International Conference on Innovative Research in Engineering, Management and Sciences
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 4 | Issue 9 | ISSN : 2456-3307

Crop Health Monitoring System Using Machine Learning

Mr. Gangadhar Immadi [*1]

[*1] Assistant Professor,Department of ISE, New Horizon College of Engineering, Bengaluru
immadi.gangadhar@gmail.com [1]

ABSTRACT

In this changing environment, appropriate and timely disease identification including early prevention has never been more important. There are several ways to detect plant pathologies. Some diseases do not have any visible symptoms, or the effect becomes noticeable too late to act, and in those situations, a sophisticated analysis is obligatory. However, most diseases generate some kind of manifestation in the visible spectrum, so the naked eye examination of a trained professional is the prime technique adopted in practice for plant disease detection. In order to achieve accurate plant disease diagnostics a plant pathologist should possess good observation skills so that one can identify characteristic symptoms

To find out whether the leaf is diseased or healthy, certain steps must be followed. i.e., Preprocessing, Feature extraction, Training of classifier and Classification. Preprocessing of image, is bringing all the images size to a reduced uniform size. Then comes extracting features of a preprocessed image which is done with the help of HOG .HoG is a feature descriptor used for object detection. In this feature descriptor the appearance of the object and the outline of the image is described by its intensity gradients. One of the advantage of HoG feature extraction is that it operates on the cells created. Any transformations doesn't affect this.

I. INTRODUCTION

Plant disease detection is an Innovative and Enlightening System helping the users to know the disease, trainings or any interesting things taking place around their Area. This Organization aids the native community to keep themselves up to date about the events around their locality or zone or in their town. There are 2 things for this method to work; one for the image processing and another is machine learning.

The user is permitted to sight the disease only of his town while user can supplement disease connected to any town. Admin will show if any misuse or inappropriate or false disease added by any users and will take specific act. The Front end used is Android Studio and backend as SQL Server. The user has to record into the system to using this app and can bring up-to-date his details as well.

The healthy leaf is shown first and so on, the user can also restore the disease resultant the latest one shown first and current disease will be shut. The user can add a image and a title connected to the leaf.

System Architecture

The user can use only 500 words to give a lecture on disease. The Appearance and texture of evaluation the disease is thrilling and amazing as the system offers swipe to move to the next or earlier disease with transition properties.

The agriculturist in provincial regions may think that it's hard to differentiate the malady which may

be available in their harvests. It's not moderate for them to go to agribusiness office and discover what the infection may be. Our principle objective is to distinguish the illness introduce in a plant by watching its morphology by picture handling and machine learning.

Pests and Diseases results in the destruction of crops or part of the plant resulting in decreased food production leading to food insecurity. Also, knowledge about the pest management or control and diseases are less in various less developed countries. Toxic pathogens, poor disease control, drastic climate changes are one of the key factors which arises in dwindled food production.

Various modern technologies have emerged to minimize postharvest processing, to fortify agricultural sustainability and to maximize the productivity. Various Laboratory based approaches such as polymerase chain reaction, gaschromatography, mass spectrometry, thermography and hyper spectral techniques have been employed for disease identification. However, these techniques are not cost effective and are high time consuming.

In recent times, server based and mobile based approach for disease identification has been employed for disease identification. Several factors of these technologies being high resolution camera, high performance processing and extensive built in accessories are the added advantages resulting in automatic disease recognition.

Modern approaches such as machine learning and deep learning algorithm has been employed to increase the recognition rate and the accuracy of the results. Various researches have taken place under the field of machine learning for plant disease detection and diagnosis, such traditional machine learning approach being random forest, artificial neural network, support vector machine(SVM), fuzzy logic, K-means method, Convolutional neural networks etc.

Random forests are as a whole, learning method for classification, regression and other tasks that operate by constructing a forest of the decision trees during the training time. Unlike decision trees, Random forets overcome the disadvantage of over fitting of their training data set and it handles both numeric and categorical data.

The histogram of oriented gradients (HOG) is an element descriptor utilized as a part of PC vision and image processing for the sake of object detection. Here we are making utilization of three component descriptors:

1. Hu moments
2. Haralick texture
3. Color Histogram

Hu moments is basically used to extract the shape of the leaves. Haralick texture is used to get the texture of the leaves and color Histogram is used to represent the distribution of the colors in an image.

