Article

DaliWS: A High-Resolution Dataset with Precise Annotations for Water Segmentation in Synthetic Aperture Radar Images

Shanshan Zhang [1], Weibin Li [1,2,*], Rongfang Wang [1], Chenbin Liang [2], Xihui Feng [3] and Yanhua Hu [4]

1   School of Artificial Intelligence, Xidian University, Xi'an 710071, China; 21171213865@stu.xidian.edu.cn (S.Z.); rfwang@xidian.edu.cn (R.W.); liangchenbin@xidian.edu.cn (C.L.)
2   Laboratory of Artificial Intelligence, Hangzhou Institute of Technology of Xidian University, Hangzhou 311231, China; liangchenbin@xidian.edu.cn
3   Key Laboratory of Coal Resources Exploration and Comprehensive Utilization, Ministry of Natural Resources, Xi'an 710021, China; fengxihui789@gmail.com
4   Department of Water Resources of Shaanxi Province, Xi'an 710004, China; sp.huyanhua@gmail.com
*   Correspondence: weibinli@xidian.edu.cn

Abstract: The frequent occurrence of global flood disasters leads to millions of people falling into poverty each year, which poses immense pressure on governments and hinders social development. Therefore, providing more data support for flood disaster detection is of paramount importance. To facilitate the development of water body detection algorithms, we create the DaliWS dataset for water segmentation, which contains abundant pixel-level annotations, and consists of high spatial resolution SAR images collected from the GaoFen-3 (GF-3) satellite. For comprehensive analysis, extensive experiments are conducted on the DaliWS dataset to explore the performance of the state-of-the-art segmentation models, including FCN, SegNeXt, U-Net, and DeeplabV3+, and investigate the impact of different polarization modes on water segmentation. Additionally, to probe the generalization of our dataset, we further evaluate the models trained with the DaliWS dataset, on publicly available water segmentation datasets. Through detailed analysis of the experimental results, we establish a valuable benchmark and provide usage guidelines for future researchers working with the DaliWS dataset. The experimental results demonstrate the F1 scores of FCN, SegNeXt, U-Net, and DeeplabV3+ on the dual-polarization data of DaliWS dataset reach to 90.361%, 90.192%, 92.110%, and 91.199%, respectively, and these four models trained using the DaliWS dataset exhibit excellent generalization performance on the public dataset, which further confirms the research value of our dataset.

Keywords: dataset construction; water segmentation; synthetic aperture radar; deep learning; GF-3

1. Introduction

Frequent global flood disasters have become a pressing issue, resulting in millions of people falling into poverty each year. These events not only impose enormous pressure on governments, but also hinder social development significantly. In July 2021, catastrophic rainfall in Zhengzhou, Henan Province, China, led to 302 deaths or disappearances [1]. In February 2022, heavy rains in Queensland, Australia, caused eight deaths. In September 2022, Sudan experienced torrential rain and flooding that resulted in 146 deaths. The economic losses caused by flood disasters worldwide in the 21st century have exceeded USD 46 billion annually [2], and approximately 70% of the deaths caused by floods in China are attributed to flash flood disasters [3,4]. These devastating impacts of flood disasters highlight the urgency and importance of effective water resource management. During floods, the prolonged duration and immense destructive power of the floodwaters necessitate the timely and accurate acquisition of flood inundation extents [5], which is crucial for minimizing disaster losses and monitoring floods.

Traditional methods for water body extraction often require extensive manpower and material resources. These methods include direct comparison approaches such as

the difference method, ratio method, regression analysis, as well as image transformation approaches such as the Normalized Difference Vegetation Index difference (NDVI), change vector analysis, principal component analysis, and texture-based analysis [6]. In the 1990s, with the emergence of machine learning, researchers began to apply more sophisticated approaches for water detection, including artificial neural networks [7], support vector machines (SVM), decision trees, random forests [7–9], multi-kernel learning, and various hybrid methods such as spectral mixture analysis, fuzzy clustering analysis, and bio-inspired evolutionary algorithms. In the early 21st century, researchers shifted their focus towards object-based image analysis, introducing techniques like Markov random fields, conditional random fields, object-level change vector analysis, and other technologies [6,10,11]. Simultaneously, object category comparison method emerged, including hybrid method at the pixel and object levels. In the 2010s, with the advent of remote sensing big data and artificial intelligence, deep learning such as autoencoders, neural networks [12], recurrent neural networks, and knowledge graphs gradually found application in water body identification.

Deep learning exhibits the ability to dynamically learn by assessing the differences between predicted outcomes and actual ground truth labels, wherein Convolutional Neural Networks (CNNs) stand out as a prominent branch of this domain. Currently, owing to ongoing innovations, CNNs have demonstrated remarkable success across a spectrum of computer vision tasks. Today, diverse standard CNN models have been widely employed for water body segmentation tasks. Several widely recognized and exceptional algorithms, such as SegNet [13], U-Net [14], RefineNet [15], PSPNet [16], Mask R-CNN [17], Deeplab series [18–21], DUPNet [22], CoANet [23], D-LinkNet [24], and PANet [25]. The advent of deep learning techniques has revolutionized the segmentation paradigm from traditional methods, significantly enhancing the accuracy and speed of water body segmentation.

As the data-driven algorithm, the deep learning model relies heavily on the quantity and quality of training data. For water body segmentation tasks, the primary data sources include optical imagery and Synthetic Aperture Radar (SAR) imagery [26]. Optical remote sensing sensors utilize the reflection of sunlight to generate images, constituting a passive imaging system. However, optical imagery is susceptible to factors such as weather conditions, lighting variations, and cloud cover, making it unable to provide all-weather remote sensing capabilities. On the other hand, SAR represents another remote sensing method that relies on active high-resolution microwave radar sensors for imaging. It actively emits electromagnetic waves towards the Earth's surface, and the surface generates echoes that are received by the radar sensor to generate SAR images. SAR images have notable advantages over optical images, since SAR satellites can perform all-weather and all-day data monitoring without being affected by factors such as clouds, fog, and illumination conditions [27]. Researchers and practitioners are more inclined to use SAR images as the preferred data source for water body segmentation tasks. Therefore, utilizing SAR images for water body segmentation is a better choice. In recent years, with the continuous development of SAR technology, the spatial resolution of SAR images has been progressively improving and has now reached sub-meter levels [12,28–30], which facilitates achieving high-precision water body extraction and provides feedback for promptly identifying disaster-stricken areas and mitigating the associated losses.

As is well known, publicly available datasets play a crucial role in driving advancements in the computer vision field and documenting its development process [31]. However, in the context of water body segmentation, there is a lack of accurately annotated high-resolution SAR image datasets. This shortage poses challenges such as unclear targets and inaccurate data, directly impacting the accuracy and generalization capability of models. To address this issue, this study constructs a precisely annotated high-resolution SAR image dataset for water body segmentation based on the GF-3 satellite imagery, termed the DaliWS dataset. Furthermore, several state-of-the-art semantic segmentation models are employed for experimental evaluation on our dataset. Through the analysis of experimental results, a valuable performance benchmark has been established.

Our contributions are summarized as follows:

1. To address the scarcity of publicly available accurately annotated high-resolution SAR image datasets for water body segmentation, this paper provides a manually annotated dataset using GF-3 satellite imagery. The dataset covers large rural water bodies in Dali County and provides pixel-level annotations for the training and validation of deep learning algorithms.

2. We explore the impact of different polarization modes, including HH polarization (single copolarization, horizontal transmit/horizontal receive), HV polarization (single copolarization, horizontal transmit/vertical receive), and HHHV (dual copolarization, horizontal transmit/horizontal receive) on water body segmentation tasks. on the water segmentation task, which provides reference for subsequent related research and facilitates the application of our dataset.

3. To further understand the characteristics of the dataset, this study extensively evaluates its performance using several state-of-the-art segmentation algorithms. The results demonstrate the dataset's inherent challenges and provide new opportunities for water segmentation research material.

4. This study conducts numerous experiments and establishes performance benchmarks specifically for our dataset, laying the foundation for future research. It is anticipated to provide valuable resources and references for the research and development of water body segmentation.

The structure of this paper is as follows: Section 2 presents the processing procedure for SAR images and the creation of the DaliWS dataset. Section 3 demonstrates the segmentation networks used for evaluation, data augmentation methods, and evaluation metrics employed in this study. Section 4 describes the experiments and presents the results. Finally, in Sections 5 and 6, we discuss and summarize the findings of this research.

2. Dataset and Material

2.1. Study Area

The GF-3 satellite captured the study area of the DaliWS dataset in Dali County, Weinan City, Shaanxi Province [32], China. Dali County is situated in the upstream region of the Yellow River, adjacent to Henan Province and Shanxi Province. It is known for its abundant water resources. In this area, multiple rivers converge, including the major Yellow River, the moderate-sized Wei River, and the smaller Beiluo River, forming a unique and rich water network. Additionally, there are numerous artificial water bodies, such as enclosed ponds and reservoirs, which provide ample irrigation and drinking water resources for the local population. Due to its special geographical location and abundant water resources, Dali County is considered a natural and typical research area for hydrological studies, and holds significant importance for water resource management and environmental protection. The satellite orbit coverage of the research area and data source is illustrated in Figure 1.

