

Technical Report: BI-17-001

딥러닝: 인공지능을 이끄는 첨단 기술

김병희, 장병탁 / 서울대학교 컴퓨터공학부 바이오지능연구실
{bhkim, btzhang}@bi.snu.ac.kr

초고: 2017-01-31

갱신: 2017-04-16

본고는 최근 학계와 산업계 전반에 걸쳐 주목받고 있는 딥러닝 기반의 인공지능의 기본 개념과 현황을 소개하고, 이러한 기술이 만들어내는 변화를 살펴볼 수 있도록 구성하였다. 1부에서는 기본 개념과 기술 동향을 소개하고 2부에서는 인공지능이 주도하는 변화를 넓게는 산업계로부터 우리의 일상 생활에까지의 범위에서 다각도로 살펴본다.

1. 기본 개념과 기술동향

가. 인공지능과 머신러닝, 딥러닝의 기본 개념

인공지능의 목표는 ‘사람처럼 생각하고 행동하는 기계/컴퓨터’를 만드는 것이다. 인공지능을 구현하고자 할 때, 지식과 규칙에서 시작해서 다양한 응용을 추구하는 top-down 방식과, 실세계의 데이터에서 시작해서 특정 응용문제를 해결하고 나아가 지식과 규칙의 발견까지도 기대하는 bottom-up 방식의 두 방향이 가능하다. 최근 주목을 받고 있는 머신러닝(machine learning)은 bottom-up의 대표적인 방식이고, 딥러닝(deep learning)은 머신러닝의 최신 기술/트렌드이다.

딥러닝을 포함한 머신러닝 과정에서는, 문제를 해결하는 기계/컴퓨터(에이전트라고 호칭)를 만들기 위해, 사람이 설정한 문제의 범주 안에서 데이터(대부분의 경우 사람이 모여서 정리)에서 시작하여 사람이 지정한 도구(알고리즘)를 이용하여 학습(learning)을 수행하고 문제를 풀게 된다. 아직은 이 과정의 대부분에 사람이 개입하여 가이드를 해주어야만 하는데, 데이터만 보고 에이전트를 만드는 핵심적인 과정의 대부분이 자동으로 이루어지도록 한 것이 인공지능 구현의 큰 발전이었다. 하지만 ‘스스로 학습한다’라고 표현을 하기에는 아직 많이 부족한 것이 인공지능의 현 주소이다.

‘에이전트를 만드는 핵심적인 과정을 자동화’하는 방법에 대해 조금 더 자세히 살펴보는 것이 현재의 인공지능을 이해하는데 도움이 된다. 딥러닝을 포함한 머신러닝의 기본

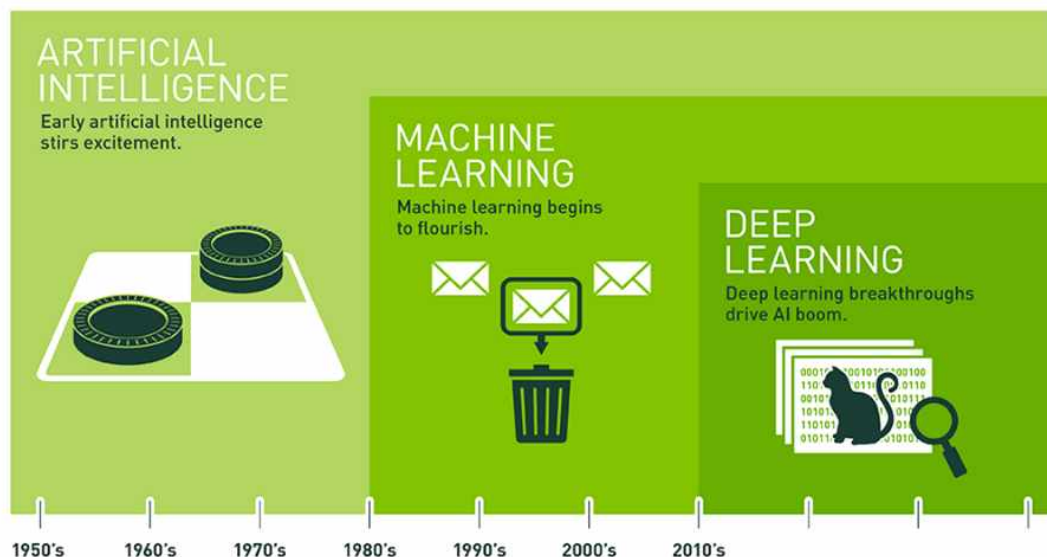


그림 1. 인공지능 기술은 1950년대에 태동하였으며, 1980년대 이후 머신러닝, 2010년 전후 딥러닝 기술이 인공지능을 구현하는 선도적 기술로 자리잡았다.

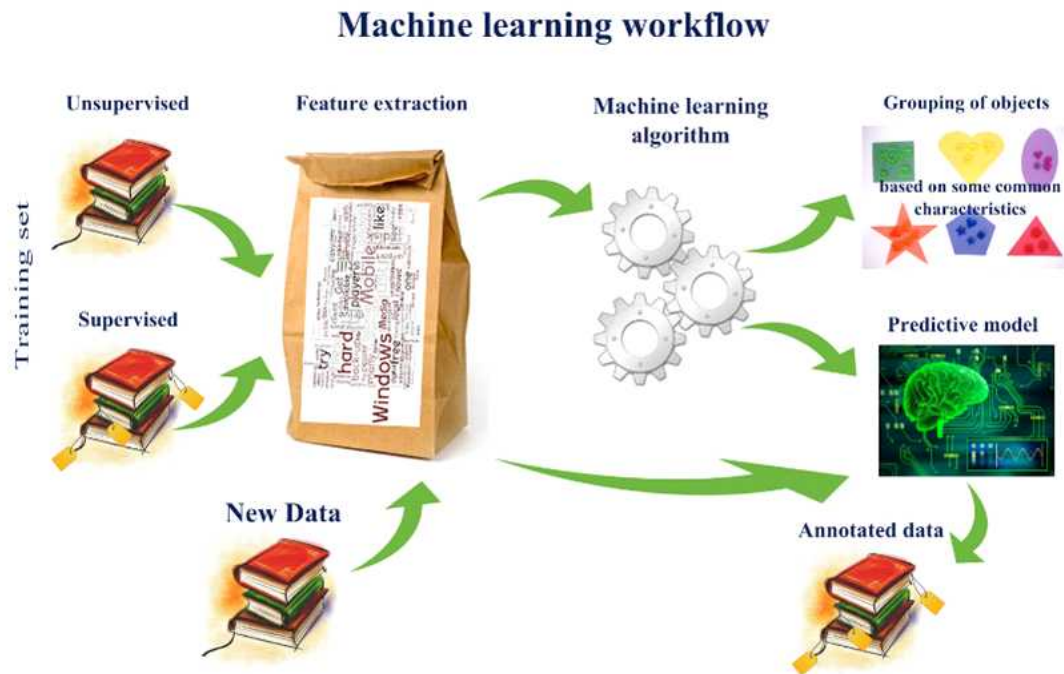


그림 2. 머신러닝의 핵심 절차는 [데이터 수집 및 정리] - [특징 추출(feature extraction)] - [머신러닝 알고리즘을 적용하여 학습 수행] - [학습한 모델을 적용]하는 네 단계로 정리할 수 있다. 학습의 종류는 데이터에 ‘표지/레이블’을 지정하는지 생략하는지 여부에 따라 비지도(unsupervised)와 지도(supervised) 학습으로 구분된다.

과정은 학습의 목표와 형태에 따라 크게 지도학습(supervised learning)과 비지도학습(unsupervised learning)으로 구분할 수 있다. 머신러닝 과정에서 가장 먼저 하는 일은, 연구자가 풀고자 하는 문제를 파악하고 명확하게 문제를 정의(define)하는 것입니다. 풀려는 문제가 인물사진을 보고 남녀를 구분하는 경우와 같이 입력받은 데이터(x)의 종류(y)를 답하는 문제로 정의 가능하면 x와 y를 이어주는 함수 $y=f(x)$ 를 데이터에서 찾는 지도학습 방식의 머신러닝 도구(알고리즘)를 사용하게 된다. 풀려는 문제가 데이터(x)에 내재된 규칙, 특성(예를 들면 인물 사진이 많이 있을 때 얼굴형태가 비슷한 사람들의 사진끼리 모으기)을 발견하려 하는 것이라면 비지도학습 방식으로 문제를 풀게 된다.

두 경우 모두 ‘문제를 푸는 데 필요한 데이터(x)’를 적어도 ‘컴퓨터가 알아볼 수 있는 형태’로 정리를 해서 입력해야 하고, 더 나아가서는 ‘문제를 잘 푸는데 도움이 되는 형태’로 변형/요약하는 과정이 필요하다. 이 ‘문제를 잘 푸는데 도움이 되는 형태로 변형/요약’하는 과정을 보통 ‘특징 추출(feature extraction)’이라고 한다. 문제에 따라, 연구자가 추구하는 해법에 따라 추출하는 특징(feature)은 얼마든지 달라지지만, 이러한 특징을 입력으로 하는 학습 과정에는 공통의 머신러닝 도구(알고리즘)를 쓸 수 있다. 지도학습의 경우에는 대표적인 도구로 결정트리, 인공신경망, 지지벡터기계(SVM) 등이 있고, 비지도학습 중 경우 비슷한 데이터를 모으는 문제의 경우는(‘군집화(clustering)’라고 부르는 비지

도학습 문제의 한 형태) 계층적 군집화, k-평균 등의 도구가 있다.

