Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 5, pp. 4070–4082, DOI 10.12694/scpe.v25i5.1861

RESEARCH ON THE CONSTRUCTION OF INTELLIGENT SYSTEM OF LANDSCAPE IN SCIENCE AND INNOVATION PARK OF SMART CITY – BASED ON THE CONCEPT OF SMART GARDEN DESIGN

KE XIE*

Abstract. With the advancement of urbanization in China, there is a growing demand for intelligent landscape management in urban science and innovation parks. To improve landscape management efficiency in science and technology innovation parks, this study adjusted the Inception module and activation function of the GoogLeNet algorithm, designed an improved GoogLeNet algorithm, and based on this, constructed a multi clue science and technology innovation park landscape vegetation intelligent recognition system. This is also the main contribution of this study. The experimental results show that the average image recognition accuracy of the "Improve GoogLe Net+MM" model designed from the study is 92.95%, which is higher than all the comparison models. Additionally, the computation time is significantly lower than the comparison models when the computation data volume is larger. The model designed in this study has a certain application value to improve the performance of the landscape intelligent management system within smart cities' science and technology parks.

Key words: Smart city; Landscape; Intelligent management; Plant recognition; GoogLe Net

1. Introduction . After entering the 21st century, artificial intelligence technology has developed by leaps and bounds, especially image recognition technology. This technology is extensively utilized in security, e-commerce, and various other industries [10]. Due to the demand for smart city construction, the landscape plant management mode at urban science and innovation park needs to become more intelligent in order to reduce the cost of manual management and improve management efficiency, and lay the foundation for subsequent intelligent plant watering and intelligent fertilization to remove pests. Therefore, the application of image recognition technology in artificial intelligence to this field is a potential research direction [13]. Intelligent recognition of landscape vegetation is a commonly utilized aspect within this area. However, issues such as inadequate recognition accuracy or speed frequently arise. Therefore, this study aims to enhance insufficient recognition accuracy and low recognition efficiency of image recognition models based on AI technology. The study will adjust the activation function and Inception module of the GoogLeNet neural network algorithm in artificial intelligence technology. Subsequently, a multi clue integration model for intelligent recognition of landscape plants in science and technology innovation parks will be designed with the improved GoogLeNet algorithm as the core. This enables GoogLeNet to obtain the possibility of improving recognition image accuracy while minimizing network parameters. This is the main academic contribution reflected in this study. In order to verify the application effect of the designed model, several comparative plant recognition experiments were arranged and carried out, in which not only the effects of using different dataset processing methods and model training methods on the final image recognition performance of the model are compared, but also the recognition accuracy of different plant recognition algorithm models built with the same dataset processing methods and model construction methods are horizontally compared.

2. Related works. Patel H et al. proposed an improved convolutional neural network called "Depth-FuseNet" [12]. This algorithm is used to fuse thermal and visible images, and extract data features from the two images to generate high-latitude features. The experiment found that this method has a recognition accuracy of 6.74 percentage points higher than the VGG16 algorithm in constructing recognition systems on the test set [12].

The research team of Hwang used an improved convolutional neural network to build a model for detecting neovascularization and age-related macula of patients. The experimental outcomes demonstrated that the

*School of Culture and Tourism, Chengdu Polytechnic, Chengdu, 610041, China (xie_ke0041@outlook.com)

approach effectively enhances clinical diagnosis efficiency [4].

The paper [7] concluded that deep learning-based detection of optic disc abnormalities in color fundus photographs is mainly limited to the field of glaucoma. However, numerous systemic and neurological diseases that pose life-threatening risks can present as deviations in the optic disc. Therefore, the authors trained a color fundus photo optic disc abnormality detection model using migration learning for ResNet-152 neural network, and the test results showed that the algorithm model has a significant advantage over the comparison algorithm with 90% recognition sensitivity and 69% specificityy [7].

The paper [2] developed a gangue recognition method using a convolutional neural network based on an auditory model. According to the experimental results, the method achieved a high recognition accuracy of 99.5%. In addition, the method offers significant noise immunity in comparison to frequently utilized recognition methods across different noise conditions [2]. Maior C S believes that with the rapid spread of SARS coronavirus type 2 around the world, the scientific community has spent a lot of energy to better understand the characteristics of the virus and the possible methods of prevention, diagnosis and treatment of COVID-19. Therefore, the author proposes an improved convolutional neural network to aid in diagnosing COVID-19. The test results show that the model achieves a balance accuracy of 87.7% when predicting one of the three categories ("no discovery", "COVID-19" and "pneumonia"), and a specific balance accuracy of 97.0% in predicting the "COVID-19" category [9].

The research team of [1] proposed an improved convolutional neural network to improve the recognition accuracy of protein images based on mass spectrometry. According to experimental results, the neural network significantly improves the recognition accuracy of protein images [1].

The paper [6] constructed an improved convolutional neural network model utilizing a two-way long-short memory network and attentional architecture to identify protein-specific spot recognition. The test results showed that the recognition performance of the algorithm is better and the computational speed is faster than the traditional convolutional neural network algorithm [6].

The author of [5] designed a convolutional neural network that combined the Gaussian statistical method. The test results show that the recognition effect of the algorithm for quantum microscopic images is significantly higher than the traditional deep learning algorithm. However, it has lower computational efficiency, being 24.2% slower than the recognition model built by the GoogLe Net algorithm and 21.9% slower than the recognition model built by the VGG16 algorithm [5].

In summary, although a lot of algorithm improvement studies have been conducted by previous people to improve the image recognition efficiency of the convolutional neural network, most of the studies failed to reduce the computational time consuming of the algorithm and improve the computational efficiency under the premise of improving the recognition accuracy of the algorithm. And identifying plants in science and technology parks is a significant and complex task that demands an intelligent system with higher efficiency. Therefore, this research seeks to improve the recognition speed of the recognition algorithm under the premise of improving its accuracy.

3. Design of intelligent plant identification system based on GoogLe Net algorithm for smart city science and innovation park.

