Tolulope Ogundeji

ITAI 1378

Object detection with transfer learning

With Pascal VOC2007 dataset

Professor McManus

**Reflection**

In this assignment, I delved into the world of object detection by working with the PASCAL VOC dataset to train a model in identifying objects within images. I have a good understanding of the differences between image classification and object detection. Image classification is basically the process of assigning a label to an entire image based on its content. Object detection is a task that involves identifying items within an image or video. Object detection is special due to bounding boxes, it uses these to provide precise locations of items in an image.

In terms of the difference in results, the output shows that the model made attempts at identifying different items within each image. Unlike image classification models such as VGG that gives a label to the entire image.

SSD MobileNet V2 is a good choice for tasks like this where computational power is limited, it provides reasonable speed and accuracy without needing too much power. However, it does have its limitations and disadvantages, such as low accuracy on small objects due to its feature extraction approach, and sometimes due to its lightweight design can deliver poor performance. The user must pay attention to detail so reasonable results can be achieved.

The role of ‘find\_images\_with\_classes’ function is to find specific images with certain classes. It is useful when the goal is to focus on certain categories rather than the entire dataset. It is useful in large datasets like COCO (Common objects in context), because running the function can reduce time and still get the desired outcome.

In the plot\_detections function, the threshold value (e.g., threshold=0.5) controls the number of detected objects displayed by filtering detections based on their confidence scores. In this notebook, it was initially set to 0.5, meaning that only predictions with confidence scores greater than or equal to 0.5 will be displayed.

The heatmap is a clear method of interpreting the model’s confidence in its detections, and this relates to why the model identifies large objects easily. The challenge is identifying small objects, because they occupy less pixel space.

My model got most of the identifications wrong, I attempted various tweaks such as changing the confidence threshold, IOU threshold, and training it repeatedly, but could not get it to perform better. My assumption is data complexity, or incompetent model. SSD MobileNet is designed for speed and not for accuracy, a different model might perform better.

I would expect the accuracy of the model to be much better, because it will learn a wider variety of features and labels which will help it generalize and reduce chances of overfitting. Like it said, the more the merrier.

If I wanted to modify the code to detect a specific set of objects, I would focus on the relevant classes, such as ‘cat’ ‘dog’ ‘horse’, etc. Then when loading the dataset, will filter it by modifying the data loading function and creating a custom dataset. Part of the code will look like this:

def load\_filtered\_voc\_data(root, year='2007', image\_set='train', transform=None):

dataset = VOCDetection(root, year=year, image\_set=image\_set, download=True) filtered\_data = [].

Despite the limitations of the model, it can be properly trained and be useful in some real-world scenarios. Examples include basic surveillance, fun projects, sports analytics, and drone applications. These are applications where the cost of error is rather insignificant and when properly trained and tested, it can be a huge time saver.

TensorFlow has a good variety of models in its library, such as Faster R-CNN. There are lots of models out there and they are used for various tasks due to their design. SSD MobileNet is good for speed and reasonable accuracy, while Faster R-CNN is known for being slower with high accuracy and being a lot more energy consuming. YOLOv5 is also known for being quite fast and is good for identifying objects in real time. It is graded as between MobileNet and Faster R-CNN.

Works Cited

CHOUDHARY, ANURAG SINGH. “Object Detection Using YOLO and Mobilenet SSD Computer Vision -.” *Analytics Vidhya*, 22 Sept. 2022, www.analyticsvidhya.com/blog/2022/09/object-detection-using-yolo-and-mobilenet-ssd/.

Gallagher, James. “The Ultimate Guide to Object Detection.” *Roboflow Blog*, 22 Aug. 2023, blog.roboflow.com/object-detection/. Accessed 30 Oct. 2024.