ISSN: 18158129 E-ISSN: 18151027
Volume 17, Issue 09, September, 2021

Apply artificial intelligence to the robot picking strawberries

Nguyen Le Dung[1,2*], Nguyen Hoang Giap[1,2], Phan Huynh Lam[1,2], Huynh Thien Nhien[1,2]

Deparment of Automation and Control, Faculty of Electrical and Electronics Engineering, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam[1]
Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc District, Ho Chi Minh City, Vietnam[2]

Corresponding author: 1*

ABSTRACT— Smart agriculture [1-10] is becoming more and more popular, from monitoring systems in greenhouses to agricultural robots. Today, the application of robots to agriculture is not unfamiliar, particularly with the development of AI, that we can create robots that automatically harvest agricultural products [13]. Some parts of the world focus on NIR cameras to process identification under models RCNN, Faster R-CNN, VOC ... in this article, we mention the use of Yolo v3 to identify strawberries by Real sense D435i 3D camera, then through the evaluation function to select a strawberry to locate the strawberry in the space for easy harvest robot, to identify the strawberry in the space we combine the 3 direction property The axis in the IMU has built-in camera and 3D image space, the harvesting process is a dynamic process, so we will use anti-noise filters, target orientation, IMU sensor noise and follow the trajectory. transfer, ensuring the picking mechanism does not occur instability when identification error and sensor impact on the system. Where the anti-noise filter for IMU sensor status is extended Kalman and for the motion process is a smooth line incorporating gauges to eliminate peak noise. In this section, we only focus on the application of strawberry identification and identification in space; In the next direction, we will evaluate the SLAM combination robot system for self-propelled problem.

KEYWORDS: D435i, Jetson nano, strawberry, artificial interligent.

1. INTRODUCTION

Image recognition is no stranger to Yolo, CNN, RCNN models [1- 12] here we use Yolo v3 to combine the existing library of 3D realsense D435i cameras; help us to quickly identify the strawberry and locate it in space; From there we easily control the robot to harvest.

We can review a system with NIR Camera [12].

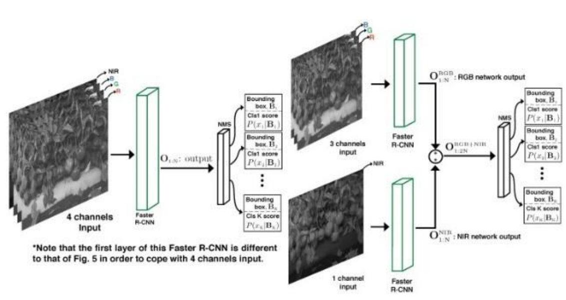

Fig 1. A diagram of the early and late fusion networks. (a) The early fusion that concatenates a 1- channel NIR image with a 3-channel RGB image; (b) The late fusion that stacks outputs, ORGB+NIR1:2N

We can see NIR camera, it has special value for food field; We can see invisible images, it helps the result more accurate, but NIR camera is more expensive than normal camera or 3D camera; so that we use 3D camera for this system.

Profit: low cost, 3D sensors, 3D point cloud, fast speed.
Nonprofit: cannot check special images out of wave signal.
Introduction of system hardware:

