

Article

Multiscale Local Feature Fusion: Marine Microalgae Classification for Few-Shot Learning

Dan Liu ^{1,2}, Ting Liu ¹, Hai Bi ³, Yunpeng Zhao ⁴ and Yuan Cheng ^{2,4,*}

¹ College of Information Engineering, Dalian Ocean University, Dalian 116023, China; liudan@dlou.edu.cn (D.L.); lting9821@163.com (T.L.)

² Key Laboratory of Environment Controlled Aquaculture (Dalian Ocean University) Ministry of Education, Dalian 116023, China

³ Hangzhou Yunxi Smart Vision Technology Co., Ltd., Dalian 116023, China; hai_bi@163.com

⁴ Ningbo Institute of Dalian University of Technology, Ningbo 315000, China; ypzhao@dlut.edu.cn

* Correspondence: chen_y_nbi@dlut.edu.cn

Abstract: In the marine ecological environment, marine microalgae is an important photosynthetic autotrophic organism, which can carry out photosynthesis and absorb carbon dioxide. With the increasingly serious eutrophication of the water body, under certain environmental factors, the rapid propagation of some algae in the water body gradually forms a harmful bloom, which damages the water environment. Therefore, how to identify the beneficial algae and harmful algae quickly and accurately has gradually become the key to solve the problem. There are more than 30,000 species of marine microalgae in the world, and the sample data are few and the characteristics are not obvious. Many of the algae are similar in shape and difficult to distinguish. The few-shot learning task is very challenging. By training very few labeled samples, the deep learning model has excellent recognition ability. Meanwhile, the few-shot classification method based on metric learning has attracted considerable attention. In this paper, in order to make full use of image features and improve the generalization ability of the model, a multi-scale local feature fusion algorithm was proposed to classify marine microalgae with few shots. First, the input image is gridded and multiscale processed, and then it is sent to the CTM category traversal module for feature extraction to obtain local features. A local feature fusion module based on the SE-NET self-attention mechanism is designed to obtain local enhanced features containing global information to improve the generalization ability of the model. Classification is realized by calculating the distance between the sample feature vector of the query set and the prototype of the support set. Under the settings of 5-way 1-shot and 5-way 5-shot, the classification accuracy of the proposed method is improved by 6.08% and 5.5%, respectively. It provides a new idea for microalgae identification and a new opportunity for the sustainable development of new energy.

Keywords: marine microalgae; few-shot learning; metric learning; feature fusion; classification



Citation: Liu, D.; Liu, T.; Bi, H.; Zhao, Y.; Cheng, Y. Multiscale Local Feature Fusion: Marine Microalgae Classification for Few-Shot Learning. *Water* **2023**, *15*, 1413. <https://doi.org/10.3390/w15071413>

Academic Editor: Guangyi Wang

Received: 28 February 2023

Revised: 28 March 2023

Accepted: 1 April 2023

Published: 5 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Marine microalgae are primitive, tiny photosynthetic organisms that can release oxygen by converting solar energy and synthesizing organic matter; this process of photosynthetic oxygen release is an important chemical reaction on Earth. Marine microalgae are tiny and can only be seen clearly with the help of a microscope. Various approaches such as computer vision and deep learning-based techniques have been frequently used for object detection [1–3]. At the same time, such methods have also been used to identify algae [4,5]. At present, few species of marine microalgae have been recorded; however, the organic matter synthesized by marine microalgae accounts for half of all photosynthetic products on Earth every year. When encountering something new, humans can usually learn it quickly with a small quantity of learning data. For example, a few pictures can be used to recognize plants that have never been seen before. The deep learning algorithms

used in modern AI systems often require a large quantity of data to be trained in order to achieve a usable model, resulting in a high training cost. Additionally, the large quantity of data for manual labeling is very time-consuming and labor-intensive. Therefore, ideally, AI systems need to have the same fast learning ability as human beings to make deep learning models with excellent recognition ability with only a very small number of labeled samples for each type of training data. This can be achieved with the few-shot classification task [6]. There are more than 30,000 species of marine microalgae worldwide, and most of them require human recognition of microalgae microscopic images. Due to the small size and dense distribution in microalgae images, most algae with similar morphology are analyzed and identified using cell detection techniques, but it takes a long time. In this study, it is important to use the few-shot learning algorithm for the classification and identification of marine microalgae microscopic images to further distinguish algae accurately with less image data and close morphology.

