CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

Pictures are the most common and convenient means of conveying or transmitting information. A picture is worth a thousand words. Pictures concisely convey information about positions, sizes and inter relationship between objects. They portray spatial information that we can recognize as object. Human beings are good at deriving information from images, because of our innate visual and mental abilities. About 75% of the information received by human is in pictorial form.

1.2 IMAGE 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.

1.2.1 PURPOSE OF IMAGE PROCESSING

The purpose of image processing is divided into 5 groups. They are:

- Visualization - Observe the objects that are not visible
- Image sharpening and restoration - To create a better image
- Image retrieval - Seek for the image of interest
- Measurement of pattern – Measures various objects in an image
- Image Recognition – Distinguish the objects in an image

1.2.2 TYPES OF IMAGE PROCESSING

The two types of methods used for Image Processing are Analog and Digital Image Processing. Analog or visual techniques of image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. The image processing is not just confined to area that has to be studied but on knowledge of analyst.

Association is another important tool in image processing through visual techniques. So analysts apply a combination of personal knowledge and collateral data to image processing.

1.2.3 APPLICATIONS

1. **Intelligent Transportation Systems** – This technique can be used in Automatic number plate recognition and Traffic sign recognition.

2. **Remote Sensing** – For this application, sensors capture the pictures of the earth's surface in remote sensing satellites or multi – spectral scanner which is mounted on an aircraft. These pictures are processed by transmitting it to the Earth station. Techniques used to interpret the objects and regions are used in flood control, city planning, resource mobilization, agricultural production monitoring, etc.

3. **Defense surveillance** – Aerial surveillance methods are used to continuously keep an eye on the land and oceans. This application is also used to locate the types and formation of naval vessels of the ocean surface. The important duty is to divide the various objects present in the water body part of the image. The different parameters such as length, breadth, area, perimeter, compactness are set up to classify each of divided objects. It is important to recognize the distribution of these objects in different directions that are east, west, north, south, northeast, northwest, southeast and south west to explain all possible formations of the vessels. We can interpret the entire oceanic scenario from the spatial distribution of these objects.

1.2.4. BENEFITS OF IMAGE PROCESSING

- Important features such as edges can be extracted from images.
- Images can be sharpness and better visual appearance.
- Minor errors can be rectified.
- Images sizes can be increased or decreased.

1.3 DIGITAL IMAGE PROCESSING

The analysis of a picture using techniques that can identify shades, colours and relationships that cannot be perceived by the human eye. Image processing is used to solve identification problems, such as in forensic medicine or in creating weather maps from satellite pictures. It deals with images in bitmapped graphics format that have been scanned in or captured with digital cameras. An image is nothing more than a two dimensional signal. It is defined by the mathematical function f(x,y) where x and y are the two co-ordinate horizontally and vertically.The value of f(x, y) at any point is gives the pixel value at that point of an image.

Figure 1.3.a Sample Image

The following figure is an example of digital image that you are now viewing on your computer screen. But actually, this image is nothing but a two

128	30	123
232	123	321
123	77	89
80	255	255

dimensional array of numbers ranging between 0 and 255.

Each number represents the value of the function f(x, y) at any point. In this case the value 128, 230, 123 each represents an individual pixel value. The dimensions of the picture is actually the dimensions of this two dimensional array.

An image may be defined as a two dimensional function f(x, y),Where x and y are spatial Co-ordinates, and the amplitude of 'f' at any pair of Co-ordinates (x, y) is called the intensity or grey level of the image that point. When x, y and the amplitude values of 'f' are all finite, discrete quantities. We call the image a digital image.

A digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements, image elements and pixels.

