MEKATRONIKA
e-ISSN: 2637-0883
VOL. 2, ISSUE 1, 38 – 46
DOI: https://doi.org/10.15282/mekatronika.v2i1.6727

ORIGINAL ARTICLE

Features Extraction of Capsicum Frutescens (C.F) NDVI Values using Image Processing

Suhaimi Puteh[1], Nurul Fadhilah Mohamed Rodzali[1], Mohd Azraai Mohd Razman[1*], Zelina Zaiton Ibrahim[2], Muhammad Nur Aiman Shapiee[1] and Mohd Azhar Mohd Razman[3]

[1]Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pahang, Malaysia.
[2]Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
[3]Sembranics Sdn Bhd, 68000 Ampang, Selangor, Malaysia.

ABSTRACT – There is yet an application for monitoring plant condition using the Normalized Difference Vegetation Index (NDVI) method for Capsicum Frutescens (C.F) or chili. This study was carried out in three phases, where the first and second phases are to create NDVI images and recognize and extract features from NDVI images. The last stage is to assess the efficiency of Neural Network (N.N.), Naïve Bayes (N.B.), and Logistic Regression (L.R.) models on the classification of chili plant health. The images of the chili plant will be captured using two types of cameras, which can be differentiated by whether or not they have an infrared filter. The images were collected to create datasets, and the NDVI images' features were extracted. The 120 NDVI images of the chili plant were divided into training and test datasets, with 70.0% training and 30.0% test. The extracted data was used to test the classification accuracy of classifiers on datasets. Finally, the N.N. model was found to have the highest classification accuracy, with 96.4 % on the training dataset and 88.9 % on the test dataset. The state of the chili plant can be predicted based on feature extraction from NDVI images by the end of the study.

ARTICLE HISTORY
Received: 18th April 2020
Revised: 30th May 2020
Accepted: 7th June 2020

KEYWORDS
Features Extraction
NDVI
Chili plant
Machine Learning
Image processing

INTRODUCTION

Agriculture is one of the most critical roles in the economy of Malaysia. Thus, it is crucial to refining the method of farming industry operation. Based on the Department of Statistics Malaysia, the agriculture sector had contributed 7.1% to the Gross Domestic Product (GDP) in 2019, which is equal to RM 101.5 billion. Several selected crops show the production increment each year, and one of them is the chili plant. Bird's eye chili, also known as *Capsicum Frutescens* (C.F), was chosen as the research subject. There is a remarkably high demand from the daily household market and manufacturing sector. The production of chilies has slightly escalated from 32.300 tonnes in 2018 to 33.900 tonnes in 2019 [1].

Farm management is an essential method to supervise the agricultural economy in the agriculture sector, whether in prices, market policy, and economic institutions. In farm management, monitoring crop is necessary to ensure crops can achieve their health and productivity level at optimum [2]. Reference to the Normalized Difference Vegetation Index (NDVI) is one of the measures used to read the state of the plant. The range of NDVI is from -1.0 to +1.0, which is derived from remotely sensed images in the red (RED) and near-infrared (NIR) bands [3]. The formula used for NDVI is as mentioned below:

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

The proposed research aims to create a system that uses image processing to assess the health of plants based on NDVI values. The emphasis will be on NDVI images of chilli plants, categorized as healthy, unhealthy, or dead. By collecting and extracting data from NDVI images of the plants, such as mean, maximum, minimum, variance, and standard deviation values, the system may calculate the plant's condition. This data is essential because it allows for estimating performance assessment needs for Machine Learning (ML) models as classifiers.

The remaining part of the paper proceeds with related works according to research followed by the methodology that discusses the image processing system design and data collection to acquire the development of NDVI images. Hence, continue by laying out the results as well as an associated discussion before concluding the paper.

*CORRESPONDING AUTHOR | Mohd Azraai Mohd Razman | ✉ mohdazraai@ump.edu.my

RELATED WORK

Many studies have been published on the use of NDVI to determine the condition of plants, differentiate between healthy and diseased plants, or determine the amount of chlorophyll in plant leaves. The NDVI was applied to several plants in the studies. In 2019, lettuce leaves were used to estimate disease intensity on the plant. Another report was published in which the leaves of wheat powdery mildew were used to forecast the plant's condition in 2020 using Unmanned Aerial Vehicles (UAV) imageries [4]. Since NDVI is closely related to the near-infrared range, a method for recognizing near-infrared using a smartphone's camera has been suggested. Three separate plant species (Japanese mock-orange cheesewood, Red gum eucalyptus, and Tea rose) were used in this analysis [5].

