

디지털시대의융합예술

Convergence of Art in Digital Age

12주차 1교시

아트테크놀로지

수업 목표

- ✓ 컴퓨터 비전(Computer Vision)의 개념과 활용에 대해 알아본다.
- ✓ 컴퓨터 비전이 활용된 예술 프로젝트와 컴퓨터 그래픽 기술을 살펴본다.
- ✓ 산업 별 컴퓨터 비전 기술이 응용되는 분야를 탐색한다.

어르신 돌봄 AI 말동무 로봇

- ▶ 자연어 처리 기술 기반 AI 기업들은 홀로 지내시는 어르신들의 말동무가 되어주는 AI 인형을 개발하여 보급하고 있다. AI 말동무 인형은 어르신 스마트 돌봄 체계 구축, 취약계층의 정서적 안정 도모 및 고독사 예방을 위한 서비스에 활용되고 있다.
- ▶ AI 말동무 인형은 안정감에 도움이 되는 봉제 인형으로 제작되었으며 인형 내부에 AI 캡슐이 탑재되어 대화 음성을 분석하는 자연어 처리 기술이 탑재되어 있다. 기존 AI 스피커 처럼 딱딱한 정보 전달이 아닌 감성적인 대화가 가능하다는 점에서 차별화를 갖는다. 어르신 대화를 통한 감정 분석을 통해 치매와 우울증, 자살, 고독사 등을 예방하는 것을 목적으로 한다.



미스터마인드(주)가 개발한 인공지능(AI) 어르신 말동무 인형. (사진=미스터마인드 제공).



비디오머그 - VIDEOMUG ✓
구독자 123만명



06:29~8:06

<https://youtu.be/jviRfd2rhyU>

교감형 AI(emotional AI)

▶ AI 전문매체 애널리틱스 인사이트((Analytics Insight)에 따르면 교감형(감정) AI 시장이 2030년 496억달러(약 70조 8000억원)으로 성장할 것으로 예상

1. 공공서비스

감정 AI 기술과 감시 카메라의 조합으로 사람들의 얼굴 표정 정보 수집이나 공공 보안 분야 활용

2. 의료 진단 케어

음성 분석을 통해 우울증, 치매를 진단하며 팬데믹 기간동안 감정 관련 케어에 대한 수요가 증가

3. 교육

학습 소프트웨어에 감정 AI를 도입하여 학습자의 감정을 인식하여 난이도를 조절하거나 자폐 아동이 타인의 감식을 인식하도록 도움을 주는 프로그램이 등장

4. 자동차 안전

감정 AI 기술을 사용해 운전자의 감정 상태를 모니터링 할 수 있음. 자동차 내부의 카메라와 마이크를 활용하여 운전자가 졸음 운전을 하거나 극단적 감정에 치우칠 때 경고 메시지를 전달

컴퓨터 비전

- ▶ 컴퓨터 비전(Computer Vision: CV)이란 컴퓨터가 상황을 인식하고 물체를 구별(객체 인식+이해)할 수 있는 시각적 능력(visual ability) 을 갖추게 하는 컴퓨터 과학 학문
- ▶ 사람의 사물을 보고 상황을 파악하여 인지하는 작업 능력을 그대로 컴퓨터가 수행
→ 컴퓨터가 이해할 수 있도록 이미지의 특징을 추출하는 행위
- ▶ 컴퓨터 비전에서 주로 활용하는 영상 정보는 주로 밝기, 색상, 텍스처 등이 있으며 이러한 정보들을 딥러닝, 특히 CNN(Convolutional Neural Network) 알고리즘을 함께 사용하여 사물을 인지함.

컴퓨터 비전



이미지를 인식하여
핸드폰으로 스캔하는 서비스



얼굴을 인식하여 다양한 효과를
적용하는 카메라 서비스



명함을 인식하여
연락처를 관리해주는 서비스



컴퓨터 비전

- ▶ 컴퓨터에 연결된 카메라가 인식한 영상에서 물체와 사람을 식별하고, 3차원의 대상의 구조를 추측하며, 각 객체의 관계, 행동, 의도, 감정 등을 이해하도록 하는 기술.
- ▶ 카메라는 센서로 들어오는 아날로그 빛의 정보를 2차원 숫자 정보 배열인 픽셀로 변환할 수 있지만 그 이상의 작업은 불가하다. 카메라는 사진을 찍을 수는 있지만 보지 못한다. 여기서 '본다'라는 말은 '이해한다'라는 뜻이다.
- ▶ 지난 5억 4천만년에 걸쳐 인류는 시각을 발전 시켜왔는데 그 대부분의 시간은 우리 뇌의 시각처리능력을 발달 시키는데 소요되었지, 생리학적인 눈(안구)를 만드는데 쓰인 것이 아니다. '본다'라고 하는 것은 눈에서 시작되나 그것의 쓰임이 발현 되는 것은 우리의 뇌 이다.



이미지넷 프로젝트

- 스탠포드 대학의 페이페이 리(Fei Fei Li)교수의 컴퓨터 비전 연구 팀은 컴퓨터(기계)에게 사람의 시각 능력을 부여할 방법을 연구했다. 그들은 아이들이 현실 세계에서 경험하며 보는 방법을 배우는 것에 착안했다. 아이의 생물학적 눈을 한 쌍의 카메라 라고 했을 때, 눈이 평균 움직이는 시간인 200밀리초(ms) 마다 한 장의 사진을 찍는 것으로 가정 했다. 아이는 3살까지 수 억 장의 방대한 학습 사진을 보게 된다. 페이페이 리는 이러한 방식으로 컴퓨터에게 보는 법을 가르치는 이미지넷(ImageNet) 프로젝트를 시작했다.



이미지넷 프로젝트

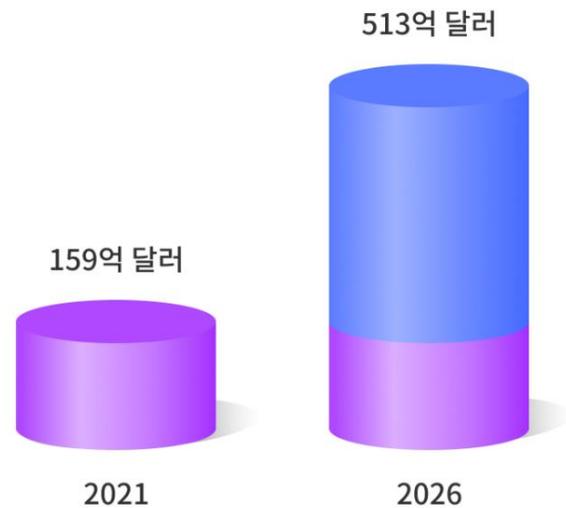
- ▶ 이미지넷 프로젝트는 약 10억 장에 이르는 이미지를 다운로드 하여 이미지에 레이블 (label)을 붙였다. 세계 167개국에서 5만명에 가까운 작업자가 10억장의 후보 이미지 정리 및 분류를 했다. 프로젝트가 시작되던 2007년은 지금처럼 빅데이터에 대한 투자가 많지 않던 시절이어서 연구팀은 연구 자금 마련에 어려움을 겪기도 했다. 그러나 포기하지 않고 연구를 계속해 2009년에 객체와 사물을 2만 2천개의 범주로 분류한 1500만 장의 이미지 데이터 베이스를 완성하여 공개한다. 이 방대한 이미지 데이터 베이스를 모든 사람들이 활용할 수 있도록 무료로 공개했다.
- ▶ 이미지넷 프로젝트의 데이터는 추후 딥러닝 기술과 만나면서 오늘날의 인공지능 기술 발전에 기초적인 자양분을 공급했다. 미래에는 컴퓨터 비전을 활용하여 의사들이 쉬지 않는 컴퓨터의 눈을 이용해 환자를 진단하고, 자동차는 똑똑하고 안전하게 도로를 주행하게 될 것이다. 재난 지역에서 위험에 처한 부상자들을 구조하고, 새로운 개척지를 탐험하게 될 것이다.

- TED 강연 *How we teach computers to understand pictures* | Fei Fei Li 중

이미지넷 프로젝트

글로벌 컴퓨터 비전 시장 규모

연평균 약 26.3%의 복합성장률(CAGR) 예상



글로벌 컴퓨터 비전 시장 규모(Source : Research and Market)

Research and Market이 발표한 자료에 의하면, 2021년 컴퓨터 비전에서의 AI 시장 규모는 159억 달러로 추정하며 2026년까지 이 시장은 513억 달러에 이를 것으로 예상

<https://blog-ko.superb-ai.com/everything-about-computer-vision-data/>

컴퓨터 비전의 전개

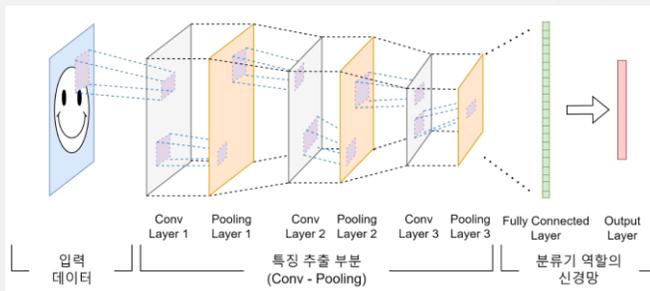
▶ 카메라가 내장된 모바일 기기의 등장으로 사진과 동영상이 기하급수적으로 증가함



