remote sensing

MDPI

Article

Paddy Rice Inundation Mapping Based on Multi-Resolution Remote-Sensing Imagery and Semantic Segmentation with a Dual-Attention and Edge-Aware Attention Network

Yuhao Jin [1,†], Qinan Bian [1], Xing Mao [2,†], Min Wang [1], Jialong Sun [3], Hengkang Zhu [4] and Jinyao Lin [5,*]

[1] College of Water Conservancy and Civil Engineering, South China Agricultural University, Guangzhou 510642, China; yuhao.jin@scau.edu.cn (Y.J.); bianqinan@stu.scau.edu.cn (Q.B.); 20253182018@stu.scau.edu.cn (M.W.)
[2] Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China; maoxing@jaas.ac.cn
[3] PowerChina Eco-Environmental Group Co., Ltd., Shenzhen 518101, China; sunjl-shj@powerchina.cn
[4] School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China; zhuhk7@mail2.sysu.edu.cn
[5] School of Geography and Remote Sensing, Guangzhou University, Guangzhou 510006, China
* Correspondence: ljy2012@gzhu.edu.cn
† These authors contributed equally to this work.

Highlights

What are the main findings?

- DAEA-Net achieved the highest OA, IoU, F1-score, and Kappa coefficient among the compared models; the edge-aware attention block (EAAB) yielded the largest single-module gain in the ablation experiments.
- Under the AOI (potential flood-prone area) constraint, PH-MCW achieved an F1-score of 0.821 on the NDRCC dataset and up to 0.764 on the AIR-PolSAR-Seg-2.0 dataset.

What are the implications of the main findings?

- Coupling high-resolution paddy rice mapping with multi-source SAR water extraction provides spatially refined evidence for verifying flood-affected paddy rice in complex terrain.
- The field-scale inundation maps provide spatially explicit evidence for flood-loss verification and support targeted post-disaster decisions, such as drainage deployment, agricultural insurance claim verification, and replanting planning.

Abstract

Timely and accurate mapping of flood inundation extents in rice-growing regions is essential for flood damage assessment. However, fragmented fields and heterogeneous land cover in complex terrain continue to pose challenges. This study proposes a two-stage mapping framework that integrates multi-resolution remote sensing imagery with deep-learning-based semantic segmentation. First, a dual-attention and edge-aware attention network (DAEA-Net) was developed to derive a high-accuracy paddy rice map from very high-resolution (VHR) imagery. Second, the flood extent in paddy fields was delineated using the prior-histogram marker-controlled watershed (PH-MCW) method on SAR imagery with area-of-interest (AOI) filtering based on terrain slope and on the permanent water bodies and potentially inundated land types of the land-type survey data. On the test set, DAEA-Net achieved the highest OA, IoU, F1-score, and Kappa among the compared semantic segmentation models, including U-Net, HRNet, and DeepLabV3+, with

Academic Editor: John Trinder

Received: 18 September 2026
Revised: 20 October 2026
Accepted: 20 October 2026
Published: 21 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

values of 0.9205, 0.8747, 0.9331, and 0.8352, respectively. PH-MCW segmentation with AOI filtering, which removes terrain-induced false alarms while retaining water surfaces, improved flood-water detection, achieving F1-scores of 0.821 and 0.764 on two flood-validation reference datasets (NDRCC and AIR-PolSAR-Seg-2.0), respectively. The proposed framework provides a semi-automated solution for mapping flood inundation over paddy rice fields in complex terrain, in which the prior water proportion is refined by local inspection before application.

Keywords: rice mapping; flood inundation; semantic segmentation; synthetic aperture radar (SAR); very-high-resolution (VHR) optical imagery

1. Introduction

Rice is a staple food crop worldwide, and its stable production underpins regional food security [1]. Extreme rainfall and the resulting floods can inundate farmland, hinder crop growth, and damage agricultural infrastructure; timely delineation of affected areas is therefore an important basis for disaster-loss assessment and post-disaster recovery [2]. In South China, mountains, hills, valley basins, and coastal plains are interleaved, and paddy rice fields are widely distributed in relatively flat areas such as river valleys, alluvial plains, and coastal lowlands. These fields are characterized by fragmentation, diverse land-cover types, and pronounced spatial heterogeneity [3]. Complex landforms and land-use patterns increase the difficulty of fine-scale paddy rice identification, while flood impacts exhibit marked spatial variation. Timely delineation of paddy rice areas and their flood inundation extents in complex terrain is therefore fundamental to spatial assessment of rice losses and agricultural disaster verification.

Optical remote-sensing imagery records surface reflectance in the visible and near-infrared spectral regions, where the spectral differences among crops and between crops and other land-cover types are directly expressed, making it an important data source for crop identification and mapping [4]. At medium and high spatial resolutions, Landsat imagery provides a stable basis for long-term crop classification [5], and model-transfer approaches based on Landsat have further broadened its applicability [6]. Sentinel-2, with its higher spatiotemporal resolution and red-edge bands, has expanded the capability of crop monitoring [7] and shows better classification performance than Landsat in fragmented landscapes [8]. For paddy rice mapping, Sentinel-2 has been widely used [9] and can distinguish complex cropping systems such as single-season rice, double-season rice, and ratoon rice [10]. Very-high-resolution (VHR) optical imagery can depict field boundaries, geometric shapes, and small-scale features such as ridges and ditches in greater detail, providing strong spatial representation in fragmented and interspersed landscapes [11,12]. In such landscapes, superpixel segmentation can preserve field-boundary integrity while suppressing complex background interference [12]; by contrast, object-based image analysis reduces intra-class heterogeneity through segmentation units [13]. However, increased spatial detail is accompanied by stronger spatial heterogeneity: within-class appearance differences and between-class similarities become more prominent, making it difficult for traditional classifiers relying on hand-crafted features to maintain stable discrimination in complex scenes [11]. More generalizable automatic feature-learning methods are therefore needed.

In contrast to traditional supervised classification methods, which rely on hand-crafted features, deep learning enables end-to-end learning of hierarchical and transferable features directly from VHR imagery, thereby improving mapping generalization in complex farmland scenes [11]. At the application level, previous studies have used rice

phenology as a key feature for regional monitoring of planting dynamics [14] and large-scale rice mapping across climatic zones [15], and have further introduced specialized semantic segmentation networks [16], temporal convolutions and recurrent networks [17,18], and phenology-guided deep learning frameworks [19,20] to progressively improve the representation of crop growth dynamics and complex farmland scenes. Recently, Transformer architectures based on self-attention have been preliminarily applied to rice identification by exploiting their capacity to model global context [21]. Furthermore, semantic segmentation models, from fully convolutional networks to various Transformer variants, have been increasingly applied in remote sensing [22–25]. However, in fragmented fields and complex landforms with pronounced scale differences, simultaneously preserving fine field boundaries and exploiting global context remains difficult.

Persistent cloud and rain during flood events limit effective optical observation. Synthetic aperture radar (SAR), with its all-day and all-weather observation capability, provides an important data source for dynamic monitoring of flood water and inundated farmland [26–30]. Existing studies have developed threshold-based segmentation, change detection, and statistical model approaches for flood-water identification from SAR imagery [26–28]. Rice is highly sensitive to inundation, and flood duration is a key determinant of yield loss [31]. The near-real-time observation capability of SAR enables flood-water information to be obtained under cloudy and rainy conditions; translating this information into timely maps of inundated paddy rice remains a key challenge for rapid disaster-loss assessment.

To address the above challenges in previous studies, we constructed a two-stage framework linking paddy rice extraction with flood inundation identification. Jiang et al. [3] previously established a two-stage framework that couples multi-resolution remotely sensed imagery with semantic segmentation, in which a cropland mask extracted from VHR imagery is used as an area of interest to constrain subsequent flood mapping in mountain-dominated terrain. Following this framework, this study extends the mapping target from a single land-use class to the extraction of paddy rice as a specific crop class within it. Singha et al. [29] combined monthly 10 m Sentinel-1 flood maps with an independently produced 10 m Sentinel-1 paddy rice map to estimate flood-affected rice areas across Bangladesh during 2014–2018. Their study demonstrated the value of integrating crop maps with SAR-derived flood maps for long-term, country-scale assessment, while assuming that no drastic changes in paddy rice distribution occurred during the study period. In contrast, the present study constructs a 0.5 m paddy rice mask from VHR optical imagery supplemented with seasonal NDVI and terrain data, and combines it with event-specific, multi-source SAR water extraction and terrain- and permanent-water-based constraints. The present framework therefore focuses on field-level delineation and false-alarm suppression in fragmented and topographically complex rice-growing landscapes, complementing the large-scale monitoring approach of Singha et al. [29]. Because the spectral and textural signatures of paddy rice partly overlap with those of other crops, distinguishing it from other crops benefits from seasonal phenological cues [15,18], together with segmentation components whose multi-scale attention and edge-aware design are adapted to the fragmented, small-field geometry of paddy rice [16,21]. To meet this requirement, a dual-attention and edge-aware attention network (DAEA-Net) for paddy rice extraction was proposed. DAEA-Net adopts DeepLabV3+ [32] as the baseline framework and introduces three modules; among them, the adaptive multi-scale attention module extends the split-fuse-select idea of Selective Kernel Networks [33] from two or three kernel branches to seven parallel atrous and pooling branches, the edge-aware attention block (EAAB) follows the MetaFormer design [34] but replaces the token mixer with multi-directional edge operators and local spatial attention instead of self-attention,

and the dual-attention module is inspired by the attention mechanisms of CBAM [35] and adapts them to the fusion of high- and low-level features in the decoder. First, Jilin-1 VHR optical imagery was used as the main data source and fused with seasonal normalized difference vegetation index (NDVI) and digital elevation model (DEM) data to extract high-resolution paddy rice areas. Second, multi-source GF-3 and Sentinel-1 SAR imagery was used to obtain water information during flood events. An area of interest (AOI) combining terrain slope with the permanent water bodies and the potentially inundated land types in the land-type survey data was applied to the extracted water, so that terrain-induced false alarms such as shadow and layover are removed while water surfaces and potentially inundated land are retained; the flood inundation extent of paddy rice fields was then obtained by differencing the pre-event and during-event water extents within the paddy rice map. By spatially associating high-spatial-resolution crop mapping with dynamic flood monitoring, the framework provides technical support for fine-scale identification of rice-affected areas in complex terrain.

2. Study Area and Data

2.1. Study Area

Guangdong Province is located in the subtropical monsoon zone, with abundant but unevenly distributed precipitation and interleaved mountains, hills, valley basins, and coastal plains. Four county-level regions, namely Yingde, Taishan, Lechang, and Pingyuan, were selected as the study areas. Together, they cover representative terrain and land-use combinations of paddy rice areas in Guangdong, including coastal plains, basin-valley alluvial plains, mountainous hills, and sloping-field-valley agricultural landforms. The contiguity, field size, and boundary morphology of paddy rice fields differ markedly among these geomorphic units. The locations and landforms of the four study areas are shown in Figure 1.

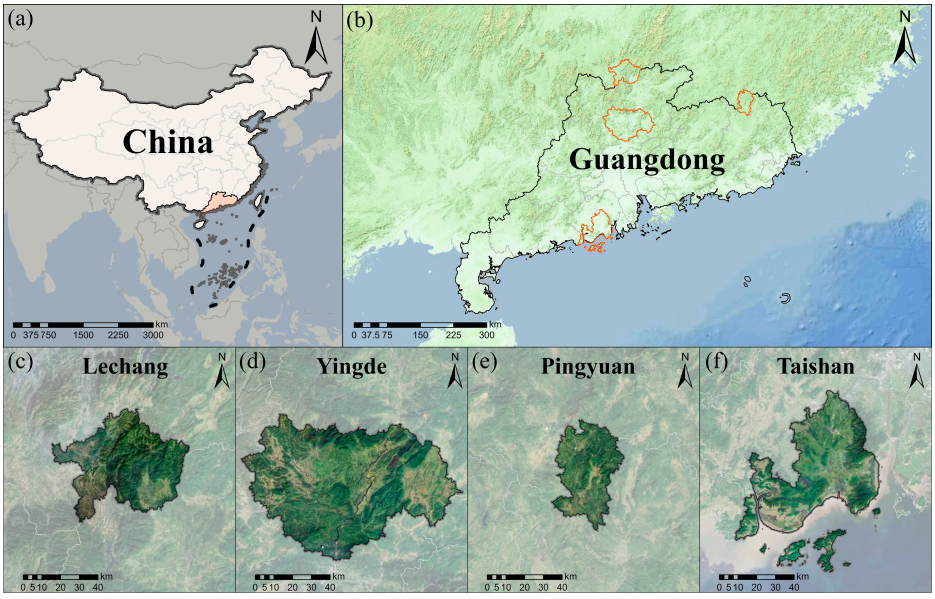


Figure 1. Location and landform overview of the study areas. (**a**) Location of Guangdong Province (highlighted in pink) within China; (**b**) distribution of the four study areas (outlined in orange) in Guangdong Province; (**c**) Lechang; (**d**) Yingde; (**e**) Pingyuan; and (**f**) Taishan, showing the terrain and landform of each study area.

Yingde is located in the southeastern part of the Nanling Mountains and the middle reaches of the Beijiang River; its paddy rice fields are concentrated along river valleys and

alluvial plains. Taishan is located in the southwest of the Pearl River Delta and adjacent to the South China Sea. Coastal plains, tidal flats, aquaculture ponds, and paddy rice fields are intermingled in Taishan, increasing the risk that ponds and tidal flats are mistaken for flood water. Lechang is located in the middle and upper reaches of the Wujiang River in northern Guangdong and is dominated by mountainous and hilly terrain; cultivated land is mostly distributed on valley terraces and hillside terraces. Pingyuan is located in the hilly valley region of northeastern Guangdong at the junction of Guangdong, Fujian, and Jiangxi provinces, where mountains and hills account for approximately 80.8% of the land area. The main regional flood hazards include Beijiang River flooding in Yingde, typhoon rainfall and storm surges in Taishan, and flash floods associated with heavy rainfall in the mountainous areas of Pingyuan and Lechang. The four study areas span coastal lowlands, the Beijiang basin, the mountainous hills of northern Guangdong, and northeastern hilly valleys. Their paddy fields differ in contiguity, field size, and boundary morphology and have different flood-response conditions, enabling evaluation of model adaptability to complex terrain, non-target low-backscattering features, and mountain shadows. The total area, cultivated land area, and mean annual precipitation of the four study areas are summarized in Table 1, with the corresponding official sources and land-survey years specified below the table.

Table 1. Total area, cultivated land area, and mean annual precipitation of the four study areas.

Study Area	Total Area (km²)	Cultivated Land Area (km²)	Mean Annual Precipitation (mm)
Yingde	5634.00	911.09	1835.80
Taishan	3308.25	482.24	1990.20
Lechang	2419.28	266.23	1522.60
Pingyuan	1381.00	109.05	1683.60

Sources: For Yingde, total area and cultivated land area were obtained from the 'Natural Environment' webpage of the Yingde Municipal People's Government [36], and mean annual precipitation from the 'Climate and Hydrology' section of Yingde Yearbook 2024 [37]. For Taishan, total area and mean annual precipitation were obtained from the 'Geography' section of Taishan Yearbook 2024 (p. 1) [38], and cultivated land area from its 'Cultivated Land Protection' section (p. 134) [38], which reports the area at the end of 2023. The precipitation value for Taishan represents the 1991–2020 average. For Lechang, total area was obtained from Lechang Yearbook 2024 (p. 42) [39], cultivated land area from the official publication 'Land Use Status Data from the 2024 Annual Land Change Survey of Lechang' [40], and mean annual precipitation from the 'Meteorology' section of Lechang Yearbook 2024 (p. 232) [39]. For Pingyuan, total area and mean annual precipitation were obtained from the 'Land Resources' and climate descriptions on the Pingyuan County People's Government website [41], and cultivated land area from the '3–2 Year-end Cultivated Land Area' table of Meizhou Statistical Yearbook 2024, which reports 163,575 mu (approximately 109.05 km²) of cultivated land at the end of 2023. The survey year for Yingde's cultivated land area is not specified on the source webpage. Areas reported in hectares and mu were converted to square kilometers.

2.2. Data Sources

2.2.1. Very-High-Resolution Optical Imagery

Paddy rice extraction was based on 0.5 m very-high-resolution optical imagery. The imagery was derived from the 2023 imagery basemap published by the Geographic Information Public Service Platform of Guangdong Province (Tianditu Guangdong, administered by the Department of Natural Resources of Guangdong Province), whose imagery is produced from the Jilin-1 satellite constellation, a Chinese high-resolution optical remote-sensing constellation composed of multiple satellites. The imagery contains red,

green and blue visible bands at 0.5 m resolution, enabling fine-grained representation of field boundaries, geometric shapes and ridge texture, and it served as the reference layer for sample annotation and model training. The imagery is used in accordance with the platform's copyright notice, and the source is acknowledged accordingly.

