**ABSTRACT**

Growing interest has been shown in computer-aided identification of malignant breast cancers in ultrasound pictures. This article suggests a deep learning strategy to solve this issue. A deep convolutional neural network was trained using the training data, which includes several hundred photos of benign and malignant cases (CNN). Three methods of training are suggested: a baseline method, which involves training the CNN architecture from scratch; a transfer-learning method, which involves training the pre-trained VGG16 CNN architecture with ultrasound images; and a fine-tuning learning method, which involves adjusting the deep learning parameters to avoid overfitting. The experimental results show that the fine-tuned model, with pre-training on US pictures, had the best performance (0.93 accuracy).

**Keywords: - reast lesion; ultrasound; convolutional neural networks; deep learning; transfer learning.**

**I. INTRODUCTION**

The most frequent cancer in women worldwide is breast cancer. The National Cancer Institute of the United States of America anticipated that there will be 268,270 new cases of breast cancer in 2018. Additionally, 38.8% of Egyptian women who have been diagnosed with cancer have breast cancer.

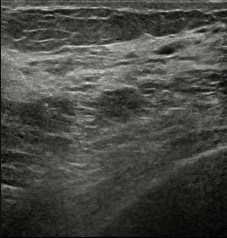
A second diagnosis can be made in the event of a suspicious result. However, mammography can occasionally detect a suspicious area that isn't cancer, causing un-needed tension and occasionally treatments. Find out how far the cancer has spread by injecting a dye into the patient prior to an MRI scan. A solid mass and a cyst filled with fluid can be distinguished one from the other using an ultrasound scan. As seen in Fig. 1, some characteristics in an ultrasound breast image aid in determining whether a breast tumour, nodule, or lump is benign or malignant. A benign ultrasound breast image can be characterised by smooth, highly regular appearance, a brighter overall appearance, and one or two smooth lobulations. On the other hand, a malignant ultrasound breast image typically has angular margins where the mass edges are difficult to discern, a mass shape with more height than width, or an increased presence of blood vessels.

Deep convolutional neural networks (CNN) are now widely used in applications for computer-aided identification and machine learning. When it comes to datasets based on nature, the depth and breadth of CNN capabilities can be controlled when the accuracy of assumptions is great. On the other hand, its value for classification, segmentation, and detection has been demonstrated in relation to medical imaging. However, there are some drawbacks because there aren't any deep learning models that have already been trained on medical data or a lot of image training data. Transfer learning is employed in two ways to aid in overcoming these disadvantages. First, using the pre-trained architecture as a feature extractor will reduce the dimensionality of the dataset when it is used as an input to a narrow classifier. Second, by carefully adjusting which layers can be specified to freeze, other layer’s learning can be enhanced, and the output size of the end layer can be modified.

We assume that more investigation into CNN models created from scratch and CNN models created using natural images is necessary due to the small number of domain-specific CNN architectures.

## (a)

## (b)

**Figure 1: Samples of the ultrasound breast lesion images used in this paper (a) Malignant, (b) Benign.**

Medical photos are different from other natural photographs in terms of texture and low contrast, and as a result, models like VGG-Net are anticipated to perform less accurately. For transfer learning on VGG-Net, they employed mammography (MG) and MRI datasets of 282 and 123 pictures, respectively. Due to this, we intend to expand on their findings in this study and conduct an experiment utilising 1,000 ultrasound images that have been enhanced using a CNN trained from scratch, VGG transfer learning, and a CNN with fine-tuned transfer learning.

**II. RELATED WORK**

Breast cancer detection has been the subject of numerous articles and numerous methods have been suggested. Brief reviews of many deep learning techniques are given here, including transfer learning, FCN-AlexNet, patch-based LeNet, U-Net, Nakagami parametric imaging with CNN, and adjusted VGG-Net.

Two ultrasound image datasets (A and B) were created by Yap et al., who demonstrated that a transfer-learning FCN-AlexNet technique had the maximum accuracy (0.92 for Dataset A, 0.88 for Dataset B). Results on alternative designs were comparable, including the U-Net technique and patch-based LeNet (0.91 for Datasets A and B) (0.89 for Dataset A, 0.78 for Dataset B). Deformable Part Models (DPM) and other deep learning techniques were used to evaluate their work. The need for such measures, according to Yap et al., was presented as a lengthy training process and unfavourable pictures.

