U_EFF_NET: High-Precision Segmentation of Offshore Farms From High-Resolution SAR Remote Sensing Images

Gang Qin, Shixin Wang, Futao Wang, Yi Zhou, Zhenqing Wang, and Weijie Zou

Abstract—Offshore aquaculture promotes the development of aquaculture industry and brings huge economic benefits to fishermen, while seriously affecting the near-coast environment. Accurate access to the range of offshore farms at home and abroad is of great significance for marine disaster warning and coastal management. Remote sensing is a very effective means of observing offshore farms. Offshore farms segmentation technology is more mature in high-resolution optical images. SAR images have the advantage of being available all day and all night. Using SAR images to extract offshore farms has become a recent research hotspot. Gaofen-3 and HISEA-1 satellites have high-resolution marine observation capabilities. Combined with their super observation capabilities, this article proposes a technical solution to extract offshore farms using Gaofen-3 and HISEA-1 images. First, the combine-crop component module is proposed to create a new SAR images dataset based on the original dataset. Next, common data augmentation methods, such as flip, rotate, and copy–paste, are used, and the online multiscale resolution image expansion is proposed to enrich the dataset and enlarge the data volume. Finally, the U_EFF_NET model is proposed. The model uses encoder–decoder structure and the lightweight feature extraction network EfficientNet-b0 as backbone. The decoder embeds the attention mechanism. The optimization strategy uses joint loss function to achieve multitask learning, and the OHEM strategy is added to the joint loss function. The method proposed in this article has high accuracy and high practicality as the frequency-weight intersection ratio of the extracted results on the test set is 98.12%. And the inference time is at the forefront.

Index Terms—2021 GaoFen challenge, high-resolution SAR remote sensing images, offshore farms segmentation, U_EFF_NET.

I. INTRODUCTION

With the over-exploitation and use of surface natural resources and the environment by human activities, global marine fishery resources are depleted under extensive and high-frequency fishing. As a result, in the face of the slowdown in the growth of the total amount of pelagic fishing, offshore aquaculture is gradually developing in order to meet the human demand for marine products, which is the most important form of aquaculture in many coastal zones. While offshore aquaculture brings a large amount of marine food [1], it also brings a series of negative impacts. Its over-intensive purse-seine farming is seriously damaging the offshore ecological environment [2], [3]. The deterioration of the near-coast environment has led to the rising labor cost of fisheries. It is very important to control the number and scale of near-coast farms. In order to reasonably plan the layout of near-coast farms, the spatial distribution of near-coast farms needs to be obtained timely and accurately, which has positive significance for resource environment, natural disaster warning, and urban planning.

At this stage, domestic and foreign offshore farm extraction methods are generally visual interpretation methods and object-oriented methods. The visual interpretation method is a way for Remote Sensing Professional Researcher to mark offshore farms based on the characteristics of offshore farming areas, so as to highlighting the farms. The method has high extraction accuracy and does not require writing algorithms to construct offshore farm segmentation models. However, the whole operation process is tedious and task-intensive, and requires the remote sensing image interpretation ability of the interpreters. In the object-oriented extraction method, the remote sensing extraction object is no longer a single image element, but the offshore farm as the extraction object. It is no longer limited only to the band spectral characteristics of individual image elements, but completely considers the shape, texture, and spectral information of offshore farms in remote sensing images for classification and extraction. With wide coverage, low imaging cost, and the advantage of all-day, all-weather image acquisition, high-resolution SAR images, as an important means of surface dynamic monitoring, can quickly and accurately extract offshore farms and obtain their spatial distribution. Common machine learning methods, such as clustering algorithms, logistic regression algorithms, convolutional neural networks, realize SAR images oil segmentation [4]. Unsupervised SAR images segmentation is based on Markov fields in higher order neighborhoods [5]. Based on feature extraction of wavelet decomposition, watershed segmentation algorithm, and

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
grayscale co-occurrence matrix preprocessing, using memetic algorithms achieve natural image and remote sensing image segmentation [6]. Using multiple grids and sparse field methods, MRF embedding of level sets implements SAR images, medical image, and natural image segmentation [7].