2.2.2. SAR Data

GF-3 is China's first civilian C-band multi-polarization SAR satellite. L1A single-look complex data in Fine Stripmap II (FSII) mode were obtained from the National Remote-Sensing Data and Application Service Platform, with a spatial resolution of 10 m; the cross-polarized HV channel was used in this study. For major flood events, L4 backscatter-coefficient products specific to each event were obtained from the National Comprehensive Earth Observation Data Sharing Platform. Sentinel-1 data [42] were obtained through the Copernicus Data Space Ecosystem (CDSE), using Ground Range Detected (GRD) products in Interferometric Wide Swath (IW) mode with VV + VH dual polarization and high resolution (HR); the cross-polarized VH channel was used in this study. These products are multi-looked (5 × 1 looks in range and azimuth) and have a spatial resolution of approximately 20 m × 22 m (range × azimuth), with a pixel spacing of 10 m × 10 m [43]. Fucheng-1 is a C-band SAR satellite with VV single polarization. The data used here were acquired in stripmap (SM) mode on 19 June 2024, with a nominal spatial resolution of approximately 3 m, and the orthorectified product was resampled to a 3 m pixel spacing. For each flood event, SAR images acquired both before and during the event were collected, covering the study area from 1 January 2022 to 31 December 2024.

2.2.3. Ancillary Data

Seasonal NDVI data and DEM data were used as supplementary inputs for rice extraction. NDVI data were retrieved from Sentinel-2 MSI multispectral imagery [44] at a spatial resolution of 10 m. Quarterly composites represented vegetation-cover dynamics at different stages of the rice growing season and provided key phenological information for distinguishing paddy rice from other dryland crops. The DEM had a spatial resolution of 5 m and was provided by the natural-resources authorities of Guangdong Province. The slope layer used in constructing the AOI was derived from the DEM. Township-level rainfall data used for the inundation-pattern analysis were obtained from CHM_PRE V2, a daily 0.1° gridded precipitation dataset for the Chinese mainland [45], and the township statistics cited in Section 4.4 were aggregated over each township area from this grid.

During the 2023 field survey, nine paddy rice sampling blocks were documented: three in Lechang, two in Pingyuan, and four in Taishan. Each block measured 2000 m × 2000 m, and the locations of the blocks are provided in Table A1. The nine sampling blocks were each confirmed by a field-survey point (nine points in total), and the on-site survey was conducted in December 2023. Field-survey data of paddy rice areas in Guangdong Province were used to annotate training samples for semantic segmentation and to cross-validate the results, providing independent ground truth for pixel-level accuracy assessment. The existing CCD-Rice and EARice10 datasets were used as public comparison products for a qualitative comparison of the spatial pattern and boundary representation of the extracted results.

To validate the accuracy and reliability of flood-inundation extraction, three datasets were adopted: two reference datasets for the pixel-level accuracy assessment and one multisensor image dataset for the cross-sensor assessment. A dataset of the June 2022 Yingde flood extent, produced by the National Disaster Reduction Center of China (NDRCC) of the Ministry of Emergency Management and distributed through the National Earth Observation Data Center (NODA), was interpreted from multi-source remote-sensing imagery acquired during the event and contains vector boundaries of inundated areas [46].

It contains three GF-3 scenes acquired on 19, 23 and 25 June 2022, and the reference water map was interpreted from the scene of 23 June 2022, the acquisition closest to the flood peak. The public AIR-PolSAR-Seg-2.0 dataset is a large-scale polarimetric SAR terrain-classification dataset built on GF-3 L1A images, of which the subset used in this study is located in the Foshan area of Guangdong Province and was acquired in November 2016 [47]. It was designed for deep-learning-based land-cover mapping and contains a water class, which was used here to support the evaluation of the water-extraction algorithm on stable water bodies. In addition, a flood-disaster remote-sensing image dataset for Meizhou, Guangdong, released by the National Earth Observation Data Center in 2024, provides contemporaneous imagery of the three sensors (GF-3, Sentinel-1 and Fucheng-1) and was used for the cross-sensor assessment. The land-type survey data provided by the natural-resources authorities of Guangdong Province cover the whole province, including both the four study areas and the Foshan validation area used for the AIR-PolSAR-Seg-2.0 assessment.

The eight input bands of DAEA-Net consist of three visible bands (red, green and blue) from the 0.5 m imagery, four quarterly NDVI layers derived from Sentinel-2 MSI, and one elevation band from the 5 m DEM (Table 2). All auxiliary layers were resampled to 0.5 m by nearest-neighbor interpolation so that the eight bands are spatially co-registered with the very-high-resolution imagery on the same grid. Nearest-neighbor interpolation was adopted because it preserves the original pixel values of the NDVI and elevation data and avoids the spurious intermediate values that bilinear or cubic interpolation would introduce into phenological and terrain gradients.

The quarterly NDVI layers were generated from Sentinel-2 MSI surface reflectance imagery in Google Earth Engine. Cloud, cloud-shadow, snow and water pixels were removed with an SCL-based mask, after which NDVI was computed as (B8 − B4)/(B8 + B4) and composited by taking the mean of all valid observations within each quarter of 2023 (January–March, April–June, July–September and October–December) at 10 m resolution. Because paddy fields flooded at transplanting can be classified as water by the SCL mask, their NDVI values could be biased during those periods; the quarterly mean compositing over all valid observations mitigates, but does not fully remove, this effect.

The 5 m DEM was provided by the natural-resources authorities of Guangdong Province and was used to derive elevation and slope terrain factors. The elevation values were normalized to [0, 1] by min–max normalization before being fed to the network.

Table 2. Composition, source and processing of the eight input bands of DAEA-Net.

Input Layer	Bands	Source	Processing	Role in the Model
VHR visible imagery	3 (R, G, B)	Jilin-1 constellation; Tianditu Guangdong 2023 basemap; 0.5 m	Provincial ortho-mosaic, 2023	Spectral and textural field-boundary features
Quarterly NDVI	4	Sentinel-2 MSI surface reflectance; 10 m	SCL-masked; NDVI = (B8 − B4)/(B8 + B4); quarterly mean, 2023	Seasonal phenology separating paddy rice from other crops
DEM	1	Natural-resources authorities of Guangdong Province; 5 m	Min–max normalization to [0, 1]	Terrain and slope features

3. Methods

This study proposes a two-stage framework for extracting flood inundation over paddy rice fields. First, Jilin-1 VHR optical imagery is fused with seasonal NDVI phenological information and DEM data. An adaptive multi-scale attention module, an edge-aware attention block (EAAB), and a dual-attention module are introduced into the

DeepLabV3+ baseline to construct DAEA-Net, enhancing multi-scale rice-feature representation and field-boundary segmentation in complex farmland scenes. Second, water and non-water seeds are marked on the multi-source SAR imagery from the prior water percentage and the histogram peaks; marker-controlled watershed segmentation is applied to extract water, and the AOI mask is then applied to the extracted water to remove terrain-induced false alarms; the flood-inundated paddy rice area is finally obtained by differencing the pre-event and during-event water extents within the paddy rice map. The overall framework is illustrated in Figure 2.

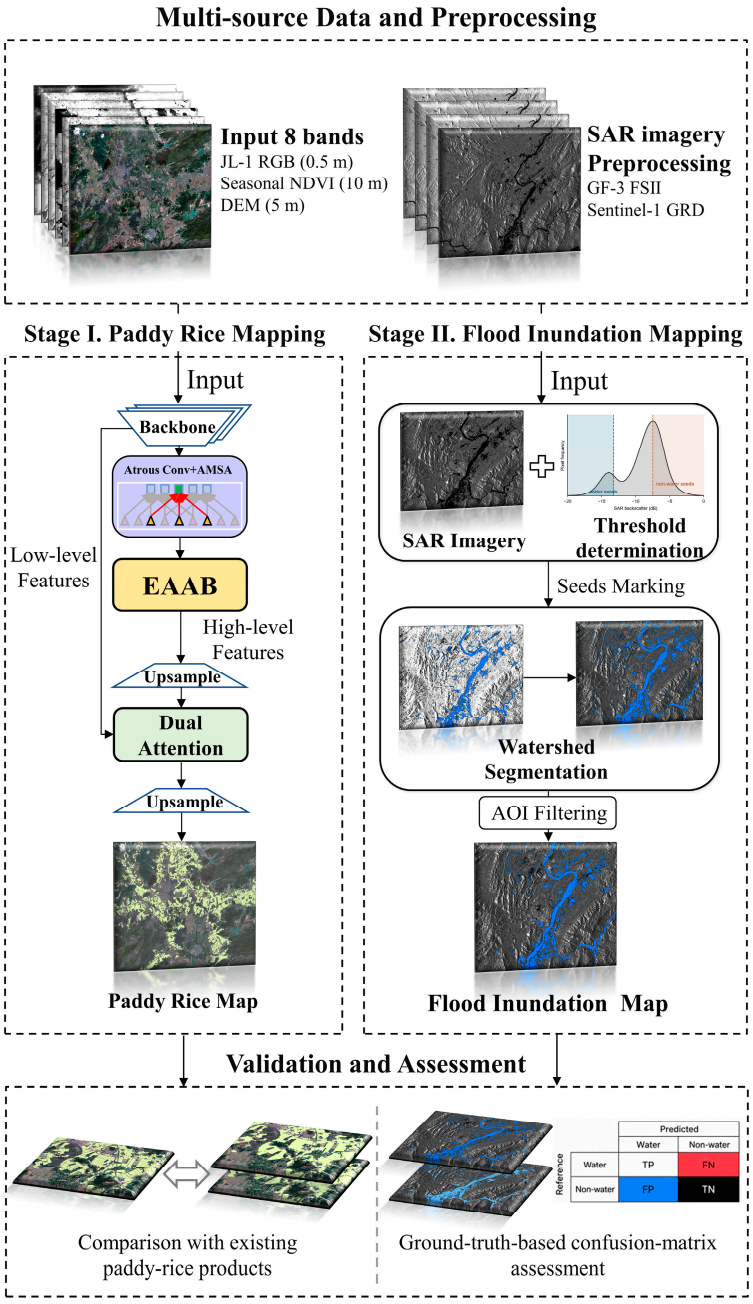


Figure 2. Framework of paddy rice inundation mapping.

3.1. Semantic Segmentation-Based Paddy Rice Extraction

3.1.1. Sample Construction and Dataset Partitioning

High-quality sample annotation is the basis for semantic-segmentation model training. Field surveys combined with expert visual interpretation were used to annotate paddy rice samples. Nine representative paddy rice sampling blocks, each 2000 m × 2000 m in size, were selected within three of the four study areas (Lechang, Pingyuan and Taishan), covering paddy field distribution patterns in mountainous and hilly terrain, hilly valleys, and coastal plains, which, together, account for the great majority of the landforms found across the four study areas. Although the blocks cover only 36 km^2, a small fraction of the more than 11,000 km^2 of the four counties, they were selected purposively to capture the main geomorphic and field-morphology types of paddy rice in the region. No sampling block was located in Yingde, so the rice map of Yingde is an extrapolation from the model trained on the other three counties; the implications are discussed in Section 5.4.

A sliding-window cropping strategy with a stride equal to the tile size was then used to cut these blocks into 1425 image tiles of 256 × 256 pixels with their corresponding annotation masks, so that adjacent tiles do not overlap and no pixel is shared between any two tiles. The dataset was randomly split in a 7:2:1 ratio into training, validation, and test sets containing 998, 285, and 142 tiles, respectively. To enlarge the training set and improve generalization [48], four geometric transformations—vertical flip, horizontal flip, vertical flip combined with a 90° counterclockwise rotation, and horizontal flip combined with a 90° counterclockwise rotation—were applied to the training set, yielding 4990 training samples. The validation and test sets were left unchanged and used solely for hyperparameter selection and final accuracy evaluation. Examples of the image tiles and their annotation masks are shown in Figure 3.

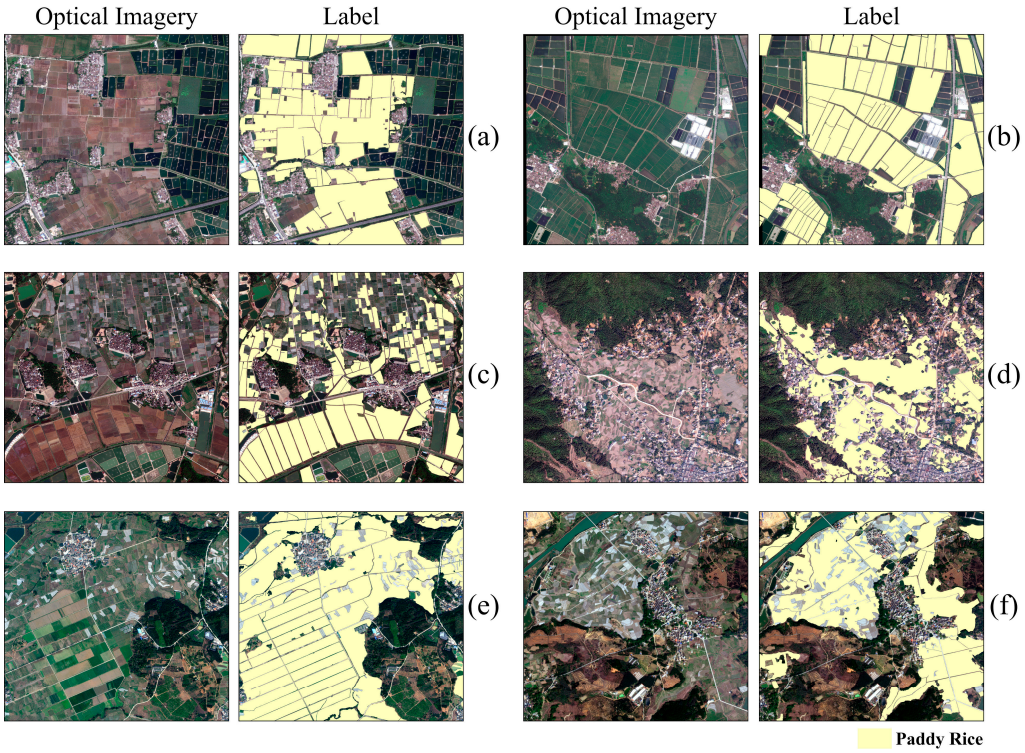


Figure 3. Typical VHR optical-image tiles and corresponding paddy rice annotation masks from the study areas. (**a**–**f**) Six representative tiles; in each pair, the left panel shows the VHR optical imagery and the right panel shows the corresponding annotation overlaid on the same image, in which yellow denotes paddy rice.

3.1.2. Semantic Segmentation Model with Dual-Attention and Edge-Aware Attention

For complex farmland remote-sensing scenes, the encoder–decoder structure of DeepLabV3+ [32] was used as the baseline framework. This segmentation paradigm can be traced back to fully convolutional networks (FCNs) [49]. DAEA-Net was constructed using a multi-branch heterogeneous feature-extraction architecture that fuses the spectral and textural information of VHR optical imagery with complementary NDVI and DEM features; the eight bands are stacked at the input and processed jointly (early fusion), and the parallel branches shown in Figure 4 are the seven AMSA branches, which work at different scales on the same fused features rather than on different data sources. DAEA-Net successively incorporates the adaptive multi-scale attention module (AMSA), the edge-aware attention block (EAAB), and the dual-attention module (DualA). DualA draws on the channel-spatial attention structure of the convolutional block attention module (CBAM) [35]. These modules strengthen adaptive multi-scale aggregation, field-boundary modeling, and refined fusion of high- and low-level features, respectively. The overall structure is shown in Figure 4.

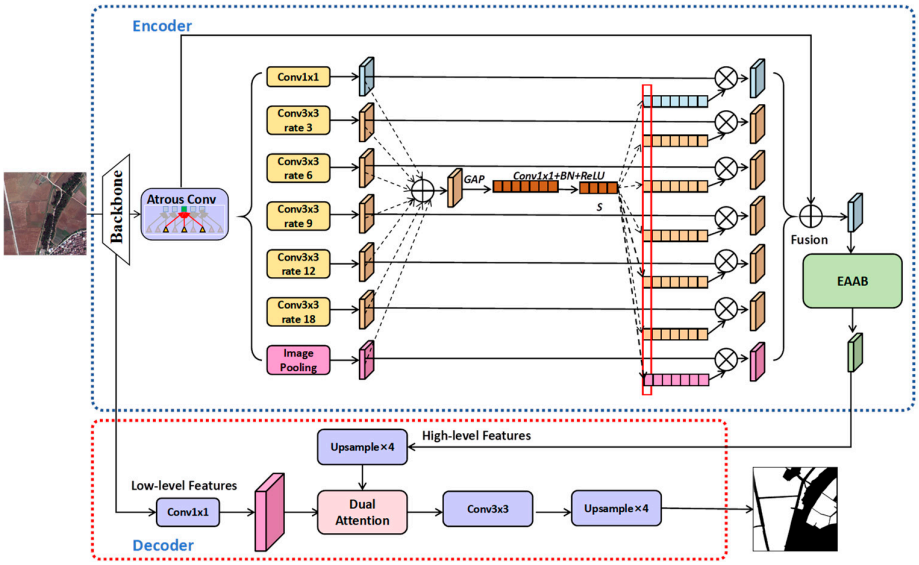


Figure 4. Overall architecture of the proposed DAEA-Net. The blue and red dotted frames indicate the encoder and the decoder, respectively. The seven parallel branches in the encoder (yellow boxes) form the AMSA, consisting of a 1 × 1 convolution, 3 × 3 atrous convolutions with dilation rates of 3, 6, 9, 12 and 18, and image pooling. ⊕ and ⊗ denote element-wise addition and multiplication, respectively. The dashed arrows show that the statistical vector S produces the channel-wise adaptive weight w_i of each branch, and the red vertical frame marks these weights.