▶ 컴퓨터 가격이 저렴해지면서 많은 사람들이 쉽게 구매하게 되었음 → 사용자의 증가



▶ CNN(Convolutional Neural Network: 합성곱 신경망)과 같은 시각적 영상 분석 인공지능 알고리즘의 활용



컴퓨터 비전 요구지식

디지털 영상 처리 및 컴퓨터 그래픽스 : 영상 편집 조작 능력

광학 : 카메라 구조 및 영상획득

선형대수학 : 영상 데이터를 행렬로 표현하기 위함

수학적 이해 : 확률과 통계 기하학 등

신호 처리 : 영상 또한 신호의 한 종류이기 때문

패턴인식 : 추출 된 특징을 분류하고 이해하기 위함

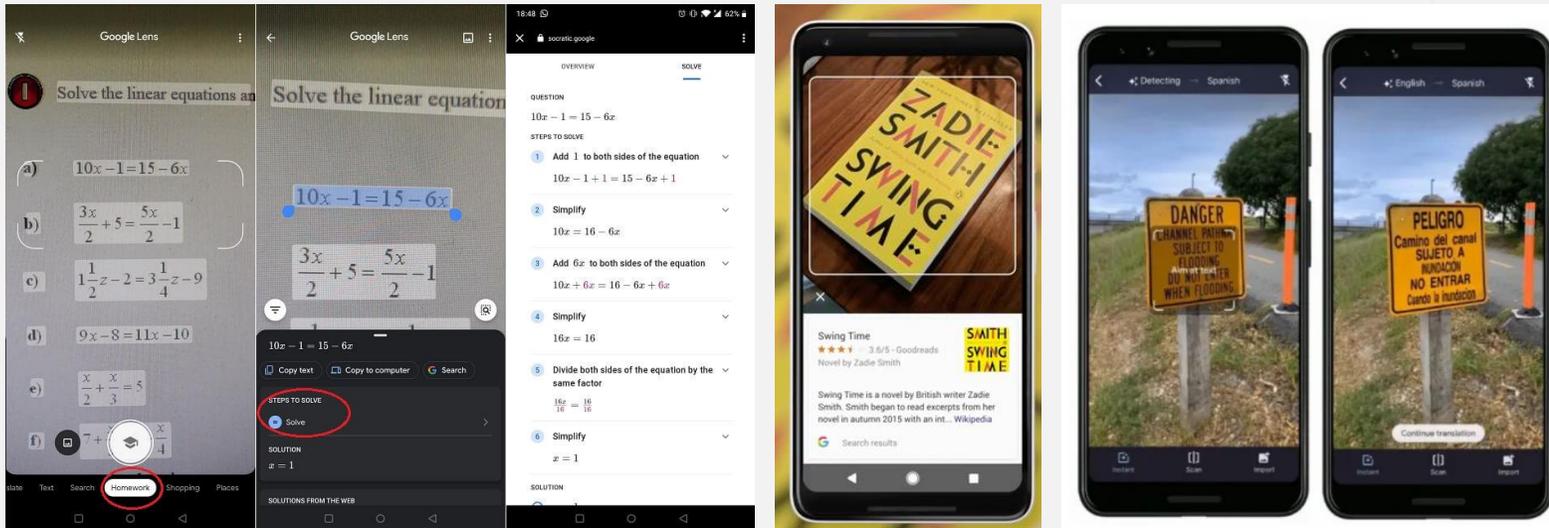
확률론 : 패턴을 찾기 위해서 데이터의 분포를 다뤄야 함

컴퓨터과학 : 수치 해석, 알고리즘 사용, 최적화 작업



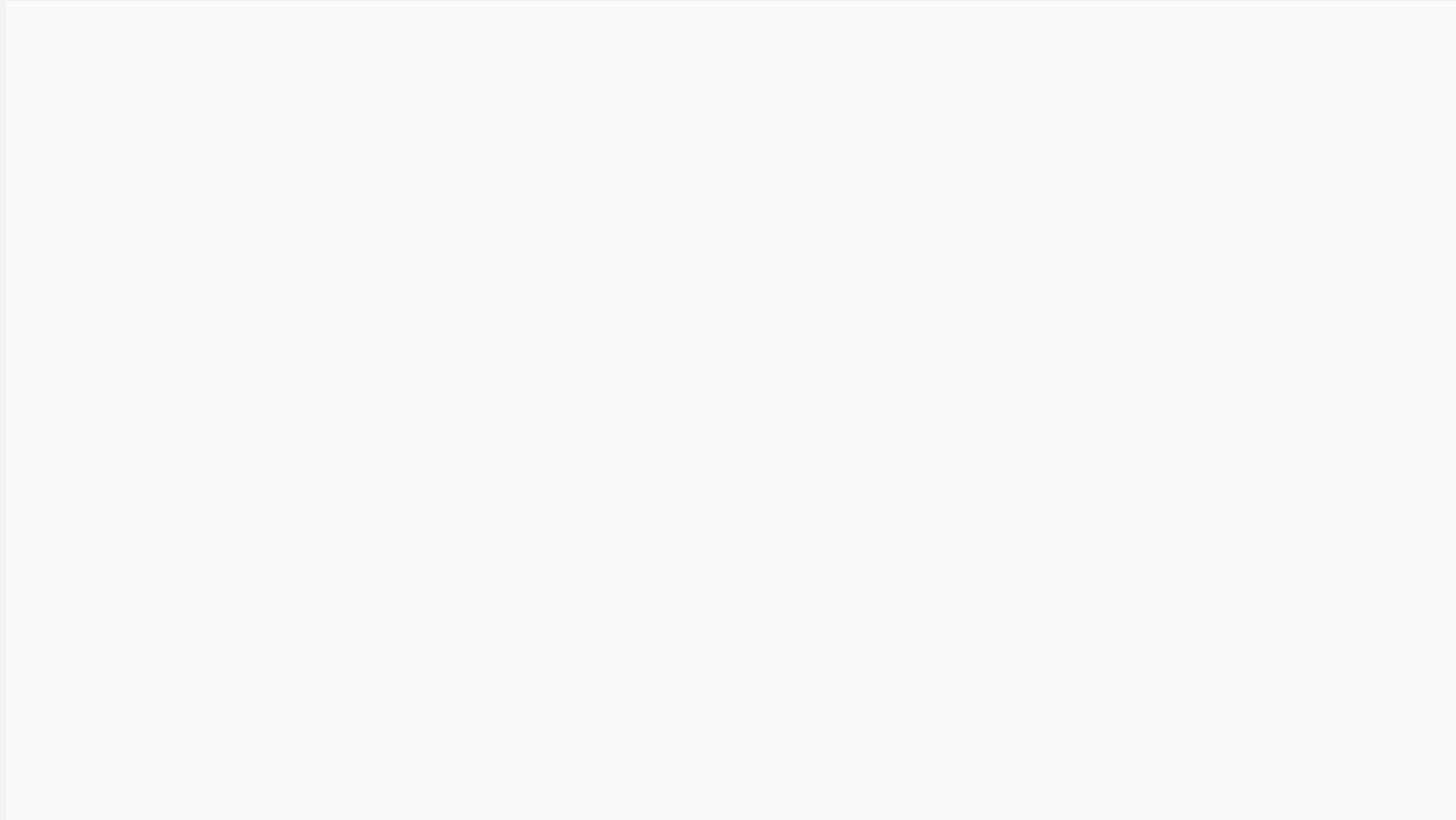
구글 렌즈

구글 어시스턴트 컴퓨터 비전 서비스 렌즈를 통해
외국어 번역 및 주변 사물 정보 확인가능



- <https://www.maketecheasier.com/awesome-google-lens-features/>
- <https://www.smartprix.com/bytes/how-to-translate-text-on-your-android-phone-using-google-lens/>
- <https://www.businessinsider.com/guides/tech/how-to-use-google-lens>

구글 렌즈



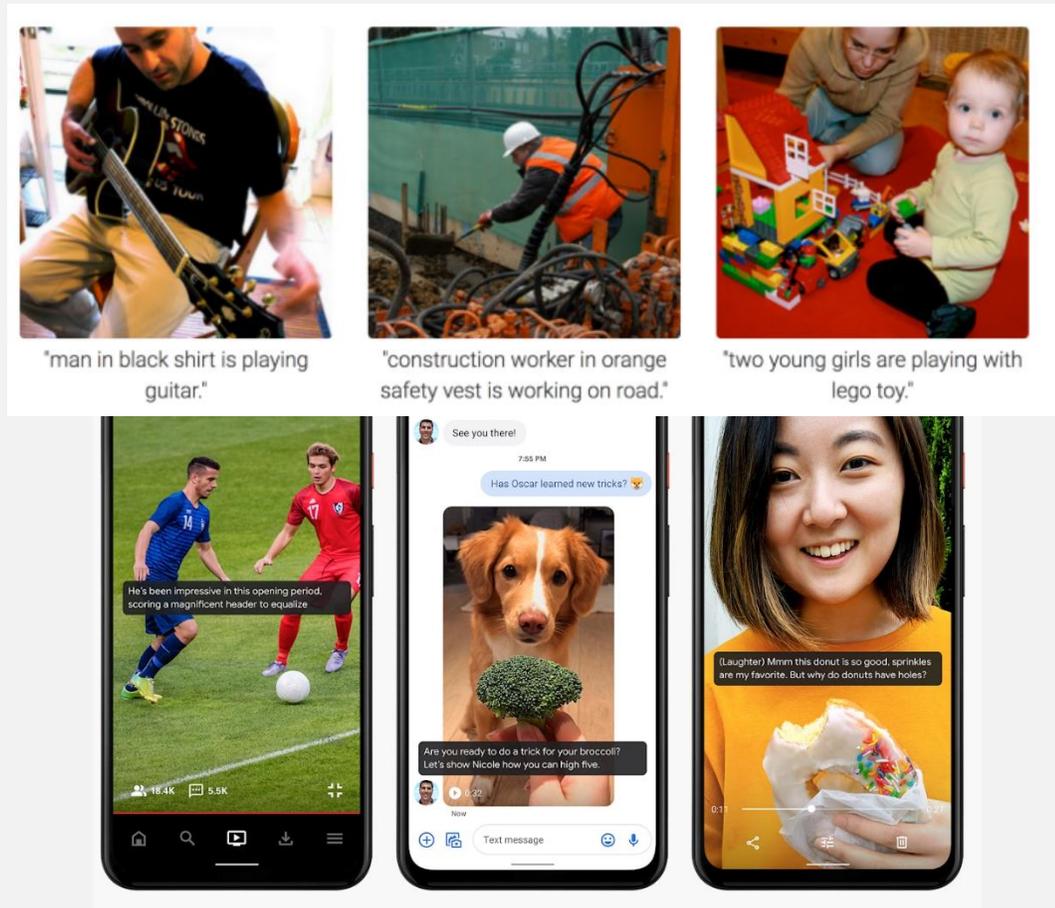
00:31~2:23

https://youtu.be/BH_6BdgTdiw



이미지넷 자동 이미지 캡션 생성 Image Captioning

컴퓨터가 이미지를 인식하여 관련 내용으로 문장을 생성 하거나, 휴대전화에서 재생되는 미디어를 자동으로 캡션하는 온디바이스(on-device) 라이브 캡션 기능이 있음



이미지넷 자동 이미지 캡션 생성 Image Captioning



https://youtu.be/6wuh4vLj_xg

픽셀 재귀 슈퍼 해상도 Pixel Recursive Super Resolution

- > 구글 인공지능 연구팀(Google Brain)이 저해상도 이미지에서 고해상도 이미지를 생성하는 픽셀 재귀 슈퍼 해상도(Pixel Recursive Super Resolution) 기술을 발표했다.

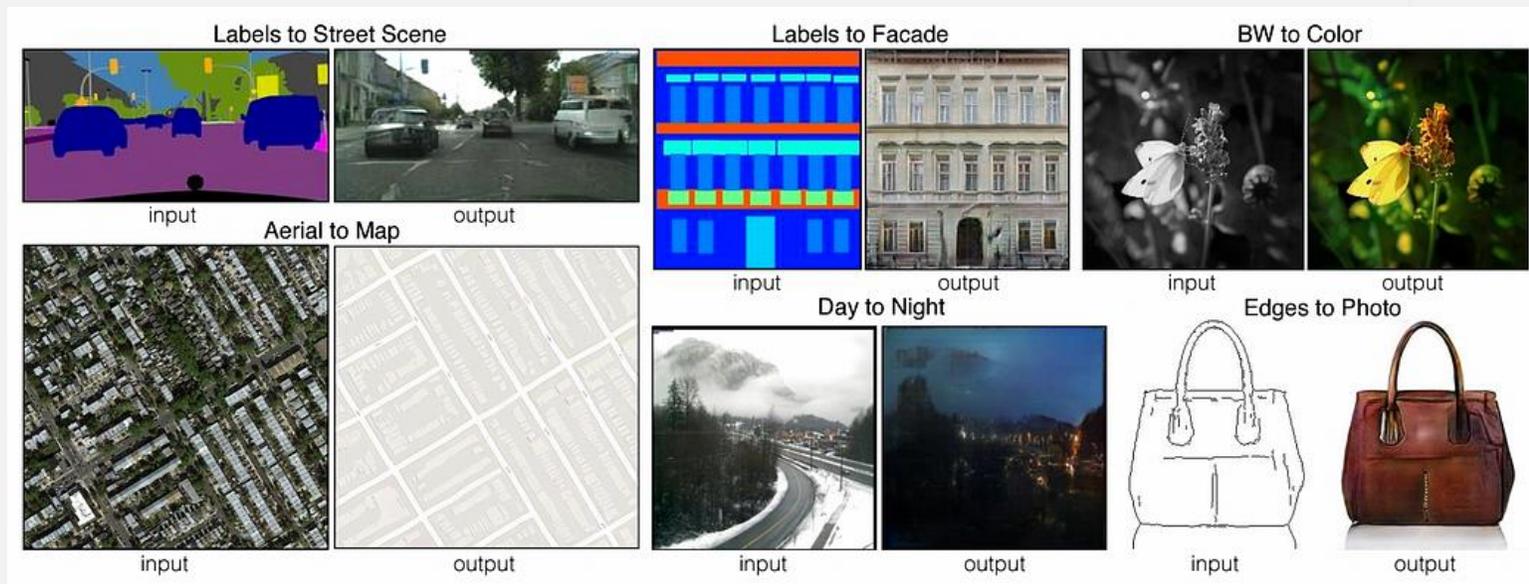
이 기술은 디지털 화상에서 모자이크를 제거할 수 있고 압축한 이미지에서 선명하게 원본 이미지를 예상할 수 있다.

다른 사진과 교차비교 하여 비슷한 색상 패턴을 보이는 사진을 비교하여 복원한다.



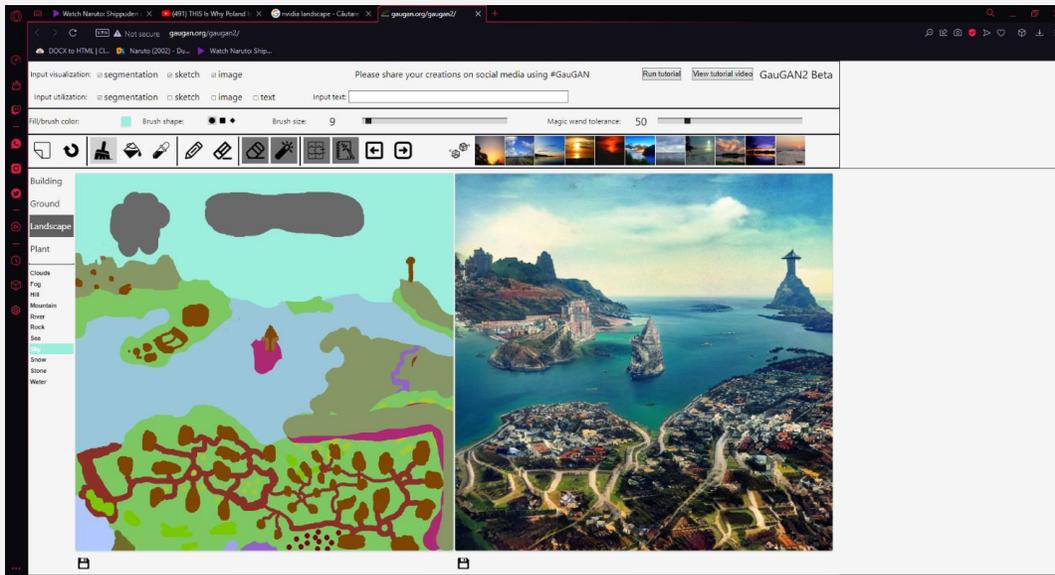
GAN Generative Adversarial Network

- GAN(Generative Adversarial Network, 생성적 적대 신경망)알고리즘 기술을 사용하여 이미지를 생성하거나 조합, 변형한다. GAN을 이용하면 적은 양의 데이터로 원본 이미지를 예측하거나 사진을 특정 방식으로 변환할 수 있으며 진짜처럼 보이는 가짜 이미지의 생성도 가능하다.

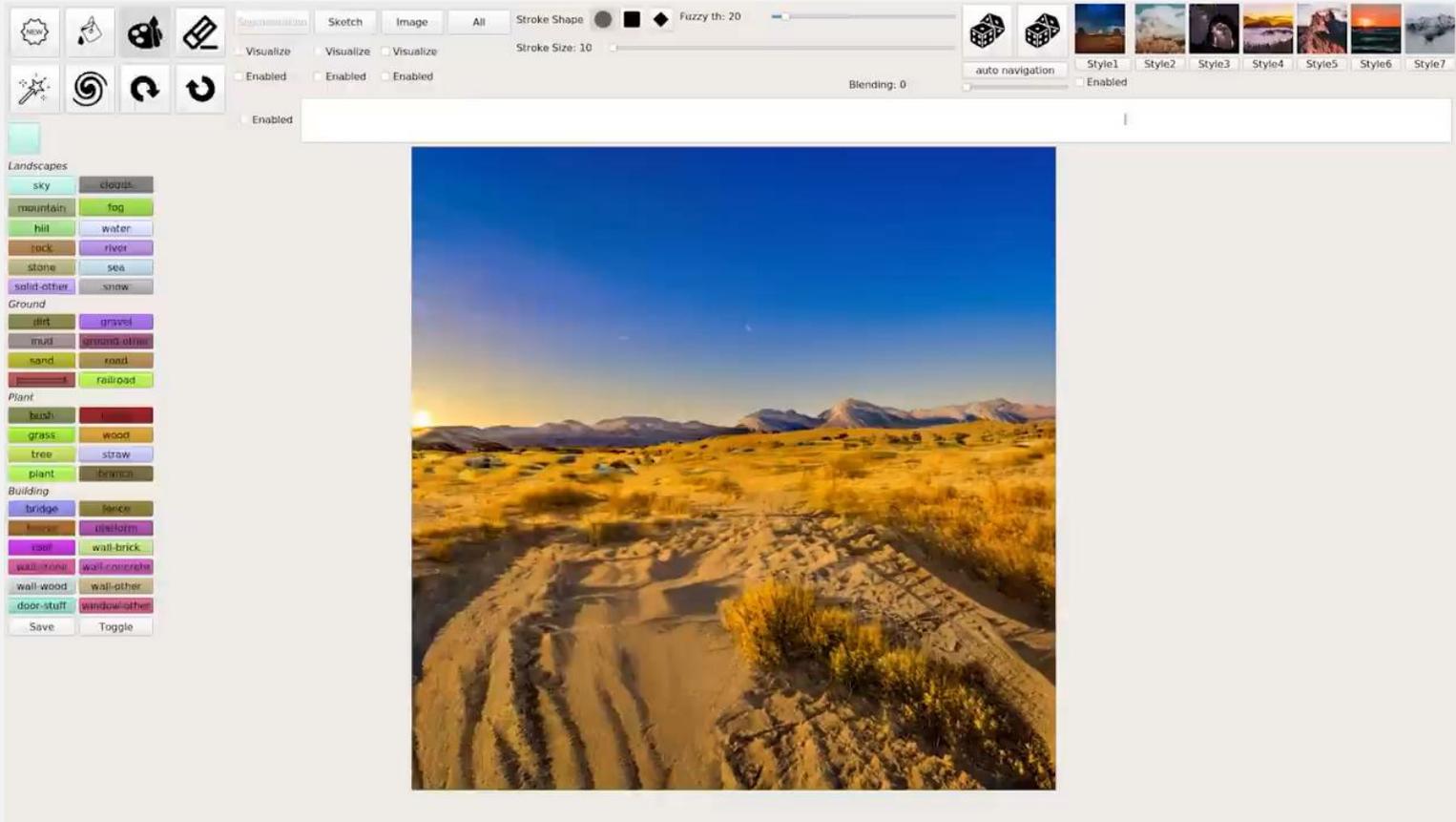


NVIDIA GauGAN2

- ▶ 엔비디아(NVIDIA)는 GAN 알고리즘을 활용한 이미지 크리에이터 GauGAN을 출시했다. 세그멘테이션 맵(Segmentation Map)의 원리로 구간을 설정하여 선을 그리고 색상을 채워 넣으면 그에 대응하는 사진 그래픽이 생성된다.
- ▶ 대략적인 스케치로 고화질의 가상 이미지를 손쉽게 만들 수 있으며 두번째 버전에서는 text-to-image기능을 제공하여 텍스트 프롬프트를 입력하여 원하는 이미지에 대한 세밀한 조정과 커스터마이징을 제공한다.



