Article

A Modeling Method for Automatic Extraction of Offshore Aquaculture Zones Based on Semantic Segmentation

Baikai Sui, Tao Jiang, Zhen Zhang *, Xinliang Pan and Chenxi Liu

College of Geomatics, Shandong University of Science and Technology, 266590 Qingdao, China;
suibaikai@yujiangtech.cn (B.S.); jiangtao@sdust.edu.cn (T.J.); panxinliang@yujiangtech.cn (X.P.);
liuchenxi@yujiangtech.cn (C.L.)

* Correspondence: zhang_L08@lzu.edu.cn

check for
updates

Abstract: Monitoring of offshore aquaculture zones is important to marine ecological environment protection and maritime safety and security. Remote sensing technology has the advantages of large-area simultaneous observation and strong timeliness, which provide normalized monitoring of marine aquaculture zones. Aiming at the problems of weak generalization ability and low recognition rate in weak signal environments of traditional target recognition algorithm, this paper proposes a method for automatic extraction of offshore fish cage and floating raft aquaculture zones based on semantic segmentation. This method uses Generative Adversarial Networks to expand the data to compensate for the lack of training samples, and uses ratio of green band to red band (G/R) instead of red band to enhance the characteristics of aquaculture spectral information, combined with atrous convolution and atrous space pyramid pooling to enhance the context semantic information, to extract and identify two types of offshore fish cage zones and floating raft aquaculture zones. The experiment is carried out in the eastern coastal waters of Shandong Province, China, and the overall identification accuracy of the two types of aquaculture zones can reach 94.8%. The results show that the method proposed in this paper can realize high-precision extraction both of offshore fish cage and floating raft aquaculture zones.

Keywords: offshore aquaculture; semantic segmentation; generative adversarial networks; high-resolution remote sensing image

1. Introduction

The aquaculture industry has developed rapidly, and aquaculture zones in coastal zones have been expanding globally. This development has brought about huge economic benefits but also negative impacts on the local offshore ecological environment and sea transportation [1]. Therefore, the timely monitoring of offshore aquaculture status is important to marine environmental protection, maritime safety, and coastal engineering construction. With the rapid development of remote sensing technology, the spatial resolution of images has continuously improved [2], thereby providing an effective means for regular monitoring of marine aquaculture. Two types of common offshore aquaculture are available. The first type is floating raft aquaculture [3], which is a long-line system composed of a floating raft with a float and rope on the surface of the shallow sea, and fixed to the bottom with a cable. This structure breeds seafood, such as kelp, seaweed, and mussels. This kind of aquaculture is dark in remote sensing images. The second type is a fish cage [3], which is composed of wood and plastic materials. This structure is used for breeding abalone, sea cucumber, and other seafood. The cage is suspended on the sea surface, and the bottom is sunk into the water to a depth of 3–6 m. This kind of aquaculture shows up as bright colors in remote sensing images.

For remote sensing, feature extraction algorithms are generally categorized into three types: traditional classification method based on statistics, advanced classification methods, and deep learning. Traditional classification methods based on spectral statistical characteristics include maximum likelihood [4], minimum distance [5], and k-means clustering [6]; in the classification of low- and medium-resolution remote sensing images, remarkable achievements have been made. However, these methods cause excessive misclassification and missing classification, which lead to difficulty in meeting the requirements of the classification of high-resolution remote sensing images. Advanced classification methods include BP neural network [7,8], support vector machine [9,10], and genetic algorithm [11,12]. Compared with traditional statistical methods, this type of algorithm improves the accuracy of ground object recognition to a certain extent. Due to the limitations of the above classification methods' learning structure, establishing complex functions is difficult [13]. Hence, these methods are not suitable for complex samples and have poor generalization ability. A deep convolutional neural network (DCNN) [14] was developed on the basis of neural networks. Due to the evident advantages of DCNN in fully mining the deeper information of data and processing complex samples, this method is widely applied in remote sensing image classification [15,16].

In extraction research on offshore aquaculture, the data sources are mainly divided into optics and synthetic aperture radar (SAR). In optics, Ma et al. extracted aquaculture zones according to the spectral characteristics of ASTER remote sensing images by constructing the water index of aquaculture zones and achieving an extraction accuracy of 86.14% [17]. Zhang et al. studied the method of using TM images to automatically draw the aquaculture map of coastal zones and used multi-scale segmentation and object relation modeling strategy (MSS/ORM) to extract the aquaculture area of TM images, which improved the classification accuracy [18]. Lu et al. established rapid detection of the spectral characteristic index in offshore aquaculture zones by using statistical average texture and threshold detection algorithm combined with offshore shape aquaculture zones in 2015 [19]. In SAR, the radar can penetrate clouds, rain, and snow and is less affected by weather. SAR images have rich polarization information. Fan et al. proposed a joint sparse representation classification method, which uses high-resolution SAR satellite remote sensing data to quickly and accurately obtain the breeding range and area of floating rafts [20]. Geng et al. proposed a deep cooperative sparse coding network (DCSCN) for ocean floating raft recognition, which effectively suppresses the influence of speckle noise and improves the target recognition accuracy of SAR images [21].

The above studies are mostly based on spectral or texture features, but the floating objects may be reticulated and contain much seawater information in high spatial resolution images. This kind of noise for the extraction task will seriously affect the accuracy of the extraction in aquaculture. At the same time, when too many suspended impurities are in the water, the background water easily confuses the extraction of floating raft aquaculture area, which seriously affects the accuracy of the algorithm [17–19].

Semantic segmentation is based on DCNN. The fully connected layer of DCNN is removed and upsampled to the same size of the input image to complete the end-to-end learning. On the basis of image empty spectrum and texture information, contextual information is fully considered, showing a strong classification ability. Currently, the excellent semantic segmentation algorithms include FCN [22], PSPNet [23], Segnet [24], and DeepLab series [25–28]. This study designed a deep network model on the basis of DeepLab V3 that can identify offshore farming in the east coastal zone of Shandong Province, China. The experimental results show that the method proposed in this paper achieved good results in the extraction of floating raft and fish cage aquaculture in offshore aquaculture zones.

The rest of the paper is structured as follows. The second part introduces the experimental methods proposed in this paper. The third part introduces the experiments and results of this paper. The fourth and the fifth parts are, respectively, the discussion and summary.

2. Materials and Methods

This paper proposes an automatic extraction method for offshore aquaculture based on DeepLab V3 [27], which includes data processing, model training, prediction extraction of aquaculture, and accuracy evaluation. The proposed method is called OAE-V3.

