DEEP LEARNING APPROACH FOR FLOOD DETECTION USING SAR IMAGE: A CASE STUDY IN XINXIANG

Bofei Zhao 1, Haigang Sui 1*, Chuan Xu 2, Junyi Liu 1

¹ State Key Laboratory of Information Engineering in Surveyinsg, Mapping and Remote Sensing, Wuhan University, China -

bowie.zhao@whu.edu.cn

² School of Computer Science, Hubei University of Technology, China - xc992002@foxmail.com

Commission III, ICWG III/IVa

KEY WORDS: SAR, Deep Learning, Flood Detection, Attention Module, Water Extraction.

ABSTRACT:

With the gradual warming of the global climate, frequent floods have caused huge losses to human life and property. Flood mapping by SAR image has been an important topic, and it is increasingly important to use deep learning method to extract flood information. In order to achieve automatic flood extent extraction, this paper proposes an attention mechanism-based water body extraction network with GF3 images, and successfully use it for flood detection in Xinxiang, Henan, China. In this paper, the proposed network incorporates the channel attention mechanism and position attention mechanism based on U-Net, to improve the efficiency and accuracy of water extraction, ignore the unimportant information by learning the weight and make the model focus toward the important information related to the water body. The OA of our method can reach 0.959 and Recall can reach 0.942 by verifying four sets of test data. Experiments show fast flood mapping in Xinxiang can be achieved by our method..

1. INTRODUCTION

Flooding affects more people than any other natural hazard and hinders social sustainable development, and a large number of casualties and property losses occurred in the affected areas (Hallegatte et al., 2016). The research shows that flood extent and population exposure for 913 large flood events from 2000 to 2018 and the total population in locations with satelliteobserved inundation grew by 58-86 million from 2000 to 2015 (Tellman et al., 2021). As the infrastructure investments (i.e. satellite and ground segments) have been made, the remote sensing image process technologies can put the new earth observation capacities to work (Stylianidis et al., 2020). It is great significance to use multi-temporal remote sensing images for flood monitoring, but the flood is often accompanied by clouds and rain, which made the optical image unusable. Consistent monitoring is achieved because the microwave energy employed by SAR can be transmitted through clouds (Zhang et al., 2022).

The traditional SAR water extraction methods include threshold segmentation method, machine learning, object-oriented segmentation method, etc. Otsu global threshold method (Zhang et al., 2020), local threshold method (Liang and Liu, 2020), minimum cross entropy threshold method (Nguyen and Tran, 2016) are mainly used to segment water bodies by setting thresholds, which is simple and fast, but due to image noise and difficulty in choosing the optimal threshold value leading to false detection; machine learning algorithms such as Markov random field (MRF) (Xinzhi et al., 2019), transfer learning (Xingli et al., 2019) require a large number of manually outlined samples, and bare soil and roads are easily confused with water bodies on SAR images, which makes it difficult to obtain water body mask data manually; object-oriented water extraction approach effectively suppresses image noise by object-level classification, but is computationally intensive (Herrera-Cruz et al., 2009).

For the open water category, the scattering is mirror-like with little energy returning back to the sensor that results in the dark image tones on SAR images (Brisco, 2015). And SAR image is seriously affected by speckle noise which brings great challenges to SAR image water extraction. Recently, image segmentation algorithms such as FCN (Long et al., 2015), SegNet (Badrinarayanan et al., 2017) and U-Net (Ronneberger et al., 2015) have been widely used in remote sensing image processing due to their great feature extraction capability.

The main goal of this study is a flood detection algorithm using Gaofen-3(GF-3) SAR imagery which is based on channel and position attention mechanism, this research produced flood distribution map using rapid water body extraction method. The method enables end-to-end water body extraction and flood extent detection by comparing the distribution of water bodies in two temporal image.

The study is presented in three main sections. In Section 2, the representative study sites and data used for analysis are described. And the proposed water detection algorithms to classify water and non-water are explained in detail in Section 2. Section 3 presents the water extraction accuracy assessments and the flood mapping results using Gaofen-3 images in Xinxiang, China. Finally, Section 4 concludes this work and presents the problems that need to be solved for future research.

