Advances in Computer, Signals and Systems (2023)
Clausius Scientific Press, Canada
DOI: 10.23977/acss.2023.070301
ISSN 2371-8838 Vol. 7 Num. 3

Research on an Improved YOLOV8 Image Segmentation Model for Crop Pests

Jichang Kang, Lianjun Zhao, Kangtao Wang, Kunpeng Zhang

Shandong University of Technology, Zibo, Shangdong, 255000 China

Keywords: YOLOV8, Simulated annealing algorithm, Crop pests, Image segmentation

Abstract: With the change of ecosystem, there are more and more kinds of crop diseases and insect pests, and the harm is becoming more and more serious. Preventing crop diseases and insect pests is the premise to ensure crop yield. Image segmentation technology is to divide a number of specific targets and regions with different characteristics in the image according to the requirements through pixel-level classification scheme, which is the first important link of image analysis. In this article, Simulated Annealing (SA) algorithm is used to optimize YOLOV8. The main purpose is to randomly find the optimal solution of the loss function in the last layer of convolutional neural network (CNN) with SA algorithm, and then update the weights and offsets of the previous layer with this solution. The CNN structure also uses the dropout regularization method to effectively reduce the influence of over-fitting. The simulation results show that compared with YOLOV7 algorithm, the average accuracy of disease identification of improved YOLOV8 is obviously higher. The pest identification model based on the improved YOLOV8 algorithm has more advantages than YOLOV7 algorithm in both accuracy and efficiency. The proposed method achieves the best detection performance on large-scale public data sets, and also performs well in the task of crop pest detection studied in this article.

1. Introduction

Although the development of agriculture has reached an unprecedented height at present, it is inevitable that China's agriculture, which is in the process of modernization transformation, has also encountered some problems that restrict its development. Among these problems, the most important one is the increasingly frequent crop pests [1]. With the change of ecosystem, there are more and more kinds of crop diseases and insect pests, and the harm is becoming more and more serious. Preventing crop diseases and insect pests is the premise to ensure crop yield. Deep learning is a machine learning method, which is realized by deep data fitting with neural network. The learning process is multi-layered and includes many nonlinear changing processes [2]. The traditional identification method is inefficient and the identification effect is not good, mainly through some simple machine learning models to judge the types of crop diseases in the actual environment. Although this method is simple and time-consuming, it does not achieve the ideal identification effect for crop diseases in actual natural conditions [3]. Through a certain combination, the original fuzzy features can be more abstracted, and the complex data relations can be more intuitively simulated, so that the original data relations that are not tight enough become closer [4]. With the help of deep learning, this advantage is more prominent. With the development of neural network and computer vision, image semantics are constantly improving, and deep learning has become the development trend in the field of artificial intelligence.

Image segmentation technology is to divide a number of specific targets and regions with different characteristics in the image according to the requirements through pixel-level classification scheme, which is the first important link of image analysis [5]. Gwo-Jiun et al. put forward a disease location diagnosis algorithm with wheat as the research object. After wavelet decomposition transformation and grain matrix calculation, the binary image of crop disease area was obtained by automatic threshold algorithm, so as to calculate the color feature value and compare it with the image in the disease database, and make a disease detection judgment [6]. Wang et al. preprocessed the corn disease image first, then constructed CNN by using the Tripletloss function to extract and learn the feature information of the disease image, then extracted the texture features of the disease image by using the SIFT algorithm, and finally classified the image by using the Softmax function [7]. At present, the automatic identification and diagnosis method of crop diseases based on machine vision has become the mainstream technology. However, in the application of related technologies in the actual production of agricultural field crops, there are still many problems in the effective and high-speed segmentation and acquisition of disease spots in crop disease images. In this article, an improved YOLOV8 based on SA algorithm is proposed for the segmentation and extraction of disease spots in crop disease images under the complicated background of agricultural field production. The SA algorithm is used to randomly find the optimal solution of the loss function in the last layer of CNN, and then the weight and bias of the previous layer are updated by this solution.

