Automatic Extraction of Green Tide From GF-3 SAR Images Based on Feature Selection and Deep Learning

Haifei Yu , Changying Wang , Jinhua Li, and Yi Sui

Abstract—Efficient and accurate monitoring of green tide is of great significance to marine disaster prevention and marine environment protection. A method is proposed in this article for the automatic extraction of the green tide from Chinese Gaofen-3 (GF-3) satellite synthetic aperture radar (SAR) images, which is based on feature selection and deep learning. In this article, since SAR images contain rich polarization information, we first employ H/A/α decomposition and other methods for the extraction of high-dimensional features from GF-3 SAR images. Second, a novel feature selection method for SAR images is designed by using the Bhattacharyya distance and the Separability index, which can select the optimal features subset with a strong ability for recognizing green tide and without correlation between features from the high-dimensional features of SAR. Then, to alleviate the model training burden and improve the prediction efficiency, a lightweight semantic segmentation network, called Mobile-SegNet, is designed based on MobileNets and SegNet. Finally, the selected optimal features and their labels are sent to Mobile-SegNet for training and obtaining the automatic recognition model of green tide, and in turn, the automatic extraction of green tide is achieved through model prediction. To verify the effectiveness of the proposed green tide extraction method, GF-3 SAR remote sensing images taken in 2020 that covered the Yellow Sea are collected and used in the green tide extraction experiment. The results show that the proposed method is available for an effective reduction of the feature dimension required for green tide extraction, and the improvement of the accuracy and efficiency of green tide detection. The overall accuracy, F1-score, mean intersection over union, and the kappa coefficient of the proposed method reached 99.52%, 95.76%, 92.19%, and 0.92, respectively.

Index Terms—Deep learning, feature selection, Chinese Gaofen-3 (GF-3) synthetic aperture radar (SAR), green tide extraction, remote sensing.

Haifei Yu and Yi Sui are with the College of Computer Science and Technology, Qingdao University, Qingdao 266071, China (e-mail: 2019025985@qdu.edu.cn; suiyi@qdu.edu.cn).

Changying Wang and Jinhua Li are with the College of Computer Science and Technology, Qingdao University, Qingdao 266071, China, and also with the Engineering Research Center for Intelligent and Accurate Measurement of Shipbuilding of Qingdao, Qingdao 266071, China (e-mail: wcing@qdu.edu.cn; lijh@qdu.edu.cn).

I. Introduction

SINCE 2007, large-scale green tide disasters have broken out in the Yellow Sea for 13 consecutive years, which not only breaks the balance of the marine ecosystem but also has negative impacts on the related industries such as marine fisheries and tourism [1], [2]. Satellite remote sensing characterized by the advantages of synchronization, large range, and rapid observation, presents an essential means for green tide monitoring [3]. Among them, synthetic aperture radar (SAR) can realize all-day and all-weather earth observation due to its lower sensitivity to cloud, rain, and fog, which has attracted more and more attention all over the world [4]. In 2016, the first self-developed C-band multipolarized SAR satellite of China, Gaofen-3 (GF-3) satellite, was successfully launched, which started to collect data on a regular basis. GF-3 satellite is characterized by high resolution, large imaging width, high radiation accuracy, and multiple imaging modes, which has been widely used in marine monitoring [5], [6], disaster monitoring [7], [8], resource exploration [9], [10], and other fields.

A large number of studies have shown that SAR images have the ability to extract green tide. Li *et al*. [11] proposed an image analysis method based on gray value and experimentally found that this method could effectively extract green tide from RADARSAT-1, ENVISAT-ASAR, and ALOS-PALSAR data. Shen *et al*. [12] proposed new green tide detection indexes for Co-polarization (HH/VV) and Cross-polarization (HV/VH) of RADARSAT-2 images, respectively. Cui *et al*. [13] employed the images from ENVISAT-ASAR and HJ-1A/B to monitor the process of green tide outbreak in the Yellow Sea and found that the average difference between the two kinds of remote sensing images is 15%. Yu *et al*. [14] presented a green tide automatic detection method based on adaptive thresholds for GF-3 SAR images and pointed out that HV is more suitable for the extraction of green tide than HH due to the lower noise level of the former. Song *et al*. [15] comprehensively used the Markov random field (MRF) and the small univalue segment assimilating nucleus to extract green tide from RADARSAT-2 images, thereby improving the extraction accuracy by 18% compared with the independent utilization of MRF. Geng *et al*. [16] selected the features of GF-3 SAR images with the help of the importance score and extracted green tide information by using Random Forest algorithm.

It can be seen that the methods for green tide extraction based on SAR images are mainly divided into threshold segmentation and machine learning classification at present. The former is simple in principle and fast in implementation, which, however, cannot make full use of the hidden information of images. In practice, it is difficult to find the optimal segmentation threshold due to the fact that threshold segmentation methods are susceptible to speckle noise and uneven gray distribution in SAR images [11]. While machine learning can be used to extract useful rules and knowledge from a large amount of incomplete random data [17]. Although traditional machine learning classification methods for green tide extraction can achieve high automation, they still show some disadvantages, such as complex implementation steps, too many hyperparameters, and great influence of human factors, moreover they usually require a combination of multiple machine learning algorithms to obtain better extraction results. In addition, some researchers extract green tide with the help of data mining [3], object-oriented segmentation [4], and level-set segmentation [18], etc. Although these methods exhibit high extraction accuracy for the green tide, they usually need post-processing of the extraction results, resulting in low efficiency. And like threshold segmentation and machine learning methods, they do not take into account the rich semantic information in SAR images. In conclusion, most existing methods fail to obtain good extraction results for large-scale green tide monitoring tasks.

In recent years, semantic segmentation technology based on deep learning has gradually become the mainstream method to solve the problem of the automatic interpretation of high-resolution remote sensing images due to its strong adaptability, high accuracy, and high automation [19], which is increasingly applied to SAR image processing, such as ground objects classification [17], [20], [21], target recognition [22], [23], and information extraction [24], [25]. However, SAR images are under-explored in the field of green tide semantic segmentation. Despite the increasingly available GF-3 SAR images, there are no large-scale and well-annotated green tide dataset now, which may hinder the further development of green tide semantic segmentation with GF-3 SAR images by using deep learning to a certain extent. Besides, the complex imaging mechanism makes SAR images contain rich information, but the high-dimensional features of SAR images are easy to cause the curse of dimensionality, which aggravates the learning burden of network models, wastes time-space cost, and computing resources. The optimization of high-dimensional features can effectively reduce the redundancy between features and improve the classification efficiency [16]. In response to the above problems, taking GF-3 dual-polarization SAR images as the data source, in this article, we propose a novel green tide automatic extraction method for GF-3 SAR images based on feature selection and semantic segmentation. The main contributions of this work are summarized as follows.

1) Combined with the Bhattacharyya distance (BD) and the Separability index (SI), a novel feature selection method for SAR remote sensing images is proposed, which can select the feature subset with the strongest ability to distinguish target ground objects from high-dimensional feature
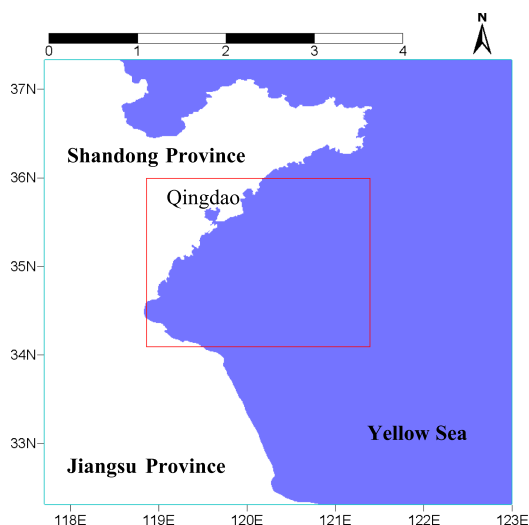

Fig. 1.   Red box shows the green tide study area.

space of SAR images, and the features in the subset have low or no correlation.

2) Based on MobileNets and SegNet, we designed a lightweight semantic segmentation network in an encoder–decoder structure, called Mobile-SegNet. The network mainly employs the depthwise separable convolution for the reduction of parameter scale of the model and adopts the down-sampling indices for connecting encoder and decoder to transmit spatial information and strengthen weak green tide boundary.

3) A novel green tide automatic extraction method is proposed for GF-3 SAR images by combining with the above feature selection method and Mobile-SegNet network.

4) We constructed a standard semantic segmentation dataset to advance green tide extraction by SAR images. The dataset contains 2694 hand-labeled 256 pixels $\times$ 256 pixels GF-3 SAR images with a 25-m resolution and 26 bands.

The rest of this article is organized as follows. Section II briefly introduces the study area, data source, and data preprocessing. Section III introduces the proposed automatic extraction method of the green tide in detail. Section IV gives the green tide extraction experiment, the result analysis, and discussion. Finally, Section V concludes this article.

II. DATA

A. Study Area and Data Source

Green tide is the major marine disaster affecting the coast of the Yellow Sea in China, which mainly occurs from May to August every year [26]. The area of this article focuses on is the coastal waters of Qingdao, China, with the coordinates range of 119°–121.5°E and 34°–36°N, as shown in Fig. 1. In this article, we collect GF-3 SAR level-1A (L1A) data as the experimental data, including six dual-polarization (VH/VV) images taken from late May to early July in 2020, as shown in Fig. 2. The imaging mode of experimental data is the standard strip with
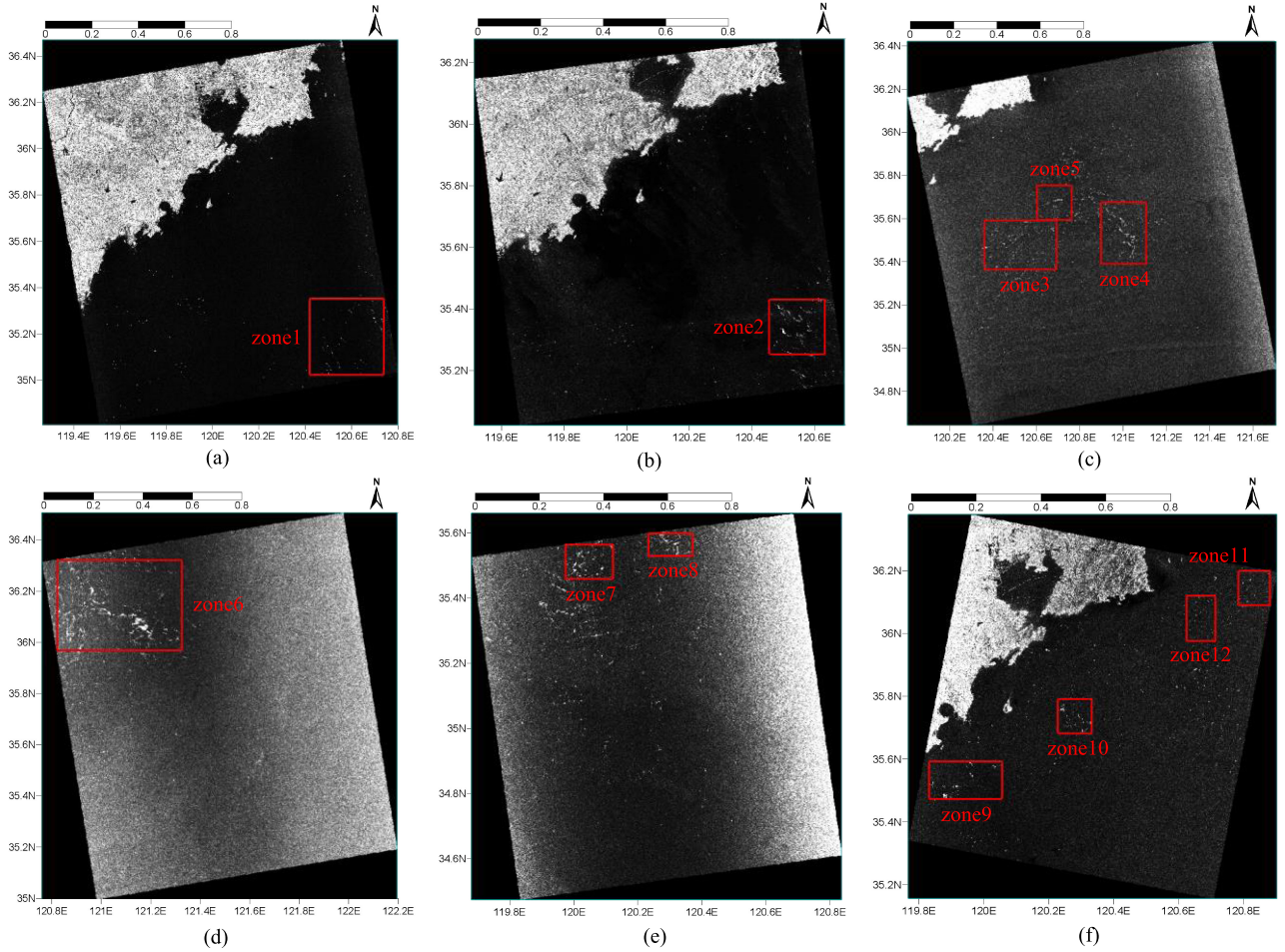
Fig. 2.   GF-3 SAR images. Red box shows the region of interest (ROI).

