WILEY | Hindawi

Research Article

Image Recognition of Crop Diseases and Insect Pests Based on Deep Learning

Mingyuan Xin [ID]^1 **and Yong Wang**^{2}

^{1}*School of Computer and Information Engineering, Heihe University, Heihe, China*
^{2}*Institute of International Education, Heihe University, Heihe, China*

Correspondence should be addressed to Mingyuan Xin; xinmy0817@126.com

Deep learning algorithms have the advantages of clear structure and high accuracy in image recognition. Accurate identification of pests and diseases in crops can improve the pertinence of pest control in farmland, which is beneficial to agricultural production. This paper proposes a DCNN-G model based on deep learning and fusion of Google data analysis, using this model to train 640 data samples, and then using 5000 test samples for testing, selecting 80% as the training set and 20% as the test set, and compare the accuracy of the model with the conventional recognition model. Research results show that after degrading a quality level 1 image using the degradation parameters above, 9 quality level images are obtained. Use YOLO's improved network, YOLO-V4, to test and validate images after quality level classification. Images of different quality levels, especially images of adjacent levels, are subjectively observed by human eyes, and it is difficult to distinguish the quality of the images. Using the algorithm model proposed in this article, the recognition accuracy is 95%, which is much higher than the basic 84% of the DCNN model. The quality level classification of crop disease and insect pest images can provide important prior information for the understanding of crop disease and insect pest images and can also provide a scientific basis for testing the imaging capabilities of sensors and objectively evaluating the image quality of crop diseases and pests. The use of convolutional neural networks to realize the classification of crop pest and disease image quality not only expands the application field of deep learning but also provides a new method for crop pest and disease image quality assessment.

1. Introduction

The correct diagnosis and prevention of various crop diseases and insect pests can create a bumper harvest in agricultural production and thus meet the daily needs of the people. Deep learning technology has great advantages in image recognition, but the basic recognition algorithm requires a large number of parameters and requires a large sample database as a comparison library. For crop diseases and insect pests, deep learning technology needs a lot of optimization in sample training.

After the rise of artificial intelligence, deep learning technology has also been widely used in saliency detection. Han proposed SRNet; the model is composed of four parts of convolutional layer modules with different functions. The BN layer and residual network are effectively used, and channel selection is added to improve the accuracy of the model's detection of steganographic algorithms [1]. Chen et al. proposed Zhu-Net, using 3&3 cores instead of traditional 5&5 cores, and optimizing the convolutional layer in the preprocessing layer to improve the detection accuracy of spatial steganography [2]. Chen et al. proposed that different spatial locations obtain saliency through competition and then obtain saliency regions through saliency map fusion and designed a winner-take-all activation network at the back end of the model to simulate the location of human visual dynamic updates [2]. Chui et al. proposed to control the two-way transfer of shallow and deep features to obtain accurate predictions. Using HED iteratively refines its output by using a reverse attention model [3]. Based on the parameter assumptions of the outdoor illumination map, Zhao and Liu use CNN to predict the sun position parameters,

atmospheric condition parameters, and camera parameters in the input image and then use the predicted parameters to synthesize the corresponding HDR environmental illumination map [4].

Zhao et al. add the residual items of local illumination on the basis of global illumination and restore the global illumination information and local illumination information through cascaded CNN through staged prediction and obtain more detailed illumination information [5]. Wu and An predict the pixel-level attention map through the context attention network and combine it with the U-Net network structure to detect the saliency area [6]. Zhai et al. proposed a saliency detection model BASNet based on boundary perception. This model integrates BCE, SSIM, and IoU three loss functions to guide the network to pay more attention to boundary quality, accurately segment saliency regions, and make the boundaries of saliency regions more clear [7]. Miao et al. use the skylight model to predict good outdoor lighting information with only a few parameters and realize the rerendering application of virtual object insertion [8]. Tao et al. use CNN to complete the prediction of outdoor HDR environmental lighting images. The method used can correctly estimate the illuminators in the scene, but the lack of illuminators in the scene can lead to incorrect predictions [9]. Although these methods have improved the standard of saliency detection, there is still much room for improvement in the quality of fine structure segments and the accuracy of the boundary. The illumination information predicted by the above method is very detailed but contains a large number of parameters. For this reason, this paper proposes a deep network model fused with Google data analysis.

Based on the existing objective quality evaluation and quality level classification of crop pests and diseases, this paper uses the classification function of deep learning to construct a deep convolutional neural network model for the classification of crop pest and disease image quality. The model realizes the classification of crop disease and insect pest images with multiple quality levels, which is more detailed and accurate than the comparison literature method. The images after the quality level classification are tested and verified by the improved network YOLO-V3 of YOLO (you only look once), and the detection accuracy is significantly improved.

2. Image Recognition of Deep Learning Convolutional Network

2.1. Three-Dimensional Image Recognition Technology. At present, a small number of studies have established projection models with different aspect ratios and sizes in advance and then use radar to sense the depth information of the surrounding environment of the object, select the projection model that best fits the current scene, and project the 3D objects close to itself to the wall of the model. In space rather than on a two-dimensional plane, the display distortion of three-dimensional objects is reduced [10]. However, this method can only roughly estimate the surrounding environ-

ment of the object. When there are multiple three-dimensional objects in the environment, it cannot optimize the display of each three-dimensional object [11]. Combining the radar information and the binocular ranging principle, after jointly estimating the accurate three-dimensional information of the environment, a specific projection model conforming to the current scene is established and then projected [12]. This method can more accurately solve the display distortion problem of each large three-dimensional object, but the real-time performance is reduced due to the need to dynamically change the shape of the projection model, and the projection model with a special shape may also cause projection distortion to other scenes [13]. In the above two methods, the variable projection model is used to solve the problem of projection distortion, and the radar sensor is used to measure the depth of field, which not only causes the increase in system cost but also complicates the processing process [14]. Through analysis, it can be seen that the three-dimensional objects that workers are most concerned about in the process of field work are only other crops and pedestrians (the two are collectively referred to as objects of interest below), so this article combines the characteristics of the application scene and proposes an enhanced three-dimensional method. The synthesis method of panoramic images, combined with object detection and coordinate ascending and inverse mapping, renders the preprepared three-dimensional model of the object of interest on the estimated position, thereby solving the above problem [15].

