remote sensing

MDPI

Article

Parcel-Level Mapping of Horticultural Crop Orchards in Complex Mountain Areas Using VHR and Time-Series Images

Shuhui Jiao [1,2], Dingxiang Hu [3], Zhanfeng Shen [1,4,*], Haoyu Wang [1,5], Wen Dong [1,4], Yifei Guo [1,2], Shuo Li [1,2], Yating Lei [1,2], Wenqi Kou [1,4], Jian Wang [3], Huimei He [3] and Yanming Fang [3]

1 National Engineering Research Center for Geomatics, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China; jiaosh@aircas.ac.cn (S.J.); wanghaoyu@aircas.ac.cn (H.W.); dongwen200436@aircas.ac.cn (W.D.); guoyifei20@mails.ucas.ac.cn (Y.G.); lishuo@aircas.ac.cn (S.L.); leiyating21@mails.ucas.ac.cn (Y.L.); 2018321010229@cau.edu.cn (W.K.)

2 College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

3 MYbank, Z Space, No. 556 Xixi Road, Hangzhou 310013, China; dingxiang.hdx@mybank.cn (D.H.); wangjian.sz@mybank.cn (J.W.); huimei.hhm@mybank.cn (H.H.); yanming.fym@mybank.cn (Y.F.)

4 University of Chinese Academy of Sciences, Beijing 100049, China

5 School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: shenzf@radi.ac.cn; Tel.: +86-186-1114-9939

Abstract: Accurate and reliable farmland crop mapping is an important foundation for relevant departments to carry out agricultural management, crop planting structure adjustment and ecological assessment. The current crop identification work mainly focuses on conventional crops, and there are few studies on parcel-level mapping of horticultural crops in complex mountainous areas. Using Miaohou Town, China, as the research area, we developed a parcel-level method for the precise mapping of horticultural crops in complex mountainous areas using very-high-resolution (VHR) optical images and Sentinel-2 optical time-series images. First, based on the VHR images with a spatial resolution of 0.55 m, the complex mountainous areas were divided into subregions with their own independent characteristics according to a zoning and hierarchical strategy. The parcels in the different study areas were then divided into plain, greenhouse, slope and terrace parcels according to their corresponding parcel characteristics. The edge-based model RCF and texture-based model DABNet were subsequently used to extract the parcels according to the characteristics of different regions. Then, Sentinel-2 images were used to construct the time-series characteristics of different crops, and an LSTM algorithm was used to classify crop types. We then designed a parcel filling strategy to determine the categories of parcels based on the classification results of the time-series data, and accurate parcel-level mapping of a horticultural crop orchard in a complex mountainous area was finally achieved. Based on visual inspection, this method appears to effectively extract farmland parcels from VHR images of complex mountainous areas. The classification accuracy reached 93.01%, and the *Kappa* coefficient was 0.9015. This method thus serves as a methodological reference for parcel-level horticultural crop mapping and can be applied to the development of local precision agriculture.

Keywords: precise parcel extraction; map-level mapping; horticultural crop orchard; deep leaning; time-series data

Citation: Jiao, S.; Hu, D.; Shen, Z.; Wang, H.; Dong, W.; Guo, Y.; Li, S.; Lei, Y.; Kou, W.; Wang, J.; et al. Parcel-Level Mapping of Horticultural Crop Orchards in Complex Mountain Areas Using VHR and Time-Series Images. *Remote Sens.* **2022**, *14*, 2015. https://doi.org/10.3390/rs14092015

Academic Editor: Brigitte Leblon

Received: 16 March 2022
Accepted: 19 April 2022
Published: 22 April 2022

1. Introduction

Horticultural crop orchards are one of the most important agricultural production types in the world [1]. China is an important horticultural crop planting area in the world [2]. According to data from China's National Bureau of Statistics, the orchards area in China reached 1.23 billion hectares with an annual output of 2.87 billion tons in 2020, and a continuous growth trend was observed [3]. In China, the planting terrain of horticultural crops is very complex. Apples, for example, are mainly planted in plains, hillsides, terraces and loess hills, at altitudes of 50–2500 m.

Methods that quickly and efficiently obtain relevant information, such as orchard planting area, is of great significance to guide fruit production and planting structure adjustment planning [4]. The traditional way of obtaining farmland information mainly depends on a large number of manual field surveys, which consumes a lot of human resources and reduces economic efficiency, and the speed of data update is slow. In this case, the development of remote-sensing technology provides an advanced and convenient technical means to solve this problem, and it has been widely used in precision agriculture [5–7].

In crop classification research on a regional scale, medium-resolution images are widely used, though they are still referred to as high-resolution (such as Sentinel-2 or SPOT of 10 to 60 m resolution or Landsat of 30 m resolution) [8–12]. Studies have shown that, among these satellite products, the spatial resolution of optical images used for classification has little effect on the accuracy of classification in areas with large parcels, while Sentinel-2 images can provide better classification results than Landsat and MODIS images for areas with fragmented farmland [12]. Many scholars have carried out research on crop classification based on Sentinel-2 images [13–16].

Crop classification methods can be divided into two types: pixel-based methods and object-based methods [12]. In the existing research, pixel-based methods seem to be dominant [12,16]. Pixel-level classification is easy to implement, but it fails to identify the parcel boundaries. In addition, the classification results obtained based on this method have single random pixels (salt and pepper effect) [16]. The results usually need to be processed with additional filters to achieve smoothing and remove noise [12].

Considering this, the development of high-resolution images led to the creation of object-oriented surface cover classification methods. Compared with traditional methods, these methods can retain more semantic information and obtain land parcel distribution information [17]. Asli Ozdarici Ok et al. [18] explored the classification accuracy of pixel- and object-based classification methods based on a machine learning algorithm and showed that the results are of higher accuracy when obtained using the parcel-based classification method compared with the pixel-based method. This view has also been confirmed by many other scholars [16,19]. Object-oriented classification methods have been widely used in the field of agricultural management [20–23]. The accurate mapping of crop planting structure using an object-oriented method mainly includes two steps: parcel extraction and parcel type recognition.

There are three main methods of parcel extraction, which are edge-based methods, image segmentation methods and deep learning algorithms [24]. In edge-based methods, an edge operator is used to obtain a continuous region with large differences in pixel values to determine which pixels are the edge pixels in the image. Wen Dai et al. [25] used a canny edge detector to extract the edge information in remote-sensing images, which was then combined with the direction of contour lines to extract terrace parcels in a mountainous area. The image segmentation algorithm is widely used [26–29], whereby pixels are clustered, according to surface texture features, into unified geographical units with homogeneous features [27]. However, this method does not take the actual geographical meaning into account in crop parcel extraction. As a result, the extracted object is only a simple geographical patch, which is different from the actual farmland parcel [28]. The third method is the use of deep learning technology for parcel extraction. In recent years, deep learning algorithms have developed rapidly in the field of computer vision. The rise of deep learning drives the rapid development of image processing [30,31]. Many scholars have used neural networks in remote-sensing tasks, such as object detection [32–35], land cover classification [36–41] and scene classification [42,43]. There are two main types of neural network models used in parcel extraction, namely, the edge-based model and the texture-based model [17]. Compared with the other two methods, parcels of higher accuracy and with more complete semantic information are extracted using deep learning methods. In addition, the appearance of very-high-resolution (VHR) images provides clearer visual features for neural networks, and carrying out parcel extraction on these images is more convenient for the deep learning algorithm [17].

The current research on farmland parcel extraction mainly occurs in areas with simple topographic conditions. There are few studies on farmland mapping in complex mountainous areas. At present, the research on parcel extraction in mountainous areas mainly combines remote-sensing images with high-resolution DEM data to extract terrace edges [44,45]. It is thus necessary to carry out research on parcel extraction in complex mountainous areas using optical remote-sensing images.

For parcel classification, a common approach is to treat the parcel as a whole. The characteristics of the parcels are determined based on pixel average, such as values for average vegetation indexes and average reflectance, and the crop types in the parcel are then inferred according to these characteristics [17,46]. However, this method is based on the assumption that only one crop is planted in the parcels. This method is suitable for areas with simple planting conditions, in which mixed planting in the parcel can be ignored. As a matter of fact, mixed planting occurs in many horticultural crop planting areas, so this parcel classification method is not suitable for mountainous areas with complex planting structures.

Joanna Pluto-Kossakowska [12] summarized the research on multitemporal classification methods for crops and arable land with optical remote-sensing images in recent years; this research found that the most frequently detected plant species are usually dominant species cultivated in the temperate climate region of the Northern hemisphere, such as cereal, rapeseed, corn, sugar beet and grassland. However, there are few studies on the extraction and classification of horticultural crops. At present, it is still a challenge to use remote-sensing images for parcel-level mapping of artificial orchards. This is mainly due to the following reasons:

(a) Many horticultural crops belong to the same family as natural forests and have similar phenological characteristics, which makes it more difficult to distinguish them from each other.
(b) Horticultural crops are mainly planted in mountainous areas. Restricted by planting conditions, there are many agricultural parcels with irregular shapes and fuzzy edges [17]. Moreover, there is a high degree of heterogeneity between mountain parcels, which makes it difficult to extract them.
(c) There is a mixed planting phenomenon in many parcels, so the conventional method of determining the parcel category is not suitable for mixed parcels in complex mountainous areas.

To solve these problems, this paper presents a mapping method for horticultural crops at the parcel-level. Firstly, based on the zonal and hierarchical strategy, the parcel can be extracted step-by-step by combining the edge-based model RCF and the texture-based model DABNet. In this process, the use of a texture model can allow differentiation of the artificial orchard from the natural forest area, because they have great texture differences in VHR images, thus avoiding interference from the natural forest during subsequent orchard classification. Then, the vegetation indexes (NDVI, EVI, SAVI) were calculated based on the Sentinel-2 images, and the time-series characteristics of crops were constructed. Due to the existence of mixed parcels, we do not take parcels as the basic unit of classification but as spatial constraints, and we determine the category of parcels by filling pixel-level classification results into parcels. We propose a parcel filling strategy to ensure the statistical accuracy of the classification results. With this filling strategy, the classification results are filled into parcels to realize the accurate orchard mapping of complex mountainous areas. We carried out experiments in Miaohou Town, Qixia City, Shandong Province, China, and successfully extracted the distribution of apple orchards and cherry orchards in the study area and verified the effectiveness and feasibility of this method.

2. Study Area and Dataset

2.1. Study Area

In order to verify the effectiveness of the method proposed in this paper, Miaohou Town, Qixia City, Shandong Province, which is a typical horticultural crop planting city in China, was selected as the study area. The geographical location of the study area is illustrated in Figure 1. Miaohou Town is located between 37°05′05″–37°29′46″N and 120°32′45″–121°15′58″E and covers an area of approximately 84.89 km^2. It is characterized by a temperate monsoon climate, with an annual average temperature of 12 °C and an annual rainfall of 754 mm.

