Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning

Hanan Aljuaid [a], Nazik Alturki [b], Najah Alsubaie [a], Lucia Cavallaro [c], Antonio Liotta [c,*]

[a] Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), PO Box 84428, Riyadh 11671, Saudi Arabia
[b] Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), PO Box 84428, Riyadh 11671, Saudi Arabia
[c] Faculty of Computer Science, Free University of Bozen-Bolzano, Piazza Domenicani, 3, Bolzano 39100, Italy

A B S T R A C T

Background and objective: Many developed and non-developed countries worldwide suffer from cancer-related fatal diseases. In particular, the rate of breast cancer in females increases daily, partially due to unawareness and undiagnosed at the early stages. A proper first breast cancer treatment can only be provided by adequately detecting and classifying cancer during the very early stages of its development. The use of medical image analysis techniques and computer-aided diagnosis may help the acceleration and the automation of both cancer detection and classification by also training and aiding less experienced physicians. For large datasets of medical images, convolutional neural networks play a significant role in detecting and classifying cancer effectively.

Methods: This article presents a novel computer-aided diagnosis method for breast cancer classification (both binary and multi-class), using a combination of deep neural networks (ResNet 18, ShuffleNet, and Inception-V3Net) and transfer learning on the BreakHis publicly available dataset.

Results and conclusions: Our proposed method provides the best average accuracy for binary classification of benign or malignant cancer cases of 99.7%, 97.66%, and 96.94% for ResNet, InceptionV3Net, and ShuffleNet, respectively. Average accuracies for multi-class classification were 97.81%, 96.07%, and 95.79% for ResNet, Inception-V3Net, and ShuffleNet, respectively.

1. Introduction

The whole world is suffering from fatal diseases. According to the World Health Organization (WHO) [1], cancer is the second leading cause of death. Specifically, female breast cancer is much higher in undeveloped countries than in developed ones. For example, in Pakistan, The rate of breast cancer diagnosed yearly is of 1.38 million patients, one third of whom die. Globally the death rate due to cancer is 9.6 million [2], with a survival rate of 1.7 million [3].

Hence, it is vital to detect cancer and start treating it at its very early stages; otherwise, cancer may spread to affect the whole breast or other body parts. Upon an accurate and effective diagnosis, a correct first treatment may increase the survival rate by up to 80% in the case of breast cancer [1]. The lesions caused by breast cancer are categorized as benign and malignant. Both categories may be cancerous or noncancerous. For example, the abnormalities in the epithelial cells are due to benign lesions, but they are unable to grow further. Hence, these will not lead to breast cancer [4]. On the other hand, malignant cells are considered dangerous due to their irregular growth in the body and may be regarded as cancerous cells. It is currently challenging to correctly analyze and classify benign and malignant lesions in microscopic images [5].

Many strategies can be adopted for the early diagnosis of breast cancer, including screening tests, self-internal feelings of distur-

bance, or visiting a nearby doctor. These strategies will diminish the pace of mortality and lift an opportunity for effective treatment [6]. Clinically speaking, many methods are being used by radiologists, such as mammography and biopsy, to detect breast cancer. Radiologists detect the initial cancer symptoms by taking mammography breast cancer images; indeed, a study confirmed that images increase the survival rate [7].

Another approach used by pathologists is a biopsy. It collects a tissue sample from the affected area of the breast and analyzes it through a microscope to detect and classify the tumor [8]. Hence, this method is considered proficient and precise for detecting and classifying breast cancer.

Previously some Science and Technology researchers have attempted to detect breast cancer by examining and analyzing the cancer cells [9]. They classified the cancerous cells into benign and malignant by performing the nuclei analysis and cell classification [10]. In medical image processing, several sound-efficient algorithms have been developed to correctly detect the area of interest and their classification [11]. However, there is still a need for a system or framework that may produce highly accurate breast cancer identification and classification results. Efficiency and accuracy in medical diagnosis results are still an open challenge due to the performance degradation. In remote places, like tribal areas or where there is a lack of resources (i.e., unavailability of expert doctors and heavy-cost machinery), there is a need for a system that may automatically suggest the type of the disease. As a result, the proper treatment may be started by the physicians.

Technology is rapidly evolving in the field of medical diagnosis. Computer-aided diagnosis (CAD) has received great acceptance worldwide due to its outstanding results in terms of accuracy and efficiency. Artificial intelligence is significantly involved, particularly machine learning and deep learning. Thanks to the use of CAD software technologies, breast cancer images can be further classified into benign and malignant lesions. Machine learning has many conventional classifiers such as Support Vector Machine (SVM), Naive Bayes (NB), Convolutional Neural Network (CNN) etc. However, CNN is one of the most remarkable classifiers due to its accuracy in terms of classification and feature extraction [12]. It performs the classification task based on pattern matching in the images contained in large datasets [13]. Keeping in mind the complex nature of integrated systems, deep learning plays an extraordinary role in medical image analysis. Indeed, for breast cancer classification, DL uses the concept of CNN, in which there are different types of layers included in the task of classification. To incorporate the CNN architecture, there is a need for many images to detect cancerous cells efficiently. There are different types of layers included in the CNN architecture and the ones are known as input layers, hidden layers, and output layers. Features may be obtained by the learning process of images with each other, which may be done on the hidden layers. The workflow and the architecture of CNN are shown in Figs. 1 and 2, respectively.

Image features can be learned from the connected layers to classify breast cancer images. In addition, features may be extracted automatically from images using pre-trained deep neural networks (ResNet, Inception-V3Net and ShuffleNet) as performed in this research.

The main contribution of the proposed method is to enhance the accuracy for binary classification of benign or malignant cancer cases by 99.7%, 97.66%, and 96.94% for ResNet, InceptionV3Net, and ShuffleNet, respectively. In addition, average accuracies for multi-class classification were 97.81%, 96.07%, and 95.79% for ResNet, Inception-V3Net, and ShuffleNet, respectively. This research has used a most efficient computer-aided diagnosis technique to correctly classify the breast cancer images into binary and multi-classes. We achieved higher accuracies by using the combination of transfer learning and deep neural networks.

The rest of the paper is organized as follows: Section 2 highlights the relevant literature related to the detection and classification of cancer, and Section 3 describes the methodology adopted by this study. Section 4 reports the experimental results obtained using pre-trained deep neural networks for breast cancer classification. Finally, the paper is concluded in Section 5.

2. Previous works

The use of probabilistic neural networks and SVM has been adopted by George et al. to detect cell nuclei and breast cancer classification. Experiments were performed on the breast cytological images and the obtained results were compared by the resulting error rate, correct detection rate, sensitivity, and specificity. They claim that the obtained results through their respective methodology are much more effective and can be applied to multiple datasets [9]. Sharma et al. presented a comprehensive study on breast cancer classification using conventional machine and deep learning (i.e., ML and DL, respectively) approaches. They extracted the image features based on color histogram and Haralick textures to classify them into benign and malignant lesions. The accuracies achieved by their proposed method were recorded between 93.25% and 93.97% [14].

