10.70711/aitr.v2i4.4865

Application of GAN-based Remote Sensing Image Segmentation Technology in Mariculture

Shuo Lian[1], Jin Zhao[1], Pu Guo[2], Taizhi Li[1]

1. School of Computer Science and Engineering, Chongqing Three Gorges University, Wanzhou District, Chongqing 404020, China

2. Department of Microelectronics and Solid-State Electronics, School of Microelectronics, Shanghai University, Jiading District, Shanghai 201800, China

Abstract: The application of remote sensing image segmentation technology in mariculture provides a new method for aquaculture monitoring and management. Generative Adversarial Network (GAN), as an emerging deep learning technology, shows great potential in remote sensing image segmentation by virtue of its unique generative adversarial mechanism. This paper summarizes the basic principle of GAN, the development of remote sensing image segmentation technology and the application of GAN in remote sensing image segmentation. Combined with the actual needs of mariculture, the application strategy of GAN-based remote sensing image segmentation technology in mariculture is discussed, including target recognition and segmentation, design and training of GAN model, and deployment of the model and real-time image processing. This paper provides detailed references and new research directions for researchers in related fields.

Keywords: GAN; Remote sensing; Image segmentation; Mariculture; Target recognition

1. Introduction

Mariculture is a crucial component of global aquaculture. In recent years, with technological advancements and increasing market demand, the scale of mariculture has been expanding. However, this growth has led to challenges such as environmental pollution, disease control, and resource utilization, creating an urgent need for efficient monitoring and management methods. Traditional monitoring techniques, which rely on manual inspections and on-site sampling, are not only time-consuming and labor-intensive but also inadequate for large-scale real-time monitoring. Remote sensing technology, with its advantages of wide coverage and real-time monitoring, has shown great potential in mariculture applications. Generative Adversarial Network (GAN)[1], as an emerging deep learning technology, possess the ability to generate high-quality images and handle complex backgrounds, offering unique advantages in remote sensing image segmentation. This paper will explore the application needs and strategies of GAN in mariculture, starting from the basic principles of GAN and considering the current development of remote sensing image segmentation technology.

2. GAN-based remote sensing image segmentation techniques

2.1 The basic principle of GAN

GAN consists of a generator and a discriminator, through the generator to generate realistic images, the discriminator to determine the authenticity of the image, the two against each other, so as to continuously improve the quality of image generation. The core idea can be expressed as a minimax game:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))]$$

where G is the generator, D is the discriminator, p_{data} is the distribution of real data and p_z is the distribution of noise. The generator deceives the discriminator by generating realistic images, and the discriminator improves its judgment by recognizing real and fake images. Through continuous adversarial training, the generator produces images of increasing quality.

The advantage of GAN is its ability to generate high quality images from noise and learn complex features of data distribution, which makes GAN perform well in tasks such as image generation, image restoration, and image segmentation. In recent years, many improved GAN models have been proposed, such as DCGAN[2], WGAN[3], CycleGAN[4], etc., which demonstrate excellent performance in different application scenarios.

2.2 Development of remote sensing image segmentation technology

Remote sensing image segmentation technology has experienced the development process from traditional methods to deep learning based methods. Traditional methods include methods based on thresholding, edge detection and region growing, which rely on expert experience and manual feature extraction and have certain limitations. With the rise of deep learning, the application of convolutional neural networks (CNN)[5] in image segmentation has made significant progress. Fully Convolutional Network (FCN)[6] is the earliest proposed end-to-end image segmentation model, which realizes segmentation of input images of arbitrary size by replacing fully connected layers with convolutional layers. FCN has achieved excellent performance on datasets such as VOC and COCO.U-Net is a classical medical image segmentation network, which achieves high-precision segmentation through a symmetric encoding-decoding structure. U-Net[7] has achieved excellent performance on many medical image segmentation tasks. SegNet[8] is a convolutional encoder-decoder architecture based on VGG16[9] with maximum pooling index for up-sampling, which improves segmentation accuracy and efficiency. In recent years, GAN-based segmentation methods have gradually become a research hotspot, and its generative adversarial mechanism gives it a unique advantage in dealing with complex backgrounds and noise interference.

2.3 Application of GAN in remote sensing image segmentation

The application of GAN in remote sensing image segmentation is mainly reflected in the segmentation effect of high precision and high efficiency. Specific applications include the recognition and segmentation of targets such as water bodies, vegetation, buildings, etc. By generating realistic segmentation results, GAN is able to deal with complex remote sensing image background and improve segmentation accuracy. In recent years, researchers have proposed a variety of GAN-based remote sensing image segmentation models. SegGAN[10] takes a pre-trained deep semantic segmentation network and installs it into a generative adversarial framework to compute better segmentation masks. CycleGAN is an unsupervised image-to-image conversion method for the case where there are no pairs of training data, and can be used for the segmentation of remote sensing images by introducing a cyclic consistency loss for the conversion between different sensor data, thus enhancing the segmentation effect. Pix2Pix[11] is a conditional GAN-based image-to-image conversion model that can realize the conversion from remote sensing images to segmentation results. PSGAN[12] realizes the segmentation of high-resolution remote sensing images by block-level processing of the images, and is suitable for large-scale remote sensing image processing. DASGAN[13] realizes the segmentation of high-resolution remote sensing images by introducing the dual-attention mechanism, utilizing the spatial attention and channel attention to weight the important regions in remote sensing images, thus realizing accurate target segmentation. CSGAN[14] realizes cross-domain segmentation by utilizing the common features of different remote sensing data sources in the training process, which is applicable to the integration and analysis of multi-source remote sensing data. In addition, GAN shows great potential in processing multispectral and hyperspectral remote sensing image segmentation.