3.1. Improved GoogLe Net algorithm design considering plant image features. Intelligent plant recognition belongs to an image recognition task, and the use of artificial intelligence (AI) algorithms is widespread in this area. The GoogLe Net neural network in artificial intelligence technology is an algorithm with good performance and fast computation speed that has been made public in recent years [15]. In addition, considering the aesthetics of the landscape and the intelligent and precise management requirements of the urban science and innovation park [3]. Figure 3.1 illustrates the typical network structure of GoogLe Net, which has demonstrated high recognition accuracy. The structure shown in Figure 3.1 has been proven to have high recognition accuracy, and the model is reasonable in terms of computation and training time, and computation consumption resources [11]. The GoogLe Net algorithm incorporates the Inception module, which expands the network width and reduces the computational effort by using asymmetric convolution. Specifically, it adds a 1×1 network in front of the 5×5 and 3×3 convolutional layers to reduce the dimensionality of the data.

The network input is a 224*224*3 RGB image, and preprocessing requires subtracting the mean data from the training set's RGB channels for each pixel. That is to subtract the mean data of the three primary color

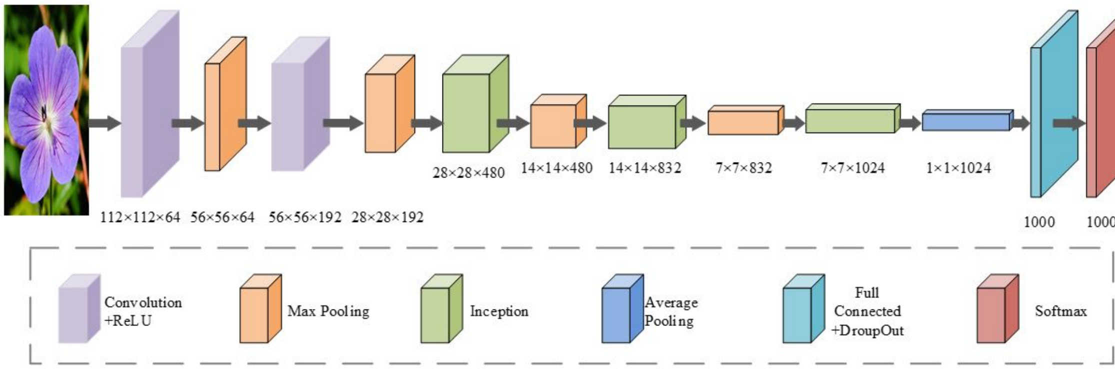

Fig. 3.1: Typical structure hierarchy of GoogLe Net

channels obtained from the training set for each pixel. And to reduce the error caused by the phenomenon of "gradient vanishing", here it sets the convolutional layer in the Inception module to use ReLU as the activation function. Specifically, 5*5. 3 * 3 convolutional layers are used to reduce the depth of the convolutional layer through a 1 * 1 filter. After maximizing the pooling layer, a 1 * 1 filter is also used to reduce the depth of the convolutional layer. After reducing the size or achieving maximum pooling, ReLU activation processing is required. The input layer of the GoogLeNet neural network is connected to a regular convolutional layer and a maximum pooling layer, followed by a reduced size convolutional layer and a regular pooling layer. Afterwards, a separate convolutional layer should not be designed. A 7-step average pooling layer of 7 * 1 should be set after the Inception module to reduce feature extraction errors caused by neighborhood size limitations during the convolution calculation process. After completing the data dimensionality reduction, the data will be input into the subsequent Dropout layer, with an output ratio of generally 60%. It will be followed by a fully connected layer containing 1024 neurons, which still uses the ReLU activation function. Finally, the data will be input into a softmax function classifier, and the predicted category data of the corresponding size will be output according to the usage requirements. After the design is completed, the Inception module structure is shown in Figure 3.2. As shown in Figure 3.2, compared to the simple Inception module, the Inception module in (b) subgraph has added a 1 * 1 convolutional filter, which can achieve the effect of changing the data dimension without changing the original feature structure, meeting the proportion invariance of features.

Considering the limited variety of cultivated landscape plants in the science park, especially the dataset vegetation types used to test the performance of the algorithm in this study are only 12, the number of neurons in the last fully connected layer of the GoogLe Net algorithm is also reduced to 12. Considering that the system needs to cope with the classification task, the cross entropy function withL_1 regular term is used here to construct the loss function and the adaptive gradient method is used to estimate the network parameters, and the loss function is calculated as shown in equation (3.1).

$$loss = -\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\left[p_m^n \log p_m^{\hat{n}} + (1-p_m^n)\log\left(1-p_m^{\hat{n}}\right)\right] + \frac{1}{2}\lambda\sum_{n=1}^{N}\sum_{l=1}^{L}\left[\|W_l\|_1 + \|b_l\|_1\right] \tag{3.1}$$

In equation (3.1), M represents the number of images and plant species in each training batch, p_m^n and$p_m^{\hat{n}}$ represent the true probability that the n image belongs to the$m$ category and the probability that it is predicted to be the category, W_l andb_l represent the weight coefficients and intercept coefficient matrix of the$l$ layer, respectively. λThe former in this study is initially set to 5×10^{-4} according to the industry experience. L represents the regularization coefficients and the number of layers containing parameters in the network, respectively. The architecture of the GoogLeNet algorithm has been established, and detailed improvement methods for the GoogLe Net algorithm used for plant recognition tasks will continue to be designed.