기존에는 문제의 종류에 따라 해당 분야의 전문가가 특징 추출을 하는 방법을 고안해내어야 했으나, 딥러닝에서는 ‘특징 추출’ 과정도 학습 과정에 포함을 하여 데이터만 보고도 문제를 표현하고 푸는 데 좋은 특징을 학습을 통해 찾는다. 그래서, 요즘 딥러닝을 이용해 지도학습 또는 비지도학습을 할 때는 주변에서 구할 수 있는 형태의 데이터(문서의 경우 단어 또는 알파벳, 사진의 경우 픽셀의 집합으로 표현된 형태의 컴퓨터 파일, 소리의 경우 spectrogram이라 부르는 가장 기본적인 파일)를 그대로 알고리즘의 입력값으로 사용한다. 이러한 학습 과정을 표현력 학습(representation learning)이라 하며, 데이터와 알고리즘이 충분히 준비된 문제에 대해 입력과 출력 사이의 전 과정을 기계가 학습을 통해 자동적으로 구축하는 종단학습(end-to-end learning) 해법이 등장하고 있다.

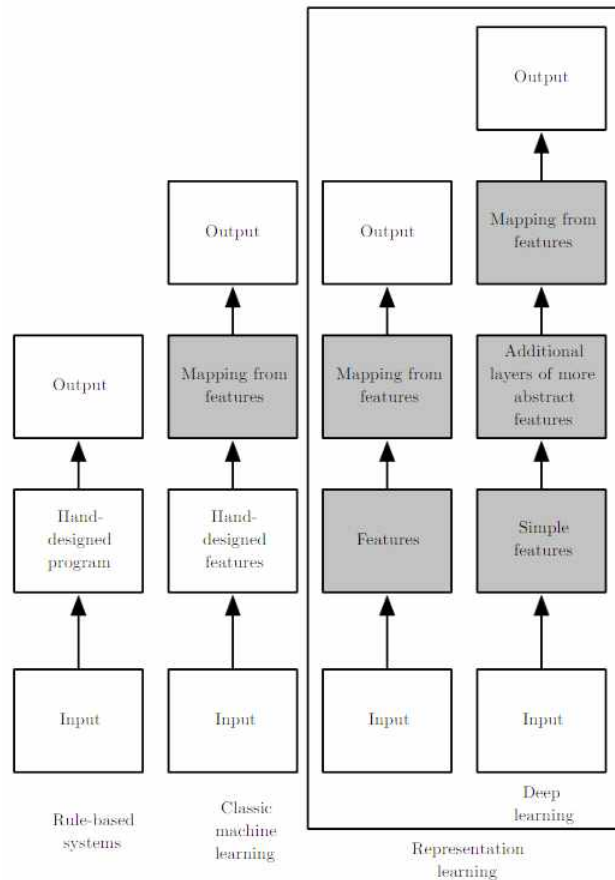


그림 3. 인공지능을 구현하는 대표적 방법론의 발전과 구성 비교. 회색 상자는 데이터에서 학습하여 구성할 수 있는 요소이다. 규칙 기반 시스템과 기존의 머신러닝과 대조적으로 딥러닝은 데이터에서 다양한 개념 수준의 특징값을 학습하여 문제를 스스로 해결하는 표현력 학습을 추구한다.

종단학습의 개념을 간략히 표현하자면, 학습에 필요한 데이터를 거의 원본 그대로(raw)



그림 4. 자율주행 문제 중 핸들 조정에 대한 종단학습의 예시. 머신러닝과 딥러닝이 발전하면서 특정 문제에 대한 해법을 단편적 모듈의 조합으로 구성하던 방법이 입력과 출력 사이의 전 과정을 단일한 학습 모듈로 해결하는 종단학습이 가능해지고 있다.

입력하여 지정한 형태의 출력이 산출되도록 학습하는 과정이다. 딥러닝이 등장하기 이전에도, 문서의 종류를 분류하거나 사진 속에 있는 물체를 맞추는 전형적인 분류(classification) 문제의 해법은 종단학습으로 얻을 수 있었다. 딥러닝 기술이 발전함에 따라 종단학습으로 해법을 얻을 수 있는 문제의 종류와 복잡도가 크게 증가하였다. 예를 들면, 자율주행 자동차에 들어가는 필수 요소 중 하나로서 전방 카메라의 입력 정보를 기반으로 한 실시간 자동 핸들 조정 방법을 살펴보자. 전통적인 방법에서는 카메라 이미지에서 자동차를 검출하는 모듈과 보행자를 검출하는 모듈을 따로 구성한 후 두 정보를 취합하여 주행 계획을 세우고 최종적으로 핸들 방향을 조정하는 복잡한 과정을 구성하였다. 그러나 딥러닝을 중심으로 하는 종단학습 해법에서는 카메라 이미지만을 입력하여 바로 핸들 방향 조정을 출력하는 것이 가능하다.

시각, 언어, 음성지능을 통합한 흥미로운 종단학습 해법은 2010년 중반을 전후하여 급격히 개발되고 있다. 대표적인 사례를 다음 표에 정리하였다. 이러한 해법은 학습에 필요한 대규모 데이터가 있어야만 구현할 수 있다. 이 중 일부는 ‘1-라’ 절에서 사례를 중심으로 조금 더 살펴볼 수 있다.

표 1. 딥러닝을 기반으로 한 ‘종단학습(end-to-end learning)’ 이 가능하게 된 문제의 사례. 기존의 해법에서는 입력과 출력 사이에 전문가의 경험을 바탕으로 구성한 여러 중간 단계가 필요하였으나, 종단학습을 통한 해법은 중간 과정이 한 종류의 딥러닝 모델 학습으로 얻게 된다. 이러한 해법이 가능하게 된 핵심 요인으로 대규모 데이터에서 개념적, 고차적 지식을 자동으로 추출하는 ‘표현력 학습(representation learning)’ 기법의 발전을 들 수 있다.

문제의 종류	입력	출력	설명
사진 태그 자동 생성 (image tagging)	사진	사진 설명 태그 집합	사진에 포함되어 있는 물체와 환경 등을 표현하는 단어의 집합을 생성
사진 설명 자동 생성 (image captioning)	사진	사진 설명 문장	사진을 설명하는 자연스러운 문장을 생성
동영상 자동 설명 (movie captioning)	동영상	동영상 설명 문장	동영상을 설명하는 자연스러운 문장을 생성
음성에서 글자 인식 (speech to text)	음성	스크립트	음성 인식. 특히, 하나의 딥러닝 모델로 영어와 중국어 동시 인식
자동 번역 (machine translation)	언어 #1 문서	언어 #2 문서	영어-프랑스어 등과 같이 두 언어 간의 번역을 위한 딥러닝 모델. 제한된 언어쌍 만으로 학습한 딥러닝 모델이 학습하지 않은 언어쌍 번역에도 활용된다.
사진/그림 자동 생성 (image synthesis)	표지, 태그	사진/그림	학습한 물체의 종류 중 하나를 지정하면, 학습 과정에서 본 적이 없지만 같은 표지를 달 수 있는 자연스러운 사진/그림을 생성

딥러닝: 인공지능을 이끄는 첨단 기술

기본적인 요소는 물체를 인식하는 능력이다. 서비스 측면에서는 특히 얼굴 인식 기능이 필수적이다.

머신러닝 연구자들은 대규모의 데이터베이스를 구축하고 공개하여 경쟁적으로 성능 향상을 도모하여왔다. 이러한 데이터베이스 중 시각지능 연구의 핵심 도구를 기존의 컴퓨터 비전 기법 춘추전국 시대에서 딥러닝 시대로 바꾸는데 기여한 ImageNet 데이터베이스¹⁾가 2009년에 구축되었다. ImageNet의 사진 데이터에서 물체 인식 성능을 겨루는 대회가 2010년부터 매년 개최되었는데, 2012년도에 기존 컴퓨터 비전 연구자들을 경악시킨 결과가 나왔다. 컨볼루션망(CNN)이라는 딥러닝 기법이 다른 모든 컴퓨터 비전 기반 팀의 결과보다 월등히 우수한 성능으로 1위를 차지하였으며, 2013년 이후 매년 획기적인 딥러닝 기법이 1위 성능을 독식하는 한편 모든 대회 참가팀들이 딥러닝 기법을 적용하게 되었다.

얼굴 인식 분야에서도 딥러닝의 기여로 큰 도약이 계속되고 있다. 음성인식 사례와 유사하게, 1990년대 이후 미국 표준연구소에서 관리하는 얼굴 인식 대회와 관련 데이터셋²⁾을 중심으로 1993년부터 2011년까지 매 2년마다 오류율이 반으로 줄어드는 발전이 있었다. 딥러닝이 등장하며 오류율 감소의 가속화와 함께, 다양한 상황에서의 얼굴 인식과 인물 인식, 동일 얼굴 탐색 등의 다각화된 기술 발전이 이어지고 있다.

