

# 유방암 조기 진단을 위한 초음파 영상의 다단계 전이 학습

겔란 아야나 · 박진형 · 최세운\*

금오공과대학교

## Multistage Transfer Learning for Breast Cancer Early Diagnosis via Ultrasound

Gelan Ayana · Jinhyung Park · Se-woon Choe<sup>\*</sup>

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 있다. 이는 연구용으로 공개된 초음파 유방 이미지를 활용하여 다양하게 개발되고 있으나, 사용자의 목적에 맞는 처리 속도 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. 실험을 위한 영상은 양성과 악성이 포함된 250장의 유방암 초음파 영상과 27,200장의 암 세포주 영상으로 구성되었다. 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

Research related to early diagnosis of breast cancer using artificial intelligence algorithms has been actively conducted in recent years. Although various algorithms that classify breast cancer based on a few publicly available ultrasound breast cancer images have been published, these methods show various limitations such as, processing speed and accuracy suitable for the user's purpose. To solve this problem, in this paper, we propose a multi-stage transfer learning where ResNet model trained on ImageNet is transfer learned to microscopic cancer cell line images, which was again transfer learned to classify ultrasound breast cancer images as benign and malignant. The images for the experiment consisted of 250 breast cancer ultrasound images including benign and malignant images and 27,200 cancer cell line images. The proposed multi-stage transfer learning algorithm showed more than 96% accuracy when classifying ultrasound breast cancer images, and is expected to show higher utilization and accuracy through the addition of more cancer cell lines and real-time image processing in the future.

### 키워드

Multistage transfer learning; breast cancer; ultrasound

### 1. Introduction

Breast cancer is screened using the state of the art diagnosis methods such as, Mammogram, Ultrasound and tissue biopsy [1]. Nevertheless, to confidently determine the presence of breast cancer, all the above screening methods must be used

together that it is impossible to conclude based on only one modality [2]. Furthermore, an experienced radiologist must verify results based on the findings from all these modalities [3]. As a result, breast cancer diagnosis is expensive, takes a long time to decide result, needs experienced professionals, and after all not accurate [4]. Ultrasound imaging is the most popular modality among the three frontline methods for the early diagnosis of breast cancer

---

\* corresponding author

especially in young women with dense breast [5]. However, the ultrasound method itself is dependent on professionals to retrieve results. In order to solve the problem of dependency on professionals, a few studies introduced the application of artificial intelligence that enabled a computer algorithm to identify if a given ultrasound breast image is cancerous or not [6]. Despite the improvements in diagnosis of breast cancer with the application of machine learning, the unavailability of large training data is a challenge [7]. With advances in the field of machine learning, researchers came up with a phenomenon called transfer learning where a model trained on large dataset for a given task is used as a pre-training model for another task [8]. The state of the art studies employed a transfer learning approach from natural images (ImageNet) not from medical images and many studies proved that transfer learning achieves better result when source and target tasks are of the same domain [8]. Here, we propose a novel multistage transfer learning algorithm for ultrasound breast cancer images two class classification as malignant and benign.

## II. Materials and method

### 2.1. Transfer learning

The proposed multistage transfer learning method involves transfer learning from an ImageNet pre-trained ResNet model to cancer cell lines microscopic images that is in turn transfer learned on ultrasound breast cancer images to classify the images as malignant and benign. In the first stage, we applied transfer learning from ImageNet, natural images, to cancer cell lines microscopic images. This stage changes the normal domain to medical image domain by extracting medical features of images automatically, and then, in stage 2, we used stage 1 as a starting point and assign weights to images of ultrasound breast cancer to categorize the images as malignant/benign, Figure 1.

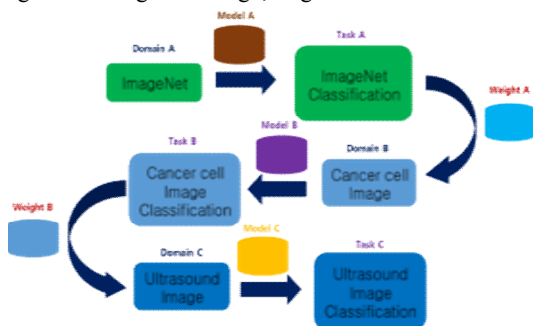


Figure 1. Multistage transfer learning for

ultrasound breast cancer early diagnosis.

### 2.2. Datasets and pre-processing

The ultrasound image data is obtained from the publicly available Mendeley Data ultrasound images [9]. The dataset is composed of 250 breast ultrasound images of which 150 is malignant case and 100 is benign case.

Cancer cell line microscopic images used in this study are, HELA (cervical cancer cell line), Huh-7 (liver cancer cell line), NCI-H1299 (lung cancer cell line), and MCF-7 (breast cancer cell line), 6,800 patches each totally being 27,200 patches. The cancer cell lines were acquired under inverted optical microscope (IX73, Olympus, Japan). The cell lines were bought from the Korean cell line bank and all the cell lines were cultured under 37°C. The pre-processing steps performed include, adaptive thresholding to binarize the acquired image followed by selecting the area of interest using OpenCV bounding box. In order to prevent the outline of a small floating object other than the cell from being detected, an object with a number of pixels less than a certain number is removed using `remove_small_objects` of Scikit-image. The selected cell region is extracted as input image for learning and validation with size  $200 \times 200$  pixels.

