Corn Diseases Recognition Method Based on Multi-feature Fusion and Improved Deep Belief Network

Xiangpeng Fan
Xinjiang University
Jianping Zhou (linkzhoujp@sina.com)
Xinjiang University
Yan Xu
Xinjiang University
Jingjing Yang
Xinjiang University

Research Article

Keywords: corn leaf diseases, deep learning, K-means cluster, features extraction, feature fusion, deep belief network, particle swarm optimization algorithm

Posted Date: May 3rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-295393/v2

Corn Diseases Recognition Method Based on Multi-feature Fusion and Improved Deep Belief Network

Xiangpeng Fan [1,2] , Jianping Zhou [1,2,3*], Yan Xu [1,2] , Jingjing Yang [4]

[1] School of Mechanical Engineering, Xinjiang University, Urumqi, 830047, China

[2] Xinjiang Agricultural Robot and Intelligent Equipment Engineering Research Center, Urumqi,, 830047, China

[3] National Key Laboratory of Mechanical Manufacturing System Engineering, Xinjiang University, Urumqi, 830047, China

[4] School of Chemistry, Xinjiang University, Urumqi, Xinjiang, 830047, PR China

* Correspondence : linkzhoujp@sina.com

Abstract: The automatic monitoring timely and accurately of crop diseases has become an important research field in precision agriculture. Aiming at the low application rate of crop disease identification models in real field environment, a disease identification method based on multi-feature fusion and improved deep belief network was proposed. We Obtained representative samples in field conditions, then we augmented the data set. K-means clustering segmentation and morphological corrosion processing were utilized to obtain segmentation maps with clear boundaries and low noises. Then color features, shape features and textures of disease images were extracted respectively and they were fused to normalize as input data. A corn disease recognition model based on deep belief network was designed, using labeled and unlabeled dual hidden layer network structure to investigate the DBN hidden layer node combination mode. We obtained the optimal hidden layer node number combination method for disease classification: [26,85,29,4]. The accuracy of optimal DBN was 92.79%. On this basis, the deep belief network recognition model was optimized by particle swarm optimization algorithm for further performance enhancement. The experiment indicated that recognition effect using multi-feature fusion as input vectors was better than a single feature. The updated PSO-DBN reached the accuracy of 96.65%, which had a faster convergence speed and higher recognition accuracy of 3.86% than the standard DBN. Compared with state-of-the-art methods including SVM, ANN and CNN models, the proposed method can effectively dig deep digital features of disease areas or lesions and has the best performance, which could meet the needs of intelligent identification of field diseases.

Key words: corn leaf diseases; deep learning; K-means cluster; features extraction; feature fusion; deep belief network; particle swarm optimization algorithm.

1. Introduction

As the climate change and environmental deterioration, crop diseases' frequent occurrence has become a severe obstacle of high production and quality [1]. The Food and Agriculture Organization (FAO) of the United Nations estimated that pests and diseases led to the loss of 20–40% of global

food production, constituting a threat to food security. The preventive measures cannot guarantee the complete disappearance of crop diseases while pesticides could minimize the damage and preserve yields. However, the utilization of pesticides is not environmentally harmless. The precise spraying of pesticide can effectively save 46-60% pesticide dosage without affecting the disease prevention and control effect [2]. Early and accurate diagnosis is a critical first step in mitigating losses caused by plant diseases. An incorrect diagnosis can lead to improper management decisions, such as selection of the wrong chemical application that could potentially result in further reduced crop health and yield. [3]. Timely and accurate knowledge of fields' phytosanitary conditions especially the disease types is a decisive factor in limiting the use of pesticides and protecting harvests. Therefore, how to automatically and accurately identify crop diseases is of great significance [4]. However, assessing the healthiness of fields is not simple, and it requires a high level of expertise and manual scouting large areas to diagnose the disease based on human experience is inefficient and restricted [5]. Indeed, a disease can be expressed differently from one plant species to another, or even from one variety to another. Even nutritional deficiencies and pests can produce symptoms similar to those of some diseases [6]. The automatic identification of diseases by imagery has the potential to solve all these issues by using automatic computer vision or artificial intelligence tools [7]. Computer vision technology has low cost and convenient operation, which has been applied widely in precision agriculture [8–13].

In the automation recognition field of precision agriculture, the fusion of color, shape and texture features has been widely used [14–19]. Their results showed that the fusion features can avoid the limitation of single feature and further improve the accuracy and robustness of disease recognition. The selection of classifiers is another emphasis of disease intelligent recognition. In the issue of classifiers, the application of relevant machine learning algorithms in the diagnostic model has achieved good results, mainly including support vector machine(SVM) [20–22], artificial neural network [1], clustering analysis [23] et al. These traditional intelligent diagnostic models mostly complete the fitting and classification of functions in one or two layers of models, which belong to the algorithm of shallow structure learning and fail to fully explore the internal characteristics of crop spot image data. In addition, when processing the feature vectors with higher dimension or more training samples, the convergence speed is slow and the accuracy is low, so their practicability and applicability are obviously insufficient. Deep learning (DL) method is a deep structure algorithm emerging in recent years, which can learn the deep features of image data. It has made brilliant achievements in image classification field et al. [24,25] , and it's application on intelligent recognition in precision agriculture is also on the rise [26–28]. By introducing DL into the field of intelligent identification of agricultural diseases and improving its identification accuracy, the practicability of the deep learning method in the field of intelligent diagnosis of diseases and pests will be further improved.

As one of the classical models of deep learning-convolutional neural networks (CNNs) has achieved good results in precision agriculture involving fruit recognition and crop disease detection et al. [29–31] . The advantage of CNN is that the original image data can be directly input to the model without manual extraction of features [32,33]. However, CNNs cost much time to adjust and modify

the parameters and CNNs count heavily on high-power graphics processing unit (GPU) and high hardware resources. To get good results, a huge number of samples (hundreds of thousands or even millions) are needed to train the model, while it is unrealistic to get so many samples. Therefore, aiming at the small sample and high pixel image data set of corn field in this paper, this study adopts the method of extracting image features first and then modeling, and uses another typical deep belief network (DBN) model, which requires less computer hardware resources and can be used as a feature classifier alone. DBN model firstly is conducted with greedy unsupervised pre-training on the network layer by layer, and then it is fine-tuned through the back-propagation algorithm, which is especially suitable for learning problems with a limited number of labeled samples. DBN has the ability of hierarchical feature learning and expression, which can automatically extract the internal features of data and exploring and analyzing the complex nonlinear relationship between input and output [34]. There is a complex nonlinear relationship between image characteristics and crop disease types. DBN is expected to be an effective tool to solve this problem.

The contributions of this paper are as follows: in this work, we proposed a effective method for corn disease recognition based on improved deep belief network which could reduce the training time and improve the recognition accuracy. The method of selecting parameters of the DBN and suitable training model structure were investigated. The proposed network was trained by fusion features obtained from field samples and then the model was optimized by PSO algorithm to optimize the parameters. Through multi-feature fusion, the problem of low accuracy of single feature of diseases was solved, and a PSO-DBN disease recognition model was established to overcome the shortcomings of shallow structure algorithm (SVM, BP et al.) with poor classification effect which also avoided the dependency of high-powered hardware condition.

2. Materials and Methods

This section introduces the materials and method that carried out in the experiment. Figure 1 illustrates the basic construction process of corn disease system in this investigation. The whole process includes data acquisition and preprocessing, image segmentation, feature extraction and fusion, model training and model testing.

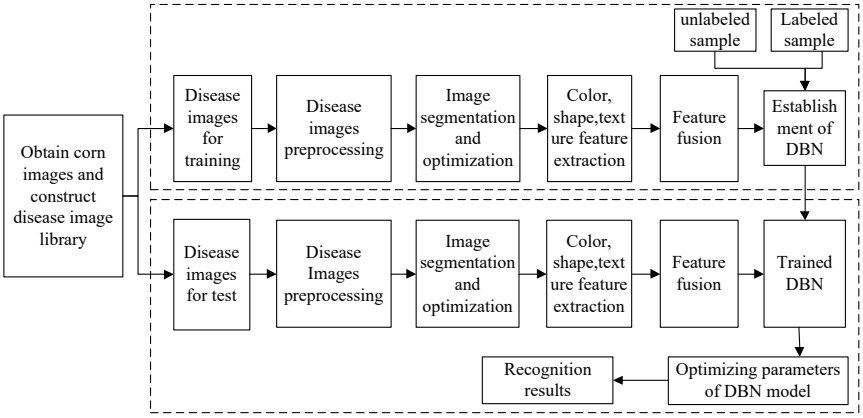


Figure 1. The process diagram of disease recognition

2.1 Image acquisition and pre-processing

The image data set needed for the experiment was captured from the corn field of Urumqi West mountain farming, Urumqi, Xinjiang, China. We utilized Huawei Honor 7 Plus smart phone (camera version: TRT-AL00A, resolution 3016 pixels×3016 pixels) to obtain images of four common corn diseases caused by fungi with various backgrounds characteristics from June 3 to 24, 2019. In order to take full consideration of the weather conditions of the natural scene, the images were collected in sunny and cloudy days respectively. It is required to reflect many complex situations including multiple background features and multiple angles in the practical application to effectively guarantee the validity of the test results. After screening, there were 200 images meeting the requirements. The original disease images database is established by using the obtained 200 images. Since the images taken had no labels and semantics, we labeled them with the help of crop pathologist. Some examples of the original sample are shown in Figure 2.

Figure 2. Some examples of corn disease images

Though various growth status of corn leaves and background factors had been considered in filming process, there was still lack of effective representative. Thus, the data augmentation technique was used to amplify the data set. Specifically, image affine transformation, perspective transformation, color dithering, adjusting the brightness and contrast, adding noise of image processing were applied to augment. The corresponding data set was extended to 8 times of original images, and the diversity of the data increased as well, which would reduce the distortion caused by lack of sample in the course of training or fitting problem. The augmented relevant sample number statistics of four different corn diseases leaves is shown in Figure 3.

