Crop Monitoring System Using Image Processing and Environmental Data

Inspiring Excellence

Shadman Rabbi - 18341015
Ahnaf Shabik - 13201013

Supervised by:
Dr. Md. Ashraful Alam

Department of Computer Science and Engineering
BRAC University

Declaration

We, hereby declare that this thesis is based on the results found by ourselves. Materials of work found by other researcher are mentioned by reference. This thesis, neither in whole or in part, has been previously submitted for any degree.

Signature of Supervisor
(Dr. Md. Ashraful Alam)

Signature of Author
(Shadman Rabbi)

Signature of Author
(Ahnaf Shabik)

Acknowledgement

Firstly we would like to thank the almighty for enabling us to initiate our research, to put our best efforts and successfully conclude it.

Secondly, we offer our genuine and heartiest appreciation to our regarded Supervisor Dr. Md. Ashraful Alam for his contribution, direction and support in leading the research and preparation of the report. His involvement, inclusion and supervision have inspired us and acted as a huge incentive all through our research.

Last but not the least, we are grateful to the resources, seniors, companions who have been indirectly but actively helpful with the research. Especially Monirul Islam Pavel, Ekhwan Islam, Touhidul Islam and Tahmidul Haq has propelled and motivated us through this journey. We would also like to acknowledge the help we got from various assets over the internet; particularly from fellow researchers' work.

Table of Contents

List of Figures

Abstract

We propose advanced crop monitoring system using image processing from environmental data for faster and better yield of agriculture. In the system, we have used colour segmentation and binary masking to analyse and differentiate between various stages of crop plantation (unripe, ripe and diseased) via aerial images. To further improve accuracy, we have used different sensors to detect important factors like the temperature and moisture of the air and soil. This is going to project precise areas of the affected regions and can be taken care of immediately after detection. Furthermore, using Support Vector Machine (SVM), we have been able to classify different types of leaf diseases up to 98% accuracy. We have also used other algorithms, like K-mean cluster and L*a*b* colour space, to detect the diseased part of the leaves. This will ensure the farmers to provide the right type of treatment for the plants. We have taken this initiative in the perspective of Bangladesh where a huge amount of population requires huge amount of food supplement and hence we need a faster and safer monitoring system to meet the agricultural needs of the nation.

Chapter 1

1. Introduction

1.1 Motivation

Bangladesh has a wide range of rural areas producing a large number of agricultural products. However the productivity has not been as optimal as of the developed countries. This is because Bangladesh has been lacking the amount of technological advancement that is required to feed such a large population, nevertheless for export purposes. Therefore technologies such as this crop monitoring system will give better yield in a minimum amount of time. Farmers working in vast fields will be able to process information faster through the data analysis of the images of the crops. Thus coming to faster conclusions and proceed with actions within a short span of time. This will boost the productivity and allow better economic results.

The system works by reading aerial images through which any anomalies within the crops can be detected and reported to the farmers for immediate actions to be taken. This will allow the farmers to be alerted at an early stage about any drastic changes that have taken place inside the field. For instance, if there are crops that have started to ripe earlier or if any part of it have started catching diseases. The affected area will be classified through colour segmentation and the system will detect which crops are healthy and which are not from given parameters allowing the farmers to act on the defected ones almost instantly. For further accuracy there will be sensors to support the detection with factors such as temperature and moisture of air and soil along with classification of leaf diseases through Support Vector Machine (SVM).

1.2 Literature Review

Gondal, et al. worked to find an automatic way to detect pests at an early stage. Agriculture is an important part of human life and the economy of a country and a lot of money is being spent for protection of crops. One of the methods to tackle the damage done to the crops by the insects is to detect them at any early stage. Constant monitoring of the crop will yield accuracy about the health of the plants. Early detection will cut huge production losses at the end because appropriate actions can be taken at the primary stage along with lowering the usage of the pesticides. It is a common research in the present because manually looking over the large fields is very difficult and requires a large human force for detailed examinations. Therefore, systems like this can not only analyse the crops to detect pest infestation but can also differentiate the type of pests. It is effective for computerised techniques to examine the images of the plants. According to them, Support Vector Machine (SVM) is used for classification of images through its features. And it is a simpler technique than other automated techniques and provides better results[1].

Prof. Bhavana Patil, A. C. Patil College, et al used their ideas in In India where 70% of of the population are under the Agriculture sector. From agriculture, 2% of raw materials like crops get destroyed by natural calamities and 98% by the pathogens. Since the traditional methods were ineffective, image processing can be used for accurate detection of disease. Different spots and patterns on plant leaves can be used to detect the disease. For better results advanced digital image processing was used. On referring various reputed IEEE, international conference and international journal papers on plant monitoring, they say that there has not been any perfect solutions to curing the diseases [2].

S. D. Khiradeand, A. B. Patil says identification of the plant disease is the key to preventing the quantity of agricultural product. Study of plant disease means the patterns that we can see on the plants which are visible. It is very critical for any type of long term or sustainable agricultural to monitoring health and disease detection. It requires so much work, need expert in works and lots of time. Image process is one of the certain processes that are used for detecting plant disease. Disease detection involves with image acquisition image-pre-processing, image segmentation and classification. These are used with their leaves images. This paper also used in some algorithm for plant disease detection [3].

Savita N et al discuss paper will help to know about different classification of plant leaf disease. One classification technique deals with getting each pattern in one of the distinct classes. Classification is based on its different morphological techniques. There are so many classification techniques like probabilistic neural network, genetic algorithm etc. Selecting a classification method is always difficult. Plant leaf disease classifications have wide applications in various fields like biological research, agricultural etc. this paper will provide different ideas of plant leaf diseases [4].

Agriculture sector plays a strategic role in the process of economic development of a country. Having disease in plants are quite natural so disease detection in agriculture field will play an important role in countries economy. Proper care can reduce the amount of damage of a plant. Detecting plants disease by an automatic technique will save working hour. Detecting plant disease when they appear on plant leaves will save plant life [5].

1.3 Thesis Orientation

- **Chapter 2** elaborations of all the algorithms and list of sensors
- **Chapter 3** presents the proposed model of our research
- **Chapter 4** demonstrates the results found in our research
- **Chapter 5**concludes the thesis and states the future research directions

Chapter 2

This chapter gives all the elaborations and details of all the algorithms and sensors that have been used for this research. Each of the algorithms and sensors play a vital role in the crop monitoring system.

2. Elaborations of Algorithms and List of Sensors

2.1 Algorithms

2.1.1 Canny Edge Detector Algorithm

The canny edge detector is an edge identification operator that uses a several-stage algorithm to recognize an extensive variety of edges in pictures.

The Canny algorithm is used in grey images and can be classified into four parts: image smoothing, gradient computation, non-maxima suppression, edge determination and connection.

Image Smoothing

Image smoothing is an important procedure to reduce noise before edge detection. The Canny algorithm uses two orders Gaussian function to smooth image. The function is as follows,

$$G(x,y) = \frac{1}{2\pi\sigma^2}\exp[-\frac{x^2 + y^2}{2\sigma^2}] \dots \dots \dots (2.1)$$

The Slope Vector is as follows,

$$\nabla G = [\frac{\partial G}{\partial x}, \frac{\partial G}{\partial y}]^T \dots \dots \dots (2.2)$$

The two orders Gaussian function is as follows,

$$\frac{\partial G}{\partial x} = kx.\exp\left|-\frac{x^2}{2\sigma^2}\right|.\exp\left|-\frac{y^2}{2\sigma^2}\right| \dots \dots \dots (2.3)$$

$$\frac{\partial G}{\partial y} = ky.\exp\left|-\frac{y^2}{2\sigma^2}\right|.\exp\left|-\frac{x^2}{2\sigma^2}\right| \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.4)$$

Gradient Computation

The finite difference is used to do gradient computation as follows:

$$E_{i,j}^{x} = \frac{1}{2}\left(I_{i,j+1} - I_{i,j} + I_{i+1,j+1} - I_{i+1,j}\right) \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.5)$$

$$E_{i,j}^{y} = \frac{1}{2}\left(I_{i,j} - I_{i+1,j} + I_{i,j+1} - I_{i+1,j+1}\right) \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.6)$$

The slope amplitude and the slope difference formulas are as follows:

$$M_{i,j} = \sqrt{(E_{i,j}^{x})^2 + (E_{i,j}^{y})^2} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.7)$$

$$\theta_{i,j} = \tan^{-1}\left|\frac{E_{i,j}^{y}}{E_{i,j}^{x}}\right| \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.8)$$

Improvement on Canny edge detection

Replace Gaussian filter

As edge and noise will be detected as high frequency signal, ordinary Gaussian filter will bring smooth effect on them. However, on the other hand, in order to get high accuracy of detection of the real edge an adaptive filter could be used to bring an accurate detection. The steps of using this adaptive filter are

1. K = 1, set the iteration n and the coefficient of the amplitude of the edge h.

2. Calculate the gradient value $G_x(x, y)$and $G_y(x, y)$

3. Calculate the weight according to the formula below:

$$d(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.9)$$

$$w(x,y) = \exp(-\frac{\sqrt{d(x,y)}}{2h^2}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.10)$$

4. The adaptive filter is:

$$f(x,y) = \frac{1}{N}\sum_{i=-1}^{1}\sum_{j=-1}^{1} f(x+i,y+j)w(x+i,y+j) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.11)$$

to smooth the image, which

$$N = \sum_{i=-1}^{1}\sum_{j=-1}^{1} w(x+i,y+j) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.12)$$

5. When K = n, stop the iterative, otherwise, k = k+1, keep do the second step

2.1.2Euclidean Distance Metric Algorithm

The Euclidean distance or Euclidean metric is the line distance between two points in Euclidean space. Euclidean space forms a metric space. The related norm is called the Euclidean norm. A general term for the Euclidean norm is the L2 norm or L2 distance.

