Metallographic Structure Recognition With Few Samples Based On Meta-transfer Learning

Kainan Zhang[1,2], Chang Liu[2] and Lixin Tang[1], Min Xiao[3]

[1]National Frontiers Science Center for Industrial Intelligence and Systems Optimization, Northeastern University, Shenyang, 110819, China
[2]Key Laboratory of Data Analytics and Optimization for Smart Industry (Northeastern University), Ministry of Education, Shenyang, 110819, China
[3]Department of Data Intelligence, Xinyu Iron & Steel Group Co., Ltd., Xinyu, Jiangxi Province 338001, China
m15044965880@163.com

Abstract. The metallographic structure reflects the specific internal morphology of metal materials and is closely related to the mechanical properties of materials. Due to the huge diversity and complexity of metallographic structures, it is difficult for traditional methods to identify them automatically and efficiently. In addition, the data samples of metallographic structures are usually small and unbalanced, which poses a great challenge to existing methods. To recognize metallographic structure effectively, a framework of metallographic structure recognition based on meta-transfer learning was proposed. Firstly, unsupervised pre-training is carried out on large-scale data to obtain the backbone architecture of feature extraction to solve the problem of few-shot learning. Then, fine-tuning is adopted to improve the performance in the meta-testing phase, and data enhancement is carried out on the support set to expand the query set, to solve the problem that the feature representation of learning cannot be generalized due to the inconsistent data distribution of new tasks. The results show that the meta-transfer learning framework can effectively identify complex few-sample metallographic structures.

Keywords: Metallographic Structure, Small and Unbalanced, Meta-transfer Learning, Few-shot Learning, Fine-tuning.

1 Introduction

Metallographic structure analysis is an important method for studying metal materials, which has a crucial role in the field of materials science. A large number of studies have shown that the macroscopic characteristics of metal materials are determined by their microstructure. Such as the mechanical properties of materials. Analysis and study of the metallographic structure of metal materials, combined with the manufacturing process and material properties, can provide the theoretical basis for researchers to guide the actual production and the development of new materials. However, the metallographic structure of metal materials is identified and evaluated by experts, and then it is connected with the technological conditions and material properties. The accuracy of the results is greatly affected by human factors. Moreover, the metallographic structure data of metal materials are few, which requires a large number of electron microscopy experiments to obtain, which is an extremely time-consuming and labor-intensive project, bringing huge difficulties to production and research. Therefore, in view of the above problems, how to apply advanced AI technology in the metallographic microstructure analysis of metal materials has become a hot research issue at present.

At present, the metallographic analysis of metal materials mainly consists of manual observation of the sample with the electron microscope, and identification by contrast with the standard metallographic structure, to further analyze and evaluate the composition and state of the tissue. A. C. Souza et al. [1] employed a gray co-occurrence matrix to extract the metallographic structure features of steel and identified each form of metallographic structure of steel by using machine learning. Martin Mueller et al. [3] utilized crystallography and machine learning to achieve the classification of steel microstructure. However, the recognition of metallographic structures based on traditional machine vision technology needs to be combined with expert experience for feature extraction. Human intervention will lead to the loss of some important information and low efficiency, thus reducing the accuracy of recognition.

In recent years, with the continuous development of computer image recognition technology, a new technical way has been opened up for research in the field of materials. Many researchers have used computer-aided metallographic analysis studies to identify and quantify metallographic structure by employing image processing techniques. The recognition of metallographic structure is the premise and foundation of metallographic structure analysis and the core and key of automatic quantitative analysis of metallographic structure. Paul et al. [3] proposed an improved stochastic forest classifier that can use a minimum number of trees for classification. A method of crystal structure recognition based on Bayesian deep learning has been proposed in the literature [4]. This method is robust to structural noise. Azimi et al. [5] proposed a classification of mild steel metallographic structures based on a deep-learning approach. Li et al. [6] used convolutional neural networks to verify the accuracy of the automatic identification of steel microstructure under different metallographic microstructure preprocessing methods. Naik et al. [7] proposed a supervised machine-learning method for the effective identification of phases (ferrite, pearlite, and martensite) in metallurgical ASTM A36 Heat-treated steel. By combining supervised and unsupervised machine learning techniques, De Cost et al. [8] utilized convolutional neural networks to classify subsets of the microstructure of high-carbon steel.

All of the above methods are based on the identification of relatively simple metallographic structures. At present, there are few reports on the automatic recognition of various complex metallographic structures with highly unbalanced data samples. To effectively identify the microstructure with few samples and imbalanced samples, a hybrid metallographic structure recognition framework based on meta-transfer learning is proposed in this paper. As the work of this paper is still in progress, the existing characteristics of this work are described as follows:

1) A meta-transfer learning framework for metallographic structure recognition of a few samples is proposed, which carries out unsupervised pre-training on large-

scale data to obtain backbone architecture for feature extraction to solve the problem of few-shot learning.

