

# 효율적인 타겟이 정해진 퓨샷 분류를 위한 베이스 클래스 선별

정 경 중\*, 신 원 용<sup>o</sup>

## Base Class Selection for Efficient Target-Guided Few-Shot Classification

Kyeong-Joong Jeong\*, Won-Yong Shin<sup>o</sup>

### 요 약

퓨샷 분류는 레이블링이 되지 않은 샘플을 아주 적은 샘플만 주어진 새로운 클래스로 분류하는 것이다. 다양한 베이스 클래스로 구성된 베이스셋을 활용한 여러 퓨샷 분류 방법들이 소개되었지만, 어떤 데이터를 베이스셋으로 활용하는 것이 더 효과적인지에 대한 문제는 거의 다루어지지 않았다. 본 레터에서는 타겟이 정해진 퓨샷 분류 문제를 효율적으로 해결하기 위해 유사도 기반 베이스 클래스를 선별하는 간단하고도 효과적인 방법을 제안한다. 실제 데이터셋을 사용한 실험을 통해 제안하는 베이스 클래스 선택 방법의 우수성을 검증한다.

**Key Words** : Base class selection, base set, efficiency, few-shot classification, similarity-based selection

### ABSTRACT

Few-shot classification aims to classify unlabeled instances into novel classes given only few-labeled instances. Although various few-shot classification methods have been developed by making use of a large amount of data, called a base set, consisting of instances with base classes, it has been largely underexplored how to select more useful base classes in a given base set for solving a specific few-shot

task. In this letter, to solve target-guided few-shot classification, we propose a simple yet efficient base set selection method based on our similarity measure. Through experiments using a real-world dataset, we demonstrate the superiority of our proposed base set selection method.

### 1. 서 론

퓨샷 분류 (few-shot classification) 과제는 레이블링이 되지 않은 샘플 (쿼리 셋; query set)을 새로운 클래스 (novel class)의 아주 적은 샘플 (서포트셋; support set) 만을 이용하여 분류하는 문제를 의미한다. 이러한 퓨샷 분류 문제 해결을 위해서는 일반적으로 다양한 베이스 클래스 (base class)에 속한 수많은 데이터로 구성된 데이터셋 (베이스 셋; base set)을 통한 학습을 필요로 한다. 대부분의 퓨샷 분류와 관련된 기존의 연구<sup>[1]</sup>에선 고정된 베이스 셋을 기반으로 이를 어떻게 활용할지에만 초점을 두었기 때문에, 베이스 셋 자체에 대한 연구 (예: 베이스 셋 내 어떤 클래스의 데이터가 특정 퓨샷 문제 해결에 주요한지)는 부족했다. 그래서 타겟이 정해진 (target-guided; 서포트 셋이 주어짐; 실제 구분하려는 클래스 레이블에 해당하는 일부 샘플이 알려짐) 상황에서의 퓨샷 분류 문제를 효율적으로 해결하기 위해 베이스 셋을 어떻게 구성해야 하는지, 더 구체적으로 베이스 셋에 속한 데이터 간의 가치를 어떻게 평가 및 비교해야 하는지에 대한 문제는 여전히 해결이 매우 어려운 실정이다. 예를 들어, 퓨샷 얼굴 인식 알고리즘<sup>[2]</sup>을 개발하고자 하는 상황에서 첫 번째 단계는 쓰임이 가능하도록 상황에 맞게 다양한 얼굴 이미지들로 이루어진 베이스 셋을 구성하는 것이다. 이 때, 가용할 수 있는 모든 얼굴 이미지를 베이스 셋으로 구성하는 것은 학습에 필요한 메모리와 계산력을 고려할 때 매우 비효율적일 것이다. 따라서 현실적인 퓨샷 분류 기술 응용에 있어서 학습의 효율을 높이기 위해 베이스 셋에 포함될 이미지를 추려내는 선별 작업과 관련한 문제는 중요하지만, 이 부분에 주목한 연구는 매우 부족하다<sup>[3]</sup>.

본 레터에서 제시한 연구는 최근에 수행된 퓨샷 분

※ 이 성과는 2021년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(No. 2021R1A2C3004345)을 받아 수행된 것임.