If we have two points P and Q, then the line segment joining them is the Euclidean distance between the points p and q. If p and q are Cartesian coordinates, then the distance between them can be found by Pythagorean formula:

$$d(p,q) = d(q,p) = \sqrt{(q_1-p_1)^2 + (q_2-p_2)^2 + \cdots + (q_n-p_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n}(q_i-p_i)^2} \ldots\ldots\ldots\ldots\ldots\ldots (2.13)$$

Euclidean vector is a position of a point in Euclidean space. p and q can be shown as Euclidean vectors, starting from the initial point of the space with their terminal points ending at the two points. The Euclidean norm or Euclidean length measures the length of the vector.

$$\|p\| = \sqrt{p_1^2 + p_2^2 + \cdots + p_n^2} = \sqrt{p.p} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.14)$$

For two dimensions

The Euclidean distance is measured by

$$d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2.15)$$

For three dimensions

The Euclidean distance is measured by

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2.16)$$

For nth dimensions

The Euclidean distance is measured by

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_i - q_i)^2 + \cdots + (p_n - q_n)^2} \ldots \ldots \ldots \ldots \ldots (2.17)$$

Squared Euclidean distance

The s Euclidean distance can be squared to place progressively greater weight on objects that are far from each other.

$$d^2(p,q) = (p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_i - q_i)^2 + \cdots + (p_n - q_n)^2 \ldots \ldots \ldots \ldots \ldots \ldots (2.18)$$

2.1.3 K-means Cluster Image Colour Segmentation

K-means cluster is a system of vector quantization, initially from signal processing. It is famous for cluster analysis in data mining. K-means cluster targets to partition n observations to k cluster where every observation is of the cluster with the closest mean. This leads to partitioning of data space into Voronoi cells.

The issue is computationally difficult (NP-hard) but there are useful heuristic algorithms that are usually used and merge rapidly to a local optimum.

Standard Algorithm

The most widely recognized algorithm works with an iterative refinement system and because of its universality it is frequently called the k-means algorithm. It is considered as Lloyd's algorithm especially in computer science. The algorithm works in two steps:

Assignment step

Appoint every observation to the cluster whose mean has the slightest squared Euclidean separation, this is naturally the "closest" mean.

$$S_i^{(t)} = \left\{ x_p : \left\| x_p - m_i^{(t)} \right\|^2 \leq \left\| x_p - m_j^{(t)} \right\|^2 \forall j, 1 \leq j \leq k \right\} \ldots \ldots \ldots \ldots \ldots (2.19)$$

Update step

Figure the new means to be the centroids of the observations in the new cluster.

$$m_i^{(t+1)} = \frac{1}{\left| S_i^{(t)} \right|} \sum_{x_j \in S_i^{(t)}} x_j \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2.20)$$

The algorithm has merged whenever the assignments do not change. There is no surety that the optimum is figured got using this step.

Hartigan-Wong method

Hartigan and Wong's technique gives a more sophisticated however more computationally costly approach to perform k means. It is as yet a heuristic technique.

Assignment Step

In order to perform k-means by Hartigan and Wong's method, we should start out by partitioning the points into random clusters.

$$\{S_j\}_{j \in \{1, \ldots, k\}} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2.21)$$

Update step

Finding the values n,m for which the whole function reaches a minimum.

$$\Delta(m,n,x) = \emptyset(S_n) + \emptyset(S_m) - \emptyset(S_n \setminus \{x\}) - \emptyset(S_m \cup \{x\}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.22)$$

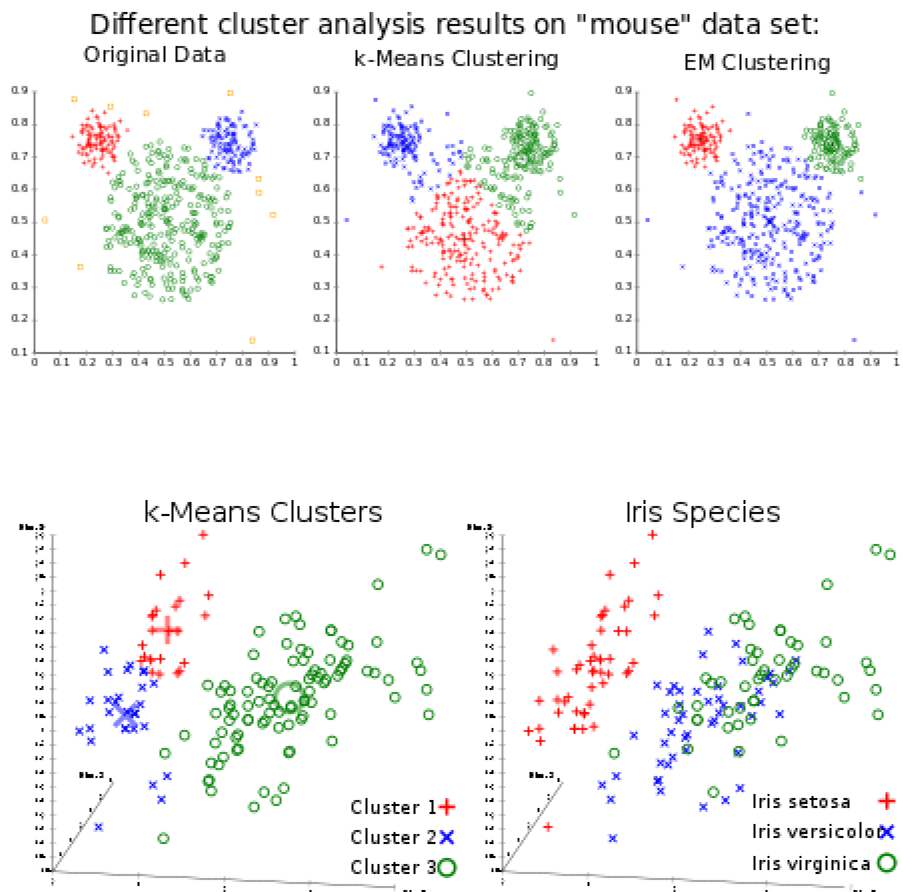


Figure 2.1 Different cluster analysis (Source: Wikipedia).

In Figure 2.1, K-means clustering result for the Iris flower data set and actual species visualized using ELKI. Cluster means are marked using larger, semi-transparent symbols and K-means clustering vs. EM clustering on an artificial dataset ("mouse"). The tendency of k-means to produce equal-sized clusters leads to bad results here, while EM benefits from the Gaussian distributions with different radius present in the data set.

2.1.4 L*A*B* Colour Segmentation

CIE L*a*b* also known as the CIELAB colour space communicates colour as three numerical qualities, L* for the softness and a* and b* for the green– red and blue– yellow colour segments. CIELAB was intended to be perceptually uniform concerning human colour vision, implying that a similar measure of numerical change in these qualities relates to about a similar measure of outwardly apparent change.

A standout amongst the most essential traits of the CIELAB show is that, regarding a given white point, it is gadget autonomous—it characterizes hues free of how they are made or shown. The CIELAB colour space is ordinarily utilized when illustrations for print must be changed over from RGB to CMYK, as the CIELAB extent incorporates both the ranges of the RGB and CMYK colour models.

The space itself is a three-dimensional genuine number space, permitting a limitless number of conceivable portrayals of hues. By and by, the space is generally mapped onto a three-dimensional whole number space for computerized portrayal, and accordingly the L*, a*, and b* values are typically outright, with a pre-characterized extend. The softness esteem, L*, speaks to the darkest dark at L* = 0, and the brightest white at L* = 100. The colour channels, a* and b*, speak to genuine impartial dark qualities at a* = 0 and b* = 0. The a* hub speaks to the green–red part, with green in the negative bearing and red in the positive course. The b* pivot speaks to the blue– yellow part, with blue in the negative bearing and yellow in the positive heading. The scaling and points of confinement of the a* and b* tomahawks will rely upon the particular usage, as depicted underneath, yet they regularly keep running in the scope of ±100 or −128 to +127 (marked 8-bit whole number).

