

AGRIVITA Journal of Agricultural Science

www.agrivita.ub.ac.id

Size Classification of Tomato Fruit Using Thresholding, Machine Learning and Deep Learning Techniques

Robert G. de Luna¹, Elmer P. Dadios, Argel A. Bandala and Ryan Rhay P. Vicerra

Gokongwei College of Engineering, De La Salle University, Philippines

ARTICLE INFO

ABSTRACT

Keywords: Image processing Kamatis Machine-vision *Solanum lycopersicum* Tomato fruit

Article History: Received: August 28, 2019 Accepted: September 30, 2019

*⁾ Corresponding author: E-mail: robert_g_deluna@dlsu.edu.ph

The size of tomato fruits is closely related to the market segment and price. Manual sorting in tomato is very dependent on human interpretation and thus, very prone to error. The study presents thresholding, machine learning and deep learning techniques in classifying the tomato as small, medium and large based from a single tomato fruit image implemented using Open CV libraries and Python programming. Tomato images with different sizes are gathered where features like area, perimeter and enclosed circle radius are extracted. The experiment shows that using thresholding, a classification accuracy of 85.83%, 65.83% and 80% was achieved for area, perimeter and enclosed circle radius, respectively. For machine learning, the training accuracy rates were recorded as 94.00%-95.00% for SVM, 97.50-92.50% for KNN and 90.33-92.50% for ANN. Comparison of models revealed that SVM is the most model without over fitting. The deep learning approach, regardless of the algorithm, produced low performances with 82.31%-78.21%-55.97% trainingvalidation-testing accuracy for VGG16, 48.17%-41.44%-37.64% for InceptionV3 and 56.05%-44.96%-22.78% for ResNet50 models. Comparative analysis showed that machine learning technique bested the performance of the thresholding and deep learning techniques in classifying the tomato fruit size in terms of accuracy performance.

INTRODUCTION

Tomato (*Solanum lycopersicum*), prevalently known as "Kamatis" in the Philippines, is a standout amongst the most important vegetable crops in the Philippines. It is a noteworthy vegetable crop that has accomplish tremendous popularity throughout the most recent century. It is developed in for all intents and purposes each nation of the world in open airfields, greenhouses and net houses. 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.

An article from Business Diary Philippines referred tomato as one of the most cultivated vegetables worldwide and is considered as a secondary crop especially in rice and corn-based farming systems (Lesaca, 2019). It is grown for both home and market in almost any community in the country. In 2016, the production value of tomatoes in the Philippines added up to roughly 3.31 billion Philippine pesos compared to 2.54 billion Philippine pesos in 2011 (Sanches, 2019).

During April to June 2018 period, the production of tomato expanded by 1.8% from 72.19 thousand metric tons in a similar quarter of the earlier year to 73.50 thousand metric tons this year. This is due to more farmers in Northern Mindanao were urged to plant because of high demand from Visayas and Metro Manila markets. Ilocos Region produced the largest tomato production this quarter at 27.52 thousand metric tons which was 37.4% of the national aggregate with Central Luzon came next with 11% followed by calabarzon with 9.8% (Phlippine Statistic Authority, 2019). This increase on tomato production is a decent open door for Filipino farmers to increase their wage. However, the greatest threat is the manner by which to give

ISSN: 0126-0537 Accredited First Grade by Ministry of Research, Technology and Higher Education of The Republic of Indonesia, Decree No: 30/E/KPT/2018

Cite this as: de Luna, R. G., Dadios, E. P., Bandala, A. A., & Vicerra, R. R. P. (2019). Size classification of tomato fruit using thresholding, machine learning and deep learning technique. *AGRIVITA Journal of Agricultural Science*, *41*(3), 586–596. http://doi.org/10.17503/agrivita.v41i3.2435

the desired volume and nature of tomatoes that the market demands.