The AMSA is a task-adapted selective-kernel aggregation unit inspired by the basic selective-kernel adaptive-fusion principle [33]. Different from the original Selective Kernel (SK) Network, which adaptively selects convolution kernel sizes, AMSA achieves multi-scale receptive-field adaptation by employing atrous (dilated) convolutions with different dilation rates together with pooling to form seven parallel branches. Multi-scale features with different receptive fields (i = 1, 2, …, 7) are extracted and denoted as F_i. The seven branch features are then summed element-wise to form the fused feature F_{fuse} (Equation (1)). Global average pooling (GAP) is applied to F_{fuse} to aggregate global spatial information and produce the global context vector F_{GAP} (Equation (2)); a 1 × 1 convolution with learnable parameters W_c, followed by batch normalization and ReLU activation δ, produces the statistical vector S (Equation (3)). Branch-specific projections parameterized by A_i and a softmax operation generate the channel-wise adaptive weight w_i for each branch

(Equation (4)). The weighted branch features are then summed and added to the original input feature F_{in} to obtain the AMSA output F_{out} (Equation (5)), enabling adaptive multi-scale context fusion. Unlike the fixed concatenation pattern used by the atrous spatial pyramid pooling (ASPP) branches of DeepLabV3+ [32], these weights vary dynamically with the input content and adjust the contribution of each receptive-field branch according to field-scale differences.

$$F_{\text{fuse}} = \sum_{i=1}^{7} F_i \tag{1}$$

$$\tilde{F}_{\text{GAP}} = \text{GAP}(F_{\text{fuse}}) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_{\text{fuse}}(i,j) \tag{2}$$

$$S = \delta(\text{BN}(W_c \cdot \tilde{F}_{\text{GAP}})) \tag{3}$$

$$w_i = \frac{e^{A_i S}}{\sum_{j=1}^{7} e^{A_j S}} \tag{4}$$

$$F_{\text{out}} = F_{\text{in}} + \sum_{i=1}^{7} w_i F_i \tag{5}$$

where F_{in} denotes the original input feature of the AMSA; F_i denotes the feature generated by the i-th multi-scale branch; F_{fuse} denotes the element-wise sum of the seven branch features; $\tilde{F}_{GAP}$ denotes the global context vector obtained by global average pooling; H and W denote the spatial height and width of the feature map, respectively; GAP denotes global average pooling; W_c denotes the learnable parameters of the 1 × 1 channel-compression layer; BN denotes batch normalization; δ denotes the ReLU activation function; A_i denotes the learnable projection parameters for the i-th branch; and w_i denotes the normalized adaptive weight of the i-th branch, with $\Sigma w_i = 1$.

The EAAB module is used to recover field boundaries and enhance edge-aware feature representation. The module first extracts multi-directional edge features using horizontal, vertical, 45°, and 135° Sobel operators, together with a Laplacian operator, and concatenates them along the channel dimension; the Sobel and Laplacian edge kernels are fixed 3 × 3 operators applied depthwise (channel-wise) to the feature maps, and the concatenated edge features are fused by a 1 × 1 convolution, followed by batch normalization and ReLU activation. This captures both coarse-edge responses from first-order gradients and fine-edge responses from second-order derivatives. To strengthen effective edges and suppress noise, average pooling and max pooling are applied separately; the outputs are concatenated and passed through a 7 × 7 convolution and sigmoid activation to generate an edge-spatial attention map (Equation (7)), which is multiplied element-wise with the edge feature to obtain enhanced edge features. The enhanced edge features are then added to the original features and passed through a MetaFormer-style residual block composed of LayerNorm and a multilayer perceptron (MLP); LayerNorm is computed over the channel dimension for each pixel position. By integrating multi-directional gradient information with spatial attention, the EAAB module adaptively highlights boundary-relevant features while suppressing irrelevant background responses, helping the decoder recover linear boundaries such as ridges and ditches and complete field contours (Figure 5).

$$F_{\text{edge}} = \text{Conv}_{1\times1}(\text{Concat}(G_{\text{x}}, G_{\text{y}}, G_{45}, G_{135}, G_{\text{lap}})), \tag{6}$$

$$M_{\text{edge}} = \sigma(\text{Conv}_{7\times7}(\text{Concat}[P_{\text{avg}}(F_{\text{edge}}), P_{\text{max}}(F_{\text{edge}})])), \tag{7}$$

where G_x, G_y, G_{45}, and G_{135} denote the edge feature maps extracted by the horizontal, vertical, 45°, and 135° Sobel operators, respectively; G_{lap} denotes the edge feature map extracted by the Laplacian operator; F_{edge} denotes the fused edge features; P_{avg} and P_{max} denote the average-pooled and max-pooled edge features, respectively; and M_{edge} denotes the edge-spatial attention map.

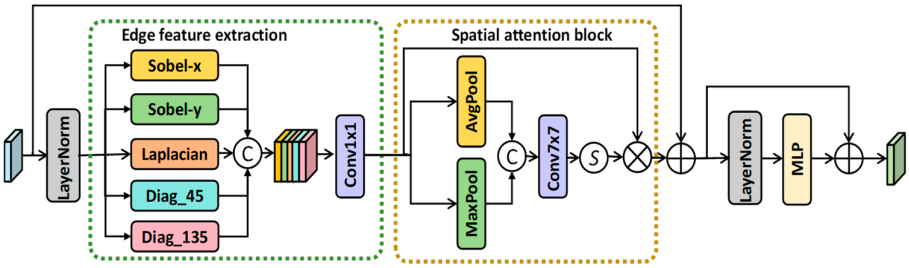


Figure 5. Architecture of the edge-aware attention block (EAAB).

DualA is inspired by the spatial and channel attention mechanisms of CBAM, but is customized for decoder-side cross-level feature fusion for paddy rice segmentation. Rather than introducing a fundamentally new attention architecture, its main improvements lie in the pre-recalibration cross-level fusion design. DualA fuses high-level semantic features and shallow spatial features in the decoder. The two features are first aligned in the channel dimension by 1 × 1 convolutions and added element-wise to obtain a fused feature F_{add} (Equation (8)); feature recalibration is then performed sequentially through spatial-attention and channel-attention stages, and the max-pooling and average-pooling operations in the CBAM spatial branch are removed. In the spatial-attention stage, the fused feature is activated by ReLU, passed through a 1 × 1 convolution, and processed by sigmoid activation to generate the two-dimensional spatial-attention map M_s (Equation (9)). The spatially weighted feature F_{spa} is then obtained by element-wise multiplication of M_s and F_{add} (Equation (10)), highlighting paddy rice regions and suppressing background interference. Unlike CBAM, which computes the channel descriptor from the raw input feature, the spatially weighted feature F_{spa} is subsequently aggregated by GAP to form the channel descriptor D (Equation (11)). Channel weights are generated through an MLP and sigmoid activation (Equation (12)), multiplied element-wise with F_{spa}, and followed by batch normalization (another component missing in CBAM) to produce the fused output, F_{out}. The decoder subsequently upsamples this feature stage by stage, restores spatial details, and uses a prediction convolution and sigmoid activation to output pixel-wise paddy rice and background probabilities; a decision threshold of 0.5 is applied to the sigmoid output to obtain the binary classification of paddy rice and background, as illustrated in Figure 6.

$$F_{\mathrm{add}} = \mathrm{Conv}_{1\times1}(F_{\mathrm{h}}) \oplus \mathrm{Conv}_{1\times1}(F_{\mathrm{l}}), \tag{8}$$

$$M_{\mathrm{s}} = \sigma(\mathrm{Conv}_{1\times1}(\mathrm{ReLU}(F_{\mathrm{add}}))), \tag{9}$$

$$F_{\mathrm{spa}} = M_{\mathrm{s}} \otimes F_{\mathrm{add}}, \tag{10}$$

$$D = \mathrm{GAP}(F_{\mathrm{spa}}) = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} F_{\mathrm{spa}}(i,j), \tag{11}$$

$$F_{\mathrm{out}} = \mathrm{BN}(F_{\mathrm{spa}} \otimes \sigma(\mathrm{MLP}(D))), \tag{12}$$

where F_{add} denotes the fused feature obtained by adding the channel-aligned high-level and shallow features; M_s denotes the two-dimensional spatial-attention map; F_{spa} denotes

the spatially weighted feature; ⊗ denotes element-wise multiplication; D denotes the channel descriptor obtained by global average pooling; and F_{out} denotes the output feature of DualA.

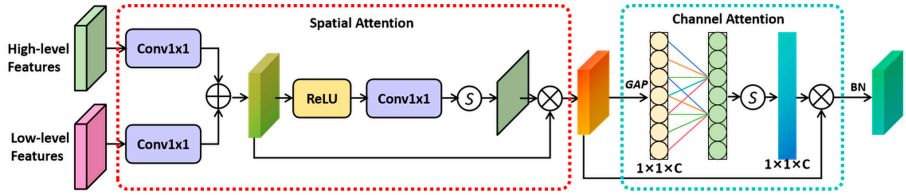


Figure 6. Architecture of the dual-attention module (DualA).

3.1.3. Experimental Settings for the Semantic Segmentation Model

In this study, MobileNetV2 was adopted as the backbone network of the proposed DAEA-Net. The model takes an eight-band input consisting of three visible bands, four quarterly NDVI layers, and one DEM band. The eight bands are concatenated at the input (early fusion), without separate branches for each data source. Because the eight-band input differs from the conventional three-band ImageNet data, ImageNet pre-trained weights were not applicable, and the MobileNetV2 backbone was therefore randomly initialized. For model training, binary cross-entropy (BCE) loss was used as the loss function, and the Adam optimizer was adopted with an initial learning rate of 1×10^{-4}. The network was trained for 120 epochs with a batch size of 8. To enlarge the training set to five-fold sample volume, the four geometric transformations described in Section 3.1.1 (vertical flip, horizontal flip, vertical flip combined with a 90° counterclockwise rotation, and horizontal flip combined with a 90° counterclockwise rotation) were applied to the training set. All experiments were implemented under the TensorFlow-Keras framework and executed on a workstation with an NVIDIA RTX 3090 GPU.

3.2. Flood Inundation Identification over Paddy Rice Fields

3.2.1. SAR Image Preprocessing

The GF-3 imagery requiring further processing was handled with the SARscape module in ENVI, including multi-looking to suppress speckle noise and homogenize pixel spatial resolution, single-channel intensity filtering, and geocoding to correct geometric distortion. The GF-3 products that had already undergone systematic radiometric correction and geocoding were imported directly for subsequent analysis. Sentinel-1 imagery was preprocessed using ESA SNAP 8.0. The standard preprocessing chain included precise orbit correction, thermal and border-noise removal, radiometric calibration, refined Lee speckle filtering, and terrain correction based on a 30 m SRTM DEM.

3.2.2. Construction of the Area-of-Interest (AOI) Mask

Two error sources significantly affect SAR water-extraction accuracy. First, the side-looking geometry of radar creates shadows on mountain backslopes and around tall buildings; their extremely low backscatter resembles water and can be misclassified as inundation. Second, permanent water bodies such as rivers, lakes, and reservoirs also exhibit low backscatter in both flood and non-flood periods and must be separated from the newly appeared flood water. We therefore constructed an area of interest (AOI), also referred to as the potential flood-prone area, which combines three elements with different roles. The first is terrain slope derived from the 5 m DEM: pixels with slopes greater than 25° are removed, because terrain-induced low backscatter such as shadow and layover is concentrated there. The second and third are the permanent water bodies and the land types that may be inundated, both taken from the land-type survey data provided by the

natural-resources authorities of Guangdong Province: these two elements are retained rather than removed, so that water surfaces and potentially inundated land in steep terrain, such as a river between high banks in a narrow valley, are not cut away by the slope constraint. The AOI is thus applied to the preliminary binary water map to remove false-alarm pixels while keeping all water present in the image, both permanent and temporary.

For the coastal county-level city of Taishan, the sea surface occupies a large share of the low-backscatter pixels and its extent varies with the tide and with the scene, so the prior percentile used for seed marking would be unstable. Before seed marking, the sea surface was therefore clipped to a fixed extent using the coastline contained in the water-body data, so that the share of low-backscatter pixels contributed by the sea is controlled; the same coastline is used when the AOI is constructed.

Stable (non-flood) water is subsequently separated from newly appearing flood water by the change-detection step described in Section 3.2.3. In the accuracy assessment of Section 3.3, the AOI was applied to the extracted water only and never to the reference data.

3.2.3. Marker-Controlled Watershed Water Segmentation with Prior-Knowledge Seeding and Histogram Analysis

Vincent and Soille [50] described an efficient algorithm for the watershed transform based on immersion simulations. The marker-controlled watershed (MCW) approach uses marker seeds to guide the flooding process and reduce the over-segmentation of traditional watershed algorithms caused by inherent SAR speckle noise [51]. MCW segmentation consists of three steps: identifying water and non-water seeds from image features, performing watershed segmentation on the backscatter intensity image, which is used as the topographic surface, and classifying unmarked pixels according to seed attributes.

For water extraction from multi-source SAR imagery, the absolute backscatter coefficients of the same object differ markedly among sensors because of differences in wavelength, polarization, incidence angle, and radiometric calibration. Fixed physical thresholds are therefore not transferable across sensors. We proposed the prior-histogram marker-controlled watershed (PH-MCW) method, which decomposes seed marking into two independent statistical-inference tasks. First, the backscatter percentile T_{water} is defined from the prior water proportion of the study region, and pixels below it are labeled as water seeds. The prior proportion was taken from publicly released land-survey data of the corresponding administrative region (inland water bodies only) [52–55] and was then refined by visual inspection of the water bodies present in each scene; the initial reference value and the value finally adopted are, respectively, 3.3% and 3.2% for Yingde, 2.2% and 2.2% for Lechang (no adjustment), 2.0% and 2.6% for Pingyuan, and approximately 3.6% and 8.4% for Taishan, where the larger adjustment reflects the residual sea surface that remains within the fixed coastal extent used before seed marking (Section 3.2.2). Second, an equi-width histogram of the valid backscatter values is computed with a fixed number of 256 bins, so that the bin width equals (max−min)/256 of the valid pixels of each scene, and the histogram is smoothed with a one-dimensional Gaussian filter (σ = 3 bins); $T_{\text{non-water}}$ is defined as the center of the bin at which the smoothed histogram reaches its global maximum, i.e., the mode, which represents the typical backscatter level of the non-water surfaces that occupy most pixels. The two thresholds are therefore derived from image statistics and prior geographic information rather than from a fixed physical value, which makes the method transferable across sensors, while the prior proportion can be refined locally where water bodies are more developed than the regional average. The cross-polarized channel was used for both sensors, VH for Sentinel-1 and HV for GF-3, because the specular return of water is weaker at cross-polarization.

In our implementation, no separate gradient image is computed. Because the signature of water in SAR is an absolute low backscatter return, the backscatter intensity itself is used as the topographic surface, so that water bodies correspond to low-elevation basins and the absolute scattering level is retained. A bidirectional watershed is then applied—a forward pass on the intensity image and a backward pass on its inverse—and pixels overlapping between the two passes are assigned by a Euclidean distance transform to the nearer seed set.

Because paddy fields are covered by irrigation water during part of the growing season and therefore appear as low-backscatter surfaces in SAR imagery, a single-date water map within the paddy rice map cannot by itself separate flood water from normal paddy water [56]. For each event, the water extent extracted from a pre-event SAR image acquired in the same growing season was therefore subtracted from that extracted from the during-event SAR image, and only water newly appearing within the paddy rice map was retained as flood-inundated paddy rice. The pre-event image for each event was the Sentinel-1 VH scene acquired closest to, and before, the event with full coverage of the study area: 2 June 2022 for Yingde, 13 September 2023 for Taishan, 22 May 2024 for Lechang, and 10 June 2024 for Pingyuan.

