Introduction

Object Detection is a task of Classification, that specifies identification of each object from the image that belongs to a certain class and Localisation, that specifies the location of an object in the image. In this project we opted for the Mask R-CNN (Region Convolutional Neural Network) which is an extension of Faster R-CNN widely used for Object Detection. Now the question arises-why use Mask R-CNN? Mask R-CNN is intended for predicting segmentation mask on Region of interest (RoI) in pixel to pixel manner whereas the Faster R-CNN is not designed for pixel to pixel alignment between network inputs & outputs.

Image segmentation is a pixel-wise mask for the object that helps in simplify or representation of an image, which assists easier to analyse more efficiently. The image segmentation are of two types Semantic Segmentation & Instance Segmentation. Semantic Segmentation specifies all the objects of the same class in the image as one instance whereas Image Segmentation specifies all the objects of the same class are assigned as different instances, which further helped us in having optimal outcomes.