A Novel Method for Identifying Crops in Parcels Constrained by Environmental Factors Through the Integration of a Gaofen-2 High-Resolution Remote Sensing Image and Sentinel-2 Time Series

Weijia Chen and Guilin Liu

***Abstract*—Accurately mapping crop cultivation types is essential for the sustainable development of precision agriculture. Environmental restrictions on crop growth, such as soil salinization in arid zones, generally lead to spatial crop growth heterogeneity within cropland fields, which in turn generates differences in the spectral responses reflected in optical remote sensing images of the same croplands and leads to pixel-scale crop-mapping misclassifications. Thus, through this article, we proposed a method to solve this problem at the geoparcel scale by integrating geometric features from a Gaofen-2 high-resolution remote sensing image and the spectral-temporal features derived from Sentinel-2 time series. The results showed that cropland parcels could be accurately extracted from Gaofen-2 images by employing the U-Net semantic segmentation model with an overall accuracy (OA) reaching 97% and a kappa coefficient of 0.95. Then, geoparcel-scale crop types were mapped based on prior crop phenology knowledge and the corresponding Sentinel-2 time series using the time-weighted dynamic time warping (TWDTW) classification algorithm. The parcel-based TWDTW algorithm had an OA of 99.64%, a kappa coefficient of 0.99, and optimal spatial homogeneity in the results, thus outperforming the pixel-based TWDTW method. These results provide a potential solution for mapping crops under spatially heterogeneous cropland conditions affected by various environmental constraints.**

***Index Terms*—Crop mapping, parcel-based, segment, spectral-temporal features, time-weighted dynamic time warping (TWDTW).**

I. Introduction

With the continuous growth of the global population and the limited availability of resources, sustainable agricultural development has become a significant challenge for human beings in today's society. Obtaining accurate crop cultivation maps serves as the foundation for agricultural management and decision-making and plays a crucial role in optimizing crop production and achieving precision agriculture management. Traditionally, mapping crop cultivation patterns has typically relied on manual surveys. However, this method is time-consuming and laborious and has certain limitations in large-scale mapping. Over time, remote sensing has shown impressive potential in mapping crop types and monitoring crop dynamics, thus gradually becoming an effective means for agricultural mapping [1], [2], [3], [4]. With the increase in available remote sensing images, such as Landsat and Sentinel images [3], some studies have reported that the crop-classification accuracies obtained based on multitemporal remote sensing data are higher than those based on single-phase remote sensing images [5], [6], [7]. This is because multitemporal satellite remote sensing images can capture the entire crop growth process and then distinguish among different crops by using their species-specific phenological variations [8]. Therefore, time-series remote sensing images and the features derived from these images can be widely used in research involving global and regional crop-type mapping [9], [10].

In previous studies, researchers have attempted to extract phenological information from time-series remote sensing data (TSRD) for crop-classification applications using machine learning algorithms, such as random forests [5] and support vector machines [11]. Although the above-mentioned methods have performed well in mapping crops or land cover, their performances are limited by several challenges, such as follows.

1) Lack of available training sample data [12].
2) Lack of efficient satellite observations for TSRD caused by the existence of clouds [13].
3) These classifiers failed to fully utilize the crop phenological information, i.e., disrupting the temporal order of TRSD has no effect on the performance of crop classification [6], [7].

An effective approach to address the above challenges is the time-weighted dynamic time warping (TWDTW) algorithm proposed by Maus et al. [14]. Compared with the traditional

Manuscript received 11 May 2023; revised 27 September 2023; accepted 28 October 2023. Date of publication 3 November 2023; date of current version 23 November 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 42271346 and Grant 41901349; in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2023A1515010736; and in part by the Startup Foundation for Talented Scholars in South China Normal University under Grant 8S0472. *(Corresponding author: Guilin Liu.)*

The authors are with the School of Geography, South China Normal University, Guangzhou 510631, China (e-mail: weijia@m.scnu.edu.cn; liuguilin@m.scnu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2023.3329987

dynamic time warping algorithm, TWDTW adds a time penalty strategy to improve the classification accuracy [5]. The TWDTW algorithm [14] has been proven to be effective in identifying the phenological patterns of crops with different agronomic management practices from TSRD [15], including vegetables [16], [17] and grains [18], as well as in assessing cropping patterns [19].

The aforementioned studies focused mainly on classifying crops at the pixel-based level using TSRD. In arid zones, the spatial heterogeneity of crop growth within croplands caused by soil salinization leads to "salt and pepper noise" in remote sensing-generated crop-mapping results using pixel-based classification methods. In response, some scholars have employed the object-based image analysis (OBIA) method [20] to generate a series of objects formed by a group of pixels with similar spectral and geometric features for mapping crops [21]. Previous studies have revealed that OBIA has a better crop-mapping performance than the pixel-based classification strategy [5], [12]. Within the OBIA framework, the optimal segmentation scheme for obtaining objects determines the accuracy of the subsequent crop classification results. The development of deep learning technology has provided a solution for the precise extraction of cropland edges [22]. Deep learning models (e.g., convolutional neural networks) can automatically learn representative discriminative features based on existing training datasets [23], which improves the efficiency and accuracy of satellite-image segmentation and classification [13]. Recently, numerous satellite image segmentation models based on DL have been proposed [24]. Among them, U-Net has been widely implemented for extracting parcels from satellite images [22], [25], [26]. Zhang et al. [27] applied the U-Net model to extract cropland boundaries from Sentinel-2 images in the Heilongjiang region of China, and the results showed an overall accuracy (OA) improvement of 3.78% over traditional OBIA. These studies show that the U-Net model has high practicability in cropland edge extraction. Meanwhile, farmers usually manage their crops based on agricultural parcels. Therefore, it is necessary to conduct parcel-scale crop mapping to provide important information with reference for farmers to develop targeted agricultural management practices and to achieve precision agricultural management for sustainable agricultural development.

In regions with heterogeneous crop growth, it is difficult to obtain high accuracy for crop mapping based on remote sensing data with a single feature band [28]. Some crops have similar phenological characteristics, resulting in small spectral differences between different crop types. Thus, it is difficult to distinguish even with high spatial resolution satellite images [29]. Although multispectral resolution satellite data (e.g., Sentinel-2, Landsat 8) can distinguish the spectral features of different crops, the spatial resolution of satellite data is relatively rough for identifying cropland edges, resulting in noncrop types (e.g., roads) that can easily be misidentified as crops [30]. Therefore, integrating the temporal-spatial and spectral features of multisource satellite images is considered an effective way to overcome the above-mentioned problems [29]. Wu et al. [29] combined high spatial and high spectral resolution remote sensing data to map crop cultivation patterns in complex heterogeneous regions and achieved much higher accuracy than that of classification using only high spatial resolution data or using only high spectral resolution data. Wu et al. [31] also demonstrated that the accuracy of crop mapping can be improved by integrating spatial and temporal spectral characteristics from multiple satellite data sources. Overall, the combination of high-resolution remote sensing data and medium-resolution TSRD meets the need to obtain accurate and timely information to map precise crop cultivation patterns and further support precision agricultural management.

To solve the problem of crop-type misidentifications in classification results due to the heterogeneous crop-growth phenomenon within croplands in arid regions, we proposed a novel method for identifying crops through the integration of spatial features and the spectral-temporal features of multisource satellite images. We extracted agricultural geoparcels in the study area based on a high-resolution image using the U-Net model, assigned temporal and spectral features derived from medium-resolution images to each geoparcel and mapped the crop types at the parcel-based level using the TWDTW algorithm. Specifically, we aimed to 1) propose a method to obtain crop types by integrating the spatial features of a high-resolution remote sensing image (GaoFen-2) and the spectral-temporal features of medium-resolution satellite images (Sentinel-2) and 2) evaluate and compare the effect of the results obtained using pixel-based and parcel-based classification methods.

II. Study Area and Data Sources

A. Study Area

The study area, which covers 7196 ha, is located at the southern foot of the Tianshan Mountains and the northern edge of the Taklamakan Desert (see Fig. 1) [32]; this region has a typical temperate continental climate with abundant sunshine and little precipitation, the average air temperature ranges from 3.8 °C to 19.3 °C, and the average precipitation ranges from 11.9 to 91.9 mm [32]. The Tarim River is the main source of irrigation water for agricultural activities in the study area. In Alar, cotton and fruit have been the main cash crops in recent decades [33]. The study area has prominent land degradation and soil salinization problems [34], which have led to crop growth heterogeneity within the local croplands.

