

영상 처리 & 컴퓨터 비전

Open CV

❖ Open CV

- OpenCV(Open Source Computer Vision Library)은 영상 처리와 컴퓨터 비전 관련 오픈 소스 라이브러리
- OpenCV는 C/C++로 구현되었고, 소스가 공개되어 있어 사용자가 재빌드 할 수 있으며 윈도우즈, 리눅스, iOS, 안드로이드 등의 다양한 플랫폼에서 C, C++, Python, JAVA, MatLab 등의 언어로 사용할 수 있음
- MMX(MultiMedia eXtension)와 SSE(Streaming SIMD Extensions) 명령어를 통해 고속의 알고리즘을 구현해 실시간 비전 응용에 더욱 강점을 가지고 있음
- OpenCV는 영상 및 비디오 입출력 모듈, 영상 처리 및 컴퓨터 비전 관련 기본 알고리즘, 기계학습 모듈 등이 구현되어 있으며 최신 논문에 발표되는 내용이 구현될 정도로 활발히 업데이트 되는 BSD(Berkeley Software Distribution) 라이선스를 따르는 Open Source 라이브러리
- 최근에는 인공지능 기계학습 분야에서 가장 주목 받는 심층 신경망(DNN, deep neural networks) 모듈을 추가하여, Tensorflow, Caffe, Torch 등의 딥러닝 프레임워크로 학습된 모델을 읽어서 순방향(forward pass) 추론 단계만 수행하여 사용하는 기능이 추가
- CUDA(Compute Unified Device Architecture)와 OpenCL(Open Computing Language) 인터페이스가 현재 개발되어 사용되고 있음

영상 처리와 컴퓨터 비전

❖ 디지털 영상 처리(digital image processing)

- 컴퓨터를 사용하여 입력 영상 데이터를 처리하는 분야

ex)

- 포토샵을 이용하여 입력 영상에 포함된 잡음(noise)을 제거 하거나
- 영상의 대비(contrast)를 개선(enhancement)하여 선명하게 하거나
- 영상의 특정 부분인 관심 영역(region of interest)을 강조(emphasizing)하거나
- 관심 영역을 분할 (segmentation)하거나 영상을 파일로 압축(compression)하거나
- 네트워크를 통해 전송하거나 영상 검색(retrieval), 분류(classification), 인식(recognition) 등의 영상 다루는 모든 분야가 영상 처리

❖ 컴퓨터 비전(computer vision)

- 카메라에 의해 캡처된 입력 영상에서 의미 있는 정보를 추출하는 분야로 주로 실시간(real time) 응용 문제를 다룸

ex)

- 컴퓨터 비전의 응용 예는 산업현장에서 자동으로 제품의 결함을 검사(industrial inspection)하거나
- 스캐너 또는 카메라 캡처 영상에서 문자 인식(character recognition), 얼굴 인식(face recognition), 지문 인식(fingerprint recognition), 사람 또는 자동차 등과 같은 움직이는 물체 검출(motion detection) 및 물체 추적(object tracking), 2개 이상의 카메라로부터 캡처한 스테레오 영상을 이용하여 깊이를 계산하거나
- 3차원 물체의 구조(structure/shape)를 계산하는 등의 스테레오 비전(stereo vision) 등이 있음

영상 처리와 컴퓨터 비전

❖ 영상 처리 분야

영상획득	저 수준 영상처리 (좁은 의미의 영상처리)
영상향상	
영상복원	
변환처리	
영상압축	
영상분할	고 수준 영상처리 (컴퓨터 비전)
영상표현	
영상인식	

