



Article Research on Cyanobacterial-Bloom Detection Based on Multispectral Imaging and Deep-Learning Method

Ze Song ¹^[6], Wenxin Xu ¹, Huilin Dong ¹, Xiaowei Wang ¹, Yuqi Cao ¹^[6], Pingjie Huang ¹, Dibo Hou ^{1,*}^[6], Zhengfang Wu ² and Zhongyi Wang ²

- State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China; sxzturbo@zju.edu.cn (Z.S.); xuwen_xin@zju.edu.cn (W.X.); 22032103@zju.edu.cn (H.D.); xw_wang@zju.edu.cn (X.W.); yuqicao1994@gmail.com (Y.C.); huangpingjie@zju.edu.cn (P.H.)
- ² City Intelligence, Cloud & AI, Huawei Technologies Co., Ltd., Shenzhen 518100, China; wuzhengfang@huawei.com (Z.W.); bigbaby.wang@huawei.com (Z.W.)
- * Correspondence: houdb@zju.edu.cn

Abstract: Frequent outbreaks of cyanobacterial blooms have become one of the most challenging water ecosystem issues and a critical concern in environmental protection. To overcome the poor stability of traditional detection algorithms, this paper proposes a method for detecting cyanobacterial blooms based on a deep-learning algorithm. An improved vegetation-index method based on a multispectral image taken by an Unmanned Aerial Vehicle (UAV) was adopted to extract inconspicuous spectral features of cyanobacterial blooms. To enhance the recognition accuracy of cyanobacterial blooms in complex scenes with noise such as reflections and shadows, an improved transformer model based on a feature-enhancement module and pixel-correction fusion was employed. The algorithm proposed in this paper was implemented in several rivers in China, achieving a detection accuracy of cyanobacterial blooms of more than 85%. The estimate of the proportion of the algae bloom contamination area and the severity of pollution were basically accurate. This paper can lay a foundation for ecological and environmental departments for the effective prevention and control of cyanobacterial blooms.

Keywords: remote-sensing technology; cyanobacterial blooms; vegetation index; deep learning

1. Introduction

Water is the source of life and an important resource for the survival and development of human beings. However, with increasing human activities and climatic changes, water resources have experienced unprecedented threats, including nutrient enrichment, and inorganic and organic pollution [1]. Water eutrophication has been occurring frequently with the intensification of nitrogen and phosphorus pollution, and one of the severely disastrous consequences is the globally increasing frequency of cyanobacterial blooms [2]. In recent years, a large number of cyanobacterial blooms have erupted in worldwide rivers and lakes, such as Taihu Lake and Dianchi Lake in China, Lake Erie in the USA, and Wood Lake in Canada [3-6], causing a series of serious ecological problems and posing a huge impact on people's production and lives. Specifically, cyanobacterial blooms can produce a variety of toxins which cause a range of diseases when ingested by organisms [7]. Moreover, the microbial degradation of cyanobacterial blooms reduce oxygen levels in lakes, resulting in the deaths of fish and inhibiting the growth of aquatic vegetation [8,9]. Mounting evidence shows that cyanobacterial blooms are highly likely to expand further, owing to ongoing eutrophication, in the future [10]. Obviously, it is especially important to perform real-time monitoring and quantification of cyanobacterial blooms.

Currently, routine methods for analyzing cyanobacterial blooms have been widely applied through field sampling and laboratory analysis [11]. However, this traditional



Citation: Song, Z.; Xu, W.; Dong, H.; Wang, X.; Cao, Y.; Huang, P.; Hou, D.; Wu, Z.; Wang, Z. Research on Cyanobacterial-Bloom Detection Based on Multispectral Imaging and Deep-Learning Method. *Sensors* **2022**, 22, 4571. https://doi.org/10.3390/ s22124571

Academic Editor: Benoit Vozel

Received: 9 May 2022 Accepted: 15 June 2022 Published: 17 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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/). method is highly time-consuming, laborious and expensive. Moreover, it cannot obtain comprehensive information on cyanobacterial blooms, so it is not intended for real-time monitoring of large-scale cyanobacterial blooms. Instead, most of the worldwide research results on cyanobacterial-bloom detection are achieved using remote-sensing hyperspectral technology [12]. In these studies, researchers mainly use a variety of vegetation indices to extract spectral features for detecting cyanobacterial blooms, including the Normalized Difference Vegetation Index (NDVI), Atmospherically Resistant Vegetation Index (ARVI), and Enhanced Vegetation Index (EVI) [13–15]. In recent years, scholars have also proposed some improved vegetation-index methods to detect cyanobacterial blooms. Hu et al. [16] proposed a new Floating Algae Index (FAI) based on NDVI and EVI to detect floating algae in the open-ocean environment. This approach addresses the sensitivity of traditional vegetation indices to changes in environmental and observational conditions. Huang et al. [17] investigated the variation in chlorophyll concentration distribution in Lake Taihu using the band ratio method, based on Geostationary Ocean Color Imager's data. Cannizzaro et al. [18] proposed a new quantitative method for cyanobacterial-bloom detection by improving the index based on the analysis of the optical properties of cyanobacterial blooms in Florida Bay. Sachidananda et al. [19] proposed a cyanobacterial index algorithm named CIcyano to detect cyanobacterial blooms in the lakes of eleven states in the United States, and verified the effectiveness of the algorithm.

However, the above applications of vegetation index are all performed with the help of satellite image data, which require a long acquisition period and are inadequate in monitoring low-concentration algal bloom areas in real time. Small UAV remote-sensing technology, which has the advantages of low cost, low risk, high timeliness, and high resolution, is widely used in various scenarios and has become a research hotspot in recent years [20–22]. It can effectively make up for the shortcomings of traditional methods, such as long acquisition time period and low detection efficiency. Nevertheless, the vegetation indices mentioned in previous literature are less capable of detecting cyanobacterial blooms in complex river-surface environments, because low-altitude remote-sensing technology will certainly amplify space interference. Moreover, there are still no commercially available preprocessing methods designed for aquatic purposes and UAV close-range remote sensing [20]. Therefore, this paper proposes a new vegetation-index method based on improved feature-band and detail fusion to robustly extract the multispectral features of cyanobacterial blooms using low-altitude airborne detection technology.

