







UNet Combined With Attention Mechanism Method for Extracting Flood Submerged Range

Wenmei Li , *Member, IEEE*, Jiaqi Wu , Huaihuai Chen , Yu Wang , *Student Member, IEEE*, Yan Jia ,
and Guan Gui , *Senior Member, IEEE*

Abstract—Synthetic aperture radar (SAR) satellite has been widely applied in real-time flood monitoring as that they are not affected by extreme weather conditions. However, there is no automatic method to quickly and accurately extract flood areas with SAR satellite images. In this article, a UNet combined with the attention mechanism (UNet-CBAM) method has been proposed for extracting flood submerged areas, and both Longgan Lake and Dahuchi in Poyang Lake Basin are selected as the test sites. Based on the polarization characteristics of two Sentinel-1A data of Poyang Lake, both UNet and UNet-CBAM extraction methods are utilized to extract the flood areas, respectively. Compared with traditional SAR image water extraction methods, simulation results demonstrate that UNet can obtain more satisfactory results in recall, precision, and F_1 -parameter, but it has no capability to guarantee continuity in edges and small bodies of water. Moreover, our proposed UNet-CBAM method can further improve recall, precision, and F_1 -parameter, respectively. Specifically, when compared with UNet, its recall is increased by 0.8% and 1.2% while F_1 -parameter is improved by 0.6% and 0.8%, respectively, in the two test sites.

Index Terms—Attention mechanism, flood area extraction, flood monitoring, synthetic aperture radar (SAR), UNet.

I. INTRODUCTION

AS ONE of the most harmful and common natural disasters, floods have high frequency of occurrence, wide coverage, and strong destructiveness. Moreover, the frequency and consequences of extreme flood events worldwide have increased rapidly over the past few decades [1]. Floods hit southern China

nearly every year because of the abundant monsoon rainfall. Increased frequency of extreme precipitation events has aggravated the occurrence of floods resulted in serious economic losses [2]. As the largest freshwater lake in China, Poyang Lake is also one of the most frequently flooded areas in China [3]. Frequent floods in Poyang Lake have caused tremendous damage to the environment and agricultural economy, threatening the lives of approximately 10 million people, who are living in the areas [4]. Since June 2020, Jiangxi Province in China has been continuously affected by extreme precipitation events. In early July, northern and central part of Jiangxi province were hit by heavy rainstorms. Many stations along the entire line of Poyang Lake and the Jiujiang section of the Yangtze River have exceeded the highest water level recorded in history. An unprecedented flood occurred in the Poyang Lake basin, as a result, 36 counties and over 500 000 residents in Jiangxi province were affected by the disaster, causing huge property and economic losses. Therefore, scientific, rapid, and accurate real-time monitoring of the influenced range and spatial distribution of flood disasters will be of great significance to provide relief and reduce the impact of disasters.

Remote sensing technology has become an important means for flood disaster monitoring thanks to its ability to sense the electromagnetic waves reflected or radiated by the target from a long distance to detect and identify the target [5]. Optical remote sensing images are often susceptible to severe weather, so that flood information can only be obtained after the clouds have dispersed. Synthetic aperture radar (SAR) remote sensing can penetrate clouds, fog, and overcome bad sunlight conditions during flood periods, which can achieve entire day and all-weather observations of the disaster-affected areas [6]–[8]. It has become the most important technical measure for large-scale flood monitoring because of its large spatial coverage and short revisit period [9]. SAR images are obtained by back-projection imaging method, the calm water surface often appears as a dark area due to low backscatter coefficient while the backscattering coefficient of nonflood area is higher, which may diffuse scattering dominates [10], [11].

With rapid development of SAR technology, scholars at home and abroad have made some achievements in flood monitoring based on SAR data [12]. Using SAR to study water extraction is one of the current focal fields of research. The methods of flood monitoring based on SAR mainly include: simple visual interpretation [13], image change detection technology [14], supervised classification [15], region growth algorithm [16],

Manuscript received 2 October 2021; revised 21 May 2022 and 22 July 2022; accepted 24 July 2022. Date of publication 27 July 2022; date of current version 19 August 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 42071414, in part by the National Science Foundation of Jiangsu Province under Grant BK20191384, in part by the Open Fund of State Key Laboratory of Remote Sensing Science under Grant OFSLRSS202202, in part by the China Postdoctoral Science Foundation under Grant 2019M661896, and in part by the Environment Protection Research Project of Jiangsu Province in 2019 under Grant 2019010. (*Corresponding author: Guan Gui.*)

Wenmei Li is with the School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing 210023, China, and also with the State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China (e-mail: liwm@njupt.edu.cn).

Jiaqi Wu, Huaihuai Chen, and Yan Jia are with the School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing 210023, China (e-mail: 1326464135@qq.com; 1195038810@qq.com; jiaayan@njupt.edu.cn).

Yu Wang and Guan Gui are with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: 1018010407@njupt.edu.cn; guiguan@njupt.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2022.3194375

object-oriented analysis technology [17], histogram threshold with interferometric SAR coherence coefficient [18], [19], and so on. Pappas et al. [20] present the extraction of rivers based on ICEYE-X2 and Sentinel-1 SAR data with super-pixel segmentation and subsequent classification methods. Lopez-Caloca et al. [21] and Lv et al. [22] separate water bodies from other land covers by using gray level co-occurrence matrix with SAR data. The development of these technologies provides effective means for accurate and automatic flood monitoring. Numerous of scholars used combinations of multiple approaches to complement each other. Shen et al. [23] combined SAR image classification algorithms with morphological algorithms, multithreshold segmentation and machine learning algorithms to obtain the range of water bodies in the study area. Westhoff et al. [24] proposed an automatic flood mapping method, which is based on the backscatter coefficient and the angle of incidence to obtain the probability that each pixel is contained in water body. Giustarini et al. [25] proposed a method jointly application of regional growth and change monitoring based on single-scene SAR image to automatically extract the flood range.

Deep learning (DL) has become a hot spot in the field of machine learning recent years. Ronneberger et al. [26] proposed the UNet model, which gains numerous successful cases in image processing [27], [28]. Many scholars have conducted many studies by improving UNet. Oktay et al. [29] added an attention mechanism to UNet, controlling the importance of different features through gate signals. Yang et al. proposed the RCNN-UNet, tackled propagation errors of road detection [30] and centerline extraction [31]. Benbahria et al. [32] used the transfer learning architecture combining UNet to map irrigated areas. Chu et al. [33] proposed the Res-UNet for sea-land segmentation. The application of DL on SAR images can be mainly divided into target recognition, image denoising, object classification, etc. Chen et al. [34] and Pei et al. [35] achieved the descent accuracy on targets recognition using DL methods based on MSTAR SAR dataset. Lattari et al. [36] proposed the UNet CNN to enhance speckle-filtering capabilities alongside texture preservation of SAR images. Chen et al. proposed feature recalibration network with multiscale spatial features, validating its generality through the classification of large-scale Gaofen-3 and TerraSAR-X images [37].

