forests

MDPI

Article

High-Resolution Mapping of Farmland Shelterbelts in an Oasis Agricultural Region Using GF-2 Imagery and Semantic Segmentation

Yingqi Xu [1], Ping Lv [2], Zhuo Zhang [2], Lanjie Li [2], Zheng Chai [2], Yuanyuan Li [1,*] and Cheng Tang [1,*]

[1] College of Urban and Environmental Sciences, Shihezi University, Shihezi 832003, China; yingqi160@gmail.com

[2] Forestry and Grassland Workstation of Xinjiang Production and Construction Corps, Urumqi 830013, China

* Correspondence: 2004liyy@163.com (Y.L.); tangcheng1983@163.com (C.T.); Tel.: +86-18935705331 (Y.L.); +86-15352608029 (C.T.)

Abstract

Farmland shelterbelts are important linear vegetation infrastructures in oasis agricultural landscapes. Their accurate extraction is essential for shelterbelt inventory and farmland management, but remains challenging because shelterbelts are narrow, elongated, locally discontinuous, and spectrally similar to croplands, orchards, roadside vegetation, bare soil, and irrigation-related features. This study developed a GF-2-based deep learning workflow for farmland shelterbelt extraction in the 11th Regiment of Alar City, Xinjiang, China. Four representative semantic segmentation models, namely U-Net, U-Net with scSE attention, U-Net++, and DeepLabV3+, were trained and evaluated using four-band GF-2 optical imagery under a unified experimental setting. Model performance was assessed using Precision, Recall, $F1$-score, Intersection over Union (IoU), overall accuracy, and Kappa coefficient. Patch-level statistical comparison and visual interpretation were further conducted to examine performance differences, shelterbelt continuity, boundary integrity, omission errors, and background confusion. The results showed that U-Net achieved the best overall performance, with a Precision of 94.58%, Recall of 94.77%, $F1$-score of 94.67%, IoU of 89.88%, overall accuracy of 99.71%, and Kappa coefficient of 0.9452. Compared with U-Net with scSE attention, U-Net++, and DeepLabV3+, U-Net better preserved the continuity and boundary integrity of narrow shelterbelts in regular field-boundary networks. The other models showed varying degrees of omission, boundary fragmentation, or confusion with spectrally similar agricultural objects. The best-performing U-Net model was then applied to the complete study area, and the extracted shelterbelt area was approximately 6.5 km^2, accounting for about 4.37% of the cultivated land area. These results indicate that GF-2 optical imagery combined with semantic segmentation can support fine-scale farmland shelterbelt mapping in oasis agricultural landscapes. They also show that model evaluation for narrow linear vegetation features should consider not only pixel-level accuracy but also spatial continuity, boundary integrity, and typical error patterns. The proposed workflow provides a practical reference for GF-2-based farmland shelterbelt inventory, high-resolution linear vegetation mapping, and shelterbelt monitoring in arid oasis agricultural landscapes.

Keywords: farmland shelterbelt; GF-2; semantic segmentation; U-Net; oasis agriculture

Academic Editor: Michael Sprintsin

Received: 2 July 2026
Revised: 13 August 2026
Accepted: 17 August 2026
Published: 19 August 2026

1. Introduction

Farmland shelterbelts are an important component of green infrastructure in oasis agricultural landscapes. In arid and semi-arid regions, cultivated land is frequently exposed to wind erosion, sand transport, soil moisture loss, and limited water availability. By reducing near-surface wind speed and modifying local microenvironmental conditions, shelterbelt and windbreak systems can help protect cropland, conserve soil moisture, and maintain the ecological stability of agricultural landscapes [1–4]. Reliable information on their spatial distribution is therefore essential for shelterbelt inventory, farmland management, and ecological protection.

Remote sensing provides an efficient means of obtaining such information over large agricultural areas, but narrow woody features remain difficult to delineate consistently. Their elongated geometry, limited width, local discontinuity, and close association with other agricultural vegetation complicate both visual interpretation and automated extraction. Similar challenges have been reported in hedgerow monitoring, where the detection of linear woody vegetation is strongly influenced by spatial resolution and landscape heterogeneity [5]. High spatial resolution is therefore particularly important for shelterbelt mapping. The GF-2 PMS sensor provides a 1 m panchromatic band and four 4 m multispectral bands covering the blue, green, red, and near-infrared spectral regions. Through pan-sharpening, the spatial detail of the panchromatic image can be integrated with multispectral information to generate a four-band image at 1 m resolution, thereby combining fine spatial detail with spectral information relevant to vegetation discrimination [6].

Earlier shelterbelt-mapping studies relied mainly on manually designed spectral, textural, morphological, and vegetation-index features, typically combined with conventional classification or object-based image analysis. Deng et al. [7] developed a belt-oriented extraction strategy incorporating classification and morphological processing, whereas Li et al. [8] investigated shelterbelt extraction from ZY1-02E multispectral imagery. Zhang et al. [9] introduced a vegetation index specifically designed for farmland shelterbelt identification. More recent work has extended farmland shelterbelt extraction toward deep learning-based multi-feature fusion [10]. Earlier feature-based approaches have provided useful information on shelterbelt distribution, but their performance can depend substantially on feature construction, sample characteristics, classifier configuration, and parameter selection [11,12]. Such dependence becomes particularly limiting when shelterbelts are fragmented or when their spectral and spatial characteristics overlap with heterogeneous agricultural backgrounds, including cropland and other woody vegetation.

Deep learning offers an alternative strategy by learning discriminative representations directly from labeled imagery rather than relying primarily on handcrafted features. Convolutional neural networks and related deep learning methods are now widely used in Earth observation for land-cover mapping, vegetation analysis, and fine-scale image segmentation [13–15]. In forest remote sensing, high-resolution UAV imagery combined with deep learning has enabled detailed tree-species discrimination and vegetation mapping in heterogeneous environments [16,17]. Comparable advances have been made in agricultural remote sensing, where semantic segmentation has been applied to field-boundary delineation and high-resolution farmland mapping [18–20]. These applications are particularly relevant to shelterbelt extraction because successful mapping requires not only accurate class discrimination but also precise boundary delineation and preservation of local linear continuity.

Deep learning has also begun to be applied specifically to farmland shelterbelt mapping. Wang et al. [21] developed an improved prototypical network using Sentinel-2 imagery for shelterbelt classification, whereas Zhou et al. [22] compared several deep learning architectures for extracting linear shelterbelt forests from high-resolution imagery.

These studies demonstrate the potential of learned feature representations for shelterbelt mapping while also indicating that extraction performance is influenced by both image characteristics and network design. This relationship is particularly relevant to shelterbelts because their narrow and elongated geometry makes accurate delineation sensitive to local omissions, discontinuities, and boundary errors.

The model configurations considered in this study represent different strategies for preserving spatial detail and incorporating contextual information during semantic segmentation. U-Net transfers encoder features to corresponding decoder stages through skip connections, facilitating the recovery of fine spatial information during reconstruction [23]. The value of multi-scale feature integration has also been demonstrated in remote-sensing segmentation [24]. Building on the U-Net framework, the spatial and channel squeeze-and-excitation (scSE) mechanism recalibrates feature responses along both spatial and channel dimensions [25], whereas U-Net++ introduces nested intermediate pathways that progressively integrate features from different encoder and decoder levels [26]. DeepLabV3+, in contrast, combines an encoder–decoder structure with atrous convolution to capture contextual information over multiple receptive fields [27]. These differences in model design are relevant to shelterbelt mapping because the task requires both precise localization of narrow linear features and sufficient contextual information to distinguish them from spectrally or structurally similar agricultural objects.

