IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 21 January 2024, accepted 16 February 2024, date of publication 14 March 2024, date of current version 20 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3375928

RESEARCH ARTICLE

Hybrid Improved Concave Matching Algorithm and ResNet Image Recognition Model

WENDA ZHENG, YIBO AI, AND WEIDONG ZHANG
National Center for Materials Service Safety, University of Science and Technology Beijing, Beijing 100083, China
Corresponding author: Weidong Zhang (zwdpaper@163.com)

ABSTRACT With the increase of organic matter content in global water bodies, timely and accurate identification of algae objects in water bodies is of great significance for rapid treatment of water eutrophication. To improve the performance of seaweed microscopic image recognition model, an improved concave matching algorithm was designed and applied to the segmentation of original microscopic images. In addition, an IResNet algorithm with self-adjusting pooling layer is designed. In this study, an improved concave matching algorithm and an improved ResNet algorithm were used to construct a microscopic algae image recognition model. The average Accuracy and Precision of the proposed algorithm in the data set are 88.6% and 90.3%, which is significantly higher than all tr4aditional and mainstream image segmentation algorithms. Among the five modified ResNet algorithms with self-adjusting pooling layers designed in this study, the fifth modified scheme has the best performance on the training set. The model's recognition accuracy and accuracy corresponding to the improved ResNet algorithm on the whole test set are 84.1% and 82.8%, respectively. The experimental data on the real data set show that the algae microscopic image recognition model has certain practical potential.

INDEX TERMS Concave point matching algorithm, neural network, ResNet, algae microscopic image.

I. INTRODUCTION

Recently, Image Recognition (IR) technology has achieved unprecedented development [1]. However, despite significant progress, the IR still face many challenges [2]. At present, the main technical routes of IR include template-based models, feature-based models and Deep Learning (DL)-based models [3], [4]. Template-based models are simple and intuitive, but error-prone when dealing with complex images. Feature-based methods perform well in processing complex images, but the feature extraction process is complex and susceptible to noise. The model with DL can automatically learn the deep features of images, but it needs a quantity of annotation data and computational resources. The issues addressed in this study are as follows: due to the small recognition objects, small pixel differences between algae and the environment, and tight contour of multi algae, precise identification of algae objects in algae images is difficult, which is not conducive to the development of related work such as algae management. So the purpose of this study is to design a method that can more accurately recognize microscopic images of seaweed. The key limitation in this issue lies in how to effectively segment algae pixels from non-algae pixels in the image. ResNet is a IR model with DL, which solves the problem of gradient disappearance in deep networks by introducing residual structure and improves the accuracy of IR [5]. However, the computational complexity of ResNet is high and requires a lot of computing resources. Concave point matching algorithm is a feature based image segmentation method, which identifies and segments images by finding concave points in images and matching similar concave points. Concave point matching algorithm has low computational complexity, but it is easy to make mistakes when processing complex images. To take into account the accuracy and efficiency of IR, a hybrid improved concave matching algorithm and ResNetIR model are proposed in this study. This model combines the low computational complexity of concave matching algorithm with the high recognition

The associate editor coordinating the review of this manuscript and approving it for publication was Yi Zhang.

accuracy of ResNet, aiming to enhance the algorithm's efficiency and accuracy, thus achieving the goal of effectively identifying algae in water bodies.

This study combines the improved concave point matching algorithm with the ResNet algorithm to develop a new type of algae microscopic image recognition model, effectively solving the difficulties encountered in the recognition of small and tightly arranged water algae images. This not only improves the technical level of accurate segmentation and identification of algae, but also contributes a scientific methodology with practical application potential to fields such as water monitoring and eutrophication treatment. In addition, our research findings have expanded the theoretical framework of combining feature-based image processing algorithms with deep learning models, providing new directions for future research in related fields.

The structure of the study is as follows: The first part describes the basic concepts and main technical routes of IR; In the second part, the improved ResNetIR model, which combines the improved concave matching algorithm and five self-adjusting pooling layer schemes, is established minutely. Thirdly, the model is verified by experiments and compared with other models. In the last part, the thesis is summarized and the future research direction is prospected.

II. RELATED WORKS

IR technology is popular in security protection, the Internet of things and other fields, can save a lot of human labor, and promote productivity. Gao P et al. constructed a multi-dimensional data model for video IR and motion capture with DL framework to improve the accuracy of human motion recognition in small and medium-sized video. First, the model extracted the moving foreground of the target and used the gradient histogram to identify human body. The proposed model fused dense trajectory features and DL features. In DL features, the model took the fusion of deep video features and video primary color features as the features of DL. Finally, the algorithm was verified that it possessed better recognition performance for both large and small-scale gesture. Aside from that, the experimental data were classified using the Imperial Computer Vision and Learning Laboratory Human Behavior dataset. The connection between this model and previous similar studies is that they all use a deep learning framework, but the disadvantage is that the model has high computational complexity and is not suitable for real-time computing tasks [6]. Liang J et al. pointed out that although the mark-semantic features in GCN showed better performance in the overall visual representation of images, they were rarely applied to local areas of images. Therefore, this paper attempted to use the GCN to learn the global and local features of the image, and to balance the two features. In this study, a multi-scale attention model was constructed, which used feature mappings of various sizes to achieve the extraction of global features. Semantic oriented attention module enabled semantic association between tags in the network. This model achieved classification accuracy of 83.4% and 94.2% respectively on two common IR datasets, which is a good recognition method [7].

Singh G believed that in some cases, DL models could also be applied to identify viruses from medical images. This was especially true in cases where patients were tested with X-rays and or nuclear magnetic scans. However, how this DL model works was crucial. In this study, a DL model with a forward reasoning process was proposed. After being trained and tested, the model gave an accuracy of more than 99% [8]. Zhang H et al. found that attention mechanisms performed well in many computer vision tasks. Therefore, an effective multi-path attention mechanism was proposed. Compared with other attention mechanisms that utilize global pooling, the main advantage of multi-path attention mechanism was that it took into account both the correlation of feature graphs and the structure information of different scales. The backbone representation was enhanced by adding a multi-path attention mechanism horizontally across two separate and separate dimensions, channel and space. Since only one simple and uniform computational block was generated, the multi-path attention mechanism could be flexibly integrated into various CNNS. With few parameters, end-to-end training could be performed together. This model was superior to various existing models [9]. Hu W J et al proposed a generative data enhancement model. Whether in agriculture or healthcare, the model outperformed the baseline model in image generation and data enhancement. The identification error rates were reduced by 4.73% and 19.59%, respectively. Meanwhile, using this model for data enhancement could significantly improve the image classification ability of the specified industry. The average accuracy of classification results on agricultural and medical images were improved by 0.96% and 1.27% over baseline [10]. Sun C et al. found that by recognizing sonar images of Marine sediments, textures in the images can be classified. The author constructed an improved CNN algorithm by combining K-means algorithm and transfer learning idea. The average accuracy of the model based on transfer learning achieved 84.39%, which was 0.97% higher than the original mode [11].

Y. J. Choi et al. designed an artificial intelligence based recognition algorithm to address the low efficiency of artificial orthodontic diagnosis. The special contribution of this algorithm compared to similar algorithms in the past lies in its integration of semi-automatic 3D visualization therapy target technology, which effectively improved the recognition accuracy in application scenarios. However, the design method also had the disadvantage of narrow applicability [12]. X. Gou et al. were influenced by the biological heuristic model developed by previous scholars and designed a new universal distributed biological heuristic model. The test results showed that the recognition accuracy of this model on iris image datasetreached95.87%, which was higher than the algorithm before improvement. However, therewas still a disadvantage of slow computation speed [13]. G. H. Lan et al. used augmented reality technology to adjust the input data of the image recognition model to reduce

image distortion and pixel loss. In the evaluation testing of multiple product datasets, it was found that the comprehensive image recognition model designed in this study had better recognition ability for low-quality images than traditional models [14]. Y. Liu et al. found that dimensionality reduction of high-dimensional nonlinear structural images as the key to improving recognition rate. To achieve the ideal effect of high-dimensional nonlinear image recognition, the author team analyzed traditional dimensionality reduction algorithms and refined their advantages, and proposed an image recognition technology based on nonlinear dimensionality reduction methods. The test results showed that the image recognition technology proposed by the author had a recognition recall rate of 98.24% in product image recognition, which washigher than traditional recognition technologies [15]. K. Ohri and M. Kumar believed that self supervised learning could effectively reduce the costs associated with obtaining labeled datasets. Therefore, they compared various self supervised image recognition technologies and summarized their applicability, advantages, and disadvantages, providing a reference for future researchers to choose self supervised learning methods [16]. M. Li et al. found that the task of recognizing few shot images still has shortcomings such as insufficient recognition accuracy and insufficient training data. Therefore, the author used a weighted imprinting strategy to design an improved method for recognizing few shot images. The test results showed that the designed method has significantly higher recognition accuracy than traditional methods in small shot image recognition tasks [17].

In the above studies, researchers have designed targeted image recognition models based on their own research problems. However, none of the 12 studies in the field of image recognition mentioned above mentioned image segmentation processing. Image segmentation is of great significance for recognizing complex scenes and multi object images, as it can enhance the application range of image recognition algorithms and expand application scenarios. Moreover, research specifically for precise recognition of microscopic images of seaweed is still quite scarce. The above two points are also the key knowledge gaps between previous research and this study. At the same time, conducting precise identification research on seaweed microscopic images also has certain reference value for improving the efficiency of human treatment of eutrophication problems in water bodies.

