

Article

Integrating Multitemporal Sentinel-1/2 Data for Coastal Land Cover Classification Using a Multibranch Convolutional Neural Network: A Case of the Yellow River Delta

Quanlong Feng ^{1,2}, Jianyu Yang ^{2,*}, Dehai Zhu ², Jiantao Liu ³, Hao Guo ²,
Batsaikhan Bayartungalag ⁴ and Baoguo Li ²

¹ College of Resources and Environmental Sciences, China Agricultural University, Beijing 100193, China; fengql@cau.edu.cn

² College of Land Science and Technology, China Agricultural University, Beijing 100083, China; zhudehai@263.net (D.Z.); guohaolys@cau.edu.cn (H.G.); libg@cau.edu.cn (B.L.)

³ School of Surveying and Geo-Informatics, Shandong Jianzhu University, Jinan 250101, China; liujiantao_1981@hotmail.com

⁴ Research Center for Ecology and Sustainable Development, Mongolian University of Science and Technology, Ulaanbaatar 14191, Mongolia; b.bayartungalag@gmail.com

* Correspondence: ycyjyang@cau.edu.cn; Tel.: +86-10- 62737554

Received: 20 March 2019; Accepted: 25 April 2019; Published: 28 April 2019



Abstract: Coastal land cover classification is a significant yet challenging task in remote sensing because of the complex and fragmented nature of coastal landscapes. However, availability of multitemporal and multisensor remote sensing data provides opportunities to improve classification accuracy. Meanwhile, rapid development of deep learning has achieved astonishing results in computer vision tasks and has also been a popular topic in the field of remote sensing. Nevertheless, designing an effective and concise deep learning model for coastal land cover classification remains problematic. To tackle this issue, we propose a multibranch convolutional neural network (MBCNN) for the fusion of multitemporal and multisensor Sentinel data to improve coastal land cover classification accuracy. The proposed model leverages a series of deformable convolutional neural networks to extract representative features from a single-source dataset. Extracted features are aggregated through an adaptive feature fusion module to predict final land cover categories. Experimental results indicate that the proposed MBCNN shows good performance, with an overall accuracy of 93.78% and a Kappa coefficient of 0.9297. Inclusion of multitemporal data improves accuracy by an average of 6.85%, while multisensor data contributes to 3.24% of accuracy increase. Additionally, the featured fusion module in this study also increases accuracy by about 2% when compared with the feature-stacking method. Results demonstrate that the proposed method can effectively mine and fuse multitemporal and multisource Sentinel data, which improves coastal land cover classification accuracy.

Keywords: convolutional neural networks; land cover classification; data fusion; Sentinel

1. Introduction

Coastal regions play an important role in social and economic development around the globe [1–5]. According to previous studies [2], about 24% of the world’s population lives in coastal areas. Meanwhile, coastal regions are home to many valuable wetland ecosystems, which perform various functions beneficial to the sustainability of human society, including flooding control, water quality improvement,

biodiversity conservation, maintaining the supply of fisheries and other resources, etc. [2–5]. However, due to anthropogenic activities and global climate change, coastal regions have experienced rapid land cover changes over the last few decades [4,5]. Therefore, accurate and timely monitoring of coastal regions by means of remote sensing is of great significance to regional, sustainable development.

In fact, accurate land cover classification of complex coastal regions is a challenging task [2,6–8]. The challenges are mainly two-fold. On the one hand, the highly fragmented landscape of coastal regions leads to large variations in the shape and scale of land objects, which increases the interclass variability and decreases the intraclass similarity. On the other hand, some vegetation classes (e.g., grassland and cropland) may have overlapping spectral reflectance at peak biomass, which also raises difficulties in accurate classification.

Many studies have been conducted on accurate coastal land cover classification [2,6–8]. In coastal areas, although some crops and natural vegetation share similar spectral features during peak growing season, they may have different seasonal variations and temporal characteristics. Therefore, the inclusion of multitemporal remote sensing data could improve classification accuracy when compared with monotemporal data alone. Davranche et al. [6] used multiseasonal SPOT-5 imagery and decision trees for coastal wetland classification in southern France. Yang et al. [7] adopted seasonal optical imagery for coastal land cover classification and demonstrated that combining multiseasonal images considerably improves classification accuracy over any single-date classification. In our previous work [8], we also utilized multitemporal Landsat data to monitor cropland dynamics of the Yellow River Delta and justified the role of multitemporal data in classification.

Meanwhile, because of the availability of diverse remote sensors, researchers have started to integrate multisensor data for better classification of coastal areas [9–13]. Specifically, fusion of optical and radar data has been widely studied [9–13]. Optical images mainly contain information regarding reflectance and emissivity characteristics of land surfaces [9], while radar data are associated with the structural, textural, and dielectric properties of land objects [10]. Therefore, integration of optical and radar data can complement each other, resulting in an improved coastal land cover classification. Rodrigues et al. [9] used multisensor data from Landsat-7 and RADARSAT-1 to identify and map tropical coastal wetlands in the Amazon of northern Brazil. Beijma et al. [10] investigated the uses of multisource airborne radar and optical data to map natural coastal salt marsh vegetation habitats.

Since the successful implementation of the European Copernicus program created by the European Space Agency (ESA), Sentinel-1 radar data and Sentinel-2 optical data are now available via open access, providing new insights for remote sensing applications, especially for large-scale environmental monitoring [14–19]. For instance, Hird et al. developed a workflow for large-area probabilistic wetland mapping based on Google Earth Engine (GEE) and Sentinel-1 and 2 data [14]. Mahdianpari et al. also adopted GEE and multisource Sentinel data to generate the first detailed (category-based) provincial-level wetland inventory map [15]. Therefore, we are highly interested in integrating multitemporal and multisensor Sentinel data for accurate coastal land cover classification.

In addition, all the above studies are based on handcrafted features and conventional machine-learning classifiers, which may fail in obtaining high-level features of complex heterogeneous coastal landscapes. Deep learning [20], on the other hand, has the ability to discover informative features with multiple levels of representation and has achieved an astonishing performance in computer vision applications [21–26], such as image classification [21], object detection [23], and semantic segmentation [24]. Recently, deep learning, especially deep convolutional neural networks (CNNs), has also been successfully applied in many remote sensing applications [27–37]. Rezaee et al. [34] applied a pre-trained AlexNet [21] for wetland mapping using monotemporal optical imagery. Rußwurm et al. [35] utilized sequential recurrent encoders and multitemporal Sentinel-2 optical data for land cover classification, which achieved state-of-the-art classification accuracies. Ji et al. [36] proposed a three-dimensional (3D) CNN for crop classification with multitemporal remote sensing images and concluded that a 3D CNN was suitable in characterizing dynamics of crop growth. Mahdianpari et al. [38] investigated state-of-the-art deep learning models for classification of complex wetland

classes and indicated that InceptionResNetV2, ResNet50, and Xception were distinguished as the top three models.

Despite improvements made by deep learning in the remote sensing field, two challenges of coastal land cover classification mentioned above still remain and need to be solved. In the context of deep learning, the two issues can be revisited as follows: (1) how to build a concise and effective deep learning model that accounts for variations in shapes and scales in fragmented coastal regions, and (2) how to design a fusion mechanism that adaptively fuses multitemporal and multisensor remote sensing data.

To address these issues, this study proposes a multibranch convolutional neural network (MBCNN) for coastal land cover classification using multitemporal and multisensor Sentinel data. First, a single-branch CNN is proposed to extract representative features from each monotemporal and single-sensor Sentinel datum. A deformable multiscale residual block is utilized in the single-branch CNN to account for shape and scale variations. Afterwards, multiple single-branch CNNs are integrated through an adaptive fusion module, which are inspired by squeeze-and-excitation networks [25], to predict the final land cover category. The selected study region is the Yellow River Delta, which is the largest natural delta of China and home to abundant coastal wetlands [39–42].

The rest of the paper is organized as follows. Section 2 introduces the study area and the dataset used. Section 3 presents the architecture and training details of the proposed multibranch neural network. Section 4 shows the experimental results and discussion, while Section 5 provides the main conclusions and suggestions for future work.

The contributions of this study are mainly two-fold. (1) We have designed a concise yet effective deep learning model for coastal land cover classification, which adopts deformable convolutional layers to account for variations of scales and shapes of coastal landscapes. (2) We have proposed a feature-level fusion module based on squeeze-and-excitation networks for multitemporal and multisensor Sentinel data fusion to boost coastal land cover classification accuracy.

2. Study Area and Dataset

2.1. Study Area

The Yellow River Delta is the largest natural delta of China and is home to many coastal wetlands (Figure 1). In this study, the Yellow River Delta refers to the Yellow River Delta National Nature Reserve [39], which is located northeast of Dongying City in Shandong Province in China. Due to deposition of abundant sediments transported by the Yellow River, the newly created wetland has increased by 30 km² per year, making the Yellow River Delta one of the fastest growing sedimentation areas around the globe [40,41].

The Yellow River Delta belongs to a temperate, continental monsoon climate, which has a hot, humid summer and a cold, dry winter. It has an annual temperature of about 11.9 °C and an annual precipitation of about 640 mm [41]. The natural vegetation includes reed, tamarisk, *Suaeda*, and *Robinia*, while the main crops include rice, lotus, corn, winter wheat, and cotton [42].

A field survey was conducted in July 2018. A total of 163 sampling sites were visited. Land cover types, photographs, and global positioning system (GPS) locations were recorded for each sampling site. According to the field survey and previous studies [8,39–42], there were 11 land cover categories in this study: forest, grassland, salt marsh, shrubs, tidal flat, bare soil, clear water, turbid water, irrigated farmland, dry farmland, and built up. Landscape descriptions for each land cover category are shown in Table 1. Training and testing samples were derived from remote sensing images through visual inspection based on sampling site GPS locations and recorded land cover types. The spatial distribution of training and testing samples are depicted in Figure 2a,b, respectively. The number of training and testing samples (in pixels) are also shown in Table 1. To make the accuracy assessment more objective and convincing, we doubled the number of testing samples in relation to training samples.

