Soybean Yield Preharvest Prediction Based on Bean Pods and Leaves Image Recognition Using Deep Learning Neural Network Combined With GRNN

Wei Lu[1], Rongting Du[1], Pengshuai Niu[1], Guangnan Xing[2], Hui Luo[1], Yiming Deng[3] and Lei Shu[1]*

[1] College of Artificial Intelligence, Nanjing Agricultural University, Nanjing, China, [2] College of Agriculture, Nanjing Agricultural University, Nanjing, China, [3] College of Engineering, Michigan State University, East Lansing, MI, United States

Soybean yield is a highly complex trait determined by multiple factors such as genotype, environment, and their interactions. The earlier the prediction during the growing season the better. Accurate soybean yield prediction is important for germplasm innovation and planting environment factor improvement. But until now, soybean yield has been determined by weight measurement manually after soybean plant harvest which is time-consuming, has high cost and low precision. This paper proposed a soybean yield in-field prediction method based on bean pods and leaves image recognition using a deep learning algorithm combined with a generalized regression neural network (GRNN). A faster region-convolutional neural network (Faster R-CNN), feature pyramid network (FPN), single shot multibox detector (SSD), and You Only Look Once (YOLOv3) were employed for bean pods recognition in which recognition precision and speed were 86.2, 89.8, 80.1, 87.4%, and 13 frames per second (FPS), 7 FPS, 24 FPS, and 39 FPS, respectively. Therefore, YOLOv3 was selected considering both recognition precision and speed. For enhancing detection performance, YOLOv3 was improved by changing IoU loss function, using the anchor frame clustering algorithm, and utilizing the partial neural network structure with which recognition precision increased to 90.3%. In order to improve soybean yield prediction precision, leaves were identified and counted, moreover, pods were further classified as single, double, treble, four, and five seeds types by improved YOLOv3 because each type seed weight varies. In addition, soybean seed number prediction models of each soybean planter were built using PLSR, BP, and GRNN with the input of different type pod numbers and leaf numbers with which prediction results were 96.24, 96.97, and 97.5%, respectively. Finally, the soybean yield of each planter was obtained by accumulating the weight of all soybean pod types and the average accuracy was up to 97.43%. The results show that it is feasible to

predict the soybean yield of plants *in situ* with high precision by fusing the number of leaves and different type soybean pods recognized by a deep neural network combined with GRNN which can speed up germplasm innovation and planting environmental factor optimization.

INTRODUCTION

Soybean is an important source of high quality protein and oil in the world, which contains about 42% protein, 20% oil, and 33% carbohydrate (Zhang et al., 2001). Soybean protein can enhance the body's immune function and play an important nutritional role in human tissues and cells. Soybean production for 2020 totaled 4.14 billion bushels, up 16% from 2019 in America due to a higher average soybean yield (Alsajri et al., 2020). Soybean yield prediction is of great importance to global food production, which is a highly complex trait determined by multiple factors such as genotype, environment, and their interactions. Accurate soybean yield prediction is important for germplasm innovation and planting environment factor improvement. Many researchers have tried to clarify the phenotype (such as yield) as an explicit function of the genotype (G), environment (E), and their interactions (G×E). In fact, the selection of individuals with good genotypic effect can further improve the yield of existing soybean varieties, which is also of great significance for high-yield soybean breeding. So, the earlier the prediction during the growing season the better. But until now, soybean yield has been popularly determined by manual weight measurement after soybean plant harvest which is time-consuming, expensive, and imprecise.

In recent years, digital image processing combined with machine learning technology has been applied for crop yield prediction in literature. The relationship between grain area and weight was studied using an image processing method (Zhao et al., 2019). A citrus fruit crop prediction algorithm based on color difference of citrus fruit and leaves was studied in citrus trees (Dorj et al., 2017). A region growing algorithm was proposed to segment cotton bolls into color images and count them and predict yield (Sun et al., 2019). Algorithms that rely on feature extraction (SVM, NN, RF) and algorithms that do not need feature extraction (GoogLeNet, VGG-16) were compared. The study found that the VGG-16 algorithm could effectively distinguish corn and soybean (Flores et al., 2020). In addition, a multi-rotor UAV system was developed to obtain high-resolution images and information related to geographical location, shooting angle, and environmental illumination, so as to carry out effective agricultural detection (Zhu et al., 2019).

More recently, deep neural networks have been employed in crop yield prediction, including the convolutional neural network (CNN), faster region-convolutional neural network (Faster R-CNN), single shot multibox detector (SSD), and You Only Look Once (YOLO), etc. The features of hyperspectral and color images was used to classify corn and estimate corn yield by CNN (Yang et al., 2021). Faster R-CNN has been modified for

detection and yield estimation of fruits (mangoes, pomegranates, tomatoes, apples, and oranges) (Behera et al., 2021), which has also been used in prediction of melon yield (Zhao et al., 2017). Comparison of Faster R-CNN and SSD in citrus counting and yield prediction has been carried out (Qin et al., 2021). A convolutional neural network combined with linear regression was used in sorghum spike identification and weight prediction (Zannou and Houndji, 2019). Lightweight YOLO was applied to predict the yield of oil palm fruit based on images collected by UAV (Junos et al., 2021). The traditional rectangular bounding box in YOLOv3 was replaced as a circular rectangular box for better positioning of tomatoes (Liu et al., 2020). A deep neural network was designed to study the influence of genotype, environment, and their interaction on yield prediction (Khaki and Wang, 2019). There are also studies using satellite images to predict small yield using machine learning methods such as the Gaussian process regression algorithm (Sharifi, 2021).

Existing deep learning detection targets such as strawberries (Yu et al., 2019), tomatoes (Hu et al., 2019), apple (Tian et al., 2019), pepper (Hespeler et al., 2021), etc. are obviously different from the background leaves and branches of plants, which brings convenience for fruit identification due to remarkable color difference. Although the color of cucumbers (Mao et al., 2020) and corn (Jin et al., 2018) ears are similar to the leaves and vines, they have large size, small number, and low density, which also reduces the difficulty of identification. In the yield prediction of wheat (Yang et al., 2019) and rice (Crisóstomo de Castro Filho et al., 2020), the dense clustering of wheat and rice grains increases the difficulty of detection, but fortunately, the leaves shield the ears of wheat and rice slightly. A mature soybean phenotype measurement algorithm called soybean phenotype measure-instance segmentation was proposed to calculate pod length, pod width, stem length, seed length, and seed width based on PCA combined with CNN (Li et al., 2021). For soybean yield detection, leaves and pods have similar color, pods are blocked by numerous leaves, moreover the pods are clustered together, which creates a huge challenge for soybean yield forecast. Moreover, the types of pods must also be identified at the same time to predict soybean production accurately because the number of grains in different types of pods are different.

Deep neural networks, unlike early shallow neural networks with a single hidden layer, have multiple hidden layers which can effectively reveal the underlying unknown and highly non-linear relationship between the input data and output variables (LeCun et al., 2015), which have been widely employed in face recognition, automatic driving, etc., but they also require more hardware and time consumption. Generally, more neural

network layers and nodes means the network is more powerful and paradoxically needs more time and hardware to run.