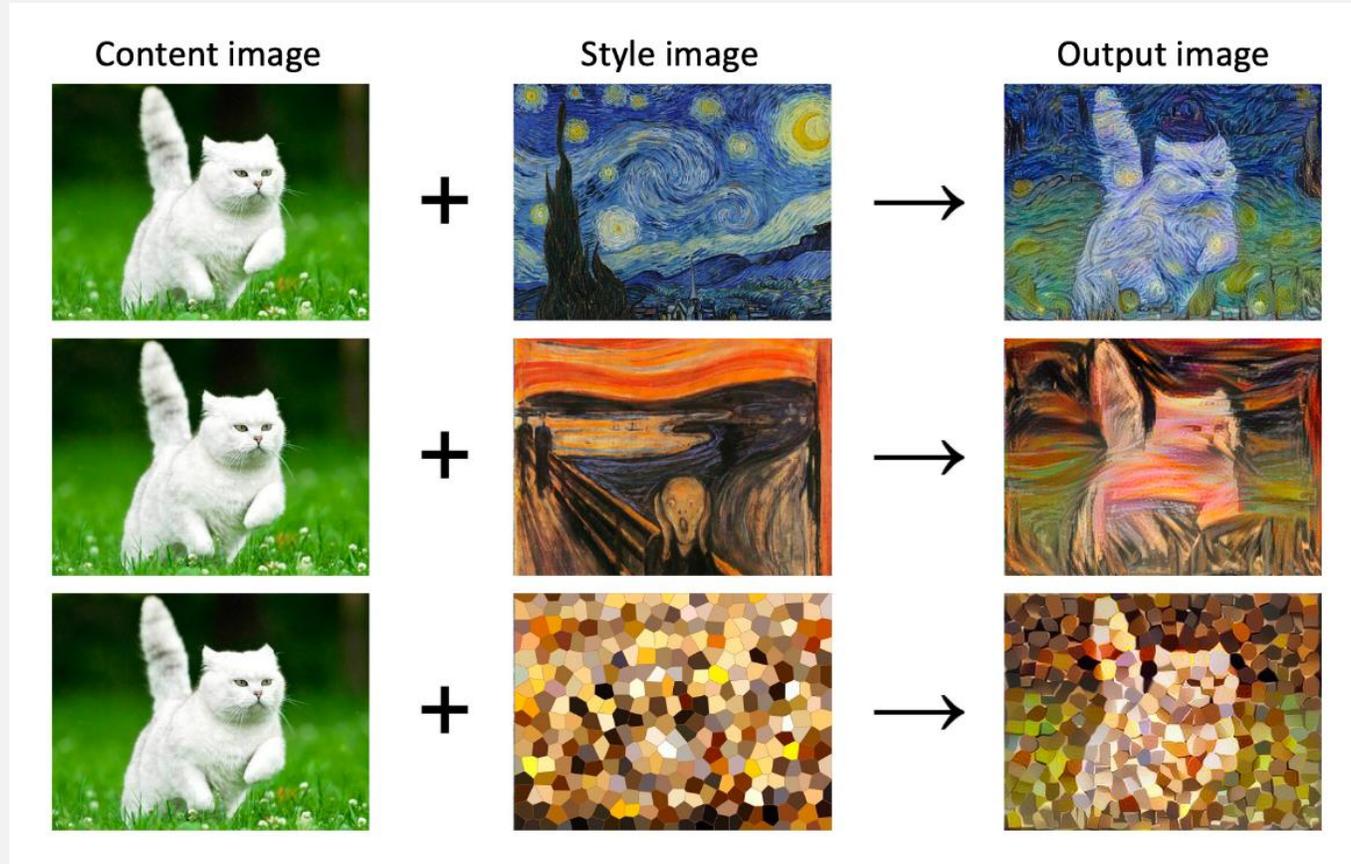
NVIDIA GauGAN2



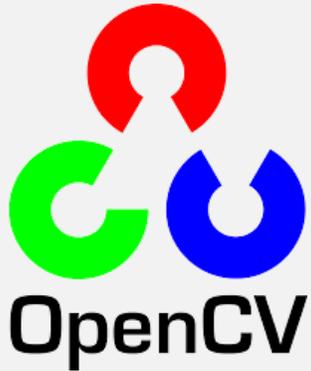
<https://youtu.be/5Qdr8AUNYqU>

스타일 전이(Neural style transfer)

원본 이미지(콘텐츠 이미지)를 스타일 이미지에 맞춰서 스타일 이미지 형태로 원본 이미지를 전이 시키는 기술



Open CV Open Source Computer Vision



- ▶ OpenCV는 인텔에서 개발한 실시간 이미지 처리와 컴퓨터 비전 애플리케이션 개발을 위한 오픈소스 프로그래밍 라이브러리이다. 2000년에 인텔에서 최초 개발되었으며 현재 비영리 단체 OpenCV.org가 관리하고 있다.
- ▶ OpenCV 기능 : 이미지 처리, 영상 스티칭 및 3D 재구성, 특징 검출, 객체 인식 및 추적, 기계학습 및 딥러닝
- ▶ 컴퓨터 비전 기술을 구현하기 위해 필요한 알고리즘 모음 기능 패키지로 텐서플로우 (TensorFlow), 파이썬(PyTorch)과 같은 딥러닝 프레임워크를 지원하여 프로그래머에게 개발 환경을 지원한다.
- ▶ OpenCV는 BSD(Berkely Software Distribution)라이선스를 따르기 때문에, 상업적 목적 사용이 가능하여 학교에서 교육용으로, 회사에서 상업용으로 제품을 만들 때 활용해도 무방하다는 장점이 있다.
- ▶ 컴퓨터 비전과 관련된 이론 및 수학 기초가 조금 부족하더라도, OpenCV 함수를 불러와 필요한 기능을 직접 구현할 수 있다는 점도 편리하다.

컴퓨터 비전 데이터셋

- ▶ 오늘날 발전된 컴퓨터 비전의 기술은 데이터셋을 공개적으로 구축한 여러 오픈소스 프로젝트의 기여로 성립될 수 있었음
- ▶ 지금처럼 이미지 데이터가 충분하지 않았던 과거부터 수집된 이미지나 영상 데이터가 모여져 공개된 데이터셋들이 오늘날 컴퓨터 비전의 발전에 지대한 역할을 함

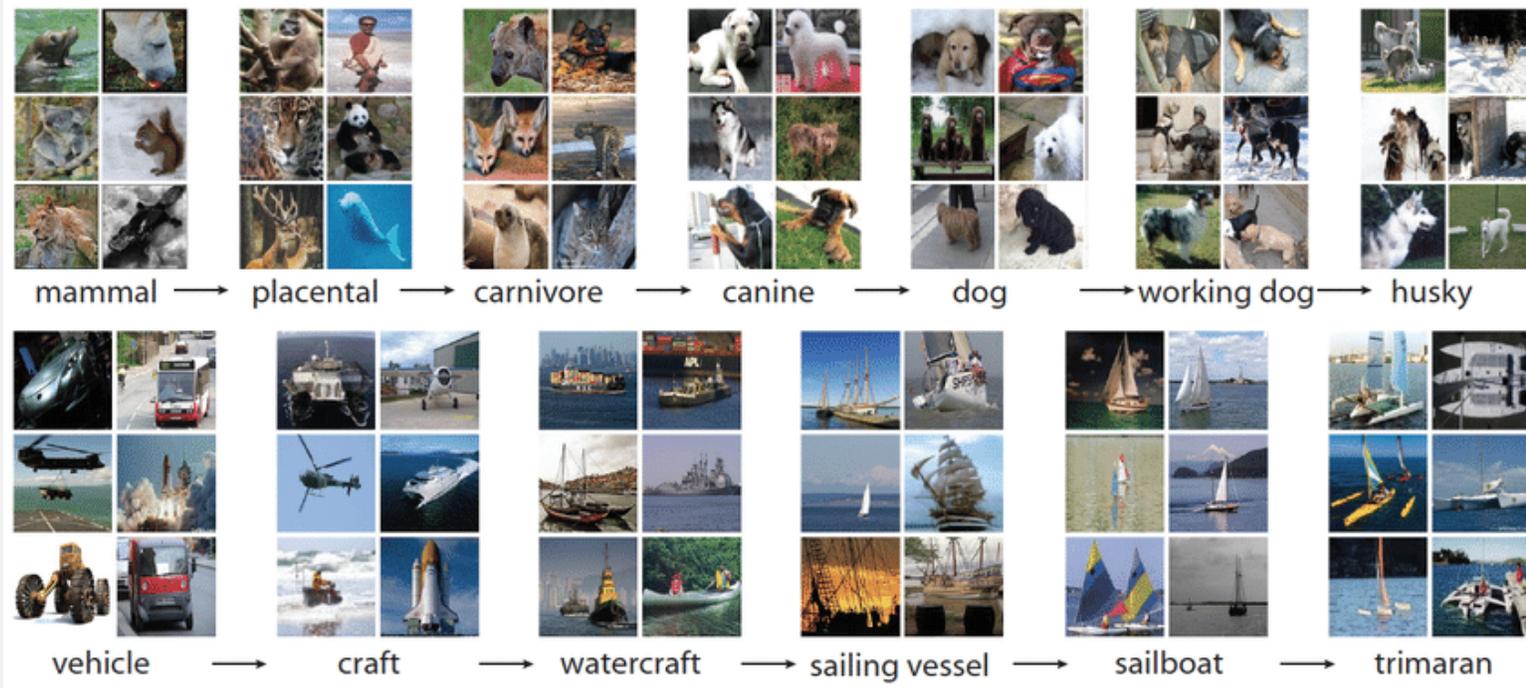


컴퓨터 비전 데이터셋 ImageNet

- ▶ ImageNet은 1,400만 개 이상의 공개 이미지 데이터셋으로, 스탠포드 대학의 페이페이 리(Fei-Fei Li) 교수와 프린스턴 대학의 카이리 교수와 컴퓨터 비전 분야에 있어 학습 데이터의 부족 문제 해결을 위해 시작한 데이터셋 프로젝트
- ▶ 분류가 매우 다양해서 일상생활에서 볼 수 있는 거의 모든 종류의 이미지를 얻을 수 있음