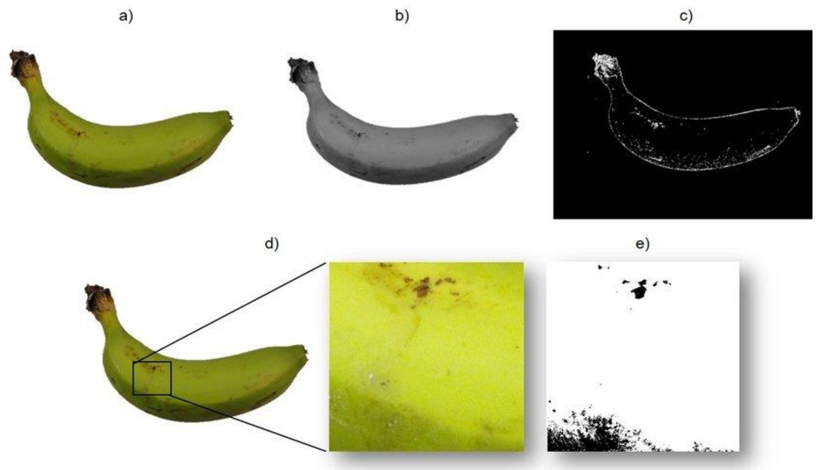


Figure 2.2 Colour Segmentation using L*A*B* [23].

Figure 2.2 shows Digital image processing to image segmentation and determine L*a*b* values. a) Digital image in JPEG format. b) Transformation of the image to grayscale. c) Contour extraction from binary image for determining the fruit contour. d) Square region inside of the fruit contour that was selected for calculating banana peel colour. e) Resultant binary image showing the spot-bruise detection on the banana peel.

Digital image processing to image segmentation and determine L*a*b* values

a) Digital image in JPEG format

b) Transformation of the image to greyscale

c) Contour extraction from binary image for determining the fruit contour

d) Square region inside of the fruit contour that was selected for calculating banana peel colour

e) Resultant binary image showing the spot-bruise detection on the banana peel

CIELAB to CIEXYZ:

Usually there are two transformations forward transformation and reverse transformation.

Forward transformation

$$L^* = 116f(\frac{Y}{Y_n}) - 16 \ldots\ldots (2.23)$$

$$a^* = 500(f\left(\frac{x}{x_n}\right) - f(\frac{y}{y_n})) \ldots\ldots (2.24)$$

$$b^* = 200(f\left(\frac{y}{y_n}\right) - f(\frac{z}{z_n})) \ldots\ldots (2.25)$$

Where

$$f(t) = \begin{cases} \sqrt[3]{t} \\ \frac{t}{3\delta^2} \end{cases} + \frac{4}{29} if \;\; t > \delta^2 \ldots\ldots (2.26)$$

Otherwise,

$$\delta = \frac{6}{29} \ldots\ldots (2.27)$$

Here $X_n, Y_{n,} Z_n$ are the CIE XYZ tristimulus values of the reference white point (the subscript n suggests "normalized").

2.1.5 Support Vector Machine (SVM)

In machine learning, support vector machines (SVMs), additionally support vector networks are directed learning models with related learning calculations that break down information utilized for arrangement and relapse examination. Given an arrangement of preparing cases, each set apart as having a place with either of two classifications, a SVM preparing calculation constructs a model that relegates new cases to one classification or the other, making it a non-probabilistic parallel straight classifier (in spite of the fact that strategies, for example, Platt scaling exist to utilize SVM in a probabilistic order setting). A SVM demonstrate is a portrayal of the cases as focuses in space, mapped so the cases of the different classifications are separated by an unmistakable hole that is as wide as could be expected under the circumstances. New illustrations are then mapped into that same space and anticipated to have a place with a class in light of which side of the hole they fall.

A standout amongst the most widely recognized errands in machine learning is classifying data. Thinking of some as given information indicates each have a place one of two classes and the objective is to choose which class another information point will be in. On account of help vector machines, an information point is seen as a ρ-dimensional vector (a list of ρ numbers), and afterward we need to know whether we can separate such focuses with a (ρ - 1) dimensional hyperplane. This is known as a straight classifier numerous hyperplanes there are to group the information. One sensible decision as the best hyper plane is the one that speaks to the biggest partition, or edge, between the two classes. So we pick the hyper plane with the goal that the separation from it to the closest information point on each side is boosted. In the event that such a hyperplane exists, it is known as the most extreme edge hyperplane and the straight classifiercharacterizes is known as a greatest edge classifier; or identically, the view of ideal soundness.

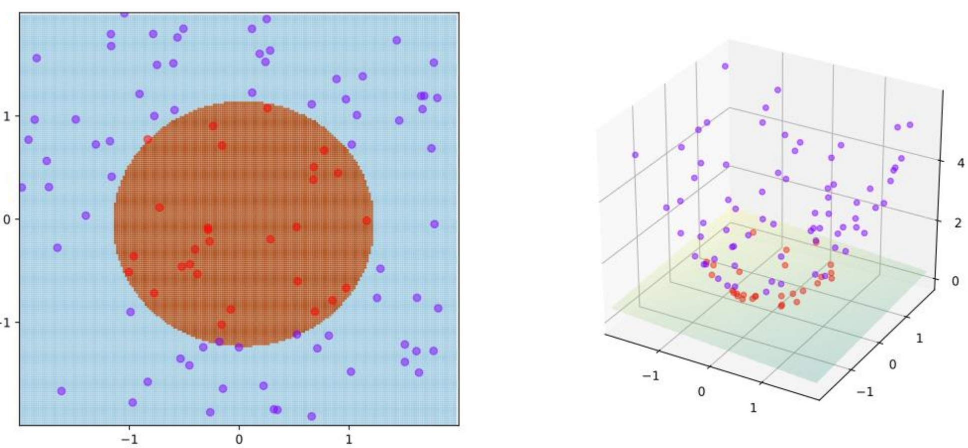


Figure 2.3A training example of SVM with kernel given by φ((a, b)) = (a, b, a2 + b2) (Source: Wikipedia).

In Figure 2.3, SVM with kernel given by φ((a, b)) = (a, b, a2 + b2) and thus K(x, y) = x•y + x2 y2. The training points are mapped to a 3-dimensional space where a separating hyperplane can be easily found.

Computing SVM classifier

Computing the (soft-margin) SVM classifier amounts to minimizing an expression of the form

$$\left[\frac{1}{n}\sum_{i=1}^{n}\max\left(0,1-y_i(\omega.x_i-b)\right)\right]+\lambda\|\omega\|^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.28)$$

Here the parameter λ determines the tradeoff between increasing the margin-size and ensuring that the $\vec{x_i}$ lie on the correct side of the margin. Thus, for sufficiently small values of λ , the second term in the loss function will become negligible. Hence, it will behave similar to the hard-margin SVM, if the input data are linearly classifiable, but will still learn if a classification rule is viable or not.

Primal

Minimizing can be rewritten as a constrained optimization problem with a differentiable objective function in the following way.

For each $i \in \{1,\ldots\ldots n\}$we introduce a variable $i=\max\left(0,1-y_i(\omega.x_i-b)\right)$.Note that i is the smallest nonnegative number satisfying $y_i(\omega.x_i-b)\geq 1-i$

So, we can rewrite the optimization problem as follows

$$\frac{1}{n}\sum_i^n i+\lambda\|\omega\|^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.29)$$

Subject to $y_i(\omega.x_i-b)\geq 1-i\ \geq 0$ for alli.

Dual

By solving for the Lagrangian dual of the above problem, one obtains the simplified maximize $(c_{1\ldots\ldots\ldots}\ c_{n)}$ $=\sum_{i=1}^{n}c_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}y_i\,c_i\ (x_i.x_j)\ y_i.c_j$,subject to $\sum_{i=1}^{n}y_i\,c_i=0$ and $0\leq c_i\leq\leq\frac{1}{2n\pi}$ for alli.

Modern methods

Recent algorithms for finding the SVM classifier include sub-gradient descent and coordinate descent. Both techniques have proven to offer significant advantages over the traditional approach when dealing with large, sparse datasets—sub-gradient methods are especially efficient when there are many training examples, and coordinate descent when the dimension of the feature space is high.

Sub-gradient descent algorithms for the SVM work directly with the expression:

$$f = (\vec{\omega}, b) = \left[\frac{1}{n}\sum_{i=1}^{n} \max\left(0, 1 - y_i(\omega . x_i - b)\right)\right] \lambda\|\omega\|^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.30)$$

Note that f is a convex function of $\vec{\omega}$ and b. As such, traditional gradient descent (or SGD) methods can be adapted, where instead of taking a step in the direction of the functions gradient, a step is taken in the direction of a vector selected from the function's sub-gradient. This approach has the advantage that, for certain implementations, the number of iterations does not scale withn, the number of data points.

2.1.6 Greyscale

In computing, colourimetry and grayscale or grayscale picture is one in which the estimation of every pixel is a solitary example speaking to just a measure of light, that is, it conveys just force data. Pictures of this sort, otherwise called highly contrasting or monochrome are made solely out of shades of dark, differing from dark at the weakest force to white at the strongest.

Grayscale pictures are distinct from one-piece bi-tonal highly contrasting pictures which, with regards to PC imaging, are pictures with just two hues: high contrast (likewise called bi-level or double pictures). Grayscale pictures have numerous shades of dark in the middle.

