ISSN 2278-3091
Volume 8, No.6, November – December 2019
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse136862019.pdf
https://doi.org/10.30534/ijatcse/2019/136862019

Cassava leaf NDVI - Artificial Neural Network (CaNDVI-ANN): A Low Cost, Portable and Non-Invasive Cassava Plant Health Monitoring Device

Edward B. Panganiban, Irma T. Plata, Bryan B. Bartolome, Allan C. Taracatac, and Freddie Rick E. Labuanan
College of Computing Studies, Information and Communication Technology, Isabela State University
Echague, Isabela, Philippines, demeters@isu.edu.ph

ABSTRACT

Numerous types of plant diseases could significantly impact the gross yield of food and agricultural assets wherein it could possibly result in a huge number of losses in terms of quantity and quality as well as in the economic aspect. Known practices to minimize loss because of plant infections includes identifying initial symptoms of infections through either based on visual examination or through laboratory investigations. However, these practices require a large workforce and take too much time. In order to solve the deficiencies of these usual and traditional approaches, this study presents a low cost, a portable and, noninvasive plant health monitoring device using an IR camera and an artificial neural network for image classification. The initial prototype was tested on a cassava plant to assess its photosynthetic activity which greatly indicates the current health status of the plant. The test result shows that almost all of the positive test samples were correctly identified, meaning it has identified low false-negative results, which obtained an accuracy of 89 %.

Key words*:* Artificial Neural Network (ANN), cassava, disease, Normalized Difference Vegetation Index (NDVI)

1. INTRODUCTION

The overall production of agricultural food products can be greatly affected by plant diseases, which can be the main reason for losses with regards to quantity, quality, and economic aspects. Laboratory tests and visual inspections are the early disease detection schemes that are commonly used to lessen losses from plant diseases. These are labor-intensive and time-consuming, however.

Many researchers have conducted non-invasive methods in order to resolve the restrictions of these conventional methods as reported in the recent findings of [1] and [2]. In addition, based on the work of [3], the physiological parameters of plants could be measured in a non-destructive manner and without direct contact with the plant using imaging techniques.

The physiological condition of the diseased tissue, like changes in transpiration, photosynthesis, accumulation of salicylic acid (SA), stomatal conductance, and even cell death, is altered during plant-pathogen infection. This gives an idea of how important early detection of plant disease is to plant physiology and its devastating effect [4].

According to studies [5] and [6], the unseen radiation patterns in the plant leaves are converted into visible images through the process of thermal remote sensing. These images are called thermal images or thermograms. These thermal images can be generated through the use of handheld, convenient thermal sensors, attached with optical systems affixed on an airplane or satellite. This technological equipment is non-contact and non-invasive. The potential use of remote sensing in agriculture includes irrigation scheduling, monitoring nursery, and greenhouse, estimation of fruit yield, detection of plant disease, evaluation of fruit maturity, and detection of fruit and vegetable bruise.

Using the same principle used in thermal imaging, a method called multispectral imaging shows promising results by analyzing near-infrared reflection, which by far has a successful experiment carried out [7] that could recognize stress in farmable crops under field environments and at an initial stage of disease growth before it can be visibly detected. This was achieved by analyzing information on spectral reflection between 450 and 900 nanometers from the plant. Multi-spectral imagery does not require an invasion of the device in the plants. The plants are scanned for the data collection of high-resolution images. Significant wavelengths combined together, as mentioned by [8], may also indicate the state of health or disease occurring within a species.

This paper presents an approach of noninvasive, portable and, low-cost multispectral imaging and artificial neural network (ANN) in stress detection to cassava leaf under field condition, which may indicate early symptoms of certain disease infection that could lead to production lost. This noninvasive cassava leaf analysis device could significantly benefit local cassava farmers in the country to conduct an intervention in terms of risk production management the soonest as possible before any disease may spread out.

2. MATERIALS AND METHODS

2.1 Development Method

To accomplish the objective of this study, the evolutionary prototyping method was employed as shown in figure 1. Evolutionary prototyping involves building a prototype and client feedback. Shortly after gathering preliminary responses from the end-users, initial prototypes are produced. Each prototype will be produced with additional features or improvements until the final device is developed. Figure 1 illustrates briefly the process that was involved in the development of the device in this paper.

