Fall Detection for Elderly People Living Alone Based on Yolov5 and ARKit

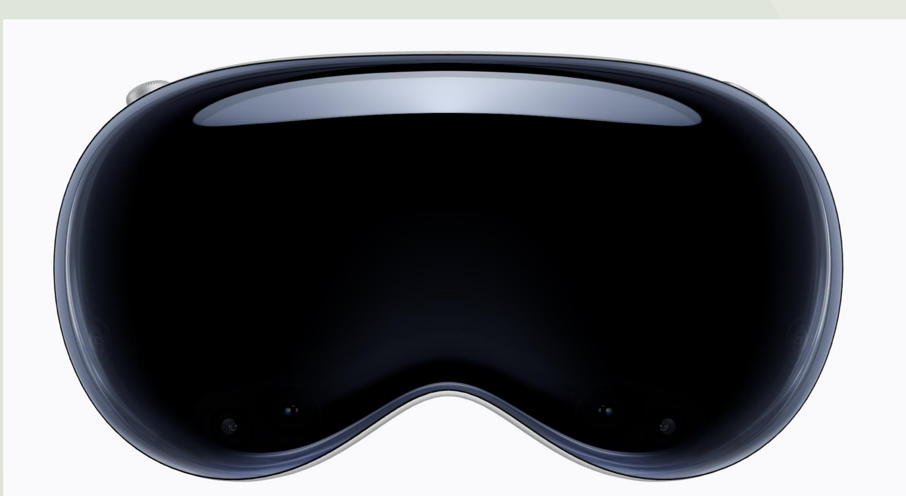
Hui Jin, Jayadithya Nalajala, Saniya Pandita, Sai Jahnavi Devabhakthuni

School of Computing and Engineering, University of Missouri-Kansas City, Kansas City, United States

Introduction

As the aging population continues to grow, the safety of elderly individuals living alone has become a pressing societal concern. The evolving social landscape presents new challenges for elderly care, particularly in addressing the high incidence of falls. These falls not only underscore the physiological vulnerability of the elderly but also highlight the deficiencies in current safety monitoring systems. Falls frequently result in severe injuries, reduced quality of life, and even life-threatening conditions. Therefore, developing a reliable and real-time fall detection technology is critical to ensuring the safety and well-being of this vulnerable demographic.

This paper presents a fall detection algorithm tailored for elderly individuals living alone, leveraging the synergistic strengths of ARKit, YOLOv5, and Apple Vision Pro. The proposed system combines YOLOv5's rapid object detection, ARKit's precise spatial tracking, and Vision Pro's immersive environmental awareness to achieve accurate and low-latency fall detection. Furthermore, we introduce a novel enhancement by integrating YOLOv10's Top-1 bounding box selection strategy into the YOLOv5 pipeline, significantly improving detection precision and reducing false positives. This fusion of technologies enables swift identification of fall events, shortens emergency response times, and ultimately minimizes the physical and psychological consequences of such incidents.



