IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 9, No. 2, June 2020, pp. 304~309
ISSN: 2252-8938, DOI: 10.11591/ijai.v9.i2.pp304-309

Nutrient deficiency detection in maize (Zea mays L.) leaves using image processing

Nurbaity Sabri, Nurul Shafekah Kassim, Shafaf Ibrahim, Rosniza Roslan, Nur Nabilah Abu Mangshor, Zaidah Ibrahim
Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM) Melaka, Kampus Jasin, 77300 Merlimau, Melaka, Malaysia

Article Info

Article history:

Received Jan 28, 2020
Revised Apr 8, 2020
Accepted Apr 25, 2020

Keywords:

Color histogram
Gray-Level Co-Occurrence
Hu-histogram
Maize leaves
Nutrient deficiency

ABSTRACT

Maize is one of the world's leading food supplies. Therefore, the crop's production must continue to reproduce to fulfill the market demand. Maize is an active feeder, therefore, it need to be adequately supplied with nutrients. The healthy plants will be in deep green color to indicate it consist of adequate nutrient. Current practice to identify the nutrient deficiency on maize leaf is throught a laboratory test. It is time consuming and required agriculture knowledge. Therefore, an image processing approach has been done to improve the laboratory test and eliminate a human error in identification process. The purpose of this research is to help agriculturist, farmers and researchers to identify the type of maize nutrient deficiency to determine an action to be taken. This research using image processing techniques to determine the type of nutrient deficiency that occurs on the plant leaf. A combination of Gray-Level Co-Occurrence Matrix (GLCM), hu-histogram and color histogram has been used as a parameter for further classification process. Random forest technique was used as classifiers manage to achive 78.35% of accuracy. It shows random forest is a suitable classifier for nutrient deficiency detection in maize leaves. More machine learning algorithm will be tested to increase current accuracy.

Corresponding Author:

Nurbaity Sabri,
Faculty of Computer and Mathematical Sciences,
Universiti Teknologi MARA (UiTM) Melaka,
Kampus Jasin, 77300 Merlimau, Melaka, Malaysia.
Email: nurbaity_sabri@uitm.edu.my

1. INTRODUCTION

Zea mays L. the scientific name of maize is a very important staple food supply in many parts of the world from maize itself to cereal. Maize can adapt in different kind of environment and growing in a wider area than other major crop such as potato, wheat and soybean [1]. The maize plant can be defined as a metabolic system whose at the end of the product will produce variety starch deposited from the maize kernels such as cereals or flour. This plant is the most prominent cereal grain all over the world, after rice and wheat. It is commonly providing full-bodied nutrients for humans and animals.

Maize plants are precisely intensive feeders and even soils need to be very fertile to fully supplied nutrients as the plant develop. Compared to other single-cell systems, maize has a long-life cycle since it easy to grow [1]. As to indications of nutrient adequacy, plants ought to be in deep green colour. Some of maize plant facing nutrient deficiency as a result of several agronomic and environment factors. Traditional nutrient deficiency diagnostic methods require comprehensive soil or plant tissue laboratory testing or manual examination by farmers [2]. This is a tedious and laborious work, though, and can be achieved only with very small samples [3]. Therefore, most of the methods of production are approximate.

Nutrients deficiency detection important to ensure all maize grown in accordance with what is requires such as Nitrogen, Iron and more. This nutrient deficiency in maize visually can be seen through the leaves of affected plants using image processing. Moreover, image processing technique can assure that the methodology may be reliable and able to be detection tools in agricultural field [4]. The use of image processing is more accurate since it can capture the differences of pattern, colour, and the surface that affected. Understanding these signs will help determine corrective action to normalize the plant [2]. Classification process in image processing consists of features and classifier. It is essential to concentrate on the feature extraction stage as it has an observably affects on the competence of the recognition system [5]. There are different technique for feature extraction such as texture, color and shape [6].