세 번째로 언어지능에서 딥러닝이 이루어낸 도약을 살펴보자. 음성과 시각지능 사례와 비교해볼 때 언어는 인류의 경험과 지식을 직접적으로 표현하는 단계의 정보라는 점에서 언어지능의 영향력은 매우 광범위하다. 문서 자동 생성, 의미 수준 단어 표현, 자동 번역

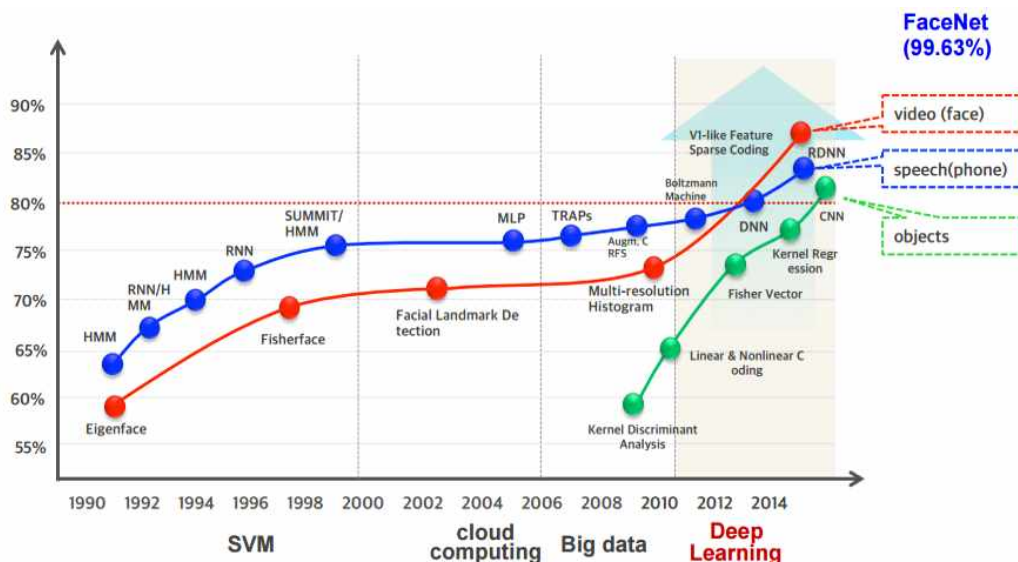


그림 6. 딥러닝이 이루어낸 시각지능의 도약 사례. 이미지에 있는 사물(object) 인식 정확도와 비디오 상에서의 얼굴 인식 정확도가 딥러닝의 도입으로 크게 향상되었다. (그림 출처: H. Choi, Recent Advances in Recurrent Neural Networks, 2015.04)

1) <http://image-net.org/>

2) <https://www.nist.gov/programs-projects/face-projects>

등과 같은 언어 단계의 지능뿐만 아니라 음성지능 및 시각지능과 결합된 사례, 예를 들면 음성 자동 번역, 사진 설명 자동 생성과 같은 기술이 딥러닝을 기반으로 발전하고 있다.

딥러닝을 핵심으로 한 문서 자동 생성 기술의 수준은 현재 보고서, 기사 등과 같은 기술적인 문서뿐만 아니라 시와 소설 등의 창작의 영역까지 넘보고 있다. 기본적인 딥러닝 기술만으로도 다양한 실질적 문서 생성을 시험해볼 수 있다. ImageNet 대회를 주관³⁾, 관리하고 2017년 1월 현재 OpenAI에서 재직 중인 Andrej Karpathy는 2015년도의 블로그 글⁴⁾에서, 알파벳 단위로 학습한 순환신경망을 이용하여 컴퓨터가 자동으로 생성한 다양한 문서를 소개하였다. 학습을 통해 자동 생성한 문서의 종류는 ‘셰익스피어의 희곡’, ‘위키피디아 문서’, ‘LaTeX 문서’, ‘Linux 소스 코드’를 포함하며, 이후 ‘가상 오바마’, ‘가상 트럼프’와 같은 수많은 확장 사례가 소셜넷에서 회자되고 있다.

다음으로 살펴볼 딥러닝 언어지능 사례는 의미 수준 단어 표현이다. 신경망을 이용하여 단어의 벡터(vector) 표현을 학습한 결과, 단어 간의 의미 수준에서의 연산, 문법적 관계, 단어의 조합을 통한 의미적 구문 생성 등과 같은 ‘언어를 이해하는’ 듯한 컴퓨터 프로그램을 구현하는 것이 가능해졌다.

마지막으로 보다 최근에 나타난 행동지능 구현의 도약을 살펴보겠다. 행동지능 도약의 배경에는 머신러닝의 기본 과정 중 하나인 강화학습(reinforcement learning)에 딥러닝이 도입이 있었다. 2015년 초에 Nature 지에 영국 스타트업 회사 DeepMind 연구진의 논문이 게재되었다. 이 논문에서는 Atari 사의 게임 화면만으로 학습하여 사람보다 게임을 더 잘 하게 된 사례가 소개되었으며, 이러한 딥 강화학습 기술을 높이 산 Google이 DeepMind를 고가에 인수하는 직접적인 계기가 되었다. Google DeepMind는 불과 1년 후에, AlphaGo를

- **semantics:** $\text{vec(Beijing)} - \text{vec(China)} + \text{vec(Japan)} = \text{vec(Tokyo)}$
- **syntactic :** $\text{vec(quick)} - \text{vec(quickly)} + \text{vec(slowly)} = \text{vec(slow)}$

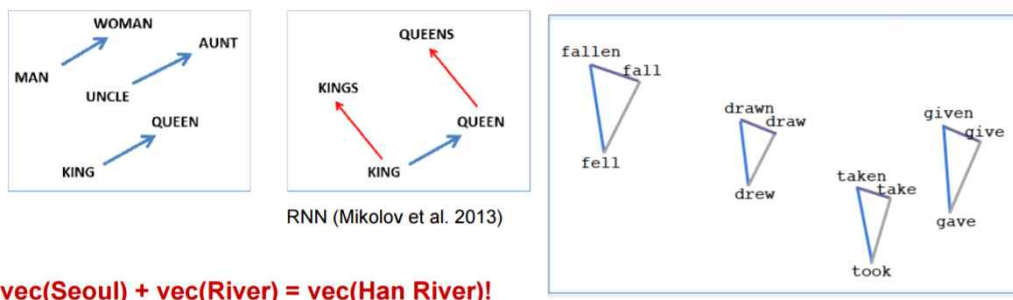


그림 7. 딥러닝이 이루어낸 언어지능 사례. 의미 수준의 단어 표현과 연산이 가능해졌다. 이러한 기술을 word2vec으로 통칭하며, sentence2vec, doc2vec 등과 같은 확장 연구로 이어지고 있다. (그림 출처: H. Choi, Recent Advances in Recurrent Neural Networks, 2015.04)

3) 스탠포드대의 Fei-Fei 교수진과 프린스턴대의 Li 교수진이 주도

4) <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

딥러닝: 인공지능을 이끄는 첨단 기술

통해 바둑의 최고수를 4:1로 꺾는 대 파란을 일으켰으며, 이는 전문가들의 예상보다 적어도 10년은 앞서 이루어낸 결과였기에 그 파급효과는 더욱 컸다.

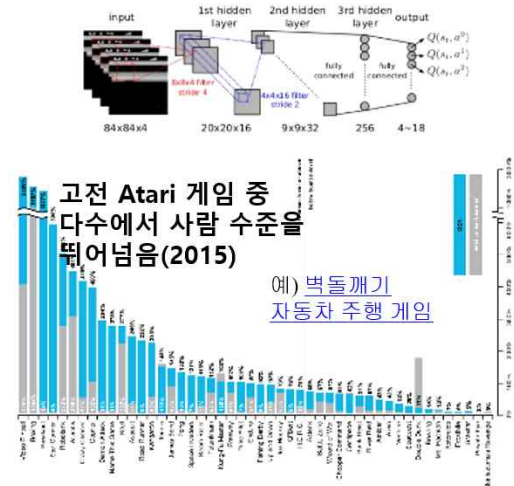


그림 8. 딥러닝이 이루어낸 행동지능의 도약 사례 1. 딥 강화학습을 통해 컴퓨터 게임의 최고수를 구현



그림 9. 딥러닝이 이루어낸 행동지능의 도약 사례 2. 딥 강화학습을 발전시켜 바둑의 최강자를 이기는 이벤트를 일구어냄

다. 딥러닝의 구현 방법

인공지능 연구자들이 추구하는 가장 이상적인 인공지능 모델은 바로 인간의 두뇌이다. 뇌과학, 인지과학, 신경과학 등의 연구를 통해 밝혀진 뇌의 구조와 작동 원리는 인공지능 연구자들에게 다양한 영감을 불러일으키는 좋은 정보원이 되어왔다. 1940년대 초에 생물 신경세포를 모사하여 구성한 인공 신경세포(퍼셉트론)에 대한 연구가 시작되었으며, 이러한 신경세포로 구성된 신경망을 인공적으로 구현한 인공신경망은 1980년대 중반부터 본격적으로 인공지능 구현을 위한 도구로 사용되기 시작하였다. 뇌의 시각 기능에 관한 1960~70년대의 뇌과학 연구는 1980년대의 컨볼루션신경망(CNN)의 발명과 2010년대 중반의 ‘주의 집중(attention)’ 기반 딥러닝 기법 개발의 토대로 이어졌다.

딥러닝의 개념을 딥러닝 기법의 구성과 동작 방식을 기준으로 간략히 표현하자면, ‘인공신경망의 다층 구조가 깊어진 모델’로 설명할 수 있다. 정보가 인공신경망에 입력되어 여러 층을 거치며 처리됨에 따라, 마치 눈을 통해 들어온 정보가 여러 층을 거치며 물체로 인식되는 것과 같은 방식으로, 여러 단계의 정보 표현과 개념, 지식이 학습될 것을 기대하는 것이다.

