

Agricultural Cultivation Parcels Extraction and Crop Mapping Using Multisource Remote Sensing Data and Intelligent Algorithms

Manjia Li^{a,b}, Jing Zhang^{a,b}, Tianjun Wu^c, Jiancheng Luo^{a,b,*}, Lijing Gao^a

^a State Key Laboratory of Remote Sensing and Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences (CAS), Beijing 100101, China, Manjia Li – limanjia21@mails.ucas.ac.cn, Jing Zhang – zhangjing221@mails.ucas.ac.cn, Jiancheng Luo – luojc@aricas.ac.cn, Lijing Gao – gaolj200869@aircas.ac.cn

^b University of Chinese Academy of Sciences, Beijing 100049, China

^c School of Land Engineering, Chang'an University, Xi'an 710064, China, Tianjun Wu – tjwu@chd.edu.cn

* Corresponding author

Abstract: The precise mapping of crop spatial distribution using remote sensing datasets is a fundamental task in precision agriculture, which has experienced profound development triggered by the continuous improvement of earth observation systems jointly with the innovation of machine learning theories. However, the extraction and classification of cultivated areas are typically accomplished simultaneously with single-task models at pixel scale, and the prior geographical knowledge was generally ignored, which may present accuracy limitation and significant separation from the monitor application. In response, the proposed work is a novel attempt to address successively the parcel extraction and parcel-wise crop classification over mountainous regions with heterogeneous and fragmented smallholder agriculture. Land parcels get precisely delineated utilizing an improved Densely Connected Link Network (D-LinkNet) with geo-knowledge as prior constraints. Each parcel is then correlated with its closest neighbors considering the environmental and temporal similarity, and classified subsequently by a proposed attention-based Network. Results show that ideal precision was attained in both stages. The incorporation of prior knowledge and neighborhood information has effectively enhanced the accuracy of parcel extraction and crop classification, respectively. Overall, the parcel-wise crop mapping framework may constrain the analysis range within the agricultural space and provides identification results corresponding to real geographic objects, contributing directly to the downstream applications such as crop monitoring, management decision-making, etc.

Keywords: Agricultural cultivation parcels, complex mountainous areas, crop classification, deep learning

1. Introduction

As the world's population continues to grow alongside limited arable land resources (FAO, 2023), timely and accurate monitoring of crop cultivation conditions appears increasingly important for comprehensive evaluation and decision-making in agricultural production management (Segarra et al., 2020). Precision monitoring of crop distribution encompasses two critical dimensions: precise spatial morphology, which means the precise delineation of the agricultural cultivation parcels (ACPs); and accurate type attributes, which means the unequivocal identification of the crop type corresponding to each unit (Beerli and Peled, 2009). Conventional methodologies for crop distribution surveys typically relies on annual farmer-reported data and the official statistical consolidation, which may suffer from considerable ambiguity and uncertainty, limiting access to detailed parcel-based information. In contrast, remote sensing has emerged as a superior alternative, offering the low-cost and wide-area coverage for precise parcel delineation and crop classification.

For the precise extraction of ACPs, numerous studies have explored the automatic derivation using various

remote sensing platforms, from medium resolution (Landsat, Sentinel-1/2) to very high resolution (Gao Fen, Planet) and unmanned aerial vehicles (Li et al., 2023; Waldner and Diakogiannis, 2020). Historically, the delineation of ACPs has predominantly relied on object-based image analysis (OBIA) methods including edge detection and region segmentation. These methods are generally efficient in open, flat areas with clear boundaries while vulnerable to noise and weak edge recognition issues for fragmented regions. Recently, with the progress in image processing, deep learning, especially convolutional neural networks (CNNs), have proven to be powerful tools for ACPs extraction. Various network architectures have been implemented for the parcel extraction, including HED, RCF and U-Net, DeepLab V3+, D-Linknet. However, despite significant advances in model architectures, the complex characteristics of mountainous agriculture pose unique challenges that cannot be adequately addressed by model refinement alone. The obstacles appear particularly evident in regions characterized by heterogeneous, fragmented smallholder farming systems with highly regionalized ACP morphology, diversified parcel types, obscure distinction between ACPs and wooded grasslands. Recent studies

have demonstrated that leveraging prior geospatial-temporal knowledge, which generally involves spatial constraints and stratification strategies, can significantly enhance the robustness and accuracy of remote sensing applications (Zhu et al., 2024b, 2024a), while there remains a critical gap in effectively integrating comprehensive geographical knowledge into ACPs extraction methodologies. In response, we designed a comprehensive framework that combines the prior knowledge with an improved Densely Connected Link Network (D-LinkNet) for ACPs extraction from the VHR images.

