Proceedings of CONF-CIAP 2026 Symposium: International Conference on Atomic Magnetometer and Applications
DOI: 10.54254/2753-8818/2026.HZ31437

A Dual-Attention Meta-Learning Approach for Fine-Grained Classification of Gastric Cancer Histopathological Images with Limited Samples

Tong Wang

Chongqing University of Arts and Sciences, Chongqing, China
wangtong2023@163.com

Abstract. The pathological diagnosis of gastric cancer relies on the precise differentiation of its fine-grained subtypes, a process often challenged by the lack of annotated data. This study proposes a novel dual-attention meta-learning framework, aiming to address the problem of fine-grained image classification under fee-shot conditions in the medical field, and is particularly suitable for histopathological images of gastric cancer. This study proposes a meta-learning method combined with a dual-attention mechanism to enhance the recognition of imperceptible inter-class differences in pathological images. The channel attention module automatically selects key features by analyzing the information density of the feature map, while the spatial attention module precisely locates the morphological structure in the image. Supported by the meta-learning framework, this model can quickly adapt to new disease classification tasks with a limited number of samples. Experimental results on a self-constructed pathological dataset containing multiple fine-grained subtypes of gastric cancer show that the proposed mean is significantly superior to several benchmark models, demonstrating obvious performance advantages. This work provides a new technical approach for developing computer-aided diagnostic systems in the absence of annotated data.

Keywords: Fine-grained image classification, Few-shot learning, Meta-learning, Attention mechanism.

1. Introduction

Gastric cancer is one of the most common and fatal malignant tumors worldwide. Gastric cancer is the fifth most common cancer worldwide, becoming a serious global health problem due to its high mortality rate (accounting for 18.0% of all cancer deaths) and poor prognosis [1]. Driven by technological breakthroughs, artificial intelligence (AI) has injected new impetus into the progress of gastric cancer research and treatment. In key scenarios that rely on image interpretation, such as endoscopy, pathological diagnosis, and radiomics analysis, artificial intelligence technology has demonstrated its advantages of high efficiency and accuracy, and has great potential for the future [2,3]. The role of artificial intelligence technology in the field of endoscopy has gone beyond the detection of gastric cancer. It can also achieve the automatic classification of endoscopic images. A

DOI: 10.54254/2753-8818/2026.HZ31437

study conducted confirmed this [4,5], in which they employed the CNN model AlexNet to carry out classification research on upper gastrointestinal organ images. This study used the conventional endoscopic examination images of patients with gastric cancer as the data basis. Ultimately, AlexNet successfully classified these images into 14 categories, covering organs such as the stomach, esophagus, and duodenum in the white light mode.

In terms of performance, the classification accuracy of this model on the training set and the validation set reached 99.3% and 96.5% respectively. This not only effectively reduces the workload of doctors in processing image data, but also provides strong support for improving clinical efficiency. According to the Global Cancer statistics of 2020, many people in China died of stomach cancer [6]. The incidence rate and the number of deaths from gastric cancer in China account for approximately 50% of the global total, and the disease burden is heavy [7]. The exact cause of gastric cancer is not yet fully understood. Current research generally holds that its occurrence is the result of the complex interaction of multiple factors, including poor dietary habits, Helicobacter pylori infection, socio-economic factors, specific lifestyles, and genetic susceptibility [8,9]. The accurate classification of different histological subtypes of gastric cancer is crucial for determining the prognosis of patients and guiding treatment plans. Pathologists observe tissue sections under a microscope and distinguish fine-grained categories such as well-differentiated adenocarcinoma, poorly differentiated adenocarcinoma and sigring cell carcinoma based on subtle morphological differences. However, this diagnostic process is time-consuming, labor-intensive, and subject to a certain degree of variability among observers.