\* First Author : (ORCID:0000-0003-2430-3551)School of Mathematics of Computing (Computational Science and Engineering), Yonsei University, jeongkj@yonsei.ac.kr, 학생(석사), 학생회원

<sup>o</sup> Corresponding Author : (ORCID:0000-0002-6533-3469)School of Mathematics of Computing (Computational Science and Engineering), Yonsei University, and Pohang University of Science and Technology (POSTECH) (Artificial Intelligence), wy.shin@yonsei.ac.kr, 교수, 종신회원

논문번호 : 202207-148-C-LU, Received July 21, 2022; Revised August 2, 2022; Accepted August 2, 2022

류와 관련한 두 가지 주요 연구의 실험 결과에 기인한다. 첫 번째 연구<sup>[4]</sup>에서는 다양한 퓨샷 분류 과제에 대한 방대한 실험을 통해 베이스 셋의 크기를 기존 대비 두 배 가까이 늘렸음에도, 퓨샷 분류 성능이 충분히 향상되지 않았다는 결과가 관찰되었다. 이러한 결과는 주어진 베이스 셋 내 데이터를 “선별적”으로 활용해 퓨샷 분류 학습의 효율을 크게 증대시킬 수 있다는 가능성을 암시한다. 두 번째 연구<sup>[5]</sup>에서는 베이스 셋에 속한 각각의 베이스 클래스들이 타겟으로 설정된 퓨샷 문제에 따라 주어진 문제 해결에 있어 그 유용한 정도가 다르다는 것을 실험적으로 보여주었다. 이러한 결과는 현실에서 타겟이 정해진 퓨샷 분류 문제의 효율적 해결을 위해 더 유용한 클래스의 데이터를 선별해 베이스 셋을 구성할 수 있다는 점을 내포한다.

본 레터에서는 효율적인 타겟이 정해진 퓨샷 분류 과제 수행을 위한 베이스 클래스 선별 방법을 제안한다. 제안된 방법에서는 서포트 셋과 베이스 셋에 속한 클래스 간의 유사도를 새롭게 정의 및 계산함으로써 정량적인 비교를 통해 베이스 클래스 선별을 수행한다. 실제 데이터셋을 사용한 실험을 통해 본 연구에서 제안된 클래스 선별 방법을 통해 베이스라인(baseline) 방법 대비 성능 저하는 줄어들면서 효율적으로 퓨샷 분류 과제를 수행할 수 있음을 확인한다.

## II. 방법론

퓨샷 분류 과제에서는 베이스 셋  $D_B$ , 서포트 셋  $D_S$ , 그리고 쿼리 셋  $D_Q$  으로 구성된 세 가지 유형의 데이터 셋이 존재한다. 베이스 셋  $D_B = \{(x_1, y_1), \dots, (x_B, y_B)\}$  은 베이스 클래스 레이블(label)  $y_i \in C_B$  을 지닌 데이터 셋을 나타낸다. 서포트 셋  $D_S = \{(\hat{x}_1, y_1), \dots, (\hat{x}_S, y_S)\}$  과 쿼리 셋  $D_Q = \{(\bar{x}_1, y_1), \dots, (\bar{x}_Q, y_Q)\}$  은 베이스 클래스와 겹치지 않는 실제 분류하고자 하는 새로운 클래스 레이블  $y_i \in C_N$  ( $C_B \cap C_N = \emptyset$ ) 을 지닌 데이터 셋을 의미한다.  $D_c$  는 클래스 레이블  $c$  를 지닌 해당 셋에 속한 모든 샘플들의 집합을 나타낸다.

앞서 언급했듯이 본 연구에서는 타겟이 설정된 상황에서 베이스 셋 선별을 통한 효율적인 퓨샷 분류를 수행하고자 한다. 이를 위해 서포트 셋  $D_S$  과 주어진 샘플을 정해진 크기의 벡터로 매핑하는 초기화된 일반적인 특징 추출 모듈(feature extractor)  $f_\phi(\cdot)$  이 주어진 경우, 전체 베이스 셋  $D_B$  에서  $M$  개의 샘플만

을 선별하여 학습을 진행함으로써 퓨샷 분류의 효과(effectiveness)를 최대화하는 데에 관심이 있다.

