Identification of the agricultural pests based on deep learning models

Yiyun Song#, XinYing Duan#, Yan Ren #
College of Resources
Sichuan Agricultural University
Chengdu, China

Lijun Luo, Dongyu Li #
College of Resources
Sichuan Agricultural University
Chengdu, China
#: These authors contributed equally to this work and should be considered co-first authors

Jingwen Xu*
College of Resources
Sichuan Agricultural University
Chengdu, China
*Corresponding author e-mail: x.j.w@163.com

***Abstract*—Pest identification is an important part of farming. It is crucial to accurately identify crop pests. In this background, based on the convolutional neural network, this paper collected 71 types of 35,000 images of pests, using the Inception-v3 and Inception-v4 model in GoogLeNet to build pests recognition model. Finally, based on several evaluation indicators, we combine the traditional methods for comparison and effect verification. The results show that the deep learning model is the best model for identifying pests and diseases in all the models involved in this study. Among them, the Inception-v4 support vector machine classification method has the highest accuracy, the overall classification accuracy rate is 97.3%, and the Inception-v3 classification accuracy is the second. The research can provide references for the identification of agricultural pests.**

Keywords-identification; pest; deep learning model; GoogLeNet

I. INTRODUCTION

Traditional artificial identification of pests is mainly based on the appearance characteristics and life habits of pests such as morphology, texture and judged by the prior knowledge of farmers. However, there are tens of thousands of kinds and large quantities of agricultural pests. It takes a lot of time to identify them manually. At the end of the 20th century, the concept of "precision agriculture" was put forward internationally[1].

China is a large agricultural country with abundant agricultural production resources, but the climate conditions are complex and diverse, and agricultural biological disasters occur frequently and seriously[2]. Every year, agricultural pests cause great losses[3]. From 2006 to 2015, the area of diseases and pests is ranged from 460.35 billion to 5075.3 billion hm^2, with an average actual loss of grain of 196.549 million tons[4], which the loss caused by pests was more serious. The main agricultural disasters in China are caused by crop pests, which have the characteristics of many kinds, great influence and the frequency of disasters. The disasters caused by crop pests interfere with the crop growth and lead to the decline in crop yield. Therefore, it is necessary to control agricultural pests[4]. Accurate pest identification is the precondition for pest control and pesticide application[5], but if it only relies on manual identification, it will be inefficient and high labor cost.

In recent years, with the rapid development of computer networks, image recognition technology based on computer vision has been widely used in the field of agricultural pest control, which can effectively reduce the cost of recognition and significantly improve the speed and efficiency of recognition[3]. The identification results of the convolutional neural network are accurate and efficient. By using the convolutional neural network, the automatic identification results of crop pests can identify the types of pests and apply pesticides accurately, which has far-reaching significance for improving China's grain production capacity and food security.

At present, recognition based on machine vision has attracted great attention of scholars all over the world and has become one of the hot research directions. Gasoumi et al. used the neural network method to classify pests in the cotton ecosystem, and achieve 90% accuracy[6]. Liang Wanjie et al. proposed a method of rice stem borer pest identification based on convolutional neural network, elaborated data preprocessing, data set construction, identification model structure design and model development and implementation technology[7]. In 2016, the Google team proposed Inception-ResNet-v2 based on the Inception-v3 model in GoogLeNet combined with the inspiration from Microsoft ResNet, and achieved the best result in ILSVRC image classification benchmark test[8]. Using Inception structure can make full use of computing resources while increasing the depth and width of the convolutional neural network, which make the difficulty of model training will not increase too much. And it can replace the last fully connected layer of other convolutional neural network with the global average pooling layer. Therefore, the research of convolutional neural network based on the GoogLeNet model in crop pest identification has important theoretical and practical significance[7].

978-1-7281-5094-9/19/$31.00 ©2019 IEEE
DOI 10.1109/MLBDBI48998.2019.00044

II. INCEPTION MODEL

A. Convolutional neural network

The convolutional neural network is a kind of backpropagation (BP) neural network, which extracts the features of objects through the interaction of the feature extraction layer and feature mapping layer. Each layer of the convolutional neural network consists of a number of small size feature maps, and the volume operations in the upper layer to lower layer are performed using a specific size filter with the upper-layer data. The most important structure of the convolutional neural network is the convolutional layer, which can extract effective features and reduce the time complexity of the images. The pooling layer is usually connected by the convolutional layer to map the features, merge similar features and greatly reduce the number of features. The pooling layer can reduce network parameters and allow the trained features to have a certain generalization ability to avoid network overfitting. Normally, the neuron nodes of each layer are connected to the upper part of the area, which has two advantages. One is that the weight coefficients can be shared and the structural parameters can be greatly reduced. The other is to preserve the position relation of local features and ensure that the position relation of local features is preserved(see Figure 1).

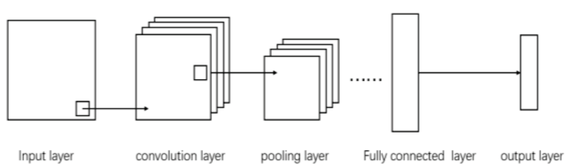

Figure 1. Convolutional neural network.

B. Inception-v3 and Inception-v4

This study uses the Inception-v3 model and the Inception-v4 model in GoogLeNet for transfer learning. Model migration refers to an image recognition system that uses tens of millions of images to train. When we encounter a new image field, the original image recognition system can be migrated to a new field, and the same effect can be achieved in the new field with fewer images. Model migration can also be combined with deep learning to distinguish the transferable degrees at different layers, and those levels with higher similarity are more likely to be migrated. For the problem of crop pests, it is more difficult to find enough training data. However, through transfer learning, the models trained from other data sources can be improved and applied to similar fields, thus the problem caused by the lack of data sources can be greatly alleviated. The basic idea of transfer learning is to use the pre-training GoogLeNet model to find the layers that can output reusable features in the pre-training model, and then the output is used as input feature to train the smaller neural network which needs fewer parameters.

The Inception-v3 model and Inception-v4 model used in this paper have been greatly improved, including convolution decomposition. Specifically, 7×7 convolution is decomposed into two one-dimensional serial convolutions, and 3×3 convolution is decomposed into two one-dimensional serial convolutions (1×3 and 3×1). In this way, the computation can be accelerated, the network depth can be further increased, the nonlinearity of the network is increased, and the Residual Blocks structure in ResNet is added to the Inception-v4 model which can improve the convergence speed of the model training.

