IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 5, 2020, accepted January 25, 2021, date of publication February 2, 2021, date of current version February 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3056516

Breast Cancer Classification From Histopathological Images Using Patch-Based Deep Learning Modeling

IRUM HIRRA[1], MUBASHIR AHMAD[1], AYAZ HUSSAIN[2], M. USMAN ASHRAF[2], IFTIKHAR AHMED SAEED[1], SYED FURQAN QADRI[3], AHMED M. ALGHAMDI[4], AND AHMED S. ALFAKEEH[5]

[1]Department of Computer Science and IT, The University of Lahore, Sargodha Campus, Sargodha 40100, Pakistan
[2]Department of Computer Science, University of Management and Technology Sialkot, Sialkot 51310, Pakistan
[3]College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China
[4]Department of Software Engineering, College of Computer Science and Engineering, University of Jeddah, Jeddah 21493, Saudi Arabia
[5]Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding authors: Mubashir Ahmad (mubashir_bit@yahoo.com), M. Usman Ashraf (usman.ashraf@skt.umt.edu.pk), and Syed Furqan Qadri (furqangillani79@gmail.com)

This work was supported by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, under Grant PF-183-42.

ABSTRACT Accurate detection and classification of breast cancer is a critical task in medical imaging due to the complexity of breast tissues. Due to automatic feature extraction ability, deep learning methods have been successfully applied in different areas, especially in the field of medical imaging. In this study, a novel patch-based deep learning method called Pa-DBN-BC is proposed to detect and classify breast cancer on histopathology images using the Deep Belief Network (DBN). Features are extracted through an unsupervised pre-training and supervised fine-tuning phase. The network automatically extracts features from image patches. Logistic regression is used to classify the patches from histopathology images. The features extracted from the patches are fed to the model as input and the model presents the result as a probability matrix as either a positive sample (cancer) or a negative sample (background). The proposed model is trained and tested on the whole slide histopathology image dataset having images from four different data cohorts and achieved an accuracy of 86%. Consequently, the proposed method is better than the traditional ones, as it automatically learns the best possible features and experimental results show that the model outperformed the previously proposed deep learning methods.

INDEX TERMS Deep Learning, deep belief network, histopathology images, classification, breast cancer.

I. INTRODUCTION

According to the American cancer society surveillance report ''Breast cancer affects one in eight women in their lifetime''[1]. After lung cancer, breast cancer stands second cause of death among women and it is the most common type of cancer [2]. According to the study, in the year 2012, the second most common cancer was breast cancer (1.7 million cases, 11.9%) [3]. In the year 2019, 268,000 cases of breast cancer and 41,760 deaths because of breast cancer were estimated by SEER (Surveillance, Epidemiology, and End Results) [4].

Breast cancer can originate from any cell, tissue, or gland of the breast. Sometimes it starts from the ducts that produce milk and sometimes it originates from the lobules which are glandular tissues. If not diagnosed at an early stage there is a possibility of spreading the cancer cells toward different parts of the body and cause damage there too. [5]. Breast tumor has two most common types: benign and malignant where the benign lesion is not cancerous, it is some kind of abnormalities in the cell and they are unable to become a cause of breast cancer and malignant is cancerous lesions. Malignant cells spread at a very fast rate by start divisions swiftly because both cells (benign and malignant) have irregular appearance and structure, it is a very difficult task to manually analyze the microscopic image [6]. Further, breast cancer is divided into two categories, one is ductal carcinoma

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

in situ (DCIS) also known as non-invasive breast cancer and the other is invasive cancer. Breast cancer could have many early indications that one should be suspicious about and approach the specialist for the detection test if found any of them. These indications are mass, micro-calcification, lopsidedness in the breasts, and architectural perversions[7]. These signs could be an indication of breast cancer. To examine the signs many medical imaging methods are there like magnetic resonance imaging (MRI), mammography, ultrasound, and tomography [8]. With the recent advancements in medical science, it is possible to detect and treat cancer at early stages. If the tumor is detected in the early stage, it can be treated and the survival rate increases. In the later stages, the treatment of cancer is very problematic and sometimes even not possible.

Mammography also called a low-dose x-ray of the breast and biopsy are two common techniques used for breast cancer detection. In mammography, radiologists use a specific type of breast images to detect early symptoms of breast cancer. Due to the use of mammography for cancer detection death ratio has decreased [9] and reading mammograms is a very crucial task and demands high expertise. It should provide accurate results otherwise wrong interpretation of the mammogram can cause a problem of false-positive cases [10]. In the second technique called biopsy pathologist analyze a tissue sample taken from an effected breast under the microscope for detection and classification of tumor [11]. Hematoxylin and Eosin (H&E) stain is usually used to stain the tissue samples that are obtained from a biopsy. These stained tissue samples are then examined by the pathologist using light microscopy. Manual examination and evaluation of the tissue images are very critical and tiresome because it involves the evaluation of the fine structure of the cell and the organization of nuclei and their complex characteristics [12]. In the present period, the practice of examining the histopathological images is based on the manual evaluation by the pathologists. By using the manual analysis method there could be many possible problems. Like underdeveloped areas or hospitals they may not have any pathologist. This problem of resource shortage could lead to an increase in cancer cases not examined at an early stage. It is declared by almost 75% of National Health Services (NHS) trusts and 76% of health boards that there is a lack of adequate medical specialists for the treatment of people who are suffering from breast cancer [13]. Now even if there is a pathologist available to evaluate and examine the histopathological images, the validity of the results still depends on the pathologist's capability, diagnostic experience, and keen knowledge of the domain. There is a possibility that results are not accurate always. Last but not the least, the histopathology images are very complex and it is a lengthy process to examine them, also this could lead to cause drowsiness to the pathologists and cause negligence. [14]. If detected at an early stage the breast cancer can be treated and the patient can be completely free from this disease which means an early diagnosis can escalate the survival rate. As discussed earlier many methods were there for the early detection and classification of breast cancer. It has been seen that these methods entirely reliant on the ability of pathologists, other medical specialists and their knowledge. These methods are time-consuming and have a high percentage of the risk factor of the wrong results because specialists have to see the images of the tissue from the affected breast through microscopes or by other manual methods. Because of all these issues, well organized and on target automated methods for detection and classification of breast cancer are required [15].

Automatic diagnosing methods are the necessity of the present. They could be helpful for the radiologists, pathologists, and other specialists by reducing their burden and by providing more accuracy in the results [12]. In the treatment of breast cancer, the main impediment is the lack of methods for early diagnosis. With the advancements in the medical field and keeping in mind the significance of the problem like breast cancer, many experts put their efforts towards the development of smart methods. With time the use of smart systems in the medical field is increasing. The use of these systems will very helpful in the early detection of breast cancer and also play a significant role in reducing the mistakes of professionals caused by the exhaustion or inexperience in diagnosing and classification of breast cancer [16].

Early detection of breast cancer can lead to recovery where the accuracy in the classification of breast cancer is of great significance. Several types of research have been organized on the implementation of machine learning for the detection and classification of breast cancer using different methods or a combination of several algorithms to increase the precision. Machine learning methods can be classified into two categories one is traditional machine learning methods and the other one is deep learning methods. Feature extraction, feature reduction, and classification, these are some of the steps that are involved during the study of histopathological images using machine learning. Considering only the stage of feature extraction here, many well-defined features are extracted from the digital images. These features demand accurate classification execution [17]. It has been seen that the recent focus of the researchers has moved towards deep learning. Deep learning uses a model architecture that is formed of different non-linear variations for the modeling of different essential or required characteristics in the data [18], [19]. In recent years, deep learning methods gained high attainment in different areas like image analysis [20], signal processing [21], speech recognition [22], facial expression recognition [23], brain MRI images analysis [24], [25], [26] and liver segmentation [27], [28], [29]. Deep learning approaches mostly dependent on a large amount of training data. In the field of medical imaging, there is a lack of available training samples. Also, the medical images involve the process of labeling and annotation which is quite an expensive task [30]. Addressing this problem of lack of sufficient data, many researchers advised many different techniques. According to different researchers, it is advised to not use the whole images as input but instead use 2D patches or 3D cubes. This process of using 2D patches and 3D cubes

also helps in the reduction of the model parameters and also reduces the problem of overfitting [31], [32]. Some researchers also suggest using data augmentation where there are plenty of techniques available that can be used for data augmentation such as flipping and rotation [33], [34]. After the augmentation of data, one could achieve a significant amount of data that is required for the classification task to train the model [35], [36]. Deep learning algorithms gained remarkable significance. Their use expands in different areas and the main factor in the increasing popularity of deep learning algorithms is their ability to feature extraction [37].

This study presents a novel DBN based classification model. Restricted Boltzmann Machine (RBM) is a generative artificial neural network (ANN) and on a specified input which used to acquire the probability distributions. DBN is a deep neural network that comprises multiple layers of hidden units. There is a connection between the layers but the units in the layers have no connection [38]. To facilitate the specialist's many kinds of researches have been done using traditional machine learning techniques, specialists do not have to scrutinize the microscopic images manually, there is still manual work involved. To train the model, there is still a need to manually feed the features. To overcome this issue and making this process completely automatic deep learning has been introduced. DBN has been used in many types of researches in medical image processing but used very rarely in breast cancer classification. In this study, application of deep learning has been exploit using the DBN model for the classification of breast cancer using histopathology images.

