2024 7th International Conference on Information and Computer Technologies (ICICT) | 979-8-3503-8562-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICICT62343.2024.00051

The Effect of Data Augmentation and Padding of the Image Dataset on Detection of Black Sigatoka Disease on Banana Leaves Using ShuffleNet V2 CNN Architecture

Jonathan M. Baguisi
School of Electrical, Electronics, and Computer Engineering
Mapua University
Intramuros, Manila, Philippines
jmbaguisi@mymail.mapua.edu.ph

Baird Rouan S. Buenaventura
School of Electrical, Electronics, and Computer Engineering
Mapua University
Intramuros, Manila, Philippines
brsbuenaventura@mymail.mapua.edu.ph

Analyn N. Yumang
School of Electrical, Electronics, and Computer Engineering
Mapua University
Intramuros, Manila, Philippines
anyumang@mapua.edu.ph

***Abstract*—In this study, image processing is paired with Shufflenet V2 CNN Algorithm to detect Black Sigatoka infected banana leaves with the aid of Data Augmentation and Padding for testing and training. Black Sigatoka disease is one of the most important diseases for banana crops since it is responsible for most yield depletion and premature ripening. Currently, some studies tackle the detection of banana crop diseases using image classification with different algorithms for differentiating infected from healthy ones. As for the result, the study has tested Shufflenet V2 Algorithm in detecting Black Sigatoka using distinct training datasets. When applying the algorithm, different setups and alterations can be made for its training dataset to fully enhance the overall performance of the algorithm through Data Augmentation and Padding. Each Data Augmentation and Padding setup is tested for the Shufflenet V2 Algorithm in this study. The results show that with this small dataset, the ShuffleNet V2 CNN algorithm trained to detect Black Sigatoka on banana leaves, augmenting the image dataset and padding every image yielding an accuracy of 95%, specificity of 93.33% and sensitivity of 96.67%.**

Keywords—*image processing, Shufflenet V2, black sigatoka, confusion matrix, data augmentation*

I. Introduction

Image Classification has been an impactful technology in different industries in different applications, especially in the agricultural and medical fields. It is a process that utilizes different algorithms to differentiate objects from their input, an image, and these algorithms' performances vary from one another. It shows that the latest ones tend to be more effective than their predecessors, especially regarding the accuracy of their output. This technology could incorporate different techniques to ensure an effective outcome of its purpose, one of which is integrating Data Augmentation into its input data. Data Augmentation is a technique used for algorithm training and dataset expansion. It utilizes existing resources to synthesize new data through alterations of the said resources which can be added to the data pool to increase the dataset's original size. It proved to be an effective method applied in processing images to increase the dataset for model training of an algorithm [1]. This procedure benefits image classification because it increases the overall accuracy of the process, further improving the algorithm's performance and preventing it from overfitting. However, different augmentation of image data undergone within an image classification process may produce different results in terms of accuracy and other factors such as sensitivity and specificity of the outcome. Such modifications like padding may have their effects differ from a different method, such as cropping, to name one, which shows that it will somehow affect the performance of the algorithm from which applied.

Data Augmentation has been an effective tool for Image Classification and is used in many fields where image classification works well for its application. A study applied Data Augmentation to improve Turkish Telephone Speech Recognition with Out of Domain Data. According to the results, data augmentation has significantly decreased the Word Error Ratio of speech recognition up to 3.5%, proving that its application can improve the performance of speech recognition [2]. Another study that utilized Data Augmentation concluded that using few authentic and augmented images improves the classification performance of ships [3]. Other studies had introduced a different approach to Data Augmentation in which the results were as effective or even more as the conventional approach. One study introduced the Circular Shift Method to Data Augmentation applied to CNN Image Classification. It provides a new variation to CNN-based models that does not mitigate its input information, which shows consistent improvement on different CNN Models. Their study allowed the CNN algorithm with circular shift operation to learn part features and shift-invariance properties of image datasets aside from the benefits of using the conventional data augmentation approach [4]. Black Sigatoka Disease had a significant impact on the banana crop industry. It is responsible for at least 35-50% of the yield depletion in most crop cultivars and the premature ripening of its fruits [5]. It is caused by a fungal pathogen called Mycosphaerella fijiensis [6]. This fungus spreads quickly through air and water, primarily through rain and irrigation waters, affecting a group of trees within the

