IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 8, 2020, accepted March 12, 2020, date of publication March 23, 2020, date of current version April 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2982443

Maize Leaf Disease Identification Based on Feature Enhancement and DMS-Robust Alexnet

MINGJIE LV[1], GUOXIONG ZHOU[1], MINGFANG HE[1], AIBIN CHEN[1], WENZHUO ZHANG[1], AND YAHUI HU[2]

[1]College of Computer and Information Engineering, Central South University of Forestry and Technology, Changsha 410004, China
[2]Institute of Plant Protection, Hunan Academy of Agricultural Sciences, Changsha 410125, China

Corresponding authors: Guoxiong Zhou (51840157@qq.com) and Yahui Hu (huyah627@163.com)

This work was supported in part by The State Bureau of Forestry 948 Project, China, under Grant 2014-4-09, in part by the National Natural Science Foundation of China under Grant 6170344, and in part by the National Key Research and Development Program of China under Grant 2017YFD0301504.

ABSTRACT The identification of maize leaf diseases will meet great challenges because of the difficulties in extracting lesion features from the constant-changing environment, uneven illumination reflection of the incident light source and many other factors. In this paper, a novel maize leaf disease recognition method is proposed. In this method, we first designed a maize leaf feature enhancement framework with the capability of enhancing the features of maize under the complex environment. Then a novel neural network is designed based on backbone Alexnet architecture, named DMS-Robust Alexnet. In the DMS-Robust Alexnet, dilated convolution and multi-scale convolution are combined to improve the capability of feature extraction. Batch normalization is performed to prevent network over-fitting while enhancing the robustness of the model. PRelu activation function and Adabound optimizer are employed to improve both convergence and accuracy. In experiments, it is validated from different perspectives that the maize leaf disease feature enhancement algorithm is conducive to improving the capability of the DMS-Robust Alexnet identification. Our method demonstrates strong robustness for maize disease images collected in the natural environment, providing a reference for the intelligent diagnosis of other plant leaf diseases.

INDEX TERMS Image enhancement, dilated convolution, multi-scale convolution, maize leaf disease, convolutional neural network.

I. INTRODUCTION

With the development of agricultural technology, image processing technology is increasingly applied to detect and classify the quality of agricultural products [1]–[3]. Extensive studies have been performed in various fields like weed identification [4]–[6], crop diseases and insect pests identification and diagnosis [8], [9]. They have become the focus of attention for modern agricultural research recently, and some remarkable results have been obtained. Maize represents a significant food crop as well as a feed source. It is also one of the most widely cultivated food crops with the highest total yield around the world. However, with the development of maize production, some problems arise. For example, due to the abuse of pesticides and the insufficiency of maize protection, the change to pathogen varieties and the variety of maize diseases are on the rise. Therefore, how to identify maize diseases in a quick and accurate way is of massive significance to the stable production of maize crops.

Maize leaf diseases could be manifested in many different symptoms [10]–[12], which requires specially trained experts to make identification and diagnosis. However, inexperienced farmers are prone to diagnostic errors. Besides, not only is artificial recognition diagnosis time-consuming and laborious, but it also shows subjectivity to some extent. In early maize disease image recognition, traditional global bottom features such as color, texture, and shape are used on a frequent basis to describe the attributes of disease spots in crop leaf disease images. For maize crops that are still in the growth, nevertheless, due to the impact made by the environment, different types of lesion changes and other factors, there is a possibility that maize leave diseases lead to the appearance of powdery matter, which makes it more difficult to perform feature extraction, thus affecting the accuracy of

The associate editor coordinating the review of this manuscript and approving it for publication was Shenghong Li.

identification. Meanwhile, the accuracy in the identification of maize disease images in practice is more easily affected by such external factors as illumination, shadow and so on. Besides, there are clear disparities in the corresponding pests and diseases in different growth stages for maize, which adds to the difficulty of identification significantly. At present, deep learning has attracted increasing attention from many researchers. It demonstrates obvious advantages over shallow models both in feature extraction and identification. The convolutional neural network (CNN) [13] structure has been widely applied in a variety of different fields. With direct input of the original data on maize image into the network, CNN structure is capable to learn features from the training data automatically. but its advantages do not necessarily ensure that it can address the previous problems, such as maize leaf disease identification referred to in this paper. Although manual feature extraction is avoided, the existence of noise and unclear features in the original image of maize could still result in error accumulation, which makes it impossible to improve the identification accuracy of the network.

In order to address the above problems, a maize disease identification method based on the combination of maize leaf diseases feature enhancement and DMS-Robust Alexnet method is proposed in this paper. Firstly, an image enhancement method for maize leaves is suggested, which has the capability to enhance the features of maize under a complex environment. Then, The DMS-Robust Alexnet neural network is introduced to classify and recognize the maize image after the above process. As demonstrated by the experimental results, in comparison with the traditional feature extraction classification method and the neural network model without the feature of image enhancement, this method is capable to achieve a higher accuracy of identification. The proposed method is clearly advantageous, and the contributions made in this paper are summarized as follows.

1) A method that combines image enhancement and deep learning for maize leaf recognition method is proposed, which can achieve high-accuracy recognition of maize leaf image in a complex environment.
2) DMS-Robust Alexnet for maize leaf image recognition and classification is proposed based on backbone Alexnet architecture, several improvements have been made to make sure our dataset can be well trained, thus our method can perform better than the Alexnet and other backbone architecture.
3) Compared with traditional classification methods and deep neural network without image enhancement, the proposed method achieves a high accuracy of classification and recognition.

II. RELATED WORK

In the most recent years, as the crop industry develops, an increasing number of experts and scholars both at home and abroad have started to explore how to identify crop diseases in a quick yet effective way. A variety of different methods like digital image processing [14], support vector machine (SVM) [15], and neural network [16] can be applied to the detection and classification of leaf diseases. Neto *et al.* [17] employed Fourier Description Factor to describe plant leaves. Then, discriminant analysis was conducted to classify four species of plants in which the average correct recognition rate reached as high as 89.4%. Ikorasaki *et al.* [18] developed an expert diagnosis system for the diagnosis of maize crop diseases with the assistance of Bayesian theorem, as a result of which the precision rate reached 90%. Despite the positive recognition results that have been obtained, the success rate of modeling recognition for non-linear data remains very low, and the success rate of modeling recognition in crop disease detection is far less than unsatisfactory. Lv *et al.* [19] applied pulse coupled neural network (PCNN) image segmentation method based on minimum cross-entropy to segment apple images. In their experiments, 50 images were used for validation, 93% of which were recognized accurately. Zhang *et al.* [20] segmented the lesions and extracted the color, shape and texture features of the lesions. Then, five maize leaves were identified by applying k-nearest-neighbor (KNN) classification algorithm, with the recognition accuracy reaching over 90%. Despite the significant improvement to recognition performance, their model remains susceptible to environmental impact. Besides, over-segmentation is easy to occur, which has a negative impact on the outcomes of recognition. Zhang *et al.* [21] employed the SVM method to classify maize disease images collected on mobile internet, with the best average recognition accuracy reaching 83.2%. Aravind *et al.* [22] applied the gray level co-occurrence matrix to extract texture features of maize leaf disease, and then used multi-class SVM to classify maize diseases. The classification method based on SVM is only suitable for small samples, and is incapable to achieve high accuracy of recognition for large samples as only an 83.7% average accuracy was reached in their experiments.

From the above literature, it can be found out that the previous approaches to crop diseases identification are based on the color and shape characteristics of the disease. They extract the disease features within the range of visible light, and then apply the classification methods to recognize diseases. These methods are heavily dependent on skilled technicians who have specialist knowledge to refine the types of disease. In addition, for most recognition tasks, the images used in experiments are frequently collected in highly demanding environments where no interference from the external environment is allowed. Therefore, picture collection in different external environments will have a negative impact on the test results. When new viruses are discovered, regardless of whether the traditional feature extraction is divided or not, the effectiveness of classification is difficult to ensure.

Deep learning has made remarkable achievements in the field of recognition [23]–[31]. It can extract useful feature representations from a large number of input images. Deep learning is capable of timely and accurate identification of crop diseases to achieve crop disease prevention. Not only does it improve the accuracy of plant disease identification,

but it also expands the scope of application for computer vision in precision agriculture. Dechant *et al.* [32] trained CNN to automatically recognize the northern leaf blight of maize. This method addresses the problem of irregular phenomena in field plant images. The precision of the scheme reached as high as 96.7%. Sibiya *et al.* [33] applied CNN to recognize and classify maize disease images captured by mobile phones. The average recognition precision reached 92.85%. Brahimi *et al.* [34] collected about 15,000 images of tomato leaf diseases and classified them into 9 diseases based on Alexnet, which led to desirable recognition results. Sladojevic *et al.* [35] collected over 4,000 disease images on the network. Following data augmentation, more than 30000 image samples were obtained. They were classified into 15 categories based on CafeNet, with the classification precision ranging between 91% and 98%. In addition, some researchers [36], [37] constructed different CNN models to change the ratio between the training set and the testing set, thus improving the accuracy in identifying maize diseases to some extent. Despite the positive results produced by the research above, the increasing sample size and the extended time of training convergence have a negative impact on the accuracy of recognition. In order to improve the accuracy in identifying maize leaf disease, it is essential to design a recognition model with a moderate sample size and high recognition accuracy.

