

Classification of Breast Cancer using Ensemble Filter Feature Selection with Triplet Attention Based Efficient Net Classifier

Bangalore Nagaraj Madhukar
Department of Electronics and
Communication Engineering,
Reva University, India
madhukar.phd@gmail.com

Shivanandamurthy Hiremath Bharathi
Department of Electronics and
Communication Engineering,
Reva University, India
bharathish@reva.edu.in

Matta Polnaya Ashwin
Department of Radio diagnosis and
Imaging, A.J. Institute of Medical Science
and Research Centre, India
Ashwinpolnaya@ajims.edu.in

Abstract: *In medical imaging, the effective detection and classification of Breast Cancer (BC) is a current research important task because of the still existing difficulty to distinguish abnormalities from normal breast tissues due to their subtle appearance and ambiguous margins and distinguish abnormalities from the normal breast. Moreover, BC detection based on an automated detection model is needed, because manual diagnosis faces problems due to cost and shortage of skilled manpower, and also takes a very long time. Using deep learning and ensemble feature selection techniques, in this paper, a novel framework is introduced for classifying BC from histopathology images. The five primary steps of the suggested framework are as follows: 1) to make the largest original dataset and then deep learning model with data augmentation to improve the learning. 2) The best features are selected by an Ensemble Filter Feature selection Method (EFFM) which combines the best feature subsets to produce the final feature subsets. 3) Then the pruned Convolution Neural Network (CNN) model is utilized to extract the optimal features. 4) Finally, the classification is done through the Triplet Attention based Efficient Network (TAENet) classifier. The suggested model produces a 98% accuracy rate after being trained and tested on two different histopathology imaging datasets including images from four different data cohorts. Subsequently, the suggested strategy outperforms the conventional ones since the ensemble filter habitually acquires the best features, and experimental results demonstrate the importance of the proposed approach.*

Keywords: *Breast cancer classification, histopathology image, ensemble filter feature selection, triplet attention based EfficientNet, TAENet.*

Received February 21, 2023; accepted June 19, 2023
<https://doi.org/10.34028/iajit/21/1/2>

1. Introduction

Breast Cancer (BC) is the most frequent malignancy in women globally there were around 2 million new cases in 2018. Various medical/diagnostic imaging techniques, such as mammography, breast ultrasonography, histopathology, and breast Magnetic Resonance Imaging (MRI), are frequently used for BC screening [9]. However, due to both clinical and technological variables, the sensitivity and specificity of all currently available imaging in conducting BC classification and prediction have proven to be relatively limited. The best mode to identify BC is through a medical imaging analysis [8]. While mammography images are generally advised, other imaging modalities such as MRI, digital mammography, ultrasound, and infrared thermography are also employed for diagnosis [27]. High-quality images from mammography can be used to see the internal anatomy of the breast. Mammography is generally said to have a low positive predictive value, sensitivity, and specificity [11]. Ultrasound imaging may be used in conjunction with mammography to

improve efficiency in detecting dense breasts [36].

Women's mortality rates may drop if BC is detected early and treated appropriately [18]. Although a surgical biopsy can determine if a breast lump is malignant or benign, it is more expensive and time-consuming [35]. A breast biopsy is advised if the screening technique reveals that the patient is at risk of developing malignant tissue. Through a biopsy examination, a pathologist can be microscopically evaluated the histological structures within the tissue. The term "histopathology analysis" refers to this process. The distinction between normal (benign) and pathological (malignant) lesions can be made through histopathology analysis. Depending on the patient and type of cancer, surgical treatment, systemic therapy, radiation therapy, and minimally invasive therapies may be used [26]. Artificial Intelligence (AI) for pattern identification now has a competitive advantage, thanks to the development of digital images in medical science. As a result, a system like this decreases the need for human dependency, boosts the rate of diagnoses, and lowers overall treatment costs by lowering false positive and false negative predictions [3, 12].

The majorities of the time, benign conditions are not hazardous to health and cannot be diagnosed as cancer. However, it can be described as a slight difference in the breast's tissue composition [32]. The two subtypes of malignant tumors are in situ and invasive [23]. Because the simple hand-crafted or semi-automatic detections based on past knowledge are incapable of handling complicated shape variations as well as the varied density distribution of the masses and their surrounding tissues, these methods still struggle to handle mass segmentation automatically [29]. In addition, manually selecting and extracting features takes a lot of time. A Convolutional Neural Network (CNN) is a collection of convolutional layers that can extract features that reflect the multiple contexts of an image without the need for feature engineering. Because of this, CNN has emerged as the technique of choice for image interpretation tasks across a wide range of industries, including the identification and categorization of BC [40].

Deep learning techniques have been effectively implemented in numerous fields, particularly in the domain of medical imaging, because of their capacity to automatically extract features. As a result of the complexity of conventional ML procedures like preprocessing, feature extraction, segmentation, and others, the system performance also suffers in terms of efficiency and accuracy [20]. The features still need to be manually fed into the model during training. Deep learning has been introduced to solve this problem and it fully automates this process. The newly developed deep learning approaches can be used to get around common ML problems. This technique can tackle image classification issues by delivering superior feature representation. Because of their much faster pace of growth than benign tumors, malignant tumors are fatal. To treat a patient with BC effectively, early tumor type identification is crucial.

In this research, a new framework for identifying BC on histopathology images is proposed. As far as we are aware, new module that is made by integrating Ensemble Filter Feature Selection Method (EFFSM) with pruned CNN is proposed for making an efficient feature representation, this is the first instance of BC classification using Ensemble Filtering Based Pruned CNN Feature Selection Method (EFPCNNFSM). With the aid of EFFSM, the suggested framework offers an automatic and precise encoding of features from images. Following that, features based on each sub-region are extracted using pruned CNN to pinpoint the breast tumor's position. To acquire more expressive low-dimensional features and to lower the computational cost, CNN is utilized to shrink the original image's dimensions. A novel module named TAENet it is formed by integrating Triplet Attention Technique (TAM) with EfficientNet is proposed for making an efficient classification. The TAM is also used to produce the feature weight distribution. Here, the space attention, channel attention, and center attention

modules are combined to form the triplet attention mechanism. Finally, this study proposes EfficientNet deep learning structure for tumor classification. The experiments are conducted with two different histopathology datasets and the performance of the suggested network is compared with the existing CNN classifier and DenseNet classifier.

The following chapters structured as follows. In section 2, the literature review is done to identify existing challenges and identify the solutions to the existing challenges. The proposed frameworks are elaborated in section 3. Results and discussions are discussed and analyzed in section 4. The Final section 6 elaborated about conclusion.

2. Literature Survey

The deep learning architecture includes CNN models as a subset that was developed for accurately classifying BC. The Deep Belief Network (DBN) is used in a unique patch-based deep learning technique to identify and categorize BC on histopathology images [14]. Using a Support Vector Machine (SVM) classifier with integrated features, an automatic BC classification system is introduced [5]. The survey examines histopathology-based cancer diagnosis. It examines machine learning techniques, deep learning algorithms, and image processing strategies

To categorize the images of histopathology, Shahidi *et al.* [32] utilize several deep learning models. In terms of two, four, and eight categorizations of BC histopathology image datasets, this study determined the best accurate models. The findings for the ImageNet dataset have been found for models like SENet, ResNeXt, DualPathNet, and Neural Search Architecture (NAS) net. In addition, these models were investigated for four classes using the Bank for the Accounts of Companies Harmonized (BACH) database. Numerous studies using various classification and image processing techniques have been conducted on the diagnosis and detection of BC. In Albalawi *et al.* [2] used the Mammographic Image Analysis Society (MIAS) dataset and the CNN classifier to identify BC. CNN has a reputation for being an effective class of approaches for image recognition issues. K-means clustering was used for segmentation, and a Wiener filter was used to remove background noise from the image. According to the comparison results, the CNN classifier outperformed other methods with an additional 0.5-4% accuracy and 3-13% specificity.

An automated radiomic analysis was utilized to distinguish between benign and malignant breast lesions. Large dimensions in the feature data may have an impact on the accuracy and efficiency of cancer classification. Therefore, Ragab *et al.* [30] devised a classification approach for dimension reduction. In particular, a dimension reduction method of Principal Component Analysis (PCA) including calculating the

variance proportion for eigenvector selection, was applied. Each subset of Principle Components (PC) had chosen using a sequential forward approach, random forest naive Bayes, and logistic regression classifiers are trained for the classification technique. Using median values for specificity and sensitivity the approach was created with the help of the random forest classifier which produces the best prediction of benign or cancerous ROIs. In Massafra *et al.* [24] suggest categorizing mammography breast scans according to their training and using interest mastering to pinpoint the distinct pixels of cancer. Features can be extracted from mammography scans using overlay CNN and then input into a recurrent neural community.

