Deep Learning
Project Report
(CSE4003)

Potato Disease Classification

By

Harsh Gupta

190076

Kartikey Choudhary

190116

BML MUNJAL UNIVERSITY™

Department of Computer Science and Engineering
School of Engineering and Technology

April, 2022

Contents

Abstract

Many plant diseases have distinct visual symptoms which can be used to identify and classify them correctly. This report presents a potato disease classification algorithm which leverages these distinct appearances and the recent advances in computer vision made possible by deep learning. The algorithm uses a deep convolutional neural network training it to classify the tubers into five classes, four disease classes and a healthy potato class. The database of images used in this study, containing potatoes of different shapes, sizes and diseases, was acquired, classified, and labelled manually by experts. The models were trained over different train-test splits to better understand the amount of image data needed to apply deep learning for such classification tasks.

Problem Definition

Classifying the potato leaves into three categories:

A. Early Blight
B. Late Blight
C. Healthy

1. Project Objectives

- To get more experiential knowledge on Convolutional Neural Network (CNN)
- To achieve good accuracy in predicting leaves whether it is Blight or healthy.

2. Challenges

- Data preprocessing and large data handling
- Training a machine learning model using less computational requirements is very difficult. Hence, we chose to optimize the network by reducing the number of learned parameters in each layer.

3. Literature Review

Year	Author	Objective	Dataset	Classifier	Result
2022	Priya Rk	Potato leaf disease detection	2500	CNN	97% of accuracy
2020	Tiwari, Divyansh and Mritunjay; Gangwar et al.	Potato leaf disease detection	1.3 million images(ILSVRC)	logistic regression, inception V3, VGG16, VGG19	97% of accuracy
2020	Sholihati and Rizqi Amaliatus et al.	Potato leaf disease detection	5,100 images	VGG16 and VGG19 convolutional neural network	91% of accuracy
2019	Suttapakti et al.	Potato leaf disease classification	300 potato leaf images	k-means clustering	91.67 % of accuracy
2018	Shima Ramesh and Mr. Ramachandra Hebbar et al.	Papaya leaves disease classifiacion	160 images	Random forest, histogram oriented gradient(HOG)	70% of accuracy
2018	S.Ramesh and D.vydeki	Rice plant about the disease	300 images	Automated plant identification	97% of accuracy
2017	Monzural Islam	Plant leaf disease detection.	300 images	SVM image segmentation	95% of accuracy
2016	Harshal Waghmare and Radha Kokare	Plant disease detection analysis and pattern detection of the leaf texture.	140 images	SVM image segmentation	96.6% of accuracy
2016	Parsad et al.	Plant leaves disease detection(5 diseases)	297 images	KNN	93% of accuracy
2015	Semary et al.	Plant leaves disease detection(12 diseases)	177 images	SVM	92% of accuracy
2015	Dandawate et al.	Plant leaves disease detection(2 diseases)	120 images	SVM	93.79% of accuracy
2015	Raza et al.	Plant leaves disease detection(2 diseases)	71 plants	SVM	89.93% of accuracy

4. Description of the Dataset

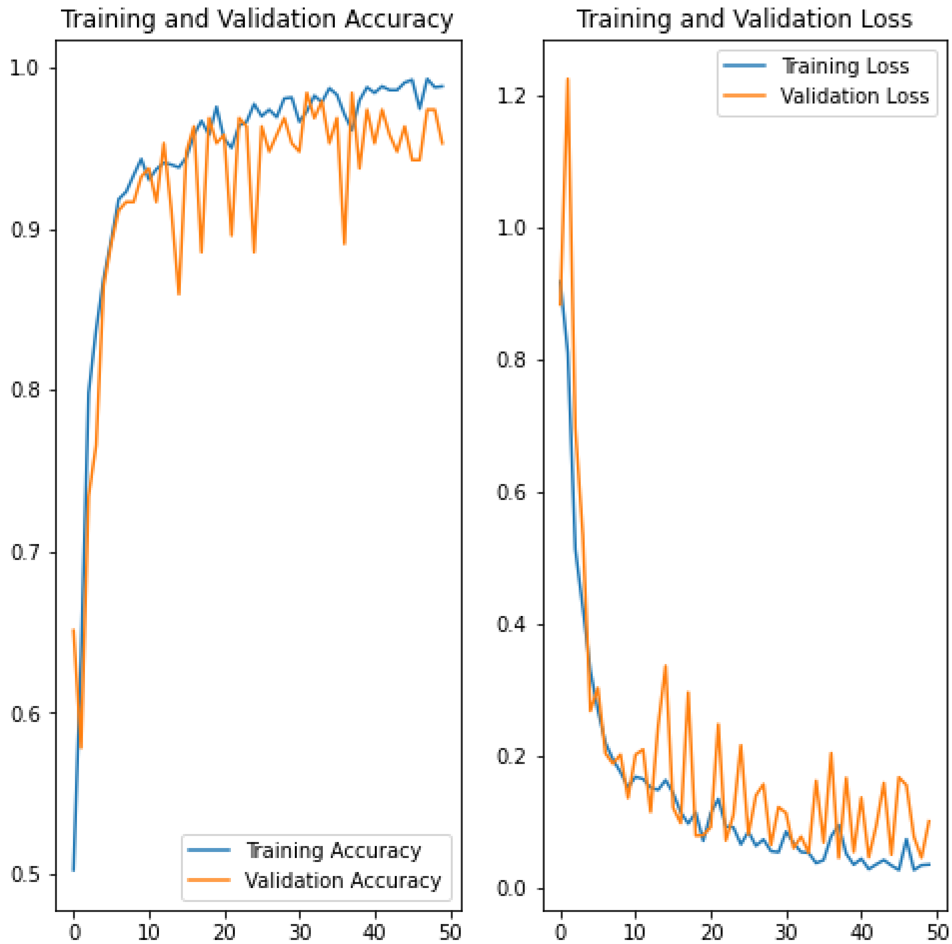
The performance of deep learning models heavily depends upon an appropriate and valid dataset. In this report, the following datasets are used.
The Potato leaf disease detection convolutional neural network (PDDCNN) method’s performance is assessed using the potato leaf images of a publicly available dataset called PlantVillage. The PlantVillage dataset was developed by Penn State University (US) and EPFL (Switzerland), which is a non-profit project. The database consists of JPG color images with 256 × 256 dimensions. It has 38 classes of diseased and healthy leaves of 14 plants. The focus of this report is on the potato crop. Therefore, 1000 leaves for late blight, 1000 leaves for early blight, and 152 images of healthy leaves were selected for the experimental purposes.

