Detection and Identification with Analysis of *Carica papaya* Leaf Using Android

John A. Bacus[1,2] and Noel B. Linsangan[1]
[1] School of Electrical, Electronics and Computer Engineering, Mapua University, Manila, Philippines
[2] College of Engineering Education, Computer Engineering Program, University of Mindanao, Davao City, Philippines
Email: johnbacus001@gmail.com, {john_bacus, nblinsangan}@umindanao.edu.ph

***Abstract*—With the increase in the usage of mobile devices such as smartphones, laptops, smartwatches, etc., access to information and communication has been effortless and convenient. Thus, making Raspberry Pi, an Android device, has been made. LineageOS is used specifically as an operating system that Konstakang developed. With CNN's MobileNet architecture and transfer learning, the classification for papaya leaf disease was a success. MobileNet Architecture was retrained using the images of the following papaya leaves such as Blackspot, Brownspot, Mealybug Infection, Powdery Mildew, Healthy, and Unknown images and employed transfer learning to create the model successfully. A total of seventy-two (72) samples were tested. The study made use of confusion matrix to compute for the accuracy of the system and got 91.667% accuracy.**

***Index Terms*—convolutional neural networks, transfer learning, MobileNet, plant disease identification, TensorFlow, lineage OS**

I. INTRODUCTION

The papaya (Carica papaya) was originated from the tropics in America. It is considered one of the most important fruit crops in the Philippines because of its enormous economic potential. Papaya is a good source of Vitamins A and C, iron, calcium, protein, carbohydrate, and phosphorous [1]. From the data of the Philippine Statistics Authority, in 2017, the Philippines had a total of 167 043 metric tons for the volume of production. SOCCSKSARGEN is known as the top producing region.

CNN has been widely used as an algorithm for image classification, including research [2]. Several researchers, such as [3], have recently studied plant disease identification based on deep learning approaches yielding a high accuracy result. A study of [4] used the kNN algorithm for image classification, which generally gives an idea about image processing. As CNN rises, different tools in different fields have yielded a high accuracy result, including research [5] and [6]. Research [7] proposed a novel identification approach for rice diseases based on deep convolutional neural networks. Recent research about feature extraction, such as [8] and [9], used support vector machines for classification. Using a dataset of 500 natural images of diseased and healthy rice leaves and stems, CNNs were trained to identify ten common rice diseases. Authors of [10] developed a Plant Identification through leaf veins utilizing CNN and Caffe as its framework. In a study by [10], CNN MobileNet architecture was used, yielding 88% and 90% overall accuracy. Unlike traditional methods, CNN can learn high-level robust features directly from the original image instead of extracting the specific features manually. It has been widely applied to various image classifications and achieved impressive results [11]. Mobile applications have been identified as the best platform for the expert system tool to reach as many users as possible. A study that has been conducted about the evaluation of banana-based on its ripeness used Ionic Framework as its tool for mobile application development. Ionic was built on Angular and Apache Cordova to support iOS, Android, and Web platforms [12].

Manual and laboratory testing are being exercised for plant leaf disease detection. Plant monitoring performed by an expert agriculturist through naked eye observation is important and necessary to control the spread of plant diseases [13], [14]. Disadvantages in identifying plant leaf disease using these features include the practice of different crop and pest terminologies that only experts can comprehend. Once a papaya tree is being infected at a very early stage, it will not produce crops resulting in the farmer's loss in profit. Hence, this study is proposed to improve plant leaf disease identification using Raspberry Pi with LineageOS for portability.

The general objective of this research study is to create a portable device that detects and identifies papaya leaf disease. To achieve this, specific goals are laid: (1) Develop a portable device using Raspberry Pi with LineageOS that can detect and identify papaya leaf diseases. (2) Use convolutional neural network to identify and detect papaya leaf diseases; (3) Test the accuracy of the classification using confusion matrix analysis. (4) Recommend treatments or control of the detected papaya leaf disease.