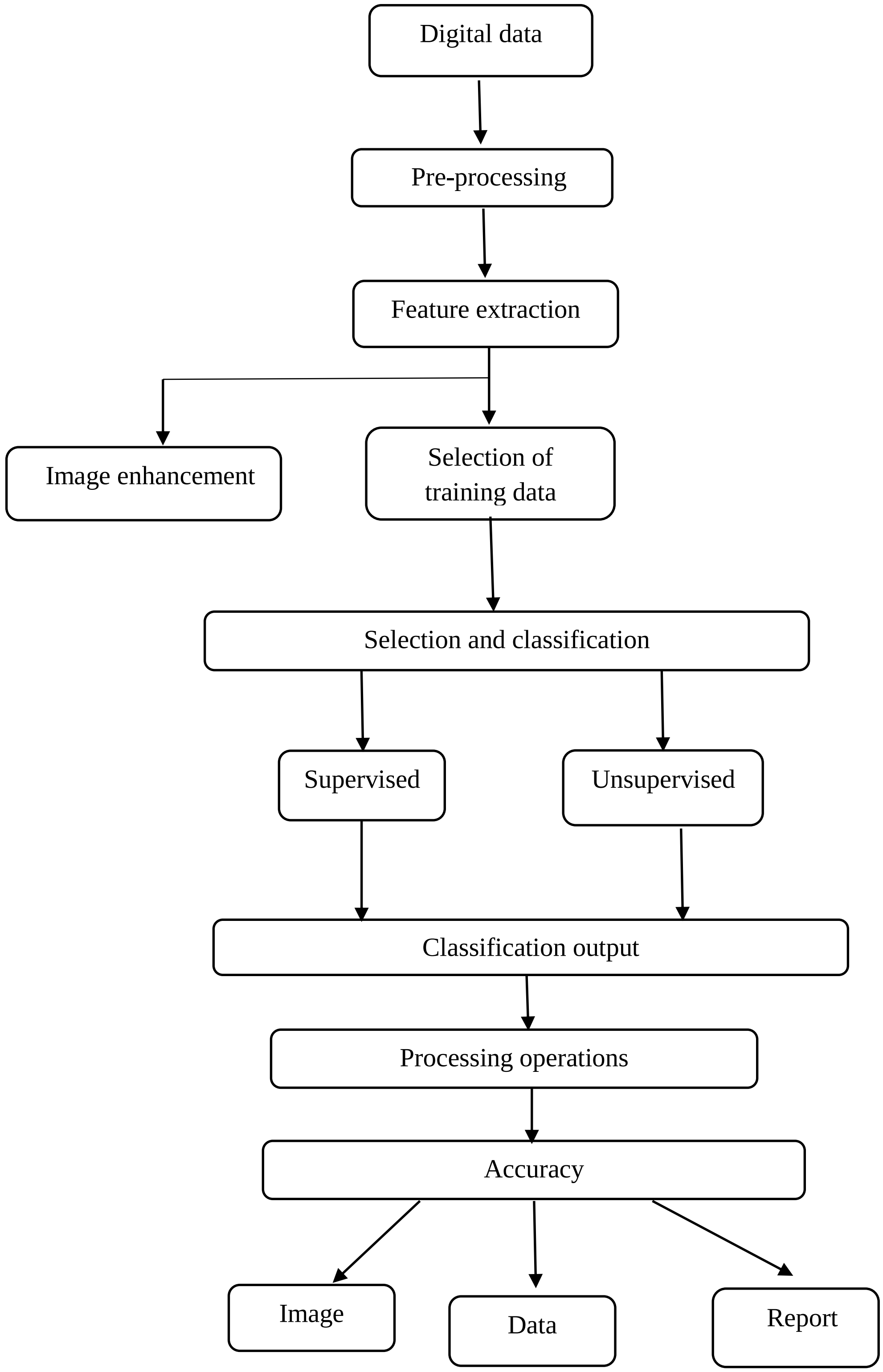


Figure 1.3.b Digital image processing

Examples of field that use Digital Image Processing

1. Gamma-Ray imaging
2. X-Ray Imaging
3. Imaging in the Ultraviolet Band
4. Imaging in the Visible and Infrared Bands
5. Imaging in the Microwave Band
6. Imaging in the Radio Band

1.3.1 IMAGE ENHANCEMENT

It is the process of manipulating an image so that the result is more suitable than the original for a specific application. The word specific is important here, because it establishes at the outset that enhancement techniques are problem oriented. Thus for example, a method that is quite useful for enhancing X-ray images may not be the best approach for enhancing the satellite images taken in the infrared band of the electromagnetic spectrum.

1.3.2 IMAGE RESTORATION

The principal goal of restoration techniques is to improve an image in some predefined sense. Image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of degradation.

Although there are areas of overlap, image enhancement is largely a subjective process, while image restoration is for the most part an objective process.

Restoration attempts to recover an image that has been degraded by using a prior knowledge of the degradation phenomenon.

Thus restoration techniques are oriented toward modelling the degradation and applying the inverse process in order to recover the original image.

1.3.3 IMAGE COMPRESSION

It refers to the process of reducing the amount of data required to represent a given quantity of information. In this definition, data and information are not the same thing. Data are the means by which information is conveyed.

Because various amounts of data can be used to represent the same amount of information, representations that contain irrelevant or repeated information are said to contain redundant data. It is one of the most useful and commercially successful technologies in the field of digital image processing. The number of images that are compressed and decompressed daily is staggering, and the compressions and decompressions themselves are virtually invisible to the user. The objective is to reduce redundancy of the image data in order to be able to store or transmit data in an efficient form.

Image compression can be lossy or lossless. Lossless compression is sometimes preferred for artificial images such as technical drawings, icons or comics. This is because lossy compression methods, especially when used at low bit rates, introduce compression artifacts. Lossless compression methods may also be preferred for high value content, such as medical imagery or image scans made for archival purposes. Lossy methods are especially suitable for natural images such as photos in applications where minor (sometimes imperceptible) loss of fidelity is acceptable to achieve a substantial reduction in bit rate.

1.3.4 IMAGE SEGMENTATION

Segmentation is the first step from low-level image processing transforming a greyscale or colour image into one or more other images to high-level image description in terms of features, objects, and scenes. The success of image analysis depends on reliability of segmentation, but an accurate partitioning of an image is generally a very challenging problem.

This techniques are either contextual or non-contextual. The latter take no account of spatial relationships between features in an image and group pixels together on the basis of some global attribute, e.g. grey level or colour.

Contextual techniques additionally exploit these relationships, e.g. group together pixels with similar grey levels and close spatial locations.

It refers to the process of partitioning a digital image into multiple segments (sets of 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 visual 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 (see edge detection). 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).

Some of the practical applications of image segmentation are:

- Locate tumors and other pathologies
- Measure tissue volumes
- Computer-guided surgery
- Diagnosis
- Treatment
- Planning
- Study of anatomical structure

1.3.5 IMAGE RECONSTRUCTION

The reconstructing an image from a series of projections, with a focus on X-ray computed tomography. This is the earliest and still the most widely used type of CT and is currently one of the principal applications of digital image processing in medicine. It is simple in principle and it can be explained quality in a straightforward, intuitive manner.

CHAPTER 2

LITERATURE REVIEW

[1] M. K. Alsmadin, K. B. Omar, S. A. Noah, and I. Almarashdeh, “Fish recognition based on robust features extraction from colour texture measurements using back-propagation classifier”, Presented the development of a system that can recognize isolated pattern of interest on image based on the combination of robust features extraction.

The system was used as classifier for fish image recognition and able to classify and categorize fish into poison or non-poison family. C. Real Life Application of Image Processing in the Fields of Agricultural Research Several attempts have been made to build automatic insect identification systems based on image analysis.

[2] S. M. Al-Saqer, “A robust recognition system for pecan weevil using artificial neural networks”, they used image descriptors as input in the neural network to recognize the pecan weevil. The authors collected different images of pecan weevil and other insects found in the paddy fields. After collecting the images they converted them into binary and resized to 114×134 pixels. To process the data, they used different image processing techniques such as Regional Properties and Zernike Moments.

Then segmentation is used to partition the image in to multiple regions. For these particular type of edge detection sobel operator is used. The drawback of this system is that if the colour of pest and leaf is almost similar then the background subtraction method cannot identify the pest.

The color change on the leaf is also identifying as pest. The accuracy of pest identification depends on the selection of threshold value.

[3] F. A. Carino, P. E. Kenmore, and V. A. Dyck, “A FARMCOP suction sampler for hoppers and predators in flooded rice fields”, The light trap, that involves varying size sample which is good for comparing seasonal and yearly catches of insects, but catches are subject to changes in insect behaviour and do not catch none flying insects; the sweep net (catching insect using fishnet), is a fast method, very economical, and good for sampling arthropods staying in canopy of rice, but it has human error due to variability and poor catch of arthropods at the base of the plant; tapping the rice, this is a sampling method that utilize a collecting pan with soap solution or oil with water to collect arthropods at the base and stem of the rice.

