Deep Learning Applied for Histological Diagnosis of Breast Cancer

YASIN YARI[1], THUY V. NGUYEN[1], AND HIEU T. NGUYEN[2]
[1]Department of Science and Industry Systems, University of South-Eastern Norway, 3616 Kongsberg, Norway
[2]Faculty of Information Technology, Posts and Telecommunications Institute of Technologies, Hanoi 122, Vietnam

Corresponding author: Hieu T. Nguyen (hieu.nguyen@usn.no)

ABSTRACT Deep learning, as one of the currently most popular computer science research trends, improves neural networks, which has more and deeper layers allowing higher abstraction levels and more accurate data analysis. Although deep convolutional neural networks, as a deep learning algorithm, has recently achieved promising results in data analysis, the requirement for a large amount of data prevents its use in medical data analysis since it is challenging to obtain data from the medical field. Breast cancer is a common cancer in women. To diagnose this kind of cancer, breast cell shapes in histopathology images should be examined by senior pathologists. The number of pathologists per population in the world is not enough, especially in Africa, and human mistake may occur in diagnosis procedure. After the evaluation of deep learning methods and algorithms in breast histological data processing, we tried to improve the current systems' accuracy. As a result, this study proposes two effective deep transfer learning-based models, which rely on pre-trained DCNN using a large collection of ImageNet dataset images that improve current state-of-the-art systems in both binary and multiclass classification. We transfer pre-trained weights of the ResNet50 and DesneNet121 on the Imagenet as initial weights and fine-tune these models with a deep classifier with data augmentation to detect various malignant and benign samples tissues in the two categories of binary classification and multiclass classification. The proposed models have been examined with optimized hyperparameters in magnification-dependent and magnification-independent classification modes. In the multiclass classification, the proposed system achieved up to 98% accuracy. As for binary classification, the proposed system provides up to 100% accuracy. The results outperform previous studies accuracies in all defined performance metrics in breast cancer CAD systems from histological images.

INDEX TERMS Breakhis dataset, breast cancer, CNN, computer-aided diagnosis (CAD), medical image classification, densNet, ResNet.

I. INTRODUCTION

Breast cancer (breast carcinoma) is the most common type of cancer in women, and it is the most dangerous cancer, together with lung cancer [1], [2]. Early detection of this type of cancer is crucial to reduce the mortality rate since breast cancer is often treatable when it is diagnosed early. Cancer starts from a benign state and, without appropriate treatment at the early stages, it becomes malignant. A common way to detect breast cancer is histological biopsy evaluation [3]. An Experienced pathologist evaluates breast histopathology images in various levels of magnification. Some times there is a need for complementary imageries like mammography to determine whether the sample tissue is malignant or not.

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

APPEARANCES OF CANCER

When the breast cells start to grow abnormally, breast cancer occurs. The infected cells divide more frequently than normal healthy cells and form a mass or lump. The cells may spread to the lymph nodes first, and then other parts of the human body.

HISTOPATHOLOGICAL EXAMINATION

A histological biopsy is a thorough examination of a sampled tissue under a microscope. The minimum magnification requirement for a proper diagnosis is $40\times$. However, the suspected areas have to be magnified to $100\times$, $200\times$, and $400\times$ in order to enable the pathologist to evaluate cell shape [2].

A. PROBLEM STATEMENT

The diagnosis procedure of breast cancer is operator dependent and requires an experienced pathologist. However, some
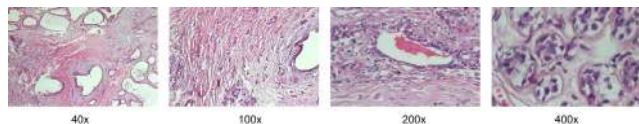
FIGURE 1. Examples of benign types of breast cancer in histopathology images in different magnifications.
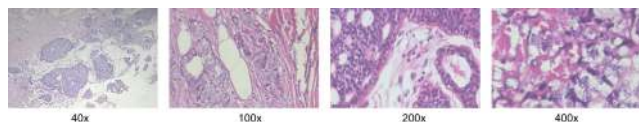

FIGURE 2. Examples of malignant types of breast cancer in histopathology images in different magnifications.

human factors like exhaustion and insufficient concentration could cause the misdetection of samples type within long and continuous procedures. In the case of misdetection, cancer may grow, and the survival rate in that condition is low. Some countries have very few pathologists per population. There is only one pathologist for every 100,000 and 130,000 in Africa and China, respectively [4]. In order to counteract the lack of experienced pathologists, the possibility of human error, the time-consuming process of screening samples, and the high cost, several Computer-Aided Diagnosis (CAD) techniques for early and automatic detection of breast cancer have been proposed and evaluated by researchers in the past [5]. These techniques can significantly help the early diagnosis of cancer. However, they are challenging to implement. Machine learning approaches nowadays are frontier in the CAD trend. With the rise of deep learning (as a part of the machine learning family), many studies have used this method in order to precisely detect samples type in histology images. However, The variety of cells size, shape, color, and scale in the histological images from one side and the complex structure of human body cells, low image quality, and similarity between benign and malignant samples from the other side can make the task challenging and prevent achieving high accuracy. Additionally, the lack of extensive and labeled datasets has also created another big challenge for the mentioned problem.

B. RELATED WORKS

HISTOPATHOLOGY DATASETS

The main source for any CAD system is both data collection and labeling within real decision-making situations by experts. There are only three datasets for breast cancer histopathological diagnosis; Mitosatypia [6], Bioimaging [7], and SSAE [8]. They are not only not fully available but also have some clinical value issues [9]. Apart from availability and clinical issues, They contain 120, 1401, and 37 images from one magnification level, respectively, which is a meager amount of data. The newest public histopathological breast cancer dataset, known as BreakHis, was released in 2016 and had the highest clinical value [5]. It has 7909 images in four magnification levels (40×, 100×, 200×, and 400×).

Following the fact that our primary research dataset is BreakHis, we reviewed the previous studies mostly on the BreakHis dataset in both binary and multiclass classification.

In the binary classification, the aim is to make an algorithm that can predict whether a sample tissue is benign or malignant. In the multiclass classification, the aim is to make an algorithm that can predict the exact subtype of tissues. These subtypes are; Adenosis (A), Fibroadenoma (F), Phyllodes Tumor(PT), and Tubular Adenoma (TA) as benign's subtypes, and Papillary Carcinoma (PC), Mucinous Carcinoma (MC), Ductal Carcinoma (DC), and Lobular Carcinoma (LC) as malignant's subtypes.

BINARY CLASSIFICATION

In the binary classification, the aim is to make an algorithm that is able to predict whether a sample tissue is benign or malignant.

MODELS

BreakHis authors were the first group that evaluated their dataset using a deep learning-based CAD system. They entrusted features extraction and classification to a deep CNN [5]. They were able to improve their results by using the AlexNet network for transfer learning. The study [10] evaluated CNN with multiple handcrafted features and compared the results with those provided with raw images. They achieved their best results by using residual blocks inspired by ResNet. Study [11] tried to find the best CNN model for this classification task and compared models such as AlexNet, ResNet, and GoogleNet. Their results are showing ResNet as the best model. The latter study also insists on the necessity of data augmentation, fine-tuning all layers, and providing large Whole Slide Images (WSI) instead of small patches. The work [9] chose the Inception-v3 model as a more efficient CNN in comparison with shallower models.

Study [12] has evaluated AlexNet and DeCAF feature extractor by transfer learning strategy. This approach enabled them to extract features from the last layer of the pre-trained AlexNet and use it for training their classifier. Authors in [13] evaluated the effect of different dimensional reduction methods on extracting features: Correlation-Based Feature Selection(CBFS), Gaussian Random Projection(GPR), and Principal Component Analysis(PCA). This study used pre-trained VGG for its purpose. Study [14] found the last convolutional layer in a model provides more important features in comparison to the final fully connected layers. Study [15] introduced a dual-stage fine-tuning that retrains a fully connected layer first and then the network thoroughly. The research [16] showed that fine-tuning on the last three layers of pre-trained AlexNet network works better than the Support Vector Machine (SVM) classification of concatenated features extracted from two pre-trained networks.

Study [17] introduced a model called deep domain knowledge-based features that mitigates the gap between the extracted features and the required specific domain that comes from using a pre-trained network on other datasets.
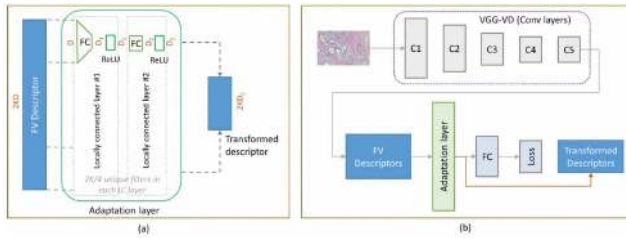
The latter study retrained the pre-trained CNN on the BreakHis dataset for efficient feature extraction. Study [18] evaluated the post encoded CNN features with Fisher Vector(FV). This study extracted a set of local features from the last convolutional layer of the model to encode local features to the FV descriptor. This approach raised the issue of high dimensionality in the data, meaning that the number of dimensions is hugely high; therefore, calculations increases in the time complexity. This issue is solved by embedding each block of the FV into a lower level dimensional feature space through a dimensionality reduction algorithm based on a multi-layer neural network model [19].

The study [20] proposed a model to capture multi-scale features by using different convolution models to ease the transition between the last convolutional and fully connected layers. Study [21] proposed a modified CNN that considers subclass information and label of each image as prior knowledge. Authors of this study claimed that their modified CNN could learn features distance better in binary classification. Research [22] introduced a CNN that is divided into a convolutional layer, pooling, and fully connected layer, and multiple hyper-parameters were tested to find the most suitable architecture for the BreakHis classification task. This new-designed convolutional model(NDCNN) was able to achieve up to 90% accuracy in binary classification. Following the fact that there may be some benign part in malignant sample images, adopting a patch-based approach becomes challenging.

Study [23] tried to overcome this challenge and introduced a Multiple Instance Learning(MIL) approach for random patches(64 × 64) extraction. Research [24] continued the MIL approach and proposed a Multiple Instance Pooling layers (MIP) to select the most distinctions feature and extract it instead of the extraction of every feature. They did it by modifying the loss function to higher activation instances.

Study [25] proposed a framework to avoid mislabeling occurs when using a patch-based approach. This framework ignores less representative samples(high confidence samples with lower entropy)and reduces the annotation cost. The remaining samples are given to the model for automatic annotation.

