International Journal of Informatics and Applied Mathematics
e-ISSN:2667-6990 Vol. 6, No. 2, 20-34

Breast Cancer Image Classification Using Convolutional Neural Networks (CNN) Models

Abdelnour Boukaache[1], Nasser Edinne Benhassine[1,2], and Djalil Boudjehem[2]

[1] Advanced Control Laboratory, University of 8 Mai 1945, Guelma, Algeria
[2] University Center, Echerif Bouchoucha, Aflou, Algeria
boukaache.anour@yahoo.fr,benhassine2691g1@gmail.com,
boudehem.djalil@gmail.com

Abstract. Breast cancer can progress silently in its early stages and frequently without noticeable symptoms. However, it poses a serious risk to women. It is imperative to recognize this potential health concern to mitigate it early. In the last few years, Convolutional Neural Networks (CNNs) have advanced significantly in their ability to classify images of breast cancer. Their capacity to automatically extract discriminant features from images has enhanced the performances and accuracy of image classification tasks. They outperform state-of-the-art techniques in this area. Furthermore, complicated models that were first learned for certain tasks can be easily adapted to complete new tasks by using transfer-learning approaches. However, deep learning-based categorization techniques could experience overfitting issues, particularly in cases where the dataset is small. The primary goal of this work is to investigate the performances of certain deep learning models to classify breast cancer images and to study the effects of data augmentation techniques, such as image rotation or displacement when utilizing a transfer learning approach. Using certain image datasets, the ResNet18, Resnet50, and VGG16 models demonstrated accuracy improvements, according to our experimental results.

Keywords: CNN · ResNet18 · Resnet50 · VGG16 · Data Augmentation · Transfer Learning

Received: 20 November 2023 | Revised: 20 December 2023 | Accepted: 17 January 2024

1 Introduction

Breast cancer is indeed one of the most common causes of death among women worldwide. Early detection plays a crucial role in reducing the number of deaths associated with this disease. Medical imaging plays a vital role in the detection and classification of breast cancer. Breast ultrasound is widely used in breast cancer detection, where this technique plays a good role in providing images of breast tissue to help classify between malignant and benign lesions. It is a valuable imaging tool for evaluating breast abnormalities [1,2,3,4]. On the other hand, Convolutional Neural Networks (CNNs) are a powerful tool of deep learning models commonly used in image classification tasks, including the classification of breast cancer images [5]. Furthermore, deep learning has achieved significant importance in the field of medical image analysis, and it has been widely applied to various tasks, including the analysis and interpretation of medical images such as breast cancer classification. [6]. However, most medical datasets suffer from limitations such as small sample sizes. In such scenarios, transfer learning is indeed a viable and effective strategy for addressing these challenges and improving the performance of machine learning models (e.g., ResNet18, Resnet50, and VGG16) [7], where the obtained knowledge is transferred from the source domain to the target domain. In recent years, various methods have discussed breast cancer detection and classification using CNN-Transformer. For example, Hasnia et al. [8], proposed a transfer learning techniques with DenseNet121, ResNet50, VGG16, and mobileNetv2, for classifying breast tumors in mammograms into benign and malignant with an accuracy of 90%. Prajapati et al. [9] applied transfer learning to the pre-trained model VGG16, (RVG) in X-ray images, with an accuracy of 88.4%. In [10], the authors employed transfer learning to the pre-trained model ResNet50, HAM10000. The model achieved 93% average accuracy and precision in the range [0.7, 0.94], which outperformed dermatologists accuracy of 84%. Sasikala et al. [11] suggested a CNN algorithm with four different transfer learning techniques: AlexNet, VGG16, ResNet50, and ResNet34, in clinical images collected from Kaggle, and an accuracy of 90.12% was achieved. Literature [12] used an application of transfer learning to the pre-trained model VGG16 for classifying diabetic retinopathy stages, they reached 86.5% of accuracy for 2-class, 80.5% for 3-class, 63.5% for 4-class, and 73.7% for 5-class. Janoria et al. [13] applied transfer learning to the pre-trained model VGG16 with the K-NN algorithm, and they achieved 99% accuracy. Transformer learning can be used to improve the classification of breast cancer ultrasound images by leveraging pre-trained models and incorporating context information. Transfer learning, a machine learning approach that allows the reuse of pre-trained models for new tasks, has been applied to breast cancer diagnosis using ultrasound images [14]. By using pre-trained models such as VGG16, VGG19, MobileNetV2, and ResNet50V2, the accuracy of breast cancer classification can be improved [15]. Furthermore, a breast cancer classification method using transformers has been proposed, which achieves high accuracy even with small datasets [16]. These approaches demonstrate the potential of transformer learning in improving the classification of breast cancer ultrasound images. Many researchers use data augmentation techniques to increase the number of samples in the dataset while preventing overfitting, which is common when training on a model that has a small set of samples [17,18]. This study focused on transfer learning approaches using pre-trained models, and we showed the performances of different models for the classication of breast tumors in ultrasound images. The primary objective of this paper is to study the potential impact of data augmentation on the performance of transfer learning models in medical imaging. We have also verified the effectiveness of using data augmentation, which further improves the performance of our model for two and three-class classification problems. With the proposed network architecture, we achieved an overall accuracy of 97.8% and 90% for 2 and 3 classes of tumor classication respectively. After applying image augmentation in our proposed network architecture, the accuracy was increased by about 3.7% with ResNet18, and 3.5% with VGG16. The rest of the paper is organized as follows. In section II, we briey present the following evaluation process. Experimental results and discussions are presented in section III. Finally, section IV concludes the paper.

