

드론과 딥러닝을 활용한 조난자 탐지 모델*

배현재⁰¹ · 장규진² · 김영민³ · 김영남² · 김진평²¹성균관대학교 소프트웨어학과²차세대융합기술연구원 컴퓨터비전 및 인공지능 연구실³인천대학교 경제학과

jason0425@snu.ac.kr, gjjang@snu.ac.kr, winston1214@naver.com, hwarangmoon@snu.ac.kr, jpkim@snu.ac.kr

Survivor Detection Model Using Drones and Deep Learning

Hyun-Jae Bae · Gyu-Jin Jang · Young-Min Kim · Young-Nam Kim · Jin-Pyung Kim

¹Department of Computer Science and Engineering, Sungkyunkwan University²Computer Vision and Artificial Intelligence laboratory, Advanced Institute of Convergence Technology³Department of Economics, Incheon National University

요 약

최근 드론과 딥러닝을 융합하여 다양한 연구를 진행한다. 사람의 시각에 의존하여 탐색하는 기존 방법은 소모적이고 반복적인 방법이다. 기존 연구는 객체 탐지만을 사용함으로써 산지 등 자연환경에서 조난자가 어떤 자세를 취하고 있는지 인식할 수 없는 문제점이 있다. 본 논문에서는 이러한 문제점들을 해결하기 위해 정확도를 개선하고, 조난자 탐지만 아니라 자세추정방법을 제안하고자 한다. 조난자를 촬영한 여러 데이터셋을 활용하여 모델을 학습하고 자세추정방법의 성능을 비교 및 평가한다. 본 논문에서는 고도별 바운딩 박스 탐지율 결과와 고도별 자세추정 정확도 결과를 확인하였다. 고도가 높아질수록 조난자 탐지성능은 0.94에서 0.38로 줄어드는 것을 확인하였고, 고도별 조난자의 자세추정 정확도성능은 0.82에서 0.33으로 줄어드는 것을 확인하였다.

1. 서 론

드론(Drone)을 활용한 산업 분야와 인구 증가로 인한 다양한 드론 활용 사례가 증가하고 있다. 2026년 드론의 시장규모는 90조에 육박할 것으로 기대되며, 2013년부터 드론 시장은 드론 기체 신고, 사용사업체, 조종자격자 등의 주요지표가 최근 3년간(2016년~2018년) 46~244%로 증가하였다. 4차 산업혁명 중 드론 시장은 가파르게 성장 중이며 2021년 현재까지도 지속해서 성장 중이다[1]. 최근 드론과 AI(Artificial Intelligence)를 융합하여 다양한 연구를 진행한다. 최근 강도 높은 이상기후 현상으로 인해 조난자 발생이 문제 되고 있다. 이와 같은 이유로 발생하는 조난자를 수색할 때 사람의 시각에 의존하여 탐색하는 방법은 소모적이고 반복적인 방식이다. 또한, 구조 요원이 육안의 피로로 인해 모든 영상을 자세히 확인하기 어렵다. 기존 연구는 객체 탐지(Object Detection)만 사용함으로써 산지 등 자연환경에서 조난자가 어떤 자세를 취하고 있는지 확인할 수 없는 문제점이 있다. 본 논문에서는 이러한 문제점들을 해결하기 위해 정확도를 개선하여, 객체 탐지만 아니라 자세추정방법(Pose Estimation)을 제안하고자 한다[2]. 본론에서는 고도별로 촬영한 조난자 데이터를 수집하여 객체탐지 모델을 학습하고 자세추정 방법의 성능을 비교 및 평가한다. 조난자 수색 시스템의 구조와 여러 지형에서 활용될 데이터 세트와 실험을 분석하고 평가하였으며, 마지막으로 결론 및 향후 방향을 제시하였다.

2. 조난자 수색 모델

본 연구에서는 YOLO[3]와 HRNet_w32 모델을 활용하여 여러 지형에서 조난자를 보다 정확하고 신속하게 수색하기 위한 목적으로 연구를 수행하였다. 기존 객체 검출의 연구들은 사람의

눈높이에 맞춰 저/중 고도에 있는 객체에 집중되어 연구되었다. 또한, 드론의 시각에서 보이는 높은 고도에 관한 객체 검출 연구는 많은 진전이 되지 않았다. 본 논문에서는 중/고고도에서의 조난자 탐지를 시도하였으며, 탐지된 조난자의 자세추정을 진행하였다.

중/고 고도에서 작은 크기의 조난자를 탐지하고, 자세추정 모델인 HRNet-w32 모델을 학습시켰다. 제안한 조난자 수색 모델 구조는 다음 [그림 1]과 같다.

