Scholar Publishing Group
Academic Journal of Agricultural Sciences
https://doi.org/10.38007/AJAS.2023.040106
ISSN 2789-8415 Vol. 4, Issue 1: 68-80

Monitoring System of Deep Learning Video Image Analysis Technology Based on Smart Agriculture

Bixuan Wan and Jing Si*

Nantong Institute of Technology, Nantong, China

**corresponding author*

Keywords: Video Images, Convolutional Networks, Immersive Learning, Disease Monitoring, Smart Agriculture

Abstract: Massive image data was generated under the background of big data. Intelligent analysis, processing, and identification of the image information of the agricultural IoT terminal can intuitively and vividly express all aspects of crop growth, development, health status, damage degree, etiology, etc. The system can respond similarly to intelligent living beings. The current image recognition technology for crops is relatively backward in terms of image recognition accuracy and time. In this case, among the millions of data at every turn, the data processing capacity is far from enough, so this paper proposes based on Smart agriculture uses deep learning video image analysis technology to monitor system research, and through the research and improvement of the deep learning video image analysis technology monitoring system, it can be better applied to smart agriculture. This model is applied to the corn disease pictures collected from farmland to accurately identify corn diseases. The experimental results show that the overall recognition accuracy of major corn diseases (maize maize leaf spot, small leaf spot, gray spot disease, smut, and black powdery mildew) using the improved convolution NNs optimization algorithm reaches 93.2% Compared with a single convolutional NNs, the accuracy is improved by a quarter, and the processing time of each picture is shortened by about one-tenth that of the traditional NNs. The accuracy and training speed of this algorithm are significantly improved compared with traditional convolutional NNss.

1. Introduction

With the rapid development of agricultural informatization, crop image information has become the main body of agricultural big data. Agriculture is a complex living system with typical ecological regionality and complexity of physiological processes. China is a large agricultural country and has huge demand for agricultural information technology and science. At present, mobile agricultural robots are used for crop monitoring in some areas of China. Mobile agricultural robots are also the main way to obtain agricultural image information. The agricultural robot is essentially an intelligent agricultural machine. Its appearance and application have changed the traditional way of agricultural labor, changed the situation of fixed-point video surveillance, achieved "patrol" of agricultural information, and was able to capture more accurate and multi-angle agricultural image information.

Therefore, with the widespread application of agricultural intelligent equipment and sensors, the Internet of Things, a large amount of valuable agricultural image data and agricultural condition information can be collected and stored. How to process these data, especially image data, and find and extract novel agricultural knowledge models from it , Become a key measure to discover the benefits of the project and promote the development of agricultural productivity [1]. Compared to the massive accumulation of agricultural data, the basic technical reserves of machine learning in the industry are seriously insufficient, and existing processing technologies in the agricultural field cannot meet the needs of such large-scale instant analysis and mining of information. How to carry out data processing and learning, and tap valuable agricultural production knowledge to effectively serve smart agriculture has become a prominent scientific and technological issue in the development of modern agriculture. Diseases will cause severe reduction in production. Therefore, diagnosis and treatment of corn diseases is an important part of corn production.

The technology proposed by Cao L et al [2] determines the best images of corn diseases by improving the clarity and quality of the images, and analyzes and recognizes these best images with the help of an improved simulation platform. Ramcharan A et al. [3] trained on the video images of the wood comb disease and got the experimental results to facilitate their deployment. Kumar S et al. [4] carried out a comparative analysis of common image processing techniques for detecting plant diseases, and the results show that image processing techniques based on deep learning have obvious advantages for detecting plant diseases. S.W. Zhang et al. [5] proposed a plant disease recognition method based on plant leaf images. Ma J et al. [6] proposed a greenhouse cucumber disease recognition system based on convolutional NNss based on deep learning and image processing.

The algorithm first adds data through data augmentation methods to improve the generalization and accuracy of the model. Transfer learning convolutional NNs model, introduce the training method of the model, extract disease picture features, speed up the training process of the convolutional NNs, reduce the degree of overfitting of the network; finally, apply the model to corn diseases collected from farmland Pictures for precise identification of corn diseases. The experimental results show that the overall recognition accuracy of major corn diseases (maize maize leaf spot, small leaf spot, gray spot disease, smut, and black powdery mildew) using the improved convolutional NNs optimization algorithm reaches 93.2%. Compared with a single convolutional NNs, the accuracy is improved by nearly a quarter, and the processing time of each picture is shortened to one tenth of that of a traditional NNs. The accuracy and training speed of this algorithm are significantly improved compared with traditional convolutional NNss, and it provides a new method for identifying plant diseases such as corn.

