Tomato Fruit Image Dataset for Deep Transfer Learning-based Defect Detection

Robert G. de Luna, Elmer P. Dadios, Argel A. Bandala, Ryan Rhay P. Vicerra
Gokongwei College of Engineering, De La Salle University Manila, Philippines
robert_g_deluna@dlsu.edu.ph, elmer.dadios@dlsu.edu.ph, argel.bandala@dlsu.edu.ph, ryan.vicerra@dlsu.edu.ph

Abstract— **Tomato is considered as one of the vegetable crops with highest demand in the Philippines. Job of the farmers does not end after harvesting since the harvested tomatoes needed to be sorted according to its size. Manual sorting is the most widely recognized strategy in sorting but is very dependent on human interpretation and thus, very prone to error. This research proposed a solution that provides sorting of tomato fruit by detection of presence of defect. The study presented the generation of image dataset for a deep learning approach detection of defects based from a single tomato fruit image. Models were implemented using OpenCV libraries and Python programming. A total of 1200 tomato images classified as no defect and with defect are gathered using the improvised image capturing box. These images are used for the training, validation, and testing of the three deep learning models namely; VGG16, InceptionV3, and ResNet50. From this, 240 images are used as testing images to assess independently the performance of the trained models using accuracy and F1-score as performance metrics. Experiment results shown that VGG16 has 95.75-95.92-98.75 training-validation-testing accuracy percentage performance, 56.38-59.24-58.33 for the InceptionV3 model, and 90.58-58.46-64.58 for the ResNet50. Comparative analysis revealed that VGG16 is the best deep learning model to be used in the detection of presence of defect in the tomato fruit based from the dataset gathered.**

keywords— tomato defect detection, deep transfer learning, image processing, solanum lycopersicum

I. Introduction

Tomato (Solanum Lycopersicum), popularly known as "Kamatis", is one of the most important vegetable crops in the Philippines. It is utilized as an ingredient in numerous food preparation and is viewed as a standout amongst the most beneficial crops for off-season production, ideally from May to September. During April to June 2018, the production of tomato increased from 72.19 thousand metric tons of the previous year to 73.50 thousand metric tons this year [1]. Farmers usually performed the segregation of tomatoes according to size and quality, whether with defect or non-defect. The accepted quality of tomatoes is separated from the one that has defects or the so-called rejects, in order to isolate the good from bad and prevent the spreading of defects. These rejects are either thrown away or sold at a lower price.

The traditional method sorting of tomatoes is manually done by persons accustomed in sorting and consumes a lot of time alone. Relying only on human inspection, the detection of defects from the crop could lead to substantial error. The tendency of shipping quality tomatoes with defects unsegregated will cause a huge effect in customer's feedback and satisfaction.

To prevent the occurrence of that possibility, the study developed a computer-vision system utilizing the concept of image processing and deep learning that will capture the image of the tomatoes and detect automatically the presence of defects. Image capturing box is designed to capture batch of tomatoes into single image. Image processing is then performed using OpenCV libraries in the enhancement and segmentation of the tomatoes. Three deep learning models are retrained for the specific purpose of defect detection using Python as the programming language. The developed system may serve as an aiding tool to segregate the quality tomato to defective tomato with a reliable accuracy that is no longer based from human assessment and subjectivity.

II. Related Works

Tomato is considered as one of the most largely grown agricultural products in the world. One of the problems in the tomato production is the presence of the disease that normally seen in the leaf and fruit of the plant. Therefore, early detection and diagnosis of plant diseases is very important in agriculture.

Machine vision nowadays play an important role in agricultural sector thru the so-called smart farming that functions to provide growth assessment [2], volume estimation of fruits, recognition of the plants [3], and detection and classification of diseases. In the study of [4], a machine-vision system that estimate the volume of tomatoes by image processing techniques was developed. Volume of the fruit is computed by estimating horizontal and vertical distance of captured images.