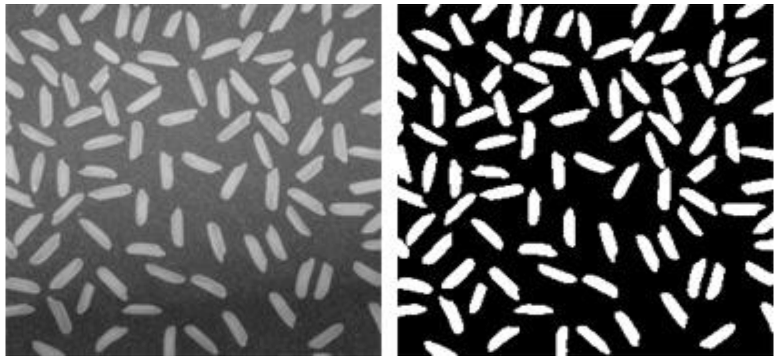


Figure 2.4Removing noise with morphological operations like image opening [25].

In Figure 2.4, Morphological erosion and dilation may also be applied on grayscale images. In that case, the morphological dilation computes for each pixel the maximum within its neighbourhood (defined by the structuring element), whereas the morphological erosion considers the minimum value within the neighbourhood.

Colour to Grayscale Converting

A common strategy is to use the principles of photometry or, more broadly, colourimetry to calculate the grayscale values (in the target grayscale colourspace) so as to have the same luminance (technically relative luminance) as the original colour image (according to its colourspace).In addition to the same (relative) luminance, this method also ensures that both images will have the same absolute luminance when displayed, as can be measured by instruments in its SI units of candelas per square meter, in any given area of the image, given equal white points. Luminance itself is defined using a standard model of human vision, so preserving the luminance in the grayscale image also preserves other perceptual lightness measures, such as L* (as in the 1976 CIE Lab colour space) which is determined by the linear luminance Y itself (as in the CIE 1931 XYZ colour space) which we will refer to here as $Y_{linear}$to avoid any ambiguity. To convert a colour from a colourspace based on a typical gamma-compressed (nonlinear) RGB colour model to a grayscale representation of its luminance, the gamma compression function must first be removed via gamma expansion (linearization) to transform the image to a linear RGB colourspace, so that the appropriate weighted sum can be applied to the linear colour components (R_{linear}, G_{linear}, B_{linear}) to calculate the linear luminance Y_{linear}, which can then be gamma-compressed back again if the grayscale result is also to be encoded and stored in a typical nonlinear colourspace.

For the basic sRGB shading space, gamma development is characterized as

$$C_{linear} = \begin{cases} \frac{C_{srgb}}{12.92} \end{cases}, C_{srgb}<0.040 \quad (2.31)$$

$$C_{linear} = \{ (\frac{C_{srgb}+0.055}{1.055})^{2.4}, C_{srgb}<0.040 \quad (2.32)$$

Here Csrgb speaks to any of the three gamma-packed sRGB primaries (R_{linear}, G_{linear}, B_{linear})each in go [0,1]) and C_{linear}, is the relating straight force esteem (R_{linear}, G_{linear}, B_{linear}), likewise in extend [0,1]). At that point, direct luminance is computed as a weighted entirety of the three straight power esteems. The sRGB shading space is characterized regarding the CIE 1931 straight luminance Ylinear, which is given by

$$Y_{linear} = 0.2126R_{linear} + 0.7152G_{linear} + 0.0722B_{linear} \quad (2.33)$$

These three particular coefficients represent the intensity (luminance) perception of typical trichromat humans to light of the precise Rec. 709 additive primary colours (chromaticities) that are used in the definition of sRGB.

2.2 Sensor Details

2.2.1 Gas Sensors

MQ135 (Carbon dioxide)

The MQ series of gas sensors or sensor modules are constructed with an electro-chemical sensor at the core of its operation. These sensors or sensor modules also have a small heater inside that is cyclically heated and let to cool when in operation to get data from the sensors. These sensors can be calibrated to some extent provided a known concentration of the gas being measured is available. These sensors produce an 8 bit analog output depending on the level of gas concentration. Later the values are converted to PPM for. The MQ135 belongs to the above mentioned MQ series gas sensors. This sensor is sensitive to the Carbon dioxide (CO2) gas.

Find Rs using the below formula:

$$R_s = \left(\frac{V_c}{VR_L - 1}\right) \times R_L \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(2.34)}$$

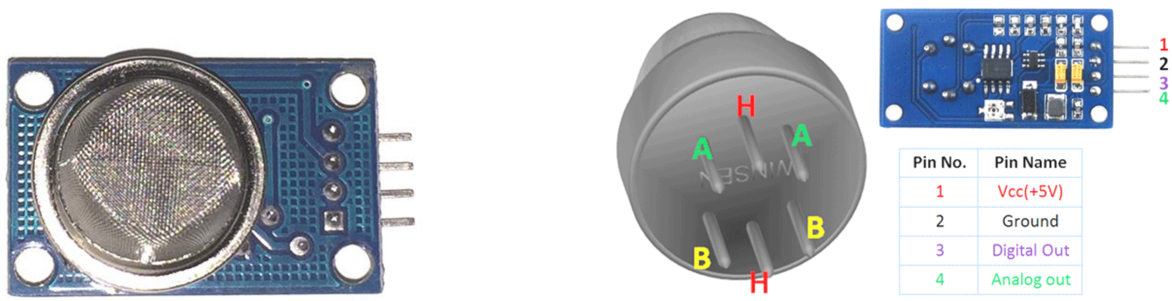


Figure 2.5 MQ135 (Carbon dioxide)(Source: Wikipedia).

MQ7 (Carbon monoxide)

This sensor has all the same characteristic and behaviour of a typical MQ series gas sensor. This certain model of MQ series gas sensor is sensitive to Carbon monoxide.

Rsgas = (5.0-sensor_volt)/sensor_volt ………………………………………………… (2.35)

Figure 2.6 MQ7 (Carbon monoxide)(Source: Wikipedia).

MQ8 (Hydrogen)

This sensor has all the same characteristic and behaviour of a typical MQ series gas sensor. This sensor is used to perceive concentration level of Hydrogen.

Figure 2.7 MQ8 (Hydrogen)(Source: Wikipedia).

Grove-ME2 (Oxygen)

The difference between this sensor and the abovementioned sensors is that the ME2 doesn't have an active heater inside. The sensor's output value has a linear relationship with the concentration level of oxygen.

Figure 2.8 Grove-ME2 (Oxygen) (Source: Wikipedia).

2.2.2 Light Sensors

Gravity Analog UV sensor V2 (Ultraviolet)

This is a UV sensor with GUVA-S12SD chip as the main chip. It has the capability of detecting wavelength of 200-370nm and a value to output linear relationship.

Figure 2.9 Gravity Analog UV sensor V2 (Ultraviolet)(Source: Wikipedia).

Gravity Analog Ambient Light sensor-SEN-0078 (Ambient light)

This sensor is somewhat like an extended version of the PT550 light sensor. It outputs an 8bit analog value depending on the light level.

Figure 2.10 Analog ambient light sensor(Source: Wikipedia).

LDR

This is the most basic light reacting sensor. It works by outputting a value of 8 bit. The output increases if the light intensity increases and vice versa. Most of the other advanced and complex light sensors are built on the basis of this sensor.

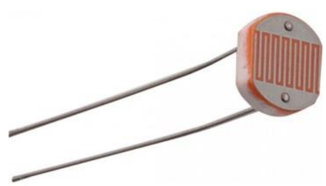


Figure 2.11 Light dependant resistor (Source: Wikipedia).

DHT22 (Moisture and Temperature)

This is a basic digital temperature and humidity sensor. It uses a capacitive humidity sensor and a thermistor to measure the surrounding air. Unlike most of the sensors used this one outputs on the digital pin. This sensor can output the value of both air temperature and humidity.

Figure 2.12 Moisture and temperature sensor (DHT22) (Source: Wikipedia).

2.2.3 Soil Sensors

DS18B20 (Temperature)

This is a 1-Wire digital temperature sensor. It reports temperature in degree with 9 to 12-bit precision having a range of -55C to 125C.

Figure 2.13 Soil temperature sensor DS18B20(Source: Wikipedia).

Soil Moisture Sensor

This a simple sensor that outputs a resistive value of it that it detects between its two probes. The presence of moisture in soil leads to a lower value as opposed to a higher value when there's absence of moisture. The change in value is linear.

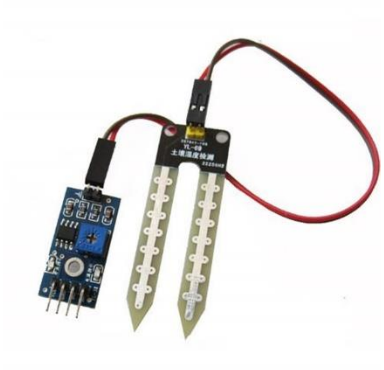


Figure 2.14 Soil moisture sensor(Source: Wikipedia).

2.2.4Justification

The CO2 and CO sensors would give us an idea about the toxicity of the air. CO is especially harmful for human as it causes CO poisoning which is quite lethal.
The Oxygen and Hydrogen values are especially useful in secluded farming environment such as greenhouse environments.