Our main contributions are as follows: This study propose a novel framework for the classification of breast cancer on histopathology images. In our knowledge this is the first time DBN based method is used in breast cancer classification. The proposed framework provides the automatic and accurate representation of features from images with the help of DBN. Patch based DBN technique not required much hardware resources which is better then other deep learning methods. The used dataset comprises data from four different data cohorts which gives the model an exceptional key because different data cohorts used different histopathology methods for staining the images and different specifications to digitize the images. The presented experiments show that the results are better than other recent deep learning-based techniques.

II. RELATED WORK

The traditional machine learning methods include sparse representation, support vector machine (SVM), and K-nearest neighbor (KNN) algorithms [39]. Authors [40] proposed a model using SVM for the detection of breast cancer where two types of SVM models have been used which are C-SVM and V-SVM. To hybridize the model a weighted area under the receiver operating characteristics curve ensemble (WAUCE) method is used. The model performed better than the other proposed studies using a single SVM structure. One limitation of this study is that it requires a large amount of time and cost. This method is quite expensive. This study [41] proposed a Bayesian networks to classify the mammographic images with each view and then used logistic regression for the final decisions on the results generated by Bayesian networks. CAD systems are bound to analyze each mammographic view separately while the radiologists have to analyze two views at a time to find the difference by comparison. An expert system for breast cancer detection (Ex-DBC) is proposed. The neuro-fuzzy method is used to find fuzzy rules. The proposed system predicts high positive values that help to avoid the wrong interpreted results that lead to biopsies. An expert system or a rule-based system is very beneficial because it contains specific and unambiguous information. The rules can be altered and updated easily whenever the need arises. However, on statistical properties, the use of rules-based has the strain and adversity that there is a possibility of deriving the similar statistical properties for some patterns of different classes, due to which there may occur a problem of incorrect recognition [42].

Mammographic images are typically acquired from the same breast with two views: a lateral view called mediolateral oblique (MLO) view and a top head-to-toe view called craniocaudal (CC) view [43]. Several pieces of research have been made to extract the features from the mammographic images with different views. In the proposed method a geometry-based region matching method is used to extract the breast masses from the mammographic images with the two views CC, MLO and then use the multi-view classifier to classify the breast masses. [44]. A CAD-based technique [45] is proposed for breast cancer classification using mammographic images. For feature selection and dimensionality reduction, a genetic algorithm is used where SVM is used for classsification. The experiments achieved better accuracy as compared to previous approaches. [46] proposed a MAD normalization based technique for breast cancer classification where the model comprises three phases. In the first phase, the MAD normalization method is used to normalize the dataset. The second step involves the weighting procedure of normalized data using K-mean clustering (KMC). The last phase is the classification phase which is done by using the AdaBoostM1 classifier. The proposed technique provides improved accuracy as compared to previously proposed methods but the limitation of this work is the cost. The proposed technique for breast cancer [47] in which the classification is done using a multi-layer perceptron (MLP) neural network. A non-dominated sorting genetic algorithm is used for the network assembly but local minimum is a problem for MLP.

The proposed model to extract the features from the mammographic images with two views MLO and CC, ANNs are used and then fused the features of resultant mammographic images from ANNs for classification of two types of cells that are benign and malignant. The proposed model can work for different views but it cannot extract features with a well-built particular performance. The reason for this is the loss of spatial information caused by the feature transformation from image to a long dimensional vector [48]. A model is proposed for the classification of benign and malignant using logistic

regression, KNN, and Ensemble learning principal component analysis (PCA). Features are extracted using PCA and transferred to the retraining phase. The model uses a small amount of data and achieves better accuracy and the accuracy could become better by using a large amount of data [49]. [50] presented a knowledge-based model to classify the cancer. In the proposed study clustering and expectation maximization is used. For the generation of fuzzy rules classification and regression trees are used wherever the results show that this model provides a promising accuracy in the classification task. This model is a hybrid intelligent system and a non-incremental.

Today deep learning became very popular in a vast variety of fields surpassing other conventional methods. Deep learning methods have shown tremendous performances in different fields including the field of medical imaging [51] like vertebrae segmentation [52], [53]. With the help of image processing techniques, it became easier to detect and classify tumors from an infected breast [54]. Authors in [55] proposed a deep learning-based model for the classification of breast cancer using CNN. The model creates patches from the images and passes them to CNN for training. This model provides simple architecture and the cost reduces due to the simplicity of architecture. But the limitation of this study is the lack of experimental data because of which the cost of the experiment increases.

The patch level classifier is very beneficial for the classification of tumor cells. A model is created using 2 stage training procedure that is a patch level network to learn from pixel-level labels along with a network learning from macroscopic breast level labels. The proposed framework contains two core modules. One module is 4-view specific columns, each based on ResNet architecture that gives a fixed dimension hidden representation for each mammography view. It uses ResNet22 columns to compute a 256 dimension hidden representation vector of each view. The second module is two fully connected layers to map from the computed hidden representation to the output prediction. Using separate pixel and breast level models differentiate their work from other approaches that use pixel-level labels in a single network [56]. The authors worked on a deep learning-based model to classify x-ray scattering images [57]. This classification is a challenging task because it is based on the attributes which can be different characteristics ranging from physical properties to the different types of angles measurement. This work is done in two stages where the first stage encounters a convolutional neural network (CNN) using 50 layers Residual Network and the second stage consists of a patch-based Convolutional autoencoder where the patch selection is random. For the classification work, SVM based classifier is used. The results show that the deep learning-based model gave far better results than the old handcrafted feature extraction methods.

The authors presented a novel technique [57] of pixel pair based on CNN called (CNN-PPF) and find the similarity between the pixels. The construction of the pixel pair is done by joining the center pixel with the pixels around and then fed to the trained multilayer CNN model to learn the features. The CNN model include ten convolutional layers, three max-pooling layers, and ReLU layers. As deep learning methods need a huge amount of data for better results while this proposed study shows that by reorganizing the data in hand we can overcome this problem. For the prediction of the pixel class, a voting strategy is used. The results of the proposed CNN-PFF model show that it can perform better than the previously proposed methods and also it is cost-effective because it used a small amount of data.

The classification of benign and malignant tumor in lung nodules framework [58] based on three types of the deep neural network is proposed which are CNN, deep neural network (DNN), and stacked autoencoder (SAE). Each layer of CNN contains multiple maps, while each map is a combination of different neural units with one convolutional kernel each. In the DNN all the input layers, hidden layers, and output layers are connected. The SAE consist of the fully connected layers and the neurons in both the input and output layers are equal. The results show that CNN outputs the best results as compared to DNN and SAE because there is limited data available that's why the size of the layer of the neural network is kept small.

Batik's pattern classification based on CNN [59] which is composed of a VGG16 for the extraction of features and used a multi-layer perceptron for the classification task. In the first step, six classifiers are trained using different features, then the best classifier is selected and trained on random images with no transformation while the last step, classifiers are tested using the transformed images. Experiments show that the VGG16 with the combination of MLP performed better on the transformed dataset as compared to other models. As the VGG16 is pre-trained so, it can be seen from the results that this model can handle transformation invariant features. More accuracy could be achieved by the better labeling of the input images and by using high-quality data only [60] presents the first deep neural network-based framework to classify the temporal sleep stage classification from multimodal and multivariate time series without incorporating handcrafted features. By using backpropagation linear spatial filtering the model gathers the data from multiple sensors and use of liner channels which boosts the input data. By using temporal convolutions it extracts the features and at the end for the classification softmax is used. The model shows the best results as compared to previous studies with a minimum computational cost. For the classification of lung cancer [61] DCNN inception v3 is used to classify between LUSC, LUAD, and normal lung tissue. The first step followed in the model is to train the inception v3 by randomly initializing the weights of the last layer then a per slide classification method is introduced to see the probabilities. The model results show that using CNN can explicitly classify between tumor types and provide better results than the old methods.

Proposed a model for the classification of Thai fast food based on deep learning [62]. The model used the GoogleNet

