Fruits Maturity Classification for Harvesting Robot

Reno Muhammad Fadilla*1, Tokuo Tsuji*2, Tatsuhiro Hiramitsu*3, Hiroaki Seki*4, Kaihei Okada*5
Kanazawa University, Kakumamachi, Kanazawa, Ishikawa, 920-1192, Japan
*Email: renofadilla01@stu.kanazawa-u.ac.jp*1, tokuo-tsuji@se.kanazawa-u.ac.jp*2, thiramitsu@se.kanazawa-u.ac.jp*3, hseki@se.kanazawa-u.ac.jp*4, kaihei2112@stu.kanazawa-u.ac.jp*5*

Abstract

This research proposes to create a machine vision using YOLOv5 model to detect maturity of the tomatoes on farm field which categorized by raw, half-ripe, and ripe. The use of multiple color space RGB, HSV, and LUV dataset aims to overcome external disturbances such as light and shadows. The results are the machine vision can differentiate fruits by its category with minor bias and obtaining a good mAP50 rates with the greatest result is LUV color spaces model. This machine vision with custom datasets in multiple color spaces can be implemented on farm field which is considered based on the purposes, environment and external disturbances. These models are developed for autonomous harvesting robot on farm field.

Keywords: Machine Vision, YOLOv5, Color Space

1. Introduction

The gradual maturity of technology has developed in many sectors, one of them is in the agricultural sector called as precision agriculture. Many agricultural systems still using traditional method, particularly in the maturity detection level. The research on precision agriculture aims to transform traditional method into digitalization precisely [1]. Robotics and automation are expected to have a significant impact on farms of the future by increasing efficiency in production for fruits grading which are gradually shifting to automation through a Computer Vision based on fruits detection. In addition, image processing can provide understanding of individual health, nutritional condition, and maturity level. However, the complexity of the field environment as well as the unstructured features of fruits resulting a great challenge to the target detection, particularly for maturity level measurement [2]..

Computer Vision (CV) is a key point for future implementation on farming technology, particularly in the fruits grading and classification, yield estimation counting, health condition, and maturity. Fundamentally, CV generated using Machine Learning (ML) algorithm which enable the analysis of massive quantities of data more rapidly and precisely, instead traditional or manual method. CV traditionally execute image by primary color Red, Green, and Blue (RGB), so the images collected by RGB based camera at RGB wavelength. In addition, grading fruits can be classified by using color, texture, and share feature descriptors. RGB color detection has a good level accuracy when using CV, however, at some condition, CV cannot detect object which has rough or irregular environment, such as on farm field. On the fam field, the objects can be disturbed by external factor, for example, sun rays, leaf, and shadows [3], [4]. To overcome this issues, execution process of CV can be implemented using multiple color space identification.

Color space basically can also describe as the way in which human eyes can perceive color vision. Color space shaped from the development of color model which the color is associated with an accurate definition of the way the components are to be inferred and the situations are viewed with the set of resulting colors. There are many different color spaces such as RGB, NTSC, LUV, HSV, CMYK, and HSI. Each color space has specified function and color model [5].

In the recent years, many researches have conducted to accomplish the best algorithm for ML, basically in image processing, particularly in the fruit's detection based on its color space. Detection of seedling using machine-transplanted rice has been developed using machine vision [6] which aimed to detect paddy seedlings using many color spaces and plant the seeds using a paddy machine. Research [7] conducted research by tomato recognition and localize the object using YOLOv5 algorithm only in RGB images. Research [8] discussed about a quality evaluation for two type of grapes using random forest model to identify quality of each fruit type. Another implementation of ML in image processing on farming activity also developed for palm fruit [9] which purposed to classified the oil palm fresh fruit based on the textures and palm's skin.

This research proposed to determine maturity level for tomato fruits based on three different color space using ML and YOLOv5 model on the real farm implementation to develop the previous similar researches. The maturity level of the tomato was divided into three class those are raw, half ripe, and ripe. In addition, this research was classified the greatest result of dataset's performance based on its color space.

