Leaf Disease Detection using Transfer Learning with Logistic Regression

Arlea Bless L. Sabarre, Alexandra Nicole S. Navidad,
Darwin S. Torbela, and Jetron J. Adtoon*

Computer Engineering Program, College of Engineering Education,
University of Mindanao, Philippines

**Corresponding author:* jetron.adtoon22@gmail.com

ABSTRACT

Many plant recognition systems are being cast off to address the serious threat of plant diseases and pathogens concerning the agricultural industry. Computer vision is a one-way method to advance the conventional way of dealing with plant disease detection. However, the result's efficiency using this method entirely depends on the leaf characteristics. Due to the popularity and successful implementation of deep learning, research directed from traditional feature-based methods to deep learning. However, training a new deep learning model from the beginning requires much data and is both costly and time-laborious. In this study, transfer learning, one efficient deep learning method, is applied to recognize and learn useful characteristics directly from the inputted data. Then, a logistic regression classifier for leaf disease classification was used. Four health conditions are addressed in this study, 25 per classification, with a total of 100 samples tested to evaluate the system's accuracy. The result showed 94% accuracy in detecting algal spots, Cercospora, leaf discoloration, and healthy leaves.

Keywords: *leaf, leaf disease, transfer learning, logistic regression, detection*

INTRODUCTION

Agriculture is one important sector in society's daily system. It provides food, opens livelihood opportunities, and supports the economy. However, threats are decreasing the growth of the agriculture sector today. This serious threat is caused by pests, plant diseases, or pathogens, which reduce crop productivity, leading to financial loss for farmers. Terrible occurrence of plant diseases worsens the current deficit of food supply, in which at least 800 million people are inadequately fed (Strange & Scott, 2005). Some crop diseases are visible in the early stage. Some crops would be visible only in the later stage, leaving the zero possibility of rescuing the crop (Surya Prabha & Satheesh Kumar, 2014). The stern effect of these diseases led to partaking a research approach to advance and promote management strategies (Daniel et al., 2014). Plant monitoring performed by an expert agriculturist through naked-eye observation is necessary to control the spread of plant diseases (Martinelli et al., 2015; Tongsri, Songkumarn, & Sangchote, 2016). This method is time-consuming, laborious, and costly, especially in dealing with large fields (Android, 2016).

Plant diseases are responsible for the significant economic losses experienced in the worldwide agricultural industry (Martinelli et al., 2015). With this, the call for advanced disease detection and crop prevention is imperative (Fang & Ramasamy, 2015). Using computer vision or machine learning techniques is the hope to solve this issue (Kamilaris & Prenafeta-Boldú, 2018; Lee, Chan, Wilkin, & Remagnino, 2015; Ramcharan et al., 2017). Researchers are already utilizing technological techniques to exceed the performance of the conventional method of plant disease detection (Cope, Remagnino, Barman, & Wilkin, 2010). However, despite the many efforts conducted, plant recognition with computer vision alone seems challenging and unsolved (Naresh & Nagendraswamy, 2016). This is because a plant in nature has very similar characteristics such as shape, color, texture, size, etc. The accuracy of the result is fully dependent on the ability of computer vision to encode morphological characteristics predefined by experts (Kamilaris & Prenafeta-Boldú, 2018; Larese et al., 2014). That is why deep learning is utilized as a new method for image processing and data analysis to show high potential and promising results. It has been successful in various fields and recently entered the field of agronomy and food products. The study is a modification of the study (Sabarre et al., 2021), where it only used transfer learning alone and attained a 90% accuracy.

Training a new deep learning model from the beginning requires sufficient data. It is very costly and time-laborious. Computing resources would take hours or days of training and may even be impossible. It is considered inefficient sometimes to apply the conventional deep learning method solely. That is why the proponents are on the grounds of applying another branch of deep learning method, transfer learning combined with logistic regression, to address the concern of making plant disease detection faster.