- Camera D435i: to capture 2D, 3D images and IMU sensor information

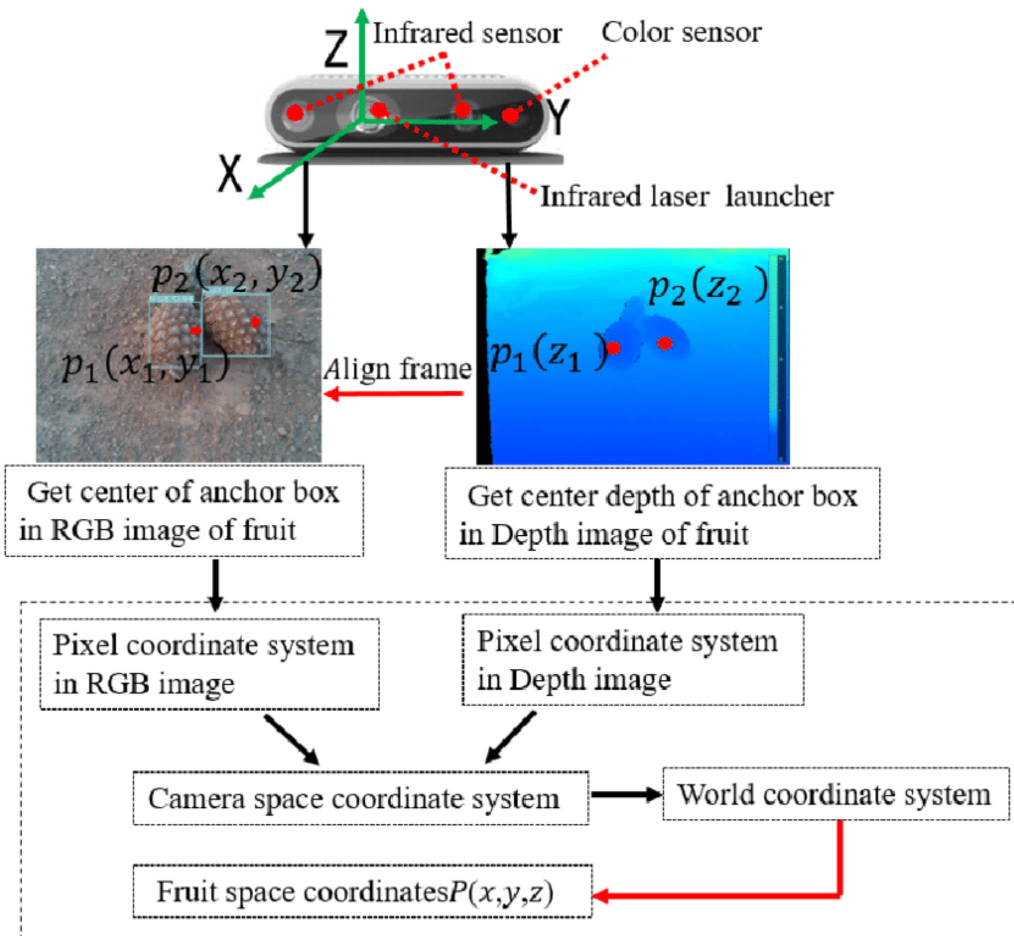

Fig 2. Reference method of convert virtual space to real space

- Jetson nano: image information processing, commands for embedded system control
- Arm robot: harvest products.

2. Data and Method

The process is divided into two phases: 2D image processing, strawberry recognition, and 3D image processing. In the framework of this paper, we stop at the 2D image recognition model by yolo, we also evaluate the 3D recognition system, but the system overload and the process urgency are not high, we will study in another article. [12]

2.1 Sampling

As is known, the basic training model will need image samples to determine parameters for the AI set. Here we train 5,300 pictures of strawberries including selfies and pictures taken from the internet. The images come in different sizes for added variety.

