Hue Saturation Value (HSV) Color Detection to Identify Nutrient Deficiencies in Chili (Capsiumannuum L) using K-Nearest Neighbor

Arie Qurania [1], Prihastuti Harsani[2], Triastinurmiatiningsih[3], Dona Putra Por[4]*

[1,2,4]Department of Computer Science, Universitas Pakuan, Indonesia.
[3]Department of Biology, Universitas Pakuan, Indonesia.

Abstract: Along with the increasing demand for the use of chili in Indonesia with increasing population growth causes high chili prices on the market. Chili is widely used as a spice and spice for cooking because it has a distinctive taste and aroma that will add flavor to the cuisine [1]. Red chili (Capsicum annuum L.) is a horticultural plant that is widely used for food needs. Chili trees need macro elements such as Nitrogen (N), Oxygen and Hydrogen, while the micro elements needed include Potassium (K), Calcium (Ca), Magnesium (Mg). Nutrient deficiency in chili plants results in disruption of plant development such as yellowing of leaves and easy to fall, fragile chili and affect the quality of chili. Nutrient deficiencies in chili plants if not immediately addressed will result in plant death and decreased quality of chili production. This study will identify nutrient deficiencies in chilli plants using color extraction on leaves with Hue Saturation Value (HSV) color characteristics and identification using k-Nearest Neighbor (KNN). Classification of nutrient deficiencies in this study is the lack of N + P and lack of P + K in chili leaves. Leaf image data is divided into 300 training data consisting of 100 normal leaf images, 100 leaf images lacking N + P elements and 100 P + K nutrient deficient images, while for training data divided into 50 normal leaf image data, 50 lacking leaf images nutrient elements N + P and 50 nutrient deficiency P + K. Accuracy results using k = 3 and 5 resulted in an accuracy value of 82% for the identification of N + P deficient leaf image and 74% for identification of P + K nutrient deficiency.

Keywords:HSV, Chili Identify, nutrient deficiency, k-Nearest Neighbor, color detecton, leaf image data

Introduction

Along with the increasing demand for the use of chili in Indonesia with increasing population growth causes high chili prices on the market. Chili is widely used as a spice and spice for cooking because it has a distinctive taste and aroma that will add flavor to the cuisine [1]. Red chili (Capsicum annuum L.) is a horticultural plant that is widely used for food needs. According to Rukmana and Oesman (2006), their use in industry makes chili a high economic value commodity. Just like other holtikura plants, chili also needs enough nutrients to grow well and produce lots of quality fruit [2].

Chili trees need macro elements such as Nitrogen (N), Oxygen and Hydrogen, while the micro elements needed include Potassium (K), Calcium (Ca), Magnesium (Mg). Nutrient deficiency in chili plants results in disruption of plant development such as yellowing of leaves and easy to fall, fragile chili and affect the quality of chili. Nutrient deficiencies in chili plants if not immediately addressed will result in plant death and decreased quality of chili production. This study will identify nutrient deficiencies in chilli plants using color extraction on leaves with Hue Saturation Value (HSV) color characteristics and identification using k-Nearest Neighbor (KNN).

Research using Red Green Blue (RGB) color features on the leaves has been carried out including on the leaves of corn to identify diseases based on the image of corn leaves using KNN with the highest accuracy rate of 91.7% [3]. Other studies using RGB color elements for the classification of flower parts were also carried out in jasmine flowers with the highest accuracy of 84% [4], while in the identification of chrysanthemum flowers with an accuracy of 75.35% [5]. Research on nutrient deficiency has been carried out by researchers on Okra plants using the deep learning method which results in an accuracy rate of 96%. This study uses data on the treatment of N, P, and K deficiency in Okra plants with hydroponic growing media [6]. The color element is the defining characteristic of a plant that has nutrient deficiency among which is the color of yellowish green leaves, brown leaf tips, and yellowish leaves [7]. RGB color elements have been studied in cucumber plants for nutrient deficiency with the combined treatment of two nutrient deficiencies resulting in the highest accuracy of 70.25% [8].