After comparing several mainstream deep neural networks, the YOLOv3 algorithm was chosen to recognize soybean pods and leaves in this paper. Moreover, in order to further improve the detection performance of the neural network, structure improvements were made to YOLOv3 to achieve accurate recognition of large leaves and small pods simultaneously. In addition, a generalized regression neural network (GRNN) model was established for prediction of seed number in a soybean plant by using the cumulative results of leaves and different type pods of four images taken at 90-degree intervals from different directions. Finally, soybean plant production was calculated based on the average grain weights of different type pods, which provided a new method and solution for soybean phenotype detection and germplasm innovation acceleration.

The remainder of this paper is organized as follows. Section Materials and Methods describes the materials and data used in this research and offers a detailed description of our improved deep neural network for soybean pods and leaves prediction, and yield modeling as well. Section Results and Discussion presents the results of our algorithm and models. Finally, the conclusion is given in section Conclusion.

MATERIALS AND METHODS

In this paper, the soybean yield prediction method included three steps as shown in **Figure 1**. First, the original imaging data of soybean plants collected were preprocessed by filtering and enhancement. Then the improved YOLOv3 model was used to identify and count leaves and different type soybean pods including one bean pods, two bean pods, three bean pods, four bean pods, and five bean pods. Finally, a GRNN model for predicting soybean yield was established based on the numbers of leaves and all categories of pods as inputs.

Materials and Image Sampling

In this study, soybean plants were grown in pots in Wanjiang Experimental Station of Nanjing Agricultural University where the latitude and longitude are 118.62°E and 31.54°N,

respectively. Each pot had four soybean plants which were evenly distributed in the pot. There were 90 pots, 360 plants, and 24 varieties in total (Xudou-18, Heidou-2, Erzaohuang, Qiyuehuang-1, Bayuehuang-4, Qingyuanxiaoqingdou, Fengdudahuangdou, Enwangheidou, Kaijiangdongdou, Huazhouhuangdou, Ganyulianmaoshao, Liyangmaojiajia, Andingxiaoheidou, Nannong1606, P06, P12, P23, P25, P53, P59, P65, GS171761, GS71244, GS71411) used for the study. All potted soybean plants were at the pod-setting stage during image sampling.

A camera (model: Intel RealSense D435 manufactured by Intel) was employed to capture soybean plant images at three different time periods (6:00–7:00, 13:00–14:00, 17:00–18:00) in order to realize the completeness of images in different light intensity environments. The camera was 1.2 m above the ground and 1.5 m horizontally from the target while image capturing. Each pot of soybean plants was photographed in four different directions at intervals of 90 degrees, so a total of 360 images were collected.

Data processing was performed on a computer with a Win10 operating system, Intel(R) Core(TM) i7-8750H processor, 8GB memory, and a NVIDIA GeForce GTX 1050 Ti display adapter. The open source deep learning framework Tensorflow2.0 was used to establish models.

Soybean Pods and Leaves Recognition
Image Processing
In order to extract clear soybean plant images, denoising and enhancement treatments were carried out on soybean plant pictures due to the light environment and haze which blurred the pictures. First, *guide filter* was selected for denoising images after comparing them with the *bilateral filter* (Yu et al., 2020) and *DWT* (Rai et al., 2012) algorithms. Then, the *gamma* algorithm was chosen for image enhancement by contrast with *laplus* (Bhairannawar et al., 2017) and *log* (Maini and Aggarwal, 2010) algorithms. The pseudocode is shown in **Figure 2**.

Since deep neural networks require a certain number of training sets to improve the accuracy of the model, a data augmentation method, a common technique in deep learning research, was employed to increase the number of existing

FIGURE 1 | Flowchart of the soybean yield prediction method.
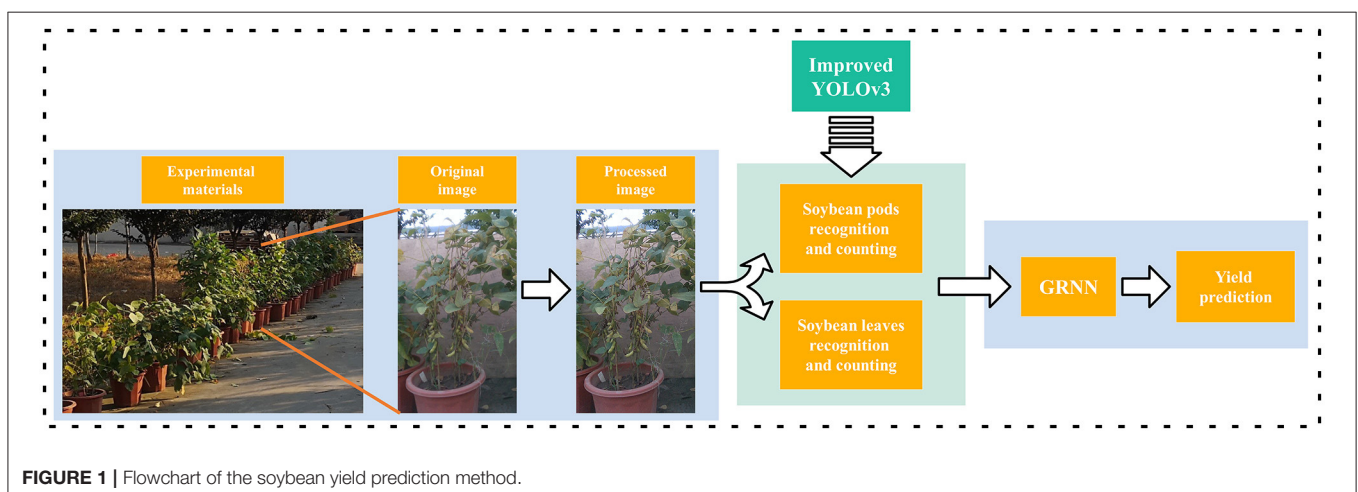
Input：filtering input image p,guidance image I,radius r, regularization ϵ, control parameter c、γ.

Output：filtering output q. enhancement output s.

Guided Filter：

1：$mean_I = f_{mean}(I)$
$mean_p = f_{mean}(p)$
$corr_I = f_{mean}(I.*I)$
$corr_{I_p} = f_{mean}(I.*p)$

2：$var_I = corr_I - mean_{I.}*mean_I$
$cov_{I_p} = corr_{I_p} - mean_{I.}*mean_p$

3：$a = cov_{I_p.}/(var_I + \epsilon)$
$b = mean_p - a.*mean_I$

4：$mean_a = f_{mean}(a)$
$mean_b = f_{mean}(b)$

5：$q = mean_{a.}*I + mean_b$

Gamma Transformation：

$$s = c*q^{\gamma}$$

FIGURE 2 | The pseudocode of denoising and enhancement algorithms.

photos to 1,800 images by rotation, scaling, mirroring, random brightness increases/decreases, and other methods. The 1,800 images were randomly divided into the training set and test set according to the ratio of 4:1, which were then used as the dataset of target detection models. The image processing process is shown in **Figure 3**.

