



## A novel framework for Breast cancer scoring based on machine learning technique using Immunohistochemistry images

Hasanain Hayder Razzaq<sup>1,2,\*</sup>, Rozaida Ghazali<sup>1</sup>, Loay E. George<sup>3</sup>, and Muhammad Zulqarnain<sup>4</sup>

<sup>1</sup>Faculty of Computer Science and Information Technology, University Tun Hussein Onn Malaysia, Batu Pahat, 86400, Johor, Malaysia.

<sup>2</sup>College of Medicine, Jabir ibn Hayyan Medical University, Najaf, Iraq.

<sup>3</sup>University of Information Technology and Communication (UoITC), 10001, Baghdad, Iraq.

<sup>4</sup>Department of Computer Science IT, Cholistan University of Veterinary and Animal Sciences, Bahawalpur, Punjab, Pakistan.

### Abstract

Breast cancer is classified as a serious disease in the medical field and there is no doubt that breast cancer detection requires effective and accurate techniques. Integrating deep learning (DL) and machine learning (ML) methods has shown promising results in this area. In this research, we introduced a hybrid approach for breast cancer diagnosis which centered on the analysis of immunohistochemical images. The proposed method encompasses algorithms for image pre-processing, segmentation, extracting informative indicators (such as relative cell area and intensity), and an algorithm for categorizing the molecular harmonic subtype of breast cancer. The number of the sample was 598 divided into training 70% and testing 30%. The 5-fold cross-validation was used to assess the proposed approach. Experimental results showcased the effectiveness of the proposed hybrid method in achieving superior performance in the detection of breast cancer, especially within breast cancer scoring systems. The accuracy of our proposed approach, which involved combining HSV integration with adaptive high boost filtering, reaches a peak at 96.5% when using SVC (linear kernel). Moreover, the precision, recall, F1-score, and specificity metrics are recorded at 95.29%, 99.99%, 95.59%, and 99.28%, respectively. Additionally, this study evaluated the efficacy of the proposed model in comparison to various other traditional breast cancer detection approaches.

**Keywords.** Breast cancer, Immunohistochemistry, Segmentation, Enhancement, RGB, HSV..

**2010 Mathematics Subject Classification.** 65L05, 34K06, 34K28.

### 1. INTRODUCTION

Breast cancer is a widespread and potentially fatal disease that impacts numerous individuals on a global scale [? ]. This disease manifests when cells within breast tissue exhibit abnormal growth, culminating in the formation of tumors. Detecting breast cancer promptly ensures that patients receive prompt and appropriate treatment, which greatly improves their chances of survival [? ]. As per international health data, breast cancer stands as the most frequently diagnosed cancer among women and contributes significantly to cancer-related fatalities worldwide [? ]. The mortality rates linked to breast cancer highlight the urgent requirement for reliable methods of detection and diagnosis [? ]. Consequently, the development of precise and effective strategies for detecting breast cancer holds paramount significance in the realm of medical research. Immunohistochemistry stands out as a valuable tool for the preoperative diagnosis of breast cancer. This technique is often referred to as IHC. IHC serves multiple purposes including the identification of specific protein markers, the observation of distinctive characteristics associated with breast cancer, the differentiation of breast cancer from other types of cancer, the detection of hormonal alterations, the exploration of prognostic factors, and the guidance of effective treatment strategies [? ]. Moreover, with technical advancements, IHC has become a convenient diagnostic method integrated into advanced protocols alongside other histochemical approaches. Ensuring standardization and maintaining quality control are vital prerequisites for the

Received: 24 August 2024 ; Accepted: 04 October 2024.

\* Corresponding author. Email: hasanain.hayder@jmu.edu.iq .

reliable application of IHC in the context of breast cancer diagnosis. Nevertheless, several practical challenges persist, such as the absence of a single, comprehensive, publicly available IHC breast cancer dataset [? ], the limited information within IHC images, and the presence of substantial background noise [? ], as well as the issue of overlapping cells in human tissue histopathologic images, especially when using the IHC technique [? ]. In recent years, significant advancements have occurred in breast cancer detection attributed to the adoption of DL and ML methodologies [? ]. DL uses an artificial neural network mechanism with several layers to autonomously capture intricate classified data illustrations [? ]. Transfer Learning (TL) emerges as a crucial aspect of DL. where (TL) made a fundamental revolution in the medical arena by utilizing pre-trained methods trained on extensive datasets for particular tasks [? ? ? ]. These pre-trained approaches have garnered profound insights from vast data, making them valuable resources for feature extraction and feature selection in diverse fields, such as skin cancer, brain tumor, and breast cancer detection [? ]. Leveraging the knowledge embedded in these pre-trained approaches facilitates the acceleration of the training process, enhancement of approach performance, and overcoming challenges associated with limited data. Moreover, using pre-trained approaches allows for a more focused refinement of the approach for the exact task at hand, resulting in a more accurate and precise breast cancer detection process. The primary hurdles in breast cancer image analysis involve the automatic identification and pinpointing of cancerous cells, as they exhibit significant variability in terms of their size, shape, and location. Furthermore, breast images may contain other anomalies, such as mastitis, adenopathy, and granuloma [? ]. Machine learning (ML) techniques have been widely applied across diverse domains, including predicting outcomes in education [? ? ? ? ? ], forecasting bankruptcy, recognizing patterns [? ? ? ? ], manipulating images, reducing feature dimensions, diagnosing faults [? ? ? ], identifying faces and micro-expressions, processing natural language, and assisting in medical diagnosis [? ? ]. Notably, Significantly, machine learning (ML) has demonstrated significant potential in the realm of breast cancer diagnosis. In the past few decades, various strategies have been put forth by scholars to tackle the issue of automated cell classification in detecting breast cancer where, some investigators have concentrated on analyzing cell nuclei, extracting features that provide crucial information for distinguishing between benign and malignant cells [? ]. Furthermore, specific approaches have utilized algorithms based on groupings, such as the circular Hough-transform and a range of statistical features, to aid in nucleus segmentation and classification. This research tries to deal with biomedical images, due to their unique characteristics which require specific handling. where initially, these images are stored in their original format within the memory system. Subsequently, an initial processing stage is carried out, encompassing tasks such as parameter identification, adaptive filtering, and adjustments to brightness and contrast. Microscope-derived images often exhibit noise and lack precise cell nucleus contours, necessitating the development of preprocessing algorithms. The primary objective at this preprocessing stage is to eliminate impulse noise and normalize the histogram. The outcome of this process leads to immunohistochemical images, which are then subjected to artificial intelligence-based diagnostic procedures.

