

# Advanced Method for Class-incremental Continual Learning

이름 김태휘

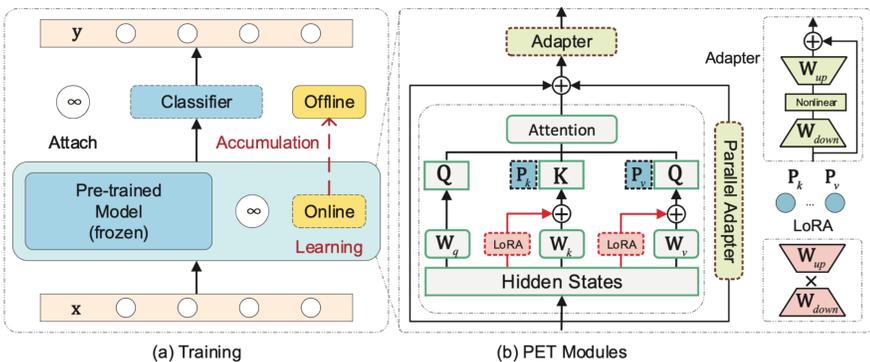
지도교수 유중빈

## 연구 배경 및 목적

본 연구는 여러 Continual Learning 방법론을 조사 및 분석한 초기 연구에 이어 진행된 연구이다. 이전 연구를 통해 사전 학습 모델 기반의 접근 방법이 등장하기 전의 큰 성과를 거둔 연구는 대부분 이미 학습한 모델의 파라미터를 지키는 것에(Stability) 큰 비중을 두었다는 점을 알 수 있었다. 한편, 최근 연구에서는 대규모 데이터셋에 대해 사전 학습한 모델을 이용한 접근 방법이 기존 제안 방법의 성능을 월등하게 뛰어 넘으면서 큰 주목을 받고 있다. 사전 학습 모델 기반의 접근 방법은 기존 접근 방법에 큰 반향을 불러왔지만, 자세히 살펴보면 이미 학습되어 있는 모델의 변화를 최소화한다는 기조에는 변함이 없다는 것을 확인할 수 있다. 이에 본 연구에서는 이러한 최신 연구 동향의 기조를 주요 연구 방향으로 삼고, 추가적인 아이디어를 구성해 성능 향상을 위한 새로운 접근 방법을 모색해보고자 한다.

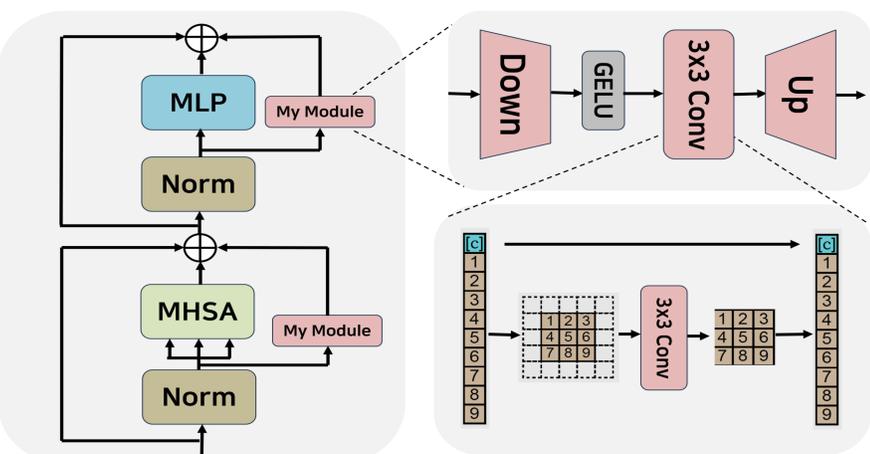
## 연구 진행 과정

본 연구에서는 사전 학습 모델을 기반으로 하는 Continual Learning 방법론 중에서, 서로 다른 데이터로 학습한 모델을 적절하게 통합하는 방식인 Model Merging 방식으로 접근하였다. 또한, 기존 방법은 다양한 Domain의 데이터에 대해 지속적으로 학습하지 않아 사전 학습 모델의 일반화 성능에 의존하는 경향을 보였다. 따라서 본 연구에서는 추가적인 Adapter를 구성해 이를 각 Task에 대해 Parameter-efficient Fine-tuning하는 방법을 사용하는 것이 모델의 견고성에 이점이 있다고 판단했다.



