

Article

An Effective Deep Learning Model for Monitoring Mangroves: A Case Study of the Indus Delta

Chen Xu ^{1,2} , Juanle Wang ^{2,3,4,*} , Yu Sang ¹, Kai Li ^{2,5}, Jingxuan Liu ^{1,2} and Gang Yang ¹

- ¹ School of Marine Technology and Geomatics, Jiangsu Ocean University, Lianyungang 222005, China
² State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
³ Jiangsu Centre for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China
⁴ China-Pakistan Earth Science Research Centre, Islamabad 45320, Pakistan
⁵ College of Geoscience and Surveying Engineering, China University of Mining & Technology (Beijing), Beijing 100083, China
* Correspondence: wangjl@igsnr.ac.cn

Abstract: Rapid and accurate identification of mangroves using remote sensing images is of great significance for assisting ecological conservation efforts in coastal zones. With the rapid development of artificial intelligence, deep learning methods have been successfully applied to a variety of fields. However, few studies have applied deep learning methods to the automatic detection of mangroves and few scholars have used medium-resolution Landsat images for large-scale mangrove identification. In this study, cloud-free Landsat 8 OLI imagery of the Indus Delta was acquired using the GEE platform, and NDVI and land use data were used to produce integrated labels to reduce the complexity and subjectivity of manually labeled samples. We proposed the use of MSNet, a semantic segmentation model fusing multiple-scale features, for mangrove extraction in the Indus Delta, and compared the performance of the MSNet model with three other semantic segmentation models, FCN-8s, SegNet, and U-Net. The overall performance ranking of the deep learning methods was MSNet > U-Net > SegNet > FCN-8s. The parallel-structured MSNet model was easy to train, had the fewest parameters and the highest validation accuracy, and provided the best results for the extraction of mangrove pixels with weak features. The MSNet model not only maintains the high-resolution features of the image and fully learns the pixels with weak features during the training process but also fuses the multiple-scale underlying features at different scales to enhance the semantic information and improve the accuracy of feature recognition and segmentation localization. Finally, the areas covered by mangroves in the Indus Delta in 2014 and 2022 were extracted using the best-performing MSNet. The statistics show an increase in mangrove-covered areas in the Indus Delta between 2014 and 2022, with a reduction of 44.37 km², an increase of 170.48 km², and a net increase of 126.11 km².

Keywords: Landsat 8; semantic segmentation; deep learning; mangrove identification; Indus Delta



Citation: Xu, C.; Wang, J.; Sang, Y.; Li, K.; Liu, J.; Yang, G. An Effective Deep Learning Model for Monitoring Mangroves: A Case Study of the Indus Delta. *Remote Sens.* **2023**, *15*, 2220. <https://doi.org/10.3390/rs15092220>

Academic Editors: Mingming Jia, Jeremy L. May and Sergio Vargas Zesati

Received: 19 February 2023

Revised: 8 April 2023

Accepted: 21 April 2023

Published: 22 April 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

Mangroves are widely found in the coastal wetlands of tropical and subtropical regions. Mangroves provide a natural habitat for many birds, fish, shrimps, and insects, as well as provide wind and wave protection, soil and water retention, and carbon storage and sequestration, serving as both an “animal paradise” and a “coastal defender” [1,2]. However, remote sensing has revealed that the global decline in mangroves from 1985–2020 exceeded 10,000 km², and the ecological functions of wetlands were degraded significantly, with the most pronounced shrinkage occurring in South Asia [3]. Direct and indirect human activities (road construction, agricultural land reclamation, aquaculture, etc.) [4] and natural factors (storms, floods, fires, etc.) [5] place mangroves at a high risk of destruction.

The Indus Delta is one of the most important areas of mangrove growth in South Asia [6]. The area is dominated by a single species, *Avicennia marina*, also known as gray or white mangrove, which accounts for over 95% of the total mangrove growth [7]. The Pakistan Forest Act 1972 was enacted by the Government of Pakistan in 1958, and afforestation activities have been carried out on various scales; however, large areas of mangrove forests in the Indus Delta are still dying out [8]. Although a very small proportion of other mangrove species was once found at the mouth of the Indus River at Keti Bunder, most of them are now extinct in Pakistan [9]. To achieve the UN 2030 Agenda faster and better (protecting and sustainably using oceans and marine resources for sustainable development (Sustainable Development Goal 14.2, SDG14) [10]), it is important to use a more efficient and accurate method for monitoring mangroves in the Indus Delta for ecological conservation in coastal wetlands.

Remote sensing technology provides powerful support for mangrove identification owing to its advantages of low cost, high efficiency, and wide range of observations [11]. The most of the common methods for mangrove extraction using remote sensing images are the index methods and supervised machine learning classification. Regular indices such as the normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and visible atmospheric resistance index (VARI) are widely used in vegetation studies [12,13]. Sudip et al. [14] used Sentinel-2 imagery to develop a discriminant normalized vegetation index (DNVI) based on two short-wave infrared (SWIR) bands and used the DNVI to map the interspecific distribution and health of mangroves in the Sundarban Delta, Bangladesh. Pujiono et al. [15] assigned NDVI images from different years to RGB colors and used the additive color theory to form an RGB-NDVI index to explain mangrove variation in the Maubesi Nature Reserve, Indonesia. Jia et al. [16] effectively extracted mangrove forests inundated by tidal flooding near Zhenzhu Harbor, Guangdong Province, China, using the reflectance of the red band and short-wave near-infrared band of Sentinel-2 as a linear baseline. Further, they combined the average reflectance values of the four red-edge bands on top of this linear baseline to establish the mangrove forest index (MFI). Supervised classification relies on computer mathematical models to learn and train manually labeled sample features to extract different types of ground objects from remote sensing images [17]. Supervised machine learning classification methods commonly used for mangrove extraction include K-nearest neighbor (KNN) [18], support vector machines (SVMs) [19,20], and random forest (RF) [21,22]. Some scholars have combined index methods with machine learning methods. Luis et al. [23] constructed several vegetation index series images (EVI2, NDVI, and VARI series images) based on Sentinel-2 time series images to effectively assess the seasonal patterns of mangroves in the Gulf of Mexico, and demonstrated the feasibility of extracting mangroves in semi-arid coastal systems based on time series vegetation indices using the Google Earth Engine (GEE) platform's random forest algorithm. However, the index methods are usually based on the spectral features of the image but ignore other underlying features (texture, color, structure, etc.); traditional supervised machine learning classification methods are susceptible to "same spectrum with different ground objects" and "same ground objects with different spectrums", resulting in low accuracy of mangrove extraction [24].

Since the successful application of the AlexNet convolutional neural network [25] in image recognition in 2012, computer vision has been widely used in the field of remote sensing, and various convolutional neural networks (CNNs) such as FCN [26], SegNet [27], and U-Net [28], etc. have been widely used for semantic segmentation of remote sensing images, and these semantic segmentation models have also made some progress in extracting mangroves. Ariel et al. [29] used UAV imagery as training data to successfully identify five mangrove species in the northern coastal area of Clarin, Bohol, by introducing convolutional neural networks using the transfer learning approach. Jamaluddin et al. [30] extracted the spatial-spectral-temporal deep features of mangroves before and after a hurricane from Sentinel-2 images and then used FCN as a classifier to effectively assess the area of degraded mangroves affected by Hurricane Irma along the southwest coast of

Florida, USA. Guo et al. [31] proposed the ME-Net semantic segmentation model inspired by deep learning techniques. The model is based on the FCN structure with global attention modules, multiple-scale context embedding, and boundary fitting units. The model was very effective at extracting mangroves with an overall accuracy of 97.48%. Li [32] obtained the physical structural features of mangroves by adding a set of convolutional neural networks and combining them with features extracted from UAV images to train an improved SegNet semantic segmentation model for the automatic extraction of mangroves in the Pearl River Delta, China. Davide and Minerva [33] proposed a cloud-based mangrove monitoring framework in which U-Net, ResNet, and RF were used to extract mangrove distributions in Southeast Asia. Finally, the F1-Score was compared to confirm that U-Net had the highest extraction accuracy. Moreno et al. [34] built a U-Net architecture with three backbone networks (ResNet-101, VGG16, and Efficient-net-B7) based on Sentinel-1 time series imagery and effectively monitored the temporal patterns of mangroves along the southeastern Brazilian coast. However, these semantic segmentation models involve a large number of parameters, lose a lot of semantic information, and are difficult to train [35].

To date, few studies have applied deep learning to medium-resolution images to extract mangroves on a large scale. Most samples required for deep learning rely on manual annotation, which is time-consuming and highly subjective. In addition, traditional semantic segmentation models not only contain a large number of parameters, but also use a large number of pooling layers in their structure, resulting in a substantial loss of image spatial information. It is easy to overlook feature-sparse pixels during the training process, and the recovered images do not correspond well to the original image after upsampling.

To solve these problems, this study set the Indus Delta as the study area to enrich the theory and methods of applying deep learning to mangrove extraction. The integration of land use and NDVI data from previous years was used to produce integrated labels to improve the efficiency and accuracy of producing mangrove label data. We propose a parallel, fewer-parameter semantic segmentation model, Multiple Scale Network (MSNet), which fuses the multiple-scale underlying features, and evaluate the effectiveness of this model with other semantic segmentation models for mangrove extraction.