위에서 정의한 문제 해결을 위해, 본 연구에서는 “클래스” 기반의 선별 방법을 고려한다. 수치로 정량화된 클래스 별 비교를 위해 서포트 셋을 구성하는 새로운 클래스와 베이스 셋을 구성하는 베이스 클래스 간의 유사도를 새롭게 정의하고 측정함으로써, 분류하려는 클래스와 유사한 베이스 클래스를 선별하는 전략을 제안한다. 클래스 간 유사도 계산을 위한 구체적 내용은 다음과 같다. 먼저, 특징 추출 모듈  $f_\phi(\cdot)$  을 활용하여, 베이스 셋과 서포트 셋에 속한 모든 샘플들의 특징 벡터  $f_\phi(x_i)$  와  $f_\phi(\hat{x}_i)$  를 도출한다. 다음으로, 특정 클래스  $c$  에 속한 클래스 별 중심 벡터를 각 클래스에 속한 특징 벡터들의 평균으로 정의하고 베이스 셋과 서포트 셋에 속한 모든 클래스 별로 각각의 중심 벡터  $\mu_c = \frac{1}{|D_c|} \sum_{(x_i, y_i) \in D_c} f_\phi(x_i)$  를 계산한다. 이후, 서포트 셋에 속한 클래스  $nc$  가 주어졌을 때 베이스 클래스  $bc$  와의 유사도  $sim(bc, nc)$  는 두 중심 벡터 사이의 유클리디안(Euclidean) 거리의 역수로 아래와 같이 주어진다.

$$sim(bc, nc) = \frac{1}{d(\mu_{bc}, \mu_{nc})} \quad (1)$$

(1)에서의 유사도 지표를 활용한 베이스 셋에서  $M$  개의 샘플을 선별하기 위한 구체적인 과정은 다음과 같다, 새로운 클래스  $nc \in C_N$  마다 선별된 베이스 클래스 내 샘플 개수가  $\left\lceil \frac{M}{|C_N|} \right\rceil$  만큼 될 수 있도록, 베이스 클래스  $bc \in C_B$  별로 계산한 유사도가 큰 순서대로 베이스 클래스에 해당하는 샘플들을 모두 선별한다. 이후, 선별된 모든 샘플들 중 최종적으로  $M$  개만을 랜덤하게 선별하는 방식을 취한다.

## III. 실험

실세계 데이터셋으로는 Caltech-USCD Birds 200-2011 (CUB) 데이터셋을 사용한다. CUB 데이터셋은 200개의 새 종류의 클래스를 지닌 11,788개의 이미지로 구성된다. 실험을 위해, 100개의 클래스로 구성된 5,891개의 이미지는 베이스 셋으로, 다른 5,897개의 이미지는 무선회된 샘플링 과정을 통해서

포트 셋과 쿼리 셋을 만들어내는 데이터셋 (노벨 셋 (novel set)) 으로 활용하였다.

퓨샷 분류 과제 설정을 위해 임의로 5개의 클래스를 노벨 셋에서 샘플링 하여 이 중 각 클래스 별로 1개의 샘플을 샘플링하여 서포트 셋으로 그리고 나머지 샘플들을 쿼리 셋으로 구성하여 5-way 1-shot 퓨샷 분류 실험을 진행하였다. 우연성을 최소화하기 위해, 무선 샘플링에 기반한 세팅 설정 과정을 60번 반복 하여 나온 성능을 평균하여 최종 성능으로 기록하였고, 성능 평가의 지표로는 분류 정확도 (accuracy)를 활용하였다. 퓨샷 분류 방법으로는 퓨샷 분류 분야에서 일반적으로 사용되는 K-NN 기반 분류<sup>[6]</sup>를 적용하였다. 베이스 셋에서 선별하려는 데이터의 크기 (즉, 샘플 개수)는 전체 베이스 셋의 크기에 대한 상대적인 비율로써 0%, 25%, 50%, 75%, 100%로 설정하여 실험을 진행하였다. 이 때, 0%는 베이스 셋을 활용하지 않는 경우, 100%는 모든 베이스 셋을 활용하는 경우를 의미한다. 실험에 사용한 특징 추출 모듈  $f_{\phi}(\cdot)$ 은 ImageNet에서 사전 학습된 Resnet18을 사용하였다. 제안 방법 (Proposed)의 성능 평가를 위해 비교 베이스라인으로 무선 샘플링을 통한 선별 방법 (Random) 및 제안된 방식과는 반대로 유사도가 작은 순서대로 선별한 방법 (Dissim)을 고려하였다.