5. Proposed Methodology

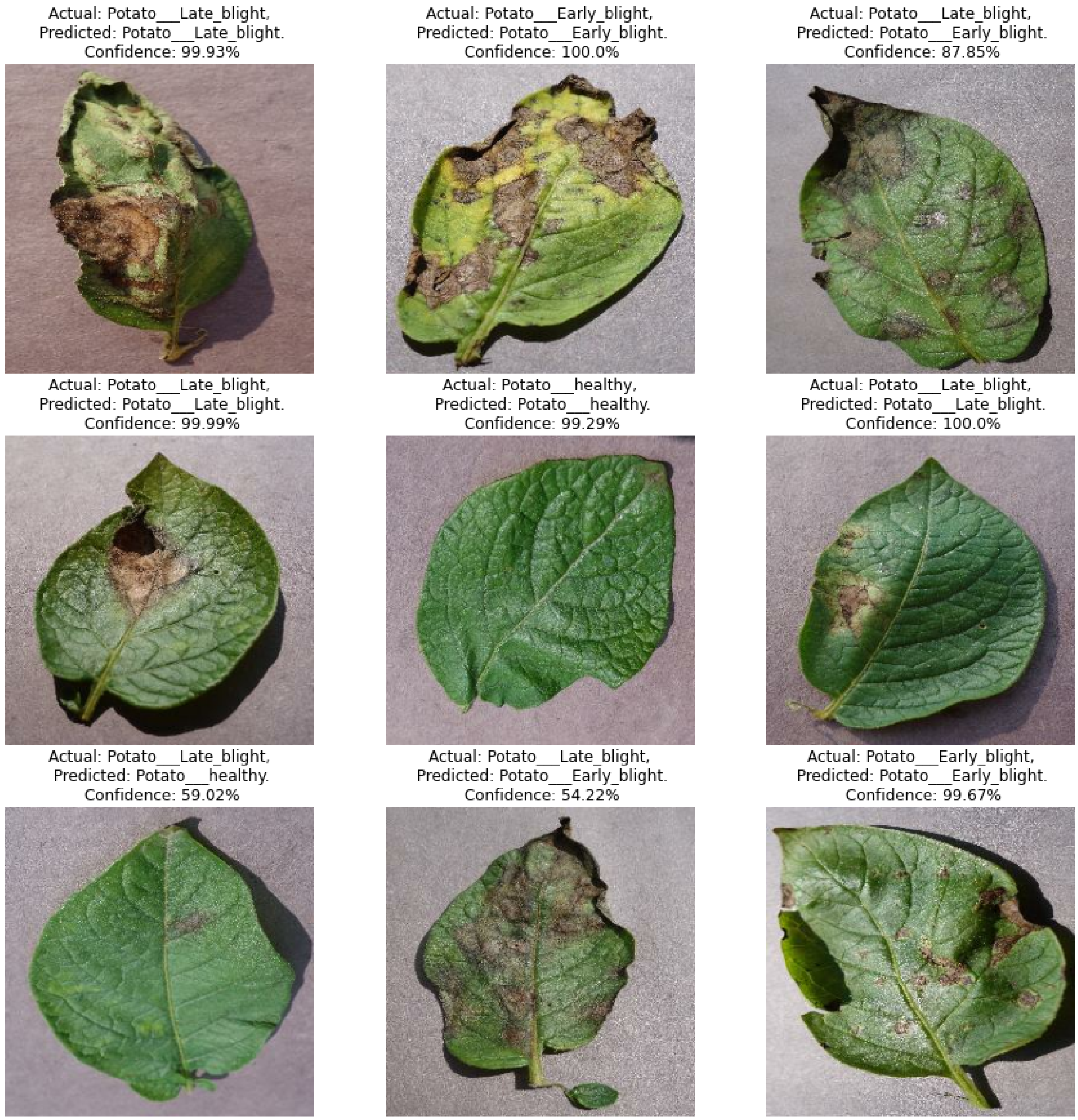
The process of detection of diseases from plant leaf images involves various steps and each of those steps can be discussed as follows.

