

PAPER • OPEN ACCESS

Detection and classification of nutrient deficiencies in plants using machine learning

To cite this article: Anu Jose *et al* 2021 *J. Phys.: Conf. Ser.* **1850** 012050

View the [article online](#) for updates and enhancements.

You may also like

- [Implication of blanket NPK application on nutrient balance of maize based on soil and tissue diagnosis approaches in the savannas of northern Nigeria](#)
Kamaluddin T Aliyu, Alpha Y Kamara, Jeroen Huising et al.
- [Maize nutrient yield response and requirement in the maize belt of Nigeria](#)
K T Aliyu, A Y Kamara, E J Huising et al.
- [Nutrient-extended input-output \(NutriO\) method for the food nitrogen footprint](#)
Azusa Oita, Kiwamu Katagiri, Tetsuya Eguchi et al.

Detection and classification of nutrient deficiencies in plants using machine learning

Anu Jose^{1,3}, S Nandagopalan^{2,4}, Vidya Ubalanka^{1,5} and Dhanya Viswanath^{1,6}

¹Department of Information Science and Engineering,
CMR Institute of Technology (VTU-RC), Bengaluru

²Department of Information Science and Engineering,
Vemana Institute of Technology (VTU-RC), Bangalore

Email: ³ anujmattam@gmail.com, ⁴ snandagopalan@vemanait.edu.in,
⁵ arnavidya@gmail.com, ⁶ dhanyaarun0705@gmail.com

Abstract: Agriculture is the major factor contributing to Indian Economy. According to the current statistics, its contribution to GDP sector is 17.9%. Technical advancement in agricultural domain will produce more agricultural products without any wastage of money, time and manpower. Nutrients play a major role in plant growth. Lack of nutrients leads to reduced crop yield and plant growth. In this work, we are trying to create an artificial neural network model to recognize and classify the nutrient deficiency in tomato by examining the leaf characteristics. This will help farmers to adjust the nutrient supply to the plant. If soil lacks a specific nutrient, it will reflect in the physical characteristics of a leaf. The color and shape of a leaf are the two major features used for identifying the nutrient deficiency. The comparison of different segmentation schemes like hue based and threshold based schemes shows their influence in the performance of the proposed system. The influence of different activation functions in the artificial neural network is also studied in this work. The results show that the proposed method was able to classify and identify nutritional deficiencies with high accuracy.

1. Introduction

According to the statistics of 2019, India is the second largest tomato producer in the world after china. India has produced 20,708,000 tons of tomatoes in the year 2019. The major Tomato producing states in the India are Andhra Pradesh, Karnataka, Odisha, Madhya Pradesh, Gujarat, West Bengal, Bihar, Chhattisgarh, Tamil Nadu, Telangana, Uttar Pradesh, Haryana, Maharashtra and Himachal Pradesh. These states are producing about 90% of the total production of the country. India exports tomatoes to Pakistan and other Central Asian Republics. According to Department Of Agriculture, Cooperation & Farmers Welfare report[1], the tomato production is given below.



Table 1: State wise tomato production in India -2019 November

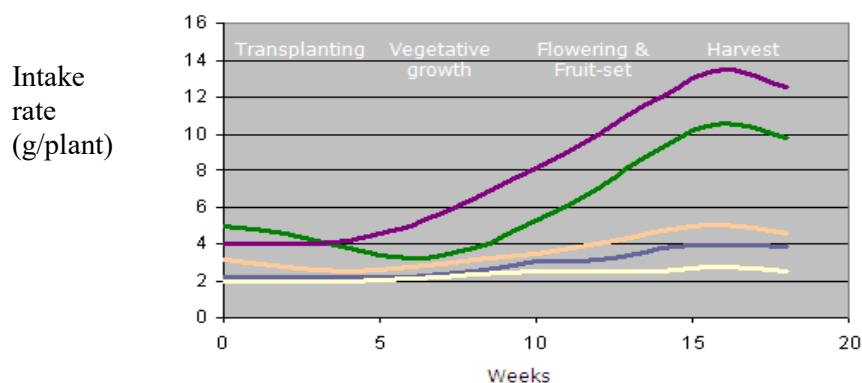
STATE/Uts	Five year Average (2013-14 to 2017-18)		2017-18		2018-19 (3 rd A.E.)	
	Production	% Share	Production	% Share	Production	% Share
ANDHRA PRADESH	2857.98	15.15	2744.32	13.89	3146.96	16.22
MADHYA PRADESH	2307.82	12.23	2419.28	12.24	2511.89	12.95
KARNATAKA	2029.47	10.76	2081.59	10.53	1775.79	9.16
GUJARAT	1321.30	7.00	1357.52	6.87	1366.57	7.05
ODISHA	1335.03	7.08	1312.07	6.64	1305.31	6.73
WEST BENGAL	1198.76	6.35	1265.25	6.40	1268.12	6.54
CHHATTISGARH	952.29	5.05	1087.33	5.50	1133.09	5.84
BIHAR	1012.07	5.36	941.56	4.77	955.57	4.93
TELANGANA	1146.26	6.08	1171.50	5.93	901.53	4.65
TAMIL NADU	564.53	2.99	887.08	4.49	845.91	4.36
UTTAR PRADESH	646.83	3.43	841.61	4.26	844.01	4.35
MAHARASHTRA	1030.04	5.46	1086.56	5.50	805.90	4.15
HARYANA	673.31	3.57	753.72	3.81	650.63	3.35