With the widespread application of deep learning in the field of digital image processing, the use of deep convolutional neural network (DCNN) methods to restore and reconstruct digital images, feature extraction, target detection, and semantic segmentation has become the mainstream research direction of many scholars [16]. Digital image processing based on DCNN employs a data-driven learning and training mode, and the accuracy and effectiveness of the processing are highly dependent on the quality and category of the training set images [17]. For crop disease and insect pest images, due to the impact of the imaging environment, it is difficult to ensure that the acquired image data sets are at the same quality level. This will not only affect the processing effects of crop disease and insect pest images of other quality levels but also affect the processing of the overall business system effect [18]. The root cause of such problems is the imbalance of data quality [19]. The problem of data imbalance is also an important issue that deep learning algorithms pay attention to [20]. Aiming at the problem of unbalanced data quality, the existing methods generally deal with the problem of massively expanding the image data set, increasing the training time and the number of iterations [21]. This method finally obtains only an average value of the processing effect and does not improve the overall effect; the other way is to restore or superresolve the image of crop diseases and insect pests and then perform top-level processing after improving the image quality, although it has a certain effect. However, the reconstruction effect will still be affected by the amount of image data of heterogeneous crop diseases and insect pests [22].

2.2. Image Quality Classification Model of Crop Diseases and Insect Pests. The classification of image quality levels belongs to the category of objective image quality evaluation and is a complicated scientific problem [23]. Automatic evaluation using NIIRS standards is mainly accomplished through the following quality equations:

$$H_{nb} = \sum_{j=2}^{k} \sum_{h=1}^{j-1} G_{jh} \left(p_j s_h + p_h s_j \right) D_{jh},$$

$$D_{jh} = \frac{d_{jh} - P_{jh}}{d_{jh} + P_{jh}}, \tag{1}$$

$$RER_{jh} = \int_0^{\infty} dF_j(y) \int_0^y (y - x) dF_h(x),$$

where D represents the coefficient of the equation; there are different assignment versions: RER (relative edge response) represents the relative edge response; G (Convolved Gain) represents the noise gain of the imaging system postprocessing; H (Edge Overshoot) represents the edge after the system postprocessing overshoot factor [24]. Since it is difficult for general researchers to obtain GSD-related parameters of crop disease and insect pest images, the calculation of RER also requires the image to have appropriate edge shape characteristics, and the calculation method of SNR is not unique. The image quality equation is used to quality crop disease and insect pest images. Hierarchical classification has great limitations:

$$f(x) = \frac{1}{Nh} \sum_{i=1}^{N} k \left(\frac{X_i - x}{h} \right),$$

$$k(x) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x^2}{2} \right), \tag{2}$$

$$h_t = \tanh \left(w_c x_t + u_c (r_t \Theta h_{t-1}) + b_c \right),$$

$$h_t = z_t \Theta h_{t-1} + (1 - z_t) \Theta h_t.$$

There are relatively few research results on the use of deep learning to achieve image quality classification. From the perspective of blind image restoration, the literature divides close-up images into two simple types, clear and blurred, due to the lack of batch normalization (BN) layer, greatly reducing the generalization ability of the network:

$$\sigma t = \frac{\sqrt{1/n \sum_{i=1}^{n} (FI_{it} - FI_{it})^2}}{FI_{it}},$$

$$u_{(j|i)} = w_{ij} A_i, \tag{3}$$

$$s_j = \sum_i c_{ij} u_{(j|i)}.$$

From the perspective of quality evaluation, a multitask quality grade prediction DCNN method is proposed. This method uses NIIRS as the subjective quality label, con- structs a quality grade classification and regression net- work, and realizes 10 quality grade classifications. The feature extraction network structure of this method is deep, and the calculation efficiency is too low for the problem of coarse quality level classification; due to the lack of random loss of connections, it is also prone to overfitting:

$$\ln \left(\frac{FI_{it}}{FI_{it} - 1} \right) = \alpha + \beta \ln FI_{it} - 1 + v_i + \mathfrak{I}_t. \tag{4}$$

At the same time, there are considerable errors and difficulties in the calculation of subjective quality labels. The image quality classification of crop diseases and insect pests is a qualitative study of images, which is a typical mathematical classification problem. The feed- forward convolutional neural network structure based on supervised learning has a powerful classification function, which is very suitable for quality level classification using this type of network structure:

$$c_{ij} = \frac{e^{b_{ij}}}{\sum_k e^{b_{ik}}},$$

$$r = \frac{\alpha}{1 - \beta},$$

$$\theta = -\frac{1}{T} \ln (1 + \beta), \tag{5}$$

$$\tau = \frac{\ln (2)}{\theta}.$$

2.3. Basic Convolutional Neural Network Structure for Image Quality Classification. Our goal is to train a classi- fier through DCNN to classify images of crop diseases and insect pests into different levels according to their quality. The classification network takes the image of crop diseases and insect pests as input and outputs a label value representing the quality level of the image. Therefore, the entire network is a frame from large to small, from coarse to fine. Since the quality of the image of crop diseases and insect pests is not necessarily related to the size of the image, the input image can be taken into a relatively fixed size. First, the input sample image is preprocessed by zero score standardization. The pur- pose is to adjust the pixel value distribution of crop pest images to a nearly normal distribution and improve the ability to activate network training. Secondly, in order to alleviate the overfitting problem and reduce additional connection parameters, multiple BN layers are added to the middle layer, and the global average pooling layer is replaced with a fully connected layer commonly used in classifiers. In order to maintain the detailed characteristics of the image from large to small, the maximum pooling (MaxPooling) method is used for data scale compression; in order to improve the convergence speed while avoiding the disappearance of the gradient, a linear correction is added after the convolution (convolution, Conv) layer

unit ReLU function:

$$\ln\left(\frac{FI_{it}}{FI_{it}-1}\right) = \alpha + \beta \ln FI_{it} - 1 + \phi X_{it} - 1 + \nu_i + \tau_t,$$

$$k_{t1}[i] = \sum_j \cos\left(w_i^1, w_j^2\right),$$

$$BCRM_{jh} = \frac{\sum_{Z=1}^{h_j}\sum_{r=1}^{n_h}\left|y_{ji}-y_{hr}\right|}{n_j n_h (u_j + u_h)},$$

