

# Novel 클래스의 가중치를 재사용하는 부분 공간 규제를 통한 소수샷 클래스 증강 학습

서주원 박민호 박경문  
경희대학교 컴퓨터공학과  
{jwseo001, pmhmist, gmpark} @ khu.ac.kr

## Few-Shot Class Incremental Learning via Subspace Regularization with Reusing Novel-Class Weights

Juwon Seo Minho Park Gyeong-Moon Park  
Dept. of Computer Science and Engineering, Kyung Hee University, Republic of Korea

### 요 약

소수샷 클래스 증강 학습 (few-shot class incremental learning)은 소수샷 학습 (few-shot)과 클래스 증강 학습 (class incremental learning)의 두 분야를 적용한 학습 기법이다. 적은 양의 데이터는 모델의 오버피팅을 야기하기 쉽고, 클래스를 점진적으로 늘려가면서 학습을 진행하다 보면 이전 클래스에 대한 지식을 잊어버리기 쉽다. 기존 연구에서는 이러한 문제를 해결하기 위해 부분 공간을 위한 규제를 제안했지만 base 클래스의 수가 적어지는 경우 성능이 하락하게 된다. 이에 본 연구에서는 이전에 학습된 novel 클래스의 가중치를 재사용하여 base 클래스의 수가 적은 상황에서도 보다 나은 성능을 보이도록 하는 방법을 제시한다.

### 1. 서 론

현대의 딥러닝 기술은 컴퓨터 비전과 자연어처리, 강화 학습 등 수많은 분야에서 매우 높은 성능을 보여주고 있다. 그러나 이러한 높은 성과는 현실 세계의 문제를 해결하기에는 다소 동떨어진 환경에서 발생한 것이다. 현실세계에서는 데이터를 수집하고 클래스를 표기하는 데 수많은 비용이 들며, 예술 분야처럼 필연적으로 사용 가능한 데이터의 수가 적은 경우가 있다. 또한, 기존의 딥러닝 기술들은 새로운 데이터가 추가되는 경우 모델을 새로 학습해야 하기 때문에 시간상으로도 저장 공간상으로도 비효율적이다.

위의 두 문제를 해결하기 위해 최근의 딥러닝 기술은 소수샷 학습 (few-shot learning)과 평생 학습 (lifelong learning)을 주제로 활발히 연구되고 있다. 소수샷 학습은 다음 문제를 해결하고자 하는 학습 기법이다. 적은 양의 데이터를 활용하여 모델을 학습시키면 주어진 데이터가 실제 데이터의 분포를 온전히 반영하지 않은 편향된 분포를 갖게 된다. 이런 편향된 분포를 이용한 학습은 모델로 하여금 주어진 적은 데이터에 대해서만 좋은 성능을 보이게 하여 오버피팅 현상을 발생시킨다. 한편, 하나의 모델에 지속적으로 데이터를 제공하여 학습을 진행하는 평생 학습 (lifelong learning)에서는 새로운 데이터가 주어질 때 기존에 학습된 가중치가 변형되기 때문에 이전에 학습했던 태스크를 제대로 수행하지 못하는 파괴적 망각 (catastrophic forgetting)을 해소하고자 노력한다. 평생 학습은 또한, 태스크를 명시하는 태스크 증강 학습 (task incremental learning)과 태스크를 명시하지 않는 클래스 증강 학습 (class incremental learning)으로 나눌 수 있다.

소수샷 학습과 클래스 증강 학습을 조합한 소수샷 클래스 증강 학습 (few-shot class incremental learning,

FSCIL)에서는 이전 태스크를 원활히 수행함과 동시에 적은 양의 데이터를 가지고 현재 태스크 또한 잘 수행해야 하기 때문에 보다 난해한 환경에서 모델을 학습시킨다.

소수샷 증강 학습을 연구한 기존의 방법 [1]에서는 사전 학습된 클래스를 활용한 부분 공간을 사용하여 가중치를 규제하는 여러가지 방법을 제시했다. 그러나 이러한 방법은 부분 공간에 의존적이어서 부분 공간이 어떻게 형성되는지에 따라 성능이 크게 변화하게 된다.

본 연구에서는 이러한 문제점을 해결하기 위해 새로 학습한 클래스 또한 부분 공간을 형성하는 데 활용하여 적은 수의 base 클래스로도 좋은 성능을 보이는 학습 방법을 제시하고자 한다.

