



## Article

# R-Unet: A Deep Learning Model for Rice Extraction in Rio Grande do Sul, Brazil

Tingyan Fu <sup>1</sup>, Shufang Tian <sup>1,\*</sup> and Jia Ge <sup>2</sup>

<sup>1</sup> School of Earth Sciences and Resources, China University of Geosciences (Beijing), Beijing 100083, China; 2001210195@email.cugb.edu.cn

<sup>2</sup> Oil and Gas Resources Investigation Center of China Geological Survey, Beijing 100083, China; gejia@mail.cgs.gov.cn

\* Correspondence: sftian@cugb.edu.cn; Tel.: +86-010-8232-2163

**Abstract:** Rice is one of the world's three major food crops, second only to sugarcane and corn in output. Timely and accurate rice extraction plays a vital role in ensuring food security. In this study, R-Unet for rice extraction was proposed based on Sentinel-2 and time-series Sentinel-1, including an attention-residual module and a multi-scale feature fusion (MFF) module. The attention-residual module deepened the network depth of the encoder and prevented information loss. The MFF module fused the high-level and low-level rice features at channel and spatial scales. After training, validation, and testing on seven datasets, R-Unet performed best on the test samples of Dataset 07, which contained optical and synthetic aperture radar (SAR) features. Precision, intersection, and union (IOU), F1-score, and Matthews correlation coefficient (MCC) were 0.948, 0.853, 0.921, and 0.888, respectively, outperforming the baseline models. Finally, the comparative analysis between R-Unet and classic models was completed in Dataset 07. The results showed that R-Unet had the best rice extraction effect, and the highest scores of precision, IOU, MCC, and F1-score were increased by 5.2%, 14.6%, 11.8%, and 9.3%, respectively. Therefore, the R-Unet proposed in this study can combine open-source sentinel images to extract rice timely and accurately, providing important information for governments to implement decisions on agricultural management.

**Keywords:** rice extraction; MFF; SAR; deep learning; residual network



**Citation:** Fu, T.; Tian, S.; Ge, J.

R-Unet: A Deep Learning Model for Rice Extraction in Rio Grande do Sul, Brazil. *Remote Sens.* **2023**, *15*, 4021. <https://doi.org/10.3390/rs15164021>

Academic Editors: Won-Ho Nam and Jingxiang Zhang

Received: 26 July 2023

Revised: 8 August 2023

Accepted: 12 August 2023

Published: 14 August 2023



**Copyright:** © 2023 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

Human life is inseparable from food crops such as rice, wheat, and corn. The output of food crops determines the economic development of a country, and ensuring food security is also one of the national strategies [1–3]. As one of the three most important food crops in the world, rice is widely planted in Asia, Africa, and America. Therefore, timely and accurate extraction of rice planting areas plays a decisive role in guiding national agricultural production [4,5]. Since the end of the last century, the successive launches of remote sensing (RS) satellites have provided abundant data sources for the application of RS technology. RS images are characterized by timeliness, periodicity, and repetition and can provide timely data for the extraction of crop planting areas at the national and provincial levels. Therefore, RS technology has become one of the most important means in the field of agricultural research, such as rice mapping [6,7].

In recent years, optical data (Landsat, Sentinel-2, etc.) have been widely used in agricultural RS fields such as rice planting extraction [8–13] and crop yield estimation [14,15]. Scholars mainly extracted rice based on the vegetation indices (VIs) of the rice growth period, such as the normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), land surface water index (LSWI) [16–19], etc. Based on the time-series EVI and LSWI of rice, ref. [17] formulated a decision tree rule to extract rice in India. However, in tropical and subtropical rice-growing areas, affected by the monsoon, a large amount of cloud cover will affect the optical image quality. In contrast, radar images

can successfully acquire microwave information about ground objects in cloudy, foggy, and rainy weather [20,21]. Meanwhile, rice contains a large amount of water during the growth period, and the dielectric constant of the water is large, which is different from other types of ground features [22,23]. Therefore, in addition to optical images, radar images are important data sources in the field of rice extraction research, such as synthetic aperture radar (SAR) Sentinel-1 [24–26]. Recently, scholars have carried out rice mapping work by obtaining time-series SAR parameters during the rice growth period, such as the backscattering coefficient ( $\sigma_0$ ) and polarization decomposition parameters [27–30]. Ref. [28] used the random forest (RF) method to perform rice mapping in Hanoi, Vietnam, based on rice time-series  $\sigma_0$ . Existing studies on rice extraction using single optical or SAR data are extensive, but few studies combine the two. A single data source cannot extract the optical and SAR features of rice at the same time, but the combination of the two can fuse optical and SAR features to improve the accuracy of rice mapping.

So far, rice extraction methods mainly include two types, one is traditional machine learning (ML) methods, such as support vector machines [13], decision trees [24,25,31], RF [13,28,32], etc., and the other is deep learning (DL) methods. Traditional ML has been widely used in rice extraction studies. The traditional ML model has a simple structure and can only learn the shallow features of rice, with poor robustness and limited precision. Recently, DL, as a popular research method in the field of computer vision, has been widely used in target detection and other fields due to its unique network structure, high robustness, and excellent fitting ability [33–36]. With the rapid development and innovation of DL, scholars have applied DL to RS image classification and obtained satisfactory results [37,38]. The semantic segmentation model is a classic model in DL that is widely used in the field of medical image classification [39]. At present, scholars have used deep semantic segmentation models, such as U-Net [38–42], SegNet [43,44], FCN [45,46], etc., to successfully complete rice mapping. U-Net was first proposed and applied to medical image segmentation by scholars. However, in the research on rice extraction, the U-Net model structure is relatively simple, and the precision of rice mapping is limited. Recently, scholars have extracted rice by improving the U-Net model [38,40]. Ref. [38] proposed a new network, MobileUNet, based on polarization decomposition parameters of rice, combining U-Net and MobileNet, and MobileUNet had a better rice extraction effect. Ref. [40] extracted rice based on time-series  $\sigma_0$ , using U-Net as the architecture, combined with VGG16 and other backbones, and the rice extraction effect was better than U-Net. The above method can extract the rice planting area, but it does not take into account the multi-scale feature information of rice. Furthermore, in the field of rice extraction research, there are limited DL models focusing on rice extraction.

In short, there are three main questions in existing studies:

Q1: The rice extraction work using only high-resolution images is difficult to obtain, has a high cost, and has poor timeliness.

Q2: Rice extraction using single optical or SAR data has been extensively studied, while studies combining the two are limited. Multi-source data can simultaneously extract the optical and SAR features of rice, and the fusion of optical and SAR features can improve the accuracy of the model to a certain extent.

Q3: The traditional ML method has been widely used, but it cannot learn the deep features of rice and has poor robustness. DL algorithms are able to solve the above problems, but there is a lack of DL models focusing on rice extraction.

Aiming at the above problems, the Sentinel-2 and time-series Sentinel-1 images of the study area were pre-processed on the GEE platform to obtain the optical index and time-series  $\sigma_0$  of rice. Meanwhile, a new deep learning model, R-Unet, focusing on rice extraction, was proposed. The multi-scale feature fusion (MFF) module and the attention-residual module were introduced in R-Unet. The main highlights of this study are as follows:

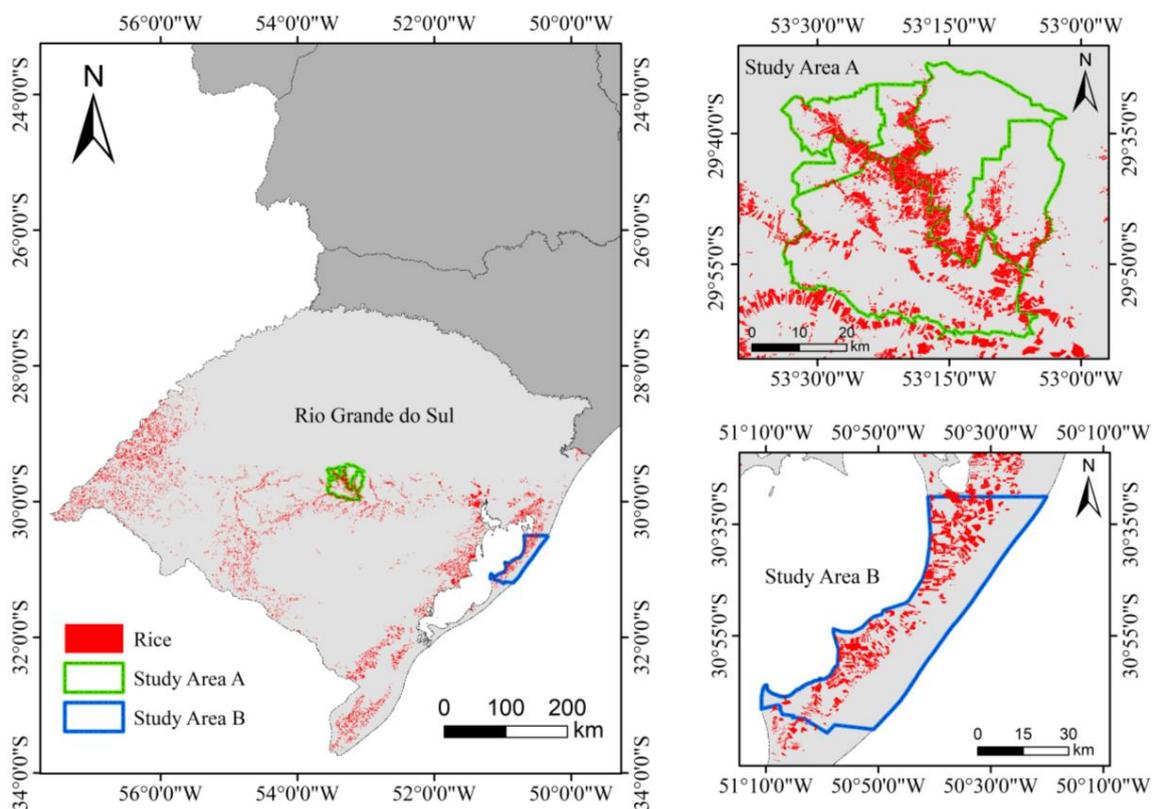
- The MFF module was introduced into the model decoder, where dilated convolution was used to increase the model receptive field;
- in order to prevent the encoder from overfitting and gradient disappearance, a residual network structure was introduced to deepen the depth of the network model;
- the channel attention mechanism was introduced into the attention-residual module to focus on the temporal features of rice;
- the use of open-source RS images (Sentinel-1 and Sentinel-2) is low in cost, easy to obtain, high in timeliness, and can integrate optical and time-series SAR features of rice.

