

Article

Collaborative Wheat Lodging Segmentation Semi-Supervised Learning Model Based on RSE-BiSeNet Using UAV Imagery

Hongbo Zhi, Baohua Yang *  and Yue Zhu

School of Information and Computer, Anhui Agricultural University, Hefei 230036, China

* Correspondence: ybh@ahau.edu.cn

Abstract: Lodging is a common natural disaster during wheat growth. The accurate identification of wheat lodging is of great significance for early warnings and post-disaster assessment. With the widespread use of unmanned aerial vehicles (UAVs), large-scale wheat lodging monitoring has become very convenient. In particular, semantic segmentation is widely used in the recognition of high-resolution field scene images from UAVs, providing a new technical path for the accurate identification of wheat lodging. However, there are still problems, such as insufficient wheat lodging data, blurred image edge information, and the poor accuracy of small target feature extraction, which limit the recognition of wheat lodging. To this end, the collaborative wheat lodging segmentation semi-supervised learning model based on RSE-BiSeNet is proposed in this study. Firstly, ResNet-18 was used in the context path of BiSeNet to replace the original backbone network and introduce squeeze-and-excitation (SE) attention, aiming to enhance the expression ability of wheat lodging characteristics. Secondly, the segmentation effects of the collaborative semi-supervised and fully supervised learning model based on RSE-BiSeNet were compared using the self-built wheat lodging dataset. Finally, the test results of the proposed RSE-BiSeNet model were compared with classic network models such as U-Net, BiSeNet, and DeepLabv3+. The experimental results showed that the wheat lodging segmentation model based on RSE-BiSeNet collaborative semi-supervised learning has a good performance. The method proposed in this study can also provide references for remote sensing UAVs, other field crop disaster evaluations, and production assistance.



Citation: Zhi, H.; Yang, B.; Zhu, Y. Collaborative Wheat Lodging Segmentation Semi-Supervised Learning Model Based on RSE-BiSeNet Using UAV Imagery. *Agronomy* **2023**, *13*, 2772. <https://doi.org/10.3390/agronomy13112772>

Academic Editors: Gniewko Niedbala and Guangsheng Zhou

Received: 12 August 2023
Revised: 2 November 2023
Accepted: 3 November 2023
Published: 6 November 2023



Copyright: © 2023 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/>).

Keywords: lodging; segmentation; UAV; wheat; semi-supervised learning

1. Introduction

Wheat is an important crop, and its production process is often affected by different disasters [1]. Among them, lodging is one of the common obstacles in the growth process of wheat. It not only affects the synthesis and transportation of wheat organic matter but also affects the yield, quality, and harvest of wheat [2]. Therefore, the rapid identification of wheat lodging is the prerequisite for wheat disaster early warning, which is of great significance for wheat growth monitoring, yield measurement, disaster assessment, and post-disaster field management.

As an affordable, time-efficient, and flexible platform, UAVs can provide fast and accurate data sources for crop growth monitoring at different scales. Studies have shown that a UAV platform equipped with multiple sensors can achieve good results in the rapid acquisition of lodging disaster information and damage estimation [3]. In particular, the use of UAVs to obtain thermal infrared images, multispectral images, and visible light images provides data sources for UAV lodging monitoring. Among them, Liu et al. successfully extracted the lodging area of rice using the thermal infrared image acquired by UAV [4]. Zhao et al. successfully identified rice lodging areas using multispectral images acquired by UAVs [5]. Chauhan et al. used UAVs to obtain multispectral images of wheat at different degrees of lodging to realize the accurate lodging assessment [6]. Dai et al. used multispectral UAV images to extract cotton lodging areas [7]. Tian et al. used multispectral

images to better estimate rice lodging areas [8]. The above research shows that crop lodging recognition based on multispectral images and thermal infrared images obtained by UAVs has achieved good results. However, the cost of carrying multispectral sensors on UAVs is relatively high. Therefore, visible light images based on UAV platforms equipped with low-cost, high-resolution digital cameras have become a current application hotspot, providing rich data for crop lodging monitoring.

In fact, UAV remote sensing images have been successfully used in the research of identifying the lodging of different crops based on machine learning methods, including support vector machines, decision trees, watershed algorithms, random forest methods, etc. Among them, Rajapaksa used the support vector machine to classify the lodging of wheat and rape with respect to UAV images [9]. Yang et al. used a decision tree to classify rice lodging areas with respect to UAV images [10]. Cao et al. used the watershed algorithm to extract wheat lodging areas from UAV images [11]. The research mentioned above shows that the texture features extracted from UAV images provide the basis for different crop lodging recognition, which is susceptible to external environmental interference and affects the accuracy of lodging recognition. Therefore, there are still many challenges in accurately extracting grain features from UAV images to identify crop lodging.

Recently, with the wide application of deep learning, effective features are extracted from UAV images based on the deep neural network framework, including the local detail features of images and advanced semantic features of images, and this has provided important contributions to improving the accuracy of crop lodging extraction. Rice lodging recognition based on different deep neural network models has obtained good results, such as LodgeNet [12], EDANet [13], and FCN-AlexNet [14]. In addition, Han et al. proposed the SMOTE-ENN-XGBoost model to effectively identify corn lodging areas at the plot scale [15]. Song et al. proposed an improved semantic segmentation network, SegNet, to effectively extract sunflower lodging [16]. The above research shows that deep learning has good development potential and trends in the lodging extraction of various crops. In fact, for the identification of wheat lodging, many scholars and experts have also proposed a variety of deep learning models and achieved excellent results. Zhang et al. proposed the automatic extraction of wheat lodging areas based on DeepLabv3+ and transfer learning methods [17]. Yu et al. proposed the Lstm_PSPNet model to accurately identify wheat lodging areas [18]. Zhang et al. compared traditional machine learning methods and deep learning methods, and they determined that the accuracy of wheat lodging extraction based on GoogLeNet performed best among several methods [19]. Mardanisamani et al. proposed a wheat lodging recognition model based on DCNN architecture [20]. In short, the method based on a deep neural network has good feasibility in the identification of field crop lodging. However, these models still have certain limitations. To this end, some studies have proposed a U-shaped network model to extract crop lodging [21]. The U-shaped structure model integrates the features of different levels of the backbone network via encoding and decoding layer by layer, which is beneficial to improving the identification of wheat lodging [22,23]. However, the U-shaped model has weak boundary detection capabilities, which may result in inaccurate boundary detection results for wheat lodging.

A bilateral segmentation network (BiSeNet) can obtain rich spatial information and a larger receptive field, providing new ideas for wheat lodging feature extraction [24]. However, there are few studies on the identification of wheat lodging areas based on the BiSeNet network. In particular, for UAV wheat field images with large intra-class variations in wheat and small inter-class differences in lodging, it is not clear whether it can be accurately segmented. In addition, model training based on deep convolutional neural networks relies on a large amount of labeled data [25]. In particular, the acquisition of pixel-level wheat lodging labels requires manual labeling, which is time-consuming and laborious, thus affecting the efficiency of wheat lodging recognition. With the application of semi-supervised learning methods, this method brings new vitality to tedious data labeling. Some studies have shown that semi-supervised learning (that is, the model is trained using both labeled and unlabeled data) achieves good results in the semantic

segmentation of crops [26]. Nong et al. used semi-supervised learning to better segment weeds and crops [27], and Pérez-Ortiz et al. used semi-supervised learning to successfully detect sunflower crops [28]. Although semi-supervised learning has achieved good results in the lodging identification of other crops, it has been rarely studied in wheat lodging monitoring. Therefore, it is necessary to explore whether BiSeNet combined with semi-supervised learning methods can improve wheat lodging segmentation accuracy.

In view of the above problems, a wheat lodging segmentation model based on semi-supervised learning is proposed to improve the segmentation effect of wheat lodging. The purpose of this research is to (1) propose an optimized BiSeNet wheat lodging segmentation model, aiming to improve the accuracy of wheat lodging segmentation; (2) compare the result of the wheat lodging segmentation model based on the cooperative training of semi-supervised learning and fully supervised learning, aiming to reduce the cost of model training; and (3) verify the RSE-BiSeNet wheat lodging segmentation model proposed in this study using the self-built wheat lodging dataset and comparative experiments.