Tomato is the most delicious vegetable. It contains vitamins like vitamin A, potassium, vitamin C, folate, and vitamin K. Tomatoes are the major source of the lycopene, an antioxidant which will reduce the risk of heart disease and cancer. It is one of the most edible commodities in the world. States like Haryana, Maharashtra, Tamilnadu, Telungana, Chhattisgarh, Odisha and Karnataka have a drastic decrease in production of tomato due to many reasons. It includes different climatic factors and nutritional deficiencies. Early detection of diseases will increase the productivity. In this paper we are proposing a machine learning model which can detect the nutritional deficiencies in tomato plant, which will help farmers to monitor the use of fertilizers.

Plants are getting nutrients and minerals for its growth from soil. The lack of nutrients and water in the soil will adversely affect its proper growth. It will be visible as nutrient deficiency symptoms. These symptoms are visible in plant parts like leaves, roots, stem and even fruits. If the growing conditions are poor, plants can't take up the nutrients present in the soil. Water logging or high alkaline and acid properties of soil will adversely affect the nutrient absorption from soil.

The major symptoms of nutrient deficiencies are leaf yellowing; brown spots on a leaf and stem. This will lead to stunted growth and poor flowering and fruiting. It will reduce the yield from the crop.

In this study, we are discussing mainly about tomato plants and their nutrients requirement [2]. The different nutrients (NPK) intake is varying during its growth stages. Nitrogen and Potassium intake are slow in the beginning and gradually increasing during the flowering time. Potassium requirement is very high during fruit development. Phosphorus (P), calcium (Ca) and magnesium (Mg) are required in relatively same amount during the lifecycle of tomato plant. Nutrients requirement during tomato growth is plotted in the figure 1.

**Figure 1:** Nutrients intake during tomato growth

The nutrients functions in plant along with its deficiency symptoms and different fertilizer sources are discussed below.

1. Nitrogen

Nitrogen is the major requirement for protein creation for tomato plant. It will increase the yield and helps to achieve proper growth. N deficiency, leads to change the leaf color from green to uniformly yellow (chlorotic). The yellowing will be uniform in the entire leaf including veins. Under extreme deficiency it will change to a yellowish white color. It will reduce the branching. Nitrogen absorption will be mainly in the form of ammonium or nitrate. Apply 2 % urea solution as fertilizer for nitrogen deficiency.

2. Phosphorus

Phosphorus is essential for cell division and formation of cell and eventually leads to proper growth. The Phosphorus deficiency symptoms include necrotic spots on older leaves, plants are shortened and petiole and the under sides of the leaves. Phosphorus is absorbed by plants in the form of phosphate. pH of the soil affects the intake of phosphorus. Foliar application of Potassium sulphate (K_2SO_4) at 1% is used as a fertilizer source for phosphorus deficiency.

3. Potassium

Potassium is a very mobile component in plant growth, which helps the transport of sugars, stomata control and act as a cofactor of many enzymes. Its deficiency symptoms include tip burn, inter-venial marginal necrosis. Older leaves may wilt. In nitrogen deficiency, chlorosis is reversible but in potassium deficiency, chlorosis is irreversible if potassium is given to the plants. Plants absorb potassium in form of ions.

4. Magnesium

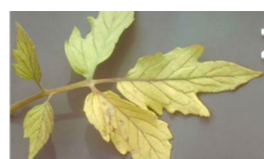
Tomatoes need an abundance of magnesium to thrive. The symptoms include mottled chlorotic areas in the interveinal tissue of leaves. Central interveinal chlorosis and green marginal bands are the major symptoms of magnesium deficiency in tomato. Foliar spray of Magnesium sulphate ($MgSO_4$) at 2% is used as a fertilizer source.

5. Calcium

Calcium is working as a major building block for creation of cell walls, and it reduces susceptibility to diseases. The classic symptoms of calcium deficiency include blossom-end rot of tomato.

6. Sulfur

It is an essential component in the creation of amino acids methionine and cystine. Plants absorb Sulfur in sulfate form. Sulfur may reduce the pH of the soil. Yellowing of leaves is the primary symptom of Sulfur deficiency. Foliar spray of Calcium sulfate (2%) at weekly intervals will reduce the Sulfur deficiency for tomato plant.



NITROGEN



MAGNESIUM



PHOSPHORUS



POTASSIUM



SULPHUR



CALCIUM

Figure 2: Different nutrient deficiency symptoms in tomato

Nutritional deficiencies in plants are a major concern which will reduce productivity and profit. Most of the time farmers can't detect the nutrient deficiency; therefore they can't take the precautions also. The work aims to classify nutritional deficiencies in tomato plant using image processing and machine learning techniques [13]. Nutritional deficiencies in tomato will be reflected in their leaves as symptoms. The colored images of tomato leaves are classified and analyzed with ANN method.