The agriculturist in provincial regions may think that it’s hard to differentiate the malady which may be available in their harvests. It's not moderate for them to go to agribusiness office and discover what the infection may be. Our principle objective is to distinguish the illness introduce in a plant by watching its morphology by picture handling and machine learning.

Pests and Diseases results in the destruction of crops or part of the plant resulting in decreased food production leading to food insecurity. Also, knowledge about the pest management or control and diseases are less in various less developed countries. Toxic pathogens, poor disease control, drastic climate changes are one of the key factors which arises in dwindled food production.

Various modern technologies have emerged to minimize postharvest processing, to fortify agricultural sustainability and to maximize the productivity. Various Laboratory based approaches such as polymerase chain reaction, gaschromatography, mass spectrometry, thermography and hyper spectral techniques have been employed for disease identification. However, these techniques are not cost effective and are high time consuming.

In recent times, server based and mobile based approach for disease identification has been

employed for disease identification. Several factors of these technologies being high resolution camera, high performance processing and extensive built in accessories are the added advantages resulting in automatic disease recognition

Plant disease detection is a portable system which runs under automaton stage which is established with native Java language. App will run smoothly and very fast.

Google Material Design is implemented because of which app looks more attractive and beautiful with good user impression and experience. Combined with admin and user it manages news, various category, notification and many things whenever we want. Main preference is good code with good design.

We can save money and time by using this app and we can create our own design and special apps or different kind of app based on our requirements.Plant disease detection is an Innovative and Enlightening System helping the users to know the disease, trainings or any interesting things taking place around their Area. This Organization aids the native community to keep themselves up to date about the events around their locality or zone or in their town.

There are 2 things for this method to work; one for the image processing and another is machine learning. The user is permitted to sight the disease only of his town while user can supplement disease connected to any town. Admin will show if any misuse or inappropriate or false disease added by any users and will take specific act. The Front end used is Android Studio and backend as SQL Server. The user has to record into the system to using this app and can bring up-to-date his details as well.

The healthy leaf is shown first and so on, the user can also restore the disease resultant the latest one shown first and current disease will be shut. The user can add a image and a title connected to the leaf.

II. RELATED STUDY

To find out whether the leaf is diseased or healthy, certain steps must be followed. i.e., Preprocessing, Feature extraction, Training of classifier and Classification. Preprocessing of image, is bringing all the images size to a reduced uniform size. Then comes extracting features of a preprocessed image which is done with the help of HOG .

HOG is a feature descriptor used for object detection. In this feature descriptor the appearance of the object and the outline of the image is described by its intensity gradients. One of the advantage of HoG feature extraction is that it operates on the cells created. Any transformations doesn't affect this.

Innovative and useful system helps the users to know the disease, traineeships or any motivating things happening around the zone. This system also helps to preserve ourselves up to date about what is trendy around the town. The user is allowed to see the disease, can input many disease connected to any place. Admin will look after if any misuse or unrelated or bogus disease is updated by any members and will take some act. The technique that is used is machine learning and image processing. The user has to record into the system to using this app and can bring up-to-date his details as well.

The up-to-date disease is shown first and so on, the user can also restore the disease resultant the latest one shown first and current disease will be shut. The user can add a image and a title connected to the news.

The user can use only 500 words to give a lecture on news. The Appearance and texture of evaluation the disease is thrilling and amazing as the system offers swipe to move to the next or earlier with transition.

III. SCOPE

In short, we can read day-to-day disease hunt, immediately it saves our time and we can keep ourselves informed on latest disease that are affecting the leaves in day to day life in short. This single project serves many users to view several disease.

To find out whether the leaf is diseased or healthy, certain steps must be followed. i.e., Preprocessing,

Feature extraction, Training of classifier and Classification. Preprocessing of image, is bringing all the images size to a reduced uniform size. Then comes extracting features of a preprocessed image which is done with the help of HOG .HoG is a feature descriptor used for object detection. In this feature descriptor the appearance of the object and the outline of the image is described by its intensity gradients. One of the advantage of HoG feature extraction is that it operates on the cells created. Any transformations doesn't affect this.