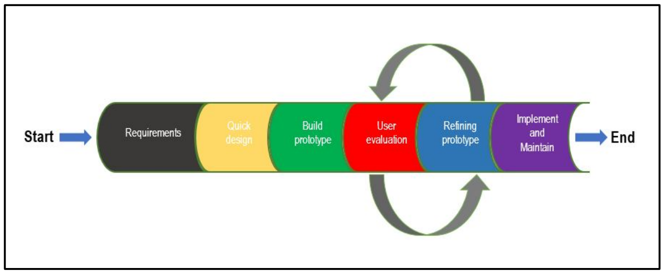

Figure 1: The evolutionary prototyping method

In accordance with the evolutionary prototyping method, the following hardware and software requirement(s) are identified:

2.2 Hardware

The hardware requirements include a laptop computer with at least core i7 7th generation, 8-Gigabytes of RAM, 320-Gigabytes of Hard disk drive - this computer will be used to train our cassava leaf NDVI - artificial neural network (CaNDVI-ANN) for image classification; Raspberry pi 3 model B – this is where the real-time cassava leaf NDVI – artificial network will be installed; Raspberry pi NoIR camera – for capturing reflectance spectra from cassava leaf; Red filter – the filter allows us to take an infrared photo in the "red" channel. This will be used to measure photosynthetic activity; 10,000mAH Power bank with micro USB cable – for a portable power source of the Raspberry device.

2.3 Software

The software listing used during the development of the system is Ubuntu 18.04 64bit, Python 3.5, TensorFlow 1.12, image annotating application and, Raspbian OS (Raspbian stretch or higher).

2.4 Methods

Near-infrared are invisible light that is beyond what a human eye could see as shown in Figure 2. This color wavelength holds ample of information that describes the current state or composition of the target object wherein our case is detecting pre-symptomatic characteristics of the photosynthetic disorder before manifesting visual symptoms on cassava leaf.

According to [9], the Normalized Difference Vegetation Index, or NDVI, presented within the early seventies and remains nowadays a really well-known instrument within the remote sensing community dealing with the agricultural observatory.

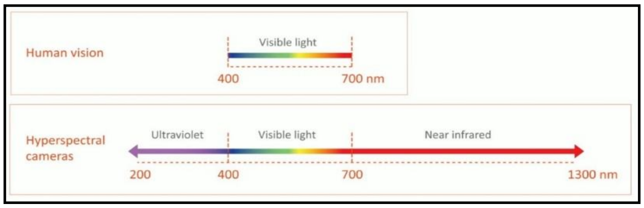

Figure 2: Color wavelengths chart

Infrared reflectance under IR-camera with an installed red filter *(refer to fig. 3)* is used to analyze photosynthetic activity on plant leaf as presented in figure 4. The NIR and the red color signifies the measured spectral reflectance obtained in both near-infrared regions and the red (visible) colored area. Each spectral band has an individual reflected ratios of the incoming ration from the spectral reflection, thus assuming values from 0.0 to 1.0. Thus, the NDVI itself varies from -1.0 to + 1 by design. The higher the photosynthetic activity, means the healthier the plant as demonstrated in figure 5.

.

Figure 3: Sample image was taken using an infrared camera with an installed red filter

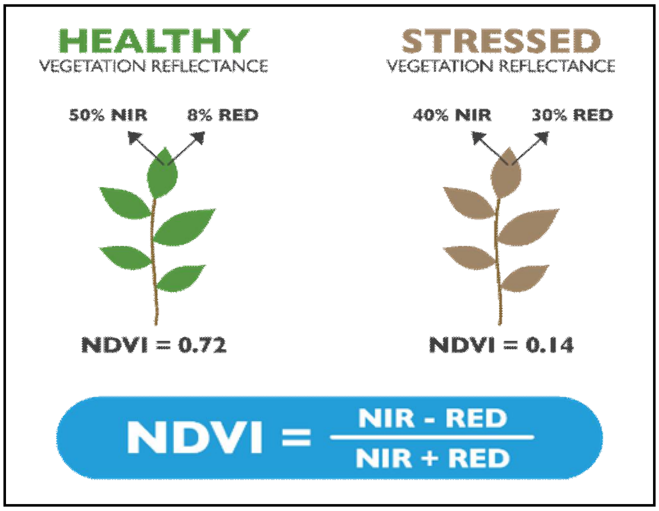

Figure 4: Sample computed NDVI value (https://www.agricolus.com/en/indici-vegetazione-ndvi-ndmi-istruzioni-luso/)

