

Classification of Macronutrient Deficiencies in Maize Plant Using Machine Learning

Leena N¹, K. K. Saju²

¹Department of Electrical and Electronics Engineering, NSS college of Engineering, India

²Department of Mechanical Engineering, Cochin University of Science and Technology, India

Article Info

Article history:

Received Apr 23, 2018

Revised Jul 15, 2018

Accepted Aug 8, 2018

Keyword:

Deep learning with auto encoders

Feature extraction

K nearest neighbors

Nutritional deficiency

Support vector machines

ABSTRACT

Detection of nutritional deficiencies in plants is vital for improving crop productivity. Timely identification of nutrient deficiency through visual symptoms in the plants can help farmers take quick corrective action by appropriate nutrient management strategies. The application of computer vision and machine learning techniques offers new prospects in non-destructive field-based analysis for nutrient deficiency. Color and shape are important parameters in feature extraction. In this work, two different techniques are used for image segmentation and feature extraction to generate two different feature sets from the same image sets. These are then used for classification using different machine learning techniques. The experimental results are analyzed and compared in terms of classification accuracy to find the best algorithm for the two feature sets.

Copyright © 2018 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Leena N,

Department of Electrical and Electronics Engineering,

NSS College of Engineering,

Palakkad, Kerala, India.

Email: leena.amulya@gmail.com

1. INTRODUCTION

The increasing cost of crop production and problems of environmental pollution caused by agrochemicals calls for the need to apply mineral fertilizers more efficiently. Plants require nutrients for their healthy growth and development. These include macronutrients like nitrogen, phosphorus and potassium as well as micronutrients like calcium, boron etc. Nutrient deficiency symptoms if left unidentified results in low yield levels and profit. Nutritional deficiencies have characteristic symptoms that are exhibited in the form of color and shape, especially on leaves. These include necrosis, Chlorosis, die-back and others. Knowledge of these symptoms can help us take corrective action to restore the plant to normal state.

The conventional techniques for nutrient deficiency diagnosis include extensive laboratory testing of soil or plant tissue or manual inspection by farmers. These are tedious and time consuming. The application of machine learning techniques using image processing offers a promising solution for identification of nutrient deficiencies. Several works have been reported on the uses of machine learning algorithms combined with image processing techniques to classify and predict abiotic and biotic stresses on plants. Table 1 provides an overview of machine learning techniques combined with image processing methods to detect nutrient deficiency symptoms in plants.

There are also several researches on the use of digital image processing combined with machine learning for detection of diseases in plants. However few literature aims to classify and identify nutritional deficiencies in plants. There have also not been any attempts to make a comparative study of feature extraction and machine learning techniques. This paper aims to make a comparative study of the performance

of feature extraction techniques and machine learning algorithms for classification of nutrient deficiency in maize plants.

Table 1. Overview of Image Processing and Machine Learning Techniques for Deficiency Detection

Application	Feature Extraction Technique	Machine learning Algorithm	Plant Species	Reference
Identification	Spectral and shape features	Support Vector Machine	Rice	[1]
Detection	Color	Clustering	Bean, pea and yellow lupine	[2]
Detection	Geometric moments		Groundnut	[3]
Detection	Texture and color features		Rice	[4]
Detection	Color and size	Artificial Neural Networks	Lettuce	[5]
Quantification	Color		Barley	[6]
Detection	Shape, Texture	Fuzzy	Oil palm	[7]
Classification	Texture features	Fuzzy k-nearest Classifier	Tomato	[8]
Classification	SIFT, Color and shape features	Random Forest	Coffee	[9]

2. RESEARCH METHOD

This work focuses on the recognition and classification of nutritional deficiencies in maize plant. The major steps of classification include selection of training samples, image preprocessing, feature extraction, selection of suitable classification approach, and accuracy assessment.

2.1. Image Data Set

Deficiencies caused by macronutrients like Nitrogen, Phosphorus and Potassium as well as normal (not affected) plants are investigated. 100 sample leaf images for the 4 classes were used in the study. 75 images for used for classification and 25 images were used for detection. Images given in Figure 1 are the samples of maize leaves suffering from nutritional deficiencies.