III. ALGAL IR MODEL DESIGN BASED ON MIXED CONCAVE MATCHING AND ResNet ALGORITHM

At present, most algae detection tasks in water are carried out by manual observation through microscopes, which is time-consuming, inefficient and heavily dependent on the personal experience of observers. Although the neural network has a good performance in the IR task, the recognition effect of the neural network algorithm is poor in the algae recognition task in the water microscopic image because of the special shape of the recognition object. Thus, an enhanced algal IR model is constructed.

A. DESIGN OF IMAGE SEGMENTATION MODEL BASED ON IMPROVED CONCAVE MATCHING ALGORITHM

After analyzing the research background and previous literature results, beginning to design the algae microscopic image recognition model proposed in this study. Before algae IR, it is necessary to segment the local image containing algae in the image to decline the interference of non-core elements in the image. Considering the characteristics of algae IR, the image segmentation task is split into the following three steps. Firstly, use the erosion and dilation techniques common in morphological operations to initially segment the slightly adherent algal cells, which can effectively reduce the number of adherent algal cells that need to be treated later. Next, the improved concave matching algorithm is used for image segmentation. The improvement of the algorithm has two aspects: First, the matching strategy is changed from the global matching of the original algorithm to the nearest neighbor satisfying concave matching with greedy strategy, so as to achieve almost no impact on the matching accuracy and greatly reduce the complexity of the model, so as to significantly enhance the operation efficiency. Second, when three adherent algal cells remain within a certain range and have three concave points, the central point of the three concave points is found and segmented as the segmentation center, so as to reduce the segmentation defect of round algal cells and improve the segmentation quality [18]. Finally, the method of distance transformation function is introduced to solve the problem of isolated pits, and the adherent algae cells are separated by the shortest distance [19], [20]. For these three tasks, an image segmentation model based on the improved concave matching algorithm is designed, and its calculation flow is shown in Figure 1.

It is required to obtain the concave data of the adhered region on the image. However, the edge information of algal cells must be extracted before the concave data can be obtained [21], [22]. In this study, three kinds of operators, namely first-order differential operator Sobel, second-order differential operator Laplacian and non-differential edge detection operator Canny, are used to extract the contour information [23].

Sobel operator performs well in processing images with gradually changing gray level and more noise, and is also very accurate in edge location. The calculation method of the horizontal edge detection image amplitude G_a is shown in Equation (1).

$$G_a = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * f\,(a, \mathrm{b}) \tag{1}$$

In Equation (1), $f\,(a, b)$ is the data expression of the image. Generally, to reduce the calculation burden, the operation method of edge detection image amplitude G_b in the vertical

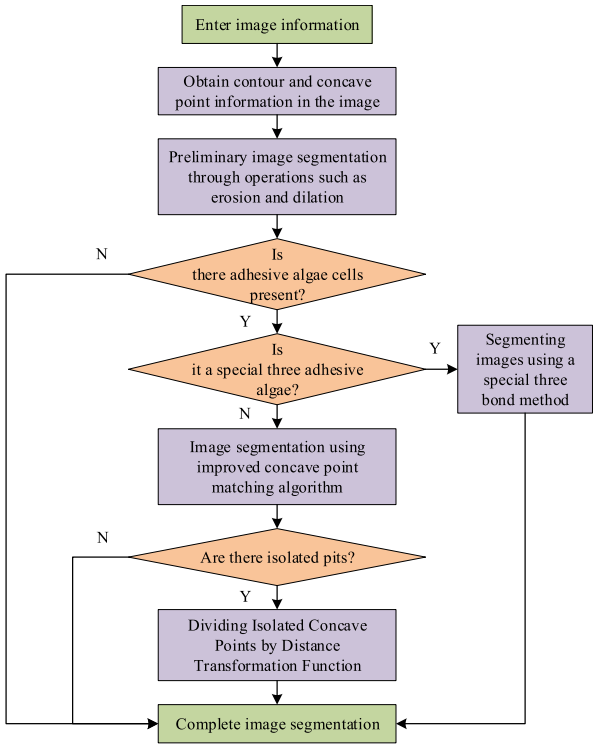

FIGURE 1. Calculation process of image segmentation model based on improved concave point matching algorithm.

direction is shown in Equation (2).

$$G_b = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * f(a, b) \tag{2}$$

Then the edge amplitude G is calculated according to Equation (3).

$$G = \sqrt{G_a^2 + G_b^2} \tag{3}$$

However, to reduce the computational complexity, this study chooses to replace Equation (3) with a more concise approximation, as shown in Equation (4), this is also one of the improvement points of the segmentation algorithm proposed in this study.

$$G = |G_a| + |G_b| \tag{4}$$

Or calculate G as in Equation (5):

$$G = max(G_a, G_b) \tag{5}$$

In order to simultaneously consider G_a and G_b, equation (4) was chosen for calculating the edge amplitude in this study. After using the Sobel operator, Laplace operator, and Canny operator to extract the edge information of algae cells, as the concave points are the points where the contours of different individual algae cells intersect, it is also necessary to extract the concave point information from the contour information in order to perform the concave point matching algorithm required for subsequent adhesive algae

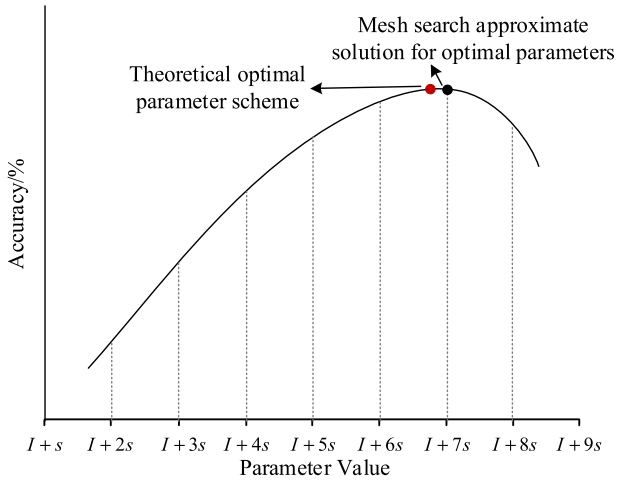

FIGURE 2. Demonstration of the computational principle of grid search method for finding the optimal parameters.

segmentation. Considering the small amount of computation, the circular template was used to detect the concave points of algal cell contour. Meanwhile, to filter out the possible small impurity profile, the circular template is described according to Equation (6). This is also a separate addition to the image segmentation model designed this time.

$$C_i = |A_i| / N_R \tag{6}$$

In equation (6), C_i represents the ratio of the internal arc length pixels of a circular template centered on point i to the total number of pixels in the circular template. Where, $|A_i|$ indicates that point i is the center of the circle on the outline; The number of pixels in the part of the arc of the circular template with radius R located inside the algae; N_R is the number of pixels forming A circle with a radius of R pixels; According to the grid search method, the corresponding R value of the optimal result is calculated as 6px. The specific calculation process is as follows: Since the unit of radius R here is pixels, the spacing of the grid is taken as 1. Since the concave points to be searched exist in the algae microscopic image, the optimization range is set to 2px∼75px according to the size of the algae microscopic image in the dataset. Calculate in sequence when R takes values of 2, 3 At 75 o'clock, the ResNet algorithm has an accuracy in identifying microscopic images of algae in the dataset. Finally, the corresponding parameter value with the highest accuracy is the set value of parameter R, which is 6px. The calculation principle of the grid search method can be described in Figure 2. In Figure 2, I and s respectively represent the allowed minimum values with solving parameters and grid spacing.

C_i represents the ratio of the inner arc length pixels of the circular template with point i as the center of the circle to the total pixels of the circular template. According to the results of multiple tests, the value of parameter α is selected as 0.55. If C_i is greater than α, C_i is a suspected concave point.

After successfully extracting the contour concave information, it is necessary to accurately match the corresponding

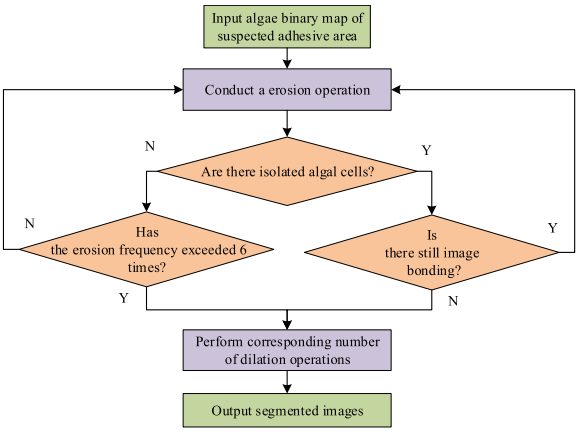

FIGURE 3. Flow chart of segmentation method for slightly adhered algae.

concave point of each concave point, so as to realize the correct segmentation of algal cells. For the slightly adhering areas in algae images, the morphological operations of erosion and expansion in image processing are used as the segmentation method for slightly adhered algae cells. The specific application steps and segmentation process of these two methods are shown in Figure 3.

Segmentation of slightly adhering algal cells is achieved by detecting whether the number of concave points in the profile is reduced. Multiple experiments have found that less than seven erosion operations are enough to separate slightly adhered algal cells. After the single algal cells were extracted, the morphology of algal cells was reduced to the greatest extent by the same number of expansion operations. The improved global concave matching algorithm is redesigned. The calculation flow of traditional global concave matching algorithm in segmentation task is shown in Figure 4.

