IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 26 June 2023, accepted 3 August 2023, date of publication 14 August 2023, date of current version 22 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3304628

RESEARCH ARTICLE

A Hybrid Dependable Deep Feature Extraction and Ensemble-Based Machine Learning Approach for Breast Cancer Detection

SELINA SHARMIN[1], (Member, IEEE), TANVIR AHAMMAD[1], MD. ALAMIN TALUKDER[1], AND PARTHO GHOSE[2]

[1]Department of Computer Science and Engineering, Jagannath University, Dhaka 1100, Bangladesh
[2]Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka 1216, Bangladesh

Corresponding authors: Tanvir Ahammad (tanvir@cse.jnu.ac.bd), Md. Alamin Talukder (alamintalukder.cse.jnu@gmail.com), and Selina Sharmin (selina@cse.jnu.ac.bd)

This work was supported by the UGC Supported Jagannath University Research Grant.

ABSTRACT Breast cancer is a prevalent and life-threatening disease that requires effective detection and diagnosis methods to improve patient outcomes. Deep learning (DL) and machine learning (ML) techniques have emerged as powerful tools in breast cancer detection, offering benefits such as improved accuracy and efficiency. However, existing methods have scalability and performance limitations, emphasizing the need for further research. In this paper, we propose a hybrid dependable breast cancer detection approach that combines the power of DL using a pre-trained ResNet50V2 model and ensemble-based ML methods. The integration of DL enables the approach to learn and extract hidden patterns from complex breast cancer images, while ML algorithms contribute interpretability and generalization capabilities. We conducted extensive experiments using a breast histopathology image-based publicly available Invasive Ductal Carcinoma (IDC) dataset comprising samples of different sizes. The results obtained from our rigorous experiments provide compelling evidence for our hybrid model's robustness and high performance. We achieved a higher accuracy rate of 95%, precision of 94.86%, recall of 94.32%, and F1 score of 94.57% compared to state-of-the-art models. We also identified Light Boosting Classifier (LGB) as the most suitable ML model in conjunction with the ResNet50V2 architecture. The results of this research offer significant contributions to breast cancer detection through an innovative approach, comprehensive performance analysis, and dependable assessment. Moreover, it has the potential to assist medical professionals in making informed decisions, improving patient care, and enhancing outcomes for breast cancer patients.

INDEX TERMS Breast cancer, breast histopathology image, deep learning, machine learning, transfer learning.

I. INTRODUCTION

Breast cancer is a prevalent and potentially life-threatening disease affecting many individuals worldwide [1]. It occurs when cells in the breast tissue grow abnormally, resulting in the development of tumors. According to global health statistics, breast cancer is women's most frequently diagnosed cancer. It accounts for many cancer-related deaths worldwide [2]. The mortality rates associated with breast cancer underscore the critical need for effective detection and diagnosis methods [3]. Early detection is crucial in improving patient outcomes and reducing mortality rates [4]. Therefore, developing accurate and efficient breast cancer detection approaches is paramount in medical research.

Deep learning (DL) and Machine learning (ML) techniques have emerged as powerful tools in breast cancer detection in recent years [5]. DL employs artificial neural networks with multiple layers to automatically learn complex hierarchical data representations [6]. Transfer learning (TL), a fundamental concept in DL, revolutionizes the field by

The associate editor coordinating the review of this manuscript and approving it for publication was Szidonia Lefkovits.

using pre-trained models trained on large datasets for specific tasks [7]. These pre-trained models have learned intricate and meaningful features from extensive data, making them valuable resources for feature extraction in various domains, including breast cancer detection [8]. By leveraging the knowledge encoded within these pre-trained models, we can expedite the training process, enhance model performance, and overcome the limitations of insufficient data. In addition, these pre-trained models empower us to focus on fine-tuning the model to the specific task at hand, enabling more efficient and accurate detection of breast cancer.

ML, on the other hand, encompasses a range of algorithms that can analyze and interpret data to identify patterns and make predictions [9], [10]. These algorithms can learn from labeled datasets to identify patterns and make precise predictions using new and unseen data [11]. DL and ML algorithms offer numerous benefits in breast cancer detection [12]. These approaches have the potential for a robust framework to improve the accuracy and efficiency of diagnostic processes, assist healthcare professionals in making informed decisions, and enhance patient care [13], [14], [15].

Feature extraction is a fundamental process in breast cancer image analysis. It involves identifying and extracting discriminative patterns and structures from medical images to represent complex data more informally [16]. By extracting relevant features, such as statistical measures, texture characteristics, or shape properties, the hidden information within the images can be revealed [10]. This process is essential as it enables the detection of image-based biomarkers, facilitates tumor characterization, and aids in distinguishing between healthy and cancerous tissues [17]. Accurate feature extraction is pivotal in improving breast cancer diagnosis and treatment planning accuracy, efficiency, and reliability.

Several influential works have significantly contributed to the literature on breast cancer detection [18], [19], [20], [21], [22]. However, these works also have their performance issues. Anjum et al. [18] proposed an efficient approach using a Histogram of Oriented Gradients (HOG) and Canny Edge features combined with Support Vector Machine (SVM) as the machine learning algorithm. While achieving a commendable accuracy rate of 94%, the method showed limitations in correctly identifying non-cancerous samples, with a relatively low accuracy rate of 22%. Kulkarni and Sundaray [23] developed a breast cancer detection model that utilized ResNet152, a deep learning architecture, for feature extraction. However, the model relied solely on fully connected layers for classification, resulting in a lower accuracy rate of 91% compared to more advanced techniques. Finally, Cruz-Roa et al. [24] proposed an automatic breast cancer detection approach using Convolutional Neural Networks (CNNs). While achieving a respectable accuracy rate of 84.23%, the model demonstrated limitations in precision, recall, and F1-score, indicating potential challenges in correctly classifying some instances of breast cancer. These limitations emphasize the ongoing need for further research and development in breast cancer detection methods to address these challenges and enhance overall performance.

Our research addresses the issues with previous methods and presents an enhanced model for breast cancer detection. We propose a breast cancer detection model that leverages the power of a pre-trained DL model, ResNet50V2, for efficient feature extraction. Then, we incorporated several ML algorithms, including Decision Tree (DT), Random Forest (RF), Extra Tree (ET), Ada Boosting (AbB), Histogram Gradient Boosting Classifier (HGBC), Gradient Boosting Classifier (GBC), Extreme Boosting Classifier (XGB), and LGB into our framework. By combining the strengths of both DL and ML, we present a hybrid dependable breast cancer prediction model which allows us to capture complex patterns and relationships in breast cancer data while benefiting from traditional machine learning algorithms' interpretability and generalization abilities.

In our experiments, we incorporated a histopathological image-based Invasive Ductal Carcinoma (IDC) dataset comprising samples of different sizes. The results obtained from our rigorous experiments provide convincing evidence for our hybrid model's robustness and high performance. Our approach achieves impressive metrics, including an accuracy rate of 95%, precision of 94.86%, recall of 94.32%, F1 score of 94.57%, and minimal error rates such as Mean Absolute Error (MAE) of 5%, Mean Squared Error (MSE) of 5%, and Root Mean Squared Error (RMSE) of 22.36%. These results demonstrate the effectiveness of our approach in accurately detecting cancer compared to existing techniques. Furthermore, our experiments suggest that the LGB is the most suitable machine learning model in conjunction with the ResNet50V2 architecture. This combination yields superior performance in breast cancer detection. Furthermore, this research has the potential to assist medical professionals in making informed decisions, improving patient care, and ultimately enhancing the outcomes for individuals affected by breast cancer.

This research offers significant contributions to breast cancer detection through a novel approach, rigorous performance analysis, and dependable assessment. Integrating DL and ensemble-based ML techniques show promising results for improved diagnostic accuracy, making it a significant contribution to breast cancer research. The research study contributes explicitly in the following ways:

1) The study introduces an innovative approach that combines deep learning and machine learning algorithms, utilizing pre-trained models for feature extraction and constructing a robust prediction model.
2) The research includes comprehensive performance analysis, employing rigorous evaluation metrics to ensure the proposed model's reliability, accuracy, and overall performance.
3) A novel aspect of this research includes a dependable assessment, evaluating the breast cancer detection system's reliability, availability, scalability, and efficiency.

4) This research can potentially improve patient care and outcomes by assisting medical professionals in making informed decisions for breast cancer detection.

The remaining sections of this paper are structured as follows: In Section II, we present a comprehensive review of related literature focusing on breast cancer detection. Section III provides a detailed explanation of our research methodology and includes a description of the dataset used. Section IV outlines the experimental setup and performance evaluation. In Section V, we conduct a thorough analysis of the dependability of our proposed approach. Section VI presents the concluding remarks and future work.

II. RELATED WORKS

Breast cancer is a primary global health concern, and the precise detection of malignant cells is essential for effective diagnosis and treatment. Histopathological image analysis and feature extraction are critical in image processing, enabling the identification of essential patterns and characteristics. Machine learning algorithms have become invaluable in this field [25], offering enhanced breast cancer detection and analysis methods. In addition, numerous related studies have contributed to the advancement of practical techniques in the domain of breast cancer detection and analysis.

Anjum et al. [18] proposed an ML approach using histopathological images of breast tissues for detecting malignant cells. They combined HOG and Canny Edge detection techniques to extract features and utilized PCA for dimensionality reduction. SVM achieved 94% accuracy in detecting cancerous cases. Mridha et al. [19] introduced a DL approach for detecting invasive ductal cancer in breast cancer images. By utilizing ANNs and CNNs, their system attained an average sensitivity rate of 81.56%. Image augmentation and regularization techniques improved the model's efficiency, with ANN achieving 82.68% accuracy and CNN achieving 86.24%. These results highlight the effectiveness of their custom architecture in classifying invasive ductal carcinoma.