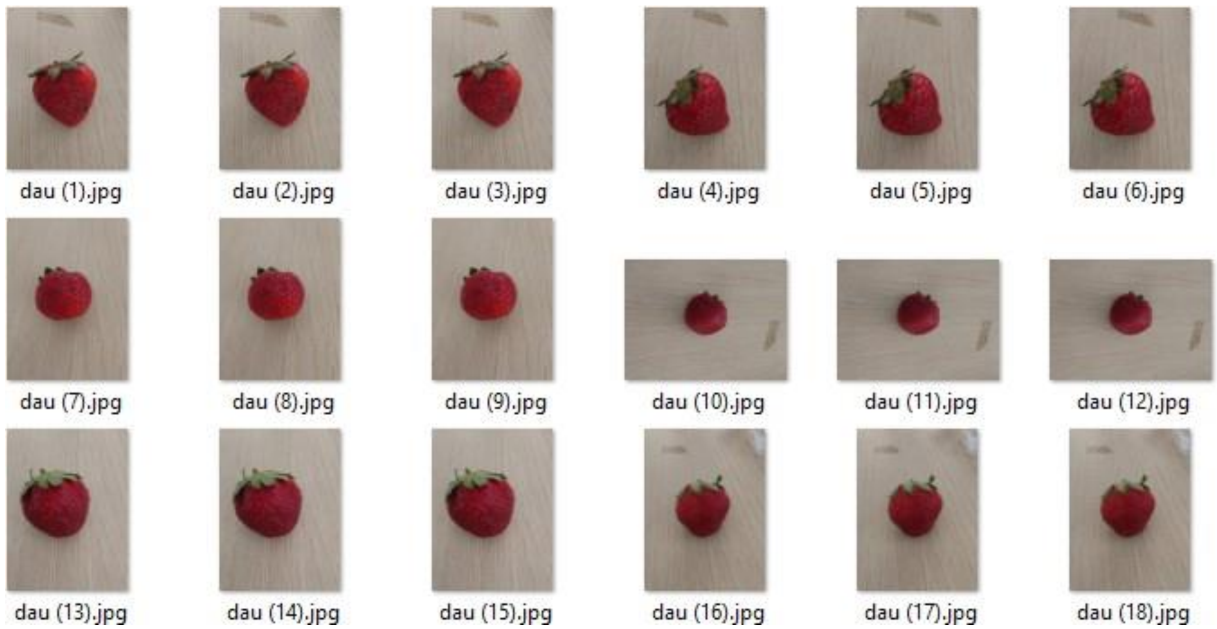

Fig 3. Data for train model

2.2 Model training

During training, when the loss or avg_loss does not change much, the train can be stopped. However, when avg_loss converges to 0, if you choose the wrong weights, it will easily cause Overfitting, so we will consider adding another parameter called mAP.

mAP (mean average precision): average value of correct predictions for each class. Where Average precision is the average value of 11 points on the Precision Recall - curve for each detection probability for the same class. Then take the average of the APs mAP.

$$Precisio = \frac{true\ positive}{true\ positive + false\ positive}$$

$$Recall = \frac{true\ positive}{true\ positive + false\ negative}$$

With a way of determining a class to be positive, Precision is defined as the ratio of the number of true positive points to those classified as positive (TP + FP). Recall is defined as the ratio of the number of true positive points to those that are actually positive (TP + FN).

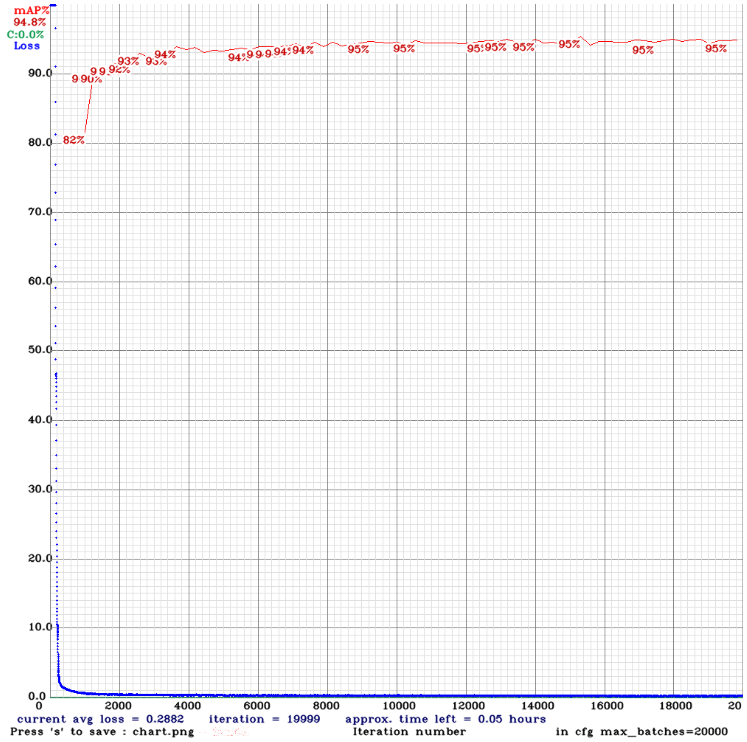

Fig 4. mAP of process trainning model strawberry

According to the results, decide to choose weights in step 14000 to detect strawberries, avg_loss about 0.25.