깊은 인공신경망을 구성하는 기본적인 방법은 크게 세 가지가 있다. 가장 기본적인 구성은 퍼셉트론이라 불리는 기본적인 인공 신경세포의 다층 구조를 입력과 출력 사이에

딥러닝: 인공지능을 이끄는 첨단 기술

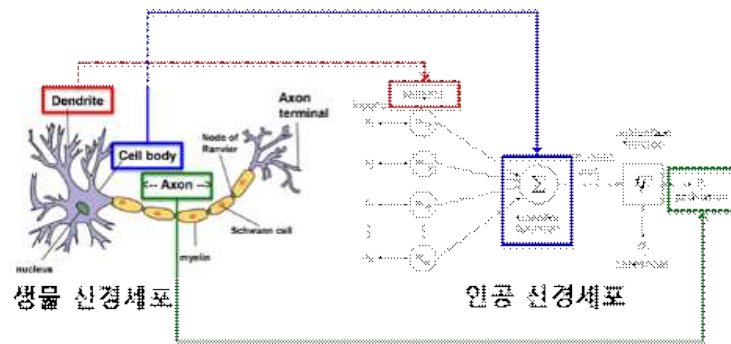


그림 10. 생물의 신경세포와 신경망을 본뜬 인공 신경세포와 인공신경망이 딥러닝 모델의 가장 기본적인 구성요소이다.

두는 다층 퍼셉트론(multilayer perceptron, MLP)이다. 다층 퍼셉트론의 이웃하는 층 간 신경세포는 모두 연결된(fully connected) 구조를 가지며, 데이터에서 학습하는 과정에서 연결선의 가중치가 전체적으로 조절이 되면서 지능을 갖추게 된다. 두 번째 기본적인 구조는 사진과 같이 공간적인 구성을 가진 정보를 처리할 때 적합한 컨볼루션망(convolutional neural network, CNN)이다. 컨볼루션망은 이웃한 공간상에서 나타나는 패턴을 학습하고 요약하는 다층 구조의 망으로 구성하며, 복잡한 문제일수록 다층 구조를 깊게 구성하고 대규모의 데이터를 통해 학습하게 된다. 세 번째의 구조는 시간에 따라 변하거나 순서가 중요한 데이터에서의 학습을 위한 순환신경망(recurrent neural network, RNN)이다. 순환신경망에서는 시간 또는 단계를 따라 펼쳐지는 망의 구조로 인해 기본적인 딥 구조가 형성되며, MLP와 CNN의 경우와 같이 다층 구조를 구성하여 다양한 스케일의 순서 정보를 학습하는 것이 가능하다.

Samoyed (16); Papillon (5.7); Pomeranian (2.7); Arctic fox (1.0); Eskimo dog (0.6); white wolf (0.4); Siberian husky (0.4)

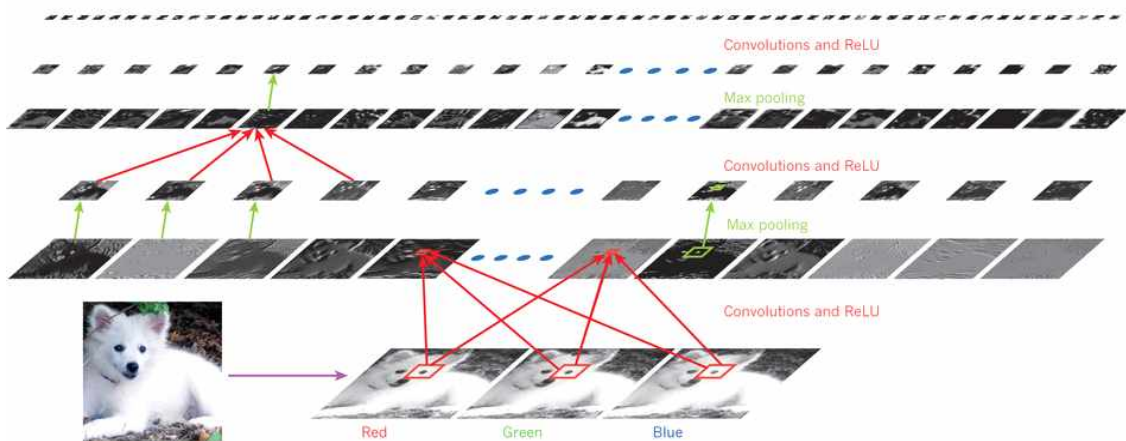


그림 11. 컬러 이미지를 위한 컨볼루션 신경망(CNN)의 작동방식. 계층별로 주어진 다수의 필터에 대해 컨볼루션과 ReLU 연산, 풀링 연산을 반복하면서 계층을 따라 올라간 후, 최종 계층에서 이미지에 나타난 물체의 후보군에 대한 가중치를 출력한다.

딥러닝: 인공지능을 이끄는 첨단 기술

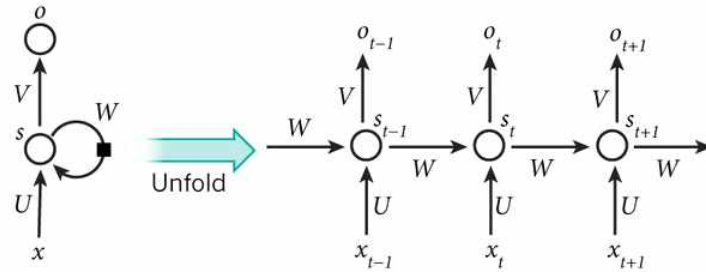


그림 12. 순환신경망 (recurrent neural network)의 개념도. 시간 또는 순서의 방향을 따라 심층 구조가 구성된다.

다음 그림에서는 인공신경망 구성에서 실제로 적용되는 다양한 사례를 종합적으로 살펴볼 수 있다.

이와 같은 다양한 방식으로 문제에 적합한 망의 구조를 정하고, 문제의 사례를 담은 대규모 데이터가 준비되면, 데이터를 살펴보고 반복학습을 통해 문제의 해법을 스스로 발견

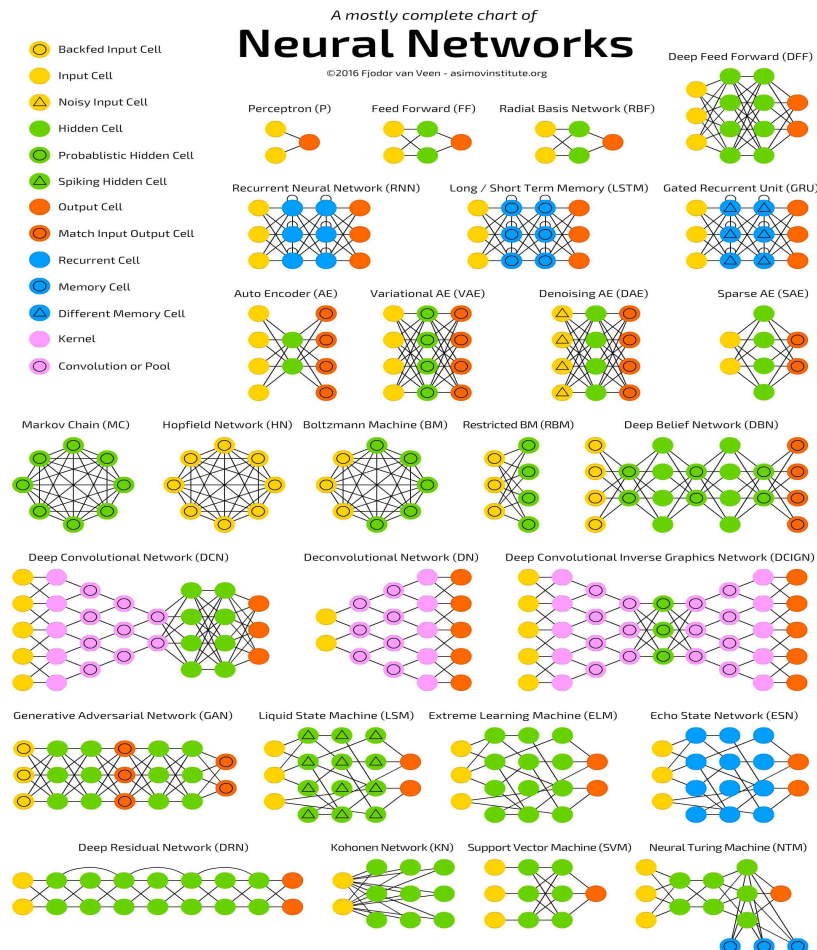


그림 13. 딥러닝의 핵심 구성 요소인 인공신경망의 다양한 사례. 인공뉴런은 원으로 표현이 되었으며, 그림에 표기된 바와 같이 다양한 종류의 인공뉴런과, 이들의 다양한 연결 방법이 가능하다.