High-resolution SAR images are important for offshore farm segmentation tasks in marine disaster warning and coastal zone management. Different ground categories have their corresponding specific values for amplitude, phase, and frequency in SAR images. According to the different information provided by amplitude and phase, a specific category can be targeted and effectively segmented, which is the traditional basic step of SAR images analysis and processing. Combining tandem-x data and amplitude and phase information from SAR images to monitor rice characteristics in different growth states [8]. Multiple features are generated based on different target decomposition theorems and different polarization methods. Then, to achieve multispecies segmentation of SAR images pixel points with the help of nearest regularized space and Markov random fields [9]. Combining texture images derived from ALOS PALSAR-L band and RADARSAT-2C band, and their different polarization methods of HH and HV, multiple machine learning methods are used to achieve land cover classification in humid tropical regions [10]. A polygonal hidden model with G0 distribution based on likelihood model and spatial texture model is used to implement the semantic segmentation of SAR images. The above methods are deficient in that they manually process, extract, and apply features based on the amplitude and phase characteristics of SAR image imaging. In the segmentation process, because of the very limited adaptation of parameters, the segmentation parameters need to be manually adjusted in different cases, which is tedious to operate during the whole extraction process. In practical applications, the visual effect of RGB band of optical images is clearer. The SAR images are combined with optical images, using their auxiliary information to decode the ground cover type. It makes up for the poor visual effect of SAR images. Researchers use both optical images and SAR images to extract the ground cover types and it is a popular research direction [11], [12], [13]. With the hardware development and deep learning technology advancement, deep convolutional neural networks have shown excellent performance in extracting image features. A series of network structures for extracting natural image feature maps have been proposed, and some researchers have successfully applied convolutional neural networks to the field of semantic segmentation. They proposed U-Net [14], SegNet [15], PSPNet [16], DeepLab [17], [18], [19], [20] series, HR-Net [21] and VGGNet [22], ResNet [23], EfficientNet [24], and other classical convolutional neural networks. With the help of these networks, the accuracy and efficiency of semantic segmentation of natural images are dramatically improved. In recent years, these methods have been applied to extract feature cover types from high-resolution optical remote sensing images, multipath residual networks for semantic segmentation of POLSAR images [25], ResNet deep learning model pretrained from optical remote sensing images enables multiclass semantic segmentation of SAR images [26]. TomoSAR point clouds and auxiliary information are used in automatically labeling buildings, and deep learning was utilized to achieve SAR images building segmentation [27]. fcn-8s, deep residual U-Net, and DeepLab v3+ multiple deep learning methods realize SAR images road segmentation [28]. Combining existing labeled data to form new labels, deep neural network implements building and road segmentation of SAR images [29]. However, the studies in SAR images are mostly limited to buildings, roads, and sea ice [30]. The extraction of offshore farms from high-resolution optical remote sensing images is technically mature, and in the traditional processing method, offshore farms are automatically identified in the optical images based on their spectral and textural features [31]. Extraction of offshore farms in optical remote sensing images based on deep convolutional neural networks methods are gradually proposed. After the expansion of floating raft farm area samples by generative adversarial networks, the deeplabv3 method was used to achieve the extraction of offshore farm area from high-resolution optical images [32]. Combining target detection with image segmentation techniques extracts coastal nets from high-resolution optical images, and then the net area is evaluated [33]. Weakly supervised semantic segmentation network extracts offshore farms from multimodal high-resolution optical remote sensing images [34]. In the last 2 years, some researchers have proposed to combine SAR images imaging mechanism and deep learning methods to extract offshore farms from high-resolution SAR images. According to offshore farm shape constraints, new components are added to the semantic segmentation model of offshore farms [35]. Based on the analysis of the polarization scattering mechanism of offshore farms from Gaofen-3 images, researchers used the c-means method to extract the offshore farm [36]. A novel semantic segmentation network is designed to realize offshore farm extraction [37]. These provide much reference value for subsequent studies.

Deep learning techniques, which have shown excellent performance in the field of natural images, have developed rapidly and started to be widely used in the field of land cover extraction from SAR images [5], [6], [7], [24], [25], [26], [27], [28]. Inspired by this, we used a fully convolutional neural network U_EFF_NET to extract offshore farms from high-resolution SAR images. A suitable network structure is essential to achieve the desired offshore farm segmentation results. The U-Net [14] structure can fuse the high-level and low-level features very well, and the feature map contains rich spatial and semantic features, so U-Net is used as the overall framework in this article. Considering the accuracy and efficiency of the model in this article, we use EfficientNet-b0 [24] as the encoder. Embedded in the decoder, the spatial channel joint attention mechanism CBAM [38] assigns differentiated weights according to the different importance of features for the purpose of intelligently correcting features. The accuracy of positive sample boundary segmentation in the segmentation results has been a challenge, and this article uses boundary-weighted cross-entropy loss, and dice loss function as joint loss function. The online hard sample mining is fused. They alleviate that problem from the aspect of model optimization.

In this article, the extraction method is proposed for the offshore farm segmentation task. EfficientNet-b0 is selected as
the feature extraction network for the U-Net model, and the main contributions of this article are as follows:

1) U_EFF_NET network structure. Using U-Net as the network skeleton and EfficientNet-b0 as the encoder. The joint spatial channel attention mechanism CBAM is embedded into the decoder to accurately correct the feature map, which makes the model more generalized.

2) In this article, a multitask learning model for offshore farms and their boundaries is proposed. The joint loss function of boundary-weighted cross-entropy loss and dice loss [39] incorporates online hard example mining [40]. They are used to enhance the segmentation ability of the farms and their borders. Label smoothing strategy is applied. That makes the segmentation results closer to the ground truth values.

3) This article uses a variety of data enhancement methods: high-resolution SAR images horizontal flip, vertical flip, 90°–270° rotation, combine-crop module, multiscale resolution (1–3 m resolution images) image expansion, and copy–paste [41].