III. MODEL VALIDATION

A. Image segmentation accuracy

The segmentation accuracy used in this paper is the ratio of the intersection and union of the binary foreground segmentation mask graph and the real foreground segmentation mask graph.

$$p=\text{mean}\left(\frac{A\cap B}{A\cup B}\right) \tag{1}$$

Where p is the accuracy rate, A means the binary foreground segmentation mask graph, and B means the real foreground segmentation mask graph.

B. Model training accuracy

Cross entropy (loss value): Cross entropy is the distance between the actual value (probability) and the predicted value (probability), that is, the smaller the value of the cross entropy, the closer the two probability distributions.

$$H_{(p,q)} = -\sum_x p(x)\log q(x) \tag{2}$$

Where p is the predicted value of the probability distribution, q is the actual value of the probability distribution, and H(p,q) is the cross entropy.

IV. THE RESULTS AND ANALYSIS

A. cross entropy

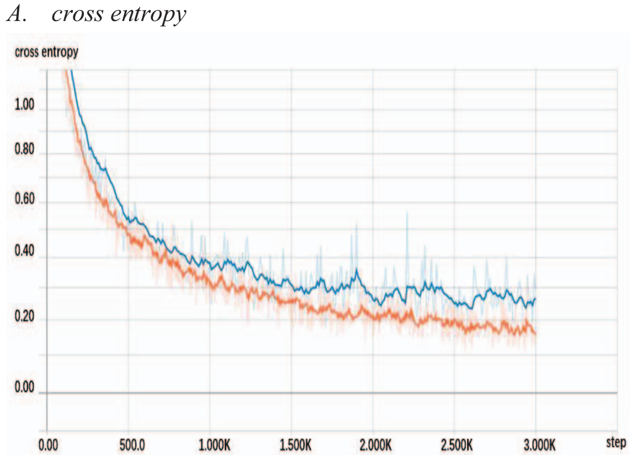

Figure 2. cross entropy

In Figure 2, the cross entropy of the model decreases as the number of steps increases. During this period, the accuracy rate has a small fluctuation, but after 2500 steps, it finally stabilizes at around 0.18. It indicates that the accuracy is steadily increasing.

B. Different pests

We used Inception-v4 model to identify 71 kinds of pests, and the total number of samples reached 35000(some of them are shown in Figure 3).

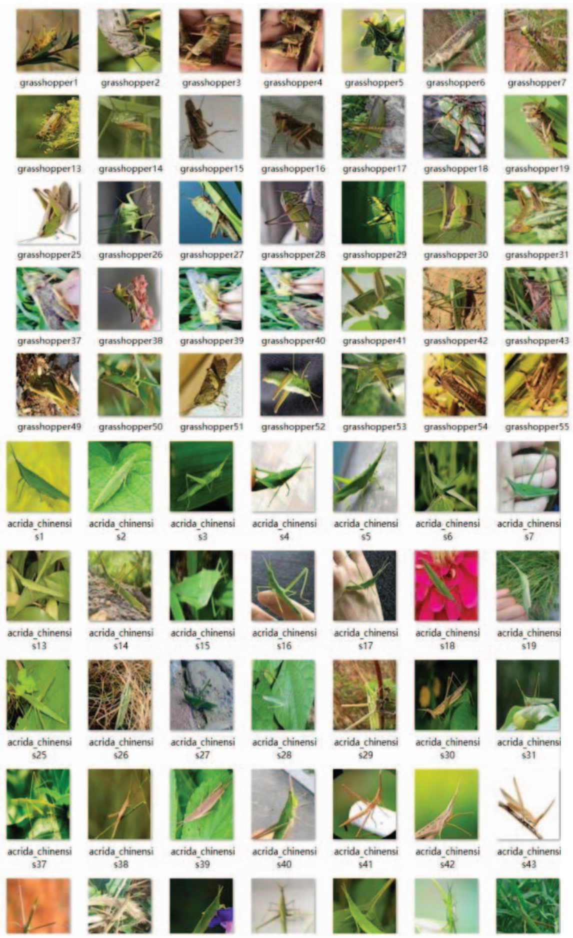

Figure 3. Images of some samples

TABLE I. IMAGE ACCURACY RATE FOR EACH TYPE OF PEST

Name	Accuracy	Name	Accuracy
silkworm moth	0.88	mole cricket	0.97
whitefly	0.99	Geometrid	0.92
Gypsophila	0.95	lappet moth	0.86
leaf beetle	0.91	Limax	1
Leafhopper	0.97	Grasshopper	0.88
scale insect	0.92	Lycaenid	0.92
carpenter moth	0.98	tussock moth	0.81
Cricket	0.95	Millipede	0.95
Weevil	0.97	Armyworm	0.85
Cutworm	0.94	notodontidae	0.77
plant hopper	0.85	Nymphalid	0.84
Earthworm	0.95	psylla	0.9
Longicorn	0.99	white butterfly	0.87
gall mite	0.93	Bark Beetle	0.92
Eucleid	0.91	bagworm moth	0.91
grasshopper	0.93	swallowtail butterflies	0.93
flatid planthopper	0.99	coconut leaf beetle	0.97
snout moth	0.75	snail	0.96
aphid	0.95	tiger moth	0.94
sawfly	0.87	hawkmoth	0.84
leaf mite	0.92	tortricidae	0.89
thrips	0.97	wireworm	0.98
lace bug	0.93	tabby moth	0.92
grub	1	termite	0.93
red imported fire ant	0.89	Red palm weevil	0.97
scarab	0.85		

It can be seen from Table I that there is a big difference in the accuracy between different pests. The grub and the limax have the highest accuracy which is 100% in the test, and the pest that has the lowest accuracy is the parasitic moth, and its accuracy is only 75%. The main reason for the result is the influence of patterns and other features. The parasitic moth lacks the distinguishing degree, and the color of the parasitic moth is brown and gray, which is easily confused with the background color. It is difficult to obtain biological edge information, and it causes lower recognition rate. At the same time, due to the seasonal influence, some pests have seasonal differences, which leads to a decrease in the reliability of the samples. In the actual test, there are differences among the pests that are affected by seasons.

C. Different models

TABLE II. THE IMAGE ACCURACY RATE FOR EACH TYPE OF MODEL

Model	Test samples	Recognized samples	Accuracy
LeNet[8]	1200	1080	0.9
CNN [9]	300	281	0.9375
SVM[10]	100	89	0.8926
BP neural network [11]	300	278	0.9267
SSGAN[12]	3500	3301	0.9432
Inception-v3	5000	4781	0.9562
Inception-v4	5000	4865	0.973

The results show that Inception-v4 has the highest accuracy among the seven models. The accuracy of SVM is the lowest and it's only 89.26%(see Table 2).