In Figure 4, in an algal image containing N pits to be matched, the distance of $N(N-1)$ sub-pits needs to be calculated. However, through analyzing a large number of microscopic images of the adherent algae, it is found that the distance of the concave pairs that can be correctly matched is generally within the diameter of the circular algae or the length of the strip algae of the adherent algae. This finding was confirmed in the second judgment. Therefore, global concave matching can be changed to local concave matching, thus eliminating the judgment of the second step and reducing the number of initial distance calculation. In an algal image containing N concave points to be matched, if M points meet the local range requirements ($M \leq N$), then only the distance of $M(M-1)$ sub-concave points needs to be calculated. Further, it is not actually necessary to calculate the distance between each concave point and the others at the beginning. If you can calculate the distance dynamically, you can further reduce the number of calculations. In an algal image containing N concave points to be matched, M points meet the local range requirement ($M \leq N$). In the optimal case, only the $M(M-1)/2$sub-concave distance needs to be calculated. In summary, Figure 5 illustrates the

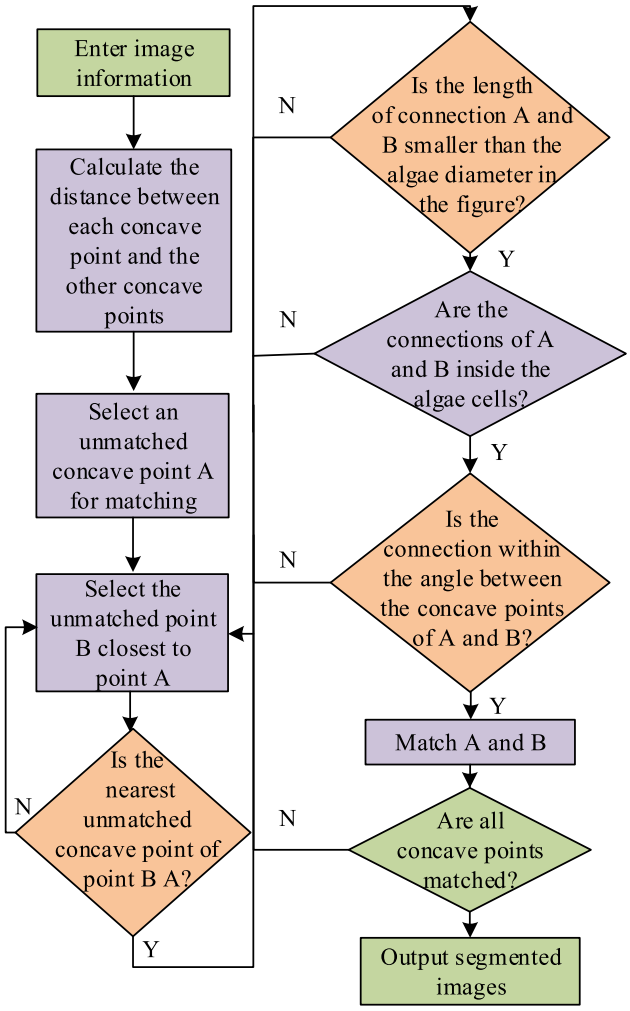

FIGURE 4. Image segmentation process using traditional global concave point matching algorithm.

image segmentation process of the improved concave point matching algorithm.

As shown in Figure 5, in the improved concave point matching algorithm, the first step is to select an unmatched concave point A and calculate the distance between the unmatched concave points in a certain range around the point. The second step is to select the nearest point B and perform the steps performed at point A in the previous step. The third step is to determine whether the shortest distance between A and B is the same. If they are the same, go to the next step. If the result is different, define B as A new point A and repeat this step. The fourth step is to determine whether the connection between two points A and B is inside the algal cell and whether the connection is within the Angle between two concave points. If either is judged to be negative, return to the previous step. If the result of both questions is ''yes'', proceed to the next step. The fifth step is to match point A to point B. The sixth step is to judge whether all concave points have been matched. If the result is no, return to the first step. Otherwise, the algorithm is finished.

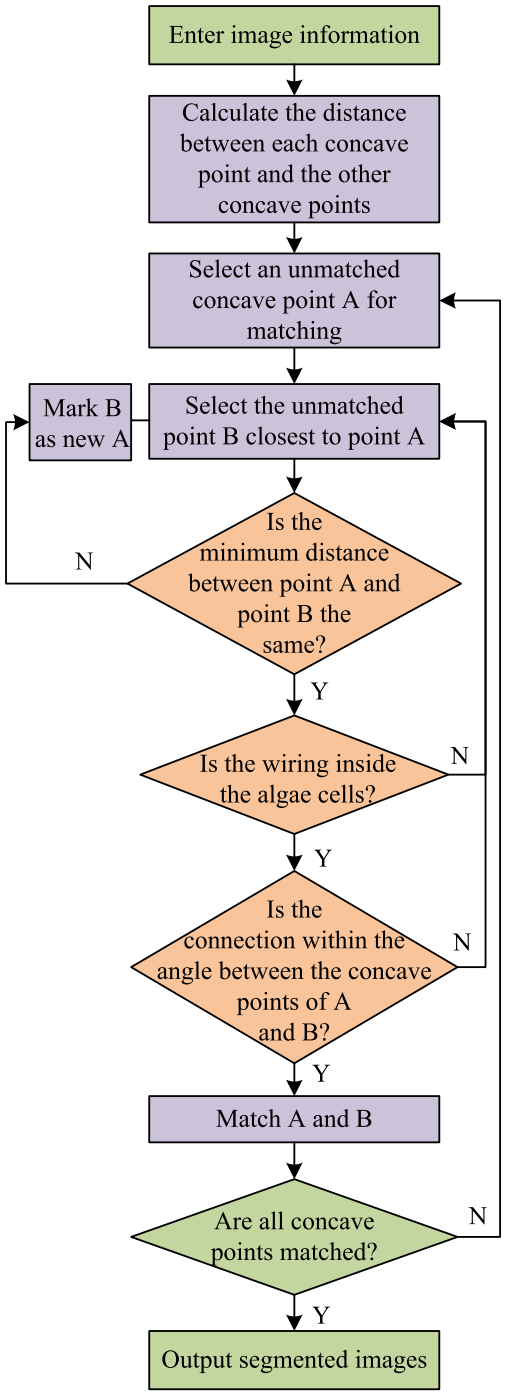

FIGURE 5. Improved concave point matching algorithm image segmentation process.

Then, an image segmentation method based on distance technology content is selected to deal with the problem of isolated concave points. In the traditional processing method, the mapping transformation is processed according to the euclidean distance, see Equation (7).

$$Dist_o(i,j) = \left((x_i - x_j)^2 + (y_i - y_j)^2\right)^{1/2} \tag{7}$$

In Equation (7), $Dist_o(i,j)$ represents the euclidean distance of two points i and j, whose coordinates are (x_i, y_i) and (x_j, y_j).

Although the euclidean distance mapping is more accurate, due to its high calculation cost, the Manhattan distance transformation with lower calculation cost is chosen here. The calculation method is shown in Equation (8), the change in this calculation method also helps to improve the accuracy of image segmentation.

$$Dist(i,j) = |x_i - x_j| + |y_i - y_j| \tag{8}$$

In Equation (8), $Dist_c(i,j)$ represents the Manhattan distance from which two points i and j are processed.

B. DESIGN OF IR MODEL BASED ON IMPROVED ResNet ALGORITHM

The seaweed microscopic image recognition model designed in this study is divided into two parts: image segmentation and image recognition. The content designed above focuses on image segmentation, and the image recognition module is designed below. Traditional neural network algorithms such as Visual Geometry Group (VGG) convolutional neural network algorithm are commonly used for microscopic IR by predecessors. Although the accuracy of the typical VGG-19 network model for algae microscopic image classification is not ideal, the classification accuracy can be further improved by deepening the depth of the network. The depth of the network is a key factor in achieving efficient networks. From AlexNet's 7 layers to VGG's 16 or even 19 layers to GoogleNet's 22 layers, the depth of the Convolutional neural network (CNN) classification network has improved significantly. However, deepening the depth of CNN network and building a deeper model will bring about the gradient dispersion problem, which will make the network more and more difficult to converge. Although this problem can be avoided by some complex methods, such as network initialization standardization, data standardization, and middle layer standardization. However, the increasing network layers also lead to the decrease of the accuracy of the training set. Given the limitations of these problems, the depth of VGG models can only reach a maximum of 19 layers, and the classification performance of VGG models beyond 19 layers has begun to decline. To further improve classification accuracy, it is necessary to choose a model that can solve the problem of classification accuracy decline caused by too deep network depth. ResNet achieves deeper network depth through the "block" convolutional structure, and ResNet50 is such a ResNet network model with 50 layers, which is more widely used. Compared with the subsequent improved CNN neural network, its structure is not too complicated to affect the training speed, so ResNet50 is chosen as the basic method to build the algal IR model in this study.