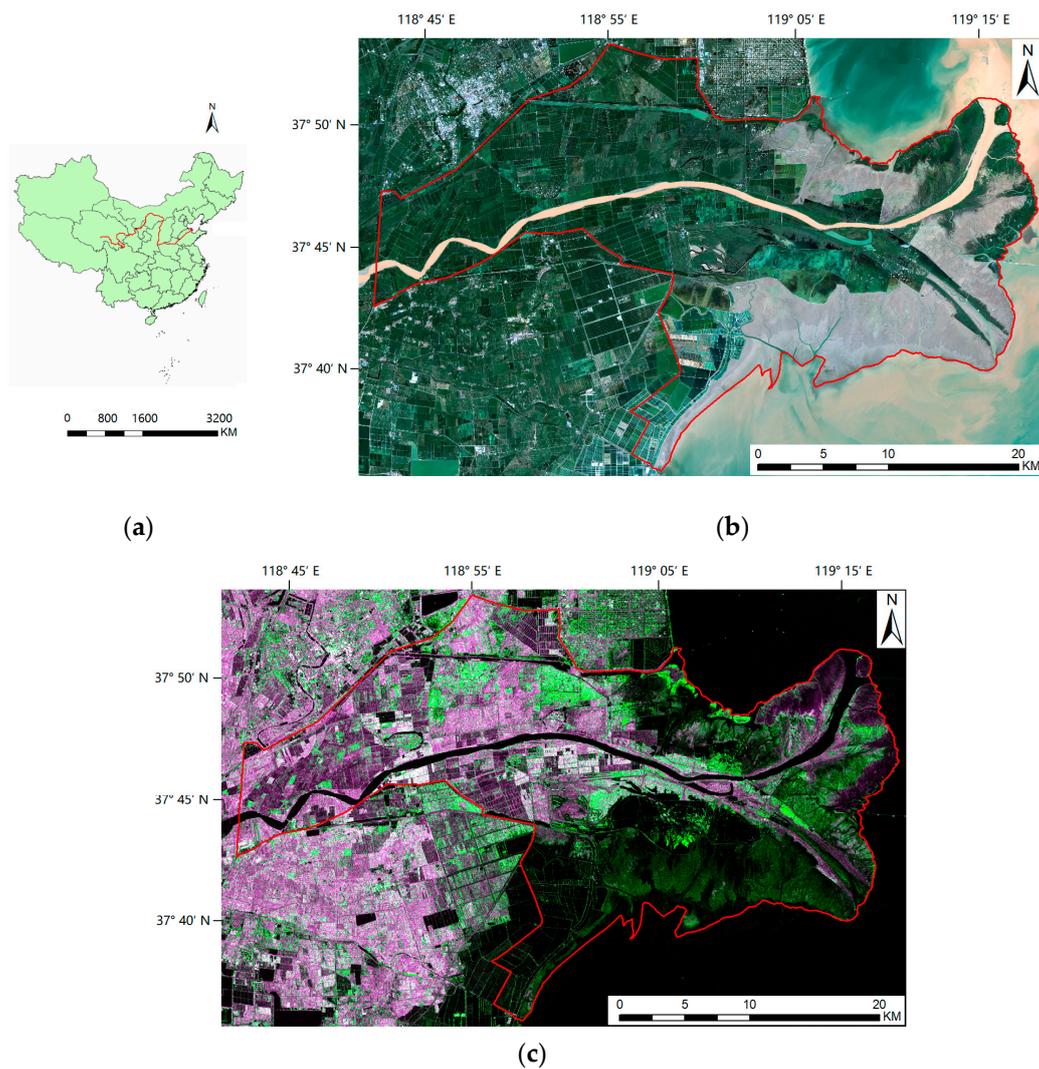


Figure 1. Study area. (a) Yellow River of China; (b) True color image of Sentinel-2 on 25 August 2018; and (c) False color image of Sentinel-1 (Red: VH, Green: VV, Blue: VH) on 26 August 2018.

Table 1. Classification scheme of the Yellow River Delta.

No.	Land Cover	Description	Training	Testing
1	Forest	Broad-leaved trees, mainly <i>Robinia</i> and willow	250	500
2	Grassland	Vegetated areas where reed is dominant	500	1000
3	Salt marsh	Vegetated areas where sea-blite is dominant	150	300
4	Shrubs	Sparsely vegetated shrubs, mainly tamarisks	75	150
5	Tidal flat	Non-vegetated foreshore areas	250	500
6	Bare soil	Non-vegetated bare land, mainly saline and alkaline land	250	500
7	Clear water	Clear water bodies, including rivers, reservoirs, aquaculture, and brine ponds	250	500
8	Turbid water	Turbid water bodies, mainly the Yellow River	150	300
9	Irrigated farmland	Including irrigated farmland, mainly rice and lotus	500	1000
10	Dry farmland	Including non-irrigated farmland, mainly winter wheat, corn, cotton, and soybean	200	500
11	Built up	Artificial surfaces including residential areas, factories, and oil fields	150	300

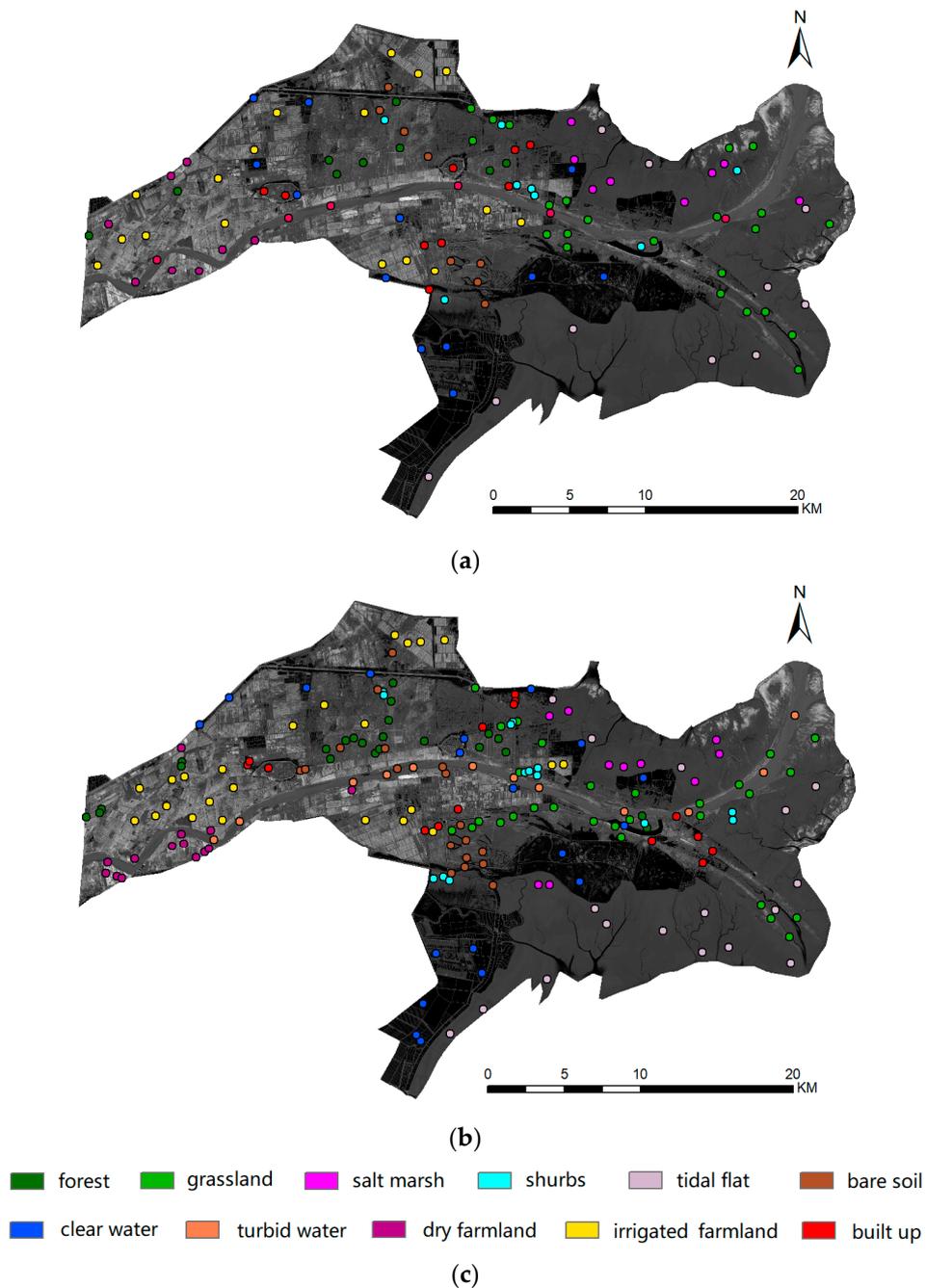


Figure 2. Spatial distribution of (a) Training samples; (b) Testing samples; and (c) Legend.

2.2. Dataset Used

Because of the availability of Sentinel-1 and Sentinel-2 data, these data were integrated for various applications such as vegetation mapping [16], soil moisture monitoring [17], and crop classification [18]. In this study, both multitemporal and multisensor Sentinel data over an entire growing season were utilized for coastal land cover classification (Table 2).

Table 2. Multitemporal Sentinel-1/2 data used in this study.

	Season	Date	Source	Product	Incidence Angle
T1	Spring	16 April 2018	S1	Level-1 GRD	38.01°
		17 April 2018	S2	Level-1C	–
T2	Summer	3 June 2018	S1	Level-1 GRD	38.01°
		6 June 2018	S2	Level-1C	–
T3	Summer	26 August 2018	S1	Level-1 GRD	38.01°
		25 August 2018	S2	Level-1C	–
T4	Autumn	25 October 2018	S1	Level-1 GRD	38.01°
		24 October 2018	S2	Level-1C	–

Note. S1: Sentinel-1; S2: Sentinel-2; and GRD: ground range detected.

Specifically, radar datasets were obtained from Sentinel-1 Level-1 ground range detected (GRD) images with a spatial resolution of 10 m × 10 m [16]. The synthetic aperture radar (SAR) onboard Sentinel-1 operates at the C-band with a revisit time of six days [17]. Preprocessing of Sentinel-1 SAR data was implemented using Sentinel-1 Toolbox provided by the ESA [18], including radiometric calibration, speckle noise reduction, and terrain correction, which outputs geo-coded backscattering coefficients of VV (for vertical transmit and vertical receive) and VH (for vertical transmit and horizontal receive) polarizations.

Optical datasets were obtained from Sentinel-2 MSI Level-1C products under cloud-free conditions. Sentinel-2 MSI has 13 bands ranging from 443 to 2190 nm and a spatial resolution from 10 to 60 m with a revisit time of five days [18]. In this study, only bands at 10 m (Bands 2, 3, 4, and 8) and 20 m (Bands 5, 6, 7, 8A, 11, and 12) resolutions were selected. Sen2Cor [19] was used to perform atmospheric correction to get the bottom-of-atmosphere (BOA) 2A product. In order to co-register with Sentinel-1 SAR data, all bands at 20 m resolution of Sentinel-2 were resampled to 10 m using a bilinear interpolation method.

3. Methods

3.1. Overview of a Multibranch Convolutional Neural Network (CNN)

Figure 3 shows the overview of the proposed multibranch CNN model for coastal land cover classification.

