CONVOLUTIONAL NEURAL NETWORKS FOR CLASSIFICATION OF SOYBEAN CROP AND WEED

Prathiba M[1*], Dr. Vathsala M K[2]
[1]Assistant Professor, Dept of CSE, Cambridge Institute of Technology, Bengaluru.
[2]Assistant Professor, Dept of ISE, Cambridge Institute of Technology, Bengaluru.

ABSTRACT

A Deep Learning algorithm has many applications, image classification is one among them. In this paper it is used in classification of soybean crop and different types of weed associated with soybean bean crop. Weed is an unwanted plant which grows by itself between the crops. These weeds will suck the moisture and nutrients from the soil, by this the crops will get less nutrients and there will be a reduction in yield produced. Weed are the main reason for poor crop harvest. To increase the yield, weeds must be removed as early as possible. Machine Learning and Deep Learning algorithms play a very important role exhibiting best performance in detection and classification of weeds associated with soybean crop. Datasets of Soybean crop and the weed specific to this crop are used for analysis. In this paper, classification of soybean crop and different kinds of weed is performed. Algorithms used is Convolutional neural network. The performance is observed by the factor accuracy. 98% of accuracy is achieved in detection of weed using CNN.

Keywords: Classification accuracy, Convolutional Neural Networks, Image processing, precision agriculture.

1. INTRODUCTION

If the population increase in the same rate as it is increasing now, by 2050 the production must be doubled to fulfill the food needs [1]. Agriculture is one field which plays a very important role in every country. Due to urbanization many farmers are moving towards urban places because of economic losses occurring in farming, therefore there is a reduction in number of farmers and hence production is reduced. Weed is one more factor which affects the rate of production. Farmers usually spray pesticides to reduce weed. The usage of pesticides will affect the yield, soil quality and also causes many health issues by using such yields and also inhaling the pesticides while spraying. Hence an efficient and optimal method is needed for weed control which is achieved via convolutional neural networks. This is happening because agriculture is done without any analysis and the traditional agricultural methods used have many loopholes in it. Machine Learning and Deep Learning has solutions to almost all these loopholes. In this paper we are mainly concentrating on weed detection, so that it can identified as soon as possible, and proper measures can be taken to remove it and to increase the production and profit. In this paper, we are going to classify weed and soybean crop using one of the efficient deep learning algorithms i.e. CNN.

The research and development in the last two decades have led to continuous extraordinary improvement in the development of intelligent system to precision agriculture. One of the key technological development requirements for progress in such areas is prediction, detection, classification using Machine Learning, Deep Learning, and Artificial intelligence. In fact, the ability to generate a performance greater than 95% is still one of the biggest challenges.

In this paper we are doing image processing on soybean and weed dataset. we are using SLIC algorithm for segmentation i.e. extracting ROI (region of interest). After segmentation we are using CNN for feature extraction and classification.

2. LITERATURE REVIEW

Development of first weed detection algorithm was done in the year 1996 [2]; this algorithm by using infrared (IR) images was able to segment crops from weeds.

Araguez et. al [3] after the analysis of the green channel histogram of plant in the image dataset, implemented a segmentation algorithm and performed the weed and crop segmentation. Hong and Lei [4] by usage of artificial neural network developed an efficient model for detection of weeds associated with maize plants in various luminosity settings with an accuracy of 92.5%. OTSU and Watershed binarization methods were used for segmentation [5]. To perform classification thresholding is done through an area analysis and this model was efficient only when crops amount is less than weed amount i.e. a farm having more weeds than crops. Green colour of the crop is to be segmented to apply fuzzy clustering and this was experimented by Romeo et. al [6]. UAV started using multispectral camera, which outputted an RGB and NIR images and these images were segmented and classified using CNN which gave an accuracy of 98%, [7].