2 Methodology Description

This section describes the data preparation processes and categorization approach employed. Fig. 1 depicts the suggested method's flow chart. Following pre-processing, pictures were modified using three pre-trained convolutional neural networks (VGG16, ResNet18, and ResNet50) to create feature vectors, which were then classified using the Support Vector Machine (SVM) classifier. The following subsections provide a quick overview of three commonly used CNN models.

2.1 Transfer learning with CNN

Transfer learning is a technique used in machine learning to reuse a pre-trained model for a new task or domain rather than starting from scratch. Pre-trained models are neural networks that have already been built on a large dataset and fine-tuned for a different dataset or problem with less data and time. Transfer learning is the art of using pre-trained models to solve deep learning problems

Pre-Trained VGG16 Deep CNN Model VGG16 (Visual Geometry Group 16) is a deep convolutional neural network (CNN) architecture that was developed at the University of Oxford and was one of the top-performing models in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2014 [19]. It consists of 16 weight layers, including 13 convolutional layers and 3 fully connected layers. The convolutional layers are followed by max-pooling layers. Moreover, it is a large network and it has about 138 million (approx.) parameters (see Fig. 2). It achieved remarkable performance on various computer vision tasks, especially image classification. VGG16 has also been used as a base model

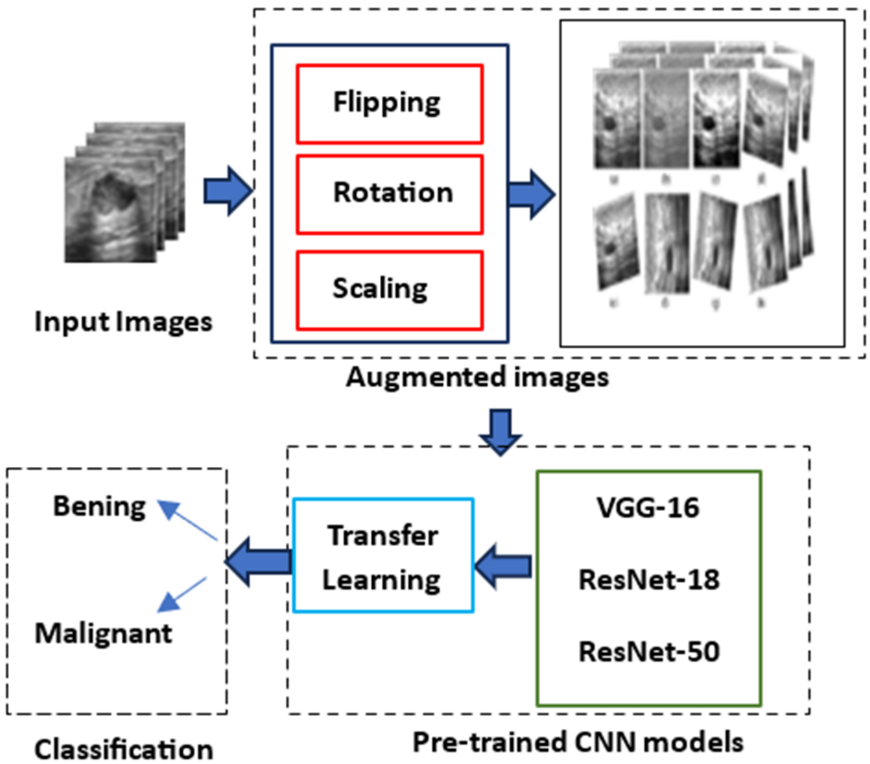

Fig. 1. Flowchart of the proposed method.

for transfer learning, where pre-trained weights from VGG16 are fine-tuned on specific tasks, making it a valuable tool in the field of deep learning for computer vision [20].