After tapping, arthropods are identified and counted immediately in the field; The visual counting and data recording can be done on field but also subject to human error and very costly in labour; sticky trap is also economical, it measures insect movement and colonization but does not catch non flying insect; yellow pan trap is also economical; it measures insect migration, easy sorting and counting of samples, but the attraction is due to colour stimulus and does not catch none flying insects. B. Feature Extraction and Image Processing Image processing is the analysis and manipulation of graphical images from sources such as photographs and videos.

[4] Gaurav Kandalkar, A.V.Deorankar and P.N.Chatur, "Identification of Agricultural Pests Using Radial Basis Function Neural Networks ", The proposed decision support system constitutes of image segmentation, feature extraction, training and testing neural network. The segmentation will be carried out for

extraction of pest from the image. Thresholding techniques are used for segmentation. After segmentation the feature extraction is done.

Classification is performed to identify the pest. Classification process done by RBF neural network. RBF neural network is an artificial neural network that uses radial basis function as activation function. The RBF is easy to train but is slower in identifying the real time images.

[5] Ganesh Bhadane, Sapana Sharma and Vijay B. Nerkar,"Early Pest Identification in Agricultural Crops using Image Processing Techniques", Images of the infected leaf are captured by digital camera and processed using image growing, image segmentation techniques to detect infected parts of the particular plants. Then the detected part is been processed for further feature extraction which gives general idea about pests.

[6] Johnny L. Miranda, Bobby D. Gerardo and Bartolome T. Tanguilig III, "Pest Identification using Image Processing Techniques in Detecting Image Pattern through Neural Network", In this system for object detection uses background subtraction method is used. Then the feature extraction and segmentation performed. To identify the insect Kohonen Self Organizing Maps neural network is used. Kohonen Self Organizing Maps neural network requires necessary and sufficient data in order to develop meaningful clusters.

The weight vectors must be based on data that can successfully group and distinguish inputs. Lack of data or extraneous data in the weight vectors will add randomness to the groupings.

[7] Rupesh G. Mundada and Dr. V. V. Gohokar,"Detection and Classification of Pests in Greenhouse Using Image Processing", In this, feature extraction was done with the help of Support Vector Machines (SVM's).

SVM is a supervised learning method. SVM classification first finds a hyper plane which separates the d-dimensional data perfectly into two classes. This is done by maximizing the margin from the hyper plane. Hyper plane is defined as, Yi = + c (1) Where w is n-dim vector and c is a constant or scalar. By adding a scalar value c it increases the margin between hyper planes and in the absence of c, hyper plane is forced to pass through the origin. Multiclass classification is also recommended and used two class SVMs to solve the problem 1) one-versus-all or one 2) kernel function. A common disadvantage of non-parametric techniques such as SVMs is the lack of transparency of results.

[8] Sushma R. Huddar, Swarna Gowri, Keerthana K.,Vasanthi S. and Sudhir Rao Rupanagudi "Novel Algorithm for Segmentation and Automatic Identification of Pests on Plants using Image Processing ", The novel algorithm is based on relative difference in pixel intensities (RDI). Let α (i, j) to be the blue chroma(Cb) value of the current pixel . The relative difference of Cb intensities between the current pixel under consideration (α) and surrounding pixels are calculated and compared with a threshold value θ.

[9] L. Zhu and Z. Zhang, “Insect recognition based on integrated region matching and dual tree complex wavelet transform”, they used integrated region matching and dual tree complex wavelet transform for image matching. The collected images of lepidopteran insects are reduced into 248×200 pixels and filtered using mean shift algorithm. The processed image is segmented in region based on k-mean algorithm using colour feature.

CHAPTER 3
SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

The existing system can only identify the type of diseases which affects the leaf using different types of classification techniques.

3.1.1 K-NEAREST NEIGHBOUR

k-Nearest Neighbour is a simple classifier in the machine learning techniques where the classification is achieved by identifying the nearest neighbours to a query examples and then make use of those neighbours for determination of the class of the query. In KNN the classification i. e. to which class the given point is belongs is based on the calculation of the minimum distance between the given point and other points. As a classifier the nearest neighbour does not include any training process. It is not applicable in case of large number of training examples as it is not robust to noisy data. For the plant leaf classification the Euclidean distance between the test samples and training samples is calculated. In this way it finds out similar measures and accordingly the class for test samples. A sample is classified based on the highest number of votes from the k neighbours, with the sample being assigned to the class most common amongst its k nearest neighbours. k is a positive integer, typically small. If $k = 1$, then the sample is simply assigned to the class of its nearest neighbour. In binary (two class) classification problems, it is helpful to choose k to be an odd number as this avoids tied votes.Nearest neighbour method is easy to implement also quite good results if the features are chosen carefully.

Disadvantages

- The main disadvantage of the KNN algorithm is that it is a slow learner, i.e. it does not learn anything from the training data and simply make use the training data itself for classification.
- Another disadvantage is this method is also rather slow if there are a large number of training examples as the algorithm must have to compute the distance and sort all the training data at each prediction.
- Also it is not robust to noisy data in case of large number of training examples. The most serious disadvantage of nearest neighbour methods is that they are very sensitive to the presence of irrelevant parameters.

3.1.2 PROBABILISTIC NEURAL NETWORKS

Probabilistic Neural Networks (PNNs) is a neural network , based on Parzen windows. In a PNN, the operations are organized into a multi-layered feed forward network with four layers. PNN is mainly used in classification problems. The first layer is input layer which calculates the distance from the input vector to the training input vectors. The second layer sums the contribution for each class of inputs and produces its net output as a vector of probabilities. Third Pattern layer contains one neuron for each case in the training data set. It stores the values of the predictor variables for the case along with the target value. The pattern neurons add the values for the class they represent. The output layer compares the weighted votes for each target category accumulated in the pattern layer and uses the largest vote to predict the target category.

As PNNs are much faster than multilayer perceptron networks their training phase requires only one pass through the training patterns. PNN can be accurate than multilayer perceptron networks also relatively insensitive to outliers.

To improve the overall performance PNNs output can be later processed by another classification system and as this happens very fast, PNNs are used in on-line applications where a real-time classifier is required.

