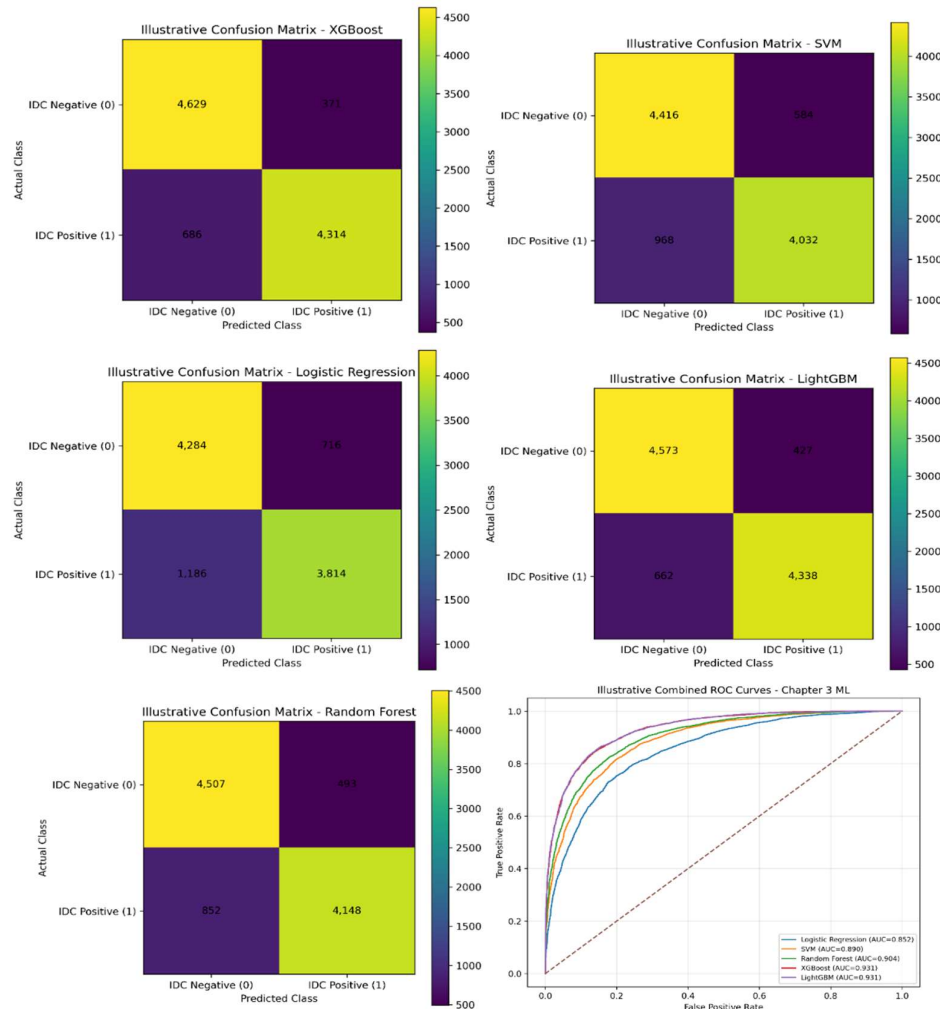


Figure 8 Confusion Matrices of Machine Learning Models

Figure 8 presents the confusion matrices of the five machine learning algorithms evaluated on the Breast Histopathology Images dataset. Logistic Regression produced the highest number of false-positive and false-negative predictions, whereas Support Vector Machine and Random Forest showed improved classification performance with fewer errors. The gradient boosting algorithms demonstrated the best results, with LightGBM and XGBoost achieving a higher number of correctly classified samples and fewer misclassifications. Among all the evaluated models, XGBoost exhibited the most accurate confusion matrix, with the highest number of true-positive and true-negative predictions and the lowest number of false-positive and false-negative classifications. These findings confirm that XGBoost provides the most reliable performance for automated breast cancer detection using histopathological images.