$$DCNN_t = \sum_{j=2}^{k}\sum_{h=1}^{j-1} G_{jh}\left(p_j s_h + p_h s_j\right) D_{jh}\left(1 - D_{jh}\right).$$

$$(6)$$

The above formula shows the DCNN architecture of quality classification, in which K is the number of BCRM, and BCRM is the abbreviation of BN layer, Conv layer, ReLU layer, and max-pooling layer. After conversion, it can be expressed by the following formula:

$$Gjj = \frac{1/2u_j\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\left|y_{ji}-y_{jr}\right|}{n_j^2},$$

$$Gw = \sum_{j=1}^{k} G_{jj}p_j s_j.$$

$$(7)$$

Convolution neural networks need to choose an appropriate loss function to transform the convolution parameter acquisition problem into an optimization problem. The label of data classification problem is often understood as a probability value, so its loss function is mainly set for this probability value. The most common loss function is the cross-entropy function, which has the advantage of faster weight update than the variance loss function. Assuming that X represents the input image and the network parameters to be trained, the DCNN quality classification network can be understood as a probability function for predicting the X quality index:

$$d_{jh} = \int_0^\infty dF_h(y)\int_0^y (y-x)dF_j(y),$$

$$G = G_w + G_{nb} + G_t.$$

$$(8)$$

In the formula, G is the number of samples in each batch, and the optimal classification level value of the segment quality level classification network is preset. Once it exceeds the preset level value, the convergence speed will be slow and the classification accuracy will be low.

2.4. Multistage Quality Classification.
Multisegment hierarchical classification is to use a shallow convolutional neural network, through adjusting the label, similar to the shape of the pyramid, from coarse to fine layer by layer classification,

can achieve fixed network depth, but not fixed number of levels of quality classification. Since any integer greater than 1 can be regarded as a linear combination of positive integers less than 1, any level of quality classification can be achieved by using a 2-level, 3-level, and 4-level quality classification network with less feature extraction layer. During the training process, the input for each segment of the network is the original training data and only a simple classification network is used to achieve the classification. In the process of testing, we need to use the judgment method to select the appropriate network structure path to achieve classification, but the judgment method is a convolutional network. The advantage of multistage network structure quality level classification is that the classification level can be set freely, the network structure is fixed, and the classification is realized layer by layer through internal circulation, so the training process is relatively simple.

The generalized camera internal parameters include the camera internal parameter matrix and the camera distortion parameters, which reflect the internal optical structure of the camera and are related to the characteristics of the camera itself. They are mainly responsible for the mapping process from the camera coordinate system to the pixel coordinate system. They are the mapping process from three-dimensional to two-dimensional, and also the conversion process from the physical length unit to the pixel length unit. The transformation of internal parameters of insect camera is as follows:

$$G = \frac{\sum_{j=1}^{k}\sum_{h=1}^{k}\sum_{t=1}^{n_j}\sum_{r=1}^{n_h}\left|y_{ij}-y_{hr}\right|}{2n^2 u},$$

$$(9)$$

where the subscript g is the camera coordinate system, K is the distortion parameter of the insect camera, and H is the coordinate point in the pixel coordinate system, which is generally half of the image length and width. Based on the above conversion principle, with the help of camera calibration chessboard, the internal parameter, external parameter matrix, and distortion parameters of four insect cameras are calibrated, respectively. The mapping projection model establishes a virtual object surrounding the environment, and the points on it are taken as the points in the world coordinate system. Through the above mapping relationship, the corresponding pixel coordinates (i.e., the points in the four images) are obtained, and the RGB information of the corresponding image points is rendered on the projection model to complete the synthesis of the basic three-dimensional panoramic image. Using this kind of projection model, the road surface can be projected to the model ground and the three-dimensional objects such as crops can be projected to the model wall in most cases, so as to ensure the correct projection relationship of the objects and avoid distortion.

It is difficult to optimize these two kinds of distortions at the same time simply by changing the size of the projection model. When the projection model is enlarged, the line bending distortion can be alleviated, but more crops, pedestrians, and other three-dimensional objects will fall into the ground area of the projection model, thus aggravating the object

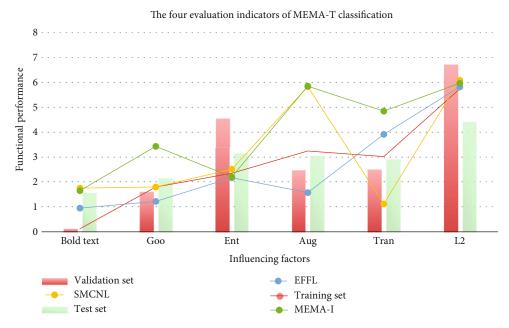The four evaluation indicators of MEMA-T classification

Figure 1: The four evaluation indicators of MEMA-T classification.

stretching distortion; on the contrary, although reducing the projection model can ensure the accuracy of most three-dimensional objects projected on the wall, it can not guarantee the straightness of the ground line. And in the process of dynamic adjustment of the model, the display screen is prone to jump, which affects the display effect.

3. Image Recognition Design of Crop Diseases and Insect Pests

3.1. Objects. In this paper, a DCNN-G model based on deep learning and Google data analysis is proposed. 640 data samples are trained by this model, and then, 5000 test samples are tested. The accuracy of this model is compared with the conventional recognition model. The 5000 images are cut into 256 pixel $\times$ 256 pixel images, and 80% of the training set and 20% of the data are selected as the test set. After one image of quality level 1 is degraded by using the above degradation parameters, nine quality level images are obtained. The images classified by quality level are tested by using the improved network YOLO-V4 of YOLO (you only look once).