The study utilized the GF-3 satellite to acquire three sets of images on 24 September 2019, 28 February 2020, and 23 May 2020, respectively. The majority of the covered areas in these images overlap, with the geographical coordinates of the image center captured on 23 May 2020, being at 34°7′ N and 110°3′ E. The original dimensions of the three images are 24,338 × 13,872, 22,553 × 14,384, and 20,312 × 18,480 pixels, respectively, employing a dual-polarization (HH, HV) mode. The imaging mode is Fine Strip I mode of the FSI system, with a resolution of 5 m and a bandwidth of 50 km, providing high-resolution support for water body extraction. The high-quality data obtained from the satellite can be applied in various fields, such as land use classification, environmental monitoring, and disaster assessment. For detailed information, refer to Table 1.
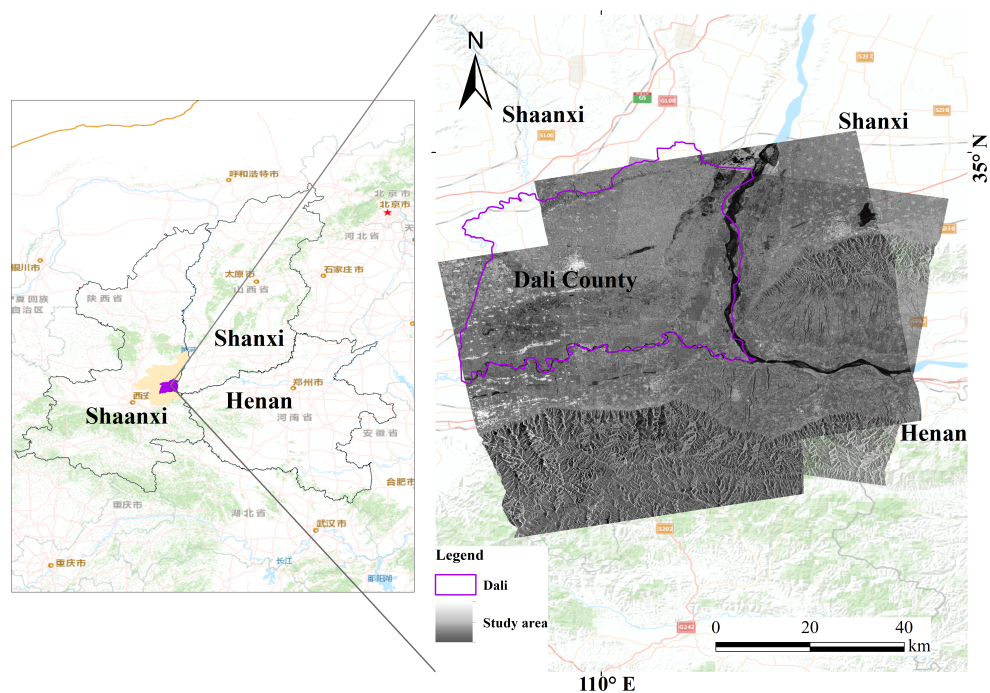
Figure 1. The DaliWS dataset. Geographic location of the study area in Shaanxi province, which is adjacent to Henan province, Shanxi province.

Table 1. Information used in the construction of DailiWS in this article, including sensor, satellite ground station, longitude, latitude, product level, incidence angle, etc.

Item	Parameter
Sensor	GF-3, China
Satellite Ground Station	Sanya Station, China
Longitude	110.3
Latitude	34.7
Imaging band	C bond
Imaging Mode	FSI
Resolution (m)	5
Product Level	L1A
Polarization mode	HH, HV
Incidence Angle (°)	46~49
Swath Width (km)	50

2.2. Data Sources

The data used in this study were obtained from the GF-3 satellite, which is China's first C-band high-resolution SAR satellite with significant importance in various fields such as global ocean observation, disaster reduction [33], glacier identification, land resource monitoring, and surface motion detection. GF-3 is also China's first low-earth orbit remote sensing satellite with a designed lifespan of 8 years, successfully launched in August 2016 [34,35]. To generate products adaptable to different environments, the SAR system of the GF-3 satellite incorporates advanced technologies such as adaptive terrain background classification, polarimetric SAR (PolSAR) [36], and cross-polarization ratio (XPOL), which significantly enhance the quality and application value of SAR data. Furthermore, the GF-3 satellite is designed with 12 different operating modes, achieving a spatial resolution of up to 1 m in spotlight mode.

SAR image products have different levels, and different processing of raw images generates products with varying levels. For different levels of products, the pixel values

represent different physical quantities, including Digital Number (DN) values, amplitude values, intensity values, and backscatter coefficient values [37]. DN values are unsigned integer values commonly used for storing complex SAR image data, while amplitude values and intensity values correspond to the magnitude and squared magnitude of complex data, respectively. Intensity values can be further processed to obtain backscatter coefficient values after radiometric correction. The data products of the GF-3 satellite are classified into L0, L1, L2, and L3 standard products, as well as L4 industry application products. Each level of product undergoes different processing to generate subsequent levels of product.

Currently, one of the important research directions in remote sensing is water body extraction. In this study, we utilized GF-3 satellite imagery to create a corresponding dataset and trained multiple deep learning models, demonstrating through experiments that GF-3 satellites can provide high-precision data for water body extraction.

2.3. Data Preprocessing

This study is based on Single-Look Complex (SLC) images from the GF-3 satellite, which belong to the L1A level product, to create a SAR water body segmentation dataset. The raw data received by the SAR sensor is in the form of raw data (L0 level), which can be processed using the focusing algorithm to generate slant-range SLC images [38]. The pixel values of this level of product are 16-bit complex data composed of real and imaginary parts [39], where the real part reflects the intensity and the imaginary part reflects the phase [26]. To facilitate subsequent operations, SLC data undergoes a series of preprocessing steps aimed at reducing image noise, correcting terrain, and improving overall image quality for better extraction and analysis of target objects.

Intensity is one of the main characteristics of SAR images, and information about objects can be extracted based on intensity images. Therefore, SAR complex data needs to be converted to SAR intensity data. Also, because radar echo signals are coherently added together, SLC data products have a lot of speckle noise that needs to be removed using multi-look processing. Multi-look processing involves averaging the SLC data in the azimuth and range directions to weaken the impact of speckle noise on object imaging. The resulting data from multi-look processing are intensity data, completing the transformation from complex data to intensity data. As multi-look processing averages the noise, it improves radiometric resolution but reduces the spatial resolution of the product image.

Speckle noise is inherent in SAR images because the SAR system itself is a coherent system. Since speckle noise can significantly affect the interpretation of SAR image objects [40], in addition to multi-look processing, specialized speckle noise filtering is required during the preprocessing stage. Commonly used SAR filters include Lee [41], Frost [42], Kuan [43], and Gamma MAP filters [44], among others. The Frost filter is one of the most popular adaptive speckle filters and is an exponentially weighted adaptive filter that suppresses speckle noise in the image using the Minimum Mean Square Error (MMSE) estimation method. Within the kernel size of n-by-n, the center pixel value is replaced by a weighted sum of the values of the neighborhood in kernel [45]. In this study, a 5×5 Frost filter was chosen to smooth the obtained intensity data, as shown in Formula (1). Among the formulas, $f(i, j)$ represents the filtered pixel value, $g(i, j)$ represents the original pixel value, $m(i, j)$ indicates the average value of surrounding pixels, $w(i, j)$ indicates the weighting factor, $V(i, j)$ is the variance of pixel values in the neighborhood, $N(i, j)$ is the number of pixels in the neighborhood, and γ and β are adjustable parameters.

$$\begin{cases} f(i, j) = w(i, j) \times g(i, j) + (1 - w(i, j)) \times m(i, j) \\ \\ w(i, j) = e^{-\left(\frac{\gamma \times V(i,j)}{N(i,j)^{\beta}}\right)} \end{cases} \tag{1}$$

After multi-look processing and Frost filtering, SAR images need to be geocoded and radiometrically calibrated [46]. After the SAR system acquires reflection information from the Earth's surface, it is encoded in the radar coordinate system, known as the slant-range

coordinate system. Generally, before applying SAR data, it needs to be transformed from the slant-range coordinate system to the geographic coordinate system. This process is known as SAR geocoding. The SAR system observes the ratio between the transmitted pulse and the received signal intensity, which is referred to as the backscattering [47]. Different SAR sensors or different receiving modes can affect the backscattering values. Radiometric calibration normalizes the backscattering values to a unified standard, mitigating differences between different SAR sensors. This process is known as SAR radiometric calibration.

The overall processing workflow of remote sensing imagery is shown in Figure 2. These steps complete the preprocessing of SAR images. The preprocessed SAR images are then input into image annotation software for the annotation task.
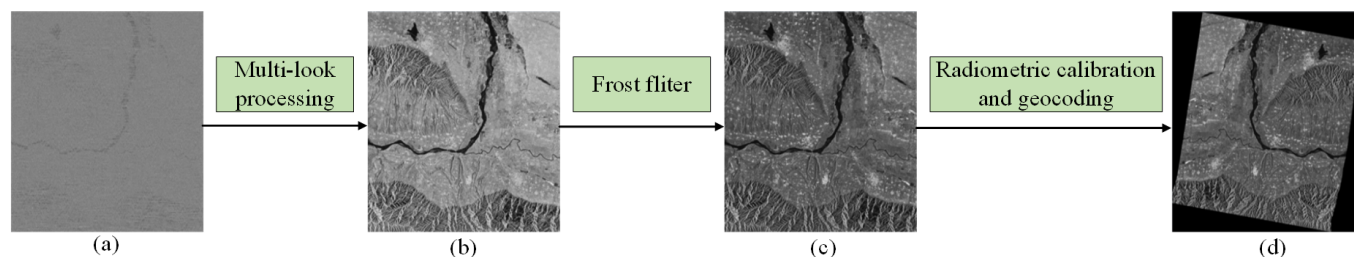

Figure 2. Processing workflow of GF-3 SLC imagery. (**a**) Original SLC imagery; (**b**) multi-view processing result imagery; (**c**) Frost filtering result imagery; (**d**) geocoding and radiometric calibration result imagery.

2.4. Label Generation

In this study, the DaliWS dataset is constructed from three scenes of GF-3 satellite images as described in Section 2.1. Due to significant overlap in most regions of the three images, we primarily use imagery from 23 May 2020 as the main reference, while the other two scenes are utilized to supplement areas not covered by the main image, ensuring comprehensive coverage of most water bodies within Dali County. The generation of dataset labels involved a purely manual annotation process, leveraging domain expertise and calibration with reference to Google Earth [39,48] and other historical images. This approach ensures more accurate annotation of the DaliWS dataset, providing crucial support for further scientific research and practical applications, particularly in the field of SAR image water body segmentation. Such a dataset not only enhances the reliability of the study, but also yields more precise results for practical applications [28].

2.4.1. Chip Creation and Sampling

Figure 3 shows the example used for dataset creation. From the figure, it can be observed that the main land features in the image are water bodies, mountains, and plains. The objective of this study is to create a high-resolution SAR image water segmentation dataset. Therefore, to reduce manual annotation effort and improve annotation accuracy, the image is divided into smaller image blocks before the formal annotation process. Additionally, in later stages of model training, large image blocks can lead to increased computational, memory, and GPU requirements, thereby reducing the parallel training capability of the models [49]. Considering these factors, in this study, the image is segmented into non-overlapping 256×256 image blocks. It is important to note that the width and height of the image are not divisible by 256, resulting in overlapping regions between the last column and last row image blocks with their preceding blocks. After completing the image segmentation process, image blocks containing water body areas are selected for annotation, while image blocks without water body areas are left unmarked. Algorithm 1 describes the process of image partitioning.

