

인공지능의 꽃, 딥러닝(Deep Learning)

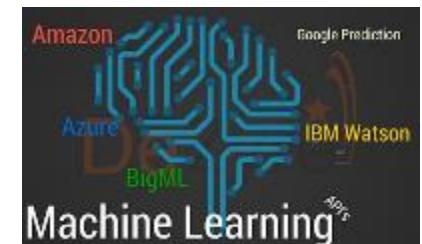
K-ICT 빅데이터센터 세미나

2016.07.26 (화)

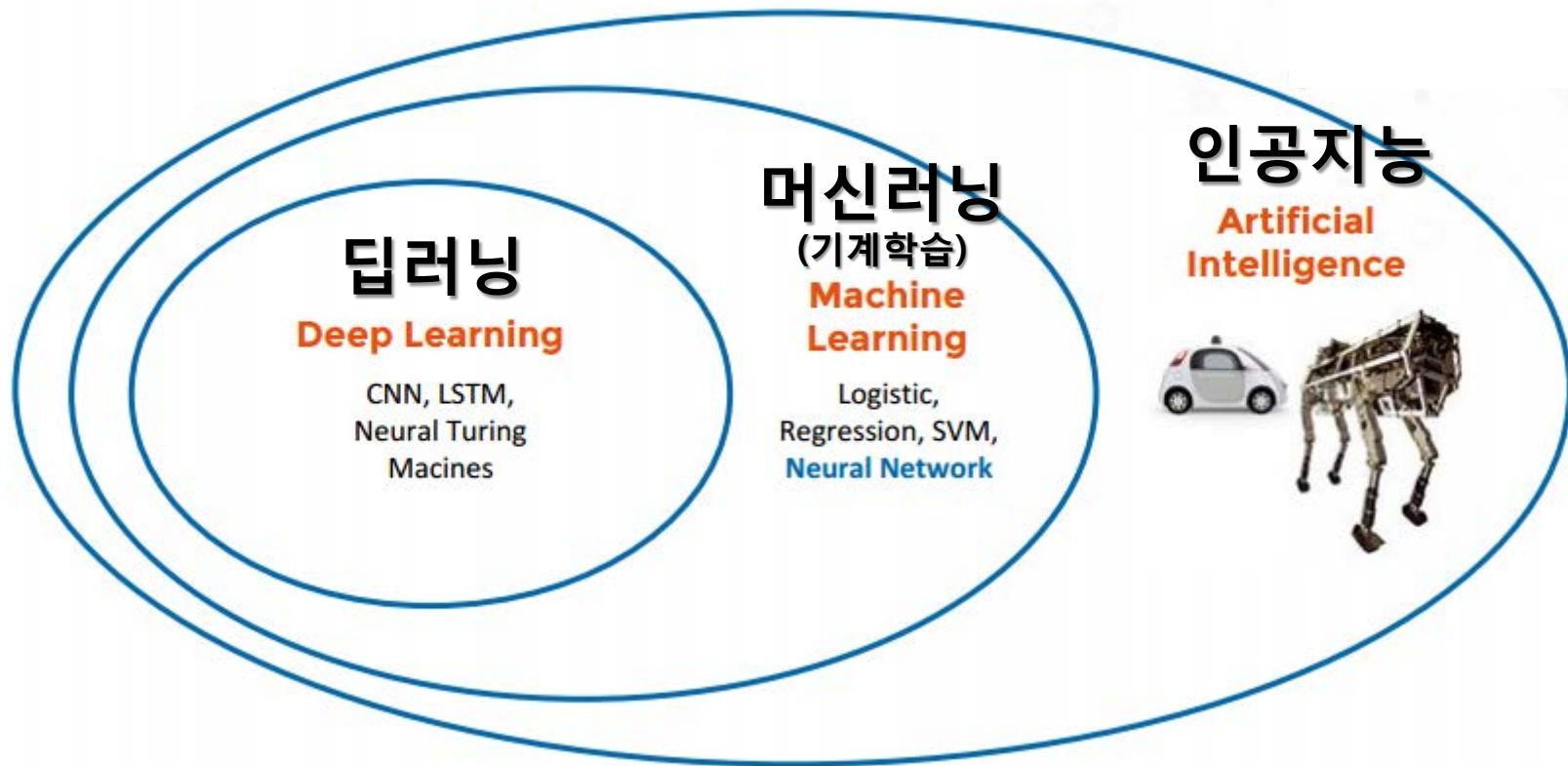
서울대학교 바이오지능연구실

김병희

bhkim@bi.snu.ac.kr



Intro



Slide by Jiqiong Qiu at DevFest 2016

<http://www.slideshare.net/SfeirGroup/first-step-deep-learning-by-jiqiong-qiu-devfest-2016>

인공지능(AI)

- “사람처럼 생각하고 사람처럼 행동하는 기계”(컴퓨터, SW)를 개발
- 사람이 기계보다 잘 하는 일을 기계가 할 수 있도록 하는 연구
- 지능을 필요로 하는 일을 기계가 할 수 있도록 하는 연구
- 1950: Turing's Paper, 1956: “Artificial Intelligence (AI)”



인공지능 구현의 전략: 학습

- AI 연구의 목표는 ‘똑똑한’ 컴퓨터/기계를 만드는 것
- 컴퓨터의 동작이 지적이고(intelligent), 상황에 따라 적절히 변하며(adaptive), 안정적(robust)이길 원함



cat

car

- 사람이 일일이 프로그램을 작성하는 것은 불가능
- 해법? 컴퓨터에게 우리가 원하는 동작의 예를 보여주고, 스스로 프로그램을 작성하도록 하자!
- 이것이 AI 구현 전략 중 ‘학습(learning)’ 기법

어떻게 똑똑한 컴퓨터 시스템을 만들 것인가?

세상(world)을 인식하고 이해할 필요가 있다

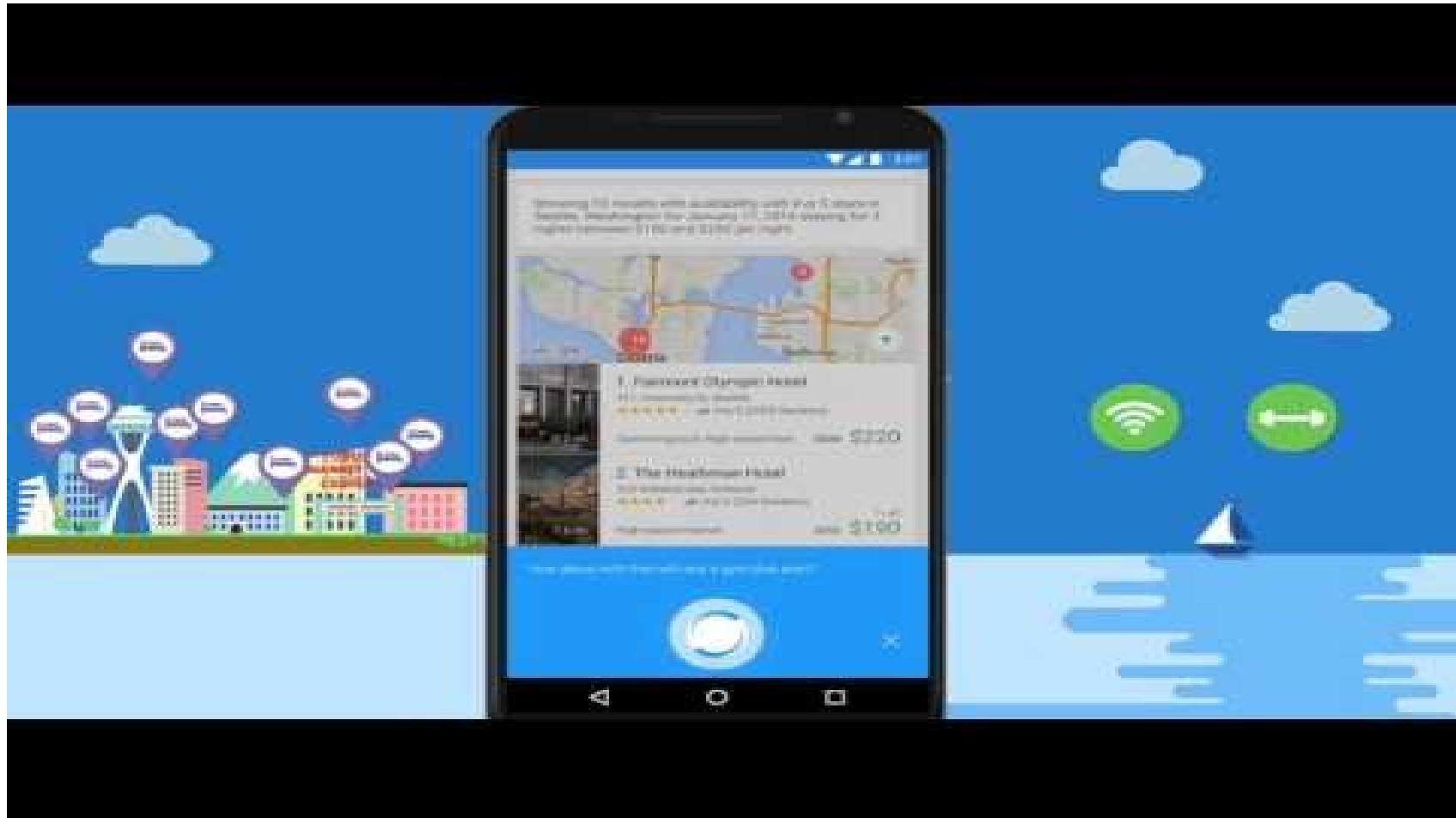
기본적인 음성 및 시각 능력

언어 이해

행동 예측

...

음성 언어를 인식하고 이해하고 대응하는 S/W

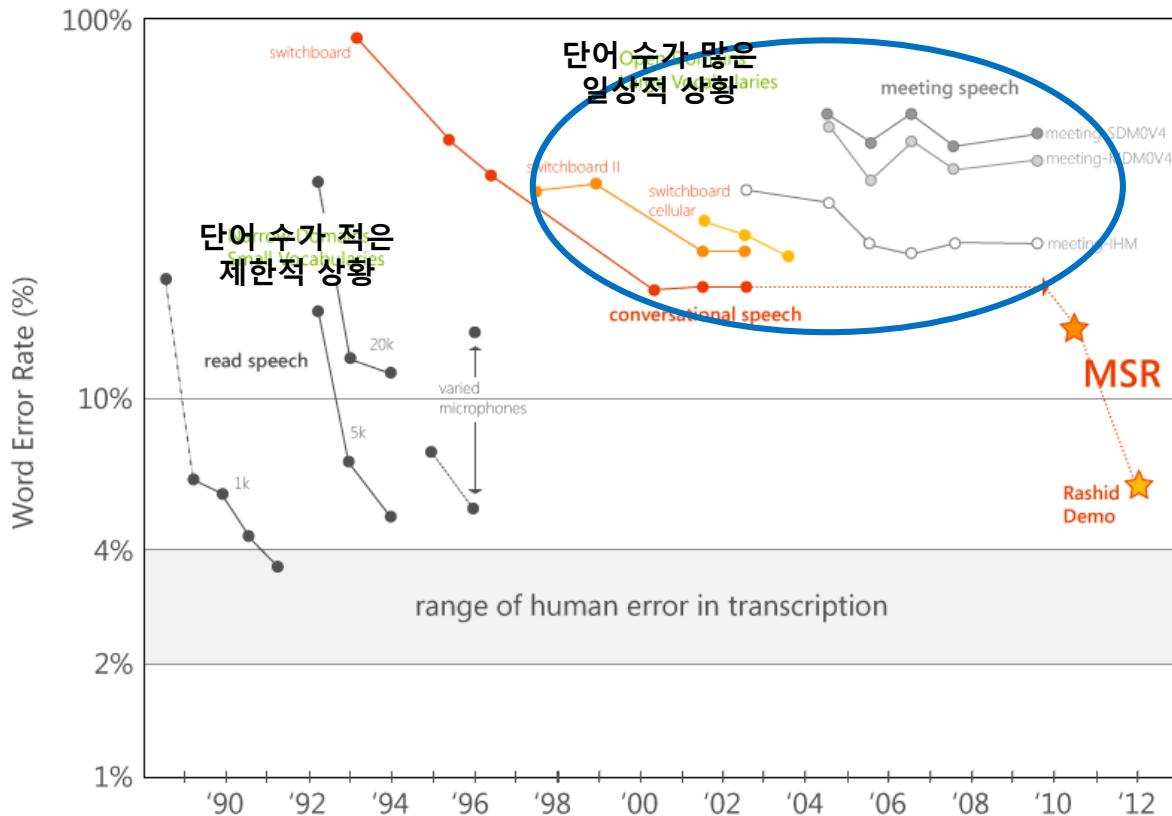


딥러닝이 있기에 가능하다

'Hound' by Soundhound

딥러닝, 음성 인식의 수준을 한 단계 올리다

NIST Evaluations of Automatic Speech Recognition



(Figure from ICML 2014 tutorial by Li Deng)

- 전화통화음성 자동 인식 문제(switchboard)에서 99년 이후 10여 년 동안 성능 정체
- Microsoft Research: 2010년 딥러닝 기술로 단어 인식 오류율을 ~23%에서 <15%로 줄임(Rick Rashid's의 [시연](#)에서는 7% 이하)!
- Google: 2012년부터 딥러닝 적용. 2015년 RNN 기반으로 성능 대폭 향상
- IBM Watson 팀은 [2015년](#) switchboard II에 대해 오류율 8%를 달성

[Geoffrey Hinton, Li Deng, Hong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Oriol Vinyals, Patrick Nguyen, Taek-Hwan Kim, and Brian Kingsbury]

Deep Neural Networks
for Acoustic Modeling
in Speech Recognition

[The shared views of four research groups]

IEEE Signal Processing
Magazine, Nov. 2012



Intelligent Personal Assistant Program

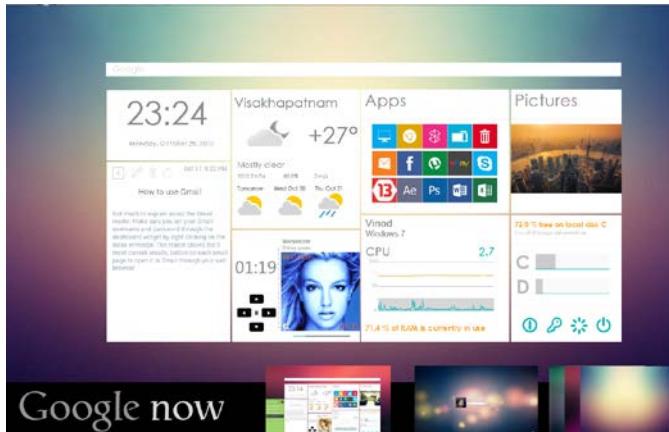


Siri

Use your voice to send messages, set reminders, search for information, and more.



<https://www.yahoo.com/tech/inside-look-microsofts-cortana-163814879.html> (2:20)



Google now



24me

quip

Personal Assistant/Robot at Home



Google Home



Pepper



Amazon Echo “Alexa”



Jibo

자율주행 자동차 (Autonomous Car)

Grand Challenge
2005



Urban Challenge
2007



Self-driving Cars
2009 ~

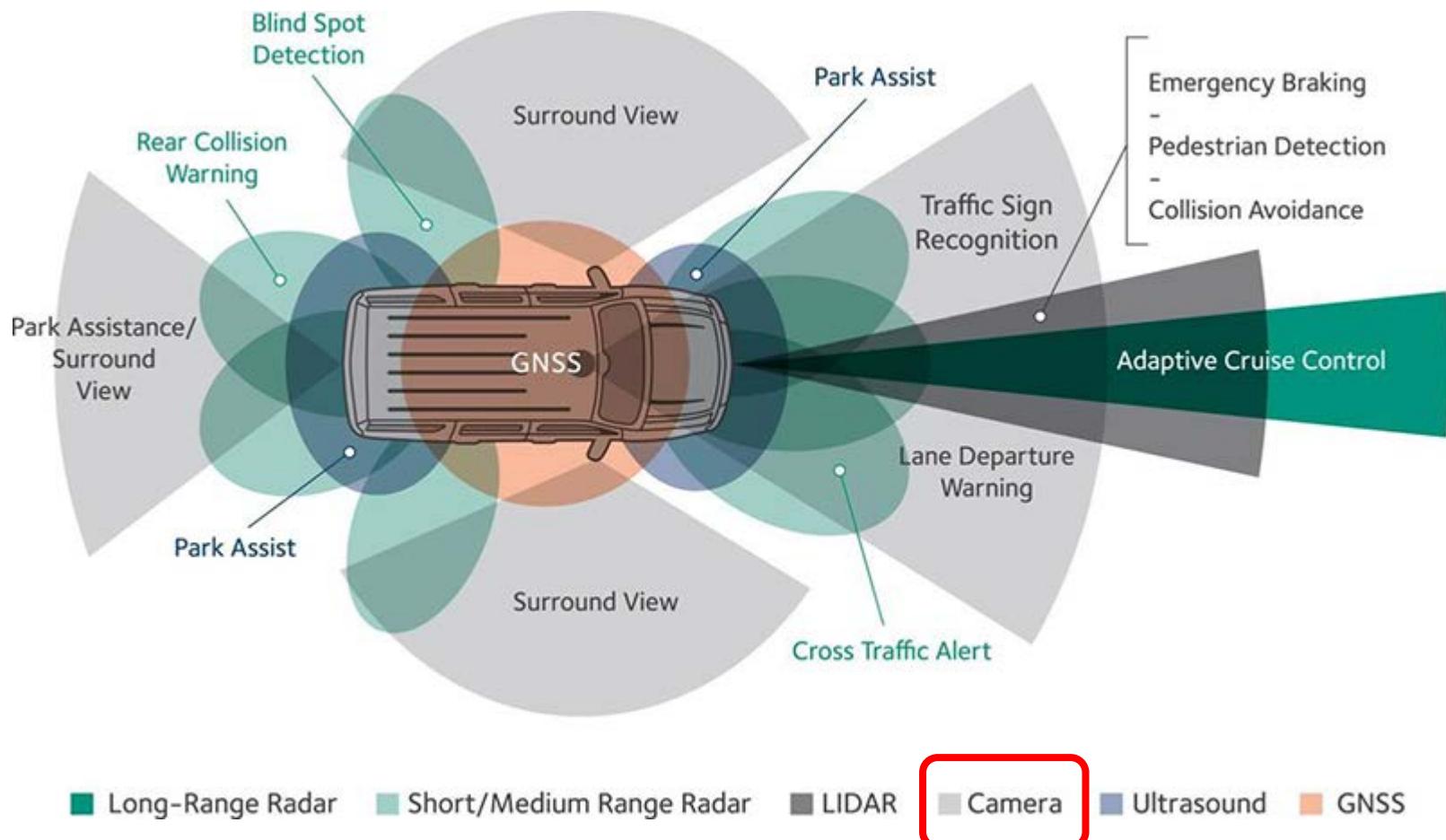


	Level 0 No-Automation	Level 1 Function-specific Automation	Level 2 Combined Function Automation	Level 3 Limited Self-Driving Automation	Level 4 Full Self-Driving Automation
설명	주행 보조 장치 없음	단일 주행 보조 기능	복수의 주행 기능 융합 보조	제한적 자율 주행	안전한 자율 주행
예시	크루즈 콘트롤, 긴급 제동, 차선 유지	차선 유지 기능 + 적응형 크루즈 콘트롤	자율 운전단, 단, 위급상황시 운전자 개입 필요	모든 환경에서 자율 주행 가능	

