Meta-transfer learning based multi-classification algorithm for breast cancer

Kaiyi Wang, Lijia Zhi*, Shaomin Zhang, Chonghua Yu, Siqing Zhang
School of Computer Science and Engineering, North Minzu University, Yinchuan 750021, Ningxia, China

ABSTRACT

The worldwide incidence of breast cancer continues to increase annually. Accurate classification of breast cancer subtypes is crucial for improving treatment precision. However, the diagnosis of breast cancer is time-consuming and labor-intensive. Although the introduction of Computer-Aided Diagnosis (CAD) systems can assist doctors in diagnosing breast cancer more effectively, establishing high-quality medical image datasets is challenging, significantly hampering the development of CAD systems. This paper proposes a meta-learning method suitable for the multi-classification of breast cancer histopathological images. This method effectively addresses the issue of limited data by adopting the concept of transfer learning and incorporates group convolution within the model to enhance efficiency and feature extraction capability. Experimental results demonstrate that this method achieves a classification accuracy of 96.30% on the BreakHis dataset across eight categories, surpassing state-of-the-art methods. Furthermore, we validated the generalization performance of this method on another publicly available dataset, Kather-CRC-2016.

Keywords: Breast cancer, histopathology, meta learning, transfer learning, multi classification

1. INTRODUCTION

Breast cancer is among the most prevalent cancers impacting women's health worldwide, as reported by the World Health Organization. In 2020, the incidence of breast cancer surpassed that of lung cancer for the first time, making it the most prevalent cancer worldwide. Early diagnosis is crucial for the treatment of breast cancer. Moreover, breast cancer has many complex subtypes (such as ductal carcinoma, mucinous carcinoma, and lobular carcinoma), each requiring different treatment plans. Correctly distinguishing breast cancer subtypes is the foundation for accurate treatment. To reduce the cost of tumor diagnosis and decrease the occurrence of errors, researchers have introduced various Computer-Aided Diagnosis (CAD) systems for detecting breast cancer, which greatly assist doctors in diagnosing the disease[1].

Due to the unique nature of the medical image processing field, accumulating and annotating large amounts of medical data is time-consuming and labor-intensive, leading to the persistent challenge of limited data availability in this field. Faced with this challenge, this paper combines the idea of transfer learning, allowing the model to be pretrained on a large-scale dataset to obtain pretrained weights and achieve a good initialization. Subsequently, using meta-learning concepts, the pretrained model is continuously optimized with a limited medical image dataset, enabling faster convergence and stronger generalization. Additionally, we introduce group convolution operations within the model to enhance feature extraction capabilities and reduce memory usage. We demonstrate the utility of our model through the multi-classification of breast cancer subtypes on the BreakHis dataset and validate the generalization of our model on the Kather-CRC-2016 colorectal cancer histopathological image dataset.

2. METHODOLOGY

2.1 Meta transfer learning

This method includes two stages. First, the model is pretrained on a large-scale dataset to obtain pretrained weights. In this stage, a feature extractor F and a classifier C are randomly initialized, and then continuously optimized using a gradient descent algorithm. The feature extractor F remains frozen in the subsequent meta-training and meta-testing phases, while the classifier C is discarded, as the pretrained weights are transferred to other classification targets in the next step.

*ljzhi_paper@163.com; phone 15909609152

International Conference on Optics, Electronics, and Communication Engineering (OECE 2024), edited by Yang Yue, Proc. of SPIE Vol. 13395, 133951P · © 2024 SPIE · 0277-786X · Published under a Creative Commons Attribution CC-BY 3.0 License · doi: 10.1117/12.3048468

The second stage is the Meta-Transfer-Learning (MTL) phase. During this phase, the pretrained weights are continuously fine-tuned using the training data. The fine-tuning process follows the gradient descent algorithm of meta-learning:

$$\mathrm{C}' \leftarrow C - \beta\left(\frac{\partial L_{\mathrm{Tr}}([F;C],\xi S_{\{12\}})}{\partial C}\right) \tag{1}$$

