MIT International Journal of Computer Science and Information Technology, Vol. 9, No. 1, January 2020 ISSN 2230-7621

Disease Classification in Maize Crop using Deep Learning

Amreen Nisha, Anurita Mishra, Bhawna Chauhan, Asheesh Tyagi, Lal Pratap Verma
Department of Computer Science and Engineering
Moradabad Institute of Technology
Moradabad, Uttar Pradesh, India.

Abstract

Convolutional Neural Networks (CNNs) are the mainstay of image classification, a deep learning marvel that takes an image and consigns it a class and a label that makes it unique. Image classification utilizing CNN sculpts a considerable part of machine learning experiments. The reason CNN is so popular is that it entails truly little pre-processing, meaning that it can read 2D images by utilizing filters that cannot be done by other conventional algorithms. This work presents a study of an experiment done to categorize diseases in Maize crop based on plant leaves by making use of Convolution Neural Network (CNN).

1. Introduction

Maize is extensively cultivated throughout the world, and a larger weight of maize is yielded each year than any other grain. Maize has become an indispensable food in many parts of the world, with the total production of maize outstripping that of wheat or rice. In 2014, total production of maize in the world was 1.04 billion tones. Maize is the most extensively grown grain crop all over the Americas.

With the increasing world population, the prerequisite to enrich the quality and quantity of crops has earned substantial importance. The use of modern tools, especially based on information technology, has enabled humans to accomplish this goal to a great extent almost all over the world. Early discovery of plant leaf stresses empowers farmers to utilize proper treatment and adjust quality and quantity accurately. In most parts of the world, leaf stresses are detected using field surveys which is a time-consuming process and is prone to error. Moreover, exceptional skills are required to point out infected leaves in the field. This gives rise to the need of automatic methods to detect infected leaf in a crop and identify their distinct types.

The automatic methods bring numerous aids such as early diagnosis of a stress facilitating farmers to take precautionary measures, detection, and identification for remedial measures and making the job of field surveyors easy. Amongst many automatic methods, image processing-based methods propose the most discernable solution to this problem. During the last decade, various types of image-processing based detection and classification problems have been successfully solved using end to end deep learning-based models.

Deep Learning has also found its way in agriculture and has been employed to a large assortment of problems in recent years involving identification of weeds, classification of land cover, identifying types of plants, fruits counting and plant leaf stress detection, etc.Deep networks allude to a class of artificial neural networks with multiple hidden layers between input and output layers. With the disposal of GPUs, the consumption of deep networks for various image processing applications has enormously amplified.

Fig. 1 exposes various building blocks of a typical deep-learning based technique for plant leaf disease recognition. The input dataset usually undergoes some preprocessing e.g., contrast enhancement, normalization or data augmentations etc. The final dataset including on average thousands of images is then divided into training and testing portions.

The training data (usually 70% to 80%) of the total dataset is utilized to train millions of constraints of a deep network. This time-consuming step is usually accomplished with the help of an easily accessible parallel computing hardware called Graphical Processing Unit (GPU). The trained model is then tested on left out test data (usually 20-30%) to assess the performance of the learned deep model.

2. Deep Learning

Deep learning has been an area of discussion in image processing-based applications for the last two decades. Most lately its use in several agricultural applications has ascended enormously and plant leaf stress sub-field is also of no exception. We find several recent efforts in this area due to many advantages posed by deep learning in the field of image processing-based applications. Some observable benefits of these techniques are listed as follows:

- The approach consents to develop end to end feature learning-based solution and hence streamlining the proposed architecture.