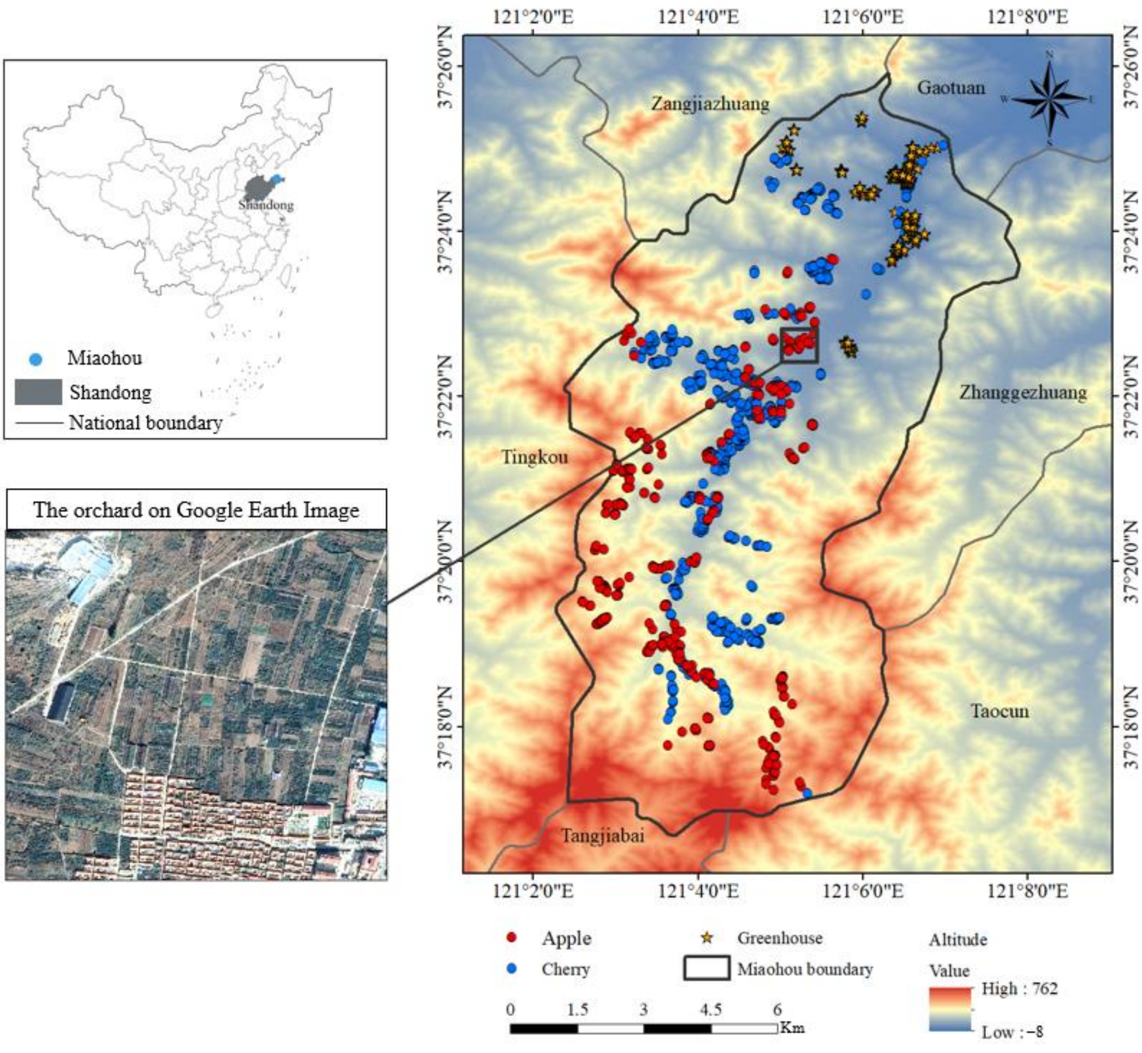

Figure 1. The location of the study area.

Miaohou is dominated by hills, and the altitude is higher in the south than in the north. The land cover types are mainly farmland, forest, woodland, buildings and water, which form a complex and diverse agricultural landscape. The local economic horticultural crops are mainly garden crops, including cherry and apple, and it is a cherry seeding base. These two horticultural crops are Rosaceae plants, and both are deciduous trees. Therefore, they have similar phenological characteristics throughout the year. In addition, a small amount of wheat, corn and peanuts are planted in the area. The phenological characteristics of these crops are shown in Figure 2 [2].

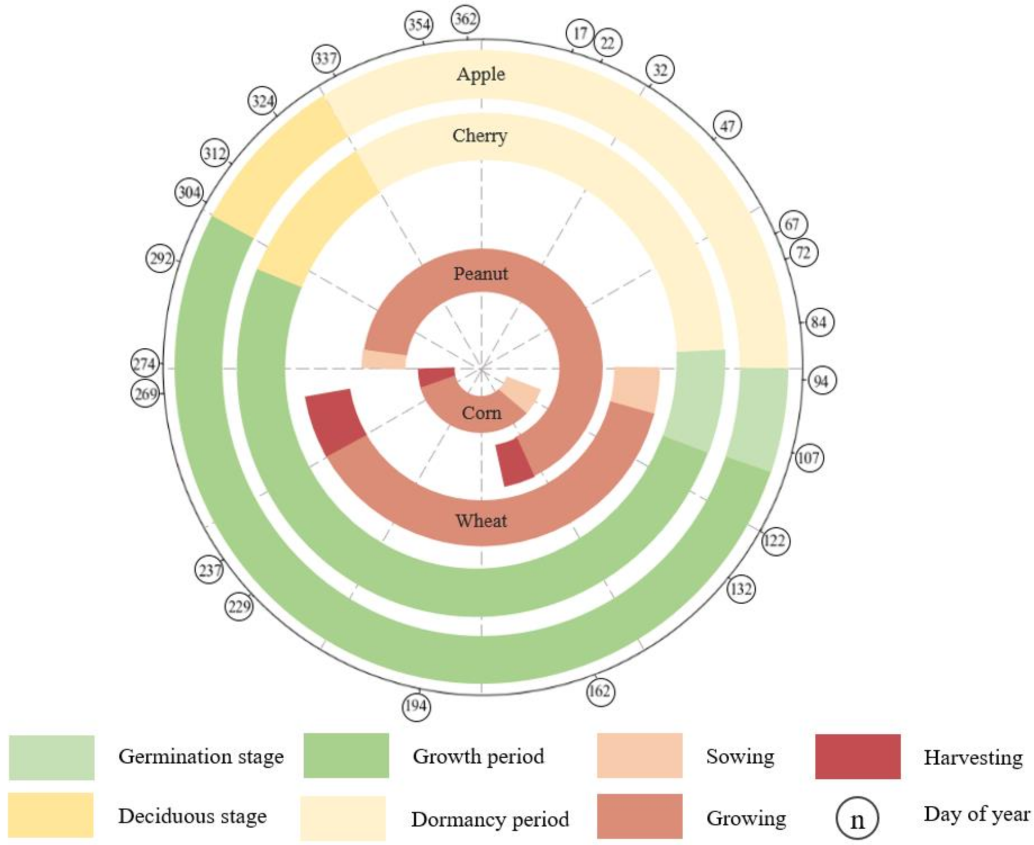

Figure 2. Crop phenological information and image acquisition time.

A large number of cherry and apple trees are planted in Miaohou Town, and the phenomenon of mixed planting is present. In addition to the traditional cherry orchard, there are many greenhouse cherries in the study area. Finally, the farmland in the study area is divided into four categories: the apple orchard, the cherry orchard, the greenhouse and the conventional cultivated land.

The best time to extract farmland parcels using VHR images is autumn. This is because the land surface is less covered by vegetation during this period, which allows more information to be obtained from the parcels.

2.2. Field Sampling Data

A field survey was conducted in the middle of May 2021 to collect field sampling points. We used GPS-enabled devices to collect the geographical coordinates of samples and corresponding crop categories. A total of 100 samples were collected, including 40 apple orchard samples, 40 cherry orchard samples and 20 greenhouse samples.

Furthermore, the sample points were further expanded by using VHR images for visual interpretation, and 1255 samples were obtained, including 430 apple orchard samples, 630 cherry orchard samples and 195 greenhouse samples. Finally, a total of 1355 sample points were obtained, and the distribution of these data are shown in Figure 1.

2.3. Remote-Sensing Data Acquisition and Preprocessing

In this study, Google satellite images were selected as VHR images for parcel extraction. The spatial resolution of the images is 0.55 m, which provides three bands of red, green and blue. The image contains rich ground spatial features and surface texture features. In terms of visual effect, the spatial distribution of agricultural parcels can clearly be seen. Taking the Sentinel-2 data for the reference image, the VHR image was geometrically corrected to ensure that the sample points have the correct correspondence of characteristic data. This permits the avoidance of subsequent data mapping errors that affect classification accuracy.

Sentinel-2 images have a spatial resolution of 10 m in visible and near-infrared bands, and the revisit period can reach 5 days. Therefore, it is very suitable for the construction of crop growth characteristics. The images are freely available from the European Space Agency (ESA) Copernicus Open Access Hub (https://scihub.copernicus.eu/without/dhus/#/home) (accessed on 13 July 2021). The product is an atmospheric apparent reflectance product that has been subjected to orthogonal correction and geometric fine correction, but not atmospheric correction. Therefore, the atmospheric corrected bottom (L2A) product was obtained by atmospheric correction of the L1C product to eliminate the atmospheric impact. When downloading data, the cloud amount of the image is set to below 10%. We browsed the images of the study area in 2019, and 24 images were finally obtained. With the exceptions of June and July, there are two or more Sentinel-2 images for each month to better simulate the temporal characteristics of crop growth. The data acquisition time is shown in Figure 2.

3. Methods

In this study, we propose a method for the precise mapping of horticultural crops in complex mountainous areas. The proposed method architecture is elaborated on in this section. The flow chart is shown in Figure 3 and mainly includes the following three parts:

1. A parcel extraction framework with zoning and hierarchical strategies based on VHR images. In this part, texture-based and edge-based deep learning models are combined to extract the parcels in the study area.
2. Crop classification based on time-series data. Based on Sentinel-2 images, the time-series characteristics of crops are constructed, and the land surface cover is classified into four categories using an LSTM algorithm.
3. For the complex agricultural planting situation in mountainous areas, we choose to take the parcel as a spatial constraint and fill it with pixel-level classification results to determine its category, rather than input the parcel into the classifier as a classification unit. A category filling strategy is designed. With this strategy, the categories of candidate parcels are determined based on the pixel-level classification results obtained in the second part. Finally, the distribution of horticultural orchards is obtained.

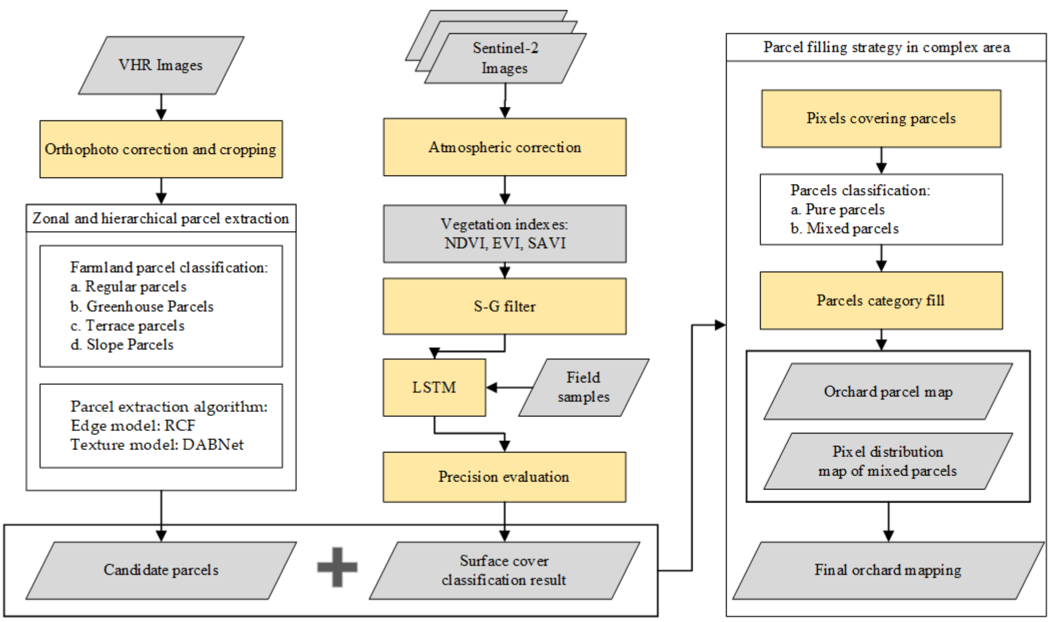

Figure 3. Workflow of the method proposed in this paper.