4) The method proposed in this article achieved the first place in the 2021 Gaofen Challenge on Automated High Resolution Earth Observation Image Interpretation—High Resolution SAR Images Offshore Farm Segmentation Track. It was the only team that evaluated the accuracy of the results over 98%. Ablation experiments and comparison experiments were conducted to evaluate the performance of the strategy proposed in this article.

II. METHODOLOGY

The 2021 Gaofen Challenge provides preprocessed single-channel high-resolution SAR images, and manually marks the offshore farms in the SAR images. For the task of segmenting offshore farms in the remote sensing images of Gaofen-3 and HISEA-1 satellites, this article proposes a semantic segmentation model U_EFF_NET, achieving high precision and high efficiency extraction of offshore farms. Meanwhile, a series of data enhancement methods are applied to expand the SAR images dataset. Then a series of model improvement methods, such as joint loss function and online hard example mining, are applied. The construction details of the model are shown in Fig. 1.

A. Data Processing Strategies

The performance of a deep learning model relies not only on the structure of the model itself, but also on the quantity and quality of the dataset. High-resolution SAR images of offshore farms are used as the training dataset.
In this article, we use common image enhancement methods to expand the high-resolution SAR images dataset by referring to and using the original SAR images. Common data enhancement methods, such as vertical flip, horizontal flip, rotation 90°, 180°, and 270°, translation and rotation transformation, copy–paste, etc., expand the training set and enhance the adaptability of the segmentation network to different angles and different spaces of offshore farms. The combine-crop module and multiscale resolution image expansion are proposed in a targeted manner.

The competition provides a high-resolution SAR images dataset with an image size of 512×512. The images partially overlap each other, resulting in some images being present in both the training and validation sets. The existence of the same images in the training and validation sets leads to the training set used by the U_EFF_NET network to contain part of the validation set. The evaluation results are inaccurate when the accuracy is calculated on the validation set extraction results of the offshore farm. The offshore farm segmentation model has a relatively high accuracy of segmentation results on the validation set, but the accuracy of offshore farm segmentation on the test set is not high. The statistical results of the validation set are confusing. To address this one problem, this article proposes the combine-crop module to stitch the high-resolution SAR images slices into whole high-resolution SAR images, and then crop the whole high-resolution SAR images into remote sensing image slices without overlap, as shown in Fig. 2. The validation set and the training set were divided in a ratio of 1:4. Using this training and validation set partitioning method, the U_EFF_NET model shows a positive correlation between the accuracy of the extraction results on the validation and test sets. In this article, the new dataset constructed on the original dataset successfully solves the problem of inaccurate accuracy evaluation during model training, obtaining and saving the model parameters with the highest accuracy on the validation set during training. The model is most likely to have the highest accuracy on the test set.

In order to increase the number of datasets, this article proposes a multiscale resolution image expansion method. The high-resolution SAR images in the training set have different resolutions of 1 and 3 m. The specific expansion process is as follows: four SAR images slices are randomly selected to form a slice of 1024×1024 size, and then it is reduced to 512×512 size; meanwhile, eight other SAR images slices are randomly selected to form a slice of 1536×1536 size, and then it is reduced to 512×512 size. The result of data expansion is shown in Fig. 3. The richness of the dataset resolution is increased.

**B. Build Model Strategy**

In this article, a new neural network structure U_EFF_NET is used, and the overall network structure is shown in Fig. 1. U-Net cleverly combines the feature extraction network EfficientNet-b0, which acts as an encoder to extract high-dimensional feature maps containing hidden complex feature information. With encoder–decoder structure to achieve segmentation effect, U-Net is an innovative network structure designed for image segmentation tasks. U_EFF_NET uses imagenet pretrained weights to achieve migration learning from natural images to domains of remote sensing images. The decoder was added with CBAM attention mechanism to highlight spatial location and morphological features of offshore farms. Its adaptive recalibration of feature maps to enhance meaningful features while suppressing weaker ones. The above strategy is beneficial to the offshore farm segmentation task and improves the segmentation effect of U_EFF_NET on offshore farms from high-resolution SAR images.

1) EfficientNet-b0: The series of EfficientNet b0-b7 deep convolutional neural networks are based on depth, width, and resolution of EfficientNet-b0. Other networks in the series search for the most suitable model building parameters. The fundamental principle is that the depth of the network depth increases to get richer and more complex features. Width increases to get more offshore farms’ feature patterns. A larger resolution captures the detailed features of offshore farms. Considering the time consumption and the segmentation efficiency of the model, this article chooses to use effintnet-b0 as the encoder of the segmentation model.

The encoder EfficientNet-b0 is shown in Fig. 4, which consists of one convolutional layer with seven blocks. Each block consists of a different number of MBConv structures stacked on
top of each other. The multiplicity factor of each MBConv and
the size of the convolution kernel are determined by different
specific parameters.

2) Convolutional Block Attention Module (CBAM): CBAM, as
shown in Fig. 5, is a joint attention mechanism that combines
the channel attention mechanism and the spatial attention mecha-
nism of the convolution module. The feature map is input to the
channel attention module, and the channel attention feature map
is output, which is wisely multiplied with the input feature map
to obtain the intelligently corrected feature map on the channel
side. The channel-corrected feature map is input to the spatial
attention module, and the spatial attention feature map is output,
which is wisely multiplied with the input feature map to obtain
the channel and spatially intelligently corrected feature maps.

