Maize Disease Recognition Based On Image Enhancement And OSCRNet

Hongji Zhang
Central South University of Forestry and Technology

Zhou Guoxiong (zhougx01@163.com)
Central South University of Forestry and Technology

Aibin Chen
Central South University of Forestry and Technology

Jiayong Li
Central South University of Forestry and Technology

Mingxuan Li
Central South University of Forestry and Technology

Wenzhuo Zhang
Central South University of Forestry and Technology

Yahui Hu
Plant protection research institute, Academy of agricultural sciences,

Wentao Yu
University of victoria

Research Article

Keywords: octave convolution, deep residual network, Self-Calibrated convolution, AReLU activation function, maize leaf pests detection, OSCRNet.

Posted Date: September 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-871678/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Maize disease recognition based on image enhancement and OSCRNet

Hongji Zhang\textsuperscript{a}, Guoxiong Zhou\textsuperscript{a,*}, Aibin Chen\textsuperscript{a}, Jiayong Li\textsuperscript{a}, Mingxuan Li\textsuperscript{a}, Wenzhuo Zhang\textsuperscript{a}, Y ahui Hu\textsuperscript{b}, Wentao Yu\textsuperscript{c}

\textsuperscript{a}College of Computer & Information Engineering, Central South University of Forestry and Technology, Changsha, 410004, Hunan, China.
\textsuperscript{b}Plant protection research institute, Academy of agricultural sciences, 410125, Changsha, Hunan, China
\textsuperscript{c}department of electrical & computer engineering, University of victoria, V8P3C8, Victoria, Canada,

* Corresponding author: Guoxiong Zhou (zhougx01@163.com)
E-mail address: 962051393@qq.com (H. Zhang), zhougx01@163.com (G. Zhou), hotaibin@163.com (A. Chen), 892889422@qq.com (J. Li), 173523919@qq.com (M. Li), 2515322296@qq.com (W. Zhang), huy ah627@163.com (Y. Hu), wentaoyu@gmail.com (W. Yu), hotaibin@163.com; kingkong148@163.com).

Abstract

Background
Under natural light irradiation, there are significant challenges in the identification of maize leaf diseases because of the difficulties in extracting lesion features from constantly changing environments, uneven illumination reflection of the incident light source and many other factors.

Results
In the present paper, a novel maize image recognition method was proposed. Firstly, an image enhancement framework of the maize leaf was designed, and a multi-scale image enhancement algorithm with color restoration was established to enhance the characteristics of the maize leaf in a complex environment and to solve the problems of high noise and blur of maize images. Subsequently, an OSCRNet maize leaf recognition network model based on the traditional ResNet backbone architecture was designed. In the OSCRNet maize leaf recognition network model, an octave convolution with characteristics to accelerate network training was adopted, reducing unnecessary redundant spatial information in the maize leaf images. Additionally, a self-calibrated convolution with multi-scale features was employed to realize the interactions of different feature information in the maize leaf images, enhance feature extraction, and solve the problems of similarity of maize disease features and easy learning disorders. Concurrently, batch normalization was employed to prevent network overfitting and enhance the robustness of the model. The experiment was conducted on the maize leaf image data set. The highest identification accuracy of rust, grey leaf disease, northern fusarium wilt, and healthy maize was 94.67%, 92.34%, 89.31% and 96.63%, respectively.

Conclusions
The aforementioned methods were beneficial in solving the problems of slow efficiency, low accuracy and image recognition training, and also outperformed other comparison models. The present 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.

Keywords: octave convolution, deep residual network, Self-Calibrated convolution, APRelu activation function, maize leaf pests detection, OSCRNet.

Introduction

Maize is a globally significant crop, serving an irreplaceable role in food, feed, medicine, fuel, industrial, and other important fields, with the planting area in China being second only to rice. However, the yield of maize often suffers due to pesticides, changes in the external environment, maize plant diseases, and insect pests. Common maize plant diseases and insect pests include: maize borer, rust, grey leaf disease, northern fusarium wilt, slime worm, maize tetranychus mite, and others. Consequently, formulating a method of prevention and control of plant diseases and insect pests of the maize leaf will effectively improve maize yield. However, there are various kinds of plant diseases and insect pests of the maize leaf, with many species being similar. Inexperienced farmers are prone to misjudgment, leading to poor governance, and time-consuming and artificial recognition, with the external environment differing in regions accommodating different maize varieties. Different planting methods and other influencing factors increase the difficulty of identifying maize leaf diseases and insect pests from the characteristics thereof. Currently, with the development of deep learning and computer technology, image recognition technology has been increasingly applied in order to detect and improve crop yield [1][2]. Extensive research has also been conducted on weed control [3][4], identification and diagnosis of crop diseases and insect pests [5], [6] and other fields. In recent years, said research has become a focal point in the field of modern agriculture, and significant results have been reported.

