

ImageNET

- AI 연구원 Fei-Fei Li 는 2006년 ImageNet에 대한 아이디어 작업을 시작했다.
 - 당시 신경망 모델 연구에 따라 신경망은 점차 복잡해졌고 overfitting문제를 피할 수 없게 되었다. 그에 AI 연구자들은 모델과 알고리즘을 훈련하는 데 사용할 수 있는 데이터를 확장하고 개선하기를 원하게 되었다.
 - 2007년 Li는 WordNet 의 창시자 중 한 명인 Princeton 교수 Christiane Fellbaum 을 만나 프로젝트에 대해 논의했다.
 - Princeton의 조교수로서 Li는 ImageNet 프로젝트에 참여하기 위해 연구원 팀을 구성하였다.
 - 이미지 분류를 돕기 위해 Amazon Mechanical Turk 를 사용했다
 - ◆ ImageNet은Amazon Mechanical Turk 서비스를 이용하여 일일이 사람이 분류한 데이터셋이다.
 - 이 데이터베이스는 명사들로 이루어진 WordNet 계통구조에 따라 조직화되어 있다.
 - WordNet 데이터베이스는 유의어집합(synset)의 개념에 기초해서 결정된 영어 단어들 사이의 관계를 담은 자료 집합이다.
 - 플로리다 에서 열린 2009년 CVPR(Computer Vision and Pattern Recognition) 컨퍼런스에서 포스터로 처음으로 데이터베이스를 발표했다
-
- 이 데이터셋은 ILSVRC (ImageNet Large Scale Visual Recognition Challenge)로 잘 알려진 국제 대회에서 사용되는 데이터셋으로도 유명하다. 이 공모전은 컴퓨터 시각 분야에서 권위가 아주 높다.
 - 이미지 응용 분야의 핵심적인 시각적 개념들을 모두 표현하기에 충분할 정도로 이미지들이 많고 다양하기 때문에 논문에도 자주 사용된다.
 - 논문에서 가장 자주 등장하는 데이터셋은 ILSVRC 2012 데이터셋이며, 최신 논문도 대개 ILSVRC 2012를 이용해 학습/평가를 진행한다.
 - 테스트 데이터셋은 공개하지 않고 있다.
 - 대개 논문을 쓰기 위한 목적으로는 평가(training) 데이터셋과 검증(validation) 데이터셋을 사용한다.
 - 14,197,122 images, 21841 synsets index 를 가지고 있다.