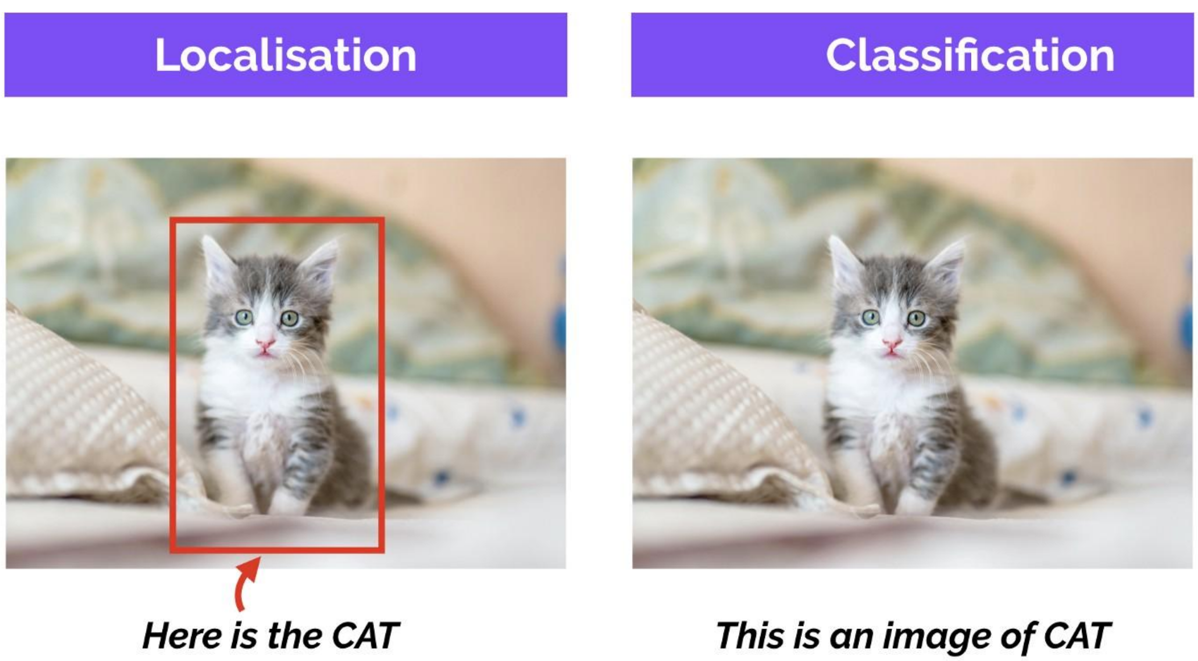
When I first started exploring the world of deep learning in my college days, I found that there were hardly any resources that could help me understand complex concepts in a simple way. My frustration inspired me to write this book. I hope my humble initiative will help students digest the complicated components of computer vision in a very easy and straightforward manner.

**Mean average precision (mAP) for Object Detection in Computer Vision :**

I assume you have already gained a basic understanding of object detection from the previous chapter and are now ready to dive deeper. If not, make sure to read the previous entries.

There are different models for object detection, so how to know which model performs better than the other? Here, comes Mean average precision or mAP into play.

mAp is the most meaningful metric which is used to evaluate object detection (Localization and classification) models such as Fast R-CNN, YOLO, Mask R-CNN, etc, and tell which model is closer to reality. Localization shows the location of an instance (e.g. bounding box coordinates) and classification determines what it is (e.g. a dog or cat).




Cat image: Photo by Kote Puerto on Unsplash

The mAP can be computed by calculating average precision (AP) separately for each class, then the average over the class. A detection is considered a true positive only if the mAP is above 0.5. metric

**Quick recap** :

**Precision & Recall :** If someone asked you to list the names of 5 presents you got on your last birthday but you couldn't exactly remember the 5 names - so you randomly guessed seven times. Out of 7 names you remembered, 5 were recalled correctly while 2 were gifts you received on last christmas. Even though you had a 100% recall (5/5), your precision was 71.4% (5/7).

Similarly, in Deep Learning, Precision estimates how many of a classifier's predictions are actually right and Recall indicates how many ground truth labels your classifier is able to pick up.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$TP = \text{True positive}$