Presently, with the development and achievements of deep learning [7] and image processing [8] [9], the convolutional neural network (CNN) [10] structure has been extensively employed in the identification of agricultural pests and diseases in terms of feature extraction and feature recognition [11][12][13]. CNN structures can automatically extract features from data by inputting the original data of images of crop diseases and pests into the network; however, the problems mentioned in the present paper, such as identification of maize leaf diseases and pests, may not necessarily be solved [14]. The ResNet neural network [22] proposed by He Keming et al. is a significant network structure model and possesses good performance in the image classification field. However, the structures of said models, including ResNet network, are complex, and the training times are long, requiring a substantial amount of data for training. Additionally, due to large noise, plant diseases, and insect pests, maize leaf images are often unclear and difficult to obtain. Based on the aforementioned traditional convolutional neural networks, including ResNet, there is substantial difficulty in achieving significant results with less data, and training
is not always possible. The identification of maize leaf pests and diseases can be bisected into maize leaf image enhancement and maize leaf image recognition. Due to external environmental factors, such as light and dust, and other influencing factors, the noise of the original maize leaf image collected from the actual agricultural environment is large, the image is blurred, and the maize leaf image presents a variety of bad feature information, resulting in a reduced dynamic recognition rate [25]. Image processing of maize leaf pests and diseases has become a crucial but challenging task. Consequently, the original image must be denoised and the image must be enhanced. Currently, Retinex [29] is widely employed in image enhancement. The Retinex algorithm is comprised of the incident image and reflection image. Simultaneously, the same logarithm operation is conducted, and then the Gaussian filter is passed to achieve the final enhancement result. However, not only does image distortion easily occur, losing part of the feature information, but the algorithm also suffers low efficiency and substantial amounts of calculation, increasing the difficulty of learning the important feature information of the network. As a result, scholars have enhanced Retinex’s algorithm to utilize MSRCR’s method of image enhancement, which has the characteristics of multi-scale analysis - a new analysis technology - and is a more efficient method for image processing, achieving significant denoising effects. During image recognition, CNN can extract more abstract feature information layer by layer through images. Typical CNN’s include Alexnet [16], VGG [15], ResNet [22], and others, which have all experienced significant progress. A vector machine was utilized to detect and classify bananas [33] and reached a recognition rate of 89.63%; and color formation was employed to classify and grade banana shapes [34], with the recognition rate of the system peaking at 95%. With regard to the identification of maize leaf diseases and insect pests, in April 2020, MAO Yandong et al. conducted research on maize disease recognition based on SVM and DS theory integrating multiple features [37] and reached a final average accuracy of 93.33%. In October 2020, Yang Mingxin et al. employed the improved transfer learning network to identify maize [36], reaching a recognition rate of 97.23%. Based on the optimized convolutional neural network [38], Li Jing et al. recognized maize borers disease through the Inception V4 network of GoogLeNet in April 2020. Additionally, a multi-scale convolutional module was employed for optimization, reaching an average recognition accuracy of 96.44%. In January 2021, Liu Aoyu et al. utilized a residual network to identify maize diseases [35], and successfully implemented the ResNet model into the TEL-Resnet network. In May 2021, Wang Can et al. proposed the use of a dual-attention semantic segmentation network for maize recognition [39], with the model’s average cross-parallel ratio and average pixel recognition accuracy reaching 94.16% and 95.68%, respectively. Although said studies contributed positively, the maize leaves were characterized by large noise, blurred images, more redundant spatial information, and similar disease features, resulting in long network training times and network learning disorders, which impeded the recognition accuracy. The previous studies and experiments have been insufficient in solving the aforementioned problems. Consequently, a new recognition model with a high learning efficiency, moderate required sample, and the ability to avoid learning disorders must be established. Although the
manual feature extraction problem was solved by the traditional CNN structure, the maize leaves and the influencing factors, such as maize leaf diseases, in the original image with large noise were still unclear, resulting in slow training efficiency, low accuracy, and an accumulation of learning disorder problems. Simultaneously, a traditional Relu activation function would suffer neuron "necrosis", meaning the network identification accuracy cannot be improved and the accuracy rate will be decreased. As a result, the primary issues in the present report are:

1) Due to a variety of interfering factors, there are difficulties in capturing images of maize leaves with obvious features in a natural environment. When the original maize leaf data set is directly employed for network training, the noise in the original maize leaf data set contributes to the learning and accumulation of incorrect features, resulting in a learning disorder. Additionally, the background image of the original maize leaf data set may lead to image blurring, resulting in a significant reduction to the recognition accuracy.

2) The convergence time of the traditional neural network [18] model is too long and the computational efficiency is too slow, with the interval time between maize growth and output being considerably short. The identification process becomes complex as the images of maize leaf diseases and insect pests suffer similar features among different diseases. Furthermore, the early and late diseases of the maize leaf will suffer feature differentiation or reduction, leading to a network learning disorder, which requires more in-depth communication regarding feature information in the network.

3) Neuronal "necrosis" may occur in the traditional Relu activation function

In the traditional CNN model, ResNet neural network is a significant network structure model, achieving superior performance during image classification. Despite such superior performance, the network structure of said models, including ResNet, is complex, and ResNet network requires substantial sample training to achieve a high-quality training model. Moreover, the considerable demand for sample training also leads to longer network training and a greater demand for large data amounts for training, meaning a substantial quantity of images of maize leaf pests and diseases are required. Due to the large noise and unclear and similar disease types, obtaining the image of maize leaves is difficult. The traditional ResNet model suffers significant limitations as the model does not prioritize learning important characteristic information, instead providing equal treatment to the characteristics of each piece of channel information, thereby allowing for the learning of wrong information to more easily occur. In view of the aforementioned problems, in the present paper, a maize leaf image enhancement algorithm based on MSRCR and the OSCRNet maize recognition network model was established, which can effectively solve the problem of low training efficiency caused by the large noise of maize leaf images and more redundant spatial information. Concurrently, the algorithm solves the issue of maize leaf diseases and insect pests possessing similar characteristics. Additionally, network learning disorders caused by the realization of different characteristics of information interaction is solved by the
algorithm. By using an improved APRelu activation, neuronal "necrosis" was avoided. Obviously, the proposed method is beneficial in maize leaf image recognition, and the contributions thereof are summarized as follows:

1) A maize leaf image enhancement algorithm based on MSRCR was proposed in order to enhance the characteristics of maize leaf images in a complex environment and solve the problems of high noise and blur of maize leaf images.

2) A maize recognition network model based on OSCRNet was established, and an octave convolution [24] with the characteristics of accelerated network training was adopted, so as to solve the problem of more redundant spatial information in maize leaf images. Further, a convolution module was combined with a multi-scale self-calibrated convolution [23] to realize the interactions between different feature information in maize images, improve the feature extraction ability, and solve the network learning disorder caused by the similarity of maize disease features. Simultaneously, APRelu activation function was adopted to introduce attention mechanisms and solve the problem of neuron "necrosis".

3) Experimental results articulate that, when compared with traditional ResNet[22], Alexnet, and other backbone architecture networks, the proposed method discerns the purpose of high-precision identification of maize leaf pests and diseases images and accelerates network learning, so that the network can be better trained. The effects of the proposed method outperformed other neural networks.

![Fig. 1. The image recognition process of maize leaf diseases and insect pests](image)

The image recognition process of maize leaf diseases and pests is articulated in Fig. 1:

1) Firstly, the data set of maize leaf pests and diseases was constructed.

2) Secondly, MSRCR maize leaf image pretreatment was conducted to obtain the enhanced images of maize leaf diseases and insect pests. Meanwhile, to obtain as many data set samples as possible, the data set was expanded.

3) Thirdly, utilizing the maize image enhanced data set, OSCRNet was trained.

4) Finally, the training model was applied to the identification of maize leaf diseases and insect pests.

