Real-Time Banana Ripeness Detection and Classification using YOLOv8

Renann G. Baldovino*, Raphael Antoine U. Lim, Patrick Reylie R. Salvador, Euri Andre P. Tiamzon
Department of Manufacturing Engineering and Management (DMEM)
Gokongwei College of Engineering (GCOE), De La Salle University (DLSU)
2401 Taft Avenue, 0922 Manila, Philippines
Corresponding author: *renann.baldovino@dlsu.edu.ph*

2024 9th International Conference on Mechatronics Engineering (ICOM) | 979-8-3503-4978-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICOM61675.2024.10652438

Abstract—**Due to farmers in the Philippines having to utilize manual methods to determine the grade and quality of freshly harvested bananas, post-harvest classification and segregation accounts for between 3% and 30% of food waste that is linked to bananas. This study focused on real-time banana object recognition and classification via a webcam using neural networks. The principles of computer vision and object recognition and how they relate to fast algorithms like CNN and YOLO are a major emphasis of this study. The researchers used the YOLOv8 algorithm to develop their model since the appropriate selection of the YOLO model determines both the accuracy and processing speed of the photographs. The method is intended to produce a detected and classified level of banana ripeness and be displayed through a live feed from the webcam. Using the confusion matrix, the model was able to detect and categorize objects with an accuracy of about 90% across all classes on the testing dataset of static images. For real-time detection of the model, samples were placed against a blue background given that the yellow and green colors of banana samples are a complementary pair according to the sRGB color wheel. Banana samples were tested both in bunches and as individual samples. However, the model was only 80% accurate during real-time individual sample testing.**

Keywords—object detection, classification, banana ripeness, postharvest, YOLO, real-time

I. Introduction

Bananas are currently one of the top agricultural crops that are being produced, having harvested more than 117M tonnes in 2019. Since bananas turn brown and produce spots once they approach the overripe stage, many people tend to throw them out even if the banana may be used for other purposes [1]. This means that bananas are factors of the largest food wastage because of their quick overripeness and the common misunderstanding in their ripening stages. The Philippines, being one of the top countries that export bananas, has managed to produce approximately 9.01M MT in 2022. Thus, the country also shares this burden of banana wastage and improper food management [2].

In the Philippines, there are three main types of bananas currently being grown. The first is the Cavendish variety, making up half of the banana production and being the primary export type. Lakatan bananas, a native variant, make up 11% of the production and are mainly used in desserts and beverages. Lastly, the saba variety contributes 37% to the total banana production and is a favored banana for cooking and in Filipino cuisine [3].

Post-harvest classification and segregation in the Philippines ranges from 3 to 30% of food wastages attributed to bananas. This is mainly due to the manual processes that farmers use to identify the grade and quality of newly harvested bananas [4]. Additionally, multiple post-harvest losses are attributed to transportation and handling by various parties before reaching the end-users. A study reported a 15.45% and 15.58% post-harvest loss of local and export banana products respectively. About 8 to 24% of the total post-harvest losses were related to quality defects and improper handling due to cuts and bruising of the fruit. However, the majority were due to the time element of transportation and segregation done by middle parties such as consolidators and wholesalers. Post-harvest losses at these stages are attributed to waste loss since most of the bananas are transported via open truck which are exposed to sunlight and can hasten the ripening process [5]. To determine the banana ripeness, a general ripening guide is used to visually determine and classify the state of the banana. The peel color is utilized to determine the ripeness (see Fig. 1).

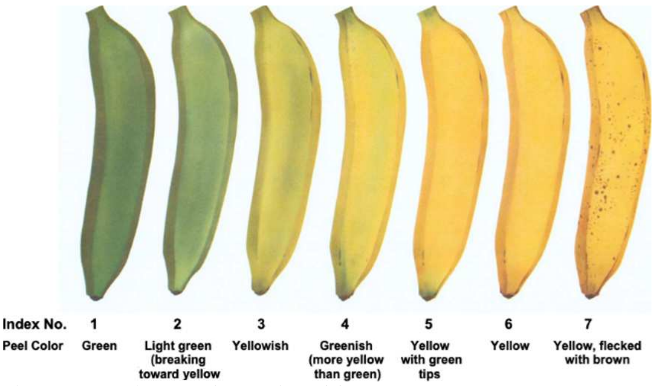


Fig. 1 Banana ripeness chart. Adapted from [6].