Despite these advances, the relative suitability of different segmentation configurations for farmland shelterbelt mapping using high-resolution GF-2 imagery remains unclear. Existing shelterbelt studies have used different image sources and extraction strategies [7,8]. Other studies have focused on vegetation-index development and multi-feature fusion for shelterbelt extraction [9,10]. More recent deep-learning studies have also differed in network design, sample construction, and evaluation settings [21,22]. Taken together, these methodological differences mean that variations in reported accuracy cannot be attributed confidently to model design alone, because they may also reflect differences in data and experimental conditions. This issue is particularly relevant in oasis agricultural landscapes, where shelterbelts occur as narrow and locally discontinuous structures embedded within heterogeneous agricultural backgrounds. A controlled comparison within a common experimental framework is therefore needed to assess the relative performance of alternative segmentation strategies for fine-scale farmland shelterbelt mapping.

Accordingly, this study evaluates four semantic segmentation model configurations—U-Net, U-Net with scSE attention, U-Net++, and DeepLabV3+—using 1 m four-band GF-2 imagery from the 11th Regiment of Alar City, Xinjiang, China. The objectives are to: (1) establish a consistent GF-2-based experimental setting for farmland shelterbelt segmentation; (2) compare the performance of the four model configurations for narrow linear shelterbelt extraction; (3) evaluate differences among the models using aggregate accuracy metrics, patch-level statistical analysis, and visual assessment; and (4) apply the best-performing model to map shelterbelt distribution across the study area. By evaluating the four models within a common experimental framework, this study aims to clarify their relative suitability for fine-scale farmland shelterbelt mapping and to provide a practical approach for high-resolution shelterbelt extraction.

2. Materials

2.1. Study Area

The study area is the 11th Regiment of Alar City, Xinjiang Uygur Autonomous Region, China, with an approximate central location of 40°37′35.7″ N and 81°33′37.1″ E (Figure 1). It is located in the eastern part of the First Division of the Xinjiang Production and Construction Corps and belongs to the oasis agricultural zone along the northern margin of the

Tarim Basin. The area is close to the Taklimakan Desert and is influenced by typical arid inland environmental conditions. This regional background is closely associated with the ecological construction of shelterbelt systems in northern China [3].

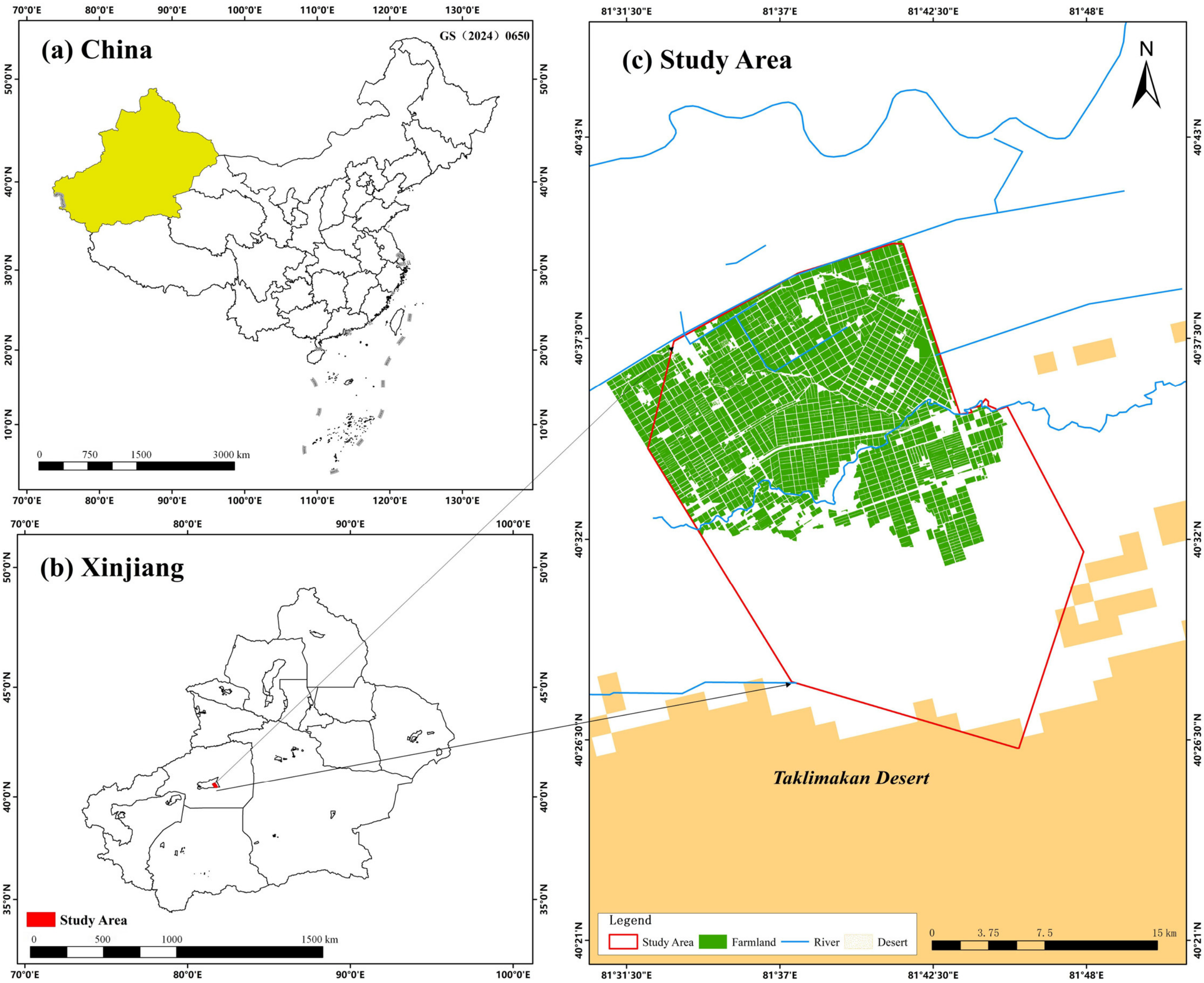

Figure 1. Geographic location and environmental context of the study area. (**a**) Location of Xinjiang within China; (**b**) location of the 11th Regiment within Xinjiang; and (**c**) detailed landscape characteristics of the study area.

The region experiences a warm-temperate continental arid climate characterized by limited precipitation, strong evaporation, and frequent wind–sand disturbance. Under these conditions, farmland shelterbelts are important for reducing wind effects, protecting cropland, and maintaining the stability of oasis agricultural landscapes [1,3,4]. Across arid and semi-arid northwestern China, vegetation configuration is strongly constrained by water availability, habitat conditions, and plant adaptation, which also influences the spatial organization of green infrastructure [28].

The local landscape is dominated by irrigated cropland, regular field blocks, rural roads, irrigation canals, and artificial shelterbelt networks. Shelterbelts are mainly composed of drought-tolerant tree and shrub species, such as *Populus euphratica*, *Populus alba var*. *pyramidalis*, and *Tamarix ramosissima*. Their distribution is closely associated with field boundaries, irrigation canals, field ditches, and rural roads. Species used in farmland plantings across southern Xinjiang oases reflect adaptation to arid agricultural conditions [29].

