Detection of wheat seedling lines in the complex environment via deep learning

Haibo Lin[1,2*], Yuandong Lu[1,2], Rongcheng Ding[1,2], Yufeng Xiu[1,2], Fazhan Yang[1,2]

(1. *Key Lab of Industrial Fluids Energy Conservation and Pollution Control (Ministry of Education), Qingdao 266520, Shandong, China*;
2. *School of Mechanical & Automotive Engineering, Qingdao University of Technology, Qingdao 266520, Shandong, China*)

Abstract: Wheat seedling line detection is critical for precision agriculture and automatic guidance in early wheat field operation. Aiming at the complex wheat field environment, a method of detecting wheat seedling lines based on deep learning was proposed in this study. Firstly, a rotated bounding box was created to improve the YOLOv3 model to predict the approximate position of the wheat seedling line; Then, according to the rotated bounding region obtained by the model, the wheat seedling line was detected by fitting the extracted center points. Finally, a comprehensive evaluation method combining angle error and distance error was proposed to evaluate the accuracy of the extracted crop line. By testing images of wheat seedlings in different environments, the results showed that the mean angle error and distance error respectively reached 0.75° and 10.84 pixels while the mean running time was 63.83 ms for a 1920×1080 pixels image. And compared to the original model the improved algorithm model improved the mAP value by 13.2%. The angle error and the distance error of the improved algorithm model were reduced by 51.4% and 39.7%, respectively. The method proposed in this study can accurately detect the wheat seedling lines at different stages and it is also suitable for the environments with weeds, shadow, bright light, and dark light. At the same time, it has a certain adaptability to wheat seedling images with a yaw angle in the shooting process. The research results could provide a reference for the automatic guidance of early wheat field machinery.
Keywords: wheat seedling lines, automatic guidance, deep learning, rotated bounding box, evaluation method
DOI: 10.25165/j.ijabe.20241705.7834

1 Introduction

Wheat is one of the main food crops in the world, and the stability and increase of its output are of great significance to world food security[1]. Researches have shown that measures such as root cutting[2], fertilization[3], and weeding[4] in early seedling rows can improve the yield of wheat[5]. However, field operations were mainly completed by manual labor or machines with low automation levels in the past, which suffered high labor intensity, low efficiency, and high cost[6]. As an important part of precision agriculture, the visual navigation technology of agricultural robots has attracted more and more attention[7,8]. It has been widely used in crop sowing[9], weeding[10], harvesting[11], and other links. The accurate and stable extraction of lines of wheat seedling rows can provide a reference for autonomous navigation and operation of wheat field agricultural machinery, which can effectively promote the intelligent development of agricultural machinery.

At present, many experts and scholars have carried out a lot of research on the extraction of lines of crops seedling rows, such as maize, wheat, rice, vegetables, etc. The traditional method was to separate crop rows from the background by image processing, and then fit the navigation line through features such as edge, parallax, brightness, and projection[12-17]. The main task was to identify the characteristics of crops and extract navigation lines using the Hough transform[18-20]. The Region of Interest (ROI) method was used to cluster the feature points[21-24], and the least square method was used to fit the navigation line for the clustering parameters. Our team[25] had done some research in this area. Through image segmentation, affine transformation, and vertical projection, the wheat seedling line was detected and extracted. Generally, the traditional image processing method has high requirements on the applicability of the environment, which is a challenge to improve the stability and real-time performance of the algorithm in complex farmland environments.

In recent years, with the increase in data volume, the innovation of algorithms, and the improvement of computing performance, deep learning has been applied to the agricultural machinery field. In the process of learning samples by combining neural networks, deep learning forms an abstract high-level to discover distributed features and attributes of data by combining low-level features[26]. Therefore, the deep learning method is characterized by high detection accuracy and strong adaptability[27,28]. Target detection algorithms in deep learning have been widely used in crop recognition. Liu et al.[29] used the Faster R-CNN model and SSD model to identify rice seedlings. Pang et al.[30] adopted MaxArea Mask scoring RCNN model to detect maize seedling rows, which could ignore the direction, shape, and intersection angle. Khan et al.[31] developed a deep learning system for weed and crop identification in farmland, the overall average identification accuracy of crop and weed was 94.73%. Liu et al.[32] adopted the YOLOv3 model to detect tomatoes, it was found that the model has

certain robustness to different light and shelter conditions in the environment.

However, there are few researches on target detection algorithms used to identify crop rows. The main reason can be attributed to that the horizontal bounding box was usually used by the target detection algorithm to annotate the target, while the crop rows show a large tilt angle during imaging. If crop rows are annotated as a whole column, it would result in too many background areas and other crop rows being included in the box, which is not conducive to training the model and extracting lines of crop rows. Aiming to solve the problem, Zhang et al.[33] adopted annotating method of crop rows segmentation. Ma et al.[34] proposed a rotatable text detection box, which could detect text in any direction. Neves et al.[35] detected underwater targets with a predicted rotated bounding box. Koirala et al.[36] evaluated the flowering level of trees by measuring panicle number with rotated bounding boxes. Yu et al.[37] annotated strawberries with a predicted rotated bounding box and determined the picking point of the strawberry based on rotated bounding boxes. Li et al.[38] positioned ships accurately with rotated bounding boxes. Compared with the horizontal bounding box, the rotated bounding box adds a rotation angle parameter, thus the position of the target can be more accurately reflected.

In order to improve the accuracy of wheat seedling line recognition, a deep learning method based on YOLO was proposed in this study. YOLOv3 has achieved a better balance in speed, accuracy, versatility, and complex scenarios than others. So in this study, firstly, the YOLOv3 model was improved to output a series of rotating recognition regions that vary with wheat seedling lines; secondly, in the rotating region, the feature points for wheat seedlings in the bounding box were extracted by the defined gray threshold value, and the center points for wheat seedling rows were extracted by moving the circular window; finally, in order to more accurately judge the accuracy of the navigation line, a comprehensive evaluation method combining angle error and distance error was proposed to evaluate the accuracy of the extracted crop line.

2   Materials

2.1   Image acquisition

The images of wheat seedlings in this research were acquired from the Base for Intelligent Wheat Root-cutting and Fertilization Technology Test in Qingdao, Shandong, China. The images were all taken from February to April 2021 when it is suitable for machine operations in the field. Figure 1 shows a schematic drawing of the image acquisition process. Images were captured by two mobile phones with different models (Play4TPro, HONOR, China; K30, MI, China). During shooting, the camera is 100-150 cm high from the ground, and the angle between the camera's optical axis and the horizontal line is 30°-60°. The field of view distance of the image would be too large and the wheat seedling rows in the image would be too dense when the angle is less than 30°, which is not conducive to the detection of the lines for the wheat seedling rows. The field of view distance of the image would be too small and the length of wheat seedling rows in the image would be too short when the angle is greater than 60°, which would affect the extraction accuracy. To fully capture the posture of wheat seedlings, the two cameras were taken along the different directions of wheat seedling rows to collect the images, as shown in Figure 1.

Wheat takes about 8 months from planting to harvest, which are divided into 12 growing cycles. After going through the overwintering period, wheat seedlings enter the turning-green stage and begin to grow. The field operations in this stage are helpful to the growth and yield of wheat. Therefore, the images of wheat used in this study were collected in the period. The images were saved in JPEG format with resolutions of 4624×3472, 4000×3000, and 1920×1080 pixels. To standardize the sample images, the resolutions of 4624×3472 and 4000×3000 pixels were converted to 1920×1080 pixels. The conversion schematic drawing is shown in Figure 2. In the conversion process, the width of the image was taken as the benchmark, the height was calculated at the ratio of 16:9 (1920:1080), and the redundant parts were removed. Finally, the image was sampled down to 1920×1080 pixels, and the bilinear interpolation method was used in this process owing to its simplicity and effectiveness.
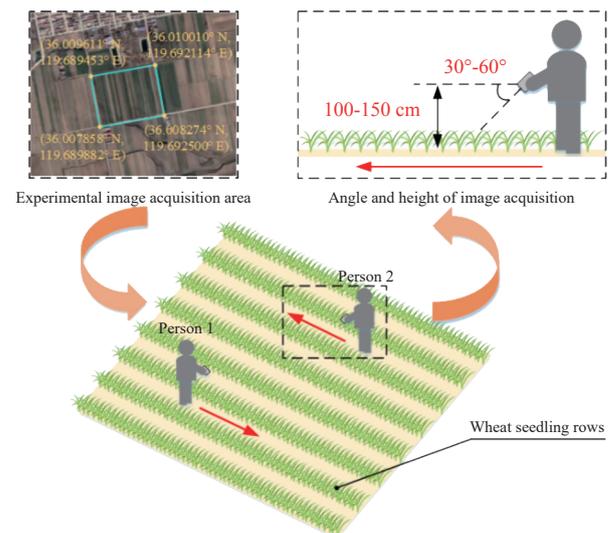

Experimental image acquisition area

Angle and height of image acquisition

Person 2

Person 1

Wheat seedling rows

Figure 1   Schematic drawing of the image acquisition process

2.2   Image set expansion

A total of 600 images were collected during the turning-green stage. To expand the image set, the images were expanded by changing brightness, rotating, mirroring, and adding noise. Some of the results are shown in Figure 3.