Here, C' is a temporary classifier used for meta-training, L_{Tr} is the training loss, and ξS_1 and ξS_2 are our fine-tuning operation coefficients. In meta-training, ξS_1 is initialized to 1 and ξS_2 is initialized to 0. They are then optimized using the gradient descent algorithm based on the loss L_T of T, as follows:

$$\xi \mathrm{S}_{\mathrm{i}} =: \xi S_i - \gamma\left(\frac{\partial L_T([F;C'],\sigma_{S_{\{12\}}})}{\partial \sigma_{\xi S_i}}\right) \tag{2}$$

Our fine-tuning operation updates the transferred weights that were frozen during the pretraining stage. For the pretrained feature extractor F, its i-th layer contains k neurons, resulting in k pairs of parameters, represented as {(Wi, k, bi, k)}. Based on MTL, $\{\xi S_{\{1,2\}}\}$ learns K pairs of scalars. Suppose X is the input, $\{\xi S_{\{1,2\}}\}$ are applied to (W, b) as follows:

$$\mathrm{SS}(X;W,b;\xi S_{\{12\}}) = (W \odot \xi S_1)X + (b + \xi S_2) \tag{3}$$

Here, $\odot$ denotes element-wise multiplication.

2.2 More efficient model

Due to the extensive parameter optimization required by meta-learning itself, this would make the model occupy a higher footprint and operate slower. Inspired by group convolutions, we applied group convolutions to all residual blocks of the model, proposing a new model with improved feature extraction capability and higher efficiency, ResNet-12G, as illustrated in Figure 1. Furthermore, through multiple sets of experiments, it has been demonstrated that the powerful feature extraction capability of this model also contributes to extracting features adequately in scenarios with minimal task loads. This, coupled with fine-tuning operations, ensures faster convergence speed and better training outcomes for the model.

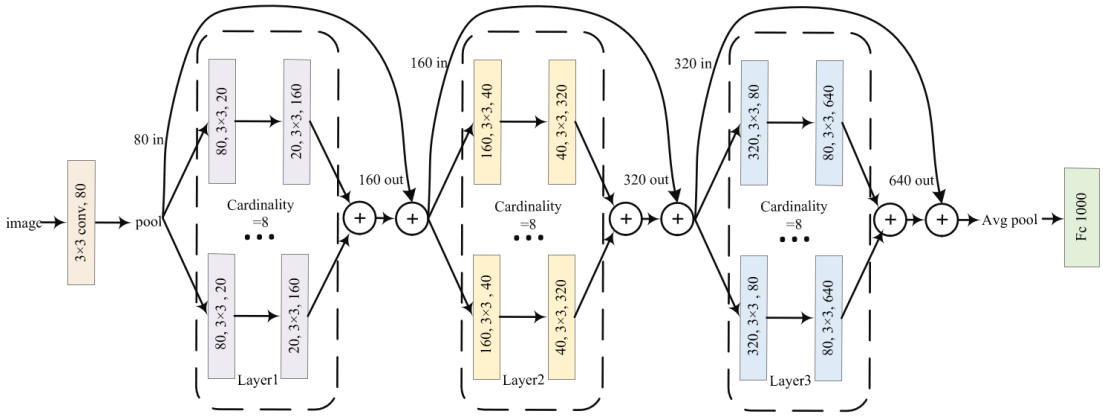

Figure 1. The architecture of our proposed ResNet-12G model.

3. DATASET

This study primarily evaluates the performance of our proposed model through multi-classification results on the BreakHis dataset and demonstrates its effectiveness by comparing it with the latest research outcomes on this dataset. Additionally, the generalization of our proposed model is validated through multi-classification experiments on another publicly available medical dataset.

BreakHis: The BreakHis dataset is a publicly available breast pathology dataset released by Spanhol et al.[2], consisting of 9,109 tumor tissue microscopic images from 82 breast cancer patients, with magnifications of 40×, 100×, 200×, and 400×, as shown in Figure 2. This dataset is divided into two classes: benign and malignant, with each class further subdivided

into four subtypes, resulting in eight categories. The four benign subtypes are A, F, PT, and TA, while the four malignant subtypes are DC, LC, MC, and PC. We perform an eight-subtype multi-classification task on this dataset.

