

## RESEARCH ARTICLE

# Modality Specific CBAM-VGGNet Model for the Classification of Breast Histopathology Images via Transfer Learning

AREESHA IJAZ<sup>1</sup>, BASIT RAZA<sup>1</sup>, IQRA KIRAN<sup>1</sup>, ABDUL WAHEED<sup>2,3</sup>, AADIL RAZA<sup>4</sup>, HABIB SHAH<sup>5</sup>, AND SULAIMAN AFTAN<sup>6</sup>

<sup>1</sup>Department of Computer Science, COMSATS University Islamabad (CUI), Islamabad 45550, Pakistan

<sup>2</sup>Department of Computer Science, Women University Swabi, Swabi 23430, Pakistan

<sup>3</sup>School of Electrical and Computer Engineering, Seoul National University, Seoul 08826, South Korea

<sup>4</sup>Department of Physics, COMSATS University Islamabad (CUI), Islamabad 45550, Pakistan

<sup>5</sup>Department of Computer Science, College of Computer Science, King Khalid University, Abha 62529, Saudi Arabia

<sup>6</sup>Department of Computer Science, Texas Tech University, Lubbock, TX 79409, USA

Corresponding authors: Basit Raza (basit.raza@comsats.edu.pk) and Abdul Waheed (abdul@netlab.snu.ac.kr)

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Large Groups RGP.2/212/1443.

**ABSTRACT** Histopathology images are very distinctive, one image may contain thousands of objects. Transferring features from natural images to histopathology images may not provide impressive outcomes. In this study, we have proposed a novel modality specific CBAM-VGGNet model for classifying H and E stained breast histopathology images. Instead of using pre-trained models on ImageNet, we have trained VGG16 and VGG19 models on the same domain cancerous histopathology datasets, which are then used as fixed feature extractors. We have added the GAP layer and Convolutional block attention module (CBAM) after the first convolutional layer of convolutional blocks. CBAM is an effective module for neural networks to focus on relevant features. We have implemented the VGG16 and VGG19 in a novel way following the configuration of state-of-the-art models with our own concatenated layers. The addition of the GAP layer in VGGNet has reduced the number of parameters, requiring less computational power. Both models are ensemble using the averaging ensemble technique. Features are extracted from the final ensemble model and then passed to the feed-forward neural network. A hybrid pre-processing technique is proposed that first uses a median filter and then contrasts limited adaptive histogram equalization (CLAHE). The median filter removes the highly significant noise and is directly related to image quality. CLAHE improves the local contrast present in an image and boosts the weak boundary edges in each image pixel. The proposed CBAM ensemble model has outperformed state-of-the-art models with an accuracy of 98.96% and 97.95% F1-score on 400X data of the BreakHis dataset.

**INDEX TERMS** Breast cancer, CBAM, CLAHE, classification, deep learning, histopathology, transfer learning.

## I. INTRODUCTION

Cancer occurs in the body due to abnormal cell growth, which starts accumulating and forms a tumor [1]. This growth can be classified as non-tumor (benign) or tumor (malignant).

The associate editor coordinating the review of this manuscript and approving it for publication was Chaker Larabi<sup>1</sup>.

In general, benign do not spread to other body parts and are hence less threatening to human health, while malignant tumor is destructive and fast-growing, which can cause serious illness in the body. To date, cancer is the most dangerous and complex genetic disease. Pathologists use morphological abnormalities of the nucleus as the main feature to distinguish between benign and malignant cells [2]. According to a global

burden of disease (GBD) study in 2018, there are 2,088,849 new cases of breast, and it caused 626,679 deaths [3]. Among 100 different types of cancer, breast cancer has significantly increased the death rate of women. It is predicted that by 2030 there will be almost 27 million new cases of breast cancer [4]. These figures indicate the demand and importance of the automated system to lower the mortality ratio in women.

Although different imaging methods are available, Pathology images are used in this study as they are considered a gold standard for cancer [5]. This is because histopathology images provide a more detailed view of the underlying disease and its effect at the tissue level. In other words, histopathology provides microarchitectural characteristics of the tissue. The disease usually evolves from the molecular level and afterward at the tissue and cell level [6]. It is crucial to locate the disease when it starts affecting the tissue. Furthermore, detailed factors of some diseases can only be inferred using the histopathology image [7]. In clinics, the biopsy is usually done by experts. In this process, the part of the tissue is removed from the patient's body and prepared on a glass slide by applying H and E staining, which provides better visualization of nuclei (purple) and cytoplasm (pink). These slides are then examined by the pathologists under the microscope to analyze the tumor [8]. Pathologists examine the crucial areas, the distribution and arrangement of cells in the tissue, cell shape and structure, nuclei management, and its irregularities because of cancer [9]. However, this process is cost-effective, laborious, and needs vast domain knowledge and qualification.

Nowadays, deep learning models have provided better outcomes in different medical imaging tasks due to their ability to learn complex features from the image automatically. This is because a lot of researchers are making use of these models to perform breast cancer classification. The automated computer-aided diagnosis systems help narrow the error caused by human interpretation and improve the accuracy of the tumor diagnosis. Data unavailability is an utmost issue when we are working on medical images. One way to deal with this problem is transfer learning. This technique is used to transfer knowledge from one problem to another. This approach has gained the attention of many researchers as it is already effectively applied to different machine learning tasks like image classification [10], [11] text sentiment classification [12], software defect classification, and others [13]. However, we can only sometimes get the expected results if the source and focused domains are dissimilar. Different convolutional neural network (CNN) architectures like GoogleNet, ResNet, and VGGNet, trained on ImageNet, are used for transfer learning in various applications [14]. However, Google brain researchers proved that the performance of medical images does not improve by using transfer learning with the help of pre-trained models [15]. This is because of the difference between learned features. If both domains are very dissimilar, it may lead to negative transfer [16]. Moreover, according to the national cancer institute, almost

20% of cancer remains unidentifiable by using mammograms and X-ray images [17]. On the other side, although magnetic resonance images (MRI) and ultrasound are high-resolution images, they provide significantly less information about the changes at the molecular level [18], [19]. When doctors fail to identify tumors using other imaging modes, a biopsy is suggested [20]. Feature extraction is a critical step in classification as it helps improve accuracy and decrease time. The given input to the algorithm may be massive to handle and contains redundancy, so it is converted into a set of effective and functional features [21]. The group of features must contain the information required to perform the final task. The useful features promote the predictive capacity of the model. On the other side, if the features are irrelevant, then it may cause misleading. To pass the relevant features and avoid the negative transfer modality-specific transfer learning technique is deployed, which helps in enhancing the classification results by providing more suitable and related features.

The paper aims to develop a model for efficiently classifying breast histopathology images. The deep learning ensemble model is proposed by embedding the CBAM and GAP layer. The main contributions of this paper are mentioned below:

- This study introduces a novel Modality specific CBAM-VGGNet model for the classification of breast histopathology images. Since our model utilize both convolutional blocks VGGNet and attention module, it can acquire the most relevant regions from both local and global levels of the pathology images.
- We implemented the VGGNet in a novel way following the configuration of state-of-the-art VGGNet with our own concatenated layers. This strategy ultimately reduces the complexity and depth of the model. The resultant model has achieved better accuracy and performed better with fewer parameters.
- A hybrid pre-processing technique is proposed by considering the complexity of the H and E stained images. So, high-quality histopathology images can be obtained.
- We proposed modality-specific trained networks that are trained specifically on the same domain cancerous histopathology datasets. These pre-trained models are feature extractors and help reduce negative transfer. This help in extracting meaningful and relevant features, ultimately boosting the model's performance.
- To perform the comparative study, two different state-of-the-art CNN models trained on ImageNet and homogenous histopathology datasets are evaluated. We have compared the proposed technique with previous studies to evaluate its strength. The evaluated results describe that our proposed CBAM-VGGNet model outperforms the previous state-of-the-art methods.

The rest of the paper is arranged into the following sections. In section II, the literature review is provided. In section III,

the proposed detailed methodology, datasets, and data augmentation are presented. In section IV experimental results, hyper-parameter setting, and discussion are provided. In the last section V, the conclusion is provided with future direction.

