**Wobot.ai Assignment**

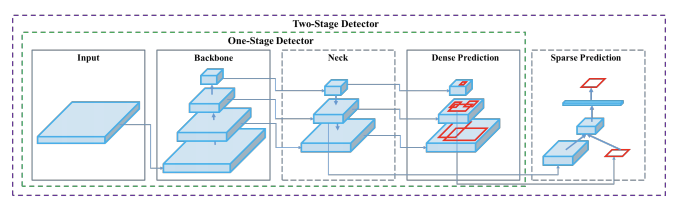
**Task: Build A Personal Safety Equipment Detection System**

In this assignment, I have built an object detector for the given Hard Hat/ head detection dataset.

**Model Used:**

**YOLOv5:**

The YOLO model was the first object detector to connect the procedure of predicting bounding boxes with class labels in an end to end differentiable network.



The YOLO network consists of three main pieces.

1) **Backbone** - A convolutional neural network that aggregates and forms image features at different granularities.

2) **Neck** - A series of layers to mix and combine image features to pass them forward to prediction.

3) **Head** - Consumes features from the neck and takes box and class prediction steps.

There are many approaches one can take to combining different architectures at each major component. The contributions of YOLOv4 and YOLOv5 are foremost to integrate breakthroughs in other areas of computer vision and prove that as a collection, they improve YOLO object detection.

**Training Procedures**

**Data Augmentation -** Data augmentation makes transformations to the base training data to expose the model to a wider range of semantic variation than the training set in isolation.

**Loss Calculations -** YOLO calculates a total loss function from constituent loss functions - GIoU, obj, and class losses. These can be carefully constructed to maximize the objective of mean average precision.

**—------------------------------------------------------------------------------------**

**Stage 1:**

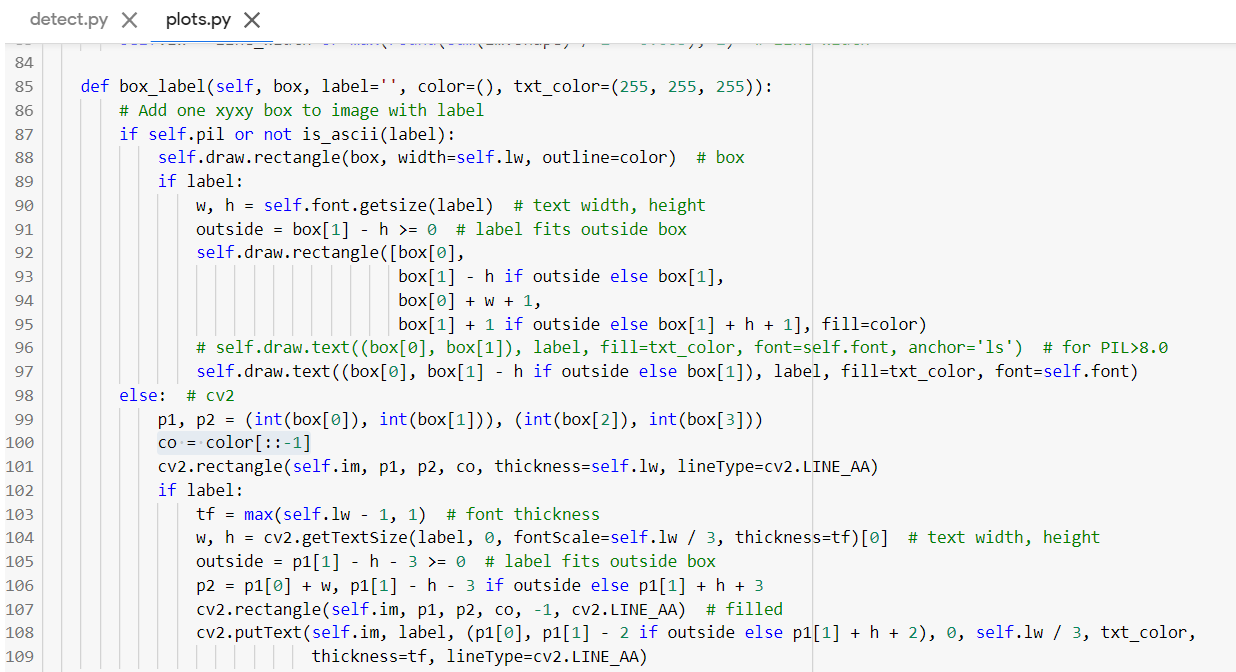
* In stage 1, I was asked to train the model on the given hard\_hat/head dataset.
* As yolo takes the annotations in .txt format, I selected equal ratios of hard\_hat/head images(48 images for training and 18 for validation) from the given dataset and annotated them using **LabelImg.**
* Then I trained the YOLOv5 model( which was pre-trained on coco dataset) with my dataset and runned it for various combinations of batch sizes and epochs.
* The best results were obtained with **400 Epochs** with **Batch Size = 4.**
* The **mAP** (mean average precision) for **Head: 0.949 (94.9%)**.
* The **mAP** (mean average precision) for **Helmet: 0.830 (83.0%)**.
* After that I did the inference on the given video and found **promising** detections.
* Output video is stored in [Stage 1](https://drive.google.com/drive/folders/1zY0inZHtkd61d0_3j_jZ2bj3fHyNUVQx?usp=sharing) folder named ‘Stage1\_Output.mp4’.

**Stage 2:**

* For this stage, I was asked to change the color of the bounding boxes according to the color of the hat.
* For this I have gone through two different files which were present in the YOLOv5 model namely, **detect.py** and **plots.py.**
* In detect.py, I wrote some lines of code which will first create an object of the source image and then save the bounding box coordinates (xmin, ymin, xmax, ymax) of the two tuple variables p1 and p2.
* Further calculated the center coordinates, shifted the y-coordinate a little above(the pixel will direct to the helmet) and using the **getpixel()** function, read the RGB values of that particular pixel. Passed the ‘**col’** variable in annotator.box\_label().



* Further in plots.py, in the box\_label function, two methods were defined to create the box labels. Since PIL uses **RGB format** only, no changes have to be made there(**if** **section**). But CV2 uses **BGR format** so there in the **else section** I just reversed the color tuple variable.



* With this I was able to achieve the colored bounding boxes according to the helmet color.
* Same thing I have to do with video by giving each frame to create the image object but I was unable to do so. This amendment works fine for images though.
* These images are also present in the [Stage 2](https://drive.google.com/drive/folders/1k52h1b2WTtmEhJsSeg-CJwt9x6Pubnxc?usp=sharing) folder.

**Stage 3:**

* In stage 3, I was asked to test my model of the given 250 Test\_Images.
* Since YOLOv5 uses .TXT format I was not able to convert the predictions into .XML format.
* I tried various methods to change the format but couldn’t achieve success.
* I have shared the
  + Test images along with the predictions. (Test\_Images)
  + Zip file of these predicted images. (Test\_Images.zip)
  + Zip file containing the annotations in .TXT format. (Annotations(.TXT Format).zip)
* It contains all the detections of the hard hat helmet and head.
* The images and the zip files are present in the [Stage 3](https://drive.google.com/drive/folders/1jnybdE0CkUl1fEbwh1ci91iGNQ8ksldq?usp=sharing) folder.

This sums up all the work done in this assignment by me. I am aware of the fact that I was not able to submit the results in the desired formats. I highly apologize for that. I tried each and every aspect to fulfill all the needs and this was the best I could perform. I am aware that it will affect my evaluations but I managed to make it up to a certain level. This project helped me to gain experience in a variety of concepts for which I will be truly thankful.

Best Regards

Sarvesh Singh