Although DL is used to classify sea and land in remote sensing images or extract water body information, it is mostly based on optical remote sensing images [38]–[40], and the attention mechanism has not been widely used in the field of remote sensing image semantic segmentation. At the same time, most of the existing flood range monitoring based on SAR images only uses single-scatter SAR images, and often only uses the traditional method for backscattering coefficients to extract the flood range. Based on the Sentinel-1A dual-polarization SAR data from ESA, images before and during the flood disaster in the Poyang Lake region of Jiangxi province during the late June 2020 and early July are selected, and UNet and UNet combined with the attention mechanism (UNet-CBAM) are built to automatically extract the flood areas, after which performance evaluation are conducted in spatial distribution and area statistics

of flood submerged range. The main contribution of this article includes the following three aspects.

- 1) UNet-CBAM is proposed for flood submerged areas extraction with dual-phase SAR images (predisaster images and images collected during the disaster). The attention mechanism can help UNet extract more important features, and the selection of dual-polarization data can more intuitively distinguish the original water area from the submerged area.
- 2) Compared with traditional extraction methods based on backscattering coefficients (Otsu method, Wishart unsupervised method), UNet has improved Recall rate by 5.2% and 5.4% in the two test sites.
- 3) The introduction of the attention mechanism effectively solves the problem of blurred boundaries between water bodies and nonwater bodies and improves the extraction accuracy of flooded areas compared with UNet.

II. TEST SITES AND DATASETS

A. Test Sites

Poyang Lake located in the northern part of Jiangxi province occupies an area of 3150 Km² at low water level, and it is the largest freshwater lake in China, including nine parts, such as Dahuchi, Banghu, Zhonghuchi, Dachahu, and the surrounding lake beaches. The lake plays a critical role in regulating water level of the Yangtze River, conserving water sources, improving the local climate, and maintaining the ecological balance of the surrounding areas. However, large areas of land reclaimed from the lake reduced the capacity of lake, resulting in a decline of the ability to regulate floods, aggravating flood disasters. On July 6, 2020, the water level of Poyang Lake in Xingzi Station reached 19.56 m, 0.56 m above the warning water level. The water level of Poyang Lake in Hukou Station rose to 19.63 m, exceeding the warning water level by 0.13 m, “Poyang Lake No. 1 Flood in 2020” was formed. The flood disaster that began on June 30 has affected 499 000 people and 35 000 hectares of crops in 36 counties (cities, districts), and has caused a direct economic loss of 490 million yuan in Jiangxi. Affected by heavy rainfall, Jiangxi Province urgently relocated 12 000 population. To monitor the spatial distribution of flood submerged ranges and their corresponding areas, Longgan Lake and Dahuchu, which have been severely affected by the flood are selected as test sites. And the test site located in the Poyang Lake Jiujiang to Nanchang basin (28°10′N–30°16′N and 115°20′E–117°05′E) includes nearly 30 counties (as shown in Fig. 1).

B. Datasets and Its Perprocessing

Two Sentinel-1A images are applied to extract flood submerged range and its areas. Sentinel-1 is an earth observation satellite in the Copernicus program of the European Space Agency, consisting of two satellites A and B. Sentinel-1A is equipped with a C-band SAR sensor and provides the following four working modes: Strip map (SM), interferometric wide swath (IW), extra-wide swath (EW), and wave mode (Wave, WV) (see Table I). Research shows that SAR scenes with

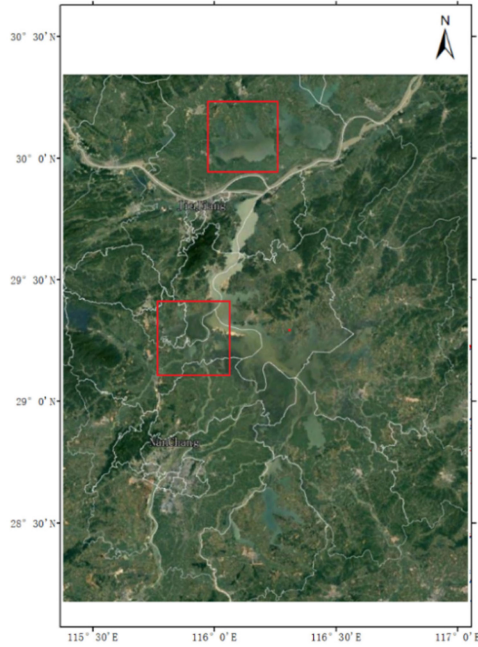


Fig. 1. Location map of the study area.

TABLE I
SENTINEL-1 IMAGING MODE

Mode	Swath width/km	incident angle/ $^{\circ}$	resolution/m	polarization
SM	80	18.3–46.8	5×5	HH+HV/VV+VH/HH/VV
IW	250	29.1–46.0	5×20	HH+HV/VV+VH/HH/VV
EW	410	18.9–47.0	20×40	HH+HV/VV+VH/HH/VV
WV	20×20	23/36	5×5	HH/VV

incident angle ranging from 30° to 50° are recommended for shoreline detection [41]. Two scenes of Sentinel-1A IW imaging mode VV and VH dual-polarization single-look complex products before the flood season on June 26 and July 8 during the flood season are collected as the research data. They are downloaded from ESA, supplemented by satellite precise orbit data of the corresponding date,¹ SRTM1 HGT, SRTM3 version4 DEM, China's county-level administrative division map, etc.

The data preprocessing consists of general orbit correction, registration, and SAR preprocessing. Initially, satellite precise orbit and digital elevation data are utilized to perform orbit correction and registration procedures for predisaster and during-disaster images, respectively. The SAR preprocessing process includes applying the orbit file, radiation calibration, multilook, filtering, terrain correction, image cropping, and so on [42], [43]. Images of the study area were first obtained by the terrain observation by progressive scan (TOPS). The TOPS is a SAR mode for wide swath acquisition, it aims at achieving the same coverage and resolution when the SCAN mode is creating scalloping-free images. The azimuth bandwidth increases with the steering of the antenna main beam and results in azimuth signal aliasing in the Doppler domain. To remove the aliasing, enhanced spectral diversity is applied in TOPS processing [42].

