

Breast Cancer Classification from Ultrasound Images using VGG16 Model based Transfer Learning

A. B. M. Aowlad Hossain*

Khulna University of Engineering & Technology /Electronics and Communication Engineering, Khulna, Bangladesh
Email: aowlad0403@ece.kuet.ac.bd
ORCID iD: <https://orcid.org/0000-0002-2559-2781>
*Corresponding Author

Jannatul Kamrun Nisha

Khulna University of Engineering & Technology /Electronics and Communication Engineering, Khulna, Bangladesh
Email: jannatulkamrunnisha1725@gmail.com
ORCID iD: <https://orcid.org/0000-0001-9033-7050>

Fatematuj Johora

Khulna University of Engineering & Technology / Biomedical Engineering, Khulna, Bangladesh
Email: fatematujjohora8002@gmail.com
ORCID iD: <https://orcid.org/0000-0003-2389-0598>

Received: 29 August, 2022; Revised: 02 October, 2022; Accepted: 19 November, 2022; Published: 08 February, 2023

Abstract: Ultrasound based breast screening is gaining attention recently especially for dense breast. The technological advancement, cancer awareness, and cost-safety-availability benefits lead rapid rise of breast ultrasound market. The irregular shape, intensity variation, and additional blood vessels of malignant cancer are distinguishable in ultrasound images from the benign phase. However, classification of breast cancer using ultrasound images is a difficult process owing to speckle noise and complex textures of breast. In this paper, a breast cancer classification method is presented using VGG16 model based transfer learning approach. We have used median filter to despeckle the images. The layers for convolution process of the pretrained VGG16 model along with the maxpooling layers have been used as feature extractor and a proposed fully connected two layers deep neural network has been designed as classifier. Adam optimizer is used with learning rate of 0.001 and binary cross-entropy is chosen as the loss function for model optimization. Dropout of hidden layers is used to avoid overfitting. Breast Ultrasound images from two databases (total 897 images) have been combined to train, validate and test the performance and generalization strength of the classifier. Experimental results showed the training accuracy as 98.2% and testing accuracy as 91% for blind testing data with a reduced of computational complexity. Gradient class activation mapping (Grad-CAM) technique has been used to visualize and check the targeted regions localization effort at the final convolutional layer and found as noteworthy. The outcomes of this work might be useful for the clinical applications of breast cancer diagnosis.

Index Terms: Breast Cancer Classification, Benign vs. Malignant Tumor, Breast Ultrasound, Deep Transfer Learning, VGG16 Model

1. Introduction

Breast cancer is a common disease of women all over the world and is considered one of the leading causes of mortality [1]. Breast tumors are classified into benign and malignant depending on their histopathology and metastasis formation. Statistics shows that about 40% breast cancer mortality reduction can be done if annual screenings are performed from age forty [2]. Timely detection and hence proper treatment of breast cancer reduces the mortality rate significantly [3,4].

Different techniques are used to detect breast cancer based on pathology test, biopsy and diagnostic imaging. However imaging based diagnosis using x-ray and ultrasound has shown proven success. Though x-ray based

mammography is a common technique to diagnose the breast cancer, it has radiation hazards and false diagnosis probability of non-cancerous mass as cancer especially for dense breast [5]. The low sensitivity and specificity limitations of mammographic scanning in women with dense breast mass have initiated utilization of other imaging modalities like ultrasound imaging and magnetic resonance imaging (MRI). Expensive MRI based screening requires dye injection as contrast agent. Medical ultrasound imaging is a popular and widely used imaging modality considering its low cost, portability, safe, and real time scanning facilities. Ultrasound screening can distinguish solid mass and fluid-filled region of breast. Furthermore, the irregular shape, intensity variation and additional blood vessels of malignant cancer are also distinguishable in ultrasound images from the benign phase [6]. Improvement of high-frequency technology, especially with probes of 8–13 MHz has opened a new window of ultrasound breast imaging applications. Recently the harmonic ultrasound imaging modality and instantaneous compounding provide improved image resolution and abnormal tissue characterization capability of ultrasound scanning [7]. Technological advancement, cancer awareness, government policies and non-government campaign leads the rise of ultrasound based breast screening applications especially for dense breast. In consequence, the breast ultrasound imaging systems marked is projected to reach over USD 2 Billion globally by the year 2027 with a compound annual growth rate of 15.5% during the projection duration years 2020 to 2027 [8]. Hence, the top famous medical imaging companies are the key market players in this demanding and growing breast ultrasound market. However, manual classification of benign and malignant cancer is difficult due to speckle noise in ultrasound images, operator skills variability, and subjective assessment of medical physicians. Therefore, modern advancement in automation with the revolutionary support of artificial intelligence through machine learning is playing a great role in the improvement of automatic breast ultrasound screening.

In this classification work, we have designed a VGG16 model based deep neural network classifier and applied transfer learning strategy from a number of state-of-art pre-trained network models to classify the breast cancer. Ultrasound images of breast from two database have been used in a mixing form with a strategic emphasis on testing the generalize performance of the network. Since the data size is still not so big even though after mixing images from two databases, transfer learning through established pretrained deep network is advantageous option as classifier learning to avoid the training of the model from scratch. Transfer learning exhibits the added advantages of reducing training computation and time as well as reducing generalization error. Ultrasound speckle noise may obscure the cancerous texture of malignant stage which leads misclassifications. Median filtering can filter speckle noise preserving the edge. Therefore, we use median filtering to preprocess the ultrasound breast images. We have trained, validated, and tested the classifier with tuning to superior performance. Visual inspection of the localization capability of the convolution layers has been done through Grad-CAM. Finally, the obtained results have been compared with related state of the arts works.