3. PROBLEM STATEMENT

Usually, weed has a major impact on yield of the crop. Broadleaf and grass are an unwanted plant which grows by itself and compete with soybean crop to utilize all the resources. These weeds as to be removed as early as possible, this can be done using AI model like CNN. In this paper using four types of images i.e. soybean, broadleaf, grass and soil we are developing a model which can classify soybean crop and weed. Using the concepts of transfer learning, image classification and deep learning this model is developed. Classification of images is basically dependent on features. For clustering SIFT descriptors are used, [8]. For feature extraction Spatial pooling, [9], Histogram encoding, [10] and recent Fisher Vector encoding, [11] are used. Even though the results of these models have higher accuracy, this can't be considered as optimal one. Since amount of time and effort required is high and domain experts are needed. Deep learning eliminates the traditional feature extraction step, which is tedious and time consuming. This is optimized from automating the learning process from raw data using graphics processing units (GPUs), [12].

4. ALGORITHMS

4.1. CNN

4.1.1. Model Architecture

One of the efficient Deep Learning model for image classification is Convolutional Networks. It is a type of Eager learner algorithm. CNN has a very good performance in classification. CNN is the most efficient algorithm currently used for image classification.
Their architectures are inspired from the mechanism of human body. Features are learned from the raw data of input images automatically. They initially identify small edges and are learnt, then parts of the leaf, then the full plant in the image is being learnt.

Using different layers of network different features are extracted. Each layer has definite number of neurons and is inputted as three dimensions i.e. height, width and number of channels. For grey scale image there will be two channels and for colored images there will be three channels RGB. This input will be in the form of matrix, for grayscale it is two-dimensional matrix and for colored image the input will be a three-dimensional matrix.

Convolutional neural network structure has four layers. The first layer of a convolutional neural network is the convolution layer, in this layer all the feature extraction takes place automatically from the input images. Feature extraction takes place by mapping the kernel on to the input image and multiplying the corresponding pixel values and taking the mean of it which reduces the size of the image without losing any information from the original image and the output will be a matrix called convolution matrix.

First layer is the Convolutive layer, which accepts an image of size [W1×H1×D1] where W1, H1, D1 represents the height, weight and number of channels respectively. The weights are multiplied with neurons which are near to it and resultant output will be in the form [W2×H2×D2] and this is called convolution matrix. Here W2, H2 and D2 represents width, height and number of channels respectively,[13]. Where

$$W2 = \frac{W1 - F + 2 * P}{S} + 1$$

$$H2 = \frac{H1 - F + 2 * P}{S} + 1$$

$$D2 = K$$

Where F is the spatial extend of the filter, K is the number of filters, P is the zero padding which defines the parameter controlling the output volume, S is the stride which defines the parameter with which we slide the filter by skipping few pixels.

Second layer is the RELU layer, MAX(0,X) activation function is being used. In this layer all the negative values are eliminated and is replaced with value zero. By this the features are not affected [14].

POOLING layer, this is the third layer in CNN. This is used between the Convolutive layers. In this layer reduction in number of parameters is done. Maximum value is taken from matrix of dimension width and height. This matrix is mapped on full input image. [W2×H2×D2] will be the output of this layer.

$$W2 = \frac{W1 - F}{S} + 1$$

$$H2 = \frac{H1 - F}{S} + 1$$

$$D2 = D1$$

At the end, the output of first layer is connected to dense fully connected layer of 512 neurons. Then again this fully connected layer is connected to one more dense layer with four neurons, which signifies four categories of input images i.e. soybean, broadleaf, grass, soil. The dense layer is nothing but multilayer perceptron, [15]. The code is executed after all layer execution is done; the output is generated from one neuron that is classification of input image to one category, the output will be of size [1×1×k].

4.1.2. CNN training

Creating CNN needs large amount of dataset which includes all possible cases which is needed for learning, creating a dataset is a tedious task, machine of higher processing capability is needed. Stable architecture is to be developed initially by selecting the right number of layers, which can do required Classification [16].

Then in training process we need to select proper loss function to reduce loss and increase the performance of the model. In the model which is being developed categorical cross entropy is selected to reduce loss.