TABLE I
BASIC INFORMATION OF GF-3 SAR IMAGES

Number	Product ID	Track Direction	Height	Width	Central Coordinates	Imaging Time
(a)	4835808	ASC	30384	28602	120.2E, 35.6N	2020.05.31
(b)	4841258	ASC	21656	18982	120.2E, 35.7N	2020.06.02
(c)	4847620	ASC	34011	32060	121.0E, 35.5N	2020.06.05
(d)	4865659	ASC	25383	23816	121.6E, 35.8N	2020.06.12
(e)	4880659	ASC	23337	22150	120.4E, 35.1N	2020.06.19
(f)	4905359	DEC	24105	23023	120.5E, 35.8N	2020.07.04

a spatial resolution of 25 m and an imaging width of 130 km. Table I lists the basic information of each image.

Note: ASC and DEC denote the ascending and descending trajectories, respectively, and both height and width are in pixels.

B. Data Pre-Processing

1) Raw Data Calibration and Format Conversion: The GF-3 SAR data collected are the L1A product, i.e., the single look complex (SLC) images. Due to the existence of various error sources, there are radiation errors in them. To accurately reflect the echo signal of ground objects, the radiometric calibration is carried out on the SLC images. After that, the calibrated SLC images are converted into the polarization scattering matrix (S_2) which can be interpreted by using the computer.

2) Polarization Matrix Conversion: The polarization scattering matrix reflects the scattering information of different linear polarization states. The S_2 of the VHVV dual-polarization image is expressed as follows:

$$S_2 = \begin{pmatrix} 0 & 0 \\ s_{vh} & s_{vv} \end{pmatrix} \tag{1}$$

where s_{vh} represents that electromagnetic waves transmitted and received are V-polarized and H-polarized, respectively, while s_{vv}

represents that both the transmitted and received electromagnetic waves are V-polarized.

The polarization scattering vector (K_2) can be obtained by vectorizing S_2 based on the Lexicographic basis, i.e.,

$$K_2 = [s_{vh}, s_{vv}]^{\mathrm{T}} \tag{2}$$

where the superscript $^{\mathrm{T}}$ refers to the matrix transpose.

Then, the polarization covariance matrix (C_2) can be expressed as follows:

$$C_2 = K_2 K_2^{*\mathrm{T}} = \begin{pmatrix} \left\langle |s_{vh}|^2 \right\rangle & \left\langle s_{vh} s_{vv}^* \right\rangle \\ \left\langle s_{vv} s_{vh}^* \right\rangle & \left\langle |s_{vv}|^2 \right\rangle \end{pmatrix} = \begin{pmatrix} c_{vh} & c_{\mathrm{real}} \\ c_{\mathrm{imag}} & c_{vv} \end{pmatrix} \tag{3}$$

where the superscript * represents the complex conjugate; c_{vh}, c_{vv}, c_{real}, and c_{imag} stand for the VH polarization component, VV polarization component, real part, and imaginary part of C_2, respectively.

3) Filtering of SAR Images: To suppress the speckle noise and enhance the interpretability of ground objects in SAR images, the refined lee filter (RLF) [27] available to effectively preserve the polarization information is utilized to filter C_2. RLF can not only suppress the noise of the main diagonal of C_2 but also filter its subdiagonal, thereby effectively solving the problem that the noise of the subdiagonal elements of C_2 is neither a multiplicative model nor an additive model.

4) Geocoding and Orthorectification: To eliminate any geometric errors that may have been introduced by the imaging, the rational polynomial coefficient attached to the GF-3 metadata file is employed for the geometric correction of images, and then the Geocoded-Ellipsoid-Corrected (GEC) L2 images are generated. After that, the global digital elevation model provided by ENVI [28] is used for the orthorectification of the GEC images before the generation of the Geocoded-Terrain-Corrected (GTC) L3 images.

5) Making Labels: Based on the optical images obtained from the coastal zone imager (CZI) of the Haiyang-1C (HY-1C) satellite in the same area and at the adjacent time, the position and shape of green tide in ROIs are roughly determined. After that, the ENVI software is used for the manual visual interpretation of 12 ROIs in the GF-3 SAR GTC images, thus obtaining the ground truth labels, as shown in Fig. 3.

III. METHOD

The new method for the automatic extraction of green tide from GF-3 SAR images proposed in this article has three stages, as shown in Fig. 4. The first stage, i.e., the feature analysis stage, mainly consists of data pre-processing, feature extraction, and feature selection of GF-3 SAR images. The second stage, i.e., the model training stage is used to divide and enhance the dataset, and to obtain the automatic recognition model of the green tide by training the Mobile-SegNet network. The last stage is the green tide extraction stage, as the final practical application stage, it is able to automatically extract the green tide from unlabeled GF-3 SAR images through model prediction.
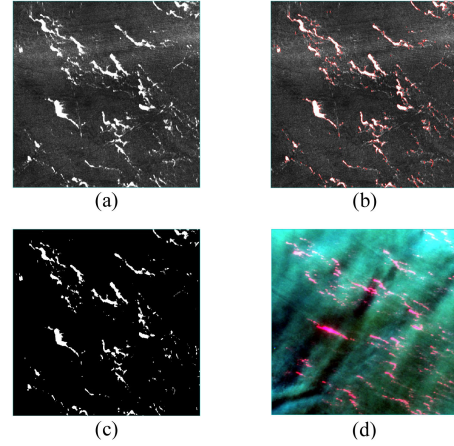

Fig. 3. Examples of visual interpretation (zone2). (a) Input: GF-3 GTC image (taken on 2020-06-02 02:10:55 UTC, 25-m resolution). (b) Process: artificially interpreted green tide shape. (c) Output: ground truth label. (d) Reference: HY-1C CZI image (taken on 2020-06-02 02:50:21 UTC, 50 m resolution).

A. Feature Extraction

The change of tone in SAR images mainly depends on the backscattering of ground objects. Each received backscattering echo will be converted into an electrical signal by the radar, which is then recorded in a specific gray tone as a digital pixel with a specific digital number used to represent the brightness [29]. Roughness is the main factor affecting radar backscattering [15]. The surface of seawater is smooth, which mainly produces surface scattering, indicating that the backscattering ability of seawater is weak. On the contrary, the rough surface of green tide mainly produces volume scattering, which provides it with a strong backscattering ability [14]. The backscattering intensity of green tide is higher than that of seawater, resulting in the fact that green tide is brighter than seawater in SAR images, which makes it possible to extract the green tide from the seawater. However, the degree of this brightness difference in various features of SAR image is inconsistent, therefore, we have to extract as many features of SAR image as possible before optimization.

1) Polarization Features: Polarization information is the most basic feature of polarimetric SAR (PolSAR), which reflects the scattering mechanism of targets. With the assistance of the pixel information expert of SAR software [30], S_2 is transformed into real part feature (Real_{vh} and Real_{vv}), imaginary part feature (Imag_{vh} and Imag_{vv}), amplitude feature (Amp_{vh} and Amp_{vv}), phase feature (Phase_{vh} and Phase_{vv}) as well as backscatter coefficient feature (DB_{vh} and DB_{vv}). The formulas of amplitude, phase, and backscatter coefficient are as follows:

$$\mathrm{Amp} = \sqrt{r^2 + i^2} \tag{4}$$

$$\mathrm{Phase} = \arctan\left(\frac{i}{r}\right) \tag{5}$$

$$DB = \left(r^2 + i^2\right) \times \frac{Q}{32767} \times \frac{1}{k} \tag{6}$$

where r and i represent the real part and imaginary part of each pixel in the S_2 complex image, respectively; Q and k represent
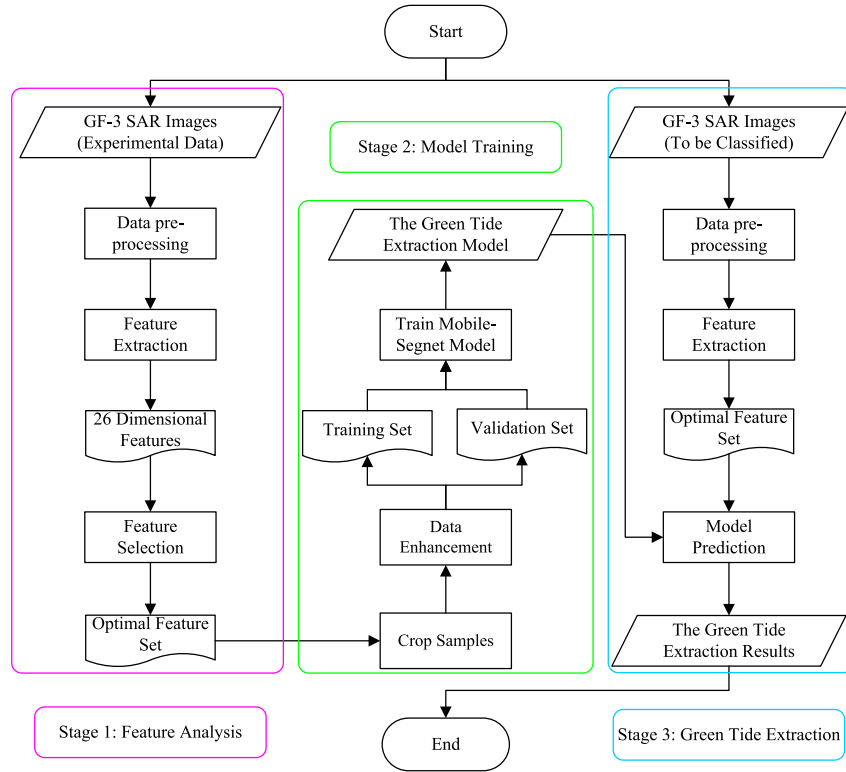
Fig. 4.　Flow of automatic extraction of the green tide from GF-3 SAR images.

the scale factor and calibration coefficient given in the GF-3 metadata file, respectively [31].