영상 처리와 컴퓨터 비전

❖ 이미지 처리

- 이미지 처리는 머신 러닝에서 가장 흥미로운 분야 중 하나.
- 이미지에서 패턴이나 물체를 인식하는 컴퓨터의 능력은 강력한 도구.
- 머신 러닝에 이미지를 사용할 때는 원본 이미지를 학습 알고리즘이 사용할 수 있는 특성으로 변환을 해야 함.
- 이미지를 다루는 라이브러리는 많은데 그 중에서 가장 유명한 라이브러리가 Open CV(Open Source Computer Vision).
- OpenCV는 BSD 라이선스를 갖는 오픈소스 컴퓨터 비전(computer vision) 라이브러리.
- 컴퓨터 비전이라는 말에 이미 시각(vision)이라는 단어가 포함되어 있어서 유추할 수 있듯이, **사람의 시각과 관련한 시스템 구조를 모방하여 컴퓨터도 물체나 상황을 식별하고 해석할 수 있도록 하는 연구 분야**로 컴퓨터 비전에서 다루는 데이터에는 이미지와 이미지 속 텍스트를 인지하는 영역 뿐만 아니라 비디오와 3D 영상도 포함됨.
- 스마트폰의 시대가 오면서 이미지뿐만 아니라 영상 데이터가 폭발적으로 증가하였고 이는 컴퓨터 비전의 필요성을 증대 시켰는데 Research and Market이 발표한 자료에 의하면 2021년 컴퓨터 비전에서의 AI 시장 규모는 159억 달러로 추정하며 2026년까지 이 시장은 513억 달러에 이를 것으로 예상했는데 이를 환산하면 연평균 약 26.3%의 복합 성장률(CAGR)로 컴퓨터 비전 시장이 확대된다는 것을 의미.

영상 처리와 컴퓨터 비전

- ❖ 딥러닝과 뉴럴 네트워크의 발전은 컴퓨터 비전 분야의 발전 속도를 가속화 시키는 계기가 되었고, 제한된 영역에서만 활용되던 컴퓨터 비전은 딥러닝의 심층 신경망의 비약적인 발전에 힘입어 그 활용 영역을 점차 확장해 나가고 있으며, 비전 분야에 있어서 대표적인 기술은 아래와 같음
 - 객체 분류(Object Classification) : 이미지 속 객체를 인지하여 그 클래스를 분류해내는 기술로 컴퓨터 비전영역에서 가장 기초적인 응용 분야
 - 객체 탐지 및 위치 식별(Object Detection & Localization) : 이미지 또는 비디오 영상에서 객체를 식별해내는 기술
 - 객체 분할(Object Segmentation) : 이미지 및 비디오 영상 프레임 내에서 객체를 따로 분할하여 의미 있는 부분만 분석할 수 있게 하는 기술
 - 이미지 캡셔닝(Image captioning) : 이미지의 상황을 텍스트로 설명할 수 있는 기술
 - 객체 추적(Object Tracking) : 비디오 영상 내의 객체의 위치 변화를 추적하는 기술로 흔히 포인트 추적(point tracking), 커널 추적(kernel tracking), 실루엣 추적(silhouette tracking) 등의 방법을 사용
 - 행동 분류(Action Classification) : 비디오 영상 내의 객체의 행동(action)을 인식하여 분류하는 기술

영상 처리와 컴퓨터 비전

❖ 영상 처리 응용 분야

- 의료 분야 : 방사선, 초음파
- 방송 통신 분야
- 공장 자동화 분야 : 품질 검사
- 출판 및 사진 분야
- 애니메이션 및 게임 분야
- 기상 및 지질 탐사 분야