3.1. Parcel Extraction Based on a Hierarchical Extraction Scheme

3.1.1. Farmland Classification System Based on Geographical Divisions

In the traditional visual interpretation process, the concepts of zoning and hierarchy simulate the cognitive image processing of human vision and consider additional spatial information suitable for the perception of large-scale geographical entities. The idea of zoning and hierarchical strategy has shown good results in many previous studies [47–49]. Experiments have demonstrated that classification with the concept of zoning and stratification can significantly improve the accuracy of crop area estimation and reduce the field sampling cost [48].

Due to the influence of complex terrain, the spatial structure characteristics of cultivated land objects in mountainous areas differ greatly. This complexity reduces the extraction accuracy of classification algorithms. Therefore, it is unreasonable to use only one extraction algorithm (whether an edge-based model or a texture-based model) to extract the parcels in the whole complex region. To solve this problem, with consideration of the terrain conditions of the study area, this paper designs a zoning and hierarchical extraction scheme. Often, when implementing a zoning strategy, regions are mainly divided by some symbolic linear elements, such as rivers, roads and topographic lines [50]. However, since the urban part of the study area is small and the parcel distribution is greatly affected by the terrain, this paper uses topographic factors, such as slope and elevation, in the regional division. The parcel extraction process is shown in Figure 4.

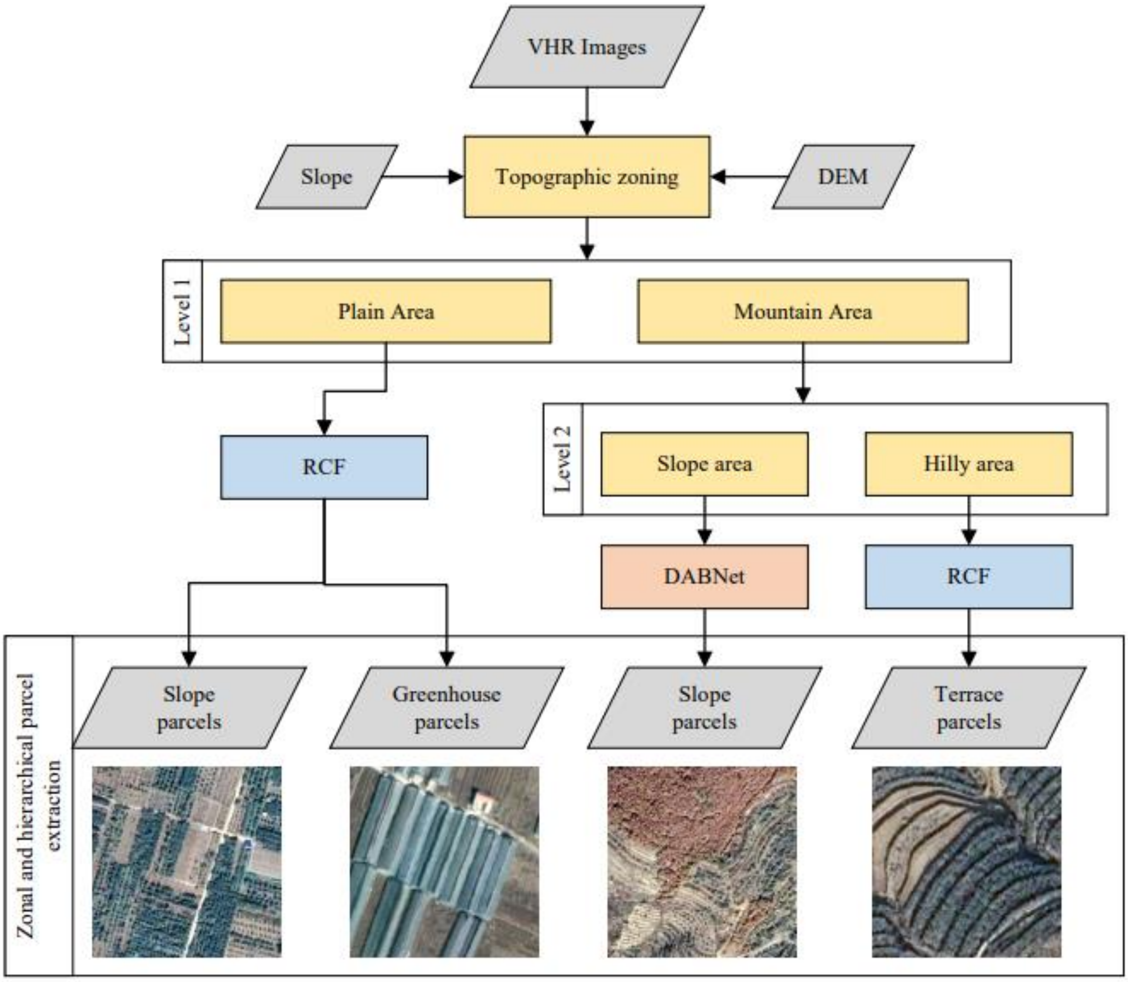

Figure 4. Parcel extraction flow chart.

The complex region can be divided into several relatively unified geographical regions by using the zoning and hierarchical extraction frame. Then, further division is carried out in these areas to extract farmland parcels. Initially, at the first level, based on the elevation and slope data, the whole study area is divided into plain area and mountainous area according to the terrain difference. With the same geographical conditions, the farmland parcels of each region have similar characteristics. After that, the plain area parcels are divided into regular parcels and greenhouse parcels. The second level is mainly for mountainous areas, which are divided into slope and terrace areas. The definition of these two areas mainly depends on whether the terrain has been artificially transformed. Crops in the slope area are directly planted on the hillside without artificially changing the terrain. The shape of the parcels in this area are irregular and have no obvious edge, but the texture features between artificially cultivated crops and naturally growing plants is obvious, and the former have a more regular texture. Terrace parcels are strip parcels or wavy sections artificially built along contour lines on hills or hillsides and have obvious edges in remote-sensing images. At the third level, slope parcels are defined by the slope area. DABNet is used to extract the slope parcels in the slope area according to the textural features. At the same time, horticultural crops and natural forests in the study area belong to the same family and species and demonstrate similar phenological characteristics. Only phenological features used for classification cause separability between horticultural crops, while the textures of artificial horticultural crops and natural forests are very different in remote-sensing images. When DABNet is used for slope parcel extraction, the natural forest and the orchards can be separated on VHR images based on spatial features. The terrace parcels are defined by the terrace area, and the parcels are extracted by the RCF model.

Finally, the crop parcels are divided into four categories: regular parcels, greenhouse parcels, terrace parcels and slope parcels. As shown in Table 1, each type has its own characteristics.

Table 1. Parcel type table.

Geographic Area	Farmland Type	Features
Plain Area	Greenhouse parcels	Regular shape, clear boundary, distributed in plain areas and different from the surrounding crop background.
	Regular parcels	Plain parcels, regular shape, clear boundary, uniform internal texture and uniform area.
Mountain Area	Terrace parcels	Long and narrow shape with uniform width, clear boundary, uniform internal texture and regular arrangement.
	Slope parcels	Fuzzy boundary, uniform internal texture, irregular shape, irregularly distributed on the hillside, great difference in area and mostly mixed parcels in Miaohou Town.

3.1.2. Parcel Extraction Based on the RCF Model

A convolutional neural network (CNN) is a multilayer perceptron with strong learning ability. Compared with traditional recognition algorithms, it can take the image as the input, avoiding the complex process of feature extraction and data reconstruction. This has great advantages for image processing. In recent years, it has been widely used in edge detection [31,51]. However, many existing CNN-based models only consider the feature of the last convolution layer when detecting the edge of objects. This results in the loss of a large amount of information [52].

In order to solve the problem of feature information loss, Yun Liu et al. [52] proposed that more accurate edge detection could be achieved by using richer convolutional features (RCF) based on VGG16. RCF combines the structural advantages of the VGG16 network [53] and the FCN [54]. Most of its network structure comes from VGG16. The convolution layers of RCF are divided into five stages that are connected through the pool layer. The main body of the network includes three parts: backbone network, deep supervision and feature fusion. Each stage performs deep supervised learning to make the networks converge as soon as possible. Then, the edge graphs of the five stages are fused, and the results are the output. Since RCF learns multiscale information, including the low layer and the target layer, and integrates the information of every layer in the network, the edge obtained through RCF, using only partial characteristics, is better than the CNN.

In this paper, the RCF model is used to extract the parcel edge. The classifier of the output layer of RCF is a sigmoid function, and the output value is the probability that the pixel is the parcel edge, ranging from 0 to 1. The larger the value, the greater the probability that the pixel is the parcel edge. Finally, the parcel edge is extracted by setting an appropriate threshold.

3.1.3. Parcel Extraction Based on DABNet

Semantic segmentation is a pixel-level prediction task. In order to improve the prediction effect, many researchers expand the convolution model depth to increase the acceptance domain of the network and capture more complex features. However, the use of more layers also requires more running time and memory. Therefore, the network makes a trade-off between prediction accuracy and speed to ensure the optimal overall performance of the model. In this paper, the DABNet model was selected to extract parcels with fuzzy edges but clear textures in mountainous areas.

DABNet is a lightweight semantic segmentation model. It can make full use of contextual information with significantly reduced parameters. It combines the advantages of bottleneck designed in RESNet and factorized convolutions in ERFNet [55,56]. A depth-wise asymmetric bottleneck (DAB) module is proposed, which achieves a balance between the speed and accuracy of the algorithm [57].

As shown in Figure 5, a 3 × 3 convolution is used in the DAB module to reduce the number of channels and to avoid establishing a deeper model. In the DAB module, two branches are used to extract features. As referred to by the non-bottleneck-1D module of ERFNet, in the first branch, the 3 × 3 depth-wise convolution is substituted for a 3 × 1 depth-wise convolution, followed by a 1 × 3 depth-wise convolution. In the second branch, only a dilated convolution is applied to the depth-wise asymmetric convolution to reduce computational cost. Then, the information of the two branches is superimposed together, and the number of channels is restored through a 1 × 1 convolution. Finally, the input features are superimposed as the output.

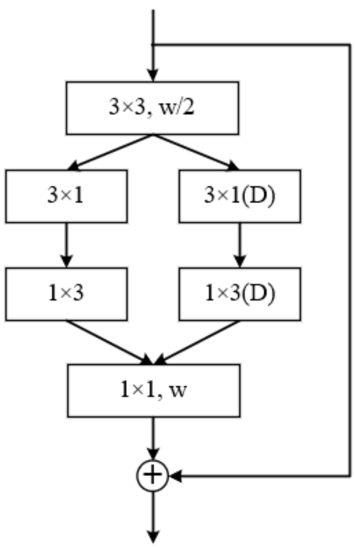

Figure 5. DAB module. W: the number of input channels; D: dilated convolution.

3.2. Horticultural Crop Classification with Time-Series Images

3.2.1. Time-Series Feature Construction

Vegetation indexes are obtained by algebraic computation between different bands of satellite images, which can enhance vegetation information and reflect the development of vegetation and soil [58]. These indexes are often used for vegetation extraction. Based on the planting situation in the study area, three indexes, namely the normalized difference vegetation index (NDVI), the enhanced vegetation index (EVI) and the soil regulation vegetation index (SAVI), are selected for analysis [59–61].

The normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI) can reflect changes in vegetation biomass. The EVI has strong anti-saturation and is more sensitive to biomass differences among crops when the biomass is high, while the NDVI is more sensitive when the biomass is low [58]. In addition, in the study area, the plant spacing of apple trees and cherry trees is generally 2–3 m, while the row spacing of conventional crops is only 20–30 cm. Therefore, in autumn, after the apple and cherry trees lose their leaves, the soil background between horticultural crops and conventional crops is very different in remote-sensing images. In light of this, we use SAVI to reflect the soil background of crops [58].

The formulas of these vegetation indexes are as follows:

$$\mathrm{NDVI} = \frac{NIR - R}{NIR + R} \tag{1}$$

$$\mathrm{EVI} = \frac{2.5 \times (NIR - R)}{NIR + 6R - 7.5B + 1} \tag{2}$$

$$\mathrm{SAVI} = \frac{(1 + L)(NIR - R)}{NIR + R + L} \tag{3}$$

where NIR is the near-infrared band, R is the red band, B is the blue band and L is the soil regulation coefficient, which is noted as 0.5 in this paper.

Due to the interference of the cloud and the atmosphere, there are irregular fluctuations in the time-series curve, which affects the classification results. In order to eliminate noise, the S–G filter is used to smooth the time-series curve to reflect the actual growth of crops [62].

The time-series curve is reconstructed using S–G filtering. The expression for S–G filtering is:

$$R_i^* = \frac{\sum_{j=-n}^{n} A_j \cdot R_{i+1}}{M} \tag{4}$$

where R is the original value of the vegetation index; R_i^* is the fitting value of R; A_j is the filter coefficient of the jth index value; R_{i+1} is the $i+1$ index value in the time-series; M is the filter length; n is the size of the moving window.

3.2.2. Classification of Parcels Based on an LSTM Model

The recurrent neural network (RNN) is a very powerful neural network model for processing and predicting sequence data [63]. RNN overcomes many limitations of traditional machine learning on input data and is widely used in the field of deep learning. However, problems still exist in the application of RNN. That is, after multiple cycles, a gradient disappearance or explosion occurs. In this case, long short-term memory (LSTM) is proposed, which solves the problem of error backtracking in traditional RNN networks [64]. It has been proved that LSTM can provide higher classification accuracy than other methods in predicting time-series data [58,65,66].

A special structure is designed in the LSTM network. By controlling the internal state of the error flow through the special structure, the problem of gradient disappearance and explosion in the training process can be solved. The unit structure is shown in Figure 6. It is composed of two cell state units and three gates, and the two states are the cell state $\boldsymbol{c}_t$ and the hidden state $\boldsymbol{h}_t$. The cell state $\boldsymbol{c}_t$ provides the network with the ability to store and load information at any point of the input sequence, which allows the network to solve the problem of long-term dependence on time-series data. The hidden state $\boldsymbol{h}_t$ conveys information from the previous event to the next unit, and it is overwritten at each step.

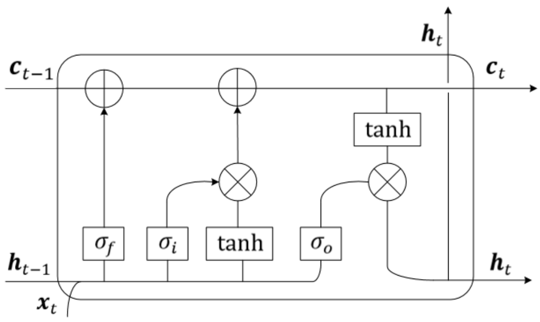

Figure 6. LSTM cell structure.

The network designs three gates to control the transmission of information to the state vector. The three gates are the input gate $\boldsymbol{i}_t$, the forgetting gate $\boldsymbol{f}_t$ and the output gate $\boldsymbol{o}_t$. Gates can be regarded as a fully connected layer, which are mainly used to control the flow of information and can facilitate the storage and update of information. The input gate $\boldsymbol{i}_t$ is used to control how much current input data $\boldsymbol{x}_t$ in the network flows into the memory unit; that is, how much information can be saved to the cell state $\boldsymbol{c}_t$. The forgetting gate $\boldsymbol{f}_t$ is a key component of LSTM, which can control the retention or forgetting of information. That is, it can control how much influence the information in the cell state at the previous time $\boldsymbol{c}_{t-1}$ will have on the current cell state $\boldsymbol{c}_t$. The output gate $\boldsymbol{o}_t$ controls which informational part of the cell state will be output at the current time. Through the combination of two state units and three gates, the network can control the flow of information throughout the network.

3.3. Classification Result Filling Strategy

The detailed flow chart of the classification strategy for candidate parcels is shown in Figure 7. First, according to the results in Sections 3.1 and 3.2, the filling pixels (Sentinel-2) of each parcel are obtained. Parcels are divided into three types according to the coverage relationship between parcels and pixels. Then, the parcels are filled accordingly based on the category of pixels covering them (i.e., their filling pixels). In this process, we carefully considered the mixed planting phenomenon in parcels. The concept of mixed parcels has been put forward by scholars in previous studies [67]. Guanyuan Shuai et al. [68] divided the parcels into pure and mixed parcels when mapping the distribution of corn and replaced parcel-level classification with pixel-level classification. Finally, they concluded that the classification method combining parcel- and pixel-level classification is applicable to many agricultural systems with small landownership in which intercropping is very common. Therefore, considering the complex planting situation, we think that it is necessary to consider the mixed parcels separately in Miaohou.

Considering the planting situation, the farmland parcels are divided into three types according to the coverage and the surrounding relationship between the parcels and pixels of the classification result: parcels surrounded by multiple pixels, multi-pixel-covered parcels and single-pixel-covered parcels.

1. Parcels surrounded by multiple pixels. This kind of parcel has a large area, so it contains many pixels (Sentinel-2). All of the pixels contained in the parcel and the pixels whose coverage area is greater than half of the pixel area at the parcel boundary are used as the filling pixels of the parcel.
2. Multi-pixel-covered parcels. When the area where the pixel intersects the parcel is greater than half of the pixel area, the pixel is regarded as a filling pixel of the parcel.
3. Single-pixel-covered parcels. For a single pixel overlay parcel, the pixel covering the parcel is the filled pixel of the parcel.

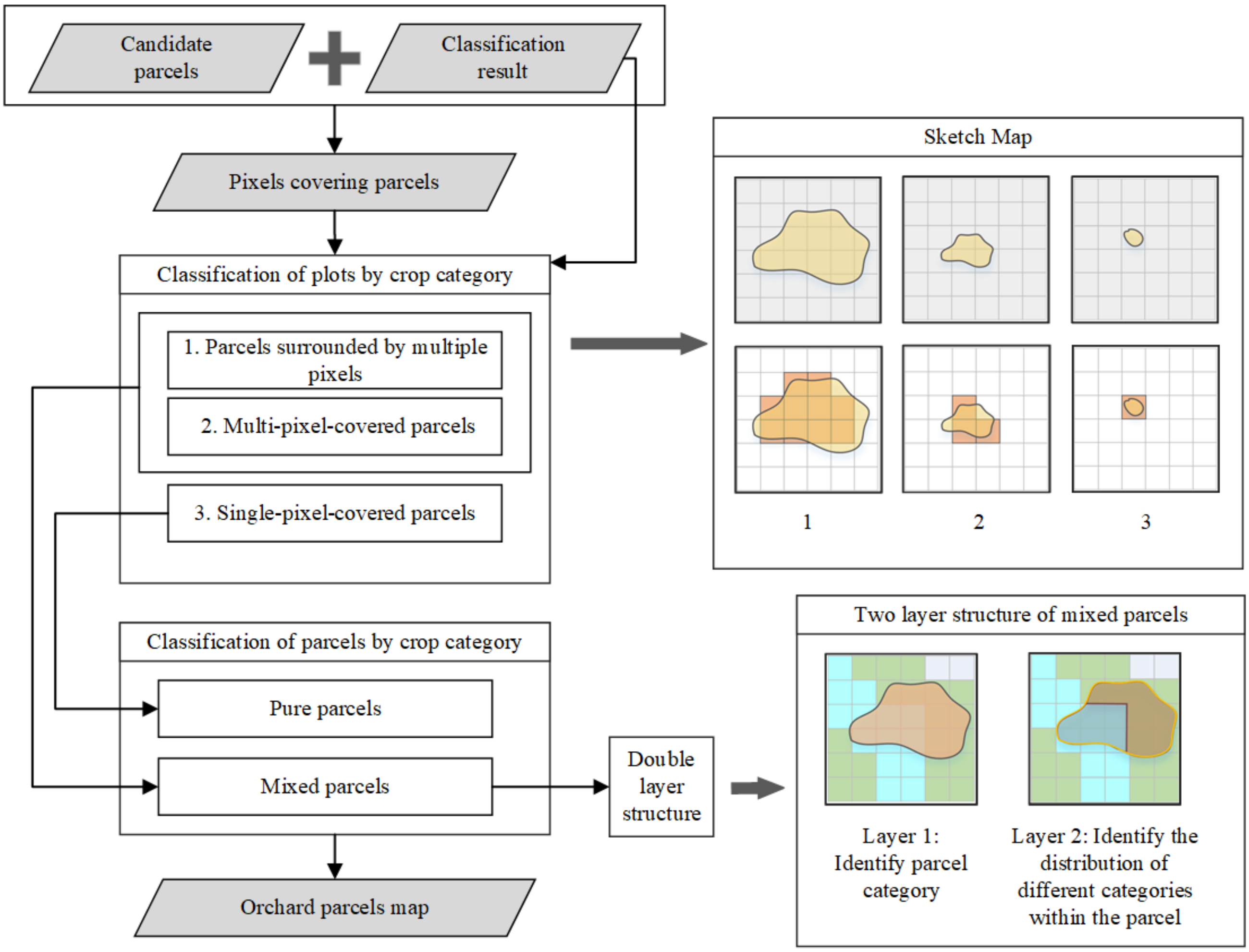

Figure 7. Parcel category filling strategy.

Subsequently, the parcel category is determined according to the pixels covering it. In this step, we divide parcels into pure parcels and mixed parcels. Pure parcels contain only one crop, while mixed parcels contain at least two crops. There are a large number of mixed parcels in the study area. The main reasons for their existence are as follows:

1. Cherry trees have gradually invaded the apple orchard and are planted in an invasive manner in apple orchards. In some parcels, the distribution of cherry trees has no regular boundary, which leads to the complexity and diversity associated with mixed planting in many parcels.
2. In VHR images, the texture of apple and cherry trees is similar. When DABNet is used to extract large slope parcels according to textural features in the mountain area, the two fruit trees are classified into one singular category. Therefore, the neural network can only obtain orchard parcels from the surface coverage, rather than detailed apple orchards or cherry orchards. If both apple and cherry orchards are forcibly used as DABNet input categories, then not only will the effectiveness of parcel extraction be greatly reduced, but the parcel will also be overly divided, becoming very different from the actual parcel and losing semantic information.

Obviously, the single-pixel-covered parcel must be a pure parcel, while the first and the second types of parcels may be pure or mixed parcels. When determining the final parcel category, we first determine whether the pixels covering them belong to the same category. If they belong to the same category, the parcel is pure, and the pixel category is the parcel category. Otherwise, the parcel is a mixed parcel. In order to facilitate the display of parcels and the orchard information statistics, we use a two-layer structure when storing and displaying mixed parcels, as shown in Figure 7. The first layer is used to identify the parcels as mixed parcels. In the second layer, the crop categories inside the mixed parcel are displayed, and the distribution positions of different crops in the parcel are outlined with pixels as the boundary. When calculating the area of crops of different categories, the pixel area of each category in the mixed parcel is taken as the statistical value.