Nakagami parameter maps were used by Byra et al. to train a CNN with an area under the receiver operating characteristic curve (AUC) of 0.91. They categorised 458 RF ultrasound pictures in their investigation. They came to the conclusion that quantitative ultrasound parameters (QUS) can influence the Nakagami method, and they verified that the CNN classification is directly connected with the image quality.

CNN was used by Hepsag et al. on the mini-MIAS and BCDR data sets. Performance measurements for the baseline tests ranged from 60 to 72%. The accuracy on the BCDR dataset increased from 65 to 85% when preprocessing methods (such as cropping and data augmentation) were performed. On the mini-MIAS dataset, similar improvements were attained.

A deep belief network (DBN) was employed by Abdel-Zaher and Eldeib to categorise breast cytology images. Their results were 99.68% accuracy, 100% sensitivity, and 99.47% specificity. They used DBN in an unsupervised manner to identify the Wisconsin Breast Cancer Dataset's features (WBCD). Additionally, a supervised phase was created using a back-propagation deep architecture that takes advantage of the conjugate gradient and Levenberg-Marquardt optimization methods. The authors verified that for both the conjugate gradient and the Levenberg-Marquardt algorithms, the accuracy of their proposed network model (DBN- NN) is superior to that of the randomly initialised weight backward propagation NN (RIW-BPNN).

Three CNN models—de novo training, cross-domain transfer learning using VGG-Net, and cross-modal transfer learning utilising mammography images—were examined by Hadad et al. for the categorization of MRI breast lesions. Accuracy values reported by the authors were 0.94, 0.90, and 0.93, respectively. Mammogram images (282 augmented to 32,064) and MRI images were employed in the authors' augmented datasets (123 augmented to 19,316). First, they used a CNN that had been completely trained by MR images. Second, they experimented with a tailored CNN through VGG-Net. Finally, they used MG-Net to develop a refined CNN.

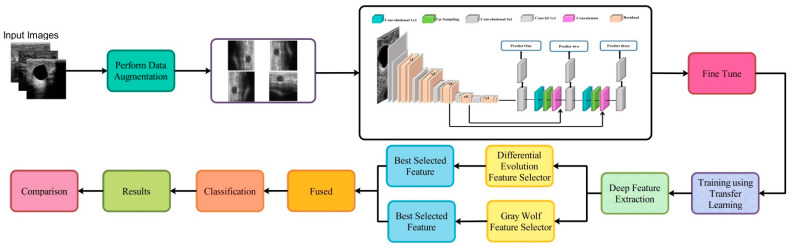
In our research, we use three different deep architectures to classify breast lesion images: a CNN architecture trained from scratch with an augmented ultrasound dataset, a transfer learning VGG16 architecture further trained with US images, and a fine-tuned architecture where the deep learning parameters are adjusted for improved performance.

**III. MATERIALS AND METHODOLOGY**

The data used in this study was gathered in Egypt at the Baheya Foundation for Breast Cancer Treatment. A broad-spectrum linear matrix array probe was used with the GE Ultrasound LOGIQ E9 XDclear device to produce ultrasound images (ML6-15-D). Expert clinicians collected the data after marking the lesions and their margins in accordance with the BI-RADS scoring criteria with the blessing of the Baheya Ethics Committee and the Egyptian Ministry of Health and Population.

The Keras library, a straightforward and effective deep learning package, was used to create the classification model in this study. In order to categorise tumours, a small network was trained from scratch (as a baseline model) using Python version 3.5 and a Theano backend. 1000 pictures, 50 epochs, a batch size of 20, and a learning rate of 0.001 were used for the training. There were two more network models employed. A deep convolutional neural network that has already been trained is the initial model, called VGG16 (CNN). More than 14 million photos from 1000 classes were used to train VGG16. The second model is a refined version created using the VGG16 pre-trained CNN model's preserved features. Rectified linear units (ReLU), pooling, dropout layers, linked layers, and loss layers are the fundamental components of CNN.

