STUDY ON ARTIFICAL INTELLIGENCE RECOGNITION METHODS FOR MAIZE LEAF LESION IMAGE
/
针对玉米叶片病斑图像的人工智能识别方法研究

Linwei LI[1)], Yanbo SONG[2)], Jie SUN[1)], Yuanyuan LU[2)], Lili NIE[1)], Fumin MA[3)], Xinyu HOU[4)], Juxia LI[1)], Yanwen LI[1)], Zhenyu LIU[5,6 *)]

[1)] College of Information Science and Engineering, Shanxi Agricultural University, Taigu 030801, China;
[2)] College of Life Sciences, Shanxi Agricultural University, Shanxi Agricultural University, Taigu 030801, China;
[3)] College of Energy and Power Engineering, Lanzhou University of Technology, Lanzhou 730050, China;
[4)] School of Information and Communication Engineering, Hainan University, Meilan 570228, China;
[5)] College of Agricultural Engineering, Shanxi Agricultural University, Taigu 030801, China;
[6)] Dryland Farm Machinery Key Technology and Equipment Key Laboratory of Shanxi Province, Taigu 030801, China;
Tel: 0354-6285940; E-mail: lzysyb@126.com. Corresponding author: Liu Zhenyu
DOI: https://doi.org/10.35633/inmateh-71-10

Keywords: *Maize lesion leaf spot, maize eyespot, image classification, neural network, AlexNet architecture*

ABSTRACT
Maize eyespot and maize curvularia leaf spot *are two diseases that often occur on maize leaves. Because of the similarity of the shape and structure, it is difficult to identify the two diseases just relying on the observation of the growers. For the harmfulness and prevention methods are different, it would cause great loss if the disease can't be identified accurately. To address this issue, this paper first employs a connected region feature recognition method to design an automated lesion cropping process after acquiring leaf images with several lesions. Subsequently, a lesion recognition model based on the AlexNet architecture is built and subjected to five-fold cross-validation experiments. The results indicate that the model achieves a comprehensive recognition accuracy exceeding 99%. To further comprehend model characteristics, an analysis of the recognition accuracy and its fluctuations is conducted, revealing that the fractal growth and biological characteristics of the lesions may influence the recognition results. Moreover, the distribution of model parameters could be a potential reason for fluctuations in recognition accuracy rates with increasing number of iterations. This paper could offer valuable reference and support for the intelligent identification and diagnosis of maize and other plant diseases.*

摘要
玉米北方炭疽病和玉米弯孢菌叶斑病是常见于玉米叶片的两种疾病，由于病斑外形结构和形状特征相似，传统的神经网络和图像识别方法较难识别两者，而两者危害程度和防治方法不同，因此误识别会造成较大损失。为此，本文首先基于连通域特征识别实现了对带病叶片图像中病斑的自动裁切。之后，基于 AlexNet 架构构建了病斑识别模型，并采用五折交叉验证法进行试验。试验结果显示，模型的综合识别准确率超过了 99%。为了进一步理解模型特性，本文还对模型的识别准确率及其波动性进行分析，分析结果显示，病斑的分形生长和生物学特性会影响识别结果，模型参数的分布可能是导致识别率随迭代次数增加而波动的潜在原因。本文研究可为玉米及其他植物病害的智能化识别诊断提供有效参考和支持。

INTRODUCTION
Maize eyespot and maize curvularia leaf spot are serious leaf diseases that pose significant threat to maize industry *(Dai et al., 1998; Fu et al., 2016)*. Both the two disease lesions can cover the whole plant leaves, which could lead to the photosynthesis loss of leaves and even cause withering of the plant. Though both the two diseases have different causes and hazards, and maize curvularia leaf spot does not infect the midrib and lacks gray or black mould layer on the surface of the lesion in the later stage in humid environment *(Xu et al., 2000)*, hardness still exists in identifying the two diseases based on intuitive phenotypic observation of lesions of various development stages *(Zhang, 2014)*.

First author. Mail address: sxrllw@126.com. Mailing address: No.1, Mingxian South Road, Taigu District, Jinzhong City, Shanxi Province, China. Tel:0354-6288587.
** Corresponding author. Mail address: lzysyb@126.com. Mailing address: No.1, Mingxian South Road, Taigu District, Jinzhong City, Shanxi Province, China. Tel:0354-6285940*

The lesions of the two diseases are initially small and always would gradually expand into a circular, oval or oblong spot in the late stage. At that time, mature curvularia leaf spot would have the size of (0.5 – 2.0) mm × (0.5 – 1.5) mm, the phenotype would be milky white core surrounded by a brown ring *(Wang et al., 2006; Zhang et al., 2010),* and there would be a translucent light yellow ring with intermittent pattern surrounding the edge of the brown ring *(Shi et al., 2000; Zhang et al., 2017).* As for the leaf spot of maize eyespot in the late stage, the size would be (2.0 – 5.0) mm × (1.0 – 2.0) mm, the core would also be milky white as well as the edge would be a brown ring, and there would be narrow ring or halo with the colour of green surrounding the brown ring's periphery *(Sun et al., 2016; Wang et al., 2010; Chen et al., 2015).* The two kinds of lesions share many similar phenotype *(Sun et al., 2015)*, so using traditional observations to distinguish the two types of lesions always results in misclassification. For the disposition schemes of maize eyespot and maize curvularia leaf spot are different, the misjudgement would lead to improper treatment, which may result in large reduction in maize production and even affect farming in the next years.