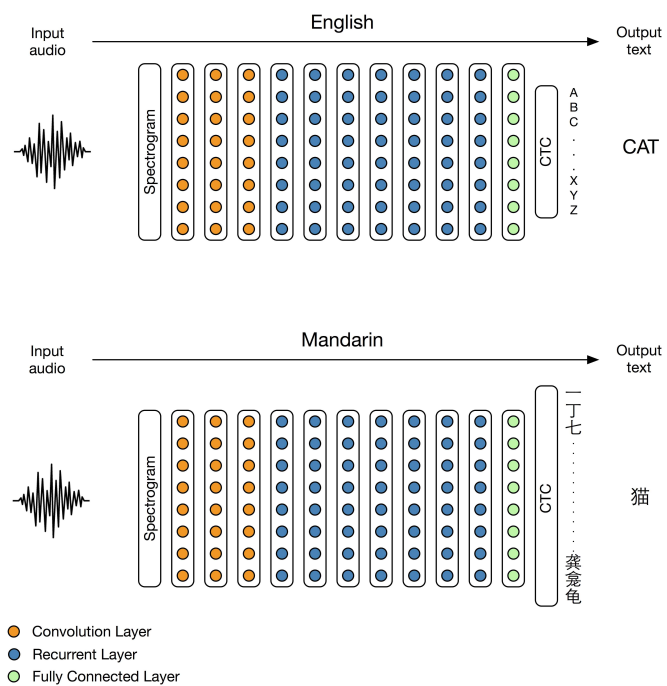
해나가는 ‘학습’ 단계와, 학습 결과를 점검하기 위한 ‘테스트’ 단계를 거치게 된다. 학습 과정에서는 학습 중인 딥러닝 모델이 내놓는 답의 오류를 층을 따라 순차적으로 줄이기 위해 오류 정보를 출력층에서 입력층 방향으로 ‘역전파(backpropagation)’를 하면서 신경망 내부의 연결선의 가중치를 조금씩 조정을 한다. 학습의 성공도를 점수화한 기준에 따라 반복학습을 통해 점수를 최대화하는 과정이 공통적으로 적용된다. 학습된 딥러닝 모델을 실전 테스트하는 단계에서는 학습에 반영하지 않은 데이터를 입력으로 주고 적절한 출력을 내놓는지를 정량적으로 평가하게 된다.

보통 학습의 궁극적 목표는 제한된 학습데이터만으로 학습한 모델이 풀고자 하는 다양한 실전 상황에서도 성공적이고 안정적으로 동작하는 것이다. 이러한 목표를 위해 보통 테스트 성능이 좋은 모델을 탐색하는 과정을 거치게 되며, 이 과정에서 수많은 계산 과정과 데이터가 필수적이다. 딥러닝 발전의 핵심 원동력은 CNN, RNN 등과 같은 새로운 알고리즘뿐만 아니라 이러한 알고리즘을 뒷받쳐줄 컴퓨팅 파워와 빅데이터가 포함된다.

라. 딥러닝의 최신 응용 사례

3절까지의 내용으로 딥러닝의 최신 응용 사례를 살펴보고 이해할 준비가 되었다.

먼저 음성지능의 사례로서 중단 학습 음성인식을 살펴보자. 다음 그림은 중국의 Google이라 불리는 Baidu의 실리콘밸리 연구진이 2015년 말에 논문으로 발표한 DeepSpeech 2의 소개이다. 입력 음성 신호와 인식한 글자 출력단 사이에 컨볼루션층, 순환신경망, 완전연결망이 각각 3층, 7층, 1층씩 순서대로 배치된 것을 확인할 수 있다. 연구진은 이 모델로 영어와 중국어를 모두 높은 성능으로 인식이 가능함을 보였다.



시각지능의 최신 사례로는 이미 지 자동 생성 사례를 살펴보겠다. 다음 그림은 2017년도 1월 현재 딥러닝 연구자들이 가장 주목하는

그림 14. 딥러닝을 적용한 음성지능의 최신 사례. Baidu에서 2015년 말에 발표한 DeepSpeech 2는 단일한 딥러닝 모델로 영어와 중국어 음성을 인식할 수 있음을 보였다.

모델 중 하나인 생성대립넷(GAN)을 이용하여 얼굴을 생성하고, 의미적 수준에서 얼굴간

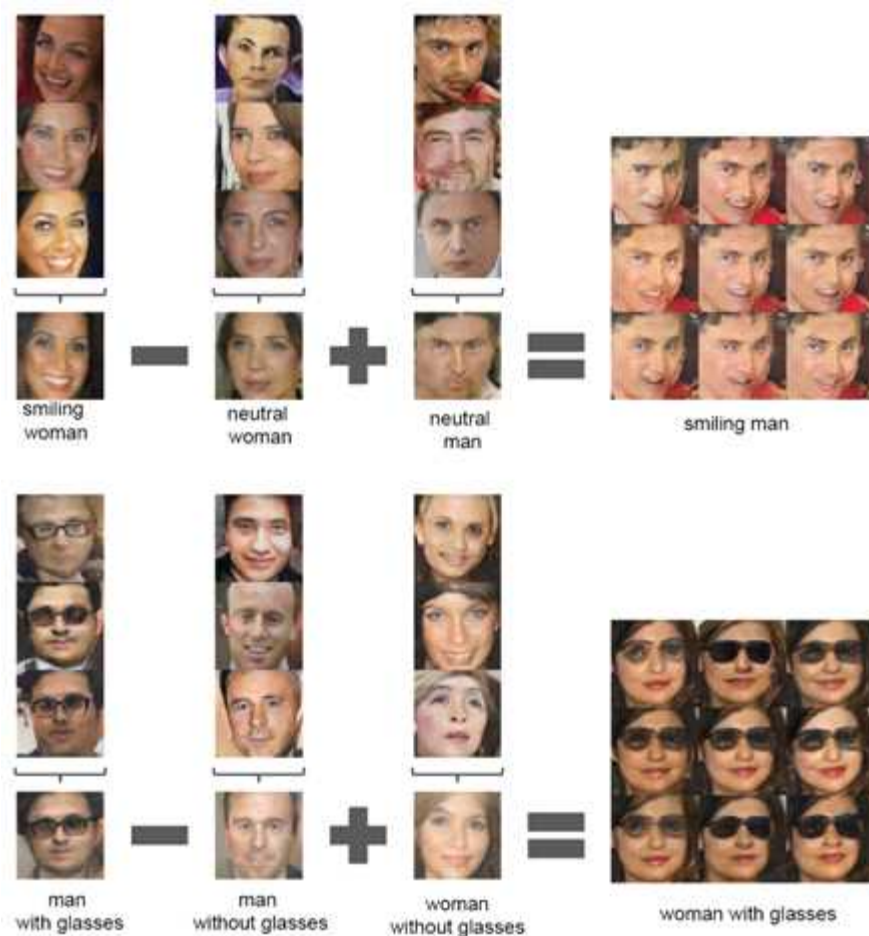


그림 15 생성대립넷(GAN)을 이용한 사람 얼굴 자동 생성 및 의미적 수준에서의 연산 사례 (Radford et al., 2015)

연상적 연산을 보인 사례이다. 컴퓨터가 생성한 얼굴의 모양은 완벽하지는 않으나 충분히 사람임을 인지할 수준이 되며, 나아가 얼굴에 나타난 고차적 정보의 가감이 가능함을 볼 수 있다.

언어지능의 사례로는 자동번역(machine translation) 시스템이 주목할 만하다. Google에 의해 2000년대 중반부터 널리 사용된 통계 기반의 기계번역은 딥러닝 기술에 힘입어 진일보하였다. 기존의 통계 기반의 기계 번역에서 사용된 메모리의 일부만 사용하면서도 깊은 표현의 학습을 통해 성능을 향상한 신경 기계 번역(Neural Machine Translation) 기법이 발전하였으며, 2016년도 하반기에 발표된 구글의 새로운 신경 기계 번역 시스템은(다음 그림), 전문가의 번역에 버금갈 정도의 뛰어난 자동 번역 성능을 보인 점과, 일부 언어 쌍으로 학습한 사례를 학습하지 않은 언어쌍에도 적용 가능한 점에서 또다른 언어지능의 마일스톤을 정립하였다. 해당 시스템은 8개의 인코더, 8개의 디코더 레이어로 구성된 잔차(residual) 연결을 갖는 깊은 LSTM 네트워크이며, 인코더와 디코더는 주의 기작(attention)을 반영하여 연결된다.

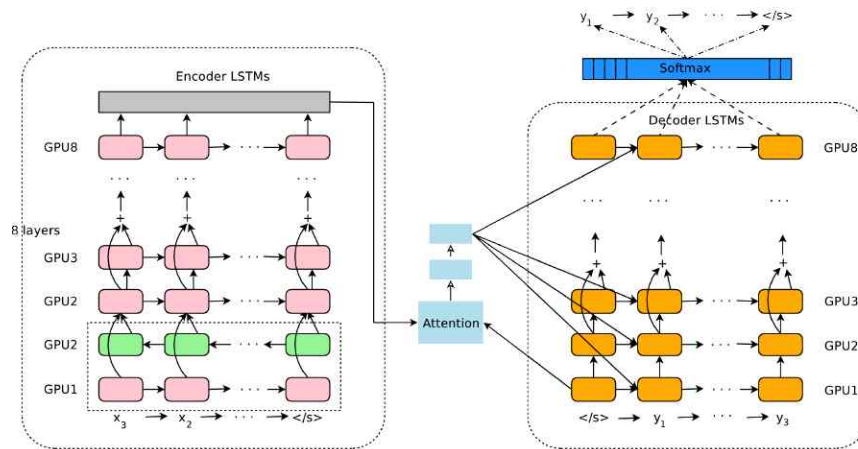


그림 16. 딥러닝을 이용한 기계 번역의 최신 사례, 구글의 신경 기계 번역 시스템(2016)

아래 그림에서는 언어지능 구현의 핵심 연구분야인 자연언어처리(natural language processing, NLP) 분야에서 딥러닝의 도입으로 인해 야기되는 해법의 변화를 소개하고 있다. 글의 종류 구분, 글에 나타난 감성 분석, 글이 다루는 주제 등과 같은 다양한 문제 해결을 위해 기존 NLP 연구자들은 특성이 다른 수많은 알고리즘과 프로세스를 만들어왔지만, 최근에는 딥러닝 기법을 공통적으로 적용하는 것이 표준적 방법으로 자리잡았다.