For the accurate crop classification, the phenological patterns get efficiently recovered based on the sequential remote sensing images and the target crops are then convincingly recognized using various series analysis methods (Niazmardi et al., 2018). Overall, the recognition unit serves as the bridge between real surface and quantitative analysis methods in the digital world, representing the elementary knowledge of the study area. Image pixels are adopted in most frameworks as the basic unit for classification (Pixel Based Image Analysis, PBIA) (Bargiel, 2017; Xu et al., 2024), which is computationally efficient but susceptible to the influence of mixed pixels and edge effects, and may fail to fully capture the spatial structural information of crops (Chen et al., 2016; Mondal and Jeganathan, 2022). Object-Based Image Analysis (OBIA) was then proposed and widely conducted, where each meaningful or homogeneous group of pixels is treated as an object with certain characteristics and properties (Cheng and Zhang, 2022; Niazmardi et al., 2018). OBIA allows for semantic analysis of the image and the relationships between objects are potentially considered, which proves to be more accurate than the PBIA approach under ideal segmentation conditions. However, most of the researches were implemented in plain areas where the land parcels exhibit typically regular geometric shapes and appear as tightly arranged quadrilateral-shaped features on high-resolution remote sensing images, and parcels are typically derived utilizing segmentation methods, leading to significant discrepancy with the actual surface features. Besides, each entity is typically regarded as an independent computing unit, ignoring its neighborhood information that may aid in feature completion and noise resistance. Consequently, a parcel-wise classification method utilizing neighborhood information was adopted to achieve better accuracy control.

Overall, a novel parcel-wise framework was proposed in this work for the accurate monitoring of crop cultivation conditions. Land parcels get precisely delineated utilizing the improved D-LinkNet with geo-knowledge as prior constraints. Multi-source datasets are then integrated within each parcel's spatial boundary to fully depict temporal cover changes and environmental conditions. Consequently, parcel relationships are established based on environmental and sequential similarities, and crop types are identified using parcel-wise time series features

and neighborhood information.

2. Study area and materials

2.1 Study area

The experiments for parcel extraction and classification were conducted in Jiangjin and Tongnan districts of Chongqing, China, as shown in Figure 1., both districts are located in the eastern part of the Sichuan Basin, downstream of the Fu River. The two districts keep a temperate subtropical humid monsoon climate with an annual average temperature of 17.9°C, 19.5°C and a total rainfall of 969.2, 1001.2 millimeters, respectively, indicating an ideal hydrothermal condition for crop cultivation. Generally, the agricultural system in mountainous areas of Southwest China is characterized by complex terrain, diverse cultivation patterns, and intricate meteorological conditions with frequent cloud cover and rainfall, which is fully reflected in both areas.

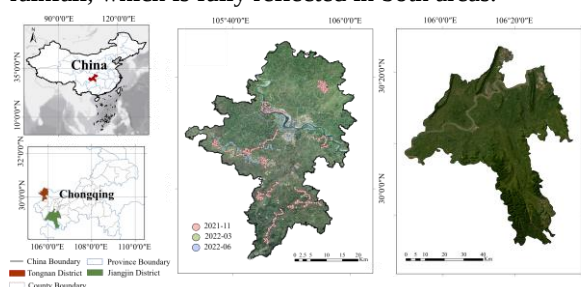


Figure 1. The location of the study areas. Distinct colored points denote the spatial locations of ground sampling across various batches.

In this study, the controlled experiments of the parcel extraction and crop classification were implemented in Jiangjin and Tongnan, respectively.

2.2 Data collection and pre-processing

1) Ground truth data:

The ground truth data were obtained through field surveys conducted separately in November 2021, March 2022, and June 2022. 2,033 sample points were annotated, as shown in Figure 1. All sample points were matched with corresponding land parcels and cross-validated based on high-resolution images. Consequently, a total of 1,502 parcels were labeled for model training and validation, including rice, corn, lemon and other types, and provide supervision signal in the classification process.

2) Remote sensing data:

Both the high- and mid-resolution images were utilized in this study for the parcel extraction and crop classification correspondingly.

a) High-resolution images from GF-2 satellite. Images with a spatial resolution of 0.8 meters were utilized as the data source to extract arable land parcels. Cloud gaps were addressed by mosaicking multi-period GF data. These images function as the data source to extract arable land parcels.

b) Mid-resolution series images from Sentinel satellites. Both radar and optical datasets were incorporated in this study.

For the former, Synthetic Aperture Radar (SAR) imagery obtained from the C-band sensor SAR on the Sentinel-1 (S1) satellite of the European Space Agency (ESA) was adopted as the data source in the assessment of sequential similarities. The essential preprocess was conducted with SNAP software to retrieve the backscattering coefficient (i.e., VV and VH) consequently.

For the latter optical images, Sentinel-2 (S2) images derived from the "COPERNICUS/S2_SR" product suite provided by the GEE platform were selected. The official Sentinel-2 cloud mask dataset ('COPERNICUS/S2_CLOUD_PROBABILITY') was also obtained and merged with the original images for the detection of potentially cloudy pixels. NDVI series reflect the land cover changes within the parcels, and are directly utilized as input features for classification.