Gabor filter has been used by researcher for various texture analysis applications [7]. Gabor filter is intended to test the whole recurrence domain of an image by portraying the center frequency and orientation parameters [8]. By deploying of the Gabor filter, image steganalysis improved scheming tool in consequence use of the steganalysis [9]. However, Gabor filters have poor execution when the image is fragmented numerous smaller texture, in this manner influencing the precision of image segmentation [10]. GLCM one of the well know texture features consist of co-occurrence matrices results are better than other texture distinct methods [11]. It is a conventional method of texture feature extraction that can be useful for image classification, segmentation, recognition and more [11]. Color moment (CM) consider as an effective and simple method for color feature [8]. There are overall three step in color moment which are mean, Variance and skewness [12]. According to previous article also mention that CM has a compact feature as it is only required three color components [12]. However, CM is low in discrimination power according to previous article. Color histogram very effective and give actual presentation visualization [12]. In addition, color histogram is useful to recognize image. The implemented color histogram approach has proven to be very easy and efficient to enforce [13]. Histogram Orientation Gradient (HOG) detects edge or gradient form that is very descriptive of local shape and does so in a local representation with only an easily managed degree of in-variance of local geometry [14]. However, it takes a long computation time [15]. A fast and very simple shape features which is hu moment has been introduce [16]. It is the best method in imge processing strategies [17]. Therefore, this research implement combination of GLCM, color histogram and hu moment to as a features to be used in classifier.

Support Vector Machine has been used to detect a leaf disease on grape leaf [18]. However, it is difficult to specify the best parameter to use if data is not separate linearly [19]. CNN is one of excellent image processing approach in Artificial Intelligent that implement general and detail task. Many CNN architectures has been used for image classification and recognition such as AlexNet and LeNet [20]. Enhancements in convolutional neural networks (CNNs) recently have made them the best in class among machine learning approaches for addressing computer vision issues, especially in image classification [21]. This architecture has also been implements on Maize leaf to identify the disease on maize plants and a high accuracy achive [22]. However, one main disadvantage of CNN based methods is that they usually need large datasets to train a feasible model [23]. Random Forests (RF) is among the most effective and efficient machine learning approaches in today's algorithms [24]. Because of their high predictive precision, random forests since then have proven to be successful in so many fields [25]. Due to a great success of random forest, implementation of this algorithm will be done to classify nutrient deficiency of maize leaf.

2. RESEARCH METHOD

Image processing consist of step by step process on classify the nutrient into three classes name Nitrogen, Potassium and Magnesium deficiency. Below shows the flow of classcation starting from input image, preprocessing, feature extraction, and classification. Figure 1 illustrates the leaf detection and classification diagram for nutrient classification on maize leaf.

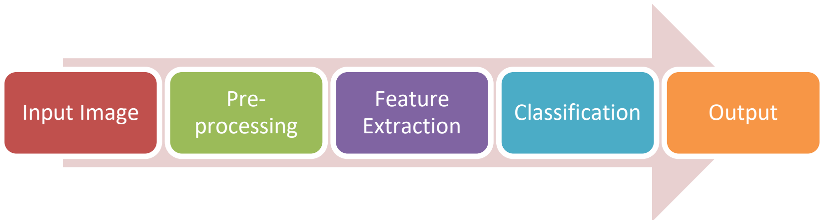

Figure 1. Leaf detection and classification

2.1. Input image

In this study, three type of nutrient deficiency which are Nitrogen, Magnesium and Potassium are being collected. Up to 30 distinct images of classes of nutrient deficiency mentioned will be processes to detect type of nutrient deficiency of the maize leaf. The dataset of maize leaf will be divided into two partition which are training and testing. Figure 2 shows example of image of nutrient deficiency.

(a) (b) (c)

Figure 2. Image of (a) Nitrogen, (b) Potassium and (c) Magnesium deficiency [26]

2.2. Preprocessing

An input image contains of unwanted noise, preprocessing phase is to remove noise and enhancement of image. Image Preprocessing is to expel the undesirable noise in image pursued by section and smoothing of the image and completed to improve the quality of the image [27]. An image of maize leaf captured then resize to 500 X 500 pixels to reduce processing time. The images needed to be enhances before go through to the next process. Images are filter using median filter. Median filter is a non-linear process to remove unwanted noise and outliers from the input images.

$$\hat{f}(x, y) = \underset{(s,t)\in S_{xy}}{\text{median}} \{g(s,t)\} \quad (1)$$

The formula above is the output of median filter where, f (x, y) is the original image and g (s, t) is the output image. S is a two-dimensional mask, where the mask size is m X m (where m is usually odd) such as 3 x 3, 5 x 5, etc.

2.3. Feature extraction

Once image preprocessing is finished, it is important to get the most important qualities of the leaves for separating them with respect of each deficiency [28]. Process of extracting related information from input image is called feature extraction. It is also to transforming input image into a set of features. In the proposed methodology, a comprehensive experimentation is carried out considering texture features from the maize leaf.