Authors in [26] introduced an interesting method that used the GoogleNet pre-trained model and trained the last layer in a magnification-specific way. Then for testing, all four magnification-specific-trained models aggregated by using the majority voting rule. Every four magnification-specific models prospect the test image, and if at least two of the models have the same decision, that decision is being used for the output of the network.

Research [27] fed handcrafted Tamura features instead of raw image to a Deep Belief Network(DBN) that consists of four Restricted Boltzmann Machines(RBM). Based on the fact that autoencoders have shown interesting results in image classification [28], the study [29] proposed a framework that has a Landmark ISOMAP(L-ISOMAP) for extracting features of histology images. This method was able to obtain high accuracy.

PREPROCESSING METHODS
Some studies decided to utilize a prepossessing technique on raw data to improve their results. Study [30] evaluated the importance of data augmentation on 40× magnification images on the result. Study [31] evaluated the effect of cluster-transformed images with the use of different clustering algorithms and compared them with the same CNN with raw images input.

Study [32] proposed a k-mean clustering on images to highlight nuclei segmentation as preprocessing. This study used a Discrete Wavelet Transform(DWT) to extract features from cluster-transformed images. Support Vector Machine(SVM) is chosen as a classifier with these features in the latter study.

Research [37] worked on color texture variation of histopathology images. This study evaluated the performance of multiple color-texture descriptors with different classifiers. The study did this step for each of the four magnification levels and made an integrated model as a magnification-independent model. Research [39] decided to remove normalizing to help model learn color-texture variabilities. This study found that performing the grayscale transformation, as a stain normalization method, decreases the accuracy of the results. Study [52] claims that conventional normalization techniques increase the noise in the image and introduce a new normalization technique that controls the noise.

MULTICLASS CLASSIFICATION
In the multiclass classification, the aim is to make an algorithm that is able to predict the exact type of sample tissues from other types. The study [56] used K-Means and autoencoder approached for image clustering and successfully classified histopathology images using InceptionResNet2. Moreover, this study evaluated the effect of the data augmentation method and was able to obtain 95.3% accuracy for its best results.

Study [57] worked on obtaining two patches (patch sampling) to prevent information loss. A CNN and K-means algorithm were combined for this purpose. With the use of ResNet50, first, essential features were extracted, and then the model was tested, and 95% of accuracy for four breast cancer types was achieved.

TABLE 1. Classification models and results [9].

							BINARY CLASSIFICATION			
Work	Preprocessing	Patch/Slide	Features extractor	Transfer Learning	Training/Test	Metrics	40X	100X	200X	400X
[33]	None	WSI	PFTAS	None	70% / 30%	PLA	83.8	82.1	84.2	82.0
[5]	SMI	RND(32*32, 64*64))	AlexNet	ImageNet	70% / 30%	F1	85.6	83.5	84.6	86.1
[34]	Binarization	WSI	FD	None	50% / 50%	ILA	97.9	16.5	16.5	25.3
[12]	None	MKV	CaffeNet	ImageNet	70% / 30%	ILA	84.6	84.8	84.2	81.6
[35]	DAB	WSI	NDCNN	ImageNet	50% / 50%	ILA	95.8	96.9	96.7	94.9
[18]	None	WSI	VGG	ImageNet	70% / 30%	ILA	87.0	86.2	85.2	89.2
[36]	None	WSI	DR	None	70% / 30%	PLA	94.9	93.6	94.5	94.4
[26]	Res	Rnd(224*224)	GoogleNet	ImageNet	80% / 20%	ILA	94.8	94.3	94.6	93.4
[21]	SMI	WSI	NDCNN	ImageNet	75% / 25%	ILA	97.8	97.6	97.5	97.9
[19]	None	WSI	ConvNet	ImageNet	70% / 30%	ILA	87.7	87.6	86.5	83.9
[37]	None	WSI	Integrated	None	70% / 30%	ILA		88.0		
[20]	None	P(228*228)	NDCNN	None	Not specified	ILA	82.7	None	None	None
[38]	None	None	PWT	None	75% / 25%	ILA	None	None	None	85.6
[39]	RGBT	WSI	JCTF	None	70% / 30%	PLA	86.8	88.4	88.8	87.5
[40]	Res(350*230)	WSI	NDCNN	None	70% / 30%	ILA	77.5	None	None	None
[41]	ETB	WSI	PFTAS	None	Not specified	ILA	81.7	81.2	80.7	81.5
[30]	DA	WSI	Inception v3	ImageNet	70% / 30%	ILA	86.0	None	None	None
[15]	DA	Rnd(224*224)	VGGNet	ImageNet	80% / 20%	ILA	91.2	91.4	88.5	84.5
[11]	Res(350*230)	WSI	ResNet	ImageNet	Not specified	PLA		95.0		
[13]	Res(224*224)	WSI	ResNet	ImageNet	75% / 25%	ILA	84.0	88.2	87.0	80.3
[42]	None	SQ	BE	None	70% / 30%	ILA	47.0	40.0	40.0	37.0
[43]	MVD	WSI	DR	None	70% / 30%	PLA	87.7	85.8	88.0	84.6
[44]	None	Rnd(64*64)	NDCNN	Camelyon	75% / 25%	ILA	96.1	None	None	None
[10]	None	WSI	NDCNN	None	Not specified	ILA	94.4	95.9	97.1	96.0
[23]	None	Rnd(64*64)	PFTAS	None	70% / 30%	ILA	87.8	85.6	80.8	82.9
[17]	Res(224*224)	WSI	ResNet152	ImageNet	80% / 20%	ILA		81.2		
[31]	KM	WSI	NDCNN	None	Not specified	ILA	85	90	90	90
[45]	GSC	WSI	HI	None	70% / 30%	ILA	85.9	80.4	78.1	71.1
[46]	SMI	WSI	Inception v3	ImageNet	70% / 30%	ILA	86.5	83.2	85.4	80.3
[32]	KM	WSI	DWT	None	Not specified	ILA		93.3		
[14]	Res(224*224)	WSI	Densenet	ImageNet	70% / 30%	PLA	94.7	95.9	96.7	89.1
[29]	GSC	WSI	DR	None	Not specified	ILA	96.8	98.1	98.2	97.5
[47]	Res(341*224)	SW	ResNet	ImageNet	70% / 30%	ILA	98.6	97.9	98.3	97.6
[25]	None	WSI	AlexNet	ImageNet	70% / 30%	ILA	90.6	90.4	90.6	90.9
[16]	Res(227*227)	WSI	AlexNet	ImageNet	Not specified	ILA	90.6	90.5	91.3	91.3
[48]	None	WSI	DR	None	70% / 30%	PLA	89.1	87.3	88.4	86.6
[49]	None	WSI	LPQ	None	70% / 30%	ILA	91.1	90.7	86.2	84.3
[27]	CE	WSI	Tam	None	70% / 30%	ILA	88.7	85.3	88.6	88.4
[24]	Res(370*230)	P(224*224)	NDCNN	None	80% / 20%	ILA	89.5	89.0	88.8	87.6
[22]	None	P(64*64)	NDCNN	None	70% / 30%	ILA	82.0	86.2	84.6	84.0
[50]	Res(224*224)	WSI	VGG19	ImageNet	98% / 2%	ILA		80.0		
[51]	DA	WSI	NDCNN	None	70% / 30%	ILA	98.3	97.1	97.8	96.1
[52]	SN	WSI	HI	None	85% / 15%	ILA	95.0	96.6	93.5	94.2
[53]	Res(460*460)	Rnd(100*100)	NDCNN	None	70% / 30%	PLA	83.0	83.1	84.6	82.1
[54]	None	WSI	JCTF	None	70% / 30%	PLA	87.2	88.2	88.8	85.8
[9]	SN and DA	WSI	ResNet	ImageNet	70% / 30%	ILA		88.9		
				MULTICLASS CLASSIFICATION						
[34]	Bin	WSI	FD	None	50% / 50%	ILA		55.6		
[35]	DAB	WSI	NDCNN	ImageNet	50% / 50%	ILA	92.8	93.9	93.7	92.9
[47]	Res(341*224)	SW(224*224)	ResNet	Im-Break	70% / 30%	ILA	95.6	94.8	95.6	94.6
[51]	DA	WSI	NDCNN	None	70% / 30%	ILA	88.2	84.6	83.1	83.9
[56]	SN	WSI	ResNet	ImageNet	80% / 20%	ILA		95.0		
[9]	SN and DA	WSI	ResNet	ImageNet	70% / 30%	ILA		63.6		

Further study [58] elevated the performance of classification with the use of a DCNN with the gradient boosting classifier. It used inception-300 × 300+GBT and achieved 93.5, 95.3, 96.1 and 91.1 accuracy on the 40X, 100X, 200X, and 400X images, respectively.

In [27], an accuracy of 91% was achieved by the use of a CNN with a mean-shift algorithm. Similarly, paper [59] generated higher dimension features for improving image classification accuracy. It achieved 87% accuracy.

Research [60] has used the nucleus-guided training of CNN to diminish the noise, which is coming from stroma to increase the CNN classification, together with reducing computational time. In [35], classification accuracy was increased with using ReLU to reduce vanishing gradient possibility. It has used filters with various sizes of 3 × 3, 5 × 5, and 7 × 7.

The proposed model obtained 93.2% accuracy in multiclass classification.

In [61], a Nottingham Grading System (NGS) was proposed to differentiate images into three subclasses. The study [62] focused on histopathological image classification limitations and successfully proposed an Enhanced Loss Function (ELF) method to increase the classifier's performance. It also shows that the ELF increases the accuracy of classification for 3% and duration of processing time up to 30-40 seconds.

In [63], two methods for the classification of histology images have been proposed. The first method is reliant on handcrafted features in which Hu moment, color histogram, and Haralick texture are used for the extraction of features from images of the BreakHis dataset. The extracted-features were then utilized for training the classifier. The second method is using transfer learning with the use of pre-trained ResNet50, VGG16, and VGG19. This research achieved its best result by transfer learning of the VGG16 network with a linear Support Vector Machine (SVM). The accuracy of the results has computed in two forms, (a) patch-based accuracies (93.97% for 40×, 92.92% for 100×, 91.23% for 200×, and 91.79% for 400×); (b) patient-based accuracies (93.25% for 40×, 91.87% for 100×, 91.5% for 200×, and 92.31% for 400×) in the magnification-dependent category of histology image classification. This study also claims that Fibro Adenoma and Mucous Carcinoma classes are the most difficult to classify.

In [64], a class Kernel Principle Component Analysis (KPCA) is introduced for extracting features, and then for each extracted feature, a KPCA model is trained. This process is repeated for all the images in the dataset, and finally, a pre-trained KPCA model makes the decision. This approach achieved 92% accuracy in binary classification.