Chugh et al. presented a detailed survey on the use of ML and DL applications to diagnose breast cancer. They thoroughly covered the literature survey and studies related to breast cancer classification. They also highlighted the attributes of these techniques in both positive and negative aspects. The conclusions the authors of this study reached by encompassing the previous studies are that deep learning techniques are much more suitable in the classification task of breast cancer images when the datasets are more extensive [15]. Houssein et al. presented a detailed and comprehensive review on deep and machine learning techniques for medical imaging-based breast cancer detection and classification. They showed all the new applications used for medical diagnosis and the rapid involvement of deep and machine learning in the medical field [16].

Hamed et al. proposed using machine learning-based models in their research work for the classification task of breast cancer. They claimed that the physicians' average accuracy obtained in the detection and classification of breast cancer is around 79%, while the accuracy obtained by their proposed model is 91% [17]. Tiwari et al. used the Wisconsin Breast Cancer Dataset to classify breast cancer images having 569 samples and 30 features. The dataset was collected from the Kaggle repository. They measured the efficiency of their work in terms of accuracy and precision. Their algorithms were logistic regression, SVM, and K Nearest Neighbor, an artificial neural network. They have used them separately to get the required results. As a result, they achieved 99.3% maximum average accuracy for the classification task of breast cancer images [18,19].

Ragab et al. proposed a multi-deep CNN for breast cancer classification. They extracted the features from the images by using the pre-trained deep neural networks. The SVM classifier further used the same features with different kernel functions. The authors of this research study further used the principal component analysis to reduce the feature vector. They claimed that their results were the highest after comparing with the other state-of-the-art CAD systems [20]. Ashraf et al. presented an efficient technique for skin cancer classification using deep learning. For this purpose, they used the real-time dataset collected from the DHQ Faisalabad, Pakistan. They classified the skin cancer images into different types, such as melanoma and non-melanoma. Their results demonstrated 93.29% accuracy of classification [21].

Khan et al. used the concept of transfer learning and deep learning to classify the breast cancer images into benign and malignant tumors. They extracted the features using GoogLeNet, VG-
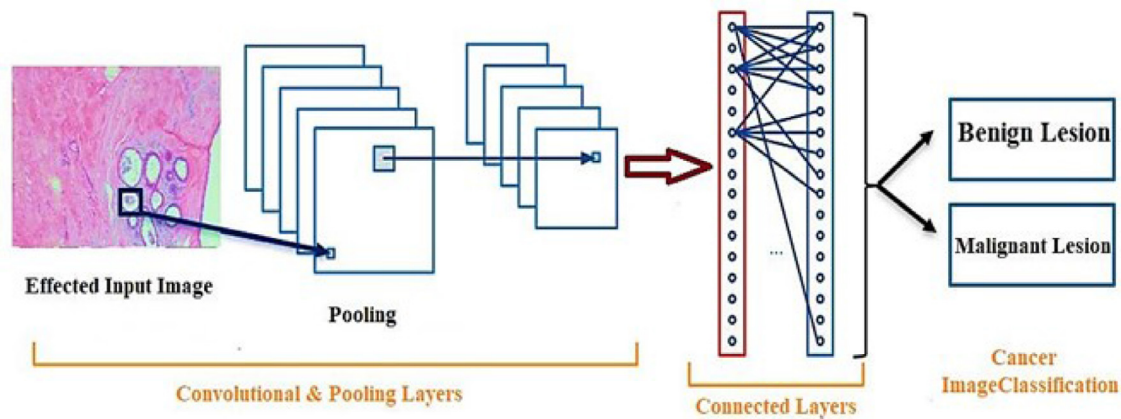
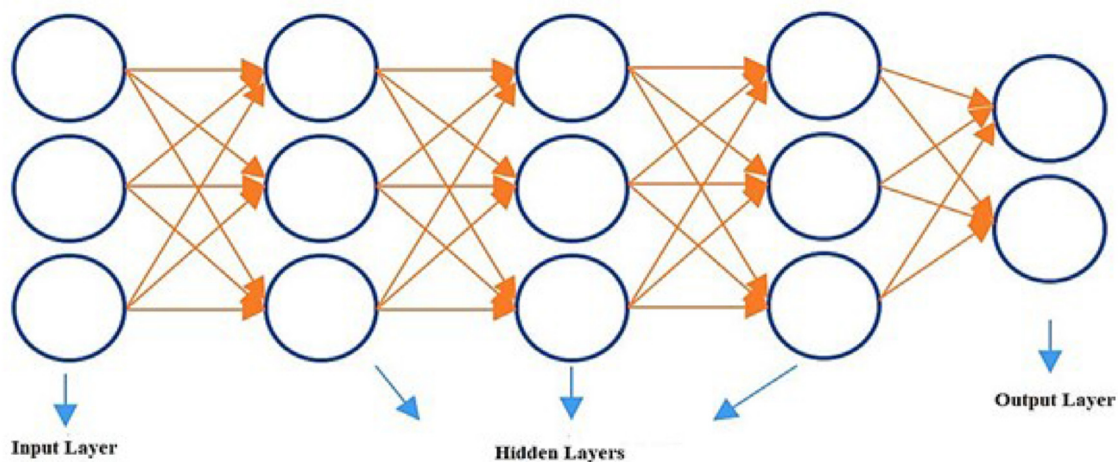Fig. 1. CNN workflow for breast cancer classification.

Fig. 2. CNN layers architecture.

GNet, and ResNet architectures pre-trained in practice. Furthermore, these features were added to the fully connected layers by incorporating the average pooling concept. The total images they used as a dataset in their work were 8000 and from these images, the network was trained with 6000 images and the remaining 2000 images were used for testing. The maximum average accuracy achieved by their proposed method was recorded as 97.25% [22]. In their research work, Khan et al. proposed using DCNN to classify the human burnt skin images. The images contained in their dataset were 450. They trained the network with 65% of images while the accuracy was tested with the remaining 35%. As a result, they achieved 79.4% accuracy in classifying the images [23]. In another work, they increased the dataset, which contained 600 images. The same technique was used for the segmentation and classification of burnt human skin. 60% and 40% ratio was used for training and testing the accuracy of the classifier. As a result, they achieved 83.4% accuracy in classifying the images [24].

Hameed et al. presented the ensembles of deep learning models to classify breast cancer images. They had taken the pre-trained VGG16 and VGG19 architectures to train four different models. Their overall accuracy was 95.29% in classifying the breast cancer images in different classes [25]. Gupta et al. classified the breast cancer images into malignant and non-malignant tumors using different supervised machine learning algorithms. Their algorithms were k-Nearest Neighborhood, Logistic Regression, Decision Tree, Random Forest, and SVM. They also used the Adam Gradient Descent Learning models and the accuracy was 98.24% [26].

Murtaza et al. presented a comprehensive study to classify breast cancer images through deep learning and imaging techniques. They took 49 different research works published in journals and conferences from 8 different repositories. They highlighted all the ML and DL algorithms used for different types of medical image classification tasks, including breast cancer classification. They presented the ten open research challenges, including the automatic detection of breast cancer for new researchers [27]. Zhang et al. presented the clinical information of breast cancer patients by the use of Bidirectional Encoder Representations (BER). They had further used the deep learning models to extract the attributes or features from the infected images of breast cancer patients. The highest results achieved by their said techniques were recorded as 96.73% [28]. Many studies have been done on image segmentation and classification using machine learning [36-38].

