International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 05 | May -2020 ISSN: 2582-3930

LEAF IDENTIFICATION USING IMAGE PROCESSING

Prashant Kumar Upadhyay (1603210134),
Ankit Sharma (1703210903),
Rajiv Ranjan (1703210908)
SubmittedTo- Mr. Anurag Mishra

Computer Science & Engineering Department

ABES Engineering College, Ghaziabad, U.P. India

Abstract

Earth may be a solely living planet exist within the system, that encompass system.Plants plays an important role in system, that helps living being (like, human, animal etc.) to survive on the planet. Leaf is that the vital a part of the plant that helps within the survival of the plant itself. Plants give food for living beings. Leaf identification and classification helps to acknowledge the species of plants that helps plants to survive within the nature. Leaf form and size are major feature to spot the species of plant.So we have applied various computer vision, methodologies and pattern recognition techniques towards automated procedure of the plant recognition.

Identification and classification of leaves is completed with many methodologies.During this paper, we tend to propose the strategies to identify the leaf and spot the leaf victimization which is a picture analysis-based approach.

Introduction

On earth there are number of plant species, which provides food and shelter for living beings. Without plants there will be no existence of environment on earth.The delay and quality of identification plant diseases is inflicting important reduction in each quality and amount of agricultural product. for instance, it's calculable that total losses are amounting to roughly 12 percent of the manufacture. So various computer vision, methodologies and pattern recognition technique have been applied towards automated procedure of the plant recognition.Automatic detection of plant and connected diseases based on a leaf image would be terribly useful for the farming world and it'll speed up readying of

remedy quickly to scale back or eliminate harm from the disease.

The input to our implementation is a picture of a diseased leaf along with the healthy and diseased portions. The output is the name of the disease that is affecting the leaf. In this project we evaluate several machine learning techniques to

i. Identify the diseased area (We used K-Means and Gaussian Mixture) and
ii. Identify the disease (We used Linear SVM, Quadratic SVM, K-Means) by classifying among different classes of diseases.

The leaf identification project is made for the farmers and agricultural scientists. The project has following scopes:

- Agricultural Field.
- Farming Related Exposure
- Medicine Research for plants

We have studied various research papers which are stated in reference section. By studying these research papers, we came to understand a lot about the processes involved in the image processing system for detection and identification of leaf. We also got idea about the steps involved in the different processes to perform it successfully. We studied about the various techniques for generating the results more efficiently and getting better results.

We got to know about the various main features (i.e., shape, color and texture) and also the various sub features under these main features which are required for the classification of the different types of leaf.

We studied about some of the different types of diseases in the leaf of various plants like blight, leaf spots, mildew and rust. Also learned how to cluster them using k-means clustering technique to classify them efficiently for better results.

We studied the binarization of image in which we can take the image in only two basic colors i.e., black and white. By doing this we can easily marks the area of the leaf and cutout its shape for identification by setting up threshold which is different for different types of leaves.

We learnt the method of feature extraction for leaf architecture to takeout the digital and morphological features of selected leaves which are further useful for plant identification and classification

We used the methods of computer vision to do feature extraction of the various medicinal plants. We came along ways to make automated systems to do feature extraction for identifying the various plant species more accurately and efficiently.

Related Previous Work

Leaf shape is a prominent feature to recognize and classify a plant stated that diameter, physiological length, physiological width, leaf area and perimeter

are basic geometry information can be extract from the leaf shape. In addition, leaf color, textures and vein are also considered as features. Cross validation is used to compare learning algorithms: MATLAB is used to compare results across means, Linear Discriminative Analysis, Linear SVM and Quadratic SVM. The optimized feature selection gave a 2% improvement over the previous set of results. It works for the best performing model (Quadratic SVM with K-means background separation with K=4), chosen with optimized feature set, the accuracy was 93.1%. The Precision Recall and F1 scores for each classifier are within 85% to 100%.

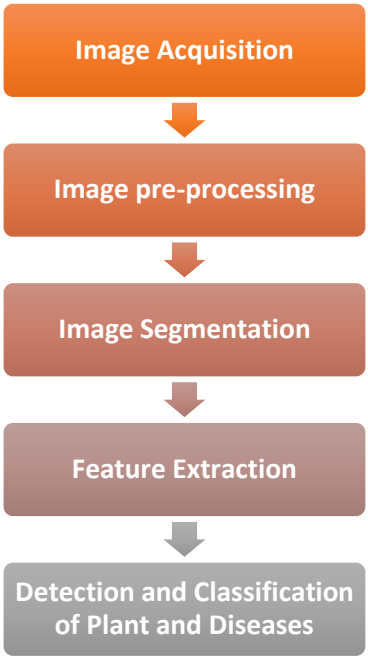
Model

We are following a methodology containing the three main basic steps. They are as follows:

i. Pre-Processing
ii. Feature Extraction
iii. Model Building and Testing

These three processes contain several sub processes and techniques to perform the image processing and identification of leaf correctly and get out final output.

