

LocationMatch: 객체 감지에 대한 일관성 있는 반지도 학습을 사용한 단순화된 접근 방식

유창성, 김형원*
충북대학교

csyoo@chungbuk.ac.kr, *hwkim@chungbuk.ac.kr

LocationMatch: A Simplifying Approach with Consistency Semi-Supervised Learning on Object Detection

Yoo Chang Seong, Kim Hyung Wom*
ChungBuk National Univ

요 약

준 지도 학습은 레이블이 있는 데이터와 레이블이 없는 데이터를 함께 활용하여 레이블이 있는 데이터에 대한 의존도를 줄이는 방법으로 주목받고 있다. 그러나 분류 분야를 주로 발전한 준 지도 학습은 다른 분야에서는 기대만큼의 성능을 보이지 못하고 있어 일반화에 한계가 있다. 특히, 준 지도 학습 알고리즘인 MixMatch는 분류 작업에 맞춰 설계되었으며, 해당 분야에서는 SOTA 성능을 달성했다. 그러나 객체 탐지 작업에 MixMatch를 그대로 적용하는 데에는 문제가 있음을 발견했다. 이 문제를 해결하기 위해 새로운 준 지도 학습 방법론인 LocationMatch를 제안했다. LocationMatch는 학습 비용 측면에서 효율적이며, 여러 객체 탐지 벤치마크 데이터셋에서 우수한 성능을 발휘했으며, 객체 탐지 분야의 새로운 가능성을 제시한다.

I. 서 론

데이터의 양과 질은 딥러닝 결과에 가장 큰 영향을 미치는 요소이다. 그러나 데이터셋 규모가 증가함에 따라, 레이블을 만드는 비용 또한 증가하고, 이에 따라 레이블이 없는 데이터를 활용하려는 준 지도 학습이 주목받고 있다. 준 지도 학습은 비용 절감과 성능 향상 측면에서 우수함을 입증했으며, 분류를 시작으로 객체 탐지, 이상 감지, 포즈 추정 등 다양한 분야에 적용되고 있다.

여러 분야에 준 지도 학습이 적용되었지만, 분류 분야에서 보여준 성공적인 결과만큼 다른 분야에서 성과를 내지 못했다. 본 논문은 준 지도 학습 중 MixMatch 방법론을 사용하여, 레이블이 있는 데이터와 레이블이 없는 데이터를 효과적으로 결합하여 성능을 높이는 방법을 제시했다. MixMatch는 분류 분야에서 당시 SOTA 결과를 보였지만, 객체 탐지 분야에서는 아직 성공적인 사례가 부족하다. 따라서 이번 논문에서는 객체 탐지 분야에 MixMatch를 적용하여 높은 성능을 달성하고자 한다. 본 연구에서는 MixMatch 방법론을 객체 탐지 분야에 적용하기 위해 몇 가지 주요 개선점을 도입했다. 첫째, 객체 탐지 특성에 맞춘 데이터 증강 기법을 새롭게 설계하여, 레이블이 없는 데이터에서도 더 풍부한 정보를 추출할 수 있도록 했다. 둘째, 탐지 네트워크 구조를 최적화하여, 다양한 크기의 객체에 대해 더 나은 성능을 발휘하도록 개선하였다. 이러한 접근 방식을 통해, 준

지도 학습의 이점을 객체 탐지 문제에서도 최대한 활용하고자 했다. 제안 방법은 객체 탐지의 벤치마크 데이터셋인 VOC 2007, 2012, COCO에서 다른 준 지도 학습보다 높은 성능 향상을 보였다.

II. 본론

1. 관련 연구

1.1 MixMatch

MixMatch는 준 지도 학습의 대표적인 방법론으로, 레이블이 있는 데이터와 레이블이 없는 데이터를 효과적으로 결합하여 학습 성능을 향상시키는 방법론이다. MixMatch는 레이블이 있는 데이터와 레이블이 없는 데이터 간의 불확실성을 줄이기 위해, 데이터 증강 및 예측 앙상블을 활용한다. 이 방법론은 특히 이미지 분류 작업에서 뛰어난 성능을 보였으며, 당시 최신 기술 수준(SOTA)을 달성했다. MixMatch는 강력한 일반화 성능을 제공하며, 레이블이 부족한 상황에서도 높은 정확도를 유지할 수 있는 장점이 있다. 이 방법은 다양한 데이터셋에서 검증되었으며, 준지도 학습의 유용성을 입증했다.

1.2 일관성 규제

일관성 규제는 준 지도 학습에서 레이블이 없는 데이터의 예측 일관성을 유지하려는 접근 방식이다. 이 방법론은 입력 데이터에 작은 노이즈를 추가하거나, 다양한 데이터 변환을 적용하여 모델의 예측이

안정적이도록 유도한다. 일관성 규제는 데이터 증강 기법과 결합되어 모델이 다양한 변형에 대해 일관된 출력을 생성하도록 학습한다. 이 방법은 특히 데이터의 다양성이 큰 상황에서 모델의 일반화 능력을 향상시키는 데 효과적이다. 다양한 연구에서 일관성 규제는 준 지도 학습의 성능을 향상시키는 데 기여했으며, 특히 객체 탐지와 같은 복잡한 작업에서도 유용한 성과를 보였다.

