

# Water-Body Type Classification in Dual PolSAR Imagery Using a Two-Step Deep Learning Method

Qiming Yuan, Jing Lu, Lin Wu, Yabo Huang, Zhengwei Guo, and Ning Li, *Member, IEEE*

**Abstract**—Water-body type classification plays a vital role in ecological conservation, water resource management, and urban planning. Accurate classification can aid decision-makers in understanding the functions of different water-body types, providing key information for urban planning and promoting harmony between human activities and the natural environment. Despite extensive research in the field of water-body segmentation, exploration in the water-body type classification community is not as widespread. Therefore, this article proposes a novel water-body type classification method based on a two-step deep learning model, decomposing water-body type classification into water-body segmentation and water-body type identification. Especially, this method constructs a unique data strategy by organically integrating backscatter features, polarimetric features, and DEM features, providing the model with rich and comprehensive information. In the first step, the segmentation network uses the fused feature to extract all water-body from Synthetic Aperture Radar (SAR) images. Subsequently, the extracted water-body are combined with the input data, forming a multi-feature input for the identification network to distinguish between natural and artificial water-body. During this process, a swarm intelligence optimization algorithm is employed to explore the optimal hyperparameters of the network, including those of the segmentation and identification networks. Finally, the proposed method is assessed using extensive experiments on water-body segmentation tasks, water-body type identification tasks, and joint water-body type classification tasks. This article not only provides a new perspective in the field of water-body type classification but also demonstrates the immense potential of deep learning network hyperparameter optimization and feature fusion in solving such problems.

**Index Terms**—Synthetic aperture radar (SAR), water-body types classification, feature fusion, polarization decomposition method, deep learning, hyperparameter optimize.

## I. INTRODUCTION

WATER, as the source of life and the carrier of civilization, plays a vital role in human life and societal development [1], [2], [3]. Natural and artificial water-body, as the two primary forms of water resources, each bear unique responsibilities [4]. Typical examples of natural and artificial water-body are shown in Figure 1. Natural water-body, such as pristine rivers and lakes, serve as essential habitats for biodiversity [5]. They existence provides irreplaceable value in maintaining ecological balance, mitigating climate change, and offering recreational areas for humans [6]. In contrast, artificial water-body, like man-made lakes and reservoirs, primarily function to store and transport water resources, ensuring basic water needs for human living and agricultural production during drought periods [7]. Additionally, they play a crucial role in flood control and power generation [8]. However, with population growth, accelerated urbanization, and environmental degradation, global water scarcity is becoming increasingly severe [9]. Given these conditions, accurate classification of natural and artificial water-body not only contributes to a comprehensive grasp of the distribution and status of water resources but also provides robust information for decision-makers in the field of water management. This support assists in the more effective management and utilization of water resources, and optimization of water use strategies, and, as a result, it addresses the increasingly severe issue of water scarcity.

Natural water-body is primarily located in ecological reserves and rural areas, while artificial water-body is predominantly found in urban settings. This distribution characteristics present challenges to traditional methods of water-body type classification, such as field surveys, which are often limited by their time-consuming nature and small spatial coverage. With the advancement of remote sensing technology, remote sensing images, possessing advantages such as spatial continuity, temporal dynamics, and extensive coverage, serve as vital data sources for water-body segmentation and water-body type classification fields [10], [11], [12], [13]. Optical remote sensing images, with their rich spectral information and shorter revisit periods, are commonly used in the fields of land use and land cover [14]. However, optical remote sensing images are easily affected by clouds and fog, which can obstruct

This work was supported by the Plan of Science and Technology of Henan Province (232102211043), the Key R&D Project of Science and Technology of Kaifeng City (22ZDYF006), the Key Laboratory of Natural Resources Monitoring and Regulation in Southern Hilly Region, Ministry of Natural Resources of the People's Republic of China (NRMSSHR2022Z01) and the Key Laboratory of Land Satellite Remote Sensing Application, Ministry of Natural Resources of the People's Republic of China (KLSMNR-202302). (*Corresponding author: Lin Wu.*)

Q. Yuan, L. Wu, Y. Huang, Z. Guo and N. Li are with the School of Computer and Information Engineering, Henan University, Kaifeng 475004, China; also with Henan Province Engineering Research Center of Spatial Information Processing, Kaifeng, 475004, China; also with Henan Key Laboratory of Big Data Analysis and Processing, Kaifeng, 475004, China, and also with Henan Engineering Research Center of Intelligent Technology and Application, Kaifeng, 475004, China (e-mail: yuanqiming@henu.edu.cn; henuwl@henu.edu.cn; hyb@henu.edu.cn; gzw@henu.edu.cn; hedalin@henu.edu.cn).

J. Lu are with the Land Satellite Remote Sensing Application Center, Ministry of Natural Resources, Beijing, 100048, China (e-mail: luj@lasac.cn).

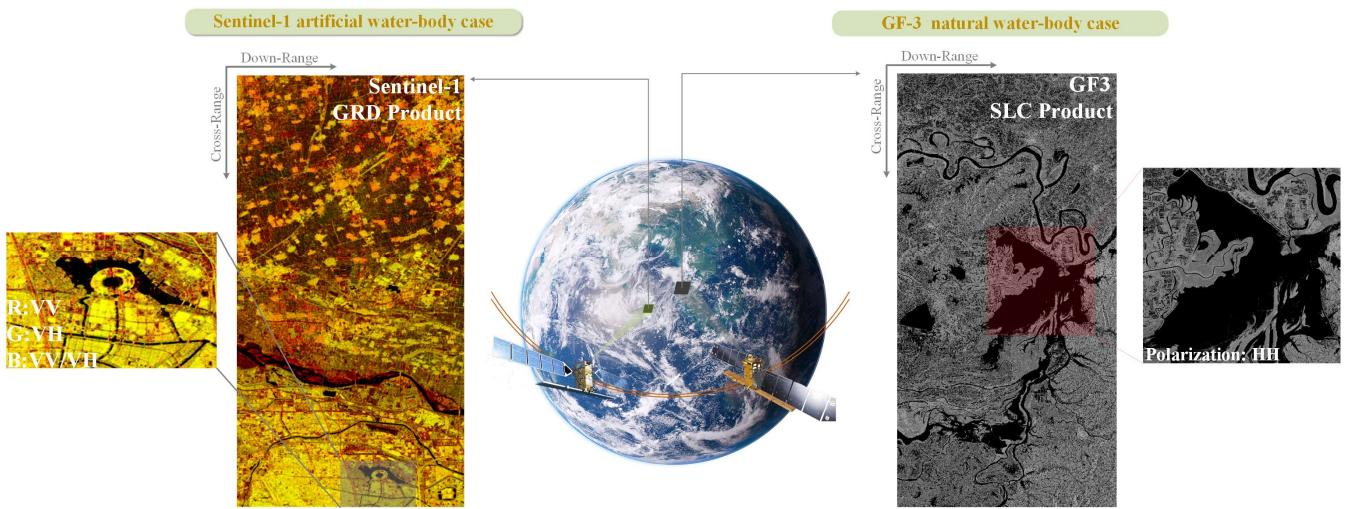


Figure 1. Typical examples of natural and artificial water-body. On the left is an artificial water-body, characterized by its relatively regular shape and clear contrast with the surrounding environment. Conversely, on the right is a natural water-body, which exhibits a relatively irregular shape and transitions naturally into the surrounding environment.

the complete coverage of the target ground object at specific times [15]. In contrast, Synthetic Aperture Radar (SAR) can perceive targets all day and in all weather conditions, thus it is widely used in fields such as agricultural management, target detection, and disaster monitoring [16]. Meanwhile, by applying polarization decomposition methods to PolSAR images, the polarization feature can be extracted. This polarization feature information serves to reveal the physical properties of water-body, thus facilitating more accurate water-body type classification.

In the community of water-body segmentation, scholars have proposed a large number of methods, which are mainly categorized into traditional machine learning methods and deep learning methods. Typical traditional machine learning methods include Markov Random Fields (MRF) [17], Fuzzy C-Means Clustering (FCM) [18], Support Vector Machines (SVM) [19], and edge detection [20]. These methods show high accuracy and efficiency in processing small scale images [21]. However, when dealing with SAR images that contain more geometric and radiometric details, the performance of water-body segmentation will significantly decrease [22]. This is mainly due to the amplitude variations caused by the various types of water-body and complex backgrounds, as well as the inability to fully utilize the contextual information. In response to these challenges, in recent years, deep learning model have demonstrated exceptional performance in the field of water-body segmentation due to their powerful ability to extract features and learn contextual information from images [23].

Although research on water-body extraction has become more in-depth, the field of classifying natural and artificial water-body has not been extensive. However, accurately identifying these two types of water-body is crucial for the entire water resource management system. Precise classification not only optimizes the allocation of water resources, and enhances the efficiency of agricultural irrigation, but also plays a key role in urban planning. Yet, in the past,

water-body type classification tasks primarily relied on traditional machine learning methods [24], [25]. While these methods have achieved certain research results in this field, there are still many problems that need to be addressed. For instance, these methods often struggle to find an appropriate threshold and are significantly affected by speckle noise [25]. Additionally, they have difficulty fully utilizing the multi-feature information of the images. Therefore, leveraging the powerful feature extraction and context learning capabilities of deep learning to enhance the accuracy of water-body type classification represents a challenge that needs to be addressed [27].

Deep learning models have shown great promise in handling complex tasks, yet their full potential has not yet been realized. Particularly in the field of model optimization, much of the scholarly work in recent years has been devoted to increasing the complexity of the model structure [28]. However, the fine-tuning and optimization of model hyperparameters, which can contribute to the stable improvement of model performance, have not been given enough attention [29]. Moreover, it is observed that the effective fusion of different features, such as polarimetric features, backscatter features, and DEM features, also serves as an effective means to improve classification accuracy [30].

In summary, there are three main challenges in the water-body type classification task. First, how to effectively apply the deep learning model for the classification of diverse water-body types. Second, how to optimize the model to enhance its performance without amplifying the model complexity and learnable parameters. Finally, how to improve water-body type classification accuracy through diversifying the input data. For the first challenge, the type classification of water-body is approached in two stages: segmentation and identification. To begin with, DeepLabV3+ is employed to segment all water-body from SAR images. Following this, ResNet50 is applied to identify them as either natural or artificial. For the second

challenge, the hyperparameters of the segmentation network and the identification network are optimized by utilizing the global search of the DBO algorithm and the iterative calculation of the fitness function. For the third challenge, the fusion of backscatter features and polarimetric features of SAR images, and DEM features is used as input data. The major contributions of this article are indicated as follows.

- 1) The proposed two-step method effectively leverages the strengths of both deep learning models. DeepLabV3+ is used for water-body segmentation and ResNet50 is used for water-body type identification. The DeepLabV3+ network can fully extract the semantic features and produce accurate probability maps. ResNet50 solves the problems of gradient vanishing and explosion by allowing gradients to be directly backpropagated through cross-layer connections.
- 2) The DBO algorithm is utilized to optimize the hyperparameters of the segmentation and identification network. As a heuristic optimization algorithm, the DBO algorithm can automatically find the optimal hyperparameter combination through global search and iterative calculation of the fitness function, thereby improving the generalization ability and performance of the model.
- 3) In terms of input data, the fusion of backscatter features and polarimetric features of SAR images, and DEM features is used as the input data of the model. This strategy effectively fuses the abundant information embedded within each feature. Specifically, the backscatter feature can provide backscatter intensity information, while the polarimetric features can reveal the disparity among scattering components of different ground objects, and the DEM features can improve the generalization ability and robustness of the model for complex terrain.

The rest of this article is organized as follows. In Section II, related work is discussed in depth and research progress is reviewed. In Section III, the study area and dataset is presented in detail. In Section IV, the proposed method is introduced. In Section V, the experiment and result are analyzed. In Section VI, an assessment and discussion of the proposed method are provided. In Section VII, the conclusions are briefly concluded.

## II. RELATED WORK

The water-body types classification can be considered as a fine-grained task of water-body segmentation, which involves extracting all water-body from images and then identifying the extracted water-body into natural water-body or artificial water-body. Given the complexity and the high-resolution nature of this task, traditional machine learning methods often fall short of providing accurate and efficient results. In recent years, with the rise of deep learning, Convolutional Neural Networks (CNNs) have first achieved great success in optical and hyperspectral remote sensing image classification.

In [31], Marmanis et al. trained the CNN using ImageNet

pertained networks. In [32], Maggiori et al. first proposed an end-to-end FCN for the classification of large-scale remote sensing images. In [33], Lyu et al. employ a Multiscale Successive Attention Fusion Network (MSAFNet) which uses a combination of modules to efficiently aggregate multiscale features, extract fine-grained water-body features, and mine semantic information for improved segmentation of various water-body. In [34], Shi et al. established a water segmentation dataset based on GF-2 images, applied the U-Net model which has a Channel-Spatial Attention Module (CBAM) for enhanced feature fusion efficiency, leading to improvements across various evaluation indicators in water-body segmentation accuracy, and demonstrated the method's generalizability on multiple datasets including GF-1, GF-6, Landsat8, and EO-1. Compared to the classification methods applies to optical or hyperspectral images, the application of CNNs to SAR image classification is restricted due to the scarcity of annotated datasets. In [16], Zhang et al. propose the Cascaded Fully-Convolutional Network (CFCN) for high-resolution SAR images, combining Upsampling Pyramid Networks (UPNs) and Fully-Convolutional Conditional Random Fields (FC-CRFs) to improve pixel-wise water-body detection and boundary accuracy, while introducing a Variable Focal Loss (VFL) function to address training inefficiency caused by imbalanced data distribution. In [35], Yuan et al. propose a Feature-Fused Encoder-Decoder Network (FFEDN) that integrates backscatter features and polarimetric features from SAR images to improve accuracy and precision in water stream extraction.