Recent advances in computer vision and image processing have provided great opportunities for intelligent identification and assessment of plant diseases *(Shi et al., 2023; Weng et al., 2019; Joseph et al., 2023; Meng et al., 2017; Toda & Okura, 2019). Ahila Priyadharshini et al (2019*) propose a maize leaf disease classification model based on CNN architecture. The architecture is trained to identify four different classes with an average accuracy of 97.89%. *Nagaraju and Chawla (2023)* propose a model named NPNet-19 based on CNN architecture to determine maize crop infections. The model's accuracy and robustness are tested by using a dataset of 15,960 images on six disease classes and one healthy class. Moreover, current researches also attempt to generate unified multi-crop-diseases models to tackle complex tasks. *Picon et al (2019)* introduce a challenging dataset containing almost equally distributed disease stages of seventeen diseases and five crops. This paper proposes three different CNN architectures with contextual non-image meta-data (such as crop information) incorporated onto the image. In addition, to increase the recognizable types of lesions and the accuracy in natural scenes with complex background noise, attention mechanisms which might add model parameters have been proven to be an essential technique *(Guo et al.,2022; Niu et al., 2021)*. *Fang et al (2022)* propose a network model named HCA-MFFNet for maize leaf disease recognition in a complex background. To reduce the influence of complex backgrounds, hard coordinated attention (HCA) in HCA-MFFNet is employed at different spatial scales to extract features from maize leaf disease images. The average recognition accuracy of the model is 97.75%. To reduce the impact of optical noise and background interference, *Li et al (2022)* propose a maize disease identification model based on wavelet threshold-guided bilateral filtering, multi-channel ResNet and attenuation factor. *Albahli et al (2022)* introduce a spatial-channel attention mechanism to focus on infected locations of maize leaf datasets which are taken under diverse and challenging environmental conditions. The approach recognizes multiple maize leaf diseases with an accuracy of 99.89%. *Li et al (2023)* propose an effective data augmentation method of auxiliary classifier GAN (ACGAN) to solve the problem of insufficient data volume. Then to extract multi-scale information on the images of infected maize leaves that are collected from the field of complex background, the DenseNet-based CNN model is developed by integrating multiple expansion modules and Convolutional Block Attention Module (CBAM). Furthermore, small and irregular lesions characterize maize leaf disease in natural scene, which increases identification difficulties of identification models *(Liu & Wang, 2021)*. *Zeng et al (2022)* enhance the model's ability to extract features from leaves with small and irregularly shaped points by replacing the convolutional kernels and activation functions. *Chen et al (2022)* propose a convolutional ensemble network to improve the capability of the model for identifying minute plant lesion features. The method attains an average accuracy of 99.37% on the local dataset and an average accuracy of 99.61% on the open-source PlantVillage dataset.

Although the models mentioned previously have achieved significant results in the identification of maize diseases, their designs focus on recognition challenges in intricate scenarios by studying comprehensive network, which consequently results in higher model complexity and parameter volume. Heavyweight networks need higher performance for equipment, while farmers always prefer to use terminal equipment to take photographs directly of diseased leaves, and human-controlled photographs do not always have too much disturbing factors such as complex lighting, so on the basis of ensuring accuracy, this paper concentrates on the model design of lightweight for image identification of common maize lesions with significant self-similarity. Firstly, automatic cropping process is investigated for lesion images to minimize the impact of subsequent recognition; then recognition model based on the AlexNet architecture is designed and the accuracy is compared to the model based on the architecture of CNN; finally, accuracy and volatility of the model as well as their possible formation mechanisms are analysed for better understanding of the recognition model.

MATERIALS AND METHODS

Image acquisition and processing

Image source

In this paper, the maize variety of Shen 135 is used as an identification host. Then *Kabatiella zeae* which causes maize eyespot and *Curvularia lunata* that causes maize curvularia leaf spot are inoculated onto the leaves of different maize plants separately. The infected maize plants are then placed in an incubator at the temperature 25 °C and the humidity of 75%. After five days of cultivation, images of the infected leaves are captured by the camera of Canon EOS2000D.

Automatic cropping of source images

During the process of image collection, grayscale and median filtering are performed to decrease the possible impact from external optical noise. In addition, the entire diseased leaf is collected by the camera, which would take too much time and manpower to crop the lesions manually. Therefore, this paper automatically cuts the lesions individually by means of the connected domain feature extraction. The automatic cropping process of lesion images is shown in Fig.1.

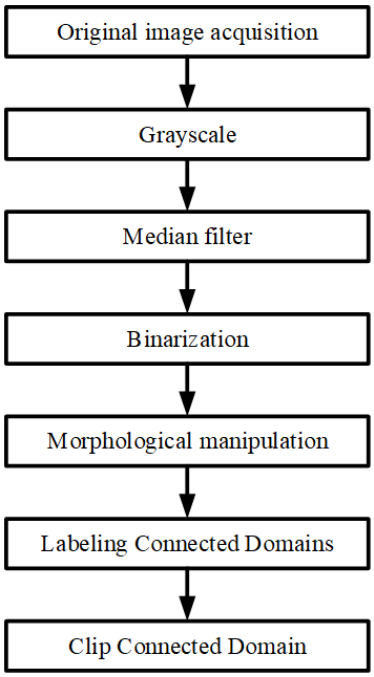


Fig. 1 – Automatic cropping process for lesion images

(1) Pre-processing before cropping

1) Grayscale conversion. Grayscale conversion can eliminate redundant information in the original colour image and simplify the image matrix. In this paper, the average of the R, G, and B parameters in the input image is utilized to determine the output image's grayscale value 'Gray':

$$\text{Gray} = (\text{R} + \text{G} + \text{B})/3 \quad (1)$$

2) Optical noise reduction. The image captured by the camera is susceptible to noise interference. Therefore, Gaussian noise in small pixel area could be eliminated by employing median filtering and the noise in larger pixel area could be reduced by applying dilation-erosion algorithm.

① Median filtering. Median filtering that can suppress noise effectively is a nonlinear signal processing technique based on ranking statistics theory. The basic principle is to replace the disturbed pixel value with the median value which is closer to the real value. The median value is chosen from the neighbourhood of the disturbed point. In this way, the isolated noise point could be eliminated. In this paper, the median filter of 3×3 window is selected for smoothing, which is:

$$f(x,y) = media_{(s,t\in S_{xy})} g\{s,t\} \quad (2)$$

where:

$f(x,y)$ - the image processed by median filtering;

S_{xy} - point (x,y) as the centre of the square field;

$g\{s,t\}$ - Input image.

