Article

Extraction of Sunflower Lodging Information Based on UAV Multi-Spectral Remote Sensing and Deep Learning

Guang Li ¹, Wenting Han ^{1,2,*}, Shenjin Huang ¹ , Weitong Ma ³, Qian Ma ⁴ and Xin Cui ¹

¹ College of Mechanical and Electronic Engineering, Northwest A&F University, Xianyang 712100, China; liguang@nwafu.edu.cn (G.L.); ShenjinHuang@nwafu.edu.cn (S.H.); cuixin321@nwafu.edu.cn (X.C.)

² Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture, Xianyang 712100, China

³ College of Water Resources and Architectural Engineering, Northwest A&F University, Xianyang 712100, China; mawt@nwafu.edu.cn

⁴ Institute of Soil and Water Conservation, Chinese Academy of Sciences, Ministry of Water Resources, Xianyang 712100, China; maqian18@mailsucas.ac.cn

* Correspondence: hwt@nwafu.edu.cn

Abstract: The rapid and accurate identification of sunflower lodging is important for the assessment of damage to sunflower crops. To develop a fast and accurate method of extraction of information on sunflower lodging, this study improves the inputs to SegNet and U-Net to render them suitable for multi-band image processing. Random forest and two improved deep learning methods are combined with RGB, RGB + NIR, RGB + red-edge, and RGB + NIR + red-edge bands of multi-spectral images captured by a UAV (unmanned aerial vehicle) to construct 12 models to extract information on sunflower lodging. These models are then combined with the method used to ignore edge-related information to predict sunflower lodging. The results of experiments show that the deep learning methods were superior to the random forest method in terms of the obtained lodging information and accuracy. The predictive accuracy of the model constructed by using a combination of SegNet and RGB + NIR had the highest overall accuracy of 88.23%. Adding NIR to RGB improved the accuracy of extraction of the lodging information whereas adding red-edge reduced it. An overlay analysis of the results for the lodging area shows that the extraction error was mainly caused by the failure of the model to recognize lodging in mixed areas and low-coverage areas. The predictive accuracy of information on sunflower lodging when edge-related information was ignored was about 2% higher than that obtained by using the direct splicing method.

Keywords: sunflower lodging; deep learning; multispectral remote sensing; UAV (unmanned aerial vehicle)

Citation: Li, G.; Han, W.; Huang, S.; Ma, W.; Ma, Q.; Cui, X. Extraction of Sunflower Lodging Information Based on UAV Multi-Spectral Remote Sensing and Deep Learning. *Remote Sens.* **2021**, *13*, 2721. <https://doi.org/10.3390/rs13142721>

Academic Editor: Justin F Moat

Received: 12 May 2021

Accepted: 9 July 2021

Published: 10 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sunflowers are planted in many countries because of their high nutritional value and economic benefits. Due to the influence of the weather, measures of field management, condition of soil, and genotypes, lodging can easily occur in the middle and late periods of sunflower growth. It significantly affects the yield and quality of the crop while complicating its automated management and harvest [1–3]. The accurate and rapid acquisition of information on sunflower lodging can provide a scientific basis for agricultural insurance and management to evaluate the extent of damage to crops and formulate corrective measures [4]. Traditional ground surveys typically use manual measurements to identify sunflower lodging, where this is time consuming, laborious, and costly [5]. Remote sensing technology has been widely used in agricultural monitoring because of its strong monitoring capability and low labor input. Due to the randomness of the time points and spatial locations of lodging, satellite remote sensing technology cannot satisfy the requirements of the spatial resolution of monitoring, which makes it difficult to obtain accurate data. Near-ground remote sensing technology is difficult to use to monitor lodging

because of its limited monitoring capability, and the long time and intensive labor required. With the advantages of easy construction, low cost, simple operation, and high spatial and temporal resolutions, UAV-based remote sensing compensates for the deficiencies of satellite remote sensing and near-ground remote sensing [6–8]. Therefore, it is important to explore this technology to develop efficient and accurate methods to identify lodging in case of agricultural disasters.

With developments in materials, inertial navigation, automatic control, and related technologies in recent years, UAVs have been intensely studied for agricultural monitoring [9]. They carry a variety of sensors to obtain different remote sensing-based features of objects on the ground. UAVs currently carry RGB cameras and multi-spectral cameras to monitor crop lodging [10–12]. Li et al. [13] used UAV RGB images to study the extraction method of maize lodging. Using the coefficient of variation and relative difference, selecting the mean values of red, green, and blue as characteristic parameters from the image texture, the lodging information of maize was extracted by combination with the threshold method. The extraction error (0.3%, 3.5%, 6.9%) of this method is much lower than the method based on RGB gray levels (22.3%, 94.1%, 32%). Li et al. [14] used wheat as the object of interest, and applied spatial domain enhancement and secondary low-pass filtering to two sets of temporal UAV RGB images acquired by a UAV, respectively, to obtain new band features, thus gaining the scatter diagram of lodging and normal wheat in different new band features combination coordinate system. Lodging extraction features were constructed according to the dividing line of scattered points. These features were combined with the k-means algorithm to extract wheat lodging information, and the yielded values of overall accuracy (OA), Kappa coefficient, and the extraction accuracy of lodging reached of 86.44%, 0.73, and 81.04%, respectively. Yang et al. [15] obtained the digital surface model (DSM) and texture information of the image through image modeling and texture analysis. The single feature probability (SFP) value was calculated to evaluate the contribution of each feature for classification and the texture and DSM were selected as lodging extraction features. Combined with the decision tree classification model, the accuracy of rice lodging information extraction was 96.71%, and the Kappa coefficient was 0.941. The use of RGB images to monitor crop growth has limitations in terms of spectral characteristics, especially a lack of near-infrared bands that can reflect the spectral differences of different canopy structures [16]. Chauhan et al. [17] used nine-band multi-spectral UAV remote sensing images combined with the nearest-neighbor algorithm to distinguish between regions of lodging and non-lodging based on differences in spectral reflectance. Kumpumaki et al. [18] used the red, green, blue, red-edge, and near-infrared bands of UAV multi-spectral images to classify and map crop lodging, and examined the linear relationship between lodging and crop yield. A few authors have also studied crop lodging by using thermal infrared images and radar data [19–21]. Information on crop lodging through UAV remote sensing is mainly acquired by combining hand-craft features and a classifier. The effect of classification needs to be improved, and the correct selection of low-level features has a significant effect on the accuracy of recognition.

Deep learning methods can automatically extract features from images. Due to the convolutional nature of deep learning methods, they take into account the characteristics of neighboring pixels. They have a wide domain of acceptance and pixel-level classification ability that can ensure more objective and accurate image processing [22,23]. Some researchers have introduced deep learning to the extraction of information on crop lodging. For example, Yang et al. [24] used FCN-AlexNet and SegNet to extract information on rice lodging from UAV RGB images, and the results yielded an F1-score of 80%. Mdya et al. [25] used edge computing with deep learning EDANet to extract information on high rates of lodging in rice fields at multiple levels. The method was 36% faster than conventional methods, with an accuracy of 99.25%. Hamidisepehr et al. [26] used computer vision and three deep learning methods (Faster R-CNN, YOLOv2, and RetinaNet) to extract and compare information on maize lodging from UAV RGB images. Zhao et al. [27] obtained RGB and multi-spectral images (RGN) by carrying different cameras on UAVs, and extracted rice

