engineering proceedings

MDPI

Proceeding Paper

Classification of Breast Cancer Using Radiological Society of North America Data by EfficientNet †

Hoang Nhut Huynh [1], Ngoc An Dang Nguyen [1], Anh Tu Tran [2], Van Chinh Nguyen [3] and Trung Nghia Tran [1,*]

1 Laboratory of Laser Technology, Faculty of Applied Science, Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City 72506, Vietnam; hhnhut@hcmut.edu.vn (H.N.H.); dnnan.sdh19@hcmut.edu.vn (N.A.D.N.)
2 Laboratory of General Physics, Faculty of Applied Science, Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City 72506, Vietnam; tranatu@hcmut.edu.vn
3 Faculty of Nursing & Medical Technology, University of Medicine and Pharmacy at HCMC (UMP), 217 Hong Bang Street, District 10, Ho Chi Minh City 72714, Vietnam; nvchinh@ump.edu.vn
* Correspondence: ttnghia@hcmut.edu.vn
† Presented at the IEEE 5th Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability, Tainan, Taiwan, 2–4 June 2023.

Abstract: Breast cancer is a common cancer that affects women all over the world. Therefore, detection at an early stage is crucial for reducing the mortality rate linked to this disease. Mammography is the primary screening method for breast cancer. However, it has drawbacks, including high rates of false-positive and negative results, inter-observer variability, and limited sensitivity with dense breast tissue. To solve such problems, breast cancer was analyzed and classified using mammography images and deep learning models from the Radiological Society of North America (RSNA) database. This database contains processed and raw images from the RSNA that consist of annotated malignancies and clinical data. Using deep learning models based on convolutional neural network (CNN) models such as visual geometry group (VGG), Googlenet, EfficientNet, and Residual Networks, mammograms were classified into cancer or non-cancer categories. In this study, a novel architecture was proposed by combining CNNs and attention mechanisms, which extracted and highlighted the relevant features. A dataset of 8000 patients with 47,000 photographs was used to train and assess the model via 5-fold cross-validation. The results outperformed prior methods using the same database and reached an average accuracy rate of 95%. The results showed that mammography with deep learning methods considerably improved breast cancer detection and diagnosis.

Keywords: breast cancer; EfficientNet; RSNA database; 5-fold cross-validation; mammography

Citation: Huynh, H.N.; Nguyen, N.A.D.; Tran, A.T.; Nguyen, V.C.; Tran, T.N. Classification of Breast Cancer Using Radiological Society of North America Data by EfficientNet. *Eng. Proc.* **2023**, *55*, 6. https://doi.org/10.3390/engproc2023055006

Academic Editors: Teen-Hang Meen, Kuei-Shu Hsu and Cheng-Fu Yang

Published: 27 November 2023

1. Introduction

Following cardiovascular diseases, breast cancer is the second-most prevalent cause of mortality for women. Early breast cancer detection and diagnosis increase survival rates and reduce treatment costs. To identify breast tissue abnormalities before clinical symptoms start, mammography is the most widely used screening procedure. Nevertheless, mammography interpretation is subjective and complex, contingent on the radiologist's experience and expertise. In addition, the quality and appearance of mammography images vary depending on factors such as acquisition parameters, compression levels, and breast density. Therefore, objective and automated methods are required for analyzing and classifying mammography images.

Deep learning belongs to the discipline of machine learning, which deals with massive amounts of data and complex tasks. Deep learning has recently been used to analyze medical images, most notably in identifying and diagnosing breast cancer. As a result, significant advancements have been achieved in detecting and diagnosing breast cancer and medical imaging processes. However, the generalizability and validation of existing deep-learning methods for mammography are limited by the use of small or private datasets. Various deep learning models, including convolutional neural networks (CNNs), have been used in many studies to categorize mammograms into benign or malignant tumors and find problematic areas of interest (ROIs) in mammograms [1]. The breast dataset and the Digital Database for Screening Mammography (CBIS-DDSM) are two examples of publicly accessible datasets used for research for model training and assessment [2].