2769-4542/24/$31.00 ©2024 IEEE
DOI 10.1109/ICICT62343.2024.00051

vicinity of the infected tree [6]. It necrotizes the leaf's surface and disables the plant from processing photosynthesis which destroys its nourishment and eventually results in the premature ripening of its fruit. It usually thrives in humid areas with frequent rainfall, especially in tropical regions [7]. Image processing is one of the most versatile tools in enhancing or extracting useful information from images. Its applications are very useful as different algorithms can utilize the data to process analog input images [8]. In agriculture, image processing can detect physical defects from plants that indicate signs of diseases. Image processed output can be fed into algorithms such as convolution neural networks, K-nearest neighbor, support vector machines, and more for various applications. SVM algorithm can detect anemia, leaf diseases, or mosquitos [9, 10, 11]. A Convolution Neural Network is a form of artificial network specializing in applications that require image recognition and processing of image data. It is a very effective algorithm in most applications, such as translating sign language for the aurally impaired, which boasted 97.6% accuracy [12]. The CNN is divided into different layers where it operates, and these neural networks extract distinct features of the data [13]. It has proven to recognize contrasting plants through their classes, diseases, and even growth rate by analyzing and learning from its dataset [14]. It allows a simplistic but effective approach to feature differentiation of leaves to distinguish crop health and map out affected areas within the farm [15]. In medicine, it shows that it can detect and count white blood cells and detect skin diseases [16, 17]. CNN is a versatile algorithm and can be used in various applications if sufficient data is available.

Significant studies suggest that data augmentation has beneficial effects on image datasets, increasing the overall performance and accuracy of algorithms. However, each modification or augmentation applied to original images may produce a different overall performance and outcome of the algorithm used, primarily when used in image classification that uses such an algorithm like the Shufflenet V2 Convolutional Neural Network model. Each input that will serve as data for either training or actual application may yield different results based on whether that input has undergone data augmentation or not and what augmentation the data went through in which one may be more favorable over than the other.

This study aims to fully utilize Data Augmentation and see its effects on an image classification algorithm such as the Convolutional Neural Network. Firstly, it needs (1) to detect Black Sigatoka Disease using Shufflenet V2 CNN Architecture. The second is (2) to run trials of the architecture with modifications of the dataset, whether augmented or padded. Lastly is (3) to compare the results of each ShuffleNet V2 CNN model trained on different dataset modifications in terms of the algorithm's performance via accuracy, sensitivity, and specificity. These objectives will specify whether data augmentation can significantly affect an algorithm's performance positively, negatively, or with no changes.

This study aims to understand further the effects of Data Augmentation on each dataset used for training algorithms and models. It will significantly benefit researchers and professionals alike in utilizing data augmentation effectively by knowing the kind of augmentation needed to acquire the most desirable results. It also shows whether using a different approach to data augmentation using ShuffleNet V2 CNN Architecture may be effective or not through the results of this study.

This study aims to utilize Data augmentation and padding on the image dataset using Shufflenet V2 Convolutional Neural Network Model in detecting Black Sigatoka Leaves on Banana leaves only. It involves applying Data Augmentation and Padding to some image data, and some will not be augmented or padded in which all will be fed to the model. Other types of crops that can be affected by the disease will be out of the scope of this study, and other diseases can also manifest in the banana or other crops. Other algorithms and models viable in detecting Black Sigatoka Disease are also out of this study's scope.

II. Methodology

A. Conceptual Framework

Figure 1 is the conceptual framework of the system. The framework has Black Sigatoka infected banana leaves as input. The ShuffleNet V2 CNN architecture will be trained on data with specific data augmentation and padding applied individually, producing different CNN models accordingly. Once the accuracy, specificity, and sensitivity of each ShufleNet V2 CNN model are obtained, the framework's final output is a tabulated parameter comparison of each of the algorithms trained on different datasets.

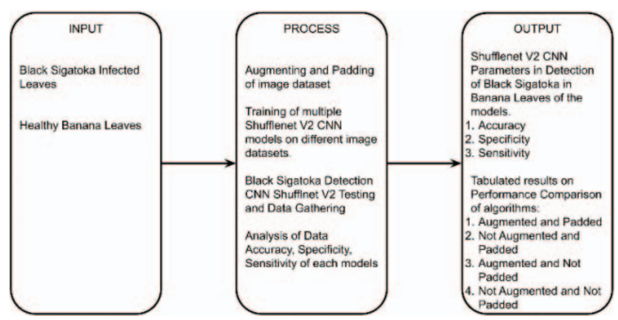


Fig. 1. Conceptual framework.

