

딥러닝 자세 추정 모델을 활용한 전면상황 위험 행동 엣지 기반 분류 기법

권태현¹, 주우현², 윤주상*

동의대학교 IT융합학과¹, 동의대학교 인공지능학과²*

dhxk9513@gmail.com¹, joowh0207@gmail.com², jsyoun@deu.ac.kr*

Front-State Dangerous Behavior Edge-Based Classification Technique Using Deep Learning Posture Estimation Model

Taehyeon Kweon¹, Woohyun Joo², Joosang Youn*

Dept. of IT Convergence., Graduate School, Dong-Eui Univ¹,

Dept. of Artificial Intelligence., Dong-Eui Univ²*

요 약

위험 행동 분류를 위한 다양한 기법이 존재하며, 그중에서도 CNN 기법이 가장 널리 활용되고 있다. 해당 기법은 대규모의 이미지 데이터를 통해 높은 정확도로 위험 행동을 분류할 수 있다는 장점이 있지만, 이러한 높은 성능을 달성하기 위해서는 대량의 데이터셋과 고사양의 컴퓨팅 디바이스가 필수적이라는 단점이 있다. 이러한 문제를 해결하기 위해 키포인트를 활용한 각도 기반 기법을 제안하였으나, 이 기법은 전면에서의 행동 분류에 있어 각도 정보만으로는 분류가 어렵다는 한계를 가지고 있다. 이를 위해 본 논문에서는 키포인트 간의 높이 차이를 활용하는 기법을 제시한다. 연구 결과 제안한 기법이 각도기반의 기법에 비해서 21.83%p 높은 정확도를 보였다.

I. 서 론

최근 안전 문제가 대두되면서 범죄와 안전사고 예방하기 위해 다양한 위험 행동 분류 기법이 제시되었다. 그중에서도 CNN(Convolutional Neural Network) 기법이 가장 널리 활용되고 있다. 해당 기법은 대규모의 이미지 데이터를 통해 높은 정확도로 위험 행동을 분류할 수 있다는 장점이 있지만, 이러한 높은 성능을 달성하기 위해서는 대량의 데이터셋과 고사양의 컴퓨팅 디바이스가 필수적이라는 단점이 있다. 특히, 학습 과정에서 많은 시간과 자원이 소요되기 때문에, 이를 효율적으로 적용하기 위해서는 상당한 기술적 인프라가 요구된다[1].

이러한 문제를 해결하고자 키포인트를 활용한 각도 기반 위험 행동 분류 기법이 제안되었다[2]. 이 기법은 YOLOv8-Pose 모델을 사용하여 키포인트를 추출한다. 해당 모델은 YOLOv8 모델에서 파생된 모델로 객체 탐지와 자세 추정을 함께 수행할 수 있도록 설계된 고성능 딥러닝 모델로, COCO-Pose 데이터셋을 사용해 사전 학습된 상태로 제공된다[3]. 별도의 학습 과정을 거치지 않기 때문에 고사양의 컴퓨팅 디바이스가 요구되지 않으며, 공공시설이나 학교와 같은 환경에서도 쉽게 적용할 수 있다는 장점이 있다. 그러나 해당 기법은 전면에서의 행동 분류에 있어 각도 정보만으로는 분류가 어렵다는 한계를 지니고 있다. 본 논문에서 기존 문제를 해결하기 위해 각도 정보로 판별이 어려운 상황에서는 키포인트 간의 높이 차이를 활용하여 위험 행동을 판단하는 기법을 제안한다.

II. 제안한 방법

기존 기법의 단점을 해결하기 위해 각도 정보로 판별이 어려운 상황에서 키포인트 간의 높이 차이를 활용하여 위험 행동을 분류하는 기법을 제안한다. 본 논문에서는 주먹질하는 행동, 발차기 행동, 넘어졌을 때의 모습 3가지의 행동 분류를 진행하였다. 표 1은 사람의 키포인트의 위치를, 표

2는 각 위험 행동별 키포인트의 높이 차이에 따른 분류 기준표를 보여준다. 표 2의 분류 기준을 정하기 위해 행동 별 100장의 이미지를 각각 키포인트 높이 차이를 구해 평균을 도출하였다.

표 1. 키포인트 항목[3]

번호	키포인트	번호	키포인트
0	Nose	8	Right-elbow
1	Left-eye	9	Left-wrist
2	Right-eye	10	Right-wrist
3	Left-ear	11	Left-hip
4	Right-ear	12	Right-hip
5	Left-shoulder	13	Left-knee
6	Right-shoulder	14	Right-knee
7	Left-elbow	15	Left-ankle
		16	Right-ankle

표 2. 위험 행동 분류 기준표

위험 행동	키포인트 높이 차이
Punch	$k_9 - k_5 < 0.32$ or $k_{10} - k_6 < 0.32$
	$k_7 - k_5 < 0.32$ or $k_8 - k_6 < 0.32$
Kick	$k_{15} > k_{14}$ or $k_{16} > k_{13}$
	$k_{15} - k_{13} < k_{16} - k_{14}$ or
	$k_{15} - k_{13} < k_{16} - k_{14}$
Fall-down	$k_{11} - k_5 < 0.07$ or $k_{12} - k_6 < 0.07$

k_n : 키포인트의 위치 값

III. 실험 및 결과

본 논문에서 제안한 기법에 대해 성능 테스트를 진행했다. 테스트는

JETSON ORIN 보드에서 진행하였으며 테스트 데이터는 AI Hub에서 제공하는 ‘시니어 이상행동 영상’과 ‘이상행동 CCTV’를 활용하였다 [2][4][5]. 또한 제안한 기법과 성능 비교를 위해 CNN 모델과 각도 기반 기법을 사용하였으며, 하이퍼 파라미터는 학습률 0.01, Epoch 200, Batch Size 16으로 설정하여 모델 학습을 진행하였다.