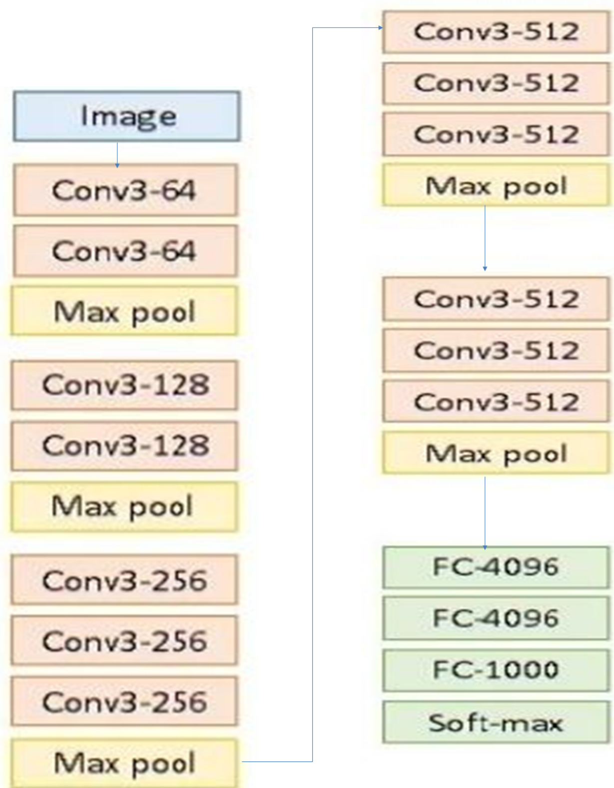

Fig. 2. VGG16 architecture. [21]

Pre-trained Residual Networks (ResNets) Residual Neural Networks (ResNets) are a form of neural network architecture developed by Kaiming He et al. in 2015. They are intended to solve the issue of vanishing gradients in deep neural networks, which occurs when the gradient signal gets too tiny to be effective for updating the network's weights during training. Furthermore, a residual block is defined as the skip connections that conduct identity mappings and are added to the layer outputs [22]. In Fig. 3, if x is the input and F(x) is the output from the layer, then the output of the residual block can be given as:

$$Y = F(x) + x \tag{1}$$

By adding non-linearity, the ReLU function makes the neural network more sophisticated and enables it to learn more intricate data representations. The residual network has multiple variations, such as ResNet18, ResNet50, and so forth.

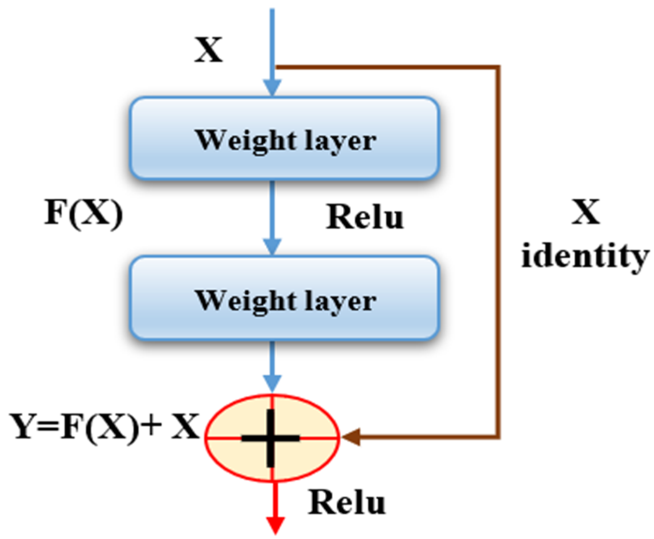

Fig. 3. Residual Basic-Block structure of ResNets.

Pre-trained ResNet18 ResNet18 is a 72-layer architecture with 18 deep layers that is trained on more than a million images from the ImageNet database (see Fig. 4). It is extremely useful and efficient in image classification. In the field of Deep Learning, Batch norm is a normalization technique done between the layers of a Neural Network instead of in the raw data. It is used to speed up training and provides some regularization [23]. Furthermore, the global average pooling is intended to take the role of fully connected layers in traditional CNNs. The network has an image input size of 224x224. As a result, the network has learned rich feature representations for a wide range of images [24,25].