The original maize images collected from the actual agricultural environment show various undesirable features due to environmental and dust factors. The CNN model is incapable to distinguish the importance of features. If such undesirable features are learned by CNN, the classification results will be severely affected. Allowing for this, a method based on image enhancement and the DMS-Robust Alexnet is proposed to identify maize leaf diseases. Our enhancement method is capable to eliminate the noise of maize image in practical applications while enhancing the maize lesion features, and can achieve high-precision identification of maize disease images.

III. MATERIALS AND METHODS

A. LESION FEATURES ENHANCEMENT AND DMS-ROBUST ALEXNET

Maize disease image processing represents a crucial and challenging task. The acquisition of high-quality disease images often requires experienced professionals to shoot. Meanwhile, it will incur a substantial amount of equipment configuration costs and time cost over the course of transmission, the image quality deteriorates due to the influence of external environmental factors and equipment. In this case, the disease image is blurred, the contrast decreases, and the details are lost. Therefore, the original image obtained in a complex environment needs to be denoised and enhanced. Traditional image processing method cannot effectively process the image information globally, and it is easy to produce halos and graying under the condition of uneven lighting.

To resolve the problem that traditional image preprocessing methods are difficult to achieve excellent global effect in a complex environment in complex, an improved Retinex enhancement method is proposed in this paper. Firstly, the maize disease image is transformed from RGB space to HSV space. Secondly, the value component V is transformed into the wavelet domain. Thirdly, in the wavelet transform the low-frequency coefficient is enhanced by applying the improved Retinex image enhancement algorithm and the high-frequency coefficient is denoised by using the Donoho threshold method. Finally, S component is processed by segmented logarithmic stretching to extract maize lesion features for further improvement to the quality of maize disease image.

The CNN is regarded as the best algorithm to extract high-level semantic features. With more abstract features extracted layer by layer from images, the CNN can achieve superior performance in image classification. Thus, achieving a superior performance in image classification. The typical CNN include Google Net [38], ResNet [39], VGG [40], etc. However, the network structure of these models is highly complex and thus requires a large scale of maize images.

Besides, a large number of maize disease images are difficult to obtain, and the above CNN based networks are too deep to be trained when only a small amount of training data is available. To tackle this issue, a novel DMS-Robust Alexnet based on Alexnet [41] backbone architecture is proposed. AlexNet deepens the structure of the network on the basis of LeNet, learning richer and higher-dimensional image features, with 5 convolutional layers, 3 fully connected layers, using Dropout to suppress overfitting, and using Relu as the activation function. The improvements in DMS-Robust Alexnet are as follows.

1) A recognition method that combines dilated convolution and multi-scale convolution fusion is proposed. Meanwhile, the structure and parameters of the DMS-Robust Alexnet are adjusted and optimized well on our dataset.
2) The generalization ability and convergence speed of network are improved with the addition of batch normalization layer.
3) PRelu activation function is used instead of Relu or Sigmoid activation function, and AdaBound is taken as the optimizer of the network to improve the effectiveness of learning convergence for the network.

The process of disease identification based on maize leaf feature enhancement and DMS-Robust Alexnet is shown in Figure 1.

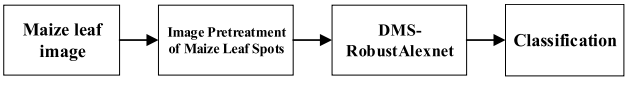

FIGURE 1. Maize leaf feature enhancement and DMS-Robust Alexnet framework flow chart.

Firstly, the data set of maize leaf disease is obtained. Secondly, all the data are enhanced to derive disease feature enhancement images. Thirdly, DMS-Robust Alexnet is trained using the enhanced data set. Finally, the trained model is applied to maize disease identification.

B. GRAPHIC GATHERING

In every single stage of maize disease identification, ranging from the initial stage of image processing to the final stage of evaluating the performance of the recognition algorithm, the significance of data sets is self-evident.

In this paper, maize disease images are collected from different sources, partly from those such as Plant Village, crop disease data set of 2018 global AI challenge and Google websites, in which different periods of the occurrence of maize leaf diseases are covered. Partly in collaboration with Hunan Academy of Agricultural Sciences of China, all images are captured with Sony ILCE-7M2 under natural illumination. The vertical distance between the camera and the maize leaves ranges from 3 to 7 centimeters, and the images are obtained from four different angles. For each image, the sub-image containing the lesion is first obtained by manual cropping, and the number of images is expanded by adjusting angles and cutting appropriately. The collected images are classified into corresponding categories by experts. The number and proportional distribution of maize disease images for each category are shown in Table 1.

TABLE 1. Number and proportion of maize diseases.

Category	Number	Proportion
Common rust	1766	24.55%
Gray leaf spot	1273	17.70%
Northern leaf blight	1129	15.69%
Zinc deficiency	496	6.89%
Round spot	375	5.21%
Fall army worm	607	8.43%
Healthy	1547	21.50%

As is shown in Table1, the distribution of disease samples obtained is very uneven. In deep learning, the uneven distribution of samples will affect the accuracy of model recognition [42]. Therefore, 5 general ways are employed to augment a small number of samples data: horizontal flip, vertical flip, horizontal-vertical flip, random rotation, perspective transformation.

To avoid severe distortion of the transformed images, the displacement of the corresponding points in the perspective transformation is limited to less than 10% of the image's side length, and the canvas size of the target image is consistent with the source image. With this method, the size of the small number of sample data is increased by five times. The data augmentation methods are shown in Table 2.

TABLE 2. Date augmentation methods.

Methods	Specific operations
Flipping	horizontal flip, vertical flip, horizontal-vertical flip
Rotation	Random rotate a certain angle with the image center as the origin (0^{o}, 90^{o}, 180^{o}, 270^{o})
Perspective transformation	Random perspective transformation (up, down, left, right)

The number and proportion distribution of each category after data augmentation are shown in Table 3.

TABLE 3. Number and proportion of maize diseases after data augmentation.

Category	Number	Proportion
Common rust	1766	14.44%
Gray leaf spot	1582	12.93%
Northern leaf blight	1631	13.34%
Zinc deficiency	1987	16.25%
Round spot	1893	15.48%
Fall army worm	1821	14.89%
Healthy	1547	12.65%

C. MAIZE LEAF LESION FEATURE ENHANCEMENT

There are some irrelevant image information in both healthy and diseased maize images. The image preprocessing function is purposed, which has following advantages: 1). Enhancing the task-related feature information. 2). Removing irrelevant information to the maximum extent. 3). Improving the reliability of image recognition. At present, there are two major image enhancement methods for maize lesion feature in complex environment, which are image fusion and image enhancement. The method of maize disease feature enhancement based on image enhancement is characterized by simplicity, effectiveness and widespread applications. It is almost applicable in color images and on non-fixed real agricultural occasions. It mainly includes histogram method, Retinex [43] principle method and wavelet-based method. Retinex algorithm is known as an image enhancement method characterized by sharpening, constant color, large dynamic range compression, and high color fidelity. However, as the traditional Retinex algorithm relies on Gauss filter to estimate the irradiation component of maize image, it is easy to blur the maize image, cause part of the information to be lost, and make the lesion features less obvious. The algorithm based on Retinex is premised on gray-world assumption. Not only is enhancement processing in RGB color space easy to cause image color distortion,

it also requires a large amount of calculation and results in low efficiency. In the wavelet domain, the global information and contour information of maize disease images are primarily distributed in the low-frequency region, while the local information, details information and noise of maize disease image are mostly distributed in the high-frequency region, as a result of which maize disease image can be transformed into the wavelet domain, before being processed in the high-frequency region and the low-frequency region with different methods. Depending on the different spectral characteristics, better processing effect can be achieved by wavelet transform.

With consideration given to the advantages of wavelet transform and the drawbacks of Retinex algorithm, a novel algorithm for maize image enhancement called WT-DIR (Wavelet transform Donoho threshold and improved Retinex) is proposed. Firstly, the image of maize disease is transformed from RGB space to HSV space, the value component V is wavelet transformed, the low-frequency coefficients are enhanced by applying the improved Retinex image enhancement algorithm, and the high-frequency coefficients are denoised by using Donoho threshold method. Finally, the noise of maize image is reduced and the lesion features are enhanced by WT-DIR. For improvement to the details of maize leaf disease features, saturation component S is processed by segmented logarithmic stretching to extract maize lesion information, which is aimed at further improving the quality of maize disease image. The flow chart of the algorithm is illustrated in Figure 2.

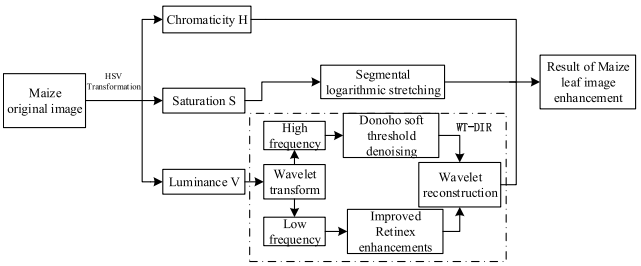

FIGURE 2. Image enhancement for maize disease images.

1) WT-DIR ALGORITHM

The WT-DIR algorithm can act as following three steps:

Step 1 (Choosing the Wavelet Basis Function): Due to the variations of ability that wavelet bases have to perform time-frequency localization and multi-resolution analysis, different wavelet bases that degrade the same maize leaf image will lead to different results, thus affecting the later recognition task. Therefore, the selection of wavelet bases plays a significant role in maize image degradation. Sym [44] wavelet demonstrates better symmetry and less phase shift in reconstruction, which makes it better suited to image processing. Therefore, sym4 wavelet is chosen in this paper to decompose image.