In recent research, Singh and Kumar [26] focused on histopathological image analysis for breast cancer detection using a cubic Support Vector Machine (SVM), utilizing histopathology-based features. Liang et al. [27] proposed an innovative Convolutional Neural Network (CNN) architecture for breast cancer classification that incorporated 2D and 3D mammograms simultaneously, improving accuracy. Zhang et al. [28] aimed to enhance the detection of malignant lesions in breast mammograms by introducing a new method called BDR-CNN-GCN, which combined two advanced neural networks, the Graph Convolutional Network (GCN) and the Convolutional Neural Network (CNN). Additionally, Ogundokun et al. [20] proposed a medical Internet of Things (IoT) based diagnostic system that efficiently identifies malignant and benign cases of breast cancer in an IoT environment, addressing the challenge of early-stage breast cancer identification. These studies highlight various approaches in breast cancer detection and classification involving histopathological analysis, multimodal imaging, advanced neural networks, and IoT-based systems, contributing to advancements in the field.

Kulkarni and Sundaray [23] proposed a classification model for detecting breast cancer using a pre-trained Residual feature extraction model in Invasive IDC datasets. Their approach employed a fold-based training approach over the entire dataset distribution, resulting in superior performance compared to other works, achieving a reliable accuracy rate of 91%. Another study by Jin and Xie [21] presented a self-defined convolutional neural network (CNN) and compared its performance with the ResNet-50 model in breast cancer detection. Through parameter alterations and four sets of experiments, they evaluated the training and validation accuracy of the two models. The results indicated that the ResNet-50 model outperformed the self-defined CNN with an accuracy rate of 86.57%, highlighting its potential in breast cancer detection.

Yu [22] developed a CNN model for automatically diagnosing Invasive Ductal Carcinoma (IDC). The model consisted of multiple 2D-convolutional layers and max-pooling layers, producing a dichotomous output of IDC or non-IDC. The model achieved satisfactory performance on a dataset of 162 patients, with an accuracy of 88% and an Area Under Curve (AUC) score of 95%, showcasing its potential in automating IDC diagnosis.

Addressing the class imbalance challenge in breast cancer datasets, Reza and Ma [29] conducted experiments on the breast cancer HPI dataset. They employed various over-sampling and under-sampling techniques to mitigate the impact of class imbalance on CNN classifier performance. The results demonstrated that synthetic over-sampling was a practical approach, improving the performance metrics of the CNN-based classifiers. As a result, the CNN framework achieved better sensitivity, specificity, F1 score, accuracy, and balanced accuracy compared to the class-balanced dataset, with F-measure and balanced accuracy values of 84.78% and 85.48%, respectively.

Cruz-Roa et al. [24] introduced a deep learning framework based on convolutional neural networks (CNN) for the automatic detection of IDC in whole slide images (WSI). Their approach involved training the CNN on a large number of image patches from WSIs to learn a hierarchical part-based representation. The method was evaluated on a dataset of 162 IDC patients with expert-delineated ground truth. The experimental results showcased the superiority of their approach, achieving an F-measure of 71.80% and a balanced accuracy of 84.23%, surpassing approaches that utilized handcrafted image features and machine learning classifiers. Furthermore, this study emphasized the potential of their method for the automatic detection of IDC regions in WSIs.

These studies emphasized the importance of utilizing advanced techniques such as feature extraction, deep learning, and pre-trained models in breast cancer detection and analysis, as summarized in Table 1. In addition, the existing

TABLE 1. Comparison of state-of-the-art studies in breast cancer detection.

Study	Algorithm	Accuracy (%)	Approach
Anjum et al. [18]	ML	94%	HOG + Canny Edge + SVM
Mridha et al. [19]	DL	86.24%	ANN + CNN
Singh et al. [26]	SVM	92.3%	Histopathology-based features
Liang et al. [27]	CNN	97% (AUC)	2D and 3D mammograms
Zhang et al. [28]	CNN	96.10%	BDR-CNN-GCN
Ogundokun et al. [20]	CNN, ANN	98.5% (CNN), 99.2% (ANN)	Medical IoT-based system
Kulkarni et al. [23]	DL	91%	Pre-trained Residual model
Jin and Xie [21]	CNN	86.57%	Self-defined CNN vs. ResNet-50
Juhan Yu [22]	CNN	88%	2D-convolutional layers
Reza et al. [29]	CNN	84.78%	Over-sampling techniques
Cruz-Roa et al. [24]	CNN	71.80%	Whole Slide Images (WSI)

works operated on different methodologies, such as HOG and Canny Edge detection and the use of custom CNN architectures and pre-trained models like ResNet-50, etc. The identified research gaps indicate that some past studies have encountered performance issues when dealing with diverse and independent datasets, including scalability concerns and dependability analysis. However, further investigation and advancements in breast cancer detection and analysis techniques can potentially address these issues and lead to significant improvements. By addressing these research gaps, we can aim for more accurate diagnoses, better treatment outcomes, and enhanced patient care in breast cancer detection.

III. METHODOLOGY

The methodology employed in this study aims to develop a hybrid dependable deep feature extraction and machine learning model for breast cancer detection. The proposed approach capitalizes on the strengths of both DL and traditional ML techniques to improve the accuracy and reliability of breast cancer detection, as illustrated in Figure 1. This section thoroughly explains the step-by-step process we used to create and test the hybrid model. The process encompasses essential stages such as data preprocessing, feature extraction, model training, and performance evaluation metrics, as outlined below:

- Dataset Acquisition: It encompasses capturing and acquiring a histopathological image-based (HPI) dataset in this model. This dataset serves as the basis for training and evaluating the models in the experiment.
- Preprocessing: To prepare the dataset for DL model training, we apply effective preprocessing techniques such as resizing images, sharpening, normalizing pixel values, and label encoding.
- Dataset Split: We evaluate the performance of the models using the k-fold cross-validation technique. This technique divides the dataset into training and test subsets, ensuring that the models are trained and tested on distinct subsets of data. This technique also ensures a reliable assessment of their capacity to generalize.
- Deep Feature Extraction: We use a pre-trained TL model to extract deep features from breast cancer images, ResNet50V2. This model has already been trained on a large dataset and has learned valuable representations that we can use to detect breast cancer.
- Machine Learning Models: We apply various ensemble-based ML models, such as DT, RF, ET, AdB, HGBC, GBC, XGB, and LGB classifiers, to analyze the deep features obtained. We aim to train these models using the extracted features and accurately classify breast cancer images.
- Performance Evaluation: We use various performance metrics to analyze the experiments to evaluate how well the suggested methodology works. Regarding detecting breast cancer, accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), and dependability analysis are the standard evaluation metrics. These metrics show how well the models can identify malignant and benign cases by assessing their performance.

By integrating deep feature extraction with machine learning algorithms, our methodology seeks to improve the accuracy and robustness of breast cancer detection, ultimately contributing to early diagnosis and effective treatment of this critical health condition.

A. DATASET ACQUISITION

This phase encompasses identifying and selecting a suitable dataset containing pertinent information on breast cancer HPI. It also involves obtaining the necessary permissions to access and utilize the data and ensuring strict adherence to data privacy, confidentiality, and ethical guidelines and regulations. This study employed the IDC dataset, representing the most prevalent subtype among all breast cancers. The dataset consists of 162 whole-mount slide images of Breast Cancer (BCa) specimens scanned at 40x magnification. These whole mount slides contain regions focusing on areas with Invasive Ductal Carcinoma (IDC). Therefore, one crucial step in automatically grading aggressiveness is to accurately identify the areas of invasive ductal carcinoma (IDC) areas within the entire sample slides [30].

The IDC dataset is publicly available and accessible from the original source [31]. It is a valuable resource for developing and evaluating automated methods for breast cancer detection. By accurately identifying and categorizing