## II. LITERATURE REVIEW

For a few years, plenty of studies have been done by different researchers for breast cancer classification using CNNs. Chuang Zhu et al. proposed a hybrid model by assembling multiple compact CNNs. This architecture contains a squeeze excitation pruning block and local and global branches. These modules help in reducing channel redundancy and provide string representation. Experiments are performed using BreakHis and BACH datasets. The multi-model assembling strategy achieves comparable results with state of the art models [22]. Togacar et al. developed a residual architecture (BreastNet) that contains attention modules for classification purposes. Attention modules help identify the important key regions in the processed histopathology images. The Hyper-column technique is utilized to achieve better results. BreastNet model with its dense, pooling, convolutional, and residual blocks achieves an accuracy of 98.80% [23]. EM Nejad et al. considered the importance of the significant features in the classification process and designed a new single-layer CNN. This model extracts salient features from the set of given histopathology images that belongs to the BreakHis dataset. This work achieves a 77.5% of recognition rate [24]. By using the statistical and structural information in the histopathology images, AA Nahid et al. classified the images into two categories. The authors propose a model that is a combination of LSTM and CNN. However, SVM and softmax are used at the final decision-making stage. The accuracy of 91.00% is attained [25]

Yan et al. proposed an automatic method for histopathology image analysis that takes advantage of both recurrent neural networks and CNN. BreakHis dataset is used during experiments, and 91.3% of accuracy is accomplished for the classification of four classes [26]. Deep convolutional neural network (DCNN) with backpropagation, ensemble, and ReLU activation function is utilized in [27] for the intra-class classification. 91.5% of accuracy is achieved for the eight-class classification using the BreakHis dataset. Xie et al. performed binary and multi-class classification. In this study, InceptionResNetV2 retrieves the input samples' features. A novel proposed autoencoder is used to transform those features into low dimensional space, which provides better clustering results [28]. By considering the good results of DCNN in previous studies, Alom et al. proposed a model of inception recurrent residual CNN [29].

Transfer always comes in handy when we are dealing with data scarcity. Sana Ullah et al. proposed a novel model to extract features from the pre-trained models, namely, VGGNet, GoogleNet, ResNet, and passed into the fully connected layers to perform the binary classification.

The proposed framework achieved 97.525% accuracy [30]. In another study performed by A Aloyayri et al. shuffleNet, InceptionV3, and ResNet18 architectures are implemented for the binary classification task. These pre-trained architectures are trained on ImageNet. The last layers of the models are fine tuned and trained on the BreakHis images. Authors achieve the highest accuracy of 98.73% [31]. Ahmad et al. used 240 training and 20 test images and classified them into four classes. Authors have applied transfer learning on ResNet, GoogleNet, and Alexnet. However, the maximum accuracy of 85% is obtained with the ResNet architecture [32]. Vesal et al. proposed a transfer learning-based approach to classify breast cancer into four sub-classes. In this study, BACH 2018 dataset is used. The extracted patches from these images are utilized to find the two well-known models, namely, ResNet50 and InceptionV3. Both models were trained on ImageNet. The best accuracy of 97.50% is achieved [33].

G Murtaza et al. proposed a precise model for the classification purpose using the BreakHis dataset and transfer learning. The AlexNet architecture is utilized, and the final layer of the model is fine-tuned to perform the binary classification. Six different machine learning classifiers then utilize the obtained features to perform classification into two classes. Furthermore, the misclassification-reducing (MR) algorithm is proposed. 81.25% accuracy is achieved in the experiments [34]. Alzubaidi et al. used the transfer learning technique and proposed a novel architecture that contains a residual link and parallel convolutional to accomplish a four-class classification task with 96.1% accuracy [21]. CA Ferreira et al. used transfer learning and deep neural network to perform breast cancer classification. The authors have used Fine-tuned Inception ResNet V2 model to perform the experiments using ICIAR 2018 dataset. The accuracy of 76% is achieved in this work [35]. Authors in [36] proposed a model that automatically classifies breast cancer metastases and uses a convolutional block attention module. The model is evaluated on the PCam dataset and achieves an accuracy of 0.976.

SH Kassani et al. proposed an ensemble deep learning-based approach for the automated binary classification of histopathology images. Three architectures named MobileNet, DenseNet, and VGG19 are ensemble and perform feature extraction. This model is tested on four publicly available datasets. MLP is utilized to make the final decision. The best accuracy is achieved with the BreakHis dataset, i.e., 98.13% [37]. In [38], authors evaluated VGGNet-16, which is trained on the ImageNet dataset to classify BreakHis and CMTHis datasets. They have achieved mean accuracy of 97% and 93%.

## III. METHODOLOGY

A hybrid pre-processing technique is used to remove the additive noise and enhance the contrast of the images. This technique is explained in section C. After improving the

image quality, we used data augmentation to increase the size of the dataset. Our main goal is to classify the breast cancer histopathology images with minimum error. We proposed a modality specific CBAM-VGGNet model that effectively helps classify histopathology images. The base models are prepared using the architectures of state-of-the-art VGG16 and VGG19 models and then training on the same domain cancerous histopathology datasets. The ImageNet weights are not utilized. These models are used as fixed feature extractors. These features are transferred to the proposed CBAM-VGGNet model, which contains VGG16 architecture but with misclassification-reducing layers and CBAM modules. The modified architecture of the VGGNet model is shown in Fig 5. Adding the GAP layer reduces the number of parameters in the model and makes it less complex. While on the other side, the CBAM module helps focus on what and where the most relevant information is present. To boost the model's performance, we have ensembled both VGG16 and VGG19 models and then performed the feature extraction. These features are then fed into the feed-forward neural network for the final classification. The proposed framework for breast cancer classification is shown in Fig 1.

Various experiments are performed with different combinations before selecting the final network. These experiments are performed to achieve the best classification results.

- The number of dense layers in the classifier has experimented with two and three. The number of filters is also modified in these layers to get the best results.
- The sigmoid and SoftMax functions are evaluated at the last classification layer.
- In models, stochastic gradient descent, ADAM, and Nadam optimizers with different learning rates are used during the pre-training stage.
- Two activation functions ReLU and tanh are also evaluated.
- We also experimented with the normal kernel initializer and without the kernel initializer.

#### A. MODALITY-SPECIFIC TRANSFER LEARNING AND FEATURE EXTRACTION STRATEGY

The shortage of substantial samples is a significant issue in training the DCNNs and continues, especially in the medical imaging field. Transfer learning is used to deal with the problem of the limited dataset [39]. It is a mechanism in which CNN's trained on one problem (large data) are used to solve another similar problem (small dataset). Nowadays, it is a famous technique in deep learning because it can train CNNs with very fewer data. In previous studies, state-of-the-art models like VGGNet, ResNet, GoogleNet, and Alexnet trained on natural images ImageNet dataset have been greatly utilized for transfer learning and expressed better results. However, these pre-trained models cannot significantly improve the performance when dealing with medical imaging tasks, especially histopathology images [15]. To test

this theory, we have employed the modality-specific transfer learning strategy in this study.

#### B. DATASETS

We employed four publicly available datasets. Three of them are used as the source datasets for the pre-training purpose, and all belong to multiple organ cancerous histopathology datasets (PCam, ICIAR 2018, lung, and colon). One of them (BreakHis) is used as the targeted dataset.

##### 1) ICIAR 2018

The ICIAR 2018 [40] dataset belongs to the BACH challenge. It is basically a widened version of the bio-imaging dataset 2015. It consists of 24 bits of RGB H and E stained breast histology images that are retrieved from whole slide image biopsies. These images are available in 200X magnification factor with a pixel size of 0.42um x .42um and .tiff format. Images are divided into four classes, namely, invasive carcinoma, benign, normal, and in situ carcinoma. Each class contains 100 images, i.e., a total of 400 images. As we performed binary classification, we grouped benign and normal images into benign classes. Invasive carcinoma and Situ are combined in the malignant class.