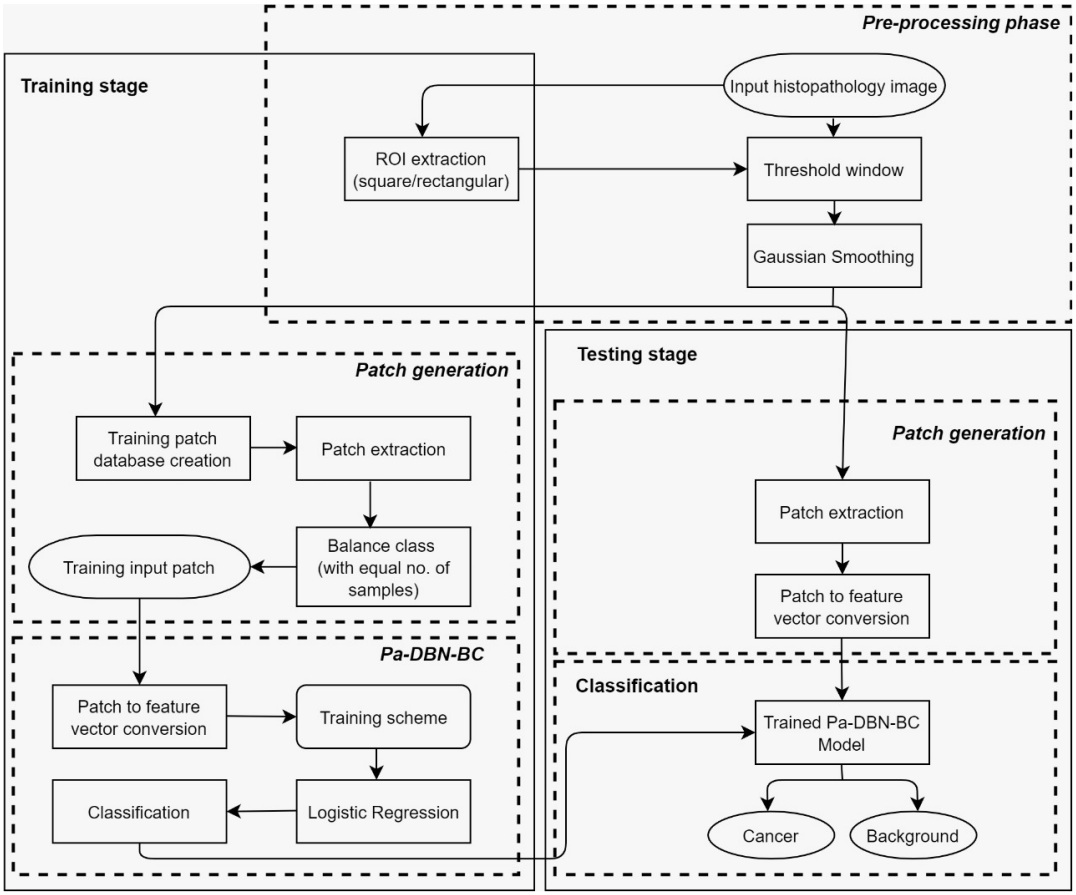

FIGURE 1. Structure of proposed Pa-DBN-BC model.

model with some convolutional, pooling layers and inception layers. The model work by the distribution of data to the preceding layers while they have different sizes and make a thin assembly of data together which minimizes the computational cost. Deep CNN based model is used in the classification of skin cancer. As the pre-trained Alexnet cannot work on a large size colored image so the extraction of the region of interest (ROI) is done. The proposed method has two plus points over conventional methods that there is no need for preprocessing and it also can use any kind of input image. The results are better compared to the earlier proposed models [63]. A model based on a deep neural network is introduced [64] for the classification of four types of brain tumors while the images are segmented using Fuzzy C-means. A discrete wavelet transform technique is used as a feature extractor and PCA is used to reduce the features. The classification is performed using a seven-layer deep neural network. This model needs quite minimal hardware resources as compared to other CNN architectures.

DBN is used in an unsupervised fashion for the initialization of the weights then fed the output matrix to the backpropagation neural network [65] The first stage is unsupervised while the second stage using a backpropagation neural network in a supervised manner. Two algorithms conjugate gradient and Levenberg Marquardt used in backpropagation neural network. In this study, the combination of supervised backpropagation neural network with the unsupervised DBN shows better results than the previous studies that used only one supervised phase. A DBN based CAD system is proposed in the study to diagnose breast cancer. For the ROI extraction, two methods are used: one is multiple mass ROI and whole mass ROI. The DBN based approach shows better accuracy in classification as compared to previous techniques [66]. A cascade deep learning-based approach is used in [67] which classify the breast cancer. A pre-trained deep learning classifier is used for classification where the model achieves high classification sensitivity.

III. METHODOLOGY

With the development of new methods, the study on deep learning received substantial attention because of its better results in different fields [68]. Like other deep learning models our model also consists of two stages: training and testing. In this study, a Pa-DBN-BC model is proposed for the classification of breast cancer on the histopathology images. The Pa-DBN-BC model comprises four main phases which are the preprocessing, patch generation, DBN, and classification phase where the model works on equal size patches of images. In preprocessing we can crop our input data around the region of interest (ROI) without worrying about the limitation of

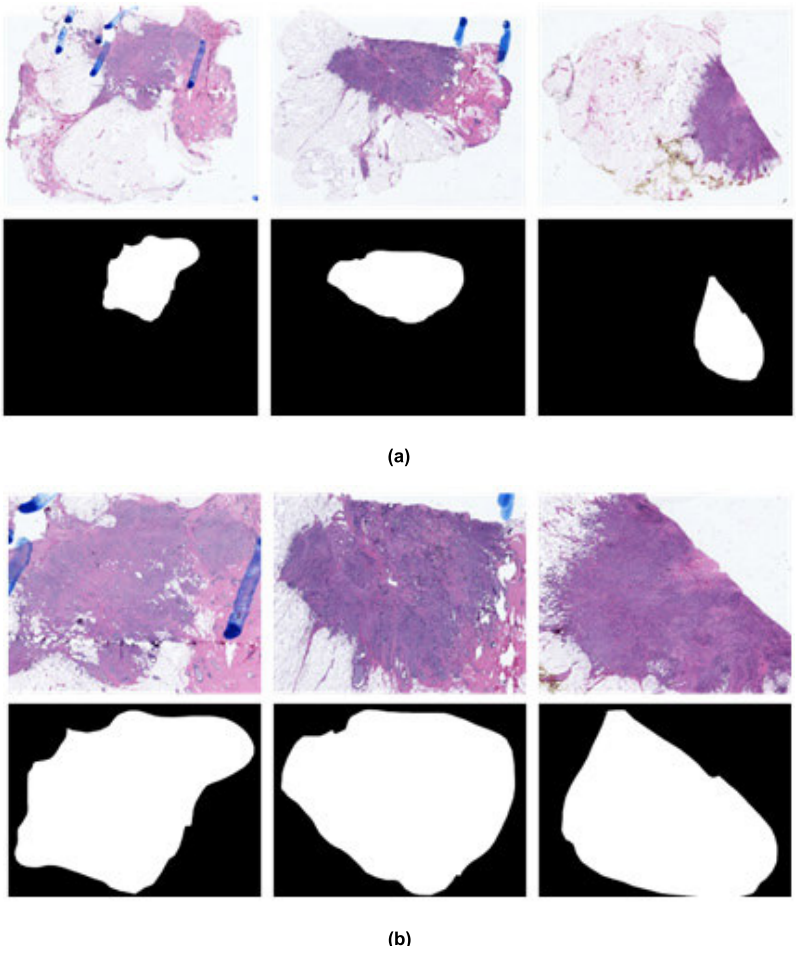

FIGURE 2. Sample images of breast cancer from the dataset used in the experiment: (a) Sample images from dataset and corresponding masks(b) Sample images from dataset and corresponding maks after cropping.

keeping the images of equal sizes. The proposed model work in an unsupervised fashion for the extraction of features from the input histopathology image patches in the form of feature vectors. The extracted features matrix is then transferred to the backpropagation neural network which is a supervised learning phase and it comprises of the conjugate gradient. A model is formed by the feature matrix of images in the training phase and the final stage is the classification stage which discriminate between cancerous and non-cancerous regions.

A. PREPROCESSING

The image preprocessing is a critical step because it can improve the input data by enhancing some features that could be important to a particular task. In this phase, each image size is reduced by cropping the images to remove extra backgrounds and focusing on the ROI. The images and the corresponding labels around the ROI are cropped randomly regardless of the sizes of different images. Original images of different pixel values with the provided labels before cropping and the images with corresponding labels after cropping are shown in Figure 2.

The next step in the preprocessing is to convert the RGB colored images from RGB color space to the grayscale. The conversion of images from RGB color space to grayscale is shown in Figure 3.

In the preprocessing phase, a threshold window is applied to the dataset. In image processing, this step is considered a basic step but having a great significance. Different tissues seem having different intensity, the threshold value can help in eliminating the non-essential segments [69]. The threshold is applied after coversion of RGB to Grayscale by manually analyzing the frequency values of both cancerous regions and background regions, the values ranges from 40 to 215 in the cancerous regions. This helps in the accurate differentiation of both regions before patch creation. Contrary to this, if the selected range is outsized, then the noise could not be turned down [70].

The smoothness of images is controlled by the use of a Gaussian filter with a fixed kernel size on the histopathology images. Gaussian filter plays an important role in achieving classification accuracy and reducing the weight of bluring pixels [71]. In this way, our model will able to learn to distinguish between a positive sample and a negative sample more

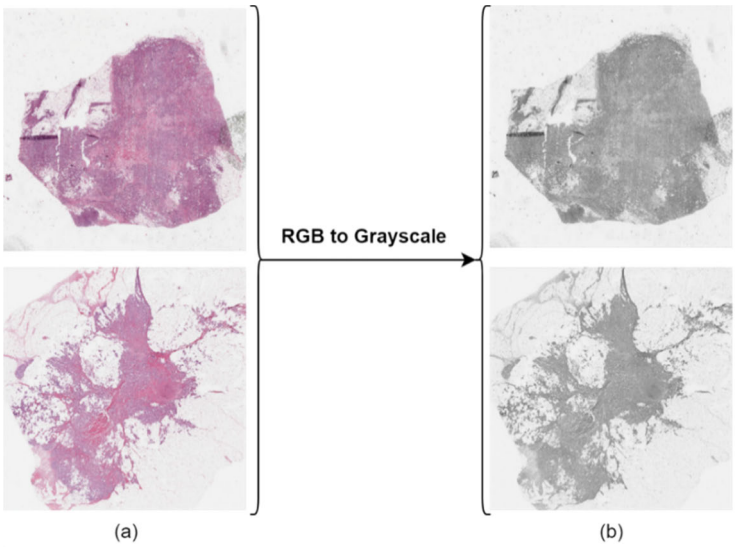

FIGURE 3. Conversion from RGB to Grayscale.