The light sensor (including LDR) & UV radiation sensor are used to measure light intensity and compare the harmful UV radiation level respectively. Keeping track of light levels could be used in flower harvesting since more colourful flowers require more light (generally) to bloom to fullest colour.

The ambient light tells us how much light the place gets.

The DHT22 helps us measure air temperature and humidity.

The soil temperature and moisture sensors help to identify the state of the soil. These values can be used to determine whether to water the crop or not. That and too high of a temperature might indicate danger and it is not wise to water crops in higher temperatures.

Chapter 3

This chapter describes how the algorithms and sensors have been used in the crop monitoring system. Starting with a flow chart of how the methods are proposed to work, it shows how the environmental data is processed and how the image detections are classified.

3. Proposed Methods

The theme is to monitor agricultural field based the sensing values of sensors and detecting healthy or affected crop field from aerial image and detecting and classify leaf diseases. The following block diagram represents the flow of the methodology of the proposed work.

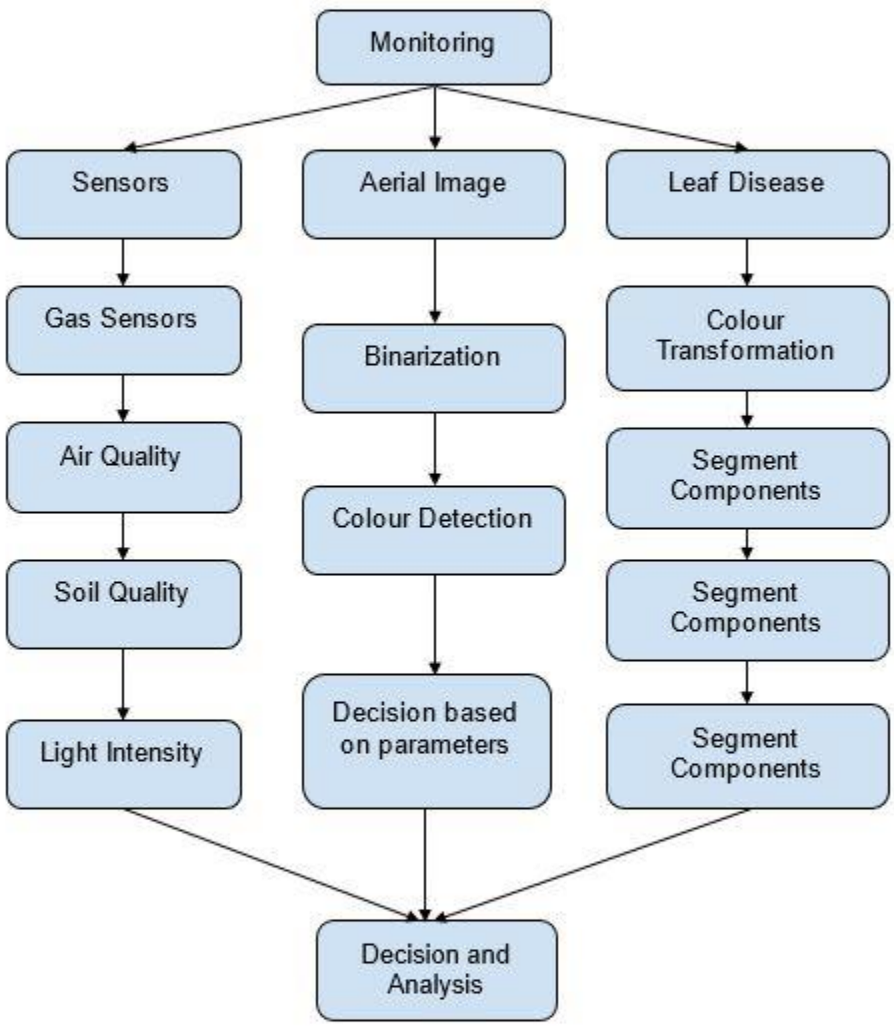


Figure 3.1Block diagram of proposed methods.

3.1 Real-Time Data Sensing

3.1.1 Gas Detection

MQ135 (Carbon dioxide): The MQ series of gas sensors or sensor modules are constructed with an electro-chemical sensor at the core of its operation. These sensors or sensor modules also have a small heater inside that is cyclically heated and let to cool when in operation to get data from the sensors. These sensors can be calibrated to some extent provided a known concentration of the gas being measured is available. These sensors produce an 8 bit analog output depending on the level of gas concentration. Later the values are converted to PPM for. The MQ135 belongs to the above mentioned MQ series gas sensors. This sensor is sensitive to the Carbon dioxide (CO2) gas.

MQ7 (Carbon monoxide): This sensor has all the same characteristic and behaviour of a typical MQ series gas sensor. This certain model of MQ series gas sensor is sensitive to Carbon monoxide.

MQ8 (Hydrogen): This sensor has all the same characteristic and behaviour of a typical MQ series gas sensor. This sensor is used to perceive concentration level of Hydrogen.

Grove-ME2 (Oxygen): The difference between this sensor and the abovementioned sensors is that the ME2 doesn't have an active heater inside. The sensor's output value has a linear relationship with the concentration level of oxygen.

3.1.2 Light

Gravity: Analog UV sensor V2 (Ultraviolet): This is a UV sensor with GUVA-S12SD chip as the main chip. It has the capability of detecting wavelength of 200-370nm and a value to output linear relationship.

Analog Ambient Light sensor-SEN-0078(Ambient light): This sensor is somewhat like an extended version of the PT550 light sensor. It outputs an 8bit analog value depending on the light level.

LDR: This is the most basic light reacting sensor. It works by outputting a value of 8 bit. The output increases if the light intensity increases and vice versa. Most of the other advanced and complex light sensors are built on the basis of this sensor.

3.1.3 Air Quality

DHT22 (Moisture and Temperature): This is a basic digital temperature and humidity sensor. It uses a capacitive humidity sensor and a thermistor to measure the surrounding air. Unlike most of the sensors used this one outputs on the digital pin. This sensor can output the value of both air temperature and humidity.

3.1.4 Soil Quality

DS18B20 (Temperature): This is a 1-Wire digital temperature sensor. It reports temperature in degree with 9 to 12-bit precision having a range of -55C to 125C.

Soil moisture sensor: This a simple sensor that outputs a resistive value of it that it detects between its two probes. The presence of moisture in soil leads to a lower value as opposed to a higher value when there's absence of moisture. The change in value is linear.

The CO2 and CO sensors would give us an idea about the toxicity of the air. CO is especially harmful for human as it causes CO poisoning which is quite lethal. The Oxygen and Hydrogen values are especially useful in secluded farming environment such as greenhouse environments.

The light sensor (including LDR) & UV radiation sensor are used to measure light intensity and compare the harmful UV radiation level respectively. Keeping track of light levels could be used in flower harvesting since more colourful flowers require more light (generally) to bloom to fullest colour. The ambient light tells us how much light the place gets. The DHT22 helps us measure air temperature and humidity. The soil temperature and moisture sensors help to identify the state of the soil. These values can be used to determine whether to water the crop or not. That and too high of a temperature might indicate danger and it is not wise to water crops in higher temperatures.

3.2 Colour Segmentation Based

Images are acquired in real time through cameras. After which each image is processed to determine to overall state of the crops as a whole. The entire field is separated in different user defined region based on colour. There is no limit to how many different types of segmentation can be done; it is entirely up to the user. To achieve this, after each image is captured it is transformed from Red-Green-Blue(RGB) colour model to Hue-Saturation-Value(HSV) colour model. After this each image is binarized using the user defined HSV parameters. To do this each pixel is matched with the user defined band of HSV value which consists of a lower limit of HSV and a higher limit of HSV. If the pixel in observation is within the user defined HSV range then it is treated as true and false otherwise. For each image captured from the camera it is separately and independently binarized using all the user defined HSV parameters. After binarizing the image with all the user defined parameters the image is effectively separated into multiple regions that represent a unique state of the crop. This is used to quickly identify which is the dominant state of the crop at the given time.

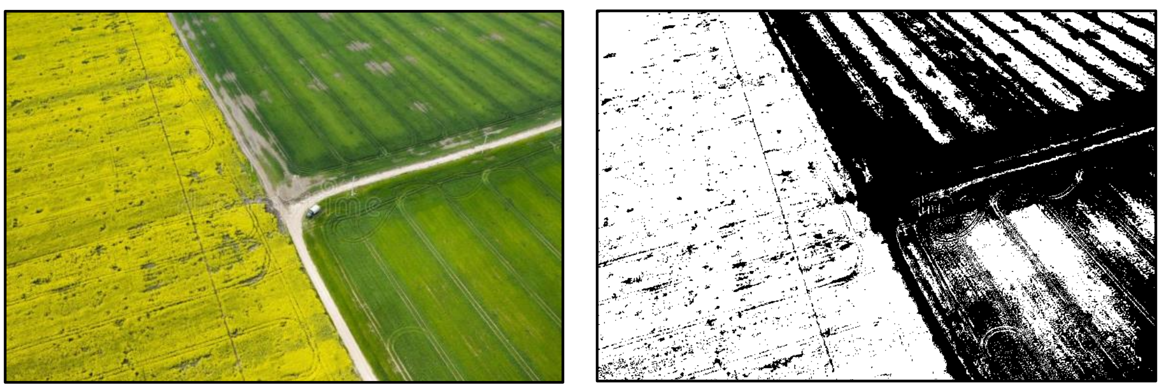


Figure 3.2Aerial Image Segmentation.