3.3. Accuracy Validation

Paddy rice extraction was evaluated at the pixel level on the test-set samples, with paddy rice samples annotated through field surveys and expert visual interpretation serving as ground truth. Six metrics were computed from the pixel-wise confusion matrix: overall accuracy (OA), precision, recall, F1-score, intersection over union (IoU), and the Kappa coefficient. Precision, recall, F1-score, and IoU were further reported for each class, whereas OA, precision, recall, IoU, F1-score, and the Kappa coefficient were used for the overall comparison among models. OA, the Kappa coefficient, and the expected agreement by chance (P_e) are defined in Equations (13)–(15), IoU in Equation (16), and precision, recall, and F1-score in Equations (17)–(19).

$$\mathrm{OA} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}, \tag{13}$$

$$\mathrm{Kappa} = \frac{\mathrm{OA}-P_e}{1-P_e}, \tag{14}$$

$$P_e = \frac{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})+(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}{(\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN})^2}, \tag{15}$$

$$\mathrm{IoU} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}}, \tag{16}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \tag{17}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \tag{18}$$

$$\mathrm{F1}-\mathrm{score} = 2 \times \frac{\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}}, \tag{19}$$

where TP, TN, FP, and FN denote the numbers of pixels correctly identified as paddy rice, correctly identified as non-paddy rice, non-paddy rice pixels misidentified as paddy rice, and missed paddy rice pixels, respectively; P_e is the expected agreement by chance, computed from the confusion-matrix counts as P_e = ((TP + FP)(TP + FN) + (TN + FP)(TN + FN))/(TP + FP + TN + FN)^2.

Flood inundation identification was validated against two reference datasets: the NDRCC dataset and the water class of the public AIR-PolSAR-Seg-2.0 dataset. Flood-water bodies extracted by PH-MCW segmentation from multi-source SAR imagery were

compared with these reference data. Five metrics were computed from the pixel-wise confusion matrix: OA, precision, recall, F1-score, and the Kappa coefficient. OA and the Kappa coefficient are calculated with Equations (13) and (14), and precision, recall, and F1-score with Equations (17)–(19), where the entries of the confusion matrix are counted in water pixels rather than paddy rice pixels. The reference water map and the SAR image used for extraction both refer to the GF-3 scene acquired on 23 June 2022, the acquisition closest to the flood peak, so that the accuracy assessment and the mapped inundation extent are temporally consistent with the reference data. Precision, recall, and F1-score reflect the reliability, completeness, and overall performance of flood-water identification, respectively. The comparison was performed over the intersection of the valid SAR data and the valid reference data, which is the same for all three methods and for both the unfiltered and filtered cases. The candidate flood-prone area was applied only to the extracted water, so that pixels outside it were treated as non-water, whereas the reference data were not clipped and retained all water classes, including rivers and reservoirs. In this accuracy assessment, the filtering was provided by the AOI (potential flood-prone area) and was applied to the extracted water only; the paddy rice mask was not involved, because the assessment is made at the level of water bodies. The paddy rice mask enters the framework only in the subsequent step of delineating the inundated area within paddy fields. To assess the effect of the AOI mask, F1-scores before and after AOI filtering were compared. In addition to the area-based metrics, the boundary accuracy of the extracted water was evaluated with the Boundary IoU (BIoU) proposed by Cheng et al. [57]. Let G and P denote the reference and predicted binary masks, respectively, and let G_d and P_d denote their boundary regions obtained by dilating the mask contours by d pixels, so that $G_d \cap G$ and $P_d \cap P$ are the boundary bands inside the two masks. The boundary bands are defined strictly inside the masks by the distance to the nearest non-water pixel, so the raster edges are not involved in the boundary-band construction and cannot be counted as boundaries. The Boundary IoU is then computed as the IoU of the two boundary bands:

$$\mathrm{BIoU}_d = \frac{|(G_d \cap G) \cap (P_d \cap P)|}{|(G_d \cap G) \cup (P_d \cap P)|} \tag{20}$$

Because the boundary band occupies only a thin ring around the water–land boundary, Boundary IoU is much more sensitive than area-based IoU to boundary errors and is therefore reported together with the area-based metrics. Two boundary-band widths, d = 2 and d = 5 pixels, were used.

4. Results

4.1. Performance of the Semantic Segmentation Model

DAEA-Net was quantitatively evaluated on the test set, which was randomly split from the 1425 tiles and is shared by all compared models. The evaluation was performed at the pixel level, and the paddy rice class and the background class were assessed separately. The resulting class-wise precision, recall, F1-score, and IoU are summarized in Table 3.

Table 3. Class-wise performance metrics of DAEA-Net on the test set.

Metric	Rice	Background
Precision	0.9297	0.9069
Recall	0.9366	0.8973
F1-score	0.9331	0.9021
IoU	0.8747	0.8216

Note: Metrics were computed from a single global confusion matrix obtained by summing the per-tile confusion matrices over all 142 (256 × 256) test tiles; precision, recall, F1-score, and IoU are reported class-wise for the rice and background classes.

The rice-class precision, recall, F1-score, and IoU were 0.9297, 0.9366, 0.9331, and 0.8747, respectively. The corresponding metrics for the background class were 0.9069, 0.8973, 0.9021, and 0.8216. The rice-class recall of 0.9366 indicates that 93.7% of the reference paddy rice pixels were correctly identified, while the precision of 0.9297 indicates that some background pixels were incorrectly classified as paddy rice. Background recall was lower than background precision, indicating that some background pixels were misclassified as paddy rice, mainly along the edges of small field ridges and field roads.

To verify the effectiveness of DAEA-Net, it was compared with U-Net [58], HRNet [59], and DeepLabV3+ on the same test set. Visual comparisons for representative sample tiles are shown in Figure 7, and quantitative results are given in Table 4. DAEA-Net achieved the highest OA, IoU, F1-score, and Kappa coefficient, with values of 0.9205, 0.8747, 0.9331, and 0.8352, respectively. Relative to U-Net, HRNet, and DeepLabV3+, its OA increased by 0.0273, 0.0148, and 0.0027, respectively, and its IoU increased by 0.0383, 0.0209, and 0.0052. Its recall was 0.9366, higher than 0.9226 for U-Net, 0.9297 for HRNet, and 0.9247 for DeepLabV3+. Its precision was 0.9297, slightly lower than the 0.9357 achieved by DeepLabV3+. In general, DAEA-Net improved classification consistency, spatial overlap, and recall, and yielded the highest Kappa coefficient.

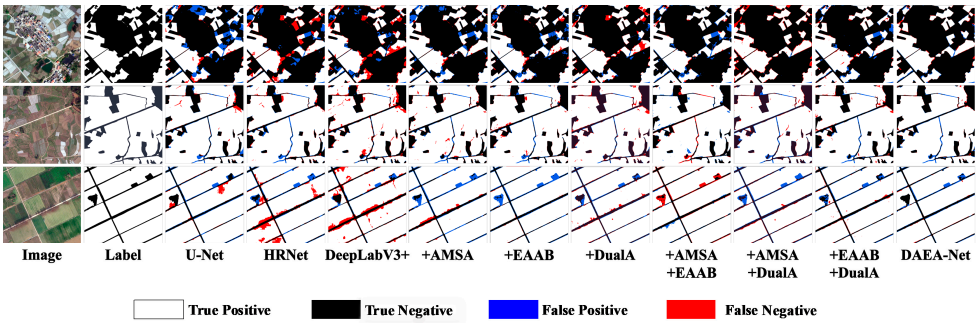


Figure 7. Comparison of paddy rice extraction results and pixel-wise classification outcomes for different model configurations on representative sample tiles.

Table 4. Comparison of semantic segmentation performance metrics among different models on the test set.

Model	OA	Precision	Recall	IoU	F1-score	Kappa
U-Net	0.8932	0.8996	0.9226	0.8364	0.9109	0.7776
HRNet	0.9057	0.9127	0.9297	0.8538	0.9211	0.8040
DeepLabV3+	0.9178	0.9357	0.9247	0.8695	0.9302	0.8303
DAEA-Net	0.9205	0.9297	0.9366	0.8747	0.9331	0.8352

Note: All metrics were computed from a single global confusion matrix obtained by summing the per-tile confusion matrices over all 142 (256 × 256) test tiles. Precision, recall, F1-score, and IoU refer to the rice (foreground) class, whereas OA and the Kappa coefficient are overall metrics over both classes.

Ablation experiments were designed to examine the independent contributions and synergies of the proposed modules (Table 5). Adding only the AMSA reduced OA from 0.9178 to 0.9093 and IoU from 0.8695 to 0.8590, although recall increased from 0.9247 to 0.9344. This indicates that multi-scale feature aggregation alone improves rice-area coverage but may introduce more false alarms.

Table 5. Ablation results of different DAEA-Net module combinations on the test set.

AMSA	EAAB	DualA	OA	Precision	Recall	IoU	F1-score	Kappa	*b*	*c*	*p*-Value
×	×	×	0.9178	0.9357	0.9247	0.8695	0.9302	0.8303	–	–	–
✓	×	×	0.9093	0.9141	0.9344	0.8590	0.9242	0.8111	355,252	434,911	$<10^{-6}$
×	✓	×	0.9191	0.9221	0.9430	0.8734	0.9324	0.8317	386,065	373,710	$<10^{-6}$
×	×	✓	0.9185	0.9445	0.9161	0.8694	0.9301	0.8324	358,992	352,304	$<10^{-6}$
✓	✓	×	0.9194	0.9210	0.9449	0.8741	0.9328	0.8322	393,015	377,893	$<10^{-6}$
✓	×	✓	0.9152	0.9293	0.9273	0.8662	0.9283	0.8245	351,489	375,686	$<10^{-6}$
×	✓	✓	0.9193	0.9399	0.9225	0.8712	0.9311	0.8335	359,729	345,950	$<10^{-6}$
✓	✓	✓	0.9205	0.9297	0.9366	0.8747	0.9331	0.8352	380,839	355,090	$<10^{-6}$

Note: All metrics were computed from a single global confusion matrix obtained by summing the per-tile confusion matrices over all 142 (256 × 256) test tiles. Precision, recall, F1-score, and IoU refer to the rice (foreground) class, whereas OA and the Kappa coefficient are overall metrics over both classes. The ××× row is the DeepLabV3+ baseline without the proposed modules; ✓ and × denote whether each module is enabled. In the McNemar columns, b and c are discordant-pair counts against the DeepLabV3+ baseline (b: misclassified by the baseline and correctly classified by the configuration; c: correctly classified by the baseline and misclassified by the configuration), and the p-value is obtained from the McNemar χ^2 test with continuity correction.

Adding only EAAB yielded an OA of 0.9191 and an IoU of 0.8734, the largest improvement among single-module configurations, demonstrating the importance of the edge-aware attention block (EAAB) for field-boundary modeling. Adding only DualA yielded an OA of 0.9185, an IoU of 0.8694, and a precision of 0.9445, the highest among single-module configurations, indicating improved semantic discrimination between paddy rice and background.

Among two-module combinations, AMSA + EAAB achieved the highest IoU and F1-score, 0.8741 and 0.9328, respectively, whereas EAAB + DualA achieved the highest precision of 0.9399. The complete DAEA-Net achieved the highest OA, IoU, F1-score, and Kappa coefficient at 0.9205, 0.8747, 0.9331, and 0.8352, respectively. These results indicate that the three modules contribute incrementally to the overall performance: the complete model achieved the highest OA, IoU, F1-score and Kappa, whereas individual modules showed distinct advantages in recall and precision. Because the differences among the configurations are small (0.001–0.003), we interpret these results as evidence of modest, consistent contributions rather than of large independent gains.

McNemar's test was conducted on the pooled predictions over the 142 test tiles to verify the statistical significance of the performance differences between each module configuration and the DeepLabV3+ baseline. Here, b is the number of pixels that the baseline misclassified but the configuration classified correctly, and c is the number of pixels that the baseline classified correctly but the configuration misclassified; $b - c > 0$ therefore indicates a net improvement over the baseline. The p-value was obtained from the McNemar χ^2 statistic with continuity correction, $\chi^2 = (|b - c| - 1)^2/(b + c)$, with one degree of freedom. The resulting discordant-pair counts and p-values are reported in Table 5, and all comparisons yielded $p < 10^{-6}$, indicating that the differences are statistically significant rather than arising from random variation. Configurations incorporating EAAB showed b values well above their c values (EAAB alone: b = 386,065, c = 373,710; AMSA + EAAB: b = 393,015, c = 377,893), and therefore corrected far more of the pixels misclassified by the baseline than they broke, whereas the AMSA-only and AMSA + DualA configurations were significantly worse than the baseline ($b < c$), consistent with their lower OA and IoU in the ablation table.

To investigate the contribution of the multi-source features, we conducted an ablation study on the input channels. Two input schemes were compared: RGB-only and the eight-band combination of RGB, NDVI, and DEM. All network configurations and training hyperparameters were kept identical, and only the input channels were modified. The quantitative results are listed in Table 6 and the visual comparison is illustrated in Figure 8. Visually, the RGB-only model produces more false-positive and false-negative pixels, especially for fragmented farmland parcels. Relative to the RGB-only scheme, the eight-band input achieved modest but consistent improvements in OA (from 0.9167 to 0.9205) and IoU (from 0.8705 to 0.8747), mainly through higher precision (from 0.9158 to 0.9297) and a reduction in false-positive pixels in fragmented parcels, at the cost of a small loss in recall (from 0.9463 to 0.9366) (Table 6).

Table 6. Ablation results of different inputs of DAEA-Net on the test set.

Input Scheme	OA	Precision	Recall	IoU	F1-Score	Kappa
RGB-only	0.9167	0.9158	0.9463	0.8705	0.9308	0.8262
DAEA-Net	0.9205	0.9297	0.9366	0.8747	0.9331	0.8352

Note: All metrics were computed from a single global confusion matrix obtained by summing the per-tile confusion matrices over all 142 (256 × 256) test tiles. Precision, recall, F1-score, and IoU refer to the rice (foreground) class, whereas OA and the Kappa coefficient are overall metrics over both classes.

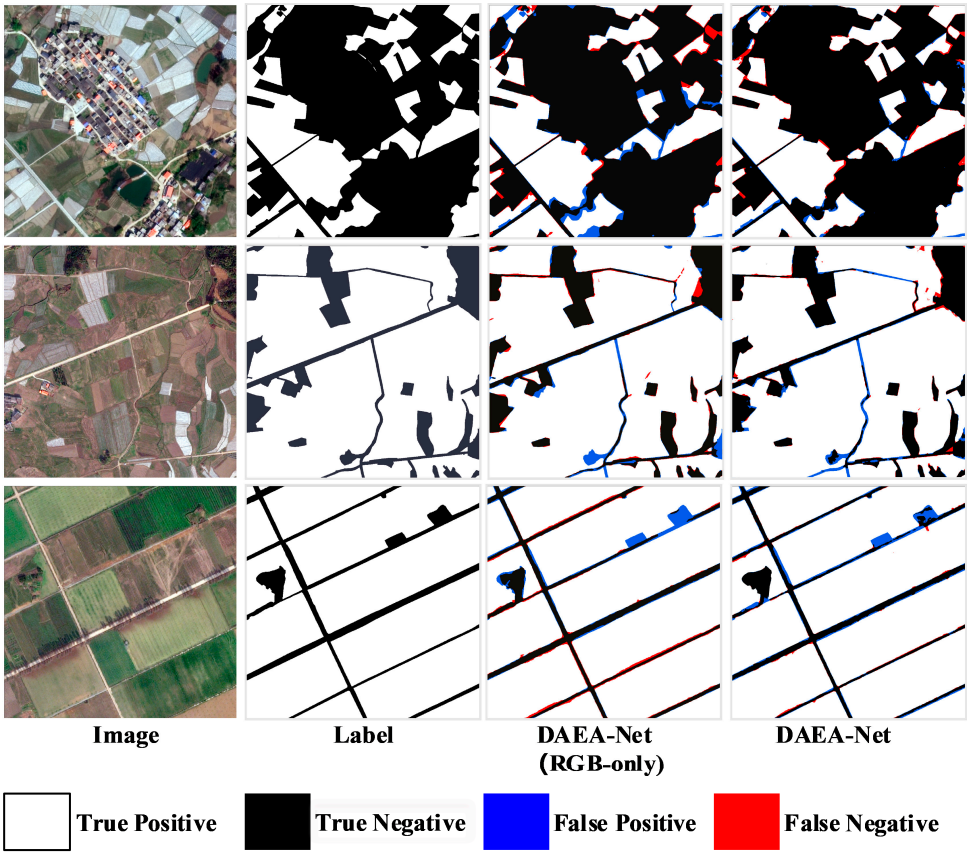


Figure 8. Comparison of paddy rice extraction results and pixel-wise classification outcomes for the RGB-only and eight-band input schemes on representative sample tiles.

