

메타학습 기반 소수샷 반려견 얼굴 식별

연수민*¹, 배지호*², 부석준*³, 최상민*⁴, 이수원*⁵

Few-Shot Dog Face Identification via Meta-Learning

Sumin Yeon*¹, Jiho Bae*², Seok-Jun Buu*³, Sang-Min Choi*⁴, and Suwon Lee*⁵

요약

최근 반려동물 개체 수 증가와 함께 실종 및 유기견 문제가 심화됨에 따라 반려견 얼굴 인식 연구는 반려견 개체 인식 데이터 세트에 대한 검증 성능을 높이는 데 집중해왔으나, 기존의 연구는 개체당 이미지 수가 절대적으로 부족한 소수샷 환경의 근본적인 문제 해결에는 한계가 있었다. 이러한 문제를 해결하고자, 본 연구는 소수의 얼굴 이미지로 반려견을 신속히 식별하는 메타학습 기반 소수샷 얼굴 인식 시스템을 제안한다. 본 논문에서는 Prototypical Networks와 Meta-DeepBDC 메타학습 기법을 DogFaceNet 데이터 세트에 적용하여 비교 실험을 수행했으며, 그 결과 Meta-DeepBDC가 ResNet-12을 Backbone으로 한 1-shot 조건에서 64.01%의 높은 분류 정확도를 달성했다. 이는 반려견 얼굴 인식 분야에 메타학습을 최초로 적용한 사례로, 데이터 부족 환경에서 개체 식별 성능을 효과적으로 개선할 수 있음을 입증했다는 점에서 의의가 있다.

Abstract

As the population of companion animals increases and the problem of lost or abandoned dogs intensifies, research in dog face recognition has focused on improving performance on individual identification datasets. However, conventional studies have shown limitations in addressing the fundamental challenge of few-shot environments, where the number of images per individual is critically low. To address this issue, this study proposes a meta-learning-based few-shot face recognition system for rapidly identifying individual dogs from a small number of images. We conducted comparative experiments on the DogFaceNet dataset using two meta-learning techniques: Prototypical Networks and Meta-DeepBDC. The results show that Meta-DeepBDC achieved a high classification accuracy of 64.01% in a 1-shot setting with a ResNet-12 backbone. This study marks the first application of meta-learning to the field of dog face recognition and is significant for demonstrating that individual identification performance can be effectively enhanced even in data-scarce environments.

Keywords

dog identification, meta-learning, few-shot learning, individual identification

* 경상국립대학교 컴퓨터공학과(*⁵ 교신저자)
- ORCID¹: <https://orcid.org/0009-0005-5598-9367>
- ORCID²: <https://orcid.org/0009-0004-5942-4972>
- ORCID³: <https://orcid.org/0000-0002-3940-3611>
- ORCID⁴: <https://orcid.org/0000-0001-5950-3081>
- ORCID⁵: <https://orcid.org/0000-0003-2603-1385>

• Received: Jul. 10, 2025, Revised: Aug. 11, 2025, Accepted: Aug. 14, 2025
• Corresponding Author: Suwon Lee
Dept. of Computer Science and Engineering, Gyeongsang National University, 501, Jinju-daero, Jinju-si, Gyeongsangnam-do, Korea
Tel.: 82+55-772-1394, Email: leesuwon@gnu.ac.kr

1. 서 론

최근 몇 년간 반려동물의 수와 길을 잃거나 버려진 유실·유기 동물의 수가 증가하면서, 마이크로칩 이식, 코 인식, 홍채 인식과 같은 다양한 반려동물 식별 방법이 제안되었다[1]. 국내의 경우 동물보호법에 따라 반려견과 그 소유자의 등록이 의무화되어 있으며, 이를 위해 내장형 마이크로칩 이식과 외장형 무선식별장치 부착 방법을 지원하고 있다. 그러나 내장형 마이크로칩의 이식에 따른 반려견 소유자의 불안감과 외장형 무선식별장치의 분실 및 파손 문제가 지속적으로 제기되고 있다. 실종된 반려견을 수색하는 체계는 아직 적립된 것이 없으며, 그 방법 또한 전적으로 인간이 직접 수색하는 것이다.

최근 기계 학습 및 딥러닝 기법은 인간 얼굴 인식에서 높은 신뢰성과 정확성을 제공하고 있으며 [2], 이를 바탕으로 다양한 제품과 서비스가 출시되고 있다. 이러한 흐름에 따라 반려견 얼굴 인식 분야에서도 딥러닝을 활용한 접근이 시도되고 있다. 이를 통해 CCTV(Closed Circuit Television)에서 포착되거나 길을 떠도는 반려견을 목격한 제보자의 사진을 사용자가 보유한 사진과의 확인하는 등의 과정을 통해 실종된 반려견을 찾는 시스템을 고려해 볼 수 있다.