Algorithm 1 Image block segmentation

Input: Assigning the original size image as "Img", the image block size is "BlockSize".

Output: ImageBlocks

  1: ImageBlocks = []

  2: H,W = Img.shape

  3: // Calculating the segmentation points for the dimensions "h" and "w".

  4: h = Range(0,H,BlockSize)

  5: w = Range(0,W,BlockSize)

  6: h[−1] = (H-BlockSize)

  7: w[−1] = (W-BlockSize)

  8: **for** $h \to i$ **do** // Iterating through "h" and assigning each iterated element to "i".

  9:      **for** $w \to j$ **do**

10:         ImageBlock = Img[i:i+BlockSize,j:j+BlockSize]

11:         ImageBlocks.add(ImageBlock)

12:      **end for**

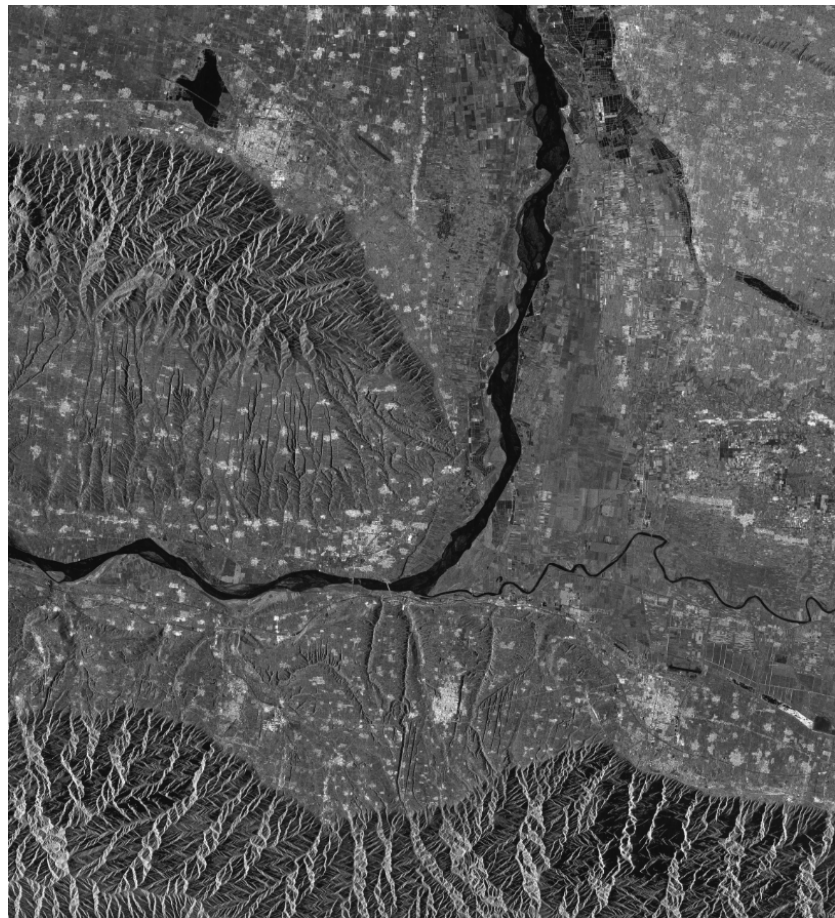13: **end for**

14: **return** ImageBlocks

Figure 3. A GF-3 remote sensing image used to create the water body segmentation dataset.

In Algorithm 1, H and W represent the height and width of the original remote sensing image, respectively. BlockSize is the size of the image block. Range (a, b, c) is a sequence generation function that generates a sequence within the range [a, b] with a step size of c. After the sequence is generated, the last segmentation point needs to be updated to avoid boundary issues.

2.4.2. Hand Labeling

This section describes the process of water area annotation using the labeling tool Labelme. Labelme [50] is an offline annotation tool developed in Python that allows annotations to be made using polygons, rectangles, circles, points, and line segments.

The water segmentation dataset created in this study is a binary classification dataset. Non-water land features such as plains, farmland, and mountains are labeled as "background", while the remaining areas are labeled as "water". In addition to rivers, the water areas also include features such as rice fields and ponds, which pose challenges for this annotation task due to their small size. To improve the accuracy of the annotations, other data sources were used for auxiliary calibration. Specifically, Google Earth [48], expert judgment, and online maps from the National Geospatial Information Sharing Service Platform were used as reference data to calibrate features that were difficult to annotate.

In addition to small water bodies, accurately annotating the transition edges between water and background is also challenging. The suitability of annotating the edge regions will affect the effectiveness of subsequent model learning. During the annotation process, special attention was paid to carefully calibrating the transition edges between the background and water.

For annotated images, Labelme records the coordinates of the polygons in JSON format files. During model training, the labels are required to be binary images, where "0" represents the background region and "1" represents the water region. Therefore, after completing the annotation work for all image blocks, the JSON format labels are batch converted to binary PNG format images using the tools provided by Labelme. An example of the DaliWS dataset is shown in Figure 4, which includes five common types of water bodies (ponds, lakes, rivers, rice fields, and reservoirs) and their corresponding ground truth annotations. The first row displays the SAR image, and the second row displays the corresponding labels.
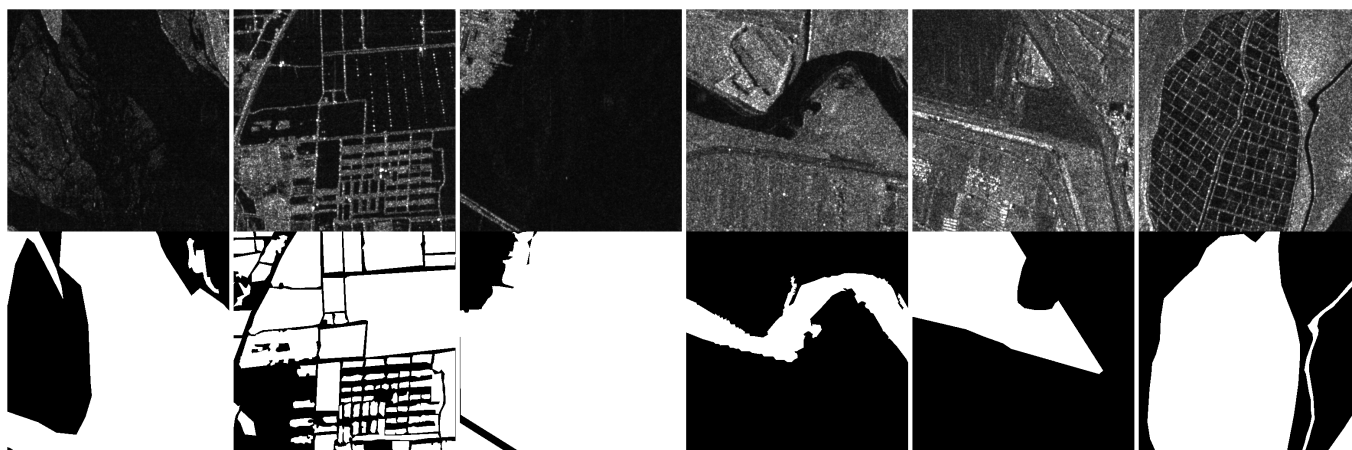

Figure 4. Examples of DaliWS dataset. The images in the top row are SAR images and those in the bottom row are corresponding annotation.

3. Experimental Study

3.1. Dataset Description

3.1.1. DaliWS Dataset

The DaliWS dataset (We provide the DaliWS dataset at: https://github.com/Dataset-RFGroup/DaliWS-Dataset) is a high-resolution SAR dataset for water body segmentation

created in this paper based on GF-3 satellite imagery. The original SLC images undergo multi-view processing, frost filtering, geocoding, and radiometric calibration to generate three SAR images. Subsequently, the image is divided into blocks to obtain 2033 images of size 256 × 256 pixels. Finally, the water bodies are accurately annotated using the Labelme annotation tool to create a complete dataset.

Prior to training, the dataset is divided approximately in a 6:2:2 ratio. Data augmentation techniques such as flipping, contrast and brightness enhancement, translation, random noise, and rotation are applied to expand the training and validation sets.

3.1.2. HISEA-1 Dataset

To validate the generalization capability of the models trained on the DaliWS dataset, we selected a publicly available dataset for testing. This dataset, named HISEA1 flooding dataset [51] and abbreviated as HISEA-1 dataset in this paper, was created using HISEA-1 satellite images for water body segmentation. The HISEA-1 satellite carries a synthetic aperture radar (SAR), which is an active remote sensing system for observing the Earth. The HISEA-1 dataset comprises 2340 non-overlapping SAR images with VV polarization. Each image has a size of 256 × 256 pixels. The dataset covers an area of 20,000 square kilometers and includes various landforms such as rivers, tributaries, reservoirs, lakes, and rice fields.

3.2. Evaluation Methods

The workflow of this study on DaliWS is illustrated in Figure 5. Four models, namely FCN-8s, SegNeXt, U-Net, and DeeplabV3+, are selected for training on the DaliWS training set. After training, the models are evaluated on the DaliWS testing set by comparing the predicted results with the ground truth labels to obtain evaluation metrics. To test the generalization performance of the models trained on the DaliWS dataset, water extraction and evaluation metric calculations are performed on the publicly available HISEA dataset.

3.2.1. FCN-8s

In 2015, Jonathan Long first applied deep learning to the field of semantic segmentation and created Fully Convolutional Networks (FCN) [52]. By adopting the fully convolutional approach, FCN can accept input images of arbitrary sizes and perform pixel-level classification tasks in an end-to-end manner. The characteristics of FCN include replacing fully connected layers with convolutional layers for end-to-end pixel-level training. Additionally, the introduction of skip connections allows for the fusion of multi-scale features [53]. The FCN architecture consists of three series: 32s, 16s, and 8s, which employ different fusion strategies for prediction. Among them, FCN-8s has more parameters and computational complexity, and it achieves higher segmentation performance. Thus, in our experiments, we selected FCN-8s as the model for water segmentation.

