ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

Estimation of Nitrogen and Chlorophyll Contents of Leaves Using SVM

Shraddha Sahurkar[1], Prof. B. J.Chilke[2]

M.Tech Student, Dept. of Electronics and Comm. Engg, SDCOE ,Wardha, Nagpur University, Maharashtra, India [1]

Assistant Professor (M.Tech), Dept. of Electronics and Comm. Engg, SDCOE , Wardha, Nagpur University, Maharashtra, India [2]

ABSTRACT: Chlorophyll and Nitrogen are dependent on each other. Estimating one will give content of other. Many approaches are discovered to estimate these two. Hyper spectral data, vegetation indices, remote sensing, multi spectral imaging data, light absorbance approach are some of the estimating approaches. These approaches are of two categories- Destructive and Non destructive. However, Image processing technique is proving to be proficient among all these; which come under non destructive method. In this paper, we are going to study one of these approach. Leaf color is usually used as a guide for estimation of nutrient status and plant health and so to determine nitrogen and chlorophyll contents also. In this paper work regarding developing an easy and proficient automatic method for finding nitrogen and chlorophyll content in a plant based on leaf color and image processing .We also proposing a new algorithm with assuring better efficiency for classification using SVM technique.

KEYWORDS: Digital image processing; k -means clustering; GLCM; Nitrogen; Chlorophyll; Support vector machine.

I. Introduction

The main tenure of India is agriculture, Indian soil is comprise of many minerals and organic elements, and inspection has resolved that soil and SVM classifier. All plants require adequate supplies of macronutrients for healthy growth, and nitrogen is a nutrient that is heavily available in Indian agricultural soil and which should not be in restricted supply. It is manual and time devastating. Plants, like all other living things, need food for their growth and development. Plants demand 16 fundamental elements. Carbon, hydrogen, and oxygen are derived from the atmosphere and soil water. The remaining 13 fundamental elements (nitrogen, iron, phosphorus, calcium, magnesium, sulphur, zinc, manganese, potassium, copper, boron, molybdenum, and chlorine) are supplied either from soil minerals and soil organic matter or by organic or inorganic fertilizers. Nitrogen is one of the ample mineral which plays an important role in yield of crops. Nitrogen (N) is a major element for plant growth and is a radical part of chlorophyll (Ch), which is primary absorber of light energy needed for photosynthesis. Ch and N affects the green color of plants and ultimately determines their biomass yield and quality. Plants adequately supplied with N are green and healthy, while plants inadequately supplied with N are pale green or yellow in color and remain small and retarded. Hence, leaf color changes have led researchers to exploit this property by using image processing analyses to detect Ch and N status in plants if there is deficiency in the content then proper measures can be taken by farmers to improve the nutrients in crops. Present way to find nitrogen content is kjeldhahl method. Nitrate test strip and nitrate meter are other existing systems.This consumes time, man power and costlier. Digital image processing is superior to manual process hence we will be able to save time and human error. Computer algorithms are used for texture analysis. Digital image processing is superior to manual process hence we will be able to save time and human error. Thus it will be helpful in guiding the need of type and the amount of the pesticide which will be very helpful in agriculture industry. The quality and quantity of crop yields are related to its nutritional availability. Over-fertilization cause environmental issues while under-fertilization cause yield reduction and poor yield quality. Various image processing tools and approaches were widely used in order to identify and detect various contents in plant leaves. The aim of this paper is to help farmers in predicting the exact value of nitrogen and chlorophyll content of leaves using support vector machine classifier so as to increase the efficiency and prediction accuracy in comparison with the other approaches.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

II. RELATED WORK

Mr. Dalgade Viren Suryakant [1] in this Contracted work, they had estimated the nitrogen content and calculating the nitrogen deficiency in pomegranate leaves. They collect different Nitrogen deficient leaves. They had measured the chlorophyll content of the collected leaves. We captured the images of collected leaves under the closed environment. These leaves are sent to the chemical analysis for the nitrogen estimation. Extracting the statistical features of images and creating the database. The captured images are compared with database and then find the nitrogen deficiency of leaf. For irrigated crops, plant analysis can be used as an aid in making decisions about nutrient applications such as nitrogen and some. To avoid this, various fertilizers are available in market which proportionates the quantity of these components. One example is petiole testing in irrigated potatoes. Nitrate nitrogen levels in the potato petiole are determined weekly, and the information is used to help make nitrogen fertilization decisions all season long. Plant analysis is also used in fruit and vegetable crops as a guide for nutrient application during the season.

