

# 인공지능 전이학습과 응용 분야 동향



강수철 || (주)롯데정보통신 정보기술연구소 수석연구원

딥러닝 알고리즘 개발로 촉발된 인공지능 기술의 급속한 발전은 기계가 인간의 대화를 이해하고 응대하며 인간보다 더 뛰어난 시각 능력으로 다양한 사물을 분류할 수 있는 수준에 이르렀다. 하지만 여전히 인공지능 기술은 새롭게 부상하는 기술(Emerging Technologies)이며, 많은 투자와 연구가 지속되고 있다. 특히, 인공지능 응용 기술 및 서비스의 상업적 확산을 위해 많은 기업들이 총력을 기울여 연구개발 및 상업화를 진행 중에 있으며, 국가 차원에서도 이를 뒷받침하고 있다. 딥러닝 기반 알고리즘들은 대부분 빅데이터와 많은 양의 연산 장치들이 필요하므로 상업화로 가는 길목이 더딜 수 밖에 없다. 따라서 적은 데이터와 좀 더 적은 자원으로 인공지능을 다양한 영역으로 확산시키기 위한 효율성을 추구하는 기술들이 개발되고 있다. 그 중에서도 한 분야에서 학습한 결과를 학습한 적이 없는 다른 분야에 적용해서 학습하는 전이학습(Transfer Learning)이 각광받고 있다. 본 고에서는 전이학습 기술에 대한 개념을 간단히 살펴보고, 이에 대한 응용 영역들에 대해 좀 더 상세히 살펴봄으로써 전이학습 기술에 대한 이해와 향후 발전 방향을 알아보도록 한다.

## I. 서론

수 년간 저성장 기조의 경제환경 속에서 G2 국가인 미국과 중국을 중심으로 4차 산업혁명을 선도하기 위해 인공지능 기술 개발에 대한 투자가 활발히 이루어지고 있다. 지난

\* 본 내용은 강수철 수석연구원(☎ 02-2028-9011, kangsc@lotte.net)에게 문의하시기 바랍니다.

\*\* 본 내용은 필자의 주관적인 의견이며 IITP의 공식적인 입장이 아님을 밝힙니다.

7월 한국을 방문한 손정의 소프트뱅크 회장도 문재인 대통령을 접견하면서 앞으로 대한민국이 집중해야 할 것은 “첫째도 인공지능, 둘째도 인공지능, 셋째도 인공지능”이란 조언을 전해주었다. 지난 11월 말 한국 바둑의 전설 이세돌 9단이 24년의 프로기사 생활을 마무리하고 은퇴했다. 그는 2016년 구글 딥마인드(DeepMind)가 만든 인공지능 바둑 프로그램인 알파고(AlphaGo)와 5번의 대국을 펼쳐 1승 4패의 성적을 거두며 인류 중에서 유일하게 알파고를 이겨본 바둑 기사가 되었다. 하지만 그는 은퇴 이유 중 하나로 알파고에게 패배한 것을 들었다. 이러한 알파고를 만든 아버지로 불리는 데미스 하사비스(Demis Hassabis)는 알파고를 업그레이드시킨 알파고 제로(AlphaGo Zero)를 2017년에 내놓고 이에 대한 의학 분야 응용을 위해 적용한 전이학습(Transfer Learning) 기술에 대해 다음과 같이 언급하였다. “전이학습이 일반 인공지능으로 가는 열쇠가 될 것이라 생각합니다. 이를 위해서는 우선 학습한 사실에서 인지적인 세부사항을 추상화해야 합니다. 그리고 여기에서 개념적인 지식을 획득하는 것이 전이학습을 할 수 있는 핵심입니다.”[1]