To the best of our knowledge, most studies concentrate on water-body segmentation, with comparatively less research focusing on the classification of water-body types. In [24], Huang et al. apply a two-level machine-learning framework for water-body types identification in urban areas, validated on high-resolution GeoEye-1 and WorldView-2 images over Wuhan and Shenzhen, two megacities in China. In [25], Xie et al. propose a novel method based on traditional machine learning for supervised water-body extraction and classification from Radarsat-2 Fully Polarimetric (FP) SAR data, validated over Suzhou and Dongguan cities in China. However, scholarly studies on dam reservoir extraction have provided water-body type classification with a new perspective and unique considerations [36], [37], [38], [39]. In [36], Fang et al. introduce a framework based on ResNet50 for identifying global artificial reservoirs from Landsat 8 satellite images. In [38], Malerba et al. train a ResNet34 to detect farm dams using high-resolution satellite images.

Intuitively, it may seem feasible to convert the binary segmentation problem of water-body segmentation directly into a multi-class segmentation problem. However, such a transformation often yields unsatisfactory results due to the negligible pixel-level differences between natural and artificial water-body. Therefore, delving deeper into the exploration and utilization of multi-features, along with the application of more robust models, is crucial for amplifying the advantages of multi-features and accurately distinguishing between types of water-body.

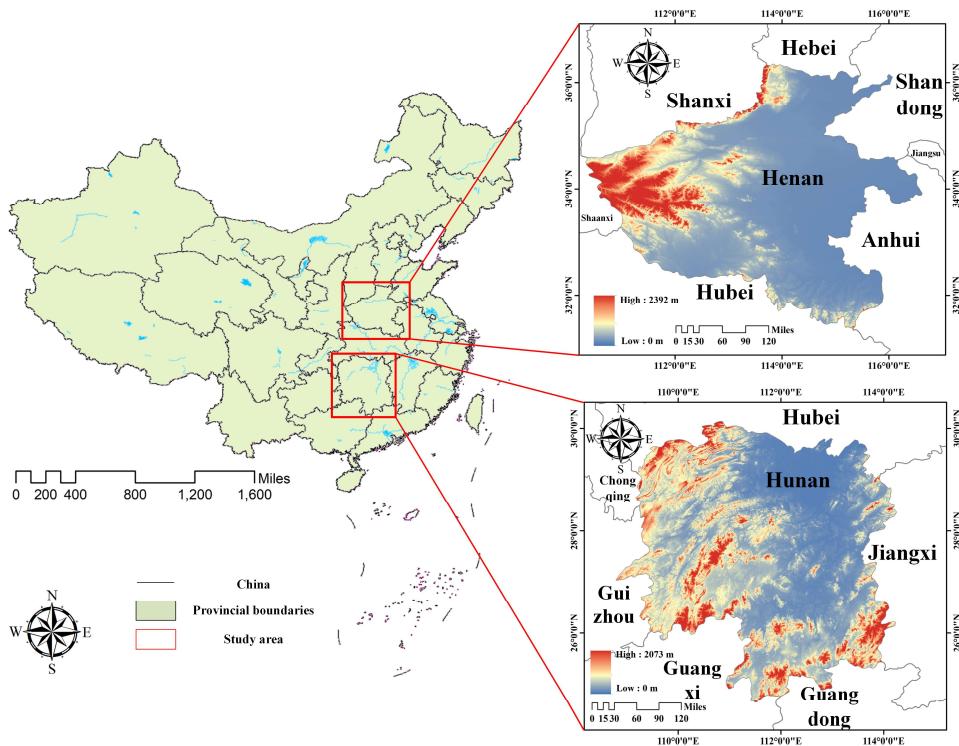


Figure 2. Location of study area.

### III. STUDY AREA AND DATA

#### A. Study area

China, grappling with a severe shortage of freshwater resources, is identified as one of the thirteen most water-deficient countries in the world. The per capita water resource in China is a mere 2200 cubic meters, which is a quarter of the global average. Within this macroscopic context, this article narrows its focus to the provinces of Henan and Hunan in China as a study area. The water resource conditions in these two provinces exemplify the contrasting situations of scarcity and abundance in China. This renders two provinces indispensable case studies for comprehending and addressing the water resource issues prevalent in China. The location of the study area is shown in Figure 2.

Henan Province, located in the central part of China (longitude 110°21' to 116°39' E, latitude 31°23' to 36°22' N), is a major agricultural province. Despite spanning the basins of the Yangtze River, Yellow River, Huai River, and Hai River, its water resources are not abundant, with per capita water resource allocation only one-sixth of the national average. In the South-to-North Water Diversion Middle Route Project, Henan Province has the longest project canal and the largest water allocation. Furthermore, Henan Province has several water irrigation projects, such as the Yellow River Diversion Project and the Huai River Diversion Project. However, these projects cannot fully resolve Henan's water shortage issue, but instead, to some extent, reflect the tense water resource situation in Henan Province.

In contrast, Hunan Province (longitude 108°47' to 114°15' E, latitude 24°38' to 30°08' N) is located in the subtropical humid

TABLE I  
DETAILED PARAMETERS OF THE S-1A SATELLITES

Parameters	S-1A
Waveband	C
Center frequency	5.405GHz
Revisit cycle	12 days
Orbital altitudes	693 km
Polarization	Single / Dual
Resolution	5 meters–20 meters
Swath	20 km–400 km
Incidence angle	20° - 45°
Number of modes	4

monsoon climate zone and boasts abundant water resources. The province has an average annual precipitation of 1427 mm, an average total water resource volume of 163 billion cubic meters, and a per capita water resource allocation of 2500 cubic meters. These figures rank 7th, 6th, and 13th in China, respectively. However, despite Hunan Province's natural water abundance, the large amount of transboundary water and the small amount of controllable water lead to long-term issues of resource and structural water shortages, posing challenges to Hunan Province's water use guarantee.

These two provinces present a stark contrast in water resource conditions: one is resource-rich, and the other is resource-scarce. At the same time, the water resource issues of both provinces are representative, reflecting the diversity and complexity of China's water resource issues. By studying the water resource conditions in these two provinces, a deeper

TABLE II  
PARAMETERS OF THE S-1A SAR DATASET

Province	Time(M/D/Y)	Range Spacing (m)	Azimuth Spacing (m)	Orbit Direction	Processing Level
Henan	7 March 2023	5	20	Ascending	L1-SLC (IW)
	7 March 2023	5	20	Ascending	L1-SLC (IW)
	12 March 2023	5	20	Ascending	L1-SLC (IW)
	12 March 2023	5	20	Ascending	L1-SLC (IW)
Hunan	1 April 2022	5	20	Ascending	L1-SLC (IW)
	1 April 2022	5	20	Ascending	L1-SLC (IW)

understanding of China's water resource issues can be achieved, providing references for future water resource management and allocation.

#### B. Satellites

The Sentinel-1 mission, an integral part of the European Space Agency's Copernicus Programme, is committed to delivering precise, timely, and readily accessible data to enhance environmental management, as well as to comprehend and alleviate the impacts of climate change. The mission comprises a two-satellite constellation, Sentinel-1A (S-1A) and Sentinel-1B (S-1B), launched in 2014 and 2016, respectively, with the primary objective of monitoring both land and oceanic regions. Notably, the introduction of the S-1A satellite signifies the inaugural deployment of the Copernicus project's C-band SAR satellites, which offer their data products at no cost with a revisit interval of 12 days. The S-1A is equipped to operate in the Interferometric Wide (IW) mode, providing either single (VV or HH) or dual polarizations (VV/VH or HH/HV) Single Look Complex (SLC) data. This model yields a nominal resolution of  $5 \times 20$  m (in range and azimuth) and encompasses a swath width of approximately 185 kilometers. This represents the highest resolution that is freely accessible for remote sensing imagery. The detailed parameters of the S-1A satellites are given in Table I.

#### C. Satellites Dataset

The SAR dataset utilized comprises seven images captured by the S-1A satellite, with four images from the Henan province and the remaining three from the Hunan province. These SLC images are procured in VV/VH polarization under the IW mode. SLC images, encompassing both amplitude and phase information, furnish us with abundant polarization details. The SAR imagery of Henan province is obtained in March, while that of Hunan province is captured in August. Detailed parameters are delineated in Table II.

#### D. S-1A Data Preprocessing

Figure 3 meticulously illustrates the preprocessing workflow necessary for the extraction of backscatter features and polarimetric matrix from the SLC image. These procedures encompass the application of the orbit file, the removal of thermal noise, calibration, deburst, generated of polarimetric matrix, multi-looking, speckle filtering, and the application of

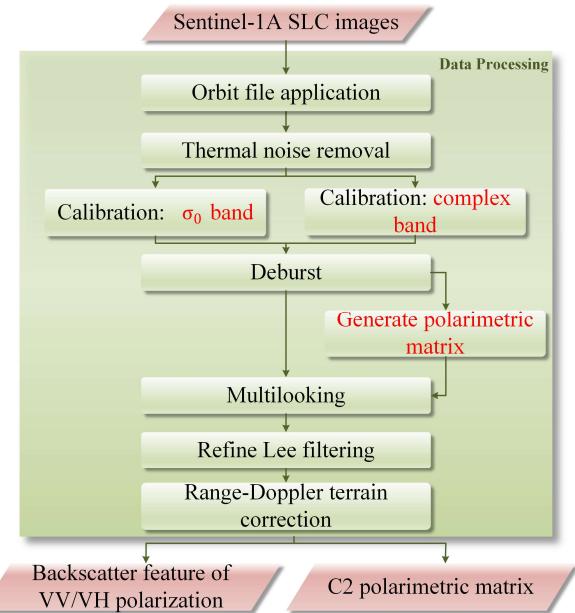


Figure 3. The pre-processing of the S1-A SLC image.

range-Doppler terrain correction. Differences in the generation of backscatter features and the polarimetric matrix arise from the selection of varied outputs during calibration. At the same time, the polarimetric matrix is generated during the deburst procedure. It is important to underscore that during the multi-looking phase, the number of range looks and azimuth looks are set to four and one, respectively. Furthermore, the approach delineated in [40] can effectively mitigate thermal noise. For precise geographical information extraction, the range-Doppler terrain correction employs Shuttle Radar Topography Mission (SRTM) data with a 30m resolution. Through the operation of deburst, each burst, containing effective signal parts, is successfully merged. The processes of multi-looking and filtering serve to minimize the impact of coherent speckle noise.

1) *Acquisition of Backscatter Feature*: Upon completion of the aforementioned preprocessing steps, the backscatter feature of the VV and VH polarization can be obtained.

2) *Acquisition of Polarimetric Feature*: After completing the step of generating the polarimetric matrix, each pixel of the

TABLE III  
FIVE CANDIDATE FEATURE COMBINATIONS

	VV	VH	DEM	$m_v$	$m_s$	$m_{dif}$
Com. A	✓	✓	✓			
Com. B	✓	✓	✓		✓	
Com. C	✓	✓	✓			✓
Com. D	✓	✓	✓			✓
Com. E	✓	✓	✓	✓	✓	✓

SLC data is represented as a  $C_{2 \times 2}$  covariance matrix. Then, the covariance matrix of  $C_{2 \times 2}$  is transformed into a Stokes vector, as shown in the Equation (1).

$$C_{2 \times 2} = \begin{bmatrix} c_{11} & c_{12} \\ c_{12}^* & c_{22} \end{bmatrix} \rightarrow \underline{S} = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \end{bmatrix} = \begin{bmatrix} c_{11} + c_{22} \\ c_{11} - c_{22} \\ 2\operatorname{Re}(c_{12}) \\ 2\operatorname{Im}(c_{12}) \end{bmatrix} \quad (1)$$

where  $\operatorname{Re}(c_{12})$  and  $\operatorname{Im}(c_{12})$  denote the real part and the imaginary part of  $C_{2 \times 2}$ , respectively. The  $c_{11}$ ,  $c_{12}$ ,  $c_{12}^*$  and  $c_{22}$  represent  $\langle |S_{VH}|^2 \rangle$ ,  $\langle S_{VH}S_{VV}^* \rangle$ ,  $\langle S_{VV}S_{VH}^* \rangle$  and  $\langle |S_{VV}|^2 \rangle$ , respectively, where  $\langle \rangle$  is multilook and/or speckle filtering.

Within the context of the model-based decomposition method in [41], the Stokes vector undergoes decomposition into a volume model and a polarized wave, as delineated in Equation (2).

$$\underline{S} = m_v s_v + m_s s_p = m_v \begin{bmatrix} \pm 0.5 \\ 0 \\ 0 \end{bmatrix} + m_s \begin{bmatrix} 1 \\ \cos 2\theta \\ \sin 2\theta \cos \delta \\ \sin 2\theta \sin \delta \end{bmatrix} \quad (2)$$

where  $s_v$  and  $s_p$  denote the partially polarized portion and the polarized portion, respectively, with  $m_v$  and  $m_s$  being the corresponding power. The angle  $\theta$  serves to assess the proportion between wave components, while  $\delta$  represents the cross-polarized phase. It merits attention that the volume model is constructed based on the random dipole cloud model [41].

So far, Equation (1) yields four known parameters, while Equation (2) determines four unknown parameters. It is shown in [41] that the parameter  $m_v$  can be obtained by solving the quadratic equation.

Then, obviously, the  $m_s$  can be calculated in Equation (3).

$$m_s = s_1 - m_v \quad (3)$$

The  $m_{dif}$  can be calculated in Equation (4).

$$m_{dif} = m_v - m_s \quad (4)$$

Ultimately, three extracted polarimetric feature components are obtained through polarization decomposition.