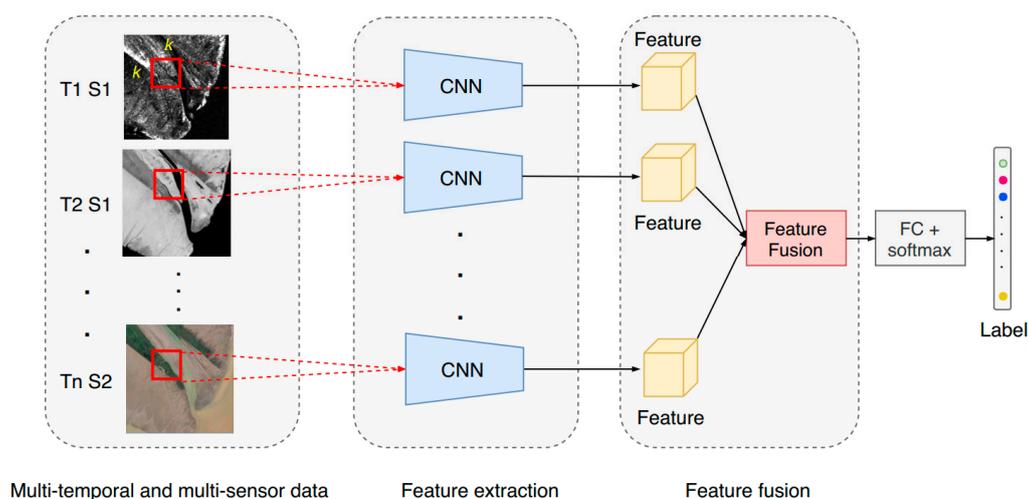


Figure 3. An overview of the proposed multibranch convolutional neural network (CNN).

As shown in Figure 3, the multibranch CNN model had two major components: (1) a feature extraction module based on single-branch CNN, and (2) a feature fusion module to aggregate the

extracted features for final land cover classification. Each single-branch CNN had the same network structure. Deformable convolutions [23,24] and multiscale residual blocks [22] were introduced to model the land surface with various shapes and scales. The extracted features from each branch were fed into an adaptive feature fusion module, through which the multitemporal and multisensor data were effectively synthesized for final classification.

3.2. Brief Introduction of CNNs

To better understand our proposal, a brief introduction of CNN is provided in this section. Generally, a typical CNN architecture is alternatively stacked by convolutional layers, pooling layers, and fully connected layers [29].

3.2.1. Convolutional Layers

Convolutional layers are of great significance in a CNN. High-level representative features can be extracted through the stacking of multiple convolutional layers. The input into a convolutional layer is a feature map x with a size of $m \times n \times c$, where $m \times n$ denotes the spatial size of the feature map, while c is the number of input channels. Supposing the convolutional layer consists of k filters, the output would be an $m' \times n' \times k$ feature map with k channels and a spatial size of $m' \times n'$. The i th output feature map of the convolutional layer, y_i , can be expressed as follows.

$$y_i = w_i * x + b_i, \quad (1)$$

where w_i and b_i denote the weights and bias of the i th filter, and $*$ is the direct convolutional operator. Afterwards, a nonlinear activation function (e.g., the rectified linear unit [43]) is usually applied to the output feature map to increase the nonlinear learning ability of the network.

3.2.2. Pooling Layers

Pooling layers are used to generalize the convolved features through down-sampling. The spatial size of the input feature map is reduced after a pooling operation, which decreases the number of parameters and computational complexity. Commonly used pooling layers include max pooling and average pooling, which use the maximum or average operator to extract values for local spatial regions, respectively.

3.2.3. Fully Connected Layers

The role of fully connected layers is to combine all input features by reshaping them into an N -dimensional vector. Simple logistic regression is used by the fully connected layers. Finally, the extracted feature vector is fed into the softmax classifier [43] to generate the probability distribution.

3.3. Single-Branch CNN for Feature Extraction

Accurate coastal land cover classification requires a set of well-established and representative features. In this study, to account for complex and fragmented coastal landscapes, we first proposed a single-branch CNN based on both deformable convolutions and multiscale residual blocks (Figure 4).

Figure 4 illustrates that the input of the proposed single-branch CNN is an image patch centered on the labeled pixel with a size of $k \times k \times c$, where k is the patch size and c is the number of channels. The proposed network consisted of several convolutional layers, max pooling layers, and deformable multiscale residual blocks; detailed information is listed in Table 3.

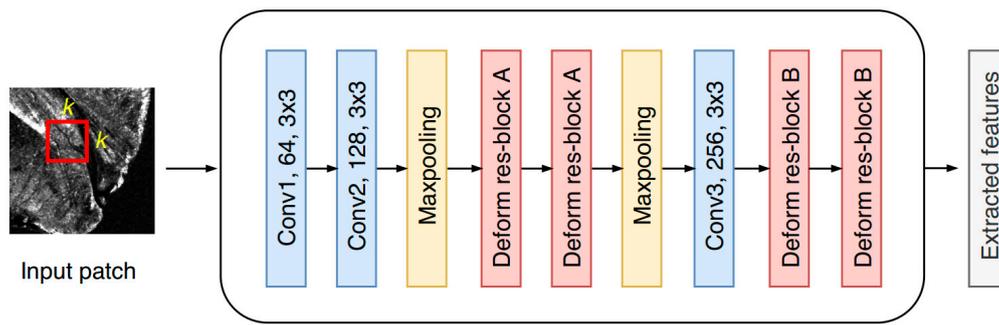


Figure 4. Architecture of the proposed single-branch CNN.

Table 3. Detailed information of the single-branch CNN.

Layer Name	Input Size	Output Size	Kernel Size	Filter Number	Stride
Input	$11 \times 11 \times 10$	–	–	–	–
Conv1	$11 \times 11 \times 10$	$11 \times 11 \times 64$	3×3	64	1
Conv2	$11 \times 11 \times 64$	$11 \times 11 \times 128$	3×3	128	1
Max-pooling1	$11 \times 11 \times 128$	$6 \times 6 \times 128$	–	–	2
Deform res-block A1	$6 \times 6 \times 128$	$6 \times 6 \times 128$	–	–	–
Deform res-block A2	$6 \times 6 \times 128$	$6 \times 6 \times 128$	–	–	–
Max-pooling2	$6 \times 6 \times 128$	$3 \times 3 \times 128$	–	–	2
Conv3	$3 \times 3 \times 128$	$3 \times 3 \times 256$	3×3	256	1
Deform res-block B1	$3 \times 3 \times 256$	$3 \times 3 \times 256$	–	–	–
Deform res-block B2	$3 \times 3 \times 256$	$3 \times 3 \times 256$	–	–	–

The deformable multiscale residual block was inspired by both deformable convolution [23,24] and a multiscale residual block [22]. Specifically, the multiscale residual block was borrowed from Bulat et al. [22], which had the merits of extracting hierarchical and multiscale features and improving gradient flow at the same time. By introducing deformable convolution into the multiscale residual block, the receptive field and sampling locations were trained to be adaptive to the shapes and scales of land objects, which enabled extraction of robust and representative features. Figure 5 shows the structure and parameters of the deformable multiscale residual blocks.

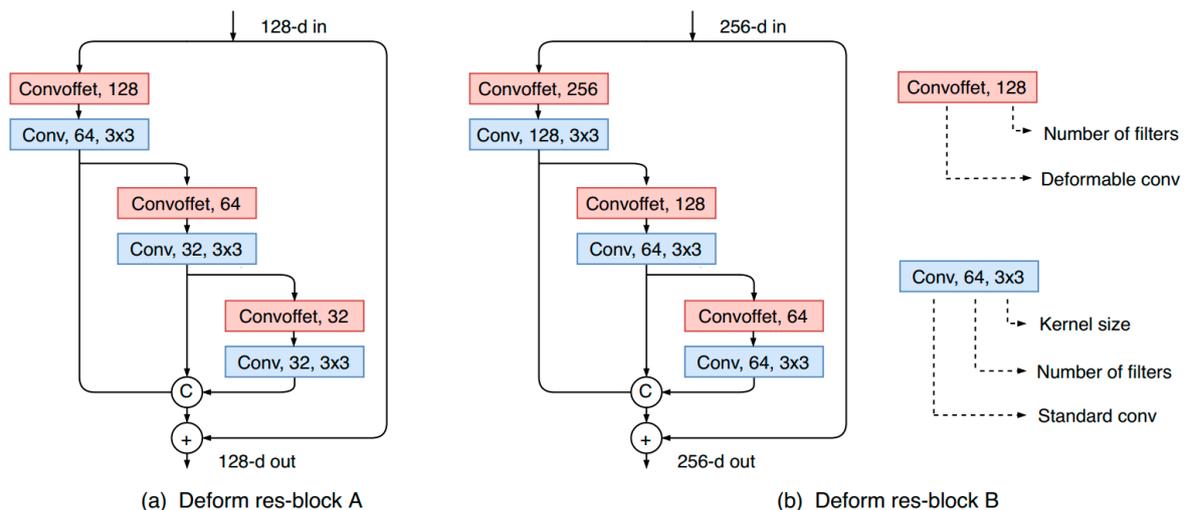


Figure 5. Architecture of the deformable multiscale residual blocks.

The mechanism of deformable convolution is illustrated in Figure 6. The offset field was derived from input feature maps, and the deformable kernel had the same resolution as the current convolutional

layer [23]. Both the kernels and offsets were learned simultaneously during the training process. Therefore, the output feature y at location p_0 can be formalized as follows:

$$y(p_0) = \sum w(p_i) * x(p_0 + p_i + \Delta p_i), \tag{2}$$

where w refers to the weights of the sampled points, x refers to the input feature map, p_i means the i th location, and Δp_i represents the offset to be learned [23,24].

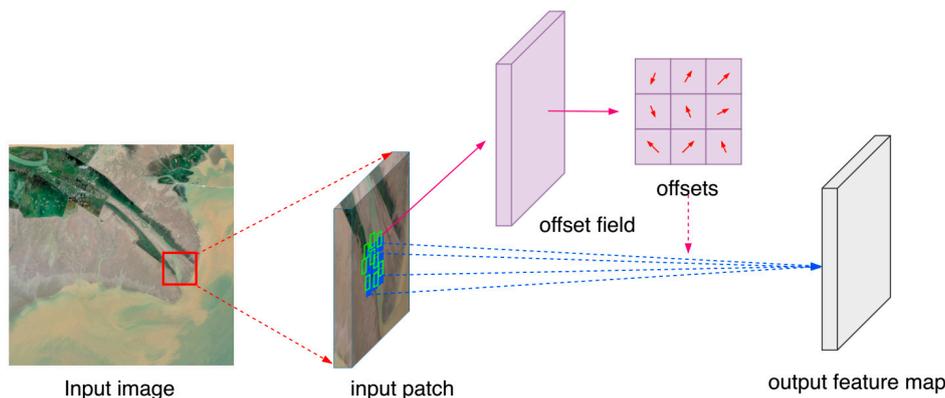


Figure 6. Illustration of a 3 × 3 deformable convolution.