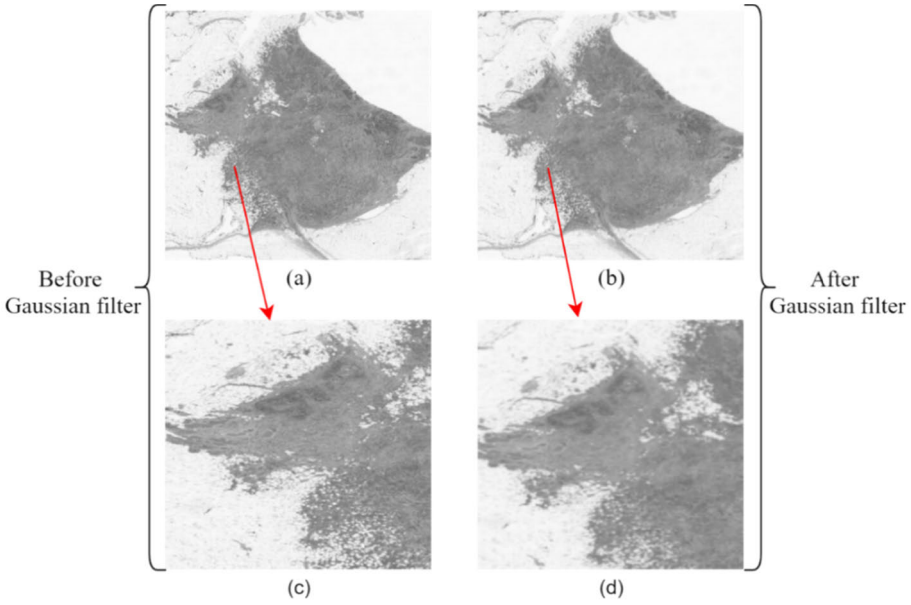

FIGURE 4. Application of Gaussian filter.

accurately. The selected value for Gaussian filter was 1.0. Figure 4 demonstrates the application of the Gaussian filter by showing a sample image before and after the application of the Gaussian filter. Same preprocessing steps are applied on both traing and testing samples. The extra step in training phase is the balanced classes.

B. PATCH BASED DATABASE CREATION

In this step, input is creared for our proposed network by creating the overlapping patches of size 32×32. All the patches having the same size with two properties either cancer region (positive samples) or background region (negative sample). A patch is considered to be cancer patch if 70 percent pixels of the patch related to abnormal region. The dataset became a binary class dataset as shown in Figure 5.

After the creation of overlapping patches with same size, the next step is to randomly shuffle the patch data. The number of positive samples is kept equal to the number of negative samples with ratio 1:1.

C. TRAINING THE PA-DBN-BC MODEL

After the creation of overlapped patches of size 32×32, now these patches are converted into the vector of size 1024 which is given as an input to PA-DBN-BC model. We choose DBN for feature extraction due to the remarkable performance of DBN in different fields like image processing, [72] and natural language processing (NLP) [73]. DBN is formed by stacking the layers of RBM and initialize the network weights by using a greedy learning approach. For the input in the training and testing phase, we use feature vectors obtained

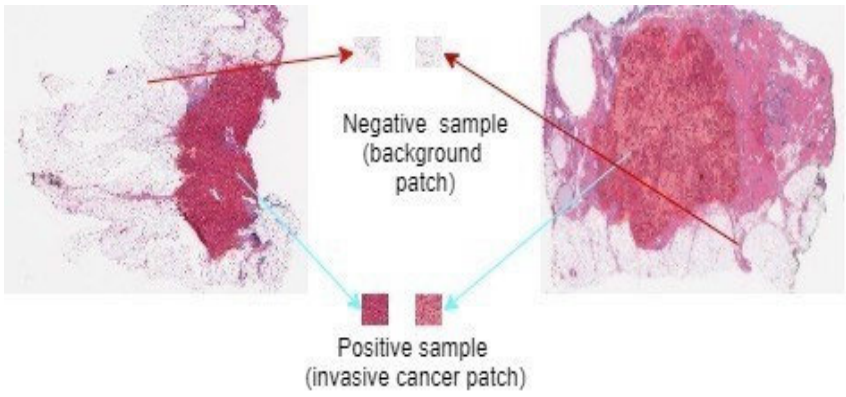

FIGURE 5. Positive and negative sample from WSI.

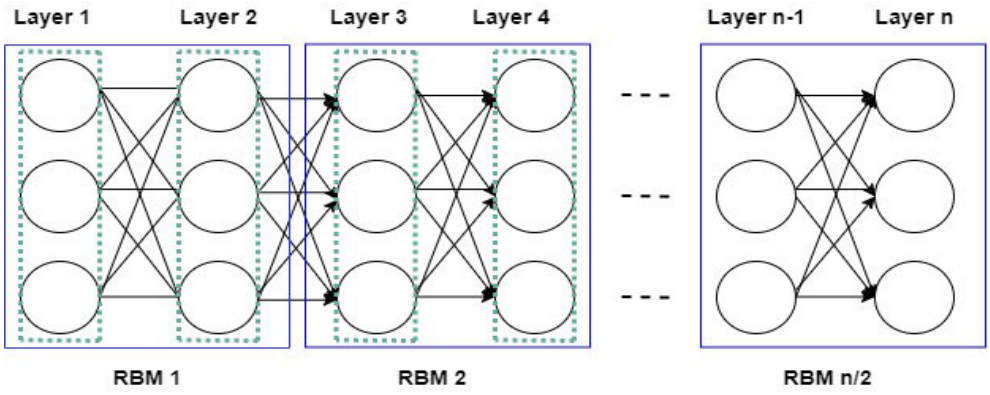

FIGURE 6. Simple DBN architecture.

from overlapping patches. Features are extracted in an unsupervised fashion. The two hidden layers are used to extract the features from histopathology images. The first RBM is formed by combining the input layer and the first hidden layer and second RBM is formed by the combination of hidden layer 1 and hidden layer 2. A simple DBN architecture is given in Figure 6.

Model is trained based on features F from patches P. In the DBN model, multiple RBMs are arranged in such a fashion that the output of one layer worked as the input of the next layer. This way the training complexity is reduced. This procedure can be represented as follows:

$$\sum_{ij}^{l} \tag{1}$$

Here i stands for the neurons present in the layer l that are connected to the neurons j of the next layer which is $l + 1$. DBN is originated from Boltzmann machines (BM), there are two types of connections permissible between neurons of each layer which are inter and intra, but in RBM only inter layer neurons connection is possible. In RBM the connection of the neurons can be formed in any fashion either directional or non-directional. So the training process in RBM is simple and fast. [74]. To initialize the parameters of DBN initially an RBM is used in an unsupervised fashion. This benefits the model in achieving better results as compared to the shallow models [75], [76]. RBM initializes the weights W and biases b of the network where more tuning of parameters was done to further improve the network. Fine-tuning is the second stage of DBN learning. In this stage, the class labels are given to the model which was not used in the pre-training. The aim of using deep learning is making such models that are generalized that can deal with new samples accurately.

For fine-tuning, NN is trained by using a labeled dataset and treated it as a classification model. After the supervised training, DBN is converted into NN by adding an output layer having two neurons for classification on the top. For the classification purpose logistic regression is considered as a proficient way. As our problem is a binary classification problem, sigmoid function is selected for logistic regression. As in binary classification problem, we always need a output with a probability between 0 and 1. The sigmoid function helps in transforming the output of logistic regression into the probability values. The sigmoid function can be represented as follows.

$$Sig\,(z) = \frac{1}{1 + e^{-z}} \tag{2}$$

where e represents the base of the natural log and z represents the output of the model trained with logistic regression which can be represented as:

$$z = w_1 x_1 + w_1 x_1 + \ldots\ldots\ldots\ldots\ldots\ldots + w_N x_N + b \tag{3}$$

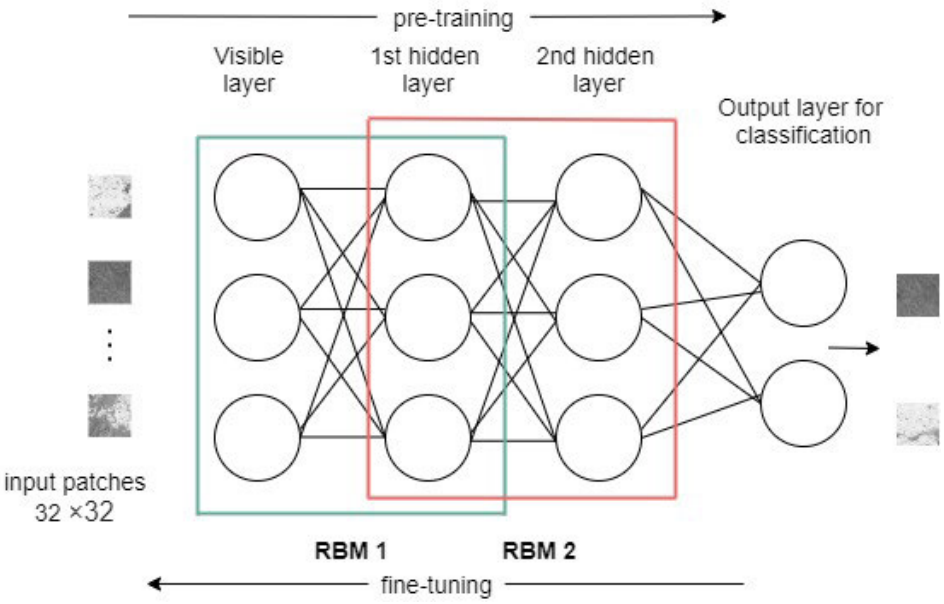

FIGURE 7. Structure of the proposed model.