#### E. DEM Feature

In addition to considering the backscatter features and polarimetric features of SAR images, DEM feature are further introduced. The introduction of DEM feature aims to enhance the model's robustness and generalization ability to more complex terrain features. Especially in alpine regions, DEM feature can provide unique and critical information about the height of the terrain. The usefulness of DEM feature may be

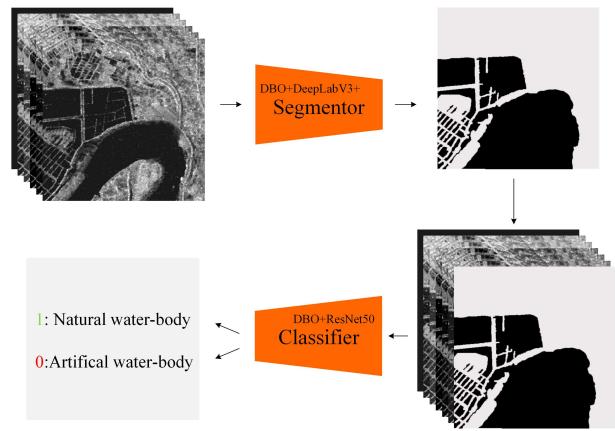


Figure 4. Overview of the proposed water-body type classification. The fusion of backscatter features, polarimetric features, and DEM features is input into the segmentor to obtain a probability map with accurate water-body locations. Subsequently, the probability map is combined with the input features and fed into the classifier to identify the extracted water-body into natural and artificial water-body.

limited in urban areas because water-body and other ground objects in the city often exhibit similar characteristics on the DEM feature, challenging feature extraction in the model. All DEM feature used in this article are from SRTM data, which has a resolution of 30m. SRTM data not only has a wide coverage, but also has a good performance in elevation accuracy and terrain representation, making it an ideal data source to be considered.

#### F. Feature Combination

The input data fuses the backscatter and polarimetric features of the SAR image, and the DEM features to form five combinations as shown in Table III. These combinations are separately applied to water-body segmentation tasks and water-body type identification tasks, aiming to select the optimal feature combination. In previous studies, water-body segmentation is largely based on backscatter features, whereas water-body type identification more heavily relied on texture features and shape features. Given this context, this article focuses on exploring possibilities beyond single backscatter, texture, and shape features. Specifically, the potential of polarimetric and DEM features is introduced, and delves into the comprehensive performance when fusing backscatter, polarimetric, and DEM features. In order to effectively validate the performance of the model upon the addition of polarimetric features and DEM features, Combination A (Com. A) includes only both backscatter features and DEM features. In Com. B-E, the introduction of various polarimetric features forms diverse combinations, enabling a deeper investigation into the specific impact of these features on water-body segmentation and identification tasks.

## IV. METHOD

This article aims to segment water-body in SAR images and classify them based on the formation of natural or artificial. First, the water-body is finely extracted by a segmentation model to obtain a probability map with accurate water-body location. Then, the probability map is added to the channel of

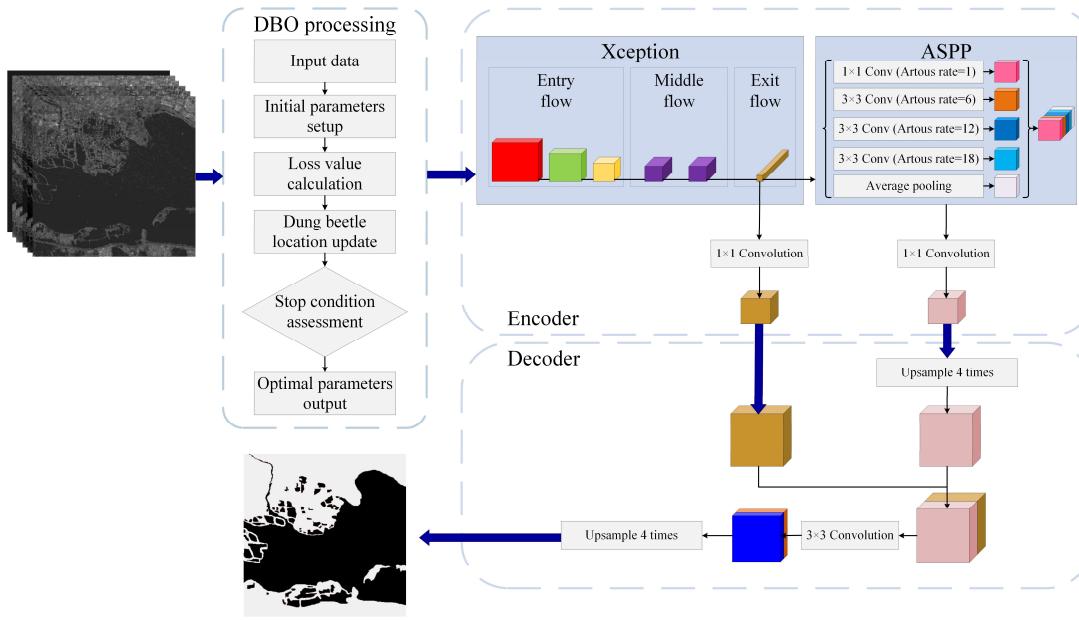


Figure 5. Architecture of the proposed segmentation model.

the input data. Finally, the identification model is utilized to learn all channel features in order to identify the water-body as either a natural or artificial water-body. The proposed method pipeline is shown in Figure 4.

#### A. Water-body Segmentation

*1) Base Model:* DeepLabV3+ is used as the base model for the water-body segmentation task [42]. It is comprised of an Encoder that amalgamates the Atrous Spatial Pyramid Pooling (ASPP) network and the Xception network, coupled with a straightforward yet efficacious Decoder. The Encoder starts with an Xception network for mining features of water-body. The Xception network uses depthwise separable convolutions, which decompose standard convolutions into depthwise convolutions and point convolutions, thereby greatly reducing computational complexity. Specifically, depthwise convolution operates spatial convolutions independently across each input channel, whereas pointwise convolutions serve to amalgamate the outputs of depthwise convolutions. Moreover, to excavate multi-scale contextual information, atrous convolutions and the ASPP are incorporated based on the Xception network, aiming to handle multi-scale information within the images. The ASPP module is comprised of three parallel atrous convolutional layers, a  $1 \times 1$  convolutional layer, and a global average pooling layer, enabling the capture of information at varying scales to procure richer contextual details. Finally, the features at different scales are concatenated into a new high-dimensional feature representation, with an additional  $1 \times 1$  convolutional layer employed to generate the final feature representation. The Decoder structure of DeepLabV3+ is ingeniously simplistic but extremely potent. It first upsamples the output of the Encoder by a factor of four, and then fuses the upsampled feature map with low-dimensional features which are the features extracted by the Xception feature extraction network. The fused features are passed through a  $3 \times 3$  convolutional layer and upsampled to the original size of the input image to produce the final water-

body segmentation result. Architecture of proposed segmentation model is shown in Figure 5.

*2) Network Loss:* Given the influence of speckle noise and considerable disparity in the water-body distribution in SAR imagery, ensuring a balanced or nearly balanced pixel count across various classes is a significant challenge. As a result, the pixel quantity among categories in SAR image datasets often displays a notable imbalance. This imbalance can detrimentally affect the stability of network training, potentially leading to an overemphasis on classes with abundant pixels, while neglecting those with fewer. Such a situation could negatively impact the final classification performance. To mitigate these issues, the application of focal loss function is incorporated for segmentation model [43]. For each pixel  $i$ , the loss is defined as

$$L_{\text{focal}}(P_i, X_i) = \begin{cases} -\alpha(1 - P_i)^\gamma \log(P_i), & X_i = 1 \\ -(1 - \alpha)P_i^\gamma \log(1 - P_i), & X_i = 0 \end{cases} \quad (5)$$

where  $X_i$  is the ground truth category label of the  $i$ th pixel, 1 denotes the water-body category, and 0 denotes the non-water-body category.  $P_i$  is the model's prediction probability for the true category. When the prediction is correct (i.e.,  $P_i$  is close to 1),  $(1 - P_i)$  is close to 0, which reduces the loss. When the prediction is wrong (i.e.,  $P_i$  is close to 0),  $(1 - P_i)$  is close to 1, which increases the loss.  $\alpha$  is an adjustable weighting parameter to balance the importance of different categories. In the water-body segmentation problem, by setting  $\alpha$ , the importance of water-body can be increased and the importance of non-water-body can be decreased, so as to alleviate the problem of an imbalance in the number of pixel points between water-body and non-water-body.  $\gamma$  is an adjustable focus parameter used to control the contribution of difficult and easy samples to the loss. When  $\gamma > 0$ , the loss of difficult-to-categorize samples increases while the loss of easy-to-categorize samples decreases. Consider an illustrative example where  $\gamma$  is set to 3. For a pixel from water-body that is highly

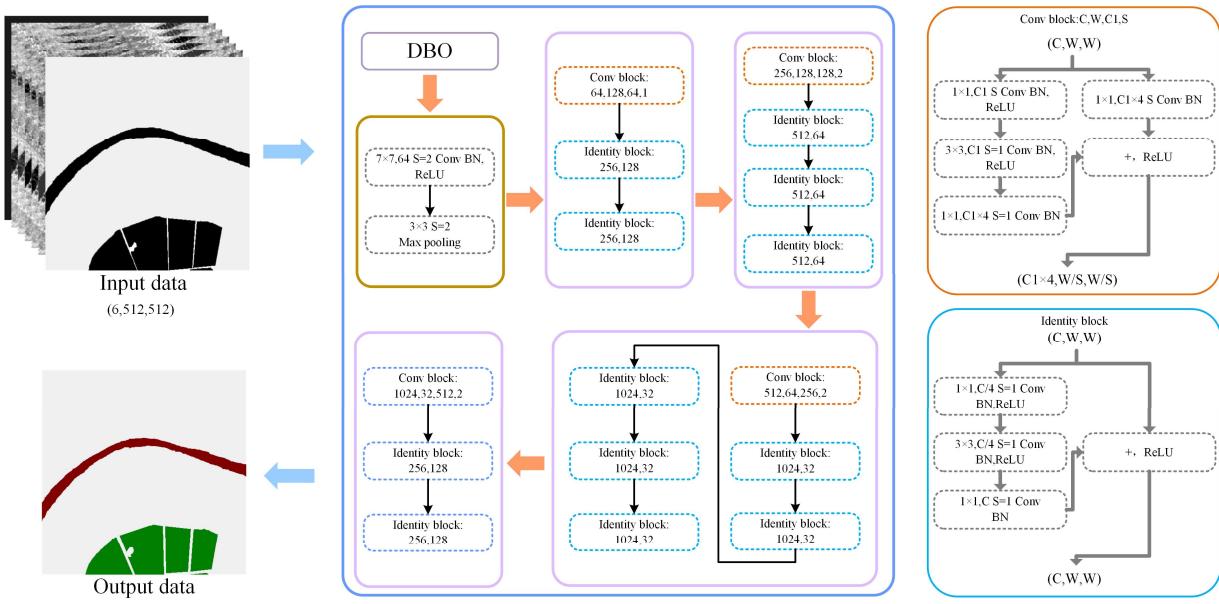


Figure 6. Architecture of the proposed identification model.

likely to be predicted as water-body ( $P_i = 0.8$ ), the modulating factor is calculated as  $(1 - P_i)^Y$ , which is 0.008. This pixel, showing a high prediction confidence, falls under the category of easy-to-predict pixels. Accordingly, its loss value diminishes significantly, by a factor of 125. On the other hand, when a water-body pixel shows low predictive probability ( $P_i = 0.2$ ), the modulating factor  $(1 - P_i)^Y$  will increase to 0.512. Being a hard-to-predict pixel, the loss value only contracts by a factor of 0.512, thus receiving a larger weight in the loss function compared to an easy-to-predict pixel.

#### B. Water-body types identification

1) *Base Model*: ResNet50 is employed as the base model for water-body type identification tasks [44]. ResNet is introduced to address the issue of accuracy degradation due to layer saturation during the training of CNNs. ResNet's architecture incorporates distinct groups of residual blocks to learn the residuals between the inputs and outputs of each network layer. Architecture of proposed identification model is shown in Figure 6. The loss function of the network can be denoted as Equation (6). When extended to a multi-layer network, the output of each layer of the network as show in Equation (7). According to the principle of the chain rule, the gradient of the  $i$ th layer can be deduced as represented in Equation (8).

$$L = \text{Loss}(X_n, Y) \quad (6)$$

where  $L$  is the loss value,  $\text{Loss}$  is the loss function,  $X_n$  is the model output,  $Y$  is the label.

$$\begin{aligned} X_n &= F_{n-1}(X_{n-1}, W_{n-1}), \\ X_{n-2} &= F_{n-2}(X_{n-2}, W_{n-2}), \dots, \\ X_2 &= F_1(X_1, W_1) \end{aligned} \quad (7)$$

where  $F$  represents model,  $W$  denotes parameter of model.

$$\frac{\partial L}{\partial X_i} = \frac{\partial L}{\partial X_n} \frac{\partial X_n}{\partial X_{n-1}} \times \dots \times \frac{\partial X_{i+1}}{\partial X_i} \quad (8)$$

From this deduction, it is observable that, with the backpropagation of error, the gradient of the preceding layer in the network progressively diminishes.