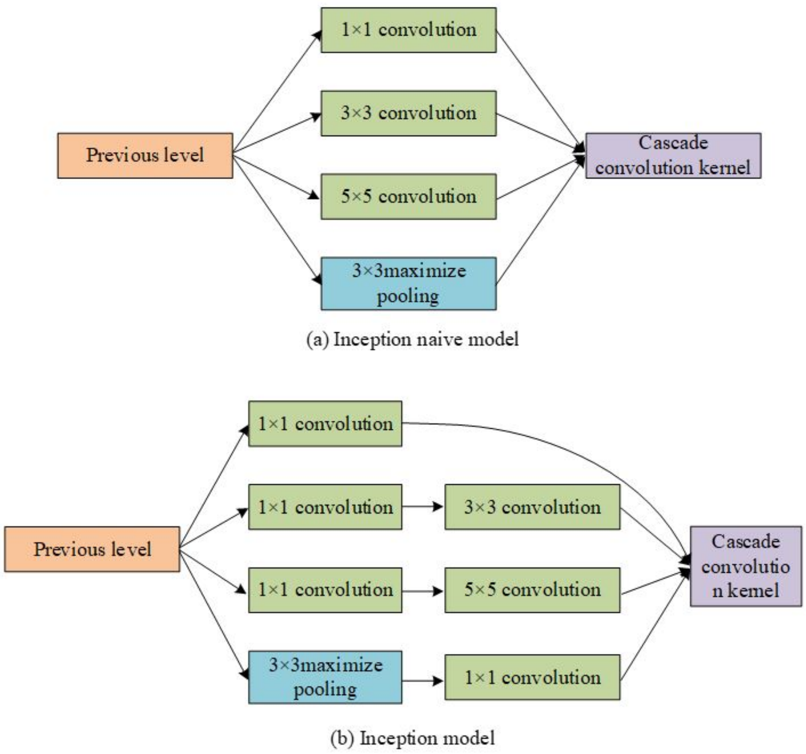

Fig. 3.2: Schematic diagram of the calculation structure of the designed Inception module

In summary, a series of structural improvements have been made to GoogLeNet to solve the problem of gradient vanishing. The GoogLeNet algorithm introduces the Inception module as its initial module to increase the network width and reduce computational complexity. This module adopts an asymmetric convolution method, which reduces the dimensionality of the data by adding a 1x1 network before the 5×5 and 3×3 convolutional layers. In the activation module, the convolutional layer in the Inception module uses ReLU as the activation function, while the 5x5 and 3x3 convolutional layers use a 1x1 filter to reduce the depth of the convolutional layer. After maximizing the pooling layer, a 1x1 filter is also used to reduce the depth of the convolutional layer. By introducing the Inception module and increasing the depth and width of the network, the network can better capture the complex features and structural information of landscape vegetation. This helps to identify different types of vegetation, their morphology, growth status, and other details, thereby improving the cognitive ability of vegetation. By using asymmetric convolution and adding a 1x1 network before the 5x5 and 3x3 convolutional layers to reduce the dimensionality of the data, the computational load can be effectively reduced. This enables the network to process large-scale image data more efficiently, improving the speed and accuracy of vegetation recognition. And ReLU, as an activation function, has nonlinear characteristics, which helps the network better learn and represent the complex features of vegetation, and can help the network better capture the nonlinear relationship of vegetation, improving the accuracy and robustness of vegetation recognition.

3.2. Design of intelligent plant recognition system based on hybrid improved GoogLe Net algorithm. . On the one hand, the deeper the hierarchy, the more training data are needed to train the neural network before it may be able to play a good application effect. However, the image data used for training the plant recognition system at the Smart City Science and Technology Park may be extremely limited, so it

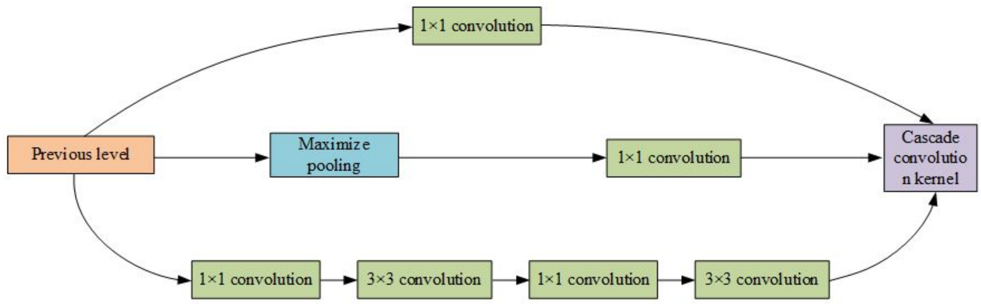

Fig. 3.3: Improved Inception module computational structure

is necessary to build a convolutional neural network with fewer network parameters to ensure its adaptability and good computational performance. Therefore, this time, various improvements are made to the classical GoogLe Net neural network algorithm to meet the demands of the intelligent plant recognition system in the science and innovation park.

In this study, the main purpose of achieving the overall number of parameters reduced while maintaining the computational accuracy of the GoogLe Net neural network is from the perspective of improving the Inception module. In the design idea of Visual Geometry Group Net (VGGNet) neural network, two 3×3 convolutional layers can achieve the same receptive field as a 5×5 kernel while reducing the number of parameters by at least 20% compared to the latter. In the process of improving the Region-Convolutional Neural Networks (R-CNN) target detection network to Fast R-CNN, the structure of reusing the convolutional layer information is used in it in order to reduce the training time of the algorithm. Specifically, the R-CNN algorithm utilizes the convolutional structure to create a feature map for each proposed region individually. While Fast R-CNN is also a convolutional feature map in the original image, resulting in all proposed boxes being formed in the same feature map output following convolutional computation. Based on the concepts of reusing the convolutional feature layer and convolutional kernel substitution, an improved Inception module is now designed by combining the two computational branches of 3*3 and 5*5 size convolutional kernels in the original Inception module; The input module first uses a 1*1 size convolutional kernel for dimensionality reduction processing, followed by an input 3 * 3 size convolutional block; The output information for this convolutional layer has two branches, one of which will directly use the output information as one of the outputs of the Inception module; The other branch will pass through convolutional layers of size 1*1 and 3*3 in order to reduce the output to one of the final module outputs. Figure 3.3 displays the enhanced computational hierarchy of the Inception module.