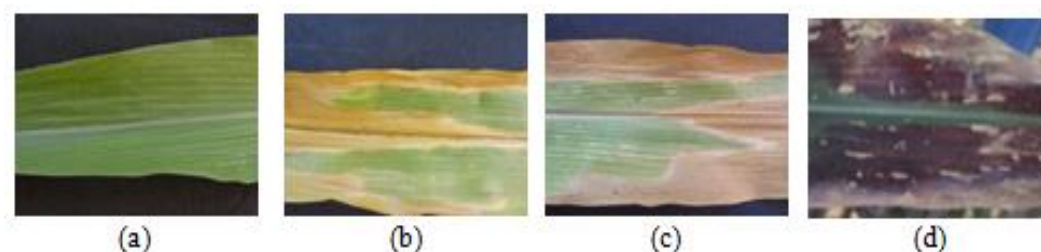


Figure 1. Images of maize leaf a) Healthy leaf b) Nitrogen deficient leaf c) Phosphorus deficiency d) Potassium deficiency

2.2. Image Preprocessing and Feature Extraction

Selecting suitable features is a critical step for successfully implementing image classification. The symptom of nutrient deficiency in plant leaves exhibit different properties in terms of colour, shape, and texture. The first step involves pre-processing the image to improve the quality of the image and to remove distortion. This is followed by segmentation techniques to segment the region of interest from the whole image. Further feature extraction is done from the segmented region having the deficient part of the leaf. This work compares the accuracy of classification using uses two main feature extraction techniques. The first technique uses colour histogram for image segmentation and feature extraction whereas the alternative method uses K-means clustering for segmentation and textural features are extracted.

MATLAB 2016 was used for designing the algorithm for image processing and classification. For histogram based image segmentation and classification, Gabor filter was used for resizing and noise filtering of images. The images were then transformed from the RGB to HSV (Hue, Saturation and Value) colour space as HSV separates the colour components (HS) from the luminance component (V) and is less sensitive to illumination changes. The deficient part of the leaves samples was segmented out from the colour image. The colour features are extracted from HSV color space of selected region based on histogram analysis. Each image added to the collection is analysed to compute a colour histogram, which shows the proportion of

pixels of each colour within the image. Thus as a result 256 features were obtained for each image. The colour histogram for each image is then stored in the database.

In the second technique K-means clustering using Euclidean distance as the minimization criteria was used for segmentation. The image is segmented into three sub-feature images with three different type of Region of Interest (ROI). The Region of Interest corresponding to nutrient deficiency is manually selected from the segmented images. The Grey Level Occurrence Matrix is then computed from the ROI after conversion from RGB to gray scale. 13 features are calculated for each image based on texture properties like Skewness, Standard Deviation, Homogeneity, Contrast, Smoothness, Correlation, Kurtosis, Energy, Entropy, Mean, Variance, RMS, and IDM. Figure 3 shows the results of K-means clustering for a leaf with nitrogen deficiency.

2.3. Machine Learning Techniques

Machine learning comprises of a set of computational modelling techniques that can learn patterns from data and perform automated tasks like identification, prediction or classification. Machine learning techniques are widely used in application like handwriting recognition, natural language processing, speech processing, consumer data predictive analysis, drug design, disease tissue classification in medicine [10],[11] network flow classification [12]. Several machine learning techniques have been used in nutrient deficiency detection in plants as shown in Table 1.

2.3.1. ANN Classifier

Artificial neural networks, usually called neural networks have emerged as an important tool for classification. Neural networks are simplified models of the biological nervous system which consists of highly interconnected network of a large number of processing elements called neurons in an architecture inspired by the brain. They have the potential to classify different forms (patterns) of arbitrary complex input/output mappings.

2.3.2. SVM Classifier

Support Vector Machine (SVM) is considered as one of the efficient machine learning method that is developed on the basis of the statistical learning theory. They are specifically implemented for the classification and regression with high dimensional space. The aim of the SVM classifier is to find an optimal hyperplane. Support vector machines are a very popular method in classification of images due to their good generalization capability even with a limited number of training datasets.

2.3.3. kNN Classifier

Another simple and very power full supervised machine learning technique widely used in statistical estimation and pattern recognition is the k Nearest Neighbour (kNN). The kNN classifier uses a non parametric and instance-based learning algorithm. In classification problems, the k-nearest neighbour algorithm finds majority vote between the k most similar instances to a given “unseen” observation. Similarity is defined according to a distance metric between two data points.

