

The Recognition of Rice Area Images by UAV Based on Deep Learning

Wei Hongyuan¹, Mao Jian¹

¹School of mechanical and automotive engineering, Shanghai University Of Engineering Science, Shanghai 201620

Abstract. Aiming at the target detection of remote sensing rice field of uav, the image of large-size uav is firstly segmented, and the type of each image is manually identified, and the image training set and verification set are made. Then, the training model of convolutional neural network is realized by python programming. The advantage and disadvantage of the two-layer convolutional neural network and ResNet50 are compared, and it is found that the training set is less and the picture feature complexity is not high in practical application. In the end, the feature recognition of rice field is realized, which has certain application value.

1 INTRODUCTION

China is a large agricultural country. As a major grain crop in China, the annual output and consumption of rice are among the top in the world. The planting of rice, removing impurity and anti-insect all have important research value. In traditional agricultural chores, operators often carry agricultural sprayers on their backs after the crop has evaporated. Time is near noon, high temperature makes pesticide volatilization acceleration, easy to cause pesticide poisoning. In recent years, with the development of unmanned aerial vehicles (uavs), agricultural uavs with high aging and mobility gradually replaced this work [1-2]. However, agricultural unmanned aerial vehicles (uavs) only rely on the rice area given by the ground station for drug application, which often leads to problems such as uneven application and drug waste, etc. and the accurate application of the medicine still needs a lot of research [3]. Domestic and foreign scholars have a lot of relevant research.

Xundong and zhang jing [4] et al. studied that the control effect of th80-1 plant protection uav on flying louse, leaf roller and rice sheath blight was 91.04%, 87.33% and 91.22%, respectively, after 14 days of application. It also confirms that drone spraying is more effective than traditional manual spraying. Xiao xiaohua, liu chun [5] et al. also compared uavs with different sprayers, and found through experiments that the three common rice disaster prevention effects were from superior to inferior: uavs, electrostatic sprayers, motor sprayers and electric sprayers. Xue xinyu and tu kang [6] et al. conducted more specific experiments, spraying crops with two pesticides, chlorpyrifos and hexazolol, and found that the effect of aerial drug application on the microstructure of rice grains was less than that of

conventional drug application. All the above studies have shown that drone spraying is more effective than the traditional method. In the area identification of rice, li Ming and huang yuqi [7] et al. established the binary classification Logistic regression model of rice. Tadasi CHOSA, Kiyoshi MIYAGAWA [8] et al. took images of the rice field near miaogao by using the normalized vegetation index to identify the rice area and quantified the growth variability among the farmland. Cesar Analide, Pankoo Kim [9] et al. proposed a color-based segmentation algorithm and used k-means clustering technology to detect the rice ear area with an accuracy up to 84.2%.

Based on the previous research, it can be seen that the use of drones to spray pesticides is faster and more effective than the traditional methods. It also has less impact on crops. In rice area measurement, quality testing and other aspects, different algorithms can be used to process the pictures taken by the uav to achieve the goal. This is also the only way for uav to achieve precision agricultural spraying. In this paper, the convolution neural network is used to study and classify the uav images so as to identify the rice and weeds in the rice field.

2 DEEP LEARNING METHOD

Hinton [10] et al. first proposed the concept of deep learning in Science, pointing out that the artificial neural network has good feature learning ability when it has multiple hidden layers. It is also pointed out that deep structure needs to be designed in order to learn complex functions which can represent high-level abstract features. However, the deep structure also means that more parameters need to be trained, which makes the image

processing slightly weak, resulting in the emergence of convolutional neural network.

