

# 메타 러닝과 방법론 연구 동향

지훈<sup>1</sup>, 이연준<sup>2</sup>

<sup>1</sup>한양대학교 ERICA ICT 융합학부 미디어테크놀로지전공 학부생

<sup>2</sup>한양대학교 컴퓨터공학과 바이오인공지능융합전공 교수

greenpea0819@hanyang.ac.kr, yeonjoonlee@hanyang.ac.kr

## A Survey on Methodology of Meta-Learning

Hoon Ji<sup>1</sup>, Yeon-Joon Lee<sup>2</sup>

<sup>1</sup>Major in Media Technology, Dept of Division of Media, Culture and Design Technology, Hanyang University ERICA

<sup>2</sup>Major in Bio Artificial Intelligence, Dept. of Computer Science and Engineering, Hanyang University

### 요 약

딥러닝은 인간이 탐지하기 어려운 데이터의 특징 및 패턴을 인지하고, 이들을 학습하여 데이터를 분류 및 예측할 수 있는 기술이다. 그러나 딥러닝 모델을 잘 학습시키기 위해서는 고품질의 대용량 데이터와 이들을 처리할 수 있는 방대한 컴퓨터 자원이 요구되는 것이 일반적이다. 따라서 소량의 데이터만이 존재하는 분야나 컴퓨터 자원이 한정되어 있는 상황에서는 딥러닝을 적용하기 어렵다. 본 논문에서는, 소량의 데이터로도 모델을 자신들의 태스크에 맞게 최적화시킬 수 있는 메타 러닝에 대해 소개하고, 메타 러닝 기법들의 방향에 따른 Metric-Based, Model-Based 및 Optimization 기반 모델들에 대해 소개하고, 앞으로 나아가야 할 연구 방향에 대해 제시한다.

### 1. 서론

딥러닝은 데이터로부터 고유의 패턴 및 특징을 추출 및 학습하여 데이터의 클래스를 분류하거나 데이터를 예측할 수 있게 하는 기술을 일컫는다. 데이터로부터 스스로 특징과 패턴을 추출하고 학습할 수 있기에, 딥러닝은 헬스케어, 컴퓨터 비전 및 사이버보안 등 넓은 영역에 걸쳐 그 적용 분야를 확대해 나가고 있다 [1].

그러나 기존의 딥러닝은 데이터의 용량이 클수록, 또 데이터의 품질이 올라갈수록 모델의 성능이 상승하는 것이 일반적이며, 대용량의 데이터와 모델의 깊은 레이어를 처리할 수 있는 방대한 양의 컴퓨터 자원 또한 요구된다 [2]. 즉, 딥러닝을 적용하고자 하는 분야의 데이터가 부족하거나, 데이터의 값이 비싼 경우, 또 컴퓨터 자원이 제한되어 있는 상황에서는 딥러닝을 적용하기가 매우 힘들다.

따라서 본 논문에서는, 적은 데이터만으로도 모델을 빠르게 학습할 수 있는 메타 러닝의 개념에 대해 소개하고 방향에 따른 방법론들을 소개하고자 한다.

### 2. 메타 러닝

‘학습하기 위한 학습’ 이라고도 불리는 메타 러닝은 데이터셋으로부터의 다양한 클래스 조합, 즉 다양한 태스크를 학습함으로써 기존에 보지 못했던 새로운 태스크에 대해서도 그 태스크에 대한 모델을 빠르게 구축할 수 있도록 하는 최적의 메타 지식을 만들어 내는 데에 초점을 둔다.

기존의 딥러닝은 특정 태스크에 대해서만 잘 작동하도록 모델을 훈련하는 데 그 목표가 있다. 이는 주어진 훈련 데이터셋에 대해서 사전에 정의된 하이퍼파라미터 (i.e. 모델 파라미터, 옵티마이저 등)를 활용하여 학습과 최적화 과정을 거쳐 모델의 가중치를 업데이트하여 해당 데이터셋에 대해 잘 작동하도록 학습한다.