¹[Online]. Available: <https://qc.sentinel1.eo.esa.int/>

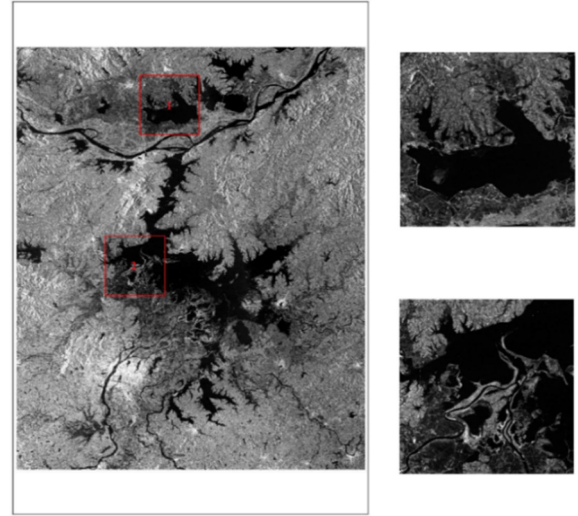


Fig. 2. Backscattering intensity of the study area.

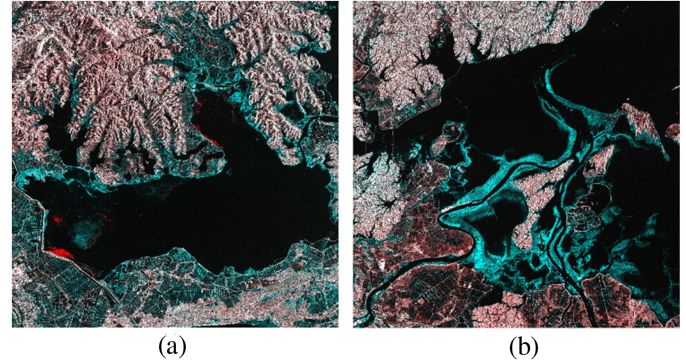


Fig. 3. False color synthesis images of test sites R, G, and B. (a) Longgan Lake. (b) Dahuchi.

The number of oblique-range views for multilook is 4, and the number of azimuth directions is 1. The refined Lee filtering method with window size of 7×7 is applied to suppress or remove the coherent speckle noise and maintain the edge information of the features. The Range–Doppler terrain correction method was used for terrain correction.

After preprocessing, the backscattering intensity image of test site is obtained (as shown in Fig. 2), and Longgan Lake (position 1) and the lake area near Dahuchi (position 2) are selected as the test sites for the subsequent evaluation of the flood monitoring performance. The RGB false color synthesis of the backscattering intensity before and after the disaster of our test site is shown in the Fig. 3, which highlights the water body for visual interpretation, that's mean the blue-green part is flooded areas, and the black part is predisaster water body.

III. METHODOLOGY

According to the imaging principle of SAR, the gray-scale value of the ground object on SAR image is determined by the intensity of backscattering. Radar parameters, object shape, complex permittivity, and surface roughness of the object are

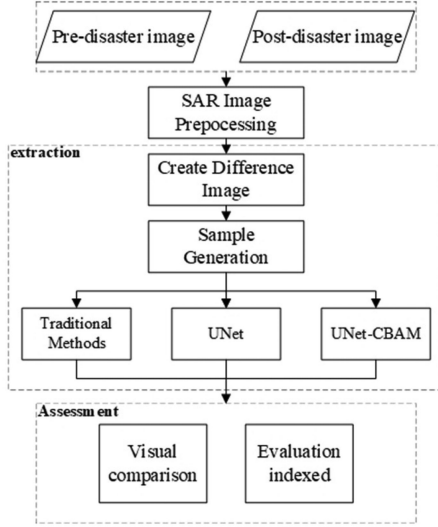


Fig. 4. Flowchart of flood monitoring process.

the four main factors that affect the backscattering intensity of target pixel. The latter three factors basically determine the backscattering characteristics of ground objects when radar parameters remained unchanged. Compared with nonwater bodies with rough surfaces, the surface of water bodies is smooth, dominated by specular reflection, and appears dark in SAR images. In this article, two approaches called UNet and UNet-CBAM are introduced to extract flood affected areas. First, SAR false color synthesis are performed with images in the order of during disaster, predisaster, predisaster corresponding to R: SAR image collected during flood, G: SAR image acquired predisaster, B: SAR image acquired predisaster (R, G, and B). Difference image between dual-phase SAR images (predisaster images and images during the disaster) are established. Then, UNet and Unet-CBAM are applied to segment the false color synthesis image to extract water information. At the same time, the traditional water extraction methods based on the backscattering characteristics is used to obtain the flooded areas, and finally the monitoring accuracy is assessed (the flowchart is shown in Fig. 4).

A. UNet

UNet is a DL approach introduced by Ronneberger et al. [26]. It is mainly composed of the following two parts: down-sampling block and up-sampling block. The up-sampling part is composed of a structure symmetrical to the encoding part. The image size is increased by deconvolution before the different decoding parts are operated, and finally the high-dimensional features with the same size as original image are obtained.

In addition to the deep abstract features obtained from the up-sampling of the previous layer, the input of each group of convolutional layers also gets the shallow local features obtained from the corresponding down-sampling layer. The deep and shallow features are merged to restore the details of the feature map and ensure that the spatial information dimension remains unchanged.

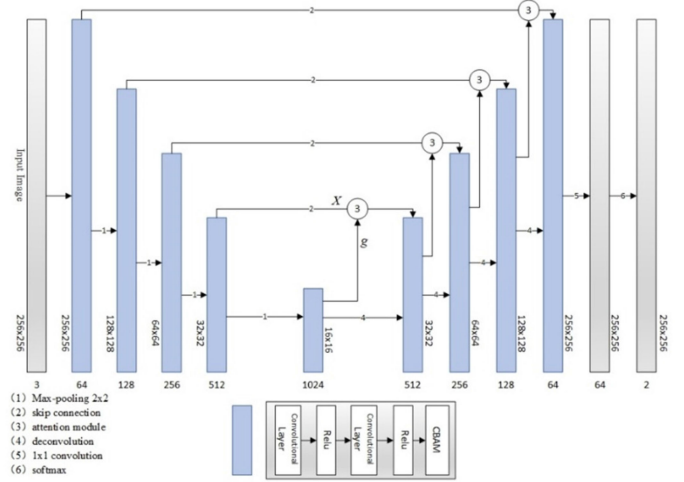


Fig. 5. UNet-CBAM structure diagram.

B. Attention Mechanism

1) *UNet-CBAM*: Convolutional block attention module (CBAM) [44] is a lightweight attention module that designed for convolution layer from channel and spatial dimensions, respectively. It can be seamlessly integrated into any convolutional neural network architecture for end-to-end training. Modified convolution unit can be added to the convolution unit to adjust weights in order to suppress unimportant features. Introducing CBAM to the end of each layer constitutes the structure of CBAM-UNet (see Fig. 5). The input image is a false-color synthesis three-band remote sensing image, the training image size is 256×256 , and the output image is an annotated image. Both the up-sampling and down-sampling parts are composed of two 3×3 convolutional layers and two ReLU. The output result of each down-sampling part is transmitted to the next one after passing through the max-pooling layer of 2×2 steps. The high-dimensional features are then subjected to a 1×1 convolution operation to obtain the dimensional features corresponding to the number of categories, and finally the annotation map is output through Softmax operation.