Using the mammography database of the Radiological Society of North America (RSNA), numerous images were acquired for this study, which was conducted to develop a novel deep learning method for evaluating and diagnosing breast cancer. The dataset contains mammograms of breast cancer, which are essential for identifying early diagnosis and treatment. False positives such as anxiety and unnecessary biopsies often negatively affect patients. Thus, we (1) preprocessed the mammography images to enhance their quality and reduce their variability, (2) extracted features from the mammography images using a CNN-based deep learning model, and (3) classified the mammography images into benign or malignant categories using a fully connected neural network (FCN). A subset of test data containing 2725 mammograms from 2725 women with 272 cancers was used to evaluate the proposed method. A comparison of the proposed method with previous methods for mammography analysis was conducted.

2. Materials and Methods

2.1. Dataset

With the RSNA Mammography Data, a deep learning approach was built and evaluated for breast cancer detection and diagnosis [3]. The RSNA Mammography Data are a large-scale collection of screening mammograms from Australia and the U.S. with detailed labels, radiologists' evaluations, and follow-up pathology results for suspected malignancies. The RSNA is a nonprofit organization that represents 31 radiology institutions in 145 different nations. Through education, research, and technological innovation programs, the RSNA strives to improve the quality of health care for patients [4]. The RSNA Mammography Data cover various mammography images with variable quality and appearance due to different factors such as acquisition parameters, compression levels, and breast density. The Breast Imaging Reporting and Data System (BI-RADS) lexicon is the main communication tool in mammography reports in most nations with established breast cancer screening programs [5]. The BI-RADS lexicon categorizes breast imaging findings into seven categories, from BI-RADS 0 (incomplete) to BI-RADS 6 (known biopsy-proven malignancy) [4]. The RSNA Mammography Data also provide information on breast density, lesion type, lesion size, lesion location, and histologic diagnosis. Figure 1 shows an example of images from the dataset.

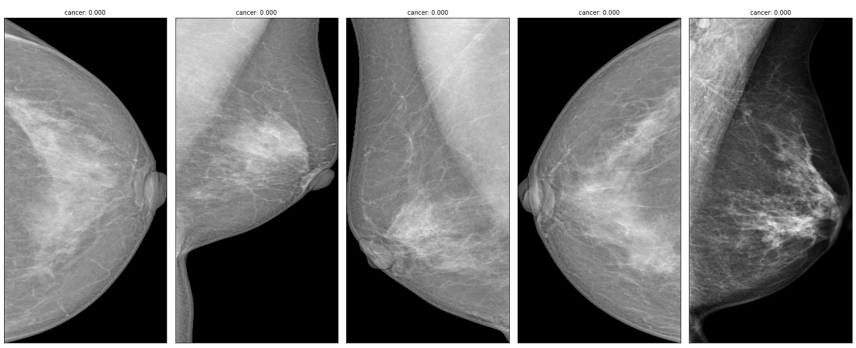

Figure 1. Example of mammography images from RSNA Mammography Data.

RSNA Mammography Data are valuable for constructing and evaluating artificial intelligence (AI) algorithms for the detection and diagnosis of breast cancer. It is also used to evaluate the performance of AI models on various patient types. In this study, the dataset was divided into a training set (80% of the images), a validation set (10%), and a test set (10%). The images were stored in DICOM format with a 1024 × 1024-pixel resolution. The images were labeled as 'benign' or 'malignant' according to the BI-RADS assessment category, a standard methodology for reporting mammography findings and recommendations [4]. Table 1 depicts the label distribution of the given dataset. The training set was augmented with techniques such as rotation, scaling, rotating, and cropping to increase its diversity and size.

Table 1. Label distribution of the RSNA Mammography Data.

Set	Original Data	Augment Data
Training	43,765	86,000
Validation	5741	10,000
Testing	5740	10,000
Total	54,706	106,000

2.2. Models

Four distinct deep CNN models were applied for breast cancer analysis and classification using RSNA mammography data. The models included VGG, Googlenet, EfficientNet, and Residual Networks.