반려견 얼굴 데이터의 경우 개체 식별을 목적으로 라벨링된 활용 가능한 공개 데이터 세트의 수가 매우 적다. 또한, 반려견의 경우 종에 따라 털의 무늬와 이목구비의 형태 등 외형적 차이가 크며, 특히 희귀종의 경우 데이터 확보가 더욱 어렵다. 따라서 반려견 개체 식별을 위한 데이터는 대체로 수량이 적고, 특정 종이나 개체에 치중될 가능성이 크다. 전통적인 딥러닝 네트워크를 이러한 불균형한 데이터 세트로 학습시키면, 샘플이 많은 클래스에 편향되어 올바른 개체 식별이 어렵다. 게다가 기존 딥러닝 네트워크는 학습 시간이 많이 소요되어 유기견 수색 등 실제 현장에서는 실용적이지 않을 수 있다. 따라서 소량의 샘플로도 신속하게 새로운 반려견 개체 클래스를 학습하고 적용할 수 있는 방법이 요구된다.

이 문제를 해결하기 위해 다양한 소수샷 분류 기법[3]이 제안되었고, 여러 방법들 중 적은 양의 데이터만으로 효과적으로 학습할 수 있는 메타학습

(Meta-learning)[4] 기법이 주요하게 사용된다. 메타학습은 소수의 예시만으로도 새로운 과제와 환경에 빠르게 적응할 수 있는 능력을 모델에 부여하며, 핵심 아이디어는 적은 데이터로 한두 번의 기울기 업데이트만으로도 새로운 과제에서 최대한의 성능을 낼 수 있도록 초기 파라미터를 학습하는 것이다. 반려견 얼굴을 통한 개체 식별 문제에서 메타학습을 비롯한 소수샷 기법을 사용한 사례는 볼 수 없었다.

본 논문에서는 Prototypical Networks[5]와 Meta-DeepBDC[6]와 같은 메타학습 기반의 소수샷 접근법을 활용하여, 데이터가 부족한 환경에서도 반려견 얼굴 인식을 통한 개체 분류 시스템을 제안한다. 제안한 시스템은 적은 수의 반려견 얼굴 데이터만으로도 개체 식별에 신속하게 적응하도록 설계되었다.

해당 시스템은 반려견 얼굴 인식을 위한 공개 데이터 세트인 DogFaceNet[7]을 사용하여 Prototypical Networks 및 Meta-DeepBDC 기반의 제안 시스템의 얼굴 식별 성능을 ResNet-12과 ResNet-18를 Backbone으로 사용하여 평가하였으며, 전이 학습(Transfer learning) 기반의 베이스라인과 성능을 비교하였다. 실험 결과 Backbone으로 ResNet-12를 사용한 Meta-DeepBDC가 적은 데이터를 사용한 소수샷 분류에서 가장 우수한 성능을 보였다.

본 논문의 기여점은 다음과 같다. 적은 수의 샘플을 가진 반려견 얼굴 데이터 세트에서 소수샷 분류 과제로 반려견 얼굴 인식을 통한 개체 식별 문제를 공식화하고, 메타학습 기반 식별 시스템을 제안했다. 또한 Prototypical Networks 및 Meta-DeepBDC를 이용하여 반려견 얼굴 인식 기법을 탐구했다. DogFaceNet 데이터 세트에서 제안한 시스템을 전이 학습 베이스라인과 비교 평가하여, ResNet-12를 사용한 Meta-DeepBDC의 우수성을 입증하였다.

본 논문의 2장에서는 동물 생체 인식을 위한 딥러닝 관련 연구와 데이터 부족 문제를 해결하기 위한 연구 동향을 검토한다. 3장에서는 반려견 얼굴 인식을 위한 메타학습 시스템의 정의와 방법론을 설명한다. 4장에서는 제안한 시스템의 실험 구현 과정과 결과를 평가하며, 5장에서는 결론과 향후 연구 방향에 대해 논의한다.

II. 관련 연구

2.1 동물 생체인식을 위한 딥러닝

동물 생체인식 연구는 주로 소[8], 말[9], 그리고 멸종위기에 처한 종[10]을 대상으로 활발히 이루어져 왔다. 소는 주둥이의 독특한 무늬(Muzzle point)를, 말·돼지·멸종위기종은 얼굴 특징을 활용해 개체를 구별하는 방식이 주로 채택되었다[11]. 반면 개의 경우에는 얼굴 인식보다는 품종 분류에 초점이 맞춰진 연구가 많았다[12][13]. 동물 생체인식 기술은 크게 주둥이 무늬, 홍채 패턴, 망막 혈관, 그리고 얼굴 영상 네 가지 방법으로 나뉜다. 주둥이 무늬 방식은 신뢰도가 높은 편이지만 전용 스캐너가 필요하고, 홍채·망막 기반 방법 역시 특수 장비 의존성이 크다. 이에 비해 개 얼굴 인식은 일반 카메라만 있으면 데이터를 수집할 수 있고, 인터넷에서 고해상도 이미지를 비교적 손쉽게 확보할 수 있다는 장점이 있다. 그러나 사람 얼굴용 모델을 그대로 적용하면 낮은 구조적 정상성(Normality) 때문에 과적합이 쉽게 발생한다는 문제가 지적되어 왔다[7].