Materials and Method

Data collection and preprocessing

Obviously, the meaning of data sets is crucial in the experimental process of deep learning, and the recognition accuracy of neural networks is seriously affected by the quality of data sets. In the present
paper, images of maize leaf diseases and pests from different sources were collected, several of which originated from the maize part of the National Data Sharing Center for Agricultural Sciences and the 2018 AI Challenger Crop Disease Detection Competition [19]. The image dataset of maize leaf pests and diseases included maize rust, maize grey spots, maize Northern blight, and healthy maize samples. Simultaneously, under natural light, the vertical distance between the lens and the maize leaves varied from 5cm-8cm, and the image data obtained from different angles and directions were classified by professionals.

Table 1 Image data set of maize leaf diseases and insect pests

| Category                  | Number | Proportion |
|---------------------------|--------|------------|
| Common_Rust               | 912    | 15.28%     |
| Gray_Leaf_Spot            | 1173   | 19.65%     |
| Northern_Leaf_Blight      | 2169   | 36.33%     |
| Healthy                   | 1716   | 28.74%     |

Several images of maize leaf diseases and pests collected in this paper are shown in Fig.2.

Fig.2 Partial images of maize leaf diseases and insect pests

As shown in Table 1, the distribution of the obtained image data of maize leaf diseases and insect pests was considerably uneven. In image recognition, sample heterogeneity will seriously affect the accuracy of model recognition [20]. As such, conducting image denoising on maize leaf image data [28] to separate and eliminate duplicate and useless data, enhancing and expanding maize leaf image data, and performing horizontal, vertical, random and reverse folding on sample images are all necessary. At the same time, the intensity of random dislocation transformation was 0.2, and the amplitude of random image scaling was set to 0.2, and then batch normalization (BN)[21] was performed. In addition to increasing the convergence rate of the network, normalization can also effectively alleviate the problem of gradient disappearance. The displacement line of the corresponding point in the transformation was 10% of the image side length, so as to avoid image distortion. Using said method, the sample data size was increased by two-fold, and the image data ratio obtained is shown in Table 2.

Table 2 Maize leaf diseases and insect pests image enhancement data set
| Category               | Number | Proportion |
|------------------------|--------|------------|
| Common_Rust            | 3450   | 26.38%     |
| Gray_Leaf_Spot         | 2729   | 20.87%     |
| Northern_Leaf_Blight   | 3738   | 28.58%     |
| Healthy                | 3163   | 24.19%     |

Table 3. Image distribution ratio

| Category               | Train (80%) | Val (20%) |
|------------------------|-------------|-----------|
| Common_Rust            | 2800        | 650       |
| Gray_Leaf_Spot         | 2183        | 546       |
| Northern_Leaf_Blight   | 2991        | 747       |
| Healthy                | 2529        | 632       |

Among the datasets, the training set accounted for 80%, and the verification set accounted for 20%, and the distribution proportion was as shown in Table 3.

Through the maize leaf image preprocessing operation, the data set had the following advantages:
1) The information of feature extraction of maize leaf images was enhanced.
2) The maize leaf data set was expanded as far as possible to provide sufficient training samples for the network.
3) Irrelevant information in maize leaf images should be eliminated to the maximum extent to reduce the impact of unnecessary information on the network.

**MSRCR maize leaf image enhancement**

In image enhancement, maize leaf disease and insect pests images have the characteristics of loud noise, fuzzy image, similar disease type characteristics and other characteristics. The traditional Retinex algorithm [29] relies on the incident image and reflected image to constitute:

\[ I(x, y) = L(x, y) \cdot R(x, y) \] (1)

Where, \( I(x, y) \) represents the received image signal of maize leaves, \( L(x, y) \) represents the light irradiation component in the external environment, and \( R(x, y) \) represents the reflection component of the target maize leaves. The formula can be obtained by using the same logarithm operation:

\[ \log[R(x, y)] = \log[I(x, y)] - \log[L(x, y)] \] (2)

From the formula, \( I(x, y) \) and a Gaussian filter can be used to obtain the result, which is expressed as:

\[ \log[R(x, y)] = [\log(I(x, y))] - [\log[I(x, y) \ast G]] \] (3)

However, in Retinex data processing, images are easily blurred, resulting in the distortion of maize leaf images and the loss of certain feature information. Retinex data processing also has the disadvantages of low efficiency and a large amount of calculation. At the same time, the loss of information in maize
Based on the Retinex algorithm, a multi-scale MSRCR image enhancement algorithm with color restoration for maize leaves was proposed. The MSRCR algorithm is an improved algorithm based on traditional Retinex algorithm. The traditional Retinex algorithm is single scale, while the MSRCR algorithm is an improved multi-scale image enhancement method based on the traditional algorithm. The basic formula is as follows:

$$R_{MSRCR}(x, y) = \frac{R_{MSRCR}(x, y) - \min_{(x,y)}(R_{MSRCR}(x, y))}{\max_{(x,y)}(R_{MSRCR}(x, y) - \min_{(x,y)}(R_{MSRCR}(x, y)))}$$ \hspace{1cm} (4)$$

Significant image polarization will occur owing to the images processed by Retinex being highly dynamic, and the data distribution being considerably wide, such that certain characteristic information in maize leaf images will be lost. For certain original image HUE graphs, if the traditional algorithm is used, color bias can more easily occur in the processed maize leaf images. Thus, the output data was processed by Retinex first, and then the data were mapped to each channel according to the original RGB proportion. On the basis of retaining the original color distribution, the maize leaf images were enhanced.