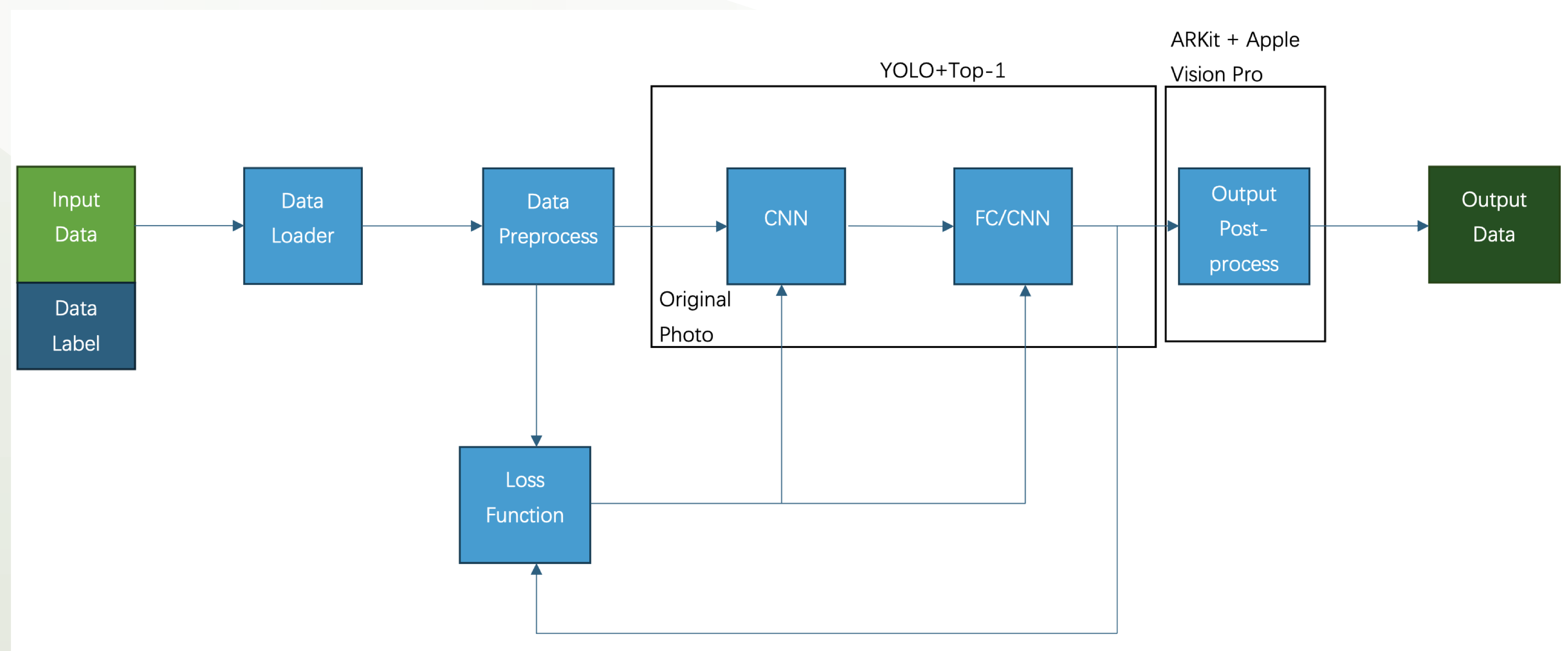
RESULTS

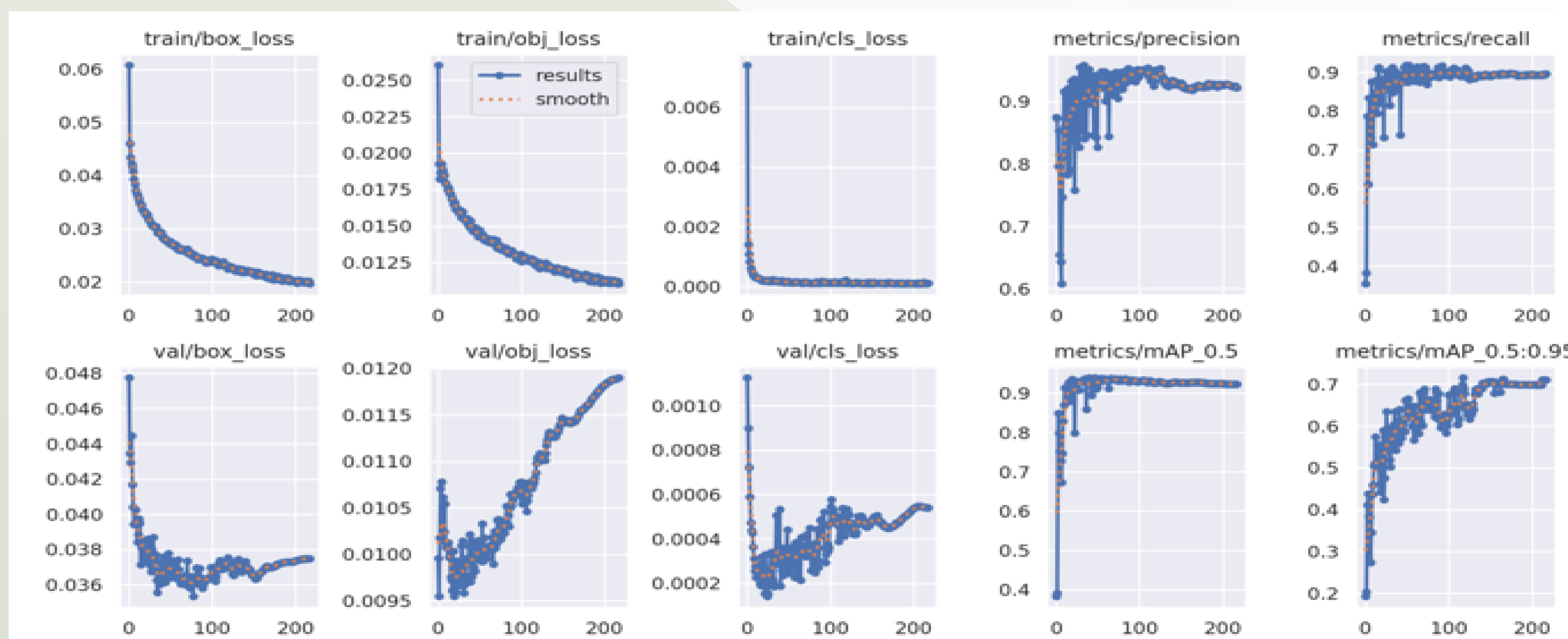
To evaluate the performance of the proposed algorithm, we conducted two comparative tests. The first test assessed the fall detection accuracy using the YOLOv5 model without ARKit, and the second incorporated ARKit and Apple Vision Pro. In the ARKit-free test, the system successfully identified fall behaviors with a prediction confidence score of 0.83. However, false positives were observed—postures such as sitting or lying down were occasionally misclassified as falls. In contrast, the system integrating ARKit and Vision Pro demonstrated superior performance, with improved environmental awareness, scene depth understanding, and precise spatial localization. The fall event was detected and annotated with a higher confidence score of 0.85, and no fall detection boxes were incorrectly assigned to partially visible individuals (e.g., only half of the body in view).