DOI: 10.5373/JARDCS/V12SP6/SP20201012
**Corresponding Author: Arie Qurania, Email: qurania@unpak.ac.id*
Article History: Received: Mar 21, 2020, Accepted: June 22, 2020

Methodology

This study uses plant data that is treated with N + P nutritional deficiencies and P+K nutritional deficiencies and normal data comparison data without nutritional deficiencies. Image data is taken with a camera that has 20MP pixels with 3 shots. The background image uses a black base in sufficient light and is centered on the leaf. Color extraction using HSV has 3 characteristics, namely Hue, Saturation and Value [9]-[10]. In question is:

a. Hue: states the actual color, such as red, violet, and yellow and is used to determine redness (redness), greenness (greeness), etc. using the formula in equation (1)
b. Saturation: sometimes called chroma, is the purity or strength of color using the formula in equation (2)
c. Value: brightness of color. The value ranges from 0-100%. If the value is 0, the color will become black, the greater the value, the brighter and new variations will appear using the formula in equation (3)

RGB color image is converted into HSV by calculating the conversion of RGB to HSV, the first way to get each value from HSV is relatively simple, namely:

$$H = \tan\left(\frac{3(G-B)}{(R-G)+(R-B)}\right) \quad (1)$$

$$S = 1 - \frac{\min(R,G,B)}{V} \quad (2)$$

$$V = \frac{R+G+B}{3} \quad (3)$$

The ease with which the first method turns out to cause problems, the first way to make hue undefined if the saturation is 0. The second solution to get each HSV value is to use the following second formula

$$r = \frac{R}{(R+G+B)};\ g = \frac{G}{(R+G+B)};\ r = \frac{B}{(R+G+B)} \quad (4)$$

$$V = max\,(r, g, b) \quad (5)$$

$$S = \begin{cases} 0,\ if\ V = 0 \\ 1 - \frac{\min(r,g,b)}{V},\ if\ V > 0 \end{cases} \quad (6)$$

$$H = \begin{cases} 0, if\ S = 0 \\ \frac{60*(g-b)}{S*V},\ if\ V = r \\ 60 * \left[2 + \frac{b-r}{S*V}\right], if\ V = g \\ 60 * \left[4 + \frac{r-g}{S*V}\right], if\ V = b \end{cases} \quad (7)$$

$$H = H + 360\ if\ H < 0 \quad (8)$$

The identification process uses KNN (k-Nearest Neighbor), a method that is included in the instance-based learning group. This algorithm is also one of the lazy learning techniques. KNN is done by finding the k group of objects in the training data that is closest (similar) to the object in new data or testing data [11]. Calculation of distances using euclidean distance and manhattan distance (city block distance), the most commonly used is euclidean distance [12].

$$\sqrt{(a1-b1)^2 + (a2-b2)^2 + \cdots + (an-bn)^2} \quad (9)$$

Where $a = a_1, a_2, ..., a_n$, and $b = b_1, b_2, ..., b_n$ represent the n attribute values of the two records. For attributes with category values, measurements with euclidean distance do not match. Instead, the following functions are used [13].

Results and Discussion

Preprocessing

Preprocessing data, is cropping and resizing the chili leaf image. Resize and cropping the original image is determined by the pixel value with a resolution of 50x50 pixels Cropping stage with the vertices of x, y, a, b is 15,15,20,20, the value is very much in accordance with the shape of the object itself which is square and only takes part green on the leaf object as shown in Fig. 1.

	Image of chili leaves resized to 50x50 pixels
	Cropping with respect to x, y, a, b angle with value [15 15 20 20]

Figure 1. Resize and Cropping Results

The nearest distance test phase is tested, the value of the distance from k-NN method with resized image 50 x 50 pixel produces different accuracy level can be seen in Table 1.Based on [14], k-Fold Cross Validation repeats k-times to divide a sample set randomly into k subsets which are mutually independent, each The test is left with a subset for testing and another subset for training with K = 5 or 10 can be used to estimate the level of errors that occur, because the training data on each fold is quite different from the original training data [15]. Overall, 5 or 10-fold cross validation are both recommended and agreed upon.

HSV Featured Extraction

HSV feature extraction stage, the image used is the result of the conversion of RGB to HSV color space, after that it is done normalization, normalization aims to change the RGB color unit with a range of values (255) into HSV color units with a range of values (0-1) so that it can be calculated and extract features from each Hue, Saturation, and Value colors.

In extracting and converting color RGB image data to HVS, the default color of the pixels in X and Y angles is the RGB color space (Red, Green, Blue). So the first thing in feature extraction, what needs to be done is to identify the RGB color units, then convert them into HSV color units. Here are some sample examples of RGB values at 50 x50 pixel cropping results on the x, y, a, b angle [15 15 20 20] with the vertices x, y [11,11] and are classified based on the effect of light on image capture as shown in Figure 3 and 4.