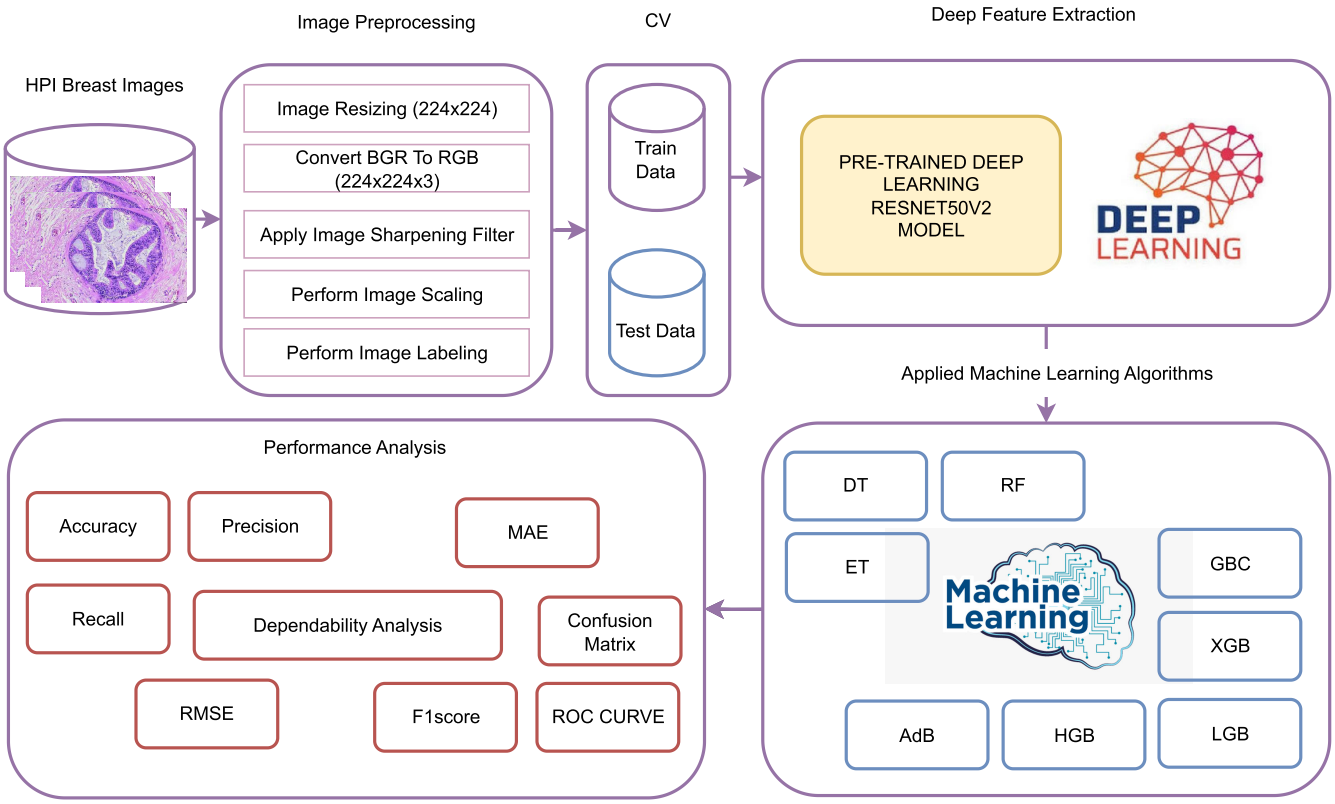


FIGURE 1. The proposed breast cancer detection architecture.

breast cancer subtypes, these automated methods can assist pathologists in saving time and reducing errors in clinical practice [32]. Our experiments specifically utilized the HPI breast cancer image dataset, which consisted of 2000 images. Among these images, 1200 were classified as positive for Invasive Ductal Carcinoma (IDC), indicating the presence of breast cancer. However, out of the total images, 800 were classified as non-cancerous cases with no evidence of IDC. The distribution of the IDC dataset utilized in our experiments is illustrated in Table 2.

To further facilitate the breast cancer detection task, 277,524 patches, each measuring 50 × 50 pixels, were extracted from the original dataset. These patches serve as the input samples for the breast cancer detection task. Among the extracted patches, 198,738 were classified as IDC negative, while 78,786 instances were classified as IDC positive. Each patch's file name follows the format: *u_xX_yY_classC.png*, where *u* represents the patient ID (e.g., 10253_*idx*5), *X* and *Y* indicate the *x* and *y* coordinates from which the patch was cropped, and *C* represents the class, with 0 denoting non-IDC and 1 denoting IDC [30].

The choice of the IDC dataset was motivated by several factors:

1) **Relevance to Breast Cancer Detection:** The dataset focuses on Invasive Ductal Carcinoma (IDC), the most common and aggressive type of breast cancer. As our study aims to improve breast cancer detection, using a dataset specifically targeting this subtype is highly relevant.
2) **Public Availability:** The IDC dataset is publicly available, ensuring our experiments' transparency and reproducibility. Researchers worldwide can access the dataset and validate our findings.
3) **Large and Diverse Sample Size:** With 2000 images and over 277,000 patches, the IDC dataset provides a substantial and diverse sample size, enabling robust evaluation and generalization of our proposed approach.
4) **Evaluation Benchmark:** As a widely used dataset in the field, the IDC dataset serves as a benchmark for breast cancer detection methods. Comparing our results with existing literature on the same dataset allows for a fair evaluation of our approach's performance.
5) **Potential Clinical Impact:** Accurate breast cancer detection can significantly impact patient outcomes and clinical practice. Using the IDC dataset, our research

TABLE 2. Distribution of IDC dataset.

IDC Class	Number of Images
Positive IDC	1200
Negative IDC	800

aims to contribute to advancing automated methods that can assist pathologists in making timely and accurate diagnoses.

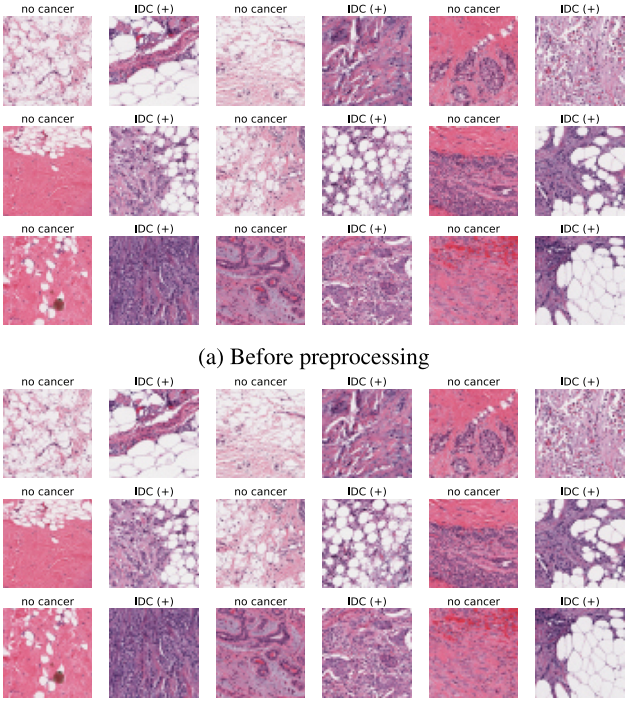
B. PREPROCESSING

In the context of our breast cancer detection methodology, image preprocessing plays a crucial role in preparing the input images for subsequent analysis. It involves a series of operations applied to the raw images to enhance their quality and extract meaningful information. Effective image preprocessing can significantly improve the performance and accuracy of the subsequent analysis.

The proposed approach employed several image preprocessing techniques to prepare the HPI images for breast cancer detection. The following techniques were applied:

- Image Resizing: We resized the images to a fixed size of 224 × 224 pixels to ensure consistent input image dimensions and reduce computational complexity. This step is crucial when using pre-trained deep learning models, as they often have specific size requirements. Resizing the images to match the model's input size allows us to leverage its learned features effectively. Additionally, resizing to a standardized size improves computational efficiency during training and inference, making the process more efficient overall.
- Convert BGR To RGB: The original HPI images were in the BGR (Blue-Green-Red) color space. We converted them to the more commonly used RGB (Red-Green-Blue) color space, better suited for subsequent image analysis tasks. This conversion step ensures consistent color representation across the dataset.
- Image Sharpening Filter: We applied an image sharpening filter to enhance the edges and fine details in the images. This filter increases the contrast between image features, making them more distinguishable and facilitating the subsequent feature extraction process.
- Image Scaling: To normalize the pixel values and bring them within a standardized range, we performed image scaling by dividing the pixel values by 255. This normalization step ensures that the pixel values are in a consistent numerical range, which aids in the convergence of subsequent machine learning algorithms.
- Image Labeling: The original HPI dataset had object labels in textual form. To facilitate the subsequent analysis, we performed image labeling by converting the object labels to numeric labels. This conversion enables the use of various ML algorithms that require numeric labels for classification tasks.

Overall, the preprocessing techniques used in our breast cancer detection methodology offer several benefits. First, image resizing ensures consistent image sizes, minimizing biases from dimension variations. Second, converting from BGR to RGB color space standardizes color representation for reliable feature extraction. Third, the image sharpening filter enhances fine details, improving the discrimination of relevant features. Fourth, scaling normalizes pixel values, aiding efficient model training. Finally, image labeling enables the use of diverse classification algorithms. By applying these techniques, we enhance the accuracy and reliability of our breast cancer detection system, preparing HPI images for deep feature extraction and machine learning.



(a) Before preprocessing

(b) After preprocessing

FIGURE 2. Comparison of original and preprocessed images for breast cancer detection.

Visual comparisons in Figure 2 further illustrate the effectiveness of these preprocessing techniques. The "Before" images show variations in dimensions, color representation, and fine details, while the corresponding "After" images demonstrate the consistent size, standardized color representation, enhanced details, and normalized pixel values. These comparisons emphasize the impact of preprocessing in preparing HPI images for subsequent analysis in our breast cancer detection methodology.

C. DEEP FEATURE EXTRACTION

This section represents the concept of deep feature extraction and its application in breast cancer detection. Deep feature extraction involves extracting meaningful and informative features from raw data using deep neural networks. These features capture high-level representations that are more effective in addressing the task at hand [10].

We employed the ResNet50V2 architecture as a deep feature extraction model for our breast cancer detection. ResNet50V2 is a convolutional neural network (CNN) that performs remarkably in various computer vision tasks. It is a variant of the ResNet architecture that uses skip connections to address the degradation problem in deep networks [33].

ResNet50V2 comprises 50 layers and is pre-trained on a large-scale image dataset, such as ImageNet. It utilizes residual blocks, enabling the network to learn residual mappings and facilitating the training of deeper models. In addition, the skip connections allow gradients to flow directly from early layers to later layers, enhancing the training process.

The ResNet50V2 architecture is beneficial for feature extraction in breast cancer detection because it can capture complex and hierarchical patterns from breast cancer images. With its deep layers, ResNet50V2 learns abstract representations of textures, shapes, and structures crucial in distinguishing between benign and malignant breast tissues. Leveraging the pre-trained ResNet50V2 model provides the advantage of TL, as it has already learned generic features from a large-scale image dataset like ImageNet. This pre-training allows ResNet50V2 to capture general image representations that can be fine-tuned for breast cancer detection. By utilizing the learned features from ResNet50V2, we can enhance the accuracy and performance of the breast cancer detection system [34].

