Script to identify the corrected bounding boxes position by Carlos Cob.

Software designed to measure the accuracy of detection in systems that they have not an identification method. To achieve this goal has been required: a large data management through dictionaries and the use of methods such as IoU (Intersection over union) and GIoU (General intersection over union).

This script needs two dictionaries as input, one with the ground truth bounding boxes and other with the predicted bounding boxes. The ground truth bounding boxes dictionary is achieved from a public source or by ourselves. In the next figure it is possible to see the structure of this dictionary (Figure 1)

```
277:
  0: [75, 112, 103, 261]
  1: [339, 102, 42, 176]
  2: [414, 92, 70, 250]
  3: [476, 87, 86, 263]
278:
  0: [65, 113, 105, 254]
  1: [338, 102, 43, 176]
  2: [413, 92, 72, 250]
  3: [476, 87, 87, 264]
```

Figure 1: Ground Truth nomenclature

The dictionary has an index number that inidcate the frame number, within it has several dictionaries with the person identification (0, 1, 2, 3), this number is as large as the maximun number of people. Within of the person identification, it is the position in the screen, the meaning of these numbers are coordinates (u,v,w,h). This type of nomenclature is the used by several dataset but it is not the standard nomenclature. Every dictionary of this program is saved in a YAML file, because this kind of files can be used either, in different types programming languages or in other programs.

On the other hand, it is necessary to save the predicted bounding boxes data in another YAML file. This file saves the data in the same way that the ground truth bounding boxes file but with a certain difference: the nomenclature is the standard used in computer vision (xA, yA, xB, yB), being xA and yA the coordinates from the up left square and xB and yB the coordinates from the down right square. In the figure 2, it is possible to see the structure of this dictionary.

```
114: {0: [247, 188, 330, 383]}, 115: {0: [246, 187, 329, 383]}, 116: {0: [247,
    190, 325, 385]}, 117: {0: [245, 190, 319, 386]}, 118: {0: [244, 189, 317,
    386]}, 119: {0: [244, 188, 314, 387]}, 120: {0: [242, 189, 310, 363]}, 121: {
  0: [243, 189, 310, 388]}, 122: {0: [242, 188, 307, 392]}, 123: {0: [242, 187,
    305, 362]}, 124: {0: [240, 198, 307, 375]}, 125: {0: [236, 201, 297, 383]},
126: {0: [233, 199, 297, 383]}, 127: {0: [229, 199, 295, 384]}, 128: {0: [226,
    196, 294, 384]}, 129: {0: [207, 190, 295, 386]}, 130: {0: [203, 193, 294,
    387]}, 131: {0: [201, 193, 294, 388]}, 132: {0: [199, 194, 294, 388]}, 133: {
  0: [203, 192, 294, 386]}, 134: {0: [202, 191, 294, 387]}, 135: {0: [204, 191,
    292, 387]}, 136: {0: [209, 208, 288, 385]}, 137: {0: [209, 207, 285, 386]},
```

Figure 2 :Predicted nomenclature

Once these two dictionaries are ready to be used it is possible to apply the IoU and the GIoU .

These methods are used to check the percentage of common area between two bounding boxes. In this case, the two areas compared will be the ground truth bounding box and the predicted bounding box. Thus, if the IoU or GIoU of both areas is superior to a X percentage, that means that the detector has a true positive.

But what is the IoU and GIoU?

The IoU or also called Inference Over Union is an evaluation metric used to measure the accuracy of an object detector on a particular dataset. The typical and most popular dataset for this type evaluation is PASCAL VOC. To apply the intersection over union the following tools are necessary (figure 3):

- **The ground-truth bounding boxes**: Normally these pictures are hand labeled and they are used to calibrate and check systems.
- **The predicted bounding boxes**: They are the bounding boxes drawn by the system that it will be evaluated.