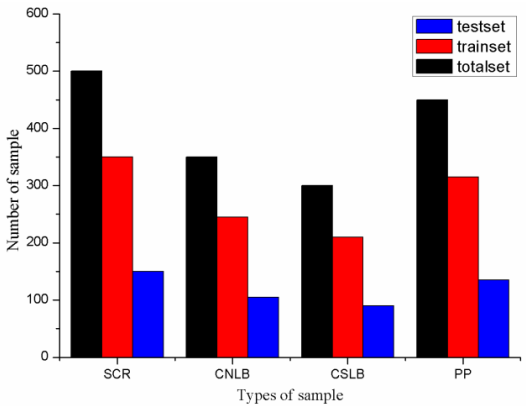
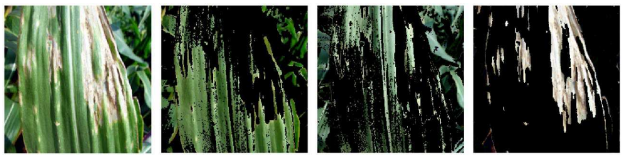


Figure 3. The Sample Statistics of Corn Disease Leaves

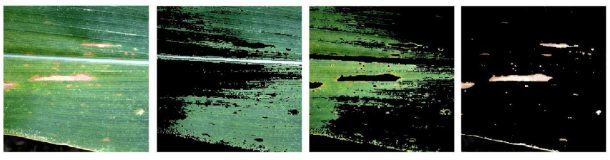
Note: SCR represents the Southern Corn Rust disease, CNLB represents the Corn Northern Leaf Blight, CSLB represents the Corn Southern Leaf Blight, PP stands for the Puccinia Polysora.

Disease spots segmentation is the crucial procedure of disease recognition based on leaf images. In this paper, k-means clustering method is adopted for segmentation. The basic idea of k-means clustering is to obtain *K* random position as the original clustering center. Then the distance between each object and each cluster center is calculated, the data object in the sample set is allocated to the nearest cluster according to the minimum distance. The mean value of the data in each cluster is recalculated and defined as a new center. This process will be repeated until the error sum of squares of each sample and the mean value of the cluster reaches the minimum, that is, the clustering center is no longer changed. After the end of the clustering, the leaves will be further divided into pathological part and normal part according to the color and shape, laying a foundation for features extraction. The segmentation results are shown in Figure 3. As a result of the complexity of the leaves, there is still some noise and outliers. Thus the morphological filtering method is used to optimize the segmentation results to remove the mis-segmentation caused by errors, such as burr, isolated points and some noise between rows [35]. In this paper, the structural elements such as matrix C are selected in the experimental process, some noise and burr are removed by corrosion operation, and the isolated points are removed to optimize the segmentation results. Figure 4 shows the segmentation results using K-means clustering and the morphological filtering method. The boundaries were clear and the objects were intact.

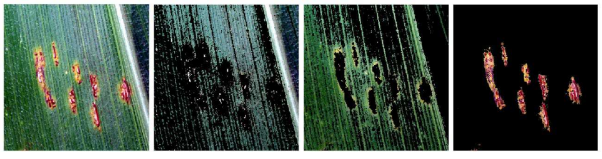


(a)

(b)



(c)



(d)

Figure 4. Contrast of segmented diseased district and morphological filtering

2.2 Feature extraction and fusion

Different diseases and symptoms of crops will produce specific symptoms of disease spots or pores in the aspects of size, shape, color and texture et al. Plant protection technicians make use of the differences of disease spots to diagnose crop diseases in the field. In the actual corn planting environment, because of the characteristics of complex leaf image structure and composition, uneven color, uncertain disease spot boundary and overlapping of different disease spots, the disease image has the characteristics of variability and complexity, which makes it more difficult to diagnose and identify. The characteristics of denaturation and complexity lead to the difficulty of diagnosis and recognition. The texture features of corn disease image are as follows: (1) the repetition characteristics of corn leaf spot texture in each dimension are different, do not follow isotropism, and there is no repetition in some dimensions; (2) the texture of corn leaf disease spot sometimes only reappears in a certain law, and the graphic complexity is very low. Thus, recognizing corn disease with leaves images need to take consideration of leaves' characteristics of color, shapes and textures comprehensively.

2.2.1 Color Features Extraction

Color features are extracted from image pixels and have the invariance advantages when they are rotated, scale altered and shifted. Because of the similar spot color of different corn diseased leaves, there is a large overlap of R, G and B color components between different disease RGB images. In color feature extraction, histogram method cannot express the information of color spatial distribution, and the excessive dimension leads to heavy computation, while the color moment method can reflect the color distribution characteristics of the image well, so the analysis of color moment features in statistical color features can improve the identification of features. *HSI* (hue, saturation, intensity) color space is similar to the principle of human eye sensory color, each component are independent and they can be processed separately, which can also eliminate the influence of different light on the image, and can greatly simplify the workload of image analysis and processing in image processing and computer vision as well, thus the *HSI* color space and *RGB* color space combined together to illustrate the color features. Because the *I* component in the *HSI* color model is independent of color (avoiding the influence of brightness on the image information), the color distribution information of the disease area is mainly concentrated in the lower order moment, so the color in the RGB, HSI color space of the diseased spot image is selected and calculated. The first rank moment (Average) and second rank moment (Variance) characteristics of components R, G, B and H, S are 10 color statistical features in total, which are used as the color feature parameters of the diseased spot image.

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} P_{ij} \tag{1}$$

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^{N} \left(P_{ij} - \mu_i \right)^2 \right)^{\frac{1}{2}} \tag{2}$$

Where μ_i represents the first rank moment of color, which defines the average intensity of each color component; σ_i represents the second rank moment of the color image, reflecting the color variance of the region to be tested, i.e., inhomogeneity; P_{ij} represents the probability of the occurrence of a pixel with a gray level of *j* or a composite pixel value in the color image *i*-th color channel component (*i* stands for R, G, B, H, S, respectively); *N* represents the total number of pixels in the image region.

2.2.2 Shape Features Extraction

The shape features of the disease spot districts are not affected by the external environmental factors (such as light, shadow, dust, et al.), but depend on the influence of the disease itself on the leaves. The morphological characteristics mainly describe the shape parameters of the object, and have a good correlation with the human visual perception system. It is an important feature of image recognition and analysis. After the disease images are segmented, the threshold value is calculated with Otsu method. The obtained disease spot regions are transformed into binary images, and the outline regions of the disease spot are obtained through the morphological operation. Because the shape features of different disease types differ greatly, and the moments has the invariance of moving, rotation and scaling when it describes the region targets. Therefore, the characteristics of Seven-Hu invariant moments and the degree of circularity, the degree of rectangle and the eccentricity, the shape complexity are selected. The four geometric characteristics are used as the characteristic parameters of the diseased leaves, and the calculation of the seven invariant moments is calculated as the method in the reference [36] , and the four geometric characteristics parameters are calculated as follows:

$$S_c = 4\pi \cdot \frac{A}{P^2} \tag{3}$$

$$S_a = \frac{A}{A_R} \tag{4}$$

$$S_e = \frac{L_{short}}{L_{long}} \tag{5}$$

$$S_s = \frac{P}{A^2} \tag{6}$$

In the formulas, where S_c represents the circular degree； A is the area of diseased spots； P stands for the parameter of diseased spots outline region, the closed curve is obtained by using the formula of general curve length $P^2 = \int_{\alpha}^{\beta} \sqrt{\varphi'^2(t) + \varphi'^2(t)dt}$, parameter t belongs to the range of [α, β]; S_a illustrates the rectangle degree; A_R means the minimum peripheral rectangular area, Function "*minboundrec*t" is utilized to calculate the A_R in matlab2016a; S_e represents the degree of eccentricity; L_{short} is the length of the short axis of the affected region's equivalent ellipse; L_{long} is the length of the long axis of lesion area's equivalent ellipse; S_s reflects the complexity of shapes.

2.2.3 Texture features extraction

The texture of leaves diseased is very different from that of healthy tissues in thickness and arrangement, which can reflect the regional characteristics of spatial distribution among pixels and the periodic changes of surface tissue structure. We utilized the gray level co-occurrence matrix to extract the texture feature, as for an image, the gray level of a pixel is represented by rows and columns of a matrix respectively [37]. Assuming that the gray level of a pixel is i and the gray level of distance δ is j, the occurrence of these two gray levels in the image is called gray level co-occurrence matrix. In the experiment, the Contrast, Correlation, Entropy, Energy and Homogeneity calculated and obtained and their equations are as follows:

$$C(i,j) = \sum_{i,j} |i-j|^2 p(i,j) \tag{7}$$

$$R(i,j) = \frac{\sum_{i,j} ijp(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{8}$$

$$\mu_x = \sum_i i \sum_j \hat{p}_\delta(i,j) \tag{9}$$

$$\mu_y = \sum_j i \sum_i \hat{p}_\delta(i,j) \tag{10}$$

$$\sigma_x = \sum_i (i-\mu_x)^2 \sum_j P_\delta(i,j) \tag{11}$$

$$\sigma_x = \sum_j (i-\mu_y)^2 \sum_i P_\delta(i,j) \tag{12}$$

$$E(i,j) = \sum_{i,j} P(i,j)^2 \tag{13}$$

$$ENT = -\sum_{i,j} P(i,j) \lg P(i,j) \tag{14}$$

$$H(i,j)=\sum_{i,j}\frac{P(i,j)}{1+|i-j|} \tag{15}$$

Where $C(i, j)$ represents the contrast, which reflects the the clarity of the image and the depth of the texture grooves; $R(i, j)$ represents the the correlation, which measures the similarity of gray level co-occurrence matrix elements in row or column directions; $E(i, j)$ stands for the energy, which illustrates the uniformity of image gray distribution and texture thickness; ENT is on behalf of Entropy, which demonstrates the texture complexity and in-homogeneity of target district; $P(i, j)$represents the probability of distribution of two pixels with gray level. $H(i, j)$stands for Homogeneity which represents how much local texture varies in an image. Some of texture feature examples are shown in Table 1.