From a remote sensing perspective, the 11th Regiment provides a challenging setting for farmland shelterbelt mapping. Field surveys conducted in the 2nd and 7th Companies of the 11th Regiment showed that farmland shelterbelts are typically about 10 m wide, with narrower belts measuring approximately 4–5 m and wider belts reaching up to approximately 30 m. At the 1 m resolution achieved after pan-sharpening, the observed shelterbelt widths correspond to roughly 4–30 image pixels, with a typical belt spanning about 10 pixels. The elongated and locally discontinuous geometry of these shelterbelts, together with the presence of surrounding croplands, orchards, roadside vegetation, bare soil, canal-side vegetation, and other agricultural features with similar spectral responses or comparable linear structures, makes accurate delineation particularly challenging. The study area therefore provides an appropriate setting for evaluating whether semantic segmentation models can extract farmland shelterbelts while preserving their spatial continuity and boundary integrity.

2.2. Remote Sensing Data Sources

GF-2 PMS2 imagery acquired on 23 June 2023 served as the primary remote sensing dataset. The imagery was obtained from the China Centre for Resources Satellite Data and Application (CRESDA, Beijing, China) on 16 January 2025. To fully cover the 11th Regiment study area, two adjacent scenes acquired on the same date were selected. The two scenes had a cloud cover of approximately 1%, with a sun elevation angle of 30.24° and a sun azimuth angle of 116.03°, providing suitable imaging conditions for high-resolution agricultural interpretation.

The GF-2 PMS sensor records panchromatic data at approximately 1 m resolution and multispectral data at approximately 4 m resolution in the blue, green, red, and near-infrared bands. GF-2 imagery has been widely used in high-resolution optical earth observation and vegetation-related land-surface feature extraction because of its fine spatial resolution and multispectral capability [6].

Panchromatic and multispectral data were subsequently fused to obtain a four-band product at 1 m spatial resolution. All four fused bands were retained as model inputs. This design differs from RGB-only inputs because the near-infrared band provides vegetation-sensitive information that is useful for distinguishing shelterbelts from croplands, orchards, roadside vegetation, and canal-side vegetation. The acquisition date corresponded to the main growing season of the oasis agricultural landscape, when vegetation signals were active and spectral similarity among agricultural vegetation types was high. Therefore, the selected imagery provided a realistic and challenging condition for evaluating farmland shelterbelt extraction from high-resolution optical imagery.

2.3. Deep Learning Model Architectures

Four architectures were chosen to represent distinct mechanisms relevant to farmland shelterbelt delineation: U-Net as the baseline encoder–decoder model, U-Net with scSE for spatial and channel recalibration, U-Net++ for progressive multi-level feature integration, and DeepLabV3+ for atrous-convolution-based contextual modeling. This selection enabled a controlled comparison of the performance of different model configurations for extracting narrow linear shelterbelts.

U-Net with scSE attention was selected to examine whether spatial and channel-wise feature recalibration could improve the discrimination of shelterbelts from complex agricultural backgrounds [25]. U-Net++ was used to test whether its intermediate decoder pathways could improve boundary representation by progressively reconciling features from different network depths [26]. DeepLabV3+ was included to test whether atrous convolution and multi-scale contextual aggregation could benefit narrow shelterbelt segmentation [27].

All four architectures were evaluated within the same experimental setup, with identical GF-2 bands, data partition, optimization procedure, inference threshold, and evaluation metrics. This design allowed performance differences to be interpreted primarily in relation to differences among the model configurations rather than variations in the experimental setup.

3. Methods

The experimental design was organized into five linked stages. First, the GF-2 imagery was preprocessed to produce a 1 m four-band mosaic covering the study area. Second, a binary shelterbelt reference mask was paired with the imagery to construct the dataset used for model development. Third, the four segmentation models were trained under a common experimental protocol. Fourth, model performance was compared using pixel-level accuracy measures, patch-level statistical analysis, and visual assessment. Finally, the selected model was applied to the complete GF-2 mosaic to generate the farmland shelterbelt map for the study area.

3.1. Remote Sensing Data Preprocessing

Raw GF-2 imagery was processed into analysis-ready data through a series of preprocessing procedures before farmland shelterbelt extraction. The preprocessing procedures were conducted in ENVI 5.6 (NV5 Geospatial Software, Broomfield, CO, USA) and comprised radiometric calibration, atmospheric correction, orthorectification, mosaicking of the two GF-2 scenes, pan-sharpening, coordinate conversion, and spatial clipping according to the boundary of the 11th Regiment, as illustrated in Figure 2.

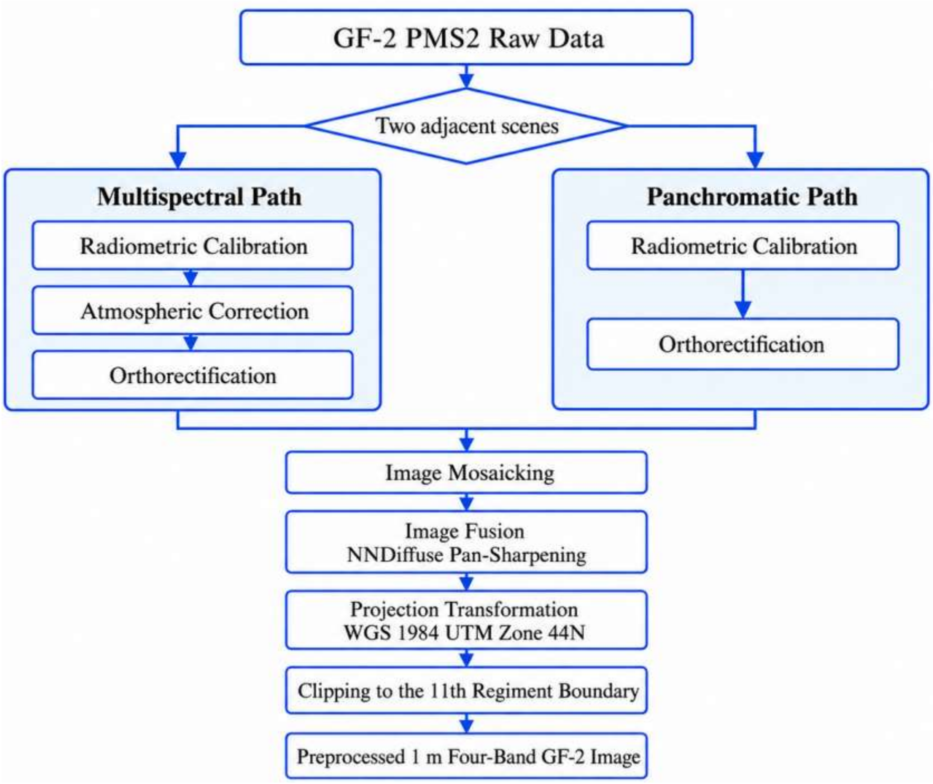

Figure 2. Workflow of GF-2 remote sensing data preprocessing.

The original digital numbers of the GF-2 panchromatic and multispectral bands were converted into radiance values using the radiometric calibration module in ENVI and the calibration parameters provided in the image metadata. The multispectral imagery was atmospherically corrected using the FLAASH module in ENVI to obtain surface reflectance data. Considering the summer acquisition period and the arid oasis conditions of the study region, the FLAASH parameters were configured with the Mid-Latitude Summer atmospheric model, the Rural aerosol model, and an initial visibility of 40 km.

Following radiometric and atmospheric correction, the GF-2 imagery was orthorectified using the image-associated rational polynomial coefficients (RPCs) together with a 30 m digital elevation model. The panchromatic and multispectral datasets were orthorectified independently to reduce geometric discrepancies and improve their spatial alignment with the reference shelterbelt mask. The two orthorectified GF-2 scenes acquired on 23 June 2023 were subsequently mosaicked into a single image covering the entire 11th Regiment.