4.2. Spatial Pattern of Paddy Rice Areas

The overall paddy rice extraction results are shown in Figure 9. Spatially, the rice-growing patterns closely matched the geomorphic characteristics of the study areas. Paddy rice fields in Taishan were concentrated and contiguous on the coastal plain, interspersed with coastal aquaculture ponds, whereas paddy rice in Yingde formed belts along the alluvial plains of the Beijiang River and its tributaries, with high planting density inside the basin. In contrast, paddy rice in Pingyuan and Lechang was strongly constrained by mountainous terrain and showed a fragmented pattern, occurring mainly on valley and hillside terraces.

Controlled jointly by terrain and river-system distribution, paddy rice in the study areas generally preferred gentle slopes. The proportions of paddy rice located on slopes of less than 5° were 95.10%, 97.61%, 82.26%, and 82.01% in Yingde, Taishan, Lechang, and Pingyuan, respectively, and the area-weighted mean slopes were 1.51°, 0.89°, 2.98°, and 3.17°. Although Lechang and Pingyuan have greater topographic relief, paddy rice there was still mainly distributed on gentle slopes and in relatively flat areas such as river valleys and intermontane basins. In terms of elevation, paddy rice in Yingde and Taishan was concentrated below 100 m, whereas paddy rice in Lechang and Pingyuan was mainly distributed between 100 and 300 m. These patterns indicate a preference for flat areas with irrigation potential, whereas the elevation range suited to rice rises as terrain relief increases. The extracted rice areas consistently followed river systems.

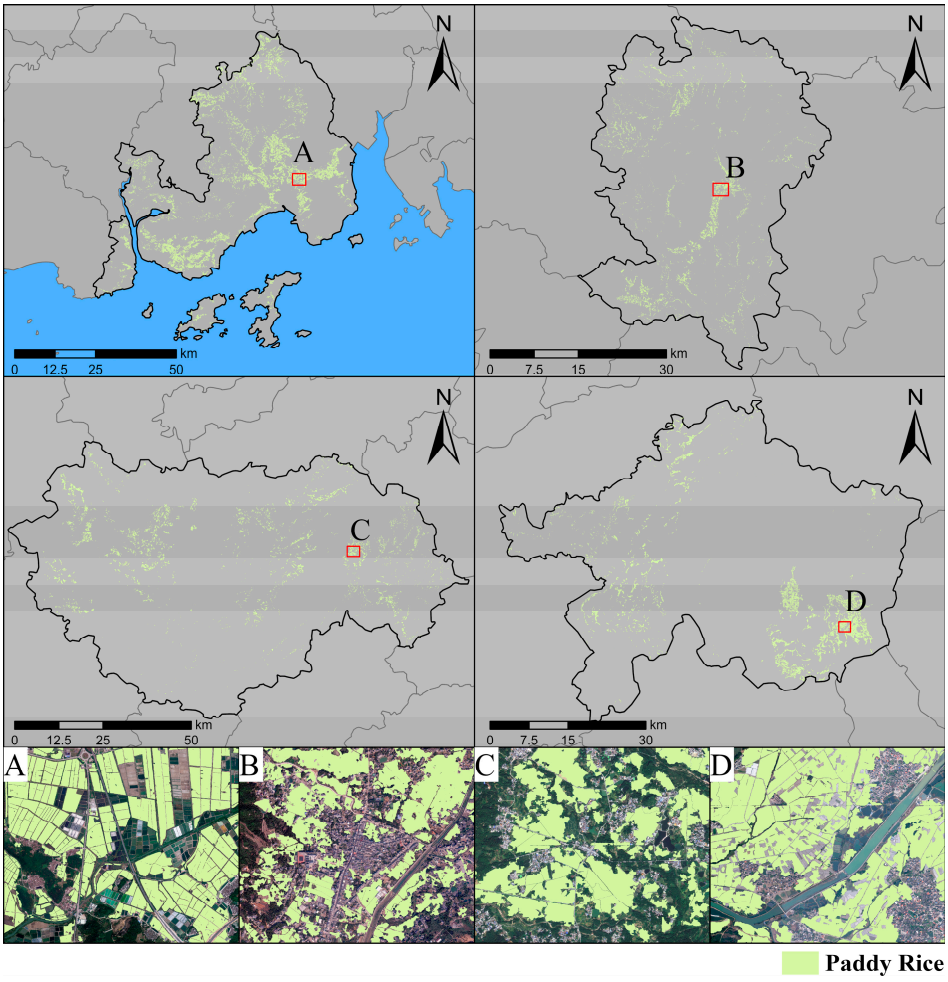


Figure 9. Spatial distribution and enlarged local details of the extracted rice-growing areas in the four study regions. (**A**) Taishan; (**B**) Pingyuan; (**C**) Yingde; (**D**) Lechang. The four maps in the upper

part show the extracted paddy rice (pale yellow) in each study region; black lines indicate the county boundaries, and grey indicates land surfaces that were not classified as paddy rice. The red squares and the corresponding letters mark the areas enlarged in the panels below, where the extracted paddy rice is overlaid on the very-high-resolution optical imagery.

4.3. Accuracy Assessment and Comparison of Flood-Water Extraction

The NDRCC dataset was used as reference validation data for flood-water extraction during the June 2022 Yingde event, and the water class of the AIR-PolSAR-Seg-2.0 dataset was used to assess the accuracy of the water-extraction algorithm itself on a general water benchmark acquired in November 2016. The NDRCC reference is a human interpretation of the same GF-3 scene of 23 June 2022; as a pixel-level reference for that acquisition, it is derived from the same scene, and it is used only for validation and never as an input to the extraction algorithm, which operates solely on the SAR backscatter. The reference and the extraction are therefore independent processes applied to the same image. Two reference methods were selected for comparison: the optimal single-threshold method and a modified Otsu method (M-Otsu) adapted from the equal-sampling strategy proposed by Cordeiro et al. [60] for optical water indices. In the present study, equal numbers of low- and high-backscatter pixels were sampled from the SAR image before applying Otsu thresholding, thereby reducing the influence of class imbalance when flood-water pixels were rare. Similar bimodal-histogram sampling strategies have been used for Sentinel-1 SAR water mapping [61]. The optimal single threshold was obtained by scanning candidate thresholds and retaining the value that gave the best accuracy on the reference data; because the single-threshold method has only this one free parameter, its optimal value represents the maximum performance attainable by that method and therefore serves as the most favorable baseline for comparison. We note that PH-MCW also uses the prior water proportion, which is refined by visual inspection of each scene (Section 3.2.3), so neither the single-threshold baseline nor PH-MCW is fully unsupervised; the comparison is between two methods that each use a reference-informed parameter. Five accuracy metrics—OA, precision, recall, F1-score, and the Kappa coefficient—were used to compare the three methods before and after AOI filtering, and the results are summarized in Table 7.

On the NDRCC dataset, PH-MCW after AOI filtering achieved an F1-score of 0.8211, which was 0.0168 and 0.0848 higher than the optimal single-threshold method and M-Otsu, respectively. AOI filtering increased the precision of PH-MCW from 0.6815 to 0.8631 but reduced its recall from 0.8723 to 0.7830, because the mask removes water pixels that are present in the reference data as well as false alarms. The increase in precision reflects the removal of terrain-induced false alarms, that is, low-backscatter pixels on slopes steeper than 25° that belong neither to the permanent water bodies nor to the land types that may be inundated; permanent water bodies and potentially inundated land are retained by the AOI and are therefore not removed. The decrease in recall reflects the removal of correctly detected flood water that lies on slopes steeper than 25° and in land types not covered by the water-body or flood-prone classes. The F1-score nevertheless increased from 0.7652 to 0.8211, indicating that the reduction in false alarms outweighed the loss of correctly detected water.

On the AIR-PolSAR-Seg-2.0 dataset, PH-MCW achieved an F1-score of 0.7640, which was 0.0240 and 0.0349 higher than the optimal single-threshold method and M-Otsu, respectively. AOI filtering produced clear accuracy gains for all three methods. For the optimal single-threshold method, the F1-score increased from 0.4283 to 0.7400, an absolute increase of 0.3117. For PH-MCW, precision increased from 0.3981 to 0.7712 while recall changed only slightly, from 0.7582 to 0.7569. This result shows that the AOI removes terrain-induced low-backscatter false alarms in complex backgrounds while keeping water surfaces, and thereby improves the reliability of the extracted water. Because the AOI

contains the permanent-water layer of the land-type survey data, that is, water information of the same kind as the class evaluated on this benchmark, the filtered accuracy on this stable-water benchmark partly reflects information shared with the water labels; the filtered F1-scores should therefore be interpreted as an upper bound on the benefit of AOI filtering, and they mainly reflect the removal of terrain-induced false alarms rather than an ability to separate permanent from temporary water. The NDRCC reference, which is not used to construct the AOI, provides the unbiased assessment.

Table 7. Accuracy of the three water-extraction methods before and after AOI filtering on the two validation datasets.

Dataset	Method	Unfiltered							Filtered						
		OA	Pre	Re	F1	Kappa	$BIoU_2$	$BIoU_5$	OA	Pre	Re	F1	Kappa	$BIoU_2$	$BIoU_5$
NDRCC	Optimal single threshold	0.9725	0.6360	0.8506	0.7278	0.7137	0.2234	0.3888	0.9838	0.8420	0.7699	0.8043	0.7959	0.2678	0.4562
	M-Otsu	0.8431	0.2099	0.9516	0.3440	0.2940	0.0690	0.1326	0.9742	0.6600	0.8325	0.7363	0.7229	0.2394	0.4028
	PH-MCW	0.9769	0.6815	0.8723	0.7652	0.7532	0.2649	0.4467	0.9853	0.8631	0.7830	0.8211	0.8134	0.2909	0.4919
AIR-PolSAR-Seg-2.0	Optimal single threshold	0.7780	0.2971	0.7676	0.4283	0.3225	0.0697	0.1612	0.9417	0.7156	0.7661	0.7400	0.7072	0.2028	0.3823
	M-Otsu	0.8535	0.3978	0.6857	0.5035	0.4246	0.0924	0.2055	0.9449	0.7798	0.6847	0.7291	0.6986	0.2012	0.3812
	PH-MCW	0.8496	0.3981	0.7582	0.5221	0.4430	0.0914	0.2088	0.9493	0.7712	0.7569	0.7640	0.7356	0.2187	0.4062

Note: OA, overall accuracy; Pre, precision; Re, recall; F1, F1-score; $BIoU_2$ and $BIoU_5$, Boundary IoU with a boundary-band width of 2 and 5 pixels, respectively.

Pixel-wise classification results for the three water-extraction methods are compared in Figure 10. Compared with the two reference methods, PH-MCW produced more complete water boundaries and fewer false-alarm pixels. The optimal single-threshold method approached PH-MCW in OA but systematically missed pixels along water–land transition zones. M-Otsu partly mitigated the histogram shift caused by class imbalance but still showed substantial omissions in low-contrast areas. The overall flood-water mapping results of the three methods are shown in Figure 11.

In both validation datasets, correctly identified water pixels were concentrated in connected water bodies such as rivers and reservoirs. The remaining false alarms and omissions were mainly distributed along water–land boundaries and small tributaries. These areas are more susceptible to SAR speckle noise and mixed-pixel effects, and boundary-labeling errors in the validation data may also cause local spatial differences. Overall, PH-MCW achieved the highest F1-score on both reference datasets and showed stable performance in boundary completeness and false-alarm control.

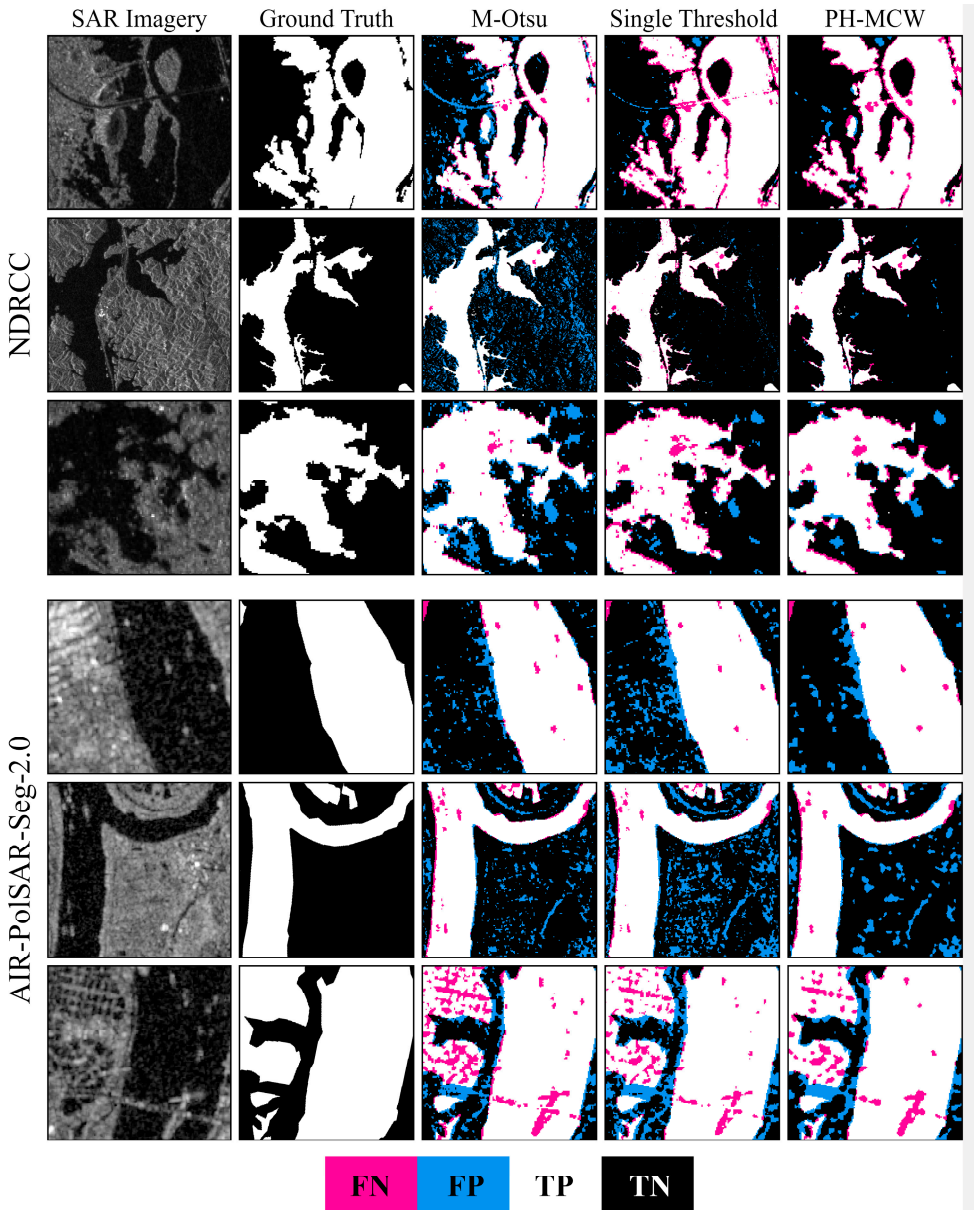


Figure 10. Comparison of water-extraction results and pixel-wise confusion categories for three methods in representative areas of the validation datasets. Both validation datasets are based on GF-3 imagery.

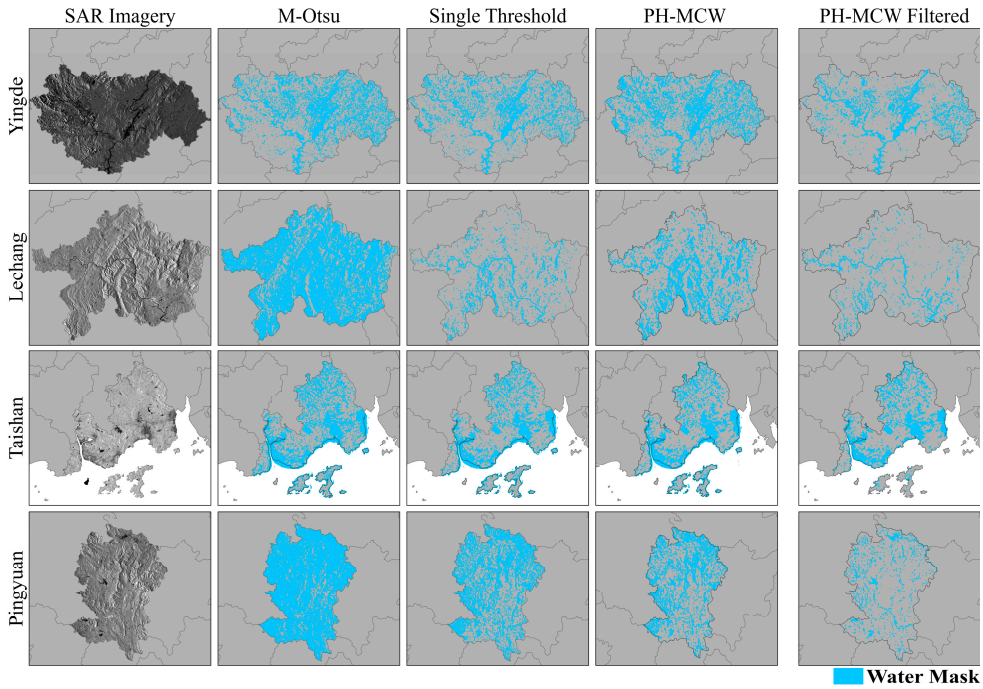


Figure 11. Comparison of water-body extraction results. The first column shows SAR imagery acquired over Yingde (GF-3, 23 June 2022), Lechang (Sentinel-1, 15 June 2024), Taishan (Sentinel-1, 25 September 2023), and Pingyuan (Sentinel-1, 17 June 2024). The second to fourth columns show the unfiltered results obtained using M-Otsu, single threshold, and PH-MCW, respectively. The fifth column shows the PH-MCW result after AOI filtering. Cyan pixels represent extracted water masks.