Following that, a number of studies were performed in which UAV was used to capture NDVI photos of plants. Bhandari's study used the PrecisionHawk Lancaster 5 UAV, which comes pre-equipped with a variety of sensors while Costa used a different kind of UAV in his research. To take NDVI images, he installed a non-calibrated RGB camera on the UAVs. [6], [7]. Sankaran suggested in his study a way of using two six-band cameras and a thermal camera instead of a UAV [8]. To summarise, there are a variety of methods for evaluating NDVI images.

Images can contain a variety of features that can be extracted. One of them is the statistical format. To estimate NDVI values of vegetation, the coefficient of determination (R^2), mean absolute error (MAE) and mean percentage error (MPE) were extracted as features [7]. In his research to determine soil colour by spectral bands and indices, Parviz compared the ANN model to multivariate regression using the percentage of mean absolute error (MAE) decline [9]. Liangju Wang used the coefficient of determination and root mean square error (RMSE) as features in his estimation of NDVI with a simple NIR sensitive RGB camera [10].

The use of ML techniques was extensive, particularly for elucidating significant factors or procedures that would lead to the most optimal classifiers for significant variables. Back Propagation (B.P.), Support Vector Machine (SVM), and Random Forest (R.F) models were trained to estimate the average values of chlorophyll material, according to Guo's study. In the analysis of Sankaran and Ehsani in 2011, Quadratic Discriminant Analysis (QDA) and Soft Independent Modeling of Classification Analogies (SIMCA) were used to identify data based on selected citrus features [11]. Li Wang also discussed the use of Support Vector Regression (SVR), Artificial Neural Network (ANN), and (R.F). The best SVR, R.F, and ANN models decreased prediction error by 29.91%, 42.74 %, and 28.21 %, respectively, using the root mean square error (RMSE) validation metric [12]. Convolutional Neural Network (CNN) models also can be used to classify datasets. Agarwal used Simplified CNN in a study in which he needed to identify tomato crop diseases. This model's accuracy is 98.4%, comparable to the accuracy level in leaf-based classification [13].

In conclusion, literature reviews on various NDVI articles were conducted to gain a thorough understanding of feature extraction and analysis. As a consequence, this will adhere to the research methodology developed in this report.

METHODOLOGY

The stages are divided into three subtasks: selecting the appropriate camera type, input variables for image processing based on NDVI values, and ML classification and parameter tuning. These sections are arranged logically to obey the complete verdicts and test the model's accuracy. The images captured by the camera must have a certain parameter, which is the infrared and red wavelength, defined in the first step. These parameters will be looked into since they are the most significant when measuring the NDVI. As a result, the NDVI image can be created. The experimental setup has been associated with image processing and machine learning techniques, and methods for capturing these parameters will also be explored. The second phase of this study will use image processing techniques to analyze the condition of the chili plants. The plant images will be analyzed first using image processing to extract the images' features using NDVI index that will decide the condition of the plants. The final step of this study will concentrate on using machine learning to classify the state of chili plants. The plant's condition will be manipulated by parameters that will be measured and checked. The research methodology's flowchart framework is clarified to illustrate the whole research operation in Figure 1.

The NDVI is a scale that ranges from -1.0 to 1.0 and is used as a plant health indicator [3]. Inanimate or dead objects, such as rocks, houses, or dead plants, are typically represented by negative values to zero. Plants with a value between 0 and 0.33 are unhealthy or depressed, whereas between 0.33 and 0.66 are moderately healthy. Finally, the range of 0.66 to 1.0 signifies a relatively healthy condition [14]. However, this number can vary depending on the type of plant and other factors. The mean values of the NDVI images were used to assess the condition of the chili plants in this study.

This study aims to use image processing based on NDVI values and machine learning techniques to classify the condition of chilli plants. The NDVI values were calculated using a camera without an infrared filter and an infrared filter. NDVI were created based on the photos taken. The state of the plants was then evaluated based on the NDVI values to decide whether they were healthy, unhealthy, or dead. The NDVI values for each plant are studied and evaluated to assess the plant's status and satisfaction by looking at the plant from a distance. The images were then transferred to a computer and analyzed by image processing. Image processing methods are used to extract the variables, which includes the NDVI values. ML algorithms such as N.B., N.N. and L.R. are being evaluated to see how accurate they are at classifying plant conditions.