3.3 Leaf Diseases Detection and Classification

The system's one of core is to detect and classification leaf diseases based on L*a*b* colour segmentation and classification using support vector machine .It contains 5 kinds of leaf diseases dataset named Anthranose, Alternata ,Bacterial Blight, Cercospora Leaf Spot, Blackspot.

3.3.1 Dataset Preprocessing

In the pre-processing stage, 30 Russian blurring and grey scaling is used to minimize segmentation error. The system approaches to convert RGB colour L*a*b colour space. http://www.aishack.in/tutorials/image-moments/After colour space transformation, filters are applied for desired enhancements, like better contrast and brightness. Noise occurrence is very general, thus such systems popularly use median and rank filters. Laplacian filter is used for sharpening. Apart from these, techniques like histogram equalization and Gabor wavelets are also used for filtering and controlling varying lighting conditions.

3.3.2 Extracting the Affected Regions

In this step we identify the mostly green coloured pixels. After that, based on specified threshold value that is computed for these pixels, the mostly green pixels are masked as follows: if the green component of the pixel intensity is less than the pre-computed threshold value, the red, green and blue components of the this pixel is assigned to a value of zero. This is done in sense that the green coloured pixels mostly represent the healthy areas of the leaf and they do not add any valuable weight to disease identification. Furthermore this significantly reduces the processing time. Now, the pixels with zeros red, green, blue values were completely removed. This is helpful as it gives more accurate disease classification and significantly reduces the processing time. From the above steps, the infected portion of the leaf is extracted. The infected region is then segmented into a number of patches of equal size. The size of the patch is chosen in such a way that the significant information is not lost. In this approach patch size of 32×32pixels is taken.

The next step is to extract the useful segments. Not all segments contain significant amount of information. So the patches which are having more than fifty percent of the information are taken into account for the further analysis. Grey-levels occur in relation to other grey levels .These matrices measure the probability that a pixel at one particular grey level will occur at a distinct distance and orientation from any pixel given that pixel has a second particular grey level. The SGDM's are represented by the function P (i, j, d, θ) where i represent the grey level of the location (x, y), and j represents the grey level of the pixel at a distance d from location (x, y) at an orientation angle of θ. SGDM's are generated for H image.

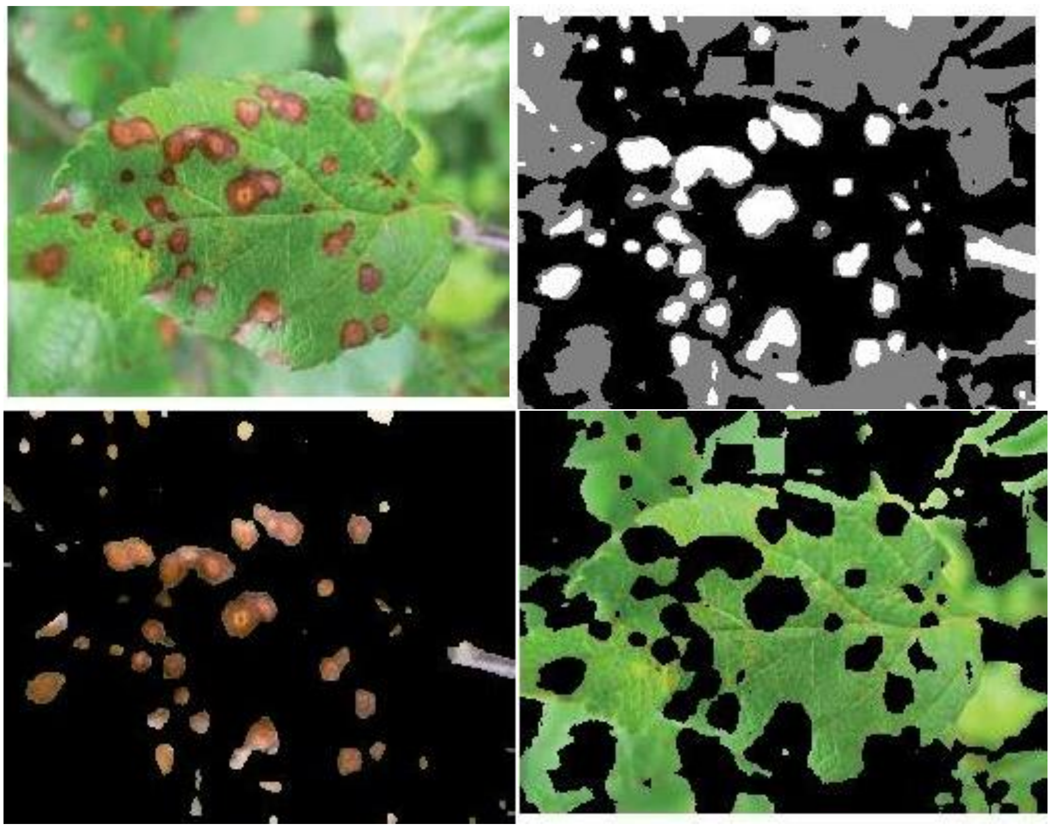


Figure 3.3Leaf Diseases detection.

3.3.3 Leaf Diseases Classification Using SVM

Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. Supervised learning involves analysing a given set of labelled observations (the training set) so as to predict the labels of unlabelled future data (the test set). Specifically, the goal is to learn some function that describes the relationship between observations and their labels. More formally, a support vector machine constructs a hyper plane or set of hyper planes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyper plane that has the largest distance to the nearest training data point of any class (so-called functional margin), in general the larger the functional margin the lower the generalization error of the classifier. In the case of support vector machines, a data point is viewed as a p-dimensional vector (a list of p numbers), and we want to know whether we can separate such points with a (p – 1)-dimensional hyper plane. This is called a linear classifier. There are many hyper planes that might classify the data. One reasonable choice as the best hyper plane is the one that represents the largest separation, or margin, between the two classes. So we choose the hyper plane so that the distance from it to the nearest data point on each side is maximized. Multiclass SVM aims to assign labels to instances by using support vector machines, where the labels are drawn from a finite set of several elements. The dominant approach for doing so is to reduce the single multiclass problem into multiple binary classification problems. Common methods for such reduction include: building binary classifiers which distinguish between (i) one of the labels and the rest (one-versus-all) or (ii) between every pair of classes (one-versus-one). Classification of new instances for the one-versus-all case is done by a winner-takes-all strategy, in which the classifier with the highest output function assigns the class.

Chapter 4

4. Results and Analysis

The quality of air and soil; density of gases and lux of light depend on area to area, as a result some areas has high rate of harvesting, some areas do not have. Which crops will go well in which area, its condition and leaf diseases detection and classification are done by making Arduino based sensor device system, Colour Segmentation based aerial image analysis and using support vector machine for leaf image classification.

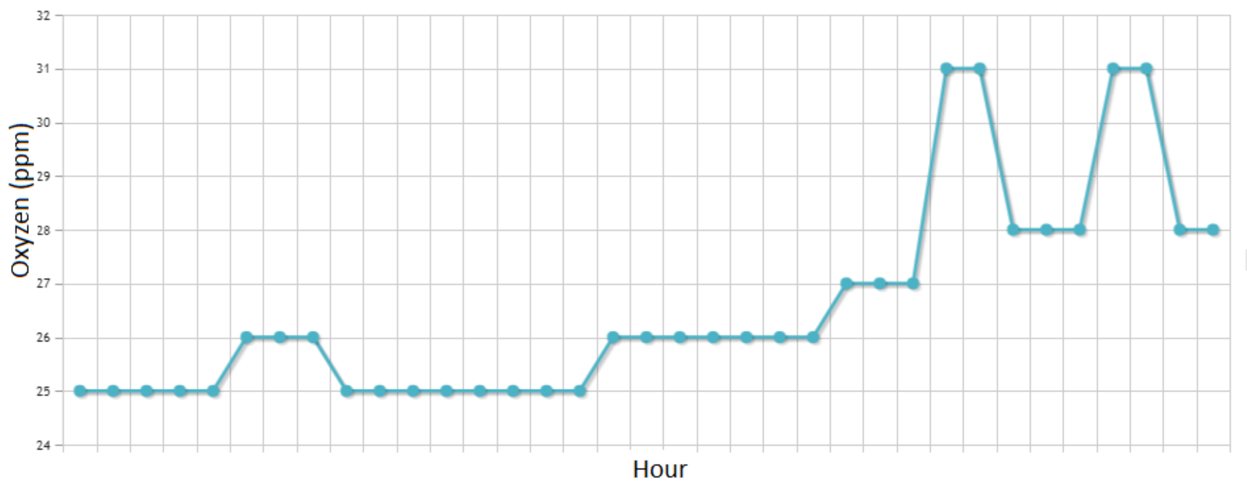


Figure 4.1 Obtain Oxygen ppm values.