2.3.4. Deep Networks Using Autoencoders

Recently, the area of deep learning is attracting widespread interest by producing remarkable research in almost every aspect of artificial intelligence [13]. Auto encoders are unsupervised machine learning techniques that apply back propagation algorithm. It tries to learn the identify function by placing constraints on the network. Auto encoders can be stacked one over the other to form deep networks called stacked auto encoders. Here learning is done by training one layer at a time. A stacked auto encoder can be used as a classifier by replacing the decoder layer with a softmax layer to classify the features extracted from the encoder layers.

3. RESULTS AND DISCUSSION

Feature extraction is done using the techniques in section 2.2. Figure 2 shows the different steps involved in the segmentation process for a leaf with nitrogen deficiency using the first technique. The segmented part is shown in Figure 2(f). The colour histogram is plotted for the segmented image and a total of 256 features are obtained. The first feature set has a dimension of 75x256 for the 75 images.

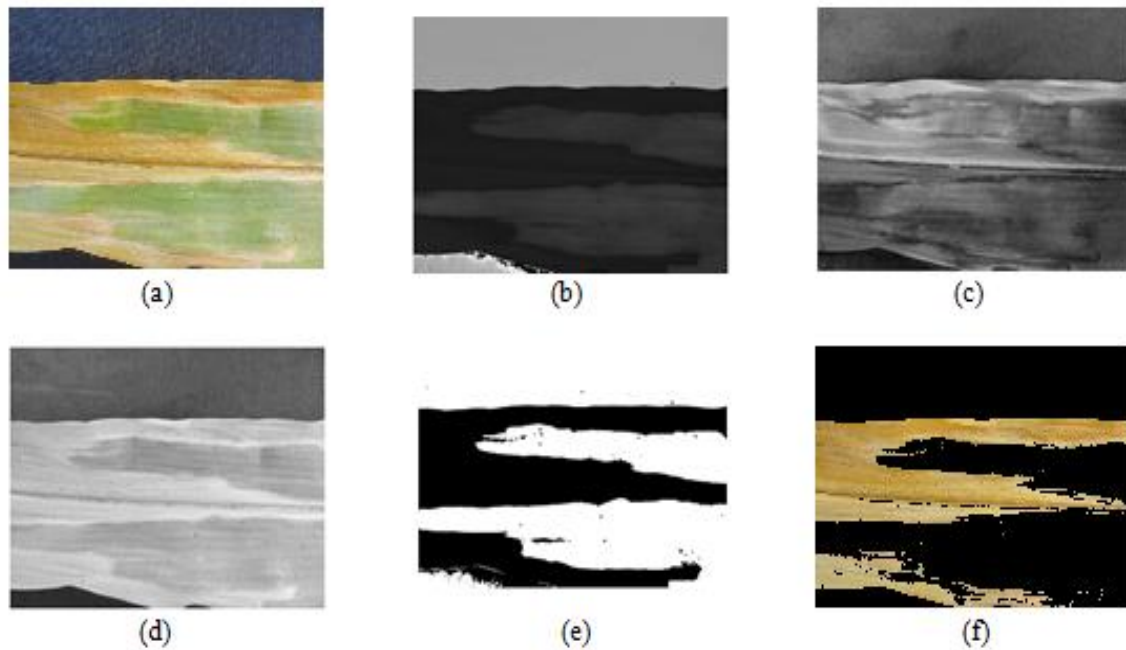


Figure 2. Processing of a leaf image with nitrogen deficiencies a) original Image b) Hue image c) Value image d) Saturation Image e) Hue mask f) segmented regions

Figure 3 shows the results of clustering to obtain the deficient parts of the image using the second method. The clustered image with deficient part is shown in Figure 3(b). This is converted to grey scale and 13 features are obtained using Grey level Occurrence matrix. The second feature set has a dimension of 75×13 . Classification was done using the classifiers in section 2.3 and the performances of the classifiers in terms of accuracy of classification were compared. The training and testing are carried out with both the feature extraction methods given in section 2.2.