영상 처리와 컴퓨터 비전

❖ 컴퓨터 비전의 어려움

➤ 사람은 이미지를 보지만 컴퓨터는 숫자를 바라봄

- 저명한 컴퓨터 비전 전문가인 UC버클리의 교수 Jitendra Malik(지텐드라 말릭)은 인공지능 전문가 Lex Fridman(렉스 프리드먼)과 진행한 인터뷰에서 사람이 이미지를 인식하는 과정이 무의식적으로 혹은 잠재 의식적으로 진행되기 때문에 흔히 컴퓨터가 비전을 처리하는 방식도 매우 쉬울 거라고 착각할 수 있다고 지적했는데 컴퓨터 비전의 신경과학을 들여다보면, 그 복잡성은 상상을 초월
- 대뇌피질의 상당 부분은 이미지 프로세싱 처리에 전념하게 되는데 이는 컴퓨터가 이미지를 처리하는 방식이 인간과 다르기 때문
- 인간이 매우 시각적이고 직관적인 방식으로 이미지를 인식하는 반면 컴퓨터는 이미지를 그 자체로 인식하는 것이 아닌 이미지의 모든 부분을 개별 픽셀로 환산하여 숫자로 인식하는데 이는 이미지를 인식할 때마다 처리해야 하는 데이터 양이 많다는 것을 의미하며 컴퓨터가 복잡한 시각적 작업을 수행하도록 하는 데 필요한 계산 및 데이터 리소스 또한 방대함

영상 처리와 컴퓨터 비전

❖ 컴퓨터 비전의 어려움

➤ 사람보다 더 많은 데이터를 필요로 하는 컴퓨터 비전 시스템

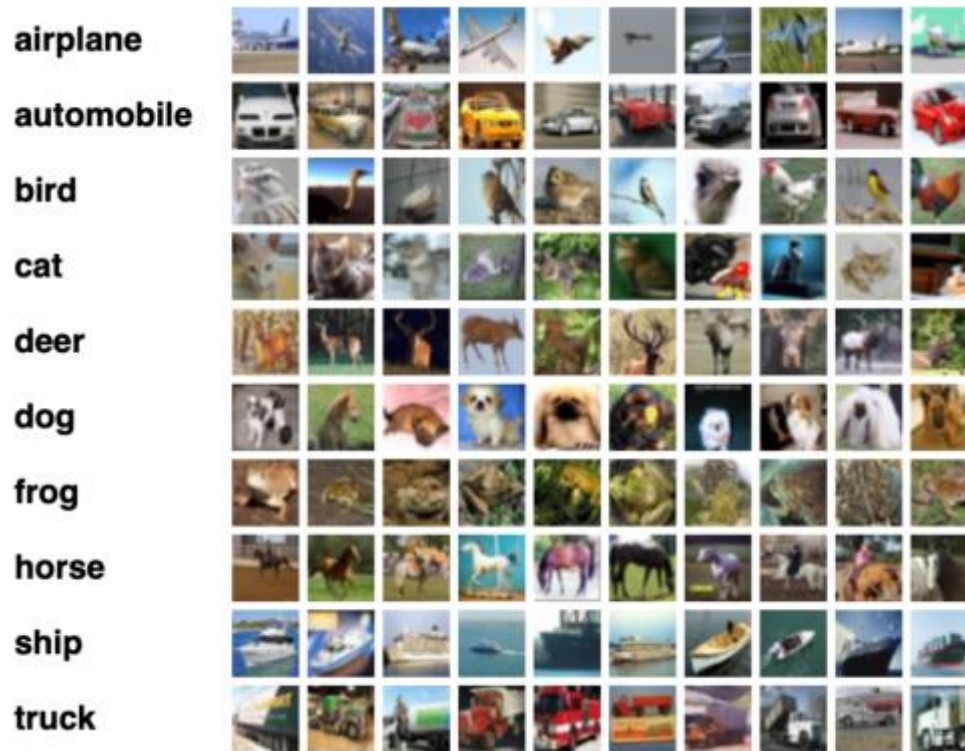
- 현재의 컴퓨터 비전 시스템이 이미지 처리에 있어서 사람과 같은 능력을 구현하려면 사람이 필요로 하는 것보다 훨씬 많은 데이터가 필요
- 현재 컴퓨터 비전 분야를 이끌어 가는 가장 고도화된 머신 러닝 기법은 지도 학습인데 지도 학습은 '정답'이라고 여겨지는 feature를 라벨링한 데이터 기반으로 학습을 이행하는 것인데 최근에는 지도 학습을 바탕으로 라벨링이 되지 않은 추가 인풋 데이터를 학습시키며 이러한 과정을 통해 모델의 성능을 개선해 나가고 있어서 학습에 사용된 데이터를 기반으로 모델의 성능이 결정되기 때문에 학습에 사용된 데이터가 실제 프로덕션 레벨에서 입력값으로 들어올 수 있는 데이터와 유사한 지가 중요한 포인트인데 쉽게 말해서 모델이 인지하려는 모든 다양한 상황에 대해 데이터로 준비해서 학습시켜야 하는 것
- 컴퓨터 비전을 사용하는 자율 주행을 예로 들어보면 자전거를 타는 사람들을 인지할 때 컴퓨터 비전 시스템은 자전거를 타는 사람들이 어떤 행동을 하는지에 대한 사전 지식이 필요하며 이를 바탕으로 동작해야 하지만 실제 자전거를 타는 사람들은 제각기 다른 행동을 할 수도 있어서 사람의 시각 시스템은 다양한 상황 속에서 이를 유연하게 인지하는 반면 컴퓨터는 실제로 발생할 수 있는 모든 상황에 대한 데이터가 주어져야 이를 학습하여 처리함
- 실제로 이런 다양한 상황의 데이터를 모두 모아 학습 시키기는 현실적으로 쉽지 않지만 대규모 자체 인력과 자본을 보유한 테슬라는 이러한 방식으로 자율주행 시장을 이끌고 있음