자료 : 미국 도로교통안전국(NHTSA)

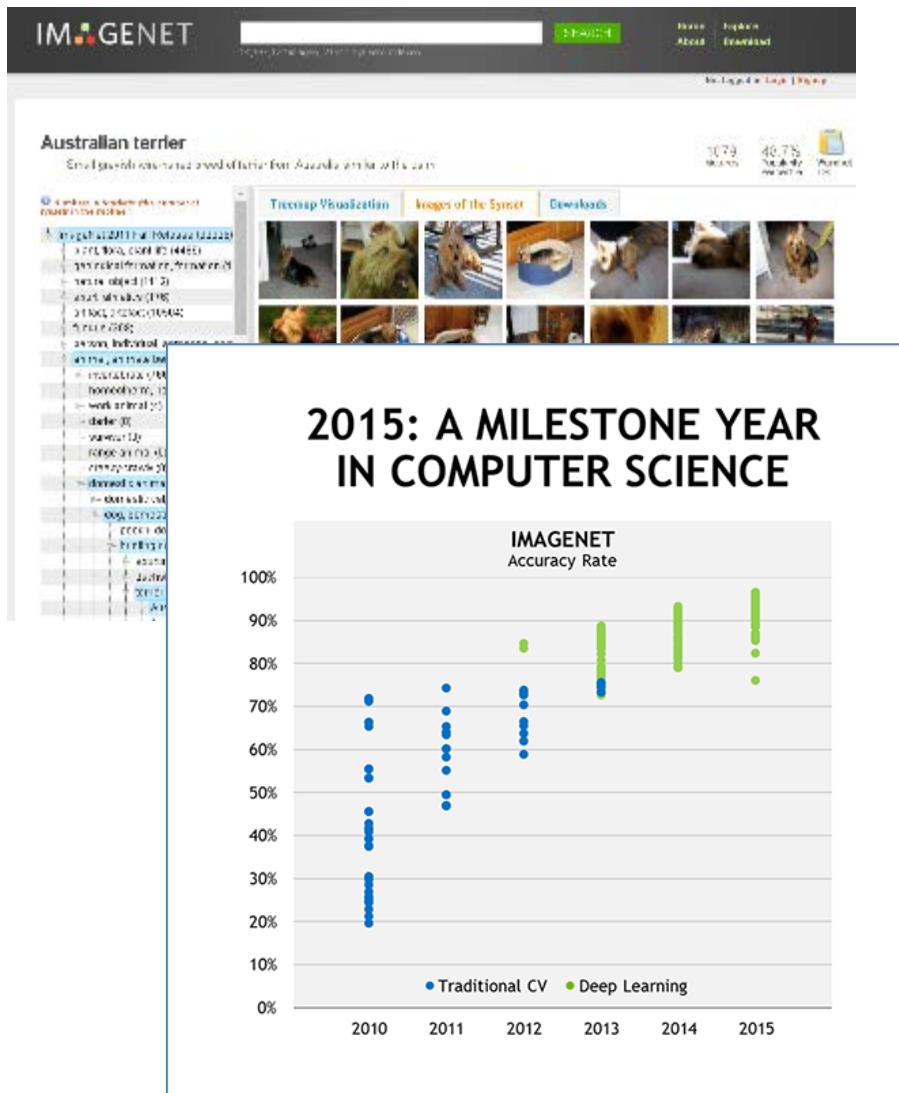


자율주행 자동차 (Autonomous Car)

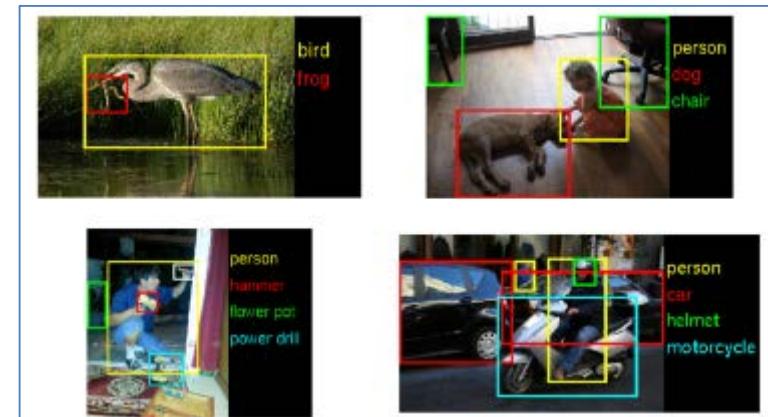


시각 인지 능력 필수 (보행자, 도로/보도, 도로 표지판, ...)

딥러닝, 시각 인지 수준을 한 단계 올리다



- ImageNet: 3만 가지 이상의 카테고리, 1천4백만여 장의 사진 DB
- 대회용으로 일부 데이터를 선별



- 2012년도: Alex Krizhevsky 등은 computer vision (CV) 기법을 사용하지 않고, 기존 CV 전문가들과 큰 격차로 1위 차지(AlexNet)
- 2015년도: 인간보다 인식율 향상. 상위 결과는 모두 딥러닝 적용

어떻게 똑똑한 컴퓨터 시스템을 만들 것인가?

세상(world)을 인식하고 이해할 필요가 있다

음성 인식
Speech Recognition

컴퓨터 비전
Computer Vision

기본적인 음성 및 시각 능력

언어 이해

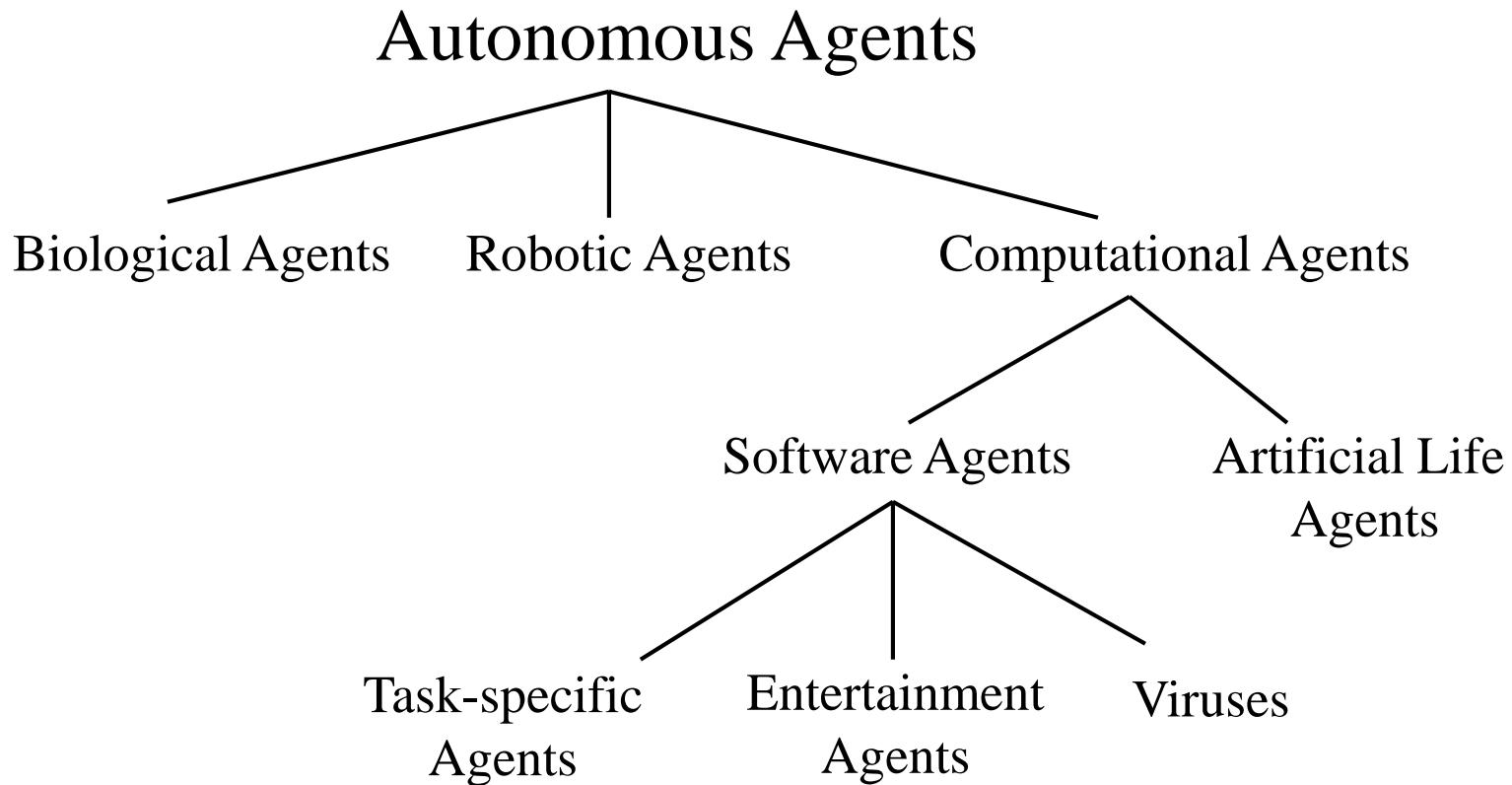
자연언어 처리/이해
Natural Language
Processing / Understanding

행동 예측

+ 지능적인 **행동**을 해야 한다

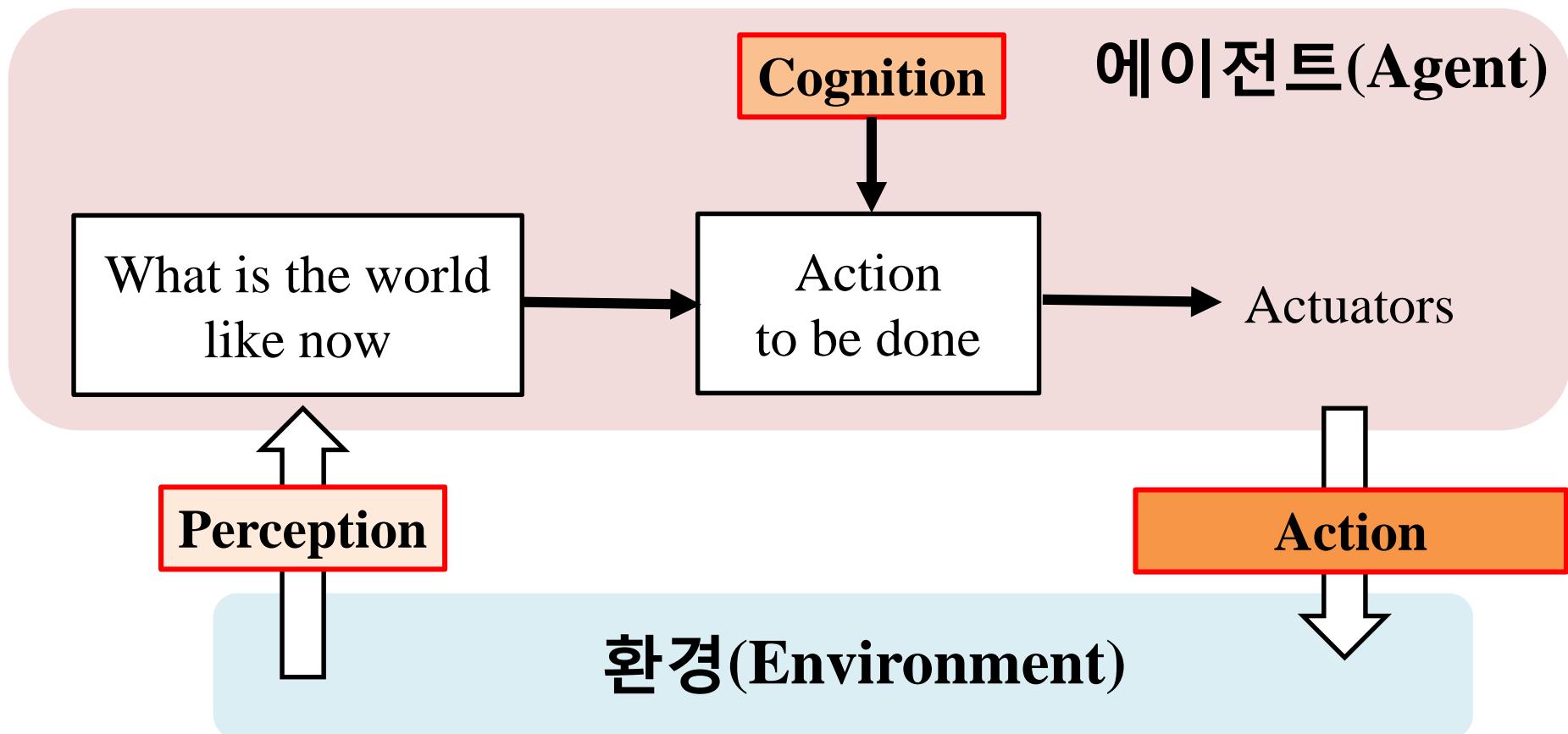
지능형 에이전트
Intelligent Agent

지능형 에이전트(Intelligent Agents)



지각, 행동, 인지 (Perception, Action and Cognition)

- An agent is anything that can
 - Perceive its environment through sensors, and
 - Act upon that environment through actuators



게임 인공지능

학습을 통해 최선의 전략과 행동 능력 습득 필요

컴퓨터 게임



Half-Life 2

- 게임에서 플레이어의 패턴을 학습하고 모방하여 인간 수준의 플레이 실현.
- 플레이어 수준에 따른 맞춤형 상대 제공

오프라인 게임

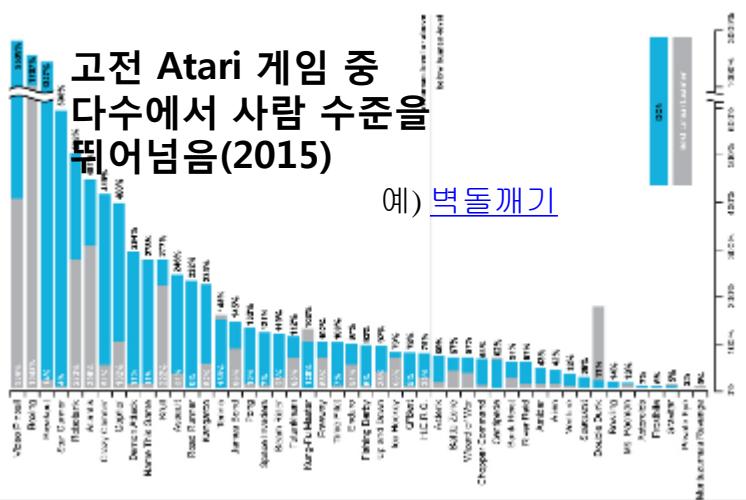
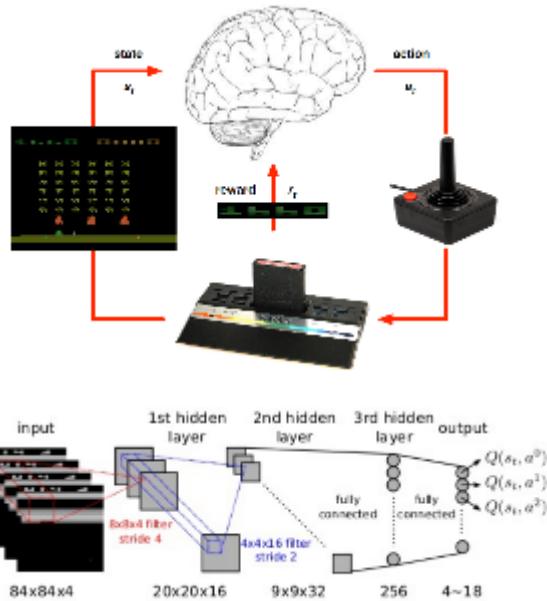


IBM Deep Blue가 체스 세계챔피언
G. Kasparov를 이김 (1997)



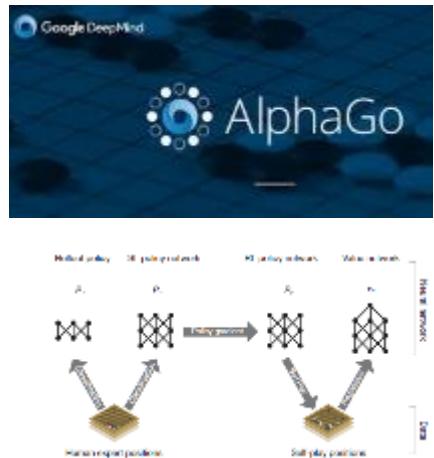
IBM Watson이 퀴즈대회(Jeopardy)에서
두 명의 챔피언을 상대로 우승 (2011)