The purpose of this study [2] is to estimate N of paddy build on leaf reflectance using Artificial Neural Network (ANN). In this study, 45 leaf samples were randomly selected under various environmental conditions. Leaf reflectance was measured by handheld spectrogram diameter while actual leaf N content was determined by Kjeldahl method. Spectral reflectance data in visible band (400–700 nm wavelength region) and actual N content were used as input and target data in ANN model building. K-fold cross validation (k=3) method was applied to select the best model and measure the overall accomplishment of model. Results indicated that ANN model with 17 neurons of hidden layer in relatively could estimate N properly. It was shown by the lowest root mean square error (RMSE) of 0.23 and the highest exactitude of 93%. This study promises to help farmers predicting N content of paddy for optimal N fertilizer application.

In this research[3] V. K. Tewari basically use to estimate the status of the plant nitrogen content / chlorophyll content in the field condition, so as to avoid the intricacies involved in other method such as chemical analysis. An experimental setup was developed to fulfill both the requirements in field condition; this was able to produce the constant illumination (artificial) and a uniform background for the image. These features were correlated with SPAD reading, which represent nitrogen / chlorophyll content of plant. Regression model were developed between various image feature and the plant nitrogen content. When this model was tested, the minimum accuracy was found to be 65% with an average accuracy of 75%, actual and predicted values of nitrogen percent were linearly correlated with *R*2 value (0.948).
These results show that the plant nitrogen content can be successfully estimated using its color image feature.

John William ORILLO's this paper[4] presents a program which identifies the 4-panel LCC equivalent of rice plants using image processing techniques and pattern recognition of the Back propagation neural network. Images of the healthy leaves were captured by digital camera and processed through RGB acquisition, color transformation, image enhancement, image segmentation and feature extraction procedures. The extricated features were computed using basic statistical methods, then served as the input to the neural network for LCC panel identification. Thirty (30) samples of IRR 82372H – Mestiso 26 variety were tested; divided into three sets with 10 leaf samples per field. The system was observed to provide an accuracy of 93.33% .

III. PREPOSED METHODOLOGY

The processing scheme consists of image acquisition through digital camera or scanner or mobile phone. Image processing includes image enhancement, filtering of image to remove noise ,image segmentation, feature extraction and comparison through SVM classifier.

Image Acquisition:-

The processing scheme consists of image acquisition through digital camera or scanner or mobile phone. Image processing includes image enhancement, filtering of image to remove noise etc .image segmentation, feature extraction and comparison through SVM classifier. Image Acquisition is one of the important phases in image classification system. The image acquisition is done using a digital camera and it is loaded and saved using JPEG format in the current directory. It supports file formats such as GIF (TIFF), JPG (JPEG), BMP (bitmap).Image should be of 256x256 dimensions, while image acquisition the image will be resized and saved in 300x400 dimensions.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