알파고는 방대한 인간의 기보와 바둑의 기본 원리를 데이터로 삼아 학습하며 정책망(Policy Network)과 가치망(Value Network)이란 두 인공 신경망에다 강화학습(Reinforcement Learning)을 결합해 최적의 승률을 계산해내는 방식으로 구동되었던 반면, 알파고 제로(AlphaGo Zero)는 어떤 데이터도 없이 아예 백지 상태에서 수많은 강화학습을 통해 바둑을 스스로 터득했다고 알려졌다. 강화학습은 머신러닝의 한 분야로 보상을 최대한 많이 받는 방법을 학습하는 것인데, 바둑이라면 상대방을 이기는 것이 보상이므로 이기려는 방법을 인공지능이 습득하도록 반복해서 학습하게 된다. 알파고와 알파고 제로의 차이는

[표 1] 알파고(AlphaGo) 시리즈 성능 비교

이름	공개시점	전적	학습법	하드웨어
알파고 판	2015년 10월	판후이 2단계에 5-0 승리	딥러닝, 강화학습	GPU 176개, TPU 4개
알파고 리	2016년 3월	이세돌 9단계에 4-1 승리	딥러닝, 강화학습	GPU 176개, TPU 4개
알파고 마스터	2017년 5월	커제 9단계에 3-0 승리	딥러닝, 강화학습	TPU 4개
알파고 제로	2017년 10월	알파고 리에 100-0, 알파고 마스터에 89-11 승리	강화학습	TPU 4개

〈자료〉 동아일보, “기보없이 독학… 최강 알파고버전 등장”, 2017. 10. 19.

기보를 익힌 상태에서 강화 학습을 했느냐, 완전한 무(無)의 상태에서 강화학습을 했느냐로 볼 수 있다. 그래서 알파고 제로의 성과는 인공지능이 인간이 해결하지 못한 난제를 풀 수도 있다는 잠재력과 강화학습이 인공지능 영역을 더 확장할 수 있다는 기대감으로 이어졌다.

그러나 모든 분야에서 알파고 제로와 같은 강화 학습 모델을 적용하는 것은 한계가 있다. 인공지능망이 문제에 대한 인간의 인지까지 모방할 수는 없기 때문이다. 게임처럼 보상이 확실한 분야에 적용하기는 쉽지만, 그렇지 않은 분야에서의 머신러닝은 충분한 데이터를 요구한다. 인류가 오랜 시간 걸쳐 쌓아놓은 지식을 배우는 시기가 필요했던 것처럼 말이다. 문제는 많은 데이터를 확보한 분야라면 머신러닝 모델을 적용하기가 쉽지만, 데이터가 부족한 분야는 그렇지 않다는 것이다. 인간과 기계의 인지 차이로 강화 학습을 적용하기 어려운 분야라면 더욱 그렇다. 그래서 주목 받는 방법이 바로 전이학습이다.

IT전문 리서치 기관 가트너(Gartner)는 2019년 새롭게 부상하는 기술(Emerging Technologies) 중 하나로 전이학습을 언급하였다. 가트너는 인공지능 기술을 실제 사용



[그림 1] 가트너 Hype Cycle for Emerging Technologies 2019

자와 기업들이 쉽게 활용하기 위해 데이터 부족, 하드웨어 자원 부족, 오버슈팅 등 여러 가지 기술 장벽들을 뛰어넘을 수 있게 해주는 필요한 기술로 전이학습을 언급하면서 한 가지 문제를 통해서나 한 영역의 데이터 셋에서 배운 것을 다른 영역에 적용할 수 있는 응용 기술이라고 소개하고 있다[2].

## II. 경량 딥러닝 연구 동향

딥러닝 기반의 인공지능 기술이 확산되면서 자율주행, 대화형 인공지능 스피커, 얼굴인식 및 생체인증 등 다양한 분야에 해당 기술이 활용되고 있다. 딥러닝의 주요 알고리즘들은 많은 양의 데이터를 통한 학습이 필요하며, 이를 위해 강력한 GPU(Graphics Processing Unit) 기반의 컴퓨팅 자원이 필요하다. 하지만 인공지능 기술의 확산과 함께 지능형 기기들의 다양한 산업적 응용을 위해서는 경량 디바이스, 모바일 디바이스, IoT 센서와 같이 저전력과 저사양의 기기에서도 딥러닝 알고리즘과 모델이 구동되어야 한다. 이러한 장벽을 극복하기 위해 기존의 학습된 모델의 정확도를 유지하면서 보다 크기가 작고, 연산을

