Indonesian Journal of Electrical Engineering and Computer Science
Vol. 22, No. 1, April 2021, pp. 179~186
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v22.i1.pp179-186

Classification of rice plant nitrogen nutrient status using k-nearest neighbors (k-NN) with light intensity data

Muliady Muliady[1], Lim Tien Sze[2], Koo Voon Chet[3], Suhadra Patra[4]
[1,4]Department of Electrical Engineering, Universitas Kristen Maranatha, Indonesia
[1,2,3]Faculty of Engineering and Technology, Multimedia University, Malaysia

Article Info

Article history:

Received Mar 19, 2020
Revised Nov 10, 2020
Accepted Dec 5, 2020

Keywords:

Consistent lighting
K-NN
Leaf color chart
Light intensity
Nitrogen nutrient
Rice plant

ABSTRACT

Crop management including the efficient use of nitrogen (N) fertilizer is important to ensure crop productivity. Human error in judging the leaf greenness when using the leaf color chart (LCC) to estimate the rice plant N nutrient status has encouraged numerous researchers to implement a machine-learning algorithm but experienced some issues in calibration and lighting. The datasets are created at 6.00-7.00AM (consistent lighting) and including light intensity, so each dataset contains RGB value and light intensity as inputs, and LCC value as a target. A system consists of a smartphone with an application that prevents user from taking an image if the light intensity is not in 2000-3500 lux, and a computer for preprocessing and classification purposes were developed. The preprocessing included cropping, splitting the rice leaf images, and calculating the average RGB values. A k-NN classifier is implemented and by using a cross-validation method is found k=5 gives the best accuracy of 97,22%. The in-site test of the system also works with an accuracy of 96.40%.

Corresponding Author:

Muliady Muliady
Department of Electrical Engineering
Universitas Kristen Maranatha
Jl. Surya Sumantri No.65, Bandung-40164, Jawa Barat Indonesia
Email: muliady@eng.maranatha.edu

1. INTRODUCTION

Food and agricultural policy research institute (FAPRI) and United Nations projected that the global rice demand from 2010 to 2035 will increase by 26%. On the other hand, in [1] showed that the rice yield from the top three rice production nations: India, China, and Indonesia grow very low, respectively only 1.0%, 0.7% and 0.4% improvement per year. Every agricultural practice that is used to improve the growth, development, and yield of crops is called crop management. It is important in maintaining the crop's health to ensure crop productivity. Advanced technology was used by numerous researchers in detecting the plant's health, mainly the plant's stress. The implementation of unmanned aerial vehicle (UAV) for spraying pesticides and crop monitoring was reviewed in paper [2], In [3] discussed on prospects of UAV technology for agricultural production management. Several types of camera sensors and imaging techniques were used to collect data and images [4-5]. Image processing and machine-learning algorithm implementation also give promising results for analyzing plant stress [6-8].

Y. Xue *et al.* [9] concluded nitrogen (N) management and irrigation using alternate wetting and moderate drying techniques used in improved crop management (ICM) could increase the rice yield, and make efficient use of N fertilizer and water compared to the local farmer practice. This paper focus on N fertilizer management, especially to estimate rice N deficiency. The implementation of advanced technology usually leads to high-cost investment, so the Irrigated rice research consortium (IRRC) introduced a new leaf color chart (LCC) and updated guidelines in many Asian countries [10]. In [11] while using this LCC had successfully guided an optimal application of N fertilizer and achieved high rice yield. The leaf color intensity is directly related to the leaf chlorophyll content which also shows the leaf N status. However, the use of N fertilizer was higher in LCC application compared to soil plant analysis development (SPAD) chlorophyll meter, due to the human error in judging the leaf greenness using LCC [12]. This human error can be reduced by using computer science that integrates with agricultural science [13].