$$I'_i(x, y) = \frac{I_i(x, y)}{\Sigma_{j=1}^{s}I'_i(x, y)}$$ \hspace{1cm} (5)$$

$$C_i(x, y) = \beta \log[\alpha I'_i(x, y)]$$ \hspace{1cm} (6)$$

$$R_{MSRCR}(x, y) = C_i(x, y)R_{MSR}(x, y)$$ \hspace{1cm} (7)$$

Through the aforementioned method processing, compared with the traditional Retinex algorithm, the MSRCR maize leaf image enhancement algorithm images were brighter. Further, due to the multi-scale characteristic, not only were the plant diseases and insect pests of maize leaf images clearer, the color was brighter. As the MSRCR parameters did not need to be too complicated, the efficiency was improved. At the same time, the quality was improved to a certain extent, effectively removing the noise in the images of maize leaf diseases and insect pests, reducing the ambiguity of maize leaf images, and retaining the necessary details in the images. The enhanced effect of maize leaf image is shown in Fig. 3.
Plant diseases and insect pests of maize leaf images have the characteristics of being complex and having a large amount of noise. To achieve the effect of rapid recognition and high precision, ResNet was introduced as the neural network framework, so as to improve the learning effect of neural network, reduce the complexity, and improve the ability of feature extraction. Octave convolution and self-calibration and combination were also introduced. The PRelu activation function was selected and configured by OSCRNet [31]. Additionally, batch normalization was introduced, which can effectively improve the robustness and identification accuracy of the network model, improve the convergence rate and prevent overfitting. The OSCRNet maize leaf image recognition network structure consists of an octave convolution module, four convolution layers, a self-calibrated convolution module and three fully connected output layers. In the present study, the octave convolution in the first layer is relied on, so as to obtain more characteristic information, reduce redundant spatial information, reduce the computational overhead of the network, and accelerate the computational efficiency of the network. Self-calibrated
convolution is used to improve the attention and generalization ability of neural network, adaptively consider the information of the surrounding environment, improve the accuracy of network, and accelerate the efficiency of network calculation. In the last convolution layer, the extracted features are all merged and the classification function is performed by the three fully connected layers. Fig. 4 presents the architecture of the OSCRNet maize leaf image recognition network. Batch normalization is used after each convolution layer to improve the robustness and identification accuracy of the model and avoid overfitting. The octave convolution module is the first convolution layer in the network, which can obtain more characteristic information and reduce redundant information. To improve the accuracy and computational efficiency of the network, the multi-scale convolution module is located before the last convolution layer. The model definition for OSCRNet is shown in Fig. 4.

**Fig. 4 OSCRNet network structure**

1) The first layer is formed by connecting the input layer and the BN layer. To enhance the generalization ability of the model and accelerate the convergence speed of the network, images are preprocessed through the BN layer.

2) The second layer is composed of convolution module 1. Conv1 has a core size of 7x7 and a depth of 64. The ability of image feature extraction is enhanced by using octave convolution, and the network convergence speed is accelerated by using APRelu operation and then BN layer processing.

3) The third layer is composed of two convolution modules. Conv2 has a core size of 3x3 and a depth of 64. With the APRelu operation, BN layer processing is performed after maximum pooling.

4) The fourth layer is composed of three convolution modules. Conv3 has a core size of 3x3 and a depth of 128. With the APRelu operation, BN layer processing is performed after maximum pooling.

5) The fifth layer is composed of four convolution modules. Conv4 has a core size of 3x3 and a depth
The sixth layer is composed of the self-calibrated convolution module. First, the results from the previous level are split into X1 and X2. The second step is to process the self-calibrated scale space. The X1 is sampled 4 times by average pooling, and then upsampling is conducted after convolution. At the same time, the original value X1 is added, and then the sigmoid activation function is used to multiply the features of Conv5_2 to obtain the output feature Y1. The third step is to process the original scale feature space. After Conv_3 convolution of feature X2, feature Y2 is obtained. In the fourth step, the output features Y1 and Y2 of the two scale spaces are concatenated to obtain the final output feature Y.

The seventh layer is composed of five convolution modules. Conv6 has a core size of 3x3 and a depth of 512. BN layer processing is performed after the APRelu operation.

The eighth layer is the first fully connected layer with 200 neurons. The APRelu operation and Dropout process are performed.

The ninth layer is the second fully connected layer with 100 neurons, which are then processed by the APRelu operation and Dropout.

The tenth layer is the last fully connected layer, composed of four neurons, representing the number of species of maize leaf diseases and pests. To determine the classification of input images, the output of the last fully connected layer will be transmitted to the output layer. The final output values add up to 1.0, with individual output limits ranging from 0 to 1. The detailed schematic diagram is shown in Fig. 5.
ResNet deep residual network

Deep Residual Network is a deep convolutional neural network model proposed by Kaiming et al. in 2015. In traditional convolutional neural networks, as the number of convolutional layers increases, degradation problems occur. Further, the accuracy of the network is saturated, and the convergence speed slows down, resulting in gradient explosion or vanishing, and a decrease in accuracy. In order to solve the problem of reduced training accuracy, a deep residual learning framework was introduced in ResNet introduced a deep residual learning framework. For deeper training models, additional layers can be built through identity mapping, or shortcut connections. Instead of each layer directly fitting the underlying mapping required, the layers fit into a residual mapping.
For the accumulation layer structure, when the input layer is $x$, the learned feature is denoted as $H(x)$, and $H(x) = F(x) + x$ can be obtained. $H(x) = F(x) + x$ can be used to conduct residual learning, namely: $F(x) = H(x) - x$, where $F(x)$ is the network map before the summation, and $H(x)$ is the network map after summation, because residual learning is easier than the original features. When the residual approaches 0, to ensure that the network performance will not decline, only identity mapping is performed for the accumulation layer. At the same time, the accumulation layer will learn new features based on the input features, so as to achieve better performance. The specific network structure is shown in Fig.6.

$$y = F(x, \{W_i\}) + x$$  \hspace{1cm} (8)$$

$$y = F(x, \{W_i\}) + W_i x$$  \hspace{1cm} (9)$$

The deep residual network can indeed stabilize the network performance while increasing the network depth, and improve the accuracy of the model to a certain extent. However, a large number of training models and residual structures increase the difficulty and training time of the training. Compared with other models, more training time is consumed, and the deep network structure also requires a larger data set. In complex natural environments, a large number of plant diseases and insect pests of maize leaf of the original images are difficult to obtain. A new OSCRNet maize leaf recognition network model was designed in the present study, which can effectively overcome the problems of small sample size, maize leaf image characteristics being too complex and leading to learning difficulties, and even the inability to learn.

**Octave convolution**

Due to the reflected light in the external environment, sand and other influencing factors, interference with the photographing of maize leaf images can more easily occur, resulting in more redundant spatial information of maize leaf pests and diseases, thereby significantly reducing the efficiency of network training. In order to solve the aforementioned problems, octave convolution with
the characteristics of accelerated convolution operation was adopted in the present study. Octave
originally refers to the octave scale, and refers to halving the sound frequency in music. The aim of
octave convolution is to halve the low-frequency information in the data, so as to accelerate the
convolution operation. For ordinary convolution, all input and output feature maps have the same spatial
resolution, but in natural images, information is transmitted at different frequencies. As shown in Fig. 7,
the higher frequencies are usually encoded with fine details, while lower frequencies usually use global
structure coding. The output feature map of convolution can also be seen as mixed information of
different frequencies. Natural images can be decomposed into low-frequency signals that capture the
global layout and coarse structure and capture fine details.

