

Deep Learning Model Aids Breast Cancer Detection

Quan Zhang^{1,*}, Guoqing Cai¹, Meiqing Cai², Jili Qian¹, Tianbo Song³

¹ Information Studies, Trine University, Phoenix AZ, USA

² Computer Science, Independent researcher, Nanchang, China

³ Computer Engineering, Arizona State University, Tempe, AZ, USA

* Corresponding author: Quan Zhang (Email: wayne168zhang@gmail.com)

Abstract: Breast cancer, a lumpy nodule or granular calcified tissue caused by cancerous changes in chest tissue, has become one of the most prevalent cancers. Due to the location and structure of the tumor, it can be detected directly by ultrasound or X-ray and is less likely to spread to other parts of the body than tumors in other parts of the body. Considering the huge number of sick people, the resources required for a full census would be enormous, but thanks to the rapid development of medical image processing technology in recent years, assisted diagnosis through deep learning models has gradually become more widely accepted. For detection models, higher accuracy means lower misdiagnosis rates and timely treatment for patients. Therefore, in this paper, we first specify the diagnose as a binary classification problem and then introduce a new pooling scheme and training method to achieve better results compared to the traditional network backbone in the past.

Keywords: Breast Cancer Detection; Medical Image Analyze; Vision Model; Deep Learning.

1. Introduction

Breast cancer stands out as one of the most prevalent cancers globally, constituting 11.7% of all new cancer cases worldwide. Characterized by clinical manifestations such as discernible breast lumps or anomalous thickening distinct from surrounding tissues, coupled with alterations in breast size, shape, or appearance, breast cancer remains a critical health concern. The disease is amenable to early to mid-stage diagnosis, primarily involving the scrutiny of burr masses and clusters of minute calcified material through x-ray diffraction analysis and multispectral ultrasound analysis.

In recent years, the paradigm of breast cancer detection has witnessed a notable shift towards visual models. This transition can be attributed to the distinctive characteristics of breast lesions, including high contrast and integrity, infrequent occurrence of fatty breast glands, and the potential for visualizing lesion locations from multiple perspectives in x-ray radiography. The efficacy of visual models hinges on the precision of lesion segmentation and classification, a fact that substantially improves with the increasing capability of the model to extract features from input images.

Neural network models, particularly those deployed for breast cancer detection, play a pivotal role in enhancing clinical diagnosis on two fronts. Firstly, in the realm of initial diagnosis, machine-led screenings outshine traditional methods where doctors individually evaluate X-ray results. Furthermore, the accurate determination and segmentation of lesion locations furnish doctors with valuable references, expediting the diagnostic process. This paper centers its focus on the latter aspect, delving into an in-depth analysis by refining the pooling layer and associated training scheme to enhance the classification performance of a conventional deep convolutional network-based model.

Anticipating a positive contribution to medical assistive systems, we posit that our work will propel the evolution of deep learning-based models, enabling swifter and more precise initial clinical examinations. This, in turn, holds promise for advancing the landscape of breast cancer diagnosis, fostering a future where cutting-edge technology

collaborates seamlessly with medical expertise for improved patient outcomes.

2. Related Work

Even before the deep learning revolution, early researchers were already working through methods like spectral analysis or traditional machine learning methods [1-2]. Given a classification task, in [3], as compared to the new model, the authors compare in detail the classification results based on SVM, KNN, etc., and likewise give some analysis for the preprocessing of this class of images. Different from innovation from model itself, lots of works proves a better feature selection will boost performance compare with the classic approach [4]: [7] proposed a method which use multi-resolution wavelets and Zernike moments to get a better representation; [8] utilize Gabor wavelet and locality sensitive discriminant analysis in feature selection process. [9] introduce a Cross- correlation analysis in both spatial and wavelet domains before using PCA as feature selection method.