The orthorectified GF-2 data were pan-sharpened with the NNDiffuse algorithm to enhance the spatial resolution of the multispectral imagery. In this process, the 1 m panchromatic band contributed the fine spatial detail, while the 4 m multispectral bands retained the spectral characteristics of the scene. The fused output comprised blue, green, red, and near-infrared bands at a spatial resolution of 1 m, providing finer delineation of narrow shelterbelts while preserving spectral information useful for distinguishing vegetation features.

Following pan-sharpening, the fused image was reprojected to WGS 1984 UTM Zone 44N and clipped to the boundary of the 11th Regiment. The resulting 1 m four-band GF-2 image provided the source data for patch generation, model training and validation, and subsequent shelterbelt mapping across the study area.

3.2. Dataset Construction

To reflect the challenges posed by narrow shelterbelts and heterogeneous agricultural surroundings, the dataset was constructed to represent both shelterbelt features and potentially confounding background objects. Image acquisition took place during the main growing season, when shelterbelts, croplands, orchards, and vegetation along roads and irrigation canals were all in active growth. Because most vegetation types were actively growing at the time of image acquisition, their spectral responses became more alike, making reliable separation among classes more difficult.

Farmland shelterbelts were represented using a binary reference mask constructed specifically for the segmentation task. Within the reference mask, pixels corresponding to shelterbelts were coded as 1, whereas all remaining pixels were coded as 0. The background class included cropland, orchards, bare soil, roads, irrigation canals, built-up surfaces, and other non-shelterbelt objects. The reference mask was manually delineated through visual interpretation of the pan-sharpened GF-2 image. Reference delineation was informed by false-color composites, local shelterbelt distribution patterns, partial field observations, and high-resolution imagery available in Google Earth (Google LLC, Mountain View, CA, USA; accessed on 5 and 21–23 May 2026). Particular attention was paid to linear vegetation features along field boundaries, rural roads, irrigation canals, and farmland edges. Linear woody vegetation belts along field boundaries, irrigation canals, roads, and farmland margins were labeled as shelterbelts when they showed continuous tree or shrub canopy patterns and were spatially associated with farmland protection structures. Orchard blocks, crop rows, isolated trees, and non-woody canal-side vegetation were excluded from the shelterbelt class. The preliminary labels were checked patch by patch by overlaying the GF-2 false-color composites, Google Earth images, and local field background information. After initial delineation, the mask was carefully checked and corrected to improve the spatial correspondence between shelterbelt labels and GF-2 image features. Before sample generation, the corrected reference mask was checked against the preprocessed GF-2 image to ensure pixel-level spatial correspondence between the image and label data.

The 1 m four-band GF-2 image and the corresponding binary mask were synchronously cropped using a non-overlapping sliding-window strategy. A 256 × 256-pixel window was paired with a 256-pixel stride, resulting in non-overlapping adjacent patches. This patch size was selected to balance local spatial detail and contextual information, allowing the models to capture narrow shelterbelt segments as well as surrounding agricultural background features.

Invalid patches with a high proportion of black or no-data pixels were removed using a threshold of 50%. Patches containing shelterbelt pixels were retained, whereas pure-background patches were randomly retained at a proportion of 5%. This filtering strategy reduced redundant background samples while preserving representative negative examples, such as croplands, orchards, roads, irrigation canals, and other linear or vegetation-covered objects that may be confused with shelterbelts.

The patch samples were randomly divided into training and validation subsets at an 80:20 ratio, yielding 5224 and 1307 patches, respectively. Because the stride equaled the patch width, adjacent patches did not overlap, thereby preventing direct pixel sharing between samples. Nevertheless, the training and validation subsets were derived from the same GF-2 image coverage, and spatial dependence between geographically proximate patches cannot be ruled out. The validation results are therefore interpreted primarily as estimates of relative model performance within the same scene, rather than as an assessment of transferability to independent scenes or geographic regions. Residual spatial dependence may still influence the absolute values of the reported accuracy metrics.

3.3. Experimental Environment and Parameter Settings

The capacity of deep semantic segmentation models to learn multilevel spatial and spectral features directly from imagery has enabled their application to remote-sensing image analysis and vegetation mapping [13–15]. To compare the extraction ability of different segmentation architectures for narrow linear farmland shelterbelts, all models were trained and evaluated under a unified experimental protocol. To support a controlled comparison, the data partition, spectral inputs, image dimensions, optimization scheme, loss formulation, and decision threshold were kept unchanged across all models.

Each model received 256 × 256-pixel image patches comprising the same four GF-2 spectral bands: blue, green, red, and near-infrared. ResNet34 was used as the encoder for U-Net, U-Net with scSE attention, and U-Net++, whereas DeepLabV3+ used ResNet50. All experiments were conducted in Google Colab (Google LLC, Mountain View, CA, USA; accessed on 20 June 2026) using an NVIDIA Tesla T4 GPU (NVIDIA Corporation, Santa Clara, CA, USA), without geometric or radiometric data augmentation. The same training and validation subsets, input size, optimizer, learning rate, loss function, thresholding rule, and model-selection procedure were used for all four architectures, and the corresponding training configuration is summarized in Table 1. Hyperparameters were selected through preliminary experiments based on training stability and validation performance. Under the final configuration, each model was trained for up to 80 epochs using the Adam optimizer, with a batch size of 8 and an initial learning rate of 1×10^{-4}. Validation loss and *IoU* were monitored throughout training to assess convergence and segmentation performance. Early stopping was applied when validation performance ceased to improve, and model checkpointing was used to retain the weights associated with the best observed validation performance. The retained checkpoint was subsequently used for inference and performance evaluation.

During inference, sigmoid outputs were converted to binary shelterbelt masks using a fixed probability threshold of 0.4. The threshold was selected empirically from the validation-set predictions after examining a range of candidate values and was then applied uniformly to all four architectures. This common decision rule avoided model-specific threshold adjustment and maintained comparability among models. Because the validation subset was also used for model monitoring, checkpoint selection, and threshold determination, the reported metrics represent validation-set performance rather than performance on an independent test set.

Table 1. Training configuration used for the four segmentation models.

Parameter	Value
Input size	256×256
Batch size	8
Maximum epochs	80
Optimizer	Adam
Initial learning rate	1×10^{-4}
Loss function	BCE-Dice loss
Output activation	Sigmoid
Model selection	Best validation checkpoint

3.4. Loss Function

Farmland shelterbelt extraction is a binary semantic segmentation task that requires distinguishing shelterbelt pixels from complex agricultural backgrounds at the pixel level. In GF-2 imagery, shelterbelt pixels usually account for a relatively small proportion of each patch, whereas background pixels, such as cropland, orchards, bare soil, roads, irrigation canals, and other non-shelterbelt objects, dominate the image. This foreground–background imbalance may cause the model to overemphasize background regions during training and reduce its sensitivity to narrow shelterbelt pixels.

