

Few-shot Class Incremental Learning

(소수샷 클래스 증가 학습)

요약

Few-shot Class Incremental Learning (FSCIL)은 적은 양의 데이터 (few-shot)으로 하나의 모델을 여러 태스크에 걸쳐 학습시키는 (class incremental) 방식이다. 적은 양의 데이터를 사용한 모델의 학습은 오버피팅을 야기하기 쉽고, 여러 태스크에 걸친 모델의 학습은 기존 태스크에서 학습한 지식을 잊어버리는 파괴적 망각(catastrophic forgetting)이 발생하기 쉽다. 본 연구에서는 기존에 학습한 지식은 잊어버리지 않은 채, 새로운 지식을 잘 습득하는 학습 방법에 대해 연구하고자 한다. 또한, 소스코드를 [GitHub](#)에 공개하여 향후 연구에 도움이 되고자 한다.

1. 서론

1.1. 연구배경

현대의 딥러닝 기술은 컴퓨터 비전과 자연어 처리, 강화 학습 등 수많은 분야에서 매우 높은 성능을 보이고 있다. 그 중, 컴퓨터 비전 분야에서는 ImageNet [1] classification 문제에서 top-1 accuracy 90.94%를 달성하는 등 [2] 아주 높은 성능을 보이고 있다.

그러나 이러한 높은 성과는 현실 세계의 문제를 해결하기에는 다소 동떨어진 환경에서 발생한 것이다. 예시로 들었던 ImageNet [1]의 경우, 1000 개의 클래스에 걸쳐 약 130 만장의 이미지를 사용하는데, 현실세계에서 이런 데이터는 수집하고 클래스를 표기하는 데 수많은 비용이 든다.

또한, 기존의 딥러닝 기술들은 대개 하나의 모델이 하나의 태스크 혹은 데이터셋을 학습하게 되는데, 새로운 데이터가 추가 되는 경우 모델을 새로 학습해야하기 때문에 비효율적이다.

위의 두 문제를 해결하기 위해 최근의 딥러닝 기술은 few-shot learning 과 lifelong learning 을 주제로 활발히 연구되고 있다. Few-shot learning 은 데이터가 적은 환경에서 모델을 효과적으로 학습하기 위한 기술이다. 대개 현존하는 딥러닝 기술들은 많은 데이터를 가정하고 있는데, 적은 수의 데이터가 주어지게 되면 오버피팅(over-fitting) 현상이 발생하게 된다.

한편, lifelong learning 은 하나의 모델이 여러개의 태스크를 순차적으로 학습시키기 위한 방법을 연구한다. 전통적인 딥러닝 방법은 새로운 태스크를 학습하기 위해 fine-tuning 방법을 적용하는데, 이러한 경우 새로운 태스크를 학습하게 되면 기존의 태스크에 대한 지식은 잊어버리는 catastrophic forgetting 이 발생하게 된다. Lifelong learning 은 태스크를 명시하는

task incremental learning 과 태스크를 명시하지 않는 class incremental learning 으로 나눌 수 있다.

1.2. 연구목표

본 연구에서는 앞서 언급했던 few-shot learning 과 lifelong learning 중 class incremental learning 을 조합하여 적은 데이터로도, 여러개의 태스크를 효과적으로 학습할 수 있는 few-shot class incremental learning (FSCIL)에 대해 연구하고자 한다.

2. 관련 연구

2.1. Few-shot learning

Few-shot learning 은 적은 수의 학습 데이터로도 모델이 좋은 성능을 내도록 하는 학습 방법이다. Few-shot learning 은 주로 모델을 pre-training 시키거나, meta-learning 방법을 이용하는 데, meta-learning 방법에는 network-based, optimization-based 그리고 metric-learning-based 방법으로 구분할 수 있다. Network-based 방법은 네트워크나 별도의 메모리를 사용하여 가중치를 생성하거나, 갱신 또는 예측을 하는 방법이다. Optimization-based 방법은 두 단계의 optimization 절차를 거쳐 가중치를 초기화 하거나 갱신하는 방법과 같은 학습 과정 자체를 학습한다. 이렇게 해서 최적의 학습 과정으로 few-shot 데이터를 학습을 시작한다 [3]. Metric-learning-based 방법은 기존에 학습했던 지식을 embedding space 에 두어 비슷한 클래스를 가까이 두도록 하는 방법이다 [4, 5, 6, 7].