3. Strategies for the application of GAN in mariculture

3.1 Target Recognition and Segmentation

GAN-based remote sensing image segmentation technology enhances the precise identification and segmentation of aquaculture areas and target species, enabling fine management, particularly in complex environments. Key applications include water-land segmentation and biometric identification. Accurate water-land segmentation in remote sensing images is crucial for mariculture planning and management. Traditional methods often struggle with complex backgrounds, while GAN, through high-dimensional feature learning, deliver superior segmentation. Typical GAN architectures like Pix2Pix and CycleGAN excel in image-to-image conversion, achieving high-precision segmentation by learning features across different spectral bands. Additionally, GAN require minimal labeled data, making them effective in data-scarce scenarios. For instance, CycleGAN's cyclic consistency loss ensures accurate unsupervised water-land segmentation. In mariculture, accurate identification of organisms like fish and shellfish is vital for fine management and monitoring. GAN, particularly conditional GAN[15], generate high-quality biometric data, improving recognition in challenging environments like turbid water or varying light conditions. By using environmental inputs to control the generator, these models enhance organism recognition, aiding in effective aquaculture planning.

3.2 Design and training of GAN model

The design of GAN models suitable for mariculture scenarios includes the optimization of the structure of the generator and the discriminator, as well as the selection of data and preprocessing methods during the training process. Common model structures include U-Net-based generators and PatchGAN-based discriminators. Data enhancement is required during training to improve the generalization ability of the model. The formula is as follows:

$$\mathcal{L}_{\mathrm{GAN}}(G,D)=\mathbb{E}_{x\sim p_{\mathrm{data}}(x)}[\log D(x)]+\mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))]$$

In addition, content loss and style loss can be introduced to further enhance the quality of the generated images:

$$\mathcal{L}_{\text{content}}(G) = \mathbb{E}_{x \sim p_{\text{data}}(x), z \sim p_z(z)}[\| x - G(z) \|_1]$$

$$\mathcal{L}_{\text{style}}(G) = \mathbb{E}_{x \sim p_{\text{data}}(x), z \sim p_z(z)}[\| \phi(x) - \phi(G(z)) \|_2]$$

where ϕ denotes the feature extraction layer of the pre-trained VGG network.

In the process of model training, the diversity and representativeness of the dataset need to be fully considered to ensure the generalization ability of the model. Meanwhile, suitable optimization algorithms and training strategies, such as Adam optimizer and learning rate decay strategy, are used to accelerate model convergence and improve segmentation accuracy.

3.3 Model deployment and real-time image processing

To ensure the efficient operation of GAN models in real-world applications, selecting the appropriate hardware platform is essential. Common platforms include graphics processing units (GPUs), tensor processing units (TPUs), and field-programmable gate arrays (FPGAs). GPUs, widely used in deep learning, offer powerful parallel computing capabilities, making them suitable for computationally intensive model inference tasks. NVIDIA's CUDA platform and TensorRT optimization tool can significantly boost GAN model inference performance. TPUs, developed by Google, accelerate matrix operations and are particularly effective for training and inferring large neural networks, providing efficient computing resources and reducing latency. FPGAs feature a highly flexible architecture that can be customized for specific needs, making them ideal for deploying lightweight GAN models while reducing power consumption and ensuring performance. In deployment, strategies such as model compression, acceleration, and distributed computing enhance real-time processing capabilities. Techniques like model pruning, quantization, and distillation reduce the number of parameters and computational complexity, speeding up inference. Distributed computing frameworks, such as TensorFlow Serving or ONNX Runtime, distribute tasks across multiple devices to improve processing efficiency.

4. Conclusion

This paper summarizes the current status and development trend of the application of GAN-based remote sensing image segmentation technology in mariculture, and puts forward specific application strategies. The GAN-based segmentation method has the advantages of high precision and high efficiency, which provides a new technical means for the modern management of mariculture. Future research can further optimize the model structure and improve its adaptability and real-time processing ability in complex environments.

5. Acknowledgement

The work described in this paper was supported in part by the Chongqing Graduate Student Research Innovation Program (CYS240759).

References

[1] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.

[2] Wu Q, Chen Y, Meng J. DCGAN-based data augmentation for tomato leaf disease identification[J]. IEEE access, 2020, 8: 98716-98728.

[3] Wang Q, Zhou X, Wang C, et al. WGAN-based synthetic minority over-sampling technique: Improving semantic fine-grained classification for lung nodules in CT images[J]. IEEE Access, 2019, 7: 18450-18463.

[4] He D, Xia Y, Qin T, et al. Dual learning for machine translation[J]. Advances in neural information processing systems, 2016, 29.

[5] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25.

[6] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.

[7] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer International Publishing, 2015: 234-241.

[8] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(12): 2481-2495.

[9] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.

[10] Zhang X, Zhu X, Zhang N, et al. Seggan: Semantic segmentation with generative adversarial network[C]//2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM). IEEE, 2018: 1-5.

[11] Isola P, Zhu J Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1125-1134.

[12] Liu Q, Zhou H, Xu Q, et al. PSGAN: A generative adversarial network for remote sensing image pan-sharpening[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(12): 10227-10242.

[13] Kapil A, Wiestler T, Lanzmich S, et al. DASGAN--Joint Domain Adaptation and Segmentation for the Analysis of Epithelial Regions in Histopathology PD-L1 Images[J]. arXiv preprint arXiv:1906.11118, 2019.

[14] Babu K K, Dubey S R. CSGAN: Cyclic-synthesized generative adversarial networks for image-to-image transformation[J]. Expert Systems with Applications, 2021, 169: 114431.

[15] Mirza M, Osindero S. Conditional generative adversarial nets[J]. arXiv preprint arXiv:1411.1784, 2014.