이러한 한계를 극복하기 위해 반려견 얼굴 식별을 위한 다양한 시도가 이어지고 있다[7][14][15]. 예를 들어 Moreira et al.은 반려견 얼굴을 통한 개체 식별을 위한 최초의 오픈 데이터 세트인 Flickr-dog 데이터 세트를 제시하고, shallow-deep CNN(Convolutional Neural Network) 구조를 설계해 Flickr-dog 데이터 세트에서 67.6%의 정확도를 보았다[14]. G. Mougeot et al.[7]은 웹에서 수집한 반려견 얼굴 데이터를 통해 공개 데이터 세트 DogFaceNet을 구축하고, 트리플렛 손실(Triplet loss)과 결합된 deep-CNN을 훈련했다. 48마리의 다른 반려견에 대한 오픈 데이터 세트에서 두 이미지가 같은 클래스인지에 대한 검증 작업에서 92%의 정확도를 달성했다. Mougeot et al.이 제시한 데이터 세트를 활용한 연구로는 B. Yoon et al.[15]의 L2-norm 레이어를 제거하고 새로운 손실 함수인 벡터 길이 손실(vector length loss)과 2단계 학습 방법을 사용하여 반려견 얼굴 식별 모델의 성능을 향상시키는 방법론이 있다. 이는 기존 학습 방법 대비 검증 작업

에서 4%p 높은 97.33%의 정확도를 달성했다. 국내 연구로는 DogFaceNet을 사용하여 ResNet-50 기반 Backbone 네트워크를 사용한 D. S. Lee et al.[16]의 연구가 있으며, 최근 실종 반려견 찾기 시스템을 설계하고 구현한 G. T. Byeon et al.[17]의 연구가 있다.

2.2 데이터 부족 문제에서의 딥러닝

데이터 부족 문제를 해결하기 위해 전이 학습[18]과 소수샷 학습[19]과 같은 여러 접근법이 제안되었다. 전이 학습은 일반적으로 관련 도메인의 대규모 라벨 데이터를 활용해 신경망을 사전 학습한 뒤, 목표 도메인의 소규모 데이터로 미세 조정(Fine-tuning)을 수행하여 성능을 향상시킨다[20]. 소수샷 기법은 극소수의 데이터만으로도 효과적인 학습이 가능하도록 초기 모델 매개변수나 임베딩 공간을 메타학습(Meta-learning) 방식 등으로 최적화하여 모델의 빠른 적응성을 확보한다. 최근에는 메타 학습 기반 접근법들이 자연 이미지 분류와 같은 일반적인 이미지 데이터 세트에 성공적으로 적용되어 뛰어난 성능을 보여 왔지만, 반려견 얼굴 인식을 통한 개체 식별 문제에서는 메타학습 기법에 대한 연구가 충분히 이루어지지 않았다.

III. 반려견 얼굴 식별을 위한 메타학습 시스템

본 연구에서는 적은 데이터 환경에서 반려견 얼굴 이미지를 이용한 개체 식별 문제를 소수샷 학습 문제로 정의하고, 최근 주목받고 있는 메타학습 기법을 적용한다. 제안하는 시스템은 소수의 샘플로 네트워크를 효율적으로 학습시키고, 전통적인 딥러닝만큼의 성능을 향상시키는 것을 목표로 한다.

그림 1은 본 논문에서 제안하는 메타학습 시스템의 전체 구조를 보여준다. 시스템은 메타 훈련(Meta-training) 단계와 메타 시험(Meta-testing) 단계로 구성된다. 본 연구에서는 거리 기반(Distance-based)의 Prototypical Networks와 Meta-DeepBDC를 메타학습 기법으로 사용하며, Backbone(Backbone) 모델로는 ResNet-12 또는

4 메타학습 기반 소수샷 반려견 얼굴 식별

ResNet-18을 사용한다. 메타 훈련 단계에서는 증강된 훈련 데이터 세트를 기반으로 이를 서포트 집합(Support set)과 쿼리 집합(Query set)으로 나누어 사용한다. Prototypical Networks의 경우 가중치가 공유되는 Backbone 네트워크를 통해 얻은 특징 벡터(Feature vector)로 프로토타입을 만들고, Meta-DeepBDC의 경우 BCD module을 추가로 거친 후 BCD matrix를 통한다는 구조적 차이가 있다. 메타 시험 단계에서는 메타 훈련을 통해 학습된 모델을 메타 훈련에서 사용하지 않은 새로운 클래스 데이터에 적용하여 분류 성능을 평가한다. 이때 시험 데이터 세트의 쿼리 집합에 대해 성능이 계산된다.