The activation function plays a crucial role in determining the final computational performance of a deep neural network algorithm. A nonlinear activation function can effectively ensure the network's fitting ability. In general, a good activation function needs to be monotonic, non-saturated, low computational complexity, with few parameters, non-linear, and differentiable everywhere. However, the ReLU or Swish activation functions used in traditional neural networks suffer from neuron "necrosis" and large computational effort, respectively. The latter is contrary to the original purpose of the research design to improve the Inception module. Therefore, in order to improve the computational accuracy and reduce the computational complexity of the GoogLe Net neural network, H_Swish function is chosen as the activation function in the algorithm, and its calculation method is shown in equation (3.2).

$$H\_Swish(x) = x \cdot ReLU(x + 3)/6 \tag{3.2}$$

In equation (3.2), x is the independent variable of the input H_Swish function. After improving and replacing the Inception module and the activation function of the GoogLe Net neural network respectively, the other structures in the neural network and the way the parameters are updated are left unchanged.

Given the limited amount of training and testing image data utilized in this study, it is necessary to use the migration learning technique to ensure high computational accuracy of the algorithm under such conditions.

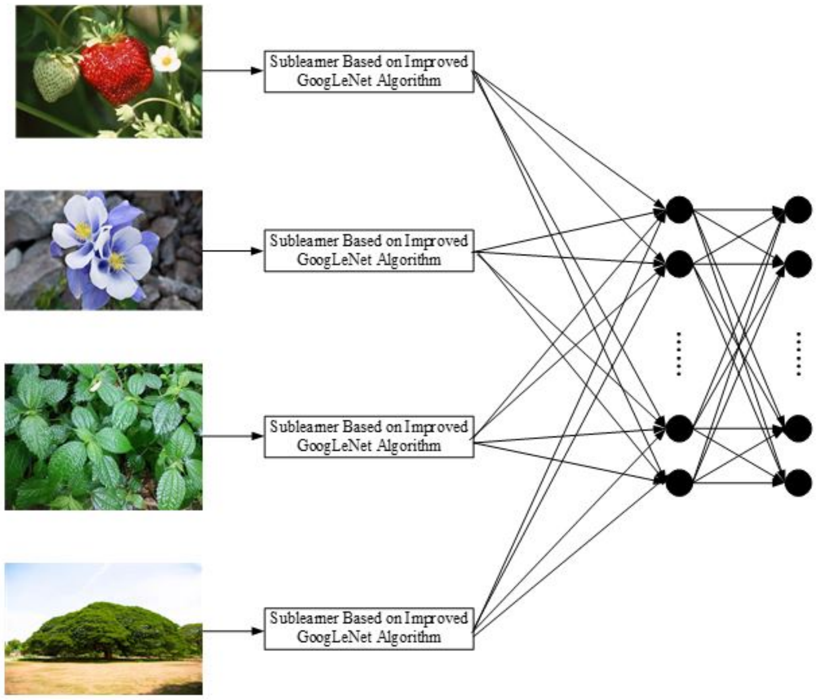

Fig. 3.4: Computational structure of plant multi-cue recognition model based on convolutional neural network algorithm

In other words, the study borrowed more complete and rich training data to train a neural network with more reasonable and mature parameters. Specifically, with reference to the general steps of migration learning, this study chose to use a richer image dataset to train the improved GoogLe Net neural network. Specifically, the last two fully connected layers in the network structure were modified to adjust the number of fully connected neurons to the number of desired classification categories, and a richer image dataset was used to train the neural network. After training, the neural network was trained again by applying the originally matched dataset to fine-tune some of the parameters to make the network more adaptable to the plant recognition needs of the science and technology park.

In plant recognition tasks, the information on recognition features that a single plant organ can provide is often more limited, especially different plant species in the same subject may have organs with similar appearance. For instance, while the flower shape and color of plum blossom and cherry blossom are relatively similar, but the leaf structure and the shape of the entire plant of plum blossom tree and cherry blossom tree differ considerably. It can be seen that the utilizing the multi-cue model for developing a plant recognition system can further improve the plant recognition ability of the system and alleviate the recognition error problem caused by the over-similarity of individual organs or local structures of plants to a certain extent. The computational layout of the plant multi-cue model based on the convolutional neural network algorithm is shown in Figure 3.4. As shown in Figure 3.4, the convolutional neural network-based plant multi-cue feature recognition model consists of two key components. Firstly, it requires training a sub-recognition model for each major recognition organ of the plant for a single organ, and the sub-recognition model in this study consists of a modified GoogLe Net neural network. Secondly, it involves efficiently merging the independent single-organ recognition models.

It is evident that each input for the plant multi-cue recognition model constructed in this study, is the final computational output of a related single-organ recognition model. This facilitates the encompassing model in acquiring an expanded insight into the plants, that are supposed to be recognized, in distinct dimensions. However, whether the final recognition results after integration can be more accurate than individual models depends mainly on the integration method of each single model. This study refers to the sublearners integration method of random forest algorithm, and uses the weighted summation of each single organ classifier to form the final classification results. The percentage of the sublearner's contribution in the final prediction is determined by the prediction category score of the single-organ prediction model. The detailed design of the integration approach is presented in this section. Since most plants can be identified by four aspects: leaves, fruits, flowers, and whole plants, only four sub-learners are also correspondingly set in the multi-cue plant identification model designed in this study, assuming that the model identification accuracy of each sub-learner isA_1 , A_2 , A_3 , andA_4 , and can be described by equation (3.3).