2. Materials

2.1. Study Area

The Indus River is the largest river in South Asia, flowing through several countries (Pakistan, China, India, etc.), and is approximately 2900 km long [36]. As the upper reaches of the Indus are mostly glaciers and snow-capped mountains, snowmelt carries a large amount of sediment that accumulates in the riverbed, resulting in a fan-shaped delta approximately 250 km wide at the estuary of the Indus, called the Indus Delta [37]. The Indus Delta is located in the city of Hyderabad, in the southern province of Sindh, Pakistan. It is bordered by the Arabian Sea to the south, and lies between $23^{\circ}30'0''$ – $25^{\circ}0'0''$ N and $67^{\circ}0'0''$ – $68^{\circ}30'0''$ E (Figure 1), with a total area of approximately 57,000 km². The Indus Delta coastline is approximately one-quarter of the total length of Pakistan's coastline [38]. The delta contains dozens of rivers and streams that support the sixth-largest mangrove forest in the world (the Indus Delta Mangroves) [39].

2.2. Data

The satellite images used in this study were selected from the Landsat 8 Operational Land Imager (Landsat 8 OLI), which has nine bands and a spatial resolution of 30 m [40]. As subtropical mangrove forests are most abundant between February and April and rainfall is minimal during this period [41], remote sensing images of the Indus Delta from 15 February 2022 to 15 April 2022 were acquired using the GEE cloud-computing platform. We selected the visible and infrared bands (Band 1–7) (Table 1) that were favorable for mangrove extraction, excluding the bright temperature bands (Band 10–11), limited the cloud content in the study area to less than 5%, and used the CFMASK masking algorithm

to eliminate the interference of extraneous factors such as clouds, water vapor, and shadows in the images [42].

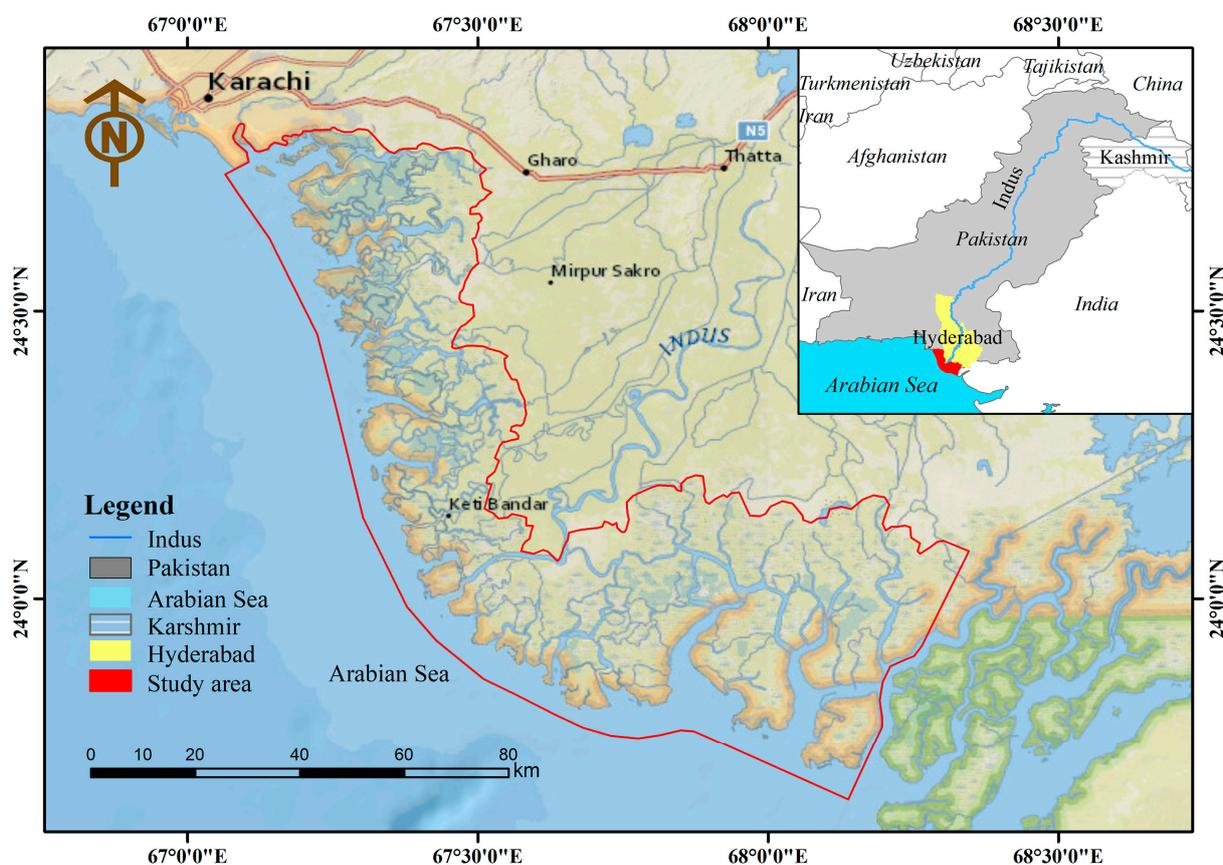


Figure 1. Location of the study area. (The red line area is the extent of study area).

Table 1. Image information of Landsat 8 OLI.

Sensor	Spectral Band	Wavelength	Resolution	Date
OLI	Band 1	0.433–0.453 μm	30 m	15 February 2022–15 April 2022
	Band 2	0.450–0.515 μm		
	Band 3	0.525–0.600 μm		
	Band 4	0.630–0.680 μm		
	Band 5	0.845–0.885 μm		
	Band 6	1.560–1.660 μm		
	Band 7	2.100–2.300 μm		

The GEE platform was used to calculate the NDVI (Formula (1)) for the study area [43] and to obtain annual-scale land use data (2020 and 2021) for the study area.

$$NDVI = (NIR - Red) / (NIR + Red) \quad (1)$$

where *Red* is Band 4 in the Landsat 8 OLI sensor and *NIR* is Band 5 in the Landsat 8 OLI sensor.

The land use data were generated by the ESA World Cover project with a spatial resolution of 10 m and contained 17 land types with an overall accuracy of approximately 75%.

3. Methods

The steps of mangrove identification in this study were divided into the following main parts: image acquisition, sample labeling, model training and prediction, and model

evaluation (Figure 2). First, the Landsat 8 OLI was used as the data source to obtain images of the study area and area of interest (AOI), and the integrated labels of the AOI were used to produce the mangrove training dataset required for the deep learning model. MSNet was then compared with three other semantic segmentation models, FCN-8s, SegNet, and U-Net, to determine the best method for extracting coastal mangroves. To evaluate the performance of the deep learning models, the training parameters of different models and the loss curves in the training were compared, and eight accuracy evaluation metrics were selected to assess the validation accuracy of the four methods. Finally, the best-performing semantic segmentation model was applied to extract the mangrove-covered areas of the Indus Delta.

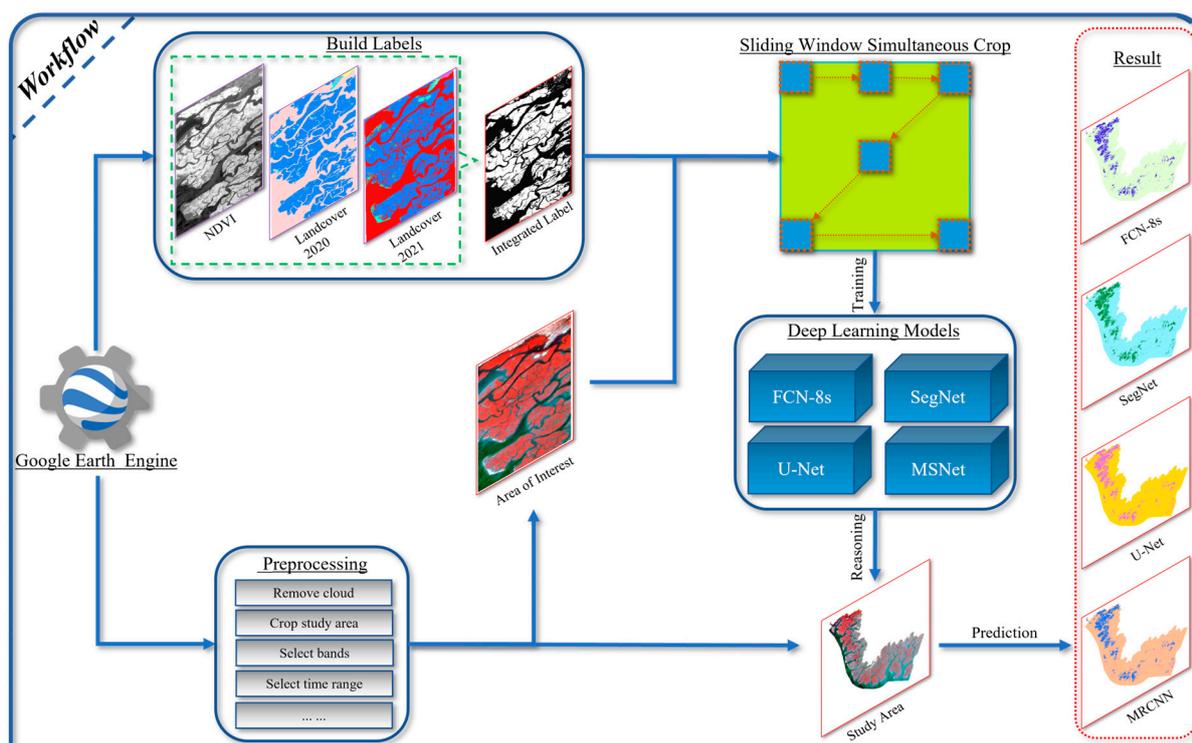


Figure 2. Flow chart of mangrove extraction.