Pre-trained ResNet50 ResNet50 is a convolutional neural network architecture that contains 48 convolutional layers, one Max-pool layer, and one average pool layer. It uses a bottleneck design for the building block [26]. Hence, 11 convolutions, which decrease the number of parameters and matrix multiplications, are used in bottleneck residual blocks. Its capability to train extremely deep networks with hundreds of layers is one of its key benefits [27].

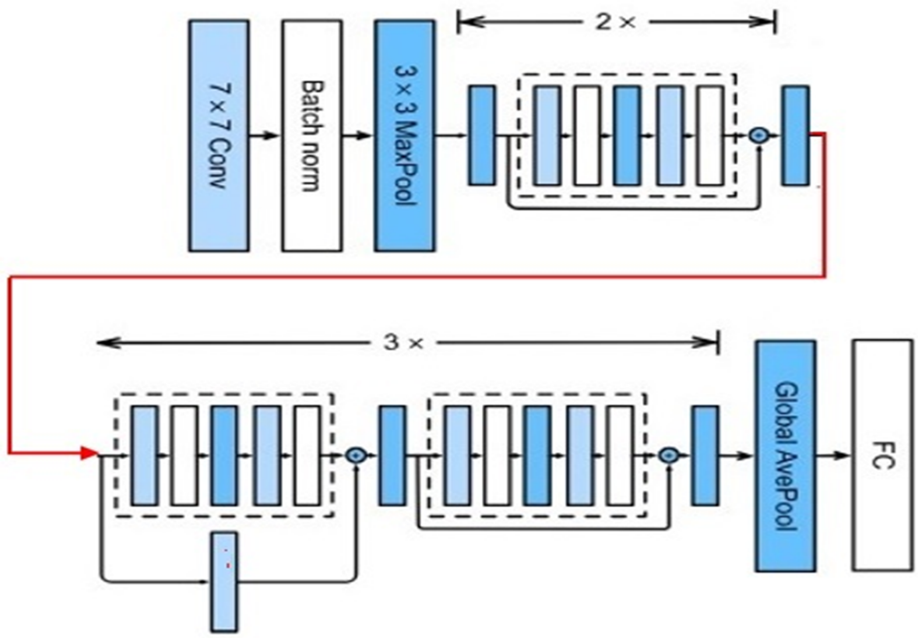

Fig. 4. Structure of ResNet18. [24]

2.2 Data augmentation

Deep neural networks require a large training set and generally perform better in the presence of more data. Finding a reliable biomedical dataset is a difficult task. Most of the datasets that are provided only contain a small quantity of data. To avoid overfitting, while maintaining a good performance, we use image data augmentation. Image augmentation is a technique to generate new batches of data from the original data. Augmentation can reduce overfitting by increasing the data set size [27,28]. In this work, to match the number of images in benign and malignant classes, we performed data augmentation on the former, using a combination of reflection and rotation.

3 Experimental Results

3.1 Database

In this study, we used two Breast Ultrasound datasets, the first (data1), contains 250 breast cancer images, 100 benign and 150 malignant. It is a database already widely used in the literature [28]. The second (data2) was US images; it is collected at baseline and includes breast ultrasound images among women in ages between 25 and 75 years old. It consists of 780 images which are categorized into three classes: 133 Normal, 487 Benign, and 210 Malignant images [29]. Examples of images from data1 and data2 are given in Fig. 5, and Fig. 6.

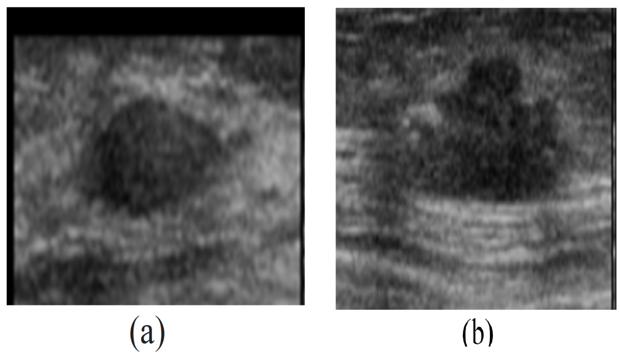

Fig. 5. Samples of Ultrasound breast images from data1: (a) benign, (b) Malignant.

3.2 Image preprocessing

The main objective of this article is to compare the accuracy of several types of CNN on tumor classification, and then the study of data augmentation in enhancing network performances. The following preprocessing steps were applied to the data before loading it to the networks.

