

YOLOv5 학습 시 바운딩 박스 개수에 따른 화재 탐지 성능 비교

성영아 · 이현섭 · 장시웅*

동의대학교

Comparison of Fire Detection Performance according to the Number of Bounding Boxes for YOLOv5

YoungA Sung · Hyoun-Sup Yi · Si-Woong Jang*

Dong-Eui University

E-mail : youngasung0@gmail.com / lhskimj@naver.com / swjang@deu.ac.kr

요 약

YOLOv5에서 객체 탐지를 위해 이미지를 학습 시 기존의 이미지에 위치 정보를 어노테이션 하는 과정이 필요하다. 가장 대표적인 방법이 이미지에 바운딩 박스를 그려 위치 정보를 메타정보로 저장하게 하는 것이다. 하지만 객체의 경계가 모호한 경우 바운딩 박스를 하는 것에 어려움을 겪게 된다. 그 대표적인 예시가 화재인 부분과 화재가 아닌 부분을 분류하는 것이다.

따라서 본 논문에서는 화재가 났다고 판단되는 샘플 100개의 이미지를 바운딩 박스 개수를 달리하여 학습시켜 보았다. 그 결과 바운딩 박스를 어노테이션 시 가장자리를 가능한 크게 잡아 하나의 박스로 어노테이션하는 것보다 조금 더 세분화 하여 박스 3개로 어노테이션하여 학습시킨 모델에서 더 뛰어난 화재 탐지 성능을 보여주었다.

ABSTRACT

In order to detect an object in yolov5, a process of annotating location information on an existing image is required when learning an image. The most representative method is to draw a bounding box on an image to store location information as meta information. However, if the boundary of the object is ambiguous, it will be difficult to make a bounding box. A representative example would be to classify parts that are not fire and parts that are fire.

Therefore, in this paper, images of 100 samples judged to have caught fire were learned by varying the number of boxes. The results showed better fire detection performance in the model where the bounding box was trained by annotating it with three boxes by segmenting it slightly more than annotating it with one box by holding the edge as large as possible during annotating it with one box.

키워드

Object Detection, Fire Detection, YOLOv5, Bounding Box

I. 서 론

최근 인공지능을 이용하여 다방면에서 인간의 삶을 개선시키기 위한 노력을 하고 있다. 그중에서도 인간의 안전과 관련된 부분에서 사람의 인

지력보다 더 뛰어난 인공지능을 사용하여 인명과 재산을 보호하는 연구들이 진행되고 있다. 화재 탐지 분야에서도 인공지능을 사용하여 빠른시간에 화재를 탐지 하려는 연구가 대표적인 사례가 될 수 있다. 화재 발생 후 '5분'을 골든타임이라고 부른다. 화재 발생 5분이 지나면 인명피해와 재산 피해가 급속하게 증가하기 때문에 초기 진화에

* corresponding author

실패한다면 많은 손실을 남길 수 있다고 한다[1]. 따라서 골든타임인 5분안에 화재를 빠르고 정확하게 인식할 수 있는 알고리즘의 개발이 중요하다.

불꽃의 크기가 상황마다 다르고 산발적이기 때문에 화재의 경계가 모호하므로 바운딩 박스의 어노테이션에 어려움이 있다. 본 논문에서는 YOLOv5를 이용하여 바운딩 박스의 개수에 따른 화재 인식의 성능을 비교 분석하였다.

II. 기존 연구

화재 탐지의 성능을 높이기 위해 여러 가지 시도들이 있었다. 실시간으로 경계가 모호한 화재를 탐지하는 것에 대한 어려움이 있기 때문이다. 그 중 이미지의 텍스처를 이용한 분석도 있다[2]. 화재 탐지를 위해 (ㄱ)전처리, (ㄴ)화재 탐지, (ㄷ)특징 추출, (ㄹ)분류 과정들이 필요하다[2]. 요즘 실시간 객체 탐지를 위해 전처리 이외의 과정들을 한 번에 처리해 줄 수 있는 YOLO(You Only Look Once)가 많은 연구에서 이용되고 있다. YOLO에서도 다른 알고리즘과 마찬가지로 학습을 위해 인간이 화재라고 인식되는 부분에 어노테이션(라벨링)을 통해 학습 데이터를 생성해야 한다. 화재 탐지에 YOLO가 굉장히 성능이 좋은 시스템이라는 것이 많은 연구에서 보여준다[3, 4, 5].