Also essential in the process of quality inspection for tomato is its maturity or ripeness grading. Several studies were produced focusing in the tomato fruit. The study [5] dealt more of the appearance of the fruit. Researchers used the color, size, and shape of the tomato fruit in creating a fuzzy logic system for maturity grading. Researchers of [6] inspected the quality of tomato according to the following features: shape, size, and degree of ripeness. An edge detection algorithm is used in estimating the shape and size of the tomato while color detecting algorithm is utilized for the ripeness determination. Considering the cost of implementation, [7] discussed a cost-effective maturity grading system for tomato that can identify the six important stages of tomato ripening based from different image processing algorithms. All algorithms were designed and developed using Simulink.

Image processing with machine learning also provided a good strategy in plant's maturity assessment. Study from [8] presented different machine learning type algorithms for tomato maturity classification based from its color. Supervised and trained classification algorithms like KNN, MLNP, and K-means clustering techniques were implemented with features in

978-1-7281-3458-1/19/$31.00

the RGB, HSI and L*a*b* color spaces. Similarly, [9] focused on representing the technique of fruit grade classification, automated machine vision-based technology. The grading process is carried out by capturing the fruit image using camera and subjected to various image processing techniques. Otsu thresholding and K-Means clustering algorithms are used to extract the significant features of the fruit and the Support Vector Machine algorithm was developed for the quality grading.

In machine-vision, information from captured images should be properly extracted and interpreted. Segmentation process is an essential part in image processing to obtain good preparation either for further process of data mining or object recognition. To lessen the effect of varying illumination, [10] introduced a method that utilized median filter to produce illumination irrelevant images was introduced to segment the ripe tomato. It uses minimum entropy criterion in calculating the illumination irrelevant angle of the given camera. Another method for segmentation was introduced in [11]. Images are gathered using three types of smartphone camera in various lighting condition with white background. The method used color transformation, applied histogram equalization, and the segmentation. Otsu combined with V channel thresholding is utilized to segment image better. Application in a picking robot was developed in the study of [12] where researchers provided an algorithm inspired by quantum-inspired ant colony for the tomato fruit picking manipulator to segment tomato from the complex background accurately. This ensures the picking activity very accurate in recognizing the tomato fruit.

The machine-vision system is also been used in the growth monitoring of plants especially in recognizing and detecting some infections or diseases. In this paper [13], tomato maturity that is based from the tomato leaf color and detection of fungal infection are determined. A more generalized and self-adapting system was produced using k-means clustering algorithm in the determination of tomato maturity and disease. Statistical and color features approach were used in the study of [14]. New approach was introduced to model the Scale Invariant Feature Transform (SIFT) texture feature by Johnson SB distribution for statistical texture information of an image. For color feature, statistical color information of an image from RGB color channel is extracted. Combination of the statistical texture and color features are used to classify tomato plant disease. Another approach was used in [15], five types of tomato diseases are classified using the classification tree algorithm. Color, shape, and texture features are extracted from healthy and unhealthy tomato plant images. These extracted features from segmented images are fed into the classification tree model to classify the disease. More complex architecture algorithm as compared to machine learning can be seen in the deep learning models.

Certain limitations of machine learning paved a way for the consideration of deep neural networks in developing intelligent models. The study of [16] provided enough disease detection in tomato leaf with a motor-controlled image capturing box was. The system was designed to identify the diseases namely Phoma Rot, Leaf Miner, and Target Spot using a deep convolutional neural network, with AlexNet as the architecture. Another attempt was made by [17] wherein the research focused on building a model based on SqueezeNet architecture to classify seven types of tomato plant diseases on the leaves including healthy leaves. The model was developed with using the Keras deep learning frameworks. Deep learning implementation in Nvidia Jetson TX1 was the focused of study in [18]. The objective of this study is to detect diseases on the leaves of tomato plants in the fields or greenhouses using deep learning approach. There are two different deep learning network architectures were tested, namely, AlexNet and SqueezeNet. Considering the heavy hardware requirements of the deep learning, [19] adopted a minimal variation of the convolutional neural network model called LeNet in detecting and identifying diseases in tomato leaves. This is a solution to the problem of tomato leaf disease detection using the simplest approach while making use of minimal computing requirement to achieve results that are comparable to state-of-the-art techniques. While several deep learning techniques are used in the detection of disease in tomato via leaf, the study of [20] focused on detection with tomato fruit as an object. This study utilized a deep neural network that recognize and predict the nutrient deficiencies in the tomato during the fruiting phase. Inception-Resnet v2 based-Convolutional Neural Network is applied. There are lots of interesting deep-learning-based work in the robotics community and some of those provided applications of computer-vision in robotics for agriculture [21] [22] [23].