2. Materials and Methods

2.1. Data Sources

2.1.1. Data Collection

The experimental site is located at the National Agricultural Science and Technology Innovation and Integration Demonstration ($31^{\circ}25' \sim 31^{\circ}42' \text{ N}$, $117^{\circ}09' \sim 117^{\circ}16' \text{ E}$), Lujiang County, Hefei, China. In this study, the experimental field wheat grain filling stage was carried out on 7 May 2021. The UAV is a DJI Phantom 3 Pro (SZ DJI Technology Co., Shenzhen, China) with a maximum flight time of about 25 min and an imaging resolution of 4000 pixels \times 3000 pixels. To obtain the wheat area accurately, the flight height of the UAV was set to 30 m, the flight speed was set to 3 m/s, the heading overlap rate was set to 70%, the side overlap rate was set to 75%, and 300 original images were obtained. Samples of lodging and non-lodging wheat are shown in Figure 1.

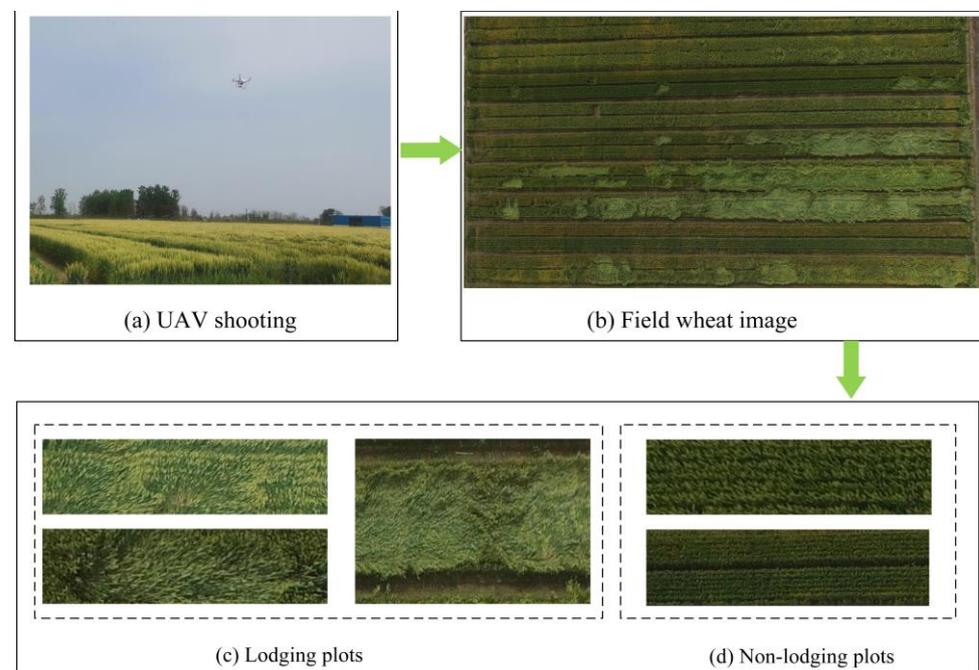


Figure 1. Example of wheat lodging and non-lodging.

2.1.2. Labeling of Wheat Lodging Areas

To mark the wheat lodging area, we used the Pix4Dmapper v4.4 software to perform orthorectification and image stitching on the collected images in order to obtain digital orthophotos of a large area of wheat lodging. Then, Labelme software (<http://labelme>.

csail.mit.edu/Release3.0/, accessed on 10 July 2023) was used to label the accommodation area, including accommodation, non-accommodation, and background, via field surveys. Among them, the background includes land, roads, etc., as shown in Figure 2.

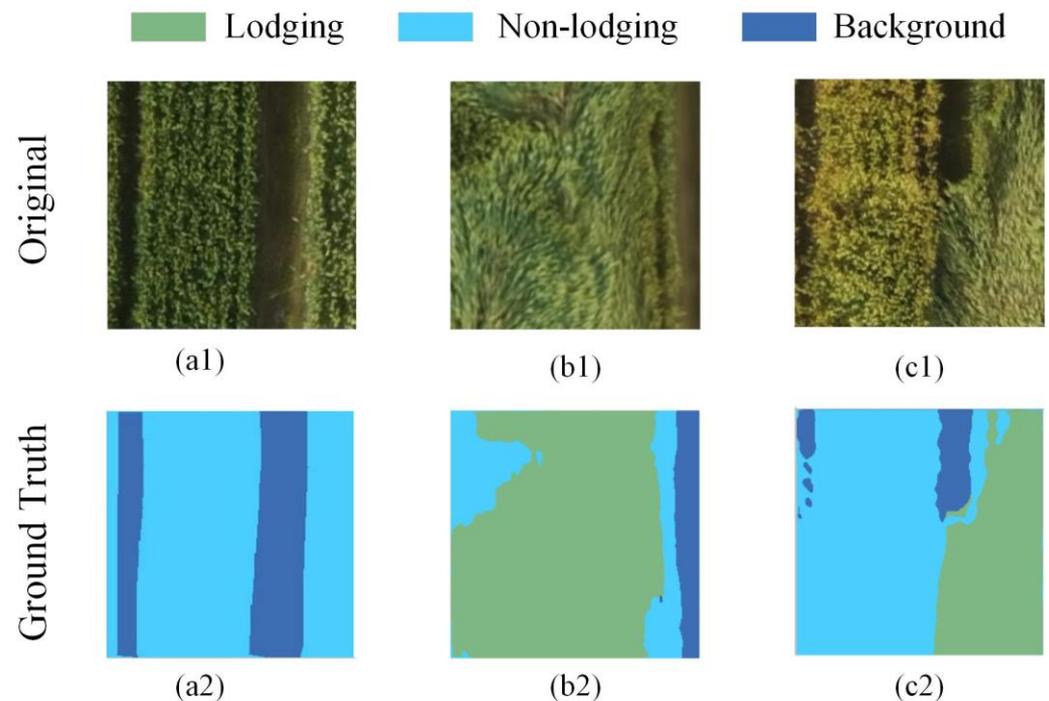


Figure 2. Wheat lodging and non-lodging distribution of study area: (a1–c1) original images and (a2–c2) ground truth images.

2.1.3. Dataset Construction

To speed up the training process of the model, the samples of the training set and the validation set were cropped to 256 pixels \times 256 pixels. Then, image enhancement processes such as noise, mirroring, rotation, and blurring were performed on the cropped training set images to improve the overall diversity of the dataset. After the data augmentation operation, the dataset images were expanded from the original 300 images to 1500 images, and the training set (1050 images), validation set (300 images), and test set (150 images) were randomly divided according to the ratio of 7:2:1.

2.2. Technical Flowchart of This Study

Figure 3 shows the technical route of this research study, which mainly includes data acquisition and processing, model training, model testing, and evaluation. Firstly, the wheat lodging images obtained by the UAV were spliced and split, and images with 256 pixels \times 256 pixels were obtained. The following processes were carried out: data annotation, data enhancement, dataset division, etc. Thus, wheat lodging training, validation, and test sets were obtained. Secondly, the model was trained using the training set, and an RSE-BiSeNet model with better weights was obtained. Finally, the test set was used for model testing and evaluation. In addition, to better obtain a robust model, ablation experiments, model comparisons, and semi-supervised learning comparison experiments were conducted.

