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

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

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

ABSTRACT

The article presents a fall detection algorithm tailored for elderly people living alone, utilizing YOLOv5, OpenPose, and ARKit. By integrating YOLOv5's rapid visual object detection, OpenPose's robust human pose estimation, and ARKit's advanced spatial perception capabilities, the proposed algorithm can accurately and swiftly detect falls among elderly individuals. The system encompasses functional modules including data collection, human pose recognition, and fall detection. Developed primarily for the iOS environment, it leverages Swift, Objective-C++, and Python programming languages to ensure comprehensive and efficient performance. Experimental evaluations indicate that the algorithm achieves excellent detection results, exhibiting high accuracy, recall, and mean Average Precision (mAP). However, certain limitations remain, primarily related to the constraints of YOLO versions utilized, as well as the size and diversity of the training dataset. Future enhancements will explore the integration of newer YOLO versions, dataset augmentation, modifications in network architecture, and incorporation of additional methodologies to further elevate the detection accuracy and overall system effectiveness.

Keywords—elderly people, ARKit, YOLOv5, OpenPose, fall detection

Introduction

As the problem of aging in society continues to intensify, the life safety of elderly people living alone has gradually attracted widespread attention from the society. The new social environment has brought new challenges to the safety management of the elderly. In the current social context, the frequent occurrence of elderly falls not only reflects the physiological vulnerability of the elderly group, but also exposes the shortcomings of the existing safety protection system in dealing with the problem of elderly falls. Fall accidents often cause serious injuries to the elderly, affect their quality of life, and even threaten their life safety. Therefore, developing a technology that can accurately and quickly detect falls of elderly people living alone and respond in time is of great significance to improving the life safety of the elderly.

At present, there have been some studies on the problem of elderly fall detection. Traditional fall monitoring methods are mainly based on sensors or video surveillance technology, but there are certain limitations and problems. The sensor is fixedly installed in a specific position and cannot perceive the dynamics of the whole body, which is prone to blind spots; and video surveillance technology is limited by camera layout, lighting conditions and privacy protection, making it difficult to achieve real-time and accurate fall detection. At the same time, traditional machine learning methods have the disadvantages of cumbersome detection steps, poor real-time performance, bloated model deployment, and lack of robustness in complex scenarios.

In order to solve the shortcomings of traditional methods, fall detection for elderly people living alone based on augmented reality and target detection algorithms has received widespread attention in recent years. AR technology can integrate virtual content with the real world, providing a more natural and intuitive way of interaction for elderly people living alone. The target detection algorithm can accurately detect fall behaviors in images or videos in real time, improving the accuracy and reliability of fall detection. The emergence of neural networks provides an effective way for feature extraction and utilization.

This paper primarily designs a fall detection algorithm for elderly people living alone based on ARKit, YOLOv5, and OpenPose, aiming to provide an effective monitoring solution. By integrating YOLOv5's fast visual detection, OpenPose's precise human pose estimation, and ARKit's robust spatial perception, the algorithm efficiently identifies fall events promptly. This timely detection will help shorten rescue response times and mitigate potential physical and mental health impacts caused by falls among elderly individuals living alone.