The second section of this paper covers literature review on some existing methods. The third section explains the dataset and proposed methodology. The obtained results as well as the performance evaluations and comparative assessment of the classifier are described in the results and discussions section. Finally, few concluding comments are provided in the conclusion section.

2. Literature Review

Computer based automatic classification is become an obvious and demanding support to the medical experts for proper diagnosis, prognosis and treatment or controlling management of breast cancer [9]. Recent rapid advances of machine learning tools especially deep learning techniques increase the strength of computer aided diagnosis significantly [10]. Modern advancement in automation with the revolutionary support of artificial intelligence through machine learning is playing a great role in the improvement of automatic breast ultrasound systems.

Various studies have been reported on breast cancer classification using conventional machine learning and deep learning methods [11-15]. Deep learning with Convolutional neural network (CNN) has proven revolutionary performance in automatic classifications of images. However, training of CNN needs huge computational load and time. Furthermore, training with small dataset causes overfitting problem. Therefore, transfer learning of established pre-trained CNNs is performs more efficiently in many scenarios with small data sets. In transfer learning, the convolution layers of the CNN mainly perform the feature extraction process and the fully connected neural network layers perform the classification job. Wan et al. have conducted a comparative study on evaluation of conventional machine learning methods, CNN and AutoML Vision in classifying the breast cancer [13]. Among different traditional approaches Random Forest classifier was found with highest accuracy of 90% whereas the accuracy of their CNN was found as 91%. Singh et. al. used conditional generative adversarial network and found 85% accuracy [13]. Xiao et al. have compared three transfer learnt models named InceptionV3, ResNet50, and Xception along with three traditional machine learning models and found that transfer learning performs better than traditional learning [15]. Among the three deep models in their study, the InceptionV3 model has individually showed the best accuracy of 85% which increases to 89% when the combination of features of all there models was considered. After the background study through literature survey it is seen that, accurate detection and classification is still challenging and there are scopes to improve the computer aided diagnosis performance. Because of different physical, anatomical or physiological nature of breasts, variations of cancers occur even in the same breasts which lead the diagnosis troublesome. Deep learning

with sufficient data can improve the detection and classification accuracy. Furthermore, speckle noises of ultrasound images also reduce the accuracy rate. Effective image preprocessing method can increase the accuracy. Being motivated by these facts, in this research work, effective preprocessed breast ultrasound data based deep learning approach was planned to develop for breast cancer classification.

3. Methodology

The flowchart of the proposed methodology for breast cancer classification is shown in Fig. 1. Firstly, the raw ultrasound breast images are preprocessed using median filtering based despeckling and then resized and normalized. Then VGG16 model is used as feature extractor and the proposed model classify the breast cancer from the extracted features using transfer learning concept. The preprocessing, network design, training strategy, and evaluation process are explained in the subsequent subsections.

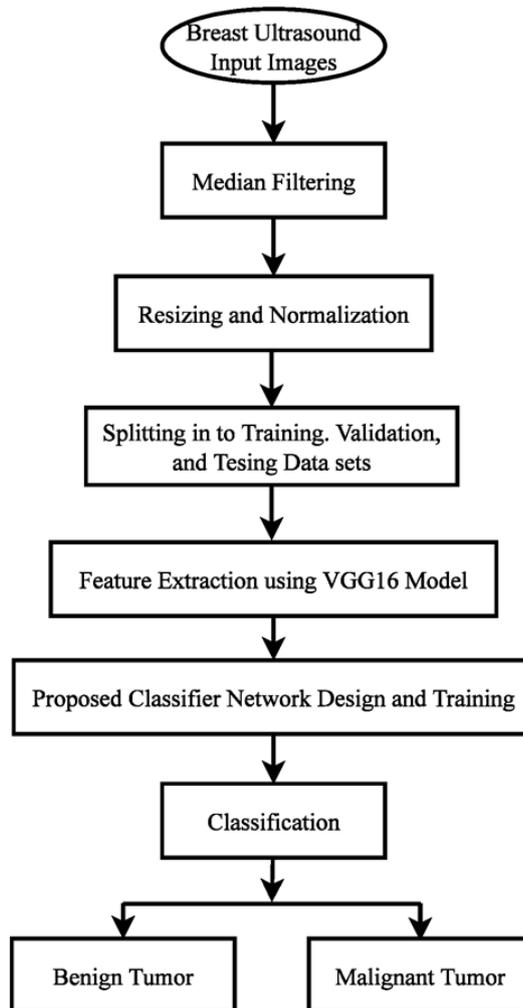


Fig. 1. Overview of the Proposed Methodology.

A. Description of Breast Ultrasound Data

The breast of female is mainly composed of fat, glandular tissues of ducts and lobules and Cooper's ligaments fibers. These look as separate layers such as skin, beneath subcutaneous fat, parenchyma with ducts and lobules, retroglandular fat, pectoralis muscles, and chest wall [16]. The volume of fat mainly defines the size of the breast. Dense breasts have more glandular and fibrous tissue and less fatty tissue. Breast cancer generated in the anatomical structures of the breast especially in the glands and fibrous tissue [17]. The fibrocystic changes that cause noncancerous breast lumps due to the fluctuation of hormone levels are called benign cancer. The benign type mammary gland tumors are usually papillomas and fibroadenomas.