3.1. Label Building

To reduce the misinterpretation and misclassification that occurs when mangrove samples are traditionally labeled manually, we used multiple sources of data to produce mangrove labels, which improved the accuracy of the mangrove sample dataset [44]. The land use data for 2020 and 2021 were reclassified into two categories, “Mangrove” and “others”, and then a suitable threshold was selected to extract the mangrove-covered areas from the NDVI images ($NDVI > 0.12$). The areas where the mangrove categories intersected in the three images were extracted using the ArcGIS software (Figure 3). To increase the readability of the “Mangrove” labels by the deep learning model, two different AOI images of the study area were acquired through the GEE platform (both AOI images are from 2022).

To verify the accuracy of the integrated labels, we used ArcGIS software to generate 1000 random verification points within the mangrove labels of each of the two AOI images (AOI-1 and AOI-2). The label values were then compared with the ground truth values, and finally, the confusion matrix of the integrated labels was calculated (Table 2). The Precisions of AOI-1 and AOI-2 were calculated to be 99.80% and 99.70%, respectively, and could be used as training labels.

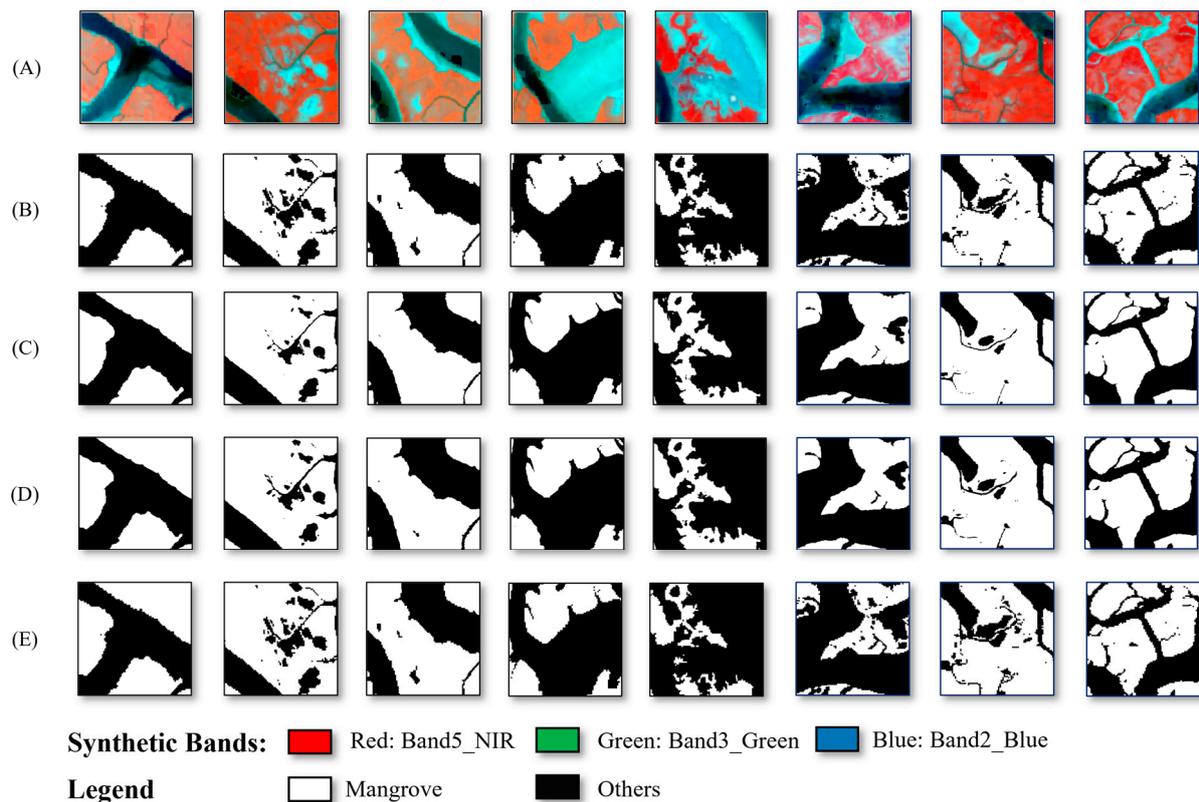


Figure 3. Examples of the original images and labels. Row (A) Landsat 8 OLI images; Row (B) NDVI reclassification images ($NDVI > 0.12$); Row (C) reclassification images of landcover in 2020; Row (D) reclassification images of landcover in 2021; Row (E) integrated labels.

Table 2. Confusion matrix of integrated labels (“M” indicates “Mangrove”; “O” indicates “Others”).

	AOI-1		AOI-2		Sum
	M	O	M	O	
M	998	0	997	0	1995
O	2	0	3	0	5
Sum	1000	0	1000	0	2000
Precision	99.80%		99.70%		99.75%

As the training sample size required for deep learning models must be uniform, and to consider the fairness of comparing the performance of different models, this study used the sliding window crop method to simultaneously crop the AOI images and labels into a $256 \times 256 \times 7$ size [45]. To obtain more training samples, the coverage repetition rate of adjacent windows was set to 0.5 during the cropping process. Using the sliding window cropping method, 1056 images of the training samples were obtained directly. Considering the heterogeneous distribution of “Mangrove” in the labeled images (the number of “Mangrove” pixels in some training labels is much smaller than the number of “Others” pixels), we excluded samples with few mangrove labels from the training data, to avoid the loss of the “Mangrove” category being ignored during the training process. In addition, deep learning models usually require large amounts of training data for training; therefore, we augmented the acquired training data (augmentation methods include horizontal flipping, vertical flipping, and mirror flipping). After data augmentation, 4224 training sample images were obtained. Some of the training sample images and label data are shown in Figure 3.

3.2. Deep Learning Models

Classical semantic segmentation models, such as FCN, SegNet, and U-Net, are based on a structure of encoding followed by decoding, where convolutional and pooling layers are usually used in the encoding part to quickly obtain the feature information of the target. Although pooling methods such as maximum pooling, hole pooling, or average pooling can increase the efficiency of the model and highlight the edges and background of the image, much spatial information is lost in the image, some tiny pixels are ignored for learning, and the outline of the objects after upsampling does not correspond well to the field conditions. Dong et al. [46] reduced the loss of image spatial information by constructing convolutional layers with different step sizes; however, this significantly increases the number of parameters in the model and requires a high level of hardware. Thus, this study proposes the MSNet model with fewer model parameters and the ability to fuse multiple scales of underlying features while maintaining high-resolution features and then extracts mangroves using four models: FCN, SegNet, U-Net, and MSNet. All these four semantic segmentation models use the rectified linear unit (ReLU) [47] activation function for the convolutional layers, the SoftMax [48] activation function for the output layer to normalize the high-dimensional vectors (calculate the probability of each pixel being of a certain class), and cross-entropy [49] for the loss function to measure the difference between the true and predicted values. The model code used in this study can be found on GitHub, <https://github.com/9MorningStar9/MSNet> (accessed on 3 February 2023).

3.2.1. FCN-8s

The FCN is based on VGG16, where the fully connected layers are replaced by convolutional layers and the network output is no longer a category but a heat map. To address the effect of convolution and pooling on the image size, an upsampling approach is proposed to recover the image (Figure 4). The FCNs are divided into FCN-32s ($\times 32$ upsampling), FCN-16s ($\times 16$ upsampling), and FCN-8s ($\times 8$ upsampling), depending on pooling result at different multiples. The upsampling process of FCN-8s is performed three times, and the skip structure is used to add the features of the third and fourth pooling layers to the upsampling layer features to improve the prediction accuracy. The 8-fold upsampling of FCN-8s is much better than the 32-fold and 16-fold upsampling models (FCN-32s, FCN-16s) [50].

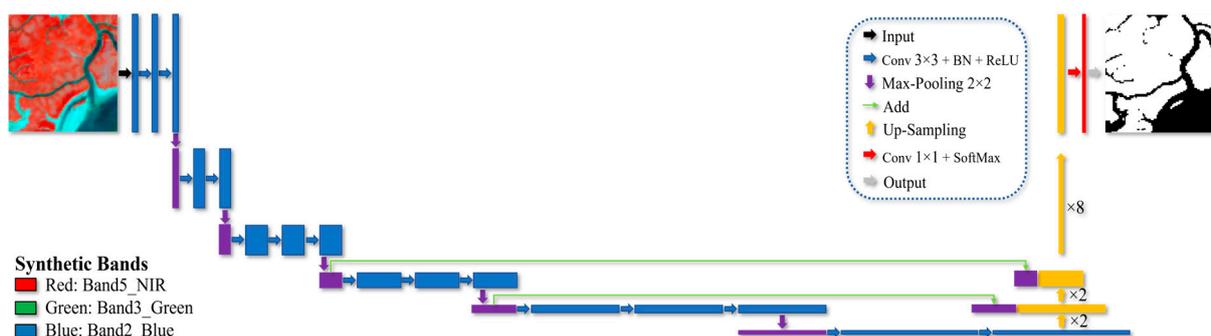


Figure 4. Structure of FCN-8s.