2.1. Data Processing

1. Band combination and normalization. As the red band is strongly absorbed in water, the ratio band (G/R) was adopted in this paper to replace the red band (R) and stretched to 0–255, which was reconstructed and normalized with the G and B bands.

2. Label making. Image processing software was used to calibrate manually the image feature categories, which were 0—background, 1—fish net cage aquaculture, and 2—floating raft aquaculture.

3. Image cropping. To prevent the model from being unable to train due to insufficient GPU memory during the training process, the image and its ground truth value map are regularly cut according to the pixel coordinate position. The size of the trained training data is 256 × 256, as shown in Figure 1.
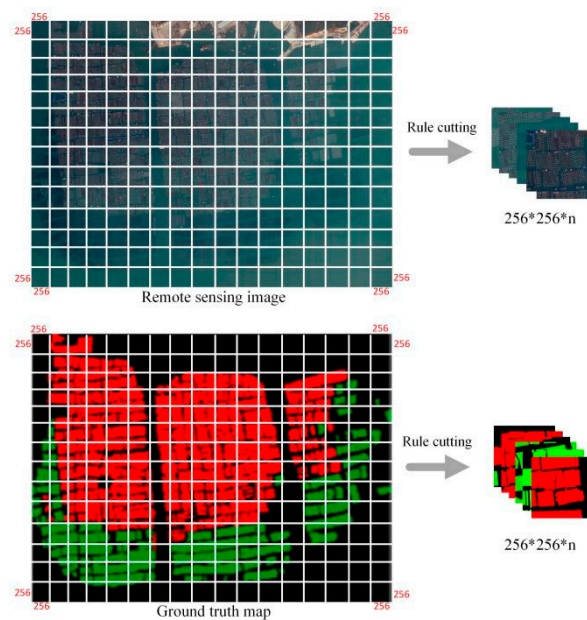
Figure 1. Image clipping. Each remote sensing image and corresponding label are cut into the same number of samples of the same size (256 × 256).

4. Data expansion. Remote sensing images are different from natural photos. Given the different shooting angles, divergent image presentation states of ground objects and limited training sample data and image processing should be expanded to prevent overfitting of the training model and enhance the generalization ability of the model. The data expansion methods used in this paper are as follows:

(1)   Ordinary image expansion: The image rotation (60°, 90°, 120°, and other angles) and adding random Gaussian noise into the image showing the expansion result is presented in Figure 2.

(a) original   (b) rotate 90°   (c) rotate 180°   (d) add gaussian noise
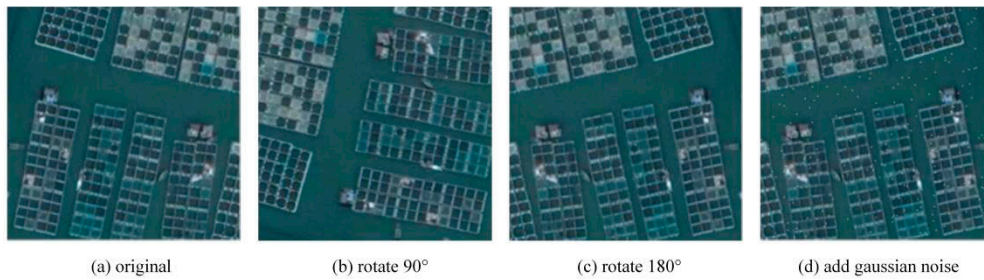
Figure 2. (**a**) The original image; (**b**) the image is obtained by rotating 90° counterclockwise; (**c**) the image is obtained by rotating 180° counterclockwise; (**d**) the image has added Gaussian noise and is rotated 180°.

GAN image expansion: The central idea of GANs [29] is to learn from existing data through the network and then generate similar data. In the process of generation, the discrimination and the generation networks are against each other until the generated image is realistic. On the basis of this idea, this paper proposed to use a condition generation network [30] to generate images for data expansion, in which the generator uses the UNET [31] network for reference for only down- and upsampling, and the discriminator consists of convolution and LeakyRelu activation layers. Figures 3 and 4 show the network framework and generated image, respectively.
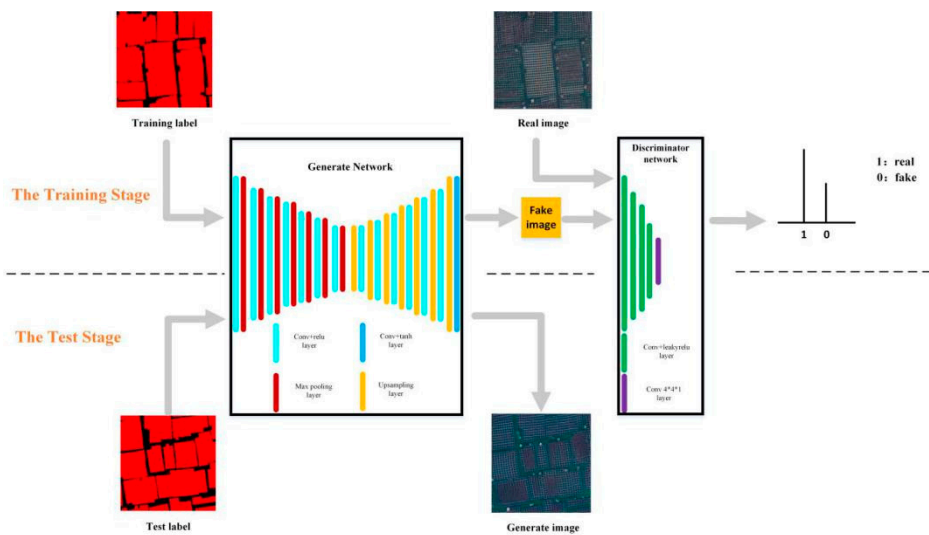
Figure 3. Condition GAN data extension process. The whole process is divided into two stages. The upper half of the dotted line is the training phase. Train and save the model. The lower half of the dotted line is the test phase, generating a new image.

(a) Real image
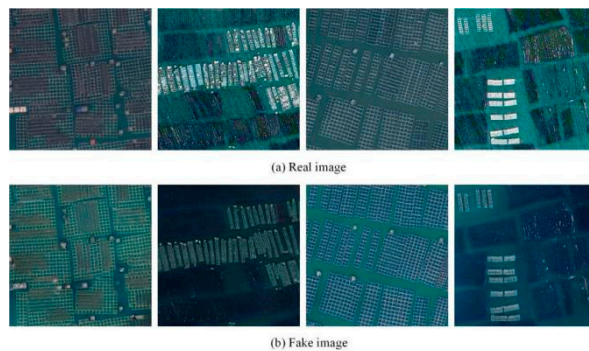
(b) Fake image

Figure 4. (**a**) Real image; (**b**) GAN-generated fake image.