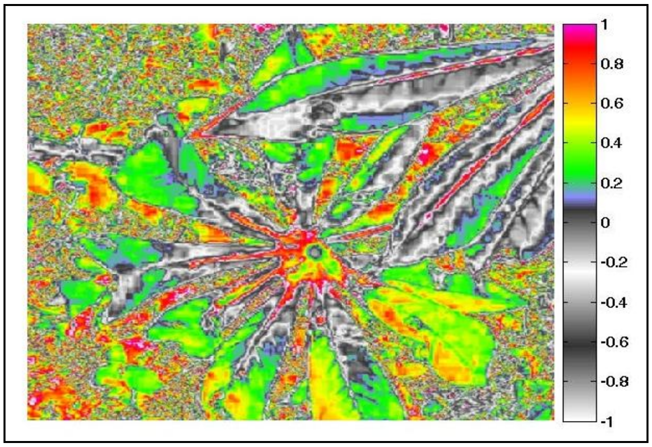

Figure 5 : Color map based on computed NDVI value
(a representation of photosynthetic activity on cassava leaf)

2.5 Training and deployment of CaNDVI-ANN

As presented in Figure 6, to train the custom cassava leaf classifier, it starts with collecting thousands of cassava leaf images that are identified as stressed *(low computed NDVI value).* The way sample images are collected was via recorded video of stress cassava leaf converted to images instead of gathering through still images. A total of 5,000 sample images were extracted and expanded to 15,000 through dataset augmentation *(flip, zoom in, and zoom out).* After acquiring enough number of images for the training dataset, image annotation comes next. Image annotation is done manually with LabelImg from https://github.com/tzutalin/labelImg, an application to map the object of interest from a cluttered image.

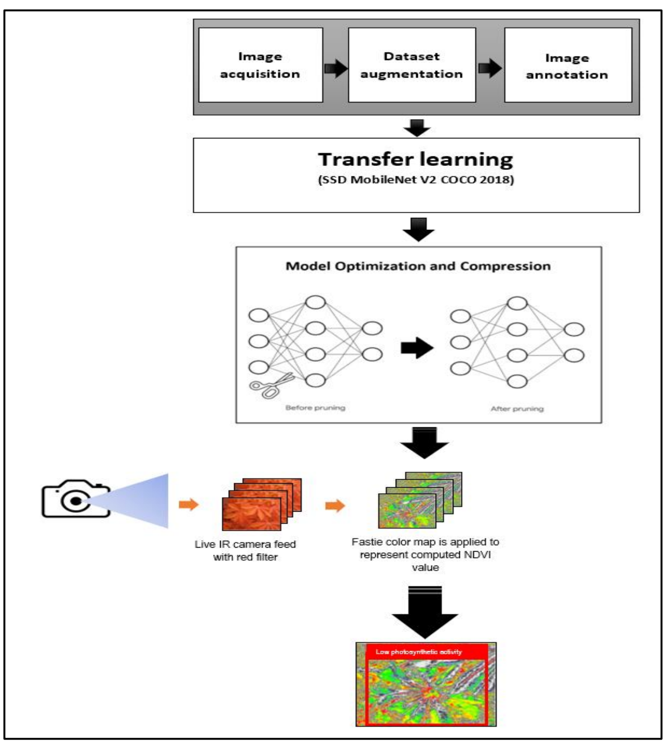

Figure 6: Cassava leaf NDVI - artificial neural network (CaNDVI-ANN) concept diagram

For the configuration of the training and evaluation procedures, protobuf files were used for the Tensorflow Object Detection API. The final layer of the SSD Mobilenet V2 COCO 2018 model which was pre-trained internally at Google will be replaced with our own custom class. The hyperparameters for the training are configured as follows; the training step is 100,000 with a training batch size of 200, 0.01 learning rate and, a validation batch size of 200. For CNN, a large number of training datasets are the basis for its accuracy [10]. Implementation of an artificial neural network with the right hardware specification is capable of detecting objects within 0.2 seconds, which makes the proposed system robust and effective[11].