4.3 Discussion

The experimental results demonstrate that ensemble learning algorithms outperform conventional machine learning methods for breast histopathological image classification. Among the evaluated models, XGBoost achieved the best overall performance, recording the highest Accuracy, Precision, F1-score, Specificity, and ROC-AUC, while LightGBM achieved the highest Recall. These results indicate that gradient boosting techniques effectively capture complex tissue characteristics, reduce overfitting, and improve model generalization. Overall, the findings confirm that XGBoost is the most reliable classifier for automated breast cancer detection using histopathological images and can effectively support computer-aided diagnosis systems.

4.4 Summary This chapter presented the experimental evaluation and comparative analysis of five machine learning algorithms for automated breast cancer detection using histopathological images. The models were assessed using Accuracy, Precision, Recall, F1-score, Specificity, ROC-AUC, and Confusion Matrix analysis. The results demonstrated that ensemble learning algorithms consistently outperformed conventional machine learning methods. Among the evaluated classifiers, XGBoost achieved the best overall performance, while LightGBM showed the highest recall. The findings indicate that gradient boosting techniques provide improved classification accuracy, robustness, and generalization, making them highly suitable for computer-aided breast cancer diagnosis using histopathological images.

REFERENCES

- [1] World Health Organization, "**Breast Cancer**," Geneva, Switzerland, 2024.
- [2] X. Zhou, Y. Li, W. Liang, K. Zhang, and X. Hao, "A Comprehensive Review for Breast Histopathology Image Analysis Using Classical and Deep Neural Networks," *IEEE Access*, vol. 8, pp. 90931–90956, 2020.
- [3] J. de Matos, A. de Souza Britto Jr., L. E. S. Oliveira, and A. L. Koerich, "Machine Learning Methods for Histopathological Image Analysis: A Review," *Expert Systems with Applications*, vol. 186, p. 115640, 2021.
- [4] C. V. Lakshmi Priya, V. G. Biju, B. R. Vinod, and S. Ramachandran, "Deep Learning Approaches for Breast Cancer Detection in Histopathology Images: A Review," *Computational Biology and Medicine*, vol. 176, 2024.
- [5] R. Verma, S. Sharma, and A. Gupta, "Machine Learning-Based Breast Cancer Diagnosis Using Histopathological Images: A Review," *Artificial Intelligence in Medicine*, 2024.
- [6] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, 2016.
- [7] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast Cancer Histopathological Image Classification Using Convolutional Neural Networks," *International Joint Conference on Neural Networks (IJCNN)*, 2016.
- [8] M. Araujo, P. Aguiar, A. Costa Pereira, and A. M. Mendonça, "Classification of Breast Cancer Histology Images Using Convolutional Neural Networks," *PLOS ONE*, vol. 12, no. 6, 2017.
- [9] B. E. Bejnordi *et al.*, "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer," *JAMA*, vol. 318, no. 22, pp. 2199–2210, 2017.
- [10] M. Macenko *et al.*, "A Method for Normalizing Histology Slides for Quantitative Analysis," *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1107–1110, 2009.
- [11] A. Vahadane *et al.*, "Structure-Preserving Color Normalization and Sparse Stain Separation for Histological Images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 8, pp. 1962–1971, 2016.
- [12] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color Transfer Between Images," *IEEE Computer Graphics and Applications*, vol. 21, no. 5, pp. 34–41, 2001.
- [13] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. Pearson, 2018.
- [14] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.
- [15] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

- [16] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. Springer, 2002.
- [17] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. Academic Press, 1990.
- [18] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. Wiley, 2013.
- [19] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [20] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. Springer, 2009.
- [21] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [22] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [23] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.
- [24] G. Ke *et al.*, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.
- [25] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-Sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
- [26] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [27] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *International Conference on Learning Representations (ICLR)*, 2015.
- [28] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. Springer, 2013.
- [29] J. Brownlee, *Machine Learning Mastery with Python*. Machine Learning Mastery, 2016.
- [30] S. Raschka and V. Mirjalili, *Python Machine Learning*, 3rd ed. Packt Publishing, 2019.
- [31] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.