Several images of tomato leaves with deficiencies like nitrogen, potassium, phosphorous, calcium, Sulfur and magnesium are used to train the artificial neural network classifier. Then the accuracy is tested with test images.

The major objectives of the work includes 1) Collect image data set of various common tomato diseases caused by macro nutrients deficiency. 2) Apply the Color Co-occurrence Method for feature extraction in tomato leaves. 3) Implement different segmentation and feature extraction schemes on dataset. 4) Develop artificial neural network for classification of the tomato leaves based on the extracted features. 5) Apply different threshold activation functions in ANN and check their accuracies with respect to the test set. 6) Compare the accuracies in different other algorithms.

2. Related Works

Balasubramaniam and Ananthi [3] have checked the deficiency of nutrients Mg, Mo, P, Zn, K, B in several plants using Fuzzy C-Means method. The proposed method was showing an efficiency of 83%. The experiment is performed in HS images. Chen et al[4] proposed an SVM based techniques for phosphorous deficiency in rice plant by inspecting RGB based images. The developed system shows efficiency of 86%. Boron deficiency in corn plant is recognized by Luz et al., [5] by inspecting the RGB images of leaves. It was applying KNN algorithm for classification. Lee and Lee [6] proposes a Multiple Linear Regression method, which is applied on rice leaves for nitrogen deficiency detection. The system was giving accuracy of 75%. Xu et al[7] proposes a system for N, K deficiency in tomato plant by evolutionary algorithm called genetic algorithm. The RGB based images are given as the input to developed system, which was showing the accuracy of 78%.

3. Proposed System Architecture

Farmers can take a photo of the leaf which has some symptoms on leaf. Farmers can upload the nutrient deficiency suspected leaf images to the local servers through internet. The uploaded images are taken from the server and processed by the image processing system [10]. The system will have an (artificial neural network) ANN [8] based classification model. The overall block diagram is given in Figure 3.

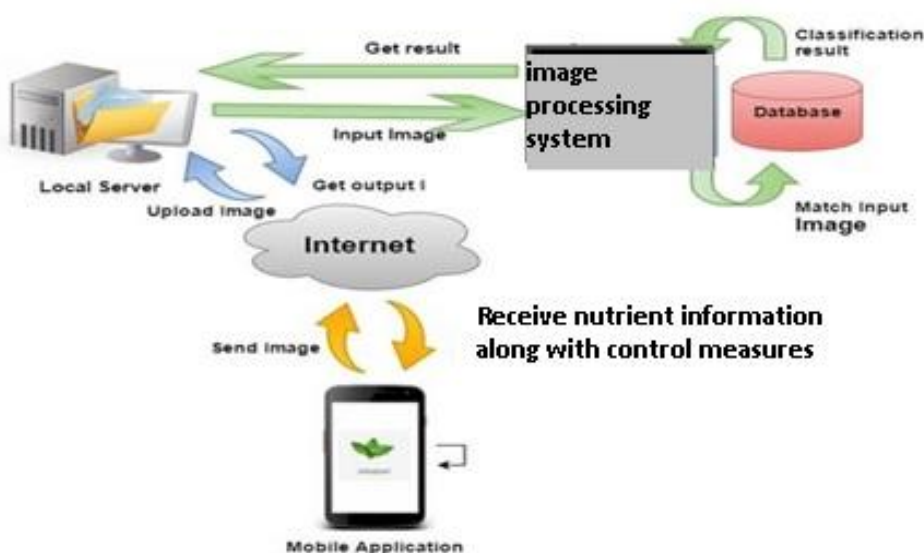


Figure 3: Overall block diagram of proposed system

It can detect the nutrient deficiency, which will be transferred back to server .The server will inform the farmer about the deficient nutrient information along with actions to be done to recover from it. The internal operations of image processing system are discussed in figure 4.