Step 2 (The Improved Retinex Algorithm): Following illumination estimation of low-frequency coefficients of maize image decomposed by wavelet transform, halo phenomenon is easy to occur. In this paper, a guided filter with edge-preserving property is employed to operate the maize disease image, based on which the illumination component image of maize image is obtained. In line with Retinex principle, the reflection component image is derived from logarithmic transformation. Then, the reflection component is corrected by Gamma transform. The above operations are repeated if the result of Gamma transform falls short of the signal noise ratio (SNR) threshold. The flow chart is illustrated in Figure 3.

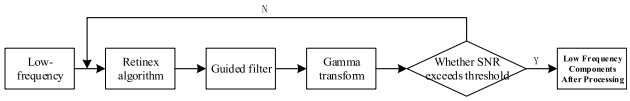

FIGURE 3. The improved Retinex algorithm.

In the improved Retinex algorithm, the low-frequency part of V in the wavelet domain (which is later collectively referred to as the low-frequency image of maize) is regarded as the product of illumination component $L(x, y)$ and reflection component $R(x, y)$ as follows.

$$I(x, y) = L(x, y) \times R(x, y) \tag{1}$$

where (x, y) represents the two-dimensional coordinate of the low-frequency component of the maize image following wavelet transform. Usually, the reflection component $R(x, y)$ is the clear image that requires solution, and there is a necessity for the estimation of $R(x, y)$ to be converted to the logarithmic domain.

$$\log \tilde{R}(x, y) = \log I(x, y) - \log \tilde{L}(x, y) \tag{2}$$

$\tilde{R}(x, y)$ and $\tilde{L}(x, y)$ represent the estimated values of reflection and irradiation components of low frequency of maize image. Equation (2) demonstrates that color constancy can be achieved when the low-frequency image of maize is independent of the influence exerted by ambient light. It is essential to calculate the illumination component of the estimated reflection component with high frequency information. The illumination component of low-frequency of maize image is estimated using the idea of center-around as follows.

$$\tilde{L}(x, y) = I(x, y)^* F(x, y) \tag{3}$$

where * indicates convolution operation, and F denotes center-around function. In general, the center-around function shows robust capability of dynamic compression, such as Gauss function, which is purposed to simulate the slow signal transformation in the original maize image.

As shown in the above calculation process, the Gauss operator of Retinex algorithm is incapable to estimate illumination well. The significant illumination changes lead to the occurrence of halo phenomenon, and the range of bright region will be compressed by the logarithmic processing, which results in partial loss. In order to address the problems, the guided filter is employed to optimize Retinex. With the idea of least square

method, Guided filtering takes into account the phenomenon of sudden change to illumination in the image, for operations through box filter and integral image technology. The speed of execution is irrelevant to the size of the filter window, and the efficiency in estimating illumination component is high, as a result of which the halo phenomenon of maize image can be reduced to a certain extent. The solution of reflection component based on guided filter can be expressed as follows.

$$\tilde{R}_V(x, y) = \log I_V(x, y) - \log [f\ (I_V(x, y))] \tag{4}$$

where (x, y) indicates the two-dimensional coordinate of the low-frequency component of the maize image, $\tilde{R}_V(x, y)$ denotes reflection component, f represents the guided filter function and $I_V(x, y)$ refers to the value component of maize image in HSV space. The guided filter function can be viewed as local linear model as follows.

$$q_j = a_k I_{V,j} + b_k \tag{5}$$

where q_j represents the result of linear transform, and k denotes the center pixel of window ω_k. In window ω_k, coefficients a_k and b_k are constant, which can be calculated as follows.

$$a_k = \frac{\frac{1}{N_{\omega_k}} \sum_{j \in \omega_k} I_{v,j}^2 - \mu_k^2}{\sigma_k^2 + \delta} \tag{6}$$

$$b_k = (1 - a_k)\, \mu_k \tag{7}$$

where μ_k and σ_k denote the mean and standard deviation in window ω_k respectively. N_{ω_k} indicates the total pixels in window ω_k. δ represents regularize, which can be used to indicate the degree of balance smoothness and margin maintenance. The bigger the δ, the better the smoothness but the worse the margin maintenance. In order to obtain stable q_j, the image needs to be averaged, as a result of which the linear model of equation (5) can be applied to the entire image of maize disease, and the guided filter function can be derived as follows.

$$f_j\,(I_V(x, y)) = \frac{1}{N_{\omega_k}} \sum_{k,j \in \omega_k} \left(a_k{}^* I_{V,j}(x, y) + b_k\right) \tag{8}$$

With Equation (8) substituted into Equation (4), the estimated value $\tilde{R}'_V(x, y)$ of reflection component of maize image is obtained from inverse logarithmic transformation.

$$\tilde{R}'_V(x, y) = \exp\left(\tilde{R}_V(x, y)\right) \tag{9}$$

After the above processing, the maize image will be made darker, and the quality of maize image needs to be improved. In this paper, a non-linear global Gamma correction method is applied to adjust the brightness of maize image as follows.

$$R_V(x, y) = \left(\tilde{R}'_V(x, y)\right)^{1/\gamma}, \tag{10}$$

where γ represents a correction parameter, the range of which is $[1, +\infty)$. In this paper, in order to make the final maize disease image the same as the original scene, there is no distortion, the empirical value is 3.

Step 3 (Donoho Soft Threshold in Wavelet Transform): In the wavelet domain, the noise of maize disease image is concentrated in the high frequency component, which corresponds to the wavelet coefficients with a lower absolute value. Generally speaking, there are two kinds of thresholding functions, which are hard thresholding function and soft thresholding function. Nevertheless, the poor continuity of the estimated wavelet coefficients resulting from hard thresholding has a potential to cause oscillation of reconstructed signals. The image will be subject to distortion as a result of ringing and pseudo-Gibbs effect, which will have a negative impact on the image recognition task of maize diseases in the later stage. The key to wavelet threshold denoising lies in threshold selection. Only by choosing the appropriate threshold, can a desirable denoising effect be achieved. In this paper, Donoho wavelet soft thresholding is involved to suppress the noise of maize disease image, which addresses the drawbacks of general wavelet threshold denoising that there is a single threshold with poor self-adaptation. Donoho *et al.* [45] came up with a threshold denoising method as selected by the following equation.

$$\delta = \sigma \sqrt{2 \ln N}, \tag{11}$$

where σ indicates standard deviation of noise, N denotes the length of signal, and δ refers to the desirable threshold.

Following wavelet transform, the V component $f(x, y)$ of maize image is decomposed into a number of different scales, and the non-linear threshold at j^{th} scale is defined as follows.

$$\delta = \sigma_j \sqrt{[2 \log(j + 1)]/j} \quad j = 1, 2, 3, \cdots, N \tag{12}$$

where σ_j indicates the standard variance of noise in high frequency image of Maize at j^{th} scale, and can represent the measurement of noise intensity at this scale. Considering the multi-scale characteristics of signal and noise, the method chooses appropriate threshold to compress the wavelet coefficients on different scales in the wavelet transform domain, based on which the denoised maize disease image is obtained by inverse wavelet transform. Not only is it effective in removing the noise of maize disease image, it also retains the necessary details of the image. This process is thus believed as conducive to carrying out further analysis of the maize disease image.

2) PIECEWISE LOG TRANSFORMATION

In order to improve the contrast of maize image while making the maize lesion image clearer, the saturation image is split into four different regions in this paper, including $0 < S < 0.25$, $0.25 < S < 0.5$, $0.5 < S < 0.75$ and $0.75 < S < 1$, which can either maintain or appropriately reduce the area with high saturation, while improving the area with low saturation to enhance image saturation. A piecewise logarithmic transformation is performed to enhance the saturation component,

as shown in equation (13).

$$S(x,y) = \begin{cases} w_1 \times \log[1 + S(x,y)] & 0 < S(x,y) \leqslant 0.25 \\ w_2 \times \log[1 + S(x,y)] & 0.25 < S(x,y) \leqslant 0.5 \\ w_3 \times \log[1 + S(x,y)] & 0.5 < S(x,y) \leqslant 0.75 \\ \log[1 + S(x,y)] & 0.75 < S(x,y) \leqslant 1 \end{cases} \tag{13}$$

where w_1, w_2 and w_3 represent coefficients. $S(x, y)$ denotes the saturation component. Finally, the S component, V component and original H component are transformed to RGB space, for the purpose of subsequent disease recognition. The maize leaf lesion feature enhancement results are presented in Figure 4.

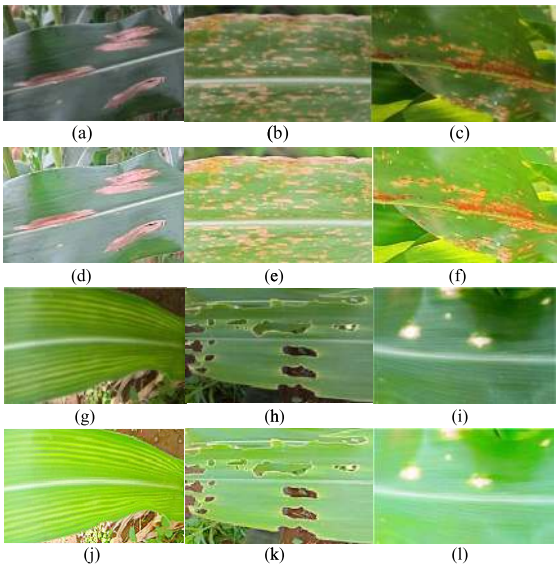

FIGURE 4. Comparison of image enhancement results with original images; (a) (d) are northern leaf blight; (b) (e) are gray leaf spot; (c) (f) are common rust; (g) (j) are zinc deficiency; (h) (k) are fall army worm; (i) (l) are round spot. where the first row are original disease images, and the second row are enhancement results.

