Indonesian Journal of Electrical Engineering and Computer Science
Vol. 28, No. 1, October 2022, pp. 183~191
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v28.i1.pp183-191

Fully automated model on breast cancer classification using deep learning classifiers

Mudhafar Jalil Jassim Ghrabat[1], Zaid Alaa Hussien[2], Mustafa S. Khalefa[3], Zaid Ameen Abduljabba[3,4], Vincent Omollo Nyangaresi[5], Mustafa A. Al Sibahee[6,7], Enas Wahab Abood[8]
[1]Department of Pharmacy, Ashur University College, Baghdad, Iraq
[2]Departments of Information Technical, Management Technical College, Southern Technical University, Basrah, Iraq
[3]Department of Computer Science, College of Education for Pure Sciences, University of Basrah, Basrah, Iraq
[4]Huazhong University of Science and Technology, Shenzhen Institute, Shenzhen, China
[5]Faculty of Biological and Physical Sciences, Tom Mboya University, Homabay, Kenya
[6]Department of Computer Technology Engineering, Iraq University College, Basrah, Iraq
[7]College of Big Data and Internet, Shenzhen Technology University, Shenzhen, China
[8]Department of Mathematics, College of Science, University of Basrah, Basrah, Iraq

Article Info

Article history:

Received Mar 11, 2022
Revised Jul 19, 2022
Accepted Aug 6, 2022

Keywords:

Breast cancer
Classification
Deep learning model
Fully automated model

ABSTRACT

Deep learning models on the same database have varied accuracy ratings; as such, additional parameters, such as pre-processing, data augmentation and transfer learning, can influence the models' capacity to obtain higher accuracy. In this paper, a fully automated model is designed using deep learning algorithm to capture images from patients and pre-process, segment and classify the intensity of cancer spread. In the first pre-processing step, pectoral muscles are removed from the input images, which are then downsized. The removal of pectoral muscles after identification may become crucial in classification systems. Finally, the pectoral musclesaredeleted from the picture by using an area expanding segmentation. All mammograms are downsized to reduce processing time. Each stage of the fully automated model uses an optimisation approach to obtain high-accuracy results at respective stages. Simulation is conducted to test the efficacy of the model against state-of-art models, and the proposed fully automated model is thoroughly investigated. For a more accurate comparison, we include the model in our analysis. In a nutshell, this work offers a wealth of information as well as review and discussion of the experimental conditions used by studies on classifying breast cancer images.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Mustafa A. Al Sibahee
College of Big Data and Internet, Shenzhen Technology University
Shenzhen, 518118, China
Email: mustafa@sztu.edu.cn

1. INTRODUCTION

Breast cancer is a major cause of death globally and isgenerally caused by the abnormal behaviour of T-cells grown in breasts. These cells may also proliferate in regions, where they are not typically seen in the human body, andthis phenomenon is clinically referred to as metastasis. Mammography is the best option for detectingbreast cancer before it spreads. Radiologists' experience determines the results of mammographic images, resulting in many false positives [1], [2].

Astley *et al.* [3] reported that subjective breast density evaluation is more accurate than automatic and semi-automated approaches in predicting the risk of breast cancer. Whether mammographic density causes more aggressive breast cancers has been a question. The effect of mammographic density onprognosis should be investigated. Breast density is one of the metrics used to measure the density or number of fibro-glandular tissues visible on mammograms [4].

The thick tissue found in the breast is of a non-fatty type and has limited effect on increasing the risk ofbreast cancer; however, it can cause difficultyindetecting abnormalities and increasingcancer risk. Breasts with high-density tissues arehighly likely to acquire cancer than the ones with reduced tissue density [5]. Breast density estimation and categorisation can be conducted using various computer approaches [6]-[11]. After removing the pectoral muscles from mammograms, researchers have presented methods for segmenting the dense breast region and dividing it by the total breast area [12], [13].

Images of breast density are segmented using various approaches, such as thresholding [14], region growth [15], clustering [16] and texture statistical variation [17], for classification and estimation. However, the poor noise ratio and the variety of densities in texture and appearance cause difficulty in segmentation and classification ofbreast density. Convolutional neural networks (CNN) have made significant progress particularly in the classification and identification of patterns in an image. In addition, DL has numerous advantages over other machine learning techniques. In literature [18]-[20], various approaches forestimatingbreast density have been reported. The steps for classification of breast cancer are as follows.

