ISSN: 2414-1895 DOI: 10.6919/ICJE.202502_11(2).0021

Few-Shot Learning with Multi-Scale Feature Fusion for Microalgae Classification

Guihong Yuan*

School of Information Engineering, Dalian Ocean University, Dalian 116000, China

*Correspondence Author: yuanguihong_work@163.com

Abstract

Phytoplankton plays a multifaceted role in the field of aquaculture, serving not only as an important food source for farmed organisms but also rich in bioactive compounds and nutrients. Due to their sensitivity to changes in the aquatic environment, phytoplankton are crucial biological indicators for assessing water quality. Traditionally, the identification of phytoplankton relies on the manual collection of samples and professional analysis under a microscope, a process that is both time-consuming and labor-intensive, and requires a high level of professional expertise and identification skills from the inspectors. Additionally, the recognition of the subtle features of microalgae is quite challenging. To address this issue, this study proposes a few-shot learning algorithm that integrates multi-scale features, aiming to enhance the efficiency of feature extraction for microalgae. Furthermore, the algorithm incorporates metric learning techniques. Under the experimental setup of 5-way 5-shot, the classification accuracy of this method reached 80.51%, which is a 3.06% improvement over the previous suboptimal model.

Keywords

Phytoplankton Microalgae; Multi-Scale Feature Fusion; Few-Shot Learning; Metric Learning.

1. Introduction

Marine microalgae are fundamental, minute single-celled organisms that are widely distributed in various aquatic environments [1]. Within marine and lake ecosystems, these microalgae exhibit rapid growth rates and vast populations, with high photosynthetic efficiency and yields per unit area [2]. Microalgae absorb solar energy and convert it into organic matter while releasing oxygen, and they have the ability to regenerate cyclically. Microalgae are not only responsible for producing nearly half of the oxygen in the atmosphere, but they also play a crucial role as producers in nature, providing essential substances and energy for numerous aquatic organisms, and have a significant impact on ecosystems and water quality [3]. Harmful algal blooms are a typical phenomenon affecting the marine environment; they not only deteriorate water quality and threaten the survival of aquatic organisms but may also lead to the extinction or death of certain marine species [4]. Therefore, in specific scenarios, on the one hand, it may be difficult or expensive to obtain a large number of labeled samples; on the other hand, for each type of training data, it is necessary to train an excellent and rapid recognition ability model using only a small number of labeled samples, which can be achieved through few-shot learning algorithms [5].

In recent years, more and more methods using meta-learning have achieved good results in few-shot learning, especially the metric-based learning method has become the mainstream of the meta-learning framework in few-shot classification algorithms. This method performs classification by

ISSN: 2414-1895 DOI: 10.6919/ICJE.202502_11(2).0021

measuring the distance between samples in the feature space. In 2017, researchers Snell proposed the concept of a Prototypical Network (PN) [6]. This network uses a special embedded representation method that clusters feature points of each category around a unique prototype. Through a simple neural network consisting of four layers, the model is able to map input data to a high-dimensional metric space and calculate the prototype for each category, which is the average of these high-dimensional feature vectors. Then, by calculating the squared Euclidean distance between test samples and these prototypes and using the softmax function to convert these distances into probability values, the category of the sample is predicted. In 2018, researchers Sung introduced the Relation Network (RN), a method that uses trainable deep metric algorithms for the classification of new image categories[7]. The Relation Network consists of two main modules: the Embedded Module (EM) and the Correlation Module (CM). The EM is responsible for extracting feature information from images, while the CM outputs the similarity between images and ultimately determines the category of the images. By 2020, researchers Tian discovered a high-performance embedding model. Their algorithm optimizes the training process of the embedding model and incorporates the concept of knowledge distillation, further enhancing the feature extraction capabilities of the embedding model through multiple iterations[8]. In the same year, researchers Zhang proposed a new distance measurement method called Earth Mover's Distance (EMD), which uses linear programming to find the best match between patches in two images, assigning different weights to patches in different positions to calculate the similarity between query and support set images[9]. In 2021, researchers Lu introduced a lightweight and effective attention method called the Pyramid Spreading Attention (EPSA) module. By replacing the PSA module in the bottleneck block of ResNet, they created a new efficient Pyramid Splitting Attention (EPSA) structure[10]. The EPSA module can be added to a mature network architecture as a plug-and-play component, significantly improving model performance. Finally, in 2022, researchers Xie proposed a few-shot classification method based on deep Brownian distance covariance, which quantifies the correlation between two few-shot image random variables by calculating the Euclidean distance between the feature function and the marginal product[11].