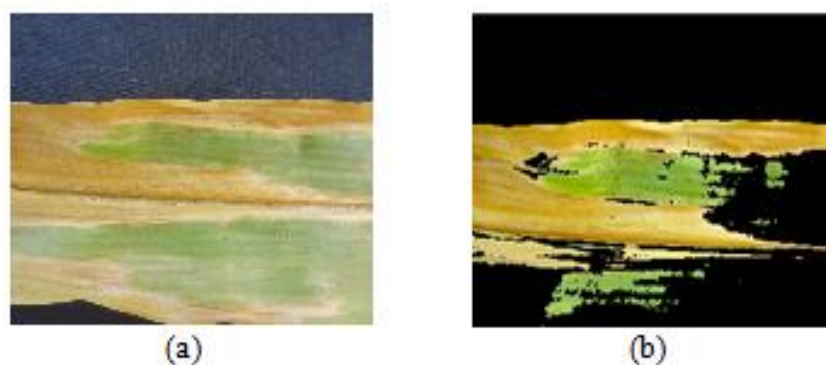


Figure 3. Processing of a leaf image with nitrogen deficiencies a) original Image b) Clustered image with deficient portion after contrast enhancement

A multilayer backpropagation neural network was chosen for the ANN classifier. The number of nodes in the input equals the number of features while the number of nodes in the output is 4 corresponding to the three deficiency classes and one normal class. Number of nodes in the hidden layer is taken as 10 and sigmoid activation function is used. Scaled conjugate gradient backpropagation technique is used for training. An overall classification accuracy of 90.9% was obtained with histogram based feature extraction technique (feature set1) whereas an overall accuracy of 81.2% was obtained for feature extraction using k-means segmentation followed by shape and texture features (feature set 2).

Figure 4 shows the graph of classification accuracy of macronutrient deficiencies affecting maize plant using the two feature sets and an SVM classifier. The effect of kernels on classification accuracy was also investigated. From the graph, it is observed that the maximum classification accuracy of 95.6% has occurred with a linear kernel for feature set 1 and 90% for feature set 2. Hence SVM with linear kernel function gives better classification accuracy compared to other kernels.

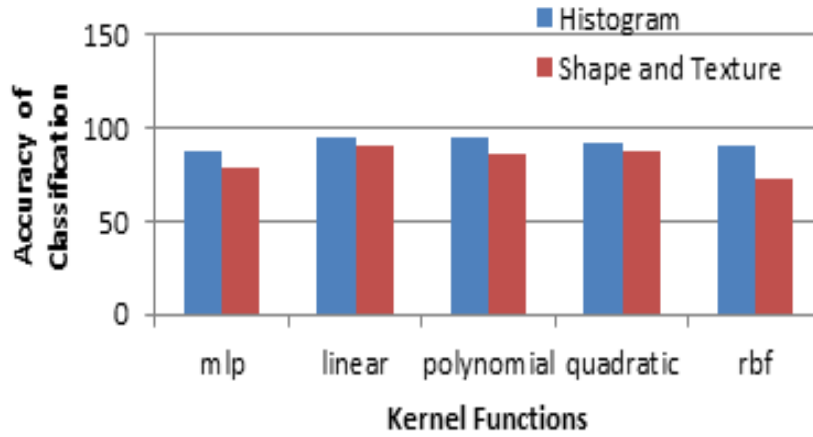


Figure 4. Classification efficiency for SVM classifier with different kernel functions

The results obtained using a KNN classifier is shown in Figure 5 (a) and (b) for the two different feature extraction techniques. The classification result in terms of efficiency was obtained with number of neighbours varied from 2 to 5. As kNN classification is based on measuring the distance between the test data and each of the training data, the chosen distance function can affect the classification accuracy. Hence effect of different distance metrics like Euclidean, Cosine, Minkowksi and Chebychev were also calculated.

