Potato Plant Image Detection Based on Deep Learning

Qiuyu Xia[1, a], Jingwen Xu[1,b*], Junfang Zhao[2], Ning Li[1] and Juncheng Wu[1]

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

[2] Chinese Academy of Meteorological Sciences, CMA, Beijing 100081, P.R. China

[b] Corresponding author: x.j.w@163.com

Abstract. Potato is one of the most important food crops in the world. The information which extraction from high resolution remote sensing image is a new way to study the potato planting distribution and growth condition. For remote sensing target detection, a lot of people were used AdaBoost algorithm, SIFT algorithm, Tamura texture feature algorithm in the past. But it's just a feature of artificial extraction. Deep learning provides an effective framework for automatic extraction of target features. The experiment uses a simple but useful deep learning method (PCANet). After image segmentation, gray, binaryzation and filtering, the 42*48 of the potato plant images are trained and tested by feature extraction. The results showed that the detection rate of potato plants could reach 82.20%, the false detection rate was 12.66%, and the detection speed is 1.22-1.31 image per second, which could be applied to high efficiency fertilization, weeding and insect pests in order to achieve the purpose of increasing potato yield.

Introduction

The potato is the fourth most important food crop in the world. Comparing with the other three main crops, potato provide not only the storage time longer, but also wide adaptability, cold, drought and barren. Many countries have already chosen to put the potato in the strategic grain reserves. However, potato production is insufficient. Therefore, the cultivation and management of potato need more technical input and support in order to realize the mechanization, automation and intelligence of potato farming.

First of all, we need to accurately identify the potato plant in the field. It is a new way to extract characteristic information from remote sensing image, and the accuracy of remote sensing image acquired by aviation aircraft and satellite has been unable to meet the needs of people. Low altitude UAV remote sensing technology to obtain high resolution remote sensing data as the goal, inherited the high-tech unmanned aircraft, remote sensing and GPS navigation, establish the remote sensing system of a special high mobility, low cost and miniaturization, integration [1]. It has advantages of real-time, high flexibility and convenience, and it not only can compensate for satellite remote sensing by cloud shield to obtain images of the shortcomings, but also to solve the traditional satellite remote sensing revisit period is too long and emergency is not timely, and it will bring more convenience for feature extraction[2].

Secondly, an excellent feature extraction algorithm can improve the efficiency of the work. The traditional feature extraction algorithms such as the 20th century in the 1980s, the United States scholar Rumelhart, Hinton and Williams proposed for the back propagation algorithm of artificial neural network (back propagation algorithm, BP algorithm) [3], compared with manual rules in the past based system, which based on statistical machine learning methods in many ways showed superiority [4]. But BP algorithm is easy to fall into local optimum when the number of layers in the neural network is increased, and it is easy to be over fitting. In the 1990s, a variety of superficial machine learning models have been put forward, the support vector machine (Support Vector Machines, SVM) is also produced in this period, and the theoretical analysis and application [5] are very successful. In 2006, Canada University of Toronto Professor Hinton and his students Salakhutdinov in "science" published the article put forward network depth and depth of learning concept [6], opened a research upsurge of the depth of learning. In contrast, deep learning algorithms have more optimization for feature extraction methods.

We have done a lot of for training, testing and analysis of the detection rate, and increasing the number and found training set the number of hidden layer nodes good can effectively improve the detection rate. When the detection rate reaches a certain precision, the corresponding recognition program can be developed. With device design, the UAV was improved to make precise fertilization and pesticide application to become a reality, not only can make chemical fertilizers and pesticides in agricultural production waste reduction, and can reduce the pesticide and Fertilizer on the environmental pollution and destruction, more to promote the development of precision agriculture.

Methods

Deep learning is a new field of artificial intelligence and machine learning, its purpose lies in the depth of the neural network structure deeper simulated human brain deep structure and ways of thinking and supplemented by some learning algorithm, so that the computer's behavior as close as possible to the human behavior [7], and its essence is to construct multi hidden layer of machine learning model and a large amount of training data, where machine learning to more useful features and methods, so as to enhance the accuracy of classification and prediction [8].

