

# REU Project: Real-Time Multi-Camera Object Detection and Similarity Matching for 3D Mapping

Emilio Montoya, David Ramirez, Andreas Spanias  
Arizona State University



## Goal

- Simultaneous live video feed from 2 cameras will run through an object detection CNN.
- The use of ORB points for feature matching across simultaneous live video feed from two cameras
- The use of a 360° camera for a broad POV for one camera
- Develop a room mapping algorithm which incorporates both object detection and feature point mapping.

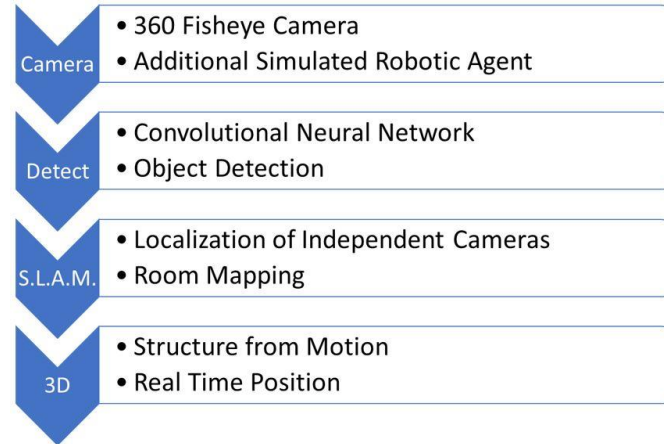


Figure 3. 3D Room Mapping SLAM Algorithm Output

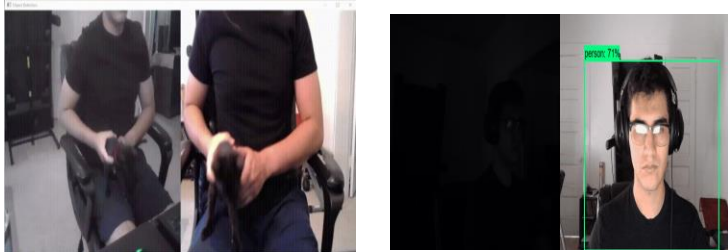


Figure 1. CNN module detecting person and cat with multiple cameras



Figure 2. Feature Point Similarity Matching (yellow lines)

CNN Module Type	Average Inference Time from Cameras
CenterNet HourGlass 104 KeyPoints 512x512	5.6910 seconds
SSD MobileNet v2 320x320	0.1129 seconds
EfficientDet D0 512x512	0.8080 seconds
SSD ResNet152 v1 FPN 640x640 (RetinaNet152)	14.8599 seconds
Faster R-CNN ResNet152 v1 640x640	22.9677 seconds

Table 1. Average frame per second of CNN TensorFlow modules performing object detection from a live video source

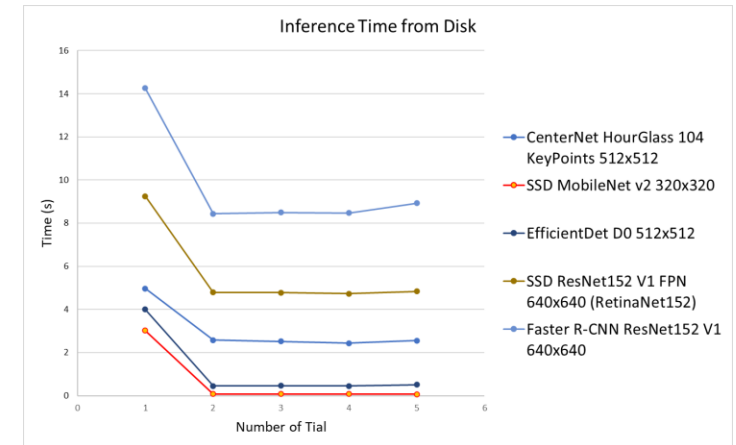


Chart 1. Comparison of CNN TensorFlow modules performing object detection from images