2) In the process of meta-testing, the method of fine-tuning is adopted to improve the identification accuracy. Moreover, the support set of metallographic data is enhanced and the query set is expanded to solve the problem that the feature representation of learning cannot be generalized.

The rest of this article is organized as follows. The second section introduces the framework and details of the proposed method. To verify the effectiveness and efficiency of the proposed method, experimental design, and preliminary experimental results are given in Section 3. Section 4 is the conclusion and future work of this paper.

2 The Proposed Method

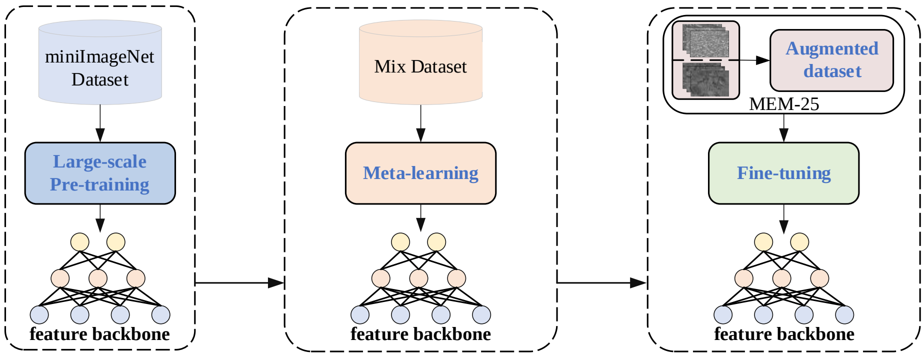

Fig. 1. Meta-Transfer Learning Framework.

We propose a meta-transfer learning framework for the metallographic structure recognition of a few samples. As can be seen from Figure 1, the implementation process of the meta-transfer learning framework mainly includes three stages, namely pre-training, meta-learning, and fine-tuning (FT). First, the self-supervised loss is used to pretrain the features of the unlabeled Benchmark dataset with few samples, to obtain the backbone network architecture of feature extraction. Then, we use ProtoNet (PN) loss in meta-learning to conduct meta-training on the network backbone architecture of feature extraction for small sample tasks of labeling simulation, so that it can quickly adapt to a few tasks. Finally, the backbone network architecture of feature extraction after meta-training is deployed on a new small sample task, and the network structure and parameters are fine-tuned according to the enhanced data set of each metallographic organization category, so that the final recognition model can achieve the desired effect.

2.1 Large-scale Pre-training of Feature Backbone

In the case of computer vision tasks, large-scale pre-training is a very important task. Self-supervised learning algorithms [9] are a more mature pre-training method that requires the network itself to learn to understand the visual world around it without any labels as a way to achieve self-supervised learning. DINO [10] is the most widely used pre-training mechanism at present. Due to the flexibility of the DINO mechanism, it can be applied in traditional convolutional networks. Therefore, the DINO mechanism is adopted in this paper for pre-training. Then ResNet is used to obtain the backbone of feature extraction.

2.2 Meta-learning with ProtoNet

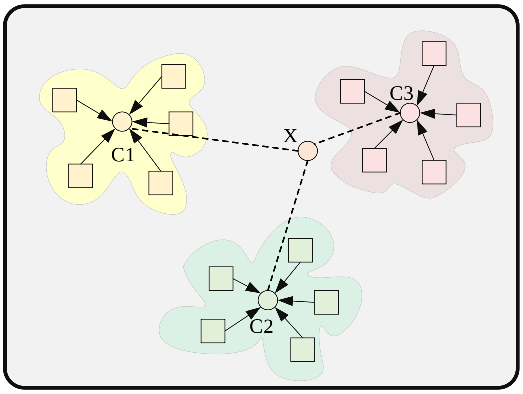

Fig. 2. Prototypical Networks in the Few-Shot Scenario.