3   Methods

3.1   Overall approach

In order to achieve accurate extraction of lines for wheat seedling rows in complex environments, the improved YOLOv3 model was proposed to extract the feature points for wheat seedlings by changing the original lateral bounding box to a rotatable bounding box. Meanwhile, the center points for wheat seedling rows were extracted by a moving circular window. Then, the center points of each wheat seedling row were fitted based on the least square method.

3.2   Prediction model of wheat seedling rows based on improved YOLOv3

3.2.1   Overview of YOLOv3 model

YOLOv3 is one of the most popular target prediction models, which is a fully convolutional neural network based on regression. The location information of the target can be predicted accurately and efficiently by the YOLOv3 model[39].

The network structure of the YOLOv3 model is shown in Figure 4. Firstly, the input image size is converted to 416×416 pixels, which is then fed into the Darknet-53 basic network for feature extraction. In Darknet-53, the convolution with a step size of
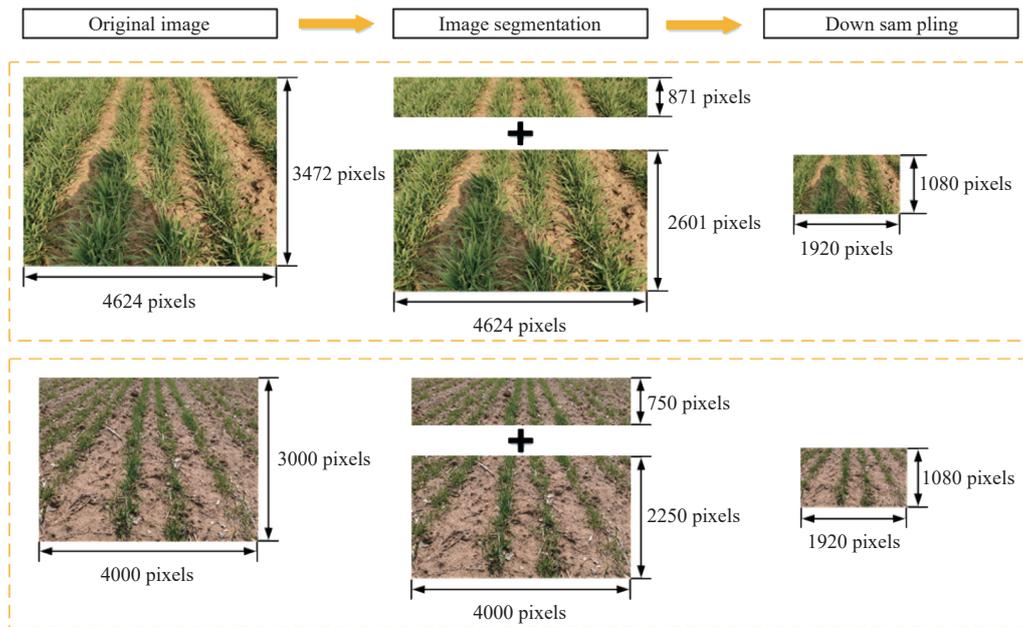
Figure 2   Schematic drawing of image conversion

a. Sample image

b. Increase brightness

c. Reduce brightness

d. 10° Clockwise rotation

e. 20° Clockwise rotation

f. 30° Clockwise rotation

g. −10° Clockwise rotation

h. −20° Clockwise rotation

i. −30° Clockwise rotation

j. Horizontal mirror

k. Add salt and pepper noise

l. Add gaussian noise

Figure 3   Examples of image set extension

Darknet 53

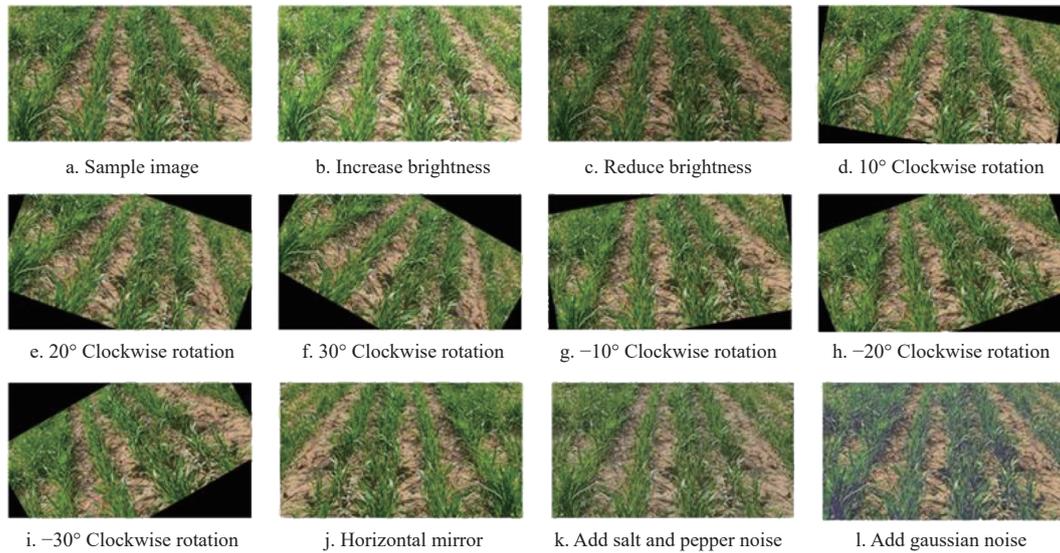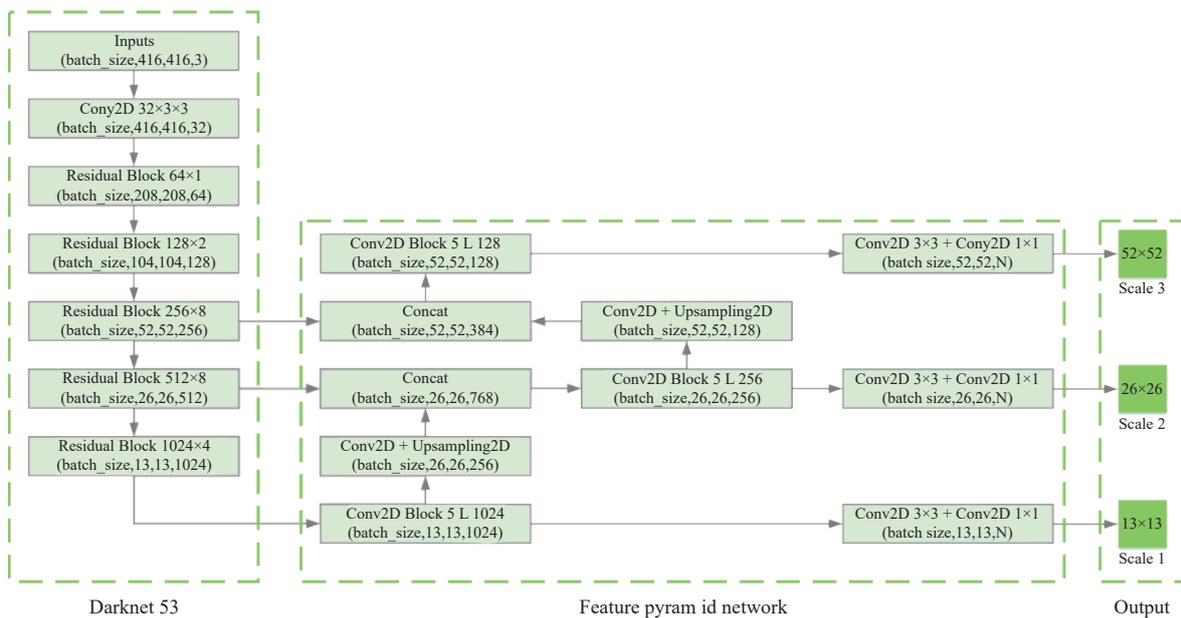Feature pyram id network

Output

Figure 4   YOLOv3 network structure used in this study

2 is used for down-sampling, which can save more information on the image while reducing the computation. To avoid the case of less location information in the middle and high levels of the convolutional neural network, a feature pyramid network is used to fuse feature maps of different levels in this YOLOv3 model. The output three-layer feature map is used to predict the targets at different scales, and the final prediction result is determined by the non-maximum suppression algorithm.

