

레드빈즈 딥러닝 스터디

모두를 위한 얇고 넓은 딥러닝

with



오 준 오

Machine Learning & Pattern Analysis laboratory

2nd October 2019

Syllabus

첫 세 스터디

첫 번째 스터디

딥러닝이란 무엇인가 / 설치부터 데이터 다루기

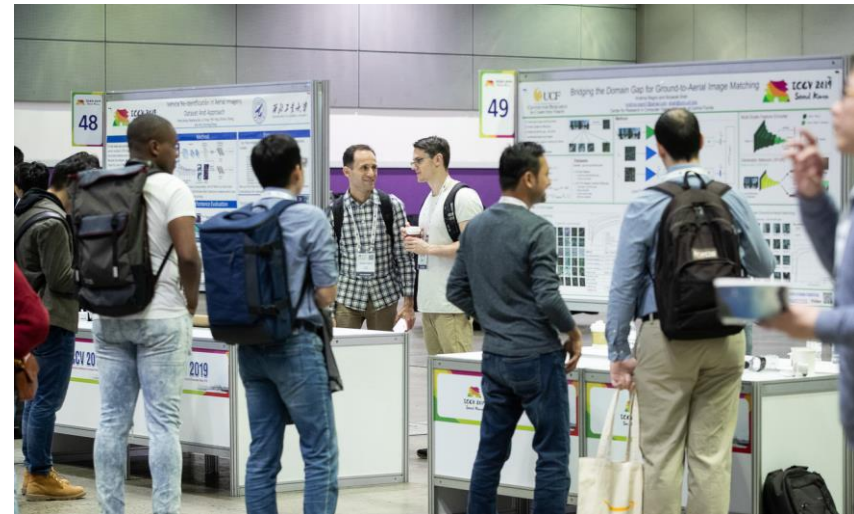
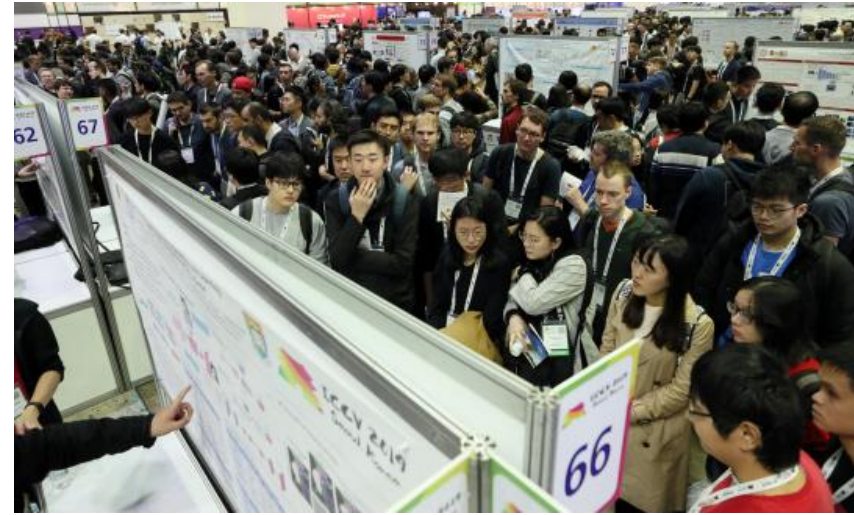
두 번째 스터디

왜 딥러닝이어야 하는가 / 모델과 학습

세 번째 스터디

컴퓨터의 눈으로 보기 / 이미지 처리하기

ICCV, International Conf. on Computer Vision



AI의 현실과 미래



AI는 정말 밝은 미래를 약속하는가?