2.2. Model Training

This section is divided into two parts. The first part introduces the OAE-V3 network structure, and the second part presents the training process of the OAE-V3 model.

2.2.1. OAE-V3 Network

The OAE-V3 model proposed in this paper is a recognition model of offshore aquaculture based on DeepLab V3 [27]. The network is mainly composed of three parts:

(1) Resnet network. The main idea of Resnet is deep residual network [32], as shown in Figure 5. The residual structure adds features (F(X)) after the direct cross-level input (x) and the output of the convolutional layer.

(2) Atrous convolution. In image semantic segmentation, the convolution neural network [33] extracts features by pooling layers to reduce the image scale, which would increase the receptive field. The final images with smaller sampling operation will need to restore the original size. This situation creates a problem—that is, the pooling operation could lose many details. To solve this problem, atrous convolution was introduced to the field of image segmentation [34]. The so-called atrous sampling is based on the original image, and the sampling frequency is set according to the rate parameter (atrous size). When rate = 1, the convolution operation is the standard convolution operation, as shown in Figure 6a. When rate > 1, sampling every pixel is done at the rate on the original image. Figure 6b shows the convolution operation when rate = 2.

(3) Atrous space pyramidization pooling. ASPP uses atrous convolution with different sampling rates and batch normalization [35] to form an atrous convolution cascade structure, which can effectively capture multiscale information.
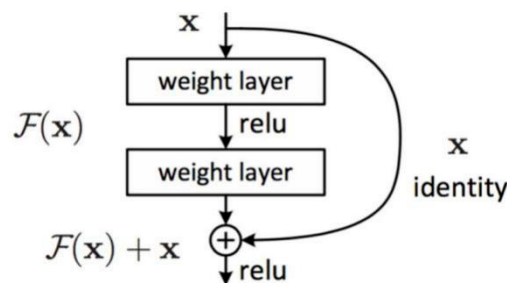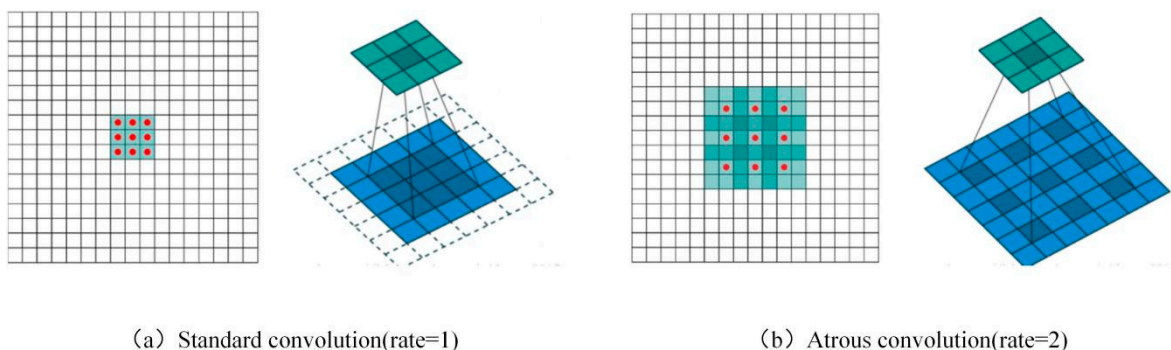

Figure 5. Residual structure.

（a）Standard convolution(rate=1) 　　　　　　　　（b）Atrous convolution(rate=2)

Figure 6. (**a**) Standard convolution form (rate = 1); (**b**) atrous convolution form (rate = 2).

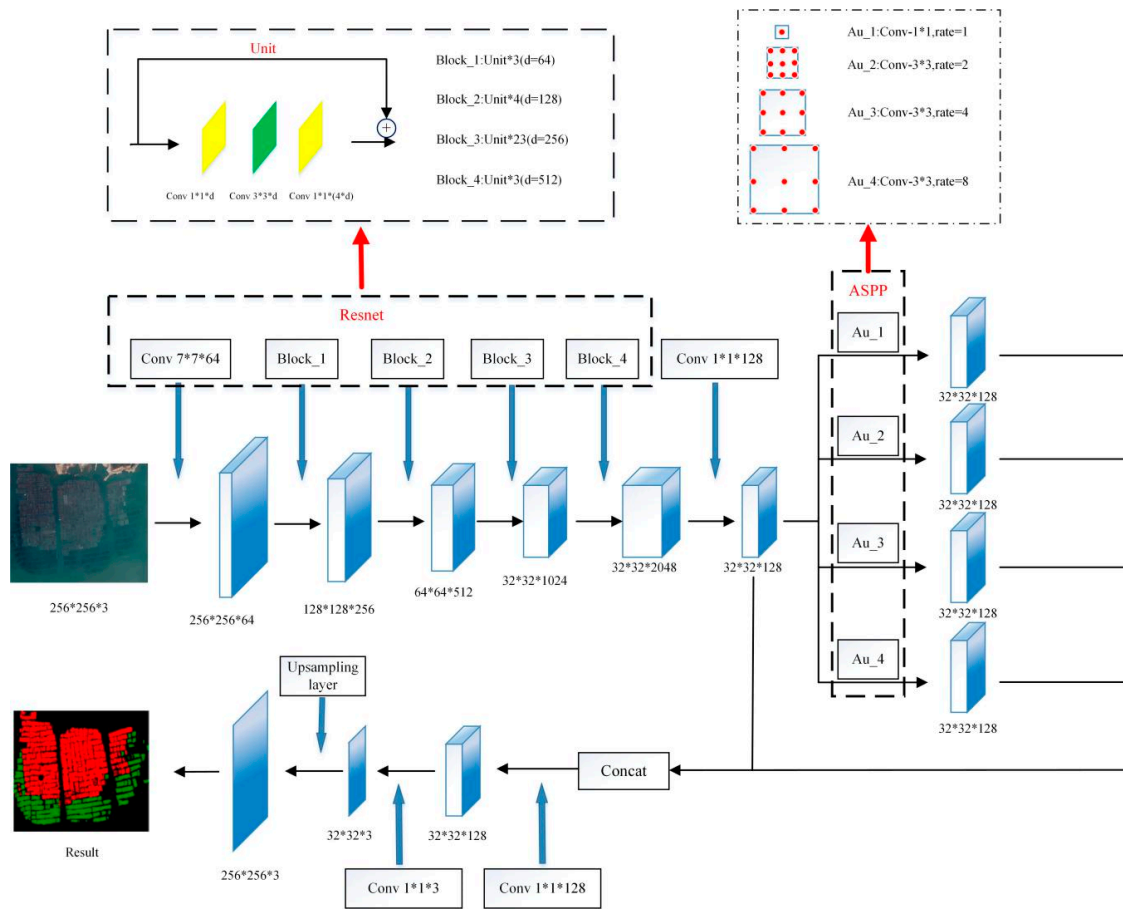Figure 7 shows the OAE-V3 network structure.