B. Data Sources

1) Satellite Remote Sensing Images: We obtained 95 Sentinel-2 Level 2A surface reflectance images with almost no cloud cover taken from March–November 2021 from the European Space Agency using the Google Earth Engine (GEE) platform (see Fig. 2). Then, we mosaicked these Sentinel-2 remote sensing images by the same acquisition date and clipped them to the scope of the study area using GEE since the study area is located at the intersection of four Sentinel-2 images (the splicing domain numbers are T44TNL, T44TNK, T44TML, and T44TMK). Meanwhile, a GaoFen-2 remote sensing image (cloud cover $< 5\%$) taken of the study area in August 2021 was obtained from the China Centre for Resources Satellite

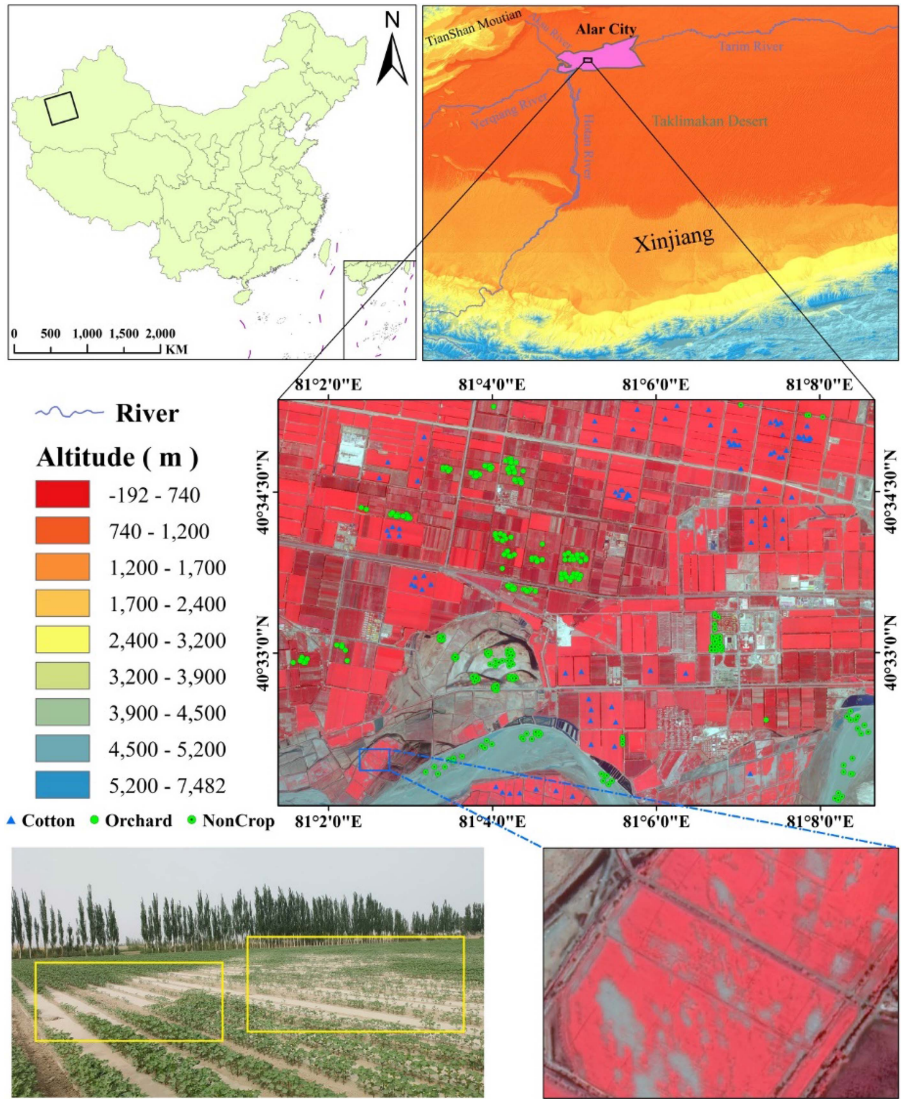

Fig. 1. Location of the study area in China (note: the Gaofen-2 red–green–blue composite image (near-infrared (NIR)/red/green) captured on August 12, 2021; the yellow boxes represent the heterogeneity of cotton growth in the croplands; and the picture of cotton seedlings was acquired by photography in May 2022).

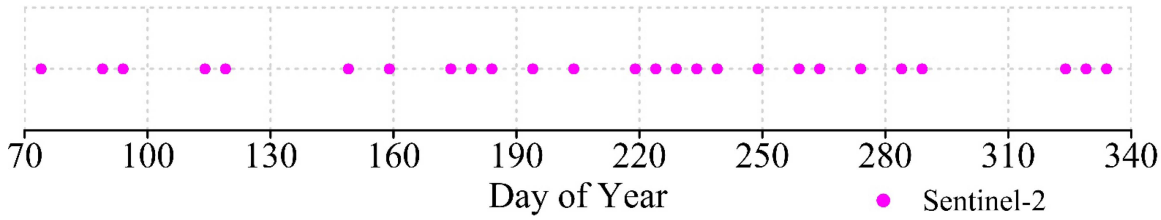

Fig. 2. Available Sentinel-2 time series covering the study area in 2021.

Data and Application (https://data.cresda.cn). The Gaofen-2 image contains information in four multispectral bands and one panchromatic band (see Table I). The high spatial resolution of GF-2 images can clearly capture the spatial characteristics of ground objects and has been successfully applied to the segmentation of agricultural fields [35], [36], [37]. Therefore, we finally selected the GF-2 image for parcel segmentation. First, we preprocessed the GF-2 image, including orthorectification and atmospheric correction of the image by using the rational polynomial coefficients model [38] and fast line-of-sight atmospheric analysis of spectral hypercubes atmospheric correction model, respectively. Then, to simultaneously preserve the high spatial resolution and rich spectral information of the image, we fused the multispectral and panchromatic bands to obtain a

TABLE I
DESCRIPTIONS OF THE GAOFEN-2 REMOTE SENSING IMAGES USED IN THIS STUDY

Bands	Spectral (μm) resolution	Spatial resolution (m)	Swath (km)	Revisit (day)
Blue	0.45–0.52	4		
Green	0.52–0.59	4		
Red	0.63–0.69	4	45	5
Near infrared	0.77–0.89	4		
Pan	0.45–0.90	1		

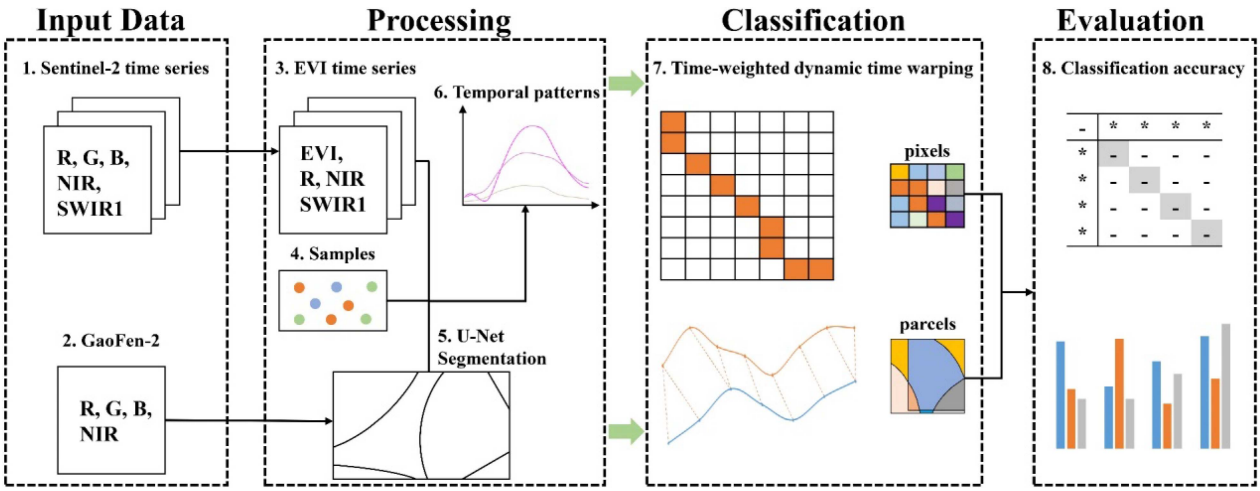

Fig. 3. Workflow of the pixel-based and parcel-based TWDTW classification methods.

multispectral image with a spatial resolution of one meter using the Gram–Schmidt pansharpening method [39].

2) Training and Validation Samples: We collected 107 ground truth samples in the Alar irrigation area using a handheld GPS instrument during the field campaign in May 2021, of which 95 samples were located in the present study area. Furthermore, we expanded ground truth samples based on geometric features of land cover types from high-resolution remote sensing images of Google Earth and temporal features from Sentinel-2 time series. Finally, we obtained 93 cotton samples, 144 orchard samples (apple, pear, and jujube trees), and 164 noncrop samples (bare lands, roads, water bodies, and construction lands) (see Fig. 1). We refer to previous studies that TWDTW requires fewer training samples [5], [7]. Therefore, among these samples, 30% of each type were used for the classification training, and 70% of each type were used for the classification validation.