[표 2] 경량 딥러닝 연구동향

구분	접근방법	연구방향
경량 알고리즘 연구	모델 구조 변경	잔여 블록, 병목 구조, 밀집 블록 등 다양한 신규 계층 구조를 이용하여 파라미터 축소 및 모델 성능을 개선하는 연구(ResNet, DenseNet, SqueezeNet)
	합성곱 필터 변경	합성곱 신경망의 가장 큰 계산량을 요구하는 합성곱 필터의 연산을 효율적으로 줄이는 연구(MobileNet, ShuffleNet)
	자동 모델 탐색	특정 요소(지연시간, 에너지 소모 등)가 주어진 경우, 강화 학습을 통해 최적 모델을 자동 탐색하는 연구(NetAdapt, MNasNet)
알고리즘 경량화 연구	모델 압축	가중치 가지치기, 양자화/이진화, 가중치 공유 기법을 통해 파라미터의 불필요한 표현력을 줄이는 연구(Deep Compression, XNOR-Net)
	지식 증류	학습된 기본 모델을 통해 새로운 모델의 생성 시 파라미터값을 활용하여 학습시간을 줄이는 연구(Knowledge Distillation, Transfer Learning)
	하드웨어 가속화	모바일 기기를 중심으로 뉴럴 프로세싱 유닛(NPU)을 통해 추론 속도를 향상시키는 연구
	모델 압축 자동 탐색	알고리즘 경량화 연구 중 일반적인 모델 압축 기법을 적용한 강화 학습 기반의 최적모델 자동 탐색 연구(PocketFlow, AMC)

<자료> 한국전자통신연구원, “경량 딥러닝 기술 동향”, 전자통신동향분석 제34권 제2호, 2019. 4.

간소화하는 연구인 경량 딥러닝 연구가 활발히 진행되고 있다.

경량 딥러닝 연구는 기존 클라우드 기반의 학습된 모델을 경량 장치에 내장하기 위한 필수 기술이며, 이를 통해 지연시간 감소, 민감한 개인정보 보호, 네트워크 트래픽 감소와 같은 다양한 이점을 가지게 된다. 경량 딥러닝 기술은 알고리즘 자체를 적은 연산과 효율적인 구조로 설계하여, 기존 모델 대비 효율을 극대화하기 위한 경량 딥러닝 알고리즘 연구와 만들어진 모델의 파라미터들을 줄이는 모델 압축(Model Compression) 등의 기법이 적용된 알고리즘 경량화 기술로 나눌 수 있다[3].