2.1 Convolutional Neural Network

Convolutional neural network is a kind of deep neural network with convolution structure. Its principle is to

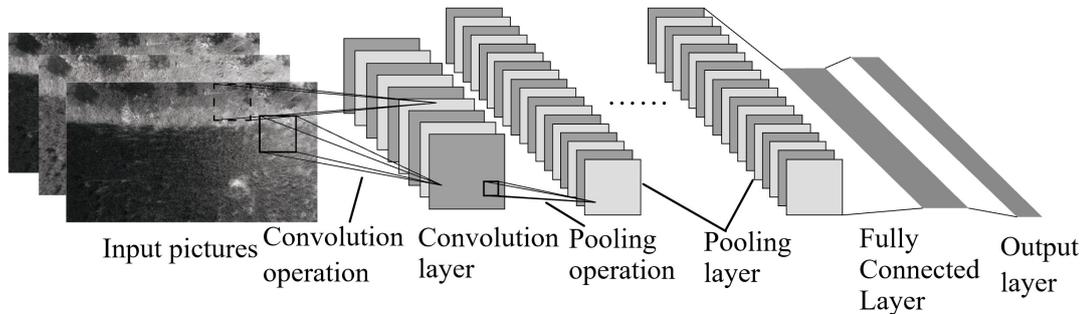


Figure 1. Work model of convolutional neural network

The convolution kernel and pooling kernel are introduced in the comparison between convolutional neural network and deep neural network. In the deep neural network, all pixels of the image are connected, and there are a lot of neuron parameters. In the convolutional neural network, the convolution kernel is just like the sensory field in human vision. It is a block of size 3*3 or 5*5, which can be used to extract the features in the input image and traverse all the information in the image by translation. Therefore, the information set in each pixel of the convolution layer is the information set in a certain region of the input layer. When the original image passes through the convolution layer, only feature extraction is realized, and the feature dimension of the convolution layer is still large. Pooling operation can solve this problem, which not only reduces the number of parameters but also maintains the local invariance of the feature. After a convolution pooling operation, it is often

introduce convolution operation into neural network and improve the shortcomings of the original neural network with a large number of parameters[11]. At the same time, it can extract deeper features and reduce the overfitting problem in common neural network. Figure1 shows the working model of a convolutional neural network.

not enough to meet the classification requirements, so the pooling layer can be continued to perform convolution operation. The more times of iteration, the deeper information can be extracted. After the iteration is completed, the data obtained before each field is integrated by the full connection layer, and all features are highly purified. Finally, softmax layer is used for image classification to determine the category of the input image.

2.2 Network selection

In deep learning, convolutional neural network has many models, among which AlexNet, GoogLeNet, VGG and ResNet[12] are successful. The parameters of several networks are shown in Table1.

Table 1. The parameters of different network

Network name	Number of parameters	Top1 accuracy	Top5 accuracy	depth
AlexNet	About $6 \cdot 10^7$	62.5%	83%	11
VGG16	About $13.8 \cdot 10^7$	71.5%	91%	23
GoogLeNet	About $5 \cdot 10^6$	72.60%	92.88%	22
ResNet50	About $2.5 \cdot 10^7$	75.9%	92.9%	168
Double convolutional layer neural network	About $2 \cdot 10^4$	None	None	7

As can be seen from the table, AlexNet proposed in 2012, as an 11-layer network, has achieved an accuracy of 83%, basically meeting the needs of object recognition. In 2014, with VGG and GoogLeNet, the complexity of the network has increased by a level, and its precision has also increased significantly. ResNet proposed in 2015 adopted residual neural network, so although the number of layers is very deep, the number of parameters is not very large compared with other networks. However, the above networks are all those that have achieved excellent results in ImageNet competition, and their training sets are relatively large. If you want to use these networks, you need to take advantage of the model parameters already trained by others, and it will also take

a lot of time. Facing the small-batch data of this paper, using a network with a shallow number of layers and a small number of parameters can also achieve the desired effect. Therefore, this paper compares the training results of ResNet50 and the dual-convolutional layer neural network in order to find the most suitable network to train the uav pictures.