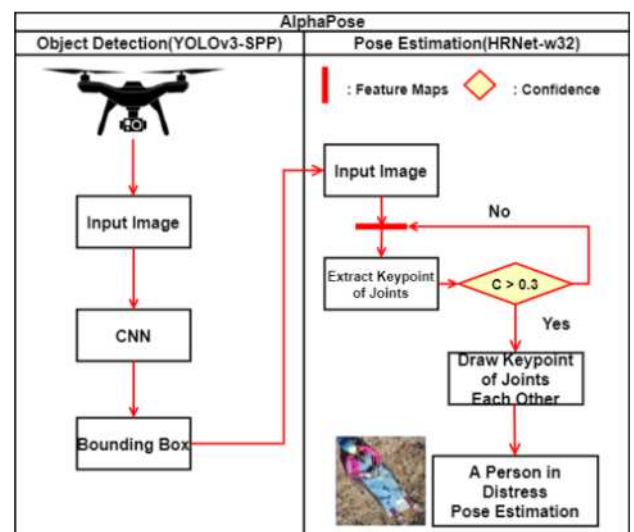


그림 1 조난자 수색 모델 시스템 구조

※ 이 논문은 행정안전부 극한재난대응기반기술개발사업의 지원을 받아 수행된 연구임(2020-MOIS31-014)

2.1 조난자 객체탐지

드론에서 촬영한 입력 이미지를 $S * S$ 그리드로 나누고, 각각의 그리드셀은 B개의 바운딩 박스와 각 바운딩 박스에 대한 컨피던스 스코어(Confidence Score)를 갖는다. 컨피던스 스코어 수식은 식 (1)과 같다.

$$Confidence Score = Pr(Object) * IoU_{pred}^{truth} \quad (1)$$

각각의 그리드셀은 C개의 클래스에 대한 조건부확률(Conditional Class Probability)을 갖는다. 조난자 클래스에 대한 조건부 확률 수식은 식 (2)이고, 조난자의 바운딩 박스를 그려주는 수식은 식 (3)과 같다. 각각의 바운딩 박스는 $x, y, w, h, confidence$ 로 구성된다. (x, y) 는 바운딩 박스의 중심점을 의미하며, (w, h) 는 입력 이미지의 width, height에 대한 상댓값이 입력된다.

$$Conditional Class Probability = Pr(Class_i | Object) \quad (2)$$

$$\begin{aligned} Score &= Conditional Class Probability * Confidence Score \quad (3) \\ &= Pr(Class_i | Object) * Pr(Object) * IoU_{pred}^{truth} \\ &= Pr(Class_i) * IoU_{pred}^{truth} \end{aligned}$$

2.2 조난자 자세추정

바운딩 박스 내에서 조난자의 관절 키포인트를 추출하는 모델이다. 조난자의 관절 키포인트 추출 후, Confidence 값이 0.3 이상이면 관절과 관절간 연결을 진행하였다. 조난자 수색 모델 구조 중 HRNet-w32 부분에서 관절 키포인트를 추출하는 수식은 식 (4)와 같다.

$$K_{Sim}(P_i, P_j | \sigma_1) = \begin{cases} \sum_n \tanh \frac{c_{i^n}}{\sigma_1} \cdot \tanh \frac{c_{j^n}}{\sigma_1}, & \text{if } k_j^n \text{ is within } \beta(k_j^n) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

K_j : Coordinate of Key Point
 P_i : Coordinate of Bounding Box
 P_j : Coordinate of Bounding Box

객체(조난자)의 모든 관절 키포인트에 대한 분포가 낮을수록 각각의 컨피던스(Confidence) 점수가 높을수록 좋다. tanh는 신뢰도가 낮은 값을 필터링하며 P_i 와 P_j 의 두 좌표가 일치할수록 tanh의 미분된 값은 1에 근접해진다.

3. 실험 결과 및 평가

3.1 실험환경





드론은 휴인스 무선형 무인항공기 드론 Blueye-820을 사용하였으며, 딥러닝 객체 탐지모델을 GPU에서 실행하기 위해 CUDA 10.1, cuDNN 8.0.4를 Ubuntu 18.04에 설치하였다. 딥러닝 프레임워크는 PyTorch 1.7.1을 활용하였다. CPU는 i7-9700, RAM 32GB, GPU는 GeForce RTX 2070 super 8GB를 사용하였다.

3.2 데이터셋

본 논문에서 활용한 조난자 데이터셋의 설계와 취득은 ‘한국정보화진흥원(NIA) 인공지능 학습용 데이터 구축 사업’ 수행 참여기관 중 한국모형항공협회에서 촬영하였다. 다양한 고도와 환경에서 조난자를 탐지하기 위해 조난자 데이터를 구축하였다.