In recent years, most methods using meta-learning have obtained good results in few-shot learning, and metric-based learning methods have become mainstream in meta-learning few-shot classification algorithms, which classify by measuring the distance between samples in the feature space. The twin network structure (Siamese neural network) proposed by Koch G et al. [7] (2015) separately extracts features from two images by the same network structure, and if the feature information of the two images is very close, then they are likely to belong to the same class of objects; otherwise, they belong to different classes of objects. The network proposed by Vinyals et al. [8] (2016) is a matching network (MN) which adopts the form of matching to achieve the few-shot classification task, and introduces the idea of the nearest neighbor algorithm to solve the overfitting problem caused by deep learning algorithms that cannot fully optimize the parameters under the condition of few-shot. It uses a neural network with an attention mechanism and a memory module to solve the problem of over-reliance on the ordinary nearest neighbor algorithm metric function, and maps the sample feature information. Snell et al. [9] (2017) used an embedding representation of the prototypical network (PN) so that the feature points corresponding to each category are clustered around a single prototypical representation. The input information is mapped to a high-dimensional metric space through a simple four-level neural network structure, and, for each category, the mean of its high-dimensional feature vector is taken as the prototype. Finally, the squared Euclidean distance between the corresponding high-dimensional feature vector of the test sample and each prototype is calculated and transformed into probability values using softmax to predict the category of that sample. Sung et al. [10] (2018) proposed the relation network (RN) to learn a portable depth metric capable of comparing the relationship between images. The network is divided into two stages: the first stage is an embedded module for extracting feature information, and the second stage is a relevance module for outputting the similarity score between two images to determine whether two images are from the same category. Li et al. [11] (2019) proposed the deep nearest neighbor neural network (DN4), which compares images with the category by comparing the local features. Zhang et al. [12] (2020) proposed a new distance metric, the Earth Mover's Distance (EMD), which finds the distance between two image blocks by means of linear programming, finding the best match between the individual blocks of two images, assigns different weights to the blocks at different locations, and expresses the similarity between the query set and the support set images by calculating the cost of the best match between the individual blocks of the two. Qi Wang et al. [13] (2021) studied a multi-attention network model, which uses semantic embedding of class tags to generate attention graphs and uses these attention graphs to create image features for one-time learning. Different from other models, the attention generators in this model can be extended to new classes. Xie et al. [14] (2022) proposed a deep Brownian distance covariance (DeepBDC) based method for few-shot classification tasks by combining the Euclidean distance between the feature function and the edge product to quantify the correlation between two random variables. Wei et al. [15] proposed a prototype network based on mixed attention to solve the problem

of relationship classification in a few-shot environment. In this model, an instance-level attention mechanism is designed to select key sample instances to support concentration, so as to alleviate the influence of noise samples on the model. A feature-level attention mechanism is also designed to highlight important feature dimensions in the feature space, so as to alleviate the problem of feature sparsity.

The above methods based on metric learning are simple, effective, and widely popular. However, the metric learning model above uses image-level global features. In the case of sparse data samples, these global features cannot effectively characterize the category distribution, which affects the final classification results to a certain extent. To effectively utilize the local distinguishable information of images and improve the feature representation and generalization ability of the model, this paper proposes a few-shot classification method based on gridded multiscale local feature fusion. First, the input algae images are processed by gridded multiscale chunking and then fed into the CTM category traversal module to obtain local features. To address the problem that the category traversal module is difficult to learn and generalize, a local feature fusion module based on the SE-NET architecture is designed to learn the relationship between multiple local features of each image to obtain a better metric feature representation. Finally, extensive experiments are conducted on a microscopic marine microalgae image dataset to verify the effectiveness of the proposed algorithm. The main contribution of our study in this work is as follows:

- The collection of microscopic image of marine microalgae;
- Trained multiscale local feature fusion, CTM, SE-NET, and a metric module on the marine microalgae dataset;
- The feasibility of the method was verified by ablation experiment;
- Comparative analysis of all the models' accuracies and performances.

2. Category Traversal Module

A category traversal module [16] can be added directly to the network of the original algorithm as a plug-and-play module. Relative to the original category traversal module, CTM extracts "intra-class commonality" and "inter-class uniqueness" in a targeted manner.

The CTM category traversal module consists of two parts: a concentrator for extracting common features within marine microalgae classes, and a projector for extracting unique features between marine microalgae classes. First, the feature information of the marine microalgae support set and query set is obtained by the category traversal module $f_{\theta}(S)$ and $f_{\theta}(Q)$. Then, the feature information of the marine microalgae support set is input into the cascade. Next, the convolutional neural network performs dimensional compression on the input marine microalgae feature. Finally, the average of each marine microalgae class sample feature map is the output as o . The process is represented as follows:

$$f_{\theta}(S) : (NK, m_1, d_1, d_1) \xrightarrow{\text{Concentrator}} o : (N, m_2, d_2, d_2) \quad (1)$$

Eq. NK denotes the input N marine microalgae categories, with K marine microalgae samples in each category, for a total of NK samples. m_1 denotes the number of channels in the marine microalgae signature map. d_1 denotes the size of the marine microalgae feature map, and after the cascader, the number of feature maps is compressed to N because the sample feature maps are averaged for each category. The dimensionality of the feature map is compressed to m_2 , and the size of the feature map is compressed to d_2 . This process removes the differences between the instances and extracts the features common to each instance of each category.

Next, the output o of the cascader enters the projector, and the feature maps corresponding to each category are cascaded along the feature map channel direction. Then, the marine microalgae feature maps are compressed using the convolutional neural network, and, finally, the mask map p corresponding to each feature map is obtained using the

softmax layer in the marine microalgae feature map channel dimension. The process is represented as follows:

$$o : (1, Nm_2, d_2, d_2) \xrightarrow{\text{Projector}} p : (1, m_3, d_3, d_3) \quad (2)$$

After obtaining the output o of the cascade in Equation (2), the dimension of the marine microalgae feature map become d_2 . Then, the CNN in the projector compresses it to d_3 , and the marine microalgae feature map dimension is also compressed to d_3 . This process extracts the unique features between classes. Finally, the initial marine microalgae feature information is obtained through the category traversal module $f_\theta(S)$ and $f_\theta(Q)$.

3. SE-NET Feature Fusion

The extracted initial marine microalgae feature information is fused by the self-attentive feature fusion module $f_\theta(S)$ and $f_\theta(Q)$. The initial feature vectors are stitched together and fused according to the weights to generate a more accurate feature representation of the marine microalgae image. Cao et al. [17] improved the efficiency of the model in processing feature information by extracting important feature information on an individual marine microalgae basis and by fusing the information as a whole. The fused feature representation enhances the association of the sample detailed features and global features. The self-attentive mechanism can reduce the dependence on external information by obtaining the internal correlation of global information features of marine microalgae.