딥러닝, 게임의 최고수를 만들어낸다

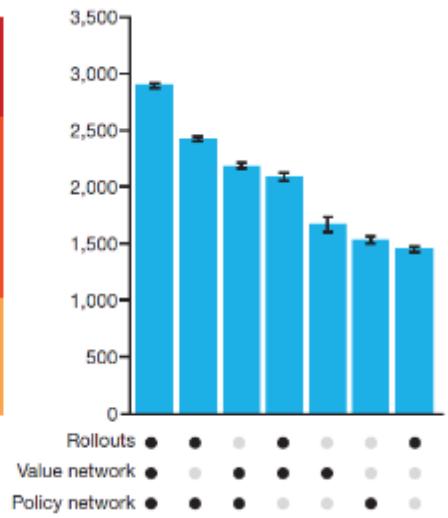
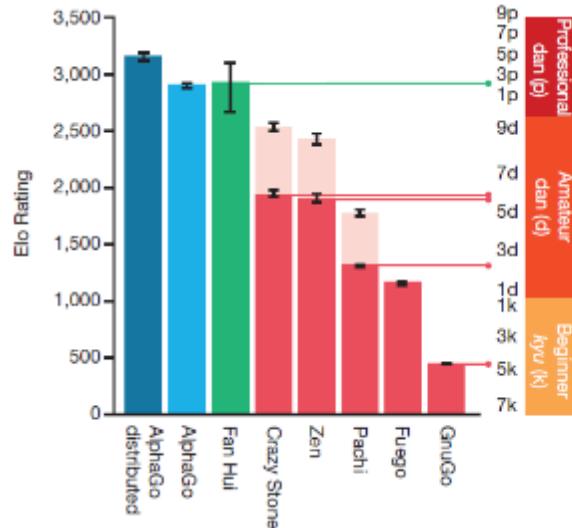


고전 Atari 게임 중
다수에서 사람 수준을
뛰어넘음(2015)

예) 벽돌깨기



바둑에서 세계 최정상 기사급의
실력을 선보임(2016)



머신러닝과 딥러닝 빠르게 이해하기

머신러닝 (Machine Learning, ML)



기계가 자동으로



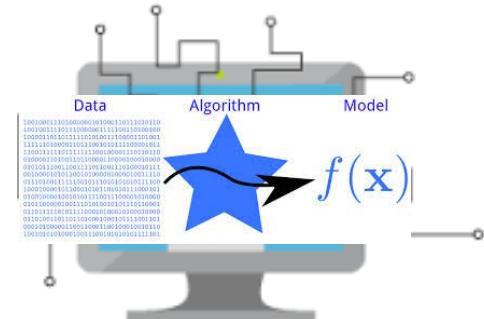
대규모 데이터에서



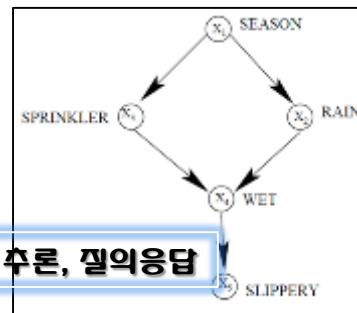
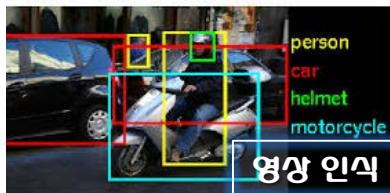
중요한 패턴과 규칙을 학습하고



의사결정, 예측 등을 수행하는 기술



Siri
음성 인식
What can I help you with?



Q. If the season is dry and the pavement is slippery, did it rain?
A. Unlikely, it is more likely that the sprinkler was ON

사진 태그 생성

Image	Generated Text
	sea, france, boat, mer, beach, river, bretagne, plage, brittany
	insect, butterfly, insects, bug, butterflies, lepidoptera
	graffiti, streetart, stencil, sticker, urbanart, street, mural, nyc, graff, sanfrancisco
	portrait, child, kid, ritratto, kids, children, boy, cute, boys, italy

사진 설명 생성

A woman is throwing a frisbee in a park.

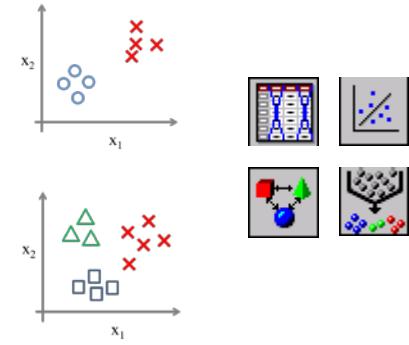
얼굴 기반 감정 인식

AD TEST FOCUS GROUP
AD TEST FOCUS GROUP
CONTENTS
100%
60%
20%

머신러닝의 세 가지 기본 학습 모드

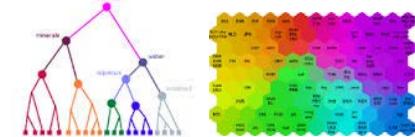
● Supervised Learning (감독학습, 지도학습, 교사학습)

- 레이블(미리 정해놓은 정답) 달린 예제로 학습하기
- 예제가 매우 많은 경우 효과적인 학습이 가능하다
- 예) 분류(classification): 레이블이 이산적인(discrete) 경우.
- 예) 회귀(regression): 레이블이 연속적 값을 가지는 경우



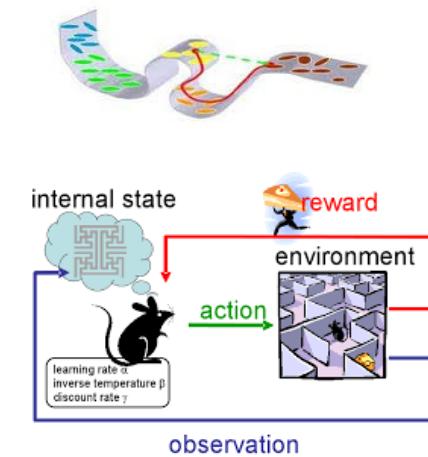
● Unsupervised Learning (무감독학습, 비지도학습, 비교학습)

- 데이터에 내재된 패턴, 특성, 구조를 학습을 통해 발견. 레이블은 고려하지 않음
- 학습 데이터는 개체에 대한 입력 속성만으로 구성됨 $D=\{\mathbf{x}\}$
- 예: 차원 축소(dimension reduction), 군집화(clustering)



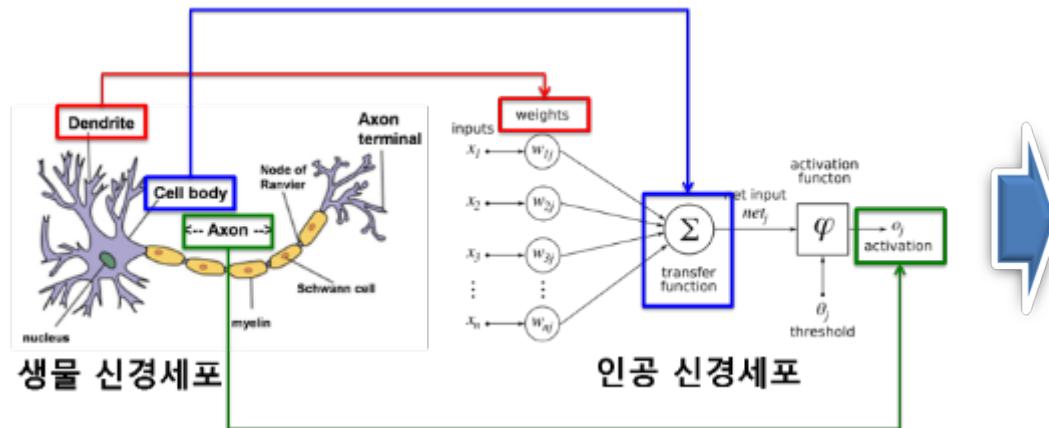
● Reinforcement Learning (강화학습)

- 시스템의 동작이 적절성(right/wrong)에 대한 피드백이 있는 학습
- 소프트웨어 에이전트가, 환경(environment) 내에서 보상(rewards)이 최대화되는 일련의 행동(action)을 수행하도록 학습하는 기법
- 환경의 상태, 에이전트의 행동, 상태 전이 규칙 및 보상, 관측 범위를 고려한 학습
- Action selection, planning, policy learning

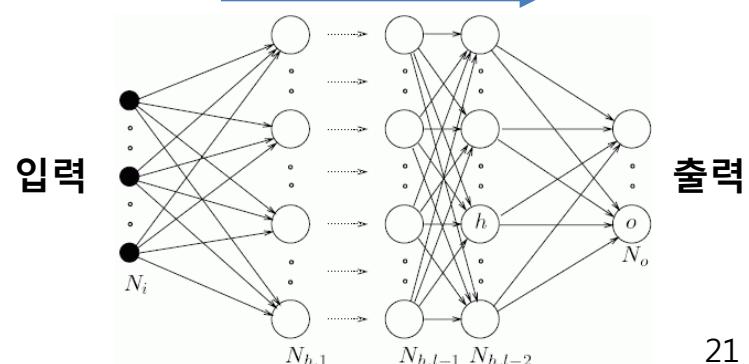


연결주의(connectionism)

- 인공지능, 인지과학 등에서 사용하는 신경 정보 처리(neural information processing)에 대한 대표적 모델 중 하나
 - 단순한 요소로 구성된 망(network)에서 지적 능력의 발현과 설명을 시도
- 인공신경망(artificial neural networks)
 - 신경세포(뉴런)의 망으로 구성된 뇌를 모사하여 계산 모델을 구성



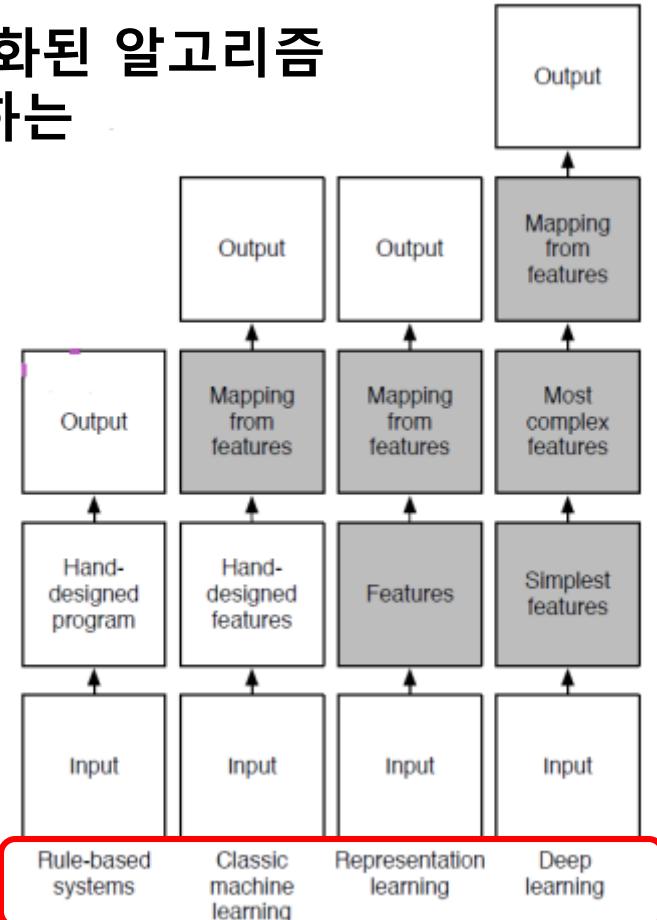
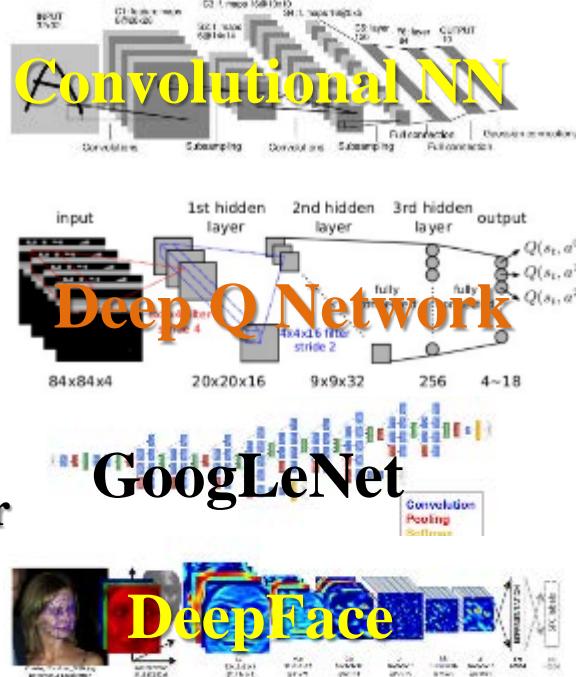
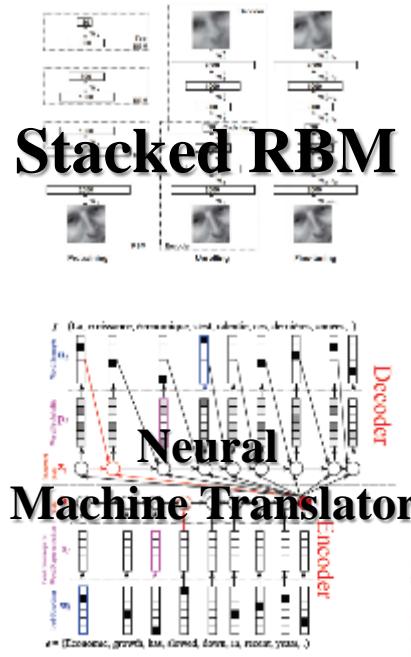
정보의 흐름(feedforward)



딥러닝(Deep Learning)

딥 러닝 기술은 사진 속 물체 인식/판별, 음성 인식, 자연언어 처리와 이해, 자동 번역 등 인공지능의 난제에서 획기적인 성능 향상을 보이고 있다.

- (좁은 의미) **인공신경망**의 다층 구조가 심화된 알고리즘
- 여러 단계의 정보 표현과 추상화를 학습하는
최신 머신러닝 알고리즘의 총칭



흰 부분: 사람이 개입, 회색 부분: 컴퓨터가 자동으로 수행

딥러닝 모델의 특징

● 다층 구조(multi layer, deep)

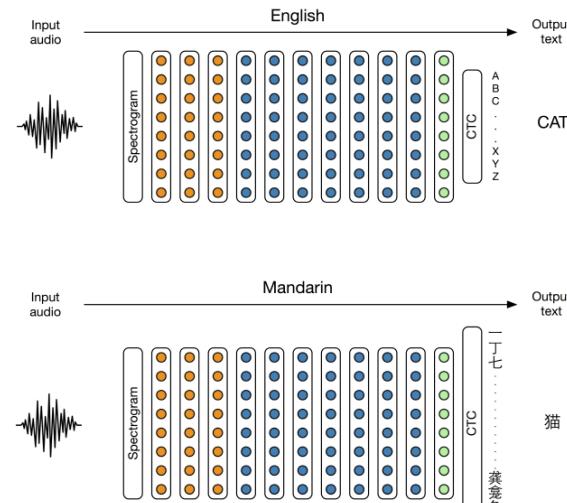
- 두뇌 신경망의 구조 모사
- 상위 층으로 갈수록 보다 추상화된 '정보가' 학습 과정에서 자동으로 생성될 것으로 기대(bottom-up, learning through layers)

● 문제 해결 과정을 학습을 통해 최대한 자동화(end-to-end)

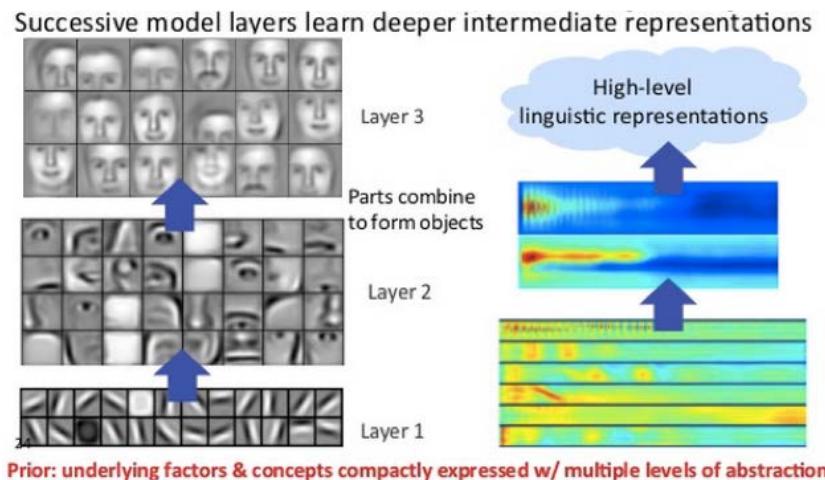
- 인간 '전문가'의 개입을 최대한 배제하고 오직 raw input과 output 사이의 모든 과정을 데이터에서 학습하는 방향 추구

● 분산 표현/정보처리(distributed representation)

- 두뇌의 또 다른 중요한 특징 모방
- 여러 뉴런이 협력하여 정보 저장/처리

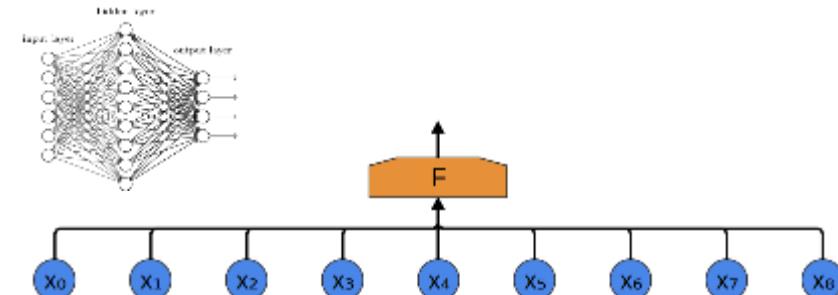


음성 인식 및 번역을 위한
딥러닝 모델 사례(Baidu)

