

# 객체 바운딩 박스와 원본 이미지 결합을 이용한 합성 데이터 생성 기법

이주혁<sup>1</sup>, 김미희<sup>2</sup>

<sup>1</sup>한경국립대학교 컴퓨터응용수학부

<sup>2</sup>한경국립대학교 컴퓨터응용수학부, 컴퓨터시스템연구소

email:{xpdlwm99, mhkim}@hknu.ac.kr

## Synthetic data generation technique using object bounding box and original image combination

Ju-Hyeok Lee<sup>1</sup>, Mi-Hui Kim<sup>2</sup>

<sup>1</sup>School of Computer Engineering & Applied Mathematics Hankyong National  
University

<sup>2</sup>School of Computer Engineering & Applied Mathematics, Computer System  
Institute Hankyong National University

### 요 약

딥러닝은 컴퓨터 비전의 상당한 발전을 기여했지만, 딥러닝 모델을 학습하려면 대규모 데이터 세트가 필요하다. 이를 해결하기 위해 데이터 증강 기술이 주목받고 있다. 본 논문에서는 객체 추출 바운딩 박스와 원본 이미지의 바운딩 박스를 결합하여 합성 데이터 생성기법을 제안한다. 원본 이미지와 동일한 범주의 데이터셋에서 참조 이미지의 객체를 추출한 다음 생성 모델을 사용하여 참조 이미지와 원본 이미지의 특징을 통합하여 새로운 합성 이미지를 만든다. 실험을 통해, 생성 기법을 통한 딥러닝 모델의 성능향상을 보여준다.

### 1. 서론

딥러닝을 통해 컴퓨터 비전 분야에서 이미지 분류, 객체 탐지, 세분화 등 다양한 문제를 해결할 수 있게 되었다. 이러한 기술들은 대량의 데이터셋을 필요로 하므로, 데이터의 양이 많을수록 성능이 향상되는 경향을 보인다. 하지만, 현실적으로 대량의 데이터를 수집하기 어려운 때가 많기 때문에, 데이터 증강 기술이 주목받고 있다.[1]

본 논문에서는 합성 데이터를 통해서 데이터 구축 문제를 해결한다. 다양한 샘플로 대규모 데이터 세트를 생성한 다음 딥러닝 모델을 교육하는 데 사용할 수 있기 때문이다. 본 논문에서는 추출된 객체 경계 상자를 원본 이미지와 결합하여 생성 모델의 입력 값으로 활용하여 합성 데이터를 생성하는 방식을 제안한다.

제안 방식은 객체 탐지와 생성 모델을 통해 새로운 생성이미지 데이터를 생성한다.

제안 방식을 검증하기 위해 이미지 분류 모델을 통해서 유효성을 판단한다.

### 2. 배경 지식

#### 2.1. 객체 탐지 모델

객체 탐지 모델은 이미지나 비디오에서 특정 객체의 위치를 찾아내는 모델을 의미한다. 이를 통해서 컴퓨터 비전 분야에서 활발하게 사용해 왔다.

객체 탐지 모델을 통해서 찾아낸 위치를 바탕으로 해당 객체의 클래스 라벨들을 결합하여 객체를 식별하고 분류할 수 있다.[2]

본 논문에는 객체 탐지 모델을 통해서 원본 이미지와 참조 이미지의 바운딩 박스를 얻고, 원본과 참조 이미지의 바운딩 박스를 생성 모델의 다중 입력값으로 사용한다.

#### 2.2. 생성 모델

생성모델은 입력값을 통해 분석한 모델이 새로운 데이터를 생성하는 모델이다. 생성하는 데이터로는 이미지, 음악, 텍스트 등 다양한 데이터 형태로 생성할 수 있으며, 딥러닝 분야에서 활발하게 사용중이다.[3]

본 논문에서는 객체 탐지 모델을 통해 얻은 바운

당 박스의 객체들을 생성모델의 입력값으로 활용하여 새로운 이미지 데이터를 생성하고자 한다.

### 3. 제안 방식

(그림 1)은 본 논문에서 제안하는 방식의 흐름도이다. 제안 방식을 이용하여 대형 데이터셋 구축 문제를 해결하고, 딥러닝 모델의 성능도 향상하고자 한다.