3. Algorithm

3.1 Flowchart

We observe the following image processing flowchart:

- Image data collected from D435i camera is divided into 2D, 3D, IMU data streams.
- Carry out strawberry identification by Yolo and choose a result for the following process.
- The next process proceeds from the 2D position on the image to 3D space for the robot to operate.

For each data, we will stack it into array (n=32, 64.) and then we do filter action, because while robot move or affect by environment, location of strawberry in camera will not stable, IMU sensor always have offset errors… so the controller will have unstable, unwanted peak, we need to smooth signal output for driver, that system is not vibration, harmful, save life devices.

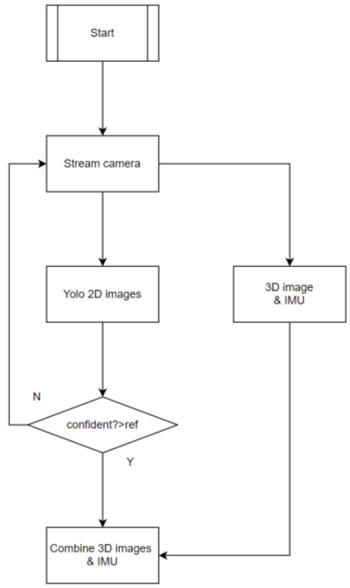

Fig 4. Flow chart image processing

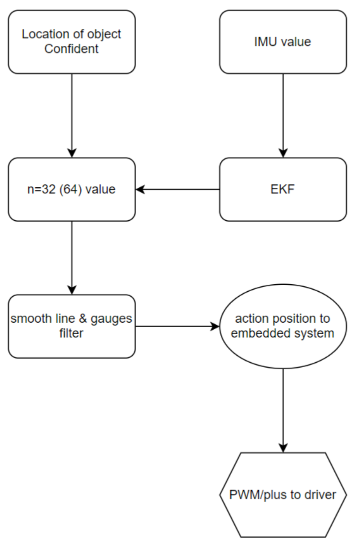

Fig 5. Flow chart filter and control embedded system

3.2 Result of detection

The basic identification result was> 95.5% (per 1000 test data), because in the greenhouse environment, separate strawberry characteristics are still missing due to the size setting removing the berries. at long range, and color properties differ when affected by the ambient light.