Considering the popular usage of smartphones and the quality of the image that can be produced by today's smartphone camera, many researchers had used it to capture the leaf images for analyzing. Combined with the machine-learning algorithm, the promising results in analyzing and classifying the crop N status based on LCC have been conducted. In [14] used k-NN algorithm to assess the printed color images of LCC and relatively calibrate its color, but need further research when used to assess a real leaf color and match its color levels with its corresponding LCC. The classification of the rice leaf color using radial basis function neural network algorithm concluded that the lighting of rice images can affect the result of image classification. The accuracy of the classification using the backlight image is 90%, higher than the one with direct sunlight which is 80% [15]. The other research used support vector machine (SVM) with radial basis function kernel [16]. This methodology had an accuracy of 98% by considering four values LCC from value 1 to value 4. The limitation of this work is not considering value 5, which indicates a high concentration of N. One of the issues that need to be considered is the image acquisition process is affected by ambient light. In [17] had successfully reduced this effect by adding a white LED with an enclosure module to illuminate the leaf while the camera took its image, but this practice needed extra effort and time if used in a larger field.

2. RESEARCH METHOD

From the data provided in the introduction section, it can be concluded that the preface works agreed in using advanced technology. In addition, based on more references, it is possible to combine advanced technology with the effective proven practice of LCC to classify the N status. So the system introduced in this work is designed using a smartphone camera to capture the rice leaf images and using machine-learning algorithm to eliminate human error in classifying the leaf color based on LCC. The former research issue in classifying the rice N status using machine-learning algorithm excludes the lighting condition in data training. This work proposes a new perspective in solving the issue by having consistent lighting of the rice leaf images and adding light intensity data in the classifying process. This research workflow is designing a system, creating a dataset, building an android application for images acquisition, and building a classifier.

2.1. System design

The system input is a rice leaves image captured by a smartphone camera. It consists of 20 leaves picked randomly from an area of a rice field to be observed. These rice leaves will be arranged and stuck on a cardboard. The system consists of two parts, the first part is a smartphone as a data acquisition device and the second part is a computer in charge of the classifying process. The smartphone camera is used to record the rice leaves images and light intensity (lux) simultaneously. After transferring the data to the computer manually, the rice leaf images will be preprocessed and its light intensity will be read. The machine-learning algorithm in the computer will determine the average LCC value. Figure 1 shows the system design and process.

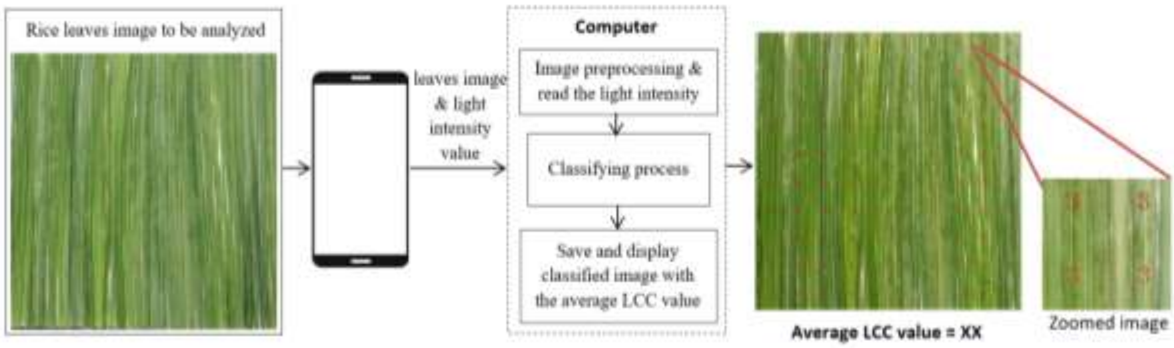

Figure 1. System design and process

2.2. Creating datasets

Numerous leaves are selected according to the international rice research institute (IRRI) guidance, which reflects each LCC value. The color of the leaves' middle part was compared with the LCC.The average LCC value is calculated from leaf images taken by a 1-megapixel smartphone camera. Direct sunlight affects leaf color readings and different people could interpret different readings, so the leaf color must be read and captured in a shading area, by the same person at the same time of the day to guarantee consistency of lighting data. According to [18], the recommended time to read and capture the leaf color is early morning (6 AM.) when the relative irradiance of the sun is higher in the visible and NIR spectrum compared to midday exposure (noon). This work had implemented this consistent lighting to all the images that were captured only from 6 to 7 AM. The light intensity in this period is generally between 2000 to 3500 lux. The leaf images then will be cropped into 100x100 pixels image as shown in Figure 2.