Comparing the test results of different baseline models on seven datasets of optical and time-series SAR features, it was found that R-Unet performed best in the rice extraction research, outperforming other models. Meanwhile, compared with other classic models, the rice extraction accuracy of R-Unet is still the highest.

## 2. Materials and Methods

### 2.1. Study Area

The state of Rio Grande do Sul in southern Brazil was selected as the study area, and study areas A and B were used as training and testing areas for rice extraction, respectively (Figure 1). As the largest country in South America, Brazil's terrain mainly includes the Amazon Plain, the Paraguay Basin, the Brazilian Plateau, and the Guyana Plateau. Brazil is a major rice producer in South America and one of the top ten rice producers in the world [40]. Rice is the main food crop in Brazil, and its main production areas are located in the south, mid-west, and northeast of the country. The southern region is dominated by Rio Grande do Sul. The study area is located in the southernmost part of Brazil, with a subtropical climate, high temperature, and rich water resources, which are extremely suitable for the growth of rice [40]. Apart from rice, soybeans, wheat, and corn are also the main crops in this area.



**Figure 1.** Geographical location map of the study area.

## 2.2. Data Preparation

### 2.2.1. Sentinel-1 Data

In our study, the time-series Sentinel-1 ground range detected (GRD) data of interferometric wide swath (IW) mode and dual-polarization (VH and VV) C-band were obtained. The time-series is from August 2019 to May 2020, covering the main growth season of rice for a total of 10 months [40]. Based on the Google Earth Engine (GEE) platform, the collected time-series dual-polarization SAR images were pre-processed. Figure 2 shows the pre-processing of the Sentinel-1 time-series data, mainly including calibration and terrain correction [10]. Meanwhile, the average value of the monthly SAR data was calculated [10], and finally, the monthly average value of the dual-polarization backscatter coefficient ( $\sigma_0$ ) in the rice growth period was obtained, with a total of 20 bands.

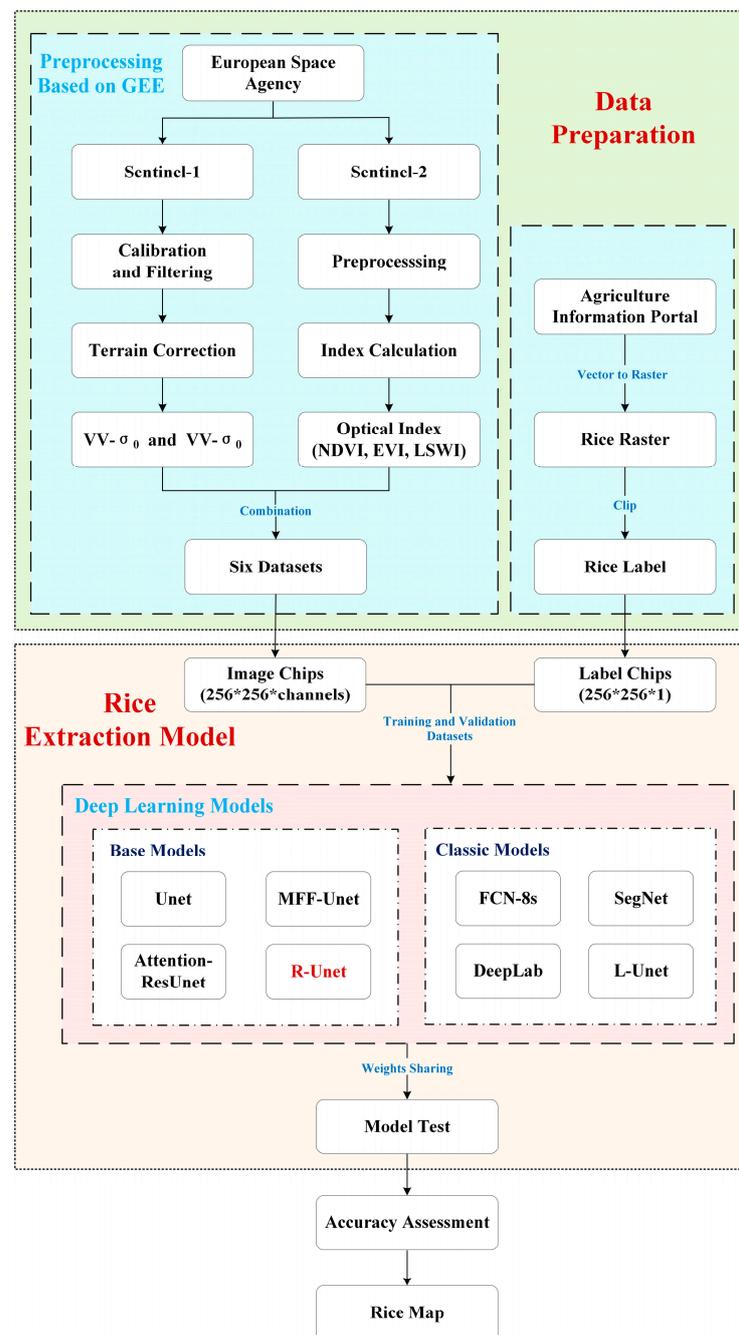


Figure 2. The research workflow.

### 2.2.2. Sentinel-2 Data

Based on the GEE platform, Sentinel-2 data with cloud cover of less than 5% in the same period were collected. The GEE platform was used to perform cloud mask processing on all images [10], and the optical indices of rice were calculated using Formulas (1)–(3) [10]. Finally, the median value of the three optical indices during the rice growth period was calculated, and the results of NDVI, EVI, and LSWI were finally obtained, with a total of three bands.

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}} \quad (1)$$

$$\text{EVI} = 2.5 * \frac{\text{NIR} - \text{Red}}{\text{NIR} + 6 * \text{Red} - 7.5 * \text{Blue} + 1} \quad (2)$$

$$\text{LSWI} = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWIR}} \quad (3)$$

where blue, red, NIR and SWIR represent Sentinel-2 band 2, band 4, band 8, and band 11, respectively.

### 2.2.3. The Ground Truth Data

The ground-truth rice map of Rio Grande do Sul was downloaded from the Brazilian Agricultural Information Portal “<https://portaldeinformacoes.conab.gov.br> (accessed on 7 July 2023)”. The rice map is shapefile data, which is converted into raster data to produce rice labels.

### 2.2.4. Training, Validation and Test Samples

Study area A was selected as the training and validation area of the rice extraction model, which contained 222 rice and non-rice pixel samples of 256 \* 256 pixels. Among them, 90% of the samples were used for model training, and the rest were used for validation. The test area was study area B, which tested the performance of different models, including 150 samples of 256 \* 256 pixels. The training, validation, and test samples each contained 7 datasets. Table 1 shows the dataset’s information. Datasets 01–03 were the SAR datasets; Dataset 04 was the optical index dataset; and Datasets 05–07 were the combination of SAR and optical indices.

**Table 1.** Datasets of rice extraction model.

Datasets	Input Bands	Channels	Description
Dataset 01	VH	10	VH: vertical–horizontal polarization
Dataset 02	VV	10	VV: vertical–vertical polarization
Dataset 03	VH + VV	20	Indices: NDVI + EVI + LSWI
Dataset 04	Indices	3	NDVI: normalized difference
Dataset 05	VH + indices	13	vegetation index
Dataset 06	VV + indices	13	EVI: enhanced vegetation index
Dataset 07	VH + VV + indices	23	LSWI: land surface water index

### 2.3. Research Technical Route

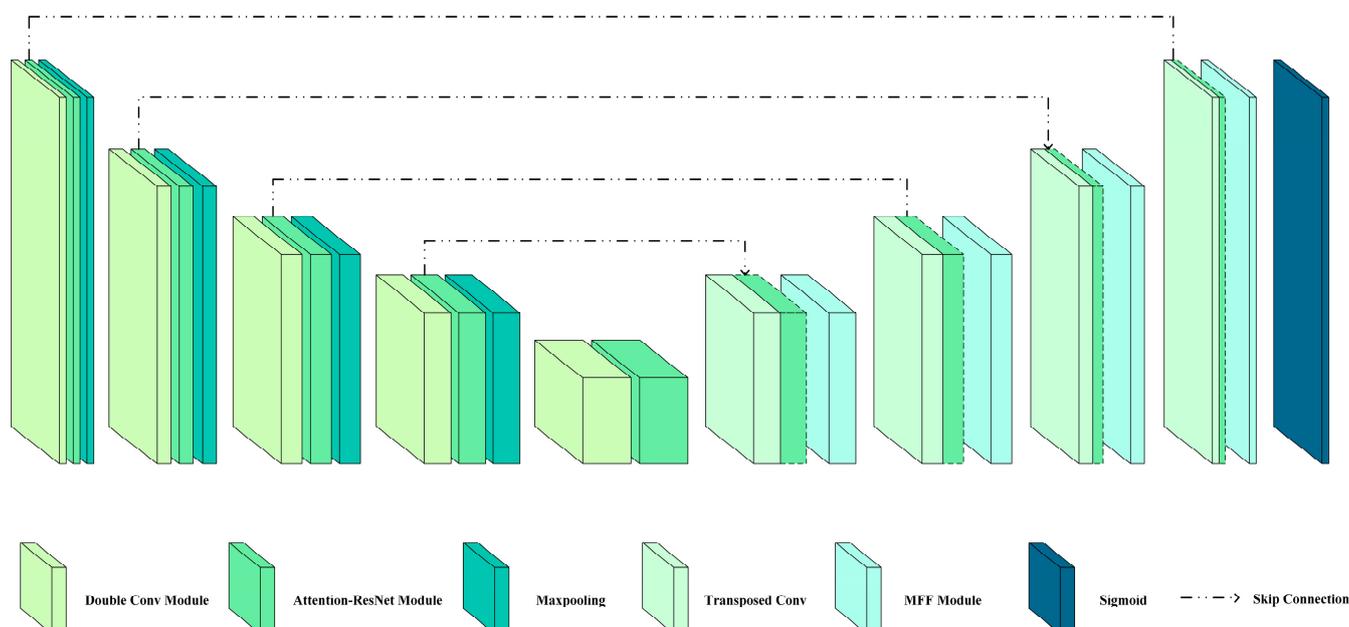
Figure 2 shows the flowchart of our study, including data preparation and the rice extraction model. Data preparation is to pre-process Sentinel-1 and Sentinel-2 data based on GEE and produce rice labels. Time-series Sentinel-1 was used to extract the dual-polarization (VH and VV) SAR features, and Sentinel-2 was used to extract the optical indices during the rice growth period. The ground-truth rice map was used to generate the rice labels. The rice extraction model was mainly based on seven datasets, and the models have been trained, validated, and tested. Finally, the accuracy evaluation and comparative analysis of the model’s performance were completed.