Additionally, a series of experiments were done to find the optimal patch size k from 9 to 29. It was found that the best classification accuracy was achieved when $k = 11$.

3.4. Adaptive Feature Fusion

The sequence of features extracted from each single-source (i.e., both single-date and single-sensor) Sentinel dataset was utilized in the proposed feature fusion module to make the final land cover prediction. As for the fusion method, many previous studies [29,33] simply stacked and concatenated all the input features without considering the importance of each feature. Inspired by squeeze-and-excitation networks (SE Nets) [25] and our previous work [30], this study proposed a fusion mechanism for feature aggregation of multibranch CNNs, which took the importance of each feature into consideration (Figure 7).

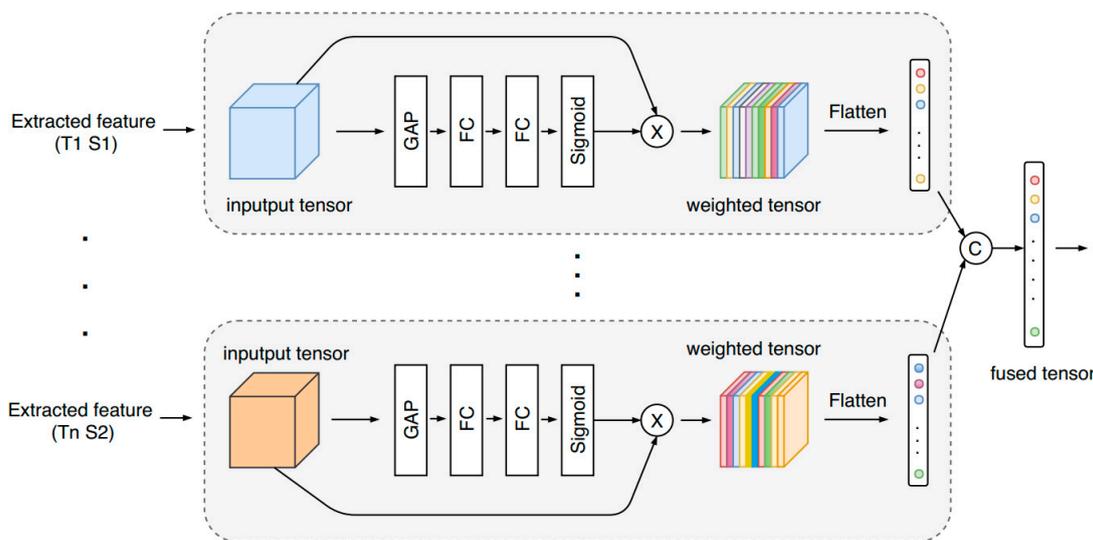


Figure 7. Structure of the adaptive feature fusion module. GAP: global average pooling; FC: fully connected; \times : element-wise multiplication; and C: concatenation.

As shown in Figure 7, the feature fusion module was used to recalibrate (or reweight) all the features extracted from each single-branch CNN through a series of squeeze-and-excitation (SE) blocks [25]. First, any input features from each branch were passed through a global average pooling (GAP) layer to generate a channel descriptor. Next, the channel-specific weight was learned with two successive fully connected layers and a sigmoid layer. After all the features from each branch were reweighted, informative features were emphasized and less useful ones were suppressed, which provided a more effective and rational method to achieve feature-level fusion of multitemporal and multisensor Sentinel data.

Finally, all the reweighted features were flattened and concatenated to generate the fused feature vectors. Then, the fused features were fed into a fully connected layer and a *softmax* layer to calculate conditional probabilities of each land cover category.

3.5. Details of Network Training

Data augmentation was utilized in this study to overcome the limited amount of training data. All the training patches were flipped up and down, left and right, and rotated 90°, 180°, and 270° to enlarge the training datasets.

All the parameters of the MBCNN should be trained. Specifically, all the weights were initialized using He normalization [43], while biases were initialized by zero. As for the optimization method, Adam [44] was utilized with a starting learning rate of 10^{-5} . An early-stop strategy was used to select the best model. Only the model with the minimum validation loss was saved.

Focal loss [26] was adopted instead of cross-entropy loss to further boost classification performance. Focal loss played the role of online hard example mining, which down-weighted loss assigned to the well-classified examples and prevented the vast number of easy examples from overwhelming the classifier during training.

In this study, about 90% of training samples were randomly selected to optimize the parameters of the proposed model. The remaining 10% of training samples were used as a validation set to evaluate classification performance during the training process. The testing set was only used to calculate final overall accuracy and the confusion matrix after the model was well trained.

The proposed MBCNN was trained with the TensorFlow library [45] on the Ubuntu 16.04 operation system with Intel CORE i7-7800 @ 3.5 GHz CPU and an NVIDIA GTX TitanX GPU with 12 GB memory.

3.6. Accuracy Assessment

To justify the effectiveness of the proposed method, both visual evaluation and a confusion matrix were utilized in this study. Visual evaluation was used to check obvious classification errors, while a confusion matrix derived from the testing samples was used to quantitatively evaluate classification performance through the following metrics: overall accuracy (OA), producer accuracy (PA), user accuracy (UA), and Kappa coefficient.

4. Results and Discussion

4.1. Results of Coastal Land Cover Classification

Figure 8 illustrates classification results of the Yellow River Delta using the proposed multibranch CNN and multitemporal, multisensor Sentinel data. From the perspective of visual inspection, the results showed good visual effect, and the spatial distributions of each classified land cover were close to field survey records. Moreover, few obvious omission and commission errors could be found in Figure 8, which also justified the effectiveness of the proposed method.

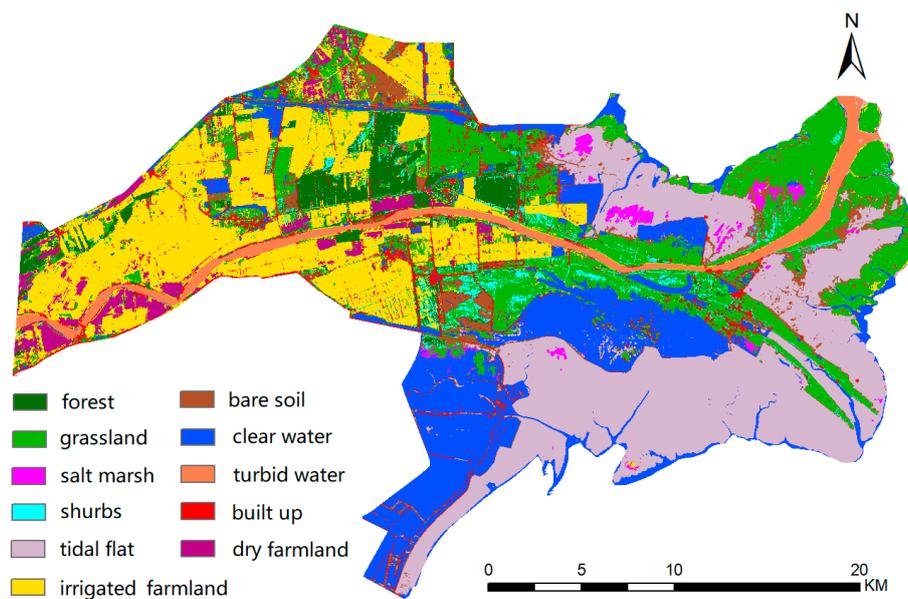


Figure 8. Classification map of the Yellow River Delta generated by the multibranch CNN.

To quantitatively evaluate performance of the proposed method, the confusion matrix, OA, and Kappa coefficient were calculated from the testing samples. The results are shown in Table 4.

Table 4. Confusion matrix of the proposed method.

Class	Ground Truth											UA%
	1	2	3	4	5	6	7	8	9	10	11	
1	445	27	0	10	0	0	0	0	3	0	0	91.75
2	40	931	0	14	0	0	0	0	44	0	8	89.78
3	0	0	294	1	0	3	0	0	0	0	0	98.66
4	0	38	0	99	0	5	0	0	16	0	5	60.74
5	0	0	5	26	500	29	12	0	0	0	0	87.41
6	8	0	1	0	0	453	14	0	0	0	0	95.17
7	0	0	0	0	0	0	474	0	0	0	0	100
8	0	0	0	0	0	0	0	300	0	0	0	100
9	0	3	0	0	0	2	0	0	932	6	0	98.83
10	0	1	0	0	0	0	0	0	4	490	0	98.99
11	7	0	0	0	0	8	0	0	1	4	287	93.49
PA%	89.00	93.10	98.00	66.00	100	90.60	94.80	100	93.20	98.00	95.67	
OA	93.78%											Kappa 0.9297

Note. 1: forest; 2: grassland; 3: salt marsh; 4: shrubs; 5: tidal flat; 6: bare soil; 7: clear water; 8: turbid water; 9: irrigated farmland; 10: dry farmland; 11: built up; PA: producer accuracy; UA: user accuracy; and OA: overall accuracy.

Table 4 indicated that the proposed multibranch CNN achieved good performance with an OA of 93.78% and a Kappa coefficient of 0.9297. Almost every class demonstrated producer accuracy of more than 89%, except for shrubs, whose PA was only 66.00%. Several shrub pixels were misclassified as forest, grassland, and tidal flat. This was understandable, because the radar backscattering properties of the shrub land cover category were similar to those of the forest category. Meanwhile, shrubs (mainly tamarisks) were sparsely distributed in the coastal wetlands surrounded by tidal flats, which caused spectral confusion between shrubs and tidal flat categories. In addition, because of the limited spatial resolution, there were hardly no pure shrub pixels, leading to spectral confusion between shrubs and grassland categories, which could also account for classification errors.

In addition, other classification errors mainly occurred between forest, grassland, and irrigated land categories as well as bare soil and tidal flat categories. This was because of the similarity of spectral and backscattering characteristics between these land cover types.

4.2. Impact of Multisensor Data on Classification

As stated earlier, inclusion of both optical and radar data would be expected to improve the accuracy of coastal land cover classification. In this section, comparison between single-sensor and multisensor classifications will be discussed. Specifically, experiments were performed for the following cases:

- (1) radar-only classification: using only multitemporal radar data from Sentinel-1 for classification;
- (2) optical-only classification: using only multitemporal optical data from Sentinel-2 for classification;
- (3) feature-stacking classification: using multitemporal radar and optical data and feature stacking for classification; and
- (4) proposed MBCNN model: using multitemporal radar and optical data and the proposed MBCNN for classification.

