

# PyTorch로 딥러닝 제대로 배우기

- 중급편 -

Part6-1. 컴퓨터 비전 이론: 컴퓨터 비전

강사: 김 동 희



# 목차

## I. 컴퓨터 비전

- 1) 개요
- 2) Recognition
- 3) Motion analysis
- 4) Image reconstruction
- 5) 컴퓨터 비전에서 어려운 점

## II. 합성곱 신경망(CNN)

- 1) CNN(Convolutional Neural Network)
- 2) CNN 구조
- 3) CNN 응용

## III. CNN 실습

- 1) 데이터
- 2) 모델
- 3) 학습
- 4) 평가

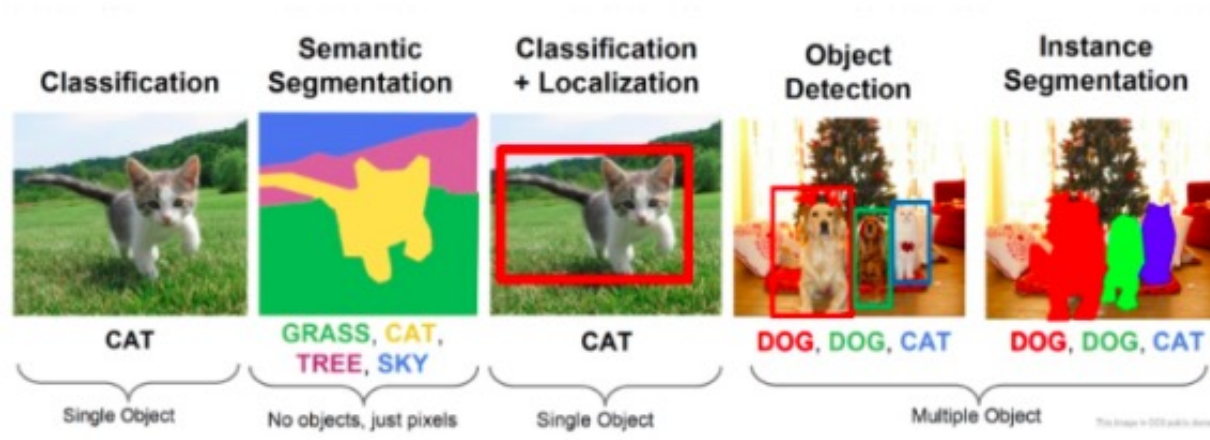


# 1. 컴퓨터 비전

# 1. 컴퓨터 비전

## □ 개요

- 컴퓨터 비전(vision)이란, 말 그대로 컴퓨터 분야에서 시각을 연구하는 분야
- 여기서 시각이란, 빛이 자극하여 일어나는 감각으로 컴퓨터에서는 일반적으로 렌즈를 통해 빛이 전달 되면서 일어남
- 인공지능 분야에서 1960년대부터 활발이 연구되었던 분야이며, 지금까지도 많이 연구되고 있음
- 인공지능에서 일반적으로 연구하고 있는 Task는 Recognition, Motion analysis, Scene or Image reconstruction, Image restoration 등이 존재



Computer Vision에서 여러 Task 예시

## 2. Recognition

### □ Classification

- 컴퓨터 비전에서 classification task는 이미지 한 장(input)에 대해서 종류(label)를 예측하는 것
- 초창기에 가장 많이 해결하고자 했던 Task이고, 정형적인 환경에서는 현재까지도 정확도가 가장 좋은 방법
- CNN 모델을 중첩해서 사용하거나, 응용하여 모델들을 구현
- LeNet, AlexNet, GoogleNet, ResNet, VGG 모델 등이 대표적



Image Classification 예시

## 2. Recognition

### □ Semantic Segmentation

- Segmentation은 모든 픽셀의 label을 예측하는 Task
- 한 장에 사진에서 여러 개의 label이 존재 가능
- Semantic Segmentation은 한 장의 이미지에서 대상을 구체적으로 정의하지 않고, 큰 분류로 구분
- 아래 예시에서 사람과 하늘, 나무는 구분하지만 사람을 한 명씩 구분하지는 않음
- DeepLabV3, FCN, LRASPP 모델이 대표적



Semantic Segmentation 예시

## 2. Recognition

### □ Instance Segmentation

- Semantic segmentation과 다르게 각각의 객체를 인식하는 것이 특징
- Mask R-CNN 모델이 대표적

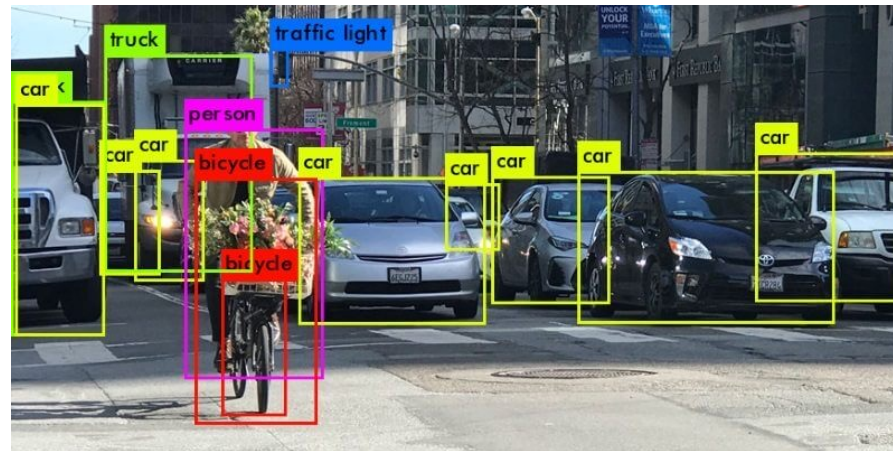


Original 이미지 vs. Instance segmentation vs. Semantic segmentation

## 2. Recognition

### ❑ Object Detection

- 하나의 이미지에서 사물(object)을 탐지하는 모델
- 이미지 입력에서 사물의 위치(localization)와 사물의 종류(label)를 동시에 예측
- 탐지 방법에 따라 One-stage detection과 Two-stage detection으로 나뉨
- 대표적인 모델로 Faster R-CNN과 SSD가 존재