In addition, due to the influence of reflections, shadows, and floating objects on the river surface, the detection thresholds of cyanobacterial blooms in different channels are not consistent. Obviously, only using a vegetation index to quantify cyanobacterial blooms is not effective enough, and a more robust algorithm is needed to assist in the detection of cyanobacterial blooms. Deep learning is a powerful and adaptive learning method that can classify or predict non-linear data accurately under complex conditions, which makes it a promising tool in performing image-level semantic segmentation of cyanobacterial blooms. Here, semantic segmentation is necessary and advantageous because of the irregular shape of cyanobacterial blooms. Jonathan et al. [23] from the University of California, Berkeley proposed the Fully Convolutional Network (FCN), pioneering the use of deep learning for image semantic segmentation, but the proposed networks were not sensitive enough to detect details and failed to consider pixels without any spatial coherence. A team from the University of Cambridge developed a deep network for image semantic segmentation named SegNet to segment the regions in an image where objects were located, with accuracy down to the pixel level [24]. Thus far, image semantic segmentation methods based on deep learning have been widely used in autonomous driving, medical image classification and other fields [25–28], as well as in the detection of cyanobacterial blooms. Yang et al. [29] proposed a deep generative adversarial network (DGAN), which demonstrated a better segmentation result on irregular cyanobacterial-bloom regions. Using an illumination processing algorithm based on a deep neural network (DCNN), Luo et al. [30] normalized the illumination intensity of images to a reasonable range, and thus, effectively improved

the accuracy of cyanobacteria detection under the condition of strong light. Xiang et al. [31] adopted the attention module and the number of residual blocks to streamline the structure based on a residual attention network model, and their detection model could discriminate various species of cyanobacterial blooms in Taihu Lake accurately and swiftly.

It is worth noting that few previous studies have paid attention to the phenomenon in which the boundary classifications of cyanobacterial blooms and other objects often give incorrect results. Among the many deep-learning algorithms, the transformer model has a strong model generalization ability. It has a flexible structure and makes essentially no assumptions about the structural errors in the input data, and can be pre-trained to handle large amounts of unlabeled data. More importantly, the attention mechanism will make the semantic association between adjacent pixels better [32]. Therefore, a new transformer model based on feature enhancement and area correction is proposed to overcome the influence of complex environmental factors and boundary misclassification in cyanobacterial-bloom detection.

In this paper, a multispectral detection model for cyanobacterial blooms based on a transformer network is proposed. The spectral features of the multispectral data of cyanobacterial blooms captured by UAV were extracted using an improved vegetationindex method. To classify the feature boundary, feature-enhancement and region-correction modules were added to the original transformer model. The proposed method not only improved the accuracy of detection of cyanobacterial blooms, but also allowed for assessment of the extent of cyanobacterial-bloom pollution and its proportion out of the total size. The results of this study can provide a reference for the control of cyanobacterial blooms in the fields of ecology and the environment.

2. Materials and Methods

2.1. Data and Evaluation

Data were collected using Phantom 4 multispectral version of the UAV made by DJI. The UAV carries a high-resolution multispectral lens and has six camera sensors, including one color sensor for visible imaging and five monochrome sensors for multispectral imaging in the blue band (B1), green band (B2), red band (B3), red edge-band (B4) and near-infrared band (B5). Healthy and flourishing plants were more reflective in these bands; thus, they can be used to detect cyanobacterial blooms.

The experimental data in this paper were collected from August to October 2021 in several areas of China with severe cyanobacterial blooms, covering many rivers and lakes in Kunming, Yunnan Province; Suzhou, Jiangsu Province; and Huzhou, Zhejiang Province. The data also covered a wide range of weather conditions and external-disturbance scenarios. The dataset included 5400 images and 10 videos, as shown in Table 1.

Province and City of Sampling	Specific Locations	Number of Images/Sheet
	Daqing River	708
Kunming, Yunnan	Baoxiang River	1080
-	Haigeng Park	558
	Xujiang River	1782
Suzhou, Jiangsu	Jinji Lake	564
	Youlian New Village	108
Lughou Zhoijong	Taihu Lake	510
Huzhou, Zhejiang	Maoer Port	90

Table 1. Distribution of cyanobacterial blooms datasets.

2.2. Methodology

In this paper, a cyanobacterial-bloom detection algorithm based on deep learning is proposed. The major steps involved image pre-processing, multi-spectral feature extraction and detection model construction, as shown in Figure 1. Among them, adaptive median filtering and light masking were used in the image pre-processing to remove the background noise generated by the sensor itself and the interference of strong light; an improved vegetation index was used in the feature extraction to extract the spectral features of cyanobacterial blooms more accurately; and a feature-enhancement module and area-correction module were added to address the misclassification of cyanobacterial-bloom boundaries during model construction.

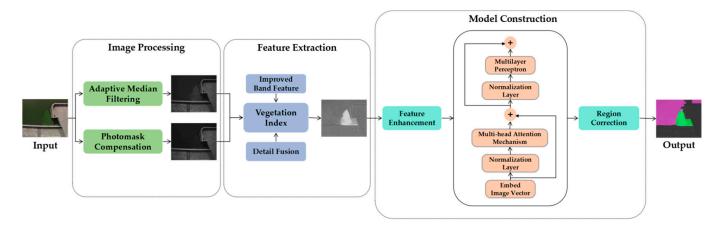


Figure 1. Framework of cyanobacterial-bloom detection algorithm.

2.2.1. Image Processing

Due to the influence of various environmental factors, multi-dimensional random noise, which interfered with the feature extraction of cyanobacterial blooms, existed in the obtained multi-spectral image data. Aimed at noises of different sizes and shapes in the images, adaptive median filtering was used to improve signal-to-noise ratio. Then, aimed at the small area of bright light spots in the image, a light mask was used to reduce the influence of bright light on the extraction of effective information.

2.2.2. Feature Extraction

As multispectral images involve multi-band spectral data, traditional vegetation indices are difficult to apply to feature extraction for all types of data. Therefore, deep-learning methods are introduced to extract data features based on pre-designed rules.