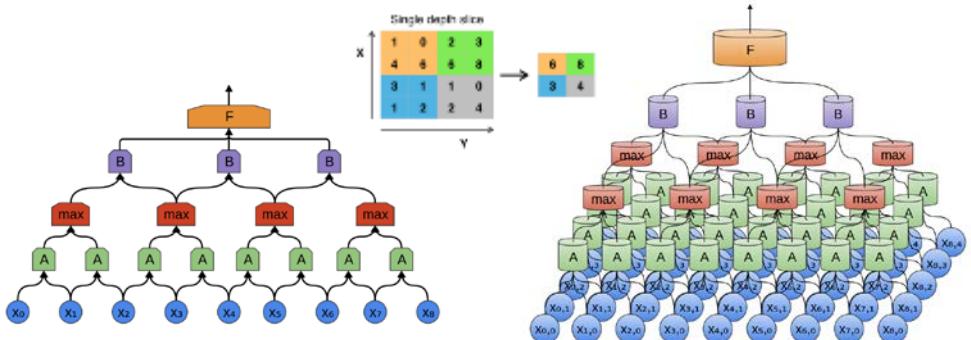


신경망의 주요 구성 형태: 망의 구조

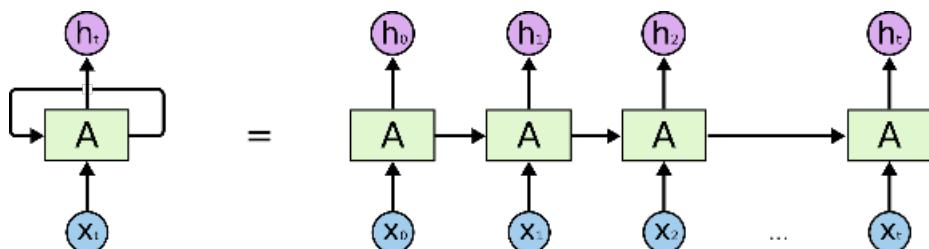
- 완전연결망(fully-connected)
 - 두 층 노드간 모든 연결선 존재



- 컨벌루션망(convolutional)
 - 여러 노드가 가중치 공유
 - Max pooling으로 줌아웃



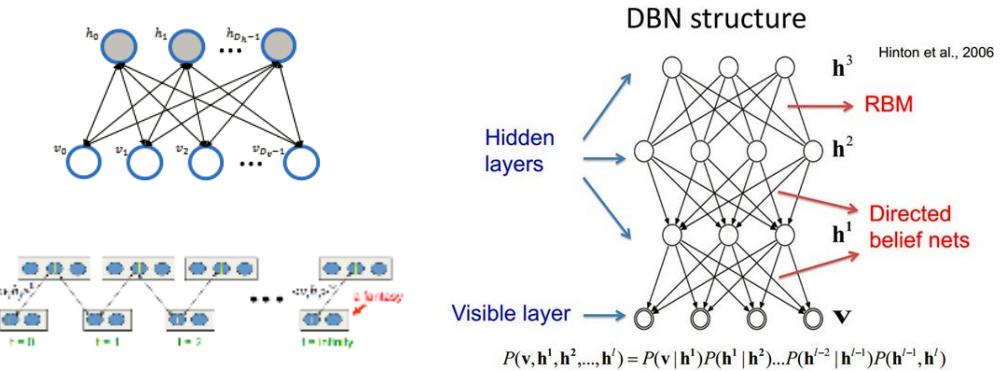
- 순환망(recurrent)
 - 연속적 데이터의 패턴 인식을 위한 신경망
 - 한 노드의 출력값, 상태값을 되돌려 추가로 입력
 - 간단한 저장장치(memory)를 장착한 효과



딥 구조를 가능하게 한 해법

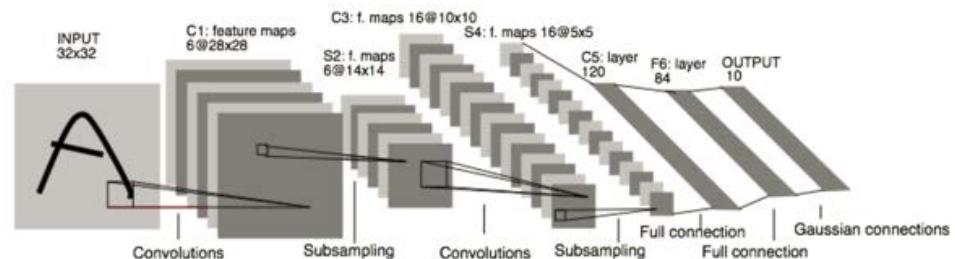
● 완전연결망(fully-connected)

- Deep Belief Net (DBN): RBM의 반복적 층 구성
- 학습기법으로 Contrastive divergence (CD) 도입
- **DBN** (deep belief network) (2006)



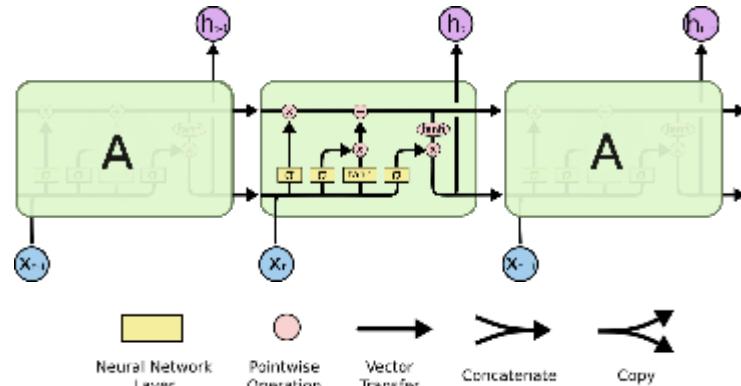
● 컨벌루션망(convolutional)

- LeNet: 컨벌루션 + Max pooling 반복 (1989)
- **CNN** (convolutional neural network)

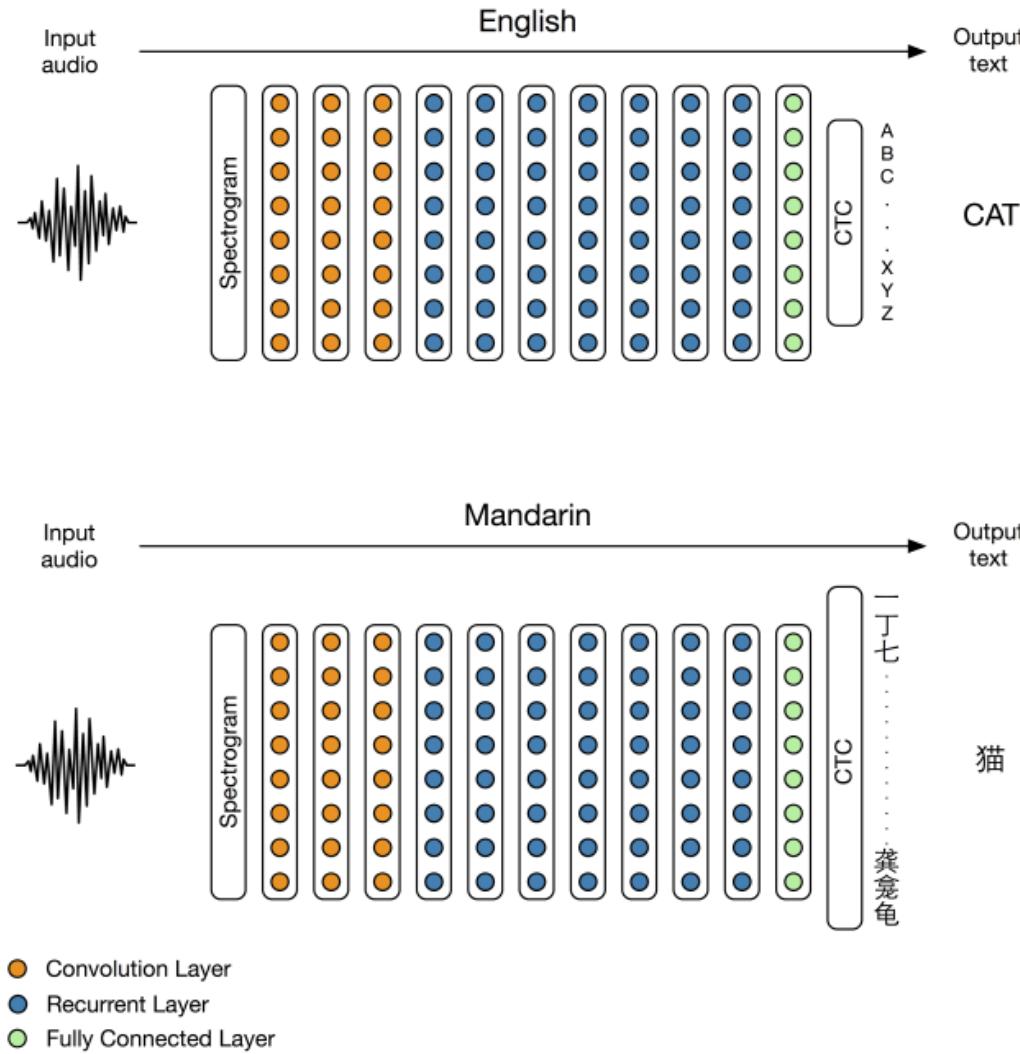


● 순환망(recurrent)

- **RNN** (recurrent neural network)
- Long Short-Term Memory (LSTM) (1997)



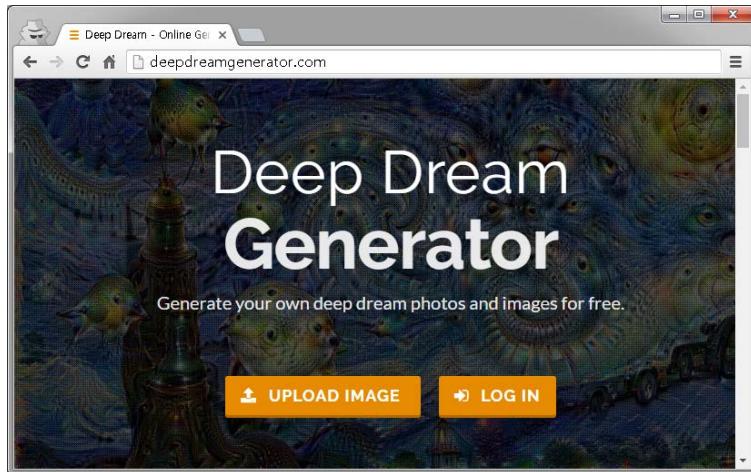
다시 보는 최신 딥러닝 모델: Baidu DeepSpeech 2



사례: Deep Learning + Image

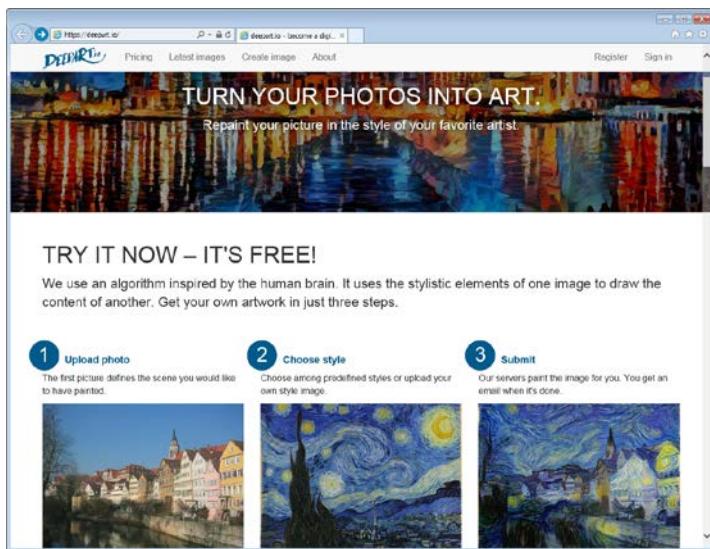
DEEP DREAM & DEEP ART 체험

Deep Dream & Deep Art 체험하기



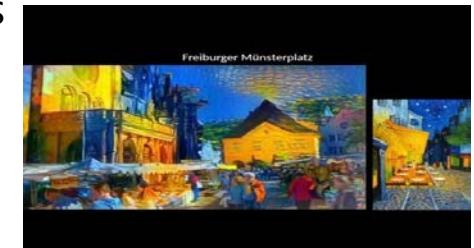
● Deep Dream

- <http://deepdreamgenerator.com/>
- 1단계: 사진 업로드
- 2단계: 온라인에서 바로 결과 확인
- 3단계: 개인 계정 생성, dream control하여 다양한 결과 확인하기



● Deep Art

- <https://deepart.io/>
- 1단계: 사진 업로드
- 2단계: 스타일을 뽑을 사진 업로드
- 3단계: e-mail로 결과 받기
- For Videos

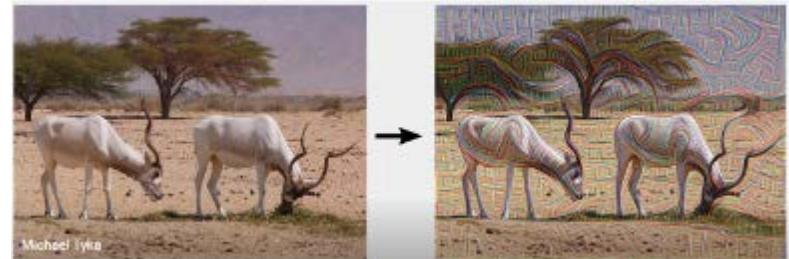


DeepDream

<https://github.com/google/deepdream>
<http://deepdreamgenerator.com/>

딥러닝은 아티스트를 위한 도구가 될 수 있다 – a new way to remix visual concepts

- neural networks that were trained to discriminate between different kinds of images have quite a bit of the information needed to generate images too
- feed the network an arbitrary image or photo and let the network analyze the picture
- pick a layer and ask the network to enhance whatever it detected

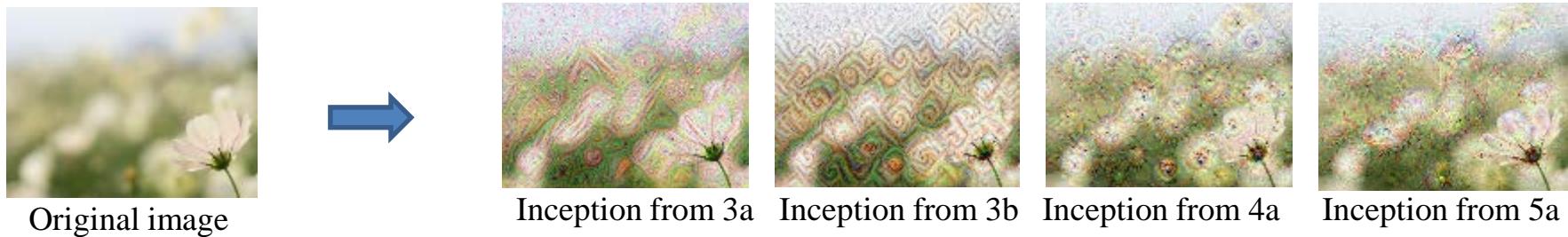
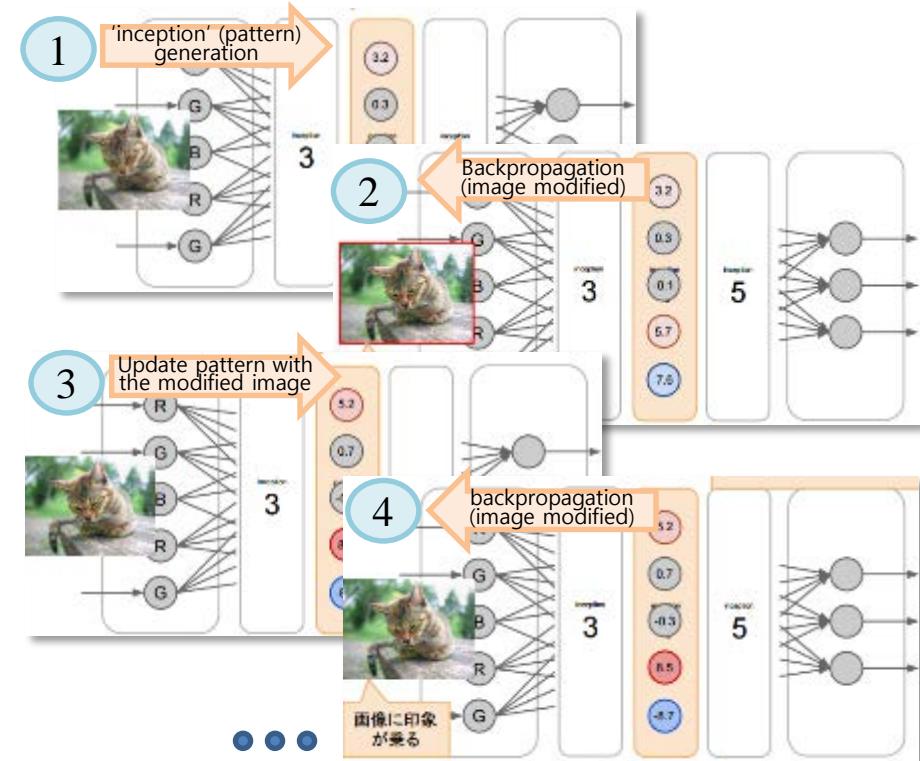
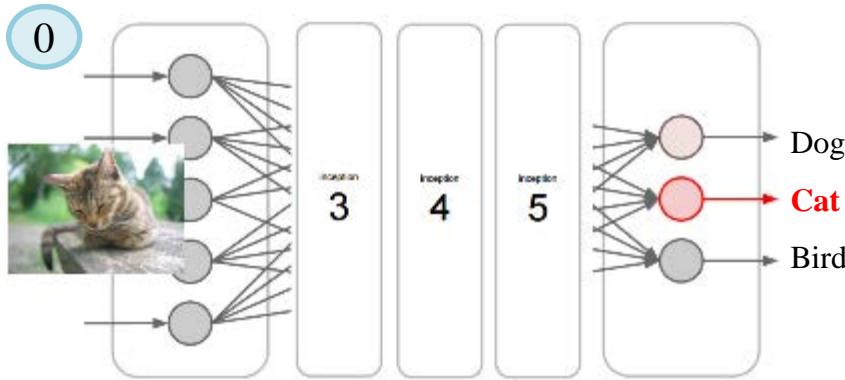