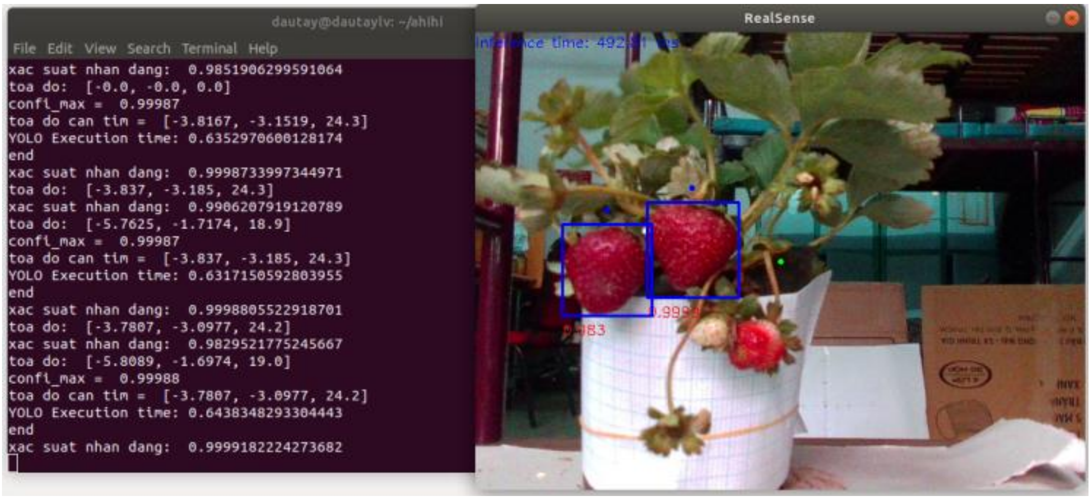

Fig 6. Results of strawberry identification through the Yolo model

Table 1. Results of reliability and processing time

Distance / **n**	**200 mm**		**300 mm**		**400 mm**	
	Reliability	Cycle	Reliability	Cycle	Reliability	Cycle
1	0.9995	0.3237	0.9916	0.3126	0.8263	0.3201
2	0.9991	0.3143	0.9907	0.3149	0.8999	0.3142
3	0.9999	0.3162	0.9432	0.2963	0.8049	0.3146
4	0.9998	0.3219	0.9599	0.2978	0.9680	0.3079
5	0.9998	0.3128	0.8121	0.2974	0.8311	0.3238
6	0.9995	0.3079	0.9907	0.3121	0.9683	0.3109
7	0.9997	0.3074	0.9342	0.3009	0.9181	0.3409
8	0.9995	0.3131	0.8940	0.3050	0.8647	0.3270
9	0.9998	0.3311	0.8621	0.3592	0.6361	0.3332
10	0.9987	0.3253	0.8215	0.3346	0.7544	0.3218
Avg	0.9996	0.3174	0.9199	0.3131	0.8472	0.3214

We found that basically 99.9% accuracy, processing time ~ 0.3s.

3.3 The positioning of strawberries in space

From 2D identification information, we can easily recognize the strawberry position in the 2D image space. Here the D435i has APIs available for 3D spatial transformation. We examine the results of the positioning in space as follows:

Table 2. Experimental results of spatial positioning

n	Target (mm)	Result (mm)	Err (mm)

	x	y	z	x	y	z	x	y	z
10	-15.6	-30	200	-16.7838	-29.173	203.9	3.9641	1.9864	4.9
10	20.2	-5.9	300	20.7794	-5.3023	302.2	2.8116	1.9825	8.6
10	16.6	-21.9	400	20.2406	-20.4708	395.7	4.5574	3.2094	11.5

Table 3. Mean error of z

Err (avg) (%)	**200 mm**	**300 m**	**400 mm**
z	2.45	2.8667	2.875

Comment:
-Generally, the error of x and y is less than 1cm (less than 5mm specific), the distance z has an error of approximately 1 cm and less than 5%, this is a number that is suitable for the requirements of the topic. The arm of the machine uses a cutting blade, so if the x and y coordinates are less than 1 cm deviation, it will not affect the final result of cutting the strawberries because there is always post-check during the cutting process.

4. Discussion and Conclusion

The authors have achieved the goal of identifying the strawberry with the rate> 95.5%, and determining the position in the space for the robot arm. The authors still realize that there are still many shortcomings to improve:

- 3fps image recognition speed ➔Convert to higher speed model & better GPU hardware.
- Add features processing conditions to cut strawberries: large enough size, color ...
- Post-check the coil when cutting ➔Deploy 01 camera at the arm end

The authors will consider combining the SLAM basis that the team has researched and published [13]; to become a complete robot system, able to autonomously in the greenhouse environment and harvest agricultural products. Harvesting strawberries is just a test, the authors will conduct on many different fruit models to evaluate the practicality of the robot.

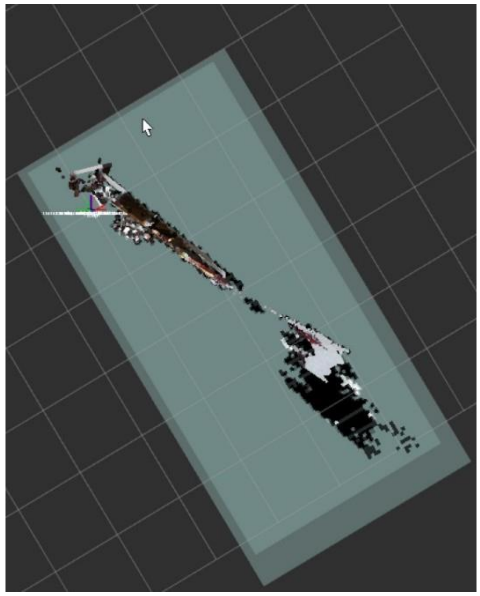

Fig 7. Map of SLAM, a prototype is process build map on Jetson Kit

5. Acknowledgments

This research is funded by Murata Science Foundation (MSF) under grant number 08-FY2020. We acknowledge the support of time and facilities from Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for this study.