## 2.4. Models and Principles

### 2.4.1. R-Unet

Based on the U-Net model, a new DL model for rice extraction, R-Unet, was proposed in this study. U-Net is the most typical model in the deep semantic segmentation model, which was first applied in the field of medical imaging [47]. U-Net is composed of an encoder and a decoder. The encoder contains four times of down-sampling, including two convolutional layers and a max-pooling layer [47]. The spatial resolution of the input data is halved, and the number of channels is increased after each down-sampling. The decoder mainly includes four up-sampling layers. In the up-sampling layer, the high-level and low-level feature information of the model is concatenated in the channel dimension through skip connections, and the size of the feature map is restored by using transposed convolution [47].

In our study, the encoder and decoder of U-Net were improved. Based on the original structure of the encoder, an attention-residual module was added after the ReLU activation function of the last convolutional layer of each down-sampling layer to deepen the depth of the model and learn the characteristic information of rice [47]. The attention-residual module mainly included two convolutional layers, each of which consisted of a  $3 \times 3$  convolutional layer, a batch normalization (BN) layer, and a ReLU activation function. In order to make the model pay more attention to rice information, an attention mechanism was added. The feature map of the attention mechanism was added to the result of the ReLU activation function of the last convolutional layer of the attention-residual module to obtain the feature map of the final output of the attention-residual module. A multi-scale feature fusion (MFF) module was added to the decoder. The up-sampled transposed convolution result was concatenated with the down-sampled attention-residual module as the input of the MFF module. The MFF module replaced the two convolutional layers in the up-sampling layer of the original decoder. The MFF module integrated high-level and low-level feature information to achieve channel and spatial scale fusion [47]. Figure 3 shows the R-Unet structure.



**Figure 3.** The structure of R-Unet.

In this study, the training, validation, and testing of all DL models were performed on a laptop with an NVIDIA GeForce RTX 3060 graphics processor with 16 GB of memory and an 8-core processor with 6 GB of memory. Hyperparameters were set as follows:

(a) learning rate = 0.0001, (b) batch size = 4, (c) epochs = 50, (d) Adam optimizer. The cross-entropy loss function was chosen as the loss function, and the formula is as follows:

$$\text{Loss} = -(y \log(P) + (1 - y) \log(1 - P)) \tag{4}$$

where Loss is the loss value,  $y$  is the real label value, and  $P$  is the predicted probability of the model.

### 2.4.2. MFF Module

The down-sampling of the encoder reduces the spatial resolution of the input image, which may lead to the loss of some rice information. Dilated convolution introduces a dilated factor to increase the receptive field while maintaining the resolution of the original image, so that the feature map contains information under a larger receptive field [47]. Dilated convolution can reduce information loss to some extent [47]. Therefore, the MFF module proposed in this study can prevent the loss of rice-related feature information and improve the learning ability of the model for rice and non-rice [47].

The MFF module proposed in this paper mainly consisted of three parts. The first part was the  $1 \times 1$  convolutional layer, BN layer, and ReLU activation function, which mainly fused the input feature map on the channel scale. The second part was the  $3 \times 3$  convolutional layer, BN layer, and ReLU activation function. Finally, the third part was the dilated convolutional layer with an expansion factor of 2, which extracted rice feature information in a wide range by expanding the receptive field. Since the rice patches selected in this paper were relatively small, the detailed rice information was lost due to the large expansion factor. Finally, the feature maps obtained from the three parts were added together, and the feature information on the channel and spatial scales was fused as the final output of the MFF module. The MFF module structure is shown in Figure 4.

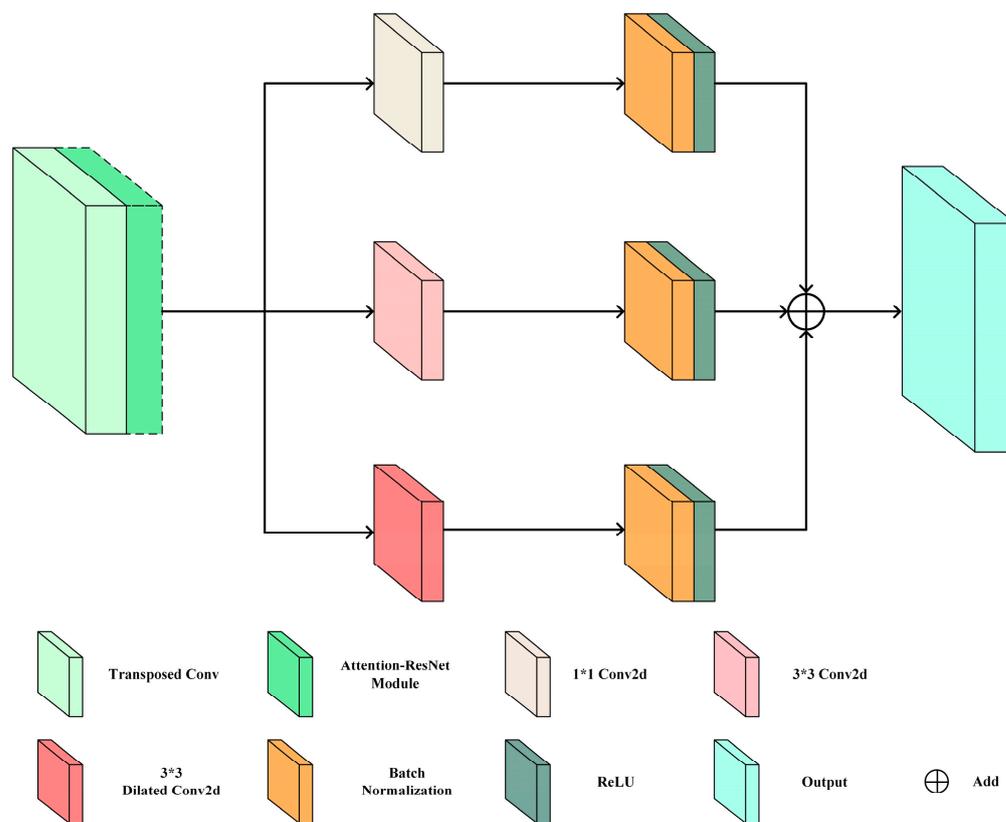
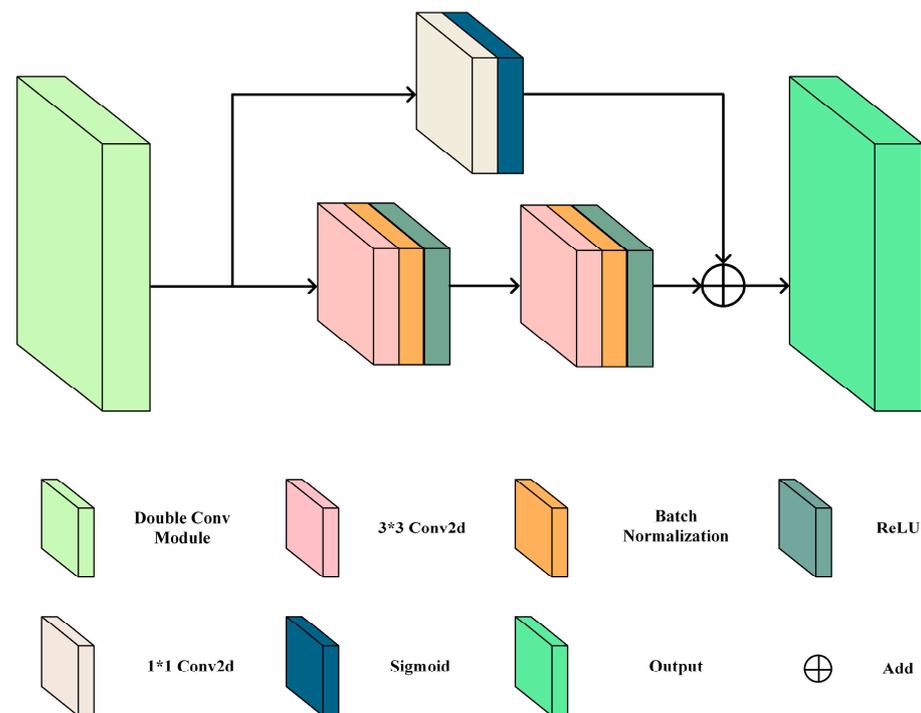


Figure 4. The structure of MFF module.

### 2.4.3. Attention-Residual Module

The down-sampling of the encoder may cause overfitting. In order to deepen the network depth of the encoder and prevent the model from overfitting, an attention-residual module was introduced in the encoder. The attention-residual module was added after the double convolution of the encoder. Since the time-series SAR data of the rice-growing period were obtained, the channel attention mechanism was introduced in the attention-residual module to extract the time-series features of rice. The feature responses of each channel of the input feature map were weighted to enhance the model's attention to rice features in different channels. The channel attention mechanism can enable the model to specifically select rice and non-rice information in different channels to optimize feature representation and model performance.

Figure 5 shows the proposed attention-residual module. The attention-residual module mainly contained two branches. The first branch included two convolutional blocks, and each convolutional layer module consisted of a  $3 \times 3$  convolutional layer, a BN layer, and a ReLU activation function. The second branch was the channel attention module, which was composed of a  $1 \times 1$  convolutional layer and a sigmoid function to enhance channel attention on the input feature map. Finally, the results of the first branch and the second branch were added as the output of the attention-residual module.