In recent years, deep learning technology has achieved remarkable success in numerous medical image analysis tasks, including conventional image classification. However, applying deep learning models to the fine-grained classification of gastric cancer faces two major challenges.Firstly, fine-grained classification requires distinguishing highly similar subcategories, where discriminative features usually exist in small local areas rather than the overall appearance of the image. Secondly, collecting a large number of high-quality histopathological images and performing pixel-level annotations for each rare cancer subtype is difficult and expensive in practice, leading to the problem of data scarcity.

How to solve these issues? This study proposes a dual-attention meta-learning framework. This framework trains the model on a large number of learning tasks, enabling it to quickly generalize to new tasks with the least amount of data. Meanwhile, the attention mechanism has been proven to effectively enhance the feature representation of fine-grained tasks by focusing the model on the information-rich regions in the image. This study hypothesizes that combining these two methods can synergistically enhance the model's ability to learn and identify subtle differences from a small number of samples.

The main contributions of this research include three aspects. Firstly, it proposes a new framework that seamlessly integrates the dual-attention module into the meta-learning process. Secondly, the proposed method is specifically designed and verified for the key and challenging applications of fine-grained gastric cancer classification with limited data.

Finally, the comprehensive experiments show that this framework outperforms the existing meta-learning and fine-grained classification baseline methods in terms of classification accuracy.

2. Related work

The authors are strongly encouraged to use this document for the preparation of your paper. Please use the .docx format (all versions after Word 2007 for PC and Word 2011 for Mac) and do not add any text to the headers (do not set running heads) nor to the footers.

DOI: 10.54254/2753-8818/2026.HZ31437

The purpose of fine-grained image classification is to distinguish different subcategories within the same basic category, such as different bird species or different car models. Early methods usually relied on strongly supervised information such as bounding boxes or partial annotations to locate the discriminative region, but the acquisition cost of these annotations was very high. Recent weakly supervised methods use deep convolutional neural networks to automatically locate key features. For instance, bilinear CNN models capture local feature interactions by calculating the external product of different feature channels. The attention-based mechanism guides the network to focus on information-rich areas without additional supervision, which is particularly applicable to medical images where pathological features are often concentrated. In the field of histopathological image classification, existing research has revealed the superiority of the ResNet series of models. For instance, Warin et al. achieved high-precision classification of oral pathology by integrating DenseNet and ResNet. Azour et al. pointed out that lightweight models like ResNet are more suitable for scenarios with limited data. The common point of these studies is that they verified the effectiveness of the ResNet architecture. Therefore, this study adopts ResNet as the core feature extractor [10].

3. Methodology

A new dual-attention element learning framework was designed to solve the problem of few-shot and fine-grained classification of gastric cancer histopathological images. The core idea of this framework is to combine the rapid adaptation ability of meta-learning with the refined feature extraction ability of the attention mechanism.

3.1. Problem definition

The relevant problem is usually described as a classification task. These tasks are presented in the form of "task sets", and each task set consists of two core datasets. One type is the support set, which contains N different categories, and each category only contains K labeled samples. The other one is the query set; Its samples also come from these N classes, but the specific samples are completely different from those in the support set and there is no overlap. The goal that the model needs to achieve is to accurately determine the class to which each sample in the query set belongs by using the limited markup information provided by the support set. In this study, the values of N and K are usually small (for example, N=5, K=1 or 5) to simulate scenarios where data is scarce.

3.2. Meta-learning backbone: MAML framework

This study adopts Model-Agnostic Meta-Learning (MAML) as the foundational learing about meta framework. The core objective of MAML is that how to discover mathematic model θ. This ensures that for any new task sampled from the task distribution p(T), the model can achieve optimal performance after only one or a few gradient descent updates. The pattern training process consists of two parts: an interior loop and an outside loop. Inner loop: It performs rapid parameter adaptation for each specific task. Outer loop: It optimizes the model's overall performance across all tasks and updates the initial parameters θ based on this optimization.

3.3. Dual-attention feature embedding part of model

To enhance the model's recognition ability in fine-grained classification, this study introduces a dual-attention module into the feature extractor. This part simulates two orders of double attention:

DOI: 10.54254/2753-8818/2026.HZ31437

channel attention and spatial attention are composed of two orders of double attention.

