Corn Crop Disease Detection using Convolutional Neural Network(CNN)

Vikash Kumar Gupta

School of Computer Science & Engineering, Lovely Professional University, Phagwara ,India

vikash.11805741@lpu.in

Abstract-Plant diseases are very common and spreading nowadays. Various Machine Learning models are used for plant disease recognition. This paper mainly focuses on the new approach for plant disease detection model, based on leaf image processing using Convolutional Neural Network(CNN) as a class of machine learning model that can train to accurately recognize objects in images. Large, multi-scale, multilayer publicly available dataset in the form of images is used, help in training the deep convolution model, provide high accuracy. The developed model is able to detect different types of plant diseases out of healthy leaves. Implementing the model on Maize Crop for recognizing a common foliage disease Northern Corn Leaf Blight(NCLB). Caffe, a deep learning framework developed by Berkley Vision and Learning Centre, was used to perform the deep CNN training. Corn is an ideal model to deploy for Farmer to monitor and manage crops. The experimental results on the developed model achieved an average accuracy of 96.3%.

Keyword - Plant disease, Maize, Corn, Disease, Phytopathology, Machine learning, Convolutional neural network, Images, Deep learning

1. Introduction

The issue of plant disease protection is an enormous problem for farmers. Plant diseases are also somewhere related to the problems of sustainable agriculture and climate change[1]. The situation is further difficult by the fact that today, diseases are transferred globally more easily than ever before. A foliage disease of maize crop, Northern Corn Leaf Blight(NCLB) is very acute in the United States and Canada. Survey results mentioned that 3billions dollar of maize crop got affected in the year 2017 in the United States and Canada. Broadcasting in a large area for early symptoms will reduce the time and also help in finding the severity in NCLB[2]. Large rates of pathogen development; it can also modify host resistance, which leads to physiological changes in host-pathogen interactions [3]. Variety in acute NCLB has no local expertise to combat them [4]. Irrelevant pesticide usage can cause the development of long-term resistance of the pathogens, causing low tendency to fight back. Providing accuracy in the diagnosis of plant diseases is one of the major challenges of agriculture. It is critical to prevent unnecessary waste of financial and resources.

Little knowledge of pesticide usage can cause the evolution of long-term resistance of pathogens, reducing the ablity to fight back. Sometimes inaccurate pesticides may cause to kill the beneficial insects, helping in pollination like Butterflies and Bees. Timely and accurate diagnosis of plant diseases is one of the strengths of precise agriculture[5]. It is important to prevent unnecessary waste of financial and other resources, for achieving healthy production. Self, field-based recognition of plant disease symptoms would be valuable for the farmer. However, the difficulty will occur because of the “noisy” nature of field imagery. This requires a computer vision approach that is certain to the target disease and insensitive to such variations.

Deep Convolutional neural networks (CNNs) model uses a massive amount of data for training and testing of data, also help in determining accurately objects in images, making them principle for object detection [6]. Unlike detection of everyday objects, plant disease symptoms require expertise and experience in the specified field to identify.

Very few, expertise for image sets of plant disease exist[7]. PlantVillage is a repository of different plants images, contains around 50,000 images of numerous crops and diseases [8]. However, these were taken with detached leaves on a plain background, and CNN failed to train and did not perform well even on-field [9]. Existing image sets are much smaller for training the model . It is difficult to collect, select and present information by experts. Variations in symptoms by diseased plants may lead to an improper diagnosis also cause difficulties to farmers in determining the disease and the pesticide for curing them.

2.Literature Survey

Computer Vision and image processing enhance the practice of plant protection and various application in the field of agriculture. Utilization of image processing techniques such as color analysis and thresholding were used for recognition and classification of plant diseases. Different approaches are used at present for plant disease detection and most common are Artificial Neural Networks(ANNs) and Support Vector Machines(SVMs). For better feature extraction theses method are combine with image preprocessing.