컴퓨터 비전 데이터셋 ImageNet

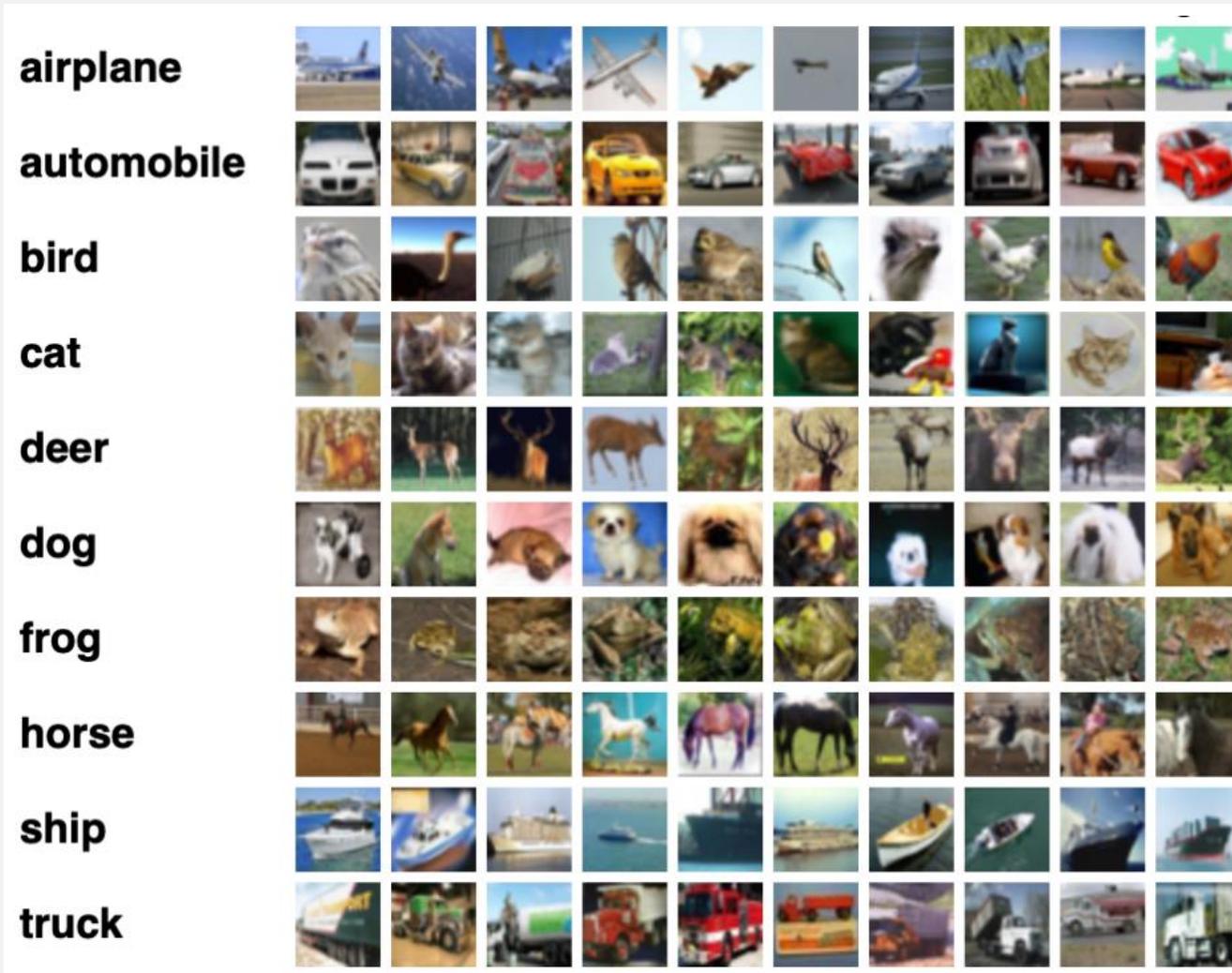


컴퓨터 비전 데이터셋 CIFAR-10

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



컴퓨터 비전 데이터셋 CIFAR-10

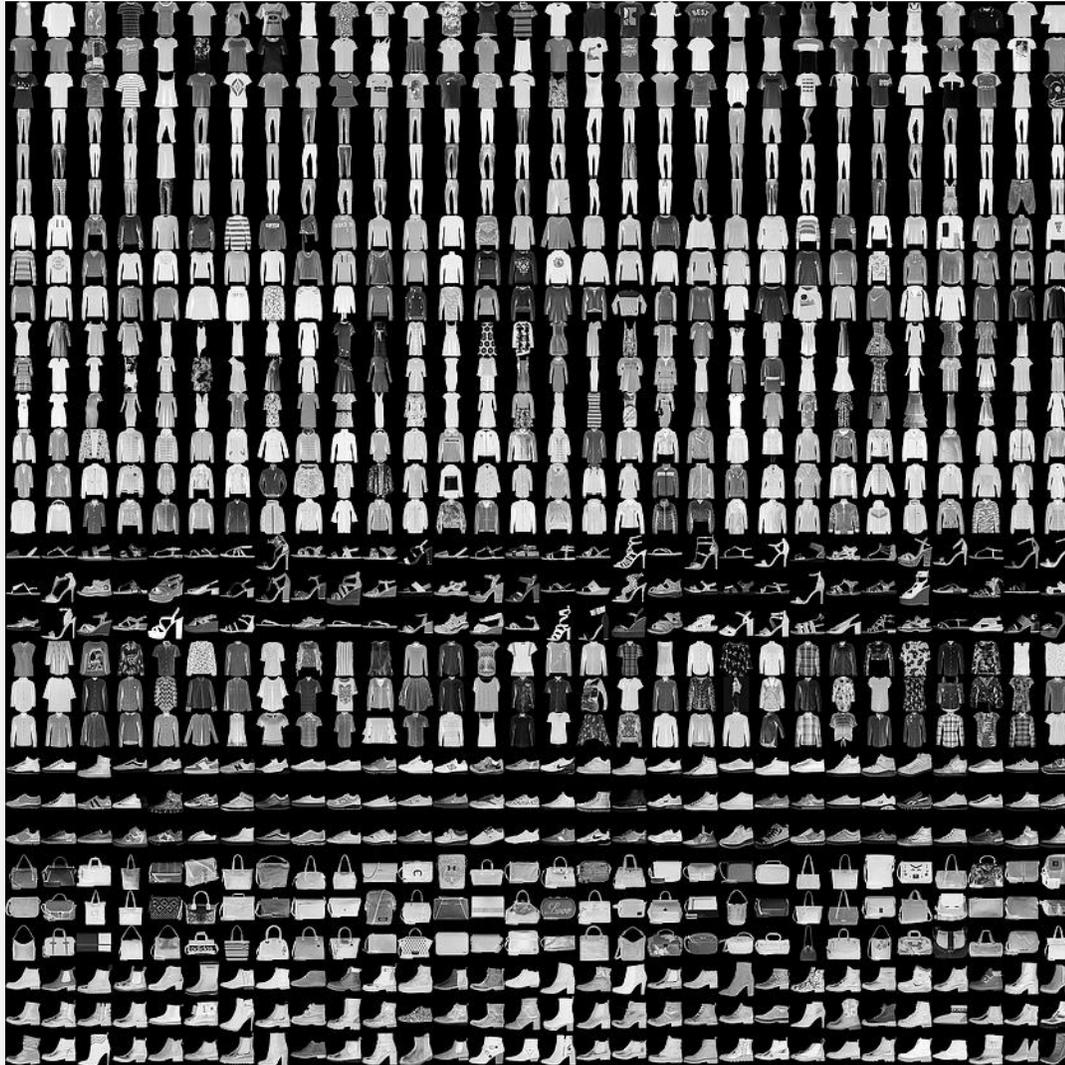


컴퓨터 비전 데이터셋 Fashion MNIST

- 손글씨 우편번호를 빠르게 읽기 위해 우편봉투로부터 숫자를 직접 추출하여 만든 손글씨 숫자 데이터셋인 NIST를 일반화 및 표준화 시킨 MNIST에 기원을 두는 데이터셋
- 손글씨 숫자를 패션과 관련된 이미지로 대체한 것임. Fashion MNIST는 10개의 카테고리 고리로 분류되는 70,000개의 흑백 이미지로 구성된 데이터셋



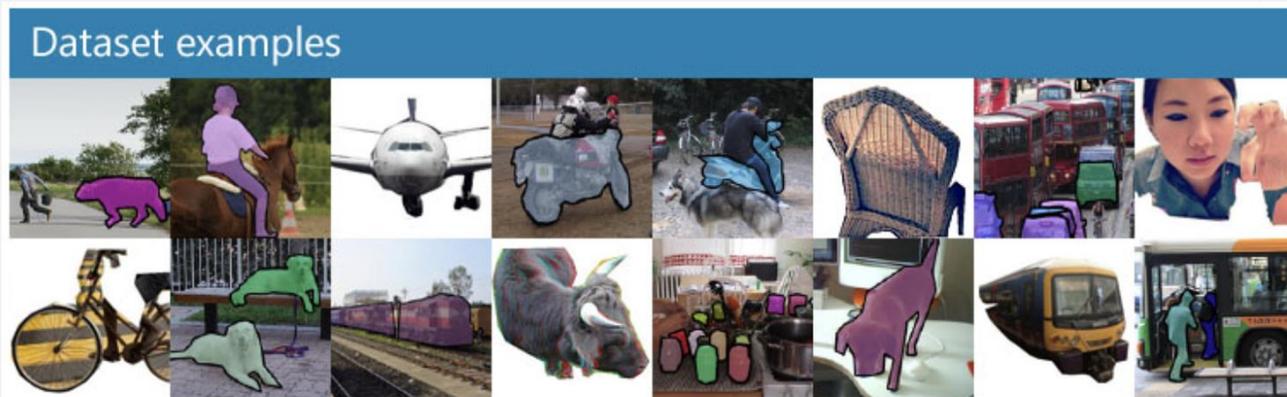
컴퓨터 비전 데이터셋 Fashion MNIST



<https://www.tensorflow.org/tutorials/keras/classification?hl=ko>

컴퓨터 비전 데이터셋 COCO dataset

- ▶ COCO(Common Objects in Context)는 딥러닝 프로그램을 교육하거나 연구논문의 성능비교 지표로 사용되는 오픈소스 데이터베이스 중 하나이다.
- ▶ Object Detection, Segmentation, Key point Detection 등에 사용되며 COCO2017의 경우 학습(training) 데이터셋 118,000장의 이미지, 검증(validation) 데이터셋 5,000장의 이미지, 테스트(test) 데이터셋 41,000장의 이미지로 구성된다.
- ▶ COCO 데이터는 이미지를 설명하는 캡션도 함께 제공하여, 이미지를 설명하는 문장을 학습시켜 유사한 이미지가 주어졌을 때 설명을 자동으로 생성할 수 있도록 하는 데 큰 기여를 하였다.



컴퓨터 비전 데이터셋 Cityscapes

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



디지털시대의융합예술

Convergence of Art in Digital Age

12주차 2교시

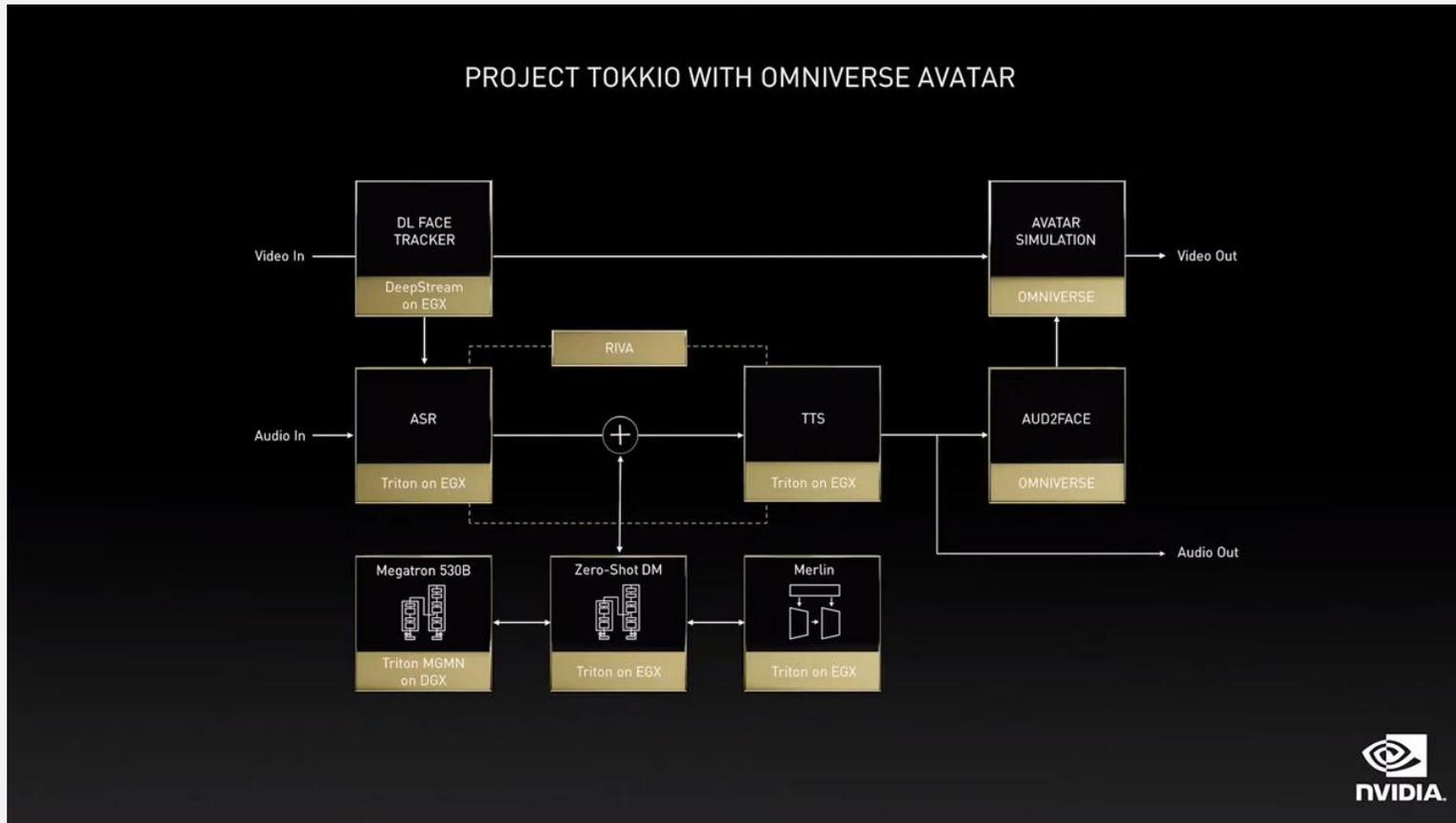
아트테크놀로지

컴퓨터 비전

- › 컴퓨터 비전 학문은 **센서 디바이스 기술, 이미지 및 영상분석 기술, 인공지능 기계학습/딥러닝 기술** 등을 종합하여 인간의 시각적 능력과 인식 능력을 복제하고자 함
- › 일반적인 컴퓨터비전 어플리케이션은 **클라우드에 있는 대량을 시각적 데이터를 이용하여 학습된 알고리즘을 구축하며 새로운 사물은 학습된 알고리즘을 이용하여 시각적 패턴을 매칭하여 유사성을 찾아 사물을 구별** 함