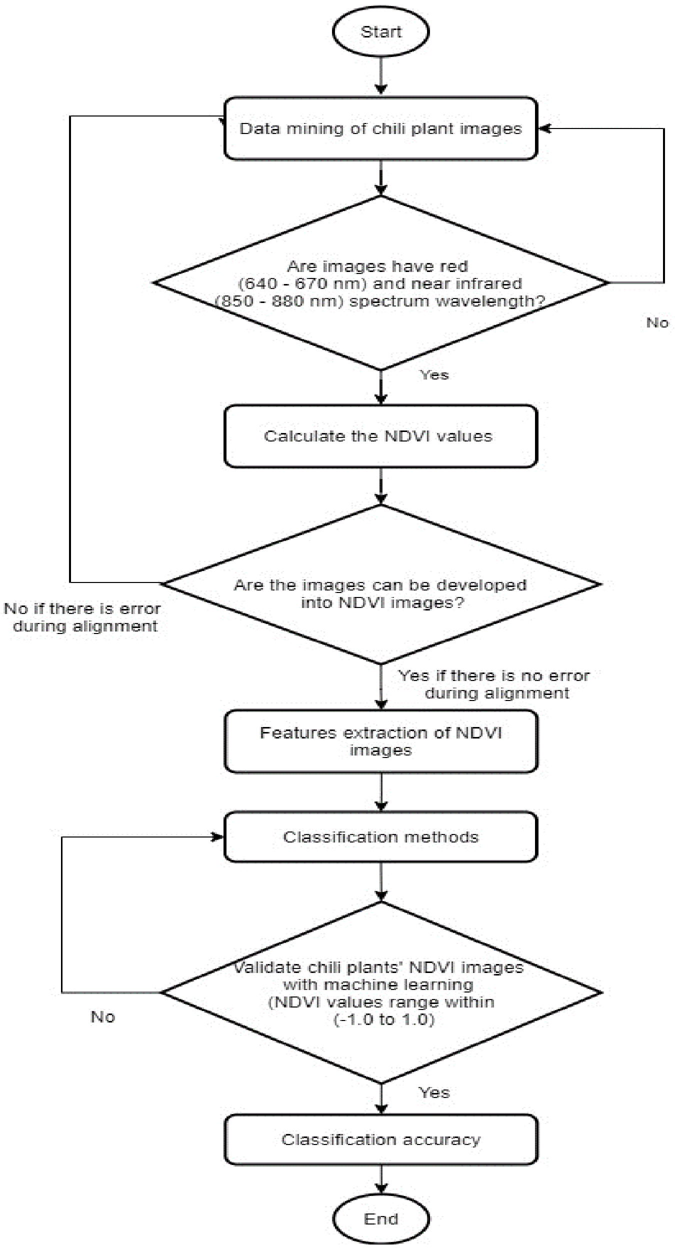


Figure 1. The experimental flowchart

Experimental setup

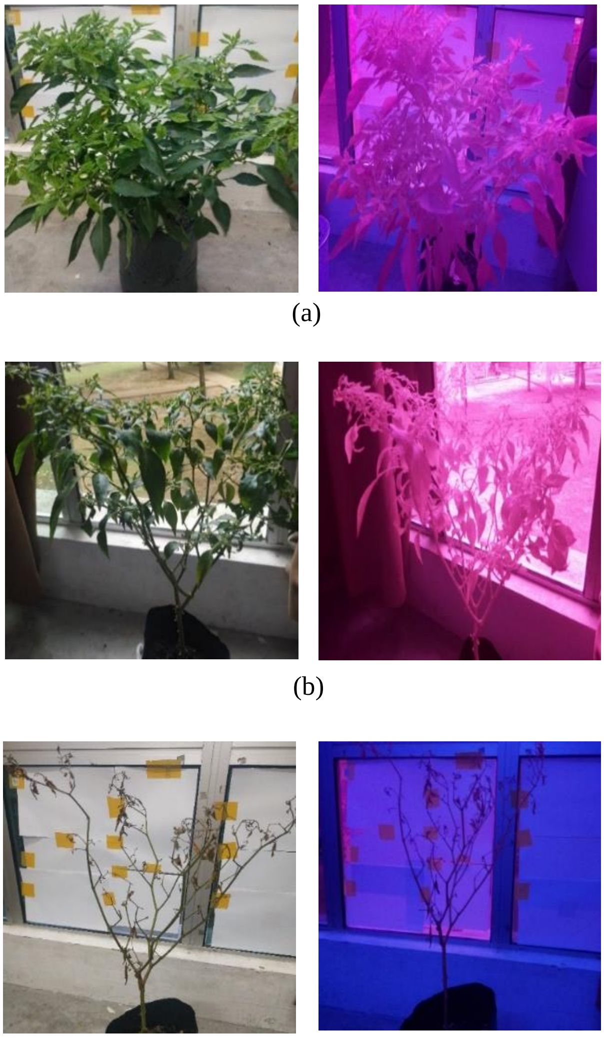
The data collection for the chili plants took place in a laboratory at Universiti Malaysia Pahang. The camera is close to the plant, but the distance between the camera and the plant should at least cover the entire chilli plant, as the camera would crop to isolate the external environment. The experiment is set up as shown in Figure 2 on the following page. The smartphone camera was positioned in the same position as the PiCamera to catch the same perspective as the PiCamera. The camera and the plant are about 20 centimetres apart. Since there are three types of plants, they are classified into good, unhealthy, and dead. The explanation for this is because the leaf's colour or pigment will vary since unhealthy leaves contain less chlorophyll than healthy leaves, whereas dead plant leaves contain no chlorophyll. After all, the pigments of the leaves have already shifted from green to brown. Before collecting plant data, the plants' images will be evaluated to measure the NDVI value of each plant. The plants would then be labelled as healthy, unhealthy, or dead depending on the values, as determined by the NDVI value index. The findings can also be justified using an existing index that has been used internationally.

