A Novel Transfer-Learning Model for Automatic Detection and Classification of Breast Cancer Based Deep CNN

Abeer Saber
dept. of computer science faculty of computers and information, Kafr El-Sheikh University
Kafr El-Sheikh 33511, Egypt
Abeer_Saber@fci.kfs.edu.eg

Mohamed Sakr
dept. of computer science Faculty of Computers and Information, Menoufia University
Menoufia 32511, Egypt
mohamed.sakr@ci.menofia.edu.eg

Osama M. Abo-Seida
dept. of computer science faculty of computers and information, Kafr El-Sheikh University
Kafr El-Sheikh 33511, Egypt
aboseida@yahoo.com

Arabi Keshk
dept. of computer science Faculty of Computers and Information, Menoufia University
Menoufia 32511, Egypt
arabi.keshk@ci.menofia.edu.eg

***Abstract*— Breast cancer (BC) is a leading cause of cancer death among women in which breast cells develop out of control is by encouraging patients to receive timely care, early detection of BC increases the likelihood of survival. In this context, a new deep learning (DL) model is presented for automatic detection and classification of the suspected area of the breast based on the transfer learning (TL) technique. A pre-trained visual geometry group (VGG)-19, VGG16, and InceptionV3 networks are used in the presented model to transfer their learning parameters for improving the performance of breast tumor classification. The main goals of this project are to use segmentation to automatically determine the affected breast tumor region, reduce training time, and improve classification performance. In the presented model, the Mammographic Image Analysis Society (MIAS) dataset is used for extracting the breast tumor features. We have chosen four evaluation metrics for evaluating the performance of the presented model accuracy, sensitivity, specificity, and area under the ROC curve (AUC). The experiments showed that transferring parameters from the model of VGG16 is a powerful for BC classification than VGG19 and Inception V3 with overall specificity, accuracy, sensitivity, and AUC 98%,96.8%, 96%, and 0.99, respectively.**

Keywords—breast cancer, deep-learning, segmentation, transfer-learning, image processing

I. INTRODUCTION

A cancer tumor is an irregular growth of cells that invades the human body's underlying tissues. Tumors are classified as benign or malignant. The benign tumor is made up of non-cancerous cells that grow only locally and cannot spread by invasion. A malignant tumor is made up of cancerous cells that can grow out of control, spread to other parts of the body, and invade the tissue around them.

Over the live cycle, about 12% of women in the USA are predicted to be diagnosed as a patient with BC. Every two minutes, on average, a woman is diagnosed with BC [1, 2]. This makes BC the foremost cancer in ladies [3]. The type of BC depends on the cells that transform into cancer.

As a result, early identification of BC is critical for improving rates of patient's survival. The high morbidity and large healthcare costs associated with cancer have prompted researchers to develop more precise cancer detection techniques. The two most popular methods for detecting BC are mammography and biopsy. In mammography, a particular form of breast imaging is used by a radiologist to identify early cancer signs in women. It has been noted that the death rate has decreased because of mammography for cancer screening usage. Another well-efficient diagnostic technique for observing BC is a biopsy.

The major BC image challenges are automatic identification and localization because of their variance in size, shapes, and locations. Also, there are other abnormalities in breast images like mastitis, adenopathy, and granuloma [4]. An important role is played in the diagnosis of BC using machine learning (ML) techniques. Different strategies for automated cell classifications for BC detection have been proposed by several researchers in recent decades.

However, due to the complexity of traditional machine learning approaches like preprocessing, segmentation, and feature extraction, system performance suffers in both efficiency and accuracy. To address traditional ML problems, the DL has recently emerged with the ability to tackle problems including image classification and object localization with outstanding feature representation. CNN's are the most common of the DL models implemented in the related works. The architecture of CNN is directly updated by the 2-D input data [5, 6].

A huge volume of data that is missing in the BC field, is needed for the CNN training task. One way to solve this is using TL to transfer the features from a large dataset like ImageNet, as illustrated, in Fig. 1. By integrating their expertise, the TL principle can be exploited for performance improvement of individual CNN architectures [7]. The key benefit of TL is that it improves classification accuracy and speeds up the training process. A model transfer is an effective TL method: first, pre-train the network parameters via the source data, then apply these learned parameters to the target model [8]. The TL technique is presented in Fig.1.

In this study, a TL-based classification model for multi-class BC detection and classification is presented. This model contains five main phases to enhance the breast images (noise removal, data augmentation, segmentation, morphological analysis, and histogram equalization). Then, a pre-trained model like VGG16, VGG19, and InceptionV3 have been applied to transmit their learned layers to the target model. The

main goals presented in this study are to automatically isolate the affected patch using segmentation, minimize the training time, and enhance the efficiency of the classification.

automatic detection and localization of cancer cells in BC images is a major challenge [4].

This paper is organized as follows: related work is described in sections II, III which contains a presented model description for BC diagnosis and classification. The presented model results over mammography data are reported in section IV. At last, in section V, the current study's conclusion is presented.