To address the foreground–background imbalance, optimization was based on a composite objective combining binary cross-entropy with Dice loss. Compound loss functions are commonly used in semantic segmentation because they can jointly consider pixel-wise classification and region-level overlap, especially under class-imbalanced conditions [30,31]. Binary cross-entropy penalizes errors at the pixel level, whereas Dice loss emphasizes agreement between the predicted and reference masks. Their weighted combination is expressed as follows:

$$L = w_b L_{BCE} + w_d L_{Dice} \tag{1}$$

$$L_{BCE} = -1 \sum_{i=1}^{N} [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \tag{2}$$

$$L_{Dice} = 1 - \frac{2\sum_{i=1}^{N} y_i p_i + \epsilon}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} p_i + \epsilon} \tag{3}$$

Here, w_b and w_d denote the weights assigned to the BCE and Dice terms, respectively; N is the number of pixels considered in the loss calculation; for pixel i, $\beta y_i \in \{0, 1\}$ is the reference label, with 1 representing shelterbelt and 0 representing background, and $p_i \in [0, 1]$ is the predicted probability of belonging to the shelterbelt class. The term ε prevents zero denominators and maintains numerical stability during loss computation. Equal weights were assigned to the two loss terms ($w_b = w_d = 1$).

3.5. Accuracy Assessment

Performance was quantified using six measures: Precision, Recall, $F1$-score, Intersection over Union (IoU), overall accuracy (OA), and the Kappa coefficient. Their calculation was based on the four entries of the binary confusion matrix—true positives (*TP*), false positives (*FP*), true negatives (*TN*), and false negatives (*FN*). Specifically, *TP* represents correctly detected shelterbelt pixels; *FP* corresponds to background pixels erroneously labeled as shelterbelts; *FN* represents shelterbelt pixels omitted by the prediction; and *TN* corresponds to background pixels correctly retained as background.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{6}$$

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

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$Kappa = \frac{p_o - p_e}{1 - p_e} \tag{9}$$

In the Kappa calculation, p_o represents the observed agreement and p_e the agreement expected by chance; *TP*, *FP*, *TN*, and *FN* retain the definitions given above.

3.6. Statistical and Visual Comparison of Model Performance

To complement the aggregate accuracy metrics, patch-wise $F1$-score and *IoU* were calculated for validation patches containing shelterbelt pixels in the reference mask. Because the same patches were evaluated for all four models, the resulting scores were paired by patch. Differences among the models were assessed separately for $F1$-score and *IoU* using the Friedman test. When the omnibus test was significant, pairwise comparisons were conducted using Wilcoxon signed-rank tests with Bonferroni correction, with the family-wise significance level set at $\alpha = 0.05$. Given the validation design described in Section 3.2, the statistical analysis was used to compare relative model performance within the study scene rather than to assess geographic generalization.

Visual assessment was conducted on selected validation patches by comparing the GF-2 imagery with the corresponding model predictions. The assessment focused on shelterbelt continuity, boundary preservation, omission errors, and confusion with visually similar agricultural features.

The statistical and visual assessments complemented the aggregate accuracy metrics and informed the selection of the best-performing model for subsequent shelterbelt mapping.

4. Results

4.1. Quantitative Comparison of Model Performance

Table 2 summarizes the quantitative performance of the four segmentation models. U-Net achieved the highest values across all reported metrics, with a precision of 94.58%, recall of 94.77%, $F1$-score of 94.67%, *IoU* of 89.88%, OA of 99.71%, and a Kappa coefficient of 0.9452. These results indicate a strong balance between precision and recall, together with the highest overlap with the reference shelterbelt masks.

Table 2. Performance metrics for the four semantic segmentation models.

Model	*Precision* (%)	Recall (%)	*F1-Score* (%)	*IoU* (%)	*OA* (%)	*Kappa*
U-Net	94.58	94.77	94.67	89.88	99.71	0.9452
U-Net scSE	91.52	93.35	92.43	85.92	99.58	0.9221
U-Net++	90.16	90.70	90.43	82.53	99.47	0.9016
DeepLabV3+	85.21	87.01	86.10	75.60	99.23	0.8571

Based on $F1$-score and *IoU*, the models ranked in the following order: U-Net, U-Net scSE, U-Net++, and DeepLabV3+. Compared with U-Net scSE, U-Net improved $F1$-score and *IoU* by 2.24 and 3.96 percentage points, respectively. The corresponding margins over U-Net++ were 4.24 and 7.35 percentage points, increasing to 8.57 and 14.28 percentage points over DeepLabV3+.

Although OA exceeded 99% for all four models, $F1$-score and *IoU* showed clearer differences among the architectures. These metrics therefore provided greater discrimination of shelterbelt segmentation performance, whereas OA was less sensitive to differences in the extraction of the comparatively sparse shelterbelt class. The aggregate results identified U-Net as the leading model, with its patch-level performance and visual behavior examined further in Sections 4.2 and 4.3.

4.2. Patch-Level Statistical Comparison of Model Performance

The patch-level distributions of $F1$-score and *IoU* for the four models are shown in Figure 3. U-Net exhibited the highest median values for both metrics, whereas DeepLabV3+ showed the lowest values, with U-Net scSE and U-Net++ exhibiting intermediate performance. The Friedman test indicated significant overall differences among the four models for both $F1$-score and *IoU* ($p < 0.05$). This pattern was consistent with the aggregate results reported in Table 2, providing additional evidence that U-Net achieved the strongest patch-level performance among the evaluated models.

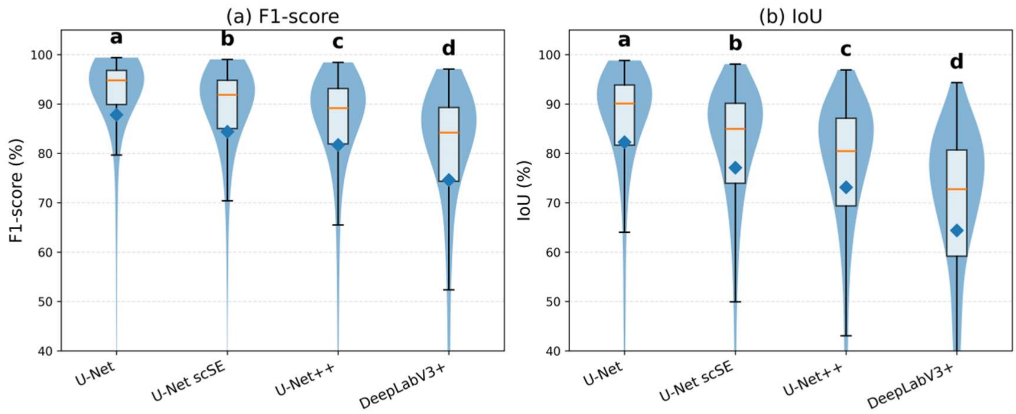

Figure 3. Patch-level distributions of $F1$-score and *IoU* for the four segmentation models. Different letters indicate significant pairwise differences based on Bonferroni-adjusted Wilcoxon signed-rank tests following the Friedman test (family-wise $\alpha = 0.05$).

Post hoc pairwise comparisons were conducted using Wilcoxon signed-rank tests with Bonferroni correction. The resulting significance groupings are indicated by letters in Figure 3, with different letters denoting significant pairwise differences at a family-wise significance level of $\alpha = 0.05$.

https://doi.org/10.3390/f17080983

4.3. Visual Comparison of Model Extraction Results

Figure 4 compares the predictions of the four models for selected validation patches. All four models captured the main shelterbelt structures, but differences were evident in feature continuity, preservation of narrow branches and intersections, and boundary delineation.