3.1. Squeeze Operation

In the convolution process, the individual feature convolution is a local spatial sensory field, and there is no channel information data. The concept of channel features is only obtained when the sum operation is performed. The purpose of the squeeze operation is to open the correlation between channels only with the sum operation to enhance the channel data correlation. The squeeze operation takes each marine microalgae feature map through the global pooling operation extracted as scalar data. Wang et al. [18] calculated the sum of each pixel value of the $N \times M$ feature map and averaged them to obtain the global average pooling result, which compresses the $N \times M$ feature map into a scalar global description feature. Then, the original $N \times M \times S$ feature map is transformed into a $1 \times 1 \times S$ feature map, and the number of channels remains unchanged.

3.2. Excitation Operation

The global description of each feature map is obtained through the squeeze operation, which opens the connection between channels, and the excitation operation is used to obtain the relationship between channels. Specifically, the corresponding channel weights are obtained by two full-connection operations and the sigmoid activation function. Among them is the ReLU function. $W1$ is the first full-connection operation, the purpose of which is dimensionality reduction, and it has a dimensionality reduction coefficient r as the hyperparameter. $W2$ is the parameter of the second full-connection operation, which restores the dimensionality to the input dimension. After two full connections, a weight normalization operation is performed on each feature map using the sigmoid activation function.

Finally, the output weight matrix of the sigmoid activation function is multiplied with the original feature map to obtain the output feature map with weights, which adds the weight coefficients of each feature map in the channel dimension by a convolution operation. This makes the channel features of the feature map more capable of extracting features, amplifying the effective features, and reducing the features. After the SE block is embedded in a layer, the received marine microalgae feature map is first pooled globally averaged to make the size of the feature map 1×1 , and the number of channels remains unchanged.

4. Materials and Methods

4.1. Problem Definition

For a few-shot marine microalgae target classification task, there exists a marine microalgae support set S , which consists of N different classes of algal images and K labeled samples of each class. Then, a query set Q of algae consists of N classes of images of the same class as the support set S and q unlabeled samples of each class. The purpose of few-shot learning is to enable the model to classify each unlabeled sample in set Q according to set S . This setup is called the N -way K -shot few-shot classification task [19].

The entire seaweed marine microalgae dataset is divided into three subdatasets according to categories, namely, the seaweed training dataset, the seaweed validation dataset, and the seaweed testing dataset, where the training set has the largest number of sample categories. The training and validation datasets are used in training, and the test dataset is used to test the classification performance of the best model generated in the training phase. There is no intersection of the categories in the three subdatasets.

The proposed few-shot classification model based on gridded multiscale local feature fusion is a neural network architecture that can be trained from end-to-end. The general framework is shown in Figure 1, which is a schematic diagram of the few-shot classification task for marine microalgae 5-way 1-shot classification. The proposed method consists of three main modules: a CTM local feature extraction module, a local feature fusion module based on the SE-NET architecture, and a metric module, as shown in Figure 1.

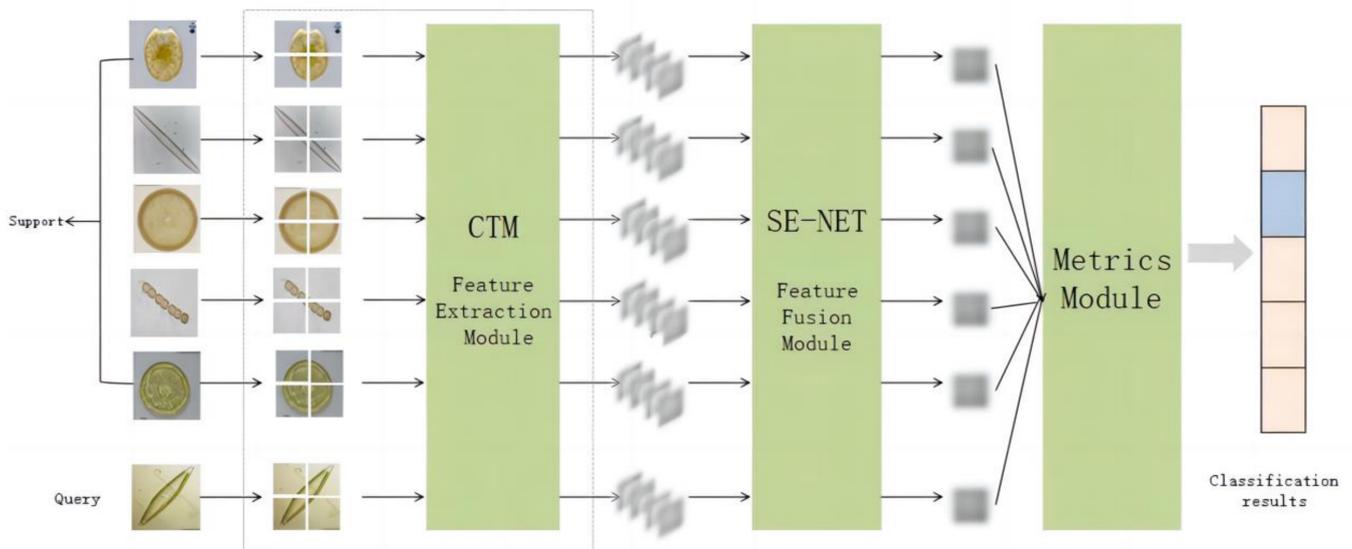


Figure 1. Schematic diagram of the 5-way 1-shot few-shot classification task.