The channel attention module contains two channel attention
methods: global maximum pooling and global average pooling.
They go through each of the two fully connected layers with
shared weights to get the adjusted two vectors, i.e., channel
attention. First, they are elementwise added operation. Second,
they are sigmoid activated to get the channel attention feature
map.

The spatial attention module, which is spatially activated
channel compression, is max-pooled and average-pooled in the
channel dimension to obtain two spatial attentions. The two
spatial attention are concatenated to a piece to form a feature
map with channel dimension of two. The convolution module
compresses it to one channel. Finally, after sigmoid activate, the
spatial attention feature map is obtained.

C. Loss Function

1) Boundary-Weighted Cross-Entropy Loss Function: In the
high-resolution SAR remote sensing images, compared with
the samples of the internal area of offshore farms, the sam-
ples of the boundary of offshore farms accounted for a small
percentage. The statistical features learned by the U_EFF_NET
model are more beneficial for extracting the internal regions of
offshore farms, and the boundary sample features are complex
and difficult to extract. The cross-entropy loss function com-
putes the similarity between the model true value and the infer-
ence result for a single pixel location. The boundary-weighted
cross-entropy loss function is designed and implemented for the
problem that the segmentation boundary is difficult to extract.
Based on its advantage of acting on individual pixel positions, the
boundary-weighted cross-entropy loss function is introduced to
offshore farm extraction to increase the offshore farm boundary
weights and thus improve the boundary extraction accuracy.
This loss function is applied to offshore farm extraction, and
improves the accuracy compared with the original cross-entropy
loss function.

The cross-entropy loss function is the most commonly used
loss function in deep learning semantic segmentation, and the
binary classification cross-entropy function is formulated as

$$\text{Loss}_{CE} = \frac{1}{N} \sum_i - [y_i \cdot \log (p_i) + (1 - y_i) \cdot \log (1 - p_i)]$$ (1)
where $y$ is the prediction result and $p$ is the ground truth. In the cross-entropy function, the weight of each pixel is equal.

The boundary area of the farm is difficult to partition. It is weighted to give a higher weigh. In this article, $w$ is taken as 4. The boundary-weighted cross-entropy function is formulated as

$$\text{Loss}_{\text{BW,CE}} = \text{Weight} \cdot \text{Loss}_{\text{CE}}$$

(2)

$$\text{Weight} = \begin{cases} 4, & \text{boundary} \\ 1, & \text{not boundary} \end{cases}$$

(3)

2) Dice Loss Function: In the task of segmenting offshore farms from high-resolution SAR images, the offshore farm area extracted by the U_EFF_NET model is considered as one whole region. The accuracy assessment metric of the extraction results is Frequency Weighted Intersection Over Union (FWIoU), which calculates the difference between the ground truth and the inference result at the region level. It both enables the model to learn the overall features of the offshore farm area, including shape and texture, and allows FWIoU to be used as a parameter for model optimization learning. Given the optimization index itself, when choosing the loss function, it is better to choose the dice loss function whose optimization index is FWIoU. The dice loss function [39] has a significant role in this offshore farm segmentation task. Extraction accuracy increases in offshore farm areas. The boundary-weighted cross-entropy loss function and the dice loss function form a joint loss function, realizing multitask learning. The joint loss function of the offshore farm segmentation model for high-resolution SAR images is as follows

$$\text{Loss}_{\text{Dice}} = 1 - \frac{\sum_{i,j} |p_{ij} \cap y_{ij}|}{\sum_{i,j} (|p_{ij}| + |y_{ij}|)}$$

(4)

$$\text{Loss} = \text{Loss}_{\text{BW,CE}} + \text{Loss}_{\text{Dice}}.$$

(5)

3) Online Hard Sample Mining: Deep learning is a black box model and its training process is agnostic. Online hard sample mining [40] can mine complex samples during black box training. And then it focuses on learning complex samples and reduce learning simple samples. The problem of category imbalance is caused by the disparity in the number of offshore farms and background categories in the sample of SAR images dataset. The background categories with high numbers are more easily recognized by the model, but the training target of the model, offshore farms, is not extracted well. The overall recognition is better and it is easy to make the model incorrectly assume that the model fit is good. As a result, deep convolutional neural networks have difficulty learning diversity and high loss samples. The core of online hard sample mining algorithm is to select some complex samples as the key optimization targets to improve the parameter optimization of the network. In the offshore farm segmentation scenario, most of the complex samples are located in the boundary region of the offshore farm, and they are selected to target mining the features of this region. In this article, we use [40] online hard sample mining, which is a more targeted online selection method to optimize the training process of black box invisibility to some extent.

III. Experiments and Results

In this section, in order to demonstrate the effectiveness of the algorithm for the task of segmenting offshore farms in high-resolution SAR images, this article will show the experimental process and experimental results from four aspects. And a series of comparison experiments are designed.