Meta-learning methods are widely used in measurement-based methodologies due to their simplicity and effectiveness. However, most measurement learning models rely mainly on the global features of images, neglecting the local features. In scenarios where microalgae sample data is scarce, using global features or local features alone cannot adequately represent the distribution of shapes and textures for different microalgae categories. In order to more effectively utilize the contextual information of local and global features in microalgae images and enhance the model's feature representation and generalization ability, this study proposes a novel multi-scale feature fusion algorithm for the classification of microalgae in few-shot learning tasks. This method first performs multi-scale feature extraction on algae images to capture both global and local features simultaneously and fuse them together. Then, a measurement module based on Deep Brownian Distance Covariance (BDC) is introduced to represent measurement features and is combined with the classic meta-learning framework algorithm, Prototypical Network. We have designed a BDC measurement module based on the meta-learning architecture, which considers not only the marginal distribution between features but also the joint distribution, thereby achieving better measurement feature representation and improving the classification accuracy of microalgae.

2. Materials and Methods

2.1 Materials

2.1.1 Data Pre-processing

Due to the insufficient number of marine microalgae microscopic image samples collected in the laboratory, it is necessary to augment the collected algae cell microscopic images through data augmentation to provide sufficient sample support for subsequent microalgae identification and small sample learning classification algorithms. At the same time, during the actual collection process, due to limitations such as acquisition equipment and methods, the original images obtained are often restricted by various factors and cannot fully reflect all the information of the original images. Denoising processing is required to obtain the processed partial typical marine microalgae microscopic image dataset as shown in Figure 1.

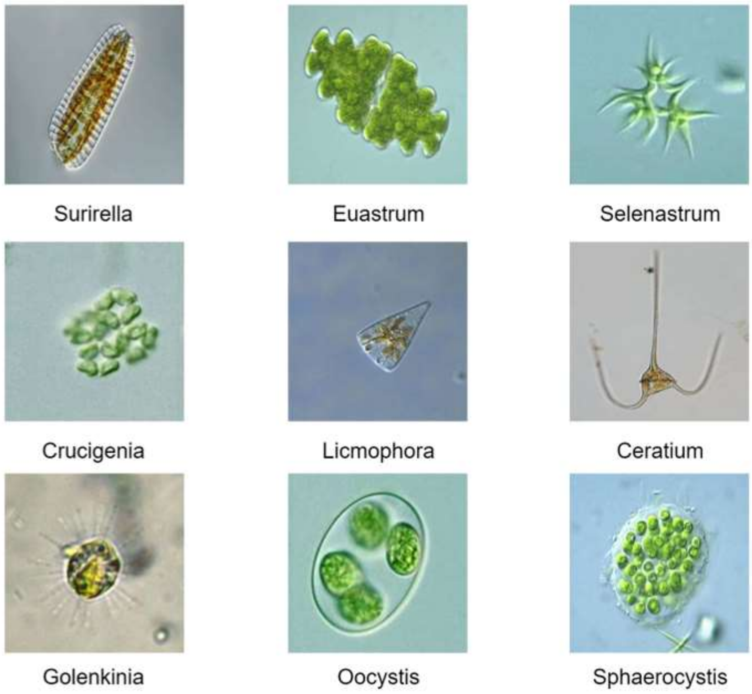

Figure 1. Partial typical microalgal image data

This article adopts a series of computer image processing technologies, including grayscale adjustment and denoising filtering, to optimize the original microalgae microscopic images. In order to enhance the usability of key information in the images, particle suppression technology is also implemented in this article to reduce unnecessary image details. Although operations such as grayscale adjustment, denoising, and particle suppression are performed in the preprocessing stage, and Gaussian noise and salt and pepper noise are introduced in the data augmentation stage, this approach may have some overlap in operations. However, this process helps to significantly enhance the robustness, generalization ability, and training effectiveness of the model, and can generate a more diverse set of training samples, effectively preventing the model from overfitting. The constructed dataset covers 150 categories of marine microalgae, with 100 samples per category, totaling 15,000 images. This article refers to the allocation method of the fine-grained small-sample public bird dataset CUB [12] and divides the dataset into training, validation, and test sets, with an approximate ratio of 2:1:1. This division strategy helps effective training, optimization, and evaluation of the model, thereby improving the model's performance and generalization ability. The specific details of the dataset division are shown in Table 1.