In a conducted interview in one of the tomato farm owners in Batangas, sorting is essential in order for the farmers to offer the tomatoes with a price value that corresponds to its size and quality. In addition, the owner told the proponents that the sorting of the tomatoes keeps going for an entire day in the wake of harvesting, particularly when there is a huge amount of harvested tomatoes. The time they spent on sorting is around 8-10 hours for a 4 ha tomato farm. When it is harvest time, the quantity of boxes of tomatoes that they accumulate is least of 50 boxes and a greatest of 250 boxes.

Manual sorting is the most widely recognized strategy for sorting the fruits. The following issues developed in quality control carried out by people: high work costs, work fatigue, inconsistency and low precision (Arjenaki, Moghaddam, & Motlagh, 2013). Agricultural industries must work with greater accuracy, consistency and efficiency to fulfill market demands (Kaur & Gupta, 2017). Automated sorting has a better accuracy that takes into consideration more noteworthy esteem recuperation through better division of various evaluations of material. Automating the sorting procedure can decrease the incorrectness that happen when the farmers are doing the sorting. This study focuses on the automatic sorting of tomatoes depending upon size. The machine can manage a standard size for the sizes of the tomatoes being sorted and sort it in like manner not at all like when in manual, the size of the tomatoes might be founded just upon the judgment of the farmer that sorts them.

Machine vision provided an innovative way in the classification and identification of plants (de Luna et al., 2017), growth monitoring and assessment (Dimatira et al., 2016; Valenzuela et al., 2017). In this study, the researchers developed a classification system that will automatically classify the size of tomato fruits using several methods like thresholding, machine learning and deep learning. The system is implemented using Python 3.0 programming with libraries intended for image processing and machine/deep learning like OpenCV, scikit-learn, and Tensorflow.

A detailed overview of the process of fruit classification and grading has been reviewed by Naik & Patel (2017). Several feature extraction methods and machine learning models for that purpose were presented. When it comes to classifications, the most common techniques in machine learning are Support Vector Machines (SVM), Artificial Neural Networks (ANN) and K-Nearest Neighbors (K-NN). The study by Yossy, Pranata, Wijaya, Hermawan, & Budiharto (2017) focused on recognition of mango using ANN. The mango sortation system can sort mango accurately with 94% success percentage. Other study by Gatica, Best, Ceroni, & Lefranc (2013) focused on olive recognition using neural network which was divided into two stages, one for classification and another one for the selection of non-overlapping olives. The first stage has a classification accuracy of 97% while the second stage achieved 88.8% performance. SVM classifier was used in the study by Jana, Basak, & Parekh (2017) which yielded an accuracy of 83.33%.

On the other hand, deep learning also paved its way in applications for classification tasks. The study conducted by de Luna, Dadios, & Bandala (2018) used deep learning approach to identify which among the tomato diseases is present in tomato plants. The model produced a classification accuracy of 95.75%. The work of Ibrahim, Sabri, & Isa (2018) also utilized the use of the pre-trained architectures for the classification of maturity level of palm oil fruit bunch. The study proves the importance of transfer learning when creating a classification system because building a new CNN model from scratch requires a large amount of data whereas the pre-trained architectures are already trained in tremendous number of datasets.

In the study of Wan, Toudeshki, Tan, & Ehsani (2018) the principles and techniques in image preprocessing are explained deeper in order to highlight the object of interest for proper extraction of features. The process will always involve the determination of the region of interest to its background and isolating them. This will ensure that the calculated area only covers the fruits that are needed, and the backgrounds are not added. Works of Iqbal, Gopal, Sankaranarayanan, & Nair (2015) provided a simple method utilizing the radius, area and perimeter in grading the size of the sweetlime fruits into three categories. Several techniques were also presented by Fellegari & Navid (2011) and Thipakorn, Waranusast, & Riyamongkol (2017) in image processing method for measuring object size and volume, but with orange and eggs respectively. The study of Gongal, Karkee, & Amatya (2018) provided a solution to apple size estimation using fusion of 2D and 3D camera. Similar studies in apple

and orange was provided by Kalantari (2014) using image-processing technique. Results are validated by comparison to the volume determined using the water displacement method, and mean fruit diameter. Computer vision for banana size was proposed in the study of Hu, Dong, Malakar, Liu, & Jaganathan (2015). They created an automatic algorithm that determine three size indicators of banana, namely length, ventral straight length, and arc height. For the tomato, Yang, Kuang, & Mouazen (2011) used visible and near-infrared (VIS-NIR) spectroscopy for size estimation of tomato fruits of three cultivars. A partial least square regression (PLSR) was adopted to establish calibration models between fruit diameter and spectra and used back-propagation artificial neural network (BPANN) for the analysis.

