DETECTION METHOD OF CORN LEAF DISEASES BASED ON CA-YOLOv8

基于CA-YOLOv8的玉米叶病检测方法

Miao XU¹, Xuan ZHANG¹, Na MA¹, Yanwen LI¹

¹ College of Information Science and Engineering, Shanxi Agricultural University, Taigu/ China;
² College of Agricultural Engineering, Shanxi Agricultural University, Taigu/ China;
Tel: +86 18434764827; E-mail: zhangxuan727@126.com

Keywords: corn leaf, disease detection, YOLOv8, attention mechanism, Coordinate Attention(CA)

ABSTRACT

In order to achieve efficient and accurate detection of common corn leaf diseases such as leaf blight, gray spot disease, and rust, a corn leaf disease detection method based on CA-YOLOv8 was proposed. In this method, the Coordinate Attention(CA) attention mechanism was added after the feature map output from the Neck part to enhance the feature extraction capability of the model. The experimental results showed that the precision, recall and mean average precision(mAP) of the CA-YOLOv8 model on the test set were 94.08%, 90.53% and 97.38%, respectively. Compared with the YOLOv8, YOLOv8+SE and YOLOv8+CBAM models, the mAP was improved by 2.15, 0.86 and 2.35 percentage points, respectively. Compared with Faster R-CNN, YOLOv5s, YOLOv7, and YOLOv8 models, the mAP has increased by 63.53, 29.24, 3.21, and 2.15 percentage points, respectively. The study showed that the CA-YOLOv8 model can provide a technical reference for the development of a portable intelligent corn leaf disease detection system.

INTRODUCTION

Corn is a globally important crop for food, feed and industrial raw materials, and the stabilization of corn production plays an important role in food security, farmers’ incomes and the national economy (Song et al., 2023; Cui et al., 2023). However, corn diseases seriously affect corn production (Zhang et al., 2021). There are many types of corn diseases worldwide that are difficult to identify, among which corn gray spot, corn leaf blight and corn rust are the most common (Zibani et al., 2022). The traditional methods for detecting corn leaf disease mainly rely on manual observation and identification, which are not only inefficient but also limited by manual experience and skills (Song et al., 2023). Therefore, efficient and accurate detection of corn leaf disease is crucial for yield improvement.

In recent years, with the generation of large-scale labeled data and the rapid improvement of computer processing capabilities, deep learning technology has achieved rapid development in the field of plant disease detection (Yang et al., 2023; Zhang et al., 2024). Due to its high extraction of high-dimensional features from targets, the effect of plant disease detection in complex situations has been significantly improved. At present, plant disease detection algorithms based on deep learning mainly include multi-stage object detection algorithms represented by Faster R-CNN (Ren et al., 2017) and single-stage object detection algorithms represented by YOLO series (Shao et al., 2022).

Sun et al. proposed a multi-scale feature fusion instance detection method based on convolutional neural network for maize leaf blight detection with an mAP of 91.83%, which is about 20% higher than the original SSD algorithm (Sun et al., 2020). Zhang et al. optimized convolutional neural network using Multi-Activation Function (MAF) module to detect maize leaf disease, and used transfer learning and warm-up methods to accelerate training, improving the accuracy of traditional artificial intelligence methods (Zhang et al., 2021). Chen et al. proposed a lightweight corn disease recognition model, DFCANet (Dual Fusion Blocks with Coordinate Attention Networks), with an average recognition accuracy of 98.47% (Chen et al., 2022).
et al. proposed the CD-Mobilenetv3 model to identify corn leaf disease, replacing the SE module of original model with the EAC module, and introducing dilated convolution into the model. The accuracy on open source datasets reached as high as 98.23% (Bi et al., 2023). Dai et al. proposed a multi-task deep-learning-based system for enhanced precision detection and diagnosis of corn leaf diseases (MTDL-EPDCLD) to enhance the detection and diagnosis of corn leaf disease, and developed a mobile application utilizing the Qt framework (Dai et al., 2023). Song et al. proposed a high-accuracy detection method based on Attention Generative Adversarial Network (Attention-GAN) and few-shot learning. GAN are used to expand data and generate more training samples. An attention mechanism was introduced to enable the model to focus more on important parts of the image, thereby improving model performance (Song et al., 2023).

