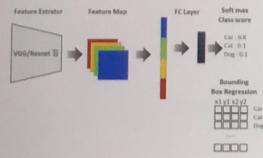


## Object Detection의 난제

Classification + Regression을 동시에

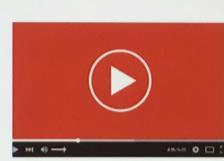
다양한 크기와 유형의 오브젝트가 섞여 있음



이미지에서 여러 개의 물체를 classification하고 동시에 위치를 찾아야 함

크기가 서로 다르고, 생김새가 다양한 오브젝트가 섞여 있는 이미지에서 이들을 Detect해야 함.

중요한 Detect 시간

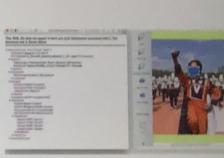


Detect 시간이 중요한 실시간 영상 기반에서 Detect해야 하는 요구사항 증대  
↓  
지연적 서비스에서 빠르게 해결해야!

명확하지 않은 이미지



데이터 세트의 부족



캡쳐 막 주면  
캡처지 알기 쉽지 않음...

오브젝트 이미지가

명확하지 않은 경우가

많음. 또한 전체

이미지에서 Detect할

오브젝트가 차지하는

비중이 높지 않음.(배경이

대부분을 차지하는 경우가

많음)

훈련 가능한 데이터 세트가 부족(MS Coco dataset 80개, Google Open Image 500개)하며 annotation을 만들어야 하므로 훈련 데이터 세트를 생성하기가 상대적으로 어려움

다양한 object에  
만장 등 학습데이터셋이 필요!

파이썬 딥러닝 컴퓨터 비전

MacBook Pro

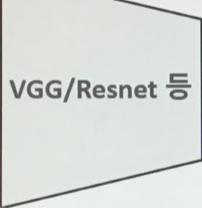
## Object Localization 개요

원본 이미지



Feature Extractor

= 원본 이미지에서 특징 추출!



Feature Map

| 차원  
| 배열화

FC Layer

fully connected layer

Soft max

Class score  
Car : 0.8  
Cat : 0.1  
Dog : 0.1

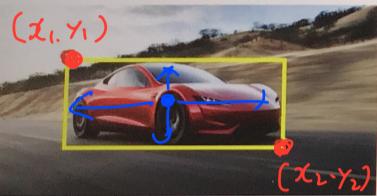
파이썬 딥러닝 컴퓨터 비전

MacBook Pro

## Object Localization 개요

*bounding box!*

원본 이미지

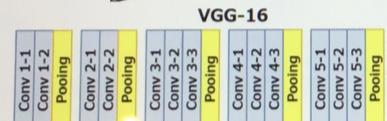


↓  
Regression으로  
좌상단, 우하단 좌표만  
찾으면 됨

Feature Extractor

VGG/Resnet 등

Feature Map



FC Layer

Soft max  
Class score

Car : 0.8  
Cat : 0.1  
Dog : 0.1

Classification layer  
Regression layer

Bounding Box Regression

x1	y1	x2	y2	
				Car
				Cat
				Dog
				.....

파이썬 딥러닝 컴퓨터 비전

MacBook Pro

## Object Localization – Bounding box 학습

원본 이미지의 Bounding box



학습된 Object의 Bounding Box



Feature Extractor

Feature Map

FC Layer

Soft max  
Class score

Bounding  
Box Regression

Car : 0.8  
Cat : 0.1  
Dog : 0.1

$$\begin{aligned} x_1 &= w_{10}f_1 + w_{20}f_2 + w_{30}f_3 + \dots \\ y_1 &= w_{11}f_1 + w_{21}f_2 + w_{31}f_3 + \dots \\ x_2 &= w_{12}f_1 + w_{22}f_2 + w_{32}f_3 + \dots \\ y_2 &= w_{13}f_1 + w_{23}f_2 + w_{33}f_3 + \dots \end{aligned}$$

가중치 update

가중치 update

파이썬 딥러닝 컴퓨터 비전

MacBook Pro

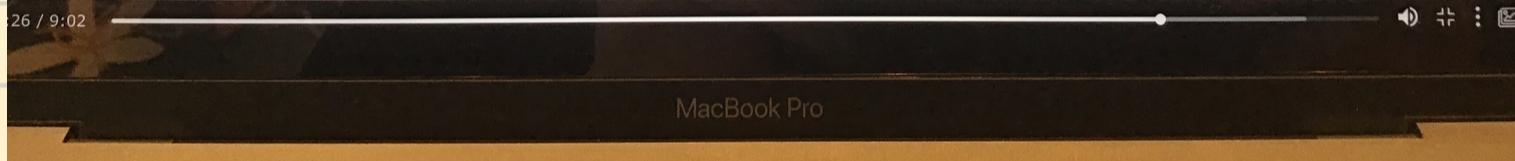
## Object Localization 예측 결과



→ 쉽게 생각하면 class별 softmax 확률

Class	Confidence Score	X1	Y1	X2	Y2
Car	0.9	50	60	220	150
Dog	0.1	210	10	240	30
Cat	0.1	160	100	180	120

파이썬 딥러닝 컴퓨터 비전



## Object Detection – 두개 이상의 Object를 검출

이미지의 어느 위치에서 Object를 찾아야 하는가?



이미지 안에 사람, 차, ...  $\Rightarrow$  어느 위치에서 Object를 찾어야 하는가?  
(= 어느 object를 찾아야 하는가?)

여러개의 object가  
있을 때,  $\Rightarrow$  object detection.

이미지 자체에서, classification  
정의가 어렵다.

↓  
detection 해야 하는 부분을  
(위치)  
찾아야 함

$\Rightarrow$  이후 classification,  
regression 실시

파이썬 딥러닝 컴퓨터 비전

MacBook Pro