[그림1] LAE의 전체적인 학습 과정

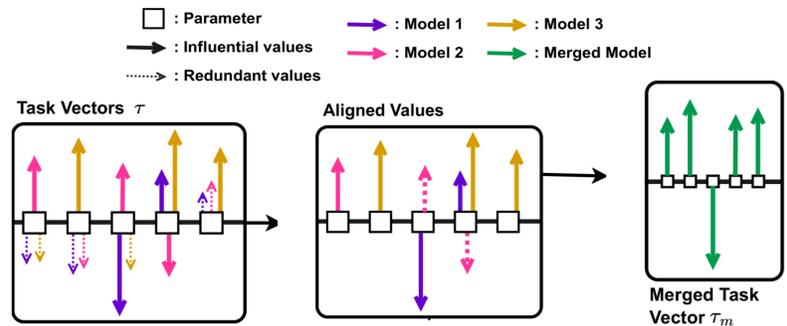
유사한 접근 방법의 LAE 프레임워크를 기반으로, 일부 구성 요소를 개선함으로써 성능을 향상하는 것을 목표로 했다. LAE는 사전 학습 모델의 파라미터를 최대한 보존하기 위해 추가적인 학습에서 제외하고, Prefix, Adapter와 같은 적은 파라미터를 가진 모듈에 대해 학습한다. 또한, 각 Task에 대해 학습된 모듈은 EMA(Exponential Moving Average)에 의해 Merge(Accumulation)되어 지속적인 학습을 가능하게 한다. 본 연구에서 개선하고자 한 부분은, 추가적인 학습이 진행되는 모듈의 구조와 학습된 모듈 간의 Merging 방법이다.



[그림2] 이미지 처리에 적합하게 설계한 새로운 모듈(Adapter): Transformer 블록에 대해서 MSA 뿐만 아니라 MLP와도 병렬적으로 연결

기존 연구에서 사용된 모듈들은 주로 LLM 모델을 Fine-tuning 하기 위해 사용되었다. 본 연구에서는 Vision 모델에 대한 최적화를 위해 두 선형 레이어 사이에 Convolutional 신경망을 추가한 새로운 구조의 Adapter 모듈을 사용했다. 또한, 모든 Block에서 MSA Layer에 대해서만 연결되어 있던 모듈을 MLP Layer에 대해서도 추가적으로 연결하도록 구성했다.

이후에는 Task에 대한 학습이 끝난 이후의 학습한 모듈(Online)과 저장되어 있는 모듈(Offline)의 병합 과정을 개선하고자 했다.

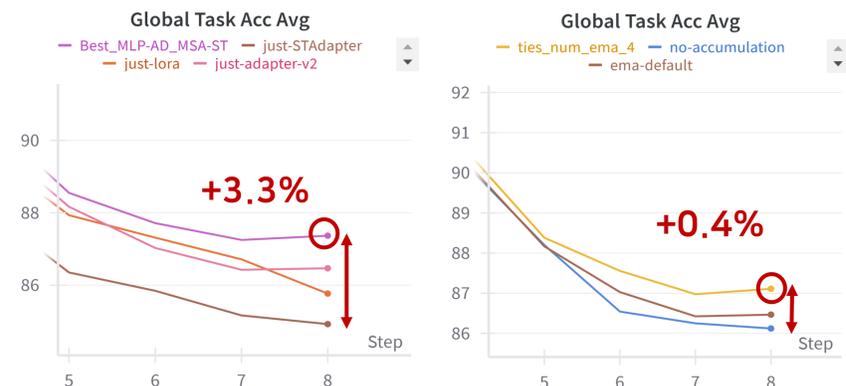


[그림3] TIES-Merging: Sign-Conflict를 해결하기 위해 제안된 Model Merging 방법

사전 학습 모델과 Fine-tuning된 모델들의 Weight를 합치는 과정에서 그 값이 서로 다른 방향으로 충돌(Sign-Conflict)하여 성능이 감소하는 현상을 해결하고자 제안된 TIES-Merging 방법을 활용했다. 이는 업데이트 값이 작은 값을 0으로 만들고, 모델 간의 업데이트 방향을 결정한 뒤, 해당 방향을 갖는 값만으로 평균 내어 모델을 병합한다. 이러한 방법은 다양한 데이터 분포에 대해서 학습하는 본 연구에 활용하기 적합했고, EMA에 대비해 견고성과 안정성을 향상시킬 수 있었다.

## 결과 및 분석

실험에 사용된 모델은 ImageNet-21k 데이터셋에 대해 사전 학습한 ViT-B/16이며, 사용된 데이터셋은 CIFAR100이다. 데이터셋의 첫 10개의 클래스를 학습한 뒤, 총 9번의 Incremental Step(Task)로 나누어 학습하였다. 아래 그림에서의 정확도는 각 Task까지 관찰된 모든 Class에 대한 정확도의 평균이다.



[그림4] LAE 프레임워크에서 개선한 아이디어에 대한 실험: (좌) 새로운 모듈 구조 적용 (우) 모듈 간의 병합 과정 개선

새로운 모듈을 LAE에 적용한 결과, 기존 연구의 정확도 (85.59%)에 비해 약 1% 상승한 86.41%를 달성했고 MLP와 병렬적으로 연결하지 않았을 때에 비해 3.3%의 정확도 향상을 기록했다. 제안한 병합 과정을 통해서 기존 연구(85.59%)에 대비 약 0.4%의 성능 향상(85.99%)을 기록했다. 향후 연구 계획으로는 두가지 개선 방안을 동시에 적용하는 것을 목표로 하고 있다. 두 방법 모두 기존 방법에 비해 성능 향상을 기록한 것으로 미루어 보아 추가적인 성능 향상이 기대된다.

## 오픈소스 URL

<https://github.com/h-wi/my-new-LAE>  
Codes are built on official source code of LAE