In ResNet50, if the input and output dimensions of the block are consistent, the calculation process of the identity mapping part will become quite simple, and the output can be obtained by directly adding the input and the output of the

convolutional layer, as shown in Equation (9).

$$Y = F(x) + x \tag{9}$$

In Equation (9), Y represents the output of the block; x represent input information; $F(x)$ denotes the result after intra-block convolution. The identity mapping of bottleneck blocks with different input and output dimensions requires 1×1 convolution to adjust the dimension of the feature graph so that they can be added.

$$Y = F(x) + W(x) \tag{10}$$

In Equation (10), $W(x)$ represents the result after 1×1 convolution. Next, the rest of the main components of ResNet50 are redesigned. First, the output calculation process of neurons is shown in Equation (11) and Equation (12).

$$z_i^{(l+1)} = w_i^{(l+1)} \tilde{y}^{(l)} + b_i^{(l+1)} \tag{11}$$

In Equation (11), $z_i^{(l+1)}$ denotes the total input corresponding to the i neuron in layer $l+1$; $w_i^{(l+1)}$ denotes the weight coefficient of neurons; Where $\tilde{y}^{(l)}$ denotes the neuron after the random discard of layer l; $b_i^{(l+1)}$ denotes the bias coefficient of the corresponding neuron. The result activation process is also required before output, as shown in Equation (12).

$$y_i^{(l+1)} = f\left(z_i^{(l+1)}\right) \tag{12}$$

In Equation (12), $y_i^{(l+1)}$ denotes the corresponding prediction result of the neuron; $f(\cdot)$ denotes the activation function. To prevent over-fitting of the neural network, it is also required to add regularization terms. Common regularization methods include one-norm regularization L_{re1} and two-norm regularization, and the calculation method of one-norm regularization is shown in Equation (13).

$$L_{re1} = L + \lambda \|\theta\| \tag{13}$$

In Equation (13), θ denotes the parameter; λ denotes the weight attenuation coefficient; L denotes the loss function. The calculation method of bi-norm regularization is shown in Equation (14).

$$L_{re2} = L + \lambda \|\theta\|^2 \tag{14}$$

In Equation (14), since binomial regular terms can suppress parameter sizes, model sparsity can also be avoided. Therefore, this paper chooses this method to design the neural network. Let the optimal network parameter found be θ^*, so that equation (15) exists.

$$\theta^* = \arg \min_{\theta} L \tag{15}$$

In this paper, gradient descent method is applied. The derivative of L is needed first to obtain the gradient. The model is then updated using gradients until the model has completed convergence.

$$\theta_{j+1} = \theta_j + \frac{lr \cdot \partial L}{\partial \theta_j} \tag{16}$$

In Equation (16), θ_j denotes the model parameter when iteration to j; $\frac{\partial L}{\partial \theta_j}$ denotes the calculated gradient; lr denotes the learning rate of the neural network. Although the classification performance of ResNet50 exceeds that of the VGG family of algorithms. However, in the actual application scenario of algae microscopic IR model, the zoom ratio of each image is different, which has little effect on algae whose length and width are nearly equal. However, for the large gap between the length and width of the strip algae, the algorithm's recognition ability will cause obvious negative effects. Therefore, it is necessary to design an improved ResNet50 algorithm which can deal with different algae image sizes adaptively.

For the convolutional portion of the CNN network, enter a picture of size (w, h). Using a convolution kernel with size (a, b) and default step size 1, a feature map of size $(w-a+1, h-b+1)$ is obtained. It can be seen that the convolution layer has no requirement on the image size, and can generate feature maps of any size without fixing the input image size. Fully connected layer neurons can set input and output parameters at will before the network is determined, but once the network structure is determined, fully connected layer neurons are fixed, each corresponding to a feature. Thus, the limitation of fixed input size comes entirely from the fully connected layer at the end of the network. To realize the network that can adapt to the image input of different sizes, it is necessary to maintain the same feature map parameters to the full connection layer when the input image size is different. The corresponding portion to be improved is located at the last pooling layer, between the last ResNet50 bottleneck block and the fully connected layer. The goal of the improvement is to output fixed-length features to fully connected layers that require fixed inputs. As long as a structure is designed so that no matter what the size of the input feature map is, a fixed number of parameters can be output to match the subsequent fully connected layer network, the function of automatically adapting to different size image input can be realized. From the position point of view, the corresponding position of this layer is the pooling layer. From a functional point of view, this layer requires dimensionality reduction of input features, so this structure is named self-adjusting pooling layer in this study.

Here, the research team first consider three relatively simple self-adjusting pooling layer structure schemes, as shown in Figure 6. Figures 6 (a), 6 (b), and 6 (c) represent the unified average pooling layer structure scheme, classification pooling structure scheme, and circular ring structure pooling scheme, respectively. Assume that the input is 2048 X x Y size image data. Scheme 1 is the simplest, regardless of the size of the input feature map of the adaptive pooling layer, it will uniformly adopt the average pooling mode to pool the corresponding size, so that the output is a parameter.

Compared with scheme 1, which simply averages the feature maps into one parameter, scheme 2 further considers to extract features from the feature maps of the input adaptive pooling layer more comprehensively. The adaptive pooling

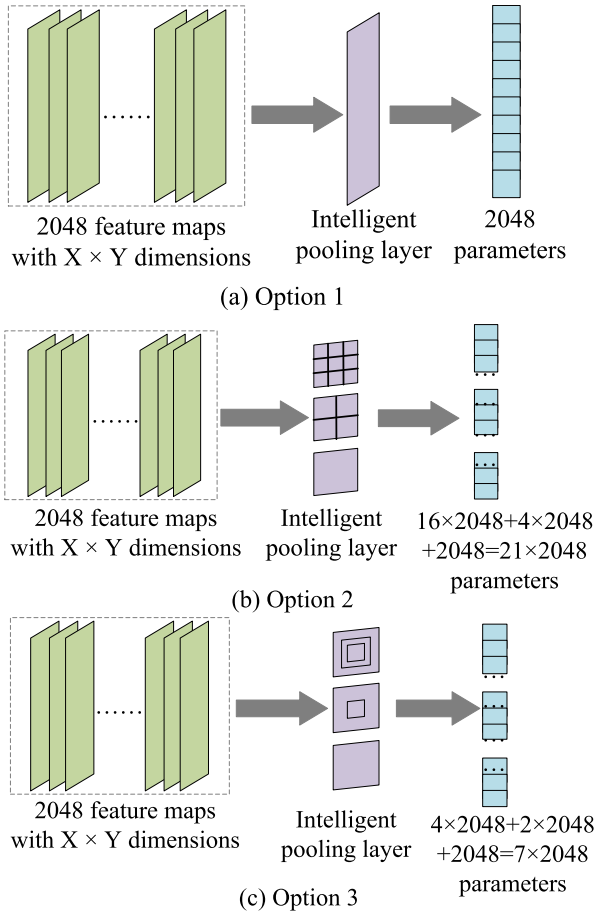

FIGURE 6. Three structural schemes of self Retention basin.

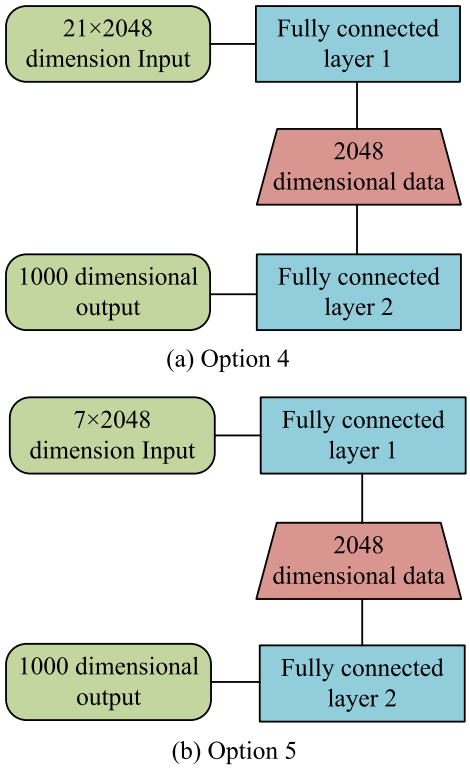

FIGURE 7. Structure comparison of scheme IV and scheme V of self retention basin chemical layer.

layer in scheme 2 is composed of three pooling structures: 1×1, 2×2 and 4×4. The input 2048 X×Y size feature maps are pooled; The output features are concatenated together, and there are $(1+4+16) \times 256$ parameters. Due to the pooling approach of this scheme with high-dimensional output vectors, the inputs of the original ResNet FC-1000 fully connected layer need to be modified accordingly, changing the original 2048-dimensional input to 21× 2048-dimensional input. The pooling structure in Option 2 May ''fragment'' the overall structure of the feature map. Although mitigated by the pooling structure of 1×1, it may have less impact on multi-target IR. However, considering that this chapter needs to identify the category of a single algae microscopic image, the whole image is a complete algae, which may have a greater adverse impact on algae feature recognition extraction. Therefore, modify the pooling structure in scheme 2, change it from ''block'' -shaped pooling structure to ''ring'' -shaped pooling structure, and form scheme 3. Due to the pooling approach of this scheme with high-dimensional output vectors, the original ResNet's fully connected layer input needed to be modified to accommodate this scheme, changing the original 2048-dimensional input to a 7× 2048-dimensional input.

Based on Plan 2 and Plan 3, Plan 4 and Plan 5 are formed respectively, which are similar in structure.

Figures 7 (a) and 7 (b) are both modified pooling layer structures based on Figure 6. ObservingFigure 7, the adaptive pooling layer of Scheme 4 is exactly the same as that of scheme 2. The reason for adding this scheme is to consider that in Scheme 2, the output of the adaptive pooling layer is 21 × 2048 parameters. If only one fully connected layer is used, its ability to learn features may be weak, for the same reason that Scheme 5 is designed.

Thus, the entire algae microscopic IR model can be obtained, and its overall operation process is shown in Figure 8. The core part of the whole model is algae image extraction and algae recognition, which are completed by using improved concave matching algorithm and improved ResNet50 algorithm respectively. The improved concave matching algorithm and the improved ResNet50 algorithm designed in this study are both supervised learning algorithms, and subsequent performance tests are also conducted on supervised data sets.