Here we made use of three feature descriptors.

Hu moments: Image moments which have the important characteristics of the image pixels helps in describing the objects. Here Hu moments help in describing the outline of a particular leaf. Hu moments are calculated over single channel only. The first step involves converting RGB to Gray scale and then the Hu moments are calculated. This step gives an array of shape descriptors.

Haralick Texture: Usually the healthy leaves and diseased leaves have different textures. Here we use Haralick texture feature to distinguish between the textures of healthy and diseased leaf. It is based on the adjacency matrix which stores the position of (I,J). Texture is calculated based on the frequency of the pixel I occupying the position next to pixel J. To calculate Haralick texture it is required that the image be converted to gray scale.

Color Histogram: Color histogram gives the representation of the colors in the image. RGB is first converted to HSV color space and the histogram is calculated for the same. It is needed to convert the RGB image to HSV since HSV model aligns closely with how human eye discerns the colors in an image. Histogram plot [8] provides the description about the number of pixels available in the given color range

IV. PROBLEM DEFINITION

The problem of efficient plant disease protection is closely related to the problems of sustainable agriculture and climate change. Research results indicate that climate change can alter stages and rates of pathogen development; it can also modify host resistance, which leads to physiological changes of host-pathogen interactions. The situation is further complicated by the fact that, today, diseases are transferred globally more easily than ever before. New diseases can occur in places where they were previously unidentified and, inherently, where there is no local expertise to combat them .

Inexperienced pesticide usage can cause the development of long-term resistance of the pathogens, severely reducing the ability to fight back. Timely and accurate diagnosis of plant diseases is one of the pillars of precision agriculture

Shortened disease extracts from plants, every story to be printed as a short disease for the viewers. The viewers can luckily flip through the short sections as they want. Share disease and exciting sections with friends and family. The user is simplified with everything going around his town. Easy to use. The user can shot to the admin if the disease is not applicable or honest.

- Because of winter season it is difficult to protect mo leaves.
- The user can't sight the disease of different towns.
- User can add only one image per plant

V. MODULE DESCRIPTION

Image processing

It is used by some researchers for detection and classification of wheat leaf disease. The steps of leaf diseases detection and classification are given below:

1. Image acquisition.
2. Preprocessing.
3. Segmentation of infected regions.
4. Feature extraction
5. Feature selection

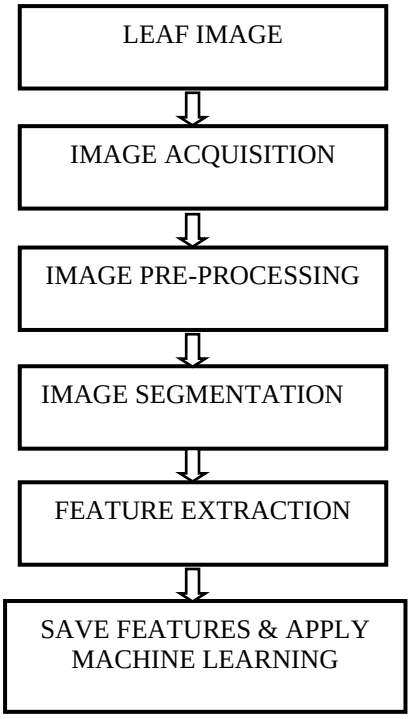


Figure : Block Diagram of Plant Disease Detection

1. Image Acquisition-

The first stage of any vision system is the image acquisition stage. After the image has been obtained, various methods of processing can be applied to the image to perform the many different vision tasks required today. However, if the image has not been acquired satisfactorily then the intended tasks may not be achievable, even with the aid of some form of image enhancement.

2. Image Pre-Processing

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image Processing system includes treating images as two dimensional signals while applying already set signal processing methods to them.

It is among rapidly growing technologies today, with its applications in various aspects of a business. Image Processing forms core research area within engineering and computer science disciplines too.