The classification maps for each experiment are illustrated in Figures 9 and 10.

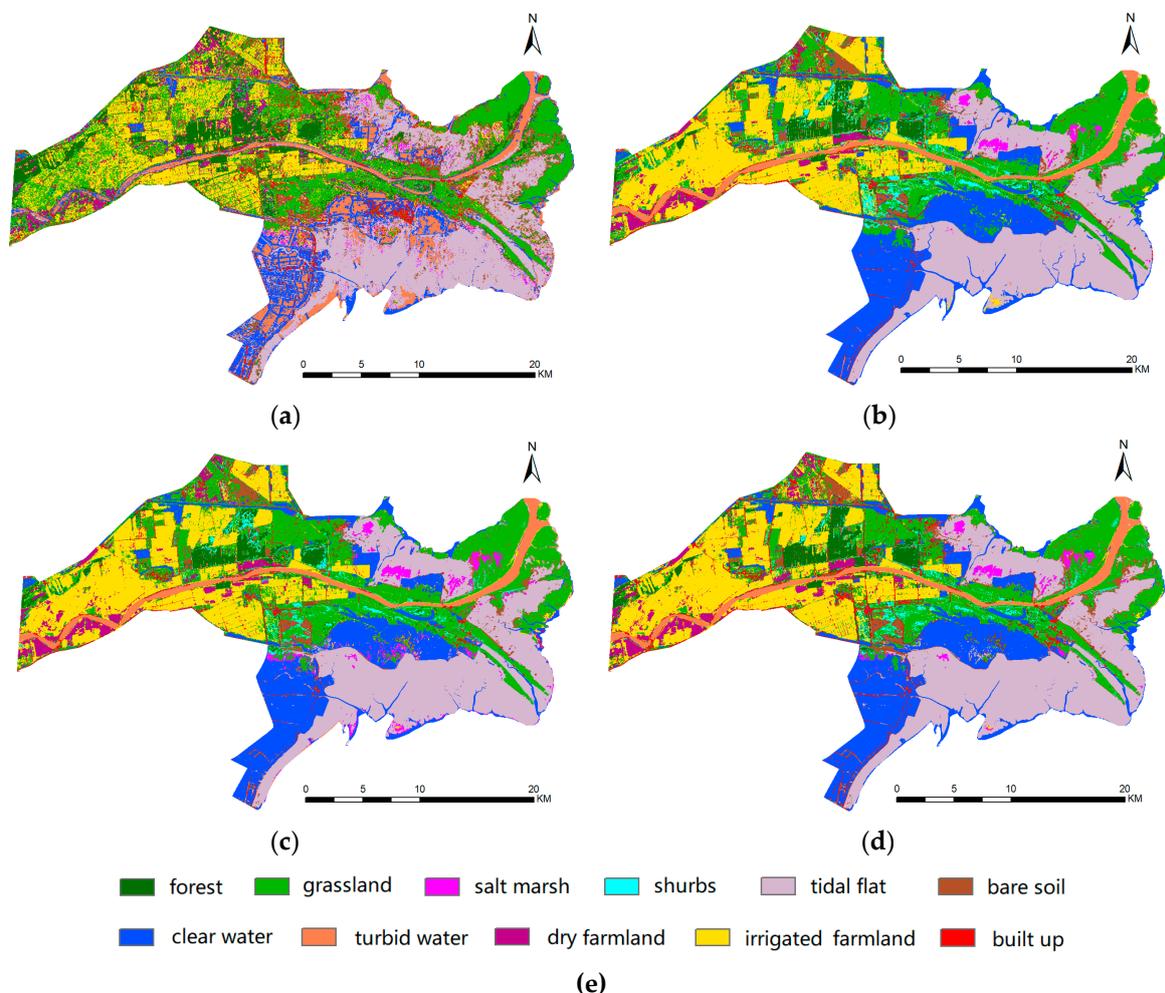


Figure 9. Classification results for the (a) radar-only, (b) optical-only, (c) feature-stacking, and (d) proposed multibranch CNN (MBCNN) classification models; (e) legend.

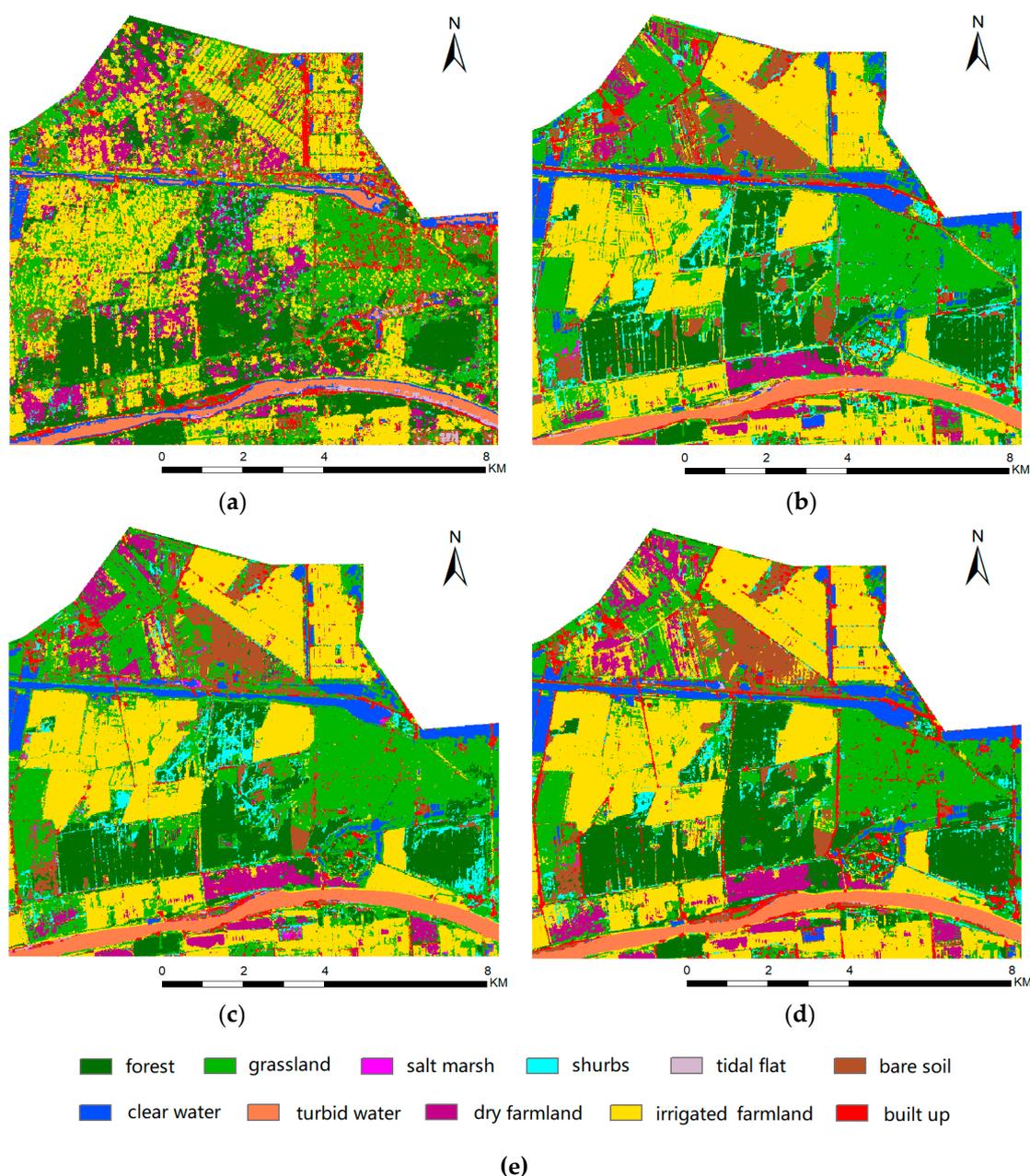


Figure 10. Details of the classification results for (a) radar-only, (b) optical-only, (c) feature-stacking, and (d) proposed MBCNN classification models; (e) legend.

Both Figures 9 and 10 illustrated that inclusion of multisensor data yielded a better classification map with fewer errors when compared with single-sensor classification. Meanwhile, it was difficult to get an accurate classification map by using Sentinel-1 radar data alone. There were many errors among various land cover categories, especially between grassland and irrigated farmland as well as between shrubs and grassland. Nonetheless, using Sentinel-2 optical data alone could achieve a much better classification map. Similar spatial patterns were found among classification maps yielded by optical-only and feature-stacking methods and the proposed MBCNN model.

In addition, Figure 10 indicated that the proposed MBCNN could effectively reduce classification errors between forest and shrubs as well as between dry farmland and forest when compared with optical-only and feature-stacking methods.

Table 5 shows detailed class-level classification accuracies (i.e., producer accuracy) for each experiment. It indicated that the proposed MBCNN achieved the highest classification accuracy with

an OA of 93.78% and a Kappa of 0.9297, which verified the effectiveness of the proposed model. Radar-only classification had the lowest OA of 64.00%, which was consistent with Figures 9 and 10. The following land cover categories had low accuracies in radar-only single-sensor classification: salt marsh, shrubs, and bare soil. Meanwhile, optical-only classification demonstrated better performance than radar-only classification. This was mainly because Sentinel-2 could provide distinctive spectral characteristics, which were essential in separating different coastal land cover categories, especially with respect to confusing vegetation types.

Table 5. Class-level classification accuracy.

No.	Class Name	Radar-only (%)	Optical-only (%)	Feature-Stacking (%)	Proposed (%)
1	Forest	70.00	85.40	84.80	89.00
2	Grassland	76.70	90.00	96.20	93.10
3	Salt marsh	14.00	85.67	97.00	98.00
4	Shrubs	2.00	61.33	45.33	66.00
5	Tidal flat	61.80	100	100	100
6	Bare soil	48.60	81.20	77.80	90.60
7	Clear water	74.00	94.20	96.00	94.80
8	Turbid water	89.33	100	100	100
9	Irrigated farmland	66.40	91.30	91.30	93.20
10	Dry farmland	64.60	95.40	96.00	98.00
11	Built up	71.00	94.00	90.33	95.67
	OA (%)	64.00	90.54	91.50	93.78
	Kappa	0.5919	0.8932	0.9037	0.9297

Table 5 also indicated that the synthetic use of Sentinel-1 and Sentinel-2 data led to an increase in classification accuracy for almost every coastal land cover category. This was rational, because integration of optical and radar features could enhance between-class separability [10,13]. Compared with Sentinel-2 data alone, inclusion of Sentinel-1 data increased OA by 0.96% and 3.24% through feature stacking and the proposed multibranch CNN, respectively.