IV. PERFORMANCE TEST OF IR MODEL BASED ON HYBRID CONCAVE MATCHING AND ResNet ALGORITHM

Two experiments need to be designed and carried out to test the performance of the algae microscopic IR model designed by the hybrid concave matching and ResNet algorithm. One is the performance verification experiment of the image segmentation model based on the improved concave matching

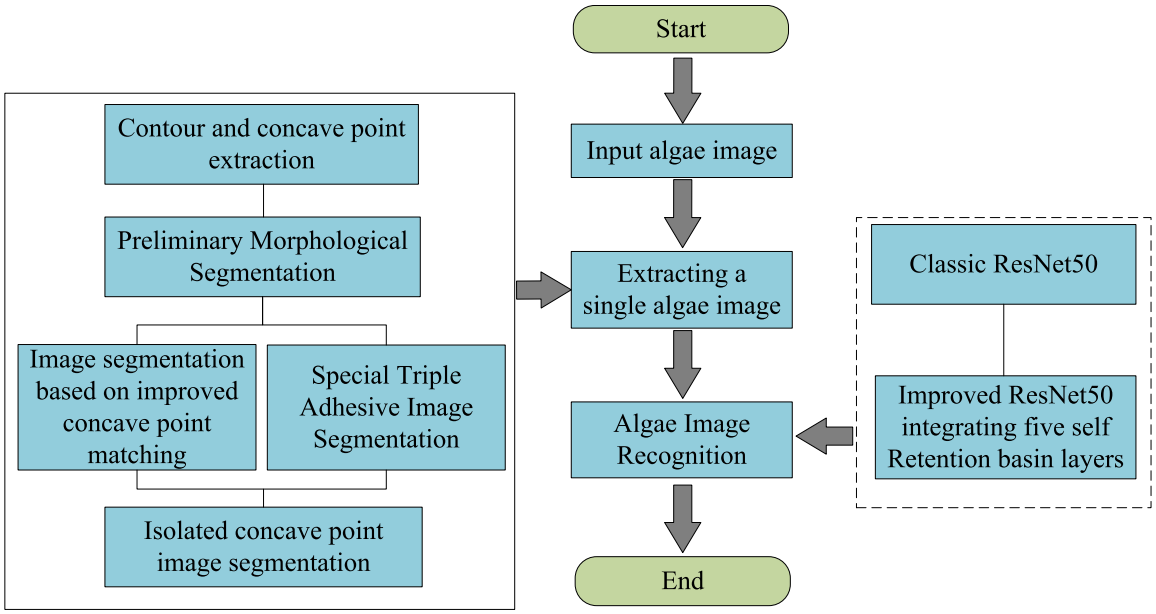

FIGURE 8. Operation process of improved algae microscopic IR model.

algorithm, and the other is the verification experiment of the whole IR model. Both experiments were conducted using Algae micro-image dataset Alalgae 6 established by previous researchers. The dataset contained six algae species and one impurity, with 14,000 images for each category. The data set was split into the training set and the test set in a 7:3 ratio.

A. PERFORMANCE TEST OF CONCAVE MATCHING IMAGE SEGMENTATION MODEL

Now choose morphology algorithm, watershed algorithm, classical Concave point matching (CPM) algorithm, Concave point matching for improving matching methods, CPM-IMM) algorithm, Concave point matching for improving distance calculation (CPM-IDC) algorithm, Meanwhile, the concave point matching (CPM-IMM-IDC) algorithm of matching method and distance transform method is improved to carry out image segmentation experiments. The reason for choosing the watershed algorithm as one of the comparative algorithms is that it is a commonly used image segmentation method that divides image regions based on features such as grayscale values, edge information, and gradients between regions. However, the watershed algorithm also has drawbacks such as high requirements for foreground pre-processing and insufficient universality for image shape and structure. Choosing it as a comparison method can improve the scientificnature of the comparison results. F-test was used to verify differences between data groups.

Here, segmentation integrity rate, segmentation Accuracy, segmentation Precision and calculation time are selected as performance evaluation indexes.

Here, segmentation integrity rate, segmentation Accuracy, segmentation Precision and calculation time are selected as performance evaluation indexes. After the experiment is completed, firstly, the segmentation quality of the segmentation algorithm is tested based on the supervised data set. Specifically, the accuracy and precision rate of each algorithm on the data set are counted, as shown in Figure 9. Figures 9 (a) and 9 (b) represent the accuracy and segmentation accuracy statistics of each algorithm after calculation on the test set, respectively. In Figure 9, the horizontal axis represents the images of different algae types; Different ICONS represent different image segmentation algorithms. It can be seen from FIG. 8 that the recognition accuracy of different models on different types of algae images is different. However, on the whole, the CPM-IMM-IDC algorithm has the highest value in the two indexes; Secondly, two CPM algorithms are partially improved. The traditional CPM algorithm has a lower value, and the morphological algorithm has the smallest value. The mean Accuracy and Precision values of various algae image based on Watershed, Morphology, CPM, CPM-IMM, CPM-IDC, CPM-IMM-IDC algorithms were 52.2%, 26.5%, 69.5%, 79.6%, 81.6%,88.6, respectively % and 44.8%, 28.5%, 64.2%, 76.0%, 83.8%, 90.3%. From the perspective of factors affecting model performance, the data shown in Figure 9 also belongs to the results of ablation experiments. And the recognition accuracy of the concave point matching algorithm, which improves the matching method and distance change method separately, is higher than that of the original concave point matching algorithm in various types of micro-algae microscopic images, while the concave point matching algorithm with both improved points has the highest recognition progress. In CPM, the matching method and distance change method are both factors that affect the performance of the model. The former can more accurately filter out useless features, while the latter can distinguish the

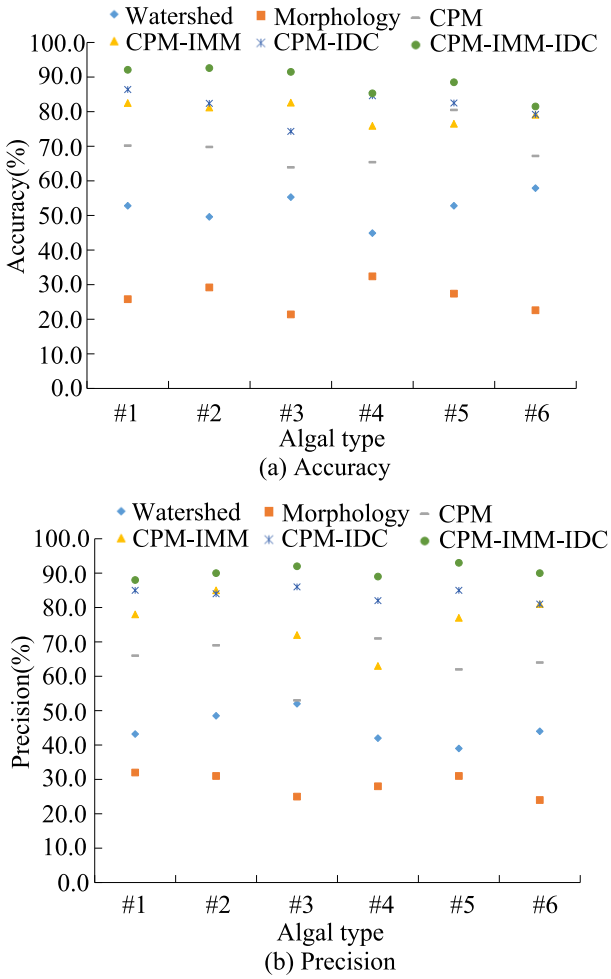

FIGURE 9. Accuracy and precision of each algorithm on the dataset.

recognition object from the environmental pixels in a clearer way. In addition, the CPM-IMM-IDC algorithm requires parameter training to be tested, so parameter settings and training iterations are also the main factors affecting model performance. From the above analysis, it can be seen that the CPM-IMM-IDC image segmentation algorithm proposed in this study has excellent segmentation quality, and the accuracy of seaweed recognition in the segmented images is significantly higher than other algorithms.

Then, the segmentation effects of each algal image segmentation algorithm on the overall data set are compared, and the comparison results are shown in Table 1. Morphological algorithm has the worst performance in segmentation accuracy and the longest computation time. Compared with traditional concave matching algorithms, the accuracy of the three improved concave matching algorithms has been significantly improved, and the overall accuracy of the improved concave matching algorithm designed in this study is the highest. The Accuracy and Precision of CPM, CPM-IMM, CPM-IDC, CPM-IMM-IDC algorithms are 72.4%, 76.5%, 81.6%, 90.3% and 73.8%, 75.7%, 79.4%, 88.2%, respectively. However, the calculation time of CPM-IMM-IDC algorithm is also higher than the other three CPM algorithms. The rationality of the experimental scheme proposed in this study is mainly reflected in the following aspects: First, in the comprehensiveness of comparative analysis, the experimental design demonstrates the performance improvement of the proposed method compared with traditional methods through comparison with various existing algorithms (such as morphology algorithm, watershed algorithm, Faster-RCNN and AlexNet). Such multi-model comparison can show the performance and reliability of the new model in different aspects, especially in the accuracy, accuracy and calculation efficiency and other important indicators. Second, in terms of the representativeness of the data set, Algae 6 data set used in this study and a large-scale data set containing 964,100 algae microscopic images contained a variety of algae. Such diversity ensured that the model generalization ability test was representative, which also increased the credibility and practicality of the research results. Thirdly, in terms of performance evaluation, multiple performance indicators, including Accuracy, precision, recall rate and F1 score, were used in this study to comprehensively evaluate the performance of the proposed method. The stability of the model was also evaluated by the method of 10-fold cross-validation, and the comparison of ROC curve and AUC value further confirmed the effectiveness of the new model in the recognition task. In short, this study designed an advanced computational model, provided sufficient data support and comprehensive evaluation test to ensure the technical rationality of the experiment.