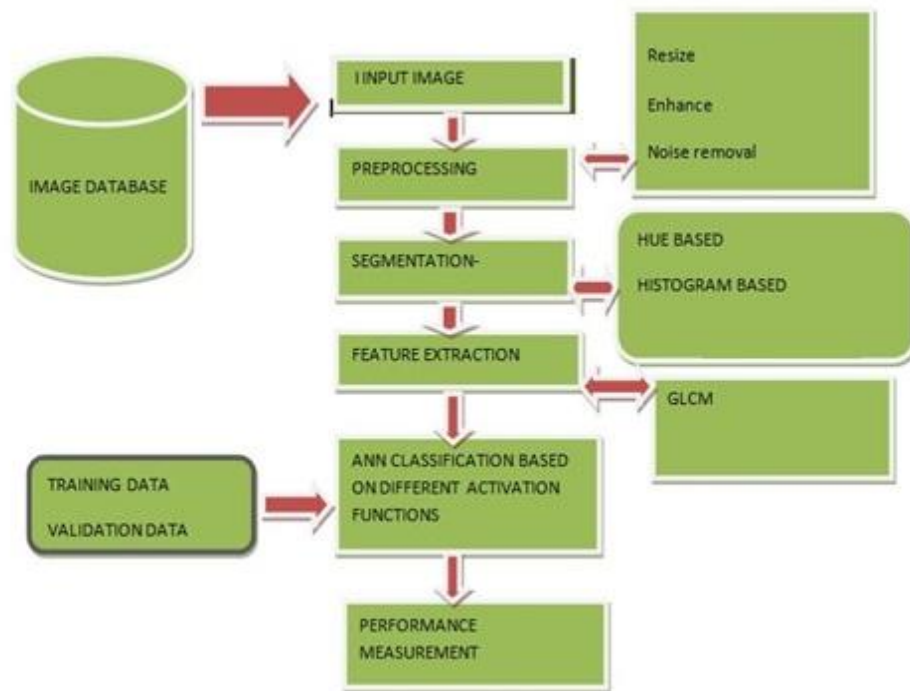


Figure 4: Basic block diagram of image processing system

Image processing is the major part of this work. It includes

3.1 Image Acquisition/Collection

Capture the details of plant leaves using camera and use bench marked datasets for study. In this study, the nutrient deficient tomato leaves are collected from various government agencies and websites. To make the dataset large enough to construct the model, different augmentation techniques are used.

3.2 Image Preprocessing

Image preprocessing is needed for train, validate and test any computer vision based model. It will suppress the unwanted information or distortions and enhance the important features needed for creating the model. The collected images will be giving to image preprocessing stage, which will enhance the accuracy of the image analysis. Three different procedures are applying as image preprocessing steps 1) Resizing the image, 2) Enhancing the image and 3) Noise removal.

a) Resizing the image

The input image size should match with the size of the stored images in the database. to make a base size for all images is a primary requirement for constructing a computer vision based model. The dimensions of the images are in the range of 560X480 for this study.

b) Enhancing the image

In this step the contrast of the images are adjusting, so that the objects will be more distinguishable from other objects and background. Histogram equalization is used to contrast in low contrast image. It is used for improving the contrast, hue and brightness of the image. Histogram representation of Magnesium deficient plant is given in Figure 5.

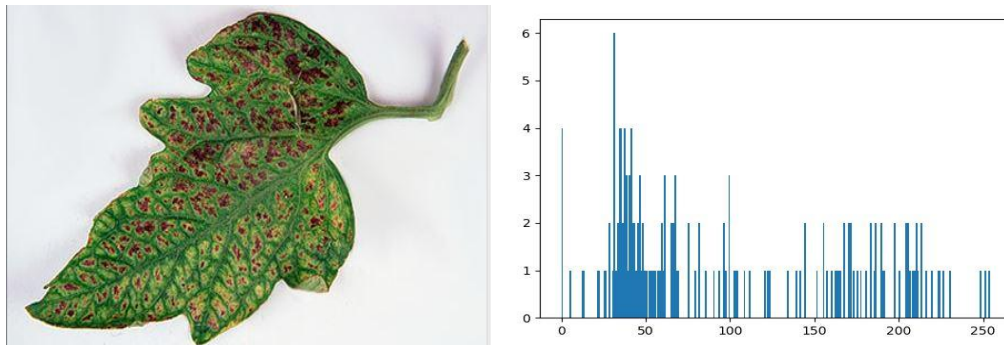


Figure5: A tomato leaf infected by Magnesium deficiency and its histogram representation
c) Noise removal

The noise in the original image will give unwanted results. Different image filters can be used for reducing the noise and enhancing the edges of the image. In this research, we are using a bilateral filtering scheme, which will preserve edges. In this smoothening scheme it replaces the intensity of each pixel's value by its neighbouring pixel's weighted average.

3.3 Image Segmentation

Image segmentation in RGB space, need lot of computational time even for small images. The dimensions of RGB are dependent on one another, but HSV is independent of one another. Another scheme is for histogram based segmentation, which is usually applicable for gray scale images.

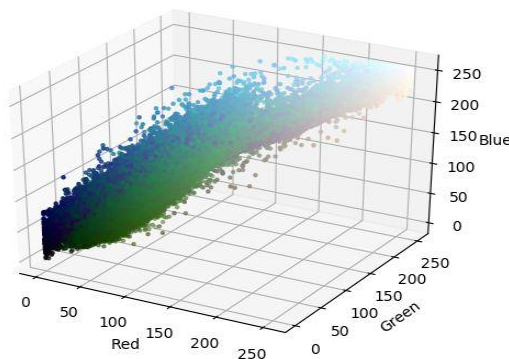


Figure 6: RGB channels of the magnesium deficient tomato leaf

Detection of the disease affected parts of the leaves using various segmentation techniques like hue based segmentation and thresholding based (OTSU)[9] are applied and compared the results.

3.3.1 Threshold based segmentation

The RGB image is converted to HSV colour space. Then the infected parts of the leaf are extracted by masking the green pixel using thresholding[15]. The parts of the leaf are divided like pixels with property value less than threshold and greater or equal to threshold. Thresholding based segmentation results of magnesium deficient leaf is shown in figure 7.