컴퓨터 비전



-
-
-
-
-

<https://youtu.be/-Sc1QIn5Dtg>

-
-
-
-
-

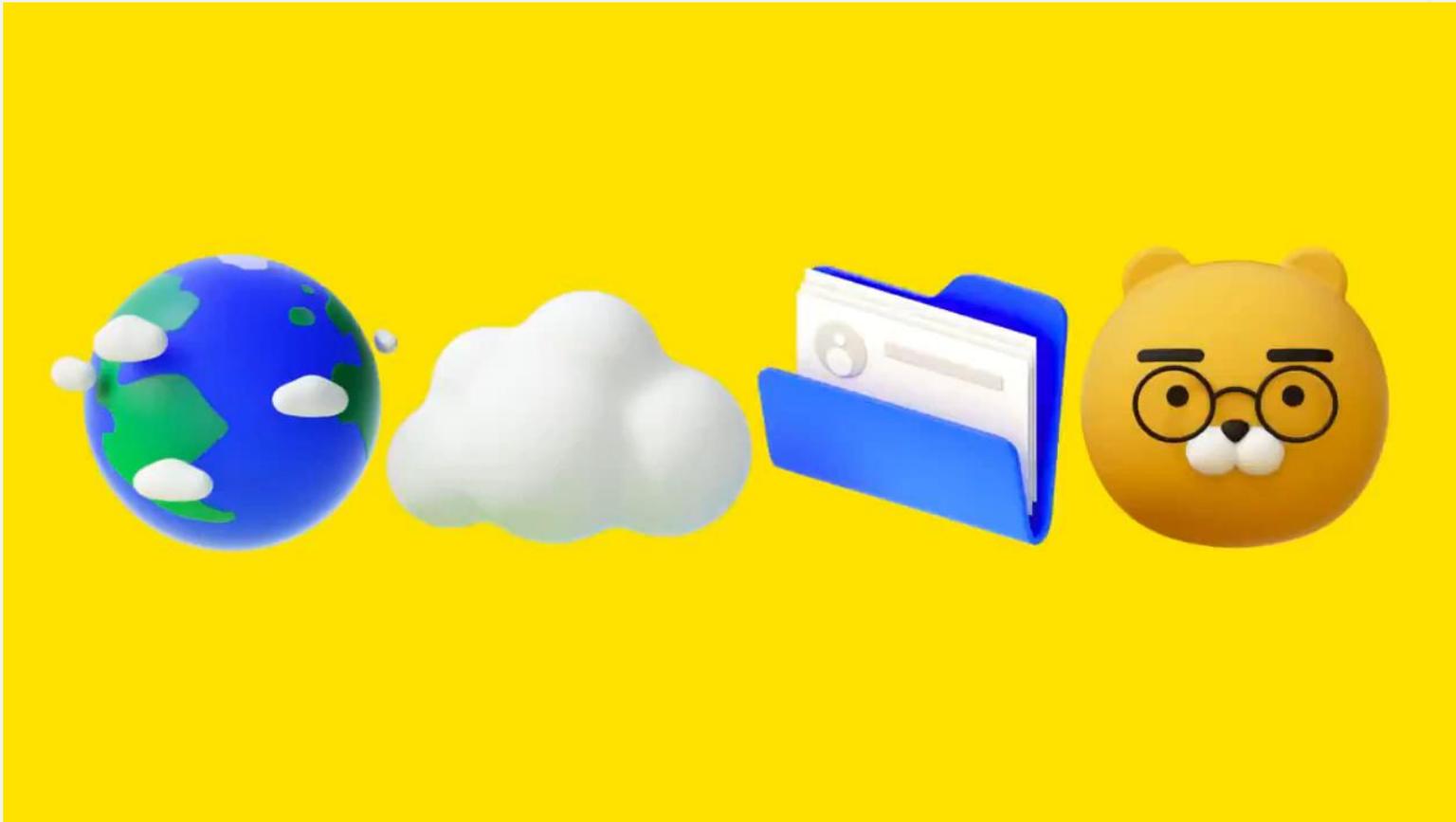
컴퓨터 비전의 작동

1. 클라우드에 있는 대량의 시각적 데이터를 기반으로 훈련시켜 학습된 알고리즘을 구축
2. 센서 디바이스로 새로운 객체를 영상을 캡처하여 데이터로 저장
3. 캡처한 영상 데이터를 학습된 알고리즘에 입력하여 특성별로 분류하고 유사성 패턴 추적
4. 유사성 패턴 매칭을 통하여 새로운 객체를 인식하고 구별

컴퓨터 비전의 작동

[Kakao i Engine] 컴퓨터 비전 Computer Vision

 Kakao Enterprise
구독자 4.01천명



00:35~2:02

<https://youtu.be/5eeRSvKYpQo>

컴퓨터 비전의 작동

1. 이미지 수집과 획득

동영상, 사진 또는 3D 등 분석할 이미지를 실시간으로 수집한다.

2. 이미지 처리

딥 러닝 모델에는 이미지 수집 프로세스가 자동화 되어 있다.
데이터 정규화를 위해 이미지 전처리(preprocessing)

3. 이미지 파악

마지막 단계는 객체를 식별하거나 분류하여 패턴을 추출하는 해석 단계

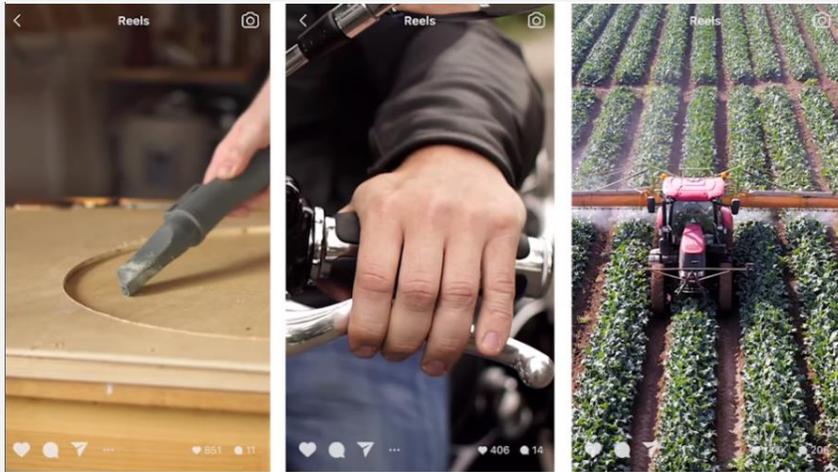
인스타그램 릴스(Reels)

- ▶ 메타(페이스북 AI 연구소 FAIR) 연구진은 AI 영상 인식기술을 기반으로 인스타그램 릴스(15~30초 이내의 짧은 영상) 추천 시스템을 구축하였다.
- ▶ 페이스북에 업로드된 비디오(공개용)에서 음성, 텍스트, 시각 정보를 자동으로 학습하도록 한 것이다. 이렇게 설계된 프로젝트를 '이미지 수집영상으로부터 학습(Learning from Videos) 프로젝트' 라 한다.
- ▶ 매일 전세계에서 업로드 되는 페이스북 영상을 AI가 학습하여 다양한 지역의 문화에서 통용되는 뉘앙스와 시각적인 식별 신호를 인식하고 그것을 분별해 내는 정확도를 향상시킬 수 있다.



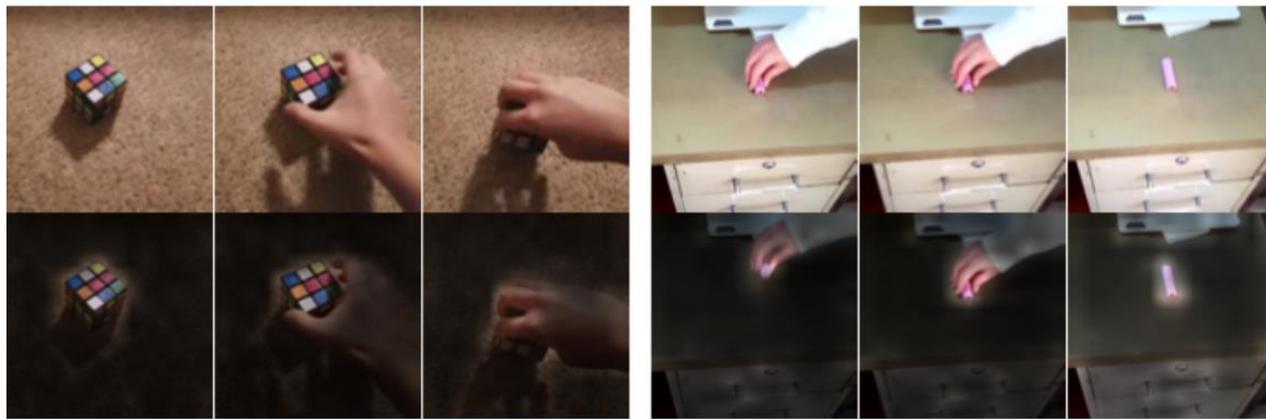
인스타그램 릴스(Reels)

- ▶ 인스타그램 릴스의 추천시스템은 영상 속 이미지와 음성을 연계해 종합적으로 분석하는 AI 모델이다.
- ▶ 릴스에는 춤을 추고 노래하는 영상이 많은데 콘텐츠 내에서 서로 다른 등장인물이 각자의 움직임 보이더라도 동일한 배경음악을 사용하는 경우가 있음
- ▶ 그렇기 때문에 영상 추천 시스템의 기준을 이미지 영상 혹은 음향으로만 선정하지 않는다.
- ▶ 효과적인 시스템 구축을 위해 시청각 데이터의 상관관계를 반영하는 GDT(Generalized Data Transformations) 모델을 적용했다.



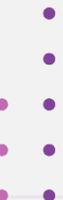
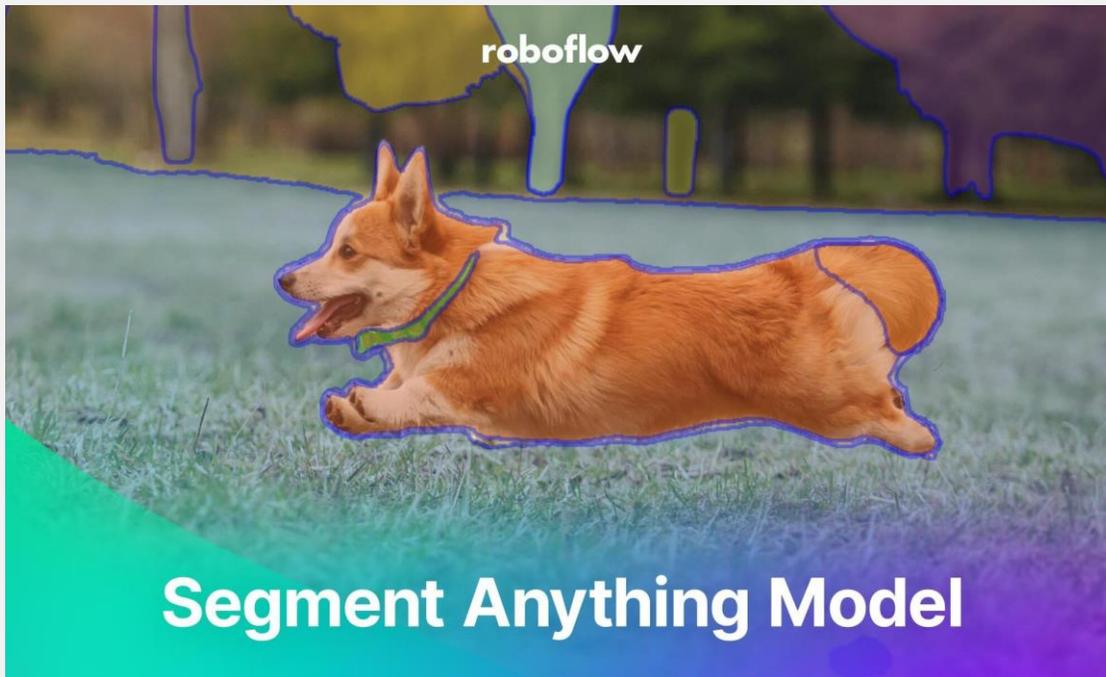
인스타그램 릴스(Reels)

- ▶ 또한 메타(페이스북 AI 연구소 FAIR) 연구진은 영상에 특화된 딥러닝 모델 '타임스포머 (TimeSformer)'을 개발했는데 이것은 영상 속 '행동 인식'이 가능하도록 설계되었다.
- ▶ 인공지능이 시간의 흐름에 따라 영상의 움직이는 대상을 감지한다. 이것은 인공지능이 시간과 공간을 구분하여 분석하지 않고 두 가지를 통합하여 인식하는 딥러닝 모델이다.
- ▶ 영상의 타임라인을 여러 단위로 쪼개어 움직임이 감지되는 곳만 인식한다. 기존에는 영상의 움직임을 인식하기 위해 프레임 전체를 학습했었으나 이제는 움직임이 발생하는 구역을 골라서 그 움직임이 얼마나 변화했는지 탐지한다.
- ▶ 연구진에 따르면 타임스포머 모델을 통해 인공지능 학습 시간과 컴퓨터 메모리 소모량이 대폭 감소되었으며 영상 분석의 정확도도 향상되었다.



컴퓨터 그래픽과 컴퓨터 비전

- ▶ 메타는 이미지와 동영상에서 객체를 분할하고 분리할 수 있는 인공지능 도구 Segment Anything을 공개했다. 이 오픈소스 툴은 대상 이미지 내부의 오브젝트에 여러 개의 선택영역을 생성할 수 있으며, 다른 시스템의 AR 또는 VR 헤드셋의 시스템 프롬프트에 응답할 수 있도록 설계되었다.



컴퓨터 그래픽과 컴퓨터 비전



• • 2:15~4:15

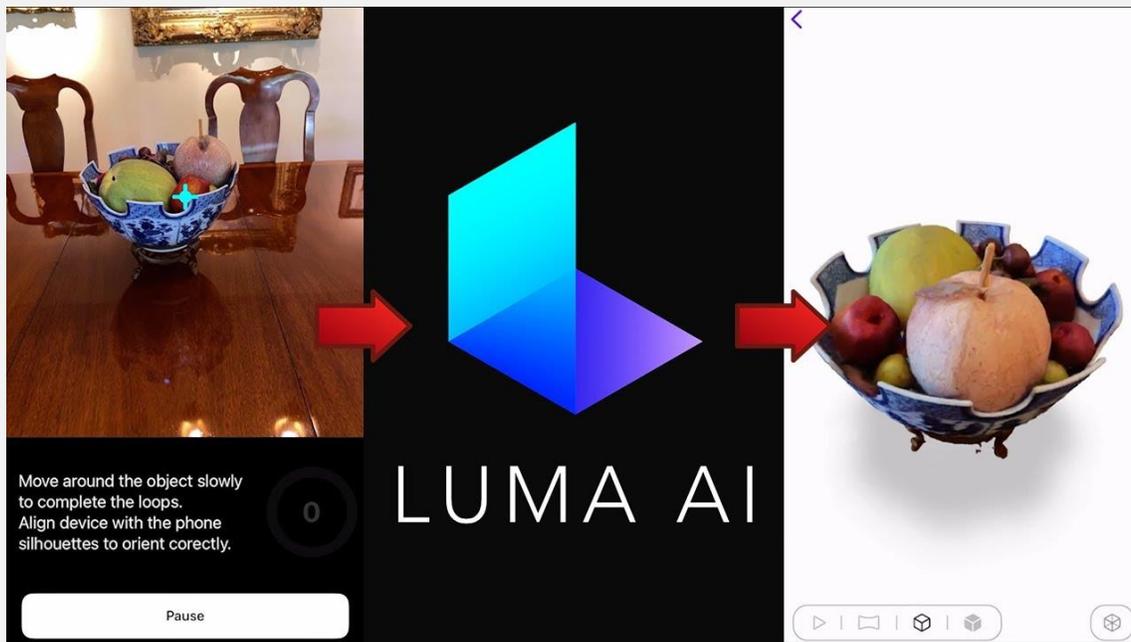
• • https://www.youtube.com/watch?v=20_vecl280A&t=378s