<https://www.youtube.com/watch?v=mhN4n0IhW0>

성장동력 확보			
인공 지능 (AI)	세계 최고 수준의 AI 기술 확보 2017년: 언어-시각 인식 기술 2019년: 언어-시각 통합지능 AI 2022년: 추론/의사 결정 AI	기술수준	70% → 100% 현재 2025
가상/증강 현실 (VR/AR)	핵심 기술확보 및 글로벌 전문 기업 육성 2018년: 디지털교과서, 평창올림픽 서비스 2019년: 휴먼팩터 기술확보 2020년: 민간주도 콘텐츠 융합서비스	국내시장 규모	1.3억불 → 40억불 현재 2025
자율 주행차	자율차 (level 3-4) 핵심부품, 실증기술 확보 2019년: 8대 핵심 부품 2021년: 자율주행차 융합기술 2024년: 자율주행차 융합서비스 실증	핵심부품 기술수준	77.5% → 100% 현재 2025
경량 소재	경량소재 (Ti, Mg, Al, 탄소섬유) 핵심기술 확보 2019년: 차체용 경량소재 (Mg, Al) 합금기술 2020년: 항공용 소재 (Ti) 원천기술 2023년: 경량소재 (Ti, Mg, Al, 탄소섬유) 양산기술	해외수출 (경량소재)	15억불 → 310억불 현재 2025
스마트 시티	스마트 시티 실증기술 구현, 해외 진출 확대 2018년: 인프라 연계 시스템 설계 2019년: 개방형 운영체제 구축 2020년: 국내 실증도시 2개 구축 2021년: 지능형 의사결정 시스템 개발 및 해외진출 확대	해외수출 비중	10% → 30% 현재 2025

실습을 시작하기 전에

개발 환경 구성하기



Anaconda Python Distribution

+



JetBrains PyCharm IDE

실습을 시작하기 전에

Anaconda Python Distribution

<https://docs.conda.io/en/latest/miniconda.html>



패키지 관리와 디플로이를 단순하게 할 목적으로 만들어진 파이썬 배포판

과학 계산(데이터 과학, 기계학습, 대규모 데이터 처리, 예측 분석 등)에 특화

Windows, Linux, macOS 용 1,400개 이상의 유명 데이터 과학 패키지가 포함

Venv(Virtual Environment) 기능을 손쉽게 관리할 수 있음

Virtual Environment 기능은 JVM(Java Virtual Machine)과 유사

실습을 시작하기 전에

JetBrains PyCharm IDE

<https://www.jetbrains.com/pycharm/>

JetBrains에서 제작한 Python 용 통합개발환경(IDE)

IntelliJ IDEA에 기반을 두고 있으며 현용 Python 개발 툴 중에 가장 완성도 높은 제품

코드 자동 완성 기능과 디버깅 기능을 지원함 - 매우 중요!

VirtualEnv는 네이티브로, Anaconda는 전용 plug-in으로 적극 지원



실습을 시작하기 전에

PyTorch

<https://pytorch.org>



Python 기반의 오픈 소스 딥러닝 라이브러리, 페이스북 AI 연구팀에 의해 개발

간결하고 빠른 구현, Tensorflow에 비해 훨씬 익히기 쉬움

Tensorflow에 비해 높은 추상화 레벨을 가지고 있기 때문

패러다임 또한 직관적으로 이해하기 쉬운 Define-by-Run 구조

기능들이 잘 모듈화되어 있고 이름 또한 직관적이라 찾고 이해하기 쉬움

딤러닝, 시작



사람은 쉽게 이 사진이 고양이임을 안다.

그러나 컴퓨터에게 물어보면?

딤러닝, 시작

고양이 검출

‘고양이’라는 개체

선, 면, 형상, 색, 크기 등 다양한 요소의 결합



크기 30cm 이하는 고양이, 이상은 고양이 아님?

갈색은 고양이, 빨간색은 고양이 아님?

단순한 선형 구분으로 식별 불가

개체를 구분할 수 있는 **특징**을 찾아야 한다!