$$A = \{A_1, A_2, A_3, A_4\} \tag{3.3}$$

In equation (3.3), A represents the sublearners recognition accuracy matrix, then the integrated output weight of each sublearners can be described by equation (3.4),

$$q_i = \frac{A_i}{\sum_i^N A_i} \tag{3.4}$$

WhereA_i represents the classification accuracy of the$i$ th plant organ recognition model on the test set, and$N$ represents the number of single-organ plant recognition sublearners. It should be noted that the inputs of each sublearners are the corresponding plant part images of the plant parts, and each sublearner operates independently of the others. The sublearners' scores weighted according to the weights of equation (3.5) will be used as the initial values of the prediction scores for each category of the multi-cue modelS_i^* ,

$$S_i^* = \frac{q_i S_i}{\sum_{i=1}^4 q_i \cdot S_i} \tag{3.5}$$

WhereP_j is the score of the final predicted category by the$i$ th sub learner, which needs to satisfy the$0 < S_i < 1$ relationship, andS_i^* the initial score of the$i$ th sub learner input to the integrated model. If the same category scores appear in the sub learners, it means that multiple sub learners predict the same plant category and they need to be combined. That is, if$P_i = P_j$, and$P_i, P_j \in \{1, 2, ..., 12\}$, $i < j$, the relationship of equations (3.6) and (3.7) exists.

$$S_i^* = S_i^* + S_j^* \tag{3.6}$$

$$P_j = 0 \tag{3.7}$$

In Eq. (3.7), P_j represents the predicted plant category output by the$j$ sub learners. The integrated learner then counts the scores of each category label of the current image to be recognized and outputs the category with the highest score as the predicted category, as shown in Eq. (3.8).

$$\hat{y} = P_i \quad (i = \arg\max(S_i^*), i = 1, 2, 3, 4) \tag{3.8}$$

Finally, the experiment presents the calculation formula for the test indicators. This study shares two indicators: accuracy and calculation time. The latter is obtained through a computer-based timer and does not require further processing. Equation (3.9) demonstrates the calculation method for accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.9}$$

In equation (3.9), *TP* and *TN* respectively represent the number of cases accurately judged as positive and negative, while *FP* and *FN* respectively represent the number of cases wrongly judged as positive and wrongly judged as negative.

Table 4.1: Statistics on the number of various types of images in the PlantCLEF2016 dataset

Number	Category ID	Blade pictures	Fruits pictures	Flower pictures	Pictures of the whole plant
#01	309265	198	96	62	119
#02	007574	120	24	37	255
#03	106245	27	23	91	122
#04	012807	41	29	24	98
#05	016730	32	61	465	135
#06	041186	25	25	28	164
#07	309347	71	51	67	230
#08	012513	115	38	38	241
#09	251599	45	117	82	256
#10	228025	30	23	33	40
#11	309295	359	27	81	269
#12	133595	162	80	50	293

4. Plant intelligent identification system performance analysis.

4.1. Performance analysis experimental program development. The following computational experiments are designed to validate the performance of the improved GoogLe Net neural network algorithm-based plant intelligent recognition system for science and technology parks designed in this study. The datasets used in the experiments are ImageNet, which is a large image database containing at least 14,000K in size and more than 20,000 categories, and PlantCLEF2016, which assisted in the initial training of the model during the migration learning phase. The latter dataset was also used for fine-tuning and performance testing of the model post-initial training. Table 4.1 displays the key components of the PlantCLEF2016 dataset.

Before using the dataset, it also needs to be preprocessed. To avoid the model parameters biased to certain categories, the image quality is enhanced by using perspective transformation, affine transformation, etc. to expand a smaller number of image types and by using contrast enhancement, color dithering, etc. To compare the computational effects of the improved algorithms, the VGG16 algorithm, the classical Google Net algorithm, and the Faster-RCNN algorithm, which are commonly used in the industry and have good performance, are selected as the comparison methods to build the plant recognition system. The hyperparameters for each neural network algorithm were determined using the dichotomous method of multiple debugging with common tuning parameters in the industry. The parameters were not duplicated. To comprehensively compare the performance of the algorithms designed in this study, a comparative experimental scheme is designed as shown in Table 4.2. It is important to note that “MM” in Table 4.2 represents the abbreviation for the multi-cue model.

The reason for choosing this method of ”expand+drop” for preprocessing is to solve the imbalance problem in the dataset and enhance the quality of training data. By expanding the dataset, the model can be exposed to more diverse images, which helps improve its generalization ability for unseen data. Discarding images with insufficient pixel ratio helps prevent the model from bias towards certain categories and ensures that relevant features are focused during training. The use of ”expand+drop” preprocessing techniques can enhance the performance of the model by providing more diverse and balanced datasets. This may lead to better generalization ability and higher accuracy during training and testing stages. By focusing on relevant features and avoiding bias towards specific categories, the model becomes more robust and effective, and can identify different plant organs or species.

Transfer learning can utilize knowledge gained from a task to improve the learning of related tasks. In this case, the source task involves model training on a large dataset such as ImageNet, which contains various images from multiple categories. By pre training on ImageNet, the model learned rich and universal features that can be transferred to the target task of plant recognition. The use of transfer learning can significantly improve the performance of models, especially when applied to tasks with limited labeled data. By using the weight initialization model learned in ImageNet, the model learns more basic features from images, which can accelerate convergence speed in target task training. This typically leads to faster training and better

Table 4.2: Comparison experimental protocol display table

Experiment number	Algorithm solutions	Program explanation	Purpose of the experiment
#01	Improve GoogLe Net+MM+New Learning	Direct training and testing with PlantCLEF2016	Exploring whether the use of migration learning is beneficial for improving recognition system performance
	Improve GoogLe Net+MM+Migration Learning	Initial training with ImageNet, fine-tuning and testing with Plant-CLEF2016	
#02	Improvements to GoogLe Net+MM	Multi-cue model by improved GoogLe Net	Exploring the effect of different model organization methods on recognition performance
	Improving GoogLe Net+ Flower Single Organ	Improving the GoogLe Net algorithm by training with only flower images	
	Improving GoogLe Net+ fruit single organ	Training a single model using only fruit images	
	Improvement of Google Net+ leaf single organ	Training a single model using only leaf images	
	Improved GoogLe Net+ whole single organ	Training a single model with only the whole image	
	Improved GoogLe Net+ hybrid recognition	Training a single model with all images	
#03	Improvements to GoogLe Net+MM	/	Comparing the recognition performance of multi-cue models composed of different algorithms
	VGG16+MM	Multi-cue model construction using the VGG16 algorithm	
	GoogLe Net+MM	Similar to the previous algorithm scheme	
	Faster-RCNN+MM		
#04	Consistent with #03	/	Compare the calculation time of each model