이어서 제안된 인식 파이프라인에 사용된 두 가지 메타학습 기반 분류 방법인 Prototypical Networks와 Meta-DeepBDC에 대해 설명한다. 두 방법 모두 에피소드(Episode) 기반의 학습 전략을 공유하지만, 특징 표현 방식과 클래스 간 유사도 계산 방식에서 차이를 보인다.

3.1 Prototypical Networks

Prototypical Networks[5]는 클래스마다 하나의 프로토타입 벡터를 추정하고, 쿼리 샘플을 해당 프로토타입과의 거리에 따라 분류하는 방법을 사용한다. 이하의 기호 체계는 양 방법 모두에 일관되게 적용된다. 무작위로 추출된 여러 개의 클래스를 하나의 학습 집합이라 할 때, 서포트 집합 $S = \{(x_i, y_i)\}_{i=1}^n$

과 쿼리 집합 $Q = \{(x_i, y_i)\}_{j=n+1}^{n+m}$ 을 생성한다. 여기서 n 은 서포트 샘플 수, m 은 쿼리 샘플 수이다. 서포트 집합에서 임베딩 함수 f_θ 를 통해 특징을 추출하면, 클래스 k 의 프로토타입 c_k 는 식 (1)과 같다.

$$c_k = \frac{1}{|S_k|} \sum_{(x_i, y_i) \in S_k} f_\theta(x_i) \quad (1)$$

쿼리 x 에 대해 거리 $d(f_\theta(x), c_k)$ 를 계산하고, 식 (2)를 통해 분류 확률을 산출한다.

$$p_\theta(y = k|x) = \frac{\exp(-d(f_\theta(x), c_k))}{\sum_k (\exp(-d(f_\theta(x), c_k)))} \quad (2)$$

에피소드마다 쿼리 집합에 대한 로그 우도(Log-likelihood)를 최대화하는 방향으로 θ 를 업데이트 한다. 식 (3)은 손실 함수를 나타낸다.

$$L_{PN} = - \sum_{(x,y) \in Q} \log p_\theta(y|x) \quad (3)$$

해당 방식을 통해 Prototypical Network는 각 클래스의 중심부터 거리를 기준으로 새로운 샘플을 분류하는 모델을 학습한다. 이는 적은 샘플에서도 빠른 수렴과 해석 가능성을 제공한다. 다만 입력 특성이 고차 통계 정보를 요구할 때는 1차 평균만으로는 표현력이 부족할 수 있다.

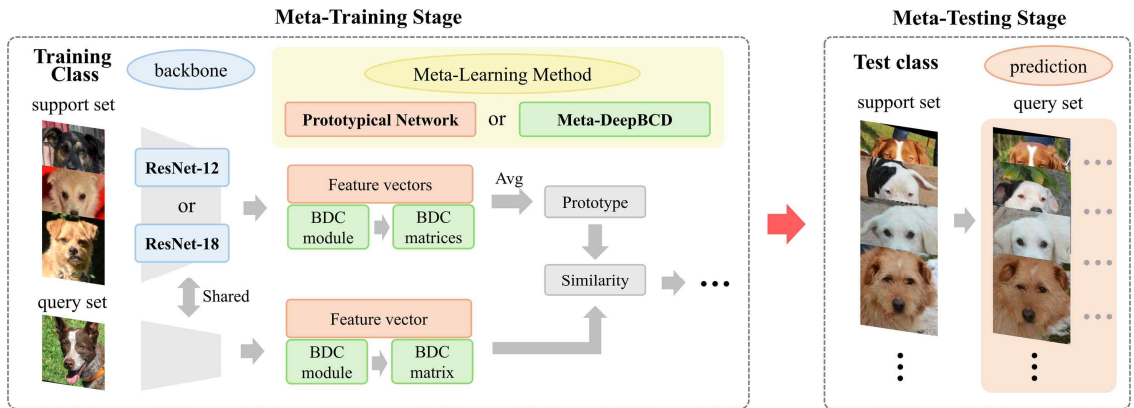


그림 1. 반려견 얼굴 식별을 위한 메타학습 시스템 구조
Fig. 1. Architecture of meta-learning system system for dog face identification

3.2 Meta-DeepBDC

Meta-DeepBDC[6]는 Prototypical Network의 프로토타입과 거리의 사용 프레임의 양방향 벡터화 공분산(BDC, Bilinear & Deep Covariance) 표현으로 확장하여 2차 통계까지 학습한다. 먼저 Backbone 네트워크가 추출한 특징 맵 $F \in \mathbb{R}^{h \times w \times d}$ 에 대해 학습 가능한 투영 행렬 $W \in \mathbb{R}^{d \times r}$ 을 적용해 축소 특징 $Z \in \mathbb{R}^{hw \times r}$ 을 얻는다. 이후 대칭 공분산 행렬의 상삼각부를 벡터화하여, 고정 길이 표현을 얻는다.