Table 1. Dataset Partitioning Indicators

Dataset	Marine Microalgae
Total number of data sets	15,000
Number of categories	150
Number of classes	100
Image size	84*84

2.2 Methods

2.2.1 The Overall Architecture of the Model

The model proposed in this paper is an end-to-end trainable neural network designed for few-shot classification. The model consists of two main components: the multi-scale local feature fusion module and the BDC metric module. The detailed process is shown in Figure 2.

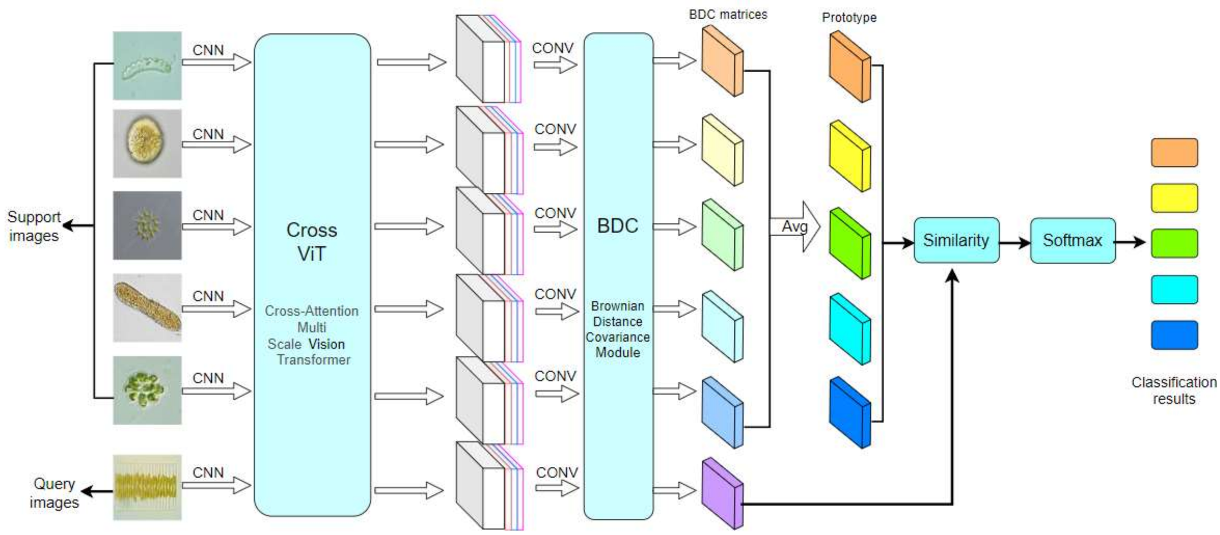

Figure 2. The overall model architecture for the example of 5way1shot

2.2.2 Feature Fusion Module

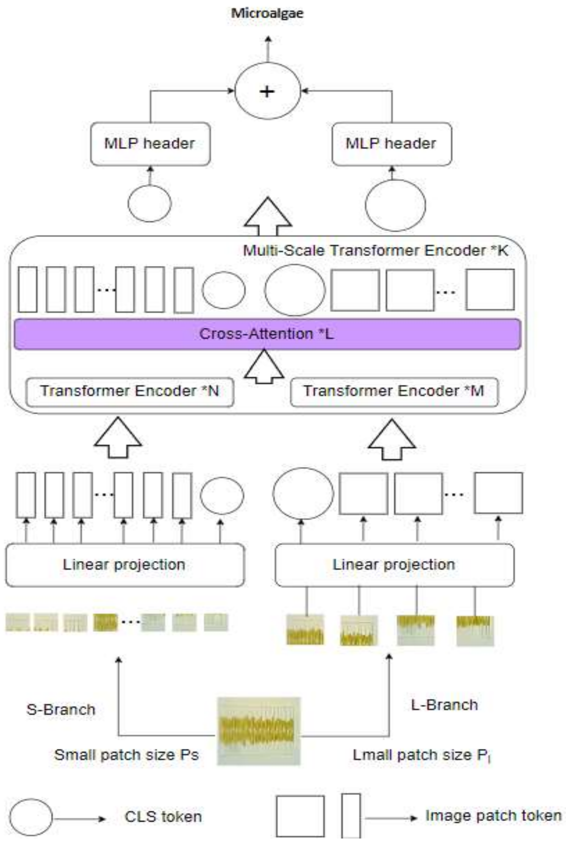

Figure 3. Feature Fusion Module

The network architecture proposed in this paper is the Cross-Attention Multi-Scale Vision Transformer (CrossViT) [13]. The model consists of K multi-scale transformer encoders, each containing two branches: L-Branch: A large (main) branch that uses a coarse-grained patch size and more encoders with a larger embedding dimension. S-Branch: A small (auxiliary) branch that uses a fine-grained patch size and fewer encoders with a smaller embedding dimension. The two branches are fused L times, and predictions are made using the class tokens from both branches. As shown in Figure 3.

2.2.3 Metric Learning

In the measurement module, this paper selects Brownian Distance Covariance (BDC) [14] for metric learning, which is defined as the Euclidean distance between the joint feature function and the product of marginal distributions. This method can naturally quantify the dependency between two random variables [15]. For discrete observations of microalgae features, the BDC metric is decoupled and plug-and-play. Therefore, this module can be understood as a pooling layer that can be seamlessly integrated into deep learning networks. It takes the processed microalgae feature maps as input and outputs the BDC matrix as the image feature representation. In this way, the similarity between two microalgae images is calculated as the inner product or Euclidean distance between the corresponding two BDC matrices. Let X ∈ Rp and Y ∈ Rq be random vectors with dimensions p and q, respectively. Under the condition that X and Y have finite first-order moments, the BDC metric is defined as follows:

$$\rho(X,Y) = \int_{R^p}\int_{R^q} \frac{|\phi XY(t,s)-\phi X(t)\phi Y(S)|^2}{C_p C_q \|t\|^{1+p}\|s\|^{1+q}} dt\, ds \tag{1}$$