2) Decomposition Features: Cloude and Pottier [32] proposed the H/A/α decomposition method. This method is a PolSAR target decomposition method based on eigenvalues and eigenvectors [33], which is available to effectively explain the radiation and scattering characteristics of target ground objects. The decomposition of C_2 can be expressed as follows:

$$C_2 = \sum_{i=1}^{2} \lambda_i c_i = \lambda_1 e_1 e_1^* + \lambda_2 e_2 e_2^* \tag{7}$$

where λ_i and e_i refer to the real eigenvalue and eigenvector of C_2, respectively; c_i represents an independent covariance matrix with rank 1, denoting a scattering mechanism (VH or VV), and its corresponding λ_i represents the intensity of this scattering mechanism [34]. Cloude and Pottier defined three parameters (H, A, and α) [16], [32] to simplify the analysis of the physical information provided by decomposition.

The polarimetric entropy H (HAa_{entropy}) represents the randomness of scattering phenomenon, which can be expressed as

$$H = -\sum_{i=1}^{2} \left(\frac{\lambda_i}{\lambda_1 + \lambda_2} \log_2 \frac{\lambda_i}{\lambda_1 + \lambda_2} \right). \tag{8}$$

The anisotropy A ($HAa_{\mathrm{anisotropy}}$) reflecting the relatively important secondary scattering mechanism can be expressed as

$$A = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}. \tag{9}$$

The scattering angle α (HAa_{alpha}) reflects the average scattering mechanism from surface scattering to dihedral angle scattering, which can be expressed as follows:

$$\alpha = \sum_{i=1}^{2} \left(\frac{\lambda_i}{\lambda_1 + \lambda_2} \cos^{-1}(e_i) \right). \tag{10}$$

The eigenvalue λ (HAa_{lambda}) [30] refer to the mean real eigenvalue of the polarization covariance matrix, which can be expressed as follows:

$$\lambda = \frac{\lambda_1 + \lambda_2}{2}. \tag{11}$$

Besides, the combined parameters of the polarimetric entropy and the anisotropy [16], [35] are as follows:

$$\begin{cases} C_{HA} = H \times A \\ C_{H1SA} = H \times (1 - A) \\ C_{1SHA} = (1 - H) \times A \\ C_{1SH1SA} = (1 - H) \times (1 - A) \end{cases}. \tag{12}$$

3) Index Features: Sentinel-1 dual-polarized water index proposed by Jia *et al.* [36] can effectively enhance the water characteristic in dual-polarization SAR images and eliminate the existence of the nonwater such as soil and vegetation, which is defined as

$$BM_{index} = \ln(10 \times c_{vh} \times c_{vv}). \tag{13}$$

Besides, band sum (BM_{sum}), band difference (BM_{sub}), and band ratio (BM_{ratio}) are our custom index features based on the
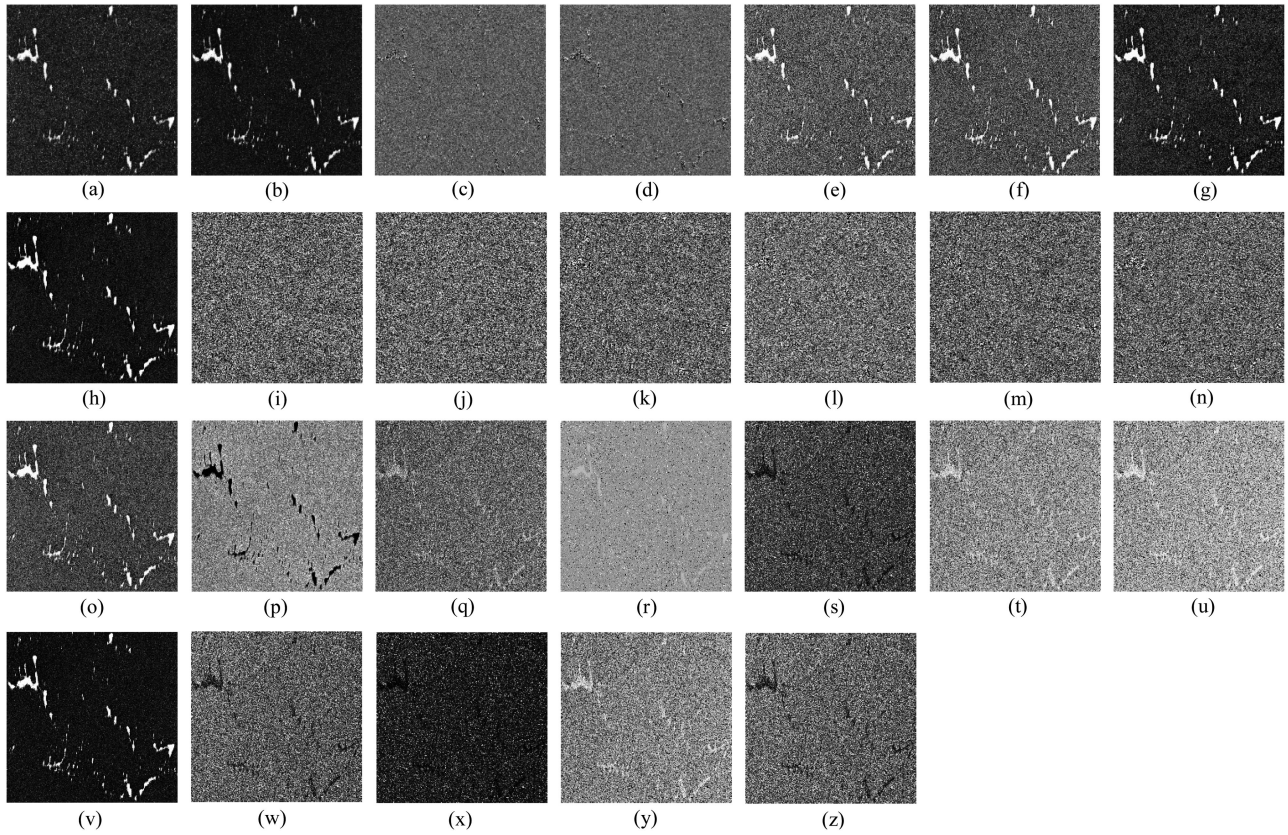
Fig. 5. Examples of 26-dimensional features extracted from GF-3 SAR images (zone10). (a)–(d) are c_{vh}, c_{vv}, c_{real}, and c_{imag}, respectively, (e)–(n) DB_{vh}, DB_{vv}, Amp$_{vh}$, Amp$_{vv}$, Phase$_{vh}$, Phase$_{vv}$, Real$_{vh}$, Real$_{vv}$, Imag$_{vh}$, and Imag$_{vv}$, respectively, (o)–(r) BM_{index}, BM_{sum}, BM_{sub}, and BM_{ratio}, respectively, (s)–(v) HAa_{entropy}, HAa_{alpha}, $HAa_{\text{anisotropy}}$, and HAa_{lambda}, respectively. (w)–(z) are C_{HA}, C_{H1sA}, C_{1sHA}, and C_{1sH1sA}, respectively.

backscattering coefficient, which are defined as follows:

$$\begin{cases} BM_{\text{sum}} = DB_{vh} + DB_{vv} \\ BM_{\text{sub}} = DB_{vh} - DB_{vv} \\ BM_{\text{ratio}} = DB_{vh}/DB_{vv} \end{cases}. \qquad (14)$$

The 26-dimensional features extracted from GF-3 SAR images are shown in Fig. 5. In order to facilitate the subsequent computer interpretation and green tide extraction, the RLF, the geocoding, and the orthorectification are performed on all features.

B. Feature Selection

Feature extraction is the premise of PolSAR image classification. Too small feature dimension may not be enough to reflect the characteristic of ground objects and make it difficult to achieve the expected classification effect; however, in the case that the feature dimension is too large, it may be prone to the curse of dimensionality [22]. Blind fusion of all features will result in data redundancy, which not only reduces the classification accuracy but also increases the computational cost. Therefore, how to choose the optimal features from the high-dimensional feature space becomes particularly important. Based on the BD [37], [38] and the SI [39], we propose a feature selection method for GF-3 SAR images in this article, as shown in Fig. 6.
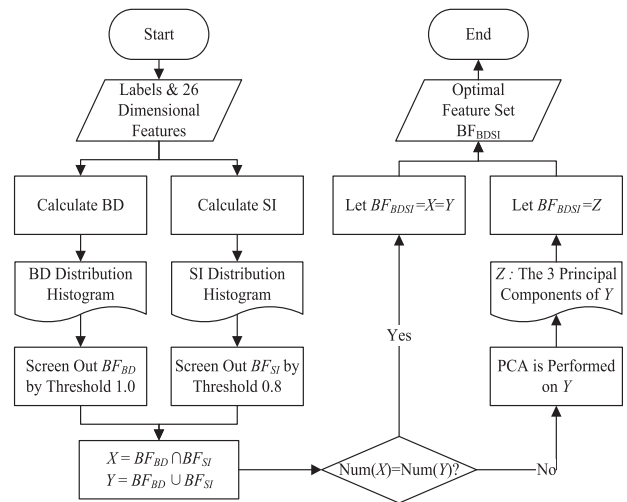

Fig. 6. Flow of the feature selection method based on BD and SI.

1) Feature Selection Method: BD reflects the degree of dispersion between two statistical samples, which can be calculated by the following equation:

$$\text{BD} = \frac{1}{4} \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} + \frac{1}{2} \log_2 \left(\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1 \sigma_2} \right) \qquad (15)$$

where μ_1, σ_1, and μ_2, σ_2 denote the mean and the standard deviation of the two categories on a feature, respectively. The larger the *BD* value of features, the higher the degree of dispersion between ground objects, and the better the divisibility; and the smaller the BD value of features, the higher the similarity between ground objects, and the harder it is to separate them [40].

SI indicates that when the number of classification categories is two or higher, if the distance between the mean values of a feature is greater than that of the standard deviation, this feature is considered to have good divisibility, which can be calculated by the following equation:

$$\mathrm{SI} = \frac{|\mu_1 - \mu_2|}{\sigma_1 + \sigma_2}. \tag{16}$$

In general, the features with SI values between 0.8 and 1.5 are beneficial to the classification, and in the case that the SI value of a feature is greater than 2, it indicates that this feature can almost separate all categories completely [35].