② Dilation-erosion treatment. The median filtering algorithm has no obvious effect on optical noise in a large area. The erosion algorithm utilizes a specific structuring element to identify the fitted regions. Dilation algorithm serves as the dual operation of erosion to conduct the complement operation of the image.

Image erosion is denoted as $g\{s,t\} \ominus B$, which is defined as:

$$g\{s,t\} \ominus B = \{x: B + x \subset g\{s,t\}\} \quad (3)$$

where:

$g\{s,t\}$ - Input image;

B - structural elements;

$g\{s,t\} \ominus B$ - Consists of translating *B* by x but still including all points x within *g*{s,t}.

Image dilation is expressed as $g\{s,t\} \oplus B$, which is defined as:

$$g\{s,t\} \oplus B = [g\{s,t\}^c \ominus (-B)]^c \quad (4)$$

$g\{s,t\}^c$ - The complement of $g\{s,t\}$.

(2) Lesion identification

1) Recognizing connected domains. The collection of several pixels which conform to the characteristics of the lesions should obey the rules, including: ① All gray values of the pixels are less than or equal to the threshold of the connected domain. ② Pixels within the same connected domain exists in pairs, which means that there exists a path between any two of the pixels.

2) Marking the image. This paper uses pipelined fast labelling method for connected domains based on zigzag and pixel-by-pixel scanning. The determining basis for the start and end in the line run is: when the value of the target pixel and the previous pixel constitutes "01", it means that the start of the run is found; when the value of the target pixel and the previous pixel constitutes "10", it means that the end of the run is determined. For 8-neighborhood connected situation, the relational expression for judging whether the run length of the current row and the run length of the previous row are connected is as follows:

$$Y_0 \leq y_1 + 1 \text{ or } Y_1 \leq y_0 - 1 \quad (5)$$

where:

Y_0, Y_1——The column coordinates of the start and end in the current line run;

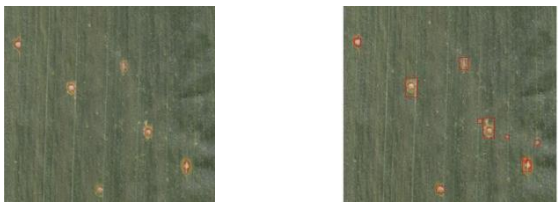
y_0, y_1——The column coordinates of the start and end in the previous line run.

When the two runs are determined as connected, the algorithm usually marks the two runs as equivalent pairs and stores the equivalence relationship in the temporary mark table. Unification of the equivalent pairs will be marked by traversing the temporary mark table. This paper uses the 8-neighborhood connectivity rule to count the pixels which meet the lesions conditions. In the pixel matrix of 3×3, compared with the 4 points of the upper, lower, left and right of the 4-neighborhood, the 8-neighborhood includes 4 extra points of the upper-left, upper-right, lower-left and lower-right. The connectivity is more comprehensive and the versatility is better. The pixels of the multiple lesions connected domains can be grouped together by using the 8-domain connectivity rule.

In this paper, the area containing lesions is extracted as a candidate, and then the grayscale image of the candidate area is gradually removed by changing the threshold value. According to a certain evaluation standard, it would be judged whether the result of each treatment is acceptable. The pixels in the image that conform to a certain connectivity rule (8-neighborhood connectivity) are marked with the same label, and only the binary image that is marked accurately can be used for quantitative extraction of the lesions.

(3) Automatic cropping of lesions

This paper adds external connected rectangle to lesion domains. The comparison between the original identification image and marked image is shown in Fig.2.



a) Before marking *b) After marking*

Fig. 2 – Comparison of the original image and the marked image

One lesion image on maize leaves would be generated by cropping directly through the size of all external connected rectangles. Cycle the operation of the cropping following the labelled numbers and writing them to specific folder will automate the creation of the database, which could replace manpower for faster extraction. Parts of the cropping image in connected domain are shown in Fig.3. In this paper, 24 leaves with lesions are collected, and 23 of the leaves are found to have diseased lesions and the lesions are extracted to form the database.

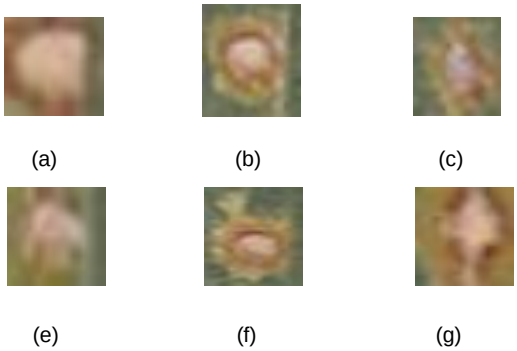
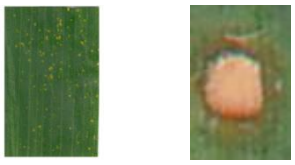


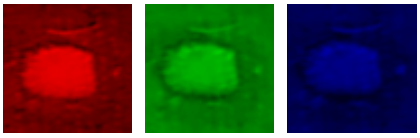
Fig. 3 – Parts of cropping image of selected region in connected domain

Data Preprocessing

After the automatic cropping process, 700 images of the lesions are obtained. Further processing is performed on these 700 images with the following specific steps: (1) All of images are reconstructed into a size of 64 × 64 × 3 (64 pixels in length, 64 pixels in width, 3 channels). Images before and after the treatment are shown in Fig.4 and Fig.5. (2) To reduce the required calculation amount and reduce overflow value which prevents the training or calculation, each pixel value that corresponds to each channel of the image is normalized. (3) In order to eliminate the potential impact of their sequence information on image recognition, the images are shuffled randomly.