•



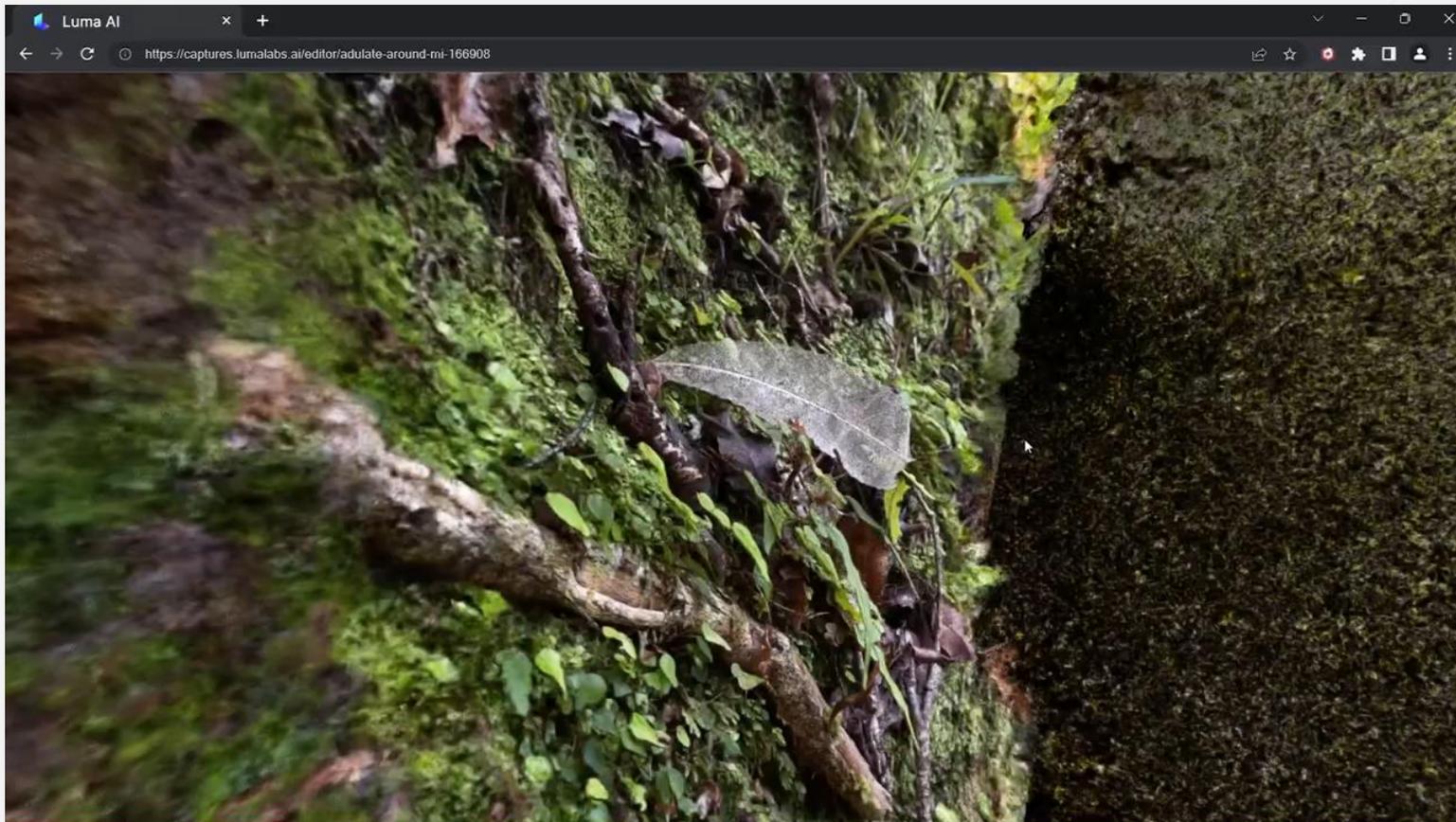
컴퓨터 그래픽과 컴퓨터 비전

- ▶ 루마 AI(Luma AI)는 이미지로 촬영한 주변 환경을 3D 버전으로 생성하는 Neural Radiance Fields (Nerf)기술을 개발하여 공개했다. 루마 AI는 언리얼 엔진(Unreal Engine)과 파트너 십을 맺고 해당 프로그램에서 사용할 수 있는 플러그인을 만들었다.
- ▶ 스마트폰과 너프(Nerf) 크리에이터 툴만 있으면 누구나 주변의 환경을 촬영하여 3D로 변환한 후 언리얼 엔진으로 불러와 생생한 입체 장면을 만들어낼 수 있다.



컴퓨터 그래픽과 컴퓨터 비전

luma Realtime NeRF on the Web



<https://youtu.be/Vt7Grog1ko>

컴퓨터 그래픽과 컴퓨터 비전

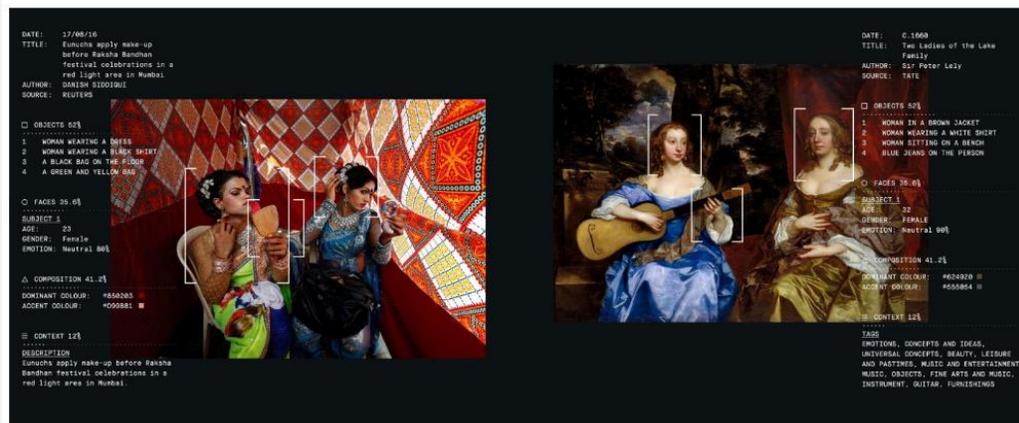
Exploring Unreal Engine 5 | NeRF in UE5



<https://youtu.be/7gfZhtokRTM>

Tate Britain X Fabrica 'Recognition'

- ▶ 영국 테이트 미술관은 2016년 마이크로소프트와 협의하여 '인공지능이 예술에 대해 대중에게 새로운 통찰력을 제공해줄 수 있는가?' 라는 문제를 다루기로 결정하였다.
- ▶ 테이트 미술관은 소장된 예술 작품에 대해 인공지능 기술을 활용하여 영국의 예술을 탐구, 조사, 이해하는데 디지털 혁신 기술을 보여줄 수 있는 디지털 크리에이터를 공모하였다.
- ▶ IK Prize 수상작으로 'Recognition(인식)'이 선정되었다. 이탈리아 베네통 그룹 커뮤니케이션 리서치 연구소 Fabrica가 출품한 이 작품은 스크린에 매일 약 1000개의 로이터 뉴스 사진과 테이트 미술관에 소장된 3만 여점의 영국 예술 작품과의 유사점을 인공지능이 찾아내어 매칭하는 알고리즘을 보여준다.



Tate Britain X Fabrica 'Recognition'



U.S. Republican presidential nominee Donald Trump speaks during a campaign event in Wilmington, Ohio, U.S. September 1, 2016. REUTERS/Carlo Allegri

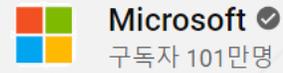
Sir Stanley Spencer *Self-Portrait* 1914, Tate © Estate of Stanley Spencer

- ▶ 관객들은 전시 현장에서 'Recognition' 작품이 만들어 지는 과정을 지켜볼 수 있었고 인공지능이 사진과 예술품의 매칭을 완료하면 그 결과물이 온라인 갤러리에 저장되고 테이트 미술관에 전시되었다.

<https://www.tate.org.uk/whats-on/tate-britain/ik-prize-2016-recognition>

Tate Britain X Fabrica 'Recognition'

Tate IK Prize: Building Recognition

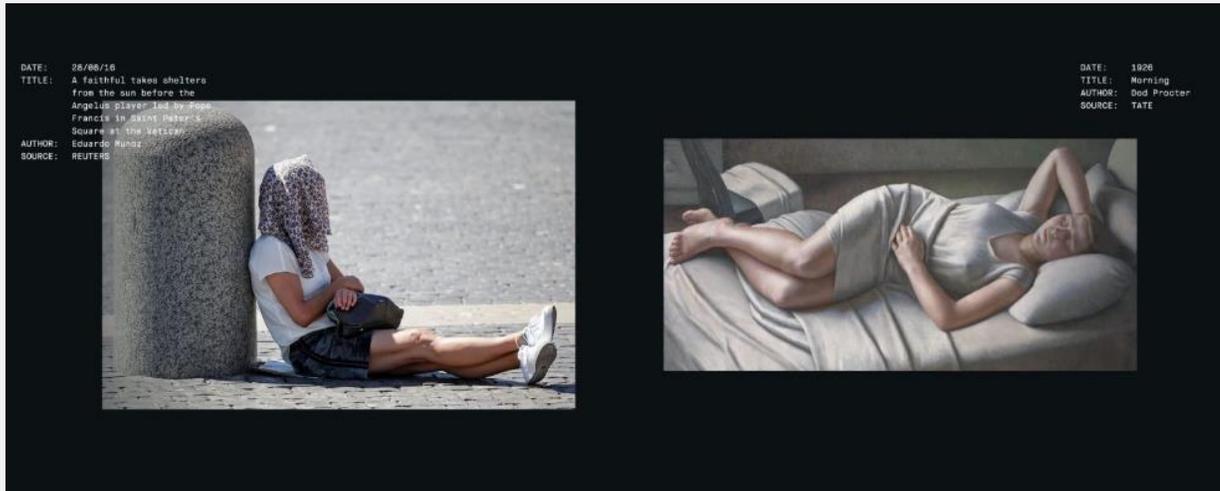


0:00~1:00

<https://youtu.be/cl1CS2qWvoQ>

Tate Britain X Fabrica 'Recognition'

- ▶ 'Recognition'작품에 사용된 인공지능은 작품 속의 패턴을 분석하고 구성 요소를 인식할 수 있도록 훈련되었으며 그것은 사물인식(object recognition), 안면인식(face recognition), 색상 및 구도 인식(color & composition recognition) 상황인식(context recognition)등의 컴퓨터 비전의 기술을 포함했다.
- ▶ 안면 인식의 경우 마이크로소프트의 컴퓨터 비전(Computer Vision) 및 감성 응용 프로그램 인터페이스(Emotion API)를 활용했다. 프로젝트에 함께 참여한 소프트웨어 개발업체 줄리브레인(JoliBrain)이 개발한 색상 및 구도 인식 기술을 통해 모양과 구조, 레이아웃과 색상의 일치 및 패턴 인식을 통해 객체 인식을 구현했다.



Tate Britain X Fabrica 'Recognition'

Tate IK Prize: Building Recognition

Microsoft  
구독자 101만명



0:00~0:57

<https://youtu.be/RNS60kwe5x4>



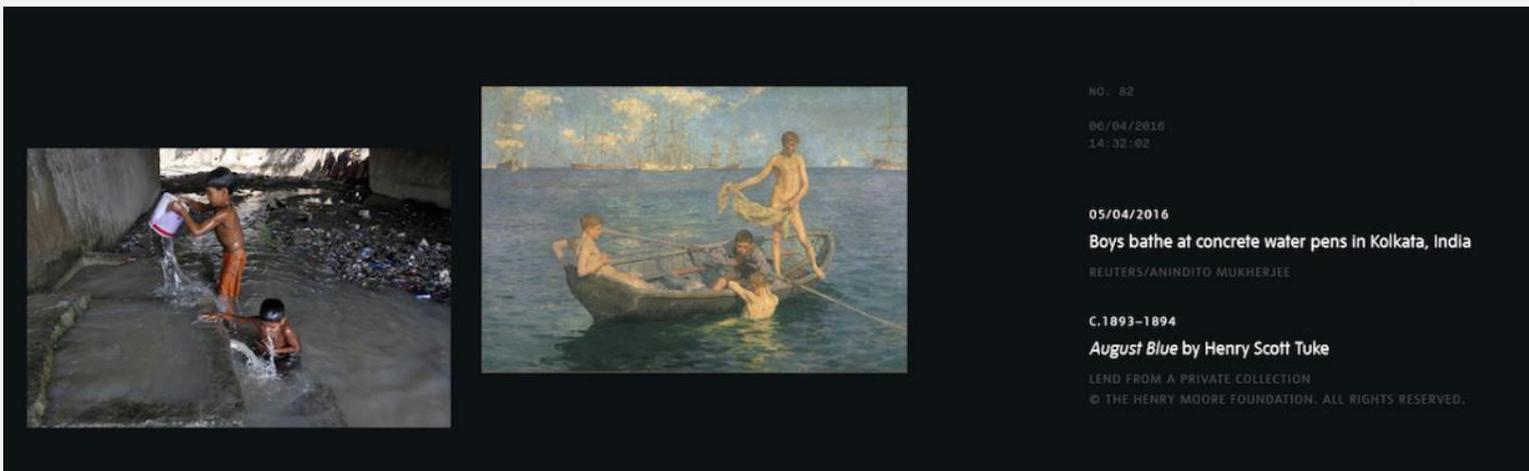
Tate Britain X Fabrica 'Recognition'

- ▶ 매일 약 1천 개의 로이터 뉴스의 이미지가 게시되었고, 인공지능은 해당 사진이 무엇에 관한 것인지 분석을 하였다. 이와 함께 테이트 온라인 아카이브 데이터베이스에 있는 3만 여점에 달하는 영국 예술 컬렉션을 검색하여 주제적 유사성과 키워드가 일치하는 작품을 매칭했다.
- ▶ 자연어 처리 기술은 이미지 캡션과 텍스트를 기반으로 상황과 주제를 분석하기 위해 사용되었다.
- ▶ 전시장에는 대화형 전시물(Interactive Installation)이 있어 관람객이 해당 뉴스 이미지와 일치하는 대안 작품을 선택하면서 인공지능을 재교육해보는 참여 형식을 가졌다.

Tate Britain X Fabrica 'Recognition'

➤ 이 프로젝트는 관객이 작품을 보면서 자연스럽게 각자의 주관적인 이해와 판단을 갖는 과정과 인공지능 알고리즘 작품을 인식하고 분석하는 접근 방식의 유사성을 탐색하는 연구였다.