Considering that in the case of multidimensional features as well as good divisibility of ground objects, the independent utilization of SI for feature selection may not effectively remove the redundancy between features, therefore, it is not easy to obtain the optimal features. In this case, we use both BD and SI to select the optimal features from the high-dimensional feature space of GF-3 SAR images. We set the threshold of BD and SI to 1.0 and 0.8, respectively. Mark the feature set with a BD value greater than 1.0 as BF_{BD}, and that with SI value greater than 0.8 as BF_{SI}. According to the maximum principle of BD and SI, the better feature sets X and Y are selected through the following equations:

$$X = BF_{\mathrm{BD}} \cap BF_{\mathrm{SI}} \tag{17}$$

$$Y = BF_{\mathrm{BD}} \cup BF_{\mathrm{SI}}. \tag{18}$$

Obviously, because X is a subset of Y, there are two kinds of quantitative relations between the features contained in X and Y. If X and Y contain the same number of features, then let the optimal feature set $BF_{\mathrm{BDSI}} = X = Y$. When the number of features contained in Y is more than X, it means that the correlation of features in Y has to be removed. The principal component analysis (PCA) [41] is employed to obtain the optimal 3-D features $Y_{\mathrm{PCA} \cdot B1}$, $Y_{\mathrm{PCA} \cdot B2}$, and $Y_{\mathrm{PCA} \cdot B3}$ from Y in this article, then let $BF_{\mathrm{BDSI}} = \{Y_{\mathrm{PCA} \cdot B1}, Y_{\mathrm{PCA} \cdot B2}, Y_{\mathrm{PCA} \cdot B3}\}$.

2) Result Analysis of Feature Selection: According to the process shown in Fig. 6, ten ROIs (zone1, zone3, zone4, and zone6–zone12) are selected for the feature selection analysis. By calculating the BD and the SI of green tide and seawater in 26-dimensional features, the ability of different features to distinguish the two kinds of ground objects is judged, with the calculation results shown in Fig. 7.

It is evident from Fig. 7 that different features have different abilities to distinguish green tide from seawater, besides, both BD and SI have excellent ability to eliminate the features with poor ability to distinguish two target categories, such as the real and imaginary parts of C_2, as well as the phase, the
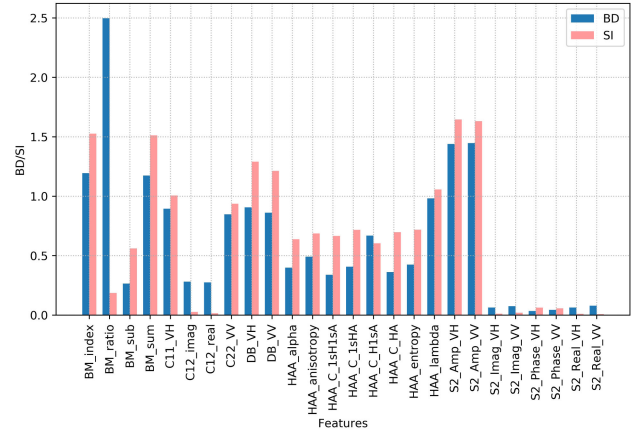

Fig. 7.   Distribution histogram of BD and SI.

real part, and the imaginary part converted from S_2. $BF_{\mathrm{BD}} = \{BM_{\mathrm{index}}, BM_{\mathrm{sum}}, BM_{\mathrm{ratio}}, \mathrm{Amp}_{vh}, \mathrm{Amp}_{vv}\}$ can be selected from the BD distribution map through the threshold of 1.0, while $BF_{\mathrm{SI}} = \{BM_{\mathrm{index}}, BM_{\mathrm{sum}}, \mathrm{Amp}_{vh}, \mathrm{Amp}_{vv}, DB_{vh}, DB_{vv}, c_{vh}, c_{vv}, HAa_{\mathrm{lambda}}\}$ can be selected from the SI distribution map through the threshold of 0.8. It can be seen that there are strong correlation feature combinations in BF_{SI}: (1) c_{vh}, c_{vv}, and BM_{index}, which is directly calculated by c_{vh} and c_{vv}; (2) DB_{vh}, DB_{vv}, and BM_{sum}, which is directly calculated by DB_{vh} and DB_{vv}. Although c_{vh}, c_{vv}, DB_{vh}, and DB_{vv} failed to appear in BF_{BD}, the nondirect related features BM_{index}, BM_{sum}, and BM_{ratio} contained in BF_{BD} could be used to reflect them.

Obviously, the ability of BD to eliminate redundancy is more potent than that of SI. However, on the other hand, SI may be able to better guarantee the integrity of features. Therefore, both BD and SI show comparative advantages, and it is of practical significance to combine the two indexes for feature selection. The intersection result and union result of BF_{BD} and BF_{SI} are $X = \{BM_{\mathrm{index}}, BM_{\mathrm{sum}}, \mathrm{Amp}_{vh}, \mathrm{Amp}_{vv}\}$ and $Y = \{BM_{\mathrm{index}}, BM_{\mathrm{sum}}, BM_{\mathrm{ratio}}, \mathrm{Amp}_{vh}, \mathrm{Amp}_{vv}, DB_{vh}, DB_{vv}, c_{vh}, c_{vv}, HAa_{\mathrm{lambda}}\}$, respectively. It is clear that Y contains more features than X. After that, the PCA algorithm converts the better feature set Y to the optimal feature set $BF_{\mathrm{BDSI}} = \{Y_{\mathrm{PCA} \cdot B1}, Y_{\mathrm{PCA} \cdot B2}, Y_{\mathrm{PCA} \cdot B3}\}$, as shown in Fig. 8.

C. Mobile-SegNet

Considering the multidimensional features and large data scale of SAR remote sensing image, we use MobileNets to improve SegNet and design a lightweight semantic segmentation network Mobile-SegNet to achieve model compression and improve the operation speed while ensuring high accuracy. MobileNets [42] is a lightweight convolutional neural network for mobile devices, in which the standard convolution layer is replaced by the depthwise separable convolution layer composed of depthwise convolution and pointwise convolution, thereby reducing the parameter quantity and computation cost, as shown in Fig. 9.
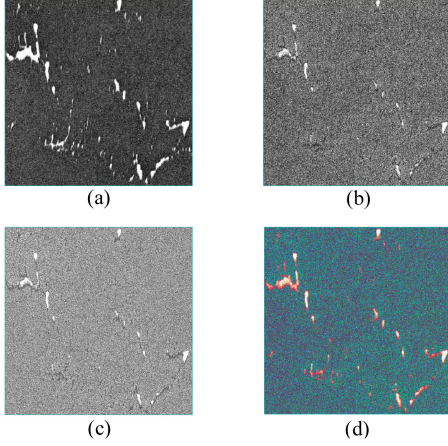
Fig. 8. Examples of the PCA results(zone10). (a) First principal component $Y_{\mathrm{PCA} \cdot B1}$. (b) Second principal component $Y_{\mathrm{PCA} \cdot B2}$. (c) Third principal component $Y_{\mathrm{PCA} \cdot B3}$. (d) False color composite image of $Y_{\mathrm{PCA} \cdot B1}$, $Y_{\mathrm{PCA} \cdot B2}$, and $Y_{\mathrm{PCA} \cdot B3}$.
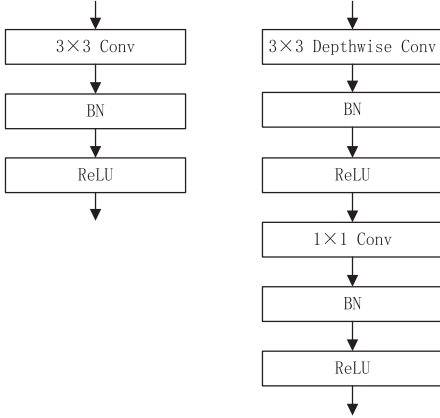

Fig. 9. Structure of the standard convolution layer (left) and the depthwise separable convolution layer (right).

The pointwise convolution is essentially a standard convolution with a convolution kernel size of 1×1, which is responsible for merging features extracted by depthwise convolution. The calculation amount of the standard convolution layer (3×3 Conv) is expressed as follows:

$$F \times F \times M \times N \times K \times K \qquad (19)$$

where F and K refer to the size of feature map and convolution kernel, respectively, M and N represent the number of input channels and output channels, respectively. In the meantime, the calculation amount of the depthwise separable convolution layer (3×3 depthwise Conv $+ 1 \times 1$ Conv) is expressed as follows:

$$F \times F \times M \times K \times K + F \times F \times M \times N \times 1 \times 1. \qquad (20)$$

Although the output results of the standard convolution layer and the depthwise separable convolution layer are identical, compared with the standard convolution layer, the calculation amount of the latter is reduced by $(1/N + 1/K^2)$ times.

Similar to SegNet, Mobile-SegNet is also a symmetric shape semantic segmentation network in an encoder–decoder structure. The encoder is used for the effective data compression and extraction of deep abstract features from the input image. While the decoder is responsible for the gradual recovery of the feature map output by the encoder to the original size of the input image. A Softmax classifier is connected behind the decoder to obtain the category probability information of each pixel, thereby outputting the segmentation results.

SegNet [43] shown in Fig. 10 uses the first 13 layers of VGG16 [44] as the encoder, including the standard convolution layer and the max-pooling layer, among them, the standard convolution layer consists of convolution, batch normalization, and ReLU activation function. And the decoder of SegNet consists of the up-sampling layer and the standard convolution layer.

In Mobile-SegNet, the encoder adopts the first 23 layers of MobileNets, including the standard convolution layer and the depthwise separable convolution layer, and the decoder uses the decoder of SegNet, as shown in Fig. 11.

Different from SegNet which uses max-pooling for image down-sampling, Mobile-SegNet sets the strides of the convolution contained in the standard convolution layer and the depthwise convolution contained in depthwise separable convolution layer as 2 to achieve the image down-sampling. As can be seen from Figs. 10 and 11, unlike the five times down-sampling of VGG16, MobileNets has only four times down-sampling, in this case, the image information loss of the encoder in Mobile-SegNet is less than that in SegNet. In order to compensate for the image detail loss caused by the down-sampling, and further reduce the number of training parameters, we only set a standard convolution layer after each up-sampling of the feature map to improve the geometric shape of the object in the decoder. In addition, the down-sampling indices of Mobile-SegNet are designed, which is similar to the pooling indices of SegNet. As a separate path, the down-sampling indices can record the spatial information during the down-sampling of the encoder and transfer it to the decoder for assisting the up-sampling, thereby reducing the training parameters and improving the smoothness of the image edge.