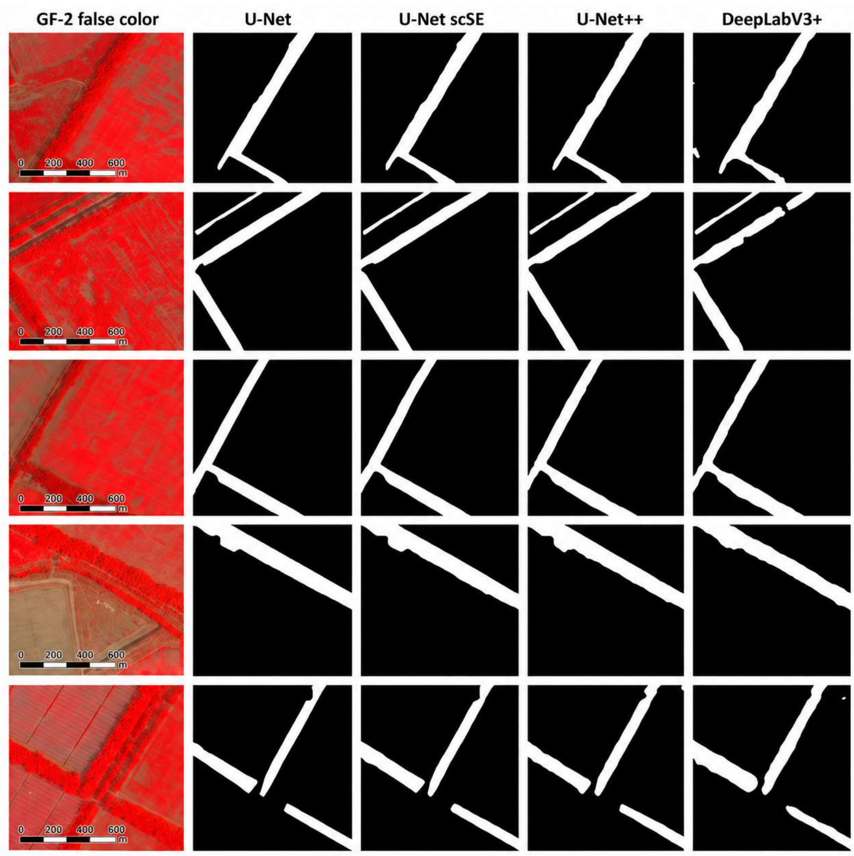

Figure 4. Qualitative comparison of representative farmland shelterbelt extraction results from different segmentation models. The first column shows GF-2 false-color image patches, and the remaining columns show binary extraction results generated by U-Net, U-Net scSE, U-Net++, and DeepLabV3+.

U-Net generally produced the most continuous predictions and preserved narrow branches and intersections more effectively than the other models. U-Net scSE retained similar overall structures but showed occasional local boundary irregularities. U-Net++ captured the main shelterbelt features but exhibited greater local shrinkage and boundary simplification in narrow or intersecting segments. DeepLabV3+ showed more discontinuities and omission errors, particularly around thin branches and complex junctions. These visual observations were consistent with the quantitative and patch-level results and supported the selection of U-Net for subsequent shelterbelt mapping across the study area.

4.4. Application of the Optimal U-Net Model for Shelterbelt Mapping

Following the model comparison, U-Net was applied to the complete preprocessed GF-2 mosaic to map farmland shelterbelts across the 11th Regiment. The mapped shelterbelts were concentrated within cultivated areas and occurred mainly along field boundaries, road margins, and irrigation canals (Figure 5). The enlarged views further show narrow, linear, and locally grid-like patterns associated with the regular arrangement of agricultural parcels. Mapped shelterbelts were comparatively sparse outside cultivated blocks.

Figure 5. Spatial distribution of farmland shelterbelts extracted by the optimal U-Net model in the 11th Regiment. Enlarged panels show representative local shelterbelt patterns.

Using the official cultivated land area of 22.29×10^4 mu (approximately 148.6 km^2) for the 11th Regiment in 2025 as the reference, the mapped shelterbelt area was approximately 6.5 km^2, equivalent to 4.37% of the cultivated land area.

5. Discussion

5.1. Applicability of Semantic Segmentation for Farmland Shelterbelt Extraction

The present results support the applicability of semantic segmentation for fine-scale farmland shelterbelt mapping from 1 m GF-2 imagery within the study scene. Among the four evaluated architectures, U-Net achieved the strongest overall performance, with an $F1$-score of 94.67% and an IoU of 89.88%, and its advantage was also evident in the patch-level and visual comparisons. The differences among the models were particularly apparent in the delineation of narrow branches, intersections, and locally fragmented segments. This pattern highlights an important characteristic of the task: reliable shelterbelt mapping depends not only on distinguishing shelterbelt pixels from the surrounding agricultural background, but also on preserving the fine spatial structure of narrow linear features.

The strong performance of U-Net is consistent with the recent high-resolution shelterbelt benchmarking study of Zhou et al. [22], in which U-Net achieved a Dice similarity coefficient of 91.45% and an IoU of 84.68% on an independent test set. Although the overlap metrics reported in the two studies are numerically similar, they should not be used for a direct performance ranking because the studies differ in image source, spatial resolution, sample construction, model configuration, and validation design. More importantly, the consistently strong performance of U-Net under different experimental settings provides converging evidence that its architecture is well suited to the delineation of narrow linear shelterbelts. A plausible explanation lies in the encoder–decoder structure of U-Net and its skip connections [23], which combine higher-level semantic information with finer spatial detail during reconstruction. For narrow shelterbelts, this ability is particularly relevant because relatively small boundary displacements, local gaps, or missed branches can substantially reduce spatial overlap with the reference mask. The present results therefore suggest that, under the conditions examined here, effective preservation of local structural information may be more important than increasing architectural complexity alone.

Compared with shelterbelt extraction approaches that rely more heavily on predefined spectral indices, handcrafted feature combinations, or rule-based processing [7–9], the present framework learns task-relevant spectral–spatial representations directly from labeled imagery. The contribution of this study, however, is not to claim universal superiority over existing extraction methods, because quantitative performance across studies is influenced by differences in sensors, spatial resolution, target definitions, reference data, and evaluation procedures. Rather, the results demonstrate that semantic segmentation can provide accurate and structurally coherent shelterbelt delineation from 1 m GF-2 imagery and that meaningful differences among alternative architectures can be identified under a common experimental protocol. Because the training and validation subsets consisted of non-overlapping patches derived from the same GF-2 image coverage, these findings primarily support within-scene model performance; geographic transferability remains to be established using spatially independent imagery or study regions.

5.2. Segmentation Behavior for Narrow Linear Shelterbelts

The comparison among models indicates that the main difficulty of this task is not simply separating shelterbelt pixels from the background, but preserving narrow, elongated, and locally discontinuous structures. Although OA exceeded 99% for all four models, $F1$-score and IoU differed much more clearly among architectures. This contrast is expected in a strongly imbalanced segmentation problem, because correct classification of the dominant background can produce a high OA even when errors remain along sparse shelterbelt features. In contrast, $F1$-score and IoU are more sensitive to omissions, fragmentation, and imperfect spatial overlap. Similar difficulties have been reported for other narrow or small targets in high-resolution remote sensing, including small-object segmentation and road extraction [32,33].

The stronger performance of U-Net appears to be closely related to its ability to retain fine spatial information while recovering higher-level semantic features. In the visual comparisons, U-Net preserved narrow branches, intersections, and continuous belt segments more consistently, whereas the other models showed varying degrees of local shrinkage, omission, or discontinuity. The fact that U-Net scSE and U-Net++ did not surpass the standard U-Net also suggests that additional architectural complexity does not necessarily improve the delineation of narrow linear targets. This should not be interpreted as evidence that attention mechanisms, nested skip connections, or multi-scale context are generally less effective. Rather, under the present dataset and training conditions, accurate recovery of local structure appears to have been more important than architectural complexity alone. Because DeepLabV3+ used a ResNet50 encoder while the other three models used ResNet34, the observed differences should be interpreted as comparisons among complete model configurations rather than as isolated effects of individual architectural components.