The above equation can also represented as follows:

$$\sum_{m=1}^{N} w_m x_m + b \tag{4}$$

where w represents the model's learned weights, x represents the features and b represents bias. The structure of a PA-DBN-BC model is given in figure 7.

D. TESTING SCHEME

For the testing stage, same preprocessing steps are applied on the testing data which were used in training. Image patches are created in the same fashion as in the training and these patches are then passed to the trained model. The patches from the histopathology images used in the testing stage are completely unseen for the model. The model classified the patches as either a positive sample (cancer) or as a negative sample (background).

IV. RESULTS AND DISCUSSION

A. DATASET SELECTION

In this study, a publicly available dataset based on histopathology images is used. The dataset includes histopathology images from the four different data cohorts, Hospital of the University of Pennsylvania (HUP), Case Western Reserve University (CWRU), Cancer Institute of New Jersey (CINJ), and The Cancer Genome Atlas (TCGA) and their corresponding binary masks of invasive breast cancer regions annotated by pathologists [77]. Dataset is available on the following link https://doi.org/10.5061/dryad.1g2nt41. Images are randomly selected from all four data cohorts for training, validation, and testing stages. Each image in the dataset having the size of 3002 × 2384.The HUP 239 images, CINJ 40 images and TCGS 195 images are digitized with Aperio scanners, while the UHCMC/CWRU 110 images are digitized with Ventana scanners. All these images are manually annotated by pathologists and delineated the cancerous regions at 2x magnification using Image scope v11.2 from Aperio and Image Viewer v3.1.4 from Ventanna. There are not such variations in the scanning that can cause potential issues to this study.

In this research the RGB images are converted into grayscale representation. The reason is simplicity. Complexed procedures can be performed in less time, while RGB images are hard to work with in many scenarios. These consist of 3 color channels and it will require more processing. We used 177 images in grayscale representation where 177 RGB images requires more computation time. In our proposed research only one color channel was enough to get the required results. The dimension of the image is very large that's why we distributed each image into equal size patches. The proposed work is implemented using Windows 10 operating system on Dell Intel ®core ™i7-8565U CPU @ 1.80 GHz and 32 GB RAM using MATLAB 2018b.

B. PARAMETRIC SETTINGS

After the preprocessing step biases and weights are initially set to zero. On several dataset there are lots of algorithms shows that random search experimental methods are more effective for hyper-parameters optimizations. In our Pa-DBN-BC model we ran our model on random search experiments to identify the best range which is stable in training and validation where figure 8-a shows the stable training and validation. It is found that random search experiments save the time to get the best parametric settings. Because all parameters are not important for tuning so random experimentation is most important based on hyper parametric

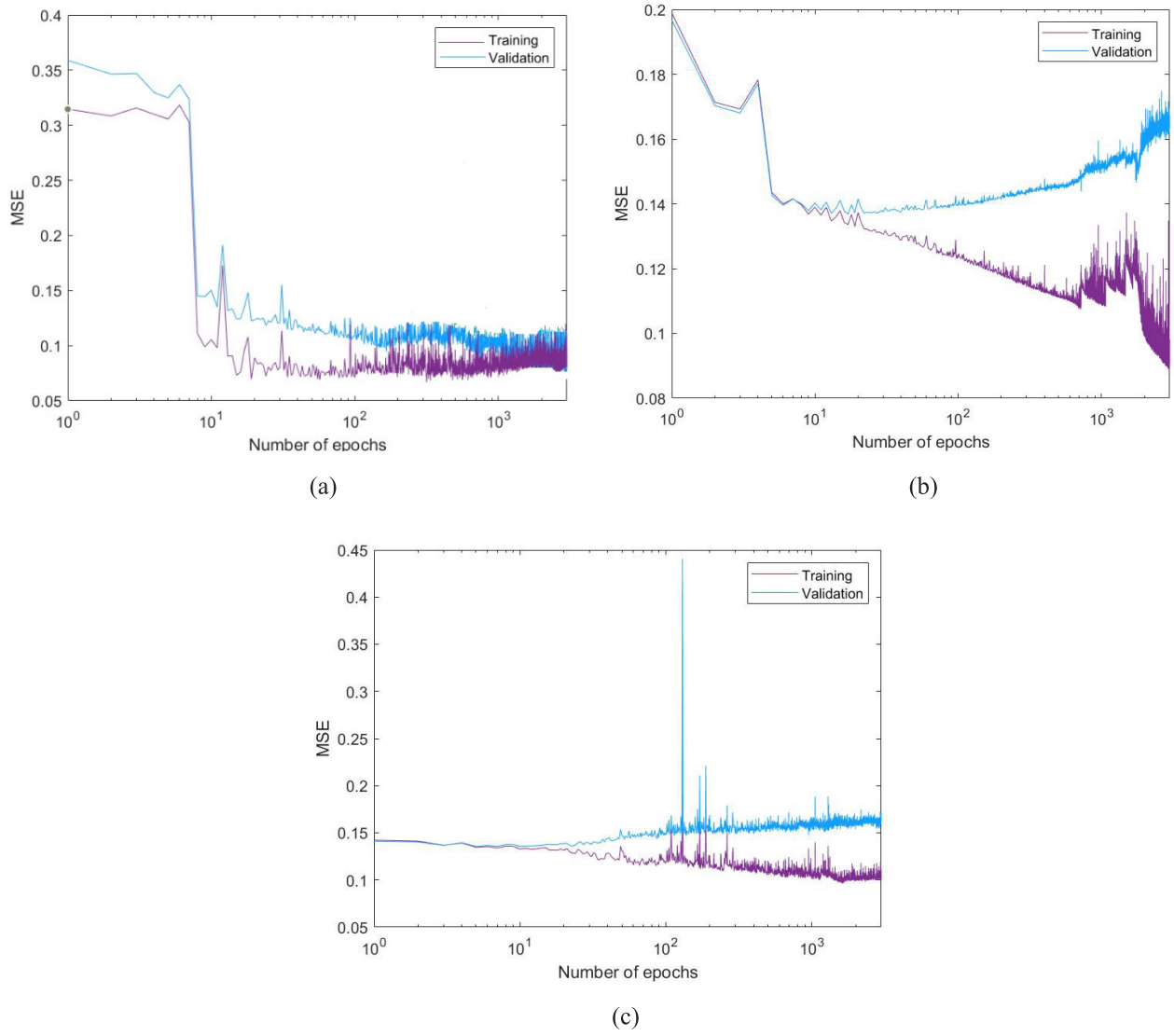

FIGURE 8. The learning curve of the proposed model (a) is the best performance on selected parameters (b) and (c) is the worst performance with different parameters.

values. During the pretraining stage, learning rate was set to 0.0001 and 0.9 was the momentum with 30 epochs for RBM training. Contrastive divergence (CD) algorithm was used to train each RBM in agreedy manner. 32 × 32 patch size was flatterned to get the 1024 vector size which is given as an input to the RBM1 in our model. The RBM2 was trained on the output of first RBM. Stochastic gradient descent (SGD) algorithm was used to train the model. Feed forward network was initialized to train the network of 100-100 was used as a hidden layers and 2 neuron were used for output of our model. Total of 177 images were used for patches dataset creation in our experiment where patch dataset is separated with a ratio of 70:15:15 for training, validation and testing respectively. In the training phase, a matrix with 519936 × 1024 dimensions is used and for validation, a matrix of 111360 × 1024 is used as input and the mini-batch size was 128. We setup the learning rate of 0.0001, momentum with 0.9 and the weight decay of 0.00001 using backpropagation neural network. To reduce the chances of overfittimg, we used a weight decay. Mini batch size in the fine tuning was used 128 where segmoid was used as an activation function. For fine tuning we trained our model on 3000 epoches. For fine tuning our model took 35 hours. Overall time was 39 hours including pre training. The parametric setting is given in Table 1.

Mean squared error (MSE) [76] was used to measure the error during the back propagation neural network training. The training error was recorded 0.09 and validation error was 0.10 on 3000 epoches wherever no convergence was recorded after these iterations. We stop the experiment here which was the best error rate at this point. We examined in our experiment with low learning rate yielded best results in figure 8-a as compared to 8-b and 8-c which have a big

TABLE 1. Parameter settings.

Hidden layers	2
Number of neurons in each layer	100
Number of epochs	30 for pre-training, 3000 for fine-tuning
Learning rate	0.0001
Momentum	0.9
Mini Batch Size	128

learning rate which cause a high variation between training and validation.

C. PERFORMANCE EVALUATION

In this section performance evaluation measures are defined that are used in this study,

ACCURACY**:** The proportion of the total number of correct predictions. It contains the proportion of both predictions of negative samples and positive samples. It can be calculated as:

$$\frac{TP + TN}{Total} \tag{5}$$

ERROR RATE: The proportion of the total number of prediction cases that were not correct. It contains the proportion of both predictions of negative samples and positive samples. It can be calculated as:

$$\frac{FP + FN}{Total} \tag{6}$$

SENSITIVITY OR RECALL: The proportion of actual positive cases that are correctly identified also known as the true positive rate. It can be calculated as:

$$\frac{TP}{ActualPositive} \tag{7}$$

FALSE POSITIVE RATE: The proportion of the total number of predictions that were negative cases but predicted as positive cases. It can be calculated as:

$$\frac{FP}{ActualNegative} \tag{8}$$

SPECIFICITY: The proportion of actual negative cases that are correctly identified also known as the true negative rate. It can be calculated as:

$$\frac{TN}{ActualNegative} \tag{9}$$

true positive (TP), which represents the cancer patches that are correctly identified; false positive (FP), represents the patches which are not the cancerous part but model identifies them as cancer regions; true negative (TN), represents the patches that are of non-cancer region and model correctly identifies them as a non-cancerious patch; and false negative (FN) represents the patches that are of cancer region but identified as the non-cancer region. The results of our model are presented in Table 2.