In order to evaluate the effectiveness of this method, a confusion matrix was constructed based on field survey data, and the overall accuracy (OA), producer accuracy (PA), user accuracy (UA) and $Kappa$ coefficient were calculated to evaluate the accuracy of the results.

4. Result

4.1. Candidate Parcel Extraction Results

The candidate parcels in the study area were obtained based on the parcel extraction framework (zoning and hierarchical). The distribution of parcels is shown in Figure 8. Different colors represent different types of parcels. Finally, 13,223 candidate parcels with a total area of 4107.7 ha were extracted. The number and area information of different types of parcels are recorded in Table 2.

Table 2. Statistical table of different parcels.

Parcel Type	Number	Total Area (m^2)	Average Area (m^2)
Regular parcels	3151	8,104,969	2572
Greenhouse parcels	207	605,478	2925
Slope parcels	903	22,332,936	24,732
Terrace parcels	8962	10,033,835	1120
Total	13,223	41,077,218	3106

We assessed the statistics of the area information of different types of parcels, as shown in Figure 9. Since there are fixed specifications when the greenhouse is built, there is no additional discussion here. Therefore, the information obtained from regular parcels, terrace parcels and sloping parcels is included. We found that the area of plain parcels is mainly below 4000 m^2, while the parcels with an area below 2000 m^2 account for nearly 70%. The area of terrace parcels is mainly below 1800 m^2, accounting for about 83.2%. It fully reflects the characteristics of small-scale farming patterns in the study area, namely that the parcel is broken, and the area is generally small. The area of sloping parcels is generally larger than that of terrace and regular parcels.

4.2. Time-Series Curve Construction Results

The time-series curves of land cover were constructed based on the Sentinel-2 images, and the Savitzky–Golay filter was used to smooth the curves. The land cover is divided into four categories: the apple orchard, the cherry orchard, the greenhouse and the non-orchard area. Non-orchard areas mainly include conventional cultivated land (peanuts, corn and wheat), natural forest and buildings. Therefore, considering the actual classification requirements, the time-series characteristics of six kinds of ground objects were collected. The curves before and after S–G filtering are shown in Figure 10.

It can be seen from Figure 10 that the time-series curves of the three vegetation indexes after filtering are in good agreement with the data before filtering. Compared with the original curves, the filtered curves are smoother and can better reflect the phenological characteristics of different objects. Through filtering, the noise in the original data can be eliminated, and the change in biomass during the process of vegetation growth can be better simulated.

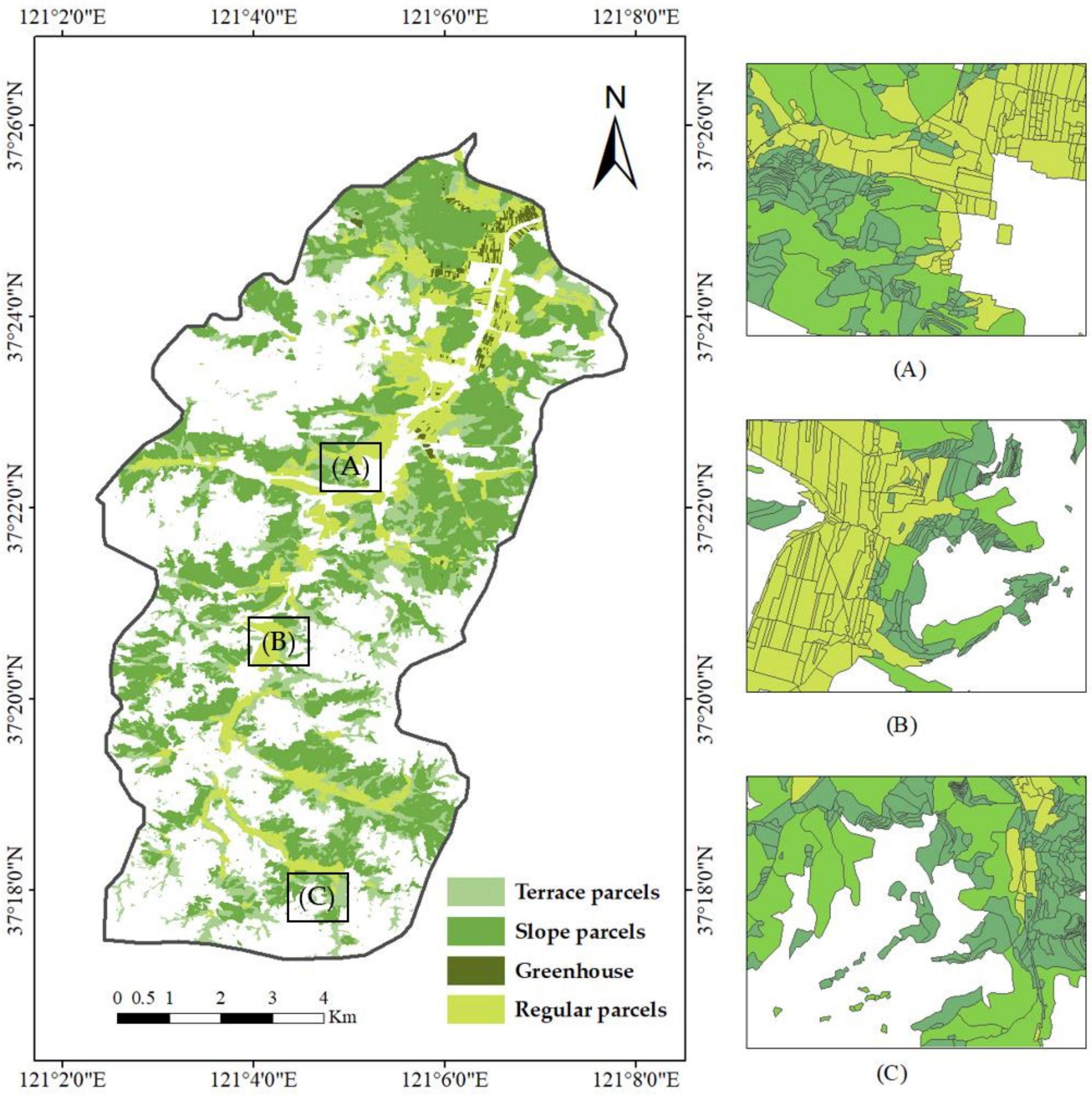

Figure 8. The distribution map of farmland parcels in the study area using the proposed method. (**A**–**C**) are subregions of the study area.

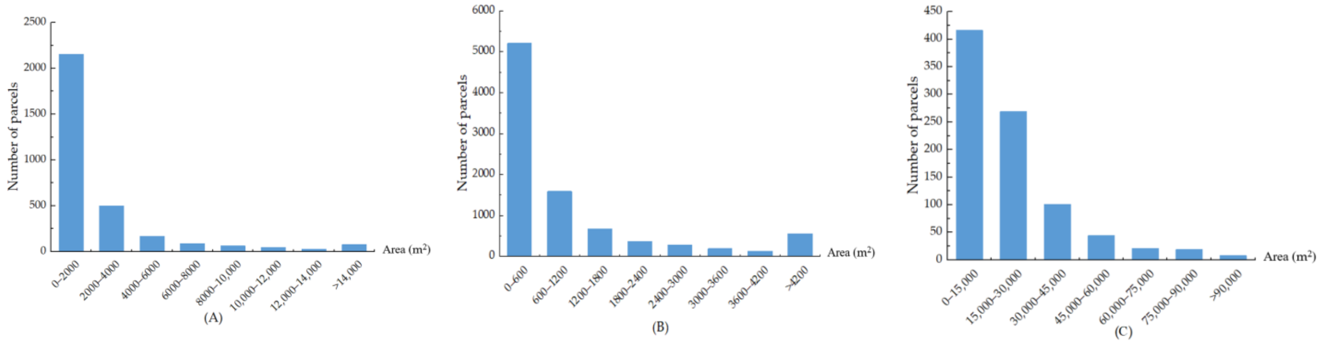

Figure 9. Area statistics of different types of candidate parcels. (**A**) is the area information of regular parcels. (**B**) is the area information of terrace parcels. (**C**) is the area information of slope parcels.

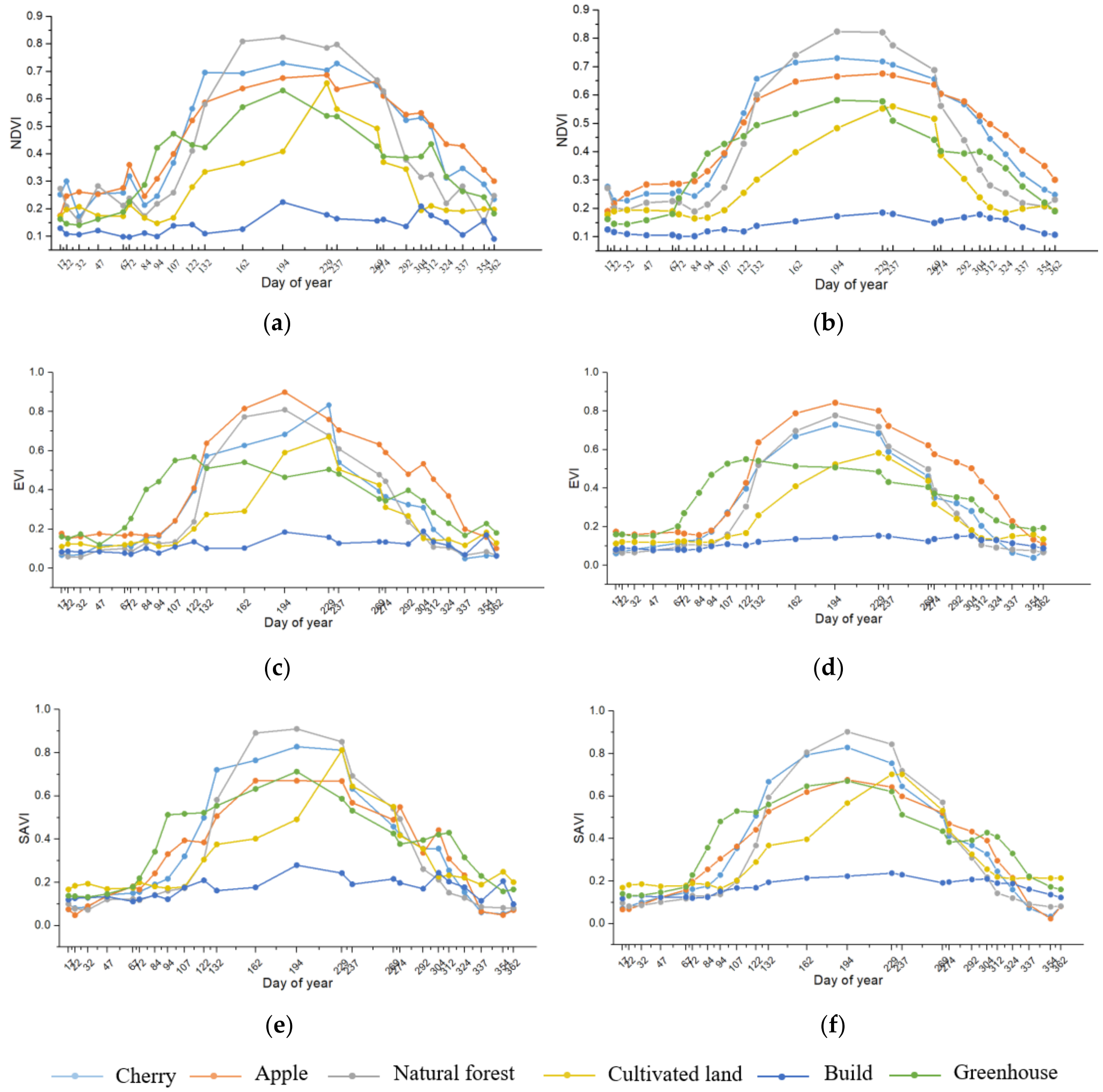

Figure 10. Vegetation index time-series before and after Savitzky–Golay (S–G) filtering. (**a**,**c**,**e**) Original vegetation index time-series curve; (**b**,**d**,**f**) Vegetation index time-series curve after S–G filtering.