더욱 강화된 컴퓨터 비전 인공지능이 등장한다면 작품 속의 사물, 얼굴, 구도, 상황과 같은 이미지 간의 시각적 일치 뿐만 아니라 작품의 함축적 의미와 상징성에 대한 정교하고 고도화된 분석을 기대할 수 있을 것이다.



산업별 컴퓨터 비전 응용 분야

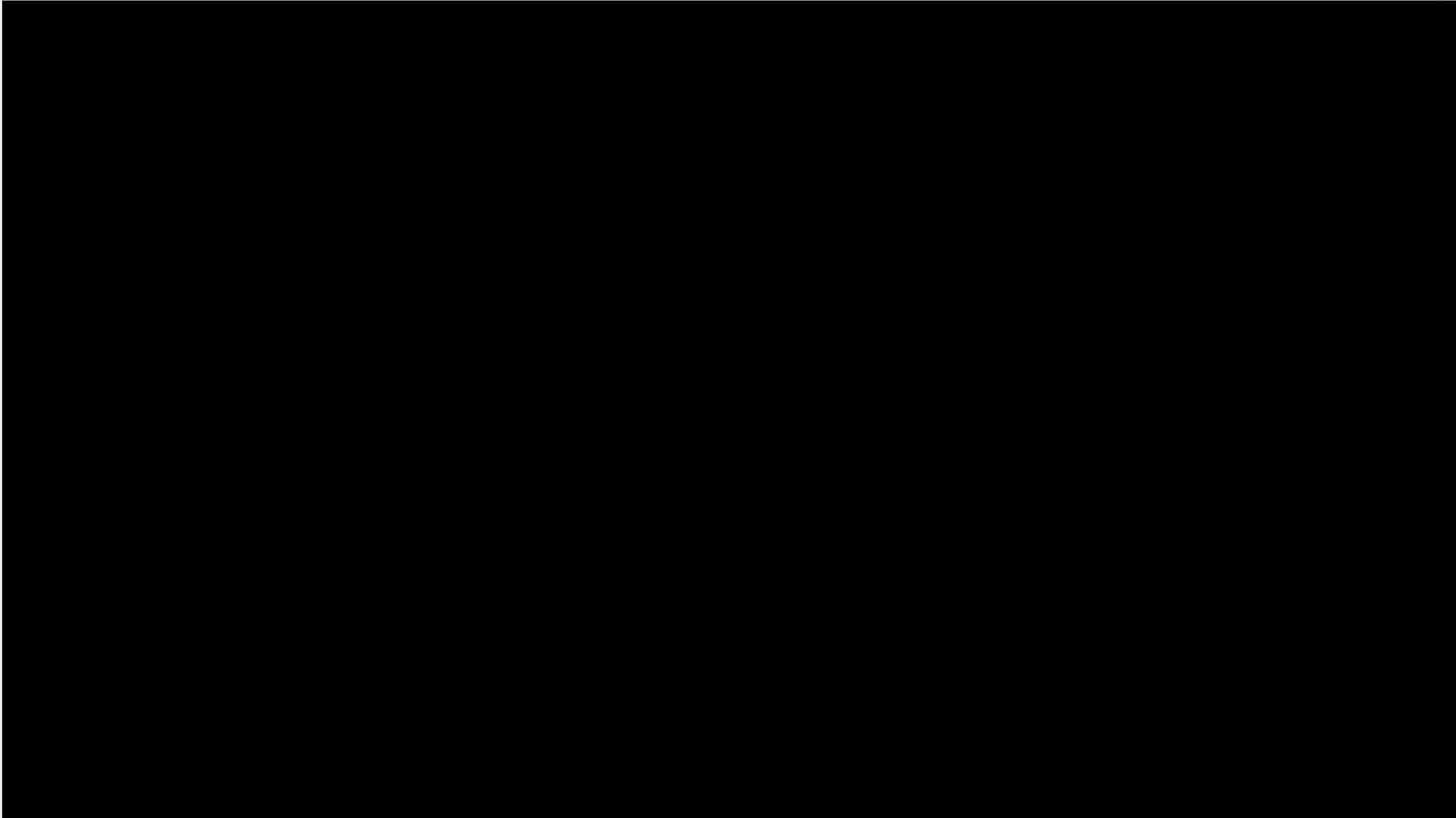
1. 항공, 자율주행 자동차, 로봇, 드론

컴퓨터 비전은 항공, 자율주행 자동차, 로봇, 드론 등에서는 실시간 객체 식별 및 추적을 통하여 주의에서 발생하는 상황에 대한 정보를 수집하고 그에 따라 동작할 수 있도록 한다.

2. 생산, 농업, 산업

컴퓨터 비전은 환경과 상황을 인식하고 생산 설비를 모니터링하고 이상 징후를 감지하고 제품의 상태를 파악할 수 있다. 제조 산업에 종사하는 기업의 경우 컴퓨터 비전을 활용하여 제품의 결함 여부를 실시간으로 식별한다. 제품의 크기가 아무리 작더라도 컴퓨터가 완제품의 이미지나 동영상을 처리하여 수십 가지의 결함을 식별한다.

산업별 컴퓨터 비전 응용 분야



0:00~1:10

<https://youtu.be/wuhbqcMzOaw>

산업별 컴퓨터 비전 응용 분야

3. 콘텐츠 구성

컴퓨터 비전을 통하여 사람이나 이미지, 영상에서 객체를 식별하고 분석한 데이터를 이용하여 콘텐츠를 구성할 수 있게 한다. 사진 인식 애플리케이션이나 소셜 미디어에서 활용된다.

4. 광학 문자 인식

컴퓨터 비전은 이미지나 영상 데이터로 부터 문자나 숫자 정보를 추출하여 활용한다.



산업별 컴퓨터 비전 응용 분야

5. 의료 분야

의료 분야에서 컴퓨터 비전 시스템은 MRI, CAT 스캔, X 레이의 이미지를 철저히 검사하여 의사처럼 정확하게 이상 징후를 감지한다. 또한 의료 전문가는 초음파 같은 3차원 이미지에 신경망을 사용하여 심박 등의 시각적 차이를 감지할 수 있다.

6. 방위 및 보안

컴퓨터 비전 기술로 사람의 얼굴을 인식하고 나이 및 감정, 행동을 인식할 수 있다. 은행이나 카지노 같은 보안이 중시되는 산업에 종사하는 기업은 컴퓨터 비전을 활용하여 거래를 교환하는 고객의 신원을 보다 정확히 식별할 수 있다. 보안 관리자와 달리, 컴퓨터 비전 알고리즘은 수백 개의 비디오 피드를 한꺼번에 분석할 수 있다.

산업별 컴퓨터 비전 응용 분야



0:00~2:00

<https://youtu.be/WXVb31y0v-Y>

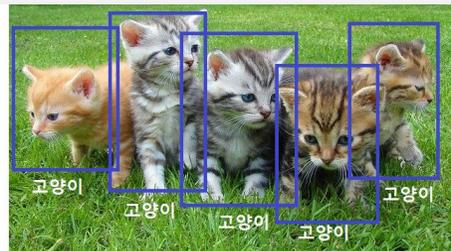
컴퓨터 비전의 목적

1. 객체 분류(Object classification)

이미지나 영상 속에서 객체가 무엇인지 판별하여 종류를 구별. 하나의 사진을 특정 개체로 분류하는 일을 수행. 사전에 정의된 집합의 이미지에 적절한 레이블(이름)을 할당하는 작업. 텍스트를 디지털화(문자인식)하고 이미지 데이터 베이스에 주석을 자동으로 생성하는 작업 가능



고양이



Single Object

Classification

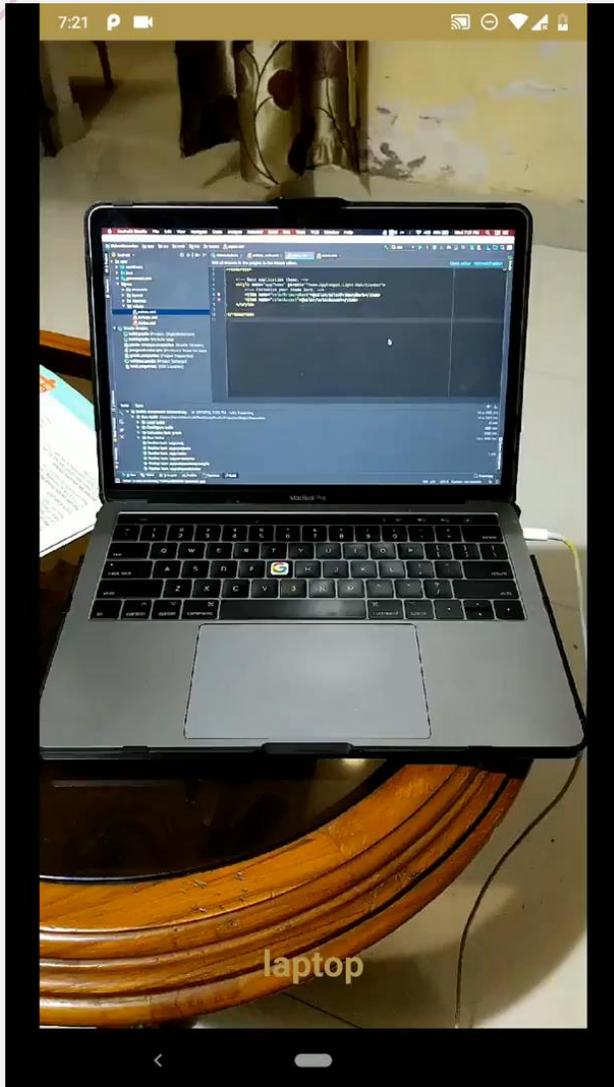


Q. "이 이미지는 무엇인가?"
 - 이미지에서 single object 감지
 A. "고양이 일 것 같습니다."
 $y = \{P_{고양이}, P_{개}, P_{오리}\}$
 $= \{0.9, 0.05, 0.05\}$

CAT

왼쪽: Image Classification, 오른쪽: Object Detection

컴퓨터 비전의 목적



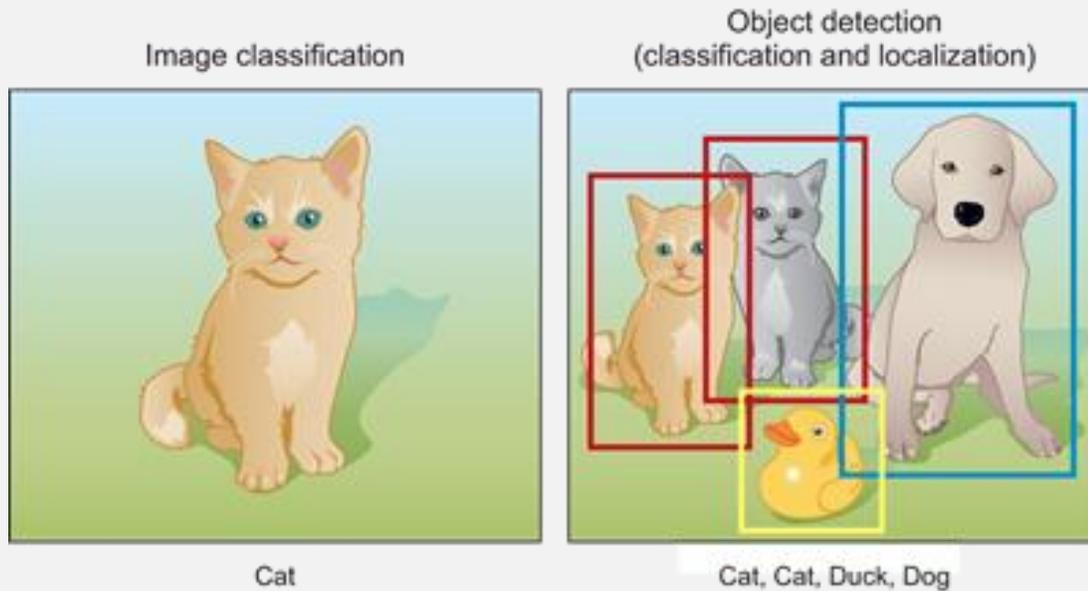
- › 이미지 속 객체를 인지하여 그 클래스를 분류해 내는 기술
 - Object classification

컴퓨터 비전의 목적

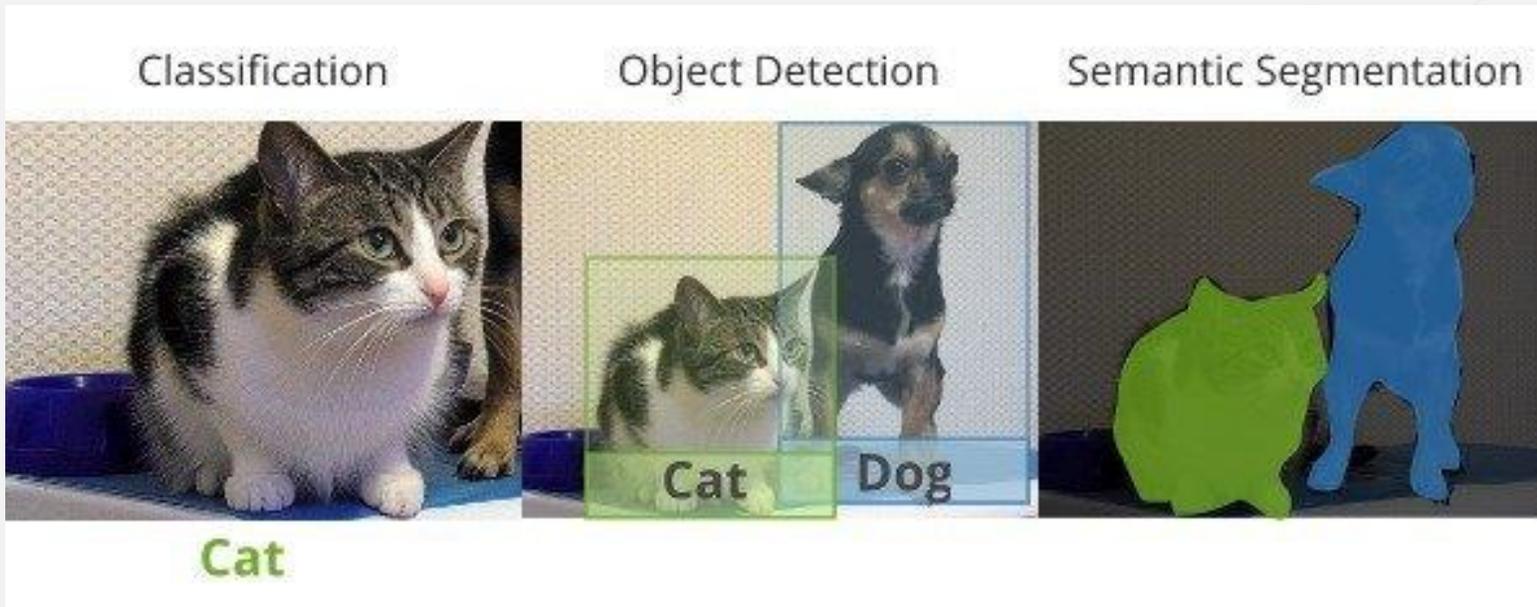
2. 객체 탐지 및 위치 식별(Object Detection and Localization)

이미지 또는 비디오 영상에서 특정 객체를 탐지하고 위치를 식별.