The adaptive feature fusion method in this study outperformed the feature-stacking method by increasing OA from 91.50% to 93.78% with an improvement of 2.28%. This was because, when simply stacking features together, the information carried by each feature may not be equally represented [30]. Nonetheless, introduction of a squeeze-and-excitation module can automatically learn the weight of each feature according to its importance, fusing multiple features in a more reasonable and effective way.

Besides, Table 5 also indicated that when using SAR data alone, it was difficult for the classification model to separate shrubs from other land cover types, which meant that image features learned from shrubs were very weak. However, those weak SAR features of shrubs still existed and were enhanced by the adaptive feature fusion method in this paper, which in turn contributed to accuracy improvement when combined with optical data.

4.3. Impact of Multitemporal Data on Classification

The role of multitemporal data in coastal land cover classification should also be verified. In this section, we conducted a series of experiments for monotemporal classification. In each single-date experiment, radar data from Sentinel-1 and optical data from Sentinel-2 were involved. The classification maps and overall accuracy for each single-date dataset are illustrated in Figure 11, Figure 12, and Table 6, respectively.

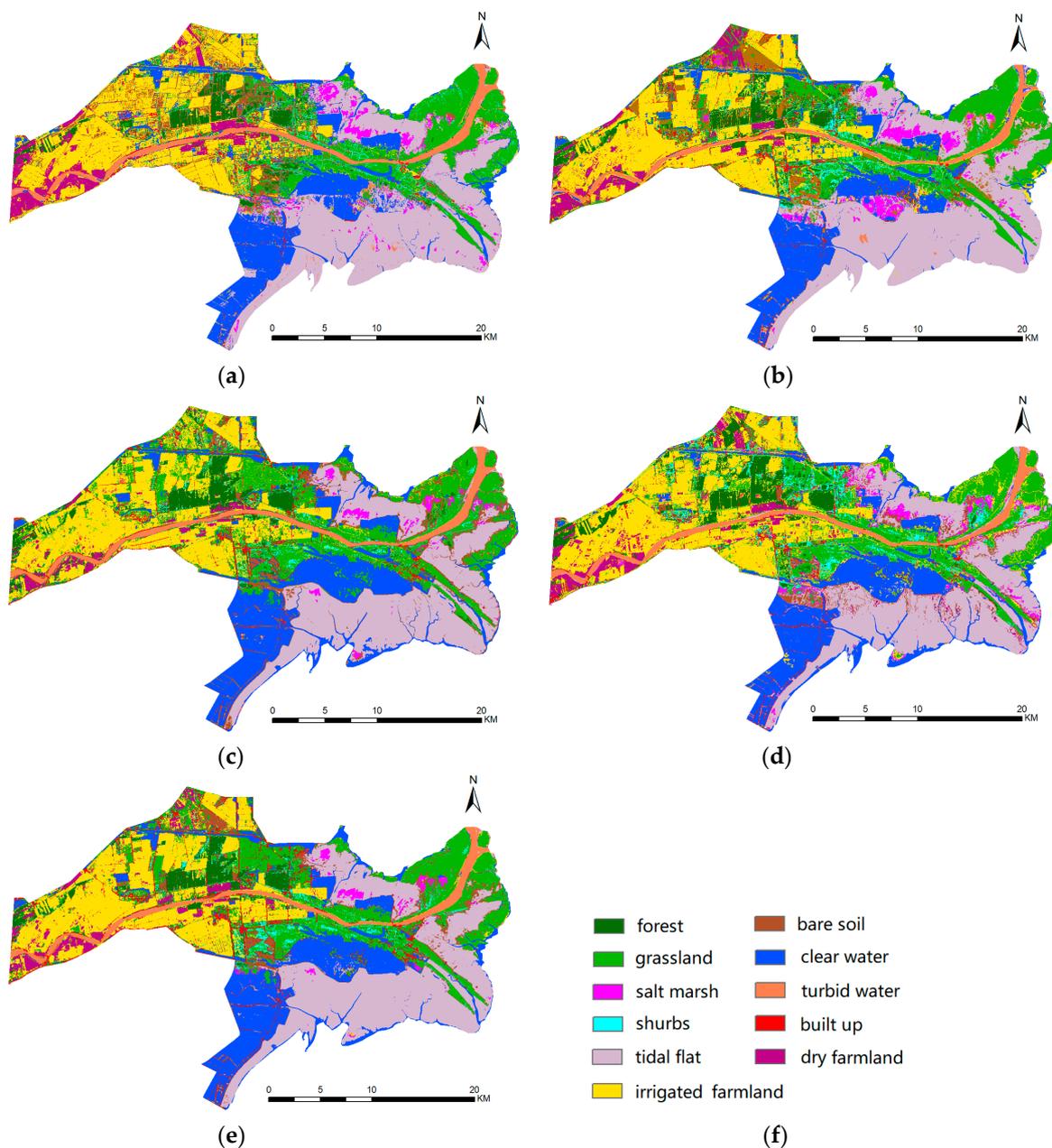


Figure 11. Classification results for (a) T1/2018-04; (b) T2/2018-06; (c) T3/2018-08; (d) T4/2018-10; and (e) multitemporal; (f) legend.

Table 6. Accuracy comparison between mono- and multitemporal classifications.

	T1 (2018.04)	T2 (2018.06)	T3 (2018.08)	T4 (2018.10)	Multitemporal
OA (%)	81.93	88.68	92.63	84.50	93.78
Kappa	0.7957	0.8719	0.9165	0.8248	0.9297

Both Figures 11 and 12 show that when compared with single-date classification, the inclusion of multitemporal data improved classification performance. The multitemporal classification map showed fewer obvious mistakes, especially between forest and shrubs, irrigated farmland and grassland, and dry farmland and bare soil. This was because phenological information conveyed by multitemporal data enhanced separability among different vegetation types [8]. This was in accordance with the quantitative evaluation shown in Table 6. By introducing temporal information, classification accuracy

was boosted by 1.15%–11.85%, with an average increase of 6.85%, which justified the importance of multitemporal data in coastal land cover classification.

Table 6 also indicated that the classification accuracy for date T1 (April 2018) was notably lower than that of other dates, with an OA of 81.93% and a Kappa of 0.7957. This was also consistent with Figures 11 and 12. This was mainly because most of the vegetation, except for winter wheat, started to turn green in April. The differences among vegetation were not distinct from either the spectral or backscattering perspectives, which resulted in low between-class separability and classification performances.

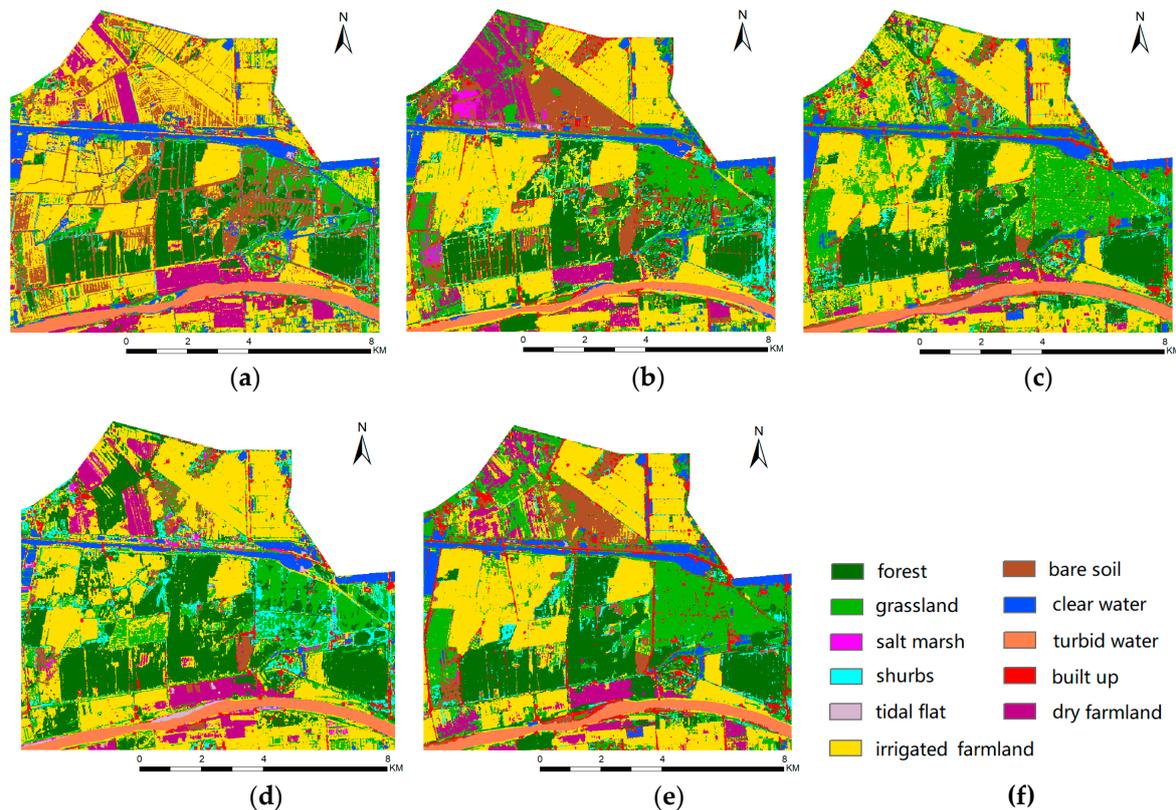


Figure 12. Details of the classification results for (a) T1/2018-04; (b) T2/2018-06; (c) T3/2018-08; (d) T4/2018-10; and (e) multitemporal; (f) legend.

4.4. Impact of Deformable Convolution on Classification

In contrast with previous land cover classification methods based on deep learning [29,30,33–37], we introduced deformable convolution to model fragmented coastal landscapes. To better interpret the impact of deformable convolution for classification, a contrast experiment was conducted in this section. In the experiment, all the deformable convolutional layers in the MBCNN were replaced by standard convolutional layers. Accuracy comparisons between standard and deformable convolution is shown in Table 7.

Table 7. Accuracy comparison between standard and deformable convolutions.

Method	OA (%)	Kappa
Standard convolution	91.69	0.9060
Deformable convolution	93.78	0.9297

Table 7 indicated that when compared with standard convolution, introduction of deformable convolution improved the OA from 91.69% to 93.78% with an increase of 2.09%, which verified the

effectiveness of deformable convolution. In fact, in complex heterogeneous landscapes such as coastal areas, a big challenge in land cover classification is the variations in the shapes and scales of land objects. Because of the fixed kernel shape, standard convolution could not capture these variations, which resulted in an inferior performance. However, by utilizing deformable receptive fields [23,24], which were adaptive to the shape and scale of input remote sensing data, deformable convolution extracted more representative features, showing better performance in complex coastal land cover classification when compared with standard convolution.