Soybean Leaves, Pods, and Types Recognition

With the rapid improvement of the computing power of computers, deep learning has made tremendous progress, and a lot of target detection algorithms based on deep learning have been proposed. Popular deep learning algorithms such as Faster R-CNN (Ren et al., 2016), feature pyramid network (FPN, Lin et al., 2017a), SSD (Liu et al., 2016), and YOLO (Redmon et al., 2016) have been applied in different areas and show very superior performance. So, these algorithms were utilized in soybean pod and type recognition in the paper.

The detection of Faster R-CNN includes two steps, the first step is region proposal network (RPN). Features of a picture are extracted by a VGG-16 (Simonyan and Zisserman, 2014) neural network, followed by foreground background classification and first prediction of the coordinate of the generated anchor. In the second step, candidate boxes with higher confidence are selected and sent to the back of the network for the second prediction
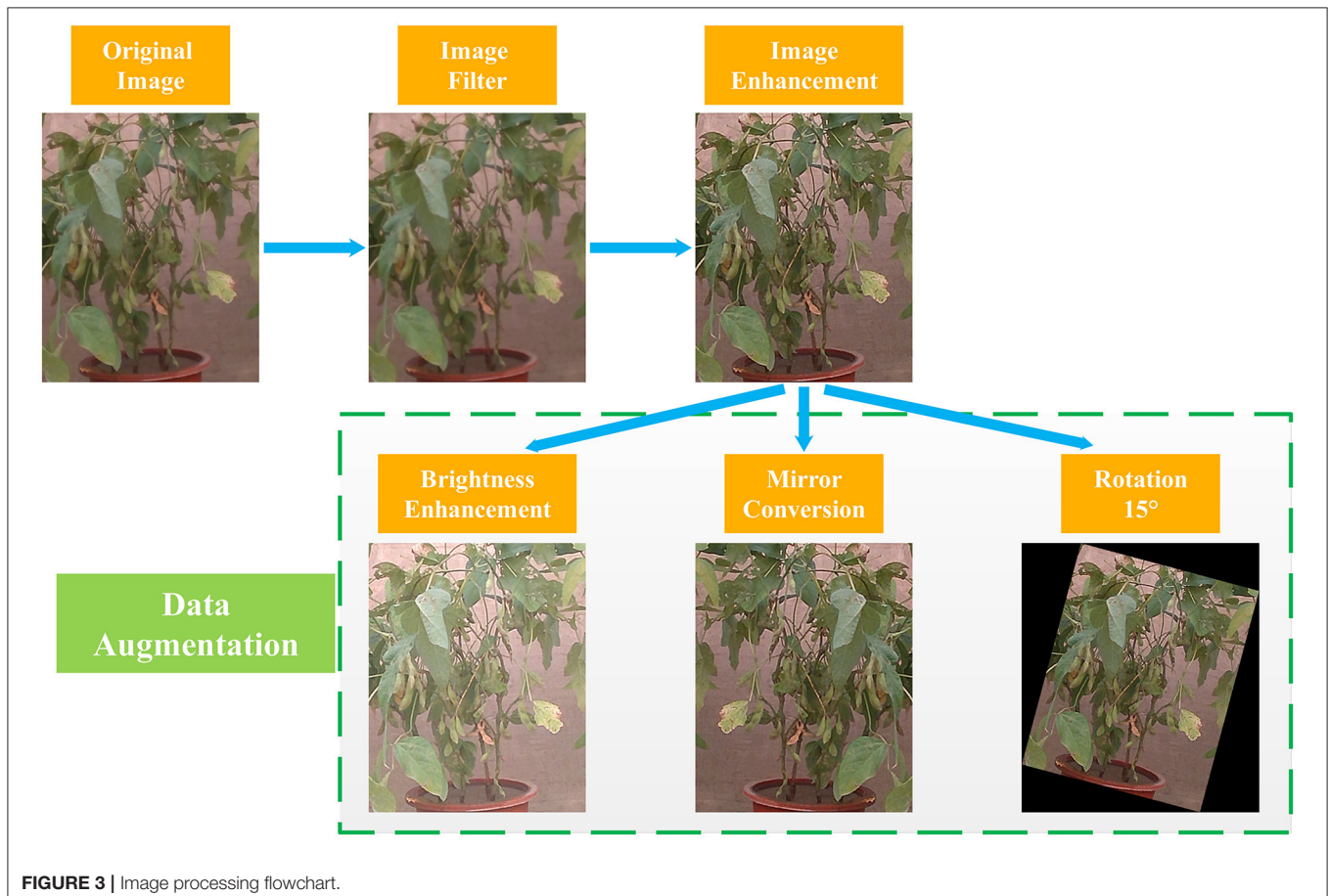
FIGURE 3 | Image processing flowchart.

of category and coordinate values, so as to predict the specific category. Faster R-CNN has extremely high prediction accuracy but is time-consuming for training (Benjdira et al., 2019).

The feature pyramid network (FPN) is proposed to alleviate the problems of multi-scale and small target detection which selects ResNet50 (He et al., 2016) as the feature network, and fuses the high-level features with the low-level features through the up-sampling process. Moreover, layers of the same class have horizontal connections and each layer is predicted independently. Therefore, the network has more abundant features. In addition, the idea of multi-scale detection is introduced into the RPN, and the anchor frame can be generated in different scales to cover different sized objects. FPN has good accuracy and precision but again, is more time-consuming.

In the single shot multibox detector (SSD) model, there is only one detection process step. The model uses VGG as the feature extraction network, and uses multiple feature layers to predict the target. As a result, the whole model is more lightweight with both good detection accuracy and speed.

The design concept of YOLO and SSD is similar, both are one stage target detection algorithms for reaching fast performance. In YOLOv3 (Redmon and Farhadi, 2018), a 28M DarkNet-53 is employed in which its parameters are only half that of ResNet101 (Lin et al., 2017b), but the performance is close to it. Pre-clustering of YOLOv2 (Redmon and Farhadi, 2017) is inherited in YOLOv3 on the anchor frame for targets clustering in the data using nine anchor frame scales that mostly fit the targets. Moreover, the feature processing is carried out in three different levels by introducing the multi-scale concept of FPN. And multiple binary classifiers are used in the calculation of classification loss to avoid competition within the class.

There are some problems such as error detection and missing detection by using YOLOv3 for soybean leaves and pods recognition because the color of the pods is similar to the leaves and the detection performance of YOLOv3 for small target like pods is not ideal. According to the above problems, an improved approach of YOLOv3 was proposed by changing the network structure and changing the clustering algorithm for increasing prediction without sacrificing too much speed. Firstly, the feature

map after one down-sampling session is superimposed to the input of the second and third residual blocks so as to increase the detail information in the deep feature map for detection. The improved network of YOLOv3 is shown in **Figure 4**. DBL stands for Darknetconv2D_BN_Leaky, and resn stands for the number of Res_units contained in the Res_block.