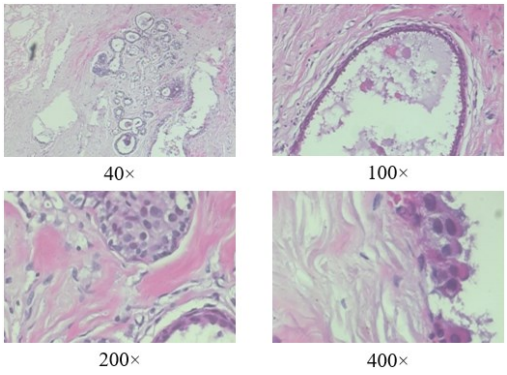

Figure 2. The relationship between the magnification factors and images in the BreakHis dataset.

Kather-CRC-2016 dataset: This dataset was created by Kather et al.[3]. using 10 anonymized H&E stained colorectal cancer (CRC) tissue slides from the archives of the Medical Center of the University of Mannheim, Germany. It contains 5,000 images of CRC tissue types, with 625 samples for each tissue type. Each sample is a non-overlapping tissue patch of 150×150 pixels. The dataset is divided into eight categories: "Tumor", "Stroma", "Complex", "Lympho", "Debris", "Mucosa", "Adipose", and "Empty". To validate the generalization of our model, we perform an eight-class classification task on this dataset.

4. RESULT

4.1 Evaluation metrics

To evaluate the performance of the classification model, we use the following four evaluation metrics: Accuracy (Acc), Precision (Pre), Sensitivity (Sen), and Specificity (Spec). Their formulas are as follows:

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$Pre = \frac{TP}{TP + FP} \tag{5}$$

$$Sen = Rec = \frac{TP}{TP + FN} \tag{6}$$

$$Spec = \frac{TN}{TN + FP} \tag{7}$$

TP represents the number of correctly predicted positive samples, known as true positives. FN denotes the number of samples predicted as negative but are actually positive, referred to as false negatives. FP represents the number of samples predicted as positive but are actually negative, known as false positives. TN stands for the number of correctly predicted negative samples, referred to as true negatives.

4.2 Experimental results

Our eight-class classification results on the BreakHis dataset are shown in Table 1. As can be seen from the table, our method achieved the best Acc at 100× magnification. In contrast, when the magnification reached 400×, the results declined significantly. This decline is due to the fact that certain features of the original images are no longer fully visible at this magnification, as illustrated in Figure 2. Moreover, some features unrelated to cancer cells start to dominate. Consequently, the features extracted by the model become more superficial, affecting the model's performance. Considering this impact, we discuss the ablation experiments using only images at 40× magnification.

Table 1. Multi-classification results at different magnification factors on the BreakHis dataset, with the best results highlighted in bold.

Magnifications	Acc (%)	Pre (%)	Sen (%)	Spec (%)
40×	95.90±0.65	98.01±0.65	97.43±0.80	97.82±0.72
100×	**96.00±0.67**	98.01±0.69	**97.47±1.00**	97.81±0.78
200×	95.65±0.73	**98.24±0.69**	96.80±0.98	**98.12±0.77**
400×	94.48±0.73	96.61±0.82	96.44±0.86	96.20±0.96

4.3 Comparison with other recent results on this dataset

We compare the results obtained by our model with the latest published research results on the BreakHis dataset. The specific details are shown in Table 2. As can be seen from the comparison, our method overall outperforms the existing research methods.

Table 2. We compare our multi-class Acc on the BreakHis dataset with the multi-class Acc of other researchers, with the best results highlighted in bold.

Method	400× (%)	100× (%)	200× (%)	400× (%)
BHCNet-6+ERF[4]	93.74	93.81	92.22	90.66
TL+Resnet-18[5]	94.49	93.27	91.29	89.56
PGLCM+SARF[6]	93.88	93.97	94.57	**94.77**
BAM[7]	93.36(average)			
STMLAN[8]	90.66(average)			
Ours	**95.90±0.65**	**96.00±0.67**	**95.65±0.73**	94.48±0.73

4.4 Ablation study

We conducted ablation experiments on images at 40× magnification to verify the effectiveness of the modules we used, including the use of pretrained weights and the use of group convolution networks. The ablation results are shown in Table 3.