Figure 7. OAE-V3 network structure framework.

2.2.2. Training

First, the training dataset is used as input to extract the feature map through the OAE-V3 Resnet network. After that, the last layer of the Resnet network is convolved through a convolution layer and inputted into the ASPP structure. The ASPP used in this paper consisted of four layers of atrous convolutional layers (rates of 1, 2, 4, and 8). Finally, the four atrous convolutional layers of ASPP were connected in series and inputted to the next convolutional layer to obtain the output characteristic graph.

The last layer is the classifier. After convolving the output feature graph, an argmax function is executed to obtain the classification result of each pixel in the sample, and the input with the label is done with the cross-entropy loss function to calculate the generation value (loss) of the sample.

The Adam optimizer based on the gradient descent algorithm is used to update the network parameters continuously and save the network parameters when the model is optimal. In the training process, the model adopts a small batch training strategy, which greatly reduces the training time. L2 regularization and drop-out function are used in the model. Figure 8 shows the entire training process.
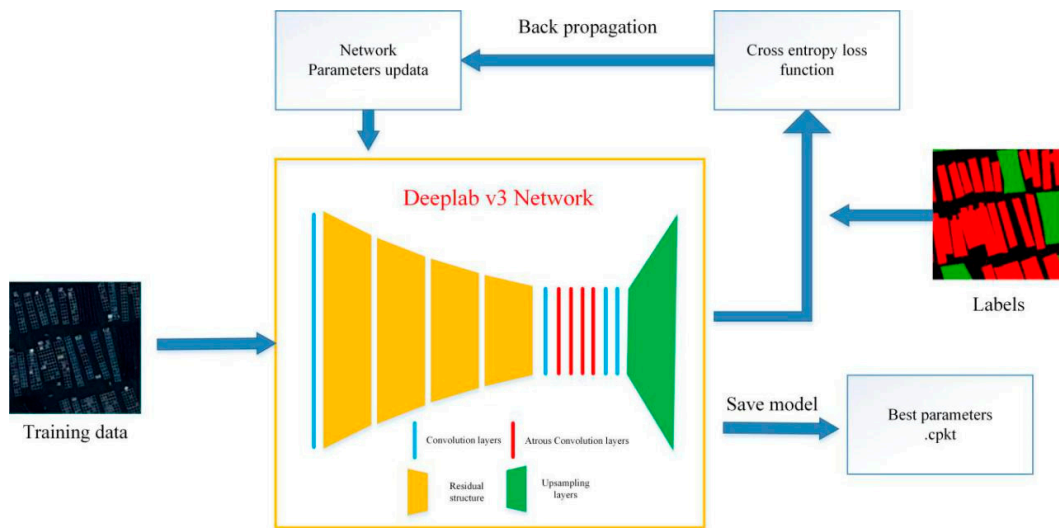
Figure 8. The training flow chart of OAE-V3 model.

2.3. Prediction Extraction of Aquaculture

Test images can be of any size, and this article will maintain the extracted image data uniform cutting for a size of 512×512 images into the training model. Finally, all test images are obtained by the argmax function of pixel level classification. The final image is obtained on the basis of the classification results of the slice in accordance with the pixel coordinates that merge to complete regional classification results. Figure 9 presents the prediction process.
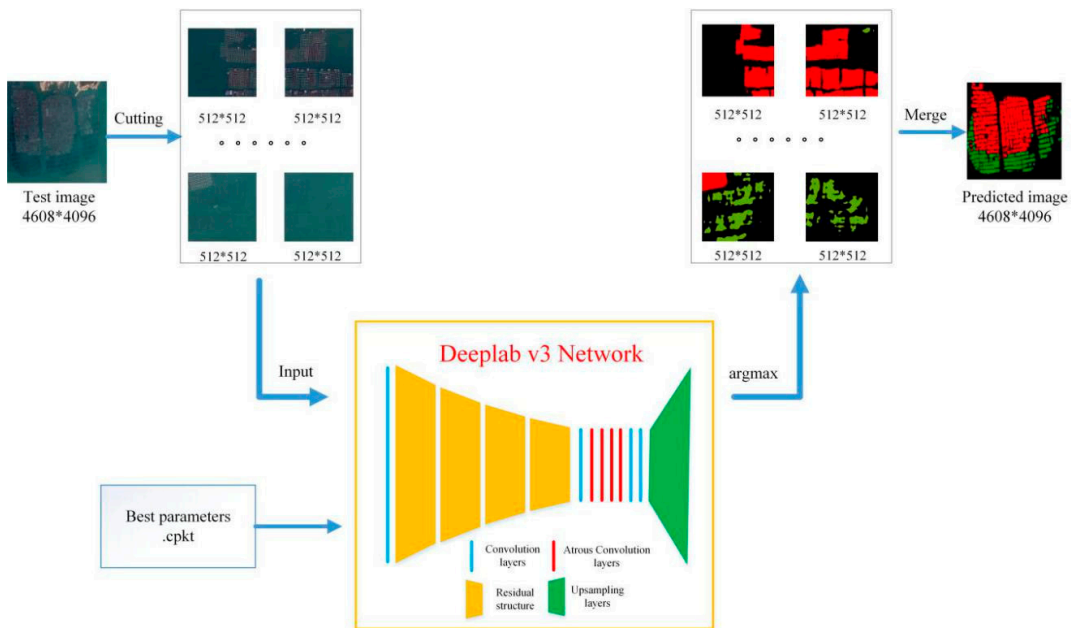

Figure 9. The prediction flow chart of OAE-V3 model.