DOI: 10.5373/JARDCS/V12SP6/SP20201012
*Corresponding Author: Arie Qurania, Email: qurania@unpak.ac.id
Article History: Received: Mar 21, 2020, Accepted: June 22, 2020

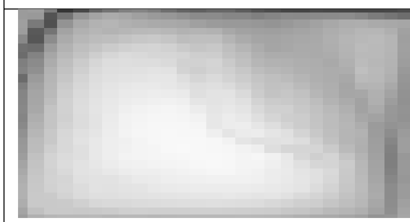
	Leaf image converted from hue color space.
	Converted leaf image with saturation color space.
	Converted leaf image with value color space.
	Converted leaf image with hue, saturation, and value (HSV) color space.

Figure 2. Results Conversion o RGB to HSV colors

Calculate the Average Value of HSV Feature Extraction

HSV value has been extracted, the next process is to take the average value of the entire HSV color pixel using the formula in equation 10

$$\bar{x} = \frac{\Sigma x}{n} \qquad (10)$$

Examples of the results of calculating the average HSV feature extraction are as in Table 1.The nearest distance test phase is tested, the value of the distance from k-NN method with resized image 50 x 50 pixel produces different accuracy level can be seen in Table 2.

Table 1 HSV feature extraction

Data training	Mean Hue	Mean Saturation	Mean Value	Output
Image_1	128.00	182.12	155.65	Defisiensi NP
Image_2	135.79	184.60	164.17	Defisiensi NP
Image_3	130.16	177.62	171.24	Defisiensi NP
Image_4	129.00	183.03	173.33	Defisiensi NP
Image_5	134.79	171.04	154.71	Defisiensi PK
Image_6	139.83	164.12	170.97	Defisiensi PK
Image_7	133.90	162.14	171.04	Defisiensi PK
Image_8	132.79	173.77	176.56	Defisiensi PK
Data testing				
Data testing	Mean Hue	Mean Saturation	Mean Value	Output
data_train_10	123.52	190.22	180.31	Defisiensi NP

Table 2. Nearest distance test using K-NN

The shortest distance (value k=3 and 5)	Level of accuracy
Deficiency N+P	82%
Deficiency P+K	74

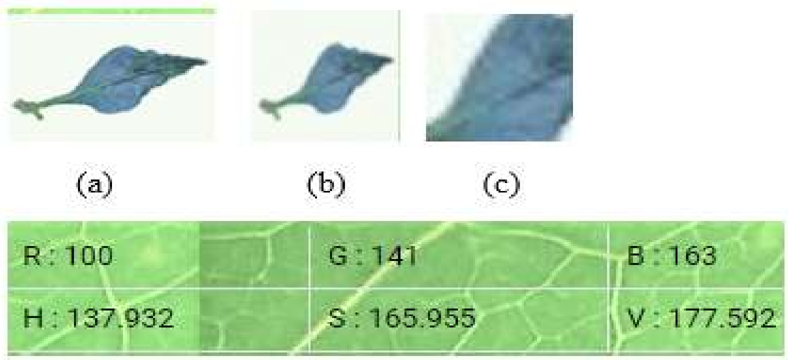


Figure 3. Extraction of RGB and HSV color features on chili leaves lacking NP.

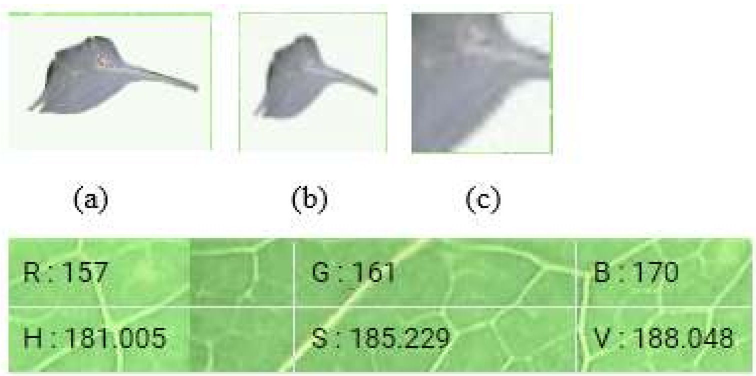


Figure4. Extraction of RGB and HSV color features on chili leaves lacking PK

Conclusion

N + P nutrient deficiency and P + K nutrient deficiency in chilli plants based on color characteristics on the surface of chili leaves that experience nutrient deficiency results in an accuracy of 82% for N + P and 74% for P + K. Color feature extraction used is to take the average values of the Hue, Saturation, and Value variables in the training and testing data. Identification using k-Nearest Neighbord with values k = 3 and k = 5, randomization of training data and testing data using k-fold validation with value k = 5. A slight difference in value between the 3 HSV parameters makes a class error to identify deficient image data more P + K nutrients compared to image data lacking N + P nutrients.