The proposed study is to develop an image capturing platform to generate dataset of tomato fruit images. These images are intended to be used in three pre-trained deep learning architectures, namely, VGG16, Inceptionv3, and ResNet50 to classify the tomato fruit as with defect or without defect. Pre-processing and image cropping of images were performed prior to the deep transfer learning training.

III. Methodology

A. Dataset Description

There are 1200 single tomato images were gathered using an improvised image capturing system for the training and testing of the deep learning models for defect detection. Fig. 1 shows a sample gathered data.

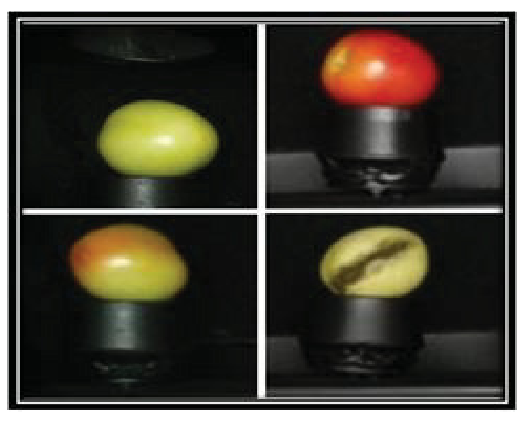


Fig 1. Sample Tomato Image

From this number, 960 images classified with defect and no defect are used in the training of the models. The defect images are collected from tomato that is infected with early blight, anthracnose, fruitworm, sunscald, and physically damaged with cracks and split. The remaining 240 images are reserved for their accuracy test.

B. System Overview

Fig. 2 below shows the overview of the system. The input image is captured from the image capturing box using digital camera, Raspberry Pi Camera Module V2. The said camera is an 8-megapixel camera module capable of 1080p video and still image with a resolution capable up to 3280 x 2464 pixels. The distance of the camera to the pile of 5 tomatoes is 15 inches with a separation of 4 inches.

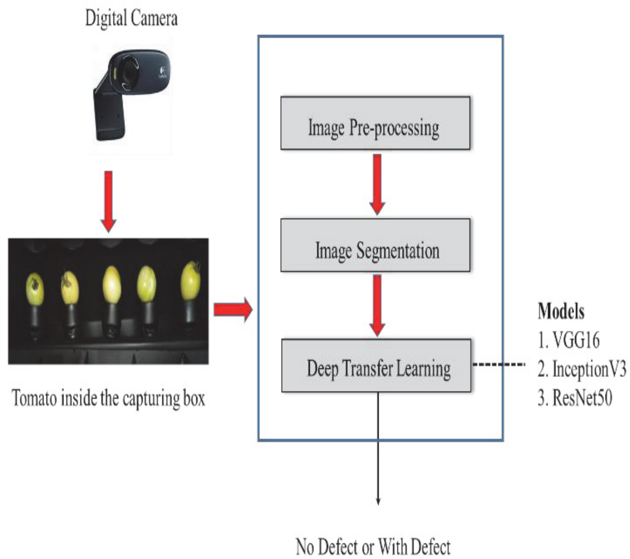


Fig 2. Overview of the Proposed System

Tomatoes are captured using the digital camera and subjected to image pre-processing using OpenCV libraries in Python. The processed image is then segmented to isolate each tomato in the image thus producing 5 cropped images of single tomato cropped to 656 x 656-pixel dimension. The cropped images are then resized to a dimension fitted to the requirement of each deep learning models, 224 x 224 for VGG16 and ResNet50 while 229 x 229 for InceptionV3. The performance of the three models are evaluated based from the accuracy and F1-score of the detection using the testing dataset in order to determine the model that will provide the best detection.