Disadvantage

- PNN is it requires large storage space

3.2 PROPOSED SYSTEM

For this study, whiteflies and aphids are chosen because this pest requires early detection and treatment to prevent durable infection. The acquired Images are given to the local machine and the image processing techniques will takes place.

We briefly explain about the experimental analysis of our methodology. Samples of 10 images are collected that comprised of different leaf disease like canker, blight, leaf spot, root rot, gall and healthy leaves. The primary attributes of the image are relied upon the shape, colour and size. The sample screenshot displays the leaf disease detection using color based segmentation model.

Advantage

- In this system, we classify leaf disease using CNN is very efficient
- This leads economical development of countries
- To identify the plant leaf disease we will increase the crop production

CHAPTER 4
TECHNICAL BACKGROUND

Agriculture is the backbone of India. Most of the scientists are doing research to increase the cultivity of crops. But one problem still exist which is a major concern of the cultivation of crop and that is crop pests. Due to these problems, the cultivation decreases and hence all the farmers and in turn the country suffers from lack of cultivation of crop. Different types of pesticides are there in market which are used to avoid the damage to fruit and vegetable, but the amount of pesticides to be used is not known due to which the cost as well as the environmental pollution gets affected. A strong demand now exists in many countries for non-chemical control methods for pests or diseases. Greenhouses are considered as biophysical systems with inputs, outputs and control process loops. Most of these control loops are automatized (e.g., climate and Fert irrigation control). However no automatic methods are available which precisely and periodically detect the pests on plants. In fact, in production conditions, periodically observes plants and search for pests. This manual method is too time consuming. Diagnosis is a most difficult task to perform manually as it is a function of a number of parameters such as environment, nutrient, organism etc. With the recent advancement in image processing and pattern recognition techniques, it is possible to develop an autonomous system for pest classification. Early detection of pest or the initial presence of a bio aggressor is a key-point for crop management. The detection of biological objects as small as such insects (dimensions are about 2mm) is a real challenge, especially when considering greenhouses dimensions (10– 100m long).

For this purpose different measures are undertaken such as manual observation of plants. This method does not give accurate measures. Hence automatic detection is very much important for early detection of pests.

However, rice may lose its quantity and quality when rice is attacked by different insect pests. Therefore, it is a top priority to find effective methods to reduce the level of their infestation in the paddy fields. In agriculture, pest control has always been considered as the most challenging task for Rice farmers [7]. Most of the farmers used the traditional pest management methods which is the regular spray program based on schedules rather than the presence of insect pests on the paddy fields. These chemicals kill useful insects which eradicate pests in crops. Assessing the density of the rice pest population in paddy fields is very important for pest forecasting decisions. Sticky traps are widely used to trap the insect pests. The trapped insects are brought to the laboratory for counting and identify manually. Usually, crop technicians identify and segregate the insects manually according to their species and count the major pests separately. the resulting counts are used to estimate the pest density in the paddy fields. However, multiple site and frequent counting of rice pests is time consuming and tedious for a crop technician. This can lead to low count accuracy and delays in obtaining accurate counts that can lead to poor decisions on rice pest management. Due to the rapid development of digital technology, there is an opportunity for image processing technology to be used in the field of agricultural research which could help the researcher to solve a complex problem. Image analysis provides a realistic opportunity for the automation of insect pest detection.

This study extends the implementation of image processing techniques to estimate pest densities in rice fields by establishing an automated detection system.

Through this system, crop technicians can easily count the pests from the collected specimens, and right pests' management can be applied to increase both the quantity and quality of rice production. Using the automated system, crop technicians can make the monitoring process easier. Rice infestation may be easily detected and monitored with the use of a camera.

Today there are more than 250 organic pesticides and thousands of formulations. The coffee industry unfortunately relies on these poisons to protect the plant and berries from insect attack and disease spread. In some advanced countries aerial spraying of these hazardous chemicals is carried out to save on labour costs. Most, coffee farmers advocate the use of BROAD SPECTRUM pesticides. These are more dangerous than systemic pesticides because they act on many insects both beneficial and harmful. There is every chance that these chemicals can easily drift or get washed or leached by heavy showers and reach groundwater or open estuaries there by contaminating the earth's precious water reserve. However, the cultivation of these crops for optimum yield and quality produce is highly technical. A lot of research has been done on greenhouse agro systems and more generally on protected crops to control pests and diseases by biological means instead of pesticides. Research in agriculture is aimed towards increase of productivity and food quality at reduced expenditure and with increased profit, which has received importance in recent time. A strong demand now exists in many countries for non -chemical control methods for pests or diseases. However no automatic methods are available which precisely and periodically detect the pests on plants. In fact, in production conditions, greenhouse staff periodically observes plants and search for pests. This manual method is time consuming.

With the recent advancement in image processing pattern recognition techniques, it is possible to develop an autonomous system for disease classification of crops. In this paper, we focus on early pest detection.

First, this implies to regularly observe the plants. Disease images are acquired using cameras or scanners. Then the acquired image has to be processed to interpret the image contents by image processing methods. The focus of this paper is on the interpretation of image for pest detection.