III. METHODOLOGY

The workflow developed herein for the identification of crop types included four major steps (see Fig. 3). First, we obtained and preprocessed Sentinel-2 and Gaofen-2 remote sensing images. Second, we segmented the cropland parcels using the U-Net model based on the Gaofen-2 images. Moreover, we constructed enhanced vegetation index (EVI) time-series patterns from the Sentinel-2 images and assigned values to each parcel. Third, we implemented pixel-based and parcel-based crop-type classification steps based on the training samples using the TWDTW approach. Finally, the accuracies of the results obtained using the pixel-based and parcel-based classification methods were evaluated.

A. Prior Knowledge of Crop Phenology

In the study area, cotton is generally sown in early April, seedlings emerge at the end of April, buds appear in mid-June, and the EVI values continue to increase throughout this period (see Fig. 4). The cotton-flowering event occurs in early July, and the EVI peaks at this time correspondingly. In mid-September, cotton enters the boll-splitting stage followed by the harvest stage in early mid-October, and the EVI decreases correspondingly. In contrast to cotton, fruit trees (apple, pear, and jujube trees) mainly undergo vegetative growth from early April to early May; during this period, the EVI value continues to increase at values above those of cotton (see Fig. 4). Reproductive growth occurs from mid-May to late August, with the EVI values peaking in July but remaining lower than those of cotton. After the harvest in September, the deciduous period starts in late October and lasts until late November, so the EVI value decreases slowly and becomes higher than that of cotton after August. The noncrop land types exhibit no significant intra-annual variations in the yearly EVI curves. In this study, the EVI was calculated from the Sentinel-2 images to construct the temporal phenological patterns of different crop types. Then, a Savitzky–Golay filter

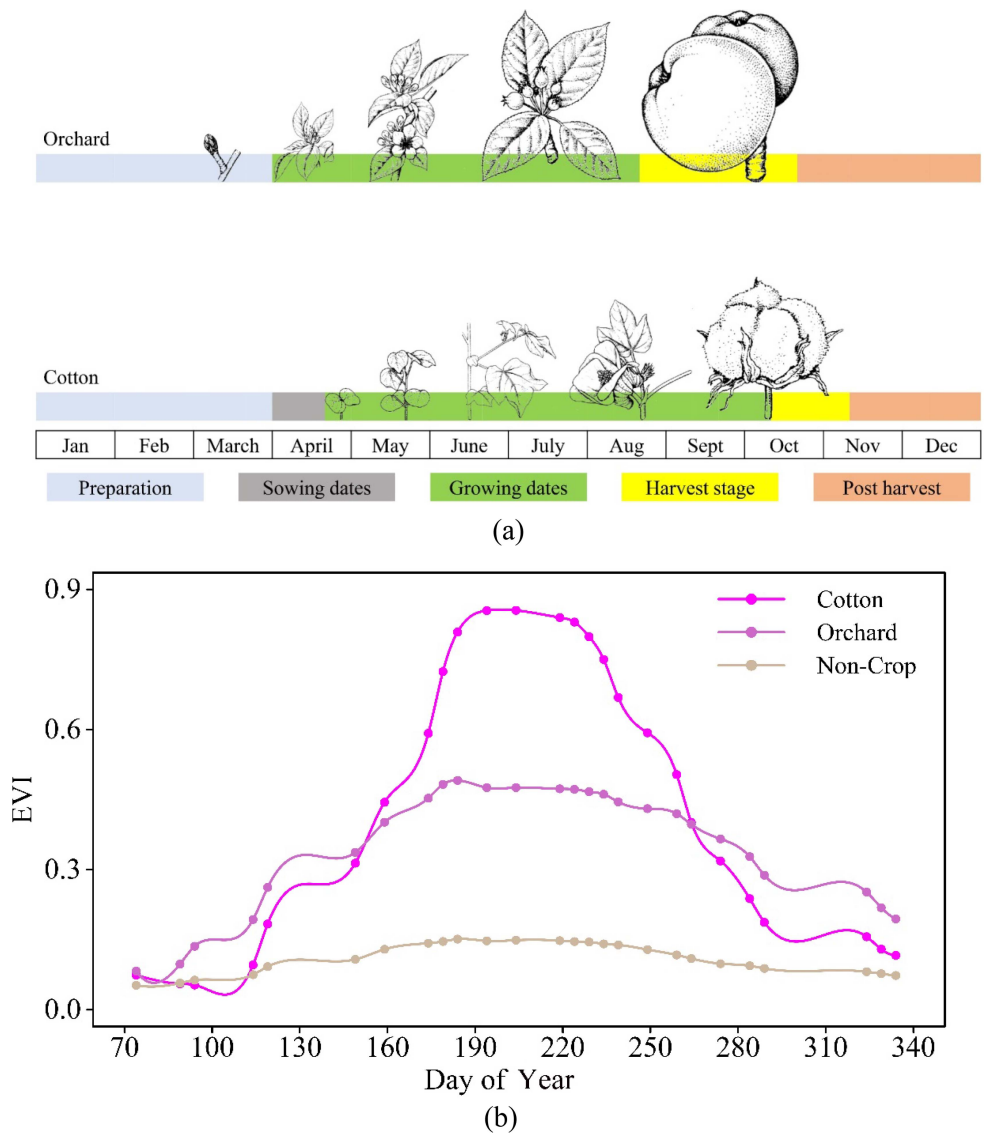

Fig. 4. (a) Phenological calendar of cotton in the study area, plant pictures obtained from the website and (b) the corresponding temporal EVI patterns in 2021 (the plant pictures were collected from Meier [42]).

was applied to reduce noise and smooth the EVI curves of various crops (see Fig. 4) since the Savitzky–Golay filter is a smoothing algorithm that is widely used for data stream smoothing and denoising [40]. The EVI is a vegetation index commonly used in studies of vegetation phenology and crop identification; the EVI is more sensitive to biomass and less likely to be saturated than the normalized difference vegetation index and is also less influenced by atmospheric and soil conditions [41]. The EVI formula is defined as follows:

$$EVI = 2.5 \times \frac{\rho NIR - \rho Red}{\rho NIR + 6 \times \rho Red - 7.5 \times \rho Blue + 1} \quad (1)$$

where $\rho Blue$, ρRed, and ρNIR are the reflectances in the blue, red, and NIR bands, respectively.

For the pixel-based classification, we used the mean curves (including the mean curves of the red, NIR, and shortwave infrared (SWIR) bands and EVI) of the Sentinel-2 time series at the location of training samples for each crop type to construct the phenological patterns for different crop types. Similarly, in the parcel-based classification, we also used the mean curves, but time series images were based on the combination of segmented parcels from Gaofen-2 and Sentinel-2 time series.

B. U-Net Architecture for Image Segmentation

When the contractive path shown on the left side of Fig. 5 extracts the high-dimensional features from the Gaofen-2 images, the U-Net model increases the spectral resolution of the Gaofen-2 images by reducing the spatial resolution and increasing the number of channels (see Fig. 5) [43]. Conversely, the expansive path shown on the right side of Fig. 5 reverts those predictions to the size of the original Gaofen-2 images; the U-Net model increases the spatial resolution of the images by reducing the spectral resolution and number of channels. The contractive

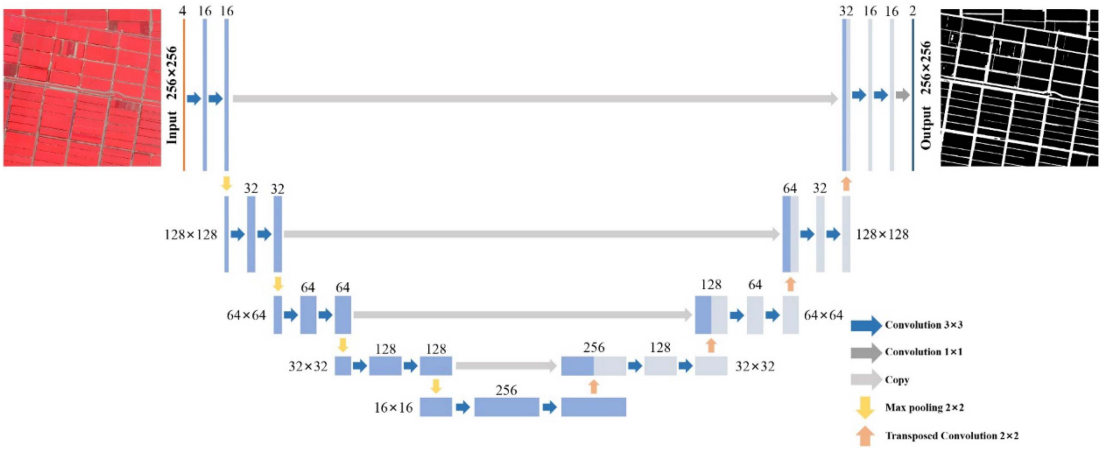

Fig. 5. U-Net architecture.

path consists of 3×3 convolutions with a rectified linear unit and a 2×2 max-pooling operation, which doubles the number of channels during each downsampling iteration. The components of the expansive path are the same as those of the contractive path, but the operation is reversed. The parcel segmentation results were finally mapped to a 1×1 convolutional layer (see Fig. 5).