Table 3. Ablation study results of our method at 40× magnification, with the best results highlighted in bold.

	Transfer	Group	Acc (%)
A	×	×	81.12±1.26
B	×	√	83.98±1.16
C	√	×	94.55±0.76
ours	√	√	**95.90±0.65**

4.5 Generalization verification

The comparison of the results obtained by our method on the Kather-CRC-2016 dataset with the latest methods is shown in Table 4. Our method not only performs excellently on the breast tumor dataset but also shows satisfactory results when applied to tumor datasets from other body parts. This demonstrates that the generalization capability of our method is not limited to a single type of histopathological classification but can be extended to the classification of pathological tissues from other parts of the human body. This will help promote the early clinical application of our method, aiding radiologists in making better diagnoses.

Table 4. Comparison of our approach and state-of-the-art methods on the SIPAKMED dataset, with the best results highlighted in bold.

Method	**Acc (%)**
Ensemble-DNN[9]	92.83
ResNet-50[10]	94.86
CRCCN-Net[11]	93.50
ResNet50+FMMN[12]	96.70
Ours	**96.72±0.58**

5. CONCLUSIONS

In this study, we combined classical transfer learning and meta-learning methods to propose a new approach suitable for multi-classification on datasets with limited data. Extensive multi-classification experiments on the publicly available BreakHis dataset demonstrated the effectiveness of our method, and ablation studies evaluated the usefulness of our proposed modules in improving model stability and efficiency. Additionally, we conducted multi-classification experiments on other publicly available datasets to validate the generalization performance of our method. We hope that this research will accelerate the development of medical CAD systems and promote progress in this field.

ACKNOWLEDGMENTS

This work was partially supported by the Ningxia Natural Science Foundation, China (Grant Nos. 2023AAC03263 and 2021AAC03198).

REFERENCES

[1] Giger, M. L. and Suzuki, K., "Computer-aided diagnosis," [Biomedical Information Technology], Academic Press, New York, 359-XXXII (2008).
[2] Spanhol, F. A., Oliveira, L. S., Petitjean, C. and Heutte, L., "A dataset for breast cancer histopathological image classification," IEEE Trans. Biomed. Eng. 63(7), 1455-1462 (2015).
[3] Kather, J. N., Weis, C. A., Bianconi, F., Melchers, S. M., Schad, L. R., Gaiser, T., Marx, A. and Zöllner, F. G., "Multi-class texture analysis in colorectal cancer histology," Sci. Rep. 6, 27988 (2016).
[4] Jiang, Y., Chen, L., Zhang, H. and Xiao, X., "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module," PLoS ONE 14(3), e0214587 (2019).
[5] Boumaraf, S., et al., "A new transfer learning-based approach to magnification dependent and independent classification of breast cancer in histopathological images," Biomed. Signal Process. Control 63, 102192 (2021).
[6] Li, J., Shi, J., Chen, J., et al., "Self-attention random forest for breast cancer image classification," Front. Oncol. 13, 1043463 (2023).
[7] Alirezazadeh, P. and Dornaika, F., "Boosted additive angular margin loss for breast cancer diagnosis from histopathological images," Comput. Biol. Med. 166, 107528 (2023).
[8] Lee, J. S. and Wu, W. K., "Breast tumor tissue image classification using single-task meta learning with auxiliary network," Cancers 16(7), 1362 (2024).
[9] Ghosh, S., Bandyopadhyay, A., Sahay, S., et al., "Colorectal histology tumor detection using ensemble deep neural network," Eng. Appl. Artif. Intell. 100, 104202 (2021).
[10] Tsai, M.-J., and Tao, Y. H., "Deep learning techniques for the classification of colorectal cancer tissue," Electronics 10, 1662 (2021).
[11] Kumar, A., Vishwakarma, A. and Bajaj, V., "Crccn-net: Automated framework for classification of colorectal tissue using histopathological images," Biomed. Signal Process. Control 79, 104172 (2023).
[12] Sharkas, M. and Attallah, O. "Color-CADx: a deep learning approach for colorectal cancer classification through triple convolutional neural networks and discrete cosine transform," Scientific Reports 14(1), 6914 (2024).