Notable research [7-10] have utilized a version of the banana ripeness chart and its stages for classification. The banana's spotting, or more accurately called senescent spotting, is the change of color of the banana as evidenced by the increasing number of brown spots which can be utilized to detect an overripe banana. This visual assessment can be utilized to better determine and classify the banana [11].

Fruit object detection is a rapidly expanding topic that uses sophisticated deep learning algorithms to discover and identify fruits automatically in a variety of settings. Deep convolutional neural networks (CNN) provide fruit detection and recognition with state-of-the-art accuracy [12]. These methods work with a broad range of fruits, including tomatoes, apples, mangoes, and more. CNN-based systems such as single shot multibox detector (SSD) and You Only Look Once (YOLO) have proven to be quite effective in obtaining high detection speed and accuracy. As a result, they work effectively in real-time applications like harvesting automation [13-14]. Moreover, the development of specialized systems such as DeepFruits has demonstrated the potential for accurate, quick, and efficient fruit recognition using deep learning techniques [15]. With that, the study aims to use neural networks to implement real time banana objects and ripeness detection and classification.

II. Theoretical and Conceptual Framework

A. *Theoretical Framework*

The study focuses on the theories of computer vision and object detection as it deals with high-speed algorithms such as YOLO and CNN. Of the many applications, these algorithms are relevant for banana ripeness detection as there are clear correlations between the colorspace of the fruit and the ripeness level. Older techniques of ripeness classification involve the use of RGB and HSV color spaces through support vector machines (SVM) which was a widely used algorithm for classification and regression analysis. However, as real-time detection became more needed in many disciplines, a faster and more reliable method was developed in the form of the YOLO algorithm.

B. *Conceptual Framework*

Figure 2 shows the different concepts crucial in the development of the study.

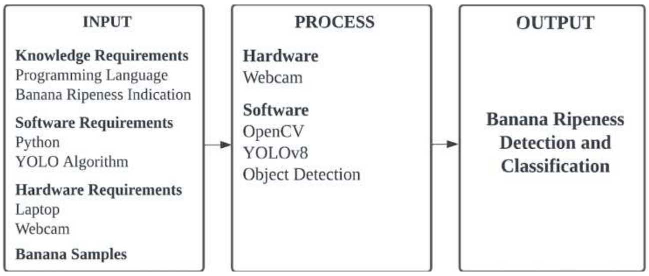


Fig. 2. Conceptual framework of the study

It organizes all the necessary inputs gained from the review of related literature and lists the respective processes that these inputs need to undergo to achieve the output program. The study will be carried out entirely in the Python language due to its compatibility with YOLO, specifically the v8 which is the latest iteration of the widely used object detection method in real-time applications. It was developed by Ultralytics which is the company behind YOLOv5 which introduced a breakthrough architecture for smaller but accurate results. Due to several modifications to its architecture, YOLOv8 is now the fastest algorithm for real-time image processing and object detection.

III. Proposed Methodology

A. *Methodological Framework*

Since object detection requires heavy training to have an accurate model, a good dataset with relatively many data points including proper labeling is crucial to the success of the model. Additionally, the proper choice of the YOLO model is another factor that will dictate the speed at which the model processes images as well as the accuracy. Figure 3 describes the methodological framework of the study which includes the research and acquisition of the banana dataset, model selection and training of the model, testing the model and obtaining the confusion matrix, finally the integration of the model with a webcam for real-time detection.

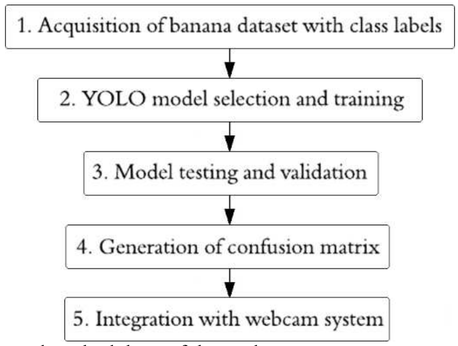


Fig. 3. Proposed methodology of the study

B. *Materials and Components*

In creating a system that indicates the ripeness of a banana automatically, knowledge of visual–based detection of banana ripeness, and proficiency in a programming language is required. The software requirements for this system are the Python programming language and importing the OpenCV library, while hardware components include a computer running Windows operating system and an external webcam.

The training of the model was done in Google Colab as it has access to more powerful systems and can train the model much faster. However, once the model has been trained and extracted, it was transferred to a Python script created locally for webcam integration since Colab does not support webcam functionality at the time of this study.

