Plant Disease Detection Using Deep Learning

Prof. Ksn Sushma[1], Nishant Upadhyay[2], Ajeet Singh[3], Prasenjeet Kr. Singh[4], Tanzeelah Firdaus[5]

[1, 2, 3, 4, 5]Department of Computer Science, HKBK college of Engineering, Bangalore, India

Abstract: Early diagnosis of plant diseases is critical since they have a substantial impact on the growth of their unique species. Many Machine Learning (ML) models have been used to detect and categorize plant diseases, but recent breakthroughs in a subset of ML called Deep Learning (DL) look to hold a lot of promise in terms of improved accuracy. A variety of developed/modified DL architectures, as well as several visualization techniques, are utilized to recognize and identify the symptoms of plant ailments. In addition, a number of performance measurements are used to evaluate various architectures/techniques. This article explains how to use DL models to display a variety of plant diseases. Furthermore, several research gaps are identified, allowing for improved efficiency in detecting plant illnesses even before issues emerge.
Keywords: Plant disease; deep learning; convolutional neural networks (CNN), Google Net Architecture, Tensorflow, and PyTorch are some of the tools that can be used;

I. INTRODUCTION

Plant disease diagnosis is difficult and time-consuming, and it is done by visually inspecting symptoms on plant leaves. Due to this complication and the vast number of farmed plants and their accompanying phytopathological difficulties, even experienced agronomists and plant pathologists frequently fail to recognize specific infections, resulting in inaccurate assumptions and remedies. Agronomists who are asked to make such diagnoses by gazing at sick plants' leaves with their eyes would greatly benefit from the availability of an integrated computational system for identifying and diagnosing plant ailments.

If the technology was simple to use and accessible via a mobile app, it may be a useful tool for farmers in locations where there are few resources for agronomic and phytopathological advice. Furthermore, in large-scale plantings, the technology might be used in conjunction with autonomous agricultural gear to discover phytopathological issues across the cultivating area using continuous image capture. When fraudsters use it in such instances, it is used until its entire usable limit is spent. As a result, we require a solution that reduces the overall permissible limit on the credit card, which is more vulnerable to fraud. Furthermore, as time passes, a Genetic algorithm creates better answers. The development of an efficient and secure electronic payment system for identifying fraud is given top priority. Artificial Intelligence (AI) Due to the advancement of computational frameworks, particularly Graphical Processing Units (GPU) embedded processors, Artificial Intelligence implementations have grown exponentially in the last few years, resulting in the advancement of novel methodologies and designs, which have now spawned a new classification, Deep Learning. Deep learning involves the use of artificial neural network structures with many functional layers, as opposed to the "swallower" topologies utilized in more traditional neural network methodologies.

The principal deep learning technology used in this study is convolutional neural networks (CNNs). In applications that need a large quantity of data, such as picture recognition, CNNs are one of the most powerful algorithms for modeling complex processes and performing pattern detection. Certain CNN architectures were trained and tested using simple images of healthy and diseased plant leaves in order to construct an automated plant disease identification and diagnostic system. The dataset comprised images from both experimental setups and real-world farming conditions. Deep learning approaches, which use less data but are unique to a few crops, may find more general answers than shallow learning approaches, which use more data but are specific to a few crops.

II. RELATED WORKS

The analysis basically deals with classifying weather the transaction is a fraudulent or not which implies that the problem is a classification problem and can be solved using classification algorithms.

A. Identitification of diseases in Plants

This research examines how image processing algorithms are developed, with a particular focus on how they might be applied to crop monitoring. The approaches are classified into two categories: 'healthy plant classification' and 'diseased plant classification,' with a focus on classification and recognition consistency, rapid stress identification, and disease extent. The paper's main focus is on hyperspectral imaging and how it's being utilized to understand more about plant health and predict disease onset. The author provides a list of ways for detecting biotic and abiotic stress in plants, as well as the level of accuracy for each method.

B. Deep Learning Models for Plant Disease Detection and Diagnosis

In this article, convolutional neural network designs were created utilizing deep learning methodologies to detect and diagnose plant diseases using simple leaf photos of healthy and diseased plants. The models were trained using 87,848 pictures from the public domain, which included 25 distinct plants in 58 different classes of [plant, disease] pairs, including healthy plants. Various model designs were trained, with the top-performing one detecting the relevant [plant, disease] pair with a success rate of 99.53 percent (or healthy plant). The model's high success rate makes it a useful advising or early warning tool, and it's an approach that might be expanded to create a complete plant disease diagnosis system that can be used in real-world scenarios.

C. Plant Disease Detection and Classification by Deep Learning

Early diagnosis of plant diseases is critical since they have a substantial impact on the growth of particular species. Many Machine Learning (ML) methods have been used to detect and categorize plant diseases, but recent breakthroughs in a subset of ML called Deep Learning (DL) appear to hold a lot of promise in terms of improved accuracy. A variety of developed/modified DL architectures, as well as several visualization techniques, are utilized to recognize and identify the signs of plant ailments. In addition, a number of performance measurements are used to evaluate various architectures/techniques. This article explains how to use DL models to display a variety of plant diseases.

