

# 화재 재난 상황 인식을 위한 객체 검출

김태성<sup>1</sup>, 방재연<sup>2</sup>, 서정운<sup>2</sup>, 손경아<sup>1,3</sup>

<sup>1</sup>아주대학교 인공지능학과

<sup>2</sup> 아주대학교 심리학과

<sup>3</sup> 아주대학교 소프트웨어학과

moggie617@ajou.ac.kr, qkdwodus777@ajou.ac.kr, tjwjddns1302@ajou.ac.kr, kasohn@ajou.ac.kr

## Object detection for Fire Disaster Situation Recognition

Tae-Seong Kim<sup>1</sup>, Jae-Yeon Bang<sup>2</sup>, Jeong-un Seo<sup>2</sup>, Kyung-Ah Sohn<sup>1,3</sup>

<sup>1</sup>Dept. of Artificial Intelligence, Ajou University

<sup>2</sup>Dept. of Psychology, Ajou University

<sup>3</sup>Dept. of Software, Ajou University

### 요 약

화재 상황에서의 빠른 현장 파악은 인명피해를 줄이는데 중요한 요소이다. 기존 연구의 화재와 관련된 데이터셋들은 대부분 불과 연기를 라벨링하여 화재의 예방에 초점을 두고 있다. 본 연구에서는 화재 상황에서 사람과 소방관, 연기, 불을 탐지하는 Object detection 모델을 만들어 현장 파악에 더욱 도움을 주고자 하였다. 이를 위해 화재 상황 이미지 약 3000장을 수집하고 라벨링하여 데이터셋을 구성하였으며 이를 이용해 객체 검출 모델인 RetinaNet을 학습하였다. 또한, 화재 상황에서 Object Detection 모델의 성능을 향상시키기 위해 기존 모델인 RetinaNet에 Dehazing(FFA-Net), Smoke augmentation, semi-supervised(ISD) 방법을 적용하였고, semi-supervised 조건에서 mAP 63.7로 가장 높은 성능을 도출하였다.

### 1. 서론

2011년부터 2020년 사이, 매년 평균 42,332건의 화재가 발생하였으며, 화재로 인해 2216명의 인명피해와, 309명의 사망자가 발생했다[1]. 화재 재난 상황을 빠르게 파악하여 인명구조에 도움을 주기 위해, 대상의 위치와 속성을 탐지하는 object detection을 활용할 수 있다. 화재와 관련해 object-detection을 수행하는 여러 연구들이 진행되어 왔지만[2][3] 몇몇 한계점을 보인다. 이전의 연구는 object의 class를 불(fire)과 연기(smoke)로 한정하기 때문에 화재 상황 대응에 불충분하다. 이에, 본 연구는 화재 이미지 데이터를 수집하고, 객체 탐지의 대상을 불, 연기, 사람, 소방관으로 설정하여 연구를 진행하였다. 또한 많은 화재 상황은 연기로 인해 object detection의 성능이 저하될 수 있다. 본 연구는 이에 대처하기 위해, smoke augmentation과 dehazing방법을 사용하였다. 추가적으로, 라벨링된 데이터를 필요로 하는 supervised 학습을 사용하기 위해서는, 라벨링에 소요되는 시간이 많다. 라벨링이 되지 않은 데이터도 최대한 활용하기 위해 semi-supervised 방법을 사용하였다.

본 연구는 이러한 기법들을 모델에 적용, 비교하여, 효과적인 방법을 탐색하였다.

### 2. 데이터셋

#### 2.1. 화재 상황 데이터셋

본 연구는 FireNET[4]과 Domestic-Fire-Smoke-Dataset[5]의 데이터와, 여러 화재 상황의 데이터 수집을 위해 비디오 플랫폼 YouTube를 활용하였다. python의 pytube, cv2, selenium 라이브러리를 활용하여 화재 동영상을 크롤링하고, 프레임을 잘라 이미지를 수집하였다. 이후 데이터셋의 품질을 위해 라벨링과 이미지를 검토하는 과정을 거쳐 총 3007장의 이미지를 최종 labeled 데이터셋으로 선정하였고, 이중 2675장을 trainset, 332장을 testset으로 분할하였다.

Semi-supervised 기법을 위한 unlabeled 데이터셋의 경우 labeled 데이터셋에 쓰이지 않은 YouTube 화재 동영상을 사용하였다. 추가로 smoke augmentation을 위해 smoke 키워드로 구글 이미지[6]에서 연기 이미지를 크롤링하였다.

