remote sensing

MDPI

Article

Deep-Learning-Based Method for the Identification of Typical Crops Using Dual-Polarimetric Synthetic Aperture Radar and High-Resolution Optical Images

Xiaoshuang Ma [1,2,3,*], Le Li [1] and Yinglei Wu [4]

[1] School of Resources and Environmental Engineering, Anhui University, Hefei 230601, China; x22301067@stu.ahu.edu.cn
[2] Information Materials and Intelligent Sensing Laboratory of Anhui Province, Anhui University, Hefei 230601, China
[3] Engineering Center for Geographic Information of Anhui Province, Anhui University, Hefei 230601, China
[4] China JIKAN Research Institute of Engineering Investigations and Design, Co., Ltd., Xi'an 710021, China; wuyinglei1013@163.com
* Correspondence: mxs.88@whu.edu.cn

Abstract: Timely monitoring of distribution and growth state of crops is crucial for agricultural management. Remote sensing (RS) techniques provide an effective tool to monitor crops. This study proposes a novel approach for the identification of typical crops, including rapeseed and wheat, using multisource remote sensing data and deep learning technology. By adopting an improved DeepLabV3+ network architecture that integrates a feature-enhanced module and an attention module, multiple features from both optical data and synthetic aperture radar (SAR) data are fully mined to take into account the spectral reflectance traits and polarimetric scattering straits of crops. The proposal can effectively address the limitations of using a single data source, alleviating the misclassification problem brought by the spectral similarity of crops in certain bands. Experimental results demonstrate that the proposed crop identification DeepLabV3+ (CI-DeepLabV3+) method outperforms traditional classification methods and the original DeepLabV3+ network, with an overall accuracy and F1 score of 94.54% and 94.55%, respectively. Experimental results also support the conclusion that using multiple features from multi-source data can indeed improve the performance of the network.

Keywords: crop classification; multi-source remote sensing; polarimetric synthetic aperture radar; deep learning

check for updates

Academic Editor: Ruyin Cao

Received: 6 October 2024
Revised: 27 November 2024
Accepted: 12 December 2024
Published: 3 January 2025

Citation: Ma, X.; Li, L.; Wu, Y. Deep-Learning-Based Method for the Identification of Typical Crops Using Dual-Polarimetric Synthetic Aperture Radar and High-Resolution Optical Images. *Remote Sens.* **2025**, *17*, 148. https://doi.org/10.3390/rs17010148

Copyright: © 2025 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Agriculture, serving as the cornerstone of human civilization, has played a crucial role in maintaining social stability and driving economic development [1,2]. With the escalating global population and the impacts of climate change, the demand for precision monitoring of the distribution of crops is in need in recent decades [3]. Remote sensing techniques, leveraging their ability to observe broad regions and continuously deliver images, have become essential tools in agricultural management [4–6]. Thus, it is meaningful work to investigate the use of remote sensing images in identifying the planting areas of typical crops.

According to imaging mechanisms, spaceborne RS sensors are mainly categorized into two classes, namely, optical sensors and synthetic aperture radar (SAR) sensors [7,8]. Due to the capability of providing high-resolution images and obtaining the spectral

reflectance information of land objects, optical RS images have been widely used for the application of crop classification [9,10]. For instance, in [11], a relief algorithm is developed to identify wheat and barley from multi-temporal Sentinel-2 images based on the analyses of phenology. Dong et al. [12] proposed a phenology time-based algorithm to use time-series data from Sentinel-2A/B satellites, aiming at exploiting the distinct phenological traits of winter wheat and mapping their planting areas. Pan et al. [13] developed an approach to accurately map different winter crops by synergistically using Sentinel-2 and Landsat-7/8 imageries within the Google Earth Engine platform. These studies confirm the successful application of optical imagery in monitoring crops; however, limited by weather conditions, the effectiveness of the above studies is hindered since the parameters of crops in some crucial growth stages may be unavailable from optical images contaminated by clouds.

SAR systems can provide reliable data, regardless of weather conditions and working in day and night [14,15]. As a relatively new technique, the polarimetric SAR (PolSAR) system can acquire images containing more polarimetric scattering information of targets. In recent years, some studies about crop classification or mapping using SAR images have been conducted. For example, the study in [16] utilized radar vegetation indices from SAR images to study crops' specific responses to radar signals during the growing season and successfully mapped specific crop types. By analyzing how field size and agricultural climate diversity impact classification performances on 14 different crop types, Arias et al. [17] used images from VH and VV polarimetric channels to accomplish the task of crop classification. Bao et al. [18] introduced some vegetation descriptors to identify crops through Sentinel-1A/B SAR data, which showed positive performances in distinguishing wheat, rapeseed, maize, and soybean. The aforementioned studies confirm the successful application of SAR imagery in crop monitoring.

Optical RS sensors can capture the spectral reflectance information, while SAR sensors can obtain the polarimetric scattering traits of crops. Therefore, in theory, combining multi-source images from both optical and SAR sensors is helpful for the accurate identification of crops. In fact, some studies about crop classification using multi-source images have been presented. For instance, Tian et al. [19] investigated the backscattering coefficients of rice at different growth stages in Sentinel-1A images and the spectral traits of vegetation in Landsat-8 images, aiming at mapping the planting areas of early, middle, and late rice. Adrian et al. [20] discovered that, by integrating the backscattering signal and texture information from Sentinel-1 SAR images and the spectral information from Sentinel-2 images, high classification accuracies in identifying ten different crop types can be obtained. Zhou et al. [21] presented a framework for crop classification, in which an accurate farmland boundary map is detected from high-resolution optical imageries and a deep learning-based algorithm is used to classify parcels using multi-sequence Sentinel-1 SAR imageries.

Although the above studies demonstrate the superiority of combining multi-source images in identifying crops, some major problems still exist: (1) Some studies still focus primarily on using medium or low spatial resolution imagery for crop monitoring, such as Landsat, MODIS, and Sentinel [22–24]. According to definitions by the United States Geological Survey (USGS) and the European Space Agency (ESA), medium resolution typically refers to spatial resolutions of 10–100 m, suitable for regional- and national-scale environmental monitoring, while low resolution refers to spatial resolutions greater than 100 m up to 1 km or more, suitable for global-scale dynamic monitoring. Within these resolution ranges, while they can meet the monitoring needs of large-scale crop-growing areas, their limitations become apparent in geographically complex regions, such as hilly and mountainous areas. In these areas, crop fields are often fragmented, small, and irregularly distributed. In such cases, the use of high-resolution remote sensing data (typically referring to resolutions of 1–10 m or higher) can enhance the distinction between

crop fields and surrounding areas, thereby improving the accuracy and reliability of crop identification. (2) As a kind of important data source, PolSAR images can provide rich polarimetric information of crops. However, to our knowledge, most studies still focus on the use of amplitude or intensity for crop identification and neglect the use of more polarimetric information (for example, polarimetric phase), making the identification models hardly mining the polarimetric scattering mechanisms of crops.

In recent years, deep learning technology has made significant progress in the field of semantic segmentation of remote sensing images, demonstrating outstanding performance, particularly in handling multi-source data and complex scene segmentation tasks [25]. Among these advancements, the core module, Atrous Spatial Pyramid Pooling (ASPP), leverages atrous convolution to flexibly expand the receptive field, enabling efficient extraction of multi-scale contextual information without reducing image resolution. Based on this, DeepLabV3+ further integrates an encoder–decoder structure, achieving a deep fusion of global semantic feature extraction and local detail representation, which significantly enhances the model's segmentation accuracy and applicability in large-scale scene segmentation tasks. For instance, Tallam et al., applied advanced deep learning image segmentation techniques (DeepLab) to analyze high-resolution images acquired by drones, monitoring the distribution dynamics of eelgrass (*Zostera marina*) in the Morro Bay estuary, California [26]. However, DeepLabV3+ still exhibits certain limitations. While its multiple downsampling operations effectively reduce computational complexity, they compromise the retention of fine-grained details, particularly in high-resolution images. This is especially evident when dealing with targets with complex boundaries or densely distributed features, where edge information may become blurred or lost. Additionally, although the ASPP module excels in multi-scale feature fusion, its ability to differentiate extremely small-scale targets remains limited, making it difficult to capture the intricate details of crops. These challenges highlight the need for further optimization of the model to enhance its performance in high-precision segmentation tasks [27].

Based on the above perspective, this study proposes an improved DeepLabV3+ model (CI-DeepLabV3+), which combines the multimodal features of high-resolution GF-2 optical images and Sentinel-1 dual polarimetric synthetic aperture radar (SAR) data to develop a high-precision method for identifying typical crops. In this model, the InceptionNeXt module is introduced to enhance the network's multi-scale feature extraction capabilities through modular design, demonstrating superior performance in the representation of complex boundaries and small targets while significantly accelerating model convergence. Additionally, the model incorporates the Convolutional Block Attention Module (CBAM), which leverages channel and spatial attention mechanisms to dynamically select critical features, effectively focusing on key regions and further improving the precision of feature representation and the interpretability of the network. Unlike traditional approaches, the CI-DeepLabV3+ model innovatively integrates SAR polarimetric features, including polarization amplitude and phase, into the training process. These polarization features not only provide unique semantic information to distinguish crops from the background but also supplement target feature representation when the optical imagery fails to adequately express detailed crop information. Under the framework of multi-scale feature fusion, the synergistic use of optical and SAR data demonstrates significant complementary advantages, substantially enhancing the model's ability to resolve small targets and complex edges.

The structure of this article is as follows. Section 2 details the study area and dataset. Section 3 introduces the proposed method. In Section 4, we describe the experimental results. In Section 5, discussions are made. Finally, Section 6 concludes the paper.

2. Study Area and Dataset

2.1. Study Area

Figure 1 illustrates the study area, located in the central part of Anhui Province, China, at the intersection of Hefei City, Chuzhou City, and Ma'anshan City. This region lies within the middle and lower reaches of the Yangtze River basin, adjacent to the Yangtze River Delta. It is both an important agricultural production base and a representative geographical transition zone. The terrain in this area is characterized by a mix of fertile plains and mountainous regions. The Yangtze River and Huai River flow through the region, depositing rich sediments that naturally enhance soil fertility, creating favorable conditions for crop cultivation. Additionally, this area is situated at the junction of the warm temperate and subtropical climate zones, with distinct seasons and mild, humid weather, showcasing typical transitional climate characteristics.