##### 2) LUNG AND COLON

The lung and colon cancer dataset [41] contains a total of 750 histopathological images, which belong to five classes. Each class contains 250 images and 768 × 768-pixel size in the jpeg file format. Originally these images are obtained from a sample of HIPAA complaints and validated sources. However, these images are augmented using the Augmentor package, and a total of 25000 images are obtained. After augmentation, each class contains 5000 images.

##### 3) PCam

A reformed category of the PCam dataset is accessible at [26] containing 327,680 images belonging to benign and malignant classes. These images are obtained from the whole slide histopathology images having binary labels indicating metastatic tissue. The pixel size of these data samples is 96 × 96. In this study, we are using the reformed version of the dataset, as the original samples contain duplicate images due to probabilistic sampling.

##### 4) BreakHis

The breast cancer histopathology image classification (BreakHis) [42] comprises 9109 microscopic images that belong to 82 patients. These images are available in (40X, 100X, 200X, and 400X) magnification factors. These images contain a pixel size of 700 × 460 having PNG format. This dataset contains two major groups i.e., benign and malignant. The benign class contains 2480 images while the malignant class consists of 5429 samples. This database has been developed in participation with the P&D laboratory.

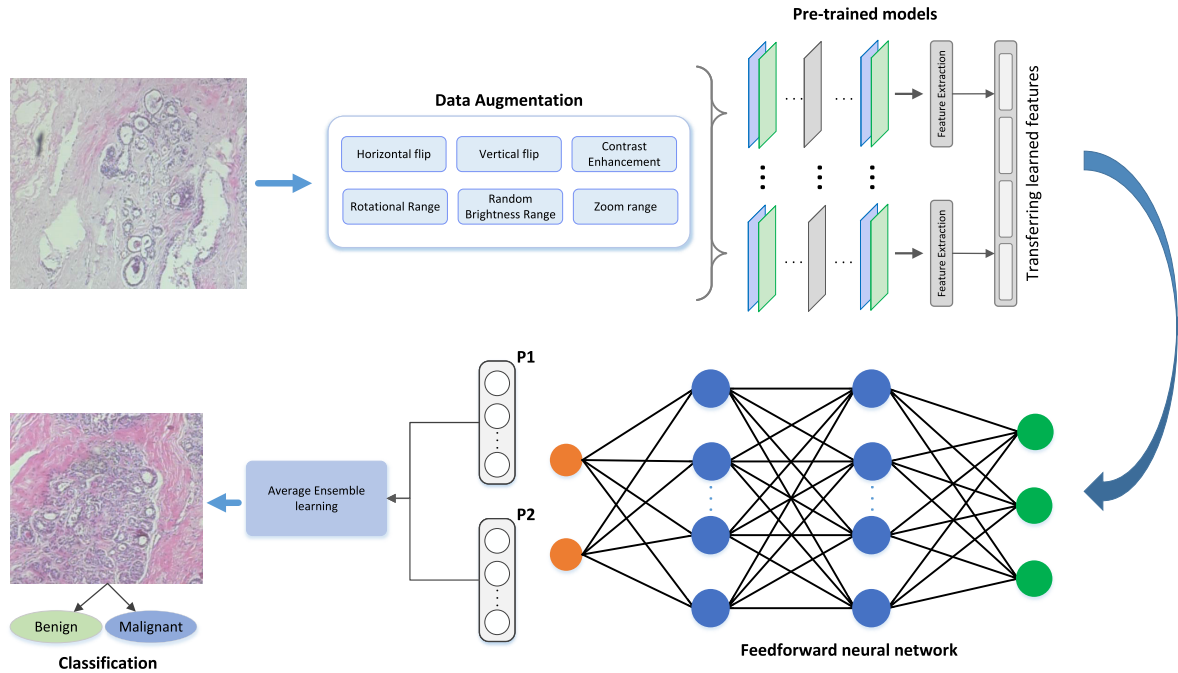


FIGURE 1. Proposed system model framework.

C. DATA PREPROCESSING

Histopathology images are very challenging for deep learning algorithms because of their complex nature. They contain a lot of information regarding image texture and hence deliver successful cancer diagnostic results [43]. The complicated background and distributing factors in histopathology images greatly slow down the decision-making process and success rate [44]. Image processing algorithms can help to deal with such unwanted conditions. In this study, we have utilized a hybrid method for pre-processing histopathology images using a median filter followed by the CLAHE technique. The results of the proposed pre-processing technique are given in Fig 2.

1) ZERO PADDING

Zero padding helps in correctly handling the edge pixels. While processing the border regions, extra rows and columns around the image must be added, so that it extends beyond the boundary of the image.

2) MEDIAN FILTER

Noise in histopathology images needs to be removed for better cancer diagnosis. Noise removal is highly significant and directly correlates with the quality of the image for further processing [45]. In histopathology images, some tissues and cells contain additive noise as they are captured with the digital camera [46]. In additive noise, a certain value is uniformly added over all the pixels of the images. Different image-denoising filters like Wiener, Median, Mean, and Gaussian are available to deal with this noise. However, we have used

the Median filter because of its better results in previous studies.

The median filter is a type of non-linear filter that is mostly used to remove noise in image processing. The main advantage of using the median filter is that it preserves the edges of the image [47], [48]. Edges are considered a significant feature for analyzing images. It moves pixel by pixel over the image, and it replaces each value with the median value of the neighboring pixels. The pattern of neighbors is known as a window that moves over the image, pixel by pixel. It also removes the noise values with larger intensity. The mathematical representation of the median filter is given below.

$$y(t) = \text{median}((x(t - T/2), x(t - T + 1), \dots, x(t), \dots, x(t + T/2))) \tag{1}$$

3) CLAHE

In histopathology images, artifacts, blurriness, overlapping, and weak boundary detection problem occur due to uneven slide staining. CLAHE algorithm helps in dealing with these issues as it focuses on improving the local contrast of the image and the weak boundary edges in each pixel of an image through the limited amplification [49]. In adaptive histogram equalization, the whole image is divided into small blocks called “tiles”. After that, histogram equalization is performed on each of these tiles.

4) PRE-PROCESSING ALGORITHM

**Inputs:** Histopathology image

**Outputs:** Pre-processed image

- Convert RGB histopathology image into grey scale

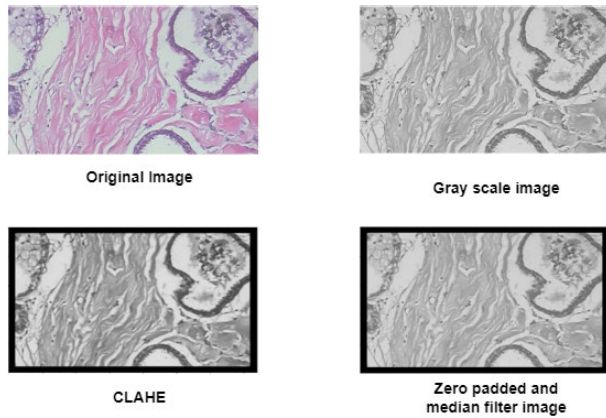


FIGURE 2. Hybrid pre-processing technique.

TABLE 1. Data augmentation metrics.

Technique	Value
Contrast enhancement	Alpha = 1.1
Horizontal flip	True
Vertical flip	True
Rotational range	90 degree
Random brightness	0.2
Zoom	0.2

- Perform zero padding
- Apply median filter on the grey scale image with  $3 \times 3$  neighborhood
- Apply contrast specific adaptive histogram equalization

**D. DATA AUGMENTATION**

In medical images, obtaining many samples is always a problem. However, it is mandatory to have enough images to obtain good results in deep learning. We have applied augmentation on our target dataset, i.e., BreakHis using various techniques to handle the data scarcity and overfitting. We have augmented our data 6 times. Table 1 shows the parameter values for these augmentation techniques.

**E. CBAM-VGGNET ARCHITECTURE**

In this section, we present the proposed modality specific CBAM-VGGNet architecture for the classification of H and E stained breast histopathology images and the theoretical details of the proposed network. The major blocks of the proposed architecture are given below.