$$Z = W^T F, \quad \Sigma = (Z^T Z) \in \mathbb{R}^{r(r+1)/2} \quad (4)$$

BDC 벡터화는 각 채널 간 2차 상관을 보존하여 배정·조명 등 복잡한 변동에 강인하며, $r \ll d$ 차원 축소 덕분에 추가 계산량은 선형적이다. 클래스별 2차 프로토타입은 식 (5)와 같이 정의된다.

$$\tilde{c}_k = \frac{1}{|S_k|} \sum_{(x_i, y_i) \in S_k} \Sigma_i \quad (5)$$

손실 함수는 Prototypical Networks과 유사한 식으로 정의되며, Backbone θ 과 투영 행렬 W 을 동시에 계산한다.

IV. 실험 및 평가

제한한 시스템의 반려견 얼굴 식별 성능 평가를 위해 Prototypical Networks 및 Meta-DeepBDC 기반 모델을 구축하고, Backbone에 따른 성능을 1, 3, 5-shot에 대해 비교하는 실험을 진행한다. 이는 전이 학습 기반의 베이스 라인과 각각 비교된다.

4.1 데이터 세트 구성

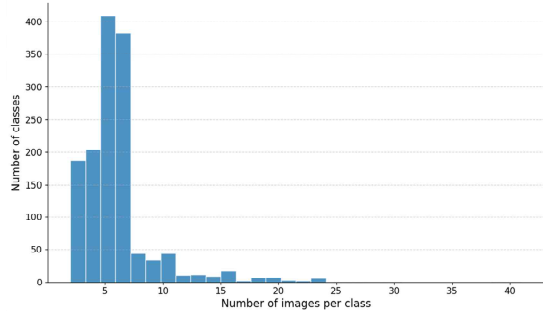
본 연구의 실험에는 DogFaceNet 데이터 세트[7]를 사용했다. 반려견 개체 식별을 목적으로 라벨링된 공개 데이터 세트는 한정되어 있고, 그 중 DogFaceNet은 관련 딥러닝 연구에서 표준 벤치마크로 사용되어 연구의 재현성 및 비교 평가에 용이하

기 때문이다. 특히 다수의 클래스와 클래스당 소수의 이미지로 구성된 해당 데이터 세트의 구조는, 본 연구가 해결하고자 하는 소수샷 학습 문제를 정의하고 평가하기에 적합한 환경을 제공하기에 선택되었다.

그림 2(a)는 DogFaceNet 데이터 세트의 샘플 이미지를 보여준다. 대부분의 이미지에서 반려견의 얼굴이 중앙에 위치하도록 정렬되어 있으며, 각 이미지는 224×224 크기로 이루어져 있다. 데이터 세트는 총 1393마리의 반려견을 클래스로 라벨링한 8363장의 얼굴 이미지로 구성된다. 그림 2(b)는 각 클래스에 해당하는 이미지 수를 시각화한 것으로, 각 클래스는 최소 2장에서 최대 41장의 이미지를 가지고 있으며, 대부분의 클래스는 4-6장의 이미지를 보유하고 있다.



(a)



(b)

그림 2. DogFaceNet 데이터 세트 구성

(a) 데이터 세트 샘플 이미지,

(b) 각 클래스에 해당하는 이미지의 수

Fig. 2. DogFaceNet datasets composition

(a) Sample images of the dataset

(b) Number of images for each class

메타 시험 단계에서는 데이터 증강 없이 원본 데이터를 사용하기 위해 10장 이상의 이미지를 가진 129개 클래스를 시험 세트로 분류했다. 10장 미만의 이미지를 가진 클래스 1264개 중 126개의 클래스를 난수 시드는 0으로 고정한 후 무작위로 선택하여 검증 세트로 설정했다. 남은 1138개의 클래스는 훈련 세트로 사용하며, 이 클래스들의 이미지 수는 각각 다양하게 분포되어 있다.

훈련 데이터셋의 클래스별 이미지 수가 매우 제한적이므로 데이터 증강 기법을 적용해 데이터를 보충한다. 각 클래스당 이미지 수가 10장이 되도록 좌우 반전(확률 50%), 무작위 회전($\pm 15^\circ$, 확률 50%), 색상 대비 밝기 선명도 중 하나를 선택하여 0.8-1.2 배로 조정, 가우시안 블러(radius=1, 확률 30%) 등의 무작위 변환을 적용하였다.

4.2 실험 세부 구현 사항

Backbone 네트워크로 ResNet-12와 ResNet-18을 사용하여 실험을 진행한다. 이는 본 연구에서 평가하는 핵심 방법론인 Meta-DeepBDC의 원본 연구[6]에서 사용된 아키텍처를 동일하게 채택하여 방법론의 성능을 공정하게 비교하고 재현성을 확보하고, 서로 다른 깊이의 네트워크를 사용함으로써 모델의 복잡도가 해당 과제에 미치는 영향을 분석하고자 했다.