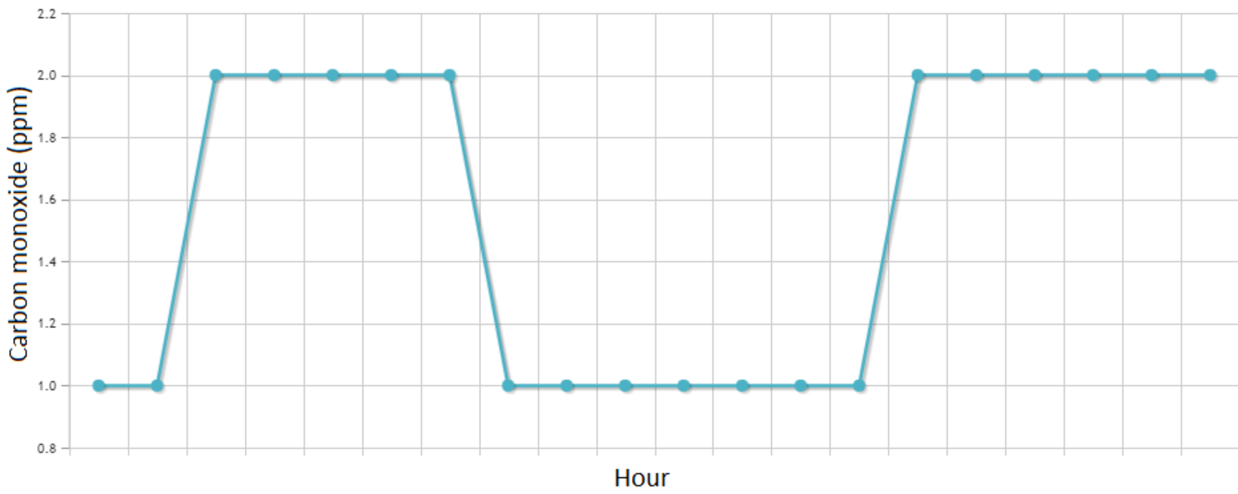


Figure 4.2 Obtain Carbon Dioxide ppm values.

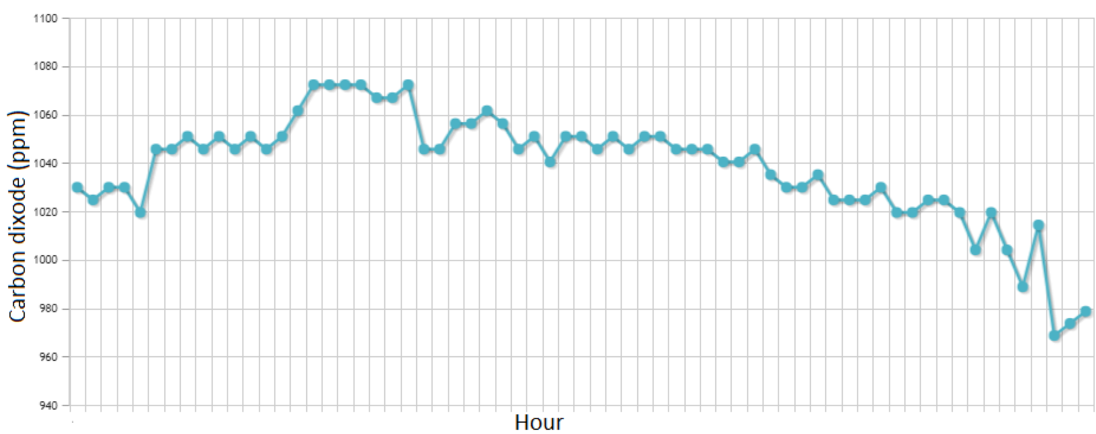


Figure 4.3 Obtain Carbon Monoxide ppm values.

Figure 4.1, Figure 4.2 and Figure 4.3show the values gathered by setting up small custom weather monitoring system around an arbitrarily chosen spot. The values were all read and logged in real tile over several days just so sufficient and stable data can be collected. Once the data were collected an arbitrary segment of the data was plotted in a graph with the Y-axis denoting the PPM (Particle per Million) and the X-axis denoting time in hour.

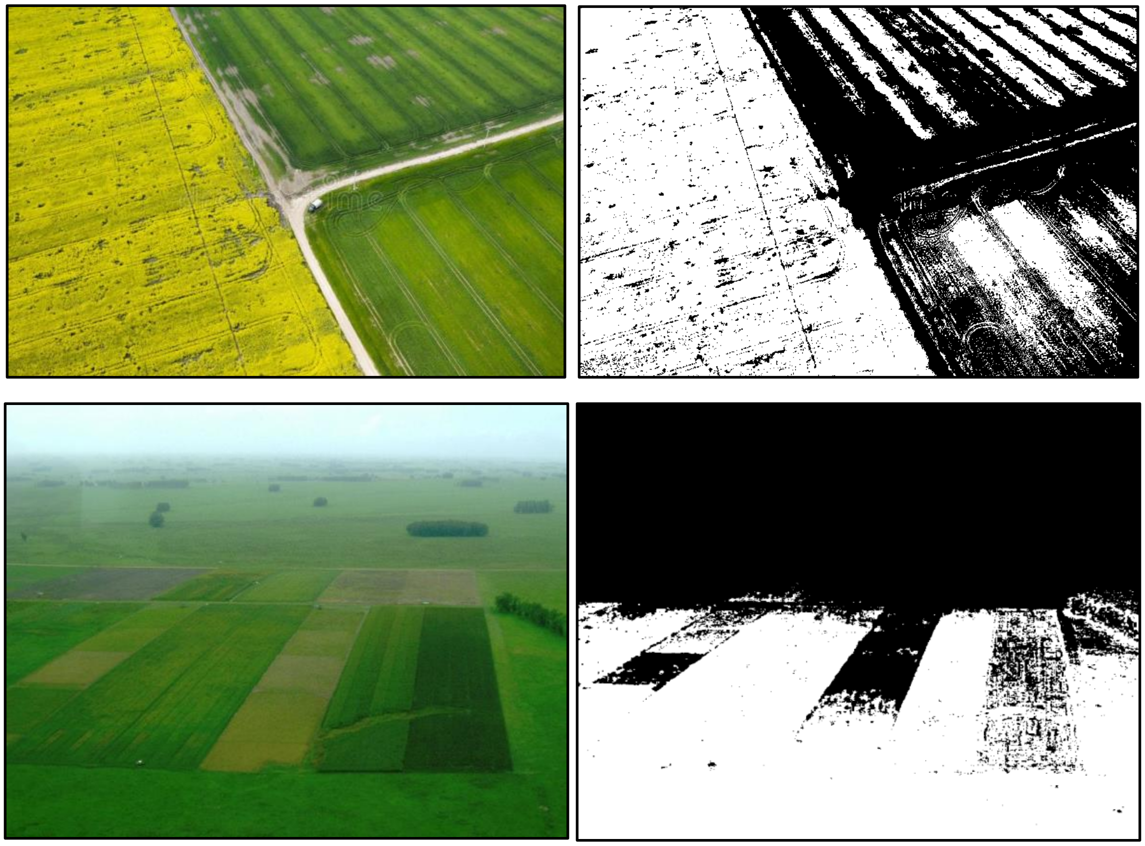
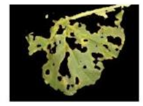
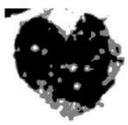
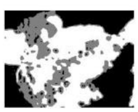
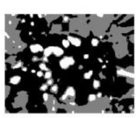
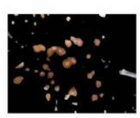
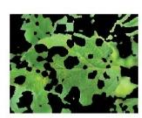


Figure 4.4 Aerial Image Segmentation (White = Dried/unhealthy,
Black = Green healthy crop field).

In Figure 4.4, there are two aerial images of large crop field with portion them ripe and portion of them not ripe. The unripe portion of the filed could indicate either just normal and plain unripe crop or it could also indicate an area affected by disease. After binarizing the image into either ripe or unripe section the images become the black and white images on the right. The binarizing process is done on each single pixel so that the proportion of ripe to unripe or affected crop region would properly scale with increased or decreased pixel settings. The apparent outliers in secluded or clustered regions were gotten rid of or smoothed out because they contribute to the overall proportion. Moreover they also could indicate the inception of a disease in a ripe region or the starting of the crops turning ripe in an unripe region. So even though the images look quite harsh it is actually useful for the overall calculation. For each image captured from the camera it is separately and independently binarized using all the user defined HSV parameters. After binarizing the image with all the user defined parameters the image is effectively separated into multiple regions that represent a unique state of the crop. This is used to quickly identify which is the dominant state of the crop at the given time.

The detected cluster leaf's affected regions are compared with a dataset using support vector machine and flowing affected results are gained, shown in Figure 4.5.

