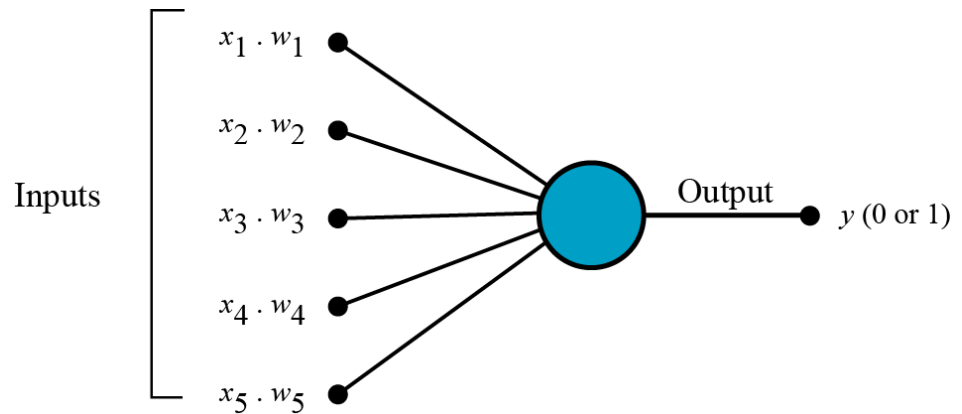


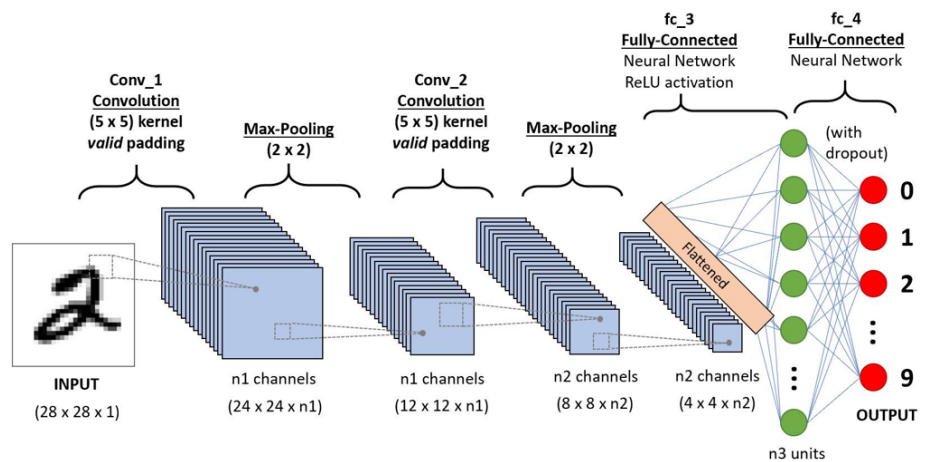
HOI

- 기술적 배경
 - 컴퓨터 비전
 - AI란?
 - AI란 마빈 민스키(Marvin Lee Minsky)가 1956년 다트머스 회의에서 최초로 언급되었다. 최근 많이 언급되는 머신러닝(Machine Learning)을 포함하고 있고 머신러닝은 딥러닝(Deep Learning)을 포함하고 있다.
 - 머신러닝이란?
 - 머신러닝은 기계학습이라고도 불리우며 두가지로 정의를 내릴 수 있다.
 - 첫번째는 아서 사무엘(Arther Lee Samuel)이 내린 정의로 “컴퓨터가 명시적으로 프로그램되지 않고도 학습할 수 있도록 하는 연구분야”이고 두번째로는 톰 미첼(Tom Mitchell)의 “만약 어떤 작업 T에서 경험 E를 통해 성능 측정 방법인 P로 측정했을 때 P가 향상된다면 컴퓨터 프로그램이 스스로 학습한다고 말한다.”라고 정의한다.
 - 머신러닝에는 대표적으로 세가지 분야가 있다. Supervised-Learning과 Unsupervised-Learning, Semi-Supervised Learning 마지막으로 Reinforcement Learning이 있다.
 - Supervised Learning
 - 데이터와 정답이 주어지고 model을 학습시키는 학습 방법
 - Classification, Regression
 - Unsupervised Learning
 - 정답이 없고 데이터만 주어진 상태에서 model을 학습시키는 방법
 - Clustering
 - Semi-Supervised Learning
 - 정답이 일부만 주어진 상태에서 model을 학습시키는 방법
 - Reinforcement Learning
 - 강화학습으로 대표적인 예로 알파고가 있다.
 - 결과값 대신 Reward를 주어 학습한다.
 - 딥러닝이란?
 - 딥러닝은 1957년 프랑크 로젠블라트(Frank Rosenblat)가 Perceptron이라는 알고리즘을 고안하면서 시작되었다.