Table 1. Some of the texture features of samples

Disease types	Northern leaf blight	Southern leaf blight	Puccinia polysora	Southern corn rust
Entropy	3.93	2.19	0.73	0.35
Contrast	2.19	1.61	0.45	0.24
Correlation	0.72	0.58	0.73	0.54
Energy	0.34	0.54	0.87	0.93
Homogeneity	0.85	0.86	0.96	0.98

2.2.4 Multi-feature fusion

Image feature fusion is the way that extracted features can fully express the relevant characteristic information about the disease target based on the feature of the target object in the image, that is, the disease region in the image. Compared with the original single feature, the accuracy of the associated feature extracted by the feature fusion method will be significantly improved, which can guarantee the discriminant information of the feature to the maximum extent and provide the basis for the classification decision of the classifier[38]. In this study, the extracted color features (a total of 10 dimensions), shape features (a total of 11 dimensions) and texture features (a total of 5 dimensions) were fused in parallel to form a 26-dimensional feature vector as the input of the model.

2.3 Deep Belief Network

Deep Belief Model was firstly proposed by Hinton et.al. in 2006 [39]. Deep belief network is a typical deep learning network, which combines low-level features into more abstract high-level representations through self-learning. It can deeply mine image information feature information, extract data high-level features, accurately fit complex functions, and have strong data classification and recognition capabilities [40]. Deep belief network is essentially made up of multiple Restricted Boltzmann Machines (RBM) stacked on top of which the data extracted from the deep features are connected to the soft-max regression model for classification, as shown in figure 5. Through

unsupervised learning of a large amount of unlabeled data, the deep level characteristics of the signal can be extracted automatically, and the dimensions reduction of the original data can be realized and the classification performance can be improved. The recognition and classification of crop diseases can be realized by combining with the soft-max regression model.

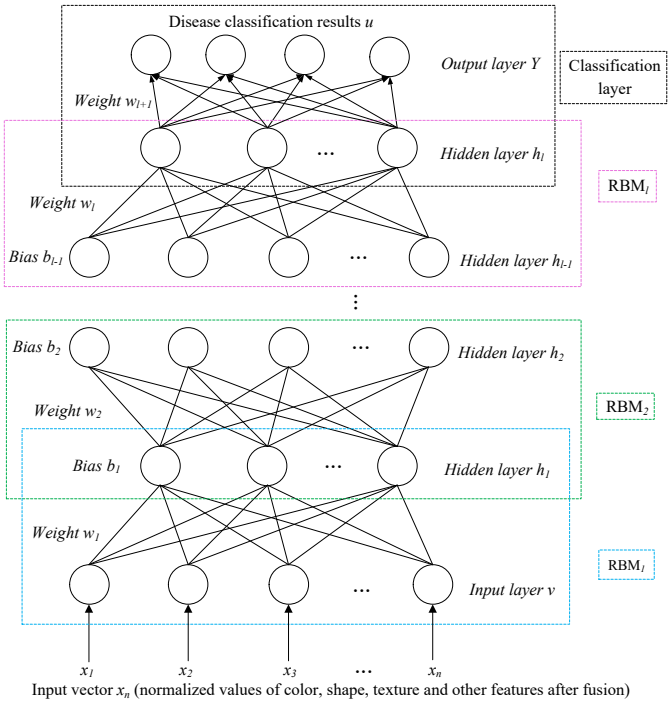


Figure 5. Structure of a standard DBN classification model

RBM consists of two layers, the hidden layer (h_l) and the visible layer (v) represents the observed data. The hidden layer h can be approximately a feature extraction layer, there is no connection in the layer and the interlayer is fully connected [41]. Parameter w is the weight value of the layer, a_i is the bias of visible layer neuron, b_j is the hidden layer neuron bias. When the number of hidden layer nodes reaches a certain value, this network model can be used to represent any discrete distribution. As a deep network structure, DBN training process mainly includes two steps: unsupervised pre-training and supervised back propagation network fine-tuning [39].

RBM is an energy-based model. The ideal state can be achieved when the model is determined to be minimized, and the purpose of network training is to minimize this energy function. In the basic RBM, visible layer and hidden layer units are binary variables, $v_i \in \{0,1\}$, $h_j \in \{0,1\}$, where 0 represents the unactivated state and 1 represents the activated state. All the units constitute a state of RBM, and these states follow a certain distribution. During the training process of RBM layers, the connection weights between the hidden layers are obtained in the form of optimal edge probability. By comparing the difference between the original sample number and the data transferred by Gibbs after RBM distribution, the reconstruction loss is obtained, and then the generated weights are evaluated, and then fine-tuned and updated by a layer of BP neural network to minimize the reconstruction error, and a new expression of the input sample is obtained. Standard RBM model is generated based on energy [42], which consists of binarization explicit layer elements

and hidden layer elements.

For a given status with *n* neurons in explicit layer and *m* neurons in hidden layer, $\theta = \{w_{i,j}, a_i, b_j\}$ is the set of parameters for the model. The energy function could be defined as follows:

$$E_\theta(v,h) = -\sum_{i=1}^{n} a_i v_i - \sum_{j=1}^{m} b_j h_j - \sum_{i=1}^{n}\sum_{j=1}^{m} h_j w_{i,j} v_i \tag{16}$$

Where v_i and h_i represent the explicit layer neuron state and hidden layer state respectively; $\boldsymbol{w}_{i,j}$ represents the the connection weight matrix element between explicit layer and hidden layer; a_i and b_i stand for the bias value of visible layer and hidden layer respectively. When the parameter is determined by above energy function, the joint probability distribution between visible and hidden layers is as follows:

$$P_\theta(v,h) = \frac{1}{Z(\theta)}\sum_h e^{-E_\theta(v,h)} \tag{17}$$

$$Z(\theta) = \sum_{v,h} e^{-E_\theta(v,h)}$$

Where $Z(\theta)$ is the partition function(dimensional factors) for the operation of normalization.

Hinton et.al. [39] Proposed a contrastive divergence (CD) method to update the weight of network easily and fast. The probability that neurons are activated in visible layer and hidden layer can be calculated as formula(18) and formula(19), where *sigmoid*(*x*) is nonlinear activation function.

$$P(h_i = 1 \mid v) = sigmoid(\sum_{i=1}^{n} W_{i,j} v_i + b_j) \tag{18}$$

$$P(v_i = 1 \mid h) = sigmoid(\sum_{i=1}^{m} W_{i,j} h_j + a_i) \tag{19}$$

$$\Delta W_{i,j} = \alpha \frac{\partial \ln P(v)}{\partial W_{i,j}} = \alpha(\langle v_i h_j \rangle_{data} - \langle v_i j_h \rangle_{model}) \tag{20}$$

weight updating of RBM model is done by taking partial derivative of logarithm of probability, where α is the learning rate and $\langle . \rangle$ is the expectation of distribution specified by subscript.

The weight obtained through RBM training can only guarantee the optimization of the feature vector in the single-layer RBM network, but it is difficult to achieve the optimization in the whole deep belief model. Therefore, the single-layer BP network is used to minimize the reconstruction error through the back propagation algorithm at the top of the multi-layer RBM, and the error information is propagated top-down to each layer RBM for fine-tuning. Global fine-tuning starts from the last layer of DBN, and a few label samples are used to fine-tune the model parameters layer by layer to the lower layer to further optimize the fitting effect of the model. At the end of the network, the cascade soft-max model is used as the classifier, and BP algorithm is selected to fine-

tune the parameters. Assuming that the DBN network has a total of *l* RBMs, the output at the highest hidden layer of the network is:

$$u^l(x) = Sigmoid(a^l + w^l u^l(x)) \tag{21}$$

The soft-max classification layer judges the output results of DBN network, and the category corresponding to the maximum probability is the recognition classification category of the model. From formula(21), it can be seen that connection weight *w* is an important parameter affecting network output. The parameter *w* is optimized by BP algorithm, so that the output of DBN network is closer to the real classification result.

The BP algorithm includes two processes, forward propagation and back propagation. In the forward propagation process, the input is mapped to the output layer by layer, and the probabilities of various categories $u^l = (u_1^l, u_2^l, \cdots, u_k^l)$, are obtained in the final classification layer. The back propagation process is based on a known category label $y^l = (y_1^l, y_2^l, \cdots, y_k^l)$. The mean square error EMS is obtained by comparing with the output data after classification. The error is propagated back from the last layer, and the weight parameters of DBN are adjusted layer by layer using gradient descent algorithm.

$$E_{MS} = \frac{1}{2}\sum_{j=1}^{k}(u_j^l - y_j^l)^2 \tag{22}$$

$$\Delta w_{i,j} = -\sigma \frac{\partial E_l}{\partial w_{i,j}} \tag{23}$$

$$w_{i,j} \leftarrow w_{i,j} + \Delta w_{i,j} \tag{24}$$

$$a_i^{l+1} \leftarrow a_i^l + \Delta a_i \tag{25}$$

$$b_j^{k+1} \leftarrow b_j^l + \Delta b_j \tag{26}$$

In the same way, other parameters in the parameter set can be adjusted to obtain the final DBN model, which has a good performance in feature mining and sample classification.

2.4 PSO algorithm and DBN optimization

In classic DBN, fixed learning rate is generally adopted for training. However, since each RBM needs multiple iterations and the direction of parameter update after each iteration is not the same, fixed learning rate may lead to "premature" phenomenon of the model or difficult convergence. According to the similarities and differences in the direction of parameter update after two successive iterations of RBM training process, an optimization method for parameter optimization based on PSO was designed, in which the algorithm controlled some parameters according to parameter update. Particle swarm optimization (PSO) algorithm originated from the behaviour of

birds hunting for food. The solution to the problem is regarded as the corresponding point in the search domain, namely the particle. The fitness value of each point determines whether the corresponding point is retained or discarded. All points have their own speed properties and position properties, which determines the corresponding position and speed of the next generation. The PSO optimization algorithm searches the current optimal solution of each particle iteratively, and evaluates the advantages and disadvantages of the solution by using the fitness value, so as to determine the global optimal solution. As a random search and parallel optimization algorithm, PSO algorithm has the advantages of simplicity, good robustness and fast convergence speed, and can find the global optimal solution with a large probability.

2.-3.4 Establishment of PSO-DBN model

Since the prediction performance of DBN is closely related to the setting of network parameters, in this paper, the realization of the DBN training and learning is based on PSO optimization algorithm. The root-mean-square error of the predicted value and the actual value of the training sample was used for the fitness function to search for the optimal parameter, so as to improve the accuracy of the prediction model. The main process of DBN model optimization by PSO algorithm is as follows:

1) Construct DBN network and generate particle swarm initialization parameters, including population size, number of iterations, learning factors and the limited interval of position and speed on the basis of DBN network parameters (number of iterations and learning rate).