## 2. METHODOLOGY

The methodology of this research focused on the construction of a hybrid and robust technique for breast cancer detection. This technique combines machine learning techniques with a feature extraction strategy to enhance the precision and dependability of breast cancer diagnosis. The purpose of this study is to develop a new framework for breast cancer scoring by using the first public IHC dataset images. Figure ?? illustrates the three main stages of the proposed method. The first step involves converting the image's color space from RGB to HSV process to enhance the image quality through the adaptive high boost filtering technique, which is applied to the S channel to improve the images contained and image brightness equalization. it involves the extraction of labeled masks from the original Whole Slide Images (WSIs) with corresponding labels. tumor patches and normal patches are randomly generated based on these label masks. Then, these patches are input into a machine learning approach to enhance binary classification as described in reference [? ]. Secondly, Gaussian blur with a threshold in conjunction with morphological operations to eliminate the problem of overlapping cells. Subsequently, binary tumor predictions are derived from this probability map using a specified threshold. The third stage of the process is done by examining the image with hybrid feature extraction in the individual image [? ? ? ? ? ]. The dynamical systems have been extensively used in real-world applications, especially biology, ecology, physics, economics, and sociology [? ? ? ? ? ]



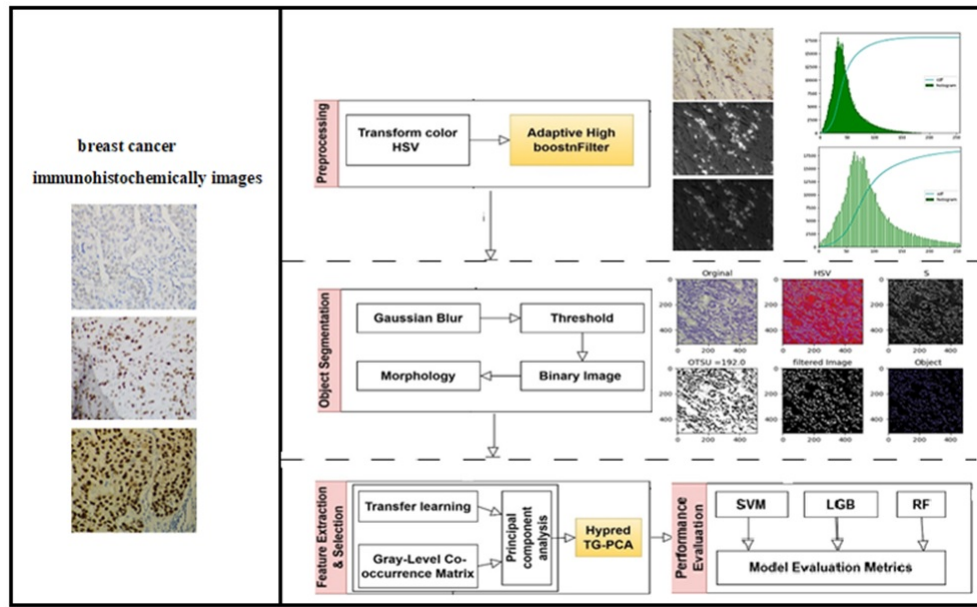


FIGURE 1. The proposed framework for breast cancer detection.

? ]. Many researchers have applied various techniques to obtain bright and dark solitons for different equations [? ? ? ? ? ]. Furthermore, the test WSI is categorized into eight subclasses: IHC 0, -2, then +3 to +8. This score aims to provide an accurate and interpretable result by employing comprehensive assessments. The procedure encompasses critical phases, including data preprocessing, feature extraction, approach training, and the evaluation of performance metrics. Finally, the proposed hybrid approach merged the feature extraction with the Principal Component Analysis (PCA) technique to generate the final result.

**2.1. Dataset acquisition.** This approach involves the acquisition and capture of an Immunohistochemistry (IHC) dataset, which forms the foundational data for training and assessing the experimental approaches. This phase involves the following key steps: Initially, the process involves the identification and selection of an appropriate dataset covering pertinent information on breast cancer HPI. The Independent Breast Cancer Immunohistochemical Images dataset gathered from patients of Iraqi origin focuses on the influence of family history and the significant role of genetic factors in breast cancer susceptibility. IHC breast cancer image collected from the Laboratories of Al Sader Hospital in Iraq in cooperation with Jabir Ibn Hayyan Medical University, University of Kufa, and University Tun Hussein Onn Malaysia (UTHM). The acquisition of images at different magnifications is performed using the Olympus MT system microscope with a relay lens with magnification of "3 × 3 coupled to a CANON EOS KISS X7i digital were used to obtain digitized images from the breast tissue slides. The dataset started by pathologists identifies tumour slides and defines a region of interest (ROI). The pathologist preferentially selects images with a single type of tumour (most cases), but some of these images also include transitional tissue, e.g., normal pathological. On average, a total of 8 images per patient is captured from each slide as shown in Table ?? . The collected dataset is composed of 598 images divided into benign and malignant tumors. Table ?? summarizes the image distribution. Both breast tumors, benign and malignant," can be sorted into different types based on the aspect of the tumoral cells under the microscope.