A. Experimental Data

The 2021 Gaofen Challenge provides 6000 high-resolution SAR images, of which 4000 images are visible as the training set and 2000 images are not visible as the test set, stored online. Serval example of training set are shown in Fig. 6. The resolution is 1–3 m. The images come from the Gaofen-3 and HISEA-1 satellites. Each SAR images was manually labeled with pixel-level offshore farms. The 4000 slices of images where overlap exists were composed into 45 whole remote sensing images. Then 45 whole remote sensing images are cropped into remote sensing image slices with no overlap rate, and there are 1369 image slices in total. A total of 1096 slices were divided into the training set and the remaining 273 slices were divided into the validation set.

The accuracy evaluation index of the results brought up by offshore farms, FWIoU, is formulated as

$$\text{FWIoU} = \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{j=1}^{N} S_{ij}S_{ii} - \sum_{j=1}^{N} S_{ij}(S_{ij} + S_{ji})}{\sum_{j=1}^{N} S_{ij}S_{ii}}$$

(6)

where $S_{ij}$ is the number of samples with the actual category $i$ and the predicted category $j$.

Of all the metrics, FWIoU is the most commonly used metric due to its simplicity and representativeness. Most researchers use this metric to report the accuracy of their assessment results, and FWIoU is the only metric specified by the competition.

B. Experimental Platform

The experimental platform uses Intel Core i7-8700@3.20GHz 6-core processor with 32.0 G RAM and Nvidia GeForce RTX 3090. For the software environment, this article uses Windows 10 Professional 64-bit operating system, the programming language is Python 3.7. Pytorch 1.7 was used as the model building tool and CUDA 11.0 was used as the GPU computing platform.

C. Experimental Results

To demonstrate the high accuracy of the offshore farm segmentation scheme proposed in this article, comparative experiments are designed in this article. The test set is stored in the cloud platform and the test set is not visible locally, so the validation set is used as the evaluation dataset during the accuracy evaluation. The extraction accuracy of offshore farms for different models of PSPNet (resnet34), Deeplabv3+ (resnet34), U_net (resnet34), and U_EFF_NET (EfficientNet-b0) are shown in Table I. To eliminate the random factor in the training process of the deep neural network, each solution was repeated
TABLE I
EXPERIMENTAL RESULTS OF DIFFERENT MODELS COMPARISON

| Method     | Feature extraction network | FWIoU(%) |
|------------|----------------------------|----------|
| PSPNet     | resnet34                   | 0.9478   |
| Deeplabv3+ | resnet34                   | 0.9527   |
| Unet       | resnet34                   | 0.9528   |
| U_EFF_NET  | EfficientNet-b0            | **0.9576** |

The bold entities mean the optimal results (on validation sets).

In the U_EFF_NET model segmentation of offshore farms for high-resolution SAR images, the voting mechanism is introduced on the test set to fuse five models with five-fold cross-validation. Meanwhile, the test-time enhanced test-time enhancement (TTA) [42] method (including horizontal flip, vertical flip, and diagonal flip) is used to obtain the final inference results of offshore farms. The image size in the test set varies from $512 \times 512$, $1024 \times 1024$, and $2048 \times 2048$ size. Considering the efficiency of farm extraction, the SAR images size in the test set is modified to $512 \times 512$ to facilitate model training because of meeting the requirement that one batch can only accept images of the same size. After U_EFF_NET extracts the offshore farms in the image, the SAR images are expanded to the original size. The U_EFF_NET inference time and inference accuracy are shown in Table I. The inference time is 1866 s and the FWIoU value is 98.12%. Since the test images are not visible, the validation set is used as an example to show the results. Fig. 7 shows some of the validation set images, their corresponding label images, and the model inference results. From the experimental results, it can be seen that the offshore farm scheme in this article gets very good results from both quantitative and qualitative perspective.
TABLE II
RESULTS OF ABLATION EXPERIMENTS

| Model      | Backbone       | Loss function | Attention mechanism | Voting mechanism | Other strategies          | FWIoU(%) |
|------------|----------------|---------------|---------------------|------------------|--------------------------|----------|
| Unet       | VGGNet         | ce            | -                   | -                | -                        | 95.87    |
| Unet       | EfficientNet-b0| ce            | -                   | -                | -                        | 96.21    |
| Unet       | EfficientNet-b0| ce            | CBAM                | -                | -                        | 96.52    |
| Unet       | EfficientNet-b0| bw_ce         | CBAM                | -                | -                        | 96.86    |
| Unet       | EfficientNet-b0| bw_ce +dice+O  | CBAM                | -                | -                        | 97.05    |
| Unet       | EfficientNet-b0| bw_ce +dice+O  | CBAM                | Model fusion     | -                        | 97.67    |
| Unet       | EfficientNet-b0| bw_ce +dice+O  | CBAM                | Model fusion     | Multi-scale resolution    | 98.12    |
|            |                |               |                     |                  | remote sensing images + tta|         |

and takes a long computational time, resulting in a long extraction time for offshore farms, which is a hindrance to practical applications. Therefore, in addition to accuracy, the inference time of the model should also be considered. In this article, we design the data preprocessing strategy and build the model with comprehensive consideration of inference time and model accuracy. Setting the batchsize achieves the highest memory utilization under the 24 GB limit of rtx titan video memory. The model inferred the offshore farm results with a uniform size of 512*512, and then reverted to the original size after the offshore farm extraction model extraction was completed. During the experiment, four models are fused to achieve the highest extraction accuracy. It saves inference time.