Zheng et al. presented a mathematically implemented Deep Learning assisted Efficient Adaboost Algorithm (DLAEABA) having the capabilities of advanced level of computations in their research work for the detection and classification of breast cancer. They also used some computer vision techniques as well. They also presented a comparison by showing the experimental results having an accuracy of 97.2% [29]. Krithiga et al. presented an intuitive approach with some fuzzy logic techniques properties to detect cell nuclei. They used deep CNN for the segmentation and classification of breast cancer. Their proposed method was much more effective in terms of accuracy and less computational time. They used ten-fold cross-validation method to check the accuracy of the classifi-
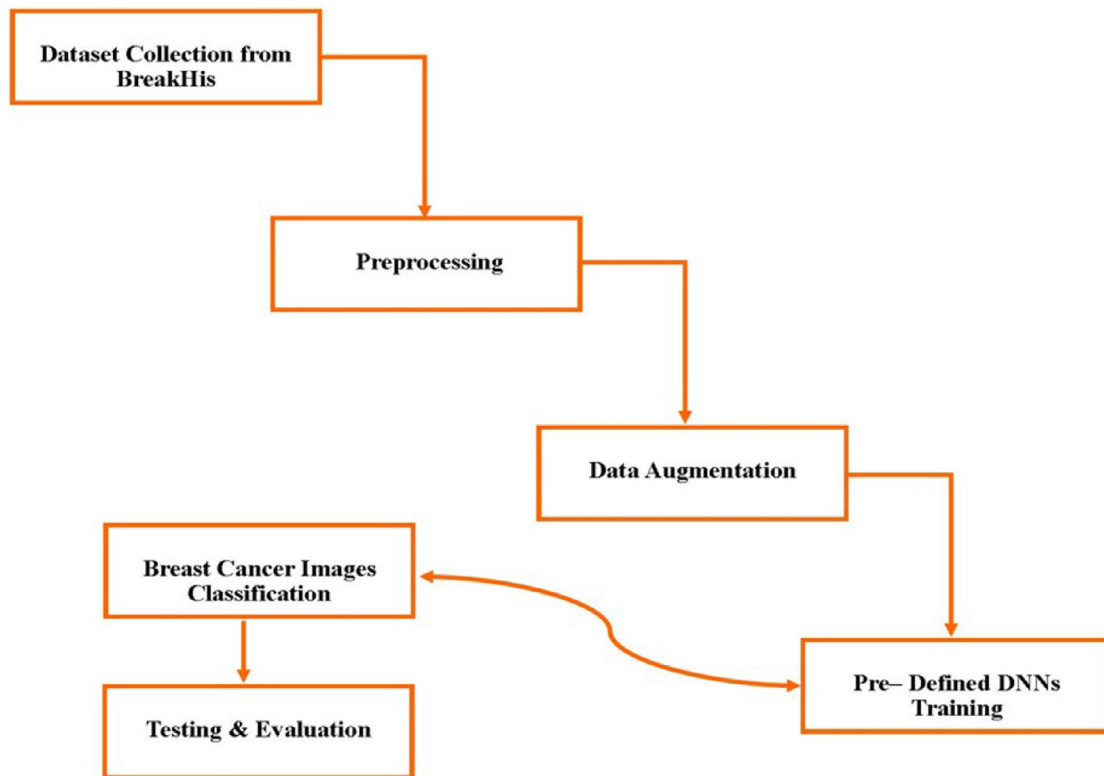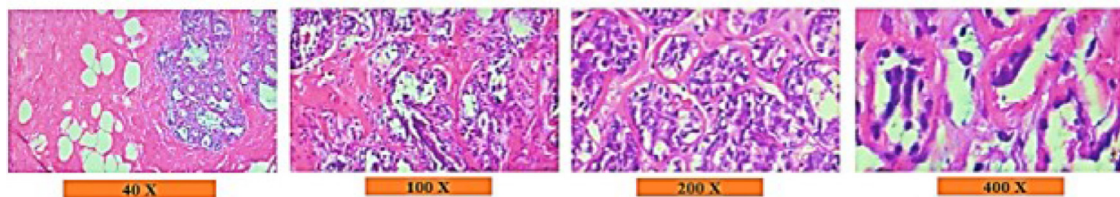
Fig. 3. Methodology workflow.

Fig. 4. Benign lesion under all magnification factors.

cation work. The accuracy they showed in their proposed research work was 98.62% [30].

The comprehensive literature survey in this section concludes that there is a high accuracy lack of early detection of cancer using deep and transfer learning. From the cited literature it can clearly complicated task due to non availability of resources or lack of experienced and qualified specialists. A lot of work in this regard has already been done by the researchers belonging to the medical field but there exists lack of accuracies in their findings. To overcome these challenges we have tried to improve the way to correctly classify the breast cancer images by incorporating the concept of deep learning as well as transfer learning so that diagnosis at the early stages of breast cancer may be done with high accuracies and promising results.

3. Methodology

This section describes the adopted methodology to classify breast cancer images into different categories. Binary and multiclass classifications have been performed using pre-trained deep neural network models. A detailed methodology is shown in Fig. 3.