Benign type breast abnormality is not cancerous spreading or life-threatening but they tend to increase the risk of developing cancer in future. Malignant tumors are cancerous which eventually spread outside the original tumor to other parts of the body. The most available malignant tumors types are carcinomas. These two subtypes are almost

40%–75% of all diagnosed cases [18]. Breast cancer is the common and life threatening diseases of breast. Accurate detection of breast cancer through different medical imaging techniques is of vital importance for proper treatment and elevating the rate of survival. Recently breast ultrasound imaging show promises in breast cancer screening especially for dense breast.

In this research work, the breast cancer datasets have been collected from two publicly available databases named Mendeley BUS dataset [19] and Breast Ultrasound Images (BUSI) dataset of Cairo University [20]. BUSI dataset consists of 133 normal breast, 437 benign phase and 210 malignant stage ultrasound images of 600 female's breast of age 25 to 75 years old. Mendeley dataset consists of 100 benign and 150 malignant tumorous breast ultrasound images. Both databases are used in different studies on detection and classification of breast cancer. The variation of BUSI dataset is more than Mendeley dataset. As mention earlier, our strategic plan is to mix these two dataset in order to test the generalization capability of our proposed classifier. Since Mendeley dataset has no normal images except the benign and malignant cases, normal images of BUSI dataset has been discarded during this study and hence target classes are set as benign and malignant.

Since ultrasound echoes are come from heterogeneous layers, the layers of breast are visible in ultrasound image. Fat looks darkish gray in ultrasound images. The Ducts and cysts inside breast are anechoic in nature, the blood vessels looks hypoechoic; on the contrary parenchyma, Cooper's ligaments, and skin surface are usually hyperechoic [21]. Figure 2 shows the typical ultrasound images of normal and tumorous breasts. Generally a benign phase of the breast is characterized as smooth, uniform and well bounded outline with hyperechoic, isoechoic or slightly hypoechoic ultrasound nature [22]. In ultrasound images, malignant tumors are usually hyperechoic region with ill-defined edges. The hyperechoic nodular shape is usually taller and having speculated borders and posterior shadow [23].

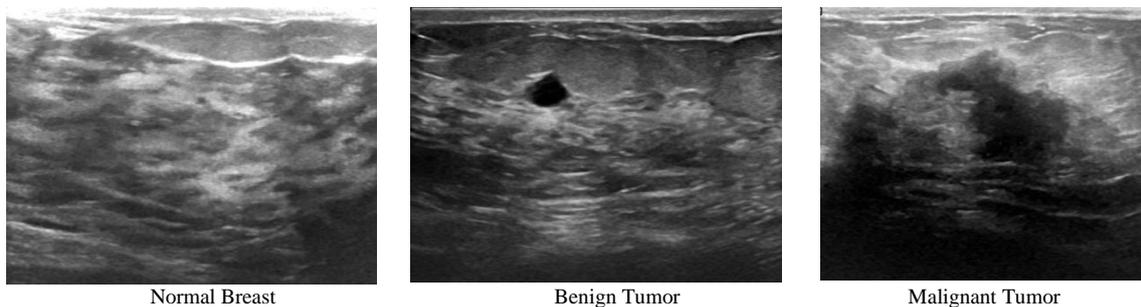


Fig. 2. Normal and Tumorous Breast Images [20].

B. Preprocessing

Breast ultrasound images have inherent speckle noise. Since the median filter is a simple but effective method for despeckling of ultrasound images preserving the edges, median filter has been used for preprocessing the ultrasound image to improve the image quality. Median filter is a nonlinear statistic filter. In median filtering of breast image, the center pixel of a 5×5 moving kernel is substituted by the median number of the corresponding kernel. The obtained median filtered images are then resized to $224 \times 224 \times 3$ considering the input image size of VGG16 model. We have then normalized the pixel intensity of the input images so that the model convergence to the solution rapidly. The normalized images are then splitted by 70% as training dataset, 15% as validation dataset, and remaining 15% as blind test dataset.

C. Feature Extraction and Network Model Design

As mentioned earlier, we have used transfer learning using VGG16 model to classify the benign and malignant tumor considering the data size limitation. Since VGG16 is trained on large scale diverse and general dataset of the order of millions of images, it will efficiently aid as a generic model of the visual realm and hence this pretrained model can detect complex visual features of breast tumor. Transfer learning exhibits the added advantages of reducing training time as well as reducing generalization error. We can take the benefits of the learned feature maps without starting from the scratch. The pretrained VGG16 model excluding the last fully connected layers is used as feature extractor which is ended by our proposed neural network for classification task.

VGG16 model is proposed by Simonyan and Zisserman as abbreviated form of Visual Geometry Group- 16 to recognize 1000 categories which is a subset of 22000 categories from over 14 million images of ImageNet dataset [24]. VGG16 model is based on deep convolutional network. Convolutional neural network is a popular and powerful deep learning technique which is mainly applied for computer vision applications like image classification, recognition, and segmentation tasks. Deep convolutional neural network is an updated form of multi-layer perceptrons. Recent revolutionary growth of GPUs has facilitated the researchers to design and train more complex and deeper network [25]. CNN consists of different layers with varying characteristics like convolutional layers for feature extraction, pooling layer to reduce dimensionality and fully connected layer for classification tasks along with different neuronal activation functions like sigmoid, hyperbolic, softmax, and rectified linear unit functions.

We used the Convolutional and Pooling layers of VGG16 architecture to create the feature dataset for input of our proposed customized fully connected network. Figure 3 represents the diagram of the VGG16 model based feature extractor and the proposed classifier model. Furthermore, the layer type and number of parameters of each layer of the model are summarized in the Table 1.