Image processing basically includes the following three steps:

- Importing the image with optical scanner or by digital photography.
- Analyzing and manipulating the image which includes data compression and image enhancement and spotting patterns that are not to human eyes like satellite photographs.
- Output is the last stage in which result can be altered image or report that is based on image analysis.

3. Image Segmentation-

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super-pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s).

4. Feature Extraction-

In machine learning, pattern recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is a dimensionality reduction process, where an initial

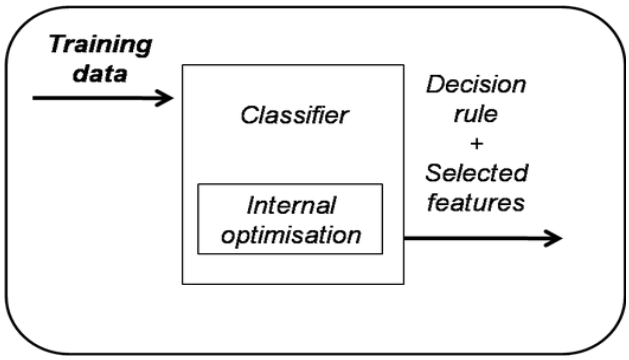
set of raw variables is reduced to more manageable groups (features) for processing, while still accurately and completely describing the original data set.

When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a reduced set of features (also named a feature vector). Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.

5. Feature Selection-

In machine learning and statistics, feature selection, also known as variable selection, attribute selection or variable subset selection, is the process of selecting a subset of relevant features (variables, predictors) for use in model construction. Feature selection techniques are used for four reasons:

- simplification of models to make them easier to interpret by researchers/users,
- shorter training times,
- to avoid the curse of dimensionality,
- enhanced generalization by reducing overfitting (formally, reduction of variance)



6. Machine Learning Algorithm-

SUPPORT VECTOR MACHINE (SVM)-

They were extremely popular around the time they were developed in the 1990s and continue to be the go-to method for a high-performing algorithm with little tuning. In machine learning, support vector machine are a set of supervised learning models with associated learning algorithms that analyses data used for Classification and regression analysis. When data are not labeled supervised learning is not possible. It constructs a hyper lane and a set of hyper lane which in a high and infinite dimensional space, which can be used for other task like outlier detection.

ADVANTAGES-

- Effective in high dimensional spaces.
- Still effective in cases where number of dimensions is greater than the number of samples.
- It is also memory efficient.
- Versatile.
- Support vector machine (SVM) it analyses the data after that it classify that data and then the regression is done.
- Two main parameters accuracy and detection time in these two the disease is detect.
- Support vector machines (SVM) increase the recognition rate

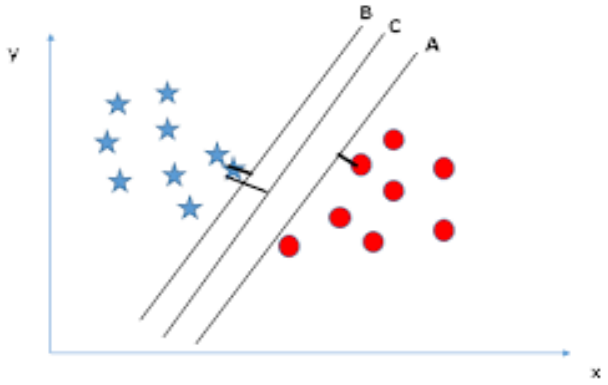


Figure : SVM Model

NEURAL NETWORKS

The disease detection techniques are based on two main neural network technologies: Self-Organising map and multilayer perceptrons (the most widely used architecture).Neural network is a system of hardware and/or software patterned after the operation of neurons in the human brain. Neural network also called artificial neural network.

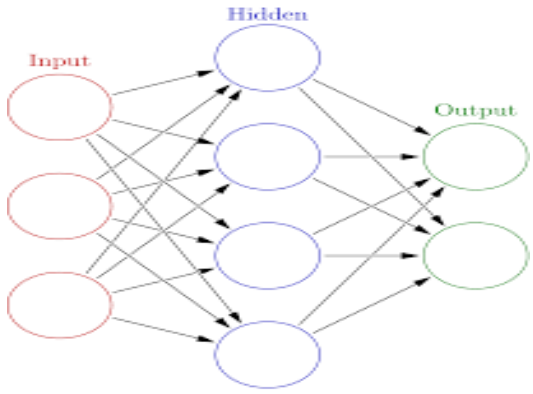


Figure: Neural Network

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyse.

