James Latimer: 4933974
Muibdeen Akinyele: 4245374

Tomato DIstance Estimation and Segmentation

Problem Definition:
As part of the agricultural robotics lab that I work in, we are running into an issue where we are unable to reliably manipulate tomatoes that appear in feed. We plan on implementing a segmentation and distance estimation layer on top of the already existing camera feed so that our end effector knows when to open and close around objects of interest. Our part in this endeavor is to create the backbone of the CV layer here - so just creating a method for our cameras to recognize objects and know the distance our end effector is from the goal. However, we are not going to be working with the end effector yet, which is why the 3D printed test apparatus is needed.

DATASET:
Data will come from the Kaggle tomatoes dataset
https://www.kaggle.com/datasets/enalis/tomatoes-dataset
That classifies tomatoes as old, damaged, ripe, unripe

How will you solve this problem:

We will use the high-level Intel RealSense API. We don't want to leverage too many of the RealSense builtins, since the goal is to have this work with any rgb camera.

The libraries we will use include: Numpy, tensorflow, and opencv. The library selection is subject to change as we get more information on what is allowed to be used for the project.

A 3D printed apparatus will be created to maintain fixed distance of the cameras. We plan to use the Ultralytics YOLOv11 library for image segmentation. However, if we cannot do that, we will look into remaking a YOLO model from scratch or look into using a less “plug and play” library.

A 3d printer will be used for creating a test apparatus to maintain a constant horizontal distance for the cameras, since RGB only stereovision methods require camera separation distance to be known. It's not a major part of the software part of this project, but it's something that needs to be done for reproducibility of test procedure.

If possible, we can directly use the Ultralytics YOLOv11 api for segmentation, else we will try and read some research papers to design yolo from scratch.

The final prototype for this project will be comprised of the following:
2 cameras that maintain a fixed horizontal distance from each other
The cameras will input a video stream of image data that will be interpreted by the model
The model will then identify ripe tomatoes and return a an estimated distance between the cameras and tomatoes

James Latimer: 4933974
Muibdeen Akinyele: 4245374

Tomato DIstance Estimation and Segmentation

What will be your contribution:

James: For this project I plan on setting up the pipeline that gets the data from the real sense pipeline into our program to work with.

Muib: Data augmentation and selection techniques. Instead of using a 3rd party to augment data, I will create functions that augment our inputs and artificially expand the dataset. If allowed, I plan to train my data using the Ultralytics training functions as well.

Beyond that, we will need to speak with our professor to see what parts of our project can be done using existing libraries, and what needs to be done from scratch. From there we can fully split all of the tasks between one another