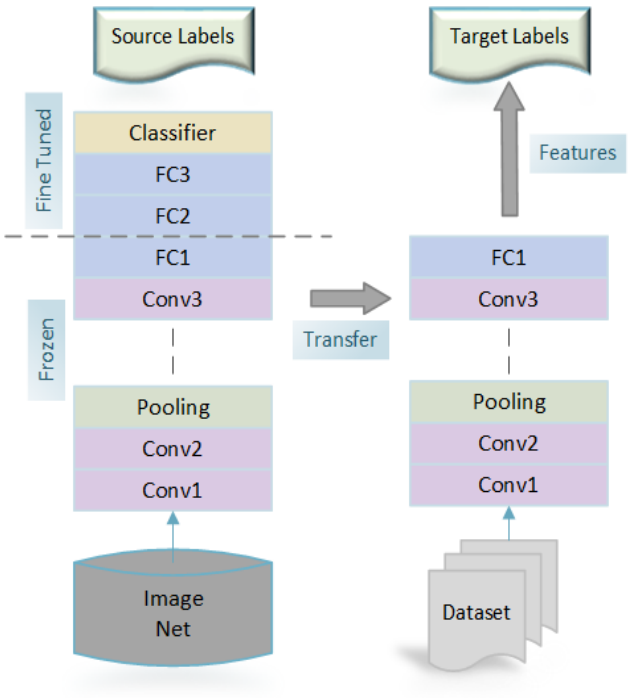


Fig. 1. Transfer learning methodology

II. RELATED WORKS

In [9] Wahab et al. implemented the TL technique from a pre-trained CNN to another presented CNN to classify the BC mitoses. Lotter et al. [10] suggested a model in which the features are retrieved using a pre-trained network ResNet50 for multi-class BC classification. Their model was able to classify lesions into five anomaly classes.

Jiang et al. [11] found that employing TL instead of constructing networks from scratch improved BC classification accuracy. The BCDR-F03 dataset was used to evaluate their model, and the accuracy was 88% for GoogleNet and 83% for AlexNet. Khan et al. [12] developed a framework in which pre-trained CNN architectures, such as GoogLeNet, VGGNet, and ResNet, are used to extract mammography features. A typical benchmark dataset is used to assess the proposed model's accuracy.

Cao et al. [13] used a random forest algorithm for integrating multiple feature sets to increase the performance of the source ResNet-125 network. They enhanced the classification accuracy to 82.90% using the "ICIAR 2018" dataset.

Using the BreaKHis dataset, Deniz et al. [14] employed a fine-tuned approach for the VGG16 and Alexnet networks for BC classification. The accuracy of the presented model is higher than that of other techniques, with a score of 91.37%. Celik et al. [15] transferred the learned parameters from the DenseNet-161 network and obtained 91.5%, 92.3% for accuracy and F-sore, respectively. BreastNet was also proposed by Toaçar et al. [16], which consists of convolutional, residual, pooling, and dense layers to obtain the much more effective characteristics from breast mammogram images. It outperforms VGG-16, AlexNet, and VGG-19 models in terms of accuracy, with a score of 98.80%.

Using the extracted features from the MIAS, Ting et al. [17] developed a CNN network for classifying BC. Their network consists of 1, 28, and 1 for input, hidden, and output layers, respectively. They are used the feature-wise data augmentation algorithm to avoid overfitting. Abbas [18] proposed a multi-layer DL structure to identify cancerous and non-cancerous areas of the breast. The proposed model contains four phases for feature extraction, feature transformation, and feature learning to obtain the decision.

Sha et al. [19] developed a method for automatically detecting and classifying the malignant regions in breast images using CNN and the grasshopper optimization algorithm. A novel CNN was trained for detecting BC by Charan et al. [20]. The presented deep CNN consists of 6 convolutional layers, 4 average-pooling layers, and 3 FCLs. The input image was 224 by 224 pixels in size, and the classification results were applied using the soft-max (SM) algorithm. The summary of related works results is shown in Table I.

TABLE I. COMPARISON BETWEEN THE EXISTING MODELS RESULTS.

Author /Year	**CNN Performance**				
	Breast Datasets	***Accuracy***	***Sensitivity***	***Specificity***	***AUC***
Wahab et al (2019)	TUPAC16	-	96.2	90.9	0.94
Jiang et al (2017)	BCDR-F03	88	-	-	-
Khan et al (2019)	benchmark	97.525	-	-	-
Cao et al (2018)	ICIAR 2018	82.90	-	-	-
Deniz et al (2018)	BreaKHis	91.37	-	-	-
Celik et al (2020)	BreaKHis	91.57	-	-	-
Toğaçar(2020)	BreakHis	98.8%.	98.7	-	-
Sha et al (2020)	MIAS	92	96	93	-
Ting (2019)	MIAS	90.5	89.4	90.7	0.90
Charan et al (2018)	MIAS	65	-	-	-
Abbas(2016)	MIAS	91.5	92	84.2	0.91
The proposed	**MIAS**	**96.8**	**96**	**98**	**0.99**

III. THE PROPOSED MODEL

The proposed method for the identification and classification of breast images has been applied and implemented. As shown in Fig. 2, the presented model contains two key aspects parts. The first part involves five steps for breast image preprocessing.

