**Image Objection Detection**

* 이미지 분류(Image Classification), 객체 탐지(Object Detecting), 이미지 분할(Image Segmentation) 영역으로 나뉜다. 특히 Segmentation의 경우에는 Semantic segmentation과 Instance segmentation으로 나누어 볼 수도 있다. 그리고 객체 탐지의 경우 이미지에 한정되지 않고 영상에서도 객체 탐지를 적용해 사용하고 있다.
* 현재는 컴퓨터가 이런 이미지와 영상에서 인간의 시각이 할 수 있는 일들을 자율적으로 수행할 수 있도록 연구하는 분야에 대해서 **'컴퓨터 비전(Computer Vision)'**이라고 부르고 있다.

1. **이미지 분류(Image Classification)**

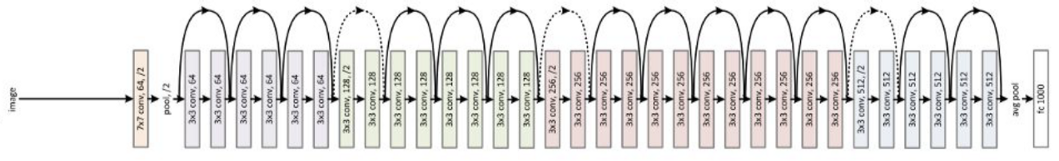
* 이미지 속 특정 대상 혹은 물체가 무엇인지 분류해내는 것
  1. **CNN(Convolutional Neural Network) - Feature extraction(특징 추출)이 핵심!**
     1. 인간의 뉴럴 네트워크를 모방해서 만드는 것으로 convolution과 pooling를 반복하여 이미지의 특성을 추출하고 Fully Connected Layer를 통해 이미지를 분류한다.
     2. Convolution : Window Sliding Matrix multiplication을 통해 특징값을 추출한다. 특정 filter의 모양과 일치하는 부분에서 큰 값을 갖는다. 전체 이미지를 순회하고 나면 유사한 모양을 가진 부분에 대한 특징만 얻을 수 있다.

이미지의 특징을 두드러지게 하기 위해서는 필터의 크기를 더 작게 하고 필터의 이동 범위를 작게 조절하면 더 세밀하게 특징을 뽑을 수 있다. 하지만 이렇게 되면 학습량이 늘고 시간이 오래 걸린다. 또한 너무 자세하게 학습하여 과적합 문제가 발생할 수 있다.

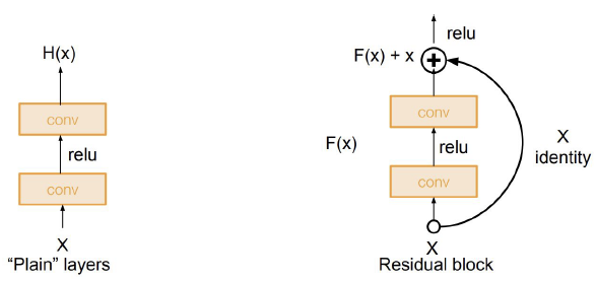
* + 1. Pooling : filter를 거쳐 얻은 feature map에서 다시 한번 특징을 표현해 낼 수 있는 값을 도출해내는 것을 말한다. MaxPooling과 AveragePooling이 있다. Pooling을 적용하게 되면은 이미지의 크기는 줄이면서 해당 영역의 특징은 그대로 유지할 수 있기 때문에 앞서 말한 문제점들을 해결 가능하다.
  1. **VGG-16(Oxford)**



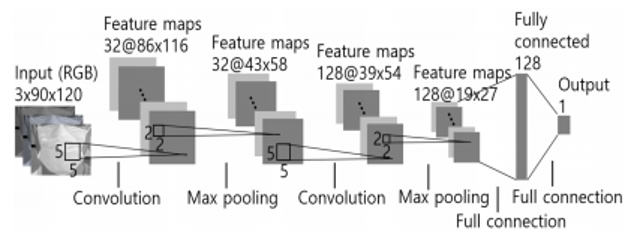
* + 1. CNN과 같이 Convolution과 Pooling을 반복하고, 마지막에 완전연결층(Fully Connected, FC)으로 구성되는 모델이다.
    2. ResNet보다 적은 층을 갖고 있어서 훈련에 소요되는 시간이 훨씬 적다.
  1. **ResNet(Microsoft)**



* + 1. Convolution과 Pooling의 반복을 통해 특징을 추출하고, 마지막에 완전연결층을 통해 분류한다. 하지만 VGG-16과 달리 이전 layer와 다음 layer을 이어주는 연결선이 존재하는 것을 볼 수 있다.
    2. 정확도가 VGG-16보다 높다
  1. **Residual Connection(Skip Connection)**



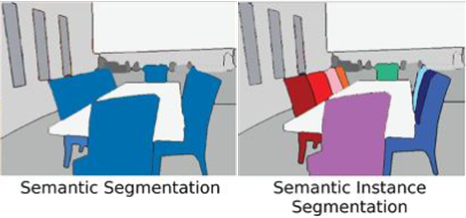
* + 1. 두 구조의 차이점은 동일한 연산을 하고 나서 Input x을 더하는 것(Residual block)과 더하지 않는 것(Plain layers)이다.
    2. **Residual block** : Input x을 더함으로써, 기존에 학습한 정보를 보존하고 추가적으로 정보를 학습한다. 즉, 이전 Layer에서 학습했던 정보를 연결함으로써 해당 층에서는 추가적으로 학습해야할 정보만을 학습하면 되어 학습할 양이 감소한다. – 오픈북이 가능한 시험
  1. **Custom CNN**



마스크를 제대로 착용했는 지 분류해주는 모델

* + 1. 어떤 데이터셋을 사용하는지, 어떻게 이 모델을 활용하는지에 따라 자유롭게 Customize 할 수 있다는 장점

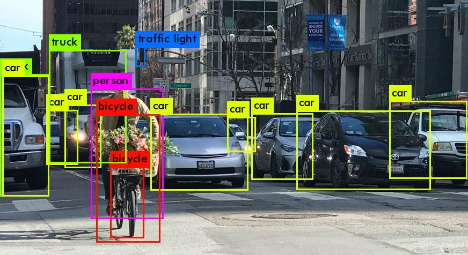
1. **이미지 분할(Image Segmentation)**



* Segmentation : 이미지 영역을 분할하여 원하는 object에 맞게 다른 영역과 구분가능하게 분할 해주는 것을 말한다.
* Semantic segmentation : 이미지 내에서 원하는 object를 분류해 냈을때, 똑같은 object로 분류한 object들에 대해 “같은” object로 인식 (사진 속 의자들을 모두 같은 object로 인식하고 모두 파란색으로 표시)
* Insatance segmentation : 이미지 내에서 원하는 object를 분류해 냈을때, 똑같은 object로 분류한 object들에 대해 “다른” object로 인식 (같은 범주의 의자임에도 각각 다른 object로 인식하여 다른 색으로 표시)
  1. Fully Convolutional Network, FCN

1. **객체 탐지(Object Detecting)**

* 이미지나 영상에서 찾고자 하는 대상에 대해 분류해내는 기술



* 1. 과정 : Object Recognition1) - Object Classification2) - Object Locallization - Object Tracking(영상에서)
  2. 전통적인 방법은 Two-shot이지만 YOLO알고리즘 등장으로 One-Shot이 가능해졌다.

1. 우리가 사용할 모델은 어떤것?
   1. <https://towardsdev.com/image-search-engine-using-deep-learning-model-c452d2637cf6>
   2. https://towardsdatascience.com/image-similarity-with-deep-learning-c17d83068f59