3.2. Steps. A high-performance computing platform (TSMC server) and nvidi card are used for model training. The training parameters were set as batch size of 32, learning rate of 0.001, and turn of 1500. Use Adam network optimizer to update the weight parameters. The model training needs about 23,400 s, and the image preprocessing uses the MATLAB r2018b platform and Python 3.7. The classification results of SMCNL, EFFL, MEMA-I, and MEMA-T in the test data set containing 640 data samples are compared. It is difficult to distinguish the quality of images with different quality levels, especially those with adjacent levels, by subjective

observation with human eyes. The results show that both one segment and multiple segments can achieve a 10-grade image quality classification of crop diseases and pests. For the three models designed by the one-stage quality classification method, accuracy, recall, and precision of n-type are significantly improved compared with S-type, and the evaluation index of n-type is basically unchanged compared with C-type. The Gaussian blur and Gaussian noise of different scales added to the original data are artificially degraded, and they are marked as 2-10 levels, respectively, with a total of 5000 images. The 5000 images are then cut into 256 pixel $\times$ 256 pixel images, and 80% of the training set and 20% of the data are selected as the test set.
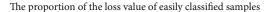
(1) Image preprocessing

Color normalization in the process of staining and scanning of aortic wall tissue samples, due to the different laboratory conditions of staining, parameter settings of a digital scanner, and illumination, the color differences of digital pest images are often caused. Color normalization can not only ensure the color consistency of the image but also preserve the biological information in the pest image, so as to improve the classification performance of the model. A normalization method based on staining separation is used to reduce the color difference of histopathological images, and the structure information in the images is saved as much as possible by generating the staining density map.

(2) 3D panorama of crop enhanced image

The adaptive projection model method is easy to take one thing into account and lose the other when solving the distortion problem. In view of the problem of display distortion and the shortage of current solutions, this paper proposes

TABLE 1: The GoogLeNet model achieved classification performance comparable.

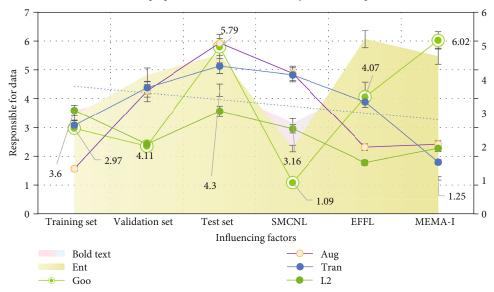| Item	Training set	Validation set	Test set	SMCNL	EFFL	MEMA-I
Bold text	0.12	0.15	1.59	1.75	0.95	1.65
Goo	1.8	1.63	2.17	1.8	1.22	3.43
Ent	2.35	4.56	3.15	2.51	2.17	2.2
Aug	3.24	2.48	3.05	5.82	1.57	5.84
Tran	3.02	2.51	2.91	1.12	3.91	4.84
L2	5.73	6.71	4.4	6.08	5.81	5.96

FIGURE 2: The proportion of the loss value of easily classified samples.

an enhanced 3D panoramic image synthesis method. With the help of the YOLO detection network and the coordinate dimension elevation inverse mapping method proposed in this paper, the 3D objects around the agricultural objects are represented in the way of virtual synthesis without dynamically changing the projection model. This paper will focus on how to detect the object of interest and estimate its accurate world space position through the image. The steps are as follows: firstly, the original image captured by four cameras is taken as the input of the object detection network, and all the interested objects are identified and selected; secondly, the object's position in the world coordinate system is estimated by using the object's pixel coordinates through the coordinate dimension ascending inverse mapping method proposed in this paper; and then, after a series of data processing for the estimated position, the accurate object is obtained. Finally, the general model of crops and pedestrians built in advance is placed and rendered on the estimated position, which will overlay and display the three-dimensional objects that are wrongly mapped on the ground. This method can reduce the line bending distortion as much as possible while solving the object stretching distortion and make the display more natural while providing accurate object position information.

TABLE 2: GoogLeNet model improves training efficiency.

Item	Bold text	Goo	Ent	Aug	Tran	L2
Training set	3.6	2.97	2.94	1.58	3.08	3.6
Validation set	4.11	2.35	4.14	4.28	4.39	2.42
Test set	4.3	5.79	4.73	5.93	5.13	3.57
SMCNL	3.16	1.09	1.95	4.88	4.83	2.96
EFFL	4.36	4.07	5.2	2.32	3.9	1.79
MEMA-I	1.25	6.02	4.69	2.42	1.81	2.28

4. Image Recognition and Analysis of Crop Diseases and Pests Based on Deep Learning

4.1. Improvement of Data Imbalance by Deep Learning. As shown in Figure 1, it is found that the four evaluation indexes of MEMA-T classification have been effectively improved, which indicates that data enhancement helps to improve the classification performance of pest images with data imbalance. L2 regularization is used to eliminate the overfitting phenomenon in the training process. This paper uses focal loss instead of cross-entropy loss function to solve the problem of data imbalance. To assess the effectiveness of
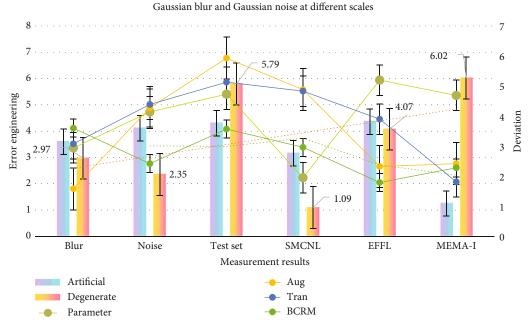
Figure 3: Gaussian blur and Gaussian noise at different scales.

Table 3: Use image information entropy as standard.

Item	Artificial	Degenerate	Parameter	Aug	Tran	BCRM
Gaussian blur	3.6	2.97	2.94	1.58	3.08	3.6
Gaussian noise	4.11	2.35	4.14	4.28	4.39	2.42
Test set	4.3	5.79	4.73	5.93	5.13	3.57
SMCNL	3.16	1.09	1.95	4.88	4.83	2.96
EFFL	4.36	4.07	5.2	2.32	3.9	1.79
MEMA-I	1.25	6.02	4.69	2.42	1.81	2.28

Table 4: For multistage quality classification.

Item	ResNet	MnasNet	SMCNL	EFFL	MEMA-I	AlexNet	GoogLeNet
MD	0.62	1.56	0.12	0.15	1.59	1.75	0.95
EFFL	1.03	1.04	1.8	1.63	2.17	1.8	1.22
MEMA-I	3.06	4.38	2.35	4.56	3.15	2.51	2.17
MEMA-T	5.18	3	3.24	2.48	3.05	5.82	1.57
VGG16	2.16	3.77	3.02	2.51	2.91	1.12	3.91
AlexNet	5.25	5.02	5.73	6.71	4.4	6.08	5.81

the different methods in the model, we used the mean classification accuracy of the four types of lesions, the training set, the validation set, and the test set, as metrics.