[10]ANN is an information-processing system based on the function of biological nervous systems, such as the brain process information. The highly interconnected neuron of the Brain helps to solve complex problems. ANN has three-layer the first layer takes input as neurons, second layer work as a hidden layer and the third layer as output layer.

[11]ANNs for image detection goes through the process of Image Acquisition, Image Preprocessing, Image Enhancement, Image Segmentation and Feature Extraction. ANN tool also has a backpropagation process
which works on a formula,

$$w = w_{old} + k.\delta x$$

where $\delta = y_{target} - y$ and k is a control parameter which controls the learning rate.

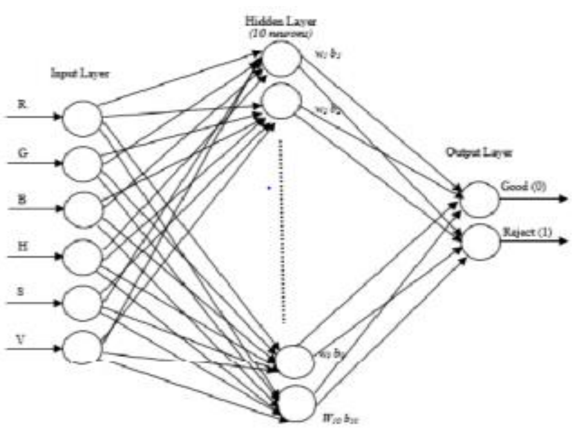


Fig. 1. Multilayer ANN with 3 layers[11]

[12] Using SVMs first read the input image and resize them, help in contrast enhancement RGB color space is converted into HSI (Hue Saturation Intensity) color space. In the segmentation step, k-means clustering operation is used to select the defected area, and it is extracted the features by using GLCM (Gray Level Co-occurrence Matrix) Compute contrast, correlation, energy, homogeneity, mean standard, entropy, root mean square. With features extraction, the median filter is used for getting noise-free feature results. Finally, the leaf disease is classified by using support vector machine (SVM) and computes the accuracy. From the obtained results, the maximum accuracy of the system is 83%.

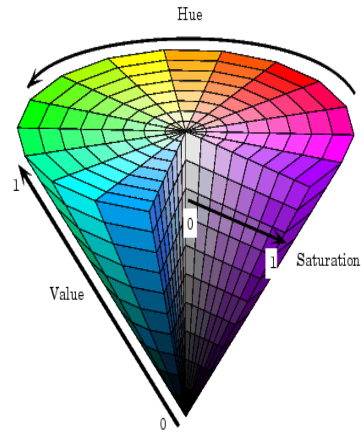


Figure 2. HSI Color Space[13]

Improvements in convolutional neural networks (CNNs) in recent years have made them the state of the art among machine learning approaches for addressing computer vision problems, particularly image classification. CNNs learn which features are most important. The linear transformations are learned during training by making small changes to the weight matrices that progressively make the transformations more helpful to the final classification task. [6]

Northern Leaf Blight (NLB) or turcium blight disease is the main focus for the paper. The first step is to distinguish NLB lesions from other damaged leaves in the field images. All visible lesions were marked using the line by the annotation feature of the Bisque image-processing platform as shown in figure 3. The infected image used to train the convolutional neural networks helps in detecting the presence of Northern Corn leaf blight(NCLB).In figure the leaf marked with A and B are the symptoms of NLB-infected maize plants. whereas, the Leaf marked with C and D are noninfected image of leaf plants. For this we collect a variety of NLB infection and other irrelevant features like soil and sky. For accuracy three-way image was taken helps to detect the presence of NLB lesions in the images.

Figure 3. NLB-infected maize plants marked with red circle[14]

The procedure for plant disease detection using Convolutional Neural Network is divided into number of steps as shown below.

Dataset- Collection of a large amount of dataset are required collecting from different sources for better training phase.Various healthy leaves images are also added in order to distinguish between disease and healthy leaves. After collection we have to filter the data by removing duplicate images.