![High and low frequency diagram](image)

Fig. 7 High and low frequency diagram

Similarly, the features of the convolution output should also have a subset of mappings that capture
spatial low-frequency variations and contain spatial redundancy information. To reduce such spatial
redundancy, octave feature representation was introduced. Scale space theory [26] provides a principled
approach to creating spatial resolution scale spaces in such a way that low frequency and high frequency
spaces can be defined, that is, the spatial resolution of the low frequency feature map is reduced by one
octave. In octave convolution, the mixed feature map is decomposed according to the frequency thereof,
and the octave convolution operation is used to store and process the feature map with low spatial
resolution and slow spatial change, thereby reducing the memory and computing cost. Different from the
existing multi-scale methods, octave convolution is represented as a single, universal convolution unit,
and the specific network structure is shown in Fig. 8.
Fig. 8 Octave Convolution structure diagram

### Octave Convolution

$$f(x^H, w^H)$$

$$f(x^L, w^L)$$

$$f(\text{pool}(x^H, 2), w(H \rightarrow l))$$

$$\text{upsample}(f(x^l, w(L \rightarrow H), 2))$$

**High frequency:**
$$Y^H = f(x^H, w^{(H \rightarrow H)}) + \text{upsample}(f(x^L, w^{(L \rightarrow H)}, 2))$$  \(\text{(10)}\)

**Low frequency:**
$$Y^L = f(X^L, W^{(L \rightarrow L)}) + f(\text{pool}(x^H, 2), W^{(H \rightarrow L)})$$ \(\text{(11)}\)

In the present study, octave convolution was adopted to expand the receptive field by two times, further facilitating the capture of more information by each layer. By reducing unnecessary and long spatial information in maize leaf images, the problem of more redundant spatial information in maize leaf images of diseases and insect pests was solved, and the learning efficiency of the network was significantly improved.

The use of octave convolution can provide the network with the following advantages:

1) The spatial resolution of the low-frequency feature map of maize leaf can be reduced to half of the original, the redundant spatial information in the maize leaf images can be reduced, and the network computing overhead can be reduced, so as to accelerate the network computing efficiency and solve the problem of slow network learning efficiency.

2) Compression of low frequency resolution of maize leaves can effectively expand the receptive field by two times, further facilitating the capture of more information by each layer.

### Self-Calibrated Convolutions

Owing to the increase of maize resistance and differentiation of disease characteristics, the characteristics of maize diseases are similar, increasing the complexity of identification, and leading to the network being prone to learning disorders. Under such circumstances, deeper and more detailed communication of feature information is required in the network. A self-calibrated convolution with multi-scale characteristics was adopted in the present study to solve the aforementioned problems. At present, the latest progress of deep learning mainly focuses on the design of more complex network structure to enhance the feature learning and expression ability thereof. By expanding the receptive field
of each convolution layer, self-calibrated convolution can enrich the output function. Meanwhile, self-calibrated convolution is different from standard convolution, which uses small convolution kernel to fuse spatial and channel direction information. Self-calibrated convolution can adaptively establish long-distance spatial and inter-channel dependence calibration operations around each spatial position, and thus, can help CNN generate more identifying feature information by explicitly merging richer information. The essence of self-calibrated convolution is a grouping convolution used for multi-scale [17] feature extraction, which is divided into two groups according to channel dimension. To increase the receptive field of the network, one path is used for conventional convolution feature extraction, and the other path is used for down-sampling operation. The result is that each spatial position can be self-calibrated by fusing information from two different spatial scales.

![Fig.9. Principle diagram of Self-Calibrated convolution](image)

In order to facilitate the network to learn better different characteristics, improve the attention and generalization ability of neural network. Self-calibrated convolution separates each channel of the convolution, but instead of each part of the channel being treated equally, each part of the channel is responsible for a different function. The self-calibrated convolutional neural network performs feature transformation at two different scales: one is the characteristics of the proportion of the original space mapping, and the other is to use the sample of potential space mapping. As such, the design of the structure is more conducive to improve the attention of the network, and help the network focus on the differentiated characteristics. The specific network structure is shown in Fig.9.

First, given input X1, the average pooling of filter size r×r and stride R is adopted:

\[ T_i = \text{AvgPool}_r(x_i) \quad (12) \]

Secondly, K2 is used for T1 to conduct feature transformation:

\[ x_i' = \text{Up}(F_2(T_i)) = \text{Up}(T_i * K_2) \quad (13) \]

There is a bilinear interpolation operator that maps the intermediate reference from the small scale space to the original feature space:

\[ Y_i' = F_3(x_i') \cdot \sigma(X_i + X_i') \quad (14) \]

Where F3(X1)= X1*K3, \( \sigma \) is the sigmoid function, and "·" means multiplying element by element. As the equation shows, using \( X_1 \) as the residual to form the weight for calibration is beneficial. The final output after calibration can be written as:
The self-calibrated convolution is a multi-scale feature extraction module, which can separate each channel of convolution, but is different from traditional convolution in that each channel is responsible for a special function. As aforementioned, the self-calibrated convolutional neural network performs feature transformation at two different scales: one is the characteristics of the proportion of the original space mapping, the other is to use the sample of potential space mapping. The interference of unrelated areas throughout the global information is avoided, and the characteristic representation of maize leaves generated by self-calibrated convolution is more discernable. The self-calibrated convolution module allows each spatial location to not only adaptively view the surrounding information environment as an embedment from the low-resolution potential space as an input in the response from the original scale space, but also model the interchannel dependencies by adopting the correction operation. Through the aforementioned methods, the receptive field with self-calibrated convolution layers can be effectively expanded, which is beneficial in improving the network attention and the differentiation of network attention, solving the problem of network learning disorder, and significantly improving the accuracy of the network.

The specific advantages of self-calibrated convolution include:

1) The self-calibrated network can locate multi-scale information in maize leaf images.
2) The feature extraction ability is improved. The interaction of different feature information in maize leaf images can avoid the interference of irrelevant information and solve the problem of network learning disorder.
3) Because of the potential mapping of the sub-sampling space, in addition to adaptively considering the information of the surrounding environment, each spatial position can also learn more accurate difference regions, thereby improving the feature extraction ability.