2. MATERIALS AND METHOD

2.1 Study site

To analysis the performance of flood detection algorithm, Xinxiang were selected as experimental areas. Xinxiang is in the north of Henan Province, east-central China. The area suffered from large-scale flooding during July 2021 due to high amount of precipitation. The SAR data used in this study concern the overflowing of the Communism Canal, due to the widespread and persistent rainfall, which occurred July 19 to July 20, 2021. About 830 mm of precipitation fell to the square meter on City of Xinxiang and caused floods. The flooding resulted in a disaster area of about 58,877.38 hectares and a crop damage area of about 15,944.43 hectares. Figure 1 showed the specific location of the experimental area, where the pink rectangle is the Xinxiang and the surrounding area.



2.2 Dataset

In total, three SAR images which acquired from GF-3 C-band SAR sensors were used in this study. The datasets are the HV polarization amplitude data under the FSII model at 10m spatial Resolution. The first non-flood dataset(01/07/2020) over the Wuhan City were used for training the proposed model and another two evaluated dataset over Xinxiang were obtained preevent(21/07/2021) and post-event(22/07/2021). These study areas included river, road, field and other landcover types. Table 1, the dataset specifications (i.e. imaging date, product type, and sensor mode, resolution and area) are provided.

Sensor	Date	Dataset	Resolution		
GF3	2020/07/01	Wuhan for training	10m		
	2021/07/21	Vinning frantsat	FSII		
	2021/07/22	Ainxiang for test	HV		
Table 2. Experimental images used in the study					

After the pre-processing, including radiation calibration, complex to amplitude, multi-look, lee filtering 7x7, geocoding, etc., we have relied on SAR and optical image to annotate the imagery on account of the quality by manual outlining. The annotation masks used in this study contain labels 0 and 1 representing no water and water area respectively.

2.3 Method

The proposed water extraction network in this paper is an improved end-to-end U-Net network structure, which incorporates two attention mechanisms, channel attention mechanism and location attention mechanism. The attention mechanism can increase the weight of important features and learn representative deep features (Fu et al., 2020). The network

consists of two parts, the encoding part and the decoding part. The encoding part is consistent with the traditional U-Net structure, which is the extraction of depth features of water bodies. The decoding restores the water body details so that the resulting map as the same scale as the input image. The difference lies in the attention mechanism model embedded at the top of the encoding part, which captures the spatial relationship between water bodies. The decoding embeds the channel attention module after each convolutional layer to enhance the learning of representative features. The model structure was shown in Figure 2.



Figure 2. The proposed water extraction model

The conventional models perform operations on all bands, which leads to a decrease in the learning ability of model. In this work, we introduce a channel attention mechanism(CAM) to give different weights to the feature channels according to their importance, so as to increase the variability between channels, accelerate the convergence of the network, and further improve the learning performance of the network. Channel attention module is show in Fig.2(A). Firstly, we reshape the feature map $A \in \mathbb{R}^{C \times W \times H}$ into $B \in \mathbb{R}^{C \times W \times H}$ is obtained by sequentially feed B into one average pooling layer, two fully connected layers and one sigmoid activation layer. Finally, we perform a matrix multiplication between the channel weight vector and reshape their result to $\mathbb{R}^{C \times W \times H}$.

Extracting representative features is crucial for scene interpretation. The position attention mechanism(PAM) encodes global contextual information into weighted features, and selectively enhances important contextual information according to the spatial attention map. Position attention module is show in Fig.2(B). Firstly, feed the feature map A into a convolution layers to generate two feature maps B and C, $\{B, C\} \in \mathbb{R}^{C \times W \times H}$. Then perform a matrix multiplication

between the transposed C and B. After that apply a softmax layer to calculate the attention map $AM \in \mathbb{R}^{N \times N}$. The feature map D was reshaped and perform a matrix multiplication between D and the transpose of S and reshape it into $\mathbb{R}^{C \times H \times W}$. Finally, it multiply by a scale parameter α and perform a element-wise sum operation with the features. The finial output is a new feature $\mathbf{E} \in \mathbb{R}^{C \times H \times W}$.

Attention mechanism has become an important research topic in deep learning, and specific attention mechanism algorithms view related literature (Chaudhari et al., 2019).

3. EXPERIMENTAL AND RESULTE

3.1 Model training

In this study, the training data were obtained from the GF3 of Wuhan in July 2020, and the training data regions and label image were shown in Figure 3. The image blocks(10240*5120) were cropped into 512*512 blocks, and some samples that were all water bodies or all non-water bodies were removed. Finally, 100 training samples and 4 validation samples with 512×512 were obtained in this study. Additionally, in order to expand the number of training samples and avoid over-fitting the model, the samples are augmented by vertical flip, horizontal flip and diagonal mirroring.