The output classifier from the transfer learning process was optimized using pruning technique for disk compression: sparse tensors are amenable to compression. Thus, by applying simple file compression to the pruned TensorFlow [12] checkpoint, the size of the model is reduced for its storage and/or transmission. This model optimization technique is very handy for well-balanced speed and classification accuracy when deploying CNN in low powered single board computer like the Model B Raspberry Pi 3.

To deploy the optimized custom trained classifier, the IR camera feed with installed red filter must be preprocessed by applying faster colormap to interpret the computed NDVI value on each frame. Detected cassava leaves with low photosynthetic activity based on computed NDVI value will have label overlay in real-time, indicating abnormalities in the leaves.

2.6 Evaluation of classification accuracy

The actual evaluation of accuracy was calculated using a simple mathematical formula adapted from the work of [13] as shown in the equation below. Equation (1) is the confusion matrix equation [14].

$$Accuracy = \frac{(TN+TP)}{(TN+TP+FN+FP)} = \frac{(\text{Number of correct assessment})}{(\text{Number of all assessments})} \quad (1)$$

where: TP – True Positive, TN – True Negative, FN – False Negative, FP – False Positive

There are, however, accuracy issues. For both types of errors, it assumes equal costs. Depending on the problem, a 99% accuracy can indicate excellent, good, mediocre, poor or terrible.

2.7 Recall

The recall is defined as the ratio of the total number of correctly classified positive examples divided with the total number of true positive examples. High Recall indicates the class is correctly recognized. It means that it has a small number of false negatives classified. Equation (2) is the recall equation.

The recall is given by the equation:

$$\textbf{\textit{Recall}} = \frac{TP}{TP + FN} \quad (2)$$

2.8 Precision

To get the precision value, the total number of correctly identified positive examples is divided by the total number of predicted positive examples. High precision implies that an example labeled as positive is certainly positive (a small number of FP). Equation (3) is the precision equation.

Precision can be determined by the relation:

$$\textit{Precision} = \frac{TP}{TP + FP} \quad (3)$$

High recall, low precision: This implies that most of the positive examples are appropriately identified (low FN), but also indicates a lot of false positives.

Low recall, high precision: This expression indicates that there is a lot of missed positive examples (high FN), but the expected positive predictions are true positive (low FP).

2.9 F-measure

Since precision and recall helps to have a measurement that denotes both of them. We calculate F-measure which then uses Harmonic Mean instead of Arithmetic Mean as it reproves the extreme values more.

The F-Measure will always be nearer to the smaller value of Precision or Recall. Equation (4) is the F-Measure equation.

$$\textit{F - measure} = \frac{2*Recall*Precision}{Recall + Precision} \quad (3)$$

3. RESULTS AND DISCUSSION

To test the classifier’s accuracy, a set of images were prepared, these consist of 50 images of a healthy cassava leaf and 50 images of disease-infected cassava leaf. In total there are 100 test images. The accuracy test was carried out using a simple python script that will loop into the test images. The test result is shown by printing the image file name including the image analysis result if it was able to detect low photosynthetic activity in the given image. Printed results in all given images are shown in the terminal window.

n = 100	Predicted No	Predicted Yes	
Actual No	TN = 50	FP = 0	50
Actual Yes	FN = 11	TP = 39	50
	61	39	

Figure 7: Confusion matrix with TP, FP, TN, FN

Based on the collected values on the classifier’s accuracy test, the computed accuracy was 0.89, with a computed recall of 0.78, precision of 0.0, and F measure of 0.0.

4. CONCLUSION

The outcome of the accuracy test indicates that most of the positive image samples are determined or recognized. It means that the result identified a low false negative (LFN). However, there are also a lot of false positives. The accuracy of the cassava leaf classifier used in this study could also be improved by collecting an additional good dataset or by upgrading the IR sensitivity of the camera to assess the wider range of the light spectrum.