### 2. 기존 연구

본 논문의 방법을 제시하기에 앞서, 소수샷 클래스 증강 학습을 다룬 기존 연구 [1]을 소개하고자 한다.

#### 2.1. 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 구성 되어있고, 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  $Q^{st}$ 으로 학습된 모델을 평가한다.  $S^t$ 의 클래스인  $C(S) = \{y | (x, y) \in S\}$ 로

정의하며  $\mathcal{C}^{(t)} = \mathcal{C}(S^{(t)})$  이다.

## 2.2. Pre-training on base classes

우선 base 클래스에 대해 아래의 손실함수를 최대화하는 방향으로 학습을 진행했다.

$$\mathcal{L}(\eta, \theta) = \frac{1}{|S^{(0)}|} \sum_{(x,y) \in S^{(0)}} \log \frac{\exp(\eta_y^T f_\theta(x))}{\sum_{c \in \mathcal{C}^{(0)}} \exp(\eta_c^T f_\theta(x))} - \alpha(\|\eta\| + \|\theta\|)^2$$

여기서  $f$ 는 특징 추출기이고 매개변수  $\theta$ 로 표현된다.  $\eta_c$ 는 클래스  $c$ 에 대한 가중치이다.

## 2.3. Fine-tuning with novel classes

Base 클래스에 대한 학습이 종료되면, 특징 추출기는 고정시킨 채, 여러개의 태스크로 이루어진 novel 클래스에 대해서 학습을 진행한다:

$$\mathcal{L}(\eta) = \frac{1}{|S^{(t)}|} \sum_{(x,y) \in S^{(t)}} \log \frac{\exp(\eta_y^T f_\theta(x))}{\sum_{c \in \mathcal{C}^{(0)}} \exp(\eta_c^T f_\theta(x))} - \alpha(\|\eta\|)^2 - \beta \left( \sum_{t' < t} \sum_{c \in \mathcal{C}^{(t')}} \|\eta_c^{t'} - \eta_c\|^2 \right) - \gamma R_{new}^{(t)}(\eta)$$

이때  $R_{new}^{(t)}$ 는 새로 학습될 novel 클래스의 가중치를 규제하기 위해 도입되었으며 후술할 subspace regularization, semantic subspace regularization 또는 linear mapping 이 적용된다.

### 2.3.1. Subspace regularization

Base 클래스의 가중치를 이용하여 부분 공간을 형성한다. 그 다음 새로 학습될 novel 클래스의 가중치를 base 클래스의 가중치로 형성된 부분 공간에 정사영 시켜준 것을 subspace target  $m_c$ 라고 정의한다:

$$m_c = P_{\mathcal{C}^{(0)}}^T \eta_c$$

Subspace target  $m_c$ 와 novel 클래스의 가중치 사이의 유클리디안 거리를 이용하여 새로 학습될 novel 클래스의 가중치가 base 클래스로 형성된 부분 공간에 위치하도록 규제한다.

$$R_{new}^{(t)} = \sum_{c \in \mathcal{C}^{(t)}} \|\eta_c - m_c\|^2$$

### 2.3.2. Semantic subspace regularization

자연어 클래스는 많은 의미를 내포하고 있다. 앞서 소개된 subspace regularization에서는 같은 부분 공간에 새로운 novel 클래스의 가중치를 위치시키는 것이 목적이지만, 어디에 위치하는지는 고려하지 않았다. 이에 자연어 임베딩을 이용하여 의미론적으로 유사한 base 클래스에 가까이 새로운 novel 클래스의 가중치를 위치시키고자 한다.

$$l_c = \sum_{j \in \mathcal{C}^{(0)}} \frac{\exp(e_j \cdot e_c / \tau)}{\sum_{j \in \mathcal{C}^{(0)}} \exp(e_j \cdot e_c / \tau)} \eta_j$$

$$R_{new}^{(t)} = \sum_{c \in \mathcal{C}^{(t)}} \|\eta_c - l_c\|^2$$

여기서  $l_c$ 는 semantic target 으로, base 클래스 임베딩과의 코사인 유사도에 소프트맥스를 적용하여 계산된다. 이러한 semantic target 과 novel 클래스 가중치 간의 유클리디안 거리를 줄이는 방향으로 규제한다.