3 ORIGINAL DATA ANALYSIS

3.1 data processing and analysis

The research area is located in tailai village, qingpu district, Shanghai (121.10842E, 31.11456N). The rice planting area in tailai village is about 440 mu. 8 paddy fields in different locations were randomly selected for shooting. Dji's genie 4pro, equipped with a FC6310 camera, was selected for direct shooting at 5 meters and 10 meters above the rice field. The size of the image was 5472*3078. This large image cannot be trained directly because it has too many pixels and contains too much information. If the convolutional neural network is used directly, the number of networks required is large, and the number of parameters must be a huge number, which makes the training impossible. If the image is reduced to the required size directly through scaling, the compression ratio is too large, and a lot of information will be lost, resulting in the failure of network learning. Therefore, the method selected in this paper is to segment each image first. The segmented image size is 456*342, and then the images are scaled according to the network input requirements, and are classified by visual

classification, so as to meet the requirements of the training set.

In the training of complex convolutional neural network, in order to reach 1000 classification categories, the data volume of standard data sets is large, and the proportion of each picture is similar, which can provide deeper features and guarantee the diversity of network. In this paper, the characteristics of rice fields in the agricultural environment are analyzed, and only a thousand pictures are enough to classify the five characteristics. A total of 1613 valid images were divided into the photos of the uav. There were 7 types of pictures after classification, of which the proportion of images of debris and boundary terrain was less than 2%, which could not be trained. Therefore, the actual number of classification is only 5. Four of them are shown in Figure 2. The ratio of training set to verification set is 7:1, and there are 1369 photos in total. The number of class images is shown in Table 2

Table 2. The number of different kinds of pictures

category	rice	weed	wasteland	rice and wasteland	weed and wasteland	Boundary topography	mixed terrain	sum total
quantity	678	400	185	168	135	23	24	1613
proportion	41.5%	24.5%	11.4%	10.5%	8.2%	1.4%	1.4%	
training set	593	350	161	147	118	0	0	1369



(A) rice



(B) weed



(C) wasteland



(D) rice and wasteland

Figure 2. Pictures of different kinds of rice fields

4 SIMULATION

4.1 comparison of different network classification

The construction of convolutional neural network often relies on some frameworks, such as TensorFlow and Caffe. These frameworks are often programmed through Python language, which is an object-oriented and literal computer language. There are applications in the fields of web production, machine learning, data mining, operation and maintenance development [13]. The implementation framework of the network in this paper is TensorFlow. The convolution network training of data in the software is divided into four steps, which are data preprocessing, input pictures, network training and testing network.

In the data preprocessing, the size of the image in the data set is changed, so that the image can meet the needs of network training and be normalized. The input image is mainly a bridge role, the function of the specific implementation is to concentrate training images in different groups were sent to the network for training, the training of the network training is a major step step, including image convolution, pooling operations, such as after the forward travel and back propagation network parameters were optimized, until meet the requirements of the iteration. The final test network is to test the training modules. Verify network accuracy. The overall flow chart of the software is shown in Figure 3.

The five images in the data set were placed under different folders and trained with two network models respectively. The results are shown in Tab.3. Among them, ResNet50's training takes a long time and is not convenient to record, so it is only represented as a long time.

Table 3. Comparison of two kinds of network training results

network model	Training set accuracy	verification set Accuracy	training time	Classification time
Double convolutional neural network	100%	94.31%	short	0.76s
ResNet50	100%	96.2%	long	1.35s

As can be seen from the table, both of them reached 100% in the training set accuracy. This is because the training set quantity is small and the characteristic quantity is not particularly complicated, so both networks can extract enough information for classification. In terms of training set accuracy, ResNet50 does not differ much from the double-layer network. Two reasons are taken into account here. The number of images of rice in the training set was too large, so it was possible that overfitting occurred in the network. Second, the number of images is too small to make ResNet50 do a good parameter update. Finally, when calling the network for

image recognition, the time of two-layer network is significantly shorter than ResNet50, because of the low complexity of two-layer network.

To sum up, ResNet50 is not suitable for the data set in this paper. When selecting the network, the actual situation of the data set should be considered. In the case of a small number of data sets and a small number of recognition types, image recognition can be carried out through a simple network. Therefore, this paper chose the two-layer convolutional neural network for subsequent experiments.