III. 화재 탐지를 위한 어노테이션 방법

3.1 학습을 위한 이미지

머신러닝을 위한 커뮤니티 Kaggle에서 화재와 관련된 이미지를 취득하였고, 화재와 관련된 755개의 이미지중 바운딩 박스 개수에 대한 성능 비교를 위해 100개의 데이터 샘플을 이용하였다[6].

이미지를 어노테이션하기 위해서 여러 가지 도구들이 존재하지만 이번 연구에서는 YOLO의 데이터 형식을 지원해 주는 파이썬 라이브러리의 하나인 LabelImg로 100개의 샘플의 어노테이션을 진행하였다.

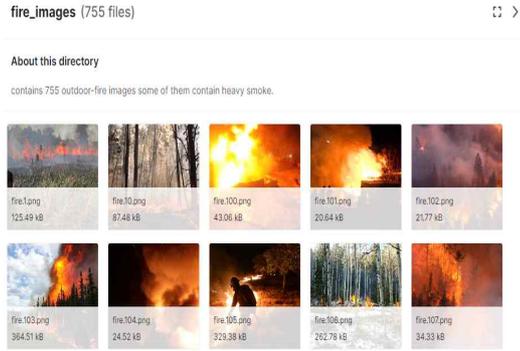


그림 1. kaggle에서 수집한 화재 이미지

3.2 바운딩 박스 1개로 어노테이션

LabelImg를 통해 100장의 학습 이미지에 예시 그림 2와 같이 화재라고 추정되는 이미지의 가장 자리를 가능한 크게 하여 1개의 바운딩 박스로 어노테이션하고 저장하였다. 표 1과 같이 한개의 사각형 각각의 모서리에 해당하는 위치 정보가 저장된다.



그림 2. 바운딩 박스 1개

표 1. 바운딩 박스 1개의 어노테이션 메타정보

위치 정보	0.508073	0.529630	0.618229	0.924074
-------	----------	----------	----------	----------

3.3 바운딩 박스 3개로 어노테이션

바운딩 박스의 개수에 따른 성능 비교를 위해 3.2절에서 사용한 이미지들과 동일한 100장의 학습 이미지에 화재 이미지를 세분화하여 3개의 바운딩 박스로 어노테이션하였다. 표 2에서와 같이 각각의 이미지마다 3개의 사각형 박스 이미지의 모서리 정보가 메타정보로 저장된다.



그림 3. 바운딩 박스 3개

표 2. 바운딩 박스 3개의 어노테이션 메타정보

위치 정보	0.621875	0.295370	0.252083	0.487037
	0.417187	0.725000	0.397917	0.322222
	0.727344	0.786111	0.193229	0.427778

IV. 어노테이션 방법에 탐지 성능 비교

어노테이션한 이미지를 학습시키기 위해서 구글 colab 환경을 이용하였다. 바운딩 박스 개수 이외에는 모든 조건이 동일한 환경에서 학습을 진행하였다. 본 연구에 사용된 YOLOv5에서의 학습 환경은 표 3과 같다. Epochs를 30이상으로 했을 때 바운딩 박스에 따른 성능 비교가 가능했다.

표 3. YOLOv5에서의 학습 환경

개발 모델	YOLOv5
개발 언어	Python-3.7.13
학습 이미지	100장
개발 환경	Colab
Epochs	30
Batch	10

본 논문에서 설계한 바운딩 박스 개수에 따른 성능을 비교하였으며, 결과는 그림 4, 그림 5와 같다. YOLOv5의 성능을 측정할 수 있는 대표적인 지표인 mAP, F1 score, Confidence 를 통해 그림 4와 그림 5를 비교할 수 있다. 바운딩 박스를 1개 했을 경우인 그림 4의 mAP는 30 epochs 할때까지 불안정하게 성능이 0.025로 향상 되는 반면 바운딩 박스를 3개 했을 경우인 그림 5를 확인해 보면 mAp가 바운딩 박스를 1개 했을 경우 보다 안정적으로 증가한다는 것을 확인할 수 있었다. Precision과 Recall 두가지의 조화를 측정하는

는 F1 score와 얼마나 객체를 잘 찾아냈는지를 측정하는 confidence 점수 또한 바운딩 박스를 3개로 어노테이션했을 경우 높게 측정되었다.