4.5. Comparison with Machine Learning Methods

As is known, machine learning-based methods have long been used for land cover mapping in the remote sensing field, such as in maximum likelihood classifier (MLC), random forest (RF), support vector machine (SVM), etc. To further justify the performance of the proposed method, it should be compared with those widely used machine learning methods. Specifically, with respect to RF [46], we involved 200 decision trees with a max depth of 13 and utilized the Gini coefficient [46] as the indicator for feature selection. With respect to SVM [47], we used radial basis function [47] as the kernel function with a gamma [47] of 0.01 and a penalty coefficient C [47] of 100. As for determining the parameters of RF and SVM classifiers, a grid-search method was utilized to find the optimal values. Specifically for RF, the number of trees were set between 100 and 300, while the max depth was between 3 to 15, respectively. For SVM, gamma was set between 0.001 to 0.1, while C had a range of 20 to 200.

In addition, all the above methods were trained and tested using the same training and testing samples as the proposed multibranch CNN in this study to maintain objectiveness. The results of accuracy comparisons are listed in Table 8.

Table 8. Accuracy comparison with machine learning methods.

Method	OA (%)	Kappa
Maximum Likelihood Classifier	74.65	0.7153
Random Forest	84.98	0.8301
Support Vector Machine	87.51	0.8541
Our Multibranch CNN	93.78	0.9297

Table 8 indicated that the traditional machine learning methods showed inferior performance to the proposed method in coastal land cover classification. The proposed multibranch CNN outperformed MLC, RF, and SVM with an increase in OA of 19.13%, 8.80%, and 6.27%, respectively. This was mainly because the proposed deep convolutional neural network could learn high-level and discriminative representations of complex and fragmented coastal landscapes, outperforming machine learning-based methods.

4.6. Comparison with Other Land Cover Classification Methods

Because the main objective of this study was to propose a deep learning-based method for coastal land cover classification, it was necessary to compare our proposed method with other classification methods (Table 9) to further demonstrate both the merits and limitations of the proposed method. It should be noted that because of the differences in the study area, number of training samples, and classified categories, it was difficult to directly compare these methods based on classification accuracies alone. Therefore, we mainly focused on the merits and shortcomings of each method.

Table 9. Overview of recently published land cover/use classification methods.

Approach	Data	Multi temporal	Multi sensor	Model	Accuracy	Number of Classes
This work	S1, S2	Yes	Yes	MBCNN	93.78	11
Rezaee et al. [34]	RE	No	No	AlexNet	94.82	8
Huang et al. [29]	WV-3	No	No	STDCNN	91.25	11
Mahdianpari et al. [38]	RE	No	No	Inception-ResNetV2 et al.	96.17	8
Rußwurm et al. [35]	S2	Yes	No	RNN	90.00	17
Ji et al. [36]	GF-2	Yes	No	3D CNN	94.70	4
Xu et al. [33]	HSI, LiDAR	No	Yes	CNN	87.98	15
Scarpa et al. [37]	S1, S2	Yes	Yes	CNN	–	–

Note. S1: Sentinel-1; S2: Sentinel-2; RE: RapidEye; GF-2: GaoFen-2; WV-3: WorldView-3; HSI: Hyperspectral Image; LiDAR: Light Detection and Ranging; L5: Landsat-5; RAST-2: RADARSAT-2; ALOS: Advanced Land Observing Satellite; MBCNN: multibranch CNN; RNN: recurrent neural networks; and STDCNN: semitransfer deep CNN.

Specifically, both Rezaee et al. [34] and Huang et al. [29] achieved good accuracies using CNN-based models and monotemporal, single-sensor data for wetland land cover and urban land use classification. Meanwhile, Mahdianpari et al. [38] investigated well-known deep learning models (e.g., ResNet, DenseNet, InceptionResNet, etc.) for wetland mapping and demonstrated that InceptionResNetV2 showed the best accuracy (96.17%). They concluded that CNN outperformed traditional machine-learning methods (e.g., random forest) in the context of complex heterogeneous landscapes, which was consistent with our findings. However, neither multitemporal nor multisensor datasets were incorporated, meaning that these methods lacked the ability to comprehensively characterize the land surface.

From the perspective of multitemporal classification, Rußwurm et al. [35] utilized recurrent neural networks (RNNs) and multitemporal Sentinel-2 data for land cover classification, and they achieved good performance. They concluded that RNN was appropriate for modeling the relationship of sequential remote sensing data, and that it showed high accuracy in multitemporal classification. Different from Rußwurm et al. [35], the MBCNN in this study utilized a feature fusion module that directly learned the importance of each temporal feature to classification performance in order to fuse temporal features. Ji et al. [36] utilized 3D CNN to learn spatio-temporal features for crop classification based on multitemporal optical data—a method that also showed high accuracy. However, when compared with MBCNN, which was based on two-dimensional (2D) convolution, 3D CNN had the drawbacks of requiring high computing complexity and having gradient vanishing along the depth channel.

In the context of multisensor fusion and classification, Xu et al. [33] adopted a two-branch CNN for urban land use classification based on hyperspectral, light detection, and ranging (LiDAR) data. As for the method of data fusion, Xu et al. [33] simply used feature stacking without considering the importance of each feature. Scarpa et al. [37] studied the fusion of Sentinel-1 and Sentinel-2 data based on deep learning. They first stacked all multitemporal Sentinel-1 and Sentinel-2 data and utilized a CNN to extract features from the stacked data. Apparently, their fusion method was more on the data-level and less on the feature-level, which may lead to a weak robustness of the fused features. Compared with these studies, we constructed a feature-level fusion method that took the importance of each feature into consideration, which could increase the representativeness and robustness of the output features.

Moreover, the above previous studies did not consider variations in shapes and scales of land objects, which was one of the most important reasons for improving limited land cover classification accuracy. To tackle this issue, deformable convolution, which could extract robust features regardless of shape and scale variations, was introduced in this study.

Overall, Table 9 indicated that the proposed MBCNN could achieve good classification performance when compared with state-of-the-art methods. Additionally, the proposed method could be used for crop type mapping through joint use of Sentinel-1 and Sentinel-2 data for crop growth monitoring and yield estimation on the regional scale [48–51].

5. Conclusions

This paper proposed a multibranch convolutional neural network for fusion of multitemporal and multisensor Sentinel data for coastal land cover classification. The proposed neural network leverages a series of single-branch CNNs for feature extraction from single-date and single-sensor Sentinel data. Deformable convolutions and multiscale residual blocks were introduced to account for the variations in shapes and scales of coastal land objects. Features extracted from each branch were then aggregated using an adaptive fusion module to make the final land cover predictions.

The experiments were performed in the Yellow River Delta, which is the largest natural delta in China. The results indicated that the proposed multibranch CNN achieved good performance with an overall accuracy of 93.78% and a Kappa coefficient of 0.9297. The introduction of deformable convolutions increased the OA by 2.09%, which justified its role in modeling complex and fragmented coastal landscapes. Meanwhile, inclusion of multitemporal data improved the OA by 1.15%–11.85%, with an average increase of 6.85%, which justified the importance of temporal information in coastal land cover classification. Moreover, when compared with optical data alone, the inclusion of radar data increased the OA from 90.54% to 93.78% with an improvement of 3.24%, which indicated that the fusion of multisensor Sentinel data could enhance the separability of coastal land cover types. However, using radar data alone cannot achieve an accurate classification result. The proposed adaptive fusion method improved the OA by an increase of 2.28% when compared with the feature-stacking method, which also justified its effectiveness in multisource data fusion.

This paper demonstrates that the proposed multibranch CNN can effectively extract and integrate features from multitemporal and multisensor Sentinel-1 and Sentinel-2 remote sensing data, which achieves good performance in coastal land cover classification. In addition, the proposed network architecture can be considered as a general framework for multitemporal and multisensor data fusion. Future work should consider more study cases to further verify the effectiveness of the proposed MBCNN.

Author Contributions: Q.F. proposed the multibranch convolutional neural network of this study and contributed to the data processing, experiments, and manuscript writing. J.Y. and D.Z. contributed to the discussion of the experiments and manuscript revision. J.L., H.G., B.B., and B.L. mainly contributed to the manuscript revision.

Funding: This study is funded and supported by the National Natural Science Foundation of China (U1706211), China Postdoctoral Science Foundation (2018M641529), and Ministry of land and resources industry public welfare projects (201511010-06).

Acknowledgments: The authors would like to thank the European Space Agency for providing the Sentinel-1 and Sentinel-2 data and SNAP software for data preprocessing. Additionally, the authors would like to give special thanks to the anonymous reviewers and editors for their very useful comments and suggestions to help improve the quality of this paper.

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

References

1. Kuenzer, C.; Klein, I.; Ullmann, T.; Georgiou, E.; Baumhauer, R.; Dech, S. Remote Sensing of River Delta Inundation: Exploiting the Potential of Coarse Spatial Resolution, Temporally-Dense MODIS Time Series. *Remote Sens.* **2015**, *7*, 8516–8542. [[CrossRef](#)]
2. Islam, M.R.; Miah, M.G.; Inoue, Y. Analysis of Land use and Land Cover Changes in the Coastal Area of Bangladesh using Landsat Imagery. *Land Degrad. Develop.* **2016**, *27*, 899–909. [[CrossRef](#)]
3. Torbick, N.; Salas, W. Mapping agricultural wetlands in the Sacramento Valley, USA with satellite remote sensing. *Wetlands Ecol. Manag.* **2015**, *23*, 79–94. [[CrossRef](#)]
4. Henderson, F.; Lewis, A. Radar detection of wetland ecosystems: a review. *Int. J. Remote Sens.* **2008**, *29*, 5809–5835. [[CrossRef](#)]
5. Mahdavi, S.; Salehi, B.; Granger, J.; Amani, M.; Brisco, B.; Huang, W. Remote sensing for wetland classification: a comprehensive review. *GISci. Remote Sens.* **2018**, *55*, 623–658. [[CrossRef](#)]
6. Davranche, A.; Lefebvre, G.; Poulin, B. Wetland monitoring using classification trees and SPOT-5 seasonal time series. *Remote Sens. Environ.* **2010**, *114*, 552–562. [[CrossRef](#)]