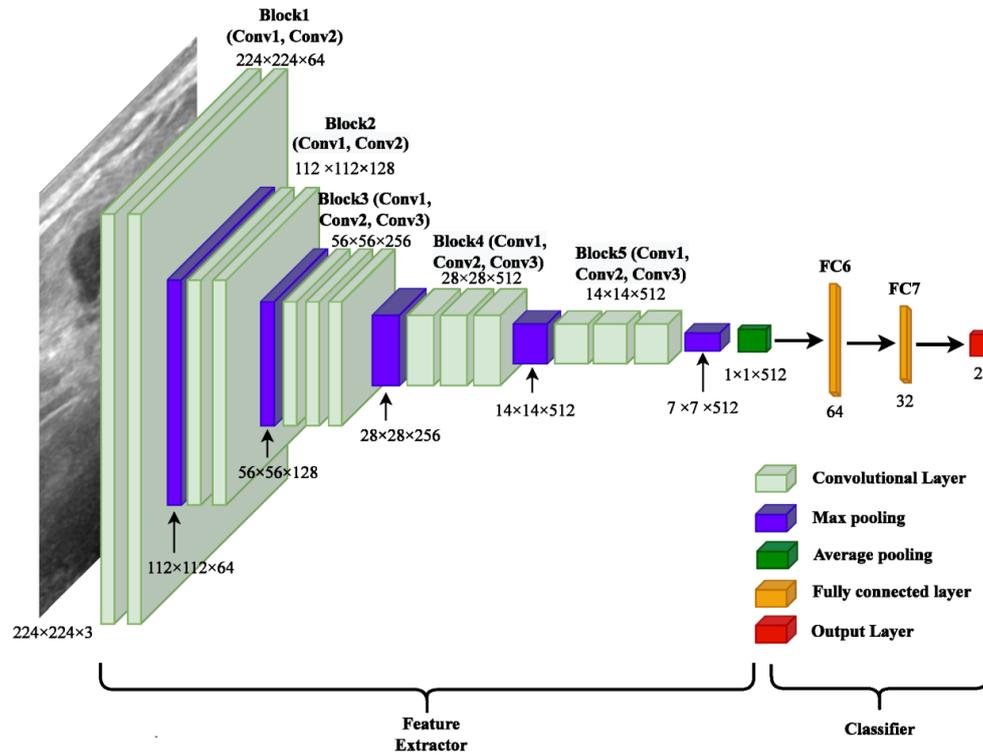


Fig. 3. VGG16 Feature Extractor and Proposed Neural Network Classifier

Table 1. Layer-wise Number of Parameters of the Proposed Model

Layer and Type	Number of Parameters
Input Layer	0
Block1 (Conv1)	1792
Block1 (Conv2)	36928
Block1 (MaxPool)	0
Block2 (Conv1)	73856
Block2 (Conv2)	147584
Block2 (MaxPool)	0
Block3 (Conv1)	295168
Block3 (Conv2)	590080
Block3 (Conv3)	590080
Block3 (MaxPool)	0
Block4 (Conv1)	1180160
Block4 (Conv2)	2359808
Block4 (Conv3)	2359808
Block4 (MaxPool)	0
Block5 (Conv1)	2359808
Block5 (Conv2)	2359808
Block5 (Conv3)	2359808
Block5 (MaxPool)	0
AveragePool	0
Flatten	0
Dense	32832
Dropout	0
Dense1	2080
Dropout1	0
Dense2	66

The three dimensional feature stack of the last max-polling layer of the VGG16 feature extractor is firstly averaged pooled and then flatten to a vector of 512 inputs. The fully connected hidden layers of our proposed network consist of 64 neurons and 32 neurons which are chosen on trial and error basis. Dropout is a simple and strong generalization method for deep networks where randomly chosen neurons are excluded during the training phase. Dropout can prevent overfitting problem. The 1st and 2nd hidden layers are dropout to 20% and 50% respectively in order to minimize overfitting. Considering our targets classes as benign and malignant, the output layer of the proposed classifier consists of two neurons.

The activation function for hidden layer neurons is set as Rectified linear unit and softmax activation function were chosen for the output layer neurons. During the training, the pretrained convolutional layers are kept freeze and non-frozen customized proposed top network layers are kept trainable. As an impact of this freezing, the total number of trainable parameters is reduced from 14,749,666 to 34,978 which decrease the training computation and time significantly. Adam optimizer is used setting 0.001 as learning rate and the binary cross-entropy is chosen as the loss function for model optimization. The settings of the dropout rate, the rate of learning, the size of batch and the number of epochs are tuned to the optimum values on trial and error basis by tuning and running the network several times. The model was trained with a dataset of 376 benign images and 252 malignant images with a batch size of 16 per epoch. The number of epoch is set as 100 considering the training and validation learning curves characteristics.

4. Results and Discussions

The preprocessed median filtered and resized images along with the original form from both databases are shown in Fig. 4. The PSNR of the filtered images with respect to the original image was found 27 dB. It was observed that the classifier performance was improved for median filtered images input in comparison with the raw images input. Therefore, the results of classifiers' performance with median filtered image input are analyzed and discussed in this section.