3.2.2 Annotation of the rotated bounding box

In this research, a rotated bounding box was used to replace the original horizontal bounding box in YOLOv3. Figure 5 shows the schematic drawing of annotating two different types of bounding boxes. Figure 5a is a horizontal bounding box, and Figure 5b is a rotated bounding box. Compared with the horizontal bounding box, the rotated bounding box contains a smaller background area and fewer other rows of wheat seedlings.

a. Horizontal bounding box annotation   b. Rotating bounding box annotation

Figure 5   Annotation of horizontal bounding box and rotated bounding box

The annotation parameters of the rotated bounding box are expressed as (x, y, w, h, α), as shown in Figure 6, where (x, y) represents the coordinates of the center point of the rotated bounding box, w, and h represent the width and height of the rotated bounding box, respectively, and α represents the clockwise rotation angle of the bounding box around the center point (x, y). The unit of α is radian, and the α value is in the range of $[0, \pi)$. In Figure 6a, $\alpha=0$. In Figure 6b, $\alpha \in \left(0, \dfrac{\pi}{2}\right)$. In Figure 6c, $\alpha \in \left(\dfrac{\pi}{2}, \pi\right)$. According to the distribution of wheat seedling rows in the captured images (Figure 6), it can be found that:

1) The α value of the bounding box annotated in the right half of the image is in the range of $\left(\dfrac{\pi}{2}, \pi\right)$. The closer the bounding box is to the right edge of the image, the smaller the α value is. The closer the bounding box is to the center of the image, the larger the α value is.

2) The α value of the bounding box annotated in the left half of the image is in the range of $\left(0, \dfrac{\pi}{2}\right)$. The closer the bounding box is to the left edge of the image, the larger the α value is. The closer the bounding box is to the center of the image, the smaller the α value is.

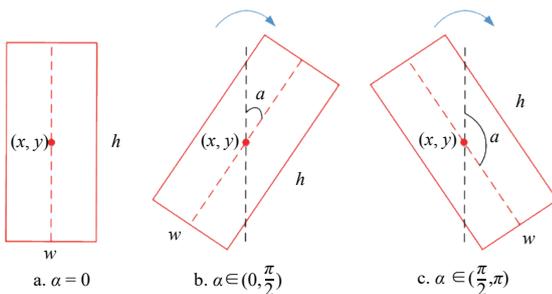a. $\alpha = 0$ b. $\alpha \in (0, \dfrac{\pi}{2})$ c. $\alpha \in (\dfrac{\pi}{2}, \pi)$

Figure 6   Three positions of the rotated bounding boxes

3.2.3 Generation of the rotated bounding box

Before the training, the *k*-means algorithm was used to cluster

the size and rotation angle of the bounding box respectively. Considering that there are three feature maps with different scales output by the YOLOv3 model and the regression of feature maps of each scale was equipped with three candidate bounding boxes, therefore nine bounding boxes with different sizes were required. In this study, 208×1117, 241×1049, and 350×987 were assigned to the 13×13 feature map to predict large wheat seedling rows. 154×479, 203×559, and 243×511 were assigned to the 26×26 feature map to predict medium-sized wheat seedling rows. 116×213, 127×299, and 243×242 were assigned to the 52×52 feature map to predict small wheat seedling rows. According to the statistics of the number of wheat seedling rows in the image, the number of candidate rotation angles was set to 6, which were 0.01 rad, 0.18 rad, 0.34 rad, 2.68 rad, 2.86 rad, and 3.05 rad, respectively.

3.2.4 Loss function

The loss function needed to be optimized to obtain the ideal model. The part related to multi-classification prediction in the loss function was removed as only wheat seedling was taken as targets in this research. At the same time, the smooth-L1 loss function was used to represent angle prediction losses[35]. The rest of the loss function was the same as that of the traditional YOLOv3 model. The loss function is shown as follows:

$$\text{Loss} = L_\alpha + L_{x,y} + L_{w,h} + L_{\text{conf}} \tag{1}$$

$$L_\alpha = \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{\text{obj}} \text{smooth}_{L1}(\alpha_i^j - \hat{\alpha}_i^j) \tag{2}$$

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \tag{3}$$

$$L_{x,y} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{\text{obj}} [(x_i^j - \hat{x}_i^j)^2 + (y_i^j - \hat{y}_i^j)^2] \tag{4}$$

$$L_{w,h} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{\text{obj}} \left[\left(\sqrt{w_i^j} - \sqrt{\hat{w}_i^j} \right)^2 + \left(\sqrt{h_i^j} - \sqrt{\hat{h}_i^j} \right)^2 \right] \tag{5}$$

$$L_{\text{conf}} = -\sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{\text{obj}} \left[\hat{c}_i^j \log\left(c_i^j\right) + \left(1 - \hat{c}_i^j\right) \log\left(1 - c_i^j\right) \right] - $$
$$\lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{noobj} \left[\hat{c}_i^j \log(c_i^j) + (1 - \hat{c}_i^j)\log(1 - c_i^j) \right] \tag{6}$$

where, S^2 represents the size of the output feature map (S=13, 26, 52). B represents the number of candidate bounding boxes for each grid in the feature map ($B = 3 \times 6$). $I_{i,j}^{\text{obj}}$ represents whether the *j*th candidate bounding box in the *i*th grid is responsible for target detection. If so, $I_{i,j}^{\text{obj}} = 1$; Otherwise, $I_{i,j}^{\text{obj}} = 0$. $I_{i,j}^{\text{noobj}}$ is the opposite to $I_{i,j}^{\text{obj}}$. α_i^j, and $\hat{\alpha}_i^j$ represent the predicted and actual rotation angle of the *j*th candidate bounding box in the *i*th grid, respectively. λ_{coord} and λ_{noobj} are used to balance the training loss of the bounding box and confidence, respectively. (x_i^j, y_i^j) and $(\hat{x}_i^j, \hat{y}_i^j)$ represent the predicted and actual coordinates of the center point of the *j*th candidate bounding box in the *i*th grid, respectively. w_i^j and h_i^j represent the predicted size of the *j*th candidate bounding box in the *i*th grid, respectively. $\hat{w}_i^j$ and $\hat{h}_i^j$ represent the actual size of the *j*th candidate bounding box in the *i*th grid, respectively. c_i^j and $\hat{c}_i^j$ represent the predicted and actual confidence of the target in the *j*th candidate bounding box in the *i*th grid, respectively.

3.3　Detection of the lines for wheat seedling rows

3.3.1　Extraction of the feature points

The line of wheat seedling row is extracted based on the predicted rotating bounding box. In this paper, the center lines of wheat seedling rows are extracted based on the predicted rotated bounding box, which can improve extraction efficiency by reducing background area. However, only a small part of the wheat seedling row is contained in some bounding box in the image, and the center lines extracted based on this region cannot accurately represent the wheat seedling row. This part of the bounding box is usually located at the edges of both sides of the image, and its height is not higher than half of the height of images. Therefore, bounding boxes are excluded when their height is less than half of the height of the images in this study.

