International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 11 | Nov 2020 www.irjet.net p-ISSN: 2395-0072

POTATO LEAF DISEASE DETECTION USING INCEPTION V3

Gunjan Chugh[1], Aman Sharma[2], Prakash Choudhary[3], Rushali Khanna[4]

[1]Assistant Professor, IT Engineering, ADGITM, New Delhi
[2, 3, 4]Student B-Tech-IT, ADGITM, New Delhi

ABSTRACT *- Agriculture is one of the major sectors in our society and from the medieval times, it is the sector that we have dwelled upon. About 60-70% of Indian population depends on the agricultural industry. Across the world, loss faced by the crop due to numerous factors like weeds, diseases and arthropods have increased to an alarming rate of about 34.9% in 1965 to about 42.1% in the late 1990s. Potato plants face various infections due to bacteria and fungi. Fungal diseases infecting plants are early blight and late blight. Our paper presents a CNN model and an algorithm to detect such diseases through the leaf of the crop. The model that we have created is trained to analyse and understand a diseased leaf and thereafter recognise the disease of the leaf. We are using InceptionV3 algorithm.*

KEYWORDS - Convolutional neural network, Deep Learning, InceptionV3, Early blight, Late Blight, Softmax, Adam, Plant Village.

1. INTRODUCTION

For any civilization, agriculture is one of the building block in the society. For many years India is one of the major countries involved in the agricultural sector [2]. With a stellar history in agriculture, this industry forms an important part of the economy. Among the various crops grown in India, potato is the most in-demand crop. In India's diverse range of subtropical climate, growing vegetables like potato yields fruitful results.

A diseased plant is one which has been hampered in terms of the normal state that it should have governed. It can also be said that a disease interferes with the yield and decreases that plants vitality. With the changing season in India, diseases tend to vary according to the environmental conditions. These diseases are governed by pathogens and the variety of the crop grown in the season. Diseases common in potatoes are Late Blight and Early Blight [2]. They can have adverse effects on potato plants and agricultural lands. Late Blight and Early Blight can be spotted on the plant leaves but it's is very time consuming if spotting is done manually. Thus, newer reforms are required.

There are many ways to detect objects and its special features from image with the help of image processing and computer vision. One of the most popular approach is deep learning CNN model [3]. In our case, the model can detect disease from the image of the leaves.

2. CONVOLUTION NEURAL NETWORKS

Convolutional neural networks are processing techniques applied to images of the jpg format. To understand and realise them [3]. They are also known as shift invariant. This is so because of their weight sharing architecture and translation invariance characteristics. Thus, they are also referred to as space invariant neural networks. They have applications in recommender systems, image and video recognition, medical image analysis, image classification, natural language processing(NLP), brain-computer interfaces that communicate between them, and financial time series. They also govern numerous more applications in varied fields possible.

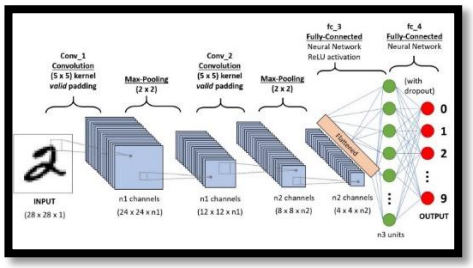

Figure 1: Convolution Neural Network

Source: Adapted from [11]

3. RELATED WORKS

Detecting diseases in plants in early stages is one of the major concerns in Agriculture field. So, there are many researchers who are already working on the issue of detecting plant diseases and diagnosing them some of the results of the research done in detecting plant diseases are as follows:

D. Tiwari et. al. [2] used transfer learning and various pre-trained model on the potato leaves image dataset and concluded that VGG 19 gave the optimal accuracy of 97.8% compared to 92% accuracy of back propagation neural network and 95% accuracy of support vector machine. Melike Sardogan et. al. [4] in their research has tried to use Convolutional Neural Network with Learning Vector Quantization algorithm to detect and classification of tomato leaf diseases. They used 500 tomato leaves images from which three input matrices have been obtained for R, G and B channel for the start of