In this study, two types of cameras were used. Smartphones have cameras capable of capturing visible spectral bands such as red, green, and blue, in which Redmi Note 5 was used. NoIR PiCamera is the second camera that was used. This camera was chosen for this study because it is a low-cost camera that can capture the NIR spectrum, which is not a visible spectral band seen through the naked eye.

Figure 2. Experiment setup

Data collection

For this research, three types of datasets are used. The data was obtained from plants that were healthy, unhealthy, and dead. These images were taken with a NoIR PiCamera and a smartphone camera, respectively, for the first and second datasets as shown in Figure 3. In total, nine photos from each class were obtained for healthy, unhealthy, and dead plants. Later, one of the data augmentation methods will be used to increase the size of these images. The process involves cropping a picture of a plant's leaf to improve the accuracy of the plant's condition.



(a)

(b)

(c)

Figure 3. Images (a) healthy (b) unhealthy and (c) dead chili plant taken with infrared filter (using smartphone's camera) and without infrared filter (using NoIR PiCamera).

Formation of NDVI image

The NDVI image was created by combining two images. The NIR spectrum was read using images from the NoIR PiCamera, while the red spectrum was read using images from smartphones. OpenCV is used to perform image alignment techniques to combine the two images. As previously mentioned, the process begins with reading the photos. After that, convert the images to grayscale, so only the NIR and RED bands need to be extracted. After that, convert the images to grayscale, so only the NIR and RED bands need to be processed. After that, a motion model must be specified. There are four different kinds of motions. However, only two motions, Translation and Homography, are involved out of four. The translation is a motion in which only the image's x and y axes must be estimated, whereas Homography is the most sophisticated transition, which requires some 3D effects [15]. Transform ECC, on the other hand, is an iterative optimization feature [16].

Consequently, a criterion must be established for the iteration to stop and the process to proceed with the find transform ECC. After the model has been calculated, it can be used to align and view two images. The NDVI equation was used to calculate the NDVI values of the images after these measures were completed. Hence, the third dataset will comprise NDVI images of chili plants, as shown in Figure 4. These values will be extracted as features and used as a classification indicator for the plant groups. The mean, maximum, minimum, variance, and standard deviation values of each NDVI image were extracted as features.

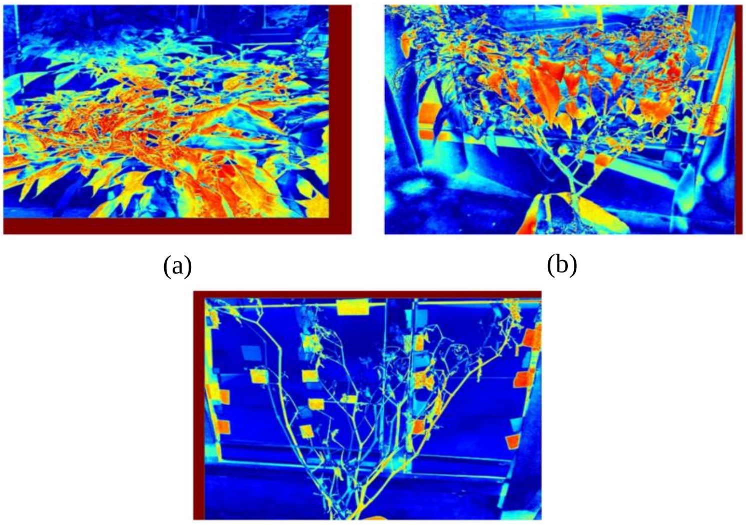


Figure 4. Illustration of images of (a) healthy (b) unhealthy and (c) dead chili plant transforms into NDVI image

Machine Learning

ML is a data analysis methodology that automates the creation of analytical models. It's a subset of artificial intelligence focused on the premise that computers can learn from data, recognize patterns, and make decisions with little to no human involvement [17]. The proposed approach of ML in this study is to classify significant models that will lead to the most optimum classifiers concerning the significant features. The training and testing datasets were split using a 70:30 ratio on 120 NDVI images, respectively.The ML models will classify each NDVI image as healthy, unhealthy or dead plant.