On average, nine images are obtained from each slide for each patient. Table ?? provides a summary of the distribution of these images. Both benign and malignant breast tumors can be categorized into different subtypes based on the characteristics of the tumors cells observed under the microscope.

**2.2. Pre-processing.**



TABLE 1. Dataset Composition.

Property	RGB
Width	5184 pixels
Height	3456 pixels
Bit depth	24
No. of images	598

TABLE 2. Image distribution by magnification factor and classes.

Magnification	Positive images	Negative images	Total images
400x	568	30	598
# Patients	78	78	78

Several pre-processing techniques were employed to prepare the dataset for training a machine learning classifier, including image resizing, sharpening, normalization of pixel values, and label encoding.

### 2.3. Dataset split.

The performance of the approach was evaluated using a fivefold cross-validation method. This involves partitioning the dataset into training and test sets, ensuring that the approaches are trained and tested on distinct data subsets. This approach enhances the reliability of assessing their ability to generalize.

### 2.4. Feature extraction.

The GLCMS approach employed in this research underwent prior training on a comprehensive dataset to extract significant features from images of breast cancer. For the process of feature extraction and segmentation, the research applied the VGG19 technique. Then, features extracted through the VGG19 technique were applied to many machine learning classifiers. This pre-trained approach has acquired valuable representations that enable identification of breast cancer. Finally, the proposed hybrid approach merged the feature extraction with the Principal Component Analysis (PCA) technique to generate the final result.

### 2.5. Image pre-processing.

Within the mechanism framework of this study for breast cancer detection, the significance of image pre-processing cannot be overstated. Its pivotal role lies in readying input images for successive analysis through a series of operations on the original images. These operations are designed to improve image quality and extract meaningful information. The impact of proficient image pre-processing extends to substantially boosting the performance and accuracy of the consequent analysis. Given that Immunohistochemistry (IHC) typically contains millions of pixels, either annotating or performing computations on IHC at high magnification scales is an exceedingly laborious and time-intensive task [? ]. Upon examination, it becomes evident that IHC consists of tissue regions and white background areas. To streamline the computational process and reduce complexity, it becomes imperative to focus exclusively on the analysis of tissue regions while disregarding the white background areas. In this study, a threshold-based segmentation approach is employed for the automatic identification of the background region. More specifically, this approach starts by converting the original image from the RGB color space to the HSV color space. Subsequently, the Otsu algorithm is utilized to determine the optimal threshold for each channel. this method led to create a final mask image by merging the masks of the S channels. From what was obtained, the Otsu method can effectively filter out approximately 75% of the background area, significantly enhancing computational efficiency. 598 IHC images were collected from the training set. where these images were used to train and assess the performance of the hybrid framework in detecting lesions. The training and validation datasets were divided into annotated patches using the sci-kit-learn package, which comprised 410 images for training and 179 images for testing. The proposed approach uses multiple image preprocessing methods to prepare the IHC images for breast cancer detection. These methods include:

**Image Resizing:** Standardized the images to a uniform size ( $256 \times 256$  pixels) in patches, then randomly allocated 179 labeled IHC to the training and testing sets. This procedure is essential when using pre-trained machine learning



classifiers, given their often-specific size prerequisites. Adjusting the images to conform with the approach’s input size allows for optimal use of its acquired features. Additionally, the uniformity in image size enhances computational efficiency in both the training and inference phases, simplifying the entire process.

**Transforming Color Space:** Converted the initial IHC images from the RGB (Red-Green-Blue) color space to the more commonly used HSV (Hue-Saturation-Value) color space. This alteration ensures uniform color representation throughout the dataset, making it more suitable for subsequent image analysis tasks. Additionally, utilized an image sharpening filter to accentuate image edges and fine details, enhancing feature contrast and aiding in subsequent feature extraction.

**Standardizing Image Value:** To standardize pixel values and ensure their alignment within a consistent numerical range conducted image scaling by dividing “pixel values by 255. This normalization process guarantees that pixel values adhere to a standardized range, facilitating the convergence of subsequent machine-learning algorithms.

**Numeric Labeling of Images:**The initial IHC dataset featured object labels presented as text. To enhance subsequent analysis, transformed these textual object labels into numerical counterparts. This conversion allows the application of diverse machine learning algorithms, especially those that necessitate numeric labels for classification tasks. Applying the above methods, this study leads to improving the precision and dependability of breast cancer detection systems, and readying IHC images for thorough feature extraction and analysis through machine learning.

### 3. EXPERIMENTAL RESULTS AND ANALYSIS

This section presents the experiments and the analysis results of the proposed hybrid approach for breast cancer detection. This research thoroughly designed the experimental framework and examined various metrics such as accuracy, sensitivity, specificity, and F1 score. These metrics served as crucial benchmarks to assess the effectiveness of the proposed approach. where scrutinized accuracy was done to gauge the overall correctness of the classification results, and evaluated sensitivity to measure the proposed method’s proficiency in identifying positive instances. At the same time, utilized specificity to gauge the method’s accuracy in identifying negative instances, and considered the F1 score for a balanced evaluation. These metrics yielded valuable insights into the approach’s performance, enabling informed assessments and comparisons with other existing methods in breast cancer detection.

#### 3.1. Environmental setup.

The research simulations were conducted using an Intel Core i7-3770 CPU running at 3.40 GHz, DDR3 memory, and 32 GB of RAM. The operating system used was Windows 10. The development environment consisted of Python 3.11 compiler along with the high-level NNs API-Keras. The necessary libraries included Tensor Flow 1.14 and Keras 2.3 frameworks were used to harness the power of deep and machine learning. The research environment was carefully configured with all the necessary dependencies and libraries, especially those tailored for pre-trained methods used in transfer learning. This streamlined environment played a pivotal role in ensuring the efficiency of the experiments, thereby simplifying the investigation and validation of the proposed hybrid mechanism for diagnosing breast cancer. the proposed approach was trained by using various parameter settings for both feature engineering and classification. In the realm of feature extraction, the hybrid approach was fine-tuned with a specific categorization: include\_top = False, weights = imagenet’, input\_tensor = None, input\_shape = (256, 256, 3), and pooling = None. This setting of ‘size’ at 256 indicated that the input images were resized to a square dimension of  $256 \times 256$  pixels. Furthermore, the primary focus was on two crucial parameters: the number of estimators and the hyperparameters, which are instrumental in managing the performance of the approaches.