B. PERFORMANCE TEST OF IR MODEL

In the experiment of IR performance, traditional ResNet, improved ResNet (IResNet), Faster-RCNN and AlexNet algorithms were used to construct recognition models respectively. All recognition models are arranged according to the frame in Figure 8, except that the IR algorithms are different. Using Python programming language to implement each IR model, and let the model run in the DL framework Tensor-flow. The operating environment of each model in the experiment is completely consistent, as shown below: The operating system is Windows8 Professional Edition; MySql is used as the database. Host processor Intel Core i7; Random access memory size 6GB; The size of the ROM is 1024 GB. The neural network algorithm parameters of each model are determined according to the optimal results obtained through multiple debugging. The final parameter setting scheme of the IResNet model is shown in Table 2.

Firstly, compare the accuracy numerical changes of IResNet during the training process under different pooling layer schemes, as shown in Figure 10. In Figure 10, the horizontal axis denotes the number of iterations, and the vertical axis denotes the accuracy of each scheme in the training process. Use the graph style to distinguish between different schemes. The convergent accuracy of all the improved schemes is superior to original ResNet algorithm, and the calculation

TABLE 1. Overall performance of each algorithm on the dataset.

Algorithm type	Algorithm number	Segmentation algorithm	Accuracy (%)	Precision (%)	Partition integrity rate (%)	Average calculation time (s)
Non-machine learning algorithms	1	Watershed	48.2	51.6	43.9	0.154
	2	Morphology	25.7	23.9	18.3	0.583
	3	Grayscale based segmentation algorithm	55.2	54.1	53.5	0.316
	4	Wavelet Transform Segmentation Algorithm	48.5	51.8	49.2	0.441
Machine learning algorithms	5	CPM	72.4	73.8	68.4	0.219
	6	CPM-IMM	76.5	75.7	72.5	0.246
	7	CPM-IDC	81.6	79.4	76.4	0.231
	8	CPM-IMM-IDC	90.3	88.2	84.2	0.274

TABLE 2. Parameter scheme of IResNetIR model.

Number	Parameter	Values and Rules	Number	Parameter	Values and Rules
#01	Training mode	Graphics processing unit	#06	Maximum number of iterations	1000
#02	Learning rate	0.0001	#07	Optimizer Type	Small batch Gradient descent
#03	BLSTM gradient descent momentum	0.1	#08	Parameter initialization method	Random initialization
#04	Sample size for single batch training	64	#09	Neuron loss rate	0.3
#05	Does the hidden layer have an offset term	Y	/	/	/

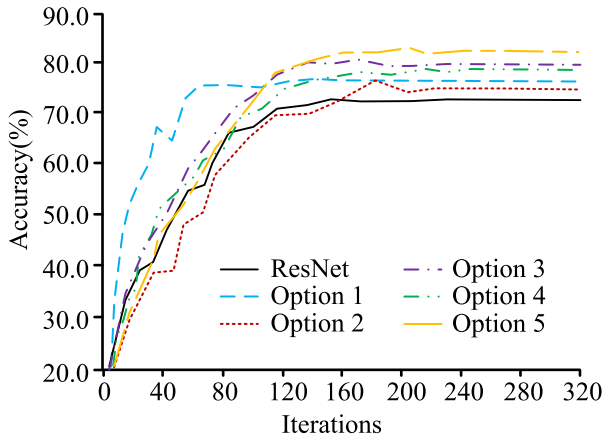

FIGURE 10. Comparison of accuracy of IResNet in training sets for different pooling layer schemes.

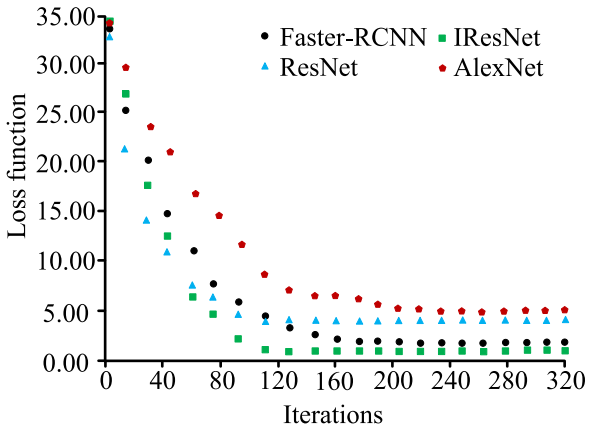

FIGURE 11. Change of loss function of each algorithm during training.

accuracy of schemes 4 and 5 based on schemes 2 and 3 is higher than that of the basic improved schemes. However, the recognition accuracy of scheme 5 was the highest after the final convergence, and subsequent studies were conducted according to Scheme 5 to form IResNet.

Figure 11 shows the change of the loss function of each algorithm in the training process. The horizontal axis denotes the number of iterations, and the vertical axis denotes the loss function value of each algorithm in the training process. The graph line style is used to distinguish different algorithms. When iteration increases, the loss function of each algorithm decreases and converges gradually. From the point of view of convergence speed, the IResNet of the fusion self-adjusting pooling layer designed in this study has the fastest convergence speed. From the perspective of training effect, the loss function values after convergence of IResNet, ResNet, Faster-RCNN and AlexNet algorithms are 0.75, 4.26, 1.39 and 5.07, respectively.

After the completion of training, the numerical changes of accuracy and precision rates of each model on the test set with different sample numbers are shown in Figure 12. Figures 12 (a) and 12 (b) represent the classification accuracy and segmentation accuracy statistical results of each algorithm under different test set calculation sample conditions, respectively. However, the horizontal axis of Figures 12 (a) and 12 (b) shows the number of seaweed images in the test set that participated in the calculation. The recognition accuracy rate and accuracy rate of each model are generally consistent with the change law of the test sample growth, and the fluctuation range is gradually reduced. When the test sample is the entire test set, that is, 25,200 images, The

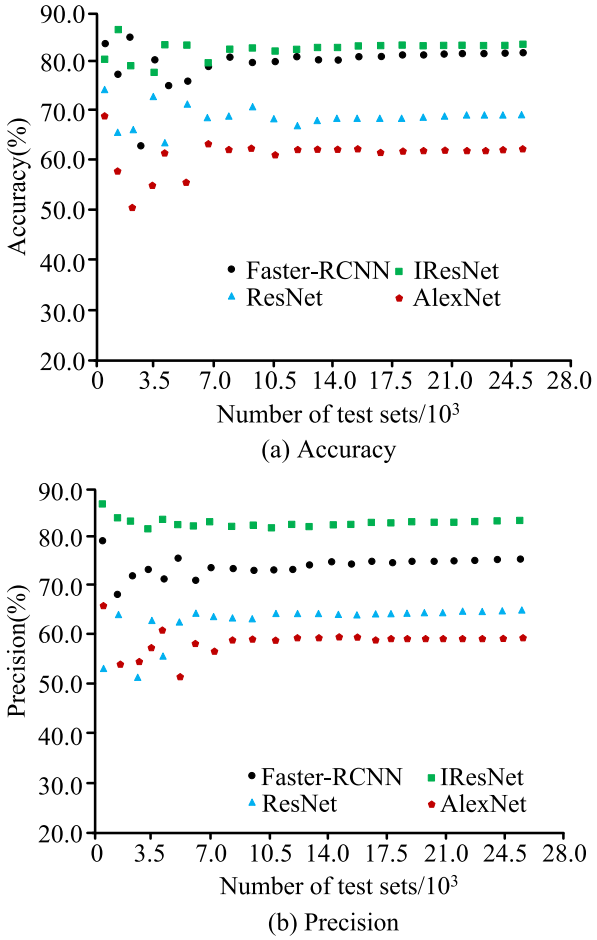

FIGURE 12. Comparison of accuracy and precision of various models on the test set.

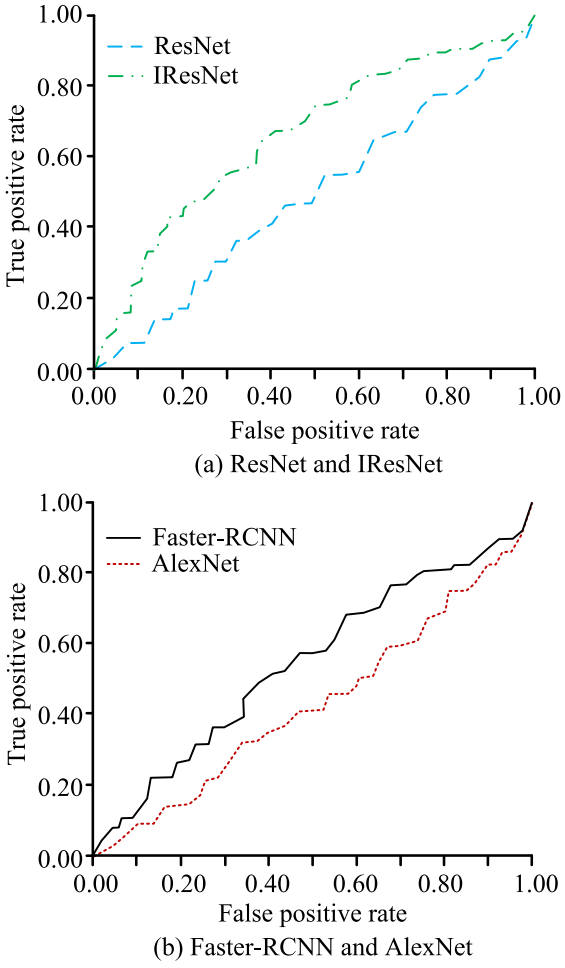

FIGURE 13. Comparison of AUC values of ROC among different models on the test set.

accuracy and precision of IResNet, ResNet, Father-RCNN and AlexNet algorithm models are 84.1%, 67.5%, 80.3%, 62.9% and 82.8%, 63.7%, 75.8%, 57.4%, respectively.