From lower layer



From higher layer

DeepDream



Images from slides by H. Nobuta from <http://www.slideshare.net/HarumitsuNobuta/deep-dream>

DeepArt: A Neural Algorithm of Artistic Style

<http://arxiv.org/pdf/1508.06576v2.pdf>

- 딥신경망으로 고급스러운 예술적 이미지를 생성 가능
- 그림의 내용(contents)과 스타일(style)을 구분하여 신경망으로 표현을 학습하고 재조합

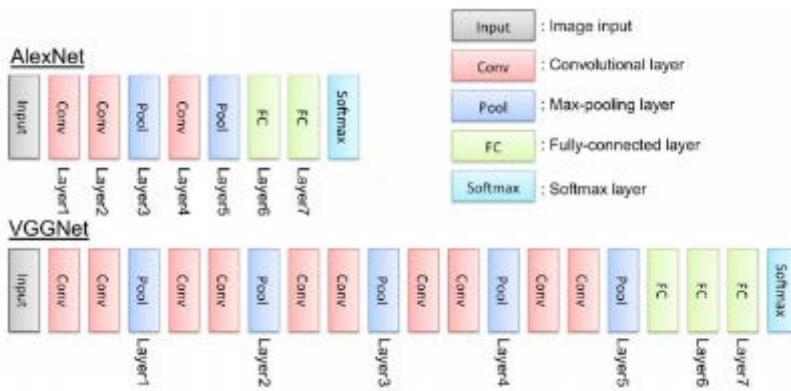
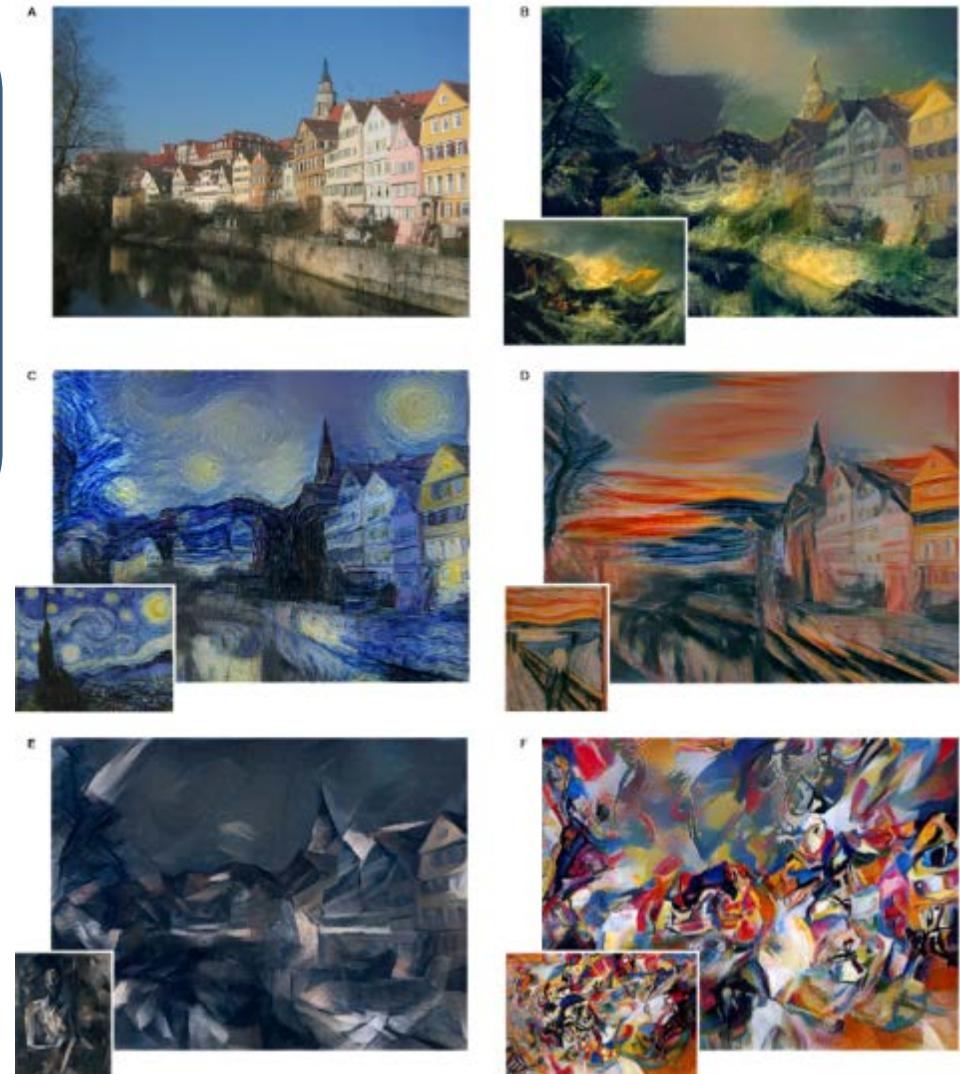
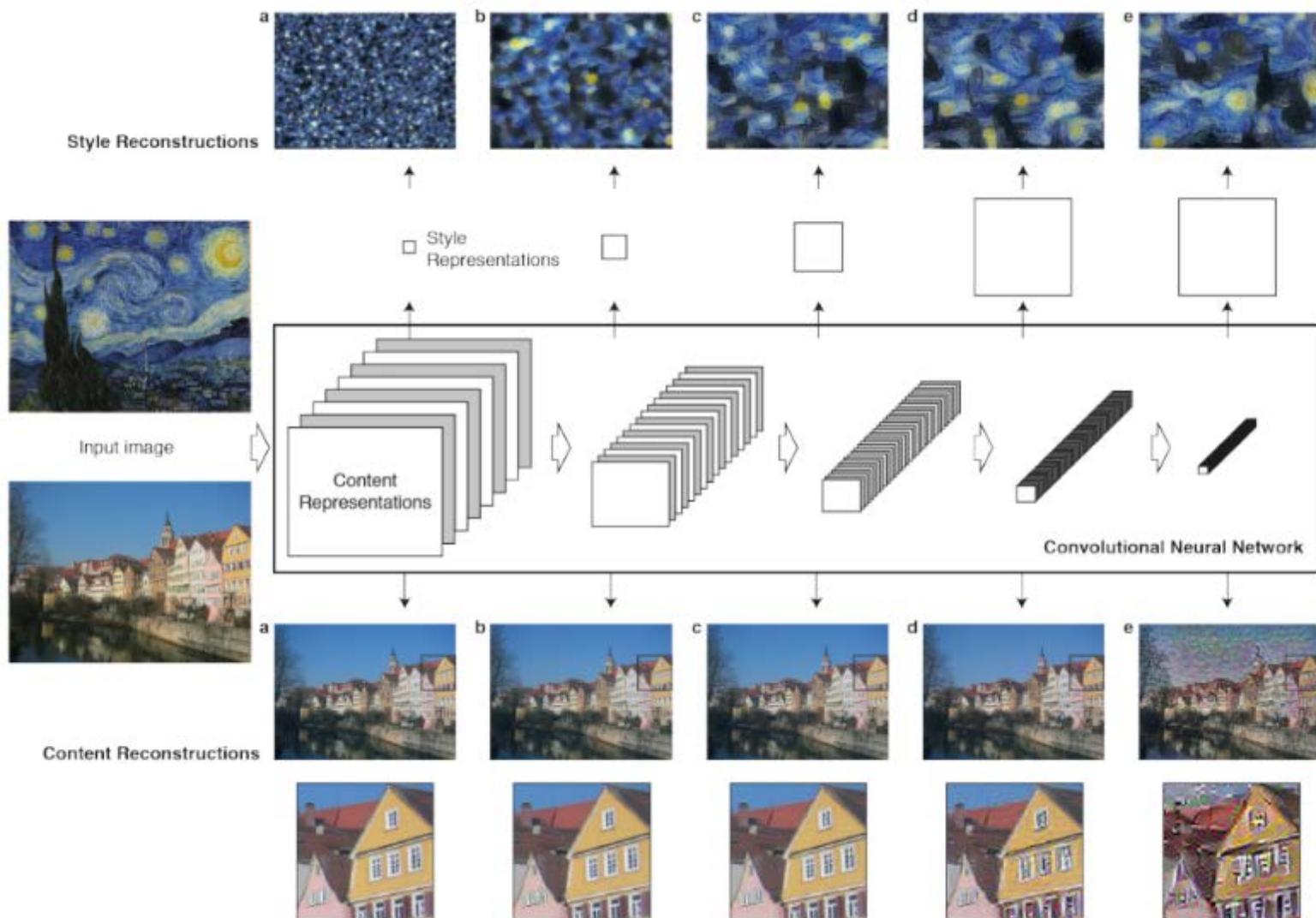


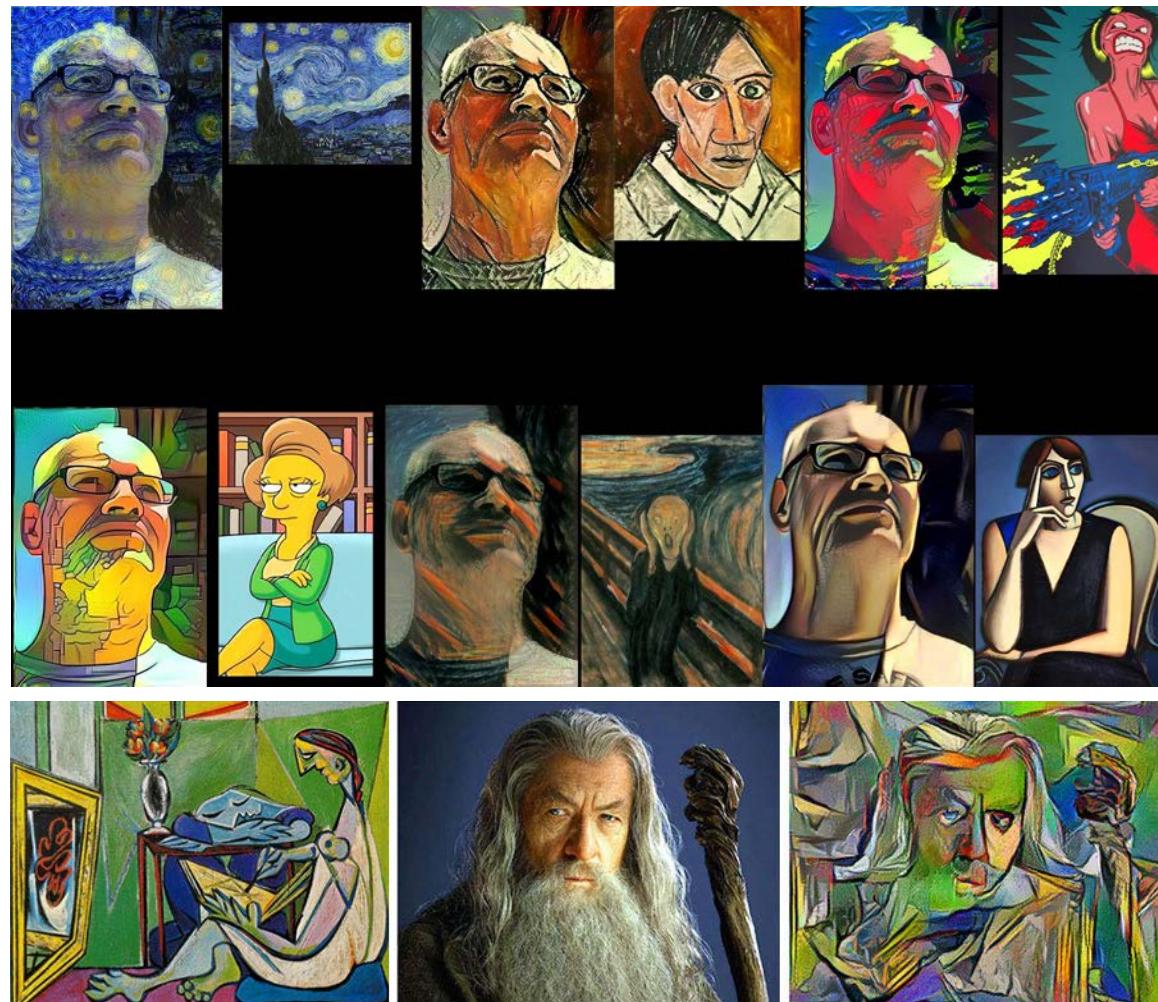
Figure from (H. Katoaka, K. Iwata, Y. Satoh, 2015)



DeepArt: A Neural Algorithm of Artistic Style



Artistic Style Image Generation



Online Demo <https://deepart.io>

사례: Deep Learning + Sequence

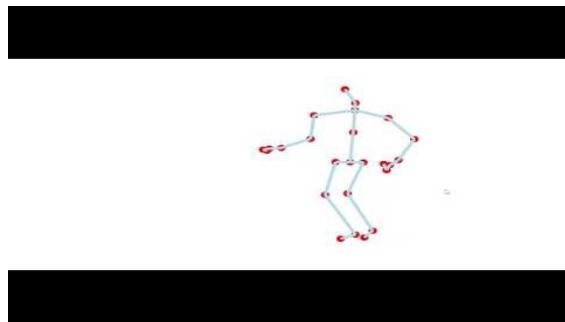
음악 생성, 댄스 생성, 필기체 생성

연속 정보 처리 Deep Learning 적용



● 음악생성

- <https://keunwoochoi.wordpress.com/2016/02/19/lstm-realbook/>
- RNN으로 재즈 코드 생성한 사례



● 댄스 생성

- <http://peltarion.com/creative-ai>
- RNN 기반의 DL로 댄스 생성

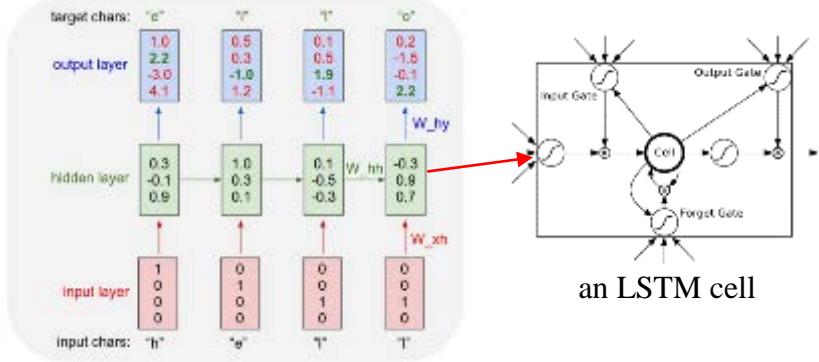
heal the world, make it a better place
yesterday, all my troubles seem so far away

● 필기체 생성

- <http://www.cs.toronto.edu/~graves/handwriting.html>
- 댄스 생성과 유사한 기법이 핵심적으로 사용됨

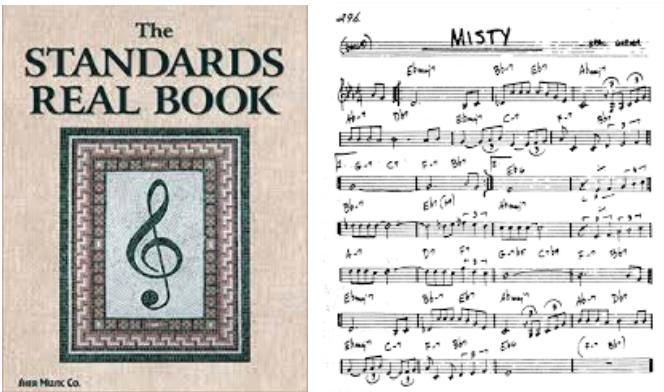
LSTM Realbook: Generation Jazz chord progressions

<https://keunwoochoi.wordpress.com/2016/02/19/lstm-realbook/>



Char-RNN / Word-RNN

Character(word)-level language models



START A:7 A:7 A:7 A:7 A:7 C:maj C:maj _END_

Preprocess into text-format scores

- 2,486 scores
- 1,259 unique chords
- 39 characters (including C,D,E,F,G,A,B,#,b,:,m,a,j,(m),i,n,s,u,(s), ...)
- 1,239.78 characters per score on average
- Maximum 10,151 characters in a score

Dataset configuration

```

1 model = Sequential()
2 model.add(LSTM(512, return_sequences=True, input_
3 model.add(Dropout(0.2))
4 model.add(LSTM(512, return_sequences=True))
5 model.add(Dropout(0.2))
6 model.add(LSTM(512, return_sequences=False))
7 model.add(Dropout(0.2))
8 model.add(Dense(num_chars))
9 model.add(Activation('softmax'))
10 model.compile(loss='categorical_crossentropy'),