The parameters will be the basis of comparison for each model used for this paper. Accuracy dictates the algorithm's error rate or how often the system will commit an error. Sensitivity depicts the ability of the system to detect true positive results or detect the disease correctly while Specificity is the ability of the system to distinguish if the specimen is healthy. The results of these parameters will show the difference between each modification of the model's performance.

B. Hardware Development

The hardware is composed of a few main components. These components are the Raspberry Pi, Camera Module, LCD Monitor, and power supply. The power supply is only connected to the raspberry pi as the peripherals are powered directly by the power supply. The input of the experimental setup is the camera module supported by a peripheral device, LED lights. The use of LED lights combined with the five (5) megapixel camera captures good quality images of specimens. The raspberry pi is the system's processor, which processes these images using the ShuflleNet V2 CNN model. The output device is the LCD monitor that displays the captured image and the classification of whether the specimen is healthy or Black Sigatoka infected made by the raspberry pi using a GUI.

C. Image Processing

Models in this study used the same training image dataset with differences only in the image processing techniques used in the raw image training dataset. There are two major types of processing techniques used in this study. These techniques are Image augmentation and padding.

Image or data augmentation is altering or modifying the existing image data to expand data points for training a learning model artificially [18]. This study used several techniques, such as rotate, flip, tilt, skew, brightness change, and contrast change. Rotation modifies images and rotates them to a random angle. Flip modifies the images to flip along their horizontal or vertical axis. Tilt modifies the images to tilt the corners. Skew modifies the image corners to tilt one of the corners forwards or backward. Brightness change modifies the raw image brightness to a random degree up or down. Finally, contrast change modifies the raw image to adjust the image contrast to a certain degree. All these augmentation techniques have a preset probability and are randomly applied.

Image padding is the process of adding data to the raw input image [19]. The most common padding is called zero-padding, where adding 0, black, around the raw image. The type of padding used in this study adds the value 255, white, where data is inserted around the raw image before being fed to the algorithm.

Figure 2 shows examples of augmented, not augmented, padded, and not padded images applied to the raw image training dataset and fed to the ShufflNet V2 CNN algorithm for each of the different models.

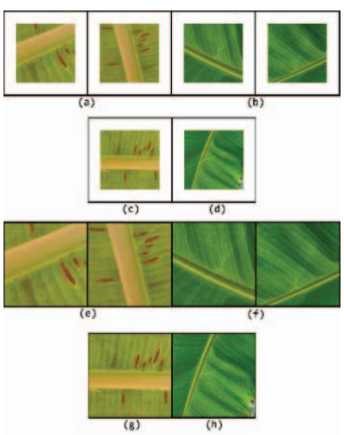


Fig. 2 (a) Augmented & Padded Infected Leaves
(b) Augmented & Padded Healthy Leaves
(c) Not Augmented & Padded Infected Leaf
(d) Not Augmented & Padded leaf
(e) Augmented & Not Padded Infected Leaves
(f) Augmented & Not Padded Healthy Leaves
(g) Not Augmented & Not Padded Infected Leaf
(h) Not Augmented & Not Padded Healthy Leaf

D. Software Development

The main flowchart of the system shown in Figure 3 is the flow of the detection program. The first step of the flow is to capture an image of the leaves using the camera. The image is then processed, which crops and resizes the image to prepare for the classification. The CNN process then classifies the image to either healthy or Black Sigatoka infected banana leaves. After the classification, it is then displayed along with the image on the GUI. The program will continue this loop until the user exits the program.