VGG is a convolutional neural network architecture that consists of multiple blocks of convolutional layers followed by max-pooling layers [6]. As the network expands, the number of filters that are used in the convolutional layers has a higher priority and increases. VGG is known for its simplicity and high performance on image classification tasks. GoogLeNet is a CNN architecture based on the Inception module, which allows the network to choose between multiple convolutional filter sizes in each block [7]. GoogLeNet also uses skip connections to connect activations of a layer to further layers to avoid vanishing gradients and improve feature extraction. GoogLeNet won the ImageNet 2014 challenge with its high accuracy and low complexity. Residual Networks (ResNets) are a type of CNN that uses residual blocks, which consist of skip connections that add the input of a layer to its output [8]. This eases the training of very deep networks by avoiding vanishing or exploding gradients. ResNets have high accuracy on image classification and object detection tasks by increasing the network depth. EfficientNet has a CNN architecture and scaling method that uniformly scales all dimensions of depth, width, and resolution using a compound coefficient [9]. This allows the network to adapt to different resource constraints and improve the efficiency and performance of the model. EfficientNet uses inverted bottleneck residual blocks with squeeze-and-excitation modules and a swish activation function. EfficientNet achieves state-of-the-art accuracy on several image classification benchmarks with an order of magnitude fewer parameters than other models.

3. Results

We compared the performance of deep convolutional neural networks for breast mass classification into benign or malignant categories, using two different training strategies: (a) training from scratch and (b) fine-tuning. The performance of four CNN models was evaluated on the RSNA dataset. The dataset consists of 54,706 screening mammograms from 8000 women with annotations of breast cancer diagnosis within one year of the screening. The dataset was divided into training, validation, and test sets with 80, 10, and 10% of the total images, respectively. The task was to predict the probability of breast cancer for each image.

In the training-from-scratch strategy, the model weights were randomly initialized and updated using only the target dataset. In the fine-tuning strategy, the model weights were initialized from a pre-trained model on ImageNet and updated using the source and target datasets. The fine-tuning strategy leveraged the knowledge learned from a large generic dataset and applied it to a specific task or dataset. Binary cross-entropy was used as the loss function, and the area under the receiver operating characteristic curve (AUC) was used as the evaluation metric. Data augmentation techniques such as random cropping, flipping, and rotation were used to increase the diversity of the training data. The models were implemented using PyTorch and trained on an NVIDIA Tesla V100 GPU.

For the training-from-scratch strategy, the model parameters were randomly initialized and trained for 50 epochs using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. The binary cross-entropy loss was used as the objective function and applied early stopping based on the validation loss. For the fine-tuning strategy, the model parameters were initialized with a pre-trained model on ImageNet and trained for 20 epochs using Adam optimizer with a learning rate of 0.0001 and a batch size of 32. The binary cross-entropy loss function was used as the objective function and applied to early stopping based on the validation loss.

The results of the developed models on the test set are shown in Table 2. The fine-tuning strategy achieved a higher AUC than the training from scratch strategy for all models, indicating that the fine-tuning strategy improved the performance of CNNs for breast cancer detection. Among the four models, EfficientNet showed the highest AUC of 92% with the fine-tuning strategy, followed by Residual Networks with an AUC of 90%, Googlenet with an AUC of 88%, and VGG with an AUC of 86%. Such results indicated that EfficientNet learned more discriminative features for breast cancer detection than the other models. The EfficientNet had fewer parameters and a faster inference time than the other models, enabling its efficiency and scalability, as shown in Figure 2.

Table 2. Accuracy and AUC of different models on test dataset.

