

# The 10-meter Winter Wheat Mapping in Shandong Province Using Sentinel-2 Data and Coarse Resolution Maps

Xiaoman Qi<sup>1</sup>, Yuebin Wang, *Member, IEEE*, Junhuan Peng<sup>2</sup>, Liqiang Zhang<sup>3</sup>, *Member, IEEE*, Wenping Yuan, and Xiaotong Qi

**Abstract**—Timely and accurate large-scale mapping of the spread of winter wheat (*Triticum aestivum*) is crucial to guarantee food security, study climate change, and monitor operational agriculture. Traditional winter wheat mapping frameworks are constrained by insufficient spatial resolution and heavy dependence on field surveys, while traditional machine learning models excessively rely on subjective judgment. Furthermore, collecting sufficient field samples covering a large area is expensive and time-consuming. In this context, an automatic label update deep learning solution is developed to produce 10-m resolution winter wheat maps using Sentinel-2 data and existing coarse-resolution (30 m) winter wheat mapping products. In particular, a label update module considering the unique phenological (seasonal) characteristics of winter wheat is designed to update labels in the training phase. The results indicate that our method yields a satisfactory classification result with an overall accuracy exceeding 92% and an  $F_1$  score greater than 0.85 for all validation samples, even when no field survey data were used for training. In addition, a 10-m spatial resolution winter wheat map for the entire Shandong province is generated, showing a significant correlation between the computed winter wheat map and the agricultural statistical land, with correlation coefficients of 0.95 and 0.78 at the municipal and county levels, respectively. The proposed methodology can serve as a viable and promising method for high-resolution, operational agricultural monitoring over large areas.

**Index Terms**—Deep learning, mapping, noise correction, sentinel-2, winter wheat.

Manuscript received 24 June 2022; revised 4 October 2022; accepted 5 November 2022. Date of publication 8 November 2022; date of current version 16 November 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 42074004; in part by Shanxi Transportation Holdings Group CO., LTD under Grant 19-JKKJ-3; and in part by the Natural Science Foundation of the Higher Education Institutions of Jiangsu Province, China under Grant 21KJB420002. (Xiaoman Qi and Yuebin Wang contributed equally to this work.) (Corresponding author: Junhuan Peng.)

Xiaoman Qi, Yuebin Wang, and Junhuan Peng are with the School of Land Science and Technology, China University of Geosciences, Beijing 100083, China (e-mail: qxm@cugb.edu.cn; xxgcdxwyb@163.com; pengjunhuan@163.com).

Liqiang Zhang is with the Faculty of Geographical Science, State Key Laboratory of Remote Sensing Science, Beijing Normal University, Beijing 100875, China (e-mail: zhanglq@bnu.edu.cn).

Wenping Yuan is with the School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai 519000, China (e-mail: yuanwpcn@126.com).

Xiaotong Qi is with the School of Marine Technology and Geomatics, Jiangsu Ocean University, Lian Yungang 222005, China (e-mail: qixiaotong08@163.com).

Digital Object Identifier 10.1109/JSTARS.2022.3220698

## I. INTRODUCTION

AS the second of the 17 Sustainable Development Goals of the United Nations, “Zero Hunger” aims to eliminate all forms of hunger and malnutrition by 2030, ensuring that all people, especially children, have continuous access to sufficient and nutritious food [1]. Given this demand, the agricultural sector now plays a more important role than ever before. Wheat, one of the world’s most extensively farmed food crops, is the primary source of carbohydrates for millions of people [2]. China is a major global wheat producer, accounting for 17.9% of the global wheat production with 11% of the worldwide wheat planting acreage [3]. Moreover, winter wheat accounts for 95% of China’s total wheat production [4]. Wheat sown areas and their yields are directly related to national food security and social stability. Thus, the assessment of winter wheat sowing areas is an integral part of wheat growth monitoring and yield estimation. Conventional methods like ground surveying are generally reliable methods of sowing area assessment, however, they are considered time-consuming and expensive methods for China, where the cropland is usually small and fragmented [5].

Due to the advancements in remote sensing technologies, satellite images offer effective means to quickly and accurately determine the spatial distribution of wheat cultivation. Significant efforts have been made to explore the distributions of crop types based on moderate resolution imaging spectroradiometers (MODIS) [6], [7], [8]. However, MODIS data have a relatively coarse spatial resolution (1000, 500, and 250 m), resulting in numerous mixed heterogeneous pixels. Thus, MODIS data are generally utilized for large-scale winter wheat mapping tasks. However, using MODIS data to perform accurate crop mapping in China, where the cropland is usually small and fragmented, is challenging [5]. To address this issue, various researchers have devised methods for distinguishing winter wheat from other various land cover types based on Landsat imagery with a spatial resolution of 30 m, which is substantially higher than that of MODIS data [9], [10], [11]. However, the temporal resolution of Landsat data is also too coarse for accurate crop mapping. Therefore, Sentinel-2 data, which allows for frequent land monitoring at a spatial resolution of 10 m with a revisit frequency of five days, have been employed in multiple studies for reliable crop mapping [12], [13], [14]. The use of sentinel-2 data could acquire

high classification performance for small agriculture fields, thus reducing the statistical errors due to low resolution. However, operationally generated annual winter wheat mapping products with a resolution of 10 m are difficult to produce on a large scale. This issue stems from the collection of field training samples not being conducted regularly. Therefore, it is necessary to explore a mapping method that does not rely on a large number of field samples.

According to the literature cited earlier, regardless of the remote sensing data used, the most prevalent strategy to map winter wheat is to link satellite-derived features and winter wheat classification methods by utilizing algorithms and training samples. Many supervised classification algorithms have been successfully implemented, yielding acceptable prediction performance, including support vector machine (SVM) [11], [15], random forest (RF) [14], [16], decision tree (DT) [9], [17], and neural network (NN) classifiers [10], [18], [19]. However, in these machine learning methods, feature extraction and processing are highly reliant on a user's subjective judgement, potentially resulting in biased and inaccurate classification [20]. Problems tied to limitations linked to subjective judgement can be circumvented using NN. NN is an end-to-end approach that can automatically extract discriminating features that have not only been used in numerous tasks but that have also consistently exhibited satisfactory performance [19], [21], [22], [23]. More importantly, these supervised classification methods rely largely on accurate and reliable training samples, typically gathered by *in situ* surveying or visual interpretation of remote sensing images with excellent spatial resolution. In general, field samples should be collected regularly for crop type mapping; however, *in situ* surveying is labor-intensive and uneconomical. Therefore, routinely obtaining ground-surveyed samples is challenging, significantly limiting the implementation of these approaches, particularly for large-area mapping [24], [25].

As an alternative approach, several studies have attempted to accomplish refined and accurate land cover mapping using low-resolution and freely available labels [26], [27], [28], [29], [30], [31]. For instance, Malkin et al. [28] introduced a label super-resolution network based on the joint distribution of low- and high-resolution labels. Although a noise-robust approach seeks to lessen the effects of noise, it is challenging to completely avoid the influence of noisy labels during training. Accordingly, Zhu et al. [31] devised a method called "Pick-and-Learn" (PAL) for automatically evaluating the relative quality of the noisy labels in the training set and tuning the network parameters to ensure that the relevant ones are used. Dong et al. [26] proposed a novel approach that jointly optimizes the model parameters and corrects the noisy label with a "synergistic noise correction loss," which reduces the impact of noisy labels, improving classification performance. However, in this study, only the class probability distribution was used to determine which labels need to be updated. When faced with a variety of real-world problems, researchers are recommended to account for the peculiarities of the individual noisy scenarios to identify the best answers, for example, using unique phenological characteristics to update noisy labels for crop mapping.

Overall, various flaws can be identified in the current winter wheat mapping framework, such as the inadequate spatiotemporal resolution of remote sensing image data, subjective assessment in traditional machine learning algorithms, and a heavy dependence on field survey samples. To address these issues, the major objective of our work is developing a novel winter wheat mapping method based on high-resolution Sentinel-2 data and existing lower-resolution labels rather than expensive field training samples. Therefore, we proposed a framework based on Sentinel-2 imagery with a 10-m spatial resolution and an existing coarse-resolution (30 m) winter wheat mapping product and applied it across China's Shandong Province. This mapping method, named "automatic label update deep learning" (ALU-DL), is automated, easy to operate, and efficient. More importantly, ALU-DL could perform high-resolution mapping of winter wheat without using field samples to train the model. In this method, we leverage the advancements in the U-Net model [32] to provide a reliable and accurate high-resolution mapping methodology for winter wheat observation and management applications. The multilayered nature of deep learning architectures facilitates the gathering of spatiotemporal information from satellite data. Notably, a label update module that considers the unique phenological (seasonal) characteristics of winter wheat was designed to update the noisy labels in the training phase. Furthermore, the experiments suggest that the proposed framework could be effectively applied to distinguish winter wheat on a broad scale, with acceptable classification accuracy.

The main contributions of this study are as follows.

- 1) Considering the unique phenological (seasonal) characteristics of winter wheat, this study develops a novel winter wheat mapping method based on high-resolution Sentinel-2 data and lower resolution products rather than expensive field training samples.
- 2) Field validation and statistical data are used to evaluate the performance of the proposed model. The results show that our model significantly outperforms other conventional approaches.
- 3) A 10-m spatial resolution winter wheat map for Shandong province is produced based on the proposed method. The derived winter wheat cultivation area is consistent with municipal and county statistical data.

## II. STUDY AREA AND DATA

In this section, we describe the study area and properties of the data used for training and evaluation. Sentinel-2 satellite data with a 10-m spatial resolution were used as training data, and the 30-m resolution winter wheat mapping product was regarded as the training label. Field samples and statistical data were used to evaluate the performance of our approach.

### A. Study Area

The study was carried out in Shandong (SD) Province (Fig. 1), one of China's major food bowls and agricultural bases. SD is located approximately between 33 °N–38 °N latitude and 114 °E–122 °E longitude, with a wide area of 158 000 km<sup>2</sup>.