Attention mechanism

The visual attention mechanism [30] is a brain signal processing mechanism unique to human vision. Human vision quickly scans the global image to obtain the target region that needs to be focused on, also known as the focus of attention, and then invests more attention resources to said region to obtain more details of the target that needs to be paid attention to, while suppressing other useless information. According to the specific task goal, the attention mechanism can adjust the attention direction and the weighted model. In the neural network, the weight of the attention mechanism can be increased to weaken or forget the content that does not conform to the attention model. Limited attention resources can be used to quickly select high-value information from a large amount of information, which significantly improves the efficiency and accuracy of information processing in deep learning.
Adaptive parameterized APRelu activation function

In the process of neural network learning, the activation function is considerably important, and can help neural networks to learn and understand. The activation function often used is the Relu activation function [32], also known as the modified linear unit. The convergence speed of Relu is faster than that of Sigmoid and TANh. At the same time, the Relu gradient will not be saturated, thereby solving the problem of gradient disappearance. However, the activation function of Relu is considerably fragile during training. Due to sunlight, dust and other interference in the environment, frequent use will slow down the gradient change in the saturated area and cause the gradient to disappear. When X <0 of Relu, the gradient will be 0, leading to the gradient being set to zero and the occurrence of neuron "necrosis" at the same time. There is also the possibility that 40 percent of the neurons in the network will die, and be unable to be reactivated.

$$Relu(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases}$$ (16)

To solve the aforementioned problems, the PRelu activation function [31], also known as Parametric Rectified Linear unit (PRelu) was established, which is an improved version of Relu. The PRelu activation function has a considerably small slope, but does not tend to 0. Compared with the Relu activation function, the PRelu activation function can solve the problem of neuron necrosis in Relu considerably well.

$$PRelu(x) = \begin{cases} x, & \text{if } x > 0 \\ q_x x, & \text{if } x \leq 0 \end{cases}$$ (17)

In the present study, the adaptive parameterized APRelu activation function was improved by combining the attention mechanism with the PRelu activation function. The APRelu activation function uses the concept of SENet for reference. For most maize leaf feature information, the importance of each feature channel in the maize leaf feature map is likely to be different. As an example, feature channel 1 of sample A is considerably important, while feature channel 2 is not important. However, feature channel 2 is important for sample B, while feature channel 1 is not. At this time, for sample A, the attention of the neural network should be focused on characteristic channel 1 and higher weight should be given to channel 1. For sample B, the focus should be on feature channel 2 and higher weight should be given to channel 2. SENet, on the other hand, can use a small, fully connected network to learn a set of weight coefficients and then weight each channel, which uses an attentional mechanism. The structure of SENet is shown in Fig.10.

Fig.10. APRelu principal diagram

The APRelu activation function can be weighted through a small fully connected network, and the
set of weights can be used as coefficients in the PRelu activation function. When the neural network uses
the APRelu activation function, each sample can have unique weight coefficients, which is a different
nonlinear variation, thereby enabling the neural network to learn more feature information in the maize
leaf images, and to learn more focused feature information of different importance. The activation
function of APRelu solves the problem of neuron "necrosis" in Relu and reduces the instability in the
process of network training. Different weight coefficients are also assigned to different channels, so that
the neural network can learn more abundant characteristic information.

Batch normalization

The images of maize leaf diseases and insect pests are considerably complex and changeable,
leading to low learning efficiency of the neural network and increasing the difficult of learning for certain
neural networks. Meanwhile, with the deepening of the network structure, the data distribution of the
hidden layer also changes a lot or even fluctuates, which will adversely affect the stability of the
network. As such, the BN algorithm was adopted in the present study to normalize the data of each layer
into an average value of 0 and a standard deviation of 1

\[
\mu = \frac{1}{n} \sum_{i=1}^{n} x_i \\
\sigma = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2
\]  
(18)

\[\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}}\]  
(19)

\(\varepsilon\) is constant, so as to prevent the occurrence of a fractal failure, so that the standard deviation is
zero.

Results

Laboratory environment

To better train and test the performance of our model, Windows10 (64-bit) operating system, Pytorch
deep learning framework, programming environment Python3.6, lab platform memory 16GB, and Intel
Core i7-9700K CPU equipped with NVIDIA GeForce GTX1660Ti 6GB GPU were used.

Experimental data and analysis

In the present study, an extended dataset consisting of four classes was used, totaling 13,078
images. The training set accounted for 80%, and the verification set accounted for 20%. The most
appropriate network configuration and parameters before training had to be conFig.ured, so as to accelerate
the speed of network training and improve the identification accuracy.
As shown in Fig.11:

1) Image recognition accuracy gradually increased with the iteration number of the ResNet34 training process.

2) The image recognition accuracy increased gradually with the iterations of the OSCRNet training process.

An observation can be made that, compared with ResNet34, OSCRNet network could extract more feature information in the data set under the training process of the same number of iterations, such that the recognition accuracy was significantly improved. Additionally, the ResNet34 model needs more iterations to achieve the ideal model accuracy. The OSCRNet model can achieve stable identification accuracy quickly and with less fluctuation.

**Ablation experiment**

In order to more clearly observe the influence of the improvement in the present study on the neural network, ablation experiments were conducted for comparison.

| Model          | ResNet34 | Octave | Self-calibration | APRelu | OSCRNet |
|----------------|----------|--------|-----------------|--------|----------|
| Time           | 431.27   | 394.51 | 442.29          | 453.36 | 413.69   |

As shown in Table 4, an observation can be made that with the improvement of the octave convolution module, the network learning speed was significantly improved by reducing a large amount of verbose spatial information in the data set. There was also improvement in the APRelu module, which could be attributed to the attention mechanism of the weight calculation, resulting in an increase in the amount of network calculation. Further, the networking speed was better improved by OSCRNet.