2. Proposed Method

2.2. Basic Principles of Deep Learning Algorithms

(1) NNs

The structure of NNs is very complicated, and the process of transmission is mainly to transmit signals by imitating human nerves. For his basic structure as shown in Figure 1:

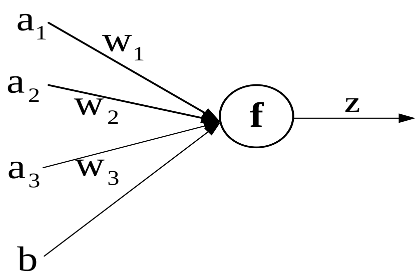

Figure 1. The basic structure of neurons in NNss

For NNs, one of the problems that we must consider in the actual design is its robustness. Only when the robustness problem is solved, its nonlinearity will not be obvious. And the image extraction effect of NNs will be even better.

Figure 2 shows the structure of a two-layer NNs.

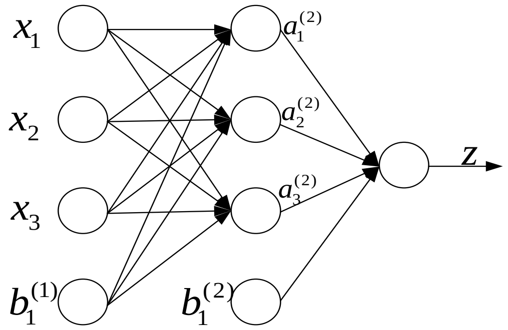

Figure 2. The basic structure of a two-layer NNs

x1, x2, x3 are the inputs of the NNs, z is the output, w is the weight value, b is the deviation of each layer.

$$\mathrm{z} = \mathrm{f}(\mathrm{a}_1 * \mathrm{w}_1 + \mathrm{a}_2 * \mathrm{w}_2 + \mathrm{a}_3 * \mathrm{w}_3 + \mathrm{b}) \tag{1}$$

Assume that the bias of the input layer is $b_1^{(1)}$, the bias of the hidden layer is $b_1^{(2)}$, and each neuron $a_i^{(2)}$ can be expressed as:

$$a_i^{(2)} = f(w_{i1}^1 x_1 + w_{i2}^1 x_2 + w_{i3}^1 x_3 + b_1^{(1)}) \tag{2}$$

$$\mathrm{z} = \mathrm{f}(w_{11}^{(2)} a_1^{(2)} + w_{12}^{(2)} a_2^{(2)} + w_{13}^{(2)} a_3^{(2)} + b_1^{(2)}) \tag{3}$$

which is:

$$\mathrm{z}_{ik} = f(\sum w_{ij}^{(k-1)} a_j^{(k-1)} + b_j^{(k-1)}) \tag{4}$$

1) Convolution layer

A simple convolution process is as follows: Suppose there is an image with a size of 5 ×5, and a 3 ×3 filter is used for convolution. The final result is a 3 ×3 Feature Map, as shown in Figure 4:

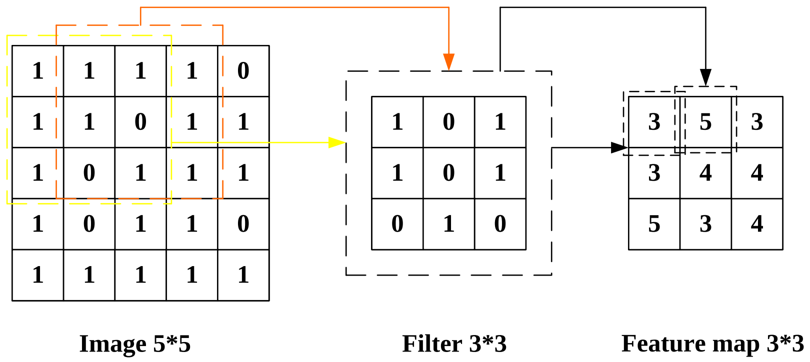

Figure 3. Convolution layer extraction feature process

According to formula (4), we assume the deviation b = 0 and the stride = 1, that is, the sliding window moves one pixel at a time. As can be seen from the above figure 3, neurons in the same layer can share a convolution kernel. Therefore, when it comes to the operation of the convolutional layer, there is no need to manually select features. The number of convolution kernels and the sliding step can let it automatically train, saving time and effort. Another feature of convolutional NNss is local perception.

