Object Detection Modeling with Thermal Imaging

Team Members: Aditya Desai, Linus Fackler, Cesar Gamez, Aryan Punjabi, Muhammad Zaid
Research lead: Abhishek Mishra Faculty Advisor: Dr. Karen Mazidi

Introduction

Since the early 2000's, the automated transportation industry has been booming in innovation and interest. This interest in automated modes of transportation has also raised concerns regarding its safety as more and more fatal incidents occur. These concerns are being addressed by an increasing focus on machine learning to aid this new technology. One such concern is detection of objects while facing poor lighting conditions.

To address this issue, we present YOLOvCAPY (Computer Analysis of Photos with YOLO)- a machine-learning approach to analyzing and detecting objects within thermal pictures and live feeds. This approach is more efficient, accurate, and robust than the current models for thermal detection allowing for better visualization of objects.

Data Collection

We collected our data from Teledyne FLIR, a thermography company that has released a data set of thermal images and videos associated with varying driving conditions. We were able to access our dataset through a direct download of the content provided by the company. Many frames were fully annotated to display 15 categories: person, bike, car, motorcycle, bus, train, truck, traffic light, fire hydrant, street sign, dog skateboard, stroller, scooter, and other vehicles. By using this dataset to train YOLOv7's default capabilities, we were able to effectively and accurately detect a multitude of object categories beyond the 15 annotated on the FLIR Dataset.

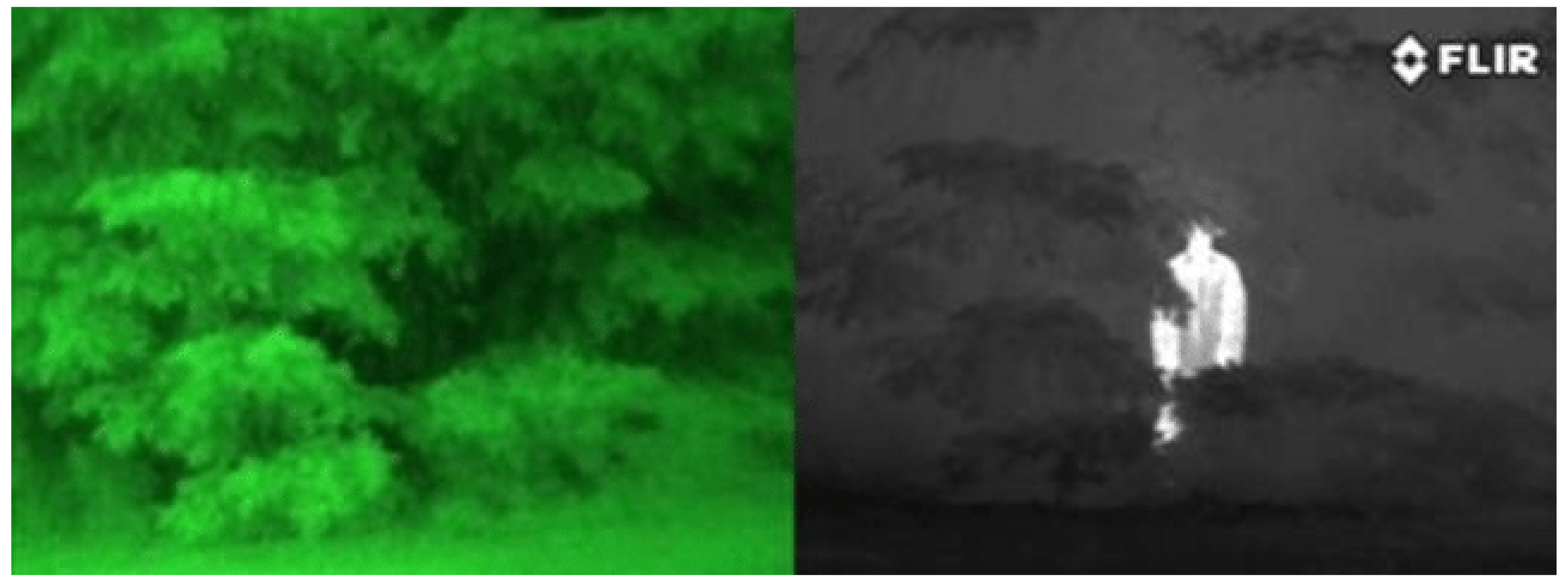

Figure 1. Comparison between normal and thermal imaging provided by Teledyne FLIR