In the target detection algorithm, the anchor frame can help the model to fit the coordinate points of the target, so that the positioning task can be transformed from finding the position of the target in the image to learning the coordinate offset of the anchor frame relative to the target. In consideration of the fact that the traditional K-means algorithm is easy to implement, but the initial clustering center needs to be set artificially, in addition, different centers have great influence on the clustering results, the K-means++ algorithm (Arthur and Vassilvitskii, 2006) was selected to cluster the length and width of the anchor frame to make it fit the soybean pod and leaves better. The clustering used Euclidean distance as a metric and set k from 6 to 11 step by step. Because a large value of k reduced the convergence rate of the model, k was set to 9 in this study after several attempts. The clustering centers obtained after the convergence of the model were: (23,31), (27,21), (41,52), (64,42), (78,90), (76,103), (91,94), (106,97), and (113,107).

In terms of loss function, traditional YOLOv3 uses IoU Loss, which is composed of coordinate regression loss, confidence loss, and classification loss, and its calculation formula is shown in Equation 1.

$$L_{IoU} = 1 - IoU \qquad (1)$$

IoU Loss has the characteristics of scale invariance. Although IoU Loss is more advantageous than mean square error, when the relationship of the prediction box and the real box is contained and being contained, IoU Loss will be a fixed value, which has great influence on the detection effect. When the two do not intersect, the value of IoU Loss is 0, which cannot be optimized. According to the above problems, DIoU Loss (Distance IoU Loss) was used to replace IoU Loss in traditional YOLOv3, which
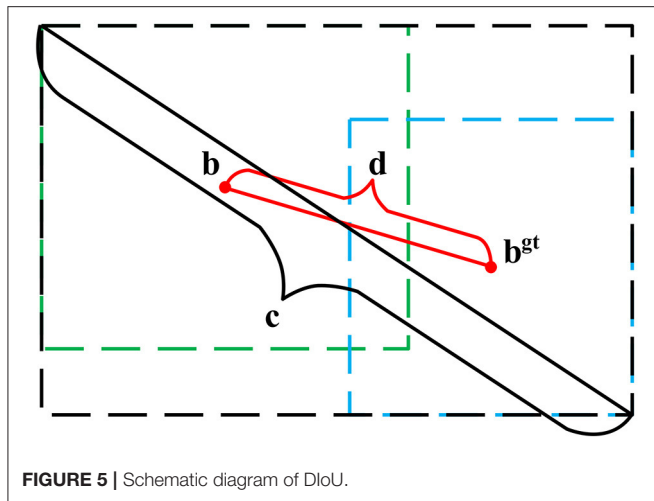

FIGURE 4 | Improved YOLOv3 for soybean leaves, pod, and type identification.
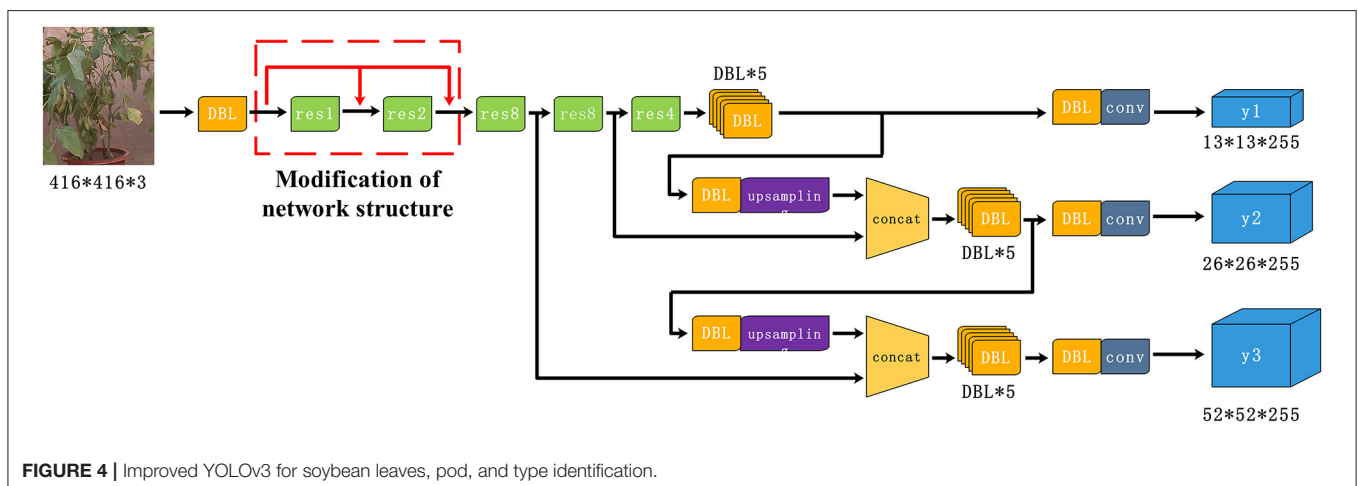
FIGURE 5 | Schematic diagram of DIoU.

made the regression of the target box more stable. The schematic diagram is shown in **Figure 5**.

DIoU is defined as,

$$DIoU = IoU - \frac{\rho^2\left(b, b^{gt}\right)}{c^2} \tag{2}$$

where b is the central coordinate of the prediction box, b^{gt} is the center coordinate of the real box, ρ is the Euclidean distance of the two center points, and c is the diagonal length of the minimum outer rectangle of the two target bounding boxes.

The final definition of the loss function DIoU Loss is shown in Equation 3.

$$L_{DIoU} = 1 - DIoU \tag{3}$$

In this paper, the improved YOLOv3 model was used to train the enhanced dataset. The size of network input was 416*416*3. The batch size was 64, the value of subbatch was 16, the momentum of dynamic parameter was 0.9, the maximum number of iterations was 14,000, the learning rate strategy was step decreasing, the initial value was 0.001, the scale parameter was 0.1, and the two step values of learning rate change were 11,200 and 12,600, respectively.

Soybean Yield Prediction

In the soybean yield prediction task based on plant images, the counting accuracy of pods was affected because some pods were occluded by leaves, and the density of leaves was positively correlated with the number of occluded pods. In addition, pods were divided into several types due to the difference number of seeds in them. The improved YOLOv3 model was used to identify the number of leaves and different types of pods, then PLSR (Geladi and Kowalski, 1986), BPNN (Hecht-Nielsen, 1992), and GRNN (Specht, 1991) models were established respectively to predict the amount of seeds. Among them, GRNN proposed by Specht has a strong non-linear mapping ability and learning speed which is an improved technique in neural networks based on non-parametric regression. It can even obtain good prediction

accuracy but only requires a small number of datasets (Izonin et al., 2021). Moreover, the network can also handle unstable data in the inputs, which is suitable for soybean yield prediction. The architecture of GRNN is shown in **Figure 6**.