Xie et al.의 구현과 동일하게 ResNet-12는 네 개의 스테이지로 구성되며, 각 스테이지마다 하나의 잔차 블록(residual block)이 있다. 각 스테이지의 블록 너비는 [64, 160, 320, 640]으로 설정되었다. 각 잔차 블록은 배치 정규화와 0.1 Leaky ReLU를 적용한 3×3 컨볼루션 레이어 세 개로 구성되며, 마지막 스테이지를 제외한 각 블록 뒤에는 2×2 맥스 풀링 레이어를 사용하여 특징 맵을 다운 샘플링한다. ResNet-12를 사용할 때는 원본 224×224 크기의 이미지를 84×84로 스케일링하여 사용한다. ResNet-18은 마지막 다운샘플링 레이어를 제거하고 사용하였다[6]. 사용한 메타학습 기법의 세부 구현 사항은 다음과 같다.

Prototypical Network는 유클리드 거리(Euclidean distance)를 사용하여 학습하였으며, 훈련 에피소드는 20-way로 구성하였다. 각 훈련 에피소드에서 클래스당 5장의 이미지를 무작위로 샘플링하여 사용

했다. 최적화 방법으로는 Adam 옵티마이저를 적용한 확률적 경사하강법 방식을 사용하였으며, 고정 학습률 10^{-4} 와 가중치 감쇠(Weight decay) 10^{-5} 을 적용하였다. 메타 시험 단계에서는 반려건 얼굴 데이터셋을 활용해 20-way 분류 과제를 수행하였다. 클래스당 서포트 집합 이미지 수를 1, 3, 5장($k=1, 3, 5$)으로 설정하고, 쿼리 집합은 5장으로 고정해 각각의 분류 성능을 평가했다. 시험 에피소드는 2000회씩 총 시험을 5회 반복하여 결과를 얻었다.

Meta-DeepBDC 역시 Prototypical Network와 동일하게 유클리드 거리를 사용하여 학습하였으며, 훈련 에피소드 구성도 20-way로 동일하게 설정했다. 학습 최적화에는 모멘텀 0.9와 가중치 감쇠(Weight decay) $5e-4$ 를 적용한 확률적 경사 하강법 알고리즘을 사용하였으며, 그 외 설정은 Prototypical Network와 동일하게 진행했다. 또한 Meta-DeepBDC의 온도 스케일 파라미터는 별도의 학습률을 사용하여 갱신했다.

Baseline 모델로는 전체 훈련 데이터 세트를 사용해 사전 학습한 모델을 사용한다. ResNet-12와 ResNet-18에 대해 전체 훈련 데이터셋을 활용하여 먼저 모델을 학습시킨 후, 메타 훈련 과정을 거친 뒤, 시험 클래스에 대해 미세 조정 및 평가를 진행했다. 미세 조정 단계에서는 클래스당 서포트 세트 이미지 수를 1, 3, 5장으로 설정하여 학습을 진행하고, 쿼리 세트를 통해 성능을 평가했다.

4.3 실험 분석

결과 분석 시 시험 과제별 평균 정확도와 95% 신뢰구간을 함께 제시한다. DogFaceNet 데이터 세트에서 다양한 기법들의 성능을 비교한 결과는 표 1에 제시되었다. 모든 경우 Meta-DeepBDC가 Prototypical Networks보다 더 우수한 성능을 나타냈다. Meta-DeepBDC는 채널 간 상관관계를 벡터화하여 2차 통계 정보를 활용하고, 배경 등의 변화에 강한 특성을 가지기 때문에 더 나은 성능을 보인 것으로 해석할 수 있다. 따라서 조명과 배경의 편차가 큰 DogFaceNet 데이터 세트에서는 Meta-DeepBDC와 같이 변화에 강한 방법이 더욱 효과적일 것으로 판단된다.

표 1. DogFaceNet 데이터 세트에서 20-way 분류 작업에 대한 정확도(%) 성능 비교
 Table 1. Performance comparison of Accuracy (in %) on the DogFaceNet dataset for 20-way classification tasks

		Baseline	Prototypical Networks	Meta-DeepBDC
ResNet-12	1-shot	18.31 ± 0.23	51.22 ± 0.30	64.01 ± 0.29
	3-shot	23.26 ± 0.25	67.14 ± 0.26	77.51 ± 0.24
	5-shot	26.83 ± 0.24	68.96 ± 0.25	82.36 ± 0.20
ResNet-18	1-shot	20.14 ± 0.29	45.12 ± 0.29	56.58 ± 0.30
	3-shot	27.06 ± 0.25	61.39 ± 0.27	67.43 ± 0.27
	5-shot	33.52 ± 0.26	65.46 ± 0.26	72.71 ± 0.25

실험 결과, k-shot 학습에서 k 값이 증가할수록 전반적인 성능 또한 향상되는 경향을 보였다. 이는 Backbone 네트워크와 학습 방법 전반에서 공통으로 나타난 현상이며, 세부 조정과 프로토타입 벡터 생성 시 더 많은 데이터를 사용할수록 모델 파라미터가 개선되어 시험 클래스에서 높은 성능을 나타내기 때문에 판단된다. Meta-DeepBDC의 경우 ResNet-12를 사용했을 때 1-shot 실험에서 64.01%, 3-shot 실험에서 77.51%, 5-shot 실험에서 82.36%의 높은 정확도를 기록했다. t-검정을 통해 p-value < 0.01임을 확인하여 모델의 우수성을 확인했다.