3) Environmental factors:

Factors related to crop phenology were also collected for a comprehensive description of environmental conditions, including the historical climate data from the WorldClim (Fick and Hijmans, 2017), current climate data from the National Tibetan Plateau Environment Data Center (TPDC) (Ding and Peng, 2021, 2020; Peng et al., 2019), the Digital Elevation Model (DEM) data sourced from the ALOS PALSAR sensor, the Hydrographic data derived from the OpenStreetMap, and the soil subtype data acquired from the Harmonized World Soil Database (*Harmonized World Soil Database version 2.0*, 2023). These datasets are employed for environmental consistency assessment, thereby facilitating natural zoning and parcel-wise correlation analysis.

3. Methods

In this study, the parcel extraction and crop classification are sequentially achieved leveraging the aforementioned multi-source datasets. Land parcels are delineated utilizing the improved D-LinkNet following a stratification strategy within different geographical zones. Parcels are then perceived as nodes in a correlated graph and classified by a proposed attention-based neural network. Here follows the detailed description of each step.

3.1 Parcel extraction with prior information and deep learning

Figure 2 illustrates the workflow of the parcel extraction. First, geographical zoning was utilized to establish a confined area for sample production and to enable parallel computing or comparative analysis, while historical data was employed to define the maximum boundary to mitigate the confusion between the target and the background in the deep learning predictions. Next, a stratification strategy was developed based on significant or unique boundaries and texture features of anthropogenic agricultural traces and structures, with ACPs extracted using the improved D-LinkNet (Xia et al., 2021). The predicted rasters were binarized to raster edges, then converted to vector lines of SHP format, and further transformed to vector polygons. Manual correction of obvious topological errors and merging of over-segmented

patches were then implemented, yielding the vectorized representation of land parcels, which serve as the fundamental spatial units for subsequent feature construction and type inference.

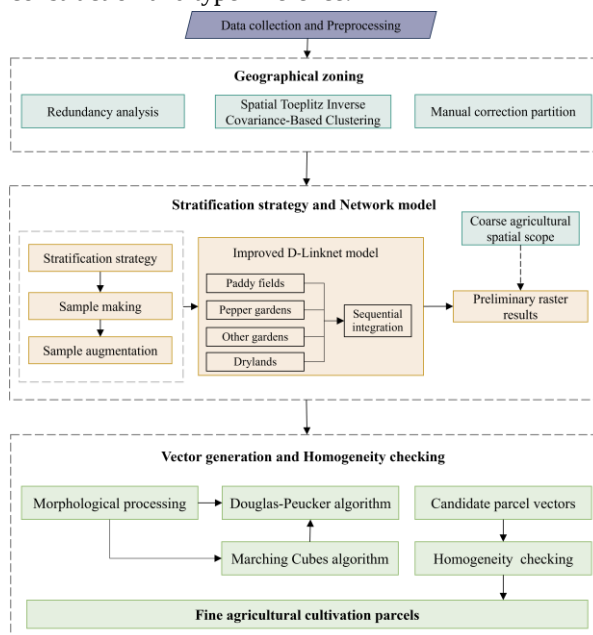


Figure 2. Workflow of the parcel extraction method.

3.1.1 Geographical zoning and Coarse spatial scope

1) Geographical zoning

Geographic data patterns follow principles that highlight correlations among proximate objects and homogeneity in similar backgrounds. Zoning refers to dividing a complex area into multiple units characterized by relatively internal homogeneity and external heterogeneity, which was implemented in this study to harmonize natural and anthropogenic influences within sub-regions for uniform parcel characteristics, facilitating sample collection and enhancing regional comprehension.

Vegetation cover status represented by the mean values of NDVI in March, August, the growing season (April–October), and the important phenological note (maturity) served as the response variables, and environmental indicators (i.e., the mean values of temperature in March, arid index, slope, altitude, and soil, along with the vegetation cover indicators) were selected as explanatory variables based on the correlation analysis. To avoid overfitting from environmental data multicollinearity, redundancy analysis (RDA) (Rao, 1964) was used to explore relationships between response and explanatory variables, with 200 points randomly selected for the analysis process using the "vegan" R package. Subsequently, the Spatial Toeplitz Inverse Covariance-Based Clustering (STICC) (Kang et al., 2022) method was employed to obtain the partitioning results based on the inferred vegetation cover status.

2) Coarse spatial scope

Geographical objects undergo scale-dependent changes, necessitating a top-down approach to image understanding that mimics human perception, which typically starts with macro perspectives and refines to specific objects of

interest. Given the potential for confusion between similar features, such as farmland boundaries and roads, it is necessary to filter out irrelevant non-agricultural areas. Considering the stability of agricultural (planting) spaces under policy-driven management, the coarse agricultural spatial scope was derived from historical land use data in this study.