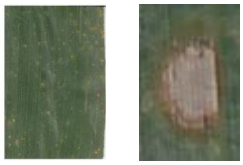


(a)Image of infected leaves (b) The captured lesion image

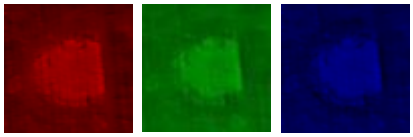


(c)The captured lesion image is reconstructed to 64 × 64 × 3

Fig. 4 – Contrast before and after preprocessing of maize eyespot



(a) Image of infected leaves (b) The captured lesion image



(c) The captured lesion image is reconstructed to 64 × 64 × 3

Fig. 5 – Contrast before and after preprocessing of maize curvularia leaf spot

Model structure

The images in the local dataset are all adjusted to a standard size of 64 × 64 × 3 and all pixel values are normalized to floating-point numbers between 0 and 1. The resized and normalized data tensors are used as the input of the neural network. The data stream passes through 3 convolutional layers and a fully connected layer, and is finally expelled by an output layer containing two neurons. The model structure is shown in Fig.6.

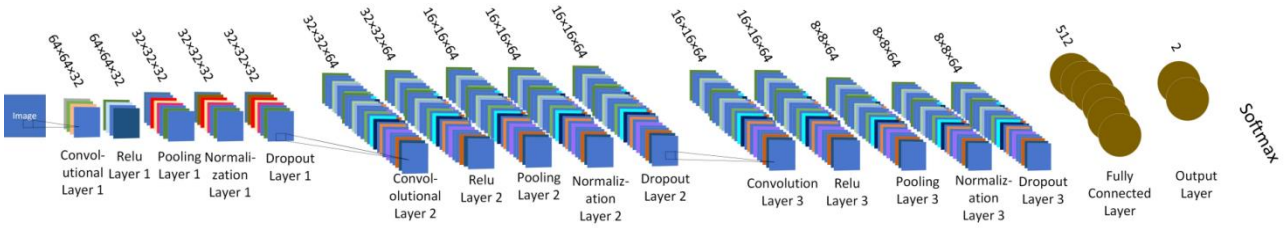


Fig. 6 – Model structure

(1) The first convolutional layer uses 32 convolution kernels of size 3 × 3 × 32 and convolution stride of 1 to perform convolution on each input image, which would extract its local texture and colour features as well as derive a tensor of 32-channel. During the convolution, the edge information of the input tensor is retained by filling in zero.

(2) By applying non-linear correction unit (ReLU) as activation function, the output of the previous convolution layer is used as an input and is mapped into the next max-pooling layer. The filter size of the pooling layer and the step size are set to 2 × 2 and 2, respectively. As a result, the magnitudes of the first and second dimensions (length and width) of the output tensor in the previous layer are reduced by half, which compresses the size of the data stream. The number of parameters required in the subsequent convolution layer and the risk of over fitting will also be reduced. The pooling layer will export an output three-dimensional tensor of size 32 × 32 × 32.

(3) The pooling layer is followed by a local response normalization layer (LRN), through which a competitive mechanism is created based on Eq. 1 *(Krizhevsky et al., 2017)*. Competitive mechanism is used for the activity of local neurons.

After processed by Eq. 1, values with larger feedback can be relatively consumed, while neurons with smaller feedback can be inhibited.

$$b^{i}_{x,\,y} = \frac{a^{i}_{x,\,y}}{\left(k + \alpha \sum_{j=\max(0,\, i-\frac{n}{2})}^{\min(N-1,\, i+\frac{n}{2})} (a^{j}_{x,\,y})^{2}\right)^{\beta}} \tag{1}$$

Where k, α, n and β are hyperparameters, $a^{i}_{x,\,y}$ represents the current activation value, N denotes the total number of feature maps generated by the previous convolution layer, and $b^{i}_{x,\,y}$ is the value of $a^{i}_{x,\,y}$, after being processed by local normalization.

(4) To prevent the possible overfitting and gradient vanishing in the neural network, a random dropout layer is added after the LRN layer. By randomly resetting 20% of the output from the previous layer to zero, the response of these neurons is discarded. The remaining 80% output continues to propagate forward. This processing realizes the neural network regularization and reduces structural risks.

(5) After the first dropout layer, the second convolutional layer follows. In the entire neural network, there are a total of three convolutional layers. The ReLU layer, max-pooling layer, LRN layer, and dropout layer also follow each convolutional layer in sequence.

Each of the second and third convolution layers contains 64 convolution kernels with one convolution step. The filter size of each pooling layer and the step size are set to 2 × 2 and 2, respectively. Each LRN layer is parameterized by setting bias to 1.0, alpha to $1.11e^{-4}$ and beta to 0.75. The drop ratio of each dropout layer is set to 0.2 during the training.

(6) The third dropout layer is connected to a fully connected layer of 512 neurons. A three-dimensional tensor of size 8 × 8 × 64 obtained after the 3 convolutions is expanded into a one-dimensional tensor containing 4,096 values. There are 2 neurons in the output layer. The activation function is set to Softmax. The output layer would export probabilities of the lesion on maize eyespot or maize curvularia leaf spot.

The whole model comprises 2,155,010 parameters, with 896 in the first, 18,496 in the second, and 36,928 in the third convolution layer. There are 2,097,664 parameters in the fully connected layer.

Furthermore, by removing all three LRN layers from the above AlexNet model, a classical CNN model is generated. The CNN model is used to compare the recognition convergence and accuracy of the model based on the AlexNet architecture.