The Flir dataset was clean and contained more than 26442 frames of phots and videos, allowing for effective machine learning. To process this amount of data we employed cloud-based GPUs to allow for efficient and accurate results.

Model

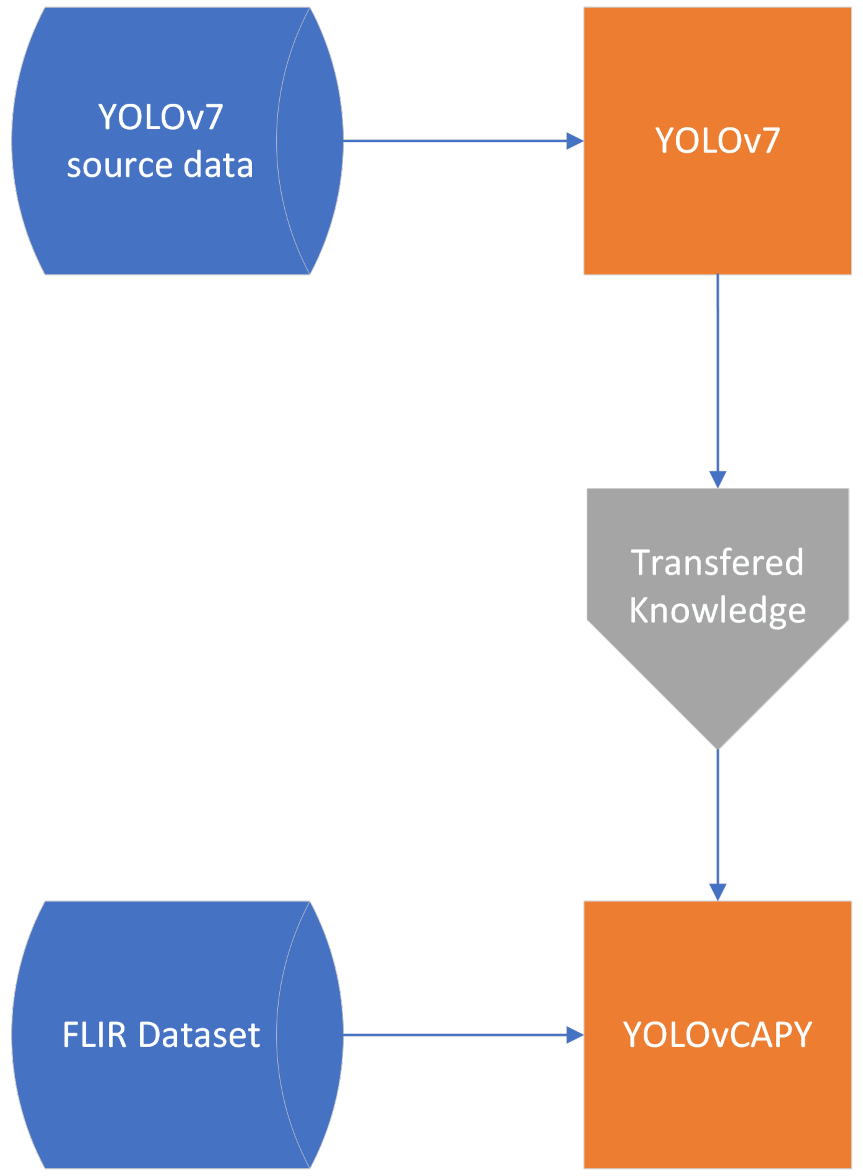

Figure 2. Transfer Learning Diagram

- **Transfer Learning**: Transfer learning is a form of AI model creation that tends to take less time to train and allows for more refinement. Essentially, an existing model is trained on the target dataset and the resulting model is refined more to work with the niche elements of the target dataset.
- **YOLOv7**: The source model utilized in the creation of YOLOvCAPY is YOLOv7. We chose this source model due to it being the newest version of one of the most powerful Convolutional Neural Network Models for object detection.
- **Development of YOLOvCAPY**: We employed transfer learning to train YOLOv7 on the FLIR Dataset and create YOLOvCAPY, which we refined to work as accurately and effectively as possible with given hardware capabilities.