2) Calculate the *fitness* value using formula (27) below.

$$fitness = \sqrt{\frac{1}{n}\sum\left[S(n) - \hat{S}(n)\right]^2} \tag{27}$$

Where n is the number of training samples; $S(n)$ and $S(\hat{n})$, output value and the expected output of the n^{th} samples respectively. The individual extremum and the population extremum were determined according to the initial fitness value of the particles, and the optimal position of each particle was regarded as its historical optimal position.

3) In the process of each iteration, the formula of particle velocity and position renewal is combined with individual extremum and global extremum to update the velocity and position of each particle.

$$V_{i,M}^{(K+1)} = \omega V_{i,M}^{K} + c_1 r_1^{K}(P_{i,M}^{K} - W_{i,M}^{K}) + c_2 r_2^{K}(P_{g,M}^{K} - W_{i,M}^{K}) \tag{28}$$

$$W_{i,M}^{K+1} = W_{i,M}^{K} + V_{i,M}^{K+1} \tag{29}$$

In the above equations, $P_{g,\,M}^{K}$, $P_{i,M}^{K}$ and $V_{i,M}^{K}$ represent the global optimal position, particle's own optimal position and velocity of particle i in the K-th iteration of M dimension. $W_{i,M}^{K}$ represents the

position of particle i in the K-th iteration of M dimension. Parameter ω represents the weight of inertia, which determines the degree of particle inheritance to the current velocity; r_1^K and r_2^K are random numbers distributed between [0,1]; c_1 and c_2 are non-negative constants, called acceleration coefficients, which adjust the moving step size of the current optimal position and the global optimal position.

The formula (30) is used to calculate the error E_{RMS} and judge whether the optimization goal is achieved. If the optimization goal is met, the optimization is stopped and the optimization results are output. If the requirements are not met, continue the optimization and repeat step 3) until the optimization goal is reached or the iteration ends.

$$E_{RMS} = \frac{\sum_{i=1}^{K} f(P_g^{(i)})}{K} \tag{30}$$

2.4 Experiment and parameters setting

In the disease image library, 70% of the images were chosen as the training set, the rest of the images were used for testing. The experiment process of images process and diseases recognition was carried out on Matlab2016a based on windows 7. The RAM is 8G with Intel Core(I7-7700) CPU.The ratio of the smallest sample(corn northern leaf blight) to the largest sample(southern corn rust disease) is 3:5, which belongs to the balance sample category [43]. DBN recognition model is constructed based on corn disease images features extraction. As a result of the types of diseases is only 4, so the structure of DBN is simple without dropout algorithm. Thus, the performance of the classifiers used in this article will not be reduced. The 26 dimensional feature parameters obtained after the extracted feature fusion were taken as the input of DBN, and the number of neurons in the input layer of DBN was selected as 26. The last layer is the output layer. According to the classification object of 4 diseases of corn, the number of neurons in the output layer is 4. The number of hidden layer nodes has a great influence on the training time and accuracy due to the need of DBN hidden layer node parameters. The unreasonable selection of the number of hidden layer neurons will lead to the decrease of DBN recognition ability and network fault tolerance. If the selection of network nodes is too small to meet the accuracy requirements; Over selection will cause the network to fall into the local optimum, which will also increase the complexity of the network. Therefore, based on experience, based on the DBN network structure formed by two-layer RBM stack, we investigated the method of selecting the number of nodes in the hidden layer of DBN. The number of hidden layer nodes in double hidden layer DBN network is 2, and there are three types of hidden layer nodes: rising type, descending type and constant type. The combination of the three types of hidden layer nodes was selected for the optimization test. In order to improve DBN training efficiency, samples in this paper were divided into 21 batches (each batch has 50 samples) for learning and training. The number of iterations was 500, and the learning rate was 0.001. The average value of 5 operations was taken as the experimental results.

The data samples used in this paper are all of 4 diseases, so the recognition accuracy is selected as the classification effect to judge the performance of the model. The recognition accuracy is expressed as: the recognition accuracy = the correct number of samples/total number of samples × 100%.The root mean square error (*RMSE*) is used to evaluate the error and consistency between the performance of the model and the standard value. Since the number of mis-classified samples is equal to the total number of samples minus the number of correctly predicted samples, the accuracy is consistent with that reflected by the *RMSE* evaluation index, and the influence of the total sample book is removed in the calculation of accuracy, so it is convenient to use different models or methods for comparison.

3. Results and analysis

3.1 Performance of DBN model with different structure

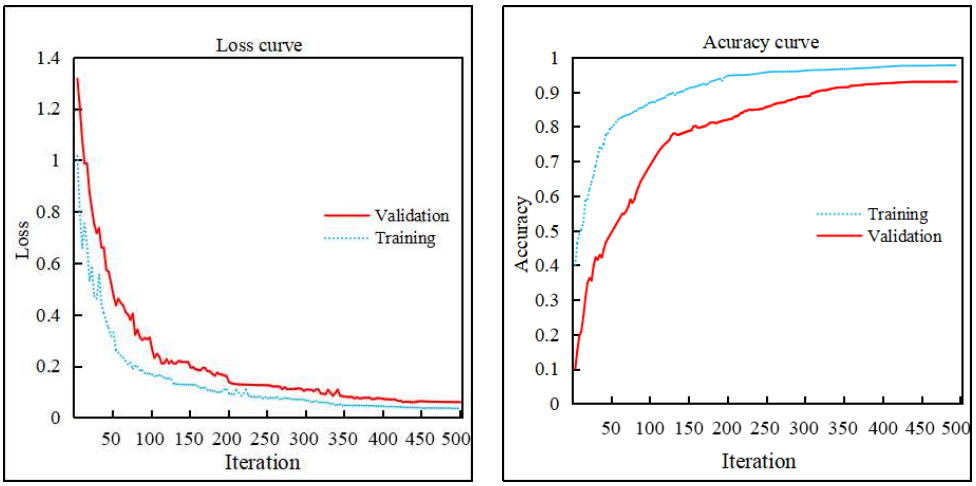
RMSE statistics of DBN model of three hidden layer node number combinations is shown in Table 2. According to Table 2, the model with a descending combination of nodes can better learn the characteristics of the original data and obtain a better classification effect. In order to determine the optimal number of nodes in the hidden layer, the number of nodes in the first hidden layer is selected in the interval [10,120] at a fixed interval of 10. The number of nodes in the second hidden layer is calculated as the integer value of 1/2, 1/3 and 1/6 of the number of nodes in the first hidden layer. Through the experimental verification of different combinations of the number of nodes, it is finally concluded that when the number of hidden layer nodes is 85-29, the classification performance of the model is optimal, and the minimum *RMSE* value is 0.0132. At this time, the network structure is [26,85,29,4].

Table 2. RMSE values of DBN under different network structures

Groups	Constant type		Rising type		Descending type	
	Nodes	**RMSE**	**Nodes**	**RMSE**	**Nodes**	**RMSE**
1	10-10	0.0604	10-15	0.0580	10-5	0.0189
2	30-30	0.0592	30-45	0.0582	30-15	0.0184
3	50-50	0.0557	50-75	0.0597	50-25	0.0171
4	70-70	0.0262	70-105	0.0542	70-35	0.0141
5	90-90	0.0598	90-135	0.0557	90-45	0.0152
6	110-110	0.0613	110-165	0.0591	110-55	0.0173

When the classification performance of DBN model had the optimal *RMSE* value, the DBN model has 26 nodes of input layer, 85 nodes of first hidden layer, 26 nodes of second hidden layer and 4 nodes of output layer. We trained and validated the optimal model and obtained the performance curve of training and validation process, which is shown in the Figure 6. From loss-curve in figure 6.a, we can find that the loss declines in training process, the model converges in

iteration of 250. In validation process, the model gets its convergence in iteration of 350. After convergence, the model has very little loss, which is lower than 0.04. In Figure 6.b, the accuracy of DBN model increases as the iteration, and the recognition accuracy of validation reaches 92.79%



a. Loss-iteration curve of DBN model b. Accuracy-iteration curve of DBN model

Figure 6. DBN model performance curve

3.2 Performance comparison of PSO-DBN and DBN

In order to effectively reduce the computational complexity and improve the recognition performance, the PSO optimization algorithm was used to search the optimal particle in the solution space of DBN by following the optimal particle in the solution space of DBN. And the initial size of optimization algorithm was set as 50, the evolution algebra for 30, accelerating factor c_1 and c_2 were set as 1.2545. The value range of the learning rate particle was [0.0001, 1]. The interval of the number of iterations of the particle was [10, 200]. Table 3 shows the specific recognition effects of DBN and PSO-DBN methods after the fusion feature training of 4 corn disease images. The average recognition rate of PSO-DBN model reached 97.65%, while the standard DBN network model recognition rate was only 92.79%. In terms of testing time, the PSO-DBN's testing time was 3.36s less than that of DBN. These indicate that DBN model optimized with PSO algorithm has better performance in training data.

Table 3. Testing performance comparison of standard DBN and PSO-DBN

Types of model	Single disease identification accuracy（%）				Average accuracy（%）	Total testing time（s）
	PP	CNLB	SCR	CSLB		
DBN	90.54	94.62	92.67	93.33	92.79	18.71
PSO-DBN	96.41	98.09	97.33	98.78	97.65	14.32

3.3 Impact on DBN and PSO-DBN with different training data

Deep belief network is a data-driven learning model, and the number of training data directly affects the recognition performance of the model. In order to verify the impact of training data in the process of multi-feature-based training of DBN, we detected the performance with the gradually increased training data from 50 to 1100. The accuracy rate was obtained every 50 samples. The change of recognition accuracy rate with data training amount was observed. PSO-DBN and DBN were compared and tested. The test results are shown in the Figure 7. When the amount of data training reached about 600, PSO-DBN has reached the recognition accuracy rate of about 90%, while the accuracy of DBN is only about 60%. When trained with full 1100 data set, the accuracy of PSO-DBN gets 97.65% and DBN gets 92.79%, which proves that DBN model optimized with PSO algorithm has higher data utilization rate and is better than DBN in accuracy.