- Resizing images to the size needed by the CNN model.
- Applying data augmentation on training images.

For the ResNet18, ResNet50, and VGG16 networks used in our experiments, we have resized all input images to 224 224.

3.3 Evaluation results

In this section, we show the evaluation results of three deep learning models (ResNet18, Resnet50, and VGG16) for the two datasets previously described. The assessment is founded on three performance metrics: the accuracy (Acc), the precision rate (Pr), and the recall rate (R) of each model using the following expressions:

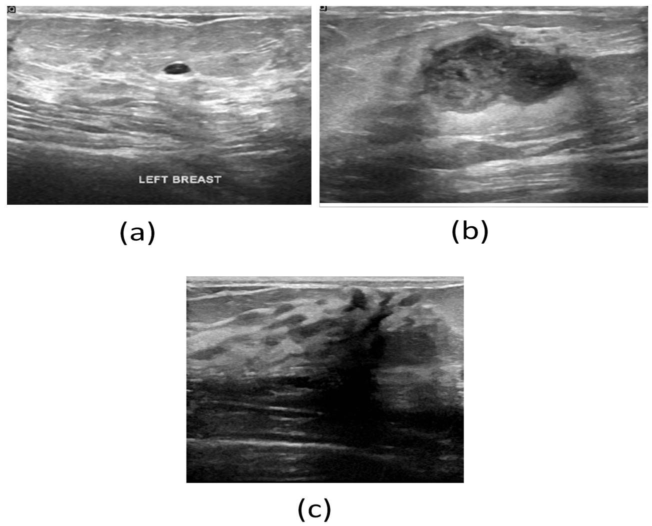

Fig. 6. Samples of Ultrasound breast images from data2: (a) benign, (b) Malignant, (c) normal.

$$R = \frac{TP}{(TP + FN)} \times 100(\%) \tag{2}$$

$$R = \frac{TP}{(TP + FP)} \times 100(\%) \tag{3}$$

$$R = \frac{TP + TN}{(TN + FN + FP + TP)} \times 100(\%) \tag{4}$$

Where: TN (True Negative) is the number of a benign classified as benign.
TP (True Positive) is the number of malignant classified as malignant.
FN (False Negative) is the number of benign classified as malignant.
FP (False Positive) is the number of malignant classified as benign.
The ratio of accurately categorized positive samples is defined as precision. The recall of the model assesses its ability to recognize positive samples. However, accuracy just evaluates how frequently the classifier predicts correctly. We present the results for the two classes (benign and malignant) and three classes (normal, benign, and malignant) classification problems.

As we can see in Table 1, the accuracy of the three models (ResNet18, Resnet50, and VGG16) for the classification of the first dataset into benign or malignant images is shown. The results are perfect and the ResNet18 deep learning model reached 99.98% accuracy. We note here that the dataset contains a low number of images, which can be used for the training process. In addition, all the images are preprocessed, which allows the model to be better trained

Table 1. Evaluation results on the first dataset (2 Classes).

Method	R (%)	Pr (%)	Acc (%)
ResNet18	100	96.8	99.98
ResNet50	100	99.7	99.96
VGG16	90.0	100	96.8

and improves the accuracy of results. In the second experiment, we will present the results of the used models on the second dataset, which contains a high number of images. Fig. 3 shows the confusion matrix of two classes, ResNet18 achieved an accuracy of 84.5%. We obtained 82% accuracy for three classes of classification.

For the same dataset, the ResNet18 network was evaluated with data augmentation and the obtained results are presented in Table 2. We see that the accuracy results were enhanced in the two cases, especially in the 2 classes by 3.7%.

Table 2. ResNet18 accuracy results.

Number of Classes	Accuracy %	
	Without data augmentation	With data augmentation
2 classes	84,5	88.2
3 classes	82	82,5

For the next step of the evaluation, we tested the VGG16 network with the second dataset. We have obtained 94.33% accuracy in the case of 2 classes and 86.5% in the case of 3 classes. The confusion matrix is shown in Fig.4. In Table 3, we summarize the results of the VGG16 network according to the used number of classes and the data augmentation technique. We remark that this technique improves the performances of the network in the two cases (2 or 3 classes). We obtained 97.8% of accuracy with 2 classes and 90% with 3 classes.