lodging with the U-Net model, respectively, with accuracy of 94.42% and 92.84%. Zheng et al. [28] used the full convolution network in deep learning to extract the area of lodging of corn from UAV images, and reported an F1-score above 90%. Deep learning combined with UAV remote sensing technology can be used to extract information on crop lodging, but the relevant studies have mainly focused on crops with high canopy density, such as rice, maize, and wheat. [15,29]. The lodging areas of such crops are mainly located in the canopy and the stem, and the features of their images are relatively simple. When lodging occurs in crops with low canopy density, such as sunflower, the lodging area is a mixture of the stem, leaves, and soil. The spectrum, texture, and spatial distribution of the images is far more complex than those of crops with high canopy density. Few studies have examined extracting information on lodging in crops with a low canopy density. Song et al. [30] obtained multi-spectral band data at a high spatial resolution by fusing UAV multi-spectral images and visible-light images, and then extracted information on sunflower lodging with an accuracy of 83.3% through the improved SegNet algorithm. Although they noted that deep learning can improve the extracted information on lodging, increasing the image characteristics of the band can improve the accuracy of extraction. To date, no focused analysis has examined the applicability of deep learning to the extraction of the lodging information of sunflowers, the influence of multi-spectral bands on the information obtained, and a vector map of lodging in a given region. The extraction of information of lodging in crops with low canopy density thus requires further research.

Sunflower crops with low canopy density are easily affected by the condition of the soil, management, and other factors that lead to significant differences in their growth, and their lodging situation is complex. In this study, random forest, SegNet, and U-Net are used to construct multiple models of sunflower lodging by combining information from the RGB, RGB + NIR, RGB + red-edge, and RGB + NIR + red-edge bands. The state of lodging of regional sunflowers was predicted by ignoring the edge-related information in the acquired images. The use of different methods in different bands to extract the lodging information of sunflowers was evaluated from classification accuracy while ignoring edge-related information. The influence of various methods and NIR and red-edge multispectral bands information on the extraction of regional lodging information was also investigated. The results here can provide theoretical and empirical support for the accurate and rapid extraction of the lodging information of regional sunflower crops.

2. Materials and Methods

2.1. Study Area and Data

2.1.1. Selection of Study Areas

The study area is located in Wuyuan County, Bayannur City, Inner Mongolia Autonomous Region (107°35'70"–108°37'50" E, 40°30'46"–41°45'16" N, altitude about 1030 m). The region has a continental monsoon climate in the middle temperate zone, which is characterized by abundant light energy, abundant sunshine (annual total radiation 153.44 CAL·cm⁻²), dry and windy, and low rainfall (annual average rainfall 170 mm). This region basically relies on Yellow River water for agricultural irrigation. Due to years of uncontrolled irrigation from the Yellow River, the groundwater level in this region is high and the soil salinity degree is high. As a result, only salinity resistant crops can be planted in this region, such as maize and sunflower, among which the largest planting scale is sunflower [31,32].

On 29 August 2020, a large number of mature sunflowers in the study area collapsed due to heavy rain and accompanied by force 4 wind. Combined with local disaster reports, the team conducted timely investigation of the affected areas, and finally selected 4 UAV sampling sites with an area of about 15 hectares (plot 1, plot 2, plot 3, and plot 4, as shown in Figure 1b). The sunflower variety “SH363” sown in early June, with a large row spacing of 1 m, a small row spacing of 0.4 m, and a plant spacing of 0.55 m. In the growing period of sunflower, cultivation, irrigation, fertilization, pest control, and weed control should be carried out according to its growth status and local farmland management measures. Sunflowers growing in different locations may have different canopy density and leaf color due to great variations in soil type, fertility, and salinity. The complex conditions of the ground features in each sampling area are different. Plot 1 contains lodging sunflowers, sunflowers with different growth conditions, wasteland, and road. Plot 2 contains lodging sunflowers, sunflowers with different growth conditions, maize, zucchini, woods, wasteland, and road. Plot 3 contains lodging sunflowers, normal sunflowers, zucchini, maize, road, and buildings. Plot 4 contains lodging sunflower, normal sunflower, zucchini, maize, and sugar beet.

Figure 1. Cont.

Figure 1. Aerial images of the study area: (a) Hetao irrigation district, Inner Mongolia, China; (b) locations of sunflower lodging plots in Wuyuan County; (c–f) multispectral pseudo-color image of fields 1, 4, 2, 3.

2.1.2. Data Acquisition

In this study, the unmanned aerial vehicle multi-spectral remote sensing system independently developed by the team is adopted to collect remote sensing data, as shown in Figure 2. The system is mainly composed of a six-rotor UAV, stabilization head, image acquisition controller, etc., which can stably obtain the multi-spectral images of the UAV without distortion and mosaic. The selected multispectral sensor is a RedEdge-M 5-channel multispectral camera produced by MicaSense, USA, which can obtain data in red (668 nm), green (560 nm), blue (475 nm), near infrared (840 nm), and red-edge (717 nm) bands. Parameters of the UAV and multi-spectral camera are shown in Table 1. The collection of UAV remote sensing images should be based on previous research experience and relevant UAV management regulations. The sky should be clear of clouds and it should be close to noon to reduce directional effects, the ground has no continuous wind direction and the wind speed is less than level 3. That is, 10:00–14:00 on 3 September 2020. The flight height was set to be 100 meters, the heading and side overlap degree was 80%, and plots 1, 2, 3, and 4 obtained by aerial photography, respectively, with 453, 526, 617, and 463 images in each band. In order to obtain accurate orthophoto images, four ground control points (GCPs) were laid on each plot, and the 3D coordinates of GCPs were accurately measured using RTK (i50, CHCNAV, Shanghai, China). Pix4DMapper software was used to splice the original image to obtain orthophoto images of these 4 plots. First, the original image was imported into Pix4DMapper software, the output coordinate system was set to China2000, and the camera was set to RedEdge mode. The wizard was completed to generate agricultural multispectral orthoimages. Before the start of the splicing program, the picture of the calibrated reflectance panel (Figure 2c) was imported before the UAV flight and the reflectance correction coefficient was set to the correct reflectance of the multispectral image. After the first stage of the splicing program is completed, geometric correction of the multi-spectral image is required. First, importing the ground control point coordinates into Pix4d software, and then the corresponding position of the ground control point coordinates would be found in the original image, and finally marked. After the splicing program is completed, the UAV multi-spectral orthophoto images can be obtained, with a ground resolution of 0.07 m. The results are in Figure 1c–f. The image was stored in TIFF format, and the gray information of five bands of red, green, blue, near red, and red-edge of the ground object was retained. Each color contains 16-bit information, and the value range is 0–65,535.

Figure 2. UAV multi-spectral remote sensing system: (a) Unmanned aerial vehicle system; (b) multispectral camera; (c) calibrated reflectance panel.

Table 1. Main parameters on Unmanned Aerial Vehicle (UAV) and camera.

Unmanned Aerial Vehicle (UAV)		Camera	
Parameters	Values	Parameters	Values
Wheelbase/mm	900	Camera model	MicaSense RedEdge-M
Takeoff mass/kg	4.7–8.2	Pixels	1280 × 960
Payload/g	820	Band	5
Endurance time/min	20	Wave length/nm	400–900
Digital communication distance/km	3	Focal length/mm	5.5
Battery power/(mA·h)	16,000	Field of view/(°)	47.2
Cruising speed/(m·s ^{−1})	5		

2.1.3. Data Set Construction

The data of the study came from normal production fields, not experimental fields after professional processing, so the complexity of crop species is different. Combined with ground investigation and unmanned aerial vehicle monitoring, the four sampling areas included sunflower, lodging sunflower, maize, zucchini, sugar beet, forest, wastelands, roads, and buildings, which brought some challenges to the lodging extraction of sunflower. In the case of different types of ground objects, to obtain the lodging information of sunflower accurately, the classification of ground objects needs to be simplified. Other non-sunflower ground objects were grouped as background, and normal sunflower and lodging sunflower were grouped separately, to unify the classification standard of different complex ground object data in lodging extraction.