ADVANTAGES-

- Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.
- Self-Organisation: An ANN can create its own organisation or representation of the information it receives during learning time.
- Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
- Fault Tolerance via Redundant Information Coding: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.

Random Forest Classifier-

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of over fitting to their training set.

Random forest classifier creates a set of decision trees from randomly selected subset of training set. It then aggregates the votes from different decision trees to decide the final class of the test object.

The advantages of random forest are:

- It is one of the most accurate learning algorithms available. For many data sets, it produces a highly accurate classifier.
- It runs efficiently on large databases.
- It can handle thousands of input variables without variable deletion.
- It gives estimates of what variables are important in the classification.
- It generates an internal unbiased estimate of the generalization error as the forest building progresses.

It has an effective method for estimating missing data and maintains accuracy when a large proportion of the data are missing.

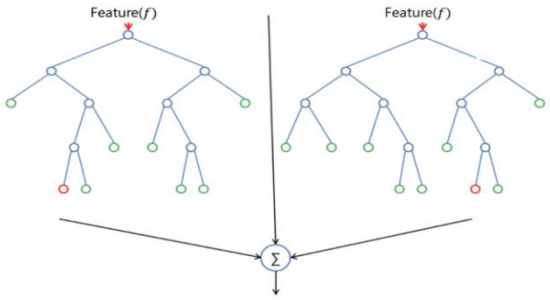


Figure : Random Forest Classifier

HOG (Histogram of Oriented Gradients)-

The histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms, scale-invariant feature transform descriptors, and shape contexts, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

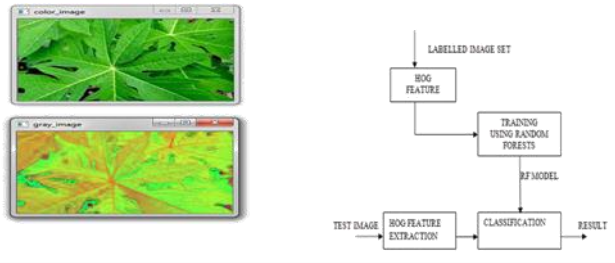
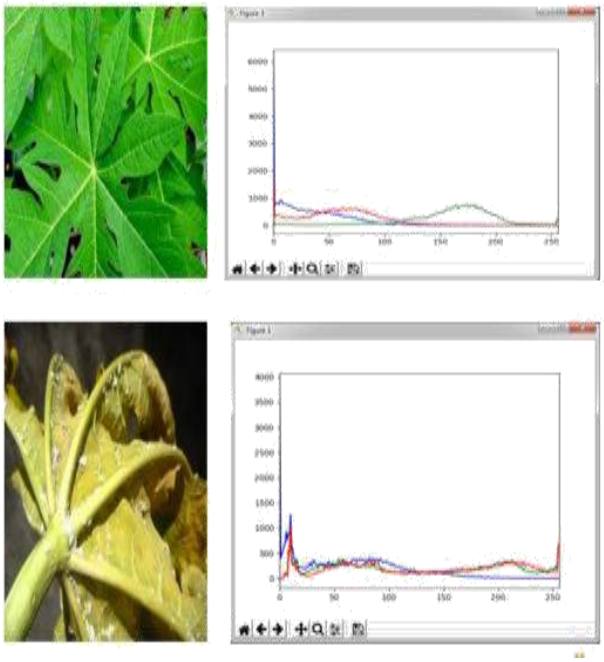
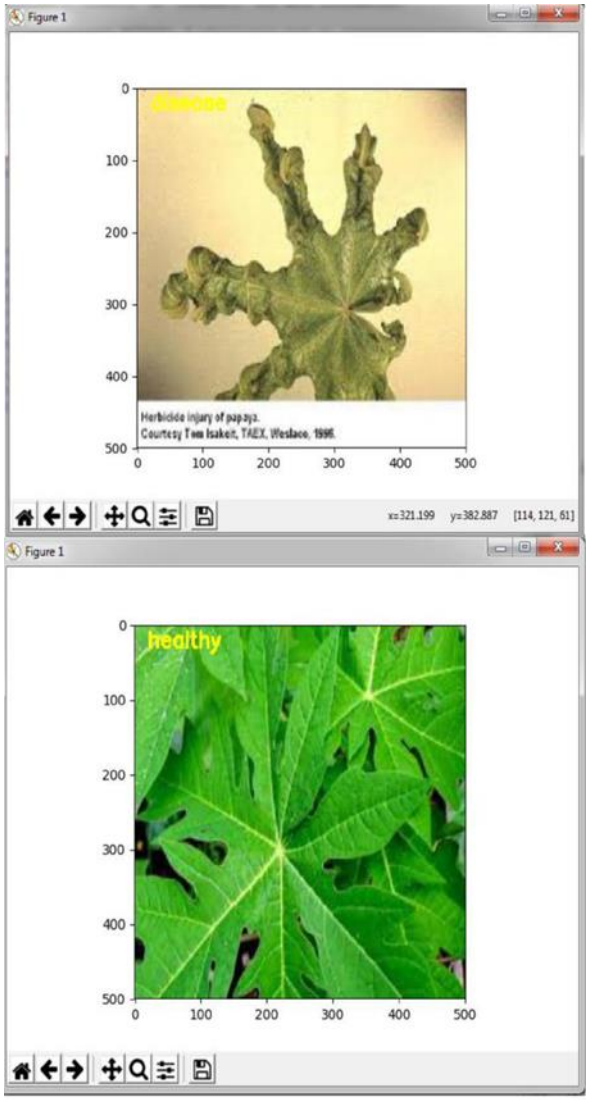