3.2.2. SegNet

The encoding part of SegNet also follows the model of VGG16 but replaces the fully connected layer with a convolutional layer in the decoding part (Figure 5). SegNet uses encoding followed by a decoding structure, but uses maximum-pooling indices to perform the upsampling process in the decoding part. This approach preserves location information at the time of pooling, reduces the scope of error during upsampling, improves the delineation of boundaries, and reduces the number of parameters for end-to-end training [51].

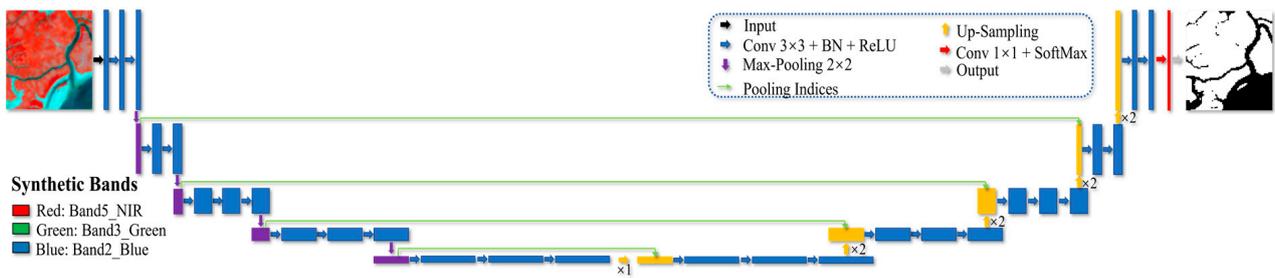


Figure 5. Structure of SegNet.

3.2.3. U-Net

U-Net uses a fully convolutional neural network to quickly extract image features using convolutional and pooling layers on the left side of the encoding part and then performs feature fusion on the right side of the decoding part (Figure 6). The decoding section directly uses the upsampling layers to convert the low-resolution image into a high-resolution image with high-level features, and then concatenates the features with the corresponding low-resolution image on the left [52].

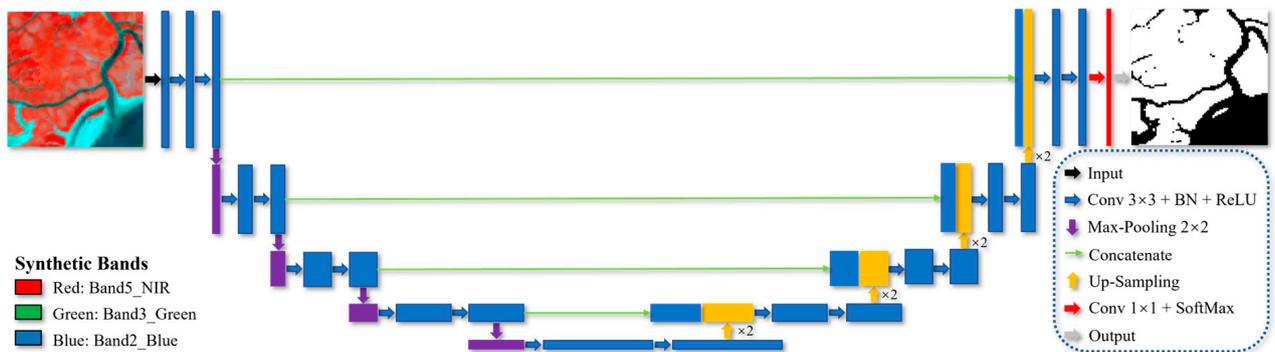


Figure 6. Structure of U-Net.

3.2.4. MSNet

The MSNet model is divided into two parts: underlying feature extraction and fused feature extraction. In underlying feature extraction, the underlying features (texture, edge, spectrum, etc.) of multiple-scale images are extracted rapidly by four pooling and convolution layers (I–IV in Figure 7) to obtain the basis for feature recognition by the model at different resolutions and are converted into high-resolution feature images by upsampling ({1}–{4} in Figure 7). In the fused feature extraction part, the spatial resolution of the image is maintained by building four sets of full-size convolutions ({1}–{4} in Figure 7). These layers ensure that the model can fully learn from feature-sparse pixels to compensate for spatial information lost during pooling. Before extracting features from each set of full-size convolutional layers, feature fusion is performed, in which multiple-scale features less than and equal to the same stage are concatenated. Taking the third stage as an example: the first convolutional layer of the full-size convolutional group {3} performs the concatenate operation with the multiple-scale features {1}–{3}. These upsampling features are used as secondary features for learning during the fusion process, thus further improving the model segmentation accuracy. Finally, the fused and re-extracted features are compressed by a dimensional transformation, and the SoftMax function is used to predict each pixel class of the image.

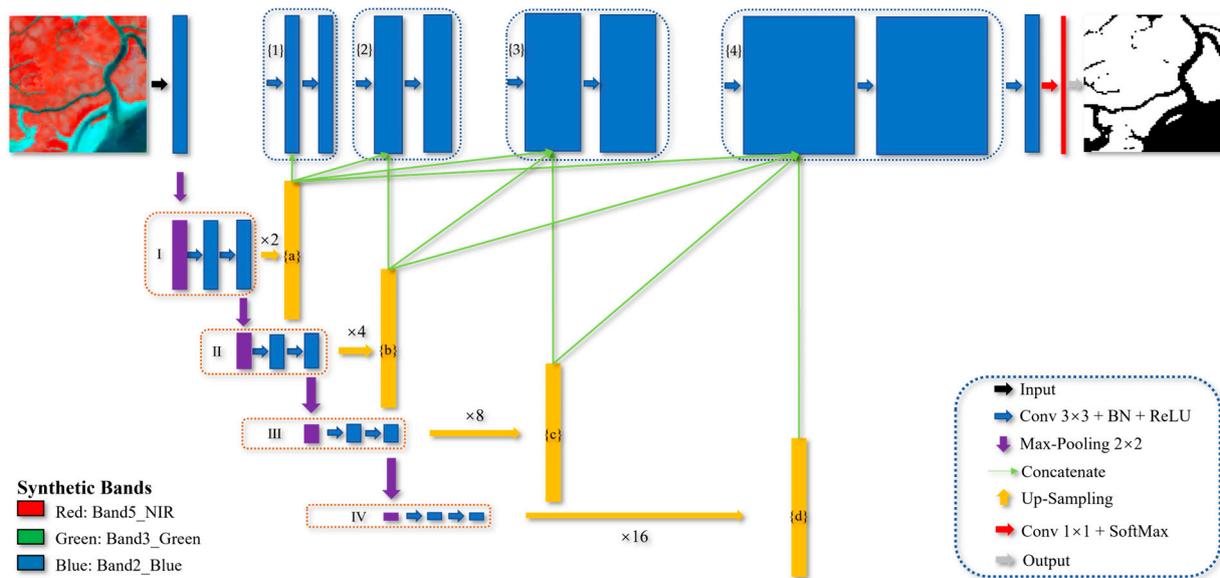


Figure 7. Structure of MSNet.

3.3. Evaluation Metrics

In this study, ten metrics were selected to evaluate the performance of the four models: Training loss curve, Model parameter count, Training time, Precision, Recall, Overall accuracy (OA), F1-Score, Intersection-over-union (IoU), Frequency weighted intersection-over-union (FWIoU), and Kappa coefficient. Precision indicates exactly how many samples whose predicted outcome is a positive class are true positives, i.e., the ratio of the number of samples accurately classified as positive to the total number of samples classified as positive (Formula (2)). Recall indicates how many samples that should have been classified as positives would have been correctly classified, i.e., the ratio of the number of samples classified as positives to the number of samples actually classified as positives in the test dataset (Formula (3)). OA is the probability that the classification result of each random sample matches the class of the test data (Formula (4)). F1-Score is a balanced value between Precision and Recall, allowing both to be maximized (Precision and Recall are contradictory and cannot be maximized simultaneously) (Formula (5)). IoU is the ratio of the intersection of the actual category samples and the predicted category samples to their concatenation (Formula (6)). FWIoU is set with weights based on the frequency of occurrence of each category. The weights are multiplied by the IoU of each category and summed (Formula (7)). The Kappa coefficient is typically calculated to be between 0 and 1. A larger Kappa value indicates a higher classification accuracy. It is a measure of classification accuracy and is often used in conjunction with the OA to check the spatial consistency of image classification (Formula (8)).

$$Precision = TP / (TP + FP) \quad (2)$$

$$Recall = TP / (TP + FN) \quad (3)$$

$$OA = (TP + TN) / (TP + TN + FP + FN) \quad (4)$$

$$F1 = (2 \times Precision \times Recall) / (Precision + Recall) \quad (5)$$

$$IoU = TP / (TP + FN + FP) \quad (6)$$

$$FWIoU = [(TP + FN)/(TP + FP + TN + FN)] \times [TP/(TP + FP + FN)] \quad (7)$$

$$Kappa = (OA - P_e)/(1 - P_e) \quad (8)$$

The formula of P_e is shown below.

$$P_e = [(TP + FN) \times (TP + FP) + (FP + TN) \times (FN + TN)] / (TP + FP + TN + FN)^2 \quad (9)$$

where true positives (TP) are pixels whose label and prediction are both “Mangrove”; true negatives (TN) are pixels whose label and prediction are both “Others”; false positives (FP) are pixels whose label is “Mangrove” but predicted as “Others”; false negatives (FN) are pixels whose label is “Others” but predicted as “Mangrove”.