Most models have an accuracy between 90% and 95%. Among them, the accuracy of every identified model in the convolutional neural network model is relatively well, and the accuracy of Inception-v3 and Inception-v4 even exceeds 95%.

The reason is that the original traditional extracted features are too simple. Compared with the convolutional neural network, it is not suitable for more complex or high resolution image recognition.

According to the analysis, the Inception-v4 model can achieve more than 99% accuracy for some typical pest images.

V. CONCLUSION

Using the Inception-v4 model in the convolutional neural network, the identification of 71 types of 35,000 images was tested. The recognition rates of different insects are obtained and compared with SVM and Inception-v3 model. The Inception-v4 model discussed in this paper has the highest precision, reaching 97.3%. The superiority of Inception-v4 model and its application in the identification of pests are fully proved.

ACKNOWLEDGMENT

This study was financially supported by National innovation training program for College Students: Research on the accurate recognition algorithm of main agricultural pests based on convolution neural network(A093).

REFERENCES

[1] Hadley J F. Precision agriculture[J]. Encyclopedia of Food Grains, 2016, 80(1):162-167.

[2] Zhou Zeyu. Research on Government Control of Crop Biological Disasters [D]. China Agricultural University, 2004.

[3] Cheng X, Zhang Y, Chen Y, et al. Pest identification via deep residual learning in complex background[J]. Computers and Electronics in Agriculture, 2017, 141:351-356.

[4] Wu Kong-ming. The development direction of science and technology of crop pest control in China [J]. Journal of Agriculture, 2018, 8 (1): 35-38.

[5] Qin Linlin, Lu Linjian, Shi Chun, etc. Design of Intelligent Greenhouse Monitoring System Based on Internet of Things [J]. Journal of Agricultural Machinery, 2015, 46 (3): 261-267.

[6] Gassoumi H, Prasad N R, Ellington J J. Neural Network-Based Approach For Insect Classification In Cotton Ecosystems, 2000[C].

[7] Liang Wanjie, Cao Hongxin. Rice pest identification based on convolution neural network [J]. Jiangsu Agricultural Science, 2017, 45 (20):241-243.

[8] Zhang Shanwen, Xie Zeqi, Zhang Qingqing The application of convolutional neural network in the identification of cucumber leaf diseases[J].Jiangsu Journal of Agricultural Sciences, 2018, 34 (1):56-61

[9] Jia Shaopeng, Gao Hongju, Hang Ying Research progress of crop pest and disease image recognition technology based on deep learning[J]. Transactions of the Chinese Society of Agricultural Machinery, 2019, 50(S1), 313-317.

[10] An Qiangqiang, Zhang Feng, Li Zhaoxing, Zhang Yaqiong Image Recognition of Plant Diseases and Pests Based on Deep Learning[J]. Agricultural Engineering, 2018,8(07),38-40.

[11] Wen Zhiyuan, Cao Leping Fourier spectrum refraction image recognition of fruit and insect pests of mandarin orange[J]. Transactions of the Chinese Society of Agricultural Engineering, 2013, 29(23), 159-165.

[12] Jia Junjie, Li Jie Image recognition algorithm based on the semi-supervised generation of anti-network pests and diseases[J].Journal of the Wuhan University of Light Industry,2019,38(04),45-52.

[13] An Qiangqiang, Zhang Feng, Li Zhaoxing, Zhang Yaqiong. Image recognition of plant diseases and insect pests based on in-depth learning [J]. Agricultural Engineering, 2018, 8 (07): 38-40.

[14] Li Guanlin, Ma Zhanhong, Wang Haiguang. Image recognition of wheat stripe rust and leaf rust based on the support vector machine [J]. Journal of China Agricultural University, 2012, 17 (02): 72-79.

[15] Dai Jianguo, Zhang Guoshun, Guo Peng, Zeng Mingjun, Cui Mina, Xue Jinli. Classification of main crops in northern Xinjiang based on visible remote sensing images of UAVs [J]. Journal of Agricultural Engineering, 2018, 34 (18): 122-129.

[16] Bi Liheng. Plant species identification method based on leaf image algorithm [J]. Zhejiang Agricultural Journal, 2017,29(12): 2142-2148.

Major crop pests identification research based on Convolutional Neural Network

Lijun Luo#, Yan Ren#, Ling Gu, Tianjiao Wang, Fenglin Zhou, Wanyi Yang and Hongyan Zhang, Dongyu Li, Jingwen Xu*

College of Resources, Sichuan Agricultural University, Chengdu, Sichuan, 611130, P.R. China

*Corresponding author e-mail: x.j.w@163.com; #:these authors contributed equally to this work and should be considered co-first authors

Abstract. There are many kinds of crop pests in China and they are prone to disasters. Agricultural pests pose a serious threat to crop growth, so how to effectively identify crop pests is crucial. With the development of computer vision technology and artificial intelligence, the combination of computer vision technology and classification and identification of pests has become a hot and difficult point for experts at home and abroad. In this paper, based on the bag-of-words model and the GoogLeNet model, 2200 pest images collected were used as experimental samples to study the identification of crop pests. The experimental results show that the average classification accuracy of the traditional bag-of-words model is about 56.41%, and the GoogLeNet model recognition accuracy can reach 96.35%. The GoogLeNet model based on transfer learning has higher precision and stronger anti-interference ability than the traditional bag-of-words model.

1. Introduction

Crop pests have always been one of the most important agricultural disasters in China. The occurrence area of crop pests in China has been increasing year by year, and has climbed from more than 2 billion acres in 1980 to nearly 7 billion acres in 2016. In China's major agricultural production areas, the economic losses caused by pests alone will reach tens of millions of yuan per year[1]. At present, the prevention and control of agricultural pests are mainly divided into three methods of biological, chemical and physical machinery[2]. However, traditional large-scale spraying of pesticides will cause waste of drugs, and at the same time bring harm to the ecological environment. The residue of pesticides may even endanger human health. The best solution is to aim at the species and the site of the pests, and to reduce pesticide pollution while effectively controlling pests. Therefore, people should be eager to study the identification and classification of crop pests.

Pattern recognition technology based on computer vision can effectively reduce the cost of recognition, and can significantly improve the recognition speed and efficiency[3]. Compared with manual method, the automatic identification of crop pests by convolutional neural network can reduce labor costs, improve work efficiency, avoid environmental pollution caused by excessive use of pesticides, and ensure food safety. In the transformation and upgrading of precision agriculture, it is of great theoretical and practical significance to promote Chinese agricultural production from labor-intensive traditional agriculture to intelligent precision agriculture.