C. Development of the Deep Learning Model

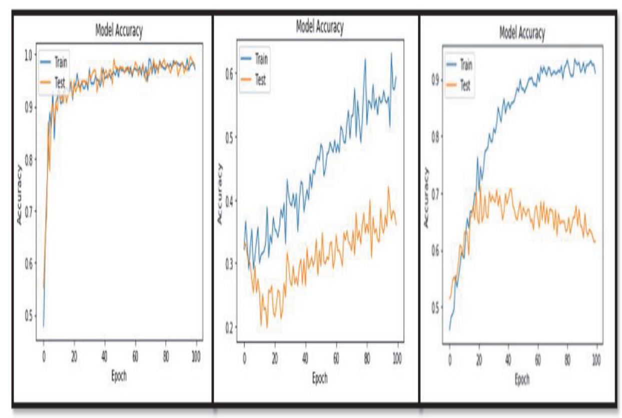
Visual Geometry Group Network (VGGNet) architecture is based from the AlexNet architecture which was improved by increasing the performance of the convolutional and the pooling layers. The cropped images are resized to 224 by 224 pixels. VGGNet16 consists of 16 convolutional layers and uses small-sized filters. ResNet or the Residual Neural Network on the other hand, modified the VGGNet architecture of VGGNet by adding layers. It consists of 152 layers of convolutional, pooling, and fully connected layers. Basically, increasing the layers will make the architecture's accuracy lower due to degradation of the quality of image. However, the GoogleNet, also known as the Inception, is a much more complex network architecture compared to the other architecture because of the introduction of a module called Inception. This module basically let the model decide the best size for each convolutional layer, thus, making it more complicated.

For consistency of the training, the same reserved 960 images are used for the training development of each model. It will then be optimized for best performance and tested using the same set of separate 240 images.

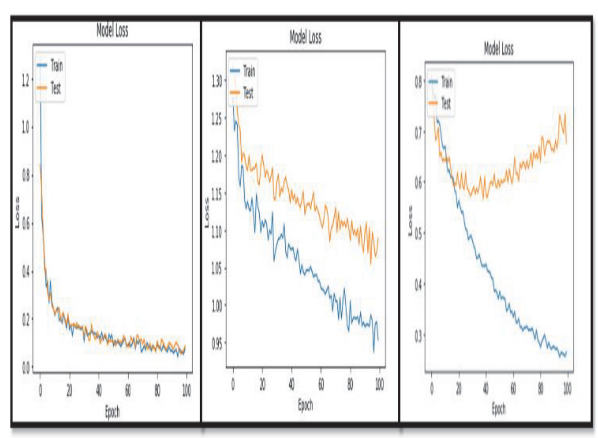
IV. Results and Discussion

The three deep learning network architecture implemented in this study are the VGG16, InceptionV3, and the ResNet50. For all models, the gathered tomato images for training are divided into 80 % to training and 20 % for the validation. Parameters are set to batch size=32, learning rate=0.0001, step size=10, and optimizers=adam. In every model, the average accuracy over the whole period of training at 100 epochs is calculated and tabulated. Separate 240 images are used to evaluate the accuracy of the model in actual implementation.

Fig. 3 and Fig. 4 show the accuracy plot and loss plot respectively of the three models in every epoch during the training. By comparison, the VGG16 had the better training and validations results that did not over fit while ResNet50 provided a worst performance.