4. Results

4.1. Performance of Models

The deep learning models in this study were trained using an NVIDIA RTX3060 GPU and TensorFlow (GPU version) 2.6.0. The optimizer for all four models was Adam, with an initial learning rate of 0.001. To better compare the performances of the models, we eliminated the early stop strategy and trained each model for 100 epochs. Owing to the memory limitations of the GPU, the training batch size was set to 16, and each epoch was iterated 264 times. During the training of the deep learning model, if the model loss did not show a decreasing trend for three consecutive epochs, the learning rate was automatically halved to fit the model better. Finally, the cross-entropy loss function was used to plot the loss curves of the four models (Figure 8). The lower the loss value and faster and smoother the decline in the loss curve, the better the adaptability and robustness of the model.

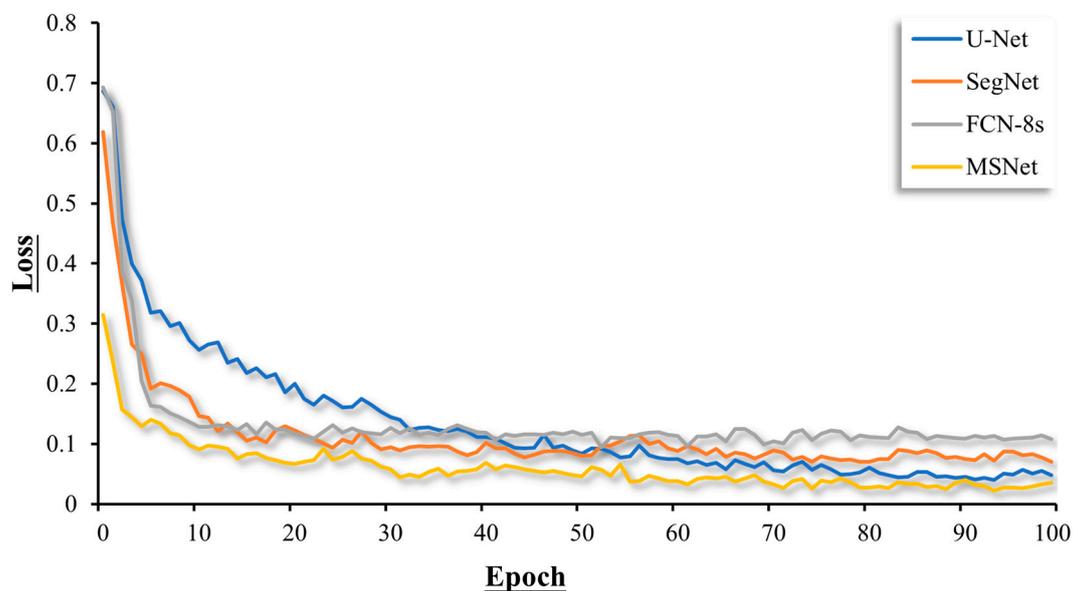


Figure 8. Training losses of FCN-8s, SegNet, U-Net, and MSNet.

As shown in Figure 6, FCN-8s, SegNet, and U-Net exhibited similar losses in the first epoch, whereas MSNet exhibited the lowest initial loss. The losses for FCN-8s and SegNet during the first 20 epochs were similar. However, the loss of FCN-8s almost stopped decreasing after the 20th epoch, whereas the SegNet loss decreased slowly and converged gradually. The loss of MSNet decreased at a rate comparable to those of FCN-8s and SegNet in the first 10 epochs. The loss value of U-Net decreased the slowest; however, its loss was lower than those of FCN-8s and SegNet at the 36th and 52nd epochs, respectively. Finally, it was the second lowest after MSNet. The MSNet model maintained the lowest loss value throughout the training process.

As shown in Table 3, U-Net has the largest number of model parameters; however, its training time is the shortest at 100 epochs and the most efficient among the four models. The training time of SegNet was the longest among the four models, and the number of model parameters was large, with the third highest loss value. Although the number of model parameters and training time of FCN-8s were similar to those of MSNet, its loss value was the highest among the four models. MSNet was average in terms of training time; however, it had the smallest number of model parameters and the lowest loss value.

Table 3. Comparison of model parameters, minimum losses, and training time (the best-performing one is shown in bold).

	Total No. of Parameters	Minimum Loss	Training Time
FCN-8s	249,594	0.0945	1 h 15 min 3 s
SegNet	463,018	0.0702	1 h 17 min 31 s
U-Net	492,560	0.0397	1 h 11 m 12 s
MSNet	161,312	0.0217	1 h 15 min 19 s

To better compare the results of the four models for mangrove extraction, we generated 2500 random validation points using ArcGIS software to validate the mangrove extraction results from the images obtained in 2014. The sample points were labeled with attributes with reference to the higher spatial resolution simultaneous-phase Sentinel-2 imagery (622 points for the category “Mangrove” and 1878 points for the category “Others”). The confusion matrices of the four models were calculated using these sample points (Table 4), and their Precision, Recall, OA, F1-Score, IoU, FWIoU, and Kappa coefficient were obtained (Table 5).

Table 4. Confusion matrix of models (“M” indicates “Mangrove”; “O” indicates “Others”).

	FCN-8s		SegNet		U-Net		MSNet		SUM
	M	O	M	O	M	O	M	O	
M	517	105	556	66	562	60	572	50	622
O	19	1859	8	1870	12	1866	9	1869	1878
SUM	536	1964	564	1936	574	1926	581	1919	2500

Table 5. Comparison of validation results (the best-performing one is shown in bold).

	Precision	Recall	OA	F1-Score	IoU	FWIoU	Kappa
FCN-8s	96.45%	83.12%	95.04%	89.29%	80.66%	20.07%	86.09%
SegNet	97.91%	89.39%	97.04%	93.76%	88.25%	21.96%	91.83%
U-Net	98.58%	90.35%	97.12%	93.98%	88.64%	22.05%	92.09%
MSNet	98.45%	91.96%	97.64%	95.09%	90.65%	22.55%	93.54%

Table 5 shows that FCN-8s is the worst performer, ranking last in all metrics. In the Precision metric, U-Net has the highest Precision at 98.58%. The Precision of MSNet is slightly lower than U-Net’s Precision, with a difference of only 0.13%. The Precisions of SegNet and FCN-8s are 97.91% and 96.45%, respectively. In the other metrics, the performances of U-Net and SegNet are similar, while MSNet outperforms the other models. In particular, the difference between MSNet and the second-ranked U-Net is significant for Recall, Kappa, and IoU. MSNet’s Recall is 91.96%, 1.61% higher than that of U-Net. The Kappa of MSNet is 93.54%, 1.45% higher than U-Net’s. The metric with the largest difference between MSNet and U-Net is IoU, which is 90.65% for MSNet, 2.01% higher than U-Net. Compared to the worst-performing FCN-8s, MSNet’s Recall and Kappa are 8.84% and 7.45% higher than those of FCN-8s, respectively. There is also a significant difference in the IoU metric, with FCN-8s’s IoU being only 80.66% and MSNet’s IoU being 9.99% higher than that of FCN-8s. Combining a total of ten metrics from Tables 3 and 5, it can be seen that the MSNet model is the best performer among the four models.

4.2. Extraction Results of Mangroves

To understand the accuracy of the four models in extracting mangroves in greater detail, we selected several typical mangrove distribution areas for comparative analysis. Figure 9 shows remote sensing images of the study area and the extraction results of the four models. The images included mangroves, bare ground, rivers, and a small areas of building land. The main areas covered by mangroves were extracted using the four models. However, mangroves in the intertidal zone or those with a low cover density did not show significant features in the images. FCN-8s was unable to effectively learn mangrove pixels with sparse features, and the boundaries of the mangrove extraction results were severely jagged, with a large amount of spatial information being lost (Row B). The mangrove boundaries extracted using SegNet were smoother and clearer, with a slight improvement in performance in identifying intertidal mangroves (Row C, Columns 2 and 3), but it was less sensitive to mangrove pixels with lower cover density (Row C, Columns 4–7). SegNet tends to ignore tiny ground objects and classify them as “Mangrove”, which results in a slight incoherence in the “Others” class of the extraction results (Row C, Columns 1 and 8). U-Net shows a large improvement in distinguishing the boundary between fine rivers and mangroves (Row D, Columns 1 and 2, 7 and 8), but has no significant advantage in identifying mangrove the pixels with low cover density (Row D, Columns 4–7). MSNet can effectively distinguish pixels of intertidal mangroves from those of water bodies because of its ability to fully learn detailed features and fuse multiple scales of the underlying features. It was more sensitive to mangrove pixels with a lower cover density and could correctly identify more mangrove-covered areas.