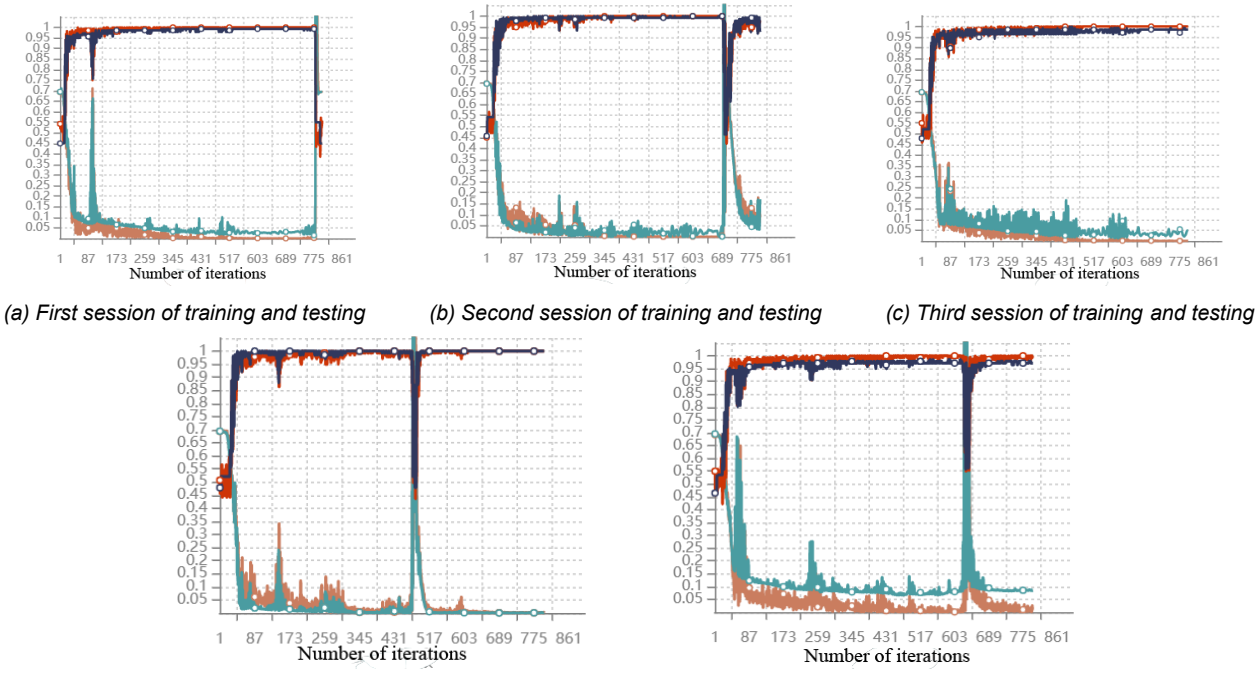
Model training

This paper employs a five-fold cross-validation experiments. Specifically, 700 preprocessed images are randomly assigned to five equal-sized groups with 140 images in the test set and 560 images in the training set. In total, AlexNet and CNN are trained at 800 iterations and then the results on convergence accuracy would be obtained. The process showed above would be conducted 5 times independently and each time the test set and training set is totally different. At the beginning of each training session, the weights and biases are initialized to 32-bit random numbers that follow a normal distribution with a standard deviation of 0.01. For both AlexNet and CNN, the cost function is set as cross entropy and Adam optimizer is used during the training process. The adjustable gradient descent method is used to adjust the parameters of the model. In every iteration of each session, the accuracy and loss values on the training and test sets as well as the maximum, minimum, mean, standard deviation, and parameter distribution of each layer are saved.

RESULTS AND ANALYSIS

Results of model training

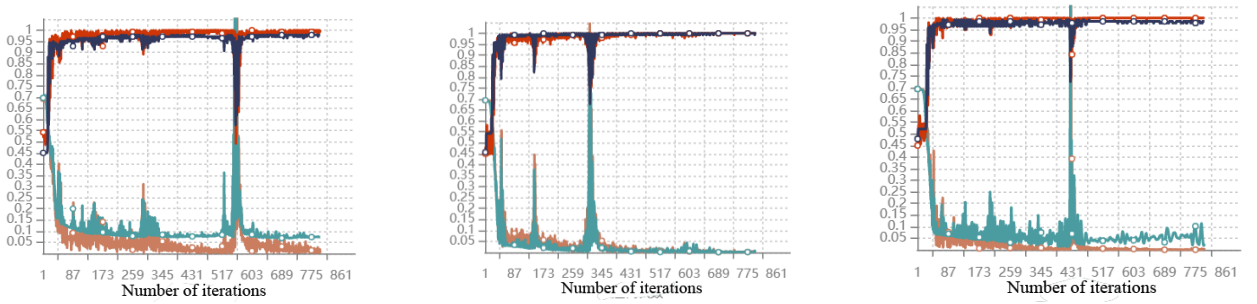
Each iteration of AlexNet or CNN would generate one model. To evaluate the generalization ability of the trained model, 800 iterations in each session is saved, so there would be 800 models for AlexNet or CNN architecture in each session. After each model is produced, it would be used to classify the 140 images in the test dataset. The best-performing model in the last 800 iterations in each time is selected as representative. The maize eyespot is treated as positive example and maize curvularia leaf spot is treated as negative example. If the maize eyespot is erroneously identified as maize curvularia leaf spot, or vice versa, the results will be wrong and will be considered a loss. The accuracy/loss results of the optimization model after each training session are summarized in Fig.7 and Fig.8, where the abscissa and ordinate represent the number of iterations and accuracy/loss value, respectively.



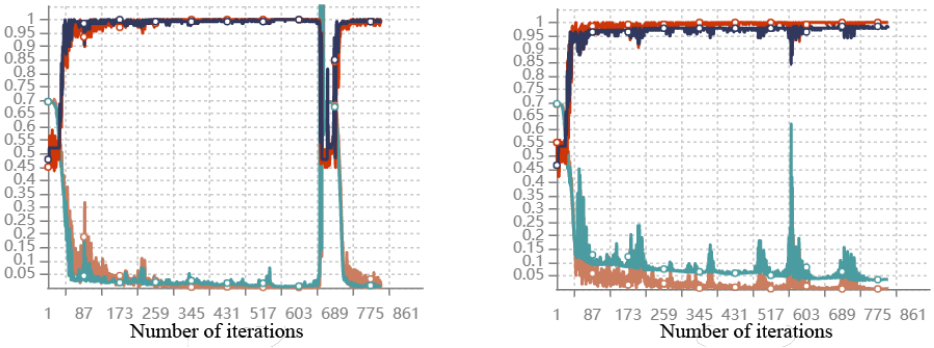
(a) First session of training and testing *(b) Second session of training and testing* *(c) Third session of training and testing*

(d) Fourth session of training and testing *(e) Fifth session of training and testing*

Fig. 7 –Training and testing results of CNN



(a) First session of training and testing *(b) Second session of training and testing* *(c) Third session of training and testing*



(d) Fourth session of training and testing *(e) Fifth session of training and testing*

Fig. 8 –Training and testing results of AlexNet

Evaluation of model generalization

The generalization ability of both models is assessed with the evaluation index shown in Eq. 2, consisting of three formulas to determine the accuracy rate P, recall rate R and F1 score, which are in turn calculated from real example number TP, false negative number FN, false positive number FP, and true negative number TN (Li Tianqi, 2019). The best-performing models produced independently in the 800 iterations of the 5 training sessions are used to calculate the P, recall rate R and F1 score.