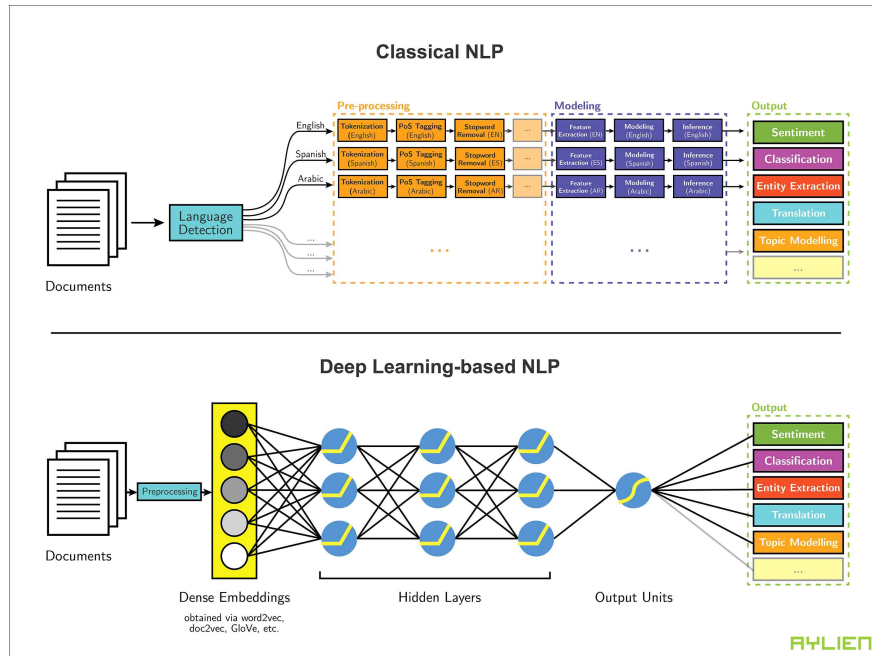


그림 17. 언어지능의 핵심인 자연언어처리(NLP) 기술에서 기존에는 출력에 따라 다양한 과정을 조합한 수많은 절차가 고안되었지만(상단), 최근에는 단일한 딥러닝 방법론으로 다양한 문제의 종단학습 기반 해법을 구성하는 것이 일반적인 해법으로 정립되고 있다(하단). (출처: <http://blog.aylien.com/leveraging-deep-learning-for-multilingual/>, 2017-1-21)

딥러닝: 인공지능을 이끄는 첨단 기술

마지막으로 다중 지능 사례를 하나 살펴보도록 하자. 다음 그림은 사진을 설명하는 문장을 자동으로 생성하는 사례이다. 딥러닝 모델을 조합하여 시각지능과 언어지능을 모두 보인 경우다. 컨볼루션망(CNN)으로 사진의 정보를 분석한 후, 이 정보를 순환신경망(RNN)에 전달하여 글자 또는 단어를 순서대로 출력하는 방식으로 설명 문장을 생성한다. 특히, 단어의 선정은 입력된 사진에서 해당 단어에 해당하는 물체에 대해 주의 집중(attention)을 함으로써 이루어지는 점에서 딥러닝 이전의 인공지능과 비교하여 큰 도약이 있었음을 확인할 수 있다.

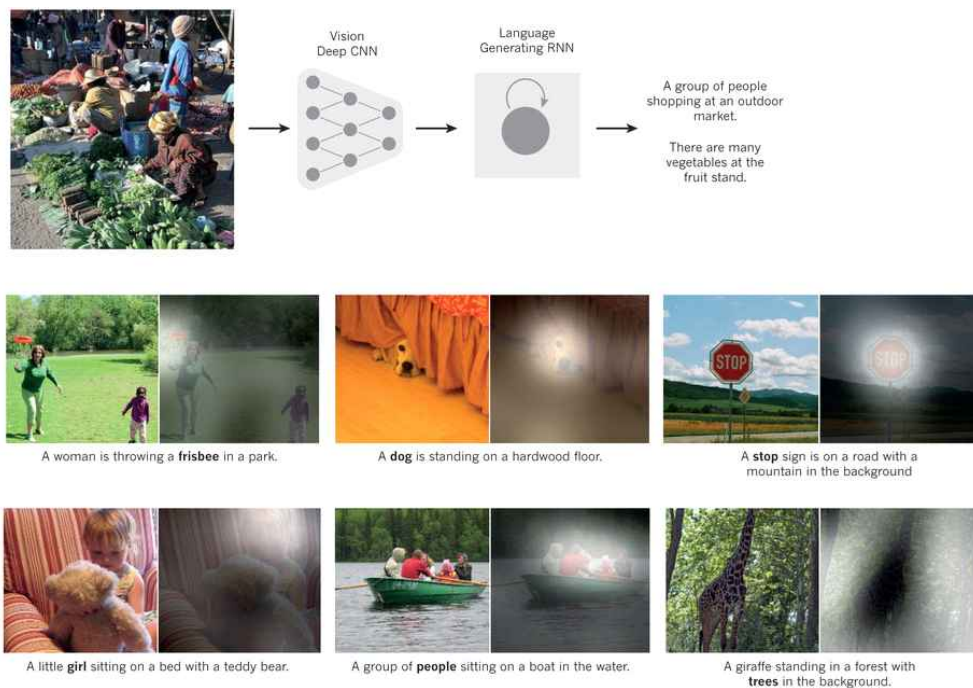


그림 18. 딥러닝 모델을 조합하여 사진을 설명하는 문장을 자동으로 생성하는 사례(2014, 2015)

2. 변화를 주도하는 인공지능 기술

가. 인공지능과 제4차 산업혁명

딥러닝 이전까지의 인공지능은 주로 텍스트 데이터에 적용된 지능형 에이전트를 개발하는데 그쳤다. 그러나 딥러닝으로 인해서 아주 복잡한 실세계의 센서 데이터도 학습할 수 있게 되었다. 최근 들어, 모든 사물들에 센서가 부착되고 인터넷에 연결되는 사물인터넷(IoT)의 시대가 도래하고 있다. 이를 통해서 물리적인 오프라인 아날로그 세계가 인터넷의 온라인 디지털 세계와 연결되는 online-offline 서비스가 시작되고 있다. 이와 같이 가상세계와 현실세계가 연결되면 지금까지 가상의 디지털 세계에만 적용되었던 인공지능 기술들이 현실의 아날로그 세계에까지 확장될 수 있고 이점에서 전세계가 인공지능화될 수 있는 시작점에 와 있다. 이 점에서 지난 2016년 2월의 다보스 포럼에서 인공지능을 제4차 산업혁명의 촉발제로 규정한 것은 큰 의미를 갖는다.

물리적인 현실세계와 디지털의 가상세계가 만나는 인공지능 연구의 대표적인 사례는 자율이동 로봇이다. 인공지능 연구자들은 1960년대부터 이동로봇 연구를 하였으나 본격적인 이동로봇 연구는 1990년대와 2000년대를 통해서 발전되었다. 그러나 인간형 휴머노이드 로봇 기술은 여전히 물체 감지와 조작을 위한 하드웨어의 발달이 늦어 생각보다 느리게 발전하고 있다. 반면에, 자율이동 로봇 연구의 파생 결과로서 자율주행 자동차가 등장하였다. 자동차는 물리 세계에서 이동 시간을 단축시켜 주는 대표적인 기계 장치이다. 그

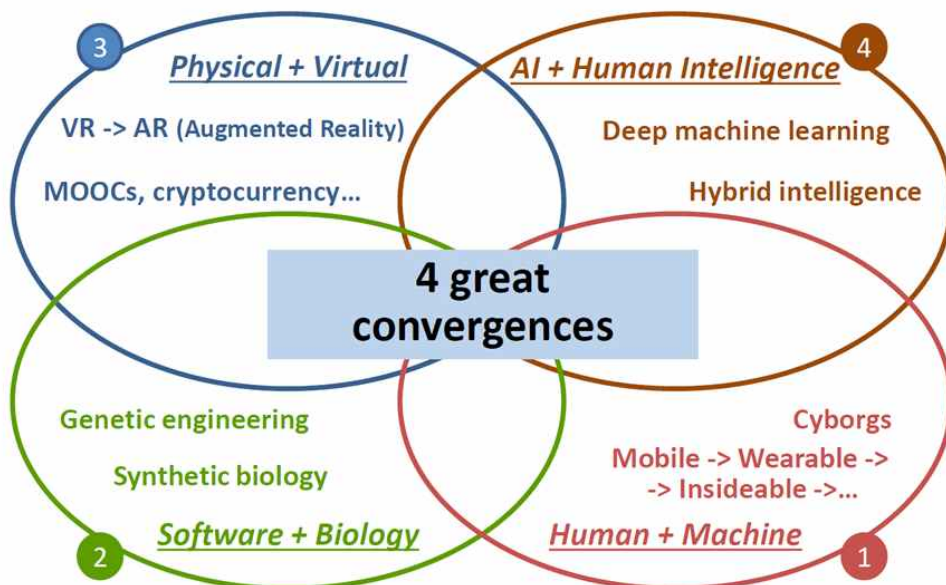


그림 19. 미래학자가 바라본 미래사회의 4대 융합 분야(David Wood, Anticipating the future of software, 2015년 서울대 초청강연 중)

러나 자동차가 구글맵과 같은 디지털 지도에 연결되고 도로 환경과 주변을 인식하며 스스로의 판단에 의해서 주행을 하는 인공지능 로봇 기술과 접목됨으로써 물리적인 세계에서 활동하는 인공지능이 탄생하게 된 것이다. 실제로 2005년의 DARPA 무인자동차 경주 대회에서 우승한 스탠포드 대학교의 인공지능랩은 자율이동 로봇을 연구해 왔으며 이 로봇 기술을 자동차에 적용하였다. 이 기술은 결국 구글의 무인차 프로젝트로 발전하였으며 오늘날 모든 자동차 회사들이 앞다투어 경쟁하고 있는 자율주행 자동차 산업으로 발전하였다.