**Figure 5.** The structure of the attention-residual module.

### 2.4.4. Performance Evaluation

In this paper, overall accuracy (OA), precision, intersection and union (IOU), recall, F1-score, and Matthews correlation coefficient (MCC) were used as the main evaluation indicators for the performance evaluation of the rice extraction model. The calculation formulas for the above evaluation indicators are shown in Formulas (5)–(10):

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$\text{IOU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (7)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

$$\text{F1-score} = 2 * \frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (9)$$

$$\text{MCC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (10)$$

where TP represents the number of pixels correctly predicted as rice, TN represents the number of pixels correctly predicted as non-rice, FP represents the number of pixels incorrectly predicted as rice, and FN represents the number of pixels incorrectly predicted as non-rice.

OA measures the model's ability to correctly classify the entire dataset [10,40]. Precision refers to the accuracy of the model for rice classification [10,40]. IOU is the ratio of the intersection and union of the number of pixels correctly predicted as rice and the true pixels [40]. Recall can measure the ability of the model to identify rice [10]. The F1-score is the harmonic mean of precision and recall, combining the performance of both [8,10,40]. MCC is the correction between predicted and true results [8].

### 3. Results

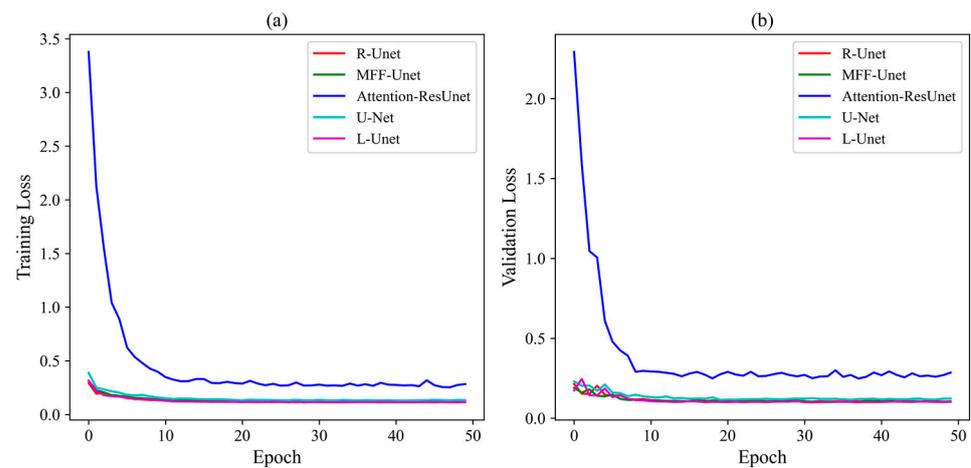
#### 3.1. Rice Extraction Model Results for Different Datasets

In order to select the best dataset suitable for the rice extraction model, four baseline models were selected for comparative analysis with R-Unet. Since R-Unet was proposed based on U-Net, U-Net was first used as one of the baseline models. Secondly, the L-Unet model contained both the residual structure and the MFF module, which was applied to the extraction of landslides by scholars. In this study, L-Unet was used as one of the baseline models. Finally, in order to compare and analyze the influence of the combined use of the MFF module and the attention-residual module on the performance of the rice extraction model, the attention-residual module and the MFF module were added to U-Net to obtain two baseline models: Attention-ResUnet and MFF-Unet.

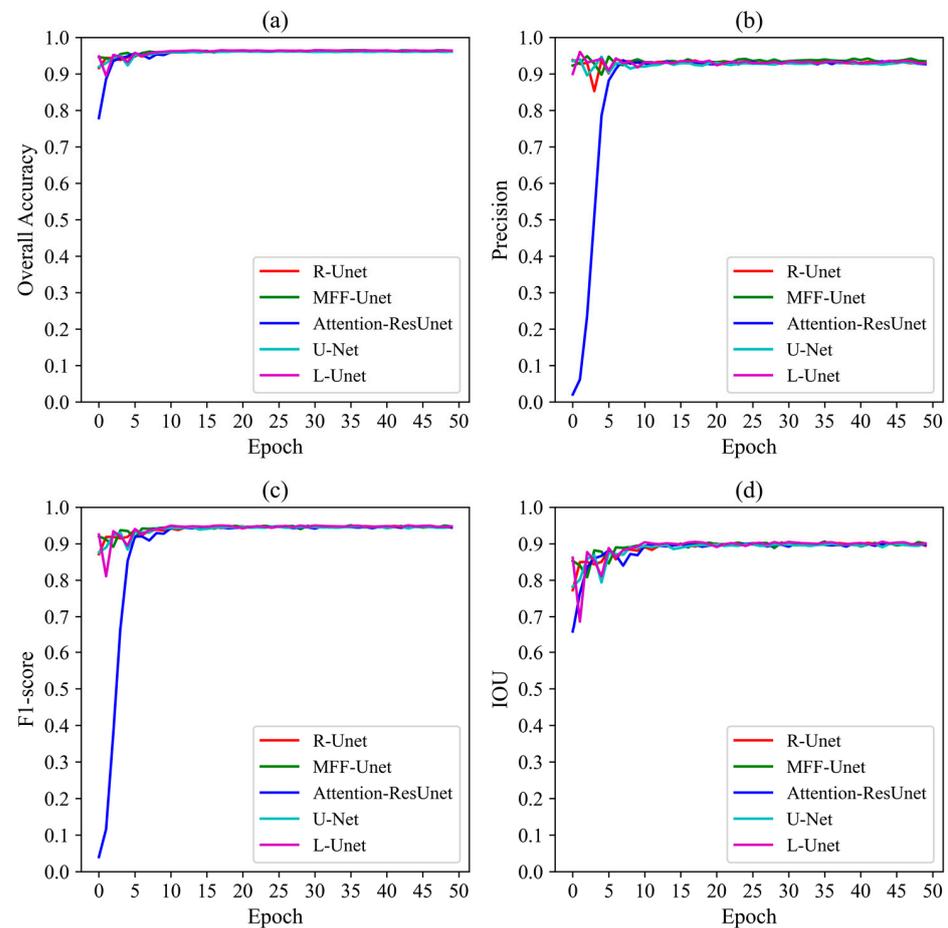
First, in study area A, five models were trained and validated on seven different datasets, and the training and validation loss values, and the OA, precision, IOU, recall, and F1-score of the validation samples were calculated. Figure 6 shows the loss values of the five baseline models in training and validation on Dataset 07. As the epoch increased, the loss value kept decreasing and eventually tended to be stable. Figure 7 shows the OA, precision, IOU, and F1-score of the five baseline models on the validation samples on Dataset 07. With the increase in epoch, the OA, precision, IOU, and F1-score tended to be stable, all above 0.9. Table 2 shows the results of the confusion matrix corresponding to the optimal accuracy of R-Unet in the validation samples. The OA, IOU, recall, and MCC of R-Unet on the validation samples were as high as 0.963, 0.902, 0.961, and 0.920, respectively.

**Table 2.** Confusion matrix results of R-Unet in validation samples.

		Predictions		
		Rice	Non-Rice	Producer Accuracy
Truth	Rice	516,645	17,602	0.967
	Non-rice	38,516	934,565	0.960
	User accuracy	0.931	0.982	
	OA	0.963		
	F1-score	0.949	MCC	0.920
	IOU	0.902	Recall	0.967



**Figure 6.** (a) Training loss values of baseline models. (b) Validation loss values of baseline models.



**Figure 7.** The OA, precision, IOU, and F1-score of the baseline models in the validation samples: where (a) is OA, (b) is precision, (c) is F1-score and (d) is IOU.

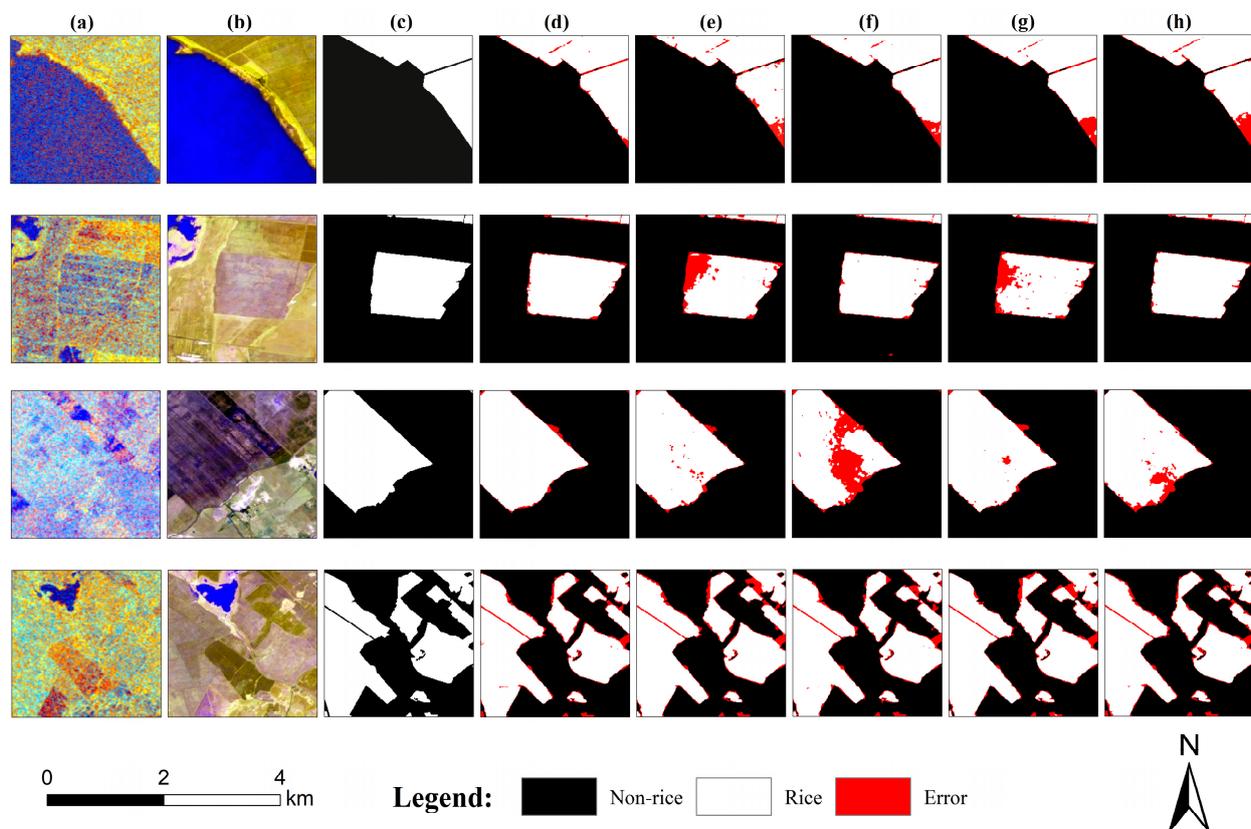
Secondly, study area B without model training was selected as the test area. The five models were tested on seven datasets, and the evaluation metrics of the test samples were calculated. In addition, the average time (Time) of five baseline models to predict rice on seven datasets was calculated in milliseconds (ms). The results showed that R-UNet had the best performance on Dataset 07. Table 3 shows the test results of the five models on Dataset 07. From Table 3, the precision, OA, IOU, F1-score, and MCC of R-UNet on Dataset 07 were 0.948, 0.952, 0.853, 0.921, and 0.888, respectively, all of which performed better than the other four baseline models. Meanwhile, R-UNet achieved a test Time of