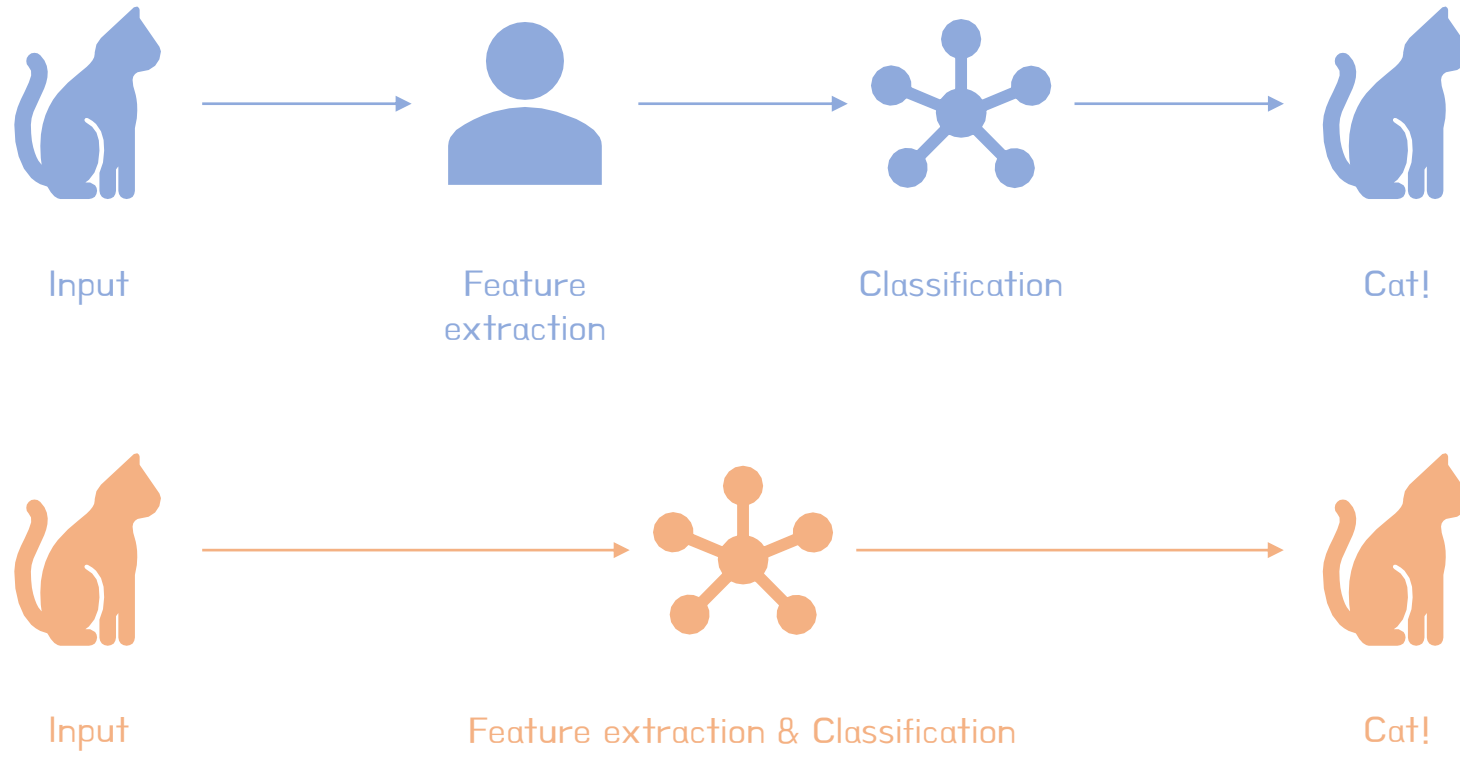
딥러닝, 시작

판별을 위한 특징을 추출하는 과정을 컴퓨터에게 시킬 순 없을까?