In view of the multispectral images collected in this paper, a vegetation-index method based on the fusion of improved band features and details was proposed. In order to better explore the differences in the reflectance of different features, multispectral images with rich information in the dataset were selected to count the image element values of cyanobacterial blooms, clean water bodies, floating vegetation, surface litter, boats and other materials in each waveband. Since the pixel values of cyanobacterial blooms and other interfering objects basically did not overlap in B1, B3 and B5, the multispectral feature-extraction method was verified based on the reflectivity of these three bands. NDVI is calculated based on the reflectance of B3 and B5 and is often used to detect vegetation growth and vegetation cover. The formula is shown in Equation (1):

$$NDVI = \frac{\rho_{B5} - \rho_{B3}}{\rho_{B5} + \rho_{B3}}$$
(1)

where ρ_{B3} and ρ_{B5} are the reflectance of the feature in the red and near-infrared bands, respectively. First, to keep the image features of the model input within the same range, the two-by-two combinations of wavebands were substituted into Equation (1) for normalization, with the goal of selecting the combination with the greater difference between the indices of each feature. In addition, to further extend the difference in indices between cyanobacterial blooms and other features, we adjusted the calculation method and

enhanced the detail. Ultimately, this paper used the vegetation index which is given in Equation (2):

$$NDVI_{sqrt} = sgn(\frac{\rho_{B5} - \rho_{B1}}{\rho_{B5} + \rho_{B1}}) \times \sqrt{\left|\frac{\rho_{B5} - \rho_{B1}}{\rho_{B5} + \rho_{B1}}\right|}$$
(2)

where ρ_{B1} and ρ_{B5} are the reflectance of the feature in the blue and the near-infrared wavelengths, respectively.

2.2.3. Model Construction

As cyanobacterial blooms vary in size and shape and are mostly irregularly distributed, it is easy to miss some important image details during feature learning and misclassify pixel points at the junction of different objects or in some boundary areas. Therefore, on the basis of the improved vegetation index, this paper proposed an improved transformer model based on feature enhancement and region correction, achieving the following two improvements to robustly detect cyanobacterial blooms.

(1) Feature-Enhancement Module

In the data processing stage, the improved vegetation index was used to calculate the NDVI_{sqrt} value of each pixel point within the image to form a single-channel image matrix, which was then converted into a three-channel two-dimensional matrix as the input to the network model. This paper also expanded the dataset to increase data diversity by randomly cropping and scaling the three-channel images whose tensor sizes were changed from 1600 \times 1300 pixels to 768 \times 768 pixels.

Boundary classification can mainly be achieved in two ways: one is to directly transfer the feature mapping of the shallow network in the model to the corresponding deep network, which is continuously transferred and superimposed; the other is to pool and convolve a certain network layer in the network structure for calculation. Both of the above methods are able to retain small amounts of spatial information, but the former greatly increases the computational load of the model, while the latter cannot guarantee the balance of the shallow-feature and deep-feature information to a certain extent. Aimed at these shortcomings, this paper designed a new feature-enhanced network architecture unit.

The design structure of the feature-enhancement module is shown in Figure 2. In order to extract more semantic features, to solve the problems of network degradation and gradient descent, and to improve the network performance, the feature-enhancement module designed in this paper contained five convolutional layers with one residual unit structure. Firstly, the stride of the first layer was chosen to be 4 and the padding to be 3. The image size was calculated to be 192×192 to further obtain the shallow features, and then used as the input for the next convolutional layer. Then, one convolution kernel with a size of 4×4 and two convolution kernels with a size of 3×3 were used to extract the multi-layer features of the image, respectively. In the residual structure unit, the 3×3 sized convolution kernel on the main channel was used to extract details of the image texture features, and the 1×1 sized one to obtain the deepest features of the graph. A layer of ReLU activation function was added between the two convolution layers to avoid gradient explosion and disappearance, allowing features in complex scenes to be learned more comprehensively. At the same time, the side channels transferred the edge features to the extracted feature maps, and the backbone network model contained rich spatial and semantic information by calculating pixel summation.

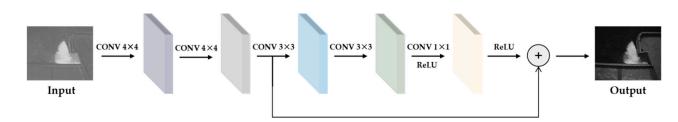


Figure 2. Structure of feature-enhancement module.

For a given multispectral feature image matrix, this feature-enhancement module performed several convolution calculations for different purposes to obtain rich, shallow and deep image features, respectively. The feature information was balanced to obtain multiscale feature mapping, which not only helped to avoid feature loss caused by increased layers in the traditional deep-learning model, but also improved the convergence rate of training and obtained more comprehensive and rich multispectral image information.

(2) Region Correction Module

Pixel similarity is scored based on the similarity of information content between images, and the number of scores is used to determine the similarity of image information content, which is also used for image semantic segmentation. Ahn et al. [33] proposed a model framework to generate image-segmentation labels from a given image-level class label. In a weakly supervised environment, the training model can segment local discriminative parts instead of the whole object region. To correct the semantic information of adjacent pixels, this paper added a pixel-similarity-based region-correction module after the original model decoder, to identify the local region where the target object was located by estimating the semantic affinity between adjacent image coordinates. Then, pixel-level region correction was performed according to the predicted affinity to obtain accurate object location information.

Given a multispectral image, the corresponding Vector Class Activation Map (CAM) was calculated. The calculation method of the target class CAM is as in Equation (3) [33]:

$$M_c(x,y) = W_c^T f^{cam}(x,y)$$
(3)

where M_c is the vector class activation mapping, W_c is the classification weight, and $f^{cam}(x, y)$ is the feature vector at the point on the feature map. The semantic similarity labels were generated based on the obtained CAM, which is calculated as in Equation (4):

$$W(x_1, x_2) = \begin{cases} 1, & \text{if } l_1 = l_2 \\ 0, & \text{otherwise} \end{cases}$$
(4)

where l_1 and l_2 are the categories of the adjacent pixels x_1 and x_2 , and $W(x_1, x_2)$ is the pixel label. Within a certain radius, the neighborhood image was connected with adjacent pixels, and the semantic affinity of the connected pixel point pairs was calculated. The information about the pixel similarity of the edges in the image would be passed to semantically similar places by random traversal and to other similar places in the surrounding area by penalty. This semantic extension would significantly correct the CAM and recover the morphology of tiny objects, especially at the boundaries of different classes. Ultimately, the class tokens of the adjusted CAM were used to synthesize segmentation labels for training the segmentation model. The pixel-similarity-based region-correction module reduced the loss of pixel details, enhanced the information correlation and improved the accuracy of cyanobacterial-bloom detection in complex scenes.

2.2.4. Evaluation Indicators

In order to evaluate the performance of the proposed algorithm and calculate the cyanobacterial bloom pollution area in the sampling area, the following evaluation indicators were used to for quantification.