In recent years, pattern recognition technology has developed rapidly. Scholars at home and abroad have published a large number of related researches, which provide a certain theoretical basis for pattern recognition of pests. At present, the machine recognition of crop pests is mainly based on images. The image-based recognition methods are roughly divided into four categories. One is based on invariant moments. For example, Diao Zhihua et al. introduced invariant moment theory into shape feature extraction, defined the shape feature of wheat leaf disease images extracted from the instantaneous parameters of 7 Hu invariant moments and applied to the designed wheat disease intelligent recognition system to obtain a high recognition rate[4]. The second is based on fuzzy clustering. For example, Vinushree et al. used KFCM to judge the density of insects in plants[5]. The third is support vector machine. For example, Wu Xiang used the corner detection to clip the original image, then segmented the Otsu algorithm's clipping image, and extracted the image features of the pest target in the segmented image by SURF algorithm. Finally, the recognition of 10 types of pest images such as Pieris rapae was realized by the SVM classifier[6]. Fourth, based on neural networks, such as Gassoumi et al. used a neural network method to classify insects in the cotton ecosystem and achieved 90% accuracy[7].

The Convolutional Neural Network (CNN) is a feedforward multi-layer neural network, which is a kind of supervised learning. It is good at dealing with related machine learning problems of images, especially large images. Through a series of methods, CNN can reduce the image with massive data recognition dimensions and finally enable it to be trained. CNN was first implemented by LeNet, which was proposed by LeCun in 1998[8]. LeNet is a classic convolutional neural network used to identify handwritten digits (the MNIST database). Although its network structure is small, it contains convolutional layer, pooling layer, and fully connected layer, which constitute the basic components of a modern CNN. In 2014, Szegedy et al. proposed GoogLeNet that can extract more features with the same amount of computation to optimizing training results[9]. In recent research work, convolutional neural networks have been widely used in image recognition for face detection and license plate recognition. However, the image recognition of pests is more complex and more difficult than other machine vision applications due to the small size, wide varieties, and large differences among species. The research on the application of convolutional neural network in pest image recognition has yet to be done.

2. Methods

2.1. Bag-of-words model (BOW model)

In recent years, with the development of computer vision technology and artificial intelligence technology, the bag-of-words model has been widely used in computer vision. In the image classification technology based on the BOW model, the three parts of feature extraction, visual dictionary construction and classifier training are generally included. The SIFT feature extraction adopted in this study mainly extracts global or local invariant features from a given image set to obtain the representation of the image. The construction of the visual dictionary is mainly to cluster the extracted image features, and the cluster center is used as a visual word, and the set of all cluster centers is the visual dictionary of this construct. Classifier training is an operation aimed at the image classification task. The classification and recognition of the image can be performed by the trained classifier.

2.2. Convolutional Neural Network(CNN) - transfer learning of the GoogLeNet model

This study uses the Inception-v3 model in GoogLeNet for transfer learning. The transfer learning of model refers to using tens of millions of images to train an image recognition system. When we encounter a new image field, we can transfer the original image recognition system to a new field, and the same effect can be achieved with less image training in the new field. It is difficult to find sufficient training data on crop pests. However, through transfer learning, the models trained from other data sources can be modified and improved to apply to similar fields, thus greatly alleviating the problems caused by insufficient data sources. The GoogLeNet model used in this experiment was based on the ImageNet dataset training. At present, there are more than 14 million images in the ImageNet dataset, covering more than 20,000 categories. Among them, there are 3,998 species in the animal category, including some insect pests. Therefore, the model trained with the ImageNet Dataset is feasible in the pest identification of this project. The transfer learning will get better results, and the training time is less. The basic idea is to use the pre-trained GoogLeNet model to find a layer in the model that can output reusable features, and then use the output of the layer as an input feature to train those smaller neural networks that require fewer parameters.

3. Results and analysis

The training results of the traditional bag-of-words model can be seen from the Figure 1. The average classification accuracy of 7 species, such as pomacea canaliculata, gryllidae, henosepilachna vigintioctomaculata, thrips, cnaphalocrocis medinalis, trialeurodes vaporariorum and acrida cinerea, was over 80%. The classification accuracy of cnaphalocrocis medinalis was 97%. However, the classification accuracy of cutworm, Spodoptera litura, Rhynchocoris humeralis, cletus punctiger dallas and achatina fulica was relatively low. The average classification accuracy of the traditional bag-of-words model can reach 56.41%.

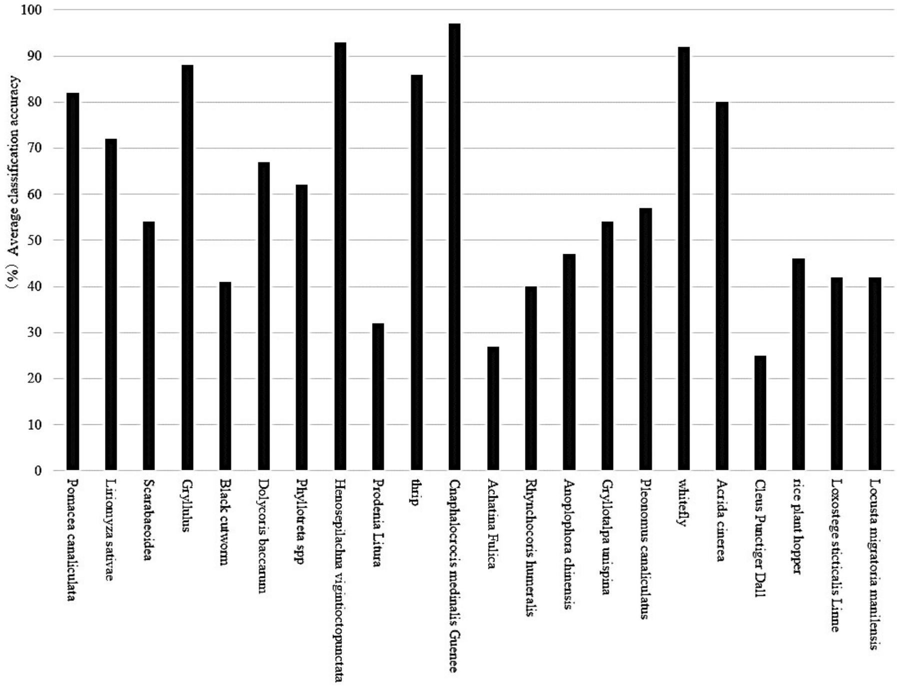

Figure 1. Classification and identification of 22 kinds of crop pests based on BOW model.

