ISDE · AIR · Taylor & Francis
Taylor & Francis Group

🔓 OPEN ACCESS | Check for updates

A generalized framework for agricultural field delineation from high-resolution satellite imageries

Yu Zhu[a,b], Yaozhong Pan[a,b,c], Tangao Hu[d], Dujuan Zhang[e], Chuanwu Zhao[a,b] and Yuan Gao[a,b]

[a]State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing, People's Republic of China; [b]Key Laboratory of Environmental Change and Natural Disaster of Ministry of Education, Faculty of Geographical Sciences, Beijing Normal University, Beijing, People's Republic of China; [c]Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining, People's Republic of China; [d]Zhejiang Provincial Key Laboratory of Urban Wetlands and Regional Change, Hangzhou Normal University, Hangzhou, People's Republic of China; [e]National Supercomputing Center in Zhengzhou, Zhengzhou University, Zhengzhou, People's Republic of China

ABSTRACT

Accurate digital data on agricultural fields are crucial for various agriculture applications. While deep learning methods have shown promise in delineating fields from high-resolution imagery, there is a lack of research evaluating key techniques for field boundary detection. Challenges also exist in converting detection results into high-quality fields. This study addresses these issues and proposes a generalized framework for agricultural field delineation (GF-AFD). First, we identify three key techniques for field boundary detection and apply them to MPSPNet model. Ablation and comparison experiments demonstrate significant performance enhancements due to techniques. They also prove effective for DeeplabV3+ model, which shares a similar architecture. The modified models outperform U-Net-based models and approach the state-of-the-art models. Second, we show that performing region segmentation on boundary results yields improved field shapes. To address weak boundary loss and unstable parameters issues in existing segmentation-based methods, we introduce the OWT method to enhance weak boundaries directionally before segmentation. We also develop a hierarchical merging method, leveraging the observational hierarchy of fields, resulting in stable parameters across regions and models. The proposed GF-AFD framework was validated cross three diverse Chinese counties. The results demonstrate the framework's robust performance, providing a valuable solution for delineating agricultural fields.

CONTACT Yaozhong Pan ✉ pyz@bnu.edu.cn 🖅 State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, People's Republic of China; Key Laboratory of Environmental Change and Natural Disaster of Ministry of Education, Faculty of Geographical Sciences, Beijing Normal University, Beijing 100875, People's Republic of China; Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining 810016, People's Republic of China; Tangao Hu ✉ hutangao@hznu.edu.cn 🖅 Zhejiang Provincial Key Laboratory of Urban Wetlands and Regional Change, Hangzhou Normal University, Hangzhou 311121, People's Republic of China

1. Introduction

Agricultural field shapes agricultural landscapes, and their accurate distribution information can assist many agricultural applications, such as crop type identification, growth monitoring, and yield prediction, in obtaining field-level data (Blaschke 2010). These field-level statistics are critical for agricultural resources management and policy-making, supporting precision agriculture (Musat et al. 2018). However, the existing high-precision field data acquisition mainly relies on field investigation and manual visual outlining, which are time-consuming and costly (García-Pedrero, Gonzalo-Martín, and Lillo-Saavedra 2017). Acquiring agricultural field data based on high-resolution satellite images has become more feasible in recent years as image resolution and accessibility have increased.

The agricultural field delineation methods can be classified into region-based and edge-based approaches. The region-based approach (RBA) focuses on cropland regions through image segmentation methods such as watershed segmentation, Mean-Shift algorithm (Belgiu and Csillik 2018; García-Pedrero, Gonzalo-Martín, and Lillo-Saavedra 2017; Ming et al. 2016; Watkins and Van Niekerk 2019a, 2019b). Nevertheless, due to the complexity of the imagery, the resulting regions are often over-segmented within the high internal variation field and under-segmented between small adjacent fields (Belgiu and Csillik 2018). The deep learning approaches have also adopted for cropland identification. For instance, D. Zhang et al. (2020) proposed a modified PSPNet model for high resolution cropland areas mapping in four provinces of China. However, these methods can only capture the overall extent of the cropland regions but may not capture the fine internal boundaries due to the lack of detailed information in the high-level category features. Therefore, directly delineating agricultural fields using region-based methods is challenging.

Due to the importance of detailed boundary information for cropland division, an increasing number of studies have used edge-based approaches (EBA) to detect field boundaries in the past decade (Cheng et al. 2020; Turker and Kok 2013; Yan and Roy 2014; 2016). Generally, EBA integrative field boundary delineation consists of two parts: field boundary detection (F-BD), and agricultural field generation (A-FG) (Figure 1). The F-BD method determines the correctness and
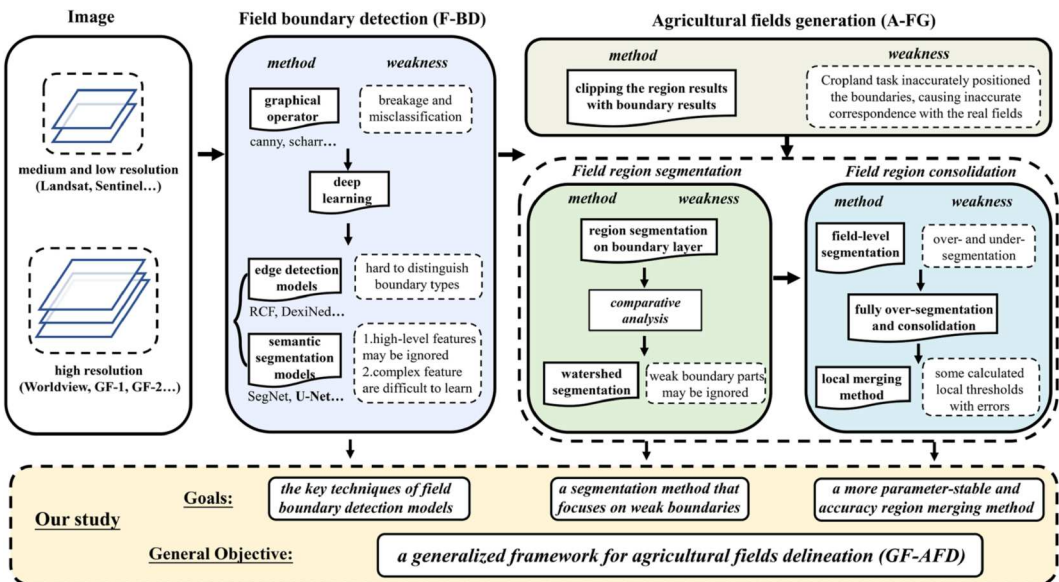
Figure 1. The steps of the edge-based approaches for field boundary delineation and the current state of research, as well as the objective of our study.

completeness of the field boundary, while the A-FG method further transforms the detected pixel-level boundary probabilities into object-level agricultural fields, emphasizing their independence and closure. Notably, the A-FG step has often been overlooked in prior studies, yet object-level results hold greater practical utility.

F-BD can also be divided into graphical operator-based and deep learning-based approaches (S. Liu et al. 2022). The graphical operator-based approach utilizes the edge operators such as Canny, Sobel to detect linear objects. But due to their limited capability, the results usually suffer from noticeable boundary breakage and misclassification. Deep learning-based methods have more powerful feature description capability and can detect accurate results, and thus are widely adopted (Crommelinck et al. 2019; Fetai, Račič, and Lisec 2021; H. Zhang et al. 2021; Masoud, Persello, and Tolpekin 2020). According to the accuracies reported in these studies, deep learning-based methods exhibit superior performance for field boundary detection.

The most suitable deep learning model for field boundary detection still needs further research. The edge detection models like RCF and DexiNed have excellent boundary awareness and can detect more potential field boundaries, yet they struggle to distinguish field boundaries from other types of boundaries due to a lack of category judgment (Xu et al. 2022). Semantic segmentation models, however, can treat field boundaries as a feature type and linking each image pixel to the field boundary/non field boundary label, thus detecting them semantically. FCN (Masoud, Persello, and Tolpekin 2020; Xia, Persello, and Koeva 2019), Segnet (Persello et al. 2019), U-Net (Fetai, Račič, and Lisec 2021; Gopidas et al. 2021), ResU-Net (Taravat et al. 2021), ResUNet-a (Waldner et al. 2021; Waldner and Diakogiannis 2020), and R2U-Net (H. Zhang et al. 2021), are widely adopted architectures. Among them, The U-Net-based architectures are most effective due to their rich low-level features, which are more suitable for recognizing linear objects with only a few pixels width (Taravat et al. 2021). Nevertheless, these U-Net-based models lack sufficient high-level feature for boundary type determination. Recently, some multitasking-based networks such as BsiNet (Long et al. 2022) and SEANet (M. Li et al. 2023) have been used for field boundary detection, and surprisingly, these models have achieved better detection results through the constraints of the additional region task signals. However, the existing models are modified on various benchmark models, and no study has yet conducted an in-depth analysis of the model design based on the characteristics of field boundaries, so as to assess the key techniques in the model that most affect the detection accuracy, and provide a baseline reference for more robust detection model design.

Most existing studies directly transform the detected continuous-valued field boundary probabilities into binarized boundaries with thresholds (Fetai, Račič, and Lisec 2021; Garcia-Pedrero et al. 2019; Marvaniya et al. 2021; Masoud, Persello, and Tolpekin 2020; Xia, Persello, and Koeva 2019). However, the direct binarization may produce many discontinuous boundaries and noisy information. Some studies have applied the gestalt rule (Turker and Kok 2013) and Suzuki85 method (Hong et al. 2021) to remove some broken boundaries. Yet, these methods can only provide a delineation of field boundary rather than object-level field results of more practical value. Recent multitask models, benefiting from the constraints of boundary signals, can help the region identification task to achieve parcel-level predictions and thus obtain object-level fields, which can be further optimized by morphological thinning and Douglas-Peucker method (M. Li et al. 2023). Some multitask models attempted to intersect the identified cropland regions with the detected boundaries (S. Liu et al. 2022; Xu et al. 2022). However, cropland region recognition tasks focus more on semantic information and may not effectively represent the transition boundaries between cropland and non-cropland areas compared to boundary tasks, potentially resulting in an inaccurate correspondence between the fields obtained from these region-based methods and the real ones. Some studies have attempted to directly obtain object-level fields from boundary detection results. For example, Watkins (Watkins and van Niekerk 2019a) performed region segmentation method

on the detected boundary layers and carried out a comparative analysis of different segmentation methods, showing that watershed segmentation is most effective in generating clear boundaries. Similarly, Waldner and Diakogiannis (Waldner and Diakogiannis 2020) used watershed segmentation to convert detected boundary information into object-level fields. However, this segmentation method can only create boundaries along high gradient, neglecting weaker boundary parts, which can lead to under-segmentation of adjacent parcels (D. Li et al. 2010). Therefore, further exploration is needed to address weak boundary loss during region segmentation.

The choice of segmentation parameters is unavoidable when using segmentation methods to transform the pixel-level boundaries results into object-level fields. One method involves additionally predicting the distance transformation of cropland and applying a threshold to obtain field-level seeds and implement the field-level segmentation (Waldner and Diakogiannis 2020). Nevertheless, the binarized seed points heavily depend on the threshold choices and do not always correspond to the fields one-to-one, which may result in over-segmentation and under-segmentation. Watkins (Watkins and Van Niekerk 2019b) served all local minima as seed points for watershed segmentation and obtained completely over-segmented regions. The final fields can be merged by removing the common edges with a given boundary strengths threshold. However, the optimal strength threshold varied across regions, and such unstable and heavily manual-dependent parameter are hardly of practical value. Therefore, it is necessary to explore a more parameter-stable region consolidation method that minimizes manual intervention.

In general, there is a lack of research assessing the key techniques for field boundary detection and challenges related to weak boundary loss and unstable parameters persist in converting detection results into high-quality fields (Figure 1). Based on feature analysis, this study aims to address these issues and develop a generalized framework for agricultural field delineation (GF-AFD). The representative study area and extensive comparison experiments were set to validate the suggested framework and conclusions. Overall, this study's contributions lie in:

(1) Summarizing and evaluating three key techniques in field boundary detection models, including multi-tasking for regions and boundaries, balancing lowest-level detail features with high-level category features, and emphasizing boundary-focused training, which provides a baseline reference for model design.
(2) Demonstrating that performing region segmentation on boundary detection results produces more regular fields compared to the region-based methods. The adopted oriented watershed method can enhance the weak boundary parts, addressing the neglected problem of weak boundary loss and yielding more independent results.
(3) Analyzing that agricultural fields exhibit an observational hierarchy and developing a hierarchy-based method for merging over-segmented regions. This method proves to maintains parameter stability across large-scale scenes and different detection models, reducing the need for manual intervention in parameter selection.

2. GF-AFD delineation framework

The GF-AFD delineation framework covers the F-BD and A-FG two parts, where the A-FG part consists of a two-step process of field region segmentation and consolidation (Figure 2). Initially, the detection model is employed to detect field boundary and region information. Subsequently, a segmentation method is applied to transition pixel-level boundaries into object-level regions, followed by a consolidation step to merge over-segmented regions into field-level results. The final field results were derived after removing a few error regions using the detected cropland masks.
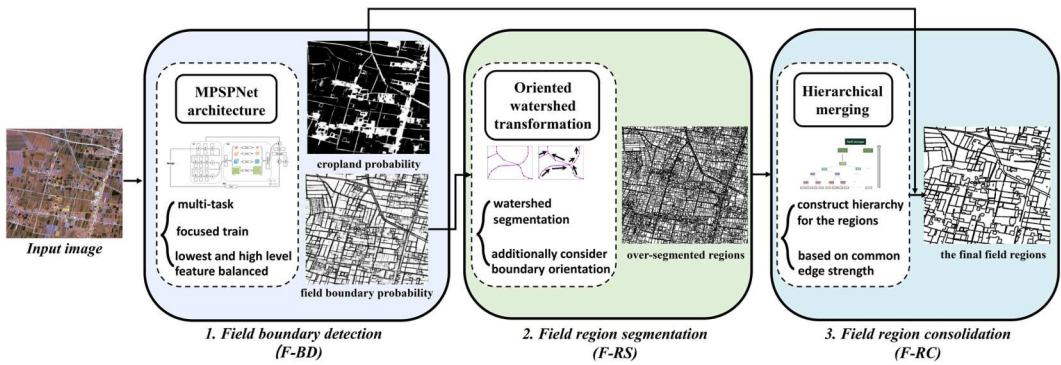
Figure 2. The flowchart of the proposed GF-AFD delineation framework.