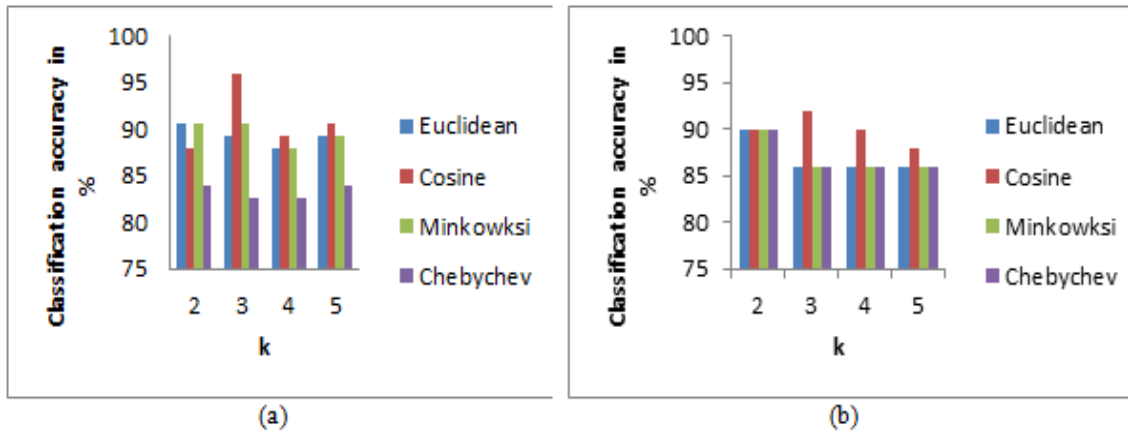


Figure 5. Classification efficiency for kNN classifier with different kernel functions (a) Feature set 1 (b) Feature set 2

The results show that best classification accuracy was obtained with number of neighbours equal to 3 with cosine distance metric for both the feature sets. The next classifier chosen was a deep network with two encoder networks for feature extraction followed by a softmax layer for classification. The two encoders have a hidden layer of size 10 and a linear transfer function. The L2 weight regularizer, sparsity regularizer and sparsity proportion were set to 0.001, 4 and 0.1 respectively. An overall classification accuracy of 100 % was obtained with histogram based feature extraction technique (feature set1) whereas an overall accuracy of 88% was obtained for feature extraction using k-means segmentation followed by shape and texture features

(feature set 2). A comparison of the classification accuracy of different machine learning classifiers for classification of macronutrient deficiencies in maize plant using the two feature sets are shown in Figure 6.

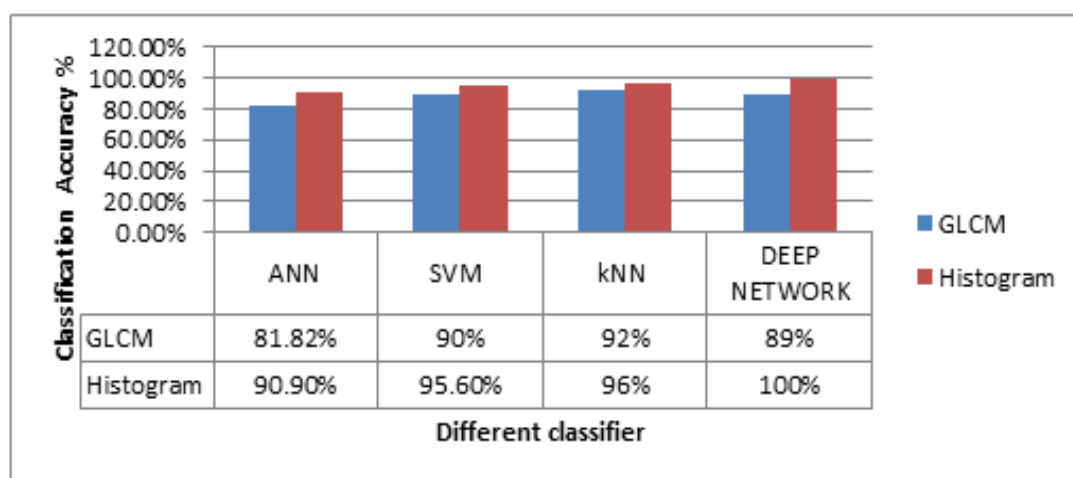


Figure 6. Comparison of efficiency of classification of different classifiers

The results indicate that Deep network using sparse auto encoders gives highest accuracy of classification for histogram based features. KNN classifier worked well with both feature sets for a cosine distance metric and k value chosen as 3. Similarly SVM with linear kernel gives good classification accuracy both the feature sets. The lowest accuracy was obtained with ANN classifier.