The above research indicates that deep learning has achieved certain results in the detection of corn leaf disease. However, most methods with high detection accuracy have problems such as multiple parameters, high computational power, and slow detection speed. Methods with low computational complexity and fast detection speed have lower detection accuracy, and the network's generalization and adaptability to complex environments are insufficient. YOLOv8 is the latest version of the YOLO (You Only Look Once) series of object detection algorithms released by Ultralytics. Compared to previous versions, YOLOv8 has faster inference speed, higher accuracy, and is easier to train and adjust (Hussain et al., 2023). Therefore, this study takes YOLOv8 as the basic model and improves it to realize efficient and accurate detection of corn leaf disease, which can not only ensure the stability of corn production, but also provide technical support for the development of portable intelligent corn leaf disease detection system.

MATERIALS AND METHODS

Dataset construction

The corn leaf disease images used in this study were from the publicly available dataset PlantDoc (https://github.com/pratikkyal/PlantDoc-Dataset) (Singh et al., 2020), which includes corn gray leaf spot, corn leaf blight, and corn rust. After screening, 70 corn gray spot images, 116 corn leaf blight images, and 114 corn rust images were obtained. Fig. 1 shows sample images of different leaf diseases in the dataset.

![Sample images of corn leaf disease](image)

Fig. 1 - Sample images of corn leaf disease

Due to the insufficient sample size of the dataset, the model cannot converge during the training process. In order to improve the effectiveness of the network training model and enhance its generalization ability, and prevent overfitting caused by insufficient training samples, image enhancement methods were used to expand the dataset, resulting in 1470 images of corn diseases including 500 images of corn gray spot, 486 images of corn leaf blight, and 484 images of corn rust. The training set, validation set and test set were divided according to the ratio of 8:1:1, and the annotation number of different leaf diseases in the divided dataset is shown in Table 1.
CA-YOLOv8 model

YOLOv8 belongs to the single-stage object detection algorithm, which only needs to extract features once to achieve object detection. Compared with previous networks, YOLOv8 improves the detection speed and accuracy while reducing the number of network parameters. For corn leaf disease images in different backgrounds, a lightweight version of YOLOv8n was selected as the base network for improvement. In this paper, CA(Coordinate attention) mechanism is added after the feature map output of the Neck part to improve the ability of the model to extract features and improve the detection accuracy of the model. The network structure name is defined as CA-YOLOv8. The CA-YOLOv8 network structure is shown in Fig. 2.

<table>
<thead>
<tr>
<th>Disease</th>
<th>training set</th>
<th>validation set</th>
<th>test set</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>corn gray leaf spot</td>
<td>529</td>
<td>56</td>
<td>72</td>
<td>657</td>
</tr>
<tr>
<td>corn leaf blight</td>
<td>813</td>
<td>71</td>
<td>96</td>
<td>980</td>
</tr>
<tr>
<td>corn rust</td>
<td>453</td>
<td>56</td>
<td>50</td>
<td>559</td>
</tr>
<tr>
<td>Total</td>
<td>1795</td>
<td>183</td>
<td>218</td>
<td>2196</td>
</tr>
</tbody>
</table>

Table 1 - Annotation number of different corn diseases

YOLOv8 model

The YOLOv8 network model uses the Backbone for feature extraction, the Neck for feature fusion, and the Head for detection and recognition. YOLOv8 integrates many excellent technologies in real-time object detection, and its main features are as follows:

1. The input end uses an adaptive scaling method to adjust the image size to 640*640, while using mosaic data augmentation to improve model robustness.
2. The backbone consists of CBS module, C2f module, and SPFF module using the CSP concept. The CBS module is a combination of convolution, normalization, and SiLU activation function, which can improve model stability, accelerate convergence speed, and prevent gradient disappearance. Compared to the C3 module, C2f adds skip connections and additional Split operations, allowing the model to obtain richer gradient flow information. SPFF adaptively fuses feature information of various scales through pooling and convolution operations, which can enhance the feature extraction ability of the model.
(3) The neck uses the idea of PAN-FPN to remove the convolution in the upsampling stage of PAN-FPN in YOLOV5, and replaces the C3 module with the C2f module.

(4) The head uses a Decoupled-Head structure to separate the classification and detection heads, capturing information on targets of different scales and improving the accuracy of object detection.

**CA attention mechanism**

Attention mechanism can help neural networks focus on information that is more critical to the current task, reduce attention to other information, and thus improve the performance and accuracy of the model. Squeeze and Excitation Networks (SENet) learn adaptive channel weights to make the model focus more on useful channel information (Hu et al., 2020), but only consider attention in the channel dimension and cannot capture attention in the spatial dimension. The Convolutional Block Attention Module (CBAM) extracts positional attention information through large-scale kernel convolution, but convolution can only extract local relationships and lacks the ability to extract remote relationships (Woo et al., 2018). Moreover, paying attention to images from both spatial and channel perspectives requires more computational resources and higher complexity. Coordinate Attention (CA) can encode horizontal and vertical position information into channel attention (HOU et al., 2021), enabling mobile networks to focus on a wide range of position information without excessive computational complexity. Fig. 3 shows a comparison diagram of different attention modules.

![Comparison of different attention modules](image)

The input of the CA attention mechanism is usually a feature map \([C, H, W]\), which is usually the output of a certain layer in a convolutional neural network (CNN). CA consists of two steps: coordinate information embedding and coordinate attention generation.

In the coordinate information embedding step, CA will perform global average pooling on the input feature map in the width and height directions, respectively, to obtain the feature maps \([C, H, 1]\) in the width direction and \([C, 1, W]\) in the height direction. This method captures remote spatial interactions with precise position information, avoids compressing all spatial information into the channel, and can help the network locate targets more accurately.

In the coordinate attention generation step, the above two feature mappings are firstly merged to obtain a new feature layer \([C, 1, W+H]\), and the merged feature layer is feature transformed using the \(1 \times 1\) convolution, BN normalization and nonlinear activation function operation to obtain a richer representation. Subsequently, the width and height direction features were separated from the above feature layers, and the two separate feature layers were transposed to restore the dimensions of width and height, resulting in two feature layers \([C, H, 1]\) and \([C, 1, W]\). Afterwards, \(1 \times 1\) convolution and sigmoid function are applied to the separated feature layers for feature transformation to make their dimensions consistent with the input, respectively. Finally, the original input feature map is multiplied by the attention scores in the width and height directions to obtain the output of the CA attention mechanism.
Experimental Platform and Network Parameter Settings
The experiment used a GPU model of NVIDIA GeForce RTX 3090, running memory of 35GB, programming language of python3.6, deep learning framework of pytorch 1.12, and GPU acceleration library of CUDA 11.4. To shorten the training time of the network, the freezing training method is adopted. When freezing the backbone network for training, the initial learning rate is set to 0.01 and the batch processing volume is set to 2. After unfreezing, the learning rate of network training is set to 0.0001, the batch size is set to 2, and the number of iterations is 100.

Model evaluation indicators
In order to verify the detection performance of the algorithm proposed in this study, Precision (P), Recall (R), Average Precision (AP), and Mean Precision (mAP) were used as evaluation indicators, and the model performance was evaluated by comparing the detection image differences between the model in this paper and other models.