27.22 ms per sample on Dataset 07, which was faster than the other four baseline models. In contrast, the accuracy of MFF-Unet, Attention-ResUnet, and U-Net was lower than that of R-Unet. Among the above three models, the accuracy of adding the MFF module or the attention-residual module was higher than that of the U-Net model. Table S1 shows the test results of the five baseline models in Datasets 01–06. From Table S1, the accuracy of the five baseline models on a single dataset was low. The precision of R-Unet on Dataset 01 and Dataset 02 was only 0.829 and 0.772, and the IOU was 0.733 and 0.506, respectively. However, regardless of any dataset, the R-Unet performed best in the test results of the baseline models.

**Table 3.** Test results of the baseline models on Dataset 07.

Model	Precision	IOU	Recall	F1-Score	MCC	OA	Time (ms)
R-Unet	0.948	0.853	0.895	0.921	0.888	0.952	27.22
MFF-Unet	0.939	0.789	0.832	0.882	0.837	0.931	30.93
Attention-ResUnet	0.936	0.806	0.854	0.893	0.850	0.937	37.75
U-Net	0.927	0.825	0.882	0.904	0.863	0.942	29.16
L-Unet	0.926	0.731	0.777	0.845	0.792	0.914	28.41

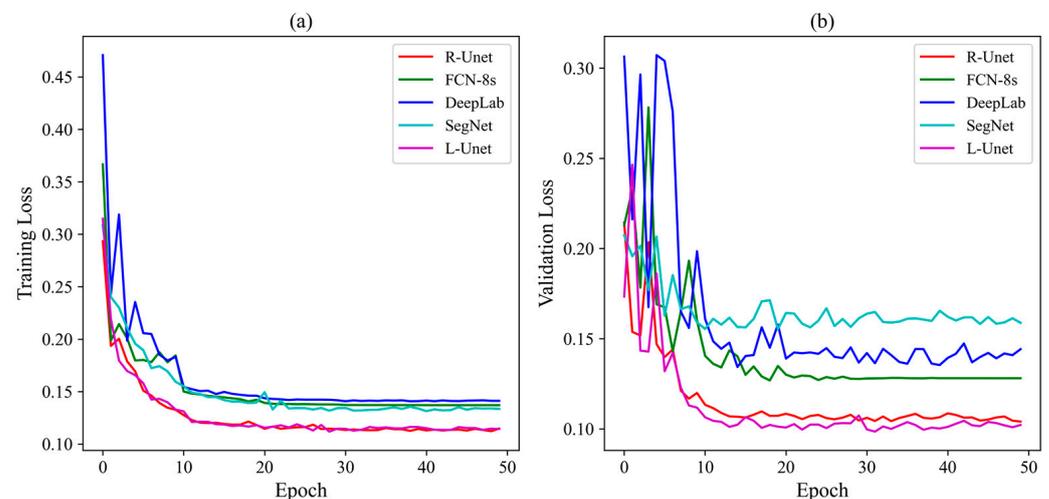
Figure 8 shows the results of rice prediction obtained by five baseline models on Dataset 07. By comparing the rice extraction results, it was found that the overall effect of R-Unet rice extraction was better, which was basically consistent with the rice ground-truth map. However, the MFF-Unet, Attention-ResUnet, and U-Net had obvious misclassifications.



**Figure 8.** Comparison of the rice extraction results of baseline models in the test samples: where (a) is the SAR images (RGB is VV, VH, and VV/VH in December 2019), (b) is the optical indices (RGB is NDVI, EVI, and LSWI, respectively), (c) is the ground-truth map, and (d–h) are the rice prediction results of R-Unet, MFF-Unet, Attention-ResUnet, U-Net, and L-Unet, respectively.

### 3.2. Rice Extraction Model Results of Different DL Models

R-UNet was compared with four classic models, including FCN-8s, DeepLab, SegNet, and L-UNet, to validate the rice extraction effect of R-UNet. Based on the results of 3.1, Dataset 07 was selected as the training, validation, and test dataset for the five models. Firstly, in study area A, five classic models were trained and validated on Dataset 07, and the loss value during training and validation, and the OA, precision, IOU, recall, and F1-score of the validation samples were calculated. Figure 9 shows the training and validation sample loss values of five classic models. As the epoch increased, the loss value continued to decrease and finally stabilized between 0.1 and 0.15. Figure 10 shows the OA, precision, IOU, and F1-score of the five classic models on the validation samples. With the increase in epoch, the scores of OA, precision, IOU, and F1-score tended to be stable, and the scores of OA, precision, and F1-score were all stable above 0.9. The IOU of R-UNet and L-UNet was stable at around 0.9, while FCN-8s, SegNet, and DeepLab were relatively low. Compared with FCN-8s, SegNet, and DeepLab, the performance of R-UNet and L-UNet in verification samples was generally better than the other three classic models.



**Figure 9.** (a) Training loss values of classic models. (b) Validation loss values of classic models.

Secondly, study area B, a test area without model training, was selected to test five classic models and calculate evaluation indicators. Table 4 shows the results of five classic models on test samples. From Table 4, the accuracy of R-UNet in test samples was better than in other classic models. The Time of DeepLab was the fastest among the five models, but the overall performance of DeepLab was low, with a precision of 0.896. In contrast, R-UNet has higher accuracy and a relatively faster Time. In addition, the IOUs of L-UNet, FCN-8s, and DeepLab were only 0.731, 0.707, and 0.719. Table 5 shows the confusion matrix of R-UNet in the test samples. The rice and non-rice user accuracies of R-UNet were as high as 0.948 and 0.954, respectively. The F1-score was as high as 0.921, and the MCC, IOU, and recall were all above 0.85.

**Table 4.** Test results of the classic models on Dataset 07.

Model	Precision	IOU	Recall	F1-Score	MCC	OA	Time (ms)
R-UNet	0.948	0.853	0.895	0.921	0.888	0.952	27.22
SegNet	0.928	0.755	0.802	0.861	0.809	0.920	27.58
L-UNet	0.926	0.731	0.777	0.845	0.792	0.914	28.41
FCN-8s	0.923	0.707	0.751	0.828	0.770	0.903	40.32
DeepLab	0.896	0.719	0.784	0.836	0.773	0.905	26.25

**Table 5.** Confusion matrix results of R-UNet in test samples.

		Predictions		
		Rice	Non-Rice	Producer Accuracy
Truth	Rice	2,717,954	318,268	0.900
	Non-rice	149,257	6,644,921	0.978
User accuracy		0.948	0.954	

