Plant Disease Identification using Leaf Images

1 Problem Statement

One of the important sectors of *Indian Economy is Agriculture*. Employment to almost 50% of the countries workforce is provided by Indian agriculture sector. *India* is known to be the world's largest producer of pulses, rice, wheat, spices and spice products. Farmer's economic growth depends on the quality of the products that they produce, which relies on the plant's growth and the yield they get. Therefore, in field of agriculture, detection of disease in plants plays an instrumental role. Plants are highly prone to diseases that affect the growth of the plant which in turn affects the ecology of the farmer. In order to detect a plant disease at very initial stage, use of automatic disease detection technique is advantageous. The symptoms of plant diseases are conspicuous in different parts of a plant such as leaves, etc. Manual detection of plant disease using leaf images is a tedious job. Hence, it is required to develop computational methods which will make the process of disease detection and classification using leaf images automatic.

2 Background Work

Despite of the challenges given in the problem statement plant disease detection is still an active area of research. Numerous approaches have been proposed over the years. In traditional systems approach for detection and differentiation of plant diseases can be achieved using Support Vector Machine algorithms. This technique was implemented for sugar beet diseases and depending on the type and stage of disease, the classification accuracy was between 65% and 90%. Another approach based on leaf images and using ANNs as a technique for an automatic detection and classification of plant diseases was used with *K*-means as a clustering procedure. ANN consisted of 10 hidden layers. The number of outputs was 6 which was the number of classes representing five diseases along with the case of a healthy leaf.

3 Materials and Methodology

Figure 1 demonstrates the framework which will be used in this work.

3.1 Data Collection and Dataset Preparation: Images can be downloaded from the Internet using the keywords plant and disease names. Subsequently, all the images can be classified into different groups.

3.2 Methods

3.2.1 Image preprocessing and Labelling: Images in the dataset may be in different formats, quality and resolution. Hence, the images need to be preprocessed, for instance, images with smaller resolution and dimension less than 500 px will not considered as valid images for the dataset. The rest will be resized to 256 × 256 in order to reduce the time for training.

3.2.2 Training: In this step, training the deep convolutional neural network for making an image classification model will be done. CaffeNet architecture will be used and adjusted to support our different categories (classes). Rectified Linear Units (ReLU) will subsequently be used as substitute for saturating nonlinearities. This activation function adaptively will learn the parameters of rectifiers and improve accuracy at negligible extra computational cost.

3.2.3 Testing: In this phase, the test set for prediction of leaf as healthy/Unhealthy with its disease name will be used to evaluate the performance of the classifier.

(a) Fine-Tuning: Fine-tuning helps to increase the accuracy of prediction by making small modifications to improve or optimize the outcome. The best suited model for plant disease detection will be achieved through the process of experimental adjustment of the parameters.

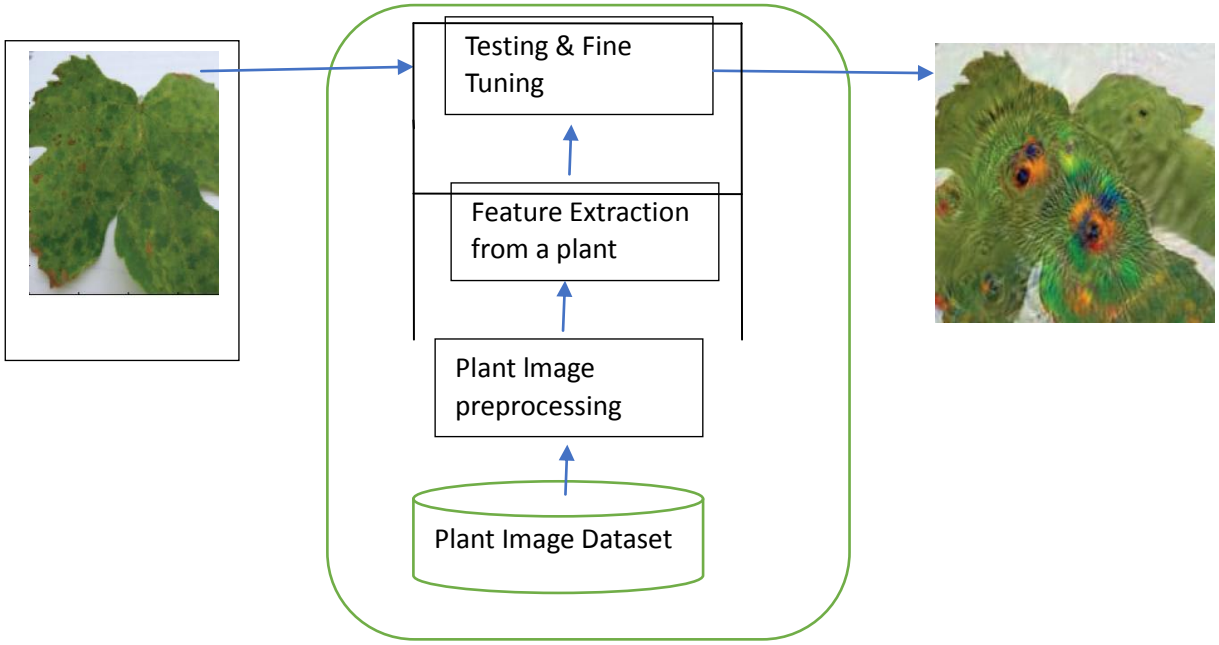

Figure 1: Architecture of Plant Disease Detection System

(b) Evaluation Measures: Measures such as accuracy and mean recall score, mean precision will be computed to evaluate the performance of the classifier.

4 Experimental Design

4.1 Software and Hardware Requirements : Python based Computer Vision and Deep Learning libraries will be exploited for the development and experimentation of the project. Tools such as Anaconda Python, and libraries such as OpenCV, Caffe integrated with CudNN will be utilized for this process. Training will be conducted on NVIDIA GPUs