Depth study of the core ideas:(1) unsupervised learning is used to training each layer; (2) each unsupervised training only one layer, learning each layer of the said as the input to the next layer; (3) using supervised learning to adjust all layers, construct multi hidden layer framework, using supervised and unsupervised reversal training way rising of sample utilization degree, and ultimately achieve the general shallow layer neural network cannot achieve the accuracy.

This experiment is applied to a simplified deep learning network for image classification. The training of image feature extraction contains the following: the principal component analysis, binary hash and block histogram [9]. The processing steps are as follows:

First of all, N piece of pictures to the average block to the average:

$$X = [\overline{X_1}, \overline{X_2}, \ldots, \overline{X_N}] \in R^{k_1 k_2 \times Nc}. \tag{1}$$

Get the matrix $\overline{X_i}$, C for the sub block matrix X_i number of columns;

Then the feature vector for XX^T, the feature vectors corresponding to the L_1 largest eigenvalues are taken as the next phase of the filter:

$$W_l^1 = q_1(XX^T) \in R^{k_1 k_2}, l = 1,2 \ldots, L_1. \tag{2}$$

The second stage and the first stage is basically the same, the difference is to filter the input image and the first step to do the second phase of the training images obtained by convolution. Get by N I_i and a picture of the L_1 filter convolution block results obtained together:

$$Y^l = [\overline{Y_1^l}, \overline{Y_2^l}, \ldots, \overline{Y_N^l}] \in R^{k_1 k_2 \times Nc}. \tag{3}$$

All the filter outputs are combined together:

$$Y = [Y^1, Y^2, \ldots, Y^{L_4}] \in R^{k_1 k_2 \times L_1 Nc}. \tag{4}$$

YY^T Eigenvector, take the largest L_2 eigenvalues and the corresponding eigenvectors as filter:

$$W_\ell^2 = q_\ell(YY^T) \in R^{k_1 k_2}, \ell = 1,2 \ldots, L_2. \tag{5}$$

Using binary hash to index and using block histogram as the combination to get the feature vector of each graph:

$$f_i = [\text{Bhist}(T_i^1), \ldots, \text{Bhist}(T_i^{L_1})]^T \in R^{(2^{L_2})L_1 B}. \tag{6}$$

The depth of learning algorithm and the traditional shallow learning different lies in:

1) Emphasizes the structure model of depth, usually have five, six layers, even ten layers of nodes in the hidden layer;

2) Clearly highlight the importance of learning characteristics, that is, through the layer by layer feature transform, the samples in the original feature representation is transformed into a new feature space to make it easier to classify or predict [10].

Results and analysis

In this paper, the identification method of potato plants is realized by feature extraction and classification. Application of high resolution computer collection and network to collect the image, through gray, compression, denoising preprocessing, finally comes to the characteristics of each training image, the characterization of each picture is 1 x n dimensional vector, then use these feature vectors to train support vector machine (SVM), is used to classify the image. The sample image data will be ready, put in the specified file directory, cutting, rotation angles and other processing, to ensure a sufficient number of samples after multiple training and testing, the results are shown in Table 1 and Table 2:

Table 1 Training time of each layer

Test item	First filter output time [s]	Second filter output time [s]	Hash histogram output time [s]	SVM classification time [s]	Total training time [s]
1	2.10367	15.4385	18.3838	1.6538	38.3347
2	4.18501	31.1080	34.4383	5.0561	74.4183
3	7.79348	61.2312	69.2103	12.0865	147.1960
4	1.94518	15.5334	17.3580	1.5987	37.1348
5	3.89381	30.5216	35.6634	5.0906	74.6340
6	7.82851	62.1656	68.9018	11.9550	147.7520

Table 2 Detection rate

Test	Training quantity	Test quantity	Detection rate
1	50	25	0.840
	50	50	0.810
	50	100	0.850
2	100	50	0.720
	100	100	0.810
	100	150	0.853
3	200	150	0.793
	200	200	0.845
	200	250	0.876
4	50	25	0.160
	50	50	0.160
	50	100	0.270
5	100	50	0.100
	100	150	0.120
	100	200	0.093
6	200	150	0.060
	200	200	0.080
	200	250	0.096

Detection rate: the percentage of the total number of positive samples. The higher the detection rate, the stronger the ability of detection. The program can accurately identify the leaf of the potato automatically, and it can satisfy the purpose of our experiment.