By analyzing the vegetation index time-series curves, we found that different objects have differences in the characteristics of long-time-series. From the diagram, it can be seen that the apple orchard, the cherry orchard and the natural forest have similar characteristics on the whole. Their index values began to increase at the end of March and started to decline rapidly around the beginning of October. However, there are differences in the peak of the curve, with the NDVI of natural forest > cherry orchard > apple orchard. In the growing season, the NDVI of apples is always slightly lower than that of cherries, but after the defoliation period (274 days), the NDVI of apples is higher than that of cherries. In the field investigation, we found that the leaves of apple trees are denser than those of cherry trees. This can be seen from the EVI curves. In the garden crop growing season, the biomass of the apple orchard is higher than that of the cherry orchard, and the EVI value is higher. The canopy coverage is higher for cherry trees than apple trees. In the SAVI index curves, the index value of the cherry orchard is higher than that of the apple orchard in the growing season. The three vegetation index values of cultivated land are lower than those of orchards in the whole growing season, and the time when the index values peak in the time curve is later for cultivated land than for orchards. At approximately mid-August, a peak is reached. There are great differences between greenhouse crops and other crops, especially in terms of EVI curves. It can be seen that the index values of greenhouses indicate that the growth season is reached earlier than for other crops, and the peak can be reached earlier. The growing season of greenhouse crops is very long. The vegetation indexes of greenhouses begin to rise rapidly starting at the beginning of March (67 days), indicating that the plants grow rapidly and reach the peak at the beginning of May (123 days). This difference is exhibited mainly because the greenhouse can provide a more favorable growth environment for premature crops. The vegetation index values of built area are relatively low, and the curve fluctuation is stable.

4.3. Accuracy Evaluation of LSTM Model Parameters and Classification Results

LSTM works well over a broad range of parameters such as learning rate, input gate bias and output gate bias [17]. There are two main parameters, those being the number of network layers and neurons in hidden layers, which have a great impact on the classification performance of LSTM. Different experiments were carried out to explore the classification accuracy and stability of the classifier with different parameters. The overall accuracy of LSTM with various parameters is shown in Figure 11. When carrying out the experiment, the number of neurons ranged from 2 to 60 with a step of 2, and the number of network layers ranged from 1 to 6 with a step of 1.

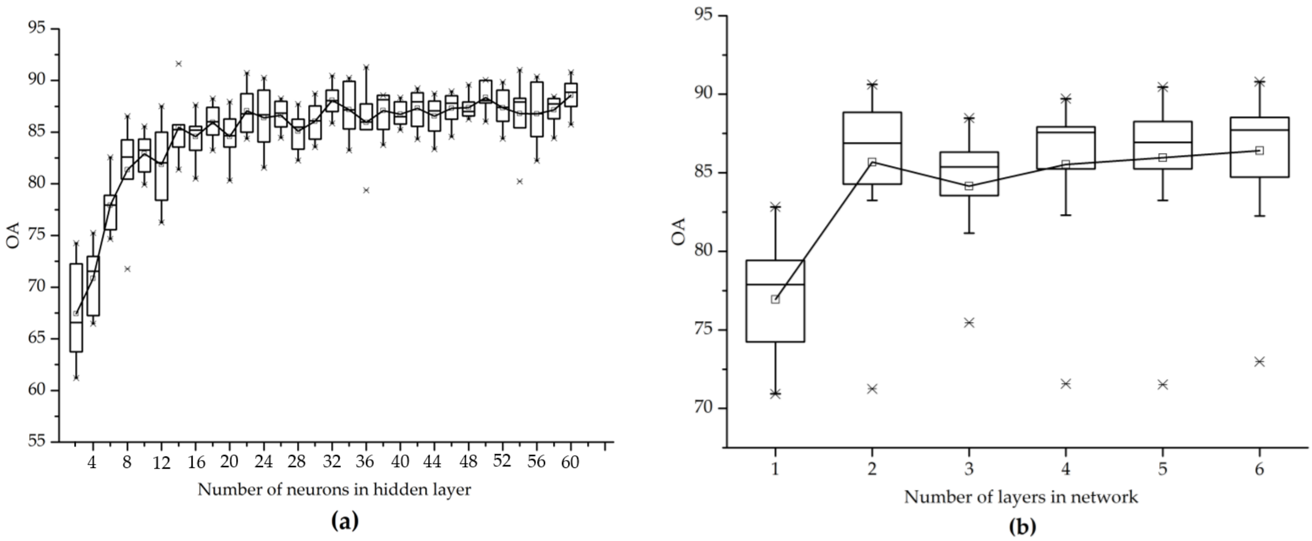

Figure 11. (**a**) The relationship between classification performance and the number of hidden layer neurons. (**b**) The relationship between classification performance and the number of layers in the network.

After performing the experiment and analysis and comprehensively considering the classification accuracy and stability of the network, 2 network layers and 32 neurons were set as the optimal network parameter configuration. In order to obtain crop mapping in the study area as early as possible in practical production, we explored the classification accuracy of time-series with different lengths. According to crop growth characteristics, seven time-series combinations were set up: day 17 to 107 with 9 images, day 17 to 132 with 11 images, day 17 to 194 with 13 images, day 17 to 237 with 13 images, day 17 to 274 with 15 images, day 17 to 304 with 19 images and the complete time-series data.

A box diagram of the classification accuracy of the different time-series is shown in Figure 12. Initially, when more images were added to the classifier, the classification accuracy increased; however, in the third and fourth groups of experiments, the classification accuracy showed a downward trend. Through analysis, we believe that this is because the study area is located by the sea and in the summer, cloudy and rainy conditions affect the quality of optical images. Although the S–G filter was used to reconstruct time-series data, the decline in data quality still reduces the data classification accuracy. After that, the accuracy continued to increase. By October (17–304), the classification accuracy tended to be stable. Although the accuracy of the last group of experiments was improved, the impact was not significant. Therefore, we believe that in the actual production, the time-series from January to October can be used to achieve good classification.

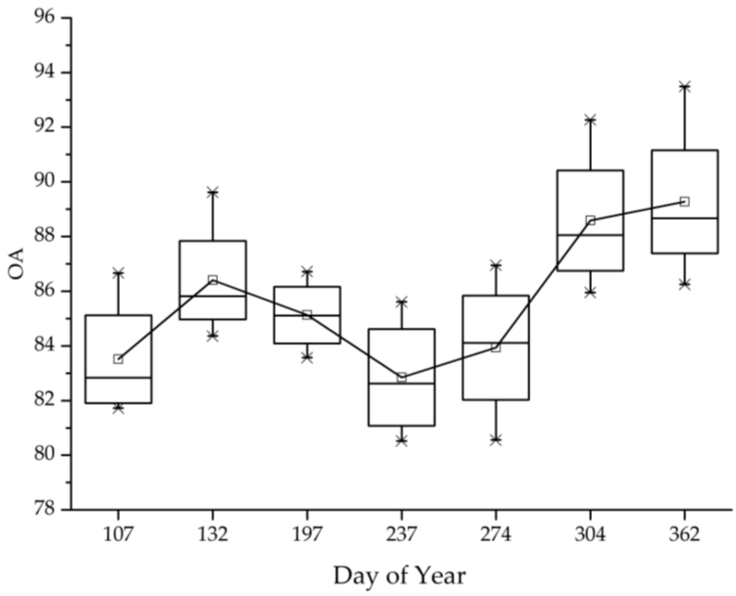

Figure 12. The relationship between classification performance and time-series data. The horizontal axis represents the time-series data combination from the seventeenth day to the indicated time.

Finally, the confusion matrix of the classification results is constructed, and the classification accuracy is calculated, as shown in Table 3. The overall classification accuracy of the results is as high as 93.01%, which demonstrates the effectiveness of the LSTM model.

4.4. Parcel Fill Result

According to the parcel filling strategy proposed in this paper, categories of candidate parcels were determined based on the classification results of Section 4.3. The parcel category distribution map is shown in Figure 13. It shows the pure apple orchard parcels, pure cherry orchard parcel, mixed parcels and non-orchard parcels. The information obtained from various parcels is shown in Table 4.

Table 3. Confusion matrix and precision table.

	Apple	Cherry	Greenhouse	Non-Orchard
Apple	163	24	0	10
Cherry	12	337	0	1
Greenhouse	0	1	191	3
Non-orchard	16	18	2	467
Producer accuracy	85.34%	88.68%	98.96%	97.08%
User accuracy	82.74%	96.29%	97.95%	92.84%
Overall accuracy	93.01%			
Kappa	0.9015			

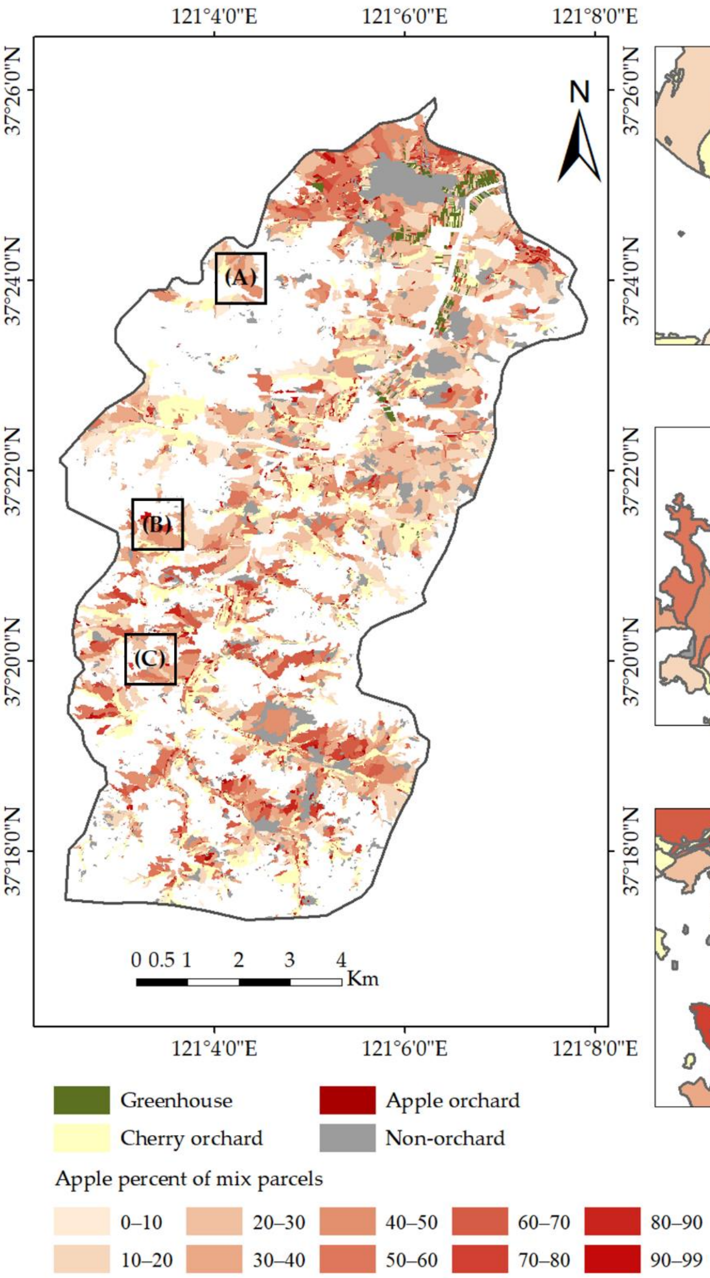

Figure 13. Parcel category fill results. (**A**–**C**) are subregions of the study area.