Table 4.3: Experimental working environment and hyperparameter setting results

Type	Number	Name	Values and Setting Results
Hardware environment	#11	Host processor	Intel Core i7-6800K
	#12	Random Access Memory Specifications	6GB
	#13	Read Only Memory Specifications	1024GB
Software environment	#21	Operating system	Windows 10 Professional Edition
	#22	Programming language	Python
	#23	Database software	MySql
Parameter settings	#31	Learning rate	0.0001
	#32	Maximum number of iterations	800
	#33	Does the hidden layer have offset items	Yes
	#34	Parameter initialization method	Random Initialization
	#35	Number of training samples in a single batch	64

generalization performance, resulting in higher accuracy and better overall performance.

The environmental settings and hyperparameter settings for this experiment are shown in Table 4.3. The environmental settings are categorized as either hardware or software. The hyperparameters are obtained by conducting multiple experiments within the conventional range or conventional setting method to select the optimal value. The remaining parameter settings were determined during the model design phase.

4.2. Analysis of experimental results. The horizontal axis of Figure 4.1 displays various strategies for preprocessing data sets, including processing method 1, processing method 2, and processing method 3, which correspond to "no expansion + no discard", "no expansion + discard", and "expansion + discard" in that order. The processing method "expand + discard" is also depicted. "Expansion" represents the use of image processing techniques to expand the number of images in a relatively small number of categories in the dataset, and "discard" represents the deletion of images in the dataset where the percentage of pixels of the

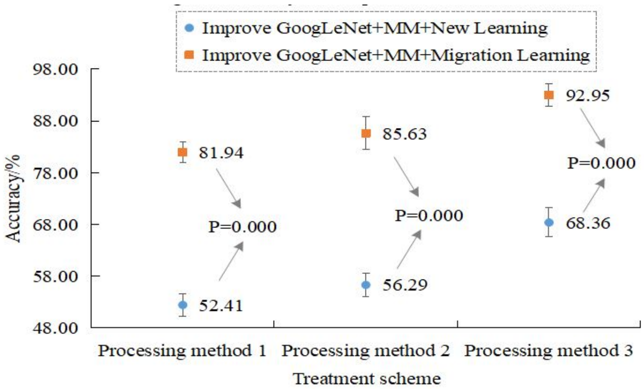

Fig. 4.1: Comparison of recognition accuracy of improved GoogLe Net under new learning and migration learning models

target plant organ is too small. The vertical axis of Figure 4.1 represents the image recognition accuracy of each model, and the different color labels represent different model training schemes. It is noted that to reduce the effect of experimental random error, each experimental scheme was repeated 20 times. The metrics of the type of measurement were presented in the form of mean ± standard deviation. The difference between groups was verified using the T difference significance test, with the level of significance set at 0.05. Observing Figure 4.1, it can be seen that the accuracy of the training scheme with the fused migratory learning is higher under the condition of using the same dataset pre-processing scheme. For example, the recognition accuracy of the migration learning scheme for treatment schemes 1 to 3 is 29.53, 29.34, and 24.59 percentage points higher than that of the brand-new learning scheme, respectively. From the perspective of processing schemes, the model using scheme 3 achieves the highest accuracy in comparison to the other conditions.

Considering the results of the study in Figure 4.2, the datasets of all model schemes of the subsequent experiments need to be pre-processed in the way of processing scheme 3, and all of them are trained with the migration learning method. The results of Experiment #02 were used to generate Figure 6, where the horizontal axis represents the model scenarios formed by different organization methods, which are explained in Table 4.2, and the vertical axis represents the image recognition accuracy of each model. As we can see in Figure 4.2, the recognition accuracy of the "Improve GoogLe Net+MM" organization scheme is the highest, with a mean value of 92.95%, while the accuracy of the models trained with whole plant images or mixed all images is lower, with a mean value of 66.38% and 71.42%, respectively. It indicates that the multi-cue model combining each major plant organ and the overall image, outperforms a single model in terms of recognition performance.

Considering the results of the study in Figure 4.2, it is necessary for all algorithms in the following experiments to create multi-cue models in order to take part in the study. The results of Experiment #03 were used to create Figure 4.2, where the horizontal axis represents the recognition models built based on each algorithm, the left vertical axis represents the recognition accuracy, and the right vertical axis represents the difference between the recognition accuracy of each algorithm model and the "Improve GoogLe Net+MM" model designed in this study, in %. As can be seen in Figure 4.3, the recognition accuracy of the "Improve GoogLe Net+MM" model designed in this study is still significantly higher than the other comparable models. The "Faster-RCNN+MM" model follows with the second-highest accuracy rate.

The average accuracy of the MM model is 90.51%, which is only 2.44 percentage points lower than that of the previous model. It shows that the recognition performance of the algorithm model designed in this study is already better than the Faster-RCNN model which has a better recognition effect in the market.