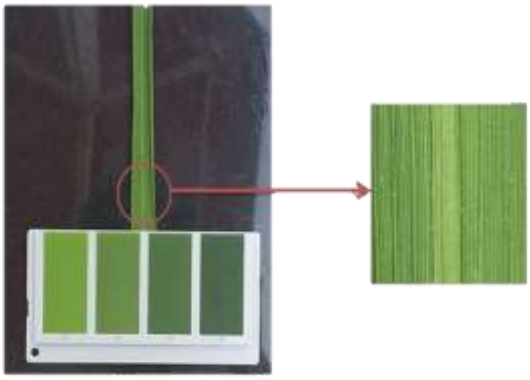

Figure 2. Image processing in creating datasets

While the leaf images are captured, a lux meter is used to measure the light intensity simultaneously. Figure 3 gives an illustration of the images in different light intensity that had been collected. The RGB value of each leaf image is read and averaged pixel by pixel. Each dataset consists of measured light intensity, average R value, average G value, average B value, and LCC value. A total of 120 leaf rice images are created as datasets for training and testing. The datasets are distributed in 40 datasets for each LCC value in 3 different light intensity. So each LCC value in specific light intensity is represented by 10 datasets. Table 1 shows the RGB average value of each LCC value.

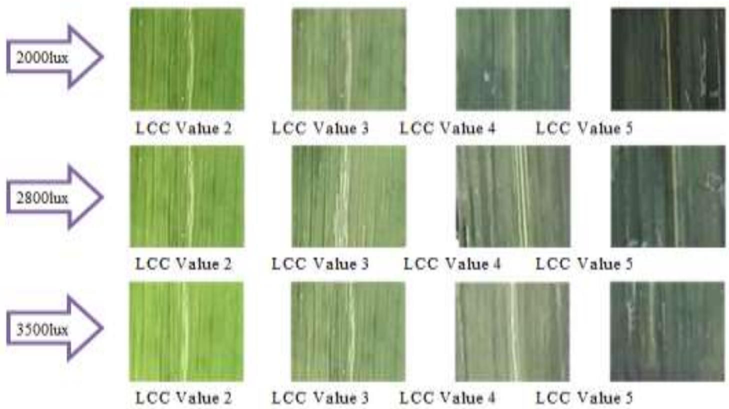

Figure 3. Process in creating datasets

The datasets in Table 1 show a significant difference in RGB value for each LCC value and the effects of the light intensity. Higher light intensity value results in higher RGB values and thus the datasets are qualified to be used as training data for machine-learning algorithm. Furthermore, these 120 datasets are separated into 84 datasets for training and 36 datasets for testing. The number of datasets collected in this work is enough to build a classifier. This is supported by the research [19] that used 75 training data and 25 testing data to classify a tomato maturity, in [20] used 115 images of rice leaf (70% for training and 30% for testing) to classify four rice diseases.

Table 1. Datasets of the classifier

Light intensity (lux)	Average value of			LCC value
	Red	Green	Blue	
2000	146	175	74	2
2000	117	152	76	3
2000	89	120	77	4
2000	88	105	79	5
2800	155	186	83	2
2800	129	165	84	3
2800	101	134	87	4
2800	99	119	90	5
3500	166	194	92	2
3500	141	173	94	3
3500	111	145	95	4
3500	110	127	97	5

2.3. Building the android application

An android application is built to capture and prevent users from taking an image if the light intensity is not in the range of 2000-3500 lux. The images should be taken in a shading lighting condition. Figure 4 shows the android application process in a flowchart. The process begins with smartphone camera and light meter initialization. The process will continue if user touches an image in this android application and then the application analyzes the light intensity. The measured light intensity value is shown on the top of the screen, along with a camera preview image. A warning message will show up if a user tries to capture the image in the light intensity that is not in the range. The application screenshots are shown in Figure 5.