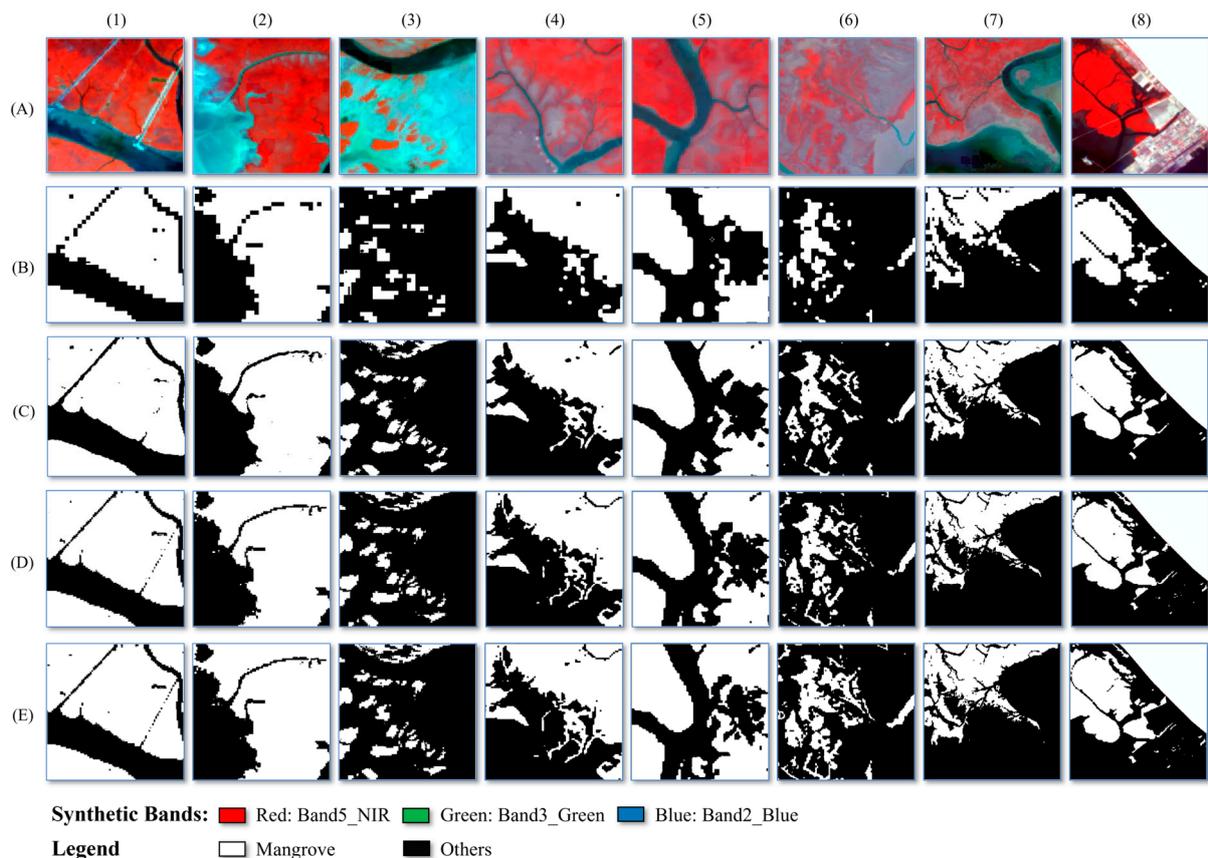


Figure 9. Detailed information on the extraction results of mangroves under different models. Row (A) Landsat 8 OLI image; Row (B) FCN-8s; Row (C) SegNet; Row (D) U-Net; Row (E) MSNet. Columns (1–8) represent groups of classification results.

4.3. Spatial Variation of Mangroves

Figure 10 shows the results of mangrove extraction in the Indus Delta using MSNet and the changes in mangrove-covered areas over 8 years. In general, mangrove area in the Indus Delta has increased. The stable growth area was mainly concentrated in the west section near Karachi, and the decreased area was also mainly concentrated in the western section with a high mangrove cover density, while the increased area was mainly concentrated near Keti Bandar, and a small part of the increased area was distributed in the eastern section near India. According to statistics, the mangrove area of the Indian Delta in 2014 was approximately 579.87 km². By 2022, the mangrove area increased to 705.98 km². During this period, the mangrove-covered area increased by approximately 170.48 km², the decreased area was 44.37 km², and the net increased area was 126.11 km².

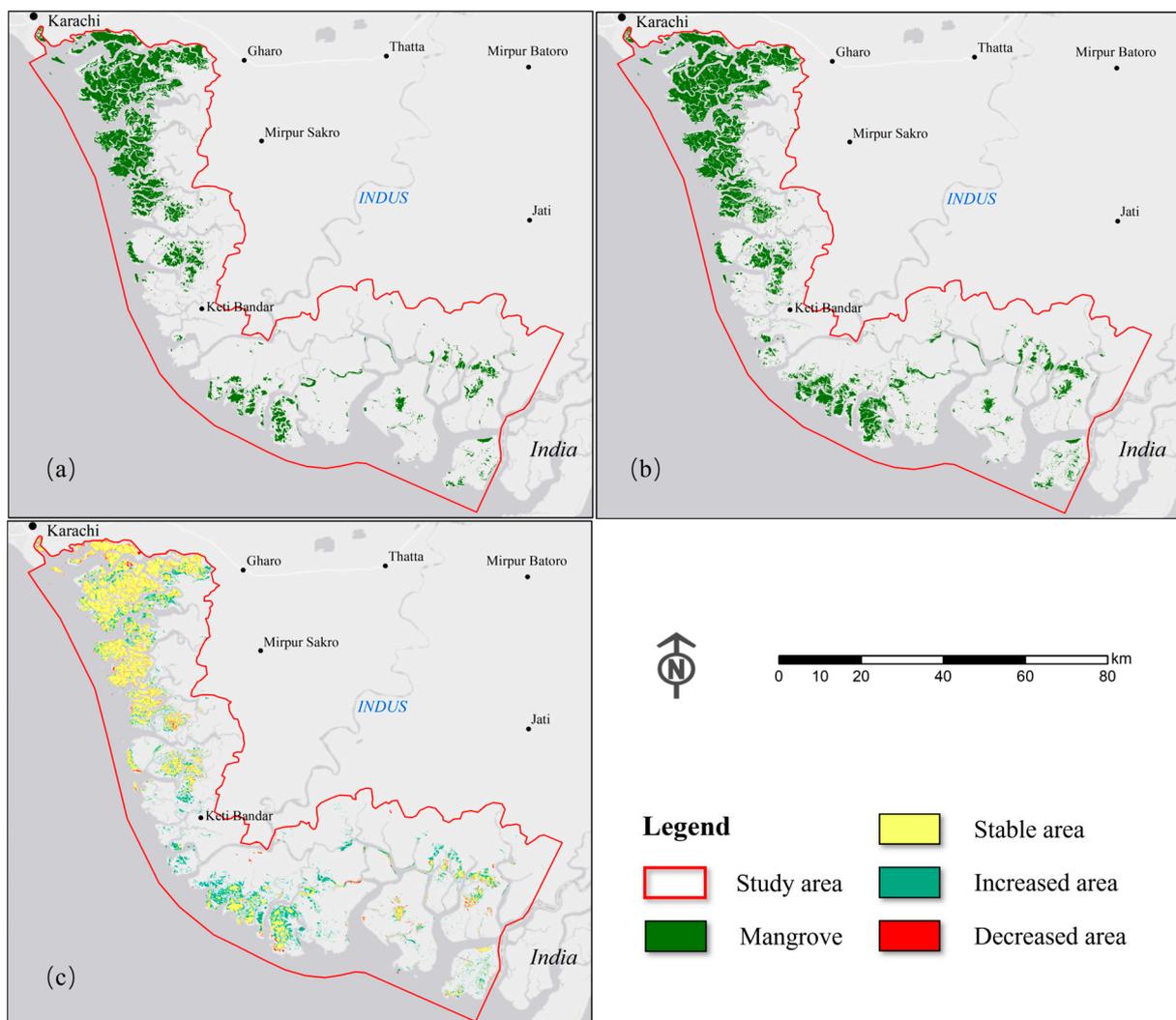


Figure 10. Mangrove extraction results of Indus Delta. (a) Mangrove-covered areas in the Indus Delta in 2014; (b) mangrove-covered areas in the Indus Delta in 2022; (c) changes in mangrove-covered areas in the Indus Delta from 2014 to 2022.

5. Discussion

Owing to the multiple pooling and downsampling of the FCN-8s model, the spatial information of the feature image was significantly lost. In the decoding part, continuous upsampling is adopted, and the upsampling feature is not learned as in other models, which results in fuzzy upsampling. Finally, FCN-8s restores the underlying features to the resolution of the original image using 8-fold upsampling. This further widens the gap

between the upsampling feature and the original image feature and leads to unsatisfactory segmentation results of the model. The boundary of the SegNet segmentation results is clearer, mainly because of the maximum pooling indices. Only the boundary information was obtained and stored in the feature mapping in the encoding part, and the accuracy of the boundary was guaranteed by the coordinate index in the decoding part. However, this structure overemphasizes the importance of object boundaries and ignores the detailed information of the image. Misclassification and missing classification often occur at the junctions of rivers and mangroves. U-Net adopts a completely symmetrical U-shaped structure and fuses low-level features with high-level features at the same resolution to improve the segmentation accuracy, which has been improved in distinguishing small rivers and mangroves. However, after the continuous downsampling of mangrove pixels with sparse features, the feature information is significantly lost, and the model cannot learn the information of these pixels. As a result, U-Net identified only a small portion of mangroves when predicting areas with low mangrove density. The extraction result of MSNet was the best, slightly better than that of U-Net in distinguishing small rivers and mangroves. It recognized most mangroves with sparse features. During the training process, the model constantly fuses the upsampling features of the multi-resolution images to fill in the underlying information of the image. This enhances the basis for ground object recognition, segmentation, and positioning. In the fused feature extraction, no pooling operation was performed; however, feature extraction was performed on the premise of maintaining the spatial resolution of the original images so that the model could learn more detailed feature information. In addition to the disadvantage of spatial information loss caused by the pooling operation, the lack of universality is one of the reasons for the poor extraction of mangroves by FCN-8s, SegNet, and U-Net. In medium-resolution remote sensing images, a single pixel typically contains several features of various ground objects. It is difficult to distinguish between these complex pixels using these models. These models are often used for indoor object recognition, unmanned driving, and medical diagnosis [52]. Objects such as human bodies, vehicles, and cells usually have high-resolution images. These images usually have clear textures, colors, and structural characteristics. Compared with the other three classical semantic segmentation models, the MSNet model with fewer parameters is more suitable for remote sensing images and is easy to train.