2.2. Class incremental learning

Class incremental learning 은 여러 태스크에 걸쳐 새로운 클래스가 주어질 때, 기존에 학습했던 클래스에 대한 성능은 유지한 채 새로운 클래스를 잘 배우도록 하는 학습 방법이다. 이때, 기존에 학습했던 클래스를 잊어버리는 현상을 catastrophic forgetting 이라고 한다. Catastrophic forgetting 을 줄이기 위한 방법으로는 generative-replay-based, expansion-based, regularization-based 방법으로 구분할 수 있다. Generative-replay-based 방법은 GAN [8]이나 VAE [9]와 같은 생성 모델을 이용하여 기존에 학습했던 데이터를 재생하는 방법이다 [10]. Expansion-based 방법은 모델의 구조 자체를 늘려 catastrophic forgetting 을 방지하는 방법이다 [11]. Regularization-based 방법은 모델의 가중치를 규제함으로써 기존 태스크에서 학습한 지식을 보존하고자 한다 [12].

3. 프로젝트 내용

3.1. 문제 정의

데이터셋의 전체 클래스는 base 클래스와 novel 클래스로 나눌 수 있다. 태스크가 T_0, T_1, T_2, \dots 의 순서로 주어진다. 0 번째 태스크 T_0 에는 base 클래스로 구성되어있고, 해당 태스크의 데이터를 이용하여 모델을 pre-train 한다. 각 태스크 T_t ($t \geq 1$)에서는, 주어진 데이터를 학습 데이터에 해당하는 support set S^t 와 query set Q^t 로 나눌 수 있다. Support set에서는 novel 클래스 중 N 개의 클래스가 각각 K 개의 데이터를 가지고 있는데, 이를 N-way, K-shot 환경이라 정의한다. t 번째 태스크의 support set에 대한 학습이 완료되면 0 번째 태스크의 클래스부터 t 번째 태스크의 클래스 모두에 걸친 query set $Q^{<=t}$ 으로 학습된 모델을 평가한다.

3.2. 사용 데이터셋

모델을 테스트 하기 위해 minilImageNet [7], tieredImageNet [14], CUB [15] 3 가지의 데이터셋을 사용할 예정이다.

minilImageNet은 기존 ImageNet의 축소버전으로 100 개의 클래스당 600 개의 이미지로 구성된다. tieredImageNet 데이터셋은 기존 ImageNet 데이터셋을 hierarchy를 기반으로 600 개의 클래스로 묶은 것이다. 각 클래스당 평균 1281 개의 이미지로 구성되어 있다. CUB 데이터 셋은 여러 종류의 새를 의미하는 200 개의 클래스와 각 클래스당 평균 60 개의 이미지로 구성된다.

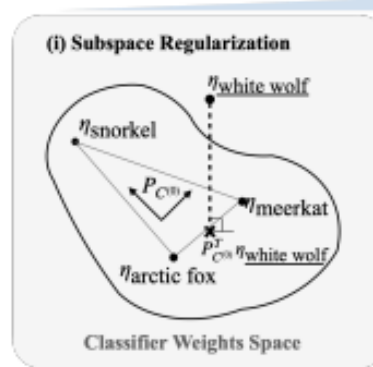
3.3. 기존 연구 분석

본 프로젝트를 진행하기 앞서 선행 연구인 "Subspace regularizers for few-shot class incremental learning" [13]을 소개하고 분석해보고자 한다.

3.3.1. 제시된 방법

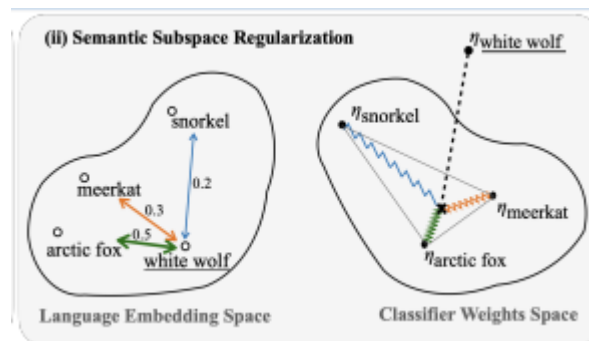
본 논문에서는 FSCIL을 다루기 위한 regularization 방법으로 subspace regularization, semantic subspace regularization, linear mapping을 소개하고 있다.

