Assignment 3- Object detection and Segmentation

Part 1- Object segmentation

Object detection is a task in computer vision that involves identifying the presence, location, and type of one or more objects in a given image. In order to have an accurate object detection model, we need to train a model based on thousands or millions of images related to that object, and this makes the task even more challenging. As was mentioned in week 2, even after finding appropriate data, training the model is not an easy task. We discussed one solution which was using pre-trained models. COCO is a large-scale object detection, segmentation, and captioning dataset.

        a) Load the image dataset oxford_iiit from tf
        b) Create the segmentation mask on the first 3 samples in the dataset. Show both the image and segmented version of it. Remember to map the images into 128x128.


Now we can start designing the model. We stop this part here and leave the rest of it to Lab 3.1.

Part 2- Annotation

One important aspect of data segmentation and object detection is to have the annotation of the images. In this part, we want to look into annotations of a sample dataset. Download the dataset from here:

```
https://github.com/experiencor/raccoon_dataset
```

a) Like before, show some samples of the images.
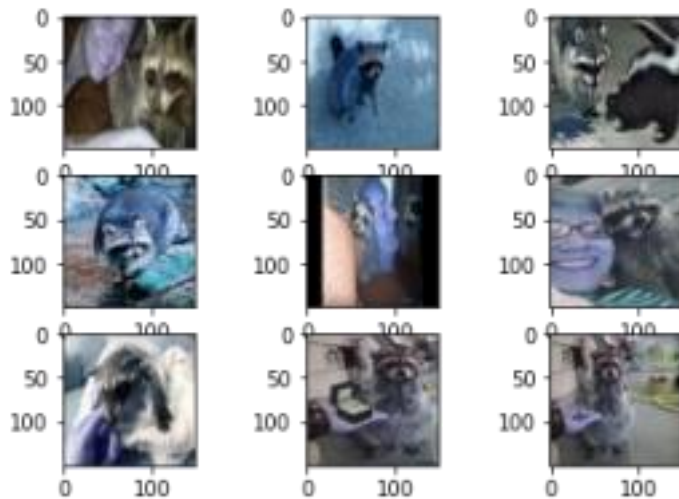
b) A class of functions is provided for you that creates masks and images. Use the provided class and create a sample of the image and its mask.

Part 3- YOLO V5

In 2016, researchers at Washington University, Allen Institute for AI, and Facebook AI Research proposed "You Only Look Once" (YOLO), a family of neural networks that improved the speed and accuracy of object detection with deep learning.

The main improvement in YOLO is the integration of the entire object detection and classification process in a single network. Instead of extracting features and regions separately, YOLO performs everything in a single pass through a single network, hence the name "You Only Look Once."

In this part of Lab, we use the latest version of YOLO (Yolo 5) to create the object detection task.

a) For the first part of the assignment, we want to see how to create the annotation for a custom dataset. You can use the provided html file (annotation_tool.html). When you open the file, upload an image of your desire and assign the proper class. Here you see an image of a raccoon and its chosen class.

b) Download required yolo5 tools and check dependencies. Now we download a pre-trained model which will be used for our object detection down the road.

```
!python detect.py --weights yolov5s.pt --img 640 --conf 0.25
-- source data/images
```

Let's check our model now. Use following commands to see how the object detection works.

```
import utils
display = utils.notebook_init() # checks
display.Image(filename='runs/detect/exp/bus.jpg', width=600)
```

There are two images in the folder, but I recommend you try different images and see the result.

c- Use train.py code to train a YOLO model. Use 640 images from COOC, in 16 batches and 3 epochs. You can use coco 128.yalm for training instruction (YOLOV5/data) and the pre-trained weights are coming from yolov5s.pt

d- For testing the model, copy any image (for Example Bike.jpg already shared with you) into data/images and run detect.py to evaluate the model. You should see the result of analysis as:

```
image 1/3 /content/yolov5/data/images/Bike.jpg: 448x640 1 person, 1
bicycle, 8 cars, 2 trucks
```

e- The results of these analyses are shared in runs/detect/exp folder (depending on how many times you run this code, you will see different numbers for the number of exp). From the utils package, use Image to see the result of analysis. Explain your understanding from this image. What is the meaning of these numbers?