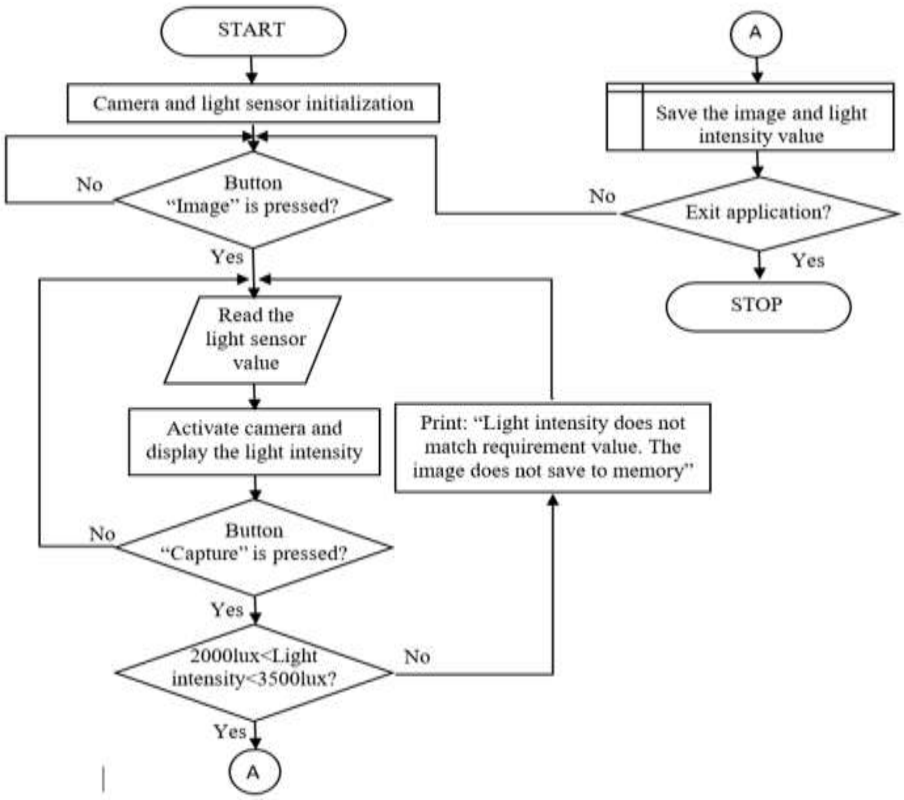

Figure 4. Flowchart of the image acquisition system

Figure 5. Screenshots of the android application

2.4. Building the classifier

The processing steps to determine the leaf LCC value in the computer are shown in Figure 6, which begins with the light intensity and the image file reading. Several preprocessing steps are needed before doing the classification using k-NN. First, the image is cropped into 1000x1000 pixels from the original smartphone image, then split into 100 pieces of images of 100x100 pixels. The RGB values of these split images are read and calculated, combined with the light intensity, both sets of values become the input. The k-NN classifier determines the leaf LCC value as output and attaches it to each piece of images. The last process is gathering all these images, calculating the average LCC value, and showing it to the user, as seen in Figure 1.

As the created datasets have RGB and light intensity values as inputs and LCC value as a target, in [21] suggests the application of supervised machine-learning algorithm to classify the leaf images' LCC value. Sometimes the literature on machine-learning classification implementation has conflict issues with each other on choosing an algorithm. In [22] introduced artificial neural networks (ANN), support vector machines (SVM), k-nearest neighbor (k-NN), single decision trees, boosted decision trees, and random forests as six relatively mature machine-learning algorithms. In [23-24] introduced classifiers that generally use in agriculture are SVM, k-NN, backpropagation neural network, and decision tree, which helps to narrow the option of choosing a suitable algorithm. Comparing k-NN and SVM use in detecting rice diseases, in [25] concluded that k-NN has better accuracy. Thus, k-NN is applied in this research because of its simplicity, effectiveness, nonparametric, and its wide range of usage in image and spatial classification [26].