Since 2017, as various types of novel network architectures have achieved good results in their respective fields [10], more and more work has utilized them for feature extraction of breast cancer images: by combine FCM (Fuzzy C-means) and NN structure, [5] illustrates a better starting feature representation is equally important for the final performance of the model. [6] fuse an attention module right after the feature selection. [8] introduce organize the data as a graph structure to apply graph convolution. Considering that healthcare datasets are generally less public and smaller in size, [9] proposes new data enhancement schemes to improve data quality.

3. Methodology

In this section, we initiate our methodology by analyzing the average age of onset and the spatial localization of breast lesions within the dataset. This involves a detailed examination of age distribution trends and the prevalence of lesions following breast visualization. Subsequently, we

provide specifics about the model employed in our study, detailing its architecture, key components, and any customized modifications made. This concise overview ensures transparency in our analytical framework, laying the groundwork for the application of the chosen model.

3.1. Data Analyze

In Figure 1, a comprehensive analysis of patient age and the corresponding number of positive cases reveals that the average age of disease onset is approximately 67 years old. Notably, the overall dataset's average age is approximately 60 years old. This detailed breakdown provides insights into the demographic characteristics of the study population.

Examining the positive samples in Figure 2, our analysis indicates an incidence rate of roughly 1 in 15 within the total population. This finding aligns with the broader disease incidence reported by the World Health Organization (WHO), bolstering the representativeness of our dataset. Given the inherent class imbalance in this binary classification problem, we employ strategies during the training phase to balance positive and negative samples. Techniques such as rotation, clipping, and other augmentation schemes are applied to address this imbalance, ensuring robust model training and performance across both classes.

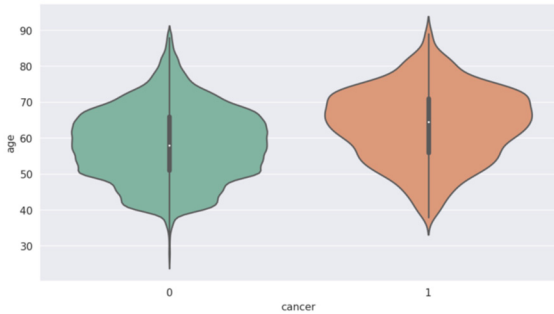


Figure 1. The age distribution of patients

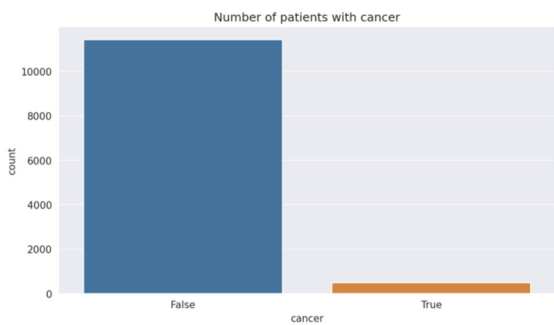


Figure 2. The Cancer distribution of patients

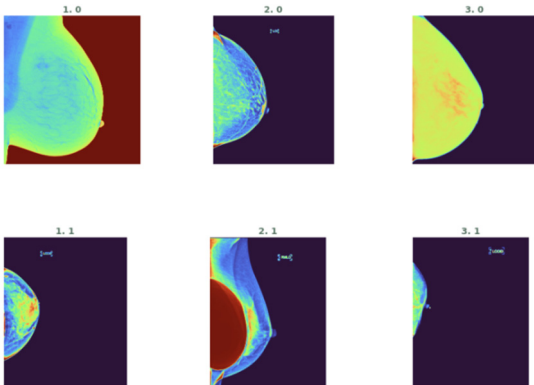


Figure 3. The images of Breast Cancer

The Figure 3 give the heatmap of the breast cancer images of the patients from different ages which are preprocessed from the X-ray images.

3.2. Model Architectures

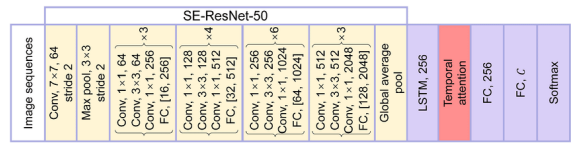


Figure 4. The architecture of SE-ResNet-50

The base model we adopted here is SE-ResNet-50 [11]. Different from the previous ResNet or VGG16, the core idea in SE (Squeeze and Excitation) method is take the relationship between different channel into consideration. To explicitly model the interdependencies between feature channels, a learning-based method was adopted here to get the importance of different channel automatically, and then enhance the useful features according to this importance and suppress the features that are not useful for the current task. The architecture of SE-ResNet-50 is shown in the Figure 4.