7. Yang, X.; Chen, L.; Li, Y.; Xi, W.; Chen, L. Rule-based land use/land cover classification in coastal areas using seasonal remote sensing imagery: a case study from Lianyungang City, China. *Environ. Monit. Assess.* **2015**, *187*, 449. [CrossRef]
8. Feng, Q.; Gong, J.; Liu, J.; Li, Y. Monitoring Cropland Dynamics of the Yellow River Delta based on Multi-Temporal Landsat Imagery over 1986 to 2015. *Sustainability* **2015**, *7*, 14834–14858. [CrossRef]
9. Rodrigues, S.W.P.; Souza-Filho, P.W.M. Use of multi-sensor data to identify and map tropical coastal wetlands in the Amazon of Northern Brazil. *Wetlands*. **2011**, *31*, 11–23. [CrossRef]
10. Beijma, S.; Comber, A.; Lamb, A. Random forest classification of salt marsh vegetation habitats using quad-polarimetric airborne SAR, elevation and optical RS data. *Remote Sens. Environ.* **2014**, *149*, 118–129. [CrossRef]
11. Corcoran, J.; Knight, J.; Gallant, A. Influence of Multi-Source and Multi-Temporal Remotely Sensed and Ancillary Data on the Accuracy of Random Forest Classification of Wetlands in Northern Minnesota. *Remote Sens.* **2013**, *5*, 3212–3238. [CrossRef]
12. Lane, C.R.; Liu, H.; Autrey, B.C.; Anenkhonov, O.A.; Chepinoga, V.V.; Wu, Q. Improved Wetland Classification Using Eight-Band High Resolution Satellite Imagery and a Hybrid Approach. *Remote Sens.* **2014**, *6*, 12187–12216. [CrossRef]
13. Franklin, S.E.; Skerries, E.M.; Stefanuk, M.A.; Ahmed, O.S. Wetland classification using Radarsat-2 SAR quad-polarization and Landsat-8 OLI spectral response data: a case study in the Hudson Bay Lowlands Ecoregion. *Int. J. Remote Sens.* **2018**, *39*, 1615–1627. [CrossRef]
14. Hird, J.N.; DeLancey, E.R.; McDermid, G.J.; Kariyeva, J. Google Earth Engine, Open-Access Satellite Data, and Machine Learning in Support of Large-Area Probabilistic Wetland Mapping. *Remote Sens.* **2019**, *11*, 43. [CrossRef]
15. Mahdianpari, M.; Salehi, B.; Mohammadimanesh, F.; Homayouni, S.; Gill, E. The First Wetland Inventory Map of Newfoundland at a Spatial Resolution of 10 m Using Sentinel-1 and Sentinel-2 Data on the Google Earth Engine Cloud Computing Platform. *Remote Sens.* **2017**, *9*, 1315. [CrossRef]
16. Erinjery, J.J.; Singh, M.; Kent, R. Mapping and assessment of vegetation types in the tropical rainforests of the Western Ghats using multispectral Sentinel-2 and SAR Sentinel-1 satellite imagery. *Remote Sens. Environ.* **2018**, *216*, 345–354. [CrossRef]
17. Hajj, M.E.; Baghdadi, N.; Zribi, M.; Bazzi, H. Synergic Use of Sentinel-1 and Sentinel-2 Images for Operational Soil Moisture Mapping at High Spatial Resolution over Agricultural Areas. *Remote Sens.* **2017**, *9*, 1292. [CrossRef]
18. Tricht, K.V.; Gobin, A.; Gilliams, S.; Piccard, I. Synergistic Use of Radar Sentinel-1 and Optical Sentinel-2 Imagery for Crop Mapping: A Case Study for Belgium. *Remote Sens.* **2018**, *10*, 1642. [CrossRef]
19. Muller-Wilm, U. Sentinel-2 MSI – Level-2A Prototype Processor Installation and User Manual. 2016. Telespazio VEGA Deutschland GmbH, Darmstadt. Available online: <http://step.esa.int/thirdparties/sen2cor/2.2.1/S2PAD-VEGA-SUM-0001-2.2.pdf> (accessed on 27 April 2019).
20. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
21. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Proc. Adv. Neural Inf. Process. Syst.* **2012**, 1097–1105. [CrossRef]
22. Bulat, A.; Tzimiropoulos, G. Binarized Convolutional Landmark Localizers for Human Pose Estimation and Face Alignment with Limited Resources. *Proc. IEEE Int. Conf. Comput. Vis.* **2017**, 3706–3714.
23. Dai, J.; Qi, H.; Xiong, Y.; Li, Y.; Zhang, G.; Hu, H.; Wei, Y. Deformable Convolutional Networks. Arxiv 2017 [1703.06211]. Available online: <https://arxiv.org/pdf/1703.06211.pdf> (accessed on 27 April 2019).
24. Jin, Q.; Meng, Z.; Pham, T.D.; Chen, Q.; Wei, L.; Su, R. DUNet: A Deformable Network for Retinal Vessel Segmentation. Arxiv 2018 [1811.01206]. Available online: <https://arxiv.org/pdf/1811.01206.pdf> (accessed on 27 April 2019).
25. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. Arxiv 2017 [1709.01507]. Available online: <https://arxiv.org/pdf/1709.01507.pdf> (accessed on 27 April 2019).
26. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal loss for dense object detection. *Proc. IEEE Int. Conf. Comput. Vis.* **2017**, 2999–3007.
27. Zhu, X.; Tuia, D.; Mou, L.; Xia, G.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geosci. Remote Sens. M.* **2017**, *5*, 8–36. [CrossRef]

28. Pan, X.; Gao, L.; Marinoni, A.; Zhang, B.; Yang, F.; Gamba, P. Semantic Labeling of High Resolution Aerial Imagery and LiDAR Data with Fine Segmentation Network. *Remote Sens.* **2018**, *10*, 743. [[CrossRef](#)]
29. Huang, B.; Zhao, B.; Song, Y. Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sens. Environ.* **2018**, *214*, 73–86. [[CrossRef](#)]
30. Feng, Q.; Zhu, D.; Yang, J.; Li, B. Multisource Hyperspectral and LiDAR Data Fusion for Urban Land-Use Mapping based on a Modified Two-Branch Convolutional Neural Network. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 28. [[CrossRef](#)]
31. Ghamisi, P.; Hofle, B.; Zhu, X. Hyperspectral and LiDAR Data Fusion Using Extinction Profiles and Deep Convolutional Neural Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3011–3024. [[CrossRef](#)]
32. Hughes, L.H.; Schmitt, M.; Mou, L.; Wang, Y.; Zhu, X. Identifying Corresponding Patches in SAR and Optical Images with a Pseudo-Siamese CNN. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 784–788. [[CrossRef](#)]
33. Xu, X.; Li, W.; Ran, Q.; Du, Q.; Gao, L.; Zhang, B. Multisource Remote Sensing Data Classification Based on Convolutional Neural Network. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 937–949. [[CrossRef](#)]
34. Rezaee, M.; Mahdianpari, M.; Zhang, Y.; Salehi, B. Deep Convolutional Neural Network for Complex Wetland Classification Using Optical Remote Sensing Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 3030–3039. [[CrossRef](#)]
35. Rußwurm, M.; Körner, M. Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 129. [[CrossRef](#)]
36. Ji, S.; Zhang, C.; Xu, A.; Shi, Y.; Duan, Y. 3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images. *Remote Sens.* **2018**, *10*, 75. [[CrossRef](#)]
37. Scarpa, G.; Gargiulo, M.; Mazza, A.; Gaetano, R. A CNN-Based Fusion Method for Feature Extraction from Sentinel Data. *Remote Sens.* **2018**, *10*, 236. [[CrossRef](#)]
38. Mahdianpari, M.; Salehi, B.; Rezaee, M.; Mohammadimanesh, F.; Zhang, Y. Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery. *Remote Sens.* **2018**, *10*, 1119. [[CrossRef](#)]
39. Yang, J.; Ren, G.; Ma, Y.; Fan, Y. Coastal wetland classification based on high resolution SAR and optical image fusion. *Proc. IEEE Int. Conf. Comput. Vis.* **2016**, 886–889.
40. Ottinger, M.; Kuenzer, C.; Liu, G.; Wang, S.; Dech, S. Monitoring land cover dynamics in the Yellow River Delta from 1995 to 2010 based on Landsat 5 TM. *Appl. Geol.* **2013**, *44*, 53–68. [[CrossRef](#)]
41. Liu, G.; Zhang, L.; Zhang, Q.; Musyimi, Z.; Jiang, Q. Spatio-Temporal Dynamics of Wetland Landscape Patterns Based on Remote Sensing in Yellow River Delta, China. *Wetlands.* **2014**, *34*, 787–801. [[CrossRef](#)]
42. Liu, J.; Feng, Q.; Gong, J.; Zhou, J.; Li, Y. Land-cover classification of the Yellow River Delta wetland based on multiple end-member spectral mixture analysis and a Random Forest classifier. *Int. J. Remote Sens.* **2016**, *37*, 1845–1867. [[CrossRef](#)]
43. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. Arxiv 2015 [1502.01852]. Available online: <https://arxiv.org/pdf/1502.01852.pdf> (accessed on 27 April 2019).
44. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. Arxiv 2014 [1412.6980]. Available online: <https://arxiv.org/pdf/1412.6980.pdf> (accessed on 27 April 2019).
45. TensorFlow. Available online: <https://tensorflow.google.cn/> (accessed on 17 November 2018).
46. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
47. Chapelle, O.; Vapnik, V.; Bousquet, O.; Mukherjee, S. Choosing Multiple Parameters for Support Vector Machines. *Mach. Learn.* **2002**, *46*, 131–159. [[CrossRef](#)]
48. Huang, J.; Ma, H.; Sedano, F.; Lewis, P.; Liang, S.; Wu, Q.; Zhang, X.; Zhu, D. Evaluation of regional estimates of winter wheat yield by assimilating three remotely sensed reflectance datasets into the coupled WOFOST-PROSAIL model. *Eur. J. Agron.* **2019**, *102*, 1–13. [[CrossRef](#)]
49. Huang, J.; Sedano, F.; Huang, Y.; Ma, H.; Li, X.; Liang, S.; Tian, L.; Zhang, X.; Fan, J.; Wu, W. Assimilating a synthetic Kalman filter leaf area index series into the WOFOST model to estimate regional winter wheat yield. *Agr. Forest Meteorol.* **2016**, *216*, 188–202. [[CrossRef](#)]

50. Huang, J.; Tian, L.; Liang, S.; Ma, H.; Becker-Reshef, I.; Su, W.; Huang, Y.; Zhang, X.; Zhu, D.; Wu, W. Improving winter wheat yield estimation by assimilation of the leaf area index from Landsat TM and MODIS data into the WOFOST model. *Agr. Forest Meteorol.* **2015**, *204*, 106–121. [[CrossRef](#)]
51. Huang, J.; Ma, H.; Su, W.; Zhang, X.; Huang, Y.; Fan, J.; Wu, W. Jointly assimilating MODIS LAI and ET products into the SWAP model for winter wheat yield estimation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4060–4071. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).