Module on channel attention: This module refers to "which functions are more important". It compresses the global spatial information, generates the weight of each feature channel, emphasizes the information channels, and suppresses the less important ones. This module first performs global mean pooling and global Max pooling on the feature maps respectively, and then generates the channel attention weight map through the shared multi-layer perceptron (MLP). Dimension Focus module: This module focuses on "which positions in the feature map are more important". It applies pooling operations along the channel dimension to generate a spatial attention weight map, thereby highlighting the key spatial positions in the image. The channel attention and spatial attention modules are connected in sequence, enabling the model to make coordinated use of channel and spatial information.

3.4. Overall framework and training process

The training process of the proposed framework is shown in Figure 1. Firstly, the feature extractor (for example, CNN backbone network) works in conjunction with the dual-attention module to map the input image into a high-dimensional feature representation. Then, within the inner loop of each episode, the model uses the support set samples to calculate the prototype of each class. Next, the features of the query set samples are classified by calculating the distance from the query set samples to each class prototype (for example, the Euclidean distance). The loss function is usually constructed based on the classification accuracy of the query set samples. In the outer loop, the meta-optimizer updates the initial parameter θ of the model based on the cumulative loss of all tasks. Through this mechanism, the model has developed the ability to be rapidly extended to new tasks that require fine-grained differentiation.

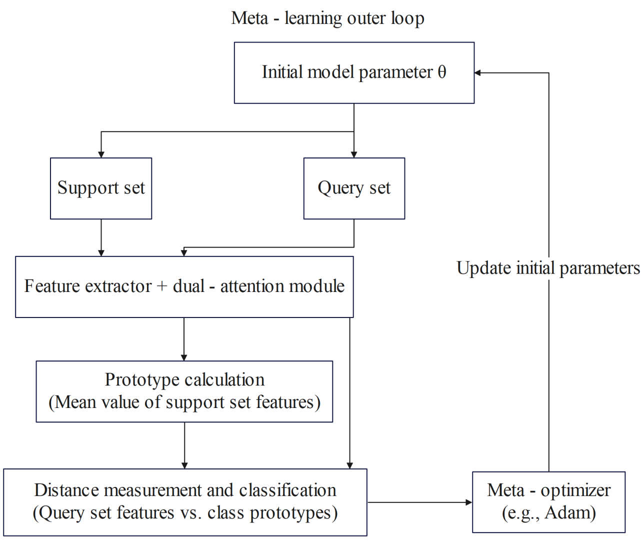

Figure 1. Flowchart (picture credit: original)

4. Experiments and results

4.1. Dataset and experimental setup

Due to the lack of a public standard dataset suitable for the classification of a few fine-grained gastric cancers, this study collected and organized histopathological image datasets containing the

DOI: 10.54254/2753-8818/2026.HZ31437

three major subtypes of gastric cancer (well-differentiated adenocarcinoma, poorly differentiated adenocarcinoma, and signet ring cell carcinoma), as shown in Figures 2, 3, and 4 respectively. All images were preprocessed to a uniform size and data augmentation was applied to increase sample diversity.

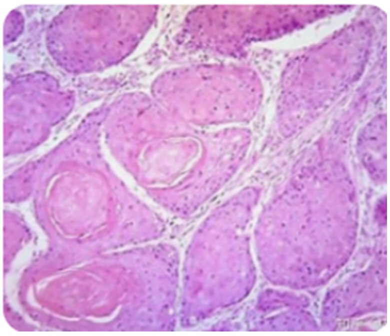

Figure 2. Well-differentiated adenocarcinoma (picture credit: original)

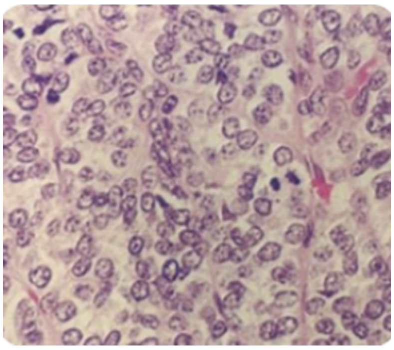

Figure 3. Poorly differentiated adenocarcinoma (picture credit: original)