In this section, a potential framework for classifying breast cancer using ultrasound pictures is described. The suggested framework's architecture is shown in Figure 2. The original ultrasound images are subjected to preliminary data augmentation before being sent for training to the honed deep network DarkNet53. TL is used for training, and the global average pool layer's characteristics are extracted. Reformed feature optimization techniques, such as the reformed RDE and reformed RGW algorithms, are used to refine the extracted features. A probabilistic method is utilised to merge the best-chosen characteristics. Finally, machine learning classifiers are employed to categorise the fused features. Below is a description of each stage in further depth.



**Figure 2: - Proposed framework for breast cancer classification using ultrasound images**

**A. Dataset**

For our research, we gathered a collection of 650 high-resolution ultrasound images in JPEG format, each measuring 960 x 720 pixels. The dataset was split into a test set of 148 images and a training set of 497 images (100 benign and 48 malignant) (337 benign and 160 malignant).

Expanding the dataset and reducing class imbalance are the first steps in our pipeline, which also include rotation (90, 180, and 270 degrees), width and height shifts, rescaling, shearing, zooming, horizontal flipping, and filling. The significance of this stage lies in its capacity to multiply our data by a factor of 24 to yield a total of 21,600 images, which will enhance the generalisation performance. 70% of the expanded dataset was used for training, while 30% was used for testing. Overfitting can be mitigated by this augmentation process and is further lessened by the drop-out layers.

**B. Pre-trained model**

A 16-layer deep learning model called VGG16 has been trained to categorise photos into 1000 different groups. Large amounts of labelled data or a lot of calculations are not required for models that use the VGG16 model as their initialization. As a result, transfer learning has been used to speed up the training of deep models for different challenges using the VGG16 pre-trained model.

**C. Pre-trained fine-tuned model**

For the dataset under consideration, a fine-tuned model is initialised using the parameters of the VGG16 model. Except for the final layer, all of the convolutional layers were frozen. Additionally, the network weights are updated using the breast tumour dataset and the Stochastic Gradient Descent (SGD) algorithm. The classifier training process was repeated 50 times.

**D. Experiments**

The tests were conducted using an NVIDIA GeForce GTX, and the baseline model took 5 minutes (for model training and testing), the VGG16 model took 100 seconds, and the fine-tuned technique took 36.6 minutes (for training and testing).

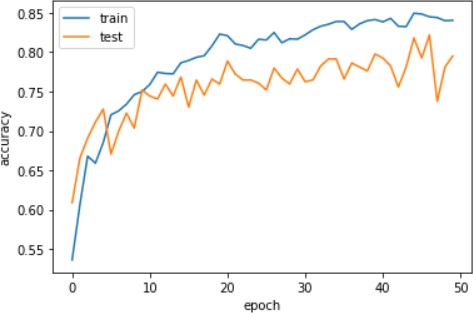
**IV. RESULTS**

Figure 3 depicts the temporal development of the baseline technique, in which the deep architecture is trained from scratch, in identifying benign and malignant breast tumours.

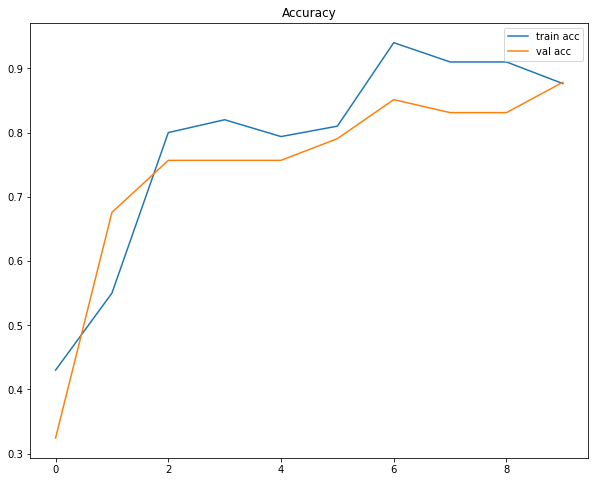
Figure 4 displays the outcomes of the second method, which is based on transfer learning and the pre-trained VGG16 model. Figure 5 displays the roc loss curve of the training data. The relative effectiveness of the three approaches is displayed in Figures 3-5.