3.2.2. SegNeXt

SegNeXt [54] network was proposed in 2022 and achieved the highest ranking in the Visual Object Classes (VOC) challenge at that time. Additionally, the network has demonstrated significant performance improvements on several mainstream semantic segmentation datasets. The major significance of SegNeXt lies in the search for a more cost-effective convolutional attention mechanism beyond transformer-based multi-head attention. SegNeXt introduces a novel Multi-Scale Convolutional Attention (MSCA) module. The MSCA module consists of three parts: deep convolution for aggregating local information, multi-branch deep stripe convolution for enlarging receptive fields, and 1 × 1 convolution for modeling relationships between different channels. SegNeXt network offers four different overall segmentation model sizes, and taking into account model size and performance, we have decided to use the SegNeXt-B model size in our dataset.
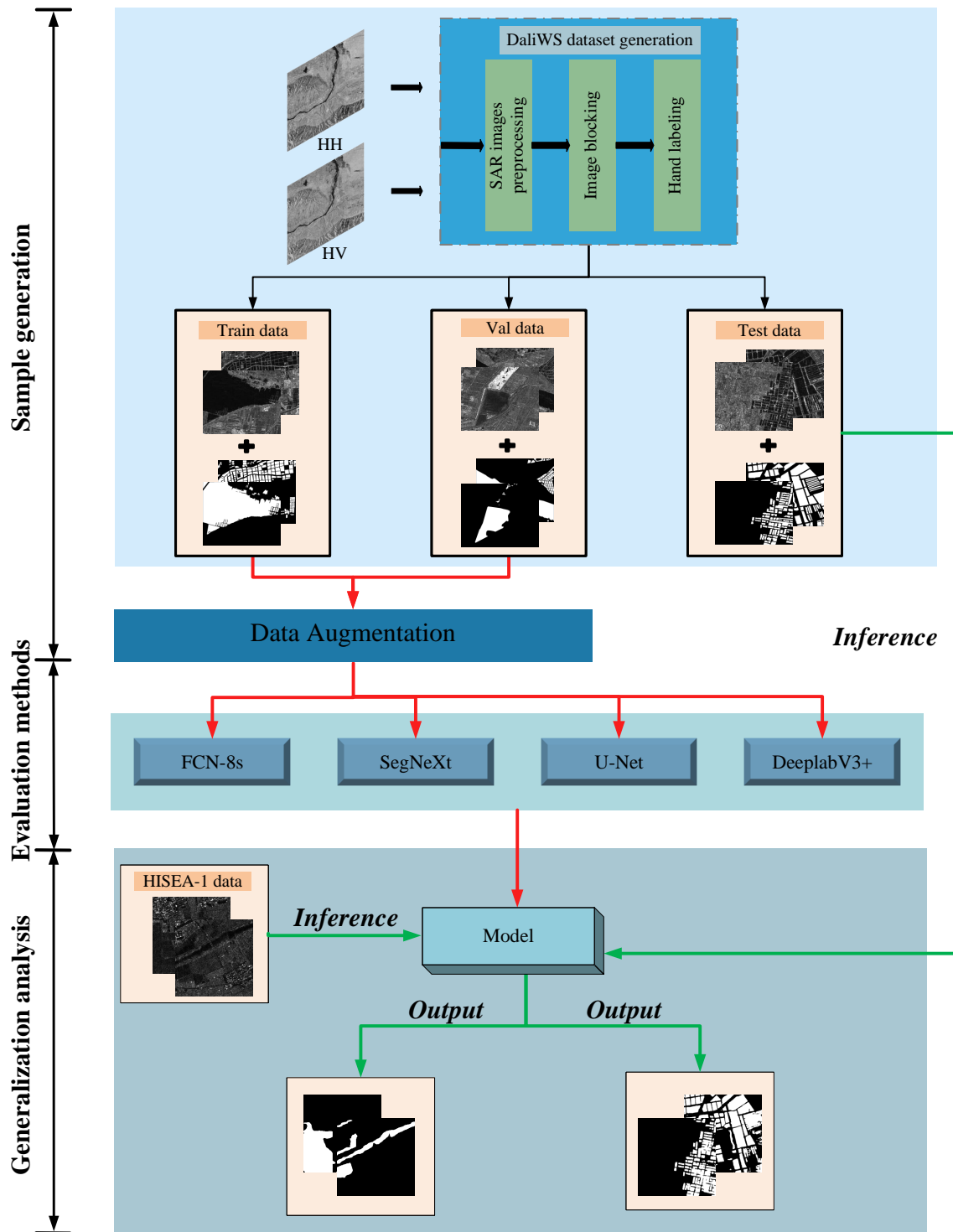
Figure 5. Model training, testing and generalization evaluation flow chart in this paper.

3.2.3. U-Net

As the name suggests, U-Net [14] is a network architecture that resembles the letter "U" and is commonly used for semantic segmentation tasks. It was initially proposed by Ronneberger for medical segmentation. The U-Net architecture consists of a contracting path and an expanding path, making it a typical encoder–decoder network. U-Net incorporates skip connections between the encoder and decoder to enable multi-scale information fusion. Due to its simplicity and effectiveness, U-Net can achieve high accuracy with

limited training samples, making it widely applicable in various image segmentation tasks, including remote sensing image segmentation.

3.2.4. DeeplabV3+

Deeplab is a series of deep learning models for image semantic segmentation, initially proposed by the Google team and continuously improved. The final version of the Deeplab series, DeeplabV3+ [21], was introduced in 2018. It utilizes the Xception [55] architecture as the backbone network for extracting abstract and high-level semantic features and incorporates new decoders to achieve more precise object boundary segmentation. In this paper, we selected ResNet50 [56] as the backbone network for DeeplabV3+. One notable feature of DeeplabV3+ is the adoption of the Atrous Spatial Pyramid Pooling (ASPP) module. Additionally, the decoder module in DeeplabV3+ is simple yet effective, gradually recovering spatial features of the targets and achieving finer segmentation of object boundaries. These advantages make DeeplabV3+ stand out among many other methods.

3.3. Experimental Settings

3.3.1. Implementation Details

As shown in Table 2, all experiments were conducted on a workstation equipped with a 64-bit Intel(R) Xeon(R) Gold 6226R CPU @ 2.90 GHz processor, RECC DDR4 128 GB memory, and NVIDIA GeForce RTX 3090Ti graphics card, using the PyTorch framework. The network's input was images of size 256×256 pixels. Additionally, each deep learning model was trained for 100 epochs with the Adam optimizer [57], using an initial learning rate of 1×10^{-5} and a batch size of 16. The training time depends on the complexity of the model.

Table 2. The experimental environment used in this study.

Item	Configuration Version
Operating system	Linux CentiOS7.9
Processor	64 Intel(R) Xeon(R) Gold 6226R CPU @ 2.90 GHz
GPU	NVIDIA GeForce RTX 3090
CUDA	CUDA 11.6
Depth framework	PyTorch 1.13.1

3.3.2. Data Augmentation

Data augmentation plays a crucial role in data preprocessing before model training. As networks become deeper, the number of parameters to learn increases, which can lead to overfitting. To address this issue, many researchers employ data augmentation methods to increase the amount of data, enhance data diversity, alleviate overfitting, and improve the generalization and robustness of deep learning models [14,58].

To enhance data diversity and improve model performance, we applied the data augmentation methods listed in Table 3 to the original image patches obtained during the image segmentation stage. Specifically, these methods include: (1) horizontal/vertical/diagonal flipping; (2) contrast enhancement; (3) brightness enhancement; (4) random noise; (5) random translation; and (6) random rotation.

Table 3. Parameters of different data augmentation methods. / indicates that no parameter is set.

Augmentation Methods	Parameter
Horizontal flip	/
Vertical flip	/
Diagonal flip	/
Contrast enhancement	[0.5, 2.5]
Brightness enhancement	[0.5, 2.5]
Random noise	[0, 0.1]
Translation	[−1,1]
Random rotation	$[-\frac{\pi}{4}, \frac{\pi}{4}]$

Diagonal flipping flips the image along the top-left to bottom-right diagonal. Contrast enhancement is applied with a factor ranging from 0.5 to 2.5, while brightness enhancement is applied with a factor ranging from 0.5 to 2.5. Random noise enhancement randomly assigns 0–10% of pixels in each image to have a value of 255. Translation augmentation shifts the image along the top-left to bottom-right diagonal, where negative values indicate shifting from the bottom-right to top-left and positive values indicate shifting from the top-left to bottom-right. Random rotation randomly selects an angle within the range of $[-\frac{\pi}{4}, \frac{\pi}{4}]$ for rotation around the center. Figure 6 shows examples of the aforementioned data augmentation techniques.
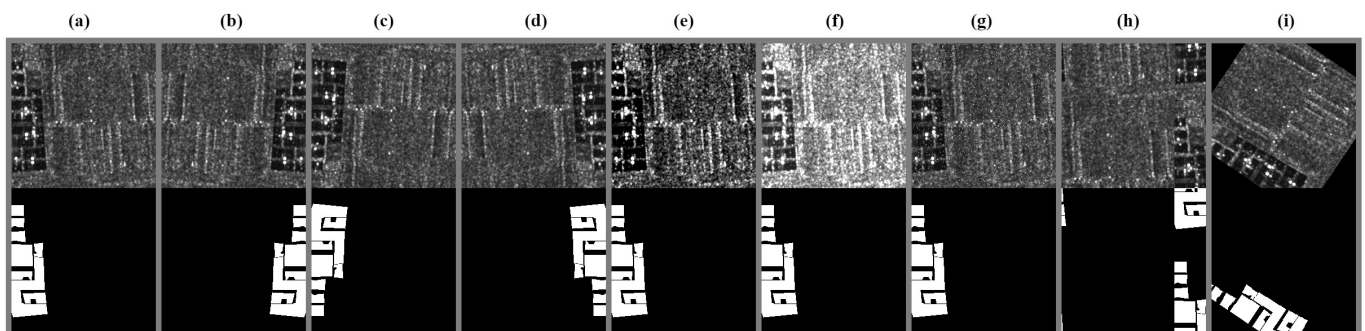

Figure 6. Diagrams of different data augmentation methods. (**a**) Original image. (**b**) Horizontal flip. (**c**) Vertical flip. (**d**) Diagonal flip. (**e**) Contrast enhancement. (**f**) Brightness enhancement. (**g**) Random noise. (**h**) Translation. (**i**) Random rotation.