2.1. Field boundary detection (F-BD)

2.1.1. Feature analysis of field boundaries

Field boundary is a semantic boundary which can be marked by various objects, such as roads or ditches, exhibiting complex features that are difficult for models to learn. Therefore, a simple field boundary/non-field boundary classification has limitations in generalizing features. However, incorporating cropland extent information significantly enhances the interpretability of field boundaries, conceptually constraining them as 'boundaries within cropland regions'. Figure 3 illustrates two types of field boundaries: outer boundaries demarcating cropland from non-cropland, often marked by wide features like roads, rivers, or drastic feature changes; and
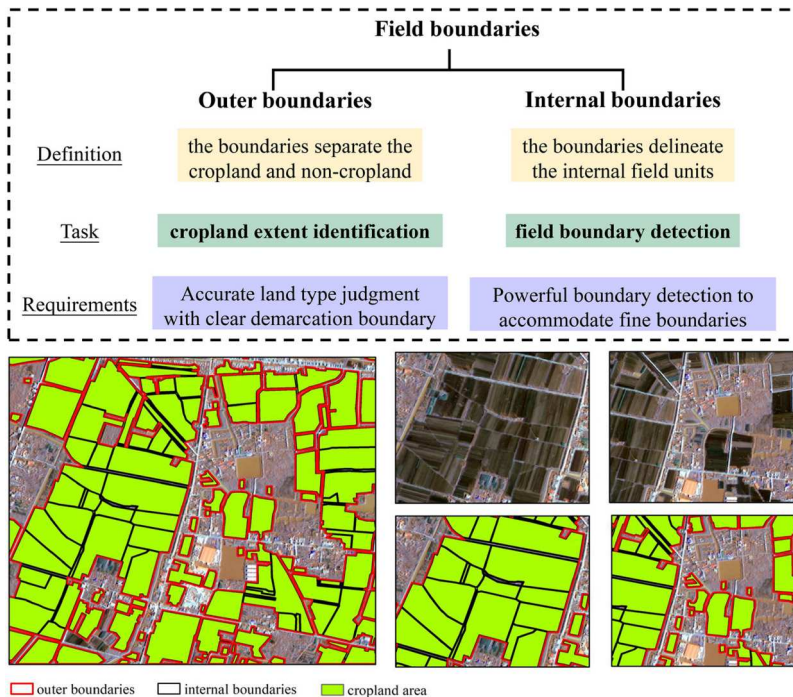
Figure 3. The conceptual and visual presentation of some types of field boundaries.

internal boundaries, typically fine field ridges dividing field units. Identifying outer boundaries corresponds to cropland region identification task, while detecting internal boundaries relates to boundary detection. Therefore, the detection model needs to be adapted to multi-tasking for regions and boundaries, with rich semantic information for region recognition, and powerful boundary perception capability.

In fact, these two tasks have a mutually reinforcing relationship. Cropland extent can constrain the presence of field boundaries, suppressing irrelevant boundaries, and field boundaries offer detailed information for finer parcel delineation within cropland, compensating for the limitations of coarse-grained analysis inherent in cropland extent identification.

Powerful high-level semantic features are necessary for cropland recognition and boundary type differentiation. However, field boundaries can be extremely narrow, sometimes as thin as one pixel, necessitating the inclusion of fine-grained, lowest-level features to maintain their integrity. Therefore, the model needs to effectively balance these two types of features.

In addition, the field boundaries occupy a very small percentage in the image, which requires the model to set a higher misclassification penalty on them and thus focus more on the boundary training. Hence, based on the feature analysis of field boundaries, it can be summarized that 'multi-tasking for regions and boundaries,' 'balancing lowest-level detail features with high-level category features,' and 'emphasizing boundary-focused training' are critical for detection.

2.1.2. Modifying MPSPNet model for assessment

In order to verify the impact of the summarized three key techniques, we adapted a common semantic segmentation network, MPSPNet, according to the techniques. This is a typical network applied for cropland region identification, consisting of the original PSPNet benchmark network (Zhao et al. 2017) and an added single convolutional layer, which has been proven to achieve high-precision cropland extent identification in four Chinese provinces (D. Zhang et al. 2020).

We adapted the MPSPNet model to multitasking for regions and boundaries, but this model requires more lowest-level features to enhance its boundary-awareness and detail-preserving capabilities. The boundary information is the first-order feature in the image, describing the discontinuity between pixels. In fact, most of the edge operators, such as Sobel, Prewitt, Scharr, etc., are designed to detect the boundary gradient by the operation of a single convolution kernel, and multi-convolution will instead weaken the integrity of boundaries. Therefore, we utilize the single convolutional layer that has been added in the MPSPNet for the boundary detail preserving in detection. Compared the original MPSPNet network, we increase the channels fusion ratio of the lowest-level features and the original deep features to 1/16, i.e. 32 and 512, to enhance the representation of the boundary information in the network, and this ratio is validated by the ablation experiments in this study (Figure 4).
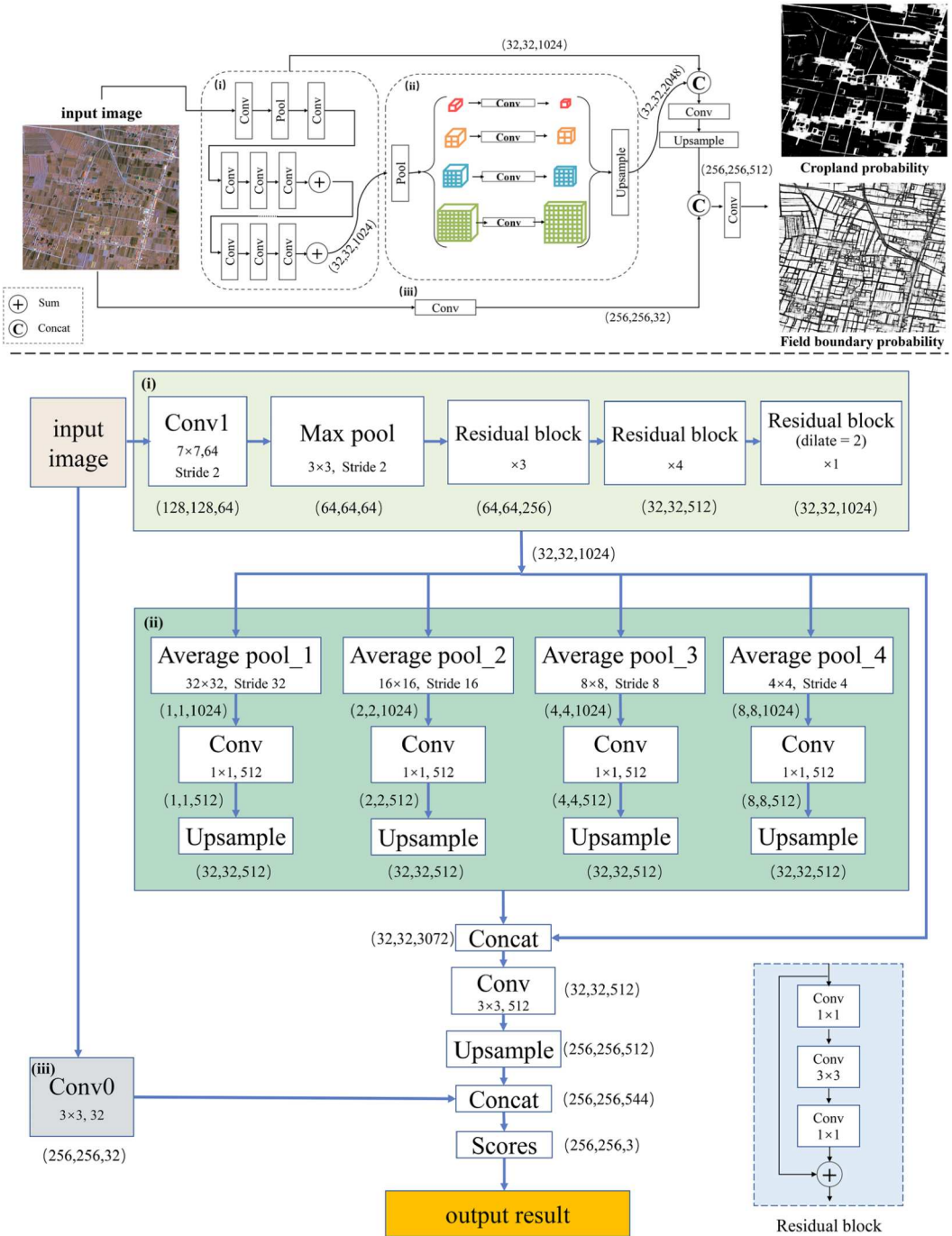
Figure 4. Modified MPSPNet network architecture. The top half is the Conceptual illustration of the network, and the bottom half describes the detailed components. (i) The ResNet with the dilated convolution; (ii) the spatial pyramid pooling module; (iii) the addition low-level information module.

In addition, the classes 'field boundary' and 'non-field boundary' are highly imbalanced in field boundary detection task. It is necessary to set a greater weight on the field boundary in the loss function to help the model training to be more focused on boundary. The weight value is calculated

from the ratio between the two class (Y. Liu et al. 2017). The loss functions are defined as follows:

$$L(bou) = \frac{1}{N}\left(-\sum_{n=1}^{N}\left(\frac{|Y-|}{|Y+\cup Y-|}y^{+\prime}\log y^{+} + \frac{|Y+|}{|Y+\cup Y-|}y^{-\prime}\log y^{-}\right)\right) \qquad (1)$$

$$L(reg) = \frac{1}{N}\left(-\sum_{n=1}^{N}(y^{+\prime}\log y^{+} + y^{-\prime}\log y^{-})\right) \qquad (2)$$

$$Loss = L(bou) + L(reg) \qquad (3)$$

where $L(bou)$ and $L(reg)$ are the loss values of the field boundary detection and cropland identification, N is the number of image pixels of image, y^{+} and y^{-} are prediction positive pixels and negative pixels, $y^{+\prime}$ and $y^{-\prime}$ are the label category of them, and $|Y+|$ and $|Y-|$ are the number of their pixels, respectively.

2.1.3. Ablation experiments setting

The following four ablation experiments were conducted to assess the performance enhancement effect results from the key techniques: (1) Detection using the PSPNet benchmark network without the lowest-level feature connection. (2) Detection using a single field boundary detection task (denoted as MPSPNet-sin). (3) Detection using a model without boundary-focued training (denoted as MPSPNet-NW). (4) An ablation experiment of the fusion ratio of the added low-level feature channels and the original feature channels.

2.1.4. Comparison with other detection models

Modified MPSPNet model was compared with existing models to assess its performance and the impact of the modified key techniques. The comparison models in this study include a modified DeeplabV3+ model, an edge detection networks DexiNed, three UNet-based networks ResU-Net, R2U-Net and ResU-Net-a, and two recent multitasking models BsiNet and SEANet. These seven methods are described below and the architectures of these models are presented in Figure 5.

DeeplabV3+ model (Chen et al. 2018): A widely-used model in cropland region recognition, it employs a similar strategy as MPSPNet model for fusing high-level features with lowest-level features. It was adapted in this study by increasing the number of lowest-level feature channels and incorporating multi-tasking for regions and boundaries, as well as loss function tuning for boundary-focused training. This comparison assesses the generalizability of the performance improvements resulting from the three key techniques.

DexiNed (Soria, Riba, and Sappa 2020): An edge detection network based on side structure, providing stronger boundary awareness compared to other edge models. This comparison helps evaluate the advantages of semantic segmentation models over edge detection networks.

R2UNet: A U-Net-based network utilizing recurrent residual blocks to enhance feature representation, recently used for field boundary detection from Sentinel-2 satellite images (H. Zhang et al. 2021).

ResUNet: A U-Net-based network with residual blocks that maintain good segmentation performance while reducing the number of parameters (Taravat et al. 2021).

ResUNet-a: An adaptation of ResU-Net for multi-tasking in region and boundary detection, known for high accuracy in field boundary detection (Waldner and Diakogiannis 2020).

The comparison with these three U-Net-based models helps to assess whether a more balanced feature allocation can lead to better performance.

BsiNet: A multi-task network based on PsiNet, using a single decoder to produce region, boundary, and distance information. BsiNet has achieved higher accuracy than ResUNet-a (Long et al. 2022).

SEANet: A recently proposed multi-task network designed with separate decoding modules for region and boundary detection, known for state-of-the-art accuracy in field boundaries from high-resolution satellite images (M. Li et al. 2023).

