Christian Mpabuka

ITAI 1378 Computer Vision

Professor: Patricia McManus

L09 Object Detection using Transfer Learning and Pascal VOC Dataset.

**Image classification aims to correctly assign a label to an image. In contrast, object detection identifies and locates specific objects within an image by drawing bounding boxes around them, providing more detailed information. This technique is utilized in applications such as self-driving cars and surveillance systems.**

The threshold value defines the minimum confidence level required to display a detection. A higher threshold means that only highly confident detections will be shown, which reduces the number of objects displayed but may increase accuracy. On the other hand, a lower threshold will display more objects, including some that could be false positives.

When running the exercise multiple times, it appears that the model is more accurate in identifying the person on the horse and the entire black car. However, due to its irregular patterns, the toucan bird tends to camouflage, making it difficult for the model to accurately place a bounding box around the bird. The toucan blends into its surroundings so well that the model struggles to distinguish it from the background. Overall, the bounding box effectively identifies most objects, but when an object blends well with its environment, the identification accuracy decreases.

Using the full Pascal VOC dataset would likely enhance the model's accuracy because it would provide increased data diversity, reduce overfitting, improve model capacity, enhance robustness, and yield more reliable evaluations. However, this improvement might come at the cost of higher computational requirements, which my computer may not be able to handle using the free service.

Detecting Specific Objects To detect specific objects such as animals or vehicles, I can either train a new model using a custom dataset of the target objects or fine-tune a pre-trained model like YOLOv8 or Efficient on a specific dataset. Additionally, I can filter the output by class to retain only the desired object classes and adjust the confidence threshold to reduce false positives.

There are several steps involved in training an object detection model. The first step is data preparation, which includes collecting, annotating, and splitting the dataset. Next, I would select an appropriate model, such as YOLOv8 or Efficient. Then, I would train the model on the dataset. After training, the model undergoes evaluation and fine-tuning to enhance its performance. To develop the performance model, I would first export and deploy it. Some challenges I may encounter include ensuring data quality and quantity. Operating on a budget can limit access to high levels of computational resources. Other challenges include model complexity, data imbalance, real-world variations, and the tuning of hyperparameters.