3.3.3. Evaluation Metrics

The distribution of water and background in the DaliWS dataset is imbalanced, making it challenging to accurately segment the water regions. To comprehensively evaluate the performance of the selected four models on the DaliWS dataset, we employ multiple evaluation metrics. The specific evaluation metrics include accuracy, recall, precision, F1 score, mean Intersection over Union (mIoU), and segmentation time. Pixel accuracy (PA) refers to the ratio of correctly classified pixels to the total number of pixels. Recall (Rec) describes the ability to find all positive samples of a specific class. Precision (Pre) represents the ability of the classifier to label negative samples as positive samples [59]. The F1 score (F1) is a composite metric that balances precision and recall. mIoU is the average of Intersection over Union for each class. Since inference time is critical for disaster monitoring [51], the speed of water extraction is also included in the evaluation. Segmentation time refers to the time required for the model to complete the water segmentation task in an image. Higher image resolutions require more time.

Based on the confusion matrix shown in Table 4, the calculation formula is as follows:

$$PA = \frac{TP + TN}{TP + FN + FP + TN} \tag{2}$$

$$Rec = \frac{TP}{TP + FN} \tag{3}$$

$$Pre = \frac{TP}{TP + FP} \tag{4}$$

$$F1 = \frac{2 \times Pre \times Rec}{Pre + Rec} \tag{5}$$

$$mIoU = \frac{1}{k+1} \sum_{i=0}^{k} \frac{TP}{TP + FP + FN} \tag{6}$$

In the formulas, TP represents the number of pixels correctly predicted as water, FN represents the number of water pixels falsely predicted as background, FP describes the number of background pixels falsely predicted as water, and TN represents the number of background pixels correctly predicted as background.

Table 4. The confusion matrix.

		Prediction	
		Water	**Background**
Ground truth	Water	TP	FP
	Background	FN	TN

4. Results

4.1. Quantitative Analysis of Four Networks on DaliWS Dataset

To comprehensively explore the DaliWS dataset, we train four segmentation networks, e.g., FCN, SegNeXt, U-Net, and DeeplabV3+ on the training set. Subsequently, we evaluate the trained models using the test set. We then perform a quantitative comparison of the segmentation results of these four models on the DaliWS dataset and conduct an in-depth investigation into the impact of polarization modes on segmentation performance. To ensure experimental fairness, we employ consistent evaluation metrics, evaluation code, and maintain identical network parameter configurations as described in Section 3.3.2.

In the first experiment, we focus on examining the water body extraction performance of different models on the DaliWS dataset. Results from Table 5 indicate that all four models achieve F1 scores of over 85%, with U-Net exhibiting the best performance. In the dual-polarization mode, U-Net achieves F1, PA, and mIoU scores of 92.110%, 98.371%, and 85.374%, respectively. DeeplabV3+ closely followed, achieving F1, PA, and mIoU scores of 91.199%, 98.188%, and 83.822% in the dual-polarization mode. In comparison, FCN-8s and SegNeXt networks demonstrate relatively poorer performance in water area segmentation, with F1 scores of 90.361% and 90.192% in the dual-polarization mode. Table 6 displays the GFLOPs and Params for these four models with an input size of (1,1,256,256), revealing that the order of GFLOPs from highest to lowest is U-Net, DeeplabV3+, SegNeXt, and FCN. U-Net has the smallest parameter count while also having the highest computational workload.

In the second experiment, we delve into the impact of different polarization modes on segmentation performance. Firstly, FCN, SegNeXt, U-Net, and DeeplabV3+ achieve F1 scores of 90.361%, 90.192%, 92.110%, and 91.199%, respectively, on HH + HV dual-polarization data, demonstrating superior performance compared to HH and HV single-polarization data. This is understandable, as the fusion of HH and HV dual-polarization data provides more comprehensive and richer spatial features for water body segmentation models. On HH and HV single-polarization data, U-Net consistently performs the best,

with F1 scores of 91.173% and 90.198%, respectively. Overall, dual-polarization outperforms single-polarization, with HH polarization being superior to HV polarization.

Table 5. Quantitative comparison of different models in terms of F1 score(%) , PA(%), Rec (%), Pre (%), and mIoU(%) on the DaliWS dataset.

Models	Polarization	F1(%)	PA(%)	Rec(%)	Pre(%)	mIoU(%)
FCN	HH	89.762	97.889	90.235	89.294	81.426
	HV	88.823	97.687	89.605	88.055	79.893
	HH+HV	**90.361**	**98.012**	**90.853**	**89.875**	**82.417**
SegNeXt	HH	90.090	97.967	**90.101**	90.080	81.967
	HV	87.894	97.605	84.779	91.246	78.403
	HH+HV	**90.192**	**98.023**	88.659	**91.780**	**82.137**
U-Net	HH	91.173	98.197	90.800	**91.549**	83.778
	HV	90.198	98.002	89.624	90.779	82.146
	HH+HV	**92.110**	**98.371**	**92.744**	91.484	**85.374**
DeeplabV3+	HH	90.898	98.170	89.126	**92.742**	83.315
	HV	89.994	97.906	**91.844**	88.216	81.808
	HH+HV	**91.199**	**98.188**	91.555	90.845	**83.822**

Bold indicates the maximum value of each evaluation index.

Table 6. The parameter sizes of each model are listed in the table below.

Model	Input-Size	GFLOPS	Params(M)	Time(s)
FCN	(1,1,256,256)	**7.78**	46.24	0.0100
SegNeXt	(1,1,256,256)	7.96	27.54	0.0324
U-Net	(1,1,256,256)	31.81	**16.69**	0.0093
DeeplabV3+	(1,1,256,256)	26.51	40.48	0.0138

"Times" refers to the inference time for a single image.

4.2. Qualitative Analysis of Four Networks on DaliWS Dataset

Figures 7 and 8 illustrate the partial water body prediction results of FCN-8s, SegNeXt, U-Net, and DeeplabV3+ on the DaliWS test dataset. From the figure, it can be observed that all models are able to predict large water bodies to some extent. However, they exhibit limitations in accurately segmenting boundaries, shadows, and small water bodies.

In the first row, it can be seen that SegNeXt and U-Net incorrectly identify a small tributary as water, while FCN and DeeplabV3+ correctly classify it as background. However, most methods erroneously label farmland as water. In the second row, U-Net misses a portion of the water body in the river. In the last row, none of the methods are able to effectively segment the water body and path boundaries in the shadow areas. Among these four models, FCN produces relatively coarse predictions due to the final feature map output not being of the same size as the original input. It requires 8-fold upsampling to restore the prediction map size. While this approach reduces computational complexity, it results in the loss of significant detail. For the DaliWS dataset, which has a spatial resolution of 5 m, such coarse boundaries severely impact the accuracy of water body extent calculations.
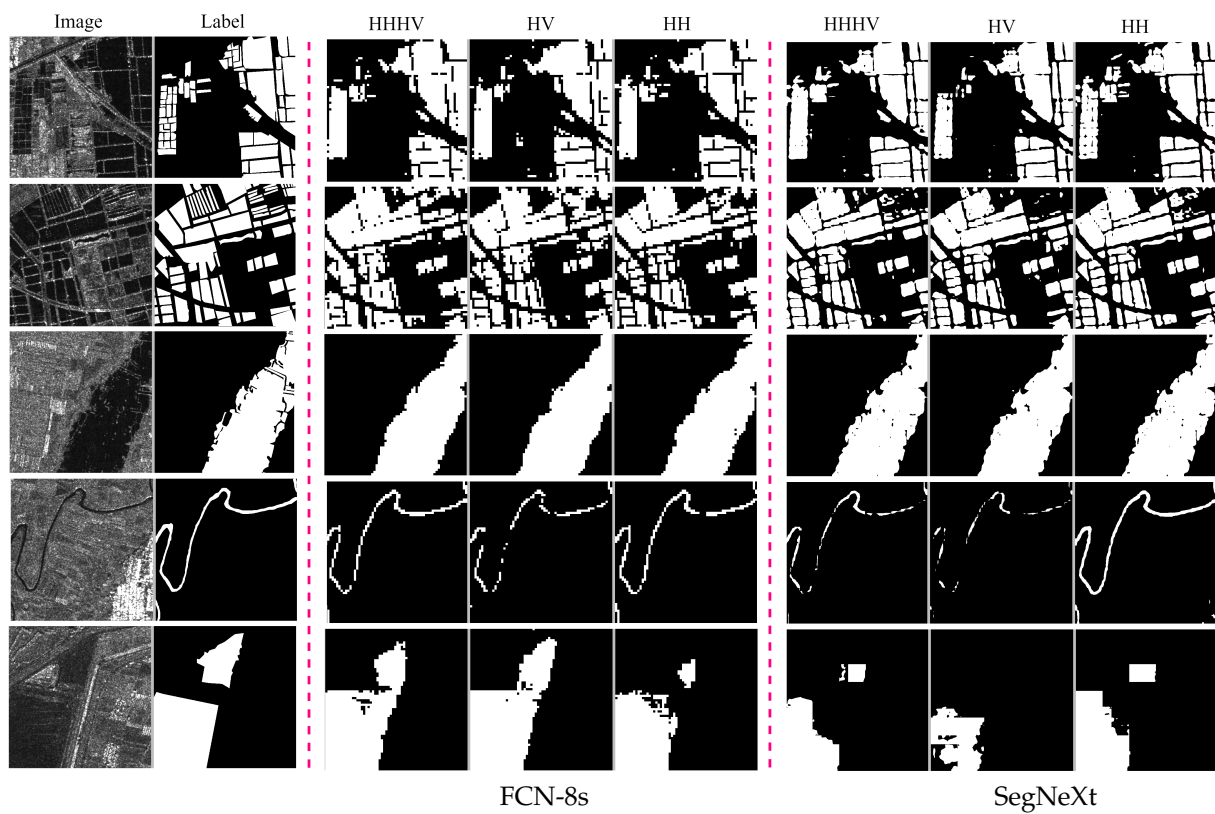
Figure 7. Comparison example images of FCN-8s and SegNeXt models trained on differently polarized data for water segmentation.
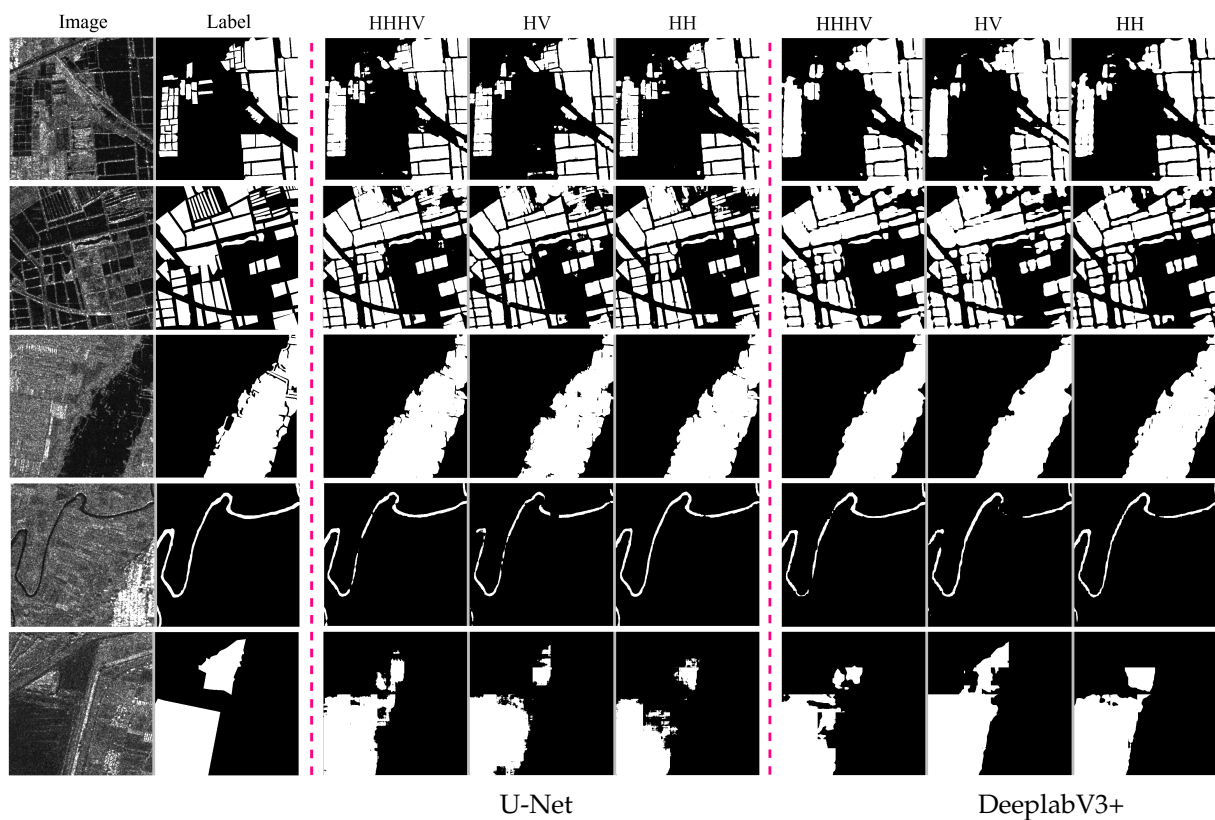