Classifiers: Neural Network, Naïve Bayes and Logistic Regression

Neural Network (N.N.) is one of the best classifiers in ML. Weighted synapses link input and output neurons in an N.N. The weights influence the amount of forward propagation that passes through the N.N. The weights can then be modified during back propagation, when the neural network starts to learn. The larger the data set and the more diverse the data set, the more the neural network will learn and the better it will be at predicting outputs [18]. Other than N.N., Naïve Bayes (N.B.) algorithm also acts as a classifier. This algorithm was created using Bayes' Theorem and the assumption of predictor independence. The existence of a specific feature in a class is assumed to be unrelated to the presence of any other feature by a N.B. classifier. It employs a similar approach to estimate the likelihood of various classes based on various attributes. This algorithm is usually used for text classification and problems with multiple groups. The N.B. model is simple to create and is particularly useful for large data sets. N.B. is considered to outperform even the most advanced classification methods due to its simplicity [19]. Lastly, Logistic Regression (L.R.) as a classifier is used in statistics to estimate (guess) the likelihood of an occurrence occurring based on previous data.

L.R. operates with binary data, in which an occurrence occurs (1) or does not occur (0) [20]. A threshold is typically defined when using it, indicating at what value the example will be classified into one of two groups. The "one vs all" approach can be used on multi-class classification problems. L.R. is a robust ML algorithm that uses a sigmoid function that works best on binary classification problems [21].

Performance Evaluation

Classification accuracy is one of the most basic yet most important criteria for evaluating classifiers because it reflects the consistency of the model that was used as a classifier in a study or analysis. The fraction of correct predictions calculated by accuracy is also known as "the ratio of correct predictions to total predictions made" [21]. The models were developed and evaluated on a Python IDE and Spyder 3.7.

EXPERIMENTAL RESULTS

The evolution of NDVI was addressed in the earlier section. Both images from the NoIR Picamera on the Raspberry Pi 4B and the camera on the smartphone were moved to the same folder on the machine for Python processing. As a result, NDVI images of chili plants based on NDVI values were produced, as shown in Figure 4 on the previous page.

Following that, the NDVI values measured and collected during the creation of NDVI images previously will be used to describe the plant's condition. Python 3 will be used for this form. As shown in Table 1 on the next page, five features of the NDVI images were extracted: mean, maximum, minimum, variance, and standard deviation.

Table 1. Features extraction and condition of chili plant based on NDVI

Image	Mean	Maximum	Minimum	Variance	Standard Deviation	Condition
A1	0.175	1	-1	0.312	0.558	Unhealthy
A2	0.273	1	-1	0.255	0.505	Unhealthy
A3	0.473	1	-1	0.160	0.400	Healthy
A4	0.292	1	-1	0.172	0.415	Unhealthy
A5	0.187	1	-1	0.308	0.555	Unhealthy
A6	0.110	1	-1	0.330	0.575	Unhealthy
A7	0.576	1	-0.195	0.041	0.203	Healthy
A8	0.489	1	-1	0.136	0.368	Healthy
A9	0.399	1	-1	0.129	0.360	Healthy
A10	0.222	1	-1	0.235	0.483	Unhealthy
A11	0.259	1	-1	0.233	0.483	Unhealthy
A12	0.260	1	-1	0.190	0.437	Unhealthy
A13	0.652	1	-1	0.061	0.247	Healthy
A14	-0.020	0.991	-1	0.150	0.387	Dead
A15	-0.053	0.840	-1	0.428	0.654	Dead
A16	-0.007	0.618	-1	0.196	0.443	Dead
A17	-0.200	0.638	-1	0.441	0.664	Dead
A18	-0.047	0.818	-1	0.336	0.580	Dead
A19	-0.081	0.809	-1	0.323	0.568	Dead

The extracted features from 120 NDVI images will be used to characterize the condition of the plant. The plants will be divided into three categories: healthy, unhealthy, and dead. Overall, there are 120 sets of data with five distinct features grouped into one of three classes using three different types of classifiers (N.B., N.N., and L.R.). Furthermore, the datasets have been split into 70.0% for training and 30.0% for testing.

The classification accuracy between the two datasets could be seen in Figure 5 on the following page: training and test datasets. It is a vital technique for evaluating the best-performing models with high classification precision. The chart shows that N.N. models have the highest classification accuracy on the training dataset with 96.4% but have an equal value with N.B. on test datasets, 88.9%. However, it cannot be identified as a stable classifier because it has a 7.5% difference between the two datasets. On the other hand, N.B. has 90.5% classification accuracy on training datasets, making this model the most stable classifiers than the other two classifiers because the difference of classification accuracy is only 1.6%. With only 78.6% and 72.2% of classification accuracy for training and test datasets, L.R. ranked the lowest. It is better to use N.N. or N.B. classifiers, as it does not achieve more than 80.0% accuracy. On top of that, the difference between training and test datasets is quite significant (6.4%). Overall, these findings suggest that ML models performed satisfactorily since all of the data in the graph is underfit since the percentage of training and datasets is higher than the percentage of test datasets.