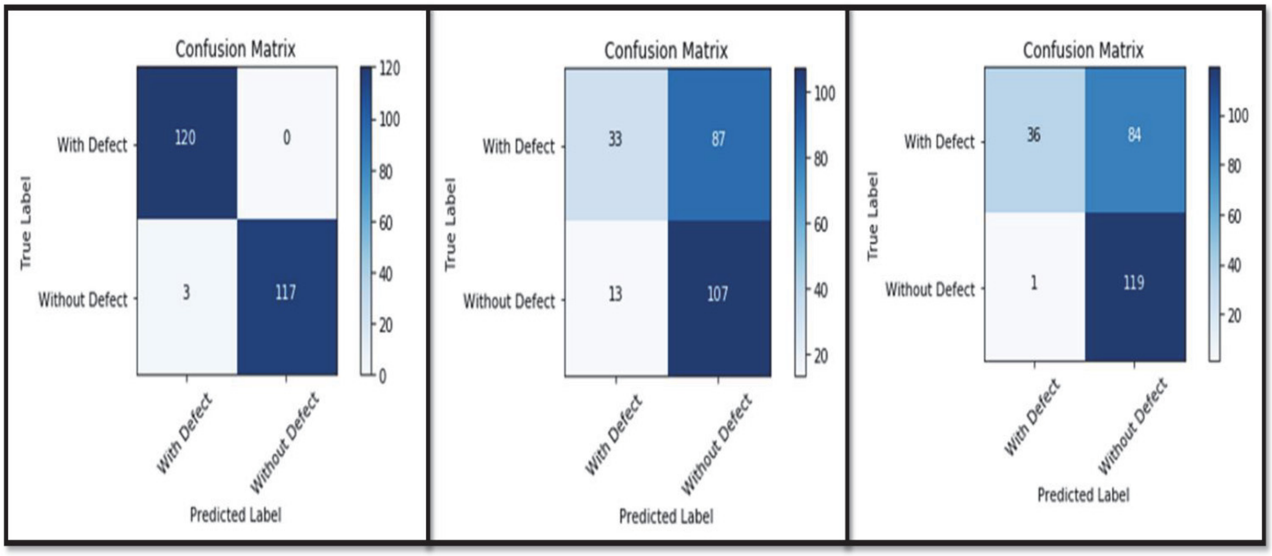


a. VGG16 b. InceptionV3 c. ResNet50
Fig 3. Accuracy Plot of the Three Deep Learning Models during Training and Validation Under 100 Epochs



a. VGG16 b. InceptionV3 c. ResNet50
Fig 4. Loss Plot of the Three Deep Learning Models during Training and Validation Under 100 Epochs

After the training development, the three models are evaluated again using separate 240 tomato images. Fig. 5 and Table 1 summarized the performance of every model according to confusion matrix, accuracy percentage and F1-score.



a. VGG16 b. InceptionV3 c. ResNet50

Fig 5. Confusion Matrix of the Three Deep Learning Models

Table 1. Accuracy Performance of the Three Deep Learning Models for Defect Detection

	Training		Validation		Testing						
Model	**No. of Samples**	**Average Accuracy, %**	**No. of Samples**	**Average Accuracy, %**	**Size**	**No. of Samples**	**Correct**	**Incorrect**	**F1-Score**	**Accuracy, %**	**Over-all Accuracy, %**
VGG16	768	95.75	192	95.92	With Defect	120	120	0	0.99	100.00	98.75
					Without Defect	120	117	3	0.99	97.50	
InceptionV3	768	56.38	192	59.24	With Defect	120	33	87	0.40	27.50	58.33
					Without Defect	120	107	13	0.68	89.17	
ResNet50	768	90.58	192	58.46	With Defect	120	36	84	0.77	30.00	64.58
					Without Defect	120	119	1	0.84	99.17	