The results of Experiment #04 were used to create Figure 4.3. The graph displays the amount of time spent on computing (vertical axis) versus the number of images to be identified (horizontal axis). These images were sourced from the ImageNet dataset. As we can see in Figure 4.4, the "Improve GoogLe Net+MM" model

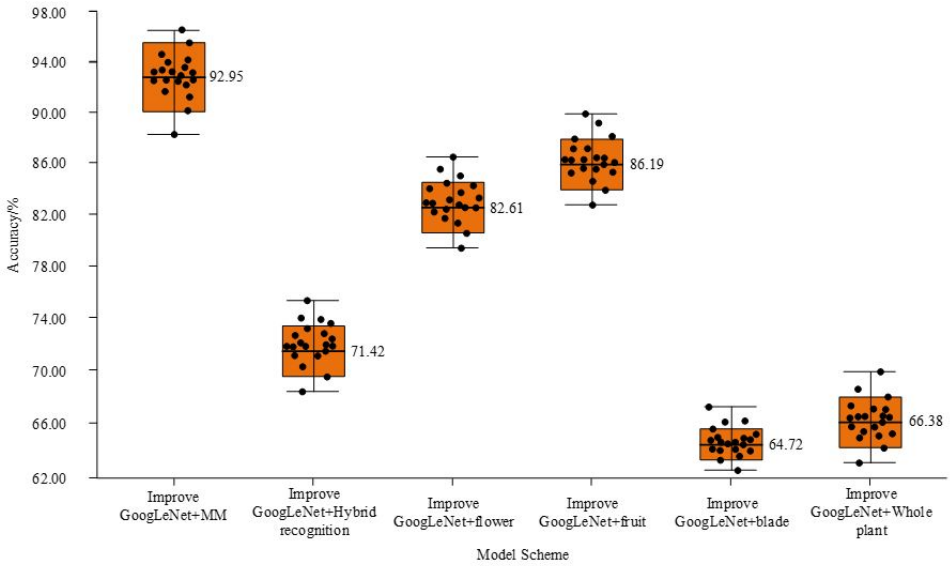

Fig. 4.2: Comparison of recognition accuracy of different model organization methods

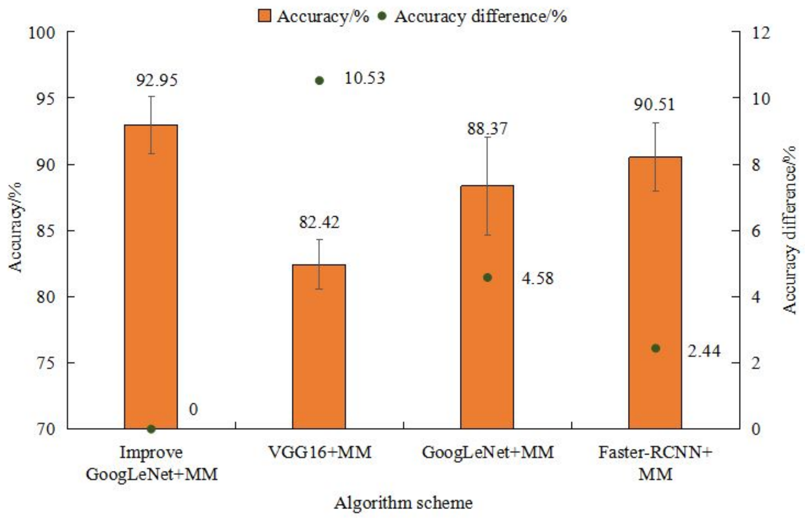

Fig. 4.3: Comparison of recognition accuracy of each algorithm model

is lower than the other three comparison models as the number of images to be recognized increases, while the computation time of the model built by the unimproved GoogLe Net algorithm is higher, but the computation time of the “VGG16+MM” model is lower than the other three comparison models when the sample size is larger. The “VGG16+MM” model requires the longest computation time of 37.86 seconds.

To further analyze the reliability and application value of this design model, 80 landscape plant experts from both domestic and foreign sources are now invited to conduct an evaluation experiment. Obtain a total of 762 real landscape plant images from a domestic science and technology innovation park, and use this design model and comparison model for landscape plant recognition. And have experts rate the recognition results of

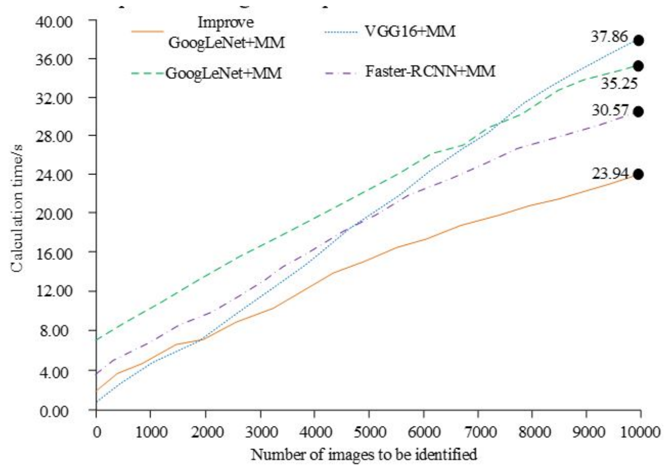

Fig. 4.4: Comparison of computational time consumption of each algorithm model

Table 4.4: Statistics of Evaluation Experiment Scoring Results

Identification	Average value	Standard deviation	Maximum value	Minimum value
Improve GoogLeNet+MM	9.18	0.43	9.77	8.81
GoogLeNet+MM	8.24	0.52	9.05	7.63
VGG16+MM	7.65	0.64	6.89	8.30
Faster-RCNN+MM	8.69	1.09	9.62	6.84

each model on a 10 point scale, and the scoring results are shown in Table 4.4. The scoring results, shown in Table 4.4, indicate that the overall average of the Improved GoogLeNet+MM recognition model designed in this study is the highest, with the most stable scoring results. The overall rating data for the Faster RCNN+MM recognition model is only lower than the former, but the stability of the rating results is the worst because the latter has the largest standard deviation of 1.09 among all models.

From the results of the above data experiments and evaluation experiments, it can be seen that the model proposed in this study exhibits favorable application outcomes. The communication between the experimental members and the expert group members in the evaluation experiment found that the expert team believes that the recognition model designed in this study better solves the problem of insufficient accuracy in identifying scarce categories in traditional landscape plant recognition models within science and technology innovation parks. This is mainly because the model designed in this study uses transfer learning to assist in training the model, enabling the model to obtain more effective local feature information that exists in different plant images.

Finally, the study will separately analyze the limitations of the design model. Firstly, the biggest limitation of this study is the inability to deploy the designed model into the plant management system for testing its effectiveness in a more realistic usage environment. Secondly, the evaluation experiment part of this study only invited experts to participate, and did not test the evaluation and application attitude of ordinary people to this model. These shortcomings will be addressed and improved in subsequent research.