The evaluation method adopted in this paper is accuracy with a confidence interval, which can be used to measure the uncertainty of the estimated values [16]. The accuracy calculated in this manner is more authentic and reliable, and the overall evaluation index for each test task is calculated as follows:

$$acc = \frac{T}{all} \tag{2}$$

In this context, "acc" represents the accuracy of classification, "T" indicates the number of samples correctly classified, and "all" represents the total number of samples. Finally, the overall accuracy is obtained by summing the average accuracy of 2000 rounds of tests for each of the 5 test tasks and then calculating the average.

3. Results and Analysis

3.1 Experimental Setup

This study divides the marine microalgae dataset into three subsets for experimentation according to their categories: the marine microalgae training dataset, the validation dataset, and the test dataset. There is no overlap in categories among these three subsets, and the training dataset has the largest number of sample categories. The study first uses the training dataset and the validation dataset for model training, and then employs the test dataset to evaluate the classification performance of the best model generated during the training phase.

Training process: The weights of the main network in the model are pre-trained, and based on this, other modules are coordinated to fine-tune the network, providing optimal parameters for subsequent N-way K-shot few-shot classification tasks. The total number of epochs in this stage is 100, with 200 sets used for training and 600 sets for validation in each epoch. In the testing phase, 2000 sets of tasks are randomly selected from the test set to evaluate the average classification accuracy. Within a set, N categories are selected, and for each selected category, a support set and a query set are chosen. The model is trained once under the support and query sets of multiple selected categories.

In each experiment, the proposed method is compared with four of the most typical and advanced few-shot classification methods currently available: RelationNet, ProtoNet, Good-Embed, Meta DeepBDC. The results of the 5-way-5-shot classification task on the marine microalgae dataset were compared, as shown in Table 2:

Table 2. displays the results of the 5-way-5-shot classification on the marine microalgae dataset.