4.2. Backbone

The backbone network selected is the ResNet network, which is modified from the VGG19 network by adding residual units through a short-circuiting mechanism [20]. The ResNet network changes are reflected in the fact that ResNet directly uses the convolution of stride = 2 for downsampling and replaces the fully connected layer with the global average pool layer. An important ResNet design principle is that the feature map doubles the feature channels when the feature map size is reduced by half, which maintains the complexity of the network layers. ResNet adds a short-circuiting mechanism between every two layers compared to the normal networks, which results in residual learning.

The ResNet-12 residual block is composed of four residual blocks of depth 3, each with three convolutional layers with a 3×3 kernel inside, and the number of channels in each stage of the network is [64, 160, 320, 640], and the number of residual blocks is [1, 1, 1, 1], respectively, and a maximum pooling layer of 2×2 is applied at the end of each block [21].

4.3. Feature Extraction Module

The CTM category traversal module is chosen as the feature extraction module, and the main CTM structure is shown in Figure 2, where concentrator represents the cascade for extracting intraclass common features of marine microalgae, and projector represents the projector for extracting interclass unique features of marine microalgae.

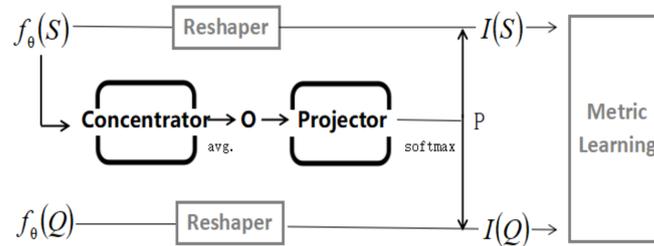


Figure 2. CTM traversal module.

CTM can extract intraclass common features and interclass unique features of marine microalgae in a targeted manner, and the intraclass common features and interclass unique features in the marine microalgae dataset are shown in Figure 3.



Figure 3. Marine microalgae data set.

In Figure 3, in the marine microalgae support set (4-shot), each category contains four samples, and the texture features of the same category samples are the same in four samples, while the shape features are different, indicating that the texture features are common to such samples, which are common features within the marine microalgae category.

The marine microalgae support set (1-shot) contains five sample categories. Each category has only one sample, and each sample contains two features: texture features and shape features. When a query sample query is given, the distance between the query sample query and the first sample is 2, and the texture features and shape features are different from the other 4 samples. The distance between the query sample query and the first sample is 2, and the texture and shape features are different, and the distance between the other 4 samples is 1, and the texture and shape features are different on only 1. Following the traditional metric learning approach, the query sample might be randomly assigned to one of the four categories. By carefully observing the five samples in the support set, it is found that the texture features are different among different categories, while the shape features have one feature in common to different categories, indicating that the texture features are the fundamental features for distinguishing the categories,

and such features with different attributes among different categories are unique features among marine microalgae.

4.4. Self-Attentive Feature Fusion Module

The global feature information of the marine microalgae image can be extracted indiscriminately by the CTM feature extractor, which may be affected by background interference and different sizes and morphologies of the target objects. To better represent the important feature information in the image, an SE-NET self-attentive feature fusion module is used to fuse the important feature information of the marine microalgae image [22]. A microalgae image sample can generate two sets of feature information after passing through the CTM feature extractor. The SE-NET self-attentive feature fusion module is used to fuse the extracted important feature information at two different scales, and, finally, these feature vectors are stitched together to generate a more accurate feature representation of this image. By fusing the microalgae image information as a whole, the efficiency of the model in processing the feature information is improved.

4.5. Metrics Module

The distance measure chosen for the metric module is the squared Euclidean distance in the Bregman scatter, where, for a series of marine microalgae feature points, the nearest point in the feature space to the average distance between all points is the mean of the series of feature points. In [23], the squared Euclidean distance was used as a distance measure, which is better than the cosine distance, and the Bergman scatter as a distance measure can transform the classification problem of few-shot classification into a linear classification model.

The loss function uses the episode small batch gradient descent training method, for which the distance d between the training sample and the prototype is taken as a negative number and transformed into a probability value p using the softmax function, and the negative logarithm of the probability value is taken as the loss. The goal is to make the samples closer and closer to the prototype of the correct category, and further and further from the prototype of other categories. The specific form of the loss function is:

$$J_{(\varphi)} = -\log P_W(y = k|x) \quad (3)$$

4.6. Evaluation Method

The evaluation method used in this paper is accuracy with a confidence interval, which can be used to measure the uncertainty of the estimated value. The accuracy rate calculated in this way is more real and reliable. The overall evaluation index is calculated as follows. First, the accuracy of classification is calculated:

$$acc = \frac{T}{all} \quad (4)$$

where T represents the number of samples correctly classified, and all represents the total number of samples. The mean avg of accuracy is obtained by averaging the sum of multiple accuracies.

5. Experimental Analysis

5.1. Image Dataset

The microalgae images from the Heshi Reef sea area in Dalian city (Liaoning province, China) were used in this study. The image information was collected by extracting seawater samples and placing the samples under the microscope of an optical microscopic imaging system. The specific steps are as follows.

Step 1: An appropriate sample of seawater is taken and it is placed under the microscope of the optical microscopy imaging system.

Step 2: The identified target in the water is magnified and imaged to a clear resolution, and then converted into an image signal through a high-speed Charge Coupled Device imaging device.

Step 3: The image signal is sent to the computer through the image acquisition card to obtain the image information.