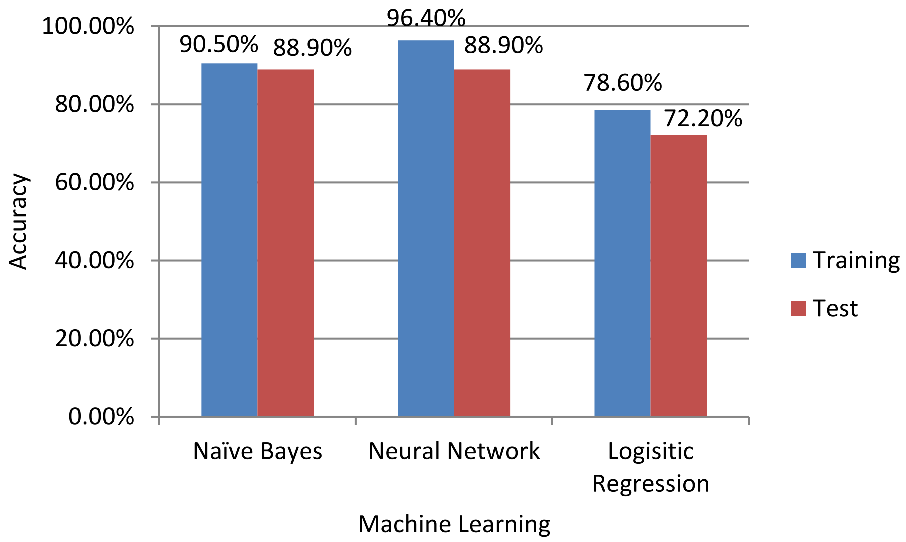
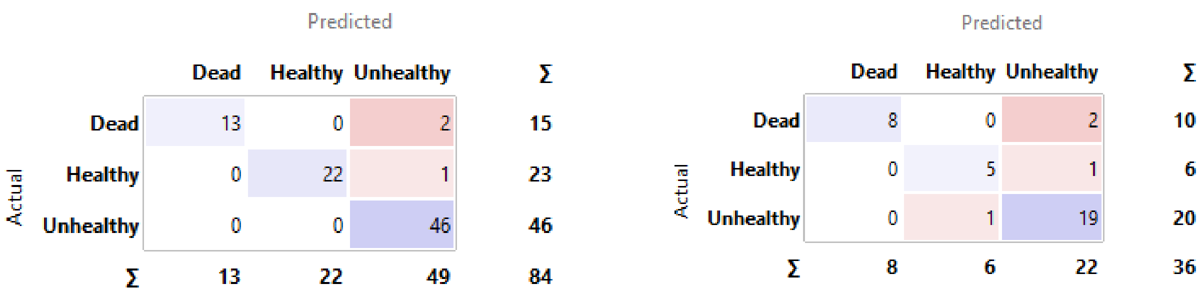


Figure 5. Classification accuracy of ML models



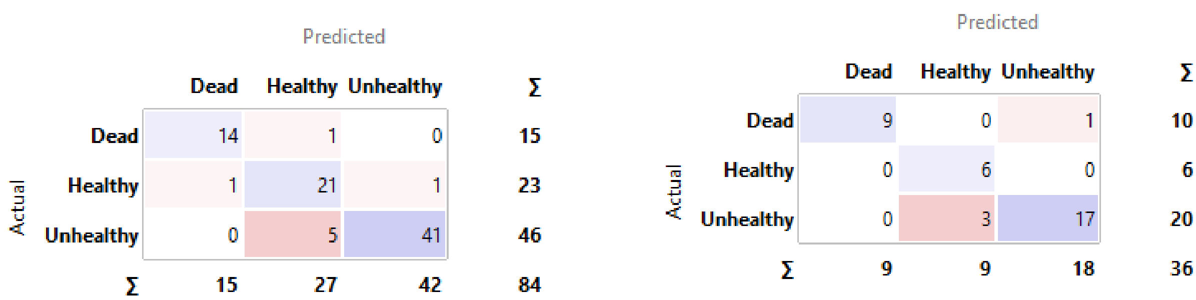
(a)

Actual \ Predicted	Dead	Healthy	Unhealthy	Σ
Dead	13	0	2	15
Healthy	0	22	1	23
Unhealthy	0	0	46	46
Σ	13	22	49	84