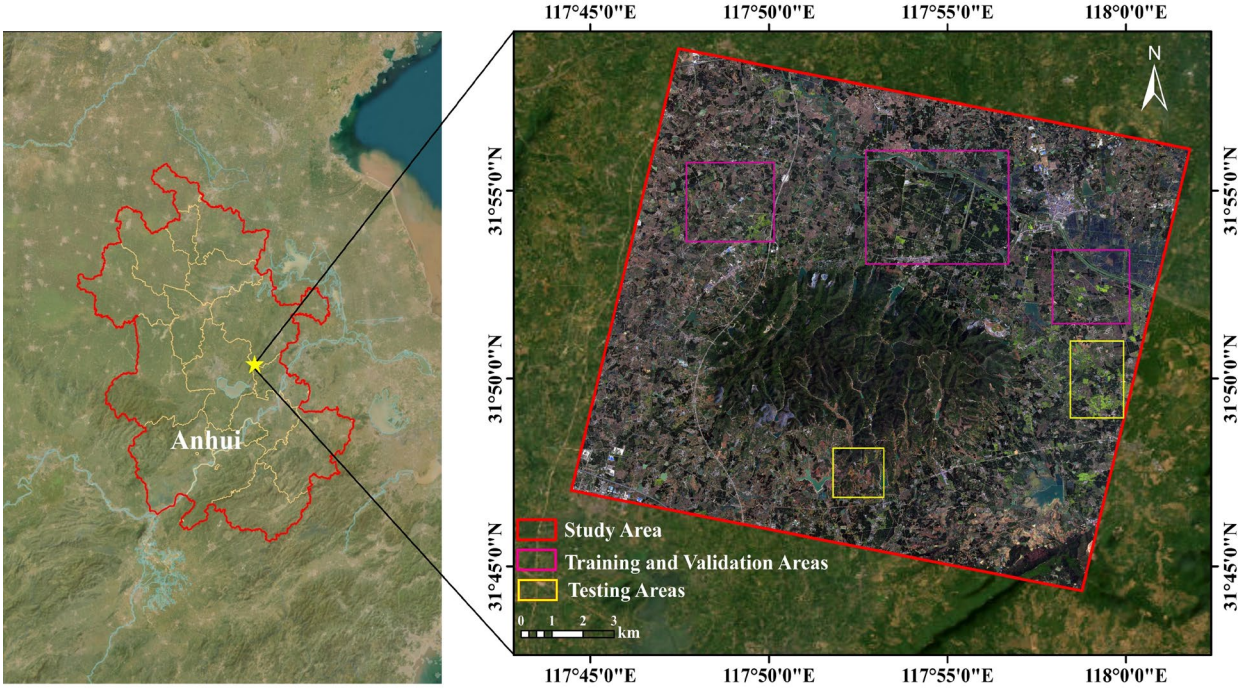


Figure 1. Map of the study area.

The study area spans latitudes from 31.739°N to 31.980°N and longitudes from 117.740°E to 118.030°E. This research utilizes one GF-2 optical high-resolution image and one Sentinel-1 dual-polarization radar image, which comprehensively cover the study area and provide robust data support for analyzing crop distribution and geographical characteristics. The region experiences an annual average temperature of approximately 17 °C and an average annual precipitation of about 1149 mm, with most rainfall occurring during the summer months. These favorable hydrothermal conditions provide an ideal environment for crop growth, further emphasizing the agricultural importance of this area. For data partitioning, the purple-marked regions shown in Figure 1 were designated as the training and validation sets, used for model training and parameter tuning. Meanwhile, the yellow-marked regions were designated as the test sets, aimed at evaluating the classification performance of the model. This data partitioning strategy ensures the independence of the training and testing processes and accurately reflects the model's classification performance across different land cover types.

2.2. Phenological Calendar of Typical Crops in the Study Area

As two typical crop types, wheat and rapeseed were widely planted in the study area, and both of them are single-season crops. The phenological calendar of wheat and rapeseed in the study area is showcased in Figure 2. Normally, wheat is cultivated at the end of October and harvested in early June of the next year, while rapeseed is cultivated in mid-October and harvested at the end of May of the next year. The growth stages of wheat mainly include the sowing stage, tillering stage, jointing stage, heading stage, and maturity stage. The growth stages of rapeseed mainly include the sowing stage, seedling stage, flowering stage, podding stage, and maturity stage.

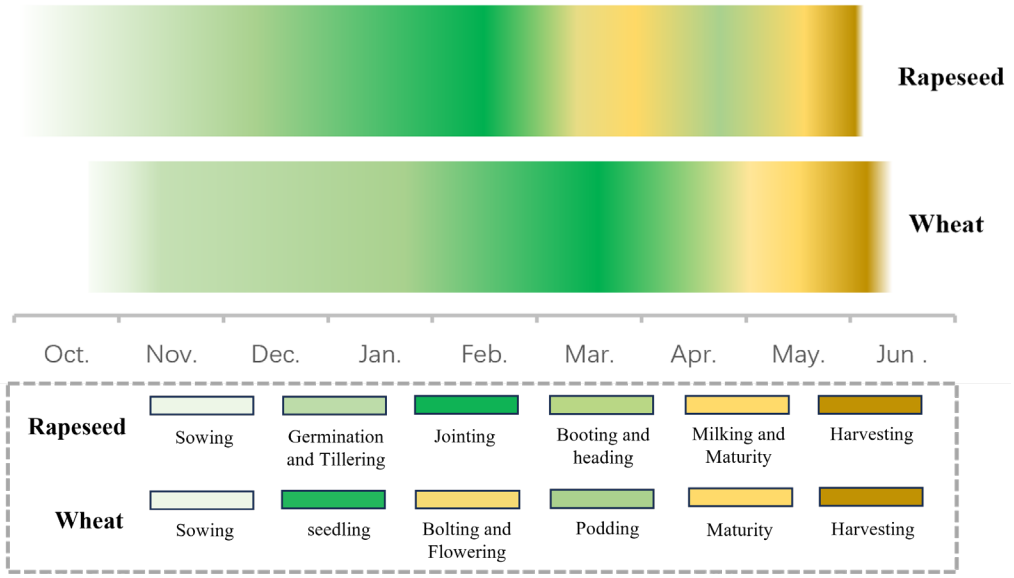


Figure 2. Phenological calendar of the two crops.

During the early phases of the growth cycle, wheat and rapeseed exhibit similar spectral reflectance characteristics, due to the fact that both crops are mainly in the periods of developing leaves and with high chlorophyll content. However, in the middle of March, rapeseed switches to the flowering phase, and the differences in the spectral reflectance characteristics between wheat and rapeseed become notable. Then, in the podding stage and maturity stage, rapeseed flowers wilt, and rapeseed becomes mature. Simultaneously, wheat progresses into the heading and maturity stages, where its spectral characteristics once again show certain similarities with mature rapeseed.

To sum up, to well distinguish rapeseed and wheat, the optimal time window is the flowering stage of rapeseed, which occurs from mid-March to mid-April. This is the principle we choose the date of remote sensing images for crop identification.

2.3. Data Sources

For the identification of crops, multisource remote sensing datasets from the GF-2 optical satellite and Sentinel-1 SAR satellite, both acquired on 7 April 2023 (i.e., the flowering stage of rapeseed), are used in this study.

(1) GF-2 images: The Gaofen-2 (GF-2) satellite is an integral component of China's advanced high-resolution Earth observation system, designed to provide precise surface data to support applications in urban planning, precision agriculture, and environmental monitoring. This satellite is capable of capturing imagery with a spatial resolution of 1 m in panchromatic imaging mode and 4 m in multispectral imaging mode. The fundamental information of the GF-2 satellite is detailed in Table 1. This satellite is capable of capturing multispectral images with a 4 m spatial resolution and panchromatic images with a 1 m

spatial resolution. The revisit cycle of the GF-2 satellite is 5 days, which means that five or six images in the optimal time window for crop identification can be acquired. In addition, the GF-2 image has a wide coverage (23.5 km × 23.5 km), ensuring the monitoring of crops in a large area.

Table 1. Basic information of GF-2 data.

Parameters	Panchromatic/Multispectral Cameras	
Spectral range (μm)	Full Color	0.45–0.90
	Multispectral	0.45–0.52
		0.52–0.59
		0.63–0.69
		0.77–0.89
Spatial resolution	Panchromatic	1 m
	Multispectral	4 m
Width	45 km (2 cameras)	
Revisit cycle (when swinging sideways)	5 days	

Radiometric calibration and atmospheric corrections were applied to reduce radiometric errors due to sensor characteristics induced as well as atmospheric scattering. Orthogonal correction was also applied to correct geometric distortions. Finally, the ortho-corrected multispectral image is fused with the panchromatic image to generate a new multispectral image with improved resolution as the panchromatic image.

(2) Dual-polarimetric SAR data: As a component of the ESA's Copernicus plan, the Sentinel-1 mission provides all-weather and global radar imagery through a C-band SAR system comprising two satellites. It is extensively utilized for crop growth monitoring, disaster management, and environmental surveillance [28]. The Sentinel-1 satellites operate in single and dual polarimetric modes and can work in four imaging modes.

The interferometric wide swath (IW) imaging mode of Sentinel-1, characterized by its extensive coverage and high resolution, is well-suited for large-scale agricultural monitoring. In this study, Single Look Complex (SLC) data acquired in IW mode was employed to monitor crops. The product's details are provided in Table 2.

Table 2. Fundamental details of the PolSAR data.

Sensor Parameter	Description
Polarimetric mode	Dual (VV + VH)
Product Type	SLC
Imaging Mode	IW
Band	C-Band
Swath	250 km
Spatial resolution	5 m × 20 m
Incidence Angle	20–46°

To fully exploit the potential of Sentinel-1 SLC data, this study conducted systematic preprocessing to enhance its applicability in crop monitoring. The preprocessing steps included orbit correction, radiometric calibration, geometric correction, and speckle filtering. First, orbit correction was applied using precise orbit files to mitigate satellite trajectory errors and improve the geometric accuracy of the imagery. Second, radiometric calibration

converted pixel values into backscatter coefficients (Sigma0), facilitating the analysis of surface scattering characteristics. Subsequently, geometric correction was performed to rectify terrain-induced distortions and project the imagery into a geographic coordinate system. Finally, speckle noise was suppressed using the Lee filter, thereby enhancing the visibility of crop features. After completing these preprocessing steps, VV and VH dual-polarization information was extracted to analyze the relationship between polarization characteristics and crop distribution.

The polarimetric information in PolSAR images inherently correlates with the targets' physical structures, offering profound insights into surface roughness, symmetry, and orientation—details not easily derivable from optical images. The Sentinel-1 data utilized in this work consist of VV and VH polarization channels, represented by the following polarimetric scattering vector [29,30]:

$$k = \left[\sqrt{2}S_{\mathrm{HV}}, S_{\mathrm{VV}}\right]^{T}, \tag{1}$$

With

$$S_{\mathrm{HV}} = |S_{\mathrm{HV}}|e^{\varphi_{\mathrm{HV}}i}, \tag{2}$$

where $|S_{\mathrm{HV}}|$ denotes the amplitude of radar signals transmitted in vertical state and received in horizontal state, φ_{HV}, denotes the polarimetric phase, and i represents the imaginary unit. To consider the phase difference between different polarization channels, the polarimetric scattering vector is often transformed to the polarimetric covariance matrix [31], which is