Pixel Accuracy (*PA*) refers to the ratio of the number of pixels whose categories are correctly predicted by the algorithm to the total number of pixels, which can be understood as the percentage of correctly classified pixels. It is the most common and intuitive evaluation index in image segmentation [34,35], and its calculation is shown in Equation (5):

$$PA = \frac{TP + TN}{TP + TN + FP + FN}$$
(5)

where *TP*, *FP*, *FN* and *TN* represent true positive, false positive, false negative and true negative cases, respectively. However, *PA* cannot determine whether the predicted cyanobacterial blooms are correctly located and may misclassify pixel points, not serving as a good indicator of the results in the case of unbalanced samples. Therefore, *PA* alone does not provide an accurate measure of algorithm performance. Intersection over Union (*IoU*) refers to the ratio of the intersection and union of the model's prediction result and the true value for a certain category. It is a standard performance measure for object category segmentation problems [36] and is calculated using Equation (6):

$$IoU = \frac{TP}{FP + TP + FN}$$
(6)

In the cyanobacterial-bloom detection task, *IoU* represents the magnitude of the difference between the algorithm's predicted cyanobacterial blooms and the true value. The higher the value, the better the detection effect. This paper uses a combination of *PA* and *IoU* to assess the improvement in the improved vegetation-index method on the overall accuracy of the algorithm.

At the same time, the proportion of the area contaminated by cyanobacterial blooms can be estimated by counting the number of cyanobacterial-bloom pixels on the river surface, and the calculation formula is shown in Equation (7):

$$P = \frac{A_1}{A} \times 100\% \tag{7}$$

where A_1 is the number of cyanobacterial-bloom pixels, and A is the total number of image elements.

3. Results

In the detection of cyanobacterial blooms, the diversity of the river and lake environment—as evidenced by distracting factors such as birds, boats, reflections and shadows on the surface of the water body—can, to some extent, affect the detection accuracy of the algorithm; this makes the detection of cyanobacterial blooms very difficult. Figure 3 indicates the presence of irregularly sized light spots and large reflections on the river surface, respectively, which can change the pixel characteristics and, thus, lead to false or missed detections. In this paper, the randomly selected samples were divided into two categories—regular scenes and complex scenes—to test the performance of the algorithm. Among them, a regular scene refers to a situation wherein there is no obvious interference, and a complex scene indicates that there is interference dominated by strong light and reflections.

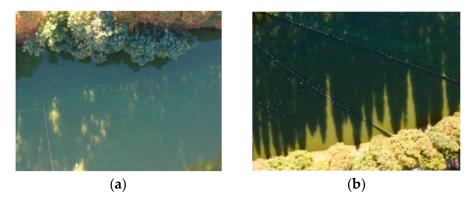


Figure 3. Complex river scenes: (a) reflection case; (b) shadow case.

3.1. Results of Feature Extraction

For different scenes, we compared the feature extraction effects before and after the improved vegetation-index method, and input the different features extracted into the same transformer model to analyze and compare the model output results.

Figure 4a shows the original visible image in a typical regular scene of the Xujiang River in Suzhou. Figure 4b shows the detection result of the multispectral features extracted from NDVI as input to the transformer model, which can basically distinguish cyanobacterial blooms from clean water; however, there is some error in the detection effect of cyanobacterial blooms in the lower-concentration areas. Figure 4c shows the detection results of the multispectral features extracted based on NDVI_{sqrt} as the input of the transformer model. There is still a small amount of misclassification at the riverbank boundary, but the model has enhanced the learning of the spectral features of cyanobacterial blooms. It can be seen intuitively that compared with NDVI, NDVI_{sqrt} identifies cyanobacterial blooms or in the early stages of growth. Provably, NDVI_{sqrt} can improve detection accuracy by extracting cyanobacterial spectral features that are weaker and more difficult to identify.

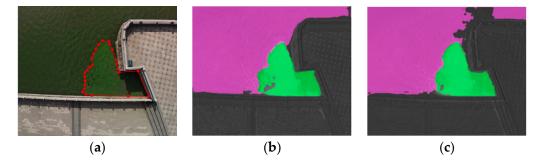


Figure 4. Detection results of different algorithms for regular scene: (a) visible light image; (b) detection results of NDVI as the feature-extraction method; (c) detection results of NDVI_{sqrt} as the feature-extraction method.

Figure 5a shows the original visible image in a typical reflection scene of the Daqing River in Yunnan. The light spot in the lower right corner covers part of cyanobacterial bloom, and brings great difficulty to accurate detection. Figure 5b,c show the detection results using the multispectral features extracted before and after the improved vegetation index as the input of the transformer model, respectively. The former had obvious detection errors and missed a large area of cyanobacterial blooms; meanwhile, the latter was able to identify cyanobacterial blooms in the reflection scene, probably because NDVI_{sqrt} was able to extract more accurate spectral features, coupled with the stronger learning ability of the attention mechanism. High-precision detection of cyanobacterial blooms can still be achieved, even under strong light conditions.

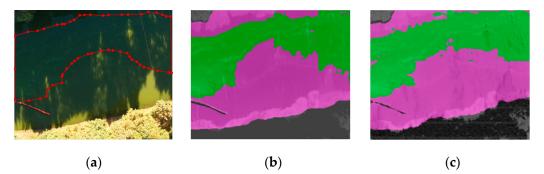


Figure 5. Detection results of different algorithms for reflection scene: (a) visible light image; (b) detection results of NDVI as the feature-extraction method; (c) detection results of NDVI_{sqrt} as the feature-extraction method.

Figure 6a shows the original visible image in a typical shadowy scene of the Daqing River in Yunnan. Figure 6b,c show the detection results of the multispectral features extracted before and after the improved vegetation index as the input of the transformer model, respectively. Both can distinguish the riverbank and water body more correctly. However, it can be seen that for mild cyanobacterial blooms under shadow coverage, NDVI extracted almost no features, and cyanobacterial blooms and partial vegetation were mistaken for water. On the contrary, NDVI_{sqrt} was able to practice efficient retrieval, even when covered.

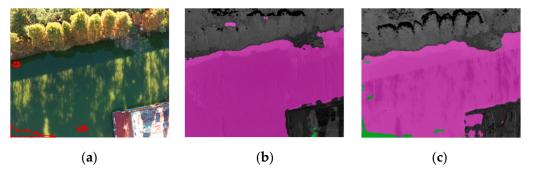


Figure 6. Detection results of different algorithms for shadow scene: (a) visible light image; (b) detection results of NDVI as the feature-extraction method; (c) detection results of $NDVI_{sqrt}$ as the feature-extraction method.