Table 4. Information obtained from different types of parcels.

Parcel Type	Number	Area (m^2)
Pure apple orchard parcel	1084	380,784
Pure cherry orchard parcel	4163	6,200,552
Greenhouse	207	605,478
Non-orchard parcel	3262	5,435,964
Mixed parcel	4507	28,454,440

In the mixed parcels, both cherry and apple trees are planted. For the mixed parcels, we use different colors to represent them according to the proportion of apple trees in them. The darker the color, the higher the proportion of apple trees. Three subregions with many mixed parcels were selected to show the two-layer structure of mixed parcels, which illustrates the effectiveness of the filling strategy. Based on the final results, we reviewed statistics on the planting information in the study area. The planting area of crops comes from two aspects: the area of pure orchard parcels and the area of orchard in mixed parcels, which can be obtained by multiplying the area of the mixed parcel by the proportion of orchard within it. The apple orchard area in the study area is calculated as 711.8 hectares, and the cherry orchard area as 1968.2 hectares.

5. Discussion

This paper proposes a method for parcel-level horticultural crop mapping in complex mountainous areas. Experiments show that this method classifies well in areas with a complex planting structure.

1. We used a hierarchical framework to extract parcels layer by layer. It is worth noting that in practical application, the geographical characteristics of the region should be fully considered using a zoning and hierarchical strategy. Inappropriate zoning policies increase the work intensity but reduce the classification accuracy. The study area of this paper is located in complex mountainous areas. Therefore, for practical consideration, we mainly use the terrain characteristics.
2. Due to the complex planting situation in mountainous areas, we selected two deep learning models for parcel extraction, and we obtained a good extraction effect. The determined parcel distribution is very close to the actual situation. However, the combination of the two models is not always needed for parcel extraction. For an area with simple planting, the edge characteristics of the parcels are very clear, and the edge model alone is sufficient to extract the parcel distribution.
3. In the case of mixed planting among crops, we choose to use the results based on pixel-level classification to determine the parcel category, rather than preferentially constructing the features of the parcel. If the parcel features are first constructed based on the mean value of pixels and then classified, the mixed parcels are likely to be misclassified because their features are not close to any crop. The parcel filling strategy in this paper can be used to avoid a situation in which the crop area is incorrectly estimated.

The results of the experiments in Miaohou Town, Qixia, show that the method proposed in this paper demonstrates good performance in parcel-level mapping of orchards. However, some problems still exist with this method, which can be improved upon in future work.

1. The parcel extraction framework mentioned in this paper is mainly dependent on VHR optical images. There is also a limitation regarding the image acquisition time, preferably measured in autumn. However, the long revisit cycle of VHR images limits the acquisition from data sources.
2. The study area is close to the sea, and there are too many clouds and rain in the summer, which affects the optical image quality. Hence, only using Sentinel-2 datasets to construct temporal features leads to large intervals between image sequences in July

and August. To solve this problem, in future work, it is necessary to fuse multisource data to construct more accurate crop characteristic curves.

6. Conclusions

It is difficult to carry out accurate mapping of horticultural crops at the parcel-level in mountainous areas due to their complex terrain and circumstances related to small-scale farming. In this study, we propose a method to meet the needs of parcel-level mapping in complex areas. This method combines the characteristics of VHR optical images and temporal optical images. The core idea includes three parts: parcel extraction based on a zoning and hierarchical framework, time-series data classification and parcel category filling. The process is as follows: firstly, based on the VHR image, the candidate parcels are extracted using the zoning and hierarchical parcel extraction framework. Then, land cover classification is achieved based on temporal optical images. Finally, the categories of parcels are obtained by using the parcel filling strategy.

A parcel-level mapping experiment was carried out in Miaohou Town, China, with VHR Google images and time-series Sentinel-2 images, and this method demonstrated good performance. Based on verification via visual inspection, it was determined that there was effective parcel extraction performed by RCF and DABNet. Based on the time-series data of different lengths, several groups of experiments were carried out. Through the evaluation of accuracy, we found that the classification accuracy using full time-series data is the highest in the study area, while the classification accuracy using data before October is also very good, demonstrating less loss when compared with the former. Therefore, we think that the orchard distribution can be extracted earlier using images from January to October for classification procedures. The classification accuracy was reduced by cloud interference in optical images from July and August. By adjusting and optimizing the parameters, the overall accuracy of the final classification results was calculated as 93.01%, and the *Kappa* coefficient was 0.9015, which demonstrates the effectiveness of time-series classification based on LSTM. For complex planting situations, we divided parcels into pure parcels and mixed parcels separately, which can avoid situations in which crop planting information is incorrectly estimated. Finally, the local planting information was obtained through statistics.

In future work, we will explore more farmland parcel extraction methods, optimize the land parcel extraction effect, fuse multisource data to build a more realistic and high-precision time-series curve and realize the land parcel-level extraction of multiple crops in mixed land parcels.

Author Contributions: S.J. proposed the methodology, prepared the data, performed and analyzed the experiments and wrote the manuscript. S.J., D.H., H.W. and Y.G. helped conduct the field investigation and classification experiments. Z.S. outlined the research topic, revised the methodology, acquired the funding and revised the manuscript. S.L. and W.D. helped improve the methodology and carry out the parcel extraction experiment. Y.L., W.K. and J.W. helped realize some algorithms related to GIS spatial analysis. H.H. and Y.F. helped prepare sample data and improve the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (2021YFB3900005), the National Natural Science Foundation of China (41971375), Third Xinjiang Scientific Expedition Program (Grant No.2021xjkk1400), the Chongqing Agricultural Industry Digital Map Project (21C00346) and the Xinjiang Tianshan innovation team project (2020D14016).

Data Availability Statement: Data sharing is not applicable to this article.

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

References

1. Kozhoridze, G.; Orlovsky, N.; Orlovsky, L.; Blumberg, D.G.; Golan-Goldhirsh, A. Classification-based mapping of trees in commercial orchards and natural forests. *Int. J. Remote Sens.* **2018**, *39*, 8784–8797. [CrossRef]
2. Zhu, Y.; Yang, G.; Yang, H.; Wu, J.; Lei, L.; Zhao, F.; Fan, L.; Zhao, C. Identification of Apple Orchard Planting Year Based on Spatiotemporally Fused Satellite Images and Clustering Analysis of Foliage Phenophase. *Remote Sens.* **2020**, *12*, 1199. [CrossRef]
3. National Bureau of Statistics. Available online: http://www.stats.gov.cn/ (accessed on 16 May 2021).
4. Yang, Y.P.; Huang, Q.T.; Wu, W.; Luo, J.C.; Gao, L.J.; Dong, W.; Wu, T.J.; Hu, X.D. Geo-Parcel Based Crop Identification by Integrating High Spatial-Temporal Resolution Imagery from Multi-Source Satellite Data. *Remote Sens.* **2017**, *9*, 1298. [CrossRef]
5. Ashourloo, D.; Shahrabi, H.S.; Azadbakht, M.; Aghighi, H.; Nematollahi, H.; Alimohammadi, A.; Matkan, A.A. Automatic canola mapping using time series of sentinel 2 images. *ISPRS J. Photogramm. Remote Sens.* **2019**, *156*, 63–76. [CrossRef]
6. Beeri, O.; Peled, A. Geographical model for precise agriculture monitoring with real-time remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 47–54. [CrossRef]
7. Meroni, M.; Marinho, E.; Sghaier, N.; Verstrate, M.M.; Leo, O. Remote Sensing Based Yield Estimation in a Stochastic Framework—Case Study of Durum Wheat in Tunisia. *Remote Sens.* **2013**, *5*, 539–557. [CrossRef]
8. Xie, D.F.; Sun, P.J.; Zhang, J.S.; Zhu, X.F.; Wang, W.N.; Yuan, Z.M.Q. Autumn crop Identification using high-spatial-temporal resolution time series data generated by modis and landsat remote sensing images. In Proceedings of the IEEE Joint International Geoscience and Remote Sensing Symposium (IGARSS)/35th Canadian Symposium on Remote Sensing, Quebec City, QC, Canada, 13–18 July 2014; pp. 2118–2121.
9. Liu, J.; Wang, L.M.; Yao, B.M.; Yang, F.G.; Yang, L.B.; Dong, Q.H. Comparative Study on Crop Recognition of Landsat-OLI and RapidEye Data. In Proceedings of the 6th International Conference on Agro-Geoinformatics, Fairfax, VA, USA, 7–10 August 2017; pp. 178–183.
10. An, R.; Li, W.; Wang, H.L.; Ruan, R.Z. Crop classification using per-field method based on ETM plus image and MODIS EVI time series analysis. In Proceedings of the 5th International Symposium on Integrated Water Resources Management/3rd International Symposium on Methodology in Hydrology, Hohai University, Nanjing, China, 19–21 November 2010; p. 674.
11. Zhang, M.; Li, Q.Z.; Wu, B.F. Investigating the capability of multi-temporal Landsat images for crop identification in high farmland fragmentation regions. In Proceedings of the 1st International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Shanghai, China, 2–4 August 2012; pp. 26–29.
12. Pluto-Kossakowska, J. Review on Multitemporal Classification Methods of Satellite Images for Crop and Arable Land Recognition. *Agriculture* **2021**, *11*, 999. [CrossRef]
13. Ramadhani, F.; Koswara, M.R.S.; Apriyana, Y.; Harmanto. The comparison of numerous machine learning algorithms performance in classifying rice growth stages based on Sentinel-2 to enhance crop monitoring in national level. In Proceedings of the 1st International Conference on Sustainable Tropical Land Management (ICSTLM), Electr Network, Bogor, Indonesia, 16–18 September 2020.
14. Baidar, T.; Fernandez-Beltran, R.; Pla, F. Sentinel-2 multi-temporal data for rice crop classification in nepal. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Electr Network, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 4259–4262.
15. She, B.; Yang, Y.Y.; Zhao, Z.G.; Huang, L.S.; Liang, D.; Zhang, D.Y. Identification and mapping of soybean and maize crops based on Sentinel-2 data. *Int. J. Agric. Biol. Eng.* **2020**, *13*, 171–182. [CrossRef]
16. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.R.; Murayama, Y.; Ranagalage, M. Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sens.* **2020**, *12*, 2291. [CrossRef]
17. Liu, W.; Wang, J.; Luo, J.; Wu, Z.; Chen, J.; Zhou, Y.; Sun, Y.; Shen, Z.; Xu, N.; Yang, Y. Farmland Parcel Mapping in Mountain Areas Using Time-Series SAR Data and VHR Optical Images. *Remote Sens.* **2020**, *12*, 3733. [CrossRef]
18. Ok, A.O.; Akar, O.; Gungor, O. Evaluation of random forest method for agricultural crop classification. *Eur. J. Remote Sens.* **2012**, *45*, 421–432. [CrossRef]
19. Gasparovic, M.; Jogun, T. The effect of fusing Sentinel-2 bands on land-cover classification. *Int. J. Remote Sens.* **2018**, *39*, 822–841. [CrossRef]
20. Huang, S.D.; Xu, W.H.; Xiong, Y.; Wu, C.; Dai, F.; Xu, H.F.; Wang, L.G.; Kou, W.L. Combining Textures and Spatial Features to Extract Tea Plantations Based on Object-Oriented Method by Using Multispectral Image. *Spectrosc. Spectr. Anal.* **2021**, *41*, 2565–2571. [CrossRef]
21. Deng, J.S.; Shi, Y.Y.; Chen, L.S.; Wang, K.; Zhu, J.X. Cotton Identification and Extraction Using Near Infrared Sensor and Object-Oriented Spectral Segmentation Technique. *Spectrosc. Spectr. Anal.* **2009**, *29*, 1754–1758.
22. Cao, X.; Li, Q.Z.; Du, X.; Zhang, M.; Zheng, X.Q. Exploring effect of segmentation scale on orient-based crop identification using HJ CCD data in Northeast China. In Proceedings of the 35th International Symposium on Remote Sensing of Environment (ISRSE35), Beijing, China, 22–26 April 2013; Institute Remote Sensing & Digital Earth: Beijing, China, 2013.
23. Jiao, X.F.; Kovacs, J.M.; Shang, J.L.; McNairn, H.; Walters, D.; Ma, B.L.; Geng, X.Y. Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 38–46. [CrossRef]
24. Hong, R.; Park, J.; Jang, S.; Shin, H.; Kim, H.; Song, I. Development of a Parcel-Level Land Boundary Extraction Algorithm for Aerial Imagery of Regularly Arranged Agricultural Areas. *Remote Sens.* **2021**, *13*, 1167. [CrossRef]