Method	Backbone	5-Way 5-Shot
RelationNet	ResNet-12	72.97±0.42
ProtoNet	ResNet-12	70.77±0.31
Good-Embed	ResNet-12	75.16±0.35
Meta DeepBDC	ResNet-12	77.45±0.37
Ours Cross ViT-BDC	ResNet-12	80.51±0.35

In this table, the proposed Cross ViT-BDC method signifies the integration of the multi-scale fusion module with the BDC metric module, which can be understood as processing the effects achieved by combining local and global features. This method has yielded excellent results on the marine microalgae classification dataset. In the 5-way-5-shot few-shot classification task, the classification accuracy of this method reached 80.51%, which is a 3.06% improvement over the best results of the four mainstream algorithms.

4. Conclusion

This study, targeting the application of few-shot learning in the identification and classification of microalgae, proposes a novel multi-scale local feature fusion strategy designed to delve into the fine-grained features of microalgae. This approach effectively integrates the local and global features of microalgae images, significantly enhancing the model's feature extraction and generalization capabilities. Ultimately, the BDC metric method from meta-learning is used to evaluate the similarity of different microalgae for classification purposes. Compared with numerous traditional few-shot learning algorithms, this method demonstrates a higher classification accuracy on microalgae datasets. Looking ahead to future research, on one hand, given the vast variety of microalgae species, there is a need to expand the dataset by collecting a wider range of marine microalgae. On the other hand, considering that most existing classification algorithms for marine microalgae are offline processes and lack a real-time classification system, which limits their applicability in practical scenarios. Therefore, converting the algorithm into a real-time classification system is one of the key issues we need to address in the future.

Acknowledgments

Natural Science Foundation.

References

[1] Mugani R, El Khalloufi F, Kasada M, et al. Monitoring of toxic cyanobacterial blooms in Lalla Takerkoust reservoir by satellite imagery and microcystin transfer to surrounding farms[J]. Harmful Algae, 2024, 135: 102631.

[2] Chorus I, Fastner J, Welker M. Cyanobacteria and cyanotoxins in a changing environment: Concepts, controversies, challenges[J]. Water, 2021, 13(18): 2463.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202502_11(2).0021

[3] Liu J Y, Zeng L H, Ren Z H. Recent application of spectroscopy for the detection of microalgae life information: A review[J]. Applied Spectroscopy Reviews, 2020, 55(1): 26-59.

[4] Li X Y, Yu R C, Richardson A J, et al. Marked shifts of harmful algal blooms in the Bohai Sea linked with combined impacts of environmental changes[J]. Harmful Algae, 2023, 121: 102370.

[5] Song Y, Wang T, Cai P, et al. A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities[J]. ACM Computing Surveys, 2023, 55(13s): 1-40.

[6] Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning[J]. Advances in neural information processing systems, 2017, 30.

[7] 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.

[8] Tian Y, Wang Y, Krishnan D, et al. Rethinking few-shot image classification: a good embedding is all you need?[C]//Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIV 16. Springer International Publishing, 2020: 266-282.

[9] Zhang C, Cai Y, Lin G, et al. Deepemd: Few-shot image classification with differentiable earth mover's distance and structured classifiers[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 12203-12213.

[10] Zhang H, Zu K, Lu J, et al. Epsanet: An efficient pyramid split attention block on convolutional neural network[J]. CoRR, 2021.

[11] Xie J, Long F, Lv J, et al. Joint distribution matters: Deep brownian distance covariance for few-shot classification[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 7972-7981.

[12] Wah, C., Branson, S., Welinder, P., Perona, P., and Belongie, S.J. 2011. The Caltech-UCSD Birds-200-2011 Dataset.

[13] Chen C F R, Fan Q, Panda R. Crossvit: Cross-attention multi-scale vision transformer for image classification[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 357-366.

[14] Székely G J, Rizzo M L. Brownian distance covariance[J]. 2009.

[15] Székely G J, Rizzo M L, Bakirov N K. Measuring and testing dependence by correlation of distances[J]. 2007.

[16] Zheng Y, Liu S, Chen H, et al. Hybrid FusionNet: A hybrid feature fusion framework for multi-source high-resolution remote sensing image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024.