#### 3.2. Evaluation metrics.

Various experiments were conducted to assess the performance of the developed hybrid approach on Immunohistochemistry Images. the efficiency of the approach was measured by using various evaluation metrics for two classes, which are referred to in Equation (??) to (??). The effectiveness of the developed method was evaluated by employing these diverse metrics, each one was designed to gauge various characteristics of the approaches’ effectiveness. The metrics are formulated as follows:

**Confusion Matrix:** Table ?? presents the confusion matrix, which dissects the model’s predictions into positive and negative classes, showcasing True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives



TABLE 3. Confusion Matrix.

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

(FN). TP indicates the number of correctly predicted positive instances, while TN represents the count of accurately predicted negative instances. FP enumerates instances incorrectly labeled as positive, and FN quantifies instances inaccurately labeled as negative. This matrix provides a thorough perspective on the approach’s performance, offering insights into both accurate and inaccurate predictions for positive and negative cases.

**Accuracy:** The approach’s overall precision is established by the proportion of accurately classified instances (TP) to the total number of instances (TP + FP) within the complete dataset, as outlined in Equation (??). However, relying solely on accuracy might not provide a thorough assessment of the approach’s effectiveness.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (3.1)$$

**Precision and Recall:** Precision and recall are among the primary assessment parameters employed to assess the precision of the developed method. Precision indicates the percentage of correctly identified positive instances, while recall or sensitivity measures the quantity of true positive occurrences (TP) relative to the total number of actual instances (TP + FN).

These metrics are described by Equations (??) and (??) as follows:

$$Precision = \frac{(TP)}{(TP + FP)} \quad (3.2)$$

$$Recall = \frac{(TP)}{(TP + FN)} \quad (3.3)$$

**F1 Score:** It computes the weighted average of precision and recall (also known as sensitivity), making it a common tool for optimizing networks with a focus on either precision or recall. Equation (??) defines the F1 score as:

$$F1 - Score = 2 * \frac{(PrecisionRecall)}{(Precision + Recal)} \quad (3.4)$$

#### 4. RESULTS ANALYSIS

The performance of several machine learning algorithms was analyzed that were used in the detection of breast cancer. A dependability analysis was carried out to assess the reliability and consistency of these algorithms. This evaluation scrutinized performance constancy across different iterations or datasets, aiming to mitigate the impact of random variations or dataset-specific traits on the results. Through a careful examination of the approaches’ capacity for generalization, obtained valuable insight into their performance resilience and pinpointed potential limitations or biases. The analysis of reliability empowers to make informed decisions about the effectiveness of these approaches in real-world scenarios. The proposed framework undergoes testing on 420 images sourced from the dataset, comparing the results with state-of-the-art approaches. These tested images, spanning PR and ER images with scores ranging from 0 to 8, are juxtaposed with pathologist-assigned scores, deemed exact by the pathologist. Accuracy for both total and PS scores is determined. where PS values are considered particularly crucial as they directly result from pre-processing and segmentation, and by adding the intensity score can have the total score for diagnosis of breast cancer. After conducting experiments, we evaluated the effectiveness of our hybrid approach by assessing its performance through various metrics such as PSNR, SSIM, MSE, and histogram.

The results are depicted in Figure ??.

A novel segmentation framework was developed to distinguish the cell from the background. Initially, the AHBF was applied to the S-Band, followed by the recombination of the three bands in the HSV model. Subsequently, a



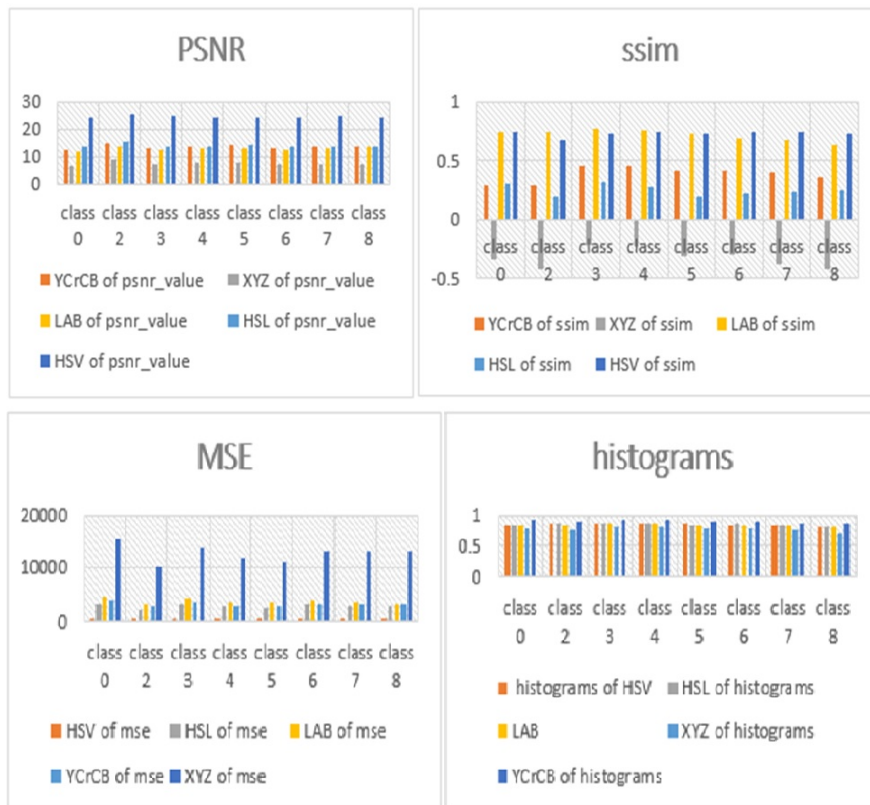


FIGURE 2. . Results of developed approaches in the terms of PSNR, ssim, MSE and histograms.