영상 처리와 컴퓨터 비전

❖ 컴퓨터 비전에서 많이 사용하는 데이터 셋

➤ CIFAR-10

- CIFAR-10 데이터셋은 머신러닝 연구에 가장 활발히 사용되는 데이터셋으로, 2009년에 제프리 힌튼 교수팀이 만들었는데 10개의 클래스(airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck)로 이루어진 6만 개의 이미지 데이터셋

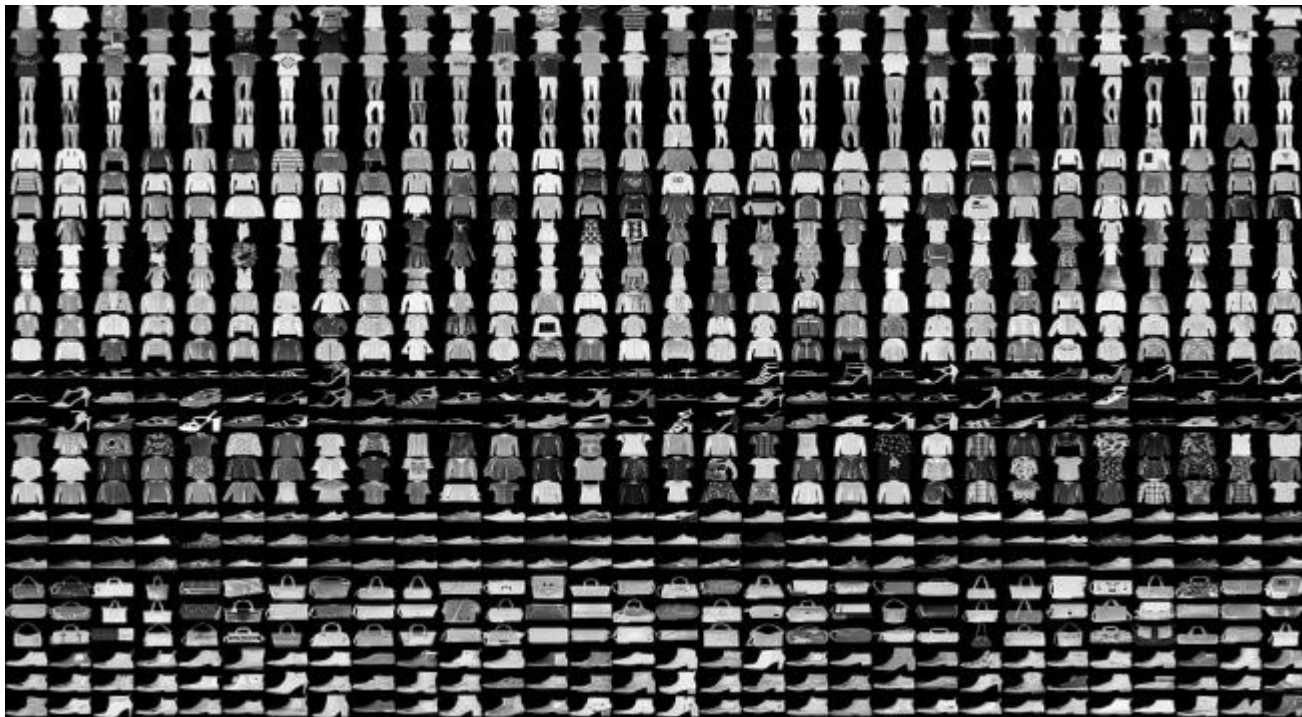


영상 처리와 컴퓨터 비전

❖ 컴퓨터 비전에서 많이 사용하는 데이터 셋

➤ Fashion MNIST

- 손글씨 우편번호를 빠르게 읽기 위해 우편봉투로부터 숫자를 직접 추출하여 만든 손글씨 숫자 데이터셋인 MNIST를 일반화, 표준화하여 