Specifically, given a feature which has c_0 in channel, we compress the feature along spatial dimension turning each two-dimensional feature channel into a real number. This number will somehow have a global receptive field and an output whose dimensionality matches the number of input feature channels. It characterizes the global distribution of the response over the feature channels, and it makes the global receptive field available to layers close to the input as well. The excitation is a mechanism like gates in recurrent neural networks. Weights are generated for each feature channel by means of parameters, which are learned to explicitly model the correlation between features channels. Considering the flexibility of the number of layers of the network, we here incorporate this design into ResNet, obtaining the structure shown above. The kernel of input layer is the largest: 7×7 , which will give the following layer a bigger reception field. The design in next 3 blocks are kept the same as original ResNeXt [15].

After that, we change the average pooling mentioned to GEM (generalized mean pooling) pooling, which can be described as the follows:

$$\mathbf{e} = \left[\left(\frac{1}{|\Omega|} \sum_{u \in \Omega} x_{cu}^p \right)^{\frac{1}{p}} \right]_{c=1, \dots, C}$$

Here x is the input of layer, p is a hyper-parameter given by users, when $p = 1$ the GEM will degrade to average pooling, and when p is infinity, GEM will degrade to max pooling. In general, compared to maximum pooling, which wants to capture local high-frequency features of an image, and average pooling, which is a global blur, GEM is a more eclectic scheme. Then we choose to adopt Adam instead of classic SGD optimizer here to accelerate the speed of convergence. The Adam makes the step size can be adaptively adjusted from both the gradient mean and gradient squared perspectives, rather than being directly determined by the current gradient.

4. Experiments

4.1. Experiments Setting and Evaluation Indexes

We obtain the data from RSNA screening Mammography Breast Cancer Detection Open Challenge. During training process, we set learning rate to 1e-5, batch size to 8 for 20 epochs. Considering that this is a binary categorization problem, we have directly adopted the weighted F1 score as the evaluation criterion here. The p here is used to control different importance of Precision and Recall, as in different tasks we may want to avoid misdiagnosis as much as possible and vice versa.

$$pF_1 = 2 \frac{pPrecision \cdot pRecall}{pPrecision + pRecall}$$

where

$$pPrecision = \frac{pTP}{pTP + pFP}$$

$$pRecall = \frac{pTP}{TP + FN}$$

4.2. Result

In our experimentation, we systematically evaluated various backbone architectures as feature extraction schemes. Beyond the conventional ResNet, we explored the efficacy of its contemporary counterparts, specifically GoogleNet [14] and the subsequently introduced EfficientNet [13]. The results, detailed in the ensuing table, underscore the impact of the implemented enhancements, revealing that our approach yielded superior performance compared to individual models. Notably, our model outperformed them by a margin ranging from 0.5% to 2.8%, signifying a noteworthy advancement in accuracy and effectiveness. This nuanced exploration of diverse backbone architectures highlights the pivotal role of feature extraction in enhancing the overall performance of breast cancer detection models.

Table 1. cancer detection models

Models	pF1 score
ResNet50	0.501
GoogleNet50	0.492
EfficientNet B2	0.515
SE-ResNeXt50	0.520

5. Conclusion

In light of recent strides in deep learning model architectures, the widespread application of medical-assisted AI has significantly expanded. Acknowledging the existing limitations in accuracy within current models, the imperative for high-quality models in subsequent applications remains paramount. In this study, we conducted a meticulous analysis of mammogram data distribution, coupled with insightful data visualization. Subsequently, our exploration extended to assessing the detection performance using diverse deep learning models. Notably, through innovative enhancements such as the reevaluation of mean pooling and the adoption of a new optimizer, our model surpassed the performance of

other existing models. This achievement is underscored by a weighted F1 score reaching 0.52, indicative of the model's efficacy in balancing precision and recall. As our ongoing optimization endeavors unfold, we posit that the landscape of breast cancer detection holds great promise for the future. The continual refinement of our model stands as a testament to the potential for further advancements in accuracy and robustness, paving the way for increasingly effective medical AI applications in the imminent future.