A comparison of the detection performance of several feature-extraction methods for cyanobacterial blooms in different test scenarios is shown in Table 2. The detection indicators of several algorithms in conventional scenarios were not significantly different, the results were all around 80%. The accuracy of the cyanobacterial-bloom detection algorithm used in this paper was much higher than the above two methods in scenarios with reflections and shadows, reaching 86.4% and 87.6%, respectively, and the IoUs were as high as 0.79 and 0.82. Compared with the vegetation index as the model input before improvement, the accuracy improved by 11.1% and 5.4%, and the IoU was 0.15 and 0.03 higher, respectively. The effectiveness of NDVI_{sqrt} was proven to extract features, so it was applied in subsequent deep-learning models.

Scenes —	NDVI		NDVI _{sqrt}	
	PA	IoU	PA	IoU
Regular	79.5%	0.71	81.6%	0.76
Reflections	75.3%	0.64	86.4%	0.79
Shadows	82.2%	0.79	87.6%	0.82

Table 2. Comparison of the detection performance of different feature-extraction methods for cyanobacterial blooms under various scenarios.

Among the more typical river channels tested, we studied the Daging River and the Haigeng Park river channel; these are more complex river scenarios, with interference from boats, birds and litter on the river surface, in addition to strong light and reflections. A comparison of the detection performance of several feature-extraction methods for cyanobacterial blooms in different typical river channels is shown in Table 3. The algorithm proposed in this paper maintained a more stable detection effect in all four river channels, unaffected by external interference factors, and the average accuracy of cyanobacterialbloom detection was above 85%, much higher than 70% before improvement. In Xujiang River, the detection accuracy increased from 85.3% to 92.3%, which extracted some information about cyanobacterial blooms in the early stage of growth. Moreover, in the Daqing River, Baoxiang River and Haigeng Park with different kinds of disturbances, the detection accuracy before and after using the improved vegetation index increased by 13.3%, 17.9% and 25.9%, respectively. The improvement in the intersection ratio showed that the overall classification task had a good effect, and the accuracy improvement caused by the misjudgment of the category was excluded. In general, the impressive improvement results support that NDVI_{sqrt} has great application value in practical scenarios.

River —	NDVI		NDVI _{sqrt}		
Kiver —	PA	IoU	PA	IoU	
Daqing River	76.1%	0.64	89.4%	0.76	
Baoxiang River	63.4%	0.58	81.3%	0.86	
Xujiang River	85.3%	0.85	92.3%	0.91	
Haigeng Park	56.4%	0.57	82.3%	0.76	

Table 3. Comparison of the performance of different feature-extraction methods for the detection of cyanobacterial blooms in typical river channels.

3.2. Detection Results

To further verify the effectiveness of the improved transformer model, we used the improved vegetation index for feature extraction and compared the detection results of the improved transformer model with those of FCN and SegNet.

Figure 7 shows a comparison of the detection results of the three deep-learning algorithms in conventional scenarios of the Xujiang River in Suzhou. There was a large number of misclassifications in the detection results of FCN, wherein the reflections on the water body, due to the fluctuations, were considered to be cyanobacteria, and some riverbanks were also mistaken for water bodies. The overall detection effect of FCN was very bad. Although there were relatively fewer false detections in SegNet than in FCN, the classification was not clear, with the boundaries of water bodies, cyanobacteria, and riverbanks all misclassified among each other. In comparison, the algorithm based on the improved transformer model was more accurate for cyanobacterial blooms and had almost no false detections.

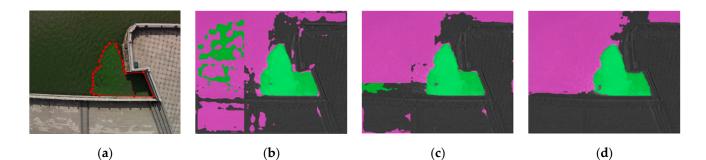


Figure 7. Detection results of different deep-learning algorithms for regular scene: (**a**) visible light images; (**b**) detection results of FCN; (**c**) detection results of SegNet; (**d**) detection results of improved transformer model.

Figure 8 shows a comparison of the detection results of the three deep-learning algorithms in the reflection scene of the Daqing River in Yunnan. Both FCN and SegNet had false and missed detections. FCN was more accurate in detecting the category of riverbanks, but some water bodies were judged to be cyanobacterial blooms, thus estimating a larger pollution area. Meanwhile, SegNet did not accurately distinguish riverbanks from clean water bodies, and some cyanobacterial blooms were identified as water bodies, causing the pollution area estimate to be smaller. The interference of reflections was severe, which caused the pixel value of the corresponding position to change so that the above two algorithms did not work correctly. In contrast, the algorithm proposed in this paper could effectively achieve accurate detection of cyanobacterial blooms in these edge areas, even under strong light interference.

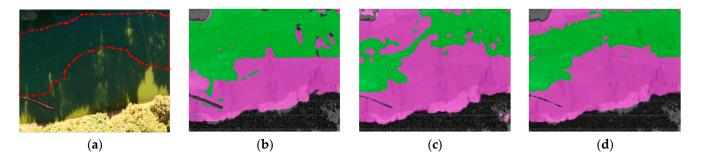


Figure 8. Detection results of different deep-learning algorithms for reflection scene: (**a**) visible light images; (**b**) detection results of FCN; (**c**) detection results of SegNet; (**d**) detection results of improved transformer model.

Figure 9 shows a comparison of the detection results of the three deep-learning algorithms in the shadowy scene of the Daqing River in Yunnan. The accuracy of FCN was slightly affected, and SegNet misidentified a large number of water bodies as cyanobacterial blooms. Such detection results were contrary to misjudgments caused by reflections; this could infer that FCN can exclude shadows interference, while SegNet can exclude reflection interference. Contrarily, the improved transformer model maintained high-accuracy detection in all scenarios. For objects of a certain color on the ship, all three algorithms misjudged, which was not an error caused by the boundary problem.

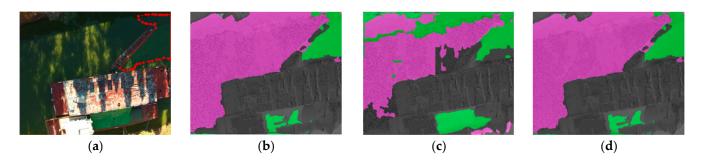


Figure 9. Detection results of different deep-learning algorithms for shadow scene: (**a**) visible light images; (**b**) detection results of FCN; (**c**) detection results of SegNet; (**d**) detection results of improved transformer model.