The comparison with these two recent multitasking models helps to assess whether some common semantic segmentation models can achieve comparable accuracy to the state-of-the-art method after modification with key techniques.
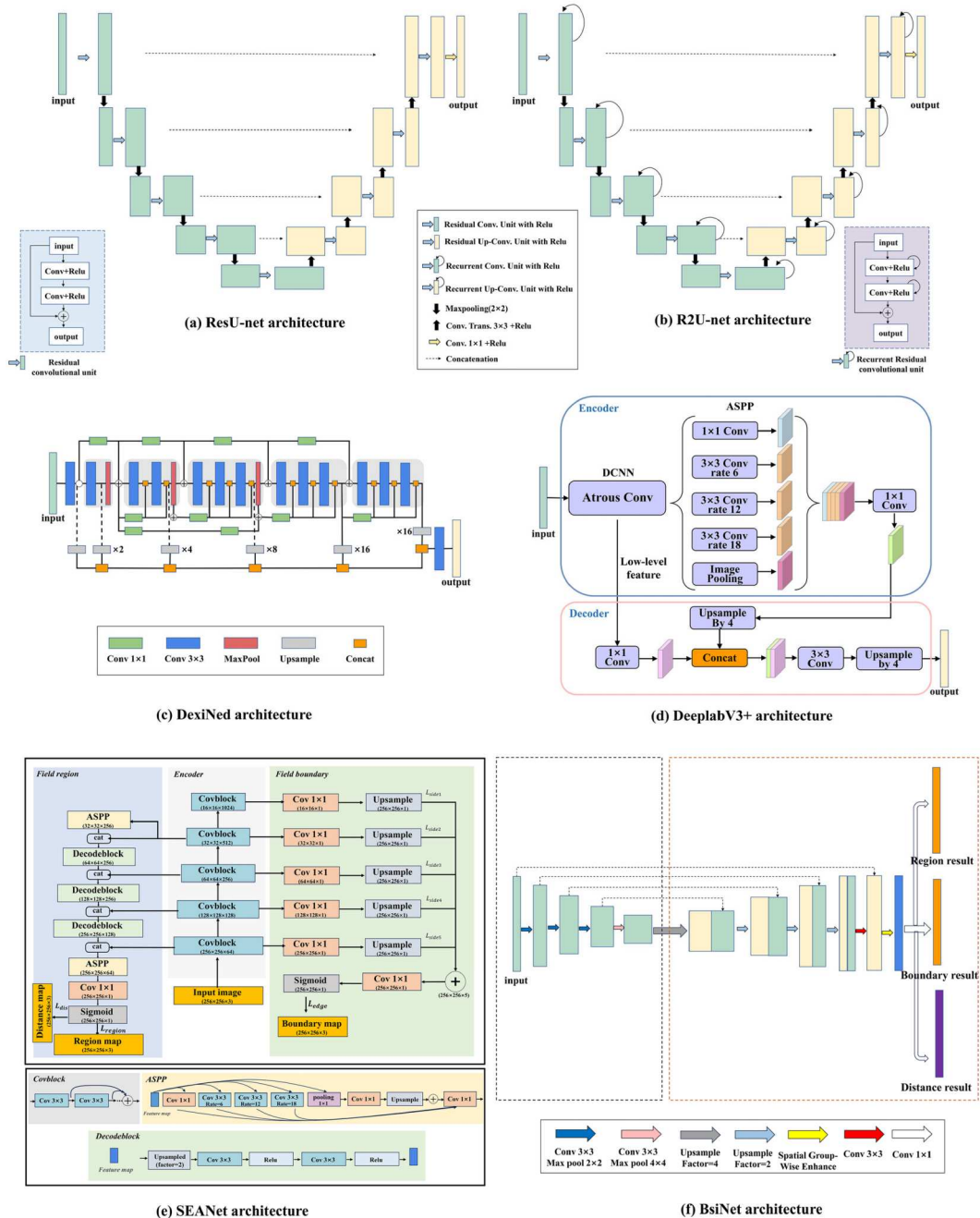
Figure 5. The architectures of ResU-Net, R2U-Net, DexiNed, DeeplabV3+, SEANet and BsiNet.
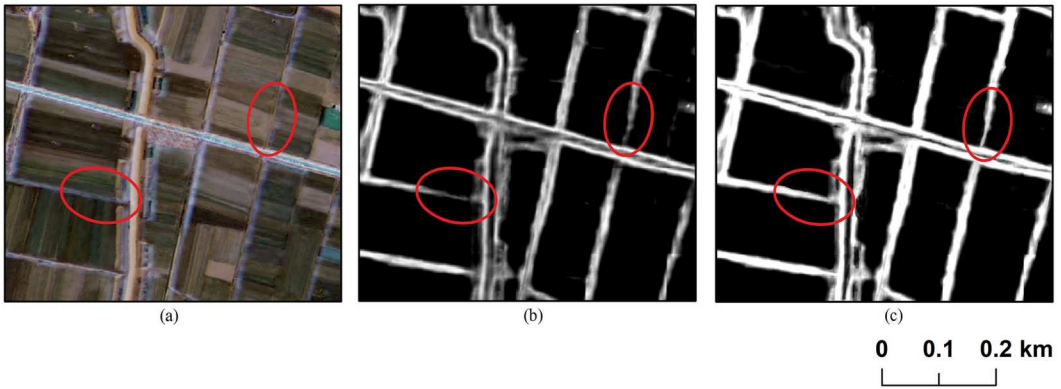
Figure 6. The presentation of some weak boundaries. (a) a local image; (b) the detected boundary probability; (c) the oriented boundary signal according to the gradient orientations. The circle shows where the weak boundary is located.

2.2. Agricultural field generation (A-FG)

2.2.1. Feature analysis of field region segmentation

Due to the blurred image boundary or the detection error, there will be some boundary parts with low probability in the detection result (Figure 6(b)). Although performing the watershed segmentation on the probability can yield object-level boundary results, the method only creates boundaries along the high-value gradients and ignores the weak boundary part, resulting in extracted adjacent parcels that are not completely separated.

It can be found that the field boundary is directional, and the detected high-value probability is also extended along the boundary direction. Therefore, the directional information can be helpful to orientationally strengthen the weak boundaries.
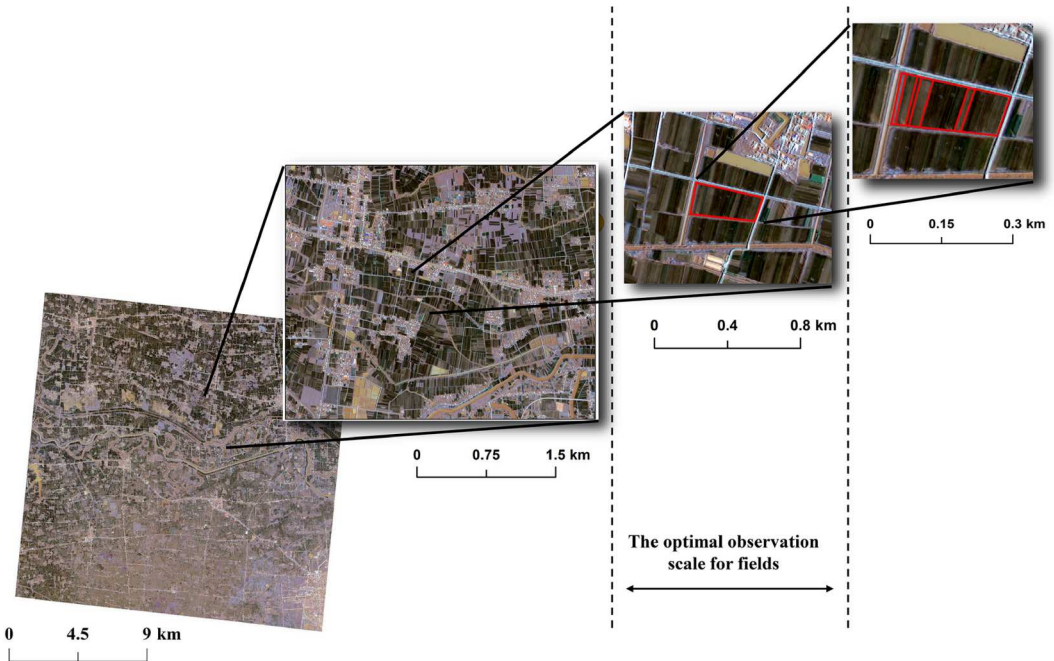
Figure 7. The presentation of different observation scale of fields.

2.2.2. Oriented watershed transformation

This paper adopted the oriented watershed transformation for region segmentation (Arbeláez et al. 2011), which additionally considers the gradient direction during segmentation. This method recursively subdivides the arcs and obtains the orientations by approximating the arcs with line segments. Then estimates the orientation of each pixel by the orientation of the arc where the pixel lies. The watershed transformation was performed on the oriented boundary signal (Figure 6 (c)), where the weak boundary is significantly strengthened and thus will be preserved in the segmentation process. To a certain extent, this method functions as a boundary connection.

2.2.3. Feature analysis of field region consolidation

The fully over-segmented regions after segmentation need to be further merged into a complete field object. However, the optimal merge parameters may vary across different regions in the image, which makes the consolidation process complex.

It can be found that fields exhibit an observation scale in the image (Figure 7). The process of outlining individual field typically involves first observing the entire image to identify the large cropland area, and then focusing on individual fields. Within each field, further subdivision attempts may be necessary to determine the optimal outline level. Thus, there is a hierarchical progression and level selection process for visual identification. In the same image, the optimal observation scale of fields is kept consistent. Therefore, it can help to achieve more stable regions merging by constructing a hierarchy for the segmented regions and simulating this observation scale.

2.2.4. Hierarchical merging

The hierarchy can be constructed by iterative greedy merging of regions (Cousty et al. 2009), and it is crucial to choose the appropriate merge weights. In this study, the strength values of common edges are calculated to measure the dissimilarity between two adjacent regions and decide the merging order. Specifically, a greedy graph $G = (R, E, W(Ei))$ was defined, where the finest regions R is the nodes of the graph, and E and $W(Ei)$ are the common edges and their weight of the adjacent regions. by iteratively merging the adjacent regions with the lowest common edge weights, a hierarchical tree structure with different segmentation levels is finally built. Figure 8 shows the construction process.

The leaf nodes of the constructed hierarchy represent all the initial segmentation regions, while the root node depicts the full image. The structure is a nested sequence of coarse to fine-segmented objects. The hierarchy levels reflect the different scales of segmentation. In this way, the local merging parameter selection is transformed into the global hierarchical-level selection for the whole
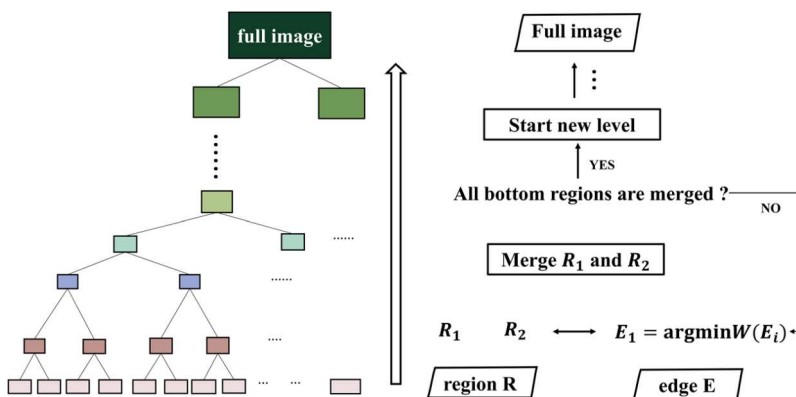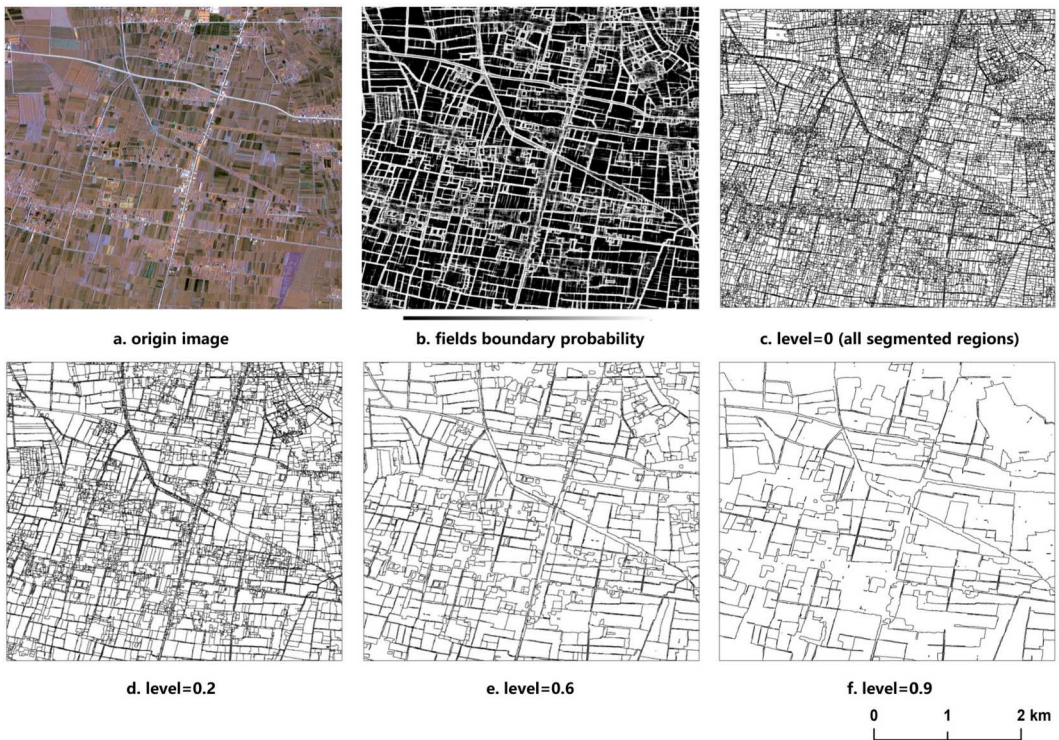

Figure 8. The process of constructing the hierarchy.

Figure 9. The delineation results at different levels in a local image.

image. Figure 9 illustrates the extraction results at various levels within a local image. By selecting the optimal level of the image, the merged field object can be obtained, and the final results were derived after removing a few irrelevant boundaries located in non-cropland areas using the detected cropland masks.