#### 2.2. 라벨링

본 연구를 진행하기 위해 labellmg[7] 프로그램을 활용하여 3007장의 화재 상황 이미지를 fire, smoke, person, firefighter의 4개의 class에 대해 라벨링 하였다. unlabeled 데이터셋과 labeled 데이터셋 각각의 개수는 <표1>, labeled 데이터셋의 trainset과 testset의 이미지 개수는 <표2>, 각 class별 개수는 <표3>과 같다.

<표 1> labeled, unlabeled 데이터셋 이미지 개수

Labeled	Unlabeled	Total
3007	1154	4161

<표 2> labeled 데이터셋 이미지 개수

Trainset	Testset	Total Labeled
2675	332	3007

<표 3> 화재 상황 데이터셋 class 개수

	Trainset	Testset	Total
Fire	2986	328	3314
Smoke	2675	375	3050
Person	926	84	1010
Firefighter	2437	419	2856

### 3. 실험 조건

본 연구에서 사용한 object detection 모델은 RetinaNet[8]이다. RetinaNet은 one-stage detector로 속도가 빠르고 여러 스케일의 object를 효율적으로 탐지하기 위한 FPN(Feature Pyramid Network)과 클래스 불균형 문제를 다루기 위한 Focal loss를 적용한 것이 주요 특징이다. 위 모델을 사용해서 세 가지 다른 조건에서 본 연구의 데이터셋으로 성능을 비교했다. <그림1은> 각 조건에서 학습되는 이미지의 예시이다.



<그림 1> 조건 별 실험 이미지 예시.

#### 3.1. Image dehazing

Image dehazing은 안개, 연기, 먼지 등이 있는 이미지에서, 선명한 이미지로 복원하는 기술이다. 화재 상황에서, object detection에 연기가 미치는 영향을 줄이기 위해 전처리 방법으로 사용하였다.

본 연구는 Image dehazing 모델 중 FFA-Net[9]을 사용하였다. FFA-Net은 CNN기반의 dehazing network로, 픽셀과, 채널마다 다른 가중치를 주는 모듈을

추가해서 중요한 정보를 가진 픽셀과, 채널에 집중하여 성능을 개선하였다.

#### 3.2. smoke augmentation

본 연구는 선행논문인 NSA[10], FPI[11]에서 사용했던 기법을 object detection에 맞게 수정하여, 원본 이미지에 연기를 더해 모델을 학습시키고자 하였다.

과정은 아래와 같다.

(1) 목표 이미지의 라벨링을 확인, 연기가 없는 것을 확인한다.

(2) 무작위로 선정된 연기 이미지로부터 일정 크기의 패치를 잘라낸다.

(3) 패치의 크기를 변환한 후, 패치를 목표 이미지의 임의의 위치에 더한다.

(4) 패치가 더해진 이미지 대하여 Poisson blending[12]을 실시한다.

(5) 추가된 연기 patch data를 라벨에 추가한다.

이 조건에서, 크롤링을 통해 얻어진 121 개의 연기 이미지를 smoke augmentation에 활용하였다.

#### 3.3. Semi-supervised Learning (ISD)

semi-supervised learning은 labeled 데이터와 unlabeled 데이터 모두 사용해서, 모델을 학습시키는 방법이다.

본 연구는 그 중, 연기의 시야 차폐가 mix-up[13] 기법을 적용한 것과 유사하여, mix-up을 활용한 semi-supervised loss인 ISD[14] (Interpolation-based Semi-supervised learning for Object Detection) loss를 사용하였다.

mix-up은 이미지 둘에 합이 1이되는 서로 다른 가중치를 곱한 뒤, 더하는 방식이다. 예시는 <그림 1>, (d)에서 확인할 수 있다.

ISD loss는 mix-up한 이미지와 원본 이미지 둘에 대한 모델의 예측을 비교하여 loss를 얻는 방식으로, object와 object, object와 background인 경우를 나누어서 loss함수를 구한다.

ISD loss의 성능을 평가하기 위해서, unlabeled 데이터 1154개를 추가한 조건과(w), 추가하지 않은 조건(w/o)을 따로 비교하였다.

### 4. 실험

#### 4.1. 실험 설정

각 조건을 RetinaNet에 적용하여 실험을 실시하였다. epoch은 100으로 설정하였고, ResNet-50 FPN 네트워크를 backbone으로 사용하였다. 평가 지표로는 AP(Average Precision)를 사용하였으며, IOU(Intersection over Union) threshold는 50%로 설정하였다.

AP는 recall 값의 변화에 따른 precision의 값을 나타낸 곡선에서, precision의 값의 평균을 구한 값이다.