Softmax is a generalization of the logic function, which converts all results into probabilities between 0 and 1. The definition of the softmax function is described in the following equation:

$$\text{Softmax}(v_i) = \frac{e^{v_i}}{\sum_1^k e^{v_k}} \quad (1)$$

where v_i is the data value at a certain pixel in the output image of the i th type; v_k is the data value of the k th type at the same position; k is the total number of classes. The softmax value is the ratio of the pixel index to the sum of all pixel indexes, through which multiclass output value can be converted into a probability distribution in the range of $[0, 1]$. The cross-entropy loss function is defined as

$$L = - \sum_{c=1}^C y_c \log(p_c) \quad (2)$$

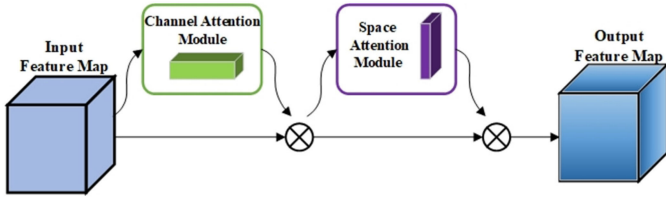


Fig. 6. CBAM structure diagram.

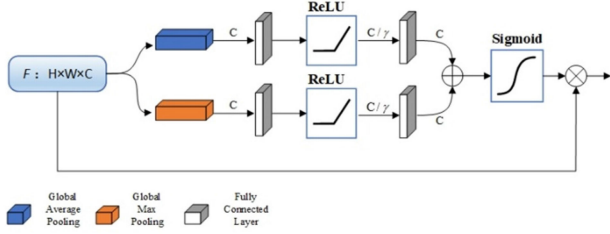


Fig. 7. Channel attention module.

where $p = [p_0, \dots, p_{c-1}]$ is a vector, and p_c represents the probability of the sample, which is predicted to be the c th category.

And the structure of CBAM is shown in Fig. 6. It consists of a channel attention module and a spatial attention module. The attention weight is calculated in the two dimensions of space and channel in turn, and then multiplied by the original feature map to adaptively adjust the features. This structure guarantees the combined usage of channel and space features.

2) *Channel Attention Module*: First, the feature map F of size $H \times W \times C$ obtained by the encoding part of feature extraction network are input into the channel attention module, and global average pooling and global max pooling operations are performed, respectively, then two $1 \times 1 \times C$ are generated as output F_{avg}^C , F_{max}^C . Second, the output results of the first step are passed through a two-layer shared fully connected network, respectively. The number of neurons in the first layer is C/γ , γ is the attenuation rate to control the intensity of attention, which is set to 8 in this article, and ReLU is used as activation function. The number of neurons in the second layer is C . The $1 \times 1 \times C/\gamma$ vector is passed through the neuron and the output feature length is C , which shows the importance of the input channels' number. Third, the two obtained features are added based on element-wise, and activated by the Sigmoid function to generate the final channel weight value $M_C(F)$. The formula is illustrated in (3), where w_0 and w_1 represent the corresponding weights of the first and second fully connected layers separately, σ represents the Sigmoid activation function

$$\begin{aligned} M_C(F) &= \sigma(\text{MLP}(F_{\text{avg}}^C) + \text{MLP}(F_{\text{max}}^C)) \\ &= \sigma(w_1(w_0(F_{\text{avg}}^C)) + w_1(w_0(F_{\text{max}}^C))). \end{aligned} \quad (3)$$

Finally, the channel weight value and the original input feature map are subjected to element-wise multiplication operation by channel dimension, and the output feature map F' of the channel attention module is obtained. The structure of the channel

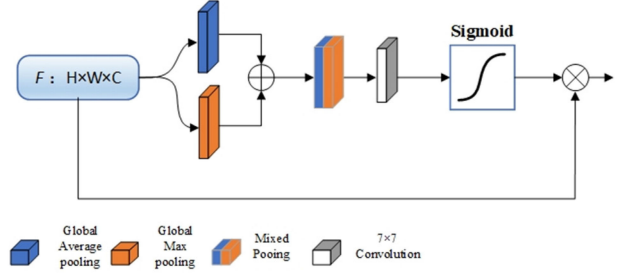


Fig. 8. Spatial attention module.

attention module is shown in Fig. 7, and its formula is shown as:

$$F' = M_C(F) \otimes F. \quad (4)$$

The channel attention module exploits both average-pooled and max-pooled features simultaneously to help obtain the noteworthy part of the input image.

3) *Spatial Attention Module*: The spatial attention module is shown in Fig. 8. As shown in Fig. 8 that the feature F' of size $H \times W \times C$ output by the channel attention module is input and performed by the global max pooling and global average pooling of the channel dimensions, respectively. This process is to obtain the channel description F_{avg}^S , F_{max}^S of size $H \times W \times 1$. After fusion of the abovementioned two, the attention weight $M_S(F)$ is obtained through a 7×7 convolutional layer with an activation function of Sigmoid. The formula is as follows in (5), where σ represents the Sigmoid activation function, and $f^{7 \times 7}$ represents the 7×7 convolution operation

$$\begin{aligned} M_S(F) &= \sigma(f^{7 \times 7}([\text{Avgpool}(F); \text{Maxpool}(F)])) \\ &= \sigma(f^{7 \times 7}([F_{\text{avg}}^S, F_{\text{max}}^S])). \end{aligned} \quad (5)$$

Finally, the channel weight and the input feature map F' are subjected to element-wise multiplication operation by space dimension to obtain the output of spatial attention module called feature map F'' , and its formula is in the following equation:

$$F'' = M_S(F) \otimes F'. \quad (6)$$

Spatial attention module generates a spatial attention map by using the spatial relationship of features, which is complementary to channel attention.