데이터셋의 상황은 맑음과 흐림이고, 시간은 오전 06-12시, 오후 2시-5시이며, 각도는 비행 촬영으로 진행하였다. 비행 촬영 시 원천 데이터의 예시는 <표 1>과 같고, 조난자 데이터셋에 대한 설명은 <표 2>와 같다.

<표 1> 지형별 조난자 데이터셋 예시

지형	조난자 데이터	지형	조난자 데이터
평지(흙)		평지(수풀)	
해상도	3,840 x 2,160	해상도	3,840 x 2,160
산악		저수지	
해상도	3,840 x 2,160	해상도	3,840 x 2,160

<표 2> 조난자 데이터셋 설명

데이터셋	설명
대상	사람, 자연
장소	평지(흙), 평지(수풀), 산악, 저수지
드론 기체	DJI Matrice 300 RTK
해상도	3,840 x 2,160(4K)
데이터 구축량	4K 영상 : 100시간 이상 4K 이미지 : 30만장 이상
데이터 형식	4K 영상 : MP4 4K 이미지 : JPG

3.3 모델 학습

본 논문에서는 모델을 학습시키기 위해 COCO(Common Objects in Context) 2017 데이터셋을 활용하였다. 실험에서 활용된 AlphaPose는 YOLOv3-spp 검출기를 통해 조난자(사람 객체)를 탐지하고, 객체의 크기만큼의 바운딩박스를 생성한다. 생성된 바운딩박스 내에서 사람의 관절 키포인트를 추출하는 HRNet 모델 학습을 진행하였다. 실험에서 활용한 HRNet은 피쳐맵(Feature Map)의 크기가 작아지고 깊이가 깊어질수록 해상도 표현을 향상하기 위해 멀티 스케일 퓨전(Multi Scale Fusions)을 수행한다. 고도가 높은 곳에서도 작은 객체의 조난자를 탐지하고 자세추정을 하기 위해 적합한 모델이라고 판단되어 실험에 활용되었다.

조난자 수색 모델의 적합한 계산을 위해 아래 <표 3>와 같이 세부 파라미터를 설정하였다. 작은 객체의 조난자를 탐지하고 자세추정을 할 수 있게 세부 파라미터를 최적화하였다.

<표 3> 제안한 모델의 파라미터 값

파라미터 값	Values(functions)
Input Size	256 x 192
Loss Function	MSE(Mean Square Error)
Optimizer	SGD
Learning Rate	0.01
Scheduler	Cosine Annealing Warmup Restarts
Epochs	100
Transforms	FLIP : True Rotation Factor : 45 Scale Factor : 0.5 Half Body of Joints Numbers : 8 Probability Half Body : 0.2

HRNet 파라미터 중 손실함수를 MSE(Mean Square Error)로 설정한 경우의 특징은 실제 정답에 대한 정답률의 오차(Error) 뿐만 아니라 다른 오답들에 대한 정답률 오차 또한 포함하여 계산하는 것이다. 스케줄러(Scheduler)는 Cosine Annealing Warmup Restarts[10]를 사용하였다. 이유는 학습 과정 중 최댓값과 최솟값 사이에서 코사인(Cosine) 함수를 활용하여 급격히 증가 또는 감소시켜 모델의 매니폴드(Manifold) 공간의 안장(Saddle Point)을 빠르게 벗어날 수 있기 때문이다. 스케줄러의 수식은 식 (5)과 같다.

$$n_t = n_{\min}^t + \frac{1}{2}(\eta_{\max}^i - \eta_{\min}^i)(1 + \cos(\frac{T_{\text{current}}}{T_i}\pi)) \quad (5)$$

Transforms는 관절 추출의 정확도(Accuracy)를 높이기 위해 데이터 증강(Data Augmentation) 3가지 방법을 사용하였다. 첫 번째는 관절의 좌우를 바꾸는 FLIP 기법을 활용하였고, 두 번째는 데이터 회전(Rotation)을 45도씩 진행하였고 세 번째는 데이터의 양을 0.5배로 줄여주는 스케일팩터(Scale Factor)를 사용하였다.