2.4. Texture feature

Gray Level Co-occurrence Matrix (GLCM) approach defines the shape of distribution of various tones intensities in the image, which are determined by acquiring the cooccurrence matrices of the image. GLCM is created from a gray-scale image. A co-occurrence matrix portrays the frequency at which a specific gray level is shown in a particular spatial relationship, in relation to another gray level in an image. In this way, the co-occurrence matrix is an outline on how the pixels values are exhibited alongside to another value in a small window. Because of their extensive dimensionality, the GLCM's are sensitive to the size of the texture samples on which they are assessed. Consequently, the quantity of dark dimensions is frequently diminished. The two-dimensional array denoted by P of both rows and columns is signifying to a set of possible gray levels in image values. P (i, j | dx, dy) is the relative frequency with which two pixels, separated by a distance d (dx, dy).

2.4.1. Contrast

It also called as inertia or sum of square variance. The function is to calculate the intensity of the contrast. The contrast statistical feature used to measure the local variations in the GLCM and restoring a calculation contrast of the intensity between the pixels and their neighbors over the whole image. N is an unknown value. Contrast is 0 if the image is constant (dimension of input image size). Equation (2) shows the equation for contrast.

$$\sum_{i,j=0}^{N-1} P_{i,j}\,(i-j)^2 \tag{2}$$

2.4.2. Correlation
Correlation texture is a return measure of linear dependency gray level between the predetermined pixels at the specified positions relative to those neighbors over the entire image. Moreover, it is provided same measurement to autocorrelation method. A correlation is calculated using this formula:

$$\sum_{i,j=0}^{N-1} P_{i,j}\left[\frac{(i-\mu_i)(i-\mu_i}{\sqrt{(\sigma\, i^2)(\sigma\, i^2)}}\right] \tag{3}$$

N is an unknown value, μ is mean based on the reference to the pixels and σ is a standard deviation. Correlation range -1 is for negative correlative image and 1 is for positive correlative image. On the other hand, for a constant image (dimension of input image size), the correlation is NaN.

2.4.3. Homogeneity
The measures of the smoothness homogeneity distribution of the gray level of the image it is approximately inversely correlated with contrast. If contrast is small, for the most part homogeneity is substantial. If weights decline away from the diagonal, the calculated texture measure will be bigger for windows with little difference [29]. Return homogeneity weights values by the inverse of the Contrast weight, with weights diminishing exponentially far from GLCM diagonal. N is an unknown value. Range of homogeneity is 1 if it is a diagonal GLCM. Homogeneity feature formula as:

$$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \tag{4}$$

2.4.4. Entropy
Entropy is a statistical measure of randomness that can be utilized to describe the texture of the maize leaf image and generally classified as a first-degree measure. The images with a larger number of distributed gray levels have bigger entropy. N is an unknown value and ln is this logarithm and uses a base close to 2.718 same with log with a base of 10. Entropy feature as following formula:

$$\sum_{i,j=0}^{N-1} P_{i,j}\left(-\ln P_{i,j}\right) \tag{5}$$

2.5. Color feature
A histogram counts the pixels number in each form and can be easily generated by reading each pixel of each image once and increasing the histogram's correct bin [13].The histogram features represent statistical-based features, whereby the histogram is often used as a representation of the distribution of likelihood of the image intensity levels [30]. The color histogram for an image is created by mapping input values in a smaller set of colors within the image from a large set of output values and counting the number of pixels of each color [31]. Scanning the image, setting color values to the histogram scale, and constructing the histogram using color attributes as indicators are easy processes for creating color histogram characteristics. For this study, to compute the color histogram is using the following parameters such as images, channels, mask, histogram size and ranges.

2.6. Shape feature
Shape is the primary source of information used to recognize objects. No visual content object can be properly recognized without shape. Moment invariants are essentially the region descriptors that are most popular and widely [32].