To solve this problem, the residual block is introduced, as shown in the Conv block and Identity block on the right side of Figure 6. In essence, the transformation modifies the output layer from  $H(x) = F(x)$  to  $H(x) = F(x) + x$ , the forward propagation transitioning from Equation (7) to Equation (9) and the backpropagation transitioning from Equation (8) to Equation (10). Therefore, despite the increasing depth of the network, the issue of vanishing gradients is avoided.

$$\begin{aligned} X_n &= F_{n-1}(X_{n-1}, W_{n-1}) + X_{n-1}, \dots, \\ X_1 &= F_1(X_0, W_0) + X_0 \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{\partial L}{\partial X_i} &= \frac{\partial L}{\partial X_n} \times \dots \times \frac{\partial X_{i+1}}{\partial X_i} \\ &= \frac{\partial L}{\partial X_n} \left[ 1 + \frac{\partial}{\partial X_i} \sum_{L=i}^{n-1} F(X_L, W_L) \right] \end{aligned} \quad (10)$$

The residual module also contributes to the reduction in parameter count. This design initially employs a  $1 \times 1$  convolution to decrease the channel dimensions from 256 to 64, before subsequently reinstating the original dimensionality via another  $1 \times 1$  convolution. This configuration results in a significant reduction in the total number of parameters as compared to conventional residual modules. Simultaneously, as the depth of the network increases, the count of ReLU activation functions escalates correspondingly, thereby enhancing the model's non-linear expressive power and generalization capability.

*2) Network Loss:* Cross-Entropy (CE) loss is used for water-body type identification [45]. For the water-body types identification task, the true label of a sample is represented as a one-hot encoded vector  $Y = [y_{natural}, y_{artificial}, y_{background}]$ . The output of the model is a predicted probability distribution  $P = [p_{natural}, p_{artificial}, p_{background}]$ , corresponding to the three classes respectively. The CE loss can be calculated as

$$\begin{aligned} \text{CE}(P_i, Y_i) = & -[y_{natural} \log(p_{natural}) \\ & + y_{artificial} \log(p_{artificial}) \\ & + y_{background} \log(p_{background})] \end{aligned} \quad (11)$$

This equation sums over all classes. The term  $Y_i \log(P_i)$  only remains for the true class because  $Y_i$  for the true class is 1 and 0 for the others. For instance, if the true label of pixel  $m$  is artificial water-body, it is represented as  $Y_m = [0 \ 1 \ 0]$ . The model prediction for this pixel is  $P_m = [0.1 \ 0.8 \ 0.1]$ . The loss value for this pixel is thus calculated through the Equation () .

$$\begin{aligned} \text{CE}_m(P_m, Y_m) = & -[0 \times \log(0.1) + 1 \times \log(0.8) \\ & + 0 \times \log(0.1)] = 0.22 \end{aligned} \quad (12)$$

The logarithmic function in the CE loss ensures a small loss when the model's prediction is close to the true class, approaching zero for perfect accuracy. Conversely, if the predictions are entirely inconsistent with the true labels, the loss will tend toward infinity.

### C.DBO

Despite the exceptional performance and irreplaceable capabilities demonstrated by deep learning networks such as DeepLabV3+ and ResNet across a variety of tasks, presenting broad application prospects when addressing diverse challenges, it is observed that their performance is often contingent upon the meticulous selection of hyperparameters. The configuration of these hyperparameters is commonly dictated by human expertise or refined through exhaustive manual tuning. This method is not only labor-intensive but also may not necessarily yield the optimal model performance. To tackle this issue, the proposed method employs the DBO algorithm to optimize the hyperparameters of segmentation and identification networks.

The DBO algorithm is a heuristic optimization algorithm, inspired by the behavioral patterns of dung beetles in nature [46]. These behavioral patterns are translated into five distinct update rules within the algorithm, serving to pinpoint the optimal solution in the hyperparameter space. After defining the hyperparameter space, the dung beetle's ball-rolling behavior is simulated to conduct a global search. This behavior enables the search process to cover the entire hyperparameter space, thereby avoiding premature convergence to local optima. Specifically, a group of ball-rolling dung beetles is initially established, with each assigned a random search direction. Then, each dung beetle starts to roll according to its search direction, updating its position according to Equation (13). This process continues until a certain number of iterations are reached.

$$x_i(t+1) = x_i(t) + \alpha \times k \times x_i(t-1) + b \times \Delta x, \quad (13)$$

$$\Delta x = |x_i(t) - X^\omega|$$

where, in the context of the current iteration denoted by  $t$ , the position of the  $i$ th dung beetle is represented by  $x_i(t)$ . The constant  $k$  signifies the deflection coefficient, while  $\alpha$  is a coefficient that takes values of either -1 or 1. The term  $X^\omega$  corresponds to the globally least favorable position.

Subsequently, the dancing behavior of dung beetles is simulated to optimize the search process. When a dung beetle encounters impediments or finds progression halted, it executes a sequence of rotations and pauses, utilizing this "dance" to recalibrate its position. Equation (14) is employed to simulate this behavior, adjusting the dance angle to modify the dung beetle's search direction.

$$x_i(t+1) = x_i(t) + \tan(\theta) |x_i(t) - x_i(t-1)| \quad (14)$$

where,  $\theta$  represents the angle of deviation.

Additionally, inspired by the behavior of female dung beetles choosing locations for egg-laying, a boundary selection strategy is proposed, defined as Equation (15):

$$\begin{aligned} Lb^* &= \max(X^* \times (1-R), Lb), \\ Ub^* &= \min(X^* \times (1+R), Ub) \end{aligned} \quad (15)$$

where,  $X^*$  represents the current local best position, while  $Lb^*$  and  $Ub^*$  denote the lower and upper bounds of the spawning region, respectively. During the iteration process, the position of the brood ball is dynamically adjusted according to the Equation (16).

$$\begin{aligned} B_i(t+1) = & X^* + b_1 \times (B_i(t) - Lb^*) \\ & + b_2 \times (B_i(t) - Ub^*) \end{aligned} \quad (16)$$

where,  $B_i(t)$  conveys the spatial data of the brood ball, while  $b_1$  and  $b_2$  are vectors of dimension  $1 \times D$ , generated randomly.

What's more, the simulation of small dung beetle's foraging behaviors is conducted to enhance search accuracy. Specifically, the optimal foraging area is defined in Equation (17).

$$\begin{aligned} Lb^b &= \max(X^b \times (1-R), Lb), \\ Ub^b &= \min(X^b \times (1+R), Ub) \end{aligned} \quad (17)$$

where  $X^b$  is indicative of the universally optimal position. Then, each small dung beetle's position is updated according to Equation (18)

$$\begin{aligned} x_i(t+1) = & x_i(t) + C_1 \times (x_i(t) - Lb^b) \\ & + C_2 \times (x_i(t) - Ub^b) \end{aligned} \quad (18)$$

where,  $C_1$  signifies a random number that adheres to a normal distribution, while  $C_2$  represents a random vector confined to the interval  $(0,1)$ .

At the same time, a simulation of the dung beetle's pilfering behavior was conducted. Assuming that there are more food resources near the global optimal position  $X^b$ , stealing dung beetles are more inclined to move in this direction. This process is simulated through Equation (19).

$$x_i(t+1) = X^b + S \times g \times (|x_i(t) - X^*| + |x_i(t) - X^b|) \quad (19)$$

where  $g$  is defined as a random vector adhering to a normal distribution, while  $S$  represents a fixed constant.

By simulating the behavior of dung beetles, effective optimization of hyperparameters in deep learning networks is achieved. The global search strategy of the DBO algorithm, particularly the simulation of dung beetles' ball-rolling and dancing behaviors, allows for a comprehensive exploration of the hyperparameter space, thereby avoiding premature convergence to local optima. This approach overcomes the labor-intensive and inefficient issues of traditional manual adjustment of hyperparameters, thereby enhancing the performance and stability of deep learning networks.

## V. EXPERIMENT AND RESULT

### A. Dataset

With the data samples introduced in Section III, these data samples are cropped into the nonoverlapping patches of  $512 \times 512$  pixels, because large-scale images would cause a huge computational burden. A total of 694 image patches with water-body are obtained. Among them, there are 291 image patches from Henan Province and 403 image patches from Hunan Province. In this article, they are divided into two subsets, which are used for water-body segmentation and water-body type identification tasks, respectively. To keep the data balanced, the segmentation and identification datasets each have 347 images. In the segmentation set, each sample is associated with a ground truth label with a pixel value of 1 for water-body and 0 for background. In the identification set, label 0 is assigned to natural water-body, label 1 is assigned to artificial water-body, and label 2 is assigned to background. For the water-body type identification task, it is performed on the segmentation set, but the goal is to segment water-body and classify them as natural water-body or artificial water-body. In the segmentation dataset, there are a total of 683 water-body, with an average of 1.96 per image. 278 images serve as the training set, while another 69 images form the testing set. The training set for the identification network consists of 300 images, containing a total of 205 natural water-body and 197 artificial water-body. The test set is composed of 47 images, which include 32 natural water-body and 29 artificial water-body.

### B. Experimental Setup

*1) Implementation Details:* The segmentation model uses DeepLabV3+ as the backbone. Use an input image patch of  $512 \times 512$  pixels. The learning rate, batch-size, and epoch of the segmentation model are all obtained by DBO iterative search, and they are 0.0003, 4, and 150 respectively. During training, data augmentation methods of horizontal and vertical flipping, and random channel offset are used. The identification model uses ResNet50 as the backbone. Also, use an input image with patches of  $512 \times 512$  pixels. The learning rate, batch-size, and epoch of the identification model are also iteratively calculated by the DBO algorithm, which are 0.006, 8, and 500, respectively. The experiments are all carried out in the Matlab environment using NVIDIA TITAN GPU. The operating system is Windows 10.

*2) Evaluation Protocol:* For the water-body type identification, a classification task, accuracy is computed as the

proportion of images correctly predicted relative to the total number of images. The tasks of water-body segmentation and water-body type classification are regarded as a multi-class segmentation problem. For water-body segmentation, it is divided into two categories: water-body and background. For the latter, it is divided into three categories: in the segmentation task, regions not predicted as water-body are the class labels of the background; for each extracted water-body, if it is identified as a natural water-body or an artificial water-body, all pixels have the same class label in this area. The widely utilized Intersection over Union (IoU) indicator is employed to assess segmentation outcomes, calculating the ratio of the intersection to the union of predicted regions and ground truth regions for a specific category. The mean IoU (mIoU) of all classes in each image is determined, with this value subsequently averaged over the entire set.

In the context of water-body segmentation and water-body type classification, the background category typically occupies a larger proportion in the image patches, thereby influencing a dominant mIoU value. Therefore, in some works, it is removed from mIoU. Furthermore, in the case of water-body type classification, the IoU of natural and artificial water-body are averaged with equal weight, while the primary objective of this article is the classification of natural and artificial water-body. To tackle these issues, the experiments perform a comprehensive evaluation. For water-body segmentation, the IoU of the water-body is computed, along with the mIoU of the water-body and the background. In the water-body type classification, the IoU for natural water-body, denoted as  $\text{IoU}^n$ , and that for artificial water-body, denoted as  $\text{IoU}^a$ , are calculated and reported. Additionally, the mIoU for both natural and artificial water-body, denoted as  $\text{mIoU}^{na}$ , as well as the mIoU for natural water-body, artificial water-body, and backgrounds, denoted as  $\text{mIoU}^{nab}$ , are also computed and presented.

TABLE IV  
QUANTITATIVE EVALUATION OF THE SEGMENTATION TASK BY DIFFERENT COMBINATIONS

	IoU (%)	mIoU (%)
Com. A	63.90	65.08
Com. B	71.14	75.69
Com. C	86.22	88.24
Com. D	71.07	72.68
Com. E	77.55	79.59

TABLE V  
QUANTITATIVE EVALUATION OF THE IDENTIFICATION TASK BY DIFFERENT COMBINATIONS

	Accuracy (%)
Com. A	85.85
Com. B	88.38
Com. C	93.63
Com. D	87.19
Com. E	90.19

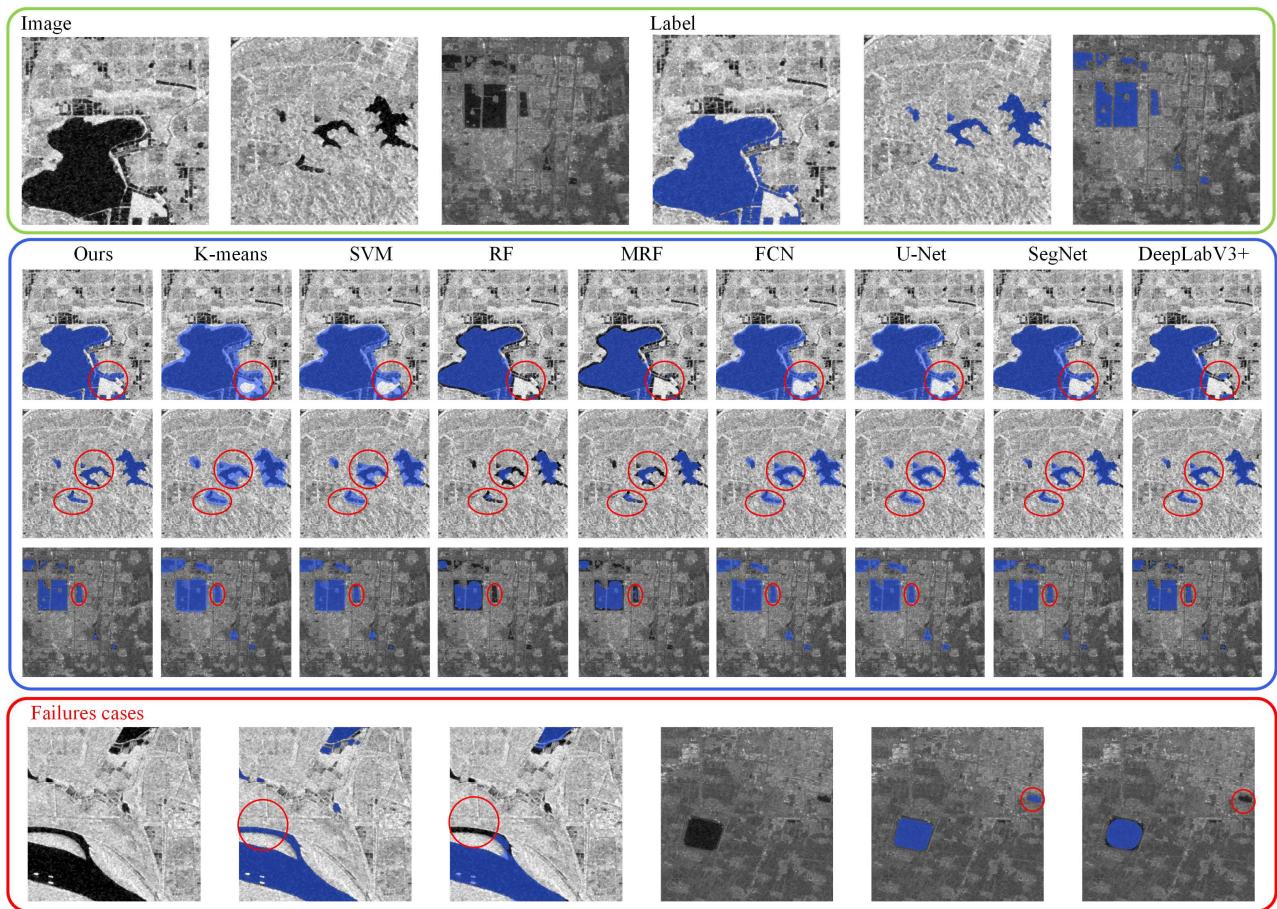


Figure 7. Visualization results of comparative methods for water-body segmentation tasks. We also provide some failure cases.

