

CCTV 영상 기반 신경망을 이용한 유기물 탐지

박민성*, 손창대, 최성진, 이현지

*(주)마크애니

*parkms@markany.com, cdson@markany.com, choisj@markany.com, hjlee@markany.com

Abandoned Object Detection using Neural Network in Video-Surveillance

Park Min Seong*, Son Chang Dae, Choi Sung Jin, Lee Hyun Ji

*MarkAny.

요 약

본 논문은 합성곱 신경망을 이용하여 CCTV 영상 환경에서의 유기물을 탐지하는 알고리즘을 제안한다. 이 알고리즘은 입력 영상에서 합성곱 신경망을 통해 물체를 탐지하고, 탐지된 물체의 정보를 이용하여 유기물을 판별하는 방법으로 설계된다. 물체 탐지를 위한 합성곱 신경망은 AI Hub 데이터셋을 통해 학습되었다. 제안하는 알고리즘은 AVSS 데이터셋을 통해 기존 유기 알고리즘에 비해 높은 성능을 보인다.

I. 서 론

유기물이란 일정 시간 동안 움직임이 없는 물체 혹은 물체의 소유자와의 거리가 가까이 있지 않은 물체로 정의된다[1]. 유기물 탐지는 매장에서 고객이 물건을 두고 갔을 때 이를 발견하여 되찾아주거나 쓰레기 무단투기 금지 장소에서 무단 투기자를 찾아내는 등 일상생활에서 유용하게 사용될 수 있다. 뿐만 아니라 테러가 빈번하게 일어나는 도시에서 유기물은 폭탄, 화학물 등 위험 물질의 후보가 될 수 있기 때문에 유기물 탐지는 매우 중요하다. 과거에는 관제 요원이 수많은 CCTV 영상을 직접 감시했지만, 최근 딥러닝의 발전으로 딥러닝을 이용한 유기물 관제 시스템이 유기물을 탐지하는데 도움을 줄 수 있다. 따라서 본 논문은 CNN을 이용하여 CCTV 영상 내 유기물을 탐지하는 알고리즘을 제안한다.

II. 본 론

유기물을 탐지하기 위해 먼저 사람과 유기물 후보가 될 수 있는 물체를 탐지해야 한다. 이를 위해 본 논문은 합성곱 신경망을 활용한 방법을 제안한다. 합성곱 신경망을 학습하기 위한 데이터는 한국지능정보사회진흥원이 운영하는 AI Hub[2]에서 제공하는 유기 관련 영상을 사용한다.

합성곱 신경망 학습을 위해 AI Hub 유기 데이터셋[3]에서 초당 2초의 이미지를 추출하여 학습데이터로 사용한다. 본 논문에서 탐지하고자 하는 유기물 후보는 캐리어와 가방이므로, 각 이미지는 사람, 캐리어, 가방 클래스와 바운딩 박스로 레이블링 된다. 추가적으로 물체와 사람 사이의 관계에 대한 정보를 얻기 위해 물체를 들고 있는 사람 클래스를 추가한다. 여러 이미지 내 지정된 물체가 중복되는 경우는 1개의 이미지를 제외하고 마스킹 된다. 라벨링 된 이미지와 구축한 학습 데이터셋 정보는 각각 그림 1과 표 1과 같다.

본 논문에서는 실시간 CCTV 환경에서 제안하는 알고리즘을 적용하기 위해 YOLO[4]를 객체 탐지 모델로 사용한다. YOLO[4]의 입력은 1920x1080 해상도의 CCTV 이미지를 608x608로 다운 샘플링된 이미지이

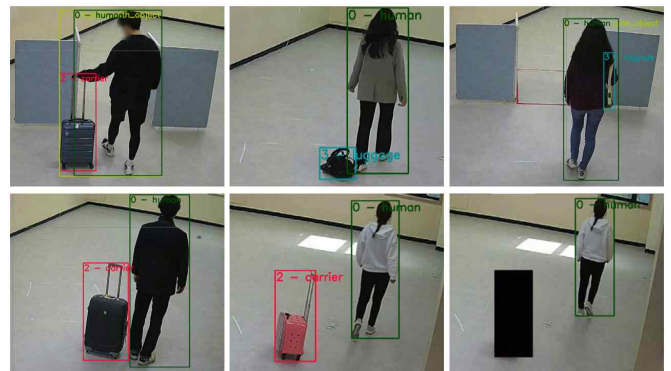


그림 1. 라벨링 된 데이터

클래스			총 이미지 수
번호	종류	개수	
0	사람	6,011	6,113
1	물체를 들고 있는 사람	2,307	
2	캐리어	1,897	
3	가방	1,566	