(b)

Actual \ Predicted	Dead	Healthy	Unhealthy	Σ
Dead	8	0	2	10
Healthy	0	5	1	6
Unhealthy	0	1	19	20
Σ	8	6	22	36

Figure 6. Confusion matrix of Neural Network from (a) training and (b) test datasets



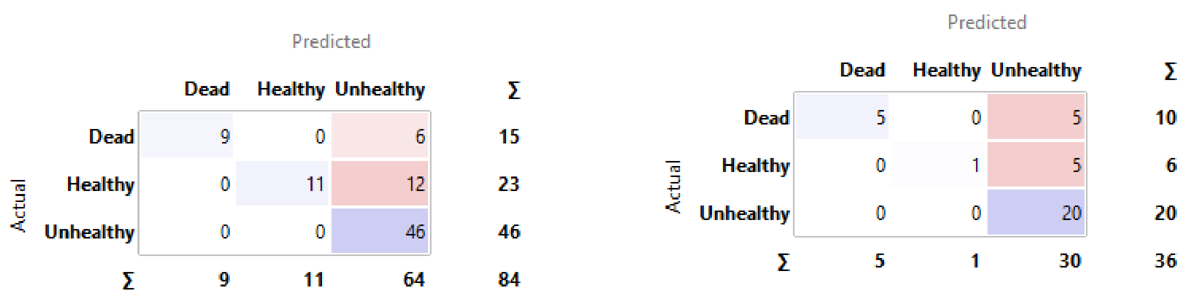
(a)

Actual \ Predicted	Dead	Healthy	Unhealthy	Σ
Dead	14	1	0	15
Healthy	1	21	1	23
Unhealthy	0	5	41	46
Σ	15	27	42	84

(b)

Actual \ Predicted	Dead	Healthy	Unhealthy	Σ
Dead	9	0	1	10
Healthy	0	6	0	6
Unhealthy	0	3	17	20
Σ	9	9	18	36

Figure 7. Confusion matrix of Naïve Bayes from (a) training and (b) test datasets



(a)

Actual \ Predicted	Dead	Healthy	Unhealthy	Σ
Dead	9	0	6	15
Healthy	0	11	12	23
Unhealthy	0	0	46	46
Σ	9	11	64	84

(b)

Actual \ Predicted	Dead	Healthy	Unhealthy	Σ
Dead	5	0	5	10
Healthy	0	1	5	6
Unhealthy	0	0	20	20
Σ	5	1	30	36

Figure 8. Confusion matrix of Logistic Regression from (a) training and (b) test datasets

According to their classifiers, the results attained from training and test datasets were shown in Figures 6, 7, and 8. In both training and test datasets, L.R. has the most significant number of misclassified instances compared to N.B. and N.N. In training datasets, L.R. misclassified 18 out of 84 instances, while N.B. and N.N. misclassified only eight and three instances, respectively. In test datasets, the number of misclassified groups for N.N. and N.B. is the same: four out of 34 cases. However, L.R. yielded ten misclassified chili plant classifications. As a result, it has been established that L.R. is not the best classifier for this analysis.

To summarize the results, the principles of implementing the proposed method provide a clearer understanding of how accuracy can be enhanced. First, the use of the NoIR PiCamera and the camera on a smartphone helps to understand the study's target. In assessing the performance of ML models, the extracted features are a vital part of classifying the plant groups. Overall, the chapter's findings show that N.N. (96.4% and 88.9%) and N.B. (90.5% and 88.9%) have the best results with training and test datasets.

CONCLUSION

This paper reports the first attempt to classify the capability of the chili plant condition by applying NDVI through two types of cameras. The creation of NDVI images of chili plants was the first significant discovery. In calculating NDVI values, the NIR and RED spectrum are extremely important. Since the accuracy of the values would be compared with the NDVI index to identify the condition of the plants, whether they are healthy, unhealthy, or already dead, the process of features extraction for each image was performed manually. N.N., N.B., and L.R. algorithms were used in this analysis. ML models can measure the extracted mean, minimum, maximum, variance, and standard deviation values.

In conclusion, N.N. and N.B. have an excellent performance in classification accuracy as they achieved 88.9%. The state of the chili plant can be predicted using features extracted from NDVI images by the end of the study. Finally, and perhaps most significantly, this study's results have a range of implications and improvements for future practice. The results of this study on ML classification accuracy can be used to guide future studies by integrating hyperparameters into the models to enhance classification accuracy.

ACKNOWLEDGEMENT

The authors would like to acknowledge Universiti Malaysia Pahang for funding this study under the Research Grant (RDU200332).