The study would significantly affect the papaya farmers. By using this app, the farmers in particular will find convenience, and they will quickly detect the disease their papaya leaf has in their papaya plantation. This will help them prevent rotting their papaya plants because an analysis of a particular papaya disease will be given, such as a recommendation or a prescription for the treatment. The disease of a papaya plant is the most productive

Manuscript received October 22, 2021; revised February 11, 2022.

doi: 10.12720/jait.13.2.162-166

factor affecting the plantation. Furthermore, this research study can benefit the global aspect, which can help future researchers to improve the identification and detection of papaya leaf diseases.

This study focuses on the development of an application classifying papaya leaf disease using Raspberry Pi as a prototype Android device. The study will mainly focus on four leaf diseases: mealybug infection, brown spot disease, black spot disease, and powdery mildew. Alongside the four diseases mentioned, it can also detect if a papaya leaf is healthy and if the object scanned/captured is unknown. The application will be used to capture the image of the papaya leaf and identify what kind of disease it has using CNN MobileNet to classify.

II. Materials and Methods

A. System Architecture

The study will work on the Raspberry Pi that has an operating system named LineageOS. TensorFlow library is used to program and build the CNN Network. The model is being implemented using Python since it offers machine learning libraries, especially Neural Networks. Keras library was used for augmenting the datasets to boost the number of training data.

Fig. 1 shows the conceptual framework of the study that represents the overall functionality of the system. Papaya leaf images are the input of this system. After the image acquisition, leaves will be automatically augmented using Keras. MobileNets, a pre-trained model with Tensorflow, will process and analyze the input. The classification of the papaya leaf falls into six categories, namely, Brownspot, blackspot, Powdery Mildew, Mealybug infection, Healthy, and Unknown. The pre-trained model MobileNets with TensorFlow will learn and process everything.

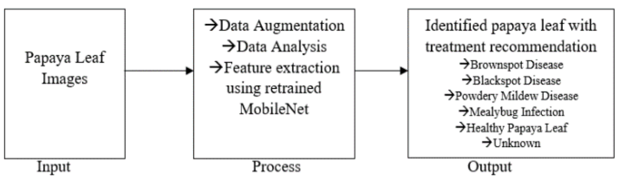

Figure 1. Conceptual framework.

The papaya leaf images are acquired from the field using the Raspberry Pi's camera module, shown in Fig. 2, or an image selected using the device's gallery will serve as the input. The Raspberry Pi with Lineage OS will now serve as the central processing device in analyzing and displaying the results.

Figure 2. Image capturing setup.

The block diagram in Fig. 3 gives the overall system for image processing to determine the disease of a papaya leaf. The device consists of five main function blocks: Raspberry Pi 4, Raspberry Pi camera module, LCD Screen, a papaya leaf, and the power supply unit. The papaya leaf will be captured using the camera module. After completing the processing and analysis, the LCD Screen will show the disease of a papaya leaf. The power supply unit will be used to provide electricity to the primary device.

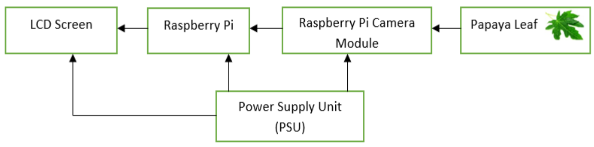

Figure 3. Block diagram of the system.

B. Methods and Procedures Collection of Leaf Images

Fig. 4 displays the training and testing images acquired from the Papaya Farms in the Davao Region together with the Department of Agriculture XI. Five health conditions were considered: 4 leaf diseases, such as Brown spot, Blackspot, Powdery Mildew, and Mealybug Infection, and one healthy leaf. The study will recommend the treatments of the four different leaf diseases, which is based on the Bureau of Plant and Industry's Information Section and Information from Diseases of fruit crops in Australia [15]. There are 611 images in total that have been randomly taken, including the unknown datasets. The images were classified and verified by experts. They were augmented, which caused to escalate the number of samples to 1394 images and was resized according to the input size of the neural network for Dataset preparation as per Table I.