Research [5] has used a pre-trained AlexNet for binary classification on the BreakHis dataset. The pre-trained AlexNet uses sliding windows and random extraction techniques.

In [12], linear regression and pre-trained CaffeNet was used to extract features and classification. Study [35] used a deep convolutional network that is able to learn discerning features. It achieved an accuracy of 92-95% on various magnification-level images in the BreakHis dataset.

C. MOTIVATION

Recent researches have shown that deep learning methods, particularly Convolutional Neural Networks (CNN), are excessively effective for image analysis [12]. In fact, CNN has become the frontier machine learning tool for computer vision and image analysis. The recent breakthrough in deep learning shows great potential to increase the performance of applications. This advancement gravitated our attention and curiosity to investigate and develop an efficient method based on deep learning to help to solve a real-world problem in medical data analysis. This work focuses on the automatic detection of breast cancer in histopathology images.

D. CONTRIBUTIONS

Give the above, this work aims to:

- Develop a high accuracy method by improving previous works' accuracies.
- Minimize false-negative results.
- Experiment image magnification-dependent and independent approaches.
- Define the exact sub-type of the sample tissues.

Achieving a high accuracy result helps the feasibility of CAD systems for breast cancer recognition in medical practice.

E. APPROACH

In order to achieve the goals, a bottom-up development approach is considered, and during each subsystem development and integration, the Agile SCRUM methodology has been chosen. The reason for choosing this methodology is the type of project which is within software development. SCRUM enables us to obtain the highest efficiency through weekly sprints [65]. In each sprint, the work that has been done in the previous week will be reviewed, and then a new sprint will be defined for the following week. The subtasks are:

- Survey the literature related to the topic.
- Evaluate image classification algorithms and state-of-the-art deep learning methods.
- Choose the most suitable algorithm for image recognition and design a new CAD system based on the structure of the efficient algorithms for our specific problem.
- Develop a preprocessing technique for the preparation of the dataset and experiments.
- Define a set of hyperparameters and optimize them for our specific problem.
- Implement the chosen models with the optimized preprocessing technique and hyperparameters for breast cancer detection.
- Perform test and analysis of the chosen models to achieve the highest accuracy.
- Conclude the study and suggest future work.

II. SYSTEM MODEL
A. TRANSFER LEARNING

The BreakHis dataset is relatively small to train a network from scratch and achieve high accuracy results. A way to alleviate this problem is using transfer learning and fine-tuning a pre-trained CNN [66]. A pre-trained CNN is trained on an extensive dataset of various domains. The use of pre-trained networks is widespread nowadays in computer vision tasks [67]. Both ResNet and DenseNet are being used in this research.

When training a plain network with a standard optimization algorithm, as the number of layers increases, the training error decreases in the beginning, but then it increases. In other words, the training error is getting worse when picking too deep networks. If we want a highly accurate model, we have to have a deep neural network. A deep model can extract

features better than shallow models by using the intermediate hidden layers [68].

The ResNet50 network is made of residual blocks. In traditional (plain) neural networks, every layer is connected to the next layer. In a network with the residual building blocks, every block is connected to the next layer, but it is also directly connected into the layer, which is 2-3 layers deeper, as a shortcut connection. Figure 4 represents a residual block. Input x is passed through a few convolutional layers (function f), and the result will become $f(x)$. Although a traditional CNN has the same logic, the ResNet then adds the original input x to the result ($f(x)$), and becomes $f(x) + x$. This addition is element-wise addition ($\oplus$). The information in x can follow a shortcut to go much deeper into the neural network. So using residual block allows us to train a much deeper network. When CNN goes deeper, the path for information from the input layer to the output becomes larger. This also happens for gradient descent in the opposite direction, and it vanishes before reaching the other side of the network. A ResNet network can be made by taking many residual blocks and stacking them together to form a network.
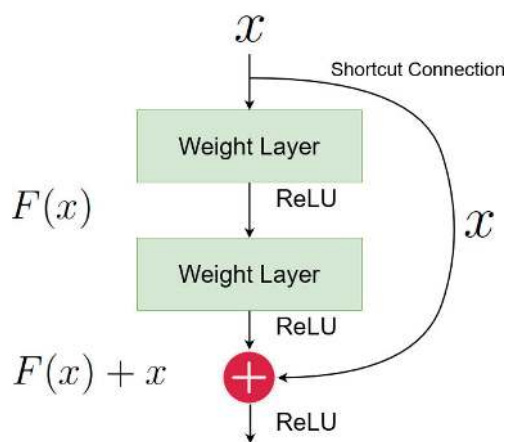

FIGURE 4. A residual block.

By taking the intermediate activations and letting them go deeper into the neural network, it helps with vanishing gradient problems and allows us to train a much deeper neural network without loss of performance [69], [70].

Resnet50 has five stages, and each of these stages has a residual block plus a convolutional block. Each residual block also has three convolutional layers, and each convolutional block also has three convolutional layers. ResNet50 has approximately 23 million parameters for training.

Having described ResNet, in DenseNet, each layer has additional inputs from all preceding layers (each layer gets information directly from all previous layers). This makes the neural network thin and compact(fewer channels). We have a copy of all the previous layers in the current layer (previous layers are concatenated into the current layer). DenseNet simplifies the connectivity pattern among layers in ResNet to ensure maximum information flow and needs
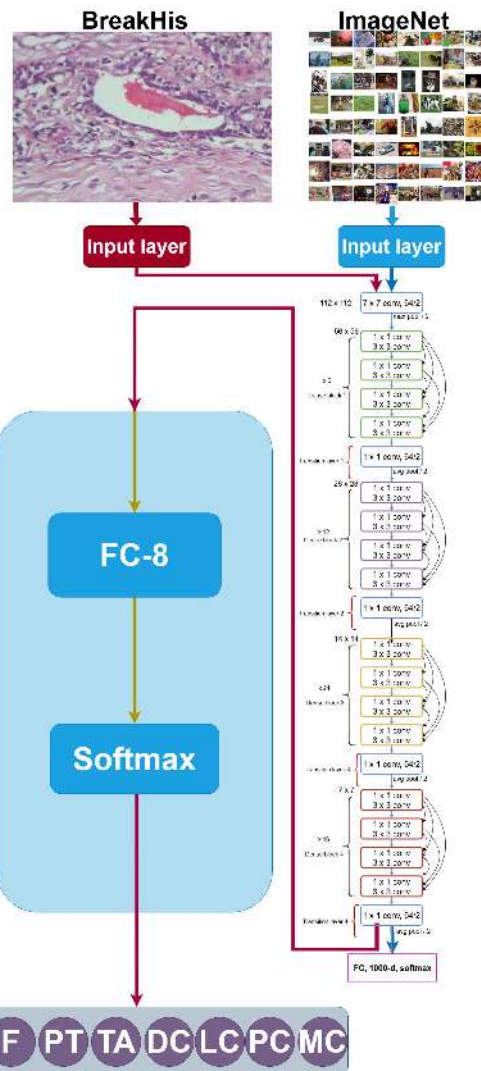

FIGURE 5. Densenet model's structure.

fewer parameters than other CNNs, and there is no place for redundant feature maps. However, its layers are narrow, and they just add a few feature maps. DenseNet has a feature layer(convolutional layer), multiple dense blocks (the concatenated layers), and a few transition layers among dense blocks [71].

DenseNet achieves similar accuracy as ResNet with less than half the amount of parameters [72]. DenseNet121 version has been chosen among other DenseNet versions, because it has around 1 million parameters which are less than the number of parameters in DenseNet196, DenseNet201, DenseNet264.

In order to transfer learn and fine-tune, we make a new layer to replace it with the fully connected layer (FC-1000) in the resNet50 and DensNet121 architectures. The new top layer has a fully connected layer (FC-8). Then we transfer the weights and fine-tune the model with a new top layer. It is done by training and backpropagation on the resnet50 and DenseNet121 with the patch-balanced dataset. Figure 5 presents our proposed model's architecture for

DenseNet121. The ResNet50 is also utilized in the same structure.

Apart from DensNet and ResNet, we utilized some other pre-trained networks to compare the results. Those networks are ResNet101, VGG19, AlexNet, and SqueezeNet.

III. DATASET PREPARATION
A. DATASET PARTITIONING
Following the fact that the BreakHis dataset is not substantially large, the partitioning process is vital to fetch the most out of the proposed model. We have divided the dataset into train, validation, and test set with 6011, 1492, and 406 images, respectively. This ratio has been chosen to increase training efficiency as much as possible. Figure 2 shows BreakHis dataset partitioning in detail.

B. DATA AUGMENTATION
In order to increase the diversity in the BreakHis dataset and boost the CAD systems' performance, a data augmentation method has been implemented. Every image in the training set is first resized to 224 × 224 pixels. Then some of the images horizontally flipped randomly. We also use the color jitter for images. It changes the tone of the original color based on Hue, Saturation, and Value (HSV). Some of the training set images are also randomly rotated and cropped. After these steps, the image is transformed into tensors (matrices of numbers) and get normalized. Figure 7 shows a batch of augmented data. As for the validation set, all images are just getting normalized without any flipping, cropping, or rotation. The test set is given to the trained model just after resizing the image to the models required input size without any change or edition (raw image).

IV. IMPLEMENTATION
A. HARDWARE AND SOFTWARE
The proposed model has been implemented, and the test results were produced with a desktop with AMD Ryzen Threadripper 1950X 16-Core Processor 3.40 GHz, 128 GB Ram, and NVIDIA 1080ti GPU has used.

As for software, we have used PyTorch in the Jupyter Notebook of Anaconda environment. The implementation code of the proposed model is available in the GitHub Repository.

B. HYPERPARAMETERS OPTIMIZATION AND SETTINGS
We define and tuned a set of hyperparameters and settings for our specific task. Table 3 shows these optimized hyperparameters and settings. Setting-3 and setting-4 are the best ones.

C. EVALUATION METRICS
In order to measure the accuracy of breast cancer CAD systems based on the chosen taxonomies, there are some metrics for fair comparison among different CAD systems.

1) IMAGE-LEVEL ACCURACY (ILA)
The total number of correctly classified images divided by the total number of images provides image-level accuracy, which is considered as the main metric for our study. Equation 1 shows the way to calculate ILA.

$$ILA = \frac{I_{correct}}{I_{total}} \tag{1}$$

2) PRECISION (POSITIVE PREDICTIVE VALUE)
The fraction of relevant samples among the retrieved samples. Equation 2 represents the formula for calculation of precision.