### 2.3.3. Linear mapping

Zero-shot 학습 분야의 몇몇 연구[2, 3, 4, 5]에서 자연어 임베딩과 분류기의 가중치를 바로 사상하는 방식이 학습에 효율적이라는 결과가 있었다. 이를 소수샷 클래스 증강 학습에 적용시키면 아래와 같다.

$$L^* = \min_L \sum_{j \in \mathcal{C}^{(0)}} \|\eta_j - L(e_j)\|^2$$

$$R_{new}^{(t)} = \sum_{c \in \mathcal{C}^{(t)}} \|\eta_c - L^*(e_c)\|^2$$

Base 클래스의 임베딩을 입력으로 하여 클래스의 가중치로 사상하도록 하는 선형 함수  $L$ 을 학습시켜 최적의 선형 함수  $L^*$ 을 얻는다. 그 다음 새로운 novel 클래스의 임베딩을  $L^*$ 에 입력으로 주어 실제 novel 클래스 가중치와 유클리디안 거리를 최소화하도록 규제한다.

## 3. 학습된 가중치의 재사용

이번 장에서는 기존 연구의 문제점을 파악하고 (3.1 절), 해당 문제점을 해결하기 위한 방안을 제시한다 (3.2 절).

### 3.1. Base 클래스의 수에 따른 성능 변화

기존 연구에서는 base 클래스의 수가 60, novel 클래스의 수는 5로, novel 클래스에 비해 base 클래스의 수가 비교적 많아 신뢰할 만한 부분 공간이 생성되었다. 그러나 이러한 접근 방법은 base 클래스의 수가 적어질 때 큰 성능 저하를 야기한다.

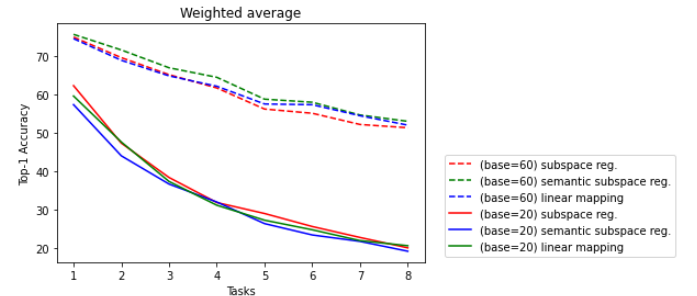


그림 1 Base 클래스의 수에 따른 성능 비교

위 그림 1은 기존 논문의 규제 기법을 base 클래스의 수가 {60, 20}, novel 클래스의 수가 {5, 10}일 때 비교한 것이다. Base 클래스의 수가 적어지면서 성능이 급격하게 하락하는 것을 확인할 수 있다.

### 3.2. 이전 가중치를 재사용하는 규제

이번 절에서는 위에서 서술한 문제를 해결하기 위해 novel 클래스를 학습할 때, base 클래스와 함께 이전 세션의 novel 클래스를 함께 활용하는 방법을 제시한다.

#### 3.2.1. Subspace regularization with reusing novel weights

Subspace regularization에서 기존의 novel 클래스를 재사용하기 위해 식을 아래와 같이 재정의한다.

$$m'_c = P_{\mathcal{C}^{(<t)}}^T \eta_c$$

$$R_{new}^{(t)} = \sum_{c \in \mathcal{C}^{(t)}} \|\eta_c - m'_c\|^2$$

여기서, novel 클래스의 가중치를 정사영시킬 부분 공간의 직교 기저를  $P_{\mathcal{C}^{(<t)}}$ 로 하여 base 클래스와 이전 novel 클래스를 활용하여 형성했다.

### 3.2.2. Semantic subspace regularization with reusing novel weights

Semantic subspace regularization 에서 이전의 novel 클래스를 재 사용하기 위해 식을 아래와 같이 재정의한다.

$$l'_c = \sum_{j \in C^{(<t)}} \frac{\exp(e_j \cdot e_c / \tau)}{\sum_{j \in C^{(<t)}} \exp(e_j \cdot e_c / \tau)} \eta_j$$

$$R'_{new}^{(t)} = \sum_{c \in C^{(t)}} \|\eta_c - l'_c\|^2$$

여기서 클래스 간의 의미론적 유사도를 측정하는데, 새로운 novel 클래스를 base 클래스에다가 이미 학습된 이전의 novel 클래스와도 함께 측정하도록 하였다.