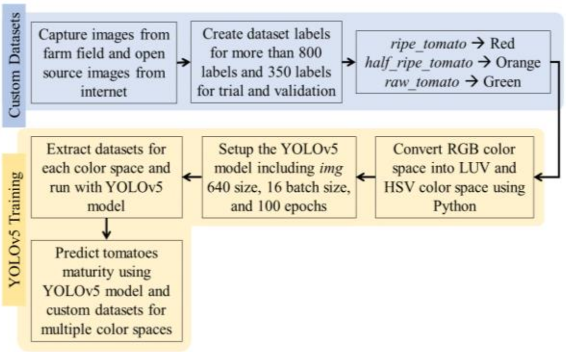


Fig. 1. Flowchart of proposed system.

2. Methodology

2.1. Proposed System

The overall flowchart of the proposed system is presented on the Fig. 1. The specified fruits which used as the main object is tomato. The system was proposed using Machine Learning based on YOLOv5 Algorithm with custom datasets. The used of YOLOv5 in this research rather than another version of YOLO has recently analyzed based on recognition and location of on-farm fruits. The previous research [7] explained about the YOLOv5 accuracy for on-farm moving robot. YOLOv5 has a high-level accuracy rate for rapid image detection with only 9.0 time/ms. Besides the accuracy level can be reach for more than 95%. In addition, the number of objects detected by YOLOv5 is the highest then other YOLO model. YOLOv7 is the latest model of YOLO version however the average detection time is 13.59 time/ms and with a little bit lower accuracy rate compared with YOLOv5. Based on this literature, originally, this research purpose is to develop a machine vision for on-farm robot. Thus, it is required a machine vision to detect fruits rapidly and precisely. All of the test has the same requirements such as 100 epochs, 32 batch with custom datasets in 640 x 640 image size, and three object classes.

This research was conducted in many trials and three sequences for identifying fruits based on three color spaces those are RGB color space trial, HSV color space trial and LUV color space trial. This research conducted to explain scientifically based on the data training. The purpose of each trial is to achieve the best success rate. Then, arrange each dataset depends on the color spaces and implement it into real robot vision on-farm for future concept.

2.2. Custom Datasets

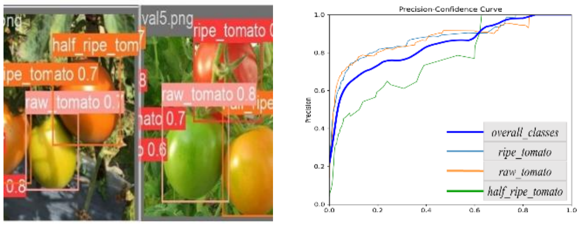
Images of tomatoes used in training and testing were collected from open-source image from internet and manually captured using phone camera on the tomato farm field located in West Java, Indonesia. The total image of the data was obtained and store basically as RGB format. The datasets for both trial and validation containing many numbers of samples which consist of single tomato for either red or green or orange tomato, group of tomato with the same color, multiple mixed color type of tomato. In addition, the datasets also consist of multiple small sized images for either red or green or orange tomato. The small sized images used to recognize fruits with far distance of camera or the fruits with a tiny shape. To increase the diversity of the samples, the datasets also contain multiple image characteristics such as blur, dim, front-facing, upper-part, bright, and overlapping objects.

The datasets consist of three classes those are *ripe_tomato*, *raw_tomato*, and *half_ripe_tomato*. The *ripe_tomato* represent to mature fruit, the *raw_tomato* represent to unripe fruit, and the *half_ripe_tomato* represent the fruit is half mature [10]. The datasets also distinguished based on the purpose those are train images and validation images. The total sample images of trial and validation for *ripe_tomato* is 192 and 136 respectively, the total samples of trial and validation for *raw_tomato* class are 243 and 101 samples respectively, and the total samples of trial and validation for *half_ripe_tomato class* are 403 and 139 respectively. The total samples from 334 images datasets are approximately 838 samples for trial and 376 samples for validation. Each class basically is RGB color space and converted to another color spaces HSV and LUV. These variety of class and color spaces appointed to determine the best color space to detect fruits maturity level.