TABLE I. Total No. of Augmented Data for Training Set

Papaya Leaf Disease	Training Set
Blackspot Disease	222
Brownspot Disease	218
Mealybug Infection	199
Powdery Mildew	204
Healthy Papaya Leaf	251
Unknown	300
Total	1394

Figure 4. Papaya leaf images samples. (a) Brownspot (b) Blackspot (c) Powdery Mildew (d) Mealybug (e) Healthy.

C. Image Analysis

After collecting image samples, training the model was possible in CNN MobileNet architecture shown in Fig. 5. The depthwise separable convolution reduces the number of parameters for the convolutional operations [16]. It also reduces the difficulty which builds the layers of the MobileNet. To mitigate the parameters, the MobileNet offers two parameters: the width multiplier and the resolution multiplier, which allows reducing its number of operations and parameters drastically. It is beneficial in mobile and devices with less computer power [16]. Width multiplier is used to thin the network whereas, the resolution multiplier changes the input dimension of the image [17].

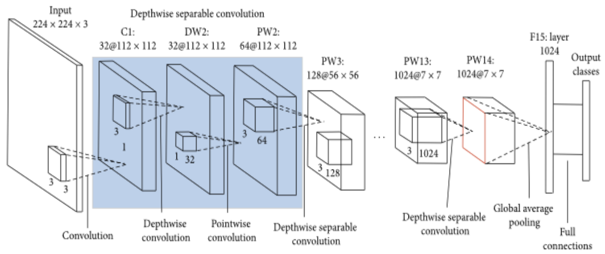

Figure 5. CNN mobilenet architecture.

D. Retraining Few Top Layers and Transfer Learning

In this process, the MobileNet network is pre-trained with a large-scale of general datasets, ImageNet, to function as a generic template for visual processing, as shown in Fig. 6. The pre-trained network transfers all the learned parameters and is set as a feature extractor to execute the target task. The final classification layer of the CNN MobileNet model is removed, then freezing the other layers and retraining the last layer with the new set of target images, and applying fine-tuning for the parameters. Then, after transfer learning, two output files are extracted, the protobuf and the label.txt, which are now embedded into the created Android application using Android Studio.

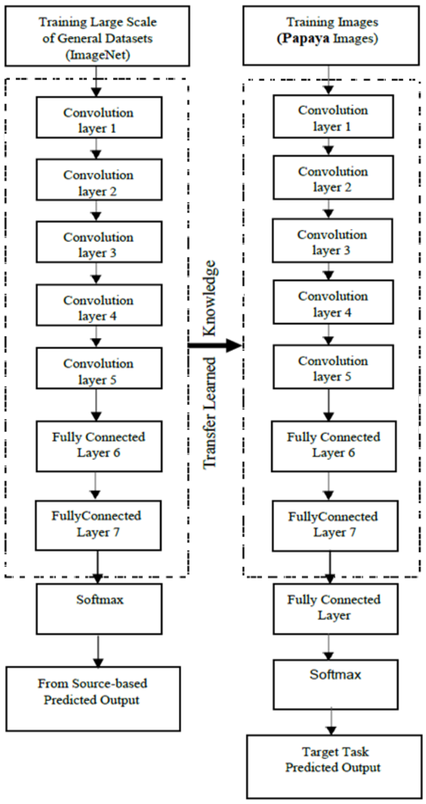

Figure 6. Transfer learning with retrained MobileNet.

E. Testing and Validation Phase

Captured leaf images for different classifications were being copied into the device to test the performance. Validation is used to authenticate the final network model specification. Confusion matrix was used to verify the overall accuracy of the system.

III. RESULTS AND DISCUSSIONS

Presented in this section is the discussion on the gathering, collection, and analysis of data and the findings and results of the study.