Initially, pectoral muscles in input photographs are eliminated as part of the initial preprocessing step, and the images are shrunk. The removal of pectoral muscles after identification may become critical in classification systems. Finally, the pectoral musclesareremoved from the image by using area expanding segmentation. All mammograms areshrunk to reduce the processing time.

In this paper, a fully automated model is designed using deep learning algorithmtocapture images from patients andpre-process, segment and classifythe intensity of cancer spread. Each stage of the fully automated model uses an optimisation approach to obtain high-accuracy results at respective stages. The simulation is conducted to test the efficacy of the model against state-of-art models.The proposed fully automated model has higher accuracy than other methods.

The remainder of this paper is organised as follows. Section 2 reviews works related. Section 3 we evaluate the results and discussion. Section 4 comparison between proposed and different deep learning pre-trained models. Section 5 presents the conclusion.

2. RELATED WORKS

The ideas of [1] were kept alive by Cumulus software, which has increased the resource and technology for finding the reasons or causes of breast cancer. The Cumulus programme [2] is a smart way to understand the risks of breast cancer; the program conductestimations based on the threshold level used tosegment the tissue (of thick ones). In this research, the area of thick breast region isclassified into six percentage categories. This strategy is less accurate and has significant drawbacks because it often lacks the precision needed for accurate segmentation. However, relying on thresholding may be less accurate.

Breast density can be classified using imaging systemsand reported by conformingtopreviously employed standards [3]. Automated procedures, including radiodensity assessment [4], area available freely. LIBRA considers 86 variables, including global characteristics, such as patient age and X-ray breast thickness, as well as parameters, such as disconnected areas and Z-score mean. This software analyses areas in breast regions by using mammography images [4]; it further helps to estimate percentage density and dense tissue area. This old, handcrafted approach hasan accuracy of 0.81 in estimatingbreast density but istime-consuming and difficult to use.

Researchers have devised a new algorithm that can accurately estimate the percentage of patients with Parkinson’s disease (PD) based on their BI-RADS density ratings. The CC-MLO-averaged accuracy of the algorithm is 0.98, which is higher than the accuracy of LIBRA. Volumetric breast density can be quantified using volume-based approaches [5]. For each pixel in a mammography image, Quantra analyses fibroglandular tissue thickness and X-ray attenuation to calculate the tissue volume in regions around the breast. The number of tissuesin each pixel, including dense and non-dense areas,is also considered.

Fully automated approacheshave been used for quantitative assessment of breast density. Volumetric estimate is used to calculate the density of breast tissues [6]-[8].The use of deep learning methods to investigate breast density has opened a new avenue in machine learning. Deep learning approaches have been applied to obtain greater results in extracting features from mammograms. Computer networks, such as CNN, help in classification via features based on pre-processing raw images and properly depicting things at varied scales and orientations due to their deep learning capabilities. CNN is a popular method among other deep learning models.

Unsupervised deep learning methods, such as those developed by Kallenberg *et al.* [9], can be used to acquire information from features, including fatty and dense tissues. CNN has been utilised to perform unsupervised feature learning on breast density areas in mammograms by using unlabelled imaging data. Sub-images based on dense or fatty regions are created from the input mammogram.

A previous study [10] that used a state-of-art deep learning model for tissue segmentation in breast reported an average accuracy of 0.9%. The FCNN developed in [11] was used to automatically segment thick fibro-glandular regions on mammograms [12], [13]. A total of 455 mammography images containing 58 instances wereused in the evaluation. ImageNet-trained VGG16 was fine-tuned for breast density estimation and segmentation.

For breast density classification, deep learning utilizes the structure of CNN and the usual BI-RADS categorisation. Leila *et al.* [14] presented two classifiers for categorizing breast density, and one of which is based on the CNN-AlexNet model. When low-quality images were excluded, the classification accuracy was increased to 98%. The study of Leila *et al.* [15] divided the dense and fatty regions of the breast into two distinct areas. Three convolution layers were employed in a deep CNN that included six phases. The first three phases were used to generate features, while the second three stages were used to forecast the chance of occurrence.