To verify the model accuracy, a few pre-labelled pictures were fed through the model and the classes obtained which generated a confusion matrix. However, since this applies to static images, the real-time capabilities of the model were not tested. Thus, a few samples of Lakatan banana were used to check the accuracy of the model through real-time detection.

The camera used for real-time detection was a Logitech C270 webcam. Additionally, the samples of Lakatan banana were placed against a dark blue background as the green and yellow colors associated with the banana are direct complements according to the sRGB color wheel (see Fig. 4).

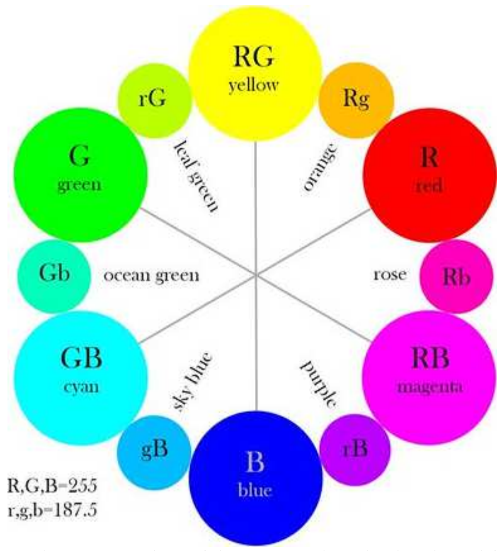


Fig. 4. Complementary colors of the sRGB color wheel. Adapted from [17].

Research has shown that specific colors which exist opposite the color wheel produce highly contrasting differences between their hues [21]. This would cause the camera and vision system to have better detection of bananas and the colors associated with ripeness levels. It is important to note that sRGB refers to the colorspace commonly used in monitors, printers, and displays not to be confused with the color wheel commonly used in artistic applications.

C. Acquisition of Banana Dataset

An open-source data set for banana ripeness entitled "Fruit Ripening Process Dataset" can be accessed through the Roboflow website under the Fruit Ripening category [16]. There are 6 data classifiers namely: unripe, rotten, ripe, overripe, freshunripe, and freshripe were used akin to the banana ripening guide discussed from [6]. The data set contains 7,546 images with 15,053 annotations wherein it averages around 2 annotations per image. These classifiers can be related through visual indicators such as colors and the presence of brown spots. Unripe bananas display characteristics as the first two bananas in the figure, fresh unripe bananas should exhibit yellowish or greenish colors, fresh ripe bananas exhibiting yellow with green tips, ripe bananas have an overall yellow color or a bit flecked with brown spots, overripe bananas have more noticeable brown spots, and lastly rotten bananas should have a dark brown color, with little to no yellow color.

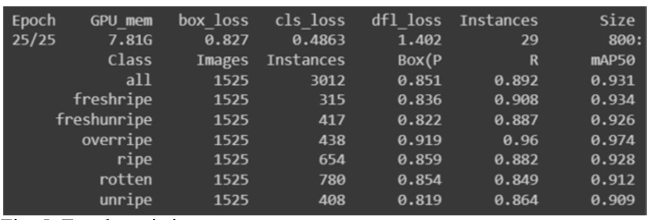
D. Model Selection and Training

The initial model used was the YOLOv8-S which was trained with 25 epochs through the banana-ripeness data set using Colab. The small variant of the algorithm was chosen based on suggestions from the acquired dataset. Additionally, it provided the best balance between complexity and accuracy as shown in the summary of variants in Table 1.

TABLE I. YOLOV8 MODEL VARIANTS. ADAPTED FROM [18].

Variant	Size	Parameters	Speed	Accuracy
YOLOv8-XS	2.4 MB	4.7M	40 FPS	41.1% mAP
YOLOv8-S	3.4 MB	7.9M	30 FPS	46.0% mAP
YOLOv8-M	8.1 MB	20.0M	20 FPS	51.1% mAP
YOLOv8-L	18.0 MB	53.6M	15 FPS	55.3% mAP
YOLOv8-X	46.2 MB	111.1M	10 FPS	61.2% mAP

If the dataset were to be trained on larger variants, it would take an exponentially longer time to train and produce a slightly better accuracy. The model took approximately 4 hrs to finish training, as shown in Figure 5, with an average training time per epoch of 9 min.



Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
25/25	7.81G	0.827	0.4863	1.402	29	800:
	Class	Images	Instances	Box(P	R	mAP50
	all	1525	3012	0.851	0.892	0.931
	freshripe	1525	315	0.836	0.908	0.934
	freshunripe	1525	417	0.822	0.887	0.926
	overripe	1525	438	0.919	0.96	0.974
	ripe	1525	654	0.859	0.882	0.928
	rotten	1525	780	0.854	0.849	0.912
	unripe	1525	408	0.819	0.864	0.909

Fig. 5. Epoch statistics.