- Perceptron은 우리 뇌에 존재하는 뉴런(Neuron)을 기반으로 고안한 알고리즘이다.
- Perceptron은 여러 입력을 받아 하나의 신호를 출력한다.



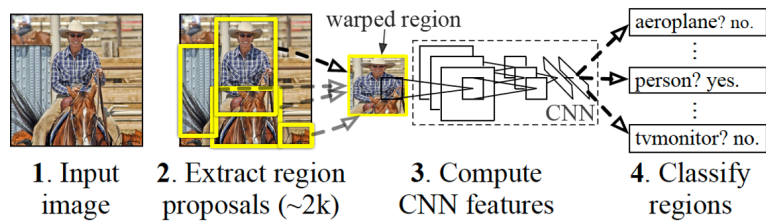
- 위와 같은 Perceptron을 여러 층으로 쌓아올려 Classification과 Regression 등 여러가지 task를 처리할 수 있는 model을 생성할 수 있다.
- 컴퓨터 비전이란
 - 신경망의 여러 구조 중에서 합성곱 신경망(Convolution Neural Network)를 사용한 CNN구조를 이용하여 이미지를 인식할 수 있는 딥러닝 기법



A CNN sequence to classify handwritten digits

- 컴퓨터 비전에는 이미지 생성, 이미지 스타일 변환, 분할 등 많이 활용되고 있고 이미지에서 사물을 검출할수도 있다.
 - 사물 검출은 이미지 속에 담긴 사물의 위치와 종류를 알아내는 기술
 - R-CNN(Regions with Convolutional Neural Network)구조를 많이 사용한다.

R-CNN: Regions with CNN features

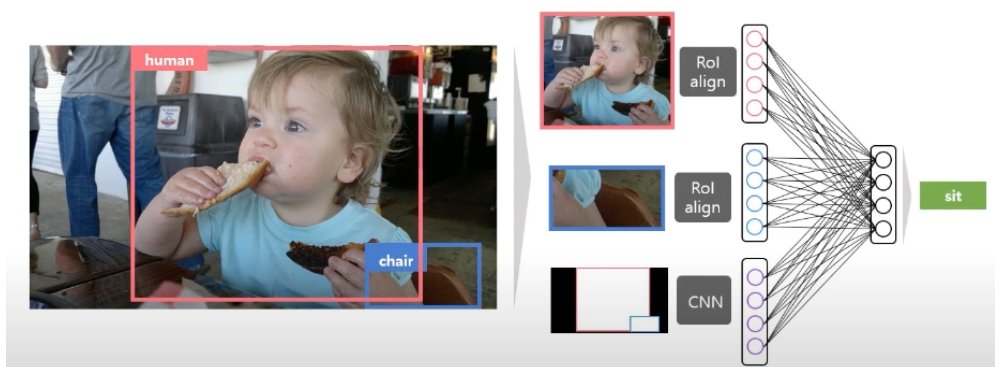


1. 이미지 입력
2. 후보 영역 추출
 - 이미지를 사각형으로 분류할 때 SVM(Support Vector Machine)을 주로 사용한다.
3. CNN 특징 계산
4. 영역 분류
 - 위와 같은 순서로 진행된다.