The confusion matrices of CNN and AlexNet are shown in Table 1 and Table 2, respectively.

$$\begin{cases} P=TP/(TP+FP) \\ R=TP/(TP+FN) \\ F1=2P\bullet R/(P+R) \end{cases} \quad (2)$$

Table 1

Confusion matrix for CNN

Model number	TP	TN	FP	FN	P	R	F1
1	76	63	0	1	0.992857	0.993506	0.993182
2	63	76	0	1	0.992857	0.992188	0.992522
3	67	72	1	0	0.992857	0.993151	0.993004
4	67	73	0	0	1	1	1
5	63	73	2	2	0.971429	0.971282	0.971355

Table 2

Confusion matrix for AlexNet

Model number	TP	TN	FP	FN	P	R	F1
1	76	60	3	1	0.971429	0.969697	0.970562
2	64	76	0	0	1	1	1
3	67	72	1	0	0.992857	0.993151	0.993004
4	67	73	0	0	1	1	1
5	64	74	1	1	0.985714	0.985641	0.985678

Comparison and analysis of model accuracy

Accuracy comparison and analysis

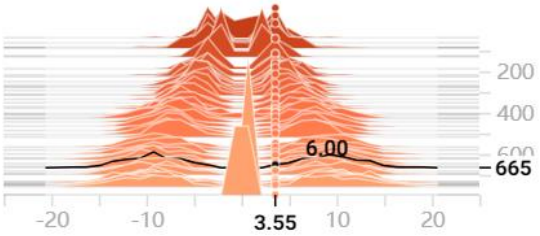
From Tab.1 and Tab.2, it could be observed that both models could achieve 100% on recognition accuracy, recall ratio and F1 score. CNN gets 100% recognition accuracy in the fourth training session, and AlexNet gets 100% recognition accuracy in the first and fourth training session. The recognition accuracy of AlexNet and CNN for maize eyespot and maize curvularia leaf spot are almost same. Though the average recall ratio and F1 score of CNN is slightly higher than AlexNet, AlexNet is obviously easier to show higher performance.

Both models demonstrate outstanding performance in recognizing lesions. This paper would analyse the appearance of 100% on recognition accuracy from the perspective of the biological characteristics and fractal theory. The formation of lesions is determined by the interactions between pathogens and host plants *(Scholthof, 2007)*. Consequently, when the same pathogen infects the same host plant (especially the same leaf), it typically leads to similar morphological and characteristics of the disease lesions. Additionally, the expansion of disease lesions follows fractal patterns over time, resulting in the emergence of similar morphological features in space *(Gong & Zhang, 2002)*. The strong similarity in morphological features in lesions of the same type enables the models to effectively learn the lesion characteristics, and then adapt to the test dataset perfectly, which leads to achieve excellent recognition performance on the test set.

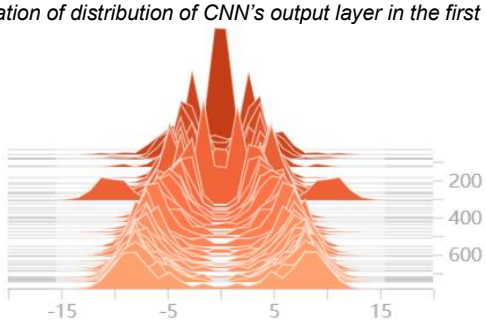
Parameter characteristic and analysis

Fig.7 and Fig.8 illustrate intuitively that during the 10 training sessions, the accuracy-iteration curve of the model generally shows a gradual increase on stability. There is also some fluctuation which is especially evident at some specific points. Taking Fig.8(b) as an example, it shows three distinct spikes where the accuracy sharply transforms from an increasing trend to a rapid decline, and then the accuracy-iteration curve is followed by a quickly increasing trend again. To better understand the reasons and implications of these dramatic fluctuations, it might be useful to analyse the phenomenon from two perspectives: parameter distribution and model principles.

(1) The parameter distribution of each iteration can be obtained as shown in Fig.9 where the X, Y and Z axis represent the parameter range for each iteration, the number of parameter, and the number of iterations, respectively.



(a) Weight variation of distribution of CNN's output layer in the first training session



(b) Weight variation of distribution of AlexNet's output layer in the second training session

Fig. 9 – Weight variation of distribution of the two models with increasement of iterations

Fig.9 demonstrates the changing trend of the variance in the output layer with the increasing iterations. From a holistic perspective, with the increasing number of iterations, the weight distribution in the output layer becomes more dispersed. Not only the output layer but also each layer in the neural network shows the similar trend that the weight values would show dispersed distributions as the number of training iterations increases.

For further analysis on model characteristics, variance was used to describe the dispersion and obtain variance-iteration curve shown in Fig.10. In Fig.10, it can be observed that the variance of the weight distribution shows increasing trend until it stabilizes in terms of the overall iterative process, which aligns with the overall changing trend in the Fig.9(b). At the same time, at certain specific number of iterations, the curve in Fig.10 exhibits clear inflection points indicated by the arrows in the image. By comparing Fig.10 with Fig.8(b), it is evident that these inflection points correspond to the positions where the model's accuracy shows dramatic fluctuations. This result might reveal an inherent connection between the abrupt fluctuations in model accuracy and the weight distribution.