The features extracted from ResNet50V2 serve as a higher-level representation of the input breast cancer images, capturing crucial information for classification and detection tasks. These features are inputs for various machine learning models, enabling us to leverage the powerful representations learned by ResNet50V2. By taking that action, we aim to enhance the accuracy and performance of our breast cancer detection system. The high-level features learned by ResNet50V2 provide a more comprehensive and informative representation of the input data, improving our ability to distinguish between benign and malignant breast tissues and ultimately leading to more effective detection of breast cancer.

D. MACHINE LEARNING MODELS

The breast cancer detection study used a mix of deep feature extraction and traditional machine learning algorithms. This approach provided several benefits, including dimensionality reduction, addressing the class imbalance, enhancing out-of-distribution detection, performing exploratory data analysis, and employing model compression techniques. Various machine learning algorithms, specially ensembled-based, were utilized in this study to achieve these objectives, as summarized below:

- Decision Tree (DT): DT is a non-parametric supervised learning algorithm that partitions the input space into regions and assigns a class label to each region [35].
- Random Forest (RF): RF is an ensemble learning method that combines multiple decision trees to improve prediction accuracy [36].
- Extra Trees (ET): ET is an extension of the Random Forest algorithm that further randomizes the tree construction process [37].
- AdaBoost (AbB): AdB is an ensemble learning method that combines multiple weak classifiers to create a strong classifier [38].
- Histogram Gradient Boosting Classifier (HGBC): HGBC is a gradient boosting classifier that utilizes histogram-based techniques for efficient training and prediction [39].
- Gradient Boosting Classifier (GBC): GBC is a boosting algorithm that builds an ensemble of weak learners, typically decision trees, by iteratively minimizing a loss function [40].
- XGBoost (Extreme Gradient Boosting): XGBoost is an optimized gradient boosting framework incorporating regularization techniques and parallel processing capabilities [39].

IV. EXPERIMENTS AND RESULT ANALYSIS

In this section, we present experiments and the results of our analysis for breast cancer detection using our proposed approach. We designed the experimental setup and analyzed metrics such as accuracy, sensitivity, specificity, and F1 score to assess the effectiveness of our method. Accuracy measures the overall correctness of the classification results, sensitivity evaluates the ability to identify positive instances, specificity measures the ability to identify negative instances, and the F1 score provides a balanced assessment. These metrics allowed us to gain valuable insights into the performance of our approach, enabling us to make informed evaluations and comparisons with other existing methods in the field of breast cancer detection.

A. ENVIRONMENT SETUP

We experimented on a high-performance machine with eight cores, 64 GB RAM, and a 100 GB disk. We also utilized TensorFlow and Keras frameworks to leverage their deep learning capabilities for the research. The environment included necessary dependencies and libraries, particularly those specific to pre-trained models for TL. Utilizing this compressed environment played a crucial role in conducting the experiments efficiently, facilitating the analysis and validation of the proposed approach for breast cancer diagnosis.

We trained our model using different parameters for feature extraction and classification purposes. In feature extraction, we configured the ResNet50V2 model using the parameters of include_top = False, weights = ''imagenet'', input_tensor = None, input_shape = (size, size, 3), pooling = None; where size is set to 224, indicating that the input images were resized to a square of 224 × 224 pixels. By setting include_top to False, we excluded the fully-connected layers at the end of the network, allowing us to use the ResNet50V2 model as a feature extractor. On the other hand, in ML algorithms, we mainly focused on the number of estimators and learning rate, which are essential for controlling the models' performances. We utilized different values of these parameters and selected the best combinations, including the estimator of 100 and the learning rate of 0.1.

B. PERFORMANCE EVALUATION METRICS

As mentioned in Section III, we assess how well our suggested method works by measuring its performance using

TABLE 3. Confusion matrix.

	Actual positive	Actual negative
Predicted positive	TP	FP
Predicted negative	FN	TN

different metrics. These metrics serve to measure different aspects of the approach's performance. The formulation of these metrics is as follows:

- **Confusion Matrix:** The confusion matrix provides a breakdown of the predicted and actual positive/negative classes, including true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The confusion matrix is presented in Table 3.
  In the table, TP represents the number of instances correctly predicted as positive, TN represents the number of instances correctly predicted as negative, FP represents the number of instances wrongly predicted as positive, and FN represents the number of instances wrongly predicted as negative. The confusion matrix allows for a detailed examination of the model's performance by providing insights into the correct and incorrect predictions made for positive and negative instances.
- **Accuracy:** To determine the model's accuracy, we calculate the ratio of correctly classified samples to the total number of samples. We did it using the following equation:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

- **Precision:** To measure the accuracy of positive predictions, this metric calculates the ratio of correctly identified positive results to the total number of positive results, including both true and false positives. The formula is as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

- **Recall:** This measures how well the model can correctly identify positive instances. It's calculated by dividing the number of true positives by the total of true and false negatives.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

- **F1-score:** The F1-score combines precision and recall into a single metric, providing a balanced measure of the model's performance.

$$\text{F1-score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{4}$$

- **MAE:** It measures the average magnitude of the errors between the predicted and actual values.

$$\text{MAE} = \frac{\sum_{i=1}^{n} |\text{predicted}(i) - \text{actual}(i)|}{n} \tag{5}$$

- **MSE:** This calculation determines the average of the squared variances between the predicted values and the actual values.

$$\text{MSE} = \frac{\sum_{i=1}^{n} (\text{predicted}(i) - \text{actual}(i))^2}{n} \tag{6}$$

- **RMSE:** RMSE is the square root of the MSE, representing the standard deviation of the errors.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\text{predicted}(i) - \text{actual}(i))^2}{n}} \tag{7}$$

  where the variable n represents the total number of data samples taken.
- **ROC Curve and AUC:** ROC curves are graphical representations used to evaluate classifier performance, with the area under the curve (AUC) indicating the separability of class labels. A higher AUC value indicates better predictive capability, while a lower value suggests a less accurate model.
- **Cross-validation:** K-fold cross-validation (CV) is widely used for evaluating machine learning models. It divides the dataset into k folds, using each fold as a validation set while the rest serve as the training set. This process is repeated k times, allowing for a comprehensive evaluation of the model's performance. In addition, by averaging the results, k-fold cross-validation provides a robust estimate of the model's generalization ability and helps identify issues like overfitting. In this study, the value of k was selected as 10 for the k-fold cross-validation. In other words, the dataset was divided into ten subsets of approximately equal size. The training process was performed iteratively, with each subset serving as the validation set once while the remaining nine were used for training.

C. RESULT ANALYSIS

This section presents the performance analysis of different ML algorithms for breast cancer detection. We conducted a dependability analysis to evaluate the reliability and consistency of the models. This analysis examined the performance stability across iterations or datasets, ensuring that random variations or dataset-specific characteristics did not influence the results. We gained insights into the models' generalization ability by assessing the performance robustness and identified potential limitations or biases. The dependability analysis allowed us to make informed decisions about the effectiveness of the models in real-world scenarios.

The performance analysis of various algorithms for breast cancer detection is presented in Table 4, which provides a detailed overview of the accuracy, precision, recall, F1-score, MAE, MSE, and RMSE metrics. Additionally, Figure 3 exhibits the corresponding graphs to represent the performance analysis visually. The results indicate that the DT algorithm achieved an accuracy of 78%, a precision of 77.39%, a recall of 79.47%, and an F1-score of 77.42%. Additionally, the DT algorithm exhibited an MAE of 22,

TABLE 4. Performance analysis of eight algorithms for breast cancer detection.

Algorithm	Accuracy	Precision	Recall	F1-score	MAE	MSE	RMSE
DT	78	77.39	79.47	77.42	22	22	46.9
RF	94.5	94.45	93.63	94.01	5.5	5.5	23.45
ET	94.5	94.18	93.92	94.05	5.5	5.5	23.45
AdB	90.5	90.6	88.73	89.53	9.5	9.5	30.82
HGB	94.5	94.18	93.92	94.05	5.5	5.5	23.45
GBC	94	93.77	93.24	93.49	6	6	24.49
XGB	94.5	93.96	94.21	94.08	5.5	5.5	23.45
LGB	95	94.86	94.32	94.57	5	5	22.36

MSE of 22, and RMSE of 46.9. On the other hand, both the RF and ET algorithm demonstrated higher accuracy than the DT, achieving an accuracy of 94.5%. The RF algorithm exhibited a precision of 94.45%, recall of 93.63%, and F1-score of 94.01%, while the ET algorithm achieved a precision of 94.18%, recall of 93.92%, and F1-score of 94.05%. Furthermore, the RF and ET algorithms had an MAE of 5.5, MSE of 5.5, and RMSE of 23.45. The AdB, HGB, GBC, XGB, and LGB algorithms also showcased competitive performance. AdB achieved an accuracy of 90.5%, while HGB, GBC, XGB, and LGB achieved accuracies of 94.5%, 94%, 94.5%, and 95%, respectively.

From Table 4, it is evident that among the analyzed algorithms, the LGB algorithm achieves the highest accuracy rate of 95%. Additionally, LGB outperforms other algorithms regarding precision, recall, and F1-score. These results demonstrate the superior performance of LGB in accurately classifying breast cancer cases. The remarkable accuracy rate achieved by the LGB algorithm contributed to its capability to handle complex datasets and effectively capture underlying patterns in breast cancer images. Furthermore, by leveraging gradient boosting techniques, LGB sequentially combines weak learners to construct a robust ensemble model. This mechanism enables LGB to learn intricate relationships and extract informative features from the dataset, improving classification performance significantly. Besides, the LGB algorithm demonstrates excellent performance based on its low MAE, MSE, and RMSE values. These metrics quantify the error between predicted and actual values, and the consistently low error rates indicate the model's reliability in making accurate predictions. Overall, the LGB algorithm achieves the highest accuracy rate and exhibits exceptional performance across various evaluation metrics, making it a promising choice for breast cancer detection.