### 2.3. Convolutional Neural Network (CNN) model

In transfer learning from ImageNet pre-trained ResNet to cancer cell line images, only the last layer is removed and global average pooling was added as well as one dense layer with Softmax. In the transfer learning from cell line image pre-trained model, the dense layer is removed and replaced with three dense layers and one drop-out layer and lastly a dense layer with sigmoid function is added to give output. The proposed multistage transfer learning for breast cancer classification is implemented based 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 at each transfer learning stages. The gradient optimizer used was Adagrad. During the training, the learning rate was 0.001, which was the same throughout the epochs. The training batch size was 16. The ultrasound image dataset was categorized in to 175 training, 50 validation, and 25 test sets with a ratio of 7:2:1, consecutively. Augmentation was used to increase the number of training images by three folds, which affected the accuracy significantly.

### III. Results and Future Work

Preliminary results of experiments with the proposed model is presented in Table 1. The accuracy without using augmentation was 87% by only replacing the Softmax layer of the cell line pre-trained model with a sigmoid function. By trying different amount and type of augmentation, shift and rotation with augmentation rate of three folds achieved the best result, accuracy of 90%. We then added regularization to the model and achieved accuracy of 92%. However, the loss was as high as 12, which needed an improvement. Here, we decided to use drop out instead of regularization that resulted in additional three dense layers and a drop out layer, which resulted in the proposed method. The proposed system achieved classification accuracy of 96% with a loss of 0.1 and training time of 400 seconds.

Table 1: Preliminary results.

Experiment	Loss	Accuracy
ResNet50 + Adagrad	4	87%
ResNet50 + Adagrad + Augmentation	0.8	90%
ResNet50 + Adagrad + Augmentation + Regularization	12	92%
ResNet50 + Adagrad + Augmentation + Drop out	0.1	96%

### Acknowledgement

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2019R1F1A1062397). This work was supported by the Brain Korea 21 FOUR Project (Dept. of IT Convergence, Kumoh National Institute of Technology)

### References

[1] A. M. Rocha García and D. Mera Fernández, "Breast tomosynthesis: State of the art," *Radiologia*, Vol. 61, Iss. 4, pp. 274 - 285, Jul. 2019.

[2] M. J. Silverstein, M. D. Lagios, A. Recht, D. C. Allred, S. E. Harms, R. Holland, D. R. Holmes, L. L. Hughes, R. J. Jackman, T. B. Julian, H. M. Kuerer, H. C. Mabry, D. R. McCready, K. M. McMasters, D. L. Page, S. H. Parker, H. A. Pass,

M. Pegram, E. Rubin, A. T. Stavros, D. Tripathy, F. Vicini and P. W. Whitworth, "Image-detected breast cancer: state of the art diagnosis and treatment," *Journal of the American College of Surgeons*, Vol. 201, No. 4, pp. 586-597, Oct. 2005.

[3] D. A. Shumway, A. Sabolch, and R. Jagsi, "Breast Cancer," In: Lo S., Teh B., Jiang GL., Mayr N. (eds) *Controversies in Radiation Oncology*. Medical Radiology, Germany, Springer, 2016.

[4] A. B. Nover, S Jagtap, W. Anjum, H. Yegingil, W. Y. Shih, W. Shih, and A. D. Brooks, "Modern breast cancer detection: A technological review," *International Journal of Biomedical Imaging*, Vol. 2009, 902326, Sep. 2009.

[5] P. Murawa, D. Murawa, B. Adamczyk, and K. Połom, "Breast cancer: Actual methods of treatment and future trends," *Reports of Practical Oncology and Radiotherapy*, Vol. 19, No. 3, pp. 165-172, Jan. 2014.

[6] S. Niu, J. Huang, J. Li, X. Liu, D. Wang, R. Zhang, Y. Wang, H. Shen, M. Qi, Y. Xiao, M. Guan, H. Liu, D. Li, F. Liu, X. Wang, Y. Xiong, S. Gao, X. Wang and J. Zhu, "Application of ultrasound artificial intelligence in the differential diagnosis between benign and malignant breast lesions of BI-RADS 4A," *BMC Cancer*, Vol. 20, No. 1, pp. 1-7, Oct. 2020

[7] G. Ayana, K. Dese, and S. Choe, "Transfer Learning in Breast Cancer Diagnoses via Ultrasound," *Cancers*, Vol. 13, No. 4, pp. 1-15, Feb. 2021

[8] M. Byra, M. Galperin, H. Ojeda-Fournie, L. Olson, M. O'Boyle, C. Comstock and M. Andre, "Breast mass classification in sonography with transfer learning using a deep convolutional neural network and color conversion," *Medical Physics*, Vol. 46, No. 2, pp. 746-755, Feb. 2019

[9] P. S. Rodrigues, "Breast Ultrasound Image", *Mendeley Data*, 2018. Available: <https://data.mendeley.com/datasets/wmy84gzngw/1>