표 1. 학습 데이터셋

고, 출력은 객체에 해당하는 클래스와 바운딩 박스의 좌표 (x,y,w,h)이다. 학습을 위한 하이퍼파라미터로 학습률(learning rate)은 0.0003, 배치 사이즈는 2, 이터레이션(iteration)은 25,000을 사용한다. 스케줄러 기법은 StepLR을 사용하고, 10,000, 13,000, 15,000, 18,000, 21,000 이터레이션마다 0.1의 gamma 값을 설정한다. 학습 데이터 셋의 클래스 불균형을 고려하기 위해 각 클래스에서 얻어지는 loss마다 0.094, 0.245, 0.300, 0.361의 비율로 곱해서 최종적인 loss를 계산한다.

YOLO[4] 성능을 확인하기 위해 545장의 이미지를 평가 데이터셋으로 이용된다. 그림 2는 YOLO의 학습/평가 loss의 그래프이다. YOLO의 성능은 평가 데이터셋에 대하여 각 클래스마다 AP(Average Precision) 값

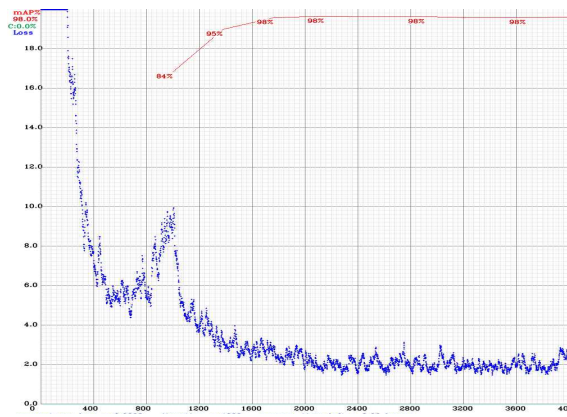


그림 2. 학습/평가 loss 그래프

이 99%, 99%, 100.00%, 98%를 보이고, 99%의 mAP(Mean Average Precision) 값을 얻었다.

본 논문은 클래스 0, 2, 3을 이용한 알고리즘 I과 클래스 0, 1, 2, 3을 이용한 알고리즘 II를 제안한다. 알고리즘 I은 입력 이미지에 대해 객체 탐지 모델이 객체를 찾았다면 그 객체가 유기 후보물인지 판별하고, 유기 후보물인 경우 이전 이미지에서의 유기물 후보와 현재 이미지에서의 유기물 후보의 IOU를 계산하여 물체가 멈춰 있는지 판별한다. 또한 사람과 유기 후보물 사이의 거리를 계산하여 주인이 있는 물체인지 아닌지 판별한다. 거리를 계산하기 위해 사람의 중심 좌표 또는 두 발의 중점 좌표와 물체의 중심 좌표가 사용된다. 최종적으로 알고리즘 I은 물체가 일정 프레임 이상 연속적으로 멈춰있고, 주인이 없는 경우일 때 유기로 판별한다.

알고리즘 II은 입력 이미지에 대해 객체 탐지 모델이 객체를 찾고, 찾은 객체와 이전 프레임의 유기 상태에 따라서 나누어서 동작한다. 알고리즘 II에서 물체가 멈춰있는지 판별하는 방법은 알고리즘 I과 같다. 주인이 있는 물체인지 아닌지 판별하기 위해서는 유기물 후보의 바운딩 박스가 물체를 들고 있는 사람의 바운딩 박스 내에 있는지를 확인한다. 이때 물체를 들고 있는 사람의 옆모습과 같이 유기물 후보가 가리워짐(occlusion)을 고려하기 위해 가리워짐 발생 시간을 계산하여 일정 시간 이상인 경우 유기로 판별한다.