3.1.2 Stratification strategy and Network model

Land reclamation and cultivation processes have created distinctive features across different agricultural land use types (i.e., paddy fields, drylands, pepper gardens, and other gardens), which are critical for translating ACPs mappings from imagery into geospatial data. Consequently, a hierarchical strategy was applied for the sequential extraction of distinct types from simplicity to complexity (i.e., paddy fields, pepper gardens, other gardens and drylands), allowing each single-task deep model to focus on edge or texture features contrapuntally for different ACPs.

The D-LinkNet (Zhou et al., 2018) was selected due to its effectiveness in handling complex structures and multi-scale features. To further enhance feature representation, a coordinate attention (CA) (Hou et al., 2021) module was integrated into the center part of the original structure, and the single mapping between the encoder and decoder parts was substituted by richer skip connections (Xia et al., 2021).

3.1.3 Stratification strategy and Network model

Based on the model's raster outputs, the closed, discrete polygons are created through morphological optimization (i.e., edge enhancement and hole elimination) and contour acquisition. Ultimately, the Marching Cubes and Douglas-Peucker algorithms (Li et al., 2023) were utilized to generate parcel vectors.

After the vector generation, under-segmentation was recognized by the certain uniformity within the correct parcel's interior. Eq. (1) was applied to assess the internal variation of each candidate parcel, utilizing the three bands of the VHR imagery as attribute features. Parcels with variability surpassing a predetermined threshold will undergo human-computer interaction and iterative optimization. Prior to these calculations, de-edge processing was performed to exclude any potential abnormal fluctuations in the edge pixels if allowed.

$$\text{Variation} = \frac{\sum (x - \bar{x})^2}{n} \quad (1)$$

where $\bar{x}$ and n refers to the mean and number of pixels in a parcel.

3.2 Parcel-wise crop classification

As shown in Figure 3., The proposed framework for crop classification consists of two interrelated steps (i.e. feature construction and correlation establishment and crop identification), and here follows the detailed description of each step.

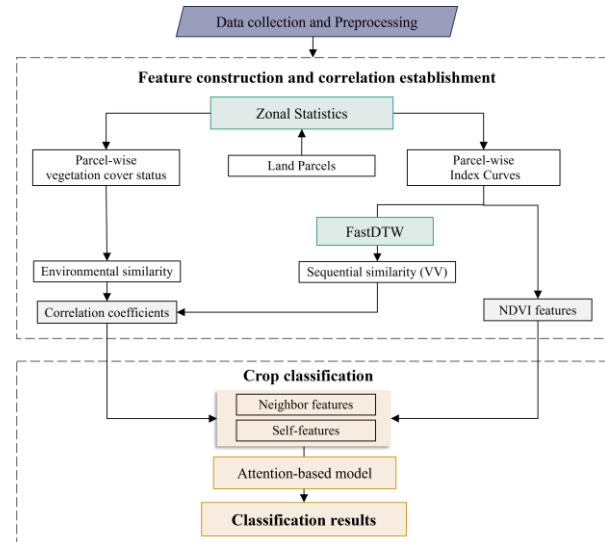


Figure 3. Flowchart for parcel-wise crop identification.

3.2.1 parcel-wise feature construction and correlation establishment

The zonal statistical analysis was implemented based on the extracted ACPs. For each parcel, the pixels within its spatial boundary were gathered and statistical indicators (the median in this study) were calculated afterward to generate parcel-wise features. Both NDVI and VH features were constructed for the subsequent parcel-wise correlation construction and type inference. Additionally, the vegetation cover status calculated in 3.1 was also mapped to each parcel as a comprehensive indicator of environment characteristics.

Subsequently, the correlation between parcels got judiciously constructed considering both environmental and temporal similarities. The former was reflected directly by the vegetation cover status, and the temporal similarity was evaluated based on the VH values utilizing the Dynamic Time Warping (DTW) algorithm. The similarity was defined as the inverse of the calculated DTW distance as follows.

$$S_{seq} = \frac{1}{1 + DTW(s_t, s_i)} \quad (2)$$

where s_t and s_i denote the VH series for the target parcel and its candidate neighbors respectively. Before computation, the original feature values are standardized to the 0-1 interval based on the minimum-maximum normalization.

Both similarities take effect in sequence during the correlation construction, reducing gradually the searching space of the candidate neighbor set. k_{env} parcels with the closest vegetation cover status were picked out, where k_{env} is a quantity threshold set manually. Consequently, k_{seq} parcels with the most similar VH series were then selected as the final neighborhood units, and the values were adopted as the correlation strength.