The mathematical representation is as below,

$$Y(x) = \frac{\sum_{k=1}^{N} y_k K(x, x_k)}{\sum_{k=1}^{N} K(x, x_k)} \tag{4}$$

where input x includes the numbers of soybean leaves and different type pods, $Y(x)$ is the predicted value of soybean yield, y_k is the activation weight for the pattern layer neuron at k, and $K(x, x_k)$ is the radial basis function kernel (Gaussian kernel) as formulated below.

$$K(x, x_k) = e^{-d_k/2\sigma^2} \tag{5}$$

$$d_k = (x - x_k)^T (x - x_k) \tag{6}$$

where d_k is the squared Euclidean distance between the training samples x_k and the input x.

Evaluation Indices

To evaluate the performance and stability of the proposed model, parameters such as prediction precision, recall, and degree of integration (IoU) were defined. Evaluation of the performance of the detection and recognition model is an essential stage. The detection accuracy and complexity are the key performance indexes in the evaluation. The basic evaluation indexes of the target detection model are accuracy rate (P) and recall rate (R). The definitions of the two indicators are shown in Equations 7, 8. Accuracy rate represents the ratio that the detected target really belongs to this class, which is used to describe the credibility of the target. Recall rate represents the ratio of the detected target to the actual total amount of the target, which is used to describe the degree of the target being found completely.

$$P = \frac{T_P}{T_P + F_P} \tag{7}$$

$$R = \frac{T_P}{T_P + F_N} \tag{8}$$

T_P–The number of correctly predicted targets
F_P–The number of wrongly predicted targets
F_N–The number of missed predicted targets

Mean average precision is used to measure the overall effect of multi-classification detection by averaging the detection precision of all categories. Its definition is shown in Formula 3.

$$MAP = \frac{\sum_{k=1}^{N} P(k)}{N} \tag{9}$$

Intersection over Union (IoU) is introduced to measure the similarity between the prediction box and real box according to the characteristics of the target detection task. Its definition is shown in Formula 4. When the prediction box is exactly
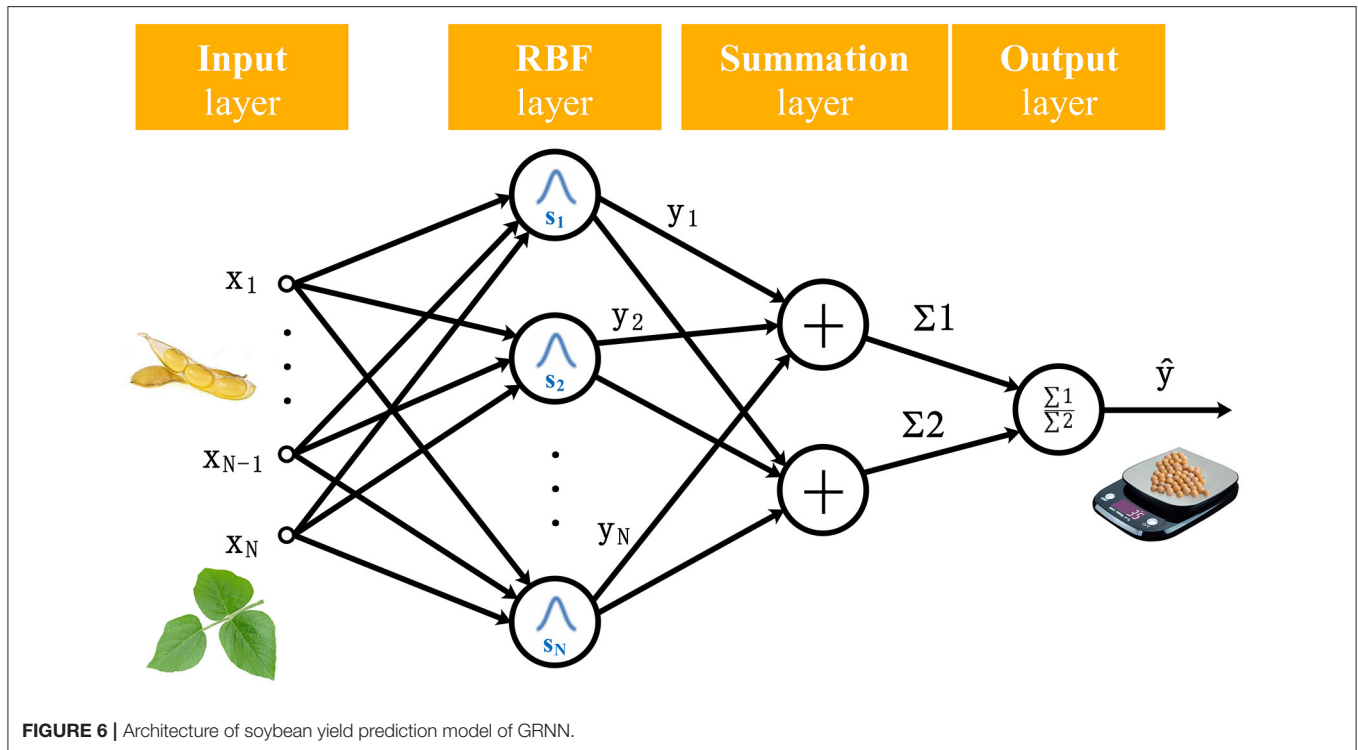
FIGURE 6 | Architecture of soybean yield prediction model of GRNN.

consistent with the real box, IoU is 1. Generally, the target is considered successfully detected when IoU is >0.5.

$$IoU = \frac{S_A \cap S_B}{S_A \cup S_B} \qquad (10)$$

Soybean detection and recognition is a dichotomous problem, which only involves foreground soybean and background soybean. Therefore, F_1 was introduced to evaluate the model accuracy comprehensively. The value of F_1 depends on the accuracy and recall. Its definition is shown in Equation 11.

$$F_1 = \frac{2PR}{P + R} \qquad (11)$$

The running speed of the detection and recognition algorithm is also an important basis for model evaluation. FPS was adopted as the evaluation standard in this experiment. Its definition is shown in Equation 12. In Equation 12, N represents the total number of samples and T represents the running time.

$$FPS = \frac{N}{T} \qquad (12)$$

RESULTS AND DISCUSSION

Soybean Leaves and Pod Types Recognition Results

The performance of popular existing algorithms such as Faster R-CNN, FPN, SSD, and YOLOv3 in soybean pod detection tasks is shown in **Table 1**.

TABLE 1 | Soybean pod detection performance of different models.

Model		Accuracy P	Recall R	F_1 value	Speed
Existing algorithms	Faster R-CNN	86.2%	80.5%	83.3%	13FPS
	FPN	**89.8%**	82.7%	86.1%	7FPS
	SSD	80.1%	74.2%	77.0%	24FPS
	YOLOv3	87.4%	81.6%	84.4%	**39FPS**
Improved YOLOv3		**90.3%**	**87.6%**	**88.9%**	36FPS

The bold values are the best results.

The table indicates that FPN showed obvious advantages over the existing popular algorithms in prediction accuracy and $F1$ index up to 89.8% and 86.1% due to its relatively large structure, followed by YOLOv3, Faster R-CNN, and SSD with 87.4, 86.2, and 80.1% in accuracy and 84.4%, 83.3, and 77.0 in $F1$, respectively. YOLOv3 had the fastest speed at 39 FPS, followed by SSD, Faster R-CNN, and FPN with 24 FPS, 13 FPS, and 7 FPS, respectively. Among them, the speed of YOLOv3 was more than five times that of FPN, but the accuracy was 2.4% lower. Therefore, YOLOv3 was the best algorithm considering the prediction accuracy and speed comprehensively.