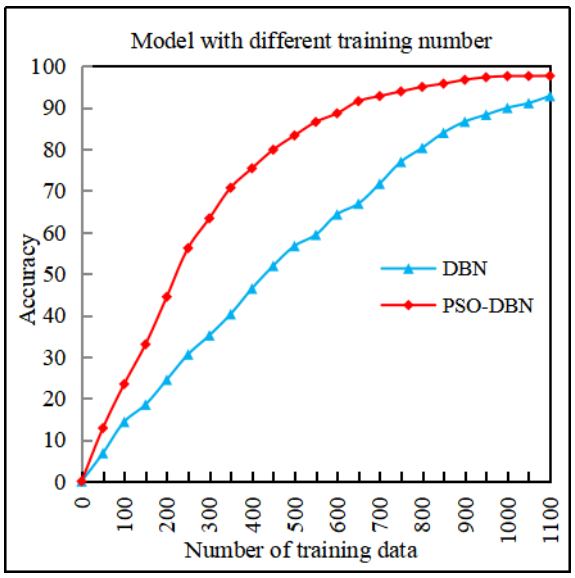
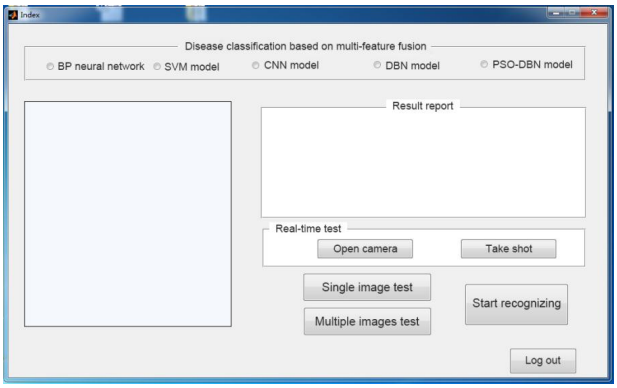


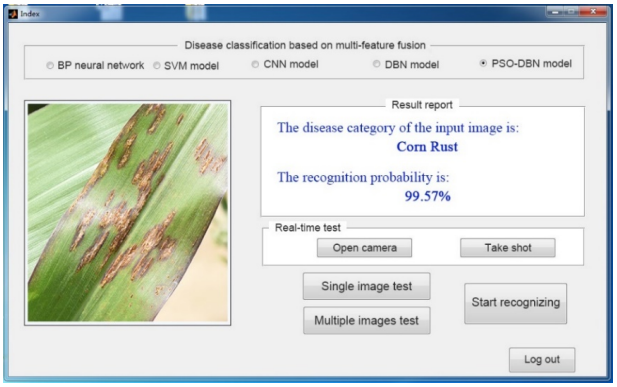
Figure 7. Accuracy comparison results

3.4 Comparison results with other models

We compared the comprehensive performance of improved DBN network model with the state-of-the-art models(BP, SVM and CNN) for accuracy and time consumption in classification test. In MATLAB 2016a, the SVM model was designed with Libsvm-Faruto Toolbox, and the BP Neural Network model was designed with the Neural Network Toolbox embedded in MATLAB. The CNN model was based on the Deep Learn Toolbox Toolbox. For various models, a MATLAB-based crop disease recognition GUI was designed, as is shown in Figure 8.



(a)



(b)

Figure 8. Crop disease recognition GUI with four models

The recognition results input with different kinds of feature are shown in Table 4 and Table 5. The experimental results show that the models using multi-feature fusion of disease images are better than using single feature type. The classification accuracy of the improved deep belief network (PSO-DBN) model proposed in this paper is obviously better than other mainstream classification algorithms. Compared with the classic DBN model, the accuracy and classification time of PSO-DBN model are also improved. The crop disease classification model based on DBN, combined with the image fusion features, can effectively extract the high-level features of the spot area information in the image, which improves the accuracy and generalization ability of the algorithm to some extent. However, the PSO-DBN model optimized by particle swarm optimization not only improves the convergence speed of the network, but also improves the classification accuracy of the model. The experimental results show that the disease recognition model based on multi-feature fusion and PSO algorithm optimized deep belief network can obtain excellent feature learning ability in the intelligent monitoring of pests and diseases in precision agriculture. This method can be used to guide the production practice of precision agriculture.

Table 4. The average recognition accuracy of disease tests based on different classifiers and different input features

Feature species	Feature dimension	Average accuracy of different models（%）				
		SVM	BP	CNN	DBN	PSO-DBN
Color	10	67.29	65.63	75.21	77.50	79.58
Shape	11	64.79	62.92	73.75	76.67	76.88
Texture	5	71.67	70.42	77.29	82.91	85.63
Fusion feature	26	81.25	78.96	85.00	89.79	97.65

Table 5. The average recognition time based on different classifiers and different input features

Types of model	Average recognition time（ms/pics）			
	T_{color}	T_{Shape}	$T_{Texture}$	T_{Fusion}
SVM	26.63	27.89	24.32	36.92
BP	28.13	29.41	25.35	39.41
CNN	26.09	27.61	23.29	35.19
DBN	27.06	28.82	24.09	38.98
PSO-DBN	23.92	24.31	20.85	29.83

4. Conclusions

Aiming at the characteristics of various diseases and insect pests and the complex background of corn disease leaves, a disease recognition and classification model based on multi-feature fusion and improved depth belief network was proposed. We obtained a variety of light, perspective and background conditions of corn plant diseases image in the field to ensure that there is a certain sample representative. The K-means clustering segmentation and the corrosion morphology processing were used to get less well-defined, noise of disease spot segmentation image. Then, color, shape and texture feature were extracted and fused respectively to normalize and set as the input vector of the training model.

1) Two disease recognition models were established based on the deep belief network, the most appropriate nodes combination of double hidden layers was explored to work out good disease classification results. The experiment illustrated that the best combination of network nodes was [26, 85,29,4], from which the DBN network model accuracy reached 92.79%. After using the PSO algorithm to optimize network parameters, the classification accuracy and computational efficiency of PSO-DBN were improved. The accuracy of PSO-DBN model reached 97.65%, which was 3.86% higher than that of DBN model. The testing calculation efficiency of PSO-DBN model on increased by 3.36 s.

2) Through comparative experiments, the classification effects of multiple classification models under the input of single feature and fused feature were studied. The research showed that, in the process of disease classification and recognition, the recognition effect obtained by the input quantity training after multi-feature fusion is better than that obtained by the single feature input. Although the recognition time after the fusion is higher than that of the single feature, the fusion feature is at least 6.88% higher than that of the single feature in recognition accuracy, which could satisfy the requirement of real-time detection of corn field disease.

3) The proposed method was compared with the state-of-the-art models, the results demonstated that the method based DBN model had higher data utilization efficiency and classification accuracy. So in the process of plant diseases automatic monitoring of precision agriculture, we should give preference to the PSO -DBN model based on multiple feature fusion.

4) In the further study, we will study the effect of the number of labeled and unlabeled samples on the accuracy of DBN model and explore the effect of multiple hidden layers (>2) on model recognition as well as the influence of the iterations of mean root square error (MRSE) to better apply PSO - DBN in precision agriculture production practice.

Author contributions: Methodology, X.F.; image acquisition, J.Y.; software, Y.X.; validation, J.Z. and X.F.; data curation, J.Y.; writing—original draft preparation, X.F. and J.Z.; writing—review and editing, Y.X.; project administration, Y.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Education of the PRC the New Engineering Education Reform Project (No. Official Letter No.17,2018), The National Students' Innovation and Entrepreneurship Training Program (No. 201810755079S), and Xinjiang Graduate Research Innovation Project (No. XJ2019G033).

Acknowledgements: I would like to express my gratitude to Professor **Jianping ZHOU**. He gives me valuable suggestions and warm encouragement. And I would like to give Special thanks to reviewers for their valuable comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ma, X. D., Guan, H. O., Qi, G. Y., Liu, G. & Tan, F. Diagnosis model of soybean leaf diseases based on improved cascade neural network. *Nongye Jixie Xuebao/Transactions Chinese Soc. Agric. Mach.* **48**, 163–168 (2017).

2. Jensen, H. G., Jacobsen, L. B., Pedersen, S. M. & Tavella, E. Socioeconomic impact of

widespread adoption of precision farming and controlled traffic systems in Denmark. *Precis. Agric.* **13**, 661–677 (2012).

3. Abdulridha, J., Ampatzidis, Y., Kakarla, S. C. & Roberts, P. Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques. *Precision Agriculture* (2019) doi:10.1007/s11119-019-09703-4.

4. Boulent, J., Foucher, S., Théau, J. & St-Charles, P. L. Convolutional Neural Networks for the Automatic Identification of Plant Diseases. *Front. Plant Sci.* **10**, (2019).

5. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D. & Stefanovic, D. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. *Comput. Intell. Neurosci.* **2016**, (2016).

6. Barbedo, J. G. A. A review on the main challenges in automatic plant disease identification based on visible range images. *Biosyst. Eng.* **144**, 52–60 (2016).

7. Geetharamani, G. & J., A. P. Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Comput. Electr. Eng.* **76**, 323–338 (2019).

8. Dhingra, G., Kumar, V. & Joshi, H. D. A novel computer vision based neutrosophic approach for leaf disease identification and classification. *Meas. J. Int. Meas. Confed.* **135**, 782–794 (2019).

9. Wang, A. C., Zhang, W. & Wei, X. H. A review on weed detection using ground-based machine vision and image processing techniques. *Comput. Electron. Agric.* **158**, 226–240 (2019).

10. Koirala, A., Walsh, K. B., Wang, Z. & McCarthy, C. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO'. *Precision Agriculture* vol. 20 1107–1135 (2019).

11. Yamamoto, K., Togami, T. & Yamaguchi, N. Super-resolution of plant disease images for the acceleration of image-based phenotyping and vigor diagnosis in agriculture. *Sensors (Switzerland)* **17**, 1–13 (2017).

12. Yan, Q. *et al.* Apple leaf diseases recognition based on an improved convolutional neural network. *Sensors* **20**, 1–14 (2020).

13. Fuentes, A., Yoon, S., Kim, S. C. & Park, D. S. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors (Switzerland)* **17**, (2017).