The model has been trained using 70% images from the mixture of both databases. The training accuracy reached to 98.2% as maximum and the validation accuracy reached to 95% as maximum and reach to a saturation stage at about 100 epochs. However, the validation curve started diverging after 100 epochs reaching its maximum. Therefore, we have saved the model for testing purpose at 100 epochs. The accuracy and loss curve during training and validation process are shown in Fig. 5. The proposed model has been implemented and verified using python language with Keras deep learning libraries in Google Colaboratory platform. The simulation task was executed in a computer with a configuration of Intel(R) Core(TM) i7-4790 CPU @3.60GHz including Google Colaboratory RAM and GPU support.

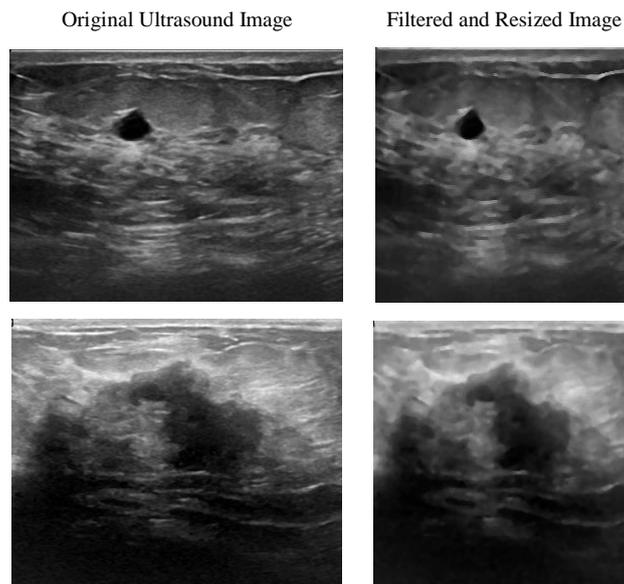


Fig. 4. Median Filtered and Resized Image along with their Original Form

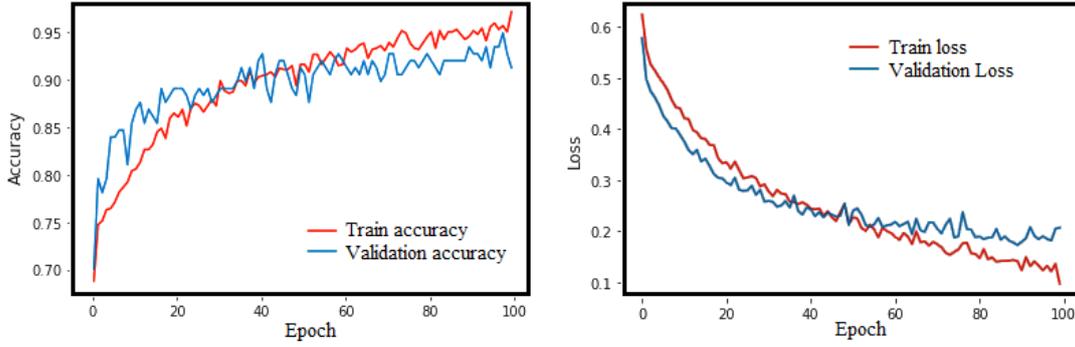


Fig. 5. Accuracy and Loss Curve during Training and Validation Process.

In computer aided medical diagnosis, different performance parameters are usually considered to assess the performance of the classifier system. We have considered the evaluation metrics [26] taking into account the obtained True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN) values and calculated the accuracy, sensitivity (recall), specificity, and F1 score to assess the performance of the proposed method and to compare with related previous works. True positive indicates the model predict properly the malignant case, false positive indicates that the malignant case is predicted wrongly, true negative means that the benign case is accurately predicted by the model, and false negative means the model predicts the benign case wrongly. A confusion matrix displays the prediction statistics of the classifier on the blind test dataset is shown in Fig. 6. It is shown that among the 80 benign and 52 malignant cases of the blind test dataset, the classifier predicts 49 cases as TP, 71 cases as TN, 9 as FP case and 3 cases as FN.

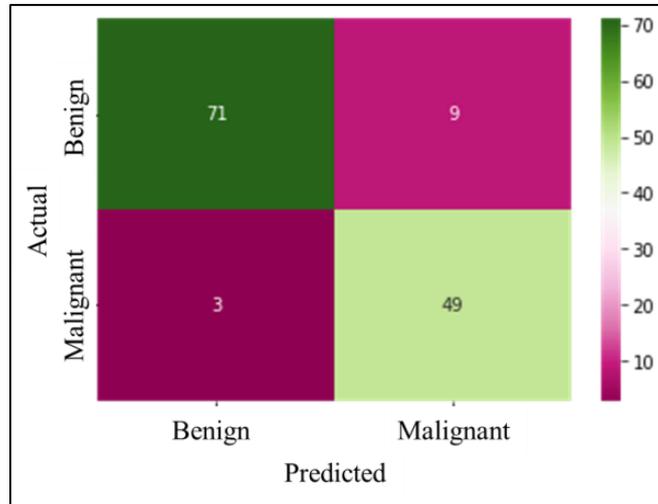


Fig. 6. Confusion Matrix of the Prediction Statistics on Blind Test Dataset

The common performance measures of the classifiers like accuracy, sensitivity, specificity, and F1 score are defined in terms of TP, FP, TN, and FN by the expression (1), (2), (3), and (4) respectively.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

$$F1\ Score = \frac{2TP}{2TP + FN + FP} \tag{4}$$

A comparative study has been showed to assess the overall performance of the proposed VGG16 model based classifier with other related works on the same and different dataset. Table 2 represents a comparative view of the performances of different relevant classifiers with the proposed model. In the study of Wan et. al. [13] with same source