The confusion matrices presented in Figure 4 offer valuable insights into the performance of various algorithms, revealing their accuracy and misclassification rates in breast cancer detection. For example, the DT algorithm's confusion matrix shows 62 actual positive cases, 11 false positive cases, 33 false negative cases, and 94 true negative cases. Similarly, the RF algorithm achieved 66 true positive cases, seven false positive cases, four false negative cases, and 123 true negative cases, as evident from its confusion matrix. The ET algorithm yielded a confusion matrix with 67 true positive cases, six false positive cases, five false negative cases, and 122 true negative cases. Finally, AdB's confusion matrix displayed 60 true positive cases, 13 false positive cases, 6 false negative cases, and 121 true negative cases.

In contrast, the HGB algorithm's confusion matrix exhibited 67 true positive cases, 6 false positive cases, 5 false negative cases, and 122 true negative cases. The GBC algorithm achieved 66 true positive cases, 7 false positive cases, 5 false negative cases, and 122 true negative cases, as observed in its confusion matrix. XGB demonstrated 68 true positive cases, 5 false positive cases, 6 false negative cases, and 121 true negative cases in its confusion matrix. Lastly, the LGB algorithm's confusion matrix revealed 67 true positive cases, 6 false positive cases, four false negative cases, and 123 true negative cases.

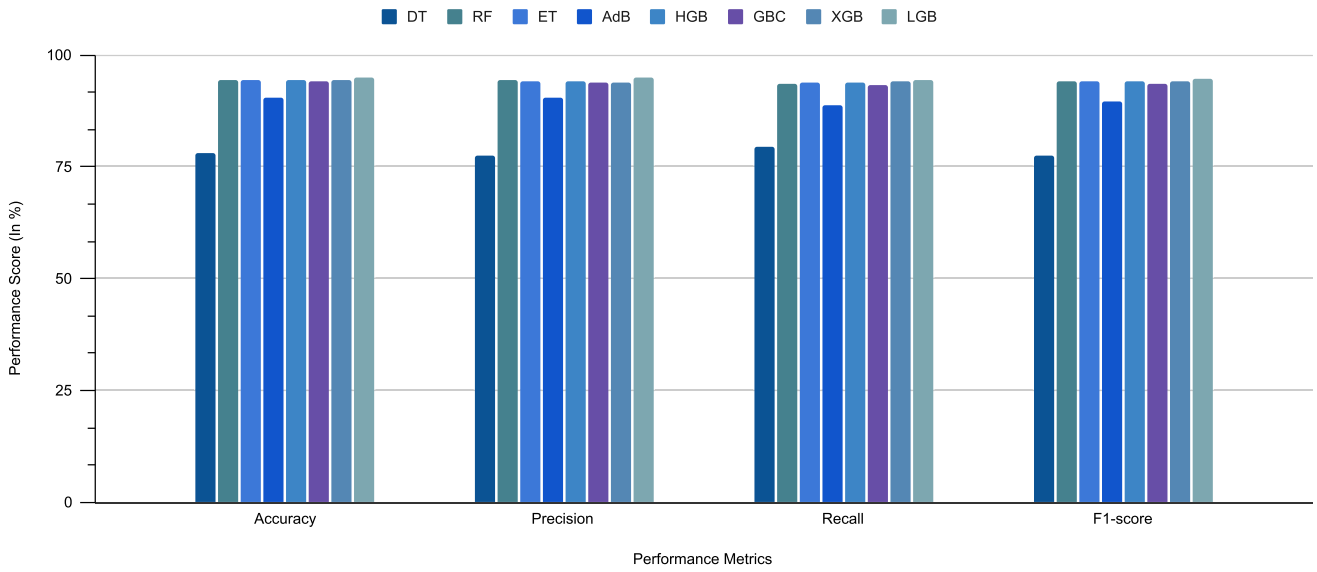
The confusion matrices reveal that the LGB algorithm achieves higher numbers of TP and TN cases than FP and FN cases, indicating better performance in breast cancer detection. Specifically, the LGB algorithm exhibits 67 TP cases, 6 FP cases, 4 FN cases, and 123 TN cases. These results suggest that the LGB algorithm has higher sensitivity in detecting TP cases and lower rates of misclassifying cases as FP or FN. Overall, the LGB algorithm performs better breast cancer detection with higher TP and TN counts and lower FP and FN counts, leading to more accurate classification results.

Table 5 presents a classification report for breast cancer detection using multiple algorithms. It provides an overview of the effectiveness of ML algorithms in accurately detecting breast cancer. The LGB algorithm stands out with promising results, as indicated by each class's precision, recall, and F1-score metrics.

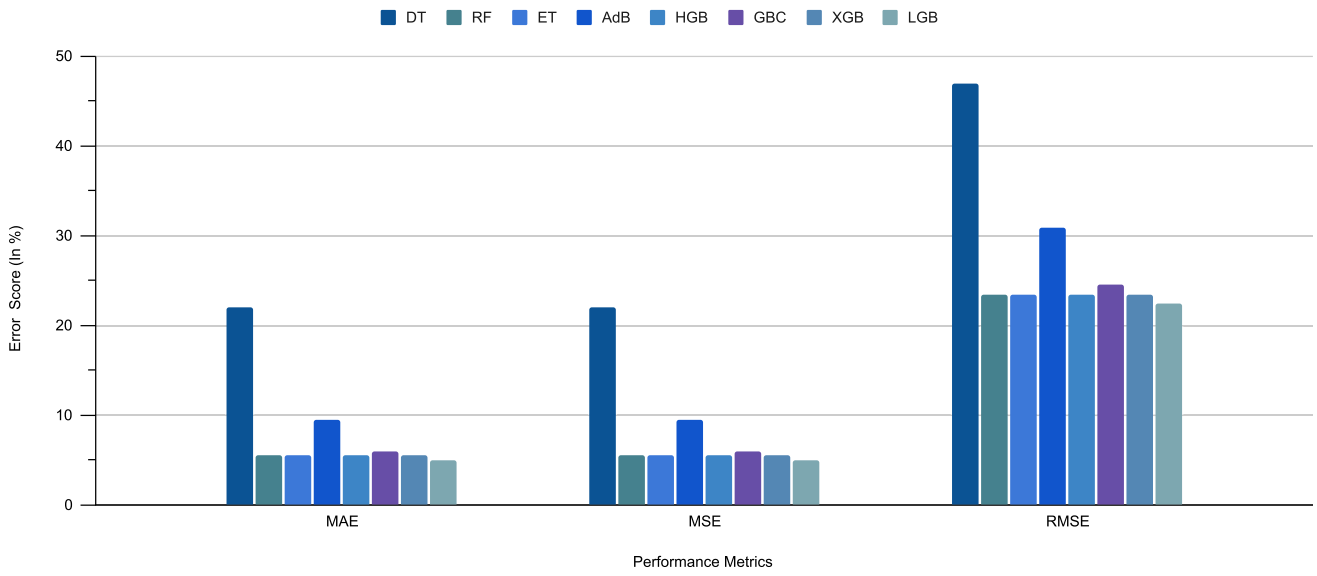
For the N_IDC class, the LGB algorithm achieves a precision of 94.37, indicating a low rate of FP predictions. On the other hand, the recall value of 91.78 indicates a high percentage of actual negative cases correctly identified. Therefore, the F1-score, calculated at 93.06, provides a balanced performance measure for this class.

Regarding the P_IDC class, the LGB algorithm achieves a precision of 95.35, demonstrating a low rate of false negatives. The recall value 96.85 indicates a high percentage of actual positive cases correctly identified. The F1-score for this class is calculated at 96.09, indicating an overall balanced performance.

These results highlight the LGB algorithm's effectiveness in accurately detecting breast cancer cases, making it a promising choice for breast cancer detection tasks.



(a) Performance



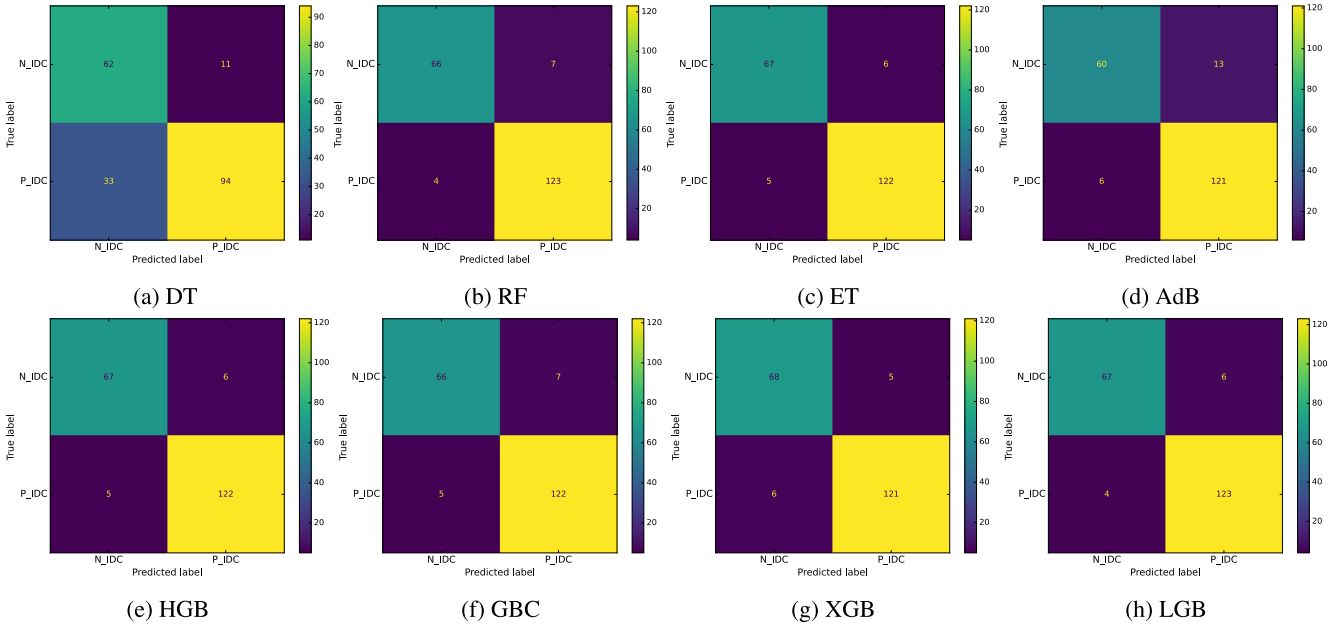
(b) Error

FIGURE 3. **Visual analysis of performance for breast cancer detection.**

In addition, the precision, recall, and F1-score metrics showcase the algorithm's ability to minimize false positives and false negatives, ensuring a reliable and balanced classification performance for both negative and positive breast cancer cases.