In addition, the classification results of MSNet were compared with those of the semi-automatic classification plugin (SCP) in QGIS, where the minimum spectral distance (MSD) and maximum likelihood (ML) classification methods were used [53,54]. MSD and ML were comparable to MSNet in their ability to distinguish between rivers and mangroves boundaries. However, these two machine learning methods have difficulty distinguishing between wetlands and mangrove pixels with sparse features, resulting in significant misclassification of wetland as mangroves. This is because they can only learn the superficial features of the image, which further highlights the superiority of MSNet over traditional machine learning methods in mangrove extraction.

Although MSNet can identify most feature-sparse mangrove pixels, we identified some problems that needed to be solved during our research. First, the image bands selected for the study may contain a large amount of redundant information and the responses of individual bands to vegetation are unclear. Second, the integrated labels obtained from the NDVI and land use maps may not be sufficiently precise (a few labeled areas contain other classes), allowing the model to learn the incorrect feature information. This integrated label is currently used only in medium-resolution images and is not necessarily applicable to images of other resolutions. In future research, we plan to combine multi-source satellite data, spatial-spectral fusion, and multi-factor features (NDVI, NDWI, MFI, etc.) to enhance the capacity for different mangrove type extractions. In addition, the applicability and accuracy of the integrated labels for images of different resolutions will be further improved, and the deep learning model structure will be optimized. We believe this approach can be applied for the interspecific classification of mangroves or other types of land cover in more delta or coastal regions with rich biodiversity.

6. Conclusions

This study proposed a deep learning model based on integrated labels, using the Indus Delta in Pakistan as the study area. The effectiveness of MSNet in extracting mangroves was demonstrated by comparing it with three deep learning methods, FCN-8s, SegNet, and U-Net, and by using MSNet to monitor mangrove changes in the Indus Delta. The following conclusions were drawn: (1) among the mangrove results extracted by the four deep learning models, MSNet extracted mangroves the best. U-Net was poor at discriminating areas with low mangrove cover densities. SegNet formed clear boundaries but had difficulty distinguishing intertidal mangroves from water. FCN-8s was the least effective at extracting mangroves. (2) The integrated labels produced using NDVI and land use data greatly reduced the cost of manually labeling the samples. This method can eliminate the potential for manual labeling of incorrect samples owing to subjective assumptions. (3) MSNet had the fewest parameters and the lowest training loss. This remains the leader in the evaluation metrics. Even with the parallel network structure, the training time of this model was not significantly different from those of the other models. MSNet not only maintains the spatial resolution of the image and learns the feature-sparse pixels but also fuses the multiple-scale features to improve the accuracy of ground object recognition and segmentation. (4) From 2014 to 2022, the area covered by mangroves increased from 579.87 km² to 705.98 km², a net increase of about 126.11 km². Overall, although the area of mangrove cover in the Indus Delta has increased, large areas of mangrove forests are still dying.

MSNet, based on integrated labels, is a deep learning method for efficient and accurate mangrove extraction suitable for large-scale mangrove extraction studies. This can help authorities in the Indus Delta to make accurate decisions regarding sustainable planning for mangrove conservation to prevent further damage to mangroves.

Author Contributions: C.X. and J.W. designed the study, processed the data, and drafted the manuscript. K.L. helped to design the study, provided knowledge and experience in technology, and reviewed the manuscript. J.L., Y.S. and G.Y. helped with image analysis, provided advice on the research, and reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China, grant number 2022YFF0711600, and the Construction Project of China Knowledge Centre for Engineering Sciences and Technology, grant number CKCEST-2022-1-41.

Data Availability Statement: The data that support the findings of this study are openly available in [MSNet] at [<https://github.com/9MorningStar9/MSNet>].

Acknowledgments: The authors are grateful for the support of the China–Pakistan Joint Research Center on Earth Sciences.

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

References

1. Wang, L.; Sousa, W.P.; Gong, P.; Biging, G.S. Comparison of IKONOS and QuickBird Images for Mapping Mangrove Species on the Caribbean Coast of Panama. *Remote Sens. Environ.* **2004**, *91*, 432–440. [[CrossRef](#)]
2. Agaton, C.B.; Collera, A.A. Now or Later? Optimal Timing of Mangrove Rehabilitation under Climate Change Uncertainty. *For. Ecol. Manag.* **2022**, *503*, 119739. [[CrossRef](#)]
3. Han, X.; Fu, D.; Ju, C.; Kang, L. 10-M Global Mangrove Classification Products of 2018–2020 Based on Big Data. *Sci. Data Bank* **2021**. [[CrossRef](#)]
4. Murdiyarso, D.; Purbopuspito, J.; Kauffman, J.B.; Warren, M.W.; Sasmito, S.D.; Donato, D.C.; Manuri, S.; Krisnawati, H.; Taberima, S.; Kurnianto, S. The Potential of Indonesian Mangrove Forests for Global Climate Change Mitigation. *Nat. Clim. Change* **2015**, *5*, 1089–1092. [[CrossRef](#)]
5. Goldberg, L.; Lagomasino, D.; Thomas, N.; Fatoyinbo, T. Global Declines in Human—Driven Mangrove Loss. *Glob. Change Biol.* **2020**, *26*, 5844–5855. [[CrossRef](#)]
6. Memon, J.A.; Thapa, G.B. Explaining the de Facto Open Access of Public Property Commons: Insights from the Indus Delta Mangroves. *Environ. Sci. Policy* **2016**, *66*, 151–159. [[CrossRef](#)]

7. Amir, S.A.; Siddiqui, P.J.A.; Masroor, R. Finfish diversity and seasonal abundance in the largest arid mangrove forest of the Indus Delta, Northern Arabian Sea. *Mar. Biodivers.* **2018**, *48*, 369–1380. [[CrossRef](#)]
8. Giri, C.; Long, J.; Abbas, S.; Murali, R.M.; Qamer, F.M.; Pengra, B.; Thau, D. Distribution and Dynamics of Mangrove Forests of South Asia. *J. Environ. Manag.* **2015**, *148*, 101–111. [[CrossRef](#)]
9. Irum, M.; Abdul, H. Constrains on mangrove forests and conservation projects in Pakistan. *J. Coast. Conserv.* **2012**, *16*, 51–62. [[CrossRef](#)]
10. Friess, D.A.; Rogers, K.; Lovelock, C.E.; Krauss, K.W.; Hamilton, S.E.; Lee, S.Y.; Lucas, R.; Primavera, J.; Rajkaran, A.; Shi, S. The State of the World's Mangrove Forests: Past, Present, and Future. *Annu. Rev. Environ. Resour.* **2019**, *44*, 89–115. [[CrossRef](#)]
11. Li, H.; Wang, C.; Cui, Y.; Hodgson, M. Mapping Salt Marsh along Coastal South Carolina Using U-Net. *ISPRS J. Photogramm. Remote Sens.* **2021**, *179*, 121–132. [[CrossRef](#)]
12. Matsushita, B.; Yang, W.; Chen, J.; Onda, Y.; Qiu, G. Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Case Study in High-Density Cypress Forest. *Sensors* **2007**, *7*, 2636–2651. [[CrossRef](#)]
13. Houborg, R.; Soegaard, H.; Boegh, E. Combining Vegetation Index and Model Inversion Methods for the Extraction of Key Vegetation Biophysical Parameters Using Terra and Aqua MODIS Reflectance Data. *Remote Sens. Environ.* **2007**, *106*, 39–58. [[CrossRef](#)]
14. Manna, S.; Raychaudhuri, B. Mapping distribution of Sundarban mangroves using Sentinel-2 data and new spectral metric for detecting their health condition. *Geocarto Int.* **2018**, *35*, 434–452. [[CrossRef](#)]
15. Eko, P.; Woo-Kyun, L.; Doo-Ahn, K.; He, L.; So-Ra, K.; Jongyeol, L.; Lee Seung, H. RGB-NDVI color composites for monitoring the change in mangrove area at the Maubesi Nature Reserve, Indonesia. *For. Sci. Technol.* **2013**, *9*, 171–179.
16. Jia, M.; Wang, Z.; Wang, C.; Mao, D.; Zhang, Y. A New Vegetation Index to Detect Periodically Submerged Mangrove Forest Using Single-Tide Sentinel-2 Imagery. *Remote Sens.* **2019**, *11*, 2043. [[CrossRef](#)]
17. Li, K.; Wang, J.; Yao, J. Effectiveness of Machine Learning Methods for Water Segmentation with ROI as the Label: A Case Study of the Tuul River in Mongolia. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*, 102497. [[CrossRef](#)]
18. Woltz, V.L.; Peneva-Reed, E.I.; Zhu, Z.; Bullock, E.L.; MacKenzie, R.A.; Apwong, M.; Krauss, K.W.; Gesch, D.B. A comprehensive assessment of mangrove species and carbon stock on Pohnpei, Micronesia. *PLoS ONE* **2022**, *17*, e0271589. [[CrossRef](#)]
19. Vidhya, R.; Vijayasekaran, D.; Farook, M.A.; Jai, S.; Rohini, M.; Sinduja, A. Improved Classification of Mangroves Health Status Using Hyperspectral Remote Sensing Data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *40–48*, 667–670. [[CrossRef](#)]
20. Jiang, Y.; Zhang, L.; Yan, M.; Qi, J.; Fu, T.; Fan, S.; Chen, B. High-Resolution Mangrove Forests Classification with Machine Learning Using Worldview and UAV Hyperspectral Data. *Remote Sens.* **2021**, *13*, 1529. [[CrossRef](#)]
21. Behera, M.D.; Barnwal, S.; Paramanik, S.; Das, P.; Bhattyacharya, B.K.; Jagadish, B.; Roy, P.S.; Ghosh, S.M.; Behera, S.K. Species-Level Classification and Mapping of a Mangrove Forest Using Random Forest—Utilisation of AVIRIS-NG and Sentinel Data. *Remote Sens.* **2021**, *13*, 2027. [[CrossRef](#)]
22. Zhao, C.; Qin, C.-Z. Identifying Large-Area Mangrove Distribution Based on Remote Sensing: A Binary Classification Approach Considering Subclasses of Non-Mangroves. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *108*, 102750. [[CrossRef](#)]
23. Valderrama-Landeros, L.; Flores-Verdugo, F.; Rodriguez-Sobreyra, R.; Kovacs, J.M.; Flores-de-Santiago, F. Extrapolating canopy phenology information using Sentinel-2 data and the Google Earth Engine platform to identify the optimal dates for remotely sensed image acquisition of semiarid mangroves. *J. Environ. Manag.* **2021**, *279*, 111617. [[CrossRef](#)]
24. He, S.; Lu, X.; Zhang, S.; Li, S.; Tang, H.T.; Zheng, W.; Lin, H.; Luo, Q. Research on classification algorithm of wetland land cover in the Linhong Estuary, Jiangsu Province. *Mar. Sci.* **2020**, *44*, 44–53. [[CrossRef](#)]
25. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
26. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651.
27. Fu, B.; Liu, M.; He, H.; Lan, F.; He, X.; Liu, L.; Huang, L.; Fan, D.; Zhao, M.; Jia, Z. Comparison of Optimized Object-Based RF-DT Algorithm and SegNet Algorithm for Classifying Karst Wetland Vegetation Communities Using Ultra-High Spatial Resolution UAV Data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102553. [[CrossRef](#)]
28. Stoian, A.; Poulain, V.; Inglada, J.; Poughon, V.; Derksen, D. Land Cover Maps Production with High Resolution Satellite Image Time Series and Convolutional Neural Networks: Adaptations and Limits for Operational Systems. *Remote Sens.* **2019**, *11*, 1986. [[CrossRef](#)]
29. Ariel, C.C.V.; Chris, J.G.A.; Larmie, T.S. Mangrove Species Identification Using Deep Neural Network. In Proceedings of the 2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, 13–14 December 2022. [[CrossRef](#)]
30. Jamaluddin, I.; Thaipisutikul, T.; Chen, Y.-N.; Chuang, C.-H.; Hu, C.-L. MDPrePost-Net: A Spatial-Spectral-Temporal Fully Convolutional Network for Mapping of Mangrove Degradation Affected by Hurricane Irma 2017 Using Sentinel-2 Data. *Remote Sens.* **2021**, *13*, 5042. [[CrossRef](#)]
31. Guo, M.; Yu, Z.; Xu, Y.; Huang, Y.; Li, C. ME-Net: A Deep Convolutional Neural Network for Extracting Mangrove Using Sentinel-2A Data. *Remote Sens.* **2021**, *13*, 1292. [[CrossRef](#)]