System Architecture

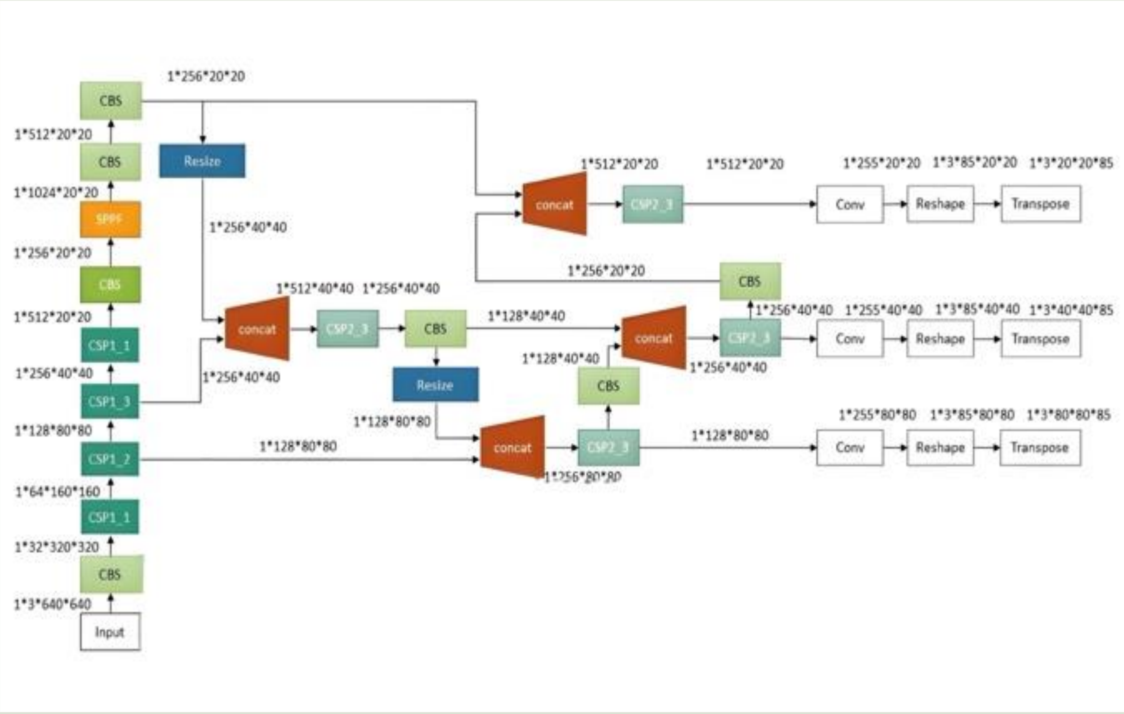
ARKit is an augmented reality development kit launched by Apple, which has powerful image recognition and tracking capabilities. ARKit integrates device camera image information and device motion sensor (including LiDAR) information.

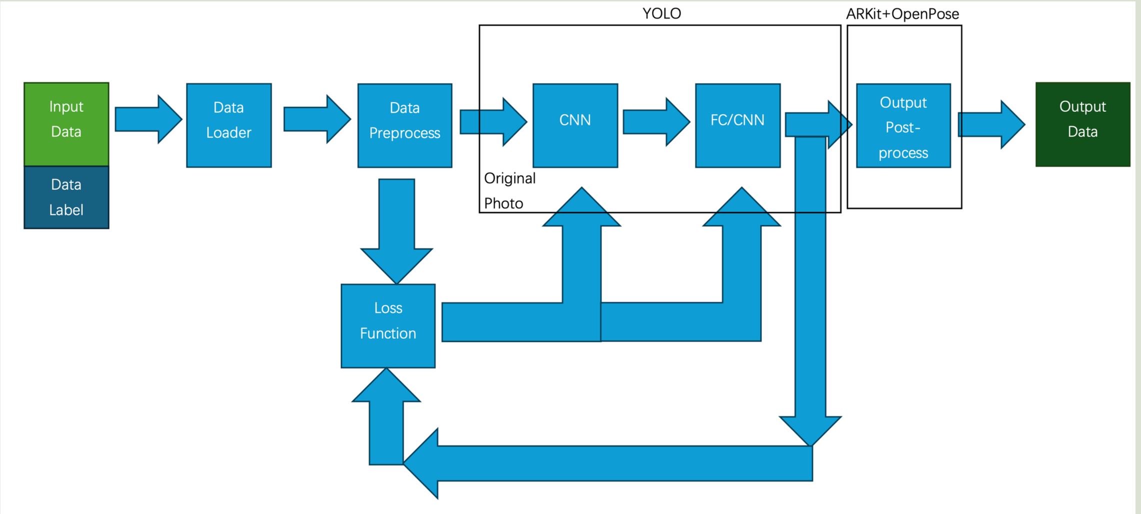
YOLOv5 (You Only Look Once) is an object detection algorithm based on deep learning. The algorithm can achieve real-time object detection and positioning by converting the object detection task into a regression problem. Compared with its predecessor YOLOv4, YOLOv5 has achieved a better balance between accuracy and speed.