The accuracy of the improved YOLOv3 algorithm was 3.32% higher than that of YOLOv3, reaching 90.3%, and also better than that of FPN, SSD, and Faster R-CNN. The speed was second only to YOLOv3, but significantly faster than SSD, Faster R-CNN, and FPN, reaching 36 FPS, which can meet the demand of real-time recognition.

Moreover, considering that the number of seeds in different pods varied greatly, the pods were further classified into single
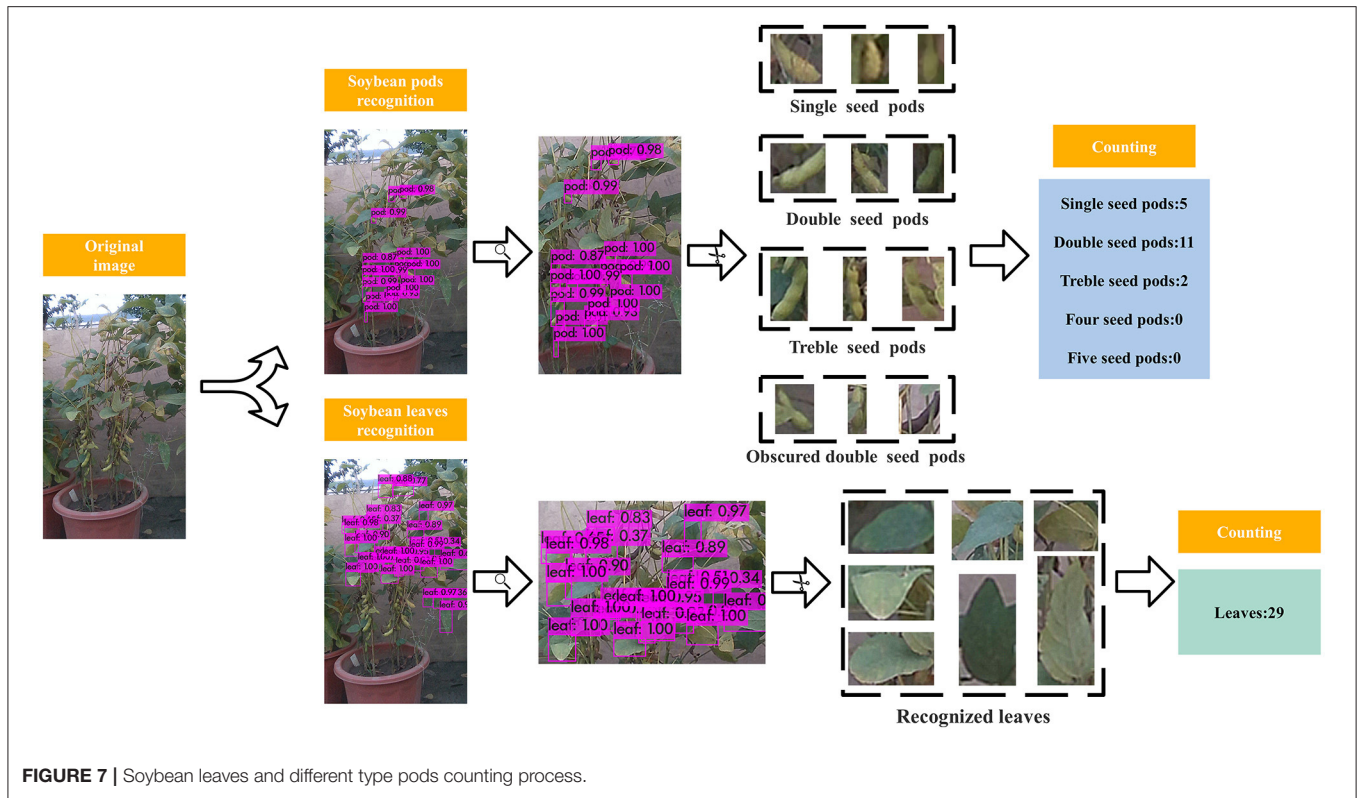
FIGURE 7 | Soybean leaves and different type pods counting process.

seed pods, double seeds pods, treble seeds pods, four seeds pods, and five seeds pods according to the number of internal seeds using the above proposed YOLOv3 model.

In order to accurately identify the pod type, pod type labeling and training were carried out on the recognized pod output anchor frame, and an accurate pod type recognition model was obtained. Soybean leaves and different type pods recognition and the counting process is shown in **Figure 7**.

Soybean Plant Yield Prediction Results

In the prediction of soybean yield, taking the non-uniform characteristics of leaves growth into account, four pictures were taken from four directions of soybean plants at 90 degrees apart, and the total number of leaves and the total number of different types of pods of the four images were extracted by the above mentioned improved YOLOv3 algorithm, and then the seed number prediction models of PLSR, BPNN, and GRNN were established with the input of leaf number and different type pods number, of which the results shown in **Table 2** indicate that the GRNN model had the highest prediction accuracy. **Figure 8** shows the comparison between the actual yield of soybean per pot and the predicted yield when the GRNN model was used for yield prediction. After the model was run three times, the average accuracy of the GRNN model was up to 97.31%.

As shown in Equation 13, y_i and y'_i represent the actual value and predicted value of soybean seeds respectively, and n stands for the number of pots of all soybeans. ACC is the accuracy calculated according to the deviation degree between

TABLE 2 | Prediction accuracy of different models.

Model	*ACC*1	*ACC*2	*ACC*3	$\overline{ACC}$
PLSR	95.57%	96.24%	95.76%	95.84%
BPNN	96.57%	96.97%	96.59%	96.71%
GRNN	97.24%	97.50%	97.20%	**97.31%**

*The bold values are the best results.*

the predicted value and the actual value of soybean seed number, which can reflect the performance of the prediction models.

$$ACC = \frac{\sum_{i=1}^{n} (y_i - \left| y_i - y'_i \right|)/y_i}{n} \quad (13)$$

To further measure the weight of soybean grains produced by plants, the average weight of 100 soybean seeds ($w_a = 0.203$ g) and the average weight of 100 soybean seeds for each pod type 5 to 10 days after soybean harvesting from random sampling were measured. Among them, the average weight of soybean grains in single seed pods w_1, double grains pods w_2, treble grains pods w_3, four grains pods w_4, and five grains pods w_5 were 0.242 g, 0.207 g, 0.196 g, 0.189 g, and 0.186 g, respectively. Based on the numbers of different type pods predicted by the improved YOLOv3, the total weight of soybean grains produced was given, as shown in **Table 3**.