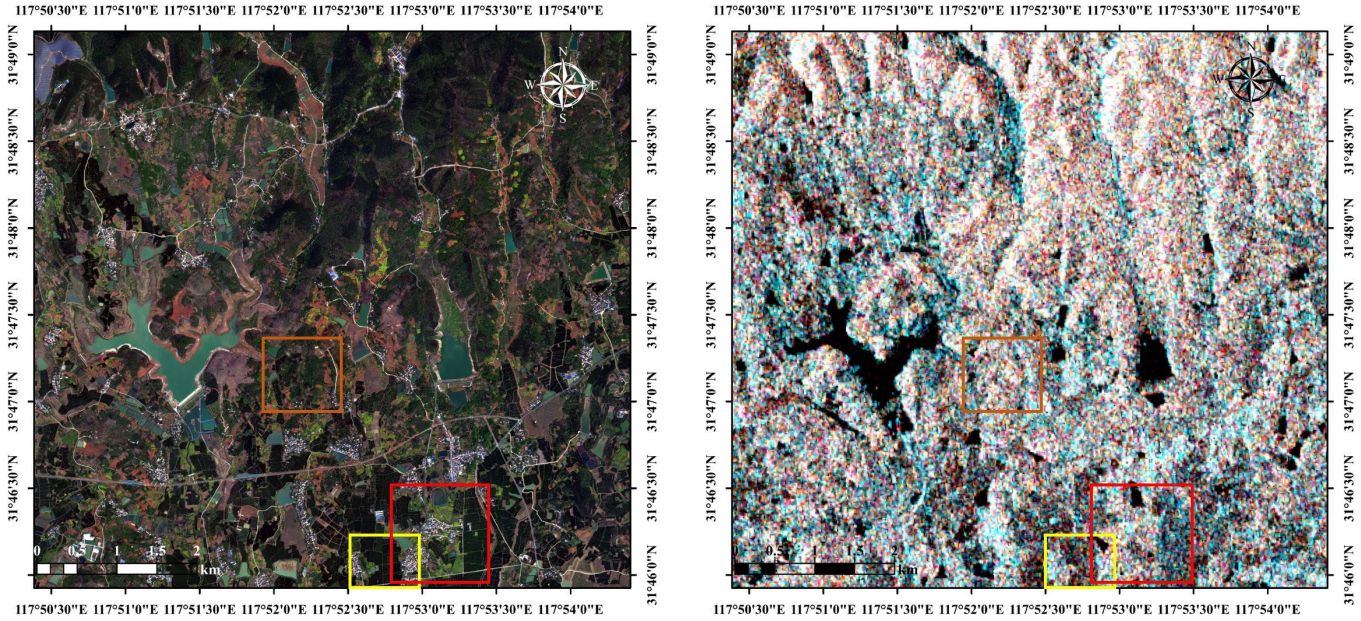
$$C_{2\times 2} = k{\cdot}k^{H} = \begin{bmatrix} 2|S_{\mathrm{HV}}|^{2} & \sqrt{2}S_{\mathrm{HV}}S_{\mathrm{VV}}^{*} \\ \sqrt{2}S_{\mathrm{VV}}S_{\mathrm{HV}}^{*} & |S_{\mathrm{VV}}|^{2} \end{bmatrix}, \tag{3}$$

where H indicates the Hermitian transpose (complex conjugate transpose) and * signifies the conjugate operator.

3. Methodology

3.1. Analyses of Spectral Reflectance and Polarimetric Scattering Traits of Crops

Figure 3 shows the optical image and the PolSAR image of the study area, which was acquired on 7 April 2023 (i.e., the flowering stage of rapeseed), where the areas planting wheats, rapeseeds, and trees are marked out. The RGB composition of the PolSAR image was constructed based on the elements of the polarimetric covariance matrix, effectively visualizing the polarimetric scattering characteristics of land cover. Specifically, the RGB channels were derived from the covariance matrix components C_{11}, C_{12r} (i.e., the real part of C_{12}), and C_{22} to highlight different scattering properties. The blue channel, built using C_{11}, primarily horizontal scattering characteristics, describes and captures the overall reflectance behavior of smooth surfaces. The green channel was calculated using the formula $C_{11} - 2 \times C_{12r} + C_{22}$, which combines horizontal and vertical scattering intensities while suppressing the correlation component C_{12r}, thereby enhancing the representation of double-scattering mechanisms, such as interactions between vegetation canopies and the ground. The red channel employs C_{22}, emphasizing vertical polarimetric scattering characteristics to enhance the visualization of vertical structures like vegetation and buildings. To improve visual clarity and enhance the contrast between land cover types, Min–Max Normalization was applied to all channels, ensuring optimal contrast enhancement. Figure 4 shows some pictures taken in the field in the same period.



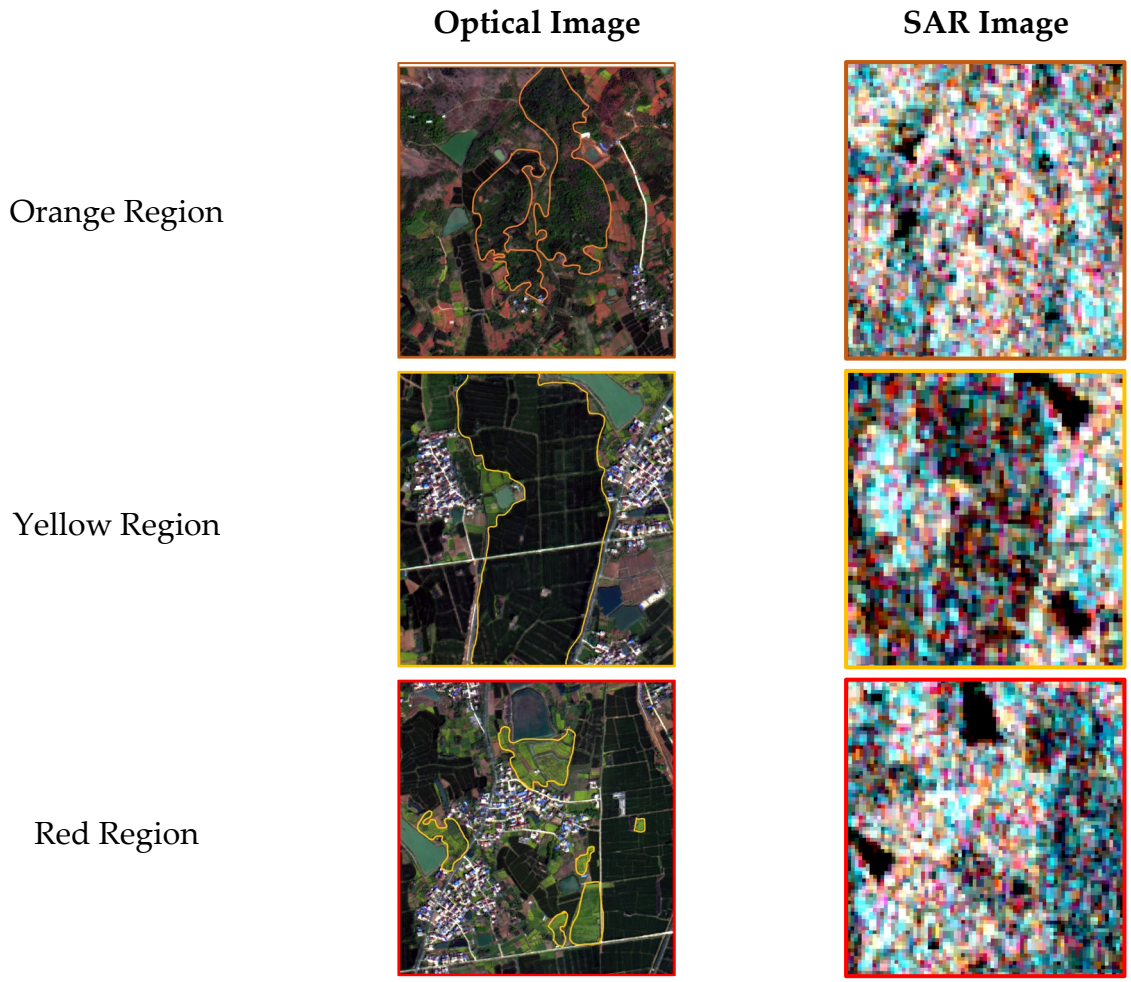


Figure 3. Sub-images of the study area (orange rectangle: forest, yellow rectangle: wheat, red rectangle: rapeseed). (**a**) GF-2 optical image. (**b**) Sentinel-1 PolSAR image.



(**a**) (**b**) (**c**)

Figure 4. Pictures taken in the field. (**a**) Wheat; (**b**) rapeseed; and (**c**) vegetation.

From Figure 3a, it can be observed that the spectral features of vegetation (marked by the orange rectangle) are visually similar to wheat (marked by the yellow rectangle) in the optical image. This similarity is mainly due to the fact that, at this growth stage, the spectral reflectance of both wheat and vegetation is dominated by the chlorophyll reflectance from green leaves. In contrast, rapeseed (marked by the red rectangle) appears visually distinct because of its bright yellow flowers, with its spectral reflectance primarily contributed by the yellow flower corolla.

To further explain this visual similarity and difference, Figure 5 provides the spectral reflectance curves in the visible and near-infrared regions. Particularly in the green band (around 0.55 μm), the reflectance characteristics align with the strong absorption of blue and red light by chlorophyll and the high reflection of green light from green leaves. However, rapeseed (orange line) exhibits significantly higher reflectance, primarily due to the surface reflectance properties of its yellow flower corolla, resulting in higher reflectance in both the green and red bands.

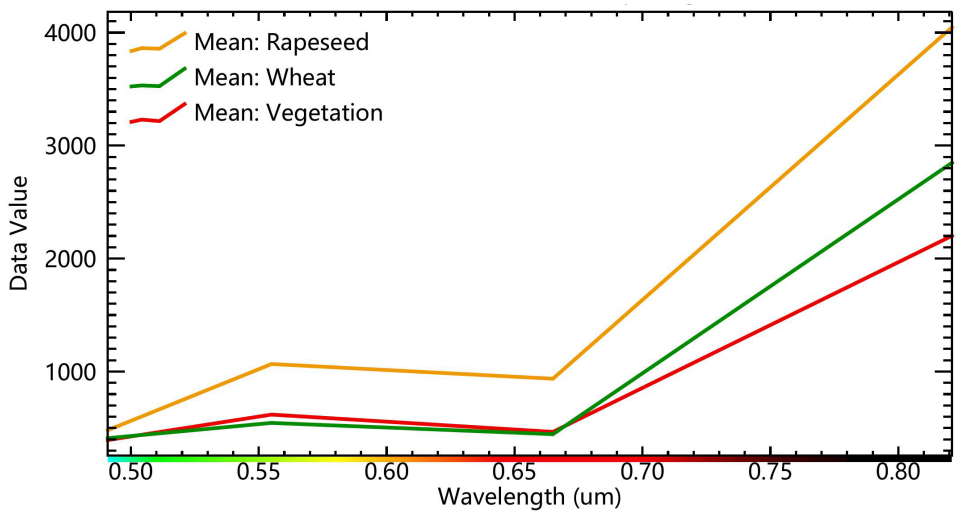


Figure 5. Spectral reflectance of different vegetation type.

From Figure 3b, it can be observed that the visual distinction between wheat and forest is more pronounced, while rapeseed and forest appear more similar in their visual representation. Combined with the profile analysis shown in Figure 6, it is further revealed that the scattering characteristics of different land cover types exhibit significant spatial variability. The profile curve of wheat indicates relatively low scattering intensity with smaller fluctuations, which aligns with its upright stem structure during the jointing stage. This structure causes more radar wave energy to be transmitted to the ground rather than directly back to the radar sensor, resulting in weaker backscatter signals. In contrast, the profile curve of rapeseed shows higher scattering intensity and more complex fluctuations,

reflecting the structural complexity of its canopy during the flowering stage. This complex canopy structure generates stronger volume scattering signals, which is consistent with the visual similarity between rapeseed and forest in Figure 3b.

