EasyChair Preprint

№ 9689

MiniDigPath: a New Standard for Pathology Images Few-Shot Learning Classification

Yan-Dong Du, Lin Feng and Xin-Lei Liu

February 9, 2023

MiniDigPath: A New Standard For Pathology Images Few-Shot Learning Classification

** [1], **[1], **[1], **[1], **[1]

**[2]

[1] **

[2] **

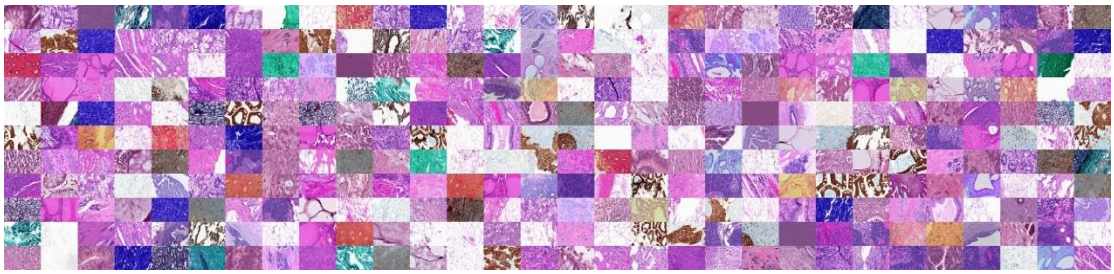

Fig.1 miniDigPath Dataset

Abstract: The miniDigPath dataset, which composed by 6 public pathology image datasets, is proposed by our work for few-shot learning (FSL). It consists 67 different diseases and tissue types, and every type has 48-500 tissue image blocks. In total, there are 21165 histopathology images. Importantly, miniDigPath is available publicly for every researcher. It explores a new idea to solve pathological images annotation using FSL, which is the importance and originality of the dataset we proposed. Experimental evaluation on the classical FSL algorithm and our method shows that the miniDigPath dataset can accomplish the task of FSL. Besides, FSL has good advantage for classification of digital pathology images.

Keywords: miniDigPath; few-shot learning; digital pathology images

1 Introduction

Histopathological analysis of tissue sections is the gold standard for assessing the presence and understanding the nature of many complex diseases, such as tumors [1], but the work of pathological analysis consumes a great deal of time and effort on the part of pathologists. With the development of computer technology, the advent of digital pathology is changing the way pathologists work and collaborate, and opening a new era for computational pathology.

In recent years, deep learning has achieved remarkable success in the digital pathology research and application[2,3,4]. A large number of annotated pathological images usually are required by deep learning methods for self-tune. However, the job of annotating images expends a lot of energy and time. Therefore, how to use a few amount of labeled data for modeling pathology image classification has

become a challenge problem. Recently, Few-Shot Learning (FSL) [5] has been proposed to solve this problem. Applying a priori knowledge，FSL can be rapidly generalized to new tasks that include only a few samples with supervised information. FSL has achieved numerous research results in image recognition since this theory was proposed[6,7,8]. However, research and application exploration of FSL methods in pathology are hampered due to the lack of the FSL digital pathology datasets.

To this end, as shown in Fig.1, we present the miniDigPath publicly dataset. By design, the miniDigPath mainly has the following four contributions:

Standards: It provides the standard and reference for the construction of future FSL pathology datasets.

Evaluation: It can be complete the performance evaluation and promotion of FSL algorithm.

Exploration: It is explore that new idea to solve the problem of annotating pathology data by using the FSL algorithm. This new idea further reduces the complexity of pathologists in performing pathology analysis.

Research: All of its data comes from publicly available datasets with Creative Commons (CC) licenses or related free licenses to facilitate research and educational purposes.

The remaining part of the paper proceeds as follows: In chapter 2th, we begin by laying out the methodology of the dataset construction and show the detailed build process. In chapter 3th, we list the current state-of-the-art models for FSL with concrete examples. In chapter 4th, we analyze the results of experimental with FSL model, and finally we conclude the whole paper.