2.2.5. Comparison with other agricultural field generation methods

This paper adopted the oriented watershed transformation (OWT) to transform the detected boundary probability to object-level boundary and used a hierarchy-based merge method (Higra) to obtain the final field results. To demonstrate its (OWT + Higra) effectiveness, four comparison experiments are conducted:

(1) Comparison with a threshold-based binarization approach, known as local adaptive thresholding (LAT). This method calculates an independent threshold for each pixel and generates field boundary results (Graesser and Ramankutty 2017).
(2) Comparison with the field results identified by the region task of a multi-task model, where the field boundaries were additional simplified by morphological thinning and Douglas-Peucker method (denoted as Region task-based method).
(3) Comparison with the method of clipping identified cropland region with boundary results. This method is commonly used in multi-task models to generate agricultural fields (denoted as Region-Boundary Fusion).
(4) Comparison with watershed segmentation (WS) method (D. Li et al. 2010), which assesses the advantage of the additional orientation information provided by the OWT method. The over-segmented regions of the WS segmentation were subsequently merged by the Higra method to allow a direct comparison of the field boundaries (denoted as WS + Higra).
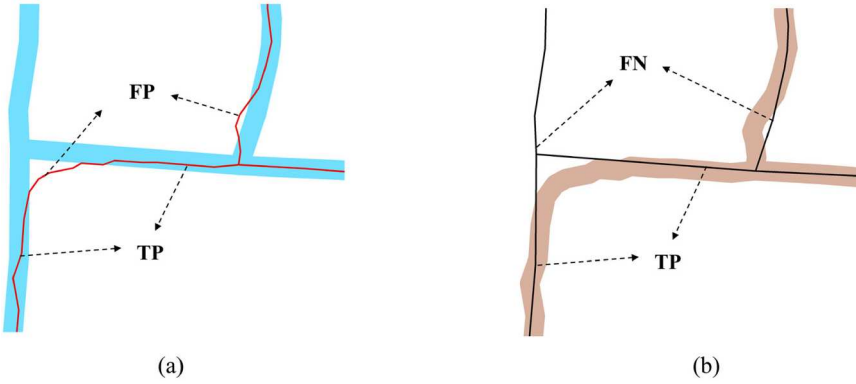
Figure 10. Schematic representation of TP, FP, and FN boundaries: (a) extracted boundary overlaid on the buffer area around the reference boundary; (b) reference boundary overlaid on the buffer area around the extracted boundary.

(5) Comparison with local merging-based method (LM), which provides an evaluation compared to our hierarchy-based merging method. local merging-based method merges objects by removing common edges whose strengths is lower than a specified height threshold, which is calculated based on local standard deviation (Watkins and Van Niekerk 2019b). This merge method was performed on the same over-segmented regions produced by the OWT method (denoted as OWT + LM).

2.4. Accuracy assessment

2.4.1. Field boundary accuracy

The boundary accuracy is the correspondence degree between the detection and reference boundaries. Considering the narrow width of the boundary, absolute correspondence can be difficult to achieve. Two pixels accuracy tolerance offsets were set in this study. True positive (TP), false positive (FP), and false negative (FN) confusion matrix metrics are used to assess the accuracy of correctly identified, incorrectly identified, and unidentified field boundaries, respectively. These metrics are computed based on the buffer polygons shown in Figure 10.

The F1 score calculated from the confusion matrix metrics is used as a measure, which is the summed average of precision and recall. The larger the F1 score, the closer to the optimal delineation. The equation is as follows:

$$F1 = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \tag{4}$$

$$\text{boundary precision(BP)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

$$\text{boundary recall(BR)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

2.4.2. Field geometric accuracy

Field geometric accuracy can be divided into three parts: area accuracy, position accuracy and shape accuracy.

a. area accuracy

This accuracy analyzes the correctness and completeness of the delineated fields. Consistent with the boundary accuracy, The F1 score was used as a metric for area accuracy.

b. position accuracy

This accuracy represents the match degree between the centroids of the extracted and reference fields. First, the Euclidean distance between the two centroids is calculated, and then it is normalized by the diameter of equal area circle:

$$P^i_{centroid} = 1 - \frac{d(C_T, C_E)}{D_{cac}} \tag{7}$$

$$D_{cac} = 2\sqrt{\frac{S_T + S_E}{\pi}} \tag{8}$$

$$P_{centroid} = \frac{P^i_{centroid}}{N} \tag{9}$$

where $P^i_{centroid}$ Pcentroididenotes the position accuracy of the ith field, $d(C_T, C_E)$ is the Euclidean distance between the centroids, $S_T + S_E$ is the combined area of two fields. and N is the field number.

c. shape accuracy

This accuracy measures the shape similarity of the delineated and reference fields. The normalized perimeter index (NPI) is applied to define the shape factor and the accuracy was expresses by
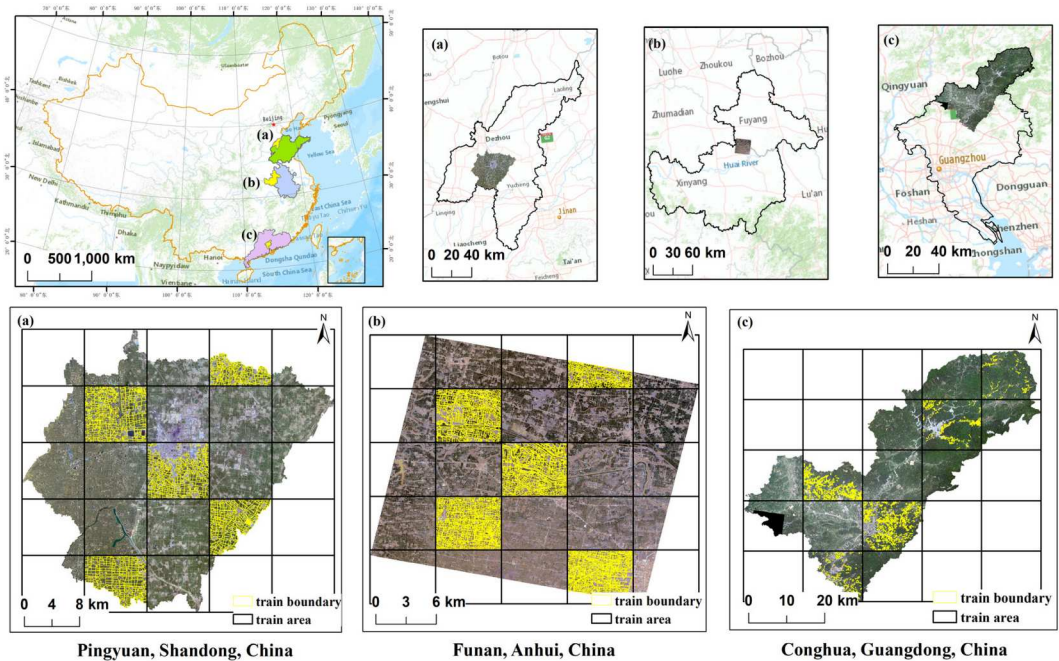
Figure 11. Study area and training area: (a) is the Pingyuan County, (b) is the Funan County, and (c) is the Conghua County.

Table 1. The details of selected study sites and their images.

Study site	Common crops	Topography	Field geometric characteristics	Cropland percentage	Image data	Image acquisition time
Pingyuan	corn, wheat, beans	plain	large and regular	63.9%	GF-1 (2 m)	July 2018
Funan	rice, wheat, fruit, rapeseed	plain	small and regular	55.3%	GF-2 (1 m)	January 2021
Conghua	rice, citrus, tea, pummelo	hilly	small and irregular	17.2%	GF-1 (2 m)	September 2019

the ratio of the two.

$$P^i shape = \frac{NPIEi}{NPITi} \tag{10}$$

$$NPI = \frac{Peac}{Pobject} \tag{11}$$

$$Pshape = \frac{P^i shape}{N} \tag{12}$$

where P^i_{shape} is the shape accuracy of the ith field, P_{object} and P_{eac} are the perimeter of the object and its equal area circle, NPI_{Ei} and NPI_{Ti} are the NPI of the extracted and reference fields.

3. Study site and data

3.1. Study site

The GF-AFD delineation framework was evaluated utilizing three study sites, covering typical regions of China from north to south (Figure 11). Table 1 summarizes the region characteristics of each area. The first site is Pingyuan County, Shandong Province, a plain area with 1181.7 km^2 in North China. This area has a flat topography with pure crops planted, mainly summer corn and winter wheat. In addition, the fields in this region are regular in shape and large in size, with minimum field complexity. The second one is a 30 × 24 km area of large-scale agricultural production in Funan County, Anhui Province. This area also has a plain terrain, but since it is located in the transition zone between temperate and subtropical zones, the cropping structure is more complex and the fields are smaller, with some cash crops mixed. The third one is Conghua County, Guangdong Province, where the main terrain is hilly and the area is 3,724 km^2, with only 17% cropland shared and mixed with forests. This is produced in a small and irregularly shaped agricultural fields, exhibiting the most complex field characteristics of the three sites. Overall, these three study sites allow for the assessment of the framework's performance in different topographies, crop types, and field characteristics.

3.2. Data processing

The base data for the Pingyuan County was acquired from Gaofen-1 satellite, featuring a 2-meter panchromatic band and 8-meter multispectral bands (RGB + NIR). In this study, we utilized 2-meter fusion images with three bands (RGB). The cloud-free image of the entire county was mosaicked from GaoFen-1 images acquired in July 2018. The second study site utilized the Gao-Fen-2 RGB fusion image with 1 m resolution, acquired on 29 January 2021. GaoFen-2 image includes a 1 m panchromatic band and 4 m multispectral bands. This choice allowed us to assess the method's applicability on different resolution and winter images. Additionally, the base data for Conghua County is also the GaoFen-1 fusion image acquired from September 2019. Further
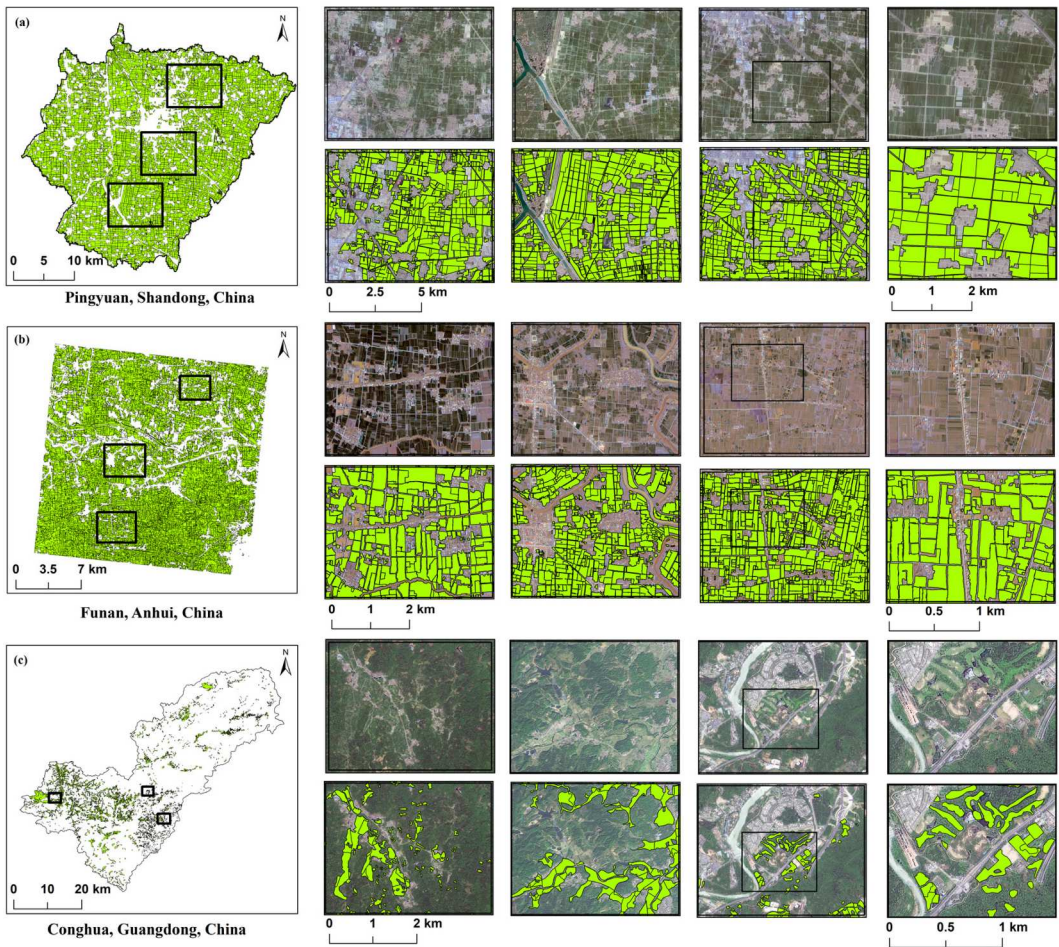
Figure 12. The delineation results of the three counties.

information about the Gaofen satellites can be found at (http://www.cheos.org.cn/n6084429/ n6084446/n6084504/n6128357/index.html). Overall, the experimental images provide an assessment of the method's performance in different image resolutions and growing seasons.

In all experimental areas, the agricultural field sample was obtained manually, and it also uses the situ identification data to validate the outlined fields, and all the areas exceeded 95% accuracy. We divided the extent of each study area into 25 equal-sized tiles and randomly selected five of them in each area for training.

The outlined vector polygons of cropland parcels were converted into rasterized region samples, assigning a pixel value of 255 to cropland and 0 to non-cropland types. This dataset was utilized for training the cropland region recognition task. Concurrently, the vector lines demarcating field boundaries were rasterized with a width of 3 pixels using a multi-ring buffer. This rasterized data was employed as the sample for field boundary task.

The Adam optimizer is used for gradient descent optimization with an initial learning rate of 0.001, a batch size = 16, weight decay = 1e-5, and 80 epochs during the training process. Momentum was set to 0.9 to regularize learning. All experiments in this article were executed with python 3.6, and pytorch 1.6, an Nvidia GTX 1080 Ti for GPU acceleration.

Table 2. The accuracy of the delineated results for the three study areas, the bolded values represent the highest accuracies.

| Area	Field boundary accuracy			Field geometric accuracy		
	F1	BP	BR	Area accuracy	Position accuracy	Shape accuracy
Pingyuan	**0.923**	0.915	0.931	**0.963**	**0.936**	**0.889**
Funan	0.883	0.862	0.905	0.922	0.903	0.871
Conghua	0.817	0.793	0.842	0.879	0.899	0.802
average	0.874	0.856	0.892	0.921	0.913	0.854

4. Result

4.1. Large-scale delineation results

The visual inspection of the delineation results confirms the good performance of the proposed method (Figure 12). In the plain areas (Pingyuan and Funan), the method can accurately distinguish cropland from other land types such as residential land, rivers and roads. Notably, the method demonstrated an ability to effectively identify fallow land in Funan. In mountainous area (Conghua), the method also effectively delineated fields from the forested areas that have similar spectral characteristics. The identified field boundaries in three areas were also accurate, including internal boundaries that are challenging to identify, and the field exhibited regular shapes. Furthermore, the method was able to identify some small and fragmented fields completely.