```

Keras and Theano codes
https://github.com/keunwoochoi/lstm_real_book



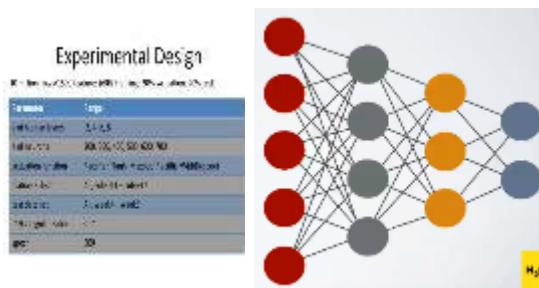
사례: Deep Learning + 금융 투자, 분석, 거래, 탐지

**빅 데이터와 딥러닝-머신러닝의 결합
은행 기술 핀테크**

- 3000여 개의 특징을 가지는 복잡한 데이터



- 딥러닝 머신러닝 모델

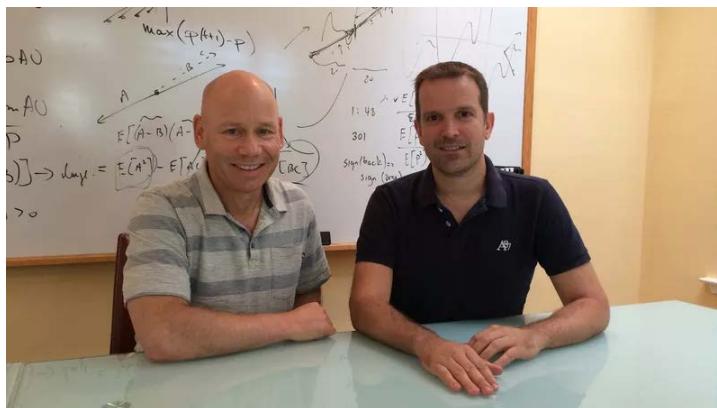
H₂O.ai

- 고객 정보로 테스트해본 결과
85%의 성능으로 사기/일반 거래 예측

- 2015년 모회사인 ebay에서 분사
시가 총액



- 딥러닝을 이용한 자기자본투자 회사 (proprietary trading company) (2011~)
 - Side 비즈니스로 헛지 펀드도 관리
- 회사 정보와 투자 활동은 철저히 비밀리에 이루어지고 있음
- “지난 3년간 매우 성공적이었습니다. 우리가 단순한 행운 덕이 아닌 맥점을 짚고 있다고 확신합니다.” (Itamar Arel, CTO, 2014년 9월 인터뷰)
- **딥러닝의 활용**
 - 다중 시간 단위 패턴 분석: 일, 주, 월, 계절, 년, ...
 - 시계열 예측 기반 주식 거래
 - 알파고에도 적용된 ‘딥 강화 학습 기법’도 적용하고 있는 것으로 예상됨



Ben-Efraim (CEO)과 Itamar Arel (CTO)



홈페이지는 1장의 그림 뿐

- Warren: S/W solution on Big Data, Natural Language, Actionable Insights
 - 빅데이터 분석을 통해 전 세계 주요 사건이 경제 시장에 미치는 영향을 분석, 시각화 (visualization), 보고서 자동 생성 (vs. Bloomberg, Thomson Reuters, Bridgewater Associates)
 - Google 스타일 검색: 자연어를 입력하여 결과를 확인
 - 9만여 가지 동향/지표 추적 (통화정책 변경, 신약 승인, 경제보고서, ...)
 - 1억 가지 탑입의 경제 관련 질문에 대답 가능 (*quant army*로 비유)
 - 예) 3등급 허리케인이 플로리다를 강타할 때 어떤 시멘트 주식이 가장 오를까?
 - 골드만삭스 및 CNBC와 협약. CIA도 사용



CNBC - 'Kensho STATS Box'
2016-4-22, Apple 주식이 1달 사이 1% 하락



CNBC, 'Ask Kensho'

Deep Learning 사례 - 음성, 사진

실시간 음성대화 번역



시연
영상

시연
영상

(시연 시작위치 3:02~)

음성인식 전 과정 자동화

Baidu DeepSpeech



얼굴 인식 시스템 고도화

- **Facebook DeepFace**
 - 2D+3D: 여러 각도에서 인식 가능
- **Google FaceNet**
 - 다양한 포즈, 조명 하에서도 인식

예시

예시

사진 자동 추상화 / 합성

Google DeepDream



- 딥러닝을 이용한 추상화, 꿈 속 영상

예시

Google DeepStereo



- 사진을 합성, 다른 각도의 사진 생성

시연
영상

Deep Learning 사례 – 추천, 금융

추천 시스템

- Netflix 영화 추천



- All-in cloud strategy: AWS + GPU

- 음악 추천



- Google, Spotify, Pandora: 딥러닝을 이용 플레이리스트 자동생성기술 모색

동영상 서비스 관리

- YouTube



- 비정상적 동영상 조회 증가 탐지/관리
- 추천의 기본 틀은 Amazon과 유사 (item-to-item 협업필터링)

사기 탐지

설명
자료

- PayPal

- 0xdata 사의 H₂O 머신러닝 플랫폼 기반



투자/거래

- Binatix



- 딥러닝으로 주식 시장의 동적 패턴 대응
- 자기자본거래, 헛지펀드

- Kensho



- 금융거래, 투자 관련 자동 자료 수집, 분석 및 보고, 의사결정 지원

인공지능 & 딥러닝 동향 및 전망

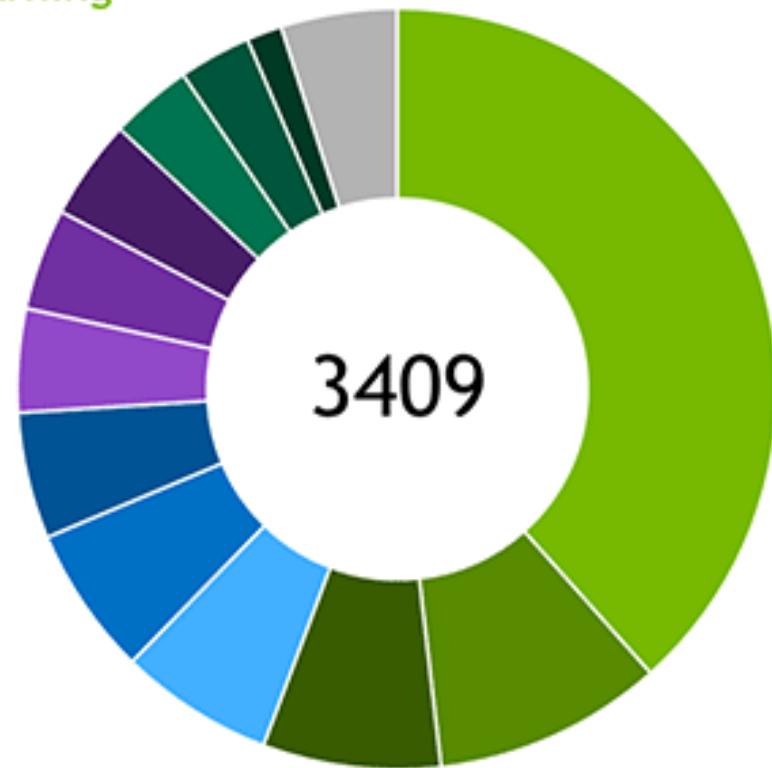
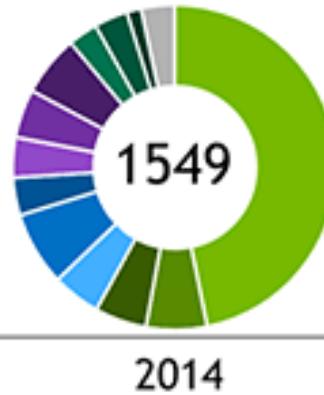
EVERY INDUSTRY WANTS INTELLIGENCE

Organizations engaged with NVIDIA on deep learning

- Higher Ed
- Internet
- Life Sciences
- Development Tools
- Finance
- Media & Entertainment
- Government
- Manufacturing
- Defense
- Automotive
- Gaming
- Oil & Gas
- Other

100

2013



인공지능 기술 회사의 폭발적 증가

Facts and Forecasts: Boom for Learning Systems

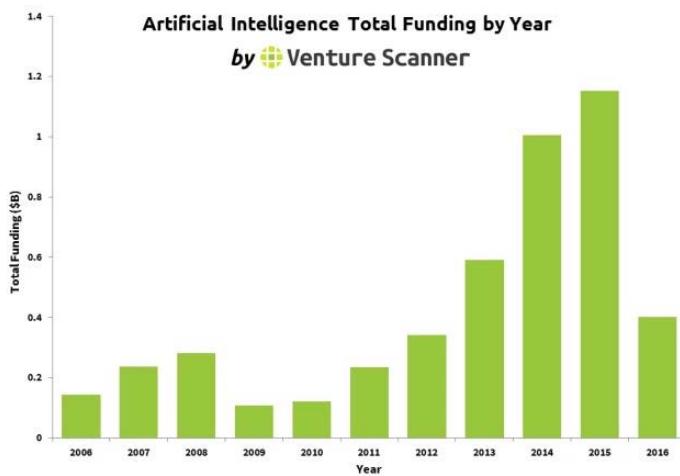
<http://www.siemens.com/innovation/en/home/pictures-of-the-future/digitalization-and-software/artificial-intelligence-facts-and-forecasts.html>

Autonomous Robots to Surpass Expert Systems: Forecast Share of the Smart Machine Market



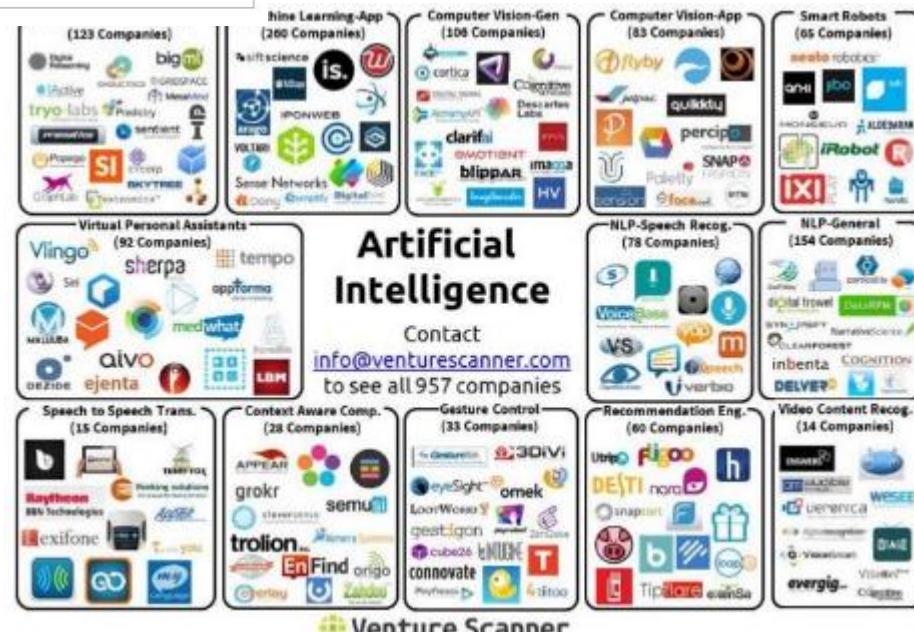
The global market for smart machines is growing by almost 20 percent annually.

‘스마트 머신’의 전세계 시장 규모는 매년 20%씩 증가하고 있다.



딥러닝을 필수 기술 요소로 포함하는 사례 증가

<http://insights.venturescanner.com/2015/01/29/going-deep-on-deep-learning-technology/>



<http://insights.venturescanner.com/tag/artificial-intelligence/>

머신러닝 인공지능·클라우드 컴퓨팅·빅데이터 결합 컴퓨팅환경이 혁신 토대

<http://www.nytimes.com/2016/03/28/technology/silicon-valley-looks-to-artificial-intelligence-for-the-next-big-thing.html?smprod=nytcore-iphone&smid=nytcore-iphone-share>



IT 핵심기술은 프로그래밍 넘어 머신러닝
컴퓨터에 작업 지시하는 것에서 학습시켜 일하는 시대로
(에릭 슈밋, 알파벳 회장)

회사를 대상으로 한 머신러닝 및 인공지능
활용법 교육이 큰 비즈니스가 될 것이다.

(Diane B. Greene, the head of Google Compute Engine)

연합뉴스, 2016/3/24

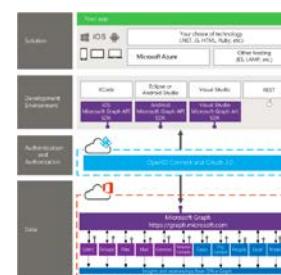


NY Times, 2016/3/27

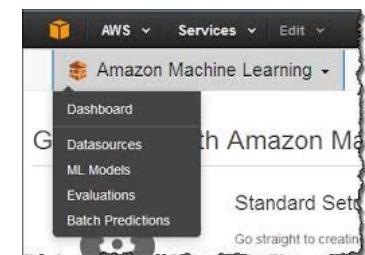
기계와 소프트웨어(주: 구글 클라우드 지칭) 대여 비즈니스 수익이
점차 Google 광고 수익을 초월할 것이다. (2020년 기준)
(Urs Hölzle, chief of technical infrastructure of Google Compute Engine)



Google
Cloud Platform + ML



Microsoft
AI on "Microsoft Graph"



Amazon
AWS + ML

Don't Buy Machine Learning

A few comments for those who are about to invest on Machine Learning* intensive project

<http://www.machinalis.com/blog/dont-buy-machine-learning/>

"Machine Learning development is like the raisins in a raisin bread: 1. You need the bread first 2. It's just a few tiny raisins but without it you would just have plain bread."

The trick is not buying "the Machine Learning thingy..."

Unless you...

- ... realize that by doing so you're not getting into a typical software development project, but a "research" and development one. Even for some better-known problems that do not strictly follow this rule, the typical level of uncertainty is higher than in traditional development projects, and much higher if the expectations about the deliverable involve identify, design and/or achieve specific values for metrics like precision or accuracy.
- ... understand that there can be a large gap between the research paper that says "it's possible" to the production ready software, sometimes so big that it can't be covered today.
- ... view your predictive models created in the process as an asset. Your competitors don't have access to this model, it is a piece of productive and potentially scalable knowledge which emerged from your very own combination of data and constraints. The team working on this is not wasting time as a "necessary evil", but investing effort in the main asset of the process.
- ... do not expect that Machine Learning will be a black box that will produce magical results. A central principle of the process is "garbage in, garbage out" and if your goals are not clear, Machine Learning won't make them any clearer.



Growing Use of Deep Learning at Google

of directories containing model description files

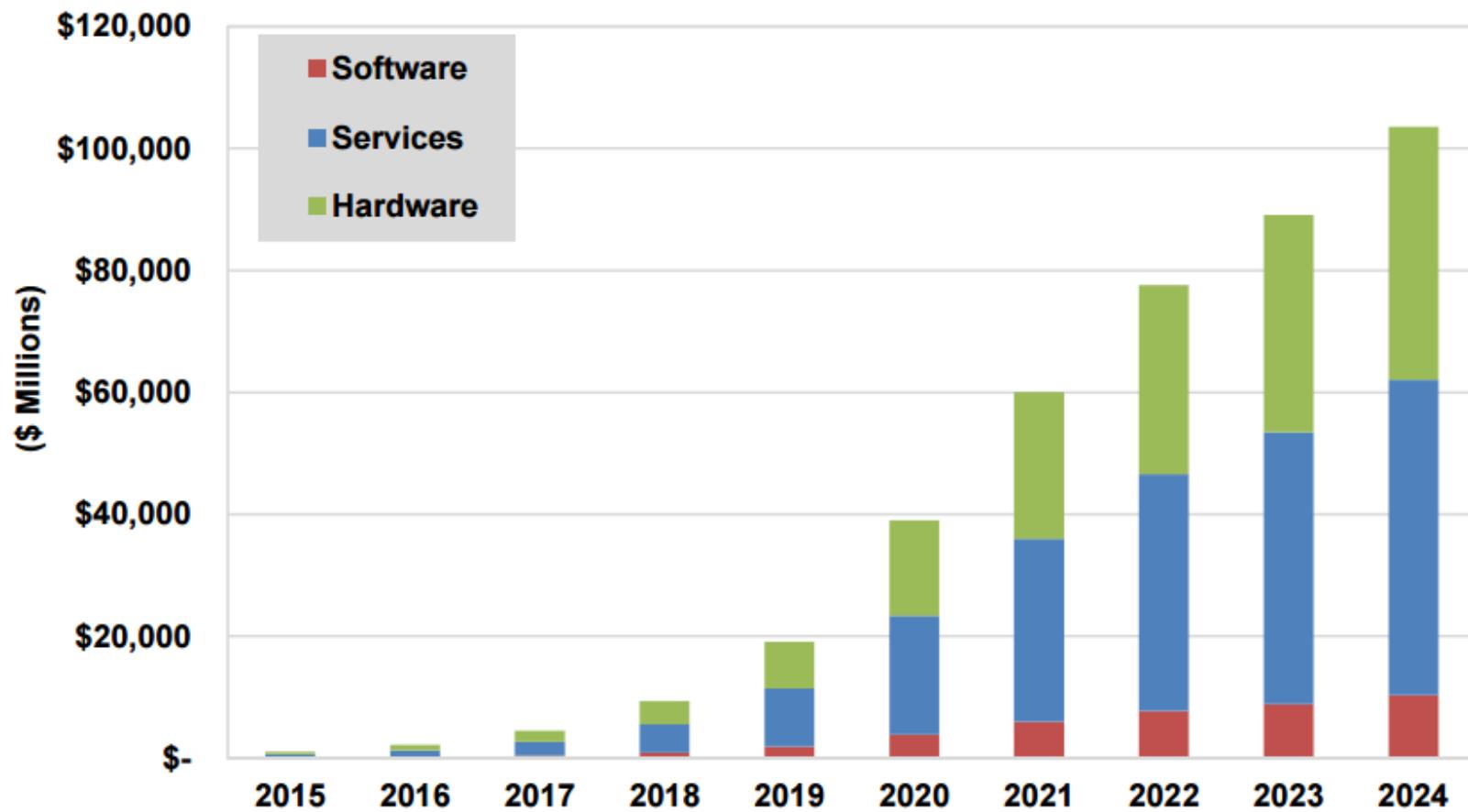


Across many products/areas:

- Android
- Apps
- drug discovery
- Gmail
- Image understanding
- Maps
- Natural language understanding
- Photos
- Robotics research
- Speech
- Translation
- YouTube
- many others ...