In order to investigate the segmentation effect of offshore farms played by the method or module proposed in the solution of this article, ablation experiments are conducted in this article. Ablation experimental results show the accuracy enhancing effect of each module, as shown in Table II. Compared with VGGNet, EfficientNet-b0 as backbone, and the FWIoU accuracy metric improves from 95.87% to 96.21%. Adding CBAM attention mechanism to the decoder of U_EFF_NET model, the extraction accuracy of offshore farms is improved from 96.21% to 96.52%. Compared with the traditional cross-entropy loss function, the boundary-weighted cross-entropy loss function improves the extraction accuracy of offshore farms from 96.52% to 96.86%. A joint loss function consisting of boundary-weighted cross-entropy loss function and dice loss function replaces the single loss function. And adding online hard sample mining strategy can improve the extraction accuracy of offshore farms from 96.86% to 97.05%. The five sets of training sets obtained by five-fold cross-validation, and then this article fuses their trained five models to soft-vote the extraction results of offshore farms implementing the model voting mechanism.

The extraction accuracy of offshore farms is improved more significantly from 97.05% to 97.67%. It has been experimentally verified that a good extraction effect can be achieved when fusing four models, and the inference time of the models can be significantly reduced. The addition of the online multiscale resolution remote sensing images expansion module and the addition of TTA improve the model accuracy from 97.67% to 98.12%.

D. Comparison of Segmentation Models in the Final Contest

Comparing with other schemes of offshore farm segmentation, as shown in Table III, it more fully illustrates the leading degree of accuracy and efficiency of the proposed scheme in this article. Since the other solutions are not visible, the difference in inference time and accuracy between the solution proposed in this article and the other solutions is analyzed at a shallow level. The accuracy of the scheme proposed in this article only breaks 98%, which is the highest accuracy scheme and greatly meets the accuracy requirements of practical applications. The inference time is 1866 s, which is faster in all schemes. It can extract offshore farms with both high accuracy and high efficiency.

IV. CONCLUSION

Facing the task of constructing offshore farm segmentation models for the high-resolution SAR images provided by Gaofen-3 and HISEA-1, this article proposes the U_EFF_NET deep convolutional neural network model. It is capable of extracting
offshore farms with high accuracy, generalizability, and efficiency. It achieved the best extraction accuracy of 98.12%. First, the combine-crop component module is proposed. Based on the original data, a new SAR images dataset without overlap rate is created so that U_EFF_NET extraction results on the validation and test sets have proportional accuracy. Second, the number of manually annotated high-resolution SAR images datasets is limited in remote sensing application studies, utilizing common data enhancement methods, such as flip, rotate, copy–paste, etc. Then online multiscale resolution image expansion is proposed, which is a data enhancement strategy proposed for different resolutions of remote sensing images. These strategies increase the number of datasets, enhance the background diversity of positive samples, and attenuate the overfitting of the model. Again, the U_EFF_NET model is proposed. The model adopts encoder–decoder structure and uses the lightweight feature extraction network EfficientNet-b0 as backbone, which is superior to other feature extraction networks such as resnet. The inference time is reduced while ensuring the extraction of useful features. The CBAM module is embedded in the decoder. Allowing the model to focus more on channels with distinct features as well as spatial locations, enabling the features to be effectively corrected. Finally, this article uses the joint loss function of the boundary-weighted cross-entropy loss function and the dice loss function. The joint loss function is better than any single loss function. And adding online hard sample mining module be able to optimize the training process without knowing the black box content.

The scheme proposed in this article has strong practicality in terms of extraction accuracy and time efficiency. It can be used as a benchmark experiment and can meet the needs of application scenarios of near-coastal farm extraction. In the future, the number of SAR images increases, and a larger range of high-resolution SAR images are provided. In order to solve the problem of large-scale offshore farm extraction, we are prepared to use a priori knowledge in the field of remote sensing and summarize a priori knowledge in the form of a knowledge graph, using knowledge graphs to construct an offshore farm segmentation model with prior knowledge.