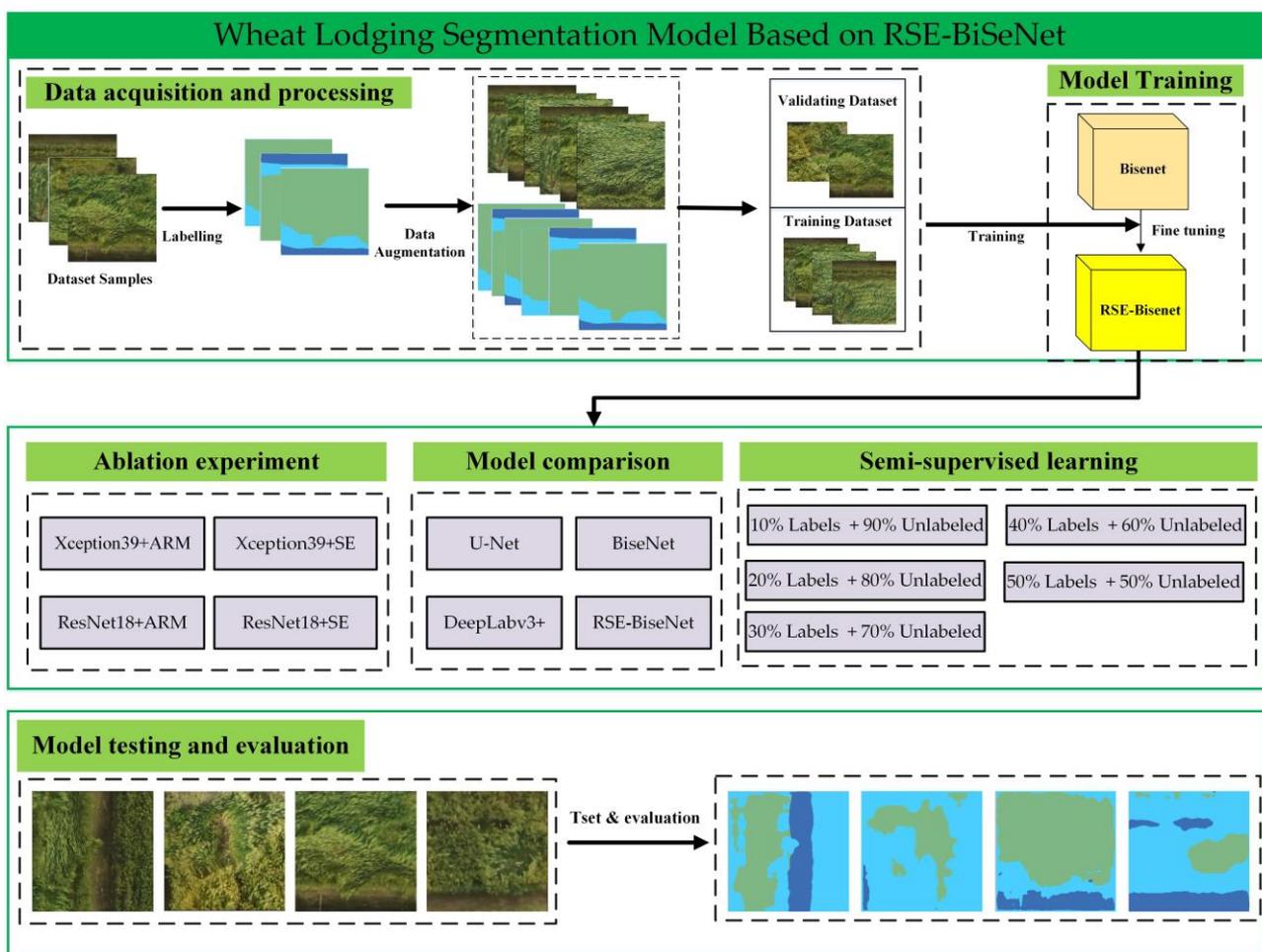


Figure 3. Technology flowchart of this study.

2.2.1. Wheat Lodging Segmentation Model Based on RSE-BiSeNet

The bilateral semantic segmentation network is an effective segmentation method proposed by Yu et al., which extracts high-dimensional nonlinear features and low-dimensional spatial features through two feature extraction paths, including spatial and context paths, so that the network has a broad receptive field and is rich in spatial feature information [24]. Among them, the key to the spatial path is to maintain the size of the input image and obtain rich spatial information. The context path mainly optimizes the output context’s semantic feature information by using the attention refinement module (ARM), aiming to obtain a sufficiently large receptive field.

To improve the extraction effect of wheat lodging features, a wheat lodging segmentation model based on the RSE-BiSeNet network was proposed in this study. Different from the original BiSeNet network, ResNet-18 is used as the backbone network in the context path of the RSE-BiSeNet network to replace the Xception39 of the original BiSeNet, and the squeeze-and-excitation (SE) [29] module is introduced to realize the fusion of channel information and multi-scale features. The specific structure of the RSE-BiSeNet network is shown in Figure 4.

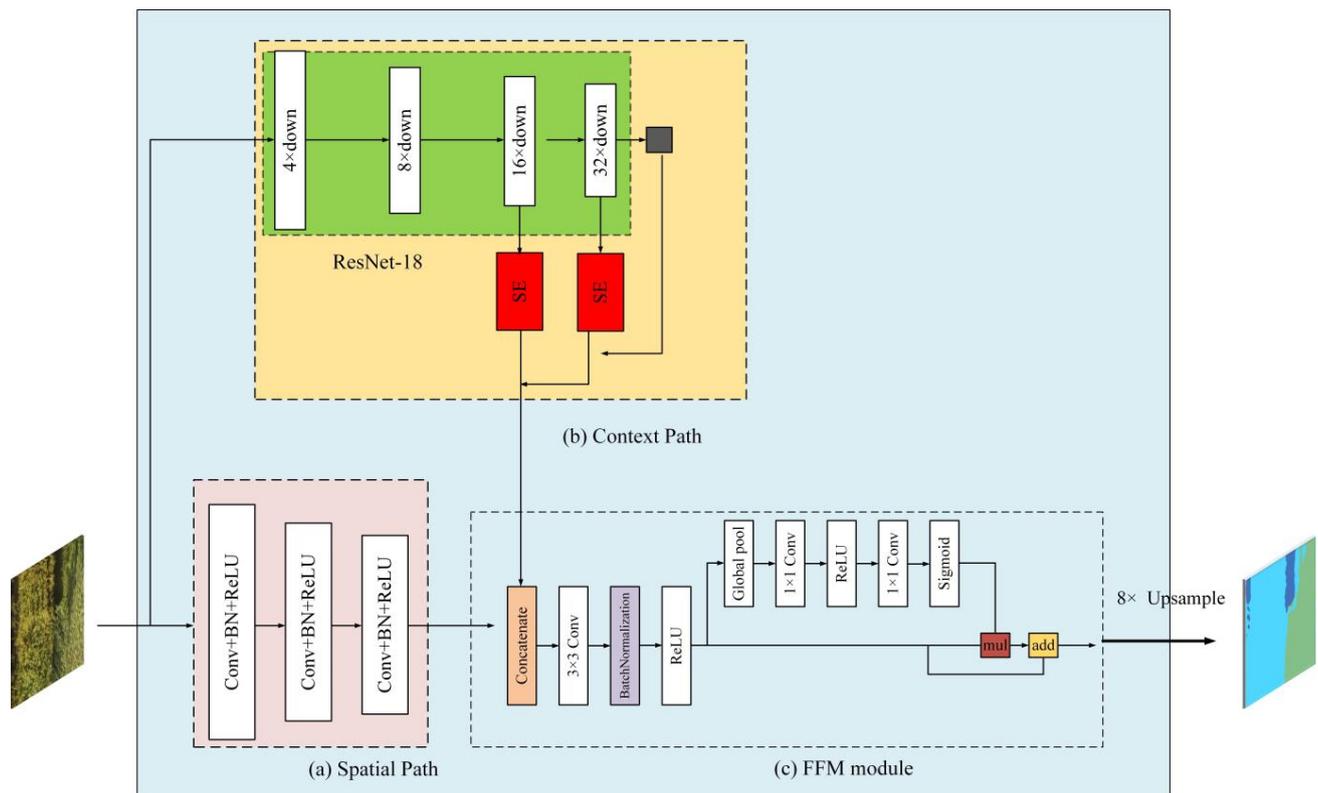


Figure 4. Overall structure of RSE-BiSeNet model.