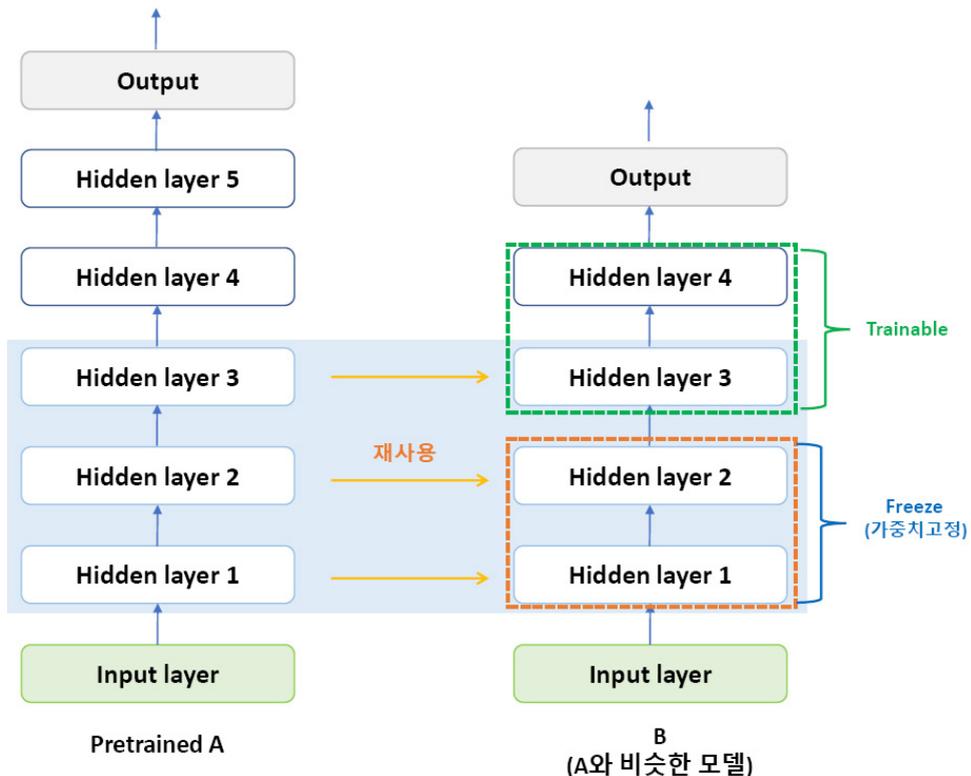
경량 딥러닝 알고리즘은 가장 일반화된 합성곱 신경망(Convolutional Neural Network: CNN)을 통해 다양한 연구가 진행 중이다. CNN 계열의 모델에서는 주로 학습 시 가장 큰 연산량을 요구하는 합성곱 연산을 줄이기 위한 효율적인 합성곱 필터기술이 일반화되고 있다. 다양한 신규 계층 구조를 설계하여 신경망 구조를 제공함으로써 우수한 추론 성능을 보이는 연구도 소개되고 있다. 이는 기본 단일 층별 연산에 그치지 않고 연산량과 파라미터 수를 줄이기 위한 잔여 블록(Residual Block) 또는 병목 블록(Bottleneck Block)과 같은 형태를 반복적으로 쌓아 신경망을 구성하는 방법이다. 그 밖에도 기존 신경망의 모델 구조를 인간에게 의존적으로 수행하지 않고 모델 구조를 자동 탐색함으로써 모델을 자동화하거나 연산량 대비 모델 압축 비율을 조정하는 등 다양한 자동 탐색 기술이 존재한다. 이는 모바일 딥러닝과 같은 다양한 기기의 성능 대비 추론 속도가 중요한 응용을 위해 정확도, 지연시간, 에너지 소모량들을 조정하여 강화 학습을 활용한 경량 모델을 탐색하는 기술이다.

이에 비해 알고리즘 경량화는 경량 딥러닝 알고리즘과 달리, 모델이 표현하는 다양한 파라미터의 크기를 줄이는 데 주목적을 가지고 있다. 파라미터가 가지는 표현력을 가능한 한 유지하면서 불필요한 가중치를 최대한 없애기 위한 방법들이다. 일반적인 딥러닝 모델은 과파라미터화(Over-parameterization) 되어 있기 때문에 모델이 가지는 가중치의 실제 값이 아주 작을 경우 모델의 정확도에 큰 영향을 미치지 못하므로, 이 값을 모두 0으로 설정하여 마치 가지치기(Pruning)를 수행하는 것과 같은 효과를 내는 가중치 가지치기(Weight Pruning)가 대표적이다. 다음으로, 일반적인 모델의 가중치는 부동 소수점 값을 가지지만, 이를 특정 비트 수로 줄이는 양자화(Quantization)를 통해 기존 딥러닝의 표현력을 유지하면서 실제 모델의 저장 크기는 줄이는 방법이 있다. 마지막으로, 0과 1로

표현하여 표현력을 많이 줄이지만, 정확도의 손실은 어느 정도 유지하면서 모델 저장 크기를 확연히 줄이는 이진화(Binarization) 기법 등이 있다. 또한, 심층 신경망으로 학습된 모델들의 숨은 지식을 계산량이 적고 얇은 신경망으로 전달하는 방법 중의 하나로 전이학습이 있다. 이것은 기존의 학습된 모델과 비슷한 유형의 다른 모델로 학습된 결과를 옮겨서 부족한 데이터를 통한 학습이나 훈련 시간을 단축시키는 방법이라 할 수 있다.