Using image segmentation and low-level preprocessing methods, BC can be found. BC automatic diagnosis is compared between the many pre-trained deep learning models with image segmentation using the thresholding technique. To effectively support the instinctive diagnosis and detection of the BC suspicious region based on two methodologies, namely 80-20 and cross-validation, Koonce [21] built a novel deep learning model based on the transfer-learning technique. With the suggested model, pre-trained CNN architectures like Inception V3, Visual Geometry Group (VGG-19), ResNet50, VGG-16, ResNet, and Inception-V2 are used to extract the features from the MIAS dataset. According to experimental findings, the VGG16 model transfer-learning is effective for diagnosing BC since it classifies mammography breast images generally accurately.

Salama and Aly [34] built and contrasted approaches for the fusion of tabular non-image data and imaging, assessing fusion at various points in the model. This study shows that combining non-image data with photos can greatly increase predictive performance and that fusing intermediate learned features is preferable to fusing final probabilities in the dataset for classifying BC. A new paradigm for the segmentation and classification of BC images was introduced by Abdelrahman *et al.* [1]. Different models are used to categorize Digital Database for Screening Mammography (DDSM), MIAS, and Curated Breast Imaging Subset of DDSM (CBIS-DDSM) datasets into benign and malignant states, including DenseNet121, ResNet50, InceptionV3, MobileNetV2, and VGG16 models. Additionally, the breast area from the mammography images is segmented using the trained modified U-Net model. This technique will increase the effectiveness of the system and serve as a radiologist's assistant in early detection. To solve this issue, transfer learning and data augmentation are used.

Ara *et al.* [4] introduced CNN based models for computer vision in mammography after surveying traditional Computer Assisted Detection (CAD). The study then goes over the most recent findings on CNNs for four separate mammography chores, including classifying breast density, classifying breast asymmetry,

classifying calcification, and classifying mass. This research also presents and compares the reported numerical outcomes for every task as well as the benefits and drawbacks of the various CNN-based methods. To treat a patient with BC effectively, early tumor type identification is crucial. The Wisconsin Breast Cancer Dataset (WBCD), which was gathered from the University of California Irvine machine learning repository (UCI) library, was used by Krishnaveni *et al.* [22] to examine the effectiveness of several machine learning algorithms for predicting BC by analyzing the dataset. Here, classifiers for separating benign from malignant tumors have been built using SVM, K-Nearest Neighbors (KNN), Logistic Regression (LR), Naive Bayes, Decision Tree, and Random Forest (RF).

Notably, despite the extensive research that has been done on the categorization of BC histopathology images, the current work is still difficult in terms of accuracy, complicated model implementation, and lengthy computation times. Additionally, the process of categorization has become simpler and more time-effective to do. Based on CNN models' effectiveness in classifying medical images, we employed EFPCNNFSM for feature extraction and an attention-based TAENet classifier to separate benign and malignant breast tissues from pathology images. With the help of ensemble feature extraction and selection, the proposed effort aims to increase the accuracy of cancer detection and classification.

3. Proposed Approach

The subject of deep learning attracts a lot of attention when many deep learning techniques are being developed which give superior outcomes in several domains. This study suggests the EFPCNNFSM-TAENet model for classifying BC based on histopathology images. Figure 1 shows the architecture of the proposed framework. The EFPCNNFSM-TAENet model comprises four stages which are preprocessing, feature selection, feature extraction, and classification. The model works on the Invasive Ductal Carcinoma (IDC) dataset and BreakHis dataset. At first, preprocessing is done by synthetic augmentation and then denoising the images using Deep Neural Network (DNN). In this work, an EFFSM is introduced which draws the best feature subsets from the provided datasets using six different filter feature selection techniques, such as Information Gain (IG), Supervised Relative Reduct (SRR), Gain-Ratio (GR), OneR, minimum Redundancy Maximum Relevance (mRMR) and Chi-Squared (CS). Then the filter-pruned CNN is used to extract the optimal features from the selected features. Additionally, rotational based triplet attention mechanisms are integrated with the EfficientNet for classification known as TAENet. The attention mechanism's main goal is to emphasize the

discriminative region of the feature map while omitting unnecessary data. Finally, the performances are assessed in terms of accuracy, precision, recall, etc.

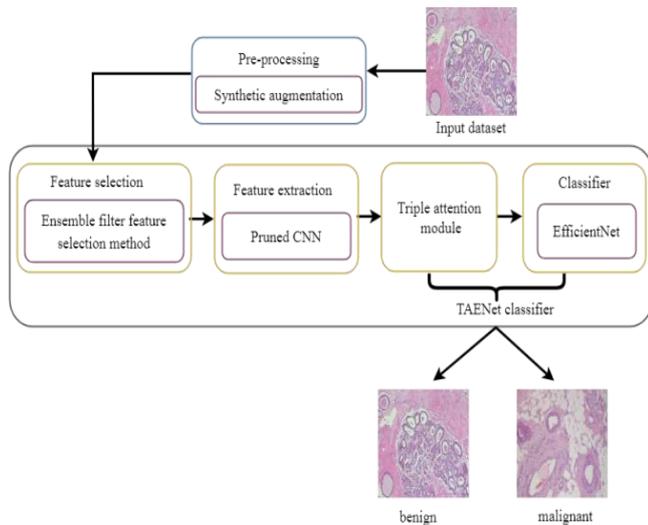


Figure 1. Architecture of proposed BC classification approach.

3.1. Preprocessing

During the deep neural network training process, the maximization of training data based over-fitting can be reduced by the data augmentation technique. The goal of augmentation is to incorporate fresh samples that match the original data supply into the initial training set. As a result, a successful augmentation technique should produce samples that are distinct from those in the initial training set while still adhering to the original data distribution. Conversely, when samples from a faulty augmentation scheme are added to the training set, they may provide samples that differ from the distribution of the original data. Simple data augmentation strategies have a little augmentation diversity while acting as an implicit regularization. Numerous efforts have been made to increase the efficacy of data augmentation to get over the limitations of traditional augmentation. Another common practice, which we refer to as synthetic augmentation, is the creation of synthetic images of original images, which boosts the quantity and diversity of the original training data. In the low-data regime, the synthetic augmentation generated samples are intended to supplement the default classifier. The medical image domain has greater problems with incomplete and unbalanced data than research done in the natural image domain. In medical image identification, researchers have begun to build synthetic augmentation to help address these problems.

3.2. Ensemble Feature Selection

The suggested approach uses a feature selection strategy based on an ensemble approach. In this work, an EFFSM is introduced which draws the best feature subsets from the provided datasets using six different filter feature selection techniques. Additionally, a

combination rule is employed, which involves combining the best feature subsets to produce the final feature subsets. The fundamental concept of the suggested feature selection approach for breast tumor classification is controlled, managed, and implemented by these functional components.

The IG [7], SRR, GR, OneR, mRMR, and CS feature selections are combined and used by the functional unit of the structure to produce the result. From the given datasets, the most relevant features are selected by this proposed EFFSM. Figure 2 shows the ensemble filter feature selection model.

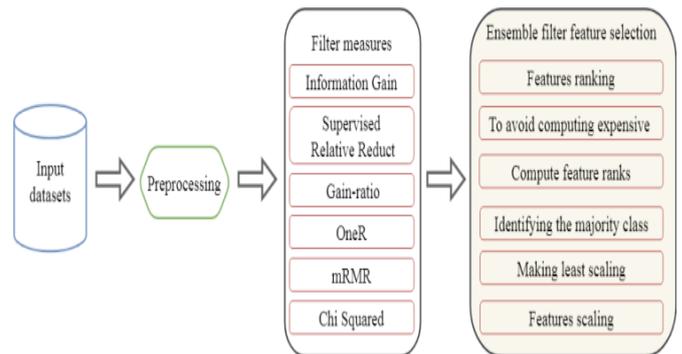


Figure 2. Ensemble filter feature selection method.