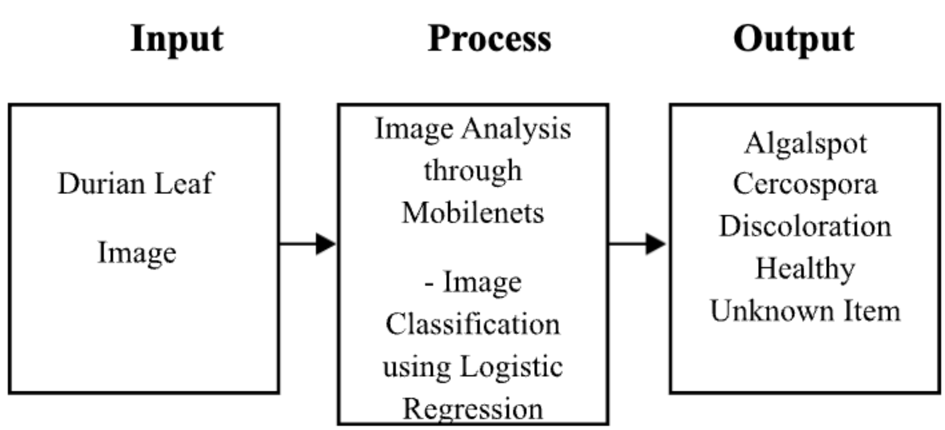
This study aimed to develop a mobile application for plant disease recognition. The significance of this study is to make plant disease detection fast, convenient, accessible, and efficient by adopting the branch of deep learning which is transfer learning, plus logistic regression. Consequently, this will escalate productivity, improve the country's agricultural sector, and bring about good quality change. The scope of the study comprised only using a specific plant or leaves for testing, the durian leaves. The inputted data or images were processed using the mobile application. The app only detected two common durian leaf diseases and durian leaf discoloration. The diseases are Cercospora Leaf Disease and Algalspot Leaf Disease. It would also show *healthy* as the detected result if none of these diseases occurred in the leaves and *unknown item* if the captured image could not be identified as a durian leaf. The app would also provide control strategies to prevent these diseases' manifestation and spread.

METHOD

Conceptual Framework of the Leaf Detection Project

Figure 1 presents the concept followed in developing the mobile application. The images were acquired from the field using an android phone. The pre-trained model, MobileNets with Tensorflow, will learn and process the details of the data. The result is then classified using Logistic Regression. The output will be displayed using the developed android mobile application.

Figure 1
Conceptual Framework of the Study



Research Methodology

Logistic regression is a statistical learning technique dedicated to classification. It has gained a tremendous reputation over the last two decades. It is primarily utilized as the generic classifier for this study. Tensorflow is an open-source software library or framework used for numerical computations and machine learning processes to retrain a special class of CNN model. MobileNets is the chosen Convolutional Neural Network model for this study. It is optimized to be executed using minimal possible computing power and enhanced latency on a smartphone device. It is considered more efficient because it uses small data to train the model and attain high-precision results with a short training time. The efficiency of this method was implemented for plant disease recognition. The mobile app was developed using Android Studio, the official Integrated Development Environment (IDE).

Figure 2 illustrates the MobileNets model utilized in this study. Mobilenets have two surface parameters: the width multiplier to thin the network and the resolution multiplier that changes the input dimensions of the image. Several studies also used MobileNet architecture to detect and identify leaf diseases and successfully provided high accuracy (Sabarre et al., 2021). Both are tuned to resolve the target task's resource and accuracy trade-off problem.