3.1. Dataset collection

The images have been collected from the Robotic Vision and Imaging Laboratory linked with the informatics department, University of Parana, Brazil [31]. There are available a vast number of images related to cancer patients. The dataset is typically known as the Breast Cancer Histopathological imaging Database (BreakHis) and has been obtained by filling out the online form available on the website of the vision laboratory. The total numbers of samples available are 7909 and have been collected from 82 different patients. The dataset used here in this research work has benign and malignant breast cancer images. All the available images in the dataset have four different magnification factors, i.e., 40×, 100×, 200×, and 400×. The quantity of benign and malignant lesions in the dataset is 2480 and 5429, respectively. Moreover, benign images address four classes: Adenosis (AE), Fibroadenoma (FA), Phyllodes Tumor (PT), and Tubular Adenoma (TA). In the same way, malignant images address four classes: Ductal Carcinoma (DC), Lobular Carcinoma (LC), Mucinous Carcinoma (MC) and Papillary Carcinoma (PC). Samples of benign and malignant lesions under all magnification factors may be viewed in Figs. 4 and 5, respectively.
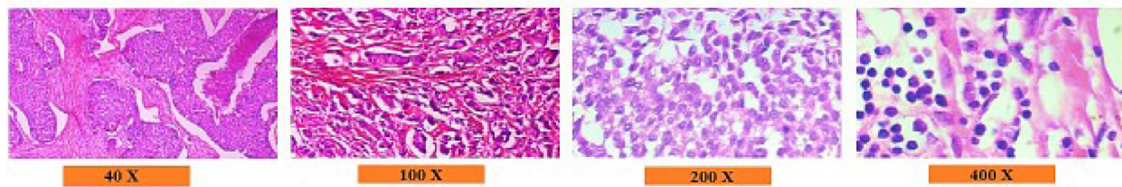
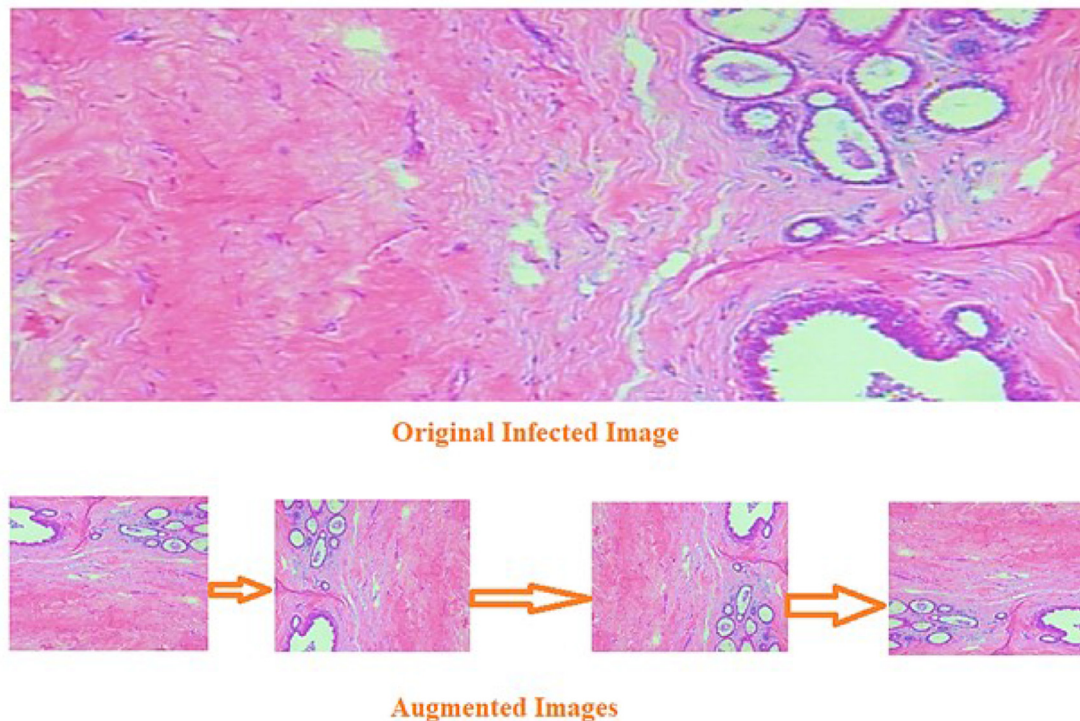**Fig. 5.** Malignant lesion under all magnification factors.

Fig. 6. Infected breast image augmentation.

Table 1
Evaluation metrics.

Sr. no.	Metrics	Mathematical expressions
1	Accuracy	$\mathcal{T}pos+\mathcal{T}neg/\mathcal{T}pos+\mathcal{F}pos+\mathcal{T}neg+\mathcal{F}neg$
2	Precision (Predictive Value in Positive sense)	$\mathcal{T}pos/\mathcal{T}pos+\mathcal{F}pos$
3	Recall (Positive Rate)	$\mathcal{T}pos/\mathcal{T}pos+\mathcal{F}neg$
4	Specificity (Negative Rate)	$\mathcal{T}neg/\mathcal{F}pos+neg$

Table 2
Classification results. B term is for binary, whereas MC represents multi-class condition.

Pre-trained DNNs	Accuracy		Precision		Sensitivity		Specificity	
	B	MC	B	MC	B	MC	B	MC
ResNet	99.7%	97.81%	99.59%	97.65%	97.53%	97.65%	97.8%	97.31%
Inception-V3Net	97.66%	96.07%	97.64%	96.05%	97.64%	96.03%	97.59%	96%
ShuffleNet	96.94%	95.79%	96.85%	95.70%	96.70%	95.70%	96.85%	95.5%

3.2. Pre-processing

Pre-processing images is an essential task in any work related to medical image analysis. However, the images obtained from the BrakeHis have different dimensions that may not fit the requirements of pre-trained deep neural networks, which we have used here in our research work for the classification purpose of breast cancer images. Thus, all the images have been passed through an editing phase to remove the noise, including the appearance of un-

desired traces and variations in the brightness or color of images and resize them into a uniform size, as per the requirements of pre-trained DNNs. To remove the noise in order to enhance the qualities of breast cancer images, two main filter has been used i.e. Median filter and gaussian filter. The main reason to use these filters are their abilities to keep maintain the original image features i.e. image edges and corners with their sharp structures remains the same. Herein, the image dimensions we used are 224 224 in
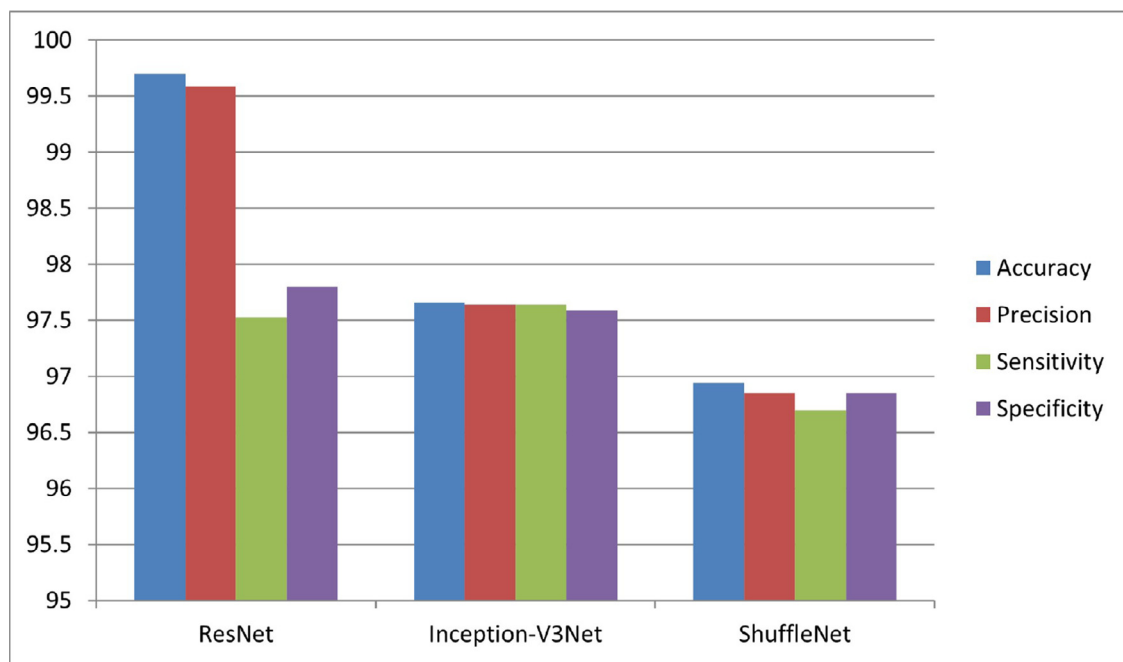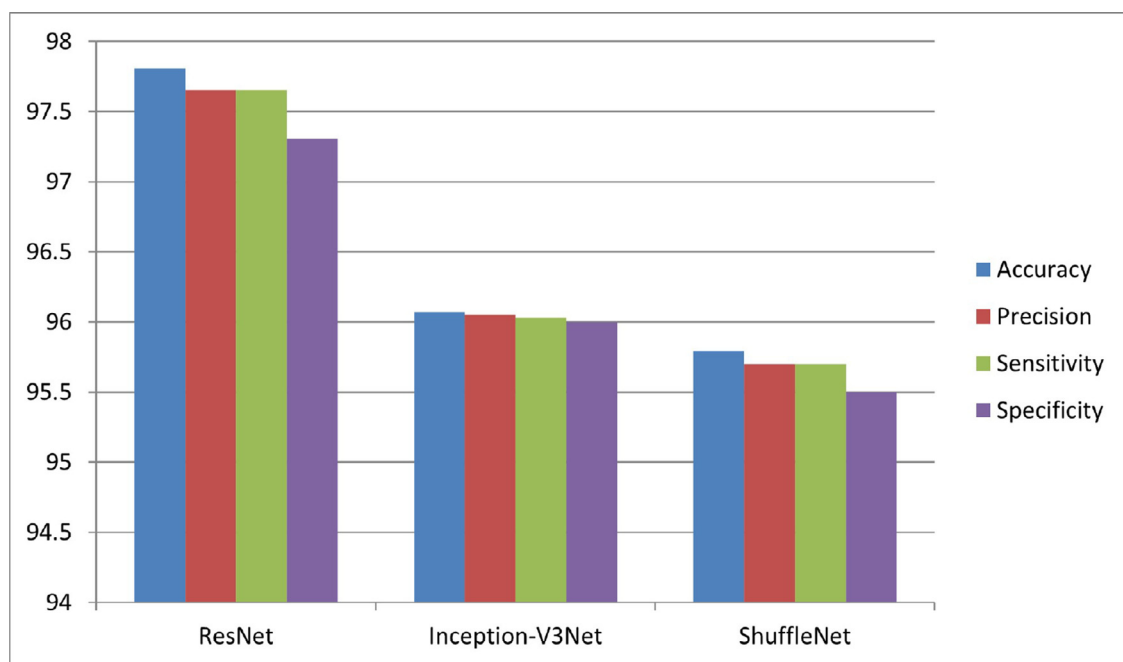
Fig. 7. Binary class classification performance.