The experimental setup used in this study were summarized at Table 1. The training process was carried out for 6 epochs with 100 iterations per each epoch. And the batch size was set to 4, adam was selected as the optimizer, the initial learning rate was set to 10^{-4} , and a multi-step learning rate decay strategy with the gamma of 0.9 was adopted to update the learning rate. The binary cross-entropy loss function was used for model training.

Operating system	Ubuntu 16.04
Deep learning framework	Pytorch 1.9
GPU	NVIDIA RTX8000
	CUDA10.1

Table 2. Experimental setup used in the study

In order to test the effectiveness of the attention mechanism and the accuracy of the proposed network, we set up some ablation experiments using the same training data. In the ablation

experiments, the three comparison models constituted by removing CAM, removing PAM and removing both CAM and PAM respectively on the proposed network structure in this paper. The network model with both CAM and PAM removed is the traditional U-Net network. The validation data were selected from four representative SAR images for qualitative and quantitative evaluation.

In this paper, we choose four quantitative evaluation indicators of water extraction accuracy including Recall, IoU, OA and F1-score. OA and Recall are binary evaluation metrics. OA is the ratio of the number of correctly classified samples to the number of all samples. Recall is the percentage of samples that are classified as true among all samples that are indeed true. The F1 is the harmonic mean of precision and recall. IoU is the ratio of the intersection of predicted and truth value and the union of predicted and truth value.

$$OA = \frac{TP + TN}{TP + TN + FN + FP}$$
(1)

$$\operatorname{Recall} = \frac{TP}{TP + FN}$$
(2)

$$F1 = \frac{1}{2} \left(\frac{1}{\Pr(1)} + \frac{1}{\operatorname{Recall}} \right)$$
(3)

$$I0U = \frac{1}{TP}$$
(4)

$$bU = \frac{TP}{TP + FN + FP}$$
(4)

Where TP means the prediction is correct, the prediction result is positive, and the true class is positive; FP is prediction wrong, prediction result is positive, and the true class is negative class; FN means the prediction is wrong, the forecast result is negative, and the true result is positive; TN means the wrong is correct, the forecast result is negative, and the true result is negative.

3.2 **Result analysis**

As shown in Table II, the results of combining different modules were evaluated respectively for same dataset. In summary, our method achieved an average OA of 95.95% and Recall of 94.221% which reveals that our network prior to the feature extractor results in a better performance compared to other methods. However, the F1(96.076%) and IOU (92.202%) of our method in this paper are slightly lower than the F1 and IOU of U-Net+PAM network, which indicates that the PAM module in this method has a greater improvement in network performance.

	Recall	F1	IoU	OA		
U-Net	93.127	95.902	91.948	95.810		
U-	93.601	96.082	92.660	95.983		
Net+PAM						
U-	91.122	95.111	90.577	95.057		
Net+CAM						
Our	94.221	96.076	92.202	95.95		
Table 3 Results in test dataset						

Table 3. Results in test data

We selected four images with complex scenes to test the performance of each network in this paper, where roads and fields are easily confused with water bodies. Fig4 shows the image results of the ablation experiments. It can be seen from the Fig4 that our method maintains the edge information better, which further proves that the PAM module has a good enhancement of the feature extraction ability. However, U-Net with CAM has a problem of missed detection in terms of small water bodies. It can be seen that it is still difficult to distinguish the boundary part between water and field.

	Image 1	Image 2	Image 3	Image 4
SAR				
Label				
U-Net+ CAM				
U-Net+ PAM				
U-Net				
Our				

Figure 4. Experimental setup used in the study



Figure 6. The flood extent mapping in Xinxiang

3.3 Model application in Xinxiang flood

The proposed flood mapping methodology consists of four major steps; pre-processing, training data generation, model training, and flood analysis (see Figure 5). Firstly, the multitemporal SAR images of the flooded area are pre-processed, and then the water of the multi-temporal is extracted sequentially by using the pre-training model, and finally the flooding mapping is carried out by overlay analysis.