Figure: HOG Diagram

HoG is a feature descriptor used for object detection. In this feature descriptor the appearance of the object and the outline of the image is described by its intensity gradients. One of the advantage of HoG feature extraction is that it operates on the cells created. Any transformation doesn't affect this.

VI. EXPERIMENTAL RESULTS





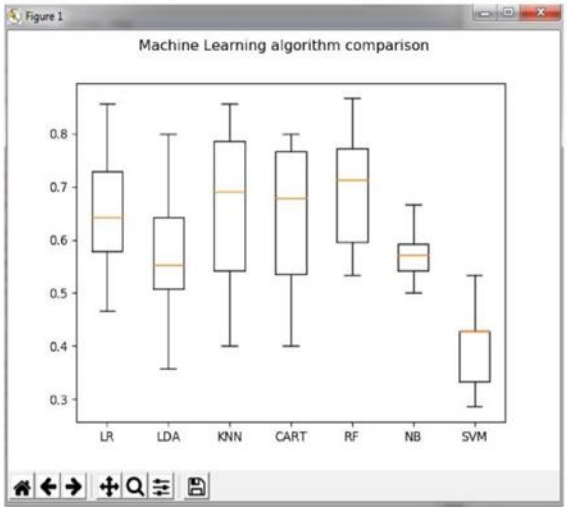


Figure: Final output of the classifier & Comparison with other models

VII. CONCLUSION

Overall idea is to detect whether it is diseased or healthy leaf with the help of a Random forest classifier. The objective of this algorithm is to recognize abnormalities that occur on plants in their greenhouses or natural environment. The image captured is usually taken with a plain background to eliminate occlusion. The algorithm was contrasted with other machine learning models for accuracy. Using Random forest classifier, the model was trained using 160 images of papaya leaves. The model could classify with approximate 70 percent accuracy. The accuracy can be increased when trained with vast number of images and by using other local features together with the global features such as SIFT (Scale Invariant Feature Transform), SURF (Speed Up Robust Features) and DENSE along with BOVW (Bag Of Visual Word).

There are many methods in automated or computer vision plant disease detection and classification process, but still, this research field is lacking. In addition, there are still no commercial solutions on the market, except those dealing with plant species recognition based on the leaves images.

Future enhancement for this is to implement a cloud storage in order to consists of the results of plant disease defect detection which has to be sent to the farmers so as to they can use the right fertilizers for that particular disease.