3.2.2 Crop classification using attention-based model

Based on the NLP transformers (Vaswani et al., 2017), this study uses Temporal Attention (TA) and Neighborhood Attention (NA) modules to explore

dependencies from different angles. In NA, the target and neighbor parcels are treated as tokens with NDVI series as original features; in TA, time-point features are tokens and the full sequence is the context. A class token is added at the start, and all tokens are projected to a high-dimension tensor before multi-attention processing. The modules are cascaded to ensure reliable pattern recognition.

NA and TA structures are defined by key parameters—number of layers and attention heads—with optimal values varying by input features. For this study's 29-time-point NDVI features and up to five neighbors, NA and TA layers are set to 2, attention heads to 3 and 2 respectively, with embedding dimensions at 64 and hidden dimensions at 128.

4. Results

4.1 Parcel extraction

Four metrics were selected including overall accuracy (OA), Kappa coefficient, F-1 score, and the mask intersection over union (IoU) (see Eq. (3) to Eq. (6)). Besides, $IoU_{Boundary}$ metric, which comprises overlap and smoothness, was complemented for the individual assessment (see Eq. (7) to Eq. (9)).

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Kappa = \frac{P_o - P_e}{1 - P_e} \quad (4)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

$$IoU = \frac{Intersection}{Union} \quad (6)$$

$$IoU_{Boundary} = Overlap \times Smoothness \quad (7)$$

$$Overlap = \frac{Length_{predicted} \cap Buffer_{ground}}{Length_{predicted}} \quad (8)$$

$$Smoothness = \begin{cases} \frac{Breakpoints_{predicted}}{Breakpoints_{ground}}, & Breakpoints_{predicted} > Breakpoints_{ground} \\ \frac{Breakpoints_{ground}}{Breakpoints_{predicted}}, & Breakpoints_{predicted} < Breakpoints_{ground} \end{cases} \quad (9)$$

where TP , TN , FP , FN refers to the true positive, true negative, false positive, and false negative, respectively; P_o refers to the overall percent agreement, and P_e refers to the hypothetical probability of change agreement; $Precision$ and $Recall$ refers to $\frac{TP}{TP+FP}$ and $\frac{TP}{TP+FN}$, respectively; $Buffer_{ground}$ refers to an inward and outward zone around the parcel boundary with a specific radius.

Fourteen validation samples, each 1000×1000 pixels, were manually annotated for the former four index, and

600 parcels were selected with a buffer zone distance of 0.8 meters (one pixel) for $IoU_{Boundary}$ calculation.

As illustrated in Table 1, the method utilized in this study outperforms the regular method based solely on the improved D-LinkNet, exhibiting enhancements of approximately 0.14 in the Kappa coefficient and 0.11 in IoU. Spatial heterogeneity in extraction performance is observed across the study area, with region 1 and region 2 exhibiting notably higher accuracy. The extraction effectiveness for different parcel types generally aligns with the stratification order, suggesting that more effective separation of targets from backgrounds leads to improved extraction outcomes. However, there are discrepancies between the IoU_{total} (total IoU for 14 validation samples) and the $IoU_{average}$ (average IoU for 14 validation samples), which stems possibly from the coexistence of small, easily overlooked parcels and large, prominent parcels—a characteristic feature particularly prevalent in garden landscapes. This heterogeneity in parcel size distribution may introduce bias in the global evaluation metrics used above, necessitating individual parcel-level precision assessment. The occurrence of low value or even outliers can be attributed to two primary factors: non-optimizable under-segmentation and incomplete region identification (Figure 4). Parcels with higher geometric complexity tend to exhibit lower $IoU_{Boundary}$.

	OA	Kappa	F-1	IoU _(total)	IoU _(average)
Regular	0.8	0.703	0.850	0.740	0.739
Our	0.9	0.847	0.921	0.854	0.841
Region 1	0.8	0.861	0.935	0.877	0.874
Region 2	0.9	0.905	0.957	0.919	0.916
Region 3	0.8	0.770	0.869	0.769	0.767
Region 4	0.8	0.830	0.890	0.802	0.767
Paddy	0.9	0.887	0.913	0.839	0.796
Drylands	0.9	0.765	0.795	0.660	0.617
Pepper	0.9	0.937	0.944	0.895	0.664
Other	0.9	0.793	0.797	0.662	0.585

Table 1. Evaluation results of the extraction.