Truth maps of four sampling areas were generated in ArcGIS (10.2) by means of a ground survey and visual interpretation, as shown in Figure 3. In plot 1, the planting area of sunflower is 69.17% and the lodging area is 12.18%. In plot 2, the planting area is 49.03% and the lodging area is 9.13%. In plot 3, the planting area of sunflower is 28.38% and the lodging area is 6.61%, and in plot 4, the planting area is 31.66% and the lodging area is 3.34% of the total. By analyzing the ground features and the area ratio of normal to lodged sunflowers in each plot, we found that plots 1, 3, and 4 contained all the ground features needed to form suitable training and validation sets. The area ratio of normal to lodged sunflowers was relatively high, and the lodging situation was complex. Plot 2 contained rich ground features, with significant differences in the growth and complex lodging of sunflowers. It was the ideal test data. The equilibrium between samples of the training and validation sets affected the results of classification. In terms of area, the proportion of collapsible sunflowers was significantly smaller than those of background and normal sunflowers. To ensure the uniformity of the sample set, areas containing different kinds

of ground features were chosen from plots 1, 3, and 4 to form the training and validation sets. The sizes of the chosen images from the plots were 5500×3037 , 4367×3761 , and 4244×3078 pixels, respectively, and the corresponding results are shown in Figure 3. The cropped UAV remote sensing image was large, and using it directly to train the model would have taken up a large amount of GPU memory. The original and the labeled images thus needed to be cropped. To ensure consistency, 10,003 images, each with a resolution of 256×256 , were cut by setting the overlap rate to 0.7 in the window slip. A total of 7259 images were randomly selected as the training set and 2744 images as the validation set. Partial data from plot 1 and data from plot 2 were used as the test data, and are respectively called test area 1 and test area 2. They were used to assess the reliability and capability of generalization of the model. The approach described above was used to obtain the dataset used for deep learning. The random forest method requires selecting pixel point sets of ground objects in the corresponding training area to train the model. A total of 380,000 pixels of images of lodged sunflowers, normal sunflowers, and the background were obtained.

Figure 3. Labeled map of each plot, and the division of training and testing areas: (a) Training and testing areas in plot 1; (b) training areas in plot 4; (c) testing areas in plot 2; (d) training areas in plot 3.

According to the truth image above, the distribution position of the data used for training and validation in the graph was obtained. Next, data sets of different bands need to be constructed. The original acquired multispectral data were 5 bands separately preserved in red, green, blue, near infrared, and red-edge. By using a composite bands module in

ArcGIS (10.2), remote sensing images containing bands of RGB, RGB + NIR, RGB + red-edge, and RGB + NIR + red-edge can be synthesized, respectively. The images were clipped according to the distribution information of training and validation data, respectively, to obtain the regional data for training and validation. Regional data containing different bands and corresponding labels were clipped to obtain data sets and labels for training and validation.

2.2. Methods

2.2.1. Extracting Sunflower Lodging

A flowchart of the research process of this study is shown in Figure 4, including UAV image acquisition, image mosaicking, dataset production, model training and validation, model testing, precision evaluation, and comparison with a conventional random forest classifier. The collected UAV image was composed of red, green, blue, near-infrared, and red-edge bands, and a mosaic was generated from matched orthophotos after refining camera positioning based on GCPs. Once the pixel labels had been obtained, the datasets used to derive the classification models were organized. The training, validation, and test datasets were created from RGB, RGB + NIR, RGB + red-edge, RGB + NIR + red-edge, and labeled images. SegNet and U-Net deep semantic segmentation models were used to combine the four band combinations to generate eight classification models. In the testing stage, the model was combined with the method of ignoring edge-related information to classify and predict the test area. The performance of deep learning and random forest in terms of extracting lodging-related information was compared and analyzed. The influence of different band combinations on deep learning-based classification was also examined, as was the effect of the method of ignoring edge-related information on the results of extraction of lodging.

Figure 4. Research flowchart.

2.2.2. Random Forest Method

Random forest is a non-parametric machine-learning algorithm composed of classification and regression trees (CART), which has high prediction accuracy, good tolerance to outliers and noise, wide application range, and is not easy to over fit [33,34]. Because of statistical learning theory, the bootstraps resampling method is used to extract multiple samples and the decision tree modeling is carried out for each sample. Multiple prediction results are integrated to construct a random forest containing N classification trees. The classifier of random forest algorithm needs to define two parameters to generate the prediction model: the number of expected classification trees (ntree) and the number of features extracted during node splitting (mtry). This paper called the random forest model on the Python platform. Many experiments showed that when ntree was 50, the error gradually converges and tended to be stable. Mtry was the square root of the total feature [32].

2.2.3. U-Net and SegNet Methods

With the proposal of end-to-end total neural network [35], Ronneberger et al. [36] and Badrinarayanan et al. [37] proposed U-Net and SegNet deep semantic segmentation network, respectively. They were applied to biomedical image segmentation and autonomous driving image segmentation. On the left side of the U-Net network is the low-resolution information after multiple subsampling, which can provide contextual semantic information of segmenting the target in the whole image, reflect the relationship between the target and its environment, and help to judge the category of the object. On the right side is a corresponding expansion path, which can provide more fine features and precise positioning for segmentation. In the middle, the low-resolution information on the left side can be combined with high-resolution information on the right side for more accurate image segmentation through jumping connection. On the left side of the SegNet, features are extracted by convolution, and the sensory field is increased by pooling, while the image becomes smaller. On the right side, deconvolution and upsampling are used to reproduce features after image classification through deconvolution, and upsampling is restored to the original size of the image. Finally, each pixel was sent to the softmax classifier to predict its category. Although U-Net and SegNet have achieved good results on some common data sets [35,37], in specific application areas, improvements are needed to achieve the purpose of application.

U-Net was first applied to the segmentation of biomedical images. Compared with remote sensing images, biomedical images have fewer bands and no coordinate information. Most U-Net network structures thus do not support multi-band input, and cannot retain the coordinate information of remote sensing images. To adapt U-Net to process remote sensing images, the authors developed a data generator that can take multi-band images as input. The development of the data generator includes four steps: firstly, using GDAL to read multi-band images. Secondly, image preprocessing, including image normalization and label one-hot coding. Next, use the color_dict function to get the color dictionary of the various categories. Finally, the generator function is used to develop the data generator, which includes batch size, image information, label information, category number, color dictionary, and image size, so as to realize the multi-band data processing.

To prevent overfitting, the dropout layer with a probability of 0.5 was added to the network after the fourth and fifth convolution operations. That is, in each training iteration, neurons were discarded with a probability of 0.5. To increase the speed of network training, batch normalization (BN) was added after each convolution. The improved U-Net is shown in Figure 5, in which the white box represents features extracted in the early stage of the path of contraction that contained abstract but rich spatial information. The yellow box represents the results of the upsampled convolution, where the features were extracted by using the entire architecture, and contained details with little spatial information. “D” in the figure represents the result of dropout processing.

Figure 5. U-Net structure image (reproduced from Wang et al. [38]).