Prototypical Networks와 Meta-DeepBDC 두 방법 모두 Baseline 보다 더 나은 성능을 보였으며, 이는 적은 데이터 환경에서 메타학습 기법의 우수성을 보여주는 결과이다.

Backbone 네트워크 성능 비교에서는 Prototypical Networks와 Meta-DeepBDC의 1-shot, 3-shot, 5-shot 설정에서 모두 ResNet-12가 더 높은 성능을 보였다. ResNet-18을 Backbone으로 학습하는 과정에서, 훈련 데이터 세트의 손실 값은 꾸준한 하락세를 보였으나, 검증 데이터 세트에서 일정 이하로 하락하지 않아 훈련 데이터 세트와의 손실 값의 격차가 나타나는 현상을 확인하였다. 이는 ResNet-18의 더 깊은 구조가 메타 훈련과 세부 조정 과정에서 과적합을 초래하여 시험 단계에서 성능이 하락한 것으로 추측된다. 반면 Pre-trained 모델의 경우 ResNet-18을 Backbone으로 사용했을 때 성능이 더 우수하게 나타났다. 이는 Pre-trained 모델이 기존 사전 학습된 성능에 의존하여 새로운 작업에 유의미한 특징을 충분히 추출하지 못했기 때문에 해석할 수 있다.

V. 결론 및 향후 연구

본 연구에서는 학습 데이터가 부족한 반려견 얼굴 인식 환경에서 개체 식별 성능을 향상시키기 위해 메타학습 기반의 소수샷 학습 시스템을 설계하고, Prototypical Networks와 Meta-DeepBDC를 적용하였다. 또한 Backbone을 달리하여 최적의 모델을 위해 비교 실험을 구축하였다. DogFaceNet 데이터 세트를 대상으로 한 20 way, 1-3-5 shot 실험 결과, 제안한 방법이 데이터가 부족한 환경에서 기존의 전통적 딥러닝 접근법보다 우수함을 보였다.

본 논문에서는 이전에 시도된 바 없던 반려견 얼굴 인식 환경에서 소수샷 학습 방법으로 문제를 정의하고, 메타학습 기법을 도입해 우수한 모델을 비교, 제시하였다는 점에서 의의를 찾을 수 있다. 그러나 본 연구는 단일 데이터 세트에 국한되어 있어, 다른 반려견 얼굴 이미지 도메인 및 실제 현장으로의 일반화 가능성을 추가로 검증할 필요가 있다. 또한, 다양한 Backbone 모델 기반의 실험과 실제 현장에서의 확장 가능성을 위해 모델 경량화와 실시간 추론 능력을 확인할 필요가 있다.

References

- [1] S. Kumar and S. K. Singh, "Monitoring of pet animal in smart cities using animal biometrics", *Future Generation Computer Systems*, Vol. 83, pp. 553-563, Jun. 2018. <https://doi.org/10.1016/j.future.2016.12.006>.
- [2] D. S. Trigueros, L. Meng, and M. Hartnett, "Face