data of our study, it was shown that among different conventional machine learning techniques Random Forest classifier outperformed with an accuracy of 90%. They also showed a CNN deep learning model given accuracy of 91%. However, it is necessary to train the CNN model from scratch which takes more times for learning. Singh et. al. found an accuracy of 85% from only Mendeley dataset using deep adversarial learning method [14]. In another comparative study with larger and different dataset, Xiao et. al showed that the transfer learning InceptionV3 given the better accuracy of 85.13% which was better than their used other traditional machine learning methods as well as a CNN model [15]. Furthermore, the Precision = $TP/(TP+FP)$ of the proposed classifier are also calculated and found as 84.5% respectively. Therefore, from the comparison, shown in the Table 2, it is seen that the proposed model shows satisfactory accuracy and noteworthy sensitivity and F1 score comparing with conventional techniques of machine learning and the state of art deep learning studies.

Table 2. Comparison of the Performance of the Proposed Method with other Studies

Ref.	Data	Method	Accuracy	Sensitivity	Specificity	F1 Score
K. W. Wan et. al. [13]	BUSI + Mendeley Dataset, 895 Images	kNN	84%	75%	88%	0.76
		Random Forest	90%	71%	100%	0.83
		SVM	73%	57%	81%	0.59
		CNN	91%	82%	96%	0.87
		AutoML Vision	85%	84%	88%	0.83
V. K. Singh et. al. [14]	Mendeley Dataset, 250 Images	Conditional generative adversarial network	85%	92%	--	--
T. Xiao et. al. [15]	Hospital of Sun Yatsen University Data, 2058 Images	Fine-tuned ResNet50	84.94%	77.39%	88.74%	0.78
		Fine-tuned InceptionV3	85.13%	77.44%	89.06%	0.78
		Xception	84.06%	77.44%	87.16%	0.76
		CNN	74.44%	63.19%	79.22%	0.60
		AdaBoost with LDA	70.55%	58.10%	74.61%	0.49
Proposed Method	BUSI + Mendeley Dataset, 897 Images	Transfer learning with VGG16	91%	94%	88.75%	0.89

The receiver operating characteristics (ROC) curves is another widely used performance measure which is a graphical representation reflecting the discriminatory strength of the classifier. The true positive rate (sensitivity) is plotted as a function of the false positive rate (1-specificity) to find the ROC curve [26,27]. Therefore a better performance characteristic is found rather the single valued performance index. Figure 7 shows the ROC curve of the proposed classifier. The area under the curve (AUC) is found as 0.97 which means the benign and malignant cancer cases are excellently separated by the classifier.

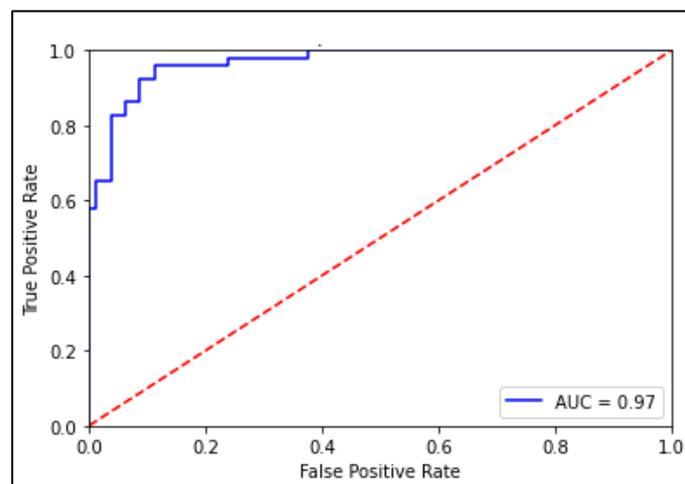


Fig. 7. ROC Curve of the Proposed Classifier.

Gradient-weighted class activation mapping (Grad-CAM) is an excellent visual support of finding the localizing capability of the targeted regions at the final convolutional layer [28]. We have computed the Grad-CAM at the final convolutional layer (Conv5) at viridis color scale. Figure 8 shows the Grad-CAM superimposed images for both benign and malignant cases along with their original input form. The Grad-CAM localizes the hyperechoic region as well as the ill-defined edges of malignant tumors which indicates that the feature extraction using the proposed model's convolution layers can successfully extract the suitable features for the fully connected network classifier. On the other hand, a Grad-CAM map for the benign cases is almost distributed indicating the non-spreading cancer with more tissue homogeneity. Therefore, the Grad-CAM visualization is showing that the convolutions layers framework of the feature extractor of the proposed pertained VGG16 model is well supported as the fully connected neural network classifier input. Considering the above-mentioned findings and discussions, it can be mentioned that the overall performance of the proposed classifier is satisfactory and hence it might be useful for clinical applications.