A. Data Gathering and Collection

It is essential to get higher accuracy and enrich the datasets. Therefore, images were augmented as studies [18] and [19] possessed. The image sample shown in Fig. 7 was augmented using Keras. Fig. 8 is an image sample of Blackspot disease that was augmented into twelve (12) images.

Figure 7. Blackspot disease.

Figure 8. Augmented blackspot disease.

Seventy-two (72) images were used for testing as shown in Fig. 9. 13 for Blackspot, 12 for Brownspot, 10 for Powdery Mildew, 8 for Mealybug Infection, 15 for Healthy Papaya Leaf, and 14 Unknown Images. These papaya leaves are taken from the papaya plantation in Davao Region, specifically in Toril, Davao City, Philippines. Additionally, the images were verified by experts in the farm together with the Department of Agriculture XI.

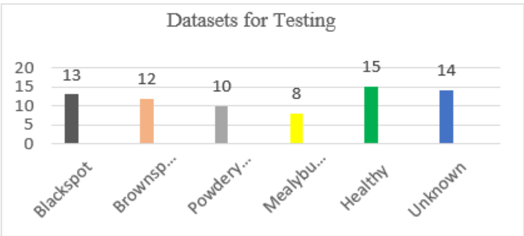

Figure 9. No. of images used for testing.

B. Actual Device Prototype

Fig. 10 shows the actual device prototype. The Raspberry Pi 4, together with an LCD screen and the camera module, was attached. The Power Supply supplies electricity to the device, enabling it to function.

Figure 10. Actual device prototype.

C. Data Analysis Result

The papaya leaf disease detector application was installed in the device. After the app was opened, the testing of the leaf diseases took place. For the papaya leaf diseases to be detected, two buttons can be tapped: the "Capture Photo" and "Select Photo." The "Capture Photo" button, as shown on the left side of Fig. 11, will redirect to the device's camera and captures the photo of a testing sample. Then, the algorithm displays the classification of the image taken. Whereas the "Select Photo" button shown on the right side of Fig. 11 will redirect to the device's storage drive, and the user can select from one of the images already captured and saved from the storage.

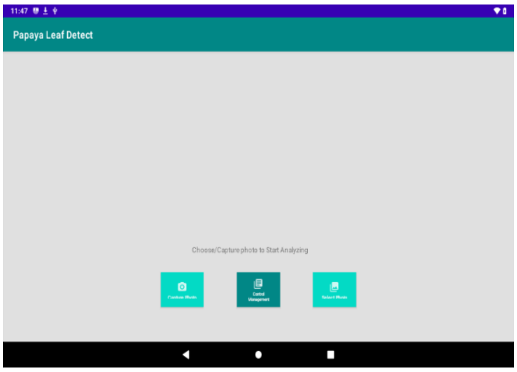

Figure 11. Graphical user interface.

The sample classification of Brownspot in the device is shown in Fig. 12.

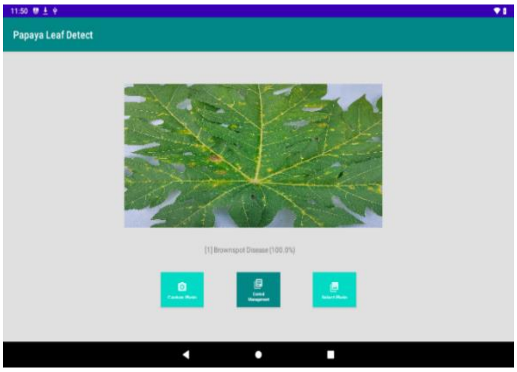

Figure 12. Detected brownspot disease.

Since the study also recommends treating each disease, Fig. 13 below shows the treatment recommendation for Blackspot, Brownspot, and Mealybug infection. The user can tap the "Control Management" button to see this tab. The user can hover to the left, right, up, and down to see the whole recommendation treatment.