Table 3. VGG16 accuracy results

Number of Classes	Accuracy %	
	Without data augmentation	With data augmentation
2 classes	94,33	97,8
3 classes	86,5	90

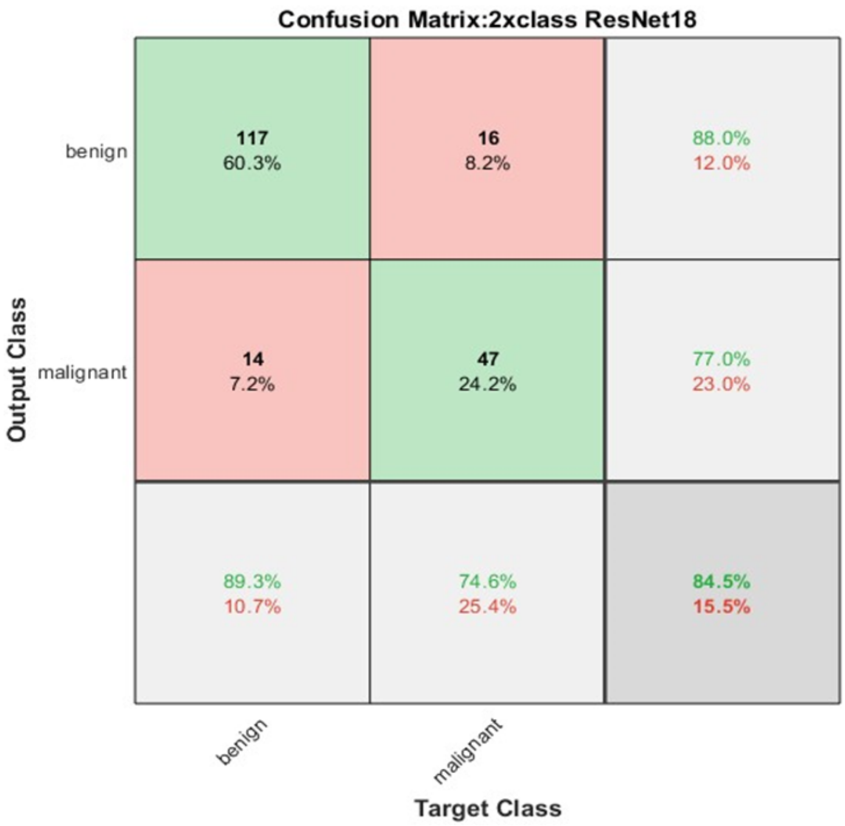

Fig. 7. Confusion matrix of ResNet18 with 2 classes.

According to the results obtained with the different deep learning models and datasets used in the investigation in our work, we can say that deep learning and transfer learning techniques can give good performances in breast cancer classification when the data are enough and well preprocessed. However, the results are not so good when the dataset is limited, and the number of images is low. For these reasons, the data augmentation technique tries to make the number of images high by rotation and translation. The image is translated by moving it either along the x- or y-axes, however, its rotation is done at random. This step will enhance the training step and then the classification performance. To show the eciency of the proposed method, we have made a comparison with other works in the literature. Table 4 shows a comparison between the results of our proposed method and previous results which are reported in the literature.