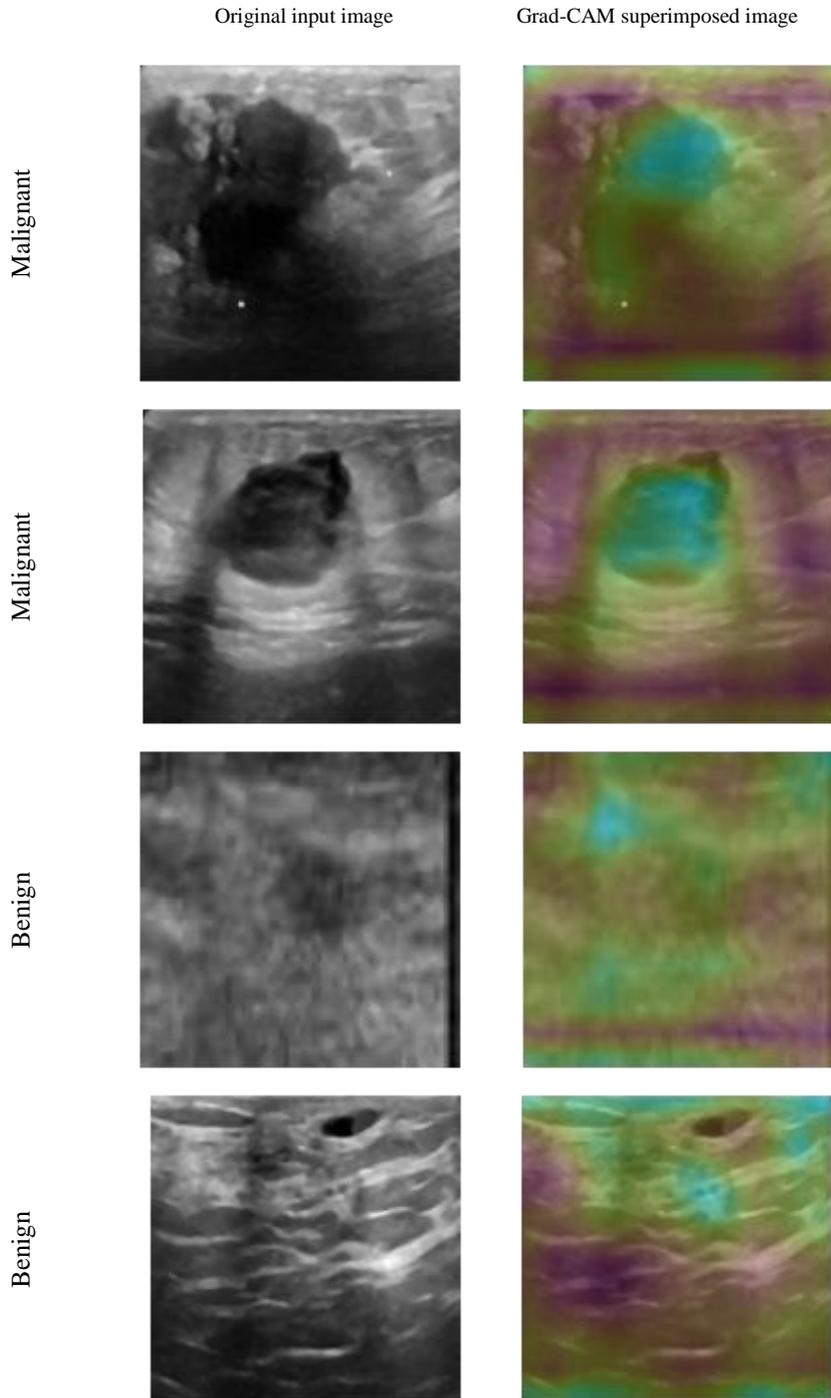


Fig. 8. Grad-CAM Superimposed Images for both Benign and Malignant Cases with their Original Input Form

5. Conclusion and Future Work

Ultrasound based automatic breast cancer screening system is of great importance for women health. In this paper, a two layers deep neural network has been proposed to classify breast cancer using the extracted features of a pretrained VGG16 model features extractor from breast ultrasound images. Median filter has been used to despeckle the raw ultrasound images which are then resized and normalized. Dropout of the fully connected layer has been used to prevent overfitting problem. Breast Ultrasound image from two databases (total 897 images) has been combined to train, validate and test the performance of the classifier. Experimental results showed satisfactory training accuracy as 98.2% and testing accuracy as 91% with blind testing dataset comparing with similar works in the literature. The Grad-CAM visualization evidences of the targeted regions localizing tendency at final convolutional layer is also found noteworthy. Though two different datasets are mixed together, still the size of the merged dataset is small. In future, the proposed network will be modified and tested using larger datasets including other cancer subtypes. Finally, this work and its future progression will be expected to help for widespread clinical applications of breast cancer diagnosis using portable and low cost ultrasound machine even in rural areas.