2. Disease Image Segmentation Based on Improved YOLOV8

The traditional dropout method forces a neural unit to work with other randomly selected neural units to achieve good results. It weakens the joint adaptability between units, thus enhancing the generalization performance of the structure [8]. In the standard neural network, the correlation between nodes enables them to cooperate to trim the noise in other nodes, but these cooperations cannot be generalized, so the over-fitting problem arises, and dropout destroys this correlation [9]. In this article, the pool layer and full connection layer in CNN structure use dropout regularization to suppress some activation units. Using dropout in max-pooling layer can avoid the above shortcomings, because dropout may suppress the values of some neurons, so the values obtained by pooling are random, instead of always being the average or the maximum of all values in the pooled area. The images of crop diseases collected under actual natural conditions are greatly influenced by illumination, because illumination is a factor that has to be considered in the external natural environment.

Border regression is used to adjust the regional suggestion box to make it closer to the real border box, which will make the positioning more accurate. The measured image will generate two boxes, one is to mark the true value, and the other is the target detection box generated by the target detection algorithm. According to the overlapping ratio of the target detection frame and the real frame, it is compared with the set threshold parameters. The detection window of diseased leaves is shown in Figure 1.

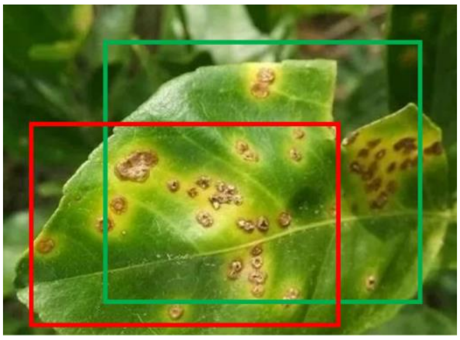

Fig.1 Detection Window of Diseased Leaves

It is known that the thermal point set of disease image is P, P_{iV} is the thermal value information of thermal point P_i, and P_i is calculated by correlation coefficient function. Calculating the thermal values of four vertices of the disease image;

$$V_i = \sum_{j=1}^{n} \frac{|Q_i P_j|^2}{\sum_{k=1}^{n} |Q_i P_k|^2} \cdot P_{iV} \quad (1)$$

Where $|Q_i P_j|^2$ is the square of the distance between vertex Q_i and thermal point P_i.

Dropout will inhibit the values of some nerve cells, and the values obtained by pooling will be random, which is no longer always the average of all values in the pooled area or the maximum value [10]. In the training process, the input after each dropout is different, and more models can be trained at the same time and with the same data. The main process of SA algorithm is to solve a large range of combinations and randomly search the global optimal solution, which has the characteristics of asymptotic convergence, parallelism, flexibility, easy implementation and fast operation speed [11]. This algorithm is the process of searching for the optimal solution, that is, the process of SA, in the process of reasonably adjusting the temperature drop. SA can not only find the optimal solution, but also operate quickly.

SA is a heuristic method to find the optimal solution randomly. Because of adding appropriate uncertain elements to the query, it will not only accept the good solution, but also receive the solution which is actually worse than the current solution according to a probability value, and this value will gradually decrease with the passage of time. This operation makes the value in the neighborhood of the solution uncertain, so this method may avoid the interval optimal solution and finally get the global optimal solution. SA is used to train CNN to get the optimal solution that meets the conditions. The objective function is:

$$Q = -\frac{1}{N} \sum_{n=1}^{N} (y^n \, log(o^n) + (1 - y^n) \, log(1 - o^n)) \quad (2)$$

Where y represents the expected output, o represents the actual output, and N represents the number of samples in one training. There are two ways to terminate the training process: one is to reach the set maximum number of iterations; Second, when the loss function is less than a certain constant. Satisfying any criterion means reaching the optimal state.