3.4 실험 결과

모델의 학습결과는 mAPval은 73.46%이고 Lossval는 0.1548이다. 조난자 수색 성능을 측정하기 위해 5m 간격으로 40m까지 고도별로 촬영된 데이터로 테스트 성능을 먼저 검증하였다. 고도별 성능 테스트를 진행한 후 사람(연기자)이 서있기, 앉아있기, 누워있기, 자세를 반복한 데이터로 테스트를 진행하였다. AlphaPose[4]를 활용한 드론 자세추정 성능은 한 사람(연기자)당 100장씩 테스트를 진행하였다. 테스트 성능 지표는 임의의 100장 중 탐지된 바운딩 박스의 개수와 탐지된 자세추정의 횟수 비율로 진행하였다. 테스트한 100장 중 바운딩 박스의 개수와 바운딩 박스의 개수 중 자세추정 횟수는 몇 번인지 검토하였다. 바운딩 박스의 평균 탐지율은 식은 (6)이고, 자세추정의 평균 탐지율 식은 (7)이다.

$$\text{Average Rate}_A = \frac{\text{Total Number of Boundingbox}}{100} \quad (6)$$

$$\text{Average Rate}_B = \frac{\text{Total Number of PoseEstimation}}{\text{Total Number of Boundingbox}} \quad (7)$$

실험 결과로는 고도가 높아질수록 바운딩 박스의 평균 탐지율은 0.61에서 0.08로 줄어드는 것을 확인하였지만, 조난자 자세추정의 평균 탐지율은 고도와 관계없이 성능이 비슷함을 확인하였다.

〈표 6〉 고도별 바운딩 박스 탐지율 결과

	10m	15m	20m	25m	30m	35m	40m
1st	-	-	0.95	0.73	0.77	0.56	0.53
2nd	0.87	0.80	0.80	0.80	0.60	0.40	0.43
3rd	0.73	0.82	0.80	0.77	0.73	0.50	0.53
4st	0.73	0.73	0.70	0.78	0.60	0.46	-
5st	0.90	0.80	0.83	0.70	0.70	0.50	0.40
6st	0.90	0.80	0.80	0.73	0.60	0.43	0.50
7st	0.86	0.87	0.83	0.71	0.72	0.57	0.30
8st	0.90	0.80	0.76	0.66	0.65	0.48	0.40
9st	0.82	0.93	0.75	0.58	0.78	0.55	0.36
10st	0.81	0.99	0.66	0.58	0.50	0.43	0.48
평균[%]	0.94	0.84	0.77	0.70	0.65	0.48	0.38

〈표 7〉 고도별 자세추정 정확도 결과

	10m	15m	20m	25m	30m	35m	40m
1st	-	-	0.76	0.33	0.75	0.63	0.10
2nd	0.87	0.83	0.60	1.0	1.0	0.0	1.0
3rd	0.98	0.73	1.0	1.0	0.0	0.45	0.4
4st	0.87	0.83	1.0	1.0	0.0	1.0	-
5st	0.97	0.80	0.70	1.0	0.67	0.30	0.7
6st	0.80	1.0	0.60	0.4	1.0	1.0	0.45
7st	0.90	0.86	1.0	0.25	1.0	1.0	0.0
8st	0.84	0.80	0.35	0.54	0.60	0.25	0.20
9st	0.98	0.86	0.70	0.38	0.54	0.30	0.10
10st	0.97	0.75	0.80	0.45	0.48	0.33	0.30
평균[%]	0.82	0.75	0.75	0.64	0.60	0.53	0.33

〈표 8〉 고도별 조난자 자세추정 시각화 결과

고도	결과[빨간색/검은색]	결과[확대]
10m		
20m		
30m		
40m		

4. 결론 및 향후 방향

국내에서 처음으로 드론을 활용하여 자세추정을 진행한 연구이다. 본 논문에서는 조난자를 더욱 빠르고 정확하게 수색하기 위해 객체탐지(Object Detection)와 자세 추정(Pose Estimation)방법을 적용하였다. 고도별로 촬영된 데이터를 활용하여 각 고도별로 조난자의 탐지율과 자세추정율은 어느 정도인지 실험을 진행하였다. 조난자 학습 모델의 결과는 mAPval은 73.46%이고 Lossval는 0.1548이다. 본 논문에서의 실험 결과로는 고도가 높아질수록 조난자 탐지성능은 0.94에서 0.38로 줄어드는 것을 확인하였지만, 조난자 자세추정의 성능은 고도에 관계없이 성능이 비슷함을 확인하였다. 실험 결과 확인 시 조난자가 서있거나 앉아있는 경우보다 누워있는 경우가 관절 키포인트 추출성능이 더 우수함을 확인할 수 있다. 이를 보완하기 위해 향후 연구로는 GAN(Generative Adversarial Network)를 활용하여 보이지 않는 조난자의 관절 키포인트까지 추출을 하여 자세추정의 성능을 높일 계획이다.

참고 문헌

- [1] 국토부_드론산업 규모 5년내 20배 육성위한 종합계획 발표, <https://www.korea.kr/news/pressReleaseView.do?newsId=156244834>
- [2] Sun, Ke, et al. "Deep high-resolution representation learning for human pose estimation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2019.
- [3] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).
- [4] Fang Hao-Shu, et al. "Rmpe: Regional multi-person pose estimation." Proceedings of the IEEE International Conference on Computer Vision. 2017.