TABLE 4. Results of classifiers with the GLCM technique.

Classifiers with GLCM	Accuracy	Precision	Recall	F1-score	Specificity
LGBM	78.53	82.21	76.81	77.82	96.82
RF	77.35	79.78	78.16	78.01	96.68
SVM	72.94	74.92	70.96	71.76	96.2
SVC (kernel-linear)	75.88	78.13	74.88	75.76	96.46
SVC (kernel-sigmoid)	75.29	80.03	73.35	75.46	96.27

Gaussian filter was employed on the amalgamated image. The Otsu threshold was then used to generate a binary image. Following this, morphological operations, namely erosion and dilation, were applied to the preceding image. Additionally, an area was implemented between these morphological operations. Lastly, our adaptive threshold was applied to the concluding image. Moreover, the performed experiment has revealed outcomes about conventional methods employing GLCM and VGG19 techniques. Our analysis indicates that the classifiers demonstrated superior performance in terms of accuracy, precision, recall, F1-score, and specificity when utilizing the VGG19 feature extraction technique. Table ??, demonstrates the results of accuracy, precision, Recall, F1-score, and specificity with the GLCM feature extraction technique, while Table ??, shows the results of the suggested classifier with the VGG19 feature extraction technique.

#### 4.1. Cross validation analysis.



TABLE 5. Results with VGG19 technique.

Classifiers with GLCM	Accuracy	Precision	Recall	F1-score	Specificity
LGBM	83.82	86.66	80.43	82.54	97.56
RF	80.22	82.51	78.46	80.24	97.04
SVM	74.41	73.77	71.36	72.14	96.19
SVC (kernel-linear)	77.65	79.34	74.20	75.99	96.69
SVC (kernel-sigmoid)	76.47	78.71	77.23	77.72	96.50

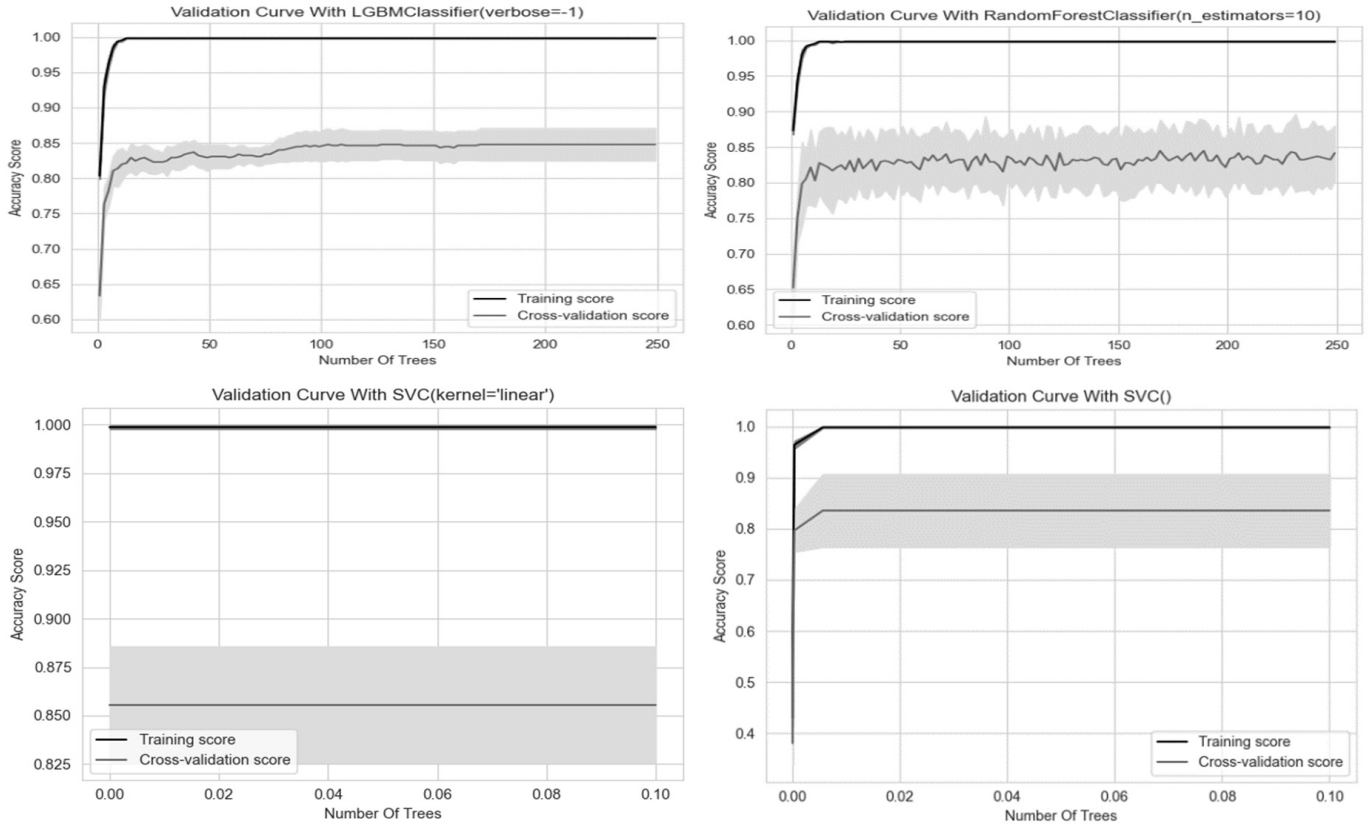


FIGURE 3. The proposed framework for breast cancer detection.