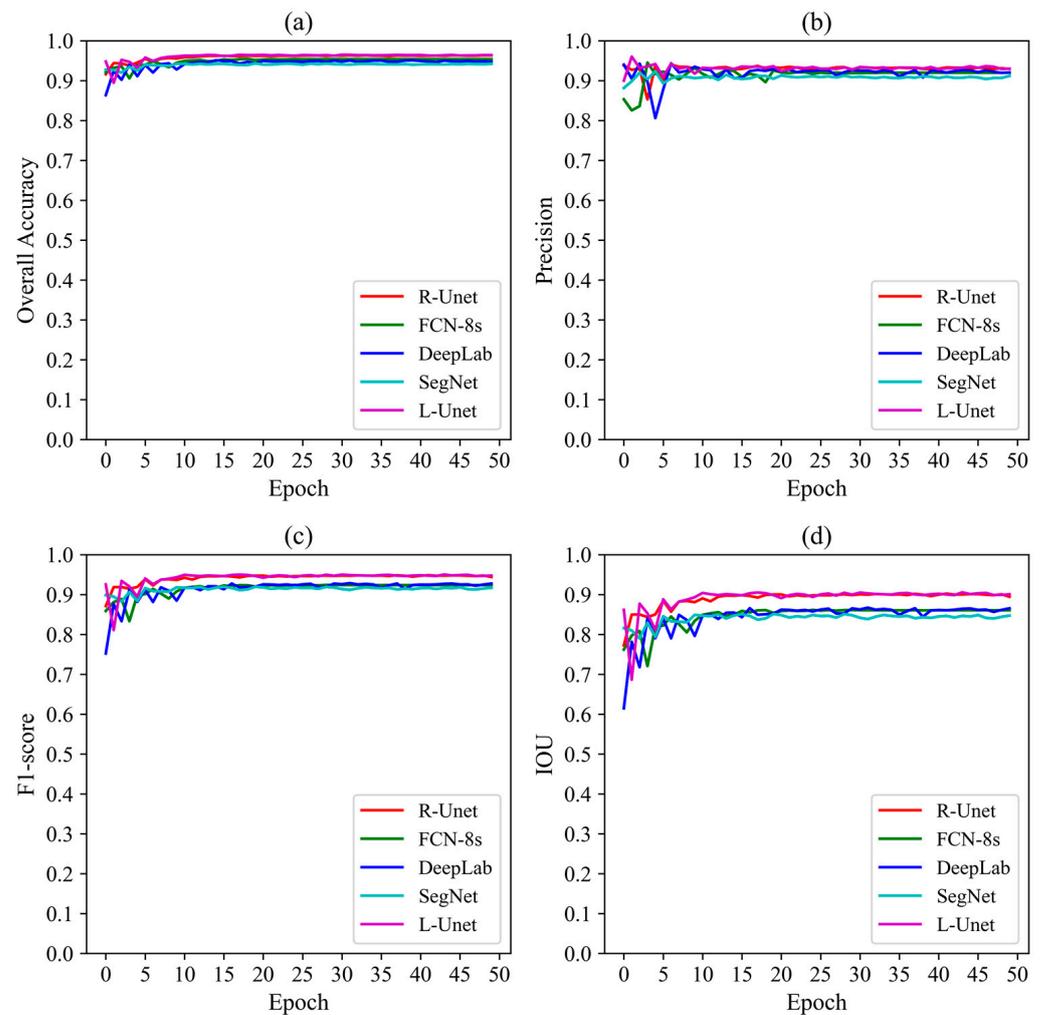
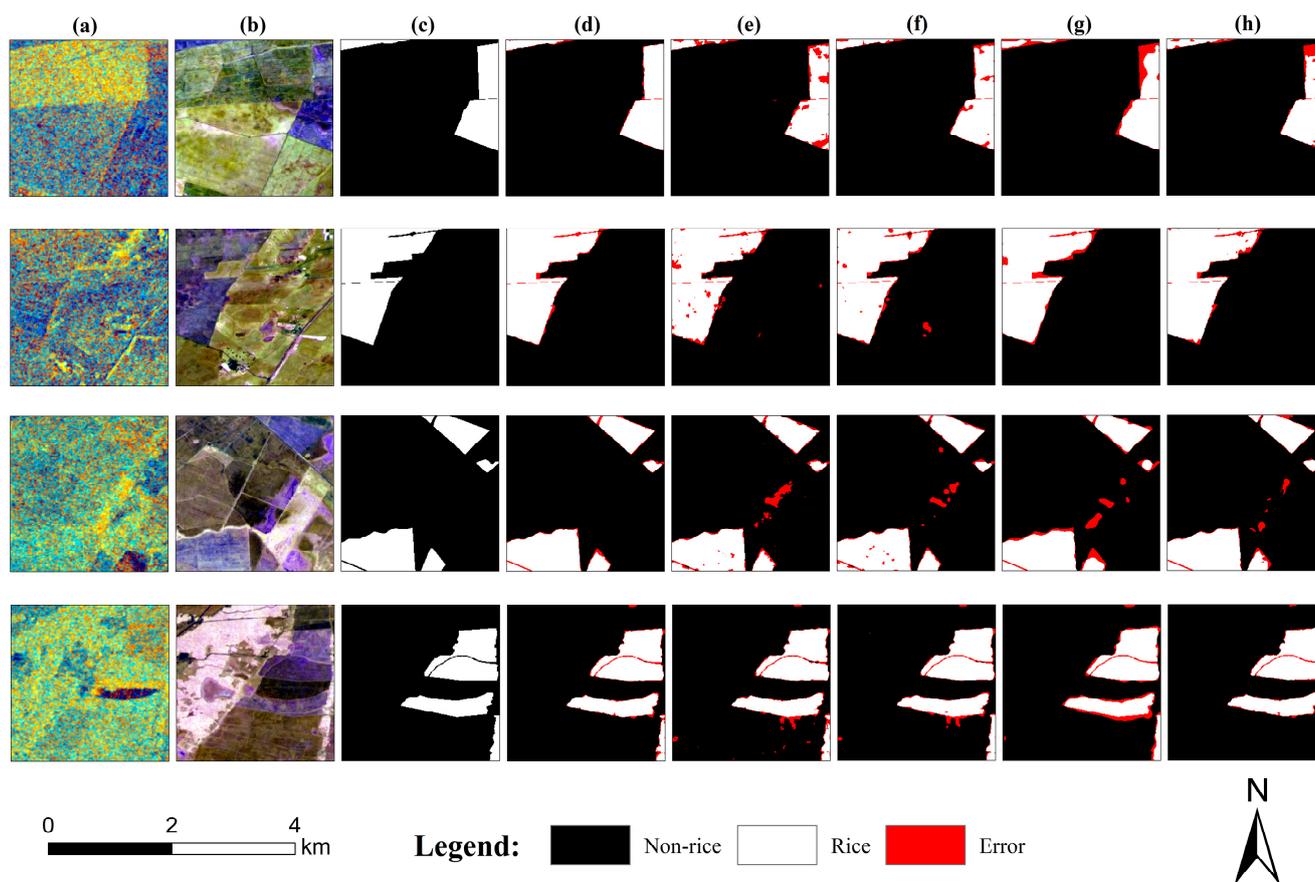
**Figure 10.** The OA, precision, IOU, and F1-score of the classic models in the validation samples: where (a) is OA, (b) is precision, (c) is F1-score and (d) is IOU.

Figure 11 shows the rice prediction results obtained by five classic models on the test samples. Comparing the rice extraction results, it was found that the overall effect of R-UNet rice extraction was better, and the prediction results were basically consistent with the actual rice map. The results of the second group in Figure 11 show that there are more misclassifications in FCN-8s and DeepLab. In addition, there were many misclassifications in FCN-8s, DeepLab, SegNet, and L-UNet. In contrast, the rice prediction results of R-UNet were better, and the distribution of misclassification was less.



**Figure 11.** Comparison of the rice extraction results of classic models in the test samples: where (a) is the SAR images (RGB is VV, VH, and VV/VH in December 2019), (b) is the optical indices (RGB is NDVI, EVI, and LSWI, respectively), (c) is the ground-truth map, and (d–h) are the rice prediction results of R-Unet, FCN-8s, DeepLab, SegNet, and L-Net, respectively.

#### 4. Discussion

##### 4.1. Optimal Dataset for Rice Extraction

With the successive launches of optical and radar satellites, optical and radar images have been widely used in agricultural RS fields such as rice extraction [8–13,24–27]. However, how to choose the optimal dataset for the rice extraction model is a major problem in the field of rice extraction research. Rice extraction using only single optical or radar images has been extensively studied [8,9,21–23,40]. In this study, to solve the above problem, seven datasets were designed to train, validate, and test five baseline models.

For Datasets 01–04 with a single data source and Datasets 05–07 with a combination of SAR and optical features, the model accuracy of the latter was better than that of the former. From Table 3 and Table S1, the precision of R-Unet on Dataset 07 and Dataset 03 was 0.948 and 0.939, respectively, and the precision on the other five datasets was all less than 0.9. For Dataset 03 and Dataset 07, the precision of the latter is 0.9% higher than that of the former, and the Time of the latter is 2.17 ms faster than the former. Except for R-Unet, the overall accuracy of the other four baseline models on Dataset 07 was better. It shows that, in the case of consistent models, the accuracy of the rice extraction model fused with SAR and optical features is better than that of a single data source.

For the single-polarization (VH or VV) and dual-polarization (VH and VV) datasets, the latter model outperformed the former in overall accuracy. It can be seen from Table 3 and Table S1 that the overall accuracy of the dual-polarization datasets (Dataset 07 and Dataset 03) was higher, and the overall accuracy of the single-polarization datasets was lower. In the single-polarization datasets, the overall accuracy of the VH polarization datasets

(Dataset 05 and Dataset 01) performed better than that of the VV polarization datasets (Dataset 06 and Dataset 02). The IOU of the R-Unet single-polarization VV dataset was only between 0.474 and 0.575. In short, the model accuracy of dual-polarization datasets performed better than that of single-polarization datasets. In the single polarization dataset, the model accuracy of VH polarization was better than that of VV polarization, which was consistent with the research results of [40].

In summary, the dual-polarization datasets that fused optical and SAR features had the highest overall accuracy. The result showed that the combination of optical and SAR data can improve the accuracy of the rice extraction model, which solved Q2. Meanwhile, the sentinel image used in this study was freely available from the European Space Agency, which solved Q1 and provided the possibility for timely and accurate rice extraction.

#### 4.2. Optimal Model for Rice Extraction

##### 4.2.1. Discussion of Results from Rice Extraction Baseline Models

The R-Unet proposed in this study was based on the U-Net, adding the attention-residual module and the MFF module in the encoder and decoder, respectively, to extract rice in the study area [47]. In order to evaluate the performance of R-Unet in rice extraction, based on U-Net, the attention-residual module and the MFF module were added to U-Net, respectively, and two baseline models of Attention-ResUnet and MFF-Unet were obtained. Meanwhile, the L-Unet applied to landslide extraction, which also contained MFF and residual modules, was also one of the baseline models [47]. From Table 3, it can be seen that the prediction Time of R-Unet was the shortest among all baseline models, at 27.22 ms. Table 6 was obtained by calculating the difference between the evaluation indicators of R-Unet and the other four baseline models. From Table 6, each indicator of R-Unet was better than the other four baseline models. The IOU, recall, and MCC of MFF-Unet were 6.4%, 6.3%, and 5.1% lower than R-Unet, respectively. The IOU, recall, and MCC of R-Unet were 4.7%, 4.1%, and 5.1% higher than those of Attention-ResUnet, respectively. Compared with U-Net, the precision, IOU, and MCC of R-Unet were increased by 2.1%, 2.8%, and 2.5%, respectively. Therefore, the attention-residual module and MFF module can improve the model performance of rice extraction to a certain extent, but a single attention-residual module or MFF module still cannot better learn the SAR and optical features of rice. The combination of the above two modules can efficiently learn rice features [47].