RESULTS
Experiment of CA-YOLOv8 model
In order to verify the performance of the CA-YOLOv8 model, 147 corn leaf disease images in the test set were tested and evaluated, and Table 2 shows the detection results of CA-YOLOv8 algorithm on different corn leaf diseases. Table 2 shows that the average accuracy of the CA-YOLOv8 algorithm can reach 97.38%, with an accuracy rate of 94.08% and a recall rate of 90.53%.

<table>
<thead>
<tr>
<th>Disease</th>
<th>P  [%]</th>
<th>R  [%]</th>
<th>AP [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>corn gray leaf spot</td>
<td>95.77</td>
<td>93.06</td>
<td>99.37</td>
</tr>
<tr>
<td>corn leaf blight</td>
<td>90.82</td>
<td>88.54</td>
<td>97.84</td>
</tr>
<tr>
<td>corn rust</td>
<td>95.65</td>
<td>90.00</td>
<td>94.93</td>
</tr>
<tr>
<td>Average</td>
<td>94.08</td>
<td>90.53</td>
<td>97.38</td>
</tr>
</tbody>
</table>

Part of the detection examples are shown in Fig. 4. It can be seen that CA-YOLOv8 algorithm proposed in this paper can accurately detect different corn leaf diseases, and the recognition results are also good for the complex background in Fig. 4 (d) ~ (f). Due to the high risk of misjudgment or loss in complex backgrounds, the CA-YOLOv8 model can focus on a wide range of positional information to achieve accurate detection of different leaf diseases. In summary, the CA-YOLOv8 model proposed in this paper can accurately detect different corn leaf diseases, and has good detection performance for small targets, multiple targets, complex backgrounds, etc.
Comparison of CA-YOLOv8 model and YOLOv8 model with different attention mechanisms

To verify the effectiveness of the attention mechanism used in this paper, the improved YOLOv8 model with CA (abbreviated as CA-YOLOv8) was compared with the original YOLOv8 model (abbreviated as YOLOv8), the YOLOv8 model with SE (abbreviated as YOLOv8+SE), and the YOLOv8 model with CBAM (abbreviated as YOLOv8+CBAM) in disease detection experiments on the test set. As shown in Table 3, the CA-YOLOv8 model has the highest AP for three different corn leaf diseases. Compared with the YOLOv8, YOLOv8+SE, and YOLOv8+CBAM models, its mAP has increased by 2.15, 0.86, and 2.35 percentage points, respectively. The CA-YOLOv8 model improved detection accuracy with almost no increase in model memory usage.

Table 3

<table>
<thead>
<tr>
<th>Model</th>
<th>AP corn gray leaf spot [%]</th>
<th>AP corn leaf blight [%]</th>
<th>AP corn rust [%]</th>
<th>mAP [%]</th>
<th>Size of the Weight Files [MB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>YOLOv8</td>
<td>98.85</td>
<td>95.52</td>
<td>91.31</td>
<td>95.23</td>
<td>42.7</td>
</tr>
<tr>
<td>YOLOv8+SE</td>
<td>97.64</td>
<td>96.49</td>
<td>95.42</td>
<td>96.52</td>
<td>42.8</td>
</tr>
<tr>
<td>YOLOv8+CBAM</td>
<td>96.82</td>
<td>95.15</td>
<td>93.12</td>
<td>95.03</td>
<td>43.0</td>
</tr>
<tr>
<td>CA-YOLOv8</td>
<td>99.37</td>
<td>97.84</td>
<td>94.93</td>
<td>97.38</td>
<td>42.9</td>
</tr>
</tbody>
</table>