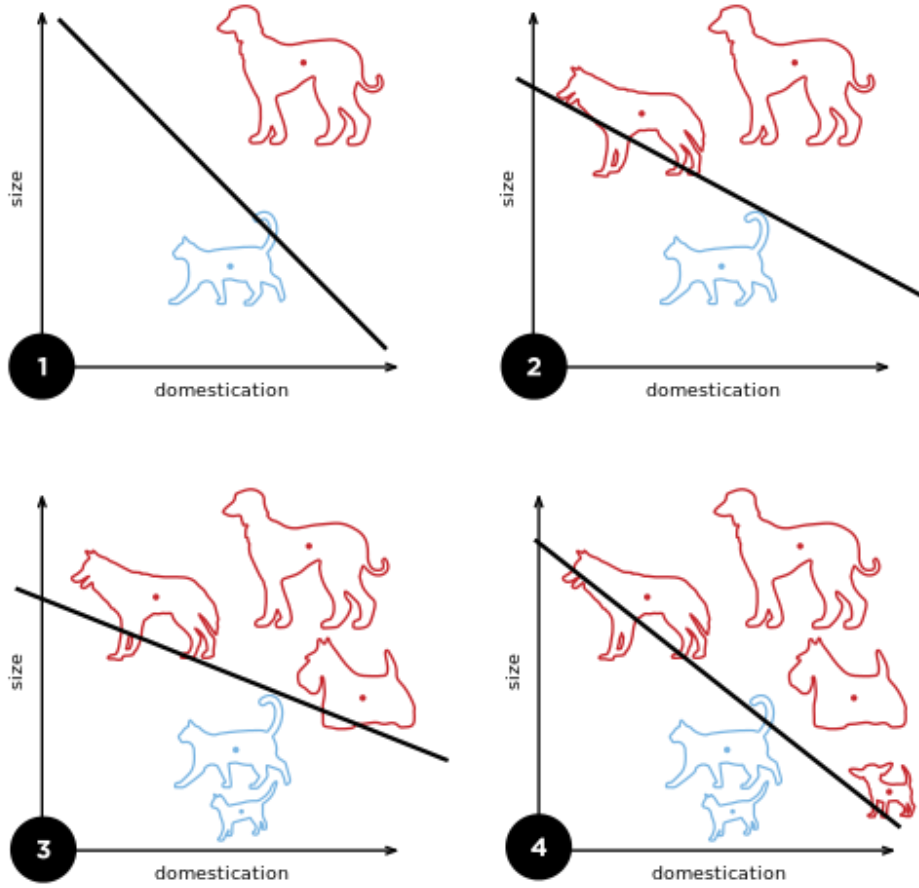


딥러닝, 시작

머신러닝 vs 딥러닝



딥러닝, 무엇인가요?



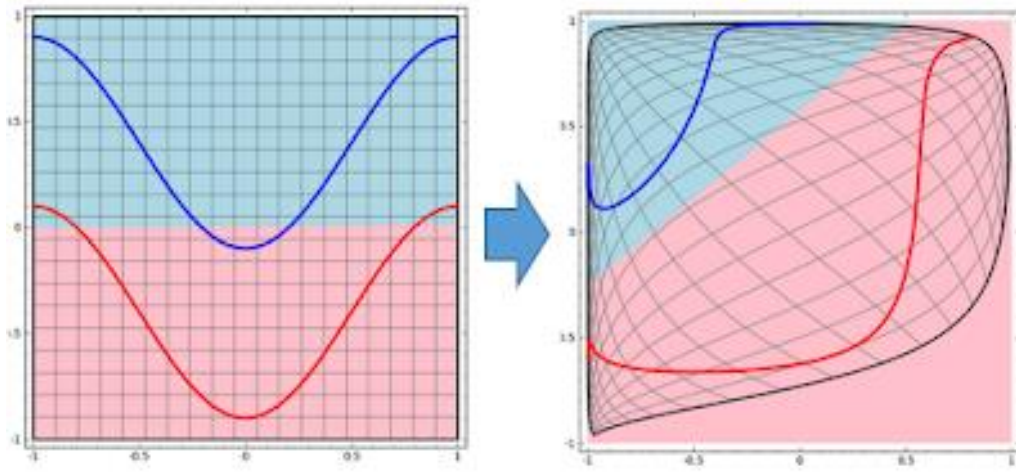
좌표 위에 점(데이터)을 구분하기 위해

모델은 직선을 그림

점이 추가될 때 마다 이를 반영하여 직선을 수정

이 과정이 사람의 학습 과정과 유사하여 학습이라 함

딥러닝, 무엇인가요?