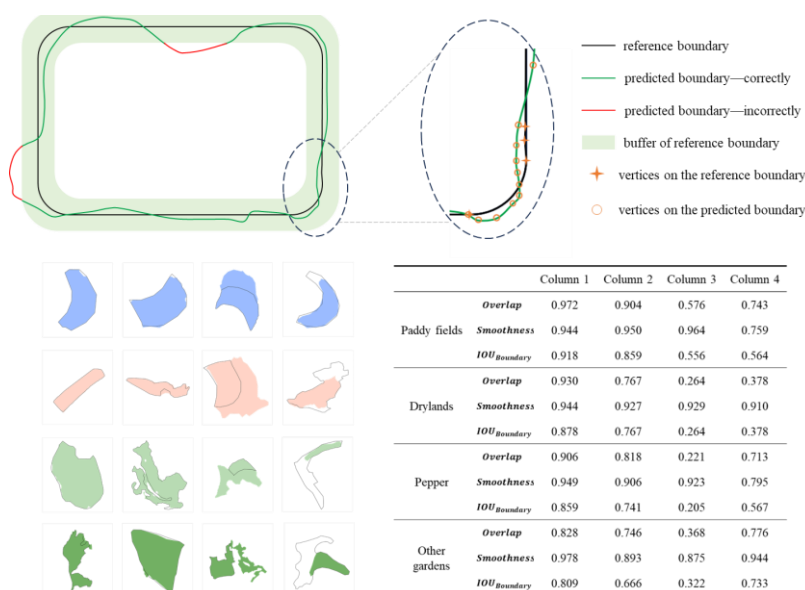


Figure 4. The relationship between ground truth and predicted result, and examples of $IoU_{Boundary}$ calculations on some parcels.

In Jiangjin district, 1.23 million agricultural parcels were identified, comprising 0.50 million paddy fields, 0.45 million drylands, 0.22 million pepper gardens, and 0.06 million other gardens, covering 50.66, 51.88, 45.56, and 11.08 million acres, respectively. The distribution analysis reveals that cultivated land and garden land constitute approximately 77% and 23% respectively of parcel count, a similar ratio reflected in total area. Notably, paddy fields exhibit a lower area-to-count ratio, indicating a predominance of smaller parcels, whereas pepper gardens display the opposite, suggesting larger, contiguous configurations (Figure 5). The areal distribution among paddy fields, drylands, and pepper gardens maintains an approximate 1:1:1 ratio, consistent with the collected statistics. Regarding the regional distribution, the ACPs are predominantly located outside of region 4, while region 2 exhibits the highest concentration of all the parcel types, attributed to its favorable topographical conditions. There are minimal differences in both area and count of paddy fields and other gardens across region 1 to region 3, while region 2 shows a marked predominance in dryland. Pepper gardens are concentrated in region 1 and region 2, with minimal presence in region 3. From these data, we can infer that the distribution of agricultural parcels in the study area exhibits certain regularities and characteristics: paddy fields and drylands are relatively evenly distributed, while pepper gardens and other gardens tend to form large-scale concentrated areas. The occurrence of low value or even outliers can be attributed to two primary factors: non-optimizable under-segmentation and incomplete region identification (Figure 4). Parcels with higher geometric complexity tend to exhibit lower $IoU_{Boundary}$.

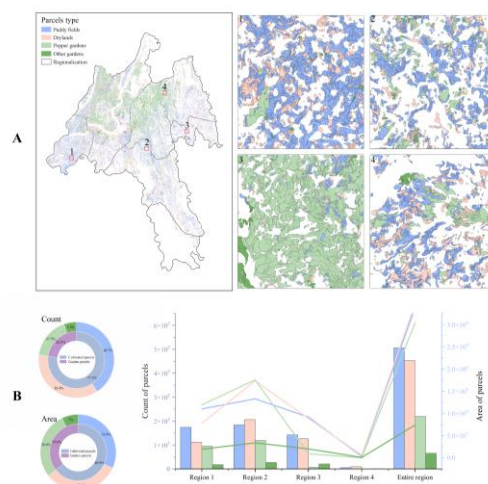


Figure 5. Distribution maps. (A) shows the spatial distribution of parcels and the detail maps, and (B) shows the statistical results of parcels.

4.2 Parcel-based crop classification

4.2.1 Parcel-wise feature reconstruction and correlation establishment

The parcel-level features of NDVI and VH were derived based on zonal statistics, and sample points were matched to the corresponding parcels. The characteristic curves of various crop types are shown in Figure 6.

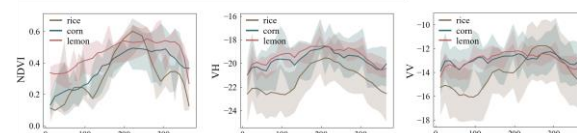


Figure 6. Time-series feature curve of major crops in the study area.

Based on the comprehensive consideration of environmental and time-series similarities, each parcel is matched with its most relevant ones. The threshold for the

correlation recognition is set at 0.7, with the maximum number of neighbors capped at 5.

4.2.2 Crop classification and accuracy assessment

Based on the parcel-wise time-series features and collected neighborhood information, the classification process was implemented using the attention-based models, as shown in Figure 7. The results were evaluated based respectively on the membership degree vectors and crop types utilizing the confusion matrix and information entropy, facilitating a holistic assessment of uncertainty distribution from individual to collective, categorical to spatial perspectives.

