Overview of Field Robot
- Geolocation of strawberries by an autonomous field robot for strawberry harvesting
- Two cameras: (1) Fixed camera at 100cm on frame of robot (Zed2Camera) (2) End effector camera inside claw (Pixy Camera)
- Fixed camera detects the mature strawberries and determines pixel coordinate. A transformation to real world coordinates for the end effector to pick up the strawberry and place in the basket.

Objectives
- Develop a computer vision algorithm capable of detecting strawberries
- Determining their coordinates within the camera’s field of view – use Depth Estimation to determine the z coordinate along with set up camera calibration to get the real-world coordinates

Model & Dataset

Base Model
- Yolov8

Initial Training Dataset
- Size: 383 training images, 100 validation
- Content: Images of ripe and unripe strawberries.

Augmented Dataset
- Content:
  - Artificial Strawberries: Simulate strawberry patch during off season and to augment the dataset (294 images).
  - Real Strawberries: Captured from a strawberry patch (28 images).
- Classes: The dataset includes three classes:
  - Mature Strawberries
  - Immature Strawberries
  - Leaves

Detection Results

Artificial Strawberry Test

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>yoloV4</td>
<td>0.90</td>
<td>0.82</td>
<td>0.86</td>
</tr>
<tr>
<td>YoloV8</td>
<td>0.96</td>
<td>0.97</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Real Strawberry Test

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>yoloV8</td>
<td>0.79</td>
<td>0.49</td>
<td>0.61</td>
</tr>
<tr>
<td>YoloV8</td>
<td>0.74</td>
<td>0.72</td>
<td>0.73</td>
</tr>
</tbody>
</table>

The yoloV8 model had better recall and F1 scores for both artificial and real strawberry data evaluation.

Method

Strawberry Detection
- Mature strawberry- over ¾ red surface area
- Immature strawberry – less than ¾ red surface area
- Only mature strawberries are harvested so need to classify strawberry maturity
- Brightness augmented by 40% in dataset to simulate varying lighting condition

Depth Estimation
- Create a dataset with artificial strawberries at varying lengths (depth) from the fixed camera at a range of 68cm to 73cm.
- Use a Depth Estimation Model (Depth-Anything) to create a depth map of image
- Yolov8 model will detect strawberry bounding boxes to isolate the depth of strawberries in depth map
- Finetune depth estimator with training and scalar calibration

Depth Estimation Results

- Test with images of strawberries at varying heights
- Compared between different pretrained models of Depth-Anything-V2 (vitb, vits, and vitl)
- The MiDas Depth Estimator had the lowest RMSE of 2.051 hence it is closer to the ground truth values

Future Work
- Set up camera calibration to map strawberry coordinates in the workspace.
- Implement coordinate transformation to determine the geolocation of strawberry coordinates.
- Assess the accuracy of strawberry detection and coordinate estimation under different scenarios and compare the results between the fixed and moving camera configurations.

References