다단계 전이 학습을 이용한 유방암 초음파 영상 분류 응용

겔란 아야나 · 박진형 · 최세운*
금오공과대학교

Proper Base-model and Optimizer Combination Improves Transfer Learning Performance for Ultrasound Breast Cancer Classification

Gelan Ayana · Jinhyung Park · Se-woon Choe*
Kumoh National Institute of Technology
E-mail : gelan@kumoh.ac.kr / 20150555@kumoh.ac.kr / sewoon@kumoh.ac.kr

요 약

인공지능 알고리즘을 이용한 유방암의 조기진단에 관련된 연구는 최근들어 활발하게 진행되고 있으나, 사용자의 목적에 맞는 처리속도 및 정확도 등에 다양한 한계점을 보인다. 이러한 문제를 해결하기 위해, 본 논문에서는 ImageNet에서 학습된 ResNet 모델을 현미경 기반 암세포 이미지에서 활용이 가능한 다단계 전이 학습을 제안하고, 이를 다시 전이 학습하여 초음파 유방암 영상을 양성 및 악성으로 분류하는 실험을 진행하였다. 제안된 다단계 전이 학습 알고리즘은 초음파 유방암 영상을 분류하였을 때 96% 이상의 정확도를 보였으며, 향후 암 세포주 및 실시간 영상처리 등의 추가를 통해 보다 높은 활용도와 정확도를 보일 것으로 기대한다.

ABSTRACT

It is challenging to find breast ultrasound image training dataset to develop an accurate machine learning model due to various regulations, personal information issues, and expensiveness of acquiring the images. However, studies targeting transfer learning for ultrasound breast cancer images classification have not been able to achieve high performance compared to radiologists. Here, we propose an improved transfer learning model for ultrasound breast cancer classification using publicly available dataset. We argue that with a proper combination of ImageNet pre-trained model and optimizer, a better performing model for ultrasound breast cancer image classification can be achieved. The proposed model provided a preliminary test accuracy of 99.5%. With more experiments involving various hyperparameters, the model is expected to achieve higher performance when subjected to new instances.

키워드
Transfer learning, breast cancer, ultrasound, classification, EfficientNetB2, Adam

Ⅰ. Introduction

Breast cancer was a reason for death of more than 600, 000 women worldwide in the year 2020 alone [1]. Death from breast cancer could be reduced by 40% if breast cancer is detected at early stage [2]. Ultrasound is a better imaging modality in detecting early breast cancer compared to other breast cancer diagnosis modalities [3]. However, ultrasound based breast cancer diagnosis is susceptible to false negative and false positive results that its analysis is too subjective [4]. Moreover, ultrasound is not a standalone breast cancer detection modality that confirmation of results with biopsy is needed. These shortcomings of ultrasound motivated researchers to sought out the application of deep learning to automatically detect breast cancer [5]. Nevertheless, the lack of large training data as well as a robust algorithm performing well on ultrasound images hindered its success [6]. This prompted for the application of transfer learning for ultrasound breast cancer

* corresponding author

classification. Transfer learning (TL) enable knowledge transfer from a model pre-trained on a large dataset, such as ImageNet, for developing a better performing model on a different target with small number of dataset for training [7]. Despite the improvement in performance from using TL, it is yet to be used in a clinical setting. The studies that applied TL for ultrasound provided test accuracy of less than 93% on publicly available dataset [7]. Therefore, there is a need for a high performing model so that it would be used in clinical setting to assist radiologists thereby decreasing the false positive and false negative rates. Here we propose an EfficientNetB2 based TL model with adam optimizer for classification of ultrasound breast cancer images as benign or malignant. In this study, we try to show that with proper combination of pre-trained base model and optimizer, a high performance model for classification of ultrasound breast cancer images could be developed.

II. Materials and method

Here we applied TL from ImageNet pre-trained EfficientNetB2 model to ultrasound images; only the last layer was removed and global average pooling was added, and one dense layer with Sigmoid was utilized, the rest is similar to the original architecture of EfficientNetB2 base model as shown in Figure 1.

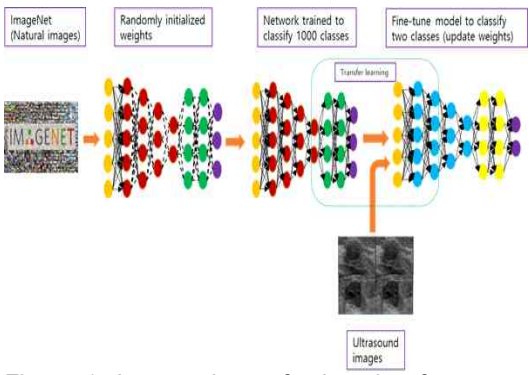

Figure 1. Improved transfer learning for ultrasound breast cancer classification.