SegNet was originally designed to segment images for autonomous driving and intelligent robots. Most SegNet network structures are based on RGB images as input, and cannot retain the coordinate information of remote sensing images. To adapt it to process remote sensing images, the authors developed a data generator that can take multi-band images as input. The SegNet is shown in Figure 6, and included an encoder composed of a convolutional layer and a pooling layer, a decoder composed of an upper sampling layer and a convolutional layer, and a softmax layer. Convolution and max pooling were performed at the encoder, and upsampling and convolution were performed at the decoder. Finally, each pixel was sent to the softmax classifier to predict its category. A key part of SegNet is the specific code and indexing structure of max pooling, which makes it useful for accurately relocating features and reducing the required end-to-end training of the parameters.

Figure 6. SegNet structure image (reproduced from Badrinarayanan et al. [37]).

2.2.4. Model Training

In all experiments, following the hyperparameter setting suggested by Kingma and Ba [39], an Adam optimizer with a learning rate of 0.001, $\beta_1 = 0.9$, and $\beta_2 = 0.999$ was used. Considering the network structures and GPU memories, the decay of 0.05, the batch size of 16, and the number of epochs equal to 50 were applied. A detailed model training, validation, and testing computing environment can be found in Table 2.

Table 2. Computation resource for model training, validation, and testing.

Operating System	Windows 10 Enterprise 64bit (DirectX 12)
CPU	Intel(R) Core(TM) i9-10920X CPU @ 3.50 GHz(12 cores/GPU node)
RAM	64GB/GPU node
Accelerator	NVIDIA GeForce GTX 3090 24GB
Image	TensorFlow-19.08-py3
Libraries	Python 3.6.13, NumPy 1.19.5, scikit-learn 0.24.1, TensorFlow-GPU 2.4.1, Keras 2.4.3, Jupiternotebook, CUDA 11.2

2.2.5. Prediction of Results

During model prediction, if a large remote sensing image to be classified is directly input to the network model, it causes memory overflow. Therefore, the images to be classified were cut into a series of smaller images to be input to the network for prediction, and the results of prediction were stitched into a final image in accordance with the Coordinate information, as shown in Figure 7. The method of direct splicing prediction often causes splicing trace, which will affect the mapping of regional remote sensing results, as shown in Figure 7a–e. In order to solve the problem of splicing trace, the prediction method of ignoring edge-related information is introduced in this paper. Marginal information of the predicted small image is ignored in the overlapping way as stitching, to obtain the regional prediction results, as shown in Figure 7d. To realize this method, it is necessary to figure out the overlap rate of adjacent images. The predicted result of the cropped image is “A,” and the result of stitching is “a.” The area percentage of “a” in “A” is “r,” and the overlap ratio of adjacent cropped images is “ $1 - r^{1/2}$.” The idea of ignoring edge-related information prediction method and the value of r is 0.5, which are obtained in this paper by referring to literature [38].

Figure 7. Flowchart of model prediction: (a) Image to be classified; (b) crop the image and complete the classification; (c) the classification results are directly stitched together; (d) stitching classification results while ignoring the edges; (e) direct splicing result; (f) the result while ignoring the edges.

2.2.6. Accuracy Evaluation Method

The concepts of precision and recall were used to evaluate the ability of eight deep semantic segmentation models and random forest models to distinguish between lodging sunflower, normal sunflower, and background. As shown in Table 3, TP (true positive) was correctly classified as the positive category, FP (false positive) was wrongly classified as the negative category, TN (true negative) was correctly classified as the negative category, and

FN (false negative) was wrongly classified as the positive category. C represents a specific category, and precision represents the ratio of the number of samples accurately classified as positive categories to the number of samples accurately classified as positive categories. Recall represents sensitivity, which is the ratio of the number of positive samples to the number of actual positive samples in the test data set. Accuracy represents the percentage of all accurate positive and negative classes of a particular class in the entire sample. F1-scores take into account both accuracy and recall, so that both are maximized and a balance is struck. Intersect-over-union (IoU) refers to the ratio of the intersection and union of actual class samples and predicted class samples. F1-score, IoU, and OA can reflect the classification, lodging extraction, and overall classification accuracy of objects in different regions, and can comprehensively reflect the lodging extraction effect. Therefore, these three indicators are selected for evaluation in this paper.

Table 3. Evaluation parameters with associated formulas.

Evaluation Matrices	Formula
Precision	$precision_c = \frac{TP_c}{TP_c + FP_c}$
Recall	$recall_c = \frac{TP_c}{TP_c + FN_c}$
Accuracy	$accuracy_c = \frac{TP_c + TN_c}{TP_c + TN_c + FP_c + FN_c}$
Overall accuracy	$OA = \sum_{c=1}^n \frac{TP_c + TN_c}{TP_c + TN_c + FP_c + FN_c}$
F1-score	$F_1 = \frac{2 \times precision \times recall}{precision + recall}$
Intersection-over-Union	$IoU = \frac{TP_c}{TP_c + FP_c + FN_c}$

3. Results

3.1. Results of Model Training and Validation

During model training with the deep learning method, model parameters were optimized by evaluating model effects with accuracy and loss. Accuracy measures the effect of the model by calculating the ratio of the number of samples correctly classified by the model to the total number of samples. Loss is calculated through the preset loss function, so as to update model parameters and achieve the purpose of reducing optimization error. The changes in the accuracy and loss values of the training and validation sets obtained by using a combination of SegNet and U-Net models in different bands are shown in Figure 8. After 40 epochs, the loss values of the training and validation sets of each model gradually decreased until leveling off, and its accuracy increased until leveling off. This indicates that a reliable model to extract lodging information could be obtained in 70 epochs.

Figure 8. Cont.

Figure 8. Accuracy of training and validation, and loss values of two deep learning methods in different bands: (a) The training and validation accuracy of SegNet; (b) the training and validation loss of SegNet; (c) the training and validation accuracy of U-Net; (d) the training and validation loss of U-Net.

The training accuracy (train_acc) and validation accuracy (val_acc) of the random forest method and the train_acc and val_acc and training loss (train_loss) and validation loss (val_loss) values of the SegNet and U-Net methods are shown in Table 4. Adding band information on the basis of RGB could improve the accuracy of model training and validation. For example, when the random forest only used RGB, the training and validation accuracy was only 81.09% and 67%. When NIR and red-edge were added on the basis of RGB, the training and validation accuracy increased to 99.96% and 85.56%. When SegNet and U-Net were used for training, it was found that adding NIR band on the basis of RGB could improve the accuracy of the model and reduce the loss value of the model. Adding red-edge band on the basis of RGB would reduce the accuracy of the model and increase the loss value. Under the RGB band, the accuracy of SegNet training set and validation set was 99.17% and 93.47%, respectively, and the loss value was 5.13% and 17.31%, respectively. When adding the NIR band on the basis of RGB, the accuracy rate increased to 99.67% and 94.07%, respectively, and the loss value decreased to 4.73% and 16.71%, respectively. When red-edge was added on the basis of RGB, the accuracy rate dropped to 98.17% and 93.17%, respectively, and the loss value increased to 5.83% and 17.39%, which showed the same result in U-Net. Different methods were used to conduct training modeling in different bands, and it was found that multi-bands had certain influence on accommodation information extraction, but the specific influence degree of each band need to be further analyzed.

Table 4. The training and validation results of the three methods in different bands.