Both PA and IoU were still used here to quantitatively evaluate the detection performance of several image-segmentation algorithms. The comparison of the detection performance of cyanobacterial blooms in different test scenarios is shown in Table 4. The difference in detection performance in conventional scenarios was not significant, but the difference in complex scenarios was obvious. FCN was the least satisfactory, with a recognition accuracy of 61% and 59.3% under reflections and shadows, respectively. SegNet had a certain degree of false detection and missed detection at the boundaries of different features, and the detection accuracy was basically maintained at 60–65%. Compared with FCN and SegNet, the average detection accuracy of the improved transformer model in this paper reached over 85% and the IoU was around 0.8, while the detection accuracy levels of FCN and SegNet were only 64.1% and 67.3%.

Scenes —	FCN		Seg	SegNet		Improved Transformer	
	PA	IoU	PA	IoU	PA	IoU	
Regular	72%	0.75	75.4%	0.72	87.6%	0.82	
Reflections	61%	0.56	60.9%	0.62	81.6%	0.76	
Shadows	59.3%	0.62	65.7%	0.73	86.4%	0.79	

Table 4. Comparison of the detection performance of different deep-learning methods for cyanobacterial blooms in various scenarios.

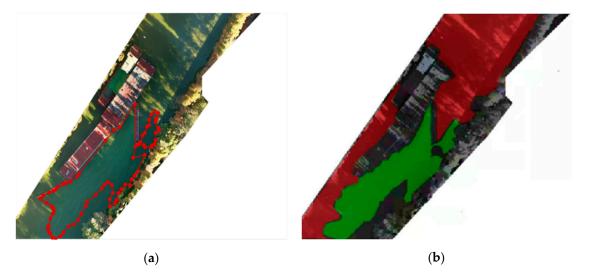
3.3. Estimation of Contaminated Area and Determination of Contamination Level

According to the quantitative assessment method of algal blooms in the Technical Specification for Remote Sensing and Ground Monitoring and Evaluation of Algal Blooms promulgated by the Ministry of Ecology and Environment of China in 2020 [37], this paper used the proportion to evaluate the degree of cyanobacterial-bloom pollution, which was classified into five levels from light to heavy: I, II, III, IV and V, as shown in Table 5.

Table 5. Grading standard of degree of cyanobacterial bloom based on the evaluation of pollution area ratio.

Levels	Proportion—P (%)	Characteristics
Ι	0	No cyanobacterial blooms
II	0 < P < 10	No obvious cyanobacterial blooms
III	$10 \le P < 30$	Mild cyanobacterial blooms
IV	$30 \le P < 60$	Moderate cyanobacterial blooms
V	$60 \le P \le 100$	Severe cyanobacterial blooms

In this paper, we used DJI SmartMap software to stitch together the Daqing River, Xujiang River, Jinji Lake and Baoxiang River to estimate the proportion, and judged the pollution level of cyanobacterial blooms. Taking the Daqing River as an example for



demonstration, Figure 10a shows the visible image of the Daqing River after stitching, and Figure 10b shows the detection results of the algorithm in this paper.

Figure 10. Daqing River Example of engineering application: (**a**) visible light image, the area surrounded by red dots is cyanobacterial blooms; (**b**) detection result of our algorithm.

As can be seen from Table 6, in the rivers with neatly shaped river surfaces, a single detection category and clearly distinguishable feature categories, such as the Xujiang River, both NDVI and the algorithm in this paper had good detection results whose detection errors were within acceptable limits. However, in scenes with complex scenarios wherein interference was serious and regional boundaries were unknown, the detection accuracy of the algorithm in this paper was much higher than that of NDVI. This was especially the case for the Daqing River, for which NDVI had a misjudgment rate of 18%, which directly led to the wrong judgment of pollution level from IV to V.

 Table 6. Estimated area of cyanobacterial-bloom contamination and determination of pollution level in typical rivers.

River	NDVI		Proposed Method		True Value	
	Percentage	Level	Percentage	Level	Percentage	Level
Daqing River	65.9%	V	48.7%	IV	47.9%	IV
Xujiang River	20.2%	III	21.8%	III	20.5%	III
Jinji Lake	91.6%	V	77.6%	V	79.4%	V
Baoxiang River	85.2%	V	79.6%	V	78.9%	V

4. Discussion

4.1. Low-Altitude Remote-Sensing Technology

The conventional method of field sampling combined with laboratory analysis has insufficient capabilities for detecting cyanobacterial blooms in large-scale river channels. Owing to the influence of cloud cover, effective data for lakes and rivers from remotesensing satellites are limited, and it is impossible to accurately monitor cyanobacterial blooms in small rivers in real time. However, low-altitude imaging remote-sensing technology can flexibly acquire multi-scene data according to demand. This technology does not require atmospheric correction of image data, but it will certainly magnify spatial interference such as shadows, lighting, and floating objects from sky light and the water's surface, so certain preprocessing work is required. Low-altitude remote-sensing technology also makes it possible to monitor daily changes in cyanobacterial blooms, which can provide more real-time information for early warnings and salvage treatments of cyanobacterial blooms.

4.2. Improved Vegetation Index

The spectral characteristics of water bodies are determined by the absorption and scattering properties of optical radiation. The chlorophyll content in water increases rapidly, resulting in changes in its spectral characteristics when cyanobacterial blooms occur. The absorption peaks and reflection peaks of different wavelength bands of reflection spectral characteristics are different, so the vegetation-index method aims to judge the cyanobacterial blooms and distinguish other objects by calculating the reflectivity matrix of specific bands. Some research groups have reported improved vegetation indices and applied them to different river scenarios [16–19]. Fernandez et al. [38] compared 26 vegetation indices and concluded that NDVI is effective and the most widely used method. However, it is difficult to have a universal index, because there are many factors that can cause changes in reflectivity, such as: sun position, sensor position, geographic location, seasons, climate changes, humidity changes, variability of the ground objects themselves, and atmospheric conditions. The selection of vegetation index is affected by the type of experimental data. As shown in the results section, $NDVI_{sqrt}$ has an accuracy of more than 85%, which is 5% higher than the 79.60% of the reference [39]. Compared to other indices, the improved vegetation index is more suitable for the feature extraction of cyanobacterial blooms in this paper.

4.3. Improved Transformer Model

Deep learning is a powerfully adaptive method with many achievements in image classification. It is necessary to introduce semantic segmentation for cyanobacterial-bloom detection, because cyanobacterial blooms are mostly granular, appearing as flocculent suspended matter, and the distribution shape is irregular. The detection accuracy of reference [29] is 93.68%, but the scene is simple and clear. Reference [30] achieves an average accuracy of 86.70% under complex lighting conditions. The improved transformer model has a test accuracy of 89.4% in the Daqing river with serious interference, which shows the effectiveness of the proposed algorithm in this paper.