4.2. Statistical accuracy assessment

The paper then evaluates the accuracy of the results in the three regions (Table 2), where the boundary accuracy (F1) and area accuracy of all areas exceed 0.8, indicating a strong ability for boundary detection and cropland identification. Additionally, the method performs well in maintaining the location and shape of the identified agricultural fields, with an average location and shape accuracy of 0.913 and 0.854, respectively.

Of the three study areas, Pingyuan County yielded the best results, with the highest accuracy in all metrics. Although the GF-2 image in Funan County has a higher resolution and provides more detailed information, its accuracy still decreased compared to that of Pingyuan County. It was observed that under the premise of boundary clarity, a higher resolution does not result in an improvement in recognition accuracy, but rather may introduce more potential misleading due to the finer detail information. In addition, all accuracy metrics in Conghua County, located in a
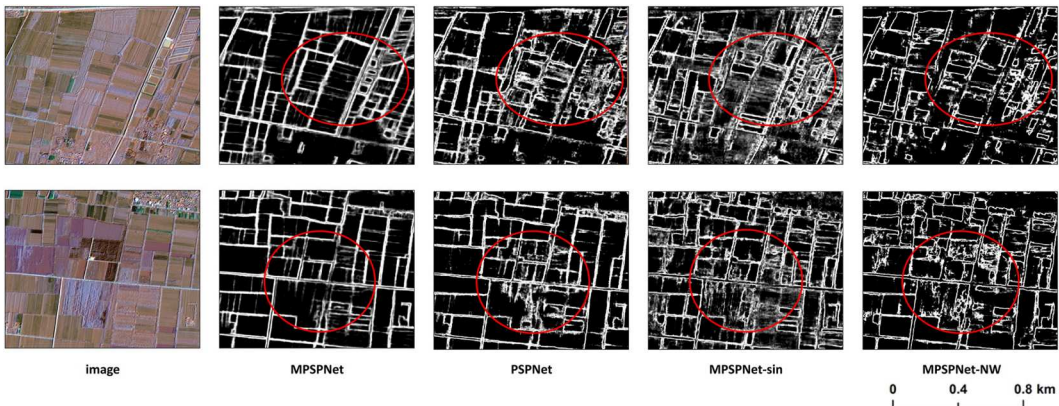
Figure 13. Detection results of different ablation experiments. The circles for regions with boundary blurring and irrelevant boundaries.

Table 3. The accuracy of the delineated results for different ablation experiments, the bolded values represent the highest accuracies and italicized values represent the accuracies of the MPSPNet.

Area	Method	Field boundary accuracy			Field geometric accuracy		
		F1	BP	BR	Area accuracy	Position accuracy	Shape accuracy
Pingyuan	*MPSPNet*	***0.938***	*0.915*	*0.962*	***0.963***	***0.936***	***0.876***
	PSPNet	0.867	0.893	0.842	0.908	0.939	0.859
	MPSPNet-sin	0.904	0.896	0.912	0.941	0.932	0.871
	MPSNet-NW	0.902	0.889	0.914	0.932	0.933	0.833
Funan	*MPSPNet*	***0.889***	*0.873*	*0.905*	***0.922***	*0.903*	*0.871*
	PSPNet	0.827	0.854	0.801	0.902	**0.912**	**0.878**
	MPSPNet-sin	0.863	0.831	0.897	0.893	0.897	0.863
	MPSNet-NW	0.852	0.842	0.861	0.908	0.883	0.853
Conghua	*MPSPNet*	***0.817***	*0.793*	*0.842*	***0.879***	*0.899*	***0.802***
	PSPNet	0.804	0.812	0.796	0.863	**0.907**	0.799
	MPSPNet-sin	0.779	0.763	0.795	0.848	0.881	0.787
	MPSNet-NW	0.808	0.791	0.826	0.874	0.864	0.795

mountainous area, were lower than those in the other two areas. However, all of these metrics were still higher than 0.8. This is because, despite the similarity in spectral features between the mountainous forest and farmland, the deep learning model can extract abstract category features, such as texture patterns, which are sufficient for cropland recognition. However, the junction between cropland and mountainous forest in this region is not clear, which is the primary reason for the accuracy gap.

4.3. Field boundary detection performance

4.3.1. Model ablation experiments

The local results detected by different ablation models are depicted in Figure 13. The detection results were transformed into the field boundary and processed for accuracy verification, and the accuracy results are presented in Tables 3 and 4.

(1) Comparison with PSPNet. Due to the lack of low-level detailed features, the PSPNet will misidentify some image pixels around the boundary, yielding ambiguous and unclear boundary results. In terms of accuracy, the detection results of the original PSPNet network exhibit a noticeable decrease in accuracy compared to the modified PSPNet (MPSPNet) network which incorporates a shallow feature module.

(2) Comparison with single-task detection. the complexity of field boundary types made it challenging to generalize accurate and comprehensive features for a single 'field boundary detection' task, as a result, the model exhibited confusing recognition in some areas. However, by incorporating the feature signals obtained during the 'cropland extent identification' task, the multi-task network was able to facilitate boundary category determination, leading to significant improvements in model accuracy and recognition performance.

Table 4. The accuracy results by the model with different number of low-level feature channels, the bolded values represent the highest accuracies.

Number of low-level feature channels	F1 accuracy				Area accuracy			
	Pingyuan	Funan	Conghua	average	Pingyuan	Funan	Conghua	average
8	0.879	0.826	0.795	0.833	0.943	0.897	0.873	0.904
16	0.923	0.883	**0.826**	0.876	0.949	**0.930**	**0.882**	0.919
32	0.938	**0.889**	0.817	**0.881**	**0.963**	0.922	0.879	**0.921**
64	**0.946**	0.847	0.789	0.860	0.916	0.889	0.836	0.880
128	0.916	0.811	0.756	0.827	0.874	0.842	0.769	0.828
256	0.893	0.760	0.721	0.791	0.832	0.807	0.728	0.789
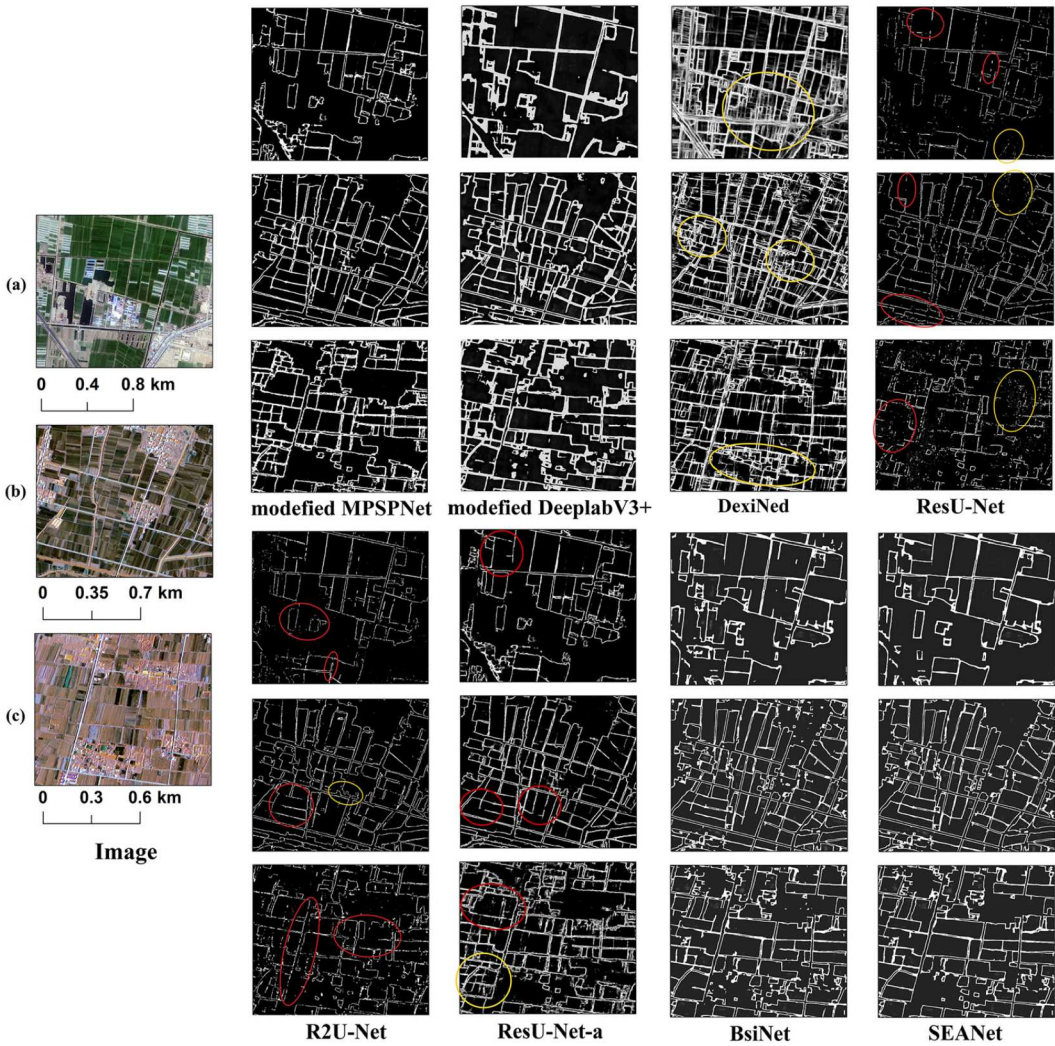
Figure 14. Detection results of different comparison experiments. The circles for regions with boundary omission and misidentification.

(3) Comparison with the model without boundary-focused training. The implementation of assigning higher misclassification cost to field boundary was found to effectively address the class imbalance issue in boundary recognition. This approach directed the model's focus toward field boundary class and extracting more accurate field boundary features during model training. This is reflected in the results as more distinct and clear boundaries.

(4) Feature channel assignment experiment. To enhance the model response to the boundary, we fused 32 additional low-level feature channels with the PSPNet model's 512 high-level feature channels, rather than the 16 channels used in the original MPSPNet model. To determine the optimal fusion configuration, we conducted an ablation experiment and evaluated the boundary and area accuracy of models with varying numbers of low-level feature channels (8, 16, 32, 64, 128, and 256). The results are presented in Table 4. In mountainous regions (Conghua) with scarce cropland, a higher number of high-level semantic features are required for accurate cropland recognition, whereas in plains with abundant cropland, shallower information is more beneficial. Ultimately, the 32 low-level channels provide a balanced representation of

Table 5. The accuracy results for the study areas of different detection model, the bolded values represent the highest accuracies and italicized values represent the accuracies of the our modified MPSPNet.

Area	Method	Field boundary accuracy			Field geometric accuracy		
		F1	BP	BR	Area accuracy	Position accuracy	Shape accuracy
Pingyuan	*Modified MPSPNet*	*0.938*	*0.915*	*0.962*	***0.963***	***0.936***	*0.876*
	Modified DeeplabV3+	0.918	0.933	0.903	0.942	0.914	0.891
	DexiNed	0.801	0.753	0.855	0.873	0.864	0.807
	ResU-net	0.833	0.873	0.796	0.909	0.882	0.831
	R2U-net	0.907	0.912	0.902	0.951	0.917	0.857
	ResU-Net-a	0.922	0.928	0.916	0.955	0.921	0.863
	BsiNet	0.942	0.937	0.947	0.949	0.905	0.877
	SEANet	**0.951**	0.956	0.946	0.961	0.930	**0.896**
Funan	*Modified MPSPNet*	*0.889*	*0.873*	*0.905*	*0.922*	*0.903*	*0.881*
	Modified DeeplabV3+	0.903	0.891	0.915	0.916	0.915	**0.906**
	DexiNed	0.719	0.674	0.771	0.798	0.876	0.723
	ResU-net	0.762	0.819	0.712	0.839	0.885	0.786
	R2U-net	0.811	0.858	0.769	0.874	0.911	0.807
	ResU-Net-a	0.864	0.871	0.857	0.897	0.917	0.822
	BsiNet	0.882	0.869	0.895	0.911	0.891	0.877
	SEANet	**0.909**	0.901	0.917	**0.931**	**0.922**	0.904
Conghua	*Modified MPSPNet*	*0.817*	*0.793*	*0.842*	*0.879*	*0.899*	*0.802*
	Modified DeeplabV3+	0.839	0.846	0.832	**0.906**	**0.913**	0.824
	DexiNed	0.731	0.689	0.778	0.759	0.853	0.738
	ResU-net	0.754	0.781	0.729	0.807	0.883	0.797
	R2U-net	0.773	0.792	0.753	0.843	0.902	0.811
	ResU-Net-a	0.786	0.798	0.774	0.859	0.907	0.817
	BsiNet	0.799	0.806	0.792	0.881	0.902	0.829
	SEANet	**0.843**	0.845	0.841	0.901	0.910	**0.833**

the model's category recognition and boundary localization abilities, resulting in a maximum average boundary and area accuracy of 0.881 and 0.921, respectively.

4.3.2. Comparison with other detection models
Based on the same field boundary/region training samples, the detection results and accuracy of the different comparison models are presented in Figure 14 and Table 5.

(1) Comparison with modified DeeplabV3+ architecture. The results demonstrate that the modified DeeplabV3+ method yields accurate and clear field boundary detection with high precision. Additionally, the utilization of Atrous Convolution and the ASPP module in the DeeplabV3+ model enhances its capability to recognize different land types. This enhancement is evident in complex areas such as Funan and Conghua, where the DeeplabV3+ model achieves even higher boundary accuracy than the MPSPNet model.
(2) Comparison with edge detection model (DexiNed). Although the DexiNed method has a powerful boundary-aware network structure and detects clear boundaries, it lacks the ability to differentiate boundary types effectively. As a result, it may detect incorrect boundary signals, such as weak linear noise within the cropland. Despite being trained on the same field boundary samples, DexiNed models are designed for generic boundary detection and struggle to extract high-level semantic features for class judgment, making it challenging to directly detect semantic boundaries like field boundaries. Consequently, the accuracies of this model are lower than those of all the semantic segmentation models adopted in this study.
(3) Comparison with three U-Net-based architectures. While U-Net-based models generate extremely fine boundary results due to abundant shallow features, the modified MPSPNet outperforms them in terms of continuous detection, fewer misclassifications, and omissions, particularly in complex land types. This is because the MPSPNet model allows for a more balanced representation of both lowest-level detail features and high-level abstract semantic

Table 6. The accuracy results for the ResU-net models with different skip connection layers, the bolded values represent the highest accuracies.