To eliminate the limitations of detecting anomalies without undue interference from a mammogram, image preprocessing is critical. Before the learning process, to reduce computing time, tumor areas are automatically detected using segmentation algorithms. Morphological analysis, histogram equalization, and Noise removal can all be applied before segmentation to improve image quality and segmentation performance.

a) Image Opening (IMGO):

$$IMGO = Input \ominus STE \oplus STE \quad (1)$$

b) Image Closing (IMGC):

$$IMGC = Input \oplus STE \ominus STE \quad (2)$$

c) White Top-hat (WHTH):

$$WHTH = Input - IMGO \quad (3)$$

d) Black Top-hat (BLTH):

$$BLTH = IMGC - Input \quad (4)$$

e) Mathematical Morphological (MAMO):

$$MAMO = Input + WHTH - BLTH \quad (5)$$

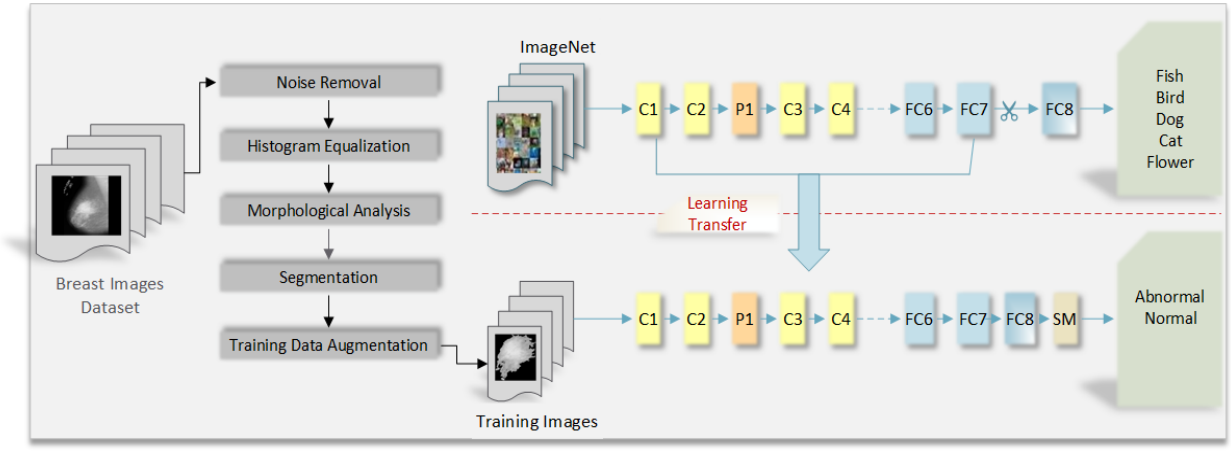


Fig.2. The proposed model for BC classification

A. *Data Preprocessing*

1) Noise removal

A 2-Dimention median filtering of 3 x 3 sizes was used to dispose noise of digitization from the mammography image.

2) Histogram equalization

To enhance the images contrast, we used the standard histogram equalization process on all levels of the original image. This is done by efficiently spreading the image's most repeated gray level, that is, by extending the image's intensity range.

3) Morphological analysis

A significant method for extracting the non-breast areas before segmentation is using the morphological analyzer. In morphological operations, after using the structuring element (STE), the related structures are retrieved from the input. The output image size is the same as the input size, and the value of each pixel is determined by the input pixel and its neighbors. The morphological operations can be estimated using Eq. 1- Eq. 5 [21]:

Where $\ominus$ and $\oplus$ relate to the operations of "erosion" and "dilation" respectively.

4) Segmentation

We used a threshold-based segmentation approach to automatically extract the affected patch to focus the analysis on the area that was most affected by cancer. So, the time for computing has been decreased.

5) Training Data Augmentation

When we employed bigger datasets, DL models performed better. One of the most commonly used methods for extending the dataset size is data augmentation, which aids in the resolve of overfitting caused by a lack of data preparation. In this research, the MIAS dataset is split into two parts: a training set of "80%" and a testing set of "20%. Then, a group of transformations is performed to increase the amount of training data. At 0o, 90o, 180o, 270o, the training mammogram data is rotated and then flipped and added to the training data. Algorithm 1 provides a comprehensive overview of training