### III. 전이학습의 개요와 적용 고려사항

전이학습이란 해결하고자 하는 문제에는 정답이 소량만 존재하는 반면, 해결하고자 하는 문제와 비슷한 문제에는 정답이 대량으로 존재할 경우 사용하는 지도학습(Supervised



<자료> Hands-On Machine Learning with Scikit-Learn and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2017, Chapter 11. Training Deep Neural Nets 참조 재구성

[그림 2] 전이학습의 개념

Learning)이다. 예를 들면, 골프의 학습 데이터는 충분히 주어졌지만 야구의 학습 데이터는 거의 없을 경우에 야구의 실력 향상을 목표로 하는 사례를 들 수 있다. 프로 야구 선수 중에 골프를 잘 하는 선수가 많기도 하고, 이러한 스포츠 종목들에는 공을 멀리 보내기 위해 배트나 골프채를 휘두르는 공통점이 있다. 그리고 그 공통점을 이용한 학습이 바로 전이학습이다. 먼저 골프의 학습 데이터로부터 골프 실력 향상을 꾀하고, 야구의 학습 데이터를 이용해서 골프의 움직임을 야구용으로 조정하는 것과 같은 것이다. 그리고 전이학습은 어떠한 분야의 지식을 다른 분야에 적용시키고자 하는 특성으로 인해 배구 등 다른 도메인 영역들로 지속적인 확장이 가능하다.

전이학습의 개념을 간단하게 살펴보면, 규모가 매우 큰 딥러닝 모델을 학습시킬 때 처음부터 새로 학습시키는 경우에 학습 속도가 느린 문제가 발생한다. 이러한 경우 기존에 학습된 비슷한 모델이 있을 때 이 모델의 하위층(lower layer)을 가져와 재사용하는 것이 학습 속도를 빠르게 할 수 있을 뿐만 아니라 학습에 필요한 훈련 셋(Training set)도 훨씬 적다. 예를 들어, 아래의 도식처럼 비행기, 자동차, 트럭, 새, 고양이 등 이미지의 데이터 셋을 분류하는 모델 A가 있다고 하자. 그런 다음 분류된 이미지에서 자동차의 종류를 분류하는 모델인 B를 학습시킨다고 할 때, 학습된 모델 A에서의 일부분(lower layer)을 재사용하여 모델 B를 학습시킬 수 있다[4].

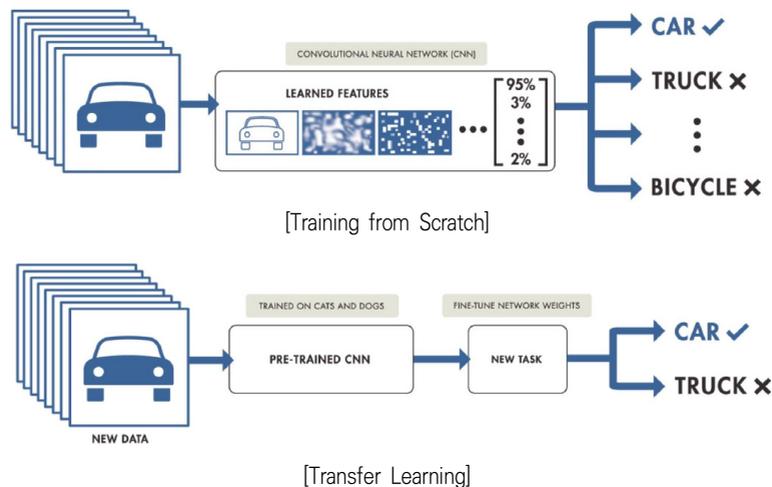
전이학습을 이미지 처리에 응용한 사례를 하나 살펴보자. CNN은 이미지 처리에 가장 널리 사용되는 알고리즘이다. 숫자 손글씨 데이터셋인 MNIST는  $28 \times 28$  픽셀로 되어 있다. 이를 일반적인 심층신경망(Deep Neural Network: DNN)으로 처리하려면, 먼저 이미지 하나를 784( $28 \times 28$ )개의 소수점 숫자로 된 벡터로 변환한 다음, 이를 뉴럴 네트워크에 집어넣어 학습시켜 분류를 한다. 꽤 괜찮은 성능이 나오기는 하지만 CNN만큼은 아니다. DNN이 이미지를 기계적으로 줄 세워서 처리한다면, CNN은 근접한 픽셀끼리 묶어서 처리한다는 점에서 데이터의 공간적인 특성을 살린 처리가 가능하다[5].