Information	Random Forest		SegNet				U-Net			
	Train_Acc (%)	Val_Acc (%)	Train_Acc (%)	Train_Loss (%)	Val_Acc (%)	Val_Loss (%)	Train_Acc (%)	Train_Loss (%)	Val_Acc (%)	Val_Loss (%)
RGB	81.09	67.00	99.17	5.13	93.47	17.31	98.99	11.96	95.10	19.20
RGB + NIR	99.10	79.77	99.67	4.73	94.07	16.71	99.01	8.07	95.21	14.24
RGB + red-edge	98.61	77.96	98.17	5.83	93.17	17.39	98.79	11.55	92.89	33.33
RGB + NIR + red-edge	99.96	85.56	99.87	4.33	94.67	16.11	99.27	10.21	96.52	15.52

3.2. Results of Sunflower Lodging Test

Figure 9 shows the results of predictions of models to extract sunflower lodging based on random forest, SegNet, and U-Net in different band combinations for test area 2. Figure 9a is the original image of the test area, from which it is clear that areas occupied by sunflowers had different canopy densities, complex areas of lodging (different proportions

of soil, film, sunflower stalks, and leaves), and background areas (other crops, buildings, and roads). Figure 9b shows the result of visual interpretation of UAV images obtained in combination with the ground survey, which was used as the truth value to evaluate the predictive accuracy of each model. Figure 9(c1–c4) shows the results of predictions of the random forest in RGB, RGB + NIR, RGB + red-edge, and RGB + NIR + red-edge. Owing to the sparse density of the canopies of sunflowers, bare soil and thin film were exposed in the images. As a result, the “salt and pepper effect” was significant in the results of classification, and most sunflowers were incorrectly classified as lodging sunflowers. As the reflectivities of other vegetation and the sunflowers overlapped significantly in each band, some background crops were incorrectly classified as sunflowers. Figure 9(d1–d4) shows the results of prediction of SegNet in different band combinations, where the “salt and pepper effect” has been suppressed. All models performed well in terms of background recognition but differed in the results of identification and extraction of sunflowers. The model that used the RGB + red-edge band was poor at predicting sunflowers and lodging. Figure 9(e1–e4) shows the results of prediction of the U-Net training model on different band combinations. Its results were similar to those of SegNet. Similarly, the model constructed by using RGB + red-edge had the poorest capability of predicting sunflowers and lodging.

Figure 9. Cont.

Figure 9. Results of prediction of sunflower lodging: (a) original image, (b) ground truth. (c1–c4) show the random forest (RGB), (RGB + NIR), (RGB + red-edge), and (RGB + NIR + red-edge), respectively. (d1–d4) represent SegNet (RGB), (RGB + NIR), (RGB + red-edge), and (RGB + NIR + red-edge), respectively. (e1–e4) represent U-Net (RGB), (RGB + NIR), (RGB + red-edge), and (RGB + NIR + red-edge), respectively.

Table 5 shows the predictive accuracy of 12 models, among which the model constructed by SegNet using the RGB + NIR band has the best result with an overall accuracy (OA) of 88.23%. Its F1-scores in terms of identifying background, sunflower, and lodging were 91.08%, 88.58%, and 51.68%, respectively, and the values of intersection-over-union (IoU) were 83.62%, 79.50%, and 34.85%, respectively. The prediction model constructed by using random forest had an OA of 43.29%, and its F1-scores for background, sunflower, and lodging were 54.14%, 48.27%, and 15.58%, respectively, with IoU values of 37.12%, 31.81%, and 8.45%, significantly lower than the predictive accuracy of the deep learning prediction model. Adding the RGB band information had different effects on the results of prediction. For random forest, the model obtained by adding the NIR and red-edge RGB bands yielded the highest predictive accuracy. For SegNet and U-Net, the model with NIR bands acquired using RGB had the highest predictive accuracy, while the model with information from the red-edge had a lower accuracy, which was consistent with the results of model training.

Table 5. Accuracy of prediction of sunflower lodging.

Model	Information	Test Area	F1-Score (%)			Intersection-over-Union (%)			OA (%)
			Background	Sunflower	Lodging	Background	Sunflower	Lodging	
Random Forest	RGB	2	54.14	46.16	15.05	37.12	30.00	8.14	41.80
	RGB + NIR	2	52.62	46.84	15.58	35.71	30.58	8.45	42.87
	RGB + red-edge	2	50.09	46.32	13.96	33.41	30.14	7.50	40.36
	RGB + NIR + red-edge	2	51.69	48.27	15.07	34.86	31.81	8.15	43.29
SegNet	RGB	2	87.71	83.08	41.71	78.11	71.05	26.35	83.98
		1	81.71	93.40	62.30	69.08	87.62	55.24	88.53
	RGB + NIR	2	91.08	88.58	51.68	83.62	79.50	34.85	88.23
		1	84.32	94.59	84.82	72.89	89.75	69.77	90.75
	RGB + red-edge	2	77.56	57.36	11.83	63.35	40.21	6.28	69.47
		1	75.99	93.37	37.12	61.27	87.56	22.79	85.61
	RGB + NIR + red-edge	2	87.87	85.73	23.93	78.37	75.02	13.59	84.49
		1	80.80	93.85	65.87	67.79	88.42	59.76	87.79
U-Net	RGB	2	84.39	79.36	35.43	72.99	65.78	21.53	80.62
		1	77.92	91.27	58.80	63.82	83.94	52.28	85.36
	RGB + NIR	2	89.58	86.43	45.37	81.12	76.11	29.34	86.47
		1	82.46	93.19	78.85	70.16	87.24	62.66	88.26
	RGB + red-edge	2	81.55	70.12	7.83	68.84	53.99	4.08	75.56
		1	72.25	90.33	48.99	56.55	82.36	30.66	84.54
	RGB + NIR + red-edge	2	89.75	84.03	39.34	81.41	72.46	24.49	84.22
		1	83.19	92.78	61.61	71.22	86.54	54.78	87.69

3.3. Direct Splicing Region Prediction Results

When lodging-related information is extracted from remote sensing images, it is necessary to know the location and area of lodging. Therefore, the results of extraction should be regional vector maps because they are useful for agricultural insurance departments. Due to the limitation of computational performance, the orthophotos of the region should be clipped into a set of images of a certain size for training the model, and the resulting model can predict only items in the clipped image. To obtain the vector map of the results of regional classification, the results of the clipped map after prediction needed to be spliced. The regional image prediction method is currently to directly splice clipped maps according to location-related information. The results of this method commonly contain traces of splicing, as shown in Figure 10a,c. Under the RGB + NIR band, SegNet and U-Net adopted the direct splicing prediction method, and the result accuracy was shown in Table 6.

Figure 10. Image of results of prediction when direct splicing and ignoring edge splicing: (a) Image of results of prediction when direct splicing; (b) image of results of prediction when ignoring edge splicing; (c) display of stitching traces; (d) image of results of prediction when ignoring edge splicing at the same position.

Table 6. Results of prediction of when edge-related information was considered.

Information	Model	F1-Score (%)			Intersection-over-Union (%)			OA (%)
		Background	Sunflower	Lodging	Background	Sunflower	Lodging	
RGB + NIR	SegNet	89.41	85.95	47.55	80.84	75.36	31.19	86.03
	U-Net	88.66	84.95	44.59	79.63	73.84	28.70	85.28

4. Discussion

4.1. Comparison of Classification of Zones of Sunflower Lodging

Compared with other crops with a high canopy density (such as wheat, rice, rapeseed, and maize), sunflowers expose more soil and plastic films after lodging that have complex textural features in the lodging area. This is because of their large planting spacing and low canopy density at maturity. The traditional classification method classifies the pixel points in the window according to the corresponding rules in the way of window slippage, without making use of the interconnection between neighborhoods [40,41]. Due to the complexity of soil texture, spectrum, and background characteristics in the sunflower lodging area, it is difficult to extract the sunflower lodging area without considering the relationship between neighborhoods. Because of the convolutional nature of deep learning methods, they take into account the characteristics of neighboring pixels. The deep learning method is convolutional and can take into account the characteristics of neighborhood pixels, which enables it to have stronger pixel-level classification ability. The experimental

results show that the deep learning method was superior to the conventional random forest method in terms of the overall accuracy of classification of the lodging area. The accuracy of the deep learning method in each experimental group was generally 20–40 percentage points higher than that of the random forest method, as shown in Table 5. Compared with the traditional design of artificial features, the deep learning method could automatically extract deep features that are difficult to hand craft to more objectively and accurately complete classification of the lodging areas. As shown in Figure 9, the random forest classification featured clearly incorrect classifications and a significant “salt and pepper” phenomenon in the results. SegNet and U-Net avoided these issues.