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives} \tag{2}$$

3) RECALL (Sensitivity)
The fraction of the total amount of relevant samples which were actually retrieved. Equation 3 shows the formula for the calculation of recall.

$$Recall = Sensitivity = \frac{TruePositives}{TruePositives + FalseNegatives} \tag{3}$$

4) F1 SCORE
As a measurement of test accuracy, the F1 score is the harmonic mean of precision and recall. Equation 4 shows the formula for the calculation of the F1 score.

$$F1_{score} = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{4}$$

V. RESULTS
We have done a set of various experiments with the most promising set of hyperparameters shown in table 3 in both binary and multiclass classification. The models have been tested with different settings to achieve the highest accuracy.

A. BINARY CLASSIFICATION
We have conducted different models with optimized hyperparameters to predict whether the sample tissues are benign or malignant. We first tested our different models in a magnification-dependent way on 40×, 100×, 200×, and 400× set of images separately.

MAGNIFICATION DEPENDENT- 40X
We tested two optimized hyperparameter settings (setting 3, setting 4) on our presented modified CNNs (2-layers designed from scratch, pre-trained ResNet50, and pre-trained DenseNet121) for our specific problem. Table 4 presents the models and achieved accuracies. As can be seen, model-3 and model-4 obtained 98.98% and 100% ILA accuracy, respectively. These results are improving state-of-the-art results on both the BreakHis dataset and every automatic breast cancer recognition studies. The result of model-4 was confirmed by repeating the experiments. The code for these experiments is available in GitHub CAD-40X-Binary Repository.
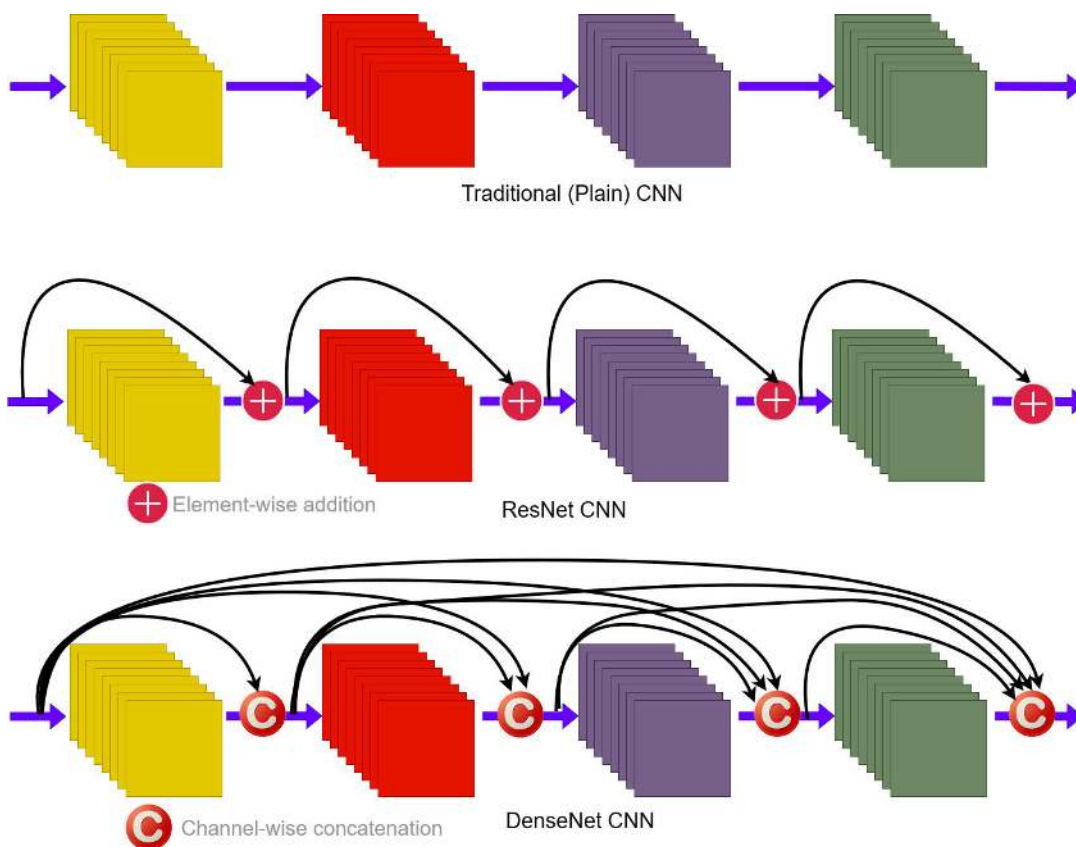
FIGURE 6. Plain, ResNet, and DensNet networks concepts.

TABLE 2. BreakHis dataset partitioning ratio in detail.

Main class		Benign				Total benign	Malignant				Total malignant	Total of both
Subclass		A	F	TA	PT		DC	LC	MC	PC		
Training set	40x	84	194	112	84	474	659	120	156	112	1047	1521
	100x	82	200	113	93	488	686	131	170	108	1095	1583
	200x	79	202	107	83	471	681	125	149	103	1058	1529
	400x	75	181	98	88	442	597	105	128	106	936	1378
Total of training set		320	777	430	348	1875	2623	481	603	429	4136	6011
Validation set	40x	21	48	28	20	117	164	29	39	27	259	376
	100x	20	49	28	23	120	172	32	42	27	273	393
	200x	20	50	26	20	116	171	31	37	25	264	380
	400x	19	45	24	22	110	149	26	32	26	233	343
Total of validation set		80	192	106	85	463	656	118	150	105	1029	1492
Test set	40x	9	11	9	6	34	41	7	10	6	64	98
	100x	11	11	9	7	36	45	7	10	7	69	105
	200x	12	12	7	7	36	44	7	10	7	68	104
	400x	12	11	8	6	36	42	6	9	6	63	99
Total of testing set		44	45	33	26	142	172	27	39	26	264	406

MAGNIFICATION DEPENDENT- 100X

The experiments for 40× images have been repeated for 100× magnification level images with optimized CNNs, hyperparameters, and data augmentations. Table 5 represents the results and settings. Model-4 has obtained 100% accuracy with setting 3. The result of the model-3 and model-4 also improving state-of-the-art results in beast cancer CAD

systems. The code for these experiments is available in GitHub CAD-100X-Binary Repository.

MAGNIFICATION DEPENDENT- 200X

The proposed models have experimented with the same approach as 40× and 100× experiments. The results are slightly lower than 40× and 100×. Model-4 achieved an
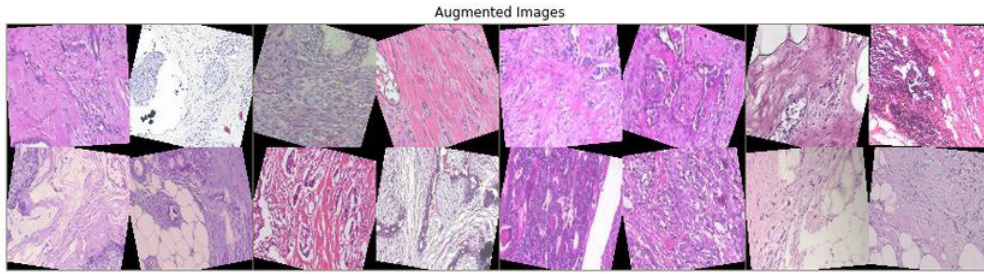
FIGURE 7. A batch of augmented data from BreakHis dataset.

TABLE 3. Different tested settings.

Hyper Parameters	Setting Range		Examples					
	From	To	Setting 1	Setting 2	Setting 3	Setting 4	Setting 5	Setting 6
α	0.00002	0.2	0.00002	0.002	0.002	0.02	0.02	0.2
γ	0.1	0.9	0.1	0.9	0.9	0.9	0.9	0.9
momentum	0.1	0.9	0.1	0.9	0.9	0.9	0.9	0.9
stepsize	1	2	2	1	2	2	2	2
batchsize	4	64	4	16	16	16	32	64
inputsize	100×100	229×229	100×100	224×224	224×224	224×224	224×224	229×229
epochs	50	200	50	100	200	200	150	200
poolingsize	2×2	7×7	2×2	5×5	7×7	7×7	7×7	7×7
Loss (lowest)	-	-	0.5478	0.0285	0.0000	0.0000	0.0001	0.0090
Accuracy (highest)	-	-	61.19%	91.55%	93.70%	92.16%	90.48%	84.65%

TABLE 4. Different tested models on 40× magnification images of BreakHis dataset.

Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121
CNN layers	2	2	40	117	40	117
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4
Data Augmentation	-	✓	✓	✓	✓	✓
Results						
Misclassified images	20	18	1	0	3	2
Accuracy(ILA)	79.59%	81.63%	98.98%	**100%**	96.94%	97.97%
Precision	78.20%	82.85%	98.38%	**100%**	95.31%	96.82%
Recall	95.31%	90.62%	**100%**	**100%**	**100%**	**100%**
F1$_{score}$	85.91%	86.57%	99.18%	**100%**	97.59%	98.38%

TABLE 5. Different tested models on 100× magnification images of BreakHis dataset.

Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121
CNN layers	2	2	40	117	40	117
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4
Data Augmentation	-	✓	✓	✓	✓	✓
Results						
Misclassified images	31	17	1	0	2	1
Accuracy(ILA)	69.52%	83.81%	99.05%	**100%**	98.10%	99.05%
Precision	73.17%	85.13%	98.57%	**100%**	98.55%	98.57%
Recall	86.95%	91.30%	**100%**	**100%**	98.55%	**100%**
F1$_{score}$	79.46%	88.10%	99.27%	**100%**	98.55%	99.27%

accuracy of 98.08%, which is the highest result in the 200× experiments. The code for these experiments is available in GitHub CAD-200X-Binary Repository.

MAGNIFICATION DEPENDENT- 400X

As for the 400X class, the tested models and settings are shown in table 7. The accuracy of 98.99% was achieved with model-4 and was confirmed by repeating the experiments.

The code for these experiments is available in GitHub CAD-400X-Binary Repository.

MAGNIFICATION INDEPENDENT- 40×, 100×, 200×, AND 400X

The proposed CNNs are trained and tested with all of the BreakHis dataset images regardless of their magnifications. Table 8 represents the models, settings, and accuracy of the

TABLE 6. Different tested models on 200× magnification images of BreakHis dataset.

Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121
CNN layers	2	2	40	117	40	117
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4
Data Augmentation	-	✓	✓	✓	✓	✓
Results						
Misclassified images	19	17	3	2	4	3
Accuracy(ILA)	81.73%	83.65%	97.12%	**98.08**%	96.15%	97.12%
Precision	84.50%	85.91%	95.77%	**97.14**%	95.71%	95.77%
Recall	88.23%	89.70%	**100**%	**100**%	98.52%	**100**%
F1$_{score}$	86.32%	87.76%	97.83%	**98.54**%	97.09%	97.74%

TABLE 7. Different tested models on 400× magnification images of BreakHis dataset.

Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121
CNN layers	2	2	40	117	40	117
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4
Data Augmentation	-	✓	✓	✓	✓	✓
Results						
Misclassified images	19	12	3	1	5	4
Accuracy(ILA)	80.81%	87.88%	96.97%	**98.99**%	94.95%	95.96%
Precision	81.42%	88.05%	95.45%	**98.43**%	93.93%	94.02%
Recall	90.47%	93.65%	**100**%	**100**%	98.41%	**100**%
F1$_{score}$	85.70%	90.76%	97.67%	**99.20**%	96.11%	96.91%

results. The model-3 achieved an accuracy of 99.26%, which improves state-of-the-art results. We also have used more pre-trained models in this section to compare more results. The code for these experiments is available in GitHub CAD-Magnification-Independent-Binary Repository.

B. MULTICLASS CLASSIFICATION

The second set of experiments have been done in multiclass classification type. First, the experiment for each magnification group has been performed and, then the models have been tested in magnification independent way.

MAGNIFICATION DEPENDENT- 40X

Table 9 represents the settings and models that have been experimented. The highest accuracies come from model-3 and model-4 (same as binary classification). However, in multiclass classification, model-3 has achieved higher accuracy than model-4. These results are also improving state-of-the-art results in CAD for breast cancer. The code for these experiments is available in GitHub CAD-40X-Multiclass Repository.

MAGNIFICATION DEPENDENT- 100X

Table 10 presents the results of the experiments on the 100× magnification images. Model-3 achieved the highest accuracy, with 97.14% ILA. The code for these experiments is available in GitHub CAD-40X-Multiclass Repository.

MAGNIFICATION DEPENDENT- 200X

Table 11 shows the tested models on 200× set of magnifications. Model 3 obtained 95.19% accuracy, which is the highest accuracy among others. The code for these

experiments is available in GitHub CAD-40X-Multiclass Repository.

MAGNIFICATION DEPENDENT- 400X

Table 12 exhibits the results of applying our proposed models on the 400× set of magnifications. The code for these experiments is available in GitHub CAD-40X-Multiclass Repository.

MAGNIFICATION INDEPENDENT- 40×, 100×, 200×, AND 400X

The proposed models have been experimented in a magnification independent way, and the results are shown in table 13. The code for these experiments is available in GitHub CAD-Magnification-Independent-Multiclass Repository.

C. EVALUATION OF RESULTS

The results of the experiments are promising. After a closer look at the misclassified images, it is revealed that they are more and less the same in most of the models, meaning the model is working quite well, and the dataset is not broad enough to increase the variety of learning. The confusion matrix is provided for the top two models in the magnification independent multiclass classification category. It shows the Lobular Carcinoma (LC) (malignant) is the most difficult tissue for classification. This difficulty comes from its very complicated cell structure in comparison to other malignant tissues [73]. Although the cell structure is more tricky in LC, there are only 626 LC images available in the dataset, which is extremely low for making a robust classifier for this category. Interestingly, most of the mislabeled images are predicted as Ductal Carcinoma (DC). This comes from a

TABLE 8. Different tested models on all of the images of BreakHis dataset.

Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121	vgg19	alexnet	squeezenet1
CNN layers	2	2	40	117	40	117	16	5	2
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4	setting 3	setting 3	setting 3
Data Augmentation	-	✓	✓	✓	✓	✓	✓	✓	✓
Results									
Misclassified images	46	40	3	5	4	5	42	142	33
Accuracy(ILA)	88.66%	90.15%	**99.26**%	98.77%	99.01%	98.77%	84.73 %	34.98%	91.63%
Precision	78.66%	87.32%	**98.87**%	98.14%	98.50%	98.50%	90.80%	68.94%	91.69%
Recall	95.93%	96.59%	**100**%	**100**%	**100**%	99.62%	93.56%	84.09%	96.21%
F1$_{score}$	86.44%	91.72%	**99.43**%	99.06%	99.24%	96.03%	92.15%	75.76%	96.19%

TABLE 9. Different tested models on 40× magnification images of BreakHis dataset in multiclass category.

Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121
CNN layers	2	2	40	117	40	117
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4
Data Augmentation	-	✓	✓	✓	✓	✓
Results						
Misclassified images	34	32	2	3	4	3
Accuracy(ILA)	54.08%	62.24%	**97.96**%	96.94%	95.92%	96.94%
Precision	69.73%	71.05%	**98.43**%	96.92%	95.45%	96.92%
Recall	82.81%	84.37%	**98.43**%	98.43%	98.43%	98.43%
F1$_{score}$	75.70%	77.13%	**98.43**%	97.66%	96.91%	97.66%

TABLE 10. Different tested models on 100× magnification images of BreakHis dataset in multiclass category.

Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121
CNN layers	2	2	40	117	40	117
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4
Data Augmentation	-	✓	✓	✓	✓	✓
Results						
Misclassified images	34	33	5	3	8	10
Accuracy(ILA)	67.61%	68.57%	95.24%	**97.14**%	92.38%	90.48%
Precision	69.66%	70.00%	97.05%	**97.14**%	94.20%	90.41%
Recall	89.85%	91.30%	95.65%	**98.55**%	94.20%	95.65%
F1$_{score}$	78.47%	79.24%	96.34%	**97.83**%	94.20%	92.95%

TABLE 11. Different tested models on 200× magnification images of BreakHis dataset in multiclass category.

Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121
CNN layers	2	2	40	117	40	117
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4
Data Augmentation	-	✓	✓	✓	✓	✓
Results						
Misclassified images	34	33	5	8	8	9
Accuracy(ILA)	67.30%	68.26%	**95.19**%	92.31%	92.31%	91.35%
Precision	71.79%	71.08%	**93.15**%	89.47%	90.54%	89.33%
Recall	82.35%	86.76%	**100**%	**100**%	98.52%	98.52%
F1$_{score}$	76.70%	78.14%	**96.45**%	94.44%	94.36%	93.70%

large number of images for ductal carcinoma in the BreakHis dataset and the similarity of malignant sample tissues. Our proposed models are solving the difficulty of classification in Fibro Adenoma (FA) and Mucous Carcinoma (MC) classes, which study [63] has mentioned. The misclassified images are mostly benign samples that are predicted as malignant. Although this is an anomaly, there are meager cases of false-negative prediction (malignant samples that are predicted as benign by the proposed models), which is the worst condition for a CAD system.

The evaluation of all the experiments is showing the Model-4 with setting-3 is the best model for binary classification. In the multiclass classification, the Model-3 with setting-3 has outperformed other models. The models are designed to be flexible, meaning it is possible to combine both model-3 and model-4 as a bigger model to boost our CAD system's performance. Figure 9 visualizes the average performance of our pre-trained models.

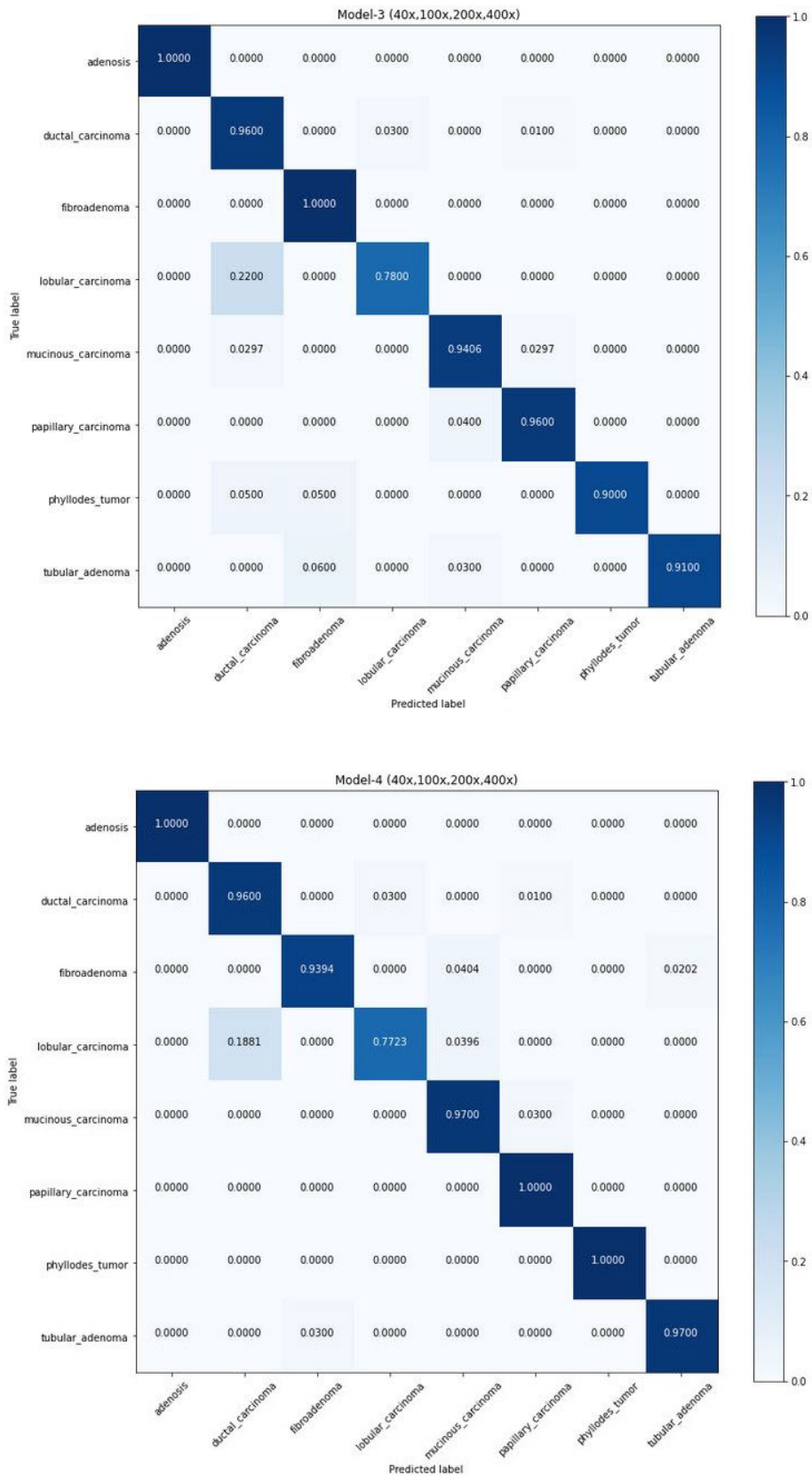As can be seen in the figure 9, the Densnet121 based models (model-4 and model-6) work slightly better than

FIGURE 8. Confusion matrix of magnification independent model-3 and model-4 on the BreakHis dataset in multiclass category.