For CNN training, [16] used an approach to classify features based on a patch-wise supervised methodology on mammography images. The raw DNN output was 0.80, and the post processed output of the deep neural network was 0.81. Many breast density classification methods have been developed, but very few literatures reached an accuracy greater than 90%; is considered a sophisticated model than the conventional methods. Breast cancer is the greatest cause of mortality among malignancies predicted in women. Despite several efforts to resolve this issue, a definitive answer has yet to be developed. The research gaps or short comings of prior studies are outlined as follows:

- Accurate detection of breast cancer by using automated approaches remains a key scientific problem.
- This problem is exacerbated by the fact that practically all accessible datasets are imbalanced, meaning that the number of instances in one class substantially out numbers those in all other classes.

By following the processes or approaches mentioned below, these research gaps can be addressed, or the recognised restrictions can be solved:

- The SMOTE approach, which identifies pictures of breast tumours to improve performance, solves the dataset imbalance problem.
- Histopathologists are well-versed in labelling lesion reports and histopathology photos. Deep learning is utilised to address the weak generalisation capacity and over-parameterised networks, which result in overfitting.

3. PROPOSED METHOD

This section explains the methodology employed in the study. As depicted in Figure 1, the entire research approach involves the classification of tissues in a mammographic breastimage that includes breast tissue, pectoral musclesand background regions. Mammograms are used in the first stage to remove pectoral muscles. Following this step, the input is rescaled to include 512×512 pixels of mammograms, which include varyingbreast density levels. Finally, a binary mask comprising dense tissue is created by processing the mammograms. Two methods are usedto classify breast densities: first, the output of binary mask is fed as input to the multi-class classifier, namely, CNN to classify the density of breast tissues into four segments, which serve a greater purpose on tissue percentagedensity. Step-by-step instructions for preparing the dataset for the proposed model are outlined in the following section.

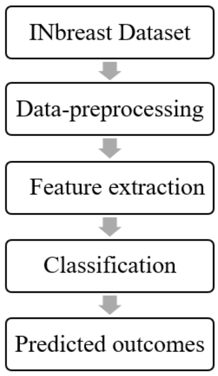
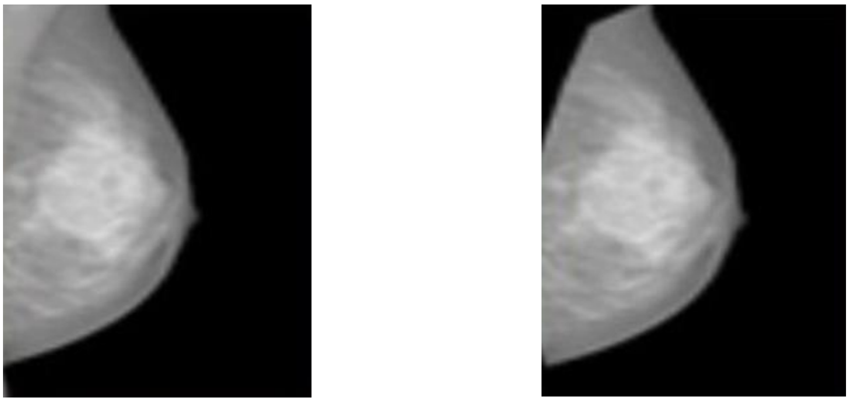


Figure 1. Proposed architecture

3.1. Preprocessing

The pectoral muscles are cleared from the input images,which are resized as an initial part of preprocessing. Processing techniques may falsely detect dense tissue areas on mammographic images due to overlap and high intensity appearing between the glandular tissue and pectoral muscles. Figure 2 the breast area and pectoral muscle are separated from the backdrop first, and then the mammography orientation is established. Finally, the pectoral muscle is removed from the picture using region growing segmentation. Figure 2(a) shows an example of eliminating the pectoral muscles. In order to save computing time, all mammograms were resized from (2,560×3,328) or (3,328×4,084) pixels to (512×512) pixels (i.e., the resolution yields the best accuracy for the segmentation stage) Figure 2(b).



