Project

190050055,190050071,190050079,1900500100

July, 2021

Object detection is one of the classical problems in computer vision where you work to recognize what and where — specifically what objects are inside a given image and also where they are in the image. The problem of object detection is more complex than classification, which also can recognize objects but doesn't indicate where the object is located in the image. In addition, classification doesn't work on images containing more than one object.

Bounding Box

A bounding box is an imaginary rectangle that serves as a point of reference for object detection and creates a collision box for that object.

Data annotators draw these rectangles over images, outlining the object of interest within each image by defining its X and Y coordinates. This makes it easier for machine learning algorithms to find what they're looking for, determine collision paths, and conserves valuable computing resources.

Bounding boxes are one of the most popular image annotation techniques in deep learning. Compared to other image processing methods, this method can reduce costs and increase annotation efficiency.

Applications of Bounding Box

Take self-driving cars as an example. An annotator will draw bounding boxes around other vehicles and label them. This helps train an algorithm to understand what vehicles look like. Annotating objects such as vehicles, traffic signals, and pedestrians makes it possible for autonomous vehicles to maneuver busy streets safely. Self-driving car perception models rely heavily on bounding boxes to make this possible.

Image annotation with the help of bounding boxes is also used to label objects seen from the viewpoint of robots and drones. Images that are annotated with the technique help robots and drones to identify the wide variety of objects seen on earth. The different types of objects that can be captured into the bounding box make it easier for drones and robots to identify similar physical objects from a distance and steer accordingly.

Non-max suppression

The objects in the image can be of different sizes and shapes, and to capture each of these perfectly, the object detection algorithms create multiple bounding boxes. Ideally, for each object in the image, we must have a single bounding box. Something that covers the whole object. To select the best bounding box, from the multiple predicted bounding boxes, these object detection algorithms use non-max suppression. This technique is used to "suppress" the less likely bounding boxes and keep only the best one. Thus leading to ideal senario from multiple bounding boxes case.

The Process The following is the process of selecting the best bounding box using non-max suppression

Step 1: Select the box with highest objectiveness score

Step 2: Then, compare the overlap (intersection over union) of this box with other boxes

Step 3: Remove the bounding boxes with overlap (intersection over union) $> 50\%$

Step 4: Then, move to the next highest objectiveness score

Step 5: Finally, repeat steps 2-4

Processing a video means, performing operations on the video frame by frame. Frames are nothing but just the particular instance of the video in a single point of time. We may have multiple frames even in a single second. Frames can be treated as similar to an image. So, whatever operations we can perform on images can be performed on frames as well.

Social distancing prediction

We have used different metrics to compute violation of social distancing between each pair of persons in the current view.

We have computed nominal distance metric by using the values of bounding box and relative positions of 2 people in the image. We created green boxes for the people in the image and red lines joining two people violating the social distance. We have efficiently extended this to videos which we processed using opencv.

we can also detect multiple objects apart from persons like car, hydrant, orange etc. We can also use this to extend for detecting efficient car parking system. which will be implemented similarly by computer vision.

Summary

During pandemic it is difficult to monitor social distancing of people. So we can monitor and detect people violating covid rules efficiently.

It can also be used to prevent further increase of people in particular area by computing density of people in that area.

This can also be useful for blind people in maintaining social distance and predicting the objects before preventing of any accidents.

A