Spatial Strawberry Detection for Fruit Quality Assessment

Nino Wegleitner
01644401@student.tuwien.ac.at

Matthias Hirschmanner
hirschmanner@acin.tuwien.ac.at

***Abstract*—This study introduces a strawberry localization system for edge computing using a mobile robot, a low-cost single board computer and a stereo camera with dedicated AI processor. The system is highly independent from the host computers processing power and able to recognize, localize and count strawberries in their different growth stages. Results show good localization and sizing performance at low strawberry-to-camera-distances and the systems ability to estimate the number of strawberries inside a raised strawberry bed.**

I. Introduction

Strawberries are highly sought-after and economically significant crops worldwide. However, their successful cultivation and optimal yield depend on accurate monitoring and management throughout their growth stages. Traditional methods of assessing strawberry crops often involve manual inspection, which is labor-intensive, time-consuming, and prone to errors.

In response to this, we evaluating the performance of a low-cost strawberry localization system consisting of a mobile robot and an OAK-D stereo camera [4] and its dedicated AI processing unit for recognizing strawberries, determining their growth stages, and precisely localizing and counting them.

II. Related Work

Similar studies [1], [9] show the popularity of lightweight single-stage neural networks like YOLO for real-time 2D object detection on edge devices and the utility of stereo cameras for fruit localization and sizing [5]. As 3D detection methods generally require more computational power and are more challenging when labeling custom datasets [7], detections obtained from 2D neural networks can be combined with depth data to acquire spatial information of detected objects [1]. While promising results could be achieved even on edge devices like Nvidia Jetson Nano [6], the system implemented in this study is highly independent from the host computers processing power and able to run on devices without dedicated GPU like a Raspberry Pi 4.

III. Material & Environment

The strawberry localization system consists of the small mobile robot Pioneer 3-DX, a Raspberry Pi 4 and an OAK-D Pro stereo camera. As a software framework the Robot Operating System (ROS) is used.

The Pioneer 3-DX by Adept MobileRobots is a lightweight two-wheeled differential drive mobile robot. The ROS library RosAria provides a ROS node to publish and subscribe to data received and send from the robot's embedded controller. This is used to control the robot and to receive odometry data.

Fig. 1. Localization system setup in its environment: Pioneer 3-DX (1), Raspberry Pi 4 (2), OAK-D Pro (3) and raised strawberry beds (4)

As a host computer, a Raspberry Pi 4 is mounted on top of the robot. This low-cost, low-power embedded device runs Ubuntu 20.04 and ROS Noetic. It is used to communicate with both the mobile robot and the stereo camera.

Luxonis OAK-D Pro combines an (active) stereo depth camera (1 MP, 120 fps), a color camera (12 MP, 60 fps) and an on-device procession unit (4 TOPS) which can be used for neural network inferencing and other computer vision tasks like object tracking. It is mounted on top of the mobile platform facing 90° to the left in regard to the robots forward driving direction.

The mobile robot drives at a low speed of $0.05\mathrm{m\,s^{-1}}$ around a raised strawberry bed and keeps a pre-defined constant distance to it until returning to its starting point. The raised strawberry beds (3.8m x 0.25m) are located outdoors on one of the terraces at TU Wien.

IV. Method

In our approach, we use YOLOv5 [2] for 2D detection and classification. In addition to utilizing images from an existing strawberry dataset [3], approximately 200 images of strawberries in various growth states were captured using the OAK-D Pro camera and labeled accordingly. The dataset size was increased to over 1200 images using image augmentation. The labels were classified in: flower, unripe and ripe strawberry. This dataset [8] was used to train a YOLOv5n [2] detection model. The input image size was set to 352 x 352 pixel and the model was trained for 150 epochs with a batch size of 4 reaching 83.2 % mAP.

The OAK-D camera pipeline uses the downscaled color camera stream at 10fps as an input for inference and feeds the output 2D detections into a spatial location calculator. This spatial calculator averages the corresponding stereo depth data of the region-of-interest (ROI) and calculates spatial coordinates for the objects center. As the devices disparity search ranges upper limit is set to 95 pixels, the minimal depth perception distance is limited to 35cm at 400 p stereo resolution.

The object tracker uses Kalman Filter and Hungarian algorithm to track up to 60 detections at once. It generates unique IDs for every tracked object. By avoiding revisiting already passed areas while driving around the raised bed, the object tracker allows counting the number of strawberries inside the bed.

To obtain the size of the detected strawberries, the intrinsic camera parameters of the camera are used to transform the detection ROI from the image space into the 3D space. Strawberries are assumed to be hanging freely with their central axis in vertical direction and to be 3D uniform around their central rotation axis. This approximation is used to create 3D bounding boxes for the detected strawberries and allows estimating size and volume.

V. Results

In this paper, the performance of the strawberry detection system is evaluated in several steps.

The measured distance accuracy is evaluated by placing a strawberry at known distances to the camera and recording around 100 samples per distance. The minimal evaluated distance was set to 30 cm and increased by 10 cm steps up to 150 cm. Results shown in Fig. 2 show low error and high accuracy only at low camera-to-strawberry-distances. The growing error at higher distances might be explained by a slight misalignment between color and depth frames.