As observed in Figure 4, the network is mainly composed of two parts: the spatial path and the context path. The role of the former is to maintain the size of the input wheat UAV image and obtain spatial information on wheat lodging. The role of the latter is to obtain a sufficiently large wheat lodging receptive field to extract high-dimensional nonlinear features. In addition, the spatial path of the RSE-BiSeNet network model contains three convolutional blocks. Each convolutional block consists of a convolutional (3×3) layer with a stride of 2, a batch normalization layer, and a ReLU activation layer. Therefore, the input is a $256 \times 256 \times 3$ wheat UAV image in this study, and the size of the feature map after the spatial path is $32 \times 32 \times 3$. Among, the feature map outputted by the spatial path pays more attention to the contour boundary of the wheat lodging area, ignoring the semantic features. Therefore, it is necessary to combine the semantic features obtained in the context path to complete the extraction of the wheat lodging area.

To obtain a larger receptive field, the lightweight ResNet-18 residual network is used as the basic network in the context path of RSE-BiSeNet for feature extraction, and a global average pooling layer is used at the end of the model. On the one hand, the global context information provides a larger receptive field, enabling the network to learn more lodging feature information. On the other hand, ResNet-18 can not only quickly determine the lodging area but also obtain rich advanced semantic features via error-reducing multi-layer backpropagation.

In addition, to refine the feature map of each stage more effectively and improve its ability to capture lodging information in wheat images, the SE module is introduced in the context path, which enhances or weakens different channels by learning the relationship between channels, so that the constructed network can pay more attention to the lodging area and significantly improve the inference efficiency of the network.

2.2.2. Semi-Supervised Learning

Semi-supervised learning mainly uses labeled data to extract information from unlabeled data, aiming to make up for the problem of low efficiency in model training when

there is less labeled data and thereby further improving the overall performance of the model [30]. Firstly, the model is trained on a small labeled dataset. Secondly, the trained model is used to predict unlabeled data, thereby creating corresponding pseudo-labeled images. Thirdly, the obtained pseudo-labels are added to the existing labeled dataset to obtain a new training dataset. The updated training set is used as input to the model for training, and the above operations are performed repeatedly until all unlabeled data generates pseudo-labeled images. Finally, a semi-supervised learning segmentation model was obtained. Figure 5 shows the process of collaborative wheat lodging segmentation semi-supervised learning in this study.

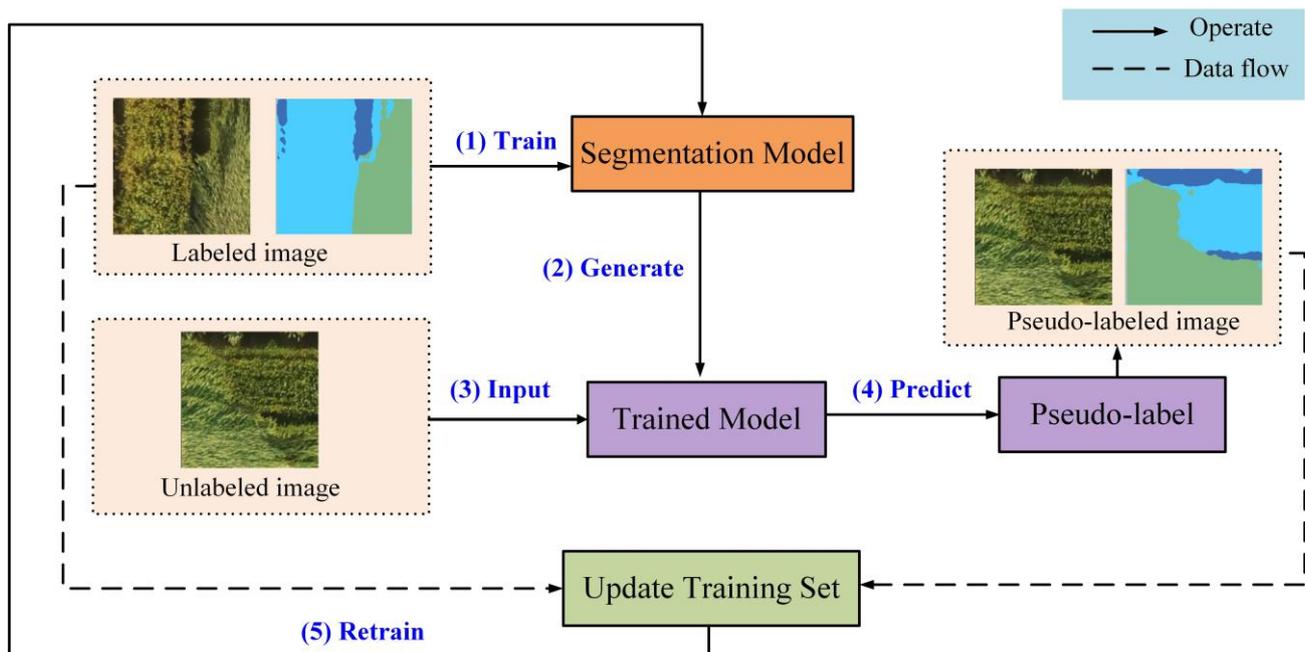


Figure 5. Flowchart of semi-supervised learning.

2.2.3. Model Evaluation Index

To effectively assess the accuracy of the wheat lodging segmentation model, on the one hand, precision and recall were used to evaluate different categories, and on the other hand, F1 (F1-score) and mean intersection over union (mIoU) were used to evaluate the model.

Among them, precision refers to the proportion of the correct number of samples predicted as the correct category, as shown in Formula (1).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

Recall refers to the proportion of correct predictions among all samples that are actually true, as shown in Formula (2).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

The F1-score is the harmonic mean of precision and recall, which reflects the comprehensive performance of wheat lodging segmentation, as shown in Formula (3).

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

mIoU is the result of summing and averaging the ratios of the intersection and union of the predicted results of each class and the true labels, as shown in Equation (4).

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN + FP + TP} \quad (4)$$

In the formula, TP refers to the correct segmentation of the wheat lodging area, which results in the actual wheat lodging area. TN refers to the correct segmentation of non-wheat lodging areas, which results in actual non-wheat lodging areas. FP refers to the correct segmentation of wheat lodging areas, which are actually non-wheat lodging areas. FN refers to the correct segmentation of non-wheat lodging areas, which are actually wheat lodging areas. $k + 1$ is the number of categories.

3. Results

3.1. Environmental Settings

In all the experiments, the training and testing of the model are based on the same hardware platform, and the experimental environment configuration parameters are as follows: The operating system is Windows 10 Professional 64-bit, the central processing unit is Intel Core i7-8700 CPU@3.20 GHz, and the graphics processor is NVIDIA GeForce RTX 2080 with 8 GB of video memory. The training of the model is based on the TensorFlow + Keras deep learning framework.

3.2. Results of Wheat Lodging Segmentation with Semi-Supervised Learning

To verify the impact of semi-supervised learning on the wheat lodging segmentation model, we reconstructed the wheat lodging training set into five different proportions of labeled data and unlabeled data. Among them, 10%, 20%, 30%, 40%, and 50% of the training set are set as labeled data, and the remaining 90%, 80%, 70%, 60%, and 50% are set as unlabeled data. The above five different training sets were input into the RSE-BiSeNet network model for training and testing, respectively. The results are shown in Table 1.

Table 1. Comparing the performance of RSE-BiSeNet model in semi-supervised learning methods.