Compared with RGB images, multi-spectral images added the NIR and red-edge bands for richer information that is more widely applicable to agriculture. Current applications of deep learning focus on the processing of RGB images, and seldom involve multi-spectral images. Two common models of deep semantic segmentation, SegNet and U-Net, were used to classify lodging regions here by combining the four band combinations of RGB, RGB + NIR, RGB + red-edge, and RGB + NIR + red-edge. Adding information from the NIR band on the basis of RGB improved the accuracy of classification, whereas adding information from the red-edge band reduced it. In order to explore the universality of the results, this study selected part of plot 1 as test area 1 (Figure 3a), and made lodging predictions. The results of prediction are shown in Table 5. As with the results in test area 2, when information on the NIR band was added on the basis of RGB, the results of prediction of SegNet and U-Net improved in terms of accuracy. When the red-edge band was added, their accuracy decreased. To find out why the red-edge band has a negative effect on the classification, we analyzed changes in the canopy spectrum, because lodging changes the structure of the canopy, resulting in changes in the canopy spectrum. In order to understand the changes in the spectra of the lodging area, spectral curves of the lodging sunflower, soil, and normal sunflower were obtained using FieldSpec Hand Held (ASD, Westborough, CO, USA) during drone monitoring. For parameters and usage of FieldSpec Hand Held, refer to [32]. The spectral curve is shown in Figure 11. The spectral curves of the soil intersect with those of the lodging sunflower near the central wavelength of the red-edge band (717 nm), and similar results are also presented in reference [30]. This indicates that the reflectance of the lodging sunflower is similar to that of the soil in the red-edge band. Figure 9(d3,e3) are the classification results of the two models under the RGB+ red-edge band. Compared with Figure 9b, it is found that most of the lodging areas in Figure 9(d3,e3) are misclassified into background. It could be concluded that the negative effect of red-edge on sunflower lodging was caused by the similar reflectance of soil and lodging sunflower in the red-edge band.

The above analysis shows that the deep learning algorithm should also select valid data for classification, and it is not acceptable to input both valid and invalid data into the model without selection, which will affect the classification accuracy. In order to further explain the problem, lodging was extracted by using red-edge and red-edge + NIR. As shown in Table A1, the red-edge band had a poor effect on information extraction of lodging sunflower, and the classification accuracy of the model was gradually improved when effective bands were added. Therefore, for the deep learning model, the input band should not be added without selection during classification. The UAV remote sensing image has high resolution, which can be used to derive canopy surface models and reflect the height of the canopy. The lodging has a direct impact on the canopy height, so some authors have studied the lodging from the extraction of the canopy height. Extraction of lodging by canopy height generally focuses on the establishment of canopy height model and the extraction of lodging and lodging degree by threshold or supervised classification methods [12,29,42]. Compared with spectral features, canopy height features can describe the lodging degree, which has certain advantages. However, the accuracy of canopy height extraction has always been a factor limiting its application. If the problem of poor accuracy of canopy height extraction is solved, the combination of canopy height and deep learning model will promote the further development of lodging monitoring.

Figure 11. Spectral curves of lodging sunflower, soil, and normal sunflower.

4.2. Comparative Analysis of Related Studies

There have been a few similar studies before this study. For example, Song et al. [30] used the hue, saturation and value (HSV) sharpening module in ENVI software to interpolate low-resolution multispectral images and obtain high-resolution multispectral images as model input, and then added a jump connection and a conditional random field to the SegNet model to extract the lodging-related information of sunflowers. The authors studied the improved effect of deep learning on the extraction of lodging sunflower from the aspects of image fusion to improve the resolution of input multispectral data and the addition of jump connection and conditional random field in the SegNet model. However, when using multispectral data for classification, all obtained multispectral bands were directly input into the model for classification in this study, without considering the negative effects of red-edge bands (as shown in Section 4.1) on the extraction of lodging sunflower. The paper also does not take into account the problem of stitching trace when the model is used to predict the regional image.

In our study, we input different bands into the deep learning model and find that when the input band contains the red-edge band, the prediction accuracy of the two deep learning models is poor, as shown in Table 5. In the SegNet model, RGB + NIR + red-edge was used as the input, and the F1-scores of background, normal sunflower, and lodging sunflower were 87.87%, 85.73%, and 23.93%, respectively. When the red-edge band was removed from the input, the F1-score increased to 91.08%, 88.58%, and 51.68%, respectively. Therefore, the classification accuracy of the deep learning model can be improved by selecting the multi-spectral band. We also discussed the effect of introducing ignoring-edge-related information on sunflower lodging monitoring in the region predicted by the deep learning model. By comparing Tables 5 and 6, the prediction method ignoring edge-related information was adopted to extract background, normal sunflower, and lodging sunflower and the F1-score was about 2–5% higher than that of the general prediction method. Therefore, the prediction method ignoring edge-related information can effectively improve the prediction accuracy of the model.

4.3. Limitations of SegNet and U-Net in Extraction Sunflower Lodging

As shown in Table 5, in test area 2, the F1-score of extraction results of background, normal sunflower and lodging sunflower by SegNet in the RGB + NIR band were 91.08%, 88.58%, and 51.68%, respectively. The F1-score of U-Net extraction results were 89.58%, 86.43%, and 45.37%, respectively. In test area 1, the F1-score of SegNet extraction results were 84.32%, 94.59%, and 84.82%, respectively. The F1-score of U-Net extraction results were 82.46%, 93.19%, and 78.85%, respectively. In literature [30], the extraction accuracy of background, sunflower, and lodging sunflower was 91.7%, 89.2%, and 83.3%, respectively, which was at the same level as the extraction results in test area 1 in our study. However,

in test area 2, the extraction accuracy of lodging sunflower was poor. The sources of error include unrecognized lodging areas and the incorrect classification of other ground objects as lodging areas. To clarify the main sources of error, the results of predictions of the deep learning models in test areas were superimposed on the label data to clarify the sources of errors and the ability of the model to extract lodging-related information. Figure 12 shows the locations and specific information concerning the incorrect classification of sunflower lodging, as they can be used to locate incorrect partitions and analyze the cause of the incorrect classification. There were a large number of unrecognized lodging areas in the Figure 12a, and there were two main reasons for this. First, the coverage of sunflowers in the lodging area was low, and the entire area contained information on soil and plastic films after lodging that led to a failure of recognition of certain areas, as shown in Figure 12c. Second, the lodging was random in terms of location and extent. The area shown in Figure 12d features different degrees of lodging, and was mixed with the normal sunflowers, which made it difficult to identify the lodging area. In the background, sunflowers were mistakenly divided into lodging areas. There were two reasons for the incorrect classification. First, there were different degrees of salinization in interior plots of the study area, leading to significant inhomogeneity in sunflower growth. Images of areas with poor growth had similar characteristics to lodging areas, and could easily be mistaken for them as shown in Figure 12e. Second, owing to the sparse canopies of sunflowers in the region, although lodging did not occur after extreme weather, the stems tilted slightly to exhibit the characteristics of culms in the orthophoto map. They were thus easily incorrectly classified as lodging areas as shown in Figure 12f.