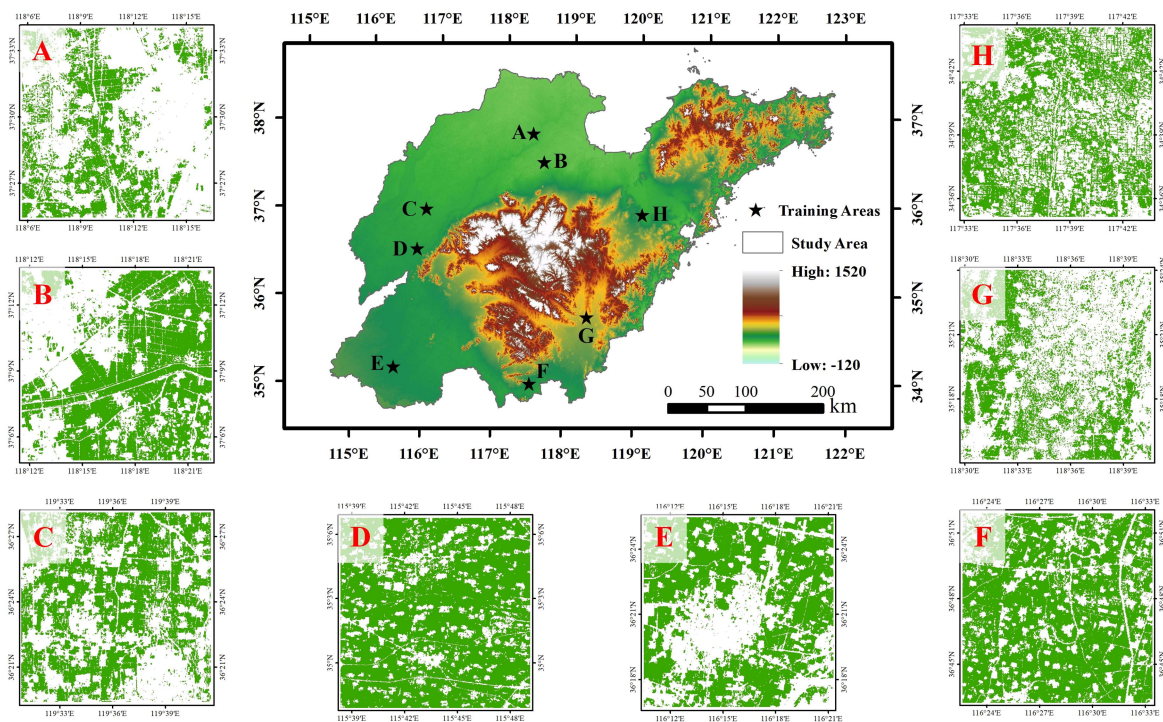


Fig. 1. Digital elevation model (DEM) maps of the Shandong Province. A–H represent the training areas. The winter wheat maps are sourced from Dong et al. [5].

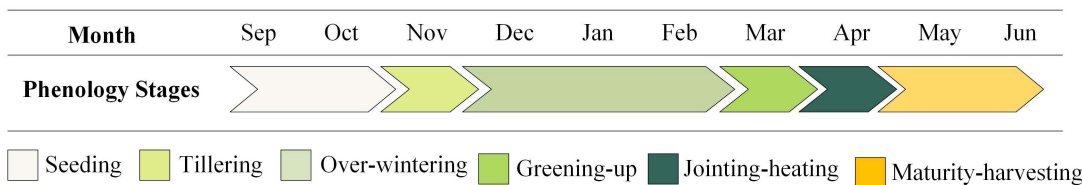


Fig. 2. Winter wheat phenological calendar of China.

SD has a warm temperate monsoon climate and annual precipitation ranging from 550 to 950 mm, with most of the rainfall concentrated in the summer. The annual average temperature of SD is 13.4 °C, with the lowest and highest temperatures typically occurring in January and July, respectively. The topography of SD is characterized by plains, with a few mountains in the center of the province. According to data from the National Bureau of Statistics (NBS), in 2018, the cultivated land area of SD was approximately 110 768 km<sup>2</sup>, accounting for approximately 70.11% of the province's land area. The primary crop planted in SD is winter wheat, with a sown area of 40 585.90 km<sup>2</sup>, accounting for approximately 17.85% of China's total winter wheat planted area.

To produce a relatively accurate and reliable model, we dispersedly selected eight 20 × 20 km training areas (A–H) representing the major grain-producing regions based on the principles of random and homogeneous distribution. The models were trained using the training areas data to learn distinctive feature representations of winter wheat fields, ensuring excellent spatial transferability and applicability across the entire province.

### B. Phenological Characteristic

In Shandong province, winter wheat is seeded from September to October and harvested from late May to June of the following year (Fig. 2) [33]. The most prominent difference between the winter wheat and other crops is that winter wheat overwinters, leading to the NDVI or EVI value of winter wheat from the day of year (DOY) initially increasing and then decreasing, and finally increasing in the spring of the following year, as shown in Fig. 3. In contrast, other crops are sown in spring and harvested in autumn [34]. Therefore, the phenological difference between winter wheat and background objects is notable. This discriminative feature can be extracted from remote sensing data, and it is feasible to discriminate winter wheat in this region.

### C. Sentinel-2 Data

Winter wheat crop area estimation was based on satellite data acquired and digitized by Sentinel-2, a high-resolution multi-spectral imaging satellite with a multispectral instrument (MSI) for land monitoring, which provides imagery of vegetation,

TABLE I  
 SPECTRAL BANDS FOR SENTINEL-2 SENSORS (MSI)

Sentinel-2 Bands	Central Wavelength ( $\mu\text{m}$ )	Resolution (m)
Band 1 – Coastal aerosol	0.443	60
Band 2 – Blue	0.490	10
Band 3 – Green	0.560	10
Band 4 – Red	0.665	10
Band 5 – Vegetation Red Edge	0.705	20
Band 6 – Vegetation Red Edge	0.740	20
Band 7 – Vegetation Red Edge	0.783	20
Band 8 – NIR	0.842	10
Band 8A – Vegetation Red Edge	0.865	20
Band 9 – Water vapor	0.945	60
Band 10 – SWIR - Cirrus	1.375	60
Band 11 – SWIR	1.610	20
Band 12 – SWIR	2.190	20

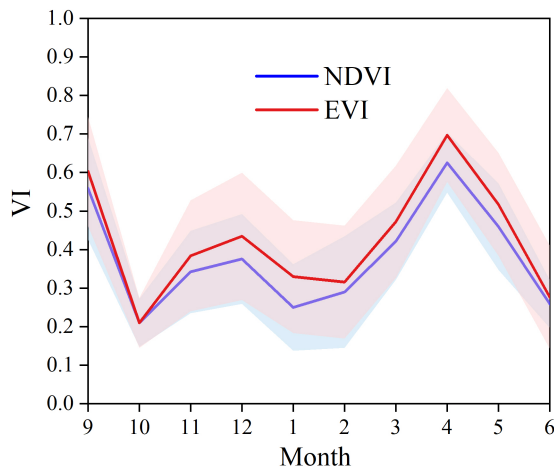


Fig. 3. Multitemporal NDVI and EVI profiles of winter wheat in Shandong province.

soil, water cover, inland waterways, coastal areas, and emergency rescue services [35]. Sentinel-2 consists of two satellites, Sentinel-2 A and Sentinel-2B. With a ten-day revisit interval for a single satellite and a five-day offset for two satellites, the observation period in the study area can be reduced to five days. The MSI, which was placed on Sentinel-2, collected 13 spectral bands in the visible and near-infrared (VNIR) and shortwave infrared (SWIR), with spatial resolutions ranging from 10 to 60 m (Table I). In this study, eight surface-related bands at 10- or 20-m resolution (B2-7, B11, and B12) and two commonly used vegetation indices (normalized difference vegetation index, NDVI; and enhanced vegetation index, EVI) were utilized for training the models. Note that all bands used were resampled at a resolution of 10 m via bicubic interpolation for further analysis.

For our study, the Google Earth Engine (GEE) platform was used to collect Sentinel-2 satellite images. Because surface reflectance (SR) data were unavailable on the GEE platform, we alternatively utilized the top of the atmosphere (TOA) reflectance data. The S2 TOA archived in the GEE was corrected radiometrically and geometrically, including orthorectification and spatial registration on a global reference system with sub-pixel accuracy [36]. Although SR data are more persuasive than TOA reflectance data, earlier research has demonstrated that TOA data are more reliable for crop type classification. For

instance, Emelyanova et al. [37] evaluated classification performance between TOA reflectance and SR data, and observed that the datasets performed similarly. Other researchers [14], [36], [38] employed TOA reflectance data to classify crops, obtaining reliable results. In this situation, in the absence of SR data, the classification results based on TOA reflectance data could be utilized to assess the performance of the proposed method.

Sentinel-2 TOA reflectance data include a QA60 band, which uses spectral criteria to detect opaque and cirrus clouds. Because the existing QA60 band cannot give reliable cloud detection results, we chose a low cloud cover criterion (20%) to filter and collect the Sentinel-2 imagery; this was done to limit omission errors in cloud/cloud shadow detection [39]. Therefore, a total of 1582 Sentinel-2 scenes were collected in the study area during the study period (9/2017–6/2018) from the GEE platform and the results are shown, as in Fig. 4. The statistical results indicated that most of the study area had a high availability of Sentinel-2 imagery, with  $> 30$  scenes during the study period [Fig. 4(a)]. More than 100 Sentinel-2 observations for each month were selected for this study [Fig. 4(b)].

#### D. Winter Wheat Distribution Map of China at 30-m Resolution

Dong et al. [5] constructed 30-m spatial resolution winter wheat distribution maps for the 2016–2018 period using the time-weighted dynamic time warping approach and monthly maximum composite NDVI from Landsat and Sentinel. The results reported based on the field samples revealed an overall accuracy (OA) of 89.88%.

In addition, the 30-m winter wheat distribution data were used as targets to train the NN. Thus, the 30-m winter wheat distribution map was resampled to a 10-m resolution to maintain consistency with satellite data. However, since resampled labels inevitably have numerous labeling errors, we treated resampled label data as noisy labels. Note that noise is caused by missing details in low-resolution labels and misidentification of different land cover categories across a large area.