The microscopic image acquisition process of marine microalgae is shown in Figure 4:

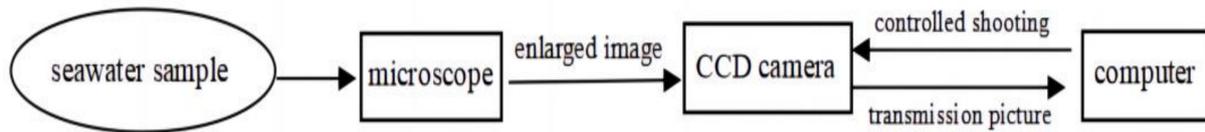


Figure 4. Marine microalgae microscopic image acquisition process.

Based on the marine microalgae microscopic image samples collected in the laboratory, the collected microscopic images of algae cells were expanded to provide sample support for the subsequent microalgae recognition and classification learning algorithm. In most cases, due to the influence of acquisition tools or means in the process of image acquisition, the acquired image, namely, the original image, is restricted by various conditions and often cannot fully reflect all the information of the original image. Some samples of the finally obtained marine microalgae microscopic image data set are shown in Figure 5:



Figure 5. Sample of data set.

The acquisition process of marine microalgae microimages is complicated. In order to avoid image overfitting, it is necessary to expand the marine microalgae microimage.

In practical problems, it is difficult for relevant researchers to collect enough relevant data in some tasks, and the sorting, screening, and the labeling of image data after collection also consumes people's mental and physical energy. In order to obtain more training data, researchers often use data augmentation methods. This method makes certain changes to the original image to make it become two different images in pixel level from the original image. However, the subject predicted in the image does not change, and this method can enrich the data set and effectively avoid overfitting.

According to the collected marine microalgae image data, the data was expanded by using such methods as data flipping, data rotation, data splicing, brightness adjustment, and noise addition. The data augmentation effect is shown in Figure 6.

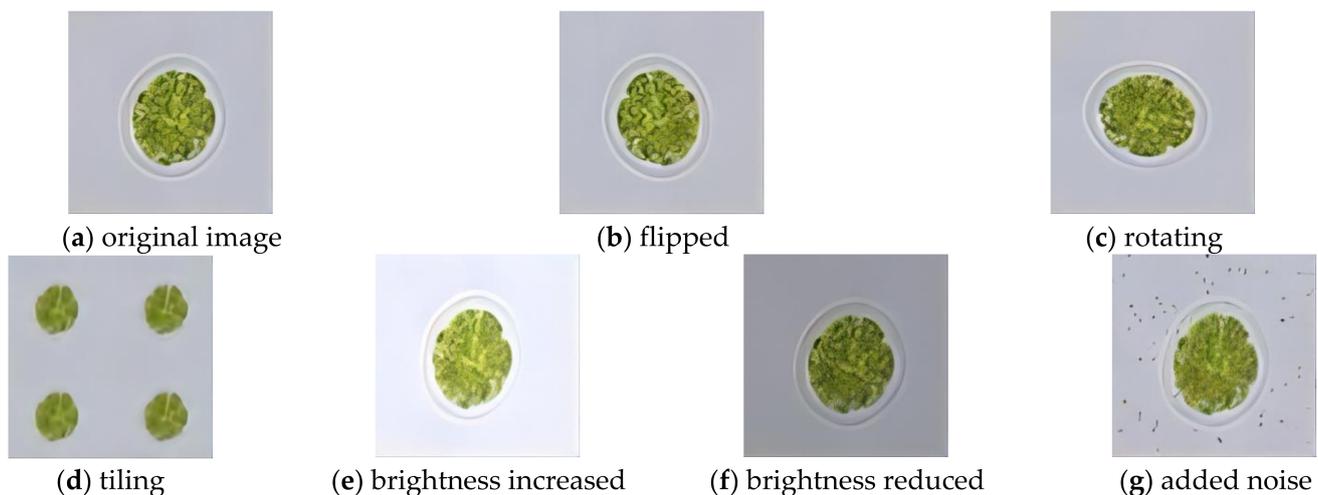


Figure 6. Algae image data broadening effect.

In most cases, the acquired image will be affected by particles due to the influence of acquisition tools or means in the process of image acquisition. The original image is limited by particles under various conditions and often cannot fully reflect all the information of the original image [24,25]. Therefore, it must be processed in the early stage of visual information processing.

By using computer technology, image processing operations such as gray correction and noise filtering are carried out on the original image. Some unnecessary image information is improved by particle suppression on the image data, and useful image information for subsequent processing is enhanced. The dataset contained 150 classes of marine microalgae images, 100 samples of each class, totaling 15,000 images. The main dataset information is shown in Table 1.

Table 1. Main dataset indicators.

Dataset	Marine Microalgae
Total number of data sets	15,000
Number of categories	150
Number of classes	100
Number of training set classes	80
Number of validation set categories	20
Number of test set classes	50
Image size	84 × 84

The overall marine microalgae microscopic image recognition framework is shown in Figure 7:

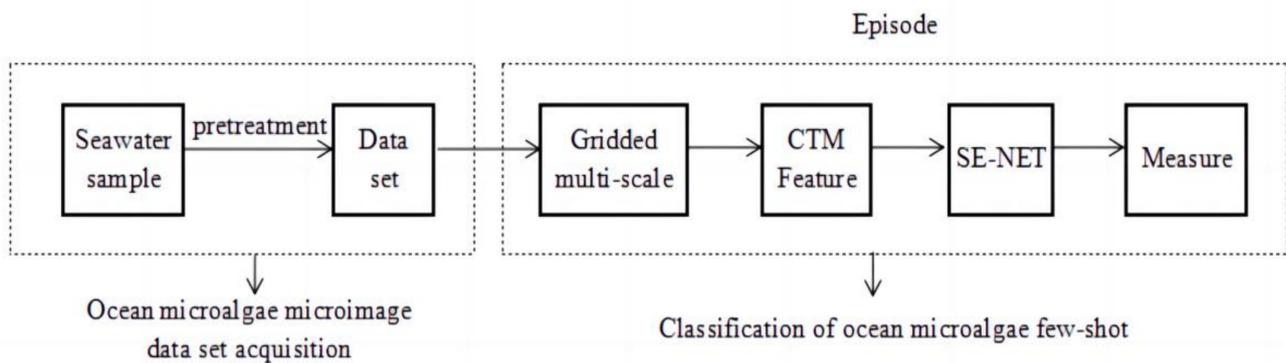


Figure 7. Marine microalgae microscopic image recognition framework.

5.2. Experimental Setup and Method

In this experiment, marine microalgae microscopic images are used as the data set, and the recognition algorithm of few-shot learning is used to identify marine microalgae images. The process is shown below.

Training process: the weight of the backbone network in the model is trained in advance, and then fine-tuned together with other modules in the model to specifically carry out the N-way K-shot few-shot classification task. There are 100 epochs in total, and there are 100 and 500 episodes in each epoch for training and verification. In the test phase, 2000 episodes tasks are randomly selected from the test set to evaluate the average accuracy of their classification. For an episode, N classes are selected, and then support and query are selected for each selected class. The model is trained once with the support and query of several selected classes.

Two scenarios commonly used in few-shot learning were selected for the experiment: 5-way 1-shot and 5-way 5-shot. The scenario of 5-way 1-shot increased the training difficulty due to the small number of support set samples. The backbone network in the model is initialized with pre-trained weights, and then fine-tuned together with other modules in the model. In particular, the situational training strategy is adopted to carry out the N-way K-shot few-shot classification task. The SGD (Stochastic Gradient Descent) optimizer is used to optimize the network. The initial learning rate is set to 100 epochs in total, and the learning rate decays by half after every 10 epochs. Each epoch has 100 and 300 episodes for training and verification.

In the test phase, 1000 episodes tasks are randomly selected from the test set to evaluate the average accuracy of their classification. In the verification and testing stage, whether the input is a whole image or a local image block, the size is adjusted to 84×84 first, and then it is input into the backbone network to extract features.

In each experiment, the proposed method is compared with eight current state-of-the-art few-shot classification methods: MatchingNet [8], ProtoNet [9], RelationNet [10], CTM [16], TADAM [26], DN4 [11], FEAT [27], and DeepEMD [12], and the results of 5-way 1-shot and 5-way 5-shot classification tasks on the marine microalgae dataset are compared, as shown in Table 2.

In Table 2, the proposed method P(1, 2, 3) indicates that the original graph is divided into a total of 14 local blocks of 1×1 , 2×2 , and 3×3 as inputs. The proposed method shows excellent performance on the marine microalgae classification dataset. On the 5-way 1-shot and 5-way 5-shot few-shot classification tasks, the results of the proposed method improved 6.08% and 5.5% over the suboptimal results, respectively.

5.3. Ablation Experiments

To verify the effect of gridded multiscale local features on the classification accuracy of the model and to select the optimal hyperparameters, the following ablation experiments were designed in the marine microalgae dataset using a different number of feature scales.

5.3.1. Multiscale Local Chunking Experiments

In the multiscale local chunking experiment, $P(,)$ indicates that the marine microalgae images are combined in multiple scale chunks, e.g., $P(1,2)$ indicates that the original image is divided into a total of 5 local chunks of 1×1 and 2×2 together as input. A total of 6 feature scales are used for the experiments, which are divided into $P2 \times 2$, $P3 \times 3$, $P(1,2)$, $P(1,3)$, $P(2,3)$, and $P(1,2,3)$, and the experimental results are shown in Table 3.

Table 2. The 5-way 1-shot and 5-way 5-shot classification results on the marine microalgae dataset.

Method	Backbone	5-Way 1-Shot	5-Way 5-Shot
MatchingNet	ResNet-12	65.12 ± 0.22	69.33 ± 0.15
ProtoNet	ResNet-12	66.33 ± 0.85	72.11 ± 0.62
RelationNet	ResNet-12	68.32 ± 0.40	75.19 ± 0.21
CTM	ResNet-12	68.78 ± 0.82	70.24 ± 0.35
TADAM	ResNet-12	58.50 ± 0.23	66.70 ± 0.11
DN4	ResNet-12	64.61 ± 0.51	69.85 ± 0.35
FEAT	ResNet-12	68.78 ± 0.30	72.80 ± 0.19
DeepEMD	ResNet-12	73.11 ± 0.82	81.16 ± 0.66
ours $P(1, 2, 3)$	ResNet-12	79.19 ± 0.45	86.66 ± 0.25

Table 3. Classification results of multiscale local chunking.

Patches	Backbone	5-Way 1-Shot	5-Way 5-Shot
$P2 \times 2$	ResNet-12	76.66 ± 0.31	84.21 ± 0.16
$P3 \times 3$	ResNet-12	76.31 ± 0.52	84.26 ± 0.62
$P(1,2)$	ResNet-12	77.31 ± 0.22	85.01 ± 0.10
$P(1,3)$	ResNet-12	75.68 ± 0.19	86.08 ± 0.82
$P(2,3)$	ResNet-12	75.92 ± 0.60	86.11 ± 0.56
$P(1,2,3)$	ResNet-12	79.19 ± 0.45	86.66 ± 0.25

From the experimental results, it can be concluded that the marine microalgae dataset with $P(1,2,3)$ scale features has the best final recognition effect. Single-scale features can only obtain partial feature information about the image, which is one-sided. However, when using multiple scales for feature extraction, not only the detailed features of the image but also the global feature information of the image can be obtained, which achieves a more accurate information representation of the image in multiple scales and can better improve the classification accuracy of the model.