In this study, we selected regions of interest (ROIs) near the study area from the Gaofen-2 image using visual interpretation in ArcGIS software. Based on the spectral features of the Gaofen-2 image, we selected nearly half of the ROIs represent parcels with homogeneous crop growth, and the other half of the ROIs represent parcels with heterogeneous crop growth. This allows the U-Net model to actively learn the spectral and morphological features within these two types of parcels so that both types of parcels can be accurately extracted from the image. These labeled ROIs were then cropped into images with sizes of 256×256, resulting in a total of 13 864 images. Among these images, 60% were used for training, 20% were used for validation, and 20% were used for testing. The U-Net model was built and trained by using the Keras 2.3.1 package and TensorFlow 1.15.0 within Python 3.6.6 software. To exploit the performance of the U-Net model, we trained the model with all four bands of fused GaoFen-2 images (i.e., the visible blue, green, red, and NIR bands). After the continuous adjustment of the model training parameters (number of model epochs, batch size, and learning rate), the optimally trained U-Net model was obtained by setting the number of model epochs to 30, the batch size to 8, and the learning rate to 0.0001. After the geoparcels were extracted, each segmented geoparcel was assigned the mean value of the Sentinel-2 time-series pixels contained in each corresponding segmented parcel, including the average reflectance values of the red, NIR, and SWIR bands and the average EVI.

C. TWDTW for Crop Classification

In this study, TWDTW was implemented using the dtwSat 0.2.6 package within R 3.6.1 software [44]. In the TWDTW method, first, the time-penalty weight distance matrix is calculated between the reference with Sentinel-2 time series $A = [A1, A2, \ldots, Am]$ and the target to be classified with Sentinel-2 time series $B = [B1, B2, \ldots, Bn]$ (see Fig. 6) as follows:

$$D_{m\times n} = \begin{bmatrix} D_{1,1} & D_{1,2} & \cdots & D_{1,n} \\ D_{2,1} & D_{2,2} & \cdots & D_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ D_{m,1} & D_{m,2} & \cdots & D_{m,n} \end{bmatrix}. \tag{2}$$

In the above-mentioned equation, each element in the distance matrix $D_{m\times n}$ is defined as follows:

$$D_{i,j} = \omega_{i,j} * |A_i - B_j| \tag{3}$$

where $\omega_{i,j}$ is the time-weighting factor that follows a logistic function. The $\omega_{i,j}$ term is given by the following formula:

$$\omega_{i,j} = \frac{1}{1 + e^{-\alpha(|i-j|-\beta)}} \tag{4}$$

where i represents the date at the ith position in the temporal pattern; j represents the date at the jth position in the time series; α is the steepness of the logical weight; and β is the midpoint of the logical weight. We used $\alpha = -0.1$ and $\beta = 50$ in this work, representing a lower penalty for time warps shorter than 50 days and a higher penalty for longer time warps; this decision was made with reference to Belgiu and Csillik [5].

Second, we calculated the cumulative distance cost matrix $C_{m\times n}$, which determines the alignment relationship between the Sentinel-2-derived temporal EVI curves of the reference and those of the target to be classified. $C_{m\times n}$ is given by formula (5)

$$C_{i,j} = D_{i,j} + \min\{C_{i-1,j-1}, C_{i-1,j}, C_{i,j-1}\} \tag{5}$$

$$C_{i,j} = \begin{cases} D_{i,j}, & i = 1, j = 1 \\ D_{i,j} + C_{i-1,1}, & 1 < i \le M, j = 1 \\ D_{i,j} + C_{1,j-1}, & i = 1, 1 < j \le N \end{cases}. \tag{6}$$

With this alignment relationship, the similarity and dissimilarity of the two-time series, the reference and target, can be determined; in this way, the time series of all targets from the Sentinel-2 images were classified.

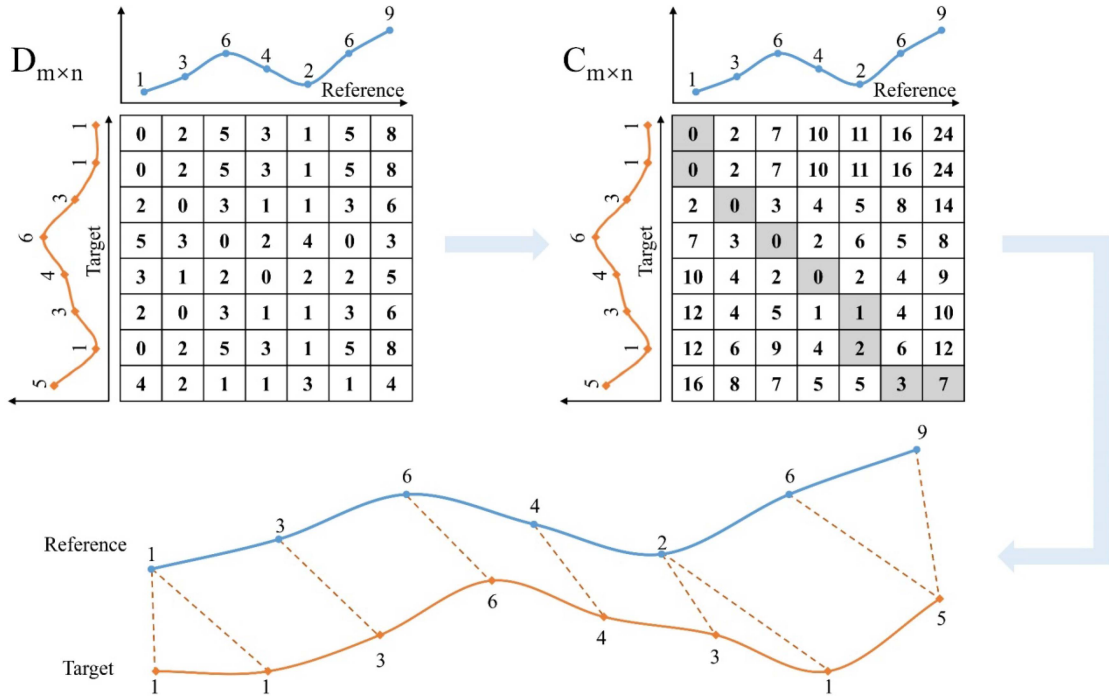

Fig. 6. Illustration of the TWDTW algorithm (adapted from Zhao et al. [18]).

D. Accuracy Assessment

To evaluate the model-training performance, we used five metrics, namely, the OA, precision, recall, F1-score, and kappa coefficient [45]. In this study, the pixel-based and parcel-based TWDTW classification results were evaluated using a confusion matrix to calculate the OA, user accuracy (UA), producer accuracy (PA) metrics, and kappa coefficient [45]. The OA is the ratio of the number of correctly classified pixels to the total number of pixels. The kappa coefficient is a metric used to test the consistency of classification results. The F1-score is a harmonic mean of the model precision and recall [36]. These metrics are defined as follows:

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

where *TP*, *TN*, *FP*, and *FN* are the number of true positive, true negative, false positive, and false negative results, respectively

$$\text{Precision} = \frac{TP}{TP + FP} \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{9}$$

$$\text{F1}-\text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

$$\text{Kappa} = \frac{\text{OA} - P_e}{1 - P_e} \tag{11}$$

$$P_e = \frac{(TP+FN) \times (TP+FP) \times (FN+FP) \times (TN+FP)}{N \times N} \tag{12}$$

TABLE II
CROPLAND SEGMENTATION ACCURACY OBTAINED USING THE U-NET ALGORITHM

Type	Precision	Recall	F1-score	OA (%)	Kappa
Cropland	0.95	0.97	0.96	97	0.95
Others	0.98	0.97	0.97		

where *TP*, *TN*, *FP*, and *FN* are the number of true positive, true negative, false positive, and false negative results, respectively.