그러나 메타 러닝은 많은 태스크에 대해서도 잘 작동할 수 있는 기반 모델을 훈련하는 것에 목표를 둔다. 구체적으로, 메타 러닝에는 미리 정의된 하이퍼파라미터가 없고, 데이터셋 전체를 한 번에 학습하는 것이 아니라 데이터셋을 여러 태스크로 나누어 학습함으로써 기존에 보지 못했던 태스크들도 빠르게 학습할 수 있도록 하는, 다양한 태스크를 아우르는 메타 지식을 찾아내는 방향으로 학습을 진행한다. 이 메타 지식은 이후에 특정 태스크로 최적화할 때 초기 파라미터로써 활용되기에 적은 데이터를 가진 새로운 태스크에 대해서도 메타 지식을 발판 삼아 빠르게 태스크 특정 모델을 구축할 수 있는 것이다. 이때, 어떤 메타 지식을 학습하느냐에 따라 거리 기반 메타 러닝, 모델 기반 메타 러닝, 그리고 최적화 기반 메타 러닝으로 방향이 나뉜다.

메타 러닝에 의해 데이터셋의 명칭과 구조도 기존과 달라지게 되는데, 크게는 메타 트레이닝셋, 메타 테스트셋으로 구성되며, 각각의 데이터셋은 다시 서포트셋과 쿼리셋으로 분할된다. 훈련 과정에서는 메타 트레이닝셋으로부터 서로 다른 클래스 조합을 가진 태스크들이 추출되며, 각 태스크의 서포트셋으로

부터 메타 지식을 훈련한 후, 서포트셋에는 없는 클래스로 구성된 퀴리셋을 통해 중간 퍼포먼스를 측정하고 업데이트를 진행한다. 그리고 테스트 단계에서 메타 테스트셋을 통해 최종 퍼포먼스를 측정한다. 이런 학습 방식을 에피소드 학습법이라고 한다. 이때, 서포트셋에 몇 개의 클래스가 포함되어 있는지에 대한 정보를 N-way 라고 하며, 각 클래스 당 몇 개의 데이터가 있는지에 대한 정보를 k-shot 으로 표기한다.

### 3. 거리 기반 메타 러닝

거리 기반의 메타 러닝의 궁극적인 목표는 다양한 태스크의 특징 벡터를 잘 표현할 수 있는 좋은 특징 공간을 만들어 내는 것이다. 이때 특징 공간은 인공 신경망의 가중치에 대응된다. 공간 기반의 메타 러닝의 방법론 중에는 Siamese-Network, Matching Network, Prototypical Network 등의 방법론이 존재한다.

최근의 연구 중 [3]은 가중치를 공유하는 CNN 모델 두 개로 구성된 Siamese-Network 를 사용하여 손바닥 이미지를 분류하는 메타 러닝 모델을 구축하여 기존의 손바닥 이미지 분류 모델 대부분을 상회하는 성취를 거두었고, [4]는 무작위 그룹 순열과 다층 컨볼루션 네트워크를 결합하여 다변량의 시계열 데이터를 처리할 수 있는 진화된 Prototypical Network 를 설계하였다.

### 4. 모델 기반 메타 러닝

모델 기반의 메타 러닝은 트레이닝셋에 대한 학습 과정에서 발생한 정보를 모델 내/외부 메모리에 저장하고 새로운 태스크를 마주할 시 저장해 놓은 정보를 활용하는 기법이다. 이와 같은 정보를 저장하기 위해 모델의 내, 외부에 별도의 저장장치가 요구된다.

모델 기반의 메타 러닝의 방법론 중에서는 Memory Augmented Neural Network(MANN), Meta Network, Recurrent Meta-Learner 와 같은 방법론이 있다

[5]는 학습 정보를 외부 메모리에 저장하는 모델인 MANN 을 설계한 연구로써, 훈련 시에 현재 입력값에 대응하는 레이블 값을 바로 다음 입력값과 함께 넣어 준다. 이처럼 라벨이 주어지지 않은 데이터를 예측하는 방법을 외부 메모리에 저장하고 후의 예측에 참고 및 업데이트하며 학습을 진행한다. 또한 [6]은 변형된 이미지를 One-Shot 러닝으로 학습하기 위해 Meta-Learner 와 이미지 변형용 서브 네트워크를 결합한 Meta-Network 를 제안하였다.