Figure 2
MobileNets Model

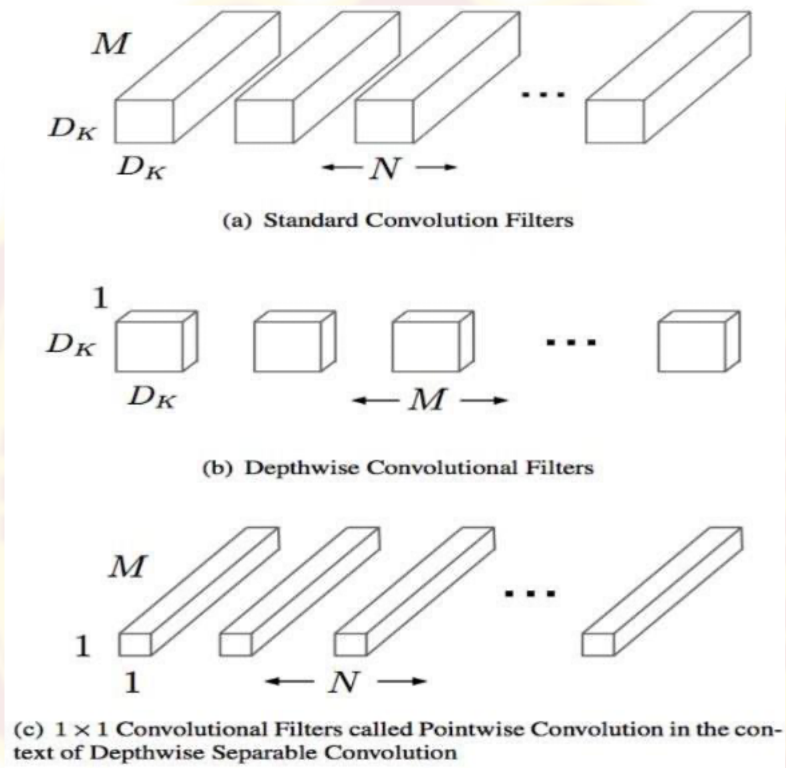
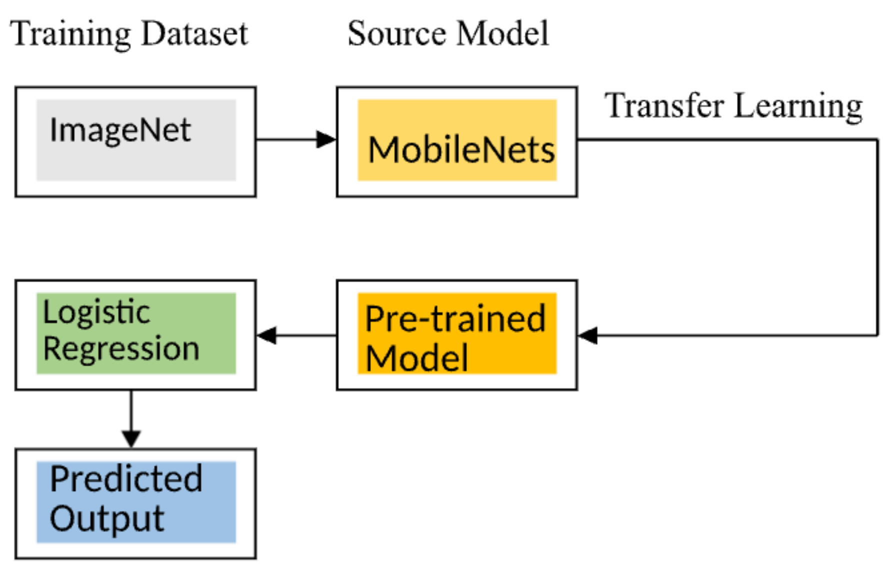


Figure 3 displays the methods and procedure process flow for the disease classification implementing the transfer learning approach with logistic regression. In this process, a network that is the MobileNets is pre-trained with a large-scale of general datasets, namely ImageNet, to effectively function as a generic template for visual processing. The pre-trained network transfers all the learned parameters and is set as a feature extractor for the target task to be executed. Logistic regression is used as a generic classifier for the classification.

Figure 3
Process Flow for Image Classification



RESULTS AND DISCUSSIONS

Displayed in Figure 4 is the developed standalone mobile application for leaf disease classification using durian leaves. MobileNets was applied as the pre-trained model for processing the inputted data alongside Tensorflow and Logistic Regression for the classification of the target task. By adding another classifier, the logistic regression, after processing the selected image, it got promising results.

