

비지도 도메인 적응 기반 위장 객체 분류

장소현, 박현희*

명지대학교

{jangso6593, hhpark*}@mju.ac.kr

Unsupervised Domain Adaptation for Camouflaged Object Classification

Jang So Hyeon, Park Hyun Hee*

Department of Applied Artificial Intelligence, Myongji University

요약

위장 객체 분류는 객체와 배경 간 시각적 유사성으로 인해 정밀한 수동 레이블 구축에 어려움이 따른다. 본 논문은 레이블이 풍부한 비위장 객체 데이터를 소스 도메인으로, 레이블이 없는 위장 객체 데이터를 타겟 도메인으로 설정한 비지도 도메인 적응 기반 위장 객체 분류 프레임워크를 제안한다. LVIS 기반 소스 데이터와 COD10K 기반 타겟 데이터를 fish, cat, bilrd, spider 4개의 class에 대한 실험을 수행한 결과, 제안한 방법은 비지도 도메인 적응을 적용하지 않은 source-only 설정 대비 평균 분류 정확도를 68.06%에서 86.86%로 크게 향상시켰으며 특히 fish 클래스에서 31.83%에서 96.06%로 현저한 성능 개선을 달성하였다. 이러한 결과는 레이블이 제한적인 위장 객체 환경에서도 비지도 도메인 적응이 효과적인 분류 성능 향상 방안이 될 수 있음을 입증한다.

I. 서론

위장 객체 분류는 주변 환경과 시각적으로 유사한 객체의 클래스를 판별하는 작업으로, 배경과의 경계가 불분명해 인식이 어렵고 정밀한 레이블 구축에 높은 비용이 요구된다[1]. 이러한 데이터 확보의 어려움은 대규모 학습 데이터를 기반으로 하는 딥러닝 모델 연구에 실질적인 제약으로 작용한다. 기존 연구는 주로 탐지 및 분할[1]에 집중되어 왔으며 객체의 범주를 식별하는 분류 문제에 대한 논의는 상대적으로 부족한 상황이다. 또한 합성곱 신경망은 객체의 형태 정보보다 배경 및 질감 특징에 편향되어 [2] 배경과 질감 및 색상을 공유하는 위장 객체는 분류 성능 저하의 원인으로 작용할 수 있다.

본 논문은 이러한 문제를 해결하기 위해 비지도 도메인 적응 기반 위장 객체 분류 기법을 제안한다. 비지도 도메인 적응은 레이블이 풍부한 소스 도메인의 지식을 레이블이 없는 타겟 도메인으로 전이하는 기법이다[3]. 본 연구에서는 레이블이 풍부한 비위장 객체 데이터를 소스 도메인으로, 레이블이 없는 위장 객체 데이터를 타겟 도메인으로 설정하여 레이블이 없는 환경에서도 안정적인 위장 객체 분류 성능을 확보하고자 한다.

II. 제안 기법

본 연구에서는 레이블이 있는 소스 도메인 데이터를 활용한 지도 학습을 통해 기본적인 특징 추출 및 분류 모델을 학습한다. 그림 1과 같이 소스 입력 이미지는 parameter-efficient fine-tuning (PEFT)[4]이 적용된 인코더를 통해 특징 맵으로 변환된 후 전역 평균 풀링(GAP)을 거쳐 전역 특징 벡터로 집계된다. 집계된 특징은 분류기를 통해 클래스 예측으로 변환되고 해당 예측은 소스 도메인의 정답 레이블과의 지도 손실을 통해 학습된다. 이 단계에서 학습된 인코더와 분류기는 교사-학생 기반 도메인 적응을 위한 초기 모델로 사용된다.

이후 교사-학생 구조를 기반으로 한 비지도 도메인 적응을 수행한다. 그림 2와 같이 교사 모델은 약하게 증강된 타겟 도메인 입력으로부터 고신뢰 의사 레이블을 생성하고 학생 모델은 강하게 증강된 타겟 데이터에 대해 FixMatch[5]를 적용하여 손실 값 손실 $L_{fixmatch}$ 로 교사 모델이

생성한 의사 레이블과의 일관성을 학습한다. 동시에 학생 모델은 소스 도메인 데이터에 대한 지도 손실 L_{source} 을 병행하여 학습함으로써 도메인 적응 과정에서 분류 기준의 붕괴를 방지한다.