References

- [1] Fear, Elise C., Paul M. Meaney, and Maria A. Stuchly. "Microwaves for breast cancer detection?." IEEE potentials 22.1 (2003): 12-18.
- [2] Moore, S. K. (2001). Better breast cancer detection. Ieee Spectrum, 38(5), 50-54.
- [3] Sutanto, D. H., & Ghani, M. K. A. (2015). A benchmark of classification framework for non-communicable disease prediction: a review. ARPN J Eng Appl Sci, 10, 9941-9955.
- [4] Wang Y, Yang K, Wan W, Mei H. Adaptive energy saving algorithms for Internet of Things devices integrating end and edge strategies. Trans Emerg Telecommun Technol. 2020..
- [5] Singh, B. K., Verma, K., & Thoke, A. S. (2016). Fuzzy cluster based neural network classifier for classifying breast tumors in ultrasound images. Expert Systems with Applications, 66, 114-123.
- [6] Chen, H., Gao, M., Zhang, Y., Liang, W., & Zou, X. (2019). Attention-based multi-NMF deep neural network with multimodality data for breast cancer prognosis model. BioMed research international, 2019.
- [7] Y. Wang, K. Yang, W. Wan, Y. Zhang and Q. Liu, "Energy-Efficient Data and Energy Integrated Management Strategy for IoT Devices Based on RF Energy Harvesting," in IEEE Internet of Things Journal, vol. 8, no. 17, pp. 13640-13651, 1 Sept.1, 2021, doi: 10.1109/JIOT.2021.3068040..
- [8] Hao Hu, Shulin Li, Jiaxin Huang, Bo Liu, and Change Che. Casting product image data for quality inspection with xception and data augmentation. Journal of Theory and Practice of Engineering Science, 3(10):42-46, 2023.
- [9] Che, C., Hu, H., Zhao, X., Li, S., & Lin, Q. (2023). Advancing Cancer Document Classification with Random Forest. Academic Journal of Science and Technology, 8(1), 278-280.
- [10] Hu H, Li S, Huang J, et al. Casting Product Image Data for Quality Inspection with Xception and Data Augmentation[J]. Journal of Theory and Practice of Engineering Science, 2023, 3(10): 42-46.
- [11] Che C, Liu B, Li S, et al. Deep Learning for Precise Robot Position Prediction in Logistics[J]. Journal of Theory and Practice of Engineering Science, 2023, 3(10): 36-41.
- [12] Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7132-7141).
- [13] Tianbo, Song, Hu Weijun, Cai Jiangfeng, Liu Weijia, Yuan Quan, and He Kun. "Bio-inspired Swarm Intelligence: a Flocking Project With Group Object Recognition." In 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE), pp. 834-837. IEEE, 2023.
- [14] Chang Che, Bo Liu, Shulin Li, Jiaxin Huang, and Hao Hu. Deep learning for precise robot position prediction in logistics. Journal of Theory and Practice of Engineering Science, 3(10):36-41, 2023.
- [15] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

- [16] Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.
- [17] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).
- [18] Lin, Q., Che, C., Hu, H., Zhao, X., & Li, S. (2023). A Comprehensive Study on Early Alzheimer's Disease Detection through Advanced Machine Learning Techniques on MRI Data. *Academic Journal of Science and Technology*, 8(1), 281–285.
- [19] S. Tianbo, H. Weijun, C. Jiangfeng, L. Weijia, Y. Quan and H. Kun, "Bio-inspired Swarm Intelligence: a Flocking Project With Group Object Recognition," 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 2023, pp. 834-837, doi: 10.1109/ICCECE58074.2023.10135464.