TABLE 2. Performance evaluation parameters.

Accuracy	**86%**
Error Rate	14%
Sensitivity	87.9%
FP Rate	15.9%
Specificity	84%

The Table shows that our proposed model achieved an accuracy of 86%.

D. COMPARISON WITH OTHER STUDIES

In this section the results of the proposed method are compared with other state of the art methods and the results show that our model performed better which are presented in Table 3.

We compared the results of the proposed method with other state of the art methods and the comparison shows that our model provides promising results. In the proposed study [78], three types of deep CNN architectures, 6-layer, 13-layer, and 17-layer are used for the detection of breast cancer mitosis using histopathology images. The dataset that is used in the study is mitos atypia. The accuracy achieved by the proposed method is 84.49% using 17-layer architecture. The CNN model is quite deep as it comprises 17 layers but still, the architecture failed to provide promising results. The error rate is 15.50% in this method as compared to our technique which is noted 14% that is much better. Comparing sensitivity with CNN model which was 80.55% that is much lower then Pa-DBN-BC model which is noted 87.9%. Authors [79] presented a deep learning-based method for the classification of breast cancer from WSI. The classification is based on learning a hierarchal part-based representation. Total 162 images were used, 84 images were used for training, 29 for validation and 49 for testing where the work is based on patches. Total 82,883 patches were used for training, 31,352 for validation and 50,963 patches were used for testing. The study shows an accuracy of 84.23%. Sensitivity is noted 87.9% as compared to the said technique which is 79.60%. Authors in [80], presented a CNN based model for classification of breast cancer using H&E stained images. Two types of classification are performed in the study, one is two class and the other

TABLE 3. Results comparison of state of the art methods with proposed Pa-DBN-BC method.

Techniques	Performance				
	Accuracy	Error rate	Sensitivity	False-positive rate	Specificity
[78] CNN (Zainudin et al., 2020)	84.49%	15.50%	80.55%	11.66%	--
[79] CNN (Cruz-Roa et al., 2014)	84.23%	--	79.60%	--	88.86%
[80] CNN (Araujo et al., 2017)	4-class=77.8% and 83.3% for 2 class	--	--	--	--
(Our) Pa-DBN-BC	**86%**	**14%**	**87.9%**	**15.9%**	**84%**

is four-class classification. Total 249 images were used for training and 36 images for testing. Total 70,000 patches were used for training and validation using ratio 75:25 from testing dataset. Both classifications achieved an accuracy of 83.3% and 77.8% respectively which is below the accuracy achieved by our proposed model.

V. CONCLUSION AND FUTURE WORK

In this study, a Pa-DBN-BC model is presented for the classification of breast cancer using histopathology images. The proposed model learns the features automatically by creating equal size patches of images. Pre-training is achieved in an unsupervised fashion and fine-tuning in a supervised fashion. After learning the features, the patch matching model is used to create a probability estimation matrix. The results show that the deep learning method improved classification accuracy in breast cancer cases. According to our results, it is observed that the patch-based model performed better than the models that work using whole images for feature extraction. This will reduce the computational cost and also provides high accuracy in the binary classification problem.

For future work, our model could perform better and could provide better accuracy if more hardware resources are available like GPU, for using a large number of patches as input. Also, this study is a binary classification study, as in this work we only classify between the cancer regions from the background regions. In future, we can work on the classification between different types of cancers using this model

REFERENCES

[1] D. A. Shumway, A. Sabolch, and R. Jagsi, "Breast cancer," *Med. Radiol.*, vol. 15, pp. 1–43, Dec. 2020, doi: 10.1007/174_2016_83.

[2] S. Ronoud and S. Asadi, "An evolutionary deep belief network extreme learning-based for breast cancer diagnosis," *Soft Comput.*, vol. 23, no. 24, pp. 13139–13159, Dec. 2019, doi: 10.1007/s00500-019-03856-0.

[3] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman, and F. Bray, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer*, vol. 136, no. 5, pp. E359–E386, Mar. 2015, doi: 10.1002/ijc.29210.

[4] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA, A Cancer J. Clin.*, vol. 69, no. 1, pp. 7–34, Jan. 2019, doi: 10.3322/caac.21551.

[5] R. John, *Global Cancer Facts & Figures-Special Section, the Obesity Epidemic*, vol. 76, 4th ed. Atlanta, GA, USA: American Cancer Society, 2018.

[6] I. Pollanen, B. Braithwaite, T. Ikonen, H. Niska, K. Haataja, P. Toivanen, and T. Tolonen, "Computer-aided breast cancer histopathological diagnosis: Comparative analysis of three DTOCS-based features: SW-DTOCS, SW-WDTOCS and SW-3-4-DTOCS," in *Proc. 4th Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Oct. 2014, p. 5, doi: 10.1109/IPTA.2014.7001986.

[7] M. M. Koo, C. von Wagner, G. A. Abel, S. McPhail, G. P. Rubin, and G. Lyratzopoulos, "Typical and atypical presenting symptoms of breast cancer and their associations with diagnostic intervals: Evidence from a national audit of cancer diagnosis," *Cancer Epidemiol.*, vol. 48, pp. 140–146, Jun. 2017, doi: 10.1016/j.canep.2017.04.010.

[8] A. B. Nover, S. Jagtap, W. Anjum, H. Yegingil, W. Y. Shih, W.-H. Shih, and A. D. Brooks, "Modern breast cancer detection: A technological review," *Int. J. Biomed. Imag.*, vol. 2009, Jan. 2009, Art. no. 902326, doi: 10.1155/2009/902326.

[9] A. Chetlen, J. Mack, and T. Chan, "Breast cancer screening controversies: Who, when, why, and how?" *Clin. Imag.*, vol. 40, no. 2, pp. 279–282, Mar. 2016, doi: 10.1016/j.clinimag.2015.05.017.

[10] Y. Ali and S. Hamed, "Early breast cancer detection using mammogram images: A review of image processing techniques," *Biosci. Biotechnol. Res. Asia*, vol. 12, no. SEMAR, pp. 225–234, Mar. 2015.

[11] A. Chekkoury, P. Khurd, and J. Ni, "Automated malignancy detection in breast histopathological images," *Med. Imag. Comput. Diagnosis*, vol. 8315, Feb. 2012, Art. no. 831515, doi: 10.1117/12.911643.

[12] S. Vesal, N. R. B, and A. Davari, *Classification of Breast Cancer Histology*, vol. 1. Cham, Switzerland: Springer, 2018.

[13] (2017). *Breast Cancer Care*. [Online]. Available: https://breastcancernow.org/about-us/media/press-releases/three-quarters-nhs-trusts-health-boards-say-not-enough-care

[14] 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, doi: 10.1016/j.ymeth.2019.06.014.

[15] T. Gaber, G. Ismail, A. Anter, M. Soliman, M. Ali, N. Semary, A. Hassanien, and V. Snasel, "Thermogram breast cancer prediction approach based on neutrosophic sets and fuzzy C-means algorithm," in *Proc. 37th Annu. Int. Conf. Eng. Med. Biol. Soc.*, Aug. 2015, pp. 4254–4257.

[16] B. Milovic, "Prediction and decision making in health care using data mining," *Int. J. Public Health Sci.*, vol. 1, no. 2, pp. 69–78, Dec. 2012.

[17] D. Komura and S. Ishikawa, "Machine learning methods for histopathological image analysis," *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 34–42, Jan. 2018, doi: 10.1016/j.csbj.2018.01.001.

[18] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

[19] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015, doi: 10.1016/j.neunet.2014.09.003.

[20] K. He and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, May 2016, pp. 770–778.

[21] O. Yildirim, U. B. Baloglu, R.-S. Tan, E. J. Ciaccio, and U. R. Acharya, "A new approach for arrhythmia classification using deep coded features and LSTM networks," *Comput. Methods Programs Biomed.*, vol. 176, pp. 121–133, Jul. 2019, doi: 10.1016/j.cmpb.2019.05.004.

[22] Y. Aytar, C. Vondrick, and A. Torralba, "SoundNet?: Learning sound representations from unlabeled video soundnet architecture," 2016, *arXiv:1610.09001*. [Online]. Available: https://arxiv.org/abs/1610.09001

[23] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in *Proc. Int. Conf. Comput. Vis.*, 2015, pp. 3730–3738.

[24] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1240–1251, May 2016.

[25] T. Brosch, L. Y. W. Tang, Y. Yoo, D. K. B. Li, A. Traboulsee, and R. Tam, "Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1229–1239, May 2016, doi: 10.1109/TMI.2016.2528821.

[26] M. Talo, U. B. Baloglu, and U. Rajendra Acharya, "Application of deep transfer learning for automated brain abnormality classification using MR images," *Cognit. Syst. Res.*, vol. 54, pp. 176–188, May 2019, doi: 10.1016/j.cogsys.2018.12.007.

[27] M. Ahmad, D. Ai, G. Xie, S. F. Qadri, H. Song, Y. Huang, Y. Wang, and J. Yang, "Deep belief network modeling for automatic liver segmentation," *IEEE Access*, vol. 7, pp. 20585–20595, 2019, doi: 10.1109/ACCESS.2019.2896961.

