

개선된 YOLOv8 자세모델을 이용한 역도 자세 감지 및 분류

남경남¹, 조인휘²

¹ 한양대학교 컴퓨터 소프트웨어학과

² 한양대학교 컴퓨터 소프트웨어학과 교수

Detection and Classification of Weightlifting Posture based on the Improved YOLOv8-Pose Model

Jingnan Nan¹, Inwhee Joe²

^{1, 2} Dept. of Computer Science, Hanyang University

요 약

최근 몇 년 동안 스포츠에서 딥 러닝 모델의 사용은 성능 분석 및 훈련 방법을 향상시킬 수 있는 잠재력으로 인해 많은 관심을 받고 있다. 이 논문에서는 역도 자세 감지 및 분류를 위해 특별히 개선된 YOLOv8 자세 모델을 제안한다. 이 연구의 주요 목표는 다양한 역도 연습, 특히 시작 및 상단 자세 구별에서 자세 추정 정확도와 견고성을 향상시키는 것이었다. 제안된 모델은 향상된 특징 추출 레이어와 역도 동작에 맞춘 최적화된 훈련 기술을 포함하여 원래 YOLOv8 자세 아키텍처에 대한 몇 가지 수정 사항을 결합한다. 저희는 역도 비디오로 구성된 사용자 지정 데이터 세트에서 모델을 분할, 주석을 달고 마지막으로 훈련하여 다양한 자세와 체형의 다양성을 보장한다. 저희의 실험 결과는 기존 YOLOv8 자세 모델에 비해 감지 정확도와 분류 성능이 크게 향상되었음을 보여준다. 개선된 모델은 더 높은 정밀도와 재현율을 달성하여 실시간 역도 자세 분석을 위한 신뢰할 수 있는 도구를 제공한다. 이 작업은 동작 분석 분야에 기여할 뿐만 아니라 자동화된 동작 모니터링 및 부상 예방에 대한 추가 연구를 위한 기반을 마련하였다.

I. 서 론

최근 몇 년간 딥러닝 기술은 스포츠 분야에서 광범위한 주목을 받았다. 딥러닝은 운동선수의 퍼포먼스를 분석하고 자세를 최적화할 수 있다. 다양한 스포츠 중에서 역도는 복잡하고 다양한 자세를 포함하고 있어 독특한 도전 과제를 안고 있다. 정확한 자세 감지 및 분류는 훈련 계획을 최적화하고, 비의도적인 부상을 방지하며, 전체적인 성과를 향상시킬 수 있다.

현재 방법과 비교하여, 전통적인 역도 자세 분석 방법은 주로 인간의 눈으로 관찰하고 주관적인 판단에 의존하는데, 이는 시간 소모가 크고 인간의 오판이 발생하기 쉽다. 그러나 컴퓨터 비전과 딥러닝 기술의 발전으로, 자동화된 역도 자세 감지 시스템은 더 이상 먼 미래의 일이 아니며, 더 정확하고 객관적인 분석을 제공하였다.

본 논문은 YOLOv8-Pose2[1] 모델을 개선하여 역도 자세 감지 및 분류의 도전 과제를 해결하는 것을 목표로 한다. 우리는 원래 모델을 기반으로 여러 가지 최적화를 수행하였으며, GAM[2] 을 추가하고 백본 네트워크를 수정하여 역도 자세의 세부 특성을 더 잘 포착하고, 특히 시작 위치와 정점 위치를 구별할 수 있도록 하였다. 주의 메커니즘은 모델이 중요한 영역에 집중할 수 있는 능력을 효과적으로 강화할 수 있으며, 수정된 백본

네트워크는 특징 추출 효과를 향상시켜 자세 추정의 정확성을 높일 수 있다.

II. 본론

A. 기본모델

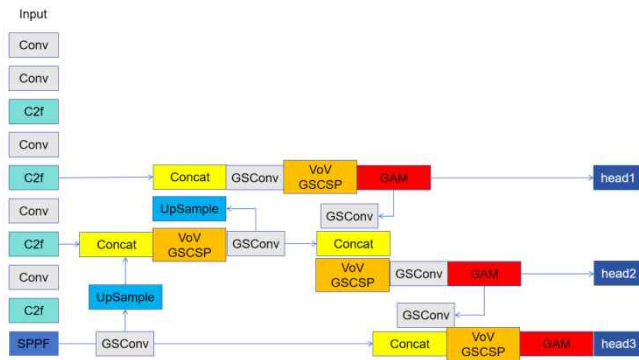
GAM[2] 메커니즘을 YOLO 모델에 통합하면 채널 및 공간 주의를 모두 강화하여 GAM의 특징 표현 기능이 활용된다. 이는 모델이 필수 객체 특징에 더 집중할 수 있게 하여 객체 검출의 배경 간섭을 완화한다. GAM은 적응형 선택 및 특징 맵 내의 중요 정보 향상을 용이하게 하여 대상 객체 특성을 보다 정밀하게 포착하고 이를 통해 검출 정확도 및 정밀도를 향상시킨다. GSConv[4]는 SC(표준 컨볼루션), DSC(깊이별 컨볼루션) 및 채널 셔플링의 조합을 활용하는 혼합 컨볼루션 접근 방식이다. 경량 작업의 경우 DSC는 탐지 속도를 크게 향상시키지만 1x1 고밀도 컨볼루션을 사용하면 더 많은 계산 리소스가 소비된다. "채널 셔플링"을 사용하려는 노력에도 불구하고 결과는 여전히 SC에 미치지 못한다. 고스트넷은 다양한 요인의 영향을 받아 SC의 경로로 어느 정도 되돌아간다. 많은 경량 모델이 유사한 설계

철학을 채택한다: 심층 신경망의 처음부터 끝까지 DSC 를 독점적으로 사용한다. 그러나 DSC 의 단점은 이미지 분류에 사용하든 탐지에 사용하든 백본 네트워크에서 직접 증폭된다.