2.7. Random forest
The random forest classifier is close to the top of the classifier rankings. Random forest could be used for classification as well as for regression. A random forest's cumulative prediction error is tightly correlated with individual trees ' intensity and density in the forest. Adding significant randomness in the base models, trees and creating subsets of the predictor system can refines bagging to separate the a tree nodes of random forest [33]. In the training phase, it automatically calculates the appropriate score for each element. After that it scales down the significance for the total of all scores to be 1. The overall explanation can be defined by using the Gini index.

$$Gini = 1 - \sum_{i=1}^{c}(pi_i)^2 \tag{6}$$

In addition, above equation includes the class and likelihood to decide which branch Gini is most likely to occur on a node. Thus, pi is the absolute frequency of the class that will be finds in the dataset, and c is the number of classes. The function of classification process is to classify image according to the type of nutrient deficiency and its accuracy percentage. In this research

3. RESULTS AND DISCUSSION

The testing result of the nutrient deficiency of maize leaf is divided into three types which are Magnesium, Nitrogen, Potassium and healthy shows in Table 1. The classifier shows the maximum probability for the class it predicts. Confusion matrix is used to calculate accuracy percentage for the overall system. The result of potassium detection effected by its similarity with magnesium features. Thus, potassium has the lowest number of correct detection as it mostly detects it as a magnesium as it has the closer features as magnesium. The accuracy achieve 78.35 percent as consequence of low detection on potassium.

Table 1. Result of the identification

Type	Number of Tested Image	Number of Image Identified Correctly	Number of Image Identified Incorrectly
Healthy	31	31	0
Magnesium	35	35	0
Nitrogen	36	26	10
Potassium	32	11	21

$$Accuracy\ \% = \frac{Total\ TRUE\ value}{Total\ images\ testing}\ x\ 100 \tag{7}$$

4. CONCLUSION

This research classifies four type of class which is healty leaf, nutrient, magnesiuam and potassium. Some of the nutrient deficiency is likely to have the same trait, therefore it is difficult to classify these nutrient. More data will be collected in the future by adding more training data for classification process. Besides, the same dataset will be tested with other available machine learning to increase the current accuracy achieve by random forest classifier.

ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Higher Education, Malaysia and Universiti Teknologi MARA for the research funding and support via grant number 600-IRMI/FRGS 5/3 (215/2019).

REFERENCES

[1] S. Hake & J. Ross-Ibarra, "Genetic Evolutionary and Plant Breeding Insights from The Domestication of Maize," *ELife*, 4, pp. 1–8, 2015.
[2] L. Nair & K. K. Saju,"Classification of Macronutrient Deficiencies in Maize Plant Using Machine Learning," *International Journal of Electrical and Computer Engineering (IJECE), 8(6)*, pp. 4197–4203, 2018.
[3] S. B. Jadhav et al., "Convolutional neural networks for leaf image-based plant,"*IAES International Journal of Artificial Intelligence (IJ-AI)*, pp. 328-341, 2018.
[4] S. Sridarane, & A. S. Vijendran,"Analysis of Maize Crop Leaf using Multivariate Image Analysis for Identifying Soil Deficiency Analysis of Maize Crop Leaf using Multivariate Image Analysis for Identifying Soil Deficiency," *Research Journal of Applied Sciences, Engineering and Technology*, 8(19), 2014.
[5] B. Chitradevi & P. Srimathi,"An Overview on Image Processing Techniques," *International Journal of Innovative Research in Computer and Communication Engineering*, 2(11 Nov 2014), pp. 6466–6472, 2014.
[6] S. Barhmi & Fatni El, O.,"Hourly Wind Speed Forecasting Based on Support Vector Machine," *IAES International Journal of Artificial Intelligence (IJ-AI)*, pp. 286-291, 2019.
[7] A. Chaudhary & S. S. Singh,"Lung Cancer Detection using Digital Image Processing," *Research Journal of Engineering and Technology*, 38(2), 1351–1359, 2012.
[8] D. Tian,"A Review on Image Feature Extraction and Representation Techniques*,"International Journal of Multimedia and Ubiquitous Engineering*, 8(4), 2013.
[9] M. Sharma & B. Singh,"Feature Extraction and Analysis using Gabor Filter and Higher Order Statistics for the JPEG Steganography," *International Journal of Applied Engineering Research*, 13(5), pp. 2945–2954, 2018.