[28] M. Ahmad, J. Yang, D. Ai, S. F. Qadri, and Y. Wang, "Deep-stacked auto encoder for liver segmentation," in *Commun. Comput. Inf. Sci.*, vol. 757, pp. 243–251, Jul. 2018, doi: 10.1007/978-981-10-7389-2_24.

[29] M. Ahmad, Y. Ding, S. F. Qadri, and J. Yang, "Convolutional-neural-network-based feature extraction for liver segmentation from CT images," in *Proc. 11th Int. Conf. Digit. Image Process. (ICDIP)*, Aug. 2019, p. 159, doi: 10.1117/12.2540175.

[30] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, and C. B. Kendall, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.

[31] H.-I. Suk, S.-W. Lee, and D. Shen, "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis," *NeuroImage*, vol. 101, pp. 569–582, Nov. 2014, doi: 10.1016/j.neuroimage.2014.06.077.

[32] J.-Z. Cheng, D. Ni, Y.-H. Chou, J. Qin, C.-M. Tiu, Y.-C. Chang, C.-S. Huang, D. Shen, and C.-M. Chen, "Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in US images and pulmonary nodules in CT scans," *Sci. Rep.*, vol. 6, no. 1, Apr. 2016, Art. no. 24454, doi: 10.1038/srep24454.

[33] F. F. Ting, Y. J. Tan, and K. S. Sim, "Convolutional neural network improvement for breast cancer classification," *Expert Syst. Appl.*, vol. 120, pp. 103–115, Apr. 2019, doi: 10.1016/j.eswa.2018.11.008.

[34] A. Setio, F. Ciompi, G. Litjens, and P. Gerke, "Pulmonary nodule detection in CT images?: False positive reduction using multi-view convolutional networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1160–1169, Mar. 2016, doi: 10.1109/TMI.2016.2536809.

[35] H. Shin, H. Roth, M. Gao, L. Lu, Z. Xu, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.* vol. 35, no. 5, pp. 1285–1298, Feb. 2016, doi: 10.1109/TMI.2016.2528162.

[36] C. Shie, C. Chuang, C. Chou, M. Wu, and E. Y. Chang, "Transfer representation learning for medical image analysis," in *Proc. 37th Annu. Int. Conf. Eng. Med. Biol. Soc.*, Aug. 2015, pp. 711–714.

[37] C. Cao, "Deep learning and its applications in biomedicine," *Genomics, Proteomics Bioinform.*, vol. 16, no. 1, pp. 17–32 2017.

[38] G. E. Hinton. (2009). *Scholorpedia*. [Online]. Available: http://www.scholarpedia.org/article/Deep_belief_networks

[39] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: Algorithms and applications," *IEEE Access*, vol. 3, pp. 490–530, 2015.

[40] H. Wang, B. Zheng, S. W. Yoon, and H. S. Ko, "A support vector machine-based ensemble algorithm for breast cancer diagnosis," *Eur. J. Oper. Res.*, vol. 267, no. 2, pp. 687–699, Jun. 2018, doi: 10.1016/j.ejor.2017.12.001.

[41] M. Velikova, M. Samulski, P. J. F. Lucas, and N. Karssemeijer, "Improved mammographic CAD performance using multi-view information: A Bayesian network framework," *Phys. Med. Biol.*, vol. 54, no. 5, pp. 1131–1147, Mar. 2009, doi: 10.1088/0031-9155/54/5/003.

[42] A. Keleá, A. Keleá, and U. Yavuz, "Expert system based on neuro-fuzzy rules for diagnosis breast cancer," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5719–5726, May 2011, doi: 10.1016/j.eswa.2010.10.061.

[43] L. Sun, J. Wang, Z. Hu, Y. Xu, and Z. Cui, "Multi-view convolutional neural networks for mammographic image classification," *IEEE Access*, vol. 7, pp. 126273–126282, 2019, doi: 10.1109/ACCESS.2019.2939167.

[44] M. Samulski and N. Karssemeijer, "Optimizing case-based detection performance in a multiview CAD system for mammography," *IEEE Trans. Med. Imag.*, vol. 30, no. 4, pp. 1001–1009, Apr. 2011, doi: 10.1109/TMI.2011.2105886.

[45] N. Zemmal, N. Azizi, N. Dey, and M. Sellami, "Adaptive semi supervised support vector machine semi supervised learning with features cooperation for breast cancer classification," *J. Med. Imag. Health Informat.*, vol. 6, no. 1, pp. 53–62, Feb. 2016.

[46] K. Polat and U. Senturk, "A novel ML approach to prediction of breast cancer: Combining of mad normalization, KMC based feature weighting and AdaBoostM1 classifier," in *Proc. 2nd Int. Symp. Multidiscip. Stud. Innov. Technol. Proc.*, Oct. 2018, pp. 1–4, doi: 10.1109/ISMSIT.2018.8567245.

[47] M. S. Ashraf, O. I. and Siti, "Intelligent breast cancer diagnosis based on enhanced Pareto optimal and multilayer perceptron neural network.," *Int. J. Comput. Aided Eng. Technol.*, vol. 10, no. 5, pp. 543–556, 2018.

[48] A. J. Bekker, H. Greenspan, and J. Goldberger, "A multi-view deep learning architecture for classification of breast microcalcifications," in *Proc. Int. Symp. Biomed. Imag.*, Jun. 2016, pp. 726–730, doi: 10.1109/ISBI.2016.7493369.

[49] R. MurtiRawat, S. Panchal, V. K. Singh, and Y. Panchal, "Breast cancer detection using K-nearest neighbors, logistic regression and ensemble learning," in *Proc. Int. Conf. Electron. Sustain. Commun. Syst.*, Jul. 2002, pp. 534–540, doi: 10.1109/icesc48915.2020.9155783.

[50] M. Nilashi, O. Ibrahim, H. Ahmadi, and L. Shahmoradi, "A knowledge-based system for breast cancer classification using fuzzy logic method," *Telematics Informat.*, vol. 34, no. 4, pp. 133–144, Jul. 2017, doi: 10.1016/j.tele.2017.01.007.

[51] J. Lee, "2017-Lee-deep learning in medical imaging_gen," *Medical*, vol. 18, no. 4, pp. 570–584, 2017.

[52] S. F. Qadri, M. Ahmad, D. Ai, J. Yang, and Y. Wang, "Deep belief network based vertebra segmentation for CT images," *Commun. Comput. Inf. Sci.*, vol. 875, pp. 536–545, Dec. 2018, doi: 10.1007/978-981-13-1702-6_53.

[53] S. F. Qadri, Z. Zhao, D. Ai, M. Ahmad, and Y. Wang, "Vertebrae segmentation via stacked sparse autoencoder from computed tomography images," in *Proc. 11th Int. Conf. Digit. Image Process. (ICDIP)*, Aug. 2019, p. 160, doi: 10.1117/12.2540176.

[54] M. Yasmin, M. Sharif, and S. Mohsin, "Survey paper on diagnosis of breast cancer using image processing techniques," *Res. J. Recent Sci.*, vol. 2, no. 10, pp. 88–98, 2013. [Online]. Available: https://www.isca.me

[55] T. D. Truong and H. T. T. Pham, "Breast cancer histopathological image classification utilizing convolutional neural network," *IFMBE Proc.*, vol. 69, pp. 531–536, Dec. 2020, doi: 10.1007/978-981-13-5859-3_92.

[56] N. Wu, J. Phang, J. Park, Y. Shen, and Z. Huang, "Deep neural networks improve Radiologists' performance in breast cancer screening," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 1184–1194, Apr. 2020, doi: 10.1109/tmi.2019.2945514.

[57] B. Wang, K. Yager, D. Yu, and M. Hoai, "X-ray scattering image classification using deep learning," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2017, pp. 697–704, doi: 10.1109/WACV.2017.83.

[58] Q. Song, L. Zhao, X. Luo, and X. Dou, "Using deep learning for classification of lung nodules on computed tomography images," *J. Healthcare Eng.*, vol. 2017, pp. 1–7, 2017, doi: 10.1155/2017/8314740.

[59] Y. Gultom, A. M. Arymurthy, and R. J. Masikome, "Batik classification using deep convolutional network transfer learning," *J. Ilmu Komputer Inf.*, vol. 11, no. 2, p. 59, Jun. 2018, doi: 10.21609/jiki.v11i2.507.

[60] S. Chambon, M. N. Galtier, P. J. Arnal, G. Wainrib, and A. Gramfort, "A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 4, pp. 758–769, Apr. 2018, doi: 10.1109/TNSRE.2018.2813138.

[61] N. Coudray, P. S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, A. L. Moreira, N. Razavian, and A. Tsirigos, "Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning," *Nature Med.*, vol. 24, no. 10, pp. 1559–1567, Oct. 2018, doi: 10.1038/s41591-018-0177-5.

[62] N. Hnoohom and S. Yuenyong, "Thai fast food image classification using deep learning," in *Proc. Int. ECTI Northern Sect. Conf. Electr., Electron., Comput. Telecommun. Eng. (ECTI-NCON)*, Feb. 2018, pp. 116–119, doi: 10.1109/ECTI-NCON.2018.8378293.

[63] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Skin cancer classification using deep learning and transfer learning," in *Proc. 9th Cairo Int. Biomed. Eng. Conf. (CIBEC)*, Dec. 2018, pp. 90–93, doi: 10.1109/CIBEC.2018.8641762.