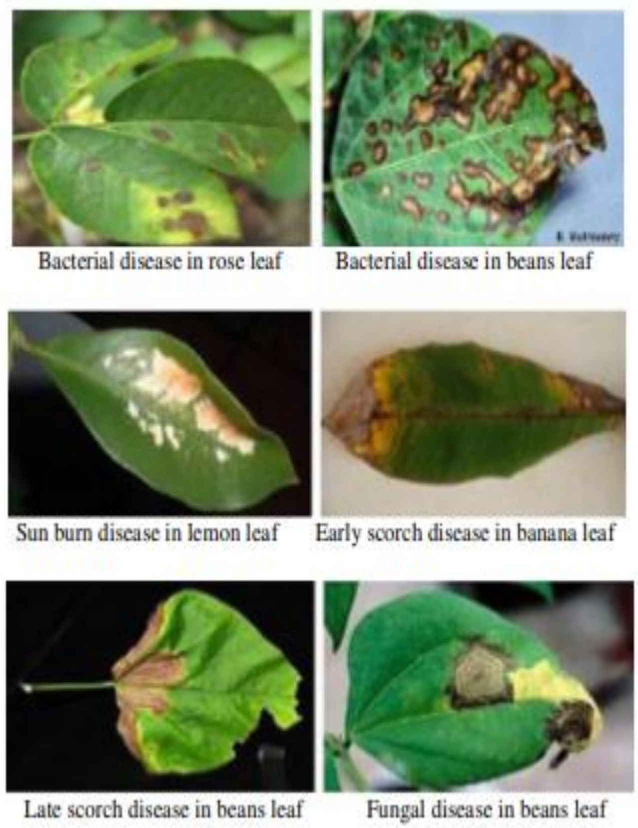


Figure 4.1 Samples of leaf disease image

Figure 4.2 Aphids on leaf

Figure 4.3 Whiteflies on leaf

Agriculture plays key role in the development of human civilization. Research in agriculture focuses on increasing the productivity and food quality. The quality and quantity of agriculture production is affected by environmental parameters and biological parameters. The major biological parameters are pests and plant diseases. Diseases and insect pest are the major problems in the agriculture. These require careful diagnosis and timely handling to protect the crops from heavy losses. The naked eye observation is the commonly used method for detection of pest and identification of plant diseases. This needs continuous monitoring.

But it is not practical in the case of large farm. Also, it is not accurate, expensive and time consuming.

Integrated pest management (IPM) method used to control pests with minimal environmental impacts. Detection, identification and application of correct management are the steps of IPM. Machine vision systems are widely used for inspection of growing plants to recognize their diseases.

The image processing can be used in various agriculture applications such as to detect diseased leaf and stem, to identify the affected area, to determine colour of affected area etc. Automatic detection of plant diseases with the help of image processing technique provides more accurate pest detection and guidance for disease management. Various papers have been published which explain different techniques used, including image processing for detection of objects, extraction of features and identification of pests based on various parameters such as colour, boundary, background colour, foreground colour, intensity of pixels etc. We have reviewed papers that have been researched on pest identification with the intention of getting a fair idea about the methods which are efficient and useful.

Basic steps of plant disease detection algorithm are

1. Image Acquisition
2. Image Pre-processing
3. Feature Extraction
4. Classification
5. Diagnosis

CHAPTER 5

SYSTEM IMPLEMENTATION

5.1 METHODOLOGY

1) Image capturing

The first step of every image processing application is image acquisition or image capturing. The images of leaves are captured by using the camera and it will store it in some formats like .PNG, .JPG, .JPEG etc.

2) Image pre-processing

Image pre-processing is used to create an enhanced and please full version of the captured image. The image pre-processing steps used in the system are: 1) Conversion of RGB image to grey image 2) Resizing of the image 3) Filtering of the image. a) Conversion of RGB to Grey Image in RGB colour model, each colour appears in its primary spectral components of red, green, and blue. The colour of a pixel is made up of three components; red, green, and blue (RGB). The disadvantages of RGB models are, it requires large space to store and it will take more time to process. So there is a need for converting the RGB model to Grey-model. b) Resizing of the Image Resizing is an important step in image pre-processing. The acquired image is resized according to the requirement of the system. Resizing of the image: Resizing is nothing but, changing the dimensions of an image. The captured image is resized using some resizing methods according to the requirement of the system. There are different methods for the resizing of images. Bilinear, Bicubic and Nearest neighbourhood interpolation are the common resizing methods. Here in our system, we are using bicubic method.

c) Filtering of the image Filtering is nothing but, eliminating the unwanted portion of the image. Different types of filters are available. Low pass filters are smoothening filters, it will pass only low frequency signals and eliminate all the high frequency signals.

High pass filters are sharpening filters, and it will eliminate all the low frequency signals and pass only high frequency signals. Band pass filters will pass the signals which is having a specific range of frequencies. In our system we are using smoothening filter. The purpose of smoothing is to reduce noise and improve the visual quality of the image. Spatial filters are applied to both static and dynamic images, whereas temporal images are applied only to dynamic images. The simplest smoothening filter is average filter.

3) Feature Extraction

Feature extraction is the most important part of this project. Some properties of the images are considered here. The different types of properties include region properties, grey covariance matrix properties etc. The properties standard deviation, entropy, contrast etc are extracted from the image and are used to train the dataset for the SVM classification. Support Vector Machines (SVM's) are a relatively new learning method used for binary classification. The basic idea is to find a hyper plane which separates the d-dimensional data perfectly into its two classes.

4) Classification

In this section we find the types of disease affected on the leaf. From the image of training and testing data, both are used to compare and give the results of affected disease and what type of disease is affected on the leaf using Convolutional Neural Network(CNN).

In CNN, the image is passed through a series of convolutional, nonlinear, pooling layers and fully connected layers, and then generates the output. A convolutional neural network consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, RELU layer i.e. activation function, pooling layers, fully connected layers and normalization layers.

Our input is a training dataset that consists of N images, each labelled with one of K different classes. Then, we use this training set to train a classifier to learn what every one of the classes looks like. In the end, we evaluate the quality of the classifier by asking it to predict labels for a new set of images that it has never seen before. We will then compare the true labels of these images to the ones predicted by the classifier.