TABLE 12. Different tested models on 400× magnification images of BreakHis dataset in multiclass category.

Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121
CNN layers	2	2	40	117	40	117
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4
Data Augmentation	-	✓	✓	✓	✓	✓
Results						
Misclassified images	34	34	5	6	7	9
Accuracy(ILA)	65.66%	65.66%	**94.95%**	93.94%	92.93%	90.91%
Precision	71.01%	71.64%	**92.64%**	91.30%	90.00%	87.50%
Recall	77.77%	76.19%	**100%**	**100%**	**100%**	**100%**
F1$_{score}$	74.23%	73.84%	**96.17%**	95.45%	94.73%	93.33%

TABLE 13. Different tested models on all of the images of BreakHis dataset.

| Model details	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Transfer learning	-	-	ResNet50	DenseNet121	ResNet50	Densenet121	ResNet101	alexnet	squeezenet1
CNN layers	2	2	48	117	48	117	99	5	2
setting	setting 3	setting 3	setting 3	setting 3	setting 4	setting 4	setting 3	setting 3	setting 3
Data Augmentation	-	✓	✓	✓	✓	✓	✓	✓	✓
Results									
Misclassified images	197	159	21	18	23	26	23	172	49
Accuracy(ILA)	51.47%	60.84%	94.83%	**95.57%**	94.33%	93.60%	94.33 %	42.36%	87.93%
Precision	59.17%	64.95%	93.54%	**94.24%**	92.57%	91.60%	92.57%	62.92%	91.82%
Recall	81.81%	86.36%	98.86%	**99.24%**	**99.24%**	**99.24%**	**99.24%**	84.84%	93.56%
F1$_{score}$	68.67%	71.14%	96.12%	**96.67%**	95.78%	95.26%	95.57%	72.25%	92.68%
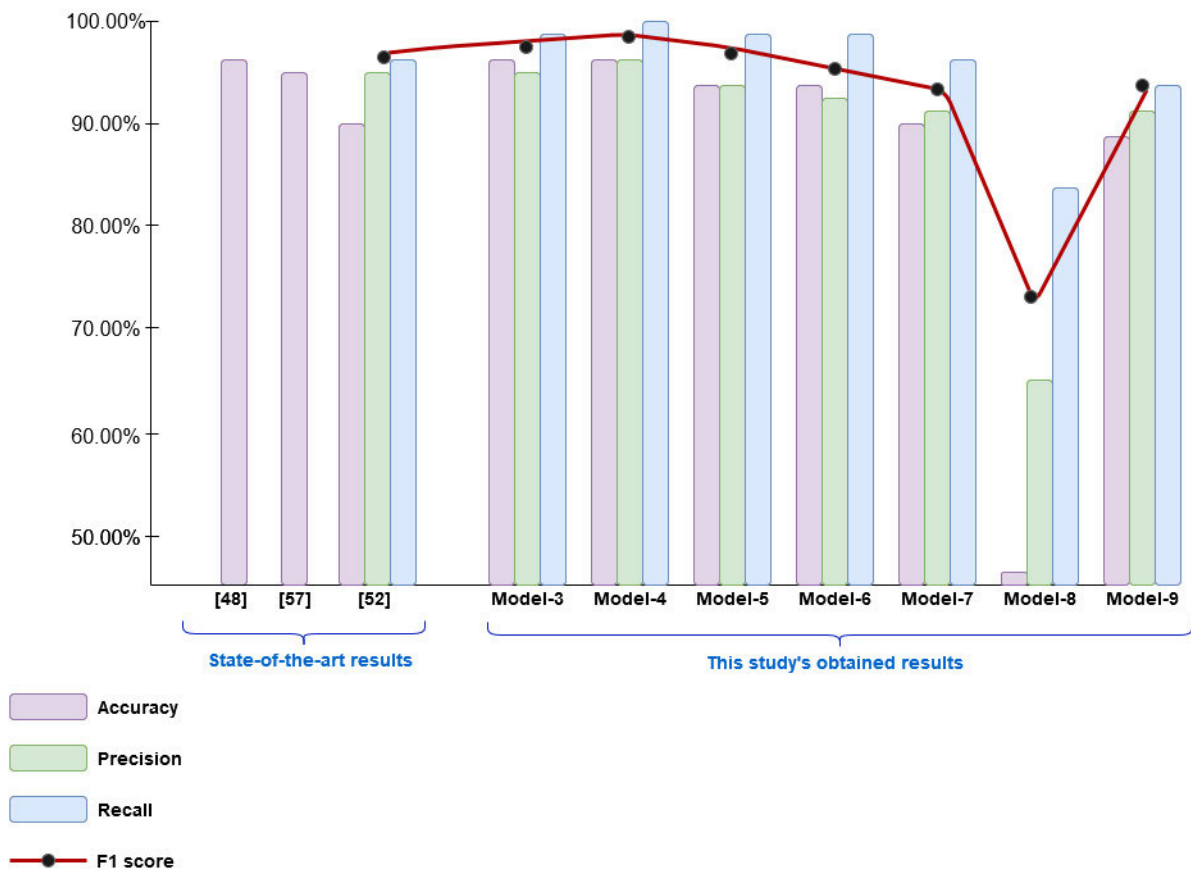
FIGURE 9. Average performance of the models.

ResNet50 based models (model-3 and model-5). The recall has the highest accuracy in all models showing that the possibility of having false-negative results is meager.

By comparing obtained-results in each magnification class, It shows a 2-layers CNN (model-1 and model-2) is outperforming pre-trained AlexNet (model-8).

Considering state-of-the-art results, which were achieved on BreakHis dataset and provided in table 1 and our best-achieved results (model-3 and model-4), our models are improving state-of-the-art results in both binary and multiclass scheme and magnification dependent and independent categories. We considered the same dataset as previous state-of-the-art studies, presented their methods, and improved the classification accuracies.

During the development, we implemented the bottom-up integration and testing approach. We split a CAD system into subsystems; Preprocessing, path/slides, feature extractor, transfer learning CNN, and postprocessing. Each part was individually tested with a range of configurations while other subsystems had a fixed configuration. By monitoring the changes in the outcome, we optimized each subsystem individually first and then integrated them to make a specialized CAD system for our specific problem. Finally, the subsystems were integrated to make a system, and the final system was evaluated. This systems engineering approach helped us to achieve our goal, together with Agile SCRUM methodology for software development.

Our main contributions in this research are the proposed models with optimized hyperparameters, which is a unique design for this specific problem. Previous studies have used ResNet networks but were not able to achieve high accuracy, mainly because of improper tuning and patch extraction approaches. Our tuning is unique to our model and has not being used in the past. Our study has utilized DesneNet121 CNN for the classification for the first time. However, most of the previous works have implemented ResNet, VGG, AlexNet, and CaffeNet for their model. The difficulty of working with the DenseNet and the large volume of the network may be the reason for not utilizing the DenseNet for the specific problem. Our study shows that not only is it feasible to utilize DenseNet for breast cancer histological diagnosis, but it is also possible to create high accuracy models. Our results are improving state-of-the-art results in all classifications (both binary and multiclass), which is very important when applying artificial intelligence in the medical domain.

VI. CONCLUSION

In this work, different methods and solutions for automatic detection of breast cancer in histopathology images have been investigated. The aim of this work was to develop a high accuracy method that can detect cancer at early stages, define the exact type of the samples, and improve previous works results.

We first surveyed the literature on the topic to find out previous approaches for solving the problem, together with state-of-the-art CAD systems for breast cancer recognition. Drew on the literature review, we proposed different models for automatic breast cancer diagnosis based on deep learning framework and transfer learning framework. Then we presented and analyzed our image preprocessing methods

(Data augmentation, dimension reduction, etc.). Moreover, the design methodologies of deep neural networks were presented. Next, the architecture of ResNet50 and DenseNet121, which were our main deep learning models, were utilized for the transfer learning framework.

Based upon an extensive study on various deep convolutional neural network techniques, we developed a very effective transfer learning architecture that consists new fully-connected classifier and an input layer that is combined with pre-trained DenseNet121 and ResNet50 models. We introduced the dataset for our study (BreakHis), the shortcomings of the dataset, and specified our training set, validation set, and test set in detail. We then implemented our proposed framework in Python with PyTorch in the Jupyter notebook of anaconda. Following the fact that hyperparameters in CNNs are very important to the efficiency of the model, we provide a set of hyperparameters (learning rate, pooling size, learning rate scheduling, etc.). We optimized a set of hyperparameters and tested them on a fraction of the BreakHis dataset. After finding the best settings for our hyperparameters, we define nine models and experiment with those models with our best hyperparameters settings.

We achieved the accuracy of 100%, 100%, 99.02%, and 99.48% for 40×, 100×, 200×, and 400× images in magnification-dependent binary classification, respectively. As for multiclass classification, 98.43%, 98.54%, 97.53%, and 97.40% accuracies for 40×, 100×, 200×, and 400× images in the magnification-dependent category has obtained. In the magnification-independent category, we achieved an accuracy of 99.50% and 97.72% for binary and multiclass classification, respectively. All of our results in all categories and magnification dependent and independent are well improving state-of-the-art results. This promising result is another leap toward digitalization and convincing medical experts to trust CAD systems for breast cancer detection.

Although this study evaluated the proposed models on the best available dataset for breast histological images, it is still may not be broad enough research. The BreakHis dataset was made from only 82 patients, which makes the data diversity extremely limited. The dataset does not provide any information about the number of images from each specific patient to calculate Patient-Level Accuracy (PLA). The proposed model is showing promising results, but before using it in real-world examples, it has to be tested with a few more datasets to compare the results and increase the variety of data.

Moreover, the proposed models are working less accurately in 400× images. This contradicts the fact that "if the training data contains great detail (such as 400× breast histopathology image), the CNN network performs better in predictions". The 400× images providing higher magnification and more details, but the system is actually working worse in this group of images. The investigation of lower accuracy in the 400× image can be another future work. However, the reason can be the total number of parameters

of ResNet50 and DenseNet121, which is not much enough to cover the great detail 400× images thoroughly.