[10] Y. Wicaksono & V. Suhartono, “Color and Texture Feature Extraction Using Gabor Filter - Local Binary Patterns for Image Segmentation with Fuzzy C-Means,” *Journal of Intelligent Systems*, 1(1), pp. 15–21, 2015.
[11] X. Zhang et al., “A Study for Texture Feature Extraction Of High-Resolution Satellite Images Based on A Direction Measure and Gray Level Co-Occurrence Matrix Fusion Algorithm,” *Sensors*, 17(7), 2017.
[12] A. Masood & M. Sharif, “Content-Based Image Retrieval Features : A Survey,” *Int. J. Advanced Networking and Applications*, 10(1), pp. 3741–3757, 2018.
[13] F. Alamdar & M. R. Keyvanpour, “A New Color Feature Extraction Method Based On Quad Histogram,” *Procedia Environmental Sciences, 10(PART A)*, pp. 777–783, 2011.
[14] T. Surasak, et al.,”Histogram Of Oriented Gradients for Human Detection in Video,” *Proceedings of 2018 5th International Conference on Business and Industrial Research: Smart Technology for Next Generation of Information, Engineering, Business and Social Science, ICBIR 2018*, pp. 172–176, 2018.
[15] M. Jamshedet, al.,”Significant HOG-Histogram of Oriented Gradient Feature Selection for Human Detection,” *International Journal of Computer Applications*, 132(17), pp. 20–24, 2015.
[16] S. Zekovich & M. Tuba, “Hu Moments Based Handwritten Digits Recognition Algorithm,” Recent Advances in Knowledge Engineering and Systems Science, pp. 98–103, 2013.
[17] M. Fernando,”Novel Approach to Use HU Moments with Image Processing Techniques for Real Time Sign Language Communication,” *International Journal of Image Processing (IJIP)*, 9(6), pp. 335–345, 2015.
[18] S. S. Kothawale, et al., “Grape Leaf Disease Detection Using SVM Classifier,” *International Journal of Innovative Research in Computer and Communication Engineering*, 2018.
[19] D. Anguita et al., “Model Selection for Support Vector Machines: Advantages and Disadvantages of the Machine Learning Theory,” *The 2010 international joint conference on neural networks (IJCNN)*, pp. 1-8, 2010.
[20] S. Wallelign,”Soybean Plant Disease Identification Using Convolutional Neural Network,” *Artificial Intelligence Research Society Conference (FLAIRS-31) Soybean*, pp. 146–151, 2017.
[21] R. J. Nelson, et al. “Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning,” *Phytopathology*, 107(11), pp. 1426–1432, 2017.
[22] C. DeChant, et al., "Automated Identification Of Northern Leaf Blight-Infected Maize Plants From Field Imagery Using Deep Learning,” pp. 1426-1432, 2017.
[23] H. Wang, et al.,“A hybrid CNN Feature Model for Pulmonary Nodule Malignancy Risk Differentiation,” *Imaging for Patient-Customized Simulations and Systems for Point-of-Care Ultrasound*, pp. 19-26, 2017.
[24] D. Nayak, et al.,“Brain MR Image Classification Using Two-Dimensional Discrete Wavelet Transform And Adaboost With Random Forests,“ *Neurocomputing*, pp. 188–197, 2016.
[25] D. Denisko & M. M. Hoffman, “Classification and Interaction In Random Forests,” *Proceedings of the National Academy of Sciences of the United States of America*, 115(8), pp. 1690–1692,” 2018.
[26] M. Anitha & K. Kaarthik,“Analysis of Nutrient Requirement of Crops Using Its Leaf*,”Journal of Chemical and Pharmaceutical Sciences*, (8), pp. 99–103, 2016.
[27] N. G. Dhakad et al., “Leaf Disease Detection Using Image Processing for Pesticides Spraying,“ *International Journal of Advance Engineering and Research Development*, 4(04), pp. 687–689, 2018.
[28] V. Marcelo, et al.,“Automatic Detection of Nutritional Deficiencies In Coffee Tree Leaves Through Shape And Texture Descriptors,” *Journal of Digital Information Management*, 15(1), 2017.
[29] G. Kaushal & R. Bala,“GLCM and KNN based Algorithm for Plant Disease Detection, “ *International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering*, 6(7), pp. 5845–5852.
[30] S. Sergyán, “Color Histogram Features Based Image Classification In Content-Based Image Retrieval Systems,” *SAMI 2008 6th International Symposium on Applied Machine Intelligence and Informatics - Proceedings*, (February 2008), pp. 221–224, 2017.
[31] K. Roy & J. Mukherjee,”Image Similarity Measure using Color Histogram, Color Coherence Vector, and Sobel Method,” *International Journal of Science and Research*, 2(1), pp. 2319–7064, 2013.
[32] A. Thomas & K. Sreekumar, “A Survey on Image Feature Descriptors-Color, Shape and Texture,” *International Journal of Computer Science and Information Technologies*, 5(6), pp. 7847–7850, 2014.
[33] D. Mohammed, et al.,”Overlapped Music Segmentation Using A New Effective Feature and Random Forests,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, pp. 181-189, 2019.