(a) (b)

Figure 2. The breast area and pectoral muscle are separated from the backdrop first, and then the mammography orientation is established, (a) original images and (b) pectoral muscle removal

3.2. Breast density classification

The CNN model is used to classify breast density; in this model, each multi-class modelling is allowed to find four classes outlined below:

3.2.1. Percentage estimation on breast density

The dense area of the tissue in a mammography image is considered the percentage of total surface area in the breast region, also known as percentage density. We discuss five stages in the traditional method:
- First, the mask images are resized to the same resolution as the input mammograms.
- The study uses non-zero pixels in an input mammography image to measure the breast area.
- To indicate the amount of thick tissue in an area, we count the number of pixels that are not zero.
- The ratio of the dense tissue area to the breast area is computed to estimate the density of the tissue area.
- The density of breast is finally classified using a multi-class classifier into four categories based on the thresholding procedures.

3.2.2. Breast density classification using DL

Breast density is difficult to categorize using most approaches because of their computational complexity. Three convolution layers havekernel sizes of 4×4, 5×5 and 9×9 with fully connected layers. The study uses the maximum pooling layer with a stride of 4×4 that includes two convolution layers. A flattening step is conducted before the output of the last convolution layer is sent into the first fully convolutional layer with 128 neurons. ReLU is used as an activation function for the four layers.

In the first FC layer, a 0.5 dropout is used to reduce the overfitting process. When all four neurons in the final FC layer are used, the soft-max function is used to determine the membership degree for a class in a binary mask input. An unbalanced dataset can be avoided using a cross-entropy loss. It is one less than the ratio of samples per class to the overall sample count.

Momentum of 0.9, learning rate of 0.001 and batch size of 16 are used to optimise the model with RMProp. Five layers of the network are randomly seeded with weights before the network is trained. During training, total layers, optimum architecture, filters and neurons are discovered. Architecture of the CNN model used in this study is shown in Figure 3. The model was slightly modified from the concept of Softmax to receive input images with a dimension.

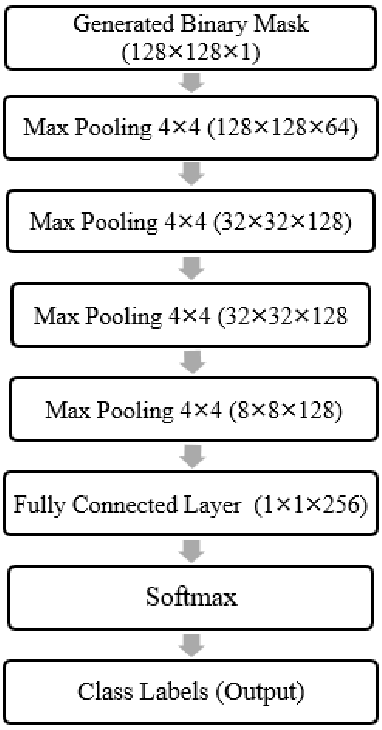


Figure 3. Proposed CNN architecture

4. RESULTS AND DISCUSSION

In this section, the programming in Python 3.5 with the PyTorch library on 64-bit Ubuntu is used to implement the proposed approach. The CPU is a 3.4GHz Intel Core-i7 CPU, the GPU isan NVIDIA 1070 with 16 GB RAM and the OS is 64-bit Ubuntu. The dataset used in the study is the INbreast dataset [21]. It is a 2D database that comprises MLO and CC 410 mammographic images that are publicly available.

INbreast breast density categorization is based on the reporting and data technique standard for breast imaging. The 3,328×4,084 pixels are the image size of a mammography. The binary masks in the ground truth is used to segment breast density missing from the INbreast dataset. As a result, the images are annotated bybreast cancer radiologists.

Among 115 patients, 80% of the images from the dataset areused for training and the remaining 20% areused as test datasets. A cross-validation is conducted on the test dataset including 82 patients (set 1–6 images, set 2–20 images, set 3–27 images, set 2–29 images) to train the CNN classifier, and the remaining images areused to validate the network. The CNN-based classification approach performs better when utilizing a balanced dataset than an imbalanced dataset [21]. CNN achieves the lowest overall accuracy of 90.29% with a size of 128×128 from the unbalanced dataset. After applying augmentation to varied input image sizes, the classification rates are 98.75% and 98.62%. However, theCNN-based classification approach for the image with size of 128×128 on the balanced dataset shows an overall accuracy. Figure 4 of 98.75%, which is improved by 0.13% compared with existing classifiers [22].