In the end, the fine-tuned learning method using US pictures performs better than the baseline, transfer-learning, and MG and MR fine-tuned learning methods.

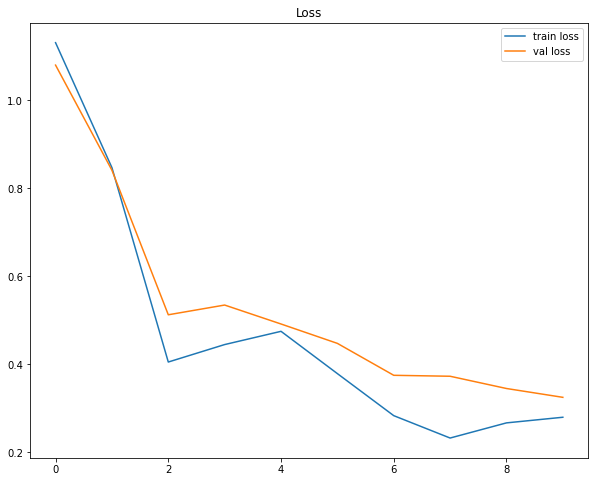
An MR-based breast cancer detection model trained from scratch has an accuracy of 0.93 percent prior to fine-tuning superior to the accuracy of the US-based detector trained from scratch, which display values of 0.79 and 0.81 respectively. The accuracy of the US-based model built using the pre-trained VGG16 model are 0.93 percent. This performance is comparable to the MR-based detector's performance after initial training.



**Figure 3: - Detection Accuracy**



**Figure 4: The system accuracy using transfer learning and the pre-trained VGG16 model.**



**Figure 5: Plotted VGG16 roc loss curve of the training data**

**V. DISCUSSION AND CONCLUSIONS**

We looked at three different deep learning methods for detecting malignant breast cancers in ultrasound pictures. We have experimentally demonstrated that the pre-trained VGG16 model's fine-tuned variation provides the best performance metrics.

The results' accuracy ranges from 85% to 94%, to put it briefly. We employed data augmentation to enhance the amount of training photos, and we used width shift, height shift, and rotation to combat overfitting. Additionally, we did fine-tuning to the bottleneck features of the VGG16 pre-trained model in order to improve accuracy.

Despite the fact that our dataset is rather small, the results claim that the fine-tuned model with medical data for pre-training has increased the accuracy of classification. This research is a step toward the development of useful and flexible deep learning algorithms for the categorization of breast cancer. The clinical assessment and treatment procedures should be complemented and supported by these systems. Modern ultrasound-based techniques have been created to achieve high-performance breast cancer screening, however there are still many obstacles to overcome. Particularly, the lack of racial and cultural diversity in the training data may have a deleterious effect on patient populations that are underrepresented in both detection and survival rates. Our research makes an effort to close this gap by making use of ultrasound breast pictures gathered from patients in North Africa and the Middle East. We intend to develop a deep learning architecture for upcoming work using pre-training data gathered from several imaging modalities. The development of new automated detection systems based on medical imaging can benefit from this pre-trained model.

