

Weed Identification in Crops Based on Convolutional Neural Networks

Shaobo Liang*

Westtown School, Pennsylvania, USA

hzhc_wjdi@163.com

*Corresponding author

Keywords: weed recognition, Weed image, model

Abstract: This study aims to develop an algorithm of transfer learning based on the Convolutional Neural Network Inceptionv3 for weed recognition by using images of weeds. Inceptionv3 shows great aptitude of image sorting as the basis of the network; classifiers are added onto Inceptionv3, weight initialization of transfer learning is implemented, and the model is trained through 1000 iterations with images downloaded from the open sources data on Kaggle. After the training, the model performed phenomenally, with a over 99% accuracy in the training set and a almost 90% accuracy in the validation set. The model also shows that it is well functioning in the actual validation, correctly recognizing the species of the weeds. This paper provides a possible solution to the need of weed recognition in precision agriculture.

1. Introduction

Removing weed in agricultural fields is the key for the success in production rate of the field, for weeds compete with the desired crops on nutrients, water, and energy; and weeding must be done in the early stages of crop growth. Currently, the primitive weeding method, manual weeding, has mostly fell out of use, with its place taken by other more efficient methods such as spraying herbicide on large pieces of farmland. One of the greatest disadvantages of herbicide spraying is the high consumption of the herbicide which leads to a great amount of chemical production power used. Another significant disadvantage of this method is the inevitable damage of herbicide on non-invasive organisms, desired crops, and the health of the humans using and consuming the crops. The reduction of herbicide usage has become one of the major focuses of technological and industrial giants and countries worldwide. Many intellectual mechanisms are being developed for this purpose, and one major problem faced is how to discern weed from crop real-time and in-time. Thus, an accurate and swift method of weed recognition has a significant application potential.

And thus, the goal for this study is to improve the ability of weed recognition by using advanced convolutional neural network to address the global need in this area.

2. Research status

2.1 Introduction to Neural Network

In the biological neural network of an organism, neurons are interlinked, and when a neuron is excited it will send chemical substances to the neuron connected to it, changing the electric potential of this new neuron; if the electric potential of this neuron is above a threshold value, it will be excited, and thus send chemical substances to the new neuron. The abstract version of this model is the neural network that we are talking about:

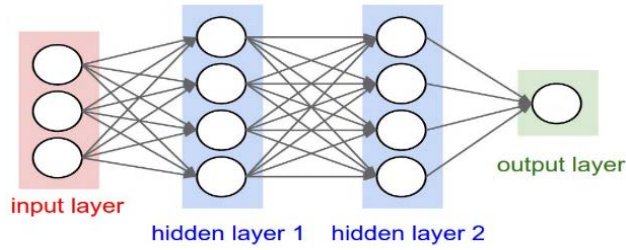


Figure 1. Neural Network Model

2.2 Convolutional neural networks

Convolutional Neural Network (CNN) is a structure of artificial neural network proposed by LeCun in 1989. Its specialty is the high efficiency of computing data with a grid like structure, in places like a set of data distributed on an axis representing time with a pattern being turned into a 2-D data represented by pixels. The difference between CNN and normal artificial neural network is that the feature abstraction in CNN involves alternating convolution layers and pooling layers. The layers in CNN usually consists of multiple characteristic planes constructed by neurons arranged in a matrix, and the same characteristic planes share the same convolution kernel for the sharing of parameters.

(1) Convolution Layer: Convolution is a mathematical algorithm with two parameters, and is the key part of CNN. Its mathematical model is shown in the following formula:

$$x_j^\ell = f\left(\sum_{i \in M_j} x_i^{\ell-1} * w_{ij}^\ell + b_j^\ell\right) \quad (1)$$

In the above formula, x_j^ℓ is the feature map of the j convolution in the ℓ layer, M_j is the set of the feature map $x_i^{\ell-1}$, w_{ij}^ℓ is the operation matrix of the feature map for the feature map, b_j^ℓ is the mapping offset of x_j^ℓ , and the function f is Activate the function.

(2) Pooling Layer: pooling layers are network structures that follows convolution layers, and they are used to compress the characteristic images in order to reduce the difficulty of the computation and to abstract the important characteristics. Common pooling methods are average pooling and max pooling. The mathematical model for average pooling is shown in the following formula:

$$x_{i,j}^\ell = \frac{1}{S_{R(i,j)}} \sum_{i,j \in R(i,j)} (x_{i,j}^{\ell-1}) \quad (2)$$

In this formula, S stands for the area. For example, if the sum of every 2×2 box and then divide the value by four, the main characteristic of the area is obtained. Usually the size of the kernels are 2×2 or 3×3 with stride equals to 2. The image will be compressed to $1/4$ of its original size after this pooling.