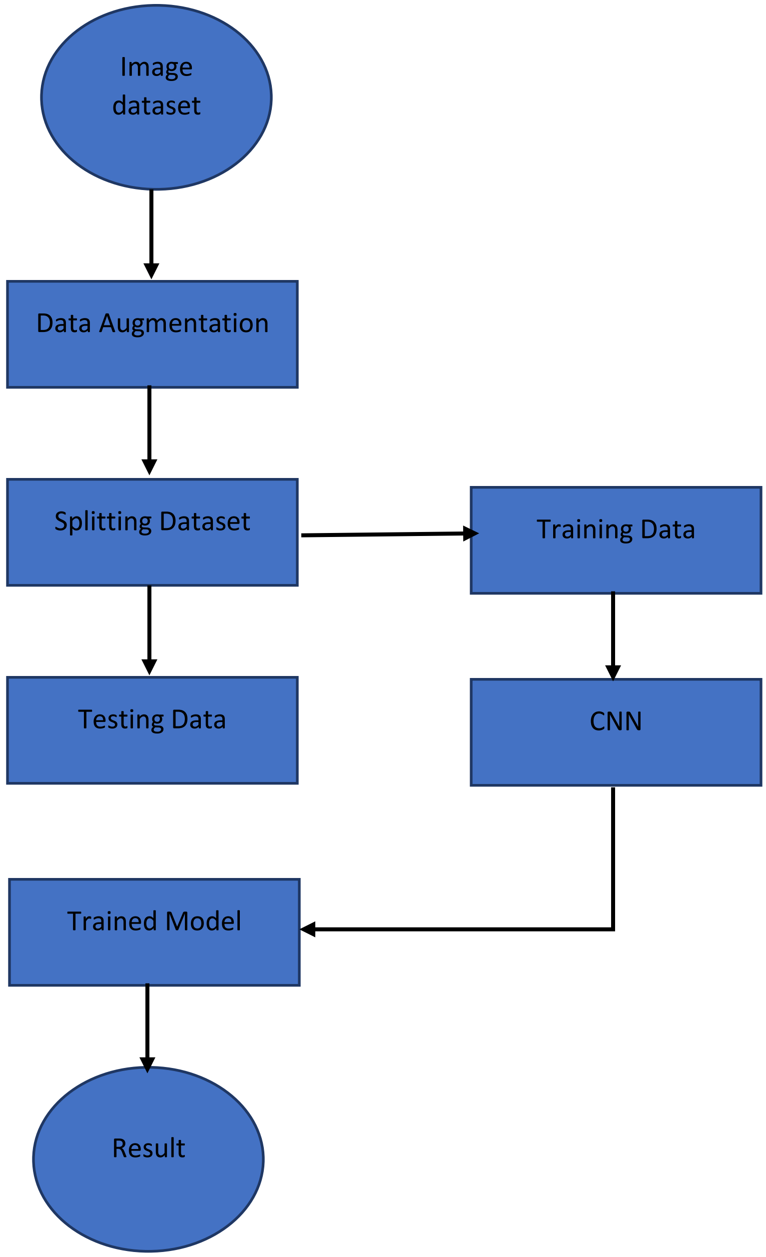

Figure 1: CNN Process

- Automatic feature learning based on accessible data instead of attempting different handcrafted features and then choosing the best for a particular application.
- Disposal of a large number of pre-trained CNN-based models makes the development work swift and easy. Just adjusting / retraining the final few layers enhances the given network according to the aimed application and hence time and effort to train the whole neural network model can be saved.
- Disposal of free software packages in addition to access to extensively used datasets makes it a pick for both beginners and experts in the field.
- Availability of high-end graphical processing units (GPUs) incorporating online access to Tesla K80 GPU catered by Google Colab. The GPU contains 12

GB RAM and is accessible free of cost for a single 12 h session.

The use of deep learning, however, has some drawbacks compared to conventional manual feature extraction-based techniques. The first and the most imperative disadvantage is the prerequisite of a huge amount of training data encompassing a large number of samples from each class (i.e., balanced). Furthermore, an enormous amount of training time even with the help of specific parallel computing hardware i.e., GPUs. Transfer learning is a technique broadly exercised to tackle the above two disadvantages. Using this approach, a pre-trained deep network is transformed such that the final few (mainly one or two) layers are re-trained preserving the remaining structure as such. Transfer learning is appropriate in cases where we have either inadequate training data or cannot meet the expense of hours/days of long training.

3. Convolution Neural network

As cited earlier, deep networks are called deep because they contain more than one hidden-layers between the input and output layers. CNNs are the most famous type of deep networks in image-processing based applications. The extensive use of parallel computing hardware and the convenience of a free development platform have granted the use of CNNs in almost all image processing applications a boost during the last decade. CNNs comprise of recurrent convolutional, non-linear Rectified Linear Unit (ReLU), and pooling layers which make it a deep network. A typical convolutional neural network architecture is displayed in Convolutional layers carry out a typical convolutional operation on the input image(s) but the coefficients of convolutional masks are learned during the training. ReLU introduces a non-linear layer that gets rid of the negative values without disturbing the convolutional or pooling layers.

Pooling layer most of the time means or finds the upper limit in a small portion of the image to prevent overfitting and for translation invariance. Pooling layers have no constraints to learn during training. Like the ReLU, even the pooling layers have to acquire nothing during training. However, the last layer(s) comprises of a conventional fully connected neural network. Fully connected signifies to the fact that every output from one layer is linked to each neuron in the next layer. Depending on the complexity, a large number of parameters are to be acquired in this layer.

4. Datasets

The use of deep learning implies a large amount of data for training the model. Generally, having thousands of images is customary to achieve good accuracy. Among other difficulties linked with plant stress identification, gathering massive amount of real-world field data is a big obstacle. The customary practice used by researchers to overcome this problem is to perform some appropriate preprocessing. This operation enhances original dataset with enormous number of translated, scales and inverted images etc. Moreover, marking field images to know the ground reality requires the advice of a professional pathologist which also incurs surplus cost.The dataset we used for our study is Plant Village dataset from Kaggle. The plant village dataset consists of images of several plant including maize, tomatoes, potato, apple, etc.