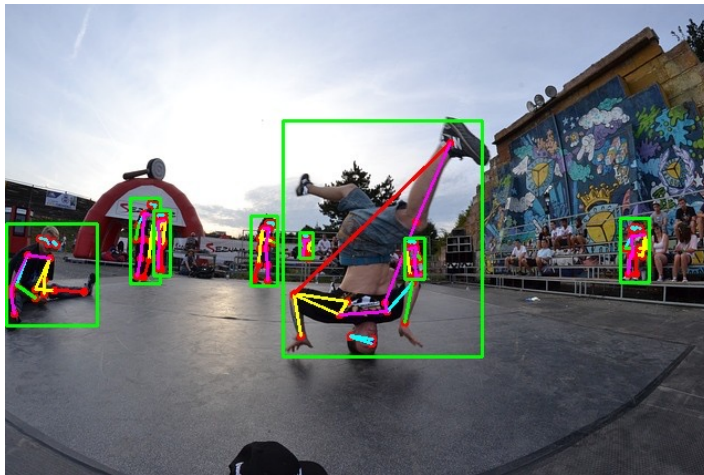
Object detection 예시



## 3. Motion analysis

### □ Keypoint analysis

- 사물의 특정위치를 학습하여 행동을 예측하거나 분석하는 컴퓨터 비전 기술
- 일반적으로는 사람을 대상으로한 연구가 많이 진행되고 있고, 동물에 대한 연구도 최근 진행되고 있음
- 대표적인 모델로는 Mask R-CNN이 있음

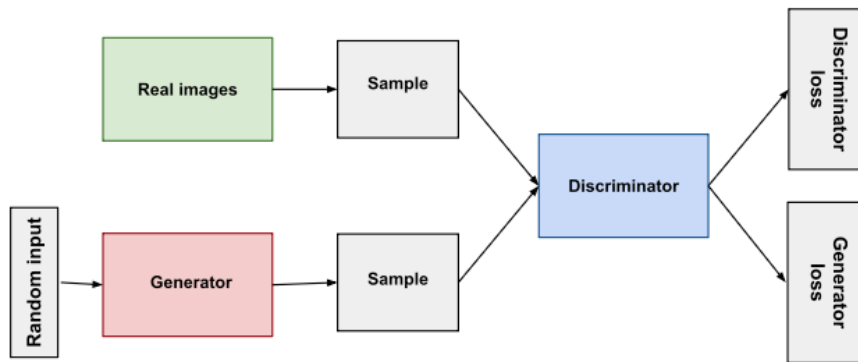


Keypoint analysis

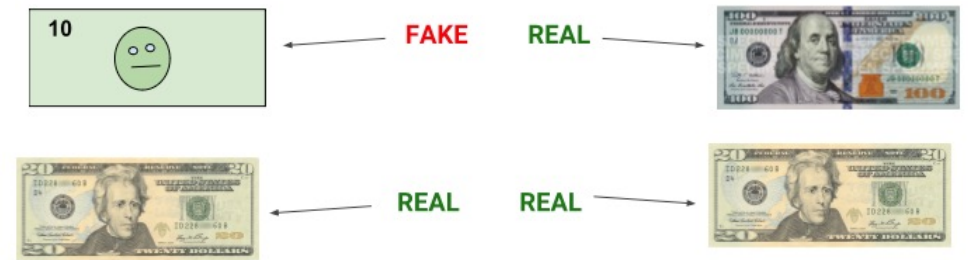
## 4. Image reconstruction

### □ GAN(Generative Adversarial Network)

- 새로운 이미지를 생성하기 위해 **생성자(Generator)**와 **구분자(Discriminator)**의 **적대적(Adversarial)** 구조로 네트워크를 구성하는 모델
- 대표 모델로 BEGAN, StyleGAN, StarGAN 등이 존재



GAN Architecture



GAN 원리

## 5. 컴퓨터 비전에서 어려운 점

### □ 데이터

- 공개된 데이터 이외에 새로운 데이터에 대해서 새로 label을 달기 매우 어려움 (고비용)
- 고해상도 이미지를 활용하거나 동영상을 활용하는 경우, 학습을 위한 고성능 컴퓨터 디바이스가 필요
- Self-supervised learning, Un-supervised learning, Transfer learning 등의 연구가 진행중

### □ 모델

- 여러가지 Task를 잘 수행 가능한 모델 개발이 어려움 (도메인이 다른 경우, label 특징이 다른 경우)
- 고성능 모델의 경우, 모바일 디바이스나 IoT 디바이스에서 활용이 어려움
- Knowledge distillation, Domain Adaptation 등의 연구가 진행중

### □ Explainable

- 인공지능 모델이 도출한 예측 결과에 대해 설명력을 부여할 수 없음
- 자동차, 의료 등 일부 산업에 활용하기 위해서는 판단 근거에 대한 설명력이 필요할 수 있음
- XAI 연구가 진행중

감사합니다.