[64] H. Mohsen, E.-S.-A. El-Dahshan, E.-S.-M. El-Horbaty, and A.-B.-M. Salem, "Classification using deep learning neural networks for brain tumors," *Future Comput. Informat. J.*, vol. 3, no. 1, pp. 68–71, Jun. 2018, doi: 10.1016/j.fcij.2017.12.001.

[65] A. M. Abdel-Zaher and A. M. Eldeib, "Breast cancer classification using deep belief networks," *Expert Syst. Appl.*, vol. 46, pp. 139–144, Mar. 2016, doi: 10.1016/j.eswa.2015.10.015.

[66] M. A. Al-antari, M. A. Al-masni, S.-U. Park, J. Park, M. K. Metwally, Y. M. Kadah, S.-M. Han, and T.-S. Kim, "An automatic computer-aided diagnosis system for breast cancer in digital mammograms via deep belief network," *J. Med. Biol. Eng.*, vol. 38, no. 3, pp. 443–456, Jun. 2018.

[67] N. Dhungel, G. Carneiro, and A. P. Bradley, "A deep learning approach for the analysis of masses in mammograms with minimal user intervention," *Med. Image Anal.*, vol. 37, pp. 114–128, Apr. 2017, doi: 10.1016/j.media.2017.01.009.

[68] C. Li, Y. Wang, X. Zhang, H. Gao, Y. Yang, and J. Wang, "Deep belief network for spectral–spatial classification of hyperspectral remote sensor data," *Sensors*, vol. 19, no. 1, p. 204, Jan. 2019, doi: 10.3390/s19010204.

[69] T. Liu, Y. Tian, S. Zhao, X. Huang, and Q. Wang, "Automatic whole heart segmentation using a two-stage U-Net framework and an adaptive threshold window," *IEEE Access*, vol. 7, pp. 83628–83636, 2019, doi: 10.1109/ACCESS.2019.2923318.

[70] M. Pop, M. Sermesant, P. M. Jodoin, A. Lalande, X. Zhuang, G. Yang, A. Young, and O. Bernard, Eds., "Statistical atlases and computational models of the heart. ACDC and MMWHS challenges," in *Proc. 8th Int. Workshop, STACOM, Held Conjunct. MICCAI*, vol. 10663. Quebec City, QC, Canada: Springer, Sep. 2018.

[71] S. Furqan Qadri, D. Ai, G. Hu, M. Ahmad, Y. Huang, Y. Wang, and J. Yang, "Automatic deep feature learning via patch-based deep belief network for vertebrae segmentation in CT images," *Appl. Sci.*, vol. 9, no. 1, p. 69, Dec. 2018, doi: 10.3390/app9010069.

[72] V. Nair and G. E. Hinton, "3D object recognition with deep belief nets," in *Proc. 22nd Int. Conf. Neural Inf. Process.*, 2009, pp. 1339–1347, doi: 10.5555/2984093.2984244.

[73] S. Zhou, Q. Chen, and X. Wang, "Active deep networks for semi-supervised sentiment classification," in *Proc. 23rd Int. Conf. Comput. Linguist. Proc. Conf.*, vol. 2, Aug. 2010, pp. 1515–1523.

[74] Y. Rizk, N. Hajj, N. Mitri, and M. Awad, "Deep belief networks and cortical algorithms: A comparative study for supervised classification," *Appl. Comput. Informat.*, vol. 15, no. 2, pp. 81–93, 2019, doi: 10.1016/j.aci.2018.01.004.

[75] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006, doi: 10.1162/neco.2006.18.7.1527.

[76] A. Fischer and C. Igel, "An introduction to restricted Boltzmann machines," *Image Anal. Comput. Vis., Appl.*, vol. 4, pp. 14–36, Sep. 2012. [Online]. Available: http://image.diku.dk/igel/paper/AItRBM-proof.pdf

[77] A. Cruz-Roa, H. Gilmore, A. Basavanhally, M. Feldman, S. Ganesan, N. Shih, J. Tomaszewski, A. Madabhushi, and F. González, "High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: Application to invasive breast cancer detection," *PLoS One*, vol. 13, no. 5, 2018, Art. no. e0196828, doi: 10.1371/journal.pone.0196828.

[78] Z. Zainudin, S. Mariyam, and Shafaatunnur, "Deep layer CNN architecture for breast cancer histopathology image detection," in *Proc. Int. Conf. Adv. Mach. Learn. Technol. Appl.*, 2020, pp. 43–51, doi: 10.1007/978-3-030-14118-9_5.

[79] A. Cruz-Roa, A. Basavanhally, F. Gonzlez, H. Gilmore, M. Feldman, S. Ganesan, N. Shih, J. Tomaszewski, and A. Madabhushi, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," *Med. Imag. Digit. Pathol.*, vol. 9041, no. 216, 2014, Art. no. 904103, doi: 10.1117/12.2043872.

[80] T. Araujo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, A. Polónia, and A. Campilho, "Classification of breast cancer histology images using convolutional neural networks," *PLoS One*, vol. 12, no. 6, 2017, Art. no. e0177544, doi: 10.1371/journal.pone.0177544.

IRUM HIRRA received the B.Sc. degree from the University of the Punjab, Pakistan, in 2015, and the M.S.C.S. degree from the Department of Computer Science, University of Lahore, Pakistan, in 2020. Her research interests include computer science, medical image processing, and artificial intelligence.

MUBASHIR AHMAD received the B.Sc. (Hons) degree from Hazara University Mansehra, Pakistan, the M.S.C.S. degree from IQRA University Islamabad, Pakistan, in 2012, and the Ph.D. degree from the Beijing Institute of Technology China, from 2014 to 2019. From 2012 to 2014, he was a php programmer in Microcom IT. He is currently an Assistant Professor with the Computer Science Department, University of Lahore, Pakistan. His research interests include deep learning, pattern recognition, and medical imaging. He was a recipient of Chinese government scholarship for the Ph.D. studies and got excellent research paper award in IGTA-Springer at Beijing, China, in 2017. He also serves as a Reviewer for IEEE Access and other Journals.

AYAZ HUSSAIN received the M.S. degree in computer engineering from the University of Engineering and Technology at Taxila (UET Taxila), Pakistan, in 2018. He is currently a Lecturer with the Department of Computer Science, Knowledge Unit System and Technology, University of Management and Technology, Sialkot, Pakistan. He has published many research articles in his field and has experience in the Internet of Things, machine learning, image processing, and fault-tolerant systems.

M. USMAN ASHRAF received the Ph.D. degree in computer science from King Abdul-Aziz University, Saudi Arabia, in 2018. He is currently an Assistant Professor and the Head of the Software Engineering Department at KUST, University of Management and Technology, Sialkot, Pakistan. His research on exascale computing systems, high performance computing (HPC) systems, parallel computing, HPC for deep learning and location based services system has appeared in the IEEE Access, *IET Software, International Journal of Advanced Research in Computer Science, International Journal of Advanced Computer Science and Applications, I.J. Information Technology and Computer Science, International Journal of Computer Science and Security*, and several International IEEE/ACM/Springer conferences. He has served as HPC Scientist at HPC Centre King Abdul-Aziz University, Saudi Arabia.

IFTIKHAR AHMED SAEED graduated in information technology with the College of Information and Electrical Engineering, China Agricultural University. He received the master's degree in computer science and engineering technology from Pakistan. He is currently working as an Assistant Professor with the Department of Computer Science, The University of Lahore. His research interests include, design and development of soil sensors, multi-sensor technique, multi-sensor fusion and the Agricultural Internet of Things (IoT).

SYED FURQAN QADRI received the Ph.D. degree from the Beijing Institute of Technology, Beijing, China, in 2019. He is currently working a Postdoc fellowship with Shenzhen University. Prior to this, he worked at GC University Faisalabad, Layyah Campus, Pakistan, as a Lecturer. His research focuses on machine learning, especially on supervised and unsupervised deep learning techniques, image segmentation, classification, pattern recognition, medical image analysis, computer vision. He has conducted extensive research and authored many peer-reviewed SCI- and EI-indexed scientific papers. He has received Best Paper Awards in ICDIP-SPIE 2019 and IGTA-Springer 2017. He also serves as a Reviewer for IEEE Access, Journal of Electronic Imaging (JEI) and for many other journals.

AHMED MOHAMMED ALGHAMDI received the B.Sc. degree in computer science and the first M.Sc. degree in business administration from King Abdulaziz University, Jeddah, Saudi Arabia, in 2005, the second master's degree in internet computing and network security from Loughborough University, U.K., in 2013, and the Ph.D. degree in computer science from King Abdulaziz University, Jeddah, Saudi Arabia. He is currently an Assistant Professor with the Software Engineering Department, College of Computer Science and Engineering, University of Jeddah, Saudi Arabia. He also has over 11 years of working experience before attending the academic carrier. His research interests include high-performance computing, big data, distributed systems, programming models, software engineering, and software testing.

AHMED S. ALFAKEEH graduated in computer science and education from King Abdulaziz University, in 2005. He received the M.Sc. degree in IT from De Montfort University, in 2010, and the Ph.D. degree from the Software Technology Research Laboratory, a research institute established within De Montfort University, Leicester, U.K., in 2015. He is currently an Assistant Professor with the Information System Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. He is currently a Vice-Dean of Research and Consultation Institute for Development. His research projects involve feature interactions detection and resolution in context-aware systems, smart homes systems, software engineering, the IoT and data science.