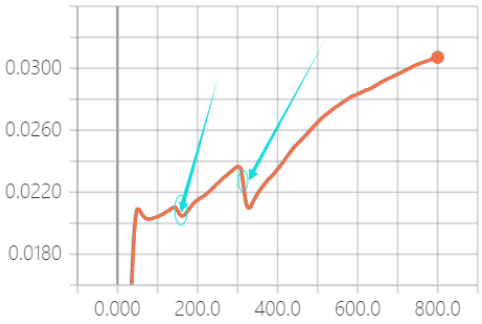


Fig. 10 – Variance-iteration curve of AlexNet *in the second training session*

(2) Besides, from the aspect of the distribution changing of activation values of the fully connected layer shown in Fig.11, the activation values are closer to zero as the number of iterations increases. For the activation function is ReLU, when the activation value of the neuron is updated to zero, it would not be updated to other numbers in subsequent iterations, which means that the node would be inactivated and would not be valid. As the number of iterations increases, more and more neurons become inactivated in the neural network. After the number of iterations exceeds the threshold, the neural network might collapse, which might lead to the corresponding decrease in modeling ability of the features and predictive ability of the neural network. After the collapse, the parameters of the model might be in the unstable state. As training proceeds, the network parameters may gradually move to more appropriate regions through adaptive adjustments on learning rate again, thus improving the performance of the model.

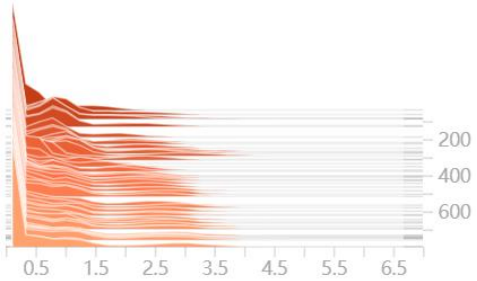


Fig. 11 – Distribution of activation values of the fully connected layer of AlexNet in the second training session

CONCLUSIONS

Misidentification of maize eyespot and maize curvularia leaf spot are found commonly on maize leaves and would lead to invalid operation of farming, which would result in large economic losses. In this paper:

(1) The automatic segmentation theory based on connectivity is utilized and the connected component labelling algorithm is employed. This methodology efficiently separates numerous concentrated lesions from maize leaves, ensuring their accurate extraction and automatic labelling.

(2) Based on AlexNet architecture, neural network model on image recognition is developed. Recognition effect of the model is also compared to image recognition model based on CNN architecture. The result shows outstanding performance on identifying the two kinds of lesions. The evaluation of generalization ability shows achieving 100% accuracy two times of the five training sessions.

(3) From perspective of the accuracy, leaf similarity in disease spots which is resulted from biological and fractal characteristics might lead to the accuracy of 100%. At the same time, from the perspective of fluctuation, activation values and weight distribution might lead to model accuracy's jitter on the way up as the iteration proceeds. This paper constructs a maize lesion recognition model and conducts an analysis of its characteristics, which will lay the groundwork for a portable intelligent lesion recognition system, subsequently contributing to advancements in smart cultivation practices.

ACKNOWLEDGEMENTS

This project is financially supported by the Basic Research Programme (Free Exploration Category) Fund of Shanxi Province, China (20210302124497), Shanxi Agricultural University, China (2021BQ113), Key Laboratory of Equipment and Informatization in Environment Controlled Agriculture, Ministry of Agriculture and Rural Affairs, China (2011NYZD2202) and Shanxi Agricultural University, China (2018023). The authors declare no competing interests.

REFERENCES

[1] Ahila Priyadharshini, R., Arivazhagan, S., Arun, M., & Mirnalini, A. (2019). Maize Leaf Disease Classification Using Deep Convolutional Neural Networks. *Neural Computing and Applications*, 31(12), 8887-8895.

[2] Albahli, S., & Masood, M. (2022). Efficient Attention-based CNN Network (EANet) for Multi-class Maize Crop Disease Classification. *Frontiers in Plant Science*, 13, 1003152.

[3] Chen, G., Chen, Z., Zheng, T., Ma, S., Yan, Z., & Liu, G. (2015). The Diagnosis and Identification of Three Fungal Leaf Spot Disease of Maize Leaf in the Hexi Corridor *(*河西走廊制种玉米 3 种真菌性叶斑病的诊断与识别*). Seed,* 34(3), 94-96.

[4] Chen, J., Zeb, A., Nanehkaran, Y. A., & Zhang, D. (2022). Stacking Ensemble Model of Deep Learning for Plant Disease Recognition. *Journal of Ambient Intelligence and Humanized Computing*, 14(9), 12359-12372.

[5] Dai, F., Wang, X., Zhu, Z., Gao, W., Huo, N., & Jin, X. (1998). Curvularia Leaf Spot of Maize: Pathogens and Varietal Resistance (玉米弯孢菌叶斑病研究). *ACTA Phytopathologica Sinica*, 28(2), 123-129.

[6] Fu, J., Jing, D., Liu, Z., & Zhou, R. (2016). Review of Epidemic Dynamics and Forecasting and Warning of Maize Leaf Disease (玉米叶部病害流行动态及预测预警研究进展). *Journal of Jilin Agricultural University,* 38(6), 651-655.

[7] Fang, S., Wang, Y., Zhou, G., Chen, A., Cai, W., & Wang, Q. (2022). Multi-channel Feature Fusion Networks with Hard Coordinate Attention Mechanism for Maize Disease Identification under Complex Backgrounds. *Computers and Electronics in Agriculture*, 203, 107486.

[8] Gong, G., & Zhang, S. (2002). Fractal analysis of patch patterns on plant diseases (植物病害病斑形状的分形研究). *Plant Protection*, 28(6), 9-12.

[9] Guo, M., Xu, T., Liu, J. , Liu, Z., Jiang, P., & Mu, T., et al. (2022). Attention Mechanisms in Computer Vision: A Survey. *Computational Visual Media*, 8(3), 331-368.