발전시킨 MNIST에 기원을 두는 데이터셋으로 손글씨 숫자를 패션과 관련된 이미지로 대체한 것인데 Fashion MNIST는 10개의 카테고리 분류되는 70,000개의 흑백 이미지로 구성된 데이터셋

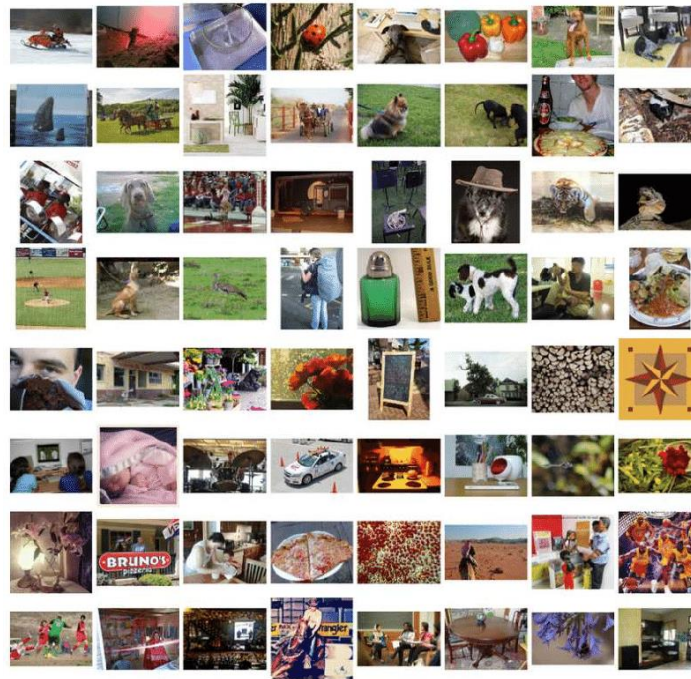


영상 처리와 컴퓨터 비전

❖ 컴퓨터 비전에서 많이 사용하는 데이터 셋

➤ ImageNet

- ImageNet은 1,400만 개 이상의 공개 이미지 데이터셋으로 스탠포드 대학의 Fei-Fei Li 교수와 그녀의 공동 작업자가 컴퓨터 비전 분야에 있어 학습 데이터의 부족 문제 해결을 위해 2009년에 시작한 데이터셋으로 그 분류가 매우 다양해서 일상생활에서 볼 수 있는 거의 모든 종류의 이미지를 얻을 수 있음