4.4. Spatiotemporal Pattern of Paddy Rice Inundation Under Intense Rainfall

The spatiotemporal pattern of paddy rice inundation was analyzed based on four representative intense-rainfall events between 2022 and 2024, including Yingde in June 2022, Taishan in September 2023, and Lechang and Pingyuan in June 2024. The four study areas differed in topographic relief, rainfall intensity, and rice-planting pattern. The inundation distribution was jointly controlled by terrain and river systems, as shown in Figure 12. Among these events, the inundated paddy rice area was the largest in Yingde, reaching 13.45 km^2; Taishan ranked second with 12.74 km^2; and Lechang and Pingyuan reached 3.87 km^2 and 2.93 km^2, respectively. In all four events, the inundated paddy rice area was calculated as the increase in water extent from the pre-flood to the during-flood image within the paddy rice map, so that irrigation water already present before each event was excluded.

Spatially, paddy rice inundation was strongly coupled with terrain conditions and rainfall centers. In Yingde, inundated paddy rice was concentrated in the flat central river basin and occurred in contiguous areas along both banks of the Beijiang main stem and its tributaries, with a smaller amount in upstream valleys to the northwest. In Pingyuan, inundated paddy rice was more scattered across intermontane valleys and river plains. In Taishan, inundated paddy rice was concentrated in fields adjacent to coastal aquaculture ponds. In Lechang, inundated paddy rice was concentrated in the flat southeastern river-valley plain. Across all four areas, inundated paddy rice tended to accumulate in low-lying terrain and extend along the river network.

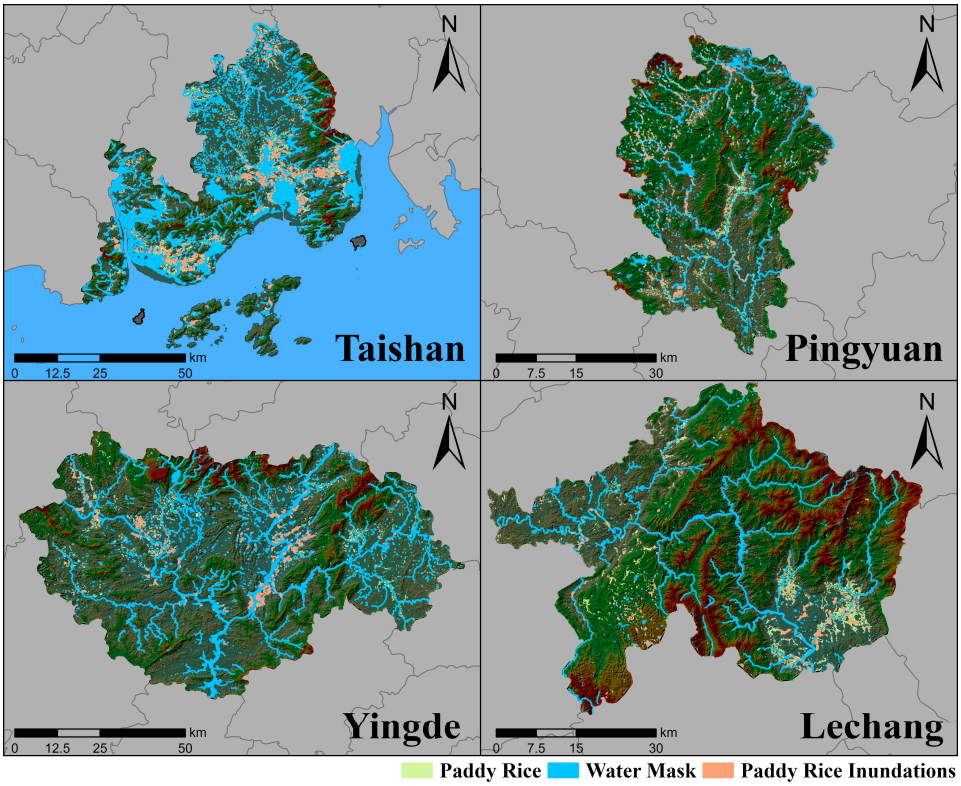


Figure 12. Spatial distribution of paddy rice inundation under intense rainfall conditions in Taishan, Pingyuan, Yingde, and Lechang.

The June 2022 intense-rainfall event in Yingde was further analyzed to characterize a single inundation process. Persistent heavy rainfall in mid-to-late June caused the Beijiang water level to rise rapidly. Figure 13 compares pre-flood, during-flood, and post-flood SAR imagery with the extracted paddy rice inundation extent. Water expanded markedly during the event, increasing from 165.34 km² before the flood to 199.20 km² during the flood, a net increase of 33.86 km² (approximately 20.5%); the inundated paddy rice area reached 13.45 km². This value was derived from the GF-3 scene of 23 June 2022, the acquisition closest to the flood peak, as the increment between the during-flood and pre-flood-water extents within the paddy rice map, so that the irrigation water already present in the paddy fields before the event is excluded; it is the mapped extent at the acquisition closest to the peak, not the maximum inundation extent. The inundation pattern closely matched the spatial distribution of rainfall. According to township-level statistics aggregated from the CHM_PRE V2 gridded precipitation dataset, the maximum 3-day cumulative rainfall across Yingde's 24 townships from 1 to 30 June 2022 ranged from 137.4 to 443.8 mm, with a mean of 322.1 mm across Yingde; the most intense rainfall occurred on June 19–21. Rainfall decreased markedly from north to south, and high values were concentrated in a belt along the middle Beijiang River. Five of Yingde's 24 township-level administrative units recorded maximum 3-day cumulative rainfall exceeding 430 mm, with Wangbu Township recording the highest value of 443.8 mm. The inundated paddy rice distribution closely matched the rainfall pattern, and the middle Beijiang valley plain was also Yingde's main rice-growing area, but its capacity to convey upstream flood peaks is limited. Most inundated paddy rice occurred on this riverside plain; Wangbu and Dazhan townships together accounted for approximately 6.6 km², about half of Yingde's inundated paddy rice area. Wangbu, upstream of the urban area and adjacent to the Beijiang main stem, was the most severely affected, followed by Dazhan along the

downstream riverbank. In contrast, Lizhen and Dadong in the south were dominated by low mountains and hills, with higher terrain and less rainfall, and experienced almost no paddy rice inundation. These results indicate that flood peaks generated by intense rainfall converging along the middle Beijiang River, combined with the low-lying valley-plain terrain, were the dominant drivers of extensive paddy rice inundation in north-central Yingde.

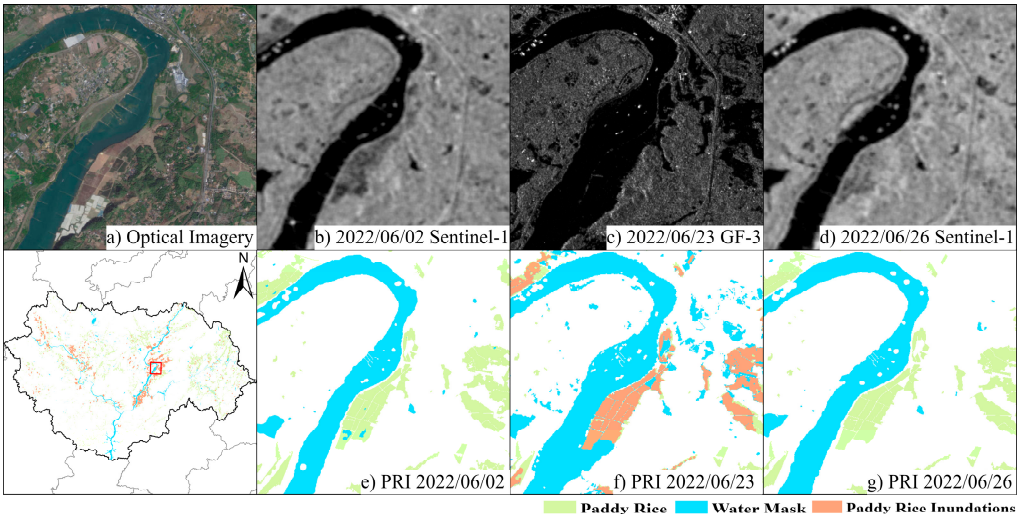


Figure 13. Example of paddy rice inundation mapping in Yingde, Guangdong Province, in June 2022, based on VHR optical imagery and multi-temporal SAR imagery, with the during-flood extent derived from the GF-3 scene of 23 June 2022.

5. Discussion

5.1. Effects of the Potential Flood-Prone Area and the Paddy Rice Mask on Inundated Paddy Rice Extraction

Based on the accuracy assessment and comparison of the water-extraction results, AOI filtering substantially reduced terrain-induced false alarms in flood-water identification, while permanent water bodies and potentially inundated land were retained. In mountainous and hilly scenes, most false alarms in flood-water extraction arose from mountain shadows and radar layover [62]. Of the candidate water pixels removed by the combined AOI mask, pixels with slopes greater than 25° accounted for 77.2–88.4%. In river-valley scenes with steep banks, the prior permanent water bodies prevented the terrain mask from over-removing the river water surfaces. In the coastal county-level city of Taishan, the sea surface also exhibits low backscatter and interferes with the percentile-based marking of water seeds in PH-MCW; the coastline contained in the water-body data was therefore used to clip the sea surface to a fixed extent before seed marking, so that the share of low-backscatter pixels contributed by the sea is controlled.

The paddy rice mask used in this study was derived from the 2023 imagery basemap, so it is one year apart from some of the flood events: it is contemporaneous with the September 2023 Taishan event and one year apart from the June 2022 Yingde event and the June 2024 Lechang and Pingyuan events. Over these intervals, the rice area changed by +0.52% in Yingde (2022–2023) [63,64] and by +0.49% in Lechang and +0.19% in Pingyuan (2023–2024) [65–73], while the provincial rice sown area varied by at most 0.45% in a single year [74], so the mask remained representative. Because the rice mask is used only as a spatial constraint that suppresses false alarms from non-target low-backscattering features—it neither generates water pixels nor enlarges the mapped inundation extent—its effect is conservative rather than inflationary. Nevertheless, aggregate statistics cannot

distinguish field-scale conversions between paddy rice and other uses, so residual uncertainty remains for fields whose use changed around an event; applying the same model to imagery of the event year could quantify this effect directly in future work. Overall, the timeliness and spatial accuracy of the prior layers directly affect the mapped inundation extent.

Previous work has also identified flood-affected rice by combining a crop map with SAR-derived flood maps. Singha et al. [29] integrated monthly 10 m Sentinel-1 flood maps with a 2017 Sentinel-1 paddy rice map to quantify flood-affected rice areas across Bangladesh during 2014–2018. This approach was designed to characterize long-term flood dynamics and affected rice areas at the national scale. The present study addresses a different spatial and operational setting: a 0.5 m paddy rice map is used to preserve field boundaries in fragmented landscapes, while event-specific flood water is extracted from multi-source SAR imagery and filtered using terrain and permanent-water constraints. Consequently, the two approaches are complementary: the former supports repeated national-scale assessment, whereas the present framework emphasizes field-level inundation delineation and false-alarm suppression in complex terrain. Because the study areas, event periods, input data, and validation datasets differ, a direct numerical comparison of accuracy is not appropriate.

5.2. Comparison with Other Rice Products

To further evaluate the extracted paddy rice map in scenes with complex land cover, its spatial pattern was compared with two widely used public products, EARice10 and CCD-Rice 2023, focusing on spatial consistency and boundary representation (Figure 14). In the main rice-distribution areas, the three products showed stable spatial correspondence, with the mapped fields aligned along the same alluvial plains, coastal lowlands, and intermontane valleys, which indicates that the dominant rice-growing regions are robustly captured by all three products.

The differences among the products were concentrated mainly at field boundaries and in scenes where paddy fields are spatially adjacent to other low-vegetation or water-covered surfaces, and they arise from two sources. First, the products reflect different rice phenology and acquisition timing, so fields at different growth stages are represented differently. Second, perennial water bodies, aquaculture ponds, and marginal vegetation are closely adjacent to paddy fields in the study areas; in coastal plains, pond surfaces, pond ridges, and paddy-field boundaries are interleaved, and the 10 m resolution products are susceptible to mixed pixels and boundary generalization. By contrast, the 0.5 m imagery used in this study preserves detailed structures such as field shapes, ridges, and ditches, reducing the possibility that adjacent water is misclassified as paddy rice. EARice10 and CCD-Rice 2023 are therefore more suitable for regional annual mapping, whereas the results of this study have advantages in fine field-boundary representation and in the subsequent construction of the paddy rice map, where precise boundaries reduce false alarms from neighboring water surfaces. Representative comparisons across different terrain and land-use scenes are shown in Figure 14.

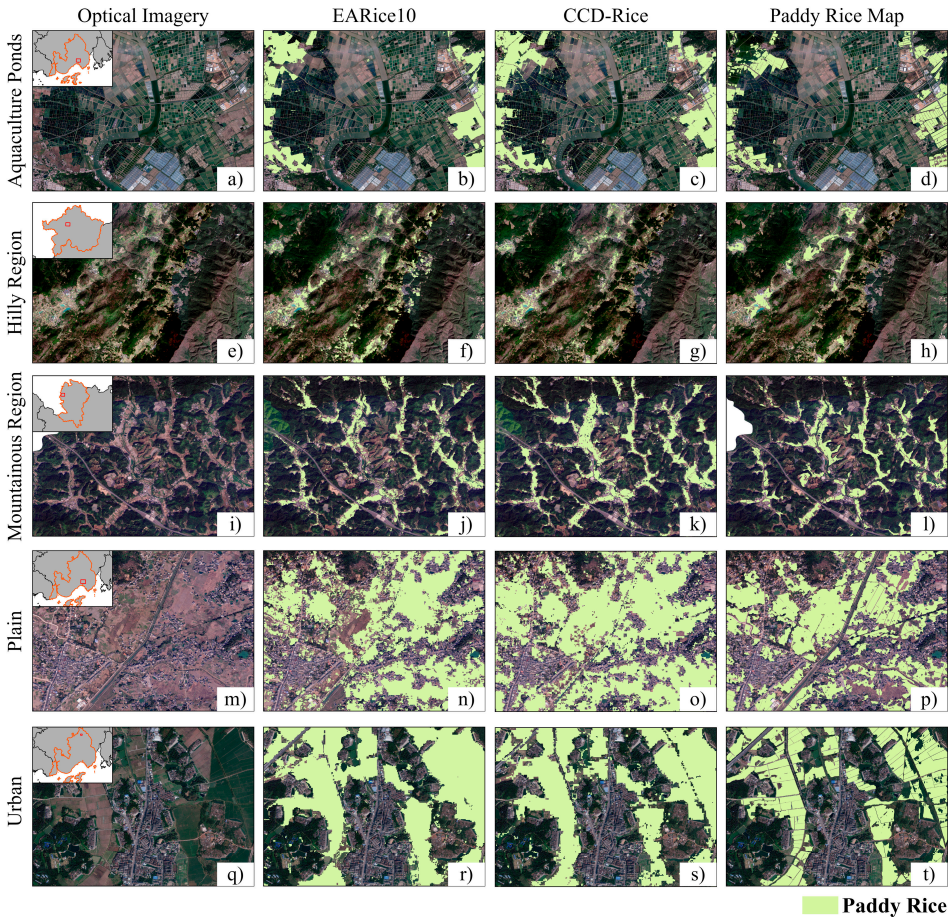


Figure 14. Comparison of rice-mapping results from this study and public rice products across representative terrain and land-use scenes. Each row corresponds to a representative scene of a different terrain or land-use type, as labelled on the left. (**a**,**e**,**i**,**m**,**q**) VHR optical imagery, with a small inset map showing the location of each scene; (**b**,**f**,**j**,**n**,**r**) EARice10 product; (**c**,**g**,**k**,**o**,**s**) CCD-Rice product; (**d**,**h**,**l**,**p**,**t**) paddy rice map extracted in this study. Pale green denotes paddy rice.