Different light intensities on crop disease images have a great influence on feature extraction of disease spots, so in order to make CNN have strong adaptability, namely generalization ability, the crop disease images collected under actual natural conditions are processed and adjusted in two aspects: image brightness and contrast [12]. Through this illumination transformation, the data volume is enriched, so that CNN can obtain diversified images of the same disease and improve the effect of subsequent image segmentation and recognition. Only when the feature intensity of a certain region in the image reaches a specified value can the convolution kernel extract features in this region, and the training of feature extraction methods will not be affected by other regions. In reality, data is usually distributed nonlinearly, and the activation layer enables CNN to learn nonlinear mapping. In an image, the relative position between features is more important than the specific position of a specific feature in the whole image. Pool layer makes the data space smaller, which indirectly prevents the occurrence of over-fitting phenomenon. For CNN, the last part is the full connection layer. The fully connected layer includes three parts: input layer, hidden layer and output layer. The output layer is usually the softmax layer. The output result of the full connection layer is a probability value, which indicates the possibility that the input of CNN belongs to each category. The CNN model of crop pest image segmentation is shown in Figure 2.

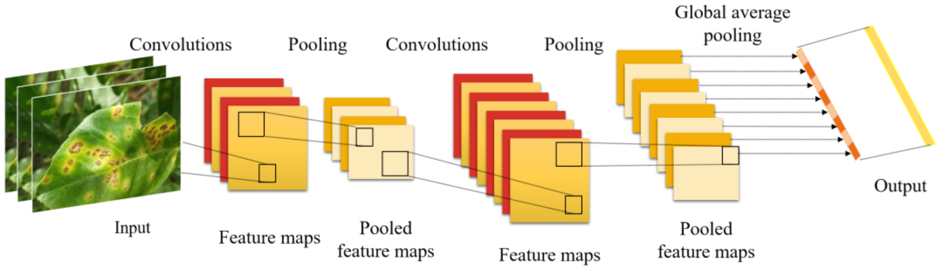

Fig.2 Cnn Model of Crop Pest Image Segmentation

In standard CNN, one of the reasons for the poor performance of small target detection comes from the receptive field with limited convolution operation. Therefore, the spatial attention module aims to make the network pay more attention to the position of small target pests by learning the target weight of the global feature map in each spatial position. Contrary to the channel attention module, the spatial attention module needs to eliminate the influence of channel information. Because the image contains the labeling information of pest targets in spatial position, the spatial attention module is trained through supervised learning. The convolution neural network function is defined as:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l\right) \quad (3)$$

Where x_i represents the input characteristic map, k represents the convolution kernel, b represents the deviation term, and the convolution output is the characteristic map x_j. Assume that the convolution layer convolves the input insect pest image with k filters, and generates k new feature maps for subsequent processing. If the output feature map is represented in a layer, then:

$$F_j^{(n)} = \sum_i w_{ij}^{(n)} * F_i^{(n-1)} + b_j^{(n)} \quad (4)$$

Where: $*$ is a two-dimensional convolution; $w_{ij}^{(n)}$ and $b_j^{(n)}$ are convolution filters and deviations, respectively; $F_j^{(n)}$ is the j output characteristic map at the n layer. The formula of active layer after convolution is as follows:

$$F_j^{(n+1)} = f\left(F_j^n\right) \quad (5)$$

Where: f is a point-by-point activation function. Convert each data item x_i in the small batch $B = \{x_1, x_2, x_3, \ldots, x_m\}$ with size m to y_i:

$$y_i = \gamma \hat{x}_i + \beta \quad (6)$$

$$\hat{x}_i = \frac{x_i - E_M(x_i)}{\sqrt{Var_M(x_i) + \varepsilon}} \quad (7)$$

Where: $E_M(x_i)$ and $Var_M(x_i)$ are the mean and variance.

When the same crop disease is collected at different time points in a day, the appearance of the disease images is different, even sometimes very different, the biggest difference is reflected in the color and brightness of the images. Dropout used in training is different every time, and multiple models are trained at the same time and with the same data. In the test stage, a new model average method is proposed to pool the probability of unit value and the influence of P value in the region to solve the average value predicted by the model.