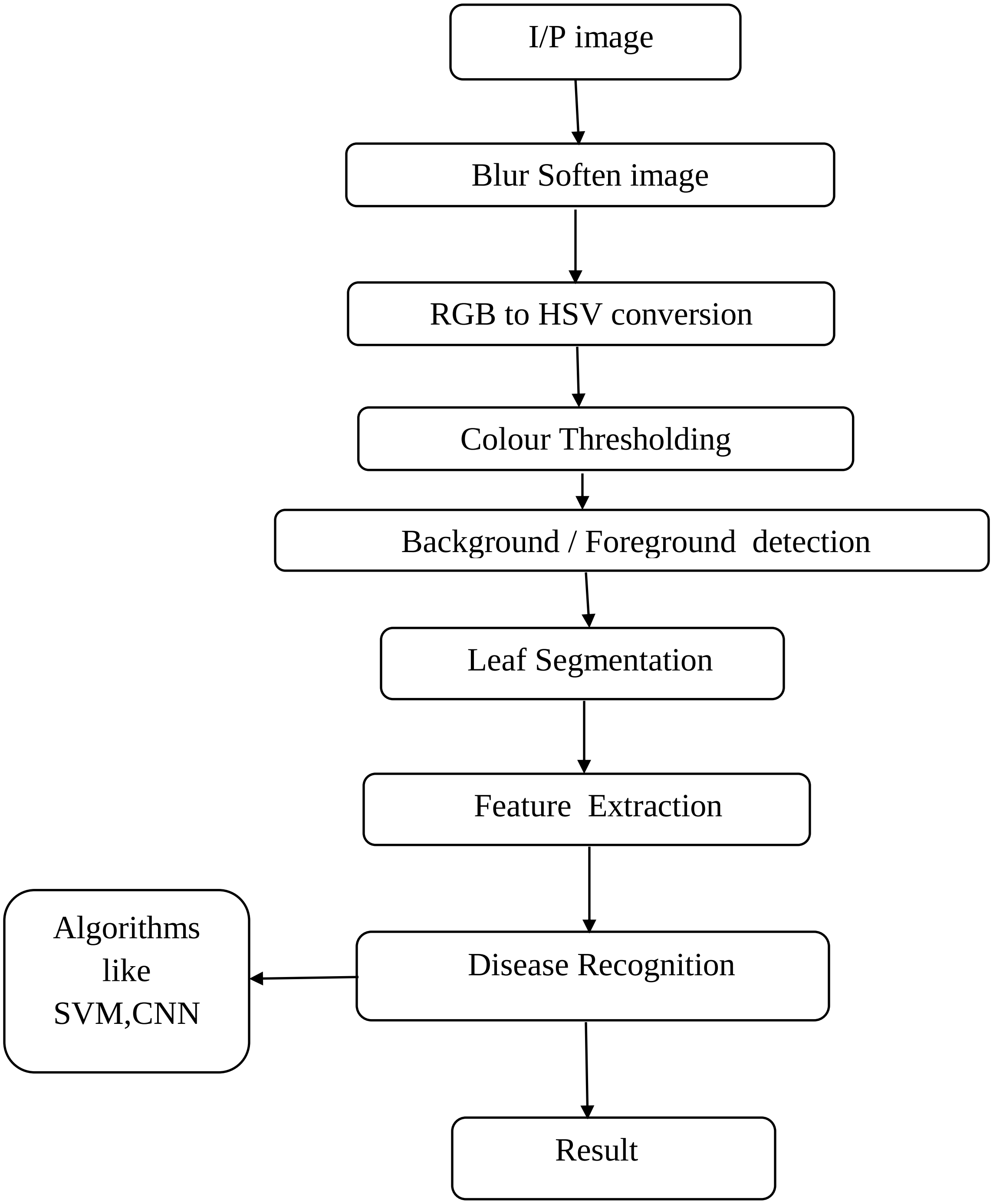


Fig 5.1 System Architecture

CHAPTER 6
SYSTEM REQUIREMENTS

6.1 HARDWARE REQUIREMENTS

The hardware requirements may serve as the basis for a contact for the implementation of the system and should therefore be a complete and consistent specification of the whole system.

System : Acer

Processor : Intel(R) core(TM) i3-3227U CPU @ 1.90GHz

RAM : 8.00 GB

Pen and Touch : Touch support with 10 Touch points

Input/output Peripheral : Mouse and Keyboard.

6.2 SOFTWARE REQUIREMENTS

The software requirements provide a basis for creating the software requirements specification .

Operating system : windows 10

Coding Language : python

Tool : Anaconda (Jupyter)

CHAPTER 7
SOFTWARE DESCRIPTION

7.1 ANACONDA

Anaconda is an open source package management system and environment management system that runs on Windows, mac OS and Linux. Conda quickly installs, runs and updates packages and their dependencies. Conda easily creates, saves, loads and switches between environments on your local computer. It was created for Python programs, but it can package and distribute software for any language.

Conda is a packaging tool and installer that aims to do more than what pip does: handle library dependencies outside of the Python packages as well as the Python packages themselves.Conda is written entirely in Python which makes it easier to use in Python virtual environments. Furthermore, we can use Conda for C libraries, R packages, Java packages and so on. It installs binaries. The conda build tool builds packages from source and conda install installs things from built conda packages.

Conda is the package manager of Anaconda, the Python distribution provided by Continuum Analytics. Anaconda is a set of binaries that includes Scipy, Numpy, Pandas along with all their dependencies.

Scipy is a statistical analysis package.

Numpy is a numerical computing package .

Pandas is a data abstraction layer that exposes a way to merge and transform data.

7.2 JUPYTER NOTEBOOK

7.2.1 INTRODUCTION

Jupyter is tool helps you to execute python code. The code which you perform will be stored in cloud and can access anytime. Jupyter is line by line execution.

The Jupyter Notebook is an interactive computing environment that enables users to author notebook documents that include: - Live code - Interactive widgets - Plots - Narrative text - Equations - Images - Video

These documents provide a complete and self-contained record of a computation that can be converted to various formats and shared with others using email, Dropbox, version control systems (like git/GitHub).

7.2.2 COMPONENTS

The Jupyter Notebook combines three components:

The notebook web application: An interactive web application for writing and running code interactively and authoring notebook documents.

Kernels: Separate processes started by the notebook web application that runs users' code in a given language and returns output back to the notebook web application. The kernel also handles things like computations for interactive widgets, tab completion and introspection.

Notebook documents: Self-contained documents that contain a representation of all content visible in the notebook web application, including inputs and outputs of the computations, narrative text, equations, images, and rich media representations of objects. Each notebook document has its own kernel.