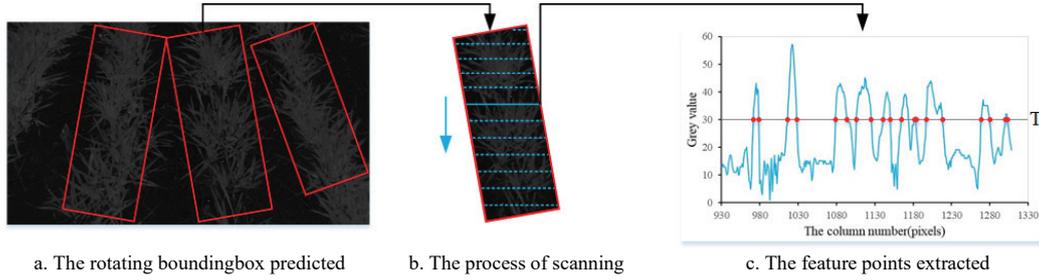The edge points of wheat leaves were taken as the feature points for wheat seedlings as the wheat leaves were an important feature of wheat seedlings. Then, the image was converted to a gray image through the H channel in HLS color space. The extraction method of the feature points is shown as follows:

$$P_{i,j} = \begin{cases} 1, & \left(\text{grey}_{i-1,j} - T\right)\left(\text{grey}_{i+1,j} - T\right) < 0 \\ 0, & \text{Otherwise} \end{cases} \quad (7)$$

where, $\text{grey}_{i,j}$ represent the gray value of the image at point (i, j). T represents the gray threshold. If $P_{i,j}=1$, it indicates that the point (i, j) is the feature point for wheat seedling.

The extraction process of feature points is shown in Figure 7. The red bounding box in Figure 7a is the predicted rotation bounding box. Figure 7b is the top-down scanning process. The red points in Figure 7c are the feature points extracted from the image.

a. The rotating boundingbox predicted　　　b. The process of scanning　　　c. The feature points extracted

Figure 7　Extraction process of feature points

3.3.2　Extraction of the center points

A large number of feature points are extracted and spread throughout the whole bounding box. In this study, the center points of wheat seedling rows were extracted to ensure the fitting accuracy of the lines. The extraction process is shown in Figure 8. A moving circular window (black circle) is defined, which moves from the middle point of the upper edge of the bounding box to the middle point of the lower edge of the bounding box along a straight line. During the moving process, the average value of the coordinates of the feature points inside the window is calculated and taken as the center point of the wheat seedling row (blue point). The specific process is as follows:
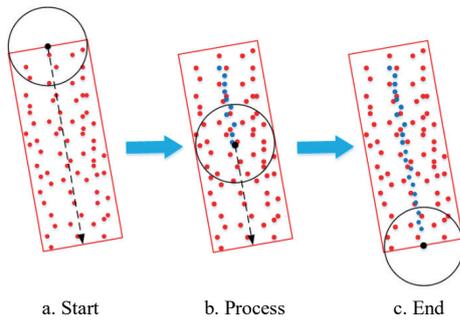

a. Start　　　b. Process　　　c. End

Figure 8　Extraction process of the feature points

Step 1: Calculate the coordinates of the four vertices of the rotated bounding box. The position coordinates of the four vertices of the bounding box before rotation were calculated based on the center point (x, y), the width w, and the height h of the rotated bounding box output by the model. The four vertices are the upper left vertex $P_1\left(x - \dfrac{w}{2}, y - \dfrac{h}{2}\right)$, the upper right vertex $P_2\left(x + \dfrac{w}{2}, y - \dfrac{h}{2}\right)$, lower right vertex $P_3\left(x + \dfrac{w}{2}, y + \dfrac{h}{2}\right)$, lower left vertex $P_4\left(x - \dfrac{w}{2}, y + \dfrac{h}{2}\right)$, respectively. Then, the coordinates of the four vertices P_1^r, P_2^r, P_3^r, P_4^r after rotation are calculated according to the rotation angle α. The calculation equations are shown as follows:

$$x_i^r = (x_i - x)\cos\alpha - (y_i - y)\sin\alpha + x_i \quad (8)$$

$$y_i^r = (x_i - x)\sin\alpha + (y_i - y)\cos\alpha + y_i \quad (9)$$

where, (x_i, y_i) is the vertex coordinate before rotation, and (x_i^r, y_i^r) is the vertex coordinates after rotation, i=1, 2, 3, 4.

Step 2: Determine the starting point, end point, and moving path of the circular window. The starting point and end point of the moving circular window, respectively, are set to the center points of the upper and lower edges of the rotated bounding box. The moving path of the circular window is set to a straight line from the start point to the endpoint. The start point and end point are dependent on the rotation angle α. When $0 \le \alpha \le \dfrac{\pi}{2}$, the start point is the midpoint of P_1^r and P_2^r, and the endpoint is the midpoint of P_3^r and P_4^r. When $\dfrac{\pi}{2} < \alpha < \pi$, the start point is the midpoint of P_3^r and P_4^r, and the endpoint is the midpoint of P_1^r and P_2^r. The equation of the moving path can be derived from the start point and end point. When α=0, the equation of the moving path is $X=x$. When $0<\alpha<\pi$, the equation of the moving path is $Y=k \cdot X+b$, $k = \tan\left(\dfrac{\pi}{2} - \alpha\right)$, $b = y - x \cdot \tan\left(\dfrac{\pi}{2} - \alpha\right)$.

Step 3: Calculate the coordinates of the center points of wheat seedling rows. The circular window was defined with a radius of $\dfrac{w}{2}$ moves from the start point to the end point along the chosen path. When the circular window moves to a new position, the average value of abscissa and ordinate of all points in the circular window are calculated respectively, which is taken as the center point for the wheat seedling row.

The feature points and center points extracted from the image are shown in Figure 9. It can be seen from Figure 9 that the distributions of center points and center points are consistent with the actual photos.

Figure 9   Feature points and center points extracted

3.3.3   Fitting of the lines of wheat seedling rows

Through optimization, the number of wheat seedling row center points is greatly reduced compared with the number of wheat seedling feature points. Here, the least square method is used to fit the center points due to its advantage of fast fitting and good effect. The best function for the data can be found using the least square method by minimizing the sum of the squares of error. The advantage of the least square method is more significant when the number of points is less.

3.3.4   Evaluation method of extraction effect

The evaluation method combining angle and distance is adopted to effectively evaluate the error between the extracted lines and the ideal lines. As shown in Figure 10, Line 1 (red dotted line) represents the extracted lines. Line 2 (red solid line) represents ideal lines. θ represents the angle between the two lines. A and B represent the intersection of two lines and the upper edge of the image, respectively. C and D represent the intersection of two lines and the lower edge of the image, respectively. In addition, d_1 represents the vertical distance from point A to Line 2, and d_2 represents the vertical distance from point C to Line 1.
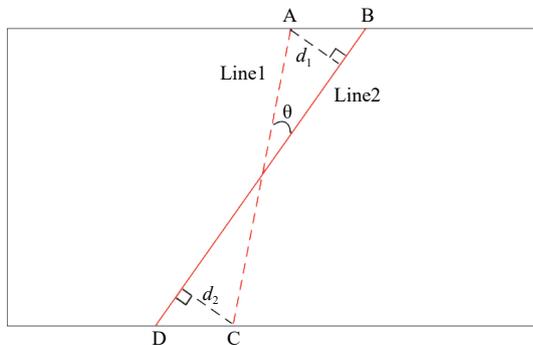

Figure 10   Schematic drawing of evaluation method

The angle θ and distance $\bar{d}$ represent the errors in position and distance between the extracted lines of the wheat seedling row and the ideal lines, respectively. The calculation equations are shown as follows:

$$\theta = \tan^{-1}\left(\frac{|k_1 - k_2|}{1 + k_1 k_2}\right) \quad (10)$$

$$d_1 = \left|\frac{k_2 x_A - y_A + b_2}{\sqrt{k_2^2 + 1}}\right| \quad (11)$$

$$d_2 = \left|\frac{k_2 x_C - y_C + b_2}{\sqrt{k_2^2 + 1}}\right| \quad (12)$$

$$\bar{d} = \frac{d_1 + d_2}{2} \quad (13)$$

where, k_1 represents the slope of Line 1. k_2 and b_2, respectively, represent the slope and intercept of Line 2. As can be seen from the formula, the smaller the angle and distance, the more accurate it is.

4   Results and analysis

4.1   Test platform

All processes in this research were run on the same computer, which was configured with Intel(R) Core(TM) I7-10700F CPU, 2.90 GHz, 16 GB operating memory, and NVIDIA GeForce RTX 2060 GPU. The algorithm was written with Python 3.7 in Windows 10.

4.2   Analysis of target prediction performance

There are 7200 wheat seedling images contained in the image set in this study. The training set accounts for 80% (5760 images) while the test set accounts for 20% (1440 images). In addition, 80% of images (4608 images) in the training set are used for the selection of model parameters, and the other 20% of images (1152 images) are used for the verification of the model. The number of iterations of training is set to 5000 to obtain a stable and accurate prediction model.