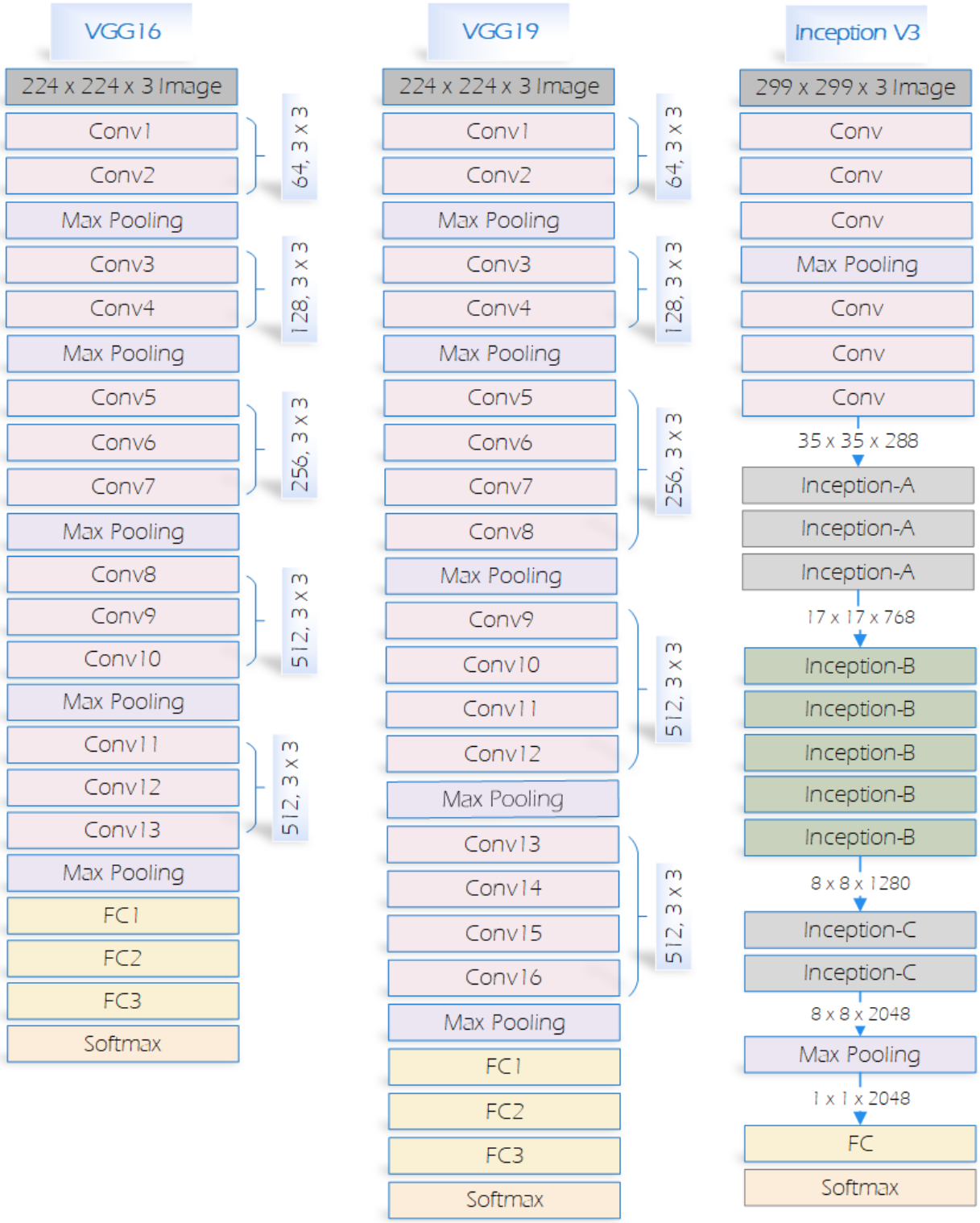


Fig.3. The basic InceptionV3, VGG16, and VGG19 architecture

data augmentation. Where R denotes the rotation process and F denotes the flipping process.

Algorithm1: Training Data Augmentation Algorithm
Input: Benign (B), Normal (N), and Malignant (M) segmented patches Processing Phases: - Phase1: ∀ B, R to (0o, 90o, 180o, 270o) - Phase2: Apply F on all step1 - Phase3: ∀ M, R to (0o, 90o, 180o, 270o) - Phase4: Apply F on all step3 - Phase5: ∀ N, R to (0o, 90o, 180o, 270o) - Phase6: Apply F on all step5 Output: - Save phases 1,2,3,4,5,6

B. Transfer the Learned Parameters from CNN

The VGG16, VGG19, and InceptionV3 networks are applied to extract features in the presented model. The network layer filters are used to identify the characteristics of the inputs, such as colors and lines. Then the identification of trivial forms and tiny pieces. The generated output can be determined by the class belonging to the input image, such as cats, birds, etc.

As illustrated in the presented model, all learned layers from the source task are frozen and transmitted to the target task, except the final three layers (Fully Connected (FC), SM, and Classification). The augmented patches are then used to continue the training process. In addition, for new class categorization, the pre-trained network parameters are coupled with the newly trained layers for new class classification. The TL method helps create a very quick training process and needs very little training data compared to building CNN from scratch. The fine-tuning was done with SGD with momentum (SGDM), which is an extension of SGD that takes into account previous gradients in each dimension. As a function of prior momentum and current gradient in that dimension, SGDM maintains momentum in each dimension. SGDM's purpose is to achieve high "velocity" in any dimension with a consistent gradient. The jittering problem in SGDM is addressed by employing high-velocity inconsistent gradient dimensions, and the saddle point problem is mitigated by using prior gradients, which offer some momentum, even when the present gradient is near zero.

VGG16 is trained by the ImageNet dataset [21]. Its structure is complex and simple. It has 13 convolutions layers, 5 pooling layers, 3 FC layers and there is an SM classifier. The RGB image is used as the input. The filters used are a 3 x 3 with a stride (S) of 1, and the pooling is a 2 x 2-pixel window with a S of 2.