IV. EXPERIMENTS AND ANALYSIS

In order to verify the effectiveness of the proposed method, the green tide extraction experiment was carried out following the process shown in Fig. 4. For one thing, the semantic segmentation network SegNet and Res-SegNet [45] based on VGG16 and ResNet50 [46], respectively, are used as the comparison network of the designed Mobile-SegNet. For another, the backscattering coefficient, as a commonly used gray feature in SAR image classification, is constructed into a 2-D feature set $F_{DB} = \{DB_{vh}, DB_{vv}\}$. We employ the 4-D feature set $F_{\mathrm{BDSI}} = X = \{BM_{\mathrm{index}}, BM_{\mathrm{sum}}, \mathrm{Amp}_{vh}, \mathrm{Amp}_{vv}\}$ and F_{DB} as the comparison features of the 3-D feature set BF_{BDSI} screened by our method. We gain three models through the training of Mobile-SegNet on BF_{BDSI}, F_{BDSI}, and F_{DB}. Similarly, we can also get three models each from SegNet and Res-SegNet. Therefore, nine
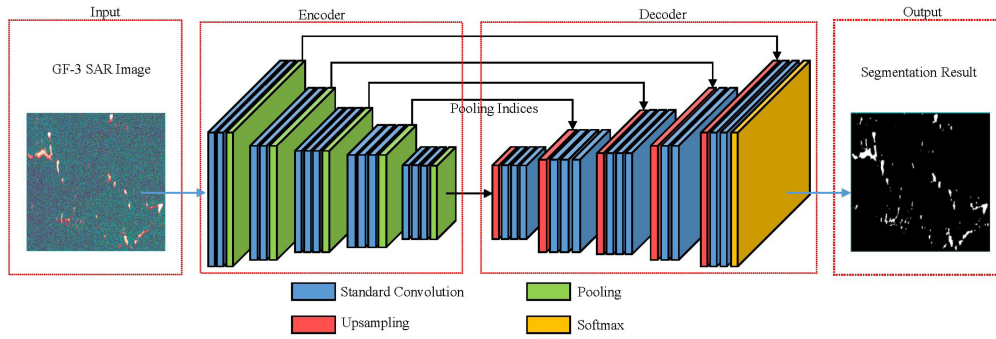
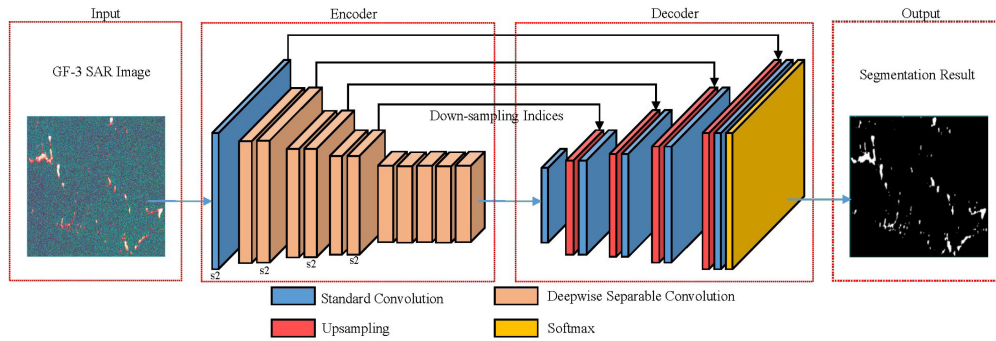Fig. 10.   Structure of SegNet semantic segmentation network.

Fig. 11.   Structure of Mobile-SegNet semantic segmentation network.

TABLE II
DATASET PARTITION

Data	Number of samples (original)	Proportion	Data enhance	Number of samples (enhanced)
Training set	1962	70%	√	7848
Validation set	490	20%	√	1960
Test set	242	10%	×	242
Total	2694			10050

green tide extraction models (*M*1–*M*9) were obtained in this experiment.

A. Experiments Settings

1) Dataset Partition and Enhancement: The 12 ROIs are cropped into subimage samples using a window size of 256 pixels × 256 pixels and a step size of 256 pixels. The ROIs of zone1, zone3, zone4, and zone6-zone12 were cut into 2452 samples, 80% of which were randomly selected for model training, and the remaining 20% were used for validation during training. The ROIs of zone2 and zone5 were cut into 242 samples, all of which were used for model prediction, as shown in Table II. The training set and verification set are enhanced by geometric transformation, including horizontal flip, vertical flip, and diagonal flip, in this manner to avoid overfitting in the process of model training, as shown in Fig. 12.

2) Experimental Environment and Related Settings: The experiments are implemented in a computer with the configuration of Intel Core i7-8750H CPU (2.20GHz) and NVIDIA GeForce GTX 1060 GPU (6GB). In this article, we adopt TensorFlow 2.1 (GPU version) [47] as the deep learning framework. And CUDA10.0 and CuDNN7.6 are selected for GPU parallel calculation and GPU acceleration, respectively. During the training process, the batch size and the number of epochs are set to 4 and 100, respectively. All models adopt the Adam [48] optimizer at the learning rate of 0.001 and use the cross-entropy as the loss function. The cross-entropy is defined as follows:

$$\text{Loss} = -\frac{1}{n}\sum_{i=1}^{n}(y_i \times \ln a) + (1 - y_i)\ln(1 - a) \qquad (21)$$

$$a = P(y_i = 1 | x_i) \qquad (22)$$

where *n* denotes the number of classification categories, and *a* represents the probability that the label category of x_i is $y_i = 1$.

3) Accuracy Evaluation Metrics: For the binary classification problem, the confusion matrix is shown in Table III. In this article, the positive/true case is green tide, while the negative/false case is seawater. We utilize four metrics based on the confusion matrix, including the overall accuracy (OA), F1-score, the mean intersection over union (MIoU), and the kappa coefficient, to compare the performance of Mobile-SegNet, SegNet,
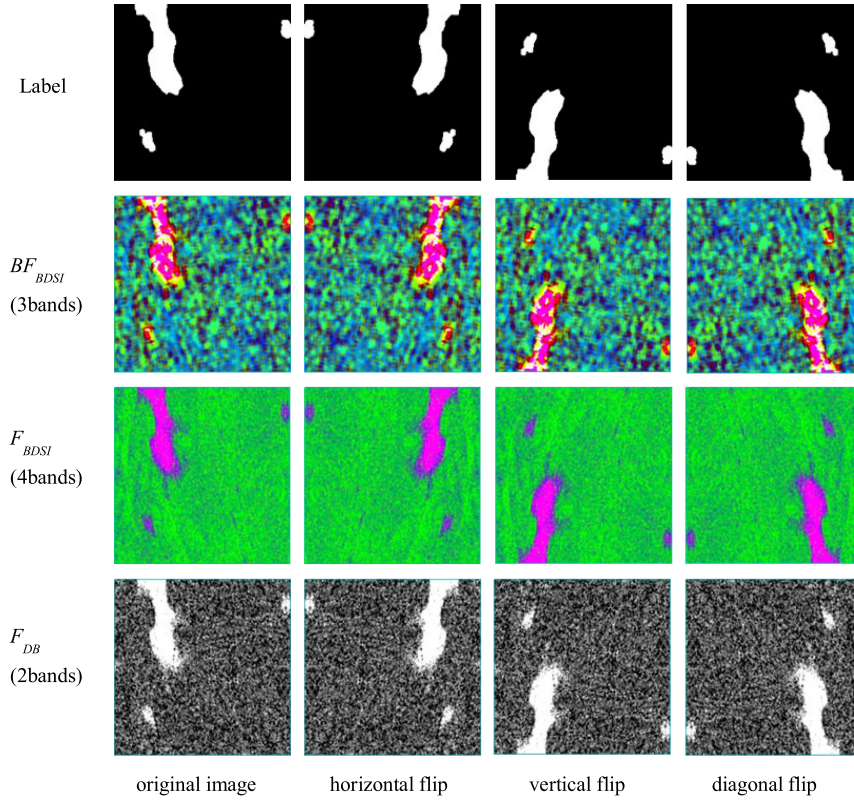
Fig. 12.   Example of dataset enhancement.

TABLE III
CONFUSION MATRIX OF BINARY CLASSIFICATION PROBLEM

Real Situation	Prediction Situation	
	Positive	Negative
True	TP	FN
False	FP	TN

and Res-SegNet on three feature sets of BF_{BDSI}, F_{BDSI}, and F_{DB}.

The overall accuracy refers to the percentage of the pixels that were correctly classified in all pixels [49], which can be calculated as follows:

$$OA = \frac{TP + TN}{TP + FP + TN + FN}. \tag{23}$$

The precision represents the correct pixels over the prediction results, and the recall represents the correct pixels over the labels [50], which can be calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{24}$$

$$\text{Recall} = \frac{TP}{TP + FN}. \tag{25}$$

And F1-score refers to the harmonic mean of precision and recall, which can be used to measure the accuracy of the binary classification model, expressed as follows:

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2TP}{2TP + FN + FP}. \tag{26}$$

MIoU is also known as the Jaccard coefficient, and the higher the value of MIoU, the better semantic segmentation performance [21], as defined by the following equation:

$$\text{MIoU} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{TP_i}{TP_i + FP_i + FN_i} \right). \tag{27}$$

Kappa coefficient is a powerful metric for the coincidence degree between two images, which can be expressed as:

$$\text{Kappa} = \frac{OA - P_e}{1 - P_e} \tag{28}$$

where P_e denotes the proportion of misclassification caused by accidental factors [51], which is defined as

$$P_e = \frac{\sum_{i=1}^{n} \left[(TP_i + FN_i) \times (TP_i + FP_i) \right]}{(TP + FP + TN + FN)^2}. \tag{29}$$

B. Experiments Results

1) Model Training Results: Fig. 13(a)–(c) shows the accuracy of SegNet, Res-SegNet, and Mobile-SegNet during the training and verification of the three feature sets with the change of epoch, respectively, and Fig. 13(d)–(f) illustrates their loss with the change of epoch.
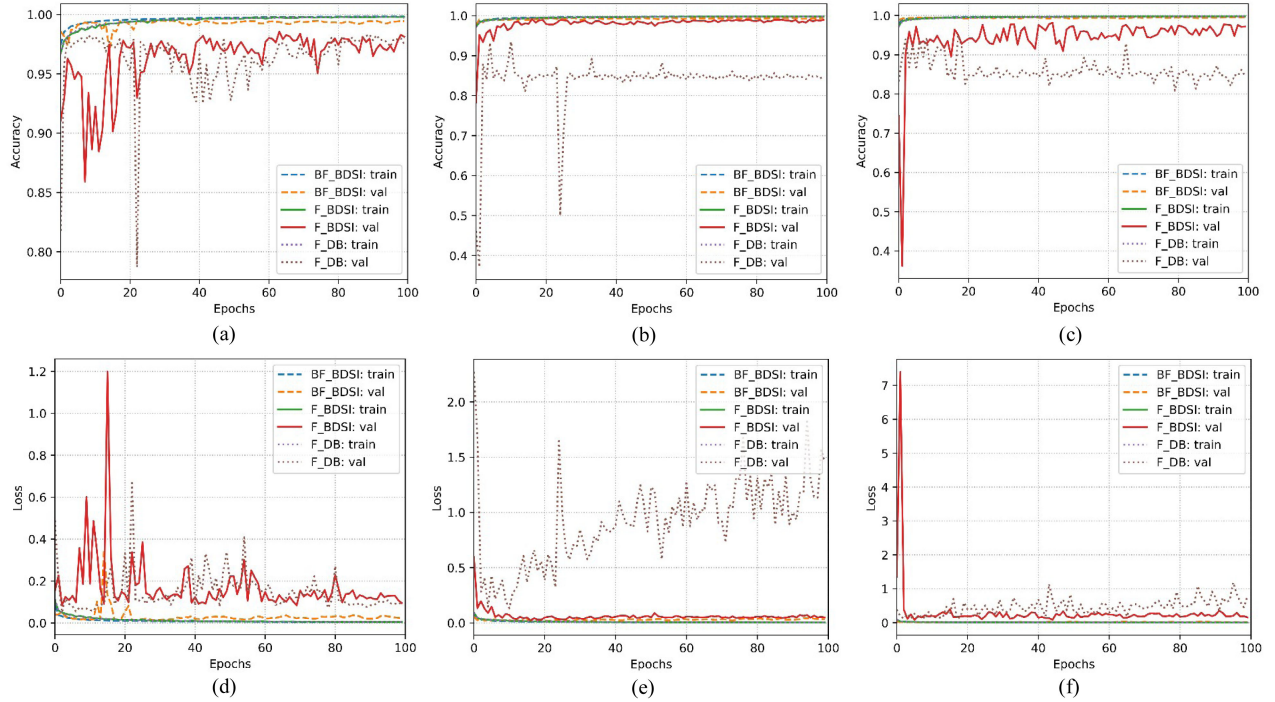
Fig. 13.　Accuracy and loss curve of model training.