6. REFERENCE

[1] Anand Koiralaa Kerry B. Walsha Zhenglin Wanga Cheryl McCarthyb, Deep learning – Method overview and review of use for fruit detection and yield estimation, Volume 162, July 2019, Pages 219-234, Computers and Electronics in Agriculture

[2] Kondo N., Monta M., Noguchi N. Agricultural Robots: Mechanisms and Practice. Trans Pacific Press; Balwyn North Victoria, Australia: 2011. [Google Scholar]

[3] Bac C.W., van Henten E.J., Hemming J., Edan Y. Harvesting Robots for High-Value Crops: State-of-the-Art Review and Challenges Ahead. J. Field Robot. 2014; 31:888–911. doi: 10.1002/rob.21525. [CrossRef] [Google Scholar]

[4] McCool C., Sa I., Dayoub F., Lehnert C., Perez T., Upcroft B. Visual Detection of Occluded Crop: For automated harvesting; Proceedings of the International Conference on Robotics and Automation; Stockholm, Sweden. 16–21 May 2016. [Google Scholar]

[5] Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., et al. Imagenet large scale visual recognition challenge. Int. J. Comput. Vis. 2015; 115:211–252. doi: 10.1007/s11263-015-0816-y. [CrossRef] [Google Scholar]

[6] Ge Z.Y., Sa I. Open datasets and tutorial documentation. 2016. [(accessed on 31 July 2016)]. Available online: http://goo.gl/9LmmOU.

[7] Wikipedia F1 Score. 2016. [(accessed on 31 July 2016)]. Available online: https://en.wikipedia.org/wiki/F1_score.

[8] Nuske S.T., Achar S., Bates T., Narasimhan S.G., Singh S. Yield Estimation in Vineyards by Visual Grape Detection; Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS ’11); San Francisco, CA, USA. 25–30 September 2011. [Google Scholar]

[9] Nuske S., Wilshusen K., Achar S., Yoder L., Narasimhan S., Singh S. Automated visual yield estimation in vineyards. J. Field Robot. 2014; 31:837–860. doi: 10.1002/rob.21541. [CrossRef] [Google Scholar]

[10] Yamamoto K., Guo W., Yoshioka Y., Ninomiya S. On plant detection of intact tomato fruits using image analysis and machine learning methods. Sensors. 2014; 14:12191–12206. doi: 10.3390/s140712191. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[11] Wang Q., Nuske S.T., Bergerman M., Singh S. Automated Crop Yield Estimation for Apple Orchards; Proceedings of the 13th Internation Symposium on Experimental Robotics (ISER 2012); Québec City, QC, Canada. 17–22 June 2012. [Google Scholar]

[12] Inkyu Sa,* Zongyuan Ge, Feras Dayoub, Ben Upcroft, Tristan Perez, and Chris McCool, DeepFruits: A Fruit Detection System Using Deep Neural Networks Sensors (Basel). 2016 Aug; 16(8): 1222.Published online 2016 Aug 3. doi: 10.3390/s16081222

[13] Yuanyue Ge 1, Ya Xiong 1, Gabriel Lins Tenorio2, And Pål Johan from 1, Fruit Localization and Environment Perception for Strawberry Harvesting Robots, special section on new technologies for smart farming 4.0: research challenges and opportunities, Digital Object Identifier 10.1109/ACCESS.2019.2946369

[14] Research on Spatial Positioning System of Fruits to be Picked in Field Based on Binocular Vision and SSD Model, Xing Zhang, Qiaoming Gao, Dong Pan, Peng Cheng Cao and Dong Hui Huang; Journal of Physics: Conference Series, Volume 1748, The 2020 5th International Seminar on Computer Technology, Mechanical and Electrical Engineering (ISCME 2020) 30 October-1 November 2020, Shenyang, China; Citation Xing Zhang et al 2021 J. Phys.: Conf. Ser. 1748 042011