Furthermore, we integrated the YOLOv10 Top-1 bounding box selection strategy into the YOLOv5 framework to replace the traditional Non-Maximum Suppression (NMS) process. This optimization not only reduced false detection rates but also improved inference speed. Empirical tests revealed that the average detection time per frame was reduced by 0.1 milliseconds, enhancing the system's real-time performance without compromising accuracy.

System Architecture

ARkit-based YOLOv5 algorithm flow chart:





CONCLUSION

This paper focuses on the design and evaluation of a fall detection algorithm for elderly individuals living alone, leveraging ARKit, YOLOv5, and Apple Vision Pro. Through in-depth analysis of fall-related risks and practical challenges faced by this demographic, the proposed system demonstrates significant potential and advantages in real-time fall detection. By combining YOLOv5's object detection capabilities with ARKit's spatial perception and Apple Vision Pro's immersive and real-time environmental sensing, the algorithm achieves accurate and comprehensive identification of fall events.

Comparative experiments against conventional fall detection approaches confirm that the proposed method outperforms in both detection accuracy and real-time responsiveness. In particular, the integration of the YOLOv10 Top-1 bounding box selection strategy into the YOLOv5 detection pipeline has enhanced the detection precision while improving inference efficiency. This strategy replaces the traditional Non-Maximum Suppression (NMS) and reduces average detection latency by 0.1 milliseconds per frame, contributing to improved system performance in time-sensitive monitoring applications.

Despite these promising outcomes, several limitations remain. In real-world deployments, the algorithm's robustness and stability must be further strengthened to handle continuous, large-scale monitoring scenarios. Performance under varied lighting, occlusion, and posture conditions remains a challenge. The system's scalability and adaptability across diverse elderly populations and environmental settings also require further investigation.

Future work will focus on optimizing the model architecture and hyperparameters to improve detection efficiency and robustness. Efforts will also be directed toward enhancing adaptability to complex indoor scenes and postural variations, leveraging the spatial context awareness from ARKit and Vision Pro. Additionally, integrating the system with wearable devices or smart home infrastructure may further improve real-time feedback and user experience. Large-scale field testing and iterative user feedback collection will be essential for validating the algorithm's reliability and practical utility in daily living environments.

REFERENCES

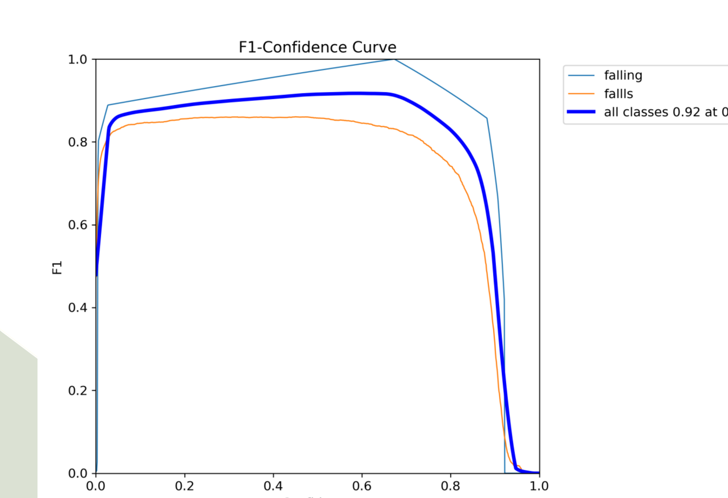
H. Le, M. Nguyen, W. Yan, and H. H. Nguyen, "Augmented Reality and Machine Learning Incorporation Using YOLOv3 and ARKit," Applied Sciences, vol. 11, no. 13, p. 6006, 2021.