하나의 필터가 동일한 가중치(weight)로 전체 이미지를 훑어 특징 맵(feature map)을 만드는 CNN의 특성은 학습해야 할 파라미터의 수를 줄여 학습 효율이 좋긴 하지만 좋은 분류기를 만드는데 꽤 오랜 학습 시간이 걸린다. MNIST는 2개의 컨볼루션 계층(Convolution Layer)으로도 금방 결과가 나오지만, 애초에 데이터의 크기가 작고 컬러 채널이 1개뿐이다. 하지만 복잡한 도심 길거리에서 자동차를 식별해내는 기능처럼 우리가 분류하기를

원하는 현실 세계의 데이터는 그보다 크고 색깔도 다양하다. 처음부터 CNN을 학습시키자니 너무 시간이 오래 걸리고 효율이 낮다면, 성능이 입증된 CNN을 가져다가 특징을 추출하고, 이를 바탕으로 우리가 원하는 분류를 수행하도록 만들 수 있다.

신경망(CNN)을 처음부터 끝까지 학습하는 경우는 별로 없다. 왜냐하면 그만큼 큰 데이터셋을 구하기가 어렵기 때문이다. 그 대신에 이미 대규모 데이터를 대상으로 학습이 끝난 CNN 모델을 가져다가 초기값으로 사용하거나 고정된 특징 추출기로 사용할 수 있다. CNN 모델을 고정된 특징 추출기로 사용하는 경우에는 CNN 끝에 달린 완전하게 연결된 계층(Fully-Connected Layer)을 없애고, 콘볼루션 계층을 통해 처리되는 값만 얻으면 된다[6].

이렇게 전이학습을 활용하기 위해서는 몇 가지 고려할 사항들이 있다. 우선 새로 훈련할 데이터가 적지만 원본 대상 데이터와 유사할 경우에는 새로 학습할 데이터는 원본 대상 데이터와 유사하기 때문에 위의 경우 최종 분류기 층(classifier layer)만 학습하게 된다. 그 다음에 새로 훈련할 데이터가 매우 많으며 원본 대상 데이터와 유사할 경우에는 전체 층에 대해서 미세 조정을 수행한다. 마지막으로 새로 훈련할 데이터가 적으며 원본 대상 데이터와 다른 경우에는 데이터의 양이 적기 때문에 최종 단계의 분류기 층을 학습시키는 게 좋다. 이렇게 층의 갯수, 활성화 정도, 하이퍼 파라미터(hyper parameters) 등 실제



〈자료〉 Training Deep Learning Models with Transfer Learning, Jon Cherrie, MathWorks, 2016, 참조 재구성

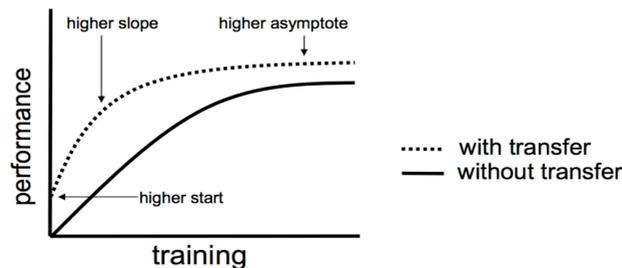
[그림 3] 전이학습을 통해 자동차 이미지 식별하기

연산 시 고려해야 할 사항들이 많기 때문에 다양한 선행 연구들을 참조하여 전이학습을 최적화하도록 해야 한다.