Experiments on the two datasets are conducted for the classification of breast density using trained CNN: one for a balanced dataset and one for an imbalanced dataset [23]. The 64×64 and 128×128 images are used to test the network [24], [25]. The generated binary image from the segmentation is given as an input to the CNN-based approach. The outputs (Figures 5-8) of this network are employed to classify the density of breast tissues in mammograms.

According to the results of the CNN-based classification approach, the data augmentation and the construction of balanced datasets from the overall unbalanced datasets increase the accuracy of classification. The use of these classification systems helps to maximize the accuracy and minimize the complexity. A strong correlation exists between the proposed approach and the radiologist manual for categorization, which is comparable with the correlation coefficients reported in literature.

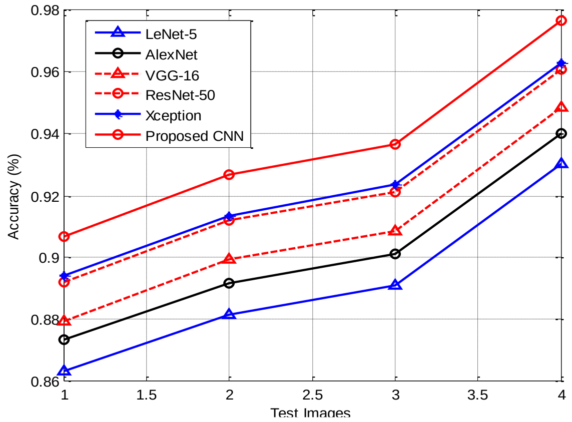


Figure 4. Accuracy

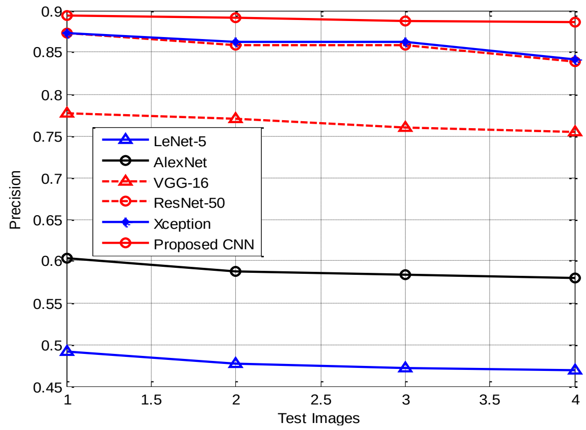


Figure 5. Precision

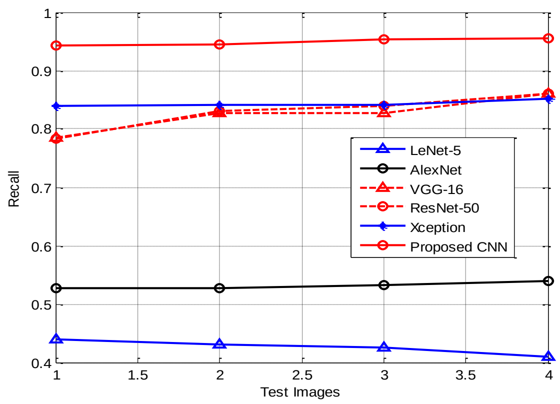


Figure 6. Recall

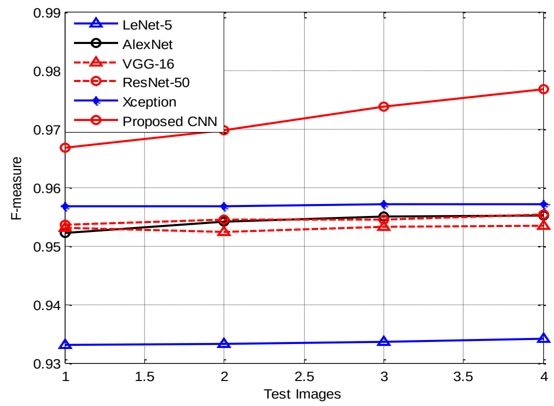


Figure 7. F-measure

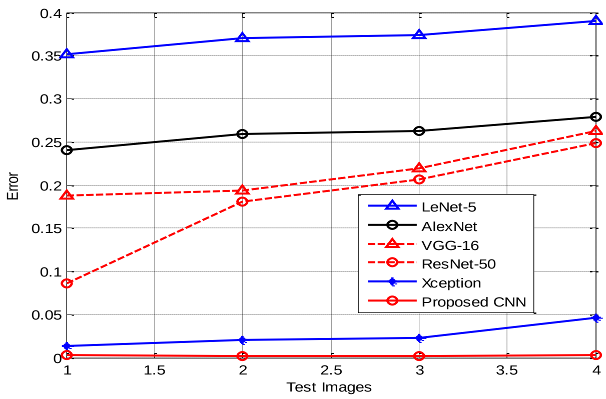


Figure 8. Percentage error

According to results from the CNN-based classification approach, the use of data augmentation and the construction of balanced datasets from the overall unbalanced datasets for increasing the accuracy of classification. The study further helps in maximizing accuracy and minimizing complexity can be attained with these classification systems. There is a strong correlation between proposed approaches in this study and radiologist manual categorization, which is comparable to the correlation coefficients reported in the literature.

5. COMPARISON BETWEEN PROPOSED AND DIFFERENT DEEP LEARNING PRE-TRAINED MODELS

We also discussed comparative findings of experiments done here on the INbreast Dataset to train accuracy and loss evaluation through this whole section, and comparative experimental results achieved using the INbreast dataset to validate loss and accuracy analysis for the VGG19, ResNet50, and CNN models. Table 1 demonstrates comparison results of performance metrics for deep learning models like-existing VGG19 and ResNet50 pre-trained model with the proposed CNN model.

Table 1. Comparison of performance

Model	Training loss	Training acc	Validation loss	Validation acc
VGG19 [23]	0.2199	0.9113	0.3675	0.8544
ResNet50 [22]	0.2722	0.8871	0.3703	0.8443
(Proposed)	0.1314	0.9489	0.3374	0.8832

6. CONCLUSION