Although the chunking process can effectively utilize the local image features, when the number of chunks is too large, a complete target may be separated into multiple local chunks, and the integrity of the target is destroyed, which adversely affects the classification. Therefore, the number of chunks needs to be traded off according to the classification results, and more chunks may not be better.

5.3.2. Grid-Based Partial Chunking Experiments

To verify the rationality of the gridded local chunking approach, a comparison experiment is conducted with the local chunking approach using random cropping by comparing the recognition results under the same number of local image blocks, as shown in Figure 8.

The performance of the grid chunking method is better than random cropping for the same number of local blocks. This is because when the image is grid chunked, multiple local block images are stitched together to contain the information of the whole image, and then each local feature is made to have global information by the local feature fusion module. In contrast, the random cropping method can only randomly take part of the local blocks on the original image, and may not be able to obtain the complete global information even through feature fusion, so its classification performance is weaker than that of the grid chunking method.

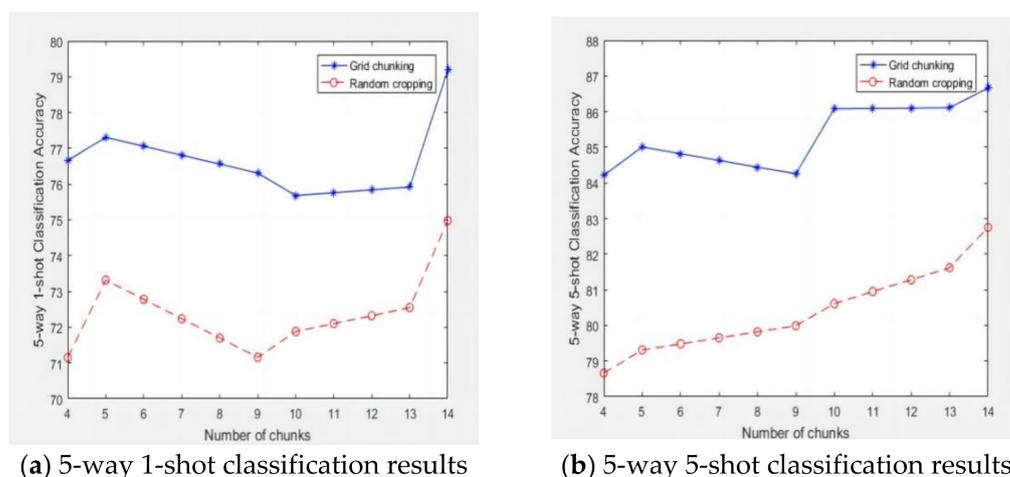


Figure 8. The 5-way 1-shot and 5-way 5-shot classification results on the marine microalgae dataset using grid chunking and random cropping.

6. Conclusions

In recent years, accurate identification of marine microalgae has provided new opportunities for sustainable photosynthetic oxygen release processes. This study is for the current few-shot learning based on the classification and identification of marine algae. In this paper, an algorithm for few-shot classification of marine microalgae based on gridded multiscale local feature fusion has been proposed. In few-shot image recognition, the gridded multiscale local features are obtained by the category traversal module, the local features of algae images are fully utilized, and the enhanced features containing global information are obtained by the SE-NET architecture local feature fusion module, which improves the feature extraction ability and generalization ability of the model. Compared with various classical networks, the proposed method has higher classification accuracy on the marine microalgae dataset, and the ablation experiments also verify the effectiveness of the proposed method. In future work, the development of marine microalgae classification technology will be promoted by enriching the data set of marine microalgae and strengthening model training and learning. For marine microalgae microscopic image data, how to use deep neural networks instead of shallow networks to mine microalgae image hidden feature information is one of the problems to be solved.

Author Contributions: D.L.: conceptualization, investigation, methodology. T.L.: methodology, writing original draft preparation, data curation. H.B. and Y.Z.: investigation, validation. Y.C.: conceptualization, writing review and editing, supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Key Laboratory of Environment Controlled Aquaculture (Dalian Ocean University) Ministry of Education Project (202213), by the Dalian Program on Key Science and Technology R&D Project (2021YF16SN013), and by the Basic Scientific Research Project for Higher Education Institutions of the Education Department of Liaoning Province (LJKMZ20221089).

Data Availability Statement: Due to confidentiality agreements, supporting data can only be made available to bona fide researchers, subject to a non-disclosure agreement.