### 3.2.3. Linear mapping with novel weights

Linear mapping 에서 기존의 novel 클래스를 재 사용하기 위해 식을 아래와 같이 재정의한다. 현재 의미론적 임베딩 공간에서 실제 가중치에 사상하는 선형 함수를 학습시키게 되는데, 이때 base 클래스와 함께 이미 학습된 이전 novel 클래스의 가중치와 그 클래스의 임베딩 모두 활용하도록 하였다.

$$L^* = \min_L \sum_{j \in C^{(<t)}} \|\eta_j - L(e_j)\|^2$$

$$R'_{new}^{(t)} = \sum_{c \in C^{(t)}} \|\eta_c - L^*(e_c)\|^2$$

## 4. 실험

실험에는 minImageNet[6]을 사용한다. Base 클래스의 수는 20, 각 세션에서 novel 클래스의 수는 10 으로 설정했다. 학습은 ResNet18[7]에서 진행되었다. 학습된 모델에 대한 평가는 top-1 accuracy 를 통해 이루어졌다. 규제 방법 중 semantic subspace regularization 에서 사용하는 GloVe[8]를 사용한 [1]의 임베딩을 사용했다.

Session	Reuse novel	1	2	8	Avg.
Model					
Subspace Reg.	X	<b>62.27</b>	47.30	18.98	34.14
	O	<b>62.27</b>	<b>47.45</b>	19.62	<u>34.25</u>
Linear Mapping	X	<u>59.87</u>	<u>47.60</u>	20.58	32.69
	O	<u>59.87</u>	46.80	<u>21.12</u>	<b>34.33</b>
Semantic Subspace Reg.	X	57.33	44.00	19.14	33.26
	O	57.33	44.50	<b>21.56</b>	33.58

표 1 Base 클래스의 수가 20 일 때 이전 novel 사용 유무에 따른 성능 비교

위 표 1 에서 보이듯이 novel 클래스의 가중치를 재사용한 경우가 그렇지 않은 경우보다 약간의 성능 향상을 보였다.

## 5. 결론 및 한계점

본 연구에서는 소수샷 증강 학습에서 기존에 존재하는 가중치를 이용한 규제 방법이 base 클래스의 수가 적은 환경에서는 큰 성능 저하가 발생했음을 밝혀냈다. 그리고, 이에 대한 해결 방안으로 이미 학습된 novel 클래스의 가중치를 이용하여 부분 공간을 다시 정의해 현재 세션에서의 novel 클래스를 학습하도록 하였고 성능 향상을 보였다.

본 연구의 한계점 또한 존재한다. 우선, 성능이 향상된 정도가 상당히 극소 하다는 점이다. 이미 학습된 novel 클래스를 재사용하는 보다 나은 방법을 강구해보아야 한다. 또한, 현재 특징 추출기는 고정되어 있는 상태이기 때문에 이를 활용한다면 보다 나은 성능을 기대할 수 있을 것이다.

## 참고 문헌

- [1] Afra Feyza Akyürek, Ekin Akyürek, Derry Wijaya and Jacob Andreas, Subspace Regularizers for Few-Shot Class Incremental Learning, International Conference on Learning Representations, 2022.
- [2] Debasmit Das and CS George Lee. Zero-shot image recognition using relational matching, adaptation and calibration. In 2019 International Joint Conference on Neural Networks (IJCNN), pp. 1–8. IEEE, 2019.
- [3] Richard Socher, Milind Ganjoo, Hamsa Sridhar, Osbert Bastani, Christopher D Manning, and Andrew Y Ng. Zero-shot learning through cross-modal transfer. In Advances in Neural Information Processing Systems, 2013.
- [4] Farhad Pourpanah, Moloud Abdar, Yuxuan Luo, Xinlei Zhou, Ran Wang, Chee Peng Lim, and Xi-Zhao Wang. A review of generalized zero-shot learning methods. arXiv preprint arXiv:2011.08641, 2020.
- [5] Bernardino Romera-Paredes and Philip Torr. An embarrassingly simple approach to zero-shot learning. In International conference on machine learning, pp. 2152–2161. PMLR, 2015.
- [6] Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Koray Kavukcuoglu and Daan Wierstra, Matching networks for one shot learning, arxiv preprint arXiv:1606.04080, 2016.
- [7] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun, Deep Residual Learning for Image Recognition, In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, 2016.
- [8] Jeffrey Pennington, Richard Socher and Christopher D Manning, Glove: Global vectors for word representation, In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp. 1532–1543, 2014.