In this study, proponents developed a machine-vision system that classify the size of the tomato fruit into small, medium, and large. The thresholding, machine learning, and deep learning techniques are created using the tomato images and its physical features like area, perimeter, and enclosed circle radius.

MATERIALS AND METHODS

Dataset Description

The dataset that was gathered and collected in some tomato farm in Batangas, Philippines last January to May 2019 which consists of 600 images of tomato plant which are categorized into different tomato size. Fig. 1 shows a sample data consisting of the three different sizes.

There are three classification of size (small, medium, and large) and each classification has

200 images distributed as 160 for the training and 40 for the testing. Maturity grade of tomatoes has no significance in the size classification so that gathered images can be green, turning, red and etc.

System Overview

The system overview is shown in Fig. 2. The input image is captured from the enclosed image capturing box using digital camera. The distance of the camera from the portion where the tomatoes will be placed is 15 inches wherein the 5 slots for the tomatoes is separated by 4 inches.

The input image from the acquisition system is a colored image. It is then subjected to image pre-processing using OpenCV libraries in Python which involves Red-Green-Blue (RGB), Hue-Saturation-Value (HSV) and Grayscale conversion. The processed image which contains 5 tomatoes is then segmented producing additional 5 images of a single tomato. Features like area, perimeter, and enclosed circle radius will be extracted from each of the cropped images thus producing a dataset for training consist of 480 rows by 4 columns where rows correspond to the number of samples while columns to the features and label. This dataset will be used in the development of classification model using the thresholding technique and machine learning. Meanwhile, cropped images are used for the deep learning approach. Separate dataset consist of 120 samples are generated for testing and validation. The performance of each models from different techniques will be evaluated from its classification accuracy to determine which approach will provide the best classification.



Fig. 1. Sample tomato image (A) small, (B) medium, (C) large



Fig. 2. Overview of the system

Development of the Size Classification Model

In the creation of size classification models, there will be three techniques to be used namely: thresholding, machine learning and deep learning.

Thresholding

The first approach is the thresholding techniques where the size classification will be based accordingly to the threshold value of some parameters. Image will be converted to binary and the contour of the tomatoes will be extracted using contour tracking based on Kalman filtering algorithm. A Kalman filter takes in information affected by some noise, error or uncertainty. The filter will take this imperfect information, sort out the useful parts and will reduce or remove those uncertainty or noise. After extracting the contour of the fruit, the radius, area and perimeter of the contour will be calculated. The generated 600 samples will be divided to 480 for the threshold determination and 120 for the testing of the model. Three models will be created under this approach using the threshold value generated from area, perimeter and radius.

Machine Learning

For the second approach, the proponents intended to adapt the use of machine learning algorithms. Machine learning models are used for different purposes, and one of them, is classifying sizes. Three popularly known models for classifying sizes, namely, SVM, K-NN, and ANN will be modelled.

Support Vector Machine (SVM), it is a controlled machine learning procedure which can be used for both classification and regression problems. The process starts by plotting each statistics item as a point in n-dimensional space where n is the number of features, with the value of each feature, being the value of a certain coordinate. Next is implementing an arrangement by finding the hyperplane that distinguishes the two classes very well.

K-Nearest Neighbors (K-NN) is a procedure that is uncertain which supplies all accessible case and classifies new cases based on a parallel measure. This procedure has been used in numerical approximation and pattern recognition since 1970's as a non-parametric technique. The algorithm assumes that same things exist in close proximity which means common things are near to each other.