TABLE IV
COMPARISON OF MODEL PERFORMANCE

Model	Method	Feature Set	Train Time(h)	Predict Time(s)	Size (MB)
M1		BF_{BDSI}	26.27	12.08	
M2	SegNet	F_{BDSI}	31.07	12.27	364
M3		F_{DB}	26.88	11.99	
M4	Res-SegNet	BF_{BDSI}	18.62	9.51	
M5		F_{BDSI}	26.22	9.97	196
M6		F_{DB}	18.98	9.38	
M7	Mobile-SegNet	BF_{BDSI}	13.61	6.45	
M8		F_{BDSI}	26.22	6.63	65
M9		F_{DB}	12.11	6.34	

It can be seen that the accuracy and loss of all models tend to be flat after training 20 epochs, and all models gradually reach the convergence state with the increase of training epoch. It is worth noting that the verification accuracy and loss of BF_{BDSI} on the three networks are almost the same as the training accuracy and loss. However, the verification accuracy and loss of F_{DB} on the three networks are quite different from the training accuracy and loss, and the verification accuracy and loss of F_{BDSI} on SegNet and Mobile-SegNet are also different from the training accuracy and loss. Therefore, BF_{BDSI} is superior to F_{DB} and F_{BDSI} in terms of the applicability of features to networks.

Table IV shows a comparison of the performance of the three networks on the three feature sets in terms of model training time, prediction time, and parameter scale. It can be seen that SegNet has the most parameters of up to 364 MB, which is followed by Res-SegNet with 196 MB. Mobile-SegNet is the lightest with only 65 MB, that is, about one-sixth of SegNet. Overall, the

training time and prediction time of the model decrease with the reduction of network parameters. In each network, F_{BDSI} takes the longest time for training, whereas the training time of BF_{BDSI} and F_{DB} is fewer and similar. In a single network, the prediction time of different feature sets does not change much, which is, however, proportional to the feature dimension of the feature set. From the perspective of time cost and space cost, the Mobile-SegNet network outperforms SegNet and Res-SegNet.

2) Green Tide Extraction Results: The nine models trained by the three networks on three feature sets are utilized for the prediction of the test set, respectively, and the predictions are restored to the shape of zone2 and zone5 by splicing, thereby obtaining the final green tide extraction results, as shown in Fig. 14.

By calculating the accuracy evaluation metrics, we quantitatively analyzed the green tide extraction results. As shown in Table V, when using SegNet to train F_{DB}, the accuracy of F_{DB} can reach the highest value, and those for OA, F1-score, MIoU, and Kappa are 98.67%, 86.24%, 78.92%, and 0.73, respectively; when using Res-SegNet for training F_{BDSI}, the accuracy of F_{BDSI} can reach the highest value, and those for OA, F1-score, MIoU, and Kappa are 98.69%, 86.45%, 79.03%, and 0.73, respectively. However, OA of F_{DB} and F_{BDSI} in their best performance is less than 99%, and Kappa is less than 0.8. In this experiment, regardless of which network BF_{BDSI} is on, OA can still reach more than 99%, and Kappa is more than 0.8, indicating the better classification effect. Therefore, compared with F_{BDSI} and F_{DB}, BF_{BDSI} has a stronger ability to distinguish green tide from seawater and a better network adaptability.

The highest OA value in Table V is 99.53%, which comes from $M1$ (BF_{BDSI} + SegNet), while the highest value of F1-score,
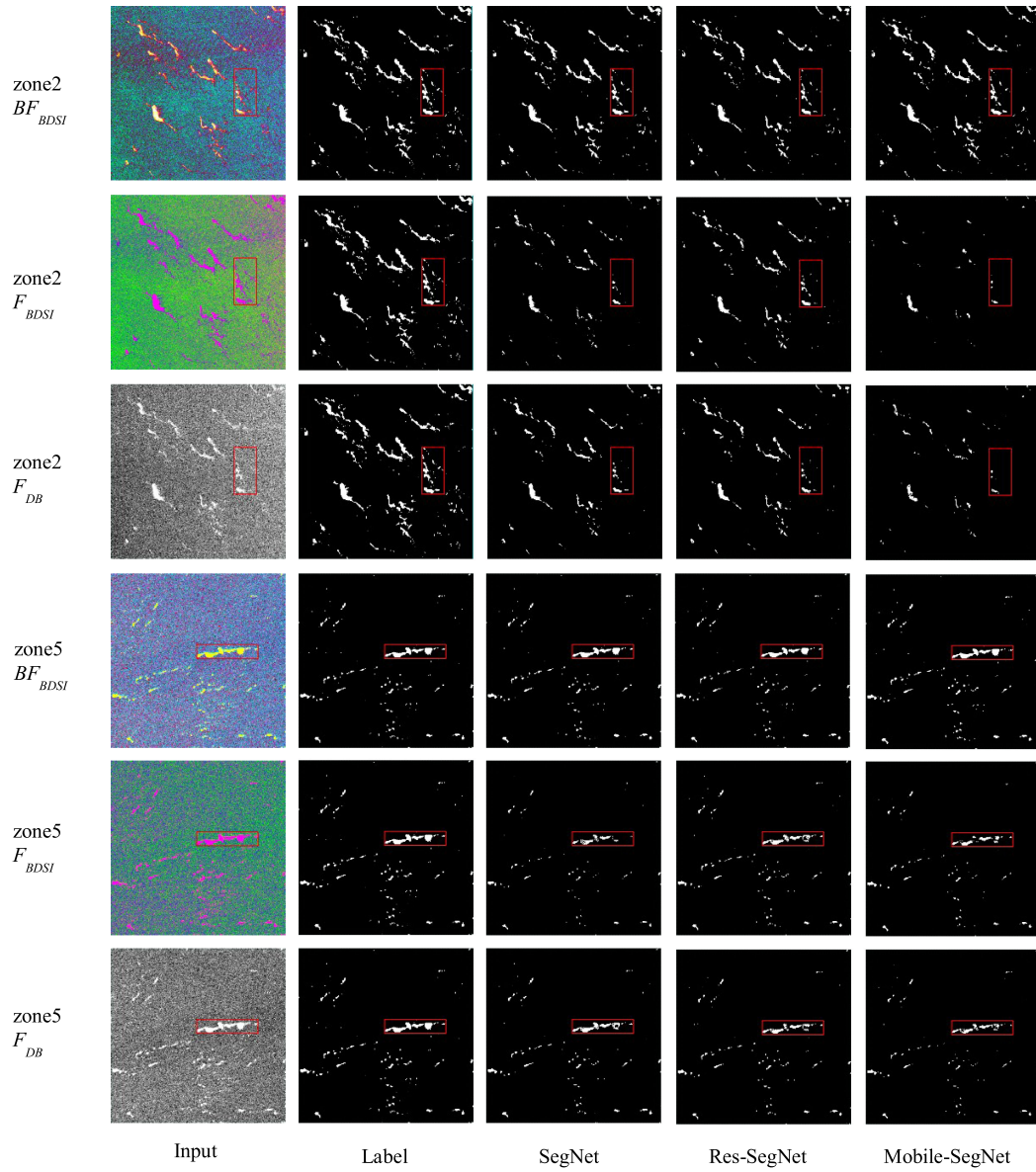
Fig. 14.   Green tide extraction results.

TABLE V
ACCURACY EVALUATION OF GREEN TIDE EXTRACTION RESULTS

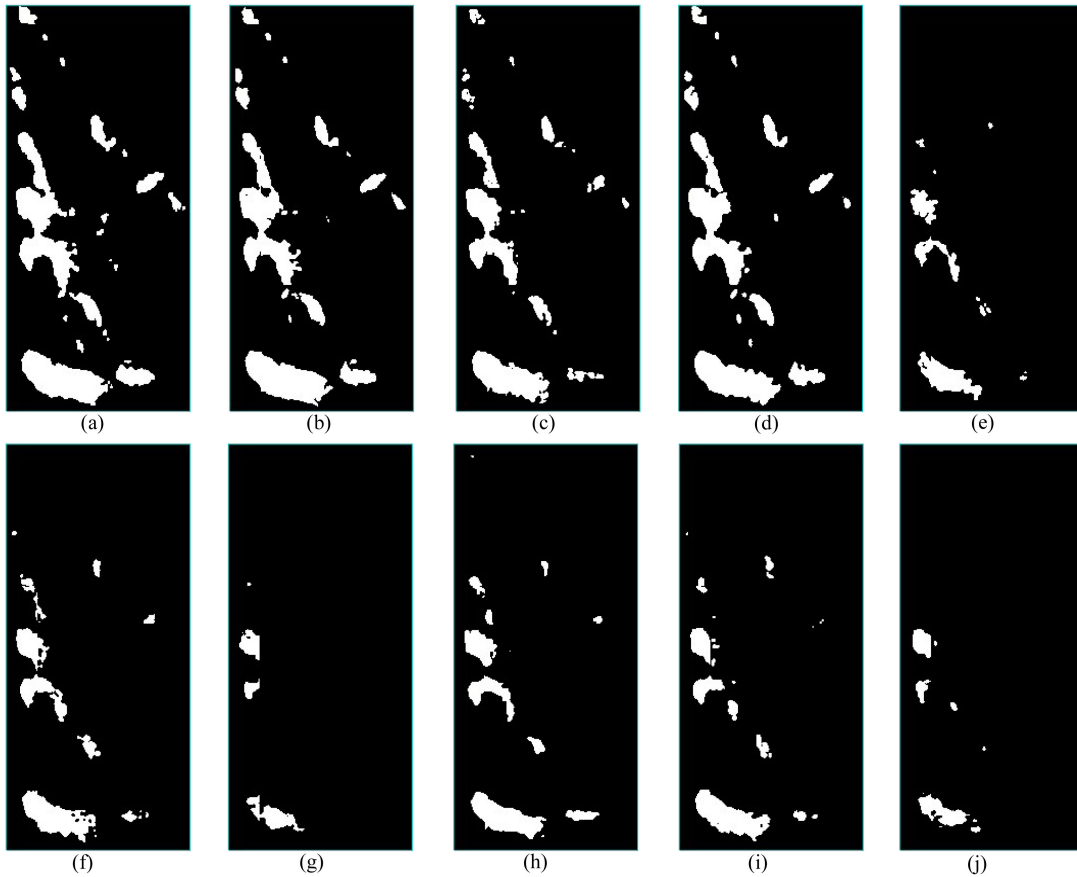| Model	Method	Feature Set	OA(%)	F1-score(%)	MIoU(%)	Kappa
M1		BF_{BDSI}	**99.53**	95.72	92.09	0.91
M2	SegNet	F_{BDSI}	98.15	77.08	68.59	0.54
M3		F_{DB}	98.67	86.24	78.92	0.73
M4		BF_{BDSI}	99.24	93.00	87.81	0.86
M5	Res-SegNet	F_{BDSI}	98.69	86.45	79.03	0.73
M6		F_{DB}	98.63	85.34	77.57	0.71
M7		BF_{BDSI}	99.52	**95.76**	**92.19**	**0.92**
M8	Mobile-SegNet	F_{BDSI}	97.90	72.89	66.01	0.46
M9		F_{DB}	98.05	75.78	68.35	0.52

Fig. 15.   Green tide extraction results of the red box in zone2. (a) Label, (b) BF_{BDSI} + SegNet, (c) BF_{BDSI} + Res-SegNet, (d) BF_{BDSI} + Mobile-SegNet, (e) F_{BDSI} + SegNet, (f) F_{BDSI} + Res-SegNet, (g) F_{BDSI} + Mobile-SegNet, (h) F_{DB} + SegNet, (i) F_{DB} + Res-SegNet, (j) F_{DB} + Mobile-SegNet.