Main Image	Clustered	ROI	Masked ROI	Disease	Accuracy
				Anthracnose	96.67%
				Anthracnose	83.94%
				Cercospora Leaf Spot	84.44%
				AlternariaAlternata	91.12%
				Black Spot	98.15%
				Black Spot	89.71%

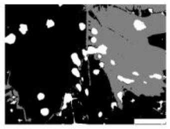
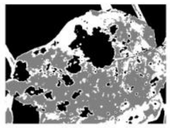
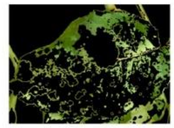
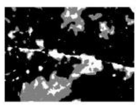
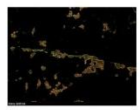
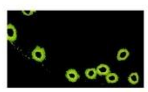
				Cercospora Leaf Spot	86.67%
				Anthracnose	86.19%
				Bacterial Blight	93.69%
				Bacterial Blight	96.54%

Figure 4.5 Different types of plants with different diseases and the accuracy of detecting those diseases.

A large number of pictures of leaves with probable diseases are taken which is called the main image. The images are then clustered using K-mean Cluster algorithm. Then we applied L*a*b* colour space to get a fixed colour, for instance grey, which have been learned from huge leaf dataset. The reason of interest is detected from the fixed colour image and masked for further accuracy. The images from the dataset are compared layer by layer using Support Vector machine (SVM) to classify what percentage of the leaf have been affected and what is the affected part's colour. Finally, the accuracy of the disease is found by comparing the affected parts of the sample leaf with many other leaf parts having the same disease. The accuracy of this system ranges from 83 to 98 percent.

On a research paper, titled "Plant Disease Identification: A Comparative Study" by S. C. Madiwalar and M. V. Wyawahare, the overall results gave a classification accuracy of 79.16% and 83.34% for Minimum Distance Classifier and Support Vector Machine respectively over a database of 86 images[28].

In this research, 86 images have been used to test the result and run them through K-means Cluster for clustering, L*a*b* for segmentation and then masking. Additional data from the MAT Lab which contains more than 5000 sets of data that have been compared layer by layer using Support Vector Machine (SVM). The algorithms ran smoothly enough to provide accuracies of 83% to 98%, where 20 images were identified having Alternaria Alternata disease, 23 having Anthracnose disease, 6 having Bacterial Blight disease, 9 having Cercospora Leaf Spot, 21 with other different types of diseases and 15 were found to be healthy with no diseases. Therefore this system has given better yield and is more trustworthy in terms of classifying diseases.

Chapter 5

5. Conclusion and Future Work

5.1 Conclusion

With increasing population, the agricultural sectors have also been increasing. It is becoming a matter of concern about how to manage such a growth in productivity. Many people have come up with various solutions to farming, but in this research, two major ways to monitor crops have been proposed to trigger accuracy. One of them being image processing through powerful algorithms like SVM and another by utilizing sensors to obtain physical data of the plants' current state directly. Early stages of disease detections will result in better yield, safer food and save a maximum amount of vegetation producing a greener country. Systems like this are not anymore just a dream, but a necessity in the world of agriculture.

5.2 Future Work

The rise of technological advancements have enabled the world to move at a faster pace and provided the space for additions in the fields of technology. Therefore focusing on agricultural industries, this system can be improved in the future by building a central server to monitor the system through means of remote access and through software applications. It can be done by operating a drone over the crop fields to capture real time aerial pictures and process them instantly and send to the central server where the results of the detections can be projected. This project can not only run on plants, but can also be used for detecting fruits diseases. An implementation of artificial intelligence can be utilized to learn more about the agriculture in particular regions by registering huge amounts of data sets for further accuracy along with profound predictions of early stages of diseases and drastic changes taking place within the plants due to alterations in climate. Moreover, machine learning methods such as add deep learning can be executed for processing data faster and more accurately with the help of robots.

References:

[1] Gondal, Danish & Niaz Yasir, “Early Pest Detection from Crop using Image Processing and Computational Intelligence”, 2015, FAST-NU Research Journal ISSN: 2313-7045.1, 2015, pp. 242-245.

[2] B. Patil and M.h. Shubham,” Plant Monitoring Using Image Processing, Raspberry Pi& Iot”, 2017, International Research Journal of Engineering and Technology,2017, pp. 326-329.

[3] S. D. Khirade and A. B. Patil, "Plant Disease Detection Using Image Processing”, *2015 International Conference on Computing Communication Control and Automation*, Pune, 2015, pp. 768-771.

[4] Savita Netal,” Detection and Classification of Plant Leaf Diseases Using Image Processing Techniques”,2015, A Review, GHRCEM, Department of Electronics and Telecommunication Engineering, Wagholi, Pune, 2015, pp. 365-366.

[5] Singh.vijai,” Detection of plant leaf diseases using image segmentation and soft computing techniques”, 2016, MNNIT Allahabad, UP, India, 2016, pp. 445-447.

[6] G. Xin, C. Ke and H. Xiaoguang, “An improved Canny edge detection algorithm for colour image”,2012, IEEE 10th International Conference on Industrial Informatics, Beijing, 2012, pp. 113-117.

[7] I. Garg and B. Kaur, "Color based segmentation using K-mean clustering and watershed segmentation," *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, 2016, pp. 3165-3169.

[8] L. Vincent, “Morphological grayscale reconstruction in image analysis: applications and efficient algorithms”, IEEE Transactions on Image Processing, vol. 2, no. 2, Apr 1993,pp. 176-201.

[9] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt and B. Scholkopf,” Support vector machines”, IEEE Intelligent Systems and their Applications, vol. 13, no. 4, July-Aug. 1998,pp. 18-28.

[10] M. D. Malkauthekar," Analysis of euclidean distance and Manhattan Distance measure in face recognition", 2013, Third International Conference on Computational Intelligence and Information Technology, Mumbai, 2013, pp. 503-507.

[11] G. A. Wilkin and X. Huang,"K-Means Clustering Algorithms: Implementation and Comparison", 2007, Second International Multi-Symposiums on Computer and Computational Sciences IMSCCS , Iowa City, IA,2007, pp. 133-136.

[12] J. Qi, Y. Yu, L. Wang and J. Liu,"K*-Means: An Effective and Efficient K-Means Clustering Algorithm", 2016, IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom), Atlanta, GA, 2016, pp. 130 - 141.

[13] S. Chebbout and H. F. Merouani, "Comparative Study of Clustering Based Colour Image Segmentation Techniques", 2012, Eighth International Conference on Signal Image Technology and Internet Based Systems, Naples, 2012, pp. 839-844.

[14] Y. A. Sari, R. V. HariGinardi and N. Suciati, "Color correction using improved linear regression algorithm", 2015, International Conference on Information and Communication Technology and Systems (ICTS), Surabaya, 2015, pp. 73-78.

[15] M. Shrivastava and P. Matlani," A smoke detection algorithm based on K-means segmentation", 2016, International Conference on Audio Language and Image Processing (ICALIP), Shanghai, 2016, pp. 301-305.

[16] Definition of wireless sensor, available from omega.com/prodinfo/wirelesssensors.html.

[17] A. Chandrakasan, R. Amirtharajah, S. Cho, J. Goodman, G. Konduri, J. Kulik, W. Rabiner and A. Wang, "Design considerations for distributed micro-sensor systems", Proceedings of the IEEE 1999 Custom Integrated Circuits Conference, San Diego, CA, May 1999, pp. 279–286.

[18] I.A. Essa, "Ubiquitous sensing for smart and aware environments",2000, IEEE Personal Communications, October2000, pp. 47–49.

[19] https://www.arduino.cc/en/Main/arduinoBoardMega2560 (Accessed on April 25, 2016)

[20] GulbagA.TemurtasF, “A study on quantitative classification of binary gas mixture using neural networksandadaptiveneurofuzzyinference systems”,2014,SensActuators B Chem 2014; pp.115, 252–256.

[21] Capone S, Siciliano P, Bârsan N, Weimar U andVasanelli L, “Analysis of CO and CH4 gas mixtures by using a micromachined sensor array”,2011, Sens Actuators B Chem 2011; pp. 78,40–48.

[22] Lee D, Jung H, Lima J, Leeb M, Banb S andHuhc J, “Explosive gas recognition system using thick film sensor array and neural network”,2010, Sens Actuators B Chem 2010, pp. 90–98.

[23] Jing Li, Yingchun Shi, Wenhui Chang andXufang Bo, “Application Research on Gas Detection with Artificial Olfactory System”,2012, Sens Actuators 2012, pp. 3, 233–254.

[25] Khalaf W andStrangio MA, “Sensor array system for gases identification and quantification”, 2013, Recent Advances in Technologies, Ed. Rijeka Croatia InTech, 2013, pp. 37-38.

[26] Y. Xie, GaoZhifei, Tang Ke, ZengMingjie and Wang Yonghai,” Wind turbine gearbox condition monitoring system based on vibration signal”, 2015, 12th IEEE International Conference on Electronic Measurement and Instruments (ICEMI), Qingdao, 2015, pp. 159-163.

[27] Kulis, Brian, Jordan, Michael I, “Revisiting k-means: new algorithms via Bayesian nonparametric", 2012, ICML, ISBN 9781450312851, pp. 131–138.

[28] S. C. Madiwalar and M. V. Wyawahare, "Plant disease identification: A comparative study," 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), Pune, 2017, pp. 13-18.