Artificial Neural Networks (ANN) is one of the primary tools used in the world of Machine Learning. These are systems which are proposed to duplicate the way how humans learn. It is, also, a relatively unpolished electronic model created based on the neural structure of the brain. The neural networks consist of input and output layers and a group of hidden layers consisting of units that transform the input into something that the output layer can use.

In this machine learning approach, the generated dataset containing features and label will be divided into 80-20 data splitting. Each splitting of dataset will be used in the three machine learning models identified. Accuracy of classification will be the basis for selecting the best splitting type and machine learning model.

Deep Learning

Visual Geometry Group Network (VGGNet) architecture improved AlexNet by increasing the convolutional and pooling layers. It also uses 3 x 3 sized filters compare to AlexNet's 11 x 11. The required image size of the VGGNet is also 224 x 224 pixels. VGGNet placed second in the ILSVRC 2014 with a top 5 test error rate of 7.3%. It consists of 16 convolutional layers. Using more filters that are small sized made it significantly better among the previous architecture because it retains finer level properties of the image. It was designed by Karen Simonyan and Andrew Zisserman. Residual Neural Network (ResNet) took the simple network architecture of VGGNet but added more layers. It has a total of 152 layers of convolutional, pooling, and fully connecter layers. Normally, increasing the layers will make the network's accuracy lower because the image quality degrades. However, designers of ResNet made a residual block in each layer so the image gradient does not degrade. This made ResNet achieves a top-5 error rate of 3.6% which beats human level performance in the ILSVRC 2015.

GoogleNet or Inception is a much more complex network architecture compared to the other architecture because of the introduction of a module called Inception. GoogleNet's inception module basically lets the model decide the best size for each convolutional layer. GoogleNet was the winning entry in ILSVRC 2014 with a top 5 test error rate of 6.67% which was very close to human level performance that had 5.1% top 5 error rate.

RESULTS AND DISCUSSION

Results of Thresholding for Size Classification

For this approach, the dataset is divided into 80% for the training and 20% for the testing. The total number of datasets are 600, which are consisted by 480 datasets for training and 120 for testing. Fig. 3 shows the scatterplot of the generated training dataset that display the generated threshold values for area, perimeter, and enclosed circle radius.



Fig. 3. Scatter plot showing the threshold values (A) area, (B) perimeter, (C) enclosed circle radius

Variable	Mean	SE Mean	Std. Dev	Min.	Q1	Median	Q3	Max.
Area	69151	891	19515	40836	53689	62407	83046	116726
Perimeter	1300.3	8.29	256.8	916.2	1091.5	1258.2	1456.1	2203.7
Radius	160.07	1.05	23.03	119.11	140.37	155.93	175.09	225.31

Table 1. Descriptive statistics generated in Minitab

Table 2. Accuracy performance of size classification models using thresholding

Threshold	Size	No. of Samples	Correct	Incorrect	Accuracy (%)	Overall Accuracy (%)
Area	Small	40	32	8	80.00	
	Medium	40	36	4	90.00	85.83
	Large	40	35	5	87.50	
Perimeter	Small	40	24	16	60.00	
	Medium	40	25	15	62.50	65.83
	Large	40	30	10	75.00	
Enclosed Circle Radius	Small	40	27	13	67.50	
	Medium	40	36	4	90.00	80.00
	Large	40	33	7	82.50	



Fig. 4. Confusion matrix for thresholding models (A) area, (B) perimeter, (C) enclosed circle radius

These threshold values are generated using the statistical tool called Minitab. Table 1 summarized the descriptive statistics generated with Q1 represent the small-to-medium threshold and Q3 for the medium-to-large threshold.

Three classification models are created using the determined threshold values. Each model is evaluated in terms of accuracy using the separate 120 testing datasets. Fig. 4 reflected the confusion matrix showing the number of tomatoes classified by the model according to its size. Table 2 summarized the accuracy percentage of the classification. From the results, Area Thresholding produced the highest overall accuracy among the three threshold parameters with a value of 85.83%.