The ImageNet database is used to train **VGG19**. VGG19 includes 19 of the deep layers divided into 16 convolutions layers and 3 max-pooling layers with a size of 244 x 244 for the input image. In VGG19, 3 x 3 is the kernel size with 1 S size, while max-pooling done in a 2-pixel by 2-pixel window with 2 S size.

The team of Google research developed **Inception-V3** as a CNN. It consists of 48 layers, each with a 299 × 299 input image and trained over the ImageNet. The InceptionV3, VGG16, and VGG19 basic architecture are presented in Fig. 3.

IV. RESULYS

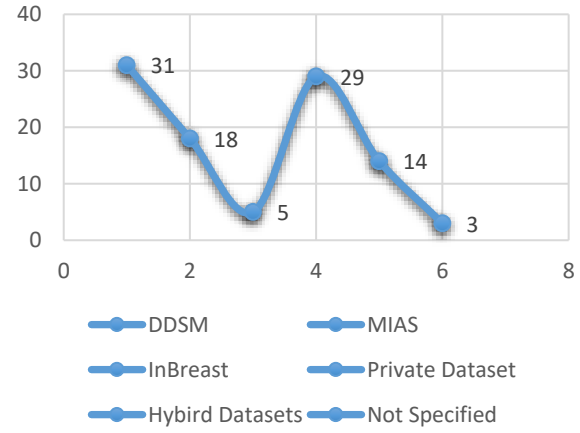


Fig.4. The most common breast databases

A. Dataset descriptopn

As illustrated above, in Fig. 4, MIAS is the second common dataset used on the basis of the statistics discussed on [22] for BC classification models. In this study, the MIAS was used to evaluate the presented model. Each image consists of 1024 x 1024 pixels. The MIAS contained 322 images, 113 images for the abnormal case (Benign and Malignant), and 209 images for the normal case. It offers a comprehensive explanation of the mammogram images for ground truth, such as abnormality present class, background tissue, abnormality center coordinates, tumor type, and abnormality circle radius.

B. Experimental Analysis

Several experiments were performed in this part to evaluate the performance of the presented model on the MIAS dataset. Fig. 5 illustrates the first stage in the provided model for preprocessing results.

Then TL is applied on the “InceptionV3, VGG16, and VGG19” models with (Min batch size = “20”, Max Epochs = “20”, Learn-rate drop factor = “0.5”, Initial learn rate = “1e-4”, and Learn-rate drop period = “5”) training parameters and compared with each other in terms of sensitivity, accuracy, AUC, and specificity.

The data was divided into three categories (normal, benign, and malignant). Then, it was split into 80% for training and 20% for testing, respectively.

The performance of the presented methodology is measured using the evaluation metrics, presented in Table II, in Eq. 6- Eq. 12. The results of the classifier after preprocessing are introduced in Table III. It can be noticed that the VGG16 obtains the best accuracy, sensitivity, specificity, and AUC with values 96.8%, 96%, 98%, and 0.99, respectively.

Category-wise receiver operating characteristic (ROC) curves of different CNNs during the testing process are shown in Fig. 7. A total comparison between InceptionV3, VGG16, and VGG19 is presented in Fig. 6, while a comparison between our model and other existing models is shown in Fig. 8.

TABLE II. THREE CLASS CONFUSION MATRIX TABLE

Classes		Predicted		
		Benign	**Malignant**	**Normal**
Actual	**Benign**	T_{BB}	T_{BM}	T_{BN}
	Malignant	T_{MB}	T_{MM}	T_{MN}
	Normal	T_{NB}	T_{NM}	T_{NN}

$$Accuracy = \frac{T_{BB}+T_{MM}+T_{NN}}{T_{BB}+T_{MB}+T_{NB}+T_{BM}+T_{MM}+T_{NM}+T_{BN}+T_{MN}+T_{NN}} \quad (6)$$

$$Sensitivity_B = \frac{T_{BB}}{T_{BB}+T_{BM}+T_{BN}} \quad (7)$$

$$Specificity_B = \frac{T_{MM}+T_{MN}+T_{NM}+T_{NN}}{T_{MM}+T_{NM}+T_{MN}+T_{NN}+T_{MB}+T_{NB}} \quad (8)$$

$$Sensitivity_M = \frac{T_{MM}}{T_{MB}+T_{MM}+T_{MN}} \quad (9)$$

$$Specificity_M = \frac{T_{BB}+T_{BN}+T_{NB}+T_{NN}}{T_{BB}+T_{BN}+T_{NB}+T_{NN}+T_{BM}+T_{NM}} \quad (10)$$

$$Sensitivity_N = \frac{T_{NN}}{T_{NB}+T_{NM}+T_{NN}} \quad (11)$$

$$Specificity_N = \frac{T_{BB}+T_{BM}+T_{MB}+T_{MM}}{T_{BB}+T_{BM}+T_{MB}+T_{MM}+T_{BN}+T_{MN}} \quad (12)$$