#### E. Field Data

To evaluate the performance of our method, we collected 730 ground reference samples (Fig. 5), comprising 65 field survey samples and 665 Google Earth samples. The georeferenced field

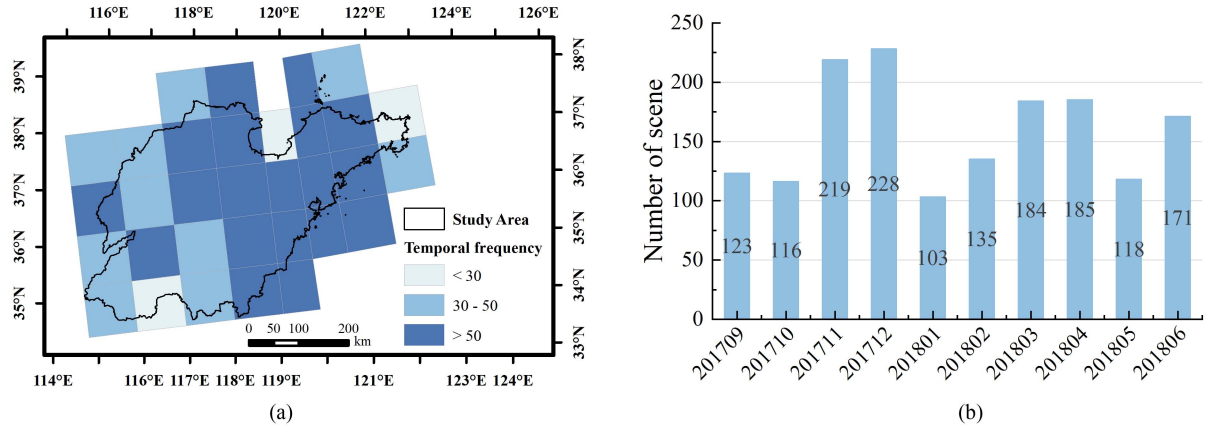


Fig. 4. Availability of Sentinel-2 data in the study area during the study period. (a) Temporal frequency of Sentinel-2 scenes used in this study. (b) Temporal distribution of Sentinel-2 scenes used in this study (total 1582).

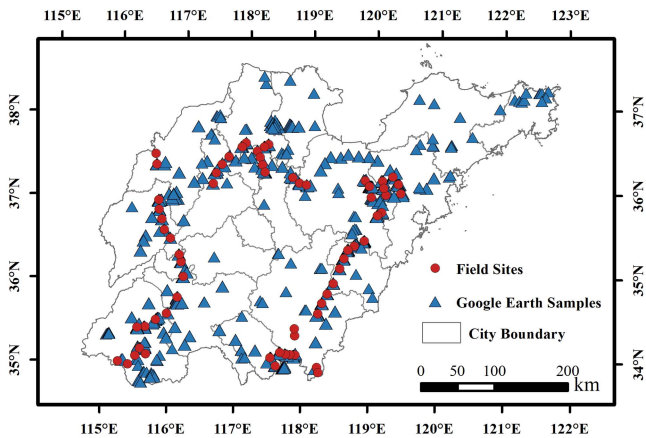


Fig. 5. Distribution of field samples over Shandong Province, China.

survey over Shandong Province was conducted in cooperation with other researchers in 2018. An MG858 handheld GPS was employed for the ground survey. Furthermore, Google Earth samples from 2018 were obtained by visual interpretation of very high-resolution images from Google Earth to replenish the field samples. To maintain the reliability of Google Earth samples, prior knowledge gathered from the *in situ* investigation was employed for visual interpretation. In total, 357 winter wheat and 353 nonwheat samples were collected.

#### F. Statistical Data

National agricultural statistical data were collected for comparison with satellite-derived crop area estimates to assess the classification results. Statistical winter wheat planting acreage (2018) at the municipality and county levels was acquired from the NBS [4]. The farming areas in the NBS of China were inferred based on the weights of the sampling croplands, which were reported by agrotechnicians who collected winter wheat growth conditions from survey samples by investigating registered farmlands or gathering estimates made by farmers. This indicated that the area's statistical data were dependable and

accurate [5]. Finally, statistical data for 16 municipalities and 137 counties in the study area were collected.

### III. METHODOLOGY

Traditional winter wheat mapping methods require manual annotation, model training, and wheat prediction when geolocations, sensor characteristics, or imaging conditions vary, which is time-consuming and inefficient [40]. Thus, this study presents a novel framework for automated and effective large-scale winter wheat mapping lacking the abovementioned repetitive operations. The framework is organized into three sequentially integrated parts (Fig. 6): 1) data preprocessing, 2) model training, and 3) validation. Consequently, the extent of winter wheat was mapped using the 10-m Sentinel-2 data and 30-m winter wheat mapping products.

#### A. Data Preprocessing

The preprocessing consisted of satellite data processing and label data processing, all of which were completed in the GEE platform using JavaScript. For label data, the processing procedures mainly involved resampling and reprojection. The 30-m winter wheat mapping production was resampled to 10-m resolution and reprojected to World Geodetic System 1984 (WGS84) for consistency with Sentinel-2 data. The preprocessing of Sentinel-2 satellite data included clouds and shadow masks, vegetation index calculation, temporal aggregation, and linear interpolation. The details of this process are described below.

*Cloud and Shadows Masks:* Cloudy observations of Sentinel-2 TOA data were eliminated based on the QA60 quality assessment band. The cloudy and shaded pixels were then masked and removed.

*NDVI and EVI Time Series Creation:* The two most regularly used spectral indicators were selected: 1) the NDVI [41] and 2) the EVI [42]. The NDVI and EVI time series have been frequently utilized to mine temporal features or phenological metrics of various crops [14], [43]. Moreover, these two indicators were superimposed on the Sentinel-2 data as two new bands.

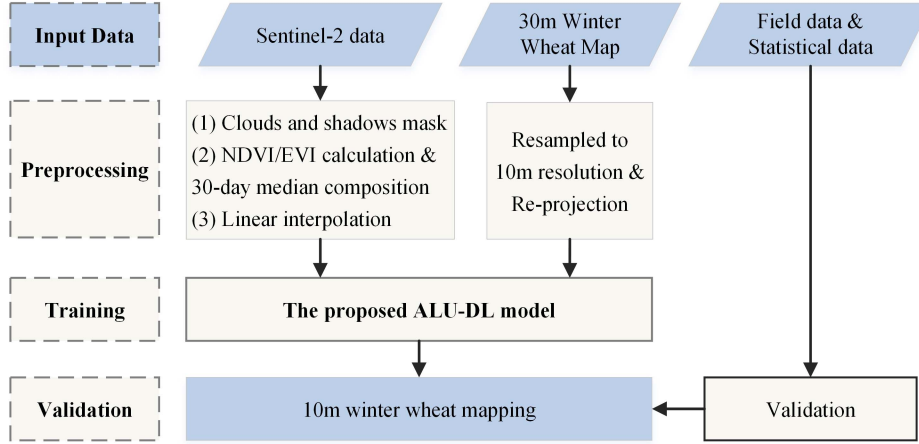


Fig. 6. Proposed framework for high-resolution winter wheat mapping based on the 10-m Sentinel-2 data and 30-m winter wheat mapping product.

TABLE II  
SUMMARY OF THE FEATURE CANDIDATES FOR WINTER WHEAT MAP GENERATION

Indices	Formulation	periods	Features
Eight bands of Sentinel-2	-	10	80
NDVI	$NDVI = \frac{NIR-RED}{NIR+RED}$	10	10
EVI	$EVI = 2.5 \times \frac{NIR-RED}{NIR+6 \times RED - 7.5 \times BLUE + 1}$	10	10
Total			100

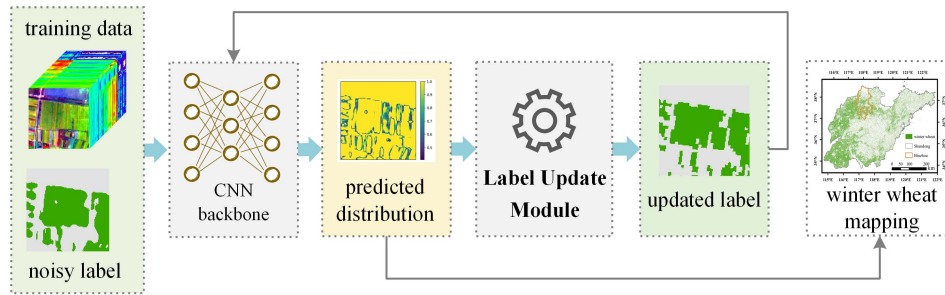


Fig. 7. Overview of the proposed ALU-DL model.

*Temporal Aggregation:* A time aggregation with regular time intervals can overcome the spatial heterogeneity of the observed data, producing consistent time series [36]. For the study area, cloud-free or near cloud-free images could be created according to the phenological stage, resulting in a total of ten time periods over ten months (Fig. 2). Finally, a 30-d composite was obtained by deriving the median value of all valid observations during each interval.

*Gap Filling With Linear Interpolation:* Owing to the effects of clouds, snow, or other conditions, it was sometimes impossible to gather continuous good-quality observations in some regions. To overcome this problem, the gaps caused by these factors were linearly interpolated over each band using the nearest valid values before and after the time step [14], [44].

Finally, monthly time series data were composited into a single image by concatenating the multitemporal spectral bands.

Here, eight bands were selected, and two vegetation indices were calculated every month, culminating in the emergence of 100 candidate features during the ten months (Table II). In this study, spatiotemporal and spectral information was leveraged simultaneously.

### B. Proposed ALU-DL Method

In this section, a novel automated winter wheat mapping framework, ALU-DL, is presented (Fig. 7). First, we describe the motivation for our method. Then, we introduce our ALU-DL method for accurate high-resolution winter wheat mapping, including the backbone segmentation model, label update module, and loss functions.