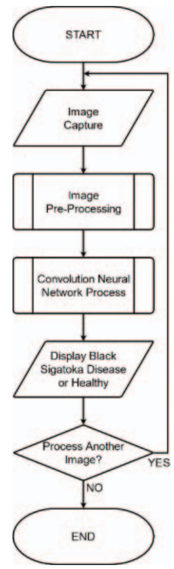


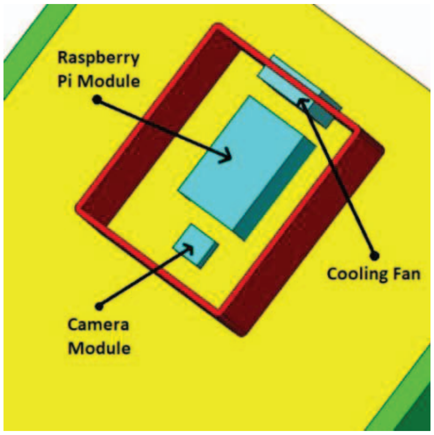
Fig. 3. System flowchart

E. Training of Different Models

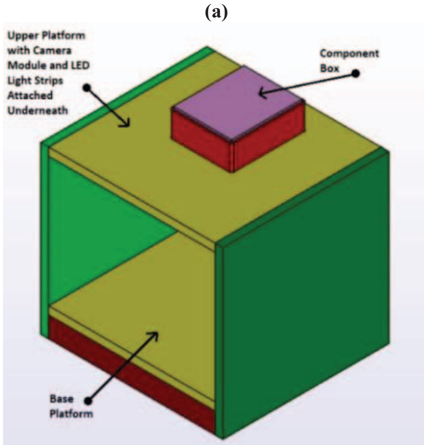
In this study, all four (4) models use the same ShuffleNet V2 CNN algorithm with the only difference being the types of image processing applied to the training dataset before feeding the data to the algorithm. Based on the types of processing applied to the dataset, each trained model will be named accordingly in this study based on various combinations of data augmentation and padding applied. The CNN model that makes use of data augmentation and padding in its training dataset is described as Model 1. Model 2 uses only padding before feeding it to the training algorithm. Model 3 makes use of augmentation, and no padding is applied to the training dataset. Finally, Model 4 is trained with data without padding and augmentation.

F. Experimental Setup

The prototype setup for this study is shown in Figure 4. The system's input device is the camera and the touch screen LCD. The camera captures images of the specimen assisted by LED lights installed in a square pattern to get a high-quality image. The captured image is then processed by the raspberry pi and classify the image using ShuffleNet V2 CNN trained algorithm. The classified images are then shown thru the GUI in the LCD touch screen of the system, along with the classification made by the CNN algorithm.



(a)



(b)

Fig. 4 Prototype Setup 3D Model (a) Top View (b) Isometric View

There are two classifications of the ShuffleNet V2 CNN is trained to detect. Thirty (30) samples of healthy leaves and (30) samples of Black Sigatoka infected leaves are used to gather data. Shown in Figure 5 are examples of the two classifications healthy and Black Sigatoka. There is a third hidden classification named unknown, which is only shown when no specimens are detected at the base of the experimental setup.

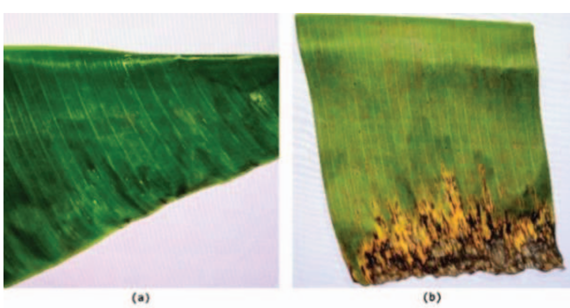


Fig. 5. Example images of specimens (a) Healthy (b) Black Sigatoka

III. Results and Discussion

Table 2 shows the results of the system's testing. The testing process for the system made use of banana leaf specimens gathered and verified by an experienced agriculturist. An expert classified the actual column, and the predicted columns are the system's predicted classification of each specimen on their model. The system classifies captured images and processes them using the raspberry pi computer. It will then output whether the specimen is healthy or Black Sigatoka infected. Thirty (30) specimens of healthy leaves were cut from three (3) banana leaves, and thirty (30) specimens of Black Sigatoka infected leaves were cut from three (3) infected banana leaves. Model 1 is trained on augmented images and has white padding on its borders. Model 2 is trained on images that are not augmented and have white padding on their border. Model 3 is trained on augmented images but has no white padding on its border. Finally, model 4 is trained on images that are not augmented and have no white padding on its border.

Several confusion matrices are produced from the classification result. These confusion matrices are shown in Tables 1, 2, 3, and 4, one for each ShuffleNet V2 CNN model trained on different image processing techniques. Model 1, shown in Table 2, utilizes augmentation and padding. Model 2, shown in Table 3, utilized padding only. Model 3, shown in Table 4, utilized augmentation but without the padding. Finally, Model 4, shown in Table 5, uses the raw image dataset, where no augmentation or padding is done before being fed to the CNN algorithm.