Five metrics were selected including OA, Precision (P), Recall (R), $F1$ — *score*, and entropy (see Eq. (3), (5), (10) to (13)) for the accuracy assessment.

$$P = \frac{TP}{TP + FP} \quad (10)$$

$$R = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (12)$$

where TP , TN , FP , and FN respectively denote the number of true positive, true negative, false positive, and false negative instances. Especially, the overall results of P , R , and $F1$ are calculated as the average of different classes weighted by the proportion of samples.

$$H(X_i) = - \sum_{k \in n} y_i^k \log p(y_i^k) \quad (13)$$

where y_i^k signifies the membership degree to category k ($k = 1, 2, \dots, n$) of an individual classification unit.

Especially, compared to the accuracy derived from the confusion matrix over all the parcels, the entropy-based method provides additional spatial information on type uncertainty at the parcel scale, which aids in understanding the spatial distribution of errors.

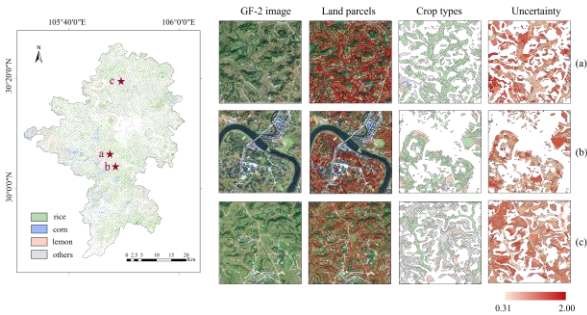


Figure 7. Overall results of parcel-wise crop classification, and detailed identification as well as uncertainty assessment for location A and B.

The OA, P , R , $F1$, and uncertainty on the test set are shown in Table 2, indicating a credible classification performance and generalization ability in areas with diverse environmental conditions.

Crop	OA	P	R	F1	Entropy
rice	-	0.985	0.930	0.956	0.409
corn	-	0.820	0.837	0.828	0.890
lemon	-	0.880	0.956	0.917	0.771
others	-	0.800	0.769	0.784	0.895
overall	0.891	0.893	0.891	0.891	0.858

Table 2. Evaluation results of the crop classification.

A controlled experiment was conducted additionally to clarify the effectiveness of neighborhood information (Table 3). As the ratio of the training set to the test set decreases, the inference capability of the models gradually decreases. The introduction of neighborhood information may lead to a certain degree of accuracy improvement, regardless of the integration approach, and the accuracy improvement becomes increasingly evident as the proportion of the training set decreases.

	Proportion of train-set					
	0.8	0.6	0.4	0.2	0.1	0.05
Non-neighbor	0.853	0.848	0.814	0.770	0.727	0.135
Neighbor	0.891	0.886	0.874	0.852	0.802	0.135

Table 3. Overall accuracy under different neighborhood spaces of different training set size.

5. Conclusion

The accurate mapping of agricultural cultivation parcels and crop types is constantly a fundamental task in precision agriculture, which may face severe obstacles in mountainous areas with significant environment heterogeneity and complex farming patterns. In response, a novel parcel-wise framework was proposed in this work to settle both problems based on multi-source remote sensing datasets.

Land parcels get precisely delineated utilizing a combined model of geographical priori information—including heterogeneous cognition, hierarchical intra-classes and parcel internal homogeneity—and a deep learning algorithm. The improved D-LinkNet model was deployed hierarchically for different land-use types under the constraints of geographical zones and the coarse agricultural spatial scope. Experimental results show that the proposed approach performed superiorly in traditional evaluation indices (e.g., OA, Kappa, $F1$, and IoU) and individual geometric precision of parcels. As a result, approximately 1.23 million parcels in Jiangjin and 0.55 million parcels in Tongnan were extracted, respectively.

The land parcels were then employed as the basic spatial units for feature construction and type inference, which constrains the analysis range within the agricultural space and provides identification results corresponding to real geographic objects. Each parcel is then correlated with its closest neighbors considering the environmental and temporal similarity, and classified subsequently by an attention-based network. Overall and individual metrics indicate a credible classification performance. Besides, the introduction of neighborhood information leads to a certain degree of accuracy improvement, especially as the proportion of the training set decreases, which is often the case for surface parameter estimations.

Reference

Bargiel, D., 2017. A new method for crop classification combining time series of radar images and crop phenology information. *Remote Sensing of Environment* 198, 369–383. <https://doi.org/10.1016/j.rse.2017.06.022>