The Following flow chart shows the flow of the process in the Leaf Identification System.



The following figure show the data flow of the proposed method.

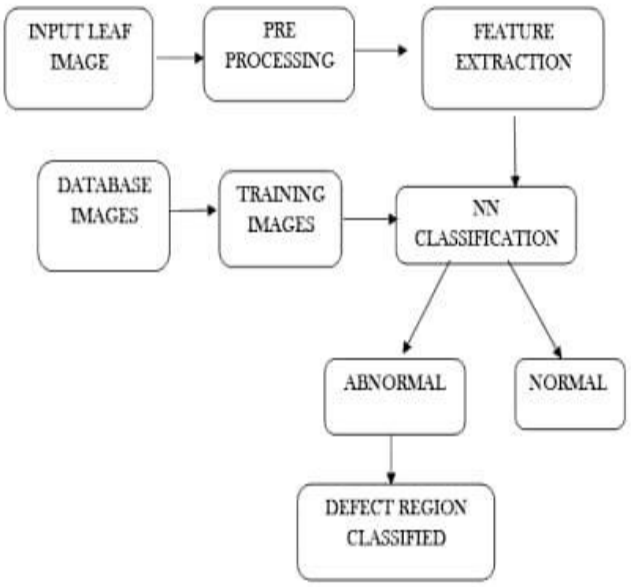


Fig-1 Proposed Method

Implementation

Software requirements for the Leaf Identification System are:

i. Python
ii. Anaconda
iii. Jupyter Notebook
iv. Browser - Chrome preferable (latest)
v. Operating System – Linux/Windows 7 or above

Hardware requirements for the Leaf Identification System are:

a. A good Laptop or Desktop with following specifications:
   i. 8 GB RAM
   ii. Intel core-i5 or above Processor
b. Digital Storage space
c. A Camera - for taking pictures of leaf (Optional)

We have following Assumptions:

I. We assume that the above requirements of software and hardware are met. There should not be any error in the operating system.
II. All the settings should be compatible and the status of the settings synchronization should be done beforehand.
III. We also assume that the leaf image should be of a fresh leaf and not a withered, crushed or damaged leaf.
IV. The leaf image should be clear.
V. We have used plant leaf datasets for the reference of the training and testing the data for images of different types of leaves and plants. These Datasets helped our model to get different features of the different types of leaves.

The datasets used are:

- Flavia leaves dataset - http://flavia.sourceforge.net
- Leaf Diseases UCI - https://archive.ics.uci.edu/ml/datasets/Leaf
- Plant Village – www.kaggle.com

Image Acquisition

In this part we collect the datasets from various sources. In this we have to take only relevant data and discard all the other parts of the data.

Importing datasets and loading image

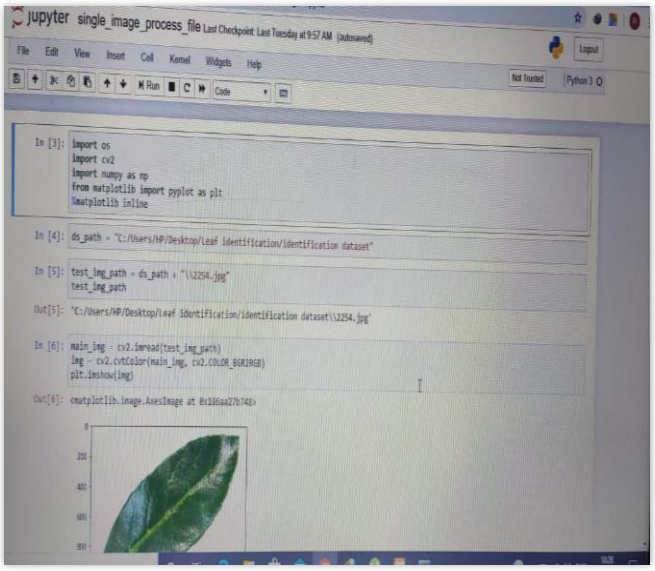


Image Pre-Processing

It is very important process in image processing which should be done in the starting after the image is collected. It helps to reduce the noise from the image and make it smoother which helps in enhancing the features of the image. In this we convert the RGB image into grayscale image and then from grayscale to binary after which the smoothening and filtering is done.