Overlap coefficient (OC) was adopted to improve the prediction accuracy. OC refers to the ratio of intersection and union between predicted and real bounding boxes, and its calculation method is shown as follows:

$$OC = \frac{A_P \cap A_T}{A_P \cup A_T} \quad (14)$$

where, A_P and A_T represent the predicted and real bounding box regions, respectively. The OC value ranges from 0 to 1. The larger the value is, the closer the predicted bounding box is to the real bounding box. Through continuous debugging and comparison results, the prediction precision of wheat seedling rows is better when the value of OC is higher than 0.85.

The evaluation indexes of the model include precision (P), recall rate (R_0), mean average precision (mAP), the harmonic average of P and R_0 (F1-score), etc. The calculation equations are shown as follows:

$$P = \frac{T_P}{T_P + F_P} \times 100\% \quad (15)$$

$$R_0 = \frac{T_P}{T_P + F_N} \times 100\% \quad (16)$$

$$mAP = \frac{1}{C}\sum_{i=1}^{N} P(i)\Delta R_0(i) \quad (17)$$

$$F1 - score = \frac{2PR_0}{P + R_0} \times 100\% \quad (18)$$

where, T_P represents the number of positive samples correctly classified; F_P represents the number of samples incorrectly classified as positive; F_N represents the number of samples incorrectly classified as negative.

The mAP is used to evaluate the performance of the model. The higher the value, the better the prediction effect of the model.

In this research, 1440 images from the test set were adopted to evaluate the performance of the model, and the results showed that the mAP was 81.2%. By analyzing the prediction results of the images in the test set, it can be found that a better prediction effect can be obtained when the edge contour of both sides of the wheat seedling row in the image is more significant. Figure 11 shows the prediction results for the wheat seedling row. It can be seen from the figure that the predicted positions of wheat seedling rows are

identical. However, the prediction results become worse when the wheat seedlings are so dense that the leaves of two adjacent rows cross. It is difficult for even people to identify the exact position of its contours. In addition, the shooting angle and height would also affect the prediction results. The smaller the angle between the optical axis of the shooting head and the horizontal plane or the lower the shooting height, the more likely the wheat contour far away from the shooting equipment is to be sheltered, which would result in lower prediction accuracy.

Figure 11　Prediction results of wheat seedling rows

The horizontal bounding box in the original YOLOv3 model was used as the control group to verify the effectiveness of the rotated bounding box in predicting wheat seedling rows. There are two annotation methods for horizontal bounding boxes: 1) annotate a row of wheat seedlings with one box; 2) annotate a row of wheat seedlings with multiple boxes. In order to ensure the comparability of the final prediction results, the training set and test set are the same as this research except for the difference in annotation methods. The results show that the mAP of the two methods was 59.2% and 76.8%, respectively. The precision of annotating a row of wheat seedlings by the horizontal bounding box with one box is the lowest. This can be explained by that most of the wheat seedling rows in the image are tilted. Annotating with a horizontal bounding box would cause the OC value of the bounding box of two adjacent wheat seedling rows to be too large. The number of prediction boxes would be smaller than the actual number of wheat seedling rows if the OC value is reduced. Conversely, the number of prediction boxes would be larger than the actual number of wheat seedling rows if the OC value is increased. Therefore, it is difficult to determine the OC values to obtain good model performance. Since the number and position of the bounding boxes in a wheat seedling row are uncertain, the effect of annotating a wheat seedling row with multiple boxes is not ideal, which is also the main reason for the high loss value of this method. In contrast, the method proposed in this study is more suitable for predicting wheat seedling rows.

4.3　Determination of extraction step sizes of feature points and center points

The extraction of feature points and center points is very important for the lines of wheat seedling rows. The detection precision of the lines of wheat seedling rows is higher and more stable in the case of all feature points and center points are extracted. However, this is unnecessary and time-consuming. Therefore, an appropriate extraction step size was selected to ensure the detection effect of lines of wheat seedling rows while reducing the amount of calculation in this study.

There were 50 images randomly selected from the test set to determine the extraction step sizes of the feature point and center point, a total of 235 wheat seedling rows were expected to be extracted. Figures 12-14, respectively, show the mean angle error, mean distance error, and mean running time of the extracted center

lines of wheat seedling rows with different step sizes during feature point extraction and center point extraction. The abscissa in Figures 12-14 is the step size in the vertical direction during the process of feature point extraction. The step length traversed in the horizontal direction was set to half of the step length traversed in the vertical direction owing to the length of most bounding boxes in the horizontal direction being lower than that in the vertical direction. In the process of center point extraction, only the step size in the vertical direction was considered. The curves of different colors in Figures 12-14 correspond to different step sizes selected during center point extraction.
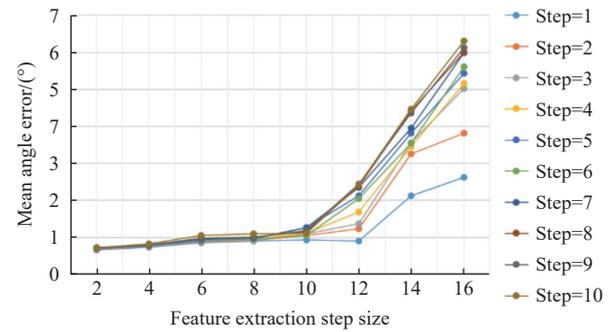

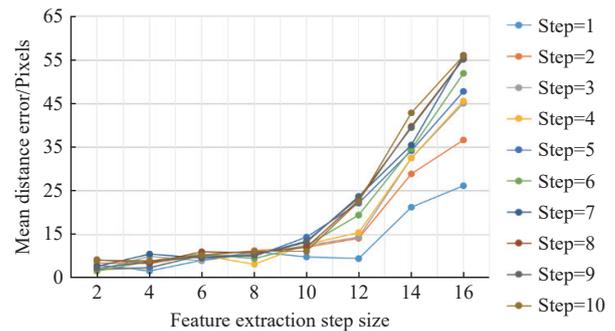Figure 12　Mean angle error corresponding to different step sizes

Figure 13　Mean distance error corresponding to different step sizes
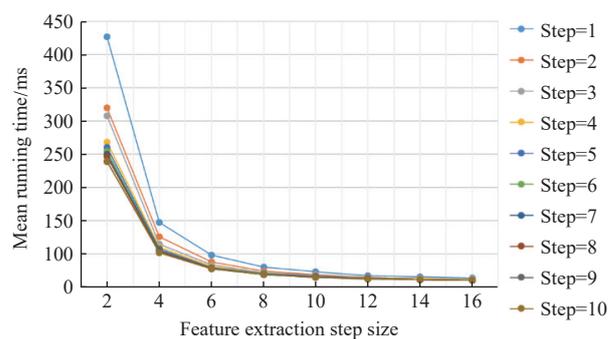

Figure 14　Mean running time corresponding to different step sizes

It can be found from Figure 12 that the corresponding mean angle error is low when the feature extraction step size is 2 to 10. The mean angle error increases greatly with the increase of feature extraction step size. The main reason is that the number of extracted feature points is too small to fit accurate lines when the step size is large. The same is true of the relationship between the step size of center point extraction process and the mean angle error.

In Figure 13, the relationship between step size and mean distance error is similar to that in Figure 12. However, when two lines are at the same angle, the distance between them may vary,

therefore, the fluctuation of the mean distance error in Figure 13 is more drastic. This also reflects that it is one-sided to evaluate the detection precision of lines only using the angle error. As a result, the detection precision of lines can be more accurately evaluated by combining angle error and distance error.

Figure 14 shows the relationship between step size and mean running time. It can be found that the mean running time decreases with the increase in step size, regardless of feature point extraction and center point extraction. The tendency is more significant when the step size is not larger than 6. Then, the mean running time gradually flattens out with the further increase in step size. The whole running time contains two parts: the running time of wheat seedling row prediction and the running time of line extraction. The increase in step size can only reduce the running time of line extraction. Therefore, the running time cannot be infinitely reduced by changing the step size.

The linear weighting method combining angle error, distance error, and running time is used to select step size. The expression can be shown as follows:

$$\min_{x \in X, y \in Y} \sum_{i=1}^{m} w_i f_i(x, y) \qquad (19)$$

where, x and y represent the traversal step in the vertical direction during the extraction of feature points and center points, respectively; $X=\{2, 4, 6, 8, 10, 12, 14, 16\}$; $Y=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$; $m=3$; w_i represents the weight of the ith variable; $f_i(x, y)$ represents the normalized value of the ith variable when the traversal step size in the vertical direction during extracting the feature point is x, the traversal step size in the horizontal direction during extracting the feature point is $\frac{x}{2}$ and the traversal step size during extracting the center point is y.