**Table 6.** Differences in evaluation indicators between R-Unet and baseline models (%).

Model	Precision	IOU	Recall	F1-Score	MCC	OA
MFF-Unet	0.9	6.4	6.3	3.9	5.1	2.1
Attention-ResUnet	1.2	4.7	4.1	2.8	3.8	1.5
U-Net	2.1	2.8	1.3	1.7	2.5	1.0
L-Unet	2.2	12.2	11.8	7.6	9.6	3.8

In addition, for R-Unet and L-Unet, which contained both the MFF module and the attention-residual module, the IOU, recall, and MCC of the former were 12.2%, 11.8%, and 9.6% higher than the latter, respectively. L-Unet contained both the attention-residual module and the MFF module, but the MFF module was only in the last layer of the encoder of L-Unet. The residual module of L-Unet was introduced into both the encoder and decoder. Although the network structure depth of the model was deepened, it may cause over-learning of the model, resulting in poor performance. Therefore, the position of the MFF module and the attention-residual module in U-Net can also affect the performance of the model. In the proposed R-Unet, the attention-residual module was only introduced into the encoder of the model. The attention-residual module can ensure the depth of the network, while the attention mechanism can pay more attention to the rice feature information and prevent information loss to a certain extent [10,38–40]. The MFF module was introduced into the model's decoder to perform multi-scale feature fusion on the high-level and low-level feature information in the up-sampling process [47,48]. While

restoring the size of the image, the output feature map contained both the channel and spatial information of the original feature map. In conclusion, the proposed R-Net had the highest rice extraction performance among the baseline models and solved Q3 to a certain extent.

#### 4.2.2. Discussion of Results from Rice Extraction Classic Models

With the rapid development of DL algorithms, deep semantic segmentation models have been widely used in agricultural RS research fields such as rice extraction [33–37]. So far, scholars have applied classic models such as FCN-8s, SegNet, and U-Net to rice extraction [38–46]. But the above model is not a proprietary model for rice extraction. In this study, the proposed R-Net was compared with FCN-8s, SegNet, DeepLab, U-Net, and L-Net. It can be seen from Tables 3 and 5 that the accuracy and effect of R-Net in the test samples outperformed the other five models, and the Time of R-Net was faster, second only to DeepLab. Meanwhile, the Time of DeepLab was the shortest, but the accuracy of rice extraction was lower than that of R-Net. Table 7 calculates the difference in evaluation metrics between R-Net and the other five models. The precision, IOU, F1-score, and MCC of R-Net were 2.1%, 2.8%, 1.7%, and 2.5% higher than U-Net, respectively. For the DeepLab model, the precision, IOU, F1-score, and MCC of the R-Net model were 5.2%, 13.4%, 8.5%, and 11.5% higher, respectively. For FCN-8s, SegNet, and L-Net, the IOU of R-Net was around 10% higher than the other three models. In summary, the overall accuracy of R-Net outperformed the other five models. It can also be seen from Figures 8 and 11 that the rice prediction results of R-Net were closer to the ground-truth rice map, while other classic models have some misclassifications.

**Table 7.** Differences in evaluation indicators between R-Net and classic models (%).

Model	Precision	IOU	Recall	F1-Score	MCC	OA
SegNet	2.0	9.8	9.3	6.0	7.9	3.2
U-Net	2.1	2.8	1.3	1.7	2.4	1.0
L-Net	2.2	12.2	11.8	7.6	9.6	3.8
FCN-8s	2.5	14.6	14.4	9.3	11.8	4.9
DeepLab	5.2	13.4	11.1	8.5	11.5	4.7

In general, the R-Net proposed in this study can combine open-access Sentinel satellite data (Sentinel-1 and Sentinel-2) to timely and accurately extract rice planting areas in the study area. The rice extraction results of R-Net can provide basic information for the country or the government to implement decisions on agricultural management, such as guiding national agricultural production and ensuring national food security.

#### 4.3. Limitations and Prospects

In this study, the rice extraction research was only carried out in Rio Grande do Sul, Brazil. Two study areas were selected as the validation and testing areas, but the rice planting pattern, growth period, and growing environment are basically consistent. Differences in rice planting patterns, rice varieties, and growing environments will affect the SAR characteristics ( $\sigma_0$ ) of rice to a certain extent. At the same time, it may affect the performance of the model when learning rice features, which may cause some differences in the results of rice extraction. Further trials in other countries were not carried out in this study, and future research will be carried out in other countries with different rice growth patterns from Rio Grande do Sul. In future research, while ensuring the performance of the model, we will try our best to improve its universality and strive to develop a general model for rice extraction research.

## 5. Conclusions

Based on the GEE platform, the optical indices and time-series dual-polarization  $\sigma_0$  data during the rice growth period in Rio Grande do Sul, Brazil, were obtained, and

seven different datasets were produced. Based on U-Net, a model R-Net focusing on rice extraction was proposed, which included both the attention-residual module and the MFF module. The attention-residual module was designed to deepen the network depth of the R-Net encoder, focusing on rice feature information while preventing information loss. The MFF module was used to perform multi-scale feature fusion processing on the high-level and low-level feature information of the decoder. In order to study the impact of different datasets on model performance, R-Net was compared with four baseline models. Meanwhile, in order to verify the performance of R-Net in rice extraction research, it was compared with four classic models. Ultimately, the following conclusions were drawn:

1. The model accuracy of the single-polarization SAR dataset was the worst, while the model accuracy of the dual-polarization SAR dataset was better. The precision of R-Net on Dataset 03 and Dataset 07 was 0.939 and 0.948, respectively, while the precision on Dataset 01 and Dataset 02 was only 0.829 and 0.772.
2. The combined dataset of SAR features and optical indices (Dataset 07) had higher accuracy than single SAR features (Dataset 03) or single optical indices (Dataset 04), and the model performance was better. In Dataset 07, the precision, F1-score, and MCC of R-Net in the test sample were 0.948, 0.921, and 0.888, respectively.
3. Compared with classic models such as FCN-8s, SegNet, U-Net, etc., R-Net had the highest test result accuracy on the best dataset, and the rice extraction effect was the best. The precision, IOU, MCC, and F1-score of R-Net rice extraction increased by 5.2%, 14.6%, 11.8%, and 9.3%, respectively.

In summary, the R-Net proposed in this study can extract the rice planting area timely and accurately, and the research results can provide crucial information for governments to implement decision-making about agricultural management.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs15164021/s1>, Table S1: Test results of the baseline models on Dataset 01–06.

**Author Contributions:** Conceptualization, T.F.; methodology, T.F.; software, T.F.; validation, T.F.; formal analysis, T.F.; investigation, T.F.; resources, T.F.; data curation, T.F.; writing—original draft preparation, T.F.; writing—review and editing, T.F., J.G. and S.T.; visualization, T.F.; supervision, T.F.; project administration, S.T.; funding acquisition, S.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by 2021 Guiding special “Double First-Class” Disciplines (Geology), China University of Geosciences (Beijing), China (project no. 64022102501).

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** We would like to thank Google Earth Engine for providing the cloud processing platform, and the Agricultural Information Portal for providing the rice map. Meanwhile, we greatly appreciate editors and reviewers for their positive comments and professional suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Laborte, A.; Gutierrez, M.; Balanza, J.; Saito, K.; Zwart, S.J.; Boschetti, M.; Murty, M.V.R.; Villano, L.; Aunario, J.; Reinke, R.; et al. RiceAtlas, A Spatial Database of Global Rice Calendars and Production. *Sci. Data* **2017**, *4*, 170074. [[CrossRef](#)]
2. Van der Velde, M.; Baruth, B.; Bussay, A.; Ceglar, A.; Garcia Condado, S.; Karetos, S.; Lecerf, R.; Lopez, R.; Maiorano, A.; Nisini, L.; et al. In-season Performance of European Union Wheat Forecasts during Extreme Impacts. *Sci. Rep.* **2018**, *8*, 15420. [[CrossRef](#)]
3. Jiang, J.; Zhang, H.; Ge, J.; Sun, C.; Xu, L.; Wang, C. Cropland Data Extraction in Mekong Delta Based on Time Series Sentinel-1 Dual-Polarized Data. *Remote Sens.* **2023**, *15*, 3050. [[CrossRef](#)]
4. Xu, L.; Zhang, H.; Wang, C.; Zhang, B.; Liu, M. Crop Classification Based on Temporal Information Using Sentinel-1 SAR Time-Series Data. *Remote Sens.* **2019**, *11*, 53. [[CrossRef](#)]
5. Arumugam, P.; Chemura, A.; Schauburger, B.; Gornott, C. Remote Sensing Based Yield Estimation of Rice (*Oryza sativa* L.) Using Gradient Boosted Regression in India. *Remote Sens.* **2021**, *13*, 2379. [[CrossRef](#)]