E. Testing and Evaluation

The test-split dataset consisted of 757 images wherein the trained model was performed on and a confusion matrix was obtained. In addition to the accuracy, the model speed was also obtained with a total computing speed of 0.7 ms for the pre-processing, 12.8 ms for the inference, and 1.6 ms for the post-processing per image. Once the model has been fully trained, it was saved as a *.pt* file for export and can be used for other applications including real-time detection [19, 20].

F. Integration with Webcam

The Python script starts by calling the video feed from the webcam. From there, the model obtained from the previous procedures will be imported and the class names specified. After all classes have been loaded and the model saved into the script, a while loop will be called which will continuously run until the user presses the '*q*' button on the keyboard which will terminate the program.

Inside the loop, the program continuously reads the video feed from the webcam as multiple images depending on the frame rate of the camera. For each of those images, it passes through the model and labels the photo based on the predicted class. This is done by creating bounding boxes around the object of interest and overlaying text containing the class name and the confidence level. After the proposed algorithm is performed, the system should output a detected and classified banana ripeness that continuously updates according to the framerate of the camera for real-time detection and classification.

Once the program has been initiated, a window will display the live feed from the webcam. The first experiment consisted of placing two bunches of bananas, one unripe and another ripe in front of the camera. After which individual bananas were then tested which consisted of 5 unripe samples, 5 ripe samples, and 5 overripe samples for a total of 15 individual samples of banana.

IV. RESULTS AND DISCUSSION

A. Confusion Matrix of Test Dataset

Upon testing the model, this yielded the following confusion matrix shown in Figure 6.

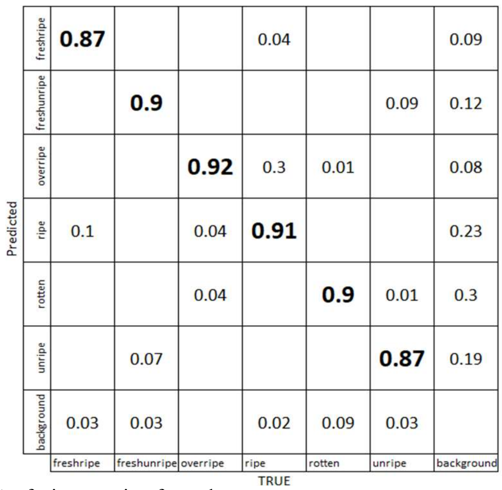


Fig. 6. Confusion matrix of test dataset

The model achieved an 89.5% testing accuracy across all classes with the highest wrong prediction was that of the "rotten" prediction misclassifying what should be just image background. The model also had difficulty differentiating ripe and overripe samples.

B. Webcam Results and Real-Time Model Accuracy

To conduct initial testing, the whole bunch of bananas was placed in front of the webcam to determine its multiobject, real-time classification. Figure 7 showcases the model's performance on the webcam feed, notably correctly predicting the entire bunch as unripe, with the bounding boxes placed at reasonable distances from the bananas.

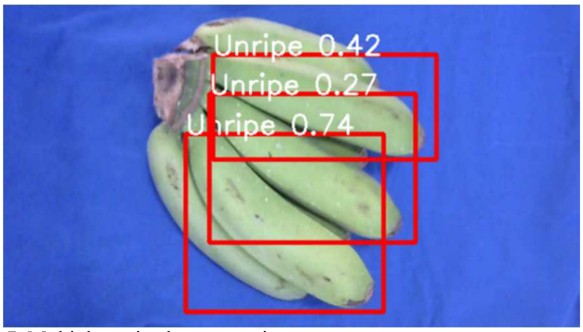


Fig. 7. Multiple unripe banana testing

Previous experiments tried utilizing simply holding up the banana to the camera with no consistent background. This caused the YOLO to have difficulties detecting the presence of a banana in the video feed. By placing banana samples against a consisted blue background, the system more accurately determined the sample and omitted any possible errors of falsely identifying objects in the background. Also, the program is capable of detecting multiple classifications of ripeness levels simultaneously as depicted in Figure 8.