In this paper, a fully automated model is designed that captures the images from patients, pre-process, segments and classifies the intensity of cancer spread using deep learning algorithm. The simulation is conducted to test the efficacy of the model against state-of-art models and the results of simulation show that the proposed fully automated model obtains increased accuracy than other methods. A total of 410 images from the dataset are used in evaluation. This technique was able to obtain a 98% success rate in the classification of breast density. In the future, we will examine deep learning with complex architectures, such as several layers and neurons, and test them on large groups of participants. The most exciting future initiatives, on the other hand, are learning and identifying visual signals associated to histological design and morphology in order to get quantitative semantic parameters such as the fraction of dead tissue or neoplasm in WSI.

CONFLICTS OF INTEREST

The authors declare that there are no conicts of interest regarding the publication of this paper. Availability of data and material. The data used to support the findings of this study are available from the corresponding author upon request.

REFERENCES

[1] X. Chen and E. Song, "Turning foes to friends: targeting cancer-associated fibroblasts," *Nature reviews Drug discovery*, vol. 18, no. 2, pp. 99-115, 2019, doi: 10.1038/s41573-018-0004-1.

[2] L. Kathryn, *et al*. "Screening performance of digital breast tomosynthesis vs digital mammography in community practice by patient age, screening round, and breast density," *JAMA network open*, vol. 3, no. 7, 2020, doi: 10.1001/jamanetworkopen.2020.11792.

[3] B. M. Keller *et al.*, "Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation," *Medical physics*, vol. 39, no. 8, pp. 4903-491, 2012, doi: 10.1118/1.4736530.

[4] Y. J. Hyun *et al*. "Automated volumetric breast density measurements in the era of the BI-RADS fifth edition: a comparison with visual assessment," *American Journal of Roentgenology*, vol. 206, no. 5, pp. 1056-1062, 2016, doi: 10.2214/AJR.15.15472.

[5] H. Ralph *et al.*, "Robust breast composition measurement-Volpara TM," in *International workshop on digital mammography*, 2010, pp. 342-349, doi: 10.1007/978-3-642-13666-5_46.

[6] K. M. Petersen, K. Nielsen, M. Ng, A. Y. Diao, and C. P. Igel, "Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1322-1331, doi: 10.1109/TMI.2016.2532122.

[7] D. M. Ufuk *et al.*, "Using deep learning to segment breast and fibroglandular tissue in MRI volumes," *Medical physics*, vol. 44, no. 2, pp. 533-546, 2017, doi: 10.1002/mp.12079.