#### IV. 전이학습의 장점 및 응용 분야

전이학습은 다른 도메인 영역에서 미리 만들어진 모델을 잘 활용하는 것이 관건이며, 이에 따라 다른 경우에 비해 초기에 빠른 학습을 시작할 수 있다. 또한, 학습의 향상도가 좀 더 가파르게 상승하며, 다른 경우에 비해 좀 더 높은 수준에서 학습 결과가 수렴된다. 이러한 장점들로 인해 다양한 응용 분야로 확장이 가능하다.

우선 인공지능의 총아라 할 수 있는 자율주행차 분야에서 전이학습 모델이 활발히 적용되고 있다. 자율주행차가 특정 상황에서 어떻게 운전을 해야 하는지 배워야 하는 상황에서 학습이 안 된 상태로 자율주행차가 돌아다닌다면 실험용 차량은 물론이고 주변 사람까지 위협에 처하게 된다. 이럴 때 자율주행차용 시뮬레이션 환경을 만들고 거기서 학습된 변수들을 이용하여 전이학습을 현실에서 하는 것이 더 안전할 것이다. 실제로 현대자동차의 자율주행 기술 연구에서 도메인 전이학습(Domain Transfer Learning)을 활용하여 인터넷 사이트나 영화 장면 등에서 수집된 관련 데이터를 학습시켜 모델로 만든 뒤 보행자의 무단횡단이나 다른 차량의 역주행 같은 돌발행동에 대처할 수 있도록 하고 있다[7]. 이에 발맞춰 인공지능 컴퓨팅 분야의 선두기업인 엔비디아(NVIDIA)에서는 최근 사전 학습된 모델과 전이학습 기법을 쉽게 활용할 수 있도록 TLT(Transfer Learning Toolkit)를 출시하였다[8].



〈자료〉 Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, Chapter 11. Transfer Learning, p.243.

[그림 4] 전이학습을 활용할 때 장점

전이학습이 유용하게 활용될 수 있는 분야 중 하나는 의료 영상 데이터의 분석 영역이다. 특정 뇌 질환이 있는 환자의 데이터가 필요한 경우 개인정보 비식별화 같은 민감한 사안이 있을 수 있고, 그 전에 전체 환자 수가 워낙 적어서 절대적인 데이터 수가 부족할 수 있다. 이러한 상황에서 모델의 변수들을 처음부터 학습하려고 하면 데이터 수가 부족해 성능이 잘 나오지 않는다. 이 때 비슷한 기전의 질환에 대한 데이터가 많거나 혹은 데이터를 분류할 때 고정 필터로 소수의 데이터를 활용할 수 있는 전이학습이 도움이 된다[9].

최근에는 BERT나 GPT2 같이 자연어 처리에서도 전이학습을 활용하고 있다. 언어모델이 어떠한 방식으로 사용될지 확실치 않더라도 미리 양질의 방대한 양의 학습 데이터로 미리 학습을 시켜두고 이렇게 학습된 모델을 간단하게 조정하여 다른 목적으로 활용하는 것이다. 즉, 영상 처리 분야에서 성공한 이미지넷(ImageNet)의 사전 훈련된 신경망 가중치 파라미터들을 여러 다른 영상 처리 분야에 활발하게 사용했던 것처럼 자연어 처리 분야에서도 수많은 문장을 수집하여 학습한 신경망을 통해 다른 문제에 성공적으로 적용하여 성능을 개선하고 있는 것이다. 대표적인 예로 프랑스어 해독 능력을 학습하고 추가적인 데이터 없이 영어 해독 능력을 갖추는 지식 전이가 가능하다[10].