D. DMS-ROBUST ALEXNET

For improvement to the learning effect of the network and reduction in the complexity of the network, dilated convolution and multi-scale convolution neural network based on Alexnet is introduced. The DMS-Robust Alexnet is constructed where PRelu activation function and Adabound optimizer are selected. The DMS-Robust Alexnet can be effective in improving the robustness and recognition accuracy of the network model, while reducing over-fitting.

The DMS-Robust Alexnet network structure consists of five convolution layers, along with a Multi-scale convolution module and three fully-connected layers. The first convolution layer relies on void convolution to ensure a wider range of feature extraction. Multi-scale convolution applies multi-scale convolution kernels to obtain features of different scales and reduce the costs incurred by network computation. The last convolution layer integrates the previously-extracted features, and the three fully connection layers perform a classification function.

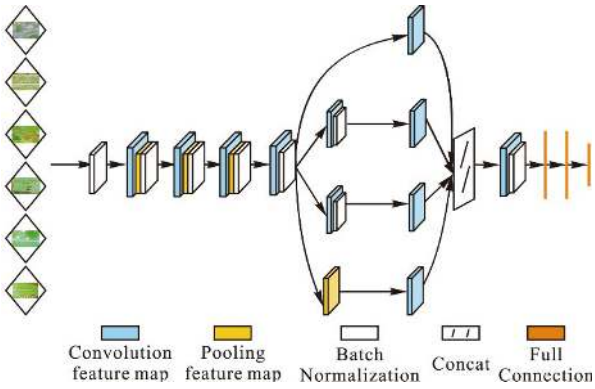

FIGURE 5. Architecture of the DMS-Robust Alexnet.

The DMS-Robust Alexnet network architecture is illustrated in Figure 5. It contains 5 convolution layers, one multi-scale convolution module and three fully connected layers. Following each convolution layer are the BN (Batch normalization) layers, which are capable to improve the generalization ability of the model, avoid over-fitting and improve the robustness of the network. In the first convolution layer, the dilated convolution is performed to derive greater receptive field. The multi-scale convolution module is set before the last convolution layer based on our summarized experience. The DMS-Robust Alexnet model is defined as follows:

1. The first layer is composed of an input layer and a BN layer. The BN layer processes the input image, which can improve the generalization ability of the model, and accelerate the convergence speed of the network.
2. The second layer is composed of convolution module 1, in which Conv1 consists of 64 filters and a kernel size of 11 × 11 pixels. Conv1 is extended by using the dilated convolution to improve the feature extraction ability of maize diseases. In addition, PReLu activation function is adopted to handle the issue of vanishing gradients. The type of pooling layer 1 is the max pooling layer, whose kernel is 3 × 3, and stride is 2. After pooling, the BN processing is carried out.
3. The third layer is composed of convolution module 2, among which Conv2 is composed of 192 filters with kernel sizes of 5 × 5 pixels, using PRelu operation. Pool layer 2 is the max pooling layer, the kernel is 3 × 3, and the stride is 2. After pooling, BN is carried out.
4. The fourth layer is composed of convolution module 3, in which Conv3 consists of 384 filters with a kernel size of 3 × 3 pixels, assisted by using PRelu operation. The type of pooling layer 3 is the max pooling layer, whose kernel is 3 × 3, and stride is 2. After pooling, BN is carried out.
5. The fifth layer is composed of convolution module 4, in which Conv4 consists of 256 filters with a kernel size

of 3 × 3 pixels. assisted by using The PRelu operation. After pooling, BN is performed.

6. The sixth layer is composed of multi-scale convolution modules, in which the number of filters (from top to bottom) of the first layer of multi-scale convolution is 96 and 16, respectively, whose kernel size is 1 × 1. The activation function is PRelu. The second layer sum of multi-scale convolution (from top to bottom) is 64, 128, 128, and 128, respectively, whose kernel size is 1 × 1, 3 × 3, 5 × 5, 1 × 1, respectively. After collection, it is processed by the concatenate layer.
7. The seventh layer is composed of convolution module 5, in which Conv5 consists of 256 filters with a kernel size of 3 × 3 pixels. As above, the PRelu operation is used, and the collection is processed by the BN layer.
8. The first fully connected layer contains 200 neurons, which is then processed by both a PReLU operation and a dropout operation.
9. The second fully connected layer contains 100 neurons which is then processed by a PReLU operation and a dropout operation.
10. The last fully connection layer contains seven neurons, standing for the number of maize leaf disease categories. The output of the last fully connection layer is then transmitted to the output layer to determine the classification of the input image. Finally, a softmax activation function is used, consequently making the sum of the output values equal to 1.0 and limiting the single output to a value between 0-1. The softmax function is an appropriate implementation into DMS-Robust Alexnet because it accounts for the relative magnitudes of all outputs. The process of DMS robustAlexnet is shown in Figure 6. Layer parameters for the DMS-Robust Alexnet are shown in Table 4.

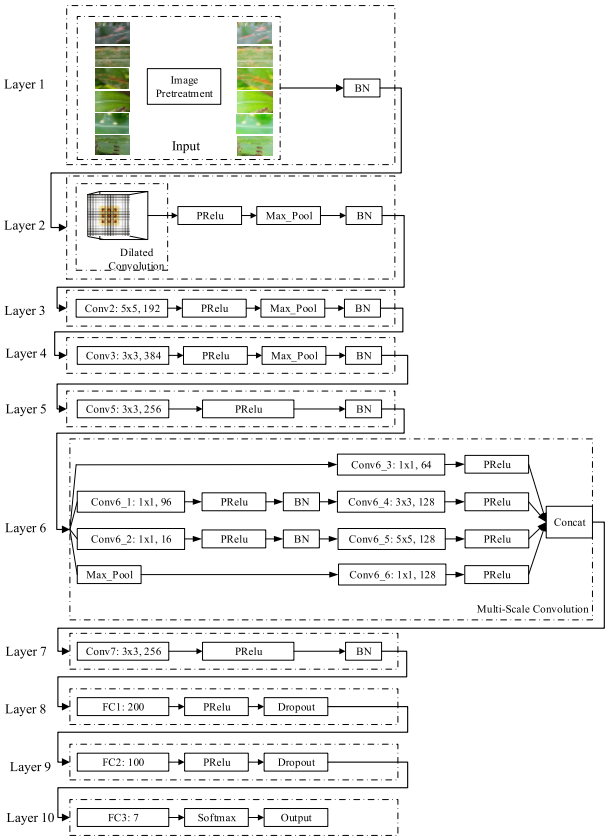

FIGURE 6. The process of the DMS-Robust Alexnet.

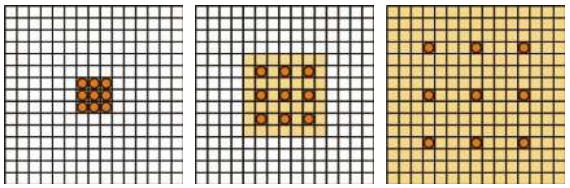

FIGURE 7. Examples of dilated convolution with different coefficients: (a) Expansion coefficient is 1; (b) Expansion coefficient is 2; (c) Expansion coefficient is 4.

1) DILATED CONVOLUTION

The features of maize diseases are changeable, and there are significant distinctions in different stages of growth. Therefore, a wider range of feature extraction is deemed necessary. In view of the above problems, this paper adopts dilated convolution in the first layer, for its capability to increase the receptive field of the model, which is conducive to enhancing the extraction ability of the model, and avoiding the significant changes in the parameters of the latter layer [46]. Dilated convolution is known as a method of data sampling on feature maps, which can increase the receptive field without any compromise on resolution and compensate for the loss of information. The receptive field is the size of the region mapped on the original image by the pixels on the feature map output from each layer of the network. The calculation of the receptive field is shown as follows.

$$r_{i+1}^2 = [(r_i - 1) + (2l + 1)]^2 \tag{14}$$

where r_i denotes the Edge length of receptive field in i^{th} layer, l denotes the expansion coefficient of dilated convolution.

The examples of dilated convolution with different coefficients are presented in Figure 7. It can be seen that with the same kernel size 3 × 3, different expansion coefficient can lead to different receptive fields. In Figure 7(a), the coefficient is 1, which shows no difference with traditional convolution. In Figure 7(b), the coefficient is 2, and the receptive field is expanded to 7 × 7. In Figure 7(c), then the coefficient is 4, and the receptive field is expanded to 15 × 15.

2) MULTI-SCALE CONVOLUTION

An improvement to the accuracy of maize disease identification is desirable. Deep learning usually improves the accuracy of recognition by increasing both network depth and network parameters. Nevertheless, only increasing the network is possible to result in over-fitting and increased computational

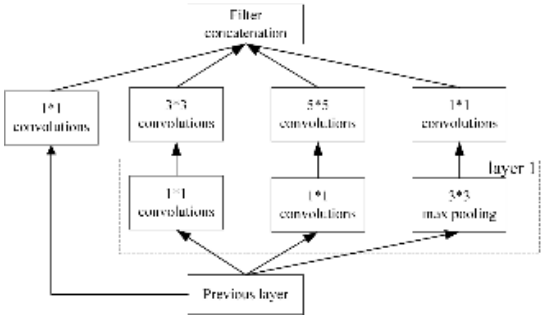

FIGURE 8. The architecture of multi-scale convolution module.

complexity. To solve this problem, Multi-scale convolution module was developed. If only convolution cores with multiple sizes are introduced, a large number of additional parameters will be introduced, thus making the model inefficient. Inspired by Network In Network [47], some 1×1 convolution kernels are used in Multi-scale convolution module.