### C. Optimal Combination Selection

Table IV demonstrates the effect on the accuracy of water-body segmentation when different combinations are used as input data to DBO + DeepLabV3+. It is obvious from the results that different features combinations of input have significant differences in the performance of the model. When solely utilizing backscatter features and DEM features as inputs, namely Com. A, the resulting IoU and mIoU values are the least desirable, with indicator values of 63.90% and 65.08%, respectively. This may be due to the fact that these two features do not fully characterize the water-body, leading to errors in the model during prediction. Conversely, in Com. B-D, the indicator values are higher than in Com. A and Com. C proves to be the most effective, with the highest IoU, and mIoU indicator values, which are 86.22% and 88.24%, respectively. This indicates that the water-body in the SAR image is still mainly dominated by the surface scattering mechanism, while  $m_v$  and  $m_{dif}$  represent non-surface scattering components, which may introduce noise in water-body segmentation and lead to the performance degradation of the model. It should be noted that even when all polarimetric features are included in the model input data, that is, when Com. E is used as input, its indicator values do not surpass those of Com. C. The values of IoU and mIoU are lower than that of Com. C by 8.67% and 8.65%, respectively. This may be due to the redundancy between some polarimetric features or the negative impact of

TABLE VI  
ACCURACY COMPARISON OF DIFFERENT METHODS ON WATER-BODY SEGMENTATION TASK

Method	IoU (%)	mIoU (%)
K-means	70.17	71.43
SVM	72.45	73.48
RF	70.52	70.55
MRF	74.04	76.07
FCN	76.69	79.63
U-Net	79.57	83.01
SegNet	81.88	83.04
DeepLabV3+	84.77	86.45
DBO + DeepLabV3+	86.22	88.24

some features on the performance of the model.

Furthermore, Table V illustrates the performance of different feature combinations in the task of water-body type identification. These results bear a significant resemblance to those from the water-body segmentation task. When utilizing Com. A as inputs, the resulting accuracy is the lowest, with indicator values of 85.85%. Among Com. B-E, Com. C also demonstrates the best performance, with indicator values of 93.63%, again proving its importance in water-body studies. This finding further strengthens our understanding of the

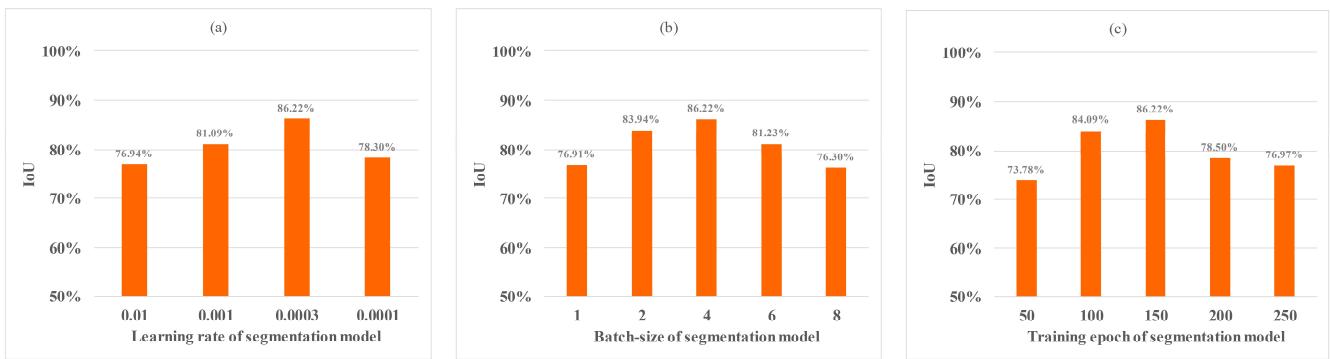


Figure 8. Parameter variation on water-body segmentation. (a) Learning rate of segmentation model. (b) Batch-size of segmentation model. (c) Training epoch of segmentation model.

dominance of the surface scattering mechanism in SAR images of water-body, and also suggests that components like  $m_v$  and  $m_{dif}$ , which represent non-surface scattering, may introduce noise in water-body type identification, leading to a decrease in model performance. Therefore, in the following experiments, only the Com. C is chosen as an input to the model.

#### D. Water-Body Segmentation

1) *Comparison to Traditional Machine Learning and State of the Art Methods:* Compare our method DBO + DeepLabV3+ with the following methods.

- 1) K-means clustering [47] posits that the proximity in "distance" between two targets correlates positively with their similarity.
- 2) Support Vector Machine (SVM) [19] aims to find an optimal hyperplane that segregates all pixels in the image to the maximum extent possible, while adhering to classification constraints, and positioning pixels at the greatest distance from the hyperplane.
- 3) Random Forest (RF) [48], predicated on a classification and regression decision tree, possesses the capacity to analyze the significance of thousands of input features.
- 4) Markov Random Field (MRF) [17] models the image as a grid, constituted by random variables, where the grayscale value of each pixel is contingent upon adjacent groups.
- 5) Full Convolutional Network (FCN) [45] supplants all the fully connected layers of traditional CNNs with convolutional layers, addressing the issue of diminished image size, a consequence of convolution and pooling, by employing transposed convolutional layers.
- 6) U-Net [49] features a U-shaped architecture, ingeniously linking the encoder with the decoder within the network structure, thereby integrating more contextual and low-dimensional information.
- 7) SegNet [50] possesses a structure in which encoders and decoders align in a one-to-one correspondence. The decoder in SegNet conducts nonlinear upsampling by directly employing the pooling index generated during the encoder's max pooling operation.
- 8) DeepLabV3+ [42] enhances the architecture of DeepLabV3 through the addition of a decoder module, aiming to optimize the network's segmentation outcomes,

particularly the delineation between individual objects. DeepLabV3+ serves as the default backbone of the proposed methodology, yielding superior performance when juxtaposed with traditional machine learning and deep learning approaches as presented in Table VI. Upon utilizing DBO for the optimization of DeepLabV3+'s hyperparameters, the model demonstrates substantial performance enhancements over other methods in the task of water-body segmentation. Specifically, it exhibits an improvement of +1.45% over DeepLabV3+ in terms of IoU metric. Therefore, it is observable that DBO not only enhances the robustness and accuracy of the model, but also demonstrates its unique strengths and potential for future development in the application of deep learning. The results for the mIoU are also delineated in Table VI. As indicated in the evaluation protocol, the values of mIoU register commonly higher compared to the values of IoU. Figure 7 presents some qualitative outcomes of the proposed method and other comparative methods, demonstrating that the prediction results generated by the proposed method are very close to the ground truth.

2) *Ablation Study:* DeepLabV3+ optimizes the hyperparameters of learning rate, batch-size, and training epoch through the DBO algorithm. If the DBO algorithm is removed, it would revert to the conventional DeepLabV3+ method, accompanied by a decrease of 1.45% relative to 86.22%. Subsequent sections will investigate the influence of these three hyperparameters on network robustness and the precision of water-body segmentation.

a) *Learning Rate:* The selection of the learning rate is a crucial factor influencing the training of deep learning models. In the context of the water-body segmentation model, the learning rate directly impacts the pace and extent of network weight updates. Theoretically, an excessively small learning rate may result in a protracted network training process, potentially causing the model to become trapped in local optima and fail to discover the global optimal solution. Conversely, an overly large learning rate, although capable of accelerating network weight updates, may induce oscillations during the training phase. This could prevent the model from converging to a stable optimal solution, or even lead to a complete failure of convergence.

Nevertheless, the conventional determination of the learning rate is often reliant on experience, with typical values being

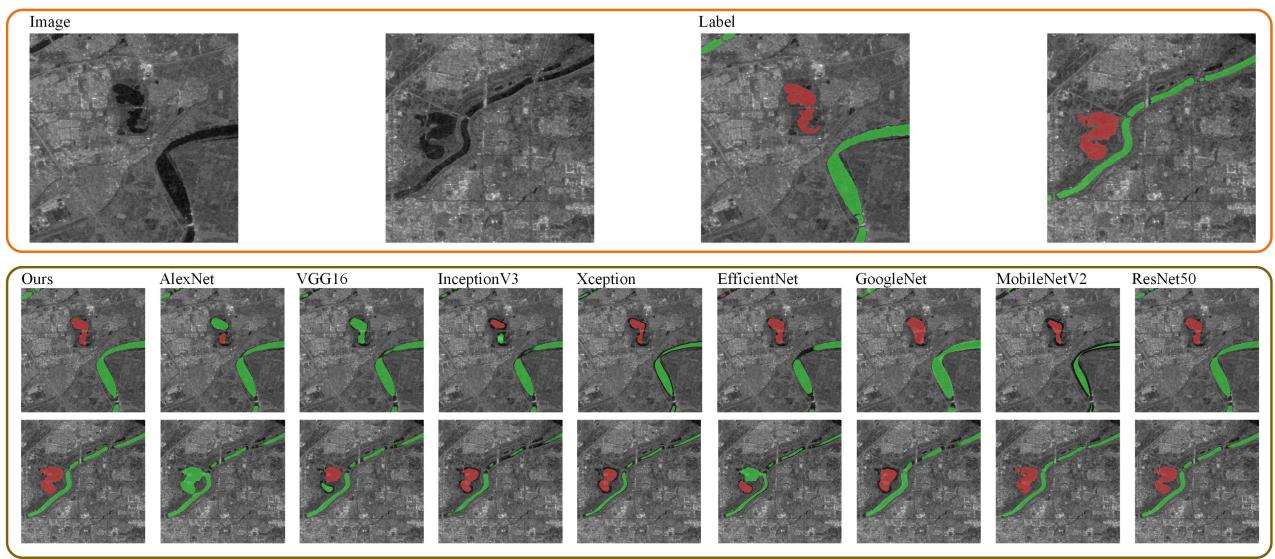


Figure 9. Visualization results of comparative methods for water-body type identification tasks. Ground truth is also provided. Red pixel denotes artificial water-body, while green pixel denotes natural water-body.

0.01, 0.001, or 0.0001. While this way is straightforward, it lacks adaptability to specific tasks and data. In the conducted experiment, Figure 8(a) displays a comparison of model performance using conventional empirical values and the optimal learning rate calculated by DBO iteration. To maintain a fair comparison, other hyperparameters (batch-size and training epoch) are held constant. When the learning rate is set to 0.01, the IoU value is at its lowest, at 76.94%. This suggests that the step size for updating network weights may be too large, potentially causing the model to overlook crucial local optima during the search for the optimal solution, thereby compromising model performance. Furthermore, a decrease in the learning rate to 0.001 results in a slight improvement in IoU, but this metric still falls short of the desired level. This could be attributed to the continued large step size for updating network weights, which may cause gradients to oscillate around the global optimum, preventing accurate convergence to the optimal solution. Moreover, a further reduction in the learning rate to 0.0001 leads to a decrease in IoU, reaching 78.30%. Finally, when the learning rate is adjusted to 0.0003, IoU reaches its peak value of 86.22%. This might be because a learning rate of 0.0003 offers an effective balance, ensuring stable updates of network weights without the slow training speed associated with a too-small learning rate. As such, at a learning rate of 0.0003, the model can effectively strike a balance between exploring the parameter space and avoiding overfitting, thereby achieving optimal model performance.

b) *Batch-Size*: Batch-size is another crucial hyperparameter, directly influencing the efficiency and stability of model learning. Primarily, the batch-size determines the quantity of samples used in each model weight update. An excessively small batch-size could induce instability in the model during training, as each update utilizes a very limited number of samples, potentially leading to high variance in gradient estimation. Moreover, smaller batch-size imply a higher number of update iterations, which could extend training

TABLE VII  
ACCURACY COMPARISON OF DIFFERENT METHODS ON WATER-BODY TYPE IDENTIFICATION TASK

Method	ACC (%)	Parameters
AlexNet	69.87	61M
VGG16	73.87	138M
InceptionV3	74.25	23.9M
Xception	78.97	22.9M
EfficientNet	82.31	5.31M
GoogleNet	85.14	7M
MobileNetV2	86.96	3.5M
ResNet18	89.17	11.7M
ResNet50	92.97	25.6M
ResNet101	90.85	44.6M
DBO + ResNet50	95.71	44.6M

time. However, an overly large batch-size presents its own potential issues. While it can reduce training time and provide more accurate gradient estimates, an excessively large batch-size might also limit the model's ability to explore the parameter space, as each update utilizes a large number of samples, potentially causing the model to become trapped in local optima too.