The simplified GoogLeNet model achieves the same classification performance as the original model, as shown in Table 1. Among them, bold text, Goo, Ent, Aug, Tran, L2, and FOC represent the best performance, respectively. The simplified GoogLeNet model has cross-entropy loss function, data enhancement, transfer learning, L2 regularization, and focal loss function. Both data enhancement and transfer learning can improve the classification performance of the

model, and the combination of the two methods can obtain higher classification accuracy than using them alone. The use of L2 regularization can restrain overfitting to a certain extent and enhance the stability of model fitting function.

As shown in Figure 2, focal loss can further improve the multiclassification performance by reducing the proportion of loss values of easy to classify samples and forcing training for difficult to classify samples. In the pathological data of this paper, MEMA-I and MEMA-T are two similar types of lesions, which are difficult to distinguish. Compared with the original GoogLeNet model, the classification accuracy
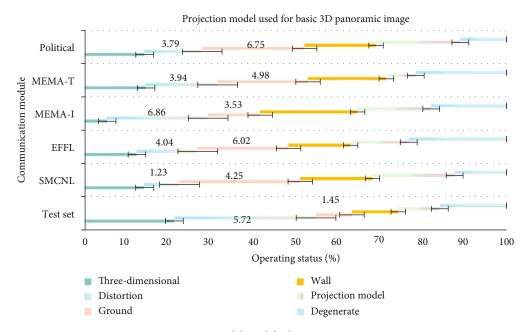
Figure 4: Projection model used for basic 3D panoramic image.

Table 5: The parameter combination is set according to the results of previous work.

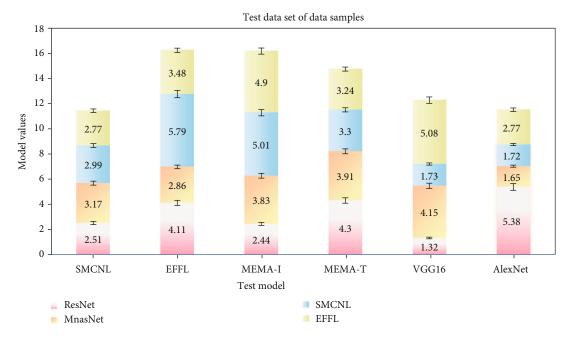| Item	Test set	SMCNL	EFFL	MEMA-I	MEMA-T	Political
Three-dimensional	3.61	2.09	3.44	1.5	3.36	3.92
Distortion	5.72	1.23	4.04	6.86	3.94	3.79
Ground	1.45	4.25	6.02	3.53	4.98	6.75
Wall	1.87	2.54	4.12	6.62	4.31	4.7
Projection model	1.68	2.89	3.85	5	1.62	5.55
Degenerate	2.7	1.83	6.5	5.12	5.01	3.05

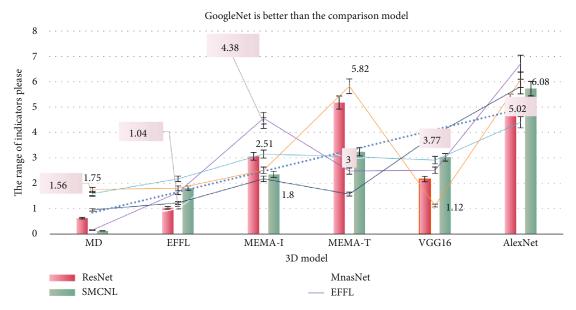Figure 5: Test data set of data samples.

FIGURE 6: GoogLeNet is better than the comparison model.

TABLE 6: Residual learning to improve the classification performance of the model.

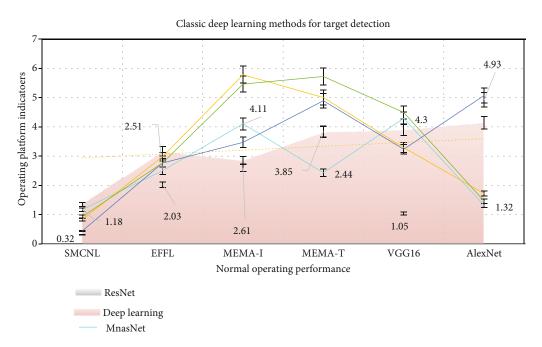| Item	ResNet	MnasNet	SMCNL	EFFL	MEMA-I	MEMA-T
SMCNL	2.51	3.17	2.99	2.77	2.78	3.12
EFFL	4.11	2.86	5.79	3.48	5.47	2.53
MEMA-I	2.44	3.83	5.01	4.9	5.73	5.32
MEMA-T	4.3	3.91	3.3	3.24	4.5	2.34
VGG16	1.32	4.15	1.73	5.08	1.47	1.88
AlexNet	5.38	1.65	1.72	2.77	2.22	6.43

FIGURE 7: Classic deep learning methods for target detection.

of simplified GoogLeNet for MEMA-I and MEMA-T is improved by about 4%.

As shown in Table 2, the simplified GoogLeNet model is used to improve the training efficiency, and the data enhancement, migration learning, and L2 regularization are effectively used to further improve the network performance. Focal loss is used to solve the problem of image data imbalances in mesomorphic diseases and pests. The experimental results show that the simplified GoogLeNet model is superior to the commonly used deep learning model in MD classification performance. This study provides a new idea for deep learning technology to classify noninflammatory aortic media lesions based on pest images. In the future work, we will get more and more kinds of pest images to further improve the generalization ability of the model.

The Gaussian blur and Gaussian noise of different scales added to the original data are artificially degraded, which are, respectively, marked as 2-10 levels, with a total of 5000 pieces. The artificial degradation parameters are shown in Figure 3. The 5000 images are then cut into 256 pixel × 256 pixel images, and 80% of the training set and 20% of the data are selected as the test set. After one image of quality level 1 is degraded by using the above degradation parameters, nine quality level images are obtained. It is difficult to distinguish the quality of images with different quality levels, especially the images with adjacent levels, by subjective observation with human eyes. Using the algorithm model proposed in this paper, the recognition accuracy is 95%, which is much higher than 84% of the basic DCNN model.