25. Dai, W.; Na, J.; Huang, N.; Hu, G.; Yang, X.; Tang, G.; Xiong, L.; Li, F. Integrated edge detection and terrain analysis for agricultural terrace delineation from remote sensing images. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 484–503. [CrossRef]
26. Jintian, C.; Xin, Z.; Weisheng, W.; Lei, W. Integration of optical and SAR remote sensing images for crop-type mapping based on a novel object-oriented feature selection method. *Int. J. Agric. Biol. Eng.* **2020**, *13*, 178–190. [CrossRef]
27. Haoyu, W.; Zhanfeng, S.; Zihan, Z.; Zeyu, X.; Shuo, L.; Shuhui, J.; Yating, L. Improvement of Region-Merging Image Segmentation Accuracy Using Multiple Merging Criteria. *Remote Sens.* **2021**, *13*, 2782. [CrossRef]
28. Hossain, M.; Chen, D. Segmentation for Object-based Image Analysis (Obia): A Review of Algorithms and Challenges From Remote Sensing Perspective. *J. Math.* **2019**, *150*, 115–134. [CrossRef]
29. Wang, S.; Chen, Y.L. The information extraction of Gannan citrus orchard based on the GF-1 remote sensing image. *IOP Conf. Ser. Earth Environ. Sci.* **2017**, *57*, 012001. [CrossRef]
30. Richter, G.M.; Agostini, F.; Barker, A.; Costomiris, D.; Qi, A. Assessing on-farm productivity of Miscanthus crops by combining soil mapping, yield modelling and remote sensing. *Biomass Bioenergy* **2016**, *85*, 252–261. [CrossRef]
31. Xie, S.; Tu, Z. Holistically-Nested Edge Detection. *Int. J. Comput. Vis.* **2017**, *125*, 1395–1403. [CrossRef]
32. Cheng, G.; Zhou, P.; Han, J. Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 7405–7415. [CrossRef]
33. Ding, P.; Zhang, Y.; Deng, W.-J.; Jia, P.; Kuijper, A. A light and faster regional convolutional neural network for object detection in optical remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2018**, *141*, 208–218. [CrossRef]
34. Rabbi, J.; Ray, N.; Schubert, M.; Chowdhury, S.; Chao, D. Small-Object Detection in Remote Sensing Images with End-to-End Edge-Enhanced GAN and Object Detector Network. *Remote Sens.* **2020**, *12*, 1432. [CrossRef]
35. Xia, G.S.; Bai, X.; Ding, J.; Zhu, Z.; Belongie, S.; Luo, J.B.; Datcu, M.; Pelillo, M.; Zhang, L.P. A Large-scale Dataset for Object Detection in Aerial Images. In Proceedings of the 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 3974–3983.
36. Akarsh, A.; Manoj, K. Image surface texture analysis and classification using deep learning. *Multimed. Tools Appl.* **2020**, *80*, 1289–1309. [CrossRef]
37. Castelluccio, M.; Poggi, G.; Sansone, C.; Verdoliva, L. Land Use Classification in Remote Sensing Images by Convolutional Neural Networks. *arXiv* **2015**, arXiv:1508.00092.
38. Emmanuel, M.; Yuliya, T.; Guillaume, C.; Pierre, A. Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 645–657. [CrossRef]
39. Paoletti, M.E.; Haut, J.M.; Plaza, J.; Plaza, A. A new deep convolutional neural network for fast hyperspectral image classification. *ISPRS J. Photogramm. Remote Sens.* **2017**, *145*, 120–147. [CrossRef]
40. Tao, H.; Li, W.; Qin, X.; Wang, P.; Yu, W.; Li, J. Terrain Classification of Polarimetric Synthetic Aperture Radar Images Based on Deep Learning and Conditional Random Field Model. *J. Radars* **2019**, *8*, 471–478. [CrossRef]
41. Fan, Y.; Ding, X.; Wu, J.; Ge, J.; Li, Y. High spatial-resolution classification of urban surfaces using a deep learning method. *Build. Environ.* **2021**, *200*, 107949. [CrossRef]
42. Cheng, G.; Han, J.W.; Lu, X.Q. Remote Sensing Image Scene Classification: Benchmark and State of the Art. *Proc. IEEE* **2017**, *105*, 1865–1883. [CrossRef]
43. Ma, L.; Liu, Y.; Zhang, X.L.; Ye, Y.X.; Yin, G.F.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [CrossRef]
44. Luo, L.H.; Li, F.Y.; Dai, Z.Y.; Yang, X.; Liu, W.; Fang, X. Terrace extraction based on remote sensing images and digital elevation model in the loess plateau, China. *Earth Sci. Inform.* **2020**, *13*, 433–446. [CrossRef]
45. Zhao, F.; Xiong, L.Y.; Wang, C.; Wang, H.R.; Wei, H.; Tang, G.A. Terraces mapping by using deep learning approach from remote sensing images and digital elevation models. *Trans. GIS* **2021**, *25*, 2438–2454. [CrossRef]
46. Blaes, X.; Vanhalle, L.; Defourny, P. Efficiency of crop identification based on optical and SAR image time series. *Remote Sens. Environ.* **2005**, *96*, 352–365. [CrossRef]
47. Wu, B.; Li, Q. Crop planting and type proportion method for crop acreage estimation of complex agricultural landscapes. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *16*, 101–112. [CrossRef]
48. Haest, B.; Borre, J.V.; Spanhove, T.; Thoonen, G.; Delalieux, S.; Kooistra, L.; Mücher, C.A.; Paelinckx, D.; Scheunders, P.; Kempeneers, P. Habitat Mapping and Quality Assessment of NATURA 2000 Heathland Using Airborne Imaging Spectroscopy. *Remote Sens.* **2017**, *9*, 266. [CrossRef]
49. Kristin, F.; Hannes, F.; Michael, F.; Marion, S.; Björn, W. Hierarchical classification with subsequent aggregation of heathland habitats using an intra-annual RapidEye time-series. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *87*, 102036. [CrossRef]
50. Sun, Y.; Luo, J.; Xia, L.; Wu, T.; Gao, L.; Dong, W.; Hu, X.; Hai, Y. Geo-parcel-based crop classification in very-high-resolution images via hierarchical perception. *Int. J. Remote Sens.* **2020**, *41*, 1603–1624. [CrossRef]
51. Wang, X.; Ma, H.; Chen, X.; You, S. Edge Preserving and Multi-Scale Contextual Neural Network for Salient Object Detection. *IEEE Trans. Image Processing* **2018**, *27*, 121–134. [CrossRef]
52. Liu, Y.; Cheng, M.-M.; Hu, X.; Wang, K.; Bai, X. Richer Convolutional Features for Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *41*, 1939–1946. [CrossRef] [PubMed]

53. Tun, N.L.; Gavrilov, A.; Tun, N.M.; Trieu, D.; Aung, H. Remote Sensing Data Classification Using A Hybrid Pre-Trained VGG16 CNN-SVM Classifier. In Proceedings of the IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg Electrotechn University, Saint Petersburg, Russia, 26–28 January 2021; pp. 2171–2175.
54. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
55. He, K.M.; Zhang, X.Y.; Ren, S.Q.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 27–30 June 2016; pp. 770–778.
56. Romera, E.; Alvarez, J.M.; Bergasa, L.M.; Arroyo, R. ERFNet: Efficient Residual Factorized ConvNet for Real-Time Semantic Segmentation. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 263–272. [CrossRef]
57. Li, G.; Yun, I.; Kim, J.; Kim, J. DABNet: Depth-wise Asymmetric Bottleneck for Real-time Semantic Segmentation. *arXiv* **2019**, arXiv:1907.11357.
58. Ren, T.; Liu, Z.; Zhang, L.; Liu, D.; Xi, X.; Kang, Y.; Zhao, Y.; Zhang, C.; Li, S.; Zhang, X. Early Identification of Seed Maize and Common Maize Production Fields Using Sentinel-2 Images. *Remote Sens.* **2020**, *12*, 2140. [CrossRef]
59. Varlamova, E.V.; Solovyev, V.S. Investigation of Eastern Siberia vegetation index variations on long-term satellite data. *Atmos. Ocean Opt.* **2018**, *10833*, 108338C. [CrossRef]
60. Richetti, J.; Judge, J.; Boote, K.J.; Johann, J.A.; Uribe-Opazo, M.A.; Becker, W.R.; Paludo, A.; Silva, L.C.D. Using phenology-based enhanced vegetation index and machine learning for soybean yield estimation in Parana State, Brazil. *J. Appl. Remote Sens.* **2018**, *12*, 026029. [CrossRef]
61. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
62. Li, T.T.; Wang, Y.F.; Liu, C.Q.; Tu, S.S. Research on Identification of Multiple Cropping Index of Farmland and Regional Optimization Scheme in China Based on NDVI Data. *Land* **2021**, *10*, 861. [CrossRef]
63. Ndikumana, E.; Minh, D.H.T.; Baghdadi, N.; Courault, D.; Hossard, L. Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sens.* **2018**, *10*, 1217. [CrossRef]
64. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
65. Arun, P.; Karnieli, A. Deep Learning-Based Phenological Event Modeling for Classification of Crops. *Remote Sens.* **2021**, *13*, 2477. [CrossRef]
66. Russwurm, M.; Korner, M. Temporal Vegetation Modelling using Long Short-Term Memory Networks for Crop Identification from Medium-Resolution Multi-Spectral Satellite Images. In Proceedings of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 1496–1504.
67. Pan, Y.Z.; Hu, T.G.; Zhu, X.F.; Zhang, J.S.; Wang, X.D. Mapping Cropland Distributions Using a Hard and Soft Classification Model. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4301–4312. [CrossRef]
68. Shuai, G.Y.; Zhang, J.S.; Basso, B.; Pan, Y.Z.; Zhu, X.F.; Zhu, S.; Liu, H.L. Multi-temporal RADARSAT-2 polarimetric SAR for maize mapping supported by segmentations from high-resolution optical image. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *74*, 1–15. [CrossRef]