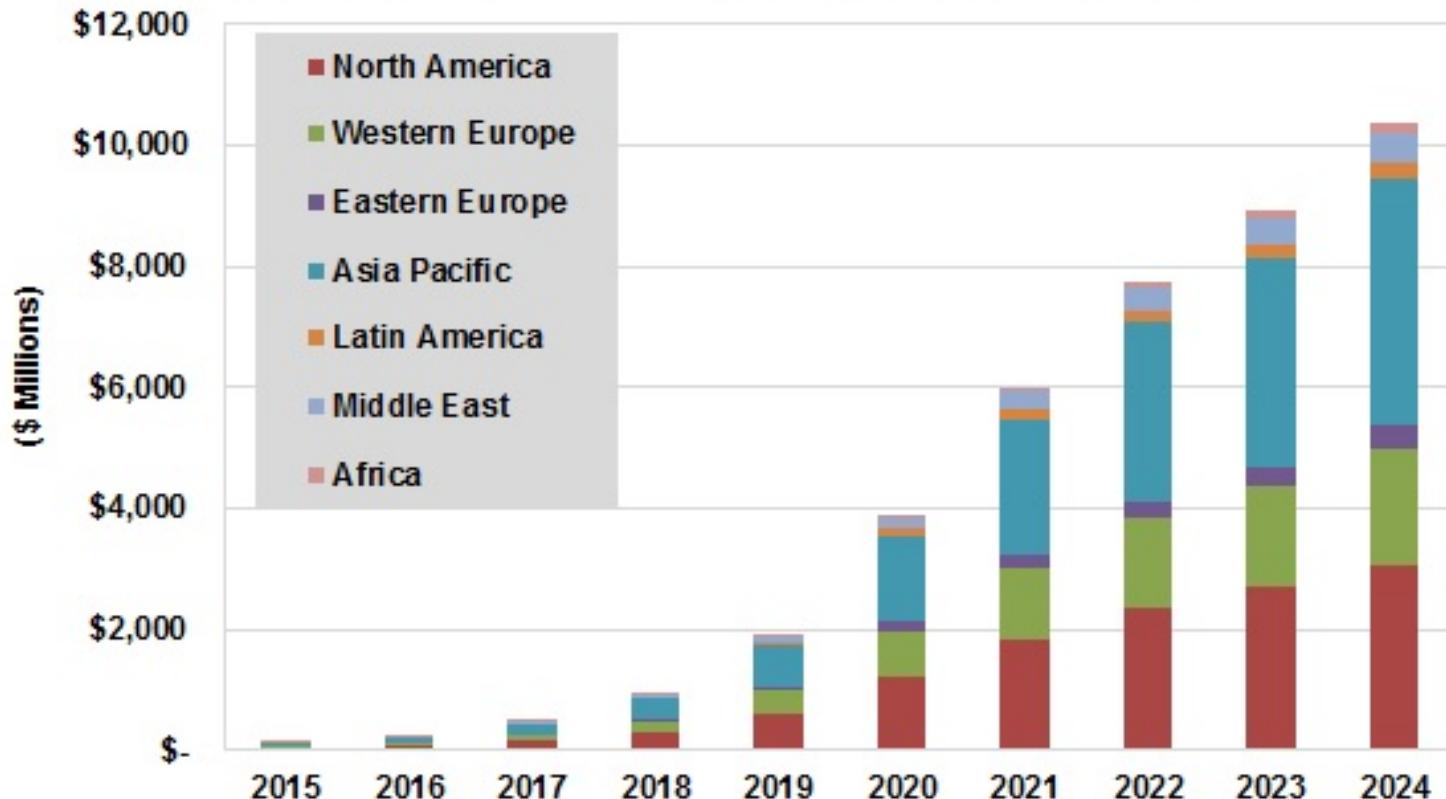


Chart 1.1 Deep Learning Total Revenue by Segment, World Markets: 2015-2024



(Source: Tractica)

Deep Learning Software Revenue by Region, World Markets: 2015-2024



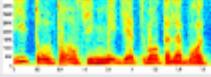
Source: Tractica

<https://www.tractica.com/newsroom/press-releases/deep-learning-software-market-to-surpass-10-billion-by-2024/>

딥러닝 기반 지능 시스템 구현의 방향

■ 현재

- Unimodal: 시각, 소리, 언어, 지식
- Multimodal: 시각+언어, 소리+언어

Input	Output
Pixels: 	"ear"
Audio: 	"sh ang hai res taur aun ts"
<query, doc1, doc2>	P(doc1 preferred over doc2)
"Hello, how are you?"	"Bonjour, comment allez-vous?"

- 실용적 응용을 위해서는 대규모 컴퓨팅 파워 필요



■ 근미래

- 알고리즘적 측면
 - Unsupervised Learning 기법과 활용 증가 (인간과 동물의 학습은 대부분 unsupervised. discovery)
 - End-to-end 방식 해법
 - CNN/RNN + 강화학습(RL)
 - Representation learning + complex reasoning

■ Unimodal

- 기존 modal에서의 성능 향상: computer vision, NLP, ...

- 보다 다양한 센서 정보 추가

■ Multimodal

- 보다 다양한 조합에의 적용
- 다양한 계층의 추상화

■ Computing power

- 고성능 컴퓨팅 기반 대규모 응용
- 모델의 경량화: 학습 모델의 압축
- 저전력용 알고리즘

Smart Machine의 시대

- 스마트어드바이저, 지능형로봇, 스마트카 등의 Smart Machine은 IT 역사상 가장 파괴적인 기술이 될 것으로 예측 (Gartner, 2014)
- 인공지능 시장 규모가 10년 뒤 70조 달러를 넘어설 것으로 예측 (YTN, 2015)

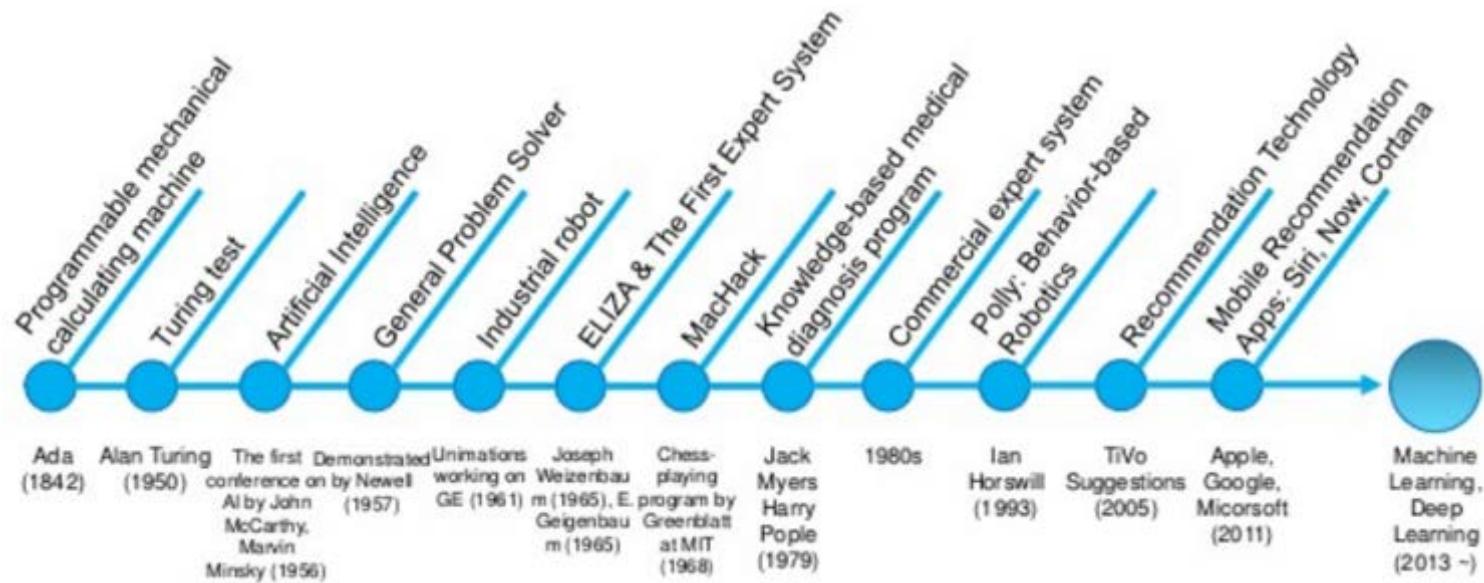


The Coming Age of Smart Machines



에필로그

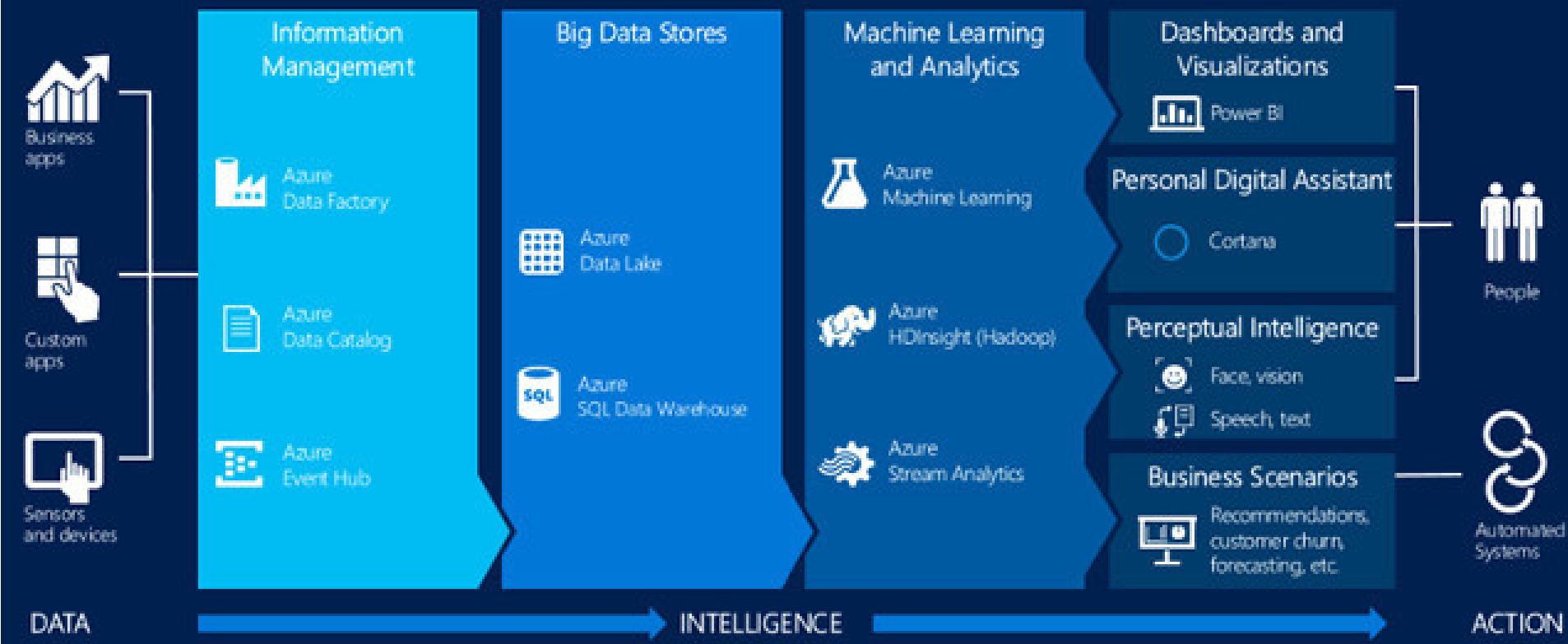
AI Timeline



부록 / 미사용 슬라이드

Cortana Analytics Suite

Transform data into intelligent action



왜 지금 딥러닝이 뜨게 되었는가?

DEEP LEARNING 기반 인공지능의 원동력

네 가지 원동력

Computing Power
(특이 GPU 기반
병렬 계산)

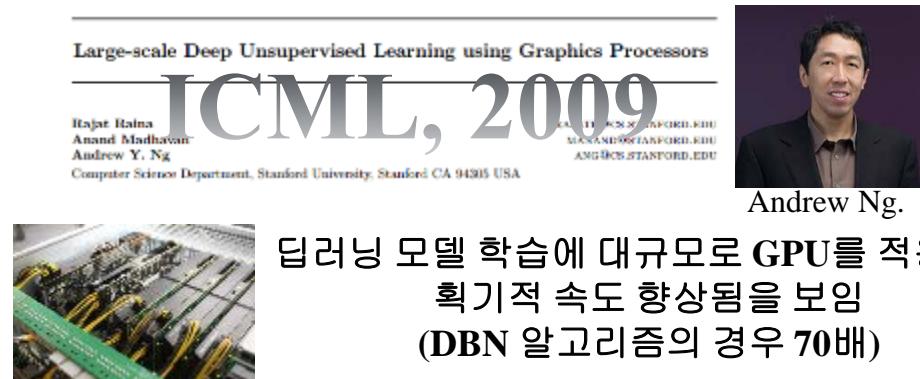
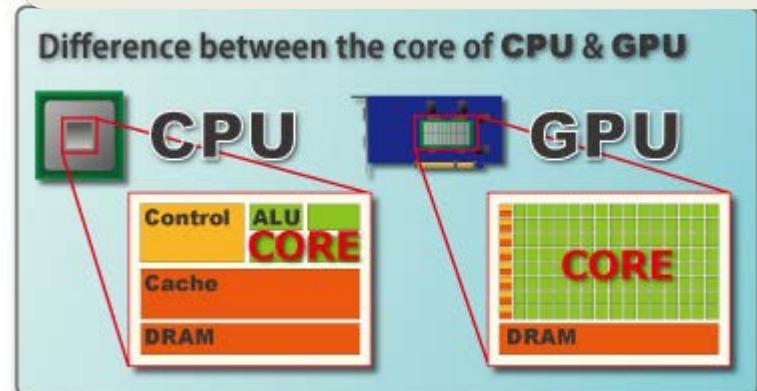
Big Data
빅데이터

획기적인
알고리즘

개발 환경의
진화

원동력 1: Computing Power

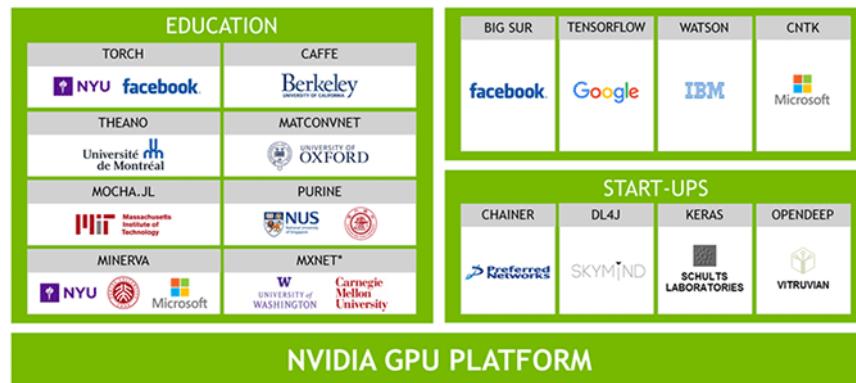
컴퓨터의 계산 속도 향상으로 대규모 모델 학습 가능
특히, Graphics Processing Unit으로 CPU보다 빠르게 병렬 계산 가능



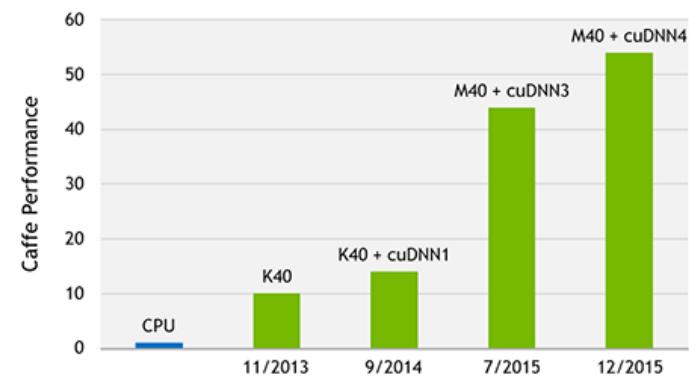
딥러닝 모델 학습에 대규모로 GPU를 적용하여
획기적 속도 향상됨을 보임
(DBN 알고리즘의 경우 70배)



ACCELERATE EVERY FRAMEWORK



50X BOOST IN DEEP LEARNING
IN 3 YEARS



*U. Washington, CMU, Stanford, TuSimple, NYU, Microsoft, U. Alberta, MIT, NYU Shanghai

AlexNet training throughput based on 20 iterations,
CPU: 1x E5-2680v3 12 Core 2.5GHz, 128GB System Memory, Ubuntu 14.04

원동력 2: 빅데이터

지능형 시스템 구성에 필요한 기초적 데이터가 얼마든지 제공된다

signal.
inary code with which the present
ls may take various forms, all of
e property that the symbol (or
representing each number (or sign
differs from the ones representi
er and the next higher number (or
atitude) in only one digit (or puls
Because this code in its primar;
built up from the conventional
a sort of reflection process and i
rms may in turn be built up fro
form in similar fashion, the c
, which has as yet no recognized
ated in this specification and
s the "reflected binary code."
a receiver station, reflected bina

글자(text)



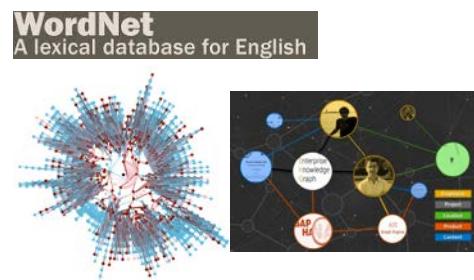
사용자 행동
(activity)



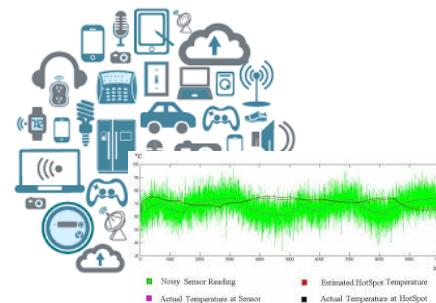
시각 정보(visual)



소리 정보(audio)



지식 체계
(knowledge graph)



센서 정보
(sensor)

원동력 2: 빅데이터

디지털 공간 전체가 AI를 보다 스마트하게 가르칠 수 있는 환경이 되고 있다.