Figure 4
Standalone Mobile Application

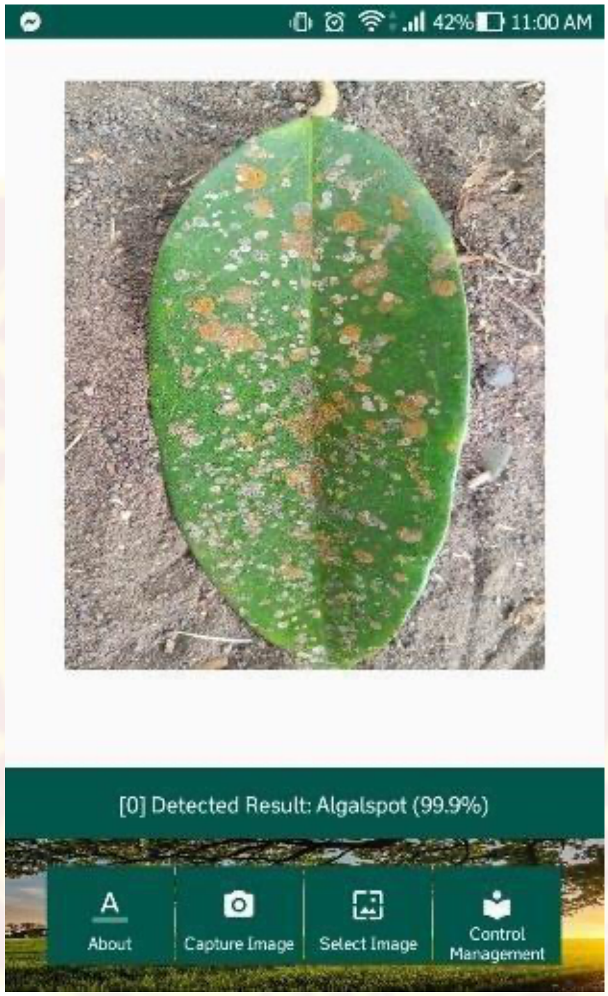


Figure 5 represents the accuracy result from the processed retrained model with the help of Tensorboard. The result produced 97% for the training dataset and 91% for the validation dataset. Shown in Table 1 is the disease prediction analysis result using durian leaves. The result is based on applying another add-on classifier, the logistic regression, conducted on 100 testing samples using the confusion matrix. The leaf disease was already classified and labeled before being subjected to testing. The computation shows that Algalspot disease yielded 88% or 22 out of 25 samples. Cercospora diseases yielded 96% or 24 out of 25 samples. Leaf discoloration yielded 100 %. Lastly, healthy durian Leaf yielded 92% or 23 out of 25 samples. As a result, from the 100 samples that were tested, 94 were correctly classified and resulted in 94% overall accuracy.

Figure 5
Accuracy Result

Table 1
Result of the Classification

N=100	Predicted Data				Classification Overall	Producer Accuracy
	Als	Crp	Lfd	Hty		
Als	22	3	0	0	25	*88%*
Crp	1	24	0	0	25	96%
Lfd	0	1	23	1	25	92%
Hty	0	0	0	25	25	100%
TOTAL	23	28	23	26	**100**	
User Accuracy	95.65%	85.71%	100%	96.15%		
Overall Accuracy				94%		

Legend:
Als = Algalspot — Crp = Cercospora
Lfd = Leaf Discoloration — Hty = Healthy Leaf
OA = Overall Accuracy

Conventional machine learning algorithms have been traditionally designed to work in isolation and are trained to solve specific tasks. The models have to be rebuilt from scratch once the feature-space distribution changes. It needs many datasets that would take days or weeks to fully finish all the training processes. The complexity of the computations in applying conventional machine learning is not suitable for developing a standalone mobile app. Internet and cloud-based transactions for computer computation and data exchanges are needed. Therefore, traditional method operation is proven to be inefficient and costly.