3.2.1. Information Gain

One of the most popular feature selection techniques based on mutual information is IG which is efficient and simple. It quantifies the information in bits about the class prediction, when that feature is present and the related class distribution is known, among the feature and the class labels C . Equation (1) expresses the IG, where $Info(D)$ represents the absolute entropy of the dataset. The sample attribute entropy is $Info_A(D)$,

$$InfoGain(A) = Info(D) - Info_A(D) \quad (1)$$

3.2.2. Supervised Relative Reduct

A feature assessment metric termed relative dependency serves as the foundation of the SRR technique [33], which is based on backward searching. This strategy was first offered as a way to avoid computing expensive Rough Set Theory (RST) discernibility functions or positive regions. The backward elimination of features is used by SRR, and if the removal of a feature causes a relative dependency equal to 1, it is removed from the set of features that are considered more redundant and low relevant features in the dataset. Each feature is taken into account one at a time, and a calculation of their relative dependence is made.

3.2.3. Gain-Ratio

The gain ratio FS technique is the discrepancy measure that provides a regularized score to enhance the IG score. The value of split information is calculated by Equation (2).

$$Split\ info_A(D) = - \sum_{j=1}^v \frac{|D_j|}{|D|} \times \log_2 \frac{D_j}{D} \quad (2)$$

Here, the configuration of v partitions indicates the *Split info*, the original dataset is represented by D , D_j represents the j th sub-dataset after being *split*, the respective numbers of samples belong to the original dataset and the sub-dataset are represented by $|D|$ and $|D_j|$, and H_j represents the entropy of the j th sub-dataset. The gain ratio is defined in Equation (3),

$$Gain\ Ratio(A) = \frac{InfoGain}{Splitinfo} \quad (3)$$

3.2.4. OneR

OneR uses a rule-based classification algorithm to rank the features [28]. In essence, the method finds a straightforward rule for each feature by identifying the majority class for each feature's value. The features are then ranked under the accuracy of the related rules after assessing each rule's accuracy.

3.2.5. mRMR

This filter employs mutual information to choose the qualities that are most distinct from one another, making them the most pertinent and least redundant for the target class [38].

3.2.6. Chi Squared

The most used statistical measure FS method that analyses the correlation between two variables is the Chi squared FS. Assessing a feature's independence from its class may be helpful. It is defined in Equation (4),

$$\chi^2 = \sum_{ij} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (4)$$

where, i and j are the variables, O , E and χ^2 represents the observed value, expected value and CS values, respectively.

3.3. Feature Extraction

Generally, image processing tasks involved in BC detection scheme takes a lot of time and is less useful for efficiently separating objects from the background during segmentation. The proposed approach reduces the number of feature vectors used in the input data transformation process to speed up processing. Feature extraction is the process of changing the supplied data. Feature vectors are used as input vectors in classification tasks because they often contain relevant information. Recent years have seen the successful implementation of deep learning techniques, including CNN because of the notable increase in accuracy for numerous applications [16]. These methods can extract hierarchical features from image data, usually referred to as objective features, and may extract ranked features from image data, which is being proved as the best

alternative to manual feature selection. In machine learning, the problem is solved by applying the classifier to a feature map of the data. Additionally, each problem has a distinct set of facts and requires a separate set of applied solutions. Therefore, to get around this, CNN is utilized to create features automatically and then mix them with the proposed attention based EfficientNet classifier [19]. In this study, deep features are extracted from the chosen feature using CNN which is shown in Figure 3. Three convolutional layers, three max pooling layers, and a single FC layer make up the seven-layer CNN architecture shown in Figure 3.

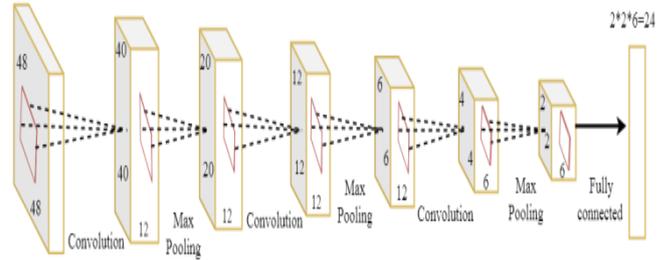


Figure 3. Architecture of CNN.

A dimension of 48x48 dimension sub-region images that is collected in earlier steps is the input for CNN. With 12 kernels of size 9x9x3, the first conv layer filters 48x48x3 input images to produce an output with the dimensions 40x40x12.

$$Conv^k(i, j) = \sum_{u, v} W^{k,l}(u, v) \times input^l(i - u, j - v) + b^{k,l} \quad (5)$$

where, $W^{k,l}$ represents the k^{th} kernel and $b^{k,l}$ represents the bias of k^{th} layer. RSigELUD is used as the activation function, which is restricted to be within the interval [1, 1]. The vanishing gradient problem affects activation functions in conventional methods like sigmoid and tangent functions. Hence, the ReLU activation function and its variations are suggested as a solution to this issue. But there is a negative area issue raised because of its effectiveness in the negative, positive, and linear activation areas, the RSigELUD activation function is suggested as a solution to the negative region and vanishing gradient issues. The RSigELUD activation function is shown in the Equation (6) [39],

$$f(x) \begin{cases} x \left(\frac{1}{1 + e^{-x}} \right) \alpha + x, & \text{if } 1 < x < \infty \\ x, & \text{if } 0 \leq x \leq 1 \\ \beta(e^x - 1), & \text{if } -\infty < x < 0 \end{cases} \quad (6)$$

Input and output are directly mapped in the region of linear activity. The values for the slope coefficients (α) and (β) be in the range of 0 to 1, x is input data. The positive and negative zones are managed by this value. Table 1 shows the parameters of CNN.

Table 1. Parameters of CNN architecture.

Parameters	Input	Con. 1	Max pooling	Con. 2	Max pooling	Con. 3	Max pooling	Fully connected
Width (W)	48	40	20	12	6	4	2	2
Height (H)	48	40	20	12	6	4	2	2
Depth/Kernel (D/K)		12	12	12	12	6	6	6
F		9×9×12	2×2×12	2×2×12	2×2×12	2×2×6	2×2×6	
Stride (S)		1	2	2	2	2	2	
Padding		0	0	1	0	1	0	
No. of Parameters while using CNN		2916		576		288		24
Parameters using FFN		3686400		12800000		576		

A max-pooling layer is coupled to the output of the first conv layer which is described in Equation (7),

$$Out^k(i,j) = \tanh(Conv^k(i,j)) \quad (7)$$

Then, until an output with the dimensions 2×2×6 is attained, the 2nd and 3rdconv/max-pooling layers are connected.

Weight pruning and filter pruning are the main focuses of current CNN model pruning techniques. From the training phase, most of these pruning methods are dispersed by establishing different threshold measures or limits based on previously learned information. Nevertheless, the filters, cannot be retrieved after they have been pruned. Furthermore, by directly regularizing on filters, the generalization capacities and training stabilities can be limited [37]. To overcome these drawbacks, the conventional method for determining a pruning measure is employed to extract a sparse structure from the original model through training. Figure 4 illustrates the CNN pruning structure.

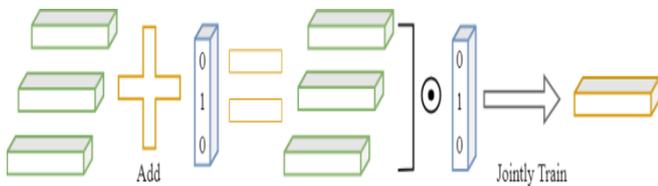


Figure 4. Pruning structure of CNN.

To create a sparse structure while maintaining the correctness of the initial model, identify the minimal subset $w \in W$. Direct regularization on w will make the batch training procedure more unstable. To designate which filter has to be pruned, a related indicator function is implemented for collaborative layers. When the weights have been quantized, the binary indicator function indicates this. The indicator function initially creates binary output weights (0 or 1) for collaboration layers, which can be described in Equation (8),

$$\sigma(v_j^i) = \begin{cases} 0, & \text{if } |v_j^i| \leq t \\ 1, & \text{if } |v_j^i| > t \end{cases} \quad (8)$$

where, i represents the i^{th} conv layer and j signifies the j^{th} filter v_j^i represents the collaborative layer parameters. The output is “1” when the total value is greater than a predetermined threshold t , and vice versa. The mask “1” denotes preservation of the relevant filter, while “0” denotes the removal of the filter. The indicator function

is combined with the threshold hyper-parameter t to allow us to regulate the rate at which each conventional layer is pruned. The value of t can vary depending on the convolutional layer. Finally, based on these pruned CNN, the optimal features are extracted.