딥러닝은 데이터를 구분하기 위해 다음과 같은 과정을 거침

데이터를 잘 구분하는 선을 **긋고**

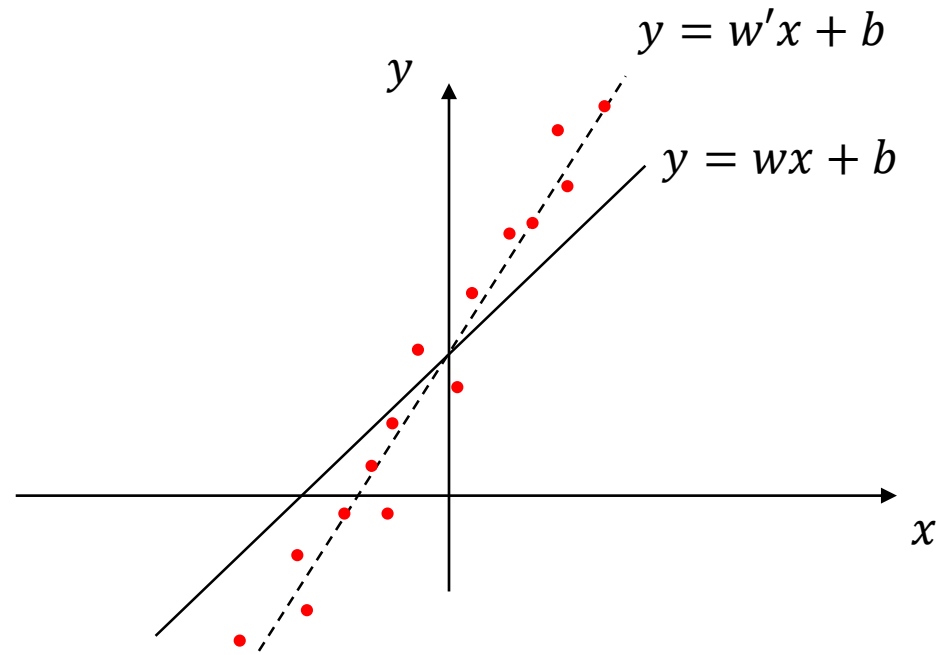
그 공간을 잘 **왜곡**한 뒤

다시 선을 **긋고**, **구기고**, **긋고**, **구기고**... 반복

이를 통해 복잡한 공간 속에서 최적의 구분선을 구함

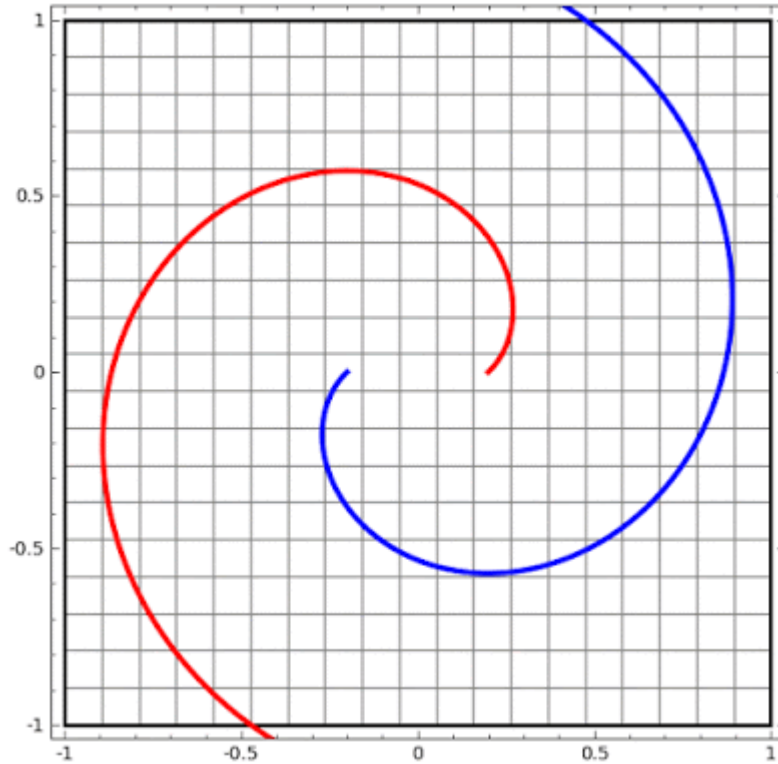
딥러닝, 무엇인가요?