The accuracy of soybean yield predicted by GRNN was 95.14% with the average weight of 100 grains, while the accuracy of
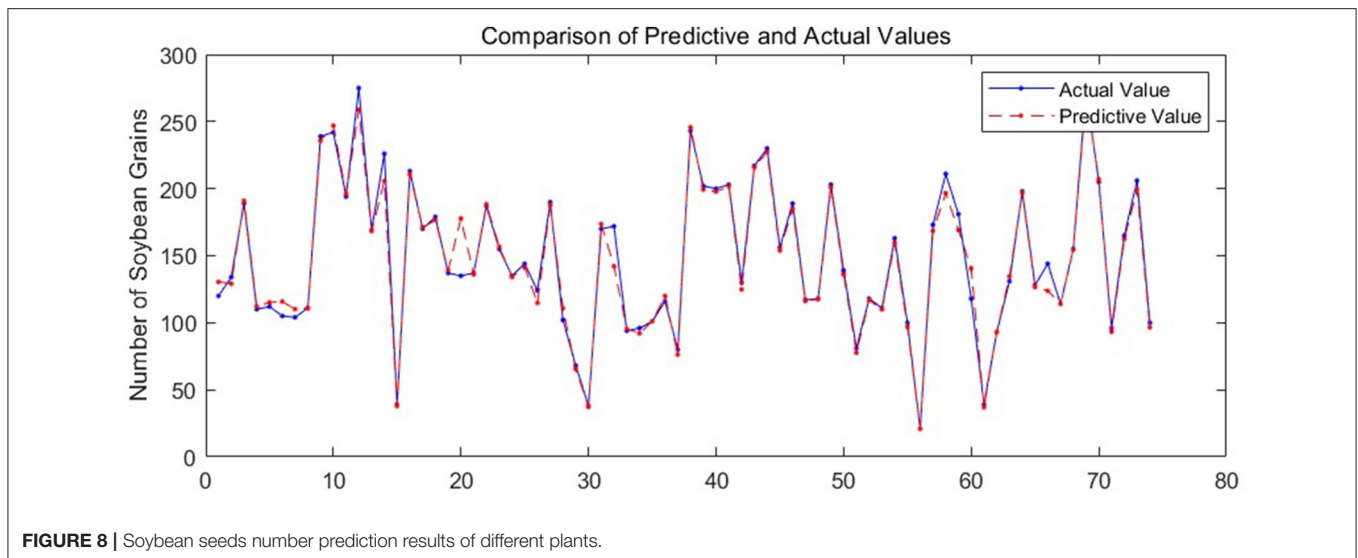
FIGURE 8 | Soybean seeds number prediction results of different plants.

TABLE 3 | Weight of single soybean grain of different pod types and total yield prediction results.

Pod type	Single (w_1)	Double (w_2)	Treble (w_3)	Four (w_4)	Five (w_5)	All types of pods (w_a)
Average weight per grain (g)	0.242	0.207	0.196	0.189	0.186	0.203
Total weight (g)			2668.670			2605.911
Accuracy			**97.43%**			95.14%

The bold values are the best results.

soybean yield predicted by GRNN with average weights of five different pods (w_1, w_2, w_3, w_4, and w_5) increased to 97.43%. The results show that it is more accurate to predict soybean yield by identifying and classifying soybean pods combined with GRNN.

CONCLUSION

Most of the existing crop yield prediction methods studied the impact of environmental changes on yield, but paid no attention to the actual yield prediction. Therefore, the most used field yield measurement method is still the traditional manual sampling process for statistical calculation, which is inefficient with low precision. Due to the high cost, time-consuming, and low accuracy of the traditional manual soybean yield measurement approach, this paper proposed a soybean yield *in situ* prediction method based on bean pods and leaves image recognition using a deep learning algorithm combined with a generalized regression neural network (GRNN). YOLOv3 is generally superior to Faster R-CNN, FPN, and SSD in terms of prediction accuracy and speed. Moreover, YOLOv3 was improved by changing the IoU loss function, using the anchor frame clustering algorithm, and utilizing the partial neural network structure in which recognition precision increased by 2.9% up to 90.3% at 36 FPS.

In this paper, we proposed to take four images of soybean plants at 90° intervals, and extract the total numbers of leaves and different type pods from the four images by improved YOLOv3. Then we established the prediction model of different type pods quantity of each plant using GRNN with inputs of the total numbers of leaves and different type pods recognized, in which

the average accuracy increased to 97.31%, which was better than PLSR and BPNN. Furthermore, the soybean grain yield was calculated using the number and average weight of each type of pod. The prediction accuracy of the yield weight was up to 97.43%, which was better than the prediction accuracy based on the total number of grains and the average weight of different type pod grains.

This study shows that the improved YOLOv3 algorithm can be used to identify the number of leaves and different type pods and, moreover, can achieve accurate soybean yield *in situ* prediction 30–40 days in advance combined with the average weight of different soybean pods, which provides a new solution for accelerating soybean germplasm innovation and phenotypic detection of other crops.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

WL proposed the conceptualization and methodology and wrote the paper. RD programmed the software. PN compared the performance of the algorithms. GX designed and carried out the experiments. HL and LS improved the methodology and conceived the experiments. All authors reviewed the manuscript.

FUNDING

REFERENCES

Alsajri, F. A., Wijewardana, C., Irby, J. T., Bellaloui, N., Krutz, L. J., and Golden, B. (2020). Developing functional relationships between temperature and soybean yield and seed quality. *Agron. J.* 112, 194–204. doi: 10.1002/agj2.20034

Arthur, D., and Vassilvitskii, S. (2006). *k-means++: The Advantages of Careful Seeding*. Stanford: Stanford University.

Behera, S. K., Rath, A. K., and Sethy, P. K. F. (2021). Yield estimation using Faster R-CNN with MIoU. *Multimed. Tools Appl.* 80, 19043–19056. doi: 10.1007/s11042-021-10704-7

Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., and Ouni, K. (2019). "Car detection using unmanned aerial vehicles: comparison between faster r-cnn and yolov3," in *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)* (Muscat: IEEE), 1–6.

Bhairannawar, S., Patil, A., Janmane, A., and Huilgol, M. (2017). "Color image enhancement using Laplacian filter and contrast limited adaptive histogram equalization," in *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)* (Vellore: IEEE), 1–5.

Crisóstomo de Castro Filho, H., Abílio de Carvalho Júnior, O., Ferreira de Carvalho, O. L., Pozzobon de Bem, P., Dos Santos de Moura, R., and Olino de Albuquerque, A. (2020). Rice crop detection using LSTM, Bi-LSTM, and machine learning models from sentinel-1 time series. *Remote Sens.* 12:2655. doi: 10.3390/rs12162655

Dorj, U. O., Lee, M., and Yun, S. (2017). An yield estimation in citrus orchards via fruit detection and counting using image processing. *Comput. Electron. Agric.* 140, 103–112. doi: 10.1016/j.compag.2017.05.019

Flores, P., Zhang, Z., Mathew, J., Jahan, N., and Stenger, J. (2020). Distinguishing volunteer corn from soybean at seedling stage using images and machine learning. *Smart Agric.* 2:61–74. doi: 10.12133/j.smartag.2020.2.3.202007-SA002

Geladi, P., and Kowalski, B. R. P. (1986). Least-squares regression: a tutorial. *Anal. Chim. Acta* 185, 1–17. doi: 10.1016/0003-2670(86)80028-9

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (San Juan, PR), 770–778.