convolution for every image in the dataset. Then they used LVQ algorithm for image classification and disease detection. In their research they concluded that LVQ algorithm with CNN effectively classify the types of tomato leaf diseases. Halil Durmus et. al. [5] in their research has tried to perform Diseases detection with the help of two architectures which are AlexNet and SqueezeNet on the hardware Nvidia Jetson Tx1. In their research they concluded that AlexNet is not suitable to perform diseases detection on mobile devices because the model developed by AlexNet is very bulky with the size of 227.6 Mbyte whereas the model developed on SqueezeNet architecture is very small with the size of 2.9Mbyte and there is also significant improvement in the inference time. So, SqueezeNet is the best suitable architecture for mobile devices like Nvidia Jetson Tx1. U. Atila et. al. [6] tried to compare state-of-the-art deep neural net architectures with the EfficientNet deep learning architecture to detect plant leaf disease detection in the Google Cloud Environment. They found that the accuracy EfficientNet architecture is better than other CNN architectures with the accuracy of 99.97%. The precision of EfficientNet is also better than other CNN architectures.

4. DATASET DESCRIPTION

A dataset is the basic step of a deep learning driven paper. It helps us not only to train the model, but can also be used to study and infer many more results from it. The dataset that is used in this paper is Plant Village Dataset [1] that is taken from Kaggle. Kaggle is an online community for data scientists and ML engineers which allows them to participate in ML competitions, work on various datasets and notebooks [2]. In this dataset there are about 20000 images of leaves of tomato plants, bell-pepper plants, potato plants and many more. The dataset contains jpg/png format of images. This dataset is a combination of healthy and diseased leaves. The diseased leaves are divided into two categories namely early blight and late blight. 80% of this dataset for healthy leaves goes into training and the rest into test; and so is the case for each of the diseased leaves. Thus, the model created will be trained on more leaves and tested on less to achieve the highest accuracy during prediction and detection. There is a different directory for each plant and each disease corresponding to that plant has separate folder.

The subset of the dataset we are using is the Potato plant dataset. In this dataset there are 1000 images of early blight leaf images, 1000 images of late blight and 152 images of heathy leaves. Some images from the dataset are shown below:

Figure 2: Early Blight Figure 3: Late Blight

Source: Adapted from [1] Source: Adapted from [1]

Figure 4: Healthy leaf

Source: Adapted from [1]

5. PLATFORM UTILIZED

The hardware used in the paper is a laptop with the specification of 8 GB DDR4-2133 SDRAM (1 x 8 GB) with Intel Core i5-7200U (2.5 GHz, up to 3.1 GHz, 3 MB cache, 2 cores) and 1 TB 5400 rpm SATA hard drive. For data visualization we have used Matplotlib library to plot graphs related to the model's performance.

Matplotlib is an open-source python library used as a plotting tool. For Machine learning we have used TensorFlow and Keras library. TensorFlow and Keras are open-source software library for deep neural nets [7].

6. PROPOSED APPROACH

In this paper, we have used a CNN based model that trains on a pre-processed dataset. This model is furthermore tested to receive the models accuracy and our desired conclusion.

6.1 FEATURE EXTRACTION

In this paper, for the purpose of feature extraction, Inception v3 architecture [7] is used. Inception is developed by google and many other researchers together. The building blocks of inception v3 are convolutions, max pooling, concats, dropouts, fully connected layers and average pooling. Batchnorm is also used throughout the model and applied to activation inputs [7]. Feature extraction helps model to clearly distinguish between all the characteristics of the image and understand them for further interpretation [3].

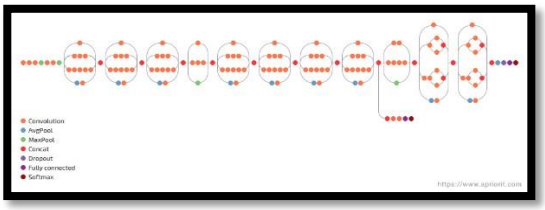

Figure 5: Inception V3 Architecture

Source: Adapted from [12]

6.2 CLASSIFICATION

In this paper we have used convolutional neural networks for the purpose of classification. After acquiring features from inception v3 [7], deep neural networks are used to train the model to detect presence of early blight and late blight in the images. During classification Adam optimiser is used to reduce the training time and quickly converge the loss [8]. Adam [8] is an extension of SGD that is now widely used for computer vision. For classification in different labels Softmax activation function is used [9]. The softmax activation function is a function that converts a vector of n real values to a vector of n real values that totals up to 1[9]. The input values can be any number but the softmax [9] transforms them into values between 0 and 1, so that they can be interpreted as probabilities.