The influence of different batch-size on model performance is studied through an ablation experiment. The experimental results, as shown in Figure 8(b), further substantiate that optimizing the batch-size allows DeepLabV3+ to achieve optimal performance in water-body segmentation tasks. When the batch-size is set to 1, the IoU value is at its lowest, at 76.91%. This could be due to the fact that a smaller batch-size lead to unstable gradient updates. As the batch-size increases to 2, there is an improvement in the IoU value, reaching 83.94%. When the batch-size is further increased to 4, the IoU value is at its highest. This suggests that a batch-size of 4 achieves an

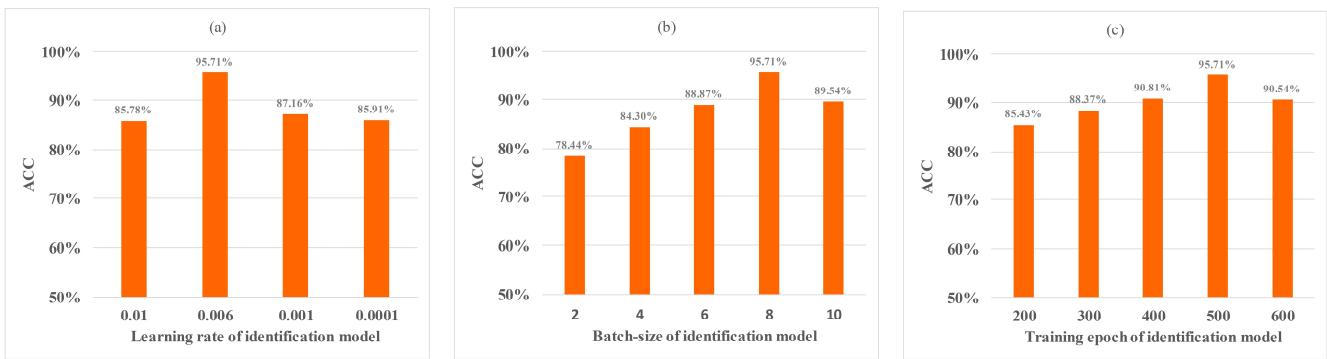


Figure 10. Parameter variation on water-body type identification. (a) Learning rate of identification model. (b) Batch-size of the identification model. (c) Training epoch of identification model.

effective balance between computational efficiency and the accuracy of gradient estimation, thereby leading to optimal model performance. However, further increasing the batch-size to 6 and 8 results in a significant decrease in the IoU value. This could be because larger batch-sizes, while providing more accurate gradient estimates, may also lead to a decrease in the model's generalization ability, as the larger batch-size reduces the model's exposure to the diversity of the training data.

*c) Training Epoch:* The training epoch is another key hyperparameter that needs to be carefully considered. Each training epoch represents one complete forward and backward propagation of the model over the entire training set. Thus, the training epoch determines how many times the model learns from the training data. Too few training epoch may result in the model not having enough opportunities to learn advanced features of the data, and as a result, the model will perform poorly. However, too many training epoch may cause the model to overfit on the training set, resulting in poor performance on new, unseen data. When determining the optimal number of training epochs, factors such as training duration, model complexity, and the model's capacity to generalize on unseen data must be carefully balanced.

In the ablation experiments, again, the effect on model performance is observed by setting multiple general iteration numbers. As shown in Figure 8(c), the epoch computed by the iterations of the DBO algorithm has the highest performance metrics. Too many or too few training epochs lead to performance degradation. When the training epoch is set to 50, the IoU value is lowest, at 73.78%. This could be due to the insufficient number of epochs, which may not provide the model ample opportunity to learn advanced features from the data, resulting in suboptimal performance. As the training epoch increases to 100, the IoU value improves to 84.09%. However, it is when the training epoch is further increased to 150 that the IoU value reaches its highest, at 86.22%. This suggests that at 150 epochs, an effective balance between computational efficiency and training duration is achieved. A further increase in the training epochs to 200 and 250 results in a decline in the IoU value, to 78.50% and 76.97% respectively. This decline could be attributed to overfitting, leading to poor performance on new data.

#### E. Water-Body Types Identification

*1) Comparison to State of the Art Methods:* The proposed model, namely DBO + ResNet50, is contrasted with the following classification models for comparative analysis.

- 1) AlexNet [51], a deep convolutional neural network, comprises five convolutional layers along with three fully connected layers. This network is notable for its introduction of ReLU activation, overlapping pooling, and dropout techniques.
- 2) VGG16 [51], is structured with 13 convolutional layers and 3 fully connected layers. Noteworthy is its utilization of  $3 \times 3$  small filters and  $2 \times 2$  maximum pooling layers.
- 3) InceptionV3 [52] incorporates batch normalization and factorization into its inception architecture. This integration enhances both computational efficiency and the accuracy of the network.
- 4) Xception [53], an extension of the inception architecture, supplants the standard inception block with depthwise separable convolutions.
- 5) EfficientNet [54] introduces a unique scaling method designed to optimize the balance between network depth, width, and resolution.
- 6) GoogleNet [55] successfully enhances the width and depth of the network by parallelizing convolution and pooling operations across different scales.
- 7) MobileNetV2 [56], in comparison to its predecessor MobileNet, introduces a linear bottleneck alongside a reverse residual structure.
- 8) ResNet50 [44], incorporates a residual module designed to address the challenges of gradient vanishing and representation bottlenecks in deep networks, achieved through the application of a short-circuit mechanism.

The visualization results of water-body type identification for the contrasting methods are shown in Figure 9. In Table VII, it is evident that ResNet50, serving as the backbone network, outperforms other comparative methods, ACC is 92.97%, which demonstrates superior efficacy. Upon optimization of ResNet50's hyperparameters utilizing DBO, a notable enhancement in performance is observed, the ACC is 95.71%, which achieves the highest accuracy. Moreover, the ACC of ResNet18 and ResNet101 is 89.17% and 90.85%, respectively. This not only attests to the appropriateness of selecting

TABLE VIII  
PERFORMANCE ON OVERALL WATER-BODY TYPE CLASSIFICATION

Method	DBO (Segmentation)	DBO (Classification)	IoU <sup>n</sup> (%)	IoU <sup>a</sup> (%)	mIoU <sup>na</sup> (%)	mIoU <sup>nab</sup> (%)
Proposed method	×	×	80.74	85.96	83.85	86.28
	✓	✓	85.12	87.05	86.09	92.98
MultiClass-DeepLabV3+	n/a	n/a	70.43	73.75	72.09	77.01

ResNet50 as the backbone network, but also underscores the effectiveness of DBO once again. Additionally, the enumeration of learnable parameters inherent to each network is documented in the Table VII, too. It should be emphasized that since DBO only optimizes hyperparameters, it does not add any additional learnable parameters to the network.

*2) Ablation Study:* When using the ResNet50 network for water-body type identification tasks, ablation experiments similar to water-body segmentation are still performed. This is because of the fact that ResNet50, too, optimizes the three hyperparameters, learning rate, batch-size, and training epoch, through the application of the DBO algorithm. The role and influence of the learning rate, batch-size, and training epoch have already been delineated in the context of utilizing DeepLabV3+ for water-body segmentation tasks, thus rendering a repetition of these details unnecessary here. The performance changes of the model caused by the ablation experiments of the three hyperparameters are shown in Figure 10.

*a) Learning Rate:* As shown in Figure 10(a), when the learning rate is set to 0.01, 0.001, 0.006, and 0.0001, the model's accuracy is 85.78%, 95.71%, 87.16%, and 85.91%, respectively. Due to the substantial depth of the ResNet50 network, subtle changes in the learning rate can significantly affect the network's convergence rate and ultimate performance. Results demonstrate that the model's accuracy reaches its highest when the learning rate is 0.006. The optimized learning rate, while ensuring the depth of the network, effectively avoids overfitting, thereby achieving the best performance in the task of water-body type identification.

*b) Batch-Size:* As shown in Figure 10(b), when the batch-size is set to 2, 4, 6, 8, 10, the model's accuracy is 78.44%, 84.30%, 88.87%, 95.71%, and 89.54%, respectively. The adjustment of batch-size influences the network's ability to capture water-body type features to a certain extent. Results show that when the batch-size is 8, the model's accuracy reaches its peak. The optimized batch-size allows the ResNet50 network to better balance memory usage and model performance when training on large-scale water body type identification tasks, thereby achieving a higher accuracy rate in water-body type identification.

*c) Training Epoch:* As shown in Figure 10(c), when the training epoch is set to 200, 300, 400, 500, 600, the model's accuracy is 85.43%, 88.37%, 90.81%, 95.71%, and 90.54%, respectively. The length of the training epoch somewhat reflects the model's depth of understanding of water-body type features. Results show that when the training epoch is 500, the model's accuracy reaches its peak. The optimized training epoch allows the ResNet50 network to ensure learning efficiency while fully

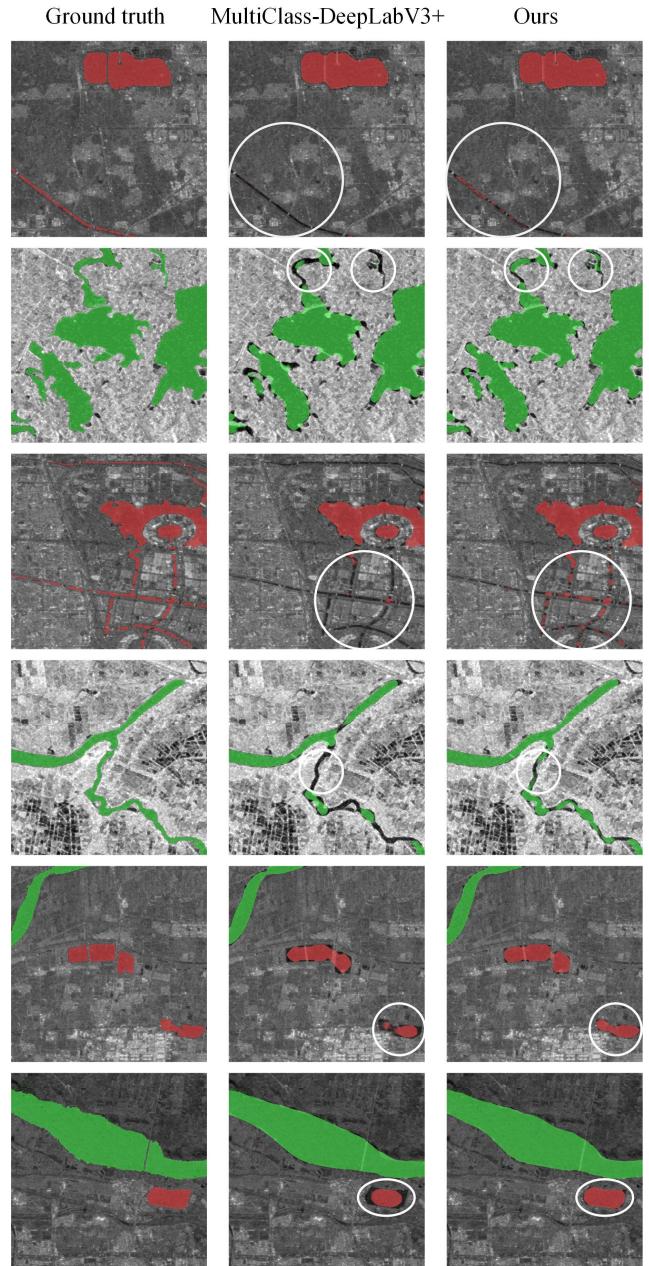


Figure 11. Visualization of water-body type classification results using our proposed method and MultiClass-DeepLabV3+. Ground truth is also provided. Red pixel denotes artificial water-body, while green pixel denotes natural water-body.

extracting and learning the complex features of water-body types, thereby achieving the best performance in the task of

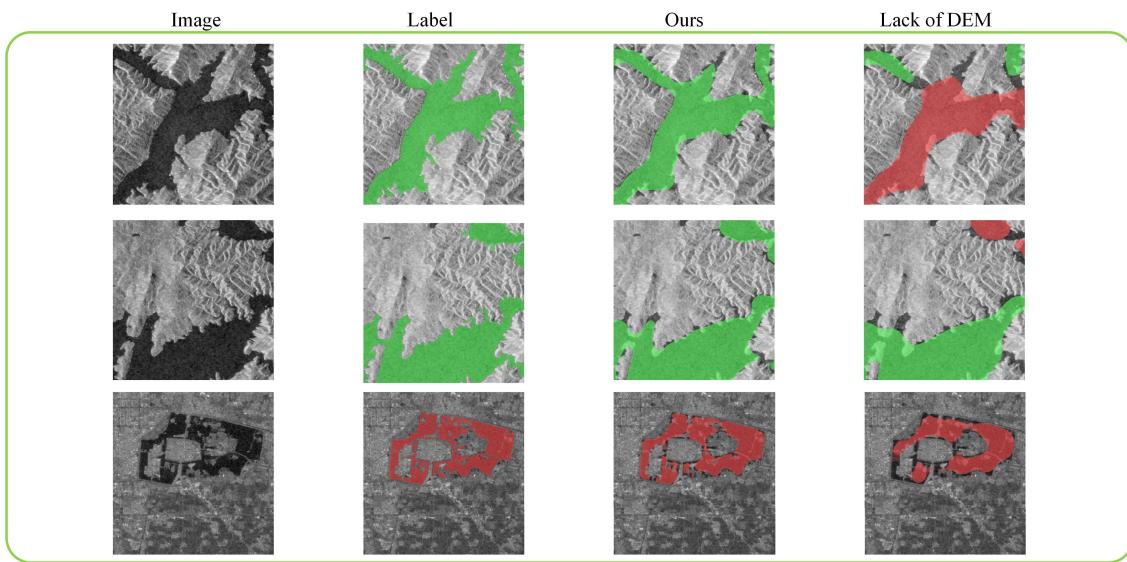


Figure 12. Visualization of water-body type classification results using our proposed method and lack of DEM feature. The SAR image and ground truth is also provided. Red pixel denotes artificial water-body, while green pixel denotes natural water-body.

water-body type identification.