3. Results and Discussion

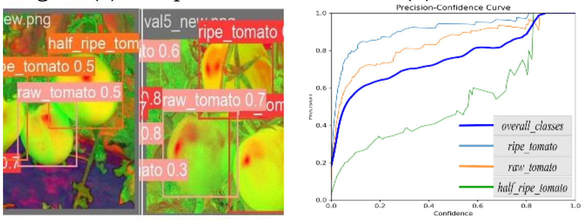
3.1. RGB Trial

This trial consists of several tests and each test has its main purpose. The first test purposed the machine vision to detect fruits and classified them into two classes *ripe_tomato and raw_tomato*. However, the test result has low quality mean Average Precision (mAP50) for only 54% and 52% respectively with total 49%. These issues happened because the machine vision cannot detect small-sized fruits for further scope of the images. In the second trial, more images added to the dataset, specifically for small-sized tomato for either red or green tomato class with total additional datasets are 90 images for both. The additional datasets including single small sized red and green tomato, and multiple objects of small-sized tomato. The results of the second trial for *ripe_tomato* and *raw_tomato* class have mAP50 92% and 83% with overall mAP50 is 87%. However, in the second trial, the machine cannot detect orange tomato which is neither categorized as red nor green tomato. Thus, the final trial conducted to add more images and one dataset labelled as *half_ripe_tomato*. The total images added for the *half_ripe_tomato* class are 133 images containing single tomato, multiple tomato, multiple mixed color tomato, and the small-sized tomato. The total additional images are 133, 328 images for overall class, and 823 samples in total for trial and 360 samples in total for



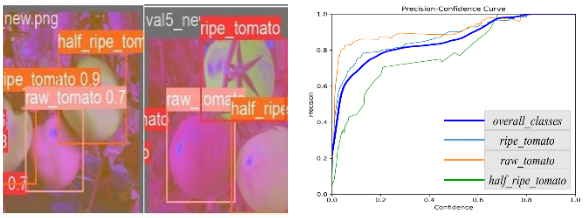
(a) **(b)**

Fig. 2. (a) Sample results RGB, and (b) P-C Curves.



(a) **(b)**

Fig. 3. (a) Sample results HSV, and (b) P-C Curves.



(a) **(b)**

Fig. 4. (a) Sample results LUV, and (b) P-C Curve.

validation. The sample images of final results shown on the Fig. 2(a) and Precision-Confidence Curve (P-C Curve) shown on the Fig. 2(b).

3.2. HSV Trial

HSV trial basically used the same number of total datasets with RGB datasets. In the HSV trial, the images converted into HSV images using Python code. The HSV trial also use three classes called as *ripe_tomato*, *raw_tomato*, and *half_ripe_tomato*. The trial test of this HSV datasets conducted in 100 epochs. Brief results of the HSV trial are the machine vision cannot differentiate *raw_tomato* and *half_ripe_tomato* class. This issue happened because the HSV color space conversion generates those tomato's skin look exactly the same. The sample images of final results shown on the Fig. 3(a) and Precision-Confidence Curve (P-C Curve) shown on the Fig. 3(b).

3.3. LUV Trial

LUV trial basically used the same datasets from the converted RGB color space to LUV color space using Python code. The LUV datasets also consist of three classes *ripe_tomato*, *raw_tomato*, and *half_ripe_tomato*. The Trial test conducted in 100 epochs with the detail information of sample final results shown on the Fig. 4(a) and Precision-Confidence Curve (P-C Curve) shown on the Fig. 4(b).

3.4. Performance Evaluation from All Datasets

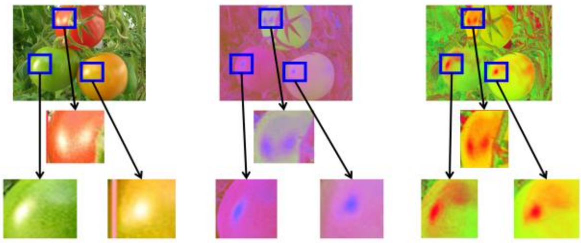
Based on this research, there are many advantages and disadvantages to use multiple color space in a dataset. Each color space has specific purpose to be used in machine vision. HSV color space can be used to detect color-based image. HSV separates luma, or the image intensity from chroma or color information. HSV can be useful for histogram equalization, removing shadows, lightning disturbance, etc. Besides, RGB color is the basic color space in every image and captured by camera. RGB color can be easily to use and formulated because this color space not necessary to be converted from the original image. However, RGB color space might be also easily distracted with external factors such as shadow, lightning, sunrays, etc. Meanwhile the LUV color space can be used to decouple color with the U and V components which represent chromaticity values of color image. Besides the L or Luminance can improve the RGB picture which distracted from over light disturbance, etc. The comparison from three color spaces which used in this research shown in Table 1.