Then, the Area Under Curve (AUC) of the Receiver operating characteristic curve (ROC) of each seaweed IR model was compared, and the statistical results were shown in Figure 13. Figure 13 (a) shows the ROC curves of the improved algorithm designed in this study and the naive algorithm before improvement, while Figure 13 (b) shows the ROC curves of other comparative algorithms. The horizontal axis represents the false positive rate and the vertical axis represents the true positive rate. The ROC curve of the IResNet question answering model designed in this study is always above the ROC curve of other models, and the AUC area is the largest. Specifically, the ROC curves corresponding to AUC values of IResNet, ResNet, Faster-RCNN and AlexNet seaweed IR models are 0.68, 0.54, 0.57 and 0.51, respectively.

To test the performance of the recognition model designed for this study on a large sample dataset, 964100 microscopic images of algae from developing countries were collected and image recognition tests were conducted. These images are all from research structures and academic organizations in the real world, and some of the images in the datasetare shown in Figure 14.

Select green algae, diatoms, blue-green algae, dinoflagellates, golden algae, and Chlorella for analysis from this microscopic image dataset. Their microscopic images are shown in Figure 15 (a), Figure 15 (b), Figure 15 (c), Figure 15 (d), Figure 15 (e), and Figure 15 (d), respectively. Figure 15 is used to visually display the contents of different algae microscopic images in the data set used in this study, and to assist in explaining the calculation process of the improved concave matching algorithm.

Green algae are a large and diverse group of aquatic plants that generally contain chlorophyll, making them principally green. They are similar to higher plants on land in terms of photosynthesis. Green algae can survive in freshwater and marine environments, as well as grow under humid land conditions. The morphology is diverse, with both single-cell and multicellular chain or sheet-like structures. Diatoms are a type of single-cell algae that contain a siliceous cell wall (called a silica shell). Silicone shells often exhibit complex porous structures, often with unique and beautiful geometric

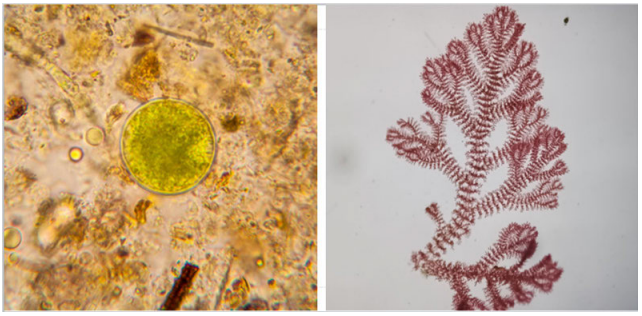

FIGURE 14. Partial image display of a large sample seaweed dataset.

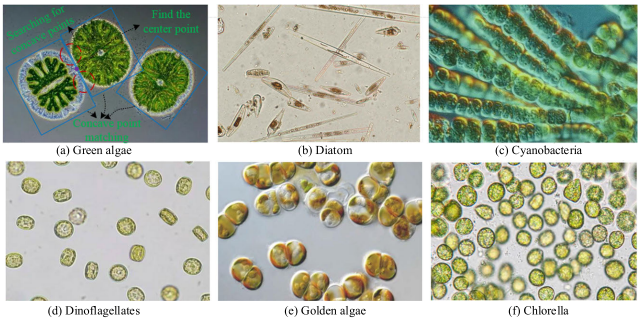

FIGURE 15. Microscopic image display of six common algae.

patterns, to distinguish different types of diatoms. They are the main primary producers in aquatic ecosystems, widely distributed in oceans and freshwater. Blue algae, also known as blue-green algae, are a group of photosynthetic bacteria. Although they are physiologically different from true green algae, they are capable of photosynthesis and releasing oxygen. Blue algae are diverse in morphology, ranging from single cells to complex multicellular structures, and can survive in extreme environments such as hot springs, saline alkali soils, and desert surfaces. Dinoflagellates are mostly single celled natural plankton, and their chloroplasts contain carotenoids, making some dinoflagellates brown or red in color. They have two unequal flagella to aid in movement and are particularly common in the ocean, with some species experiencing deadly red tide under certain conditions. Golden algae usually have golden chloroplasts and are important members of freshwater ecosystems. They can be single-celled or form filamentous structures. Gold algae are highly complex in biological classification, with some species exhibiting photosynthetic behavior closer to that of plants, while others resemble that of animals. Chlorella vulgaris is a green single celled micro-algae known for photosynthesis, with a simple spherical or elliptical cell structure. In a well lit environment, they can reproduce in large numbers and are a food source for many aquatic organisms. Chlorella vulgaris is also widely used in biotechnology research, including the development of bio-energy.

It is extremely difficult to identify artificial algae species through microscopic images, and the main difficulties are reflected in the following aspects:

Morphological diversity: There are significant differences in the morphology of different types of algae, and even within the same type of algae, their different life stages may lead to morphological differences, increasing the difficulty of identification; Morphological similarity: Although there are many types of algae, different types of algae often have similar morphological characteristics, such as cell size, shape, and chloroplast configuration. Especially in microscopic images, there may be a lack of sufficient resolution details, making distinguishing different types a challenge; Species variation: Even within the same species of algae, there may be morphological variations among individuals due to environmental factors such as light, nutrients, and temperature, resulting in uncertainty in identification. Staining and imaging quality: The quality of microscopic images directly affects recognition accuracy. The preparation of samples, staining techniques, and microscopic imaging settings (such as light sources, objective lenses, cameras, etc.) may all affect recognition. Uneven staining or unclear imaging may mask important recognition features. Human factors: require recognizers to have rich professional knowledge and practical experience. Artificial recognition requires long-term alignment with the microscope and continuous concentration, which can lead to visual fatigue and affect recognition efficiency and accuracy. These difficulties indicate that even for experienced algae experts, manual identification of algae species through microscopic images is a time-consuming, complex, and challenging task. Therefore, researchers and practitioners often turn to automation tools and computer vision technology to improve recognition speed and accuracy, especially for large-scale water quality monitoring projects.

Now, taking Figure 15 (a) as an example, we will demonstrate the matching process of the improved concave point matching algorithm. Firstly, the algorithm extracts the edge information of algae cells through Sobel operator, Laplace operator, and Canny operator, and obtains the concave points of the image. The concave points in the images of different algae represent the local concave points formed on the boundary of multiple algae individuals after multiple algae individuals squeeze and close to each other in the microscopic images of algae. The next step is to calculate the center point of the image with concave points. Finally, concavity and center point information are used for concavity matching, and different algae individuals are segmented based on the matching results.

The test results show that the recognition accuracy of IResNet, ResNet, Faster RCNN, and AlexNet models is 87.59%, 83.25%, 81.66%, and 79.18%, respectively. This indicates that the model designed in this study performs better than traditional models on large datasets. Finally, the performance of all the main evaluation indicators of each model in each part of the test set is compared.

Finally, compare the performance of all main evaluation indicators of each model on each part of the test set, and the statistical results are shown in Table 3. In the experimental results of Table 3, a significant difference test was also added.

TABLE 3. Comparison of main evaluation indicators of each model on the overall test Set.

Algorithm name	Test Set Usage Ratio	Accuracy (%)	Precision (%)	Recall (%)	F1 value	Number of iterations required for stability	Average recognition time (ms)
IResNet	50%	84.1*	82.8*	86.2*	83.9*	229*	4.5*
	100%	83.5*	81.9*	84.3*	82.7*	235*	4.4*
ResNet	50%	67.5*	63.7*	71.5*	65.4*	157*	3.2*
	100%	66.8*	63.8*	70.8*	63.5*	162*	3.1*
Faster-RCNN	50%	80.3*	75.8*	82.4*	78.3*	243*	4.6*
	100%	80.1*	76.0*	80.7*	77.1*	251*	4.6*
AlexNet	50%	62.9*	57.4*	68.5*	61.8*	128*	2.8*
	100%	62.4*	57.7*	66.1*	60.3*	134*	2.8*

The test method is as follows: after the operation of each algorithm test set is completed and the indicator data is counted, multiple sets of F difference tests are conducted on the data of each indicator in each recognition algorithm, and "*" is used as a marker to indicate significant differences between the data of this group and other groups of data of the same indicator. Observing Table 3, it can be seen that the improved ResNet algorithm based on Scheme 5's self-adjusting pooling layer designed in this study is significantly superior to other comparative recognition algorithms in terms of accuracy, Precision, Recall, and F1, and there is a significant difference in F-test data between the former and other algorithms. However, the IR speed of this algorithm is slightly slower than that of ResNet and AlexNet, and almost the same as that of Faster-RCNN algorithm. From the calculation time and number of iterations completed in Table 3, it can be seen that the IResNet algorithm designed in this study has the highest computational complexity. However, considering that the image segmentation and recognition module of this model did not make targeted improvements based on the pixel characteristics of the research object, but only improved from the perspective of distinguishing the image recognition object from the environment and the performance of the feature extraction structure itself, it can be considered that the designed algorithm has good scalability in the field of microscopic image recognition.