최근 인공지능이 상용화에 성공한 또 다른 사례는 아마존 에코 비서 로봇이다. 에코는 스피커 형태를 가진 장치로서 사람과 대화하며 질문에 답하거나 음악을 틀어주거나 주문을 받아주는 역할을 하는 디지털 비서이다. 이에 앞서서 애플이 스마트폰의 앱으로 출시한 대화 에이전트 시리가 있으나 에코는 시리와는 달리 물리적인 형태를 갖춘 탁상형 로봇이다. 가상 세계에만 존재하던 시리가 몸체를 갖추고 물리적인 세계로 등장한 것으로 볼 수 있다. 2016년 5월에 구글도 구글홈이라는 탁상형 로봇을 출시하였다. 2016년 말에 국내의 SKT사에서 인공지능 홈비서 누구(Nugu)를 선보였다. MIT에서 창업한 지보사에서는 카메라를 갖추고 고개를 돌리는 탁상형 로봇 지보를 개발 중에 있다. 더 나아가 일본의 소프트뱅크는 페퍼라는 휴머노이드 로봇을 이미 상용화하여 일본에서는 상점 안내에 사용하고 있다.

최근 사용자 인터페이스가 터치에서 인공지능 대화로 옮겨가고 있다. 챗봇 기술의 발전으로 인해서 사람과 장치가 음성과 텍스트 대화로 상호작용하는 추세이다. 페이스북은 M이라는 챗봇을 개발하여 텍스트로 사람들과 대화하며 선물도 추천해 준다. 마이크로소프트는 최근에 테이라는 챗봇을 개발하여 트위터에 공개하였다. 사람들이 테이를 인종차별



그림 20. Smart Machine의 시대. Gartner는 스마트어드바이저, 지능형로봇, 스마트카 등의 Smart Machine은 IT 역사상 가장 파괴적인 기술이 될 것으로 예측(2014)하였으며, 인공지능 시장 규모는 2025년까지 70조 달러를 넘어설 것으로 예측된다 (YTN, 2015).

적인 발언을 하도록 가르쳐서 사고를 낸 적도 있다. 이는 앞으로 인공지능 기술이 윤리적 사회적 법적인 문제를 야기할 수도 있음을 경고하는 중요한 사건이 되기도 하였다. 최근에는 챗봇을 사용하여 금융 자산관리를 도와주는 로보어드바이저가 등장하고 있다. 또한 뉴욕타임즈는 2015년부터 신문 기사를 대신 써 주는 로봇을 사용하고 있다. 이와 같이 지능형 에이전트와 인공지능 로봇은 우리의 일상생활에 점차 가까이 다가오고 있다.

가트너 그룹은 가상비서, 개인서비스 로봇, 자율주행차, 드론 등을 가리켜서 스마트 머신이라고 하고 지금까지 보지 못한 가장 혁명적인 IT 기술이 될 것이라고 예측하였다. 다가오는 제4차 산업혁명의 시대에는 이러한 스마트 머신들의 새로운 종들이 등장하여 온라인과 오프라인을 연결하는 새로운 서비스와 사업 생태계를 만들어 갈 것이다.

나. 인공지능시대 인간의 삶

인공지능은 단순한 기술의 변화를 넘어서 타 산업과 사회전반에 걸친 변화를 야기할 새로운 패러다임으로 이해될 필요가 있다. 딥러닝 기술은 스스로 학습 진화하며 더욱 똑똑해지는 지능 폭발현상을 초래하였다. 똑똑한 기계를 이용하여 더욱 똑똑한 기계를 만들고 이는 다시 새로운 똑똑한 기계로 이어지는 지능폭발 현상이 이미 1965년에 예견된 바 있다. 2006년에는 레이 커즈와일에 의해서 인간의 지능을 능가하는 인공지능이 출현할 것이라는 특이점 논의가 시작되었다. 2014년에는 닉 보스트롬에 의해서 인간의 지능을 뛰어넘는 수퍼지능의 시대를 논하고 있다. 최근에 국제인공지능학회(AAAD)에서 발표한 AI100 Study 그룹의 보고서에 의하면 앞으로 2030년까지 자율주행, 홈로봇, 헬스, 교육, 사회 안전, 오락 등 인간의 모든 삶에 있어서 인공지능의 영향이 더욱 커질 것으로 내다보고 있다.

자율주행자동차는 생각보다 훨씬 빨리 상용화된 인공지능 기술 중 하나이다. 2004년에 미국의 DARPA는 무인자동차 경주대회를 모하비 사막에서 처음으로 개최하였다. 그러나 이때는 모든 참가팀들이 7마일도 못가고 실패하였다. 두 번째 실시된 2005년 그랜드 챌린지에서는 세바스찬 쓰런 교수가 이끈 스탠포드 대학교의 인공지능연구소가 우승하였다. 쓰런 교수는 후에 결국 구글 무인자동차를 개발하게 된다. 세 번째의 무인차 경주대회는 앞의 대회와는 달리 도시 환경에서 이루어진



그림 21. 인간의 삶을 바꾸는 스마트 머신의 사례

다. 어떤 챌린지로 불리는 이 대회는 실제의 마을을 하나 선정하여 무인차 경주를 하였으며, 다른 차들도 주행하고 신호등도 지켜야 하는 실세계 상황을 반영한 자율 주행 대회이다. 이 대회에서는 카네기멜런 대학의 무인차 보스가 일등을 차지하였고 스탠포드 대학의 주니어가 2등을 차지하였다. 이 대회에서는 실제로 무인차간 충돌 사고가 발생하고 교통 체증 현상도 일어났다.

자율주행차의 확대는 앞으로 인간의 삶과 일의 형태를 완전히 바꾸어 놓을 수도 있을 것이다. 우버와 리프트와 같은 온라인 택시 사업 회사들이 무인차를 개발하고 있으며 앞으로는 차를 소유하지 않게 될 수도 있다. 또한 산업이 재편될 수도 있다. 구글, 애플 등의 IT 회사들이 자율주행 신기술을 개발하여 기존의 자동차 회사들을 위협하고 있다. 전통적인 자동차 회사들은 이를 막기 위해서 안간힘을 쓰고 있다. 일본의 자동차 회사 토요타는 무인차 개발을 위해서 1조원의 연구비를 투입하여 MIT와 스탠포드 대학교에 인공지능 연구소를 설립하였다. 벤츠와 BMW 등도 미래형 자율주행 자동차를 선보이고 있다. 국내 회사들도 최근 자율주행 인공지능 기술과 스마트카 개발에 크게 투자하고 있다. 한편 삼성전자나 LG전자 및 SKT 등도 IT 기술의 장점을 활용하여 자동차 산업에 뛰어들고 있다.

인공지능은 홈서비스 로봇 분야에서 향후 15년 동안에 많은 발전이 이루어질 것으로 예상되고 있다. 2015년 12월에 전기자동차 회사 테슬라 사장인 일란 머스크는 1조 원짜리 인공지능 회사 OpenAI를 설립하였다. OpenAI는 인공지능 기술을 개발하여 모두 공개하는 것을 원칙으로 하고 있으며, 홈로봇을 위한 인공지능 기술 개발에 박차를 가하고 있다. 올 초에 페이스북 최고경영자인 마크 저커버그는 개인의 신년 결심으로 홈 환경에서 개인서비스를 하는 로봇을 개발하겠다고 하였다. 집안을 돌아다니면 가사를 돌보고 경비를 하는 다양한 형태의 개인서비스 로봇들이 개발되고 있다. 서울대 인지로봇인공지능 연구센터에서는 뽀로로 만화영화를 보고 아이들과 질의응답하며 놀면서 영어를 가르치는 교육용 로봇 뽀로로봇을 개발하였다. 또한 직장을 가진 엄마를 대신하여 아이들의 하루 일정을 챙겨주는 보모로봇 오페어(Aupair)를 개발하고 있다. 국내 회사로는 네이버랩스가 최근 로보틱스 연구팀을 구성하여 인공지능 연구자를 모으고 있다.