학생 모델이 최종 분류를 위한 목적 함수는 다음과 같다.

$$L_{total} = L_{fixmatch} + \lambda_u L_{source} \quad (1)$$

교사 모델은 학생 모델 파라미터의 지수 이동 평균(exponential moving average, EMA)으로 갱신되며, 최종적으로 학습된 학생 모델을 사용하여 타겟 도메인에서 위장 객체의 클래스를 예측한다.

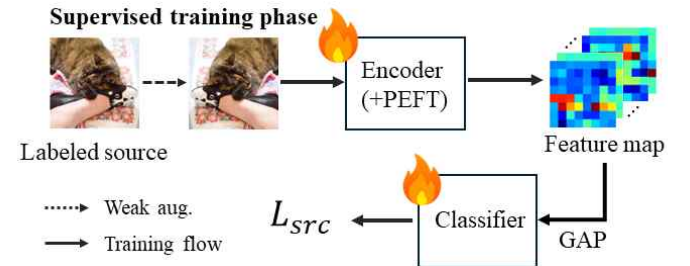


그림 1. 레이블 기반 소스 도메인 지도 학습

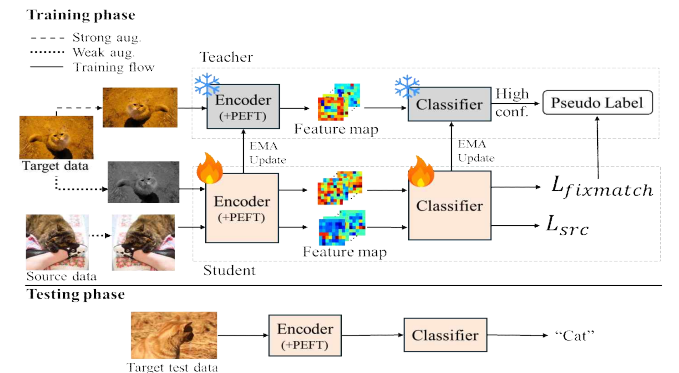


그림 2. 교사-학생 기반 비지도 도메인 학습 파이프라인

III. 실험 및 결과

소스 도메인 데이터는 LVIS[6] 데이터셋을 사용하여 fish, spider, bird, cat 클래스에 해당하는 샘플을 추출하였고 데이터 수가 부족한 fish와 spider 클래스는 Roboflow 플랫폼을 통해 추가 수집하여 보완하였다. 타겟 도메인으로는 COD10K-v3[7] 데이터셋에서 fish, spider, bird, cat에 해당하는 샘플을 추출하여 사용하였으며, 타겟 데이터의 클래스 레이블은 학습 단계에서는 사용되지 않고 평가 단계에서만 활용하였다.

제안한 비지도 도메인 적응 기반 위장 객체 분류 기법의 유효성을 검증하기 위해 비지도 도메인 적응을 수행하지 않은 source-only 모델을 비교 기준으로 설정했다. 표 1은 각 모델의 클래스별 분류 정확도를 나타낸다.

표 1. 클래스별 분류 정확도

class	fish	cat	bird	spider	mean
source-only	31.83	54.11	94.17	92.12	68.06
proposed	96.06	57.53	97.50	96.36	86.86

실험 결과, source-only 모델은 위장 객체에 대해 전반적으로 성능 저하가 발생하였으며, 특히 fish 클래스의 경우 정확도가 31.83%로 매우 낮게 나타났다. 반면 제안한 기법을 적용한 경우 fish 클래스에서 96.06%로 크게 개선되어 가장 높은 향상폭(+64.23%p)을 보였으며 평균 분류 정확도는 68.06%에서 86.86%로 18.80%p 향상되었다. 다만 cat 클래스는 57.53%로 상대적으로 낮은 성능을 보였는데, cat 클래스는 57.53%로 상대적으로 낮은 성능을 보였는데, 이는 표 2에 제시된 클래스별 데이터 분포에서 확인할 수 있듯이 cat 클래스의 데이터 수가 상대적으로 적어 충분한 도메인 적응이 이루어지지 못했기 때문으로 분석된다.