표 3. 성능평가표

위험행동	이미지	CNN	Angle	Height
		정확도	정확도	정확도
Punch	500	97.3	63.3	95.4
Kick	500	97.2	68.7	94.3
Fall-down	500	98.9	89.6	97.4
Average[%]		97.80	73.87	95.70

제안한 기법은 평균적으로 95.7%의 정확도를 보여주었으며, CNN 기법은 97.80%의 정확도를 보여주었다. 반면 각도 기반 기법이 평균적으로 73.87%로, 상대적으로 낮은 정확도를 보여주었다.

각도 기반 기법이 정면과 측면이 혼합되어 있는 테스트 이미지에서 73.87%의 낮은 정확도를 보였지만, 이를 개선한 제안된 기법은 정면 모습에서도 우수한 성능을 발휘하여 정확도를 크게 향상했다. 특히, 제안된 기법은 정면과 측면 모두에서 높은 정확도를 유지하며, 평균적으로 95.7%의 정확도를 보여주었다. 이는 정면 모습을 키포인트의 높이차를 활용해서 효과적으로 분석하여 위험 행동을 정확하게 분류할 수 있는 능력을 보여준다. 또한 CNN 모델과 비교했을 때 정확도는 조금 낮은 수준이지만 모델 학습을 위한 고성능 디바이스가 필요 없다는 것과 모든 픽셀을 추론하는 것이 아닌 키포인트의 값만 가지고 분류를 진행하기 때문에 실시간 환경에서 원활하게 작동한다는 이점을 가진다. 그림 1은 제안한 기법의 구현 결과 이미지이다.

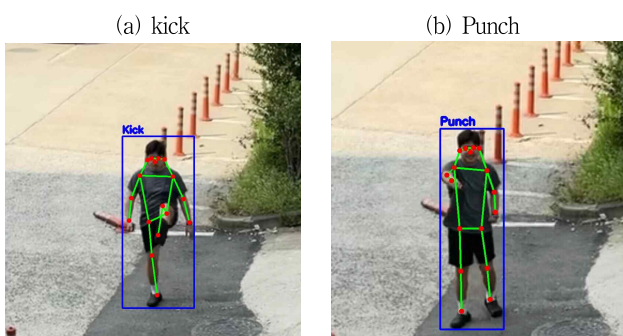


그림 1. 제안한 기법 구현 결과 이미지

IV. 결론

본 논문에서는 딥러닝 자세 추정 모델을 사용한 위험 행동 분류 기법을 제안하였다. 제안한 기법은 각도 기반 모델의 단점을 보완하고자 측면에서는 각도 기법을 사용하고 정면에서는 키포인트의 높이 차이를 활용하여 위험 행동 분류하였다. 제안된 기법은 영상 데이터에서 측면과 아울러 다양한 방향에서도 효과적으로 위험 행동을 분류할 수 있기 때문에 실시간 모니터링 시스템의 신뢰성을 크게 향상시킨다. 이를 통해 작업 현장에서 다양한 각도에서 발생할 수 있는 위험 행동을 포괄적으로 감지할 수 있으며, 안전 관리의 사각지대를 최소화하여 전반적인 사고 예방 효과를 높일 수 있을 것으로 기대된다. 추후 연구에서는 추가적인 위험행동을 분류하기 위한 연구를 진행 할 것이다.

ACKNOWLEDGMENT

이 논문은 과학기술정보통신부 및 정보통신기획평가원의 정보통신방송 표준개발지원사업(RS-2024-00397768)과 부산광역시 및 (재)부산테크노파크의 BB21plus 사업으로 지원된 연구결과임.

참 고 문 헌

- [1] S. Ji, W. Xu, M. Yang, and K. Yu, "3D Convolutional Neural Networks for Human Action Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 1, pp. 221-231, 2012.
- [2] T. Kweon and J. Youn, "An angle-based dangerous behavior classification edge system implementation using deep learning pose estimation model", Journal of Korea Multimedia Society, Vol. 27, No. 8, 2024.
- [3] Ultralytics, YOLOv8, 2023, <https://github.com/ultralytics/ultralytics>, (accessed Aug 12, 2024).
- [4] AI Hub, "AI Hub Data Set: Traffic Signs Data", 2024, <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=&topMenu=&aihubDataSe=data&dataSetSn=171>, (accessed Jul 11, 2024).
- [5] AI Hub, "AI Hub Data Set: Korean Traffic Signs Data", 2024, <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=&topMenu=&aihubDataSe=data&dataSetSn=167>, (accessed Jul 11, 2024).