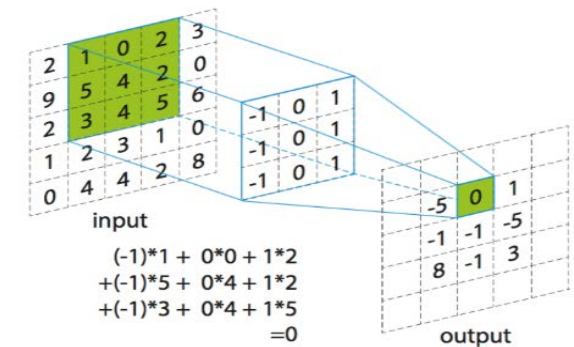


Figure 2. Convolutional neural network structure

3. Experimental design

3.1 Pre-processing of Data

The data being used is a set of 3500 images of 12 categories (species of plants) from the website of the Kaggle competition, in which a randomly picked 80% of the images are used as the training set and the rest 20% are used as the validation set. And so, the training and validation sets would have similar patterns statistically, which insures the accuracy of our model.



Figure 3. Sample Images of Weeds

3.2 Procedure of the Experiment

The experiment is started after the organization of the data. First, hyper-parameters of the model are set: the epochs are set as 300, 500, and 1000 to test difference in stability of the model with different number of epochs. The batch size is set as 20, and the number of training is set by dividing the total number of the samples over the batch size (20). After every iteration is completed, the weights and the bias parameter are reset. The number of validations made using the validation can also be calculated by dividing the total number of samples by the batch size (20).

At the same time, the loss function is calculated with rmsProp optimizer in backpropagation, and the learning rate is set as 0.001, while the learning decay rate is set as 0.0001. A well designed learning rate will lead to the loss between the anticipated value and the output of the model be in a gradient descent, which not only ensures the possibility of obtaining a converging loss function in a reasonable about of time, but also can avoid the learning rate to be too large. Here, the method of implementation of a learning decay rate is used to achieve this goal, and indeed this made the model avoid the greatest loss that were anticipated, and realized the reasonable update of the weights and bias in backpropagation. The effect of rmsProp is demonstrated by the figure below:

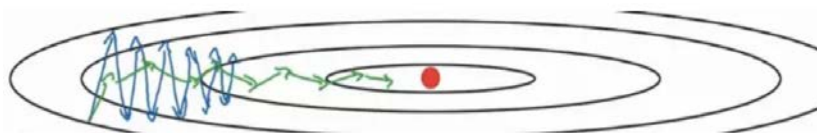


Figure 4. RMSProp Optimizer Formula and Effect

We can observe from the experiment that every iteration takes about 70 seconds, and the training of each batch takes about 700-800ms. The loss value becomes less as more iterations are gone through, while the acc (accuracy) gets gradually higher. The loss and acc shows that the structure of the model and the hyper-parameters are well set.

4. Results and Analysis

This experiment is done on a windows 10 platform with 32GB Ram, Nvidia GTX 1080Ti GPU. Through multiple trials of adjusting parameters, after 1000 times of training, the model have the ability of recognizing different types of weeds. The accuracy and loss of the training set and the validation set through the training is shown in the following figures:

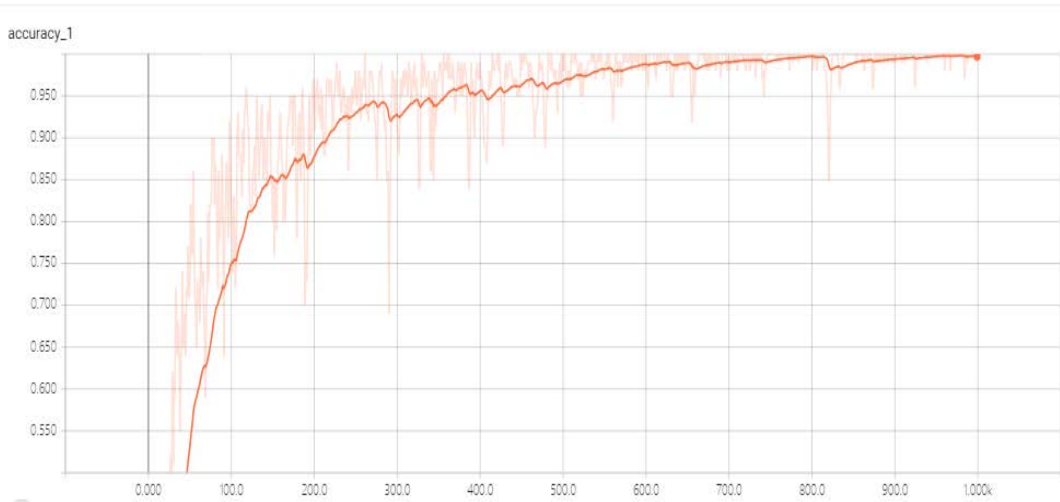


Figure 5. Accuracy in the Training Set

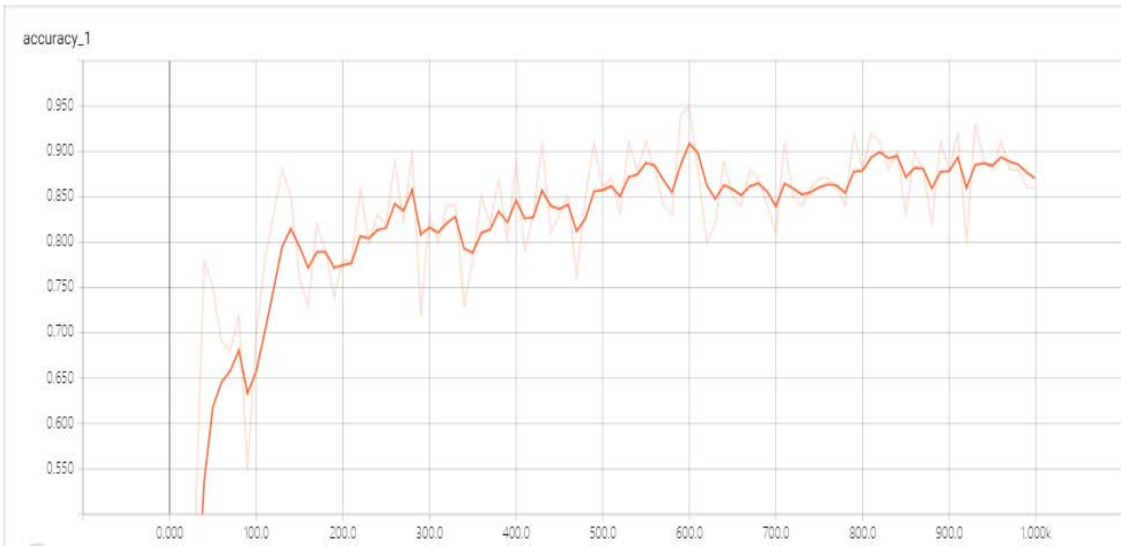


Figure 6. Accuracy in the Validation Set

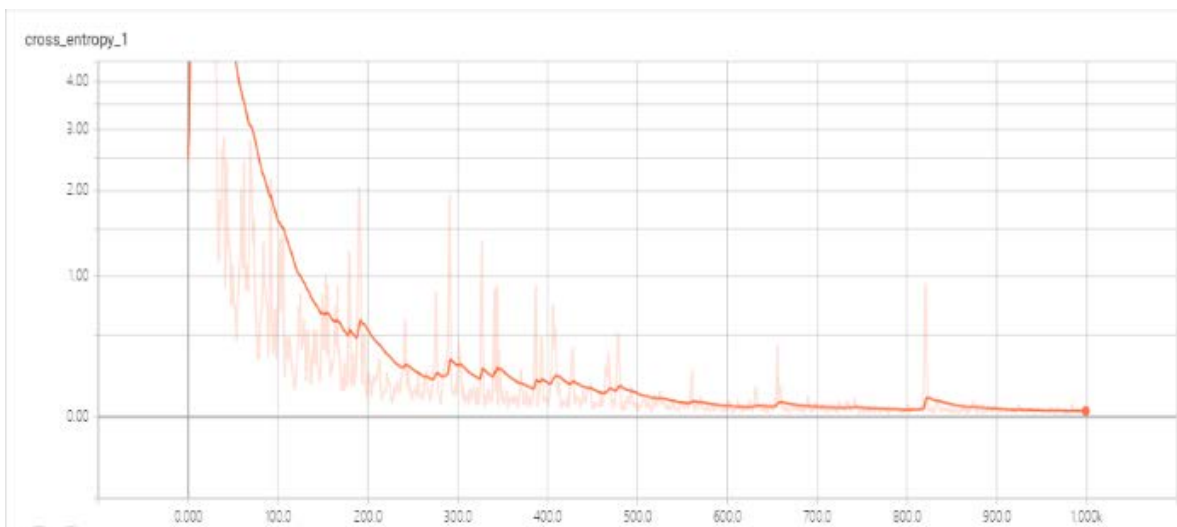


Figure 7. Loss in the Training Set

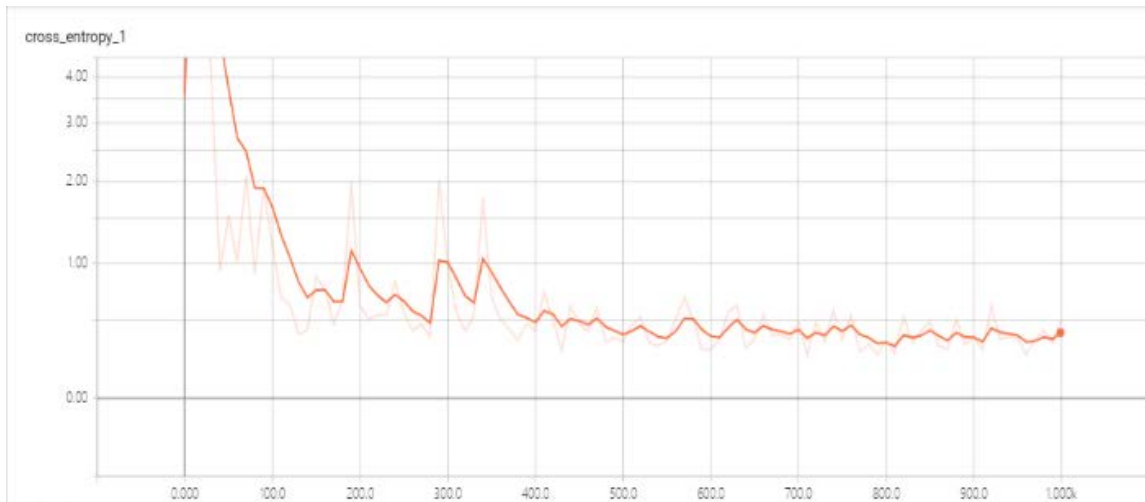


Figure 8. Loss in the Validation Set

Figure 4.1 and 4.2 shows the accuracy of the training and validation sets throughout the training. The horizontal axis represents the iterations of training, from 0 to 1000; the vertical axis represents accuracy, from 0 to 1 represents from 0% to 100%. A higher accuracy demonstrates a better performance. Figure 4.3 and 4.4 shows the loss of the training and validation sets throughout the training. The horizontal axis represents the iterations of training, from 0 to 1000; the vertical axis represents loss, and a value closer to 0 represents greater stability and thus a higher reliability.

From the above four figures, it can be observed that after 1000 iterations of training the accuracy is approaching 99% and the loss is approaching 0 in the training set; it can also be observed that after 1000 iterations of training the accuracy is approaching 90% and the loss is oscillating around 0.4 in the validation set. From the trend of the curves in the figures one can observe that the network is not overfitting or underfitting. With the trained model, the anticipated results of weed recognition is attained.

The actual result can be shown in the figures below:



Figure 9. Result of the Experiment

With the observations from the experiment, the model after 1000 iterations of training has great robustness. It can recognize the 12 types of well—in the trial none of the weeds were recognized incorrectly, and the recognition speed reaches 10 images/second, which would be sufficient for real agricultural application.

5. Conclusions and Future Prospects