Fig. 8. Multiple unripe and ripe banana testing

However, as the sample increases, the computing power was not sufficient to label all samples wherein two overripe samples were not detected, and one sample was misclassified as an unripe banana. Afterwards, the bananas were separated into individual samples which were then brought in front of the webcam one after the other for individual testing. Figure 9 shows the samples used, consisting of 5 unripe samples, 5 ripe samples, and 5 overripe samples.


Fig. 9. Individual samples utilized in the experimentation: unripe (top left), ripe (bottom left), and overripe (right) samples

The presence of green color throughout the individual samples were used to gauge the current ripeness level of the banana. Overripe samples were left for several days until browning and speckling was observed.

Samples 1 to 5 were determined to be unripe due to their green color and firmness. However, the program failed to correctly classify one sample. This is consistent with the testing confusion matrix which showed that the model had a low accuracy for unripe samples. The next 5 samples were perfectly classified as ripe due to the samples having an almost yellow color throughout except for the tips which is a common occurrence in freshly ripe bananas. The program correctly classified 3 out of the 5 overripe samples which may be also due to the dataset used for training as multiple bananas can be labeled as ripe despite the presence of speckling in the samples.

The following tables summarize the results from the experimentation. Table 2 displays the individual sample results along with the respective confidence levels of classification as determined by the YOLOv8 algorithm. Table 3 on the other hand, indicates the ability of the model to correctly classify each sample. Since there was no alternative technique to determine the subtle differences between ripeness levels (i.e. between fresh ripe and fresh unripe), if the model correctly identified the general ripeness levels of unripe, ripe, and overripe correctly as indicated in Figure 9, it would still be considered as a successful identification.

TABLE II. RESULTS FROM INDIVIDUAL SAMPLE TESTING

Sample No.	Model Classification	Confidence Level
1	RIPE	0.78
2	UNRIPE	0.63
3	UNRIPE	0.7
4	UNRIPE	0.43
5	UNRIPE	0.65
6	RIPE	0.78
7	RIPE	0.67
8	RIPE	0.49
9	RIPE	0.78
10	RIPE	0.49
11	ROTTEN	0.83
12	RIPE	0.42
13	ROTTEN	0.71
14	ROTTEN	0.84
15	RIPE	0.81

TABLE III. REAL-TIME CLASSIFICATION OF INDIVIDUAL SAMPLES

Sample	Actual Classification	Model Classification	Remarks
1	UNRIPE	RIPE	FAIL
2	UNRIPE	UNRIPE	PASS
3	UNRIPE	UNRIPE	PASS
4	UNRIPE	UNRIPE	PASS
5	UNRIPE	UNRIPE	PASS
6	RIPE	RIPE	PASS
7	RIPE	RIPE	PASS
8	RIPE	RIPE	PASS
9	RIPE	RIPE	PASS
10	RIPE	RIPE	PASS
11	OVERRIPE	ROTTEN	PASS
12	OVERRIPE	RIPE	FAIL
13	OVERRIPE	ROTTEN	PASS
14	OVERRIPE	ROTTEN	PASS
15	OVERRIPE	RIPE	FAIL

V. Conclusion and Recommendations

The study was able to use YOLOv8 for training a banana-ripeness data set for banana detection and classification. The model was able to accurately detect and classify at around 90% across all classifications through the confusion matrix of the testing dataset but in real-time testing the model was only 80% accurate. It is then recommended that further improvements should be done on the model such as fine tuning hyperparameters or changing the YOLO model used.

Integration of agriculture and neural networks will continue to evolve and improve in the implementation as computing costs lessen and algorithms become even more efficient. The research and model implemented can be utilized by banana plantations and small farms alike in streamlining their process.

References

[1] Emily, "Browning bananas create nearly 50 million tonnes of food waste," *Open Access Government*, May 11, 2022. https://www.openaccessgovernment.org/browning-banana-food-waste-household/135392/ (accessed Dec. 04, 2023).

[2] "Philippines: production volume of bananas 2022 | Statista," Statista, 2022. https://www.statista.com/statistics/751577/philippines-banana-production/#:~:text=In%202022%2C%20the%20volume%20of,reached%20its%20peak%20in%202018. (accessed Dec. 04, 2023).

[3] "Banana – Industry Strategic Science and Technology Plans (ISPs) Platform," Dost.gov.ph, 2014. https://ispweb.pcaarrd.dost.gov.ph/isp-commodities/banana/ (accessed Dec. 04, 2023).

[4] Eduardo Jr Piedad and June Anne Caladcad, "Post-harvested Musa acuminata Banana Tiers Dataset," *Data in Brief*, vol. 46, pp. 108856–108856, Feb. 2023, doi: https://doi.org/10.1016/j.dib.2022.108856.