In the overlay analysis, the part of the water body range intersected by the two temporal represents the unchanged water body range, and the other parts mean that the water body has changed. Compared with the traditional change detection algorithm, this method is simpler and faster. By the superposition relationship between layers, it can quickly reflect the flood development status. Finally, this paper uses two GF3 images on July 21 and July 22, 2021, combined with river and building vector data provided by the National Basic Geographic Information Center (NGCC), to mapping the floods that occurred in the Xinxiang region of Henan Province, China.

As shown in Figure 6, many rivers in Xinxiang were overflowing, with the upstream section of the Sha river in the Communist canal being the most serious, and a large amount of farmland on both sides of the river was flooded and some residential areas were invaded by the flood.

4. CONCLUSION AND FUTURE WORK

In this paper, we propose a water body extraction method for GF3 SAR images and implement flood mapping in Xinxiang, Henan, China. The experimental results show that our model suppresses the SAR image noise and improves the mixed

separation of roads and water bodies. The water body extraction of GF3 images using our method is able to maintain the detail information and has good performance on Recall and OA.

However, there are some problems with the proposed method in this work that remain to be solved. The model training relies on high quality datasets, and the model is only applicable to remote sensing images with the same sensors. In addition, the phenomenon that targets such as roads and farmlands are mixed with water bodies has not been completely solved.

5. ACKNOWLEDGEMENTS

This research is supported by the National Natural Science Foundation of China (Grant No. 42192583)

6. REFERENCES

Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39, 2481-2495.

Brisco, B., 2015. Mapping and Monitoring Surface Water and Wetlands with Synthetic Aperture Radar.

Chaudhari, S., Polatkan, G., Ramanath, R., Mithal, V., 2019. An Attentive Survey of Attention Models.

Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., Lu, H., 2020. Dual Attention Network for Scene Segmentation, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Hallegatte, S., Vogt-Schilb, A., Bangalore, M., Rozenberg, J., 2016. Unbreakable: Building the Resilience of the Poor in the Face of Natural Disasters.

Herrera-Cruz, V., Koudogbo, F., Herrera, V., 2009. TerraSAR-X rapid mapping for flood events. Proceedings of the international society for photogrammetry remote sensing, 170-175.

Liang, J., Liu, D., 2020. A local thresholding approach to flood water delineation using Sentinel-1 SAR imagery. ISPRS Journal of Photogrammetry and Remote Sensing 159, 53-62.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis Machine Intelligence 39, 640-651.

Nguyen, D.B., Tran, G.H.T., 2016. Extraction of water body information from high spatial resolution Sentinel-1 SAR images using Li's minimum cross entropy threshold method: a case study of Thac Ba Lake Yen Bai Province Vietnam. Journal of Mining Earth Sciences 56, 67-77.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net:

Convolutional Networks for Biomedical Image Segmentation, in: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp. 234-241.

Stylianidis, E., Akca, D., Poli, D., Hofer, M., Gruen, A., Sanchez Martin, V., Smagas, K., Walli, A., Altan, O., Jimeno, E., Garcia, A., 2020. FORSAT: a 3D forest monitoring system for cover mapping and volumetric 3D change detection. International Journal of Digital Earth 13, 854-885.

Tellman, B., Sullivan, J.A., Kuhn, C., Kettner, A.J., Doyle, C.S., Brakenridge, G.R., Erickson, T.A., Slayback, D.A., 2021. Satellite imaging reveals increased proportion of population exposed to floods. Nature 596, 80-86.

Xingli, Q., Yang, J., Li, P., Sun, W., 2019. Research on Water Body Extraction from Gaofen-3 Imagery Based on Polarimetric Decomposition and Machine Learning. Geomatics and Information Science of Wuhan University, 6903-6906.

Xinzhi, G., Qingwei, Z., Hua, S., Weiguo, H., Wenjun, X., 2019. Study on water information extraction using domestic GF-3 image. Journal of Remote Sensing 23, 555-565.

Zhang, B., Wdowinski, S., Gann, D., Hong, S.-H., Sah, J., 2022. Spatiotemporal variations of wetland backscatter: The role of water depth and vegetation characteristics in Sentinel-1 dual-polarization SAR observations. Remote Sensing of Environment 270, 112864.

Zhang, M., Chen, F., Liang, D., Tian, B., Yang, A., 2020. Use of Sentinel-1 GRD SAR Images to Delineate Flood Extent in Pakistan. Sustainability 12, 5784.