2 The miniDigPath dataset

In this section, we describe how to establish the miniDigPath dataset. The creation of the dataset is made of three parts: 1) Collection of the source datasets, 2) Data sampling and 3) Renumbering of the sampled data, as shown in Figure 2. In the first step, we collected 6 pre-processed datasets as source datasets, (e.g., ET, BACH, KIMIA_path24, KIMIA_path960, NCT-CRC-HE-100K and CCGC). All source datasets have Creative Commons (CC) licenses or related free licenses, which allows us to develop derived datasets based on them. In the second step, due to the different amount of data in different source datasets, the principles of random or complete sampling were applied to the source datasets with sufficient or deficient data. In the third step, the sampled images are first uniformly converted to JPG, and then the images in JPG format are renumbered as n001-n067 according to the class sampling order.

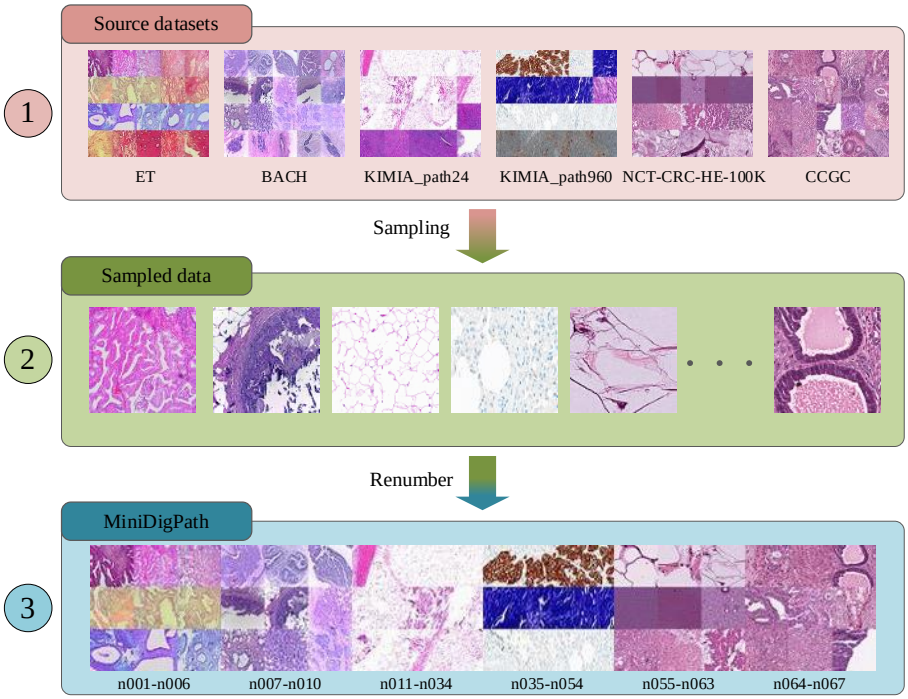

Fig.2 Building Process

The details of miniDigPath are shown in Table 1. The **n001-n006** of miniDigPath is based on ET [9], a pathology image dataset of endometrial diseases which contains normal, endometrial polyps, endometrial hyperplasia and endometrial adenocarcinoma, etc. We select 2700 source images in 6 classes to use in sampling work. The **n007-n010** of miniDigPath is based on a prior study [10] for classifying and localizing from breast cancer pathology image dataset, which provides a dataset (BACH) of two parts, microscopic images and whole-slides. The microscopic images consisted in 400 hematoxylin & eosin stained histological images into four classes: normal, benign, carcinoma in situ and invasive cancer. The whole-slides consisted in pixel-wise labeling of histology images as the same four classes. We use all microscopic images as sampled data. The **n011-n034** of miniDigPath is based on KIMIA_path24 [11], a digital pathology dataset of 24 full-scan images from different textured tissues of the human body. We randomly sample 10605 source images from 24 classes. The **n035-n054** of miniDigPath is based on KIMIA_path960 [12]. It is consisted of 960 pathological images belonging to 20 different tissue types. Like the sampling process of the n007-n010, we sample all source images in 20 types. The **n055-n063** of miniDigPath is based on NCT-CRC-HE-100K [13]. It is used to predicting survival of colorectal cancer histology slides, which included of 100,000 no-overlapping image blocks stained by hematoxylin & eosin. The **n064-n067** of miniDigPath is based on a prior study for classification from gastrointestinal cancer pathology images, which used a dataset (CCGC) of tens of thousands of pathology images belonging to four classes. The sampling work for NCT-CRC-HE-100K and CCGC, we select 500 images in every class to ensure the miniDigPath have sufficient task data.