Because the imaging parameters of the above images can not be obtained, the NIIRS quality equation can not be used for classification. The image information entropy is used as the standard for classification and evaluation. It is also difficult to distinguish the quality level of 30 images selected from the test set with the quality level labels of 1, 5, and 10 through the value of information entropy and curve, as shown in Table 3.

In the DCNN architecture of quality classification, there are three kinds of network models. The first one is called normal type (referred to as n-type). There are three BCRM layers in the middle of the network; the second one is called simple type (referred to as S-type). Compared with n-type, there is only one BCRM layer in the middle of the network. The third type is called complex type (C-type for short). Compared with n-type, there are five BCRM layers in the middle of the network. For the multistage quality classification, the 2-fork structure is adopted in all layers, and the n-type network structure is adopted in all layers. The input image size of crop diseases and insect pests is 256×256, the number of iterations is set to 200, and the network parameters are shown in Table 4.

4.2. Image Recognition Effect Analysis of Training DCNN-G. As shown in Figure 4, because the projection model used in the basic three-dimensional panoramic image is established in advance according to experience, it is impossible to adapt to all scenes, so when the real three-dimensional environment of the crop does not match the projection

Table 7: Method is limited to certain types of resolution data.

Item	ResNet	MnasNet	Deep learning	Remote sensing	Crop	Pests and diseases
SMCNL	0.32	1.18	1.35	0.84	0.45	0.94
EFFL	2.03	2.51	3.17	2.99	2.77	2.78
MEMA-I	2.61	4.11	2.86	5.79	3.48	5.47
MEMA-T	3.85	2.44	3.83	5.01	4.9	5.73
VGG16	1.05	4.3	3.91	3.3	3.24	4.5
AlexNet	4.93	1.32	4.15	1.73	5.08	1.47

model created by the simulation, the display will be distorted.

As shown in Table 5, the best training effect can be obtained by manually adjusting the parameters of L2 regularization ($\lambda = 0.001$). Two key parameters involved in focal loss are $\alpha = 0.25$ and $\gamma = 2$. The parameter combination is set according to the results of previous work to ensure the best classification performance of the model.

As shown in Figure 5, the classification results of SMCNL, EFFL, MEMA-I, and MEMA-T in the test data set contain 640 data samples, in which the bold text represents the best result. Two samples of SMCNL and EFFL were misclassified into other types. The model has good recognition and resolution ability for SMCNL and EFFL. Compared with SMCNL and EFFL, MEMA-I and MEMA-T have more misclassification samples, especially about 16% of them are identified as MEMA-I. The reason is that the two types of lesions show similar pathological features. In addition, due to the serious imbalance of data, MEMA-T data samples are less, resulting in poor classification results.

As shown in Figure 6, all models adopt the same data enhancement, migration learning, L2 regularization, and focal loss. Through observation, it is found that the improved GoogLeNet is better than the comparison model in MD classification performance. EFFL got the best results among all the classification indexes. MEMA-I achieves the best results in accuracy, accuracy, and F value. SMCNL has the highest accuracy and precision. Due to the imbalance of data and the small number of MEMA-T samples in the test data set, the MEMA-T classification results are inadequate.

As shown in Table 6, the ResNet model also performs well. It has an 18 layer network structure, which is close to the network depth of the GoogLeNet model used in this paper. In addition, it introduces residual learning to improve the classification performance of the model. However, due to the large difference of lesion size among different lesion types of MD, GoogLeNet introduces convolution kernels of different sizes, which can better extract the context information of MD, so its performance is better than ResNet. Compared with other networks, the performance of AlexNet is poor. The main reason is that AlexNet is a shallow network with an eight-layer network structure, which reduces the accuracy of MD classification tasks due to its limited ability to capture
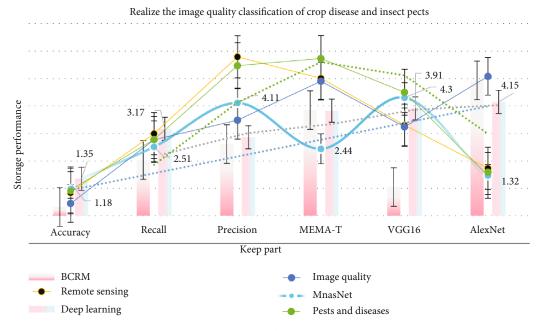
FIGURE 8: Realize the image quality classification of crop diseases and insect pests.

TABLE 8: Training and detection of crop disease and insect pest image data.

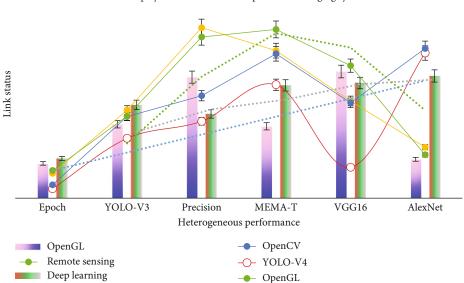| Item	BCRM	MnasNet	Deep learning	Remote sensing	Image quality	Pests and diseases
Accuracy	0.32	1.18	1.35	0.84	0.45	0.94
Recall	2.03	2.51	3.17	2.99	2.77	2.78
Precision	2.61	4.11	2.86	5.79	3.48	5.47
MEMA-T	3.85	2.44	3.83	5.01	4.9	5.73
VGG16	1.05	4.3	3.91	3.3	3.24	4.5
AlexNet	4.93	1.32	4.15	1.73	5.08	1.47

TABLE 9: The average value of mAP detected by 3 types of different data targets.

Item	Epoch	YOLO-V3	Precision	MEMA-T	VGG16	AlexNet
YOLO-V4	0.03	1.82	0.28	0	1.09	1.89
OpenGL	3.76	2	1.94	3.9	1.05	1.1
Deep learning	3.46	3.64	2.68	3.22	5.24	3.15
Remote sensing	5.44	3.26	3.94	4.01	3.1	5.82
OpenCV	2.16	2.35	4.16	1.65	3.18	3.82
OpenGL	5.89	6.34	2.22	6.89	2.57	6.38
SMCNL	1.01	3.3	3	6.43	6.87	1.62

image features of diseases and pests. The comparison results show that the improved GoogLeNet model can recognize specific histopathological changes from the images of diseases and insect pests and help to improve the automatic classification of four types of MD lesions. The results show that the classification results of MEMA-T are poor, mainly due to the small number of samples and serious data imbalance.

