MECHANICAL ENGINEERING STUDENT SEMINAR

Monday, January 13 2025 at 10:00, D. Dan and Betty Kahn Building, Room 217.

Object Identification, Segmentation, and Dimension Extraction from RGBD Images Using Deep Learning

Elran Mizrahi

Adviser: Prof. Anath Fischer

This research proposes a comprehensive methodology for identifying, segmenting, and extracting dimensions of objects from RGBD images in real-time using deep learning techniques. The system integrates two primary components: a segmentation model for detecting and extracting objects, and a regression-based dimension extraction model to predict the dimensions of the detected objects. While the methodology is applicable to general objects, cylinders were chosen in this work.

The segmentation was implemented by using the YOLO algorithm, which operates on RGB channels and enables real-time instance-based segmentation. For dimension extraction, a ResNet architecture model was selected, modified to process RGBD inputs with fusion techniques, and was optimized for predicting cylinder dimensions through regression.

Datasets for this study were created by using synthetic images which were generated in the Blender software. Moreover, real RGBD images were captured with an Intel RealSense Lidar Camera L515. Synthetic datasets included cluttered scenes with cylinders partially occluded by geometric primitives, while real datasets incorporated common objects in complex scenarios. Augmentation techniques were applied to ensure robustness across datasets. Fine-tuning on real-world data further enhanced the system's feasibility.

The system was evaluated through an integrated pipeline that combined segmentation and dimension extraction. The feasibility was demonstrated on cylindrical objects in complex scenes in real-time.

The segmentation model achieved a mAP50 of 0.99 and ran at 15 FPS. The dimension extraction model reached 4.6% MRE for diameter and 14.6% MRE for height after training on real-world RGBD images.

This study highlights the integration of synthetic and real-world datasets for model training and the exploration of RGBD fusion for feature enrichment. The developed system provides a framework for object detection and dimension analysis for industrial applications

Note: the seminar will be given in Hebrew

`