14. Deng, X. W. *et al.* Recognition of weeds at seedling stage in paddy fields using multi-feature fusion and deep belief networks. *Trans. Chinese Soc. Agric. Eng.* **34**, 165–172 (2018).

15. Liu, N. & Kan, J. M. Plant leaf recognition based on multi-feature fusion and deep belief networks method. *J. Beijing For. Univ.* **38**, 114–123 (2016).

16. Zhao, P. & Wei, X. Z. Weed recognition in agricultural field using multiple feature fusions. *Trans. Chinese Soc. Agric. Mach.* **45**, 275–281 (2014).

17. Ahmed, F., Al-Mamun, H. A., Bari, A. S. M. H., Hossain, E. & Kwan, P. Classification of crops and weeds from digital images: A support vector machine approach. *Crop Prot.* **40**, 98–104 (2012).

18. He, D. J. *et al.* Weed recognition based on SVM-DS multi-feature fusion. *Trans. Chinese Soc. Agric. Mach.* **44**, 182–187 (2013).

19. Dang, M. Y., Meng, Q. K., Gu, F., Gu, B. & Hu, Y. H. Rapid recognition of potato late blight based on machine vision. *Trans. Chinese Soc. Agric. Eng.* **36**, 193–200 (2020).

20. Xu, L. F. *et al.* Corn leaf disease identification based on multiple classifiers fusion. *Trans. Chinese Soc. Agric. Eng.* **31**, 194–201 (2015).

21. Wei, L. R., Yue, J., Li, Z. B., Kou, G. J. & Qu, H. P. Multi-classification Detection Method of Plant Leaf Disease Based on Kernel Function SVM. *Trans. Chinese Soc. Agric. Mach.* **48**, 166–171 (2017).

22. Ebrahimi, M. A., Khoshtaghaza, M. H., Minaei, S. & Jamshidi, B. Vision-based pest detection based on SVM classification method. *Comput. Electron. Agric.* **137**, 52–58 (2017).

23. Zhu, J. F. & Li, X. Descending and clustering methods for color image recognition of maize leaf diseases. *Jiangsu Agric. Sci.* **44**, 350–354 (2016).

24. Liu, G. L. & Yu, J. B. Machining Roughness Prediction Based on Knowledge-based Deep Belief Network. *J. Mech. Eng.* **55**, 94–106 (2019).

25. Liu, Y. *et al.* A Survey of Research and Application of Small Object Detection Based on Deep Learning. *Acta Electron. Sin.* **48**, 590–601 (2020).

26. Wang, Y. X., Zhang, Y., Yang, C. Y., Meng, Q. L. & Shang, J. Advances in new nondestructive detection and identification techniques of crop diseases based on deep learning. *J. Zhejiang Agric.* **31**, 669–676 (2019).

27. Nturambirwe, J. F. I. & Opara, U. L. Machine learning applications to non-destructive defect detection in horticultural products. *Biosyst. Eng.* **189**, 60–83 (2020).

28. Zhang, L. X., Chen, Y. Q., Li, Y. X., Ma, J. C. & Du, K. M. Detection and Counting System for Winter Wheat Ears Based on Convolutional Neural Network. *Trans. Chinese Soc. Agric. Mach.* **50**, 144–150 (2019).

29. Barbedo, J. G. A. & Castro, G. B. Influence of image quality on the identification of psyllids using convolutional neural networks. *Biosyst. Eng.* **182**, 151–158 (2019).

30. Arnal Barbedo, J. G. Plant disease identification from individual lesions and spots using deep learning. *Biosyst. Eng.* **180**, 96–107 (2019).

31. Too, E. C., Yujian, L., Njuki, S. & Yingchun, L. A comparative study of fine-tuning deep learning models for plant disease identification. *Comput. Electron. Agric.* **161**, 272–279 (2019).

32. Zhou, F. Y., Jin, L. P. & Dong, J. Review of Convolutional Neural Network. *Chinese J. Comput.* **40**, 1229–1251 (2017).

33. Zhang, X., Guo, S. S., Li, Y. B. & Jiang, L. Semi-supervised Fault Identification Based on Laplacian Eigenmap and Deep Belief Networks. *J. Mech. Eng.* **56**, 69–81 (2020).

34. Lecun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).

35. Bi, K., Jiang, P., Li, L., Shi, B. & Wang, C. Non-destructive measurement of wheat spike characteristics based on morphological image processing. *Nongye Gongcheng Xuebao/Transactions Chinese Soc. Agric. Eng.* **26**, 212–216 (2010).

36. Ming-Kuei, H. Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory* **8**, 179–187 (1962).

37. Sharif, M. *et al.* Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. *Comput. Electron. Agric.* **150**, 220–234 (2018).

38. Liu, J., Wang, X. & Wang, T. Image recognition of tree species based on multi feature fusion and CNN model. *J. Beijing For. Univ.* **41**, 76–86 (2019).

39. Holden, A. J. *et al.* Reducing the dimensionality of data with neural networks. **313**, 504–507 (2006).

40. Schmidhuber, J. Deep Learning in neural networks: An overview. *Neural Networks* **61**, 85–117 (2015).

41. Zhang, S. W. & Zhang, C. L. Disease and insect pest forecasting model of greenhouse winter jujube based on modified deep belief network. *Trans. Chinese Soc. Agric. Eng.* **333**, 202–208 (2017).

42. Hinton, G. E. A practical guide to training restricted boltzmann machines. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **24**, 599–619 (2012).

43. Weiss, G. M. & Provost, F. Learning when training data are costly: The effect of class distribution on tree induction. *J. Artif. Intell. Res.* **19**, 315–354 (2003).