Results

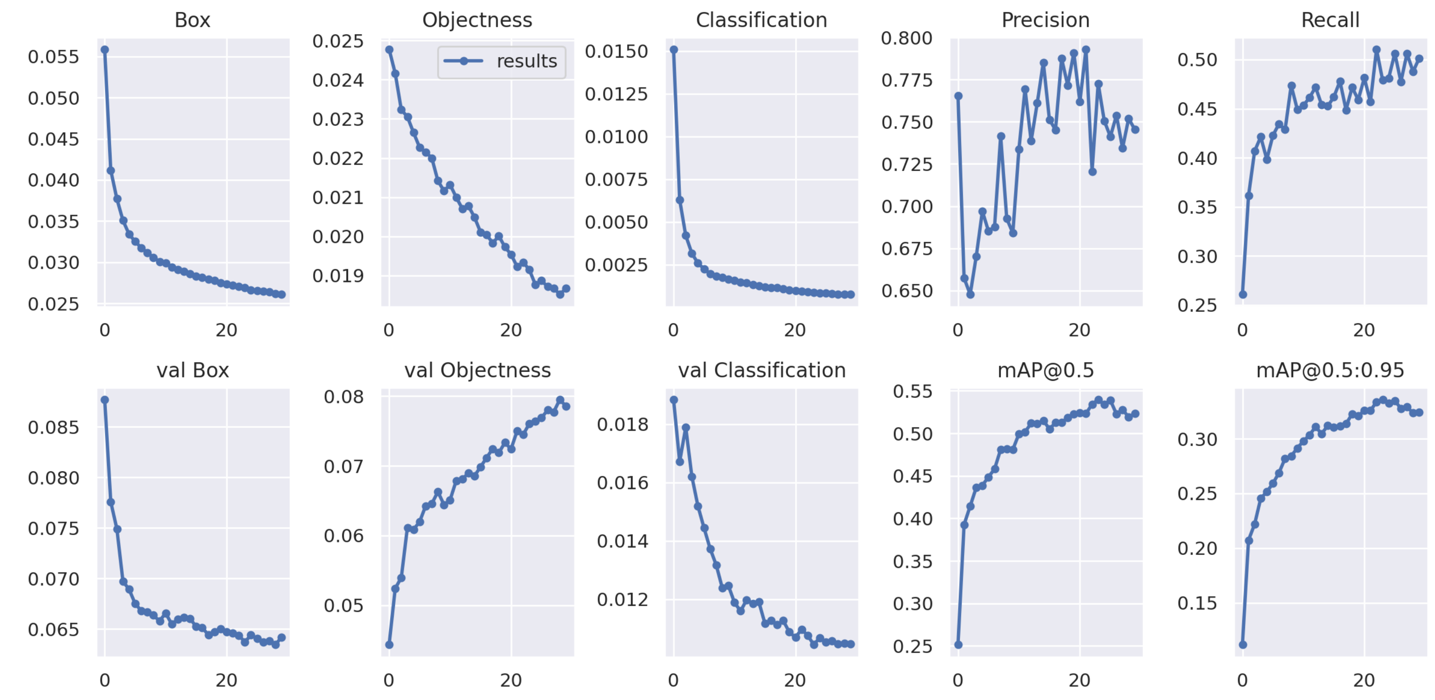

Figure 3. Resulting Metrics of YOLOvCAPY

Analysis and Comparison

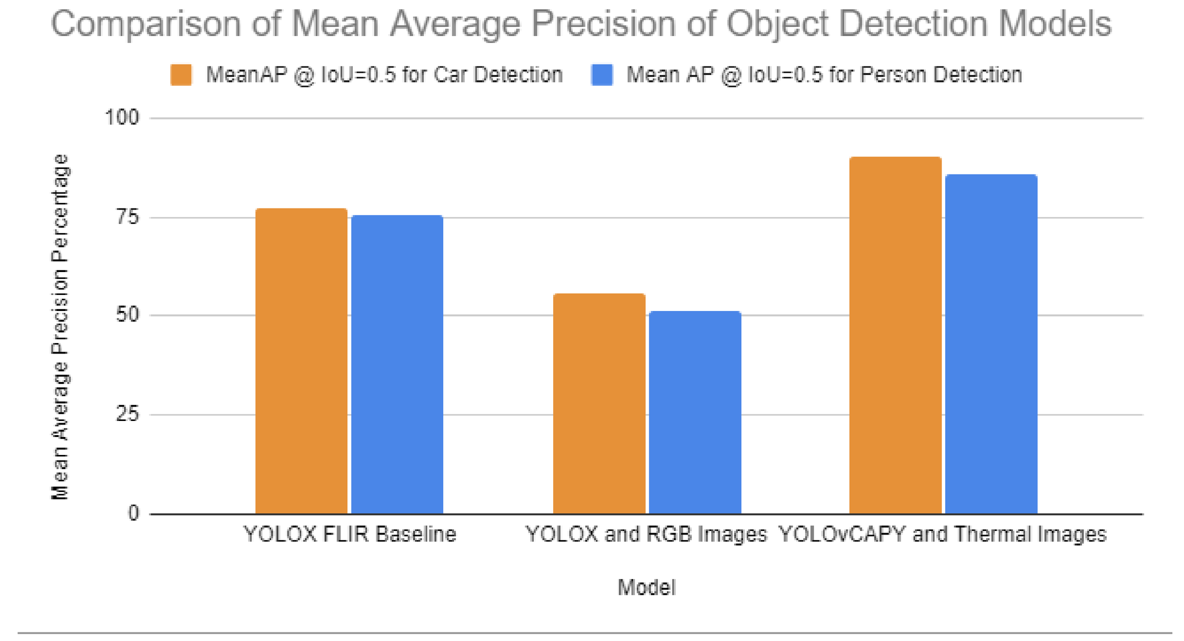

Figure 4.

As demonstrated by graph, YOLOvCAPY displays a significantly higher Mean Average Precision at an Intersection over Union of .5 (mAP @ IoU .5) for the detection of cars and people in comparison to other models. This higher mAP @ IoU .5 reveals that the developed YOLOvCAPY model is more accurate and precise than other popular models.

Future Endeavors

With the success of our model, it is understood thermal images could be used effectively for object detection in poor lighting conditions and improve current video camera systems in vehicles. To better the relevance of our model for future works, we would improve:

1. **Data diversity**: The quality of our model's effectiveness and accuracy was highly limited to the FLIR Dataset and is something we would like to work on in the future to improve its relevance. Despite the extent of the data, we would like to include some edge cases to make our recognition algorithm more accurate.
2. **Features**: We would add more features to our model in order to improve its usability in the industry. Some of these features include detection of distance between objects and the source of the video and velocity of said objects in a certain direction towards or away from the camera.

References

Carion, Nicolas, et al. End-To-End Object Detection with Transformers. Moses, Daniel A. "Deep Learning Applied to Automatic Disease Detection Using Chest X-Rays." Journal of Medical Imaging and Radiation Oncology, vol. 65, no. 5,