$TN = \text{True negative}$

$FP = \text{False positive}$

$FN = \text{False negative}$

**True Positive (TP):** Bounding has an overlap with the ground truth higher than the set threshhold.

**False Positive (FP):** Bounding box does not overlap well enough with the ground truth.

**False Negative (FN):** A ground truth not detected although the image actually contains the object.

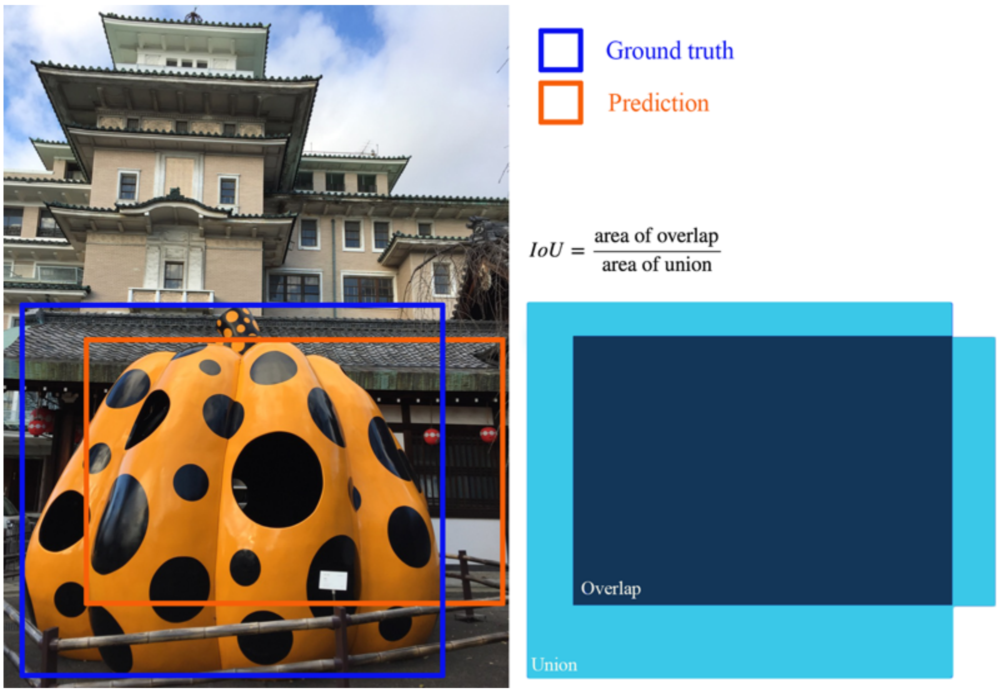
**True Negative (TN):** Non-detection of bounding boxes. For example, no bounding box in a free space that also has no objects and no ground truth.

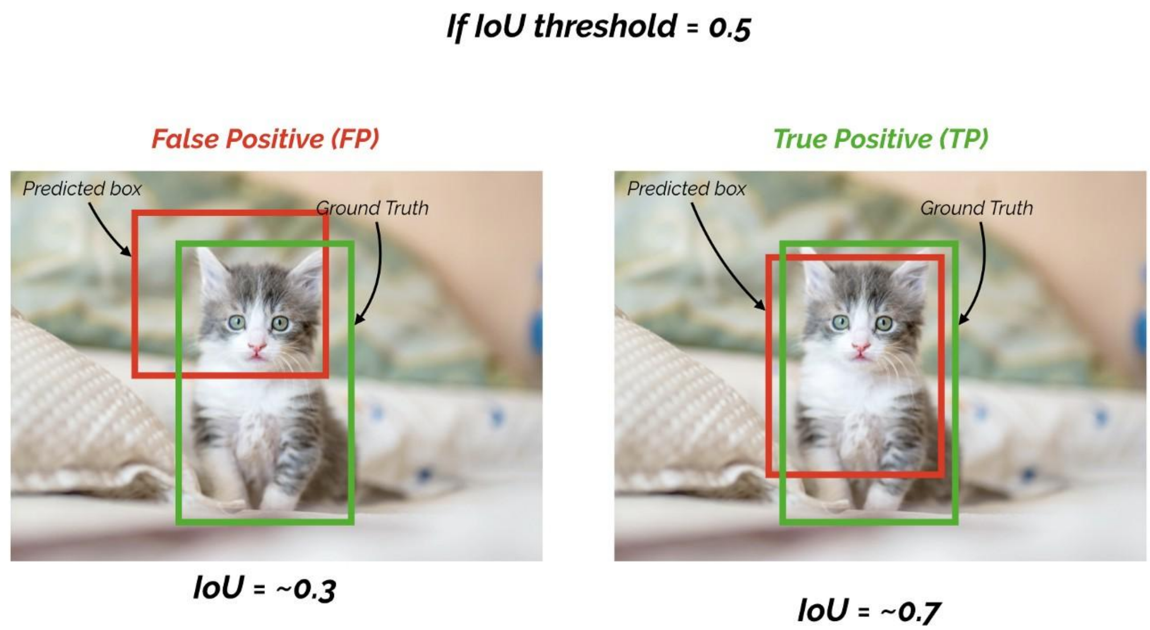
The testing for cancer as an example

$$Precision = \frac{TP}{\text{total positive results}}$$

$$Recall = \frac{TP}{\text{total cancer cases}}$$

**IoU (Intersection over Union) :** IoU is a detection evaluation metric that measures the overlap between the predicted bounding box and the ground truth bounding box. The value of IoU is always between 0 and 1. When both boxes perfectly overlap, i.e the IoU between the two boxes becomes equal to 1.