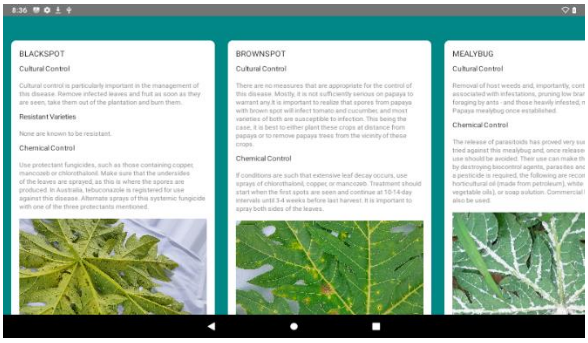

Figure 13. Treatment recommendation for the leaf diseases.

The test results are then plotted in a confusion matrix shown in Table II to confirm the accuracy of the data. Each row in the confusion matrix represents the actual results, whereas the columns represent the predicted results.

TABLE II. CONFUSION MATRIX

		Predicted Result					
		Bls	Brs	PwM	MbI	HpL	Un
Actual Result	Bls	**12**	1	0	0	0	0
	Brs	1	**9**	0	0	0	0
	PwM	0	0	**10**	0	0	0
	MbI	0	0	0	**6**	0	0
	HpL	0	2	0	0	**15**	0
	Un	0	0	0	2	0	**14**
TRUTH OVERALL		13	12	10	8	15	14
ACCURACY		92.3%	75%	100%	75%	100%	100%
TOTAL NO. OF SAMPLES		72					
OVERALL ACCURACY		**91.667%**					

Legend:
Bls = Blackspot
Brs = Brownspot
MbI = Mealy bug infection
OA = Overall Accuracy
HpL = Healthy papaya leaf
PwM = Powdery Mildew
Un = Unknown

A total of seventy-two (72) samples were tested. Blackspot disease has shown 92.3% accuracy, Brownspot has 75%, Powdery Mildew has 100%, Mealybug Infection has 75%, Healthy Papaya Leaf has 100%, including the Unknown.

The overall accuracy of the system is 91.667% calculated by using Equation (1).

$$Overall\ Accuracy = \frac{(Bls,Bls)+(Brs,Brs)+(PwM,PwM,Mbl,Mbl+HpL,HpL+Un,Un)}{TOTAL\ NO.OF\ SAMPLES} \quad (1)$$

IV. CONCLUSION AND FUTURE WORKS

The portable device Raspberry Pi with LineageOS was successfully made, which can classify papaya leaf diseases. With the use of CNN MobileNet architecture, transfer learning and retraining the network was made possible. The device was able to classify papaya leaf diseases such as Blackspot, Brownspot, Powdery Mildew, Mealybug Infection, healthy papaya leaf, and unknown. The classifier detected 66 out of 72 samples that give the system's overall accuracy to 91.667%.

For future works, proper selection of datasets during training must also be considered for augmentation to avoid image distortion and for better network learning. It is also highly recommended to use more training data to give a higher accuracy result and correct classification. Also, the latest CNN architecture that is suitable for less powerful devices is recommended to achieve a high-performance result. More of the papaya leaf diseases should also be added.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

John Bacus gathered the dataset, conducted the research, and wrote the paper. Noel Linsangan supervised all the work and approved the final revision.

References

[1] The Philippine Recommends for Papaya Committee 2005, *The Philippine Recommends for Papaya*, Los Banos, Laguna: PCARRD-DOST, 2006, p. 91, (Philippines Recommends Series no. 27-B).

[2] S. Bunrit, N. Kerdprasop, and K. Kerdprasop, "Improving the representation of CNN based features by autoencoder for a task of construction material image classification," *Journal of Advances in Information Technology*, vol. 11, no. 4, pp. 192-199, 2020.

[3] R. E. Angelia and N. B. Linsangan, "Fermentation level classification of cross cut cacao beans using k-NN algorithm," in *Proc. 5th International Conference on Bioinformatics Research and Applications*, 2018.