According to the time used for feature extraction, it can be calculated that 1.22-1.31 images can be processed per second, and the time is not long, which can be applied to a large number of data training and testing, and estimated time required.

According to test 1, 2, 3 shows that the classifier has a certain learning ability, such as test 2, training 100 samples, respectively, 50, 100, 150 of the test, the detection rate increased with the increase in the number of tests; and according to the ratio of the number of training and test number, 2:1, 1:1, 1:2, the higher the ratio, the corresponding detection rate has decreased.

According to the test 1 and test 2, test 2 and test 3, the number of training respectively 50,100,200 and consistent test number, number of training but low detection rates, although the feature extraction and variable, not more accurate detection. In tests 4, 5, 6, when the negative samples were tested, the number of training was the same, the higher the number, the higher the false positive rate was. In the test 4 and test 5, test 5 and test 6, the training number is different, and the test number is the same, but the number of training is more and the false detection rate is lower.

Conclusions

This study analyzes the detection performance of deep learning classifier based on feature extraction algorithm of potato plants. We training and testing different numbers of positive and negative samples, and using the method of exchanging positive and negative samples to eliminate some deviation. We found that if a small number of samples for training and testing were used, it causes the detection rate to be unstable, sometimes high and sometimes low. If the sample is added, the situation will be stable. In order to better detection results, we need adjust the structure of network and set the proper number of hidden layer nodes through a large number of testing in the experimental process. In this way, it can be ensured that the identification procedure can be widely used.

The feature of deep learning is unknown and random, it is the most suitable for the characteristics of object recognition, but internal operations are not visible. If we can extract the best feature of potato image, the recognition ability of the target will have greatly improved. On the other hand, we could also compare the algorithm of multiple features, then combined with several algorithms to improve the detection accuracy.

Acknowledgements

References

[1] Houyi D L: Development of UAV and UAV technology [J]. World Science and Technology Research and Development, 1998, 20(6):113-116.

[2] Wang B, Wang W: Application of UAV Remote Sensing in the Now War Military Operations [J]. Wireless Internet Technology, 2015.

[3] Rumelhart D, Hinton G, and Williams R: Learning Representations by Back-Propagating Errors [J]. Nature, 1986, 323(6088):533-536.

[4] Yu Kai, Jia Lei, Chen Yuqiang: Depth Study of Yesterday, Today and Tomorrow [J]. Computer research and development, 2013, 50(9):1799-1804.

[5] Guo Lili, Ding Shifei: Research Progress of Deep Learning. Computer science, 2015, 5:28-33.

[6] Hintio G, Salakhutdinov R: Reducing the Dimensionality of Data with Neural Networks [J].Science, 2006, 313(5786):504-507.

[7] Arel I, Rose D C, Karnowski T P: Deep Machine Learning - A New Frontier in Artificial Intelligence Research [Research Frontier] [J]. IEEE Computational Intelligence Magazine, 2010, 5(4):13-18.

[8] Collobert, Ronan, Weston, et al: A Unified Architecture for Natural Language Processing [J]. Journal of Parallel & Distributed Computing, 2008.

[9] Chan T H, Jia K, Gao S, et al. PCANet: A simple Deep Learning Baseline for Image Classification [J]. arXiv preprint arXiv: 1404.3606, 2014.

[10] Chen Xianchang: Deep learning Algorithm and Application Research Based on Convolutional Neural Network [D]. Zhejiang Gongshang University, 2014.