

# 자동 라벨링 기술을 활용한 흡연자 객체 감지 모델에 관한 연구

정현수<sup>1</sup>, 길준민<sup>2</sup>

<sup>1</sup>대구가톨릭대학교, <sup>2</sup>제주대학교

<sup>1</sup>jhs3215563@gamil.com, <sup>2</sup>jmgil@jejunu.ac.kr

## Study on Smoker Object Detection Model using Automatic Labeling Technology

Hyunsu Jeung<sup>1</sup>, Joon-Min Gil<sup>2</sup>

<sup>1</sup>Deagu Catholic University, <sup>2</sup>Jeju National University.

### 요약

본 논문은 흡연자와 비흡연자 간 구별을 목표로, 기존 수작업 라벨링 과정의 시간 소모와 비용 문제를 극복하기 위해 자동 라벨링 기술을 도입했다. autodistill을 사용한 자동 라벨링은 기존 방식에 비해 신뢰성은 높지 않지만, 추가 검수를 통해 신뢰성을 높였다. 또한, YOLOv8n 모델을 활용하여 객체 탐지의 흡연 행위를 식별하는 데 있어서 높은 정확도를 보여주었다.

### I. 서론

현대 사회에서 흡연은 여전히 중요한 공중보건 문제로 남아 있으며, 흡연자와 비흡연자를 구분하는 것은 흡연 관련 정책(예: 금연 구역)을 효과적으로 수행하는 데 있어 중요하다. 이를 위한 컴퓨터 비전 기술의 적용을 해보고자 한다. 이를 위해, 효율적인 데이터 가공 방식을 도입하여 객체 탐지 시스템의 개선을 목표로 삼는다.

본 논문에서는 흡연하는 사람과 흡연하지 않는 사람을 구별하는 객체 탐지를 위해 효율적인 데이터 가공 방식을 도입한다. 컴퓨터 비전 분야에서 객체 탐지 시스템은 원천 데이터에 대한 라벨링을 사람의 수작업으로 진행해왔다. 이 방식은 높은 신뢰도를 제공하지만, 많은 시간과 비용을 요구하며 지속적인 집중력을 필요로 한다. 이에 대한 대안으로, 본 연구에서는 자동 라벨링 기술인 autodistill[1]을 사용하여 데이터 가공 과정을 개선한다. 그리고 자동 라벨링한 학습 결과와 이후 사람이 검수한 후의 데이터를 이용해 객체 감지 모델 중 YOLOv5를 기반으로 최적화되어 있고 실시간 객체 감지에서 우수한 성능을 제공하는 YOLOv8n[2]을 사용하여 학습 결과 간 성능 차이를 분석한다.

### II. 본론

#### 1. 데이터 세트 준비

본 논문에서 사용된 데이터세트는 과학기술정보통신부의 재원으로 한국지능정보사회진흥원의 지원으로 구축된 “공원 주요시설 및 불법행위 감시 CCTV 영상 데이터”[3]이며, 이 데이터세트는 공원에서 다양한 활동을 포착한 3fps의 영상으로 구성되어 있다. 본 논문에서는 분석을 위해 10fps마다 한 장씩의 이미지를 추출하여 저장하였다.

#### 2. 자동 라벨링

자동 라벨링 과정은 pytorch 2.1.2+cu118, NVIDIA 3060 ti 환경에서 진행하였다. 여기서는 autodistill[1]에 사용되는 Detection and Segmentation 모델인 Grounded-SAM[4] 사용하여 자동 라벨링 과정을 수행하였다. 이 모델은 객체 탐지(object detect)와 분할(segmentation) 기능을 통합하여 이미지 내 객체를 정확하게 감지하고 분류한다. 이 모델은 일반적인 객체

감지 사례에 대한 데이터는 가공할 수 있지만, 흡연과 같은 특수한 행위를 포착하는 데에는 아직 한계가 있다. 본 연구의 주된 라벨링 대상은 ‘사람’이므로 필요한 라벨링 대상을 사람으로 설정하여 진행하였으며, 이 과정에서 각 이미지 처리에는 평균 2.5초가 소요되었다.

#### 3. 자동 라벨링 결과 및 검수

자동 라벨링의 이상적인 결과는 그림(1)과 같다. 그러나, 일부 데이터에는 그림(2)와 같이 다른 물체가 라벨링되거나 라벨링이 불완전한 경우도 발견되었다. 이러한 예외 사항들은 추가적인 검수를 통해 해결하였다.