These differences are important for shelterbelt mapping because relatively small local errors can alter the apparent continuity of an entire linear network. A missed branch, a short break, or boundary simplification may involve only a limited number of pixels but can still reduce the structural fidelity of the mapped shelterbelt system. Related work on fine-scale vegetation segmentation has likewise emphasized the importance of precise structural delineation [34], while boundary-aware remote sensing studies have shown that explicit preservation of edge information can improve the representation of complex object boundaries [35,36]. The present findings therefore support the use of overlap-based metrics together with patch-level and visual assessments when evaluating farmland shelterbelt segmentation, particularly when the practical objective is to preserve the continuity and geometry of narrow linear features.

5.3. Implications and Future Perspectives

The study-area mapping demonstrates the potential of 1 m GF-2 imagery and semantic segmentation for fine-scale farmland shelterbelt inventory. The mapped shelterbelts occurred predominantly along field boundaries, road margins, and irrigation canals and covered approximately 6.5 km^2, corresponding to 4.37% of the cultivated land area. Beyond providing an estimate of shelterbelt extent, the resulting map retains the characteristic linear organization of the shelterbelt network within the agricultural landscape. Such spatial information can provide a useful basis for shelterbelt inventory, network maintenance, and subsequent assessment of shelterbelt condition and configuration.

The interpretation of model performance should, however, take the validation design into account. Training and validation samples were generated as non-overlapping 256 × 256-pixel patches, thereby preventing direct pixel sharing between the two subsets. This design reduces the risk of direct information leakage but does not guarantee spatial independence, because both subsets were derived from the same GF-2 image coverage and geographically adjacent patches may remain spatially correlated. The use of the same validation samples for all four models nevertheless provides a consistent basis for comparing their relative performance within the study scene, because differences in sample composition are held constant across architectures. Accordingly, the reported metrics should be interpreted primarily as comparative within-scene validation estimates rather than as independent measures of geographic generalization. Residual spatial dependence may influence the absolute magnitude of the accuracy estimates, and transferability to other scenes or regions therefore remains to be established.

Future studies should address this limitation through spatially separated hold-out areas, independent GF-2 scenes, and cross-region validation. Such designs would provide a more rigorous assessment of model transferability under variation in shelterbelt species, planting density, field configuration, phenology, irrigation conditions, and surrounding land cover. Multi-temporal imagery may further improve discrimination between shelterbelts and spectrally similar cropland, orchard, roadside, and canal-side vegetation. In addition, field measurements and complementary UAV, LiDAR, or hyperspectral observations could extend the present framework from shelterbelt occurrence mapping to the characterization of structural attributes such as belt width, tree height, porosity, orientation, and network continuity. Recent work integrating UAV LiDAR and hyperspectral data illustrates the potential of multi-source remote sensing for linking remotely sensed information with forest inventory and management variables [37]. Such developments would extend fine-scale shelterbelt mapping from spatial inventory toward structural condition assessment and windbreak-system management.

6. Conclusions

This study compared four semantic segmentation models for farmland shelterbelt extraction from 1 m GF-2 imagery in an oasis agricultural region. Under a common experimental protocol, U-Net achieved the best overall performance, with an $F1$-score of 94.67% and an IoU of 89.88%. The quantitative, patch-level, and visual evaluations further showed that U-Net more effectively preserved narrow branches, intersections, and continuous shelterbelt segments than the other evaluated models. These findings indicate that the preservation of fine spatial structure is particularly important for mapping narrow linear vegetation features.

Application of U-Net across the study area yielded a detailed map of farmland shelterbelt distribution. Shelterbelts occurred predominantly along field boundaries, road margins, and irrigation canals, with a mapped area of approximately 6.5 km^2, equivalent to 4.37% of the cultivated land area. The results demonstrate the potential of combining high-resolution GF-2 imagery with semantic segmentation for detailed shelterbelt inventory in oasis agricultural landscapes.

The reported accuracy should be interpreted in the context of the validation design, because the training and validation subsets were derived from non-overlapping patches within the same GF-2 image coverage. The present results therefore support within-scene model comparison but do not establish geographic transferability. Further evaluation using spatially independent scenes and study regions is needed to assess model performance beyond the present area. Future work could also incorporate multi-temporal and multi-source remote sensing data to improve discrimination of spectrally similar vegetation and to characterize shelterbelt structural attributes.

Author Contributions: Y.X. designed the study, performed GF-2 image preprocessing, constructed the shelterbelt reference mask and patch dataset, implemented deep learning model training and evaluation, generated the figures, and drafted the manuscript. Y.L. supervised the study, contributed to its design and the interpretation of the findings, and made substantive revisions to the manuscript. C.T. contributed to study conceptualization and design, advised on the methodological framework, assisted with funding acquisition and result interpretation, and substantially revised the manuscript. P.L., Z.Z., L.L. and Z.C. provided GF-2 imagery, field background information, and technical assistance with data management and computing support. They assisted in data checking and participated in discussions regarding the analysis results. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Science and Technology Program of XPCC (Grant No. 2026YD024), the National Natural Science Foundation of China (Grant No. 32560372), and the Plan for Tackling Key Scientific and Technological Problems in Key Areas of the Xinjiang Production and Construction Corps, under the project "Research and demonstration on structural optimization techniques for desert–oasis shelterbelt systems" (Grant No. 2023AB017-02).

Data Availability Statement: The GF-2 PMS2 data analyzed in this study were supplied by the Forestry and Grassland Workstation of the Xinjiang Production and Construction Corps and originate from the China Centre for Resources Satellite Data and Application (CRESDA), China's official satellite-data distribution platform. The source imagery cannot be released publicly because its use is subject to access restrictions. Researchers seeking the original data should apply through the authorized provider under the applicable data-use regulations. Code for the deep learning models can be obtained from the corresponding author on reasonable request.

Acknowledgments: The authors thank the Forestry and Grassland Workstation of the Xinjiang Production and Construction Corps for supplying the GF-2 imagery used in this research.

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

References

1. Aili, A.; Xu, H.; Waheed, A.; Bakayisire, F.; Xie, Y. Synergistic windbreak efficiency of desert vegetation and oasis shelter forests. *PLoS ONE* **2024**, *19*, e0312876. [CrossRef] [PubMed]
2. Enescu, C.M.; Mihalache, M.; Ilie, L.; Dinca, L.; Constandache, C.; Murariu, G. Agricultural benefits of shelterbelts and windbreaks: A bibliometric analysis. *Agriculture* **2025**, *15*, 1204. [CrossRef]
3. Zhai, J.; Wang, L.; Liu, Y.; Wang, C.; Mao, X. Assessing the effects of China's Three-North Shelter Forest Program over 40 years. *Sci. Total Environ.* **2023**, *857*, 159354. [CrossRef] [PubMed]
4. Podhrázská, J.; Kučera, J.; Doubrava, D.; Doležal, P. Functions of windbreaks in the landscape ecological network and methods of their evaluation. *Forests* **2021**, *12*, 67. [CrossRef]