4. CONCLUSION

Classification of macronutrient deficiencies in maize plant was done using different feature extraction methods and different classifiers. The algorithms were tested on the three plant macronutrient deficiencies in maize plants, namely nitrogen, potassium and phosphorus. Two feature extraction techniques were used to develop two different feature sets for the same leaf images. The results reveal that deep network with auto encoders gives highest accuracy and has superior performance for histogram based feature extraction compared to the other classification methods. The best classification accuracy with shape and texture features was obtained with KNN classifier. The effect of different kernels for SVM was also investigated for both the feature sets. It was seen that SVM with linear kernel gave the highest accuracy of classification in both cases. Similarly the effect of distance metric and number of neighbours on accuracy of kNN was also checked for both the feature sets. It was seen that kNN with cosine metric and k value of 3 gives highest accuracy. The study can be extended to micronutrient deficiency and disease infestations also.

REFERENCES

- [1] Chen L. *et al.*, "Identification of nitrogen, phosphorus, and potassium deficiencies in rice based on static scanning technology and hierarchical identification method," *PLoS ONE*, vol. 9, 2014.
- [2] Wiwart M., *et al.*, "Early diagnostics of macronutrient deficiencies in three legume species by color image analysis," *Computers and Electronics in Agriculture*, vol. 65, 2009.
- [3] Chandan J. B., *et al.*, "Detection and Analysis of Deficiencies in Groundnut Plant using Geometric Moments," *World Academy of Science. Engineering and Technology International Journal of Biological, Agricultural, Biomolecular, Food and Biotechnological Engineering*, vol/issue: 5(10), 2011.
- [4] Sanyal P., *et al.*, "Color texture analysis of rice leaves diagnosing deficiency in the balance of mineral levels towards improvement of crop productivity," *10th International Conference on Information Technology, IEEE*, pp. 85-90, 2007.
- [5] Hetzroni A., *et al.*, "Machine vision monitoring of plant health," *Adv Space Res*, vol/issue: 14(11), pp. 203–212, 2009.
- [6] Pagola M., *et al.*, "New method to assess barley nitrogen nutrition status based on image colour analysis," *Comput Electron Agric*, vol/issue: 65(2), pp. 213–218, 2009.

- [7] Hairuddin M. A., *et al.*, "Overview of image processing approach for nutrient deficiencies detection in *Elaeis Guineensis*," *IEEE International Conference on System Engineering and Technology*, pp. 116-120, 2011.
- [8] Xu G., *et al.*, "Use of leaf color images to identify nitrogen and potassium deficient tomatoes," *Pattern Recognit Lett*, vol/issue: 32(11), pp. 1584-1590, 2011.
- [9] Monsalve D., *et al.*, "Automatic Classification of Nutritional Deficiencies in Coffee Plants," *6th Latin-American Conference on Networked and Electronic Media*, 2015.
- [10] E. I. Zacharaki, *et al.*, "Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme," *Magn. Reson.*, vol. 62, pp. 1609–1618, 2009.
- [11] M. Hamiane, *et al.*, "SVM Classification of MRI Brain Images for Computer – Assisted Diagnosis," *International Journal of Electrical and Computer Engineering*, vol/issue: 7(5), 2017.
- [12] A. Munther, *et al.*, "A Preliminary Performance Evaluation of K-means, KNN and EM Unsupervised Machine Learning Methods for Network Flow Classification," *International Journal of Electrical and Computer Engineering*, vol/issue: 6(2), pp. 778-784, 2016.
- [13] S. Narejo, *et al.*, "EEG Based Eye State Classification using Deep Belief Network and Stacked AutoEncoder," *International Journal of Electrical and Computer Engineering*, vol/issue: 6(6), pp. 3131-3141, 2016.

BIOGRAPHIES OF AUTHORS



Leena.N is pursuing her doctoral degree under the guidance of Dr.K.K. Saju at Cochin University of Science and Technology, India. She is an Assistant Professor in the Department of Electrical and Electronics Engineering at NSS College of Engineering, Kerala, India. Her research interests include soft computing; Machine learning and Electric drives and control.



K K Saju is Professor at the Department of Mechanical Engineering at Cochin University of Technology, India. He is also the Director of the International Relations and Academic Admissions at Cochin University. His interests lie in material science, robotics and automation. He is author of more than 20 international publications and has handled several funded projects.