Based on the Table 1, the LUV and RGB color space has similar mAP50 rates with the overall value 85% and 83% respectively. On the other hand, the LUV datasets have a good performance in object detection proven by the mAP50 rate shown on the Table 1 because LUV color space are fundamentally suitable for restoration error estimation, fast detection, and impulsive noise removal in real-time color imaging. Based on the Table 1, *ripe_tomato* class and *raw_tomato* class for LUV color space's mAP50 rate is higher than RGB color space because some of the images in RGB overlapped by shadows and light as shown in the Fig. 6(a) and Fig. 5(a) respectively. The explanation of comparison between RGB and LUV datasets also visualized in P-C Curve as shown in the Fig. 2(b) and Fig. 4(b) which shows the RGB curve relatively fluctuating than LUV graph. Nevertheless, the graph shows a great stability and raise gradually with small fluctuations. In contrary, the HSV has the lowest of all datasets training. The HSV datasets conversion from RGB color space has the typical color tone with its backgrounds such as leaf, trees, and grass. The HSV datasets, either *ripe_tomato* or *raw_tomato* class has a high mAP50 rates for more than 70%. However, there is an issue regarding the *half_ripe_tomato* class, the machine vision cannot detect the object significantly resulting only 57% of mAP50 value as shown on the Table 1. This value is nearly with 50% and the results should be evaluated to increase the rates [11]. The P-C curves Fig. 3(b) showing the relatively unarranged graph for the *half_ripe_tomato* and the prediction line increased gradually.

The shadow and lightness disturbances occurred during the dataset collection in this research as shown in Fig. 5 and Fig. 6. RGB image reflect lightness, such as sunrays,

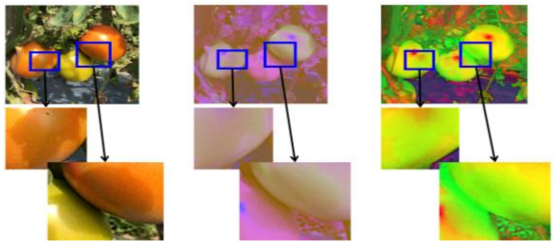
Table 1. Final results comparison.

Class	Total Samples		mAP50 *(Epochs = 100)*		
	Trial	Validation	RGB	HSV	LUV
ripe_tomato	192	136	88%	83%	95%
raw_tomato	243	101	81%	77%	84%
half_ripe_tomato	403	139	80%	57%	77%
Overall	838	376	83%	72%	85%



(a) **(b)** **(c)**

Fig. 5. Lightness disturbance: (a) RGB, (b) LUV, (c) HSV.



(a) **(b)** **(c)**

Fig. 6. Shadow disturbance: (a) RGB, (b) LUV, (c) HSV.

into white spot, whereas LUV image converts the white spot into warm blue color as shown in the Fig. 5(b) which was not distract the real color of the image and affecting the machine vision detection more precisely. HSV images turn the white spot into dark red color as shown in Fig. 5(c) and affect the real color of the images. HSV images turn the shadow into lighter color compared with the basic RGB image as shown in the Fig. 6(c) and Fig. 6(a) consecutively. This result influence the machine vision to detect object more precisely. Moreover, LUV also turn the shadows into lighter color than RGB shadow, however, the shadows still configurable and not affected much to fruit's basic color as shown on the Fig. 6(b). Basically, YOLOv5 model trained using a pre-training model originally from RGB color which trained in another color space LUV and HSV. Hence, the model generated results in relatively similar to RGB color space. Overall, LUV color space was the greatest datasets for tomato classification. Therefore, the LUV can be suitable for harvesting robot despite each color space has an advantages and disadvantages in certain occasions and purposes.

4. Conclusion

This research has shown the model of ML which used to determine the best datasets differentiated from color space RGB, HSV, and LUV. According to the final results, RGB and LUV datasets generate an outstanding typical result for the machine vision, on the contrary, the HSV datasets obtain a relatively poor result. The best class in this research from overall datasets evaluation were *ripe_tomato* and *raw_tomato* class. Meanwhile, the *half_ripe_tomato* class acquire intermittent results for all datasets color space depends on the backgrounds and external distraction from the objects. Moreover, HSV datasets need some additional images and adjustment of the "Hue" value as required for increasing the success rate. In conclusion, all of the datasets in this research can be used depending on the purposes, and object's circumstances which correlated with external disturbances such as backgrounds, lightness, and shadows of the images. LUV color space has shown its performance as the best datasets for rapid segmentation of tomato on farm and more suitable in many external disturbances such lightness and shadows. On the other hand, HSV datasets can be used to recognized tomatoes on farm which has high light intensity and shadow disturbance. Therefore, this color space can be more suitable for harvesting robot with rapid segmentation.