The k-NN classifier is built using python programming language and scikit learn library. The k-NN works by calculating the distance of k nearest points to the training data. The hardest part of k-NN algorithm implementation is choosing the appropriate k value. In [26] recommended an experienced rule to determine k which is smaller or equal to the square root of the number of training data. Almost the same suggestion comes from paper [27] to select a small and odd k value, so this experiment chooses and compares k=5, 7, and 9. The final k value is obtained using a cross-validation confusion matrix discussed in [28], and the best accuracy within these three k values will be used in this research. The confusion matrix of each k value created by the testing data is shown in Figure 7. The accuracy, precision, and recall of each k value are obtained with the equation from [28] and the confusion matrix, is shown in Table 2. The best k value is 5, which succeeds in predicting 35 data correctly with only 1 failure. The accuracy of k-NN classifier with k=5 is 97.22%.

Table 2. Accuracy, precision, and recall of k=5, 7, and 9

k	Accuracy (%)	Precision value of the LCC value of (%)				Recall value of the LCC value of (%)			
		2	3	4	5	2	3	4	5
5	97.22	100	90	100	100	88.89	100	100	100
7	83.33	81.82	85.71	72.73	100	100	66.67	88.89	77.78
9	69.44	72.72	66.67	62.50	88.89	88.89	44.44	71.42	88.89

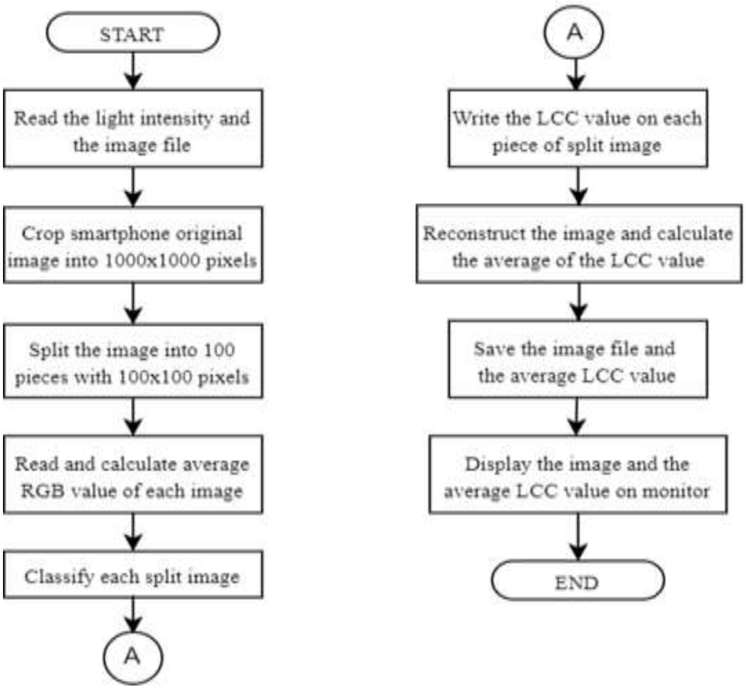

Figure 6. The process to determine the LCC value

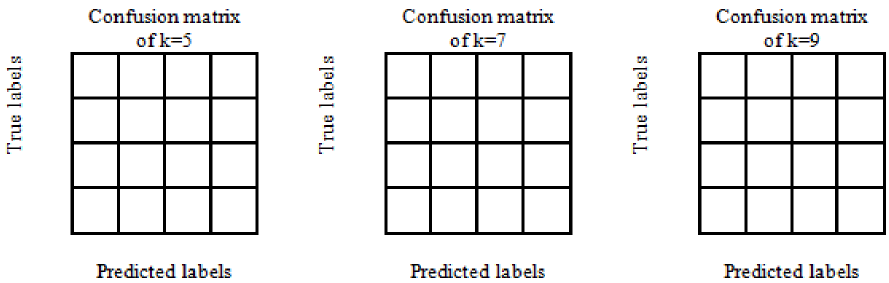

Figure 7. Confusion matrix of k=5, 7, and 9, there are missing numbers in the confusion matrix