그럼 어떻게 학습을 시키나요?



x 는 입력, y 는 출력, w 는?

딥러닝, 무엇인가요?



인공신경망은 입력을 받아

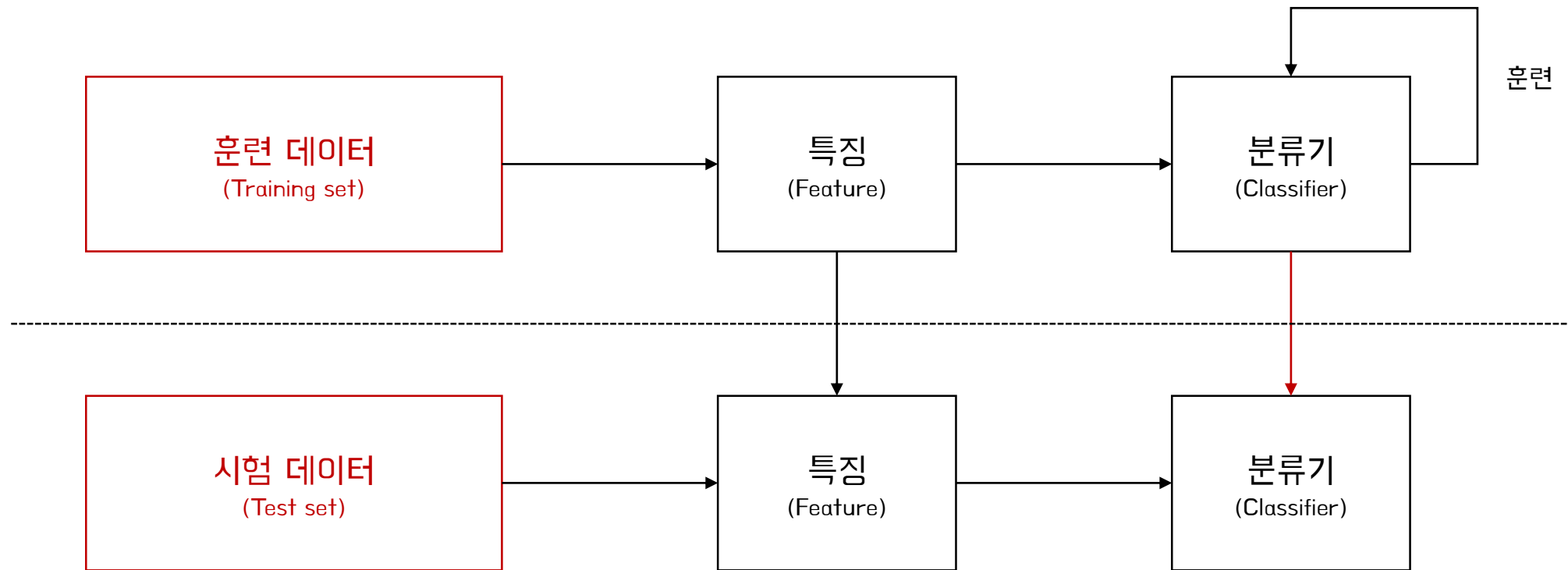
$$y = wx + b$$

위의 수식을 계산하고(i), 여기에 활성화 함수를 씌워

비선형적으로 공간을 변화시킴(ii)