Figure 7: Thresholding based segmentation results

3.3.2 Hue Based segmentation

3.3.2.1 HSV conversion

The first step for image segmentation is convert the image form RGB to HSV color format[12].Hue is representing the original color of the object, whereas saturation represent the gray level content of image and value represents the brightness of the image.Figure8 will show the conversion of RGB to HSV conversion of defected leaves.. Besides segmentation in the hue space is computationally less expensive than in the RGB space.

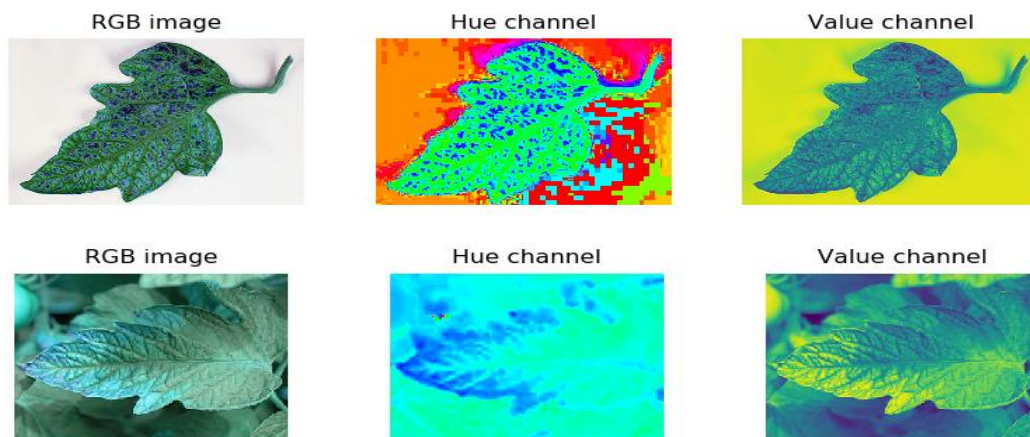


Figure 8: RGB to HSV conversion of magnesium deficient and Pottasium deficient leaves.

3.3.2.2 Thresholding

Applying thresholding[14] will help to remove background from foreground. Thresholding can be applied on hue channel to separate the leaf from background. We finally perform an additional thresholding on the Value channel to partly remove the shadow of the leaf. The result of operation is shown in figure 9.

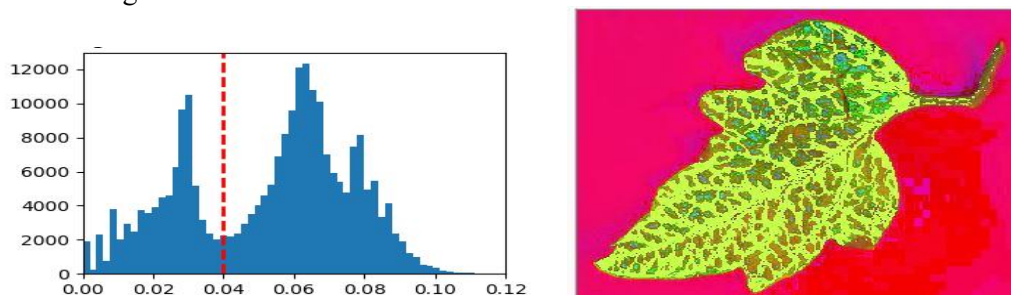


Figure 9: Histogram of Hue channel with threshold and the resultant image after thresholding.

3.3.2.3. Blob detection

Blob detection is a technique which is used to detect regions in the digital image that are differ in properties including brightness and color. All points inside the blob has similar properties. Convolution is the common scheme used for blob detection[12]. Laplacian of Gaussian (LoG) is the most accurate method for the detection of blob. It calculates the Laplacian of Gaussian images successively by increasing standard deviation and stacks them up in a cube. The local maximas in this cube is called Blobs. Gaussian filter is used for smoothening the image. After that we are finding the zero crossings using Laplacian. Laplacian of Gaussian (LoG) operation include both the above steps..

The two steps in the laplacian operation is obtained by taking the Laplacian of the Gaussian kernel after that we should convolve it with the image.

$$\nabla^2(f*g)=f*\nabla^2g$$

Where g is the Gaussian kernel and f is the image. LoG can be expressed as

$$\text{LoG}(x,y)=-\frac{1}{\pi\sigma^4}\left[1-\frac{x^2+y^2}{2\sigma^2}\right]e^{-\frac{x^2+y^2}{2\sigma^2}}$$

The LoG kernel weights can be calculated from the above equation for a given standard deviation σ . The figure 10 shows the blobs detected in the above method.



Figure10: Blob detection using LOG

3.4. Feature Extraction

The input given to the algorithm is huge and can lead to complex processing. The inputs given are compact bonded together so that it represents as set of features. If the features of the image are extracted wisely, it gauges proper information from the input in order to perform relevant task.