TABLE I. Confusion Matrix OF Model 1

ShuffleNet V2 CNN Model 1 (Augmented & Padded)		PREDICTED	
		BLACK SIGATOKA	**HEALTHY**
ACTUAL	**BLACK SIGATOKA**	**29**	**1**
	HEALTHY	**2**	**28**

TABLE II. Confusion Matrix of Model 2

ShuffleNet V2 CNN Model 2 (Not Augmented & Padded)		PREDICTED	
		BLACK SIGATOKA	**HEALTHY**
ACTUAL	**BLACK SIGATOKA**	**26**	**4**
	HEALTHY	**1**	**29**

TABLE III. Confusion Matrix of Model 3

ShuffleNet V2 CNN Model 3 (Augmented & Not Padded)		PREDICTED	
		BLACK SIGATOKA	**HEALTHY**
ACTUAL	**BLACK SIGATOKA**	**28**	**2**
	HEALTHY	**15**	**15**

TABLE IV. Confusion Matrix of Model 4

ShuffleNet V2 CNN Model 4 (Not Augmented & Not Padded)		PREDICTED	
		BLACK SIGATOKA	**HEALTHY**
ACTUAL	**BLACK SIGATOKA**	**30**	**0**
	HEALTHY	**29**	**1**

The formulas to prepare the comparison table used are equations (1), (2), and (3). The True Positive (TP) refers to the Black Sigatoka specimens correctly predicted, and True Negative (TN) as the Black Sigatoka incorrectly classified as healthy. The True Negative (TN) is the healthy specimens correctly classified, and the False Negative (FN) is the healthy specimens incorrectly predicted as Black Sigatoka. Table 6 compares the accuracy, sensitivity, and specificity of the different ShuflleNet V2 CNN trained models. The model with the highest parameters is Model 1, which achieved 95% accuracy, 9.67% sensitivity, and 93.33% specificity. Meanwhile, the ShuffleNet V2 CNN model with the lowest parameters is model 4, which only achieved 51.67% accuracy, 100% sensitivity, and 3.33% specificity.

$$\%Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \tag{1}$$

$$\%Sensitivity = \frac{TP}{TP+FN} \times 100 \tag{2}$$

$$\%Specificity = \frac{TN}{TN+FP} \times 100 \tag{3}$$

TABLE VI. Parameter Comparison Table

ShuffleNet V2 CNN	**Model 1 (Augmented & Padded)**	**Model 2 (Not Augmented & Padded)**	**Model 3 (Augmented & Not Padded)**	**Model 4 (Not Augmented & Not Padded)**
Accuracy	**95.00%**	**91.67%**	**71.67%**	**51.67%**
Sensitivity	**96.67%**	**86.67%**	**93.33%**	**100%**
Specificity	**93.33%**	**96.67%**	**50%**	**3.33%**

IV. Conclusion

The outcome based on the results shows that the ShuffleNet V2 CNN algorithm has significant differences between all the trained models. Model 1 is trained on augmented and padded images, resulting in the best accuracy, sensitivity, and specificity. The best results came from Model 1, with accuracy achieved is 95%, sensitivity at 96.67%, and specificity at 93.3%. With a very small dataset, augmentation and white padding significantly affect a trained model's accuracy, sensitivity, and specificity. Throughout the testing of the specimens, it has been observed that the models that are trained on non-augmented images produce different results when rotated, while models trained on augmented images have more ease producing the same results even when the specimen is rotated in various orientations. Models that are not padded, such as Model 3 and Model 4, have difficulties classifying specimens correctly that are small, while these models can still classify leaf specimens correctly if the specimen covers the whole area of the camera. Padding the border of the image dataset trains the model in focusing on classifying the specimen and augmenting the images allowing the model to classify the specimens in various orientations. Thus, it is evident that with this specific small dataset, the ShuffleNet V2 CNN algorithm that is trained to detect Black Sigatoka on banana leaves, applying data augmentation and padding produces the best accuracy, specificity, and sensitivity.

References

[1] L. Perez and J. Wang, "The Effectiveness of Data Augmentation in Image Classification using Deep Learning," *arXiv,* 13 12 2017.

[2] Z. G. Uslu and T. Yıldırım, "Improving Turkish Telephone Speech Recognition with Data Augmentation and Out of Domain Data," *2019 16th International Multi-Conference on Systems, Signals & Devices (SSD'19),* pp. 176-179, 2019.