**1) VGGNet**

The original VGGNet model is a CNN developed by the Visual Geometry Group at the University of Oxford [50]. VGGNet architecture is characterized by  $3 \times 3$  convolutional kernels and  $2 \times 2$  pooling layers. The network depth can be improved using small convolutional layers that enhance feature learning. VGG has three fully connected layers. The first two contain 4096 channels, and the third has 1000 channels for each class. All VGG hidden layers use ReLU. The two

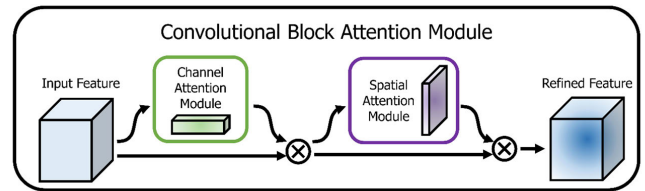


FIGURE 3. CBAM Module.

most commonly used VGGNet architectures are VGG16 and VGG19, which we have used in this study.

**2) STAINED P LAYER**

The global average pooling layer averages each feature map separately and provides a single value. It performs the average operation per channel. We can think of each feature map as the final feature representation per category over which we want to do the classification.

**3) CBAM MODULE**

CBAM given by Woo et al. [51] performs a spatial operation and channel size in the CNN network. It contains two major modules named spatial attention module and channel attention module. The complete CBAM module is given in Fig 3. While both modules are separately given in Fig 4.

*a: CHANNEL ATTENTION MODULE:*

It is developed by utilizing the relations of the features among channels. It focuses on what is crucial in the given input. At first, to achieve the values of  $F_{cavg}$ , and  $F_{cmax}$ , the spatial information available in feature maps is obliged to average and maximum pooling processes. These values are then fed to the network  $McRc \times 1 \times 1$  containing one hidden layer. Where C indicates the channel number. Channel attention is obtained by using equations 2 and 3, which are given below.

$$M_c(F) = (MLP(AvgPool(F)) + MLP(MaxPool(F))) \tag{2}$$

$$M_c(F) = (W1(W0(F_{cavg})) + W1(W0(F_{cmax}))) \tag{3}$$

Here is the sigmoid function, W0 and W1 shows the weight values, MLP indicates multi-layer perceptron.

*b: SPATIAL ATTENTION MODULE:*

This module uses the relationship of properties among space to create a spatial attention map. Spatial module focus on where crucial information is present in the input. Maximum and average pooling operations are performed to achieve the spatial attention maps. These are the values of  $F_{savg} R1 \times H \times W$  and  $F_{smax} R1 \times H \times W$  respectively. The obtained values are combined, and a convolution layer such as  $M_s(F) R H \times W$  is applied. Equation 4 and 5 is used to obtain

the spatial map.

$$Ms(F) = (f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \quad (4)$$

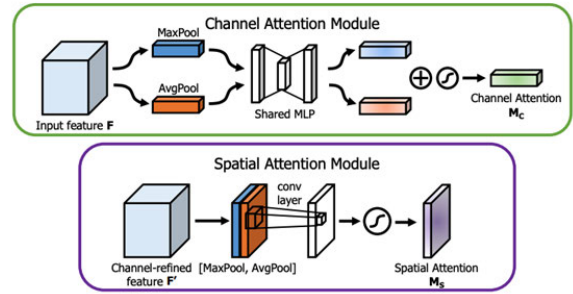
$$Ms(F) = (f^{7 \times 7}([F_{\text{savg}}; F_{\text{smax}}])) \quad (5)$$

Here specifies the sigmoid function, and  $f^{7 \times 7}$  shows the  $7 \times 7$  convolution process. The significance of attention has been studied extensively in the literature [52], [53], [54], [55], [56]. Attention not only tells where to focus but also improves the representation of interests. The attention mechanism has improved the representation power of our model: focusing on essential features and suppressing unnecessary ones. The CBAM module used in this paper emphasizes meaningful features along both spatial and channel axes. These are applied sequentially so that every branch can learn what and where to attend in the channel and spatial axes, respectively. As a result, our model helps information flow within the network by learning the information to emphasize or suppress. In the channel attention module, both average-pooled and max-pooled features are used simultaneously and greatly improve the representation power of the model rather than using each independently. On the contrary, in the spatial attention module, first, we applied the average pooling and then the max pooling operation along the channel axis and concatenated them to generate an efficient feature descriptor. After that, a convolutional layer is applied to generate a spatial attention map that encodes where to emphasize or suppress. We have used the channel first and then the spatial module. The sequential arrangement provides better results than the parallel arrangement.

**F. MODIFIED CBAM-VGGNet MODEL**

We have added a GAP layer after every first layer of the convolutional block that reduces the number of parameters by reducing the network depth compared to the original VGGNet to avoid overfitting and underfitting problems during training. The GAP layer acts as the downsampling. The downsampling layer is mainly used to improve the anti-distortion ability of the network to the image while retaining the main features of the sample and reducing the number of parameters [57]. The GAP layers of all convolutional blocks are connected. After every GAP layer, we added the CBAM module to refine the feature maps. In the end, we have concatenated the output of the CBAM blocks.

We have compared the results of our proposed modality specific transfer learning technique with the state-of-the-art models trained on ImageNet datasets. PCam, ICIAR 2018, lung, and colon datasets are used for the pre-training. However, the BreakHis dataset is used for evaluation purposes. In this study, we have utilized transfer learning as feature extraction. Although, deep learning can automatically retrieve the features. However, we still need to know which features are useful according to our problem. CNN’s have the potential to figure out which features are relevant and need to be considered. The learned features can then be further used for another similar problem



**FIGURE 4. Channel and spatial attention module.**

**TABLE 2. Running time.**

Model	Parameters	Running time
VGG16	9,995,072	1.056 hours
VGG19	12,944,960	1.208 hours

**TABLE 3. Hyper parameters.**

Property	Value
Batch size	32
Epoch	1000
Learning rate	0.0001, beta1 0.6, beta2 0.2
Optimizer	Adam
Loss function	Binary cross entropy
Activation	Softmax

**IV. EXPERIMENTAL RESULTS AND DISCUSSIONS**

**A. EXPERIMENTAL PLATFORM**

The images are trained and tested using the Keras package with Tensorflow as the deep learning framework backend. The NVIDIA Tesla K80 16GB memory is utilized. The 128GB RAM was available to perform these experiments.

**B. HYPER PARAMETER OPTIMIZATION**

Hyper-parameters are variables with pre-determined values that govern the whole training process of CNN’s. Although neural networks learn the links between outputs and inputs, hyper-parameter tuning is necessary to get better performance for the model. To find out the best hyper-parameters for final training, we have performed a number of experiments. We have evaluated two different optimizers i.e., Adam (Adaptive moment estimation) and SGD (stochastic decent gradient), with three distinct learning rates (0.01, 0.001, and 0.0001). SoftMax classifier is utilized to perform all the classification experiments. It gives the probability for each class label. The probability range is between 0 and 1, and the sum of all probabilities will equal 1. Binary cross-entropy is used as we are working on a binary classification problem. The mathematical representation of the binary cross entropy is given in equation 6. The fully connected layer has a ReLU activation function and 256 hidden neurons. It is followed by a dropout layer having a probability of 0.4 utilized. The dense layer is added to solve the binary classification problem. The final parameters that we have selected for training are given in Table 3. While the running time of both models is given

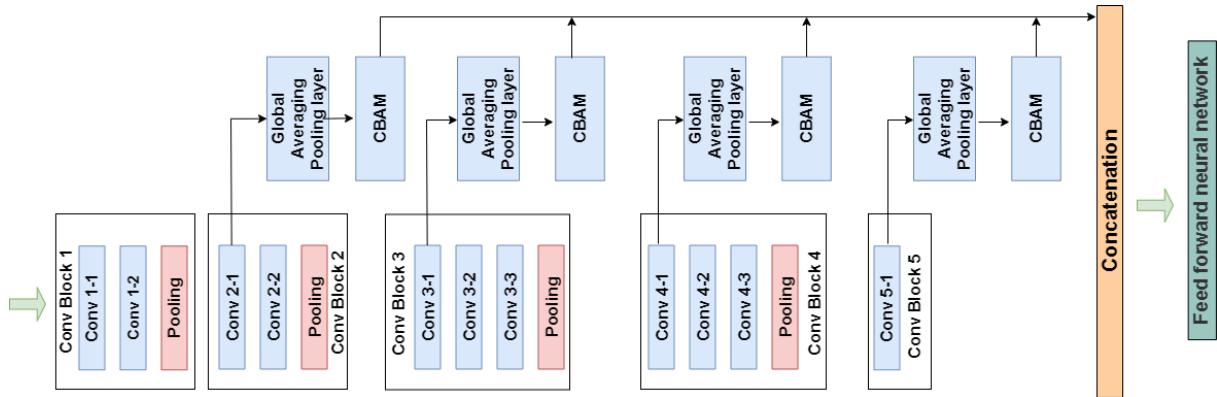


FIGURE 5. Modified VGGNet Model.