In addition, Leaky ReLU, ELU, etc. are often used in different types of NNss. The commonly used centralized activation functions are shown in Table 1:

Table 1. Several commonly used activation functions

Activation function	formula	shortcoming	advantage
sigmoid	$\sigma(x)=1/\left(1+e^{-x}\right)$	Gradient dispersion	-
tanh	$\tanh(x)=2\sigma(2x)-1$	Gradient dispersion is not resolved	Solved the origin symmetry
relu	$f(x)=\max(0,x)$	Gradient dispersion is not completely resolved	Fast convergence

The Sigmoid activation function is shown in formula (5):

$$\sigma(\mathrm{x}) = \mathrm{sigm}(\mathrm{x}) = \frac{1}{1+e^{-x}} \tag{5}$$

Using the Sigmoid activation function has a large amount of calculation and the saturation region changes slowly. The hyperbolic tangent function Tanh is shown in formula (6):

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{6}$$

The linear correction unit ReLU activation function is shown in formula (7):

$$\mathrm{ReLU}(\mathrm{x}) = \max(0,\mathrm{x}) = \begin{cases} x, x \geq 0 \\ 0, x < 0 \end{cases} \tag{7}$$

The calculation formula for the Leaky ReLU activation function of the leakage correction linear unit is shown in (8):

$$\mathrm{LReLU}(\mathrm{x}) = \begin{cases} 1, x \geq 0 \\ ax, x < 0, 0 < a < 1 \end{cases} \tag{8}$$

The exponential linear unit ELU activation function is shown in the following formula (9):

$$\mathrm{ELU(x)} = \begin{cases} 1, x \geq 0 \\ a(e^x - 1), x < 0, 0 < a < 1 \end{cases} \tag{9}$$

Therefore, in actual applications, it has not been fully proved that using ELU activation functions is always better than ReLU activation functions.

2) Pooling layer

Figure 4 shows the extraction feature of the pooling layer. In this paper, the maximum pooling is used, that is, the maximum value of all numbers is taken within the selected sliding frame range.

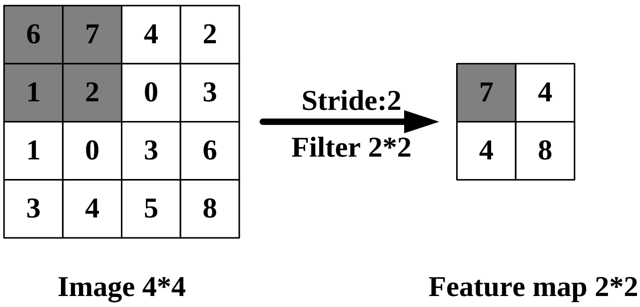

Figure 4. Feature extraction from pooling layer

If the input unit size of the pooling layer in the figure above is not an integer multiple of two, or the input and output sizes are expected to remain unchanged, the zero-padding method is generally used to make up multiples of two, and then pooled. Zero padding is to add zeros around the image to obtain the image size needed for the experiment.

3) Fully connected layer

Before the Softmax classifier classifies, the task of extracting features is basically completed. You only need to transform the extracted features into an $n \times 1$ vector. You can also reduce the dimension of the vector through several fully connected layers to correspond to the category.

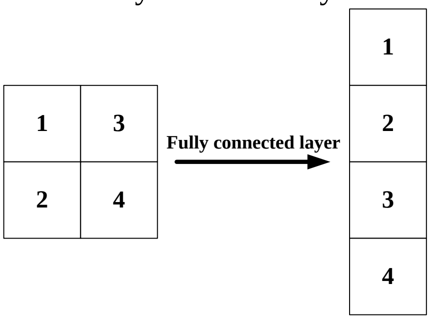

Figure 5. Working principle of the full connection layer

(2) Analysis of Convolutional NNs Model

Among them, the input of the network is a picture of size 32×32, and the convolution layer and pooling layer contain multiple two-dimensional feature maps, and these feature maps are composed of multiple units of warp warp elements. Each feature map is from the past. A feature extracted from a layer. The C1 layer in LeNet is a convolution operation layer. According to the principle of weight sharing, multiple neurons in each feature map of this layer simultaneously use a 5×5 filter to perform a convolution with a stride of 1 on the input layer. After operation, the size of the new feature map is 28×28. Because the number of filters is 6, each filter gets a feature map. Next is the S2 layer as the pooling layer, which is composed of 6 feature maps obtained from the convolution operation. The S2 layer performs a downsampling operation with a step size of 2 to obtain 6 feature maps with a size of 14×14. The convolution operation of the C3 layer is different from the C1 layer, which is obtained by convolving the feature maps in the S2 layer according to different combinations. Six filters are used in this layer, and the size of the convolution kernel is 5 × 5. After the convolution, the C3 layer obtains 16 feature maps with a size of 10 × 10. The S4 operation is similar to the S2 operation. C5 and F6 are two fully connected layer operations. They pull the feature vector into a one-dimensional vector and reduce the dimension. It can be seen from Fig. 7 that the network structure of the model is simple and the parameters are few. It can increase the training efficiency and prevent the occurrence of overfitting, which is beneficial for processing images with large data such as HSI.