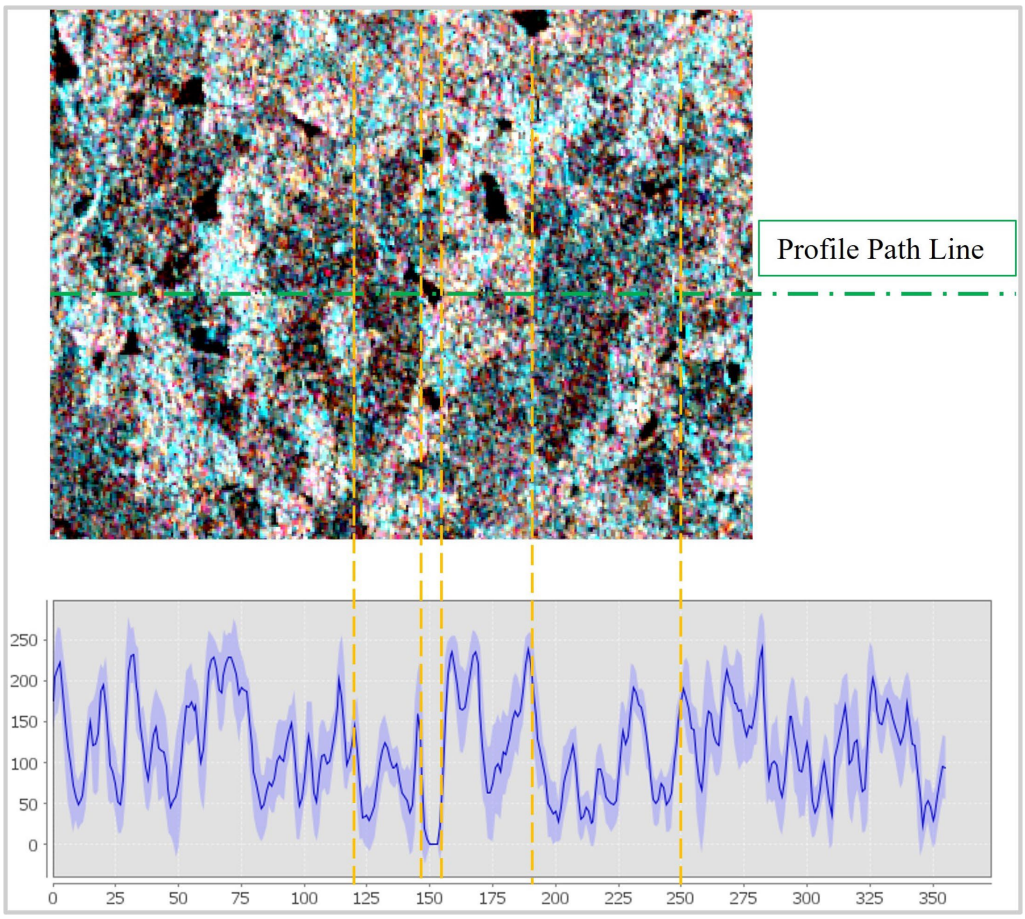


Figure 6. The scattering characteristics of different land cover types.

3.2. Crop Identification Model Based on Improved DeeplabV3+ Network

Among the deep learning-based semantic segmentation networks, DeepLabV3+ stands out as a quintessential model, primarily owing to its innovative use of atrous convolutions. These specialized convolutions adeptly increase the receptive field, capturing a wider array of traits while preserving the resolution of the feature map [32–34]. Considering the above advantages, this study employs the DeepLabV3+ model to classify crop planting areas.

In the DeepLabV3+ architecture, the encoder is a fundamental component that analyzes input images and distills them into comprehensive and high-level representations. These representations capture the essential qualities and details needed for precise image segmentation. On the other side of the architecture, the decoder plays a pivotal role in refining these features, progressively reconstructing the spatial dimensions to generate precise segmentation maps.

While the classical DeepLabV3+ model is recognized for its strengths, it encounters limitations when fully mining high-level semantic features to distinguish different types of crops with similar image characteristics. To overcome these challenges and enhance the model's ability to recognize crops, we have developed an advanced and refined model named CI-DeepLabV3+. CI-DeepLabV3+ builds on the inherent strengths of the DeepLabV3+ framework and introduces two key modules designed to enhance its performance. First, the InceptionNeXt module is embedded within the low-level feature outputs of the encoder, designed to extract and amplify the subtle differences in the original input

features, making them more suitable for subsequent segmentation tasks [35]. This enhancement helps the model achieve a deeper understanding of the image content. Second, a Convolutional Block Attention Module (CBAM) is added after the decoder concatenation to highlight cropping information [36]. The basic structure of the proposed CI-DeepLabV3+ model is showcased in Figure 7. The details of each module are as follows:

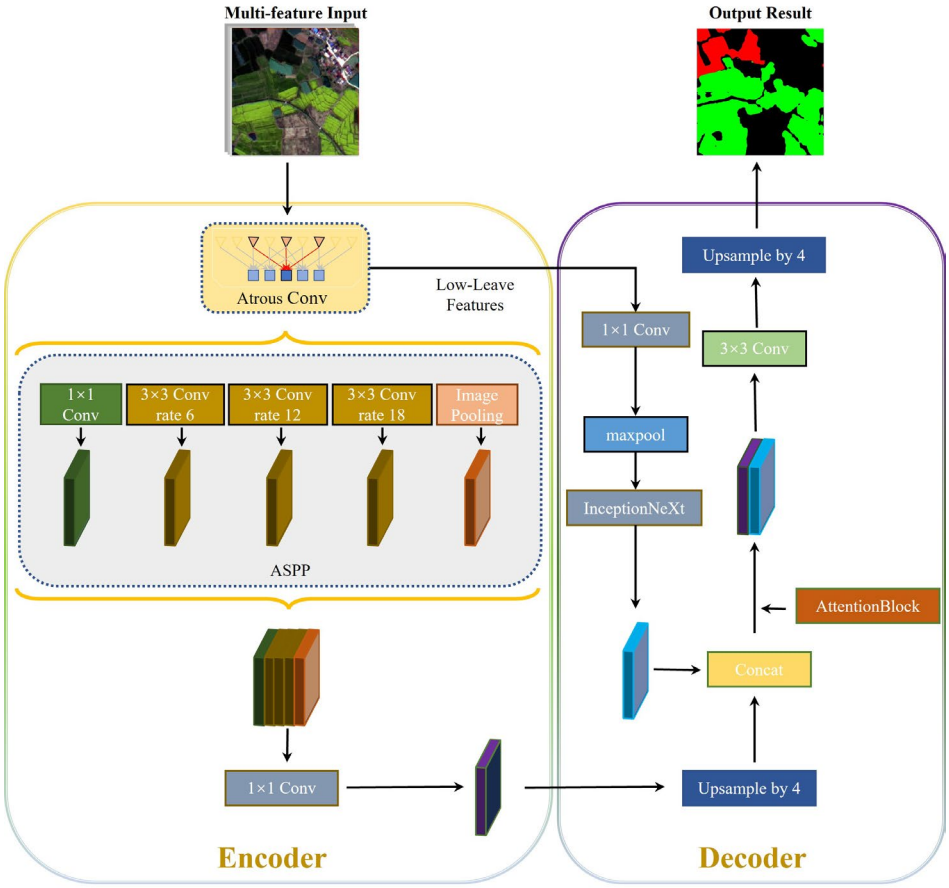


Figure 7. Diagram of the CI-DeepLabV3+ network.

(1) InceptionNeXt module: The operational workflow of this module is shown in Figure 8. The InceptionNeXt module is an advanced convolutional module specifically designed to enhance the efficiency and feature extraction capabilities of deep learning models. Its core innovation lies in decomposing large-kernel depthwise convolutions into four parallel branches: a small square kernel, two orthogonal band kernels, and an identity mapping. This design not only improves the model's ability to capture diverse spatial features but also maintains computational efficiency. Among these branches, the small square kernel captures local spatial features, while the horizontal and vertical band kernels extract features in their respective directions. The identity mapping branch retains essential channel information without adding computational cost, enhancing the interpretability of the network. Depthwise convolutions are applied independently to each input channel, significantly reducing the number of parameters and computational complexity compared to traditional convolutions. The convolution outputs are then combined using pointwise (1×1) convolutions, which integrate features across channels, further enhancing feature fusion. Additionally, the module incorporates batch normalization after convolution to stabilize and accelerate the training process, reducing internal covariate shift. The application of the rectified linear unit (ReLU) activation function introduces non-linearity into the network, enabling it to learn complex feature representations more effectively.

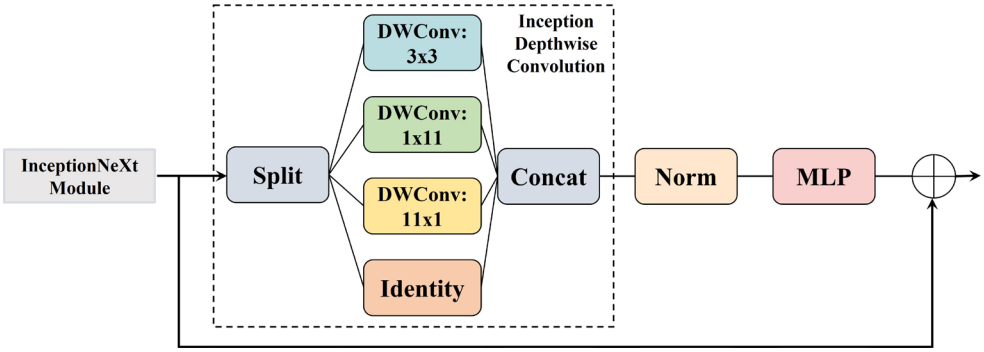


Figure 8. Schematic diagram of the enhanced module structure.

(2) Attention module: Attention mechanisms facilitate prioritization of the most prominent features within a visual scene in human perception [37–40]. CBAM is a module designed to enhance the performance of convolutional neural networks (CNNs), comprising primarily of two sub-modules: channel attention and spatial attention. The channel attention computes global average pooling of feature maps, followed by a series of convolutional and activation layers to generate channel attention weights, dynamically adjusting responses across different channels of the feature map. Spatial attention, on the other hand, generates spatial attention maps through convolutional operations to enhance spatial details within the feature map. The architecture of attention module is illustrated in Figure 9.

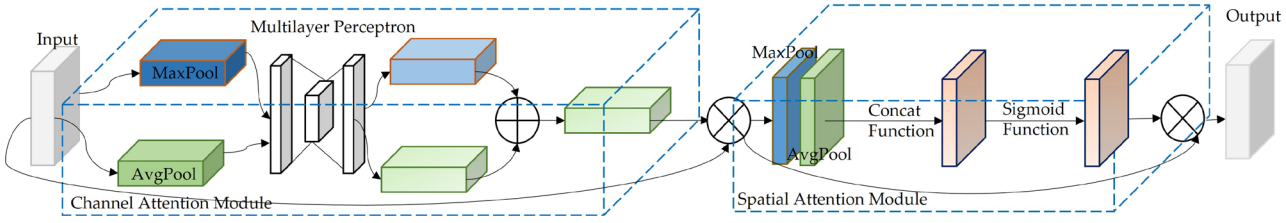


Figure 9. Schematic diagram of the structure of the attention module.

4. Experimental Part

4.1. Experimental Settings

To ensure the network was adequately trained, the study area was divided into training and test sets. The purple regions included three plots, and the sample size was allocated based on their proportional area. A total of 3800 images were randomly selected from the image database to form the training set. Simultaneously, the validation set comprised 20% of the training sample size (760 images), randomly sampled from the training regions. This proportion was determined to balance sample utilization and the ability to validate the model during the training process. For the test set, samples were randomly selected from two distinct regions, each representing different geographical and crop characteristics within the study area.

All samples were cropped to a size of 128 × 128 pixels for the following reasons: Firstly, as the optical images were sourced from GF-2 high-resolution satellite data, 128 × 128-pixel samples effectively captured fine details such as crop textures, boundary features, and other high-resolution characteristics, aiding the model in accurately extracting land-cover details. Secondly, this size encompassed both the target land features and their surrounding context, enabling the network to learn contextual information effectively. Furthermore, it adhered to the input requirements of deep learning models, ensuring computational efficiency while minimizing memory usage. Consequently, the selection of 128 × 128 pixels

as the sample size represents a well-considered balance between the need for detailed data representation and the demands of model training. The experimental hardware and software setups are detailed in Table 3.

Table 3. Computer configurations.