Results of Machine Learning for Size Classification

There are three machine learning models used under this approach, namely, Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and the Artificial Neural Network (ANN). The same dataset was used and a standard 80%-20% training-testing data splitting was used. The training accuracy

performance was evaluated using 10-fold stratified cross-validation per model. Optimization was also performed using GridSearchCV in finding optimal values for hyper-parameter tuning. Models were tested the using the testing dataset to establish the best machine learning model. Table 3 summarized the accuracy percentage of every model in different dataset splitting both during training and testing.

There is significant increase in all machine learning models once optimization using GridSearchCv is performed as compared to the performance using the default value of parameters. Result also depicted that regardless of whatever combinations of data splitting is used, every model has consistent performance in the training phase.

In Fig. 5, testing accuracy revealed that the performance of all the models using independent dataset for testing is high also consistent with its

performance during the training phase. Comparison of the three model's performance favored SVM as the best optimized machine learning model with training performance of 95.00% accuracy. The said model is best in a sense that the model is consistent in training and testing as supported by its accuracy performance. The SVM model therefore is not over fitted, an indication of a good model.

Results of Deep Learning for Size Classification

There are three deep learning models implemented under this approach, namely, VGG16, InceptionV3, and the ResNet50. All models used a batch size of 32, learning rate of 0.00001 and with "adam" as an optimizer. Unlike with the first two methods, the deep learning model utilized the use of the images itself and not from the extracted features dataset.

Table 3. Accuracy performance of size classification models using machine learning

	Splitting		Training		Testing Using Optimized Model				
Model		No. of Samples	Default Parameter Accuracy (%)	Optimized Parameter Accuracy (%)	No. of Samples	Correct	Incorrect	Accuracy (%)	
SVM	80-20	480	90.16	94	120	114	6	95.00	
KNN	80-20	480	88.33	97.5	120	111	9	92.50	
ANN	80-20	480	89.5	90.33	120	111	9	92.50	



Training and Testing Accuracy



Copyright © 2019 Universitas Brawijaya



Robert G. de Luna et al: Size Classification of Tomato Fruit

Fig. 6. Accuracy plot of the three deep learning models during training and testing under 100 epochs (A) VGG16, (B) InceptionV3, (C) ResNet50



Fig. 7. Loss plot of the three deep learning models during training and testing under 100 epochs (A) VGG16, (B) InceptionV3, (C) ResNet50

For all models, the gathered 600 tomato images are divided into 80% to training and 20% for the validation. In the training phase, the accuracy at regular intervals (for every epoch) is generated. Separate 320 images are used to evaluate the accuracy of the model in actual implementation. Fig. 6 and Fig. 7 depicted the performance of the three models showing the accuracy plot and loss plot respectively in every epoch during the training and testing. It can be shown that the VGG16 model performs well compared to the other two deep learning models as the training and testing accuracy plots are increasing exponentially.



Robert G. de Luna et al: Size Classification of Tomato Fruit

Fig. 8. Confusion matrix for deep learning models (A) VGG16, (B) InceptionV3, (C) ResNet50

Table 4.	Accuracy	performance	of size	classification	models	using deep	learning
----------	----------	-------------	---------	----------------	--------	------------	----------

	Training		Validation		Testing											
Model	No. of Samples	Average Accuracy (%)	No. of Sam- ples	Aver- age Ac- curacy (%)	Size	No. of Sam- ples	Cor- rect	Incor- rect	Accuracy (%)	Over-all Accu- racy (%)						
					Small	120	45	75	37.50							
VGG16	480	82.31	120	78.21	Medium	120	104	16	86.67	55.97						
					Large	80	35	45	43.75							
					Small	120	95	25	79.17							
InceptionV3	480	48.17	120	41.44	Medium	120	0	120	0.00	37.64						
															Large	80
ResNet50		56.05	120	44.96	Small	120	5	115	4.17							
	480				Medium	120	41	79	34.17	22.78						
					Large	80	24	56	30.00							

The three trained deep learning models are evaluated again in actual set up using separate 320 tomato images. Fig. 8 and Table 4 summarized the confusion matrix and the accuracy percentage of every model.