**Table 5. Recognition accuracy rate under each individual module**
As shown in Table 5, under the improvement of the octave convolution module, the accuracy of the network decreased, which can be attributed to the octave convolution module excluding the redundant spatial information and a small part of the useful feature information when processing the feature information. Such circumstances led to an improved network with only the octave convolution module. Although the receptive field was increased, the loss of part of the feature information in the maize leaf images would still lead to a decrease in accuracy. The accuracy of maize leaf image recognition was improved by the self-calibrated convolution module and the APrelu module, while there was steady improvement with OSCRNet. Through the aforementioned experiments, an observation can be made that although the model was improved under the separate octave convolution module, self-calibration convolution module, and APrelu module, there were still defects. At the same time, in the absence of octave convolution operation, in which only the self-calibration convolution module was used, the recall rate would be reduced in several cases. Such reduction could be attributed to, without the screening information of octave convolution, the self-calibration convolution module being more likely to learn irrelevant or even wrong information from the maize leaf images, which can lead to disordered learning of maize leaf image feature information and a reduced recognition rate. Under the combined effect of octave convolution, self-calibration convolution and APrelu, the learning efficiency and the accuracy of the network were improved, in addition to ensuring the robustness of the network to a certain extent.

**Comparison with other networks**

To enhance the comparability of the present experiment, the same data set was used to train different networks under the same conditions:

| Schemes  | Model  | 精度   |
|----------|--------|--------|
| Scheme1  | ResNet34 | 76.42% |
| Scheme2  | Vgg    | 94.38% |
| Scheme3  | Alexnet | 83.59% |
| Scheme4  | OSCRNet | 97.68% |

As shown in Table 6, OSCRNet could achieve higher identification accuracy than the other networks. Such high accuracy could be attributed to the self-calibrated convolution module, in which feature transformation was conducted at two different scales. Such feature transformation avoided the interference of irrelevant regions in the global information, made the generated feature representation more discernable, and improved the attention of the network, thereby enabling the network to learn more
critical feature information. Meanwhile, in order to increase the number of training samples, horizontal folding, vertical folding, random folding and reverse folding were performed on the sample images. The intensity of random dislocation transformation was 0.2, and the amplitude of random image scaling was set to 0.2, and then normalized processing was conducted. A variety of color features was provided to increase image diversification, reduce complexity, and avoid network overfitting.

Table 7. Network training speed of different networks

| Model       | ResNet34 | Vgg  | Alexnet | OSCRNet |
|-------------|----------|------|---------|---------|
| Time        | 431.27   | 428.53 | 486.67  | 413.69  |

As shown in Table 7, compared with other network models, OSCRNet had the highest convergence rate. Regarding ResNet34, octave convolution was introduced to reduce the redundant spatial information in the maize leaf images, such that the convergence rate was also improved. The octave convolution module could reduce the spatial resolution of the low-frequency maize leaf feature map to half of the original one while expanding the receptive field by two times, reduce the redundant spatial information in the maize leaf images, and significantly reduce the computing overhead of the network.

Table 8 OSCRNet's recognition accuracy of various types of maize leaves

| Classification      | Original Dataset | Enhancement Dataset |
|---------------------|------------------|---------------------|
| Common_Rust         | 93.38%           | 95.67%              |
| Gray_Leaf_Spot      | 91.27%           | 93.86%              |
| Northern_Leaf_Blight| 88.63%           | 89.94%              |
| Healthy             | 92.54%           | 94.63%              |

Table 9. OSCRNet's histogram of the recognition accuracy of various types of maize leaves

As shown in Table 8, by analyzing the different identification accuracy of OSCRNet for each different species, the identification accuracy of OSCRNet for common maize rust was found to be higher
than that of the other maize leaf pests and diseases. OSCRNet has obvious advantages over other networks in terms of Gray_Leaf_Spot and Northern_Leaf_Blight identification.

Table 10. Accuracy and training efficiency of different improved networks

| Model                        | Original Dataset | Enhancement Dataset | Time   |
|------------------------------|------------------|---------------------|--------|
| GoogLeNet[38]                | 84.64%           | 85.38%              | 417.96 |
| MobileNetV2[36]              | 83.27%           | 87.41%              | 423.75 |
| DMS-Robust Alexnet[40]       | 85.91%           | 87.39%              | 425.59 |
| TEL-ResNet[35]               | 89.65%           | 91.69%              | 415.26 |
| OSCRNet                      | 91.21%           | 93.53%              | 413.69 |

As shown in Table 10, OSCRNet had the highest training efficiency compared with other improved network models under the same conditions. When octave convolution was introduced into OSCRNet, redundant spatial information in maize leaf images was reduced, and the convergence rate was also improved. In the octave convolution module and the self-calibration module, the sensing field was expanded by two times, and the spatial resolution of the low-frequency feature map of maize leaf was reduced to half of the original, which reduced the redundant spatial information in the maize leaf images, significantly reduced the network computing overhead, and improved the network computing speed. At the same time, OSCRNet had the highest training accuracy compared with other improved models under the same conditions, reaching 93.53% in the enhanced data set. Compared with other models, the robustness of the model was impromptu, rendering the learning process more reliable and significantly avoiding the phenomenon of network learning disorder.

Discussion

We have made oscrnet for maize image recognition, which covers a variety of disease features of maize leaves. Through the comparison and analysis of multiple groups of experiments, the availability of the oscrnet model proposed in this paper for maize image recognition is effectively verified. Many plant images are used in the field of deep learning. However, maize image recognition technology is not perfect, and the improvement of efficiency and recognition rate is slow. Therefore, our work is conducive to promote the application of depth learning based method in plant protection. Analyzing the error information learned from the model is a very useful insight to improve the network performance proposed by us. In the design of oscrnet model, we can classify and avoid the problem of similar plant disease features in plant images. However, in our data set, such problems will still affect the model learning, which is a very serious problem in the field of plant recognition. The similar characteristics of maize diseases are shown in Fig.12. Although oscrnet uses the information interaction of self calibrated revolutions to solve this problem and has been greatly improved, it still can not completely avoid it. Therefore, plant recognition methods still need to be improved.
In order to solve such difficulties, we will consider introducing attention mechanism to improve and alleviate this problem in our future work. At the same time, expand the number of data sets as much as possible to ensure that the data sets are more balanced and perfect.