Fig. 8. Multi class classification performance.

both ResNet18 and ShuffleNet, whereas we opted for a size of 299 × 299 in the case of Inception-V3Net.

3.3. Data augmentation

To correctly classify the breast cancer images into benign (resp., malignant) lesions, CNN needs massive data to learn its parameters. Data Augmentation is a standard benchmark or methodology to extend the data we take as a training data set [32]. Therefore, system performance may be increased by using data augmentation. It also reduces redundancy and overfitting by arranging the dataset in a balanced manner [33]. Random reflection, multiple rotations, and translations regarding horizontal or vertical aspects of breast

cancer images may be considered with different techniques to augment the data. Therefore, we performed the augmentation only on the training dataset instead of the testing one. 90 degree and 180 degree rotations as well as flipping has been performed during image augmentation process by maintaining the original image features. The size of augmented images has been reduced in order to enhance the accuracy as well as to speed up the process. A sample of data augmentation operations is shown in Fig. 6.

3.4. Pre-defined DNNs training

In medical imaging works, deep neural networks (DNNs) play a significant role as they can handle the complex nature of datasets.

A characteristic of DNNs is to learn the features robustly from images so that classification may be accomplished efficiently. We have used three pre-trained and defined DNN classifiers, that are ResNet, Inception-V3Net and ShuffleNet. All of them are CNNs, on which training has been done by using more than two million images from the Image net database and more than two thousand categories can be predicted by the image classification [34].

Our used DNNs such as ResNet 18, ShuffleNet, and Inception-V3Net have 18, 48, and 50 hidden layers can accomodate breast cancer images of different input sizes. These DNNs may also be developed using directed acyclic graphs. The concept of transfer learning has also been incorporated to classify breast cancer images with these DNNs. The main advantage of using the concept of transfer learning is to boost the classification process and reduces the time to complete the task.

3.5. Breast cancer images classification

Breast cancer image classification may be done by using two different kinds of approaches that are imaging-based or patient-based [35]. On the one hand, in an imaging-based approach, the data in the form of images may only be collected from a specific cancer patient that may be used further for training and testing purposes. On the other hand, a patient-based approach is a general approach in which imaging data of different patients may come under consideration for the training and testing purpose.

We have opted for the imaging-based technique because of the optimistic results obtained by using this approach. As we have discussed in the previous section that our used DNN's have multiple hidden layers to receive the input images of breast cancer in different sizes, so the extraction of different image properties may got easiness to correctly classify the images into the different classes of cancer. The BrakeHis dataset has been used in this research work and the images have been collected randomly, i.e., 65% of images were used for training and the remaining 35% images were used for testing purposes considering all the parameters.

3.6. Testing and evaluation

Testing and evaluation are the necessary matrices to determine the success rate of classifiers used herein to classify breast cancer images. We have expressed the success rate of pre-trained DNNs in terms of accuracy, precision, sensitivity and specificity. In addition, there will be different sets of observations and predictions typically referred to as true positive, true negative, false positive, and false negative. Using these metrics, all the measurements to determine the success rate is computed and these metrics may are shown in Table 1.

Eq. (1) computes the convolutional layers at different stages

$$O(i, j) = \sum k = 1 f\left(\sum l = 1 f input(i + k - 1, j + l - 1) kernel(k, l)\right) \tag{1}$$

where $i = 1, 2, \ldots, m - f + 1$ and $j = 1, 2, \ldots, n - f + 1$.

4. Results

This section explains the experimental results derived for breast cancer classification using three pre-trained DNNs and transfer learning. The experiments have been done on the publicly available BrakeHis dataset consisting of 7909 images derived from the diagnoses of 82 patients and are further sub-categorized into benign and malignant lesions. Furthermore, some images have also been collected from the Punjab Institute of Nuclear Medicine and Radiotherapy (PINUM), Faisalabad, for experimentation purposes. Implementation has been made in the Python framework. Binary

Table 3
Confusion matrix for ResNet (binary).

Result Class	C-I	182	10	99.96
		47.5%	2.2%	
	C-II	8	188	99.4
		2.0%	48.5%	
		99.97	99.43	99.7
		C-I	**C-II**	
	Target Class			

Table 4
Confusion matrix for InceptionV3Net (binary).

Result Class	C-I	178	4	98.7
		47.4%	1.0%	
	C-II	8	184	96.6
		2.0%	49.3%	
		96.6	98.6	97.6
		C-I	**C-II**	
	Target Class			

Table 5
Confusion matrix for ShuffleNet (binary).

Result Class	C-I	183	9	96.60
		48.9%	2.1%	
	C-II	2	180	97.2
		0.7%	48.5%	
		96.59	97.3	96.94
		C-I	**C-II**	
	Target Class			

and multi-class classification results in the calculated form are displayed in Table 2. Binary and multi-class classification results in graphical form are shown in Figs. 7 and 8, respectively.