3. Result Analysis and Discussion

Before the target detection task, image processing technology can be used to preprocess the image of crop disease leaves. Image preprocessing technology mainly weakens or even removes the interference information in the picture through mathematical operation and image transformation, strengthens the expression of important information, and makes the subsequent image feature extraction more rapid and accurate, and enhances the detectability. The essence of pest detection task is the subtask of general target detection. Aiming at AgriPest pest detection data set, this article adopts the evaluation index of general target detection as the basis. In this article, the deep learning pretreatment of crop disease images collected under actual natural conditions can increase the number of data sets, for example, an image will generate at least two more images after illumination transformation and normalization, which is reflected in the data set construction of this article. On the other hand, the deep learning network can better extract the lesion features in crop disease images during training, so that CNN with superior performance can be trained and the ability of segmentation and recognition of crop diseases can be improved.

The leaves of crop diseases usually have the characteristics of uneven illumination and different leaf sizes, so the selection of feature extraction, image segmentation and recognition methods is very important, which directly affects the detection speed and accuracy of later image processing. Compared with the high-level statistical features, the image model of crop pests constructed by using the bottom features can get higher accuracy. The time-consuming of pest image segmentation using different methods is compared and analyzed, as shown in Figure 3.

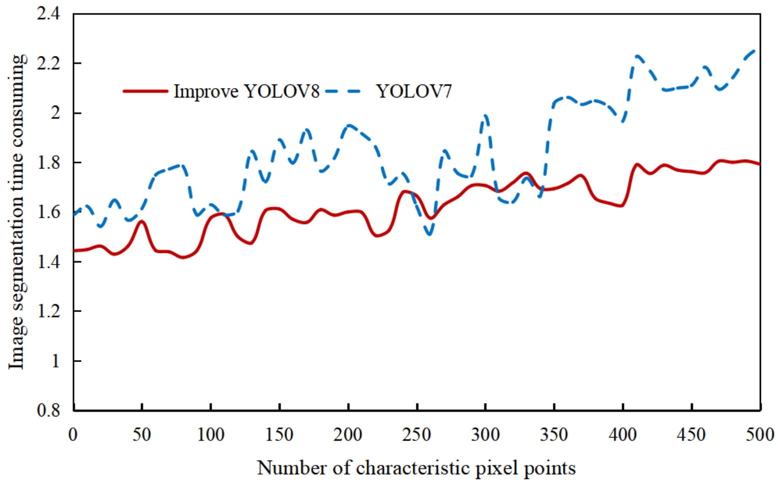

Fig.3 Time-Consuming Segmentation of Pest Images by Different Methods

As can be seen from Figure 3, the time-consuming of pest image segmentation processing based on YOLOV7 algorithm increases with the increase of the number of pixel points of feature information, which takes a long time. However, the time-consuming of pest image segmentation based on improved YOLOV8 has an upward trend, and it has obvious advantages compared with YOLOV7 algorithm. In computer operation, the time required for each matching is the same, so reducing the number of matching times can reduce the time required for image recognition and make the recognition results display faster.

The image feature information with differences is fused, so that the fused feature information can be easily distinguished, and the expressive ability of the image is improved. The fused features are washed by channels, so that each region of the feature map contains the features of different channels, and the obtained image fusion information is more robust. 9610264 takes the average accuracy of crop pest image segmentation as the test index, and selects YOLOV7 algorithm as the comparison object. The experimental results are shown in Table 1 and Table 2.

Table 1 Average Accuracy of Pest Image Segmentation Based on Improved YOLOV8 Algorithm

Sample size	Accuracy (%)
15	99.74
30	99.45
45	98.76
60	98.15
75	97.59
90	96.89
105	95.74

Table 2 Average Accuracy of Pest Image Segmentation Based on Yolov7 Algorithm

Sample size	Accuracy (%)
15	97.75
30	96.69
45	95.48
60	94.37
75	92.77
90	91.71
105	90.52

From the experimental data, it can be seen that when the number of test samples begins to increase, the average accuracy of disease identification of the two methods has a certain downward trend. However, compared with YOLOV7 algorithm, the average accuracy of disease identification of the improved YOLOV8 algorithm is obviously higher.