Table 1 Sampling breakdown of miniDigPath dataset

Name	Source	Serial number	Quantity per class
ET	[9]	n001-n006	200-500
BACH	[10]	n007-n010	100
KIMIA_path24	[11]	n011-n034	78-500
KIMIA_path960	[12]	n035-n054	48
NCT-CRC-HE-100K	[13]	n055-n063	500
CCGC	[14]	n064-n067	500

3 Few-shot learning methods

Few-shot learning (FSL) is proposed in response to that the deep learning has the poor adaptation ability in the face of new classes. Research on FSL is currently divided into three general directions as follows:

Metric-based FSL approach, which focuses on how to use the comparison of relationship score between samples for classification, with common models such as Graph Neural Networks (GNN) [15], Prototypical Networks (PN) [16], and Relational Networks (RN) [6],In recent years, the proposal of (TPN) [20] has also improved the performance of metric learning.

Meta-learning based FSL is to use the meta-knowledge obtained during training to learn how to quickly adapt meta-learning model to different types of learning tasks. Common models are Model-agnostic Meta-Learning (MAML) [7], Memory- Augmented Neural Networks (MAMNs) [17];

Knowledge transfer-based FSL approach, the main idea of this method is that the FSL model is first pre-trained by a base class with sufficient amount of data, and then fine-tuned in novel class with a few samples.Such as Baseline++[18] and SSL[21].

On the other hand, some researchers focus on task enhancement during the FSL algorithm. Based on the three mainstream directions mentioned above, the method of task enhancement can enhance the classification accuracy of FSL algorithms. The common models are ATA [22] and STATUP [23].