- Beerli, O., Peled, A., 2009. Geographical model for precise agriculture monitoring with real-time remote sensing. *ISPRS J. Photogramm. Remote Sens.* 64, 47–54. <https://doi.org/10.1016/j.isprsjprs.2008.07.007>
- Chen, Y., Song, X., Wang, S., Huang, J., Mansaray, L.R., 2016. Impacts of spatial heterogeneity on crop area mapping in Canada using MODIS data. *ISPRS J. Photogramm. Remote Sens.* 119, 451–461. <https://doi.org/10.1016/j.isprsjprs.2016.07.007>
- Cheng, J., Zhang, F., 2022. A Novel Crop Classification Method Based on the Tensor-GCN for Time-Series PolSAR Data. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING* 60.
- Ding, Y., Peng, S., 2021. Spatiotemporal change and attribution of potential evapotranspiration over China from 1901 to 2100. *Theor. Appl. Climatol.* 145, 79–94. <https://doi.org/10.1007/s00704-021-03625-w>
- Ding, Y., Peng, S., 2020. Spatiotemporal Trends and Attribution of Drought across China from 1901–2100. *Sustainability* 12, 477. <https://doi.org/10.3390/su12020477>
- Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. <https://doi.org/10.1002/joc.5086>
- Harmonized World Soil Database version 2.0, 2023. . FAO; International Institute for Applied Systems Analysis (IIASA); <https://doi.org/10.4060/cc3823en>
- Hou, Q., Zhou, D., Feng, J., 2021. Coordinate Attention for Efficient Mobile Network Design, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 13708–13717. <https://doi.org/10.1109/CVPR46437.2021.01350>
- Kang, Y., Wu, K., Gao, S., Ng, I., Rao, J., Ye, S., Zhang, F., Fei, T., 2022. STICC: a multivariate spatial clustering method for repeated geographic pattern discovery with consideration of spatial contiguity. *Int. J. Geogr. Inf. Sci.* 36, 1518–1549. <https://doi.org/10.1080/13658816.2022.2053980>
- Li, M., Long, J., Stein, A., Wang, X., 2023. Using a semantic edge-aware multi-task neural network to delineate agricultural parcels from remote sensing images. *ISPRS J. Photogramm. Remote Sens.* 200, 24–40. <https://doi.org/10.1016/j.isprsjprs.2023.04.019>
- Mondal, S., Jeganathan, C., 2022. Effect of scale, landscape heterogeneity and terrain complexity on agriculture mapping accuracy from time-series NDVI in the Western-Himalaya region. *Landscape Ecol.* 37, 2757–2781. <https://doi.org/10.1007/s10980-022-01533-6>
- Niazmardi, S., Homayouni, S., Safari, A., McNairn, H., Shang, J., Beckett, K., 2018. Histogram-based spatiotemporal feature classification of vegetation indices time-series for crop mapping. *Int. J. Appl. Earth Obs. Geoinf.* 72, 34–41. <https://doi.org/10.1016/j.jag.2018.05.014>
- Peng, S., Ding, Y., Liu, W., Li, Z., 2019. 1 km monthly temperature and precipitation dataset for China from 1901 to 2017. *Earth Syst. Sci. Data* 11, 1931–1946. <https://doi.org/10.5194/essd-11-1931-2019>
- Rao, C.R., 1964. The Use and Interpretation of Principal Component Analysis in Applied Research. *Sankhyā: Indian J. Stat.* (1961-2002) 26, 329–358.
- Segarra, J., Buchailot, M.L., Araus, J.L., Kefauver, S.C., 2020. Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications. *Agronomy* 10, 641. <https://doi.org/10.3390/agronomy10050641>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention Is All You Need.
- Waldner, F., Diakogiannis, F.I., 2020. Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network. *Remote Sens. Environ.* 245, 111741. <https://doi.org/10.1016/j.rse.2020.111741>
- Xia, L., Zhang, J., Zhang, X., Yang, H., Xu, M., 2021. Precise Extraction of Buildings from High-Resolution Remote-Sensing Images Based on Semantic Edges and Segmentation. *Remote Sens.* 13, 3083. <https://doi.org/10.3390/rs13163083>
- Xu, Y., Zhou, J., Zhang, Z., 2024. A new Bayesian semi-supervised active learning framework for large-scale crop mapping using Sentinel-2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 209, 17–34. <https://doi.org/10.1016/j.isprsjprs.2024.01.023>
- Zhou, L., Zhang, C., Wu, M., 2018. D-LinkNet: LinkNet with Pretrained Encoder and Dilated Convolution for High Resolution Satellite Imagery Road Extraction, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Presented at the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 192–1924. <https://doi.org/10.1109/CVPRW.2018.00034>
- Zhu, Q., Li, Z., Song, T., Yao, L., Guan, Q., Zhang, L., 2024a. Unrestricted region and scale: Deep self-supervised building mapping framework across different cities from five continents. *ISPRS J. Photogramm. Remote Sens.* 209, 344–367. <https://doi.org/10.1016/j.isprsjprs.2024.01.021>
- Zhu, Q., Ran, L., Zhang, Y., Guan, Q., 2024b. Integrating geographic knowledge into deep learning for spatiotemporal local climate zone mapping derived thermal environment exploration across Chinese climate zones. *ISPRS J. Photogramm. Remote Sens.* 217, 53–75. <https://doi.org/10.1016/j.isprsjprs.2024.08.004>