6. Islam, M.D.; Di, L.; Qamer, F.M.; Shrestha, S.; Guo, L.; Lin, L.; Mayer, T.J.; Phalke, A.R. Rapid Rice Yield Estimation Using Integrated Remote Sensing and Meteorological Data and Machine Learning. *Remote Sens.* **2023**, *15*, 2374. [[CrossRef](#)]
7. Fu, T.; Tian, S.; Zhan, Q. Phenological Analysis and Yield Estimation of Rice Based on Multi-spectral and SAR Data in Maha Sarakham, Thailand. *J. Spat. Sci.* **2023**, *68*, 2184428. [[CrossRef](#)]
8. Xia, L.; Zhao, F.; Chen, J.; Yu, L. A Full Resolution Deep Learning Network for Paddy Rice Mapping Using Landsat Data. *ISPRS J. Photogramm. Remote Sens.* **2022**, *194*, 91–107. [[CrossRef](#)]
9. Wayan Nuarsa, I.; Nishio, F.; Hongo, C. Spectral Characteristics and Mapping of Rice Plants Using Multi-Temporal Landsat Data. *J. Agric. Sci.* **2011**, *3*, 54–67. [[CrossRef](#)]
10. Onojeghuo, A.O.; Miao, Y.; Blackburn, G.A. Deep ResU-Net Convolutional Neural Networks Segmentation for Smallholder Paddy Rice Mapping Using Sentinel 1 SAR and Sentinel 2 Optical Imagery. *Remote Sens.* **2023**, *15*, 1517. [[CrossRef](#)]
11. 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. [[CrossRef](#)]
12. Du, M.; Huang, J.; Wei, P.; Yang, L.; Chai, D.; Peng, D.; Sha, J.; Sun, W.; Huang, R. Dynamic Mapping of Paddy Rice Using Multi-Temporal Landsat Data Based on a Deep Semantic Segmentation Model. *Agronomy* **2022**, *12*, 1583. [[CrossRef](#)]
13. Shan, J.; Qiu, L.; Tian, M.; Wang, J.; Wang, Z.; Huang, X. Study on Extraction Methods of Paddy Rice Area Based on GF-6 Satellite Image. In Proceedings of the 2021 9th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Shenzhen, China, 26–29 July 2021. [[CrossRef](#)]
14. Nazir, A.; Ullah, S.; Ahmad Saqib, Z.; Abbas, A.; Ali, A.; Shahid Iqbal, M.; Hussain, K.; Shakir, M.; Shah, M.; Usman Butt, M. Estimation and Forecasting of Rice Yield Using Phenology-Based Algorithm and Linear Regression Model on Sentinel-II Satellite Data. *Agriculture* **2021**, *11*, 1026. [[CrossRef](#)]
15. Liaqat, M.U.; Cheema, M.J.M.; Huang, W.; Mahmood, T.; Zaman, M.; Khan, M.M. Evaluation of MODIS and Landsat Multiband Vegetation Indices Used for Wheat Yield Estimation in Irrigated Indus Basin. *Comput. Electron. Agric.* **2017**, *138*, 39–47. [[CrossRef](#)]
16. Chen, N.; Yu, L.; Zhang, X.; Shen, Y.; Zeng, L.; Hu, Q.; Niyogi, D. Mapping Paddy Rice Fields by Combining Multi-Temporal Vegetation Index and Synthetic Aperture Radar Remote Sensing Data Using Google Earth Engine Machine Learning Platform. *Remote Sens.* **2020**, *12*, 2992. [[CrossRef](#)]
17. Teluguntla, P.; Ryu, D.; George, B.; Walker, J.P.; Malano, H.M. Mapping Flooded Rice Paddies Using Time Series of MODIS Imagery in the Krishna River Basin, India. *Remote Sens.* **2015**, *7*, 8858–8882. [[CrossRef](#)]
18. Zhong, L.; Hu, L.; Zhou, H. Deep Learning Based Multi-temporal Crop Classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [[CrossRef](#)]
19. Jiang, T.; Liu, X.N.; Wu, L. Method for Mapping Rice Fields in Complex Landscape Areas Based on Pre-trained Convolutional Neural Network from HJ-1 A/b Data. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 418. [[CrossRef](#)]
20. Yang, L.; Huang, R.; Zhang, J.; Huang, J.; Wang, L.; Dong, J.; Shao, J. Inter-Continental Transfer of Pre-Trained Deep Learning Rice Mapping Model and Its Generalization Ability. *Remote Sens.* **2023**, *15*, 2443. [[CrossRef](#)]
21. Lin, Z.; Zhong, R.; Xiong, X.; Guo, C.; Xu, J.; Zhu, Y.; Xu, J.; Ying, Y.; Ting, K.C.; Huang, J.; et al. Large-Scale Rice Mapping Using Multi-Task Spatiotemporal Deep Learning and Sentinel-1 SAR Time Series. *Remote Sens.* **2022**, *14*, 699. [[CrossRef](#)]
22. Xu, L.; Zhang, H.; Wang, C.; Wei, S.; Zhang, B.; Wu, F.; Tang, Y. Paddy Rice Mapping in Thailand Using Time-Series Sentinel-1 Data and Deep Learning Model. *Remote Sens.* **2021**, *13*, 3994. [[CrossRef](#)]
23. Yang, H.; Pan, B.; Li, N.; Wang, W.; Zhang, J.; Zhang, X. A Systematic Method for Spatio-temporal Phenology Estimation of Paddy Rice Using Time Series Sentinel-1 Images. *Remote Sens. Environ.* **2021**, *259*, 112394. [[CrossRef](#)]
24. Inoue, S.; Ito, A.; Yonezawa, C. Mapping Paddy Fields in Japan by Using a Sentinel-1 SAR Time Series Supplemented by Sentinel-2 Images on Google Earth Engine. *Remote Sens.* **2020**, *12*, 1622. [[CrossRef](#)]
25. Clauss, K.; Ottinger, M.; Leinenkugel, P.; Kuenzer, C. Estimating Rice Production in the Mekong Delta, Vietnam, Utilizing Time Series of Sentinel-1 SAR Data. *Int. J. Appl. Earth Obs.* **2018**, *73*, 574–585. [[CrossRef](#)]
26. Phung, H.; Nguyen, L. Crop Monitoring in the Mekong Delta, Vietnam Using Multi-Temporal Sentinel-1 Data with C-Band. *Lect. Notes Civ. Eng.* **2020**, *80*, 979–986. [[CrossRef](#)]
27. Singha, M.; Dong, J.; Zhang, G.; Xiao, X. High Resolution Paddy Rice Maps in Cloud-prone Bangladesh and Northeast India Using Sentinel-1 Data. *Sci. Data* **2019**, *6*, 26. [[CrossRef](#)] [[PubMed](#)]
28. Lasko, K.; Vadrevu, K.P.; Tran, V.T.; Justice, C. Mapping Double and Single Crop Paddy Rice with Sentinel-1A at Varying Spatial Scales and Polarizations in Hanoi, Vietnam. *IEEE J. Select. Top. Appl. Earth Observat. Remote Sens.* **2018**, *11*, 498–512. [[CrossRef](#)] [[PubMed](#)]
29. Mansaray, L.R.; Yang, L.; Kabba, V.T.S.; Kanu, A.S.; Huang, J.; Wang, F. Optimising Rice Mapping in Cloud-prone Environments by Combining Quad-source Optical with Sentinel-1A Microwave Satellite Imagery. *GIScience Remote Sens.* **2019**, *56*, 1333–1354. [[CrossRef](#)]
30. Dai, X.; Chen, S.; Jia, K.; Jiang, H.; Sun, Y.; Li, D.; Zheng, Q.; Huang, J. A Decision-Tree Approach to Identifying Paddy Rice Lodging with Multiple Pieces of Polarization Information Derived from Sentinel-1. *Remote Sens.* **2023**, *15*, 240. [[CrossRef](#)]
31. He, Z.; Li, S.; Deng, Y.; Zhai, P.; Hu, Y. Rice Paddy Fields Identification Based on Backscatter Features of Quad-Pol RADARSAT-2 Data and Simple Decision Tree Method. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 6765–6768. [[CrossRef](#)]

32. Peng, W.; Li, S.; He, Z.; Ning, S.; Liu, Y.; Su, Z. Random Forest Classification of Rice Planting Area Using Multi-Temporal Polarimetric Radarsat-2 Data. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 2411–2414. [[CrossRef](#)]
33. 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. [[CrossRef](#)]
34. Parelus, E.J. A Review of Deep-Learning Methods for Change Detection in Multispectral Remote Sensing Images. *Remote Sens.* **2023**, *15*, 2092. [[CrossRef](#)]
35. Miller, J.; Nair, U.; Ramachandran, R.; Maskey, M. Detection of Transverse Cirrus Bands in Satellite Imagery Using Deep Learning. *Comput. Geosci.* **2018**, *118*, 79–85. [[CrossRef](#)]
36. Palafox, L.F.; Hamilton, C.W.; Scheidt, S.P.; Alvarez, A.M. Automated Detection of Geological Landforms on Mars Using Convolutional Neural Networks. *Comput. Geosci.* **2017**, *101*, 48–56. [[CrossRef](#)]
37. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. An Object-based Convolutional Neural Network (OCNN) for Urban Land Use Classification. *Remote Sens. Environ.* **2018**, *216*, 57–70. [[CrossRef](#)]
38. Ning, S.; Li, S.; He, Z.; Zhai, P. Extraction of Rice-planted Area Based on MobileUnet Model and Radarsat-2 Data. In Proceedings of the 2019 SAR in Big Data Era (BIGSARDATA), Beijing, China, 5–6 August 2019; pp. 1–4. [[CrossRef](#)]
39. Wei, P.; Chai, D.; Huang, R.; Peng, D.; Lin, T.; Sha, J.; Sun, W.; Huang, J. Rice Mapping Based on Sentinel-1 Images Using the Coupling of Prior Knowledge and Deep Semantic Segmentation Network: A case Study in Northeast China from 2019 to 2021. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102948. [[CrossRef](#)]
40. De Bem, P.P.; De Carvalho Júnior, O.; De Carvalho, O.; Gomes, R.; Guimarães, R.; Pimentel, C. Irrigated Rice Crop Identification in Southern Brazil Using Convolutional Neural Networks and Sentinel-1 Time Series. *Remote Sens. Appl. Soc. Environ.* **2021**, *24*, 100627. [[CrossRef](#)]
41. Zhao, F.; Xia, L.; Kylling, A.; Li, R.Q.; Shang, H.; Xu, M. Detection Flying Aircraft from Landsat 8 OLI Data. *ISPRS-J. Photogramm. Remote Sens.* **2018**, *1*, 176–184. [[CrossRef](#)]
42. Wei, P.; Huang, R.; Lin, T.; Huang, J. Rice Mapping in Training Sample Shortage Regions Using a Deep Semantic Segmentation Model Trained on Pseudo-Labels. *Remote Sens.* **2022**, *14*, 328. [[CrossRef](#)]
43. Sun, Z.; Di, L.; Fang, H.; Burgess, A. Deep Learning Classification for Crop Types in North Dakota. *IEEE J. Select. Top. Appl. Earth Observat. Remote Sens.* **2020**, *14*, 2200–2213. [[CrossRef](#)]
44. Zhu, S.; Li, S.; Yang, Z. Research on the Distribution Map of Weeds in Rice Field Based on SegNet. 3D Imaging—Multidimensional Signal Processing and Deep Learning. *Smart Innov. Syst. Technol.* **2022**, *298*, 91–99. [[CrossRef](#)]
45. Wang, M.; Wang, J.; Cui, Y.; Liu, J.; Chen, L. Agricultural Field Boundary Delineation with Satellite Image Segmentation for High-Resolution Crop Mapping: A Case Study of Rice Paddy. *Agronomy* **2022**, *12*, 2342. [[CrossRef](#)]
46. 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. [[CrossRef](#)]
47. Dong, Z.; An, S.; Zhang, J.; Yu, J.; Li, J.; Xu, D. L-Unet: A Landslide Extraction Model Using Multi-Scale Feature Fusion and Attention Mechanism. *Remote Sens.* **2022**, *14*, 2552. [[CrossRef](#)]
48. Cai, B.; Wang, M.; Chen, Y.; Hu, Y.; Liu, M. MFF-Net: A Multi-feature Fusion Network for Community Detection in Complex Network. *Knowl.-Based Syst.* **2022**, *252*, 109408. [[CrossRef](#)]

**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.