This study used the Inception-v3 model in GoogLeNet for transfer learning, which was trained by the ImageNet dataset. The Figure 2 shows the identification of two example crop pests, planthopper and prodenia litura, using the Inception-v3 model transferred to mobile phones. The judgment result in the Figure 2 of planthopper means that the probability of being planthoppers is 0.6094699, and the probability of being whitefly is 0.33967218. That is, the recognition result of this image is planthoppers, and the judgment result is correct. In the same way, the judgment result in the figure of prodenia litura means that the probability of the pest being prodenia litura is 0.98734725, and the judgment result is also correct. It can be concluded from the training results that the GoogLeNet model has a high correct rate, and the effect of transfer learning is good. The overall mean recognition accuracy can reach 96.35%.

Figure 2. The result of identifying 2 kinds of crop pests with GoogLeNet model.

Comparing the above training results, it can be found that the traditional bag-of-words model is susceptible to complex background and other factors, which is not conducive to determining the correlation between features and information redundancy, and is not conducive to subsequent classifier training. The GoogLeNet model is based on ImageNet dataset, which has large amount of data, a wide range of image categories, including most pests, and then using our pest dataset for transfer learning. Compared with the traditional bag-of-words model, the model we trained has higher precision and better anti-interference ability.

4. Conclusion

In this study, 2200 pest images collected were used as experimental samples. We mainly adopted neural network as our research idea. Based on the Tensorflow deep learning framework, an image database of crop pests was established through image flipping, random clipping and zero-padding, color jittering and saliency image segmentation in Pycharm. Through transfer learning, the features of crop pests were extracted from the optimized GoogLeNet model, and a classifier for crop pests image recognition was established. The experimental results show that the average classification accuracy of the traditional bag-of-words model can reach more than 56.41%, and the GoogLeNet model recognition accuracy can reach 96.35%. The GoogLeNet model has higher precision and stronger anti-interference ability than the traditional bag-of-words model.

Acknowledgments

This study was financially supported by National innovation training program for College Students: Research on the accurate recognition algorithm of main agricultural pests based on convolution neural network(A093).

References

[1] Zihe Li. (2012) Natural Control of Agricultural Pests and Diseases. J. Heilongjiang Science and Technology Information., 36: 246.

[2] Ruihua Wang. (2017) Techniques and Recommendations for Crop Pest Control. J. Agricultural Science-Technology and Information., 16: 72.

[3] Cheng, X., Zhang, Y., Chen, Y., et al. (2017) Pest identification via deep residual learning in complex background. J. Computers and Electronics in Agriculture., 141: 351-356.

[4] Jingjing Wang, Zhang Wu, Lianzhong Liu, Shuai Huang. (2014) Summary of crop diseases and pests image recognition technology. J. Computer Engineering & Science., 36(07): 1363-1370.

[5] Vinushree, N., Hemalatha, B., Kaliappan, V. K. (2014) Efficient Kernel-Based Fuzzy C-Means Clustering for Pest Detection and Classification. In: Proceedings of the 2014Computing and Communication Technologies (WCCCT). Tamilnadu.
[6] Xiang Wu. (2016) Study on Identification of Pests Based on Machine Vision. http://www.zju.edu.cn/.
[7] Gassoumi, H., Prasad, N.R., Ellington, J.J. (2000) Neural Network-Based Approach For Insect Classification In Cotton Ecosystems. In: International Conference on Intelligent Technologies (InTech 2000). Bangkok.
[8] Lecun, Y., Bottou, L., Bengio, Y., et al. (1998) Gradient-based learning applied to document recognition. J. Proceedings of the IEEE., 86(11): 2278-2324.

Research on the Optimization of the Spatial Pattern and Layout of Mountainous Rural Residents Area under the Background of Rural Revitalization： Taking Tianquan County of Ya'an as an example

LI Junhui, LUO LIjun, LI Hao, WEI Yali
College of Resources, Sichuan Agricultural University, Sichuan Province, China
Correspondent Author： WEI Yali, Email:weiyali@sicau.edu.cn

Abstract Exploring the spatial pattern and optimization strategy of rural settlements in mountainous areas is conducive to enhancing the rural human settlement environment and carrying out national land space governance and planning. This article took Tianquan County of Ya'an, a mountainous and hilly areas as an example; conducted spatial statistics supported by ArcGIS; used Moran index, kernel density analysis, and settlement layout suitability evaluation to make an integrated analysis on the spatial pattern of mountainous rural settlements and the layout optimization strategy. The results showed that: ① Rural settlements in the study area were generally scattered on a large scale and aggregated on a small scale. with Renyi Township and Hongjun Village as the center, the density decreased from the inside to the outside. ② Through the evaluation of layout suitability, the study area was divided into four types of renovation patterns: leading demonstration type, optimized development type, restricted development type and ecological migration type. The research results can provide decision reference for the layout optimization of rural residential areas and village planning in mountainous areas.
Key words Mountainous and hilly areas; Rural Residential Area; Spatial pattern; Layout optimization

1 Introduction

Rural residential area is the key element that constitute urban and rural construction land due to their long-term production and life functions of agricultural population[1-2] . Exploring strategies for optimizing the layout of the rural settlements in the vast but underdeveloped mountainous areas would be helpful to guide the layout of rural areas in special areas, help national land space governance and ecological protection and promote the rural revitalization strategy, which have important practical significance.

Foreign scholars started to study the distribution of rural residential area earlier. They mostly focus on the factors that affect the distribution and location of settlements, and put forward relevant theories about the layout optimization of settlements[3-6] . Domestic scholars have also carried out extensive exploration of the spatial pattern of settlements, such as introducing Voronoi diagram[7] , geographic concentration index[8] , kernel density[9] , geographic detector[10] and other models for analysis. Taking a wide view of the existing achievements, most of the research has focused on the study of the spatial pattern and influencing factors of traditional agricultural areas, but the corresponding optimization strategies of mountainous and hilly areas are still insufficient. In view of this, this article was guided by rural revitalization and took Tianquan County of Ya'an as an example to explore and summarize the layout experience of special landform areas in mountains and hills, integrating a variety of measurement and spatial analysis methods to carry out a study on the optimal layout of mountainous rural residential areas.

2 Study area and data

2.1 Study Area Overview

Tianquan County is belonging to Ya'an City and is located on the western edge of the Sichuan Basin, at the foot of the Erlang Mountain, the upstream of the Qingyi River (102°16′E~102°55′E, 29°49′N~30°21′N). The terrain is high in the northwest and low in the southeast. The mountainous area of the county accounts for 98.5% and the hills are 1.5%, which shows that Tianquan County is a veritable mountainous and hilly area. Influenced by the subtropical monsoon climate, Tianquan County was warm, rainy and with distinct seasons. This region has 2 towns and 13 townships, covers an area of 239029.05hm^2 in 2017, and was run through by Sichuan-Tibet Railway and Ya'an-Kangding Expressway.