Configuration	Details
CPU	10th Gen Intel(R) Core(TM) i5-10500 @ 4.50 GHz
RAM	16 GB
GPU	NVIDIA GeForce RTX 1080Ti
Language	Python 3.8
Programming	PyCharm 2022.01.01
Framework	PyTorch 1.12.1

To optimize the configuration of model hyperparameters throughout the training phase, this research undertook a comprehensive series of experiments for the meticulous adjustment of various hyperparameters. The outcomes suggest that optimal performance is achieved with a learning rate of 0.0001, a batch size of 10, and the employment of the Adam optimizer. We provided the well-trained CI-DeeplabV3+ model for reproducing our work at GitHub—lile13955267907/CI-DeeplabV3-plus.

4.2. Evaluation Metrics

To objectively evaluate the effectiveness of the model in recognizing wheat and rapeseed, this study used the micro-averaging method to calculate several evaluation metrics, such as precision, recall, and F1 scores. Micro-averaging is a common method for calculating evaluation metrics and has great advantages in assessing the overall performance, especially when dealing with unbalanced datasets or when a comprehensive comparison of the performance of multiple models is required.

Overall accuracy (OA) denotes the proportion of samples accurately classified by the model relative to the overall sample count. The formula to compute OA is

$$\mathrm{OA} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN} + \mathrm{TN}}, \tag{4}$$

where TP, TN, FP, and FN correspond to the counts of true positive, true negative, false positive, and false negative samples, respectively.

Micro-precision measures the classification accuracy of positive samples for all classes by calculating the proportion of pixels correctly classified out of all correctly classified pixels

$$\mathrm{Micro-Precision(M-P)} = \frac{\sum_{i=1}^{N} \mathrm{TP}_i}{\sum_{i=1}^{N} (\mathrm{TP}_i + \mathrm{FP}_i)}, \tag{5}$$

where N denotes the class number.

Micro-Recall measures the recognition ability of the model for positive samples across all classes, i.e., the proportion of samples correctly classified as positive out of all true positive samples

$$\mathrm{Micro-Recall(M-R)} = \frac{\sum_{i=1}^{N} \mathrm{TP}_i}{\sum_{i=1}^{N} (\mathrm{TP}_i + \mathrm{FN}_i)}. \tag{6}$$

The micro-average F1 Score makes a compromise between micro-average recall and micro-average precision, which is

$$\mathrm{F1} = \frac{2 \times \mathrm{M-P} \times \mathrm{M-R}}{\mathrm{M-P} + \mathrm{M-R}}. \tag{7}$$

The intersection over union (IOU) indicator measures the extent of congruence between the predicted and actual results. The formula for calculating the IOU for each category i is as follows:

$$\mathrm{IOU}_i = \frac{\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FP}_i + \mathrm{FN}_i}. \tag{8}$$

MIOU is the IOU averaged over all categories, i.e.,

$$\mathrm{MIOU} = \frac{1}{\mathrm{N}} \sum_{i=1}^{N} \mathrm{IOU}_i. \tag{9}$$

This study employs multiple evaluation metrics to comprehensively assess the model's performance in crop classification tasks. The choice of these metrics is based on their ability to measure different aspects of the model's classification effectiveness, thus providing a more comprehensive performance analysis. OA is the most basic metric, used to evaluate the model's overall classification accuracy across all categories. However, due to the class imbalance in crop types, relying solely on OA may obscure classification errors in certain classes. Therefore, we introduced the micro-average F1 score and MIOU for more specific class-level evaluation.

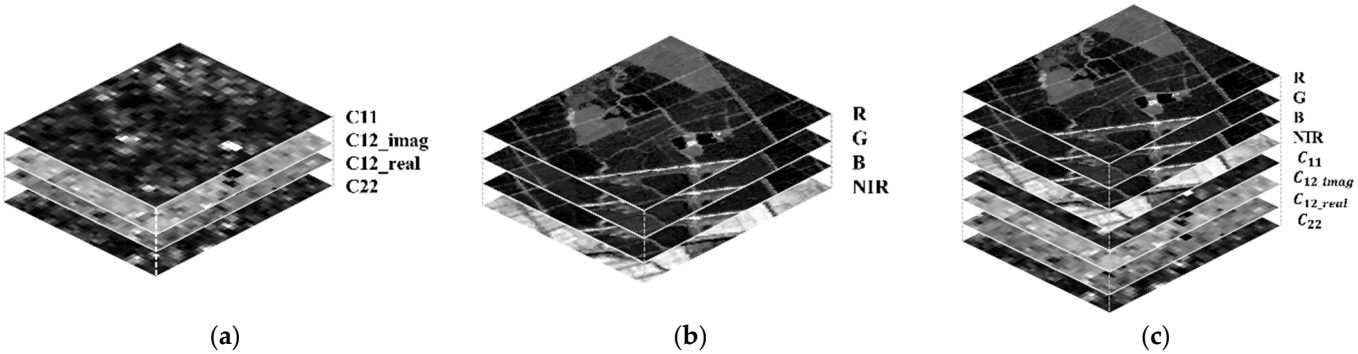
The F1 score balances precision and recall by combining M-P and M-R. In crop classification, certain small patches of crop classes may lead to inconsistencies between precision and recall. The F1 score effectively balances these two metrics, ensuring that the model performs more robustly in complex scenarios and in recognizing small patches. MIOU, on the other hand, focuses on pixel-level classification accuracy and is particularly suited for crop classification tasks requiring precise boundary delineation. MIOU quantifies the model's ability to distinguish boundaries between different crops, ensuring more accurate performance in scenarios with complex boundaries. This is especially crucial for crops like wheat and rapeseed, which may exhibit similar characteristics along their boundaries.

4.3. Experimental Results

(1) Comparison between different feature sets: This study used GF-2 optical images and Sentinel-1 dual-polarimetric SAR images to construct a feature set for crop classification. To explore the advantages of multiple feature sets in classification performance, we constructed three feature sets: a PolSAR feature set, an optical feature set, and an integrated optical and PolSAR feature set, as shown in Figure 10.

The PolSAR feature set primarily includes dual-polarization information (VV and VH channels) and their scattering characteristics, which effectively capture the structural information of crops. During the tillering stage of wheat, its vertical structure results in a lower radar backscatter coefficient, enabling effective differentiation from other crop types. However, due to the similarity in polarization scattering mechanisms between rapeseed and forests, the PolSAR feature set has limitations in distinguishing these two targets. The optical feature set derived from the spectral information of GF-2 imagery performs excellently in distinguishing different crop types, particularly in classifying crops based on their spectral reflectance characteristics. However, this feature set has limitations in detecting small-scale crop patches, especially in regions with complex boundaries, where classification accuracy tends to be lower. The integrated optical and PolSAR feature set combines the spectral information from optical imagery and the structural features from

PolSAR imagery, fully exploiting the complementary advantages of multi-source data. The optical features provide the spectral reflectance characteristics of crops, helping to distinguish crops at different growth stages, while the PolSAR features capture the geometric shape and sparsity of the crops, addressing the shortcomings of optical data in representing structural information. This integrated feature set effectively distinguishes wheat from rapeseed and substantially enhances classification accuracy in boundary regions.



(**a**) (**b**) (**c**)

Figure 10. (**a**) Classification result with Feature Set 1: utilizing solely PolSAR feature information; (**b**) classification result with Feature Set 2: utilizing solely optical feature information; and (**c**) classification result with Feature Set 3: concurrently leveraging both optical and PolSAR feature information.

The classification results based on different feature sets are presented in Figure 11, and the corresponding quantitative assessment results are listed in Table 4. Through the comparison of classification results from different feature sets, it was found that the model trained with the PolSAR feature set has limitations in recognizing rapeseed, mainly due to the similarity in scattering mechanisms between rapeseed and forest. This results in poor performance in distinguishing these two classes using the PolSAR feature set. The optical feature set, on the other hand, achieves an overall accuracy of 88% by exploiting spectral differences but performs poorly in detecting small-scale crop patches and handling boundary blurring issues. In contrast, integrating optical features with PolSAR features for combined classification fully utilizes the complementary advantages of both, significantly improving the classification performance in small-scale crop regions and demonstrating excellent accuracy in identifying complex boundaries. This further demonstrates the critical role of multi-source data fusion in enhancing classification accuracy and overcoming the limitations of traditional feature sets.

Table 4. Results of the CI-DeepLabV3+ model trained on different feature sets.

Dataset	OA (%)	F1 (%)	MIOU (%)
Feature Set 1	62.32	54.09	29.18
Feature Set 2	88.76	88.42	77.79
Feature Set 3	94.54	94.55	85.11

(2) Comparison of CI-DeepLabV3+ with other methods: For a thorough assessment of the CI-DeepLabV3+ model's performance in classification tasks, it was compared with some classical supervised classification techniques and deep learning-based semantic segmentation methods. Specifically, we explored the Support Vector Machine (SVM) [41], a traditional supervised classification technique, alongside U-Net and the canonical DeepLabV3+ deep learning semantic segmentation algorithms [42]. SVM is renowned for its robust small sample learning capabilities and stable handling of high-dimensional data, attributed to the model's emphasis on margin maximization. On the other hand, U-Net and DeepLabV3+, as classic semantic segmentation algorithms in the field of deep learning, owe their efficiency

to their complex network architecture and unique data augmentation methods. Their powerful feature extraction and learning mechanisms have demonstrated outstanding performance across numerous domains. The comparison of these algorithms aims to highlight the advantages of the CI-DeepLabV3+ in managing complex classification scenarios.