A Few-Shot Learning (FSL) task can be used to quickly establish the ability to recognize new concepts from just one or a few examples. At present, meta-learning is the main method to deal with small sample learning problems. A prototypical network (ProtoNet) [11] is one of the mainstream methods of meta-learning. As shown in Figure 2, ProtoNet in the Few-Shot Scenarios. This network can identify new categories that have never been seen in the training process and requires only a small amount of sample data for each category. By constructing the feature mapping function f_ϕ, ProtoNet maps the sample data of dimension D to the space of dimension M and extracts their "mean value" to represent it as the primitive form c_k of this class. Meanwhile, using Euclidean distance as the distance metric, Protonet trains the data of this class to be the closest to the primitive form representation of this class, while being far away from the primitive form representation of other classes. The training process is to minimize the objective function by stochastic gradient descent. The formula is expressed as follows:

$$f_\phi : \mathbb{R}^D \to \mathbb{R}^M \tag{1}$$

$$c_k = \frac{1}{|C_k|} \sum_{(\mathrm{X}_i, y_i) \in C_k} f_\phi(\mathrm{X}_i) \tag{2}$$

$$L(\theta) = -\log p_\theta(y = k \mid \mathrm{X}) \tag{3}$$

where k is the real label of the training sample. c_k is the prototype of class k. $L(\theta)$ is the objective function.

During the test, softmax is applied to the distance between the test data query and the original data of each category to determine the category label of the test data. The expression of X belonging to class k is:

$$p_\phi\left(y=k\middle|\mathrm{X}\right)=\frac{\exp\left(-d\left(f_\phi\left(\mathrm{X}\right),c_k\right)\right)}{\sum_{k'}\exp\left(-d\left(f_\phi\left(\mathrm{X}\right),c_{k'}\right)\right)} \tag{4}$$

where d is the cosine distance.

2.3 Meta-testing with Fine-tuning

To maintain consistency with meta-training, the ProtoNet network of meta-training is directly deployed on the new task. Since the data of the new task is a category that has never been seen before, it may lead to the failure of generalization of the feature representation of learning. Different from the previous use of support set data to fine-adjust the weight of the model, we adopt the data enhancement method to expand the support set data as query data, update the whole backbone network architecture, and achieve the ultimate purpose of identifying the metallographic organization structure.

3 Experiments

3.1 Dataset Description

Metallographic structure image data was provided by a steel plant and obtained by data mining. Figure 3 shows part of the metallographic structure data. Due to the small scale of the obtained metallographic data, the metallographic structure data used in the paper was obtained by manual selection and cropping, and named MEM-25. The dataset contains about 335 metallographic images in a total of 25 types. Among them, individual metallographic structure charts only a few maps. To facilitate training, set the size of the metallographic image is 224×224 pixels. We randomly divided 25 types of MEM-25 datasets into 15 kinds of training, 5 kinds of validation, and 5 kinds of tests.

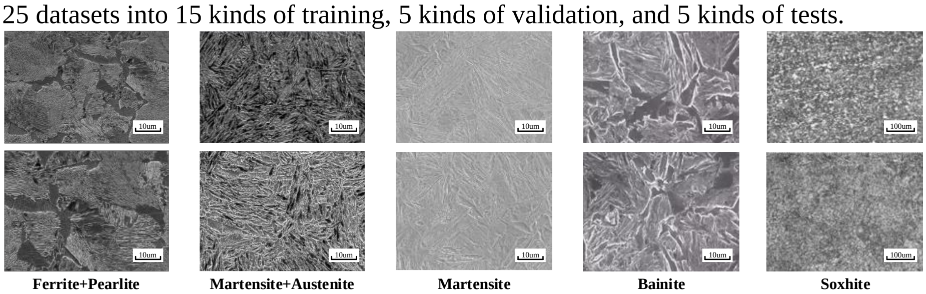

Fig. 3. Partial Chart of Metallographic Structure Data.

3.2 Experimental Settings

To validate the performance of the proposed few-sample metallographic tissue recognition framework, two sets of experiments were conducted with mixed data (CIFAR-FS [12] and MEM-25) and MEM-25, respectively. The whole experimental process is divided into three stages. Firstly, miniImageNet [13] is used as the pre-training dataset, and self-supervised loss is used to pre-train the unlabeled data to generate the feature extractor. Then, two sets of experiments were conducted for the meta-learning process. The difference between the two experiments is that different datasets CIFAR-FS (64 training, 16 validation, and 20 testing classes) and MEM-25 were selected respectively in the meta-training phase. The number of training tasks was set to 100 and validation tasks to 30. In the meta-testing phase, the data are chosen from the test data divided by MEM-25. For CIFAR-FS, the convention is to evaluate 5-way-1-shot (5w1s) or 5-way-5-shot episodes, and the size of the query set for each episode is fixed to 15×5. Due to limited EME-25, we selected the 5-way-1-shot (5 represents the number of categories in the task and 1 represents the number of samples contained in each category) problem as the training and evaluation benchmark. The training epoch was set to 100. To evaluate the small-sample classification performance of the framework, 100 evaluation tasks were assigned from the divided test set. The evaluation indicator is the average classification accuracy of all test tasks. Finally, fine-tuning is utilized in the meta-testing phase to achieve further improvement in the framework performance. In this process, the meta-trained feature backbone is deployed on a new few-sample learning task, and the support set for each task is expanded using data augmentation to achieve performance improvement.