영상 처리와 컴퓨터 비전

❖ 컴퓨터 비전에서 많이 사용하는 데이터 셋

➤ MS COCO

- COCO는 Common Objects in COntext의 약자로 딥러닝 프로그램을 교육하는 데 사용되는 가장 인기 있는 오픈소스 데이터베이스 중 하나로 Object Detection, Segmentation, Keypoint Detection 등에 쓰임
- COCO 데이터는 이미지를 설명하는 캡션도 함께 제공하여 이미지를 설명하는 문장을 학습시켜 유사한 이미지가 주어졌을 때 설명을 자동으로 생성할 수 있도록 하는 데 큰 기여를 함
- COCO 데이터셋에 대한 관심도 많지만 이 데이터들을 저장하는 COCO 데이터셋의 방식인 데이터 포맷에 대한 관심도 많음



영상 처리와 컴퓨터 비전

❖ 컴퓨터 비전에서 많이 사용하는 데이터 셋

➤ Cityscapes

- 자율주행에서 사용될 수 있는 도로 환경과 관련된 가장 유명한 데이터셋으로 50개 도시의 거리 장면에서 기록된 다양한 스테레오 비디오 시퀀스 세트를 포함하는 새로운 대규모 데이터셋으로 도시 장면에 대한 의미론적 이해에 중점을 둠

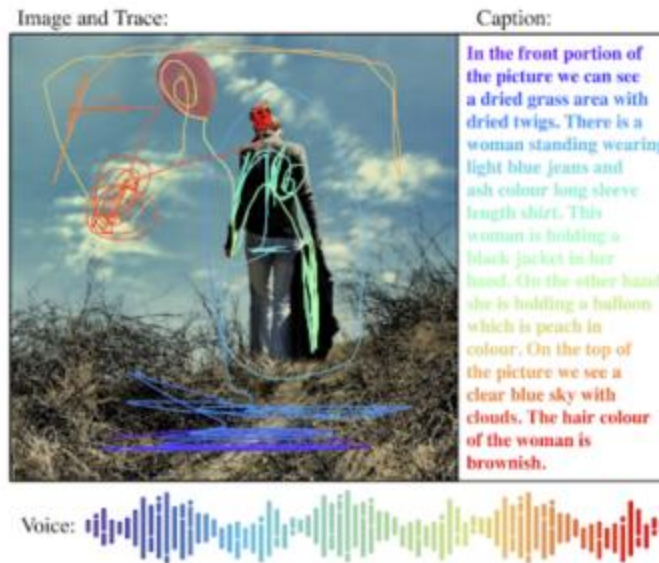


영상 처리와 컴퓨터 비전

❖ 컴퓨터 비전에서 많이 사용하는 데이터 셋

➤ Open image

- 구글이 머신러닝을 위해 2016년에 공개한 이미지에 주석이 달린 데이터셋으로 약 190만개에 전문 라벨러들이 라벨링을 검수한 이미지들을 포함하고 있음
- 데이터셋은 V1부터 계속 업데이트 되어 2020년 2월 기준으로 가장 최신 버전인 V6버전까지 공개되었음
- 이미지 상의 객체간 관계를 보여주는 라벨링, 이미지를 설명하는 캡션과 음성 나레이션이 추가되어 이미지와 캡션 사이를 연결하는 복잡한 모델을 만드는 데 유용하게 사용되고 있음



Open CV

❖ Open CV

➤ Open CV 설치

- Python: pip install opencv-python
- Anaconda: conda install --channel <https://conda.anaconda.org/menpo/opencv3>

➤ Open CV 버전 확인

```
import cv2  
cv2.__version__
```

'4.5.1'

Open CV의 주요 모듈

❖ Open CV의 주요 모듈