TABLE 4. Results without transfer learning.

Model	Accuracy	Specificity	Sensitivity	F1-score
VGG16	83.26	84.86	78.75	81.94
VGG19	81.69	83.63	73.32	82.35

in Table 2.

$$\text{Binary cross entropy} = (y_i \times \log(p(y_i)) + (1 - y_i) \times \log(1 - p(y_i))) \quad (6)$$

### C. RESULTS

This study used a modality specific transfer learning strategy for the classification task. We have also analyzed the performance of state-of-the-art CNN architectures to compare with the proposed strategy.

We have conducted various experiments to perform the binary classification of breast cancer histopathology images. To prove the effectiveness of our strategy, we have given the results of each model and also a comparative analysis. First, we have performed experiments on two different models without transfer learning which is given in Table 4. Then, we analyzed the results of the state-of-the-art models trained on ImageNet, shown in Tables 5 and 6. We have given the detailed results of the proposed modality specific strategy in Tables 6 and 7. Experiments are performed separately on the complete dataset and on four distinct magnification factors (40X, 100X, 200X, and 400 X). The results using the modality-specific transfer learning on CBAM-VGGNet are given in Table 8. We also compared the results with some previous studies in Table 11. Accuracy, F1 score, sensitivity, and specificity are used as the evaluation metrics.

#### 1) RESULTS OF STATE-OF-THE-ART MODELS TRAINED ON ImageNet

In the second part of the experiments, two state-of-the-art models trained on ImageNet are used to perform experiments. Among these models, VGG16 achieved the best accuracy

TABLE 5. Results of state-of-the-art models trained on ImageNet.

Model	MG factor	Acc	Spec	Sens	F1-score
VGG16	40X	88.22	86.26	95.43	91.99
	100X	90.47	88.06	93.64	91.30
	200X	90.03	89.56	90.57	90.54
	400X	87.02	89.43	88.60	89.24
VGG19	40X	84.96	88.21	77.31	89.61
	100X	87.17	86.33	88.21	88.09
	200X	86.59	84.82	88.79	87.54
	400X	87.37	87.2	87.55	87.77

TABLE 6. Results of ImageNet on complete BreakHis dataset (40X, 100X, 200X, 400X).

Model	Acc	Spec	Sens	F1-score
VGG16	86.42	90.80	80.00	90.84
VGG19	84.45	87.30	77.41	88.87

of 90.47% on 100X data. While the accuracy of VGG16 on 40X, 200X, and 400X is 88.22%, 90.03%, and 87.02% on BreakHis dataset. The lowest performance, in this case, is observed by the VGG19 model at 40X magnification factor i.e., 84.96. Accuracy, specificity, sensitivity, and F1-score of both models on different magnification factors are shown in above Table 5. In Table 6, the results for complete BreakHis dataset are given.

After that, we performed experiments using the proposed methodology. The customized data is used for pre-training models instead of those trained on ImageNet. These models perform modality specific transfer learning as they are trained on cancerous histopathology datasets. In this case, when we are taking advantage of the same domain transfer learning, VGG16 has provided the best results. However, these classification results are better than the results of the ImageNet.

VGG16 provides the best accuracy of 91.36% on a 100X magnification factor. While in the case of ImageNet, the highest accuracy is 90.47%. This proves that better features are learned and transferred while performing modality-specific transfer learning instead of the cross-domain. These results are shown in Table 7.



**TABLE 7. Results of proposed modality-specific strategy.**

Model	MG factor	Acc	Spec	Sens	F1-score
VGG16	40X	88.71	94.99	87.35	92.00
	100X	91.36	90.15	88.48	89.99
	200X	90.51	93.05	87.66	90.46
	400X	86.67	84.64	88.91	87.54
	Complete dataset	95.00	97.90	92.26	96.30

**TABLE 8. Results of proposed CBAM-VGGNet VGG16 model.**

Model	MG factor	Acc	Spec	Sens	F1-score
VGG16	40X	93.98	95.95	89.76	95.60
	100X	95.42	96.14	94.63	95.63
	200X	95.21	93.33	97.46	95.51
	400X	98.76	98.78	98.72	98.34
	Complete dataset	95.00	97.90	92.26	96.30

**TABLE 9. Results of modality specific CBAM-VGGNet VGG19 model.**

Model	MG factor	Acc	Spec	Sens	F1-score
VGG19	40X	95.73	96.05	95.00	96.92
	100X	97.98	98.96	96.95	98.07
	200X	96.24	96.75	95.08	97.27
	400X	98.22	99.18	98.73	98.98
	Complete dataset	96.01	98.75	90.44	97.00

**TABLE 10. Ensemble modality specific CBAM-VGGNet model results.**

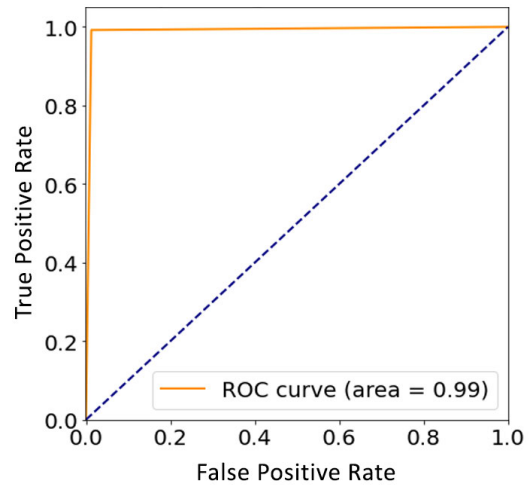
Model	MG factor	Acc	Spec	Sens	F1-score
Proposed	40X	94.44	94.36	94.78	96.05
	100X	97.61	98.24	96.93	97.77
	200X	98.70	96.42	99.17	97.82
	400X	98.96	98.76	97.08	97.95
	Complete dataset	97.66	98.51	95.80	98.29

2) RESULTS OF PROPOSED MODALITY-SPECIFIC STRATEGY ON COMPLETE BreakHis DATASET (40X, 100X, 200X, 400X) To further boost the classification capacity of our best-performing model, we have utilized the Modality specific CBAM-VGGNet model described above. In this case, the highest accuracy (98.76%) is attained on the 400X magnification factor. The results of the Modality-specific CBAM-VGGNet VGG16 model on four magnification factors and the complete dataset are shown in Table 8 below. While for VGG19, the results are shown in Table 9.

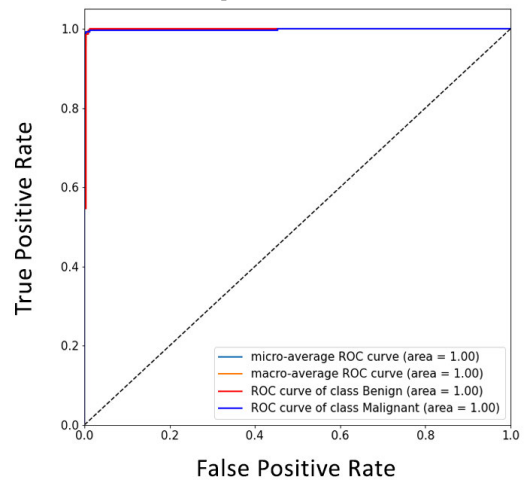
After performing this experiment, we combined VGG16 and VGG19 using averaging ensemble method which is our final model. Researchers have argued that ensemble models provide more generalized results than single models, so we have adopted the ensemble architecture [58]. The results indicate that combining these two models provide better results which are given in Table 10. The ROC curves for the 400X and complete BreakHis dataset are given in Fig 6 and 7.