Figure 12. The error distribution of predicted lodging for two test areas and easily misclassified features images: (a) Test area 2 (SegNet with RGB + NIR); (b) test area 1 (SegNet with RGB + NIR); (c) lodging of low-coverage sunflowers; (d) mixed zone; (e) low-coverage sunflowers; (f) dew-stem area.

Compared with test area 2, test area 1 was more accurate in extracting lodging sunflower. By comparing the remote sensing images of the two test areas, it was found that in test area 1, the growth of sunflower was uniform, and the information of lodging sunflower could be basically extracted by the model. In test area 2, there were great differences in the growth of sunflower. Due to soil salinity, there were many plots of sunflower with poor growth in this area, and information of both normal sunflower and lodging sunflower existed in some lodging areas, which had a great impact on the lodging extraction of the model. By comparing and analyzing the results of extracting sunflower lodging in test area 1 and 2, we may conclude that the currently commonly used SegNet and U-Net models are not applicable to extracting lodging information in the area with more complex sunflower growth. In order to improve the applicability of the deep learning model for lodging extraction in the heterogeneous region of sunflower growth. Heterogeneity of sunflower in different plots can be obtained by evaluating the salinity of plots, and lodging extraction models can be constructed according to different heterogeneity, which may improve the applicability of the model for lodging extraction in plots with sunflower growth heterogeneity caused by salinity and alkali. The applicability of the deep learning model in lodging extraction in the mixed area can be improved by increasing the image resolution and adjusting the training samples in the mixed area. Due to the limited types of data acquired in the monitoring of sunflower lodging in this study, and the absence of land salinity and remote sensing image data with different resolutions, further studies on improving the applicability of the deep learning model in lodging extraction in heterogeneous areas of sunflower growth are restricted.

4.4. Summary and Prospect

The lodging-related information of sunflowers should include not only the area of lodging but also its degree, which is generally described by the lodging angle. Different degrees of lodging have different effects on the growth and yield of sunflowers. To identify the degree of lodging, some scholars have used structure-from-motion (SfM) technology to reconstruct crops in 3D and obtain their height to judge the severity of lodging [29,42]. Our experiments, in conjunction with previous studies, show that the reconstructed model of the study area was accurate over only a small range, and the error in the calculated crop elevation increased with the study area. In case of a limited spatial resolution, it is difficult to determine the specific height and lodging angle of sunflowers from remote sensing images. The study area was located in the Hetao irrigation area of Inner Mongolia Autonomous Region, and is part of the Yellow River irrigation area. The soil was rich in salt and alkalis, which enhanced the differences in growth among sunflower crops and increased the difficulty of identifying the degree of lodging.

In future work, research on sunflower lodging should be conducted from the following perspectives: (1) Collaborate the canopy height data with the deep learning model to establish a deep learning model that can evaluate the lodging degree. (2) A feature-screening method is needed to obtain more useful features and improve the accuracy of extraction.

5. Conclusions

In this study, UAV multi-spectral remote sensing technology and deep learning algorithms were used to extract lodging-related information on sunflower crops. The following conclusions can be drawn from the results:

1. Compared with the random forest method, the deep learning method has advantages in terms of the accuracy of classification of areas with sunflower lodging. Deep learning can be used to mine deep features of images and avoid the "salt and pepper phenomenon" in pixel-level classification. The classification accuracy of such methods was about 40% higher than that of the random forest method in experiments. However, the commonly used SegNet and U-Net models are not adequate for generalizing areas of sunflower lodging in cases of complex growth of the crop.

2. By using UAV multi-spectral images, the influence of multi-spectral band information based on RGB images on the extraction of lodging-related information for sunflowers was studied for a few combinations of bands. The results of extraction of two deep learning methods showed that the addition of the NIR band can increase the accuracy of classification whereas the addition of the red-edge band reduces it. Thus, while accuracy is improved by using more classification-related information, not all information can be directly used to classify, and inhibiting data need to be filtered out.
3. Compared with the traditional method, the proposed method for predicting lodging-related information for regional sunflowers that ignores edge-related information in images removed traces of stitching and improved the accuracy of classification by 2%. The results here can provide technical support for the accurate prediction of lodging-related information on regional sunflowers.

Author Contributions: Conceptualization, G.L. and W.H.; data curation, G.L., S.H., W.M., Q.M., and X.C.; formal analysis, G.L.; investigation, G.L., S.H., W.M., and X.C.; writing—original draft, G.L.; writing—review and editing, W.H., S.H., and Q.M. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Science Foundation of China (51979233), and the 111 Project (No. B12007).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank Haipeng Chen, Jiandong Tang, Jianyi Ao, and Chaoqun Li for their work on field data collection.

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

Abbreviations

UAV	Unmanned aerial vehicle
RGB	Red band, green band, blue band
NIR	Near-infrared band
GCPs	Ground control points
IoU	Intersect-over-union
OA	Overall accuracy
Val	Validation
Acc	Accuracy

Appendix A

Table A1. Results of deep learning for prediction of sunflower lodging.

Model	Information	F1-Score (%)			Intersection-over-Union (%)			OA (%)
		Background	Sunflower	Lodging	Background	Sunflower	Lodging	
SegNet	Red-edge	81.84	69.58	5.69	69.27	53.35	2.93	75.80
	Red-edge + NIR	90.48	87.22	30.97	82.61	77.34	18.32	87.06

References

1. Berry, P.M. Lodging Resistance cereal lodging resistance in Cereals cereal. In *Sustainable Food Production*; Christou, P., Savin, R., Costa-Pierce, B.A., Misztal, I., Whitelaw, C.B.A., Eds.; Springer: New York, NY, USA, 2013; pp. 1096–1110.
2. Rajkumara, S. Lodging in cereals—A review. *Agric. Rev.* **2008**, *1*, 55–60.
3. Berry, P.M.; Sterling, M.; Spink, J.H.; Baker, C.K.; Sylvester-Bradley, R.; Mooney, S.J.; Tams, A.; Ennos, A.R. Understanding and Reducing Lodging in Cereals. *Adv. Agron.* **2004**, *84*, 217–271.