MIoU, and Kappa are 95.76%, 92.19%, and 0.92 respectively, which all come from $M7$ (BF_{BDSI} +Mobile-SegNet). Besides, it can be seen from Table IV that Mobile-SegNet is superior to Seg-Net in terms of time and space. Therefore, from the perspective of comprehensive conditions, the green tide extraction ability of $M1$ is the strongest among nine models. In other words, in the three kinds of networks, Mobile-SegNet is the most suitable network for green tide extraction on BF_{BDSI}.

C. Analysis and Discussion

The goal of lightweight networks lies in the reduction of the model size and the improvement of the model speed while maintaining the model performance. At present, the lightweight network can be divided into two directions, one is to compress the trained complex model to obtain the small model, and the other is to design and train the small model directly. The Mobile-SegNet designed in this article belongs to the latter, which is a low-weight and efficient semantic segmentation network. The major difference among the Mobile-SegNet network, the SegNet, and Res-SegNet network used for comparison is the encoder. Specifically, the encoders of both SegNet and Res-SegNet use the standard convolution for deep feature extraction and use max-pooling for image down-sampling. However, the encoder of Mobile-SegNet uses the depthwise separable convolution for deep feature extraction and realizes down-sampling using the depthwise separable convolution.

Generating thumbnails corresponding to images is the main function of down-sampling. Although the max-pooling down-sampling is able to reduce the estimated mean shift due to the parameter error of the convolution layer to a certain extent, the estimation variance increases due to the limited size of the pooling neighborhood, which may lead to the loss of some information extracted by convolution. When the standard convolution is executed, the receptive field and channel will be considered at the same time. While the depthwise separable convolution considers the receptive field first and then the channel, realizing the separation calculation of the receptive field and channel, which makes the required parameters and computation of the depth separable convolution much less than those of the standard convolution.

However, for the small-scale network model, in the case that the standard convolution is replaced by the depthwise separable convolution, the model size will be significantly reduced and the operation speed will be greatly improved, but the ability of the model may become less ideal [52]. As can be seen from Table V, Figs. 15 and 16, for feature sets F_{BDSI} and F_{DB}, the models obtained by Mobile-SegNet are suboptimal, which may be due to the compromise between the accuracy and speed of Mobile-SegNet network.
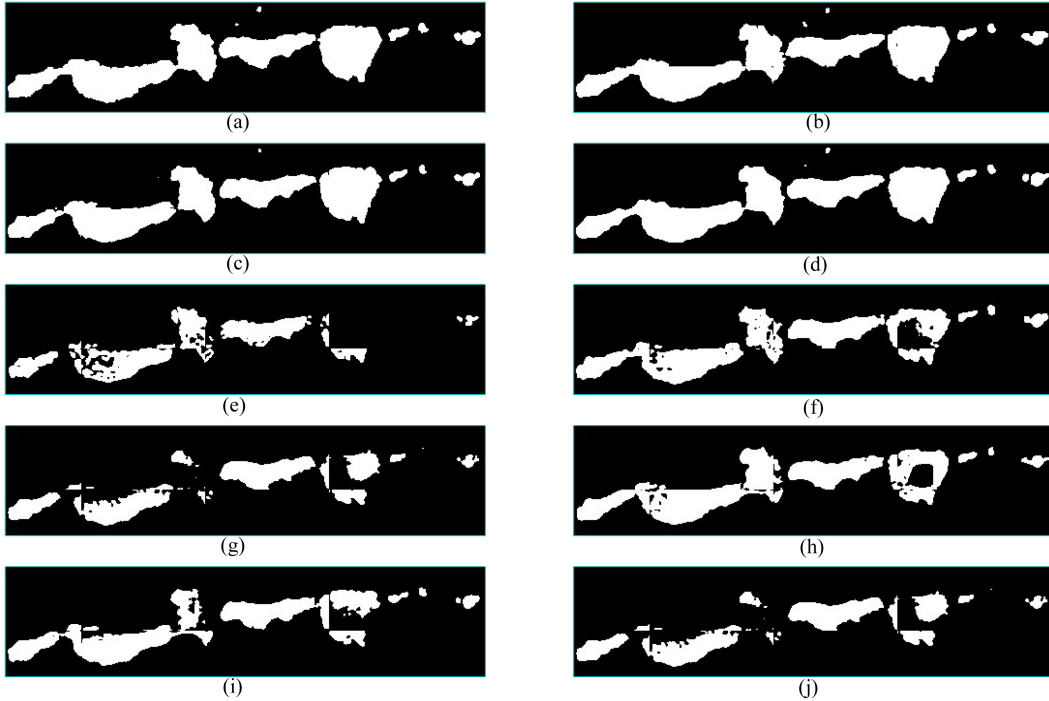
Fig. 16. Green tide extraction results of the red box in zone5. (a) Label, (b) BF_{BDSI} + SegNet, (c) BF_{BDSI} + Res-SegNet, (d) BF_{BDSI} + Mobile-SegNet, (e) F_{BDSI} + SegNet, (f) F_{BDSI} + Res-SegNet, (g) F_{BDSI} + Mobile-SegNet, (h) F_{DB} + SegNet, (i) F_{DB} + Res-SegNet, (j) F_{DB} + Mobile-SegNet.

The network structure depth of Mobile-SegNet, SegNet, and Res-SegNet are 22 layers, 37 layers, and 53 layers, respectively. The residual block can strengthen the transmission of upper and lower information flow and improve the feature reuse rate [46], [50]. Res-SegNet introduced the residual block on the basis of SegNet to expand the network depth, but it may not be available to extract deeper semantic features for the green tide in an irregular shape.

The speckle noise, generated by the defects of the coherent principle on which SAR imaging is based, is inevitable, but it can be suppressed by using the methods, such as wave filtering and wavelet transform [14]. Although we have performed the RLF on SAR data in experiments, there is still some noise that cannot be eliminated. In the areas with dense noise distribution, the brightness of the seawater covered by noise will be improved, thereby reducing the discrimination between green tide and seawater. These noises may be presented in different ways on different feature images, and the band superposition of different features in the feature set will further amplify the negative impact of noise, resulting in a complex background in green tide extraction. In the process of green tide information extraction, the complex backgrounds interfere with the network model perception of the multiscale contextual features (semantic information), which increases the difficulty of network learning.

As can be seen from Fig. 15, for scattered and small green tide targets, the extraction results of feature set F_{BDSI} and F_{DB} on the three kinds of networks all show apparent missing detection and false detection, which makes it unable to accurately express the distribution of green tide. However, the extraction results of BF_{BDSI} on the three networks are good, and there is little

difference from the visual point of view. It can be seen from Fig. 16 that for densely distributed and large green tide targets, the extraction results of feature set F_{BDSI} and F_{DB} on the three kinds of networks are hollowed out and fragmented in varying degrees, and the final extraction results are not completely due to the loss of green tide details. In contrast, BF_{BDSI} is available for a complete interior, clear outline, and fine edge of green tide patches on these three networks. Under the coexistence of complex background and multiscale green tide, the feature set provided by our feature selection method can effectively improve the representation ability of semantic information of green tide and fully reflect the typical characteristics of the green tide.

In the field of green tide monitoring using SAR images, most researchers choose full-polarization SAR data containing more information, whereas dual-polarization SAR has a wider strip width than full-polarization SAR, which can satisfy the requirement of wide ocean area monitoring. However, there are few research works made on dual-polarization SAR at present. In view of this, we will pay more attention to and discuss the temporal-spatial distribution of green tide in coastal areas and the application of dual-polarization SAR in the future study.

V. CONCLUSION

SAR images show high resolution and high timeliness, but in the meantime, they produce massive data and face "data disaster." When sending the high-dimensional feature data of SAR images into the deep learning model for training blindly, it will not only result in a higher learning burden of the model

but also a huge waste of time, space, and computing resources, which makes it difficult to achieve the ideal effect of green tide extraction. In this article, we propose an automatic green tide extraction method for GF-3 SAR images based on feature selection and deep learning, in this manner to provide a technical support for coastal green tide monitoring and promote the application and development of GF-3 satellite. According to the study, sending the optimal features data screened from high-dimensional features into the semantic segmentation network can effectively avoid the impact of redundant features on the model, thus significantly reducing the learning burden and training cost of the model, which improves the efficiency of green tide detection while ensuring high accuracy.

The proposed feature selection method is available for a reduction of the required feature dimension and data amount of GF-3 SAR images in green tide extraction. It can be seen from the results that the selected optimal features available to effectively reflect the real situation of the target ground objects not only outperformed the traditional features, such as backscattering coefficient but also show strong adaptability to different semantic segmentation networks. Mobile-SegNet, a lightweight semantic segmentation network designed in this article, can obtain the rich spatial information and context features from GF-3 SAR images with high noise and weak boundary, which has good extraction ability for both large and small patches of green tide. The experiments show that Mobile-SegNet outperformed SegNet and Res-SegNet in terms of parameter scale, execution efficiency, and segmentation accuracy.

ACKNOWLEDGMENT

REFERENCES

[1] D. A. Lyons *et al.*, "Macroalgal blooms alter community structure and primary productivity in marine ecosystems," *Global Change Biol.*, vol. 20, no. 9, pp. 2712–2724, Dec. 2014.

[2] M. Zhou, D. Liu, D. M. Anderson, and I. Valiela, "Introduction to the special issue on green tides in the Yellow Sea," *Estuar Coastal Shelf Sci.*, vol. 163, no. 2015, pp. 3–8, Jun. 2015.

[3] R. Wang, C. Wang, and J. Li, "An intelligent divisional green tide detection of adaptive threshold for GF-1 image based on data mining," *Acta Oceanol. Sin.*, vol. 41, no. 4, pp. 131–144, Apr. 2019.

[4] X. Jiang, Y. Zou, H. Wang, and H. Zhu, "Application study on quick extraction of entermorpha prolifera information using SAR data," *Acta Oceanol. Sin.*, vol. 31, no. 2, pp. 63–68, Mar. 2009.

[5] P. Chen, J. Yang, and J. Wang, "Marine targets detection using GF-3 SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2017, pp. 1978–1980.

[6] J. Fan, X. Wang, X. Wang, X. Liu, J. Zhao, and Q. Meng, "GF-3 PolSAR marine aquaculture recognition based on complex convolutional neural networks," in *Proc. 10th Int. Conf. Intell. Control Inf. Proces.*, Feb. 2019, pp. 112–115.