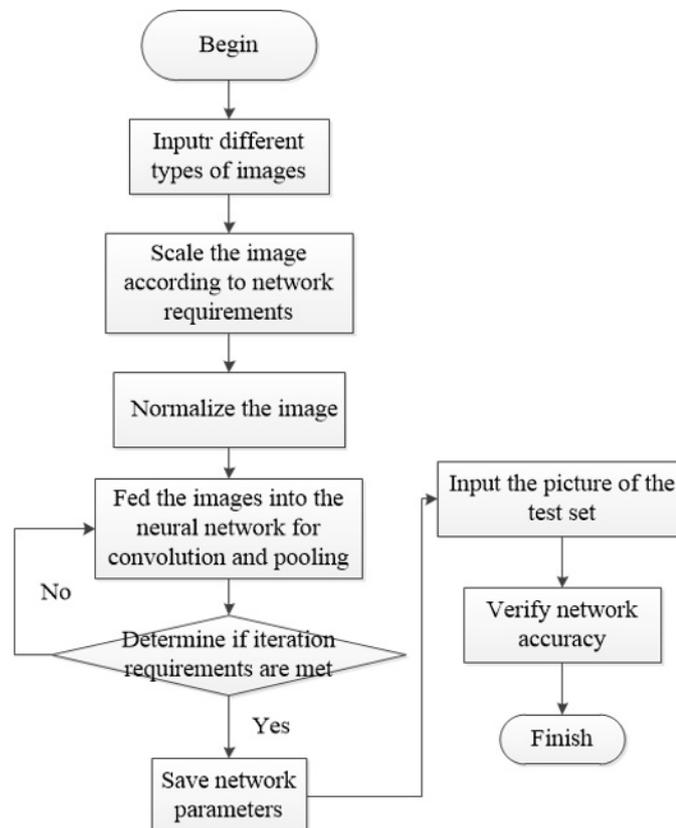


Figure 3. Software flow pattern

4.2 field picture area identification

With a trained network, the images captured by the drone can be identified. The steps are as follows: first, the whole picture is segmented; Identify each image using the previously saved network; Reassemble the whole

picture. The effect of the implementation is shown in Figure 4.

It can be seen from the figure that the large size pictures taken by the drone were divided into small pieces, and the color in each square was identified according to the neural network, thus determining the specific characteristics of each area in the rice field.

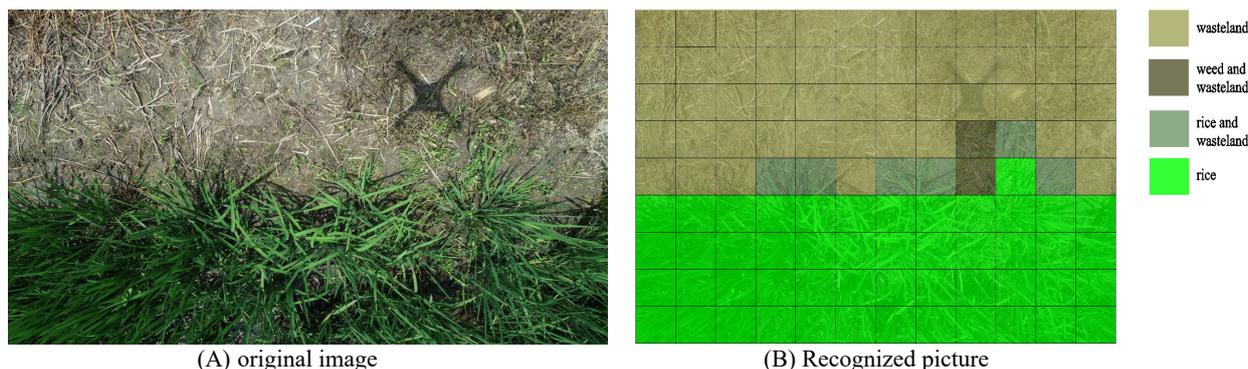


Figure 4. The convolution network identifies the rice field area

5 CONCLUSION

With the progress and development of science and technology, unmanned aerial vehicles have gradually entered people's field of vision. To use unmanned aerial vehicles for precision agricultural work, it is necessary to have precise positioning of crops. Unmanned aerial vehicles have a wide range of shooting range and a fast flight speed. Traditional k-means clustering technology and vegetation normalization index all require a lot of time for calculation and processing, which cannot play a role in the guidance of uav, and the precision is not high, so only the estimation of rice area can be done.