[10] Joseph, D. S., Pawar, P. M., & Pramanik, R. (2023). Intelligent Plant Disease Diagnosis Using Convolutional Neural Network: A Review. *Multimedia Tools and Applications*, 82(14), 21415-21481.

[11] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, 60(6), 84-90.

[12] Li, E., Wang, L., Xie, Q., Gao, R., Su, Z., & Li, Y. (2023). A Novel Deep Learning Method for Maize Disease Identification Based on Small Sample-size and Complex Background Datasets. *Ecological Informatics*, 75, 102011.

[13] Liu, J., & Wang, X. (2021). Plant Diseases and Pests Detection Based on Deep Learning: a Review. *Plant Methods*, 17, 1-18.

[14] Li, Z., Zhou, G., Hu, Y., Chen, A., Lu, C., & He, M. (2022). Maize Leaf Disease Identification Based on WG-MARNet. *PLoS One*, 17(4), e0267650.

[15] Meng, Y., Chen, G., Lu, J., & Xu, R. (2017). Simulink Platform in Video Image Real-time Diagnosis of Maize Disease (Simulink 平台在玉米病害视频图像中的实时诊断). *Journal of Jilin Agricultural University*, 39(4), 483-487.

[16] Nagaraju, M., & Chawla, P. (2023). Maize Crop Disease Detection Using NPNet-19 Convolutional Neural Network. *Neural Computing and Applications*, 35(4), 3075-3099.

[17] Niu, Z., Zhong, G., & Yu, H. (2021). A Review on the Attention Mechanism of Deep Learning. *Neurocomputing*, 452, 48-62.

[18] Picon, A., Seitz, M., Alvarez-Gila, A., Mohnke, P., Ortiz-Barredo, A., & Echazarra, J. (2019). Crop Conditional Convolutional Neural Networks for Massive Multi-crop Plant Disease Classification over Cell Phone Acquired Images Taken on Real Field Conditions. *Computers and Electronics in Agriculture*, 167, 105093.

[19] Shi, J., Liu, Y., & Wei, L. (2000). On Pathogens of Maize Curvularia Leaf Spot (玉米弯孢菌叶斑病病原菌的研究). *Journal of Shenyang Agricultural University,* 31(5), 479-481.

[20] Sun, J., Xiao, S., Xu, J., Liu, K., Ma, C., Xue, C., & Chen, J. (2016). Occurrence Condition and Chemical Control of Maize Eyespot in Liaoning Province (辽宁省玉米北方炭疽病发生条件与药剂防治). J*ournal of Maize Sciences, 24*(5), 147-151.

[21] Sun, J., Xiao, S., Lu, Y., Tu, G., Xue, C., & Chen, J. (2015). Isolation Identification and Biological Characteristics of Aureobasidium Zeae in Liaoning Province (辽宁省玉米北方炭疽病菌的分离鉴定及生物学特性). *Journal of Plant Protection,* 42(6), 927-934.

[22] Scholthof, K. B. G. (2007). The Disease Triangle: Pathogens, the Environment and Society. *Nature Reviews Microbiology*, 5(2), 152-156.

[23] Shi, T., Liu, Y., Zheng, X., Hu, K., Huang, H., Liu, H., & Huang, H. (2023). Recent Advances in Plant Disease Severity Assessment Using Convolutional Neural Networks. *Scientific Reports*, 13(1), 2336-2349.

[24] Toda, Y., & Okura, F. (2019). How Convolutional Neural Networks Diagnose Plant Disease. *Plant Phenomics*, 2019, 9237136.

[25] Wang, C., Chen, L., Guo, X., Lu, Z., & Gao, J. (2006). Study on Some Problems of Maize Curvularia Leaf Spot Caused by Curvularia Lunata (玉米弯孢菌叶斑病发生和防治若干问题研究). *Journal of Maize Sciences, 14(2),* 144-146+149.

[26] Wang, X., Shi, J., Jin, Q., Li, X., & Sun, S. (2010). Field Manual of Corn Pests and Diseases (玉米病虫害田间手册). Beijing/*China: China Agricultural Science and Technology Press.*

[27] Weng, Y., Zeng, R., Wu, C., Wang, M., Wang, X., & Liu, Y. (2019). A Survey on Deep-learning-based Plant Phenotype Research in Agriculture (基于深度学习的农业植物表型研究综述). *Scientia Sinica Vitae,* 49(6), 698-716.

[28] Xu, X., Dong, H., Jiang, Y., Qiao, Y., Liu, Z., & Sun, J. (2000). Preliminary Studies on the Northern Anthracnose Disease in Maize (辽宁省玉米新病害——北方炭疽病研究初报). *Journal of Shenyang Agricultural University,* 31(5), 507-510.

[29] Zhang, S., & Zhang, C. (2014). Maize Disease Recognition based on Local Discriminant Algorithm (基于局部判别映射算法的玉米病害识别方法). *Transactions of the Chinese Society of Agricultural Engineering*, 30(11), 167-172.

[30] Zeng, W., Li, H., Hu, G., & Dong, L. (2022). Identification of Maize Leaf Diseases by Using the SKPSNet-50 Convolutional Neural Network Model. *Sustainable Computing: Informatics and Systems*, 35, 100695.

[31] Zhang, X., Su, Q., Song, S., Du, J., Liu, Y., Zhang, W., Li, H., & Jin, Q. (2010). Pathogenicity Differentiation and RAPD Analysis of Curvularia Lunata in Northeast China (东北地区玉米弯孢菌致病性分化与 RAPD 分析). *Journal of Maize Sciences,* 18(6), 122-126.

[32] Zhang, Y., Zhang, S., Yang, S., Zhou, Y., & Zhao, W. (2017). Pathogen Identification of Curvularia Leaf Spot of Maize in Heilongjiang Province (黑龙江省玉米弯孢霉叶斑病病原鉴定). *Journal of Northeast Agricultural University,* 48(6), 17-23.