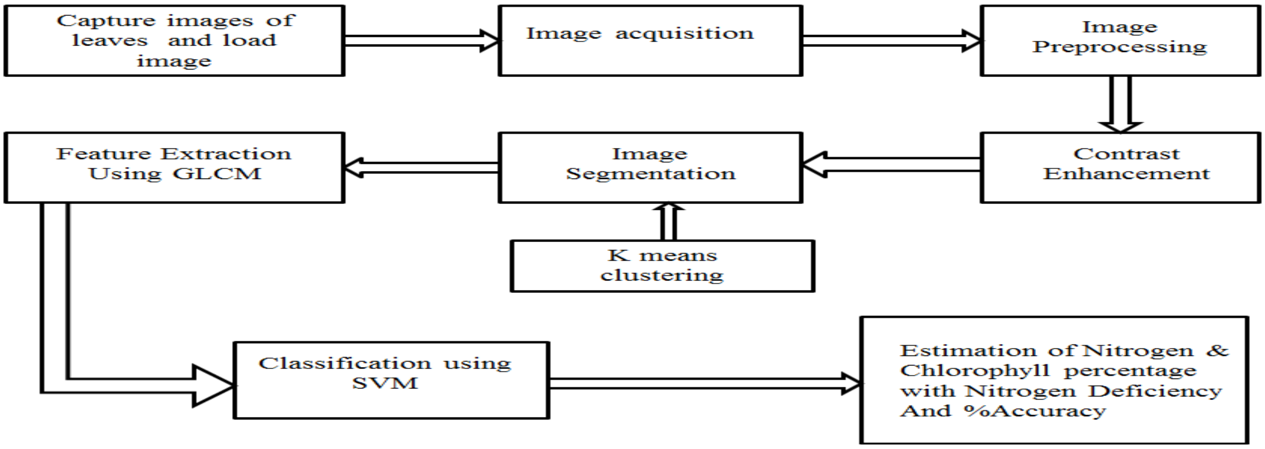


Image Preprocessing:-

Second step is to improve the database of images that suppress undesired distortion. Enhance image feature is important for further processing and analysis task .It includes color space conversion, image enhancement for contrast improvement, image resizing, filtering to remove noise etc.

Steps Used here are:

- Color space transformation: That is RGB image is transferred in to gray scale image using formulae
- Where *R, G, B correspond to the colour of the pixel, respectively.*
- $Gray = 0.2989 * R + 0.5870 * G + 0.1140 * B$ *(7)*
- ***Image Resizing***: In this work, all images must be with the same size and equal dimension. So, the gray image should be resized to equal dimensions.
- ***Image enhancement***: The Image enhancement process involves noise reduction and contrast adjustment. The value colour plane had its intensity adjusted so that the darker parts of the image goes Darker to aid the image segmentation process .

Image Segmentation:-

Segmentation means partitioning of images into various part or region and extracting meaningful region known as region of interest (ROI).The level to which subdivision is carried depends on the problem being solved .Segmentation can be stopped when the region of interest in an application have been separated. Segmentation exactitude determines success or failure of computerized analysis operation. So algorithm picked for segmentation should perform best for given requirement. The segmentation can be done using various methods like Otsu' method, k-means clustering, converting RGB image into HIS model etc.

K-means clustering

Clustering is a method to divide a set of data into a specific number of groups. It's one of the popular method is *k*-means clustering. In *k*-means clustering, it partitions a collection of data into a *k* number group of data11, 12. It classifies a given set of data into *k* number of disjoint cluster. *K*-means algorithm consists of two separate phases.

In the first phase it calculates the *k* centroid and in the second phase it takes each point to the cluster which has nearest centroid from the respective data point. There are different methods to define the distance of the nearest centroid and one of the most used methods is Euclidean distance. Once the grouping is done it recalculate the new centroid of each cluster and based on that centroid, a new Euclidean distance is calculated between each center and each data point and assigns the points in the cluster which have minimum Euclidean distance. Each cluster in the partition is defined by its member objects and by its centroid. The centroid for each cluster is the point to which the sum of distances from all the objects in that cluster is minimized. So *K*-means is an iterative algorithm in which it minimizes the sum of distances from each object to its cluster centroid, over all clusters.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

Feature Extraction using GLCM:-

Feature extraction plays important role for identification of object .After segmentation the area of interest Image feature includes color texture. Texture means how the color is distributed in the image, roughness, hardness of the image.

1. **GLCM generation.** Before applying GLCM method to input images, we must be sure that we can get texture feature for leaf part only. So we convert all pixels outside the leaf using mask to (NaN) "Not a Number", to be ignored during GLCM calculation. After that, GLCM has been calculated using the following angles (0◦ , 45◦ , 90◦, 135◦). So that, we have got 4 GLCM.

2. **Texture feature calculation.** The GCLM that generated in the last step were then used to calculate the texture features. A number of texture features may be extracted from the GLCM . In this work 7 texture features, as shown in equation (1) to equation (7), have been extracted as follow: Calculate GLCM using the following angles (0◦ , 45◦ , 90◦, 135◦), so we got 4 GLCM.

– Contrast:

$$\text{CONTRAST} = \sum_{i,j} |i-j|^2 p(i,j) \quad (1)$$

– Entropy:

$$\text{ENTROPY} = -\sum_{i=0}^{G-1} \times \sum_{j=0}^{G-1} P(i,j) \times \log(P(i,j)) \quad (2)$$

– Mean:

$$x = \sum x/N \quad (3)$$

Where: x =(Sometimes call the x-bar) is the symbol for the mean.
$\sum$=is the symbol for summation.
x =is the symbol for the scores.
N =is the symbol for the number of score.