3) COMPARATIVE ANALYSIS WITH PREVIOUS METHODS

A relative analysis of the proposed technology with five previous methods is performed and given in Table 11 to show the supremacy of the modality specific transfer learning technique. It is analyzed that our proposed technique provides



(a) Empirical ROC curve



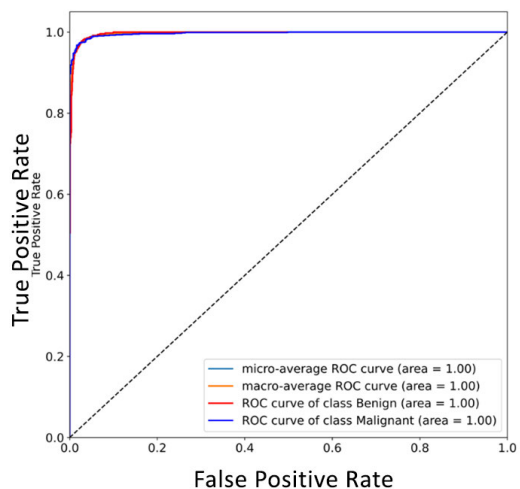
(b) Extension of ROC curve for every class

**FIGURE 6. ROC curve for Modality specific CBAM VGGNet (400X).**

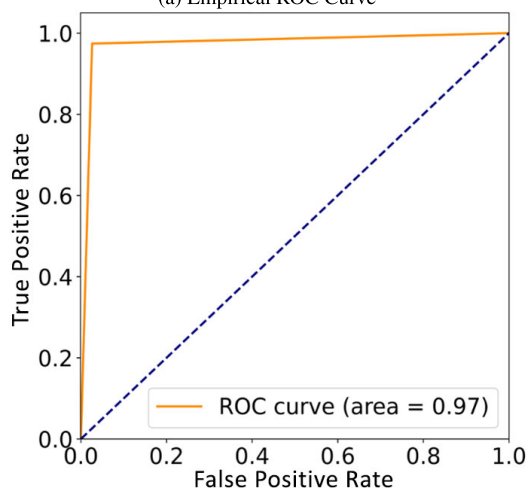
higher accuracy than all these methods. This table shows the strength of our proposed strategy in terms of accuracy compared to similar methods. Our proposed strategy has outperformed these previous studies.

Erfankhah et al. [59] have used BreakHis along with KimiaPath24 and invasive ductal carcinoma (IDC) datasets to perform feature extraction. The texture features are extracted from the circularly symmetric pixel neighborhood. Histogram values are also utilized to differentiate between the heterogeneous and homogeneous regions of the image tissue. In this study, the LBP method is used in which feature extraction was based on the homogenous and heterogeneous status of the tissue regions. In the end, they used SVM as the classifier and achieved an accuracy of 88.30% using the BreakHis dataset.

In Another study by Lichtbalu and Stoean [60], the authors used six machine learning techniques to perform the classification task. AlexNet model is utilized; however, no fine-tuning is performed. Five classifiers named (random forest, LR, Naïve Bayes, SVM, and nearest neighbor) containing



(a) Empirical ROC Curve



(b) Extension of ROC curve for every class

**FIGURE 7. ROC curve for Modality-specific CBAM VGGNet on complete BreakHis dataset.**

various features are used. Classification is performed using each classifier individually. Furthermore, the Fourier trig transform and the principal component analysis were also used in the model. The highest accuracy of 86.67% was achieved in this study.

Gu et al. [61] have generated the discriminative binary codes using the DCMM model’s histopathology images. To fully utilize the advantage of cross-magnification information DCMM model with mutual guidance learning paradigm is conducted. It is based on the densely connected architecture and low-high magnification data pairs. Different magnification levels are used to learn the binary codes from the BreakHis data. The best accuracy is achieved using a 400X magnification factor, i.e., 96.31%

Nahid et al. [62] have used CNN containing residual block to classify histopathology images. In this study, CNN takes the raw images as the input and extracts the Global features. The authors used CNN in five cases to retrieve more information from the images. (A) CNN Raw image (B) CNN CT

**TABLE 11. Comparative analysis with previous studies.**

Method	Model	MG factor	Acc
Erfankhan Hamed, et al	LBP	40X	88.30
		100X	88.30
		200X	87.10
		400X	83.40
Daniel Lichtblau et al	AlexNet	40X	81.61
		100X	84.47
		200X	86.67
		400X	83.15
Yun Gu et al	DCMM	40X	95.62
		100X	95.03
		200X	97.04
		400X	96.31
Nahid et al	Deep CNN	40X	90.00
		100X	90.00
		200X	91.00
		400X	90.00
Togacar mesut et al	Deep CNN	40X	97.99
		100X	97.84
		200X	98.51
		400X	95.88
Yan Hao et al	DenseNet 201	40X	96.75
		100X	95.21
		200X	96.57
		400X	93.15
Pin Wang et al	FE-BkCapsNet	40X	92.71
		100X	94.52
		200X	94.03
		400X	93.54
Proposed (VGG16 + VGG19)	2022	40X	94.44
		100X	97.61
		200X	98.70
		400X	98.96

histogram (C) CNN CT LBP (D) CNN discrete Fourier transform (E) CNN discrete cosine transform. All experiments are performed using the BreakHis dataset. The best accuracy is achieved using CNN CT histogram on 200X magnification factor, i.e., 91.00%

Togacar et al. [23], proposed a novel model named as BreasNet that contains residual architecture built on attention modules. BreakHis dataset is used to perform experiments; however, no increase in data is performed. Data augmentation is performed to enhance the images. After augmentation, each image is processed through the model to identify the important key regions present in the data. The Hypercolumn technique is also utilized to perform better classification in this study. This novel model contains convolutional, residual, dense, and pooling blocks.

Hao et al. [63] used DenseNet201 as the basic model to extract the deep semantic features. These features are fused with the gray-level co-occurrence matrix (GLCM) features. The authors have used a support vector machine for classification and BreakHis dataset for all experiments. Two different experimental sets are considered: magnification specific binary (MSB) and magnification independent binary (MIB), and the highest accuracy of 96.57% is achieved.

Wang et al. [64] proposed the FE-BkCapsNet model, which takes advantage of CNN and CapsNet (to focus on detailed information about position and posture). By combining this information, more discriminative features can be obtained. The model was tested on the BreakHis dataset, and the best accuracy of 94.52% was obtained.

## V. CONCLUSION AND FUTURE WORK

This work proposes a CBAM-VGGNet model for the binary classification of histopathology images. The main objective is to classify malignant images effectively. The results section shows that the proposed modality-specific transfer learning provides better classification outcomes as compared to state-of-the-art models trained on ImageNet. The same domain pre-trained models have outperformed the previous state-of-the-art models as more relevant features are transferred by using our approach. Furthermore, the CBAM focuses on important features, while the GAP layer makes the model less complex. In future work, we extend it to perform the grading of breast cancer. Also, using different feature selection techniques after feature extraction can further help reduce collinear features.