Figure 8. Comparison example images of U-Net and DeeplabV3+ models trained on differently polarized data for water segmentation.

The four models are trained using three different modes: HH, HV, and HH+HV. By observing the results in the first and second rows of Figures 7 and 8, it can be seen that U-Net achieves the best segmentation performance across all three modes. The models trained with HV single-polarization data performed the worst. In the fourth row of Figures 7 and 8, the models trained with HH single-polarization data show better segmentation performance compared to HH+HV dual-polarization data, particularly in accurately segmenting elongated rivers. In the fifth row of Figures 7 and 8, all four models struggled to segment water bodies with indistinct boundaries.

4.3. Generalization Analysis of Four Models Trained on the DaliWS Dataset

To assess the generalization capabilities of models trained on the DaliWS dataset, we conduct generalization experiments on the HISEA-1 dataset. To do this, different polarization data from DaliWS are used to train segmentation models like FCN-8s, SegNeXt, U-Net, and DeeplabV3+. These models are then tested on the HISEA-1 dataset to obtain segmentation metrics. The DaliWS dataset has two polarization modes, HH and HV, while the HISEA-1 dataset only has VV polarization. For each model, we perform three sets of experiments: (1) HH testing on VV, (2) HV testing on VV, and (3) HHHV testing on VV. It is noteworthy that, in experiment (3), we duplicate the VV data to create a dual-channel image.

The partial water body prediction results of FCN-8s, SegNeXt, U-Net, and DeeplabV3+ on the HISEA-1 dataset are shown in Figures 9 and 10. From the figures, it can be observed that HH polarization data overall outperformed HV and HH+HV, with the latter two polarization data showing instances of interrupted flow in predicting water bodies. The higher rate of false detections in the U-Net prediction map in Figure 10 demonstrates that for HH+HV polarization data; the models do not efficiently extract complementary features from both data types.
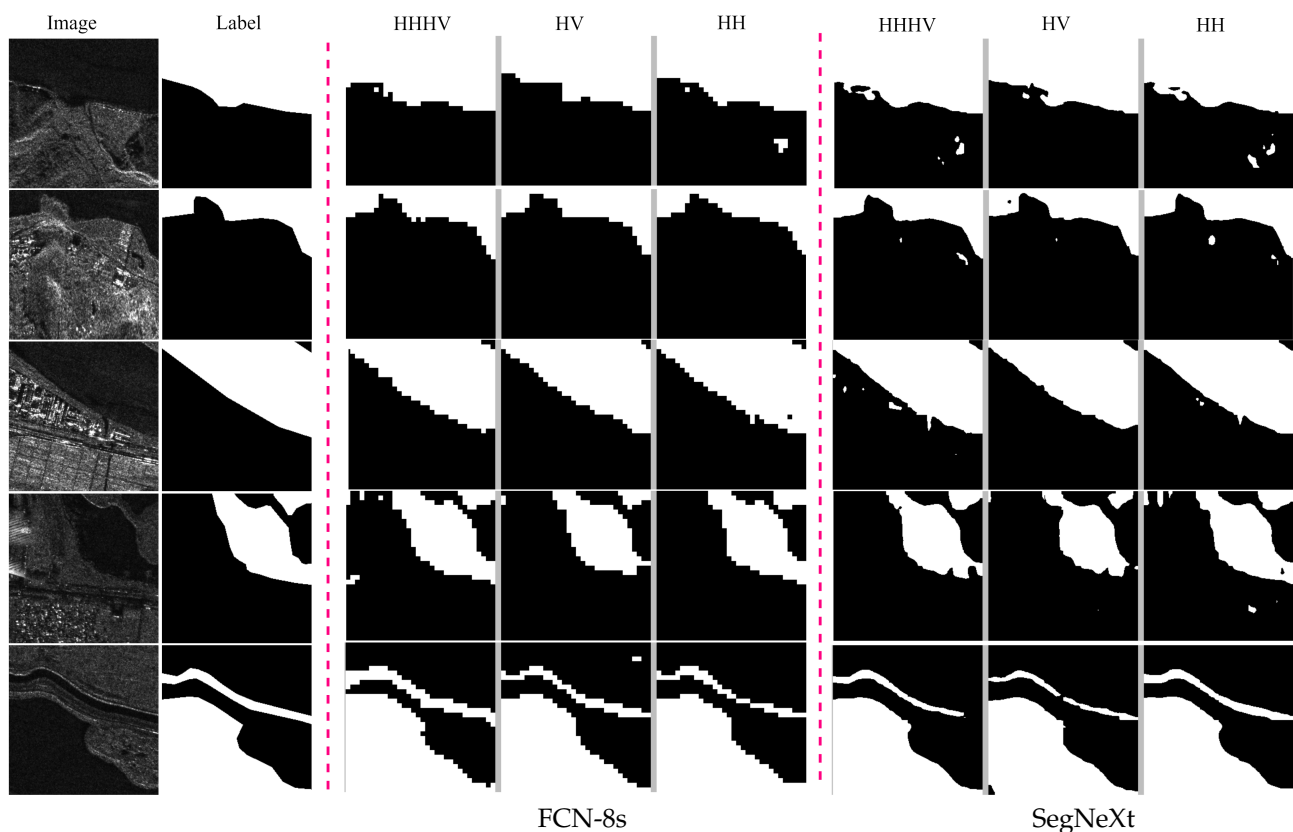

Figure 9. Comparison of example images for generalization predictions by FCN-8s and SegNeXt models.
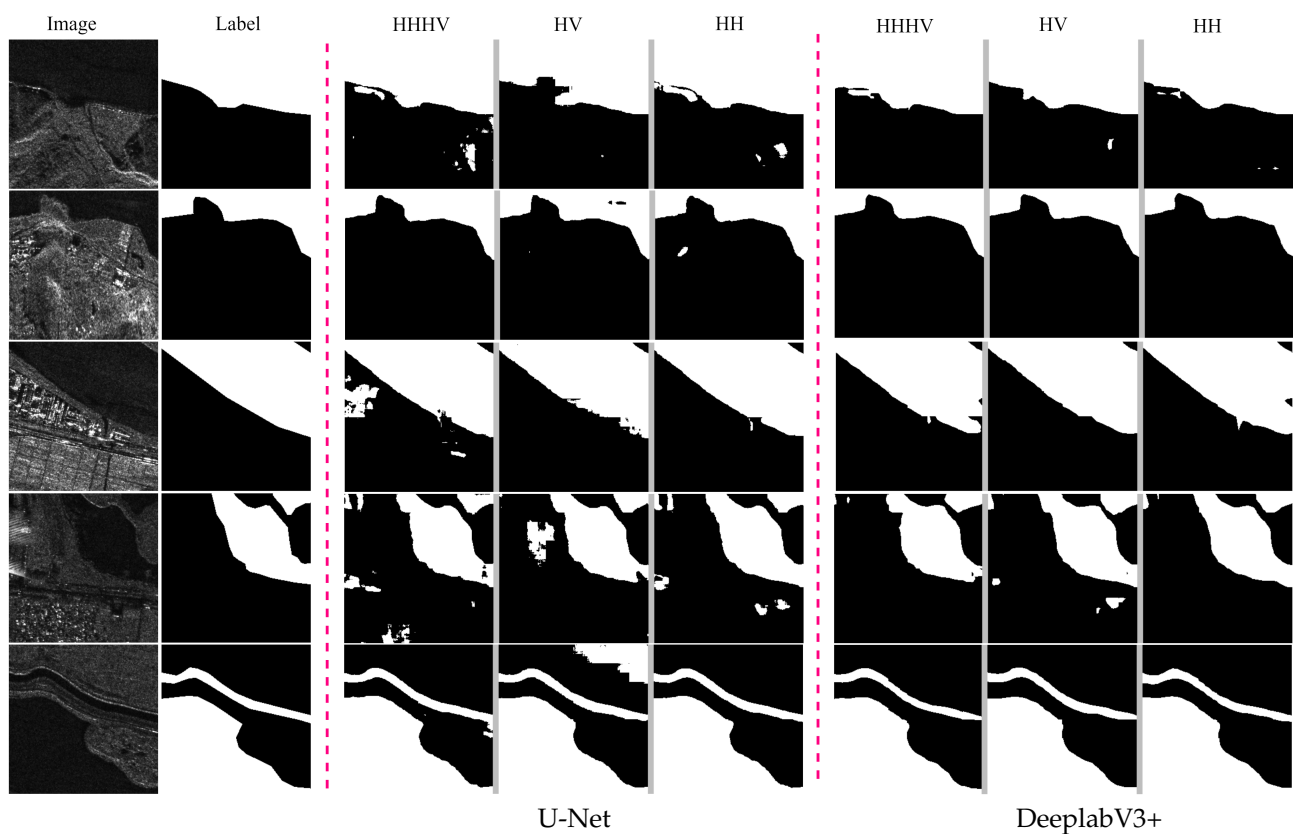
Figure 10. Comparison of example images for generalization predictions by U-Net and DeeplabV3+ models.