3.3.1.1. Subspace regularization



Novel 클래스에 대한 overfitting 을 방지하기 위해 base class 의 weight vector 를 통해 만들어지는 subspace 에 feature 들을 두는 기법이다. novel 클래스의 값을 base 클래스 subspace 에 정사영 시켜준 것을 subspace target 으로 정의하며 이 값과 원래 novel 클래스 vector 간의 유클리디안 거리 값을 regularization 할 때 사용해주다면 지나치게 novel 클래스위주로 학습되는 것을 방지하는 효과가 있다.

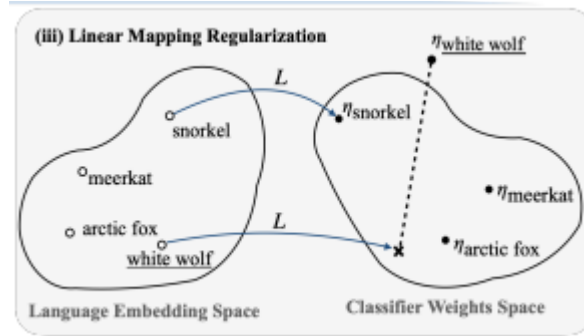
3.3.1.2. Semantic subspace regularization



Subspace Regularization 기법에서는 이미 있는 base 클래스와 novel 클래스의 관계에 대해 알 수 있는 정보가 부족하여 classification 할 때 손해를 볼 수도 있다는 단점이 있다. 위 이미지를 보면 novel 클래스인 white wolf 는 base 클래스에 있는 요소들 중 arctic fox 와 유사하다는 것을 알 수 있는데 이러한 정보가 누락되었음을 의미한다.

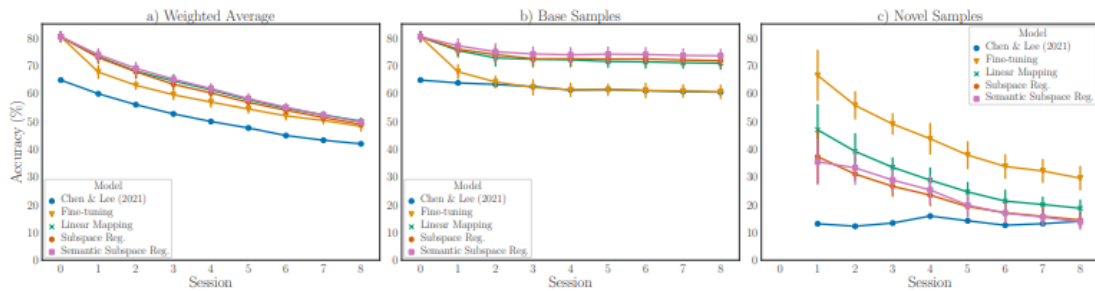
따라서 language embedding 을 통해 클래스간 관계를 파악하고 이 값을 softmax 함수에 적용시켜 나온 결과를 semantic target 이라고 정의한다. 이 semantic target 과 novel class vector 간의 유클리디안 거리 값을 regularization 할 때 사용한다면 기존 model 에 비해 성능 향상이 가능하다.

3.3.1.3. Linear mapping



Language embedding vector 와 classifier weight 를 바로 mapping 시키는 방식이 효율적인 학습에 효과가 있다는 연구 결과가 존재한다. 이 방식을 few-shot learning 을 할 때 적용시켜 줌으로써 기존의 방식 대비 성능 향상이 가능하다.