Incase if there are limited images in the dataset or if complexity of the model is to be reduced we can adapt pre-trained networks available in keras and this process is called transfer learning[17], which has a pretrained information stored in a knowledge base which can reused while created a new model whose job is to perform similar task.

The pre-trained model is capable of extracting features from raw data in the image automatically and can be given as an input to the classifier to perform the required classification. This convolutive part is made frozen, to make this made ready to use for testing without again training it. Now the model can accept an image as input and feature vector is generated as output and this is given as input to the classifier, [18]. In our model 4 convolutional layers are used.

By using this model feature extraction is done automatically, this has eliminated the tedious and time-consuming task of traditional feature extraction. And there is also reduction in number of parameters, the one neuron is not connected to all neurons in the next layer, it only connects to the nearby neighbor neurons. Because of this the model can run on CPU with decent amount of time consumption, [20].

Fine tuning is a method of taking the advantage of knowledge base and using a pretrained model and retraining that model to fit to current classification task, and dataset. The only change to be done is in the last layer of the model, i.e. Fully connected layer. The initially freeze layers of the model can be used for training and by the addition of new last layer, we can adapt the model for any new classification problem. The final layer is for classifier, a preinitialized multilayer perceptron is used for this purpose.

To learn features for our soybean and weed dataset and to perform classification, we use Convolution neural network along with fine tuning technique. The pre trained network used is LeNet.

5. METHODOLOGY

We first create a stable model using a pretrained model from the knowledge base of the proposed model for classification of weed and soybean crop using the dataset using transfer learning.

5.1. Dataset

The dataset has been downloaded from Kaggle.com and it has more than 15,000 images which includes all category of images. Out of which 3000 images of all category is segregated for testing and 3000 images of all categories for validation. The categories are soybean, grass, soil, broadleaf. Training, testing, validation data are distributed similarly for all categories.

5.2. Convolutional Neural Network developed architecture

Our method is based on Deep learning concept, LeNet which is simple and earliest CNN. LeNet is trained by backpropagation algorithms. This was initially developed to recognize handwritten characters. This was later called as LeNet. This network contains four convolutional layers Conv1, Conv3, Conv5, Conv7, four pooling layers P2, P4, P6, P8; dropout layers, and two fully connected layers FC9, FC10. The input is given from the dataset and a classification output is expected which belongs to either of the four categories.

The image of [150×150×3] size is inputted. In the first Convolutional Layer of LeNet 32 convolution kernels with matrix size (3,3) and strider S=0 is applied. (3 channels*32 nodes*3pix*3pix)+32 bias= 896 parameters after first convolutional layer. In second convolutional layer 32 posterior nodes*64 nodes*3pix*3pix)+64 bias= 18496 parameters. In third convolution layer (64 nodes posterior*128 nodes*3pix*3pix)+128 bias= 73856 parameters. In fourth convolutional layer (128 nodes posterior*128 nodes*3pix*3pix)+128 bias= 147584 parameters. For the calculation of loss Categorical Cross entropy is being used. Since we have inputs which belongs to only one category. The loss as to be minimized to increase the performance. The architecture and layers in LeNet is showed in fig.1.