2.2. Construction of Improved Convolutional NNs Model

(1) Improving ideas

Examples are images, sounds, and text. As a representative of deep learning technology, CNN has made rapid development in recent years. The structural composition of CNN has strong hierarchical correlation and spatial correlation to meet the high requirements in the field of image recognition and classification. The most convenient advantage is that it can autonomously extract the features of the image for training and learning. In order to solve the problems of long model convergence time and large sample similarity, the following two improvements were made on the basis of traditional CNN: 1) The training method of transfer learning was used to accelerate the training process of CNN and reduce the network's problem. The degree of fitting; 2) The method of data enhancement is used to enhance the generalization of the model. Its technical route is shown in Figure 6.

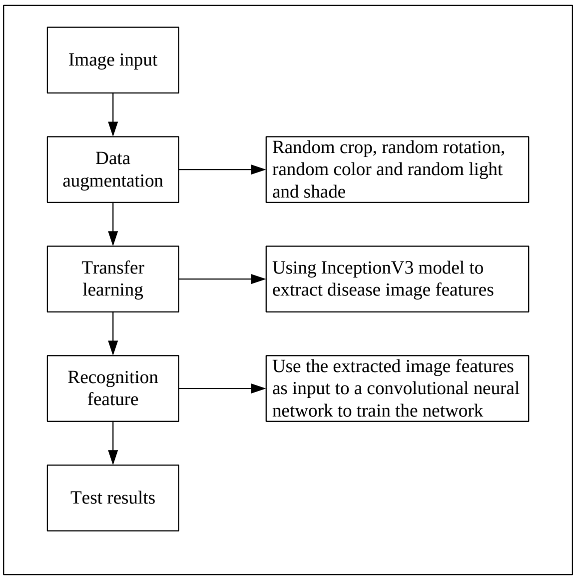

Figure 6. Improve dimage recognition process

The main process is as follows: the data is augmented by random cropping, random rotation, random color, and light and dark; the InceptionV3 model is used to maintain the parameters unchanged and extract the disease image features; and use the extracted image features as convolution Input the NNs, train the network, and finally get the recognition result.

(2) Data enhancement and transfer learning

CNN needs to optimize the generalization ability of deep learning models[9-10]. Because the image is high-dimensional, such as a 100 × 100 pixel image, it corresponds to a 10,000-dimensional high-dimensional space. In this study, four methods, random cropping, random rotation, random color, and light and shade were used to transform the data, and more training samples were obtained.

Transfer learning considers that most of the data and tasks are related. Models obtained from other large batches of data are used to train new data, thereby reducing training steps, shortening training time, and increasing the amount of data, which reduces the degree of overfitting[11-12]. In order to achieve a good classification effect on the target data set after training, traditional machine learning methods require the training set and test set to have the same feature space and data distribution. Problems such as local optimal solutions and overfitting. The transfer learning process is shown in Figure 7.

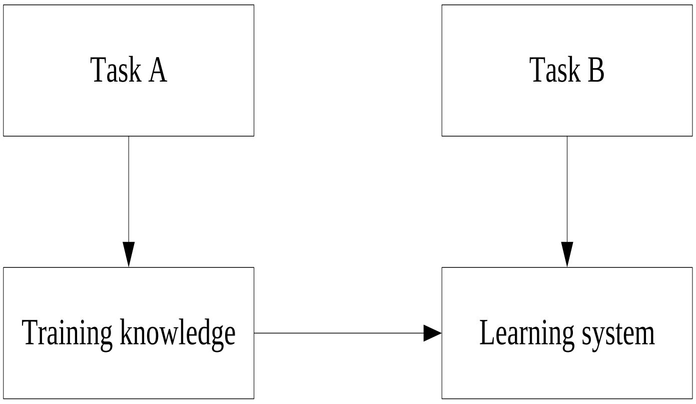

Figure 7. Transfer learning flow chart

The transfer learning process generally involves applying the things learned from the source domain (Task A in Figure 8) to the target domain (Task B in Figure 8).