ACKNOWLEDGMENT

Gratitude and acknowledgment go to the Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (PCAARRD), one of the sectoral councils under the Department of Science and Technology (DOST), for approving the project and for providing funding support and assistance. Special thanks also are dedicated to the Isabela State University for providing resources and technical assistance as well.

REFERENCES

[1] A. Y. Khaled, S. Abd Aziz, S. K. Bejo, N. M. Nawi, I. A. Seman, and D. I. Onwude, **“Early detection of diseases in plant tissue using spectroscopy–applications and limitations,”** *Appl. Spectrosc. Rev.*, vol. 53, no. 1, pp. 36–64, 2018. https://doi.org/10.1080/05704928.2017.1352510

[2] S. Han and F. Cointault, **“Early detection of the disease on leaves by image processing,”** no. June 2013, pp. 24–27.

[3] H. Xu, S. Zhu, Y. Ying, and H. Jiang, **“Early detection of plant disease using infrared thermal imaging,”** vol. 6381, no. 2006, pp. 1–7. https://doi.org/10.1117/12.685534

[4] E. E. Omran, **“Early sensing of peanut leaf spot using spectroscopy and thermal imaging,”** *Arch. Agron. Soil Sci.*, vol. 00, no. 00, pp. 1–14, 2016.

[5] R. Ishimwe, K. Abutaleb, and F. Ahmed, **“Applications of thermal imaging in agriculture — A review,”** no. September, pp. 128–140, 2014. https://doi.org/10.4236/ars.2014.33011

[6] L. Molinero-ruiz, “**Use of blue-green fluorescence and thermal imaging in the early detection of sunflower infection by the root parasitic weed Orobanche,”** vol. 8, no. May, pp. 1–10, 2017. https://doi.org/10.3389/fpls.2017.00833

[7] D. Moshou *et al.*, **“Plant disease detection based on data fusion of hyperspectral and multi-spectral fluorescence imaging using Kohonen maps,”** *Real-Time Imaging*, vol. 11, no. 2, pp. 75–83, 2005. https://doi.org/10.1016/j.rti.2005.03.003

[8] A. Lowe, N. Harrison, and A. P. French, **“Hyperspectral image analysis techniques for the**

detection and classification of the early onset of plant disease and stress," *Plant Methods*, pp. 1–12, 2017.

[9] G. Rabatel, N. Gorretta, and S. Labbé, "**Getting NDVI spectral bands from a single standard RGB digital camera: A methodological approach,"** *14th Conf. Spanish Assoc. Artif. Intell.*, vol. 7023 LNAI, no. 1, pp. 333–342, 2011. https://doi.org/10.1007/978-3-642-25274-7_34

[10] Y. Chunjing, Z. Yueyao, Z. Yaxuan, and H. Liu, **"Application of convolutional neural network in classification of high resolution agricultural remote sensing images,"** *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 42, no. 2W7, pp. 989–992, 2017.

[11] K. Dimililer and E. Kiani, "**Application of back propagation neural networks on maize plant detection,"** *Procedia Comput. Sci.*, vol. 120, pp. 376–381, 2017. https://doi.org/10.1016/j.procs.2017.11.253

[12] E. B. Panganiban, A. C. Paglinawan, W. Y. Chung, and G. L. S. Paa, "**A Novel Technique in Classifying Heart Diseases based on Electrocardiogram (ECG) Signals using Deep Learning and Spectrogram Image Analysis**," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 4, pp. 1734–1740, 2019.

[13] N. W. Wen Zhu, Nancy Zeng, W. Zhu, N. Zeng, and N. Wang, "**Sensitivity, specificity, accuracy, associated confidence Interval and ROC analysis with practical SAS® implementations,"** *Rev. Asoc. Med. Argent.*, vol. 60, pp. 239–241, 2010.

[14] I. T. Plata, A. C. Taracatac, and E. B. Panganiban, "**Development and testing of embedded system for smart detection and recognition of witches' broom disease on cassava plants using enhanced viola-jones and template matching algorithm,**" *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 5, pp. 2613–2621, 2019. https://doi.org/10.30534/ijatcse/2019/113852019