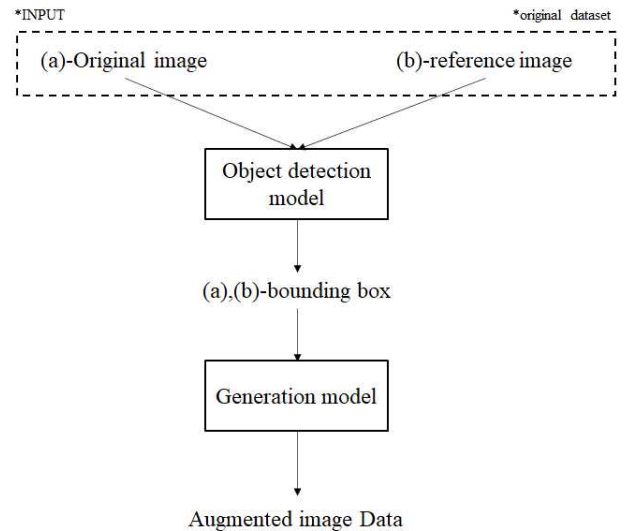
제안 방식은 기존에 존재하는 임의의 원본 데이터셋(original dataset)을 활용한다. 원본 데이터 셋의 이미지를 순차적으로 증강한다.(그림 1)-(a)는 순차적으로 원본 이미지(original image)를 지정한다. (그림 1)-(b)인 참조 이미지(reference image)는 원본 이미지를 제외하고 같은 카테고리의 원본 데이터 셋의 랜덤으로 하나를 지정한다. 이렇게 얻은 원본 이미지와 참조 이미지는 객체 추출 과정을 거친다.

#### 3.1. 객체 추출

원본 이미지와 참조 이미지에서 객체 바운딩 상자를 추출해야 한다. 이를 YOLO[4]와 같은 객체 탐지 모델 구조를 기반으로 설계했다. object detection model을 통해 얻은 바운딩 박스로 객체 이미지를 추출했다. 추출된 이미지는 생성 모델의 입력값으로 사용된다.

#### 3.2. 생성모델

3.1. 객체 추출을 통해 얻은 원본 및 참조 객체 이미지를 원본 객체 이미지를 기반으로 크기를 조정한다. DCGAN[5] 생성모델 구조를 기반으로 생성모델



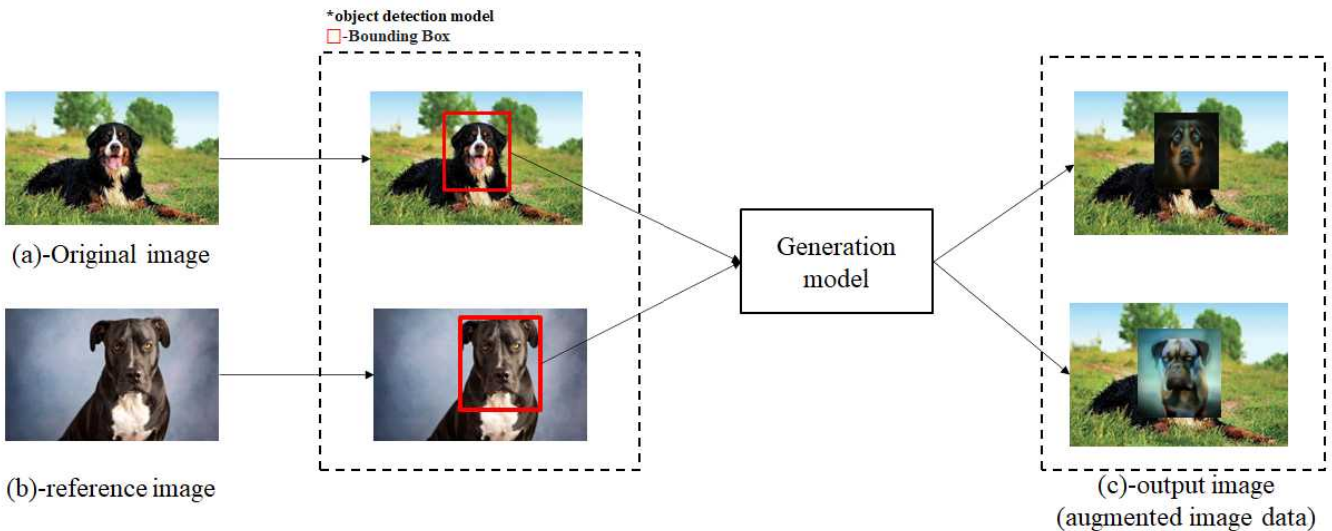
(그림 1) 객체 바운딩 박스와 원본 이미지 결합을 이용한 합성 데이터 생성 기법 흐름도