References

[1] E. Mavridou, E. Vrochidou, G. A. Papakostas, T. Pachidis, and V. G. Kaburlasos, "Machine vision systems in precision agriculture for crop farming," Journal of Imaging, vol. 5, no. 12. 2019. doi: 10.3390/jimaging5120089.

[2] W. Jia, Y. Tian, R. Luo, Z. Zhang, J. Lian, and Y. Zheng, "Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot," Comput Electron Agric, vol. 172, 2020, doi: 10.1016/j.compag.2020.105380.

[3] A. Bhargava and A. Bansal, "Fruits and vegetables quality evaluation using computer vision: A review," Journal of King Saud University - Computer and Information Sciences, vol. 33, no. 3. 2021. doi: 10.1016/j.jksuci.2018.06.002.

[4] N. Ismail and O. A. Malik, "Real-time visual inspection system for grading fruits using computer vision and deep learning techniques," Information Processing in Agriculture, vol. 9, no. 1, 2022, doi: 10.1016/j.inpa.2021.01.005.

[5] H. Rubesh, D. Hema, and S. Kannan, "Interactive Color Image Segmentation using HSV Color Space," Science and Technology Journal, vol. 7, p. 1, 2020, doi: 10.22232/stj.2019.07.01.05.

[6] H. Li et al., "An automatic approach for detecting seedlings per hill of machine-transplanted hybrid rice

utilizing machine vision," Comput Electron Agric, vol. 185, 2021, doi: 10.1016/j.compag.2021.106178.

[7] T. Li et al., "Tomato recognition and location algorithm based on improved YOLOv5," Comput Electron Agric, vol. 208, 2023, doi: 10.1016/j.compag.2023.107759.

[8] D. Pietro Cavallo, M. Cefola, B. Pace, A. F. Logrieco, and G. Attolico, "Non-destructive and contactless quality evaluation of table grapes by a computer vision system," Comput Electron Agric, vol. 156, 2019, doi: 10.1016/j.compag.2018.12.019.

[9] A. Septiarini, A. Sunyoto, H. Hamdani, A. A. Kasim, F. Utaminingrum, and H. R. Hatta, "Machine vision for the maturity classification of oil palm fresh fruit bunches based on color and texture features," Sci Hortic, vol. 286, 2021, doi: 10.1016/j.scienta.2021.110245.

[10] Ifmalinda, Andasuryani, and I. Rasinta, "Classification of tomato (Lycoersicon Esculentum Miil) ripeness levels based on HSV color using digital image processing," in IOP Conference Series: Earth and Environmental Science, 2022. doi: 10.1088/1755-1315/1116/1/012062.

[11] Q. Zheng, Z. Peng, P. Zhu, Y. Zhao, R. Zhai, and W. Ma, "An Object Recognition Grasping Approach Using Proximal Policy Optimization with YOLOv5," IEEE Access, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3305339.

Authors Introduction

Mr. Reno Muhammad Fadilla

He is currently a Master Student in the Division of Frontier Engineering, Kanazawa University, Japan since 2023 with research theme in the fields of computer vision, and robotics. He received his Bachelor of Engineering from UIN Bandung, Indonesia in 2022.

Dr. Tokuo Tsuji

He is currently an Assoc. Professor in the Faculty of Frontier Engineering, Kanazawa University with research area in the fields of informatics, and intelligent robotics. He received his Doctoral Degree in Engineering.

Dr. Tatsuhiro Hiramitsu

He is currently an Assistant Professor in the Division of Frontier Engineering, Kanazawa University, Japan since 2019 with a research area in the fields of soft robotics and intelligent robotics. He received his Doctoral Degree in Engineering.

Dr. Hiroaki Seki

He is currently a Professor in the Faculty of Frontier Engineering, Kanazawa University, specialty in the fields of robotics and mechatronics automation. He received his Doctoral Degree in Engineering from The University of Tokyo in 1994.

Mr. Kaihei Okada

He is currently a master student in the Division of Frontier Engineering, Kanazawa University, Japan since 2023 with research theme in the fields of computer vision and robotics. He received his Bachelor Degree in Engineering from Kanazawa University, Japan.