

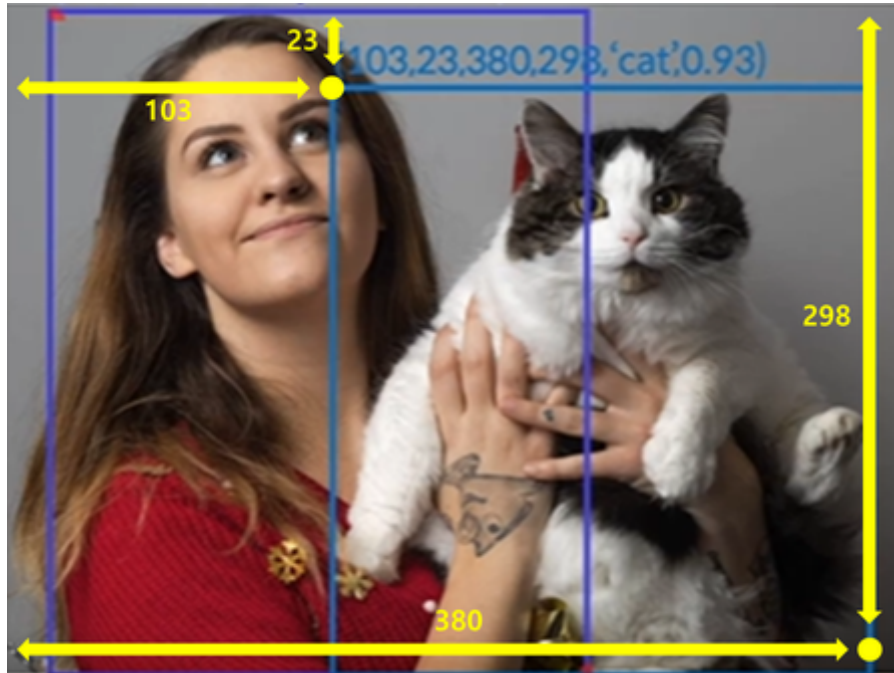
## 섹션 2. Object Detection 문제영역 소개

### 1. Object Detection 문제영역 소개

#### Computer Vision의 문제 영역

- image classification (이미지 분류)
- semantic image segmentation : 전체 픽셀에 대한 분류를 수행
- Object Detection (물체 검출) : 이미지 내에 오브젝트를 포지션 잡아서 바운딩 박스를 그리고, 이미지 내에 대응되는 클래스 등을 검출해내는 것

#### Object Detection 문제 영역의 출력값



(x\_min, y\_min, x\_max, y\_max, class, confidence) 순으로 나타난 것임.

ex) 고양이 Bounding Box

- 절대 좌표 : (103, 23, 380, 298, 'cat', 0.93)
- 상대 좌표 : 이미지의 가로 총 길이=400, 세로 총 길이=300이라고 하면,  $103/400 = 0.25$ ,  $23/300=0.07$ ,  $380/400=0.95$ ,  $298/300=0.99$ 이므로 (0.25, 0.07, 0.95, 0.99, 'cat', 0.93)으로도 나타낼 수 있다.

## 2. Object Detection Metric (IoU, mAP, MS COCO Metric)

Object Detection은 Supervised learning을 통해 진행하고, Bounding box의 포지션 값과 class label로 평가한다.

#### Metric 1 - IoU(Intersection over Union) Metric

- 한 개의 바운딩 박스와 한 개의 바운딩 박스가 얼마나 일치 하는가를 0~1 사이로 표현한다.
- 두 개의 바운딩 박스가 일치할수록 1이 되고, 일치하지 않을수록 0이 된다.
- 계산 식 :  $IoU = \text{두 박스의 교집합의 크기} / \text{두 박스의 합집합의 크기}$
- 이미지 하나에 여러 개의 물체가 있기 때문에 실제로는 잘 사용하지 않는다.

#### Metric 2 - mAP(mean Average Precision)

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$F1Score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}}$$

- Precision(정밀도) : 모델이 True라고 분류한 것 중에서 실제 True인 것의 비율입니다.
- Recall(재현율) : 실제 True인 것 중에서 모델이 True라고 예측한 것의 비율입니다.

=> 좋은 Object Detector model은 Precision과 Recall이 모두 높게 나와야 한다. 그러나 Precision과 Recall이 서로 반대로 가는 경향이 있다. 예를 들어, 100명의 사람이 있는 이미지 중 한 사람만 사람으로 감지했다면 precision은 1이고, recall은 1/100=0.01이 된다. 그렇기 때문에 F1 score를 통해 확인한다.

- F1 score : Precision과 Recall의 조화평균
- mAP : Recall별 Precision의 평균(AP)를 전체 클래스로 확장하여 평균을 낸 값. 이때, Positive 판단 기준은 특정 임계치(ex, 0.5)이상의 IoU이다.

### Metric 3 - Ms COCO Metric

- positive의 판단 기준에서 IoU의 임계치를 0.5로 정의하는 것에서 확장하여, 0.5, 0.55, 0.6 등으로 바뀌가면서 측정한 AP들의 평균을 낸 것.
- 최근에 가장 많이 사용되는 metric이다.

## 3. Object Detection Datasets - Pascal VOC Dataset, MS-COCO, KITTI, Open Images Dataset v4

### Pascal VOC Dataset

- 20개의 클래스로 구성.
- 2007년, 2012년 두 개의 데이터 셋이 있으며, 두 개를 합쳐서 사용할 수 있음.
- 예전에 만들어진 데이터 셋이기 때문에 한 장의 이미지에 많은 바운딩 박스가 들어있지는 않음
- Person, animal, vehicle, indoor 등의 카테고리.

### MS-COCO

- 비교적 최근에 구축되었기 때문에 파스칼보다 많은 데이터 셋을 가지고 있음.

### KITTI

- 자율 주행 연구를 위한 이미지 셋으로, 동영상 데이터 셋임.

### Open Images Dataset v4

- 가장 최근에 나타난 데이터 셋으로, 가장 많은 데이터 셋.

- 600개의 레이블 제공
- 필요한 이미지가 있을 때, 구글 오픈 이미지 데이터 셋에서 레이블을 검색해서 사용하면 좋음.