The entire experiment was implemented via Python 3.7, and Pytorch 1.8.0, on an RTX 3060 NVIDIA GPU.

3.3 Results and Analysis

As this work is still in progress, the adequacy of the experimental design is still insufficient. This section gives and analyzes the preliminary experimental results for the two groups of experiments.

Mixed Samples Experiment. In this subsection, training samples from CIFAR-FS and test samples from MEM-25 were used for meta-learning to realize the recognition of metallographic structures. Firstly, the training data of the CIFAR-FS is used for meta-training, and then the test data divided by MEM-25 is used for meta-testing. The feature extractor adopts a fixed feature backbone architecture, which is combined with ProtoNet to obtain the test results of metallographic structure recognition. To further improve the recognition accuracy, the feature extractor in the meta-testing phase is fine-tuned to optimize the recognition accuracy, to realize the purpose of accurately identifying the metallographic structure. The experimental results are shown in Table 1.

Table 1. Experiment results of mixed samples.

Architecture	Pre-Training	Meta-Training(CIFAR-FS)	MEM-25(5w1s)
ResNet50	DINO(IN1K)	ProtoNet(PN)	56.72%
		PN+FT	75.63%

In the process of fine-tuning, the model is optimized by adjusting the learning rate and other hyperparameters to adapt to the new small sample task. It can be seen from Table 1 that the few-shot learning based on the fine-tuning method has higher recognition accuracy than meta-learning without fine-tuning method. Experimental results show the effectiveness of fine-tuning in the meta-testing phase.

MEM-25 Samples Experiment. Firstly, meta-training is conducted based on the training data of MEM-25, and then meta-testing is conducted using the test data divided by MEM-25 to obtain the test results of metallographic structure recognition. To further improve the recognition accuracy, the feature extractor in the meta-testing phase is fine-tuned to optimize the recognition accuracy, to realize the purpose of accurately identifying the metallographic structure. The experimental results are shown in Table 2.

Table 2. Experiment results on MEM-25.

Architecture	Pre-Training	Meta-Training(MEM-25)	MEM-25(5w1s)
ResNet18	-	ProtoNet(PN)	59.63%
		PN+FT	83.24%
ResNet50	-	ProtoNet(PN)	58.20%
		PN+FT	85.08%
ResNet50	DINO(IN1K)	ProtoNet(PN)	65.36%
		PN+FT	90.72%

As shown in Table 2, in the meta-testing phase, fine-tuning can effectively improve identification accuracy. Compared with Table 1, the experimental results show that in the meta-learning phase, image features can be effectively learned by using the same data set for training, and the prior knowledge learned in the training process can be effectively utilized to obtain higher accuracy in the testing process. For Table 2, different feature backbones have different recognition accuracy. The more features that can be learned from the network backbone architecture used, the better the fine-tuning will be during the meta-testing process.

Comparison with ProtoNet without DINO. To verify the effectiveness of the DINO method, sample data of MEM-25 is used to test the prototype network without DINO, and the experimental results are shown in Table 2. As can be seen from the table, the experimental result is lower than that of ProtoNet with DINO, which proves the effectiveness of the DINO pre-training method.

Comparison with traditional ProtoNet. To verify the effectiveness of the framework, the traditional ProtoNet is used to experiment on MEM-25, and the experimental results are shown in Table 3. As can be seen from the table, the accuracy of metallographic structure recognition is 65.47%.

Table 3. Experiment results compared with traditional ProtoNet on MEM-25.

Method	MEM-25(5w1s)
traditional PN	65.47%
PN+FT	90.72%

4 Conclusion

For the complexity of the material metallographic organization, few samples, and many unbalanced categories, this paper proposes a few-shot learning method for metallographic organization recognition. The method introduces transfer learning, which is effectively combined with meta-learning to achieve accurate recognition of metallographic organization under a few sample conditions. To further improve the recognition accuracy, a fine-tuning approach is utilized to optimize the feature extraction backbone architecture in the meta-testing phase. The experimental results show that the framework can effectively identify metallographic structures, which is important for guiding the actual production and development of new materials.

The work in this paper is still in progress, and in subsequent research, mathematical methods and optimization methods will try to be integrated into the framework to improve the identification accuracy.

Acknowledgments

This work was supported in part by the Major Program of the National Natural Science Foundation of China (72192830, 72192831), the National Natural Science Foundation of China (62176049), and the 111 Project (B16009).