Area	The skip connection layers	Field boundary accuracy			Field geometric accuracy		
		F1	BP	BR	Area accuracy	Position accuracy	Shape accuracy
Average (Pingyuan, Funan, Conghua)	No connection	0.701	0.743	0.663	0.821	0.843	0.798
	S1	**0.815**	0.819	0.811	**0.855**	0.879	**0.831**
	S1 + S2	0.811	0.827	0.796	0.848	**0.885**	0.825
	S1 + S2 + S3	0.794	0.818	0.771	0.836	0.874	0.813
	S1 + S2 + S3 + S4	0.783	0.824	0.746	0.852	0.883	0.805

features, while the U-Net architecture fused the all features from low to high, which will affect the representation of the important features and lack balanced feature allocation. In fact, some middle-level features with low detail accuracy for boundary location and poor semantic information for category judgement have little benefit for model. Despite the efforts of ResU-Net and R2U-Net to enhance feature extraction by adopting more powerful convolutional units, ResU-Net-a further bolsters its performance by imparting interpretability to field boundaries via cropland region related tasks. However, the allocation of feature channels in these models remains unchanged, resulting in an insufficient representation of critical high-level category features and lowest-level detail features. Consequently, ResU-Net-a shows a lower boundary F1 accuracy compared to the modified MPSPNet model in three study areas.

An additional ablation experiment on ResU-Net for multi-level feature fusion was conducted, with results in Table 6 showing that connecting the lowest-level features from the encoder to the decoded features (S1) yields the best delineation performance. This approach outperforms the method without any skip-layer connections (no connection) and the method that connects more features, reinforcing the necessity of ensuring balanced feature allocation of lowest-level detailed features and high-level semantic information for field boundary detection.

(4) Comparison with two recent multi-task networks. Both networks produce clear and accurate field boundary results. In terms of accuracy, the modified MPSPNet model performs comparably with them. The modified MPSPNet model achieves higher boundary and area accuracies than the BSiNet model in Funan and Conghua. When compared with the state-of-the-art SEANet model, the average boundary and area accuracies of the modified MPSPNet model in the three areas only differ by 0.019 and 0.009.

Specifically, SEANet employs dedicated decoding modules for field boundaries and cropland regions. The boundary module adopts the side structure of HED to generate boundary outputs and weighted merge the multi-scale boundary predictions through a 1×1 convolution. This weight assignment is adaptive and differs from U-Net, where merging is done with equal weights. Statistically, we averaged the weights of multiscale boundary features in the three study areas, showing that

Table 7. The boundary accuracies (F1 score) of different types of predictions based on DeeplabV3+ model in Pingyuan county.

Area	The type of prediction (DeeplabV3+)	Boundary accuracy (F1 score)
Pingyuan	Multi-task model's boundary	0.918
	Multi-task model's region	0.879
	Single-task model's boundary	0.836
	Single-task model's region	0.549

Figure 15. The boundary and region results detected by multi-task and single-task DeeplabV3+ models: (a) the local image, (b) the ground truth of field boundary, (c) the field boundaries detected by multi-task model, (d) the field boundaries detected by single-task model; (e) the ground truth of cropland region, (f) the cropland region detected by multi-task model; (g) the cropland region detected by single-task model. The circles show areas where boundary predictions can provide more fine-grained parcel delineation information than region predictions.

lowest-level and high-level features carry significant weights. This observation validates the proposed technique of accommodating and balancing lowest-level detail features and high-level category features and offers a potential theoretical explanation for SEANet's substantial accuracy improvement over existing U-Net-based networks. Additionally, SEANet incorporates other techniques like task-dependent uncertainty and ASPP, contributing to its improved accuracy compared to the modified MPSPNet model.

4.3.3. Comparison with region and boundary detection results

Given our model's transformation for multitasking in predicting field boundaries and regions, it is crucial to validate which results possesses superior granularity and can offer more accurate and comprehensive cropland delineation information, thus serving as a benchmark reference for subsequent field generation. Employing the modified DeeplabV3+ model, we compared the boundary granularity of four predictions: multitasking model's region and boundary predictions, and single-task model's region and boundary predictions. Both single-task models were trained using the same field region or boundary labels. All results were transformed into boundary-level representations for accuracy assessment. Table 7 displays their boundary accuracies in the Pingyuan image.

As shown in Figure 15, applying a single-task semantic segmentation model directly to predict field boundaries (Figure 15(d)) leads to significant fragmentation, struggling to extract object-level parcels. Similarly, direct application of the model for cropland region prediction leads to region results lacking fine-grained delineation of internal boundaries and only roughly describing the overall cropland distribution (Figure 15(g)). In contrast, the multitasking network, integrating both region and boundary tasks, conceptually constrains field boundaries as 'boundaries within cropland regions.', significantly enhancing interpretability and producing more accurate and continuous boundaries (Figure 15(c)). The multitasking model's region prediction (Figure 15(f)) also outperforms the single-task model.

However, it is challenging to take multitasking model's region prediction directly as field results. As shown in the red-circled area of Figure 15, some internal field boundaries exhibit noticeable fragmentation, leading to parcel under-segmentation. In contrast, boundary predictions are

Table 8. The field accuracies of different agricultural field generation methods on a boundary layer detected from modified MPSPNet model, the bolded values represent the highest accuracies.

| Area	Method	Field boundary accuracy			Field geometric accuracy		
		F1	BP	BR	Area accuracy	Position accuracy	Shape accuracy
Pingyuan	LAT	0.848	0.871	0.826	0.903	0.929	0.771
	Region task-based	0.881	0.868	0.894	0.908	0.927	0.795
	Region-Boundary	0.893	0.872	0.915	0.917	0.921	0.804
	WS + Higra	0.907	0.886	0.928	0.938	0.918	0.845
	OWT + LM	0.817	0.811	0.823	0.907	0.897	0.798
	OWT + Higra	**0.938**	0.915	0.962	**0.963**	**0.936**	**0.876**
Funan	LAT	0.793	0.775	0.811	0.821	0.899	0.764
	Region task-based	0.837	0.834	0.840	0.883	0.906	0.817
	Region-Boundary	0.849	0.841	0.857	0.893	**0.915**	0.833
	WS + Higra	0.854	0.823	0.887	0.908	0.912	0.869
	OWT + LM	0.802	0.796	0.809	0.864	0.875	0.796
	OWT + Higra	**0.889**	0.873	0.905	**0.922**	0.903	**0.881**
Conghua	LAT	0.782	0.763	0.802	0.826	0.859	0.767
	Region task-based	0.795	0.786	0.804	0.836	0.897	0.768
	Region-Boundary	0.801	0.794	0.808	0.842	0.901	0.772
	WS + Higra	0.809	0.793	0.826	0.853	**0.906**	0.793
	OWT + LM	0.743	0.751	0.735	0.783	0.819	0.727
	OWT + Higra	**0.817**	0.793	0.842	**0.879**	0.899	**0.802**

more refined and provide more complete and clear fields delineation, which is also reflected in their superior boundary accuracy. Therefore, multi-task model's boundary prediction is a better baseline reference choice for generating field results.

4.4. Agricultural field generation performance

Tables 8 and 9 show the accuracy results of the different field generation methods performed on the boundary layer detected from the modified MPSPNet and the state-of-the-art SEANet model respectively. Figure 16 shows the field results generated by different methods, while Figure 17 illustrates the generation process of Region-Boundary Fusion, WS + Higra, and our OWT + Higra methods.

Table 9. The field accuracies of different agricultural field generation methods on a boundary layer detected from SEANet model, the bolded values represent the highest accuracies.

Area	Method	Field boundary accuracy			Field geometric accuracy		
		F1	BP	BR	Area accuracy	Position accuracy	Shape accuracy
Pingyuan	LAT	0.869	0.892	0.847	0.893	0.912	0.786
	Region task-based	0.921	0.917	0.926	0.926	0.917	0.849
	Region-Boundary	0.928	0.914	0.942	0.933	**0.929**	0.857
	WS + Higra	0.936	0.906	0.947	0.942	0.918	0.872
	OWT + LM	0.811	0.825	0.797	0.914	0.905	0.826
	OWT + Higra	**0.945**	0.932	0.958	**0.957**	0.927	**0.891**
Funan	LAT	0.822	0.796	0.849	0.843	0.891	0.757
	Region task-based	0.872	0.862	0.882	0.902	0.903	0.861
	Region-Boundary	0.879	0.868	0.889	0.909	0.909	0.868
	WS + Higra	0.886	0.879	0.893	0.917	0.907	0.882
	OWT + LM	0.795	0.806	0.784	0.871	0.883	0.806
	OWT + Higra	**0.897**	0.903	0.891	**0.929**	**0.919**	**0.887**
Conghua	LAT	0.806	0.778	0.836	0.849	0.873	0.747
	Region task-based	0.819	0.811	0.827	0.871	0.901	0.794
	Region-Boundary	0.824	0.819	0.829	0.878	0.907	0.797
	WS + Higra	0.829	0.817	0.841	0.881	0.902	0.809
	OWT + LM	0.756	0.749	0.763	0.796	0.849	0.744
	OWT + Higra	**0.838**	0.827	0.849	**0.885**	**0.911**	**0.816**
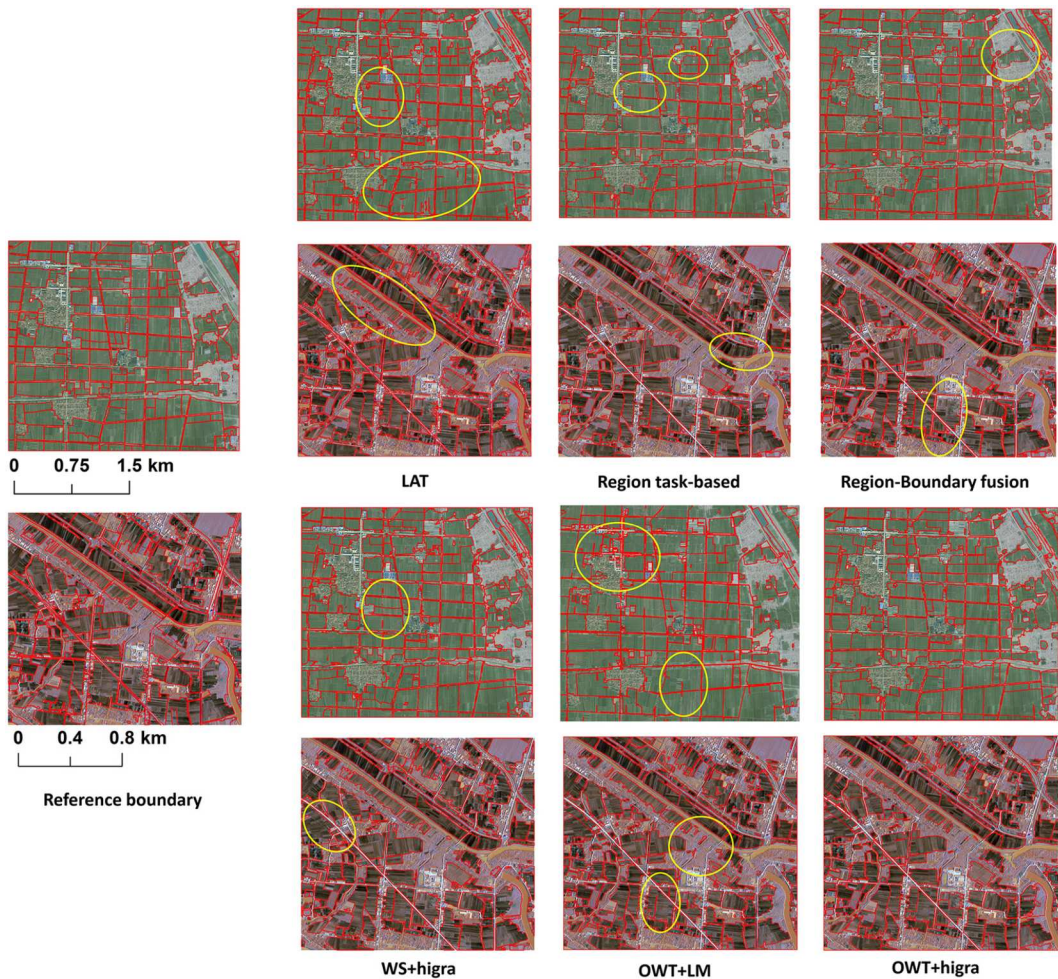
Figure 16. The results of different agricultural fields generation methods. The circles for regions with boundary breakage and misidentification.

(1) Comparison with threshold-based boundary binarization (LAT). The results of the local adaptive threshold (LAT) method exhibit some boundary breaks and noise within the field, as indicated by the yellow circle area. This occurs because the LAT method retains part of the detected linear signals after binarization. Since LAT operates on a pixel-by-pixel basis, it cannot eliminate all points on the 'fake boundaries' comprehensively. In contrast, the adopted region segmentation-based method identifies all potential boundary lines and subsequently determines whether to retain or remove the boundaries. This approach allows for the removal of fake boundaries as a whole or the preservation of entire boundaries when necessary, resulting in fields with continuous and internally pure boundaries.