IV. EXPERIMENT

A. Data Preparation

Samples are produced using the LabelMe labeling tool based on Python 3.6. First, the selected Sentinel-1A SAR image is cropped to a size of 1024×1024 , leaving an overlapping area of 200 pixels between adjacent images. Second, three semantic categories of predisaster water bodies, postdisaster submerged area, and nonwater bodies for labeling are set up, and the labeled map is obtained (see Fig. 9). After labeling, the training and the validation samples are allocated and divided into 256×256 small images. To prevent the model from being unable to

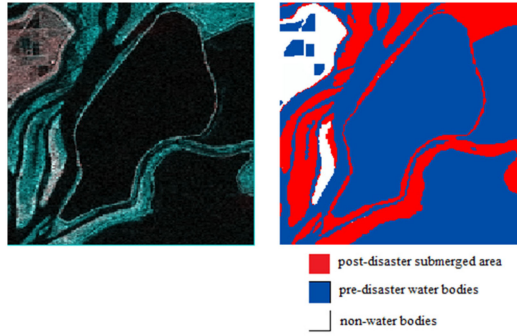


Fig. 9. Sample labeling examples.

converge and underfitting due to too few images in the training set, the cropped sample images are rotated by 90° , 180° , 270° , horizontal flip and vertical flip. After the data enhancement processing, the final sample number is 1572, of which the number of training samples is 510, and 1062 is used as the verification set.

B. Model Training and Accuracy Assessment

The model training process used cross entropy as the loss function, Adam as the optimization algorithm, the initial learning rate is 0.001, and the number of epochs is set to 100. The training curve of UNet-CBAM is shown as Fig. 10, and the network finally converged.

Traditional methods, such as Otsu and Wishart unsupervised method based on backscattering coefficients are also used as comparison with the methods we introduced. The experimental accuracy evaluation uses evaluation indicators commonly used in semantic segmentation: recall, precision, and F_1 measure value

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (7)$$

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (8)$$

$$F_1 = \frac{2PR}{P + R} \quad (9)$$

where N_{TP} represents the number of changed pixels predicted to be changed; N_{FN} represents the set of unchanged pixels predicted to be changed; N_{FP} represents the set of unchanged pixels predicted to be unchanged. R represents recall; P represents precision, F_1 represents F_1 measure value. The precision and the recall rate represent the correct frequency of samples predicted to be positive. The precision is based on the prediction result while recall rate is based on the original samples. F_1 is the harmonic average of recall and precision. Higher the F_1 is, more ideal the experimental method is.

V. RESULTS

Since the input size of the model is fixed at 256×256 , it is necessary to define a test picture generator to convert the image into a size that can be input into the model, and then the sliding

TABLE II
FLOOD MONITORING ACCURACY OF LONGGAN LAKE TEST SITE

Method	Recall(%)	Precision(%)	F_1
Otsu	86.4	81.6	83.9
Wishart	86.2	82.2	84.2
UNet	91.6	83.1	87.1
UNet-CBAM	92.4 (0.8% \uparrow)	83.4 (0.3% \uparrow)	87.7 (0.6% \uparrow)

TABLE III
FLOOD MONITORING ACCURACY OF DAHUCHI LAKE TEST SITE

Method	Recall(%)	Precision(%)	F_1
Otsu	86.5	84.5	85.5
Wishart	86.0	84.6	85.3
UNet	89.4	85.0	87.2
UNet-CBAM	90.6 (1.2% \uparrow)	85.6 (0.6% \uparrow)	88.0 (0.8% \uparrow)

window is applied to block the test site image into the test picture generator. Next, an empty matrix of the image size covered the study area is established, and the results of each small piece obtained are spliced into this matrix, and finally a classification map of the entire study area is obtained. Figs. 11 and 12 are the prediction map of Longgan Lake and Dahuchi obtained by using UNet and UNet-CBAM, respectively. In visual interpretation, UNet-CBAM distinguishes more small water bodies in the land part, and the continuity of water body edges is better maintained compared with that of UNet.

In quantitatively aspect, the accuracy of flood monitoring results of Longgan Lake with several methods is evaluated by recall, precision, and F_1 measure value (as shown in Table II). The banks of the Longgan Lake test site are dominated by terraced farmland and river beaches. Table II shows that the UNet precision is 83.1%, which is 1.5% and 0.9% higher than that of the Otsu algorithm and Wishart unsupervised classification, respectively. The recall of UNet method is 91.6%, which presents an increase of 5.2% and 5.4%, compared to 86.4% of the Otsu algorithm and 86.2% of the Wishart unsupervised classification. The F_1 parameter of UNet is 87.1%, which is 3.2% and 2.9% higher than that of the Otsu algorithm and Wishart unsupervised classification method, respectively. The recall of UNet-CBAM after introducing the attention mechanism module is increased by 0.8%, the precision is increased by 0.3%, and the F_1 parameter is increased by 0.6% compared with those of UNet.

Dahuchi test site has high requirements for continuity in the extraction of the riverbank. The results of the recall, precision, and F_1 in the Dahuchi test site are shown in Table III. The accuracy rates of the Otsu algorithm, Wishart unsupervised classification, and UNet are 84.5%, 84.6%, 85.0%, the recall rates are 86.5%, 86.0%, 89.4%, and F_1 parameter are 85.5%, 85.3%, 87.2%. It can be seen that UNet increases precision, recall, and F_1 parameter by up to 0.5%, 2.9%, and 1.9%, respectively, using deep features to extract water information through sample iteration, it obtains relatively more reliable results. After the introduction of the attention mechanism module, the UNet-CBAM method further improves the recall rate by 1.2%, while the precision rate and F_1 parameters are increased by 0.6% and 0.8%, respectively.

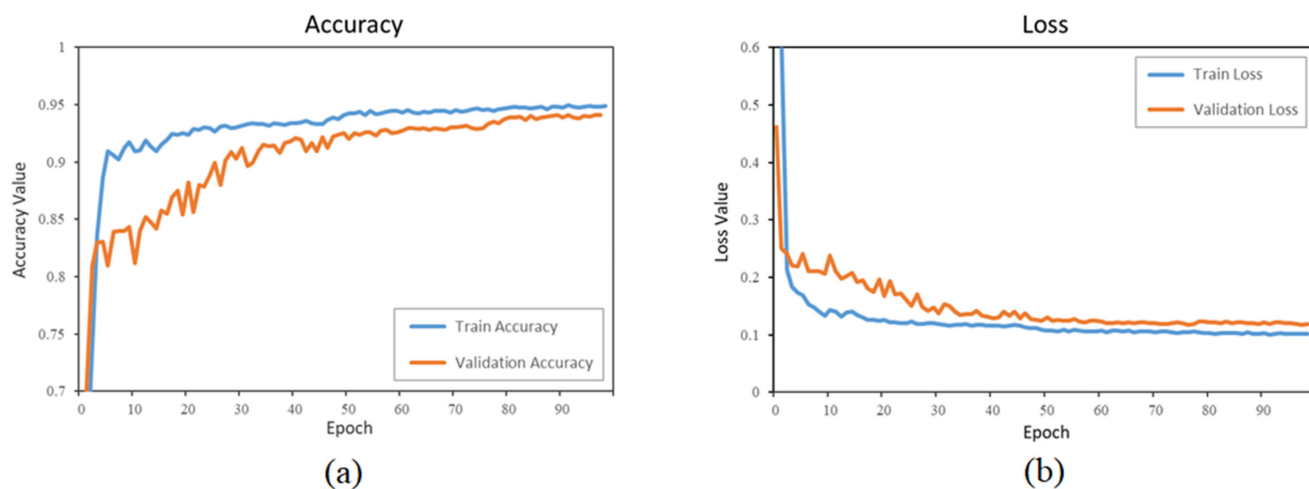


Fig. 10. Training curve of UNet-CBAM. (a) Accuracy. (b) Loss.