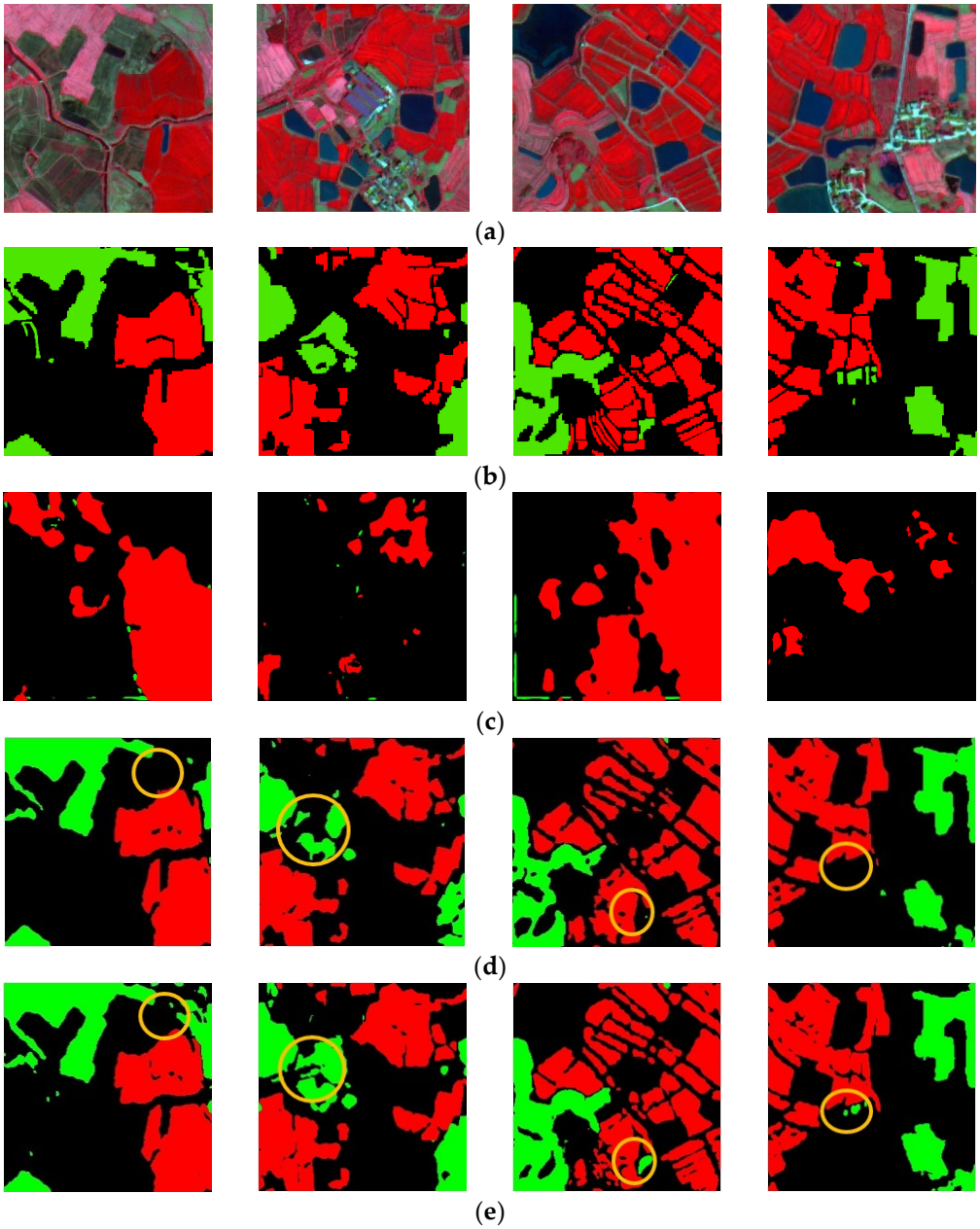


Figure 11. Classification results of the CI-DeepLabV3+ model with different feature sets (wheat is marked in red and rapeseed is marked in green): (**a**) GF-2 optical image (false color); (**b**) ground truth label image; (**c**) classification result with Feature Set 1; (**d**) classification result with Feature Set 2; and (**e**) classification result with Feature Set 3.

The semantic segmentation algorithm was trained utilizing feature sets from both optical and PolSAR data. The classification diagram is presented in Figure 12, while the quantitative evaluation results of the various methods are detailed in Table 5.

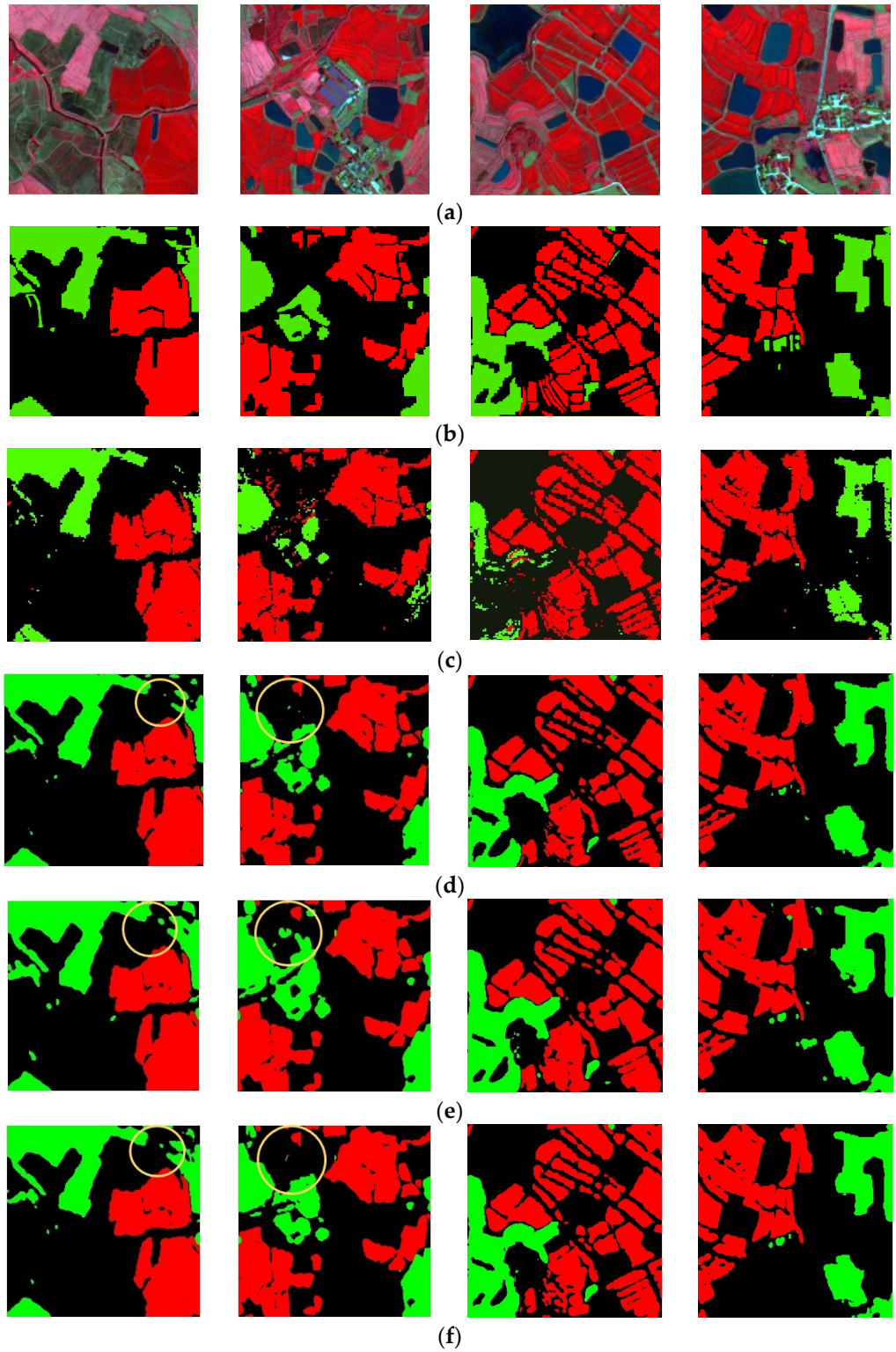


Figure 12. Classification results of different methods. (**a**) GF-2 optical image (false color display); (**b**) ground truth label image; (**c**) classification result with SVM; (**d**) classification result with U-net; (**e**) classification result with DeepLabV3+; and (**f**) classification result with CI-DeepLabV3+.

Table 5. Classification results of different methods.

Method	OA (%)	F1 (%)	MIOU (%)
SVM	89.12	88.61	64.47
U-net	90.74	90.53	81.81
DeeplabV3+	91.06	90.87	82.44
CI-DeepLabV3+	94.54	94.55	85.11

The experimental results demonstrate that the CI-DeepLabV3+ model exhibits superior overall performance, significantly surpassing other comparative models. In contrast, the traditional supervised classification method, SVM, lacks automatic feature extraction capabilities and struggles to handle high-dimensional, multi-feature remote sensing data. When dealing with non-linear relationships and complex backgrounds, SVM is susceptible to noise and environmental factors, resulting in imprecise classification boundaries, with its mean intersection over union (MIOU) barely reaching 64%. This highlights the limitations of SVM in complex scenarios. Another comparative model, U-Net, mainly relies on convolutional downsampling and upsampling operations in its network structure, which weakens its ability to capture local details of the image. Consequently, its performance in classifying edge and fragmented regions is suboptimal. Despite the fact that traditional methods use filtering techniques to suppress certain types of noise, the inherent speckle noise in optical and SAR images continues to challenge these approaches. In contrast, the CI-DeepLabV3+ model enhances the original DeepLabV3+ architecture by introducing advanced feature extraction mechanisms and classical attention modules, significantly improving its overall performance. Specifically, the model integrates the InceptionNeXt and CBAMs, which further augment its capabilities. The InceptionNeXt module is built upon the latest advancements in deep learning, combining the design strengths of the Inception module and the ConvNeXt framework [43,44]. By decomposing large kernel depthwise convolutions into multiple parallel branches, this module retains the superior multi-scale feature extraction capabilities of the Inception module while inheriting the efficient convolutional architecture of ConvNeXt. This design enables the module to capture intricate image details and rich spatial features while substantially reducing computational cost. Furthermore, the CBAM, as a well-established attention mechanism, effectively enhances feature selection capabilities by combining channel attention and spatial attention. Its design dynamically assigns higher weights to key regions in feature maps, increasing the model's sensitivity to target areas while reducing its reliance on redundant information. The combination of these two modules not only improves the model's robustness to noise but also maximizes the utilization of amplitude and phase information in SAR data. These improvements significantly boost the model's overall accuracy (OA) and F1 scores, achieving more precise boundary delineations for wheat and rapeseed, as demonstrated by experimental results.

(3) Impact of optimizers and learning rates on model performance: In the training process of neural network models, the accurate configuration of hyperparameters is crucial because it directly impacts the model's effectiveness and the precision of the results. Particularly, underfitting and overfitting are frequent challenges in training, and precise hyperparameter tuning can effectively alleviate these issues. Through extensive experiments and the exploration of various parameter combinations, we delved into the impact of optimizer choice on model performance. In this study, we specifically examined the Adam and RMSprop optimizers, which have high flexibility and efficiency with different data types and network architectures and showcase broad adaptability in deep learning. Concurrently, the learning rate indicates the extent of network weight updates in the optimization procedure. To ensure that gradient descent achieves optimal performance, it is imperative

to set the learning rate within an appropriate range. Experiments were conducted with various combinations of optimal optimizers and learning rates. The quantitative evaluation results are displayed in Figure 13. Evidently, the optimal outcomes are obtained when employing the Adam optimizer with an initial learning rate of 0.0001.



Figure 13. The results for different combinations of hyperparameters.

(4) Challenges and improvements in crop identification under dispersed farming patterns: The study area is primarily composed of dispersed farmlands independently managed by small-scale farmers. The lack of commercialization and large-scale management has resulted in diverse and scattered cropping patterns. This dispersed farming model, coupled with varying planting times chosen by different farmers, leads to significant differences in crop growth stages, posing substantial challenges to the accurate identification of crop types and growth conditions. To address these challenges, the CI-DeepLabV3+ model integrates the InceptionNeXt and CBAMs, effectively capturing the diverse spatial and spectral features of crops while leveraging attention mechanisms to focus on critical areas. Moreover, multi-source data fusion further exploits the complementary advantages of optical and SAR data, enhancing the overall performance of the model. As shown in Table 6, the classification results indicate that rapeseed slightly outperforms wheat in terms of overall accuracy (OA), intersection over union (IOU), and F1 score. This may be attributed to the unique polarimetric scattering characteristics of rapeseed, which the model can effectively capture. These findings validate the adaptability and robustness of the CI-DeepLabV3+ model in complex agricultural scenarios characterized by diversity, heterogeneity, and noise.

Table 6. Classification accuracy of wheat and rapeseed.