1) *Motivation:* Because the most traditional winter wheat classification method relies largely on accurate and trustworthy

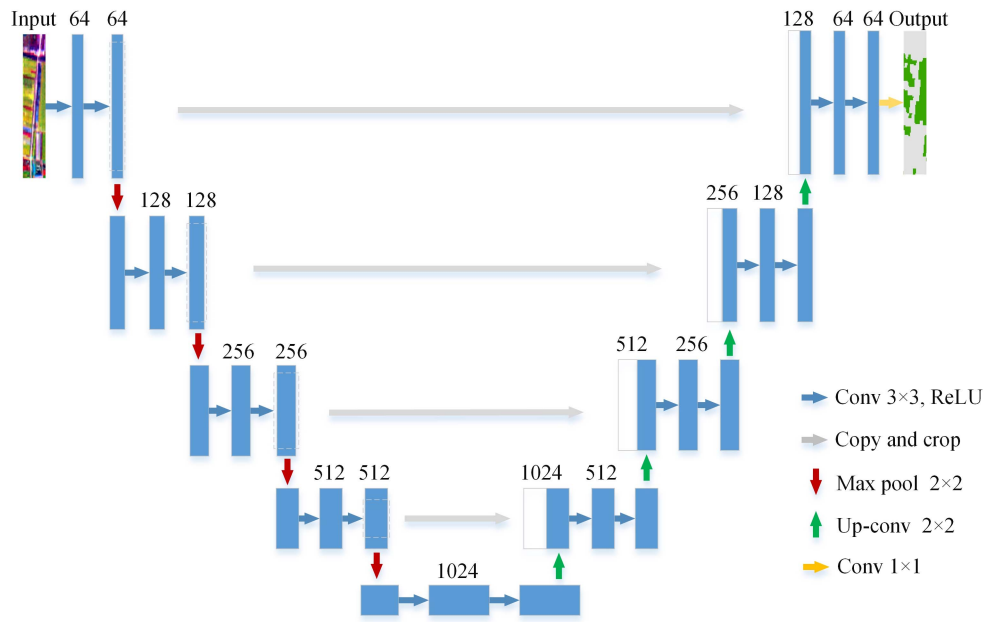


Fig. 8. Schematic overview of U-Net architecture. Figure modified from Ronneberger et al. [32].

training samples, which are costly and time-consuming, mapping of accurate large-scale winter wheat distribution results based on the existing low-resolution mapping product was a promising method. To alleviate the influence of noisy labels on accurate winter wheat mapping, we proposed the ALU-DL approach, which contains a label update module and can iteratively correct noisy labels based on the special phenological (seasonal) characteristics of winter wheat. In the proposed ALU-DL method, we exploited U-Net architecture with a strong generalization ability to extract crop types from satellite data.

2) *Segmentation Model*: Ronneberger et al. [32] proposed a U-Net model with encoder–decoder architecture for biomedical image segmentation (Fig. 8). Intermediate feature fusion was presented by U-Net, promoting the reuse of features in image segmentation tasks by concatenating multilevel feature maps with the same dimensionality via shortcut connections. Heller et al. [45] proved that deep learning models perform robustly on data with random perturbations. In particular, the superior performance of U-Net considering massive choppy perturbations is noteworthy. Thus, we selected U-Net as the convolutional neural network (CNN) backbone as it has been widely used for segmentation tasks, yielding favorable segmentation performance compared to state-of-the-art approaches in land cover/use mapping [12], [46], [47], [48].

In this article, the U-Net model analyzed multiband satellite observations throughout the winter wheat growing season and estimated the class probability distribution, which was used to input the label update module. Notably, the 30-m winter wheat mapping product was resampled to a 10-m resolution and considered the noisy label for model training.

3) *Label Update Module*: As shown in Fig. 7, the predicted distributions obtained by the CNN backbone were used as inputs for the label update module, which determined the accuracy

of pixelwise labeling using phenological information and then updated labels according to the accuracy. Finally, the updated labels were used as the label input for the next epoch. Note that noisy labels were the input of the label update model at the beginning of the training.

The proposed label update module was adapted from the framework designed by Dong et al. [26], which was applied to produce high-resolution land cover maps from a lower-resolution product in China. The primary steps of the noise correction and label update modules are illustrated in Fig. 9. First, the predicted probability distribution obtained by the U-Net backbone was used as the input for the label update module. The pixels predicted to be winter wheat with a probability greater than 0.9 were perceived as reliable predictions for winter wheat, and their observation band (OB) and/or vegetation index (VI) time series curves were averaged as the standard seasonal change curve of winter wheat. We then measured the distance/similarity between the standard seasonal change curve and the curve for each pixel in the original training samples. A greater distance value indicated that a pixel was more likely to be annotated as “nonwheat.” In this article, the update ratio threshold was set at 0.7, and the details are discussed in Section V-A. Finally, we sorted the distances in ascending order according to numerical magnitude. The pixels whose value of distance (similarity) was greater (less) than the threshold were considered nonwheat pixels, and those labeled “wheat” were regarded as noise labels following label update processing. Similarly, pixels with a distance (similarity) less (greater) than the threshold were considered wheat, and those labeled “nonwheat” were also updated. Therefore, the updated labels were acquired and utilized for training for the next epoch.

In this study, class separability between the winter wheat and the nonwheat type in the time series OB and/or VI data

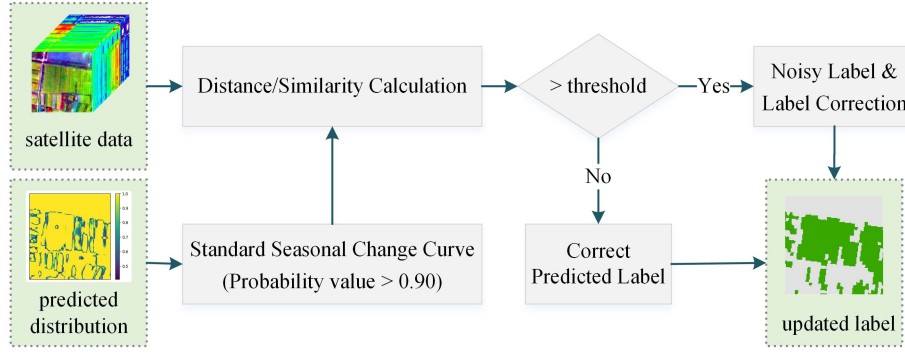


Fig. 9. Process of the label update module in the ALU-DL method.

was investigated using three commonly used distance/similarity statistics: 1) Euclidean distance, 2) cosine similarity, and 3) dynamic time warping, which are reportedly effective measures for this task [49], [50], [51], [52], [53], [54]. A greater distance or lower similarity between two classes indicated more different characteristics, encouraging successful class discrimination.

We denote the spectrum vector of each pixel as  $S_k (k = 1, 2, \dots, H \times W)$ ;  $H$  is the height of the image, and  $W$  is the width. Let  $S_a = \langle a_1, \dots, a_n \rangle$  and  $S_b = \langle b_1, \dots, b_n \rangle$  be two spectrum sequences, where  $n$  is the length of the spectrum vector and  $\delta$  is the distance between two sequence members, i.e.,  $\delta(a_i, b_j)$  represents the distance of the element  $a_i \in S_a \forall i = 1, \dots, n$  and  $b_j \in S_b \forall j = 1, \dots, n$ .

For each image, the standard seasonal change curve can be represented as

$$\bar{S} = \frac{1}{N_w} \sum_{i=1}^{N_w} S_i \quad (1)$$

where  $N_w$  denotes the number of pixels annotated as reliable winter wheat prediction results.

Because each satellite image had a different scenario complexity, we strived to update the labels of each image individually rather than using statistical information from all samples. Thus, the training efficiency was improved by skipping an extra inference step on the entire training dataset. It is worth noting that every training epoch updated all labels across the entire dataset. The ALU-DL models based on three distance/similarity-based update strategies were termed ED-UNet, CS-UNet, and DTW-UNet.

4) *Loss Function*: An illustration of ALU-DL is shown in Fig. 7. To determine the labels to be updated, an adaptive label update module was introduced (Fig. 9). In each mini-batch phase, forward computation and backward propagation were used to simultaneously update the network parameters and training labels.

Referring to the work of Dong et al. [26], we assume that high-resolution satellite data and noisy labels are available for training the models. The high-resolution training dataset can be denoted as  $X = \{x_i | x_i \in R^{H \times W \times C}, i = 1, 2, \dots, N\}$ , where  $N$  is the number of training images and  $x_i$  represents each image with  $H$  (height),  $W$  (width), and  $C$  (channels).

$Y = \{y_i | y_i \in [0, 1]^{H \times W}, i = 1, 2, \dots, N\}$  denotes the associated training label set, where  $[0, 1]$  indicates whether  $x_i$  denotes winter wheat. In general, the optimization problem on reliable labels implies minimizing a standard loss function  $L$  concerning the parameters  $\theta$  of the network, i.e.,  $\min_{\theta} L(\theta | X, Y)$ . Nevertheless, the models trained using this standard loss function were susceptible to being misled by the wrong labels. Therefore, label update information was considered to improve model performance.

As previously described, the network parameters and noisy labels were optimized simultaneously in this study, i.e.,  $\min_{\theta, Y} L(\theta, Y | X)$ . The predicted distribution is defined as  $\hat{Y}^d = f(X; \theta)$ , where  $f$  denotes the prediction output using the Softmax function. The updated label was set as  $\hat{Y} = \{\hat{y}_i | \hat{y}_i \in [0, 1]^{H \times W}, i = 1, 2, \dots, N\}$ . In training epoch  $t$ ,  $\hat{Y}^t$  was obtained by correcting the label  $\hat{Y}^{(t-1)}$  from the previous epoch,  $t - 1$ . At the start of the training,  $\hat{Y}$  was obtained by updating noisy labels  $Y$ .

To acquire an accurate winter wheat map based on the noisy labels, we applied a joint loss function  $L(\theta, Y | X)$ , which is composed of two parts and can be expressed as

$$L(\theta, Y | X) = L_{\text{update}}(\theta, \hat{Y}) + \alpha L_{\text{initial}}(\theta, Y) \quad (2)$$

where  $L_{\text{update}}(\theta, \hat{Y})$  and  $L_{\text{initial}}(\theta, Y)$  represent the cross-entropy loss with the updated label and the original noisy label, respectively, and  $\alpha$  is a hyperparameter that balances the two loss terms during training.