APPENDIX
ABBREVIATION

AI	Artificial Intelligence
ANN	Artificial Neural Network
CAD	Computer Aided Diagnosis
CBFS	Correlation-Based Feature Selection
CNN	Convoloutinal Neural Network
ConvNet	Convolutional Networks
DBN	Deep Belief Network
DCNN	Deep Convolutional Neural Network
DWT	Discrete Wavelet Transform
ELF	Enhanced Loss Function
F1	F1 Score
FC	Fully Connected
FV	Fisher Vector
GPR	Gausian Random Projection
HSV	Hue, Saturation,and Value
ILA	Image-Level Accuracy
KPCA	Kernel Principle Component Analysis
L-ISOMAP	Landmark ISOMAP
MIB	Magnification Independent Binary classification
MIL	Multiple Instance Learning
MIM	Magnification Independent Multiclass classification
MIP	Multiple Instance Pooling layer
MSB	Magnification Specific Binary classification
MSM	Magnification Specific Multiclass classification
NDCNN	New-Designed Convolutional Neural Network
NGS	Nottingham Grading System
PCA	Principal Component Analysis
PLA	Patient-Level Accuracy
RBM	Restricted Boltzmann Machines
ReLU	Rectified Linear Unit
SGD	Stochastic Gradient Descent
SVM	Support Vector Machine
TL	Transfer Learning
WSI	Whole Slide Image
	Preprocessing:
CE	Contrast Enhancement
DAB	Data Augmentation with Balancing
DA	Data Augmentation
ETB	E-AHE and TB-HAT techniques
GSC	Gray Scale Conversion
KM	K-Mean clustering
MVD	Multilevel Variational mode Decomposition
RGBT	RBG channel Transformation
Res	Resizing
SMI	Subtract Mean Image
SN	Stain Normalization
	Patch Extraction:
GSC	Gray Scale Conversion
P	Patch extraction
RND	Random
SQ	deviding to non-overlapping Square tiles
SW	Sliding Window
	Features Extractors:
BE	Binarization Encoding
DR	Dimension Reduction
DWT	Discrete Wavelet Transform
FD	Fractal Dimension
HI	Histogram Information
JCTF	Joint Color Texture Features
LPQ	Local Phase Quantization
PFTAS	Parameter-Free Threshold Adjacency Statistics
PWT	Pyramid Structure Wavelet Transform
Tam	Tamura features

REFERENCES

[1] World Health Organization (WHO). (2020). *Breast Cancer*. [Online]. Available: https://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/

[2] ECA-Z Team. (2020). *Breast Cancer*. [Online]. Available: https://www.cancer.org/cancer/breast-cancer.html

[3] C. Wang, J. Shi, Q. Zhang, and S. Ying, "Histopathological image classification with bilinear convolutional neural networks," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 4050–4053.

[4] M. L. Wilson, K. A. Fleming, M. A. Kuti, L. M. Looi, N. Lago, and K. Ru, "Access to pathology and laboratory medicine services: A crucial gap," *Lancet*, vol. 391, no. 10133, pp. 1927–1938, May 2018.

[5] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 2560–2567.

[6] M. Veta *et al.*, "Assessment of algorithms for mitosis detection in breast cancer histopathology images," *Med. Image Anal.*, vol. 20, no. 1, pp. 237–248, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841514001807?casa_token=MnqqSVI4YAgAAAAA:S3AHjUcjQSF6oYtuvB9pnf1Ge0XwwXGNvw-0UfPldbW_Ec2M3S0QXH4oHOk7lRMmEM3mAmIaV5Pb

[7] T. Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, A. Eloy, A. Polónia, and A. Campilho, "Classification of breast cancer histology images using convolutional neural networks," *PLoS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0177544.

[8] J. Xu, L. Xiang, R. Hang, and J. Wu, "Stacked sparse autoencoder (SSAE) based framework for nuclei patch classification on breast cancer histopathology," in *Proc. IEEE 11th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2014, pp. 999–1002.

[9] Y. Benhammou, B. Achchab, F. Herrera, and S. Tabik, "BreakHis based breast cancer automatic diagnosis using deep learning: Taxonomy, survey and insights," *Neurocomputing*, vol. 375, pp. 9–24, Jan. 2020.

[10] A.-A. Nahid and Y. Kong, "Histopathological breast-image classification using local and frequency domains by convolutional neural network," *Information*, vol. 9, no. 1, p. 19, Jan. 2018.

[11] J. Sun and A. Binder, "Comparison of deep learning architectures for H&E histopathology images," in *Proc. IEEE Conf. Big Data Analytics (ICBDA)*, Nov. 2017, pp. 43–48.

[12] F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte, "Deep features for breast cancer histopathological image classification," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2017, pp. 1868–1873.

[13] S. Cascianelli, R. Bello-Cerezo, F. Bianconi, M. L. Fravolini, M. Belal, B. Palumbo, and J. N. Kather, "Dimensionality reduction strategies for cnn-based classification of histopathological images," in *Proc. Int. Conf. Intell. Interact. Multimedia Syst. Services*. Springer, 2018, pp. 21–30. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-59480-4_3

[14] V. Gupta and A. Bhavsar, "Sequential modeling of deep features for breast cancer histopathological image classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2018, pp. 2254–2261.

[15] W. Zhi, H. W. F. Yueng, Z. Chen, S. M. Zandavi, Z. Lu, and Y. Y. Chung, "Using transfer learning with convolutional neural networks to diagnose breast cancer from histopathological images," in *Neural Information Processing—ICONIP* (Lecture Notes in Computer Science), vol. 10637, D. Liu, S. Xie, Y. Li, D. Zhao, and E. S. El-Alfy, Eds. Cham, Switzerland: Springer, 2017. [Online]. Available: https://link.springer.com/chapter/10.1007%2F978-3-319-70093-9_71, doi: 10.1007/978-3-319-70093-9_71.

[16] E. Deniz, A. engür, Z. Kadiro lu, Y. Guo, V. Bajaj, and U. Budak, "Transfer learning based histopathologic image classification for breast cancer detection," *Health Inf. Sci. Syst.*, vol. 6, no. 1, p. 18, Dec. 2018.

[17] G. Zhang, M. Xiao, and Y. Huang, "Histopathological image recognition with domain knowledge based deep features," in *Intelligent Computing Methodologies—ICIC* (Lecture Notes in Computer Science), vol. 10956, D. S. Huang, M. Gromiha, K. Han, and A. Hussain, Eds. Cham, Switzerland: Springer, 2018. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-95957-3_38, doi: 10.1007/978-3-319-95957-3_38.

[18] Y. Song, J. J. Zou, H. Chang, and W. Cai, "Adapting Fisher vectors for histopathology image classification," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 600–603.

[19] Y. Song, H. Chang, H. Huang, and W. Cai, "Supervised intra-embedding of Fisher vectors for histopathology image classification," in *Medical Image Computing and Computer Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), vol. 10435, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. Collins, and S. Duchesne, Eds. Cham, Switzerland: Springer, 2017. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-66179-7_12, doi: 10.1007/978-3-319-66179-7_12.

[20] S. Akbar, M. Peikari, S. Salama, S. Nofech-Mozes, and A. Martel, "Transitioning between convolutional and fully connected layers in neural networks," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support—DLMIA* (Lecture Notes in Computer Science), vol. 10553, M. Cardoso *et al.*, Eds. Cham, Switzerland: Springer, 2017. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-67558-9_17, doi: 10.1007/978-3-319-67558-9_17.

[21] B. Wei, Z. Han, X. He, and Y. Yin, "Deep learning model based breast cancer histopathological image classification," in *Proc. IEEE 2nd Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA)*, Apr. 2017, pp. 348–353.

[22] K. V. K. S. S. Kumar, B. V. Rao, and G. V. E. S. Kumar, "Fractional order PLL based sensorless control of PMSM with sliding mode observer," in *Proc. Int. Conf. Power, Instrum., Control Comput. (PICC)*, Jan. 2018, pp. 1–6.

[23] P. J. Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte, and P. Honeine, "Multiple instance learning for histopathological breast cancer image classification," *Expert Syst. Appl.*, vol. 117, pp. 103–111, Mar. 2019.

[24] K. Das, S. Conjeti, A. G. Roy, J. Chatterjee, and D. Sheet, "Multiple instance learning of deep convolutional neural networks for breast histopathology whole slide classification," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 578–581.

[25] B. Du, Q. Qi, H. Zheng, Y. Huang, and X. Ding, "Breast cancer histopathological image classification via deep active learning and confidence boosting," in *Artificial Neural Networks and Machine Learning—ICANN* (Lecture Notes in Computer Science), vol. 11140, V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, and I. Maglogiannis, Eds. Cham, Switzerland: Springer, 2018. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-01421-6_11, doi: 10.1007/978-3-030-01421-6_11.

[26] K. Das, S. P. K. Karri, A. G. Roy, J. Chatterjee, and D. Sheet, "Classifying histopathology whole-slides using fusion of decisions from deep convolutional network on a collection of random multi-views at multi-magnification," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 1024–1027.

[27] A.-A. Nahid, A. Mikaelian, and Y. Kong, "Histopathological breast-image classification with restricted Boltzmann machine along with backpropagation," School Eng., Macquarie Univ., Sydney, NSW, Australia, Tech. Rep., 2018. [Online]. Available: https://www.biomedres.info/biomedical-research/histopathological-breastimage-classification-with-restricted-boltzmann-machine-along-with-backpropagation-10287.html, doi: 10.4066/biomedicalresearch.29-17-3903.

[28] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li, and A. M. Dobaie, "Facial expression recognition via learning deep sparse autoencoders," *Neurocomputing*, vol. 273, pp. 643–649, Jan. 2018.

[29] S. Pratiher and S. Chattoraj, "Diving deep onto discriminative ensemble of histological hashing & class-specific manifold learning for multi-class breast carcinoma taxonomy," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1025–1029.

[30] J. Chang, J. Yu, T. Han, H.-J. Chang, and E. Park, "A method for classifying medical images using transfer learning: A pilot study on histopathology of breast cancer," in *Proc. IEEE 19th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Oct. 2017, pp. 1–4.

[31] A.-A. Nahid, M. A. Mehrabi, and Y. Kong, "Histopathological breast cancer image classification by deep neural network techniques guided by local clustering," *BioMed Res. Int.*, vol. 2018, pp. 1–20, Mar. 2018.

[32] R. Karthiga and K. Narasimhan, "Automated diagnosis of breast cancer using wavelet based entropy features," in *Proc. 2nd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Mar. 2018, pp. 274–279.

[33] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016.

[34] A. Chan and J. A. Tuszynski, "Automatic prediction of tumour malignancy in breast cancer with fractal dimension," *Roy. Soc. Open Sci.*, vol. 3, no. 12, Dec. 2016, Art. no. 160558.

[35] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, "Breast cancer multi-classification from histopathological images with structured deep learning model," *Sci. Rep.*, vol. 7, no. 1, pp. 1–10, Dec. 2017.