The large-scale data set ImageNet and the trained convolutional NNs have strong generalization ability and can effectively classify and identify image data sets. The InceptionV3 model is trained through ImageNet, which contains information that can identify 1000 species. The structure of the InceptionV3 model is shown in Figure 8:

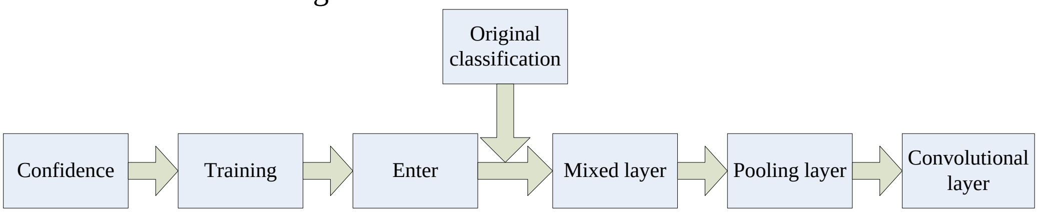

Figure 8. InceptionV3 model structure

In order to solve the problem of labeling data and training time, this study transfers the trained InceptionV3 model on ImageNet to the corn disease identification data set. The parameters of all convolutional layers in the trained In-ceptionV3 model are retained, and the trained NNs is used to extract the image features of corn disease plants. Inceptionmodel is used to process corn disease images, and its structure is shown in Figure 9. Gradually decompose the convolutional layer, and finally pass the pooling layer. Using this model can speed up calculations and reduce calculation costs.

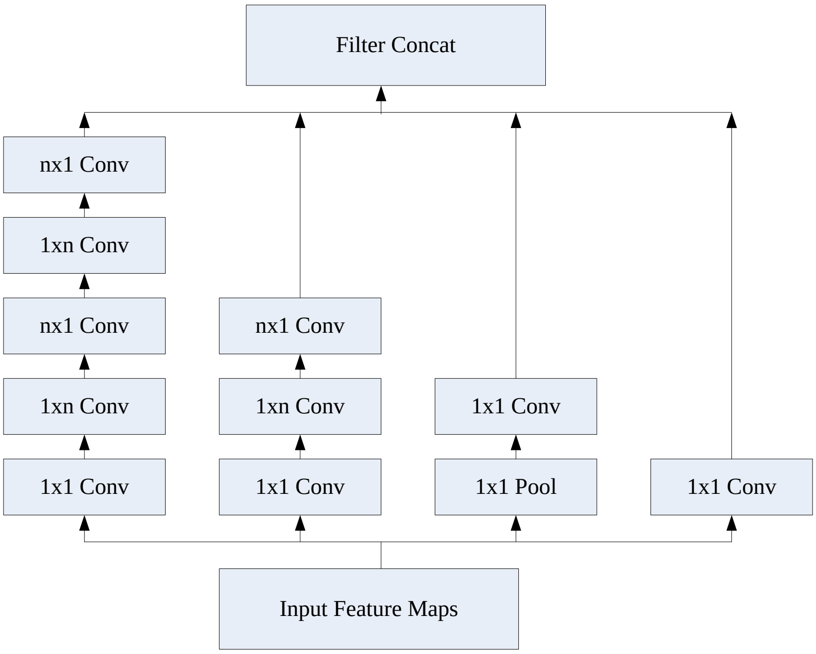

Figure 9. Inception model structure

3. Experiments

3.1. Subject

The data set of this experiment is five major diseases (maize maize leaf spot, small leaf spot, gray leaf spot, smut, and smut) on maize growing season caused by the NikonCOOLPIX5700 digital camera in the field environment. The image (pixel resolution: 2560 × 1920) is transferred to the computer in JPG format. The images were preprocessed and converted to a resolution of 640 x 480 pixels, for a total of 93 images. The collected disease symptoms were concentrated in corn leaves and ears, and the diseased parts were significantly different from normal leaves, that is, the normal leaves were green or dark green, while the lesions were yellow-brown or dark brown, and the shape of the lesions was round. Shape, ellipse, rectangle or shuttle. Figure 1 shows images of five major diseases of corn.The data set statistics are shown in Table 2:

Table 2. Statistics of the data set

symptom	Clear number	Unclear number
Leaf Spot	7	17
Gray leaf spot	9	16
Smut	11	14

