

## 준지도 학습 기반 객체 탐지 모델에서 데이터셋 변화에 따른 성능 변화

유승수, \*황원준

아주대학교

ssyu@ajou.ac.kr, \*wjhwang@ajou.ac.kr

### Performance Change according to Data Set Size Change in Semi-Supervised Learning based Object Detection

Seungsoo Yu \*Wonjun Hwang

Ajou University

#### 요 약

Semi Supervised Learning 은 일부의 data 에는 labeling 을 하고 나머지 data 에는 labeling 을 안한채로 학습을 진행하는 방법이다. Object Detection 은 이미지에서 여러개의 객체들의 대한 위치를 여러개의 바운딩 박스로 지정해서 찾는 Computer Vision task 이다. 당연하게도, model training 단계에서 사용되는 data set 의 크기가 크고 객체가 많을 수록 일반적으로 model 의 성능이 좋아 질 것이다. 하지만 실험 환경에 따라 data set 을 잘 확보하지 못하던가, 실험 장치가 데이터 셋을 감당하지 못하는 등의 문제가 발생 할 수 있다. 그렇기에 본 논문에서는 semi supervised learning based object detection model 을 알아보고 data set 의 크기를 조절해가며 model 을 training 시킨 뒤 data set 의 크기에 따라 성능이 어떻게 변화하는 지를 알아 볼 것이다.

#### 1. 서론

Computer Vision 분야에서는 Image 를 Input 으로 주고 model 이 학습을 거쳐 어떤 특정한 task 를 수행하기 때문에 Image 의 양과 label 이 중요하다. 하지만 labeling 작업은 일일이 하기에 굉장한 시간과 비용이 들고 labeling 만으로 실제 세상에 존재하는 모든 data 들의 분포나 특징을 포함 시킬 수는 없다. 따라서, 최근에는 일부 data 는 labeling 을 하고 나머지 data 에는 labeling 을 하지 않으면서도 성능을 더 좋게 만드는 Semi-Supervised Learning[1]이 연구되고 있다.

Object Detection 분야도 Semi-Supervised Learning 을 기반으로 많은 model 들이 연구되고 있다. 본 논문에서 쓰일 unbiased teacher model[2] 또한 대표적인 Semi-Supervised learning 기반 model 이다. 본 논문에서는 unbiased teacher

model 에 data set 비율을 달리하여 학습을 하고 성능 변화를 볼 것이다.

#### 2. Unbiased Teacher

많은 Semi-Supervised 기반 object detection model 을 보면, teacher student model 을 사용하고 있다. Unbiased Teacher 또한 Object Detection 을 수행하는 teacher student model 이다. 그림 1 에서 보는 것처럼, Unbiased teacher 는 학습을 하는데 있어 크게 2 stage 로 나뉘어 진다.

첫번째 Stage 에서는 보통의 Supervised Learning 처럼 label 이 되어있는 data 들을 object detector 에 학습시킨다. pre-training 을 통해 model 은 어느정도 객체 탐지에 필요한 weight 값을 가지게 된다.

두번째 Stage 에서는 앞서 training 시켰던 model 을 복제하

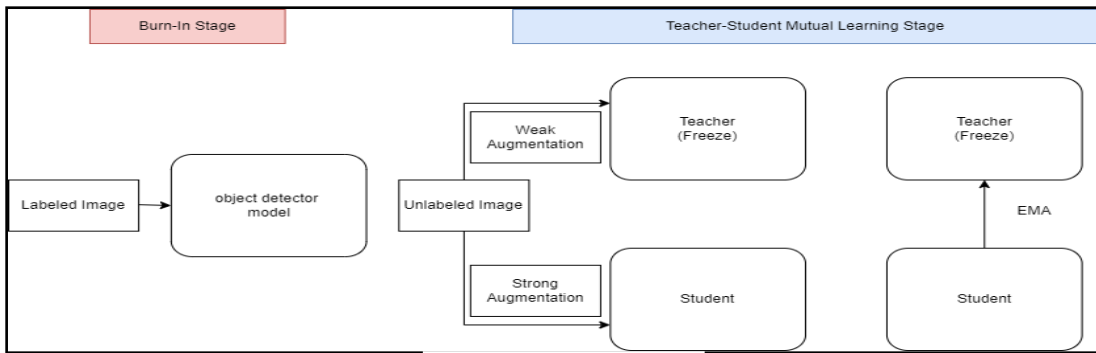


그림 1. Unbiased Teacher

여 Teacher, Student 2 개로 만들고, Teacher 는 Student 가 배울 수 있게 unlabeled image 에 pseudo labeling 을 한다. 두 model 에 들어갈 image 에는 augmentation 이 적용되는데, teacher 에는 weak augmentation 이, student 에는 strong augmentation 이 적용된다. pseudo labeling[3]이란 model 이 예측을 통해 unlabeled image 에 가짜로 label 을 부여하는 기술을 말한다. student 는 pseudo labeling 된 teacher 의 image 로 학습을 진행하고 student 가 배운 지식이 EMA 를 통해 teacher 로 전이된다. EMA[4]는 최신 data 에 더 높은 가중치를 부여하여 평균을 구하는 방법이다. 따라서 이 과정을 반복하면 teacher 가 student 에게 EMA 방식으로 update 되어지기 때문에 다른 시간 때의 student model 들의 ensemble 로 여겨질 수 있다. 결과적으로, teacher model 이 정확도가 높은 pseudo label 을 만들어 낼 수 있게 된다.