Seen in Table 4 is the listing of the training, validation, and testing accuracy. Even though VGG16 provided good performance in the training and validation, testing accuracy that uses independent 320 images provided only 55.97 %

in classifying tomato in three different sizes. In general, all three models used in deep learning were not able to meet the desired 80 % and above accuracy.

CONCLUSION

Different algorithms for classifying the tomato size into small, medium and large from an image, using the combination of image processing, thresholding, machine learning and deep learning

techniques, are proposed in this paper. Images of tomato fruits categorized in three sizes were gathered using a defined image capturing system. Three geometrical features were extracted, namely, area, perimeter and enclosed circle radius. These are compiled in a csv file and used in threshold and machine learning modelling. Deep learning on the other hand, utilized the gathered images and subjected to several image pre-processing before the training. Results shown that in the thresholding method, using the area as the parameter yielded the highest percentage of accuracy of 85.83 %, while for machine learning method, KNN registered the highest accuracy of 97.50% in the training phase and 92.50% in testing but it is the SVM to be considered as the most stable model with 94% and 95% training-testing accuracy, an indication of a good fit for a model. For the deep learning, no models reached the target 80% and above performance with only 55.97% as the highest for VGG16. It is therefore concluded that machine learning model, specifically the SVM, is the best model to implement in the classification of tomato size. Although deep learning provided significant contributions in classification tasks, application of which is limited to size classification as modelling results shown poor performance regardless of what architecture to use.

ACKNOWLEDGEMENT

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

REFERENCES

- Arjenaki, O. O., Moghaddam, P. A., & Motlagh, A. M. (2013). Online tomato sorting based on shape, maturity, size, and surface defects using machine vision. *Turkish Journal of Agriculture* and Forestry, 37, 62–68. https://doi.org/10.3906/ tar-1201-10
- de Luna, R. G., Baldovino, R. G., Cotoco, E. A., de Ocampo, A. L. P., Valenzuela, I. C., Culaba, A. B., & Gokongwei, E. P. D. (2017). Identification of philippine herbal medicine plant leaf using artificial neural network. 2017IEEE

9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 1–8. https://doi. org/10.1109/HNICEM.2017.8269470

- de Luna, R. G., Dadios, E. P., & Bandala, A. A. (2018). Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition. *TENCON 2018-2018 IEEE Region 10 Conference*, 1414–1419. https://doi.org/10.1109/TENCON.2018.8650088
- Dimatira, J. B. U., Dadios, E. P., Culibrina, F., Magsumbol, J., Cruz, J. D., Sumage, K., ... Gomez, M. (2016). Application of fuzzy logic in recognition of tomato fruit maturity in smart farming. 2016 IEEE Region 10 Conference (TENCON), 2031–2035. https://doi.org/10.1109/TENCON.2016.7848382
- Fellegari, R., & Navid, H. (2011). Determining the orange volume using image processing. Interbational Conference on Food Engineering and Biotechnoly (ICFEB 2011), 180–184. Retrieved from http://ipcbee.com/vol9/34-B10018.pdf
- Gatica, G., Best, S., Ceroni, J., & Lefranc, G. (2013). Olive fruits recognition using neural networks. *Procedia Computer Science*, *17*, 412–419. https://doi.org/10.1016/j.procs.2013.05.053
- Gongal, A., Karkee, M., & Amatya, S. (2018). Apple fruit size estimation using a 3D machine vision system. *Information Processing in Agriculture*, 5(4), 498–503. https://doi.org/10.1016/j. inpa.2018.06.002
- Hu, M.-H., Dong, Q.-L., Malakar, P. K., Liu, B.-L., & Jaganathan, G. K. (2015). Determining banana size based on computer vision. *International Journal of Food Properties*, 18(3), 508–520. https://doi.org/10.1080/10942912.2013.833223
- Ibrahim, Z., Sabri, N., & Isa, D. (2018). Palm Oil Fresh Fruit Bunch Ripeness Grading Recognition Using Convolutional Neural Network. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(3–2), 109–113. Retrieved from http://journal.utem.edu.my/index. php/jtec/article/view/4720
- Iqbal, S. M., Gopal, A., Sankaranarayanan, P. E., & Nair, A. B. (2015). Estimation of size and shape of citrus fruits using image processing for automatic grading. 2015 3rd International Conference on Signal Processing, Communication and Networking (ICSCN), 1–8. https://doi. org/10.1109/ICSCN.2015.7219859