5.3. Cross-Sensor Applicability and Prior-Parameter Sensitivity of PH-MCW

To verify the applicability of PH-MCW to imagery from different sensors, we used the three scenes from different sensors in the Guangdong Meizhou flood-disaster remote-sensing image dataset, namely GF-3, Sentinel-1 and Fucheng-1, and evaluated the extraction accuracy against reference water annotations digitized by the authors through visual interpretation from each image and cross-checked by two other authors. The three Meizhou scenes were acquired on 17 June 2024 (GF-3), 19 June 2024 (Fucheng-1), and 17 June 2024 (Sentinel-1). In addition, a sensitivity experiment was conducted over the prior parameter of PH-MCW to examine how the choice of the prior percentile affects the extraction accuracy. The accuracy metrics at different prior percentiles are reported in Table 8.

Table 8. Accuracy of PH-MCW at different prior percentiles for the three sensors in the cross-sensor validation domain.

Prior Percentile (%)	GF-3			Sentinel-1			Fucheng-1		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score
0.1	0.957	0.701	0.809	0.892	0.580	0.703	0.903	0.402	0.556
0.3	0.899	0.815	0.855	0.811	0.852	**0.831**	0.817	0.672	0.737
0.5	0.869	0.855	**0.862**	0.747	0.912	0.821	0.731	0.777	**0.753**
1.0	0.808	0.899	0.851	0.640	0.970	0.772	0.574	0.870	0.692
2.0	0.728	0.932	0.817	0.524	0.992	0.686	0.467	0.922	0.620
5.0	0.584	0.966	0.728	0.398	0.999	0.569	0.350	0.961	0.513
15.0	0.351	0.991	0.518	0.267	1.000	0.421	0.270	0.989	0.424

Note: Bold indicates the maximum F1-score of each sensor in this table, i.e., the value obtained at the optimal prior percentile.

The sensitivity curves show a single-peaked shape: the F1-score first increases and then decreases as the prior percentile grows, and the peak is located in a narrow low-percentile interval for all three sensors (0.3–0.6%, with the peak F1-scores of 0.862, 0.831 and 0.753 for GF-3, Sentinel-1 and Fucheng-1, respectively). Because the optimum percentile is identified on the same reference annotations used for evaluation, these peak F1-scores are optimistic; they serve to compare the three sensors and to locate the optimum interval rather than as absolute accuracy estimates. The three sensors have nearly the same peak position and reach a reasonable accuracy at the peak, which indicates that seed marking based on water percentiles and histogram morphology has good adaptability to the radiometric differences among sensors. In addition, near the peak, the accuracy decreases only slightly when the prior percentile changes, whereas it decreases markedly once the percentile departs sufficiently from the optimum: when the prior percentile moves from 1% to 5%, the F1-score drops by 0.123, 0.203 and 0.179 for the three sensors, respectively. This indicates that the influence of the prior parameter on the extraction accuracy is moderate within the peak interval and grows strongly outside it. The corresponding water-extraction results of the three sensors are compared in local enlarged views in Figure 15.

In multi-source remote-sensing analysis, defining the threshold by percentiles rather than by fixed values adapts to the radiometric differences among sensors; in particular, some GF-3 L1A products are distributed without calibration parameters, and their DN-value images can be used directly by PH-MCW.

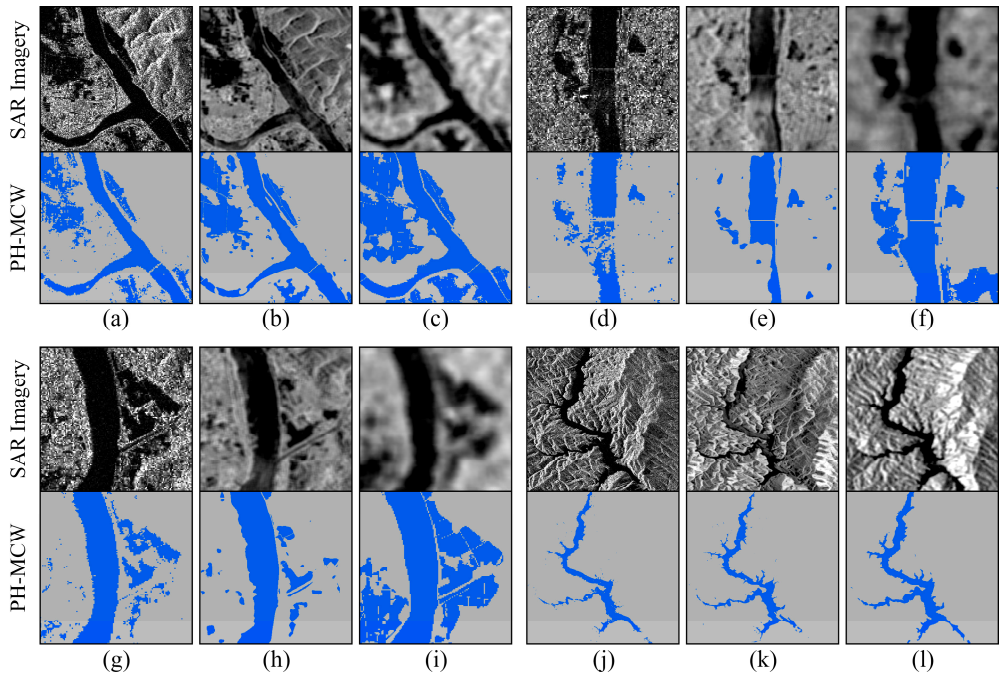


Figure 15. Comparison of water extraction results among the three SAR sensors in the cross-sensor validation domain. (**a**,**d**,**g**,**j**) GF-3; (**b**,**e**,**h**,**k**) Fucheng-1; (**c**,**f**,**i**,**l**) Sentinel-1.

In summary, PH-MCW achieves reasonable pixel-level accuracy on all three sensors within a consistent single-peaked optimum interval, and its accuracy degrades only slightly near the optimum. The prior percentile should therefore be regarded as a transferable initial constraint rather than a fixed universal constant. Because the peak percentile differed slightly among the three sensors, and because the radiometric characteristics of the sensors differ, the specific prior percentile for a new sensor or a new region cannot be determined a priori; it should be taken from the land-survey water proportion of the target region or, where such data are not available, located with a local sensitivity test—0.3–0.6% for the cross-sensor domain examined here—and confirmed locally before large-scale application. The prior percentile in Table 8 and the prior water proportion used in the main experiments (Section 3.2.3) are the same quantity—the fraction of low-backscatter pixels identifiable as water within the image extent—but they refer to different extents: the main experiments use the land-survey water proportion of each administrative region, whereas Table 8 sweeps the prior on the smaller region where the three sensor images overlap, where the water proportion is inherently lower, so its optimum is specific to that validation domain.

5.4. General Applicability, Practical Implications, and Limitations

The proposed two-stage framework is general in its construction rather than being tied to the specific study areas. The DAEA-Net segmentation model and the PH-MCW water-extraction method both rely on generic mechanisms—multi-scale attention and edge-aware features for rice mapping, and prior-percentile and histogram-morphology seed marking for water extraction—so the framework can, in principle, be transferred to other rice-growing regions and to other sensors, provided that field samples and the prior water proportion are re-established locally. The cross-sensor experiment in Section 5.3 has already shown that PH-MCW maintains reasonable accuracy across GF-3, Sentinel-1 and Fucheng-1.

Beyond its general applicability, the framework has practical value in that it provides a spatially explicit, field-scale delineation of flood-inundated paddy fields. The present

method locates the specific fields that were inundated at the acquisition closest to the flood peak, together with their terrain and river-system context. This field-scale spatial pattern supports targeted post-disaster operations—such as the deployment of drainage and rescue resources, the verification of agricultural insurance claims, and the planning of replanting and recovery—which require knowing where the affected fields are. Because the inundation distribution is tightly coupled with terrain and river systems, the map also provides a basis for flood-risk zoning and for siting drainage infrastructure in the study region.

The method nevertheless has limitations that should be kept in mind when interpreting the results. First, the mapped inundation is a single-date snapshot: SAR acquisitions are separated by revisit intervals, and the image used for each event is the acquisition closest to the flood peak, which may not coincide with the maximum inundation. Second, flood water in paddy fields recedes rapidly, and shallow or rapidly draining inundation may already have disappeared by the time of imaging, so the mapped extent is a lower bound on the peak inundation. Third, flood water is not always visible on SAR: wind roughening of the water surface, the rice canopy overtopping the water, and speckle can all reduce the contrast that the extraction relies on. For these reasons, the extracted inundation extent is not directly comparable with official disaster statistics of the affected area, which are compiled over the entire event and under administrative definitions. The method is therefore best suited to analyzing the spatial pattern and field-scale distribution of paddy rice inundation and to supporting spatially targeted post-disaster decisions, rather than to reproducing an aggregate area that matches official statistics. Fourth, the 0.5 m basemap is a provincial ortho-mosaic without per-scene acquisition dates, so the exact phenological stage of each field at the time of imaging is unknown; although the quarterly NDVI layers compensate for this to some extent, residual uncertainty in the phenological alignment of the rice map remains. Fifth, among the four study areas, we selected representative areas in Lechang, Pingyuan and Taishan to construct the dataset, and no separate training block was established in Yingde; the rice map of Yingde is therefore obtained by extrapolation. Yingde's paddy fields are concentrated on the Beijiang alluvial plains, a landscape type also represented by the plain and valley fields in the training blocks, which supports the transfer, but field-scale differences may still introduce residual error. Sixth, fields that had not yet been transplanted, and therefore did not appear as paddy rice in the rice mask, are excluded from the mapped inundation area, so late-planted or temporarily fallow fields may be omitted from the inundation estimate. Seventh, the test set was randomly split at the tile level from the same nine blocks as the training set, so nearby tiles from the same field may be spatially correlated. Although the stride equals the tile size (no pixel overlap) and all compared models share the same split, this can slightly inflate the absolute accuracy of models that memorize local patterns and, because the differences among configurations are only 0.001–0.003, could, in principle, also affect their ranking; a spatially independent (block-wise or leave-one-county-out) split would provide a more rigorous assessment in future work.

6. Conclusions

This study tackled the challenge of mapping flood-affected paddy rice areas in complex terrain by constructing a two-stage mapping framework consisting of paddy rice extraction and flood-water identification. First, seasonal NDVI and DEM data were fused with Jilin-1 optical imagery, and DAEA-Net, a semantic-segmentation model integrating dual-attention and an edge-aware attention block (EAAB), was used to extract paddy rice areas. Second, multi-source GF-3 and Sentinel-1 SAR imagery were analyzed with PH-MCW; the AOI was applied to the extracted water to remove terrain-induced false alarms,

and the inundation extent of paddy fields was obtained by differencing the pre-event and during-event water extents within the paddy rice map.

On the test set, DAEA-Net achieved an OA of 0.9205, a Kappa coefficient of 0.8352, and an IoU of 0.8747, thereby capturing the spatial distribution and field boundaries of paddy rice under different geomorphic backgrounds. The AOI (potential flood-prone area), which removes steep-slope false alarms while retaining permanent water bodies and potentially inundated land, suppressed the false alarms caused by mountain shadows and radar layover. PH-MCW achieved an F1-score of 0.821 on the NDRCC dataset and up to 0.764 on the AIR-PolSAR-Seg-2.0 dataset, and maintained pixel-level accuracy across GF-3, Sentinel-1 and Fucheng-1 when each sensor was evaluated against its own reference annotations, with a sensitivity analysis showing that the prior percentile is a transferable initial constraint that should be confirmed locally before large-scale application.

Author Contributions: Conceptualization, Y.J.; data curation, Q.B., M.W. and H.Z.; methodology, Y.J. and X.M.; formal analysis, Q.B. and J.L.; writing—original draft preparation, Y.J. and Q.B.; writing—review and editing, Q.B. and X.M.; supervision, Y.J. and J.L.; visualization, Q.B. and J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the GuangDong Basic and Applied Basic Research Foundation (grant no. 2025A1515011668), the Jiangsu Agricultural Science and Technology Independent Innovation Fund [grant number CX(23)1025], the Development Program of Powerchina (Grant No. DJ-ZDXM-2023-39) and by institutional support from the Guangdong Engineering and Research Center for Unmanned Aerial Vehicle Remote Sensing of Agricultural Water and Soil Information.

Data Availability Statement: The Sentinel-1 SAR and Sentinel-2 MSI imagery used in this study are publicly available through the Copernicus Data Space Ecosystem (https://dataspace.copernicus.eu, accessed on 12 April 2026), and the GF-3 SAR imagery was downloaded from the China Platform of Earth Observation System (https://www.cpeos.org.cn/, accessed on 17 June 2026). The 0.5 m very-high-resolution optical imagery was derived from the 2023 imagery basemap published by the Geographic Information Public Service Platform of Guangdong Province (Tianditu Guangdong, https://guangdong.tianditu.gov.cn/, accessed on 20 May 2026), which has been openly shared on the platform since April 2026 and can be accessed, within the platform's terms of use, through its standard OGC services (WMS, WMTS, WCS, Tile and GeoMapService); users should credit the source as required by the platform's copyright notice. The flood reference data used for validation include the Remote-Sensing Monitoring Dataset of the June 2022 Flood Disaster in Yingde, Guangdong Province, produced by the National Disaster Reduction Center of China (NDRCC) and distributed by the National Earth Observation Data Center (NODA), and the Multi-Satellite Remote-Sensing Imagery Data of the 16 June 2024 Flood Disaster in Meizhou City, Guangdong Province, obtained from the National Earth Observation Data Center (https://www.chinageoss.cn/, accessed on 27 July 2026). The EARice10 and CCD-Rice paddy rice products are available from the Earth System Science Data repository, and the AIR-PolSAR-Seg-2.0 dataset is publicly accessible at https://doi.org/10.12000/JR24237, accessed on 27 July 2026. The 5 m DEM and the land-type survey data are not publicly available. The code for DAEA-Net and PH-MCW, the final paddy rice maps (as polygon shapefiles), and the final flood-inundated paddy rice maps are available from the corresponding author upon reasonable request and from the GitHub repository at https://github.com/bianqinanedu/paddy-rice-inundation-mapping, accessed on 13 October 2026. The annotated training tiles contain non-public data and are not shared; only the coordinates of the sampling blocks (Table A1) are provided.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Geographic coordinates of the nine paddy rice sampling blocks. Each block is 2000 m × 2000 m. Block IDs: LC = Lechang, PY = Pingyuan, TS = Taishan.

No.	Block ID	County	Center Longitude (°E)	Center Latitude (°N)	Bounding Box (°E/°N)
1	LCA	Lechang	113.4633	25.1284	113.4532–113.4733/25.1190–25.1378
2	LCB	Lechang	113.4256	25.1150	113.4155–113.4356/25.1056–25.1243
3	LCC	Lechang	113.4430	25.1472	113.4325–113.4535/25.1372–25.1573
4	PYA	Pingyuan	115.9445	24.6851	115.9341–115.9549/24.6754–24.6948
5	PYB	Pingyuan	115.9397	24.6473	115.9293–115.9501/24.6375–24.6570
6	TSA	Taishan	112.9218	22.0637	112.9119–112.9318/22.0545–22.0730
7	TSB	Taishan	112.9003	22.0538	112.8904–112.9103/22.0445–22.0631
8	TSC	Taishan	112.8555	22.0387	112.8456–112.8655/22.0294–22.0479
9	TSD	Taishan	112.7927	21.9848	112.7828–112.8027/21.9755–21.9941

References

1. Muthayya, S.; Sugimoto, J.D.; Montgomery, S.; Maberly, G.F. An overview of global rice production, supply, trade, and consumption. *Ann. N. Y. Acad. Sci.* **2014**, *1324*, 7–14. https://doi.org/10.1111/nyas.12540.
2. Bofana, J.; Zhang, M.; Wu, B.; Zeng, H.; Nabil, M.; Zhang, N.; Elnashar, A.; Tian, F.; da Silva, J.M.; Botão, A.; et al. How long did crops survive from floods caused by Cyclone Idai in Mozambique detected with multi-satellite data. *Remote Sens. Environ.* **2022**, *269*, 112808. https://doi.org/10.1016/j.rse.2021.112808.
3. Jiang, H.; Xu, J.; Zhang, X.; Zhou, X.; Liu, Y.; Ku, M.; Jia, K.; Dai, X.; Sun, Y.; Chen, S. Cropland inundation mapping in a mountain dominated region based on multi-resolution remotely sensed imagery and active learning for semantic segmentation. *J. Hydrol.* **2024**, *639*, 131609. https://doi.org/10.1016/j.jhydrol.2024.131609.
4. Weiss, M.; Jacob, F.; Duveiller, G. Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* **2020**, *236*, 111402. https://doi.org/10.1016/j.rse.2019.111402.
5. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R.; et al. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172. https://doi.org/10.1016/j.rse.2014.02.001.
6. Chen, R.; Chen, X.; Ren, Y. Automatic multi-temporal land cover mapping with medium spatial resolution using the model migration method. *Remote Sens.* **2025**, *17*, 37. https://doi.org/10.3390/rs17010037.
7. Mulualem, G.M.; Unnisa, Z.; Paramanik, S.; Dash, J. A decade of remote sensing for vegetation monitoring with Sentinel-2. *Remote Sens.* **2026**, *18*, 2448. https://doi.org/10.3390/rs18152448.
8. Pluto-Kossakowska, J. Review on multitemporal classification methods of satellite images for crop and arable land recognition. *Agriculture* **2021**, *11*, 999. https://doi.org/10.3390/agriculture11100999.
9. Fernández-Urrutia, M.; Arbelo, M.; Gil, A. Identification of paddy croplands and its stages using remote sensors: A systematic review. *Sensors* **2023**, *23*, 6932. https://doi.org/10.3390/s23156932.
10. Liu, S.; Chen, Y.; Ma, Y.; Kong, X.; Zhang, X.; Zhang, D. Mapping ratoon rice planting area in central China using Sentinel-2 time stacks and the phenology-based algorithm. *Remote Sens.* **2020**, *12*, 3400. https://doi.org/10.3390/rs12203400.
11. Zhang, D.; Pan, Y.; Zhang, J.; Hu, T.; Zhao, J.; Li, N.; Chen, Q. A generalized approach based on convolutional neural networks for large area cropland mapping at very high resolution. *Remote Sens. Environ.* **2020**, *247*, 111912. https://doi.org/10.1016/j.rse.2020.111912.
12. Yang, L.; Wang, L.; Abubakar, G.A.; Huang, J. High-resolution rice mapping based on SNIC segmentation and multi-source remote sensing images. *Remote Sens.* **2021**, *13*, 1148. https://doi.org/10.3390/rs13061148.
13. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004.
14. Wang, S.; Huang, C.; Huang, L.; Xu, X.; Shi, H.; Gu, Q.; Xue, Q.; Liu, S.; Zhang, C. Monitoring paddy rice cultivation adjustments in Northeast China through time series remote sensing and deep learning. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *142*, 104739. https://doi.org/10.1016/j.jag.2025.104739.