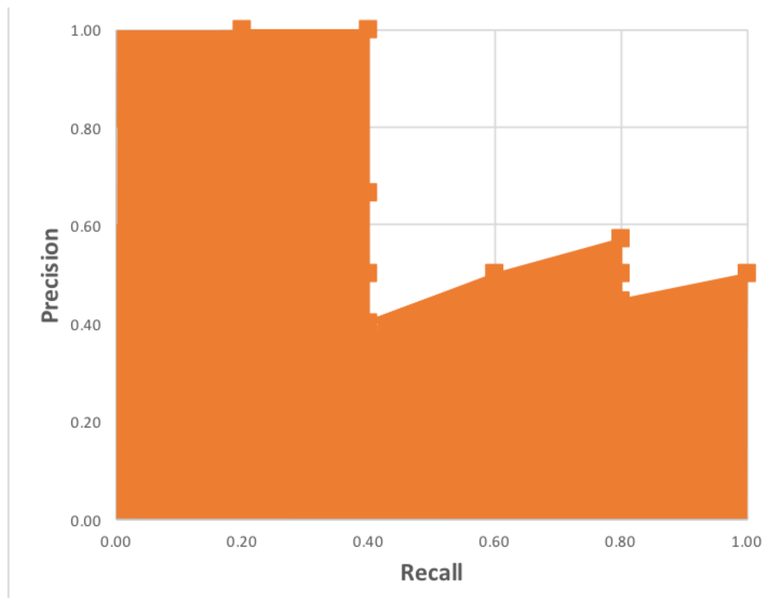

For example, in the above image if IoU = 0.5, and the IoU value for a prediction is 0.7, then we mark the prediction as True Positive. On the other hand, if IoU is 0.3, we mark it as False Positive.

---

**Average Precision definition & interpretation :**

The Average Precision is the area under Precision-Recall Curve. This gives the average precision of the detector. Let's consider an image with 5 apples where our detector provides 10 detections.

| Rank | Correct | Precision | Recall |
|---|---|---|---|
| 1 | True Positive | 1.00 | 0.20 |
| 2 | True Positive | 1.00 | 0.40 |
| 3 | False Positive | 0.67 | 0.40 |
| 4 | False Positive | 0.50 | 0.40 |
| 5 | False Positive | 0.40 | 0.40 |
| 6 | True Positive | 0.50 | 0.60 |
| 7 | True Positive | 0.57 | 0.80 |
| 8 | False Positive | 0.50 | 0.80 |
| 9 | False Positive | 0.44 | 0.80 |
| 10 | True Positive | 0.50 | 1.00 |




The Precision-Recall Curve represents the conflict between precision and recall. Typically, high precision and low recall models produce very confident predictions but miss a part of the instances. Models with low precision and high recall can find most objects, but the predictions are false positives to a certain degree, and confidence decreases.

In simple words, Average precision basically calculates the average of the precision values across the data. A model that have an AP of 1.0. would be considered as perfect.

**Example :** Let's say we have a set of bounding box predictions. Along with that, we have the IoU score which we calculated by comparing these bounding box predictions with the actual bounding boxes. Assume, we have a threshold of 0.5.

**Precision**

| Predicted | IoU | TP/ FP |
|---|---|---|
| Bounding Box 1 | 0.7 | TP |
| Bounding Box 2 | 0.2 | FP |
| Bounding Box 3 | 0.9 | TP |
| Bounding Box 4 | 0.8 | TP |
| Bounding Box 5 | 0.4 | FP |

So we can see in the above image example, that we have five bounding boxes with their IoU scores, and based on the IoU score we can define if this bounding box is a TP or a FP.

Now, Lets calculate the precision for this specific scenario where we are only considering the bounding box1.

**Average Precision**

| Predicted | IoU | TP/ FP | $\text{Precision}_i$ |
|---|---|---|---|
| Bounding Box 1 | 0.7 | TP | 1 |
| Bounding Box 2 | 0.2 | FP | 1 |
| Bounding Box 3 | 0.9 | TP | 0.66 |
| Bounding Box 4 | 0.8 | TP | 0.75 |
| Bounding Box 5 | 0.4 | FP | 0.75 |

$$AP = \frac{1}{N} \sum precision_i$$

$$= \frac{1}{5} (4.16)$$

$$= 0.832$$

Analytics Vidhya

If we have k classes, then for each individual class, we'll calculate this average precision, and finally take an average across all the classes. This way you can find the **mean average precision**.

**mean average precision formula :**

$$mAP = \frac{\sum_{i=1}^{K} AP_i}{K}$$

Average across all class

Mean average precision (mAP) is the mean of the average precision (AP) for all classes of objects.