2.2 Data Source

①This study used the 2017 Tianquan County Land Use Change Survey Database (1:10000) to obtain data on various types of areas such as rural settlements, roads, and water areas in the study area. ②The study used 30m DEM of Tianquan County to obtained the elevation, slope, etc. (The data set is provided by Geospatial Data Cloud site, Computer Network Information Center, Chinese Academy of Sciences. (http://www.gscloud.cn)) ③ Based on the ArcGIS 10.2 platform, the study used the "Tianquan County Overall Land Use Planning Map(2006-2020)(2014 revised edition)" to spatially register and correct the collected data, then used 3D analysis to obtain the elevation and slope data of the study area.

3 Methods

① This article introduced the kernel density model to reflect the spatial layout of the neighborhood of rural settlements [11] . The Rosenblatt-Parzen kernel estimation model was used to measure the local density characteristics of rural settlements through the distance decay function, and the peaks and cores of the elements were used to create a continuous surface for visual expression [12] . ②This article also used the Global Moran's I to analyze the aggregation or difference of rural settlements [13], and used the Anselin Local Moran's I to make up the insufficiency that global spatial autocorrelation analysis can't visualize the spatial distribution characteristics of the target objects, and analyze whether there are gather-disperse characteristics in the rural settlements. ③ To establish a suitability index system for the layout of the rural residential area in Tianquan County, this article evaluated the layout suitability of the rural settlements by using the range standardization and entropy method [14-15], and selected 4 factor levels including natural resources, production and life, socio-economic and residential patch factors in 9 index levels(Table 1).

Table 1 Index System of Suitability Evaluation of Residential Area Distribution in Tianquan County

Object level	Factor level	Index level	Index classification and data standardization division					Weight
			5	4	3	2	1	
Comprehensive influence evaluation	Natural resource	Elevation	＜300	300～500	500～800	800～1200	＞1200	0.070
		Aspect	flat, south	southeast, southwest	east, west	northeast, northwest	north	0.134
		Slope	＜3	3～8	8～15	15～25	＞25	0.128
	Socio-economy	Distance to road	＜300	300～800	800～1500	1500～2500	＞2500	0.222
		Distance to designated town	＜1000	1000～2000	2000～3000	3000～4000	＞4000	0.225
	Production&living	Distance to river&reservoir	＜500	500～1000	1000～1500	1500～2000	＞2000	0.054
		Distance to tillage	＜500	500～1000	1000～1500	1500～2000	＞2000	0.026
	Residential patch	Residential fractal dimension index	＜1	1～2	2～3	3～4	＞4	0.003
		Residential patch shape index	＜1	1～5	5～10	10～15	＞15	0.137

4 Results and analysis

4.1 Residential area density analysis results

Rural residential area in Tianquan County were generally scattered, with poor urban-rural integration, and only clustered around individual villages with superior comprehensive conditions. The high kernel density estimated values of the rural residential area were mostly spatially distributed in the east. The distribution of the estimated value presented a trend of decay from the surrounding villages around the designated towns such as Laochang Township, Renyi Township, and Maxi Village as the center to the surroundings, whose values could reach 33~43 pieces/km². Lianglu Township in the southwest is far away from the designated towns, but due to the development of its characteristic tourism, a certain concentration has been formed, with a kernel density value of 15-20 pieces /km². Leying Township, Xinchang Township and Duogong Township had high kernel density, mainly concentrated at 26~33 pieces /km2. And the kernel density value of the north was the least, mainly at 0~7 pieces /km2.

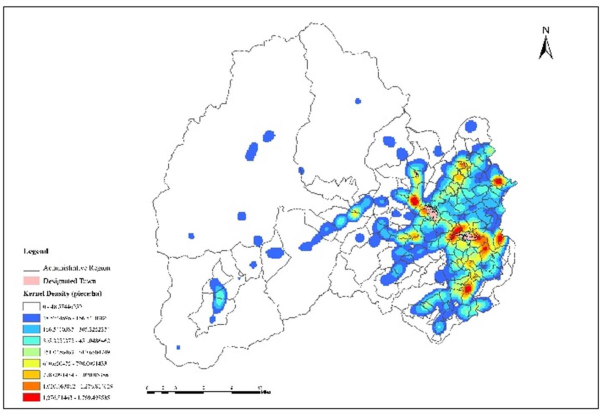

Figure 1 Analysis of Tianquan County kernel density

4.2 Moran's I Analysis

4.2.1 Global Moran's I Analysis

The global Moran's I analysis showed that: Since the P value was less than 0.001 with a 99% confidence level and Moran's I was equal to 0.377> 0, it can be considered that the residential area showed the distribution feature of spatial autocorrelation. And Z(I) was greater than the critical value equal to 2.58, showing a spatial aggregating feature with high autocorrelation and significant structure differentiation[16] .

4.2.2 Local Moran's I Analysis

This article used LISA cluster map to determine the specific location and spatial difference characteristics of every settlement. The analysis result showed that the rural settlements was spatially high-value agglomeration. The P(d) was less than 0.01, and the score of Z(d) was 26.75, which was much greater than the critical value of 2.58, indicating that the test results showed highly significant with a confidence level greater than 99% and wasn't randomly generated. Therefore, the spatial scale of the rural residential area in Tianquan County had a significant high value on macroscale, and the agglomeration was distributed in the eastern county, while the southwestern area showed a small amount of high-value agglomeration.