3.4. Triplet Attention Module

The attention scheme is a technique for distributing probabilities. To improve the high-dimensional features quality of the hidden layer, the method computes features at various times to facilitate the features through further information which has greater weighting coefficients. In the proposed model, a triplet attention mechanism is introduced for BC classification that draws inspiration from the Convolutional Block Attention Module (CBAM). The proposed triplet attention module is shown in Figure 5.

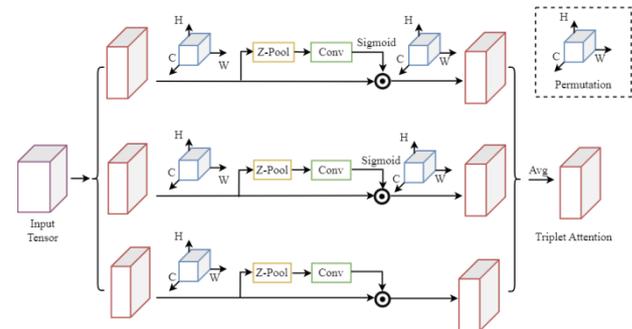


Figure 5. Structure of triplet attention module.

Using the spatial dimension W and the channel dimension C compute the attention weight is the responsibility of the top branch. The middle branch is also in charge of the C and spatial dimension H . The bottom branch is utilized to record spatial dependencies (W and H). To create links among the spatial and channel dimensions in the first two branches, the rotation operation is utilized. The weights are then simply averaged to combine them. Three parallel branches make up the triplet attention, two of which are in charge of recording interactions concerning the spatial dimensions H or W and C . The last branch, which is used to develop spatial attention, is comparable to the CBAM. The results of all three branches are combined using direct averaging [25].

coefficient. EfficientNet produces more effective results. The initial stage in the compound scaling strategy is to look for a grid in order to ascertain the association concerning the various scaling magnitudes of the standard network in a specific resource restriction. This makes it possible to choose a proper scaling factor for the breadth, depth, and resolution dimensions. The target network is then scaled using these coefficients from the baseline network. The Mobile Inverted Bottleneck Convolution (MBConv) serves as the foundation of EfficientNet. Direct connections are utilized between bottlenecks that connect much fewer channels than expansion layers because blocks in MBConv consist of a layer that expands and consequently compresses the channels. In comparison to conventional layers, this architecture uses in-depth separable convolutions that almost double the calculation efficiency [6]. The kernel size (k) specifies the dimensions of the 2D convolution window. In compound scaling, depth, width, and resolution are evenly scaled using the compound coefficient and the guidelines in Equation (11).

$$\begin{aligned} \text{depth:} \quad & d = \alpha^\Psi \\ \text{width:} \quad & w = \beta^\Psi \\ \text{resolution:} \quad & r = \gamma^\Psi \\ & \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \end{aligned} \quad (11)$$

where, grid search can be used to identify the constants α , β , and γ . The amount of resources that can be used for model scaling is controlled by the user-defined coefficient φ , and the network width, depth, and resolution are assigned to these additional resources in accordance with α , β , and γ , respectively.

4. Result and Discussion

In the proposed technique, filter feature selection with TAENet mechanism-based tumor classification is carried out in the BC image dataset. The experiments were conducted on Google Colab utilizing Python programming, a core i3 processor, and a 4GB RAM system to determine the effectiveness of the suggested network-based tumor classification approach. The two datasets used in the suggested method have respective weights of 70%, 10%, and 20% for training, validation, and testing. The suggested technique is used to increase the performance of trained model from the dataset, and after that, the models are fine-tuned using loss functions. With the training dataset, the network is trained with 200 epochs before having its parameters optimized. The models are assessed using the training and validation data at each epoch. Finally, predictions for the test data are made using the trained models, and the resulting scores are determined using real segmentation and assessment metrics. To create various scores and determine the general distribution of metrics, the average number of runs for each network is set to 15. The number of epochs is 250, the optimizer is Adam, the batch size is 4, and the bias initialization is 0.

In this experiment, the proposed deep learning models' performances are evaluated based on varied epoch sizes. Figure 8 shows the results training accuracy comparison evaluation. Observing the two conventional classifiers CNN and DenseNet, indicate that their classification performances are unaffected by the number of training epochs. When 70% of the training samples are used for these approaches, promising results are obtained, and serving more training examples to these methods slightly improves the performance. The proposed TAENet method often performs the best compared to other methods. The accuracy curve is displayed in Figure 9.

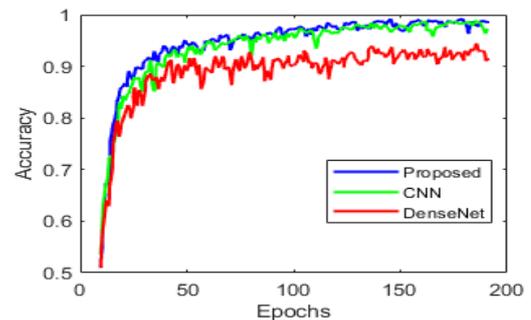


Figure 8. Training accuracy curve.

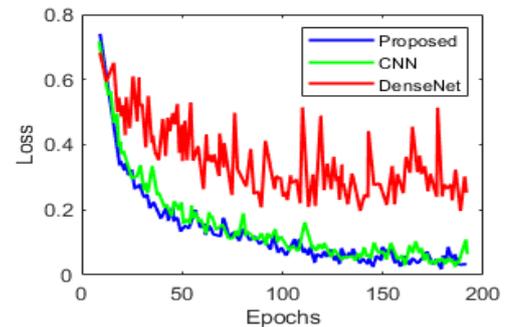


Figure 9. Training loss curve.

The proposed technique has a higher initial loss value, which effectively decreases as epochs increases. The curve converges rapidly and achieves the minimal value after 20 epochs, according to the observation. With 200 training epochs completed on the sample, the loss value is relatively small. Because of this, the suggested network model has a high accuracy rate and a low loss value. From the comparative analysis, the minimum validation loss is obtained for the proposed model than the CNN and denseNet. To achieve better results, the system runs with numerous tests and various learning rates. The proposed method chooses the default learning rate as 0.0001.

4.1. Performance Evaluation

The proposed tumor classification scheme is evaluated by the measures of F1 score, recall, precision accuracy, sensitivity, specificity and classification rate. These parameter metrics are defined as below,

$$accuracy = \frac{TP + TN}{N'} \in (0, 1) \tag{12}$$

where, number of test images is denoted as N' , number of false positive, false negative, true positive and true negative is denoted as FP, FN, TP and TN , respectively. Sensitivity refers to the ability of correctly identifying the images, while the specificity refers to the ability to correctly classify normal images. For these measurements, higher values indicate the best classification effect.

The $F1$ score is known as the dice score which is interrelated to the IoU. It is expressed in Equation (13),

$$F1_{score} = \frac{2 \times Precision \times recall}{recall + precision} \tag{13}$$

The consistent average of F1 score, precision, and recall is so far the most appropriate for unbalanced datasets by its definition. According to the formula, the F1 score result should also be zero.

$$precision = \frac{TP}{TP + FP} \tag{14}$$

$$recall = \frac{TP}{TP + FN} \tag{15}$$

It is necessary to assess the effectiveness of self-supervised techniques on particular downstream classification tasks.

4.2. IDC Dataset Analysis

The most prevalent subtype of BC is IDC. When grading the aggressiveness of a whole mount sample, pathologists frequently concentrate on the areas that contain the IDC [10]. 162 whole mount slide photos of BC specimens that were scanned at 40x made up the original dataset. A total of 277,524 patches measuring 50 by 50 were taken from it (198,738 IDC negative and 78,786 IDC positive). 0 denotes a non-IDC and 1 denotes IDC.

Breast histology images are divided into two categories in this study: tumor and non-tumor. The sample classified images are shown in Table 2. Purple (nuclei), white (background), pale pink (cytoplasm in stroma), red (red blood cells), and dark pink are the five colors that dominate a histology slide (cytoplasm in squamous epithelium). The findings obtained using the suggested strategy are extremely encouraging and show the strong discriminatory potential of this trait. To validate the method suggested in this paper, the test set's preprocessed breast histopathology images are employed. 512x512 patches with a 50% overlap and contiguous 128x128 non-overlapping patches should be extracted from the test images. The triplet attention based EfficientNet method is employed to compute the final feature of each image and make a final classification.