[5] "Assessment of the postharvest handling systes and losses of Cardava banana in the Philippines," *ResearchGate*, 2018, doi: https://doi.org/10.13140//RG.2.2.19939.40489.

[6] U.S. Department of Agriculture, "Bananas visual aid," Rertieved from https://www.ams.usda.gov.

[7] T. Ringer and M. Blanke, "Non-invasive, real-time in-situ techniques to determine the ripening stage of banana," *Journal of Food Measurement and Characterization*, vol. 15, no. 5, pp. 4426–4437, 2021. doi: 10.1007/s11694-021-01009-2.

[8] F. M. A. Mazen and A. A. Nashat, "Ripeness classification of bananas using an artificial neural network," *Arabian Journal for Science and Engineering*, 2019, doi: 10.1007/s13369-018-03695-5.

[9] M. Santoyo-Mora, A. Sancen-Plaza, A. Espinosa-Calderon, A. I. Barranco-Gutierrez, A. I. and J. Prado-Olivarez, J. "Nondestructive quantification of the ripening process *in banana (Musa AAB Simmonds) using multispectral imaging." Journal of Sensors*, 2019, pp. 1–12, doi: 10.1155/2019/6742896.

[10] N. Saranya, K. Srinivasan, and S. K. P. Kumar, "Banana ripeness stage identification: a deep learning approach," *Journal of Ambient Intelligence and Humanized Computing*, 2021. doi: 10.1007/s12652-021-03267-w

[11] R. Quevedo, F. Mendoza, J. M. Aguilera, J. Chanona, and G. Gutiérrez-López, "*Determination of senescent spotting in banana (Musa cavendish) using fractal texture Fourier image*," *Journal of Food Engineering*, vol. 84, no. 4, pp. 509–515, 2008. doi: 10.1016/j.jfoodeng.2007.06.013.

[12] F. Xiao, H. Wang, Y. Xu, and R. Zhang, "Fruit Detection and Recognition Based on Deep Learning for Automatic Harvesting: An Overview and Review," Agronomy, vol. 13, no. 6, p. 1625, Jun. 2023. [Online]. Available: https://www.mdpi.com/2073-4395/13/6/1625.

[13] K. Bresilla, G. D. Perulli, A. Boini, B. Morandi, L. C. Grappadelli, and L. Manfrini, "Single-shot convolution neural networks for real-time fruit detection within the tree," *Frontiers in Plant Science*, vol. 10, p. 611, May 2019. Retrieved from https://www.frontiersin.org

[14] W. Jia, Y. Xu, Y. Lu, X. Yin, N. Pan, R. Jiang, and X. Ge, "An accurate green fruits detection method based on optimized YOLOX-m," *Front. Plant Sci*., vol. 14, p. 1187734, May 2023. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10200941/

[15] I. Sa, Z. Ge, F. Dayoub, B. Upcroft, T. Perez, and C. McCool, "DeepFruits: a fruit detection system using deep neural networks," *Sensors (Basel)*, vol. 16, no. 8, p. 1222, Aug. 2016. [Online]. Available: https://www.mdpi.com/1424-8220/16/8/1222

[16] F. Ripening, 'Fruit Ripening Process Dataset', Roboflow Universe. Roboflow, Oct-2022.

[17] M. S. Abeln, "A visually-uniform digital color wheel," Blogspot.com, 2015. https://therefractedlight.blogspot.com

[18] M. Abdullah, "YOLO working principle, difference between its ddifferent variants and versions," *Medium*, Oct. 08, 2023. https://medium.com.

[19] M. A. A. Felipe, T. V. Olegario, N. T. Bugtai and R. G. Baldovino, "Vision-based liquid level detection in amber glass bottles using OpenCV," 2019 7th International Conference on Robot Intelligence Technology and Applications (RiTA), Daejeon, Korea (South), 2019, pp. 148-152, doi: 10.1109/RITAPP.2019.8932807.

[20] R. B. G. Luta, A. C. L. Ong, S. J. C. Lao, R. G. Baldovino, N. T. Bugtai and E. P. Dadios, "A noncontact pH level sensing indicator using computer vision and knowledge-based systems," 2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Manila, Philippines, 2017, pp. 1-5, doi: 10.1109/HNICEM.2017.8269474.

[21] R. W. Pridmore, "Complementary colors: A literature review," *Color Research & Application*, vol. 46, no. 2, Sep. 2020, doi: https://doi.org/10.1002/col.22576.