2.4. Prediction Evaluation

The evaluation indexes of extraction accuracy in this paper include overall pixel classification accuracy, classification accuracy of each class, F1 score, and Kappa coefficient. The four indices can be calculated by the confusion matrix. In the following formula, N represents the total number of pixels, $n = 2$ (n refers to the number of categories), x_{ii} represents the number of pixels correctly classified

for each category, x_{+i} represents the number of pixels predicted for each category, and x_{i+} presents the number of pixels of real value for each category (where i = 0, 1, 2).

The overall pixel classification accuracy is the ratio of the number of correctly classified samples to the total number of samples. The calculation method is expressed as in Equation.

$$\text{pre\_pixel} = \frac{\sum_{i=0}^{n} x_{ii}}{N} * 100\%$$

Accuracy P measures the classifier's ability not to misclassify true negative samples into positive samples. The calculation method is expressed as in Equation.

$$P = \sum_{i=0}^{n} \frac{x_{ii}/x_{i+}}{x_{i+}} * 100\%$$

Accuracy R measures the ability of the classifier to find all positive samples. The calculation method is expressed as in Equation.

$$R = \sum_{i=0}^{n} \frac{x_{ii}/x_{+i}}{x_{+i}} * 100\%$$

F score is the weighted harmonic mean of accuracy and accuracy. F1 means that both values are equally important. The calculation formula is expressed as in Equation.

$$F1 = \frac{2 * P * R}{P + R} * 100\%$$

The Kappa coefficient is an index for determining the degree of coincidence or precision between two images. The closer the coefficient is to 1, the better the classification effect. The statistical method is expressed as in Equation.

$$K = \frac{N \sum_{i=0}^{n} x_{ii} - \sum_{i=0}^{n} (x_{i+} x_{+i})}{N^2 - \sum_{i=0}^{n} (x_{i+} x_{+i})}$$

2.5. Data

The experimental zone of this paper comprises three different coastal zones of Yantai and Weihai in Shandong Province, China. Each zone has two different kinds of breeding zones: a fish row cage and floating raft breeding zones. The data used in this paper are Unmanned Aerial Vehicle (UAV) aerial photography and Quick Bird satellite data, which were acquired on May 23, 2019 with a spatial resolution of 1.2 m. The sizes of the three images are 4608 × 4096, and the images and visual interpretation of the ground truth map are shown in Figure 10.
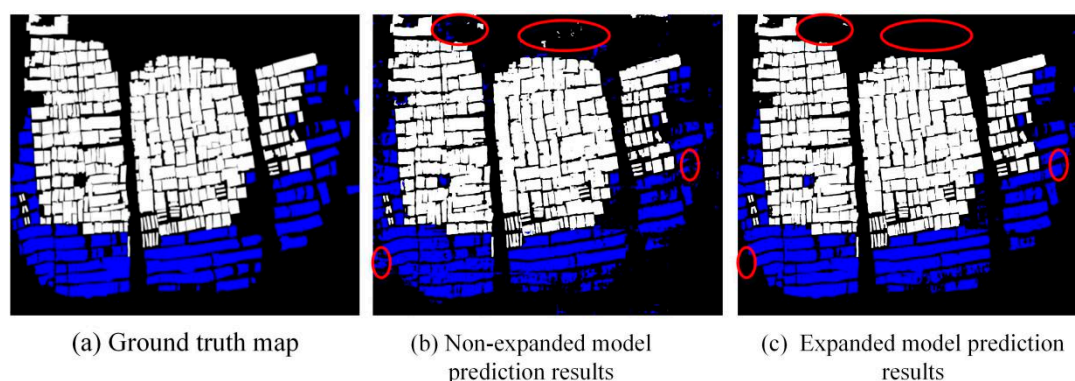
(a) Research area (b) Remote sensing images (c) Ground truth map
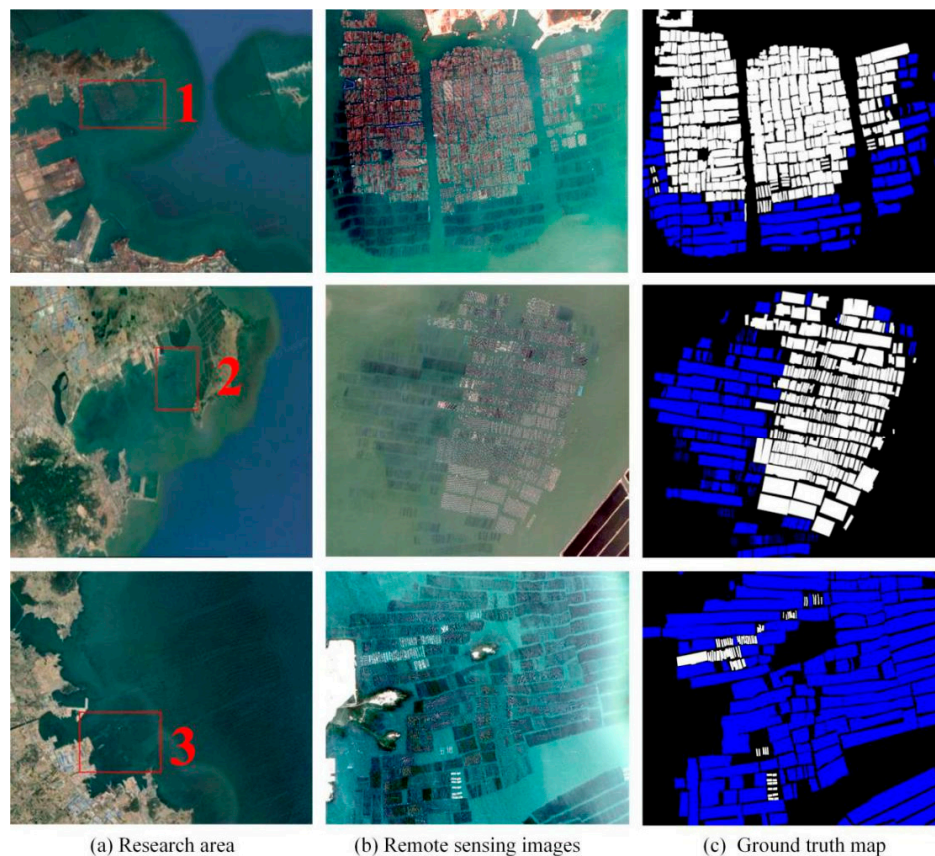
Figure 10. (**a**) The selected part in the red box is the study area; (**b**) remote sensing images corresponding to the study area; (**c**) ground truth map, in which black is the background, white is fish row cage aquaculture, and blue is floating raft aquaculture.