15. Liu, Z.; Liu, J.; Su, Y.; Xiao, X.; Dong, J.; Liu, L. A general model for large-scale paddy rice mapping by combining biological characteristics, deep learning, and multisource remote sensing data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 14705–14717. https://doi.org/10.1109/JSTARS.2025.3573750.
16. Xia, L.; Zhao, F.; Chen, J.; Yu, L.; Lu, M.; Yu, Q.; Liang, S.; Fan, L.; Sun, X.; Wu, S.; et al. A full resolution deep learning network for paddy rice mapping using Landsat data. *ISPRS J. Photogramm. Remote Sens.* **2022**, *194*, 91–107. https://doi.org/10.1016/j.isprsjprs.2022.10.005.
17. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. https://doi.org/10.1016/j.rse.2018.11.032.
18. Zhu, A.-X.; Zhao, F.-H.; Pan, H.-B.; Liu, J.-Z. Mapping rice paddy distribution using remote sensing by coupling deep learning with phenological characteristics. *Remote Sens.* **2021**, *13*, 1360. https://doi.org/10.3390/rs13071360.
19. Thorp, K.R.; Drajat, D. Deep machine learning with Sentinel satellite data to map paddy rice production stages across West Java, Indonesia. *Remote Sens. Environ.* **2021**, *265*, 112679. https://doi.org/10.1016/j.rse.2021.112679.
20. Li, H.; Pan, B.; Xu, R.; Liu, Z.; Pan, Y.; Zhang, X. Phenology-guided deep learning for automated rice mapping using SAR time-series imagery. *Int. J. Digit. Earth* **2026**, *19*, 2620881. https://doi.org/10.1080/17538947.2026.2620881.
21. Zhang, X.; Wei, H.; Shao, Y.; Luan, H.; Wang, D.-H. Transformer-based dual-branch spatial–temporal–spectral feature fusion network for paddy rice mapping. *Remote Sens.* **2025**, *17*, 1999. https://doi.org/10.3390/rs17121999.
22. Minaee, S.; Boykov, Y.; Porikli, F.; Plaza, A.; Kehtarnavaz, N.; Terzopoulos, D. Image segmentation using deep learning: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 3523–3542. https://doi.org/10.1109/TPAMI.2021.3059968.
23. Zhu, X.X.; Tuia, D.; Mou, L.; Xia, G.-S.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. https://doi.org/10.1109/MGRS.2017.2762307.
24. Yuan, Q.; Shen, H.; Li, T.; Li, Z.; Li, S.; Jiang, Y.; Xu, H.; Tan, W.; Yang, Q.; Wang, J.; et al. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sens. Environ.* **2020**, *241*, 111716. https://doi.org/10.1016/j.rse.2020.111716.
25. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. https://doi.org/10.1016/j.isprsjprs.2019.04.015.
26. DeVries, B.; Huang, C.; Armston, J.; Huang, W.; Jones, J.W.; Lange, M.W. Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. *Remote Sens. Environ.* **2020**, *240*, 111664. https://doi.org/10.1016/j.rse.2020.111664.
27. McCormack, T.; Campanyà, J.; Naughton, O. A methodology for mapping annual flood extent using multi-temporal Sentinel-1 imagery. *Remote Sens. Environ.* **2022**, *282*, 113273. https://doi.org/10.1016/j.rse.2022.113273.
28. Guan, H.; Huang, J.; Li, L.; Li, X.; Miao, S.; Su, W.; Ma, Y.; Niu, Q.; Huang, H. Improved Gaussian mixture model to map the flooded crops of VV and VH polarization data. *Remote Sens. Environ.* **2023**, *295*, 113714. https://doi.org/10.1016/j.rse.2023.113714.
29. Singha, M.; Dong, J.; Sarmah, S.; You, N.; Zhou, Y.; Zhang, G.; Doughty, R.; Xiao, X. Identifying floods and flood-affected paddy rice fields in Bangladesh based on Sentinel-1 imagery and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2020**, *166*, 278–293. https://doi.org/10.1016/j.isprsjprs.2020.06.011.
30. Tong, X.; Luo, X.; Liu, S.; Xie, H.; Chao, W.; Liu, S.; Liu, S.; Makhinov, A.N.; Makhinova, A.F.; Jiang, Y. An approach for flood monitoring by the combined use of Landsat 8 optical imagery and COSMO-SkyMed radar imagery. *ISPRS J. Photogramm. Remote Sens.* **2018**, *136*, 144–153. https://doi.org/10.1016/j.isprsjprs.2017.11.006.
31. Panda, D.; Barik, J. Flooding tolerance in rice: Focus on mechanisms and approaches. *Rice Sci.* **2021**, *28*, 43–57. https://doi.org/10.1016/j.rsci.2020.11.006.
32. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Computer Vision—ECCV 2018*; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2018; Volume 11211, pp. 833–851. https://doi.org/10.1007/978-3-030-01234-2_49.
33. Li, X.; Wang, W.; Hu, X.; Yang, J. Selective Kernel Networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019*; IEEE: Piscataway, NJ, USA, 2019; pp. 510–519. https://doi.org/10.1109/CVPR.2019.00060.
34. Yu, W.; Luo, M.; Zhou, P.; Si, C.; Zhou, Y.; Wang, X.; Feng, J.; Yan, S. MetaFormer is actually what you need for vision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022*; IEEE: Piscataway, NJ, USA, 2022; pp. 10819–10829. https://doi.org/10.1109/CVPR52688.2022.01055.

35. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional block attention module. In *Computer Vision—ECCV 2018*; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2018; Volume 11211, pp. 3–19. https://doi.org/10.1007/978-3-030-01234-2_1.
36. Yingde Municipal People's Government. Natural Environment. Available online: http://www.yingde.gov.cn/ydgk/zrhj/ (accessed on 4 October 2026).
37. Yingde Municipal People's Government. Yingde Statistical Yearbook 2024. Available online: http://www.yingde.gov.cn/zwgk/tjsj/content/post_2149998.html (accessed on 4 October 2026).
38. Local Chronicles Office of the Taishan Municipal People's Government. *Taishan Yearbook 2024*; Zhongzhou Ancient Books Publishing House: Zhengzhou, China, 2024; pp. 1, 134.
39. Editorial Committee of Lechang Yearbook. *Lechang Yearbook 2024*; Zhongzhou Ancient Books Publishing House: Zhengzhou, China, 2024; pp. 42, 232.
40. Lechang Municipal People's Government. Land Use Status Data from the 2024 Annual Land Change Survey of Lechang. Available online: http://www.lechang.gov.cn/lczhiyuanju/content/post_2793617.html (accessed on 4 October 2026).
41. Pingyuan County People's Government. About Pingyuan. Available online: https://www.pingyuan.gov.cn/zjpy/index.html (accessed on 4 October 2026).
42. Torres, R.; Snoeij, P.; Geudtner, D.; Bibby, D.; Davidson, M.; Attema, E.; Potin, P.; Rommen, B.; Floury, N.; Brown, M.; et al. GMES Sentinel-1 mission. *Remote Sens. Environ.* **2012**, *120*, 9–24. https://doi.org/10.1016/j.rse.2011.05.028.
43. ESA. Sentinel-1 Products. SentiWiki. Available online: https://sentiwiki.copernicus.eu/web/s1-products (accessed on 4 October 2026).
44. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026.
45. Hu, J.; Miao, C. *CHM_PRE V2: An Upgraded High-Precision Gridded Precipitation Dataset for the Chinese Mainland Considering Spatial Autocorrelation and Covariates*; National Tibetan Plateau/Third Pole Environment Data Center: Beijing, China, 2023. https://doi.org/10.11888/Atmos.tpdc.300523.
46. Liu, M. National Disaster Reduction Center of China. *Remote Sensing Monitoring Dataset of the June 2022 Flood Disaster in Yingde, Guangdong Province*; National Earth Observation Data Center: Beijing, China, 2026. https://doi.org/10.11878/db.202602.000056.
47. Wang, Z.; Zhao, L.; Wang, Y.; Zeng, X.; Kang, J.; Yang, J.; Sun, X. AIR-PolSAR-Seg-2.0: Polarimetric SAR ground terrain classification dataset for large-scale complex scenes. *J. Radars* **2025**, *14*, 353–365. https://doi.org/10.12000/JR24237.
48. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 60. https://doi.org/10.1186/s40537-019-0197-0.
49. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015*; IEEE: Piscataway, NJ, USA, 2015; pp. 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965.
50. Vincent, L.; Soille, P. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 583–598. https://doi.org/10.1109/34.87344.
51. Meyer, F.; Beucher, S. Morphological segmentation. *J. Vis. Commun. Image Represent.* **1990**, *1*, 21–46. https://doi.org/10.1016/1047-3203(90)90014-M.
52. Pingyuan County Third National Land Survey Leading Group Office; Pingyuan County Natural Resources Bureau; Pingyuan County Bureau of Statistics. *Main Data Bulletin of the Third National Land Survey of Pingyuan County*; Pingyuan County People's Government: Pingyuan, China, 2022. (In Chinese)
53. Lechang Municipal Natural Resources Bureau. *Land Use Status Data of the 2024 Annual National Land Change Survey of Lechang City*; Lechang Municipal People's Government: Lechang, China, 2025. (In Chinese)
54. Taishan City Third National Land Survey Leading Group Office; Taishan City Natural Resources Bureau; Taishan City Bureau of Statistics. *Main Data Bulletin of the Third National Land Survey of Taishan City*; Taishan City People's Government: Taishan, China, 2022. (In Chinese)
55. Qingyuan Municipal Natural Resources Bureau. *Land Use Status Data of the 2020 Annual National Land Change Survey of Qingyuan City*. Qingyuan Municipal People's Government: Qingyuan, China, 2022. (In Chinese).
56. Bai, Y.; Wu, W.; Yang, Z.; Yu, J.; Zhao, B.; Liu, X.; Yang, H.; Mas, E.; Koshimura, S. Enhancement of detecting permanent water and temporary water in flood disasters by fusing Sentinel-1 and Sentinel-2 imagery using deep learning algorithms: Demonstration of Sen1Floods11 benchmark datasets. *Remote Sens.* **2021**, *13*, 2220. https://doi.org/10.3390/rs13112220.

57. Cheng, B.; Girshick, R.; Dollár, P.; Berg, A.C.; Kirillov, A. Boundary IoU: Improving Object-Centric Image Segmentation Evaluation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021*; IEEE: Piscataway, NJ, USA, 2021; pp. 15329–15337. https://doi.org/10.1109/CVPR46437.2021.01508.
58. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Navab, N., Hornegger, J., Wells, W., Frangi, A., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2015; Volume 9351, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
59. Wang, J.; Sun, K.; Cheng, T.; Jiang, B.; Deng, C.; Zhao, Y.; Liu, D.; Mu, Y.; Tan, M.; Wang, X.; et al. Deep high-resolution representation learning for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 3349–3364. https://doi.org/10.1109/TPAMI.2020.2983686.
60. Cordeiro, M.C.R.; Martinez, J.-M.; Peña-Luque, S. Automatic water detection from multidimensional hierarchical clustering for Sentinel-2 images and a comparison with Level 2A processors. *Remote Sens. Environ.* **2021**, *253*, 112209. https://doi.org/10.1016/j.rse.2020.112209.
61. Markert, K.N.; Markert, A.M.; Mayer, T.; Nauman, C.; Haag, A.; Poortinga, A.; Bhandari, B.; Thwal, N.S.; Kunlamai, T.; Chishtie, F.; et al. Comparing Sentinel-1 Surface Water Mapping Algorithms and Radiometric Terrain Correction Processing in Southeast Asia Utilizing Google Earth Engine. *Remote Sens.* **2020**, *12*, 2469. https://doi.org/10.3390/rs12152469.
62. Colacicco, R.; Refice, A.; Nutricato, R.; Bovenga, F.; Caporusso, G.; D'Addabbo, A.; La Salandra, M.; Lovergine, F.P.; Nitti, D.O.; Capolongo, D. High-resolution flood monitoring based on advanced statistical modeling of Sentinel-1 multi-temporal stacks. *Remote Sens.* **2024**, *16*, 294. https://doi.org/10.3390/rs16020294.
63. Yingde Municipal Bureau of Statistics; Survey Office of the National Bureau of Statistics in Yingde. *Yingde Statistical Yearbook 2023*; Yingde Municipal Bureau of Statistics: Yingde, China, 2023.
64. Yingde Municipal Bureau of Statistics; Survey Office of the National Bureau of Statistics in Yingde. *Yingde Statistical Yearbook 2024*; Yingde Municipal Bureau of Statistics: Yingde, China, 2024.
65. Lechang Municipal Bureau of Statistics. *Statistical Bulletin on National Economic and Social Development of Lechang City 2023*; Lechang Municipal People's Government: Lechang, China, 2023. Available online: http://www.lechang.gov.cn/ (accessed on 3 September 2026).
66. Lechang Municipal Bureau of Statistics. *Statistical Bulletin on National Economic and Social Development of Lechang City 2024*; Lechang Municipal People's Government: Lechang, China, 2024. Available online: http://www.lechang.gov.cn/ (accessed on 3 September 2026).
67. Lechang Municipal Bureau of Statistics. *Statistical Bulletin on National Economic and Social Development of Lechang City 2025*; Lechang Municipal People's Government: Lechang, China, 2025. Available online: http://www.lechang.gov.cn/ (accessed on 3 September 2026).
68. Jiangmen Municipal Bureau of Statistics. *Jiangmen Statistical Yearbook 2023*; China Economic and Social Big Data Research Platform (CNKI): Beijing, China, 2023.
69. Jiangmen Municipal Bureau of Statistics. *Jiangmen Statistical Yearbook 2024*; China Economic and Social Big Data Research Platform (CNKI): Beijing, China, 2024.
70. Jiangmen Municipal Bureau of Statistics. *Jiangmen Statistical Yearbook 2025*; China Economic and Social Big Data Research Platform (CNKI): Beijing, China, 2025.
71. Meizhou Municipal Bureau of Statistics; Survey Office of the National Bureau of Statistics in Meizhou. *Meizhou Statistical Yearbook 2023*; Meizhou Municipal Bureau of Statistics: Meizhou, China, 2023.
72. Meizhou Municipal Bureau of Statistics; Survey Office of the National Bureau of Statistics in Meizhou. *Meizhou Statistical Yearbook 2024*; Meizhou Municipal Bureau of Statistics: Meizhou, China, 2024.
73. Meizhou Municipal Bureau of Statistics; Survey Office of the National Bureau of Statistics in Meizhou. *Meizhou Statistical Yearbook 2025*; Meizhou Municipal Bureau of Statistics: Meizhou, China, 2025.
74. Guangdong Rural Statistical Yearbook Editorial Committee. *Guangdong Rural Statistical Yearbook 2025*; China Statistics Press: Beijing, China, 2026.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.