Image processing and Labelling- For better feature extraction dataset are preprocessed. Preprocessing process include cropping if image for highlighting the region of intrest.Image having lower resolutin are not considerinig in dataset. The main focus is to take the image in dataset having high resoultion at diseased region. Images for training the dataset were resized to 256 × 256 for reducing the time which was automatically computed by python script,with the help of OpenCV framework[15]. Agricultural experts are required for analysing leaf images and labelled all the images with appropriate disease acronym. For training and validation dataset it is important to accurately classification of images. All visible lesions were marked using the line by the annotation feature of the Bisque image-processing platform.

Augmentation Process - Augmentation is required for reducing the overfitting during training stage. It also increases the dataset and introduce slight distortion in images. The image augmentation contained the transformation techniques including affine transformation, perspective transformation, and simple image rotations. Simple image rotation as well as rotation on different axis were applied for the Augmentation process.

Neural Network Training - For image classification there should be training of deep conolutional neural network. Caffe a Open source deep learning framework containing Caffenet model developed by BVLC[16]. This framework was used with the large dataset if images called Imagenet[17]. Caffenet is a deep CNN having multiple layer helping in extracting feature from the images. Network contains Eight layers in which five are convolutional and remaining are connected layers.The Convolutional Neural Network are build using Convolutional layer. As shown in figure 4, In stage 1, all the three convolutional neural networks(CNNs) are used to train the images. In stage 2, heat maps are created by using each of the three CNNs individually. In stage 3, CNNs trained model takes the heat maps as input and give whole image containing diseased tissue as output.

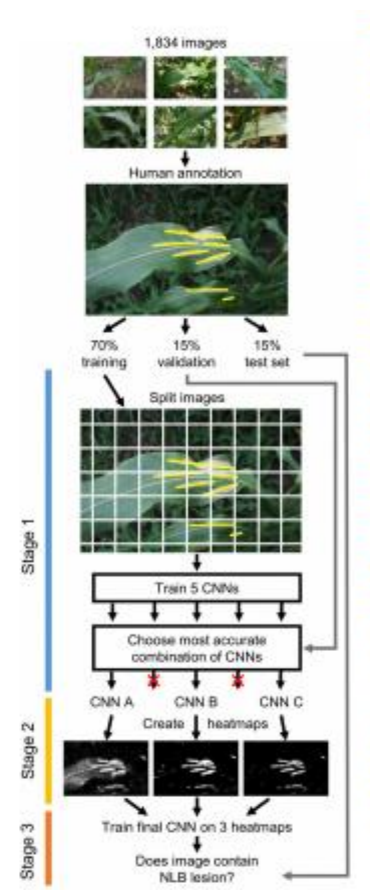


Figure 3. Three stages of the classification pipeline[14]

Performed Tests - Dataset is divided into training set and the testing set for measuring the performance. Training set is used for training Neural Network and testing set is used for prediction.hence,the accuracy can be calculated. Cross validation technique are used to evaluate the predictive model and the procedure was repeated at thousand iteration.

Fine-Tuninig - To increase the efficiency and effectiveness of a process we use Fine-Tuning. It also help in optimizing the outcome. Fine-tuned learning experiments require learning but it is much easier and faster than learning from scratch[18]. From scratch a new softmax classifier was trained using back-propagation algorithm with data from the dataset. This backpropagation algorithm ran for one-lakh iterations. This process was repeated by changing parameter of hidden layers and hyperparameter.

Equipment - A PC is required for overall process of testing and training of the model. Training of the model was performed in Graphics Processing Unit(GPU) mode. Each training iteration will take approximately eight hours.

After training the whole dataset of both original and augmented images.Using larger dataset convolutional networks able to learn features fastly. At last after 100th training, an overall accuracy of 96.3% was acheived.

