COE379L - Project 04 Proposal
Martin Infante, mbi244
Version History:

- Predator-Prey Danger Detector deployed on Multi-Model Inference Server (old)

Danger Proximity with Nearest Civilian Validation in response to Armed Threats

Introduction

My proposal aims to develop an inference server capable of identifying armed threats in images/videos, assessing the proximity of civilians to those threats, and highlighting the civilian nearest to said perpetrator. Hopefully, as AR glasses develop, the inference server will be able to connect to smart glasses to provide officers with an optimized awareness of a perpetrator's whereabouts when in crowded environments. The project will leverage object detection and monocular depth estimation to find midpoints of objects and approximate the objects' actual distances from one another. This project also plans to build upon the research paper titled "Detection of weapon possession and fire in Public Safety surveillance cameras" [1], but will focus more on the linkage of detected weapons to detected persons via IoU logic of bounding boxed objects.

Datasets

The following two datasets contain images of guns used in hostile environments.

1. GitHub - MarcusLimJunYi/Monash-Guns-Dataset
2. GitHub - Deepknowledge-US/US-Real-time-gun-detection-in-CCTV-An-open-problem-dataset

Technologies

1. For objection detection of guns and humans, I will follow suit with the research paper and use the YOLOv5 architecture due to its specialization in real-time applications.
2. For estimating danger proximity, a neural net model using monocular depth estimation will be used. Midpoints of bounding boxes for the armed person and surrounding civilians will be calculated. Then the pixel distance between these midpoints will be scaled using the depth map generated by the monocular depth estimation model, as a means of approximating real-world distances.
3. IoU logic of bounding boxed objects will also be applied, however the research paper did not specify their methodology when applying it to the model.

Products to be delivered

A 2-stage model inference server that will be able to handle POST requests, and returns a JSON formatted output classifying civilians, armed threats, and approximated real-world proximities.

```JSON
{
  "threat_detected": ["Yes", "No"]
  "threat": ["person#", null]
  "threat_bbox": [x1, x2, y1, y2],

  "civilian_count": int
  "nearest_civilian": ["person#", null]
  "nearest_civilian_bbox": [x1, x2, y1, y2]

  "nearest_dist_from_threat": int
}
```

References

[1] N. S. Moura, J. M. Gondim, D. B. Claro, M. Souza, and R. de C. Figueiredo, "Detection of weapon possession and fire in Public Safety surveillance cameras," in *Anais do Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, Nov. 2021, doi: 10.5753/eniac.2021.18261.