

YOLOv8-SLG 를 이용한 열화상을 통한 인간 키포인트 감지 개선

장우건, 조인휘*
한양대학교, *한양대학교

yuqian_zhang@foxmail.com, *iwjoe@hanyang.ac.kr

Improving Human Pose Recognition with YOLOv8-SLG Using Thermal Images

Zhang Yu Qian, Joe Inw Hee*
Hanyang Univ., *Hanyang Univ.

요 약

본 연구는 야간 또는 저조도 조건에서 열화상 인체 키포인트 인식을 위한 키포인트 검출 기술을 기반으로 한 새로운 방법(YOLOv8-SLG)을 제안한다. YOLOv8-SLG 는 YOLOv8 과 LSKA 어텐션 메커니즘을 결합하고 기존 모델의 컨볼루션을 대체하여 저조도 환경에서 기존 방법의 인식 정확도 부족 문제를 해결한다. 실험 결과 보면, 이러한 개선 사항이 복잡한 환경에서 인체 키포인트 인식 성능을 크게 향상시키는 것으로 나타났다.

I. 서 론

인체 키포인트 인식은 이미지나 비디오에서 인체 부위와 자세를 감지하고 인식하는 기술로, 보안 모니터링, 운동 분석, 인간-컴퓨터 상호작용 및 의료 보건 등 다양한 분야에서 널리 사용된다 [1]. 정확한 키포인트 인식은 모니터링의 품질과 사용자 경험을 크게 향상시킬 수 있다. 그러나 전통적인 방법은 저조도 환경에서 정확도가 낮고, RGB 이미지는 조명이 부족하거나 역광인 경우에 효과가 떨어지며, 프라이버시 문제도 존재한다. 따라서 가시광선에 의존하지 않는 열화상을 이용한 키포인트 인식이 특정 응용 분야에 더 적합하다. 하지만 열화상 이미지의 낮은 해상도와 환경 온도 변화는 키포인트 인식에 도전 과제를 안겨준다 [2]. 이러한 문제를 해결하기 위해, 본 연구에서는 YOLOv8-SLG 모델을 제안하고, YOLOv8 의 백본 네트워크, 넥, 그리고 디텍션 헤드를 개선하였다. 저조도와 열화상 조건에서 인체 키포인트 인식을 최적화하였다. 공유 GroupNorm 컨볼루션을 통해 모델의 다중 스케일 처리 능력을 강화하고, 목표 위치 추적 및 분류 성능을 향상시켰다. 본 연구는 인체 키포인트 인식 기술이 프라이버시 민감하거나 저조도 환경, 예를 들어 야간 모니터링 및 특정 의료 보건 모니터링 장면에 더 잘 적용할 수 있도록 한다 [3].

II. 본론

1. 인체 키포인트 인식

인체 키포인트 인식은 2D 와 3D 두 가지로 나눌 수 있다. 2D 키포인트 인식은 단일 이미지나 비디오 프레임에서 인체 키포인트를 감지하고 위치를 지정하여 인체 골격을 얻는 것이다 [4]. 반면 3D 키포인트 인식은 삼차원

공간에서 키포인트를 위치시켜 3D 좌표를 결과로 얻는다. 그러나 다중 인물 장면에서는 RGB 이미지의 가림 문제로 인해 모델이 모든 키포인트를 정확하게 감지하기 어려울 수 있다. 따라서 가림 문제를 효과적으로 처리할 수 있는 키포인트 인식 알고리즘이 절실히 필요하며, 이는 우리가 열화상을 이용한 인체 키포인트 인식을 고려하게 만든다. 열화상은 가림 문제를 처리하는 데 더 효과적이다 [5]. YOLOv8-SLG 모델은 열화상에서 인체 키포인트를 실시간으로 감지할 수 있으며, 기존 방법에 비해 더 빠른 수렴 속도와 더 나은 성능을 가진다.

2. YOLOv8-SLG 소개

YOLO(You Only Look Once) 모델은 목표 감지와 키포인트 인식 모델로, 인체 키포인트를 빠르고 정확하게 감지할 수 있다 [6]. 그러나 YOLOv8 은 가림 문제를 처리하는 데 결함이 있어, 본 논문에서는 YOLOv8 의 백본 네트워크, 넥, 그리고 디텍션 헤드를 개선하였다. 먼저, 계산 복잡도와 메모리 요구 사항 측면에서 향상된 C2f 모듈과 개선된 SPPF 레이어를 제안하여 로컬 특징 추출 능력을 향상시키는 동시에 정확성을 유지하였다. SCConv 로 일반 컨볼루션을 대체하여 특징 맵 처리를 최적화함으로써 전체 계산량과 파라미터 수를 크게 줄였다 [7]. 특징 추출 측면에서는 큰 분리 가능 커널 어텐션 메커니즘(LSKA)을 개선된 SPPF 레이어에 도입하여 큰 컨볼루션 커널을 분해함으로써 공간 어텐션 메커니즘을 강화하고, 고저층 의미 정보의 정확한 통합을 가능하게 하였다 [8]. 마지막으로, 다중 스케일 처리 측면에서는 공유 GroupNorm 컨볼루션으로 기존 컨볼루션을 대체하여 열화상에서 인체 세부 사항의 감지 능력을 강화하고, 다중 스케일 처리 능력을 개선하며 계산 비용을 낮췄다. 이러한 개선 사항은 열화상에서 인체 감지 및 위치 추적과 특징 추출 성능을

향상시키는 데 기여하였다. YOLOv8-SLG 의 구조는 그림 1 에 나와 있다 [9].