IV. RESULTS

A. U-Net Model Performance on Cropland Segmentation

Fig. 7(a) shows the excellent parcel-extraction performance of these training results obtained using the U-Net model. When the U-Net model was trained to the 30th epoch, the training accuracy curve tended to be stable, and the value was close to 1. Moreover, the training loss value decreased from 0.71 to close to 0. To test the effectiveness of the model in the geoparcel extraction application, we validated the accuracy of the model with segmentation results using 2773 GaoFen-2256 × 256 images. The OA and kappa coefficient of the geoparcel extraction results reached 97% and 0.95, respectively (see Table II). The above-mentioned evidence showed the excellent performance of the trained U-Net model in extracting cropland parcels. The trained U-Net model mentioned above was then employed to segment the geoparcels of the croplands and other land-cover types using the Gaofen-2 image. Finally, 5971 geoparcels were obtained, of which 4320

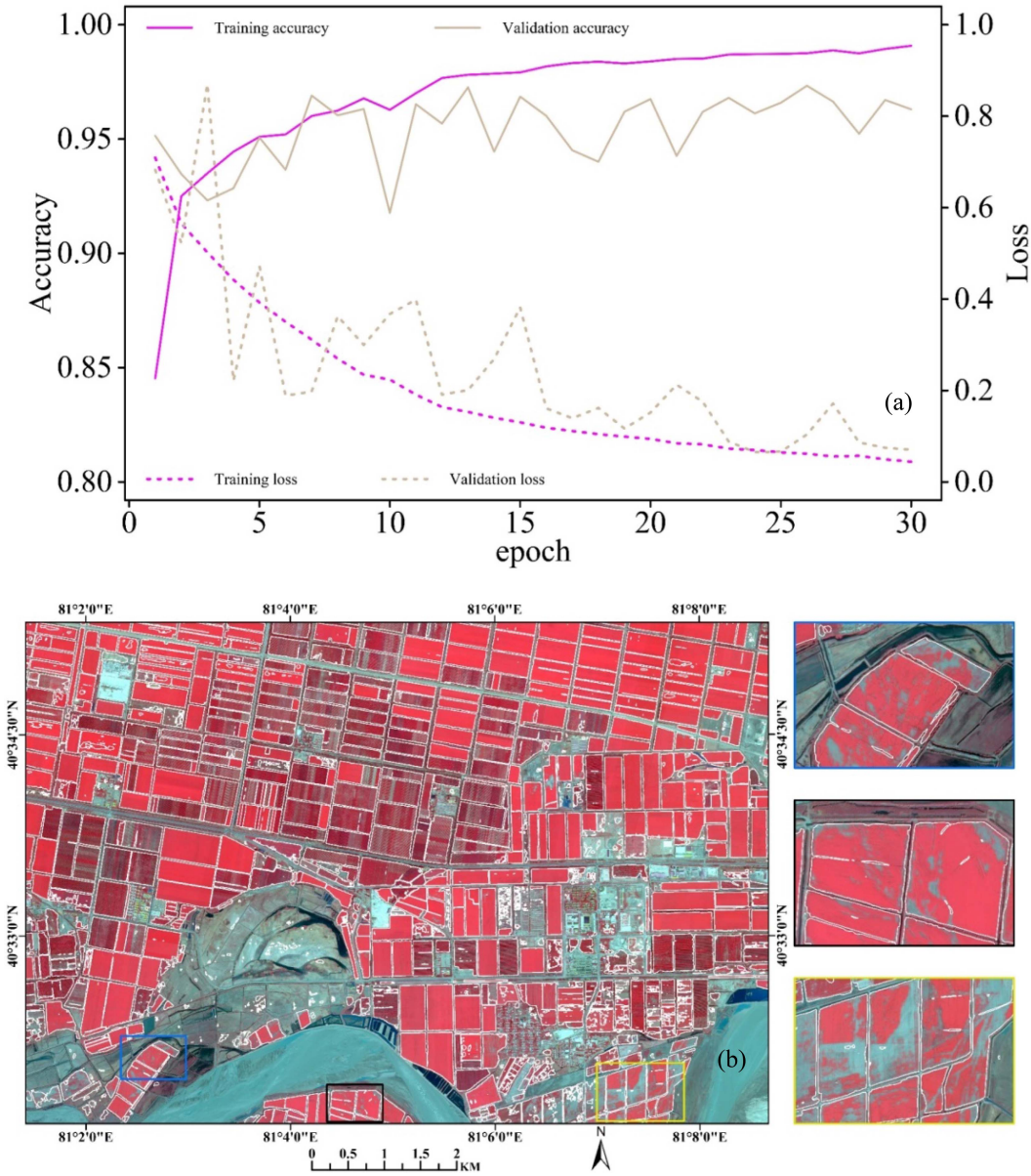

Fig. 7. (a) Accuracy and loss of the training set and validation set. (b) Parcel segmentation based on a Gaofen-2 fused image using the trained U-Net model (the white lines represent the boundary of segmented parcels; the Gaofen-2 image was captured on August 12, 2021).

were cropland parcels and 1651 belonged to the other land-cover types [see Fig. 7(b)].

B. Crop Mapping Using the TWDTW Algorithm

We mapped crop types using the parcel-based and pixel-based TWDTW algorithms, and the results exhibited OAs of 99.64% and 94.26%, respectively (see Tables III and IV). The OA of the crop-classification results obtained using the parcel-based TWDTW algorithm was improved by 5.38% compared to that obtained using the pixel-based TWDTW classification method. Such an improvement confirms that the parcel-based TWDTW classification method has a better performance when identifying crops from remote sensing images than the pixel-based TWDTW classification method. Meanwhile, the kappa coefficient of the crop-classification results obtained using the parcel-based TWDTW method (0.99) was improved by 8% compared to that obtained with the pixel-based TWDTW method (0.91).

The UA obtained for the cotton-type identification results using the parcel-based TWDTW method (UA = 100%) was improved by 11% compared to that obtained using the pixel-based TWDTW method (UA = 89%). In contrast, the PA of the parcel-based TWDTW method was only 1% lower than that of the pixel-based TWDTW method (see Tables III and IV). The UA and PA of the parcel-based TWDTW method in identifying orchard types improved by 7% and 1%, respectively, and both were higher than those of the pixel-based TWDTW method. We found no significant difference between the UAs of these two approaches when identifying noncrop types (UA = 99%)

TABLE III
ASSESSMENT OF THE CROP-CLASSIFICATION RESULTS OBTAINED WITH THE PIXEL-BASED TWDTW METHOD

	Cotton	Non-crop	Orchard	Unclassified	Total	UA(%)
Cotton	65	8	0	0	73	89.00
Non-crop	0	99	1	0	100	99.00
Orchard	0	7	99	0	106	93.00
Unclassified	0	0	0	0	0	100.00
Total	65	114	100	0	279	
PA(%)	100.00	85.00	99.00	100.00		
OA(%)						94.26
Kappa						0.91

TABLE IV
ASSESSMENT OF THE CROP-CLASSIFICATION RESULTS OBTAINED WITH THE PARCEL-BASED TWDTW METHOD

	Cotton	Non-crop	Orchard	Unclassified	Total	UA(%)
Cotton	64	0	0	0	64	100.00
Non-crop	1	114	0	0	115	99.00
Orchard	0	0	100	0	100	100.00
Unclassified	0	0	0	0	0	100.00
Total	0	114	100	0	279	
PA(%)	99.00	100.00	100.00	100.00		
OA(%)						99.64
Kappa						0.99

but observed a PA increase of 15% when using the TWDTW algorithm from the pixel-based to the parcel-based levels. The UAs and PAs of all types derived from the parcel-based TWDTW method improved compared to those derived from the pixel-based TWDTW method (see Tables III and IV).

The crop-mapping results obtained by incorporating Gaofen-2 imagery with spatial features and Sentinel-2 time series with spectral-temporal features were spatially finer and more accurate than the crop-classification results obtained using Sentinel-2 data alone (see Fig. 8). The pixel-based approach may misclassify areas with sparse crop growth within croplands into other types although these misclassified areas are part of the cropland (see Fig. 8). In contrast, by averaging the EVI values within each parcel, combining pixels as objects can help reduce salt-and-pepper noise in the pixel-based classification process and improve the spatial fragmentation of the classification results (see Fig. 8). Fig. 8 demonstrates that the parcel-based crop-type mapping results were more spatially homogeneous than the results obtained from the pixel-based method. Thus, the developed approach effectively prevents the noise problem caused by the presence of pixels with different attribute information within the same cropland parcel (see Fig. 9).

V. DISCUSSION

A. Geoparcel Segmentation Based on the U-Net Method and Gaofen-2 Image

For large-scale geoparcel-based crop mapping, accurately segmenting agricultural parcel objects from satellite imagery is an unavoidable difficulty [5], [12]. This difficulty comes mainly from two aspects: 1) the accessibility of high-spatial-resolution remote sensing images for site-specific crop mapping and 2) the indispensability of excellent segmentation algorithms. In this study, we selected the Gaofen-2 image to segment agricultural land parcels and obtained precise cropland edges. In addition, to address this difficulty, we employed the U-Net model to segment the cropland parcels from the Gaofen-2 image covering the study area. It is trained from scratch using a training dataset collected near the study area. After testing, the U-Net model achieved remarkable accuracy in agricultural parcel segmentation (see Table II). The high precision and recall scores indicate the model's capability to correctly identify agricultural parcels. Furthermore, the U-Net model is trained faster due to the use of cross-layer connectivity and pooling operations, which reduce the number of parameters in the network [see Fig. 7(a)]. Moreover, the U-Net model produced well-delineated boundaries for regions characterized by homogeneous parcels and heterogeneous parcels and was able to deal with the spectral heterogeneity between different parcels [see Fig. 7(b)]. The result of agricultural parcel boundaries extracted from a high-resolution image using U-Net shows that the edges of the parcels can be precisely distinguished from roads and farmland shelterbelts (see Fig. 7). Thus, the U-Net model exhibited robustness to data heterogeneity commonly encountered in agricultural parcel segmentation tasks. Our findings also share similarities with those of Waldner and Diakogiannis [22], whose results show that U-Net can actively learn to distinguish edges that do not belong to cultivated land. These results highlight the potential of the

Fig. 8. (a) Crop-type mapping using the pixel-based TWDTW method and (b) the parcel-based TWDTW method (the blue dotted rectangles represent the comparison of the classification effect on the same position under two approaches).