5. Conclusion. This study focuses on the recognition accuracy and speed issues of traditional landscape vegetation intelligent recognition models. The GoogLeNet algorithm is improved and a landscape plant recognition system for science and technology innovation parks is designed. The experimental results show that the model's recognition accuracy is significantly improved by pre-processing the training data set with "expand+drop" and training the model with transfer learning. The recognition model constructed with multiple cues achieved an average classification accuracy of 92.95%, which surpasses that of all single models significantly. Compared with other models based on different algorithms but built in the same way, the recognition accuracy of the "Improve GoogLe Net+MM" model designed in this study is 92.95%, which is higher than all the comparison models. The "Faster-RCNN +MM" model follows with an average accuracy of 90.51%. Moreover, the computation time of the "Improve GoogLe Net + MM" model is lower than other models when processing larger data. The research data show that the improved GoogLe Net algorithm can improve the accuracy of plant recognition. However, due to limited research energy, this study was unable to combine the designed recognition system with the plant intelligent management system to analyze its application value. Subsequent research can develop real-time monitoring and feedback systems to continuously monitor the health status of plants and provide timely feedback. Real time data is fed back to the factory's intelligent management system for real-time decision-making and adjustment through sensor networks, cameras, or other IoT devices. Utilize machine learning and intelligent algorithms to deeply integrate plant recognition systems and factory intelligent management systems. By continuously learning and optimizing, the system can adapt to different environments and plant species, and provide more accurate predictions and recommendations.

REFERENCES

[1] A. R. Basharat, X. Ning, and X. Liu, "EnvCNN: A Convolutional Neural Network Model for Evaluating Isotopic Envelopes in Top-Down Mass-Spectral Deconvolution," *Analytical Chemistry*, vol. 92, no. 11, pp. 7778–7785, 2020.

[2] X. Chen, S. Wang, H. Liu, et al., "Coal Gangue Recognition using Multichannel Auditory Spectrogram of Hydraulic Support Sound in Convolutional Neural Network," *Measurement Science and Technology*, vol. 33, no. 1, pp. 015107–015116, 2022.

[3] J. Y. Choi and B. Lee, "Ensemble of Deep Convolutional Neural Networks with Gabor Face Representations for Face Recognition," *IEEE Transactions on Image Processing*, vol. 29, no. 3, pp. 3270–3281, 2020.

[4] D. J. Hwang, S. Choi, J. Ko, et al., "Distinguishing Retinal Angiomatous Proliferation from Polypoidal Choroidal Vasculopathy with a Deep Neural Network Based on Optical Coherence Tomography," *Scientific Reports*, vol. 11, no. 1, pp. 1–12, 2021.

[5] O. Ken, J. Liu, and H. Zoltán, " TNG: effect of baryonic processes on weak lensing with IllustrisTNG simulations," *Monthly Notices of the Royal Astronomical Society*, vol. 502, no. 4, pp. 5593–5602, 2021.

[6] Z. Li, J. Fang, S. Wang, et al., "Adapt-Kcr: a novel deep learning framework for accurate prediction of lysine crotonylation sites based on learning embedding features and attention architecture," *Briefings in Bioinformatics*, vol. 23, no. 2, pp. 1467–1469, 2022.

[7] T. Liu, J. Wei, H. Zhu, et al., "Detection of Optic Disc Abnormalities in Color Fundus Photographs Using Deep Learning," *Journal of Neuro-Ophthalmology*, vol. 41, no. 3, pp. 368–374, 2021.

[8] B. Lla, C. Sc, Xz. D, et al., "Deep Convolutional Neural Network for Accurate Segmentation and Quantification of White Matter Hyperintensities," *Neurocomputing*, vol. 384, no. 7, pp. 231–242, 2020.

[9] C. S. Maior, J. M. Santana, I. D. Lins, et al., "Convolutional Neural Network Model Based on Radiological Images to Support COVID-19 Diagnosis: Evaluating Database Biases," *PLoS ONE*, vol. 16, no. 3, pp. e0247839, 2021.

[10] N. B. Marya, P. D. Powers, L. Fujii-Lau, et al., "Application of Artificial Intelligence using a Novel EUS-based Convolutional Neural Network Model to Identify and Distinguish Benign and Malignant Hepatic Masses," *Gastrointestinal Endoscopy*, vol. 93, no. 5, pp. 1121–1130, 2020.

[11] K. B. Meena and V. Tyagi, "Distinguishing Computer-Generated Images from Photographic Images using Two-Stream Convolutional Neural Network," *Applied Soft Computing*, vol. 100, no. 4, pp. 107025–107033, 2021.

[12] H. Patel and K. P. Upla, "DepthFuseNet: An Approach for Fusion of Thermal and Visible Images using a Convolutional Neural Network," *Optical Engineering*, vol. 60, no. 1, pp. 013104–013132, 2021.

[13] X. Tan, K. Li, J. Zhang, et al., "Automatic Model for Cervical Cancer Screening Based on Convolutional Neural Network: A Retrospective, Multicohort, Multicenter Study," *Cancer Cell International*, vol. 21, no. 1, pp. 1–10, 2021.

[14] Y. Wan, X. Wang, Q. Chen, et al., "A Disease Category Feature Database Construction Method of Brain Image based on Deep Convolutional Neural Network," *PLoS ONE*, vol. 15, no. 6, pp. e0232791, 2020.

[15] B. Yu, L. Xu, J. Peng, et al., "Global Chlorophyll-a Concentration Estimation from Moderate Resolution Imaging Spectroradiometer using Convolutional Neural Networks," *Journal of Applied Remote Sensing*, vol. 14, no. 3, pp. 15–23, 2020.

Edited by: Zhengyi Chai
Special issue on: Data-Driven Optimization Algorithms for Sustainable and Smart City
Received: Nov 16, 2024
Accepted: Jun 26, 2024