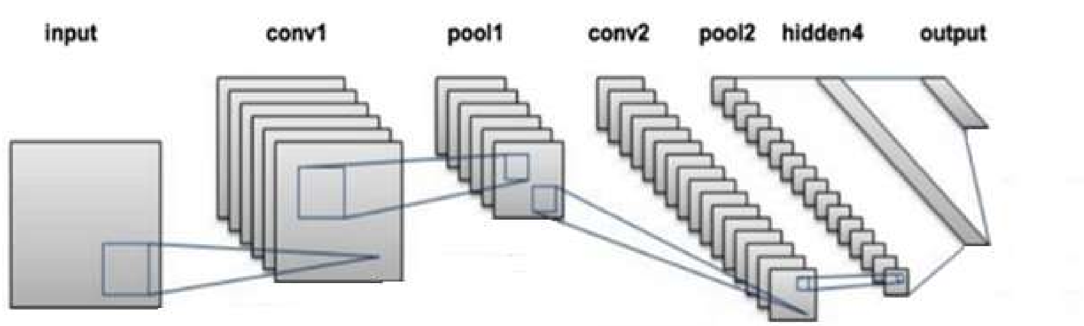


Figure 1. LeNet architecture

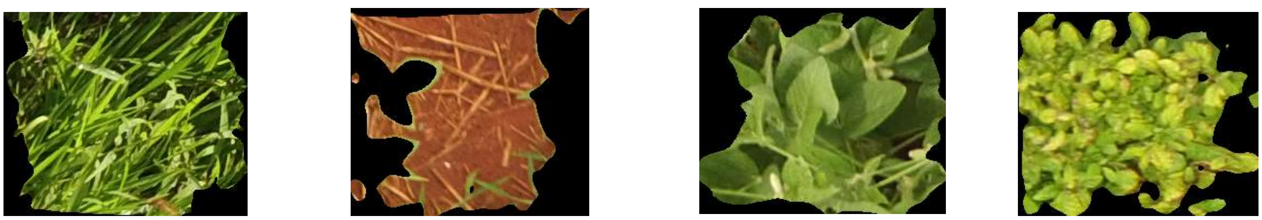
6. RESULTS

For the task of classification and evaluating results, we exploit LeNet to classify weed and soybean crop and by varying the epochs the experiment is conducted by 2, 5, 6, 8, 10 epochs and learning rate of 0.001 and a Batchsize of 31.

TABLE 1. Classification results.

Epochs	Time taken for training (s)	Accuracy (%)
2 Epochs	22767.75	0.6451
5 Epochs	26226.67	0.8491
6 Epochs	27266.94	0.8562
8 Epochs	29777.20	0.8642
10 Epochs	45556.30	0.9230

The model was run for many times with varying epochs and batch size it showed that as number of epochs increases, the accuracy also increased. As batch size decreased the accuracy has increased. Currently 92.30% of accuracy is achieved with 10 epochs and batchsize of 31.



Grass Soil Soybean Broadleaf

Figure 2. Classification based on category.

Many parameters affect the accuracy of the model like size of the dataset, i.e. number of images which we can get for training, number of convolutional layers where the feature extraction is done, Batch size defines how many iterations are there in each epoch. In this model as batch size decreases accuracy increases and the learning rate.

7. DISCUSSION

LeNet along with fine-tuning technique has reduced the complexity of designing and developing model from scratch. And facilitates on automatic feature extraction, which has reduced the time of feature extraction as compared to traditional method. The change which is to be adapted is inclusion of fully connected layer. The number of neurons in fully connected layer depends on interest of classification task. Classification of soybean and weed has four categories, hence the last fully connected layer should have four neurons. Initial all layers are freeze and output is obtained by training the fully connected layers. By doing this the model fits to classification problem and the efficiency of the algorithm is increased.

8. CONCLUSION

Deep neural networks are very efficient in image-based classification because of its automatic feature extraction capability from raw data, but with normal machines training the model requires more time. Here in this paper we have developed a model for classifying soybean and weed images. We are using convolutional neural networks for building a classifier, which is being trained on 9000 images from dataset downloaded from Kaggle.com. LeNet architecture is used for transfer learning for classifying soybean crop and weeds associated with it like grass and Broadleaf. And, fine tuning is used for reducing the number of parameters while training i.e. the neuron in one layer connects with only the nearest neurons in the next layer, not to all neurons in the next layer. By this number of weights are reduced and training time is reduced without reducing the precision or accuracy of the result. And, batchsize is very important. Larger the batchsize causes memory problems and training time is increased.