32. Li, R.; Shen, X.; Zhai, C.; Zhang, Z.; Zhang, Y.; Jiang, B. *A Method for Automatic Identification of Mangrove Plants Based on UAV Visible Light Remote Sensing*; Peking University Shenzhen Graduate School: Shenzhen, China, 2021.
33. Lomeo, D.; Singh, M. Cloud-Based Monitoring and Evaluation of the Spatial-Temporal Distribution of Southeast Asia's Mangroves Using Deep Learning. *Remote Sens.* **2022**, *14*, 2291. [[CrossRef](#)]
34. Moreno, G.M.D.S.; de Carvalho Júnior, O.A.; de Carvalho, O.L.F.; Andrade, T.C. Deep Semantic Segmentation of Mangroves in Brazil Combining Spatial, Temporal, and Polarization Data from Sentinel-1 Time Series. *Ocean. Coast. Manag.* **2023**, *231*, 106381. [[CrossRef](#)]
35. Xu, X. Research on Remote Sensing Image Feature Classification Algorithm of Island Coastal Zone Based on Deep Learning. Master's Thesis, China University of Mining & Technology, Xuzhou, China, 2022.
36. Ahmed, W.; Wu, Y.; Kidwai, S.; Li, X.; Mahmood, T.; Zhang, J. Do Indus Delta Mangroves and Indus River Contribute to Organic Carbon in Deltaic Creeks and Coastal Waters (Northwest Indian Ocean, Pakistan)? *Cont. Shelf Res.* **2021**, *231*, 104601. [[CrossRef](#)]
37. Giosan, L.; Constantinescu, S.; Clift, P.D.; Tabrez, A.R.; Danish, M.; Inam, A. Recent Morphodynamics of the Indus Delta Shore and Shelf. *Cont. Shelf Res.* **2006**, *26*, 1668–1684. [[CrossRef](#)]
38. Gilani, H.; Naz, H.I.; Arshad, M.; Nazim, K.; Akram, U.; Abrar, A.; Asif, M. Evaluating Mangrove Conservation and Sustainability through Spatiotemporal (1990–2020) Mangrove Cover Change Analysis in Pakistan. *Estuar. Coast. Shelf Sci.* **2021**, *249*, 107128. [[CrossRef](#)]
39. Kidwai, S.; Ahmed, W.; Tabrez, S.M.; Zhang, J.; Giosan, L.; Clift, P.; Inam, A. The Indus Delta—Catchment, River, Coast, and People. In *Coasts and Estuaries*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 213–232. ISBN 978-0-12-814003-1.
40. Chai, D.; Newsam, S.; Zhang, H.K.; Qiu, Y.; Huang, J. Cloud and Cloud Shadow Detection in Landsat Imagery Based on Deep Convolutional Neural Networks. *Remote Sens. Environ.* **2019**, *225*, 307–316. [[CrossRef](#)]
41. Khan, M.U.; Cai, L.; Nazim, K.; Ahmed, M.; Zhao, X.; Yang, D. Effects of the Summer Monsoon on the Polychaete Assemblages and Benthic Environment of Three Mangrove Swamps along the Sindh Coast, Pakistan. *Reg. Stud. Mar. Sci.* **2022**, *56*, 102613. [[CrossRef](#)]
42. Joshi, P.P.; Wynne, R.H.; Thomas, V.A. Cloud Detection Algorithm Using SVM with SWIR2 and Tasseled Cap Applied to Landsat 8. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *82*, 101898. [[CrossRef](#)]
43. Ezaidi, S.; Aydda, A.; Kabbachi, B.; Althuwaynee, O.F.; Ezaidi, A.; Haddou, M.A.; Idoumskine, I.; Thorpe, J.; Park, H.-J.; Kim, S.-W. Multi-Temporal Landsat-Derived NDVI for Vegetation Cover Degradation for the Period 1984–2018 in Part of the Arganeraie Biosphere Reserve (Morocco). *Remote Sens. Appl. Soc. Environ.* **2022**, *27*, 100800. [[CrossRef](#)]
44. Lu, Y.; Wang, L. How to Automate Timely Large-Scale Mangrove Mapping with Remote Sensing. *Remote Sens. Environ.* **2021**, *264*, 112584. [[CrossRef](#)]
45. Wang, Z.Q.; Zhou, Y.; Wang, S.X.; Wang, F.T.; Xu, Z.Y. House building extraction from high-resolution remote sensing images based on IEU-Net. *Natl. Remote Sens. Bull.* **2021**, *25*, 2245–2254. [[CrossRef](#)]
46. Dong, Z.; Wang, G.; Amankwah, S.O.Y.; Wei, X.; Hu, Y.; Feng, A. Monitoring the Summer Flooding in the Poyang Lake Area of China in 2020 Based on Sentinel-1 Data and Multiple Convolutional Neural Networks. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102400. [[CrossRef](#)]
47. Nair, V.; Hinton, G.E. Rectified Linear Units Improve Restricted Boltzmann Machines. In Proceedings of the 27th International Conference on International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010.
48. Gao, B.; Pavel, L. On the Properties of the Softmax Function with Application in Game Theory and Reinforcement Learning. *arXiv* **2018**, arXiv:1704.00805.
49. Kevin, P. *Murphy Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, MA, USA, 2012.
50. Jayanthi, P.; Murali Krishna, I.V. A Comparative Study on Fully Convolutional Networks—FCN-8, FCN-16, and FCN-32. In *Deep Learning for Medical Applications with Unique Data*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 19–30. ISBN 978-0-12-824145-5.
51. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)]
52. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lect. Notes Comput. Sci.* **2015**, *9351*, 234–241.
53. Latifa, P.; Gilles, D.; Alex, S.L.; Paulo, T.; Thierry, F. Coastal Land Use in Northeast Brazil: Mangrove Coverage Evolution Over Three Decades. *Trop. Conserv. Sci.* **2019**, *12*, 1–15. [[CrossRef](#)]
54. Luca, C. Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS. *J. Open Source Softw.* **2021**, *6*, 3172. [[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.