마지막으로 로봇틱스 분야에서도 전이학습을 많이 활용하고 있다. 기존의 인공지능 기술을 탑재한 로봇틱스는 동적인 환경에서 자율로봇을 사용하여 다양한 임무를 수행하기 위해 온라인 학습, 강화학습과 같은 적응형 기계학습 기법을 시도하여 왔으나, 초기 학습 단계의 제한적, 과적합 모델 구조로 인해 로봇 구성, 대상 작업, 환경 등이 다른 이종의 도메인에 대한 학습 한계를 극복하지 못하고 있다. 예를 들어, 생김새가 다른 두 휴머노이드 로봇이 있다고 가정해보면, 먼저 한 로봇이 택배를 나르는 작업을 7일이 걸려 학습하였다고 한다. 학습을 마친 로봇은 더 이상 사람의 개입이 필요하지 않으며 효율적으로 일을 잘 한다. 또 다른 로봇에게 주어진 같은 작업을 학습하게 하려면, 기존의 인공지능 기술로는 다시 7일을 소요 하에 로봇에게 학습시켜야 한다. 만약, 사람처럼 이전의 로봇이 지도자가 되어 다른 로봇에게 지식을 전이할 수 있다면, 일의 효율성이 좋아질 것이다[11].

## V. 결론

사람도 무언가를 배울 때 무(無)에서 시작하지 않는다. 평생 쌓아온 지식을 바탕으로

새로운 사실을 이해하며, 딥러닝 역시 같은 방식이 필요하다. 특히, 라벨 없이 스스로 패턴을 파악하며 학습하는 것이 중요하며, 이렇게 비지도학습(Unsupervised Learning)과 전이학습이 결합된다면, 지금까지와 비교할 수 없을 만큼 인공지능이 널리 활용될 것으로 예상된다. 하지만 현재까지 전이학습은 만능이 아니다. 적용하려는 범위에 따라 유용하지 않을 수 있다는 한계도 있다. 또한, 다른 환경에서 만들어진 전이학습 모델이 제대로 작동해야 다른 범위에도 적용할 수 있다는 전제가 필요하다. 전이학습은 인공지능이 특정 문제를 해결하기 위해 누적한 지능을 다른 문제에 적용할 수 있는지 확인하는 분야이자 인공지능을 현실에 유용한 형태로 도입할 수 있게 해주는 통로로서 활발히 연구되고 있다.

[ 참고문헌 ]

- [1] Martin Ford, Architects of Intelligence: The truth about AI from the people building it, 2018, pp.174-180.
- [2] Hype Cycle for Emerging Technologies 2019, Gartner Inc., G00370466, 06 August 2019.
- [3] 이용주, 문용혁, 박준용, 민옥기, “경량 딥러닝 기술 동향”, 한국전자통신연구원, 2019 Electronics and Telecommunications Trends, 전자통신동향분석 제34권 제2호 2019. 4, pp.40-50.
- [4] Transfer Learning, Lisa Torrey and Jude Shavlik, Handbook of Research on Machine Learning Applications, IGI Global, 2009.
- [5] Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, Stefan Carlsson, CNN Features off-the-shelf: an Astounding Baseline for Recognition, 2014 IEEE conference on computer vision and pattern recognition workshops, March, 2014.
- [6] Jason Yosinski, Jeff Clune, Yoshua Bengio, Hod Lipson, How transferable are features in deep neural networks?, Advances in Neural Information Processing Systems 27, 2014, pp.3320-3328.
- [7] 변영현, 광근창, 자동차 주행환경에서 보행자 분류를 위한 딥러닝 모델의 전이학습 및 성능비교, Journal of KIIT. Vol.16, No.10, Oct. 31, 2, 2018, pp.83-92.
- [8] Holger Roth, Poonam Chitale and Mahendra Roopa. Fast AI Assisted Annotation and Transfer Learning Powered by the Clara Train SDK, NVIDIA DevBlog, April 26, 2019.
- [9] Dipanjan Sarkar, Raghav Bali, Tamoghna Ghosh, Hands-On Transfer Learning with Python Implement advanced deep learning and neural network models using TensorFlow and Keras, Packt Publishing, 2018.
- [10] Jacob Devlin Ming-Wei Chang Kenton Lee Kristina Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, Google AI Language, 2018.
- [11] Helwa, M. K., & Angela P. S. Multi-robot transfer learning: A dynamical system perspective. IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS), IEEE, 2017.