◦ HOI(Human Object Interaction detection)

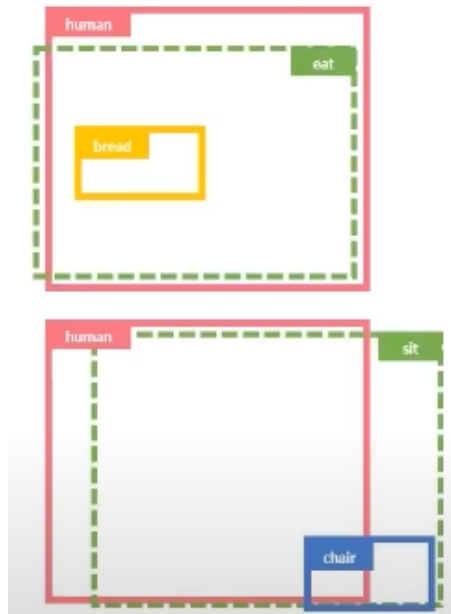
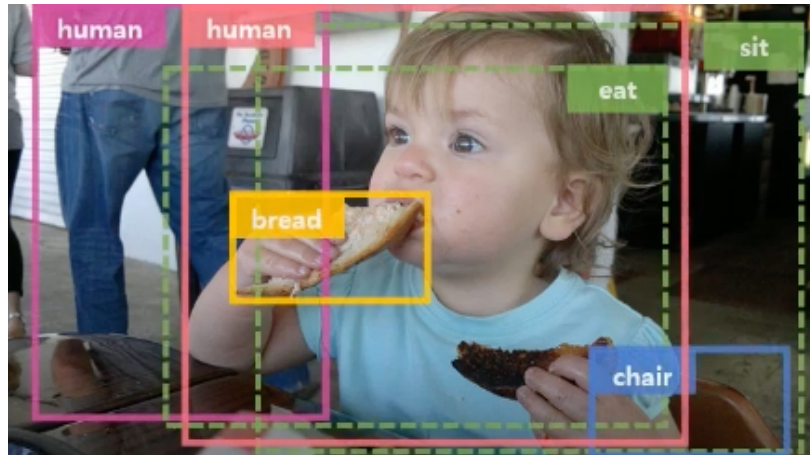
- 컴퓨터 비전이 발전함에 따라 우리 삶에 더 가까이 다가오고 있다.
- HOI이란?
 - 컴퓨터 비전을 활용하여 사람과 사물의 상호작용을 식별할 수 있는 기술입니다.
 - 초기에는 Sequential HOI Detectors를 사용하여 각각 사물과 사람을 인식하는 network와 사물과 사람의 상호작용을 확인하는 network를 따로 설계하여 학습하였다.

◦ Sequential HOI Detectors

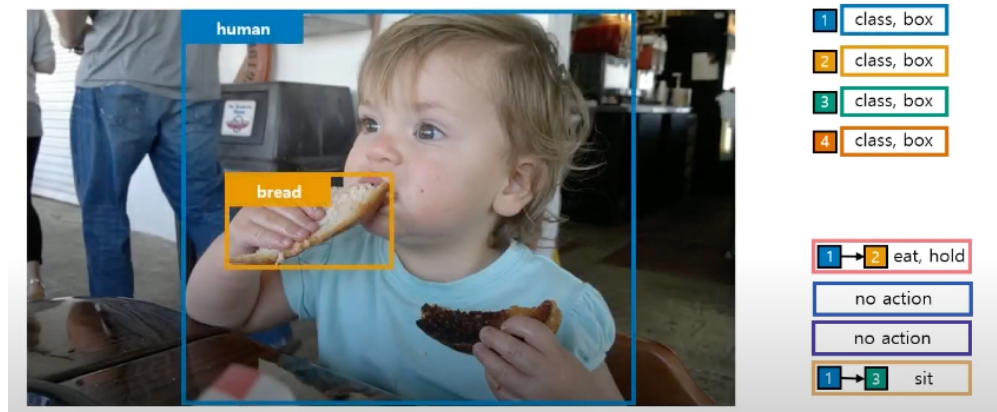
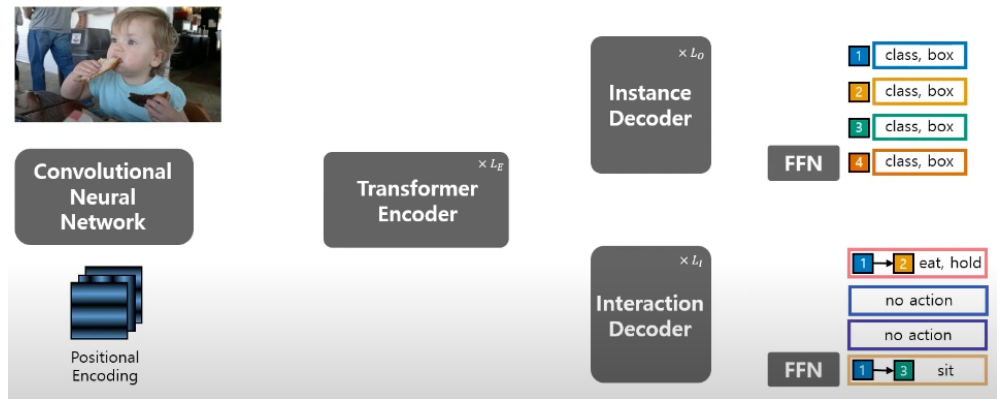