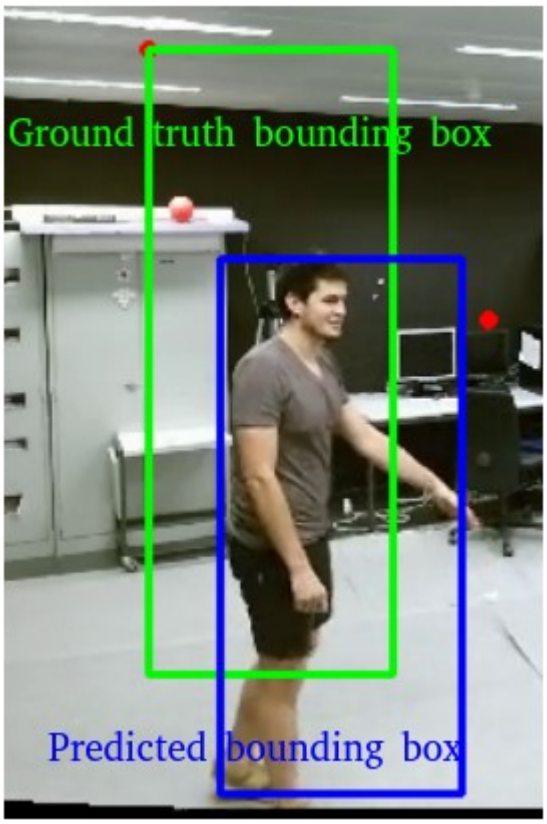

Figure 3: Ground truth bounding box vs Predicted bounding box

The IoU would be the Area of Overlap divided by the Area of Union, which is the area enclosed between the two bounding boxes. In the figure 4, it is easier to understand this concept:

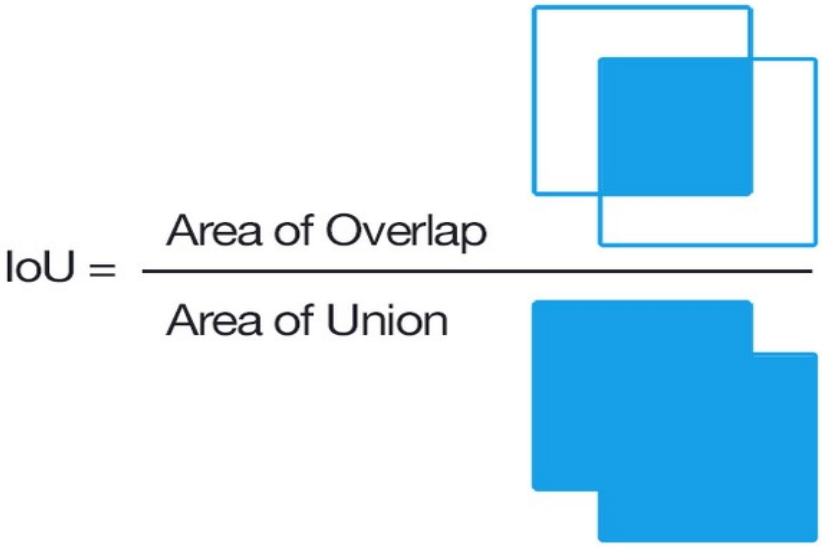

Figure 4: Intersection over union

With this in mind and with this IoU result, it is possible to link a predicted bounding box with a ground-truth bounding box when there is a high IoU and thus, to guarantee that the predicted bounding box is in the correct place. This is a reliable form to test if the system detects correctly .

IoU method provides good results, but also, a remarkable problem: when there are more of one bounding box in the system and it exists overlap between bounding boxes it is difficult to obtain a correct result even with the use of IOU. For this reason, it is necessary to use a trusted method. In the literature exists a newer version of the IoU, which is called GIoU (General Intersection Over Union). This method is explained in the next article [1] and uses the IoU but also it adds a normalized component that achieves better results that IoU. In the figure 5, it is possible to difference the loss function IoU and GioU .

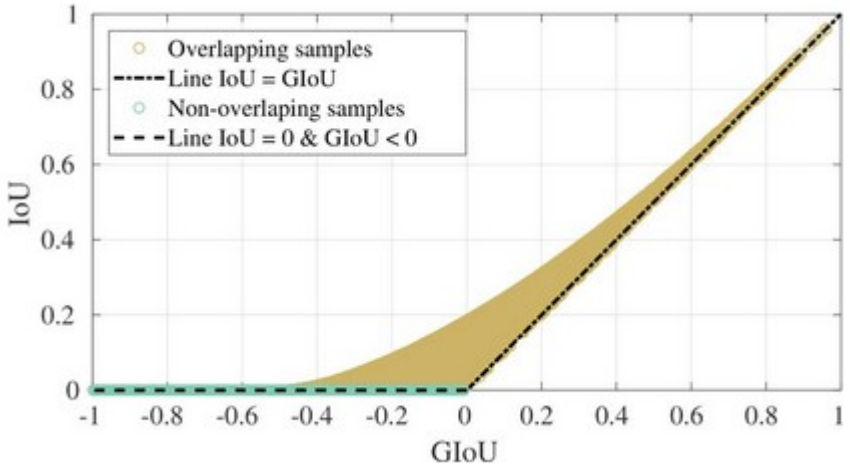

Figure 5: IoU vs GIoU

These mechanisms will be used to generate a metric of the values obtained in the detector.