For our experiment, we used the images of leaves of Maize Crop. There are more than 7000 images of leaves of maize crop belonging to 4 classes namely Gray Leaf Spot, Common Rust, Healthy and Northern Leaf Blight.

5. Experiment and Result

The data was divided into training, validation and testing sets. Each set has images of leaves of Maize Corn under four classes:

I. Gray Leaf Spot

II. Common Rust

III. Healthy

IV. Northern Leaf Blight

- Training set- 6312 images belonging to 4 classes.
- Validation set- 1829 images belonging to 4 classes.
- Testing set- 1004 images belonging to 4 classes.

A model was built using the Sequential model from Keras models and several layers. The model would take input as arrays of images of size (150,150).

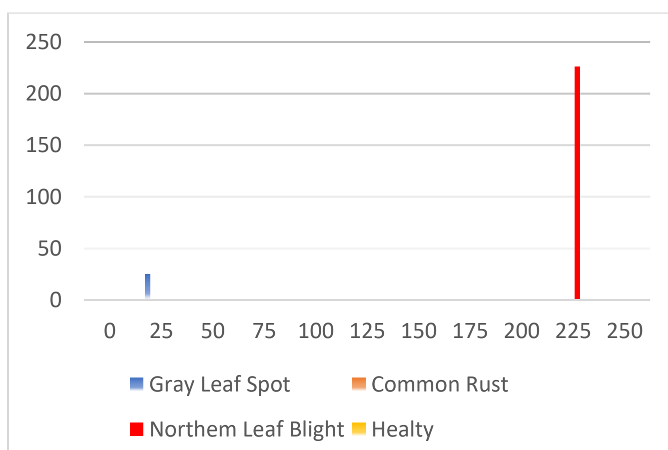

Figure 2: Testing Accuracy

5.1 Training and Validation

The model was trained with the training set and the validation set was used for validation. The model was trained over 25 epochs. After 25 epochs, the model was able to achieve a: Training accuracy of 97.08% and Validation accuracy of 96.61%

The training and validation loss of the model over 25 epochs are shown in the below graph:

5.2 Testing

The model was tested for prediction accuracy using the testing set which consists of 1004 images belonging to 4 classes.

Upon testing, the model was able to achieve an accuracy of 96.71%.

The model was further tested on a **custom testing set** which consisted of **252 images belonging to 1 class** to see the extent to which the model could accurately predict. All the images were from the Class Northern Leaf blight. The model was able to predict 241 images out of 252 images accurately which makes up to an **accuracy of 95.63%.**

6. Conclusion

This study presents a plant disease recognition model based on deep learning. The model has high training and high validation accuracy based on the experimentation results. The model also gave promising results while testing on customary data. The model makes use of Convolutional Neural Networks which makes the overall process fruitful due to high efficiency of CNNs with images.

References

[1] Chopda, J., Raveshiya, H., Nakum, S. and Nakrani, V., Cotton Crop Disease Detection using Decision Tree Classifier, International Conference on Smart City and Emerging Technology (ICSCET), IEEE, pp. 1-5, January 2018.

[2] Kamble, J.K., Plant Disease Detector, International Conference on Advances in Communication and Computing Technology (ICACCT), pp. 97-101, February 2012.

[3] Kulkarni, O., Crop Disease Detection Using Deep Learning, Proceedings of Journal of Global Research in Computer Science, Volume 4, No.4, April 2013.

[4] Mique Jr, E. L. and Palaoag, T. D., Rice Pest and Disease Detection Using Convolutional Neural Network, International Conference on Information Science and System, ACM, pp. 147-151, April 2018.

[5] Serosh Karim Noon, Muhammad Amjad, Muhammad Ali Qureshi, Abdul Mannan, Use of deep learning techniques for identification of plant leaf stresses: A review, Sustainable Computing: Informatics and Systems (ELSEVIER), 2020.

[6] HE Dun-chun, ZHAN Jia-sui, XIE Lian-hui, Problems, challenges and future of plant disease management: from an ecological point of view, ScienceDirect (ELSEVIER), Journal of Integrative Agriculture 2016, 15(4): 705–715.

[7] Parul Sharma, Yash Paul Singh Berwal, Wiqas Ghai, Performance analysis of deep learning CNN models for disease detection in plants using image segmentation, Information Processing In Agriculture, Received 8 July 2019 Received in revised form 7 November 2019 Accepted 14 November 2019.

[8] M. Akila, P. Deepan, Detection and Classification of Plant Leaf Diseases by using Deep Learning Algorithm, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, 2018.

[9] Aayush Singh, Deepanshu Sonparote, Nitish Talekar, Dilip Kale, A Comparative Study on Crop Disease Detection using Deep Learning Techniques, International Research Journal of Engineering and Technology (IRJET), ISSN: 2395-0056, Volume 7 Issue: 05, May 2020