This paper primarily designs a fall detection algorithm for elderly people living alone based on ARKit, YOLOv5, and OpenPose, aiming to provide an effective monitoring solution. It combines the 3D spatial modeling capabilities of ARKit, object detection technology of YOLOv5, and precise human pose estimation from OpenPose to accurately model the monitoring scene. Utilizing ARKit's depth perception and scene tracking, the algorithm establishes reliable spatial awareness. YOLOv5 enables real-time analysis of video streams to detect the posture and behavior patterns of elderly individuals, while OpenPose contributes detailed human pose data, significantly improving fall event identification accuracy. By effectively integrating these technologies, the algorithm promptly recognizes falls, facilitating immediate alerts and reducing rescue response times, thereby minimizing the physical and mental health impacts on elderly individuals living alone.

The algorithm consists of multiple modules. First, the input image or video is loaded through the data loader, and then enters the preprocessing module for preliminary processing. The processed data enters the feature extraction network (CNN) the YOLO network to extract key features, and is classified and regressed through the structured output network. The loss function is used for feedback optimization to improve detection accuracy. Finally, the output is processed in combination with the spatial information of ARKit and OpenPose to generate the fall detection result.

YOLOv5 network structure diagram and ARkit & OpenPose-based YOLOv5 algorithm flow chart:



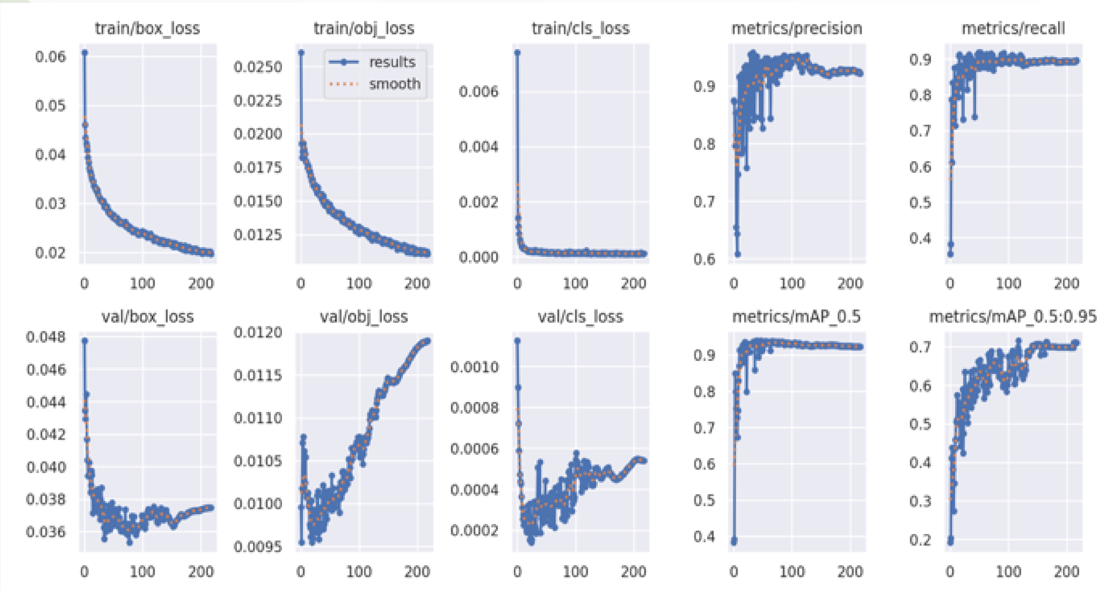


RESULTS

The bounding box loss of the training set and validation set gradually decreases, indicating that the accuracy of target localization is continuously improving. Object loss also decreased steadily, indicating that the model performed better at detecting the presence of fall events. The classification loss is close to zero, which is expected since the dataset only has two classes (standing and falling). As training progresses, both precision and recall improve significantly, with precision approaching 0.9 and recall reaching 0.9, indicating that the model's ability to correctly detect fall events is gradually increasing.