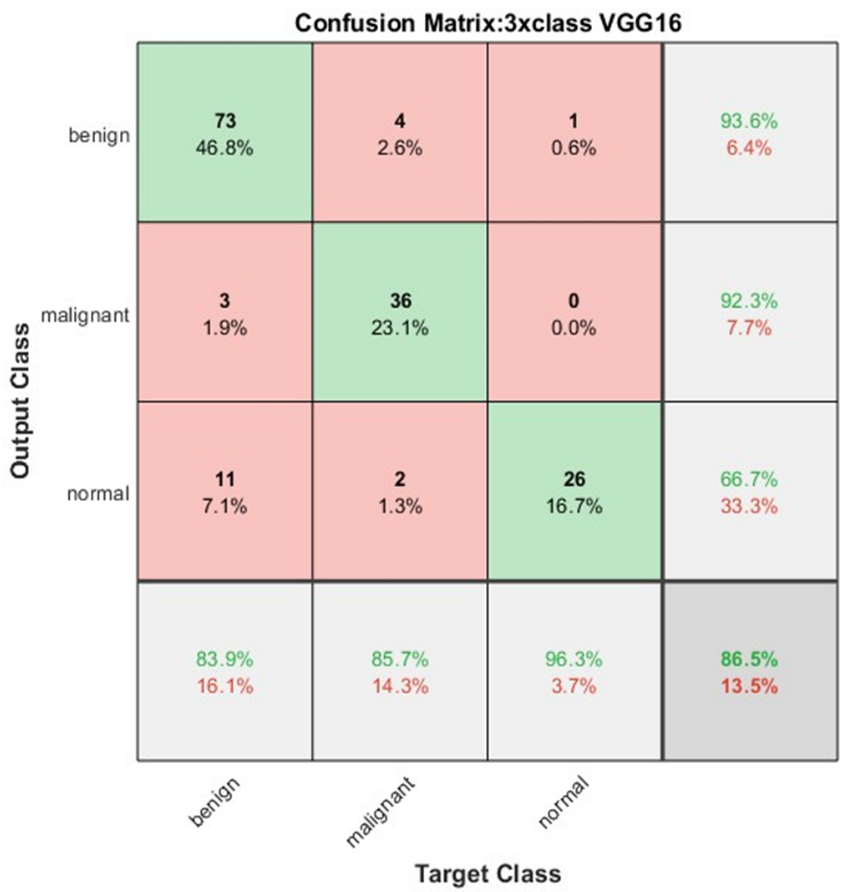

Fig. 8. Confusion matrix of VGG16 with 3 classes.

Table 4. VGG16 accuracy results

Author	Classification methods	Accuracy%
Gour et al. [30]	ResNet50	84.34
Wang et al. [31]	VGG16	80.6
Deniz et al. [32]	VGG16 and AlexNet	91.30
Gonalves et al. [33]	DenseNet	91.67
The proposed method	ResNet18	88.2
	VGG16	97.8

4 Conclusion

Early-stage breast cancer classification poses a significant role in successful treatment. In this study, we classified breast cancer using a few Convolutional Neural Network (CNN) models. In this study, two distinct datasets are used to test the ResNet18, Resnet50, and VGG16 models. To address the overfitting issue, data augmentation approaches are investigated concerning the problem of data limitation. The obtained results demonstrate the effectiveness of deep learning models in transfer learning. Furthermore, large datasets lead to better performances, elsewhere small datasets result in lower performances. However, by increasing the training dataset, the data augmentation strategy will raise the network's overall accuracy. The outcomes of using data augmentation demonstrated significant improvement. With ResNet18 and VGG16, the accuracy was improved by 3.7% and 3.5%, respectively. In future works, a thorough investigation of other specialized databases and pre-trained CNN will be carried out to generalize our results.

References

1. Mohammed Amine Naji, Sanaa El Filali, Kawtar Aarika, EL Habib Benlahmar, Rachida Ait Abdelouhahid, and Olivier Debauche. Machine learning algorithms for breast cancer prediction and diagnosis. *Procedia Computer Science*, 191:487–492, 2021.
2. Yahya Alqahtani, Umakant Mandawkar, Aditi Sharma, Mohammad Najmus Saquib Hasan, Mrunalini Harish Kulkarni, and R Sugumar. Breast cancer pathological image classification based on the multiscale cnn squeeze model. *Computational Intelligence and Neuroscience*, 2022, 2022.
3. Fei Gao, Teresa Wu, Jing Li, Bin Zheng, Lingxiang Ruan, Desheng Shang, and Bhavika Patel. Sd-cnn: A shallow-deep cnn for improved breast cancer diagnosis. *Computerized Medical Imaging and Graphics*, 70:53–62, 2018.
4. Nasser Edinne Benhassine, Abdelnour Boukaache, and Djalil Boudjehem. Classification of mammogram images using the energy probability in frequency domain and most discriminative power coefficients. *International Journal of Imaging Systems and Technology*, 30(1):45–56, 2020.

5. Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do, and Kaori Togashi. Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, 9:611–629, 2018.
6. June-Goo Lee, Sanghoon Jun, Young-Won Cho, Hyunna Lee, Guk Bae Kim, Joon Beom Seo, and Namkug Kim. Deep learning in medical imaging: general overview. *Korean journal of radiology*, 18(4):570–584, 2017.
7. Shallu Sharma and Rajesh Mehra. Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology imagesa comparative insight. *Journal of digital imaging*, 33:632–654, 2020.
8. Hasnia Merzoug, Hayat Yedjour, and Samira Chouraqui. A deep system for breast tumor classification from mammograms. *International Journal on Communications Antenna and Propagation (IRECAP)*, 12(6), 2022. doi:10.15866/irecap.v12i6.22300.
9. Shreyansh A. Prajapati, R. Nagaraj, and Suman Mitra. Classification of dental diseases using cnn and transfer learning. In *2017 5th International Symposium on Computational and Business Intelligence (ISCBI)*, pages 70–74, 2017.
10. Duyen NT Le, Hieu X Le, Lua T Ngo, and Hoan T Ngo. Transfer learning with class-weighted and focal loss function for automatic skin cancer classification. *arXiv preprint arXiv:2009.05977*, 2020.
11. S Sasikala, SN Shivappriya, et al. Towards improving skin cancer detection using transfer learning. *Biosci Biotechnol Res Commun*, 13(11):55–60, 2020.
12. Enas MF El Houby. Using transfer learning for diabetic retinopathy stage classification. *Applied Computing and Informatics*, 2021.
13. Honey Janoria, Jasmine Minj, and Pooja Patre. Classification of skin disease from skin images using transfer learning technique. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pages 888–895, 2020.
14. Weonsuk Lee, Hyeonsoo Lee, Hyunjae Lee, Eun Kyung Park, Hyeonseob Nam, and Thijs Kooi. Transformer-based deep neural network for breast cancer classification on digital breast tomosynthesis images. *Radiology: Artificial Intelligence*, 5(3):e220159, 2023.
15. Mobarak Zourhri, Soufiane Hamida, Nouhaila Akouz, Bouchaib Cherradi, Hasna Nhaila, and Mohamed El Khaili. Deep learning technique for classification of breast cancer using ultrasound images. In *2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, pages 1–8, 2023.
16. Wen Yu and Yanqiu Li. Breast cancer classification from histopathological images using transformers. In *Thirteenth International Conference on Information Optics and Photonics (CIOP 2022)*, volume 12478, pages 562–569. SPIE, 2022.
17. Abeer Saber, Mohamed Sakr, Osama M Abo-Seida, Arabi Keshk, and Huiling Chen. A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique. *IEEE Access*, 9:71194–71209, 2021.
18. Khalid M Hosny, Mohamed A Kassem, and Mohamed M Fouad. Classification of skin lesions into seven classes using transfer learning with alexnet. *Journal of digital imaging*, 33:1325–1334, 2020.
19. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
20. Alaa MA Barhoom, MOHAMMED RASHEED J Al-Hiealy, and SAMY S Abu-Naser. Bone abnormalities detection and classification using deep learning-

vgg16 algorithm. *Journal of Theoretical and Applied Information Technology*, 100(20):6173–6184, 2022.

21. Srikanth Tammina. Transfer learning using vgg-16 with deep convolutional neural network for classifying images. *International Journal of Scientific and Research Publications (IJSRP)*, 9(10):143–150, 2019.
22. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
23. Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. pmlr, 2015.
24. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
25. Bo Huang, Jianhong Liu, Qian Zhang, Kang Liu, Kun Li, and Xinyu Liao. Identification and classification of aluminum scrap grades based on the resnet18 model. *Applied Sciences*, 12(21):11133, 2022.
26. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
27. Nurul Ashikin Samat, Mohd Najib Mohd Salleh, and Haseeb Ali. The comparison of pooling functions in convolutional neural network for sentiment analysis task. In *Recent Advances on Soft Computing and Data Mining: Proceedings of the Fourth International Conference on Soft Computing and Data Mining (SCDM 2020), Melaka, Malaysia, 2020*, pages 202–210. Springer, 2020.
28. SR Paulo. Breast ultrasound image. mendeley data, 2017.
29. Walid Al-Dhabyani, Mohammed Gomaa, Hussien Khaled, and Aly Fahmy. Dataset of breast ultrasound images. *Data in brief*, 28:104863, 2020.
30. Mahesh Gour, Sweta Jain, and T Sunil Kumar. Residual learning based cnn for breast cancer histopathological image classification. *International Journal of Imaging Systems and Technology*, 30(3):621–635, 2020.
31. Yaqi Wang, Lingling Sun, Kaiqiang Ma, and Jiannan Fang. Breast cancer microscope image classification based on cnn with image deformation. In *Image Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27-29, 2018, Proceedings 15*, pages 845–852. Springer, 2018.
32. Erkan Deniz, Abdulkadir Şengür, Zehra Kadiroğlu, Yanhui Guo, Varun Bajaj, and Ümit Budak. Transfer learning based histopathologic image classification for breast cancer detection. *Health information science and systems*, 6:1–7, 2018.
33. Caroline B Gonçalves, Jefferson R Souza, and Henrique Fernandes. Classification of static infrared images using pre-trained cnn for breast cancer detection. In *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 101–106. IEEE, 2021.