[4] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378-384, 2017.

[5] S. Bhadoria, P. Aggarwal, C. G. Dethe, and R. Vig, "Comparison of segmentation tools for multiple modalities in medical imaging," *Journal of Advances in Information Technology*, vol. 3, no. 4, pp. 197-205, November 2012.

[6] B. Xie, X. Zhu, C. Han, Y. Wang, X. Li, and Y. Zhang, "Research of the space-borne infrared ship target recognition technology based on the complex background," *Journal of Advances in Information Technology*, vol. 10, no. 2, pp. 48-53, May 2019.

[7] N. B. Linsangan and R. S. Pangantihon, "FPGA-Based plant identification through leaf veins," in *Proc. 5th International Conference on Biomedical and Bioinformatics Engineering*, 2018.

[8] S. Lahmiri, "Features extraction from high frequency domain for retina digital images classification," *Journal of Advances in Information Technology*, vol. 4, no. 4, pp. 194-198, November 2013.

[9] J. A. C. Moreano and N. B. L. S. Palomino, "Global facial recognition using gabor wavelet, support vector machines and 3D face models," *Journal of Advances in Information Technology*, vol. 11, no. 3, pp. 143-148, August 2020.

[10] D. Padilla, A. Yumang, A. L. Diaz, and G. Inlong, "Differentiating atopic dermatitis and psoriasis chronic plaque using convolutional neural network MobileNet architecture," in *Proc. IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management*, 2019.

[11] A. A. M. Al-Saffar, H. Tao, and M. A. Talab, "Review of deep convolution neural network in image classification," in *Proc. International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications*, 2017.

[12] K. Kipli, *et al.*, "Image processing mobile application for banana ripeness evaluation," in *Proc. International Conference on Computational Approach in Smart Systems Design and Applications*, 2018.

[13] N. Petrellis, "Plant disease diagnosis based on image processing, appropriate for mobile phone implementation," *CEUR Workshop Proc.*, vol. 1498, pp. 238-246, 2015.

[14] V. Singh and A. K. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," *Inf. Process. Agric.*, vol. 4, no. 1, pp. 41-49, 2017.

[15] T. Cooke, D. Persley, and S. House, *Information from Diseases of Fruit Crops in Australia*, CSIRO Publishing, 2009.

[16] M. J. Macawile, V. V. Quinones, A. Ballado, J. D. Cruz, and M. V. Caya, "White blood cell classification and counting using convolutional neural network," in *Proc. 3rd International Conference on Control and Robotics Engineering*, 2018.

[17] A. Howard, *et al.*, "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.

[18] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image," *Computational Intelligence and Neuroscience*, article ID 3289801, 2016.

[19] C. Chuang and C. Fan, "Deep-Learning based joint iris and sclera recognition with YOLO network for identity identification," *Journal of Advances in Information Technology*, vol. 12, no. 1, pp. 60-65, February 2021.

John A. Bacus was born in Kidapawan City, Philippines. He received his degree of Bachelor of Science in Computer Engineering in 2017 from the University of Mindanao, Philippines. He went on to earn his Master of Science in Computer Engineering in 2019 at Mapua University. His academic research interests include image processing, machine learning and neural networks. He is now currently working as a faculty member of the Computer Engineering department at the University of Mindanao, Philippines.

Noel B. Linsangan, at present, is the Program Head of the Computer Engineering (CpE) program of the Mapua University in Manila, Philippines. Mapua University is the largest engineering school in the Philippines. He obtained his baccalaureate degree in Computer Engineering from the Mapua Institute of Technology in 1988 and his Master in Engineering in Computer Engineering from the University of the City of Manila in 2000. Right after graduating from college, he joined the faculty of the Computer Engineering program and in 2000 he was appointed as its Program Head. Majority of his research articles fall under the category of Computer Vision. Other research articles include topics in machine learning, wireless communication applications, embedded system applications, and ICT applications.