REFERENCES

[1] B. Cheng and E. T. Matson, “A Feature-Based Machine Learning Agent for Automatic Rice and Weed Discrimination,” Springer, Cham, pp.517–527, 2015.

[2] R. Brivot and J. Marchant, “Segmentation of plants and weeds for aprecision crop protection robot using infrared images,”IEE Proceedings-Vision, Image and Signal Processing, vol. 143, no. 2, pp. 118–124, 1996.

[3] S. Equipo T´ecnico, “Proceedings of vii seae congress spanish societyfor organic farming,” 2006.

[4] H. Y. Jeon, L. F. Tian, and H. Zhu, “Robust crop and weed segmentation under uncontrolled outdoor illumination,”Sensors, vol. 11, no. 6, pp.6270–6283, 2011.

[5] S. H. Hlaing and A. S. Khaing, “Weed and crop segmentation and classification using area thresholding,”IJRET, vol. 3, pp. 375–382, 2014

[6] J. Romeo, G. Pajares, M. Montalvo, J. Guerrero, M. Guijarro, andA. Ribeiro, “Crop row detection in maize fields inspired on the humanvisual perception,”The Scientific World Journal, vol. 2012, 2012.

[7] C. Potena, D. Nardi, and A. Pretto, “Fast and accurate crop and weedidentification with summarized train sets for precision agriculture,” inInternational Conference on Intelligent Autonomous Systems. Springer,2016, pp. 105–121.

[8] P.Sermanet, D.Eigen, X.Zhang, M.Mathieu, R.Fergus,and Y.LeCun. “Overfeat: Integrated recognition, localization and detection using convolutional networks.” arXiv:1312.6229, 2013.

[9] H. B. Burke, “Artificial neural networks for cancer research: Outcome prediction.” Sem. Surg. Oncol, vol. 10, pp. 73–79, 1994.

[10] H.B. Burke, P.H. Goodman, D.B. Rosen, D.E. Henson, J.N. Weinstein, F.E. Harrell, J.R. Marks, D.P. Winchester, D.G. Bostwick, “Artificial neural networks improve the accuracy of cancer survival prediction.” Cancer, vol. 79, pp. 857-862, 1997.

[11] J. Lampinen, S. Smolander, and M.Korhonen, “Wood surface inspection system based on generic visual features.” Industrial Applications of Neural Networks, F. F. Soulie and P. Gallinari, Eds, Singapore: World Scientific, pp. 35-42, 1998.

[12] T. Petsche, A. Marcantonio, C. Darken, S. J. Hanson, G. M. Huhn, I. Santoso, “An autoassociator for on-line motor monitoring.”, Industrial Applications of Neural Networks, F. F. Soulie and P. Gallinari, Eds, Singapore: World Scientific, pp. 91-97, 1998

[13] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel. “Backpropagation applied to handwritten zip code recognition.” Neural Computation, 1(4):541–551, 1989.

[14] Y.Boureau, F. Bach, Y. LeCun, andJ. Ponce. “Learning midlevel features for recognition.” CVPR, 2010.

[15] Marc Parizeau. “Le perceptron multicouche et son algorithme de rétropropagation de l’erreur.” Département de génie électrique et de génie informatique, Université Laval, 10 Septembre 2014.

[16] D. Lowe. “Distinctive image features from scale-invariant keypoints.” IJCV, 60(2):91–110,2004.

[17] Y. LeCun, L. Bottou, and J. HuangFu. “Learning methods for generic object recognition with invariance to pose and lighting.” CVPR, 2004

[18] F. Perronnin, J. S´anchez, and T. Mensink. “Improving the fisher kernel for large-scale image classification.” ECCV, 2010.

[19] P.Sermanet,D.Eigen,X.Zhang,M.Mathieu,R.Fergus,and Y.LeCun. “Overfeat: Integrated recognition, localizationand detection using convolutional networks.” arXiv:1312.6229, 2013.

[20] Schmidhuber, J. “Multi-column deep neural networks for image classification.” CVPR. 2012.