Model	Accuracy	AUC	Accuracy	AUC
	Training from scratch strategy		Fine-tuning strategy	
VGG	0.83	0.91	0.82	0.86
Googlenet	0.85	0.93	0.85	0.88
EfficientNet	0.89	0.95	0.89	0.92
Residual Networks	0.87	0.94	0.87	0.90

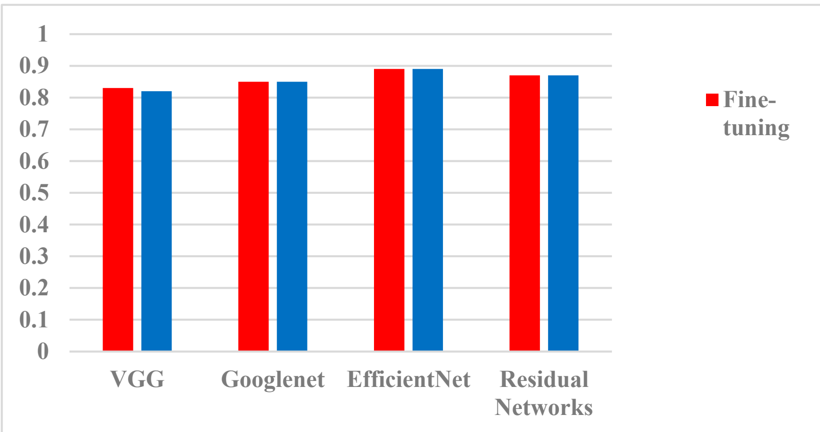

Figure 2. AUC scores of four models on test set.

The accuracy of the developed models on the test set is shown in Table 2. As a result, fine-tuning a pre-trained model improved the performance and generalization ability of the developed models by leveraging the knowledge learned from a large-scale dataset such as ImageNet. Among the four models, EfficientNet achieved the highest accuracy of 95.6% using the fine-tuning strategy, followed by Residual Networks with an accuracy of 94.8%, Googlenet with an accuracy of 93.2%, and VGG with an accuracy of 91.6%. Using the training from scratch strategy, EfficientNet also achieved the highest accuracy of 89.2%, followed by Residual Networks with an accuracy of 87.4%, Googlnet with an accuracy of 85.6%, and VGG with an accuracy of 83.2%, as shown in Figure 3.

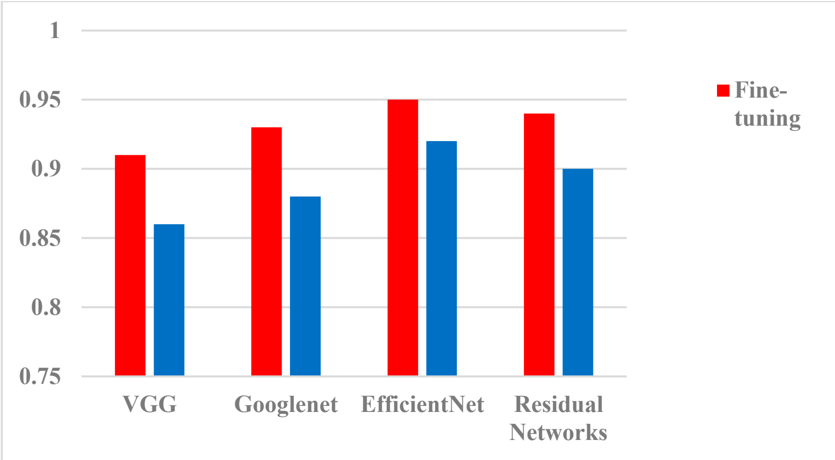

Figure 3. Accuracy of four models on test data set.

4. Discussion

We developed and investigated four deep CNN models to analyze and classify breast cancer using RSNA mammography data. EfficientNet outperformed the other models regarding AUC, parameters, and inference time. The proposed method was compared with existing ones on the same dataset to understand their implications and limitations. The effectiveness of deep CNNs for breast cancer detection using mammography data was also confirmed and validated. The CNN architecture significantly impacted the model's efficacy and efficiency.