5. Pirbasti, M.A.; McArdle, G.; Akbari, V. Hedgerows monitoring in remote sensing: A comprehensive review. *IEEE Access* **2024**, *12*, 156184–156207. [CrossRef]
6. Chen, L.; Letu, H.; Fan, M.; Shang, H.; Tao, J.; Wu, L.; Zhang, Y.; Yu, C.; Gu, J.; Zhang, N.; et al. An introduction to the Chinese High-Resolution Earth Observation System: Gaofen-1~7 civilian satellites. *J. Remote Sens.* **2022**, *2022*, 9769536. [CrossRef]
7. Deng, R.; Guo, Q.; Jia, M.; Wu, Y.; Zhou, Q.; Xu, Z. Extraction of farmland shelterbelts from remote sensing imagery based on a belt-oriented method. *Front. For. Glob. Change* **2023**, *6*, 1247032. [CrossRef]
8. Li, Y.F.; Sun, B.; Gao, Z.H.; Wang, B.Y.; Yan, Z.Y.; Su, W.S.; Gao, T.; Yue, W. Farmland shelterbelt information extraction based on multispectral image of the ZY1-02E satellite. *Natl. Remote Sens. Bull.* **2024**, *28*, 624–634. [CrossRef]
9. Zhang, X.; Liu, J.; Meng, L.; Qin, C.; An, Z.; Wang, Y.; Liu, H. Enhanced blue band vegetation index (the re-modified anthocyanin reflectance index, RMARI) for accurate farmland shelterbelt extraction. *Remote Sens.* **2024**, *16*, 3680. [CrossRef]
10. Zhang, Q.; Zhou, Y.; Zhao, H.; Wu, W.; Huang, Y. A multi-feature fusion based remote sensing inversion method for farmland shelterbelts. *Front. Remote Sens.* **2026**, *7*, 1730222. [CrossRef]
11. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [CrossRef]
12. Talukdar, S.; Singha, P.; Mahato, S.; Shahfahad; Pal, S.; Liou, Y.-A.; Rahman, A. Land-use land-cover classification by machine learning classifiers for satellite observations—A review. *Remote Sens.* **2020**, *12*, 1135. [CrossRef]
13. Hoeser, T.; Kuenzer, C. Object detection and image segmentation with deep learning on Earth observation data: A review—Part I: Evolution and recent trends. *Remote Sens.* **2020**, *12*, 1667. [CrossRef]
14. Kattenborn, T.; Leitloff, J.; Schiefer, F.; Hinz, S. Review on convolutional neural networks (CNN) in vegetation remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 24–49. [CrossRef]
15. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [CrossRef]
16. Zhang, C.; Xia, K.; Feng, H.; Yang, Y.; Du, X. Tree species classification using deep learning and RGB optical images obtained by an unmanned aerial vehicle. *J. For. Res.* **2021**, *32*, 1879–1888. [CrossRef]
17. Zhang, C.; Zhang, L.; Zhang, B.Y.J.; Sun, J.; Dong, S.; Wang, X.; Li, Y.; Xu, J.; Chu, W.; Dong, Y.; et al. Land cover classification in a mixed forest-grassland ecosystem using LResU-net and UAV imagery. *J. For. Res.* **2022**, *33*, 923–936. [CrossRef]
18. Waldner, F.; Diakogiannis, F.I. Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network. *Remote Sens. Environ.* **2020**, *245*, 111741. [CrossRef]
19. Zhu, Y.; Pan, Y.; Hu, T.; Zhang, D.; Zhao, C.; Gao, Y. A generalized framework for agricultural field delineation from high-resolution satellite imageries. *Int. J. Digit. Earth* **2024**, *17*, 2297947. [CrossRef]
20. Tang, Z.; Pan, X.; She, X.; Ma, J.; Zhao, J. Detail and deep feature multi-branch fusion network for high-resolution farmland remote-sensing segmentation. *Remote Sens.* **2025**, *17*, 789. [CrossRef]
21. Wang, Y.; Li, Q.; Wang, H.; Zhang, Y.; Du, X.; Shen, Y.; Dong, Y. Improved prototypical network model for classification of farmland shelterbelt using Sentinel-2 imagery. *Forests* **2024**, *15*, 1995. [CrossRef]
22. Zhou, W.; Liu, L.; Liu, R.; Chen, F.; Yang, L.; Qin, L.; Lyu, R. Precise mapping of linear shelterbelt forests in agricultural landscapes: A deep learning benchmarking study. *Forests* **2026**, *17*, 91. [CrossRef]
23. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer: Cham, Switzerland, 2015; pp. 234–241. [CrossRef]
24. Diakogiannis, F.I.; Waldner, F.; Caccetta, P.; Wu, C. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 94–114. [CrossRef]
25. Roy, A.G.; Navab, N.; Wachinger, C. Recalibrating fully convolutional networks with spatial and channel squeeze and excitation blocks. *IEEE Trans. Med. Imaging* **2019**, *38*, 540–549. [CrossRef] [PubMed]
26. Zhou, Z.; Rahman Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans. Med. Imaging* **2020**, *39*, 1856–1867. [CrossRef] [PubMed]
27. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Computer Vision—ECCV 2018*; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Springer: Cham, Switzerland, 2018; pp. 833–851. [CrossRef]
28. Xu, B.W.; Liu, H.; Wang, J.M.; Cao, S. Ecological design approach for urban green space in arid and semi-arid areas in Northwest China based on habitat and vegetation characteristics thereof. *Landsc. Archit.* **2024**, *31*, 12–18. [CrossRef]
29. Xue, Z.; Wang, X.; Zhou, W. Selection of optimal farmland tree species in southern Xinjiang oasis based on the process of photosynthesis. *Pol. J. Environ. Stud.* **2025**, *34*, 5077–5083. [CrossRef]
30. Jadon, S. A survey of loss functions for semantic segmentation. In *Proceedings of the 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Viña del Mar, Chile, 27–29 October 2020*; IEEE: Piscataway, NJ, USA, 2020; pp. 1–7. [CrossRef]

31. Sudre, C.H.; Li, W.; Vercauteren, T.; Ourselin, S.; Cardoso, M.J. Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Cardoso, M.J., Arbel, T., Carneiro, G., Syeda-Mahmood, T., Tavares, J.M.R.S., Moradi, M., Bradley, A., Greenspan, H., Papa, J.P., Madabhushi, A., et al., Eds.; Springer: Cham, Switzerland, 2017; pp. 240–248. [CrossRef] [PubMed]
32. Kampffmeyer, M.; Salberg, A.-B.; Jenssen, R. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, USA, 26 June–1 July 2016*; IEEE: Piscataway, NJ, USA, 2016; pp. 680–688. [CrossRef]
33. Zhang, Z.; Liu, Q.; Wang, Y. Road extraction by deep residual U-Net. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 749–753. [CrossRef]
34. Buchner, L.; Eisen, A.-K.; Jochner-Oette, S. How precise is precise enough? Tree crown segmentation using high resolution close-up multispectral UAV images and its effect on NDVI accuracy in *Fraxinus excelsior* L. trees. *J. For. Res.* **2025**, *36*, 137. [CrossRef]
35. Marmanis, D.; Schindler, K.; Wegner, J.D.; Galliani, S.; Datcu, M.; Stilla, U. Classification with an edge: Improving semantic image segmentation with boundary detection. *ISPRS J. Photogramm. Remote Sens.* **2018**, *135*, 158–172. [CrossRef]
36. Li, X.; Xie, L.; Wang, C.; Miao, J.; Shen, H.; Zhang, L. Boundary-enhanced dual-stream network for semantic segmentation of high-resolution remote sensing images. *GISci. Remote Sens.* **2024**, *61*, 2356355. [CrossRef]
37. Quan, Y.; Shao, G.; Hao, Y.; Li, M. Bridging forestry practice and remote sensing: Scaling up forest composition with integrated UAV LiDAR and hyperspectral data. *J. For. Res.* **2026**, *37*, 33. [CrossRef]