In the context of dependability analysis, we compared the ML algorithms' performances using the AUC score as an essential metric, as depicted in Figure 5. The AUC score assesses each algorithm's breast cancer detection performance, with higher scores indicating a better ability to distinguish between positive and negative cases. Based on our experiments, the AUC scores for the different algorithms were as follows: DT (79.47), RF (97.91), ET (97.94), AdB (97.58), HGB (99.08), GBC (98.81), XGB (98.77), and LGB (99.09). These scores highlight the strong performance of the algorithms in accurately classifying breast cancer cases. The high AUC values achieved by the HGB, GBC, XGB, and LGB algorithms are particularly noteworthy.

Among these high-performing algorithms, the LGB algorithm stands out with an exceptionally high AUC score of 99.09. This score indicates its strong discriminatory power



(a) DT (b) RF (c) ET (d) AdB

(e) HGB (f) GBC (g) XGB (h) LGB

FIGURE 4. Confusion matrix for breast cancer detection.

TABLE 5. Classification report for breast cancer.

Algorithm		DT			RF			ET			AdB	
Class	precision	recall	f1-score	precision	recall	f1-score	precision	recall	f1-score	precision	recall	f1-score
N_IDC	65.26	84.93	73.81	94.29	90.41	92.31	90.91	82.19	86.33	90.91	82.19	86.33
P_IDC	89.52	74.02	81.03	94.62	96.85	95.72	90.3	95.28	92.72	90.3	95.28	92.72
accuracy	78	78	78	94.5	94.5	94.5	90.5	90.5	90.5	90.5	90.5	90.5
macro avg	77.39	79.47	77.42	94.45	93.63	94.01	90.6	88.73	89.53	90.6	88.73	89.53
weighted avg	80.67	78	78.4	94.5	94.5	94.47	90.52	90.5	90.39	90.52	90.5	90.39
Algorithm		HGB			GBC			XGB			LGB	
Class	precision	recall	f1-score	precision	recall	f1-score	precision	recall	f1-score	precision	recall	f1-score
N_IDC	93.06	91.78	92.41	92.96	90.41	91.67	91.89	93.15	92.52	94.37	91.78	93.06
P_IDC	95.31	96.06	95.69	94.57	96.06	95.31	96.03	95.28	95.65	95.35	96.85	96.09
accuracy	94.5	94.5	94.5	94	94	94	94.5	94.5	94.5	95	95	95
macro avg	94.18	93.92	94.05	93.77	93.24	93.49	93.96	94.21	94.08	94.86	94.32	94.57
weighted avg	94.49	94.5	94.49	93.98	94	93.98	94.52	94.5	94.51	94.99	95	94.98

in distinguishing between positive and negative breast cancer cases. In addition, the high AUC score suggests that the LGB algorithm has a higher probability of assigning higher predicted probabilities to positive and negative instances. The results show that the LGB algorithm has the potential to perform better in detecting breast cancer, making it a reliable and accurate option for predicting outcomes.

Figure 5 also reveals that the LGB algorithm demonstrates high precision and recall values for negative and positive breast cancer cases, as the classification report indicates. Its F1 scores show a good balance between precision and recall, indicating a robust overall performance. Additionally, the LGB algorithm stands out with an exceptional AUC score of 99.09, reflecting its discriminatory solid power in distinguishing between positive and negative cases. This analysis suggests a higher probability of assigning higher predicted probabilities to positive instances. Based on the evaluation metrics, the LGB algorithm is the top choice for breast cancer detection, offering superior accuracy, precision, recall, F1 score, and AUC score.

D. DISCUSSION

The comparison presented in Table 6 provides insights into different breast cancer detection approaches and their corresponding accuracy rates. For example, Anjum et al. [18] utilized HOG and Canny Edge techniques with SVM, achieving an accuracy rate of 94%, albeit with lower accuracy for non-cancerous cases. Mridha et al. [19] employed a CNN-based approach, achieving an accuracy rate 86.24% for invasive ductal cancer detection. Kulkarni and Sundaray [23] utilized a pre-trained ResNet152 model with FC layers, achieving an accuracy rate of 91%. Jin and Xie [21] compared a self-defined CNN model with ResNet-50, with the latter achieving an accuracy rate of 86.57%. Other studies, such as those by Yu [22], Reza and Ma [29], and Cruz-Roa et al. [24], also employed CNN-based approaches, achieving accuracy rates ranging from 84.23% to 88%. Importantly, our proposed

TABLE 6. Comparison of breast cancer detection approaches.

Author	Feature Extraction	Machine/Deep Learning	Accuracy Rate
Anjum et al. [18]	HOG and Canny Edge	SVM	94% (C) 22% (NC)
Mridha et al. [19]	-	CNN	86.24%
Kulkarni et al. [23]	ResNet152	FC Layer	91%
Jin and Xie [21]	-	ResNet-50	86.57%
Juhan Yu [22]	-	CNN	88%
Reza et al. [29]	-	CNN	85.48%
Cruz-Roa et al. [24]	-	CNN	84.23%
Our Model	ResNet50V2	LGB	95%

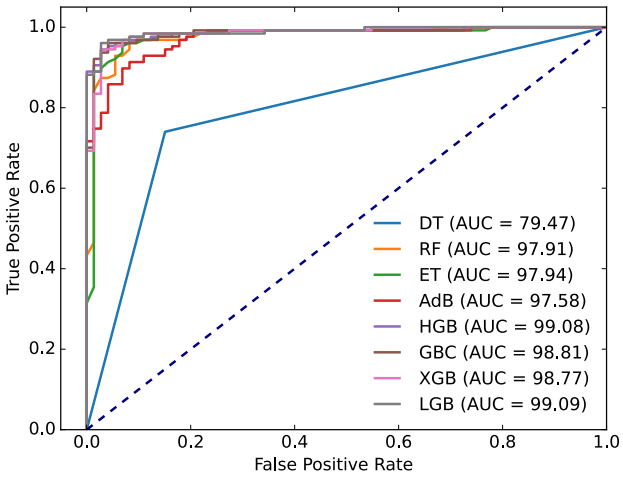


FIGURE 5. ROC Curve for evaluating classifiers.

model, which utilizes ResNet50V2 for feature extraction and LGB for classification, achieved the highest accuracy rate of 95% in breast cancer detection among the listed approaches, as shown in Table 6.

Using LGB in our model provides several benefits for breast cancer detection. Firstly, LGB is a gradient-boosting framework renowned for its high efficiency and speed, making it suitable for efficiently processing large datasets. This efficiency allows for faster and more timely detection and diagnosis of breast cancer, crucial for early intervention and improved patient outcomes. Additionally, LGB can handle complex and non-linear relationships within the data, enabling the model to capture intricate patterns and subtle variations that may indicate breast cancer. This ability to capture complex relationships enhances the accuracy and reliability of our model, ensuring more accurate predictions and reducing the risk of false negatives or positives.

E. LIMITATIONS

While our hybrid model shows promising results, it is essential to acknowledge its limitations for a balanced interpretation of our findings. Firstly, our study focuses specifically on Invasive Ductal Carcinoma (IDC), which may restrict the generalizability of our approach to other breast cancer subtypes. Further research is needed to assess its performance across a broader range of cancer types.

Secondly, the performance of our model is impacted by the quality and variety of the training dataset. Although we utilized a comprehensive dataset, the availability of more extensive and diverse datasets could further enhance the model's generalization capabilities.

Lastly, while our model demonstrates high accuracy, there are still areas for improvement, particularly in scenarios with challenging image conditions, such as noise and artifacts. Further optimization and refinement of the model could enhance performance in real-world clinical settings.

Despite these limitations, our hybrid breast cancer detection model significantly advances the field and offers valuable insights for accurate IDC diagnosis. As with any new technology, further research and validation on more extensive and diverse datasets will be crucial to assess its potential clinical impact and extend its applicability to a broader spectrum of breast cancer cases.

V. DEPENDABILITY ANALYSIS

The dependability of our developed breast cancer detection model was evaluated based on the study conducted by Talukder et al. [11]. Their reliability evaluation assessed the availability, efficiency, and scalability aspects. Our model employed deep learning for feature extraction and utilized the LGB algorithm for reliable differentiation between positive and negative IDC breast scenarios. This choice ensured high efficiency without significant loss, enhancing the availability of our proposed methodology. Furthermore, through comprehensive analysis and productivity assessments, we compared our framework to existing techniques, and it outperformed them with lower error rates and computational loss.