5-fold cross-validation is a widely employed method in the assessment of machine and deep learning approaches. The procedure involves dividing the training dataset into 5 subsets or folds. One-fold was designated as a validation set in each iteration and the remaining folds were for training. This iteration was repeated five times, enabling a thorough evaluation of the approach’s performance.

Additionally, averaging the results from these iterations provides a reliable estimate of the approach’s generalization ability and helps identify issues like overfitting. In this research, a value of k equal to 5 was selected for the 5-fold cross-validation, resulting in the dataset being split into ten roughly equal subsets. Subsequently, the training process was carried out iteratively, with each of the ten subsets taking turns as the validation set, while the remaining nine subsets were utilized for training. Figure ?? illustrates the cross-validation of the developed technique. The breast cancer scoring performance of the developed hybrid technique and traditional machine learning classifiers was assessed using numerous assessment parameters such as precision, recall, F1-score, and specificity. Precision measures the accurate





TABLE 6. Breast cancer scoring results (%) of proposed and existing classifiers on the Immunohistochemistry Images dataset.

Classifiers with GLCM	Accuracy	Precision	Recall	F1-score	Specificity
LGBM	94.10	97.59	99.98	98.78	99.64
RF	94.25	88.04	99.92	93.64	98.04
SVM	89.44	97.59	98.32	97.51	99.64
SVC (kernel-linear)	96.5	95.29	99.99	95.59	99.28
SVC (kernel-sigmoid)	77.9	82.75	88.88	85.71	97.33

TABLE 7. Computational time &amp; accuracy of the proposed method compared to the state-of-the-art approaches.

Author(S),(Year), Ref.	Time(s)	Accuracy	#Images
Mouelhi et al., 2013, [? ]	80	55	120
Mouelhi et al., 2018, [? ]	77	56	84
Mahanta et al., 2021, [? ]	6.25	81	40
Our proposed method	2.5	96.5	598

identification of true positive observations within the positive instances. Recall measures the ability to classify all actual positive instances correctly. To calculate the final precision and recall values, these measures were individually computed for each class and then aggregated the ratios across all classes. The test accuracy was assessed using the F-measure, which is a combination of test precision and recall. In this section, the performance of the suggested hybrid classifiers is presented and compared to existing approaches in Table ???. On the other hand, Table ??? presents the experimental results, indicating that our hybrid approach outperforms existing studies in both score accuracy and time complexity. This superior performance is attributed to the creation of valuable features, which enrich the dataset and consequently diminish the significance of the original dataset.

## 5. CONCLUSION

This paper presents a novel machine-learning method aimed at improving the classification accuracy of breast cancer in the Immunohistochemistry Images dataset. The main goal of the proposed approach is to assist health-care professionals in identifying and diagnosing breast cancer. To achieve the objective of this study leveraged the first publicly available IHC dataset images and established a new framework for breast cancer scoring by collecting immunohistochemistry images. The proposed approach is divided into four distinct phases. In the initial stage, it performs a color space conversion from RGB to HSV for the input image. Subsequently, the dataset images are enhanced using the AHBF within the S channel. Secondly, we proposed an enhanced hybrid feature extraction method to address the issue of overlapping in the improved image. This approach integrates a segmentation algorithm with hybrid feature extraction, allowing for the calculation of areas and the detection of cell staining intensity. The third stage examined the image with several classifiers. To assess the performance of the proposed approach for breast cancer diagnosis, this approach utilized five evaluation metrics: accuracy, precision, recall, specificity, and F1-score. experimental findings demonstrate that the capabilities of a hybrid approach based on feature extracting mechanisms are highly effective in classifying microscopy breast cancer images. The experimental results of the developed hybrid approach have demonstrated that our framework outperformed existing state-of-the-art approaches.

## CONFLICTS OF INTEREST

The authors declares no conflict of interest.



## AUTHOR CONTRIBUTIONS

Conceptualization, H. H. Razzaq and R. Ghazali; methodology, H. H. Razzaq; software, M. Zulqarnain; validation, and L. E. George; formal analysis, M. Zulqarnain; investigation, R. Ghazali; resources, L. E. George; data curation, H. H. Razzaq; writing–original draft preparation, R. Ghazali; writing–review and editing, R. Ghazali; supervision, R. Ghazali; project administration, H. H. Razzaq; funding acquisition, M. Zulqarnain.