The classification effect of CNN on crop leaf images depends on the size of convolution kernel. The larger the convolution kernel, the wider the receptive field of CNN, which is more convenient for analyzing the global information of crop leaf images, but at the same time, the detailed features of the images will be ignored. The smaller the convolution kernel is, the easier it is to obtain the detailed features of the image, but the global information of the crop leaf image cannot be obtained. Therefore, in order to obtain accurate feature information of images, multi-convolution kernels are needed. The result of precision test using YOLOV7 algorithm is shown in Figure 4. The result of precision test using the improved YOLOV8 algorithm is shown in Figure 5.

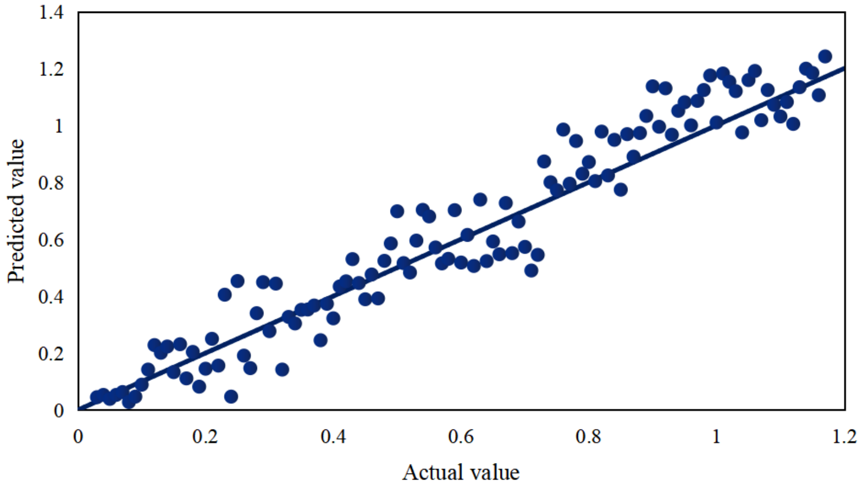

Fig.4 Accuracy Test Results of YOLOV7 Algorithm

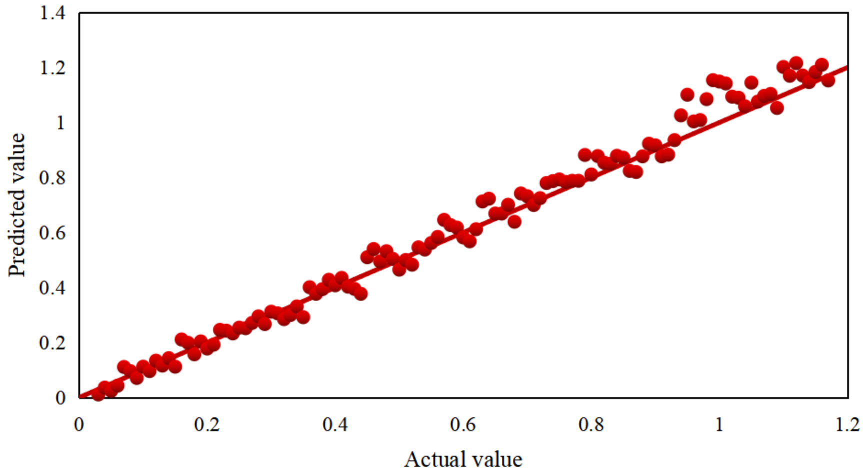

Fig.5 Accuracy Test Results of Improved YOLOV8 Algorithm

The points on the graph represent the ratio of the predicted value to the actual value. The closer the predicted value is to the actual value, the closer the square point is to the straight line y=x. On the contrary, the greater the difference between the predicted value and the actual value, the farther the square point deviates from the straight line y=x. The comprehensive experimental results show that the pest identification model based on the improved YOLOV8 algorithm is superior to YOLOV7 algorithm in both accuracy and efficiency.