Method of Supervised Learning	Number of Labels	Number of Unlabeled	Labels Proportion	Precision (%)	Recall (%)	F1-Score (%)	mIoU (%)
Semi-supervised learning	120	1080	10%	76.9	73.1	74.4	68.9
	240	960	20%	81.3	74.9	80.1	76.2
	360	840	30%	83.7	80.1	83.9	81.4
	480	720	40%	89.5	83.8	89.3	84.1
	600	600	50%	92.9	89.4	90.3	87.6

As observed in Table 1, when the labeled data account for 10%, 20%, 30%, 40%, and 50% of the training set, the F1-score of the wheat lodging segmentation model based on the semi-supervised learning method reaches 74.4%, 80.1%, 83.9%, 89.3%, and 90.3%. Experimental results show that as the number of labels increases, the effect of the wheat lodging segmentation model gradually improves. To visually display the comparative results of semi-supervised learning methods, the effects of several test data are randomly selected, as shown in Figure 6. As observed in Figure 6, when the label data accounted for 10%, the wheat lodging segmentation results exhibited obvious boundary blurs, and there is also the problem of misidentifying the non-lodging areas of wheat as lodging areas of wheat. This may be because there are less label data, which leads to an increase in the ratio of incorrect information to correct information, causing the edge of the segmentation result of wheat lodging to deviate from the boundary of ground truth. When the proportion of label data is 20%, the boundary of wheat lodging area segmentation is not fine enough, and some under-segmentation occurs. When the proportion of label data is 30% and 40%, the segmentation results are relatively similar, and the segmentation results of the wheat

lodging area are relatively clear, with only a little over-segmentation. When the label data accounted for 50%, the addition of pseudo-label data and the expansion of the training set significantly improved the segmentation effect of wheat lodging areas via the method based on semi-supervised learning, and the results of wheat lodging segmentation were closer to the ground truth.

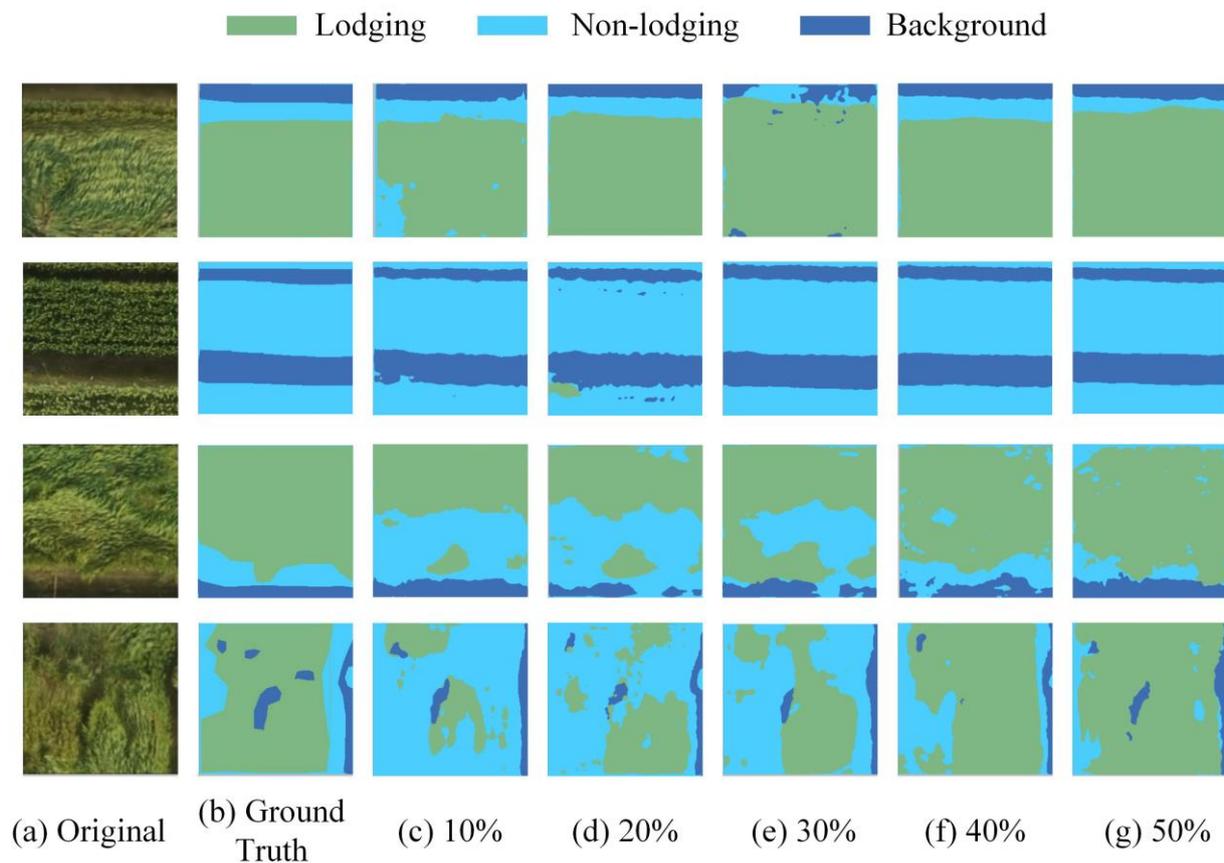


Figure 6. Visualization of segmentation results with different label proportions.

4. Discussion

4.1. The Influence of Attention Mechanism and Backbone Network on Model Performance

To verify the effectiveness of the proposed RSE-BiSeNet model, it is necessary to conduct ablation experiments to verify the impact of different modules on the model. The following experiments were designed: that is, the ablation experiments were performed on the backbone network and the attention mechanism module. In the experiment, 1200 images from the self-built wheat lodging dataset were used as the training set and inputted to the RSE-BiSeNet model. The training process based on the RSE-BiSeNet model was run in the same experimental environment, and the training parameters were also the same. The experimental results are shown in Figure 7.

It can be observed in Figure 7 that in the context path of the BiSeNet network, the results of wheat lodging segmentation based on different backbone networks and different attention mechanism modules are different. Regardless of the backbone network with ResNet18 or Xception39 as the context path, the segmentation effect of adding the SE attention module is better than that of the ARM attention module. Among them, when Xception39 is used as the backbone network, the F1-score of the wheat lodging segmentation model with SE is 90.6%, and the mIoU value is 86.8%, which are 0.6% and 3.2% higher than those with ARM. When ResNet18 is used as the backbone network, the F1-score of the wheat lodging segmentation model with SE is 93.1% and the mIoU is 89.2%, which are 1.5% and 0.9% higher than those with ARM. In fact, the ARM module mainly guides

feature learning via global average pooling, which can easily lose the details of wheat lodging, resulting in a decrease in accuracy. However, the SE module can learn features of different scales, especially the wheat lodging location information and detail features, thereby improving the ability of the network to extract wheat lodging areas.