7.3 ANACONDA PYTHON

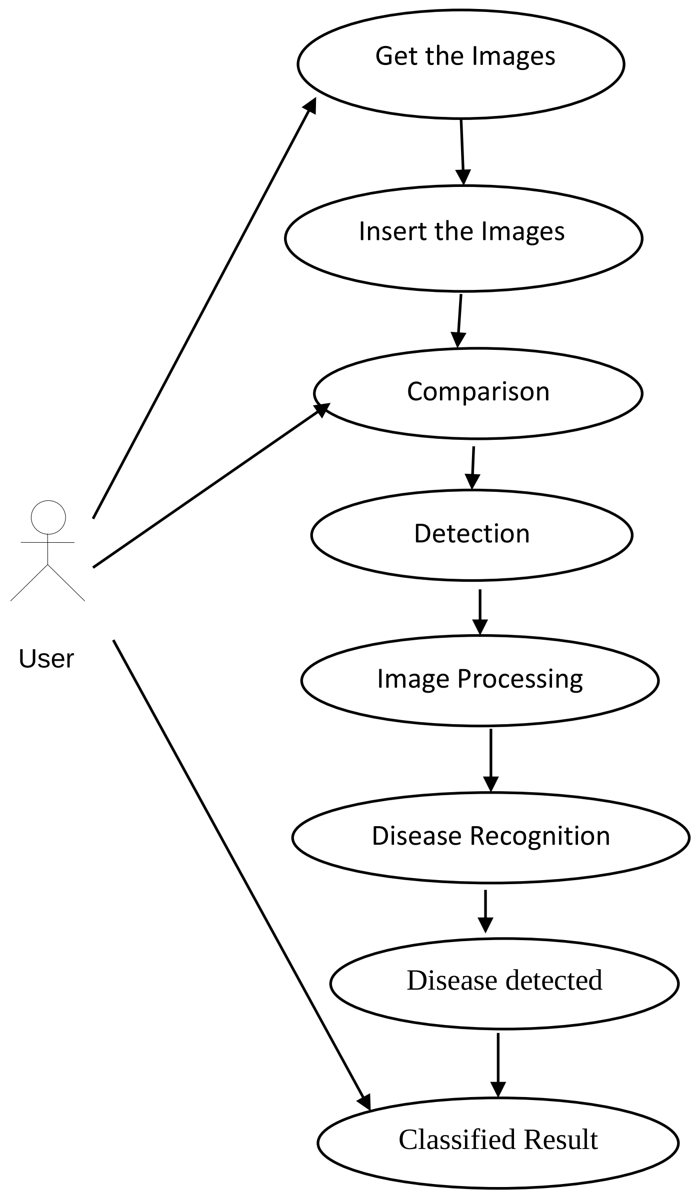
Anaconda python is faster than vanilla python: they bundle Intel MKL and this does make most numpy computations faster. You can easily do a local user install, no need to ask permission from your admin in many cases (you may face web proxy issues though)Under Windows, you don't have a lot of choices and anyway you need a python package installer.

Anaconda inc. is a company. This is a plus in a corporate setting. You can get support contracts for instance. Anaconda Inc. has historically made a lot of efforts to please the personal, students and academic users. This has basically worked out pretty well for all concerned.

