

Introduction to Convolutional Neural Networks for Visual Recognition

History of Computer Vision

1. Development of the Vision
 - a. Visual processing starts with simple structure of the visual world
 - b. Brain builds up the complexity of the visual information until it can recognize the complex visual world
2. Object Composition
 - a. Every object is composed of simple geometric primitives
 - b. For accurate detection, objects are need to reduce the complex structure into a collection of simpler shapes
 - i. generalized cylinder (e.g., 나무 인체 모형)
 - ii. pictorial structure (e.g., 졸라맨)
3. Image Classification Tasks
 - a. Image segmentation (Normalized cut)
 - i. 주어진 이미지에서 사람과 배경을 구분하는 태스크
 - ii. 명확하게 표현하는 것이 목적이 아니라 구분만이 목적이므로 이미지 픽셀이 뭉개져도 상관없음
 - iii. e.g.,



b. Face detection

- i. 주어진 이미지에서 사람의 얼굴을 감지하는 태스크

c. Object recognition

- i. 주어진 이미지에서 물체를 감지하는 태스크
- ii. diagnostic 하고 invariant 한 특징을 잘 추출하는 것이 중요함
- camera angles, occlusion, lighting, viewpoint, intrinsic variation of the object itself 등 같은 물체가 여러 요인에 의해 이미지에서 다르게 표현 될 수 있기 때문임
 - e.g., 두 이미지가 주어졌을 때, 이미지 내에서 주요 특징을 추출하여 물체를 감지함

d. Human recognition

- i. 주어진 이미지에서 사람을 감지하는 태스크로 object recognition 과 유사함

4. Benchmark Datasets

a. PASCAL Visual Object Challenge

b. ImageNet

- i. visual data는 매우 복잡하므로 모델이 high dimension of input 과 a lot of parameters to fit 하다는 특징을 가짐
- ii. 그러나 초기에는 데이터의 양이 적었으므로 overfitting이 많이 발생함

iii. 따라서 이를 해결하기 위해 ImageNet 프로젝트가 시행됨

5. Other Fields of Computer Vision

a. Object detection

i. 이미지가 주어졌을 때, 이미지 내 여러 물체들을 감지하는 기법

b. Image captioning

i. 이미지가 주어졌을 때, 해당 이미지에 대한 설명을 생성하는 기법

c. Action classification

6. Recent Quests

a. Semantic segmentation

i. 하나의 물체를 감지하는 것이 아니라 이미지를 픽셀 단위로 분석함

b. 3D understanding

c. Activity recognition