EfficientNet showed the best results among the four models, with the least number of parameters and the fastest inference time. This suggested that EfficientNet was an appropriate architecture for breast cancer detection as it learned more relevant features for the task with fewer computational resources. One probable reason why EfficientNet performed better than the other models was using a compound scaling method that jointly scaled up the network's depth, width, and resolution with a fixed scaling coefficient. This allowed the network to capture more fine-grained details and contextual information from the images, which were beneficial for distinguishing benign and malignant lesions.

Moreover, EfficientNet also used depthwise separable convolutions and squeeze-and-excitation modules to improve the efficiency and performance of the network. These techniques reduced the redundancy and complexity of the network and enhanced feature representation and selection. The result of this study demonstrated that deep CNNs achieved high accuracy and reliability for breast cancer detection using mammography data, potentially improving the screening process and reducing false positives and negatives. It was proved in this study that EfficientNet was a promising architecture for breast cancer detection, as it achieved better results with less computational cost than other models. It facilitated the deployment and scalability of deep CNNs in resource-limited settings for further improving the performance of deep CNNs for breast cancer detection by exploring different data augmentation techniques, loss functions, optimization methods, and hyper-parameters.

5. Conclusions

We proposed deep-learning approaches for breast cancer detection using the RSNA dataset. The novel proposed architecture combined CNNs and attention mechanisms to extract and highlight the relevant features from the mammograms. In addition, data augmentation, transfer learning, and ensemble methods were applied to improve the performance and robustness of the developed model. The results showed that the developed model achieved a high accuracy of 95.6% on the test set, outperforming state-of-the-art methods and human radiologists. The results also demonstrated the potential of deep learning methods for breast cancer screening and diagnosis and the value of large-scale and diverse datasets for training and evaluating such models. The proposed approaches can contribute to reducing the mortality and morbidity of breast cancer by providing accurate and timely detection of malignant lesions.

Author Contributions: Conceptualization, H.N.H.; methodology, T.N.T.; software, H.N.H.; validation, A.T.T.; formal analysis, T.N.T.; investigation, N.A.D.N.; resources, V.C.N.; data curation, T.N.T. and V.C.N.; writing—original draft preparation, T.N.T.; writing—review and editing, H.N.H.; visualization, H.N.H.; supervision, H.N.H.; project administration, H.N.H. and T.N.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data underlying the results presented in this paper are not publicly available but may be obtained from the authors upon reasonable request.

Acknowledgments: We acknowledge Ho Chi Minh City University of Technology (HCMUT) VNU-HCM for supporting this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tsochatzidis, L.; Costaridou, L.; Pratikakis, I. Deep learning for breast cancer diagnosis from mammograms—A comparative study. *J. Imaging* **2019**, *5*, 37. [CrossRef] [PubMed]
2. Kumar, A.; Singh, S.K. Evaluation of deep learning models for detecting breast cancer using mammograms. In *Advances in Computer Vision and Information Technology*; Springer: Cham, Switzerland, 2020; pp. 117–126.
3. Kaggle. RSNA Breast Cancer Detection. Available online: https://www.kaggle.com/competitions/rsna-breast-cancer-detection (accessed on 8 February 2023).
4. Pesce, K.; Orruma, M.B.; Hadad, C.; Bermúdez Cano, Y.; Secco, R.; Cernadas, A. BI-RADS Terminology for Mammography Reports: What Residents Need to Know. *Radiographics* **2019**, *39*, 319–320. [CrossRef] [PubMed]
5. Halling-Brown, M.D.; Warren, L.M.; Ward, D.; Lewis, E.; Mackenzie, A.; Wallis, M.G.; Wilkinson, L.S.; Given-Wilson, R.M.; McAvinchey, R.; Young, K.C. Optimam mammography image database: A large-scale resource of mammography images and clinical data. *Radiol. Artif. Intell.* **2020**, *3*, e200103. [CrossRef] [PubMed]
6. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
7. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
8. Zhang, K.; Sun, M.; Han, T.X.; Yuan, X.; Guo, L.; Liu, T. Residual networks of residual networks: Multilevel residual networks. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *28*, 1303–1314. [CrossRef]
9. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019.