4.3. Use YOLO-V4 to Test and Analyze the Model after Training. As shown in Figure 7, the model tested by

YOLO-V4 can not only realize the classification of fixed number of grades but also realize the classification of non-fixed number of grades, and the classification results are more detailed and accurate. After the preprocessing of quality level classification, the classical deep learning method is used for target detection, and the detection accuracy is significantly improved, which can effectively solve the problem of imbalanced data quality in the training set. A convolution neural network is used to classify the image quality of crop diseases and insect pests, which not only expands the

The display effect of the basic 3D panoramic imaging system

FIGURE 9: The display effect of the basic 3D panoramic imaging system.

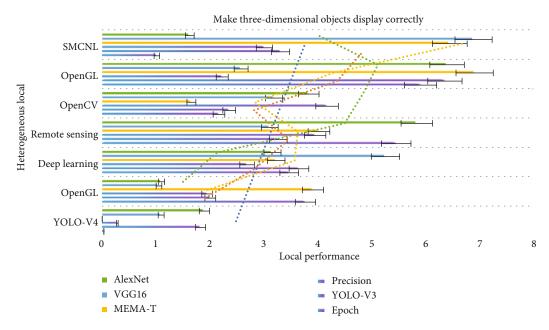Make three-dimensional objects display correctly

FIGURE 10: Make three-dimensional objects display correctly.

application field of deep learning but also provides a new method for image quality evaluation of crop diseases and insect pests.

As shown in Table 7, many image processing methods of crop diseases and insect pests are based on the same image quality level, such as crop disease and insect pest image registration, landmark detection, and target recognition; generally, different data sources, different spatial resolution, and different spectral resolution data are distinguished, and different technical means are used for processing, and even some methods are used. The method is limited to some kind of resolution data. Another example is image restoration; many methods are known that the image is degraded under

the premise of restoration processing. This makes people have to design different methods or adjust the corresponding parameters for different quality levels of images in order to obtain the corresponding processing effect, which is obviously inconsistent with the technical requirements of the fully automated intelligent era. The quality classification of crop disease and insect pest images can not only provide important prior information for understanding crop disease and insect pest images but also provide scientific basis for testing the imaging ability of sensors and objectively evaluating the image quality of crop diseases and insect pests.

As shown in Figure 8, both one segment and multisegment can well achieve 10 levels of image quality classification

(a)   (b)

FIGURE 11: Sample image of processing sample (camera live shooting, left (a), right (b)).

of crop diseases and insect pests. For the three models of one-stage quality classification, accuracy, recall, and precision of n-type are significantly improved compared with S-type, and the evaluation indexes of n-type are basically unchanged compared with C-type, which indicates that BCRM has an impact on the results, and it also means that n-type can fully meet the application requirements for 10 quality classification. For the multistage quality level classification, it also achieves the effect of one-stage classification and does not need to estimate the number of BCRM. Each stage adopts the n-type classification structure of 10 quality levels, which can fully ensure the accuracy of two quality levels in each stage and further reduce the error accumulation. However, the training time and detection time are relatively long, which is the reason, that closely related to the number of parameters and the complexity of intermediate computation.

The original 610 image data of agricultural crop diseases and pests are directly trained and detected, as shown in Table 8. Quality classification refers to training and testing according to different quality classes. That is, the training set and the test set belong to the same quality class. The number ratio of the training set data and test set data is 4 : 1, and the test set data is not in the training set. Because the similar quality level has little influence on the target detection, the images of crop diseases and insect pests of levels 1-4, levels 5-8, and levels 9-10 are classified into the same quality class in the experiment, and the final map value is the average value of map of these three types of different data target detection, and the results are shown in Table 9.

As shown in Figure 9, the display effect in the same perspective and scene is compared when the method is used and not used. It shows the display effect of the basic three-dimensional panoramic image system. When the pests and diseases are in the ground area of the projection model, the system will cause projection distortion to these objects, so that the workers can not accurately obtain the position information of the objects around the objects, or even can not see

the objects clearly. Using the enhanced three-dimensional panoramic image synthesis method in this paper, the images of diseases and pests are clearly presented in the scene through the three-dimensional model, which can greatly facilitate workers to observe.

As shown in Figure 10, compared with the existing solutions, this method does not need to reduce the projection model to make the three-dimensional object display correctly, so it can maximize the ground part of the model, so it can better weaken the line bending distortion, that is, to maintain the flatness of the road as far as possible.

As shown in Figure 11, (a) shows the display effect of this method. The projection model size is set to 3000 mm, and the light is basically straight by observing the vertical line. However, the existing solution (adaptive model method) can only make 3D objects mapped on the wall by reducing the projection model, as shown in (b). At this time, the size of the projection model is 700 mm; although the lateral pests are well displayed, the image obviously causes great bending to the road.

5. Conclusions

This paper first introduces the basic three-dimensional panoramic image synthesis algorithm and analyzes the causes of its display distortion, then proposes the enhanced three-dimensional panoramic image synthesis method, and finally verifies the performance indicators of this method through experiments. Based on the existing algorithms of 3D panoramic image synthesis of pests and diseases, this paper proposes an enhanced 3D panoramic image synthesis method of pests and diseases based on coordinate ascending inverse mapping, which is used to solve the display distortion problem of 3D objects in the original system. Firstly, the position of the object of interest in each image is detected by using the YOLO-V4 network, and then, the inverse mapping method of coordinate elevation from the pixel coordinate system to

the world coordinate system is derived by using the insect camera calibration parameters combined with the supplementary conditions, so as to preliminarily estimate the position of the object of interest in the world coordinate system. Then, by merging and filtering the estimated positions, you get the final estimated position, place the preset model in the corresponding position, and render it to complete the highlighting. Compared with the existing solutions, this method has many advantages, such as low cost, less computation, and good display effect. At the same time, the generation speed meets the real-time requirements, and the position estimation also meets the accuracy requirements. It can further improve the display quality and use value of panoramic images of diseases and pests. In the future, the experiment can be extended to real pests to test the display performance of the system on real roads.