(2) Comparison with Region task-based method. As depicted in Figure 16, the field results predicted by region task still exhibit noticeable internal boundary fragmentation or loss. These boundaries are completely preserved in the results obtained through the Region-Boundary Fusion method. This is also evident in terms of accuracy across all three study areas (Pingyuan, Funan, and Conghua) and for both boundary layers (detected by the modified MPSPNet and SEANet). The field results derived from the Region task-based method consistently yielded lower accuracies compared to the Region-Boundary Fusion method. This underscores the
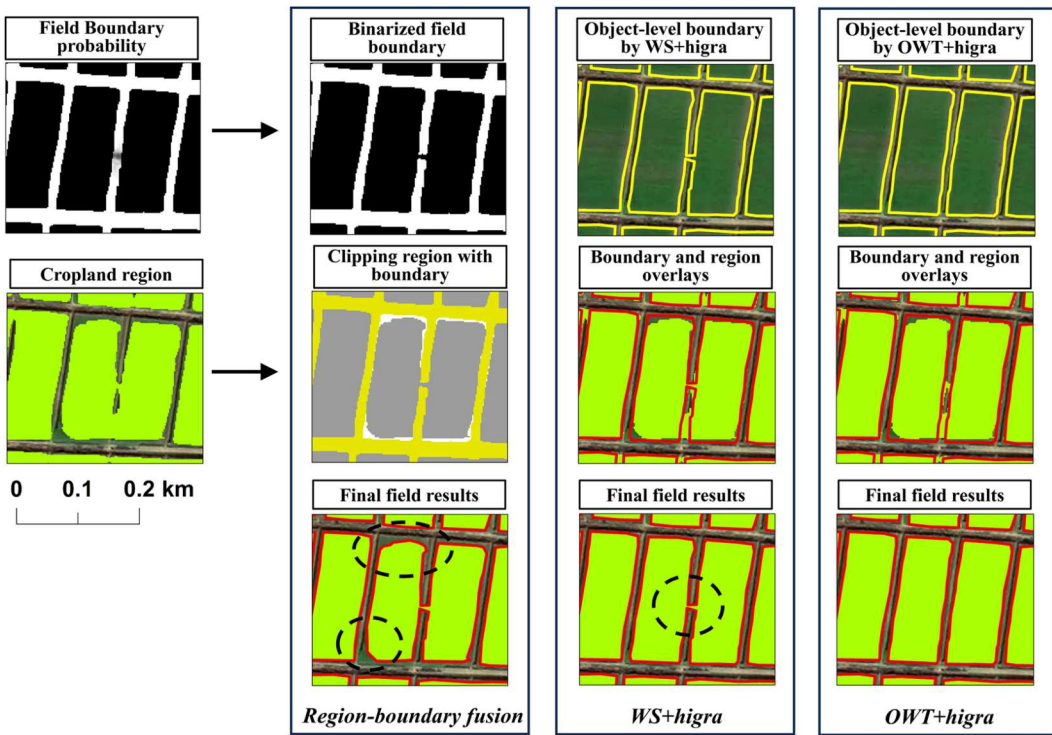
Figure 17. The generation process of Region-Boundary Fusion, WS + Higra and our OWT + Higra method. The black circles show obvious inaccurate boundary correspondence and field under-segmentation.

potential for enhancing result quality to a certain extent by optimizing the identified region results with finer and more complete boundary information.

(3) Comparison with Region-Boundary Fusion method. Across all three study areas and for both boundary layers, the Region-Boundary Fusion method exhibited lower boundary and geometric accuracies compared to the two segmentation-based methods (WS + Higra and OWT + Higra). This can be attributed to the method's implementation process. As depicted in Figure 17, the Region-Boundary Fusion method relies on identified cropland regions as a reference and integrates the boundary results to segment the fields. Consequently, the generated results are dependent on the quality of the identified regions. However, in comparison to the boundary detection task, the Region task has a weaker ability to precisely locate boundaries. This limitation makes it challenging to accurately represent the transformed boundary between cropland and non-cropland, resulting in difficulty achieving a complete correspondence between the generated fields and the real ones. Additionally, the shapes of the generated fields are not sufficiently regular. In contrast, in segmentation-based methods, the final field results are entirely derived from the detected boundary information, with region information primarily used to eliminate non-cultivated objects. Therefore, these methods produce more accurate and regular results with higher accuracies.

(4) Comparison with watershed segmentation (WS). The results obtained from watershed segmentation exhibit a noticeable issue of field under-segmentation. This problem arises because the boundary probabilities detected for pixels along the same boundary are not identical, leading to some weak gradient boundary parts being unsegmented when only considering gradient strength in segmentation. However, through the incorporation of orientation information using the OWT method, these weak boundaries can be effectively reinforced.
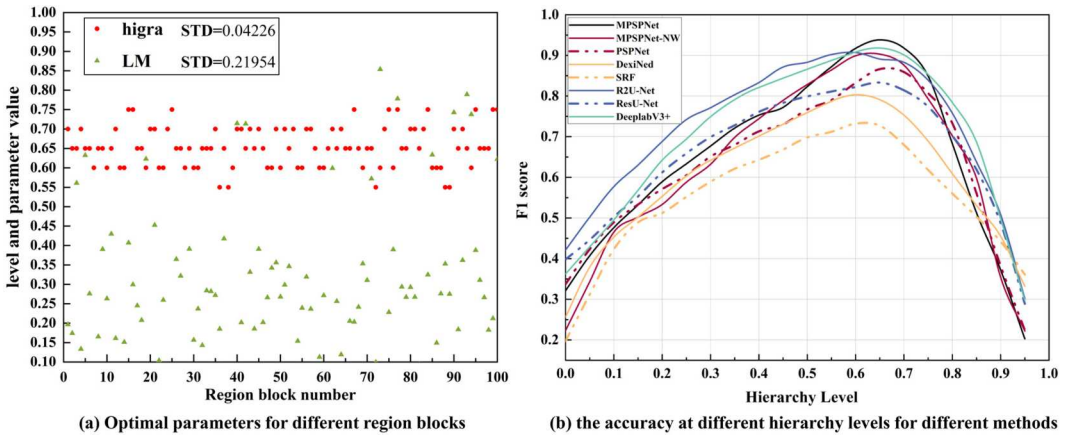
Figure 18. (a) Scatter plot of the optimal parameters for different region blocks, (b) The accuracy curve at different hierarchy levels for different methods.

This orientation-based approach functions as boundary repair, resulting in field boundaries that are continuous and more accurate across all three regions.

(5) Comparison with local merging-based method (LM). The local merging-based method is affected by the issue of suboptimal calculations of local parameters in certain regions, leading to noticeable boundary misclassifications and omissions in the results. Although both merging methods aim to remove common edges with weak strength for merging, the hierarchical merging method leverages the relative relationship of strengths at the scale of the entire image, thus possessing stronger generalization capabilities. Consequently, the challenge of locally selecting unstable parameters is transformed into a stable optimal hierarchical-level selection for the entire image. This transformation results in a reduction of local errors and an improvement in boundary accuracy, as measured by the F1 score, by 0.121, 0.087, and 0.074 in the three regions. These findings underscore the effectiveness of the hierarchical merging method.

(6) Effect of different detection results. The performance-enhancing effect of the developed OWT + Higra method is evident when applied to boundary layers detected by different models (the modified MPSPNet and SEANet), proving its generalizability. Moreover, a more powerful model can yield higher-quality field results. The SEANet + OWT + Higra method achieves the highest accuracy across all areas, indicating that this field generation method can serve as a generic framework and is compatible with stronger detection models, exhibiting extensibility. Additionally, it is worth noting that the accuracy enhancement effect of OWT + Higra diminishes with more powerful detection models. Compared to the region task-based method, OWT + Higra shows an average boundary F1 improvement of 0.04 on the MPSPNet layer and 0.02 on the SEANet boundary layer.

4.5. Stability validation of segmentation parameters

To further verify the threshold stability of the hierarchical merging method, two supplementary experiments were conducted. The first experiment involved dividing the Pingyuan County image into 100 smaller blocks and calculating the optimal hierarchical level that maximizes boundary accuracy for each block, as well as the optimal height parameter under the local merging method. As seen from Figure 18(a), the optimal hierarchy level for all region blocks of the hierarchical merging method is concentrated around 0.65, with a standard deviation of only 0.042, demonstrating greater stability compared to the local merging method with discrete data, which had a standard deviation of 0.21. This confirms the stability of the hierarchical merging method across different image localities.

The second experiment aimed to verify the threshold stability across different detection methods. As shown in Figure 18(b), all methods exhibit a pattern of initially increasing and then decreasing F1 scores as the level increases. The optimal level for all techniques is approximately 0.65, remaining unchanged across different detection methods. This is because, despite variations in the probability values detected by different methods, the relationship between boundary strengths remains constant. The hierarchical structure allows for relatively fixed global threshold across various detections in the same image. This method can be used to explore the range of optimal thresholds for different types of agricultural landscapes and establish a set of empirical threshold rules, potentially leading to significant improvement in detection efficiency.

5. Discussion

This study primarily focuses on two aspects of agricultural field delineation: field boundary detection and agricultural field generation. Specifically, we summarize three key techniques for the field boundary detection model. It's important to note that our contribution lies in providing a baseline reference for model design rather than developing a new detection model. Two common semantic segmentation models, MPSPNet and DeeplabV3+, achieves an area accuracy of over 0.9 in the plains in the plains after modified by the key techniques. Among the techniques, multitasking provides significant accuracy effect as it greatly improves the interpretability of field boundaries. Another crucial aspect was the balance between lowest-level detail features and high-level category features, which is often overlooked in current research but proved to be vital in our study. By optimizing the allocation of these two features in the model and reducing mid-level features, our model surpassed all existing U-Net-based models. This observation also partly explains why the current state-of-the-art SEANet outperforms the ResU-Net-a model. Furthermore, boundary-focused training also significantly influences detection performance, which can be achieved by assigning a larger misclassification penalty to field boundaries or employing some category-imbalance loss functions, but most existing studies have overlooked this technique (Crommelinck et al. 2019; Fetai, Račič, and Lisec 2021; H. Zhang et al. 2021; Taravat et al. 2021; Xia, Persello, and Koeva 2019; Yang et al. 2020).

Additional techniques may be available to further enhance model performance. For instance, replacing the commonly used VGG16 and ResNet backbones with more advanced self-attention transformer-based backbones could be considered. Incorporating more additional auxiliary tasks may provide more relevant task constraints. Moreover, some attention mechanisms-based modules and atrous convolutional layers can be integrated into the model to enhance feature description capabilities. Overall, within the proposed GF-AFD framework, the detection model component acts as an interface that is adaptable to future, more robust detection models.

This study combines the OWT segmentation method with a hierarchy-based merging method to convert detection results into object-level agricultural fields. In contrast to certain end-to-end methods, such as outputting the identified result of region task or refining identified cropland regions using detected boundaries, our method does involve parameter setting, which may introduce some inconvenience. Parameter selection is a necessary aspect of using segmentation methods to transform pixel-level boundary results into object-level fields. However, our method can produce more regular and accurate field delineations since the boundary results provide a finer depiction of field shapes compared to region results. Based on detected results with the currently most powerful SEANet model, employing our segmentation-based generation method still yields noticeable performance improvements. Therefore, a trade-off must be considered between delineation accuracy improvement and human involvement in real-world applications. In the future, if detection models can maintain continuous detected field boundaries and accurately correspond identified cropland regions to real cropland, the accuracy enhancement effect of our method will be weakened, and the parameter-less region-based methods may be more practical.

Some automatic parameters calculation methods deserve further exploration. For instance, Pareto optimization methods could offer the optimal balance between over- and under-

segmentation and potentially provide the best segmentation parameters. However, it's important to note that these optimization methods can be computationally intensive. In this study, we developed a hierarchy-based merging approach and demonstrated that the parameters of our merging method remained stable across a wide range of imagery and different detection models. Therefore, constructing a lookup table of empirical thresholds may be a more efficient and advantageous approach from an application standpoint, as it minimizes the need for manual intervention.

Due to the heterogeneous factors that must be considered, it is difficult to make a direct comparison with the relevant studies' reported accuracies. First, there were significant differences in the experimental setup among methods, such as image data (e.g. resolution, spectral bands, and sensors), landscape complexity, and training sets. Second, there is no standard method for assessing field boundary delineation accuracy. Therefore, the results reported in previous studies cannot be directly compared, as they used numerous statistical metrics, including mean absolute error (Meyer and Van Niekerk 2016; Watkins and van Niekerk 2019a; Watkins and Van Niekerk 2019b), F1 score (Crommelinck et al. 2019; Graesser and Ramankutty 2017; Masoud, Persello, and Tolpekin 2020; Wagner and Oppelt 2020; Yang et al. 2020), overall accuracy (Fetai, Račič, and Lisec 2021; Vlachopoulos et al. 2020; Waldner et al. 2021), boundary displacement error (Freixenet et al. 2002; Garcia-Pedrero et al. 2019), and the Jaccard Index (Taravat et al. 2021; Tetteh, Gocht, and Conrad 2020). Furthermore, a comprehensive accuracy assessment of the delineated field results should include both boundary and geometric levels. The area, location, and shape accuracies comprise the geometric accuracy evaluated in this study. In addition, the number of field vertices and the subdivision rate are also potential evaluation metrics that can be considered. Although the selected comparison models are representative in our study, there is always a limited number of models that can be compared through an experiment. This highlights the need for common data sets and a shared set of evaluation metrics to enable systematic benchmarking of methods.

In addition, only single-date image was used in this study, as high-resolution images are not easily available, yet their detailed information is more suitable for field delineation in China. It is recommended to use multi-temporal images in the further study when the images are sufficient, as they can better highlight some potential boundaries (Cheng et al. 2020). The choice of image resolution for the agricultural field delineation is also a potential research topic. Existing studies have adopted various spatial resolutions for delineation, including 30 m(Graesser and Ramankutty 2017), 10 m (Gopidas et al. 2021; H. Zhang et al. 2021), and ≤2 m (S. Liu et al. 2022; Persello et al. 2019). The choice of image resolution was mainly to provide a clear representation of cropland division. For instance, high-resolution images are typically applied in smallholder farming regions like China, while 10-meter images are commonly used in areas such as the United States and Australia. However, in cases where both 10-meter and 2-meter resolution images clearly show the boundaries, the 10-meter resolution image may result in better delineation performance because lower resolution can lead to smoother and more uniform textures within cropland areas. It is worth noting that no study has quantitatively assessed the impact of image resolution on delineation accuracy, indicating a potential avenue for future research.