[36] M. A. Kahya, W. Al-Hayani, and Z. Y. Algamal, "Classification of breast cancer histopathology images based on adaptive sparse support vector machine," *J. Appl. Math. Bioinf.*, vol. 7, no. 1, p. 49, 2017.

[37] V. Gupta and A. Bhavsar, "An integrated multi-scale model for breast cancer histopathological image classification with joint colour-texture features," in *Computer Analysis of Images and Patterns—CAIP* (Lecture Notes in Computer Science), vol. 10425, M. Felsberg, A. Heyden, and N. Krüger, Eds. Cham, Switzerland: Springer, 2017. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-64698-5_30, doi: 10.1007/978-3-319-64698-5_30.

[38] A. A. Samah, M. F. A. Fauzi, and S. Mansor, "Classification of benign and malignant tumors in histopathology images," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Sep. 2017, pp. 102–106.

[39] V. Gupta, A. Singh, K. Sharma, and A. Bhavsar, "Automated classification for breast cancer histopathology images: Is stain normalization important?" in *Computer Assisted and Robotic Endoscopy and Clinical Image-Based Procedures—CARE* (Lecture Notes in Computer Science), vol. 10550, M. Cardoso *et al.*, Eds. Cham, Switzerland: Springer, 2017. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-67543-5_16, doi: 10.1007/978-3-319-67543-5_16.

[40] E. M. Nejad, L. S. Affendey, R. B. Latip, and I. B. Ishak, "Classification of histopathology images of breast into benign and malignant using a single-layer convolutional neural network," in *Proc. Int. Conf. Imag., Signal Process. Commun. (ICISPC)*, 2017, pp. 50–53.

[41] M. Sharma, R. Singh, and M. Bhattacharya, "Classification of breast tumors as benign and malignant using textural feature descriptor," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Nov. 2017, pp. 1110–1113.

[42] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, Y. Ma, H. Shi, and Y. Zhao, "Size-scalable content-based histopathological image retrieval from database that consists of WSIs," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 4, pp. 1278–1287, Jul. 2018.

[43] S. Chattoraj and K. Vishwakarma, "Classification of histopathological breast cancer images using iterative VMD aided Zernike moments & textural signatures," 2018, *arXiv:1801.04880*. [Online]. Available: http://arxiv.org/abs/1801.04880

[44] B. S. Veeling, J. Linmans, J. Winkens, T. Cohen, and M. Welling, "Rotation equivariant CNNs for digital pathology," in *Medical Image Computing and Computer Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), vol. 11071, A. Frangi, J. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, Eds. Cham, Switzerland: Springer, 2018. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-00934-2_24, doi: 10.1007/978-3-030-00934-2_24.

[45] D. Sanchez-Morillo, J. González, M. García-Rojo, and J. Ortega, "Classification of breast cancer histopathological images using KAZE features," in *Bioinformatics and Biomedical Engineering—IWBBIO* (Lecture Notes in Computer Science), vol. 10814, I. Rojas and F. Ortuño, Eds. Cham, Switzerland: Springer, 2018. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-78759-6_26, doi: 10.1007/978-3-319-78759-6_26.

[46] Y. Benhammou, S. Tabik, B. Achchab, and F. Herrera, "A first study exploring the performance of the state-of-the art CNN model in the problem of breast cancer," in *Proc. Int. Conf. Learn. Optim. Algorithms: Theory Appl. (LOPAL)*, 2018, pp. 1–6.

[47] Z. Gandomkar, P. C. Brennan, and C. Mello-Thoms, "MuDeRN: Multi-category classification of breast histopathological image using deep residual networks," *Artif. Intell. Med.*, vol. 88, pp. 14–24, Jun. 2018.

[48] P. Alirezazadeh, B. Hejrati, A. Monsef-Esfahani, and A. Fathi, "Representation learning-based unsupervised domain adaptation for classification of breast cancer histopathology images," *Biocybern. Biomed. Eng.*, vol. 38, no. 3, pp. 671–683, 2018.

[49] J. A. Badejo, E. Adetiba, A. Akinrinmade, and M. B. Akanle, "Medical image classification with hand-designed or machine-designed texture descriptors: A performance evaluation," in *Bioinformatics and Biomedical Engineering—IWBBIO* (Lecture Notes in Computer Science), vol. 10814, I. Rojas and F. Ortuño, Eds. Cham, Switzerland: Springer, 2018. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-78759-6_25, doi: 10.1007/978-3-319-78759-6_25.

[50] E. Mahraban Nejad, L. S. Affendey, R. B. Latip, I. B. Ishak, and R. Banaeeyan, "Transferred semantic scores for scalable retrieval of histopathological breast cancer images," *Int. J. Multimedia Inf. Retr.*, vol. 7, no. 4, pp. 241–249, Nov. 2018.

[51] D. Bardou, K. Zhang, and S. M. Ahmad, "Classification of breast cancer based on histology images using convolutional neural networks," *IEEE Access*, vol. 6, pp. 24680–24693, 2018.

[52] A.-A. Nahid and Y. Kong, "Histopathological breast-image classification using concatenated R–G–B histogram information," *Ann. Data Sci.*, vol. 6, no. 3, pp. 513–529, Sep. 2019.

[53] N. Bayramoglu, J. Kannala, and J. Heikkila, "Deep learning for magnification independent breast cancer histopathology image classification," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 2440–2445.

[54] V. Gupta and A. Bhavsar, "Breast cancer histopathological image classification: Is magnification important?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jul. 2017, pp. 17–24.

[55] M. A. Nawaz, A. A. Sewissy, and T. H. A. Soliman, "Automated classification of breast cancer histology images using deep learning based convolutional neural networks," *Int. J. Comput. Sci. Netw. Secur.*, vol. 4, pp. 152–160, Apr. 2018.

[56] J. Xie, R. Liu, J. Luttrell, IV, and C. Zhang, "Deep learning based analysis of histopathological images of breast cancer," *Frontiers Genet.*, vol. 10, p. 80, Feb. 2019.

[57] Y. Li, J. Wu, and Q. Wu, "Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning," *IEEE Access*, vol. 7, pp. 21400–21408, 2019.

[58] D. M. Vo, N.-Q. Nguyen, and S.-W. Lee, "Classification of breast cancer histology images using incremental boosting convolution networks," *Inf. Sci.*, vol. 482, pp. 123–138, May 2019.

[59] K. Roy, D. Banik, D. Bhattacharjee, and M. Nasipuri, "Patch-based system for classification of breast histology images using deep learning," *Comput. Med. Imag. Graph.*, vol. 71, pp. 90–103, Jan. 2019.

[60] Y. Zheng, Z. Jiang, F. Xie, H. Zhang, Y. Ma, H. Shi, and Y. Zhao, "Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification," *Pattern Recognit.*, vol. 71, pp. 14–25, Nov. 2017.

[61] A. Hamidinekoo, E. Denton, A. Rampun, K. Honnor, and R. Zwiggelaar, "Deep learning in mammography and breast histology, an overview and future trends," *Med. Image Anal.*, vol. 47, pp. 45–67, Jul. 2018.

[62] S. Acharya, A. Alsadoon, P. W. C. Prasad, S. Abdullah, and A. Deva, "Deep convolutional network for breast cancer classification: Enhanced loss function (ELF)," *J. Supercomput.*, 2020. [Online]. Available: https://link.springer.com/article/10.1007%2Fs11227-020-03157-6, doi: 10.1007/s11227-020-03157-6.

[63] S. Sharma and R. Mehra, "Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—A comparative insight," *J. Digit. Imag.*, vol. 33, pp. 632–654, 2020. [Online]. Available: https://link.springer.com/article/10.1007%2Fs10278-019-00307-y, doi: 10.1007/s10278-019-00307-y.

[64] Y. Zhang, B. Zhang, F. Coenen, J. Xiao, and W. Lu, "One-class kernel subspace ensemble for medical image classification," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, p. 17, Dec. 2014.

[65] H. Guang-yong, "Study and practice of import scrum agile software development," in *Proc. IEEE 3rd Int. Conf. Commun. Softw. Netw.*, May 2011, pp. 217–220.

[66] Y. Yari, H. Nguyen, and T. Nguyen, "Accuracy improvement in binary and multi-class classification of breast histopathology images," in *Proc. IEEE 8th Int. Conf. Commun. Electron. (ICCE)*, 2020, pp. 1–6.

[67] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection," *Construct. Building Mater.*, vol. 157, pp. 322–330, Dec. 2017.

[68] J. Ba and R. Caruana, "Do deep nets really need to be deep?" in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2654–2662.

[69] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Computer Vision—ECCV* (Lecture Notes in Computer Science), vol. 9908, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham, Switzerland: Springer, 2016. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-46493-0_38, doi: 10.1007/978-3-319-46493-0_38.

[70] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[71] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 4700–4708.

[72] T. Tan, Z. Li, H. Liu, F. G. Zanjani, Q. Ouyang, Y. Tang, Z. Hu, and Q. Li, "Optimize transfer learning for lung diseases in bronchoscopy using a new concept: Sequential fine-tuning," *IEEE J. TransIEEE J. Transl. Eng. Health Med.*, vol. 6, pp. 1–8, 2018.

[73] M. C. Staff. (2018). *Invasive Lobular Carcinoma*. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/invasive-lobular-carcinoma/symptoms-causes/syc-20373973

YASIN YARI received the M.Sc. degree in systems engineering with embedded systems from the University of South-Eastern Norway, Campus Kongsberg, in 2020. His research interests include deep learning, image and video processing, algorithms, segmentation and classification, embedded real-time systems, the IoT, and data science.

THUY V. NGUYEN received the B.Sc. degree in electrical engineering from the Hanoi University of Science and Technology (HUST), Hanoi, Vietnam, the M.Sc. degree from New Mexico State University, Las Cruces, NM, USA, and the Ph.D. degrees from the University of Texas at Dallas, Richardson, TX, USA. He is currently a Lecturer with the Faculty of Information Technology, Posts and Telecommunications Institute of Technology (PTIT), Hanoi, Vietnam. Before joining PTIT, he was a Member of Technical Staff with Flash Channel Architecture, Seagate, Fremont, CA, USA. His research interest includes coding theory and its applications in next-generation communication systems.

HIEU T. NGUYEN received the B.Sc. degree in electrical engineering from the Hanoi University of Science and Technology, the M.Sc. degree from the University of Saskatchewan, Canada, and the Ph.D. degree from the Norwegian University of Science and Technology. He is a Faculty Member with the Faculty of Technology, Natural Sciences, and Maritime Sciences, University of South-Eastern Norway (USN).

• • •