Table 2. Sample classified breast images.

	Image 1	Image 2	Image 3	Image 4	Image 5
Tumor					
Non-tumor					

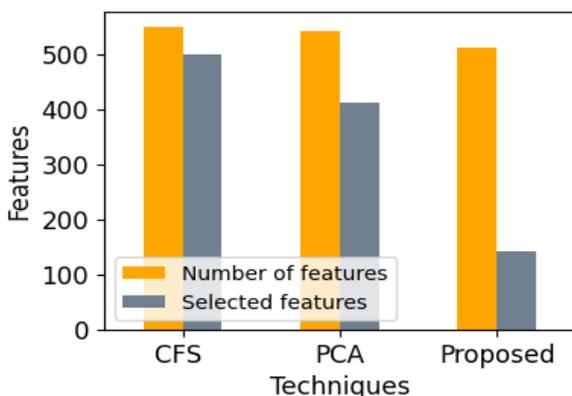


Figure 10. Total features vs selected features.

The process of feature selection is illustrated in Figure 10 using different techniques. Correlation-based Feature Selection (CFS) [13] used 550 characteristics, of which 500 are ultimately acquired after selection. There are 550 features available for PCA [15] from that 400 are chosen. In this case, all of the algorithms used the same number of features, but different features are chosen. There are 520 features in the suggested filter feature selection process, and 150 of them are chosen.

The results of the suggested TAENet image classification are described in Table 3. The accuracy (%), sensitivity (%), and specificity (%) for each pre-processed image are calculated. Five sample images are used in this investigation. It has been determined

whether the condition is normal or cancer using several medical images. Some of the preprocessed images are displayed in Table 3 all of the pre-processed images are expected to fall into one of two classes: normal or cancer. The first image analysis predicted the class as cancer for breast image with 93.34% accuracy, 87.56% sensitivity with 95.76% specificity for the first image. With 95.66% accuracy, 89.34% sensitivity, and 97.37% specificity, the second image is found to be normal and belonged to the anticipated class. The prediction for the third and fourth images is accurate. The fifth set of image is identified as cancer, while the expected category is normal. The classification rate analysis for various images is shown in Figure 11.

Table 3. Classification results.

Preprocessed target image	Original class	Predicted class	Accuracy	Sensitivity	Specificity
Image 1 tumor	Cancer	Cancer	93.34	87.56	95.76
Image 2 non tumor	Normal	Normal	95.66	89.34	97.37
Image 3 tumor	Cancer	Cancer	87.23	82.96	89.48
Image 4 non tumor	Normal	Normal	91.42	87.45	93.64
Image 5 tumor	Cancer	Normal	89.77	82.28	91.42

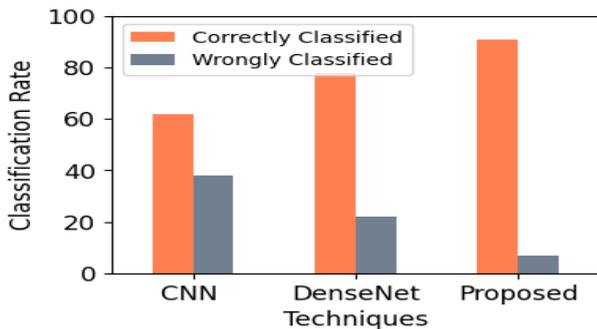


Figure 11. Classification rate.

This section addresses whether or not image classification, utilizing classification techniques, is accurate. The TAENet classifier properly classified the chosen images with a classification rate of 95.5%, while the classification rate for incorrect classification was 0.9%. In CNN, the image had a 62% correct classification rate and a 36% incorrect classification rate. The classification rate in DenseNet [31] is 75% for correctly classified items and 20.85% for incorrectly classified items.

Table 4. Classification output with different training ratios.

Sample image	Training/Testing	TAENet classifier		
		Accuracy	Sensitivity	Specificity
Image 4 non tumor	90%-10%	91.42	87.45	93.64
	80%-20%	88.38	81.04	91.36
	70%-30%	98.42	94.45	94.42
	60%-40%	81.37	71.53	84.55
	50%-50%	78.71	79.22	80.34
	40%-60%	72.34	63.49	75.91

method are explained in Table 4. Here, an examination of the sample images is taken into account. Specificity, accuracy, and sensitivity are determined after training the image. With 90% training and 10% testing, the accuracy is 91.42%, with 93.64% specificity and 87.45% sensitivity. With 80% and 20%, the precision is 88.38%, the specificity and sensitivity are 91.36% and 81.04%, respectively. 98.42% accuracy, 94.42% specificity, and 63.49% sensitivity are achieved by 70% and 30% of training and testing images.

The confusion matrix of the proposed classifier and existing CNN and DenseNet classifier is displayed in Figure 12-a), (b), (c), and (d) respectively. From this, it can be noted that a total of 80800 images is utilized for testing the proposed classifier. Except for CNN, the DenseNet classifier and proposed classifier have accuracy levels above 90%. The proposed classifier has a higher accuracy of 98% in these classifiers. The TAENet classifier beats other classifiers by forecasting the test data using the ensemble classifiers' highest accuracy.

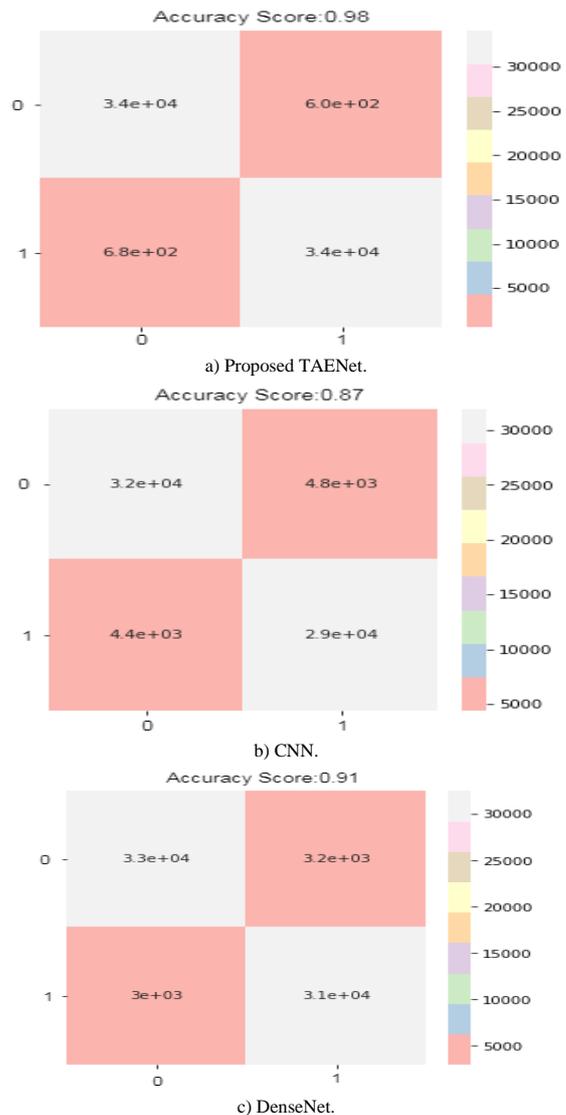


Figure 12. Confusion matrix.

The training and testing results for the suggested

Table 5. Comparative analysis.

Method	Accuracy	Precision	Recall	F1 Score
CNN	87.31	86.05	84.42	85.34
DenseNet	91.55	91.17	89.74	90.53
Proposed	98.84	98.32	97.18	96.86

The efficacy of the suggested method can be compared to several cutting-edge approaches that are employed for the classification of BC on histopathology images. The majority of these cutting-edge deep learning techniques are dataset-based. The proposed classifier is evaluated against CNN and DenseNet. Table 5 presents the comparison analysis.