This experiment used the open source data of weed images from Kaggle and Inceptionv3 transfer learning based on convolutional neural network, and the model shows an ability to accurately recognize the 12 different types of weeds which it was trained for, thus showing its great application potential and reached its anticipation.

However, due improvement must be realized if this model is to be applied in real life, for in the real weeding process there are more than 12 types of weeds. If more images of other types of weeds can be collected and if the model can be trained also with these images, the model would be more

apt for real application. In addition, neural network is only a mathematical model—an algorithm—and weeding must be done by hardware such as weeding drones or robots that can implement this algorithm. If these needs can be fulfilled, humans agriculture can be turned into a much more efficient, environmentally friendly, safe, and sustainable process.

Acknowledgments

In the end, I would like to thanks the teachers and peers who have supported me greatly through the process, especially in data collection and marking, of this experiment. Thanks should also be given to the data source of Kaggle for providing free and useful data which has been essential to this research.

References

- [1] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [C]. Advances in neural information processing systems. 2012: 1097-1105.
- [2] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 1-9.
- [3] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv:1409.1556, 2014.
- [4] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.
- [5] Sze V, Chen Y H, Yang T J, et al. Efficient processing of deep neural networks: A tutorial and survey [J]. arXiv preprint arXiv: 1703.09039, 2017.
- [6] Duchi J, Hazan E, Singer Y. Adaptive subgradient methods for online learning and stochastic optimization [J]. Journal of Machine Learning Research, 2011, 12 (Jul): 2121-2159.
- [7] Zeiler M D. ADADELTA: an adaptive learning rate method [J]. arXiv preprint arXiv:1212.5701, 2012.
- [8] Kingma D, Ba J. Adam: A method for stochastic optimization [J]. arXiv preprint arXiv: 1412.6980, 2014.