The generalization test results in Table 7 show that models trained with HH polarization data from the DaliWS dataset obtain F1 scores higher than 80% on the VV polarization data from the HISEA-1 dataset, showing better generalization performance. The models that are trained with HH+HV dual-polarization data come in second. Except for FCN, the HH+HV dual-polarization data exhibits commendable performance on the other three models. The generalization capability of HV polarization data are the weakest, significantly lower than that of HH and HH + HV. In terms of overall model generalization performance, SegNeXt exhibits the best results, with F1 scores of 85.276%, 77.307%, and 82.608% for HH, HV, and HH+HV, respectively. DeeplabV3+ follows closely, achieving F1 scores of 85.982%, 77.446%, and 81.650%. FCN demonstrates the poorest generalization performance. In summary, the four models trained on the DaliWS dataset in this study display excellent generalization capabilities on the public dataset, affirming the dataset's value for research purposes.

Table 7. Generalization results of four models trained on the DaliWS dataset.

Models	Polarization	F1(%)	PA(%)	Rec(%)	Pre(%)	mIoU(%)
	HH	**83.890**	**93.543**	**82.770**	85.040	**72.250**
FCN	HV	62.070	83.645	65.886	58.672	45.001
	HH+HV	75.867	91.597	65.028	**91.042**	61.117
	HH	**85.276**	**94.177**	**83.023**	87.654	**74.331**
SegNeXt	HV	77.307	92.064	66.553	**92.206**	63.008
	HH+HV	82.608	93.468	76.379	89.944	70.369

Table 7. *Cont.*

Models	Polarization	F1(%)	PA(%)	Rec(%)	Pre(%)	mIoU(%)
U-Net	HH	**85.391**	**94.313**	**81.827**	**89.280**	**74.506**
	HV	68.426	87.234	68.110	68.746	52.006
	HH+HV	80.365	91.937	81.240	79.508	67.175
DeeplabV3+	HH	**85.982**	**94.553**	**82.241**	90.080	**75.411**
	HV	77.446	90.715	78.487	76.433	63.194
	HH+HV	81.650	93.297	73.428	**91.945**	68.990

Bold indicates the maximum value of each evaluation index.

5. Discussion

In order to support research on Deep Neural Network (DEN)-based water body segmentation algorithms, this paper creates a high-resolution SAR image water body segmentation dataset. The dataset consists of L1A-level SLC images from the GF-3 satellite, captured on 24 September 2019, 28 February 2020, and 23 May 2020, respectively. The original images underwent multi-view processing, Frost filtering, geocoding, and radiometric calibration to generate the final images used for dataset creation. Subsequently, the images are partitioned and sampled, dividing the large image into non-overlapping 256 × 256 image blocks and selecting the parts containing water bodies for the dataset. Finally, the Labelme tool was used for image annotation.

After completing the construction of the DaliWS dataset, this study selects four segmentation networks, namely FCN, SegNeXt, U-Net, and DeeplabV3+, for training and evaluation on the dataset. Figure 11 shows a full comparison of how well these four models can segment and generalize across three polarization modes on the DaliWS dataset. All four models demonstrate outstanding segmentation performance on the DaliWS dataset. Among them, U-Net exhibits the best segmentation accuracy and inference time, achieving an F1 score of 92.110%. However, there is still room for improvement in the segmentation results, particularly in capturing details at the edges.

The performance of SegNeXt in Section 4.1 prompted deep reflection. Despite ranking highest in the VOC challenge, SegNeXt shows the poorest segmentation performance on the DaliWS dataset. In response, we provide a reasonable explanation by comparing the differences between the two datasets. Firstly, although the SegNeXt model excelled on the VOC dataset, it cannot be assumed that it would achieve the same performance on SAR images due to domain differences between natural images and SAR images, which may hinder effective knowledge transfer. Secondly, water body segmentation is a binary classification task, while SAR images are primarily composed of black and gray. A complex SegNeXt network may struggle to allow each parameter to learn sufficiently, whereas simpler segmentation networks like U-Net and Deeplabv3+ are more suitable for binary segmentation tasks on SAR images.

In Table 6, U-Net has the highest GFLOPS among the four models and the shortest inference time per image. In contrast, SegNeXt has significantly lower GFLOPS than U-Net, yet it has the longest inference time per image. Additionally, by comparing the GFLOPS and inference time of other models in Table 6, we conclude that GFLOPS is only a theoretical indicator for measuring model speed and cannot represent the actual inference speed of the model.

In the generalization experiments shown in Table 7, the HH polarization mode performs the best in generalization testing, while models trained using HV polarization data show significantly lower performance than HH polarization. We provide an explanation for this phenomenon: the HISEA-1 dataset used VV polarization, corresponding to HH, making the data distribution of these two modes more similar. On the other hand, although the HH+HV polarization mode integrates HH polarization information, the presence of HV polarization interferes with the overall segmentation results.
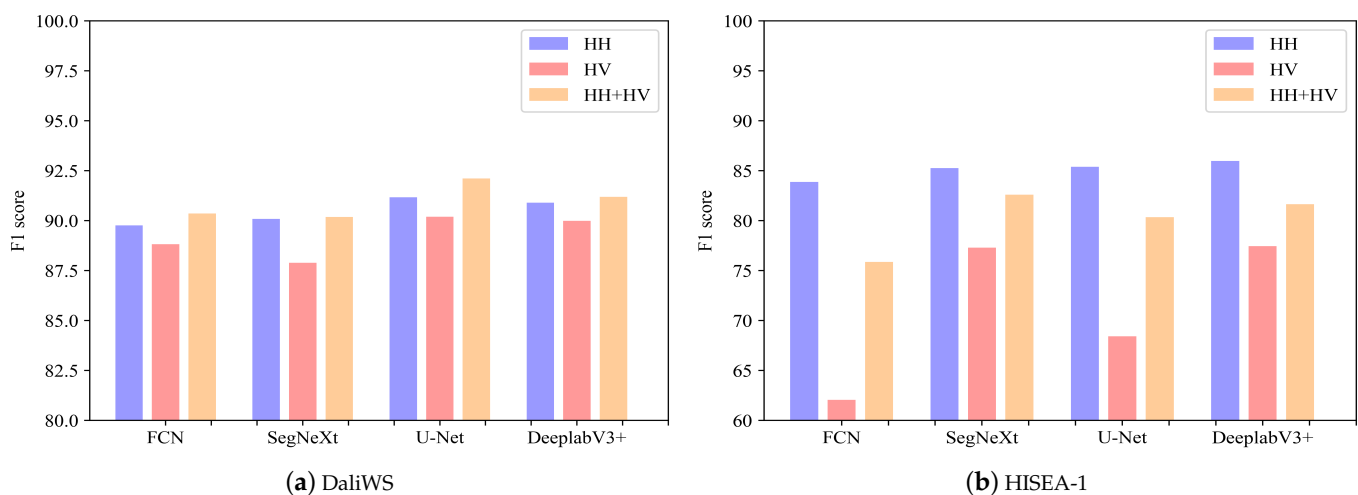
Figure 11. (**a**) F1 score histograms for four models trained on the DaliWS dataset with three different polarization modes; (**b**) F1 score histograms for four models trained on the DaliWS dataset with three polarization modes, evaluated on the HISEAR-1 dataset.

In summary, the experimental results of this study validate the effectiveness of the DaliWS dataset and compare and analyze the water body segmentation performance under different segmentation networks and polarization modes. This research holds significant implications for water remote sensing image analysis and water resource management, providing valuable insights for related studies and applications.

6. Conclusions

In conclusion, despite the rapid advancements in computer vision technology, its high-precision performance still relies heavily on precisely annotated datasets. This study aims to contribute significantly to SAR image water body segmentation tasks and offer valuable insights for future research endeavors. We meticulously create and extensively explore the DaliWS dataset, assessing it through a range of network models. While achieving satisfactory segmentation results across diverse network models, there remains room for enhancement in capturing finer details at water body edges.

Our investigation also delves deeply into the impact of different polarization modes on segmentation performance. And the experimental results demonstrate a certain correlation between datasets sharing similar polarization modes, thereby contributing to superior generalization performance. Notably, models trained on HH+HV datasets perform worse on the VV dataset. This shows that more research is needed to figure out how to combine multi-polarization data effectively.

DaliWS provides multi-polarized information and precise annotations, facilitating the extraction of practical disaster-related information. However, the dataset's sources are not yet comprehensive enough. In the future, we will continue to upgrade DaliWS by collecting multi-source remote sensing data and conducting pixel-level annotations, thereby offering greater convenience for flood disaster emergency responses. To sum up, our work establishes a high-resolution SAR image water body segmentation dataset, analyzes the performance of various segmentation networks and polarization modes, and provides crucial practical insights and references for water body remote sensing image analysis and water resource management.