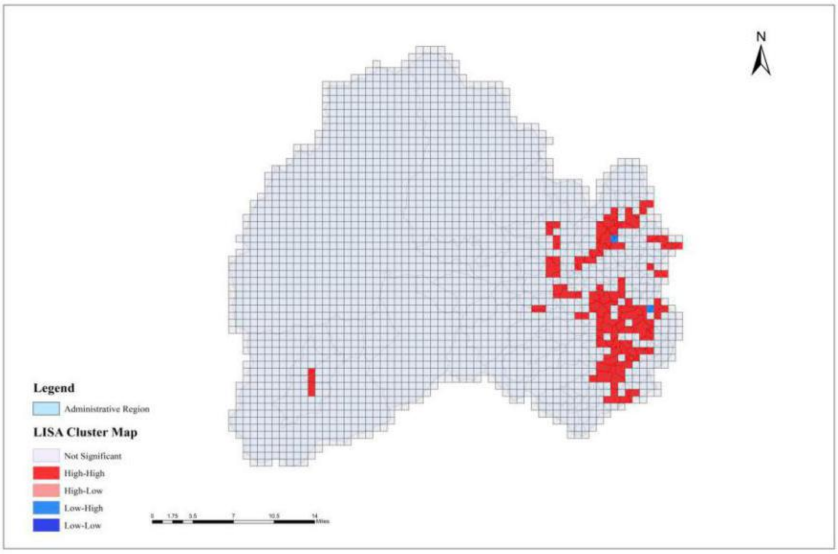

Figure 2 LISA Cluster Map of Tianquan County

4.3 Layout Suitability Analysis

Rural residential areas of different suitability levels should adopt different renovation methods. Areas with suitable natural conditions and good infrastructure such as transportation should carry out on-site renovation pattern; however, areas with poor ecological environment and weak infrastructure should focus on ecological protection and carry out ecological migration renovation pattern. Using the GIS spatial analysis function, the single-factor layers were weighted and superimposed to obtain a comprehensive evaluation result, and the reclassification was performed by the equal-interval method to obtain a suitability distribution map for rural residential land comprehensive renovation (Figure 3). According to the development and restriction conditions of areas with different suitability levels, the rural settlements renovation patterns were divided into four types: leading demonstration type, optimized development type, restricted development type and ecological migration type (Table 2).

Table2 layout suitability patterns of rural settlements and area percentages

Index	Highly suitable area	Moderately suitable area	Low suitable area	Unsuitable area
Score interval	5～4	4～3	3～2	2～1
Optimization mode	Leading demonstration type	Optimsized development type	Restricted development type	Ecological migration type
area/hm²	94.650	569.070	920.637	170.761
pecentage/％	5.4	32.4	52.5	9.7

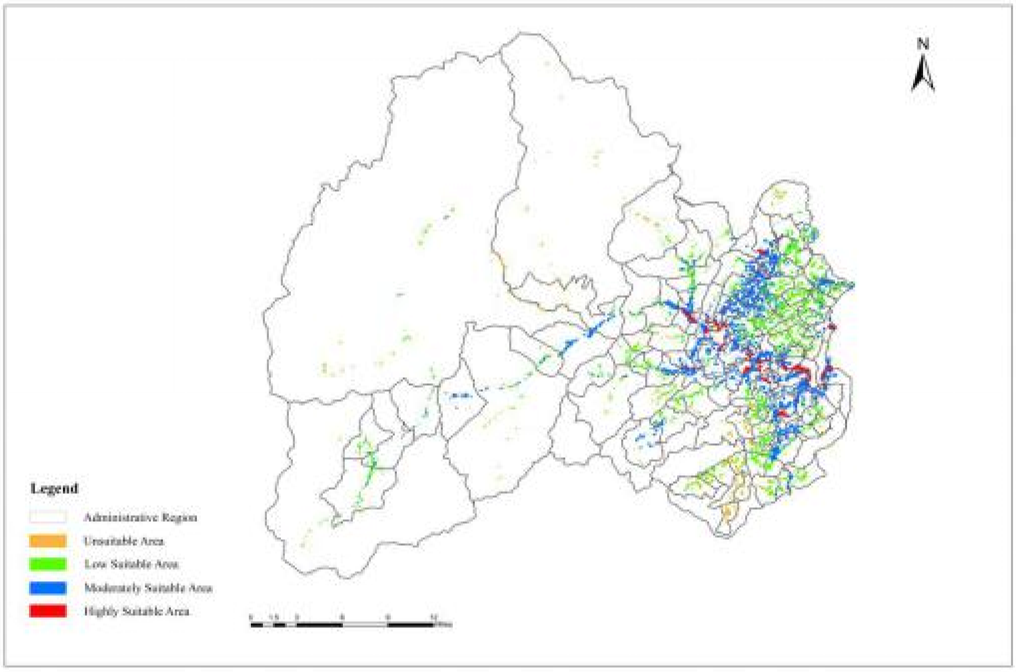

Figure 3 Evaluation of residential area in Tianquan County

(1) Leading demonstration type: This type of rural residential land had the highest suitability, with an area of 94.650 hm^2, accounting for 5.4% of the total area, and most of it was located in the hilly valley near

the market town. With superior geographical location, rich natural resources, relatively complete infrastructure, strong industrial support and collective economic strength, this area could promote the adjustment and upgrading of rural industrial structure, help the industry prosper, strengthen the construction of rural ecological civilization and human settlement environment, promote farming culture to integrate with modern civilization and provide protection for ecological livability.

(2) Optimized development type: The rural residential area of this type covered an area of 569.070hm^2, accounting for 32.4% of the total area. This area was mainly distributed in the secondary area radiated by the center of the town and mostly performed zonal distribution along the highway, which had great potential for renovation. It is necessary to sort out the traffic network, consolidate farmland, improve the infrastructure configuration, coordinate the land use structure, and reserve the construction land for the resettlement of the villagers. It is the key area of renovation at this stage.

(3) Restricted development type: This type of rural residential area covered an area of 920.637 hm^2, accounting for 52.5% of the total area, and was mostly located in areas with high altitudes, complex terrain, and relatively weak infrastructure. In the process of renovating such rural settlements, attention should be paid to scale control, with the emphasis on the protection of the ecological environment. By Reinvigorating the idle homesteads in the villages and promoting renovation and reconstruction, the rural residential environment will be gradually improved.

(4) Ecological migration type: This type had the lowest suitability, with an area of 170.761hm^2, accounting for 9.7% of the total area. Most of them were located at the area whose slope is greater than 20°, with high altitude, fragile ecological environment, and prone to geological disasters, which was not suitable for planning rural settlements. Through measures such as gradual poverty alleviation and relocation, ecological immigration and so on, rural residents should be actively encouraged to relocate, and attach importance to former site reclamation and ecological restoration to achieve the good ecological and environmental value functions.

5 Conclusion

The renovation of rural residential area in mountainous and hilly areas is of great significance to the optimize human settlements and construct ecological corridors in underdeveloped areas. This study took Tianquan County as an example, combing the spatial pattern of rural residential area, and evaluated the layout suitability from four aspects to explore the layout optimization model of settlements. The main conclusions are as follows.

(1) Rural residential area in mountainous and hilly areas were characterized by generally scattered on a large scale and aggregated on a small scale, with decreasing radiation from inside to outside; and the eastern, southeastern and southwestern areas showed radial H-H type aggregation; and the remaining settlements didn’t present regular distribution affected by ecological conservation and socio-economic restrictions.