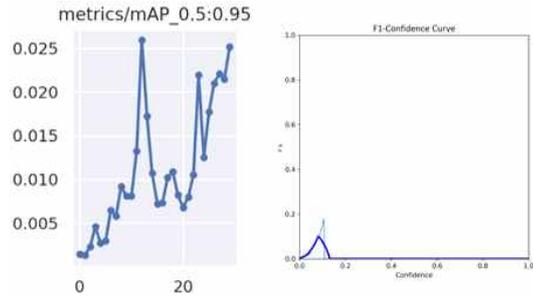


그림 4. 바운딩 박스 1개시 성능 결과

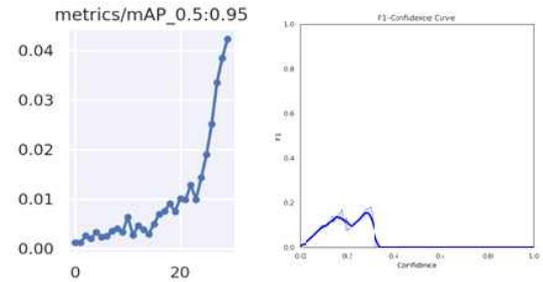


그림 5. 바운딩 박스 3개시 성능 결과

YOLOv5의 성능을 측정할 수 있는 대표적인 지표인 mAP, F1 score, Confidence 를 통해 그림 4와 그림 5를 비교할 수 있다. 바운딩 박스를 1개 했을 경우인 그림 4의 mAP는 30 epochs 할때까지 불안정하게 성능이 0.025로 향상 되는 반면 바운딩 박스를 3개 했을 경우인 그림 5를 확인해 보면 mAp가 바운딩 박스를 1개 했을 경우 보다 안정적으로 증가한다는 것을 확인할 수 있었다. Precision과 Recall 두가지의 조화를 측정하는 F1 score와 얼마나 객체를 잘 찾아냈는지를 측정하는 confidence 점수 또한 바운딩 박스를 3개로 어노테이션했을 경우 높게 측정되었다.

V. 결론

본 논문에서는 화재와 같이 경계가 애매한 이미지를 어노테이션하기 위해 박스를 바운딩할 때 바운딩 박스의 개수가 1개일 때보다 3개일 경우 화재 탐지의 성능이 향상된다는 것을 확인할 수 있었다.

실제로 학습에서 사용되지 않는 새로운 화재 이미지로 테스트해 보았을 때 바운딩 박스를 1개

로 학습시킨 weight 값을 가진 모델의 성능은 바운딩 박스 3개로 학습시킨 weight 가진 모델보다 현저히 떨어짐을 확인할 수 있었다. 어노테이션시 바운딩 박스의 가장자리를 가능한 한 크게 하여 바운딩 박스를 하나로 만드는 것보다 하나의 이미지를 세분화하여 3개의 바운딩 박스로 어노테이션 하는 것이 화재 탐지 성능을 많이 향상시킬 수 있었다.

이번 연구를 통해서 바운딩 박스 한개를 한 경우와 3개를 한 경우만을 비교 분석하였기 때문에 한계가 존재한다. 추후 더 많은 케이스의 바운딩 박스 개수에 따른 성능 연구를 수행할 필요가 있다.

Acknowledgement

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 지역지능화혁신인재양성(Grand ICT연구센터) 사업의 연구결과로 수행되었음” (IITP- 2022-2020-0-01791). 또한, 본 논문은 부산광역시 및 (재)부산인재평생교육진흥원의 BB21플러스 사업으로 지원된 연구임.

References

- [1] There is also a golden time for fires?
[Internet]. Available: <https://tech.hyundaimotorgroup.com/kr/developers-blog/ai-based-fire-identification-and-early-suppression-system/>
- [2] Chunyu, Yu, et al. “Texture analysis of smoke for real-time fire detection,” *2009 second international workshop on computer science and engineering*, Vol. 2. IEEE, 2009.
- [3] Wilson, Sneha, et al. “A Comprehensive Study on Fire Detection,” *2018 Conference on Emerging Devices and Smart Systems (ICEDSS)*, IEEE, 2018.
- [4] Feng, Jiang, et al. “Design and experimental research of video detection system for ship fire,” *2019 2nd International Conference on Safety Produce Informatization (IICSPI)*. IEEE, 2019.
- [5] Hongyu, Huang, et al. “An improved multi-scale fire detection method based on convolutional neural network,” *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*. IEEE, 2020.
- [6] Kaggle, AHMED SAIED, FIRE Dataset [Internet]. Available: <https://www.kaggle.com/datasets/phylake>

1337/fire-dataset

- [7] Wu, Shixiao, and Libing Zhang. “Using popular object detection methods for real time forest fire detection.” *2018 11th International symposium on computational intelligence and design (ISCID)*. Vol. 1. IEEE, 2018.