3. RESULTS AND ANALYSIS

The in-site test was done by using several combinations of 20 rice leaves that stuck on a cardboard and the images are taken under the light intensity of 2000-3500 lux. The classified result is shown in Table 3. The average accuracy of the classification system is 96.40%. The image with a uniform leaves gives higher accuracy than the mixed one. Figure 8 shows several classified leaves images. Compare with the work in [17], which both of the works concern in the light intensity but using a different algorithm. This work has contributed to achieving higher accuracy with no additional module/cost.

Table 3. Classification test

Leaves sample combination	Light intensity	Average LCC	LCC value classification	Accuracy (%)
12 leaves of LCC value 2, 8 leaves of LCC value 3	2338 lux	2.4	2.31	96.25
5 leaves of LCC value 3, 15 leaves of LCC value 4	2864 lux	3.75	3.46	92.27
2 leaves of LCC value 3, 9 leaves of LCC value 4, 9 leaves of LCC value 5	3256 lux	4.35	4.67	92.64
20 leaves of LCC value 4	2023 lux	4	4.05	98.75
20 leaves of LCC value 4	2768 lux	4	4.03	99.25
20 leaves of LCC value 4	3567 lux	4	4.03	99.25

Classified LCC value 2.31

Classified LCC value 3.46

Classified LCC value 4.67

Figure 8. Classified leaves images

4. CONCLUSION

The research work was done successfully in classifying the rice plant N nutrient based on LCC. The high accuracy of prediction is achieved by including light intensity value in the datasets and have consistent lighting between 2000-3500 lux. The system can be improved by eliminating onsite work to pick and stick the leaves on cardboard and has a wider range of lighting intensity.

REFERENCES

[1] D. K. Ray, N. D. Mueller and P. C. West, and J. A. Foley, "Yield Trends Are Insufficient to Double Global Crop Production by 2050," *PLoS One*, vol. 8, no. 6, 2013, doi: 10.1371/journal.pone.0066428.

[2] U. R. Mogili and B. B. V. L. Deepak, "Review on Application of Drone Systems in Precision Agriculture," in *Procedia Computer Science*, vol. 133, pp. 502-509, 2018, doi: 10.1016/j.procs.2018.07.063.

[3] N. S., M. P., J. P. Sinha, H. L. Kushwaha and Dilip Kushwaha, "Prospect of Unmanned Aerial Vehicle (UAV) Technology for Agricultural Production," *Int. Conf. Emerg. Technol. Agric. Food Eng. 27-30th December, 2016, Agric. Food Eng. Dep. IIT Kharagpur*, no. December, 2016.

[4] S. Khanal, J. Fulton, and S. Shearer, "An overview of current and potential applications of thermal remote sensing in precision agriculture," *Computers and Electronics in Agriculture*, vol. 139. Elsevier B.V., pp. 22-32, 15-Jun-2017, doi: 10.1016/j.compag.2017.05.001.

[5] L. Deng, Z. Mao, X. Li, Z. Hu, F. Duan, and Y. Yan, "UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras," *ISPRS J. Photogramm. Remote Sens.*, vol. 146, no. March, pp. 124-136, 2018, doi: 10.1016/j.isprsjprs.2018.09.008.

[6] S. Ghosal, D. Blystone, A. K. Singh, B. Ganapathysubramanian, A. Singh, and S. Sarkar, "An explainable deep machine vision framework for plant stress phenotyping," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 18, pp. 4613-4618, May 2018, doi: 10.1073/pnas.1716999115.

[7] A. Shah, P. Gupta, and Y. M. Ajgar, "Macro-Nutrient Deficiency Identification in Plants Using Image Processing and Machine Learning," *2018 3rd Int. Conf. Converg. Technol. I2CT 2018*, pp. 1-4, 2018, doi: 10.1109/I2CT.2018.8529789.