$L_{\text{update}}(\theta, \hat{Y})$  is the major loss used to govern the update of the network parameters  $\theta$ . The loss function is the cross-entropy loss between the predicted class probability distribution and the updated labels from the last epoch. The updated labels used in each epoch were obtained from the previous epoch using the label update module.  $L_{\text{update}}(\theta, \hat{Y})$  is defined as

$$L_{\text{update}}(\theta, \hat{Y}) = - \sum_{m=1}^{H \times W} [\hat{y}_m \log f_m(x_m; \theta) + (1 - \hat{y}_m) \log(1 - f_m(x_m; \theta))]. \quad (3)$$

Meanwhile, to prevent increasingly coarse boundaries, the original noisy label  $Y$  is also used for training, and  $L_{\text{initial}}(\theta, Y)$



is formulated as

$$L_{\text{initial}}(\theta, Y) = - \sum_{m=1}^{H \times W} [y_m \log f_m(x_m; \theta) + (1 - y_m) \log(1 - f_m(x_m; \theta))]. \quad (4)$$

### C. Model Training

The network was set up to use multispectral, spatiotemporal, mosaic satellite data as inputs and predict the distribution of winter wheat. The mosaic satellite images were cropped into chips of  $128 \times 128$  pixels with 20% overlap and matched with the corresponding 10-m resolution resampled winter wheat map through their geographic coordinates. Thus, paired satellite images and labels were obtained as the original training dataset. A total of 3200 experimental patches were acquired, all of which were used as training samples. The test dataset consisted of reliable point-based field samples collected *in situ* or from visual interpretations.

Although the original labels contained a considerable portion of correct labels, the U-Net model was pretrained using the original noisy labels to boost the representational capacity of the backbone model. Therefore, we pretrained the backbone network using loss  $L_{\text{initial}}(\theta, Y)$  to obtain the initial network parameters to be used in the next phase of the label update. To avoid overfitting noisy labels, the baseline U-Net model was trained for 5 epochs at a fixed learning rate of 0.01.

To correct the noisy label and iteratively train the U-Net model using the updated label, we obtained the network parameters for wheat mapping by training the backbone network with joint loss  $L(\theta, Y|X)$ . In this phase, the inputs of the U-Net backbone were satellite data and noisy labels. The label update module then determined the accuracy of the labels on a pixel-by-pixel basis based on phenological information and reversed the labels deemed inaccurate. Hence, we acquired a relatively clean label set by performing the label update process. For the label update process, the network parameters acquired by the pretraining phase were utilized to initiate the model, and the fixed learning rate was set to 0.01. The network was trained for 10 epochs when there were no noticeable changes in test precision.

The Adam optimizer was used to train all the networks from scratch using the PyTorch framework with a batch size of eight. The algorithms were implemented on an NVIDIA GeForce RTX 2080Ti GPU.

### D. Accuracy Assessment

The performance of the proposed method was assessed using 730 field locations obtained through field surveys and Google Earth visual interpretation. To objectively evaluate the performance of our method, we used the following evaluation indicators to measure the mapping results: the OA, the producer's accuracy (PA), the user's accuracy (UA), the  $F_1$  score, and the Kappa coefficient (Kappa). The OA value could assess the overall model performance with which all samples were classified; this approach has been used extensively in previous research [55]. The  $F_1$  score is the harmonic average of PA and

UA; while Kappa is a ratio that represents the error reduction of the classification model's misclassification versus the misclassification of a random classification procedure.

### E. Comparison With Other Methods

To prove the efficiency of the proposed approach, we compared the performance of ALU-DL with that of other frequently used classifiers by conducting winter wheat mapping in the study area. In this study, two machine learning methods and three deep learning models were selected: SVM, RF, the baseline U-Net model, the method proposed by Dong et al. [26], and the PAL model.

The SVM and RF are supervised learning methods proven useful for land use and cover classification. SVM uses kernel functions to map samples from the original space to the high-dimensional feature space and then searches the mapped space for the partition hyperplane with "the greatest margin" [56]. RF is an ensemble learning strategy that uses bagging to combine many DT classifiers to minimize the model prediction variance. To successfully restrict the risk of overfitting, each tree is formulated using random sample selection and feature number selection [57]. SVM and RF are widely used as baseline models for remote sensing tasks because they can handle high-dimensional input variables.

Moreover, the baseline U-Net model without the label update module was applied to map winter wheat, and the results were compared with those of the proposed ALU-DL model to verify the effectiveness of the label update module. Since the method proposed in this study was motivated by Dong et al. [26], we compared the mapping performance of our model with the initial noise correction (INC) model proposed by Dong et al. [26] to evaluate whether phenological characteristics would improve the performance of the mapping method. The PAL model, which is a significant segmentation network for noisy labeled datasets, is also regarded as one of the baseline methods. All the baseline methods were implemented with their default settings.

## IV. RESULTS

### A. Winter Wheat Mapping Accuracy of ALU-DL

As there is no official reference crop data layer, it is impossible to validate the accuracy of large-scale high-resolution winter wheat mapping pixel by pixel. Thus, pixel-level validation samples were used to quantify the classification performance of different models in terms of the aforementioned indices. The winter wheat in Shandong Province was classified using the proposed ALU-DL method with three update strategies, and the assessment metrics are presented in Table III.

As shown in Table III, the CS-UNet model achieved outstanding classification performance, with an OA of 92.19% and a Kappa of 0.827. Furthermore, the CS-UNet model's PA, UA, and  $F_1$  scores all exceeded 90%. The results show that the CS-UNet model was adequate and robust for winter wheat mapping in this area, even when no field samples were used for training. The accuracies for the other two label updated models were considerably lower, while OA decreased by 2.05–8.08% (Kappa

TABLE III  
ACCURACIES OF WINTER WHEAT USING ALU-DL IN SHANDONG PROVINCE

Models	Class	Ground truth samples		PA(%)	UA(%)	OA(%)	F <sub>1</sub>	Kappa
		Wheat	Nonwheat					
ED-UNet	Wheat	324	39	90.76	89.26	90.14	0.877	0.777
	Nonwheat	33	334	90.76	89.26			
DTW-UNet	Wheat	296	55	82.91	84.33	84.11	0.825	0.671
	Nonwheat	61	318	85.25	83.91			
CS-UNet	Wheat	326	26	91.32	92.61	<b>92.19</b>	<b>0.914</b>	<b>0.827</b>
	Nonwheat	31	347	93.03	91.80			

The best results are shown in bold.

TABLE IV  
ACCURACIES OF WINTER WHEAT USING ALU-DL IN SHANDONG PROVINCE

Models	Class	Ground truth samples		PA(%)	UA(%)	OA(%)	F <sub>1</sub>	Kappa
		Wheat	Nonwheat					
SVM	Wheat	243	29	68.07	89.34	80.41	0.801	0.606
	Nonwheat	114	344	92.23	75.11			
RF	Wheat	274	30	76.75	90.13	84.52	0.844	0.689
	Nonwheat	83	343	91.96	80.521			
Baseline U-Net	Wheat	329	70	92.16	82.46	86.58	0.863	0.728
	Nonwheat	28	303	81.23	91.54			
INC model	Wheat	325	66	91.03	83.12	86.57	0.862	0.714
	Nonwheat	32	307	82.30	90.56			
PAL	Wheat	327	63	91.60	83.85	87.24	0.870	0.712
	Nonwheat	30	309	83.06	91.15			
ALU-DL	Wheat	326	26	91.32	92.61	<b>92.19</b>	<b>0.914</b>	<b>0.827</b>
	Nonwheat	31	347	93.03	91.80			

The best results are shown in bold.

reduced by 0.05–0.156) compared with CS-UNet, implying that some nonwinter wheat pixels were mistakenly categorized as winter wheat. The OA of DTW-UNet was slightly lower than that of the other models, implying that DTW-UNet might introduce additional noise information in the label update module.

As a result, the CS-UNet model, which achieved the best classification performance among the three update strategies, was selected as the ALU-DL model to compare with other models and perform the follow-up evaluation.

### B. Comparisons of the Performance With Other Mapping Methods

Table IV compares ALU-DL to the other five winter wheat mapping models previously described. The ALU-DL model produced the most competitive performance with an OA of 92.19% and a Kappa of 0.827, demonstrating the stable and robust capacity of ALU-DL to distinguish winter wheat from other land cover types.

Compared to the four deep learning models, the accuracy of the two machine learning methods, SVM and RF, was considerably lower, with 80.41% and 84.52% OA, respectively. Given that no reliable high-resolution training labels were used in any of the models, the agreement between the field samples and the mapping results of the deep learning models demonstrated that deep learning approaches have considerable potential to build maps by overcoming the influence of noisy labels.

In contrast, the mapping accuracy of the ALU-DL model increased dramatically compared to that of the baseline U-Net model, with the OA rising by 5.61%; this increase confirmed the advantages of the label update module with the CS update strategy in correcting wrong labels. The classification performance of the INC model was comparable to that of the baseline

U-Net model and inferior to that of ALU-DL. This indicates that phenological characteristics, rather than the class probability distribution, were better suited for determining which labels needed to be updated for this task. Compared with the ALU-DL, the performance of PAL is poor, with the OA decreasing by 4.95%. In general, these results indicate that the proposed method outperforms all the adopted baselines.

### C. Visual Assessment

As no complete high-resolution winter wheat validation exists for this task, visual inspection was used to qualitatively evaluate the results. Because of the better mapping performance over the validation samples, ALU-DL was employed for winter wheat mapping in Shandong Province, and all pixels were divided into two categories, 1) winter wheat and 2) nonwinter wheat. The resulting winter wheat maps are shown in Fig. 10; Fig. 10(a) shows the produced 10-m binary winter wheat/nonwinter wheat maps of Shandong Province. The distribution of winter wheat in this area was consistent with the geographical distribution; major winter wheat zones were frequently located in plains with ample water supply. More specifically, winter wheat planting areas in Shandong Province were primarily concentrated in the western plains, southern region, and northeastern areas of the middle mountainous regions, whereas planting areas were scattered and less abundant in central hilly regions and eastern coastal areas. This result is generally consistent with the distribution of winter wheat mapping for Shandong Province [34].