Feature extraction is the major part of texture analysis. GLCM (Grey-level Co-occurrence Matrices)[11] based color co-occurrence method is used for extracting statistical features of the leaves. The selected features are Contrast, Variance, IDM (Image difference-measure), Correlation, Energy, Smoothness, Homogeneity, Mean, Entropy, Standard Deviation, RMS (root mean square) contrast, Kurtosis, skewness[16] and shade[9]. All these are calculated from GLCM using its corresponding formulas.

3.5. Classification

The extracted features are stored in a.csv file. Based on the extracted features an artificial neural network model is created. This multi label classification model is used for classifying the images into different deficiency classes. For multi class artificial neural network has two stages. One is training stage and the other one is testing stage. 5 fold cross validation is applied on dataset. To evaluate classifier output quality using cross-validation, Receiver Operating Characteristic (ROC) metric is used.

The activation function will decide whether signal should pass or not. It will map the output value to a range (e.g., between 0 and 1).The usual choice of activation function for multiclass classification is softmax function, which is a generalized form of logistic function .

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

for $j = 1 \dots K$. We can see that the softmax function normalizes a K dimensional vector z of arbitrary real values into a K dimensional vector $\sigma(z)$ whose components sum to 1 (in other words, a probability vector), and it also provides a weighted average of each z_j relative to the aggregate of z_j 's in a way that exaggerates differences (returns a value close to 0 or 1) .

In this Paper, we have compared the performance of the developed system using the following non linear activation functions like Rectified Linear Unit (ReLU), Swish and Softmax. Their influence in the accuracy is also studied separately [18].The graphical representation of activation functions are plotted in figure11.

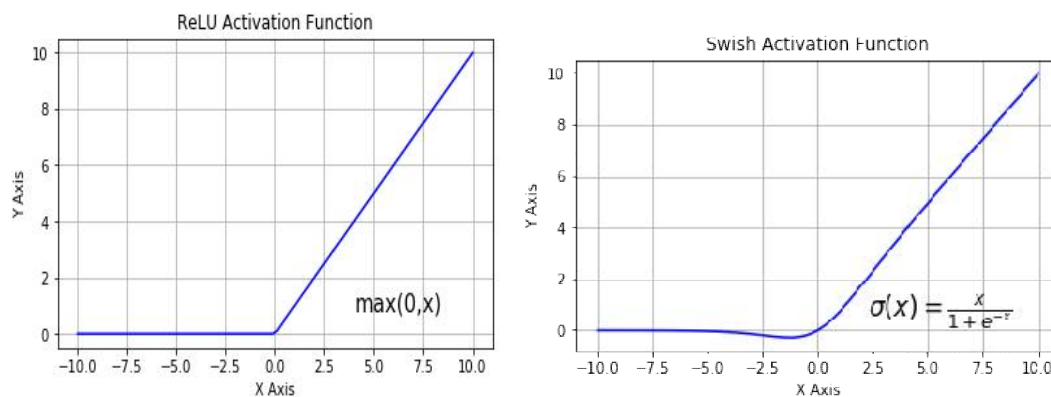


Figure11: Activation functions graphical representation

Based on the output of artificial neural network classification value, the deficiency is detected.



Figure12: Nutrition deficiency detection results

The total percentage of accuracy can be calculated by

$$\text{Accuracy} = (\text{Number of correct outcome} * 100) / \text{Number of test images}.$$

4.Experimental Results and Observation

The experiment was conducted in a dataset of 4049 images, which was extracted from IPNI dataset and other online available datasets [17].The distribution of images with different deficiency classes are given below in figure13.

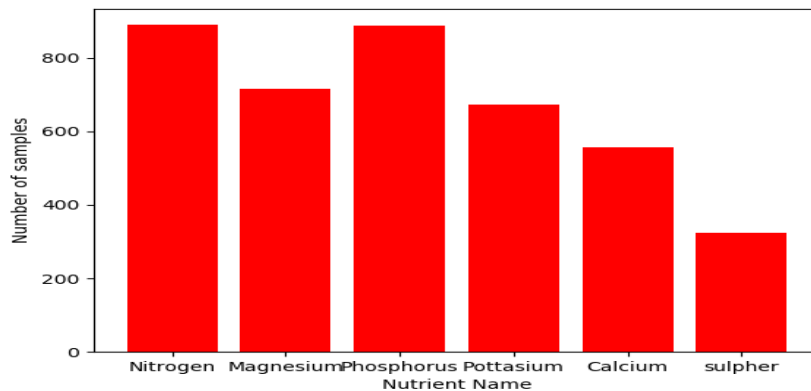


Figure13: Data distribution chart for the experiment

In this study, we are implemented 2 segmentation schemes. The first one is threshold based green color masking scheme and the second one is hue based segmentation. We have extracted feature from both the schemes and checked the accuracy. Hue based scheme work better than the threshold based scheme. Thresholding scheme gives 77% accuracy and Hue based scheme gives 88%.