4.3. BreakHis Dataset Analysis

9109 microscopic images of breast tumor tissue taken at various magnifications (40X, 100X, 200X, and 400X) are gathered from 82 patients who make up the BC Histopathological image classification (BreakHis). It now has 2,480 benign and 5,429 cancerous samples (700X460 pixels, 3-channel RGB, 8-bit depth in each channel, Portable Network Graphic (PNG) format). The

two primary categories of the BreakHis dataset are benign tumors and malignant cancers. If there are no signs of malignancy, a lesion is said to be histologically benign, such as significant cellular atypia, mitosis, breakdown of basement membranes, metastasis, etc. Normal benign tumors are slow-growing, confined growths that are considered “innocent.” Cancer is regarded as a malignant tumor because the lesion can spread to distant areas (metastasize) and infect other structures, causing damage and eventual death. The dataset is available in (<https://www.kaggle.com/code/nasrulkim86/breast-cancer-histopathology-images-classification/data>). In this study, the breast histology images are classified into two categories:

- a) Benign.
- b) Malignant.

Table 6 shows the sample classified images for the BreakHis dataset. The process of feature selection is illustrated in Figure 13 using several techniques.

Table 6. Sample classified image.

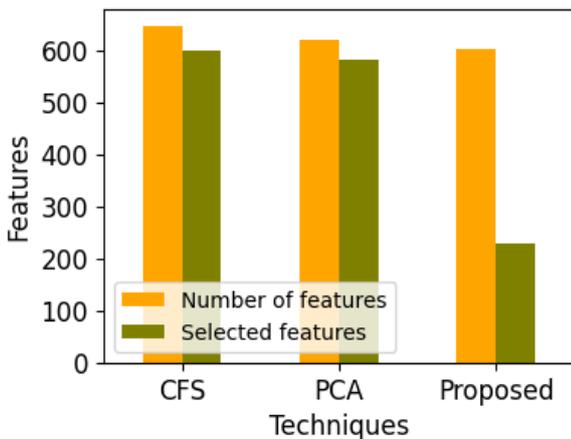
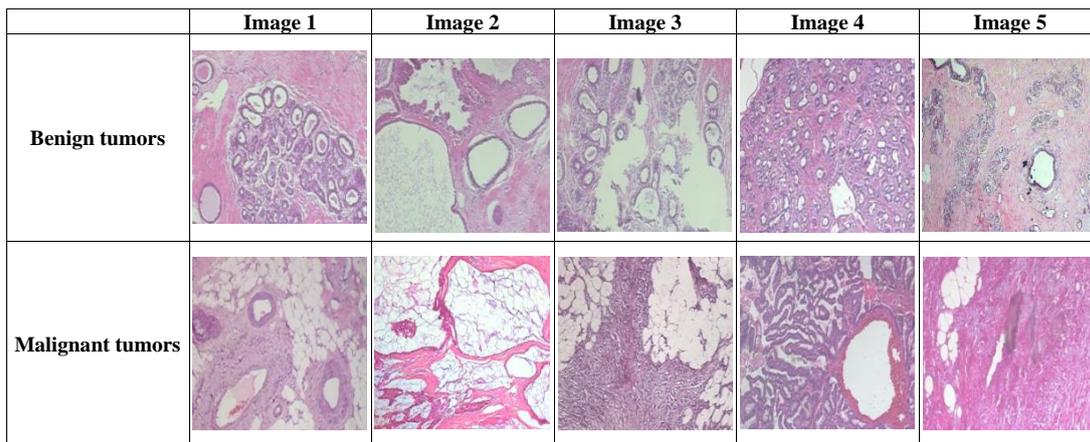


Figure 13. Total features vs. selected features.

In CFS, there are 650 characteristics in total, of which 600 are acquired after selection. The chosen feature out of 620 features for PCA is one of 580. In this case, all of the algorithms used the same number of features, but different features are chosen. In the suggested filter

feature selection procedure, there are 600 features from that 200 are chosen.

Table 7. Image classification outcome.

Preprocessed input images	Original class	Predicted class	Accuracy	Sensitivity	Specificity
Image 1	benign tumors	benign tumors	98.48	98.90	98.04
Image 2	malignant tumors	malignant tumors	97.03	97.79	96.26
Image 3	benign tumors	benign tumors	96.87	97.48	96.25
Image 4	malignant tumors	benign tumors	95.75	97.12	94.38
Image 5	malignant tumors	Normal	93.41	87.16	91.43

The results of the suggested TAENet image classification are shown in Table 7. The accuracy (%), sensitivity (%), and specificity (%) for each pre-processed image are calculated. Five sample images are used for this investigation. The diagnosis of the illness as benign or malignant has been made using several

medical images. Some of the preprocessed images are displayed in Table 7 all of the pre-processed images are expected to fall into one of two classes: normal or cancer. With 98.48% accuracy, 98.90% sensitivity, and 98.04% specificity for the first image and correctly identified the class as benign tumors. The second image proved to be malignant tumors with 97.03% accuracy, 97.79% sensitivity, and 96.26% specificity. The third and fourth image is correctly predicted. The fifth image is classified to be cancer and the predicted class is normal. The classification rate analysis for a diverse image is shown in Figure 14.

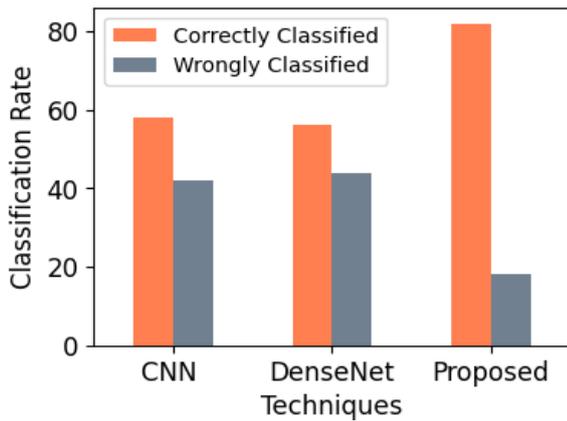


Figure 14. Classification rate.

This addresses whether or not the image classification techniques, are accurate. The TAENet properly classified the chosen images with a classification rate of 80%, whereas the classification rate for incorrect classification is 18%. CNN accurately classified the image at 58% and incorrectly classified it at 42%. The classification rate in DenseNet was 55% for correctly classified items and 44% for incorrectly classed items. Table 8 shows the classification output with different training ratios. From the analysis, 70%-30% training and testing, the proposed model achieves better performance results. The confusion matrix of the proposed classifier and existing CNN and DenseNet classifier is displayed in Figure 15-a), (b), (c), and (d) respectively.

Table 8. Classification output with different training ratios.

Sample Image	Training/Testing	Proposed Model		
		Accuracy	Sensitivity	Specificity
Image 5 malignant tumors	90%-10%	92.22	91.26	94.43
	80%-20%	90.81	91.79	88.36
	70%-30%	97.63	98.12	98.12
	60%-40%	88.64	87.50	92.95
	50%-50%	87.39	85.24	87.24
	40%-60%	84.56	83.45	84.56

From this, it can be noted that a total of 6376 images is utilized for testing the proposed classifier. Except for CNN, the DenseNet classifier and proposed classifier have accuracy levels above 90%. The proposed classifier has a higher accuracy of 98% in these classifiers. The TAENet classifier beats other classifiers by forecasting the test data using the ensemble

classifiers' highest performance. In Figure 16, a comparison analysis is illustrated.

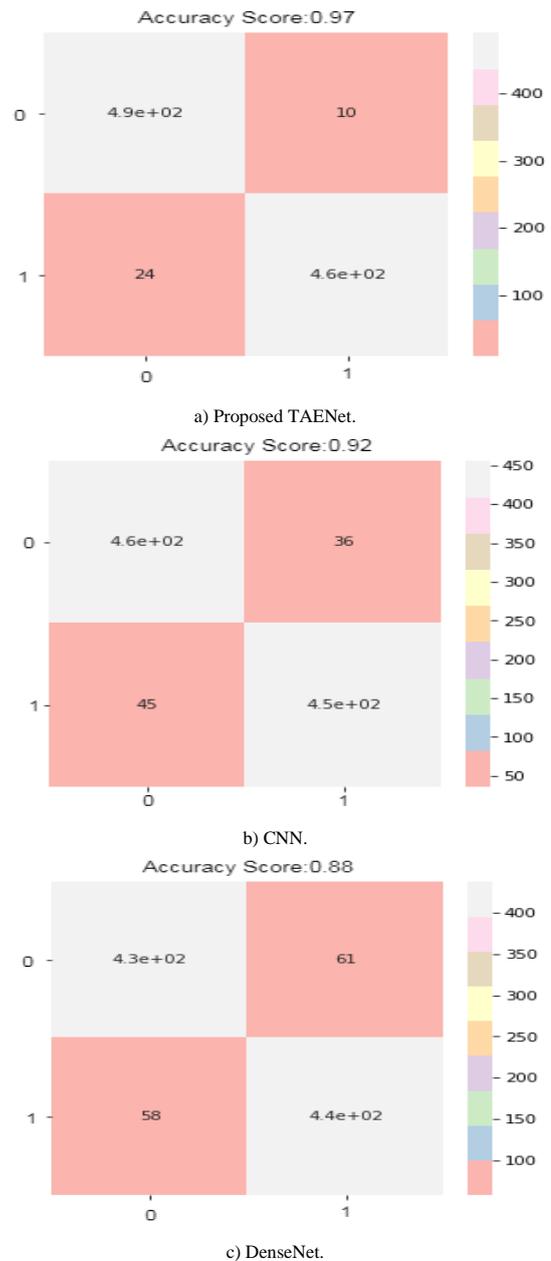


Figure 15. Confusion matrix.