### 5. 최적화 기반 메타 러닝

최적화 기반의 메타 러닝은 메타 러닝을 내/외부 단계에 걸친 최적화 과정으로 구현한다. 내부 단계는 특정 태스크에 적합하도록 파라미터를 최적화하는 것이 목표이고, 외부 단계는 적은 단계만을 거쳐 세부 태스크에 맞게 업데이트될 수 있는 대표 파라미터(메타 지식)를 학습하는 것을 목표로 한다.

[7]은 초기 가중치 파라미터를 가진 모델을 개별 태스크에 대해 업데이트했을 때, 각 태스크에 대한 손

실의 합이 최소가 되도록 초기 가중치 파라미터를 업데이트하는 방향으로 학습하는 Model-Agnostic Meta-Learning(MAML)을 연구한 논문으로서, 경사도 기반의 학습 방식을 사용한다. 또 이에 대한 개선 연구로서 2019 년 발표된 [8]은 MAML 의 내부 단계의 최적화 스택이 많은 경우, 최적화 경로를 저장하고 차별화하는 과정에서 발생할 수 있는 계산 및 메모리 부담과 경사도 소실을 방지하기 위해 최적화 경로를 배제하고 최적화 결과값만을 사용하여 메타 러닝을 진행하는 Implicit Meta-Learning 을 제시했다.

### 6. 결론

본 논문은 적은 데이터 혹은 제한된 컴퓨터 자원을 가진 상황에서 적용할 수 있는 메타 러닝에 대해 다루었다. 이러한 연구를 통해 메타 지식에 따른 다양한 메타 러닝의 방향에 대해 이해하고, 상황에 맞는 모델을 적용할 수 있을 것으로 기대된다. 추가적으로, [1]에서의 언급 및 예시 논문의 특성과 같이 트레이닝 셋과 테스트셋 간의 유사도가 높은 연구 외에도 트레이닝셋과 테스트셋 사이의 유사도가 낮은 연구가 더욱 활발해져야 한다.

### 7. ACKNOWLEDGEMENT

"본 연구는 2023 년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학지원사업의 연구결과로 수행되었음"(2018-0-00192)

### 참고문헌

- [1] Mike Huisman "A survey of deep meta-learning" 54 4483-4541 2021
- [2] Timothy Hospedales "Meta-Learning in Neural Networks : A Survey" IEEE Transactions on Pattern Analysis and Machine Intelligence 44 5149-5169 2021
- [3] Huikai Shao "Few-Shot Learning for Palmprint Recognition via Meta-Siamese Network" IEEE Transactions on Instrumentation and Measurement" 70 5009812 2021
- [4] Xuchao Zhang "TapNet: Multivariate Time Series Classification with Attentional Prototypical Network" Proceedings of the AAAI Conference of Artificial Intelligence 34 6845-6852 2020
- [5] Adam Santoro "Meta Learning with Memory-Augmented Neural Networks" 33<sup>rd</sup> International Conference on Machine Learning Research New York, USA 2016 1842-1850
- [6] Zitian Chen "Image Deformation Meta-Networks for One-Shot Learning" IEEE/CVF Conference on Computer Vision and Pattern Recognition Long Beach, CA, USA 8680-8689
- [7] Chelsea Finn "Model-Agnostic Meta-Learning for Fast Adaption of Deep Networks" 34<sup>th</sup> International Conference on Machine Learning Sydney, Australia 2017 1126-1133
- [8] Aravind Rajeswaran "Meta-Learning with Implicit Gradients" 33<sup>rd</sup> Conference on Neural Information Processing Systems Vancouver, Canada 2019 113-124