In this paper, for the rice field pictures taken by the drone, the segmentation is carried out first, and the images are classified by visual method, and the training set for the rice field area is made. The convolution neural network is used to identify the characteristics of the rice field. The comprehensive situation of recognition is better than the traditional detection method. Combined with the autonomous navigation technology of uav [14], the fixed-point quantitative pesticide spraying of uav can be realized. Therefore, uav precision agriculture can be achieved.

Reference

1. Liu baohua, Chen xiping, cui zhenling, meng qingfeng, zhao Ming. Research progress on yield potential and yield difference of three grain crops [J]. Chinese Journal of Eco-Agriculture, **23** (05):525-534 (2015)
2. Ning guoyun, xu qinfang, bai chao. Experiments on the application and control of rice diseases and insect pests by uavs for plant protection [J]. ACTA agriculturae zhejiangensis, **59** (05):765-766 (2008)
3. Qi wenzhe, Wang Feifei, Meng Zhen, Zhang Dianli, Wang Hongyan, Qiao Kang, Ji Xiaoxue. Eppo uav application status in China [J]. agrochemicals , **57** (4) : 247-254 (2018)
4. Xun dong, zhang jing, he keja, Yang haona, liu Yang, cao jia, zhou qing. Research on the effects of application of th80-1 plant protection uav on the control of major diseases and pests of rice [J]. Hunan Agricultural Sciences, (08): 39-42 (2015)
5. Xiao xiaohua, liu chun, Yang changhong, wu honghua, liu xianggui, wu jinzhong. Effect analysis of uav control on rice diseases and pests [J]. South China Agriculture, **10**(07): 5-8(2016)
6. Xue xinyu, tu kang, LAN yubin, qin weicai, zhang ling. Effects of high concentration application of uav on rice quality [J]. Transactions of The Chinese Society of Agricultural Machinery, **44**(12): 94-98+79 (2013)
7. Li Ming, huang yue qi, li xumeng, peng dongxing, xie jingxin. Extraction of rice planting information based on remote sensing image of uav [J]. Transactions of the Chinese Society of Agricultural engineering, **34**(04): 108-114(2014)
8. Tadasi CHOSA, Kiyoshi MIYAGAWA, Shigeto TAMURA, Kengo YAMAZAKI, Shinichi IYOSHI, Masami FURUHATA, Kota MOTOBAYASHI. Monitoring Rice Growth over a Production Region Using an Unmanned Aerial Vehicle: Preliminary Trial for Establishing a Regional Rice Strain [J]. IFAC Proceedings Volumes, **43**(26)(2010)
9. Cesar Analide, Pankoo Kim, Md Nasim Reza, In Seop Na, Sun Wook Baek, Kyeong-Hwan Lee. Automatic Rice Yield Estimation Using Image Processing Technique [M] (2017)
10. Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. Science, **313** (5786):504-507 (2006)
11. Chang liang, deng xiaoming, zhou mingquan, wu zhongke, yuan ye, Yang shuo, wang hongan. Convolution neural network in image understanding [J]. ACTA automatic sinica, **42**(09):1300-1312(2016)
12. Li yandong, Hao zongbo, Lei hang. Summary of research on convolutional neural network [J]. Journal of Computer Applications, **36**(09):2508-2515+2565 (2016)
13. Han xiaodong, Wang haosen, Wang shuo, Wang jianwen, Wang jie. Application of Python in image processing [J]. Beijing Surveying and Mapping, **32**(03):312-317 (2008)
14. Wang yuxi. Study on autonomous navigation method of agricultural uav based on beidou [D]. Northwest A & F University (2017)