Fig. 10(b) and (c) show the 10-m winter wheat map of Binzhou City derived by the ALU-DL model and the baseline U-Net model, respectively. Two local zones (A and B) are shown in Fig. 10(A-I, A-II, A-III B-I, B-II, B-III) to highlight more details, displaying the mapping results and the original Sentinel-2

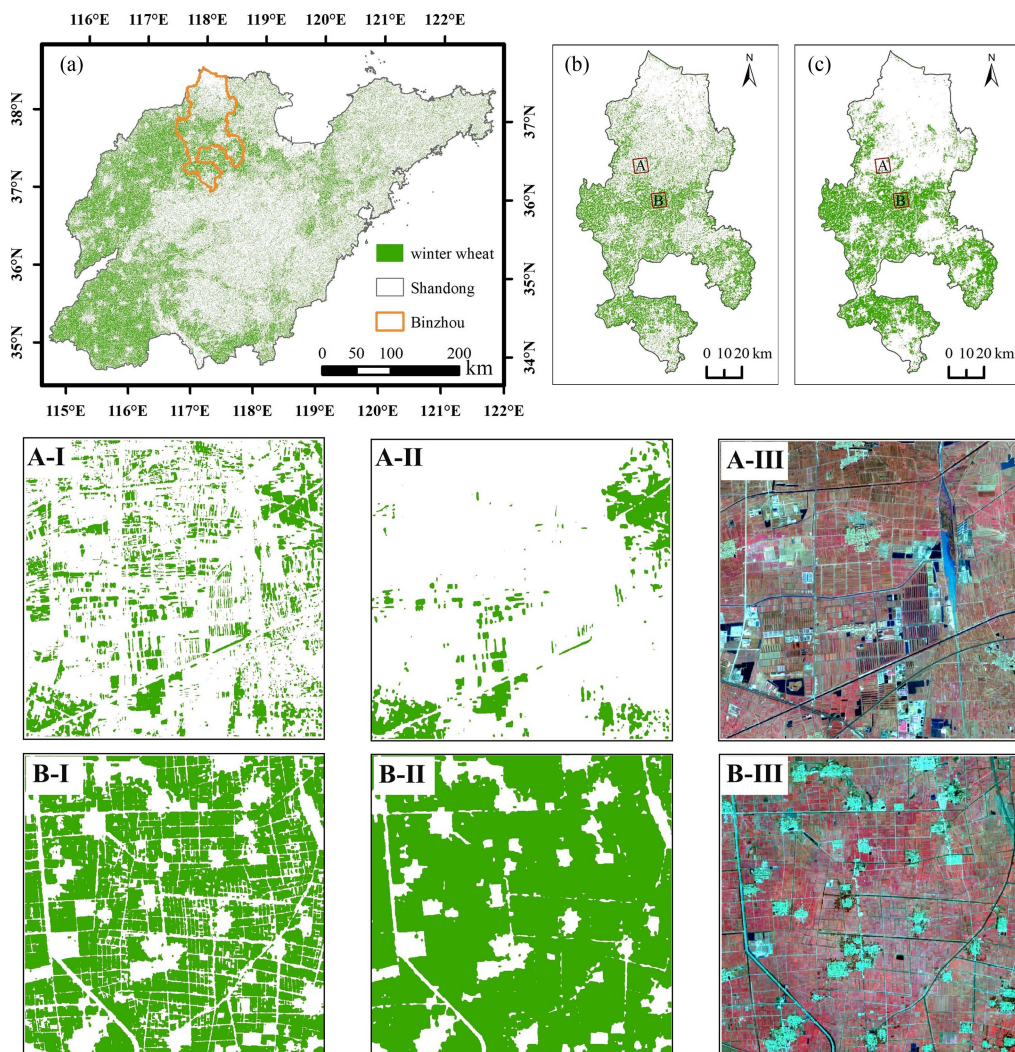


Fig. 10. Winter wheat maps with 10-m spatial resolution derived with ALU-DL in the (a) Shandong Province and (b) Binzhou City. (c) Winter wheat map with 10-m spatial resolution generated by the baseline U-Net model. A-I and B-I represent the local 10-m winter wheat mapping results of the ALU-DL method, and A-II and B-II are the 10-m winter wheat mapping results of the baseline U-Net model. A-III and B-III are the Sentinel-2 optical image calculated in GEE (the median value through the grown stage). The Sentinel-2 images are false-color composited with near-infrared (B8), red (B4), and green (B3) bands displayed in RGB channels.

images. Sentinel-2 images were false-color composites with near-infrared (B8), red (B4), and green (B3) bands displayed in RGB channels. Generally, winter wheat distributions derived with the ALU-DL model [Fig. 10 (A-I, B-I)] were consistent with optical images [Fig. 10 (A-III B-III)] in the two local zones, indicating that the ALU-DL model could effectively distinguish winter wheat from other land cover types. In particular, because of the higher resolution input images and label updating module, the suggested technique not only accurately maps the fragmented winter wheat parcel (Zone A) but also detects the scattered field paths (Zone B).

#### D. Areal Comparison

Ideally, large-scale field or pixel-level validation information should be available; however, this is unlikely in fact due to the lack of relative productions. The NBS does, however, publish

municipal- and county-level crop acreage statistics, which were used as a comparison to evaluate the performance of our method. Using the previously computed ALU-DL model, we estimated the cultivated areas for winter wheat at the municipal and county levels. Fig. 11 shows the scatterplots and linear fits for the mapping results at the municipal and county levels. Several municipalities lacked available statistical areas for winter wheat, such as Laiwu City and Rizhao City. Thus a total of 14 municipalities and 123 counties were used for evaluation and are presented in Fig. 11.

A correlation coefficient ( $R^2$ ) of 1.0 indicates a perfect agreement between the area of the mapping product and the agricultural statistical areas. At the municipal level [Fig. 11(a)], the  $R^2$  between the classified and statistical areas was 0.95, indicating a significant correlation. However, the approach performed slightly less efficiently at the county level, with an  $R^2$  value of 0.78 [Fig. 11(b)]. Overall, there was a significant agreement

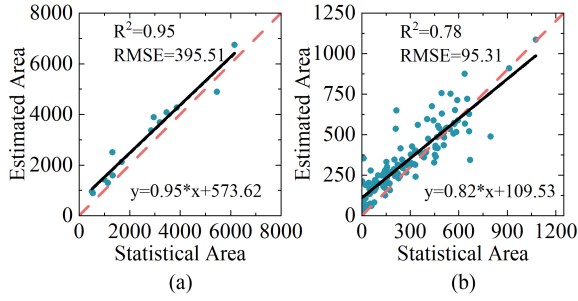


Fig. 11. Comparison between the estimated planting area ( $\text{km}^2$ ) of winter wheat and agricultural statistical area at (a) the municipal and (b) county level for 2018. The red dotted line denotes the 1:1 line. The unit of root mean square error (RMSE) is  $\text{km}^2$ .

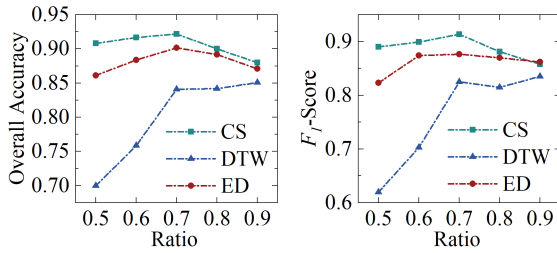


Fig. 12. Classification performance of the ALU-DL method with different update ratios for the noisy label in the label update module when both optical bands and vegetation indices are used.

between our results and statistical data. These findings provide compelling evidence of the reliability of the proposed model.

The slope and intercept of the linear model for each sample were also relevant for understanding the utility of the output [58]. The average linear regression slopes for winter wheat at the municipal and county level were 0.95 and 0.78, respectively, and the average intercepts were 395.51 and 95.31  $\text{km}^2$ , respectively. Ideally, the linear regression slope and intercept term should be close to 1 and 0, respectively. Thus, a slope  $< 1$  and positive intercepts imply that the ALU-DL model analysis under-classifies large-area regions and over-classifies small-area regions. Based on the above analysis, the expected crop area of winter wheat was slightly overestimated by the ALU-DL model, implying that although our method could mitigate the effects of noisy labels, acquiring a completely clean label set is still challenging.

## V. DISCUSSION

### A. Performance Comparison of Different Threshold Values for Label Update Module

Determining the threshold value for the label update module is associated with assumptions that should be further discussed. In this section, we discuss the five tested threshold values in 0.1-step intervals from 0.5 to 0.9; OA and  $F_1$  scores were used as classification accuracy measures with alternative thresholding. Validation samples were used to determine the optimal ratio parameters.

As shown in Fig. 12, satisfactory performance was achieved by the CS-UNet model when the threshold value was 0.7, and the OA value and  $F_1$  score were maintained at approximately

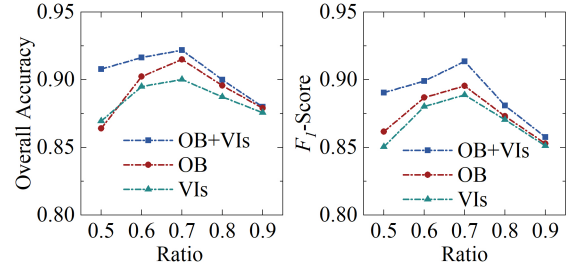


Fig. 13. Classification performance of the ALU-DL with different combinations of bands used.

92% and 0.91, respectively. Moreover, a threshold that was too low or too high for both CS-UNet and ED-UNet models led to slightly poorer mapping accuracy, whereas the highest mapping accuracy was achieved when the threshold value was set to 0.7. Regarding the DTW-UNet model, when the thresholds were less than 0.7, the mapping accuracy decreased significantly. Satisfactory performances of the DTW-UNet model were achieved with a threshold value ranging from 0.7 to 0.9; within this range, the mapping accuracy remained almost steady.

The above results demonstrate that the accuracy of the raw labels was close to 70%; in other words, the percentage of noisy labels in the original labels approached 30%. However, the field survey results of Dong et al. [5] revealed an OA of 89.88%. The decrease in the accuracy of the raw labels was mainly caused by the large number of labeling errors introduced by the data resampling process. In a word, the noise was not only sourced from the misclassification of the initial 30-m resolution data product but also from the missing details in the low-resolution labels.