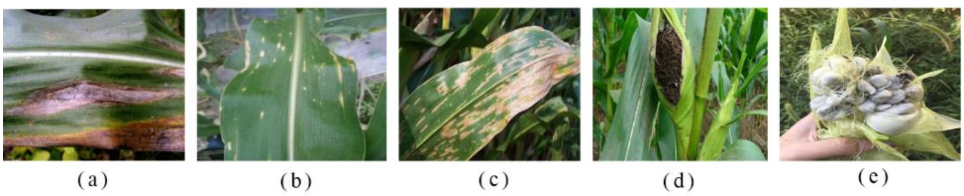

Figure 10. 5commoncorndiseases

The main symptoms of maize leaf spot: spindle-type lesions. The lesions are generally 5 to 10 cm in length, 1 to 2 cm in width, sometimes more than 20 cm in length, and more than 3 cm in width, as shown in Figure 10 (a). The main symptoms of maize small spot: the spot is oval and round, and the size is (5-10mm) × (3-4mm). The spot is often densely connected to form a large patch, as shown in Figure 10 (b). The main symptoms of gray spot disease in corn: rectangular lesions, the size of which is (2 ~ 4cm) × (1 ~ 6mm), as shown in Figure 10 (c). The main disease symptoms of corn smut: damage to the ears, forming a "grey bag", the diseased plants are generally short, the diseased ears are short, the base is large, and the apex is sharp, as shown in Figure 10 (d). The main symptoms of corneal powdery mildew: the initial appearance of the powdery mildew is white or off-white film, the fleshy interior is white, and the white tumor is white inside. Later, the tumor is filled with black powder, and it will disperse after the outer membrane is ruptured. , As shown in Figure 10 (e).

3.2. Experimental Environment and Experimental Parameters

The experimental software environment is a Windows 64-bit system, using the Tensorflow framework for training, and Python as the programming language. The computer memory is 16GB, equipped with IntelCoreTMi7-6700KCPU@4.00GHzx8 processor. The collected corn disease plant images were divided into three parts, of which training data accounted for 80%, verification data accounted for 10%, and test data accounted for 10%. This study uses the AlexNet structure, and solving the 7 × 7 convolution integral into two one-dimensional convolutions (1 × 7, 7 × 1), which can speed up model calculations and reduce 1 Deconvolution into two convolutions increases the network's depth and non-linear capabilities. The specific parameters are shown in Table 3:

Table 3. Computer related parameters

<table>
<tr><td colspan="4">Parameter configuration</td></tr>
<tr><td rowspan="4">computer</td><td>RAM</td><td>16GB</td><td>Graphics card</td></tr>
<tr><td>frame</td><td>Tensorflow</td><td rowspan="3">IntelCortMi7-6700KCPU@4.00GHzx8</td></tr>
<tr><td>Programming language</td><td>Python</td></tr>
<tr><td>Number of system bits</td><td>64</td></tr>
</table>

4. Discussion

4.1. Comparative Analysis of Image Recognition Accuracy

(1) Image enhancement processing

The test data collected images of five diseased plants, such as common maize leaf spot, small leaf spot, maize gray spot, smut, and black powdery mildew. Based on this principle, we use the semantic segmentation task in Python to enhance the transformation of the picture, and use the four methods of random cropping, random rotation, random color, and light and dark to enhance the data of the collected corn plant disease pictures, and the corn plant spot disease image Extensions have been made to improve the robustness of the model.

(2) Analysis of identification results of various diseases

Table 1 shows the comparison of the accuracy rates of the traditional model and the improved model for the five types of corn disease image recognition. Figure 11 shows this comparison visually.

Table 4. Comparison of the accuracy rates of five disease image recognition

	Large spot	Small spot	Maize gray spot	Smut	Powdery mildew
Original model	83.2%	81.3%	78.9%	76.8%	85.5%
Improved model	91.6%	92.6%	93.7%	88.8%	96.8%

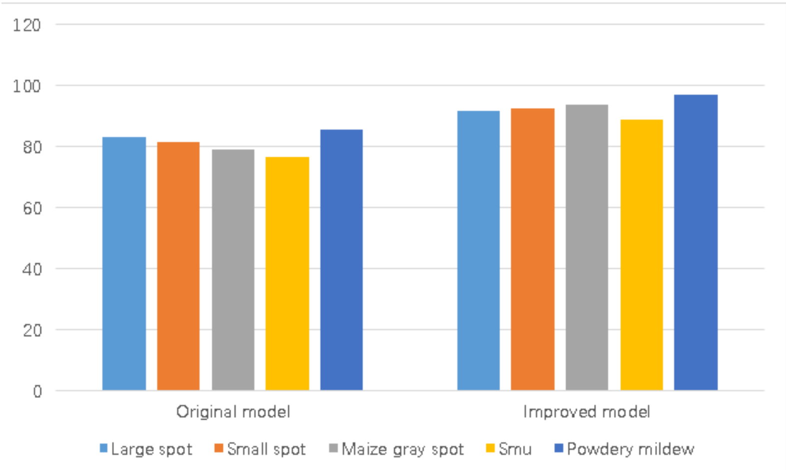

Figure 11. Comparison of the accuracy rates of five disease image recognition

The accuracy of recognition is shown in Table 5:

Table 5. Recognition accuracy rate

disease	Oridinal model	Improved model
Large spot	82	92
Small spot	81	93
Maize gray spot	79	94
smu	76	87
Powdery mildew	86	97

(2) Comparative analysis of overall accuracy and rate

This experiment compared the accuracy of the data before and after the enhancement, and conducted model training in two. Table 6 shows the original corn plant lesion image data set and the enhanced corn plant lesion image data set and compared their performance. It can be seen from Table 6 that the data volume of the corn plant disease image data set has increased from 90 to 4510, and the accuracy of the data has increased from 81.5% to 93.2%. The test results show that the data enhancement is important for the effect of corn plant image recognition Impact.

Table 6. Comparison of the accuracy of the test training set

	Dimension	Activation function	The amount of data	Accuracy(%)	Time consuming
Original model	7×7	Relu	90	81.5	2.72
Improved model	1×7,7×1	Relu	4510	93.2	0.26

In order to achieve more accurate results, this study will perform 4000 iterations of the two models respectively, and compare the accuracy of the original model of the improved optimal model. From Figure 12, it can be clearly seen that after 4,000 iterations of the training of the two models, the accuracy of the traditional CNN is stable after 3,800 iterations, stable at about 81.5%, and the accuracy of the improved CNN model is 1300 times It tends to be stable, and finally can reach about 93.2%, which proves that the improved model has a significant improvement in accuracy than the traditional CNN. Figure 13 shows the observed error rate as a function of the number of iterations based on a comparison of recognition performance experiments.

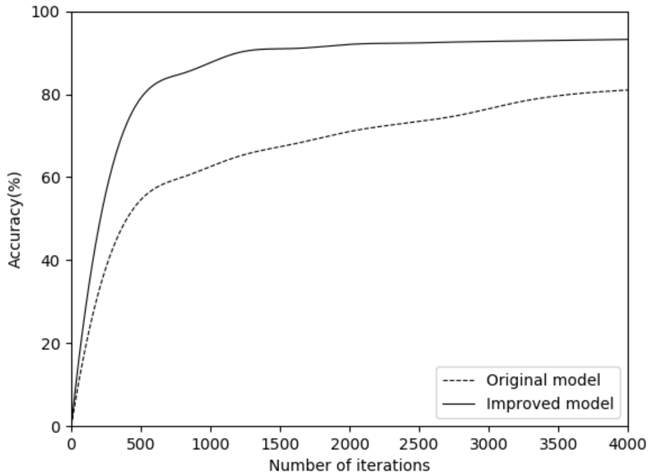

Figure 12. Plot of accuracy as iterations

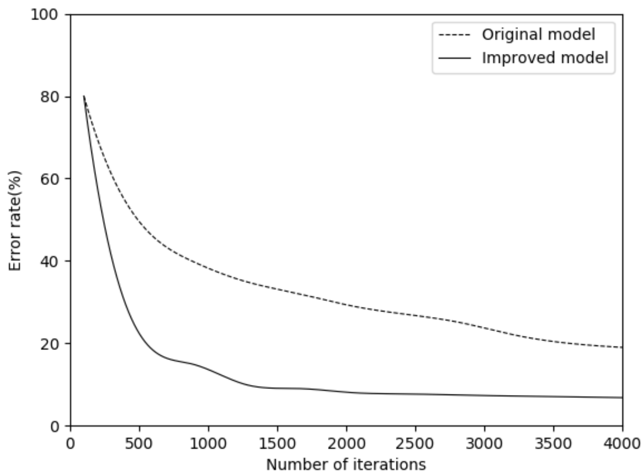

Figure 13. Variation of error rate with number of iterations

Through the method of transfer learning, first use the large public image data set on the Internet to pre-train the deep network, and then pre-train the weights on ImageNet to the corresponding part of the new network using the features extracted by CNN. Training on the network layer, stop training when the training performance reaches the best. In this paper, the Incep-tionV3 model trained on ImageNet is transferred to a corn plant disease identification data set. The parameters of all convolutional layers in the trained InceptionV3 model are retained. It can be seen that compared with traditional deep learning, the use of pre-trained parameter weights through transfer learning reduces training time and improves the efficiency of identifying maize plant diseases. Figure 14 shows the overall accuracy and rate comparison between the traditional model and the improved model.