Figures

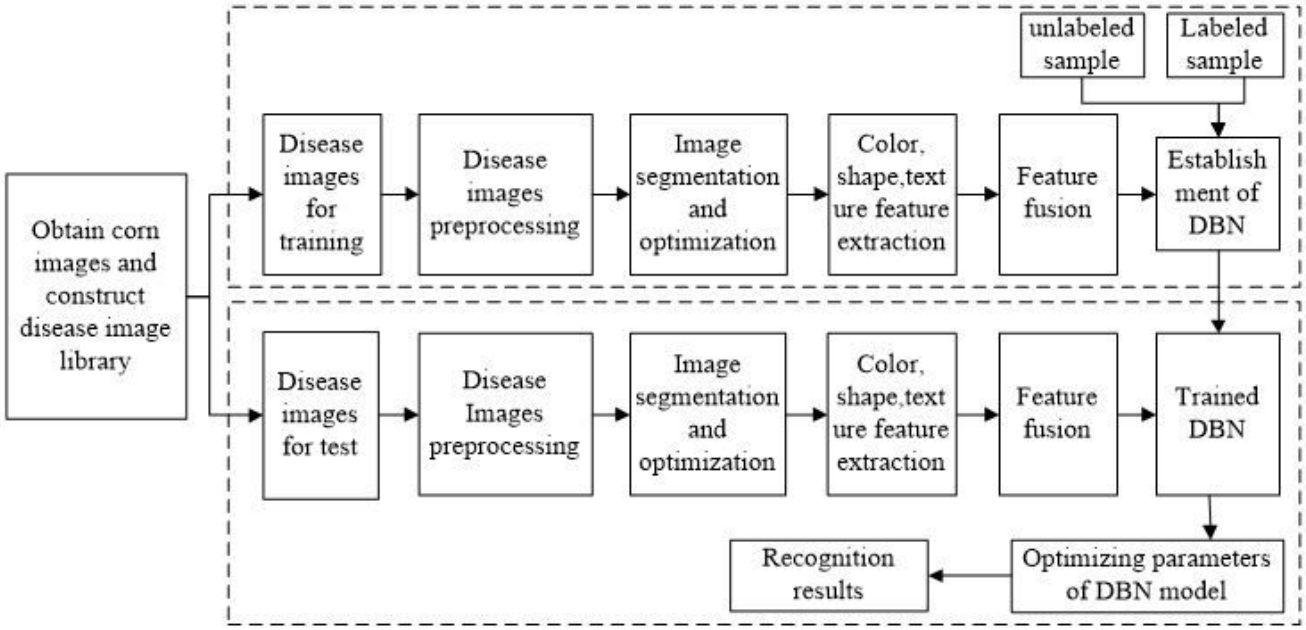


Figure 1

The process diagram of disease recognition

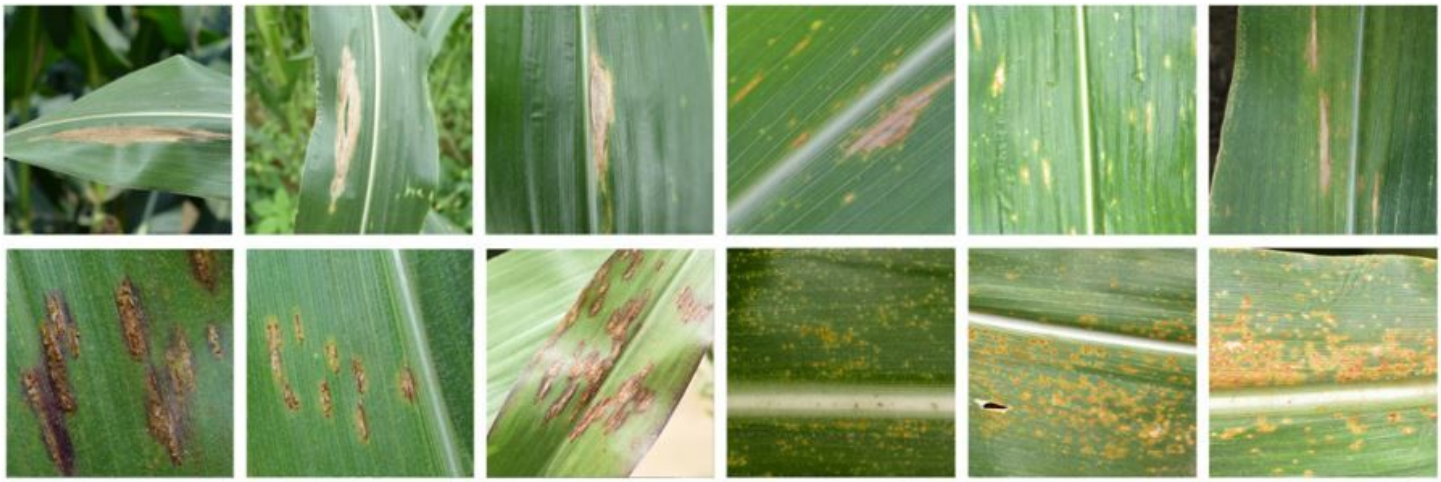


Figure 2

Some examples of corn disease images

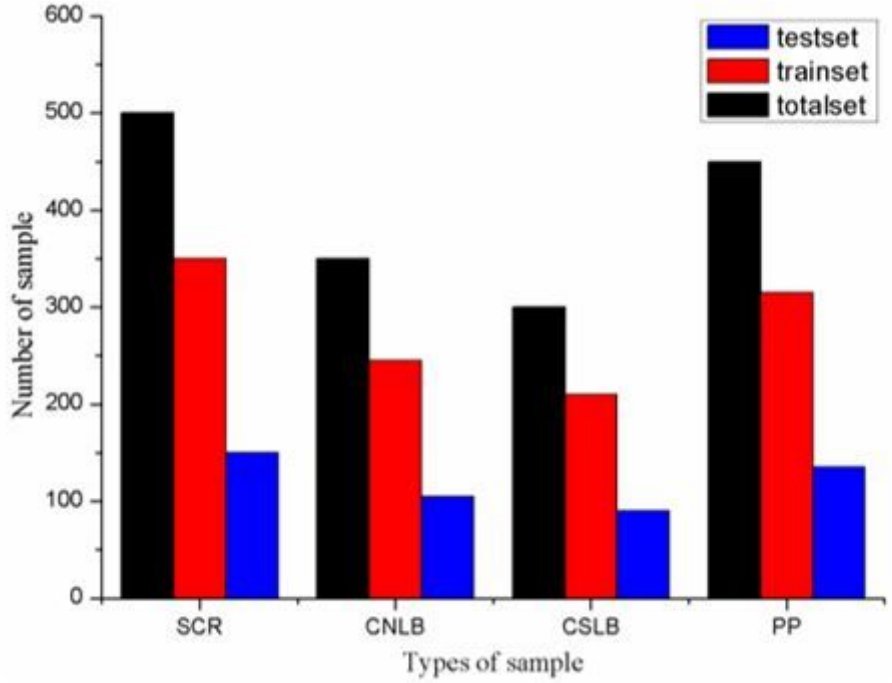


Figure 3

The Sample Statistics of Corn Disease Leaves

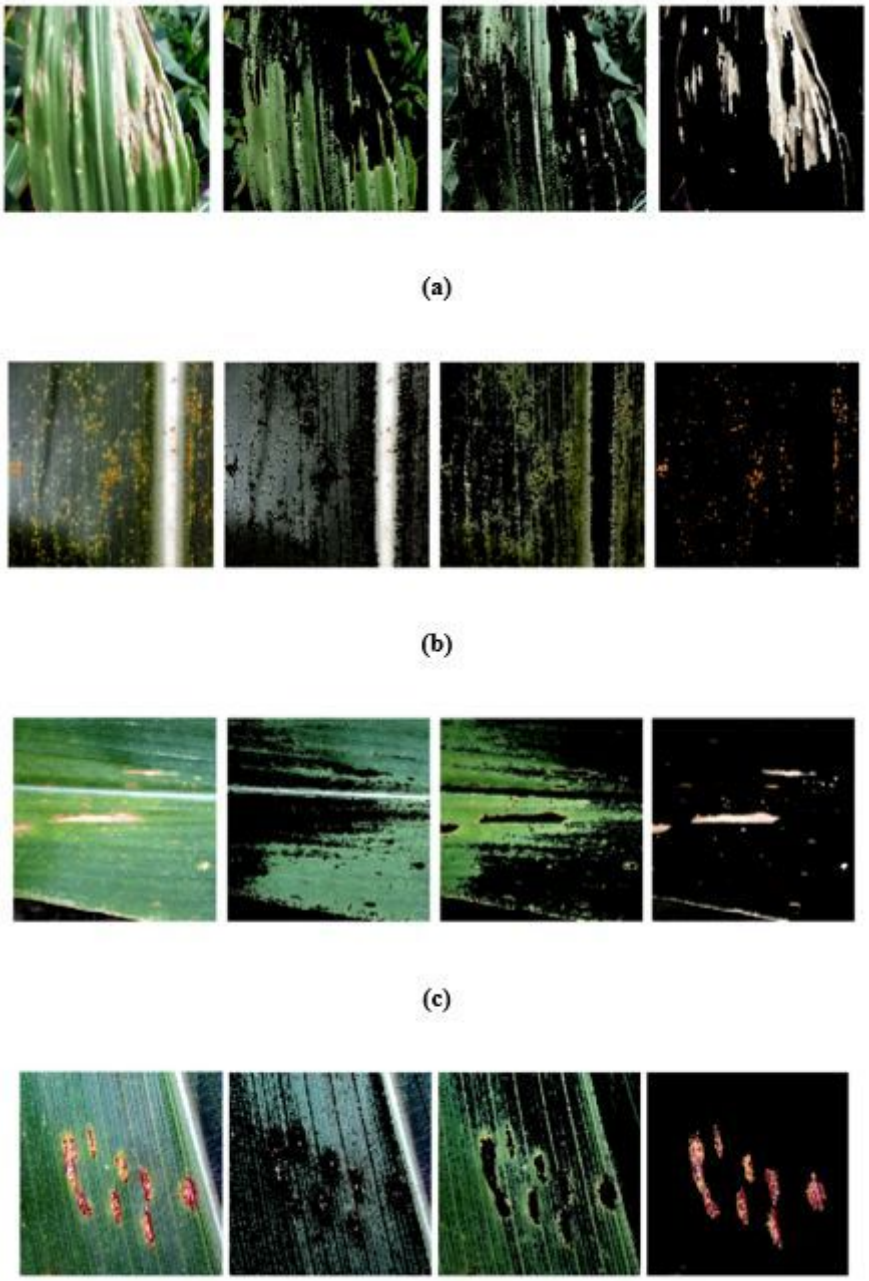


Figure 4

Contrast of segmented diseased district and morphological filtering

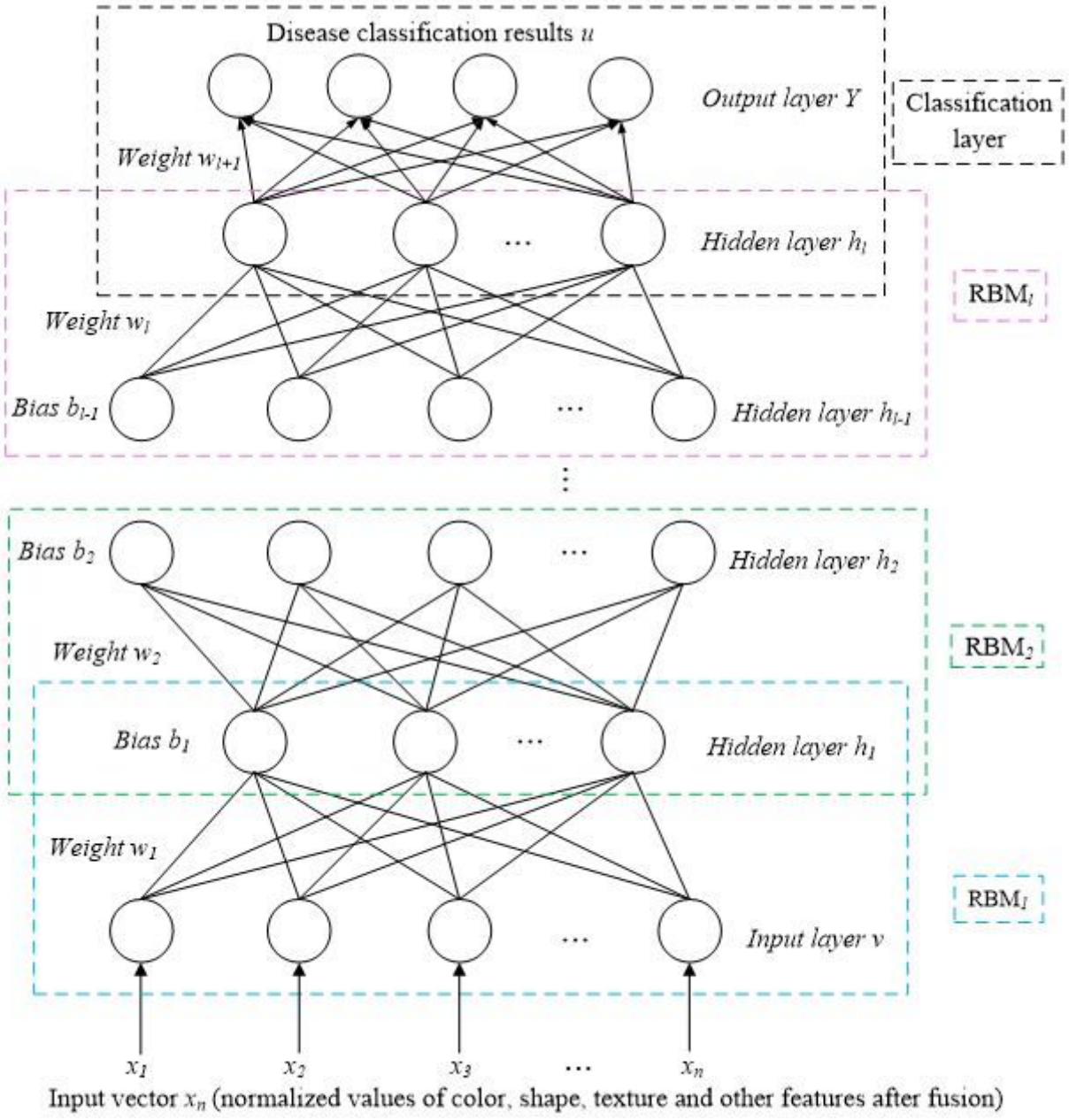
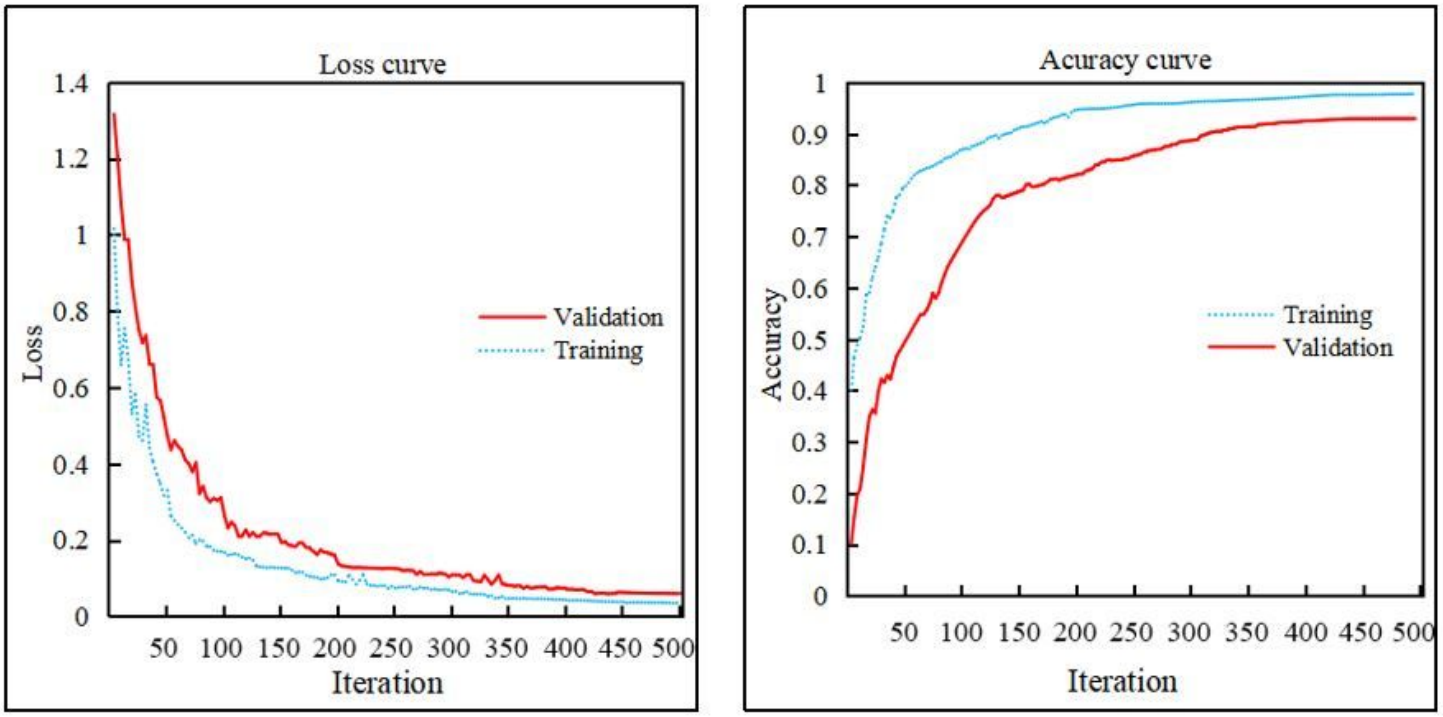