### 3. Data Set Size

보통 computer vision 에서 학습을 할 때, GPU 를 사용하여 학습을 진행한다. 하나의 이미지에서조차 수많은 pixel 에 대해서 반복적인 연산을 하기 때문이다. 그런 이미지를 많이 모아놓은 data set 을 학습하려면 많은 수의 GPU 가 distributed data parallel system 을 통해 학습을 진행해야 한다. 하지만 상황에 따라서는, 많은 수의 GPU 를 확보 할 수 없거나 data set 을 잘 모으지 못하는 경우가 발생 할 수 있다. 그렇기에 본 논문에서는 data set 을 작게하여 실험 결과를 얻고, 이 결과를 바탕으로 model 의 성능을 개선한다면 후에 data set 의 크기를 늘렸을 때, 비슷하게 성능이 좋아질 수 있는 가능성이 있는지 확인하기 위해 data set 의 size 를 조절해가며 성능을 실험해보려 한다.

### 4. Unbiased teacher with Data Set Change

data set 은 object detection 에서 많이 사용하는 data set 인 PASCAL VOC 원본(100%), 50%, 25%, random drop (50%)를 이용하였다.

Labeled Data	Unlabeled Data	Batch Size	AP 50
VOC07	VOC12	8	80.51

표 1. Original Research Result

위의 표 1 은 논문에 나와있는 실험 결과이다. 8 대의 GPU 를 사용하였고 batch size 를 8 로 하였을 때 AP 50 에서 80.51 의 결과를 얻었다고 한다. 본 논문에서는 data set 은 똑같이 하고 GPU 1 대, batch size 2, iteration 18000 으로 조정하여 data set 의 크기를 바꿔가면서 실험해보았다.

data set size	training time	Batch Size	AP 50
100%	7 hour	2	64.84
50%	3hour	2	64.41
50% random	3hour	2	63.98

표 2. Data Set Change Result

위의 표 2 는 data set 의 변화에 따른 unbiased teacher model 의 결과를 보여주고 있다. 당연히 data set 의 크기, batch size, gpu 개수를 모두 줄였기 때문에 100%를 썼을 때, 원래 결과보다 떨어 진 것이 보인다. 하지만 IOU > 0.5 일 때 맞는 걸로 평가한 AP 50 수치를 보면 100%나 50%, random drop 이 별 차이를 보이지 않고 학습 시간도 많이 줄어들어 여건이 안될 때에는 data set 의 size 를 줄여도 어느정도 원래 data set 을 반영한다는 것을 알 수 있다. 따라서, 줄인 data set 의 결과로부터 model 을 개선하고 성능을 끌어올리면 나중에 data set 의 크기를 원복하거나 확보하였을 때, 성능이 비슷하게 올라갈 가능성이 있다는 것을 알 수 있다.

### 5. 결론 및 향후 연구

이 논문에서는 data set 변화에 따른 object detection model 의 성능 변화를 알아보았다. 표 2 에서 보이는 것과 같이

여러 제약이 있을 때 data set 의 크기를 줄여도 원래 data set 을 어느정도 반영한다는 결과를 볼 수 있었다. 물론 data set 은 클 수록 좋지만 사전에 data set 을 작게 하여 결과를 얻고 그것을 바탕으로 model 의 성능을 개선 시킨다면, 향후 data set 을 원복시켰을 때도 마찬가지로 성능 향상을 기대해 볼 수 있겠다.

단순히 data 비율 말고도 data 의 객체 수 비율에 따라 어떻게 결과가 반영 되는 지를 알아보는 연구로도 확장 할 수 있을 것이다.

## 6. 사사

본 연구는 2022 년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음 (2022-0-01077)

## 7. Reference

- [1] Samuli Laine, Samuli Laine. TEMPORAL ENSEMBLING FOR SEMI-SUPERVISED LEARNING. In ICLR, 2017
- [2] Yen-Cheng Liu, Chih-Yao Ma, Zijian He, Chia-Wen Kuo, Kan Chen, Peizhao Zhang, Bichen Wu, Zsolt Kira1, Peter Vajda. UNBIASED TEACHER FOR SEMI-SUPERVISED OBJECT DETECTION. In ICLR, 2021.
- [3] Eric Arazo, Diego Ortego, Paul Albert, Noel E. O'Connor, Kevin McGuinness. Pseudo-Labeling and Confirmation Bias in DeepSemi-Supervised Learning. In IJCNN, 2020
- [4] Zhaowei Cai, Avinash Ravichandran, Subhansu Maji, Charless Fowlkes, Zhuowen Tu, Stefano Soatto. Exponential Moving Average Normalization for Self-supervised and Semi-supervised Learning. In CVPR, 2021