Conclusions

With the progress of deep learning and computer vision technology, image recognition, image analysis and other fields have been extensively studied. Deep learning-based image recognition and management of crop pests and diseases has also been extensively applied in recent years. For better improving the yield of maize crop, a maize leaf image enhancement algorithm based on the MSRCR and OSCRNet networks were proposed to identify maize leaf diseases and insect pests. More feature information was provided through the maize leaf image enhancement processing of MSRCR, which solved the problem of maize leaf disease and insect pests image feature extraction in a more complex environment. In OSCRNet, the octave convolution and self-calibrated convolution was combined to solve the problems of slow learning efficiency, low accuracy and easy learning disorder in the image recognition of maize leaf diseases and insect pests. The activation function of APRelu was improved to avoid the phenomenon of neuron "necrosis". Compared with the data of other network models, the present results show that OSCRNet can achieve better performance in the image recognition of maize leaf diseases and insect pests. At the same time, the algorithm of OSCRNet is not only limited to the field of maize leaf diseases and insect pests, but can also detect and identify other crop diseases. The OSCRNet algorithm is of significant value to crop yield and crop disease and insect pest control, and opens up new possibilities for the prevention and management of crop diseases.

Abbreviations

MSRCR: Multi-Scale Retinex with Color Restoration
OSCRNet: Octave and Self-Calibrated ResNet
Statement

- Ethics approval and consent to participate
  Not applicable
- Consent for publication
  Agree to publish
- Availability of data and materials
  During the current study, data sets were obtained from the authors at reasonable request
- Competing interests
  The authors declare that they have no competing interests
- Funding
  This work was supported by Changsha Municipal Natural Science Foundation (Grant No. kq2014160); in part by the National Natural Science Foundation in China (Grant No. 61703441); in part by the key projects of Department of Education Hunan Province (Grant No. 19A511); in part by Hunan Key Laboratory of Intelligent Logistics Technology (2019TP1015).
- Authors' contributions
  ZHJ is a major contributor to this manuscript
- Acknowledgements
  Not applicable

Reference

[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] 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.
[3] 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
[4] 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.
[5] 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.
[6] 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.

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

[8] 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.

[9] 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.

[10] A. Khan, A. Sohail, U. Zahoor, 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

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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

[16] 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.

[17] M. Lin, Q. Chen, and S. Yan, “Network in network,” 2013, arXiv:1312.4400. [Online]. Available: http://arxiv.org/abs/1312.4400

[18] 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.

[19] SINGKEK SCI-TECH.Agricultural Disease.2018 AI Challenger[DB/OL]. (2018-12-19) [2020-3-20.
https://challenger.ai/competition/pdr2018.

[20] 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.

[21] Zhu Wei, Qu Jingyi, Wu renbiao. Straight through convolution neural network image classification algorithm combined with batch normalization [J]. Journal of computer aided design and graphics, 2017, 29 (9): 1650-1657.

[22] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] / IEEE. 2016 IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 770-778.

[23] Jiang-Jiang, Liu Qibin, Hou Ming-Ming Cheng, Changhu Wang, Jiashi Feng, CS, Nankai University NUS ByteDance AI Lab “Improving Convolutional Networks with Self-Calibrated Convolutions” 2020 doi: 20cvprSCNet.pdf (mmcheng.net)

[24] Yunpeng Chen, Haoqi Fan, Bing Xu, Zhicheng Yan, Yannis Kalantidis, Marcus Rohrbach, Shuicheng Yan, Jiashi Feng Facebook AI, National University of Singapore, “Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks with Octave Convolution” 2019, doi: 1904.05049 (arxiv.org)

[25] Tang, Jing Lei, et al. “Weed identification based on K-means feature learning combined with convolutional neural network”, Computers and electronics in agriculture vol. 135, pp. 63-70, Apr. 2017. doi: 10.1016/j.compag.2017.01.001.

[26] Cong Ge. Nonlinear scale space theory in computer vision [D]. Institute of automation, Chinese Academy of Sciences. Institute of automation, Chinese Academy of Sciences, 1997 doi: http://ir.ia.ac.cn/handle/173211/5667

[27] Wenjie Luo, Yujia Li, Raquel Urtasun, Richard Zemel “Understanding the effective receptive field in deep convolutional neural networks” doi: Understanding the effective receptive field in deep convolutional neural networks | Proceedings of the 30th International Conference on Neural Information Processing Systems (acm.org)

[28] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, Timo Aila” Noise2Noise: Learning Image Restoration without Clean Data” 2018 doi: [1803.04189]

[29] Chen Wei, Wenjing Wang, Wenhan Yang, Jiaying Liu “Deep Retinex Decomposition for Low-Light Enhancement” 2018 doi: 1808.04560.pdf (arxiv.org)

[30] Xiaolong Wang, Ross Girshick, Abhinav Gupta, Kaiming He “Non-local Neural Networks” doi: [1711.07971v1] Non-local Neural Networks (arxiv.org)

[31] Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun “Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification” 2015 doi: Delving Deep into Rectifiers:
[32] Zhao M, Zhong S, Fu X, et al. Deep residual networks with adaptively parametric rectifier linear units for fault diagnosis[J]. IEEE Transactions on Industrial Electronics, 2020. doi:10.1109/TIE.2020.2972458.

[33] Fu L, Duan J, Zou X, et al. Computers and Electronics in Agriculture Banana detection based on color and texture features in the natural environment[J]. Computers and Electronics in Agriculture, 2019. doi:167.

[34] Mazen, FMA, Nashat. AAUse artificial neural network to classify bananas by their maturity. Arab J Science and Engineering 44, doi: 6901-6910 (2019).

[35] Liu Aoyu, Wu Yunzhi, Zhu Xiaoning. “Maize disease recognition based on deep residual network” 2021 doi:10.3969/j.issn.1000-4440.2021.01.009

[36] Yang Mingxin, Zhang Yaoguang, Liu Tao. “Maize disease recognition based on the Convolutional Neural Network with a small sampling size” 2020. doi: 10.13930/j.cnki.cjea.200375

[37] Mao Yandong, Gong He. “Maize disease identification study based on SVM and DS evidence theory fusion multi-features[J]” 2020. doi:10.13733/j.jcam.issn.2095-5553.2020.04.024

[38] Li Jing, Chen Guifang, An Ning, “Image recognition of Pyrausta nubilalis based on optimized convolutional neural network” 2020 doi: 10.7671/j.issn.1001-411X.201907017

[39] Wang Can, Wu Xinhui, Zhang Yanqing, Wang Wenjun. “Recognition and segmentation of maize seedlings in field based on dual attention semantic segmentation network[J]” 2021. doi: 10.11975/j.issn.1002-6819.2021.09.024

[40] Lv, Mingjie, Guoxiong Zhou, He, Mingfang, etc. Maize Leaf Disease Identification Based on Feature Enhancement and DMS-Robust Alexnet, IEEE Access, 2020. doi:8(2):57952-57966