6. Conclusions

This article presents a framework for agricultural field delineation, which summarizes and validates three key techniques for the field boundary detection model and addresses issues related to weak boundary loss and unstable parameters in the agricultural field generation process. The delineation framework demonstrates strong performance across three counties in China, surpassing comparison methods. Experimental results allowed us to conclude that: (1) The three key techniques for the detection model are multi-tasking for regions and boundaries, balancing lowest-level detail features with high-level category features, and boundary-focused training. These techniques were validated by modifying the MPSPNet model, whose ablation experiments validate the significant

performance improvement resulting from the techniques, and the comparison experiments validate the generalization of the techniques and the performance that surpasses the U-Net-based network and approaches the state-of-the-art model. (2) Performing region segmentation on detected probabilities yields more accurate field results compared to the method of clipping region results with detected boundaries. Additionally, the oriented watershed transformation provides more individual fields with continuous boundaries. (3) The developed hierarchical merge method transforms locally unstable segmentation thresholds into global stable thresholds, and yields more regular and accurate field results. In general, the proposed GF-AFD framework provides a valuable solution to the challenges of agricultural field delineation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Data availability statement

The high-resolution images are provided by China Centre for Resources Satellite Data and Application (https://www.cresda.com/zgzywxyyzx/index.html). Accounts with permissions are required to download high-resolution images on this site. The outlined reference data and code in this study are available by contacting the corresponding author.

References

Arbeláez, P., M. Maire, C. Fowlkes, and J. Malik. 2011. "Contour Detection and Hierarchical Image Segmentation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33:898–916. https://doi.org/10.1109/TPAMI.2010.161.

Belgiu, M., and O. Csillik. 2018. "Sentinel-2 Cropland Mapping Using Pixel-Based and Object-Based Time-Weighted Dynamic Time Warping Analysis." *Remote Sensing of Environment* 204:509–523. https://doi.org/10.1016/j.rse.2017.10.005.

Blaschke, T. 2010. "Object Based Image Analysis for Remote Sensing." *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (1): 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004.

Chen, L.-C., Y. Zhu, G. Papandreou, F. Schrof, and H. Adam. 2018. "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation." In *Eccv*, edited by V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, 833–851. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-01234-2_49.

Cheng, T., X. Ji, G. Yang, H. Zheng, J. Ma, X. Yao, Y. Zhu, and W. Cao. 2020. "DESTIN: A New Method for Delineating the Boundaries of Crop Fields by Fusing Spatial and Temporal Information from WorldView and Planet Satellite Imagery." *Computers and Electronics in Agriculture* 178:105787. https://doi.org/10.1016/j.compag.2020.105787.

Cousty, J., G. Bertrand, L. Najman, and M. Couprie. 2009. "Watershed Cuts: Minimum Spanning Forests and the Drop of Water Principle." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31:1362–1374. https://doi.org/10.1109/TPAMI.2008.173.

Crommelinck, S., M. Koeva, M. Y. Yang, and G. Vosselman. 2019. "Application of Deep Learning for Delineation of Visible Cadastral Boundaries from Remote Sensing Imagery." *Remote Sensing* 11 (21): 2505. https://doi.org/10.3390/rs11212505.

Fetai, B., M. Račič, and A. Lisec. 2021. "Deep Learning for Detection of Visible Land Boundaries from UAV Imagery." *Remote Sensing* 13 (11): 2077. https://doi.org/10.3390/rs13112077.

Freixenet, J., X. Muñoz, D. Raba, J. Martí, and X. Cufí. 2002. "Yet Another Survey on Image Segmentation: Region and Boundary Information Integration." In *Computer Vision – ECCV*, edited by A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, 408–422. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 2352. Berlin: Springer. https://doi.org/10.1007/3-540-47977-5_27.

García-Pedrero, A., C. Gonzalo-Martín, and M. Lillo-Saavedra. 2017. "A Machine Learning Approach for Agricultural Parcel Delineation Through Agglomerative Segmentation." *International Journal of Remote Sensing* 38 (7): 1809–1819. https://doi.org/10.1080/01431161.2016.1278312.

Garcia-Pedrero, A., M. Lillo-Saavedra, D. Rodriguez-Esparragon, and C. Gonzalo-Martin. 2019. "Deep Learning for Automatic Outlining Agricultural Parcels: Exploiting the Land Parcel Identification System." *IEEE Access* 7:158223–158236. https://doi.org/10.1109/ACCESS.2019.2950371.

Gopidas, D. K., P. D. Research, S. Narayana, G. College, K. Kchavadi, R. Priya, and A. P. Head. 2021. "Integrated Deep Learning Based Segmentation and Classification Method for Boundary Delineation of Agricultural Fields in Multitemporal Satellite Images." *International Journal of Modern Agriculture* 10 (2): 1804–1822.

Graesser, J., and N. Ramankutty. 2017. "Detection of Cropland Field Parcels from Landsat Imagery." *Remote Sensing of Environment* 201:165–180. https://doi.org/10.1016/j.rse.2017.08.027.

Hong, R., J. Park, S. Jang, H. Shin, H. Kim, and I. Song. 2021. "Development of a Parcel-Level Land Boundary Extraction Algorithm for Aerial Imagery of Regularly Arranged Agricultural Areas." *Remote Sensing* 13 (6): 1167. https://doi.org/10.3390/rs13061167.

Li, M., J. Long, A. Stein, and X. Wang. 2023. "Using a Semantic Edge-Aware Multi-Task Neural Network to Delineate Agricultural Parcels from Remote Sensing Images." *ISPRS Journal of Photogrammetry and Remote Sensing* 200:24–40. https://doi.org/10.1016/j.isprsjprs.2023.04.019.

Li, D., G. Zhang, Z. Wu, and L. Yi. 2010. "An Edge Embedded Marker-Based Watershed Algorithm for High Spatial Resolution Remote Sensing Image Segmentation." *IEEE Transactions on Image Processing* 19:2781–2787. https://doi.org/10.1109/TIP.2010.2049528.

Liu, Y., M. M. Cheng, X. Hu, K. Wang, and X. Bai. 2017. "Richer Convolutional Features for Edge Detection." In *Proceedings of 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 2017-Janua*, 5872–5881. Honolulu, HI: IEEE. https://doi.org/10.1109/CVPR.2017.622.

Liu, S., L. Liu, F. Xu, J. Chen, Y. Yuan, and X. Chen. 2022. "A Deep Learning Method for Individual Arable Field (IAF) Extraction with Cross-Domain Adversarial Capability." *Computers and Electronics in Agriculture* 203:107473. https://doi.org/10.1016/j.compag.2022.107473.

Long, J., M. Li, X. Wang, and A. Stein. 2022. "Delineation of Agricultural Fields Using Multi-Task BsiNet from High-Resolution Satellite Images." *International Journal of Applied Earth Observation and Geoinformation* 112:102871. https://doi.org/10.1016/j.jag.2022.102871.

Marvaniya, S., U. Devi, J. Hazra, S. Mujumdar, and N. Gupta. 2021. "Small, Sparse, but Substantial: Techniques for Segmenting Small Agricultural Fields Using Sparse Ground Data." *International Journal of Remote Sensing* 42 (4): 1512–1534. https://doi.org/10.1080/01431161.2020.1834166.

Masoud, K. M., C. Persello, and V. A. Tolpekin. 2020. "Delineation of Agricultural Field Boundaries from Sentinel-2 Images Using a Novel Super-Resolution Contour Detector Based on Fully Convolutional Networks." *Remote Sensing* 12 (1): 59. https://doi.org/10.3390/RS12010059.

Meyer, H. P., and A. Van Niekerk. 2016. "Assessing Edge and Area Metrics for Image Segmentation Parameter Tuning and Evaluation." In *Proceedings of GEOBIA 2016: Solutions and Synergies, 14–16 September 2016*, edited by N. Kerle, M. Gerke, and S. Lefevre. Enschede: Netherlands University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC). https://doi.org/10.3990/2.440.

Ming, D., X. Zhang, M. Wang, and W. Zhou. 2016. "Cropland Extraction Based on OBIA and Adaptive Scale Pre-Estimation." *Photogrammetric Engineering & Remote Sensing* 82:635–644. https://doi.org/10.14358/PERS.82.8.635.

Musat, G. A., M. Colezea, F. Pop, C. Negru, M. Mocanu, C. Esposito, and A. Castiglione. 2018. "Advanced Services for Efficient Management of Smart Farms." *Journal of Parallel and Distributed Computing* 116:3–17. https://doi.org/10.1016/j.jpdc.2017.10.017.

Persello, C., V. A. Tolpekin, J. R. Bergado, and R. A. de By. 2019. "Delineation of Agricultural Fields in Smallholder Farms from Satellite Images Using Fully Convolutional Networks and Combinatorial Grouping." *Remote Sensing of Environment* 231:111253. https://doi.org/10.1016/j.rse.2019.111253.

Soria, X., E. Riba, and A. Sappa. 2020. "Dense Extreme Inception Network: Towards a Robust CNN Model for Edge Detection." In *Proceedings of 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020*, 1912–1921. Snowmass, CO: IEEE. https://doi.org/10.1109/WACV45572.2020.9093290.

Taravat, A., M. P. Wagner, R. Bonifacio, and D. Petit. 2021. "Advanced Fully Convolutional Networks for Agricultural Field Boundary Detection." *Remote Sensing* 13:1–12. https://doi.org/10.3390/rs13040722.

Tetteh, G. O., A. Gocht, and C. Conrad. 2020. "Optimal Parameters for Delineating Agricultural Parcels from Satellite Images Based on Supervised Bayesian Optimization." *Computers and Electronics in Agriculture* 178:105696. https://doi.org/10.1016/j.compag.2020.105696.

Turker, M., and E. H. Kok. 2013. "Field-Based Sub-Boundary Extraction from Remote Sensing Imagery Using Perceptual Grouping." *ISPRS Journal of Photogrammetry and Remote Sensing* 79:106–121. https://doi.org/10.1016/j.isprsjprs.2013.02.009.

Vlachopoulos, O., B. Leblon, J. Wang, A. Haddadi, A. LaRocque, and G. Patterson. 2020. "Delineation of Crop Field Areas and Boundaries from UAS Imagery Using PBIA and GEOBIA with Random Forest Classification." *Remote Sensing* 12 (16): 2640. https://doi.org/10.3390/RS12162640.

Wagner, M. P., and N. Oppelt. 2020. "Deep Learning and Adaptive Graph-Based Growing Contours for Agricultural Field Extraction." *Remote Sensing* 12 (12): 1990. https://doi.org/10.3390/rs12121990.

Waldner, F., and F. I. Diakogiannis. 2020. "Deep Learning on Edge: Extracting Field Boundaries from Satellite Images with a Convolutional Neural Network." *Remote Sensing of Environment* 245:111741. https://doi.org/10.1016/j.rse.2020.111741.

Waldner, F., F. I. Diakogiannis, K. Batchelor, M. Ciccotosto-Camp, E. Cooper-Williams, C. Herrmann, G. Mata, and A. Toovey. 2021. "Detect, Consolidate, Delineate: Scalable Mapping of Field Boundaries Using Satellite Images." *Remote Sensing* 13 (11): 2197. https://doi.org/10.3390/rs13112197.

Watkins, B., and A. van Niekerk. 2019a. "A Comparison of Object-Based Image Analysis Approaches for Field Boundary Delineation Using Multi-Temporal Sentinel-2 Imagery." *Computers and Electronics in Agriculture* 158:294–302. https://doi.org/10.1016/j.compag.2019.02.009.

Watkins, B., and A. Van Niekerk. 2019b. "Automating Field Boundary Delineation with Multi-Temporal Sentinel-2 Imagery." *Computers and Electronics in Agriculture* 167:105078. https://doi.org/10.1016/j.compag.2019.105078.

Xia, X., C. Persello, and M. Koeva. 2019. "Deep Fully Convolutional Networks for Cadastral Boundary Detection from UAV Images." *Remote Sensing* 11 (14): 1725. https://doi.org/10.3390/rs11141725.

Xu, L., D. Ming, T. Du, Y. Chen, D. Dong, and C. Zhou. 2022. "Delineation of Cultivated Land Parcels Based on Deep Convolutional Networks and Geographical Thematic Scene Division of Remotely Sensed Images." *Computers and Electronics in Agriculture* 192:106611. https://doi.org/10.1016/j.compag.2021.106611.

Yan, L., and D. P. Roy. 2014. "Automated Crop Field Extraction from Multi-Temporal Web Enabled Landsat Data." *Remote Sensing of Environment* 144:42–64. https://doi.org/10.1016/j.rse.2014.01.006.

Yan, L., and D. P. Roy. 2016. "Conterminous United States Crop Field Size Quantification from Multi-Temporal Landsat Data." *Remote Sensing of Environment* 172:67–86. https://doi.org/10.1016/j.rse.2015.10.034.

Yang, R., Z. U. Ahmed, U. C. Schulthess, M. Kamal, and R. Rai. 2020. "Detecting Functional Field Units from Satellite Images in Smallholder Farming Systems Using a Deep Learning Based Computer Vision Approach: A Case Study from Bangladesh." *Remote Sensing Applications: Society and Environment* 20:100413. https://doi.org/10.1016/j.rsase.2020.100413.

Zhang, H., M. Liu, Y. Wang, J. Shang, X. Liu, B. Li, A. Song, and Q. Li. 2021. "Automated Delineation of Agricultural Field Boundaries from Sentinel-2 Images Using Recurrent Residual U-Net." *International Journal of Applied Earth Observation and Geoinformation* 105:102557. https://doi.org/10.1016/j.jag.2021.102557.

Zhang, D., Y. Pan, J. Zhang, T. Hu, J. Zhao, N. Li, and Q. Chen. 2020. "A Generalized Approach Based on Convolutional Neural Networks for Large Area Cropland Mapping at Very High Resolution." *Remote Sensing of Environment* 247:111912. https://doi.org/10.1016/j.rse.2020.111912.

Zhao, H., J. Shi, X. Qi, X. Wang, and J. Jia. 2017. "Pyramid Scene Parsing Network." In *Proceedings of 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 2017-Janua*, 6230–6239. Honolulu, HI: IEEE. https://doi.org/10.1109/CVPR.2017.660.