References

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA: A Cancer Journal for Clinicians*, vol. 70, no. 1, pp. 7-30, 2020.
- [2] E. K. Arleo, R. E. Hendrick, M. A. Helvie, and E. A. Sickles, "Comparison of recommendations for screening mammography using CISNET models," *Cancer*, vol. 123, pp. 3673–3680, 2017.
- [3] M. L. Giger, N. Karssemeijer, and J. A. Schnabel, "Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer," *Annual Review of Biomedical Engineering*, vol. 15, pp. 327–357, 2002.
- [4] H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: a survey," *Pattern Recognition*, vol. 43, no. 1, pp. 299–317, 2010.
- [5] E. Devoli-Disha, S. Manxhuka-K"erliu, H. Ymeri, and A. Kutllovci, "Comparative accuracy of mammography and ultrasound in women with breast symptoms according to age and breast density," *Bosnian Journal of Basic Medical Sciences*, vol. 9, no. 2, pp. 131–136, 2009.
- [6] B. Sahiner, H.-P. Chen, M. A. Roubidoux, L. A. Hadjiiski, M. A. Helvie, et al., "Computer-aided diagnosis of malignant and benign breast masses in 3D ultrasound volumes: Effect on radiologists' accuracy," *Radiology*, vol. 242, no. 3, pp. 716–724, 2007.
- [7] C. R. Merritt, "Technology update," *Radiologic Clinics of North America*, vol. 39, pp. 385-397, 2001.
- [8] Marketwatch, Automated Breast Ultrasound System Market Share, Upcoming Trends, Size, Key Segments, Growth Status and Forecast 2027, Retrieved from <https://www.marketwatch.com> on March 23, 2022.
- [9] M. H. Yap, E. Edirisinghe, and H. Bez, "Processed images in human perception: A case study in ultrasound breast imaging," *European Journal of Radiology*, vol. 73, no. 3, pp. 682–687, 2010.
- [10] J. Yanase and E. Triantaphyllou, "A systematic survey of computer-aided diagnosis in medicine: Past and present developments," *Expert Systems with Applications*, vol. 138, pp. 112-821, 2019.
- [11] H.D. Cheng, J. Shan, W. Ju, Y. Guo, L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern Recognition*, vol. 43, pp. 299-317, 2010.
- [12] C. D. L. Nascimento et. al., "Breast tumor classification in ultrasound images using support vector machines and neural networks," *Research on Biomedical Engineering*, vol. 32, no. 03, 2016.
- [13] K. W. Wan, C. H. Wong, H. F. Ip, D. Fan, P. L. Yuen, H. Y. Fong, and M. Ying, "Evaluation of the performance of traditional machine learning algorithms, convolutional neural network and AutoML Vision in ultrasound breast lesions classification: A comparative study," *Quantitative Imaging in Medicine and Surgery*, vol. 11, pp. 1381-1393, 2021.
- [14] V. K. Singh, H. A. Rashwan, M. Abdel-Nasser, M. M. K. Sarker, F. Akram, N. Pandey, S. Romani, and D. Puig, "An efficient solution for breast tumor segmentation and classification in ultrasound images using deep adversarial learning," Preprint arXiv:1907.00887, 2019.
- [15] T. Xiao, L. Liu, K. Li, W. Qin, S. Yu, and Z. Li, "Comparison of transferred deep neural networks in ultrasonic breast masses discrimination," *BioMed Research International*, vol. 9, 2018.
- [16] M. D. Barber, *Breast Cancer: An Atlas of Investigation and Management*, Clinical Publishing, 2008.
- [17] N. Harbeck et al., "Breast cancer," *Nature Reviews Disease Primers*, vol. 5, no. 1, 2019.
- [18] E. A. Rakha et al., "Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation," *The Journal of Pathology*, vol. 208, no. 4, pp. 495-506, 2006.
- [19] P. S. Rodrigues, "Breast ultrasound image," *Mendeley Data*, V1, 2017.
- [20] W. Al-Dhabyani, M. Goma, H. Khaled, and A. Fahmy, "Dataset of breast ultrasound images." *Data Brief*, vol. 28, no. 104863, 2019.
- [21] J. A. Rapelyea and C. G. Marks, *Breast Imaging: Breast Ultrasound Past, Present, and Future*. Intech Open, 2017.
- [22] A. T. Stavros, D. Thickman, C. L. Rapp, M. A. Dennis, S. H. Parker, and G. A. Sisney, "Solid breast nodules: use of sonography to distinguish between benign and malignant lesions," *Radiology*, vol. 196, pp. 123-134, 1995.
- [23] A. T. Stavros, "Benign Solid Nodules: Specific pathologic diagnosis," *Breast Ultrasound*, vol. 13, pp. 528–596, 2004.
- [24] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Proceedings of the third International Conference on Learning Representations*, 1-14, 2015.
- [25] A. Krizhevsky I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp.1106-1114, 2012.

- [26] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168-192, 2021.
- [27] K. H. Zou, J. O'Malley, and L. Mauri, "Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models," *Circulation*, vol. 115, no. 5, pp. 654-657, 2007.
- [28] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336-359, 2020.

Authors' Profiles



A. B. M. Aowlad Hossain is a Professor in the Electronics and Communication Engineering Department of Khulna University of Engineering & Technology (KUET), Bangladesh since 2015. He obtained Doctor of Philosophy in Biomedical Engineering with a major in medical imaging from the Kyung Hee University of South Korea in the year 2012. His MSc in Electrical and Electronic Engineering (EEE) was completed in Bangladesh University of Engineering & Technology (BUET) in 2005 and BSc in EEE was completed from KUET in 2002. His research interests are on Medical Imaging Systems, Machine Learning, Computer Aided Diagnosis, Modeling and Simulation of Physiological Systems etc. About 60 research papers have been published by Prof. Hossain in different peer reviewed journals and proceedings. Prof. Hossain has got Awards for Best Papers in two IEEE co-sponsored international conferences. Prof. Hossain serves in technical committees of various conferences and journals. Prof. Dr. Hossain is senior member of IEEE and fellow of the Institution of Engineers Bangladesh.



Jannatul Kamrun Nisha obtained BSc in Electronics and Communication Engineering from KUET, Bangladesh in 2022. She is conducting researches on the applications of artificial intelligence for computer aided disease diagnosis.



Fatematuj Johora obtained BSc in Biomedical Engineering from KUET, Bangladesh in 2020. Currently, she is continuing her MSc Engineering at the same department and. Her research interest is on Medical Imaging Systems and Automated Diseases Detection.

How to cite this paper: A. B. M. Aowlad Hossain, Jannatul Kamrun Nisha, Fatematuj Johora, "Breast Cancer Classification from Ultrasound Images using VGG16 Model based Transfer Learning", *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, Vol.15, No.1, pp. 12-22, 2023. DOI:10.5815/ijigsp.2023.01.02