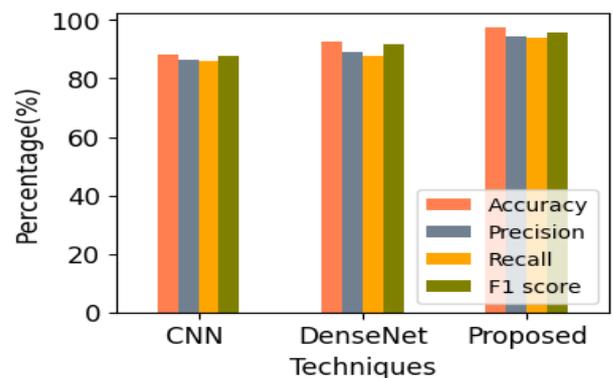


Figure 16. Comparative analysis.

The efficacy of the suggested method can be compared to several cutting-edge approaches that are

employed to categorize histopathological images of BC. The majority of these cutting-edge deep learning techniques are dataset-based. The proposed classifier is evaluated against CNN and DenseNet.

5. Conclusions

BC is one of the most terrible diseases which affect women nowadays. In this research, the International Classification of Diseases (ICD) and BreastHist histopathology datasets are used, and a novel EFPCNNFSM based TAENet classification network is applied to estimate the effectiveness and to discover the maximum precision of classifying malignant and benign BC. For the feature selection approach, the association between various dataset features has been examined by using EFPCNNFSM. Then, the TAENet is used to classify the tumors based on non-tumor and tumor, benign and malignant in the experiments. From the study, it is concluded that TAENet classifier achieves better accuracy (98%) than the CNN and DenseNet classifiers. In the future, this work can be enhanced by managing a fairly large dataset and including more features like identifying the stage of BC. When employing a lot of inputs, this model might perform better and deliver greater accuracy. Additionally, the study uses binary classification, to distinguish between cancer and non-cancerous conditions. Using this model, other kinds of cancer can also be classified in the future.

References

- [1] Abdelrahman L., Al Ghamdi M., Collado-Mesa F., and Abdel-Mottaleb M., "Convolutional Neural Networks for Breast Cancer Detection in Mammography: A Survey," *Computers in Biology and Medicine*, vol. 131, pp. 104248, 2021. <https://doi.org/10.1016/j.compbimed.2021.104248>
- [2] Albalawi U., Manimurugan S., and Varatharajan R., "Classification of Breast Cancer Mammogram Images Using Convolution Neural Network," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 13, 2022. <https://doi.org/10.1002/cpe.5803>
- [3] Alzubaidi L., Al-Shamma O., Fadhel M., Farhan L., Zhang J., and Duan Y., "Optimizing the Performance of Breast Cancer Classification by Employing the Same Domain Transfer Learning from Hybrid Deep Convolutional Neural Network Model," *Electronics*, vol. 9, no. 3, pp. 1-21, 2020. <https://doi.org/10.3390/electronics9030445>
- [4] Ara S., Das A., and Dey A., "Malignant and Benign Breast Cancer Classification Using Machine Learning Algorithms," in *Proceeding of the International Conference on Artificial Intelligence*, Islamabad, pp. 97-101, 2021. DOI:10.1109/ICAIS2203.2021.9445249
- [5] Aswathy M. and Jagannath M., "An SVM Approach towards Breast Cancer Classification from H and E-Stained Histopathology Images Based on Integrated Features," *Medical and Biological Engineering and Computing*, vol. 59, no. 9, pp. 1773-1783, 2021. DOI: 10.1007/s11517-021-02403-0
- [6] Atila Ü., Uçar M., Akyol K., and Uçar E., "Plant Leaf Disease Classification Using Efficientnet Deep Learning Model," *Ecological Informatics*, vol. 61, pp. 101182, 2021. <https://doi.org/10.1016/j.ecoinf.2020.101182>
- [7] Bania R. and Halder A., "R-HEFS: Rough Set Based Heterogeneous Ensemble Feature Selection Method for Medical Data Classification," *Artificial Intelligence in Medicine*, vol. 114, pp. 102049, 2021. <https://doi.org/10.1016/j.artmed.2021.102049>
- [8] Buchberger W., Geiger-Gritsch S., Knapp R., Gautsch K., and Oberaigner W., "Combined Screening with Mammography and Ultrasound in a Population-Based Screening Program," *European Journal of Radiology*, vol. 101, pp. 24-29, 2018. <https://doi.org/10.1016/j.ejrad.2018.01.022>
- [9] Conti A., Duggento A., Indovina I., Guerrisi M., and Toschi N., "Radiomics in Breast Cancer Classification and Prediction," *Seminars in Cancer Biology*, vol. 72, pp. 238-250, 2021. <https://doi.org/10.1016/j.semcancer.2020.04.002>
- [10] Cruz-Roa A., Basavanhally A., González F., Gilmore H., Feldman M., Ganesan S., Shih N., Tomaszewski J., and Madabhushi A., "Automatic Detection of Invasive Ductal Carcinoma in Whole Slide Images with Convolutional Neural Networks," in *Proceedings of the SPIE-The International Society for Optical Engineering*, San Diego, pp. 904103, 2014. <https://doi.org/10.1117/12.2043872>
- [11] Feuer E., Wun L., Boring C., Flanders W., Timmel M., and Tong T., "The Lifetime Risk of Developing Breast Cancer," *Journal of the National Cancer Institute*, vol. 85, no. 11, pp. 892-897, 1993. DOI: 10.1093/jnci/85.11.892
- [12] Ghiasi M. and Zendejboudi S., "Application of Decision Tree-Based Ensemble Learning in the Classification of Breast Cancer," *Computers in Biology and Medicine*, vol. 128, pp. 104089, 2021. <https://doi.org/10.1016/j.compbimed.2020.104089>
- [13] Gopal V., Al-Turjman F., Kumar R., Anand L., and Rajesh M., "Feature Selection and Classification in Breast Cancer Prediction Using IoT and Machine Learning," *Measurement*, vol. 178, pp. 109442, 2021. <https://doi.org/10.1016/j.measurement.2021.109442>
- [14] Hamed G., Marey M., Amin S., and Tolba M.,

