

Application of SlowFast deep learning model and skeleton-based action recognition in violence detection problem

Dang Quang Huy

Abstract: The study proposes an intelligent video surveillance solution by combining the SlowFast model and skeleton recognition. The system uses SlowFast to capture the dynamic features of actions and YOLO-Pose to track body movements. This combination enables the system to accurately analyze violent behaviors (punching, kicking) in real time while minimizing false alarms from normal activities. This is a highly practical application for security protection in schools and public places.

Keywords: Violence detection, SlowFast model, Skeleton recognition

References

- [1] C. Feichtenhofer, H. Fan, J. Malik, and K. He, "SlowFast Networks for Video Recognition," in Proc. IEEE/CVF ICCV, 2019. [2] D. Maji, S. Nagori, M. Mathew, and D. Poddar, "YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation," arXiv preprint arXiv:2204.06806, 2022. [3] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning Spatiotemporal Features with 3D Convolutional Networks," arXiv preprint arXiv:1412.0767, 2015.