As shown in Table 2, the ResNet DNNs provide the maximum average accuracy among other DNNs for both binary and multi-class classification. In addition, all eight sub-classes of malignant lesions have also been classified in multi-class classification. These are the highest results compared to the state-of-the-art studies, as shown in Fig. 9.

Binary class classification results using all the matrices are shown in Fig. 7. In addition, multi-class classification results using all the matrices have also been presented graphically by considering the magnifications 100× and 200×, respectively, as shown in Fig. 8. ResNet DNN provides the maximum average accuracies in both aspects of binary as well as multi-class classification and may be viewed in their graphical representations.

Some studies applied deep learning models to diagnostics of breast cancer disease utilizing genuine histopathological images from the BreakHis dataset. The vast majority of the past investigations focused on binary classification, but some studies also considered multi-class classification. All of the related studies used different pre-trained deep neural networks.

In this study, we applied three pre-trained (incorporating the concept of transfer learning) neural networks (i.e., ResNet18, Inception-V3Net, and ShuffleNet) in image-based binary classification (i.e., benign or malignant) and multi-class classification (eight classes) on real images from openly accessible BreakHis dataset. A comparison with state-of-the-art techniques is shown in Fig. 9.

In terms of binary classification, confusion matrices for all the pre-trained DNNs such as ResNet, InceptionV3Net and ShuffleNet used here in this task may be viewed in Tables 3–5, respectively. Similarly, confusion matrices in multi-class classification for all the pre-trained DNNs such as ResNet, InceptionV3Net and ShuffleNet used here in this task may be viewed in Tables 6–8, respectively.

In light of the results displayed in Table 6, ResNet among all the pre-trained DNNs provides the maximum classification results and is further followed by InceptionV3 Net and ShuffleNet results.
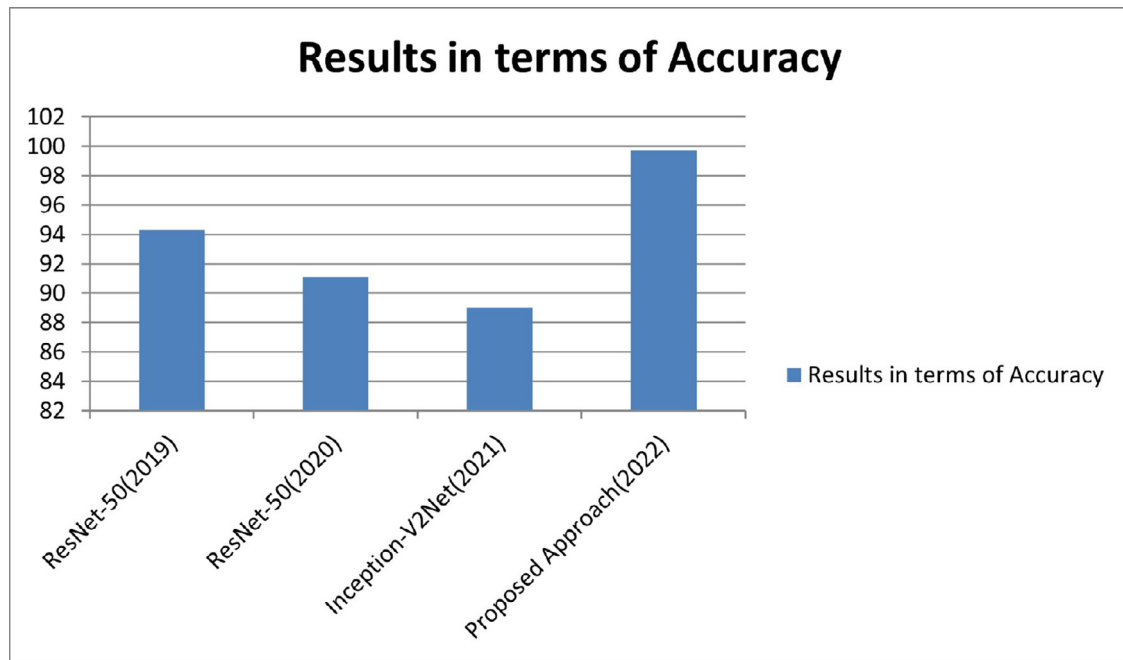
Fig. 9. Comparison with state of the art techniques.

Table 6
Confusion matrix for ResNet (multi-class).

Result Class										
	C-I	32	0	0	0	0	1	0	0	97.4
		11.8%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	
	C-II	0	33	0	1	0	1	0	0	98.17
		0.0%	12.1%	0.0%	0.4%	0.0%	0.4%	0.0%	0.0%	
	C-III	0	0	31	1	0	1	0	0	98.15
		0.0%	0.0%	11.4%	0.4%	0.0%	0.4%	0.0%	0.0%	
	C-IV	0	0	0	33	0	1	0	1	98.17
		0.0%	0.0%	0.0%	12.1%	0.0%	0.4%	0.0%	0.4%	
	C-V	0	0	0	0	33	1	0	1	98.17
		0.0%	0.0%	0.0%	0.0%	12.1%	0.4%	0.0%	0.4%	
	C-VI	1	0	0	0	0	32	0	0	97.5
		0.4%	0.0%	0.0%	0.0%	0.0%	11.8%	0.0%	0.0%	
	C-VII	1	1	1	0	0	0	33	0	98.6
		0.4%	0.4%	0.4%	0.0%	0.0%	0.0%	12.7%	0.0%	
	C-VIII	0	0	0	0	1	0	0	32	96.5
		0.0%	0.0%	0.0%	0.0%	0.4%	0.%	0.0%	11.8%	
		97.4	98.17	98.15	98.17	98.17	97.5	98.5	96.5	97.81
		C-I	**C-II**	**C-III**	**C-IV**	**C-V**	**C-VI**	**C-VII**	**C-VIII**	
					Target Class					

Table 7
Confusion matrix for InceptionV3Net (multi-class).

Result Class										
	C-I	31	0	0	1	0	1	0	0	97.4
		11.4%	0.0%	0.0%	0.4%	0.0%	0.4%	0.0%	0.0%	
	C-II	0	33	0	1	0	1	0	3	97.99
		0.0%	12.5%	0.0%	0.4%	0.0%	0.4%	0.0%	1.1%	
	C-III	0	0	32	1	0	1	0	0	91.895
		0.0%	0.0%	12.1%	0.4%	0.0%	0.4%	0.0%	0.0%	
	C-IV	0	0	0	33	0	1	0	1	91.895
		0.0%	0.0%	0.0%	12.5%	0.0%	0.4%	0.0%	0.4%	
	C-V	0	0	3	0	33	1	0	1	100
		0.0%	0.0%	1.1%	0.0%	12.5%	0.4%	0.0%	0.4%	
	C-VI	1	0	0	0	0	32	0	0	97.98
		0.4%	0.0%	0.0%	0.0%	0.0%	11.8%	0.0%	0.0%	
	C-VII	1	1	1	0	0	0	32	0	95.93
		0.4%	0.4%	0.4%	0.0%	0.0%	0.0%	11.8%	0.0%	
	C-VIII	0	0	0	0	1	0	0	33	97.99
		0.0%	0.0%	0.0%	0.0%	0.4%	0.%	0.0%	12.5%	
		100	94.89	97.99	91.895	91.895	97.99	95.92	97.99	96.07
		C-I	**C-II**	**C-III**	**C-IV**	**C-V**	**C-VI**	**C-VII**	**C-VIII**	
					Target Class					

Table 8
Confusion matrix for ShuffleNet (multi-class).