V. CONCLUSION

In view of the poor performance of the current mainstream seaweed microscopic IR, this study designed an improved ResNet neural network algorithm and an improved concave point matching algorithm, and will use both to build a seaweed microscopic IR model. Validation experiments based on real seaweed microscopic image data sets show that The mean Accuracy and Precision values of various algae image based on Watershed, Morphology, CPM, CPM-IMM, CPM-IDC, CPM-IMM-IDC algorithms were 52.2%, 26.5%, 69.5%, 79.6%, 81.6% and 88.6, respectively, 44.8%, 28.5%, 64.2%, 76.0%, 83.8%, 90.3%. The convergent accuracy of all self-adjusting pooling layer improvement schemes is superior to the original ResNet, and the calculation accuracy of schemes 4 and 5 based on schemes 2 and 3 is higher than that of the basic improvement schemes. However, scheme 5 has the highest recognition accuracy after final convergence. From the point of view of convergence speed, the IResNet of the fusion self-adjusting pooling layer designed in this study has the fastest convergence speed. From the perspective of training effect, the loss function values after convergence of IResNet, ResNet, Faster-RCNN and AlexNet algorithms are 0.75, 4.26, 1.39 and 5.07, respectively. The improved concave matching algorithm owns the highest overall accuracy. The Accuracy and Precision of CPM, CPM-IMM, CPM-IDC and CPM-IMM-IDC algorithms are 72.4%, 76.5%, 81.6%, 90.3% and 73.8%, 75.7%, 79.4%, 88.2%, respectively. However, the calculation time of CPM-IMM-IDC algorithm is also higher than the other three CPM algorithms. When the test sample is the whole test set, the accuracy and accuracy of IResNet, ResNet, father-RCNN and AlexNet algorithm models are 84.1%, 67.5%, 80.3%, 62.9% and 82.8%, 63.7%, 75.8%, 57.4%, respectively. The corresponding AUC values of ROC curve were 0.68, 0.54, 0.57 and 0.51, respectively. The improved concave matching algorithm designed in this study has obvious advantages over the traditional algorithm in segmentation accuracy, and its seaweed microscopic IR accuracy is also higher than that of common recognition models. However, due to the limitations of research conditions, the designed model could not be applied to market products for testing in this study. Future studies will improve this aspect.

REFERENCES

[1] Y. Zhao, C. Wang, J. Pei, and X. Yang, "Nonlinear loose coupled non-negative matrix factorization for low-resolution image recognition," *Neurocomputing*, vol. 443, pp. 183–198, Jul. 2021, doi: 10.1016/j.neucom.2021.02.068.

[2] Y. Wang, H. Liu, M. Guo, X. Shen, B. Han, and Y. Zhou, "Image recognition model based on deep learning for remaining oil recognition from visualization experiment," *Fuel*, vol. 291, May 2021, Art. no. 120216, doi: 10.1016/j.fuel.2021.120216.

[3] X. Yang, Y. Zhang, W. Lv, and D. Wang, "Image recognition of wind turbine blade damage based on a deep learning model with transfer learning and an ensemble learning classifier," *Renew. Energy*, vol. 163, pp. 386–397, Jan. 2021, doi: 10.1016/j.renene.2020.08.125.

[4] Y. Yang and X. Song, "Research on face intelligent perception technology integrating deep learning under different illumination intensities," *J. Comput. Cognit. Eng.*, vol. 1, no. 1, pp. 32–36, Feb. 2022, doi: 10.47852/bonviewjcce19919.

[5] Z. Li, W. Wang, Y. Sun, S. Wang, S. Deng, and Y. Lin, "Applying image recognition to frost built-up detection in air source heat pumps," *Energy*, vol. 233, Oct. 2021, Art. no. 121004, doi: 10.1016/j.energy.2021.121004.

[6] P. Gao, D. Zhao, and X. Chen, "Multi-dimensional data modelling of video image action recognition and motion capture in deep learning framework," *IET Image Process.*, vol. 14, no. 7, pp. 1257–1264, May 2020, doi: 10.1049/iet-ipr.2019.0588.

[7] J. Liang, F. Xu, and S. Yu, "A multi-scale semantic attention representation for multi-label image recognition with graph networks," *Neurocomputing*, vol. 491, pp. 14–23, Jun. 2022, doi: 10.1016/j.neucom.2022.03.057.
[8] G. Singh, "Think positive: An interpretable neural network for image recognition," *Neural Netw.*, vol. 151, pp. 178–189, Jul. 2022, doi: 10.1016/j.neunet.2022.03.034.
[9] H. Zhang, G. Peng, Z. Wu, J. Gong, D. Xu, and H. Shi, "MAM: A multipath attention mechanism for image recognition," *IET Image Process.*, vol. 16, no. 3, pp. 691–702, Feb. 2022, doi: 10.1049/ipr2.12370.
[10] W. J. Hu, T. Y. Xie, B. S. Li, Y. X. Du, and N. N. Xiong, "An edge intelligence-based generative data augmentation system for IoT image recognition tasks," *J. Internet Technol.*, vol. 22, no. 4, pp. 765–778, Jul. 2021, doi: 10.53106/160792642021072204005.
[11] C. Sun, L. Wang, N. Wang, and S. Jin, "Image recognition technology in texture identification of marine sediment sonar image," *Complexity*, vol. 2021, pp. 1–8, Mar. 2021, doi: 10.1155/2021/6646187.
[12] Y. J. Choi and K.-J. Lee, "Possibilities of artificial intelligence use in orthodontic diagnosis and treatment planning: Image recognition and three-dimensional VTO," *Seminars Orthodontics*, vol. 27, no. 2, pp. 121–129, Jun. 2021, doi: 10.1053/j.sodo.2021.05.008.
[13] X. Gou, Q. Liu, H. Rong, M. Hu, P. Paul, F. Deng, X. Zhang, and Z. Yu, "A novel spiking neural P system for image recognition," *Int. J. Unconv. Comput.*, vol. 16, nos. 2–3, pp. 121–139, Jun. 2021.
[14] G. Lan, Z. Liu, Y. Zhang, T. Scargill, J. Stojkovic, C. Joe-Wong, and M. Gorlatova, "Edge-assisted collaborative image recognition for mobile augmented reality," *ACM Trans. Sensor Netw.*, vol. 18, no. 1, pp. 1–31, Oct. 2021, doi: 10.1145/3469033.
[15] Y. Liu, J. Wang, and W. Bai, "Commodity price recognition and simulation of image recognition technology based on the nonlinear dimensionality reduction method," *Adv. Math. Phys.*, vol. 2021, pp. 1–9, Nov. 2021, doi: 10.1155/2021/1045342.
[16] K. Ohri and M. Kumar, "Review on self-supervised image recognition using deep neural networks," *Knowledge-Based Syst.*, vol. 224, Jul. 2021, Art. no. 107090, doi: 10.1016/j.knosys.2021.107090.
[17] M. Li, R. Wang, J. Yang, L. Xue, and M. Hu, "Multi-domain few-shot image recognition with knowledge transfer," *Neurocomputing*, vol. 442, pp. 64–72, Jun. 2021, doi: 10.1016/j.neucom.2021.01.123.
[18] Z. Gao, Y. Liu, G. Xu, and X. Wen, "Pairwise attention network for cross-domain image recognition," *Neurocomputing*, vol. 453, pp. 393–402, Sep. 2021, doi: 10.1016/j.neucom.2020.06.147.
[19] G. A. Traeger, M. H. Teichmann, B. Schröder, and M. Wenderoth, "Combining grating-coupled illumination and image recognition for stable and localized optical scanning tunneling microscopy," *Rev. Sci. Instrum.*, vol. 94, no. 2, p. 23702, Feb. 2023, doi: 10.1063/5.0123604.
[20] Z. Zhong, Z. Hou, J. Liang, E. Lin, and H. Shi, "Multimodal cooperative self-attention network for action recognition," *IET Image Process.*, vol. 17, no. 6, pp. 1775–1783, May 2023, doi: 10.1049/ipr2.12754.
[21] G. Zhang, L. Wang, H. Wang, Y. Chen, and J. Dang, "Theoretical and experimental research on two-phase flow image reconstruction and flow pattern recognition," *Rev. Sci. Instrum.*, vol. 94, no. 3, p. 34709, Mar. 2023, doi: 10.1063/5.0131667.
[22] R. Dhawan, K. Moyal, and A. Choudhary, "Optical filter-less photonic FMCW radar for multi-target detection," *IEEE Photon. Technol. Lett.*, vol. 35, no. 2, pp. 81–84, Jan. 2023, doi: 10.1109/LPT.2022.3224219.
[23] M. Zhao, L. Yue, J. Hu, S. Du, P. Li, and L. Wang, "Salient target detection in hyperspectral image based on visual attention," *IET Image Process.*, vol. 15, no. 10, pp. 2301–2308, Aug. 2021, doi: 10.1049/ipr2.12197.

WENDA ZHENG was born in Hubei. He received the B.S. degree from the National University of Defense Technology, Hunan, China, in 2001, and the M.S. degree from Nanjing Political College, Jiangsu, China, in 2006. He is currently pursuing the degree in control science and engineering from the University of Science and Technology Beijing. His research interests include data mining, natural language, machine vision, and knowledge graphics.

YIBO AI received the B.S., M.S., and Ph.D. degrees in control science and engineering from the University of Science and Technology Beijing, Beijing, China, in 2004, 2007, and 2018, respectively. Her research interests include big data fusion, complex system modeling, nondestructive detection, system reliability analysis, and remaining life prediction. She has been engaged in the application of high-speed train gear transmission systems, high-accuracy motorized spindle, biometrics feature recognition, and multi-level system service safety assessment.

WEIDONG ZHANG received the Ph.D. degree in control science and engineering from the University of Science and Technology Beijing, in 2004. He has presided over more than 20 scientific research projects, such as the National Science and Technology Infrastructure Program, the National Key Research and Development Plan Project, the National Natural Science Foundation of China, and enterprise projects. His research interests include monitor testing technology and data integration, system safety evaluation and life prediction, and intelligent optimization and control for systems.