Author Contributions: Conceptualization, S.Z. and C.L.; methodology, S.Z.; validation, S.Z. and R.W.; investigation, S.Z. and R.W.; resources, W.L., X.F. and Y.H.; data curation, S.Z. and R.W.; writing—original draft preparation, S.Z., C.L., W.L. and R.W.; writing—review and editing, S.Z. and C.L.; visualization, S.Z., C.L., W.L. and R.W.; supervision, W.L., X.F. and Y.H.; project administration, S.Z.,

References

1. Liu, S.; Wang, J.H. Analysis of the monitoring ability of high-resolution satellites for the "21·7" heavy rain in Henan. *Acta Meteorol. Sin.* **2022**, *80*, 765–776.
2. Li, C.; Xie, J.G.L. A Review of the Study on Flash Flood Early Warning in China. *Pearl River* **2017**, *38*, 29–35.
3. Xu, J.; Wang, Y. Water Problems and Their Countermeasures in the Protection and Developing of the Yangtze River. *J. Yangtze River Sci. Res. Inst.* **2020**, *37*, 1–6.
4. Liu, C.;Nie, R. Research Conception and Achievement Prospect of Key Technologies for Forecast and Early Warning of Flash Flood and Sediment Disasters in Mountainous Rainstorm. *Adv. Eng. Sci.* **2020**, *52*, 1–8.
5. Wu, Y.; Ma, M. Urban flood prevention countermeasures in Nanjing of Jiangsu Province. *China Flood Drought Manag.* **2023**, *33*, 53–56.
6. Qi, X. Overview of Object-oriented Change Detection in Remote Sensing Images. *Beijing Surv. Mapping.* **2021**, *35*, 427–431.
7. Sarp, G.; Ozcelik, M. Water body extraction and change detection using time series: A case study of Lake Burdur, Turkey. *J. Taibah Univ. Sci.* **2017**, *11*, 381–391. [CrossRef]
8. Guo, Z.; Wu, L.; Huang, Y.; Guo, Z.; Zhao, J.; Li, N. Water-Body Segmentation for SAR Images: Past, Current, and Future. *Remote Sens.* **2022**, *14*, 1752. [CrossRef]
9. Rokni, K.; Ahmad, A.; Selamat, A.; Hazini, S. Water feature extraction and change detection using multitemporal Landsat imagery. *Remote Sens.* **2014**, *6*, 4173–4189. [CrossRef]
10. Zhang, X.; Zhang, H.; Wang, C. Water-change detection with Chinese Gaofen-3 simulated compact polarimetric SAR images. In Proceedings of the 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA), Beijing, China, 13–14 November 2017; pp. 1–4.
11. Fan, K.; Wang, Z. Change detection of remote sensing images through DT-CWT and MRF. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *21*, 375–385. [CrossRef]
12. Xu, Z.; Wang, R.L.N. A novel approach to change detection in SAR images with CNN classification. *J. Radars* **2017**, *6*, 483–491.
13. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]
14. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18; Springer: Berlin, Germany, 2015; pp. 234–241.
15. Lin, G.; Milan, A.; Shen, C.; Reid, I. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1925–1934.
16. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2881–2890.
17. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
18. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv* **2014**, arXiv:1412.7062.
19. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [CrossRef] [PubMed]
20. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv* **2017**, arXiv:1706.05587.

21. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.

22. Liu, Z.; Chen, X.; Zhou, S.; Yu, H.; Guo, J.; Liu, Y. DUPnet: Water Body Segmentation with Dense Block and Multi-Scale Spatial Pyramid Pooling for Remote Sensing Images. *Remote Sens.* **2022**, *14*, 5567. [CrossRef]

23. Mei, J.; Li, R.J.; Gao, W.; Cheng, M.M. CoANet: Connectivity attention network for road extraction from satellite imagery. *IEEE Trans. Image Process.* **2021**, *30*, 8540–8552. [CrossRef]

24. Zhou, L.; Zhang, C.; Wu, M. D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–23 June 2018; pp. 182–186.

25. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path aggregation network for instance segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8759–8768.

26. Lei, Y.; Leng, X.Z. Construction and recognition performance analysis of wide-swath SAR maritime large moving ships dataset. *J. Radars* **2022**, *11*, 347–362.

27. Huang, Q.; Zhu, W.Y. Summary of research on construction of SAR image ship target detection dataset. *Telecommun. Eng.* **2021**, *61*, 1451–1458.

28. Zhu, D.; Geng, Z.X. SAR Database Construction for Ground Targets at Multiple Angles and Target Recognition Method. *J. Nanjing Univ. Aeronaut. Astronaut.* **2022**, *54*, 985–994.

29. Sun, X.; Wang, Z. AIR-SARShip-1.0: High-resolution SAR Ship Detection Dataset. *Leidaxuebao* **2019**, *8*, 852–863.

30. Wang, Y.; Wang, C.; Zhang, H.; Dong, Y.; Wei, S. A SAR dataset of ship detection for deep learning under complex backgrounds. *Remote Sens.* **2019**, *11*, 765. [CrossRef]

31. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [CrossRef]

32. Ke, H.L.H. Present state and change of phreatic water quality in Dali County, Shaanxi, China. *Geol. Bull. China* **2008**, *27*, 1196–1204.

33. Wen, N.; Zeng, F.; Dai, K.; Li, T.; Zhang, X.; Pirasteh, S.; Liu, C.; Xu, Q. Evaluating and Analyzing the Potential of the Gaofen-3 SAR Satellite for Landslide Monitoring. *Remote Sens.* **2022**, *14*, 4425. [CrossRef]

34. Zhao, L.; Zhang, Q.; Li, Y.; Qi, Y.; Yuan, X.; Liu, J.; Li, H. China's Gaofen-3 satellite system and its application and prospect. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 11019–11028. [CrossRef]

35. Kang, W.; Xiang, Y.; Wang, F.; Wan, L.; You, H. Flood detection in gaofen-3 SAR images via fully convolutional networks. *Sensors* **2018**, *18*, 2915. [CrossRef]

36. Gao, H.; Wang, C.; Wang, G.; Zhu, J.; Tang, Y.; Shen, P.; Zhu, Z. A crop classification method integrating GF-3 PolSAR and Sentinel-2A optical data in the Dongting Lake Basin. *Sensors* **2018**, *18*, 3139. [CrossRef] [PubMed]

37. Freeman, A. SAR calibration: An overview. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 1107–1121. [CrossRef]

38. Choi, H.; Jeong, J. Despeckling images using a preprocessing filter and discrete wavelet transform-based noise reduction techniques. *IEEE Sens. J.* **2018**, *18*, 3131–3139. [CrossRef]

39. Xu, C.; Su, H. RSDD-SAR: Rotated Ship Detection Dataset in SAR Images. *J. Radars* **2022**, *11*, 581–599.

40. Jaybhay, J.; Shastri, R. A study of speckle noise reduction filters. *Signal Image Process. Int. J. (SIPIJ)* **2015**, *6*, 71–80. [CrossRef]

41. Wang, X. Lee filter for multiscale image denoising. In Proceedings of the 2006 8th International Conference on Signal Processing, Guilin, China, 16–20 November 2006; Volume 1.

42. Saxena, N.; Rathore, N. A review on speckle noise filtering techniques for SAR images. *Int. J. Adv. Res. Comput. Sci. Electron. Eng. (IJARCSEE)* **2013**, *2*, 243–247.

43. Deepa, N.; Nagarajan, N. Kuan noise filter with Hough transformation based reweighted linear program boost classification for plant leaf disease detection. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *12*, 5979–5992. [CrossRef]

44. Mullissa, A.; Vollrath, A.; Odongo-Braun, C.; Slagter, B.; Balling, J.; Gou, Y.; Gorelick, N.; Reiche, J. Sentinel-1 sar backscatter analysis ready data preparation in google earth engine. *Remote Sens.* **2021**, *13*, 1954. [CrossRef]

45. Kulkarni, S.; Kedar, M.; Rege, P.P. Comparison of Different Speckle Noise Reduction Filters for RISAT-1 SAR Imagery. In Proceedings of the 2018 International Conference on Communication and Signal Processing (ICCSP), Tamilnadu, India, 3–5 April 2018; pp. 0537–0541.

46. Frulla, L.; Milovich, J.; Karszenbaum, H.; Gagliardini, D. Radiometric corrections and calibration of SAR images. In Proceedings of the IGARSS'98. Sensing and Managing the Environment. 1998 IEEE International Geoscience and Remote Sensing. Symposium Proceedings (Cat. No. 98CH36174), Seattle, WA, USA, 6–10 July 1998; Volume 2, pp. 1147–1149.

47. Manjusree, P.; Prasanna Kumar, L.; Bhatt, C.M.; Rao, G.S.; Bhanumurthy, V. Optimization of threshold ranges for rapid flood inundation mapping by evaluating backscatter profiles of high incidence angle SAR images. *Int. J. Disaster Risk Sci.* **2012**, *3*, 113–122. [CrossRef]

48. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

49. Yuan, K.; Zhuang, X.; Schaefer, G.; Feng, J.; Guan, L.; Fang, H. Deep-learning-based multispectral satellite image segmentation for water body detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7422–7434. [CrossRef]

50. Torralba, A.; Russell, B.C.; Yuen, J. Labelme: Online image annotation and applications. *Proc. IEEE* **2010**, *98*, 1467–1484. [CrossRef]

51. Lv, S.; Meng, L.; Edwing, D.; Xue, S.; Geng, X.; Yan, X.H. High-Performance Segmentation for Flood Mapping of HISEA-1 SAR Remote Sensing Images. *Remote Sens.* **2022**, *14*, 5504. [CrossRef]

52. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.

53. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin, Germany, 2006; Volume 4.

54. Guo, M.H.; Lu, C.Z.; Hou, Q.; Liu, Z.; Cheng, M.M.; Hu, S.M. Segnext: Rethinking convolutional attention design for semantic segmentation. *arXiv* **2022**, arXiv:2209.08575.

55. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.

56. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.

57. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.

58. Fan, Z.; Hou, J.; Zang, Q.; Chen, Y.; Yan, F. River Segmentation of Remote Sensing Images Based on Composite Attention Network. *Complexity* **2022**, *2022*, 7750281. [CrossRef]

59. Wu, X.; Zhan, C.; Lai, Y.K.; Cheng, M.M.; Yang, J. Ip102: A large-scale benchmark dataset for insect pest recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–17 June 2019; pp. 8787–8796.