REFERENCES

[1] Y. Wu, “Research on coastal aquaculture detection using remote sensing images,” Chin. Acad. Sci., Beijing: Institute of remote sensing and digital earth, 2017, pp. 1–3.
[2] M. Gai and Q. Che, “Analysis on the influence of Liaoning’s economic development on the coastal sea environment [J],” Ocean Develop. Manage., vol. 24, no. 5, pp. 49–52, 2007.
[3] Y. Shu, L. Luo, and Y. Wen, “Influence of marine aquaculture on coastal ecological environment [J],” Mar. Environ. Sci., vol. 21, no. 3, 2002, Art. no. 75.
[4] D. Cantorna et al., “Oil spill detection in SAR images using convolutional neural networks: A comparative analysis with clustering and logistic regression algorithms,” Appl. Soft Comput., vol. 54, 2017, Art. no. 105716.
[5] F. Wang, Y. Wu, Q. Zhang, W. Hao, M. Li, and G. Liao, “Unsupervised SAR scene image segmentation using higher order neighborhood-based triplet Markov fields model,” IEEE Trans. Geosci. Remote Sens., vol. 52, no. 8, pp. 5193–5205, Aug. 2014.
[6] L. Liao, M. Gong, S. Wang, B. Hou, Z. Zheng, and Q. Wu, “Natural and remote sensing image segmentation using memetic computing,” IEEE Comput. Intell. Mag., vol. 5, no. 2, pp. 78–91, May 2010.
[7] X. Yang, X. Gao, D. Yao, X. Li, and J. Li, “An efficient MRF embedded level set method for image segmentation,” IEEE Trans. Image Process., vol. 24, no. 1, pp. 9–21, Jan. 2015.
[8] C. Gao and E. Ertin, “Paddy-rice monitoring using Tandem-X,” IEEE Proc. Geosci. Remote Sens., vol. 53, no. 2, pp. 900–910, Feb. 2015.
[9] F. Zhang et al., “Nearest-regularized subspace classification for PolSAR imagery using polarimetric feature vector and spatial information,” Remote Sens., vol. 9, 2017, Art. no. 1114.
[10] G. Li, D. Lu, and E. Moran, “A comparative analysis of ALOS PALSAR L-band and RADARSAT-2 C-band data for land-cover classification in a tropical moist region,” ISPRS J. Photogramm. Remote Sens., no. 70, pp. 26–38, 2012.
[11] M. A. Belenguer-Plomer et al., “CNN-based burned area mapping using radiance and optical data,” Remote Sens. Environ., vol. 260, 2021, Art. no. 112468.
[12] W. Zhao et al., “Deeply synergistic optical and SAR time series for crop dynamic monitoring,” Remote Sens. Environ., vol. 247, 2020, Art. no. 11952.
[13] P. Zhang, Y. Ban, and A. Nascetti, “Learning U-Net without forgetting for near-real-time wildfire monitoring by the fusion of SAR and optical time series,” Remote Sens. Environ., vol. 261, Aug. 2021, Art. no. 112467.
[14] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv., Springer, Cham, 2015, pp. 234–241.
[15] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A deep convolutional encoder–decoder architecture for image segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 12, pp. 2481–2495, 2017.
[16] H. Zhao et al., “Pyramid Scene Parsing Network,” IEEE Trans. Geosci. Remote Sens., 2016.
[17] L.-C. Chen, G. Papandreou, L. Kokkinos, K. Murphy, and A. L. Yuille, “Semantic image segmentation with deep convolutional nets and fully connected CRFs,” 2014, arXiv:1412.7062.
[18] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834–848, 2017.
[19] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, “Rethinking atrous convolution for semantic image segmentation,” 2017, arXiv:1706.05587.
[20] L. C. Chen et al., “Encoder–Decoder With Atrous Separable Convolution for Semantic Image Segmentation,” Cham, Switzerland: Springer, 2018.
[21] J. Wang et al., “Deep high-resolution representation learning for visual recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 10, pp. 3349–3364, 2020.
[22] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large scale image recognition,” 2014, arXiv:1409.1556.
[23] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 770–778.
[24] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in Proc. Int. Conf. Mach. Learn., PMLR, 2019, pp. 6105–6114.
[25] L. Ding et al., “MP-ResNet: Multipath residual network for the semantic segmentation of high-resolution PolSAR images,” IEEE Geosci. Remote Sens. Lett., vol. 10, Feb. 2021, pp. 1–5.
[26] W. Wu, H. Li, X. Li, H. Guo, and L. Zhang, “PolSAR image semantic segmentation based on deep transfer learning-realizing smooth classification with small training sets,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 6, pp. 977–981, Jun. 2019.
[27] M. Shazhad, M. Maurer, F. Fründorfer, Y. Wang, and X. X. Zhu, “Buildings detection in VHR SAR images using fully convolutional neural networks,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 2, pp. 1100–1116, Feb. 2019.
[28] C. Henry, S. M. Azimi, and N. Merkle, “Road segmentation in SAR satellite images with deep fully-convolutional neural networks,” IEEE Geosci. Remote Sens. Lett., vol. 15, no. 12, pp. 1867–1871, Dec. 2018.
[29] X. Wang, L. Cavaglini, M. Eggimann, M. Magno, and L. Benini, “HR-SARNet: A deep neural network for urban scene segmentation from high-resolution SAR data,” in Proc. IEEE Sensors Appl. Symp., 2020, pp. 1–6.
[30] A. G. Roy, N. Navab, and C. Wachinger, “Concurrent spatial and channel ‘squeeze & excitation’ in fully convolutional networks,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention, Springer, Cham, 2018, pp. 421–429.
[31] Z. Changming et al., “Extract enclosure culture in coastal waters based on high spatial resolution remote sensing image,” J. Dalian Maritime Univ., vol. 37, no. 3, pp. 67–69, Aug. 2011.
B. Sui, T. Jiang, Z. Zhang, X. Pan, and C. Liu, “A modeling method for automatic extraction of offshore aquaculture zones based on semantic segmentation,” ISPRS Int. J. Geo-Inf., vol. 9, 2020, Art. no. 145.