#### F. Water-Body Types Classification

In reference to Section V-B, the  $\text{IoU}^n$ ,  $\text{IoU}^a$ ,  $\text{mIoU}^{na}$ , and  $\text{mIoU}^{nab}$  of the entire water-body type classification pipeline are reported in Table VIII.  $\text{mIoU}^{na}$  is conducted equal measurements on the segmentation accuracy of natural and artificial water-body, resulting in a value of 86.09%. This value is marginally lower than the  $\text{IoU}$  value of DBO + DeepLabV3+ in Table VI, underscoring the robustness of the water-body classifier. Ablation studies furnished in Table VIII substantiate the effectiveness of DBO optimization, too. For instance, when DBO is not utilized, the pipeline is downgraded to a combination of DeepLabV3+ and ResNet50. The  $\text{mIoU}^{na}$  and  $\text{mIoU}^{nab}$  generated in this scenario are lower by 2.24% and 6.7%, respectively, compared to the full version (i.e., 86.09% and 92.98%). The proposed method addresses the water-body type classification task in two stages: the acquisition of water-body in the first stage, followed by their identification as natural or artificial water-body in the subsequent stage. To validate this design, comparative experiments are provided. The proposed method will be compared with a single-stage end-to-end multi-class semantic segmentation network. A multi-class DeepLabV3+ network (MultiClass-DeepLabV3+) is employed to forecast pixel-level labels directly, namely predict as either natural water-body, artificial water-body, and backgrounds. The performance of MultiClass-DeepLabV3+, as showcased in Table VIII, is markedly inferior to that of the two-stage pipeline. As delineated in Section I, a single-stage multi-class segmentation network experiences suboptimal performance due to the minimal pixel-level distinction between natural and artificial water-body. Figure 11 presents qualitative results from the two-step pipeline and MultiClass-DeepLabV3+. When juxtaposed with the ground truth, the results derived from the two-stage pipeline exhibit a considerable performance

enhancement over MultiClass-DeepLabV3+.

## VI. DISCUSSION

In this section, the effectiveness and performance of the proposed method are extensively discussed, along with the impact of key parameter variations on model performance. Initially, comparative experiments are conducted to reveal the significance of DEM feature on outcomes. Subsequently, adjustments are made to several key parameters within DBO, aiming to analyze their influence on the loss value of the fitness function. Finally, a detailed analysis is carried out regarding the strengths and weaknesses of the proposed method, followed by a projection of potential directions for future improvements.

#### A. The Influence of DEM Features on Model Performance

To further substantiate the efficacy of the proposed method, comparative experiments are carried out. The DEM feature is removed from the input data to observe its influence on the results. As depicted in Figure 12, the model's performance in identifying water-body is notably superior in mountainous regions compared to flat terrain. This could be due to the complex terrain in mountainous regions, where the large variations in DEM features aid the model in more accurately extracting and conducting contextual learning. Moreover, the overlay shadow of mountainous regions in SAR data could potentially influence the results. Without the DEM features, the model might erroneously recognize the shadow as a water-body. However, the addition of DEM features reduces the impact of shadows on the model. Therefore, the integration of DEM features allows the model to distinguish more accurately between water-body and shadows, further enhancing the model's capability to identify water-body types in mountainous terrain. However, in urban areas, the improvement in the model's performance for water-body type identification is not as pronounced. This could be attributed to the relatively flat terrain in urban areas, where the changes in DEM features are minimal, thus having a smaller impact on the model's

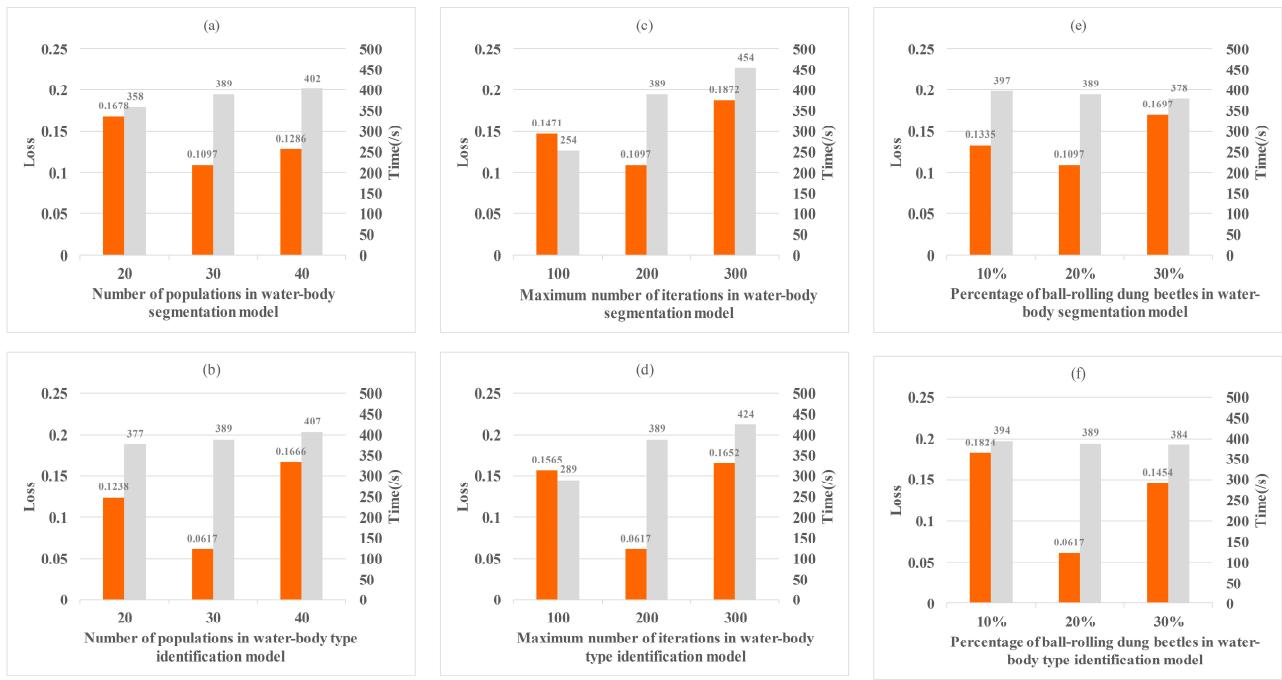


Figure 13. Parameter variation on DBO. (a) Number of populations in water-body segmentation model. (b) Number of populations in water-body type identification model. (c) Maximum number of iterations in water-body segmentation model. (d) Maximum number of iterations in water-body type identification model. (e) Percentage of ball-rolling dung beetles in water-body segmentation model. (f) Percentage of ball-rolling dung beetles in water-body type identification model.

recognition accuracy. The significant advantage of DEM features in water-body types identification in mountainous regions further corroborates the superiority of the proposed method, namely, the enhancement of water-body type classification accuracy in different terrains through the fusion of multi-source data.

#### B. The Influence of Changes in Parameters in DBO

Additionally, some key parameters of DBO are altered separately in the segmentation and identification networks. This is done to observe their impact on the loss value of the fitness function. These parameters include the number of populations, the maximum number of iterations, and the proportion of ball-rolling dung beetles. First, the number of populations in the segmentation network changes from the original setting of 30 to 20 and 40, which is to explore the effect of a number of population changes on the loss value. As shown in Figure 13(a), when the number of population is 20, the loss value increases. However, when the number of population is 40, the loss value decreases, but the computation time also increases. This result indicates that increasing the number of population can improve the model's performance but also increase the computation time. The same experiment is validated in the identification network. As shown in Fig. 13(b), the trend of the loss value of the identification network is similar to that of the segmentation network. Next, the maximum number of iterations changes from the original setting of 200 to 100 and 300. This is done to reveal its impact on the model's performance. As illustrated in Figures 13(c) and 13(d), whether in the segmentation network or identification network, when the maximum number of iterations is 100, the loss value tends to increase. However, when the maximum number of iterations is 300, the loss value

decreases, but the computation time also increases. Finally, the percentage of ball-rolling dung beetles is adjusted from the original setting of 20% to 10% and 30% to explore its effect on the model performance. As shown in Figures 13(e) and 13(f), regardless of the segmentation network or the identification network, the loss value tends to rise. This suggests that maintaining an appropriate proportion of ball-rolling dung beetles is crucial for the robustness of the model. Based on these experiments, it becomes apparent that by properly adjusting the key parameters of DBO, an ideal balance can be found between model performance and computational burden.

#### C. Advantages, Disadvantages, and Future Prospects

The main advantage of the proposed method is that it uses a combination of deep learning and optimization algorithms, as well as the fusion of multi-source data. And innovatively divide the water-body type classification task into two stages to solve the problem. First, two deep learning models, DeepLabV3+ and ResNet50, are adopted, both demonstrating superior performance in image segmentation and classification tasks. Second, DBO is introduced for hyperparameter optimization, a step that notably enhanced the performance and generalization ability of the model. Furthermore, backscatter features, polarimetric features, and DEM features are also fused. The fusion of these features enables the model to comprehensively understand the characteristics of water-body from multiple perspectives, thereby improving the accuracy of water-body type classification.

Although this article has achieved certain results, there are still some limitations. First, the ambiguity of the boundary between water-body and land, and the pixel-level similarity between natural water-body and artificial water-body are still

relatively large obstacles. Second, proposed method may face the problem of excessive computational burden when dealing with large-scale image data.

In view of the above deficiencies, the following directions for future improvement are considered. Firstly, the exploration of additional feature fusion and optimization techniques can enhance the model's ability to identify specific types of water-body. Secondly, the investigation of more efficient computational methods and model structures can reduce the model's computational burden and enable it to handle larger scale image data.

Overall, the method proposed in this article shows superior performance on the task of water-body type classification, but there is still room for improvement. Anticipation is held for further enhancement of the model's performance and provision of more effective solutions for water body type identification through the introduction of advanced techniques and methods in future research.

## VII. CONCLUSION

This article presents a novel approach to water-body types identification based on two-step deep learning model, supplemented by DBO for network hyperparameter optimization. In this article, the effectiveness of the proposed method in classifying between natural and artificial water-body is investigated. The results underscore the potential of deep learning combined with DBO optimization to address challenges in this community. In the first step, the DeepLabV3+ network is used for feature extraction and accomplished precise water-body segmentation. In the second step, the ResNet50 network is employed for classifying the segmented water-body into natural or artificial.

Furthermore, the DBO is employed to optimize hyperparameters within the deep learning network models. Through the tuning of parameters such as learning rate, batch-size, and training epochs, the goal is to improve the model's performance and its ability to generalize. In addition, a strategy is proposed for fusing backscatter features, polarimetric features of SAR images, and DEM feature for input.

The results of the experiment confirmed the effectiveness of our proposed method. In particular, the DBO optimization and the fusion of backscatter, polarimetric, and DEM features significantly improved the accuracy of water-body type classification, especially in complex terrain conditions. This research not only contributes to the field of water-body type classification but also demonstrates the value of deep learning and DBO in advancing remote sensing technology.

This article sets a foundation for future work in this community. As water resources become increasingly crucial in the face of climate change and urban development, it is believed that the techniques developed in this article will have broad and significant implications for water resource management, ecological protection, and urban planning.