[8] L. Juhun and R. M. Nishikawa, "Automated mammographic breast density estimation using a fully convolutional network," *Medical physics*, vol. 45, no. 3, pp. 1178-1190, 2018, doi: 10.1002/mp.12763.

[9] M. Mario, M. Grgic, and R. M. Rangayyan, "Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms," *Medical & biological engineering & computing*, vol. 54, no. 7, 2016, doi: 10.1002/mp.12763.

[10] T. Stylianos, M. E. Mavroforakis, H. V. Georgiou, N. Dimitropoulos, and S. Theodoridis. "A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry," *Computer methods and programs in biomedicine*, vol. 102, no. 1, pp. 47-63, 2011, doi:10.1016/j.cmpb.2010.11.016.

[11] M. Aly *et al*., "Understanding clinical mammographic breast density assessment: a deep learning perspective," *Journal of digital imaging*, vol. 31, no. 4, pp. 387-392, 2018.

[12] L. Songfeng *et al.*, "Computer-aided assessment of breast density: comparison of supervised deep learning and feature-based statistical learning," *Physics in Medicine & Biology*, vol. 63, no. 2, 2018, doi: 10.1088/1361-6560/aa9f87.

[13] Y. Keping, L. T. L. Lin, X. Cheng, Z. Yi, and T. Sato, "Deep-learning-empowered breast cancer auxiliary diagnosis for 5GB remote e-health," *IEEE Wireless Communications*, vol. 28, no. 3, pp. 54-61, 2021, doi:10.1109/MWC.001.2000374.

[14] D. Leila *et al.*,"Breast cancer risk genes-association analysis in more than 113,000 women," *The New England journal of medicine*, vol. 384, no. pp. 428-439, 2021, doi: 10.1056/NEJMoa1913948.

[15] D. Leila *et al.*,"Breast cancer association consortium. breast cancer risk genes-association analysis in more than 113,000 women," *New England Journal of Medicine*, vol. 384, no. 5, pp. 428-439, 2021.

[16] B. Kara, J. Cuzick, and K. A. Phillips, "Key steps for effective breast cancer prevention," *Nature Reviews Cancer*, vol. 20, no. 8, pp. 417-436, 2020.

[17] H. Ariella, D. R. Sudhan, C. L. Arteaga, "Overcoming endocrine resistance in breast cancer," *Cancer Cell*, vol. 37, no. 4, 2020.

[18] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, "INbreast: toward a full-field digital mammographic database," *Academic radiology*, vol. 19, no. 2, pp. 236–248, Feb. 2012, doi: 10.1016/j.acra.2011.09.014.

[19] C. Tejalal, V. Mishra, A. Goswami, and J. Sarangapani, "A transfer learning with structured filter pruning approach for improved breast cancer classification on point-of-care devices," *Computers in Biology and Medicine*, vol. 134, 2021, doi: 10.1016/j.compbiomed.2021.104432.

[20] S. Rishav, T. Ahmed, A. Kumar, A. K. Singh, A. K. Pandey, and S. K. Singh. "Imbalanced breast cancer classification using transfer learning," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 18, no. 1, pp. 83-93, 2020, doi: 10.1109/TCBB.2020.2980831.

[21] M.J. J. Ghrabat, G. Ma, P. L. P. Avila, M. J. Jassim, and S. J. Jassim. "Content-based image retrieval of color, shape and texture by using novel multi-SVM classifier," *International Journal of Machine Learning and Computing*, vol. 9, no. 4, 2019.

[22] M. J. J. Ghrabat, G. Ma, and C. Cheng, "Towards efficient for learning model image retrieval," in *2018 14th International Conference on Semantics, Knowledge and Grids (SKG)*, 2018, pp. 92-99, doi: 10.1109/SKG.2018.00020.

[23] A, M, Shohel *et al.*, "Medicine prediction based on doctor's degree: a data mining approach," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 26, no. 2, pp. 1125-1134, 2022, doi: 10.11591/ijeecs.v26.i2.pp1125-1134..

[24] M. F. A. Kadir, A. F. A. Abidin, M. A. Mohamed, and N. A. Hamid, "Spam detection using machine learning based binary classifier," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 26, no. 1, pp. 310-317, 2022, doi: 10.11591/ijeecs.v26.i1.pp310-317.