Hecht-Nielsen, R. (1992). *Theory of the Backpropagation Neural Network[M]//Neural Networks for Perception*. Cambridge, MA: Academic Press, 65–93.

Hespeler, S. C., Nemati, H., and Dehghan-Niri, E. (2021). Non-destructive thermal imaging for object detection via advanced deep learning for robotic inspection and harvesting of chili peppers. *Artif. Intell. Agric.* 5, 102–117. doi: 10.1016/j.aiia.2021.05.003

Hu, C., Liu, X., Pan, Z., and Automatic, L., i. P. (2019). detection of single ripe tomato on plant combining faster R-CNN and intuitionistic fuzzy set. *IEEE Access* 7, 154683–154696. doi: 10.1109/ACCESS.2019.2949343

Izonin, I., Tkachenko, R., Zub, K., and Tkachenko, P. (2021). A GRNN-based approach towards prediction from small datasets in medical application. *Procedia Comput. Sci.* 184, 242–249. doi: 10.1016/j.procs.2021.03.033

Jin, S., Su, Y., Gao, S., Wu, F., Hu, T., and Liu, J. (2018). Deep learning: individual maize segmentation from terrestrial lidar data using faster R-CNN and regional growth algorithms. *Front. Plant Sci.* 9:866. doi: 10.3389/fpls.2018.00866

Junos, M. H., Khairuddin, A. S. M., Thannirmalai, S., and Dahari, M. (2021). Automatic detection of oil palm fruits from UAV images using an improved YOLO model. *Vis. Comput.* 1–15. doi: 10.1007/s00371-021-02116-3

Khaki, S., and Wang, L. (2019). Crop yield prediction using deep neural networks. *Front. Plant Sci.* 10:621. doi: 10.3389/fpls.2019.00621

LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Li, S., Yan, Z., Guo, Y., Su, X., Cao, Y., Jiang, B. (2021). SPM-IS: an auto-algorithm to acquire a mature soybean phenotype based on instance segmentation. *Crop J.* doi: 10.1016/j.cj.05.014. [Epub ahead of print].

Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017a). "Feature pyramid networks for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (San Juan, PR), 2117–2125.

Lin, T. Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017b). "Focal loss for dense object detection," in *Proceedings of the IEEE International Conference on Computer Vision* (Venice), 2980–2988. doi: 10.1109/ICCV.2017.324

Liu, G., Nouaze, J. C., Touko Mbouembe, P. L., and Kim, J. H. (2020). YOLO-tomato: a robust algorithm for tomato detection based on YOLOv3. *Sensors* 20:2145. doi: 10.3390/s20072145

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., and Fu, C. (2016). "Ssd: single shot multibox detector," in *European Conference on Computer Vision* (Cham: Springer), 21–37. doi: 10.1007/978-3-319-46448-0_2

Maini, R., and Aggarwal, H. (2010). A comprehensive review of image enhancement techniques. *arXiv preprint arXiv:4053.*

Mao, S., Ma, L., i,. Y., Zhang, Y., Zhou, B., Wang, J., et al. (2020). cucumber recognition algorithm for harvesting robots in the natural environment using deep learning and multi-feature fusion. *Comput. Electron. Agric.* 170:105254. doi: 10.1016/j.compag.2020.105254

Qin, X., Huang, R., and Hua, B. (2021). Research and implementation of yield recognition of Citrus reticulata based on target detection. *J. Phys. Conf. Ser.* 1820:012128. doi: 10.1088/1742-6596/1820/1/012128

Rai, R. K., Asnani, J., and Sontakke, T. R. (2012). Review of shrinkage techniques for image denoising. *Int. J. Comput. Appl.* 42, 13–16. doi: 10.5120/5799-8009

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV), 779–788.

Redmon, J., and Farhadi, A. (2017). "YOLO9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition* (Honolulu, HI), 7263–7271.

Redmon, J., and Farhadi, A. (2018). Yolov3: an incremental improvement. *arXiv preprint arXiv:02767.*

Ren, S., He, K., Girshick, R., and Sun, J. (2016). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 1137–1149. doi: 10.1109/tpami.2016.2577031

Sharifi, A. (2021). Yield prediction with machine learning algorithms and satellite images. *J. Sci. Food Agric.* 101, 891–896. doi: 10.1002/jsfa.10696

Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1556.*

Specht, D. F. (1991). A general regression neural network. *IEEE Trans. Neural Netw.* 2, 568–576. doi: 10.1109/72.97934

Sun, S., Paterson, L. I. C., Chee, A. H., Robertson, P. W., and Image, J. S. (2019). processing algorithms for infield single cotton boll counting and yield prediction. *Comput. Electron. Agric.* 166:104976. doi: 10.1016/j.compag.2019.104976

Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., and Liang, Z. (2019). Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Comput. Electron. Agric.* 157, 417–426. doi: 10.1016/j.compag.2019.01.012

Yang, W., Nigon, T., Hao, Z., Paiao, G. D., Fernández, F. G., and Mulla, D. (2021). Estimation of corn yield based on hyperspectral imagery and convolutional neural network. *Comput. Electron. Agric.* 184:106092. doi: 10.1016/j.compag.2021.106092

Yang, Y., Huang, X., Cao, L., Chen, L., and Huang, K. (2019). "Field wheat ears count based on YOLOv3," in *2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)* (Dublin: IEEE), 444–448.

Yu, H., He, F., Pan, Y. (2020). A scalable region-based level set method using adaptive bilateral filter for noisy image segmentation. *Multimedia Tools Appl.* 79, 5743–5765. doi: 10.1007/s11042-019-08493-1

Yu, Y., Zhang, K., Yang, L., and Zhang, D. (2019). Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. *Comput. Electron. Agric.* 163:104846. doi: 10.1016/j.compag.2019.06.001

Zannou, J. G. N., and Houndji, V. R. S. (2019). "Using machine learning," in *2019 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART)* (Paris: IEEE), 1–4.

Zhang, J., Mungara, P., and Jane, J. (2001). Mechanical and thermal properties of extruded soy protein sheets. *Polymer* 42, 2569–2578. doi: 10.1016/S0032-3861(00)00624-8

Zhao, S., Zheng, H., Chi, M., Chai, X., and Liu, Y. (2019). Rapid yield prediction in paddy fields based on 2D image modelling of rice panicles. *Comput. Electron. Agric.* 162, 759–766. doi: 10.1016/j.compag.2019.05.020

Zhao, T., Wang, Z., Yang, Q., and Chen, Y. (2017). Melon yield prediction using small unmanned aerial vehicles[C]//Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping II. *Int. Soc. Optics Photon.* 10218:1021808. doi: 10.1117/12.2262412

Zhu, J., Cen, H., He, L., and He, Y. (2019). Development and performance evaluation of a multi-rotor unmanned aircraft system for agricultural monitoring. *Smart Agric.* 1:43. doi: 10.12133/j.smartag.2019.1.1.201812-SA011