In the study zones, 864 images with the size of 256 × 256 are obtained through regular cutting and are randomly shuffled. Among the samples, 200 small images are taken for data expansion to 2400 as training data, and the remaining 664 small images are taken as verification data. The learning rate of the model is set as 0.0001. The training iterated for a total of 30,000 times. From the method, 32 samples are randomly selected from 2400 samples for each iteration for training.

3. Results and Discussion

3.1. Analysis of the Data Expansion

In the training process, first, the data expansion experiments, including ordinary rotation, noise addition, and generation of "false" images are based on GAN network. Figure 11 is the classification result of the OAE-V3 model before and after data expansion in study area 1. Compared with the results before the expansion, the model extracted after the expansion has higher recognition rate and fewer noise spots. The model trained with the expanded data is obviously better than that without.

<div align="center">(a) Ground truth map　　　(b) Non-expanded model　　　(c) Expanded model prediction
prediction results　　　　　　results</div>

Figure 11. (**a**) The ground truth map; (**b**) the prediction result of non-expanded model; (**c**) the prediction result of expanded model (black is the background, white is fish row cage aquaculture, and blue is floating raft aquaculture).

3.2. Analysis of the Bands

In addition, the band analysis of remote sensing images is carried out and revealed that the spectral information of fish row cage aquaculture zone is clearly different from other categories in the R, G, and B bands. However, the spectral information of floating raft aquaculture in the R band is similar to that of seawater, which is difficult to distinguish, but is the easiest to distinguish in the G band. Therefore, this paper removes and replaces the R band with ratio band G/R to select the data composed of G, B, and G/R bands for training and prediction. In this paper, a small area with more floating raft aquaculture is selected for classification and comparison, as shown in Figure 12. The ratio band G/R is the better option to replace the R band for classification.
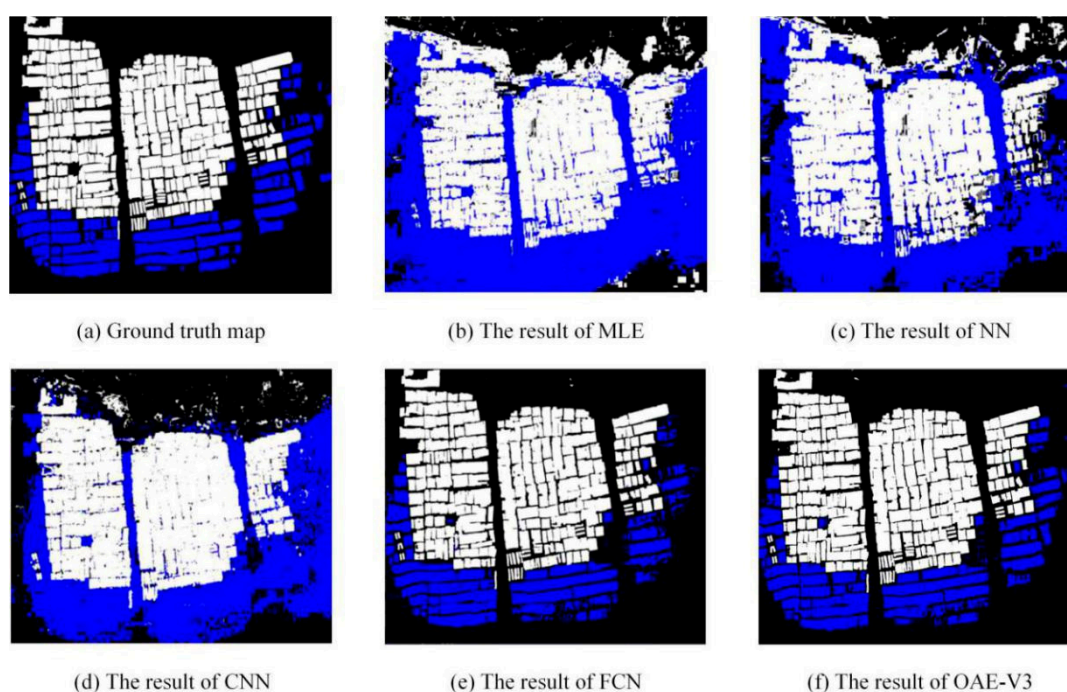

<div align="center">(a) Test images　　　　　(b) Test result of RGB bands　　　(c) Test result of G/R bands</div>
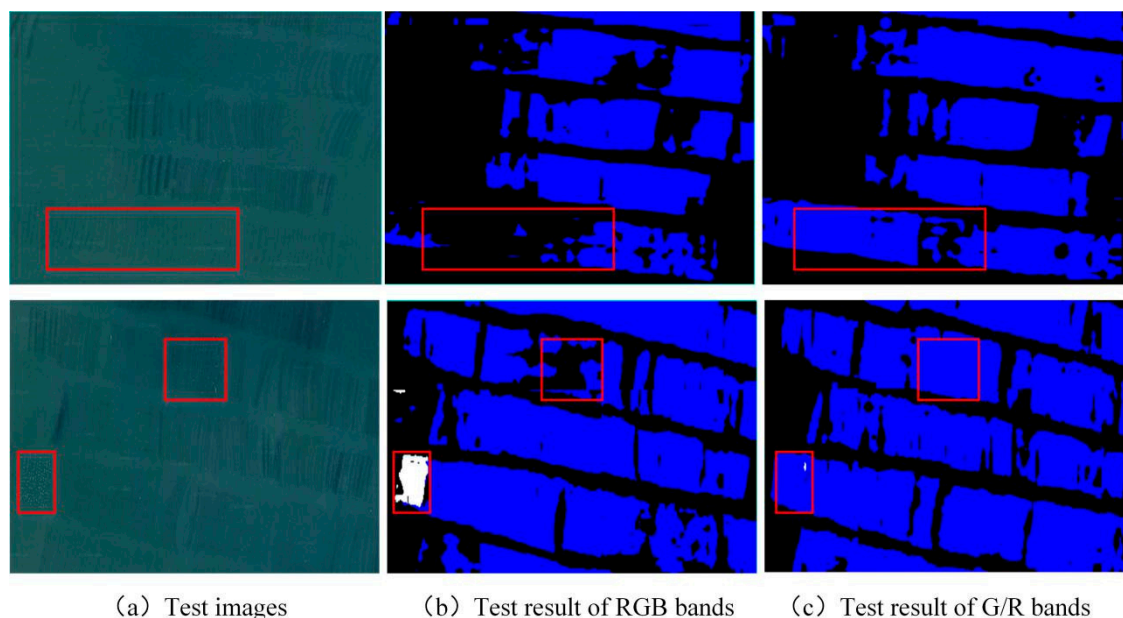
Figure 12. (**a**) The test images; (**b**) prediction result of RGB data training model; (**c**) prediction result of the ratio band(G/R) data training model (black is the background, white is fish row cage aquaculture, and blue is floating raft aquaculture).