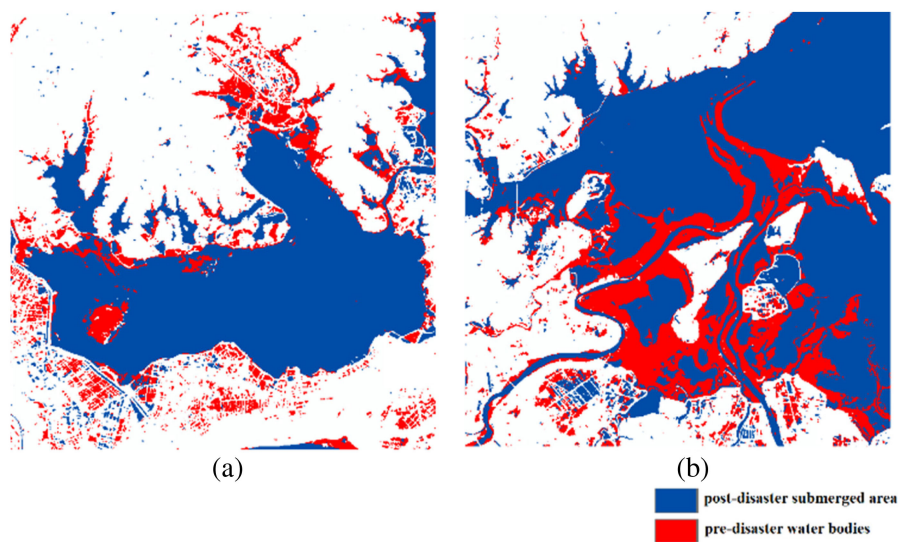


Fig. 11. UNet prediction map of test site. (a) Longgan Lake prediction result. (b) Dahuchi prediction result.

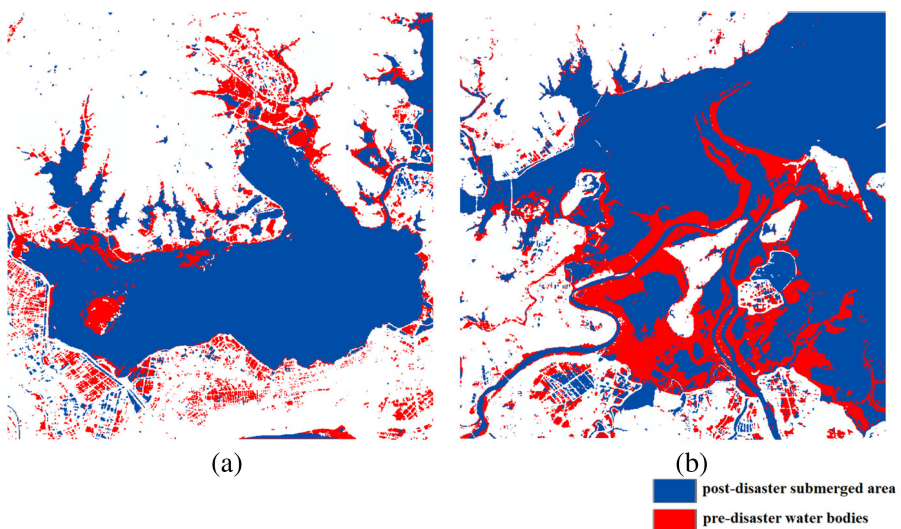


Fig. 12. UNet-CBAM prediction map of test site. (a) Longgan Lake prediction result. (b) Dahuchi prediction result.

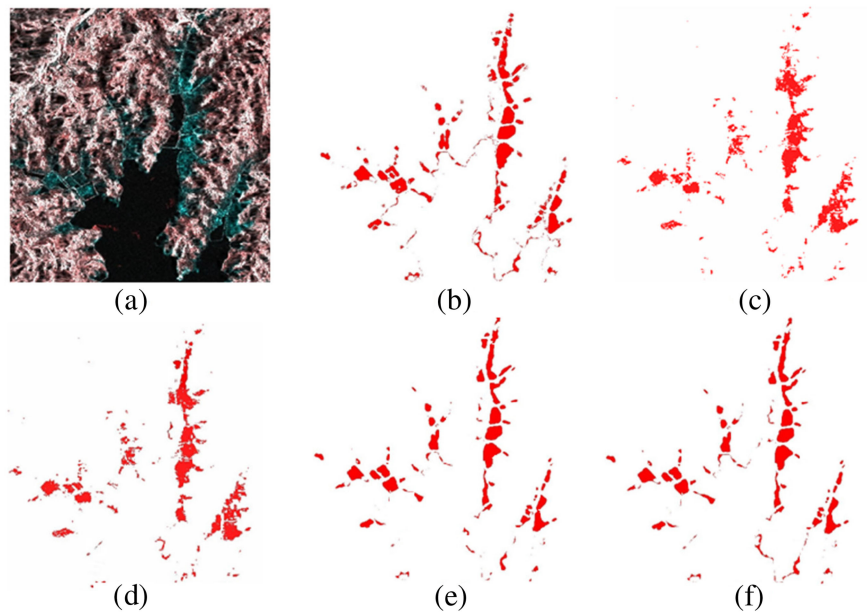


Fig. 13. Partial flood interested areas, (a) ground truth, (b) and extraction results with (c) Otsu, (d) Wishart unsupervised method, (e) UNet and (f) UNet-CBAM in Longgan Lake test site.

According to the extraction results of partial flooding in the Longgan Lake test site (see Fig. 13), the water bodies extracted by traditional water extraction methods Otsu method (c) and Wishart unsupervised method (d) have more jagged edges. Compared with the traditional method, the extraction results of water edges and small water bodies predicted by the UNet method (e) is closer to the ground truth (b), but the continuity of the extraction results at the junction of the water bodies and the nonwater bodies is not well. The attention mechanism enables the network to increase the calculation weight of the changed areas and has a certain inhibitory effect on noise caused by complex backgrounds. Therefore, the results of UNet-CBAM (f) are less likely to be broken.