(2) According to the evaluation results of the layout suitability of the rural residential area, the areas are divided into four types of renovation patterns: leading demonstration type, optimized development type, restricted development type and ecological migration type.

References

[1]Zhou Guohua, He Yanhua, Tang Chengli, et al. On Rural Settlement Patterns in New Era, progress in Geography, 2010,29(2):186-492(in Chinese)

[2]Wang Ting, Zhou Guohua, Yang Yan. Study on the Rational Distribution of the Rural Residential Land in Nanyue District of Hengyang. progress in Geography, 2008, 27(6): 25-31(in Chinese)

[3]Skinner G W. Marketing and social structure in rural China, Part I. The Joumal of Asian Studies, 1964,24 (01): 3-43

[4]Tomida M, Hoshino K,Tomida M,et al. An introduction to rural geography. E. Arnold,1985:331-341

[5]Peter S Robinson. Im plications of rural sett lement patterns for development: a historical case study in Qaukeni, Eastern C ape, South Africa. Development Southern Africa, 2003,20(3): 405-421

[6]Inge Tiwrsenl, Jan Ub. Modeling Residential Location Choice in An Area with Spatial l3artiers. Regional Society. 2002，36(6):13- 64

[7]Cao Yingchun, Zhang Yukun. Study on the Rational Distribution of the Rural Residential Land in Nanyue District of Hengyang. Journal of Hebei University (Natural Science Edtion).2014, 34(02): 129-136(in Chinese)

[8]Zheng Xiandi, Zheng Huawei, Liu Youzhao. Spatial Distribution and Influencing Factors of Rural Residential Areas Based on GIS: A Case Study of Xuyi County. Chinese Journal of Soil Science.2019, 50(03): 550-554(in Chinese)

[9]Kou Xiangwei, Li Weifang, Jiang Wenda, et al. Spatial Pattern and Influence Factors of Rural Settlements: a Case of Siming Mountain Area in Ningbo City. Mountain Research.2018, 36(05): 126-133(in Chinese)

[10]Huang Cong, Zhao Xiaomin, Guo Xi, et al. Study on the rural residential layout optimization of Yujiang County based on kernel density. Joumal of China Agricultural University.2016,21(11):165-174(in Chinese)

[11]Zhang Shuai, Xu Jian, Yang Ping. Research on Space-time Evolution Characteristics of Rural Settlements in Multi-ethnic Areas: A Case Study of Binchuan County, Yunnan Province. Yunnan Geographic Environment Research.2019, 31(03): 30-37(in Chinese)

[12]Liu Dajun, Hu Jing, Chen Junzi, et al. The Study of Spatial Distribution Pattern of Traditional Villages in China. China Population, Resources and Environment. 2014,24(4):157- 162(in Chinese)

[13]Wu Jian, Cheng Peng, Wen Chaoxiang, et al. Ecological risk assessment of land use based on exploratory spatial data analysis (ESDA) : A case study of Haitan Island, Fujian Province. Chinese Journal of Applied Ecology.2014, 25(7) : 2056- 2062.(in Chinese)

[14]Cao Zhi, Liu Yansui, Li Yurui, et al. The spatial pattern of Chinese specialty villages and its influencing factors. Acta Geographical Sinica:1-18[2020-06-13]. http://kns.cnki. net/kcms/detail/11.1856.p.20191113.0847.002.html.

[15]Cai Yusheng, Lu Jingwei. Evaluation of Regional Development Quality in Beijing-Tianjin-Hebei Region Based on Entropy Method. Journal of Industrial Technological Economics. 2018, 37(11): 67-74. (in Chinese)

[16]Chen Zongfeng, Li Yurui, Liu Yansui. Distribution pattern characteristic and type classification of rural settlements in loess hilly-gully region. Chinese Society of Agricultural Engineering.2017, 33(14): 266-274+316(in Chinese)

中华人民共和国国家版权局

计算机软件著作权登记证书

证书号： 软著登字第4654991号

软 件 名 称： 配肥宝小程序系统
［简称：配肥宝小程序］
V1.0

著 作 权 人： 罗莉钧;肖逸;王天骄

开发完成日期： 2019年09月12日

首次发表日期： 2019年09月20日

权利取得方式： 原始取得

权 利 范 围： 全部权利

登　记　号： 2019SR1234234

根据《计算机软件保护条例》和《计算机软件著作权登记办法》的规定，经中国版权保护中心审核，对以上事项予以登记。

中华人民共和国国家版权局
计算机软件著作权
登记专用章

No. 04875963

2019年11月29日

中华人民共和国国家版权局

计算机软件著作权登记证书

（副本）　　证书号：　软著登字第4655016号

软 件 名 称：　配肥宝网站开发系统
[简称：配肥宝网站]
V1.0

著 作 权 人：　王天骄;罗莉钧;肖逸

开发完成日期：　2019年09月17日

首次发表日期：　2019年09月21日

权利取得方式：　原始取得

权 利 范 围：　全部权利

登　　记　　号：　2019SR1234259

根据《计算机软件保护条例》和《计算机软件著作权登记办法》的规定，经中国版权保护中心审核，对以上事项予以登记。

中华人民共和国国家版权局
计算机软件著作权
登记专用章

No. 04875983

2019年11月29日

中华人民共和国国家版权局

计算机软件著作权登记证书

证书号：软著登字第5283720号

软 件 名 称：土壤识别app内置系统
V1.0

著 作 权 人：肖逸;王天骄;罗莉钧

开发完成日期：2020年03月01日

首次发表日期：2020年03月10日

权利取得方式：原始取得

权 利 范 围：全部权利

登 记 号：2020SR0405024

根据《计算机软件保护条例》和《计算机软件著作权登记办法》的规定，经中国版权保护中心审核，对以上事项予以登记。

中华人民共和国国家版权局
计算机软件著作权
登记专用章

No. 05627492

2020年04月30日

中华人民共和国国家版权局

计算机软件著作权登记证书

证书号：软著登字第4996069号

软件名称：基于无人机遥感影像的土地利用类型识别系统
[简称：土地利用类型识别系统]
V1.0

著作权人：四川农业大学;古玲;徐精文;罗莉钧;张怀巍

开发完成日期：2019年11月30日

首次发表日期：未发表

权利取得方式：原始取得

权利范围：全部权利

登记号：2020SR0117373

根据《计算机软件保护条例》和《计算机软件著作权登记办法》的规定，经中国版权保护中心审核，对以上事项予以登记。

中华人民共和国国家版权局
计算机软件著作权
登记专用章

No. 05292917

2020年01月22日