The traditional neural network is usually a stack of convolution layers, with each layer using only one size for convolution core. In fact, the same feature map can use multiple convolution kernels of different sizes to obtain the features of different scales, before these features are combined. The resulting features are often better than if a single convolution core is used. As the size of convolution kernel of 3×3 and 5×5 is large and convolution, a lot of parameters will be involved in the convolution. Therefore, in order to reduce the parameters of 3×3, 5×5 convolution, a 1×1 convolution is added to reduce the dimension of the feature, and the max pooling is applied to capture different features.

The architecture of multi-scale convolution is illustrated in Figure 8. The main idea of the multi-scale convolution is based on determining how an optimal local sparse structure in a convolutional vision network can be approximated and covered by readily available dense components. In order to prevent patch alignment issues, current incarnations of the architecture are restricted to filter sizes 1×1, 3×3 and 5×5. It means that the suggested architecture is a combination of all those layers with their output filter banks concatenated into a single output vector providing the input for the next layer.

3) BATCH NORMALIZATION

The features of maize lesions are complex and variable. The neural network learning speed is low or even difficult to learn. Meanwhile, as the neural network structure continues to deepen, the distribution of hidden layer data has undergone significant changes and even fluctuations, which will have a negative impact on the stability of the network. On this paper, the BN algorithm, which normalizes the data of each layer to a mean of 0 and a standard deviation of 1, is applied. This is purposed to ensure data stability and makes it easier and more stable to train deep network models, while improving the capability of network generalization. The BN algorithm calculates the mean and variance of each batch samples as follows.

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{15}$$

$$\sigma = \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^2 \tag{16}$$

where μ and σ represent the mean and standard deviation of each batch samples X. It is followed by batch normalization.

$$\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}} \tag{17}$$

where ε is constant to prevent the fractal from failing in case the standard deviation is zero.

4) ACTIVATION FUNCTION

Due to the interference caused by sunlight, fog and dust, the intensity range of pixel signal in maize disease image is extremely wide. Sigmoid function, which is used in traditional networks on a frequent basis, shows a tendency to slow down the gradient change in the saturated region and cause it to approach zero gradually, thus leading to the disappearance of gradient. Relu function exhibits a faster convergence speed than Sigmoid and tanh function, and the gradient will not be saturated. It is possible to cause the phenomenon of ''gradient disappearance'' in the training of neural network. A problem is that, the gradient after a neuron is always zero and ceases to respond to any data. To resolve the problems as mentioned above, PRelu activation function rather than Sigmoid and Relu activation function is utilized in the DMS-Robust Alexnet. Its mathematical expression is shown as follows

$$PRelu\,(x) = \begin{cases} x & \text{if } x > 0 \\ ax & \text{if } x \leqslant 0 \end{cases} \tag{18}$$

where x indicates the output value of a neuron, and a denotes the hyper parameter, which is set to 0.25 in our experiments. The difference between PRelu function and Relu function is that, when the input signal is less than 0, the value of PRelu function is a function with a smaller slope, which changes the distribution of data and retains some values of the negative axis. As a consequence of that, the negative information will not be lost completely. Thus, the information loss of maize disease image is avoided.

5) OPTIMIZER OF NETWORK

At present, SGD and Adam are commonly applied as optimizers in the field of deep learning. However, in the process of training network, the convergence speed of SGD is low in the early stage. Meanwhile, SGD presents difficulty in the selection of an appropriate learning rate for the task of maize leaf disease identification. When training data is made sparse, SGD has a possibility to result in poor performance and limited training speed. Moreover, SGD is susceptible to the convergence to local optimum and could be trapped at saddle point. In addition, Momentum and Nesterov [48] optimization methods are capable to make gradient updating

more flexible. Despite this, these artificial set learning rates are difficult to operate compared with adaptive learning rates. Among the methods of adaptive learning rate, Adam has been accepted as the default algorithm for many deep learning frameworks due to its fast training speed [49]. However, due to the complex and constant-changing shooting environment of maize image and the significant difference of disease in different periods, the learning rate of Adam adaptive method is lacking in consistency.

In order to address the low convergence speed of SGD and the poor generalization ability of Adam, Adabound [50] optimizer is selected according to the characteristics of maize disease image. AdaBound is an optimizer that behaves like Adam at the start to training, and gradually transforms to SGD at the end. Adabound demonstrates apparent advantages in processing sparse data and dealing with non-linear targets. More crucially, it requires less memory. For each parameter, the self-adaptive learning rate is calculated, which also has a positive effect on non-convex optimization, large data sets and high-dimensional space.

IV. APPLICATION AND RESULTS ANALYSIS

A. EVALUATION METRICS

In order to better evaluate the performance of our model, F_1, the average precision, average recall rate and average accuracy are taken as the evaluation metrics, the details of which are shown as follows.

$$F_{1(i)} = 2P_iR_i/(P_i + R_i) \tag{19}$$

$$P_i = \mathrm{TP}_i/(\mathrm{TP}_i + \mathrm{FP}_i); \quad R_i = \mathrm{TP}_i/(\mathrm{TP}_i + \mathrm{FN}_i) \tag{20}$$

$$P = \frac{\sum_{i=1}^{n_{\mathrm{cl}}} P_i}{n_{\mathrm{cl}}}; \quad R = \frac{\sum_{i=1}^{n_{\mathrm{cl}}} R_i}{n_{\mathrm{cl}}}; \quad F_1 = \frac{\sum_{i=1}^{n_{\mathrm{cl}}} F_{1(i)}}{n_{\mathrm{cl}}};$$

$$\mathrm{AA} = \frac{1}{n_{c1}} \sum_{i=1}^{n_{c1}} \frac{n_{ii}}{n_i} \tag{21}$$

where P denotes precision, R means recall, and AA indicates the average accuracy. n_{cl} refers to the total number of instances, i represents the category index, and n_i denotes the number of the i^{th} class. n_{ii} indicates the number of accurate prediction in i^{th} class, where the first subscript i represents the class index, and the second subscript i refers to the prediction result. TP_i stands for TP in the i^{th} class. TP_i and FN_i are identical to TP_i.

B. EXPERIMENTS SETUP

The data set used in this paper consists of 7 classes, involving a total of 12227 pieces. The memory of the experimental platform is 16 GB. It is equipped with Core i7-7770K CPU @ 4.00 GHz X8 processor, Nvidia Geforce GTX 1080Ti GPU, Windows 10 64bit operating system and pytorch deep learning framework. In addition, in order to improve recognition effect, the input image is set to 256 × 256.

TABLE 4. The configuration of DMS-Robust Alexnet.

Layer	Parameter	Follow-up actions
Input	256 × 256 × 3	BN
Convolution1(Conv1)	64 convolution filters (11 × 11), 4 strides, 2 padding, 2 coefficients of expansion	PRelu+BN
Pooling1	Max pooling (3 × 3), 2 strides	-
Convolution2 (Conv2)	192 convolution filters (5 × 5), 5 strides, 2 padding	PRelu+BN
Pooling2	Max pooling (3 × 3), 2 strides	-
Convolution3 (Conv3)	384 convolution filters (3 × 3), 3 strides, 1 padding	PRelu+BN
Pooling3	Max pooling (3 × 3), 2 strides	-
Convolution4 (Conv4)	256 convolution filters (3 × 3), 3 strides, 1 padding	PRelu+BN
Multi-Scale Convolution	96 convolution filters (1 × 1), 1 stride, 1 padding	PRelu+BN
	16 convolution filters (1 × 1), 1 stride, 1 padding	PRelu+BN
	Max pooling (3 × 3), 2 strides	-
	64 convolution filters (1 × 1), 1 stride, 1 padding	PRelu
	128 convolution filters (3 × 3), 1 stride, 1 padding	PRelu
	128 convolution filters (5 × 5), 1 stride,1 padding	PRelu
	128 convolution filters (1 × 1), 1 stride, 1 padding	PRelu
	448 concat	-
Convolution5 (Conv5)	256 convolution filters (3 × 3), 3 strides, 1 padding	PRelu+BN
Fully Connect1 (fc1)	200 nodes	PRelu+Drop
Fully Connect2 (fc2)	100 nodes	PRelu+Drop
Fully Connect3 (fc3)	7 nodes	PRelu+Drop
Output	1 node	Softmax

The weights of all layers are initialized by the Gauss distribution in the first place. With the hardware performance and training time taken into account, the number of batch samples for validation and training is set to 64, and the initial learning rate is 0.001. Each 200 epochs serve as a training session. The configuration of DMS-Robust Alexnet is shown in Table 4.

C. PERFORMANCE AND ANALYSIS

In this section, the experimental results obtained by applying our method are introduced in detail, and the performance under different experimental conditions is analyzed. The main work is as follows. Firstly, the improved DMS-Robust Alexnet model is compared with the original Alexnet model. Secondly, experiments on the effects of image enhancement algorithms on model performance. Thirdly, an experiment is conducted to validate the effects of the choice of activation function and optimization method. Finally, a comparison is performed with other methods used in the same field.

1) COMPARISON DMS-ROBUST ALEXNET WITH ORIGINAL ALEXNET

To enhance the contrast of experiments, two indexes of loss process for the two models. Both model inputs are processed by using the improved image enhancement method. Expanded data sets with a scale of 8:2 is trained and validated.