## REFERENCES

- [1] D. Alhajim, H. Fahama, and T. Mohammed, *Enhancing and securing a real-time embedded face recognition system using Raspberry Pi*, J. Al-Qadisiyah Comput. Sci. Math., *16*(1) (2024), 92-104.
- [2] W. Al-Yaseen, A. Jehad, Q. Abed, and A. Idrees, *The Use of Modified K-Means Algorithm to Enhance the Performance of Support Vector Machine in Classifying Breast Cancer*, Int. J. Intell. Eng. Syst., *14*(2), 2021.
- [3] M. Ashrafzadeh, 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., *2023* (2023), 1-12.
- [4] A. Budianto, S. Andarini, T. Hariyanti, and N. Muslihah, *The prediction of colorectal cancer: Perspective of smoking and socioeconomic influence of culture in East Java with SEM analysis*, Nurture, *18*(2) (2024), 315-327.
- [5] P. Chhabra and S. Goyal, *A thorough review on deep learning neural network*, in Proc. Int. Conf. Artif. Intell. Smart Commun. (AISC), (2023), pp. 220-226.
- [6] P. Donahue, A. MacKenzie, A. Filipovic, and L. Koelmeyer, *Advances in the prevention and treatment of breast cancer-related lymphedema*, Breast Cancer Res. Treat., *200*(1) (2023), 1-14.
- [7] P. Ghahremani, Y. Li, A. Kaufman, R. Vanguri, N. Greenwald, M. Angelo, and S. Nadeem, *Deep learning-inferred multiplex immunofluorescence for immunohistochemical image quantification*, Nat. Mach. Intell., *4*(4) (2022), 401-412.
- [8] Global World Health Organization, *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, 2023.
- [9] N. Goodarzi, N. Chalakinia, and A. Talaei, *Clarifying the Role of Emotional Self-regulation Strategies and Health Anxiety in Predicting Negative and Positive Reactions to Stress, Psychosomatic Symptoms, and Quality of Life Indicators in Women with Breast Cancer: Clarifying the role of emotional self-regulation strategies in women with breast cancer*, Int J Body Mind Cult., *11*(2), 118-126.
- [10] Y. Gu, S. Malmir, J. Manafian, O. A. Ilhan, A. Alizadeh, and A. J. Othman, *Variety interaction between k-lump and k-kink solutions for the (3+ 1)-D Burger system by bilinear analysis*, Results in Physics, *43* (2022), 106032.
- [11] G. Hiremath, J. Mathew, and N. Boraiah, *Hybrid Statistical and Texture Features with DenseNet 121 for Breast Cancer Classification*, Int. J. Intell. Eng. Syst., *16*(2), 2023.
- [12] D. M. Isroilovna, H. A. Altememy, A.M. Jabbar, M. A. Majeed, H. J. Jumaa, S. S. Sadeq, and et al. *The Effects of Cognitive Behavioral Therapy on the Lifestyle of Breast Cancer Patients: Effects of CBT on the lifestyle of patients with breast cancer*, Int J Body Mind Cult. *10*(4) (2023), 531-541.
- [13] M. Kadhim and A. Radhi, *Heart disease classification using optimized machine learning algorithms*, Iraqi Journal For Computer Science and Mathematics, *4*(2) (2023), 3-42.
- [14] A. A. Kassem, *Snake Optimization with deep learning enabled disease detection model for colorectal cancer*, Journal of Smart Internet of Things, *2022*(1) (2023), 178-195.
- [15] W. Liu, X. Chang, L. Chen, D. Phung, X. Zhang, Y. Yang, and A. Hauptmann, *Pair-based uncertainty and diversity promoting early active learning for person reidentification*, ACM Trans. Intell. Syst. Technol., *11*(2) (2020), 1-15.
- [16] J. Luo, H. Chen, Q. Zhang, Y. Xu, H. Huang, and X. Zhao, *An improved grasshopper optimization algorithm with application to financial stress prediction*, Appl. Math. Model., *64* (2018), 654-668.



- [17] E. Mahanta, N. Hussain, N. Das, L. Kakoti, and M. Chowdhury, *IHC-Net: A fully convolutional neural network for automated nuclear segmentation and ensemble classification for Allred scoring in breast pathology*, Applied Soft Computing, *103* (201), 107136.
- [18] J. Manafian and M. Lakestani, *Application of  $\tan(\phi/2)$ -expansion method for solving the Biswas–Milovic equation for Kerr law nonlinearity*, Optik, *127*(4) (2016), 2040-2054.
- [19] J. Manafian and M. Lakestani, *Optical soliton solutions for the Gerdjikov–Ivanov model via  $\tan(\phi/2)$ -expansion method*, Optik, *127*(20) (2016), 9603-9620.
- [20] J. Manafian, and M. Lakestani, *Abundant soliton solutions for the Kundu–Eckhaus equation via  $\tan(\phi/2)$ -expansion method*, Optik, *127*(14) (2016), 5543-5551.
- [21] J. Manafian and M. Lakestani, *N-lump and interaction solutions of localized waves to the  $(2+ 1)$ -dimensional variable-coefficient Caudrey–Dodd–Gibbon–Kotera–Sawada equation*, Journal of Geometry and Physics, *150* (2020), 103598.
- [22] J. Manafian, *An optimal galerkin-homotopy asymptotic method applied to the nonlinear second-order bvps*, Proc Inst Math Mech, *47*(1) (2021), 156-182.
- [23] M. Lakestani, J. Manafian, A. R. Najafzadeh, and M. Partohaghighi, *Some new soliton solutions for the nonlinear the fifth-order integrable equations*, Computational Methods for Differential Equations, *10*(2) (2022), 445-460.
- [24] J. Manafian, L. A. Dawood, and M. Lakestani, *New solutions to a generalized fifth-order KdV like equation with prime number  $p = 3$  via a generalized bilinear differential operator*, Partial Differential Equations in Applied Mathematics, *9* (2024), 100600.
- [25] J. Manafian and N. Allahverdiyeva, *An analytical analysis to solve the fractional differential equations*, Adv Math Models Appl, *6*(2) (2021), 128-61
- [26] V. 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., *21* (2023), 1372-1382.
- [27] E. Morrison, R. Lefever, N. Lewis, J. Kapadia and R. Bauer, *Conventional histological and cytological staining with simultaneous immunohistochemistry enabled by invisible chromogens*. Laboratory Investigation, *102*(5) (2022), 545-553.
- [28] A. Mouelhi, M. Sayadi, F. Fnaiech, K. Mrad, and K. B. Romdhane, *Automatic image segmentation of nuclear stained breast tissue sections using color active contour model and an improved watershed method*, Biomedical Signal Processing and Control, *8*(5) (2013), 421-436.
- [29] A. Mukhlif, B. Al-Khateeb, and M. Mohammed, *Classification of breast cancer images using new transfer learning techniques*, Iraqi Journal For Computer Science and Mathematics, *4*(1) (2023), 167-180.
- [30] A. Mouelhi, H. Rmili, J. Ali, M. Sayadi, R. Doghri, and K. Mrad, *Fast unsupervised nuclear segmentation and classification scheme for automatic allred cancer scoring in immunohistochemical breast tissue images*. Comput. Methods Programs Biomed., *165* (2018), 37-51.
- [31] S. H. M. Mohammed, and A. Çinar, *Lung cancer classification with Convolutional Neural Network Architectures*, Qubahan Academic Journal, *1*(1) (2021), 33-39.
- [32] S. K. Mohapatra, M. Abebe, L. Mekuanint, S. Prasad, P. K. Bala, and S. K. Dhala, *Pneumonia and tuberculosis detection with chest x-ray images and medical records using deep learning techniques*, Review of Computer Engineering Research, *10*(4) (2023), 136-149.
- [33] Y. Pang, Y. Wei, and C. Kartsonaki, *Associations of adiposity and weight change with recurrence and survival in breast cancer patients: A systematic review and meta-analysis*, Breast Cancer, *29*(4) (2022), 575-588.
- [34] M. M. Salih, A. M. Abdulazeez, and O. M. S. Hassan, *Gender Classification Based on Iris Recognition Using Artificial Neural Networks*, Qubahan Academic Journal, *1*(2) (2021), 156-163.
- [35] R. Saturi and P. Premchand, *Multi-objective feature selection method by using ACO with PSO algorithm for breast cancer detection*, Int. J. Intell. Eng. Syst., *14*(5), (2021).
- [36] R. Shamsi, J. Manafian, and S. Esmaeili, *Ranking extreme efficient decision making units in stochastic dea*, Adv Math Models Appl, *7*(1) (2022), 38-43.