4. Conclusions

At present, the automatic identification and diagnosis method of crop diseases based on machine vision has become the mainstream technology. However, in the application of related technologies in the actual production of agricultural field crops, there are still many problems in the effective and high-speed segmentation and acquisition of disease spots in crop disease images. Convolutional network model has the characteristics of partial receptive field, hierarchical structure, combined extraction process and classification operation, and has many achievements in image processing and recognition research. In the traditional network, the learning rate is a global constant. Choosing a large learning rate is not conducive to getting the minimum value of the loss function, while a small learning rate will consume a lot of training time. In this article, an improved YOLOV8 based on SA algorithm is proposed for the segmentation and extraction of disease spots in crop disease images under the complicated background of agricultural field production. The comprehensive experimental results show that the pest identification model based on the improved YOLOV8 algorithm is superior to YOLOV7 algorithm in both accuracy and efficiency. This method is effective and more practical, and it can quickly converge and gradually reach the best, so it has better feasibility and effectiveness. The CNN based on adaptive learning rate algorithm proposed in this article is compared with the traditional CNN. In the future research, more adaptive learning rate algorithms can be studied, and compared with the methods proposed in this article, we can find the differences of various methods and get a more efficient model.

References

[1] Jiang X, Zhen J, Miao J, et al. Assessing mangrove leaf traits under different pest and disease severity with hyperspectral imaging spectroscopy. Ecological Indicators, vol. 129, no. 2, pp. 107901, 202.

[2] MN Rodríguez-García, F García-Sánchez, R Valencia-García. Knowledge-Based System for Crop Pests and Diseases Recognition. Electronics, vol. 10, no. 8, pp. 905, 2021.

[3] Shajahan S, Sivarajan S, Maharlooei M, et al. Identification and Counting of Soybean Aphids from Digital Images Using Shape Classification. Transactions of the ASABE (American Society of Agricultural and Biological Engineers), vol. 60, no. 5, pp. 1467-1477, 2017.

[4] Zhang M, Zhang W, Liang X, et al. Detection of fatigue crack propagation through damage characteristic FWHM using FBG sensors. Sensor Review, vol. 40, no. 6, pp. 665-673, 2020.

[5] Kim J Y, Bellotti A, Alapati P, et al. Use of a non-collinear wave mixing technique to image internal microscale damage in concrete. Journal of Applied Physics, no. 14, pp. 131, 2022.

[6] Gwo-Jiun, Horng, Min-Xiang, et al. The Smart Image Recognition Mechanism for Crop Harvesting System in Intelligent Agriculture. IEEE Sensors Journal, vol. 20, no. 5, pp. 2766-2781, 2019.

[7] Wang B. Identification of Crop Diseases and Insect Pests Based on Deep Learning. Scientific Programming, vol. 2022, pp. 1-10, 2022.

[8] Raja R, Slaughter D C, Fennimore S A, et al. Crop signalling: A novel crop recognition technique for robotic weed control ScienceDirect. Biosystems Engineering, vol. 187, pp. 278-291, 2019.

[9] Zou W, Shen C, Yin G. Application of image recognition technology in agricultural production process. International Agricultural Engineering Journal, vol. 27, no. 2, pp. 318-326, 2018.

[10] Yu D, Zhang B, Zhao C, et al. Scene classification of remote sensing image using ensemble convolutional neural network. Journal of Remote Sensing, vol. 24, no. 6, pp. 717-727, 2020.

[11] Wang D, Mao K. Task-generic semantic convolutional neural network for web text-aided image classification. Neurocomputing, vol. 329, no. 3, pp. 103-115, 2019.

[12] Zhao M, Hu C, Wei F, et al. Real-Time Underwater Image Recognition with FPGA Embedded System for Convolutional Neural Network. Sensors, vol. 19, no. 2, pp. 350, 2019.