2.1. Dataset

For this study, we used a publicly available dataset called Mendeley breast ultrasound dataset, which is composed of 250 images (150 malignant and 100 benign cases) [8]. The data size used is 200 images, 100 from each class, to prevent class imbalance and bias from using unbalanced data. the data was categorized into 6:2:2 ratio for training, validation, and test, consecutively. We employed augmentation to increase the training dataset to 360 via vertical flip and rotation.

2.2. Implementation

The proposed method was implemented on the Keras on TensorFlow framework using Python. Two pieces of RTX 3090 GPUs were employed to accelerate the training. The model was trained for 30 epochs. Adam was utilized as an optimizer. We fine tuned all the weights from the ImageNet pre-trained model with a learning rate that decays exponentially, starting from 0.001. the training batch size was 16.

III. Results and Future Work

Average nested five-fold cross validation test accuracy was used to evaluate the performance of the proposed method. Consequently, the proposed method provided test accuracy of 99.5% averaged over five-fold cross validation results, see table 1. the model also provided a small loss ranging from 0.009 to 0.3, which implies that the model is performing well in generalizing for the dataset it was not trained on.

Table 1. Five-fold cross validation results.

Cross validation	Test accuracy (%)	Average 5-fold cross validation test accuracy (%)
Fold 1	100	99.5%
Fold 2	100	
Fold 3	100	
Fold 4	97.5	
Fold 5	100	

Additionally, we made a cross-sectional comparison with TL based studies published for classifying ultrasound breast cancer images. The proposed model outperformed all the published works in terms of test accuracy as shown in Table 2.

Table 2. Comparison of the prosed method with state-of-the-art methods.

Paper	CNN	Pre-training dataset	Application	Dataset	Pre-processing	Performance
Byra et al. [9]	VGG19	ImageNet	Classification	UDIAT and OASBUD	Matching layer	• Accuracy=0.830± 0.026 (OASBUD)
Hijab et al. [10]	VGG16	ImageNet	Classification	Baheya Foundation	Augmentation	• Accuracy=97.39 %
Yap et al. [11]	AlexNet	ImageNet	Classification	Dataset A and B	Splitting in to patches	• F1 measure=0.91 (Dataset A) • F1 measure=0.89 (Dataset B)
The proposed method	EfficientNet B2	ImageNet	Classification	Mendeley	Augmentation	• Accuracy=99.5%

Acknowledgement

이 논문은 2020년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2019R1F1A1062397). 본 논문은 4단계 BK21 사업(금오공과대학교 IT융복합학공학과)에 의하여 지원되었음.

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA. Cancer J. Clin.,* vol. 70, no. 1, pp. 7-30, Jan. 2020.

[2] D. A. Berry et al., "Effect of Screening and Adjuvant Therapy on Mortality From Breast Cancer," *Obstet. Gynecol. Surv.,* vol. 61, no. 3, pp. 179-180, Mar. 2006.

[3] R. Guo, G. Lu, B. Qin, and B. Fei, "Ultrasound Imaging Technologies for Breast Cancer Detection and Management: A Review," *Ultrasound Med. Biol.,* vol. 44, no. 1, pp. 37-70, Jan. 2018.

[4] J. Geisel, M. Raghu, and R. Hooley, "The Role of Ultrasound in Breast Cancer Screening: The Case for and Against Ultrasound," *Semin. Ultrasound, CT MRI,* vol. 39, no. 1, pp. 25-34, Feb. 2018.

[5] S. Liu et al., "Deep Learning in Medical Ultrasound Analysis: A Review," *Engineering*, vol. 5, no. 2, pp. 261-275, Apr. 2019.

[6] N. Houssami, G. Kirkpatrick-Jones, N. Noguchi, and C. I. Lee, "Artificial Intelligence (AI) for the early detection of breast cancer: a scoping review to assess AI's potential in breast screening practice," *Expert Rev. Med. Devices,* vol. 16, no. 5, pp. 351-362, May 2019.

[7] G. Ayana, K. Dese, and S. Choe, "Transfer Learning in Breast Cancer Diagnoses via Ultrasound Imaging," *Cancers (Basel).*, vol. 13, no. 4, p. 738, Feb. 2021.

[8] P. S. Rodrigues, "Breast Ultrasound Image," Mendeley Data, 2018.

[9] M. Byra et al., "Breast mass classification in sonography with transfer learning using a deep convolutional neural network and color conversion," *Med. Phys.,* vol. 46, no. 2, pp. 746-755, Feb. 2019.

[10] A. Hijab, M. A. Rushdi, M. M. Gomaa, and A. Eldeib, "Breast Cancer Classification in Ultrasound Images using Transfer Learning," *2019 Fifth International Conference on Advances in Biomedical Engineering (ICABME),* Vol. 2019-Octob, pp. 1-4, Oct. 2019.

[11] M. H. Yap et al., "Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks," *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 4, pp. 1218-1226, Jul. 2018.