Due to the need to display the surrounding conditions of diseases and pests in real time in the process of field work, it is required that the screen should not appear the phenomenon of jam, so it puts forward high requirements for the real-time performance of the image generation method in this paper, so the real-time performance of this method will be verified first. Then, the accuracy of object detection and position estimation in this method is quantitatively tested. The scene generation is divided into two parts: using OpenCV to generate the ground part and using OpenGL to render the scene part. When OpenCV is used to generate the ground point by point, the LUT look-up table method is used in the former. The calculation of the mapping relationship is completed and saved when the program is initialized, and only look-up table is needed in real-time rendering, thus reducing the processing time of a single frame.

Compared with the existing solutions, this method has many advantages, which can intuitively and prominently display the position of other crops and pedestrians around the object. When the object of interest is displayed, it will not cause projection distortion to other scenes, and the overall view is comfortable. In addition, this model does not need other depth sensors. Because a fixed projection model can be used, LUT can be used to ensure real-time performance. The algorithm in this paper can keep the straightness of the ground lines to the maximum in the display. In conclusion, this method has a high application value.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] T. Han, "Design and application of multicolor image identification in soil pollution component detection," *Arabian Journal of Geosciences*, vol. 13, no. 18, 2020.

[2] B. Chen, H. Li, W. Luo, and J. Huang, "Image processing operations identification via convolutional neural network," *SCIENCE CHINA Information Sciences*, vol. 63, no. 3, 2020.

[3] T.-K. Chui, J. Tan, Y. Li, and H. A. Raynor, "Validating an automated image identification process of a passive image-assisted dietary assessment method: proof of concept," *Public Health Nutrition*, vol. 23, no. 15, pp. 2700–2710, 2020.

[4] Z. Zhao and N. Liu, "the recognition and localization of insulators adopting SURF and IFS based on correlation coefficient," *Optik*, vol. 125, no. 20, pp. 6049–6052, 2020.

[5] Z. Zhao, N. Liu, and L. Wang, "Localization of multiple insulators by orientation angle detection and binary shape prior knowledge," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 22, no. 6, pp. 3421–3428, 2019.

[6] Q. Wu and J. An, "an active contour model based on texture distribution for extracting inhomo geneous insulators from aerial images," *IEEE Transactions on Geoscience & Remote Sensing*, vol. 52, no. 6, pp. 3613–3626, 2019.

[7] Y. Zhai, R. Chen, Q. Yang, X. Li, and Z. Zhao, "Insulator fault detection based on spatial morphological features of aerial images," *IEEE Access*, vol. 122, no. 6, pp. 35316–35326, 2019.

[8] X. Miao, X. Liu, J. Chen, S. Zhuang, J. Fan, and H. Jiang, "Insulator detection in aerial images for transmission line inspection using single shot multibox detector," *IEEE Access*, vol. 125, no. 7, pp. 9945–9956, 2019.

[9] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, and D. Xu, "Detection of power line insulator defects using aerial images analyzed with convolutional neural networks," *IEEE Transactions on Systems Man & Cybernetics Systems*, vol. 50, no. 4, pp. 1486–1498, 2018.

[10] Y. W. Tian, I. Cheng, and X. Q. Wang, *Feature vectors determination for pest detection on apples based on hyperspectral imaging*, Transactions of the Chinese Society of Agricultural, 2014.

[11] *Engineering (Transactions of the CSAE)*, vol. 30, no. 12, pp. 132–139, 2019.

[12] R. P. Haff, S. Saranwong, W. Thanapase, A. Janhiran, S. Kasemsumran, and S. Kawano, "Automatic image analysis and spot classification for detection of fruit fly infestation in hyperspectral images of mangoes," *Postharvest Biology and Technology*, vol. 86, no. 1, pp. 23–28, 2013.

[13] D. E. Purcell, M. G. O'Shea, R. A. Johnson, and S. Kokot, "Near-infrared spectroscopy for the prediction of disease ratings for leaf gall in sugarcane clones," *Applied Spectroscopy*, vol. 63, no. 4, pp. 450–457, 2009.

[14] R. Teixeira, J. I. R. Fernandéz, J. Pereira, and L. B. Monteiro, "Identification of grapholita molesta (Busk) (Lepidoptera: Tortricidae) biotypes using infrared spectroscopy," *Neotropical Entomology*, vol. 44, no. 2, pp. 129–133, 2019.

[15] L. Z. Cheng, X. C. Zhu, and L. L. Gao, "Hyperspectral estimation of phosphorus content for apple leaves based on the random forest model," *Journal of Fruit Science*, vol. 33, no. 10, pp. 1219–1229, 2019.

[16] S. J. Liu, Y. W. Tian, and D. Feng, "Nondestructive detection method of hyperspectral imaging for apple disease," *Journal*

of Shenyang Agricultural University, vol. 47, no. 5, pp. 634–640, 2016.

[17] R. Van Beers, B. Aernouts, L. L. Gutiérrez et al., "Optimal illumination-detection distance and detector size for predicting Braeburn apple maturity from Vis/NIR laser reflectance measurements," *Food and Bioprocess Technology*, vol. 8, no. 10, pp. 2123–2136, 2019.

[18] X. L. Wang, "Extraction and modeling of physical and chemical properties of fruit trees flowers and leaves based on spectral analysis technique," *Beijing Forestry University,2017WU Fei, LU Xinyan, SONG Jun. Learning of multimodal representations with random walks on the click graph, IEEE Trans. Image Process*, vol. 25, no. 2, pp. 630–642, 2019.

[19] L. U. Xiaoqiang, W. A. N. G. Binqiang, and Z. H. E. N. G. Xiangtao, "Exploring models and data for remote sensing image caption generation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 4, pp. 2183–2195, 2018.

[20] N. MERKLE, S. AUER, R. MULLER, and P. REINARTZ, "Exploring the potential of conditional adversarial networks for optical and SAR image matching," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 6, pp. 1811–1820, 2018.

[21] J. T. Barron and J. Malik, "Shape, illumination, and reflectance from shading," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 8, pp. 1670–1687, 2019.

[22] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2018.

[23] X. Li, Y. Dong, P. Peers, and X. Tong, "Modeling surface appearance from a single photograph using self-augmented convolutional neural networks," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, pp. 1–11, 2017.

[24] D. Guarnera, G. C. Guarnera, A. Ghosh, C. Denk, and M. Glencross, "BRDF representation and acquisition [C]//computer graphics," *Forum*, vol. 35, no. 2, pp. 625–650, 2019.