**VI.** **REFERENCES**

1. Cancer screening in the United States, 2013, CA: A Cancer Journal for Clinicians, vol. 63, no. 2, pp. 87-105, 2013. R. A. Smith, D. Brooks, V. Cokkinides, D. Saslow, and O. W. Brawley
2. Cancer Incidence in Egypt: Results of the National Population-Based Cancer Registry Program, Journal of Cancer Epidemiology, vol. 2014, pages. 1-18, 2014. A. S. Ibrahim, H. M. Khaled, N. N. Mikhail, H. Baraka, and H. Kamel.
3. Journal of Clinical Oncology, vol. 20, no. 3, pp. 719–726. C. Vogel, "Efficacy and Safety of Trastuzumab as a Single Agent in First-Line Treatment of HER2-Overexpressing Metastatic Breast Cancer."
4. New Targeted Agents and Immunotherapy in Breast Cancer, American Society of Clinical Oncology Educational Book, vol. 37, pp. 65-75, 2017. I. Mayer, R. Dent, T. Tan, P. Savas, and S. Loi.
5. Pembrolizumab (pembro) monotherapy for previously treated metastatic triple-negative breast cancer (mTNBC): KEYNOTE-086 cohort A., Journal of Clinical Oncology, vol. 35, no. 15 suppl, p. 1008-1008, 2017. S. Adams, P. Schmid, H. S. Rugo, E. P. Winer, D. Loirat, A. Awada, D. W. Cescon
6. C. V. Poznak, M. Gnant, and L. American Society of Clinical Oncology Educational Book, vol. Schnipper, "Therapeutic Bone- Modifying Agents in the Nonmetastatic Breast Cancer Setting: The Controversy and a Value Assessment," p. 37, pp. 116–122, 2017.
7. Image-Guided Procedures for Breast Masses, H. W. Hazard and N. M. Hansen, Advances in Surgery, vol. 41, pp. 257-272, 2007.
8. Dropout: a straightforward method to avoid neural networks from overfitting, in N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, J. Machine Learning Res, vol. 15, pp. 1929–1958, 2014.
9. Fast R-CNN, 2015 IEEE International Conference on Computer Vision (ICCV), paper by R. Girshick.
10. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, 2017, pp. 84–90.
11. A.Zisserman and K. Simoyan. Large-scale image recognition using very deep convolutional networks. 2014, arXiv preprint 1409.1556.
12. Deep Learning for Health Informatics, IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 1, pp. 4-21, 2017. D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.-Z. Yang.
13. Deep learning for visual understanding: A review, Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. Lew, Neurocomputing, vol. 187, pp. 27–48, 2016.
14. Breast tumour detection and identification using deep convolutional neural networks, Research Group on Mathematical Modeling School of Mathematical Sciences University EAFIT Medelln, Colombia, J.D. Gallego-Posado, D.A. Montoya-Zapata, and Q.L. Quintero-Monteroya.
15. The 9th International Symposium on Chinese Spoken Language Processing, Singapore, 2014, pp. 446-449. Z. You and B. Xu, "Improving training time of deep neural network with asynchronous averaged stochastic gradient descent."

Cite this URL: 10.1109/ISCSLP.2014.6936596.

1. Combining Nakagami imaging with convolutional neural network for breast lesion classification, 2017 IEEE International Ultrasonics Symposium (IUS), Washington, DC, pp. 1-4. M. Byra, H. Piotrzkowska-Wróblewska, K. Dobruch-Sobczak, and A. Nowicki.
2. Classification of breast lesions using cross-modal deep learning, IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, 2017, pp. 109–112. O. Hadad, R. Bakalo, R. Ben-Ari, S. Hashoul, and G. Amit.
3. Automatic Breast Ultrasound Lesions Detection Using Convolutional Neural Networks by M. H. Yap et al., IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 4, July 2018, pp. 1218–1226.
4. Using deep learning for mammography classification: P. U. Hepsa, S. A. zel, and A. Yazc, 2017 International Conference on Computer Science and Engineering (UBMK), Antalya, pp. 418–423.
5. Breast cancer categorization using deep belief networks, Expert Systems with Applications, vol. 46, pp. 139–144, 2016. A. Abdel-Zaher and A. Eldeib.
6. Breast cancer identification in digital mammograms, 2015 IEEE International Conference on Imaging Systems and Techniques (IST), Macau, pp. 1-6. K. L. Kashyap, M. K. Bajpai, and P. Khanna.

Cite this URL: 10.1109/IST.2015.7294523

1. Automatic Breast Ultrasound Lesions Detection Using Convolutional Neural Networks by M. H. Yap et al., IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 4, July 2018, pp. 1218–1226.
2. Joint Weakly and Semi-Supervised Deep Learning for Localization and Classification of Masses in Breast Ultrasound Images, S. Y. Shin, S. Lee, I. D. Yun, S. M. Kim, and K. M. Lee, IEEE Transactions on Medical Imaging, vol. 38, no. 3, March 2019, pp. 762–774.
3. Race, ethnicity, and the diagnosis of breast cancer: B. Daly and O. I. Olopade, JAMA, vol. 313, no. 2, pp. 141–142, 2015.