Figure 5

Structure of a standard DBN classification model



a. Loss-iteration curve of DBN model

b. Accuracy-iteration curve of DBN model

Figure 6

DBN model performance curve

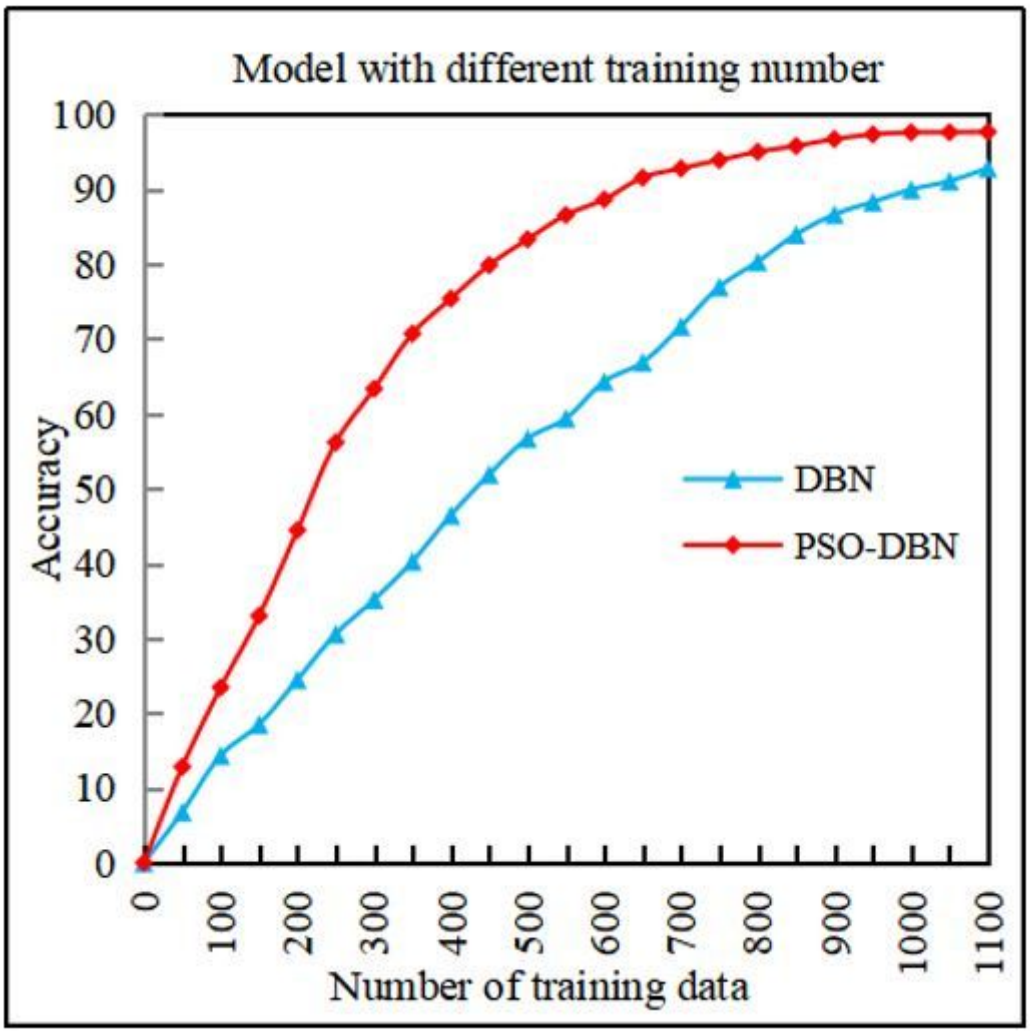
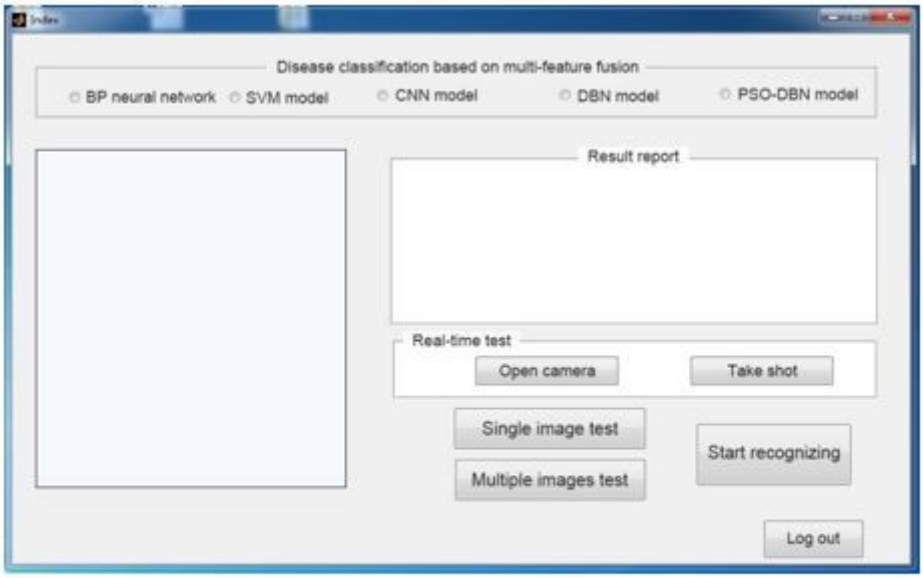
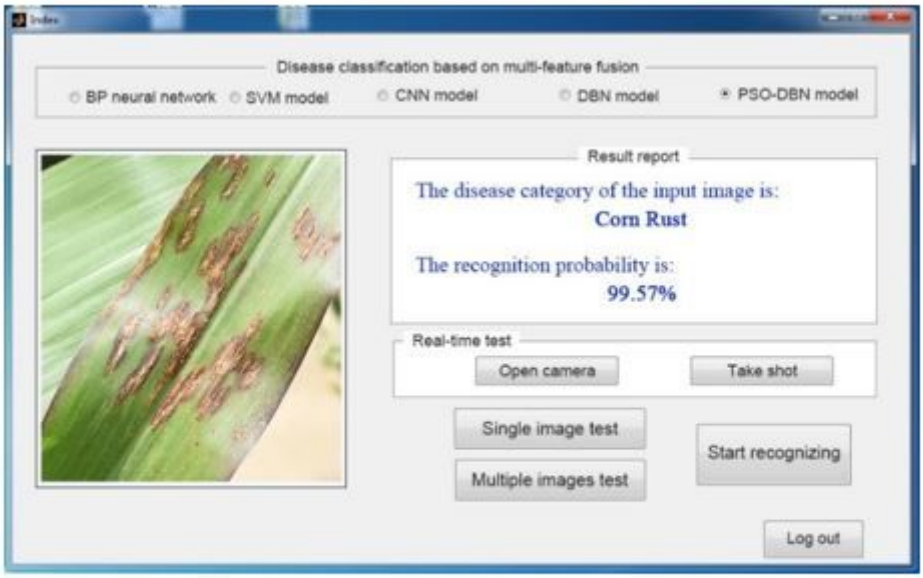


Figure 7

Accuracy comparison results



(a)



(b)

Figure 8

Crop disease recognition GUI with four models