대용량 DB
대용량 저장 장치



클라우드 컴퓨팅



이십여 년 이상 누적된
전세계 검색 기록



인터넷 사용 기록
(web cookies, online footprints, ...)



사물인터넷



Wikipedia 등의 무료 정보원

원동력 3: 획기적인 알고리즘

학계와 기업의 최고 인력이 새로운 ML/DL 기반 AI 해법을 쏟아내고 있다.



원동력 4: 개발 환경의 진화

[클라우드 + ML/DL] [공개 소프트웨어: 개방, 공유, 협업]



<https://github.com/Microsoft/CNTK>



<https://github.com/tensorflow/tensorflow>

theano

<https://github.com/Theano/Theano>



<https://github.com/torch/torch7>

Caffe

<https://github.com/BVLC/caffe>

클라우드 기반 머신러닝 기술

클라우드(Cloud) 환경에서 빅데이터(Big Data)를 활용하는 머신러닝(Machine Learning) 기술이 지능형 서비스/솔루션의 핵심 포맷으로 부각되고 있다.



Dashboard
Datasources
ML Models
Evaluations
Batch Predictions

Get Started with Amazon Machine Learning

Standard Setup

Go straight to creating

<https://aws.amazon.com/ko/blogs/aws/amazon-machine-learning-make-data-driven-decisions-at-scale/>

Microsoft Azure

Machine Learning

Powerful cloud-based machine learning service

Model your data
Train your models
Deploy your models

Get started now

<http://azure.microsoft.com/ko-kr/services/machine-learning/>



API

Integrate your data to add the following resources to your application:

- Cloud detection
- Opportunity analysis
- Diagnosis
- Suspicious activity detection
- And much more...

<https://cloud.google.com/prediction/docs>

IBM Bluemix

Instant Runtimes

App-centric runtime environments based on Cloud Foundry

Personality Insights
Alchemy API
Concept Expansion
Tradeoff Analysis

Data Management & Analytics

Cloudant
Insights for Twitter
Redis Cloud
MongoLab

Mobile & Internet of Things

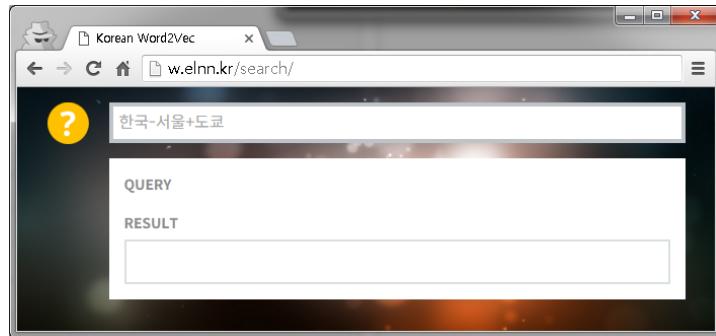
Push
Mobile Data
Internet of Things
Twilio

<http://www.ibm.com/cloud-computing/bluemix/>

공개 Deep Learning 개발도구 확산

개발도구명	Core language	Library type	Auto. gradient computation	CPU	GPU	Multi-GPU 지원	라이선스 종류	상업화 가능성
Theano	Python	Numerical computation	Flexible	✓	✓	x	BSD	✓ 저작권 고지
Google TensorFlow	Python	Numerical computation, deep learning models	Flexible	✓	✓	✓	Apache 2.0	✓ 저작권 고지, 라이센스 사본 첨부
Microsoft CNTK	C++	Numerical Computation	Flexible	✓	✓	✓	MIT	✓ 저작권 고지, 라이센스 사본 첨부
Samsung Veles	Python	Deep learning models	Restricted	✓	✓	Not specified	Apache 2.0	✓ 저작권 고지, 라이센스 사본 첨부
Caffe	C++	Deep learning models	Restricted	✓	✓	✓	BSD	✓ 저작권 고지
Torch	Lua	Numerical computation, deep learning models	Flexible	✓	✓	✓	BSD	✓ 저작권 고지
Keras	Python (using Theano or Tensorflow)	Deep learning models	Restricted	✓	✓	✓	MIT	✓ 저작권 고지, 라이센스 사본 첨부
Blocks	Python (using Theano)	Deep learning models	Restricted	✓	✓	x	MIT	✓ 저작권 고지, 라이센스 사본 첨부
Lasagne	Python (using Theano)	Deep learning models	Restricted	✓	✓	x	MIT	✓ 저작권 고지, 라이센스 사본 첨부

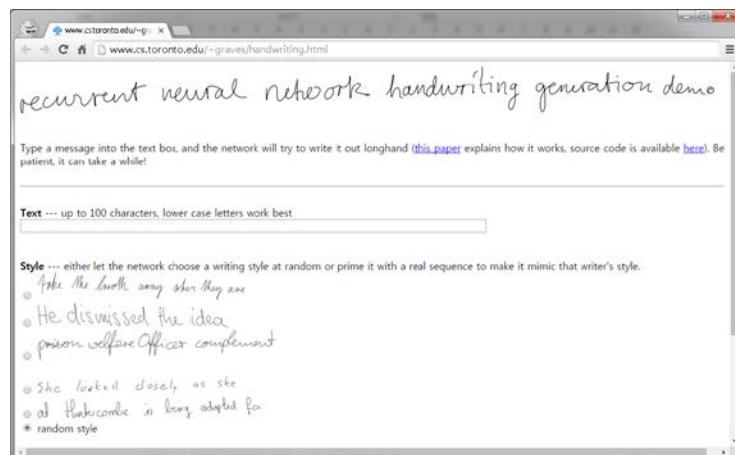
심심할 때 즐기는 딥러닝 데모



한글 Word2Vec: 한국-서울+도쿄=일본?
(<http://w.elnn.kr/search/>)

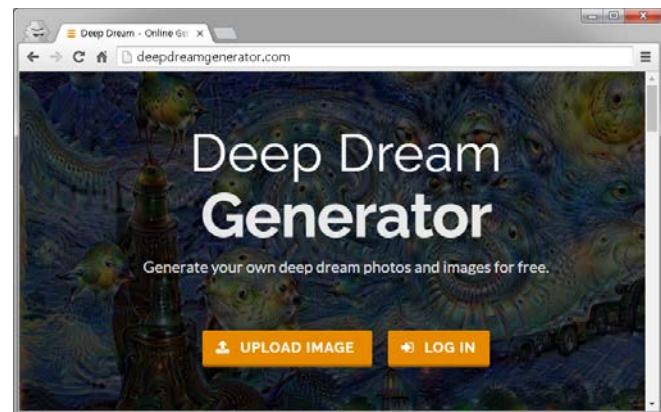


Eyescream Project: 컴퓨터가 자동으로 사진 생성
(<http://soumith.ch/eyescream/>)



필기체 문장 자동 생성

(<http://www.cs.toronto.edu/~graves/handwriting.html>)



Deep Dream: 사진을 추상화로 만들기
(<http://deepdreamgenerator.com/>)

설명으로 사진 찾기

<http://cs.stanford.edu/people/karpathy/deepimagesent/>

<http://cs.stanford.edu/people/karpathy/deepimagesent/rankingdemo/>

Query: A city street corner from a park view

[그림 테두리 색]

- 빨강: 부적절한 검색 결과
- 초록: 성공적 검색
- 노랑: top 5에 포함되지 않은 정답 사진



1.07 city
0.64 street
0.32 corner
-0.60 from
-0.23 park
0.36 view

top sentences for this image:

- (4.51) A view of a city street with motorcycles and taxi cabs.
- (3.37) a busy street with taxi's parked down the road.
- (3.34) many motorcycles parked on a city street near buildings
- (3.32) A bunch of cycling people riding bikes on a street.
- (3.16) A crowd of people riding bikes down a city street.

The Next Rembrandt

Bringing Rembrandt back to life.

By analysing his past work - a ‘new’ painting can be created.

Deep learning and various machine learning techniques are applied to get Rembrandt’s style in content, shapes, and lighting



Rembrandt van Rijn (1606-1669),
‘Master of light and shadow’



Deep Learning
Pattern Analysis
Facial Recognition
3D Printing
...



New Rembrandt-like painting
(3D printed, April 2016) ⁶⁸

The Next Rembrandt



Step 1: Gathering the data

- The entire collection of Rembrandt's work, pixel by pixel analysis
- High resolution 3D scan, digital files → upscaled by **deep learning** to maximize resolution and quality
- 163,263 Rembrandt painting fragments, 150 GB of digitally rendered graphics



Step 2: Determining the subject

- 346 paintings were studied and the 'quintessential' Rembrandt subject was determined as follows
- a portrait of a Caucasian male with facial hair, between the ages of thirty and forty, wearing black clothes with a white collar and a hat, facing to the right



Step 3: Generating the features - Mastering the style of Rembrandt into a 2D image

- Generating features based on data: use of geometry, composition, painting materials. **Facial recognition algorithms**
- Distances between the facial features. Affine transformation on the features + accurate placing
- 500+ Hours of rendering – a digital file true to Rembrandt's style in content, shapes, and lighting



Step 4: Bringing it to life – 3D printing

- Paintings are not just 2D – 3D texture, brushstrokes and layers of paint
- Creating a height map – **patterns of canvas surfaces and layers of paint**
- 13 layers of ink, 148 mega pixels

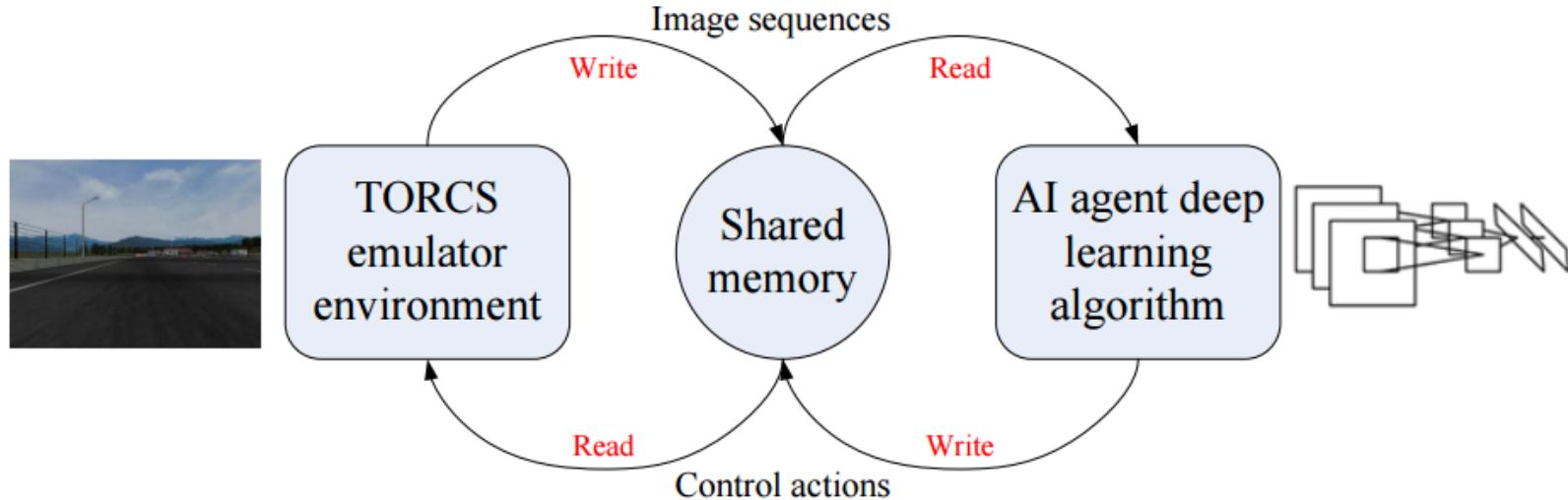
Steer Control

● Deep Learning for Steer Control (CNN)



Feature
Extraction

- angle: the angle between the car's heading and the tangent of the track;
- toMiddle: the distance between the center of the car and the center line of the track;
- min_dist: the distance between the car and the 1st preceding car;
- lane: the lane of the 1st preceding car



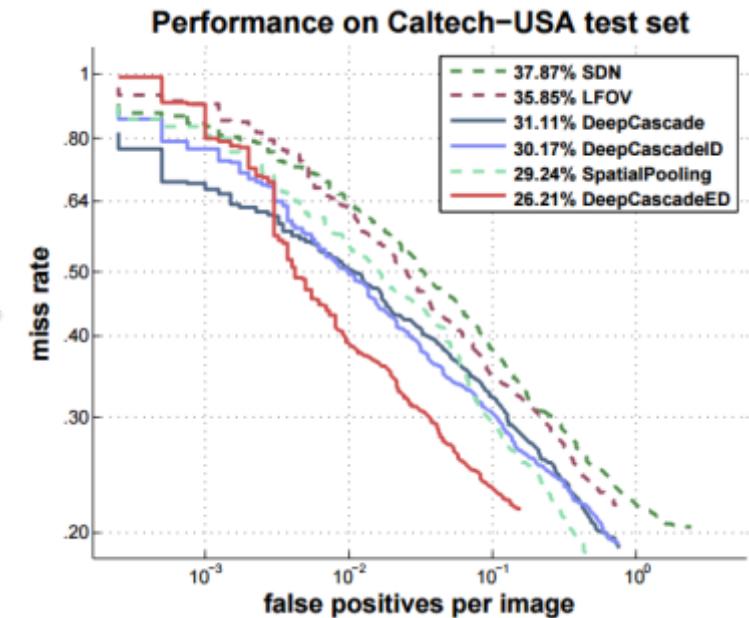
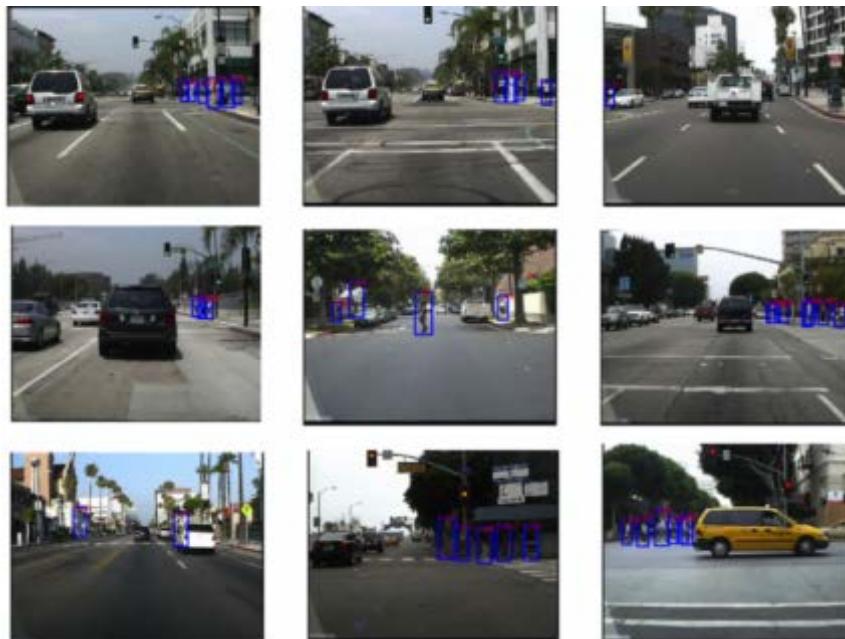
<http://www.princeton.edu/~alaink/Orf467F14/Deep%20Driving.pdf>

Pedestrian Recognition

● Deep Learning for Pedestrian Recognition (CNN)



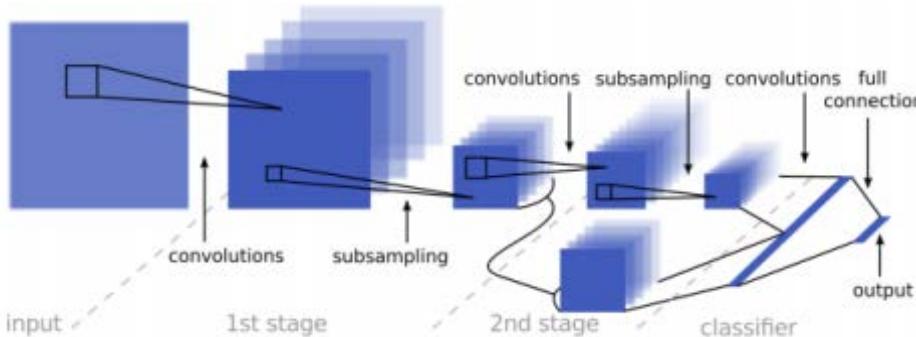
- DeepCascade Network
- Pretrained on Imagenet
- 5x5 , 1x1 convolution filter
- Caltech dataset



Angelova, Anelia, et al. "Real-Time Pedestrian Detection With Deep Network Cascades."

Traffic Sign Recognition

● Deep Learning for Traffic Sign Recognition (CNN)



- 2-Stage CNN
- Multi-scale Features
- Non-Linearity: Tanh()
- GTSRB traffic sign dataset



#	Team	Method	Accuracy
197	IDSIA	EBLearn 2LConvNet ms 108-108 + 100-feats CF classifier + No color	98.98%
196	IDSIA	cnn_hog3	98.98%
178	sermanet	EBLearn 2LConvNet ms 108-108	98.97%
		EBLearn 2LConvNet ms 108-200 + 100-feats CF classifier + No color	98.85%
		EBLearn 2LConvNet ms 108-200 + 100-feats CF classifier + No color + Random features + No jitter	97.33%

Sermanet, Pierre, and Yann LeCun. "Traffic sign recognition with multi-scale convolutional networks," IJCNN 2011.