Step-1: Image Acquisition: It deals with the acquisition of data from reliable sources to maintain the standard and stability so that it can be compared or extended for future studies.
Step-2: Image Pre-processing: It is a very essential phase of the framework. In this phase mainly deals with the denoising of the image, enhancement of the image, and maintaining standard image size for all the images. Denoising and enhancement of images are essential to get a better result while segmenting the images.
Step-3: Image Segmentation: In this phase, the image will be segmented according to the region of interest. Here, in this case, the region of interest is the regions on the leaf which are affected by various diseases that need to be separated from the existing images.
Step-4: Extraction of Features: Depending on the obtained region of interest need to identify the patterns that exist. A different region of interest will have different patterns, from that scenario, one can able to extract features that are crucial in deciding the detection as well as classification.
Step-5: Processed Data: All the information related to processed image data by using the steps-1 to 5 will be gathered into a single location.
Step-6: Training Data: The training data will be obtained from the processed data. About 75% of the data with random indexing was considered to train the classifier model.
Step-7: Testing Data: The testing data will also be obtained from the processed data. About 25% of the data with random indexing was considered to test the classifier model.
Step-8: Classification: Test data will be provided to the trained classifier to classify the images into various categories such as Late Blight, Early Blight, and Healthy.

6. Experimental Results/Comparison



This is the graphical representation between training and validation accuracy/loss.



This is the prediction for the sample images.

7. Justification of novelty of the proposal

The disease identification using the integration of CNNs, machine vision, and assessing potential and feasibility to have integrated them into a smart sprayer adds novelty and innovation to this project is being reported here. Due to better accuracy our model performance is better as compared to the existing models.

8. Role & Contributions of each group member

We have equally divided our work to do this project. We together collected the datasets and read all the necessary papers. Kartikey looked into preprocessing of the images and I looked into the machine learning architecture that we can use and then Kartikey compared all the models and suggested which model performed the best.

9. Conclusions and Future Scope

Crop diseases are a common threat to food security in all nations, but with the deep learning advances in detection and classification allow these threats to be maintained and eliminated at early stages rapidly and accurately. This report presented a proposed deep learning architecture for potato leaf blight classification. The proposed architecture consists of 16 layers: two main convolutional layers for feature extraction with different convolution window sizes followed by two fully connected layers for classification. The proposed architecture achieved an overall mean testing accuracy of 95%.The proposed architecture achieved better results than the related works in terms of overall testing accuracy. One prospect for future work is to apply transfer learning based on more advanced pre-trained deep neural network architectures such as AlexNet, VGG-16, and VGG-19. Using pre-trained architectures may reduce the computation time in the training phase and may lead to better testing accuracy.

10. References

1. Iqbal, Z.; Khan, M.A.; Sharif, M.; Shah, J.H.; ur Rehman, M.H.; Javed, K. An automated detection and classification of citrus plant diseases using image processing techniques: A review. Comput. Electron. Agric. 2018, 153, 12–32. [CrossRef]

2. Oerke, E.C. Crop losses to pests. J. Agric. Sci. 2006, 144, 31–43. [CrossRef]

3. Golhani, K.; Balasundram, S.K.; Vadamalai, G.; Pradhan, B. A review of neural networks in plant disease detection using hyperspectral data. Inf. Process. Agric. 2018, 5, 354–371. [CrossRef]

4. Chouhan, S.S.; Kaul, A.; Singh, U.P.; Jain, S. Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: An automatic approach towards plant pathology. IEEE Access 2018, 6, 8852–8863. [CrossRef]
5. Sasaki, Y.; Okamoto, T.; Imou, K.; Torii, T. Automatic diagnosis of plant disease-Spectral reflectance of healthy and diseased leaves. IFAC Proc. Vol. 1998, 31, 145–150. [CrossRef]
6. Henson, J.M.; French, R. The polymerase chain reaction and plant disease diagnosis. Annu. Rev. Phytopathol. 1993, 31, 81–109. [CrossRef]
7. Koo, C.; Malapi-Wight, M.; Kim, H.S.; Cifci, O.S.; Vaughn-Diaz, V.L.; Ma, B.; Kim, S.; Abdel-Raziq, H.; Ong, K.; Jo, Y.K.; et al. Development of a real-time microchip PCR system for portable plant disease diagnosis. PLoS ONE 2013, 8, e82704. [CrossRef]
8. Singh, U.P.; Chouhan, S.S.; Jain, S.; Jain, S. Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease. IEEE Access 2019, 7, 43721–43729. [CrossRef]

Appendix

I am attaching the python file in the zip file.