

ImageNet 데이터셋

Dataset 구성

- 총 21,841개 classes의 14,197,122개 데이터로 구성되어 있음
- 다양한 해상도의 image size로 구성되
- 학습 데이터셋 용량은 약 138GB, 검증 데이터셋 용량은 약 6GB
- 227×227 pixels로 crop된 sub-sampled images를 사용하는 것이 일반적

Dataset 특징-

Stanford University의 Li Fei-Fei 교수를 중심으로 제작되었으며, Google, Microsoft 와 같은 기업 또한 함께 참여하여 제작한 데이터셋이다.

- ImageNet Dataset의 경우, ILSVRC: ImageNet Large Scale Visual Recognition Competition 에서 사용되는 데이터셋으로 유명하다.

- 최근까지 이미지 분류뿐 아니라 객체인식(object detection), 의미 분할(semantic segmentation), 자세 추정(pose estimation) 등의 다양한 컴퓨터 비전 문제에 활용 될 수 있는 공통 convolutional neural network (CNN) 백본 모델 학습에 널리 활용 되어 왔다.

- 유명한 Amazon Mechanical Turk 서비스를 이용하여 일일이 사람이 분류한 데이터셋이다. 이 데이터셋은 ILSVRC (ImageNet Large Scale Visual Recognition Challenge)로 잘 알려진 국제 대회에서 사용되는 데이터셋으로도 유명하다.

- 논문에서 가장 자주 등장하는 데이터셋은 ILSVRC 2012 데이터셋이다. 최신 논문도 대개 ILSVRC 2012를 이용해 학습/평가를 진행한다.

- 사실 일반적으로 딥러닝에 사용되는 데이터셋은 평가(training) / 검증(validation) / 테스트(test) 데이터셋으로 나누어지는데, ILSVRC 2012는 실제 대회에서 사용되

었던 데이터셋이며 테스트 데이터셋은 공개하지 않고 있다. 그래서 대개 논문을 쓰기 위한 목적으로는 평가(training) 데이터셋과 검증(validation) 데이터셋을 사용한다.

- 이 데이터가 유명해진 이유는 ImageNet Large Scale Visual Recognition Challenge(ILSVRC) 공모전이 이 자료 집합을 사용하기 때문이다. 이 공모전은 컴퓨터 시각 분야에서 권위가 아주 높다.

- 또 다른 이유는 응용 분야의 핵심적인 시각적 개념들을 모두 표현하기에 충분할 정도로 이미지들이 많고 다양하다는 점이다. 그래서 흔히 합성곱 신경망(CNN, Convolutional Neural Network)들은 이 자료 집합으로 훈련되며, 미리 훈련된 모델을 임의의 이미지에서 특징들을 추출하는 용도로 사용할 수 있다. 이 자료들의 시각적 개념들이 다른 응용 분야의 자료 객체들로 전달하는 일종의 전이 학습으로 사용한다.