- 이후에 Parallel HOI Detectors로 발전하였고 Region of Interaction을 설정하여 해당 영역에 포함되는 사물과 사람의 상호작용을 확인할 수 있다. Sequential에 비해서 속도가 매우 향상되었다.

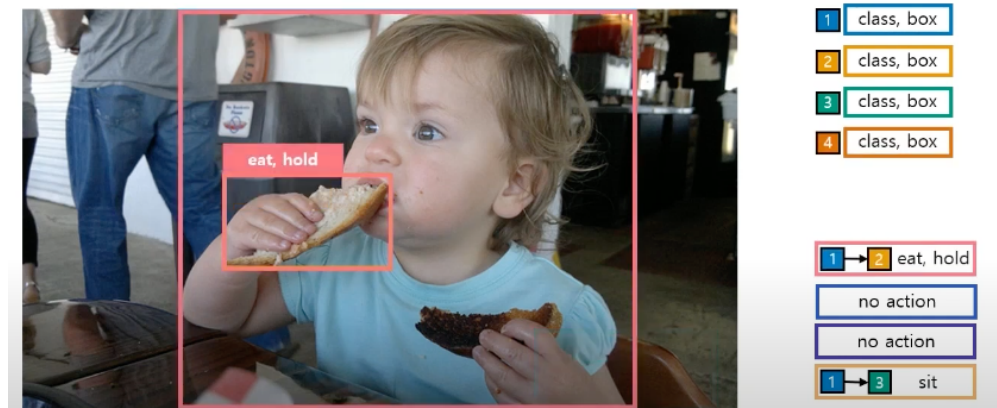
하지만 병목현상이 발생하여 알고리즘의 발전이 여전히 필요한 상황



- 최근 고려대학교 컴퓨터학과 김현우 교수님의 연구실에서 발표한 HOCR(End to End Human Object Interaction Detection with Transformers)에서
- CNN을 통과한 이미지를 4등분하여 Transformer Encoder를 이용하여 encoding 후 instance decoder와 interaction decoder를 각각 따로 통과하여



instance decoder를 통과한 결과는 각각의 사물과 사람의 객체를 인식하고
interaction decoder를 통과한 결과는 instance decoder를 통과한 결과간의
interaction을 보여준다.



<https://www.youtube.com/watch?v=hmfRkkmt0WE>

- 대표적인 예로는 아마존고(Amazon-Go)가 있는데 무인 매장에서 사람과 물건의 상호작용을 CCTV를 통해 분석하고 사람이 물건을 고르면 자동으로 인터넷 장바구니에 물건이 들어가게 되고 사람이 매장을 나서면 등록된 카드로 장바구니에 담은 물건이 결제되는 매장이다.

- HOI라는 기술 자체도 개발되지 얼마되지 않은 기술이기 때문에 현재도 관련된 논문이 많이 publish되고 있는 상황이고, 최근 국내에도 아마존고와 같은 매장을 여러 기업에서 상용화하려고 노력하고 있다. 따라서 주목받고 있는 기술중에 하나라고 볼 수 있다.

•

◦ 홈 CCTV

■ cctv란

- Closed-circuit Television의 줄임말로 반대말로는 흔히 알고 있는 TV가 있다. CCTV는 다수에게 보여주는 TV가 아닌 특정인들에게 제공되는 TV로 목적에 따라 다양하게 사용될 수 있다.
- 대표적으로 변화가, 범죄 다발 지역 등 방범을 위해 사용되는 경우가 많은데 IOT 기술이 발달함에 따라서 CCTV가 우리 생활에 직접적으로 다가왔고 집안으로까지 들어올수 있게 되었다.
- CCTV는 크게 아날로그와 디지털 방식으로 구분된다.