– Correlation:

$$\text{CORRELATION} = \sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma i \sigma j} \quad (4)$$

– Variance:

$$S^2 == \frac{\sum_{i=1}^{n} (X_1{}^{n} - X_{av})^2}{n-1} \quad (5)$$

– Inverse Difference Moment:

$$\text{IDM} = \frac{\sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} Pij}{1+(i-j)^2} \quad (6)$$

– Skewness:

$$S = \frac{E(x-\mu)^3}{\sigma^3} \quad (7)$$

SVM Classification:-

The support vector machine (SVM) is a type of classifier that is originally a binary classification method developed by Vapnik and colleagues at Bell laboratories. The main advantages of the SVM are that it can obtain current optimal solution under finite samples; it can obtain the global optimal solution without falling into local optimums that normal algorithms have; it transforms nonlinear problems into linear problems in a higher dimension space, and the algorithm complexity is irrigated with space dimension.

SVM is one of the most useful learning algorithms that are considered as the alternative to the neural network approach. This algorithmic approach is built on distinctive characteristics analysis by analyzing the expected error minimization. This approach considered the empirical risk to improve the training procedure. The risk estimation is here build on the structural analysis so that the generalization error will be reduced. The error margin is analyzed under class deviation and build on its nearest training patterns are obtained. SVM is proficient to analyzing the separating planes and to identify the largest margin so that the support to the data points will be identified. This model is also build on the polynomial kernel representation so that the clever learning to the elements will be done and more exactitude will be obtained.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

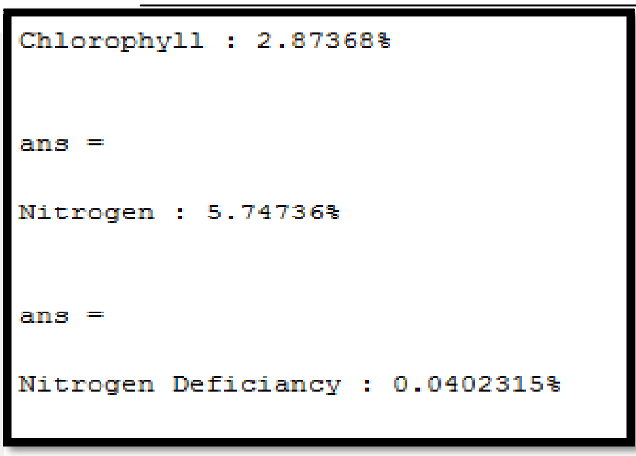
Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

The final step of this work is the classification phase. The inputs of this stage are training dataset feature vectors and their corresponding classes, whereas the outputs are the decision that determine type of input image (healthy of infected). To achieve good results, SVM was trained and tested using different kernel functions that are: Linear kernel, radial basis function (RBF) kernel, Multi-Layer Perceptron (MLP) kernel and Polynomial kernel . For the process of evaluating the obtained results, we used N-fold cross-validation technique with N = 10. Firstly,the dataset is divided into equally (or nearly equally) N-subsets. Then the cross validation process is performed N times with each sub-set being once the test dataset and all the others being the training dataset. This process is repeated N times. Hence we take the average of performance of N runs. The algorithm performance measure can be calculated as the average of the performances of 10 runs.

IV. RESULT AND DISCUSSION

As we previously mentioned, the aim of this work is to detect the nitrogen and chlorophyll content of leaf input image . In this article, firstly, we used 500 of healthy leaf images. Then, the experiment repeated used 500 of leaf images. After pre-processing phase, we calculated GLCM matrix for each of training and testing image. The GLCM are represented by the function $P(i, j, d, \theta)$, where I represents the gray level of the location (x, y), and j represents the gray level of the pixel at a distance d from location (x, y) at an orientation angle of θ. Here, we take d=1 and $\in 0^\circ$, 45° , 90°, 135°. So, we have got 4 GLCM matrixes for each image. Then we used these values to calculate 7 features. So, at the end of this phase we have got 28 features per single image. Hence, we used these feature vectors as input to the SVM classifier. SVM was trained and tested using different kernel functions that are: (Linear kernel, (RBF) kernel, (MLP) kernel and Polynomial kernel). After cropping, we convert the image from RGB to gray level color space, then resize the image to (400) pixels then removing the background. Its followed by do more enhancement on the image like fill the holes that may exist in the image and do some image erosion. Then, black space around gray image and binary mask had been cropped. Finally, final results are displayed on the command window as shown below.From the cross validation accuracy it is noticed that there is significant improvement in the accuracy if the number of training samples increases.