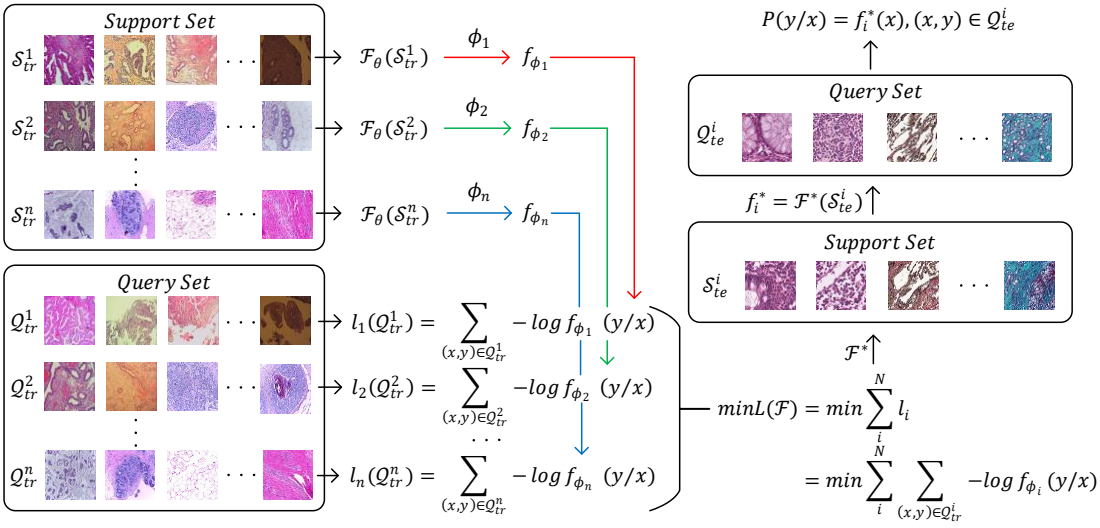

Fig.3 Few-shot learning model

In terms of training, the training and testing process of FSL model can be seen as the ability of human which quickly adapt and learn new tasks after acquiring some basic skills. For example, in the infant stage, human can quickly classify objects by learning a small number of photos. Meanwhile, FSL is trained and tested with a task as the basic unit, and every task has training and testing set, called the support and query set. Take supervised learning in FSL as an example, as shown in Fig.3, considers a task set $\mathcal{T} = \{\mathcal{T}^{\mathrm{tr}}, \mathcal{T}^{\mathrm{te}}\}$ that satisfies distribution $P(\mathcal{T})$, the goal of FSL is to make the distribution $P(\mathcal{T})$ be adapted by the few-shot model through two stages of training. In the first stage, a task-level classification function $\mathcal{F}^*$ is obtained by FSL through training on $\mathcal{T}^{\mathrm{tr}}$; and in the second stage, FSL learns to discover a sample-level classification functions f^* that can complete the classification of the testing task query set on $\mathcal{T}^{\mathrm{te}}$. The details are as follows:

Training phase: First, FSL defines a classification function $\mathcal{F}$ and the loss function L, and it randomly samples a few-shot training task $\mathcal{T}_i \in \mathcal{T}^{tr}$ on $P(\mathcal{T})$. Then, for the training task support set $\mathcal{S}_{tr}^i$ in $\mathcal{T}_{\mathrm{i}}$, the function $\mathcal{F}$ is trained to learn a sample-level classification function $f_i = \mathcal{F}(\mathcal{S}_{tr}^i)$. Using the training task query set $\mathcal{Q}_{tr}^i$, the loss l_{i} of f_{i} is computed, and the total loss $L(\mathcal{F}) = \sum_i^{|N|} l_i$ is calculated by the training tasks set which sampled N times on $P(\mathcal{T})$. Finally, the loss $L(\mathcal{F})$ is minimized to generates a task-level classification function $\mathcal{F}^* = \arg\min_{\mathcal{F}} L(\mathcal{F})$ by optimizers such as Adam, SGD and SVGD.

Testing phase: As same as the training phase, FSL first randomly chooses a few-shot testing task $\mathcal{T}_{\mathrm{j}} \in \mathcal{T}^{\mathrm{te}}$ on $P(\mathcal{T})$, and the function $\mathcal{F}^*$ obtained in the training phase is trained to learn a sample-level classification function $f_{\mathrm{i}}^* = \mathcal{F}^*(\mathcal{S}_{te}^i)$ on the testing task support set $\mathcal{S}_{te}^i$ in $\mathcal{T}_{\mathrm{j}}$. Then, in order to understand the performance of the FSL model, the function f_{i}^* is evaluated on the testing task query set $\mathcal{Q}_{te}^i$. In particular, during training $\mathcal{F}^*$, the training task set can also be split into a training task set and a validation task set.

In terms of model architecture, FSL is generally divided logically into a feature extraction network part and a classifier part. The feature extraction network part usually consists of convolutional neural networks, which maps the data into a high-dimensional vector space and highly abstracts the data information. The classifier part usually involves a fully connected layer with a nonlinear activation functions, and some researchers have also constructed the classifier part as a shallow neural network module through which the distance relationship between samples is measured for the purpose of classification.

4 Experiment

The miniImageNet is a classical evaluation dataset for the FSL algorithm, which contains 60,000 images from different common classes, including animals, plants, and so on. It is can perform the evaluation task of the FSL algorithm excellently and effectively. Therefore, to verify the effectiveness of miniDigPath as an FSL evaluation dataset, we compared the evaluation results of FSL models on miniDigPath and miniImageNet.