Result Class		C-I	C-II	C-III	C-IV	C-V	C-VI	C-VII	C-VIII	
	C-I	33	1	0	0	0	1	0	0	96.94
		12.1%	0.4%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	
	C-II	0	33	0	1	0	1	0	0	95.78
		0.0%	12.1%	0.0%	0.4%	0.0%	0.4%	0.0%	0.0%	
	C-III	0	0	31	1	0	1	1	0	96.84
		0.0%	0.0%	11.4%	0.4%	0.0%	0.4%	0.4%	0.0%	
	C-IV	0	1	0	32	0	1	0	1	95.70
		0.0%	0.4%	0.0%	11.8%	0.0%	0.4%	0.0%	0.4%	
	C-V	0	0	0	0	30	1	0	0	96.70
		0.0%	0.0%	0.0%	0.0%	11.0%	0.4%	0.0%	0.0%	
	C-VI	1	0	0	2	0	31	0	0	95.72
		0.4%	0.0%	0.0%	0.7%	0.0%	11.4%	0.0%	0.0%	
	C-VII	1	1	1	0	0	0	31	0	96.86
		0.4%	0.4%	0.4%	0.0%	0.0%	0.0%	11.4%	0.0%	
	C-VIII	0	0	0	0	1	0	0	32	91.8
		0.0%	0.0%	0.0%	0.0%	0.4%	0.%	0.0%	11.8%	
		96.94	95.79	96.85	95.70	96.70	95.70	96.85	91.8	95.79
		C-I	C-II	C-III	C-IV	C-V	C-VI	C-VII	C-VIII	
					Target Class					

Confusion matrices for all the eight classes have been derived and tested as well and has been considered in both cases of classification such as binary and multi class classifications. Number of images in each case is different for testing purpose and from the mentioned confusion matrices all obtained results may be viewed in a brief manner.

5. Conclusion and future works

This paper presented a computer-aided diagnosis for breast cancer image classification by using deep neural networks and transfer learning. The publicly available breast cancer images by BrakeHis have been used, consisting of 7909 images derived from the diagnoses of 82 different patients. Different image magnification factors were considered, along with the data augmentation techniques to boost the classification process. Three DNNs were used to classify breast cancer images by applying the imaging-based method. In addition, 65% of images were used for training purposes and the remaining 35% were used for testing purposes. Performance evaluation was based on different matrices such as accuracy, precision, sensitivity, and specificity. We found ResNet to be the most accurate and efficient classifier by proving the maximum average accuracies between 97.81% and 99.70% for binary and multi-class classification.

In the future, we plan to incorporate more robust datasets or patient-based image datasets to extend our model's functionality and aim to improve breast cancer detection accuracy further. We are additionally intending to stretch out our framework to make it equipped for handling more complex datasets to diagnose the breast cancer images with higher accuracies by keeping in views the state of the art techniques.

Statements of ethical approval

Not required for this study because no human or animals directly participated in this study.

Declaration of Competing interests

Authors declare that they have no conflict of interest.

Acknowledgments

References

[1] World Health Organization, WHO Position Paper on Mammography Screening, World Health Organization, 2014.

[2] S. Ahmad, S. Ur Rehman, A. Iqbal, R. K. Farooq, A. Shahid, M. I. Ullah, Breast cancer research in Pakistan: a bibliometric analysis:, 11(3) (2021). https://journals.sagepub.com/doi/full/10.1177/21582440211046934 10.1177/21582440211046934

[3] K.D. Miller, R.L. Siegel, . Chun, C. Lin, A.B. Mariotto, J.L. Kramer, J.H. Rowland, K.D. Stein, R. Alteri, A. Jemal, Cancer treatment and survivorship statistics, 2016, CA 66 (4) (2016) 271–289, doi:10.3322/CAAC.21349. https://onlinelibrary.wiley.com/doi/full/10.3322/caac.21349

[4] I. Pollanen, B. Braithwaite, T. Ikonen, H. Niska, K. Haataja, P. Toivanen, 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: 2014 4th International Conference on Image Processing Theory, Tools and Applications, IPTA 2014, 2015, doi:10.1109/IPTA.2014.7001986.

[5] C. López, M. Lejeune, R. Bosch, A. Korzynska, M. García-Rojo, M.T. Salvadó, T. Álvaro, C. Callau, A. Roso, J. Jaén, Digital image analysis in breast cancer: an example of an automated methodology and the effects of image compression, Stud. Health Technol. Inform. 179 (2012) 155–171, doi:10.3233/978-1-61499-086-4-155. https://ebooks.iospress.nl/doi/10.3233/978-1-61499-086-4-155

[6] A.P. Pandian, Identification and classification of cancer cells using capsule network with pathological images, J. Artif. Intell. Capsule Netw. 01 (01) (2019) 37–44, doi:10.36548/jaicn.2019.1.005.

[7] A. Chetlen, J. Mack, T. Chan, Breast cancer screening controversies: who, when, why, and how? Clin. Imaging 40 (2) (2016) 279–282, doi:10.1016/J.CLINIMAG.2015.05.017.

[8] A. Chekkoury, P. Khurd, J. Ni, C. Bahlmann, A. Kamen, A. Patel, L. Grady, M. Singh, M. Groher, N. Navab, E. Krupinski, J. Johnson, A. Graham, R. Weinstein, Automated malignancy detection in breast histopathological images, 8315 (2012) 332–344. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/8315/831515/Automated-malignancy-detection-in-breast-histopathological-images/10.1117/12.911643.full 10.1117/12.911643

[9] Y.M. George, H.H. Zayed, M.I. Roushdy, B.M. Elbagoury, Remote computer-aided breast cancer detection and diagnosis system based on cytological images, IEEE Syst. J. 8 (3) (2014) 949–964, doi:10.1109/JSYST.2013.2279415.

[10] Z. Jan, K.S. U., I. N., A.M. A., B. B., Automated detection of malignant cells based on structural analysis and naive Bayes classifier, 2016. https://sujo-old.usindh.edu.pk/index.php/SURJ/article/view/2348.

[11] M.T. McCann, J.A. Ozolek, C.A. Castro, B. Parvin, J. Kovačević, Automated histology analysis: opportunities for signal processing, IEEE Signal Process. Mag. 32 (1) (2015) 78–87, doi:10.1109/MSP.2014.2346443.