TABLEIII. THE PRESENTED MODEL CLASSIFICATION PERFORMANCE

Deep Network	**Class**	**Deep Network Classifier Performance**			
		Accuracy	***Sensitivity***	***Specificity***	***AUC***
VGG16	Benign	0.971	0.99	0.97	0.992
	Malignant	0.974	0.95	0.98	0.991
	Normal	0.959	0.94	0.99	0.992
	Average	0.968	0.96	0.98	0.99
VGG19	Benign	0.941	0.801	0.982	0.961
	Malignant	0.9549	0.94	0.956	0.98
	Normal	0.934	0.95	0.905	0.97
	Average	0.943	0.897	0.9476	0.97
Inception V3	Benign	0.9689	0.96	0.97	0.99
	Malignant	0.96	0.86	0.98	0.98
	Normal	0.9567	0.96	0.95	0.99
	Average	0.9619	0.926	0.967	0.99

V. CONCLUSION

In this study, the proposed model focuses on improving the classification outcomes of the MAIS dataset to assist medical doctors in the identification and diagnosis of BC. The images from the MAIS were divided into three groups. For noise reduction, improving contrast in mammogram images, removal of non-breast regions, and detecting the tumor location, original dataset of MIAS were pre-processed. We have mentioned the idea of data augmentation to increase the size of a dataset to enhance the CNN structure's performance. The freezing and fine-tuning techniques are then applied to enhance the tumor classification performance on the aforementioned dataset. In comparison with the other models that discussed, the VGG16 network obtained the best sensitivity, accuracy, AUC, specificity of 96%, 96.8%, 0.99, 98%, respectively. Finally, we can conclude that the VGG16 performs better than InceptionV3 and VGG19 as the generalization ability of networks without residual structure becomes worse with the increase of depth.

REFERENCES

[1] Lowry, Kathryn P., et al. "Long-term outcomes and cost-effectiveness of breast cancer screening with digital breast tomosynthesis in the United States." JNCI: Journal of the National Cancer Institute, vol. 112, pp. 582-589, June 2020.

[2] George, Sharon A. "Barriers to breast cancer screening: an integrative review." Health care for women international, vol. 21, pp. 53-65, 2000.

[3] American Cancer Society. Cancer Facts & Figures 2020. Available at: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2020/cancer-facts-and-figures-2020.pdf.

[4] Zhou, Juan, et al. "Weakly supervised 3D deep learning for breast cancer classification and localization of the lesions in MR images." Journal of Magnetic Resonance Imaging, vol. 50, pp. 1144-1151, March 2019.

[5] Zhang, Qingchen, et al. "A survey on deep learning for big data." Information Fusion, vol. 42, pp. 146-157, July 2018.

[6] A. Saber, A.M. Al-Zoghby, and S. Elmougy, "Big-data aggregating, linking, integrating and representing using semantic web technologies," International Conference on Advanced Machine Learning Technologies and Applications. Springer, Cham, Egypt, vol. 723, p. 331-342 , January 2018.

[7] Weiss, Karl, Taghi M. Khoshgoftaar, and DingDing Wang. "A survey of transfer learning." Journal of Big data, vol. 3, pp. 1-40, May 2016.

[8] Samala, R. K., Chan, H. P., Hadjiiski, L. M., Helvie, M. A., Cha, K. H., & Richter, C. D." Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms". Physics in Medicine & Biology, vol. 62(23), pp. 8894–8908, Nov 2017.

[9] Wahab, Noorul, Asifullah Khan, and Yeon Soo Lee. "Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images." Microscopy, vol. 68(3), pp. 216-233, 2019.

[10] Lotter, William, et al. "Robust breast cancer detection in mammography and digital breast tomosynthesis using annotation-efficient deep learning approach." arXiv preprint arXiv:1912.11027, pp. 1-16, Dec 2019.

[11] Jiang, Fan, et al. "Breast mass lesion classification in mammograms by transfer learning." Proceedings of the 5th international conference on bioinformatics and computational biology, pp. 59-62, January 2017.

[12] Khan, SanaUllah, et al. "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning." Pattern Recognition Letters, vol. 125, pp. 1-6, 2019.

[13] Cao, Hongliu, et al. "Improve the performance of transfer learning without fine-tuning using dissimilarity-based multi-view learning for breast cancer histology images." International conference image analysis and recognition. Springer, Cham, pp. 779-787, 2018.