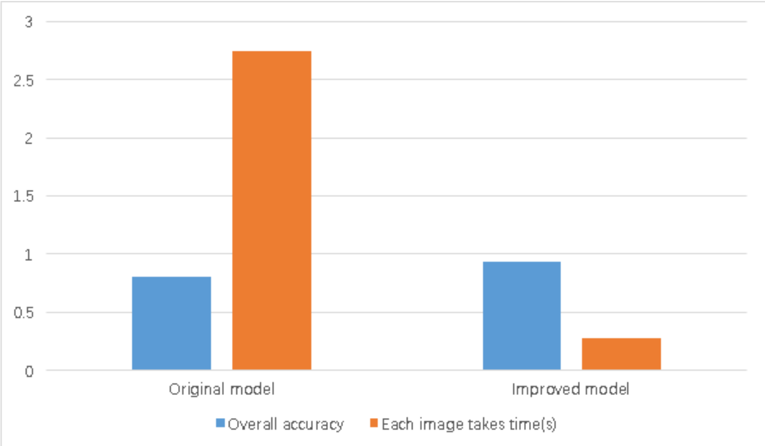

Figure 14. Overall accuracy and rate comparison

5. Conclusion

The research idea is that this article adopts an improved deep learning-based video image processing technology to increase the training data set through four data enhancement methods: random cropping, random rotation, random color, and light and dark, and uses the InceptionV3 model for training.

This model avoids the problems of traditional convolutional NNs classification that is easy to fall into the local optimal solution and overfitting due to the small amount of data. The feature pre-selection operation allows the network to automatically learn and identify the class features of the corn plant lesion image target. The improved model has an accuracy rate of 93.2% in image recognition of corn plant lesions, and the model recognition time is nearly 10 times shorter than that of the traditional convolutional NNs model, which meets the practical application of corn plant disease recognition and improves the robustness of the experiment. Sex.

In general, the improved convolutional NNs can be used not only in the recognition of digital pictures of plant diseases such as corn, but also in remote sensing images, especially in the recognition of plant diseases in hyperspectral images, and has broad application prospects in the agricultural field. In the integrated learning network, feature extraction and pattern recognition are combined, and a set of learning mechanisms are shared to complete the training of the recognition components and feature extraction components, solve the problem of "process and target" mismatch, and systematically simulate and verify Its effectiveness has a constructive application prospect in the field of intelligent and informatized agricultural production technology.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Baker M A , Robinson J S.(2016). "The Effects of Kolb's Experiential Learning Model on Successful Intelligence in Secondary Agriculture Students.", Journal of Agricultural Education, 3(2),pp.57-63.

[2] Cao L , Meng Y , Lu J , et al.(2015). "Research on Video Image Recognition Technology of Maize Disease Based on the Fusion of Genetic Algorithm and Simulink Platform.", Springer International Publishing,12(1),pp.147-155.

[3] Ramcharan A , Baranowski K , Mccloskey P , et al.(2017). "Using Transfer Learning for Image-Based Cassava Disease Detection.", Frontiers in Plant Science,13(9),pp.89-96.

[4] Kumar S , Kaur R .(2015) "Plant Disease Detection using Image Processing- A Review.", IJCA,124(2),pp.528-535.

[5] S.W. Zhang, Y.J. Shang, L. Wang.(2015). "Plant disease recognition based on plant leaf image.", Journal of Animal & Plant Sciences, 25(3),pp.42-45.

[6] Ma J , Du K , Zheng F , et al.(2018). "Disease Recognition System for Greenhouse Cucumbers Based on Deep Convolutional NNs.", Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, 34(12),pp.186-192.

[7] Hong, Seunghoon, You, Tackgeun, Kwak, Suha.(2015). "Online Tracking by Learning Discriminative Saliency Map with Convolutional NNs.",Computer Science, 43(13),pp.597-606.

[8] Zeng Haoyang, Edwards Matthew D, Liu Ge.(2016). "Convolutional NNs Architectures for Predicting DNA–protein Binding.", Bioinformatics, 25(11),pp.12-20.

[9] Vinayak R , Lizhen L , Dunson D B.(2016). "Data Augmentation for Models Based on Rejection Sampling.", Biometrika, 13(2),pp.2-10.

[10] Rosario R R.(2017). "A Data Augmentation Approach to Short Text Classification.", Proquest Llc, 17(4),pp.267-275.

[11] Jayaram, Vinay, Alamgir, Morteza, Altun, Yasemin.(2015). "Transfer Learning in Brain-Computer Interfaces.", IEEE Computational Intelligence Magazine, 11(1),pp.20-31.

[12] Rojas-Carulla, Mateo, Schölkopf, Bernhard, Turner, Richard.(2015) "Invariant Models for Causal Transfer Learning.", Statistics, 15(2),pp.67-73.