(그림1) 이상적인 자동 라벨링 결과



(그림2) 잘못된 자동 라벨링 결과

라벨링이 완료된 후, 라벨이 없는 데이터셋을 제거하여 총 2,732장의 학습 이미지와 라벨링 데이터셋을 확보하였다. 추가로 라벨링된 데이터셋을 라벨링 툴을 사용하여 사람이 검수한 이미지 2,406장의 데이터셋을 별도로 구성하여, 두 데이터셋의 성능을 비교 분석하기 위한 준비를 하였다.

#### 4. 학습 및 성능 평가

학습 세트와 검증 세트는 각각 8:2로 나누어 구글 코랩 pytorch-2.1.0+cu121, Tesla T4 환경에서 학습을 진행하였다. 에포크(epoch)는 100, 배치 크기는 32, 이미지 크기는 640px로 각각 설정하여 학습을 진행하였다.

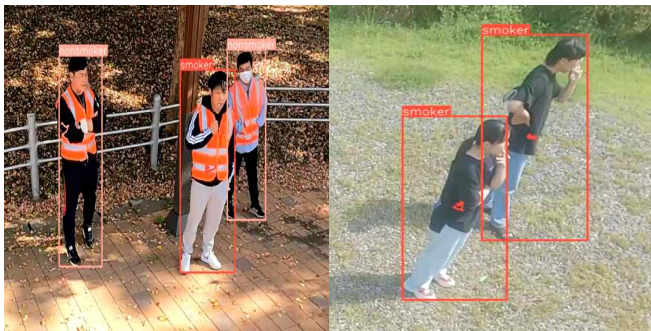
학습이 완료된 후 모델의 성능을 평가하기 위해 mAP50을 측정하였으며 이는 객체 탐지 모델의 정확도를 평가하는 한 지표로, <표 1>은 학습이 완료된 모델의 mAP50의 결과를 보여준다.

<표 1> YOLOv8n mAP50

	mAP50	smoke	non-smoke
auto-label	0.78	0.944	0.62
auto-label+inspects	0.922	0.954	0.89

#### 5. 테스트 및 결과 분석

이 모델 성능을 평가하기 위해 학습에 사용하지 않은 테스트 영상을 사용했다. 이 테스트 영상은 실제 공원 환경에서 촬영된 영상으로 흡연자와 비흡연자의 행동 패턴이 포함된 영상이다. (그림 3)에서 보이는 바와 같이, 테스트 영상이 효과적으로 작동한다는 검증 기회를 제공하였다.



(그림3) 테스트 결과 시각화

### III. 결론

본 논문에서는 흡연자와 비흡연자를 구별하는 객체 탐지를 위한 효율적인 데이터 가공 방식 적용과 그 성능에 관해 탐구하였다. auto distill[1]을 활용해 데이터를 일반적인 객체 감지 사례에 대한 데이터는 가공할 수 있지만 흡연과 같은 특별한 사례에 대한 데이터는 가공이 아직은 불가능한 것으로 나타났다. 이는 autodistill 기술이 모든 상황에서 완벽하게 적용되기에는 한계가 있다. 그러나 본 연구는 자동 라벨링 기술과 사람의 검수를 혼합하면 시간 절약과 데이터 가공의 효율성과 신뢰성을 향상시킬 수 있다는 점을 입증하였다.

본 연구의 결과는 향후 객체 탐지와 분할 모델을 더욱 세밀하게 조정(fine-tuning)하여, 흡연과 같은 특별한 사례에 대한 데이터 가공에 적합한 모델을 학습시켜 적용할 예정이다.

### ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2022R1A2C1092934).

### 참고 문헌

- [1] Autodistill. (2024). Autodistill [Software].GitHub. <https://github.com/autodistill/autodistill>. Accessed January 10, 2024.
- [2] 이태희, 박천수 “YOLOv8을 이용한 실시간 화재 검출 방법”, 반도체 디스플레이 기술 학회지, 제22권 제2호, pp. 70-80, 2023년.
- [3] 공원 주요시설 및 불법행위 감시 CCTV 영상 데이터 (<https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=477>)
- [4] IDEA-Research. (2024). Grounded Segment Anything [Software]. GitHub. <https://github.com/IDEA-Research/Grounded-Segment-Anything>. Accessed January 10, 2024.