3.3.2. 발생가능한 문제점



위 자료는 기존 방식인 fine-tuning 의 정확도와 관련 연구에서 제시된 방식들의 정확도를 비교한 표이다. x 축은 session 을 의미하며 1 session 이 증가할 때 마다 새로운 novel 클래스를 추가적으로 학습하게 된다. 위 데이터 기준 base 클래스가 novel 클래스에 비해 더 많기 때문에 클래스들 수 간 가중치를 고려해 평균을 낸다면 기존 방식에 비해 성능향상이 존재한다는 것은 사실이다. 그러나 base 클래스, novel 클래스 각각에 대하여 살펴본다면 base 클래스에 관련해서는 성능 향상이 있었지만 novel 클래스에 대해서는 오히려 성능 저하가 있었음을 알 수 있다.

본 연구에서는 성능 저하가 일어난 부분을 보완하여 보다 더 높은 정확도로 학습할 수 있는 모델로 개선한다.

4. 향후 일정

3/10 ~ 3/28 : 선행 연구 조사 및 실험 결과 reproduce

3/29 ~ 4/12: 비교군 선정 및 reproduce

4/13 ~ 4/27: 기존 연구들의 문제점 분석 및 해결 방안 도출

4/28 : 중간보고서 제출

4/29 ~ 6/15 : 아이디어 구현 및 실험으로 효과 입증

6/16 : 최종보고서 제출

5. 결론 및 기대효과

기존 few-shot class incremental learning 의 연구들의 문제점들을 파악하고 다음 두가지 측면에서 개선된 모델을 제안하는 것이 목표이다. 첫째, 새로운 novel 클래스에 대해서 더 잘 학습하여 새로운 태스크에 대해 좋은 성능을 보이고, 둘째, 기존에 학습했던 지식을 최대한 보존하여 catastrophic forgetting 을 완화한다. 그 다음, 본 논문의 방법이 기존 연구보다 개선되었음을 실험을 통해 입증하고, 소스코드를 공개하여 향후 연구에 도움이 되고자 한다.

6. 참고 문헌

- [1] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). IEEE.
- [2] Wortsman, M., Ilharco, G., Gadre, S. Y., Roelofs, R., Gontijo-Lopes, R., Morcos, A. S., ... & Schmidt, L. (2022). Model soups: averaging weights of multiple fine-tuned models improves accuracy without increasing inference time. *arXiv preprint arXiv:2203.05482*.
- [3] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In ICML, 2017.
- [4] Gregory Koch, Richard Zemel, and Ruslan Salakhutdinov. Siamese neural networks for one-shot image recognition. In ICMLW, 2015
- [5] Jake Snell, Kevin Swersky, and Richard Zemel. Prototypical networks for few-shot learning. In NIPS, 2017.
- [6] Flood Sung, Yongxin Yang, Li Zhang, Tao Xiang, Philip H.S. Torr, and Timothy M. Hospedales. Learning to compare: Relation network for few-shot learning. In CVPR, 2018.

[7] Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Koray Kavukcuoglu, and Daan Wierstra. Matching networks for one shot learning. In NIPS, 2016.

[8] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.

[9] Diederik P. Kingma, Max Welling, Auto-encoding Variational Bayes, In IEEE Conference on Learning Representations, 2014.

[10] Chensen Wu, Luis Herranz, Xialei Liu, Yaxing Wang, Joost van de Weijer, Bogdan Raducanu, Memory replay GANs: learning to generate images from new categories without forgetting. In NeurIPS 2018.

[11] Andrei A Rusu, Neil C Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, and Raia Hadsell. Progressive Neural Networks. arXiv preprint arXiv:1606.04671, 2016.

[12] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. Overcoming Catastrophic Forgetting in Neural Networks. Proceedings of the National Academy of Sciences, 2017.

[13] Afra Feyza Akyürek, Ekin Akyürek, Derry Wijaya, Jacob Andreas, Subspace regularizers for few-shot class incremental learning, In ICLR 2022.

[14] Mengye Ren, Eleni Triantafillou, Sachin Ravi, Jake Snell, Kevin Swersky, Joshua B Tenenbaum, Hugo Larochelle, and Richard S Zemel. Meta-learning for semi-supervised few-shot classification. arXiv preprint arXiv:1803.00676, 2018.

[15] Welinder P., Branson S., Mita T., Wah C., Schroff F., Belongie S., Perona, P. "Caltech-UCSD Birds 200". California Institute of Technology. CNS-TR-2010-001. 2010