## REFERENCES

- M. F. Akay, "Support vector machines combined with feature selection for breast cancer diagnosis," *Exp. Syst. Appl.*, vol. 36, no. 2, pp. 3240–3247, 2009.
- E. G. Fischer, "Nuclear morphology and the biology of cancer cells," *Acta Cytologica*, vol. 64, no. 6, pp. 511–519, 2020.
- F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, 2018.
- F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte, "Deep features for breast cancer histopathological image classification," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2017, pp. 1868–1873.
- S. R. Lakhani, "WHO classification of Tumours of the breast," WHO, France, Tech. Rep., 2012. [Online]. Available: <https://espace.library.uq.edu.au/view/UQ:8984059>
- A. Sikandar, "Histopathology: An old yet important technique in modern science, in *Histopathology—An Update*. London, U.K.: IntechOpen, 2018.
- M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 147–171, 2009.
- Z. Hameed, S. Zahia, B. Garcia-Zapirain, J. J. Aguirre, and A. M. Vanegas, "Breast cancer histopathology image classification using an ensemble of deep learning models," *Sensors*, vol. 20, no. 16, p. 4373, Aug. 2020.
- R. Rubin and D. S. E. Strayer Rubin, *Rubin's Pathology: Clinicopathologic Foundations of Medicine*. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2008.
- L. Duan, D. Xu, and I. Tsang, "Learning with augmented features for heterogeneous domain adaptation," 2012, *arXiv:1206.4660*.
- B. Kulis, K. Saenko, and T. Darrell, "What you saw is not what you get: Domain adaptation using asymmetric kernel transforms," in *Proc. CVPR*, Jun. 2011, pp. 1785–1792.
- C. Wang and S. Mahadevan, "Heterogeneous domain adaptation using manifold alignment," in *Proc. 22nd Int. Joint Conf. Artif. Intell.*, 2011, pp. 1541–1546. [Online]. Available: <https://dl.acm.org/doi/10.5555/2283516.2283652>
- S. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
- E. Baykal, "Transfer learning with pre-trained deep convolutional neural networks for serous cell classification," *Multimedia Tools Appl.*, vol. 79, pp. 15593–15611, Jun. 2020.
- M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, "Transfusion: Understanding transfer learning for medical imaging," 2019, *arXiv:1902.07208*.
- W. Zhang, L. Deng, L. Zhang, and D. Wu, "A survey on negative transfer," 2020, *arXiv:2009.00909*.
- N. F. Boyd, "Mammographic densities and breast cancer risk," *Cancer Epidemiol. Prevention Biomarkers*, vol. 7, no. 12, pp. 1133–1144, 1998.
- P. A. Carney, C. J. Kasales, A. N. A. Tosteson, J. E. Weiss, M. E. Goodrich, S. P. Poplack, W. S. Wells, and L. Titus-Ernstoff, "Likelihood of additional work-up among women undergoing routine screening mammography: The impact of age, breast density, and hormone therapy use," *Preventive Med.*, vol. 39, no. 1, pp. 48–55, Jul. 2004.
- T. A. Huisman, "Fetal magnetic resonance imaging of the brain: Is ventriculomegaly the tip of the syndromal iceberg?" in *Seminars in Ultrasound, CT and MRI*. Amsterdam, The Netherlands: Elsevier, 2011.
- N. Maroof, A. Khan, S. A. Qureshi, A. U. Rehman, R. K. Khalil, and S.-O. Shim, "Mitosis detection in breast cancer histopathology images using hybrid feature space," *Photodiagnosis Photodynamic Therapy*, vol. 31, Sep. 2020, Art. no. 101885.
- L. Alzubaidi, O. Al-Shamma, M. A. Fadhel, L. Farhan, J. Zhang, and Y. Duan, "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, p. 445, Mar. 2020.
- C. Zhu, F. Song, Y. Wang, H. Dong, Y. Guo, and J. Liu, "Breast cancer histopathology image classification through assembling multiple compact CNNs," *BMC Med. Informat. Decis. Making*, vol. 19, no. 1, pp. 1–17, Dec. 2019.
- M. Togaçar, K. B. Özkurt, B. Ergen, and Z. Cömert, "BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer," *Phys. A, Stat. Mech. Appl.*, vol. 545, May 2020, Art. no. 123592.
- E. M. Nejad, L. S. Affendey, R. B. Latip, and I. B. Ishak, "Classification of histopathology images of breast into benign and malignant using a single-layer convolutional neural network," in *Proc. Int. Conf. Imag., Signal Process. Commun.*, Jul. 2017, pp. 50–53.
- A.-A. Nahid, M. A. Mehrabi, and Y. Kong, "Histopathological breast cancer image classification by deep neural network techniques guided by local clustering," *BioMed Res. Int.*, vol. 2018, pp. 1–20, Mar. 2018.
- R. Yan, F. Ren, Z. Wang, L. Wang, T. Zhang, Y. Liu, and X. Rao, "Breast cancer histopathological image classification using a hybrid deep neural network," *Methods*, vol. 173, pp. 52–60, Feb. 2019.
- S. A. Adeshina, A. P. Adedigba, A. A. Adeniyi, and A. M. Aibinu, "Breast cancer histopathology image classification with deep convolutional neural networks," in *Proc. 14th Int. Conf. Electron. Comput. Comput. (ICECCO)*, Nov. 2018, pp. 206–212.
- J. Xie, R. Liu, J. Luttrell, and C. Zhang, "Deep learning based analysis of histopathological images of breast cancer," *Frontiers Genet.*, vol. 10, p. 80, Feb. 2019.
- M. Z. Alom, C. Yakopcic, M. S. Nasrin, T. M. Taha, and V. K. Asari, "Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network," *J. Digit. Imag.*, vol. 32, no. 4, pp. 605–617, Aug. 2019.
- S. Khan, N. Islam, Z. Jan, I. U. Din, and J. J. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognit. Lett.*, vol. 125, pp. 1–6, Jul. 2019.
- C. H. Muntean and M. Chowkkar, "Breast cancer detection from histopathological images using deep learning and transfer learning," in *Proc. 7th Int. Conf. Mach. Learn. Technol. (ICMLT)*. Cham, Switzerland: Springer, Mar. 2022, pp. 164–169.
- H. M. Ahmad, S. Ghuffar, and K. Khurshid, "Classification of breast cancer histology images using transfer learning," in *Proc. 16th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2019, pp. 812–819.
- S. Vesal, N. Ravikumar, A. Davari, S. Ellmann, and A. Maier, "Classification of breast cancer histology images using transfer learning," in *Proc. Int. Conf. Image Anal. Recognit.* Cham, Switzerland: Springer, Jan. 2018, pp. 812–819. [Online]. Available: [https://link.springer.com/chapter/10.1007/978-3-319-93000-8\\_92](https://link.springer.com/chapter/10.1007/978-3-319-93000-8_92)
- G. Murtaza, L. Shuib, A. W. A. Wahab, G. Mujtaba, G. Mujtaba, G. Raza, and N. A. Azmi, "Breast cancer classification using digital biopsy histopathology images through transfer learning," *J. Phys., Conf.*, vol. 1339, no. 1, Dec. 2019, Art. no. 012035.
- C. A. Ferreira, "Classification of breast cancer histology images through transfer learning using a pre-trained inception ResNet V2," in *Proc. Int. Conf. Image Anal. Recognit.* Cham, Switzerland: Springer, 2018, pp. 763–770.
- Y. Liang, J. Yang, X. Quan, and H. Zhang, "Metastatic breast cancer recognition in histopathology images using convolutional neural network with attention mechanism," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2019, pp. 2922–2926.
- S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Classification of histopathological biopsy images using ensemble of deep learning networks," 2019, *arXiv:1909.11870*.
- A. Kumar, S. K. Singh, S. Saxena, K. Lakshmanan, A. K. Sangaiah, H. Chauhan, S. Shrivastava, and R. K. Singh, "Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer," *Inf. Sci.*, vol. 508, pp. 405–421, Jan. 2020.