REFERENCES

[1] Mohd Yusrizal bin Ab. Razak, "Department of Statistics Malaysia Press Release," *Dep. Stat. Malaysia*, no. May, pp. 5–9, 2020.

[2] M. E. Bliss, "Farm management | agriculture | Britannica," Mar. 10, 2017. https://www.britannica.com/topic/farmmanagement

[3] "What is NDVI (Normalized Difference Vegetation Index)? - GIS Geography", 2020. https://gisgeography.com/ndvinormalized-difference-vegetation-index/

[4] F. Pham, A. Raheja, and S. Bhandari, "Machine learning models for predicting lettuce health using UAV imageries," no. May 2019, p. 24, 2019, doi: 10.1117/12.2519157.

[5] S. Chung, L. E. Breshears, and J. Y. Yoon, "Smartphone near infrared monitoring of plant stress," Comput. Electron. Agric., vol. 154, no. April, pp. 93–98, 2018, doi: 10.1016/j.compag.2018.08.046.

[6] S. Bhandari, A. Raheja, D. Do, and F. Pham, "Machine learning techniques for the assessment of citrus plant health using UAV-based digital images," no. May, p. 21, 2018, doi: 10.1117/12.2303989.

[7] L. Costa, L. Nunes, and Y. Ampatzidis, "A new visible band index (vNDVI) for estimating NDVI values on RGB images utilizing genetic algorithms," Comput. Electron. Agric., vol. 172, no. March, p. 105334, 2020, doi: 10.1016/j.compag.2020.105334.

[8] S. Sankaran, J. M. Maja, S. Buchanon, and R. Ehsani, "Huanglongbing (Citrus Greening) detection using visible, near infrared and thermal imaging techniques," Sensors (Switzerland), vol. 13, no. 2, pp. 2117–2130, 2013, doi: 10.3390/s130202117.

[9] L. Parviz, "Performance evaluation of remote sensing data with machine learning technique to determine soil color," Polish J. Soil Sci., vol. 53, no. 1, pp. 97–116, 2020, doi: 10.17951/pjss/2020.53.1.97.

[10] L. Wang, Y. Duan, L. Zhang, T. U. Rehman, D. Ma, and J. Jin, "Precise estimation of NDVI with a simple NIR sensitive RGB camera and machine learning methods for corn plants," Sensors (Switzerland), vol. 20, no. 11, pp. 1–15, 2020, doi: 10.3390/s20113208.

[11] S. Sankaran and R. Ehsani, "Visible-near infrared spectroscopy based citrus greening detection: Evaluation of spectral feature extraction techniques," Crop Prot., vol. 30, no. 11, pp. 1508–1513, 2011, doi: 10.1016/j.cropro.2011.07.005.

[12] L. Wang, Q. Chang, J. Yang, X. Zhang, and F. Li, "Estimation of paddy rice leaf area index using machine learning methods based on hyperspectral data from multi-year experiments," PLoS One, vol. 13, no. 12, pp. 1–16, 2018, doi: 10.1371/journal.pone.0207624.

[13] M. Agarwal, S. K. Gupta, and K. K. Biswas, "Development of Efficient CNN model for Tomato crop disease identification," Sustain. Comput. Informatics Syst., vol. 28, p. 100407, 2020, doi: 10.1016/j.suscom.2020.100407.

[14] "NDVI Plant Health and Your Farm | Sentera," 2020. https://sentera.com/understanding-ndvi-plant-health/

[15] S. Mallick, “Image Alignment (ECC) in OpenCV (C++ / Python) | Learn OpenCV,” 2015. https://learnopencv.com/imagealignment-ecc-in-opencv-c-python/

[16] G. D. Evangelidis and E. Z. Psarakis, "Parametric image alignment using enhanced correlation coefficient maximization," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 10, pp. 1858–1865, 2008, doi: 10.1109/TPAMI.2008.113.

[17] "Machine Learning: What it is and why it matters | SAS," 2018. https://www.sas.com/en_my/insights/analytics/machinelearning.html

[18] L. Bermudez, "Overview of Neural Networks. What is a neural network? | by Luis Bermudez | machinevision | Medium," Nov. 15, 2017. https://medium.com/machinevision/overview-of-neural-networks-b86ce02ea3d1

[19] Sunil Ray, "Learn Naive Bayes Algorithm | Naive Bayes Classifier Examples," Sep. 11, 2017. https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/

[20] T. Rymarczyk, E. Kozłowski, G. Kłosowski, and K. Niderla, “Logistic regression for machine learning in process tomography,” *Sensors (Switzerland)*, vol. 19, no. 15, pp. 1–19, 2019.

[21] "Logistic Regression For Machine Learning and Classification - Kambria," KAMBRIA, Jul. 09, 2019. https://kambria.io/blog/logistic-regression-for-machine-learning/