Fig. 5 (a) shows the detection results of multiple corn gray spot disease leaves under complex background, Fig. 5 (b) shows the detection results of single corn leaf blight disease leaf under complex background, and Fig. 5 (c) shows the detection results of single corn rust disease leaf. It can be seen that the YOLOv8, YOLOv8+SE, YOLOv8+CBAM, and CA-YOLOv8 models can correctly detect corn leaf disease in different backgrounds, while the CA-YOLOv8 model has higher detection accuracy.
In order to explore the detection effects of different models on three types of corn leaf diseases, the performance of different models on the test set were compared, as shown in Table 4. The results showed that the CA-YOLOv8 model had the highest average accuracy in detecting three types of corn leaf diseases compared with other models. Compared with the Faster R-CNN, YOLOv5s, YOLOv7 and original YOLOv8 models, the mAP of CA-YOLOv8 model improved by 63.53, 29.24, 3.21 and 2.15 percentage points respectively. The memory usage of the CA-YOLOv8 model is only 0.1MB higher than that of the original YOLOv8 model.

**Comparison with other models**

![Detection results of YOLOv8 model with different attention mechanisms](image)

*Fig. 5 - Detection results of YOLOv8 model with different attention mechanisms*

<table>
<thead>
<tr>
<th>Model</th>
<th>AP corn gray leaf spot [%]</th>
<th>AP corn leaf blight [%]</th>
<th>AP corn rust [%]</th>
<th>mAP [%]</th>
<th>Size of the Weight Files [MB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>YOLOv8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YOLOv8+SE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YOLOv8+CBAM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA-YOLOv8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4**
To better evaluate the performance of detection models, the results of different models were visualized, as shown in Fig. 6. Fig. 6 (a) shows the detection results of single corn gray spot disease leaf under complex background. It can be seen that the Faster R-CNN model failed to detect gray spot disease leaf, while the YOLOv5s, YOLOv7, YOLOv8, and CA-YOLOv8 models were able to correctly detect gray spot disease leaf. The CA-YOLOv8 model has the highest confidence level. Fig. 6 (b) shows the detection results of multiple corn leaf blight diseases. It can be seen that the Faster R-CNN model has false positives and omissions, the YOLOv5s model has false positives, the YOLOv8 model has false positives, and both the YOLOv7 and CA-YOLOv8 models can correctly detect multiple corn leaf blight diseases. However, the confidence of the CA-YOLOv8 model is better than that of YOLOv7. Fig. 6 (c) shows the detection results of single corn rust leaf under simple background. It can be seen that the Faster R-CNN model failed to detect corn rust leaves, while the YOLOv5s, YOLOv7, YOLOv8, and CA-YOLOv8 models were able to correctly detect corn rust leaves. The CA-YOLOv8 model has the highest confidence level. In summary, the CA-YOLOv8 model can accurately detect three types of corn leaf diseases under different backgrounds, and has the highest confidence compared to other models.
CONCLUSIONS

(1) A CA-YOLOv8 corn leaf disease detection model was proposed for the detection of corn gray spot, corn leaf blight, and corn rust. The CA attention mechanism was added to the feature map output from the Neck part of the network, which helps the model to better locate and identify the target, improve the detection accuracy of the model, and hardly increase the memory occupation of the model.

(2) In order to verify the performance of the CA-YOLOv8 model, four groups of experiments were set up for comparative analysis. The four groups of experiments were YOLOv8, YOLOv8+SE, and YOLOv8+CBAM, and the CA-YOLOv8 model proposed in this paper. The experimental results showed that the mAP of the CA-YOLOv8 model was the highest.

(3) Under the same experimental conditions, compared with Faster R-CNN, YOLOv5s, YOLOv7, and YOLOv8 models, the CA-YOLOv8 model achieved better results on the corn leaf disease dataset, with mAP improvements of 63.53, 29.24, 3.21, and 2.15 percentage points, respectively.

The experiment fully proves that the CA-YOLOv8 model not only improves the evaluation indicators, but also achieves good visual effects. This study can provide technical reference for the development of a portable intelligent corn leaf disease detection system.

ACKNOWLEDGEMENT

This research, titled “Detection of corn leaf diseases based on CA-YOLOv8”, was funded by the Basic Research Program of Shanxi Province (Research on wheat grain appearance quality detection based on computer vision).

REFERENCES