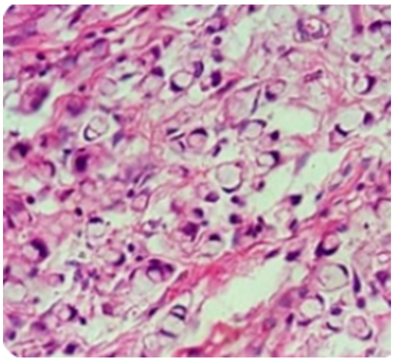

Figure 4. Signet ring cell carcinoma (picture credit: original)

The experiment adopted the scenario training method. During the training phase, a large number of N-way K-shot tasks are randomly selected from the dataset. During the testing phase, test tasks are constructed by sampling from the hold-out categories in the dataset to evaluate the generalization performance of the model. This study compared this method with classic few-shot learning algorithms (such as Prototypical Networks, Relation Networks and baseline MAML).

DOI: 10.54254/2753-8818/2026.HZ31437

4.2. Results and analysis

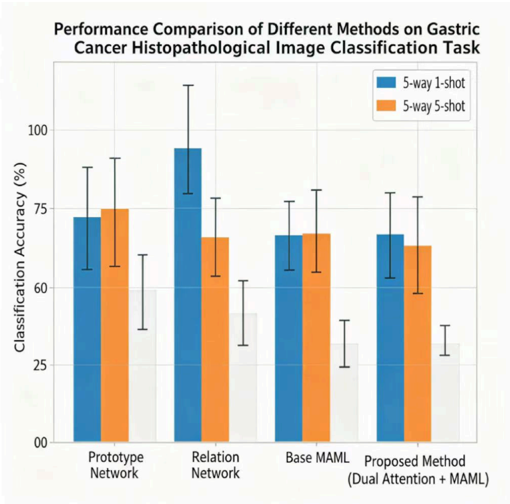

Figure 5. Performance comparison of different methods on the gastric cancer histopathological image classification task (picture credit: original)

The experimental results (Figure 5) show that this method achieves the best classification performance in both Settings. Compared with the baseline MAML, the performance improvement is mainly attributed to the dual-attention module, which enhances the model's ability to capture subtle differentiating features, enabling it to make more accurate judgments with limited data. Furthermore, the performance of all methods under the 5-shot setting is superior to that under the 1-shot setting, which is expected because more support samples provide richer class information.

4.3. Ablation study

To verify the effectiveness of each component in the dual-attention module, an ablation study was conducted. The results are shown in Figure 6. Removing either the channel attention or spatial attention module will lead to a decline in model performance, which indicates the necessity of the two modules and their complementary effects.

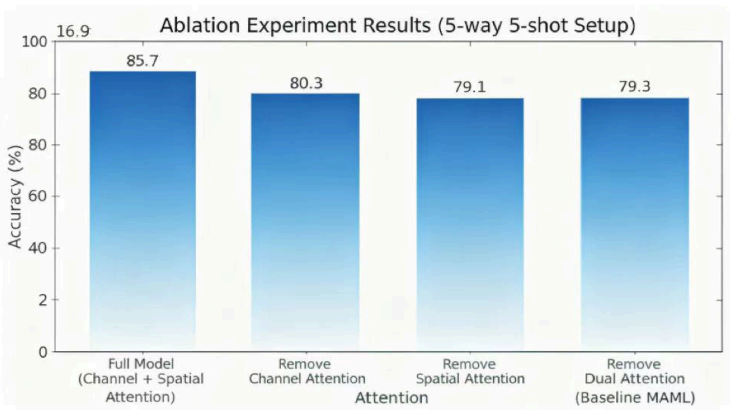

Figure 6. Ablation study results (picture credit: original)

DOI: 10.54254/2753-8818/2026.HZ31437

5. Conclusion

This study proposes a new framework that combines the dual-attention mechanism and meta-learning to address the scarcity of annotation data in the fine-grained classification of gastric cancer histopathological images. This framework acquires the ability to quickly adapt to new tasks through meta-learning and focuses on the key feature channels and spatial regions related to diagnosis through a dual-attention mechanism, thereby effectively improving the classification accuracy under the condition of few shots. Experiments on the self-built dataset of gastric cancer histopathological images verified the effectiveness of this method and its advantages over the baseline model.