3.3. Comparative Analysis of Multiple Supervised Classification Methods

As the classification method proposed in this paper belongs to supervised classification, other supervised methods were used, namely, traditional supervised maximum likelihood (MLE) [36], artificial neural network (NN) [37] classification, convolutional neural network (CNN) [34], and full

convolutional neural network (FCN) [22] semantic segmentation, to conduct offshore aquaculture zones classification in study area 1.

Figure 13 shows the classification diagram of offshore aquaculture extraction obtained by the OAE-V3 method and other supervised classifications. MLE, NN, and CNN cannot well distinguish the floating raft aquaculture zones and its adjacent seawater, resulting in a poor classification effect and more noise points after classification. In comparison, FCN has evident improvement in classification effect, which can distinguish the floating raft zones from sea water and reduce noise points. However, seawater in floating raft cultivation zones cannot be identified and extracted, and thus, results in incomplete identification. These zones include floating raft cultivation zones with no prominent spectral information in the research area. The classification map obtained by the method of this paper can be accurately identified in the two types of aquaculture zones. Compared with FCN, the edge is more evident, the noise point is less, and the identification of floating raft breeding area is more comprehensive. By enlarging the part of floating raft aquaculture zones with no prominent spectral information in the study area, Figure 13 indicates that the extraction effect of OAE-V3 model of floating raft aquaculture zones was significantly better than that of FCN.

(a) Ground truth map   (b) The result of MLE   (c) The result of NN

(d) The result of CNN   (e) The result of FCN   (f) The result of OAE-V3

Figure 13. (**a**) The ground truth map; (**b**) the results of MLE; (**c**) the results of NN; (**d**) the results of CNN; (**e**) the results of FCN; (**f**) the results of OAE-V3. (The black is the background, white is fish row cage aquaculture, and blue is floating raft aquaculture.) The method proposed in this paper (OAE-V3) is best at extracting sea surface aquaculture zones.

Table 1 shows the comparison of classification accuracy of each supervised classification method. The classification accuracy of extraction of aquaculture zones by MLE, NN, and CNN methods is lower, but compared with the most traditional MLE, each evaluation index of CNN considerably improved. F1 score increased from 58% to 74%, Kappa coefficient increased from 0.399 to 0.615, accuracy of fish row cage aquaculture zones exceeded 80% and reached 82.1%, but the accuracy of floating raft aquaculture zones was only 35.8%. This result shows that deep learning is very effective in the classification and extraction of remote sensing images of offshore aquaculture zones. The emergence of semantic segmentation based on the training of a large number of samples increases deep learning classification. FCN, which was first proposed, has achieved very good results in the classification of offshore aquaculture zones. Compared with CNN, FCN has made a significant breakthrough, with

the classification accuracy of fish row aquaculture zones reaching 90.5% and the extraction accuracy of floating raft aquaculture zones reaching 89.7%.

Table 1. Comparison of classification accuracy of different supervised classification methods for offshore aquaculture zones.

Method	Accuracy of Fish Steak Cage	Accuracy of Floating Raft	Pre_pixel	F1 Score	Kappa
MLE	69.8%	23.9%	57.6%	58%	0.399
NN	74.1%	33.8%	72.9%	69%	0.547
CNN	82.1%	35.8%	76.8%	74%	0.615
FCN	90.5%	89.7%	92.5%	91%	0.885
OAE-V3	94.5%	92.0%	94.8%	93%	0.925

The OAE-V3 extraction method proposed in this paper obtained the best score among all the indicators. Compared with FCN, the classification accuracy of fish cage aquaculture zones increased from 90.5% to 94.5%, the classification accuracy of floating raft aquaculture zones reached 92.0%, the overall pixel classification accuracy reached 94.8%, and the F1 score and Kappa coefficients were the highest at 93% and 0.925, respectively.

In conclusion, the OAE-V3 method proposed in this paper has the best overall extraction effect in offshore aquaculture zones. Figure 13 shows the extraction results of offshore aquaculture zones in three study zones obtained by OAE-V3.

4. Conclusions

Spectral and spatial information are the key features of offshore aquaculture zones extraction, and contextual information is a high-level summary of spectral and spatial information. Based on this idea, this paper proposes a method (OAE-V3) to identify offshore fish cage and floating raft aquaculture zones of remote sensing images.

The advantages of the OAE-V3 method proposed in this paper are: (a) Using the data expansion method on the basis of GAN effectively made up for the problem of insufficient samples. (b) Using residual network structure can solve the problem of vanishing gradients and neural network degradation, thereby extracting the complex (local) information of remote sensing image. (c) Using atrous convolution instead of partial convolution layer improved the resolution of computing feature response while maintaining the number of parameters and calculation amount to obtain more contextual information. (d) Using ASPP to cascade the void convolution of different sampling rates can extract multiscale features of images.

In this paper, the method is applied to a high-resolution multispectral remote sensing image dataset for automatic extraction of offshore aquaculture zones. The results showed that the overall identification accuracy of offshore aquaculture could reach 94.8%, for fish row cage aquaculture zones it could reach 94.5%, and for floating raft aquaculture zones it could reach 92.0%. This proves that the OAE-V3 method proposed in this paper fully accounts for the context's semantic information and improves the recognition accuracy of offshore fish cage and floating raft aquaculture zones greatly, especially when the floating raft is submerged in areas with weak signal.

In the future, we will continue to study the effectiveness of the model in identifying surface obstacles and focus on the following: 1) The extraction effect of the model on floating raft aquaculture zones on remote sensing images with extremely weak spectral information will be explored. 2) Using remote sensing datasets with different sources and different resolutions to investigate the effectiveness of the model on different datasets. 3) Investigating the effectiveness of the model in identifying other obstacles on the sea surface.