6.3 MODEL

The CNN model contains an input layer where all images are fed in the start. After which the images are fed to the inception v3 architecture [7] to extract features. These feature extracted images are sent to deep neural nets that classify our images on pre-trained knowledge. Finally, the images are sent to the output layer.

The model has a couple of layers that are first created and after that they are compiled together using functions from the TensorFlow library [7].

Figure 7 : CNN Model

Source: Adapted from [13]

7. RESULTS

The model that we have proposed in this paper has used the plant village dataset [1] which has about 1000 leaf images of early blight and 152 images of healthy plants. For this model, the dataset has been divided into 2 parts that are the training set and the test set. Training set comprises of 80% of the dataset while the test set comprises of 20% of the dataset. The pre-trained model used on this dataset for feature extraction is Inception V3 [7]. Here for classification our CNN model provides an accuracy of about 90% on the basis of the training and testing done on it.

8. CONCLUSION

In this research, we have made a CNN model with the help of Inception V3 architecture and Adam Optimiser to diagnose and classify disease of potato plants such as early and late blight where we achieved an accuracy of 90% over the test dataset in classification. With the help of our model, a farmer can build a computer setup from where he can monitor the plants health issues efficiently, enhance his crop yield and detect and diagnose diseases in the early stages itself.

9. ACKNOWLEDGEMENT

We would like to thank our college Dr. Akhilesh Das Gupta Institute of Technology and Management for providing us with this opportunity to write a research paper along with our minor project. We would also like to thank Ms. Gunjan Chugh for being our mentor and mentoring us all through the paper. Without them, it wouldn't have been possible to successfully carry out this research paper.

10. FUTURE SCOPE

In recent years, there are many healthy ways to tackle diseases but still if the disease is detected at an early stage, it causes much less of a damage and can be cured more easily. Whereas on the other hand, if it is detected on a later stage it can lead to a very poor yield and hence also incur extra cost of pesticides and manure. Only a detection of disease is not going to help the farmer from preventing all sorts of losses. Thus, it could be done that we install another AI driven bot to take the necessary steps for the cure of these diseased leaves.

11. REFERENCES

[1] Tairu Oluwafemi Emmanuel, " PlantVillage Dataset ", [online] Available: https://www.kaggle.com/emmarex/plantdisease.

[2] D. Tiwari, Adem Tuncer, S. Bhardwaj M. Ashish, N. Gangwar, A. Sharma, S. Patel, "Potato Leaf Diseases

Detection Using Deep Learning", International Conference on Intelligent Computing and Control Systems, 2020.

[3] Chee Seng Chan, Sue Han Lee, Simon Joseph Mayo, Paolo Remagnino, "How Deep Learning Extracts and Learns Leaf Features for Plant Classification", Pattern Recognition, 2017.

[4] Adem Tuncer Melike Sardogan, and Yunus Ozen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm", International Conference on Computer Science and Engineering, 2018.

[5] H. Durmus¸, E. O. Gunes ¨ ¸ and M. Kırcı, "Disease detection on the leaves of the tomato plants by using deep learning", In Agro-Geoinformatics, IEEE 6th International Conference on, pp. 1-5, 2017.

[6] Ü. Atila, M. Uçar, K. Akyol, et al., Plant leaf disease classification using EfficientNet deep learning model, Ecological Informatics, 2019.

[7] Xiaoling Xia, Cui Xu and Bing Nan, "Inception-v3 for flower classification," 2017 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, 2017, pp. 783-787.

[8] Kingma, Diederik & Ba, Jimmy. (2014). Adam: A Method for Stochastic Optimization. International Conference on Learning Representations.

[9] Nwankpa C, Ijomah W, Gachagan A, Marshall S (2018) Activation functions: comparison of trends in practice and research for deep learning. Preprint arXiv:1811.03378

[11] Saha, S. (2018). *A CNN sequence to classify handwritten digits* [Online image]. Towards Data Science. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

[12] Boyko, S. (2019). The architecture of the Inception neural network [Online image]. Apriorit. https://www.apriorit.com/dev-blog/606-inception-v3-for-video-classification

[13] Tabian, I.; Fu, H.; Sharif Khodaei, Z. A Convolutional Neural Network for Impact Detection and Characterization of Complex Composite Structures. Sensors 2019, 19, 4933.