We obtained the accuracy of our model 95%, indicating its ability to classify breast cancer cases accurately. Precision, recall, and F1-score measures were 94.86%, 94.32%, and 94.57%, respectively, demonstrating the model's effectiveness in minimizing false positives and false negatives. Additionally, the MAE, MSE, and RMSE values were 5, 5, and 22.36, respectively, affirming our model's overall accuracy and the small magnitude of errors. These results validate the dependability of our breast cancer detection framework, highlighting its potential for reliable and accurate diagnosis.

Table 7 demonstrates the scalability of our proposed LGB model by evaluating its performance on different numbers of images. The results show that as the number of images increases, the accuracy rate of our model improves.

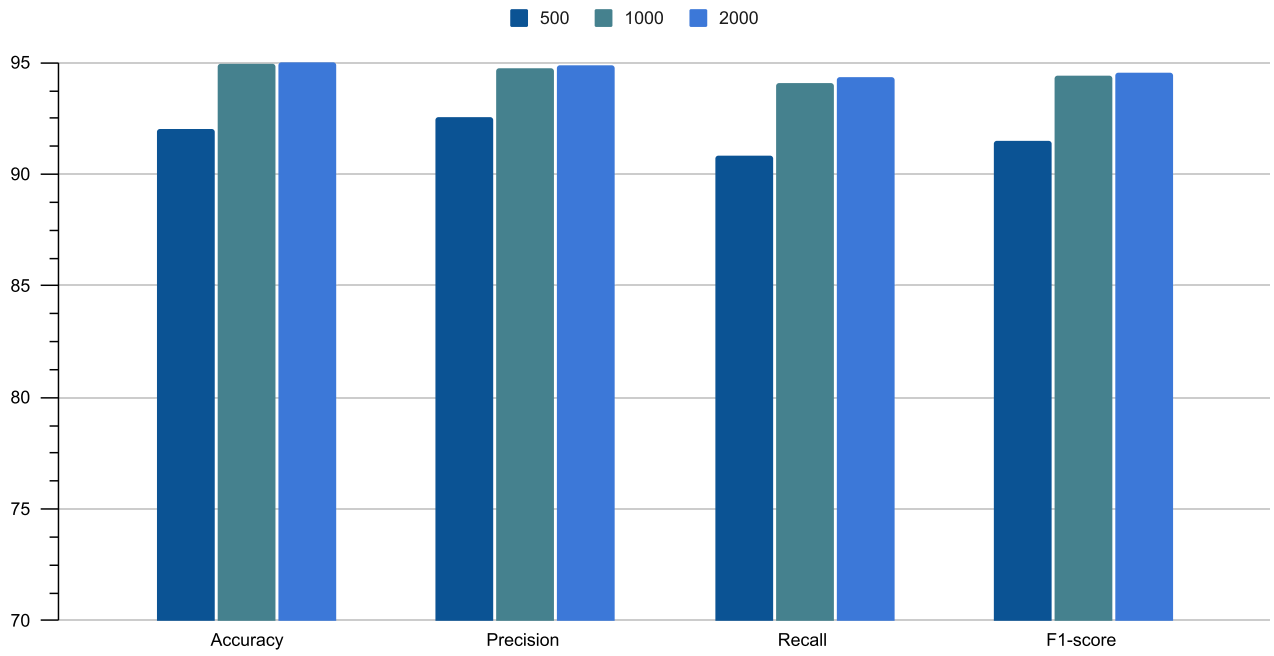


FIGURE 6. Scalability of our proposed model with respect to different sample sizes.

TABLE 7. Performance of proposed approach on different sampling sizes.

Num of Images	Accuracy	Precision	Recall	F1-score
500	92	92.53	90.83	91.51
1000	95	94.73	94.07	94.39
2000	95	94.86	94.32	94.57

For example, with 500 images, the accuracy rate achieved was 92%, which increased to 94% with 1000 images. Finally, with 2000 images, the accuracy rate reached 95%. These findings highlight the scalability of our model, indicating its ability to handle larger datasets without compromising accuracy. In addition, the precision, recall, and F1-score measures also show consistent improvements as the number of images increases.

The scalability of our proposed LGB model is crucial in real-world applications where the volume of data continues to grow. Maintaining high accuracy rates with larger datasets ensures the model's reliability and effectiveness in handling diverse and expanding datasets. Figure 6 visually represents the performance variation of our model for the number of images, further emphasizing its scalability. The increasing trend of accuracy, precision, recall, and F1-score reinforces the scalability of our model and validates its suitability for practical deployment in breast cancer detection scenarios.

These findings highlight the scalability of our proposed LGB model and reinforce the dependability of our developed breast cancer detection model. The improved performance, high accuracy rates, and consistent results across different dataset sizes demonstrate the efficiency and reliability of our model. Moreover, our model outperformed several existing techniques, exhibiting lower error rates and computational loss. This comprehensive evaluation and analysis contribute to the trustworthiness and robustness of our proposed methodology, further establishing its dependability in breast cancer detection.

VI. CONCLUSION

This paper presents a hybrid dependable model for breast cancer detection that combines DL and ML techniques. Our model leverages the strengths of the DL-based pred-trained ResNet50V2 TL model in extracting complex patterns and representations from publicly available HPI-based IDC images. Through extensive experimentation on a comprehensive dataset of IDC, we demonstrated the effectiveness of our approach by achieving high accuracy, precision, recall, and F1 score.

Our study contributes significantly to breast cancer detection by addressing several limitations of existing works. By integrating DL and ML techniques, our hybrid model offers a reliable and accurate approach to IDC detection. The incorporation of DL empowers the model to learn and extract meaningful features from complex image data, while the ML algorithms contribute interpretability and generalization capabilities. Moreover, by utilizing publicly available breast histopathology images, mainly targeting IDC, we ensure the accessibility and reproducibility of our research findings.

By focusing on this dataset, we provide valuable insights into the potential of our model for enabling early and accurate detection of IDC. Furthermore, the results obtained from our rigorous experiments serve as compelling evidence for our hybrid model's robustness and high performance. With an improved accuracy rate of 95%, a precision of 94.86%, recall of 94.32%, F1 score of 94.57%, and minimal error rates including MAE of 5%, MSE of 5%, and RMSE of

22.36%, our approach demonstrates its effectiveness in accurately detecting IDC. Moreover, our study holds promise for assisting medical professionals in making informed decisions and ultimately contributing to improved patient care. Furthermore, based on our findings, we propose that LGB is the most suitable ML algorithm for breast cancer classification after performing feature extraction using a pre-trained Resnet502 model.

In future research, several areas can be explored to enhance breast cancer detection and improve the impact of diagnosis. Firstly, integrating additional data sources, such as genomic or clinical data, can provide a more comprehensive understanding of breast cancer and contribute to the refinement of our model. In addition, by incorporating these additional data types, we can uncover critical insights into the genetic and clinical factors that influence breast cancer development, progression, and response to treatment.

Furthermore, our study provides a foundation for future research endeavors in breast cancer detection. By exploring these potential directions, we can continue to advance the accuracy, accessibility, and impact of breast cancer diagnosis, ultimately benefiting patients and healthcare systems worldwide.

ACKNOWLEDGMENT

We gratefully acknowledge the financial support provided by the UGC-supported Jagannath University Research Grant. This research was made possible by the generous funding from this grant, which significantly contributed to the successful completion of our study.

DECLARATIONS

The authors declare that they have no competing interests.

REFERENCES

[1] M. Ashrafizadeh, A. Zarrabi, A. Bigham, A. Taheriazam, Y. Saghari, S. Mirzaei, M. Hashemi, K. Hushmandi, H. Karimi-Maleh, E. N. Zare, E. Sharifi, Y. N. Ertas, N. Rabiee, G. Sethi, and M. Shen, "(Nano)platforms in breast cancer therapy: Drug/gene delivery, advanced nanocarriers and immunotherapy," *Medicinal Res. Rev.*, vol. 2023, pp. 1–12, May 2023.

[2] Q. Huang, L. Chen, J. Aiqun, F. Shi, and D. Feng, "Effectiveness of the hospital-community-family nursing model in breast cancer patients discharged with PICC lines," *J. Community Health Nursing*, vol. 2023, pp. 1–11, May 2023.

[3] P. M. C. Donahue, A. MacKenzie, A. Filipovic, and L. Koelmeyer, "Advances in the prevention and treatment of breast cancer-related lymphedema," *Breast Cancer Res. Treat.*, vol. 200, no. 1, pp. 1–14, Jul. 2023.

[4] *Global Breast Cancer Initiative Implementation Framework: Assessing, Strengthening and Scaling-up of Services for the Early Detection and Management of Breast Cancer*, World Health Organization, Geneva, Switzerland, 2023.

[5] V. B. Mathema, P. Sen, S. Lamichhane, M. Orešič, and S. Khoomrung, "Deep learning facilitates multi-data type analysis and predictive biomarker discovery in cancer precision medicine," *Comput. Struct. Biotechnol. J.*, vol. 21, pp. 1372–1382, Jan. 2023.

[6] P. Chhabra and Dr. S. Goyal, "A thorough review on deep learning neural network," in *Proc. Int. Conf. Artif. Intell. Smart Commun. (AISC)*, Jan. 2023, pp. 220–226.

[7] M. A. Talukder, M. M. Islam, M. A. Uddin, A. Akhter, M. A. J. Pramanik, S. Aryal, M. A. A. Almoyad, K. F. Hasan, and M. A. Moni, "An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning," *Expert Syst. Appl.*, vol. 230, Nov. 2023, Art. no. 120534.