Considering that there is a need to have massive data resources, a huge amount of time for training that would probably take days or weeks, need the use of computers and an internet connection. With this, the proponents decided to adopt an alternative way or a new way approach to machine learning, which is the transfer learning approach, where one utilizes the knowledge acquired from training a specific model for the target task to solve related ones. All the processes and computations are embedded in a mobile phone, making the app a standalone mobile application. The conventional method shows high latency performance, while in adopting transfer learning, since all the processes are embedded within the device, a low latency performance is expected.

CONCLUSION AND FUTURE WORKS

In this paper, various procedures were conducted that the combination of the pre-trained MobileNet deep neural network architecture with the logistic regression classifier shows good accuracy for durian leaf disease detection and classification. MobileNet was trained for the new task as a feature extractor machine and logistic regression classifier trained on the target dataset, confirming that the proposed method outperforms similar state-of-the-art methods for plant recognition using transfer learning. The classifier correctly detected 94 out of 100 samples, which accumulated an overall accuracy of 94%.

REFERENCES

Cope, J. S., Remagnino, P., Barman, S., & Wilkin, P. (2010). The extraction of venation from leaf images by evolved vein classifiers and ant colony algorithms. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-642-17688-3_14

Daniel, R., Borines, L. M., Soguilon, C., Montiel, C., Palermo, V. G., Guadalquiver, G. A., … Guest, D. (2014). Development of disease management recommendations for the durian and jackfruit industries in the Philippines using farmer participatory research. *Food Security*. https://doi.org/10.1007/s12571-014-0352-6

Fang, Y., & Ramasamy, R. P. (2015). Current and prospective methods for plant disease detection. *Biosensors*. https://doi.org/10.3390/bios5030537

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*. https://doi.org/10.1016/j.compag.2018.02.016

Larese, M. G., Namías, R., Craviotto, R. M., Arango, M. R., Gallo, C., & Granitto, P. M. (2014). Automatic classification of legumes using leaf vein image features. *Pattern Recognition*. https://doi.org/10.1016/j.patcog.2013.06.012

Lee, S. H., Chan, C. S., Wilkin, P., & Remagnino, P. (2015). Deep-plant: Plant identification with convolutional neural networks. In *Proceedings - International Conference on Image Processing, ICIP*. https://doi.org/10.1109/ICIP.2015.7350839

Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., … Dandekar, A. M. (2015). Advanced methods of plant disease detection. A review. *Agronomy for Sustainable Development*. https://doi.org/10.1007/s13593-014-0246-1

Naresh, Y. G., & Nagendraswamy, H. S. (2016). Classification of medicinal plants: An approach using modified LBP with symbolic representation. *Neurocomputing*. https://doi.org/10.1016/j.neucom.2015.08.090

Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., & Hughes, D. P. (2017). Deep learning for image-based cassava disease detection. *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2017.01852

Sabarre, A. B. L., Navidad, A. N. S., Torbela, D. S., & Adtoon, J. J. (2021). Development of durian leaf disease detection on Android device. *International Journal of Electrical and Computer Engineering*. https://doi.org/10.11591/ijece.v11i6.pp4962-4971

Strange, R., & Scott, P. (2005). Plant disease: A threat to global food security. *Annual Review of Phytopathology*, 43, 83-116. https://doi.org/10.1146/annurev.phyto.43.113004.133839.

Surya Prabha, D., & Satheesh Kumar, J. (2014). Study on banana leaf disease identification using image processing methods. *International Journal of Research in Computer Science and Information Technology*.

Tongsri, V., Songkumarn, P., & Sangchote, S. (2016). Leaf spot characteristics of Phomopsis durionis on durian (durio Zibethinus Murray) and latent infection of the pathogen. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*. https://doi.org/10.11118/actaun201664010185