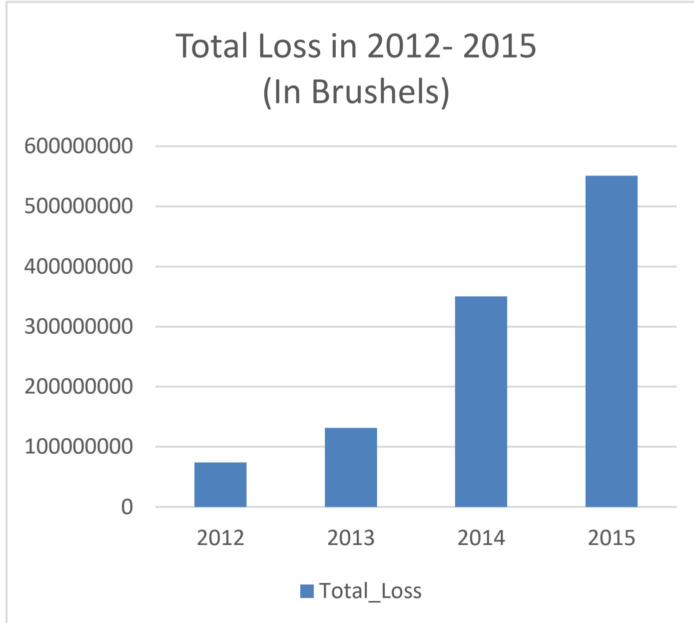
3. Data Description

We have gathered image data from various platforms with a distinct view. Dataset is collected in three way- Using hand-held cameras, with a camera mounted on a boom, and with a drone camera. Annotation in each image is done by the expertise for better classification results.[19] The three data sets comprise of 18,222 images annotated with 105,705 NLB lesions, making this the largest publicly available image set annotated for a single plant disease. All these images are taken from Musgrave Research Farm.[20]

Northern Leaf Blight (NLB) or turcium blight disease has been growing at severe rate causes estimated yield losses rising steadily from 1.9 million metric tons in 2012 to 14 million metric tons in 2015. The estimated mean economic loss was $76.51 USD per acre in United States and Ontario[21].

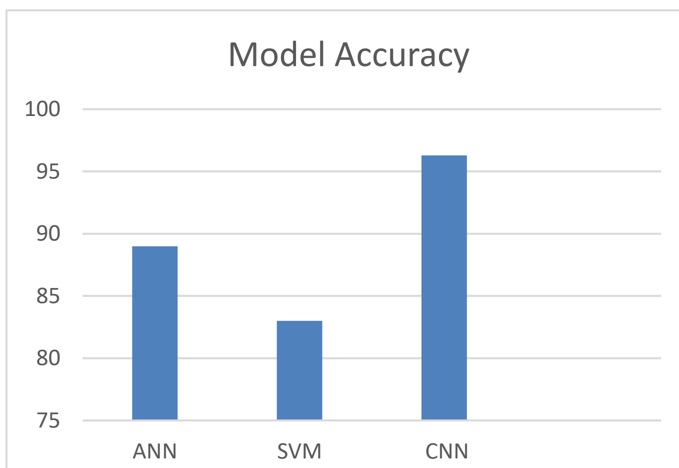
The large size of NLB lesions makes this disease an attractive type for image-based phenotyping. After the formation of lesion having typically the size of 1cm width and 5cm length. Possiblity of finding the large lesions is more than the smaller lesions or pustules. Total estimated corn loss in Brushels by Northern corn leaf blight or Setosphaeira turcica disease in 2012, 2013, 2014 and 2015 are 73993728, 131554916, 350068129 and 551054156 respectively causing the total loss of 1106670929(Brushels)

[14]. It is represented in graph below



4. Result and Discussion

The result of the three different method are SVM, ANN and CNN are 83%, 89% and 96.3% respectively. As we know that the convolution neural network were able to learn the feature on the larger datasets, results got explored with different types of images. After fine-tuning the accuracy of the CNNs model got increased.



We analyzed total of 1796 images of maize leaves, in which 1028 images are NLB disease infected and 768 are noninfected leaves.The images of infected leaves were annoted by NLB leasions are on average of 6.7 lines/image. One important benefit of three-stage pipeline was to make use of full-resolution images.Compared them by scaling the image down, cropping into smaller full-resolution section.