B. GAM&GSConv 기반모델

본 연구에서는 YOLOv8 특징 융합 네트워크 내에서 주의 메커니즘을 통합하는 다양한 방법을 탐구하였다. 이러한 방법은 역도에서 폐색 및 사소한 형태 위반과 같은 문제를 해결하기 위해 설계되었다. GAM 을 통합함으로써 네트워크는 특징 맵의 채널 및 공간 특징에 가중치를 할당했다. 이 접근 방식은 관련 없는 특징의 가중치를 억제하면서 유용한 특징의 가중치를 향상시켜 중요한 정보를 포함하는 대상 영역에 더 집중하고 관련 없는 세부 사항을 억제했다. 이는 객체 검출의 정확도를 전반적으로 향상시켰다.

네트워크 구조는 GAM 주의 메커니즘을 원래의 YOLO 모델에 통합하는 것이 이상적인 결과를 도출했음을 보여준다. 그러나 이러한 통합은 모델의 복잡성을 증가시키고 처리 속도를 늦추었다. 이를 해결하기 위해 우리는 이 통합 네트워크에서 표준 컨볼루션 계층을 경량 컨볼루션 방식인 GSConv 로 대체했다. GSConv 는 유사한 학습 능력을 유지하면서 원래 컨볼루션의 약 60%에서 70%의 계산 비용을 갖는다.



(그림 2) YOLOv8-GG 의 아키텍처 다이어그램.

C. 경량화결과

공식 문서에서 다양한 척도의 모델을 제공하고 있음을 감안하여 YOLOv8n 과 YOLOv8m 두 가지 사양을 예로 선택하였다. 통합 전과 통합 후의 결과를 비교해 보면 모델이 상당히 간소화되었음을 알 수 있다.

<표 1> 매개변수와 GFLOPs 통계

	YOLOv8n	YOLOv8n-GG	YOLOv8m	YOLOv8m-GG
Params.	3.2M	5.02M	26.4M	19.51M
GFLOPs	8.83	13.2	86.2	47.4

이 실험의 주요 목적은 역도 자세를 정확하게 포착하는 것이다. 저희는 주의 메커니즘(GAM)을 YOLO 모델에 통합한 후 모델 경량화에 중점을 두었다. 마지막으로 원래 YOLO 모델과 비교했다. 실험은 YOLOv8m-Pose-GG 모델과 기본 모델 YOLOv8m-Pose 를 비교하여 맞춤형 역도 데이터 세트에서

수행되었다. 실험은 일관된 훈련 조건을 유지하기 위해 에포크를 100 으로 설정하여 동일한 실험 환경에서 수행되었다.

<표 2>역도 포즈 데이터 세트의 포즈 결과

Model	Params.	GFLOPs	mAP50	mAP50-90
YOLOv8m-pose	26.4M	86.2	0.962	0.8111
YOLOv8m-pose-GG	19.51M	47.4	0.980	0.884

기준 YOLOv8m-Pose-GG 모델이 기준 YOLOv8m-Pose 모델에 비해 가벼울 뿐만 아니라 포즈 검출에서도 더 나은 결과를 얻는다는 것을 관찰할 수 있다. YOLOv8m-Pose-GG 모델은 mAP50 에서 2% 향상, mAP50-90 에서 6.3% 향상, Params 와 GFLOPs 가 모두 감소하였다.

III. 결론

현재 자세 추정 응용 분야에서 새로운 자세 감지 기술은 매우 중요한 역할을 한다. 그러나 특정 인체 자세를 분할하는 것은 여전히 어려운 과제이며, 관련 데이터 세트를 개발하는 데 상당한 인력과 시간이 필요하다. 정확한 자세 분류를 위해서는 딥 러닝과 기계 학습 방법을 결합하는 것이 필수적이다.

또한 전통적인 컨볼루션 연산은 많은 수의 특징 맵을 생성하므로 상당한 계산 리소스가 필요하다. 본 연구에서는 역도 포즈에 초점을 맞추어 정밀도를 향상시키기 위해 딥 러닝에 주의 메커니즘을 도입했다. 또한 멀티 스케일 컨볼루션 모듈을 통합하여 다양한 조합을 탐색하여 기본 모델에 비해 가볍고 정확한 YOLO-GG 모델을 개발했다. 이 모델은 효율성을 유지하면서 포즈 인식 정확도를 향상시킨다. 전통적인 객체 감지 작업에도 적용할 수 있어 사용자 지정 데이터 세트에 대한 고품질 교육을 달성할 수 있다.

참 고 문 헌

- [1]. Qu W, Zhong S, Wu Y, et al. Development of a real-time pen-holding gesture recognition system based on improved YOLOv8[C]//2023 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML). IEEE, 2023: 1035-1039.
- [2]. Schulze, Kirsten E. "The Free Aceh Movement (GAM): anatomy of a separatist organization." (2004).
- [3]. Torralba, Antonio, Bryan C. Russell, and Jenny Yuen. "Labelme: Online image annotation and applications." Proceedings of the IEEE 98.8 (2010): 1467-1484.
- [4]. Li, Hulin, et al. "Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles." arXiv preprint arXiv:2206.02424 (2022).