### B. Performance Comparison of Input Bands

An experiment was conducted to quantify the effect of multispectral information in the input time series of the update module on the classification performance of the ALU-DL model. All bands and individual OB or VIs were compared to determine how limited spectral lengths affected winter wheat mapping in the study area. In this section, the OA and  $F_1$  scores were used as performance indicators.

As shown in Fig. 13, the complete time series (OB+VIs) input yielded the highest OA of 92.19% for the CS-UNet model compared to individual OB or individual VIs input. The  $F_1$  scores demonstrated a similar superiority. The results showed that incorporating multitemporal information could enhance the classification performance of the ALU-DL model. A possible explanation could be that detailed spatiotemporal information enables better accuracy when distinguishing between winter wheat and nonwinter wheat, which has been proven by other studies [59].

### C. Visual Comparison of Models by T-Distributed Stochastic Neighbor Embedding

After dimensionality reduction, high-dimensional features could be presented in a plane for feature separability analysis [55]. Using multilayer processing modules, input characteristics of crop categorization were translated into high-level

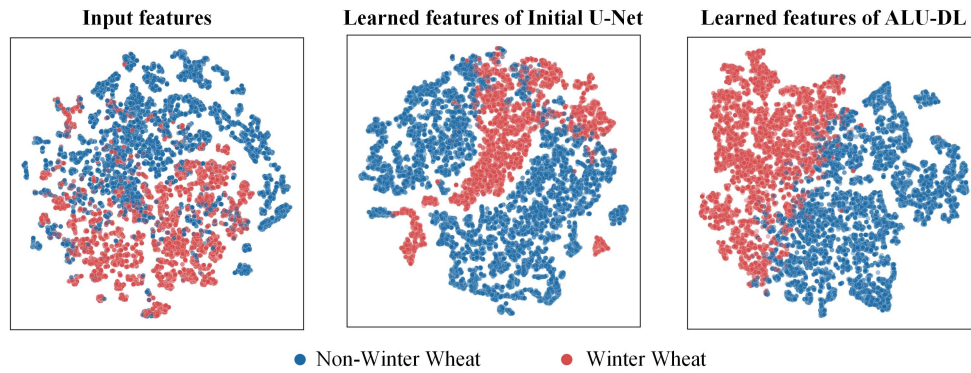


Fig. 14. Feature separability comparison based on t-SNE visualization of input features and learned features. Six thousand samples of each crop type are randomly selected for illustration; each point represents a sample.

feature representations, also referred to as hidden features in deep neural networks. Hidden features contained fine-grained temporal information and exhibited complicated patterns, challenging intuitive interpretation. Therefore, we used t-distributed stochastic neighbor embedding (t-SNE), a powerful nonlinear dimension reduction approach, to project high-dimensional input and learned features into a 2-D space [60]. This approach facilitated the visual comparison of two feature types. The ability of feature representations to distinguish winter wheat could be assessed by comparing the separability of features between winter and nonwinter wheat samples. Feature separability analysis can be considered a general complement to model-specific interpretation methods to understand feature transformations within deep neural networks. Although a visual comparison could not provide a quantitative explanation, it could aid in the intuitive appraisal and investigation of the consequences of sophisticated feature learning.

In this study, we used the output of the hidden feature from the last hidden layer as the learned spatiotemporal feature of two deep learning models. After being projected into a 2-D space using t-SNE, a separability comparison of high-dimension features for input features and the features derived by two models was conducted, as shown in Fig. 14. Notably, using input features, a substantial number of winter wheat samples (red circles) were indistinguishable from nonwinter wheat samples (blue circles). However, learned features of the winter and nonwinter wheat samples showed better separation. Expectedly, the best separability was exhibited by the features retrieved by the ALU-DL model than another two types of features, showing that the ALU-DL model retrieved more valuable information for winter wheat mapping than the raw input features and the baseline U-Net model.

## VI. CONCLUSION

This study aimed to develop a new classification method to automatically recognize winter wheat across large areas for operational agricultural monitoring. We selected Shandong Province as the study area using images acquired by the Sentinel-2 satellite. Therefore, an automated winter wheat mapping method, ALU-DL, based on Sentinel-2 data for large-scale applications

was developed. More importantly, no field-level samples were adopted in the framework used to train the model. The method could withstand the harmful effects of noise caused by the lower resolution (30 m) label, benefiting from a label update module based on the special phenological (seasonal) characteristics of winter wheat. The efficiency and validity of our proposed framework were validated using field survey samples. The winter wheat mapping results derived using the ALU-DL model achieved an OA of  $> 92\%$ , demonstrating the effectiveness and accuracy of the proposed method. Our proposed method significantly outperformed other widely used methods in terms of a variety of evaluation indicators, yielding sufficient classification performance. Furthermore, certain impact factors regarding the applicability of models for Sentinel-2-based winter wheat monitoring were discussed. Using the proposed method with the existing 30-m resolution winter wheat maps, we produced refined 10-m resolution winter wheat maps across Shandong province without *in situ* samples for training. The accuracy of the mapping product was evaluated by crop acreage statistics at the municipal and county levels. The  $R^2$  values between estimated areas of winter wheat and the agricultural statistical area were 0.95 and 0.78 at the municipal level and the county level, respectively. The results of this study indicate that the proposed methodology could offer a viable and promising method for high-resolution, operational agricultural monitoring over large areas.

For future research, we aim to derive a method to automatically identify the ratio of correct labels in the label update module and characterize winter wheat at the field level, rather than the pixel level. Field-level classification may lead to better classification performance. Notably, the label update technique enabled the application of current knowledge and products to remote sensing scenarios, such as the extraction of water body borders using high-resolution images.

## ACKNOWLEDGMENT

The authors would like to thank the European Space Agency for providing the Sentinel-2 data. The authors would also like to thank the National Bureau of Statistics of China for providing the statistical data and the team led by Prof. Wenping Yuan at Sun

Yat-sen University, for providing 30-m winter wheat mapping products and field validation data. The authors would like to thank Editage (www.editage.cn) for English language editing.