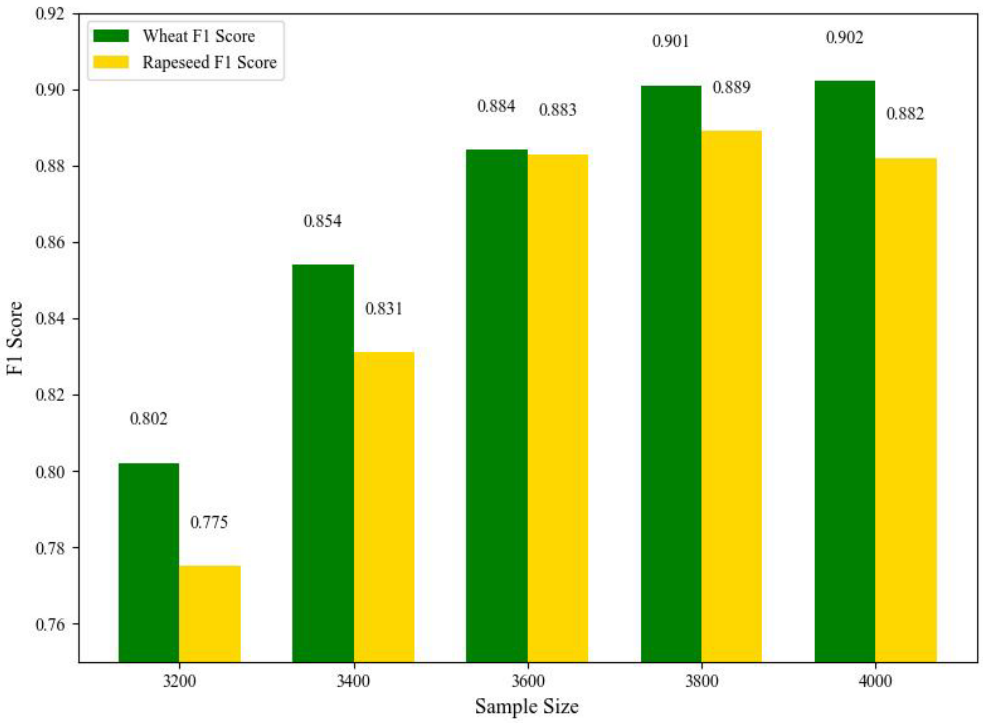
	OA (%)		**IOU (%)**		**F1 (%)**	
Improved DeeplabV3+	wheat 94.62	rapeseed 93.45	wheat 81.94	rapeseed 80.55	wheat 90.08	rapeseed 88.99

5. Discussion

5.1. Analysis of the Impact of Sample Size on Classification Performance

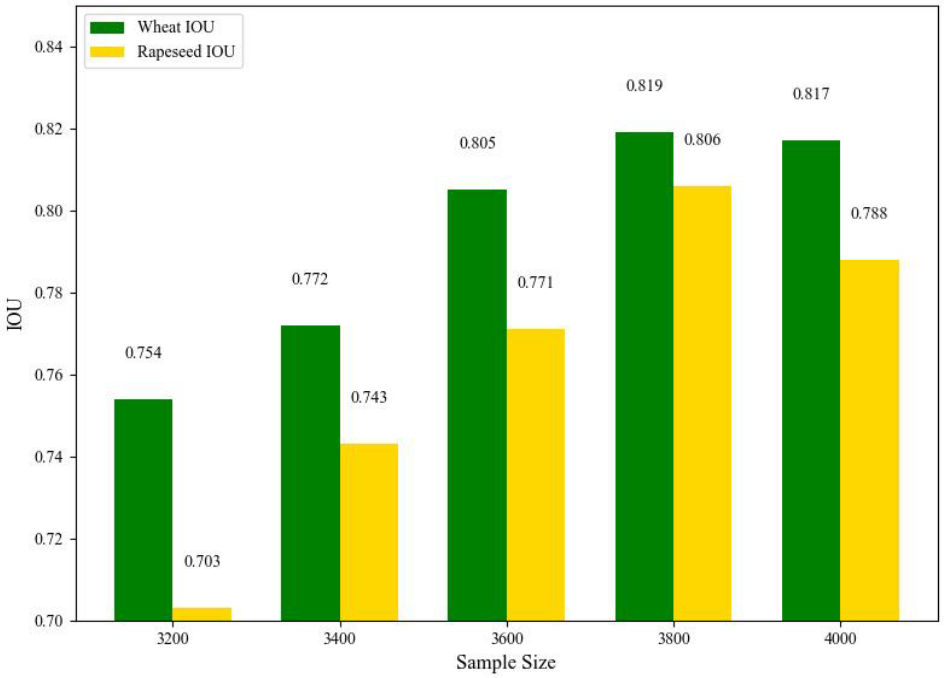
We conducted an in-depth analysis of the impact of different sample numbers on model performance. The results indicate that as the sample size increased from 3200 to 3800, the model showed continuous improvement across various metrics, including overall accuracy (OA), F1 score, and intersection over union (IoU). This trend underscores the critical role of larger sample sizes in enhancing model training. However, when the sample size further increased to 4000, the rate of improvement leveled off, suggesting diminishing marginal returns from adding more samples.

Specifically, as shown in Figure 14, panels (a), (b), and (c), respectively, illustrate that when the sample size increased from 3200 to 3800, the improvement in classification performance was particularly significant. For example, the overall accuracy (OA) of wheat increased from 85.5% to 94.6%, while the F1 score rose from 80.2% to 90.1%. However, when the sample size further increased to 4000, although both the OA and F1 score for wheat exhibited slight increments (0.2% and 0.1%, respectively), the improvement was marginal. Similarly, for rapeseed, the OA and F1 score reached near-peak levels at a sample size of 3800, after which the growth trend flattened or even exhibited slight fluctuations. This indicates that at a sample size of 3800, the model effectively leveraged the spatial and spectral features in the data while avoiding the increased model complexity caused by redundant samples.

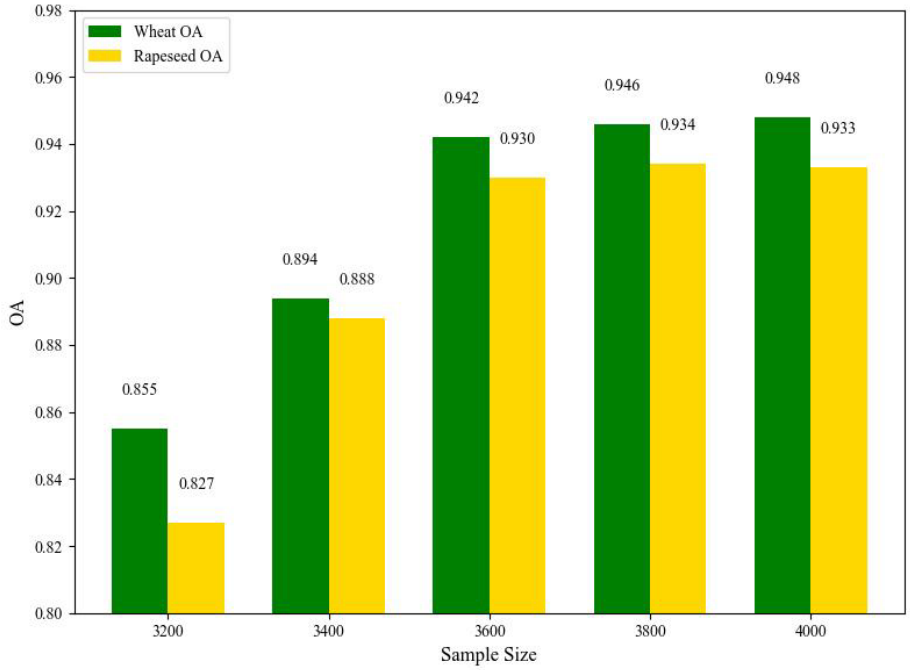


(**a**)

Figure 14. *Cont.*



(**b**)



(**c**)

Figure 14. Performance metrics for varying sample sizes. (**a**) F1 score for wheat and rapeseed; (**b**) IOU for wheat and rapeseed; and (**c**) OA for wheat and rapeseed.

5.2. Generalization Ability Evaluation

In this study, to validate the replicability and scalability of the proposed model, we tested it using optical and SAR data acquired in the same study area during the second year. These new datasets encompass dynamic variations in crop growth conditions, aiming to assess the model's adaptability in multi-temporal scenarios. Experimental results demonstrate that the model maintains high classification performance on the second-year data, reflecting its robustness and potential for cross-temporal applications. Specifically, as shown in Table 7, the model exhibits good classification accuracy and consistency,

particularly for the wheat category, showcasing strong reliability and stability. Although the classification performance for rapeseed is slightly lower than that of wheat, it still achieves a high level of accuracy. This difference in rapeseed classification performance can be attributed to the reduced area of rapeseed cultivation in the second year, as small-scale farmers in the region opted to decrease rapeseed planting.

Table 7. Performance of the model on the second-year dataset.

	OA (%)		**IOU (%)**		**F1 (%)**	
Improved DeeplabV3+	wheat 90.44	rapeseed 88.63	wheat 80.67	rapeseed 76.61	wheat 86.43	rapeseed 80.08

6. Conclusions

Wheat, a principal cereal crop, and rapeseed, a leading oilseed crop, are pivotal in securing global food security and enhancing agricultural productivity. Accurate mapping of their distribution is essential for formulating effective agricultural policies, guaranteeing food security, and coordinating regional development strategies. Distribution patterns of wheat and oilseeds are generally continuous on a large scale but fragmented at the local level. The complexity of these patterns, coupled with the diversity of neighboring environments in the growing areas, renders the precise recognition of these areas challenging. This paper presents a CI-DeepLabV3+ model for identifying wheat and rapeseed cultivation areas. In particular, the images from the wheat nodulation period and the rape flowering period, which are the most distinguishable periods of the two crop growth stages, were utilized for model training. By combining the intensity and phase information of PolSAR data with optical data, the limitations of a single data source in crop identification were successfully overcome. This fusion approach not only improves the sensitivity of the model to crop growth stages and scattering characteristics but also enhances the robustness of the model under complex environmental conditions. The experimental results validate the pleasing precision of the CI-DeepLabV3+ model in identifying wheat and rapeseed planting areas. However, certain limitations in this study should be acknowledged. First, the research was conducted in a geographically limited area, which may not fully reflect the diversity of global agricultural conditions. The applicability of the model to regions with different climatic, soil, or crop management conditions remains to be evaluated. Additionally, the current approach relies on labeled datasets, which require considerable time and effort for annotation. This could hinder the scalability of the methodology in large-scale or real-time applications.

Future research will focus on extending this methodology to include additional data sources, such as multi-temporal and high-resolution SAR data, to further enhance classification accuracy. Efforts will also be made to incorporate automated annotation techniques and self-supervised learning strategies to reduce dependency on manual labeling and improve efficiency. These advancements aim to enable more effective and scalable crop monitoring systems, particularly in regions with fragmented and small-scale farming.