- Jana, S., Basak, S., & Parekh, R. (2017). Automatic fruit recognition from natural images using color and texture features. 2017 Devices for Integrated Circuit (DevIC), 620–624. https://doi. org/10.1109/DEVIC.2017.8074025
- Kalantari, D. (2014). Size determination of apple and orange fruits using the image processing technique. Advances in Agriculture, Sciences and Engineering Research, 4(2), 1537–1543. Retrieved from https://www.cabdirect.org/ cabdirect/abstract/20143381993
- Kaur, K., & Gupta, O. (2017). A machine learning approach to determine maturity stages of tomatoes. Oriental Journal of Computer Science and Technology, 10(3), 683–690. https://doi. org/10.13005/ojcst/10.03.19
- Lesaca, P. R. A. (2019). There is gold in tomatoes. Retrieved from Business Diary Philippines website: https://businessdiary.com.ph/6293/ gold-tomatoes/?fbclid=IwAR0K5ezJ_ T 6 W e x L m m i Q z Y y f X 6 0 t -5sbWs8VGdUScO1hpbSmALz8MYDOYT00
- Naik, S., & Patel, B. (2017). Machine Vision based Fruit Classification and Grading-A Review. International Journal of Computer Applications, 170(9), 22–34. Retrieved from https://pdfs.semanticscholar.org/4aa0/ a0cdfe036b512b9c6a0e9f34c7f47382a27d.pdf
- Phlippine Statistic Authority. (2019). *Major Vegetables and Root Crops Quarterly Bulletin, April-June 2019.* Retrieved from https://psa.gov.ph/vegetableroot-crops-main/tomato
- Sanches, M. J. (2019). Production value of tomatoes in the Philippines from 2011 to 2018 (in billion Philippine pesos). Retrieved from https://www. statista.com/statistics/752489/philippines-

tomato-production-value/?fbclid=lwAR1I0IDFgw wFcMsWB5AANdhhC732sgcMf7wYC5CtOFi0w RkIVQjzQ61HSeY

- Thipakorn, J., Waranusast, R., & Riyamongkol, P. (2017). Egg weight prediction and egg size classification using image processing and machine learning. 2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 477–480. https://doi.org/10.1109/ ECTICon.2017.8096278
- Valenzuela, I. C., Puno, J. C. V, Bandala, A. A., Baldovino, R. G., de Luna, R. G., De Ocampo, A. L., ... Dadios, E. P. (2017). Quality assessment of lettuce using artificial neural network. 2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 1–5. https://doi. org/10.1109/HNICEM.2017.8269506
- Wan, P., Toudeshki, A., Tan, H., & Ehsani, R. (2018). A methodology for fresh tomato maturity detection using computer vision. *Computers and Electronics in Agriculture*, 146, 43–50. https:// doi.org/10.1016/j.compag.2018.01.011
- Yang, H. Q., Kuang, B. Y., & Mouazen, A. M. (2011). Size estimation of tomato fruits based on spectroscopic analysis. Advanced Materials Research, 225–226, 1254–1257. https://doi. org/10.4028/www.scientific.net/AMR.225-226.1254
- Yossy, E. H., Pranata, J., Wijaya, T., Hermawan, H., & Budiharto, W. (2017). Mango fruit sortation system using neural network and computer vision. *Procedia Computer Science*, *116*, 596– 603. https://doi.org/10.1016/j.procs.2017.10.013