References

1. Souza A C, Silva G C, Caldeira L, et al.: An enhanced method for the identification of ferritic morphologies in welded fusion zones based on gray-level co-occurrence matrix: a computational intelligence approach. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 235(7), 1228-1240 (2021).
2. Müller M, Britz D, Ulrich L, et al.: Classification of bainitic structures using textural parameters and machine learning techniques. Metals 10(5), 630 (2020).
3. Paul A, Mukherjee D P, Das P, et al.: Improved random forest for classification. IEEE Transactions on Image Processing 27(8), 4012-4024 (2018).
4. Leitherer A, Ziletti A, Ghiringhelli L M.: Robust recognition and exploratory analysis of crystal structures via Bayesian deep learning. Nature Communications 12(1), 6234 (2021).
5. Azimi S M, Britz D, Engstler M, et al.: Advanced steel microstructural classification by deep learning methods. Scientific Reports 8(1), 2128 (2018).
6. Fan L. Li W, Chen J.: Automatic identification of microstructure of iron and steel material based on convolutional neural network. Journal of Iron and Steel Research 32(1), 33-43, (2020).
7. Naik D L, Sajid H U, Kiran R.: Texture-based metallurgical phase identification in structural steels: a supervised machine learning approach. Metals 9(5), 546 (2019).
8. DeCost B L, Francis T, Holm E A.: Exploring the microstructure manifold: image texture representations applied to ultrahigh carbon steel microstructures. Acta Materialia 133, 30-40 (2017).
9. Goyal P, Mahajan D, Gupta A, et al.: Scaling and benchmarking self-supervised visual representation learning. In: Proceedings of the IEEE/CVF International Conference on computer vision 2019, pp.6391-6400. IEEE, Seoul, Korea (South) (2019).

10. Caron M, Touvron H, Misra I, et al.: Emerging properties in self-supervised vision transformers. In: Proceedings of the IEEE/CVF international conference on computer vision 2021, pp. 9650-9660. IEEE, Montreal, QC, Canada (2021).
11. Snell J, Swersky K, Zemel R.: Prototypical networks for few-shot learning. Advances in neural information processing systems 30, 1-11 (2017).
12. Oreshkin B, Rodríguez López P, Lacoste A.: Tadam: Task dependent adaptive metric for improved few-shot learning. Advances in neural information processing systems 31, 1-13 (2018).
13. Vinyals O, Blundell C, Lillicrap T, et al.: Matching networks for one shot learning. Advances in neural information processing systems 29, 1-9 (2016).

Reviewer #1 Summary
The present paper is dedicated to the interesting topic of metallographic structure recognition. There are several questions which need to be resolved before the paper can be accepted for publication.
Response to Reviewer #1
We would like to express our appreciation to you for your insightful comments and valuable suggestions. Following your comments and suggestions, we have carefully revised our paper, and we think the revised paper has been much improved.

Comment 1: Traditionally during microstructure investigations microstructures are analyzed with different magnification. However, authors are not mentioning in the text how the micrographs obtained with different magnification were treated. It's also unclear how the particle sizes were taken into account.
Response: Thank you for the valuable comment. In the revised manuscript, we have supplemented the data description section. In addition, the scale is annotated on the metallographic structure image in Figure 3.

Comment 2: It's unclear what kinds of units were used for the "size" 224x224.
Response: Thank you for the valuable comment. In the revised manuscript, we have redescribed the size of the metallographic image. The "size" 224x224 refers to the pixel size of the image.

Reviewer #3 Summary
The paper is in principle on scope for DACOMSIN, addressing the theme of "knowledge discovery and machine learning in materials research." Specifically, it applies machine learning to the microstructure of metal materials by image analysis (presumably from electron microscopy).
Response to Reviewer #3
We would like to express our appreciation to you for your insightful comments and valuable suggestions. The following are our point-to-point responses to your comments.

Comment 1: Instead of "Literature [5]" it should be "Azimi et al. [5]".
Response: Thank you for the valuable comment. We have replaced Literature [5] with Azimi et al. [5] in the third paragraph of the Introduction.

Comment 2: The phrase "microstructure structures" appears redundant.
Response: Thank you for the valuable comment. We have replaced microstructure structures with microstructure in the fourth paragraph of the Introduction.

Comment 3: The use of the word "experiment" in the paper is confusing, since first there were electron microscopy experiments (not done by the authors), and then "experiments" with the data.
Response: Thank you for the valuable comment. We have replaced experiments with electron microscopy experiments in the fourth paragraph of the Introduction.

All other experiments in this paper refer to machine learning data experiments.