표 1은 베이스 클래스 선별 방식 및 선택된 샘플들의 비율에 따른 퓨샷 분류 성능 비교 결과를 보여준다. 베이스 셋에서 데이터를 선별함에 있어 본 연구에서 제안한 유사도 기반 방법을 사용할 경우 두 가지 베이스라인인 Random과 Dissim 대비 베이스 셋 선별 비율에 상관없이 항상 우수한 분류 정확도를 보임을 확인할 수 있다. 또한, 제안한 방법 사용 시 단지 25%의 베이스 셋 내 데이터만을 사용하더라도 분류 정확도가 여전히 70% 이상을 유지할 수 있고, 전체 베이스 셋을 활용하는 경우 (즉, 100%) 대비 정확도 손실이 6.26%에 불과함을 알 수 있다. 이는 제안한 방법이 학습 데이터 축소에 따른 성능 저하를 최소화하면서

도 효과적으로 베이스 셋 내 데이터를 선별함을 내포한다. 다시 말해, 타겟이 정해진 퓨샷 분류 과제의 해결을 위해 제안된 방법을 통해 베이스 셋을 효율적 그리고 효과적으로 구성할 수 있음을 시사한다.

#### IV. 결론

본 레터에서는 타겟이 설정된 퓨샷 분류 과제의 효율적인 수행을 위해 간단하면서도 효율적인 유사도 기반 베이스 클래스 선별 방법을 제안하였다. 실세계 데이터셋을 사용한 실험을 통해, 제안한 방법을 사용해 선별한 전체 베이스 셋의 단지 1/4만을 활용하면서도 70% 이상의 분류 정확도를 취득할 수 있음을 보였다.

#### References

- [1] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few shot learning," *ACM Computing Surv.*, vol. 53, no. 3, pp. 1-34, May 2020. (<https://doi.org/10.1145/3386252>)
- [2] Y. Wu, H. Liu, and Y. Fu, "Low-shot face recognition with hybrid classifiers," in *Proc. IEEE/CVF ICCV*, pp. 1933-1939, Venice, Italy, Oct. 2017. (<https://www.dio.org/10.1109/ICCVW.2017.228>)
- [3] K.-J. Jeong and W.-Y. Shin, "Target-guided base subset selection for efficient few-shot classification," in *Proc. KICS Summer Conf.*, Jeju Island, Korea, Jun. 2022. (<http://www.kics.or.kr>)
- [4] E. Triantafillou, T. Zhu, V. Dumoulin, P. Lamblin, U. Evci, K. Xu, R. Goroshin, C. Gelada, K. Swersky, P.-A. Manzagol, and H. Larochelle, "Meta-dataset: A dataset of datasets for learning to learn from few examples," in *Proc. ICLR*, pp. 1-24, Virtual Event, Apr. 2020. (<https://openreview.net/pdf?id=rkgAGAVKPr>)
- [5] J. Yan, D. Zhai, J. Jiang, and X. Liu, "Target-guided adaptive base class reweighting for few-shot learning," in *Proc. ACM Int. Multimedia Conf. (MM)*, pp. 5335-5343, Chengdu, China, Oct. 2021. (<https://doi.org/10.1145/3474085.3475656>)

표 1. 베이스 클래스 선별 방법 및 선택된 샘플들의 비율에 따른 분류 정확도

Table. 1. Classification accuracy according to the type of base class selection methods and the ratio of selected samples.

방법	베이스 셋 선별 비율				
	0%	25%	50%	75%	100%
Random	54.73	68.89	73.08	75.96	78.68
Dissim	54.73	62.80	69.41	74.83	78.68
Proposed	54.73	<b>72.42</b>	<b>74.35</b>	<b>77.21</b>	78.68

- [6] O. Sbai, C. Couprie, and M. Aubry, “Impact of base dataset design on fewshot image classification,” in *Proc. ECCV*, pp. 1-17, Virtual Event, Aug. 2020.  
([https://www.ecva.net/papers/eccv\\_2020/papers\\_ECCV/papers/123610579.pdf](https://www.ecva.net/papers/eccv_2020/papers_ECCV/papers/123610579.pdf))