[8] A. Singh, B. Ganapathysubramanian, A. K. Singh, and S. Sarkar, "Machine Learning for High-Throughput Stress Phenotyping in Plants," *Trends Plant Sci.*, vol. 21, no. 2, pp. 110–124, 2016, doi: 10.1016/j.tplants.2015.10.015.

[9] Y. Xue, H. Duan, L. Liu, Z. Wang, J. Yang, and J. Zhang, "An Improved Crop Management Increases Grain Yield and Nitrogen and Water Use Efficiency in Rice," no. february, pp. 271-284, 2013, doi: 10.2135/cropsci2012.06.0360.

[10] B. C. Witt, R. Mutters, and R. J. Buresh, "New Leaf Color Chart for Effective," *Better Crop.*, vol. 89, no. 1, pp. 36-39, 2005.

[11] N. Ahmad, A. Zada, M. Junaid, and A. Ali, "Bridging the Yield Gap in Rice Production by Using Leaf Color Chart for Nitrogen Management," *J. Bot.*, vol. 2016, 2016, doi: 10.1155/2016/2728391.

[12] M. Z. Hussain, S. A. Khan, and T. M. Thiyagarajan, "Increasing Nitrogen Use Efficiency in Rice (Oryza Sativa L.) with Chlorophyll Meter and Leaf Color Chart," *IUP J. Soil Water Sci. Vol II No 4 pp 3654 Novemb. 2009*, no. November 2009, 2009, doi: 10.2139/ssrn.1513199.

[13] S. J. Leghari *et al.*, "Modern Leaf Color Chart Successfully Prepared and Used in Crop Production of Sindh, Pakistan," *Eur. Acad. Res.*, vol. 4, no. May, pp. 900-916, 2016.

[14] Y. A. Sari, R. V. H. Ginardi, and R. Sarno, "Assessment of Color Levels in Leaf Color Chart Using Smartphone Camera with Relative Calibration Assessment of Color Levels in Leaf Color Chart Using Smartphone Camera with Relative Calibration," no. April, pp. 2-4, 2016.

[15] R. F. Rahmat, Chairunnisaq, E. B. Nababan, and Onrizal, "Classification of rice plant fertilizer needs based on leaf color chart using radial basis function neural network," *J. Phys. Conf. Ser.*, vol. 1116, no. 2, 2018, doi: 10.1088/1742-6596/1116/2/022037.

[16] P. K. Sethy, B. B. Nayak, N. K. Barpanda, and A. K. Rath, "Rice Nitrogen Status Estimation Of Western Tract Of Odisha Using SVM Based On Color Feature: A Comparative Analysis With LCC," *Int. J. Res. Advent Technol.*, vol. 7, no. 4, pp. 397-400, 2019, doi: 10.32622/ijrat.742019152.

[17] R. D. Tipones and J. C. Dela Cruz, "Nutrient Management with Automated Leaf Color Level Assessment and Ambient Light Neutralizer using SVM Classifier," *2019 Int. Symp. Multimed. Commun. Technol. ISMAC 2019*, no. Lcc, 2019, doi: 10.1109/ISMAC.2019.8836149.

[18] D. Barolet, F. Christiaens, and M. R. Hamblin, "Infrared and skin: Friend or foe," *J. Photochem. Photobiol. B Biol.*, vol. 155, no. January 2016, pp. 78-85, 2016, doi: 10.1016/j.jphotobiol.2015.12.014.

[19] W. D. N. Pacheco and F. R. J. López, "Tomato classification according to organoleptic maturity (coloration) using machine learning algorithms K-NN, MLP, and K-Means Clustering," *2019 22nd Symp. Image, Signal Process. Artif. Vision, STSIVA 2019 - Conf. Proc.*, 2019, doi: 10.1109/STSIVA.2019.8730232.