[8] A. Subasi, A. D. Kandpal, K. A. Raj, and U. Bagci, "Breast cancer detection from mammograms using artificial intelligence," in *Applications of Artificial Intelligence in Medical Imaging*. Amsterdam, The Netherlands: Elsevier, 2023, pp. 109–136.

[9] M. M. Hassan, M. M. Hassan, F. Yasmin, M. A. R. Khan, S. Zaman, K. K. Islam, and A. K. Bairagi, "A comparative assessment of machine learning algorithms with the least absolute shrinkage and selection operator for breast cancer detection and prediction," *Decis. Anal. J.*, vol. 7, Jun. 2023, Art. no. 100245.

[10] M. A. Talukder, M. M. Islam, M. A. Uddin, A. Akhter, K. F. Hasan, and M. A. Moni, "Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning," *Expert Syst. Appl.*, vol. 205, Nov. 2022, Art. no. 117695.

[11] M. A. Talukder, K. F. Hasan, M. M. Islam, M. A. Uddin, A. Akhter, M. A. Yousuf, F. Alharbi, and M. A. Moni, "A dependable hybrid machine learning model for network intrusion detection," *J. Inf. Secur. Appl.*, vol. 72, Feb. 2023, Art. no. 103405.

[12] K. M. M. Uddin, N. Biswas, S. T. Rikta, S. K. Dey, and A. Qazi, "XML-LightGBMDroid: A self-driven interactive mobile application utilizing explainable machine learning for breast cancer diagnosis," *Eng. Rep.*, vol. 2023, May 2023, Art. no. e12666.

[13] H. Long, S. Li, and Y. Chen, "Digital health in chronic obstructive pulmonary disease," *Chronic Diseases and Translational Medicine*, vol. 2023, pp. 1–20, Jun. 2023.

[14] M. A. Uddin, M. M. Islam, M. A. Talukder, M. A. A. Hossain, A. Akhter, S. Aryal, and M. Muntaha, "Machine learning based diabetes detection model for false negative reduction," *Biomed. Mater. Devices*, vol. 2023, pp. 1–17, Jun. 2023.

[15] N. Ahmed, R. Ahammed, M. M. Islam, M. A. Uddin, A. Akhter, M. A. Talukder, and B. K. Paul, "Machine learning based diabetes prediction and development of smart web application," *Int. J. Cognit. Comput. Eng.*, vol. 2, pp. 229–241, Jun. 2021.

[16] M. M. Srikantamurthy, V. P. S. Rallabandi, D. B. Dudekula, S. Natarajan, and J. Park, "Classification of benign and malignant subtypes of breast cancer histopathology imaging using hybrid CNN-LSTM based transfer learning," *BMC Med. Imag.*, vol. 23, no. 1, pp. 1–15, Jan. 2023.

[17] S. Gupta, S. Agrawal, S. K. Singh, and S. Kumar, "A novel transfer learning-based model for ultrasound breast cancer image classification," in *Proc. ICCVBIC*. Cham, Switzerland: Springer, 2023, pp. 511–523.

[18] R. Anjum, R. R. Dipti, H. O. Rashid, and S. Ripon, "An efficient breast cancer analysis technique by using a combination of HOG and Canny edge detection techniques," in *Proc. 5th Int. Conf. Trends Electron. Informat. (ICOEI)*, Jun. 2021, pp. 1290–1295.

[19] K. Mridha, S. Kumbhani, S. Jha, D. Joshi, A. Ghosh, and R. N. Shaw, "Deep learning algorithms are used to automatically detection invasive ducal carcinoma in whole slide images," in *Proc. IEEE 6th Int. Conf. Comput., Commun. Autom. (ICCCA)*, Dec. 2021, pp. 123–129.

[20] R. O. Ogundokun, S. Misra, M. Douglas, R. Damaševičius, and R. Maskeliūnas, "Medical Internet-of-Things based breast cancer diagnosis using hyperparameter-optimized neural networks," *Future Internet*, vol. 14, no. 5, p. 153, May 2022.

[21] C. Jin and W. Xie, "Effective detection for invasive ductal carcinoma histopathology images based on ResNet," in *Proc. 3rd Int. Conf. Electron. Commun. Artif. Intell. (IWECAI)*, Jan. 2022, pp. 380–383.

[22] J. Yu, "Predicting invasive ductal carcinoma based on convolutional neural network," in *Proc. 3rd Int. Conf. Electron. Commun. Artif. Intell. (IWECAI)*, Jan. 2022, pp. 388–391.

[23] S. Kulkarni and A. Sundaray, "Detection of invasive ductal carcinoma using transfer learning with deep residual network," in *Proc. 19th OITS Int. Conf. Inf. Technol. (OCIT)*, Dec. 2021, pp. 115–120.

[24] A. Cruz-Roa, A. Basavanhally, F. González, H. Gilmore, M. Feldman, S. Ganesan, N. Shih, J. Tomaszewski, and A. Madabhushi, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," *Proc. SPIE*, vol. 9041, May 2014, Art. no. 904103.

[25] M. A. Naji, S. E. Filali, K. Aarika, E. H. Benlahmar, R. A. Abdelouhahid, and O. Debauche, "Machine learning algorithms for breast cancer prediction and diagnosis," *Proc. Comput. Sci.*, vol. 191, pp. 487–492, Jan. 2021.

[26] S. Singh and R. Kumar, "Histopathological image analysis for breast cancer detection using cubic SVM," in *Proc. 7th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Feb. 2020, pp. 498–503.

[27] G. Liang, X. Wang, Y. Zhang, X. Xing, H. Blanton, T. Salem, and N. Jacobs, "Joint 2D–3D breast cancer classification," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Nov. 2019, pp. 692–696.
[28] Y.-D. Zhang, S. C. Satapathy, D. S. Guttery, J. M. Górriz, and S.-H. Wang, "Improved breast cancer classification through combining graph convolutional network and convolutional neural network," *Inf. Process. Manage.*, vol. 58, no. 2, Mar. 2021, Art. no. 102439.
[29] M. S. Reza and J. Ma, "Imbalanced histopathological breast cancer image classification with convolutional neural network," in *Proc. 14th IEEE Int. Conf. Signal Process. (ICSP)*, Aug. 2018, pp. 619–624.
[30] A. Janowczyk and A. Madabhushi, "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases," *J. Pathol. Informat.*, vol. 7, no. 1, p. 29, Jan. 2016.
[31] P. T. Mooney. (2021). *Breast Histopathology Images*. Accessed: May 23, 2023. [Online]. Available: https://www.kaggle.com/datasets/paultimothymooney/breast-histopathology-images
[32] S. Sarkar, M. Ghosh, and P. Dutta, "Breast cancer detection and diagnosis: A review of recent trends," *IEEE Syst. J.*, vol. 11, no. 1, pp. 12–24, Apr. 2017.
[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[34] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 630–645.
[35] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Boca Raton, FL, USA: CRC Press, 1984.
[36] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[37] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, Apr. 2006.
[38] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of online learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.
[39] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.
[40] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001.

SELINA SHARMIN (Member, IEEE) received the B.Sc. (Hons.), M.S., and Ph.D. degrees in directional wireless sensor networking from the Department of Computer Science and Engineering, University of Dhaka, Bangladesh, in 2007, 2009, and 2018, respectively.

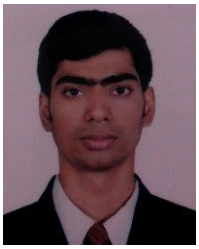
She is currently an Associate Professor with the Department of Computer Science and Engineering, Jagannath University, Bangladesh. Her current research interests include wireless sensor networks, directional wireless sensor networks, ad hoc networks, coverage problems, mac protocols, performance analysis, cloud computing, and machine learning. She is a member of the Green Networking Research (GNR) Group and the Director of Bangladesh Women in Technology (BWIT).

TANVIR AHAMMAD received the B.Sc. and M.Sc. degrees in computer science and engineering from Jagannath University, Dhaka, Bangladesh, in 2013 and 2015, respectively.

From 2011 to 2013, he was an Undergraduate Research Fellow with Jagannath University, under the Re-Engineering of the Information Technology Research and Resource Center (ITRRC) Project, which was funded by HEQEP, UGC, GoB, and the World Bank. His outstanding performance in the role earned him recognition for his research fellow activities. He is currently a Lecturer with the Department of Computer Science and Engineering, Jagannath University. His research interests encompass data mining, text mining, machine learning, and cloud computing. He has published four indexed journal articles and five conference papers, contributing to the academic community with his research findings.

MD. ALAMIN TALUKDER received the B.Sc. degree in computer science and engineering from Jagannath University, Dhaka, Bangladesh, in 2022, where he is currently pursuing the M.Sc. degree in computer science and engineering.

He is a Research Assistant with the School of Information Technology, Deakin University, Australia, and the Department of CSE, Jagannath University. He is a passionate data scientist and researcher specializing in machine learning, deep learning, and AI-based decision-making systems. His work focuses on developing interpretable and clinically relevant models to boost productivity and efficiency. His research interests encompass neuroscience, natural language processing, security, and big data analysis. He received several awards and honors, including the ICT Research Fellow (ICT Division in Bangladesh) and Merit Scholarship (Department of CSE, Jagannath University).



PARTHO GHOSE received the B.Sc. and M.Sc. degrees in computer science and engineering from Jagannath University, Dhaka, Bangladesh.

He is currently a Lecturer with the Department of Computer Science and Engineering, Bangladesh University of Business and Technology (BUBT). He is proficient in prominent frameworks, including TensorFlow, Pytorch, NumPy, Matplotlib, and so on. He is passionate about learning fundamental computer science algorithms, including data structures, dynamic programming, and game theory. He has experience working in programming languages, including Python, Java, C, and C++. His research interests include deep learning algorithms, computer vision, machine learning, and health informatics.