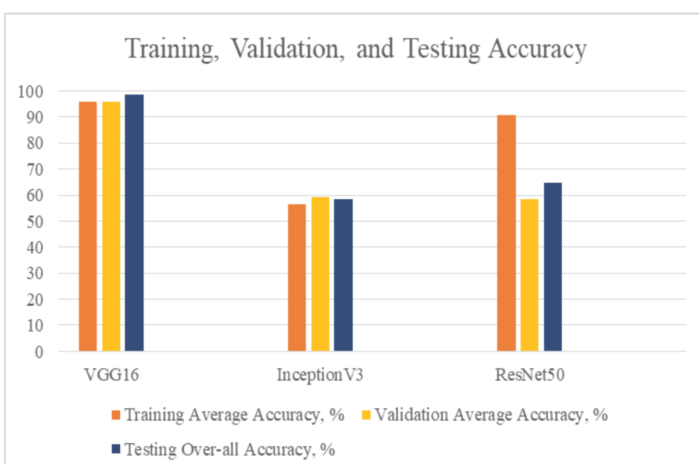


Fig 6. Comparison of Training, Validation, and Testing Accuracy of Three Different Deep Learning Model

Fig. 6 reflected the performance of the three models in the training, validation, and testing stage. Comparison of the three model's performance revealed that the architecture of VGG16 is the best model to be used in the detection of defect for tomato fruit being consistent with high percentage of accuracy.

V. CONCLUSIONS

There are three deep learning models implemented in the detection of defects in the generated image dataset of tomato fruits. These are the VGG16, InceptionV3, and the ResNet50. All the three common network architecture in deep learning are retrained using the gathered images of tomato fruits classified as with defect and without defects. A total of 1200 images were used in the training, validation, and testing of all models. VGG16 registered a training-validation-testing accuracy of 95.75%-95.92%-98.75%, InceptionV3 with 56.38%-59.24%-58.33%, while 90.58%-58.46%-64.58% for ResNet50. VGG16 outperformed the other two models and thus considered as the best model due to its best performance in the detection of presence of defect in the tomato fruit.

VI. ACKNOWLEDGEMENT

The authors would like to acknowledge the Department of Science and Technology (DOST) through the Engineering Research and Development for Technology (ERDT) for funding this research. The authors also recognize the assistance and technical expertise of members of the Intelligent Systems Learning (ISL) Laboratory of the De La Salle University (DLSU) Manila.

VII. REFERENCES

[1] "Philippine Statistics Authority," 2018. [Online]. Available: https://psa.gov.ph/vegetable-root-crops-main/tomato.

[2] I. C. Valenzuela, J. C. Puno, A. A. Bandala, R. G. Baldovino and R. G. D. E. P. de Luna, "Quality assessment of lettuce using artificial neural network," in *2017 IEEE 9th International Conference on Humanoid Nanotechnology Information Technology Communication and Control Environment and Management*, Manila, 2017.

[3] R. G. de Luna, R. G. Baldovino, E. A. Cotoco, A. L. de Ocampo, I. C. Valenzuela, A. B. Culaba and E. P. Dadios, "Identification of philippine herbal medicine plant leaf using artificial neural network," in *2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management* , Manila, 2017.

[4] S. UluiŞik, F. Yildiz and A. T. Özdemİr, "Image processing based machine vision system for tomato volume estimation," in *2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting*, Istanbul, 2018.

[5] J. B. U. Dimatira, E. P. Dadios, F. Culibrina, J.-A. Magsumbol, J. D. Cruz, K. Sumage, M. T. Tan and M. Gomez, "Application of fuzzy logic in recognition of tomato fruit maturity in smart farming," in *2016 IEEE Region 10 Conference (TENCON)*, Singapore, 2016.

[6] R. R. Mhaski, P. B. Chopade and M. P. Dale, "Determination of ripeness and grading of tomato using image analysis on Raspberry Pi," Mathura, 2015.

[7] S. R. Rupanagudi, B. S. Ranjani, P. Nagaraj and V. G. Bhat, "A cost effective tomato maturity grading system using image processing for farmers," in *2014 International Conference on Contemporary Computing and Informatics*, 2014, 2014.