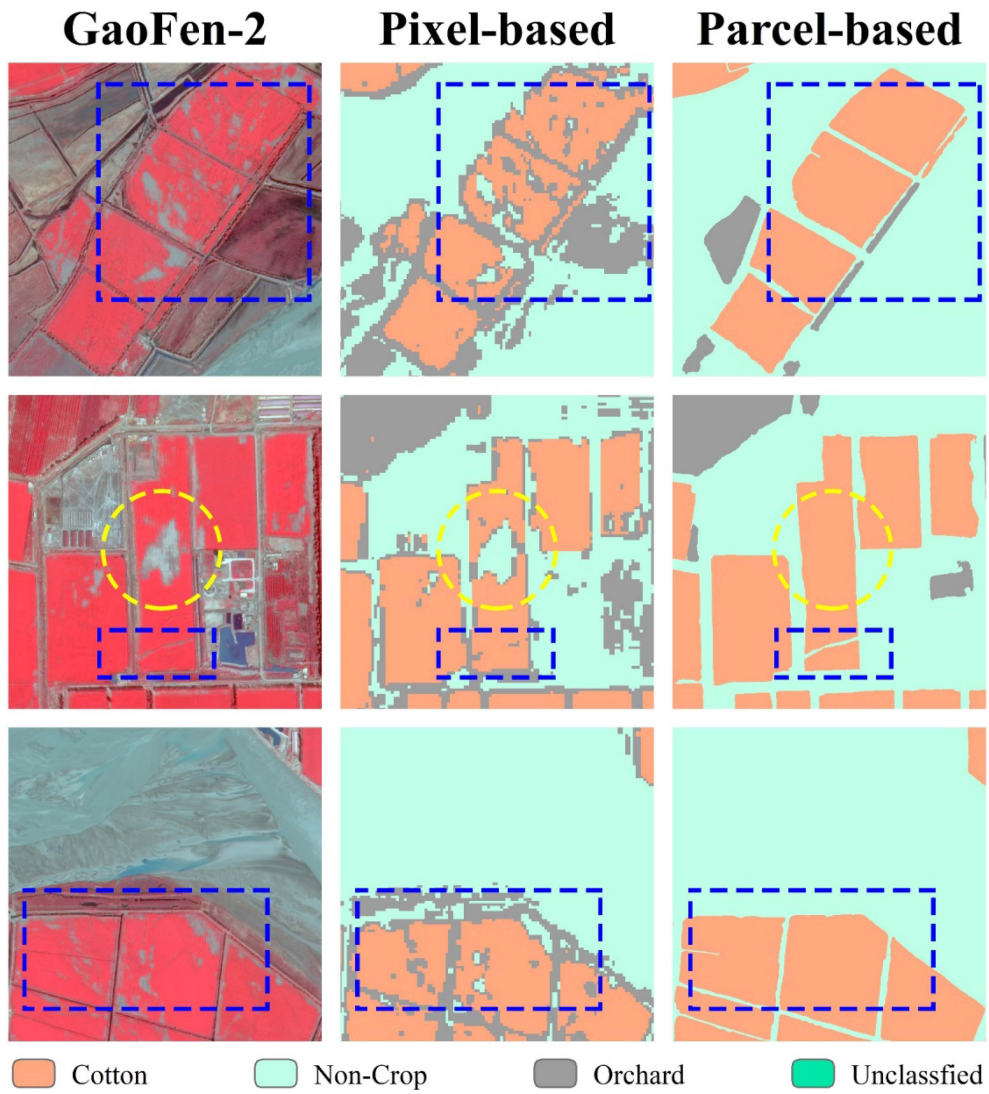

Fig. 9. Comparison of the crop-mapping results obtained using the pixel-based and parcel-based TWDTW methods (the dashed boxes indicate the comparison of different classification effects in the same position).

U-Net model as a reliable tool for precise parcel segmentation in agricultural applications.

The key difference from the study by Belgiu and Csillik [5] is the different methods implemented to segment agricultural parcels. They employed eCognition software to automatically segment agricultural parcels from Sentinel-2 images; however, the method did not work well in extracting agricultural parcels in this study. Similarly, Wu et al. [29] found that it was difficult to extract pure cropland edges from high spatial resolution images using eCognition and found that the extracted cropland parcels were likely to contain noncrop types (e.g., roads) and other problems. This incomplete segmentation of cropland may cause confusion in the crop classification results.

In contrast to the high-resolution image used in this study, Zhang et al. [27] used U-Net to extract cropland and found that the spatial resolution of Sentinel-2 images could not accurately depict fragmented parcels and could not clearly distinguish edges between adjacent and homogeneous fields. Masoud et al. [30] similarly believe that the 10 m resolution of Sentinel-2 images is relatively coarse for extracting cropland edges, which may also lead to parcel segmentation results appearing to be confused with noncrop types.

B. Advantage of the Parcel-Based TWDTW Method for Crop Mapping

The spectral heterogeneity of crop growth within a cropland parcel in remote sensing imagery is affected by various environmental constraints and can lead to crop misidentifications. Thus, we proposed an approach to address the above-mentioned crop-classification problem by integrating high-resolution remote sensing image-generated field edges based on U-Net segmentation and the precise spatiotemporal features derived from Sentinel-2 time series based on the TWDTW classification algorithm. The results showed that using cropland parcels as the basic analysis unit to map crop types not only yielded finer and more homogeneous results than the pixel-based approach in terms of the spatial distribution (see Fig. 9) but also outperformed the

pixel-based approach in terms of the final classification accuracy (see Tables III and IV). Moreover, the crop-classification results obtained using the parcel-based approach reflected fine semantic features. Our findings are generally consistent with the results reported by Liu et al. [46], who extracted precise cropland parcels in mountainous areas using high-resolution remote sensing images to effectively obtain the actual field edges of croplands. Thus, this study provides a promising approach that fully considers spatial and spectral-temporal features from multisource remote sensing images when mapping crop types.

In addition to the problem of spectral heterogeneity within croplands, the misclassification problem caused by spectral confusion at cropland edges is very prominent. At field edges, the presence of farmland shelterbelts with phenological characteristics similar to those of orchards can lead to pixels that should be classified as noncrops being misidentified as orchards when using the pixel-based TWDTW method (see Fig. 9). The parcel-based TWDTW method can not only accurately extract the field edges of croplands but can also differentiate the noncrop planted pixels outside the cropland edges from the crop-planted pixels within the cropland, thus minimizing the spectral confusion problem.

The largest differences between the parcel-based and pixel-based TWDTW methods were observed in the cotton and noncrop classification results (11% difference in the UA for cotton and 15% difference in the PA for noncrops) (see Tables III and IV). As mentioned previously, the spatial heterogeneity arising from crop growth conditions in cotton parcels raises the issue of significant spectral heterogeneity occurring within the same crop type. Moreover, phenological pattern confusion among farmland shelterbelt types belonging to the noncrop type and orchards at cropland edges can cause errors in which farmland shelterbelts are misidentified as orchard land. Unfortunately, the pixel-based approach cannot handle these problems well since it does not consider the relationships among surrounding pixels [5]. In contrast, the parcel-based approach effectively alleviates this problem by averaging the spectral values of all pixels within each parcel. Ultimately, the crop-classification performance of the TWDTW method at the parcel level is significantly superior to that at the pixel level. This finding is in line with the results reported by Belgiu and Csillik [5], who argued that grouping pixels as objects reduces spectral heterogeneity within the same type.

In addition, the TWDTW approach calculates the dissimilarity value between each pixel or parcel and the time-series pattern of a given land type [14]. A TWDTW dissimilarity value closer to 0 means that the pixel or parcel has better similarity to the given time-series pattern and vice versa [12], [44]. In this study, the dissimilarity values obtained for the pixel-based TWDTW approach ranged from 0.63 to 16.32 (see Fig. 10), whereas the dissimilarity values for the parcel-based TWDTW approach ranged from 0.007 to 0.101 (see Fig. 10). These results show that the parcel-based TWDTW approach smoothed the spectral values of pixels with high heterogeneity by averaging the spectral values of all pixels within and at the edges of each parcel. Thus, the dissimilarity value at the parcel-based level was relatively close to the given time-series pattern, thus significantly improving the resulting crop-mapping accuracy. The TWDTW dissimilarity value results at the parcel-based level were like those derived from the object-based approach reported by Csillik et al. [12], whose object-based TWDTW dissimilarity values were also spatially smooth.