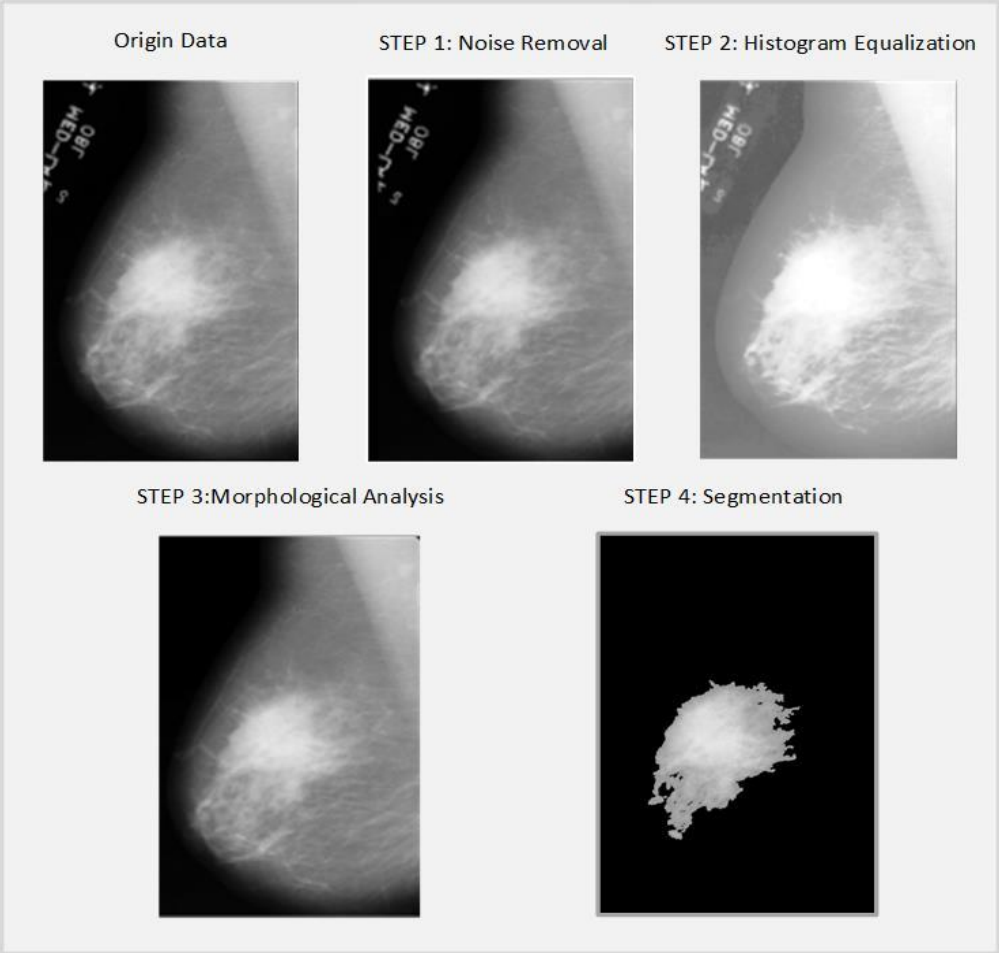


Fig.5. Data preprocessing results

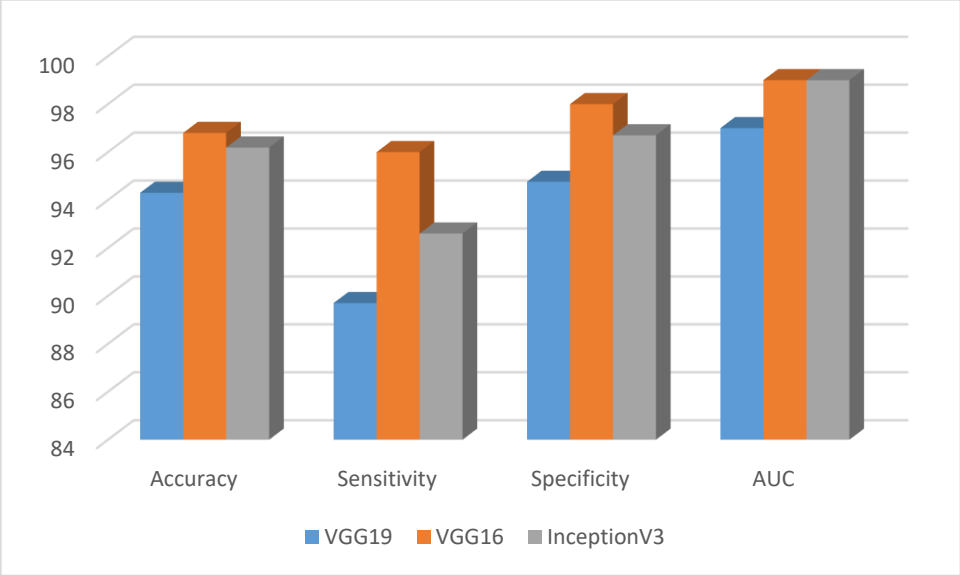
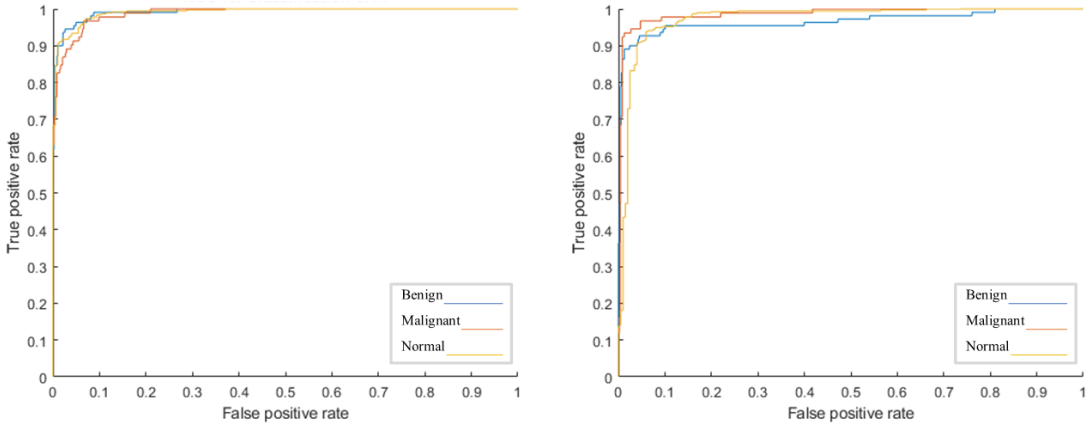
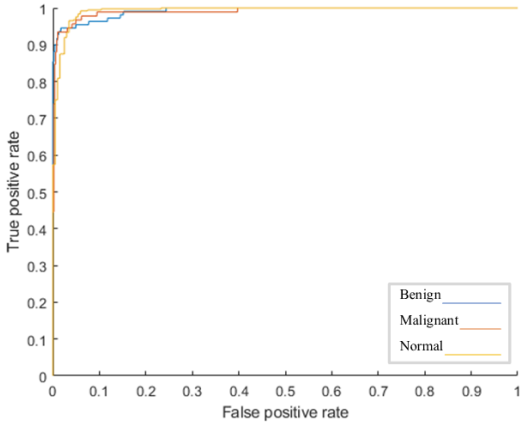