The results of validation after each training iteration using Alexnet and the DMS-Robust Alexnet models are compared in Figure 9.

From Figure 9(a) and Figure 9(c), when accuracy shows a tendency to converge, Alexnet's recognition accuracy reaches 91.83%. The recognition accuracy of DMS-Robust Alexnet reaches 98.62%. Under the identical conditions, the DMS-Robust Alexnet is clearly superior to Alexnet. It can be seen from Figure 9 that Alexnet model achieves the accuracy of regional convergence in the 60^{th} iteration. However, the accuracy still experiences significant fluctuations in the following iterations, which indicates that Alexnet is inconsistent in training. In addition, Alexnet model requires more iterations to achieve ideal recognition accuracy. With the DMS-Robust Alexnet model concerned, as the number of iterations is on the rise, the model is continuous and quick to reach a higher recognition accuracy. In the 60^{th} iteration, the recognition rate of regional convergence is reached, and the fluctuation is gradually reduced, suggesting that the network structure is capable to converge in a quicker and smoother way.

From Figure 9(b) and Figure 9(d), it can be seen that the minimum loss of Alexnet model is approximately 0.20. In the training process, however, the loss will suddenly increase in some cases, such as in the 46^{th} and 92^{th} iterations. The loss continues to fluctuate considerably as training progresses. Finally, the training is terminated after the 200^{th} iteration is complete. The DMS-Robust Alexnet model achieves a minimum loss of about 0.02. As the process of training proceeds, the loss is reduced gradually, the fluctuation gradually

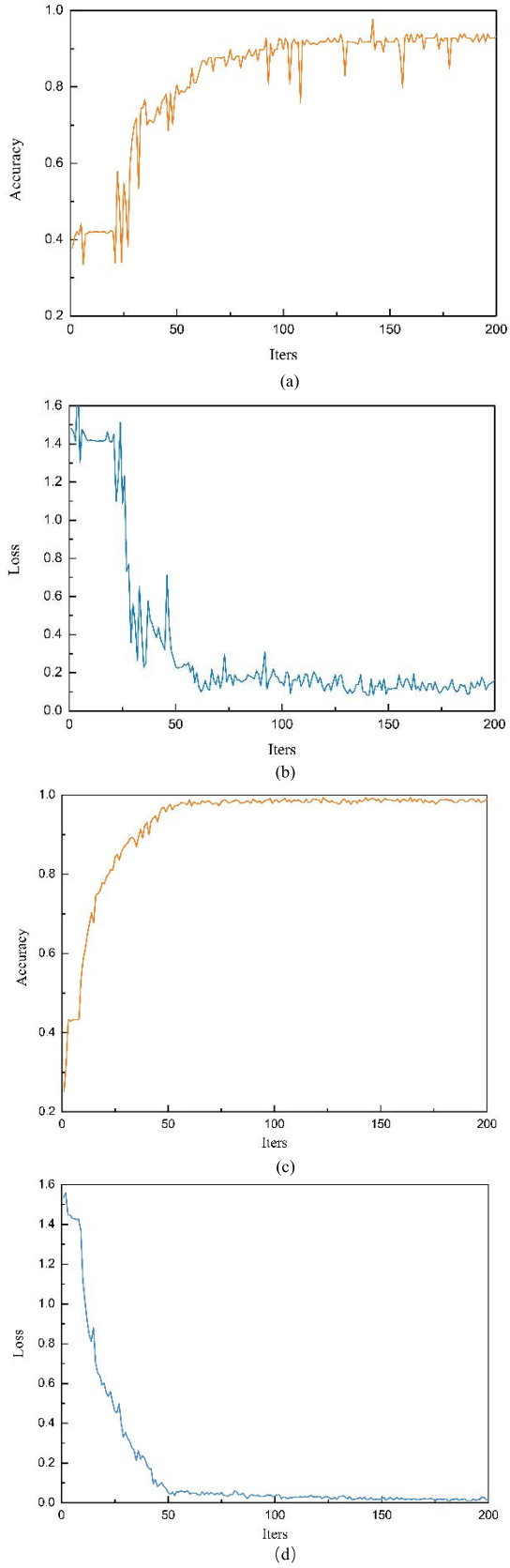

FIGURE 9. Comparison between Alexnet and DMS-Robust Alexnet: (a) Accuracy of Alexnet; (b) Mean loss of Alexnet; (c) Accuracy of DMS-Robust Alexnet; (d) Mean loss of DMS-Robust Alexnet.

stabilizes, and the training is also terminated after 200 iterations, which demonstrates that DMS-Robust Alexnet has faster convergence speed and more robustness than Alexnet model.

Generally speaking, compared with Alexnet, not only does the DMS-Robust Alexnet achieve higher recognition accuracy and consistent learning process, it also shows faster convergence and more robustness under the identical training conditions.

From Table 5, it can be seen that the average precision, recall and value of the DMS-Robust Alexnet for seven types of maize images are higher than those of the original Alexnet model, and the average F1 value is 6 percentage points higher than that of Alexnet model, with a maximum of 98.80%. The experimental results are considered as compliant with the practical application requirements of effective identification of maize diseases. As demonstrated by the results, the DMS-Robust Alexnet can be adequate to extract the abstract features, and more capable of feature extraction and generalization.

TABLE 5. Comparison of different network iterations and training time.

Model	Categories	Average precision(%)	Average recall(%)	F_1 (%)
Alexnet	Common rust	92.72	94.87	93.78
	Northern leaf blight	93.51	93.46	93.48
	Gray leaf spot	94.39	94.28	94.33
	Zinc deficiency	87.62	88.95	88.27
	Round spot	89.72	90.34	90.02
	Fall army worm	88.26	89.54	88.90
	Healthy	95.31	94.27	94.78
DMS-Robust Alexnet	Common rust	98.95	98.47	98.67
	Northern leaf blight	98.03	98.00	98.01
	Gray leaf spot	99.16	98.21	98.68
	Zinc deficiency	98.24	97.17	97.70
	Round spot	96.78	97.52	97.15
	Fall army worm	97.41	95.12	96.25
	Healthy	99.20	98.44	98.80

2) EFFECT OF IMAGE ENHANCEMENT ALGORITHMS ON MODEL PERFORMANCE

In order to analyze the effects of the proposed image enhancement algorithm, in this experiment, 80% standard datasets are used to conduct DMS-Robust Alexnet training with enhanced and no enhancement experiments respectively, and it is validated with 20% standard datasets. The result is presented in Figure 10.

In Figure 10, the recognition ability without image enhancement is only capable to reach 96.36%. When the conditions are identical, after applying the improved enhancement algorithm, the highest accuracy can increase to 99.12%, and the lowest accuracy can rise to 97.82%. Seven kinds of maize images are recognized by the enhanced algorithm,

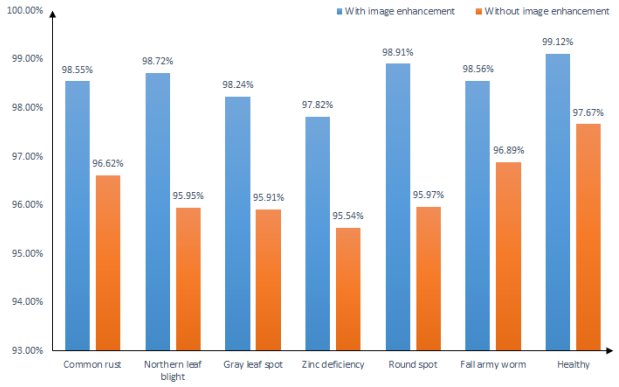

FIGURE 10. The effect of image enhancement.

and the accuracy is 2.0, 2.9, 2.4, 2.4, 3.0, 1.7 and 1.5 percentage points higher. The recognition accuracy is improved significantly by applying the enhancement algorithm, which demonstrates that the model will accumulate noise in the process of learning due to the limitation of its anti-jamming ability. When the accumulation reaches a certain level, it will have a significant impact on the recognition results. Therefore, our image enhancement algorithm is validated effectively in improving recognition accuracy in practice.

3) EFFECTS OF OPTIMIZER AND ACTIVATION FUNCTION

Normally, the original Alexnet model uses Relu function or Sigmoid function and the optimizer uses Adam. Considering the complexity of the experiment, this paper chooses different optimizers (Adam, Adabound) and different activation functions (Relu, Sigmoid, PRelu) to validate the performance of DMS-Robust Alexnet.

Schemas with different optimizer and activation function can be viewed in Table 6. To validate the effect of optimizer selection and unify activation function, the comparison schemes include 1 - 5, 3 - 4 and 2 - 6. From the results of loss and F_1, it can be seen that although the F_1 of scheme 1 is 0.53 percentage points higher than that of scheme 5, the loss remains higher than that of scheme 5. Other schemes 4 and 6 using Adabound are superior to those using Adam. In addition, the influence of activation function is compared to unify the optimizer. The comparison scheme includes 1-2-3 and 4-5-6. It can be seen from the table that the loss of scheme 3 using PRelu function is reduced as compared to that of

TABLE 6. The configuration of DMS-Robust Alexnet.

Schemes	Parameters		Loss	F_1(%)
	Optimizer	Activation Function		
Scheme 1	Adam	Relu	0.041	97.14
Scheme 2	Adam	Sigmoid	0.037	96.35
Scheme 3	Adam	PRelu	0.027	97.47
Scheme 4	AdaBound	PRelu	**0.012**	**97.89**
Scheme 5	AdaBound	Relu	0.035	96.63
Scheme 6	AdaBound	Sigmoid	0.029	96.71

scheme 1 and 2. The F_1 and loss of scheme 4 are better than those of scheme 5 and 6. Compared with other schemes, the loss of scheme 4 is closer to 0, and F_1 reaches 97.89%, indicating excellent performance. The performance of the model is improved substantially. The model in this paper finally decides to take Adabound as the optimizer and PRelu as the activation function.