인공지능은 미래의 교육 방식에 큰 혁신을 가지고 올 것으로 예상된다. 동영상 강의 자료를 인터넷에서 온라인 서비스하는 MOOC가 활발해지고 있다. 초중고 수준의 다양한 온라인 강의를 제공하는 칸 아카데미를 비롯하여 대학교 강의를 온라인화 한 Coursera와 edX, 그리고 기업체 요구의 실제적인 IT 교육을 해주는 Udacity 등 다양한 모델의 온라인 교육 회사들이 등장하였다. 인공지능과 머신러닝 기술을 사용하면 피교육자들의 학습 특성을 파악하여 개인맞춤형 교육을 제공할 수 있고 교사들의 교육의 질을 높일 수 있다. 특히 가상현실이나 로봇과 결합하여 모바일 상황에서 교육이 이루어질 경우 획기적인 새로운 교육 모델이 등장할 수도 있다.

헬스케어는 고령화되어가는 선진 사회에서는 그 중요성이 점차 커지고 있는 새로운 산업이다. 모바일 기술과 웨어러블 장치의 등장을 통해서 개인의 일상 기록을 습득할 수 있을 뿐만 아니라 머신러닝 기반의 데이터분석을 통해서 맞춤형 서비스가 가능해 지고 있다. 의료뿐만 아니라 스포츠나 피트니스 분야에서 최근 인공지능 기술의 도입이 시작되고 있다. IBM은 왓슨을 사용하여 스포츠웨어 회사 언더아머와 협력으로 개인 맞춤형 건강관리 서비스를 추진하고 있다.

사회 안전과 보안 시스템 및 소외 계층에 대한 배려 등 사회 시스템을 개선하는 데에도 인공지능 기술이 사용된다. 또한 새로운 미디어를 통한 엔터테인먼트 산업에도 인공지능 기술이 기여하고 있다. 이미 오래전부터 아마존이나 넷플릭스 같은 인터넷 서비스 업체들은 머신러닝 기법을 사용하여 책이나 영화, 음악 등의 개인 맞춤형 추천서비스를 실시하였다. 최근에는 왓스앱이나 스냅챗 같은 채팅 프로그램이 등장하여 소셜넷을 통해서 서로 소통할 수 있게 해 준다. 많은 사용자생산콘텐츠가 자연언어, 음성, 영상 데이터를 포함하고 있고 머신러닝 인공지능 기법들은 이를 분석하여 서비스를 자동화하고 더욱 편리하게 사용할 수 있도록 해준다. 앞으로는 소프트웨어뿐만 아니라 더욱 저렴한 가격의 센서와 장치들이 등장함으로써 가상현실, 촉각 장치, 반려 로봇 등과 결합하여 보다 대화 기반이고 인간같은 상호작용이 이루어질 것이며 인지 능력과 감성, 교감 등의 정서가 더욱 중요해 질 것으로 보인다.

인공지능이 가져올 변화 중에는 많은 사람들이 우려하는 것도 있다. 특히 일과 직업에서의 변화이다. 인공지능은 오랜 동안 사람을 닮은 지능형 에이전트 또는 도우미 로봇 기술을 연구해 왔으며 최근 들어서 로봇 어드바이저, 로봇 저널리스트 등 인간의 일의 영역을 침범하고 있다. 심지어는 음악이나 미술 등과 같은 예술의 영역에서도 인공지능이 사



그림 22. 인간과 기계와의 중복 영역이 생기는 분야의 사례

람의 일을 일부 대신하기 시작하였다. 영국의 이코노미스트지는 최근 인공지능에 의해서 타격을 받기 쉬운 직업들을 정리하는 기사를 낸 적이 있다. 지식을 필요로 하면서 반복적인 일들이 우선 인공지능에 의해서 빨리 대체될 직업들에 속하고, 반면에 아주 창의적인 일이나 면대면으로 상담을 해야 하는 컨설팅 등은 인공지능으로 대체되는데 많은 시간이 필요로 할 것으로 내다보고 있다.

다. 영화 속의 미래 인공지능

인공지능이 최근 아주 빠른 속도로 산업화되고 있다. 때로는 인공지능 연구자들조차도 놀라게 한다. 자율주행 자동차의 등장이 그러하고 대화하는 챗봇의 보급이 그러하며 개인 서비스 로봇의 부상이 그러하다. 인공지능의 미래를 예측하기란 쉬운 일이 아니다. 아마도 미래를 예측하기 보다는 미래는 만들어 가는 것이 더 쉬운 일일 수도 있을 것이다. 많은 과학기술의 발전이 그러했듯이 과학공상영화가 미래를 일견하게 해준다. 인공지능이 이슈화되면서 관련 영화들도 최근 많이 등장하였다.

2013년에 나온 Her라는 영화는 로맨틱 공상과학 코미디 드라마로 분류되는 인공지능 영화다. 주인공은 스마트폰의 인공지능 운영체제 형태로 서비스되는 사만타라는 이름의 인공지능 가상로봇과 사랑에 빠진다. 이혼 후 외로이 살아가는 남자의 여자 친구 역할을 인공지능이 해 줄 수도 있음을 암시한다. 그러나 결국 남자 주인공은 인공지능 사만타가 자기만을 사랑하는 것이 아니라 수천 명의 고객을 서비스하는 기계라는 것을 알고 크게 실망한다. 2012년에 나온 로봇과 프랭크라는 영화는 노인 돌보미 로봇과의 생활을 그린 공상과학영화이다. 직장일로 바빠서 아버지를 돌 볼 시간이 없는 아들이 휴머노이드 로봇을 한대 선물한다. 아버지 프랭크는 처음에는 로봇을 싫어하지만 점차 친해져서 딸보다도



Her (2013)



Robot & Frank (2012)

Bicentennial Man (1999)



Ex Machina (2015)



그림 23. 영화 속에 등장한 미래의 인공지능 사례

편함을 느낀다. 도벽이 있는 프랭크는 마음에 들지 않는 부부의 집을 털어서 보물을 훔치기로 결심하고 로봇과 함께 도둑질을 모의하게 된다. 도우미로서 친구로서 반려자로서 노인과 함께 살아가는 기계와 사람의 공존의 일면을 엿볼 수 있는 미래의 모습이다. 2009년에 나온 영화 바이센테니얼 맨에서는 인간형 AI로봇 앤드류가 창의성을 가지며 물건을 만들어 돈을 벌고 경제력도 가진다. 기술의 발전에 힘입어 사람의 피부와 내장을 가져 음식을 먹고 감정을 가지며 이성에 대한 사랑을 느낀다. 결국은 인간의 권리를 갖게 해달라고 법정에 소송까지 하게 되며, 인간의 권리를 얻기 위해서 인간은 죽는다는 것을 인정하고 사랑하는 사람과 함께 죽음을 택한다.

철학적으로 인공지능은 1950년대에 기능주의를 기반으로 출발하였다. 즉 반도체로 만든 지능도 인간의 지능과 대등할 수 있다는 입장이었으며 일부 인공지능 연구자들은 이를 굳게 믿는 강인공지능 주의자였다. 그러나 대부분의 인공지능 연구자는 인공지능이 인간의 지능과 동등하지는 않더라도 유용하며 인간을 이해하는데 도움이 되는 좋은 도구라고 생각하는 약인공지능 주의자들이었다. 그러나 최근 구글 자율주행 자동차, 딥마인드의 인공지능 알파고, 그리고 보스턴 다이내믹스에서 만든 눈길을 넘어지지 않으면서 걷고 물건을 두 팔로 들어 나르는 휴머노이드 로봇 아틀라스를 보면서 다양한 문제를 풀도록 학습될 수 있는 범용인공지능 즉 강인공지능에 대한 두려움을 느끼는 사람들이 늘어나고 있다. 스스로 학습하는 기계의 등장은 선순환 되먹임 사이클을 반복할 경우 더욱 가속화되어 지능폭발 현상을 유발하고 이를 통해서 어느 날에는 인간을 능가하는 인공지능이 등장하는 특이점이 도래하고 결국은 의식을 가진 수퍼지능도 등장할 수 있다고 보는 두려움을 야기한다. 따라서 앞으로의 인공지능 연구는 단순한 과학기술적 경제산업적 파급효과를 넘어서 인문학적, 윤리적, 사회적, 법적인 이슈들을 포함하는 다각도의 논의가 함께 이루어져야 할 것이다.

<참고문헌>

- Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798–1828.
- Bengio, Y. (2009). Learning Deep Architectures for AI. *Foundations and Trends® in Machine Learning* (Vol. 2).
- Bishop, C. M. (2006). *Pattern Recognition And Machine Learning*. Springer.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Haykin, S. (2008). *Neural Networks and Learning Machines* (Vol. 3). New Jersey USA: Pearson Prentice Hall.
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science*, 313(5786), 504–507.
- Koller, D., & Friedman, N. (2009). *Probabilistic Graphical Models: Principles and*

Techniques. MIT Press.

- Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, 350(6266), 1332–1338.
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2016). Building Machines that learn and think like people. *Behavioral and Brain Science*, (in press).
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- Mackay, D. J. C. (2003). *Information Theory , Inference , and Learning Algorithms*.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw Hill.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460.
- Wachsmuth, I., Lenzen, M., & Knoblich, G. (2008). *Embodied Communication in Humans and Machines*. Embodied Communication in Humans and Machines. Oxford University Press.
- Widrow, B., & Lehr, M. A. (1990). 30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation. *Proceedings of the IEEE*, 78(9), 1415–1442.
- Zhang, B.-T. (2008). Hypernetworks: A molecular evolutionary architecture for cognitive learning and memory. *IEEE Computational Intelligence Magazine*, (August), 49–63.