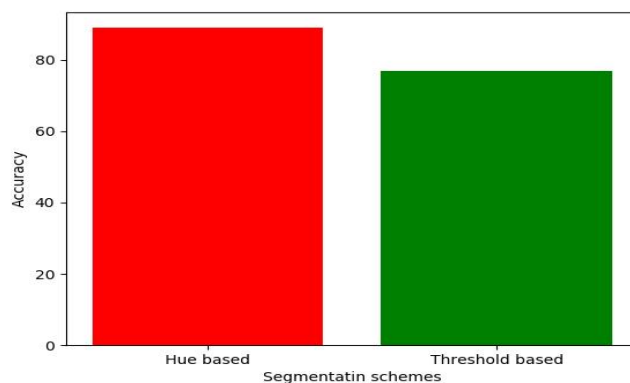


Figure14: Comparison of segmentation schemes with respect to accuracy

A few extracted features of different leaves with deficiency is listed in the table2 below. The features are extracted using GLCM color occurrence statistical schemes.

Table 2: Feature extracted from different leaves

Nutrient Deficiency name	Mean	Standard deviation	Variance	Skewness	Kurtosis
Calcium	3.133	3.817	1.007	1.917	1.777
Magnesium	0.480	0.062	0.002	0.247	5.491
Nitrogen	2.631	4.607	1.280	2.573	4.171
Phosphorus	3.312	3.373	9.122	2.751	6.389
Potassium	2.094	9.977	4.944	6.126	7.927
Sulphur	2.845	4.272	1.329	4.399	4.188

Accuracy of developed system with respect to different nutritional deficiency in the proposed system is shown below. The proposed system shows 93% accuracy for detecting magnesium deficiency as highest and 68% accuracy for detecting Sulphur lowest. Figure 15 shows the accuracy comparison.

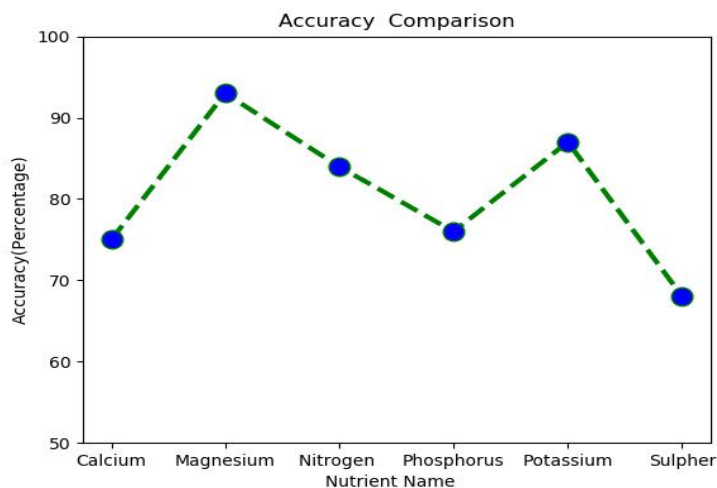


Figure15: Accuracy comparison with respect to different nutrients.

Receiver Operating Characteristic (ROC) metric for multi label classification is carried out through macro averaging and the resultant graph is plotted below in Figure16.

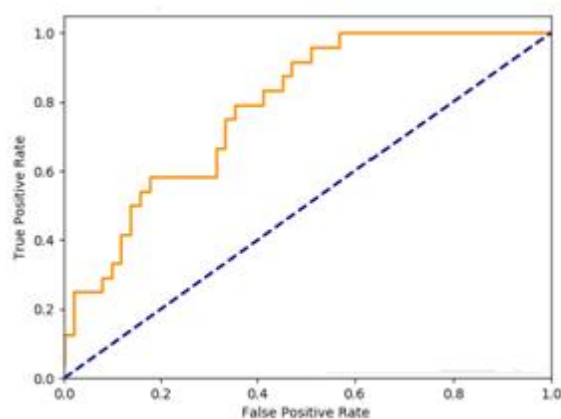


Figure16: ROC Curve for the classifier

Different activation functions like RELU, Softmax and SWISH are applied in neural network. The accuracy comparison is given in the below diagram. SWISH will work better than RELU and softmax for the multi label classification. From figure 17, we can conclude that the proposed system works good in SWISH function.

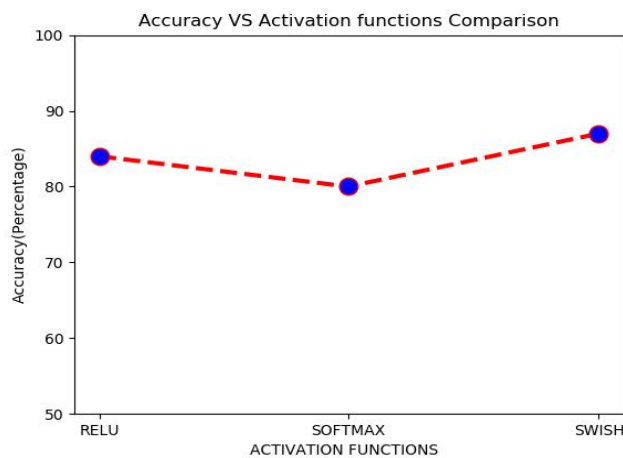


Figure17: Accuracy comparison with respect to different activation functions.