[20] A. A. Joshi and B. D. Jadhav, "Monitoring and controlling rice diseases using Image processing techniques," *Int. Conf. Comput. Anal. Secur. Trends, CAST 2016*, pp. 471-476, 2017, doi: 10.1109/CAST.2016.7915015.

[21] M. Mohammed, M. B. Khan, and E. B. M. Bashie, "Machine learning: Algorithms and applications," no. July. 2016.

[22] A. E. Maxwell, T. A. Warner, and F. Fang, "Implementation of machine-learning classification in remote sensing: An applied review," *Int. J. Remote Sens.*, vol. 39, no. 9, pp. 2784-2817, 2018, doi: 10.1080/01431161.2018.1433343.

[23] B. S. Anami, N. N. Malvade, and S. Palaiah, "Classification of yield affecting biotic and abiotic paddy crop stresses using field images," *Inf. Process. Agric.*, 2019, doi: 10.1016/j.inpa.2019.08.005.

[24] C. O. F. Healthy, D. V. Leaves, and U. C. Features, "Classification of healthy and diseased vine leaves," vol. 3, no. 4, pp. 296-302, doi: 10.15547/artte.2015.04.002.

[25] K. Jagan, M. Balasubramanian, and S. Palanivel, "Detection and Recognition of Diseases from Paddy Plant Leaf Images," *Int. J. Comput. Appl.*, vol. 144, no. 12, pp. 34-41, 2016, doi: 10.5120/ijca2016910505.

[26] J. Sun, W. Du, and N. Shi, "A Survey of kNN Algorithm," *Information Engineering and Applied Computing*, vol. 1, no. 1, pp. 1-10, 2018.

[27] M. J. Islam, Q. M. J. Wu, M. Ahmadi, and M. A. Sid-Ahmed, "Investigating the Performance of Naive- Bayes Classifiers and K- Nearest Neighbor Classifiers," pp. 1541-1546, 2008, doi: 10.1109/iccit.2007.148.

[28] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, "An improved method to construct basic probability assignment based on the confusion matrix for classification problem," *Inf. Sci. (Ny).*, vol. 340-341, pp. 250-261, 2016, doi: 10.1016/j.ins.2016.01.033.

BIOGRAPHIES OF AUTHORS

Muliady Muliady received his S.T.(B.Eng.) degree in Electrical Engineering from Universitas Kristen Maranatha, Indonesia. Received his M.T.(M.Eng.) degree in Instrumentation and Control from Institut Teknologi Bandung, Indonesia. Currently, he is in the second year of Ph.D. degree at Multimedia University, Melaka-Malaysia. His research interests include robotics, UAV in precision agriculture, and machine learning. He is a senior lecturer at the Electrical Engineering Department, Universitas Kristen Maranatha, also coaching secondary, and undergraduate students for robotics competition

Associate Professor Dr Tien Sze LIM is currently an Associate Professor with the Faculty of Engineering and Technology, Multimedia University. His research interests include remote sensing technologies, drone design, signal processing and embedded system design. Assoc. Prof Lim has been a principal consultant for various government agencies in remote sensing projects and has published more than 50 papers in refereed journals, presented at international conferences, written 2 books, and holds 2 patents. Dr. Lim is a member of The Institute of Electrical and Electronics Engineers (IEEE) and member of Board of Engineers Malaysia.

Prof. Koo is currently a full Professor of Multimedia University, Malaysia. His research interests include remote sensing technologies, signal processing, data analytics and embedded system design. He has published more than 100 papers in refereed journals, international conferences, 4 books, and 11 patents. Prof. Koo is a registered Professional Engineer, a Fellow of the ASEAN Academy of Engineering and Technology (AAET), a senior member of IEEE, and the past Chair of the IEEE Geoscience and Remote Sensing Society Chapter, Malaysia Section.

Suhadra Pattra was born in Bandung, Indonesia, in 1997. He received the S.T. (B.Eng.) degree in electrical engineering from the University Kristen Maranatha, Bandung, Indonesia, in 2020. Since then, he has been with Pattra Aksa Jaya, Bandung, Indonesia, where he is currently the staff of R and D.