1.3 엔트로피 최소화

엔트로피 최소화는 모델이 더 확신 있는 예측을 하도록 유도하는 기법이다. 이 방법은 모델의 출력 확률 분포에서 불확실성을 줄이는 것을 목표로 한다. 레이블이 없는 데이터에서 특히 중요하며, 모델이 애매한 예측 대신 특정 클래스에 높은 확신을 가지도록 도와준다. 이를 통해 학습 과정에서 레이블이 부족한 상황에서도 더욱 견고한 성능을 발휘할 수 있다. 엔트로피 최소화는 준 지도 학습에서 일관된 예측을 유지하고 성능을 높이는 데 중요한 역할을 한다.

2. LocationMatch

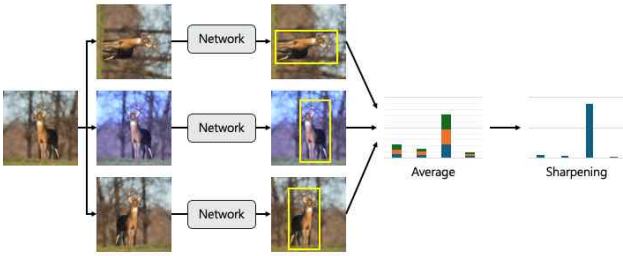


그림 1. Location Guessing 의 흐름도

2.1 Location Guessing

LocationMatch 는 레이블이 없는 데이터에 대해 K 번 증강을 수행한 후 각 증강된 이미지에 대해 개별적으로 객체탐지기의 예측을 수행한다. 이때 예측된 바운딩 박스들을 평균을 내고 더 정확한 위치를 추정하며, sharpening 기법을 적용하여 모델의 예측을 더 극대화하여 확실하게 만든다. Sharpening 은 모델이 예측한 확률 분포를 더욱 극단적으로 만들어, 엔트로피가 높은 애매한 예측을 줄이고 엔트로피를 낮추어 확신 있는 예측으로 만드는 방법이다.

$$Sharpen(p, T)_i = \frac{p_i^{\frac{1}{T}}}{\sum_{j=1}^L p_j^{\frac{1}{T}}}$$

각 증강된 이미지에 대해 예측된 바운딩 박스의 좌표와 클래스 확률을 합산하고 평균 내어 일관성 규제를 적용한다. 또한, 평균을 낸 예측된 바운딩 박스의 좌표와 클래스 확률에 sharpening 을 적용하여 entropy minimization 의 효과를 얻을 수 있다. 이러한 과정은 레이블이 없는 데이터에서도 신뢰할 만한 바운딩 박스를 추정할 수 있게 한다.

2.2 데이터 증강

다른 전형적인 준 지도 학습 방법론들과 마찬가지로, 레이블이 있는 데이터와 레이블이 없는 데이터 모두에 대해 데이터 증강을 진행한다. 사용한 데이터 증강 기법으로는 회전, 뒤집기, 색상 변화, 노이즈 추가, 가우시안 노이즈 그리고 컷아웃이 적용되었다.

III. 실험 및 결과

Method	VOC 2007	VOC 2012	COCO
Pseudo-Label	69.8 mAP	72.3 mAP	72.7 mAP
LocationMatch	70.6 mAP	73.2 mAP	73.5 mAP

본 연구에서는 객체 탐지 모델로 YOLOv5 를 사용했으며, 제안된 LocationMatch 방법론의 성능을 평가하기 위해 Pseudo-Label 방법론과 비교를 진행했다. 성능 평가 지표로는 평균 정확도(mAP)를 사용하였으며, 실험에는 VOC2007, VOC2012, COCO 데이터셋이 활용되었다. 각 데이터셋에서의 결과, LocationMatch 는 모든 데이터셋에서 Pseudo-Label 방법론을 능가하는 성능을 보였다. VOC2007, VOC2012 데이터셋에서는 기존 방법론 대비 mAP 가 각각 0.8, 0.9 향상되었으며, COCO 데이터셋에서도 유의미한 성능 향상을 확인할 수 있었다.

III. 결론

본 연구에서는 객체 탐지 분야에서 준 지도 학습의 성능을 극대화하기 위해 새로운 방법론인 LocationMatch 를 제안하였다. LocationMatch 는 레이블이 없는 데이터에 대해 효과적인 바운딩 박스 추정을 통해 기존의 Pseudo-Label 방법론을 능가하는 성능을 보여주었다. 이 연구는 객체 탐지 모델에서 준 지도 학습의 적용 가능성을 확대하는 중요한 발걸음이 될 수 있으며, 레이블 비용을 줄이면서도 우수한 성능을 유지할 수 있는 새로운 접근법을 제시한다.

ACKNOWLEDGMENT

Put This Work was supported by Regional Leading Research Center (RLRC) of the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No.2022R1A5A8026986) and supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2020-0-01304, Development of Self-Learnable Mobile Recursive Neural Network Processor Technology). It was also supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Communication Technology Research Center support program (IITP-2024-2020-0-01462) supervised by the IITP (Institute of Information & communications Technology Planning & Evaluation).

참 고 문 헌

- [1] Xie, Q., Dai, Z., Hovy, E., Luong, M. T., & Le, Q. (2019, April 29). *Unsupervised Data Augmentation for Consistency Training*. arXiv.org. <https://arxiv.org/abs/1904.12848>
- [3] Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A., & Raffel, C. (2019, May 6). *MixMatch: A Holistic Approach to Semi-Supervised Learning*. arXiv.org. <https://arxiv.org/abs/1905.02249>