DEEP LEARNING FOR COMPUTER VISION - ASSIGNMENT 02

Write about any one YOLO version.

Aneeta Sara Shaw
MTech CSE
Software Engineering

April 2, 2023.

YOLOv5

On June 25th, 2020, the first official version of YOLOv5 was released by Ultralytics, a computer vision model used for detecting objects.
YOLOv5 is a model in the You Only Look Once (YOLO) family of computer vision models. YOLOv5 can be used to do object detection on the images of the dataset, basically image segmentation.
YOLO is a method to do object detection, which is the algorithm/strategy behind how the code is going to detect objects in image. YOLO looks at the entire image only once and goes through the network once and detect objects. Hence the name. It is very fast. That's the reason it has got so popular.

Object detection, a usecase for which YOLOv5 is designed, involves creating features from input and images. These features are then fed through a prediction system to draw boxes around objects and predict their classes. The YOLO network consists of three main components:-

(i) A convolutional neural network that aggregates and forms image features at different granularities. (Backbone)

(ii) A series of layers to mix and combine image features to pass them forward to prediction. (Neck)

(iii) Consumes features from the neck and takes box and class prediction steps (Head).

YOLOv5 for object Detection

(i) **Environment Setup** ⇒ → Enable GPU in Google Colab
→ Mounting your personal drive
→ Cloning YOLOv5 repository made & maintained by Ultralytics.

Input ⇒ !git clone https://github.com/ultralytics/yolov5.
→ Install required package in order to run YOLOv5 detector.

(ii) **Interference YOLOv5** ⇒ → python code initializes our detector and contains code to make predictions.
→ Path of image/video/youtubelink on which detection process is to be performed is provided.
→ weight file is provided for the model.
→ minimum confidence value for the model to consider a prediction as viable.

(iii) **Example of YOLOv5x** ⇒ → biggest model with 476 layers and 87 million parameters along with FLOPs value of 218.6 billion. (floating point operations)
→ Model is able to predict accurately even on complex images.

(iv) **YOLOv5 object detection on Youtube videos** ⇒ - added capability to perform object detection on a Youtube video by just passing the URL.

YOLOv5 is very user friendly and comes ready to use on custom objects out of the box. as it is a single-stage detector.