The limitation of this study lies in the relatively limited scale of the dataset and the current focus on only three major subtypes. Future work will focus on expanding the dataset to include more rare subtypes to further verify the generalization of this method. In addition, exploring more effective feature fusion methods and introducing the Transformer architecture into the meta-learning pipeline are also directions worthy of research. This study provides a useful reference for the development of intelligent auxiliary diagnostic tools for rare diseases and rare subtypes.

References

[1] Sun, W., Shi, H., & Huang, Z. (2024). Value of convolutional neural network model based on gastric histopathological image dataset for early prediction of gastric cancer. Journal of North Sichuan Medical College, 39(7), 877–881. https: //doi.org/CNKI: SUN: NOTH.0.2024-07-003

[2] Elemento, O., Leslie, C., Lundin, J., & Tourassi, G. (2021). Artificial intelligence in cancer research, diagnosis and therapy. Nature Reviews Cancer, 21(12).

[3] Hirasawa, T., Aoyama, K., Tanimoto, T., Ishihara, S., Shichijo, S., Ozawa, T., Ohnishi, T., Fujishiro, M., Matsuo, K., Fujisaki, J., & Tada, T. (2018). Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. Gastric Cancer, 21(4).

[4] Cho, B.-J., Bang, C. S., Park, S. W., Yang, Y. J., Seo, S. I., Lim, H., Shin, W. G., Hong, J. T., Yoo, Y. T., Hong, S. H., Choi, J. H., Lee, J. J., & Baik, G. H. (2019). Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. Endoscopy, 51(12).

[5] Zhang, J., Shen, L., Hu, S., Ding, Q., Mu, G., Yin, A., Huang, X., Liu, J., Jiang, X., Wang, Z., Deng, Y., Liu, M., Lin, R., Ling, T., Li, P., Wu, Q., Jin, P., Chen, J., & Yu, H. (2019). A deep neural network improves endoscopic detection of early gastric cancer without blind spots. Endoscopy, 51(6).

[6] Shen, Y., Liu, H., Wang, S., Zhang, X., Hu, X., Lou, Y., Liu, Z., Xie, X., & Zhao, Y. (2025). Comparative study on predicting gastric precancerous status based on machine learning methods. Chinese Archives of Traditional Chinese Medicine, 43(6).

[7] Wang, X. (2021). Epidemiological study on Helicobacter pylori infection in the whole age group population in high incidence area of gastric cancer (Doctoral dissertation, Lanzhou University).

[8] Collatuzzo, G., Pelucchi, C., Negri, E., López-Carrillo, L., Tsugane, S., Hidaka, A., Hamada, G. S., Hernández-Ramírez, R. U., López-Cervantes, M., Malekzadeh, R., Pourfarzi, F., Mu, L., Zhang, Z.-F., Lunet, N., La Vecchia, C., & Boffetta, P. (2021). Exploring the interactions between Helicobacter pylori infection and other risk factors of gastric cancer: A pooled analysis in the Stomach Cancer Pooling (StoP) Project. International Journal of Cancer, 149(6).

[9] Sitarz, R., Skierucha, M., Mielko, J., Offerhaus, G. J. A., Maciejewski, R., & Polkowski, W. P. (2018). Gastric cancer: epidemiology, prevention, classification, and treatment. Cancer management and research, 239-248.

[10] Zhang, W., Jia, Z., & Li, Q. (2024). Gastric cancer pathological image diagnosis system based on ResNet and UNet. Journal of Hebei University (Natural Science Edition), 44(4), 433–440.