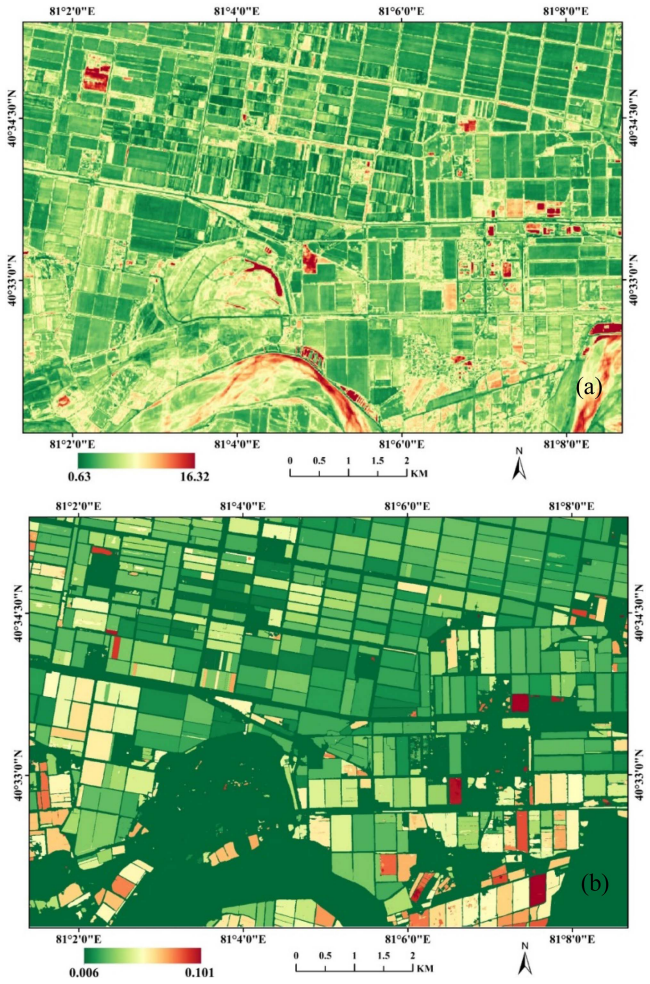

Fig. 10. TWDTW dissimilarity values at (a) the pixel-based level and (b) the parcel-based level.

C. Recommendations and Limitations

The results showed that the PA of the cotton-identification results obtained using the parcel-based TWDTW method was only 1% lower than that obtained using the pixel-based TWDTW method. The reason for this difference was that the average spectral-temporal patterns of all pixels in a few cotton parcels with extremely high spectral heterogeneity were like the patterns of the noncrop land type, thus leading to cotton being misidentified as the noncrop land type when using the parcel-based TWDTW method (see Fig. 11). In contrast, some cotton pixels could still be identified using the pixel-based TWDTW method (see Fig. 11). Although applying the mean value of pixels within a parcel can effectively eliminate heterogeneous disturbances in the cropland parcel, it does not appear to be an effective

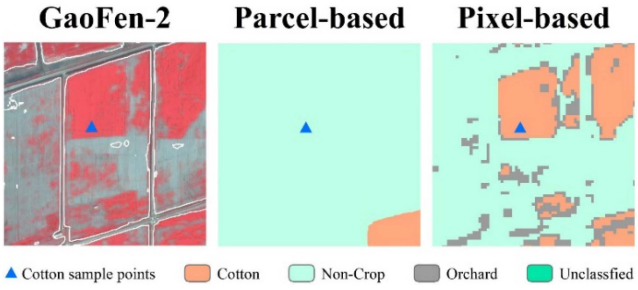

Fig. 11. Comparison of the cotton-mapping results obtained using the pixel-based and parcel-based TWDTW methods (the white line represents the cropland edges extracted using the U-Net model).

solution for a few parcels with extreme heterogeneity. Therefore, in addition to the mean values, other additional statistical values of phenological features (including maximum, percentiles, and standard deviation values calculated from EVI data) can also be considered for parcel-based crop classification. In addition, we mainly selected pixels with relatively low heterogeneity among the cropland parcels as classification training and validation samples and lacked samples closer to the edge of the cropland parcels (see Fig. 1). Therefore, the classification errors on the edges of cropland parcels could not be verified regardless of how the training and validation samples were randomly assigned, which led to the inability to quantitatively characterize the apparent accuracy gap between the parcel-based and pixel-based crop classifications. However, we compared the results obtained by the two classification methods regarding the spatial morphology. The results showed that the misclassification problem on parcel edges could be effectively resolved by using the parcel-based TWDTW method. Here, we recommend selecting some sample points closer to the edges of the farmland, which can better quantify the accuracy gap between parcel-based and pixel-based crop classification.

Although a combination of high-spatial-resolution and multispectral data effectively improved the accuracy and fineness of crop-type mapping in this study, we found that this approach significantly increased the number of pixels to be processed (equivalent to 100 times the number of pixels) and thus increased the computation time required for the TWDTW classification process; this increase can seriously affect the crop-mapping efficiency. Despite this shortcoming, our research takes full advantage of multisource satellite platforms and provides a potential method for finely mapping crops over the areas affected by various environmental constraints. In future work, we will focus on large-scale applications of this method by effectively overcoming this computational inefficiency disadvantage.

VI. Conclusion

To eliminate spectral heterogeneity effects on the inside and at the edges of croplands on the crop identification results, in this article, we proposed a method that can fully utilize spatial-temporal-spectral features from multisource satellite data to map crop types in arid regions. In this method, first, cropland parcels are segmented and extracted using the U-Net segmentation model and a Gaofen-2 high-resolution remote sensing image; then, the TWDTW algorithm is employed to perform crop mapping at the segmented parcel scale using time-series Sentinel-2 multispectral images. The results showed that the proposed approach had a better crop-mapping performance than the pixel-based method, with an increased OA of 5.38% compared to the results of the pixel-based method. In addition, the approach at the parcel-based scale yielded a map with relatively fine and homogeneous crop types and effectively solved the crop misidentification problem on the inside and at the edges of croplands. This study provides a practical solution for fine-scale crop mapping by reducing the interference effects of the spatial crop-growth heterogeneity that are caused by various environmental constraints within croplands on the corresponding remote sensing spectral features. The proposed method exhibited robust performance and can be applied to other regions in the world.

Acknowledgment

The authors are grateful to the European Space Agency for sharing Sentinel-2 remote sensing images and the China Centre for Resources Satellite Data and Application for sharing Gaofen-2 remote sensing images. The authors are also grateful for four anonymous reviewers.

References

[1] G. Li et al., "Crop type mapping using time-series Sentinel-2 imagery and U-Net in early growth periods in the Hetao irrigation district in China," *Comput. Electron. Agriculture*, vol. 203, Dec. 2022, Art. no. 107478.

[2] H. X. Zhang et al., "Crop classification and acreage estimation in North Korea using phenology features," *GISci. Remote Sens.*, vol. 54, no. 3, pp. 381–406, Jun. 2017.

[3] R. Sonobe, Y. Yamaya, H. Tani, X. F. Wang, N. Kobayashi, and K. I. Mochizuki, "Assessing the suitability of data from Sentinel-1A and 2A for crop classification," *GISci. Remote Sens.*, vol. 54, no. 6, pp. 918–938, 2017.

[4] Y. N. Zhou, J. C. Luo, L. Feng, Y. P. Yang, Y. H. Chen, and W. Wu, "Long-short-term-memory-based crop classification using high-resolution optical images and multi-temporal SAR data," *GISci. Remote Sens.*, vol. 56, no. 8, pp. 1170–1191, Nov. 2019.

[5] M. Belgiu and O. Csillik, "Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis," *Remote Sens. Environ.*, vol. 204, pp. 509–523, Jan. 2018.

[6] G. W. Gella, W. Bijker, and M. Belgiu, "Mapping crop types in complex farming areas using SAR imagery with dynamic time warping," *ISPRS J. Photogramm. Remote Sens.*, vol. 175, pp. 171–183, May 2021.

[7] W. S. Moola, W. Bijker, M. Belgiu, and M. M. Li, "Vegetable mapping using fuzzy classification of dynamic time warping distances from time series of Sentinel-1A images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 102, Oct. 2021, Art. no. 102405.

[8] Y. M. Wang et al., "Exploring the potential of multi-source unsupervised domain adaptation in crop mapping using Sentinel-2 images," *GISci. Remote Sens.*, vol. 59, no. 1, pp. 2247–2265, Dec. 2022.

[9] D. Ashourloo et al., "Automatic canola mapping using time series of Sentinel 2 images," *ISPRS J. Photogramm. Remote Sens.*, vol. 156, pp. 63–76, Oct. 2019.