[12] S.M. McKinney, M. Sieniek, V. Godbole, J. Godwin, N. Antropova, H. Ashrafian, T. Back, M. Chesus, G.C. Corrado, A. Darzi, M. Etemadi, F. Garcia-Vicente, F.J. Gilbert, M. Halling-Brown, D. Hassabis, S. Jansen, A. Karthikesalingam, C.J. Kelly, D. King, J.R. Ledsam, D. Melnick, H. Mostofi, L. Peng, J.J. Reicher, B. Romera-Paredes, R. Sidebottom, M. Suleyman, D. Tse, K.C. Young, J. De Fauw, S. Shetty, International evaluation of an AI system for breast cancer screening, Nature 577 (7788) (2020) 89–94, doi:10.1038/s41586-019-1799-6. https://www.nature.com/articles/s41586-019-1799-6

[13] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, Breast cancer histopathological image classification using convolutional neural networks, in: Proceedings of the International Joint Conference on Neural Networks, 2016-October, 2016, pp. 2560–2567, doi:10.1109/IJCNN.2016.7727519.

[14] S. Sharma, R. Mehra, Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology imagesa comparative insight, J. Digit. Imaging 33 (3) (2020) 632–654, doi:10.1007/S10278-019-00307-Y/TABLES/10. https://link.springer.com/article/10.1007/s10278-019-00307-y

[15] G. Chugh, S. Kumar, N. Singh, Survey on machine learning and deep learning applications in breast cancer diagnosis, Cognit. Comput. 13 (6) (2021) 1451–1470, doi:10.1007/S12559-020-09813-6/TABLES/5. https://link.springer.com/article/10.1007/s12559-020-09813-6

[16] E.H. Houssein, M.M. Emam, A.A. Ali, P.N. Suganthan, Deep and machine learning techniques for medical imaging-based breast cancer: a comprehensive review, Expert Syst. Appl. 167 (2021) 114161, doi:10.1016/J.ESWA.2020.114161.

[17] G. Hamed, M.A.E.-R. Marey, S.E.-S. Amin, M.F. Tolba, Deep learning in breast cancer detection and classification, in: A.-E. Hassanien, A.T. Azar, T. Gaber, D. Oliva, F.M. Tolba (Eds.), Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020), Springer International Publishing, Cham, 2020, pp. 322–333.

[18] A.U. Rehman Butt, W. Ahmad, R. Ashraf, M. Asif, S.A. Cheema, Computer aided diagnosis (CAD) for segmentation and classification of burnt human skin, in: 1st International Conference on Electrical, Communication and Computer Engineering, ICECCE 2019, 2019, doi:10.1109/ICECCE47252.2019.8940758.

[19] M. TIWARI, R. Bharuka, P. Shah, R. Lokare, Breast cancer prediction using deep learning and machine learning techniques, SSRN Electron. J. (2020), doi:10.2139/SSRN.3558786.

[20] D.A. Ragab, O. Attallah, M. Sharkas, J. Ren, S. Marshall, A framework for breast cancer classification using Multi-DCNNs, Comput. Biol. Med. 131 (2021) 104245, doi:10.1016/J.COMPBIOMED.2021.104245.

[21] R. Ashraf, I. Kiran, T. Mahmood, A. Ur Rehman Butt, N. Razzaq, Z. Farooq, An efficient technique for skin cancer classification using deep learning, in: Proceedings - 2020 23rd IEEE International Multi-Topic Conference, INMIC 2020, 2020, doi:10.1109/INMIC50486.2020.9318164.

[22] S.U. Khan, N. Islam, Z. Jan, I. Ud Din, J.J. Rodrigues, A novel deep learning based framework for the detection and classification of breast cancer using transfer learning, Pattern Recognit. Lett. 125 (2019) 1–6, doi:10.1016/J.PATREC.2019.03.022.

[23] F.A. Khan, A.U.R. Butt, M. Asif, W. Ahmad, M. Nawaz, M. Jamjoom, E. Alabdulkreem, Computer-aided diagnosis for burnt skin images using deep convolutional neural network, Multimed. Tools Appl. 79 (45–46) (2020) 34545–34568, doi:10.1007/S11042-020-08768-Y/FIGURES/12. https://link.springer.com/article/10.1007/s11042-020-08768-y

[24] F.A. Khan, A.U.R. Butt, M. Asif, H. Aljuaid, A. Adnan, S. Shaheen, I. ul Haq, Burnt human skin segmentation and depth classification using deep convolutional neural network (DCNN), J. Med. Imaging Health Inform. 10 (10) (2020) 2421–2429, doi:10.1166/JMIHI.2020.3258.

[25] Z. Hameed, S. Zahia, B. Garcia-Zapirain, J.J. Aguirre, A.M. Vanegas, Breast cancer histopathology image classification using an ensemble of deep learning models, Sensors 20 (16) (2020) 4373, doi:10.3390/S20164373. https://www.mdpi.com/1424-8220/20/16/4373/

[26] P. Gupta, S. Garg, Breast cancer prediction using varying parameters of machine learning models, Procedia Comput. Sci. 171 (2020) 593–601, doi:10.1016/J.PROCS.2020.04.064.

[27] G. Murtaza, L. Shuib, A.W. Abdul Wahab, G. Mujtaba, G. Mujtaba, H.F. Nweke, M.A. Al-garadi, F. Zulfiqar, G. Raza, N.A. Azmi, Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges, Artif. Intell. Rev. 53 (3) (2020) 1655–1720, doi:10.1007/S10462-019-09716-5/FIGURES/4. https://link.springer.com/article/10.1007/s10462-019-09716-5

[28] X. Zhang, Y. Zhang, Q. Zhang, Y. Ren, T. Qiu, J. Ma, Q. Sun, Extracting comprehensive clinical information for breast cancer using deep learning methods, Int. J. Med. Inform. 132 (2019) 103985, doi:10.1016/J.IJMEDINF.2019.103985.

[29] J. Zheng, D. Lin, Z. Gao, S. Wang, M. He, J. Fan, Deep learning assisted efficient adaboost algorithm for breast cancer detection and early diagnosis, IEEE Access 8 (2020) 96946–96954, doi:10.1109/ACCESS.2020.2993536.

[30] R. Krithiga, P. Geetha, Deep learning based breast cancer detection and classification using fuzzy merging techniques, Mach. Vis. Appl. 31 (7–8) (2020) 1–18, doi:10.1007/S00138-020-01122-0/TABLES/7. https://link.springer.com/article/10.1007/s00138-020-01122-0

[31] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, A dataset for breast cancer histopathological image classification, IEEE Trans. Biomed. Eng. 63 (7) (2016) 1455–1462, doi:10.1109/TBME.2015.2496264.

[32] S. Hosseinzadeh Kassani, P. Hosseinzadeh Kassani, A comparative study of deep learning architectures on melanoma detection, Tissue Cell 58 (2019) 76–83, doi:10.1016/J.TICE.2019.04.009.

[33] A. Aloyayri, A. Krzyżak, Breast Cancer Classification from Histopathological Images Using Transfer Learning and Deep Neural Networks, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12415 LNAI, 2020, pp. 491–502, doi:10.1007/978-3-030-61401-0_45/TABLES/5. https://link.springer.com/chapter/10.1007/978-3-030-61401-0_45

[34] X. Zhang, X. Zhou, M. Lin, ShuffleNet: an extremely efficient convolutional neural network for mobile devices, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 6848–6856.

[35] Y. Gu, J. Yang, Densely-connected multi-magnification hashing for histopathological image retrieval, IEEE J. Biomed. Health Inform. 23 (4) (2019) 1683–1691, doi:10.1109/JBHI.2018.2882647.