- [37] M. Talukder, M. Islam, M. Uddin, A. Akhter, M. Pramanik, S. Aryal, M. Almoyad, K. Hasan, and M. Moni, *An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning*, Expert Syst. Appl., 230 (2023), 120534.
- [38] R. Torrisi, E. Marrazzo, E. Agostinetti, R. De Sanctis, A. Losurdo, G. Masci, C. Tinterri, and A. Santoro, *Neoadjuvant chemotherapy in hormone receptor-positive/HER2-negative early breast cancer: When, why and what?*, Critical Reviews in Oncology/Hematology, 160 (2021), 103280.
- [39] M. Wang, H. Chen, H. Li, Z. Cai, X. Zhao, C. Tong, J. Li, and X. Xu, *Grey wolf optimization evolving kernel extreme learning machine: Application to bankruptcy prediction*, Eng. Appl. Artif. Intell., 63 (2017), 54-68.
- [40] Y. Wei, N. Ni, D. Liu, H. Chen, M. Wang, Q. Li, X. Cui, and H. Ye, *An improved grey wolf optimization strategy enhanced SVM and its application in predicting the second major*, Math. Problems Eng., 2017 (2017), 1-12.
- [41] X. Zhang, M. Fan, D. Wang, P. Zhou, and D. Tao, *Top-k feature selection framework using robust 0-1 integer programming*, IEEE Trans. Neural Netw. Learn. Syst., early access, (2020).
- [42] X. Zhang, T. Wang, J. Wang, G. Tang, and L. Zhao, *Pyramid channel-based feature attention network for image dehazing*, Comput. Vis. Image Understand., 197-198 (2020), 103003.
- [43] X. Zhang, R. Jiang, T. Wang, P. Huang, and L. Zhao, *Attention-based interpolation network for video deblurring*, Neurocomputing, (2020).
- [44] X. Zhang, T. Wang, W. Luo, and P. Huang, *Multi-level fusion and attention-guided CNN for image dehazing*, IEEE Trans. Circuits Syst. Video Technol., early access, (2020).
- [45] X. Zhang, R. Jiang, T. Wang, and J. Wang, *Recursive neural network for video deblurring*, IEEE Trans. Circuits Syst. Video Technol., Early access, (2020).
- [46] Y. Zhang, R. Liu, A. Heidari, X. Wang, Y. Chen, M. Wang, and H. Chen, *Towards augmented kernel extreme learning models for bankruptcy prediction: Algorithmic behavior and comprehensive analysis*, Neurocomputing, 430 (2021), 185-212.
- [47] M. Zhang, X. Xie, J. Manafian, O. A. Ilhan, and G. Singh, *Characteristics of the new multiple rogue wave solutions to the fractional generalized CBS-BK equation*, Journal of Advanced Research, 38 (2022), 131-142.
- [48] D. Zhao, C. Huang, Y. Wei, F. Yu, M. Wang, and H. Chen, *An effective computational model for bankruptcy prediction using kernel extreme learning machine approach*, Comput. Econ., 49(2) (2017), 325-341.
- [49] J. Zhou, L. Luo, Q. Dou, H. Chen, C. Chen, G. Li, Z. Jiang, and P. Heng, *Weakly supervised 3D deep learning for breast cancer classification and localization of the lesions in MR images*, J. Magn. Reson. Imag., 50(4) (2019), 1144-1151.
- [50] W. Zhu, C. Ma, X. Zhao, M. Wang, A. Heidari, H. Chen, and C. Li, *Evaluation of sino foreign cooperative education project using orthogonal sine cosine optimized kernel extreme learning machine*, IEEE Access, 8 (2020), 61107-61123.
- [51] M. Zulqarnain, H. Shah, R. Ghazali, O. Alqahtani, R. Sheikh, and M. Asadullah, *Attention aware deep learning approaches for an efficient stress classification model*, Brain Sci., 13(7) (2023), 994.
- [52] Y. Yang, H. Zhao, L. You, R. Tu, X. Wu, and X. Jin, *Semantic portrait color transfer with internet images*, Multimedia Tools Appl., 76(1) (2017), 523-541.
- [53] C. Yu, M. Chen, K. Cheng, X. Zhao, C. Ma, F. Kuang, and H. Chen, *SGOA: Annealing-behaved grasshopper optimizer for global tasks*, Eng. Comput., 2 (2021), 1-28.