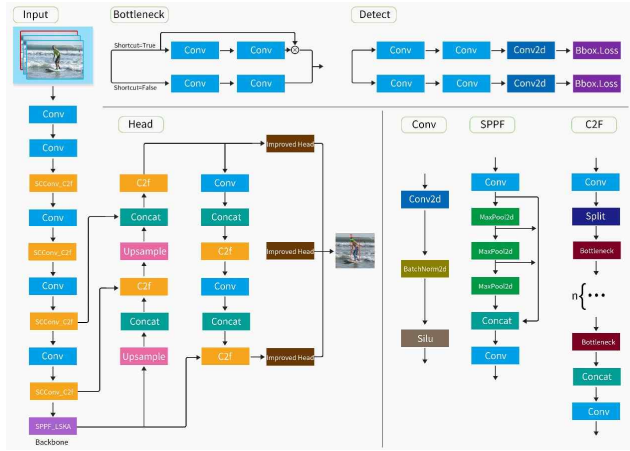


그림 1: YOLOv8-SLG 모델 구조도

3. 실험 결과 분석

개선된 모델의 성능을 검증하기 위해 "다중 가림 장면"과 "부분 가림 인체"를 포함한 열화상 이미지를 선택하여 동일한 하드웨어 조건에서 검출을 수행하였다. 결과는 그림 2 에 나와 있으며, 왼쪽 그림은 원래 YOLOv8 알고리즘을 사용한 검출 결과를, 오른쪽 그림은 개선된 YOLOv8-SLG 알고리즘을 사용한 검출 결과를 보여준다.

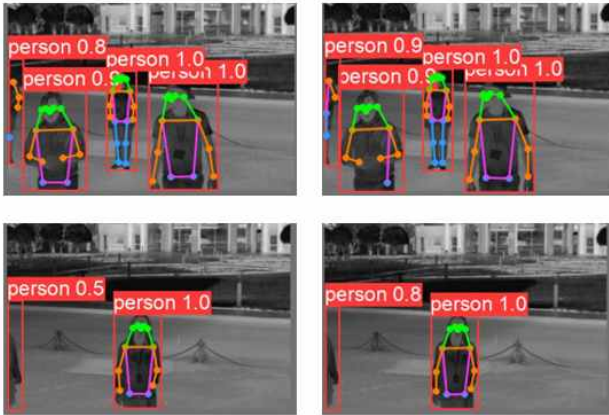


그림 2: YOLOv8 와 YOLOv8-SLG 의 동일 이미지에 대한 탐지 결과

Model	mAP50-95(B)	Recall(P)	mAP50(P)	GFLOPs
Yolov5	0.912	0.916	0.931	8.7G
Yolov6	0.921	0.933	0.945	12.6G
Yolov8	0.936	0.932	0.951	9.3G
Yolov8-PPA	0.926	0.942	0.953	10.5G
Our Model	0.937	0.946	0.967	8.6G

표 1: Our Model 과 주요 모델 실험 결과 비교

비교 결과 보면, 원래 YOLOv8 알고리즘은 검출 시 신뢰도가 낮았으며, 특히 부분적으로 가려진 열화상

이미지에서 두드러졌다. 반면, 개선된 YOLOv8-SLG 알고리즘은 세 가지 주요 강화 요소를 통합하여 특징 추출 능력을 향상시키고, 파라미터 수를 줄이며, 가림 및 부분적으로 보이는 장면에서 열화상 인체 키포인트 인식 성능을 크게 향상시켰다.

III. 결론

본 연구는 키포인트 검출 기술을 기반으로 한 YOLOv8-SLG 모델을 제안하였다. 이는 야간 또는 저조도 조건에서 열화상 인체 키포인트 인식을 위한 것이다. YOLOv8 모델의 개선을 통해 더 높은 정확도와 더 적은 계산 메모리로 열화상 이미지에서 인체를 효율적으로 검출할 수 있다. YOLOv8-SLG 는 부분 가림 및 복잡한 환경에서 다른 알고리즘보다 우수한 검출 성능을 보이며, 실시간 성능과 정확성 요구를 충족하여 이동 및 엡지 디바이스에 적합하고, 스마트 시티 발전에서 잠재력을 보여준다.

참 고 문 헌

- [1] CHEN, I.-Chien, et al. Multi-person pose estimation using thermal images. IEEE Access, 2020, 8: 174964-174971.
- [2] GUO, Yongping, et al. Identity-preserved human posture detection in infrared thermal images: A benchmark. Sensors, 2022, 23.1: 92.
- [3] ZENG, Jiayi; ZHONG, Han. YOLOv8-PD: an improved road damage detection algorithm based on YOLOv8n model. Scientific reports, 2024, 14.1: 12052.
- [4] MARTINEZ, Julieta, et al. A simple yet effective baseline for 3d human pose estimation. In: Proceedings of the IEEE international conference on computer vision. 2017. p. 2640-2649.
- [5] WANG, Huan; ZHOU, Luping; WANG, Lei. Miss detection vs. false alarm: Adversarial learning for small object segmentation in infrared images. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019. p. 8509-8518.
- [6] VARGHESE, Rejin; SAMBATH, M. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness. In: 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). IEEE, 2024. p. 1-6.
- [7] LI, Jiafeng; WEN, Ying; HE, Lianghua. Sconv: Spatial and channel reconstruction convolution for feature redundancy. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023. p. 6153-6162.

- [8] LAU, Kin Wai; PO, Lai-Man; REHMAN, Yasar Abbas Ur. Large separable kernel attention: Rethinking the large kernel attention design in cnn. *Expert Systems with Applications*, 2024, 236: 121352.
- [9] ZHU, Yean, et al. In-bed human pose estimation using multi-source information fusion for health monitoring in real-world scenarios. *Information Fusion*, 2024, 105: 102209.