- recognition: From traditional to deep learning methods", arXiv preprint, arXiv:1811.00116, Nov. 2018. <https://doi.org/10.48550/arXiv.1811.00116>.
- [3] T. B. Sekou, M. Hidane, J. Olivier, and H. Cardot, "From patch to image segmentation using fully convolutional networks—Application to retinal images", arXiv preprint, arXiv:1904.03892, Apr. 2019. <https://doi.org/10.48550/arXiv.1904.03892>.
- [4] A. Nichol, J. Achiam, and J. Schulman, "On first-order meta-learning algorithms", arXiv preprint, arXiv:1803.02999, Mar. 2018. <https://doi.org/10.48550/arXiv.1803.02999>.
- [5] J. Snell, K. Swersky, and R. S. Zemel, "Prototypical networks for few-shot learning", Proc. Adv. Neural Inf. Process. Syst., Vol. 30, pp. 4077-4087, Dec. 2017.
- [6] J. Xie, Y. Zhang, X. Xu, Y. Shen, M. Tan, and C. Gong, "Joint distribution matters: Deep Brownian distance covariance for few-shot classification", Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), New Orleans, LA, USA, pp. 16693-16702, Jun. 2022. <https://doi.org/10.1109/CVPR52688.2022.01628>.
- [7] G. Mougeot, D. Li, and S. Jia, "A deep learning approach for dog face verification and recognition", Proc. 16th Pacific Rim Int. Conf. Artif. Intell. (PRICAI), Cuvu, Yanuca Island, Fiji, pp. 493-505, Aug. 2019. https://doi.org/10.1007/978-3-030-30508-6_40.
- [8] S. Kumar and S. K. Singh, "Cattle recognition: A new frontier in visual animal biometrics research", Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci., Vol. 90, No. 4, pp. 689-708, Dec. 2020. <https://doi.org/10.1007/s40010-019-00610-x>.
- [9] I. Jarraya, W. Ouarda, and A. M. Alimi, "A preliminary investigation on horses recognition using facial texture features", Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC), Hong Kong, China, pp. 2954-2959, Oct. 2015. <https://doi.org/10.1109/SMC.2015.520>.
- [10] D. Deb, J. W. Bohnacker, A. K. Jain, and D. White, "Face recognition: Primates in the wild", Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS), Los Angeles, CA, USA, pp. 1-10, Oct. 2018. <https://doi.org/10.1109/BTAS.2018.8698502>.
- [11] P. Cihan, S. C. Inan, E. A. Cansu, and S. Kalkan, "Identification and recognition of animals from biometric markers using computer vision approaches: a review", Kafkas Univ. Vet. Fak. Derg., Vol. 29, No. 6, pp. 581-593, Dec. 2023. <https://doi.org/10.9775/kvfd.2023.29936>.
- [12] M. H. Khan, Y. Zhang, M. Hayat, J. Shen, L. Shao, and T. Tuytelaars, "AnimalWeb: A large-scale hierarchical dataset of annotated animal faces", Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Seattle, WA, USA, pp. 6939-6948, Jun. 2020. <https://doi.org/10.1109/CVPR42600.2020.00796>.
- [13] A. Ayanzadeh and S. Vahidnia, "Modified deep neural networks for dog breeds identification", Preprints, Oct. 2018. <https://doi.org/10.20944/preprints201810.0087.v1>
- [14] T. P. Moreira, D. Menotti, W. R. Schwartz, and A. Rocha, "Where is my puppy? Retrieving lost dogs by facial features", Multimedia Tools Appl., Vol. 76, pp. 15325-15340, Jun. 2017. <https://doi.org/10.1007/s11042-016-3864-3>.
- [15] B. Yoon, H. So, and J. Rhee, "A methodology for utilizing vector space to improve the performance of a dog face identification model", Appl. Sci., Vol. 11, No. 5, pp. 2074, Mar. 2021. <https://doi.org/10.3390/app11052074>.
- [16] D. S. Lee and G. M. Park, "A Study on Improving Facial Recognition Performance to Introduce a New Dog Registration Method", JBE, Vol. 27, No. 5, pp. 794-807, Dec. 2022. <https://doi.org/10.5909/JBE.2022.27.5.794>.
- [17] G. T. Byeon and C. G. Kang, "Design and Implementation of a Missing Pet Search

Systemusing Deep Learning-based Image Analy", Journal of KIIT, Vol. 23, No. 5, pp. 203-209, May 2025. <https://doi.org/10.14801/jkiit.2025.23.5.203>.

- [18] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer", Journal of Machine Learning Research, Vol. 21, No. 140, pp. 1-67, Oct. 2020.
- [19] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition", Proc. ICML Deep Learning Workshop, Lille, France, Vol. 2, No. 1, pp. 1-8, Jul. 2015.
- [20] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer", Journal of Machine Learning Research, Vol. 21, No. 140, pp. 1-67, Oct. 2020.

저자소개

부 석 준 (Seok-Jun Buu)



2023년 3월 : 연세대학교
컴퓨터과학과(공학박사)
2023년 9월 ~ 현재 :
경상국립대학교 컴퓨터공학과
조교수
관심분야 : 인공지능, 딥러닝

최 상 민 (Sang-Min Choi)



2015년 2월 : 연세대학교
컴퓨터과학과(공학박사)
2022년 3월 ~ 현재 :
경상국립대학교 컴퓨터공학과
조교수
관심분야 : 추천시스템, 알고리즘

이 수 원 (Suwon Lee)



2017년 7월 : 한국과학기술원
전산학과(공학박사)
2018년 3월 ~ 현재 :
경상국립대학교 컴퓨터공학과
부교수
관심분야 : 증강현실, 컴퓨터비전,
인간-컴퓨터 상호작용

연 수 민 (Sumin Yeon)



2022년 3월 ~ 현재 :
경상국립대학교 컴퓨터공학과
학사과정
관심분야 : 인공지능, 딥러닝

배 지 호 (Jiho Bae)



2025년 2월 : 경상국립대학교
컴퓨터공학과(공학사)
2025년 3월 ~ 현재 :
경상국립대학교 컴퓨터공학과
석사과정
관심분야 : 인공지능, 컴퓨터비전,
증강현실