4.1 Setting of experimental parameters

We split miniImageNet and miniDigPath into training task，validation task and testing task set ratio of 64:16:20 and 44:10:13. Meanwhile, to ensure fair comparison with the miniImageNet, the input resolutions is 224×224, we following Baseline++[18],all models is evaluated on 5-way 1-shot/5-shot settings and report the mean and 95% confidence interval over 2000 few-shot tasks.The all models are trained for 400 epochs, using a cross entropy loss and an Adam optimizer with an initial learning rate 0.001. Meanwhile, to ensure fair comparison with the miniImageNet, we use the classification accuracy as the evaluation criterion and get a lot of experimental results on the widely used 5-Way 1-Shot and 5-Way 5-Shot setting. (“Way” and “Shot” represents the number of classes in the support set and the number of instances included in every class).

The hardware environment used for the experiments is NVIDIA GeForce RTX 3090 GPU platform; the software environment is Linux system, Python programming language, and Pytorch deep learning framework.

4.2 Experimental results

The experimental results are reported in Table 2. It is show that on the 5-Way 1-Shot setting of the miniDigPath and miniImageNet dataset, except TPN, the classification accuracy ranges of the FSL methods are 40% to 50% and 50% to 60%. In the miniDigPath dataset on the 5-Way 5-shot setting, the highest accuracy rate can only reach 67.65%, the classification accuracies of the other methods are

fluctuated between 50% and 68%, and on the miniImageNet, the evaluation results are basically distributed between 63% and 80%. Therefore, like the miniImageNet dataset, the FSL algorithm can be well distinguished and evaluated by miniDigPath. On the other hand, the evaluation results show that on the 5-Way 1-Shot and 5-Way 5-shot settings of the FSL methods, the classification accuracy on the miniImageNet is improved by 10% and 13% compared with the proposed dataset, and it is indicates that miniDigPath has better evaluation potential than miniImageNet.

Table 2. classification accuracy of FSL model under minidigpath and miniimagenet

Model	**Backbone**	**5-Way**			
		1-Shot		***5-Shot***	
		miniImageNet	***miniDigPath***	***miniImageNet***	***miniDigPath***
RN[6]	Conv-4	50.44±0.82%	43.86±0.76%	65.32±0.70%	54.74±0.35%
GNN[15]	Conv-4	50.33±0.36%	40.02±0.98%	66.41±0.63%	50.84±0.54%
PN[16]	Conv-4	49.42±0.78%	42.07±0.21%	68.20±0.66%	54.63±0.57%
MAML[7]	Conv-4	48.70±1.84%	40.01±0.12%	63.10±0.92%	51.03±0.45%
TPN[20]	Resnet-10	65.41±0.49%	55.96±0.58%	79.22±0.31%	67.65±0.50%
Baseline++[18]	Resnet-10	57.10±0.81%	48.90±0.53%	77.65±0.65%	62.22±0.56%
SSL[21]	Resnet-10	58.10±0.82%	49.87±0.43%	79.20±0.65%	66.09±0.41%

5 Conclusions

We present miniDigPath, a public and available dataset for few-shot digital pathology images. the miniDigPath mainly has the following four contributions: 1) **Standards**, 2) **Evaluation**, 3 **Exploration and** 4) **Research**. We have compared the evaluation potential of miniDigPath and miniImageNet dataset, including the 5-Way 1-Shot and 5-Way 5-Shot settings of the FSL methods. The experimental results indicate that the miniDigPath has better prospect, and on the digital pathology image domain, FSL methods has a good recognition rate. In the next step, we will continue to explore the performance of other classic models on miniDigPath dataset. And we hope the proposed by miniDigPath could facilitate the future few-shot digital pathological image classification research in medical analysis.