Each value of the variable is normalized to eliminate the influence of different units and value ranges on the results. Finally, in the process of feature point extraction, the step size in the vertical direction is set to 8, the step size in the horizontal direction is set to 4, and the step size in the process of central point extraction is set to 10.

4.4　Analysis of the detection effect of lines

The wheat seedling, wheat field environment, and shooting angle were discussed to verify the detection effect of center lines in this paper. As for the wheat seedling itself, the influence of different morphology of wheat seedlings caused by different growth stages on the extraction method of the center line was analyzed. As for the wheat field environment, the influence of weeds, shadows, and different brightness of light in the process of shooting on the extraction method of the center line was analyzed. As for the shooting angle, the influence of the yaw angle during shooting on the extraction method of the center line was analyzed.

Figure 15 shows the detection effect of center lines of representative images in four time periods. The four images were taken on February 25, 2021, March 11, 2021, March 25, 2021, and April 7, 2021, respectively. It can be found from Figure 15 that the
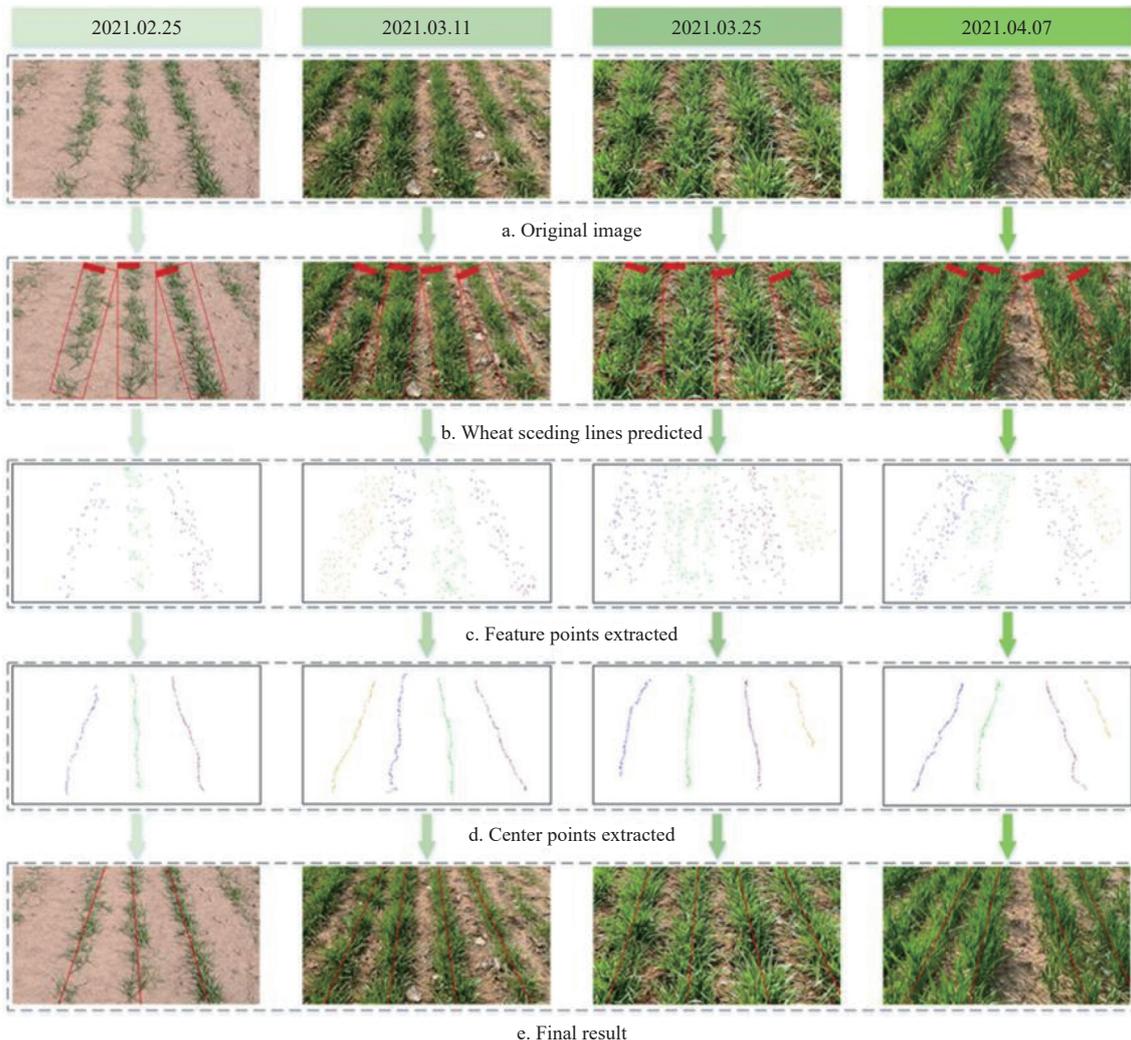

Figure 15   Detection effect of lines in different periods

contour of the wheat seedlings is clear in the first period when wheat seedlings begin to grow after winter. The position of wheat seedling rows can be accurately predicted by the improved YOLOv3 model. The seedlings' rows become wider as the wheat grows, resulting in more leaves of two adjacent wheat seedlings' rows being crossed. Although the bounding box inevitably contains other wheat seedling rows, the approximate position of wheat seedling rows can still be predicted.

It can be found from Figure 15 that the color of wheat seedlings in different periods is slightly different. Therefore, gray threshold T plays an important role in feature point extraction. The number of feature points would decrease if the value is too large or too small, and then the precision of the fitted line would decrease. As a result, T values would vary with different periods. After a large number of tests, it is found that the T value is appropriate to be set as 30 to 35 in the early turning-green stage, 35 to 50 in the middle turning-green stage, and 50 to 55 in the late turning-green stage. In this study, the T values of the four time periods in Figure 15 are set to 30, 40, 40, and 50, respectively.

Figure 15d is the extracted center point, which reflects the central position of the wheat seedling row. Figure 15e shows the center lines of wheat seedling row fitting based on the least square method. It can be seen from Figure 15 that the precision of the lines of the wheat seedling row is high, while the precision of the lines on both sides is low. This is because the scale of the predicted bounding boxes on both sides of the image is smaller, and the number of extracted center points is less, leading to a larger error in the predicted lines of wheat seedling rows.

During the image acquisition, the shooting environment changes with the change of shooting location and time. The extraction of lines of wheat seedling rows is also affected by the complex environment. Figure 16 shows the detection effect of the center lines in case of the situations of weeds, shadow, strong light, and dark light. As can be seen from Figure 16, wheat seedling row prediction and center point extraction are not affected by weeds, shadow, strong light, and dark light. Therefore, the method proposed in this research has a certain anti-interference to weeds, shadow, and light.