- [39] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Jan. 2009.
- [40] G. Aresta, "BACH: Grand challenge on breast cancer histology images," *Med. Image Anal.*, vol. 56, pp. 122–139, Aug. 2019.
- [41] A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wilson, L. A. DeLand, and S. M. Mastorides, "Lung and colon cancer histopathological image dataset (LC25000)," 2019, *arXiv:1912.12142*.
- [42] F. A. Spanhol, L. S. Oliveira, L. Heutte, and C. Petitjean, "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016.
- [43] J. Shi, J. Wu, Y. Li, Q. Zhang, and S. Ying, "Histopathological image classification with color pattern random binary hashing-based PCANet and matrix-form classifier," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 5, pp. 1327–1337, Sep. 2017.
- [44] S. Öztürk and B. Akdemir, "Effects of histopathological image pre-processing on convolutional neural networks," *Proc. Comput. Sci.*, vol. 132, pp. 396–403, Jan. 2018.
- [45] W. Wu, C. Parmar, P. Grossmann, J. Quackenbush, P. Lambin, J. Bussink, R. Mak, and H. J. W. Aerts, "Exploratory study to identify radiomics classifiers for lung cancer histology," *Frontiers Oncol.*, vol. 6, p. 71, Mar. 2016.
- [46] A. Danielyan, V. Katkovnik, and K. Egiazarian, "BM3D frames and variational image deblurring," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1715–1728, Apr. 2012.
- [47] J. C. Church, Y. Chen, and S. V. Rice, "A spatial median filter for noise removal in digital images," in *Proc. IEEE SoutheastCon*, Apr. 2008, pp. 618–623.
- [48] D. Maheswari and V. Radha, "Noise removal in compound image using median filter," *Int. J. Comput. Sci. Eng.*, vol. 2, no. 4, pp. 1359–1362, 2010.
- [49] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graph. Gems*, vol. 2, no. 4, pp. 474–485, 1994. [Online]. Available: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=ba974749d478e87b3874aab643d481595cd399aa>
- [50] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
- [51] S. Woo, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.
- [52] V. Mnih, N. Heess, and A. Graves, "Recurrent models of visual attention," in *Proc. Neural Inf. Process. Syst. (NIPS)*, 2014, pp. 1–9.
- [53] J. Ba, V. Mnih, and K. Kavukcuoglu, "Multiple object recognition with visual attention," 2014, *arXiv:1412.7755*.
- [54] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*.
- [55] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, "Show, attend and tell: Neural image caption generation with visual attention," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 2048–2057.
- [56] K. Gregor, I. G. A. Danihelka, D. J. Rezende, and D. Wierstra, "DRAW: A recurrent neural network for image generation," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 1462–1471.
- [57] J. Xiao, J. Wang, S. Cao, and B. Li, "Application of a novel and improved VGG-19 network in the detection of workers wearing masks," *J. Phys., Conf.*, vol. 1518, no. 1, Apr. 2020, Art. no. 012041.
- [58] C. U. Chiagoziem, "Multi-classification of breast cancer lesions in histopathological images using DEEP\_Pachi: Multiple self-attention head," *Diagnostics*, vol. 12, no. 5, p. 1152, 2022.
- [59] H. Erfankhah, M. Yazdi, M. Babaie, and H. R. Tizhoosh, "Heterogeneity-aware local binary patterns for retrieval of histopathology images," *IEEE Access*, vol. 7, pp. 18354–18367, 2019.
- [60] D. Lichtblau and C. Stoean, "Cancer diagnosis through a tandem of classifiers for digitized histopathological slides," *PLoS ONE*, vol. 14, no. 1, Jan. 2019, Art. no. e0209274.
- [61] Y. Gu and Y. Jie, "Densely-connected multi-magnification hashing for histopathological image retrieval," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 4, pp. 1683–1691, Jul. 2019.
- [62] A.-A. Nahid and Y. Kong, "Histopathological breast-image classification using local and frequency domains by convolutional neural network," *Information*, vol. 9, no. 1, p. 19, Jan. 2018.
- [63] Y. Hao, L. Zhang, S. Qiao, Y. Bai, R. Cheng, H. Xue, Y. Hou, W. Zhang, and G. Zhang, "Breast cancer histopathological images classification based on deep semantic features and gray level co-occurrence matrix," *PLoS ONE*, vol. 17, no. 5, May 2022, Art. no. e0267955.
- [64] P. Wang, J. Wang, Y. Li, P. Li, L. Li, and M. Jiang, "Automatic classification of breast cancer histopathological images based on deep feature fusion and enhanced routing," *Biomed. Signal Process. Control*, vol. 65, Mar. 2021, Art. no. 102341.



**AREESHA IJAZ** received the M.S. degree in computer science from COMSATS University Islamabad (CUI), Islamabad, Pakistan. She is a part of the Medical Imaging and Diagnostics MID Laboratory, National Center of Artificial Intelligence (NCAI). Her research interests include medical image processing, computer vision, machine learning, and deep learning.



**BASIT RAZA** received the master's degree in computer science from the University of Central Punjab, Lahore, Pakistan, and the Ph.D. degree in computer science from International Islamic University Islamabad (IIU), Islamabad, Pakistan, in 2014. He conducted his Ph.D. research at the Faculty of Computing, Universiti Teknologi Malaysia (UTM), Malaysia. Currently, he is an Assistant Professor with the Department of Computer Science, COMSATS University Islamabad (CUI), Islamabad. His research interests include data science, database management systems, the IoT, image analysis, security and privacy, data mining, data warehousing, machine learning, deep learning, and artificial intelligence. He authored several papers in refereed journals and has been serving as a Reviewer for prestigious journals, such as the journal of *Knowledge and Information Systems, Network and Computer Applications, Physical Communication, Information Systems, Applied Soft Computing, Swarm and Evolutionary Computation, Swarm Intelligence, Artificial Intelligence Review, and Future Generation Computer Systems*.



**IQRA KIRAN** received the B.Sc. degree from Quaid-i-Azam University, Islamabad, Pakistan, the M.S. degree in computer science from the COMSATS Institute of Information Technology, Islamabad, and the M.C.S. degree in computer science from the National University of Modern Languages, Islamabad. Her research interests include medical image, deep learning, computer vision, and machine learning.



**ABDUL WAHEED** received the master's degree in computer science, in 2014, and the Ph.D. degree in computer science from the Department of Information Technology, Hazara University Mansehra, in 2021. He completed his Ph.D. research at NetLab-INMC under the School of Electrical and Computer Engineering (ECE), Seoul National University (SNU), South Korea, in 2019, under the HEC Research Program. He was an Assistant Professor at the Department of Computer Sciences and the Dean of the Faculty of Engineering and Information Technology (FEIT), Northern University, Nowshera, Khyber Pakhtunkhwa. He is an Assistant Professor of computer science with the Department of Computer Science, Women University Swabi, Khyber Pakhtunkhwa. He has numerous publications in journals and international conferences. His research interests include information security, secure and smart cryptography, heterogeneous communications within the IoT, mobile Adhoc networks (MANETs), wireless sensor networks (WSNs) security, and fuzzy logic-based decision-making theory. He is a member of the Crypto-Net Research Group, Hazara University.



**AADIL RAZA** received the M.Sc. degree in electronics from Quaid-i-Azam University, Islamabad, Pakistan, in 2005, the M.S. degree in nanoelectronics and nanomechanics from the University of Sheffield, U.K., in 2009, and the Ph.D. degree in electrical engineering from COMSATS University Islamabad, Pakistan, in 2018. He was a Research Fellow for one year in the Photonic Research Laboratory, The Hong Kong Polytechnic University (PolyU), Hong Kong. Currently, he is working as an Assistant Professor with the Department of Physics, COMSATS University Islamabad. His research interests include radio over fiber (RoF) systems, free space optical (FSO) communication, distributed acoustic sensor (DAS) technology, machine learning for optical communication, and image processing and advanced digital signal processing (DSP) to realize high speed data center interconnect (DCI).



**HABIB SHAH** received the Ph.D. degree from the Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, in 2013. He is an Assistant Professor and the Head of the Research Unit, College of Computer Science, King Khalid University, Saudi Arabia. He has successfully published more than 50 articles in various international SCI and Scopus journals and conference proceedings. His research interests include artificial intelligence, learning algorithms, data mining techniques, the IoT, and time series analysis and optimization. He is an editorial board member, the guest editor and act as a reviewer for various journals and conferences as well. He has also served as a program committee member and co-organizer for numerous international conferences/workshops. Currently, he is working on three research projects of KKU and KSA.



**SULAIMAN AFTAN** received the master's degree in information security from Lewis University Chicago, IL, USA, in 2015. He is currently pursuing the Ph.D. degree with Texas Tech University, USA. He is a Research Assistant with the Computer Science Department, Texas Tech University. He got many professional certificates, such as ACCES DATA Certified, Security and Mobile Penetration, and Leadership Certifications. He was honored by Gold Key International and President of the Saudi Student Club from Lewis University. He has published more than two articles in various international conference and Scopus journal. His research interests include artificial intelligence, data mining techniques, machine learning, and deep learning. Currently, he is working on more than three research projects at Texas Tech University.

...