HoloNav: HoloLens as a surgical navigation system

Ee Xuan Tan
extan@student.tudelft.nl
Supervisors: Ricardo Guerra Marroquim, Mohamed Benmahdjoub

Detecting optical reflective spheres using YOLOv5 and the Hololens' grayscale cameras

1. Research Question

Can we use an object detection algorithm like YOLOv5 to precisely detect optical reflective spheres using the Hololens' grayscale cameras?

- How can we create bounding boxes on the grayscale images from the large data sets consisting of 1000+ images?
- How do we generate the text files to feed YOLOv5 the correct data?
- How accurate does YOLOv5 perform against the data sets? What are the limitations?

2. Background information

Object detection, in short, allows for the localization of objects of interest in an image. There are many object detection algorithms; however, YOLOv5 was chosen. YOLO (You Only Look Once) is an object detection algorithm that can be easily trained to identify objects accurately and at fast speeds.

	YOLOv5	Faster R-CNN	SSD	YOLOv4
Speed	✓			✓
Accuracy	✓			✓

3. Methodology

The YOLOv5s model will be trained with 70% of the data set, validated with 20% and tested with 10%. The data set consists of images and label files that have the dimensions of the bounding boxes

Bounding boxes

Bounding boxes can be

1. Manually annotated using Roboflow
2. Automatically annotate based on distance

Trade-off between accuracy and scalability. With data sets with over 1000+ images option 2 was preferred.

Metrics

Some important metrics that will be used to test accuracy:

1. The number of correctly identified, falsely identified and miss-identified. True positive (TP), False positive (FP), False negative (FN).
2. mAP is the measure of the average precision values calculated over recall values from 0 to 1
3. Average pixel distance, this is the average distance between the annotated bounding boxes to the predicted boxes.

4. Results & Discussion

- 2 Different data sets
  - A: Little background noise
  - B: Different elements and noise in background
- 3 Different models trained
  - A: Trained on data set A
  - B: Trained on data set B
  - C: Trained on data set B then on data set A
- Performance of models A and B was good with mAP values of 0.9948 and 0.9380 when tested against respective validation sets. Performed poorly against opposite test sets
- Model C had the best results. Where it performed well against set A and less good against set B. Although, outperformed both models A and B. See Table 1 and 2.

Table 1: Model C results of TP, FP, FN and mAP@0.5

Data set	TP	FP	FN	mAP@0.5
A: validation set	807	4	9	0.9948
B: validation set	772	157	492	0.4797

Table 2: Average pixel distance of model C

Data set	Avp	Var	Std
A: validation set	0.8938	0.3784	0.5327
B: validation set	7.9719	46.2933	5.8924

5. Conclusion

The possibility of using an object detection algorithm YOLOv5, to precisely detect optical reflective spheres using images from the HL2 gray scale cameras was explored. The results gathered show that YOLOv5, a state-of-the-art object detection algorithm, can be used to precisely detect spheres. The trained models showed promising results when tested against validation sets. However, are limited when exposed to other data sets.

6. Future work

- Train model on a combined dataset with images from both data set A and data set B. See if the model is able to accurately detect spheres in both data set.
- Improve annotation on data set B as this uses a fixed bounding box for all spheres in this data set.
- Train data sets using YOLOv4 and compare the results with that of YOLOv5.