[25] A, Mustofa *et al*, "Simulating the COVID-19 epidemic event and its prevention measures using python programming,". *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 1, pp. 278-288, 2022, doi: 10.11591/ijeecs.v26.i1.pp278-288.

BIOGRAPHIES OF AUTHORS

Mudhafar Jalil Jassim Ghrabat He received a Ph.D. degree in Center for Biomedical Imaging and Bioinformatics, Huazhong University of Science and Technology, Wuhan, China. He received his master degree of Information Technology from SRM University, India, Chennai 2015, and got B.S. degree of computer science from Al Mustansiriyah University, Baghdad, Iraq, in 2005. His current research interests include Artificial Intelligence, Machine Learning, Data Mining and Image Processing. He is program committees members for 33 international conferences. His current teacher at Ashur university in Baghdad-Iraq. He can be contacted at email: mudhsafar.jalil@au.edu.iq.

Zaid Alaa Hussien received the B.Sc. degree in Computer Engineering from University of Basrah, Iraq in 2004, the M.Tech. degree in Computer Science and Engineering from Guru Gobind Singh Indraprastha University, India in 2009 and the Ph.D. degree from School of Computer Science and Technology, Huazhong University of Science and Technology, China in 2017. His current research interests include focus on security and integrity issues in cloud computing, big data, and internet of things. He can be contacted at email: zaid.alaa@stu.edu.iq.

Dr. Mustafa S. Khalefa received the Science degree in Computer Science from university of Basra in 2005, and the Aggregation in Computer Science, Software Engineering from University Technology Malaysia (UTM) from Malaysia, Rabat, in 2009. He received the Master degree in Software engineering from Faculty of Computer Science and Technology; in 2017 he completed His Ph.D. in Information System and Software Engineering from University Putra Malaysia from faculty of Computer Science and Information Technology department of Information system and Software Engineering. Mustafa Currently work as lecturer and researcher for University of Basra, Faculty of Pure Science, Computer Science department, in addition he is consultant for many companies and training center with many skills in top management, currently research interest are in: software engineering, information system, information technology, ontology, knowledge management, enterprise resources planning, internet of things, business analysis and business development, globule virtual team, leadership. Where he is the author for Many research publications. He can be contacted at email: mustafakhalefa@gmail.com.

Zaid Ameen Abduljabbar received the bachelor's and master's degrees in computer science from University of Basrah, Iraq, in 2002 and 2006, respectively, and the Ph.D. degree in computer engineering from the Department of Computer Science and Technology, Huazhong University of Science and Technology, China, in 2017. His research interests include cloud security, searchable encryption systems, similarity measures, Internet of Things, secure computation, biometric, and soft computing. He has published regular articles for more than 40 IEEE International Conferences and High-quality articles in SCI journals, and he holds 3 international patents and 2 International Computer Software Copyright. He has always served as a Reviewer for several prestigious journals, and has served as the PC Chair/PC member for more than 25 international conferences. He can be contacted at email: zaid.ameen@uobasrah.edu.iq.

Vincent Omollo Nyangaresi received his bachelors degree in telecommunication and information engineering in 2010 and a masters degree in information technology security and audit in 2018. His first doctorate degree is in information technology security and audit. He is currently pursuing his second doctorate degree in computer science. His research interests include machine learning, computer networks and security protocols, telecommunication engineering and systems modeling. He has published over 40 research articles in peer reviewed journals, conferences and symposiums. He can be contacted at email: vnyangaresi@tmuc.ac.ke.

Mustafa A. Al Sibahee current position at college of big data and internet, Shenzhen Technology University, Shenzhen-China. Department of Computer Technology Engineering, Iraq University College, Basrah, Iraq. Received his, Ph.D. 2018, from Huazhong University of Science and Technology, Wuhan-China. From April-2019 to March-2021, was a postdoctor at Shenzhen Huazhong University of Science and Technology, Research Institute, Shenzhen-China. His research interests include Computer Networks and Information Security, Internet of Things and Wireless Sensor Networks (wsns). He can be contacted at email: mustafa@sztu.edu.cn.

Enas Wahab Abood received the bachelor's and master's degrees in computer science from Basrah University, Iraq, in 2005 and 2011, respectively. Her research interests include image processing, sound processing, encryption systems, similarity measures, NLP, Graph theory. She has many articles in different aspects of computer science. She can be contacted at email: enas.abood@uobasrah.edu.iq.