산업용에서는 주로 아날로그 방식을 사용했지만 최근에는 주로 디지털 방식을 사용한다. 디지털은 네트워크로 연결되기 때문에 'IP 카메라'라고 부르기도 한다.

- 아날로그 카메라는 메인 시스템인 DVR과 CCTV카메라가 1:1로 연결되어 영상 케이블과 전원선, 제어선 등 복잡한 설치가 필요하여 설치비용이 높고 카메라를 추가하는 것도 제한적이다.



아날로그 방식의 CCTV

- 디지털 카메라는 각각의 카메라와 메인 시스템인 NVR과 네트워크로 연결되기 때문에 네트워크가 존재하는 환경이라면 카메라만 설치한다면 설치 및 업그레이드가 간편하다.



http://it.chosun.com/site/data/html_dir/2015/04/03/2015040385008.html

- 홈 CCTV는 일반적인 IP 카메라와 거의 동일하게 동작한다. 네트워크만 연결되어 있다면 카메라에 담긴 영상을 어디에서든 접근하여 볼 수 있고, 여러 카메라가 존재한다면 영상을 동시에 녹화, 관리가 가능하다.

- 시스템 구성

- 용도 및 대상에 따른 분류

- CCTV의 사용 목적에 따라 cctv가 학습해야 할 상황 혹은 사람, 사물이 다양해 질 수 있지만 사람이 사용한다는 큰 범주를 벗어나지 않기 때문에 크게 가정과 사업체로 구분하여 더욱 상세하게 학습을 진행한다.
 - 가정 - 사람(어린이 및 독거노인)/반려동물
 - 최근 보호자가 없는 상황에서 반려동물 혹은 어린이가 가정에 남아있을 때 홈 cctv를 많이 사용한다. 이 때는 사람뿐만이 아니라 반려동물에게도 위험한 상황이 처해질 수 있으므로 동물과 사물의 상호작용 추가적인 학습이 필요하다.
 - 또한 독거노인의 경우 고독사가 증가하고 있기 때문에 가정에 움직임이 없을 경우 알림서비스 혹은 병원과의 연락도 추가적으로 필요할 수 있다.
 - 사업체 - 식당 및 가게/영유아 위주의 여가시설
 - 사업체의 경우 가정에 비해서 사고보단 사건이 발생할 확률이 높다. 따라서 사업체에서 발생할 수 있는 범죄를 좀 더 상세히 학습을 진행한다.

- 홈 CCTV 핵심 기능 개요

- HOI기술을 활용한 홈 cctv의 핵심 기능
 - 현재 홈 CCTV는 가정은 물론 개인 사업체에서도 모든 내용을 녹화하거나 실시간으로 화면을 보고 있지 않으면 사건 사고를 미연에 방지할 수 없고, 사건사고 발생 후 시간을 특정하여 다시 확인해야 한다.
 - cctv에 HOI기술이 적용된다면 cctv가 사람과 사물의 상호작용을 학습하여 사람이 위험상황에 놓여있거나, 사건사고가 발생한다고 판단하면 사용자에게 알림서비스를 제공하고 녹화를 시작한다.

- 어플 알림서비스

- 위험을 감지했을 때 자동적으로 어플 알림 및 녹화진행
 - 사람과 사물의 상호작용을 파악하고 특정 위험한 상황에 놓여있다고 판단이 되면 사용자의 모바일에 설치되어있는 홈 cctv의 어플을 통해 알림 서비스를 송신한다.
 - 현재 사용되고있는 사업체의 cctv는 사고가 발생하면 시간을 특정하여 모든 녹화 내용을 확인해야 하지만, cctv가 사람과 사물의 상호작용을 이해하고 위험하거나 사고가 발생한다고 판단하면 자동으로 녹화가 시작된다.

- 사업 아이템 알고리즘

- 객체인식 알고리즘 구현
 - HOI기술의 첫번째 단계 파이썬을 이용한 구현
- Human Pose Estimation
 - 특정 자세 인식
 - 자료조사
 - 알고리즘 구상
- 모바일 연계 (IOT)
 - 어플에 추가기능