복습

훈련 과정



시험 과정

실습

MNIST Dataset을 이용한 학습 과정 살펴보기

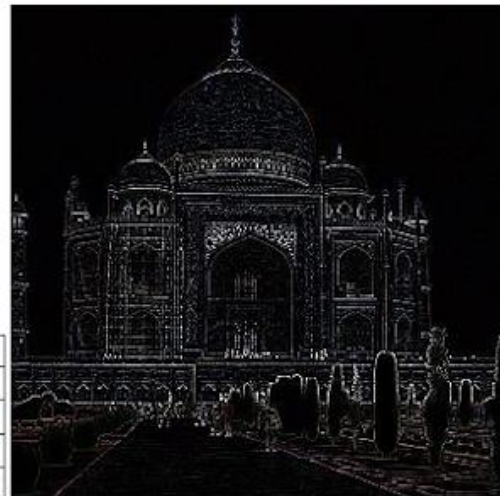
CIFAR-10 Dataset에서는 왜 성능이 안 나올까?

Convolutional Neural Network(CNN)

0	0	0	0	0
0	1	1	1	0
0	1	1	1	0
0	1	1	1	0
0	0	0	0	0



	0	1	0	
	1	-4	1	
	0	1	0	



CNN은 (i)번 과정이 다름

CNN의 Convolution filter는 이미지를 변화시킴

목적하는 작업의 성공률을 높이도록 이미지를 변형

Convolutional Neural Network(CNN)

CNN 시각화

1 _{x1}	1 _{x0}	1 _{x1}	0	0
0 _{x0}	1 _{x1}	1 _{x0}	1	0
0 _{x1}	0 _{x0}	1 _{x1}	1	1
0	0	1	1	0
0	1	1	0	0

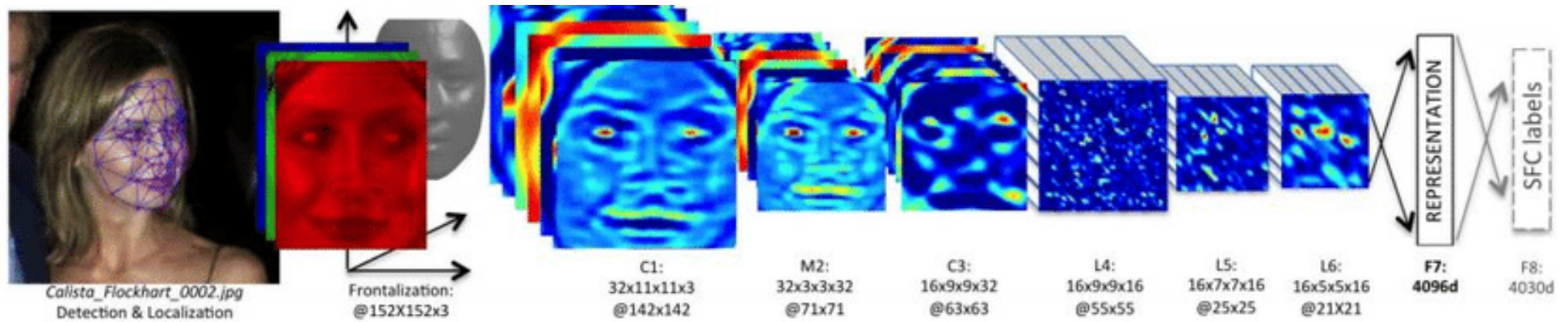
Image

4		

Convolved
Feature

Convolutional Neural Network(CNN)

CNN 시각화



수고하셨습니다
감사합니다