As shown in Figure 4, this paper also tests the common semantic segmentation models CNN and DNN. FCN is insensitive to details and the size of the pixel blocks is much smaller than the whole image, which is not fully applicable to unstructured data. SegNet ignores adjacent information, and image resolution is reduced during network processing; thus, features are not sufficiently trained, causing interference to the contextual semantic understanding. To enhance the understanding of contextual semantics and exploit the multispectral features of images more effectively, the transformer model based on feature enhancement and region correction in this paper pays attention to the associations between adjacent pixels, and uses a multi-headed attention mechanism, which improves the classification accuracy of cyanobacterial-bloom boundary regions.

5. Conclusions

This work introduced a multispectral detection model for cyanobacterial blooms based on deep learning. Images were preprocessed with adaptive median filtering and light masks. The traditional vegetation-index method was improved to extract the multispectral features of cyanobacterial blooms, and then the transformer neural network was used for the first time to perform pixel-level semantic recognition of cyanobacterial blooms. Finally, the algorithm was applied to multiple rivers to accurately evaluate their pollution area and pollution levels. This method not only greatly reduced the sampling time, but also improved the accuracy and efficiency of the cyanobacterial-bloom detection algorithm, providing the possibility for online monitoring of cyanobacterial blooms. The research results of this paper contribute to the fields of water-quality anomaly detection, remotesensing technology and deep learning. On the one hand, the improved vegetation index showed a better feature extraction effect than other indices in the dataset of this paper, and may expand the applicability of vegetation indices in the field of low-level multispectral cyanobacterial-bloom monitoring. On the other hand, the transformer neural network

acterial blooms, and showed good results

was first applied in the detection of cyanobacterial blooms, and showed good results when modifying the corresponding network structure. The work of this paper has certain enlightenment significance for the application of the transformer model in the multispectral detection of cyanobacterial blooms, and provids strong support for the application of neural networks in the field of environmental monitoring.

For future work, research on preprocessing methods can be introduced to eliminate insignificant environmental interference before feature extraction. In addition, higher sampling frequency may reveal the trends of cyanobacterial blooms, and a corresponding early-warning model, with the help of time series analysis, is recommended to develop a method of predicting future growth trends of cyanobacterial blooms; this is significant for early warnings and the rapid treatment of cyanobacterial blooms.

Author Contributions: Conceptualization, D.H. and Z.W. (Zhengfang Wu); methodology, W.X. and Z.S.; validation, H.D. and Z.W. (Zhongyi Wang); formal analysis, X.W. and Z.W. (Zhongyi Wang); investigation, Z.S. and Z.W. (Zhengfang Wu); resources, D.H.; writing—original draft preparation, Z.S.; writing—review and editing, Y.C.; supervision, Y.C.; project administration, P.H.; funding acquisition, D.H., Z.W. (Zhengfang Wu) and Z.W. (Zhongyi Wang). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Technology Research and Development Program of Zhejiang Province (2021C03177, 2022C03078), the National Key R&D Program of China (2017YFC1403800), and the National Natural Science Foundation of China (61803333).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

- Winslow, L.A.; Leach, T.H.; Rose, K.C. Global lake response to the recent warming hiatus. *Environ. Res. Lett.* 2018, 13, 054005. [CrossRef]
- Qin, B.; Paerl, H.W.; Brookes, J.D.; Liu, J.; Jeppesen, E.; Zhu, G.; Zhang, Y.; Xu, H.; Shi, K.; Deng, J. Why Lake Taihu continues to be plagued with cyanobacterial blooms through 10 years (2007–2017) efforts. *Sci. Bull.* 2019, *64*, 354–356. [CrossRef]
- 3. Huang, C.; Wang, X.; Yang, H.; Li, Y.; Wang, Y.; Chen, X.; Xu, L. Satellite data regarding the eutrophication response to human activities in the plateau lake Dianchi in China from 1974 to 2009. *Sci. Total Environ.* **2014**, *485*, 1–11. [CrossRef] [PubMed]
- Paerl, H.W.; Xu, H.; McCarthy, M.J.; Zhu, G.; Qin, B.; Li, Y.; Gardner, W.S. Controlling harmful cyanobacterial blooms in a hyper-eutrophic lake (Lake Taihu, China): The need for a dual nutrient (N & P) management strategy. *Water Res.* 2011, 45, 1973–1983.
- 5. Michalak, A.M.; Anderson, E.J.; Beletsky, D.; Boland, S.; Bosch, N.S.; Bridgeman, T.B.; Chaffin, J.D.; Cho, K.; Confesor, R.; Daloğlu, I.; et al. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6448–6452. [CrossRef]
- 6. Binding, C.E.; Greenberg, T.A.; Bukata, R.P. Time series analysis of algal blooms in Lake of the Woods using the MERIS maximum chlorophyll index. *J. Plankton. Res.* 2011, *33*, 1847–1852. [CrossRef]
- Sun, D.Y.; Hu, C.M.; Qiu, Z.F.; Shi, K. Estimating phycocyanin pigment concentration in productive inland waters using Landsat measurements: A case study in Lake Dianchi. *Opt. Express* 2015, 23, 3055–3074. [CrossRef]
- Zhang, Y.L.; Jeppesen, E.; Liu, X.H.; Qin, B.Q.; Shi, K.; Zhou, Y.Q. Global loss of aquatic vegetation in lakes. *Earth-Sci. Rev.* 2017, 173, 259–265. [CrossRef]
- 9. Juttner, F.; Watson, S.B. Biochemical and ecological control of geosmin and 2-methylisoborneol in source waters. *Appl. Environ. Microb* 2007, 73, 4395–4406. [CrossRef]
- 10. Shi, K.; Zhang, Y.L.; Qin, B.Q.; Zhou, B.T. Remote sensing of cyanobacterial blooms in inland waters: Present knowledge and future challenges. *Sci. Bull.* **2019**, *64*, 1540–1556. [CrossRef]
- 11. Qin, B.Q.; Li, W.; Zhu, G.W.; Zhang, Y.L.; Wu, T.F.; Gao, G. Cyanobacterial bloom management through integrated monitoring and forecasting in large shallow eutrophic Lake Taihu (China). *J. Hazard. Mater.* **2015**, *287*, 356–363. [CrossRef] [PubMed]
- 12. Hu, B.Y.; Zhao, L.; Zhou, W.J.; Wang, Y.W. Current Situation, Potential Development Issues and Countermeasures of Bloomforming Cyanobacteria in China. *J. Hydrol.* **2012**, *33*, 138–143.
- 13. Huang, S.; Tang, L.N.; Hupy, J.P.; Wang, Y.; Shao, G.F. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *J. For. Res.* **2021**, *32*, 1–6. [CrossRef]