5. Conclussion

There are many method in Computer Vision for plant disease detection but still they are not giving the full accuracy. Convolutional Neural Network model reaches toward but still there are some gap in the result.In this paper Deep CNNs approaches gives a good result.For the future we have to increase the accuracy by making the model more capable in detection and classification.

Limitation have to be overcome in the future like, All photograph were taken at a single place so, generalizability of the dataset should be eliminated through which the symptoms of the same disease in other region able to classify and detect. Some lesion are visible by naked eyes but some are difficult by the model due to heavy shade, washed out by bright light.

6. References

[1] K. A. Garrett, S. P. Dendy, E. E. Frank, M. N. Rouse, and S. E. Travers, "Climate Change Effects on Plant Disease: Genomes to Ecosystems," *Annu. Rev. Phytopathol.*, vol. 44, no. 1, pp. 489–509, 2006.

[2] J. A. Poland and R. J. Nelson, "In the eye of the beholder: The effect of rater variability and different rating scales on QTL mapping," *Phytopathology*, vol. 101, no. 2, pp. 290–298, 2011.

[3] S. Chakraborty, A. V. Tiedemann, and P. S. Teng, "Climate change: Potential impact on plant diseases," *Environ. Pollut.*, vol. 108, no. 3, pp. 317–326, 2000.

[4] J. R. Rohr, T. R. Raffel, J. M. Romansic, H. McCallum, and P. J. Hudson, "Evaluating the links between climate, disease spread, and amphibian declines," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 45, pp. 17436–17441, 2008.

[5] S. A. Miller, F. D. Beed, and C. L. Harmon, "Plant Disease Diagnostic Capabilities and Networks," *Annu. Rev. Phytopathol.*, vol. 47, no. 1, pp. 15–38, 2009.

[6] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[7] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, pp. 70–90, 2018.

[8] D. P. Hughes and M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," 2015.

[9] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.

[10] M. Majumder, "Artificial Neural Network," pp. 49–54, 2015.

[11] N. Shah and S. Jain, "Detection of Disease in Cotton Leaf using Artificial Neural Network," *Proc. - 2019 Amity Int. Conf. Artif. Intell. AICAI 2019*, pp. 473–476, 2019.

[12] K. Ko Zaw, Z. Ma Ma Myo, and D. Thae Hsu Thoung, "Support Vector Machine Based Classification of Leaf Diseases," *Int. J. Sci. Eng. Appl.*, vol. 7, no. 08, pp. 143–147, 2018.

[13] MATLAB, "Convert from HSV to RGB Color Space," p. 1, 2015.

[14] C. DeChant *et al.*, "Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning," *Phytopathology*, vol. 107, no. 11, pp. 1426–1432, 2017.

[15] י. גרינבלט, "Opencv," *עלון הנוטע*, vol. 66, pp. 37–39, 2012.

[16] Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," *MM 2014 - Proc. 2014 ACM Conf. Multimed.*, pp. 675–678, 2014.

[17] Jia Deng, Wei Dong, R. Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," pp. 248–255, 2009.

[18] N. Pittaras, F. Markatopoulou, V. Mezaris, and I. Patras, "Comparison of fine-tuning and extension strategies for deep convolutional neural networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10132 LNCS, pp. 102–114, 2017.

[19] T. Wiesner-Hanks *et al.*, "Image set for deep learning: Field images of maize annotated with disease symptoms," *BMC Res. Notes*, vol. 11, no. 1, pp. 10–12, 2018.

[20] Cornell Univsersity/College of Agriculture and Life Sciences, "Cornell University Agricultural Experiment Station Research Farms," pp. 1–4, 2015.

[21] D. S. Mueller *et al.*, "Corn yield loss estimates due to diseases in the United States and Ontario, Canada from 2012 to 2015," *Plant Heal. Prog.*, vol. 17, no. 3, pp. 211–222, 2016.

8. Data Citation

1. Wiesner-Hanks, Tyr. Image set for deep learning: field images of maize annotated with disease symptoms. Open Science Framework. April 2018. https://osf.io/p67rz/.