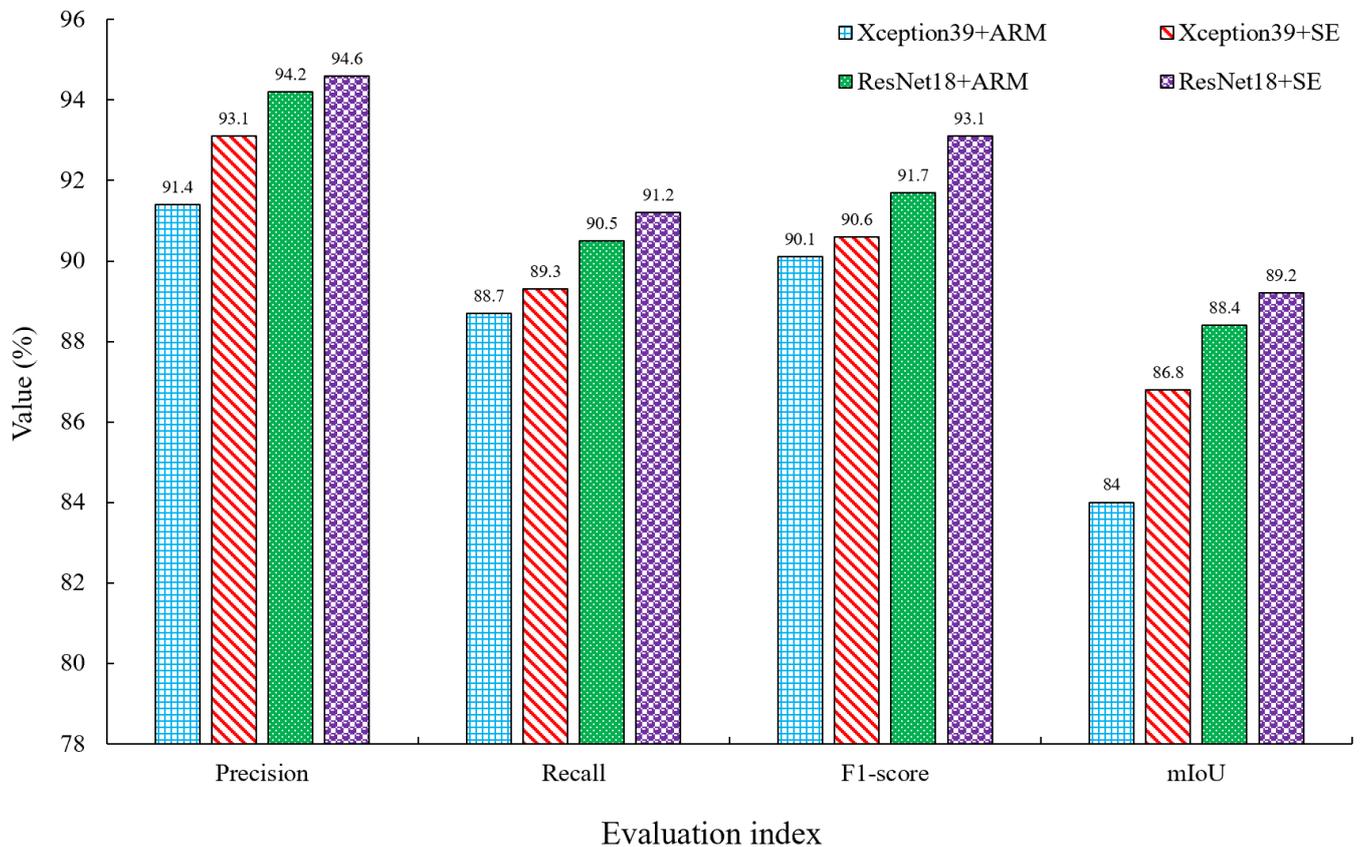


Figure 7. Comparing the performance of different backbone networks.

4.2. Comparison of Wheat Lodging Segmentation Results Based on Different Models

Semantic segmentation models with different network structures can affect the segmentation accuracy of wheat lodging and non-lodging areas. Therefore, the performance difference of classical semantic segmentation models applied on the wheat lodging image dataset (training set: 1200 images) was evaluated, and the effectiveness of the RSE-BiSeNet model was further demonstrated, which was compared with classical semantic segmentation algorithms, including U-Net [21,22], and DeepLabv3+ [23], BiSeNet [24]. All experiments were carried out on the same configuration of the experimental platform, and the hyperparameters in the models were the same. The results of wheat segmentation based on different models are shown in Table 2.

Table 2. Comparison of results of different segmentation models.

Models	Precision	Recall	F1-Score	mIoU
U-Net	89.7	86.2	88.3	80.5
BiseNet	91.4	88.7	90.1	84
DeepLabv3+	92.1	89.4	91.9	86.7
RSE-BiSeNet	94.6	91.2	93.1	89.2

It can be observed in Table 2 that the F1-score value of the wheat lodging segmentation model based on the U-Net network is 88.3%, the mIoU is 80.5%, and the F1-score is lower among the four segmentation models. The effect of the wheat lodging segmentation model

based on the BiSeNet is better than that based on U-Net. Among them, the F1-score of the segmentation model based on BiSeNet is 90.1% and the mIoU value is 84%, which are 2% and 4.2% higher than those based on U-Net. The effect of the segmentation model based on DeepLabv3+ is close to that based on BiSeNet. Among them, the F1-score value of the segmentation model based on DeepLabv3+ is 91.9%, and the mIoU value is 86.7%, which are 2% and 3.1% higher than those based on BiSeNet. Compared with BiSeNet and DeepLabv3+, the performance of the segmentation model based on RSE-BiSeNet has increased by 3.2% and 1.3% relative to the F1-score and 5.8% and 2.8% relative to mIoU, respectively. The F1-score and mIoU of the wheat lodging segmentation model based on RSE-BiSeNet are superior to those of the other models; the F1-score reaches 93.1%, and the mIoU reaches 89.2%, indicating that the model proposed in this study is more suitable for extracting wheat lodging regions.

To demonstrate the effect of wheat lodging segmentation based on different networks, four images were randomly selected to compare the differences in model performances. The results are shown in Figure 8.

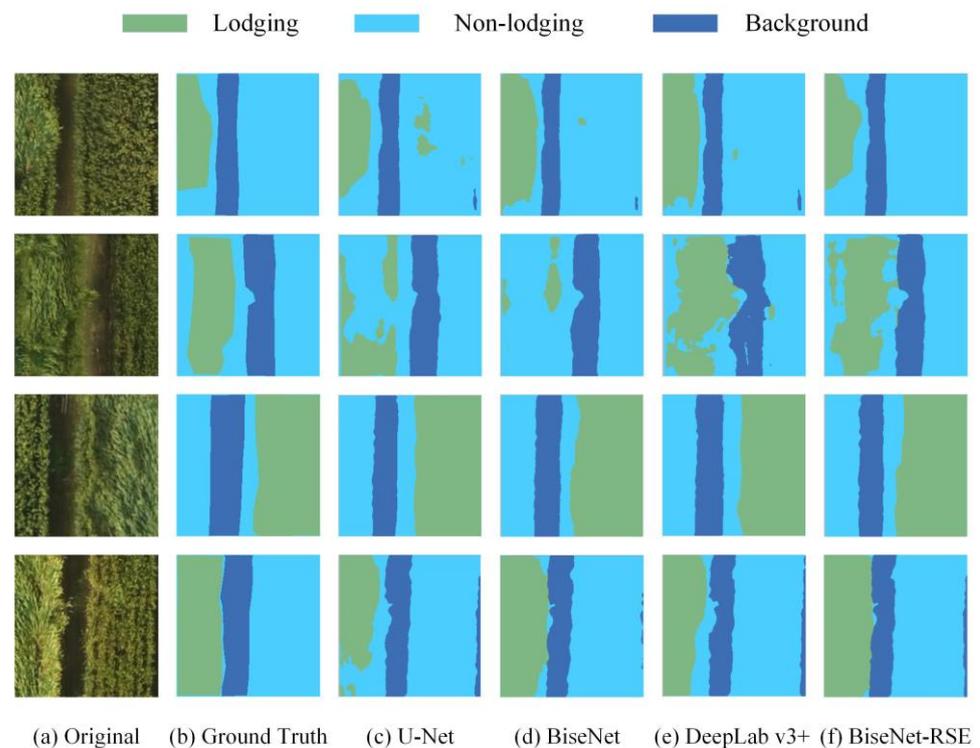


Figure 8. Visualization of segmentation results for different models.

It can be observed in Figure 8 that the segmentation models based on U-Net performed poorly, and the segmentation effect in the edge detail area was not ideal, and there were many mis-segmentation phenomena. For example, many edge parts of the wheat field are mis-segmented into lodging. The segmentation effect of the model based on BiSeNet is better than U-Net. Although there are fewer false segmentations, the edges of the lodging area are still not accurately identified. The wheat lodging segmentation results based on BiSeNet and DeepLabv3+ are similar, and they still need to be further improved in the extraction of wheat lodging details. The wheat lodging segmentation effect based on RSE-BiSeNet proposed in this study is the best among several models, and the segmentation results for the lodging wheat area, non-lodging wheat area, and other background areas are relatively accurate, which can make up for the limitations of the BiSeNet model. The RSE-BiSeNet model proposed in this study not only extracts the spatial detail information of wheat lodging but also extracts semantic context information from images acquired by drones. On the one hand, the small-step spatial path of the RSE-BiSeNet model can generate

higher-resolution feature maps by retaining spatial position information. On the other hand, the fast down-sampling semantic path of the RSE-BiSeNet model can obtain considerable receptive fields. Therefore, the model performs well in the semantic segmentation task of wheat lodging in natural scenes.

4.3. Comparison of Wheat Lodging Segmentation Results Based on Semi-Supervised Learning and Fully Supervised Learning

Figure 9 shows that with the same amount of labeled data, the evaluation index of semi-supervised learning is close to that of fully supervised learning, and it can effectively utilize a large amount of unlabeled data, thereby further improving the segmentation performance of the model. When the proportion of labeled data is small, the advantages of the semi-supervised learning method are more obvious, and when the proportion of labeled data is high, the improvement of the semi-supervised learning method is smaller than that of the fully supervised learning method; moreover, the four indicators of the model are relatively close to each other. As the number of labels gradually increases, the accuracy of the segmentation models of both methods gradually increases. The possible reason is that when there is a large amount of labeled data, the segmentation network is well trained, and the segmentation accuracy of the model also tends to be stable.