C. Yu et al., “Segmentation and density statistics of mariculture cages from remote sensing images using mask R-CNN,” Inf. Process. Agriculture, vol. 9, pp. 417–430, 2022, doi: 10.1016/j.ipa.2021.04.013.

C. Liang et al., “Semi-weakly-supervised semantic segmentation method and its application for coastal aquaculture areas based on multi-source remote sensing images—Taking the Fujian coastal area (mainly Sanduo) as an example,” Remote Sens., vol. 13, 2021, Art. no. 1083.

Y. Zhang, C. Wang, J. Chen, and F. Wang, “Shape-constrained method of remote sensing monitoring of marine raft aquaculture areas on multi-temporal synthetic sentinel-1 imagery,” Remote Sens., vol. 14, 2022, Art. no. 1249.

J. Fan, J. Zhao, W. An, and Y. Hu, “Marine floating raft aquaculture detection of GF-3 PolSAR images based on collective multikernel fuzzy clustering,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 12, no. 8, pp. 2741–2754, Aug. 2019.

Y. Zhang et al., “Combining segmentation network and nonsubsampled contourlet transform for automatic marine raft aquaculture area extraction from sentinel-1 images,” Remote Sens., vol. 12, 2020, Art. no. 4182.

S. Woo, J. Park, J.-Y. Lee, and I.-S. Kweon, “CBAM: Convolutional block attention module,” in Proc. Eur. Conf. Comput. Vis., 2018, pp. 3–19.

F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully convolutional neural networks for volumetric medical image segmentation,” in Proc. IEEE 4th Int. Conf. 3D Vis., 2016, pp. 565–571.

A. Shrivastava, A. Gupta, and R. Girshick, “Training Region-based object detectors with online hard example mining,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 761–769.

G. Ghiass et al., “Simple copy-paste is a strong data augmentation method for instance segmentation,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2021, pp. 2918–2928.

G. Wang, W. Li, S. Ourselin, and T. Vercauteren, “Automatic brain tumor segmentation using convolutional neural networks with test-time augmentation,” in Proc. Int. MICCAI Brainlesion Workshop, Springer, Cham, 2018, pp. 61–72.

J. Kang et al., “Decoding the partial pretrained networks for sea-ice segmentation of 2021 gaofen challenge,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 15, 2020, pp. 4521–4530.

Gang Qin was born in Henan, China, in 1997. He received the undergraduate academic degree in software engineering from the Zhengzhou University, Zhengzhou, China, in 2020. He is currently working toward the Ph.D. degree in cartography and geographic information systems with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. His research interests include remote sensing image classification and knowledge graph based on deep learning.

Shixin Wang received the B.S. degree in natural resources from the Nanjing University, Nanjing, China, in 1987, and the M.S. degree in cartography and geographical information system from the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 1990. He is currently a Researcher with the Aerospace Information Research Institute, Chinese Academy of Sciences. His research interests include remote sensing of environment and disaster.

Futao Wang received the B.S. degree in resource and environmental engineering from the Shandong University of Technology, Zibo, China, in 2005, the M.S. degree in mapping and geographic information from the Guilin University of Technology, Guilin, China, in 2008, and the Ph.D. degree in cartography and geographical information system from the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2011. He is currently an Associate Research Fellow with the Aerospace Information Research Institute, Chinese Academy of Sciences. His research interests include remote disaster remote sensing monitoring and assessment.

Dr. Wang is currently a member of the Youth Innovation Promotion Association of Chinese Academy of Sciences, and a Reviewer for International Journal of Disaster Risk Science and Journal of Remote Sensing.

Yi Zhou received the B.S. degree in natural resources from the Nanjing University, Nanjing, China, in 1986. She is currently a Researcher with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. Her research interests include remote sensing of environment and disaster, remote sensing of urban systems and urbanization, and quantitative monitoring and evaluation of environmental remote sensing.

Zhengqing Wang was born in Shandong, China, in 1997. He received the B.S. degree in geographic information science from the China University of Petroleum, Qingdao, China, in 2019. He is currently working toward the Ph.D. degree in cartography and geographic information systems with the Airspace Information Research Institute, Chinese Academy of Sciences, Beijing, China and the University of Chinese Academy of Science, Beijing, China. His research interests include remote sensing and GIS and object detection based on deep learning.

Weijie Zou was born in Hunan, China, in 1998. He received the B.S. degree in surveying engineering from the Wuhan University, Wuhan, China, in 2020. He is currently working toward the master’s degree in resources and environment with the Airspace Information Research Institute, Chinese Academy of Sciences, Beijing, China. His research interests include remote sensing semantic segmentation and object detection based on deep learning.