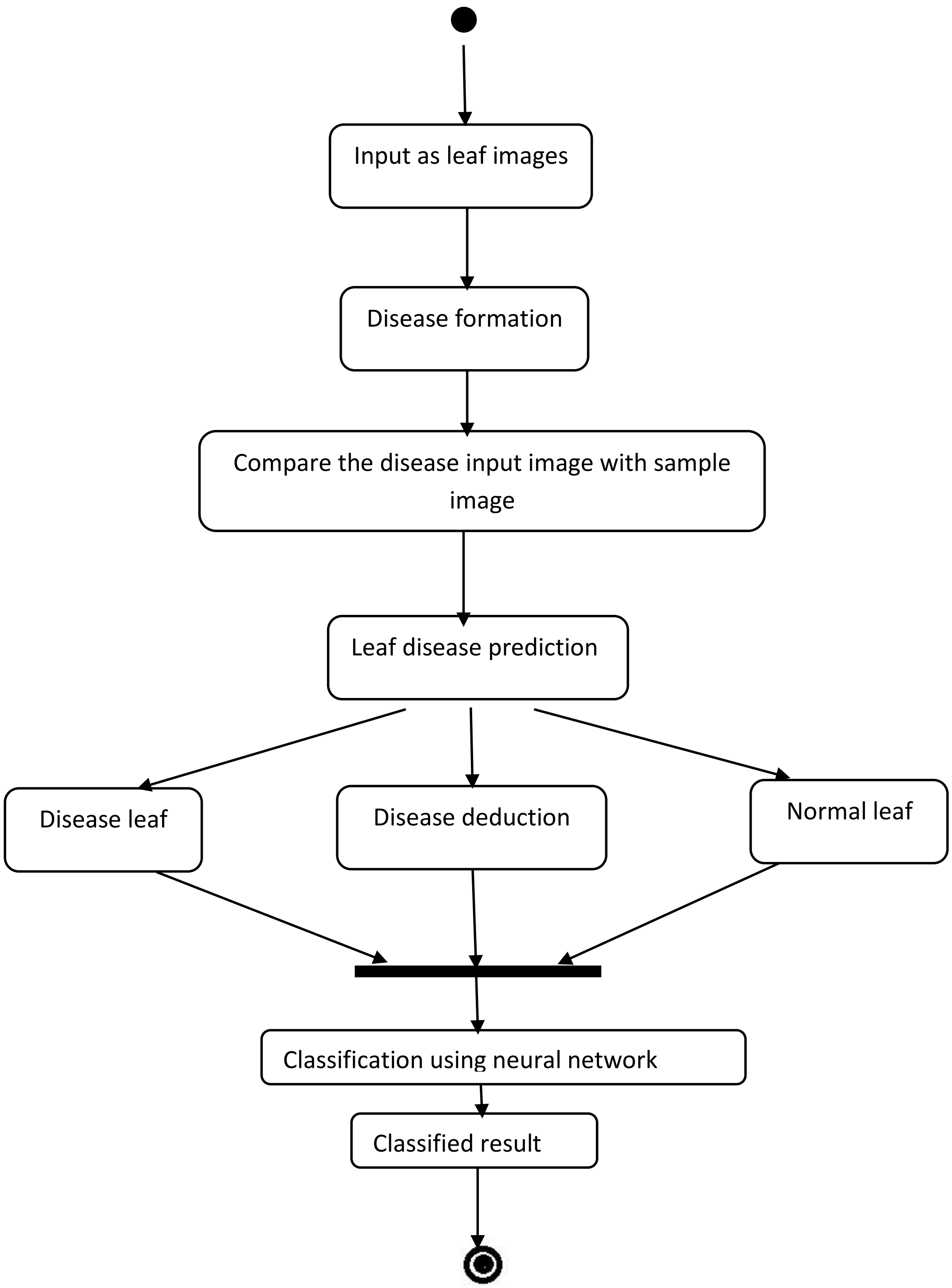
CHAPTER 8

UML DIAGRAMS

8.1 USE CASE DIAGRAM



8.2 ACTIVITY DIAGRAM



CHAPTER 9

SOURCE CODE FOR CNN_TRAIN

```
print("starting")
from sklearn.datasets import load_files
from keras.utils import np_utils
import numpy as np
from glob import glob
import cv2
import numpy as np
from keras.preprocessing import image
from tqdm import tqdm
# In[2]:
print("done import")
# define function to load train, test, and validation datasets
def load_dataset(path):
    data = load_files(path)
    dog_files = np.array(data['filenames'])
    dog_targets = np_utils.to_categorical(np.array(data['target']), 3)
    return dog_files, dog_targets
# In[3]:
print("starting load data")
# load train, test, and validation datasets
```

```
train_files, train_targets = load_dataset('../data/sunflower/training')

test_files, test_targets = load_dataset('../data/sunflower/training')

print("done load data")

import random

random.seed(9)

def path_to_tensor(img_path):

    # loads RGB image as PIL.Image.Image type

    try:

        img = image.load_img(img_path, target_size=(64,64))

    # convert PIL.Image.Image type to 3D tensor with shape (224, 224, 3)

    except IOError:

        pass

    x = image.img_to_array(img)

    # convert 3D tensor to 4D tensor with shape (1, 224, 224, 3) and return 4D tensor

    return np.expand_dims(x, axis=0)

def paths_to_tensor(img_paths):

    list_of_tensors = [path_to_tensor(img_path) for img_path in tqdm(img_paths)]

    return np.vstack(list_of_tensors)

# pre-process the data for Keras

train_tensors = paths_to_tensor(train_files).astype('float32')/255

test_tensors = paths_to_tensor(test_files).astype('float32')/255

from keras.layers import Conv2D, MaxPooling2D, BatchNormalization
```

```
from keras.layers import Dropout, Flatten, Dense

from keras.models import Sequential

model = Sequential()

model.add(Conv2D(filters=32,kernel_size=(3,3),strides=(1,1),input_shape=(64,64,
3),activation='relu'))

model.add(BatchNormalization())

model.add(MaxPooling2D(pool_size=(2,2)))

model.add(Conv2D(filters=64,kernel_size=(3,3),strides=(1,1),activation='relu'))

model.add(BatchNormalization())

model.add(MaxPooling2D(pool_size=(2,2)))

model.add(Flatten())

model.add(Dense(512,activation='relu'))

model.add(BatchNormalization())

model.add(Dense(3,activation='softmax'))

model.summary()

# ### Compile the Model

# In[20]:

model.compile(optimizer='rmsprop',loss='categorical_crossentropy',
metrics=['accuracy'])

# ### (IMPLEMENTATION) Train the Model

#

# Train your model in the code cell below.  Use model checkpointing to save the
model that attains the best validation loss.

#
```

You are welcome to [augment the training data](https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html), but this is not a requirement.

In[21]:

```
from keras.callbacks import ModelCheckpoint

### TODO: specify the number of epochs that you would like to use to train the model.

epochs = 4

### Do NOT modify the code below this line.

checkpointer = ModelCheckpoint(filepath='weights.best.from_scratch.hdf5',

            verbose=1, save_best_only=True)

model.fit(train_tensors,train_targets,epochs=epochs,batch_size=64, callbacks=[checkpointer], verbose=1)

#

score=model.evaluate(test_tensors,test_targets)

print('test loss:',score[0])

print('test accuracy:',score[1])
```

CHAPTER 10
CONCLUSION

Detecting the disease is main purpose of the proposed system. The result shows the valuable approach which find out accurate detection of the diseased leaf. Image processing technique is applied to detect the affected part of leaf from the input image. CNN algorithm used to classify diseased leaf. In this system we detecting which type of disease is affected on the leaf. In future, this project can be extended for developing application. This technique would be useful for saving the farmers from a huge loss. Hence this project is a small contribution to farmers to with stand in their life.

REFERENCES

[1] M. K. Alsmadin, K. B. Omar, S. A. Noah, and I. Almarashdeh, “Fish recognition based on robust features extraction from color texture measurements using back-propagation classifier,” Journal of Theoretical And Applied Information Technology, 2010.

[2] S. M. Al-Saqer, “A robust recognition system for pecan weevil using artificial neural networks,” American Journal of App. Sci., 2012.

[3] F. A. Carino, P. E. Kenmore, and V. A. Dyck, “A FARMCOP suction sampler for hoppers and predators in flooded rice fields,” The International Rice Research Newsletter, vol. 4, ch. 5, pp. 21-22, 1979.

[4] Gaurav Kandalkar, A.V.Deorankar and P.N.Chatur, "Identification of Agricultural Pests Using Radial Basis Function Neural Networks ", International Journal of Engineering Research and Applications, April 2014, pp.52-56.

[5] Ganesh Bhadane, Sapana Sharma and Vijay B. Nerkar,"Early Pest Identification in Agricultural Crops using Image Processing Techniques", International Journal of Electrical, Electronics and Computer Engineering 2(2), 2013, pp. 77-82.

[6] Johnny L. Miranda, Bobby D. Gerardo and Bartolome T. Tanguilig III, "Pest Identification using Image Processing Techniques in Detecting Image Pattern through Neural Network", Advances in Computer and Electronics Technology, 2014.

[7] Rupesh G. Mundada and Dr. V. V. Gohokar,"Detection and Classification of Pests in Greenhouse Using Image Processing", IOSR Journal of Electronics and Communication Engineering, Vol 5, Issue 6 (MarApr) 2013, pp. 57-63.

[8] Sushma R. Huddar, Swarna Gowri, Keerthana K.,Vasanthi S. and Sudhir Rao Rupanagudi "Novel Algorithm for Segmentation and Automatic Identification of Pests on Plants using Image Processing ", Third International Conference Computing Communication & Networking Technologies(ICCCNT) , July 2012, pp. 26-28.

[9] L. Zhu and Z. Zhang, “Insect recognition based on integrated region matching and dual tree complex wavelet transform,” Journal of Zhejiang University-SCIENCE C (Computers & Electronics), 2011.