a. Weeds
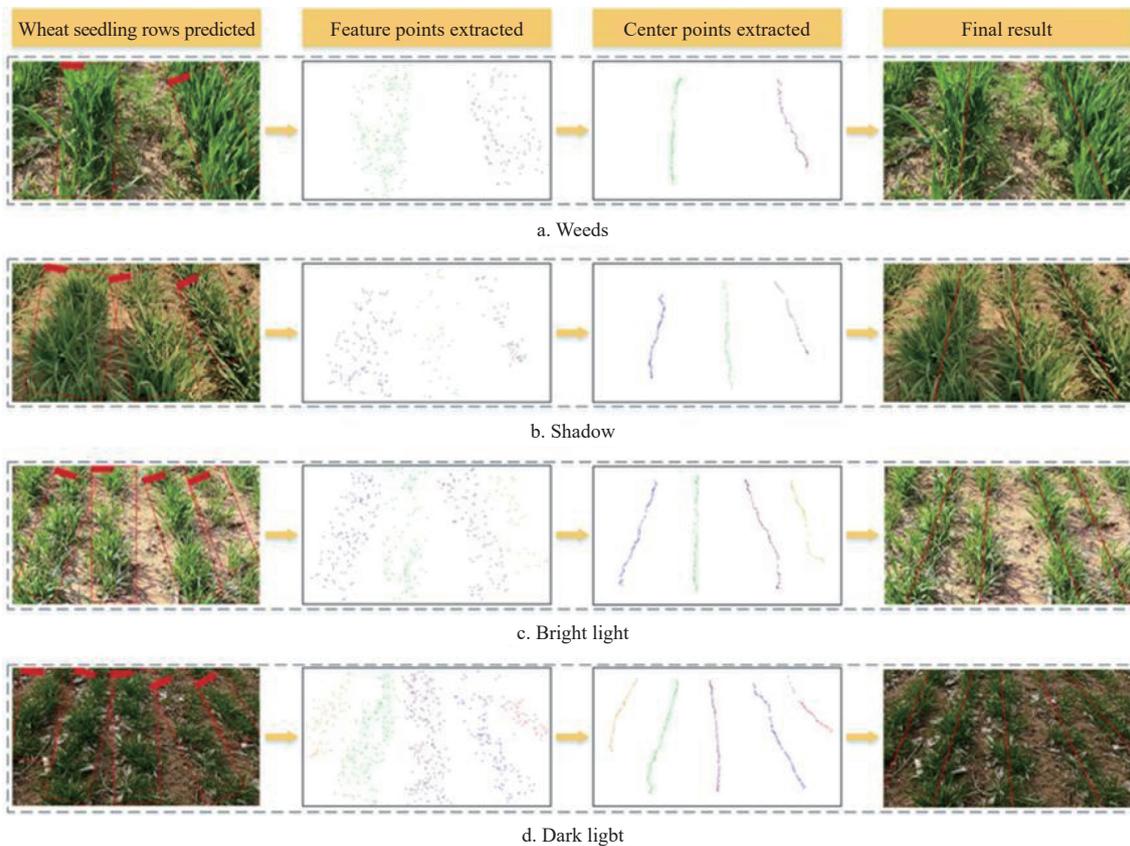
b. Shadow

c. Bright light

d. Dark ligbt

Figure 16   Extraction effect of lines in different wheat field environments

In reality, there is inevitably a certain yaw angle in the process of image acquisition due to the uneven ground, mechanical vibration, and linear error of the crop line. Therefore, a stable extraction method of center lines should also be suitable for images collected when the yaw angle of the camera is not 0. The images with certain yaw angles are simulated by images rotated at different angles to verify the influence of the yaw angle on the extraction accuracy of the method proposed in this research. Figure 17 shows the detection effect of center lines when the image is rotated at 10°, 20°, and −30°. As can be seen from Figure 17, image rotation has no significant influence on the extraction of center lines. Therefore, the extraction method of center lines based on the improved YOLOv3 model is effective and has certain adaptability to wheat seedling images with yaw angle.

Table 1 lists the detection results of the center lines of wheat seedling rows according to the specific classification mentioned above. The images for each specific classification contained 50 images from the test set. It can be seen from Table 1 that the extraction effect is significantly affected by the shooting time. The extraction accuracy of the center lines decreases as well as the running time increases slightly with the growth of wheat seedlings. However, the proposed method shows certain stability for images with weeds, shadows, bright light, dark light, and yaw angles. Finally, the mean angle error of the extraction results is 0.75°, the mean distance error is 10.84 pixels, and the mean running time is 63.83 ms.
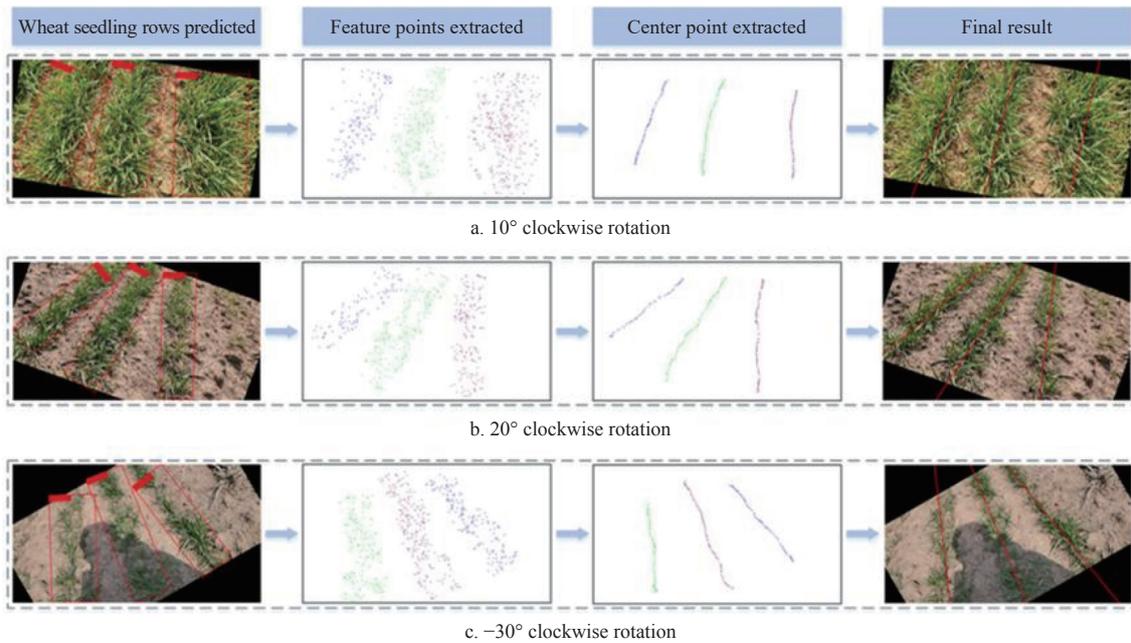
a. 10° clockwise rotation

b. 20° clockwise rotation

c. −30° clockwise rotation

Figure 17   Extraction effect of lines at different yaw angles

Table 1   Extraction results of lines for wheat seedling rows

Classification	The specific classification (date style: Year.Month.Day)	Mean angle error/(°)	Mean distance error/pixels	Mean running time/ms
Shooting time	2021.02.25-2021.02.27	0.50	6.34	62.42
	2021.03.11-2021.03.12	0.58	6.19	63.16
	2021.03.25-2021.03.26	0.92	12.39	67.59
	2021.04.09-2021.04.12	1.32	21.48	73.82
Shooting environment	Weeds	0.81	13.63	64.73
	Shadow	0.65	10.26	65.91
	Bright light	0.57	6.81	63.29
	Dark light	0.77	9.46	64.51
The yaw angle simulated	Rotated 10°	0.61	8.24	58.19
	Rotated 20°	0.69	9.69	57.94
	Rotated 30°	0.85	14.75	60.67

Table 2 lists the evaluation statistics of 50 wheat seedling images in different periods of the turning-green stage. From Table 2, the extraction accuracy of the proposed algorithm is better than that of the previous algorithm, especially in the late greening period, where the distance error and angle error of the previous algorithm are twice as much as that of the proposed algorithm.

Table 2   Evaluation of wheat seedling lines at different times under two models

Model	Category	Distance error/pixel	Angle error/(°)
Improved model	Early greening	10.52	1.06
	Mid-greening	9.36	0.89
	Late greening	13.61	1.15
	Mean value	11.16	1.03
Original model	Early greening	22.67	1.79
	Mid-greening	16.34	1.53
	Late greening	31.64	3.06
	Mean value	18.50	2.12

5   Conclusions

Aiming at the complex wheat field environment, a stable and accurate detection method for the center lines extraction of wheat seedling rows was proposed. The following conclusions were obtained by testing the wheat seedling images:

1) A rotated bounding box was created to improve the YOLOv3 model to predict the position of the wheat seedling line. The test results showed that the method proposed in this research could accurately detect the center lines of wheat seedling rows at the turning-green stage. The mAP of the improved model was improved by 13.2%. The method was also suitable for wheat seedling rows detection in the presence of weeds, shadow, bright light, and dark light. The method had certain adaptability to wheat seedling images with yaw angle in the shooting process.

2) An evaluation approach of row lines based on angle and distance was proposed. The results showed that the mean angle error was 0.75°, the mean distance error was 10.84 pixels, and the mean running time was 63.83 ms for a 1920×1080 pixels image. The angle error and distance error of the improved model were reduced by 51.4% and 39.7%, respectively.

At present, there are still some problems that need to be solved, such as whether the accuracy of wheat seedling row center line extraction detection method can be improved for crop rows with crop defects. In the future, we should further optimize the model and conduct in-depth research on crop rows with crop defects to improve the accuracy of detection.