- “Deep Learning in Breast Cancer Detection and Classification,” in *Proceedings of the International Conference on Artificial Intelligence and Computer Vision*, Cairo, pp. 322-333, 2020. https://doi.org/10.1007/978-3-030-44289-7_30
- [15] Haq A., Li J., Saboor A., Khan J., Wali S., Ahmad S., Ali A., Khan G., and Zhou W., “Detection of Breast Cancer through Clinical Data Using Supervised and Unsupervised Feature Selection Techniques,” *IEEE Access*, vol. 9, pp. 22090-22105, 2021. DOI: 10.1109/ACCESS.2021.3055806
- [16] Hirra I., Ahmad M., Hussain A., Ashraf M., Saeed I., Qadri S., Alghamdi A., and Alfakeeh A., “Breast Cancer Classification from Histopathological Images Using Patch-Based Deep Learning Modeling,” *IEEE Access*, vol. 9, pp. 24273-24287, 2021. DOI: 10.1109/ACCESS.2021.3056516
- [17] Holste G., Partridge S., Rahbar H., Biswas D., Lee C., and Alessio A., “End-to-End Learning of Fused Image and Non-Image Features for Improved Breast Cancer Classification from MRI,” in *Proceedings of the International Conference on Computer Vision Workshops*, Montreal, pp. 3287-3296, 2021. DOI: 10.1109/ICCVW54120.2021.00368
- [18] Houssein E., Emam M., Ali A., and Suganthan P., “Deep and Machine Learning Techniques for Medical Imaging-Based Breast Cancer: A Comprehensive Review,” *Expert Systems with Applications*, vol. 167, pp. 114161, 2021. <https://doi.org/10.1016/j.eswa.2020.114161>
- [19] Jones C. and Murugamani C., “Malaria Parasite Detection on Microscopic Blood Smear Images with Integrated Deep Learning Algorithms,” *The International Arab Journal of Information Technology*, vol. 20, no. 2, pp. 170-179, 2023. <https://doi.org/10.34028/iajit/20/2/3>
- [20] Kiliçarslan S. and Celik M., “RSigELU: A Nonlinear Activation Function for Deep Neural Networks,” *Expert Systems with Applications*, vol. 174, pp. 114805, 2021. <https://doi.org/10.1016/j.eswa.2021.114805>
- [21] Koonce B., *Convolutional Neural Networks with Swift for Tensorflow*, Apress, 2021. https://doi.org/10.1007/978-1-4842-6168-2_10
- [22] Krishnaveni S., Sivamohan S., Sridhar S., and Prabakaran S., “Efficient Feature Selection and Classification through Ensemble Method for Network Intrusion Detection on Cloud Computing,” *Cluster Computing*, vol. 24, no. 3, pp. 1761-1779, 2021. DOI: 10.1007/s10586-020-03222-y
- [23] Krithiga R. and Geetha P., “Breast Cancer Detection, Segmentation and Classification on Histopathology Images Analysis: A Systematic Review,” *Archives of Computational Methods in Engineering*, vol. 28, no. 1, pp. 2607-2619, 2021. DOI: 10.1007/s11831-020-09470-w
- [24] Massafra R., Bove S., Lorusso V., Biafora A., Comes M., Didonna V., Diotaiuti S., and Fanizzi A., “Radiomic Feature Reduction Approach to Predict Breast Cancer by Contrast-Enhanced Spectral Mammography Images,” *Diagnostics*, vol. 11, no. 4, pp. 1-15, 2021. DOI: 10.3390/diagnostics11040684
- [25] Misra D., Nalamada T., Arasanipalai A., and Hou Q., “Rotate to Attend: Convolutional Triplet Attention Module” in *Proceedings of the Winter Conference on Applications of Computer Vision*, Waikoloa, pp. 3139-3147, 2021. DOI: 10.1109/WACV48630.2021.00318
- [26] Murtaza G., Shuib L., Abdul Wahab A., Mujtaba G., Nweke H., and Al-Garadi M., “Deep Learning-Based Breast Cancer Classification through Medical Imaging Modalities: State of the Art and Research Challenges,” *Artificial Intelligence Review*, vol. 53, no. 1, pp. 1655-1720, 2020. DOI: 10.1007/s10462-019-09716-5
- [27] Pavithra M., Rajmohan R., Kumar T., and Ramya R., *Machine Vision Inspection Systems, Volume 2: Machine Learning-Based Approaches*, Scrivener Publishing LLC, 2021. <https://doi.org/10.1002/9781119786122.ch12>
- [28] Peng H., Long F., and Ding C., “Feature Selection Based on Mutual Information Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226-1238, 2005. DOI:10.1109/TPAMI.2005.159
- [29] Pes B., “Ensemble Feature Selection for High-Dimensional Data: A Stability Analysis across Multiple Domains,” *Neural Computing and Applications*, vol. 32, no. 3, pp. 5951-5973, 2020. DOI: 10.1007/s00521-019-04082-3
- [30] Ragab D., Attallah O., Sharkas M., Ren J., and Marshall S., “A Framework for Breast Cancer Classification Using Multi-DCNNs,” *Computers in Biology and Medicine*, vol. 131, pp. 104245, 2021. <https://doi.org/10.1016/j.combiomed.2021.104245>
- [31] Rane N., Sunny J., Kanade R., and Devi S., “Breast Cancer Classification and Prediction Using Machine Learning,” *International Journal of Engineering Research and Technology*, vol. 9, no. 2, pp. 576-580, 2020. DOI:10.17577/IJERTV9IS020280
- [32] Sabate J., Clotet M., Torrubia S., Gomez A., Gurrero R., “Radiologic Evaluation of Breast Disorders Related to Pregnancy and Location,” *Rodiographics*, vol. 27, no. suppl_1, pp. S101-S124, 2007. DOI:10.1148/rg.27si075505

- [33] Saber A., Sakr M., Abo-Seida O., Keshk A., and Chen H., "A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique," *IEEE Access*, vol. 9, pp. 71194-71209, 2021. DOI: 10.1109/ACCESS.2021.3079204
- [34] Salama W. and Aly M., "Deep Learning in Mammography Images Segmentation and Classification: Automated CNN Approach," *Alexandria Engineering Journal*, vol. 60, no. 5, pp. 4701-4709, 2021. <https://doi.org/10.1016/j.aej.2021.03.048>
- [35] Sarkar T., Hazra A., and Das N., "Classification of Colorectal Cancer Histology Images Using Image Reconstruction and Modified DenseNet," in *Proceedings of the International Conference on Computational Intelligence in Communications and Business Analytics*, Santiniketan, pp. 259-271, 2021. https://doi.org/10.1007/978-3-030-75529-4_20
- [36] Shahidi F., Daud S., Abas H., Ahmad N., and Maarop N., "Breast Cancer Classification Using Deep Learning Approaches and Histopathology Image: A Comparison Study," *IEEE Access*, vol. 8, pp. 187531-187552, 2020. DOI: 10.1109/ACCESS.2020.3029881
- [37] Tian G., Chen J., Zeng X., and Liu Y., "Pruning by Training: A Novel Deep Neural Network Compression Framework for Image Processing," *IEEE Signal Processing Letters*, vol. 28, pp. 344-348, 2021. DOI: 10.1109/LSP.2021.3054315
- [38] Wang Z., Li M., Wang H., Jiang H., Yao Y., Zhang H., and Xin J., "Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion with CNN Deep Features," *IEEE Access*, vol. 7, pp. 105146-105158, 2019. DOI: 10.1109/ACCESS.2019.2892795
- [39] Yeom S., Seegerer P., Lapuschkin S., Binder A., Wiedemann S., Müller K., and Samek W., "Pruning by Explaining: A Novel Criterion for Deep Neural Network Pruning," *Pattern Recognition*, vol. 115, pp. 107899, 2021. <https://doi.org/10.1016/j.patcog.2021.107899>
- [40] Zhang Y., Satapathy S., Guttery D., Górriz J., and Wang S., "Improved Breast Cancer Classification through Combining Graph Convolutional Network and Convolutional Neural Network," *Information Processing and Management*, vol. 58, no. 2, pp. 102439, 2021. <https://doi.org/10.1016/j.ipm.2020.102439>



Bangalore Nagaraj Madhukar is pursuing his Ph.D in the field of Image Processing in the School of ECE, REVA University, Bengaluru, India under the supervision of Bharathi.S.H. He has obtained in B.E. in ECE from Bangalore University, Bengaluru, MSc [Engg] in Digital Signal and Image Processing from Coventry University, U.K., and M.Tech in Geo-Informatics from VTU, Belgaum, India. His research areas of interest are Signal Processing, Image Processing, Video Processing, Remote Sensing, Radar Imaging, and Communication Systems.



Shivanandamurthy Hiremath Bharathi is a Professor in the school of ECE at REVA University, Bengaluru, India. Her main research area includes Video processing, Image processing and Electromagnetism, VLSI, Image Processing, and Video Processing. She has published more than 56 papers in national, international journals and conferences. She is a member of IEEE and Fellow of IETE societies. She is also a member of IEEE Women in Engineering. She is member of Amateur Radio Club.



Matta Polnaya Ashwin is working as Associate Professor in Department of Radiology, Interventional Radiology Division, AJ Institute of Medical Science and Research Centre (AJIMS and RC), Mangaluru, India. He has done his MBBS and MD in Radio-Diagnosis. He has rich experience of 15 years. His Research interests are Cross Sectional Body Imaging, Vascular and Non Vascular Interventional Radiological procedures.