VI. CONCLUSION

In this article, a DL method UNet combined with attention mechanism called UNet-CBAM is proposed to extract flood submerged range, and its performance is evaluated by several indicators. Two Sentinel-1A images of the first flood of Poyang Lake that occurred from the end of June to the beginning of July 2020 are applied to monitor the flooded areas and the corresponding ground truth is utilized to test the accuracy and performance of our proposed method. Meanwhile, traditional methods Otsu and Wishart unsupervised method are also applied to compare with the performance of UNet and UNet-CBAM. Although UNet provides promising result for flood mapping on the overall scale, it fails to achieve sufficient accuracy when it comes to small water bodies and edges. UNet-CBAM combines UNet with attention mechanism could guarantee the continuity of edges and detection capability of small water bodies. The results show that UNet achieves better results in flood submerged range extraction than traditional SAR water extraction methods

based on backscattering coefficients. By adding CBAM, the extraction result of UNet-CBAM has been significantly improved in recall, precision, and F_1 parameter, meanwhile, the problem of discontinuity of UNet flood extraction results has also been solved.

Nonetheless, there are several drawbacks needed to be resolved further. First, the generalization of UNet-CBAM, in our experiment, the samples are produced based on the two test sites. If the experiment can be carried out on larger scale areas, and compared with some basic DL methods, the performance of UNet-CBAM would be more objective. Second, although the samples are enhanced by a few methods, the number of samples is still not enough for UNet. Few-shot learning methods combined with UNet-CBAM may reduce the reliance of DL on samples' amount and improve the accuracy.

REFERENCES

- [1] L. M. Bouwer et al., "Confronting disaster losses," *Science*, vol. 318, no. 5851, pp. 753–753, 2007.
- [2] C. Nie et al., "Spatial and temporal changes in flooding and the affecting factors in China," *Natural Hazards*, vol. 61, no. 2, pp. 425–439, 2012.
- [3] S. Qi et al., "Inundation extent and flood frequency mapping using LANDSAT imagery and digital elevation models," *Mapping Sci. Remote Sens.*, vol. 46, no. 1, pp. 101–127, 2009.
- [4] X. Li and Q. Zhang, "Variation of floods characteristics and their responses to climate and human activities in Poyang Lake, China," *Chin. Geographical Sci.*, vol. 25, no. 1, pp. 13–25, 2015.
- [5] J. F. Pekel et al., "High-resolution mapping of global surface water and its long-term changes," *Nature*, vol. 540, no. 7633, pp. 418–422, 2016.
- [6] D. C. Mason et al., "Detection of flooded urban areas in high resolution synthetic aperture radar images using double scattering," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 28, pp. 150–159, 2014.
- [7] M. Jiang, X. Zhao, and X. Shi, "Kinematic behavior analysis of the wadi landslide from time-series sentinel-1 data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 127–135, 2022.

- [8] R. Xiao, M. Jiang, Z. Li, and X. He, "New insights into the 2020 Sardoba dam failure in Uzbekistan from Earth observation," *Int. J. Appl. Earth Observation Geoinformation*, vol. 107, Mar. 2022, Art. no. 102705, doi: 10.1016/j.jag.2022.102705.
- [9] F. Cian, M. Marconcini, and P. Ceccato, "Normalized difference flood index for rapid flood mapping: Taking advantage of EO big data," *Remote Sens. Environ.*, vol. 209, pp. 712–730, 2018.
- [10] S. F. Sherpa, M. Shirzaei, C. Ojha, S. Werth, and R. Hostache, "Probabilistic mapping of august 2018 flood of Kerala, India, using space-borne synthetic aperture radar," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 896–913, 2020.
- [11] J. Cohen et al., "Implications of boreal forest stand characteristics for X-band SAR flood mapping accuracy," *Remote Sens. Environ.*, vol. 186, pp. 47–63, 2016.
- [12] R. Hostache et al., "Near-real-time assimilation of SAR-derived flood maps for improving flood forecasts," *Water Resour. Res.*, vol. 54, no. 8, pp. 5516–5535, 2018.
- [13] J. Sanyal and X. Lu, "Application of remote sensing in flood management with special reference to monsoon asia: A review," *Natural Hazards: J. Int. Soc. Prevention Mitigation Natural Hazards*, vol. 33, no. 2, pp. 283–301, 2004.
- [14] M. A. Clement, C. G. Kilsby, and P. Moore, "Multi-temporal synthetic aperture radar flood mapping using change detection," *J. Flood Risk Manage.*, vol. 11, no. 2, pp. 152–168, 2018.
- [15] L. Pulvirenti et al., "Monitoring flood evolution in agricultural areas using COSMO-SkyMed data: Analysis of the Tuscany inundation of December 2009," *Proc. SPIE*, vol. 8179, pp. 142–153, Oct. 2011.
- [16] D. C. Mason et al., "Automatic near Real-time selection of flood water levels from high resolution synthetic aperture radar images for assimilation into hydraulic models: A case study," *Remote Sens. Environ.*, vol. 124, no. 9, pp. 705–716, 2012.
- [17] V. Herrera-Cruz and F. Koudogbo, "TerraSAR-X rapid mapping for flood events," in *Proc. ISPRS (Earth Imag. Geospatial Inf)*, Hannover, Germany, 2009, pp. 170–175.
- [18] L. Pulvirenti et al., "Use of SAR data for detecting floodwater in urban and agricultural areas: The role of the interferometric coherence," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1532–1544, Mar. 2016.
- [19] M. Jiang and A. M. Guarnieri, "Distributed scatterer interferometry with the refinement of spatiotemporal coherence," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 3977–3987, Jun. 2020.
- [20] O. A. Pappas et al., "River planform extraction from high-resolution SAR images via generalized gamma distribution superpixel classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 3942–3955, May 2021.
- [21] A. A. López-Caloca et al., "Analyzing short term spatial and temporal dynamics of water presence at a basin-scale in Mexico using SAR data," *GIScience Remote Sens.*, vol. 57, no. 7, pp. 985–1004, 2020.
- [22] W. Lv, Q. Yu, and W. Yu, "Water extraction in SAR images using GLCM and support vector machine," in *Proc. IEEE 10th Int. Conf. Signal Process.*, Beijing, China, 2010, pp. 740–743.
- [23] X. Shen et al., "Near-real-time non-obstructed flood inundation mapping using synthetic aperture radar," *Remote Sens. Environ.*, vol. 221, pp. 302–315, 2019.
- [24] R. S. Westerhoff et al., "Automated global water mapping based on Wide-swath orbital synthetic aperture radar," *Hydrol. Earth Syst. Sci. Discuss.*, vol. 17, no. 2, pp. 651–663, 2013.
- [25] L. Giustarini et al., "Probabilistic flood mapping using synthetic aperture radar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 6958–6969, Dec. 2016.
- [26] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput. Assist. Interv. Soc.*, Cham, Germany, 2015, pp. 234–241.
- [27] D. Hordiiuk et al., "Semantic segmentation for ships detection from satellite imagery," in *Proc. IEEE 39th Int. Conf. Electron. Nanotechnol.*, 2019, pp. 454–457.
- [28] X. Zhao et al., "Use of unmanned aerial vehicle imagery and deep learning UNet to extract rice lodging," *Sensors*, vol. 19, no. 18, 2019, Art. no. 3859.
- [29] H. Hwang, H. Z. U. Rehman, and S. Lee, "3D U-Net for skull stripping in brain MRI," *Appl. Sci.*, vol. 9, no. 3, 2019, Art. no. 569.
- [30] O. Oktay et al., "Attention U-Net: Learning where to look for the pancreas," 2018. [Online]. Available: <https://arxiv.org/abs/1804.03999>
- [31] X. Yang et al., "Road detection and centerline extraction via deep recurrent convolutional neural network U-Net," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7209–7220, Sep. 2019.
- [32] Z. Benbahria et al., "Land cover intelligent mapping using transfer learning and semantic segmentation," in *Proc. 7th Mediterranean Congr. Telecommun.*, 2019, pp. 1–5.
- [33] Z. Chu et al., "Sea-land segmentation with Res-UNet and fully connected CRF," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 3840–3843.
- [34] S. Chen and H. Wang, "SAR target recognition based on deep learning," in *Proc. Int. Conf. Data Sci. Adv. Anal.*, 2014, pp. 541–547.
- [35] J. Pei, Y. Huang, W. Huo, Y. Zhang, J. Yang, and T.-S. Yeo, "SAR automatic target recognition based on multiview deep learning framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2196–2210, Apr. 2018.
- [36] F. Lattari et al., "Deep learning for SAR image despeckling," *Remote Sens.*, vol. 11, no. 13, 2019, Art. no. 1532.
- [37] L. Chen et al., "A new deep learning algorithm for SAR scene classification based on spatial statistical modeling and features re-calibration," *Sensors*, vol. 19, no. 11, 2019, Art. no. 2479.
- [38] L. Yu et al., "Convolutional neural networks for water body extraction from Landsat imagery," *Int. J. Comput. Intell. Appl.*, vol. 16, no. 1, 2017, Art. no. 1750001.
- [39] W. Jiang et al., "Multilayer perceptron neural network for surface water extraction in Landsat 8 OLI satellite images," *Remote Sens.*, vol. 10, no. 5, 2018, Art. no. 755.
- [40] Z. Miao, K. Fu, H. Sun, X. Sun, and M. Yan, "Automatic water-body segmentation from high-resolution satellite images via deep networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 602–606, Apr. 2018.
- [41] L. Wu et al., "Study on characteristics of SAR imagery around the coast for shoreline detection," *Coastal Eng. J.*, vol. 61, no. 2, pp. 152–170, 2019.
- [42] M. Jiang, "Sentinel-1 TOPS co-registration over low coherence area and its application to velocity estimation using the all pairs shortest path algorithm," *J. Geodesy*, vol. 94, no. 95, pp. 1–15, 2020.
- [43] F. Filippini, "Sentinel-1 GRD preprocessing workflow," in *Proc. Multi-disciplinary Digit. Publishing Inst.*, vol. 18, no. 1, 2019, Art. no. 11.
- [44] S. Woo et al., "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 3–19.