4. Chauhan, S.; Darvishzadeh, R.; Lu, Y.; Boschetti, M.; Nelson, A. Understanding wheat lodging using multi-temporal Sentinel-1 and Sentinel-2 data. *Remote Sens. Environ.* **2020**, *243*, 111804. [\[CrossRef\]](#)
5. Han, L.; Yang, G.; Feng, H.; Zhou, C.; Yang, H.; Xu, B.; Li, Z.; Yang, X. Quantitative Identification of Maize Lodging-Causing Feature Factors Using Unmanned Aerial Vehicle Images and a Nomogram Computation. *Remote Sens.* **2018**, *10*, 1528. [\[CrossRef\]](#)
6. Zhang, L.; Zhang, H.; Niu, Y.; Han, W. Mapping Maize Water Stress Based on UAV Multispectral Remote Sensing. *Remote Sens.* **2019**, *11*, 605. [\[CrossRef\]](#)
7. Zhang, L.; Niu, Y.; Zhang, H.; Han, W.; Li, G.; Tang, J.; Peng, X. Maize Canopy Temperature Extracted from UAV Thermal and RGB Imagery and Its Application in Water Stress Monitoring. *Front. Plant Sci.* **2019**, *10*, 1270. [\[CrossRef\]](#)
8. Niu, Y.; Zhang, L.; Zhang, H.; Han, W.; Peng, X. Estimating Above-Ground Biomass of Maize Using Features Derived from UAV-Based RGB Imagery. *Remote Sens.* **2019**, *11*, 1261. [\[CrossRef\]](#)
9. Yao, X.; Liu, W.; Han, W.; Li, G.; Ma, Q. Development of Response Surface Model of Endurance Time and Structural Parameter Optimization for a Tailsitter UAV. *Sensors* **2020**, *20*, 1766. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Dai, J.; Zhang, G.; Guo, P.; Zeng, T.; Xue, J. Information extraction of cotton lodging based on multi-spectral image from UAV remote sensing. *Trans. Chin. Soc. Agric. Eng.* **2019**, *35*, 63–70.
11. Zhao, L.; Duan, Y.; Shi, Y.; Zhang, B. Wheat lodging identification using DSM by drone. *Chin. Agric. Inf.* **2019**, *31*, 36–42.
12. Mao, Z.; Deng, L.; Zhao, X.; Hu, Y. Extraction of Maize Lodging in Breeding Plot Based on UAV Remote Sensing. *Chin. Agric. Sci. Bull.* **2019**, *35*, 62–68.
13. Li, Z.; Chen, Z.; Wang, L.; Liu, J.; Zhou, Q. Area extraction of maize lodging based on remote sensing by small unmanned aerial vehicle. *Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 207–213.
14. Li, G.; Zhang, L.; Song, C.; Peng, M.; Zhang, Y.; Han, W. Extraction Method of Wheat Lodging Information Based on Multi-temporal UAV Remote Sensing Data. *Trans. Chin. Soc. Agric. Mach.* **2019**, *50*, 211–220.
15. Yang, M.; Huang, K.; Kuo, Y.; Tsai, H.; Lin, L. Spatial and Spectral Hybrid Image Classification for Rice Lodging Assessment through UAV Imagery. *Remote Sens.* **2017**, *9*, 583. [\[CrossRef\]](#)
16. Li, H.; Zhao, C.; Yang, G.; Feng, H. Variations in crop variables within wheat canopies and responses of canopy spectral characteristics and derived vegetation indices to different vertical leaf layers and spikes. *Remote Sens. Environ.* **2015**, *169*, 358–374. [\[CrossRef\]](#)
17. Chauhan, S.; Darvishzadeh, R.; Lu, Y.; Stroppiana, D.; Boschetti, M.; Pepe, M.; Nelson, A. Wheat Lodging Assessment Using Multispectral Uav Data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-2/W13*, 235–240. [\[CrossRef\]](#)
18. Kumpumaki, T.; Linna, P.; Lipping, T. Crop Lodging Analysis from Uas Orthophoto Mosaic, Sentinel-2 Image and Crop Yield Monitor Data. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 7723–7726.
19. Chauhan, S.; Darvishzadeh, R.; Boschetti, M.; Nelson, A. Estimation of crop angle of inclination for lodged wheat using multi-sensor SAR data. *Remote Sens. Environ.* **2020**, *236*, 111488. [\[CrossRef\]](#)
20. Liu, T.; Li, R.; Zhong, X.; Jiang, M.; Jin, X.; Zhou, P.; Liu, S.; Sun, C.; Guo, W. Estimates of rice lodging using indices derived from UAV visible and thermal infrared images. *Agric. For. Meteorol.* **2018**, *252*, 144–154. [\[CrossRef\]](#)
21. Yang, H.; Chen, E.; Li, Z.; Zhao, C.; Yang, G.; Pignatti, S.; Casa, R.; Zhao, L. Wheat lodging monitoring using polarimetric index from RADARSAT-2 data. *Int. J. Appl. Earth Obs.* **2015**, *34*, 157–166. [\[CrossRef\]](#)
22. Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 778–782. [\[CrossRef\]](#)
23. De, S.; Bhattacharya, A. Urban classification using PolSAR data and deep learning. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 353–356.
24. Yang, M.D.; Tseng, H.H.; Hsu, Y.C.; Hui, P.T. Semantic Segmentation Using Deep Learning with Vegetation Indices for Rice Lodging Identification in Multi-date UAV Visible Images. *Remote Sens.* **2020**, *12*, 633. [\[CrossRef\]](#)
25. Mdya, B.; Jgb, C.; Hui, P.; Hhta, B.; Ycha, B.; Ccs, C. Adaptive autonomous UAV scouting for rice lodging assessment using edge computing with deep learning EDANet. *Comput. Electron. Agric.* **2020**, *179*, 105817.
26. Hamidisepehr, A.; Mirnezami, S.; Ward, J. Comparison of object detection methods for crop damage assessment using deep learning. *Trans. ASABE* **2020**, *63*, 1969–1980. [\[CrossRef\]](#)
27. Zhao, X.; Yuan, Y.; Song, M.; Ding, Y.; Lin, F.; Liang, D.; Zhang, D. Use of Unmanned Aerial Vehicle Imagery and Deep Learning UNet to Extract Rice Lodging. *Sensors* **2019**, *19*, 3859. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Zheng, E.; Tian, Y.; Chen, T.; Branch, J. Region Extraction of Corn Lodging in UAV Images Based on Deep Learning. *J. Henan. Agric. Sci.* **2018**, *8*, 155–160.
29. Wilke, N.; Siegmann, B.; Klingbeil, L.; Burkart, A.; Kraska, T.; Muller, O.; Doorn, A.V.; Heinemann, S.; Rascher, U. Quantifying Lodging Percentage and Lodging Severity Using a UAV-Based Canopy Height Model Combined with an Objective Threshold Approach. *Remote Sens.* **2019**, *11*, 515. [\[CrossRef\]](#)
30. Song, Z.; Zhang, Z.; Yang, S.; Ding, D.; Ning, J. Identifying sunflower lodging based on image fusion and deep semantic segmentation with UAV remote sensing imaging. *Comput. Electron. Agric.* **2020**, *179*, 105812. [\[CrossRef\]](#)

31. Zhang, Z.; Wei, G.; Yao, Z.; Tan, C.; Wang, X.; Han, J. Soil Salt Inversion Model Based on UAV Multispectral Remote Sensing. *Trans. Chin. Soc. Agric. Mach.* **2019**, *50*, 151–160.
32. Ma, Q.; Han, W.; Huang, S.; Dong, S.; Chen, H. Distinguishing Planting Structures of Different Complexity from UAV Multispectral Images. *Sensors* **2021**, *21*, 1994. [[CrossRef](#)]
33. Svetnik, V. Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947–1958. [[CrossRef](#)] [[PubMed](#)]
34. Liaw, A.; Wiener, M.; Liaw, A. Classification and Regression with Random Forest. *R News* **2002**, *2*, 18–22.
35. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal.* **2015**, *39*, 640–651.
36. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
37. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)] [[PubMed](#)]
38. Wang, Z.; Zhou, Y.; Wang, S.; Wang, F.; Xu, Z. House building extraction from high resolution remote sensing image based on IEU-Net. *J. Remote Sens.* **2021**, in press.
39. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
40. Chamorro-Martinez, J.; Martinez-Jimenez, P. A comparative study of texture coarseness measures. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009; pp. 1337–1340.
41. Bi, J.; Chen, Y.; Wang, J.Z. A Sparse Support Vector Machine Approach to Region-Based Image Categorization. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; pp. 1121–1128.
42. Chu, T.; Starek, M.; Brewer, M.; Murray, S.; Pruter, L. Assessing Lodging Severity over an Experimental Maize (*Zea mays* L.) Field Using UAS Images. *Remote Sens.* **2017**, *9*, 923. [[CrossRef](#)]