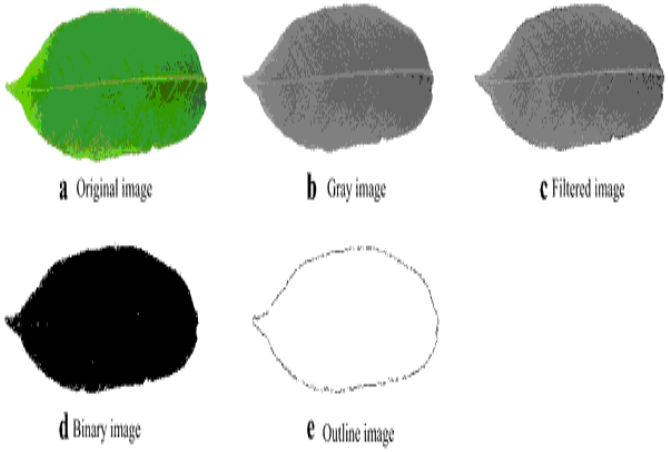


Fig-2 Image Pre-Processing

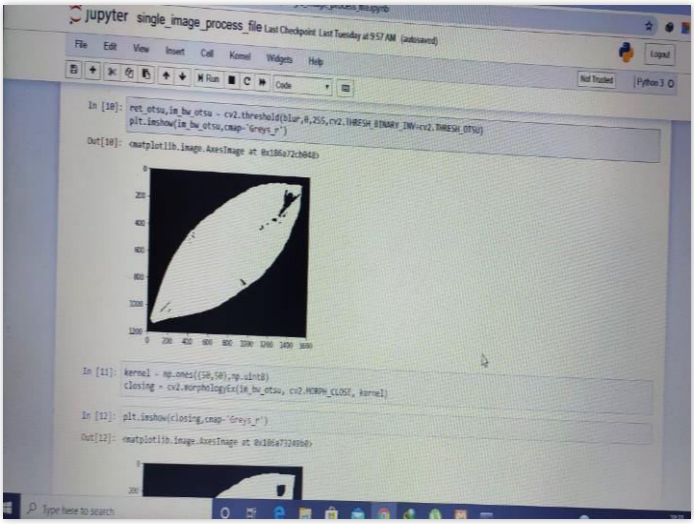
Image Segmentation

It is a very critical step in image analysis, which is basically done for extracting the region of interest (ROI). In this the each pixel of image is segmented or grouped in different part or segments on the basis of their color, texture, gray level, intensity of pixel. Here we have used k-means clustering technique for segmentation.

Feature Extraction

It is one of the most important steps in image processing and pattern analysis. In this various chosen (right) set of features or pixels are taken reducing the number of variables to perform computation to increase efficiency.

Feature Extraction



Some main machinelearning techniques used in are:

- K-means clustering
  - K-means clustering is a machine learning technique in which, the n points of observation are partitioned

into k clusters or groups by observing their nearest mean clusters or centroids. It is very helpful in image segmentation.

- Otsu Thresholding method
    - It is an algorithm which was given by Nobuyuki Otsu, which is used for getting a single intensity threshold to separate the image pixels into two basic classes i.e., foreground and background. It is an efficient techniques of image binarization.
- SVM classifier
    - Support Vector Machine is a very useful supervised learning technique for image processing. In this the SVM works to analyse the data or image and create linear decision boundaries to classify the image in multiple classes.

Results

Our Leaf Identification System was able to detect and identify the leaf and its disease or not with very high accuracy.

The result is shown in the form of a graph which describes the accuracy of the image identification of the leaf with the help of the project model. The results are compared with measurements of grid count method. Experiments were carried upon 70 leaves of different species.

The accuracy resulted to approximately 96%-97% which is better than other such systems present currently.

It will be very helpful to the farmers to identify the disease and then they can apply the correct remedy for it.

Conclusion

The disease detection feature of this project is highly useful as it predicts the disease instantly without the need of visiting any testing lab.This Leaf Identification System can be used as an ideal farming assistant platform as there are very less such softwares which are beneficial for the farmers.

Our project provides more accurate results then other such projects or models at the same level present currently in the market.

The leaf identification software in much efficient as compared to other such softwares as it is lite weight.

This can also help in Minimum Support Price prediction which can provide a surety to the farmers before or during plantation of their crops.

Future to this it is needed to compute disease severity percentage on plant leaf and integrate both results for actual prediction of yield loss and plant growth.

The disease detection feature of this project can be enhanced with more knowledge database about

remedies for different plant diseases and in future it can be used to suggest the suitable medicines for the particular disease in plants.

References

Research Papers/Journals

i. Ab Jabal, Mohamad Faizal. (2013). LEAF FEATURES EXTRACTION AND RECOGNITION APPROACHES TO CLASSIFY PLANT. Journal of Computer Science. 9. 1295-1304. 10.3844/jcssp.2013.1295.1304.
ii. Farm: Leaf based disease identification in farms CS 229 Project Report Fall 2015 (Authors: Kushal Chawda, Chanchal Hazra)
iii. Image Binarization using Otsu Threshold Algorithm (Author: Jamileh Yousefi) Published in the year 2011
iv. A Method of Feature Extraction from Leaf Architecture (Authors: V. R. Patil, R. R. Manza) Published in the year 2015
v. Computer Vision Based Feature Extraction of Leaves for Identification of Medicinal Values of Plants (Authors: D Venkataraman, Mangayarkarasi N), Published in the year 2016

Sites

- www.google.com
- www.semanticscholar.org
- www.researchgate.net
- https://en.wikipedia.org
- http://www.jmis.org/archive/view_article?pid=jmis-6-2-49