Acknowledgments: This work was supported by the Key Laboratory of Environment Controlled Aquaculture (Dalian Ocean University) Ministry of Education Project (202213), by the Dalian Program on Key Science and Technology R&D Project (2021YF16SN013), and by the Basic Scientific Research Project for Higher Education Institutions of the Education Department of Liaoning Province (LJKMZ20221089).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hamza, I.; Habibuulah, A.; Ahmed, K.; Chughtai, A.H.; Jiskani, M.A. Geospatial analysis of wetlands based on land use/land cover dynamics using remote sensing and GIS in Sindh, Pakistan. *Sci. Prog.* **2021**, *104*, 368504211026143. [[CrossRef](#)]
2. Abdullah; Ali, S.; Khan, Z.; Hussain, A.; Athar, A.; Kim, H.-C. Computer Vision Based Deep Learning Approach for the Detection and Classification of Algae Species Using Microscopic Images. *Water* **2022**, *14*, 2219. [[CrossRef](#)]
3. Li, Z.; He, J.; Ni, T.; Huo, J. Numerical computation based few-shot learning for intelligent sea surface temperature prediction. *Multimed. Syst.* **2022**, 1–13. [[CrossRef](#)]
4. Liu, D.; Wang, P.; Cheng, Y.; Bi, H. An improved Algae-YOLO Model Based on Deep Learning for Object Detection of Ocean Microalgae Considering Aquacultural Lightweight Deployment. *Front. Mar. Sci.* **2022**, *9*, 1070638. [[CrossRef](#)]
5. Wang, Y.; Cheng, Y.; Bi, H.; Yu, Q.; Liu, D. Learning VGG network model based on the depth of the ocean single-celled algae recognition algorithm. *J. Dalian Ocean Univ.* **2021**, *4*, 334–339.
6. Qiu, Y. Research on Small Sample Image Classification Method Based on Neural Network. Master's Thesis, Lanzhou University of Technology, Lanzhou, China, 2020.
7. Koch, G.; Zemel, R.; Salakhutdinov, R. Siamese neural networks for one-shot image recognition. In Proceedings of the ICML Deep Learning Workshop, Lille, France, 6–11 July 2015; Volume 2.
8. Vinyals, O.; Blundell, C.; Lillicrap, T.P.; Kavukcuoglu, K.; Wierstra, D. Matching Networks for One Shot Learning. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, abs/1606.04080.
9. Snell, J.; Swersky, K.; Zemel, R. Prototypical networks for few-shot learning. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.
10. Sung, F.; Yang, Y.; Zhang, L.; Xiang, T.; Torr, P.; Hospedales, T. Learning to compare: Relation network for few-shot learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 1199–1208.
11. Li, W.; Wang, L.; Xu, J.; Huo, J.; Gao, Y.; Luo, J. Revisiting local descriptor based image-to-class measure for few-shot learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 7260–7268.
12. Zhang, C.; Cai, Y.; Lin, G.; Shen, C. Deepemd: Few-shot image classification with differentiable earth mover's distance and structured classifiers. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 12203–12213.
13. Wang, Q.; Jin, Z.; Li, W.; Li, Q. Few-shot image classification algorithm based on Multi-scale channel attention mechanism. *Hubei Univ. Technol.* **2022**, *37*, 34–39.
14. Xie, J.; Long, F.; Lv, J.; Wang, Q.; Li, P. Joint Distribution Matters: Deep Brownian Distance Covariance for Few-Shot Classification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 7972–7981.
15. Wei, S.; Zhang, J.; Chen, L. A few-shot learning method for adaptive local relational networks. *J. Shenyang Polytech. Univ.* **2021**, *40*, 35–41.
16. Li, H.; Eigen, D.; Dodge, S.; Zeiler, M.; Wang, X. Finding task-relevant features for few-shot learning by category traversal. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 1–10.
17. Cao, Y.; Xu, J.; Lin, S.; Wei, F.; Hu, H. GCNet: Non-local Networks Meet Squeeze-Excitation Networks and Beyond. *arXiv* **2019**, arXiv:1904.11492.
18. Wang, Q.-S.; Wang, F.-S.; Chen, J.-G. Faster R-CNN target detection algorithm incorporating adaptive attention mechanism. *Adv. Lasers Optoelectron.* **2022**, *59*, 1215016.
19. Wang, X.; Zhang, H. Few shot classification network based on attention mechanism and graph convolution. *Comput. Eng. Appl.* **2021**, *57*, 164–170.
20. Zhang, Y.; Zhang, F.; Liu, Y. Improved handwritten digit recognition technique for ResNet network. *J. Xianyang Norm. Coll.* **2021**, *36*, 7–10.
21. Wang, H.; Zhang, M.; Liu, H. A food image recognition method based on improved ResNet network. *J. Shaanxi Univ. Sci. Technol.* **2022**, *40*, 154–160.
22. Wang, R.; Han, M.; Yang, J. Few-shot learning of multi-level attentional feature networks. *J. Electron. Inf.* **2020**, *42*, 772–778.
23. Zhan, Z.; Cheng, D. Text classification research based on LDA and distance metric learning. *J. Nat. Sci. Hunan Norm. Univ.* **2016**, *39*, 70–76.
24. Kramer, S.J.; Siegel, D.A. How Can Phytoplankton Pigments Be Best Used to Characterize Surface Ocean Phytoplankton Groups for Ocean Color Remote Sensing Algorithms? *J. Geophys. Res. Oceans* **2019**, *124*, 7557–7574. [[CrossRef](#)] [[PubMed](#)]
25. Kramer, S.J.; Siegel, D.A.; Graff, J.R. Phytoplankton Community Composition Determined from Co-variability Among Phytoplankton Pigments from the NAAMES Field Campaign. *Front. Mar. Sci.* **2020**, *7*, 215. [[CrossRef](#)]

26. Oreshkin, B.; Rodríguez López, P.; Lacoste, A. Tadam: Task dependent adaptive metric for improved few-shot learning. *arXiv* **2018**, arXiv:1805.10123.
27. Gao, F.; Xu, J.; Lang, R.; Wang, J.; Hussain, A.; Zhou, H. A Few-Shot Learning Method for SAR Images Based on Weighted Distance and Feature Fusion. *Remote Sens.* **2022**, *14*, 4583. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.