## REFERENCES

- [1] "The 17 sustainable development goals," United Nations, New York, NY, USA, 2015. [Online]. Available: <https://sdgs.un.org/goals>
- [2] L. Liu et al., "Uncertainty in wheat phenology simulation induced by cultivar parameterization under climate warming," *Eur. J. Agronomy*, vol. 94, pp. 46–53, 2018.
- [3] FAOSTAT, "FAO statistical databases," Food Agriculture Org., Rome, Italy, 2018. [Online]. Available: <http://www.fao.org/faostat/en/>
- [4] "National statistical yearbook," National Bureau of Statistics of China, Beijing, China, 2018. [Online]. Available: <http://www.stats.gov.cn/tjsj/>
- [5] J. Dong et al., "Early-season mapping of winter wheat in China based on landsat and Sentinel images," *Earth Syst. Sci. Data*, vol. 12, no. 4, pp. 3081–3095, 2020.
- [6] R. Massey et al., "Modis phenology-derived, multi-year distribution of continuous us crop types," *Remote Sens. Environ.*, vol. 198, pp. 490–503, 2017.
- [7] B. D. Wardlow and S. L. Egbert, "Large-area crop mapping using time-series modis 250 m NDVI data: An assessment for the us central great plains," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 1096–1116, 2008.
- [8] B. D. Wardlow, S. L. Egbert, and J. H. Kastens, "Analysis of time-series modis 250 m vegetation index data for crop classification in the us central great plains," *Remote Sens. Environ.*, vol. 108, no. 3, pp. 290–310, 2007.
- [9] A. Khan et al., "Evaluating landsat and rapideye data for winter wheat mapping and area estimation in Punjab, Pakistan," *Remote Sens.*, vol. 10, no. 4, 2018, Art. no. 489.
- [10] T. T. Nguyen et al., "Monitoring agriculture areas with satellite images and deep learning," *Appl. Soft Comput.*, vol. 95, 2020, Art. no. 106565.
- [11] S. Wang, G. Azzari, and D. B. Lobell, "Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques," *Remote Sens. Environ.*, vol. 222, pp. 303–317, 2019.
- [12] M. Li and A. Stein, "Mapping land use from high resolution satellite images by exploiting the spatial arrangement of land cover objects," *Remote Sens.*, vol. 12, no. 24, 2020, Art. no. 4158.
- [13] A. Nasrallah et al., "A novel approach for mapping wheat areas using high resolution Sentinel-2 images," *Sensors*, vol. 18, no. 7, 2018, Art. no. 2089.
- [14] N. You et al., "The 10-m crop type maps in northeast China during 2017–2019," *Sci. Data*, vol. 8, no. 1, 2021, Art. no. 41.
- [15] A. Asgarian, A. Soffianian, and S. Pourmanafi, "Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: A case of central Iran using multi-temporal landsat 8 imagery," *Comput. Electron. Agriculture*, vol. 127, pp. 531–540, 2016.
- [16] J. Liu, Q. Feng, J. Gong, J. Zhou, J. Liang, and Y. Li, "Winter wheat mapping using a random forest classifier combined with multi-temporal and multi-sensor data," *Int. J. Digit. Earth*, vol. 11, no. 8, pp. 783–802, 2018.
- [17] F. Waldner et al., "Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity," *Int. J. Remote Sens.*, vol. 37, no. 14, pp. 3196–3231, 2016.
- [18] J. Adrian, V. Sagan, and M. Maimaitijiang, "Sentinel SAR-optical fusion for crop type mapping using deep learning and Google Earth engine," *ISPRS J. Photogrammetry Remote Sens.*, vol. 175, pp. 215–235, 2021.
- [19] L. Zhong, L. Hu, H. Zhou, and X. Tao, "Deep learning based winter wheat mapping using statistical data as ground references in kansas and northern texas, us," *Remote Sens. Environ.*, vol. 233, 2019, Art. no. 111411.
- [20] F. P. W. Lo, Y. Sun, J. Qiu, and B. Lo, "Image-based food classification and volume estimation for dietary assessment: A review," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 7, pp. 1926–1939, Jul. 2020.
- [21] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021.
- [22] D. Hong et al., "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340–4354, May 2021.
- [23] D. Hong et al., "Spectralformer: Rethinking hyperspectral image classification with transformers," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–15, 2022.
- [24] H. Zhang, J. Kang, X. Xu, and L. 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, 2020, Art. no. 105618.
- [25] H. Zhang, Y. Song, C. Han, and L. Zhang, "Remote sensing image spatiotemporal fusion using a generative adversarial network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4273–4286, May 2021.
- [26] R. Dong, W. Fang, H. Fu, L. Gan, J. Wang, and P. Gong, "High-resolution land cover mapping through learning with noise correction," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2021.
- [27] R. Dong et al., "Improving 3-m resolution land cover mapping through efficient learning from an imperfect 10-m resolution map," *Remote Sens.*, vol. 12, no. 9, 2020, Art. no. 1418.
- [28] K. Malkin et al., "Label super-resolution networks," in *Proc. Int. Conf. Learn. Representations*, 2018, pp. 1–22.
- [29] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.
- [30] N. Malkin, C. Robinson, and N. Jojic, "High-resolution land cover change from low-resolution labels: Simple baselines for the 2021 IEEE GRSS data fusion contest," Tech. Rep., 2021.
- [31] H. Zhu, J. Shi, and J. Wu, "Pick-and-learn: Automatic quality evaluation for noisy-labeled image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2019, pp. 576–584.
- [32] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2015, pp. 234–241.
- [33] F. Xu et al., "Mapping winter wheat with combinations of temporally aggregated Sentinel-2 and landsat-8 data in shandong province, China," *Remote Sens.*, vol. 12, no. 12, 2020, Art. no. 2065.
- [34] H. Zhang, H. Du, C. Zhang, and L. Zhang, "An automated early-season method to map winter wheat using time-series Sentinel-2 data: A case study of shandong, China," *Comput. Electron. Agriculture*, vol. 182, 2021, Art. no. 105962.
- [35] M. Drusch et al., "Sentinel-2: Esa's optical high-resolution mission for GMES operational services," *Remote Sens. Environ.*, vol. 120, pp. 25–36, 2012.
- [36] N. You and J. Dong, "Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth engine," *ISPRS J. Photogrammetry Remote Sens.*, vol. 161, pp. 109–123, 2020.
- [37] I. Eremyanova, O. Barron, and M. Alaiakhsh, "A comparative evaluation of arid inflow-dependent vegetation maps derived from landsat top-of-atmosphere and surface reflectances," *Int. J. Remote Sens.*, vol. 39, no. 20, pp. 6607–6630, 2018.
- [38] L. Yin, N. You, G. Zhang, J. Huang, and J. Dong, "Optimizing feature selection of individual crop types for improved crop mapping," *Remote Sens.*, vol. 12, no. 1, 2020, Art. no. 162.
- [39] X. Yang et al., "Monthly estimation of the surface water extent in France at a 10-m resolution using Sentinel-2 data," *Remote Sens. Environ.*, vol. 244, 2020, Art. no. 111803.
- [40] D. Zhang et al., "A generalized approach based on convolutional neural networks for large area cropland mapping at very high resolution," *Remote Sens. Environ.*, vol. 247, 2020, Art. no. 111912.
- [41] C. J. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sens. Environ.*, vol. 8, no. 2, pp. 127–150, 1979.
- [42] A. Huete, H. Liu, K. Batchily, and W. Van Leeuwen, "A comparison of vegetation indices over a global set of TM images for EOS-modis," *Remote Sens. Environ.*, vol. 59, no. 3, pp. 440–451, 1997.
- [43] L. Zhong, L. Hu, and H. Zhou, "Deep learning based multi-temporal crop classification," *Remote Sens. Environ.*, vol. 221, pp. 430–443, 2019.
- [44] S. Kandasamy, F. Baret, A. Verger, P. Neveux, and M. Weiss, "A comparison of methods for smoothing and gap filling time series of remote sensing observations—application to modis LAI products," *Biogeosciences*, vol. 10, no. 6, pp. 4055–4071, 2013.
- [45] N. Heller, J. Dean, and N. Papanikolopoulos, "Imperfect segmentation labels: How much do they matter?," in *Intravascular Imaging and Computer Assisted Stenting and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*, Berlin, Germany: Springer, 2018, pp. 112–120.
- [46] L. Garg, P. Shukla, S. K. Singh, V. Bajpai, and U. Yadav, "Land use land cover classification from satellite imagery using mUnet: A modified Unet architecture," in *Proc. Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2019, pp. 359–365.

- [47] A. Stoian, V. Poulain, J. Inglada, V. Poughon, and D. Derksen, "Land cover maps production with high resolution satellite image time series and convolutional neural networks: Adaptations and limits for operational systems," *Remote Sens.*, vol. 11, no. 17, 2019, Art. no. 1986.
- [48] P. Zhang, Y. Ke, Z. Zhang, M. Wang, P. Li, and S. Zhang, "Urban land use and land cover classification using novel deep learning models based on high spatial resolution satellite imagery," *Sensors*, vol. 18, no. 11, 2018, Art. no. 3717.
- [49] 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, 2019, Art. no. 1257.
- [50] M. R. da Silva, O. A. de Carvalho, R. F. Guimarães, R. A. Trancoso Gomes, and C. Rosa Silva, "Wheat planted area detection from the MODIS NDVI time series classification using the nearest neighbour method calculated by the euclidean distance and cosine similarity measures," *Geocarto Int.*, vol. 35, no. 13, pp. 1400–1414, 2020.
- [51] A. Fatikhunnada et al., "Assessment of pre-treatment and classification methods for Java paddy field cropping pattern detection on MODIS images," *Remote Sens. Appl. Soc. Environ.*, vol. 17, 2020, Art. no. 100281.
- [52] H. Kerner, C. Nakalembe, and I. Becker-Reshef, "Field-level crop type classification with k nearest neighbors: A baseline for a new Kenya smallholder dataset," in *Proc. Int. Conf. Learn. Representations Workshop Comput. Vis. Agriculture*, 2020, pp. 1–6.
- [53] W. S. Moola, W. Bijker, M. Belgiu, and M. Li, "Vegetable mapping using fuzzy classification of dynamic time warping distances from time series of Sentinel-1a images," *Int. J. Appl. Earth Observation Geoinformation*, vol. 102, 2021, Art. no. 102405.
- [54] L. Xu, H. Zhang, C. Wang, B. Zhang, and M. Liu, "Crop classification based on temporal information using Sentinel-1 SAR time-series data," *Remote Sens.*, vol. 11, no. 1, 2018, Art. no. 53.
- [55] J. Xu et al., "Towards interpreting multi-temporal deep learning models in crop mapping," *Remote Sens. Environ.*, vol. 264, 2021, Art. no. 112599.
- [56] M. Pal and P. M. Mather, "Support vector machines for classification in remote sensing," *Int. J. Remote Sens.*, vol. 26, no. 5, pp. 1007–1011, 2005.
- [57] M. Belgiu and L. Drăguț, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogrammetry Remote Sens.*, vol. 114, pp. 24–31, 2016.
- [58] D. M. Johnson, "Using the landsat archive to map crop cover history across the United States," *Remote Sens. Environ.*, vol. 232, 2019, Art. no. 111286.
- [59] J. Brinkhoff, J. Vardanega, and A. J. Robson, "Land cover classification of nine perennial crops using Sentinel-1 and-2 data," *Remote Sens.*, vol. 12, no. 1, 2019, Art. no. 96.
- [60] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, 2008.



**Xiaoman Qi** received the B.S. degree in engineering in 2018 from the China University of Geosciences, Beijing, China, where she is currently working toward the Ph.D. degree in engineering with the School of Land Science and Technology.

Her research interests include deep learning and remote sensing image processing.



**Yuebin Wang** (Member, IEEE) received the Ph.D. degree in engineering from the School of Geography, Beijing Normal University, Beijing, China, in 2016.

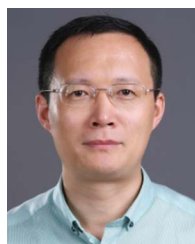
He was a Postdoctoral Researcher with the School of Mathematical Sciences, Beijing Normal University. He is currently an Associate Professor with the School of Land Science and Technology, China University of Geosciences (Beijing), Beijing, China. His research interests include remote sensing imagery processing and 3-D urban modeling.



**Junhuan Peng** was born in Chongqing, China, in 1964. He received the B.S. degree in surveying and mapping and the M.S. degree in mine surveying from Central South University, Changsha, China, in 1985 and 1988, respectively, and the Ph.D. degree in geodesy and surveying engineering from Wuhan University, Wuhan, China, in 2003.

From 1988 to 2003, he was with the Guilin University of Technology, Guilin, China, as a Lecturer and then as a Professor. From 2003 to 2005, he was a Postdoctoral Researcher in geodesy and surveying engineering with Shanghai Astronomical Observatory, China Academy of Sciences, Shanghai, China. From 2005 to 2007, he was with Chongqing University as a Professor. Since 2007, he has been working as a Full Professor with the China University of Geosciences, Beijing, China. He has authored/coauthored more than 40 papers in Chinese and English in both international and national journals.

Dr. Peng served as a Reviewer for several famous surveying and mapping and related academic journals.



**Liqiang Zhang** (Member, IEEE) received the Ph.D. degree in geoinformatics from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2004.

He is currently a Professor with the Faculty of Geographical Science, Beijing Normal University, Beijing, China. His research interests include remote sensing image processing, 3-D urban reconstruction, and spatial object recognition.

**Wenping Yuan** received the Ph.D. degree in ecology from the Institute of Botany, Chinese Academy of Sciences, Beijing, China, in 2007.

He is currently a Professor and Doctoral supervisor with the School of Atmospheric Science, Sun Yat-Sen University, Guangzhou, China. His research interests include terrestrial ecosystem models based on processes-based analyses at multiple sites and ecosystem types.



**Xiaotong Qi** received the Ph.D. degree in photogrammetry and remote sensing from the School of Geosciences and Surveying Engineering, China University of Mining and Technology, Beijing, China, in 2020.

She is currently a Lecturer with the School of Marine Technology and Geomatics, Jiangsu Ocean University, Lianyungang, China. Her research interests include hyperspectral remote-sensing image processing and application.