[10] D. Bargiel, "A new method for crop classification combining time series of radar images and crop phenology information," *Remote Sens. Environ.*, vol. 198, pp. 369–383, Sep. 2017.

[11] H. Y. Zhang, J. Z. Kang, X. Xu, and L. P. Zhang, "Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yi'an County, Heilongjiang Province, China," *Comput. Electron. Agriculture*, vol. 176, Sep. 2020, Art. no. 105618.

[12] O. Csillik, M. Belgiu, G. P. Asner, and M. Kelly, "Object-based time-constrained dynamic time warping classification of crops using Sentinel-2," *Remote Sens.*, vol. 11, no. 10, May 2019, Art. no. 1257.
[13] C. Pelletier, G. I. Webb, and F. Petitjean, "Temporal convolutional neural network for the classification of satellite image time series," *Remote Sens.*, vol. 11, no. 5, Mar. 2019, Art. no. 523.
[14] V. Maus, G. Camara, R. Cartaxo, A. Sanchez, F. M. Ramos, and G. R. de Queiroz, "A time-weighted dynamic time warping method for land-use and land-cover mapping," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 8, pp. 3729–3739, Aug. 2016.
[15] C. C. Li, G. Xian, Q. Zhou, and B. W. Pengra, "A novel automatic phenology learning (APL) method of training sample selection using multiple datasets for time-series land cover mapping," *Remote Sens. Environ.*, vol. 266, Dec. 2021, Art. no. 112670.
[16] M. Belgiu, W. Bijker, O. Csillik, and A. Stein, "Phenology-based sample generation for supervised crop type classification," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 95, Mar. 2021, Art. no. 102264.
[17] M. M. Li and W. Bijker, "Vegetable classification in Indonesia using dynamic time warping of Sentinel-1A dual polarization SAR time series," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 78, pp. 268–280, Jun. 2019.
[18] F. Zhao et al., "Determination of key phenological phases of winter wheat based on the time-weighted dynamic time warping algorithm and MODIS time-series data," *Remote Sens.*, vol. 13, no. 9, May 2021, Art. no. 1836.
[19] S. Moharana, B. Kambhammettu, S. Chintala, A. S. Rani, and R. Avtar, "Spatial distribution of inter- and intra-crop variability using time-weighted dynamic time warping analysis from Sentinel-1 datasets," *Remote Sens. Appl., Soc. Environ.*, vol. 24, 2021, Art. no. 100630.
[20] N. Zhou, X. Li, Z. Shen, T. Wu, and J. Luo, "Geo-parcel-based change detection using optical and SAR images in cloudy and rainy areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, no. 1, pp. 1326–1332, Jan. 2021.
[21] J. A. Long, R. L. Lawrence, M. C. Greenwood, L. Marshall, and P. R. Miller, "Object-oriented crop classification using multitemporal ETM plus SLC-off imagery and random forest," *GISci. Remote Sens.*, vol. 50, no. 4, pp. 418–436, Aug. 2013.
[22] F. Waldner and F. I. Diakogiannis, "Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network," *Remote Sens. Environ.*, vol. 245, Aug. 2020, Art. no. 111741.
[23] C. Persello, V. A. Tolpekin, J. R. Bergado, and R. A. de By, "Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping," *Remote Sens. Environ.*, vol. 231, Sep. 2019, Art. no. 111253.
[24] W. Zhang, P. Tang, and L. J. Zhao, "Fast and accurate land cover classification on medium resolution remote sensing images using segmentation models," *Int. J. Remote Sens.*, vol. 42, no. 9, pp. 3277–3301, May 2021.
[25] W. C. Xu, P. C. Chen, Y. L. Zhan, S. D. Chen, L. Zhang, and Y. B. Lan, "Cotton yield estimation model based on machine learning using time series UAV remote sensing data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 104, Dec. 2021, Art. no. 102511.
[26] L. J. Wang, J. Y. Wang, X. W. Zhang, L. G. Wang, and F. Qin, "Deep segmentation and classification of complex crops using multi-feature satellite imagery," *Comput. Electron. Agriculture*, vol. 200, Sep. 2022, Art. no. 107249.
[27] H. X. Zhang et al., "Automated delineation of agricultural field boundaries from Sentinel-2 images using recurrent residual U-Net," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 105, Dec. 2021, Art. no. 102557.
[28] J. O'Connell, U. Bradter, and T. G. Benton, "Wide-area mapping of small-scale features in agricultural landscapes using airborne remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 109, pp. 165–177, Nov. 2015.
[29] M. Q. Wu et al., "Fine crop mapping by combining high spectral and high spatial resolution remote sensing data in complex heterogeneous areas," *Comput. Electron. Agriculture*, vol. 139, pp. 1–9, Jun. 2017.
[30] K. M. Masoud, C. Persello, and V. A. Tolpekin, "Delineation of agricultural field boundaries from Sentinel-2 images using a novel super-resolution contour detector based on fully convolutional networks," *Remote Sens.*, vol. 12, no. 1, Jan. 2020, Art. no. 59.
[31] M. Q. Wu et al., "Reconstruction of daily 30 m data from HJ CCD, GF-1 WFV, Landsat, and MODIS data for crop monitoring," *Remote Sens.*, vol. 7, no. 12, pp. 16293–16314, Dec. 2015.
[32] Q. Q. Li, G. L. Liu, and W. J. Chen, "Toward a simple and generic approach for identifying multi-year cotton cropping patterns using Landsat and Sentinel-2 time series," *Remote Sens.*, vol. 13, no. 24, Dec. 2021, Art. no. 5183.
[33] L. Gong, G. X. He, and W. G. Liu, "Long-term cropping effects on agricultural sustainability in Alar oasis of Xinjiang, China," *Sustainability*, vol. 8, no. 1, Jan. 2016, Art. no. 61.
[34] F. M. Chen et al., "The status quo of desertification and the prevention strategy in Xinjiang," *J. Food Agriculture Environ.*, vol. 11, no. 2, pp. 1025–1032, 2013.
[35] J. Long, M. M. Li, X. Q. Wang, and A. Stein, "Delineation of agricultural fields using multi-task BsiNet from high-resolution satellite images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 112, Aug. 2022, Art. no. 102871.
[36] Y. W. Sun et al., "Geo-parcel-based crop classification in very-high-resolution images via hierarchical perception," *Int. J. Remote Sens.*, vol. 41, no. 4, pp. 1603–1624, Feb. 2020.
[37] L. B. Yang, L. M. Wang, G. A. Abubakar, and J. F. Huang, "High-resolution rice mapping based on SNIC segmentation and multi-source remote sensing images," *Remote Sens.*, vol. 13, no. 6, Mar. 2021, Art. no. 1148.
[38] J. Liu et al., "Geometric correction of GF-1 satellite images based on block adjustment of rational polynomial model," *Trans. Chin. Soc. Agricultural Eng.*, vol. 31, no. 22, pp. 146–154, 2015.
[39] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," U.S. Patent 6 011 875, Jan. 4, 2000.
[40] J. Chen, P. Jonsson, M. Tamura, Z. H. Gu, B. Matsushita, and L. Eklundh, "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter," *Remote Sens. Environ.*, vol. 91, no. 3/4, pp. 332–344, Jun. 2004.
[41] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sens. Environ.*, vol. 83, no. 1/2, pp. 195–213, Nov. 2002.
[42] U. Meier, *Growth Stages of Mono- and Dicotyledonous Plants: BBCH Monograph*. Quedlinburg, Germany: Julius Kühn-Institut, 2018.
[43] H. X. Li, C. Z. Wang, Y. X. Cui, and M. Hodgson, "Mapping salt marsh along coastal South Carolina using U-Net," *ISPRS J. Photogramm. Remote Sens.*, vol. 179, pp. 121–132, Sep. 2021.
[44] V. Maus, G. Camara, M. Appel, and E. Pebesma, "dtwSat: Time-weighted dynamic time warping for satellite image time series analysis in R," *J. Statist. Softw.*, vol. 88, no. 5, pp. 1–31, Jan. 2019.
[45] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, pp. 159–174, 1977.
[46] W. Liu et al., "Farmland parcel mapping in mountain areas using time-series SAR data and VHR optical images," *Remote Sens.*, vol. 12, no. 22, Nov. 2020, Art. no. 3733.

Weijia Chen received the B.E. degree in geographic information science from Chuzhou University, Chuzhou, China, in 2020. He is currently working toward the M.S. degree in cartography and geographic information system with the School of Geography, South China Normal University, Guangzhou, China.

Guilin Liu received the Ph.D. degree in environmental remote sensing and geoinformatics from Universität Trier, Trier, Germany, in 2017.

He is currently an Associate Professor with the School of Geography, South China Normal University, Guangzhou, China. His research interests include agricultural remote sensing, land use change, crop model and remote sensing data assimilation, and time-series analysis.