Author Contributions: Conceptualization, Baikai Sui and Tao Jiang; Data curation, Xinliang Pan; Formal analysis, Baikai Sui; Investigation, Zhen Zhang, Xinliang Pan and Chenxi Liu; Methodology, Baikai Sui and Zhen Zhang;

References

1. Lipton, D.W.; Kim, D.H. Assessing the economic viability of offshore aquaculture in korea: An evaluation based on rock bream, oplegnathus fasciatus, production. *J. World Aquacult. Soc.* **2007**, *38*, 506–515. [CrossRef]

2. Al-Nasrawi, A.K.M.; Hopley, C.A.; Hamylton, S.M.; Jones, B.G. A Spatio-Temporal Assessment of Landcover and Coastal Changes at Wandandian Delta System, Southeastern Australia. *J. Mar. Sci. Eng.* **2017**, *5*, 55. [CrossRef]

3. Ferreira, J.G.; Sequeira, A.; Hawkins, A.J.S.; Newton, A.; Nickell, T.D.; Pastres, R.; Forte, J.; Bricker, S.B. Analysis of coastal and offshore aquaculture: Application of the FARM model to multiple systems and shellfish species. *Aquaculture* **2009**, *292*, 129–138. [CrossRef]

4. Dai, H.; Bao, Y.; Bao, M. Maximum likelihood estimate for the dispersion parameter of the negative binomial distribution. *Stat. Probab. Lett.* **2013**, *83*, 21–27. [CrossRef]

5. Forney, G. Generalized minimum distance decoding. *IEEE Trans. Inf. Theory* **1966**, *12*, 125–131. [CrossRef]

6. Chen, C.W.; Luo, J.; Parker, K.J. Image segmentation via adaptive k-mean clustering and knowledge-based morphological operations with biomedical applications. *IEEE Trans. Image Process.* **1998**, *7*, 1673–1683. [CrossRef]

7. Heermann, P.D.; Khazenie, N. Classification of multispectral remote sensing data using a back-propagation neural network. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 81–88. [CrossRef]

8. Shih-Chung, B.L.; Freedman, M.T.; Lin, J.S.; Mun, S.K. Automatic lung nodule detection using profile matching and back-propagation neural network techniques. *J. Digit. Imaging* **1993**, *6*, 48–54. [CrossRef]

9. Saunders, C.; Stitson, M.O.; Weston, J.; Holloway, R.; Bottou, L.; Scholkopf, B. Support vector machine. *Comput. Intell. Neurosci.* **2002**, *1*, 1–28.

10. Zhou, X.D.; Yang, C.C.; Meng, N.N. Method of remote sensing image fine classification based on geometric features and svm. *Constr. Build. Mater.* **2012**, *500*, 562–568. [CrossRef]

11. Maulik, U.; Bandyopadhyay, S. Genetic algorithm-based clustering technique. *Pattern Recognit.* **2000**, *33*, 1455–1456. [CrossRef]

12. Guo, Y.; Wu, Y.; Ju, Z.; Wang, J.; Zhao, L. Remote sensing image classification by the chaos genetic algorithm in monitoring land use changes. *Math. Comput. Model* **2010**, *51*, 1408–1416.

13. Bianchini, M.; Scarselli, F. On the Complexity of Neural Network Classifiers: A Comparison between Shallow and Deep Architectures. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *25*, 1553–1565. [CrossRef]

14. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2012**, *60*, 84–90. [CrossRef]

15. Maggiori, E.; Tarabalka, Y.; Charpiat, G.; Alliez, P. Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Trans. Geosci. Remote Sens.* **2016**, *55*, 645–657. [CrossRef]

16. Chen, S.W.; Tao, C.S. Polsar image classification using polarimetric-feature-driven deep convolutional neural network. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 627–631. [CrossRef]

17. Ma, Y.; Zhao, D.; Wang, R.; Su, W. Offshore aquatic farming areas extraction method based on aster data. *Trends Parasitol.* **2010**, *2011*, 59–63.

18. Zhang, T.; Li, Q.; Yang, X.; Zhou, C.; Su, F. Automatic mapping aquaculture in coastal zone from TM imagery with OBIA approach. *Int. Geol. Rev.* **2010**, *2*, 1–4.

19. Lu, Y.; Li, Q.; Du, X.; Wang, H.; Liu, J. A method of coastal aquaculture area automatic extraction with high spatial resolution images. *Remote Sens. Technol. Appl.* **2015**, *30*, 486–494.

20. Fan, J.; Chu, J.; Jie, G.; Zhang, F. Floating raft aquaculture information automatic extraction based on high resolution SAR images. *IEEE Int. Geosci. Remote Sens. Symp.* **2015**. [CrossRef]

21. Geng, J.; Fan, J.; Chu, J.; Wang, H. Research on marine floating raft aquaculture sar image target recognition based on deep collaborative sparse coding network. *Acta Autom. Sin.* **2016**, *42*, 593–604.

22. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *39*, 640–651.

23. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid scene parsing network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Hawaii, HI, USA, 21–26 July 2017. [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]

25. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Semantic image segmentation with deep convolutional nets and fully connected crfs. *Comput. Intell. Neurosci.* **2014**, *4*, 357–361.

26. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *40*, 834–848. [CrossRef] [PubMed]

27. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In Proceedings of the European conference on computer vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 833–851.

28. Zhang, Z.; Huang, J.; Jiang, T.; Sui, B.K.; Pan, X.L. Semantic segmentation of very high-resolution remote sensing image based on multiple band combinations and patchwise scene analysis. *J. Appl. Remote Sens.* **2020**, *14*, 016502. [CrossRef]

29. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Bing, X.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014.

30. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Hawaii, HI, USA, 21–26 July 2017; pp. 5967–5976. [CrossRef]

31. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI, Munich, Germany, 5–9 October 2015; pp. 234–241.

32. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2015; pp. 770–778.

33. Li, H.; Lin, Z.; Shen, X.; Brandt, J.; Hua, G. A convolutional neural network cascade for face detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 5325–5334. [CrossRef]

34. Papandreou, G.; Kokkinos, I.; PierreAndré, S. Modeling Local and Global Deformations in Deep Learning: Epitomic Convolution, Multiple Instance Learning, and Sliding Window Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 390–399.

35. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on International Conference on Machine Learning, Lille, France, 6–11 July 2015.

36. Siskind, J.M.; Morris, Q. A maximum-likelihood approach to visual event classification. *IET Comput. Vis.* **1996**, *96*, 347–360.

37. Dreiseitla, S.; Ohno-Machadob, L. Logistic regression and artificial neural network classification models: A methodology review. *J. Biomed. Inform.* **2002**, *35*, 352–359. [CrossRef]