Wenmei Li (Member, IEEE) received the M.S. degree from Nanjing University, Nanjing, China, in 2010, and the Ph.D. degree from the Chinese Academy of Forestry, Beijing, China, in 2013. She is currently working toward the postdoctoral degree with the Nanjing University of Posts and Telecommunications, Nanjing. She is currently and Associate Professor with the School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications. Her research interests include deep learning, optimization, image reconstruct, and their application in land remote sensing.



Jiaqi Wu received the B.S. degree in geographic information science in 2019 from the Nanjing University of Posts and Telecommunications, Nanjing, China, where he is currently working toward the M.A. degree. His research interests include spatial-temporal fusion, time series, deep learning, and their applications in land remote sensing.



Huaihuai Chen received the B.S. degree in geographic information science in 2020 from the Nanjing University of Posts and Telecommunications, Nanjing, China, where she is currently working toward the M.A. degree. Her research interests include information extraction, image fusion, change monitoring, and their application in remote sensing.



Yu Wang (Student Member, IEEE) received the B.S. degree in communication engineering in 2018 from the Nanjing University of Posts and Telecommunications, Nanjing, China, where he is currently working toward the Ph.D. degree. He has authored or coauthored more than 15 papers in peer-reviewed IEEE journal/conferences. His research interests include deep learning, optimization, and its application in wireless communications. Mr. was the recipient of two best paper awards (i.e., CSPS 2018 and ICEICT 2019).



remote sensing, and antenna design.

Yan Jia received the double M.S. degree in telecommunications engineering and computer application technology from Politecnico di Torino, Turin, Italy, and Henan Polytechnic University, Jiaozuo, China, in 2013, and the Ph.D. degree in electronics engineering from Politecnico di Torino in 2017. She is currently with the Nanjing University of Posts and Telecommunications, Nanjing, China. Her research interests include microwave remote sensing, soil moisture retrieval, Global Navigation Satellite System Reflectometry (GNSS-R) applications to land



Guan Gui (Senior Member, IEEE) received the Dr. Eng degree in information and communication engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2012. From 2009 to 2014, he was with the Wireless Signal Processing and Network Laboratory (Prof. Adachi laboratory), Department of Communications Engineering, Graduate School of Engineering, Tohoku University, as a Research Assistant and Postdoctoral Research Fellow. From 2014 to 2015, he was an Assistant Professor with the Department of Electronics and Information System, Akita Prefectural University. Since 2015, he has been a Professor with the Nanjing University of Posts and Telecommunications, Nanjing, China. He has authored or coauthored more than 200 international peer-reviewed journal/conference papers. His research interests include deep learning, compressive sensing, and advanced wireless techniques. Dr. Gui was the recipient of the best paper awards, e.g., ICNC 2018, ICC 2017, ICC 2014 and VTC 2014-Spring. He was also the recipient of the Member and Global Activities Contributions Award in 2018, the Top Editor Award of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in 2020. He was selected as Jiangsu Specially-Appointed Professor in 2016, Jiangsu High-level Innovation and Entrepreneurial Talent in 2016, Jiangsu Six Top Talent in 2018, Nanjing Youth Award in 2018. He was an Editor of *Security and Communication Networks* (2012–2016). He has been the Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY since 2017, IEEE ACCESS since 2018, *Physical Communication* since 2019, *KSII Transactions on Internet and Information Systems* since 2017, *Journal of Communications* since 2019, and the Editor-in-Chief of *EAI Transactions on Artificial Intelligence* since 2018.