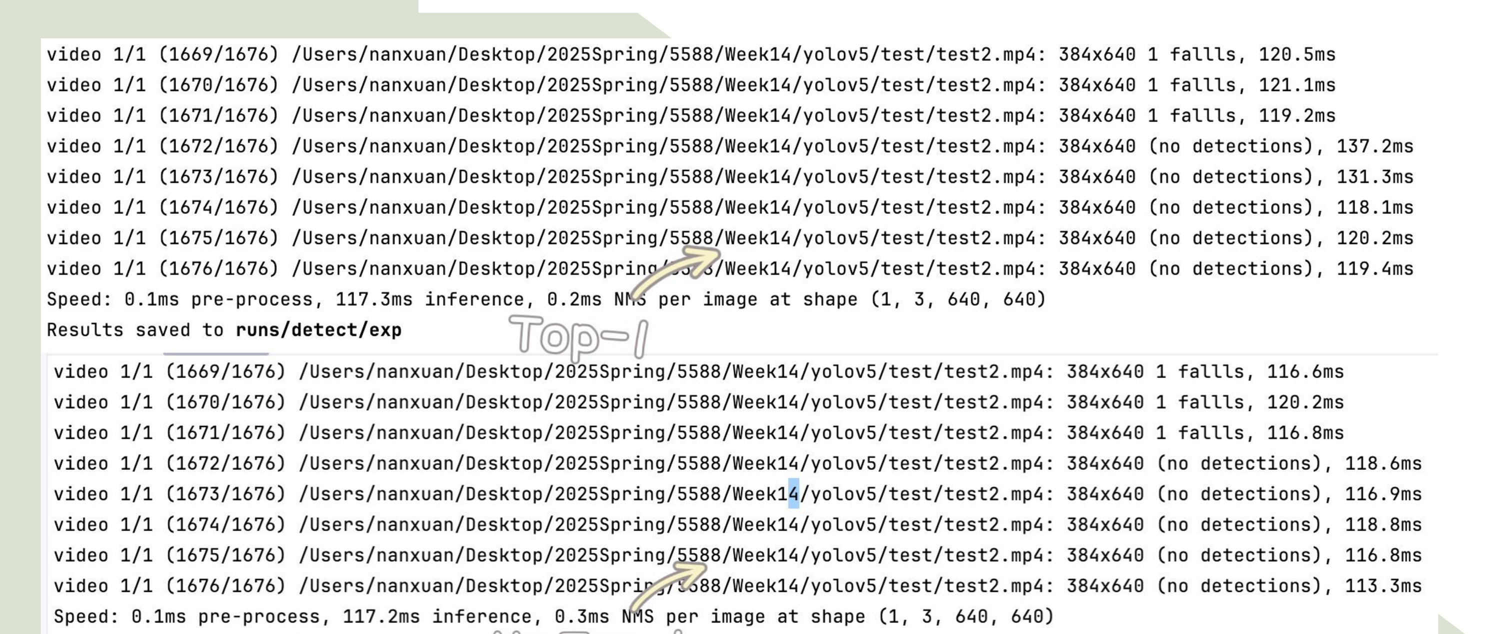
A. Gupta, R. Srivastava, H. Gupta, and B. Kumar, "IoT based fall detection monitoring and alarm system for elderly," in 2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), 2020, pp. 1-5.

K. M. Sharook, A. Premkumar, R. Aishwaryaa, J. M. Amrutha, and L. R. Deepthi, "IFall Detection Using Transformer Model," Lecture Notes in Networks and Systems, pp. 58-66, 2022.

ACKNOWLEDGMENT

This research was conducted as part of the Computer Science CS 5588 Data Science Capstone course under the guidance of Yugyung Lee, Ph.D., Professor. Additionally, it was supported by the technical resources by Sanda University and University of Missouri-Kansas City.





```
video 1/1 (1669/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 1 fallls, 120.5ms
video 1/1 (1670/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 1 fallls, 121.1ms
video 1/1 (1671/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 1 fallls, 119.2ms
video 1/1 (1672/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 137.2ms
video 1/1 (1673/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 131.3ms
video 1/1 (1674/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 118.1ms
video 1/1 (1675/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 120.2ms
video 1/1 (1676/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 119.4ms
Speed: 0.1ms pre-process, 117.3ms inference, 0.2ms NMS per image at shape (1, 3, 640, 640)
Results saved to runs/detect/exp
```

Top-1

```
video 1/1 (1669/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 1 fallls, 116.6ms
video 1/1 (1670/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 1 fallls, 120.2ms
video 1/1 (1671/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 1 fallls, 116.8ms
video 1/1 (1672/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 118.6ms
video 1/1 (1673/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 116.9ms
video 1/1 (1674/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 118.8ms
video 1/1 (1675/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 116.8ms
video 1/1 (1676/1676) /Users/nanxuan/Desktop/2025Spring/5588/Week14/yolov5/test/test2.mp4: 384x640 (no detections), 113.3ms
Speed: 0.1ms pre-process, 117.2ms inference, 0.3ms NMS per image at shape (1, 3, 640, 640)
Results saved to runs/detect/exp2
```

No Top-1