III. MATERIALS AND METHODS

A. Convolutional Neural Network Models

Artificial neural networks (ANNs) are computer models that, through their neurons and synapses, simulate the basic principles of brain activity. Their main distinguishing characteristic is their ability to be trained through guided learning. In this method, neural networks are "trained" to represent a system using available data that includes particular matchings of the system's inputs and outputs.

CNNs are a development of standard artificial neural networks that are primarily utilized for applications involving repeated patterns in various areas of the modeling space, such as image recognition. AlexNet and GoogLeNet were the two fundamental CNN architectures used in this work, which aimed to recognize plant illnesses from photographs of its leaves. The Torch71 machine learning computational framework, which runs on the LuaJIT2 programming language, was used to develop these algorithms, as well as their training and testing methods.

B. Training and Testing Datasets

A publicly accessible set of 54,306 photos of healthy and diseased plant leaves was used to train and assess the CNN models. Hughes and Salathé present a database sample with fewer images (2015). The database utilized here is divided into 38 different classes, each with a pair of plants and diseases connected with it, with some classes featuring healthy plants.

By randomly separating the 54,306 images into three datasets, the training set, and the testing set, the entire data was initially divided into 3 datasets, the training set, and the testing set. The most common training/testing dataset dividing ratio in neural network applications is 50/50; the second of similar partition ratios (60/40); and the third ratio (80/20) should have no major influence on the functionality of the models developed. With these 38 classes, two studies were conducted, and two CNN prototypes were established: one that was practiced with only research lab surroundings pictures and tested on field conditions images, and the other that was trained and tested in a reverse way, that is, on-field photos and then tested on lab images. The training/testing ratio was not optimum in both circumstances, as the percentage of laboratory settings images in those 38 classes was 55.8% and the percentage of field scenarios images was 44.2 percent, significantly below the 80/20 ratio employed in the construction of the basic model.

IV. RESULTS

To initialize, we should take note that for a dataset with 38 class labels, random guessing will only achieve an overall accuracy of 2.6 percent. On the PlantVillage dataset, our total accuracy ranged from 85 percent to 99 percent across all of our experimental setups, which is comprised of three visual representations of the image data, indicating that the deep-learning approach has a lot of potential for similar prediction challenges.

To avoid over-fitting, we modify the test set to train set ratio and discover that in case of GoogLeNet-TransferLearning, Color-20–80, the model achieves an overall accuracy of 98.21%, despite training on only 20% of the data and testing the trained model on the remaining 80%. The overall performance of both AlexNet and GoogLeNet diminishes as the test set to train set ratio is increased as we expected, but not as significantly as we would expect if the model was over-fitting.

The three versions of the dataset (color, gray-scale, and segmented) show a consistent variation in performance across all tests while the rest of the experimental arrangement is kept the same. The models show best performance with the colored version of dataset. When designing the research, we were concerned that the neural networks would only learn to detect inherent biases in the illumination, data collection method, and instrument.

We examined the model's performance on pictures from reliable internet sources, such as academic agricultural extension agencies, and found that these methods provide good results on the PlantVillage dataset, which was collected in a controlled environment.

We were able to predict the correct class label accurately from among 38 candidate class labels using the best model on these datasets, with an overall accuracy of 31.40 percent in dataset-1 and 31.69 percent in dataset-2. The average accuracy of a random classifier was only 2.63 percent. The correct class was in the top five predictions in 52.89 percent of cases in dataset 1, and it was in the top five predictions in 65.61 percent of cases in dataset 2. The best models for both datasets were GoogLeNet:Segmented:TransferLearning:80–20 for dataset 1 and GoogLeNet:Color:TransferLearning:80–20 for dataset 2.

V. ACKNOWLEDGMENT

REFERENCES

[1] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (surf). Comput. Vis. Image Underst. 110, 346–359. doi: 10.1016/j.cviu.2007.09.014

[2] Garcia-Ruiz, F., Sankaran, S., Maja, J. M., Lee, W. S., Rasmussen, J., and Ehsani R. (2013). Comparison of two aerial imaging platforms for identification of huanglongbing-infected citrus trees. Comput. Electron. Agric. 91, 106–115. doi: 10.1016/j.compag.2012.12.002

[3] He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition. arXiv:1512.03385.

[4] Huang, K. Y. (2007). Application of artificial neural network for detecting phalaenopsis seedling diseases using color and texture features. Comput. Electron. Agric. 57, 3–11. doi: 10.1016/j.compag.2007.01.015

[5] Mokhtar, U., Ali, M. A., Hassanien, A. E., and Hefny, H. (2015). "Identifying two of tomatoes leaf viruses using support vector machine," in Information Systems Design and Intelligent Applications, eds J. K. Mandal, S. C. Satapathy, M. K. Sanyal, P. P. Sarkar, A. Mukhopadhyay (Springer), 771–782.

[6] https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

[7] https://en.wikipedia.org/wiki/Transfer_learning

[8] https://towardsdatascience.com/alexnet-the-architecture-that-challenged-cnns-e406d5297951

[9] https://www.mathworks.com/help/deeplearning/ref/googlenet.html

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)