을 구축한다. 구축한 생성 모델을 통해 새로운 이미지 데이터를 생성한다. 생성한 이미지 데이터를 원본 객체 바운딩 박스 위치에 삽입하여 새로운 이미지 데이터를 생성 완료한다.

생성 모델을 통해 새로운 데이터 증강이 많으면 많을수록 새로운 이미지 데이터를 얻을 수 있다.

### 4. 실험 결과 및 평가

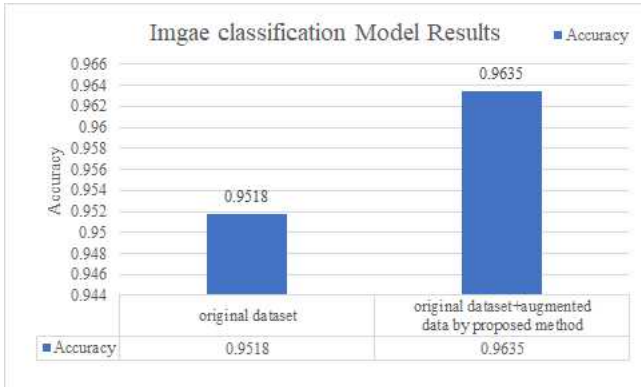
(그림 2)에서는 제안 방식을 통해 생성된 이미지 예시를 보여주고 있다. 생성된 이미지는 (그림 2)-(c)와 같이 합성된 이미지 데이터가 생성됐다. 본 논문에서는 강아지 데이터셋을 구축했으며, 원본 이미지로는 약 5,000장을 증강된 데이터셋으로는 약 10,000장으로 원본 이미지 데이터셋 보다 2배 많은 데이터를



(그림 2) 제안 기법을 통해 생성된 증강 이미지 데이터 예시

증강했다.

이미지 분류 모델은 딥러닝 모델인 합성곱 신경망의 구조를 바탕으로 설계했다. 본 논문에서 제안된 생성 기법을 통해, 이미지 분류 모델 결과에서도 많은 데이터를 학습했기에 원본 이미지 데이터 셋보다 정확도가 약 1.3%성능 향상 된 것을 볼 수 있다.



(그림 3) 이미지 분류 모델 결과

### 5. 결론 및 향후 연구

추출된 원본 및 참조이미지의 객체 바운딩 박스를 생성 모델의 입력 값으로 사용하여 합성된 이미지 데이터를 생성하는 방식을 제안했다. 제안된 방식을 통해 다양한 샘플로 대규모 데이터 세트를 생성할 수 있으며, 이를 딥러닝 모델을 교육하는 데 사용할 수 있다.

향후연구로서 바운딩 박스의 부분만 이미지를 생성하게 되어 (그림2)와 같이 매우 부자연한 이미지로 증강된다. 이미지 내의 다른 요소들을 바탕으로 생성된 이미지의 배경과 원본 이미지의 특징을 유사하게 개선하고자 한다.

### 6. Acknowledgement

이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2018R1A2B6009620), 교신저자 김미희.

### 참고문헌

[1] Connor Shorten & Taghi M. Khoshgoftaar , Data Augmentation Techniques for Deep Learning: A Review, Journal of Big Data, 6, 60, 1-48, 2019.  
 [2] Jiao, Licheng, et al. "A survey of deep learning-based object detection.", IEEE access, 7, 128837-128868, 2019.  
 [3] Iqbal, Touseef, and Shaima Qureshi. "The

survey: Text generation models in deep learning.", Journal of King Saud University-Computer and Information Sciences, 34. 6, 2515-2528, 2022

[4] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788, 2016.

[5] Fang, Wei, et al. "A method for improving CNN-based image recognition using DCGAN." Computers, Materials and Continua 57.1, 167-17, 2018.