## REFERENCES

- [1] B. K. R. Kadapala and A. H. K., "Region-Growing-Based Automatic Localized Adaptive Thresholding Algorithm for Water Extraction Using Sentinel-2 MSI Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-8, 2023.
- [2] Q. Liu, Y. Tian, L. Zhang and B. Chen, "Urban Surface Water Mapping from VHR Images Based on Superpixel Segmentation and Target Detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 5339-5356, 2022.
- [3] P. Nie, X. Cheng, Z. Song, M. Mao, T. Wang and L. Meng, "Rethinking BiSeNet: A Lightweight Network for Urban Water Extraction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-10, 2023.
- [4] B. C. Jr, C. E. Robinson, M. J. Hinton, and H. A.J. Russell, "A framework for conceptualizing groundwater-surface water interactions and identifying potential impacts on water quality, water quantity, and ecosystems," *Journal of Hydrology*, vol. 574, pp. 609-627, 2019.
- [5] S. Yao, C. Chen, Q. Chen, J. Zhang and M. He, "Combining process-based model and machine learning to predict hydrological regimes in floodplain wetlands under climate change," *Journal of Hydrology*, pp. 130193, 2023.
- [6] Y. Zhuo, Z. Wu, H. Xu, D. Yan, M. Jiang, X. Zhang, H. Wang, "Adaptive selection and optimal combination scheme of candidate models for real-time integrated prediction of urban flood," *Journal of Hydrology*, pp. 130152, 2023.
- [7] X. Li, X. Ye, C. Yuan, C. Xu, "Can water release from local reservoirs cope with the droughts of downstream lake in a large river-lake system?" *Journal of Hydrology*, vol. 625, pp. 130172, 2023.
- [8] W. Feng, H. Sui, W. Huang, C. Xu and K. An, "Water Body Extraction From Very High-Resolution Remote Sensing Imagery Using Deep U-Net and a Superpixel-Based Conditional Random Field Model," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 4, pp. 618-622, 2019.
- [9] M. Lu, L. Fang, M. Li, B. Zhang, Y. Zhang and P. Ghamisi, "NFANet: A Novel Method for Weakly Supervised Water Extraction From High-Resolution Remote-Sensing Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-14, 2022.
- [10] Z. Miao, K. Fu, H. Sun, X. Sun and M. Yan, "Automatic Water-Body Segmentation From High-Resolution Satellite Images via Deep Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 4, pp. 602-606, 2018.
- [11] M. Lu, L. Fang, M. Li, B. Zhang, Y. Zhang and P. Ghamisi, "NFANet: A Novel Method for Weakly Supervised Water Extraction From High-Resolution Remote-Sensing Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-14, 2022.
- [12] Y. Ren, X. Li, X. Yang and H. Xu, "Development of a Dual-Attention U-Net Model for Sea Ice and Open Water Classification on SAR Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [13] M. Helleis, M. Wieland, C. Krullkowski, S. Martinis and S. Plank, "Sentinel-1-Based Water and Flood Mapping: Benchmarking Convolutional Neural Networks Against an Operational Rule-Based Processing Chain," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 2023-2036, 2022.
- [14] N. Li, S. Niu, Z. Guo, L. Wu, J. Zhao, L. Min, D. Ge, J. Chen, "Dynamic Waterline Mapping of Inland Great Lakes Using Time-Series SAR Data From GF-3 and S-1A Satellites: A Case Study of DJK Reservoir, China," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 11, pp. 4297-4314, 2019.
- [15] W. Tan, J. Li, L. Xu and M. A. Chapman, "Semiautomated Segmentation of Sentinel-1 SAR Imagery for Mapping Sea Ice in Labrador Coast," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 5, pp. 1419-1432, 2018.
- [16] J. Zhang, M. Xing, G. Sun, J. Chen, M. Li, Y. Hu, Z. Bao, "Water Body Detection in High-Resolution SAR Images With Cascaded Fully-Convolutional Network and Variable Focal Loss," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 1, pp. 316-332, 2021.
- [17] H. Deng and D. A. Clausi, "Unsupervised segmentation of synthetic aperture Radar sea ice imagery using a novel Markov random field model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 528-538, 2005.
- [18] Y. Le, N. Li, "Improved Change Detection Method for Flood Monitoring," *Journal of Radars*, vol. 6, no. 2, pp. 204-212, 2017.
- [19] W. Lv, Q. Yu and W. Yu, "Water extraction in SAR images using GLCM and Support Vector Machine," *IEEE 10th international conference on signal processing proceedings*. 2010.
- [20] J. Liang, D. Liu, "A local thresholding approach to flood water delineation using Sentinel-1 SAR imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 53-62, 2020.
- [21] N. Li, Z. Guo, J. Zhao, L. Wu and Z. Guo, "Characterizing Ancient Channel of the Yellow River From Spaceborne SAR: Case Study of Chinese Gaofen-3 Satellite," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, no. 1502805, pp. 1-5, 2022.

- [22] Y. Ren, X. Li, X. Yang and H. Xu, "Development of a Dual-Attention U-Net Model for Sea Ice and Open Water Classification on SAR Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, no. 4010205, pp. 1-5, 2022.
- [23] Y. Gao, F. Gao, J. Dong and S. Wang, "Transferred Deep Learning for Sea Ice Change Detection From Synthetic-Aperture Radar Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 10, pp. 1655-1659, 2019.
- [24] X. Huang, C. Xie, X. Fang and L. Zhang, "Combining Pixel- and Object-Based Machine Learning for Identification of Water-Body Types From Urban High-Resolution Remote-Sensing Imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 5, pp. 2097-2110, 2015.
- [25] L. Xie, H. Zhang, C. Wang and F. Chen, "Water-Body types identification in urban areas from radarsat-2 fully polarimetric SAR data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 50, pp. 10-25, 2016.
- [26] J. Wei, M. Wang, K. Mikelsons, L. Jiang, S. Kratzer, Z. Lee, T. Moore, H. M. Sosik and D. V. Zande, "Global satellite water classification data products over oceanic, coastal, and inland waters," *Remote Sensing of Environment*, vol. 282, pp. 113233, 2022.
- [27] S. Qing, T. Cui, Q. Lai, Y. Bao, R. Diao, Y. Yue and Y. Hao, "Improving remote sensing retrieval of water clarity in complex coastal and inland waters with modified absorption estimation and optical water classification using Sentinel-2 MSI," *International Journal of Applied Earth Observation and Geoinformation*, vol. 102, pp. 102377, 2021.
- [28] N. Abid, M. Shahzad, M. I. Malik, U. Schwancke, A. Ulges, G. Kovács and F. Shafait, "UCL: Unsupervised Curriculum Learning for water body classification from remote sensing imagery," *International Journal of Applied Earth Observation and Geoinformation*, vol. 105, pp. 102568, 2021.
- [29] Y. Li, B. Dang, Y. Zhang and Z. Du, "Water body classification from high-resolution optical remote sensing imagery: Achievements and perspectives," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 187, 2022, pp. 306-327, 2022.
- [30] N. Nasir, A. Kansal, O. Alshalte, F. Barneih, A. Shanableh, M. Al-Shabi and A. A. Shammaa, "Deep learning detection of types of water-bodies using optical variables and ensembling," *Intelligent Systems with Applications*, vol. 18, pp. 200222, 2023.
- [31] D. Marmanis, M. Datcu, T. Esch and U. Stilla, "Deep Learning Earth Observation Classification Using ImageNet Pretrained Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 105-109, 2016.
- [32] E. Maggiori, Y. Tarabalka, G. Charpiat and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Transactions on geoscience and remote sensing*, vol. 55, no.2, pp. 645-657, 2016.
- [33] X. Lyu, W. Jiang, X. Li, Y. Fang, Z. Xu and X. Wang, "MSAFNet: Multiscale Successive Attention Fusion Network for Water Body Extraction of Remote Sensing Images," *Remote Sensing*, vol.15, no.12 pp. 3121, 2023.
- [34] T. Shi, Z. Guo, C. Li, X. Lan, X. Gao, X. Yan, "Improvement of deep learning Method for water body segmentation of remote sensing images based on attention modules," *Earth Science Informatics*, pp. 1-12, 2023.
- [35] D. Yuan, C. Wang, L. Wu, X. Yang, Z. Guo, X. Dang, J. Zhao and N Li, "Water Stream Extraction via Feature-Fused Encoder-Decoder Network Based on SAR Images," *Remote Sensing*, vol.15, no. 6, pp. 1559, 2023.
- [36] W. Fang, C. Wang, X. Chen, W. Wan, H. Li, S. Zhu, Y. Fang, B. Liu and Y. Hong, "Recognizing global reservoirs from Landsat 8 images: A deep learning approach," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 9, pp. 3168-3177, 2019.
- [37] R. Balaniuk, O. Isupova, S. Reece, "Mining and tailings dam detection in satellite imagery using deep learning," *Sensors*, vol. 20, no. 23, pp. 6936, 2020.
- [38] M.E. Malerba, N. Wright and P. I. Macreadie, "A continental-scale assessment of density, size, distribution and historical trends of farm dams using deep learning convolutional neural networks," *Remote Sensing*, vol. 13, no. 2, pp. 319, 2021.
- [39] R. Nagaraj and L. S. Kumar, "Surface water body extraction and Change Detection Analysis using Machine Learning Algorithms: A Case study of Vaigai Dam, India," *IEEE International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT)*, 2023.
- [40] L. Mascolo, J. M. Lopez-Sanchez and S. R. Cloude, "Thermal noise removal from polarimetric Sentinel-1 data," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2021.
- [41] L. Mascolo, S. R. Cloude and J. M. Lopez-Sanchez, "Model-based decomposition of dual-pol SAR data: Application to Sentinel-1," 2021.
- [42] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," *Proceedings of the European conference on computer vision (ECCV)*, 2018.
- [43] T. Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, "Focal loss for dense object detection," *Proceedings of the IEEE international conference on computer vision*, 2017.
- [44] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.
- [45] J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015.
- [46] J. Xue, B. Shen, "Dung beetle optimizer: A new meta-heuristic algorithm for global optimization," *The Journal of Supercomputing*, vol. 79, no. 7, pp. 7305-7336, 2023.
- [47] Z. Liu, F. Li, N. Li, R. Wang and H. Zhang, "A novel region-merging approach for coastline extraction from sentinel-1A IW mode SAR imagery," *IEEE Geoscience and remote sensing letters*, vol. 13, no .3, pp. 324-328, 2016.
- [48] L. Xie, H. Zhang and C. Wang, "Water-body types classification using Radarsat-2 fully polarimetric SAR data," *2015 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology (ICARES)*, 2015.
- [49] Z. Zhang, Q. Liu, Y. Wang, "Road extraction by deep residual u-net," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 749-753, 2018.
- [50] V. Badrinarayanan, A. Kendall and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481-2495, 2017.
- [51] W. Yu, K. Yang, Y. Bai, T. Xiao, H. Yao and Y. Rui, "Visualizing and comparing AlexNet and VGG using deconvolutional layers," *Proceedings of the 33 rd International Conference on Machine Learning*, 2016.
- [52] X. Xia, C. Xu, B. Nan, "Inception-v3 for flower classification," *2017 2nd international conference on image, vision and computing (ICIVC)*, 2017.
- [53] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.
- [54] L. T. Duong, P. T. Nguyen, C. D. Sipio and D. D. Ruscio, "Automated fruit recognition using EfficientNet and MixNet," *Computers and Electronics in Agriculture*, vol. 171, pp. 105326, 2020.
- [55] R. Anand, T. Shanthi, M. S. Nithish and S. Lakshman, "Face recognition and classification using GoogleNET architecture," *Soft Computing for Problem Solving: SocProS 2018*, vol. 1, Springer Singapore, 2020.
- [56] Q. Xiang, X. Wang, R. Li, G. Zhang, J. Lai and Q. Hu, "Fruit image classification based on Mobilenetv2 with transfer learning technique," *Proceedings of the 3rd international conference on computer science and application engineering*, 2019.



**Qiming Yuan** was born in Henan, China, in 1998. He received the B.S. degree in College of Information Technology from Shanghai Jian Qiao University, Shanghai, China, in 2016, where he is currently pursuing the M.S. degree with the School of Computer and Information Engineering from Henan University, Henan, China. His research interests include synthetic aperture radar (SAR) image processing and deep learning.



Jing Lu received the B. S. degree from Nankai University, Tianjin, China, in 2008, the M. S. degree from Beijing Normal University, Beijing, China, in 2013. In June 2014, she worked on remote sensing applications in Land Satellite Remote Sensing Application Center, Ministry of Natural Resources, Beijing, China.



Lin Wu (M'18) received the B.S. degree in Electrical Technology Education from Henan University, Kaifeng, China, in 2001, and the M.S. degree in Applied Mathematics from Henan University, Kaifeng, China, in 2007.

Since 2001, she has been a teacher at the College of Computer and Information Engineering, Henan University, Kaifeng, where she was an associate professor in 2015. Her research interests include intelligent processing algorithm, SAR image processing techniques, and SAR image application of

water environment.



**Yabo Huang** received the B.S. degree in computer science and technology from Henan University, Kaifeng, China, in 2004, and the M.S. degree in Applied Mathematics from Henan University, Kaifeng, China, in 2007, and the doctor of science degree in Cartography and geographical information system from Henan University, Kaifeng, China, in 2017. Since July 2007, she has been a teacher at the College of Computer and Information Engineering, Henan University, Kaifeng, where she was an associate professor in 2020. Her research interests include intelligent processing algorithm, SAR and optical image processing techniques and water environment.



**Zhengwei Guo** (M'19) received her B.S. degree in radio technology engineering from Huazhong University of Science and Technology, Wuhan, China, in 1984 and then joined Henan University, Kaifeng, China, as an Assistant Professor.

Since 2008, Zhengwei Guo served as a professor in School of Computer and Information Engineering of Henan University. Her research focus on management system software, data processing system, and information security technology. Her projects on SAR image applications and information processing system

have been supported by National Natural Science Foundation of China and Department of Science and Technology of Henan Province. Guo has published several papers including peer reviewed journals, conference proceedings, patents, software copyrights, and received the First-Prize Scientific and Technological Progress Award of Henan Province.



**Ning Li** (S'15-M'16) received the B.S. degree in electronics information engineering from Northeast Forestry University, Harbin, China, in 2009, the M.S. degree in communication and information system from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2012, and the Ph.D. degree at the Institute of Electronics, Chinese Academy of Sciences (IECAS), Beijing, China, in 2015. In July 2015, he joined the Department of Space Microwave Remote Sensing System, IECAS, Beijing, China, where he was an assistant professor. Since

December 2017, he has been a full professor with the School of Computer and Information Engineering, Henan University, Kaifeng, China. His research interests include synthetic aperture radar (SAR) and inverse SAR imaging algorithms and autofocusing techniques, SAR polarimetric theory, and SAR image processing. Dr. Li was a recipient of the Special Prize of President Scholarship for Postgraduate Students from the University of Chinese Academy of Sciences in 2015.