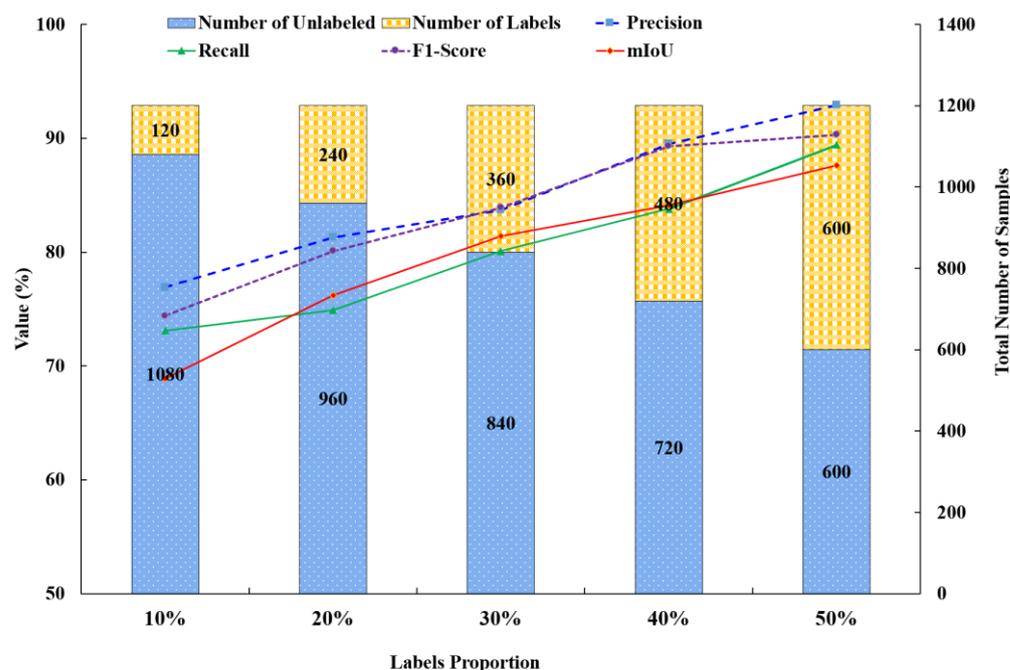


Figure 9. Visualization of segmentation results for semi-supervised learning.

In fact, with the increase in the cost of labeled data, semi-supervised learning methods have gradually received the attention of many scholars. Research has shown that semi-supervised learning achieves good results with respect to semantic segmentation in the agricultural field. Fourati et al. utilized semi-supervised learning to achieve good results in the detection of wheat ears [31]. Najafian et al. only used a small amount of label data to achieve the precise monitoring of wheat ears [32]. In the process of semantic segmentation, non-supervised and semi-supervised learning also achieved good results [33]. Pauletto et al. used semi-supervised learning to conduct model training in order to obtain a better segmentation effect [34]. Ubbens et al. used non-supervised learning relative to semantic segmentation and counting organs [35]. The above research shows the efficiency of semi-supervised learning.

To verify the effect of fully supervised learning, we compared the performance of wheat lodging segmentation under fully supervised learning methods. All experiments

were based on the same experimental environment parameter settings and used the same amount of labeled data as in semi-supervised learning. The test results are shown in Table 3.

Table 3. The performance of the RSE-BiSeNet model in fully supervised learning methods.

Method of Supervised Learning	Number of Labels	Number of Unlabeled	Labels Proportion	Precision (%)	Recall (%)	F1-Score (%)	mIoU (%)
Fully supervised learning	120	—	—	74.5	71.1	72.8	67.8
	240	—	—	79.2	73.4	78.8	75.4
	360	—	—	81.8	78.7	81.6	79.2
	480	—	—	86.4	81.8	85.6	82.4
	600	—	—	91.4	87.1	88.5	86.9
	1200	—	—	94.6	91.2	93.1	89.2

It can be observed in Table 3 that under the condition of the same number of labels, the evaluation index of the wheat lodging segmentation model based on fully supervised learning showcased superior performance. In fact, it is clear from Tables 1 and 3 that when the number of labels is the same, the evaluation index of the collaborative wheat lodging segmentation semi-supervised learning method is better than that of the fully supervised learning method. When label data accounted for 10%, the F1-score based on the semi-supervised learning method was 74.4%, and the mIoU was 68.9%, which were 2.2% and 1.6% higher than those based on the fully supervised learning method. When the proportion of labeled data is 20–40%, the models based on two different supervised learning methods perform well. As the amount of labeled data increases, the gap between the performance of models based on semi-supervised learning and fully supervised learning methods gradually decreases. In particular, when the proportion of labeled data reaches 50%, the F1-score of the model based on the fully supervised learning method is 88.5%, and the mIoU value is 86.9%. The F1-score of the model based on the semi-supervised learning method reaches 90.3% and the mIoU reaches 87.6%, which were 2% and 0.8% higher than those of fully supervised learning. Experiments show that model training with semi-supervised learning only requires half of the manually labeled image samples to achieve good segmentation performance.

5. Conclusions

To realize the accurate identification of wheat lodging based on the UAV platform, a wheat lodging segmentation model based on RSE-BiSeNet was proposed, which was trained, verified, and tested using the self-built wheat dataset. Moreover, the segmentation effects of wheat lodging were compared with different supervised learning methods, including semi-supervised learning and fully supervised learning. The experimental results showed that collaborative semi-supervised RSE-BiSeNet learning can effectively improve the semantic segmentation accuracy of wheat lodging images. In particular, the collaborative BiSeNet segmentation semi-supervised learning method can make full use of unlabeled wheat lodging images, and the model has good robustness in the case of different proportions of labeled data. Among them, the number of labeled images reaches 50% of the wheat lodging training set, and the F1-score of the wheat lodging segmentation model based on RSE-BiSeNet can reach 90.3%, which provides technical support for growth monitoring and the disaster prediction of other field crops.

At present, the collaborative wheat lodging segmentation semi-supervised learning model proposed in this study can not only reduce the cost of manual labeling but also utilize a large amount of unlabeled data. In the future, we will carry out unsupervised learning and weakly supervised learning to meet the needs of mining useful information from unlabeled data in complex agricultural scenarios, thereby improving model representation and generalization capabilities. We will also enhance the universality of the model, combine it with digital cameras and other sensors carried by UAV, and apply the model to more spatial scales and more crop growth monitoring fields.

Author Contributions: Writing and software: H.Z.; writing—review and editing: B.Y.; methodology: Y.Z. and H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Major Science and Technology Projects in Anhui Province (202203a06020007), the Select the Best Candidates to Undertake Key Research Project of common technologies in Hefei City (GJ2022QN03), and the Open Project of Jiangsu Key Laboratory of Information Agriculture (Grant No. 15266).