## 4.2. 실험 결과

&lt;표 4&gt; Smoke Augmentation, Dehazing 조건 실험 결과

	RetinaNet 원본	Smoke Augmentation	Dehazing (FFA)
Fire	53.7	<b>56.2</b>	51.7
Smoke	70.4	<b>72.6</b>	65.5
Person	<b>52.5</b>	52.2	38.7
Firefighter	70.9	<b>71.9</b>	67.2
mAP	61.8	<b>63.2</b>	55.7

원본 조건과 비교한 smoke augmentation과 FFA-Net을 활용한 Dehazing 조건의 클래스별 AP와 평균AP(mAP)를 <표 4>에 정리하였다. 평균 AP로 볼 때, smoke augmentation을 적용한 조건이 가장 높았고, person의 class AP로 봤을 때는 원본 RetinaNet이 가장 높았다.

&lt;표 5&gt; semi-supervised 조건 실험 결과

	RetinaNet 원본	ISD(w/o)	ISD(w)
Fire	53.7	55.5	<b>59</b>
Smoke	70.4	72.5	<b>75.2</b>
Person	52.5	44.0	47.6
Firefighter	70.9	70.6	<b>73.1</b>
mAP	61.8	60.6	<b>63.7</b>

다음은 원본 조건과 비교해서, semi-supervised loss인 ISD loss를 적용한 조건을 <표 5>에 정리하였다. w/o 조건은 unlabeled data를 추가하지 않은 조건이고, w/ 조건은 unlabeled data 1154장을 추가한 조건이다. ISD loss를 추가했을 때, 원본과 비교하여 person의 class AP가 많이 떨어졌지만, fire와 smoke의 class AP가 높아졌다. ISD의 w/ 조건과 w/o 조건을 비교하였을 때, w/ 조건에서 모든 class의 AP가 높아진 것을 볼 수 있었다.

종합적으로, 가장 mAP가 높은 조건은 unlabeled 데이터를 추가한 ISD조건이며, person class의 AP가 크게 차이 나지 않고 mAP가 높아진 조건은 Smoke augmentation 조건이다.

## 5. 결론

본 연구는 화재 상황에서의 object detection 적용을 위해, 불(fire), 연기(smoke), 사람(person), 소방관(firefighter)으로 라벨링한 화재 상황 데이터셋을 구축하였다. 또한 Smoke augmentation, dehazing (FFA), Semi-supervised Learning (ISD)을 RetinaNet에 적용, 비교 분석하였다. 그 중 Semi-supervised Learning

(ISD)조건의 mAP가 가장 높았지만 person class의 정확도가 떨어지는 단점을 보였다. 추후 연구를 통해 제안된 모델들의 person class AP를 향상시키고, 기존 화재 예방을 목적으로 한 모델들과 통합할 예정이다.

## 사사문구

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2022-2018-0-01431)

## 참고문헌

- [1] e 나라지표, [https://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx\\_cd=1632](https://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=1632)
- [2] 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). Hangzhou, China. 2018. Vol. 1, pp. 280-284.
- [3] Li, Pu, and Wangda Zhao. "Image fire detection algorithms based on convolutional neural networks." Case Studies in Thermal Engineering. Vol.19. Article. 100625. 2020.
- [4] OlafenwaMosesRoger/FireNET, 2019. <https://github.com/OlafenwaMoses/FireNET>
- [5] datacluster-labs/Domestic-Fire-and-Smoke-Dataset, 2021. <https://github.com/datacluster-labs/Domestic-Fire-and-Smoke-Dataset>
- [6] Google Images, <https://www.google.co.kr/imgph?hl=ko>
- [7] heartexlabs/labelImg, 2022. <https://github.com/heartexlabs/labelImg>
- [8] Lin, Tsung-Yi, et al. "Focal loss for dense object detection." Proceedings of the IEEE international conference on computer vision. Venice, Italy. 2017. pp. 2980-2988.
- [9] Qin, Xu, et al. "FFA-Net: Feature fusion attention network for single image dehazing." Proceedings of the AAAI Conference on Artificial Intelligence. New York, USA. 2020. Vol. 34, No. 07, pp. 11908-11915.
- [10] Schlüter, Hannah M., et al. "Self-supervised out-of-distribution detection and localization with natural synthetic anomalies (nsa)." arXiv preprint arXiv:2109.15222. 2021.
- [11] Castro, Eduardo, Jaime S. Cardoso, and Jose Costa Pereira. "Elastic deformations for data augmentation in breast cancer mass detection." 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). Nevada, USA. 2018. pp. 230-234.
- [12] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. ACM Transactions on graphics (TOG). Vol.22. pp. 313-318. 2003.
- [13] Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. mixup: Beyond empirical risk minimization. arXiv preprint arXiv:1710.09412. 2017.
- [14] Jeong, Jisoo, et al. "Interpolation-based semi-supervised learning for object detection." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Tennessee, USA. 2021. pp. 11602-11611.