Acknowledgment

The researchers would like to thank the Ministry of Research, Technology, and Higher Education Institutions and LLDIKTI region IV for fully funding this research through the Higher Education Collaboration Research Grant with contract number 2889/L4/PP/2019. Acknowledgements are also conveyed to the Department of Computer Science, Faculty of Mathematics and Natural Sciences, Pakuan University and Department of Informatics Engineering, Bina Nusantara University for their moral support and facilities for infrastructure in research.

DOI: 10.5373/JARDCS/V12SP6/SP20201012
**Corresponding Author: Arie Qurania, Email: qurania@unpak.ac.id*
Article History: Received: Mar 21, 2020, Accepted: June 22, 2020

References

[1] Sherman, P. W., and Billing, J. Darwinian gastronomy: why we use spices: spices taste good because they are good for us. *BioScience*, vol. 49, no. 6, 1999, 453-463.

[2] Rukmana, R., and Oesman, Y. Y.*Bertanam cabai dalam pot*. Yogyakarta: Kanisius Publisher, 2006.

[3] Harsani, Prihastuti, Qurania, A., andDamayanti, K. Maize Plant Desease Identification (Zea Mays L. Saccharata) Using Image Processing and K-Nearest Neighbor (K-Nn). *International Journal of Engineering and Technology*. vol. 7,2018, 402-405.

[4] Qur'ania, A., and Sarinah, I. Identification of jasmine flower (Jasminum sp.) based on the shape of the flower using sobel edge and k-nearest neighbour. In *IOP Conference Series: Materials Science and Engineering*, vol. 332, no. 1, 2018, 012008.

[5] Qur'ania, A., Harsani, P., and Putri, V. K. Identification of Chrysanthemum Flower Based on Color and Flower Form using Sobel Edge and Artificial Neural Network. *International Journal of Engineering and Technology*, vol. 8, no. 2S7, 2018, 71-75.

[6] Wulandhari L, Gunawan A, Qurania A, Harsani P, Triastinurmiatiningsih, Tarawan F, and Hermawan R. Plant nutrient deficiency detection using deep convolutional neural network. *ICIC Express Letters*,vol. 13, no. 10, 2019, 971-977.

[7] Triastinurmiatiningsih, Harsani, P., Qurania, A., and Hermawan, R. F. Effects of Deficiency Nitrogen Phosphorus Potassium Calcium in Okra (Abelmoschus esculentus L. Moench) Through Hydroponics. *International Journal of Recent Technology and Engineering*, vol.8, no. 3, 2019, 4393-4396.

[8] Qur'ania, A., Harsani, P., Triastinurmiatiningsih, T., Wulandhari, L. A., and Gunawan, A. A. S. Color Extraction and Edge Detection of Nutrient Deficiencies in Cucumber Leaves Using Artificial Neural Networks. *CommIT (Communication and Information Technology) Journal*, vol. 14, no. 1, 2020, 23-30.

[9] Trisnawan, Sambas, A., Jannah, M., Rahayu, D. S., Mamat, M., Mohamed, M. A., Sanjaya, W. S . M. Design of 4 Dof Robot ARM Based on Adaptive Neuro-Fuzzy (ANFIS) using Vision in Detecting Color Objects. *International Journal of Recent Technology and Engineering,* vol. 8, no. 2s7, 2019, 224-227.

[10] Gonzalez, R. C., Woods, R. E., and Eddins, S. L. *Digital image processing using MATLAB*. New Delhi: Pearson Education India, 2004.

[11] Wu, X., and Kumar, V. *The Top Ten Algorithms in Data Mining*. New York: CRC Press, 2009.

[12] Bramer, M..*Principles of Data Mining*. London: Springer, 2007.

[13] Larose, D. T. *Discovering Knowledge in Data*. New Jersey: John Willey & Sons, Inc, 2005.

[14] Fu, L.*Neural Network In Computer Intelligence*. Singapore: McGraw Hill, 1994.

[15] Hastie, T, Tibshirani, R., and Friedman, J.*The elements of statistical learning: data mining, inference, and prediction*. Canadian: Springer Science & Business Media, 2009.