Project photos:- pre-processing of image Fig. 8 RESULT OF CH AND

V. CONCLUSION

This research was basically undertaken to estimate the status of the plant nitrogen content / chlorophyll content in the leaf, so as to avoid the intricacies involved in other method such as chemical analysis.Digital image processing was selected as the tool to estimate plant nitrogen content in leaf. We aimed to evaluate the applicability of a fast and noninvasive method for the estimation of total chlorophyll and nitrogen content of plants using a smart phone camera. The review of literature says about how to estimate nitrogen content in plant leaf. The estimation of nitrogen content in

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

leaf is done based on color and texture features .The review outlines several methods for estimation of nitrogen content in the leaf. The estimation methods are costlier and time consuming. The image processing methods reviewed is this paper gives a cost effective and speedier approach for estimating nitrogen content in leaves. The proposed system consisted of three main phases; pre-processing, feature extraction, and classification phases. This approach has began with image acquisition, background removal, texture feature extracting. Finally, SVM model is developed for classification stage using the extracted features and has been evaluated using N-fold cross-validation. From the experimental results we found that, the highest classification accuracy of 99.83% has been achieved using linear kernel function.

Further work can be carried out as the extension of the outlined work in the paper. Leaf image can be capture under Different lighting condition and their effect can be analyzed. Also, other features of the leaf images can be explored and evaluated for their improved representation of nitrogen content in leaves.

REFERENCES

1. Mr. Dalgade Viren Suryakant. "Nitrogen Deficiency Calculation of Leaves Using Artificial Neural Network" Proceedings ofVescomm-2016 4th National *Conference on "Recent Trendes in VLSI, Embeded System, Signal Processing and Communication*.
2. Whine Ayu Lestari1,Yeni Herdiyeni1, Lilik Budi Prasetyo "Nitrogen Estimation of Paddy Based on Leaf Reflectance Using Artificial Neural Network" 978-1-4673-9360-7/15/$31.00 2015 IEEE.
3. V. K. Tewari*, Ashok Kumar Arudra, Satya Prakash Kumar,Vishal Pandey, Narendra Singh Chandel , " Estimation of plant nitrogen content using digital image processing " July, 2013 Agric Eng Int: CIGR Journal
4. John William Orillo, Gideon Joseph Emperador#2, Mark Geocel Gasgonia, Marifel Parpan, Jessica Yang "Rice Plant Nitrogen Level Assessment through Image Processing using Artificial Neural Network" 7th IEEE *International Conference Humanoid*, Nanotechnology, Information Technology Communication and Control, Environment and Management (HNICEM) *The Institute of Electrical and Electronics Engineers Inc. (IEEE)* – Philippine Section 12-16 November 2013 Hotel Centro, Puerto Princesa, Palawan, Philippines.
5. M. M. Ali, Ahmed Al-Ani, Derek Eamus and Daniel K. Y. Tan. "An Algorithm Based on the RGB Colour Model to Estimate Plant Chlorophyll and Nitrogen Contents", 2013 *International Conference on Sustainable Environment and Agriculture IPCBEE* vol.57 (2013) © (2013) IACSIT Press, Singapore DOI: 10.7763/IPCBEE. 2013. V57. 10
6. Vasudev B. Sunagar And Team "Estimation Of Nitrogen Content In Leaves Using Image Processing",Computer Science & Engineering Department, BVBCET, Hubli – 580031, India.
7. PrajaktaPataneand AnupVibhute,"Chlorophyll and Nitrogen Estimation Techniques: A Review"*International Journal of Engineering Research and Reviews* ISSN 2348-697X Vol. 2, Issue 4, pp: (33-41), Month: October - December 2014.
8. RangaswamyH,M.G.Suraj ,"A Review of Approaches for Estimating Nitrogen Content in Plant Leaves using Image Processing Techniques"Second National Conference on Emerging Trends in Computer Science & Engineering (NCETCSE-2016), Department of CSE, BGSIT.
9. Vanschoenwinkel, B., Manderick, B.: Appropriate kernel functions for support vector machine learning with sequences of symbolic data. In: Winkler, J.R., Niranjan, M., Lawrence, N.D. (eds.) Machine Learning Workshop. LNCS (LNAI), vol. 3635, pp. 256–280. Springer, Heidelberg (2005)