Fig. 6. Comparison between VGG19, VGG16, and InceptionV3 networks



(A) InceptionV3 (B) VGG19



(C) VGG16

Fig. 7. ROC curves and corresponding AUC values of InceptionV3, VGG16, and VGG19

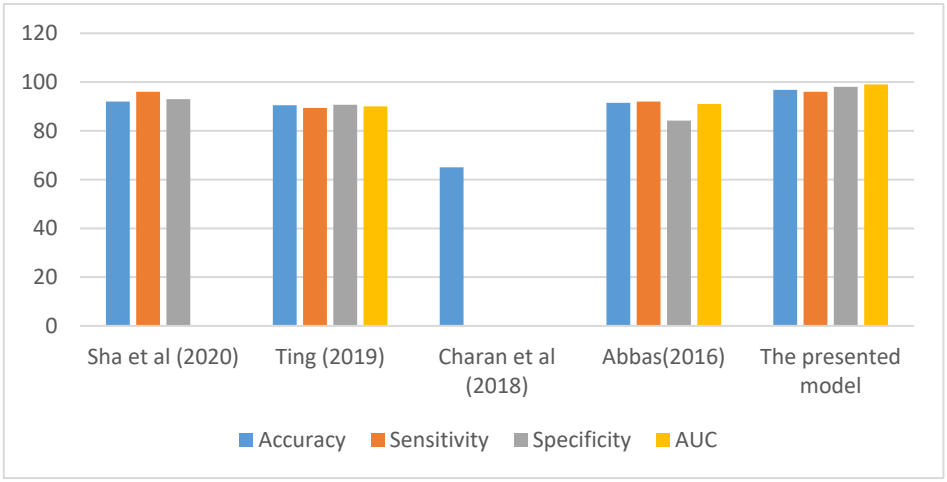


Fig. 8. Comparison between our model and existing models

[14] Deniz, E., Şengür, A., Kadiroğlu, Z., Guo, Y., Bajaj, V., & Budak, Ü. Transfer learning based histopathologic image classification for breast cancer detection. Health information science and systems, vol. 6(1), pp. 6-18, 2018.

[15] Celik, Yusuf, et al. "Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images." Pattern Recognition Letters, vol. 133, pp. 232-239, May 2020.

[16] Toğaçar, Mesut, et al. "BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer." Physica A: Statistical Mechanics and its Applications, vol. 545, 123592, May 2020.

[17] Ting, Fung Fung, Yen Jun Tan, and Kok Swee Sim. "Convolutional neural network improvement for breast cancer classification." Expert Systems with Applications, vol. 120, pp. 103-115, 2019.

[18] Abbas, Qaisar. "Deepcad: A computer-aided diagnosis system for mammographic masses using deep invariant features." Computers, vol. 5(4), pp. 1-15, October 2016.

[19] Sha, Zijun, Lin Hu, and Babak Daneshvar Rouyendegh. "Deep learning and optimization algorithms for automatic breast cancer detection." International Journal of Imaging Systems and Technology, vol. 30(2), pp. 495-506, 2020.

[20] Charan, Saira, Muhammad Jaleed Khan, and Khurram Khurshid. "Breast cancer detection in mammograms using convolutional neural network." 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). IEEE, p. 1-5, March 2018.

[21] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556, pp. 1-14, Apr 2014.

[22] Yengec Tasdemir, Sena B., Kasim Tasdemir, and Zafer Aydin. "A review of mammographic region of interest classification." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery : e1357, vol. 10(5), pp. 1-15, Feb 2020.