표 2. 클래스별 데이터 분포

class		fish	cat	bird	spider	total
source data	train	2,451	1,918	2,451	4,135	12,780
	test	114	376	470	781	1,734
target data	train	563	212	260	186	1,221
	test	355	146	165	120	786

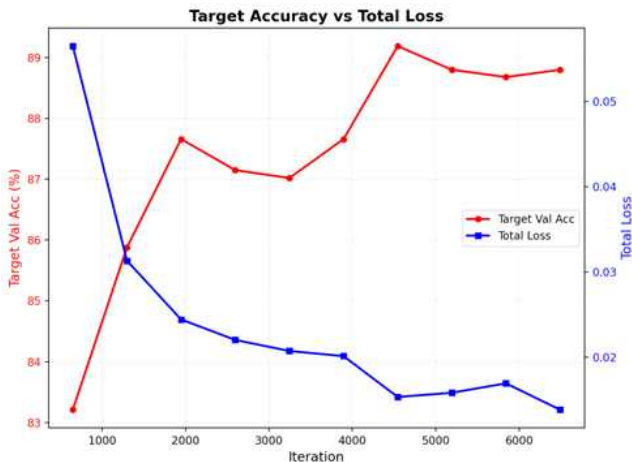


그림 3 비지도 도메인 적응 과정에서의 성능 및 손실 변화
그림 3은 제안 기법의 학습 과정을 보여준다. 학습 손실은 초기 0.056에서

0.014 수준으로 안정적으로 감소하였으며 검증 정확도는 83.2%에서 시작하여 최고점인 89.2%에 도달한 뒤 안정적인 성능을 유지하였다. 이는 본 모델이 위장 객체의 특징을 효과적으로 학습하며 도메인 간 간극을 안정적으로 극복하고 있음을 입증한다.

IV. 결론

본 논문에서는 배경과의 높은 시각적 유사성으로 인해 식별이 어려운 위장 객체 분류 문제를 비지도 도메인 적응 관점에서 접근하는 프레임워크를 제안하였다. 제안한 방법은 레이블이 풍부한 비위장 객체의 지식을 교사-학생 구조를 통해 전이함으로써 위장 객체 데이터셋의 클래스 레이블 부족 문제를 해결하였다.

실험 결과 source-only 모델과 비교하여 모든 클래스에서 일관된 분류 성능 향상을 달성하였다. 특히 도메인 간 분포 차이가 큰 클래스에서도 안정적인 성능 개선을 보였으며 추가적인 수동 레이블링 없이도 평균 분류 정확도를 크게 향상시킴으로써 위장 객체 분류 환경에서 제안 방법의 효과성과 실용성을 입증하였다.

ACKNOWLEDGMENT

이 논문은 산업통상자원부 및 산업기술기획평가원 (KEIT)과 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1A2C2005705, 분산 머신러닝 기반 지능형 플라잉 기지국을 위한 AI-MAC 프로토콜, No. RS-2024-00469138, 국산 AI반도체 기반 비전인식기술 소프트웨어개발도구 개발)

참고 문헌

- [1] C. Lei, et al., "Towards real zero-shot camouflaged object segmentation without camouflaged annotations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 47, no. 12, pp. 11990-12004, Dec. 2025.
- [2] D. Kang, et al., "Progressive and multi-task learning based salient detection method for camouflage object in defense boundary surveillance," in *Proc. Inst. Electron. Eng. Korea Summer Conf.*, pp. 2599-2601, Aug. 2020.
- [3] B. H. Ngo, et al., "Learning CNN on ViT: A hybrid model to explicitly class-specific boundaries for domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 28545-28554, Seattle, WA, USA, June 2024.
- [4] Q. Li, et al., "Parameter efficient fine-tuning on selective parameters for transformer-based pre-trained models," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, pp. 1 - 6, Niagara Falls, ON, Canada, June 2024.
- [5] K. Sohn, et al., "FixMatch: Simplifying semi-supervised learning with consistency and confidence," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 33, pp. 596-608, 2020.
- [6] A. Gupta, et al., "LVIS: A dataset for large vocabulary instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 5351-5359, Long Beach, CA, USA, June 2019.
- [7] D.-P. Fan, et al., "Camouflaged object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 2774-2784, Seattle, WA, USA, June 2020.