[7] W. Jia, W. Mengfei, and D. Jiang, "Detecting recent landslide activities in Yigong and surrounding areas in eastern Tibet of China based on GF-3 SAR amplitude imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Oct. 2020, pp. 5151–5154.

[8] D. Tang, F. Wang, Y. Xiang, H. You, and W. Kang, "Automatic water detection method in flooding area for GF-3 single-polarization data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Nov. 2018, pp. 5266–5269.

[9] R. Yu, G. Wang, T. Shi, W. Zhang, C. Lu, and T. Zhang, "Potential of land cover classification based on GF-1 and GF-3 data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Feb. 2020, pp. 2747–2750.

[10] S. Wei, H. Zhang, C. Wang, F. Wu, and B. Zhang, "Corn fine classification with GF-3 high-resolution SAR data based on deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Nov. 2019, pp. 6397–6400.

[11] Y. Li, G. Liang, S. Yu, and P. Chen, "Selection of microwave remote sensing data for monitoring enteromorpha disaster," *Mar. Environ. Sci.*, vol. 30, no. 5, pp. 133–136, Oct. 2011.

[12] H. Shen, W. Perrie, Q. Liu, and Y. He, "Detection of macroalgae blooms by complex SAR imagery," *Mar. Pollut. Bull.*, vol. 78, no. 2, pp. 190–195, Jan. 2014.

[13] T. W. Cui *et al.*, "Satellite monitoring of massive green macroalgae bloom GMB: Imaging ability comparison of multi-source data and drifting velocity estimation," *Int. J. Remote Sens.*, vol. 33, no. 17, pp. 5513–5527, Sep. 2012.

[14] H. Yu, C. Wang, Y. Sui, J. Li, and J. Chu, "Automatic extraction of green tide using dual polarization Chinese GF-3 SAR images," *J. Coastal Res.*, vol. 102, no. 1, pp. 318–325, Dec. 2020.

[15] Q. Song and C. Wang, "A green tide extraction method based on Markov random field improved algorithm of SAR images," *Hydro. Surv. Charting*, vol. 50, no. 5, pp. 49–52, Sep. 2020.

[16] X. Geng, P. Li, J. Yang, L. Shi, X. Li, and J. Zhao, "Ulva prolifera detection with dual-polarization GF-3 SAR data," in *Proc. IOP Conf. Ser.: Earth Environ. Sci.*, 2020, pp. 1–10, doi:10.1088/1755-1315/502/1/012026.

[17] B. Cui, W. Jing, L. Huang, Z. Li, and Y. Lu, "SANet: A sea–land segmentation network via adaptive multiscale feature learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* vol. 14, pp. 116–126, Nov. 2020, doi: 10.1109/JSTARS.2020.3040176.

[18] Y. Zhang, Z. Pan, C. Cao, J. Duan, and J. Lu, "Information extraction of enteromorpha green tide area based on variational level set method," *Acta Oceanol. Sin.*, vol. 39, no. 9, pp. 121–132, Sep. 2017.

[19] Y. Qin and M. Chi, "RSImageNet: A universal deep semantic segmentation lifecycle for remote sensing images," *IEEE Access*, vol. 8, pp. 68254–68267, Apr. 2020, doi: 10.1109/ACCESS.2020.2986514.

[20] 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.

[21] X. Shi, S. Fu, J. Chen, F. Wang, and F. Xu, "Object-level semantic segmentation on the high-resolution Gaofen-3 FUSAR-map dataset," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3107–3119, Mar. 2021, doi: 10.1109/JSTARS.2021.3063797.

[22] Y. Wang, L. Luo, and Z. Zhou, "Polarized SAR orchard classification based on improved deeplab," *J. Image Graph.*, vol. 24, no. 11, pp. 2035–2044, May 2019.

[23] X. Sun, P. Wang, C. Wang, Y. Liu, and K. Fu, "PBNet: Part-based convolutional neural network for complex composite object detection in remote sensing imagery," *ISPRS J. Photogramm.*, vol. 173, pp. 50–65, Mar. 2021, doi: 10.1016/j.isprsjprs.2020.12.015.

[24] G. Wei, H. Guo, J. An, and H. Liu, "Image segmentation for SAR oil spill on sea surface based on fully convolutional neural network," *J. Comput. Appl.*, vol. 39, no. 1, pp. 182–186, Jul. 2019.

[25] Q. He, X. Sun, Z. Yan, and K. Fu, "DABNet: Deformable contextual and boundary-weighted network for cloud detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: 10.1109/TGRS.2020.3045474.

[26] C. Wang, J. Chu, M. Tan, F. Shao, Y. Sui, and S. Li, "An automatic detection of green tide using multi-windows with their adaptive threshold from landsat TM/ETM plus image," *Acta Oceanol. Sin.*, vol. 36, no. 11, pp. 106–114, Jul. 2017.

[27] Z. Sun, D. Lin, W. Wei, M. Woźniak, and R. Damaševičius, "Road detection based on shearlet for GF-3 synthetic aperture radar images," *IEEE Access*, vol. 8, pp. 28133–28141, Jan. 2020, doi: 10.1109/ACCESS.2020.2966580.

[28] 2015. [Online]. Available: http://envi.geoscene.cn

[29] H. Wakabayashi and K. Yamakawa, "Backscattering characteristics of persistent scatterers extracted from PALSAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 425–428.

[30] 2008. [Online]. Available: http://www.piesat.com.cn/product/pie-sar/

[31] K. Chen, X. Xie, and M. Lin, "An adaptive Gaofen-3 SAR wind field retrieval algorithm based on information entropy," *IEEE Access*, vol. 8, pp. 204494–204508, Nov. 2020, doi: 10.1109/ACCESS.2020.3037023.

[32] S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 498–518, Mar. 1996.

[33] M. Salehi, Y. Maghsoudi, and A. Mohammadzadeh, "Assessment of the potential of H/A/Alpha decomposition for polarimetric interferometric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2440–2451, Apr. 2018.

[34] J. Wang, X. Xing, J. Pan, L. Dong, D. Yang, and Y. Wang, "Application study on the target decomposition method of dual-polarization SAR," *Remot. Sens. Inf.*, vol. 28, no. 2, pp. 106–109, Apr. 2013.

[35] P. Li, X. Xu, H. Dong, and X. Deng, "Polarimetric SAR image feature seletion and multi-layer SVM classification using divisibility index," *J. Comput. Appl.*, vol. 38, no. 1, pp. 138–142, Jan. 2018.

[36] S. Jia, D. Xue, C. Li, J. Zhang, and W. Li, "Study on new method for water area information extraction based on Sentinel-1 data," *Yangtze River*, vol. 50, no. 2, pp. 217–221, Feb. 2019.

[37] E. Choi and C. Lee, "Feature extraction based on the Bhattacharyya distance," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2000, pp. 2146–2148.

[38] M. H. Shah and X. Dang, "Novel feature selection method using Bhattacharyya distance for neural networks based automatic modulation classification," *IEEE Signal Proc. Lett.*, vol. 27, pp. 106–110, Dec. 2020, doi: 10.1109/LSP.2019.2957924.

[39] I. G. Cumming and J. Zyl, "Feature utility in polarimetric radar image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 1989, pp. 1841–1846.

[40] H. Deng, C. Cui, W. Shan, and M. Xu, "Urban building area extraction based on GF-3 satellite SAR images," *Geomatics Word*, vol. 25, no. 6, pp. 79–84, Dec. 2018.

[41] J. Shlens, "A tutorial on principal component analysis," *Int. J. Remote Sens.*, vol. 51, no. 2, pp. 1–12, Apr. 2014.

[42] A. G. Howard *et al.*, "MobileNets: Efficient convolutional neural networks for mobile vision applications," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Apr. 2017, pp. 1–9.

[43] 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, Oct. 2017.

[44] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[45] D. Daimary, M. B. Bora, K. Amitab, and D. Kandar, "Brain tumor segmentation from MRI images using hybrid convolutional neural networks," *Procedia Comput. Sci.*, vol. 167, pp. 2419–2428, Jan. 2020, doi: 10.1016/j.procs.2020.03.295.

[46] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

[47] M. B. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Syst. Des. Implement*, 2016, pp. 265–283.

[48] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[49] S. Wang, X. Mu, D. Yang, H. He, and P. Zhao, "Attention guided encoder-decoder network with multi-scale context aggregation for land cover segmentation," *IEEE Access*, vol. 8, pp. 215299–215309, Nov. 2020, doi: 10.1109/ACCESS.2020.3040862.

[50] Q. Wu, F. Luo, P. Wu, B. Wang, H. Yang, and Y. Wu, "Automatic road extraction from high-resolution remote sensing images using a method based on densely connected spatial feature-enhanced pyramid," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3–17, Dec. 2020, doi: 10.1109/JSTARS.2020.3042816.

[51] S. M. Vieira, U. Kaymak, and J. M. C. Sousa, "Cohen's Kappa coefficient as a performance measure for feature selection," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, Jul. 2010, pp. 1–8.

[52] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 1800–1807.

Changying Wang was born in Tongliao, Nei Monggol Autonomous Region, China, in 1980. She received the B.S. and M.S. degrees in computing mathematics from the University of Inner Mongolia, Nei Monggol Autonomous Region, Hohhot, China, in 2002, and the Ph.D. degree in environmental science and engineering from Ocean University of China, Qingdao, Shandong, in 2009.

From 2010 to 2014, she was a Lecturer with the College of Software Engineering, Qingdao University, Qingdao, China. Since 2015, she has been an Assistant Professor with the College of Computer Science and Technology, Qingdao University. She has authored or coauthored more than 40 articles and 2 inventions. Her research interests include information extraction of satellite remote images, marine science, data mining, and complex networks.

Jinhua Li was born in Changsha, Hu Nan Province, China in 1963. He received the B.S. degree in computer software from Huazhong University of Science and Technology, Wuhan, Hubei, in 1985, the M.S. degree in computer software from China Ship Research and Development Academy, Beijing, China, in 1988, and the Ph.D. degree in computer science from University of Stuttgart, Stuttgart, Germany, in 1998.

From 1985 to 1993, he was a Research Assistant with Wuhan 709 Research Institute of China Shipping Research Institute. From 1993 to 1998, he was a Research Professor with the Computer Science Department, University of Stuttgart. From 1999 to 2003, he was a Research Professor with Henan Xuji Group. Since 2003, he has been a Professor with the College of Computer Science and Technology, Qingdao University, Qingdao, China. He has authored or coauthored more than 50 articles and more than five inventions. His research interests include theory and technology of software engineering, computer science, and software intensive application technology.

Yi Sui was born in Qingdao, Shandong province, China in 1984. She received the B.S. and M.S. degrees in software engineering from the University of Qingdao, Shandong, China, in 2009, and the Ph.D. degree in systems theory from Qingdao University, Qingdao, Shandong, in 2012.

From 2012 to 2019, she was a Lecturer with the College of Computer Science and Technology, Qingdao University. Since 2020, she has been an Assistant Professor with the College of Computer Science and Technology, Qingdao University. She has authored or coauthored more than 20 articles and two inventions. Her research interests include data mining and complex networks.

Haifei Yu was born in Weifang, Shandong, China, in 1996. He received the B.S. degree in software engineering from Harbin University, Harbin, China, in 2019. He is currently working toward the Master's degree in software engineering with Qingdao University, Qingdao, China.

His research interests include deep learning and remote sensing image processing.