- Kaufman, Y.J.; Tanre, D. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Trans. Geosci. Remote Sens.* 1992, 30, 261–270. [CrossRef]
- 15. Zhang, Y.L.; Song, C.H.; Band, L.E.; Sun, G.; Li, J.X. Reanalysis of global terrestrial vegetation trends from MODIS products: Browning or greening? *Remote Sens. Environ.* **2017**, *191*, 145–155. [CrossRef]
- 16. Hu, C.M. A novel ocean color index to detect floating algae in the global oceans. *Remote Sens. Environ.* **2009**, *113*, 2118–2129. [CrossRef]
- Huang, C.; Shi, K.; Yang, H.; Li, Y.; Zhu, A.-X.; Sun, D.; Xu, L.; Zou, J.; Chen, X. Satellite observation of hourly dynamic characteristics of algae with Geostationary Ocean Color Imager (GOCI) data in Lake Taihu. *Remote Sens. Environ.* 2015, 159, 278–287. [CrossRef]
- Cannizzaro, J.P.; Barnes, B.B.; Hu, C.M.; Corcoran, A.A.; Hubbard, K.A.; Muhlbach, E.; Muhlbach, E.; Sharp, W.C.; Brand, L.E.; Kelble, C.R. Remote detection of cyanobacteria blooms in an optically shallow subtropical lagoonal estuary using MODIS data. *Remote Sens. Environ.* 2019, 231, 111227. [CrossRef]
- 19. Mishra, S.; Stumpf, R.P.; Schaeffer, B.; Werdell, P.J.; Loftin, K.A.; Meredith, A. Evaluation of a satellite-based cyanobacteria bloom detection algorithm using field-measured microcystin data. *Sci. Total Environ.* **2021**, 774, 145462. [CrossRef]
- Cillero Castro, C.; Domínguez Gómez, J.A.; Delgado Martín, J.; Hinojo Sánchez, B.A.; Cereijo Arango, J.L.; Cheda Tuya, F.A.; Díaz-Varela, R. An UAV and Satellite Multispectral Data Approach to Monitor Water Quality in Small Reservoirs. *Remote Sens.* 2020, 12, 1514. [CrossRef]
- Barruffa, A.S.; Pardo, A.; Faggian, R.; Sposito, V. Monitoring cyanobacterial harmful algal blooms by unmanned aerial vehicles in aquatic ecosystems. *Environ. Sci.-Water Res.* 2021, 7, 573–583.
- Sharp, S.L.; Forrest, A.L.; Bouma-Gregson, K.; Jin, Y.F.; Cortes, A.; Schladow, S.G. Quantifying Scales of Spatial Variability of Cyanobacteria in a Large, Eutrophic Lake Using Multiplatform Remote Sensing Tools. *Front. Environ. Sci.* 2021, 9, 612934. [CrossRef]
- Shelhamer, E.; Long, J.; Darrell, T. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017, 39, 640–651. [CrossRef]
- 24. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]
- Cai, Y.; Huang, X.G.; Zhang, Z.A.; Zhu, X.N.; Ma, X. Real-Time semantic segmentation algorithm based on feature fusion technology. *Laser Optoelectron. Prog.* 2020, 57, 10435.
- 26. Hou, Y.N.; Ma, Z.; Liu, C.X.; Loy, C.C. Learning lightweight lane detection cnns by self attention distillation. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October–2 November 2019.
- Cheng, O.Y.; Carlo, B.; Chen, C.; Turkay, K.; Qiu, H.Q.; Daniel, R. Self-supervision with superpixels: Training few-shot medical image segmentation without annotation. In Proceedings of the ECCV: European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020.
- Mina, R.; Yang, H.J.; Christoph, M. Recurrent generative adversarial network for learning imbalanced medical image semantic segmentation. *Multimed. Tools. Appl.* 2020, 79, 15329–15348.
- 29. Yang, S.; Chen, L.F.; Shi, Y.; Mao, Y.M. Semantic segmentation of blue-green algae based on deep generative adversarial net. *J. Comput. Appl.* **2018**, *38*, 1554–1561.
- 30. Luo, A.N.; Zheng, J. Semantic Segmentation of Cyanoacteria Pictures Based on DCNN Under Complicated Illumination. *Comput. Appl.Softw.* **2018**, *35*, 254–259.
- 31. Xiang, H.Y.; Zou, B.; Tang, L.; Chen, W.G.; Rao, K.F.; Liu, Y.; Ma, M.; Yang, Y. Phytoplankton recognition based on residual attention network. *Acta Ecol. Sin.* **2021**, *41*, 6883–6892.
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st Annual Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.
- Ahn, J.; Kwak, S. Learning pixel-level semantic affinity with image-level supervision for weakly supervised semantic segmentation. In Proceedings of the 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018.
- 34. Zhang, A.; Wang, K.C.P.; Fei, Y.; Liu, Y.; Chen, C.; Yang, G.; Li, J.Q.; Yang, E.; Qiu, S. Automated pixel-pevel pavement crack detection on 3D asphalt surfaces with a recurrent neural network. *Comput.-Aided Civ. Inf.* **2019**, *34*, 213–229. [CrossRef]
- 35. Sun, H.; Zheng, X.T.; Lu, X.Q.; Wu, S.Y. Spectral-spatial attention network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 3232–3245. [CrossRef]
- Cai, Z.W.; Vasconcelos, N. Cascade R-CNN: High quality object detection and instance segmentation. *IEEE Trans. Pattern. Anal. Mach. Intell.* 2021, 43, 1483–1498. [CrossRef] [PubMed]
- Technical Specifications for Monitoring and Evaluating Algal Bloom Based on Remote Sensing and Field Monitoring. Available online: https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/other/qt/202002/t20200213_762889.shtml (accessed on 12 April 2020).
- 38. Fernandez-Figueroa, E.G.; Wilson, A.E.; Rogers, S.R. Commercially available unoccupied aerial systems for monitoring harmful algal blooms: A comparative study. *Limnol. Oceanogr.-Meth.* **2022**, *20*, 146–158. [CrossRef]
- Zhao, H.; Li, J.S.; Wang, C. Monitoring Cyanobacteria Bloom in Dianchi Lake Based on Ground-Based Multispectral Remote-Sensing Imaging: Preliminary Results. *Remote Sens.* 2021, 13, 3970. [CrossRef]