4) COMPARISON WITH BASELINE METHODS

For further analysis of the performance of DMS-Robust Alexnet, a comparison is performed with GA-SVM [51], SEG-KNN [20], SIMPLE-CNN [33], VGGNet [40], GoogleNet [52] and ResNet [53], maize disease recognition baselines. The comparison results are indicated in Table 7. SEG-KNN employs a binarization method to segment the image, and then extracts the color features, shape features and texture features as the input of the KNN classifier to identify the lesions of maize leaves. GA-SVM extracted 20-dimensional maize leaf statistical characteristics, and the genetic algorithm was applied to obtain penalty factor and kernel function type of SVM. As revealed by the experimental results, the improved SVM using genetic algorithm is superior to the ordinary SVM. SIMPLE-CNN constructs two convolution layers and two fully connected layers to identify maize leaf lesions. The structure is simple and easy to train, which makes it suitable for small data sets. VGGNet improves Alexnet by replacing large kernel-sized filters (11 and 5 in the first and second convolutional layer, respectively) with multiple 3×3 kernel-sized filters one after another. In this paper, the VGG16 network model is used for comparison. GoogleNet has 9 inception modules stacked linearly, with a global average pooling at the end of the last inception module which is 22 layers deep. To deal with the vanishing gradient issue for very deep architecture, ResNet is proposed, which utilizes the identity shortcut connection to skip one or more layers. In this paper, considering the size of our maize leaf disease dataset, ResNet50 is adopted. SEG-KNN and GA-SVM are involved to identify maize leaf lesions through a combination of traditional feature extraction and classifier. The proposed method and other method adopt the deep convolution neural network model.

TABLE 7. Comparison with baseline methods.

Methods	Accuracy (%)
SEG-KNN	80.12
GA-SVM	82.71
SIMPLE-CNN	87.46
ResNet50	95.47
GoogleNet	94.06
VGGNet	93.98
Our proposed method	**98.62**

From the result, it can be seen that the deep neural network model is capable of achieving higher recognition accuracy than traditional classifiers (SEG-KNN, GA-SVM). SIMPLE-CNN has higher accuracy than SEG-KNN and GA-SVM but has a lower accuracy than other deep models. For ResNet50, the accuracy is 95.47%, with 3.29% lower than the proposed method. For GoogleNet, the result is 94.06%, which is a bit lower than ResNet50. For VGG16, it can only reach 93.98%. Specifically, proposed the DMS-Robust Alexnet relies on dilated convolution and multi-scale convolution to better extract the complex and changeable lesions. Finally, 98.62% of accuracy is achieved in the experiment. The DMS-Robust Alexnet is thus clearly superior to other baselines. We analyze why our proposed DMS-Robust Alexnet is better than other deep models for the following reasons: 1). Because of our maize leaf disease dataset only with 12227 samples, compared with other large-scale open datasets, which is very small. ResNet, GoogleNet, and VGGNet are designed for the large-scale dataset. The DMS-Robust Alexnet is a well-designed structure for our dataset. 2). DMS-Robust Alexnet adopts dilated convolution in the first layer to increase the receptive field of the model, thus enhancing the extraction ability of the model and avoiding the significant changes in the parameters of the latter layer. In addition, the introduction of Multi-Scale Convolution extraction features can more accurately characterize different diseases, which helps improve the accuracy of disease recognition at different stages of diseases 3). the proposed model is combined with a designed maize leaf feature enhancement framework.

V. CONCLUSION

As computer technology and machine vision technology advance, experts and scholars both at home and abroad have conducted extensive research into image analysis technology. In the most recent years, the leaf disease recognition based on deep learning has been met with increasingly widespread applications in the field of crop disease detection. In order for improved accuracy and effectiveness in image recognition for maize leaf disease, a method devised based on image enhancement and DMS-Robust Alexnet is proposed in this paper to identify maize leaf diseases. This method demonstrates the capability to identify and distinguish between healthy leaves and six different diseases of maize leaves. Firstly, the characteristics of maize leaf diseases are enhanced. Then, DMS-Robust Alexnet is constructed for the purpose of recognition and classification, with the accuracy of recognition reaching as high as 98.62%. As revealed by the experimental results, the suggested method eliminates the need to select the specific features. Therefore, it is considered as suited to the identification of maize leaf disease images. More crucially, it is capable to achieve the desired accuracy and effect, and the model after training is easy to use. Therefore, it opens up a new possibility to prevent and control crop diseases.

In the future, more types of maize pests and diseases will be identified, Meanwhile, in order to assist agricultural practitioners in making a quick and reasonable judgment of the information on crop disease, this method will be extended to the cloud, and obtain crop images through UAV cameras, with the aim to facilitate a rapid and reasonable judgment of crop disease information.

ACKNOWLEDGMENT

The authors would like to thank Gao Yuda's help and support.

REFERENCES

[1] R. D. L. Pires, D. N. Gonçalves, J. P. M. Oruê, W. E. S. Kanashiro, J. F. Rodrigues, B. B. Machado, and W. N. Gonçalves, "Local descriptors for soybean disease recognition," *Comput. Electron. Agricult.*, vol. 125, pp. 48–55, Jul. 2016, doi: 10.1016/j.compag.2016.04.032.

[2] R. D. Nayagam, "Implementation of external defects detection system to classify the fruits," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 4, no. 2, 2016, Art. no. 1850003. [Online]. Available: http://www.ijircce.com/upload/2016/spl2/15_deivanayagam%20journal_NEW.pdf.

[3] D. Puerto, D. Gila, J. García, and J. Ortega, "Sorting olive batches for the milling process using image processing," *Sensors*, vol. 15, no. 7, pp. 15738–15754, Jul. 2015, doi: 10.3390/s150715738.

[4] I. H. Yano, W. E. Santiago, J. R. Alves, L. T. M. Mota, and B. Teruel, "Choosing classifier for weed identification in sugarcane fields through images taken by UAV," *Bulg. J. Agricult. Sci.*, vol. 23, no. 3, pp. 491–497, 2017. [Online]. Available: https://www.agrojournal.org/23/03-21.pdf

[5] S. Sabzi, Y. Abbaspour-Gilandeh, and G. García-Mateos, "A fast and accurate expert system for weed identification in potato crops using Metaheuristic algorithms," *Comput. Ind.*, vol. 98, pp. 80–89, Jun. 2018, doi: 10.1016/j.compind.2018.03.001.

[6] J. Tang, D. Wang, Z. Zhang, L. He, J. Xin, and Y. Xu, "Weed identification based on K-means feature learning combined with convolutional neural network," *Comput. Electron. Agricult.*, vol. 135, pp. 63–70, Apr. 2017, doi: 10.1016/j.compag.2017.01.001.

[7] D. Majumdar, D. K. Kole, A. Chakraborty, and D. D. Majumder, "An integrated digital image analysis system for detection, recognition and diagnosis of disease in wheat leaves," in *Proc. 3rd Int. Symp. Women Comput. Inform.*, vol. 15, Aug. 2015, pp. 400–405, doi: 10.1145/2791405.2791474.

[8] S. Zhang, X. Wu, Z. You, and L. Zhang, "Leaf image based cucumber disease recognition using sparse representation classification," *Comput. Electron. Agricult.*, vol. 134, pp. 135–141, Mar. 2017, doi: 10.1016/j.compag.2017.01.014.

[9] K. Du, Z. Sun, Y. Li, F. Zheng, J. Chu, and Y. Su, "Diagnostic model for wheat leaf conditions using image features and a support vector machine," *Trans. ASABE*, vol. 59, no. 5, pp. 1041–1052, 2016, doi: 10.13031/trans.59.11434.

[10] H. Z. Wang, Y. W. Dong, and T. Meng, "Epidemic dynamics of maize main leaf spot disease in Hegang region," *Heilongjiang Agricult. Sci.*, vol. 23, pp. 60–63, Sep. 2015, doi: 10.11942/j.issn1002-2767.2015.09.0060.

[11] M. W. Harding, K. Jindal, J. T. Tambong, F. Daayf, R. J. Howard, H. Derksen, L. M. Reid, A. U. Tenuta, and J. Feng, "Goss's bacterial wilt and leaf blight of corn in Canada–disease update," *Can. J. Plant Pathol.*, vol. 40, no. 4, pp. 471–480, Oct. 2018, doi: 10.1080/07060661.2018.1506502.

[12] D. K. Berger, "Grey leaf spot disease of maize and food security research," *South Afr. J. Botany*, vol. 100, no. 109, p. 327, Jan. 2017, doi: 10.1016/j.sajb.2017.01.028.

[13] A. Khan, A. Sohail, U. Zahoora, and A. Saeed Qureshi, "A survey of the recent architectures of deep convolutional neural networks," 2019, *arXiv:1901.06032*. [Online]. Available: http://arxiv.org/abs/1901.06032

[14] S. D. Khirade and A. B. Patil, "Plant disease detection using image processing," in *Proc. Int. Conf. Comput. Commun. Control Autom.*, Feb. 2015, pp. 768–771, doi: 10.1109/ICCUBEA.2015.153.

[15] C.-L. Chung, K.-J. Huang, S.-Y. Chen, M.-H. Lai, Y.-C. Chen, and Y.-F. Kuo, "Detecting Bakanae disease in rice seedlings by machine vision," *Comput. Electron. Agricult.*, vol. 121, pp. 404–411, Feb. 2016, doi: 10.1016/j.compag.2016.01.008.