Author Contributions: Conceptualization, X.M. and L.L.; methodology, X.M.; software, X.M. and L.L.; validation, L.L.; formal analysis, L.L.; investigation, X.M. and Y.W.; resources, Y.W.; data curation, X.M.; writing—original draft preparation, X.M.; writing—review and editing, X.M. and L.L.; visualization, X.M.; supervision, X.M.; project administration, Y.W.; funding acquisition, X.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Natural Science Foundation of Anhui Province, China, under Grant 2308085MD116, and in part by the Hefei Municipal Natural Science Foundation, China, under Grant 2021041.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: Author Yinglei Wu was employed by the company China JIKAN Research Institute of Engineering Investigations and Design, Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Li, S.P.; Gong, Q.X.; Yang, S.L. A Sustainable, Regional Agricultural Development Measurement System Based on Dissipative Structure Theory and the Entropy Weight Method: A Case Study in Chengdu, China. *Sustainability* **2019**, *11*, 5313. [CrossRef]
2. Liu, Y.S.; Zhou, Y. Reflections on China's food security and land use policy under rapid urbanization. *Land Use Policy* **2021**, *109*, 105699. [CrossRef]
3. Qiu, Z.J.; He, W.; Li, W.M. A Note on the Adoption of Precision Agriculture in Eastern China. *Outlook Agric.* **2007**, *36*, 255–257. [CrossRef]
4. Tian, H.F.; Qin, Y.C.; Niu, Z.; Wang, L.; Ge, S.S. Summer Maize Mapping by Compositing Time Series Sentinel-1A Imagery Based on Crop Growth Cycles. *J. Indian Soc. Remote Sens.* **2021**, *49*, 2863–2874. [CrossRef]
5. Chang, Z.; Li, H.; Niu, Z.; Chen, D.H.; Liu, Y.F.; Zou, C.; Chen, J.; Han, W.J.; Liu, S.S.; Zhang, N.M. Crop Type Identification Using High-Resolution Remote Sensing Images Based on an Improved DeepLabV3+ Network. *Remote Sens.* **2023**, *15*, 5088. [CrossRef]
6. Ren, S.J.; Guo, H.; Wu, X.; Zhang, L.G.; Ji, M.; Wang, J.; Chen, J. Winter wheat planted area monitoring and yield modeling using MODIS data in the Huang-Huai-Hai Plain, China. *Comput. Electron. Agric.* **2023**, *182*, 106049. [CrossRef]
7. Tian, H.F.; Pei, J.; Huang, J.X.; Li, X.C.; Wang, J.; Zhou, B.Y.; Qin, Y.C.; Wang, L. Garlic and Winter Wheat Identification Based on Active and Passive Satellite Imagery and the Google Earth Engine in Northern China. *Remote Sens.* **2020**, *12*, 3539. [CrossRef]
8. Chen, J.Y.; Han, H.F.; Xu, M.M.; Wan, J.H.; Liu, S.W. MMF-CNN: A multimodal fusion CNN network for winter wheat extraction incorporating active and passive time series data. *Int. J. Remote Sens.* **2020**, *45*, 1796–1825. [CrossRef]
9. Li, Q.J.; Liu, J.; Mi, X.F.; Yang, J.; Yu, T. Object-oriented crop classification for GF-6 WFV remote sensing images based on Convolutional Neural Network. *Natl. Remote Sens. Bull.* **2021**, *25*, 549–558. [CrossRef]
10. Zhou, Q.B.; Yu, Q.Y.; Liu, J.; Wu, W.B.; Tang, H.J. Perspective of Chinese GF-1 high-resolution satellite data in agricultural remote sensing monitoring. *J. Integr. Agric.* **2017**, *16*, 242–251. [CrossRef]
11. Ashourloo, D.; Nematollahi, H.; Huete, A.; Aghighi, M.; Shahrabi, H.S.; Azadbakht, M.; Goodarzdashti, S. A new phenology-based method for mapping wheat and barley using time-series of Sentinel-2 images. *Remote Sens. Environ.* **2022**, *280*, 113206. [CrossRef]
12. Dong, Q.; Chen, X.H.; Chen, J.; Zhang, C.S.; Liu, L.C.; Cao, X.; Zang, Y.Z.; Zhu, X.F.; Cui, X.H. Mapping Winter Wheat in North China Using Sentinel 2A/B Data: A Method Based on Phenology-Time Weighted Dynamic Time Warping. *Remote Sens.* **2020**, *12*, 1274. [CrossRef]
13. Pan, L.; Xia, H.M.; Zhao, X.Y.; Guo, Y.; Qin, Y.C. Mapping Winter Crops Using a Phenology Algorithm, Time-Series Sentinel-2 and Landsat-7/8 Images, and Google Earth Engine. *Remote Sens.* **2021**, *13*, 2510. [CrossRef]
14. Clauss, K.; Ottinger, M.; Kuenzer, C. Mapping rice areas with Sentinel-1 time series and superpixel segmentation. *Int. J. Remote Sens.* **2018**, *39*, 1399–1420. [CrossRef]
15. Ma, X.S.; Wang, C.; Yin, Z.X.; Wu, P.H. SAR Image Despeckling by Noisy Reference-Based Deep Learning Method. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 8807–8818. [CrossRef]
16. Gella, G.W.; Bijker, W.; Belgiu, M. Mapping crop types in complex farming areas using SAR imagery with dynamic time warping. *ISPRS J. of Photogramm. Remote Sens.* **2021**, *175*, 171–183. [CrossRef]
17. Arias, M.; Campo-Bescós, M.A.; Álvarez-Mozos, J. Crop Classification Based on Temporal Signatures of Sentinel-1 Observations over Navarre Province, Spain. *Remote Sens.* **2020**, *12*, 278. [CrossRef]
18. Bao, X.; Zhang, R.; Lv, J.C.; Wu, R.Z.; Zhang, H.S.; Chen, J.; Zhang, B.; Ouyang, X.Y.; Liu, G.X. Vegetation descriptors from Sentinel-1 SAR data for crop growth monitoring. *ISPRS J. Photogramm. Remote Sens.* **2023**, *203*, 86–114. [CrossRef]
19. Tian, H.F.; Wu, M.Q.; Wang, L.; Niu, Z. Mapping Early, Middle and Late Rice Extent Using Sentinel-1A and Landsat-8 Data in the Poyang Lake Plain, China. *Sensors* **2018**, *18*, 185. [CrossRef]
20. Adrian, J.; Sagan, V.; Maimaitijiang, M. Sentinel SAR-optical fusion for crop type mapping using deep learning and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 215–235. [CrossRef]
21. Zhou, Y.N.; Luo, J.C.; Feng, L.; Yang, Y.P.; Chen, Y.H.; Wu, W. Long-short-term-memory-based crop classification using high-resolution optical images and multi-temporal SAR data. *GISci. Remote Sens.* **2019**, *56*, 1170–1191. [CrossRef]
22. Segarra, J.; Buchaillot, M.L.; Araus, J.L.; Kefauver, S.C. Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy* **2020**, *10*, 641. [CrossRef]

23. Hu, Q.; Yin, H.; Friedl, M.A.; You, L.; Li, Z.; Tang, H.; Wu, W. Integrating coarse-resolution images and agricultural statistics to generate sub-pixel crop type maps and reconciled area estimates. *Remote Sens. Environ.* **2021**, *258*, 112365. [CrossRef]
24. Choudhary, K.; Wenzhong, S.; Mukesh, S.B.; Samuel, C. Agriculture phenology monitoring using NDVI time series based on remote sensing satellites: A case study of Guangdong, China. *Opt. Mem. Neural Netw.* **2019**, *28*, 204–214. [CrossRef]
25. Wang, Y.; Wang, C.; Wu, H.; Chen, P. An improved Deeplabv3+ semantic segmentation algorithm with multiple loss constraints. *PLoS ONE* **2022**, *17*, e0261582. [CrossRef]
26. Tallam, K.; Nguyen, N.; Ventura, J.; Fricker, A.; Calhoun, S.; O'Leary, J.; Walter, R.K. Application of Deep Learning for Classification of Intertidal Eelgrass from Drone-Acquired Imagery. *Remote Sens.* **2023**, *15*, 2321. [CrossRef]
27. Anilkumar, P.; Venugopal, P. An adaptive multichannel DeepLabv3+ for semantic segmentation of aerial images using improved Beluga whale optimization algorithm. *Multimed. Tools Appl.* **2024**, *83*, 46439–46478. [CrossRef]
28. Ma, X.S.; Huang, Z.Y.; Zhu, S.Y.; Fang, W.; Wu, Y.L. Rice Planting Area Identification Based on Multi-Temporal Sentinel-1 SAR Images and an Attention U-Net Model. *Remote Sens.* **2022**, *14*, 4573. [CrossRef]
29. Sun, W.D.; Li, P.X.; Du, B.; Yang, J.; Tian, L.L.; Li, M.Y.; Zhao, L.L. Scatter Matrix Based Domain Adaptation for Bi-Temporal Polarimetric SAR Images. *Remote Sens.* **2020**, *12*, 658. [CrossRef]
30. Ma, X.S.; Xu, J.G.; Pan, J.; Yang, J.; Wu, P.H.; Meng, X.C. Detection of marine oil spills from radar satellite images for the coastal ecological risk assessment. *J. Environ. Manag.* **2023**, *325*, 116637. [CrossRef]
31. Lee, J.S.; Pottier, E. *Polarimetric Radar Imaging: From Basics to Applications*; CRC Press: Boca Raton, FL, USA, 2009.
32. Das, S.; Fime, A.A.; Siddique, N.; Hashem, M.M.A. Estimation of Road Boundary for Intelligent Vehicles Based on DeepLabV3+ Architecture. *IEEE Access* **2021**, *9*, 121060–121075. [CrossRef]
33. Du, S.J.; Du, S.H.; Liu, B.; Zhang, X.Y. Incorporating DeepLabv3+ and object-based image analysis for semantic segmentation of very high resolution remote sensing images. *Int. J. Digit. Earth* **2021**, *14*, 357–378. [CrossRef]
34. Wang, X.F.; Kang, M.L.; Chen, Y.; Jiang, W.X.; Wang, M.Y.; Weise, T.; Tan, M.; Xu, L.X.; Li, X.L.; Zou, L.; et al. Adaptive Local Cross-Channel Vector Pooling Attention Module for Semantic Segmentation of Remote Sensing Imagery. *Remote Sens.* **2023**, *15*, 1980. [CrossRef]
35. Yu, W.; Zhou, P.; Yan, S.C.; Wang, X. InceptionNeXt: When Inception Meets ConvNeXt. In Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024; pp. 5672–5683. Available online: https://ieeexplore.ieee.org/document/10657355/authors%E2%80%9D%5Cl%E2%80%9Cauthors (accessed on 1 November 2024).
36. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Convolutional block attention module. In *Proceedings of the European Conference on Computer Vision*; Springer: Cham, Switzerland, 2018; pp. 3–19. Available online: https://arxiv.org/abs/1807.06521 (accessed on 12 April 2020).
37. Chen, P.; Zhou, H.; Li, Y.; Liu, B.X.; Liu, P. Oil Spill Identification in Radar Images Using a Soft Attention Segmentation Model. *Remote Sens.* **2022**, *14*, 2180. [CrossRef]
38. Seydi, S.T.; Amani, M.; Ghorbanian, A. A Dual Attention Convolutional Neural Network for Crop Classification Using Time-Series Sentinel-2 Imagery. *Remote Sens.* **2022**, *14*, 498. [CrossRef]
39. Yin, M.M.; Chen, Z.B.; Zhang, C.J. A CNN-Transformer Network Combining CBAM for Change Detection in High-Resolution Remote Sensing Images. *Remote Sens.* **2023**, *15*, 2406. [CrossRef]
40. Wang, C.S.; Zhang, R.; Chang, L.L. A Study on the Dynamic Effects and Ecological Stress of Eco-Environment in the Headwaters of the Yangtze River Based on Improved DeepLab V3+ Network. *Remote Sens.* **2022**, *14*, 2225. [CrossRef]
41. Löw, F.; Michel, S.; Dech, S.; Conrad, C. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using Support Vector Machines. *ISPRS J. Photogramm. Remote Sens.* **2013**, *85*, 102–119. [CrossRef]
42. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241. Available online: http://link.springer.com/10.1007/978-3-319-24574-4_28 (accessed on 9 July 2023).
43. Li, X.; Zhang, Y.; Liu, Y.; Yao, X.; Zhou, X. Arbitrary shape text detection fusing InceptionNeXt and multi-scale attention mechanism. *J. Supercomput.* **2024**, *80*, 25484–25509. [CrossRef]
44. Liu, Z.; Mao, H.; Wu, C.Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A convnet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 11976–11986. Available online: https://arxiv.org/abs/2201.03545 (accessed on 22 January 2023).

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.