Reference

1. Rorke L B. Pathologic diagnosis as the gold standard [J]. 1997, 79(4): 665-667.
2. Campanella G, Hanna M G, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images[J]. Nature medicine, 2019, 25(8): 1301-1309.
3. Noorbakhsh J, Farahmand S, Namburi S, et al. Deep learning-based cross-classifications reveal conserved spatial behaviors within tumor histological images[J]. Nature communications, 2020, 11(1): 1-14.
4. Riasatian A, Babaie M, Maleki D, et al. Fine-Tuning and training of densenet for histopathology image representation using TCGA diagnostic slides[J]. Medical Image Analysis, 2021, 70: 102032.
5. Fe-Fei L. A Bayesian approach to unsupervised one-shot learning of object classes[C]//Proceedings Ninth IEEE International Conference on Computer Vision. IEEE, 2003: 1134-1141.

6. Sung F, Yang Y, Zhang L, et al. Learning to compare: Relation network for few-shot learning[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 1199-1208.
7. Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adap- tation of deep networks[C]//International Conference on Machine Learning. PMLR, 2017: 1126-1135.
8. Garcia V, Bruna J. Few-shot learning with graph neural networks [J]. arXiv preprint arXiv:1711.04043, 2017.
9. Sun H, Zeng X, Xu T, et al. Computer-aided diagnosis in histopathological images of the endometrium using a convolutional neural network and attention mechanisms[J]. IEEE journal of biomedical and health informatics, 2019, 24(6): 1664-1676.
10. Aresta G, Araújo T, Kwok S, et al. Bach: Grand challenge on breast cancer histology images[J]. Medical image analysis, 2019, 56: 122-139.
11. Babaie M, Kalra S, Sriram A, et al. Classification and retrieval of digital pathology scans: a new dataset[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017: 8-16.
12. Kumar M D, Babaie M, Zhu S, et al. A comparative study of CNN, BoVW and LBP for classification of histopathological images[C]//2017 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2017: 1-7.
13. Kather J N, Krisam J, Charoentong P, et al. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study[J]. PLoS medicine, 2019, 16(1): e1002730.
14. Kather, Jakob Nikolas. Histological images for MSI vs. MSS classification in gastrointestinal cancer, FFPE samples [Data set]. Zenodo. http://doi.org/10.5281/zenodo.2530835,2019.
15. Garcia V, Bruna J. Few-shot learning with graph neural networks [J]. ArXiv preprint arXiv: 1711.04043, 2017.
16. Snell J, Swersky K, Zemel R S. Prototypical networks for few-shot learning [J]. ArXiv preprint arXiv: 1703.05175, 2017.
17. Santoro A, Bartunov S, Botvinick M, et al. Meta-learning with memory-augmented neural networks [C]//International conference on machine learning. PMLR, 2016: 1842-1850.
18. Chen W Y, Liu Y C, Kira Z, Wang Y C F and Huang J B.2019. A closer look at few-shot classification//Proceedings of the 7rd International Conference on Learning Representations.
19. Mangla P, Kumari N, Sinha A, et al. Charting the right manifold: Manifold mixup for few-shot learning[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2020: 2218-2227.
20. Yanbin Liu, Juho Lee, Minseop Park, Saehoon Kim, Eunho Yang, Sung Ju Hwang, and Yi Yang. Learning to propagate labels: Transductive propagation network for few-shot learning. In 7th International Conference on Learning Representations, ICLR 2019, 2019.
21. Wang C, Liu R Z, Ye H J and Yu Y. 2020.Novelty-prepared few-shot classification[EB/OL].[2022-06-15].https://arxiv.org/pdf/2003.00497v1.pdf
22. Wang H, Deng Z H. Cross-Domain Few-Shot Classification via Adversarial Task Augmentation [J]. arXiv preprint arXiv:2104.14385, 2021.
23. Phoo C P and Hariharan B.2021.Self-training for few-shot transfer across extreme task differences//Proceedings of the 9rd International Conference on Learning Representations.