제안하는 알고리즘의 성능은 AVSS[5] 데이터셋의 유기 영상을 이용하여 평가된다. 알고리즘 II가 [6]에서 제안하는 알고리즘에 비해 높은 성능을 보인다. 제안하는 알고리즘 I, II의 자세한 성능은 표 2와 같다.

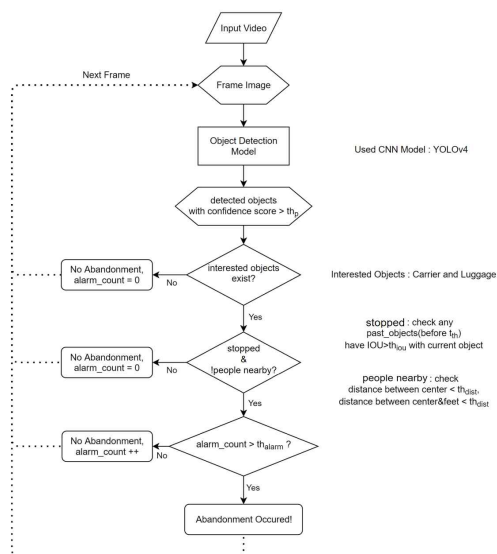


그림 3. 알고리즘 I

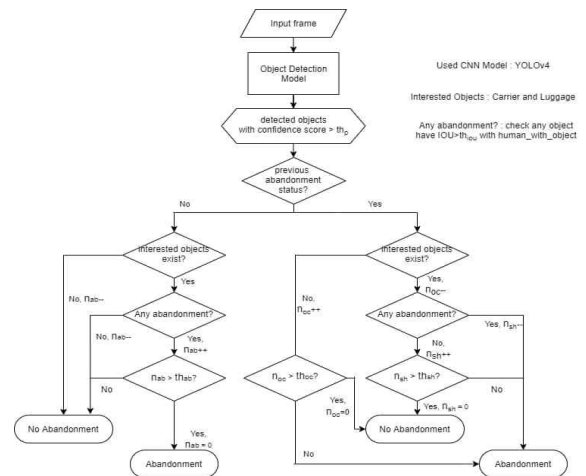


그림 4. 알고리즘 II

	precision	recall	f1 score
SOD+CCNN [6]	97.77%	51.89%	67.80%
SOD+CCNN [6] (Generated Sample)	97.48%	66.59%	79.13%
알고리즘 I (ours)	86.64%	80.66%	83.50%
알고리즘 II(ours)	98.49%	85.23%	91.38%

표 2. AVSS[5] 데이터셋에 대한 성능 비교

III. 결 론

본 논문에서는 CCTV 영상 기반 합성곱 신경망을 이용한 유기물 탐지 알고리즘을 제안한다. 제안한 알고리즘은 실시간 CCTV 환경의 유기물 발생 지역에서 사용될 수 있다.

ACKNOWLEDGMENT

본 연구는 2021년도 과학기술정보통신부(IITP)의 재원으로 지원을 받아 수행되었다. (No. 2021-0-00348, 2D/3D 영상 통합 분석을 이용한 클라우드 기반 무인점포 환경 대응형 영상보안시스템 개발)

참 고 문 헌

- [1] Elena L., Juan C. S., Diego. O., and Jose M. M. "Abandoned Object Detection in Video-Surveillance: Survey and Comparison", MDPI, 2018.
- [2] 한국지능정보사회진흥원, AI Hub (<http://aihub.or.kr/>).
- [3] AI Hub 이상행동 CCTV 영상. (<http://aihub.or.kr/aidata/139>).
- [4] Bochkovskiy. A., Wang, C. Y., and Liao H. Y. M., "YOLOv4: Optimal speed and accuracy of object detection." [Online]. Available: <http://arxiv.org/abs/2004.10934>.
- [5] Bayona A., SanMiguel J. C., and Martínez J. M., "Comparative evaluation of stationary foreground object detection algorithms based on background subtraction techniques". In Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 25 - 30, September 2009.
- [6] Sorina S., and Radu T. I., "Real-Time Deep Learning Method for Abandoned Luggage Detection in Video". [Online] available: <http://arxiv.org/abs/1803.01160>.