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

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

References

1. Yang, B.; Gao, Z.; Gao, Y.; Zhu, Y. Rapid detection and counting of wheat ears in the field using YOLOv4 with attention module. *Agronomy* **2021**, *11*, 1202. [[CrossRef](#)]
2. Li, Y.; Yang, B.; Zhou, S.; Cui, Q. Identification lodging degree of wheat using point cloud data and convolutional neural network. *Front. Plant Sci.* **2022**, *13*, 968479. [[CrossRef](#)] [[PubMed](#)]
3. Biswal, S.; Chatterjee, C.; Mailapalli, D.R. Damage Assessment Due to Wheat Lodging Using UAV-Based Multispectral and Thermal Imageries. *J. Indian Soc. Remote Sens.* **2023**, *51*, 935–948. [[CrossRef](#)]
4. 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](#)]
5. 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](#)]
6. Chauhan, S.; Darvishzadeh, R.; Lu, Y.; Stroppiana, D.; Boschetti, M.; Pepe, M.; Nelson, A. Wheat lodging assessment using multispectral UAV data. The International Archives of the Photogrammetry. *Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 235–240.
7. Dai, J.G.; Zhang, G.S.; Guo, P.; Zeng, T.J.; Cui, M.; Xue, J.L. Information extraction of cotton lodging based on multi-spectral image from UAV remote sensing. *Trans. CSAE* **2019**, *35*, 63–70.
8. Tian, M.; Ban, S.; Yuan, T.; Ji, Y.; Ma, C.; Li, L. Assessing rice lodging using UAV visible and multispectral image. *Int. J. Remote Sens.* **2021**, *12*, 1–18. [[CrossRef](#)]
9. Rajapaksa, S.; Eramian, M.; Duddu, H.; Wang, M.; Shirtliffe, S.; Ryu, S.; Josuttis, A.; Zhang, T.; Vail, S.; Pozniak, C. Classification of crop lodging with gray level co-occurrence matrix. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 251–258.
10. Yang, M.-D.; Huang, K.-S.; Kuo, Y.-H.; Tsai, H.P.; Lin, L.-M. Spatial and Spectral Hybrid Image Classification for Rice Lodging Assessment through UAV Imagery. *Remote Sens.* **2017**, *9*, 583. [[CrossRef](#)]
11. Cao, W.; Qiao, Z.; Gao, Z.; Lu, S.; Tian, F. Use of unmanned aerial vehicle imagery and a hybrid algorithm combining a watershed algorithm and adaptive threshold segmentation to extract wheat lodging. *Phys. Chem. Earth Parts A/B/C* **2021**, *123*, 103016. [[CrossRef](#)]
12. Su, Z.; Wang, Y.; Xu, Q.; Gao, R.; Kong, Q. LodgeNet: Improved rice lodging recognition using semantic segmentation of UAV high-resolution remote sensing images. *Comput. Electron. Agric.* **2022**, *196*, 106873. [[CrossRef](#)]
13. Yang, M.D.; Boubin, J.G.; Tsai, H.P.; Tseng, H.H.; Hsu, Y.C.; Stewart, C.C. Adaptive autonomous UAV scouting for rice lodging assessment using edge computing with deep learning EDANet. *Comput. Electron. Agric.* **2020**, *179*, 105817. [[CrossRef](#)]
14. Yang, M.D.; Tseng, H.-H.; Hsu, Y.-C.; Tsai, H.P. Semantic Segmentation Using Deep Learning with Vegetation Indices for Rice Lodging Identification in Multi-date UAV Visible Images. *Remote Sens.* **2020**, *12*, 633. [[CrossRef](#)]
15. Han, L.; Yang, G.; Yang, X.; Song, X.; Xu, B.; Li, Z.; Wu, J.; Yang, H.; Wu, J. An explainable XGBoost model improved by SMOTE-ENN technique for maize lodging detection based on multi-source unmanned aerial vehicle images. *Comput. Electron. Agric.* **2022**, *194*, 106804. [[CrossRef](#)]
16. Song, Z.H.; Zhang, Z.T.; Yang, S.Q.; Ding, D.Y.; Ning, J.F. Identifying sunflower lodging based on image fusion and deep semantic segmentation with UAV remote sensing imaging. *Comput. Electron. Agric.* **2020**, *179*, 105812. [[CrossRef](#)]
17. Zhang, D.; Ding, Y.; Chen, P.; Zhang, X.; Pan, Z.; Liang, D. Automatic extraction of wheat lodging area based on transfer learning method and deeplabv3+ network. *Comput. Electron. Agric.* **2020**, *179*, 105845. [[CrossRef](#)]
18. Yu, J.; Cheng, T.; Cai, N.; Zhou, X.-G.; Diao, Z.; Wang, T.; Du, S.; Liang, D.; Zhang, D. Wheat Lodging Segmentation Based on Lstm_PSPNet Deep Learning Network. *Drones* **2023**, *7*, 143. [[CrossRef](#)]
19. Zhang, Z.; Flores, P.; Igathinathane, C.; LNaik, D.; Kiran, R.; Ransom, J.K. Wheat lodging detection from UAS imagery using machine learning algorithms. *Remote Sens.* **2020**, *12*, 1838. [[CrossRef](#)]
20. Mardanisamani, S.; Maleki, F.; Hosseinzadeh Kassani, S.; Rajapaksa, S.; Duddu, H.; Wang, M.; Shirtliffe, S.; Ryu, S.; Josuttis, A.; Zhang, T.; et al. Crop lodging prediction from UAV-acquired images of wheat and canola using a DCNN augmented with handcrafted texture features. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019.

21. Yang, B.; Zhu, Y.; Zhou, S. Accurate wheat lodging extraction from multi-channel UAV images using a lightweight network model. *Sensors* **2021**, *21*, 6826. [[CrossRef](#)]
22. Bai, H.; Liu, L.; Han, Q.; Zhao, Y.; Zhao, Y. A novel UNet segmentation method based on deep learning for preferential flow in soil. *Soil Tillage Res.* **2023**, *233*, 105792. [[CrossRef](#)]
23. Baheti, B.; Innani, S.; Gajre, S.; Talbar, S. Semantic scene segmentation in unstructured environment with modified DeepLabV3+. *Pattern Recognit. Lett.* **2020**, *138*, 223–229. [[CrossRef](#)]
24. Yu, C.; Wang, J.; Peng, C.; Gao, C.; Yu, G.; Sang, N. Bisenet: Bilateral segmentation network for real-time semantic segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 325–341.
25. Zualkernan, I.; Abuhani, D.A.; Hussain, M.H.; Khan, J.; ElMohandes, M. Machine Learning for Precision Agriculture Using Imagery from Unmanned Aerial Vehicles (UAVs): A Survey. *Drones* **2023**, *7*, 382. [[CrossRef](#)]
26. Shorewala, S.; Ashfaq, A.; Sidharth, R.; Verma, U. Weed density and distribution estimation for precision agriculture using semi-supervised learning. *IEEE Access* **2021**, *9*, 27971–27986. [[CrossRef](#)]
27. Nong, C.; Fan, X.; Wang, J. Semi-supervised learning for weed and crop segmentation using uav imagery. *Front. Plant Sci.* **2022**, *13*, 927368. [[CrossRef](#)] [[PubMed](#)]
28. Pérez-Ortiz, M.; Peña, J.M.; Gutiérrez, P.A.; Torres-Sánchez, J.; Hervás-Martínez, C.; López-Granados, F. A semi-supervised system for weed map** in sunflower crops using unmanned aerial vehicles and a crop row detection method. *Appl. Soft Comput.* **2015**, *37*, 533–544. [[CrossRef](#)]
29. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
30. Reddy, Y.C.A.P.; Viswanath, P.; Reddy, B.E. Semi-supervised learning: A brief review. *Int. J. Eng. Technol.* **2018**, *7*, 81. [[CrossRef](#)]
31. Fourati, F.; Mseddi, W.S.; Attia, R. Wheat head detection using deep, semi-supervised and ensemble learning. *Can. J. Remote Sens.* **2021**, *47*, 198–208. [[CrossRef](#)]
32. Najafian, K.; Ghanbari, A.; Stavness, I.; Jin, L.; Shirdel, G.H.; Maleki, F. A Semi-Self-Supervised learning approach for wheat head detection using extremely small number of labeled samples. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 11–17 October 2021.
33. Schmarje, L.; Santarossa, M.; Schröder, S.-M.; Koch, R. A survey on semi-, self- and unsupervised learning for image classification. *IEEE Access* **2021**, *9*, 82146–82168. [[CrossRef](#)]
34. Pauletto, L.; Amini, M.-R.; Winckler, N. Self semi supervised neural architecture search for semantic segmentation. *arXiv* **2022**, arXiv:2201.12646.
35. Ubbens, J.R.; Ayalew, T.W.; Shirtliffe, S.; Josuttis, A.; Pozniak, C.; Stavness, I. Autocount: Unsupervised segmentation and counting of organs in field images. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 391–399.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.