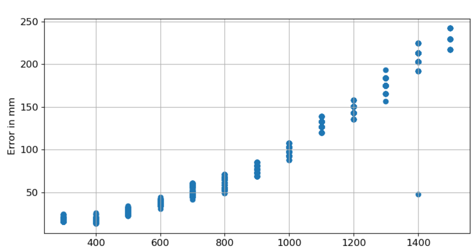

Fig. 2. Measured strawberry distance residuals

Furthermore, the size of two ripe sample strawberries (small: 29 mm x 27 mm, big: 43 mm x 41 mm) is measured at different camera-to-strawberry-distances in a range from 30 cm to 150 cm. Best results with low estimation error and low variance are achieved at low distances at around 50cm as can be seen in Fig. 3. At higher distances the variance of the recorded data increases noticeably and a linear growth of the height error can be observed. The reason for this is unknown and might be explained by errors regarding camera calibration or in the calculation process.

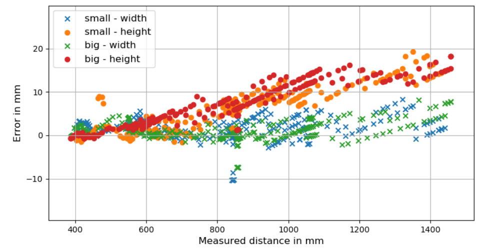

Fig. 3. Measured size residuals for two different sized ripe strawberries

Both size and distance evaluation show the best results at low distances. As measurements under the minimal depth perception distance of 35 cm tend to generate oversized outliers, a distance of 50 cm was defined as the preferable distance when driving around the raised strawberry beds. To evaluate counting, the mobile robot was driven around the raised strawberry bed and the number of flowers, unripe and ripe strawberries were counted. The experiment was repeated 4 times and the results were compared with a manual count. The results are shown in Tab. I and show similar numbers as the manual count for both flower and unripe strawberries, but higher numbers of detections for ripe strawberries. This is mainly caused by false positive detections and by assignment of multiple tracker IDs to single objects. The difference in number of detections in the runs can be attributed to the changing weather conditions causing changes in lighting and contrast.

TABLE I
Number of Detections per Class

Experiment	Flower	Ripe	Unripe
1	3	146	64
2	1	146	86
3	0	150	90
4	1	161	68
5	2	151	78
$\mu \pm \sigma$	1.4±1.04	152.8±6.85	77.2±10.01
Manual count	1	119	85

VI. Conclusion

We introduced a spatial strawberry localization system which realizes strawberry growth classification, sizing and counting using YOLOv5n object detection model running on OAK-D stereo cameras dedicated AI processing unit. The experiments show high size and positioning accuracy at low distances. To further enhance it, the overall detection accuracy of the YOLOv5n model could be improved by increasing the number of training samples. Furthermore, a short-range stereo camera, like the OAK-D SR model, which is better fit for distance measurement of under 1m, would allow using higher image resolution and therefore increase detection and depth accuracy.

REFERENCES

[1] Zhang H D Xu Y W Imou K Li M et al. Hu H M, Kaizu Y. Recognition and localization of strawberries from 3d binocular cameras for a strawberry picking robot using coupled yolo/mask r-cnn. *Int J Agric Biol Eng*, 15(6):175–179, 2022.

[2] Glenn Jocher. Yolov5 by ultralytics. https://github.com/ultralytics/yolov5, 2020.

[3] Stefan Lechner. Huggingface strawberry dataset. https://huggingface.co/datasets/St333fan/StrawberryRedGreenFlower. [Online; accessed 17-July-2023].

[4] Luxonis. Oak-d pro. https://shop.luxonis.com/products/oak-d-pro?variant=42455252369631. [Online; accessed 16-July-2023].

[5] Chiranjivi Neupane, Anand Koirala, Zhenglin Wang, and Kerry Brian Walsh. Evaluation of depth cameras for use in fruit localization and sizing: Finding a successor to kinect v2. *Agronomy*, 11(9):1780, Sep 2021.

[6] Muhammad Fauzan Ridho and Irwan. Strawberry fruit quality assessment for harvesting robot using ssd convolutional neural network. In *2021 8th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, pages 157–162, 2021.

[7] Yangfan Wang, Chen Wang, Peng Long, Yuzong Gu, and Wenfa Li. Recent advances in 3d object detection based on rgb-d: A survey. *Displays*, 70:102077, 2021.

[8] Nino Wegleitner. Strawberry growth state classification dataset. https://universe.roboflow.com/ninoweg/strawberrygreenredflower, 2023. [Online; accessed 13-July-2023].

[9] Yanchao Zhang, Jiya Yu, Yang Chen, Wen Yang, Wenbo Zhang, and Yong He. Real-time strawberry detection using deep neural networks on embedded system (rtsd-net): An edge ai application. *Computers and Electronics in Agriculture*, 192:106586, 2022.