[3] S. Moon, J. Lee, J. Lee, A. R. Oh, D. Nam and W. Yoo, "A Study on the Improvement of Fine-grained Ship Classification through Data Augmentation Using Generative Adversarial Networks," *2021 International Conference on Information and Communication Technology Convergence (ICTC),* pp. 1230-1232, 2021.

[4] K. Zhang, Z. Cao and J. Wu, "Circular Shift: An Effective Data Augmentation Method For Convolutional Neural Network On Image Classification," *2020 IEEE International Conference on Image Processing (ICIP),* pp. 1676-1680, 2020.

[5] R. S. Bennet and P. A. Arnerson, "Black sigatoka of bananas and plantains," *The Plant Health Instructor,* 2003.

[6] L. I. S. Gomes, G. W. Douhan, L. B. J. Bibiano, L. A. Maffia and a. E. S. G. Mizubuti, "Mycosphaerella musicola Identified as the Only Pathogen of the Sigatoka Disease Complex Present in Minas Gerais State, Brazil," *Plant Disease,* vol. 97, no. 12, pp. 1537-1543, 12 2013.

[7] A. Muimba-Kankolongo, Food Crop Production by Smallholder Farmers in Southern Africa, Elsevier, 2018.

[8] B. Sridhar, "Applications Of Digital Image Processing In Real Time World," *International Journal Of Scientific & Technology Research,* vol. 8, no. 12, pp. 3354-3357, December 2019.

[9] C. C. Hortinela, J. R. Balbin, J. C. Fausto, P. D. C. Divina and J. P. T. Felices, "Identification of Abnormal Red Blood Cells and Diagnosing Specific Types of Anemia Using Image Processing and Support Vector Machine," *2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM),* pp. 1-6, 1 11 2019.

[10] D. A. Padilla, G. V. Magwili, A. L. A. Marohom, C. M. G. Co, J. C. C. Gaño and J. M. U. Tuazon, "Portable Yellow Spot Disease Identifier on Sugarcane Leaf via Image Processing Using Support Vector Machine," *2019 5th International Conference on Control, Automation and Robotics (ICCAR),* pp. 901-905, 4 2019.

[11] A. M. M. De Los Reyes, A. C. A. Reyes, J. L. Torres, D. A. Padilla and J. Villaverde, "Detection of Aedes Aegypti mosquito by digital image processing techniques and support vector machine," *2016 IEEE Region 10 Conference (TENCON),* pp. 2342-2345, 11 2016.

[12] J. R. Balbin, D. A. Padilla, F. S. Caluyo, J. C. Fausto, C. C. Hortinala, C. O. Manlises, C. K. S. Bernardino, E. G. Fiñones and L. T. Ventura, "Sign language word translator using Neural Networks for the Aurally Impaired as a tool for communication," *2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE),* pp. 425-429, 25 11 2016.

[13] N. Ma, X. Zhang, H.-T. Zheng and J. Sun, "ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design," *arXi,* 30 7 2018.

[14] D. A. Padilla, R. A. I. Pajes and J. T. De Guzman, "Detection of Corn Leaf Diseases Using Convolutional Neural Network With OpenMP

Implementation," *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM),* pp. 1-6, 3 12 2020.

[15] H. S. Abdullahi, R. E. Sheriff and F. Mahieddine, "Convolution neural network in precision agriculture for plant image recognition and classification," *2017 Seventh International Conference on Innovative Computing Technology (INTECH),* pp. 1-3, 8 2017.

[16] M. J. Macawile, V. V. Quiñones, A. Ballado, J. Dela Cruz and M. V. Caya, "White blood cell classification and counting using convolutional neural network," *2018 3rd International Conference on Control and Robotics Engineering (ICCRE),* pp. 259-263, 4 2018.

[17] D. Padilla, A. Yumang, A. L. Diaz and G. Inlong, "Differentiating Atopic Dermatitis and Psoriasis Chronic Plaque using Convolutional Neural Network MobileNet Architecture," *2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM),* pp. 1-6, 11 2019.

[18] M. S. Alam, D. Wang and A. Sowmya, "Image data augmentation for improving performance of deep learning-based model in pathological lung segmentation," *2021 Digital Image Computing: Techniques and Applications (DICTA),* pp. 1-5, 2021.

[19] M. A. Islam, M. Kowal, S. Jia, K. G. Derpanis and N. D. B. Bruce, "Position, Padding and Predictions: A Deeper Look at Position Information in CNNs," 2021.