5. Conclusion and Future Scope of Work

An image processing model consisting of features from hue is the best model for the task of tomato leaf classification. Elimination of intensity in texture feature calculation is the major advantage. It nullifies the effect of lighting variations in an outdoor environment. The comparison of different segmentation schemes like hue based and threshold based schemes shows their influence in the performance of the proposed system. Artificial neural network has given an accuracy of 88.27% for detecting and classifying nutrient deficiencies. ANN can be designed with different activation functions and their performance is analyzed and evaluated. In this study, the visible nutritional deficiency symptoms in the leaves are considered, even images of fruits can also be included in dataset which may increase the precision and accuracy of the classification. Design of convolution neural network for the nutrition deficiency detection may give better results.

References

- [1] "Horticulture reports "available at <http://agricoop.gov.in/horticulture-reports>, accessed on 12/12/2019.
- [2] "Crop Guide: Tomato Plant Nutrition" available at <https://www.haifa-group.com/tomato-fertilizer/crop-guide-tomato-plant-nutrition> accessed on 12/12/2019
- [3] Balasubramaniam, P., Ananthi, V.P., 2016. Segmentation of nutrient deficiency in incomplete crop images using intuitionistic fuzzy C-means clustering algorithm. *Nonlinear Dynamics* 83 (1–2), 849–866.
- [4] Chen, L.S., Zhang, S.J., Wang, K., Shen, Z.Q., Deng, J.S., 2013. Identifying of rice phosphorus stress based on machine vision technology. *Life Science Journal* 10 (2), 2655–2663.
- [5] Luz, P.H.C., Marin, M.A., Devecchio, F.F.S., Romualdo, L.M., Zuñiga, A.M.G., Oliveira, M.W.S., Herling, V.R., Bruno, O.M., 2018. Boron Deficiency Precisely Identified on Growth Stage V4 of Maize Crop Using Texture Image Analysis. *Communications in Soil Science and Plant Analysis* 49 (2), 159–169
- [6] Lee, K.-J., Lee, B.-W., 2013. Estimation of rice growth and nitrogen nutrition status using color digital camera image analysis. *European Journal of Agronomy* 48, 57–65.
- [7] Xu, G., Zhang, F., Shah, S.G., Ye, Y., Mao, H., 2011. Use of leaf color images to identify nitrogen and potassium deficient tomatoes. *Pattern Recognition Letters* 32 (11), 1584–1590

- [8] Kai, S.2011. A research of maize disease image recognition of corn based on BP networks, In: Measuring Technology and Mechatronics Automation (ICMTMA), 2011 Third International Conference on, 2011, p. 246-249.
- [9] Marcelo Vassallo-Barco, Luis Vives-Garnique, Victor Tuesta-Monteza, Heber I. Mejía-Cabrera,2017, Automatic Detection of Nutritional Deficiencies In Coffee Tree Leaves Through Shape And Texture Descriptors, Journal of Digital Information Management % Volume 15 , 2017,265-272
- [10] Goodridge, W., Bernard, M., Jordan, R., Rampersad, R. (2017). Intelligent diagnosis of diseases in plants using a hybrid Multi-Criteria decision making technique, Computers and Electronics in Agriculture, 133, p. 80-87.
- [11] Abdolvahab Ehsanirad, Sharath Kumar Y.H., 2010, Leaf recognition for plant classification using GLCM and PCA methods, 3, pp. 31-36, 2010.
- [12] Piyush Chaudhary, Anand K. Chaudhari, Dr. A. N. Cheeran and Sharda Godara,—”Color Transform Based Approach for Disease Spot Detection on Plant Leaf”, IJCST, 2012, 3(6), 65-70.
- [13] S. Sridevy and Anna Saro Vijendran, 2016, “An Evolving Expert System for Maize Plant Nutrient Deficiency using Image Processing Techniques”, International Journal of Science and Research, Vol. 5, No. 2, pp. 70-77.
- [14] Jianlun Wang, Jianlei He, Yu Han, Chanqui Ouyang and Daoliang Li, 2013, “An Adaptive Thresholding Algorithm of Field Leaf Image”, Computers and Electronics in Agriculture, Vol. 96, pp. 23-39.
- [15] Parviz Ahmadi Moghaddam, Mohammadali Haddad Derafshi and Vine Shirzad, 2011 “Estimation of Single Leaf Chlorophyll Content in Sugar Beet using Machine Vision”, Academic Journals, Vol. 35, No. 6, pp. 563-568.
- [16] H. Muhammed Asraf, Nooritawati Md Tahir, S.B. Shah Rizam, R. Abdullah, 2012. “Elaeis Nutritional Lacking Identification based Statistical Analysis And Artificial Neural Network”, Proceedings of Recent Advances in Systems Science and Mathematical Modeling, pp. 144-149.
- [17] “Nutrient Deficiencies” available at <https://growabundant.com/nutrient-deficiencies/> accessed on 16/12/2019.
- [18] E Trentin, 2001, “Networks with trainable amplitude of activation functions”, Neural Networks, Elsevier, Vol 14 (4).