[8] W. D. N. Pacheco and F. R. J. López, "Tomato classification according to organoleptic maturity (coloration) using machine learning algorithms K-NN, MLP, and K-Means Clustering," in *2019 XXII Symposium on Image, Signal Processing and Artificial Vision*, Bucaramanga, Colombia, 2019.

[9] S. S. Deulkar and S. S. Barve, "An Automated Tomato Quality Grading using Clustering based Support Vector Machine," in *2018 3rd International Conference on Communication and Electronics Systems*, Coimbatore, India, 2018.

[10] X. Liang and Z. Z. Jiang, "Segmenting ripe tomato pictures based on the illumination irrelevant images," in *2013 Ninth International Conference on Natural Computation*, Shenyang, 2013.

[11] Y. A. Sari and S. Adinugroho, "Tomato ripeness clustering using 6-means algorithm based on v-channel otsu segmentation," in *2017 5th International Symposium on Computational and Business Intelligence*, Dubai, 2017.

[12] D. Yongjun and L. Jiying, "The application of Quantum-inspired ant colony algorithm in automatic segmentation of tomato image," in *2017 2nd International Conference on Image, Vision and Computing*, Chengdu, 2017.

[13] T. Mehra, V. Kumar and P. Gupta, "Maturity and disease detection in tomato using computer vision," in *2016 Fourth International Conference on Parallel, Distributed and Grid Computing*, Waknaghat, 2016.

[14] C. S. Hlaing and S. M. Maung Zaw, "Tomato Plant Diseases Classification Using Statistical Texture Feature and Color Feature," in *2018 IEEE/ACIS 17th International Conference on Computer and Information Science*, 2018, 2018.

[15] H. Sabrol and K. Satish, "Tomato plant disease classification in digital images using classification tree," in *2016 International Conference on Communication and Signal Processing*, Melmaruvathur, 2016.

[16] R. G. de Luna, E. P. Dadios and A. A. Bandala, "Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition," in *TENCON 2018 - 2018 IEEE Region 10 Conference*, Jeju, Korea, 2018.

[17] A. Hidayatuloh, M. Nursalman and E. Nugraha, "Identification of Tomato Plant Diseases by Leaf Image Using Squeezenet Model," in *2018 International Conference on Information Technology Systems and Innovation*, Bandung-Padang, Indonesia, 2018.

[18] H. Durmus, E. O. Günes and M. Kirci, "Disease detection on the leaves of the tomato plants by using deep learning," in *2017 6th International Conference on Agro-Geoinformatics*, Fairfax, VA, USA, 2017.

[19] P. T. A. Pranathi, K. SaiAshritha, N. B. Chittaragi and S. G. Koolagudi, "Tomato Leaf Disease Detection Using Convolutional Neural Networks," in *2018 Eleventh International Conference on Contemporary Computing*, Noida, 2018.

[20] C. Jae-Won, T. T. Trung, T. L. Huynh Thien, P. Geon-Soo, C. Van Dang and K. Jong-Wook, "A Nutrient Deficiency Prediction Method Using Deep Learning on Development of Tomato Fruits," in *2018 International Conference on Fuzzy Theory and Its Applications*, Daegu, South Korea, 2018.

[21] A. Milioto, P. Lottes and C. Stachniss, "Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs," in *IEEE International Conference on Robotics and Automation 2018*, 2018.

[22] T. Parhar, H. Baweja, M. Jenkins and G. Kantor, "A deep learning-based stalk grasping pipeline," in *IEEE International Conference on Robotics and Automation 2018*, 2018.

[23] C. Wu, R. Zeng, J. Pan, C. C. L. Wang and Y. Liu, "Plant Phenotyping by Deep-Learning-Based Planner for Multi-Robots," *IEEE Robotics and Automation Letters,* vol. 4, no. 4, pp. 3113 - 3120, 2019.