Image에서 다수의 객체(object)를 모두 찾아 bounding box로 표현하는 기술 (Detection) Image에서 발견된 class가 어디쯤 위치하는지 bounding box로 표현하는 기술(Localization)



컴퓨터 비전의 목적



이 이미지는
무엇인가??

어떤 것들이
어디쯤에 있는가?

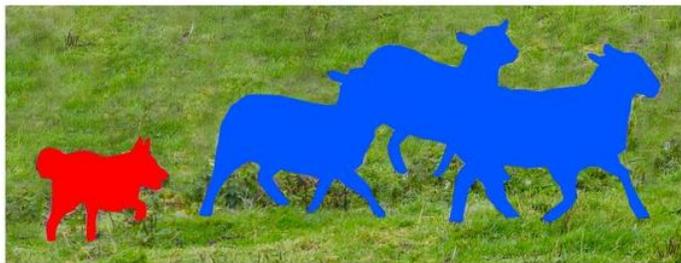
어떤 것들이
어디에 있는가?



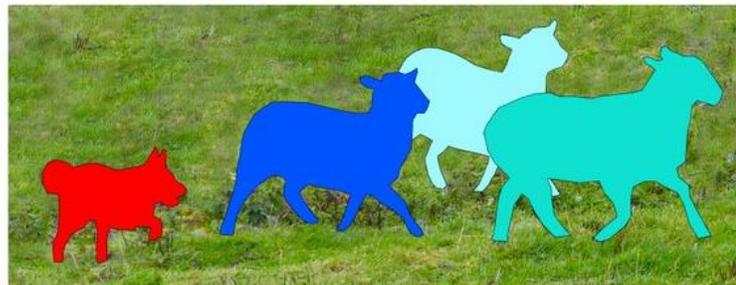
컴퓨터 비전의 목적

3. 분할(Segmentation) 기술은 이미지 혹은 비디오 영상 데이터에서 의미 있는 객체를 따로 추출하여 마스크로 변환

- ▶ **Semantic Segmentation** '의미적 분할' 이미지나 영상의 영역을 물체의 종류에 따라 '분할'하는 작업. 예를 들어 사진에서 이 영역은 도로, 여기는 배경, 여기는 사람..
- ▶ **Instance Segmentation** semantic segmentation과 상당히 유사하지만, 고양이와 두 마리 있을 때 각각 두 마리의 고양이를 구분



Semantic Segmentation



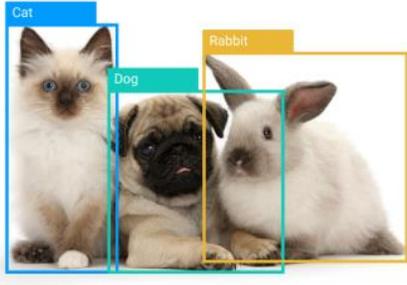
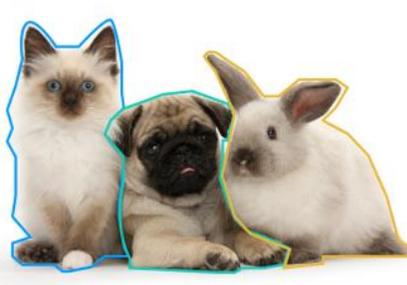
Instance Segmentation

컴퓨터 비전의 목적

OLACT: Real-Time Instance Segmentation [ICCV Trailer]

The following results
were generated in
Real-time

컴퓨터 비전의 목적

Classification	Classification + Localization	Object Detection	Instance Segmentation
			
Cat	Cat	Cat, Dog, Rabbit	Cat, Dog, Rabbit



컴퓨터 비전의 목적

Classification

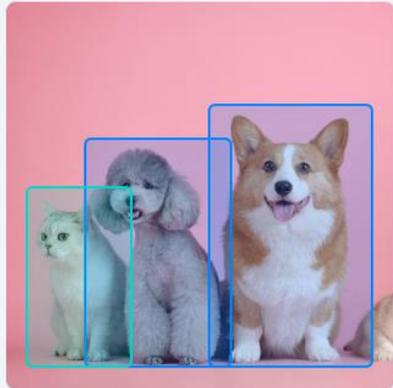


Cat



Single Object

Detection

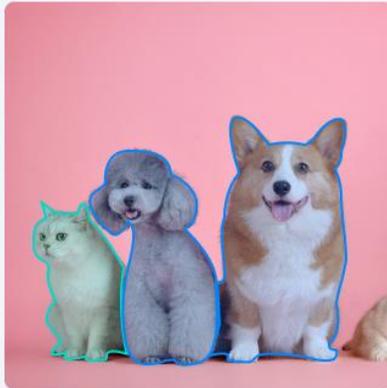


Cat Dog



Multiple Objects

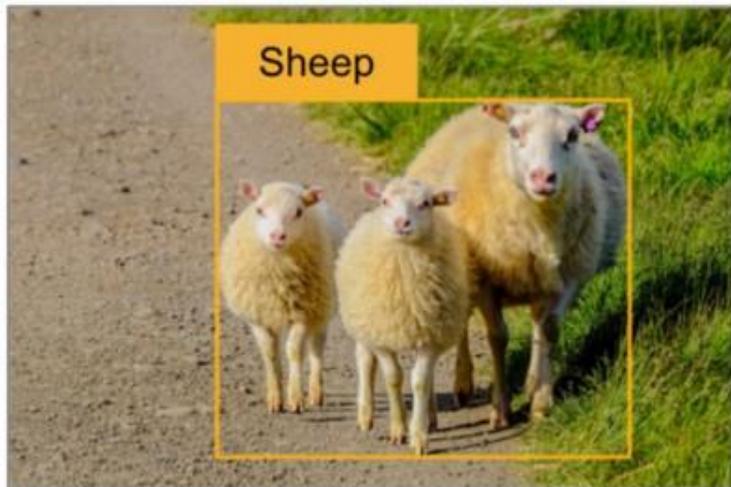
Segmentation



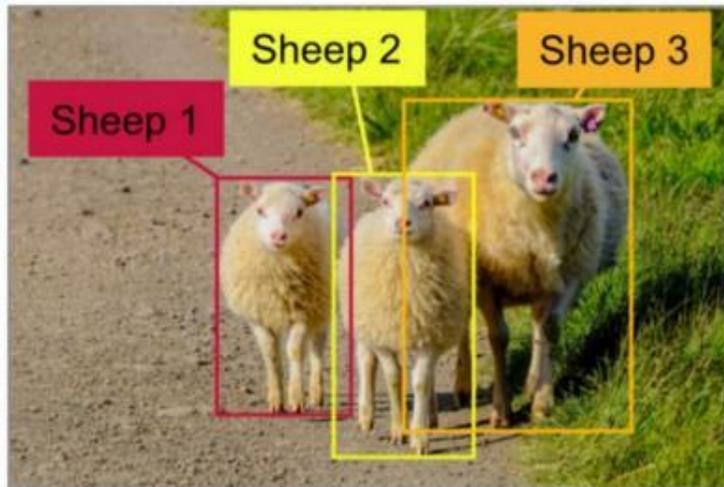
Cat Dog

V7 Labs

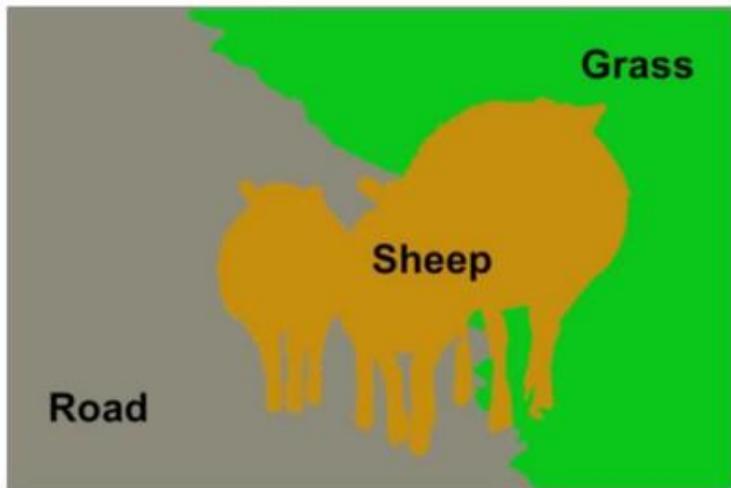
컴퓨터 비전의 목적



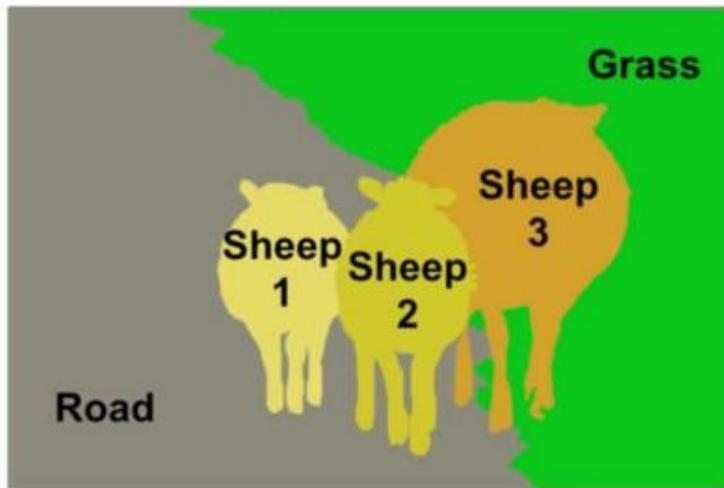
Classification + Localization



Object Detection



Semantic Segmentation



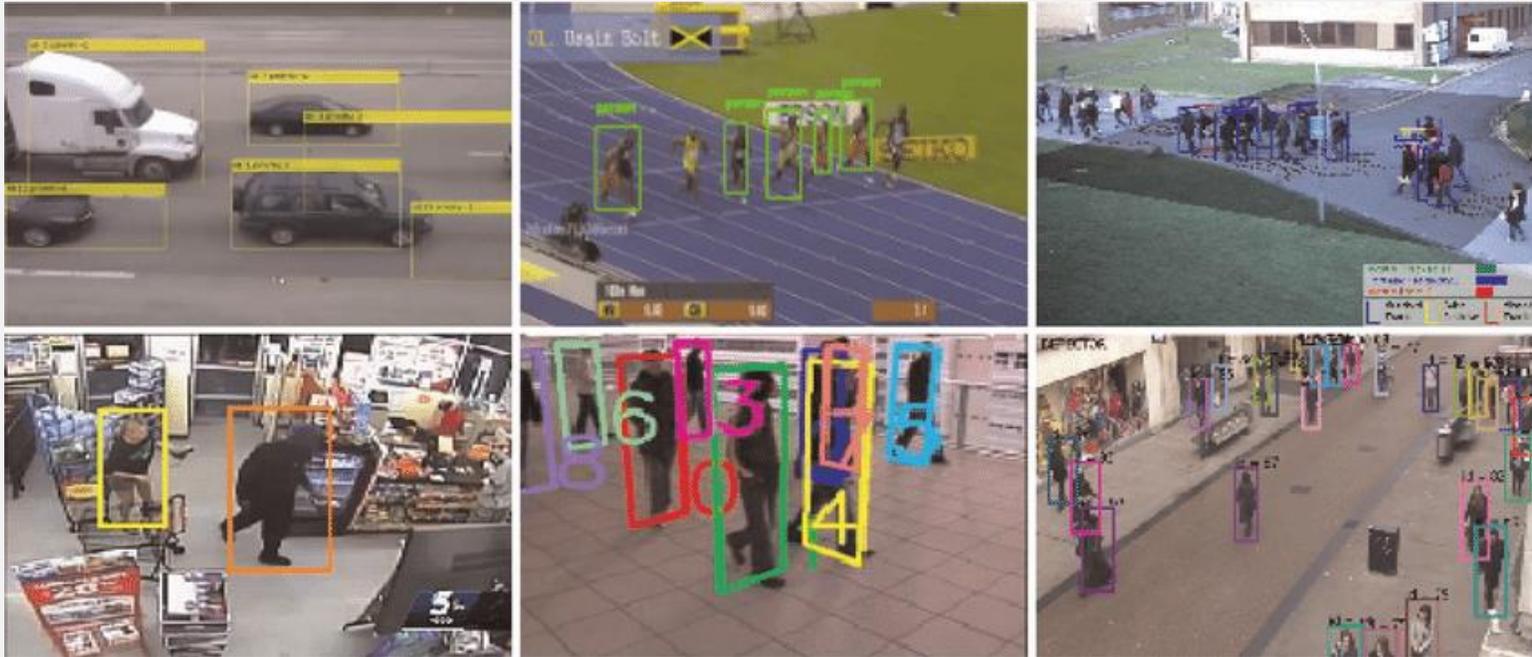
Instance Segmentation

컴퓨터 비전의 목적

4. 객체 추적(Object tracking) 기술은 비디오 영상 내의 객체의 위치와 변화를 추적하는 것으로 포인트 기반에 객체 추적(Point-based object tracking, 실루엣 객체 추적(Silhouette object tracking) 등의 기술 등이 있음



컴퓨터 비전의 목적



<https://becominghuman.ai/what-is-the-main-purpose-of-video-annotation-in-machine-learning-and-ai-11805710bd95>



컴퓨터 비전의 목적

5. 비디오 캡셔닝(Video captioning) 이미지나 영상의 상황을 텍스트로 설명

Video



Caption

“Two boys are playing baseball in the ground”



A female weight lifter bends at the knees.

She lifts a barbell to her chest.

She then lifts it over her head before dropping it heavily to the ground.