[References]

[1]   Mueller N D, Gerber J S, Johnston M, Ray D K, Ramankutty N, Foley J A. Correction: Corrigendum: Closing yield gaps through nutrient and water management. Nature, 2013; 494: 390.

[2]   Yu S L, Qi X H, Liu X Y, Dong Q Y, Xu Y M. Researches on the effect on yield increase by deep cultivation-root cutting in winter wheat. Scientia Agricultura Sinica, 1985; 4: 30–35. (in Chinese)

[3]   Li X H, Yang S X, Yu Z W, Yu S L. Effects of organic manure application on growth and senescence of root in winter wheat. Journal of Plant

Nutrition and Fertilizers, 2005; 4: 467–472. (in Chinese)

[4] Khaliq A, Shakeel M, Matloob A, Hussain S, Tanveer A, Murtaza G. Influence of tillage and weed control practices on growth and yield of wheat. Philippine Journal of Crop Science, 2013; 38(3): 54–62.

[5] Lv Z Q, Li R S, Dong Q Y, Yin K R, Yu S L. Mechanical root-cutting models for increasing the yield of winter wheat. Transactions of the CSAE, 2006; 4: 103–106. (in Chinese)

[6] Xie B, Wu Z B, Mao E R. Development and prospect of key technologies on agricultural tractor. Transactions of the CSAM, 2018; 49: 1–17. (in Chinese)

[7] Zhang M, Ji Y H, Li S C, Cao R Y, Xu H Z, Zhang Z Q. Research progress of agricultural machinery navigation technology. Transactions of the CSAM, 2020; 51(4): 1–18. (in Chinese)

[8] Aravind K R, Raja P, Pérez-Ruiz M. Task-based agricultural mobile robots in arable farming: A review. Spanish Journal of Agricultural Research, 2017; 15: 1–16.

[9] Li J B, Chen B Q, Liu Y. Image detection method of navigation route of cotton plastic film mulch planter. Transactions of the CSAM, 2014; 45(1): 40–45. (in Chinese)

[10] Meng Q K, Qiu R C, He J, Zhang M, Ma X D, Liu G. Development of agricultural implement system based on machine vision and fuzzy control. Computers and Electronics in Agriculture, 2015; 112: 128–138.

[11] Zhao T, Noboru N, Yang L L, Ishii K, Chen J. Fast edge detection method for wheat field based on visual recognition. Transactions of the CSAM, 2016; 47(11): 32–37. (in Chinese)

[12] Zhang X Y, Li X N, Zhang B H, Zhou J, Tian G Z, Xiong Y J, et al. Automated robust crop-row detection in maize fields based on position clustering algorithm and shortest path method. Computers and Electronics in Agriculture, 2018; 154: 165–175.

[13] Gong J L, Wang X X, Zhang Y F, Lan Y B. Extraction method of corn rhizome navigation lines based on edge detection and area localization. Transactions of the CSAM, 2020; 51(10): 26–33. (in Chinese)

[14] Zhai Z Q, Zhu Z X, Du Y F, Song Z H, Mao E R. Multi-crop-row detection algorithm based on binocular vision. Biosystems Engineering, 2016; 150: 89–103.

[15] Basso M, Pignaton de Freitas E. A UAV guidance system using crop row detection and line follower algorithms. Journal of Intelligent and Robotic Systems: Theory & Application, 2020; 97(3-4): 605–621. doi: 10.1007/s10846-019-01006-0.

[16] Guan Z H, Chen K Y, Ding Y C, Wu C Y, Liao Q X. Visual navigation path extraction method in rice harvesting. Transactions of the CSAM, 2020; 51(1): 19–28. (in Chinese)

[17] Miguel Guerrero J, Jaime Ruz J, Pajares G. Crop rows and weeds detection in maize fields applying a computer vision system based on geometry. Computers and Electronics in Agriculture, 2017; 142: 461–472.

[18] Jiang G Q, Wang X J, Wang Z H, Liu H M. Wheat rows detection at the early growth stage based on Hough transform and vanishing point. Computers and Electronics in Agriculture, 2016; 123: 211–223.

[19] Zhang Q, Chen S J M E, Li B. A visual navigation algorithm for paddy field weeding robot based on image understanding. Computers and Electronics in Agriculture, 2017; 143: 66–78.

[20] Gong J L, Wang X X, Zhang Y F, Lan Y B, Mostafa K. Navigation line extraction based on root and stalk composite locating points. Computers and Electrical Engineering, 2021; 92: 107115.

[21] Jiang G Q, Wang Z H, Liu H M. Automatic detection of crop rows based on multi-ROIs. Expert Systems with Applications, 2015; 42(5): 2429–2441.

[22] Yang Y, Zhang B L, Zha J Y, Wen X, Chen L Q, Zhang T, et al. Real-time extraction of navigation lines between corn rows. Transactions of the CSAE, 2020; 36(12): 162–171. (in Chinese)

[23] He J, Zang Y, Luo X W, Zhao R M, He J, Jiao J K. Visual detection of rice rows based on Bayesian decision theory and robust regression least squares method. Int J Agric & Biol Eng, 2021; 14(1): 199–206.

[24] Zhou Y, Yang Y, Zhang B L, Wen X, Yue X, Chen L Q. Autonomous detection of crop rows based on adaptive multi-ROI in maize fields. Int J Agric & Biol Eng, 2021; 14(4): 217–225.

[25] Lu Y D, Lin H B, Ding R C, Xiu Y F. A method for extracting identification lines of early wheat seedling based on machine vision. Journal of Agricultural Mechanization Research, 2023; 45(8): 1–9. (in Chinese)

[26] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015; 521: 436–444.

[27] Kamilaris A, Prenafeta-Boldú F X. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 2018; 147: 70–90.

[28] Koirala A, Walsha K B, Wang Z L, McCarthy C. Deep learning - Method overview and review of use for fruit detection and yield estimation. Computers and Electronics in Agriculture, 2019; 162: 219–234.

[29] Liu F C, Yang Y, Zeng Y M, Liu Z Y. Bending diagnosis of rice seedling lines and guidance line extraction of automatic weeding equipment in paddy field. Mechanical Systems and Signal Processing, 2020; 142: 106791.

[30] Pang Y, Shi Y Y, Gao S C, Jiang F, Veeranampalayam-Sivakumar A-N, Thompson L, et al. Improved crop row detection with deep neural network for early-season maize stand count in UAV imagery. Computers and Electronics in Agriculture, 2020; 178: 105766.

[31] Khan S, Tufail M, Khan M T, Khan Z A, Anwar S. Deep learning-based identification system of weeds and crops in strawberry and pea fields for a precision agriculture sprayer. Precision Agriculture, 2021; 22: 1711–1727.

[32] Liu G X, Nouaze J C, Mbouembe P L T, Kim J H. YOLO-tomato: A robust algorithm for tomato detection based on YOLOv3. Sensors, 2020; 20(7): 2145.

[33] Zhang Q, Wang J H, Li B. Extraction method for centerlines of rice seedings based on YOLOv3 target detection. Transactions of the CSAM, 2020; 51(8): 34–43. (in Chinese)

[34] Ma J Q, Shao W Y, Ye H, Wang L, Wang H, Zheng Y B, et al. Arbitrary-oriented scene text detection via rotation proposals. IEEE Transactions on Multimedia, 2018; 20(11): 3111–3122.

[35] Neves G, Ruiz M, Fontinele J, Oliveira L. Rotated object detection with forward-looking sonar in underwater applications. Expert Systems with Applications, 2020; 140: 112870.

[36] Koirala A, Walsh K B, Wang Z L, Anderson N. Deep learning for mango (*Mangifera indica*) panicle stage classification. Agronomy, 2020; 10(1): 143.

[37] Yu Y, Zhang K L, Liu H, Yang L, Zhang D X. Real-time visual localization of the picking points for a ridge-planting strawberry harvesting robot. IEEE Access, 2020; 8: 116556–116568.

[38] Li L H, Zhou Z Q, Wang B, Miao L J, Zong H. A novel CNN-based method for accurate ship detection in HR optical remote sensing images via rotated bounding box. IEEE Transactions on Geoscience and Remote Sensing, 2021; 59(1): 686–699.

[39] Huang J H, Zhang H Y, Wang L, Zhang Z L, Zhao C M. Improved YOLOv3 model for miniature camera detection. Optics & Laser Technology, 2021; 142: 107133.