In this study, we designed an intelligent fall detection algorithm for elderly people living alone, combining the 3D spatial perception of ARKit and the real-time object detection capabilities of YOLOv5. The algorithm combines ARKit's precise mapping of the physical environment with YOLOv5's fast detection to ensure that fall events can be reliably identified, thereby improving safety and response efficiency.

We conducted two different tests to compare the algorithm. The first test evaluated the accuracy of fall detection without ARKit, while the second test evaluated the accuracy of fall detection with ARKit. As shown in Figure 5, the fall behavior can be detected without ARKit, and the model's confidence in this prediction is 0.83. As shown in Figure 6, there may be a problem of false positives, and the model may misjudge some non-fall postures (such as sitting) as falls. In Figure 7, the image of the detection results shows that the algorithm can effectively detect the fall event and annotate the event with a high confidence of 0.85. Using ARKit enhances the model's ability to detect falls in real-world environments, especially in understanding scene depth and spatial position. In Figure 8, a falls detection box is not assigned just because the person is only half.



CONCLUSION

This paper takes the design of a fall detection algorithm for elderly people living alone based on ARKit, YOLOv5, and OpenPose as the research direction and conducts an in-depth analysis and discussion on the fall problem of elderly people living alone. Through the analysis of the research results, it can be seen that the algorithm has certain advantages and potential in fall detection. By integrating ARKit's spatial perception, YOLOv5's object detection, and OpenPose's human pose estimation, the algorithm achieves more accurate and comprehensive fall identification. Comparative experiments with traditional fall detection technology show that the fall detection algorithm based on ARKit, YOLOv5, and OpenPose offers higher accuracy and real-time performance. However, there are still some problems and challenges in the research results. In practical applications, the robustness and stability of the algorithm need to be further improved. The operating efficiency and real-time performance of the algorithm also need to be optimized to meet the needs of large-scale real-time monitoring. The scalability and universality of the algorithm also need further research and improvement to meet the needs of different fall scenarios and elderly groups.

Future work includes optimizing the model structure and parameter settings of the algorithm to improve its accuracy and efficiency; further enhancing the robustness and stability of the algorithm, particularly in adapting to complex scenes and posture variations by leveraging OpenPose's pose estimation capabilities; exploring the combination of the algorithm with hardware devices to improve real-time performance and usability; and conducting more practical application cases and user feedback to further verify the reliability and effectiveness of the algorithm.

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.

Y. Yin, L. Lei, M. Liang, X. Li, Y. He, and L. Qin, "Research on Fall Detection Algorithm for the Elderly Living Alone Based on YOLO," 2021 IEEE International Conference on Emergency Science and Information Technology (ICESIT), vol. 12, no. 7, pp. 1515, Dec. 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.

G. Sun, S. Wang, and J. Xie, "An Image Object Detection Model Based on Mixed Attention Mechanism Optimized YOLOv5," Electronics, vol. 2023, pp. 403-408, 2023.

Y. Sun, J. Song, Y. Li, Y. Li, S. Li, and Z. Duan, "IVP-YOLOv5: An intelligent vehicle-pedestrian detection method based on YOLOv5s," Connection Science, vol. 35, no. 1, p. 2168254, 2023.

R. Y. Lee, H. Kamaludin, N. Z. M. Safar, N. Wahid, N. Abdullah, and D. Meidelfi, "Intelligence Eye for Blinds and Visually Impaired by Using Region-Based Convolutional Neural Network (R-CNN)," JOIV: International Journal on Informatics Visualization, vol. 5, no. 4, pp. 409-409, 2021.

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.