[16] W.-J. Liang, H. Zhang, G.-F. Zhang, and H.-X. Cao, "Rice blast disease recognition using a deep convolutional neural network," *Sci. Rep.*, vol. 9, no. 1, pp. 2045–2322, Dec. 2019, doi: 10.1038/s41598-019-38966-0.

[17] J. C. Neto, G. E. Meyer, D. D. Jones, and A. K. Samal, "Plant species identification using elliptic Fourier leaf shape analysis," *Comput. Electron. Agricult.*, vol. 50, no. 2, pp. 121–134, Feb. 2006, doi: 10.1016/j.compag.2005.09.004.

[18] F. Ikorasaki and M. B. Akbar, "Detecting corn plant disease with expert system using Bayes theorem method," in *Proc. 6th Int. Conf. Cyber IT Service Manage. (CITSM)*, Aug. 2018, pp. 1–3, doi: 10.1109/CITSM.2018.8674303.

[19] L. Xu and J. Lv, "Recognition method for apple fruit based on SUSAN and PCNN," *Multimedia Tools Appl.*, vol. 77, no. 6, pp. 7205–7219, Mar. 2018, doi: 10.1007/s11042-017-4629-6.

[20] S. W. Zhang and Y. J. Shang, and L. Wang, "Plant disease recognition based on plant leaf image," *J. Animal Plant Sci.*, vol. 25, no. 3, pp. 42–45, 2015. [Online]. Available: http://www.thejaps.org.pk/docs/Supplementary/v-25/07.pdf

[21] L. N. Zhang and B. Yang, "Research on recognition of maize disease based on mobile Internet and support vector machine technique," *Adv. Mater. Research.*, vol. 905, pp. 659–662, Apr. 2014, doi: 10.4028/Www.scientific.net/AMR.905.659.

[22] K. R. Aravind, P. Raja, K. V. Mukesh, R. Aniirudh, R. Ashiwin, and C. Szczepanski, "Disease classification in maize crop using bag of features and multiclass support vector machine," in *Proc. 2nd Int. Conf. Inventive Syst. Control (ICISC)*, Jan. 2018, pp. 1191–1196, doi: 10.1109/ICISC.2018.8398993.

[23] B. Alipanahi, A. Delong, M. T. Weirauch, and B. J. Frey, "Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning," *Nature Biotechnol.*, vol. 33, no. 8, pp. 831–838, Aug. 2015, doi: 10.1038/nbt.3300.

[24] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016, doi: 10.1038/nature16961.

[25] X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, "Identification of maize leaf diseases using improved deep convolutional neural networks," *IEEE Access*, vol. 6, pp. 30370–30377, Jun. 2018, doi: 10.1109/ACCESS.2018.2844405.

[26] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[27] A. dos Santos Ferreira, D. Matte Freitas, G. Gonçalves da Silva, H. Pistori, and M. Theophilo Folhes, "Weed detection in soybean crops using ConvNets," *Comput. Electron. Agricult.*, vol. 143, pp. 314–324, Dec. 2017, doi: 10.1016/j.compag.2017.10.027.

[28] N. Zeng, Z. Wang, B. Zineddin, Y. Li, M. Du, L. Xiao, X. Liu, and T. Young, "Image-based quantitative analysis of gold immunochromatographic strip via cellular neural network approach," *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1129–1136, May 2014, doi: 10.1109/TMI.2014.2305394.

[29] W. Ding and G. Taylor, "Automatic moth detection from trap images for pest management," *Comput. Electron. Agricult.*, vol. 123, pp. 17–28, Apr. 2016, doi: 10.1016/j.compag.2016.02.003.

[30] Y. Wen, K. Zhang, Zh. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 499–515, doi: 10.1007/978-3-319-46478-7_31.

[31] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[32] C. DeChant, T. Wiesner-Hanks, S. Chen, E. L. Stewart, J. Yosinski, M. A. Gore, R. J. Nelson, and H. Lipson, "Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning," *Phytopathology*, vol. 107, no. 11, pp. 1426–1432, Nov. 2017, doi: 10.1094/PHYTO-11-16-0417-R.

[33] M. Sibiya and M. Sumbwanyambe, "A computational procedure for the recognition and classification of maize leaf diseases out of healthy leaves using convolutional neural networks," *AgriEngineering*, vol. 1, no. 1, pp. 119–131, Mar. 2019, doi: 10.3390/agriengineering1010009.

[34] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep learning for tomato diseases: Classification and symptoms visualization," *Appl. Artif. Intell.*, vol. 31, no. 4, pp. 299–315, Apr. 2017, doi: 10.1080/08839514.2017.1315516.
[35] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–11, Jun. 2016, doi: 10.1155/2016/3289801.
[36] R. A. Priyadharshini, S. Arivazhagan, M. Arun, and A. Mirnalini, "Maize leaf disease classification using deep convolutional neural networks," *Neural Comput. Appl.*, vol. 31, no. 12, pp. 8887–8895, Dec. 2019, doi: 10.1007/s00521-019-04228-3.
[37] Z. Lin, S. Mu, A. Shi, C. Pang, and X. Sun, "A novel method of maize leaf disease image identification based on a multichannel convolutional neural network," *Trans. ASABE*, vol. 61, no. 5, pp. 1461–1474, 2018, doi: 10.13031/trans.12440.
[38] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.
[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
[40] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: http://arxiv.org/abs/1409.1556
[41] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105, doi: 10.1145/3065386.
[42] F. J. Pulgar, A. J. River, F. Charte, and M. J. del Jesus, "On the impact of imbalanced data in convolutional neural networks performance," in *Proc. Int. Conf. Hybrid Artif. Intell. Syst.* Cham, Switzerland: Springer, 2017, pp. 220–232, doi: 10.1007/978-3-319-59650-1_19.
[43] M. Elad, "Retinex by two bilateral filters," in *Proc. Int. Conf. Scale-Space Theories Comput. Vis.* Berlin, Germany: Springer, 2005, pp. 229–2005, doi: 10.1007/11408031_19.
[44] R. E. Masumdar and R. G. Karandikar, "Comparative study of different wavelet transforms in fusion of multimodal medical images," *Int. J. Comput. Appl.*, vol. 146, no. 11, Jul. 2016. [Online]. Available: https://www.semanticscholar.org/paper/Comparative-Study-of-Different-Wavelet-Transforms-Masumdar-Karandikar/ddf1c7569c8084a37cc7e5186e8f899e08367603
[45] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Sep. 1994, doi: 10.1093/biomet/81.3.425.
[46] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: https://arxiv.org/abs/1502.03167
[47] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, *arXiv:1312.4400*. [Online]. Available: http://arxiv.org/abs/1312.4400
[48] A. Botev, G. Lever, and D. Barber, "Nesterov's accelerated gradient and momentum as approximations to regularised update descent," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 1899–1903, doi: 10.1109/IJCNN.2017.7966082.
[49] A. C. Wilson, R. Roelofs, M. Stern, N. Srebro, and B. Recht, "The marginal value of adaptive gradient methods in machine learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 4148–4158. [Online]. Available: https://www.papers.nips.cc/paper/7003
[50] L. Luo, Y. Xiong, Y. Liu, and X. Sun, "Adaptive gradient methods with dynamic bound of learning rate," 2019, *arXiv:1902.09843*. [Online]. Available: http://arxiv.org/abs/1902.09843
[51] Z. Y. Zhang, X. Y. He, X. H. Sun, L. M. Guo, J. Wang, F. Wang, "Image recognition of maize leaf disease based on GA-SVM," *Chem. Eng. Trans.*, vol. 46, pp. 199–204, Dec. 2015. [Online]. Available: https://www.semanticscholar.org/paper/Image-Recognition-of-Maize-Leaf-Disease-Based-on-Zhang-He/abf6d7438d0904a829b5e7c63e3be7a287090b13
[52] W.-S. Jeon and S.-Y. Rhee, "Plant leaf recognition using a convolution neural network," *Int. J. Fuzzy Log. Intell. Syst.*, vol. 17, no. 1, pp. 26–34, Mar. 2017, doi: 10.5391/IJFIS.2017.17.1.26.
[53] P. Ouppaphan, "Corn disease identification from leaf images using convolutional neural networks," in *Proc. 21st Int. Comput. Sci. Eng. Conf. (ICSEC)*, Nov. 2017, pp. 1–5, doi: 10.1109/ICSEC.2017.8443919.

MINGJIE LV graduated from the Central South University of Forestry and Technology, in 2017. He is mainly engaged in the study of deep learning and image processing.

GUOXIONG ZHOU received the B.Sc. degree from Hunan Agricultural University, in 2002, and the M.Sc. and Ph.D. degrees from Central South University, in 2006 and 2010, respectively. He is currently an Associate Professor with the Central South University of Forestry and Technology. His main research interests include forest fire prevention and robot.

MINGFANG HE was born in 1985. She received the Ph.D. degree. Her main research interests include control science and engineering, and artificial intelligence.

AIBIN CHEN was born in Min County, Hunan, in 1971. He received the Ph.D. degree. He was a Professor. His research interest includes image processing.

WENZHUO ZHANG was born in 1996. He received the bachelor's degree from the Central South University of Forestry and Technology, in 2018. His main research interests include direction of forest fire prevention and graphic image processing.

YAHUI HU was born in 1979. He received the Ph.D. degree. His main research interests include control of biological diseases and insect pests.