WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 7, Number 2, Year 2025
DOI: https://doi.org/10.62051/ijcsit.v7n2.05
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Design and Implementation of an Intelligent Elderly Fall Detection and Alert System Based on YOLOv10

Zedong Fu, Xinyi Li, Jiayao Yang, Xin Xiao, Yichen Deng, Zi Ye, Li Pang *

College of Computer and Software, Southwest Petroleum University, Chengdu, China
*Corresponding Author: Li Pang

ABSTRACT

With the accelerating global trend of population aging, frequent falls among older adults have become an increasingly serious social issue. To address this challenge, this paper proposes and designs an intelligent elderly fall detection and alert system based on the YOLOv10 deep learning framework. Leveraging advanced computer vision, the system achieves real time, accurate recognition of fall events and employs multi channel alert mechanisms to promptly notify caregivers after a fall occurs. Experimental results show that the system delivers high detection accuracy along with strong real time performance and robustness, effectively reducing both false alarms and missed detections. Compared with traditional methods and several deep learning approaches, it demonstrates significant advantages, providing solid technical support for home based eldercare and smart caregiving.

KEYWORDS

Elderly fall detection; YOLOv10; Deep learning; Intelligent alerting; Visual recognition

1. INTRODUCTION

Population aging has become a global phenomenon, particularly in developing countries where the proportion of older adults continues to rise. Consequently, eldercare challenges are mounting, and fall incidents among seniors are a critical concern [1]. Traditional caregiving methods are often inefficient and delayed, making rapid response at the moment of a fall difficult. Therefore, developing an intelligent system capable of real-time, accurate fall detection with timely alerts is of great practical significance for safeguarding seniors' lives, improving quality of life, and reducing the burden on families and society.

Fall detection technologies can be broadly categorized into wearable, environment-sensor-based, and vision-based approaches [2]. With advances in image processing and pattern recognition, vision-based fall detection has become the mainstream research direction. Early methods mainly relied on traditional image processing, but their robustness and accuracy suffer under real-world conditions such as illumination changes, complex backgrounds, and occlusions. In recent years, deep learning has brought transformative progress to vision-based fall detection. Building on the efficiency of the YOLO series, YOLOv10 introduces innovations such as NMS-free training and consistency optimization, further improving detection accuracy and inference speed, and providing a stronger algorithmic backbone for real-time fall detection.

In response, this paper designs and implements an intelligent elderly fall detection and alert system centered on the YOLOv10 algorithm and integrated with IoT technologies, forming a complete solution that combines video acquisition, fall detection, intelligent decision-making, and multi-channel alerting.

2. SYSTEM OVERVIEW

Functionally, the system uses YOLOv10 to accurately detect human targets in video frames and assess their posture. Once a fall is detected, it immediately triggers local audible/visual alarms and sends SMS messages to notify emergency contacts remotely. The system also logs historical events, enabling users to review alerts and related video clips.

In terms of performance, the system aims for high detection accuracy with a false alarm rate below 5% and a miss rate below 2% to minimize unnecessary disturbances while avoiding missed critical events. It is designed for robustness across different lighting conditions, background environments, and mild occlusions.

The video acquisition module (perception layer) captures real-time video streams via camera and feeds them to subsequent processing while ensuring stable, high-quality input for the detection algorithm. The fall detection module, the system's core, integrates a YOLOv10-based model to analyze incoming frames in real time, detect human targets, and—using features such as body posture, bounding-box aspect ratio, and inter-frame state changes—determine whether a fall has occurred. The decision module applies temporal logic to the detector's preliminary results to curb single-frame coincidences and reduce false alarms, thereby improving reliability. The communication module supports remote alerting by sending preset SMS notifications to caregivers via an SMS gateway. The data storage & management module records detailed information about each alert, including time, location, duration, and screenshots, and stores logs and key configuration parameters to aid system maintenance and optimization.

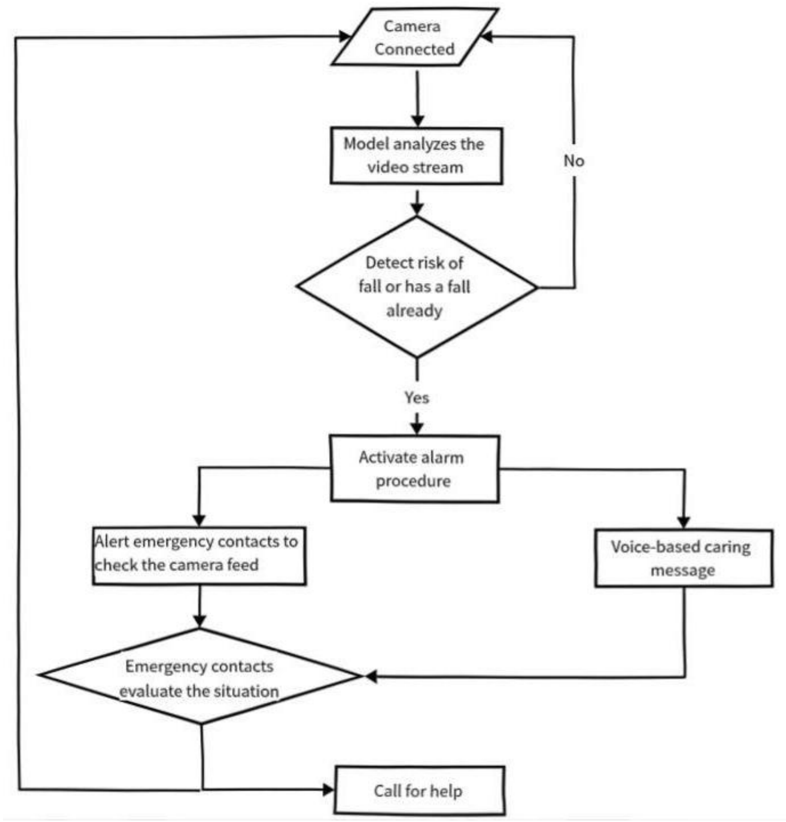

Figure 1. System workflow diagram

3. HARDWARE DESIGN

The hardware design considers YOLOv10's computational demands alongside real-time performance and system stability. As the system's "eyes," the camera's performance directly affects detection accuracy; thus a high-resolution, high-frame-rate camera is selected, with the specific model chosen based on deployment environment and budget.

For communication and alerting, a mobile communication module is integrated to send SMS notifications to caregivers and, where applicable, push alert information and data to a designated mobile app over cellular networks—ensuring timely notification wherever mobile coverage is available. The remote alerting mechanism uses the communication module to transmit event time, location, type, and snapshots or short video clips to preset phone numbers.

4. SOFTWARE DESIGN

The software stack implements the system's core functions—leveraging the YOLOv10 model, data processing, fall-posture inference, alert logic, and communication.

4.1. Core Detection Algorithm: YOLOv10 Principles and Architecture

YOLO (You Only Look Once) is known for its end-to-end detection paradigm and excellent speed-accuracy balance. YOLOv10 maintains high efficiency while further boosting detection performance [3]. The key idea is to frame object detection as regression: a single neural network directly predicts bounding boxes and class probabilities, simplifying the pipeline and reducing computation [4].

The architecture typically comprises Backbone, Neck, and Head. The Backbone extracts multi-scale feature maps from the input image using an efficient CNN to balance representation and compute. The Neck (e.g., FPN, PANet) fuses multi-scale features to enhance detection across object sizes. The Head makes final predictions of bounding-box coordinates, class confidences, and objectness. Through careful component design and optimization, YOLOv10 strikes a favorable balance among parameter count, computation, and detection performance [5].

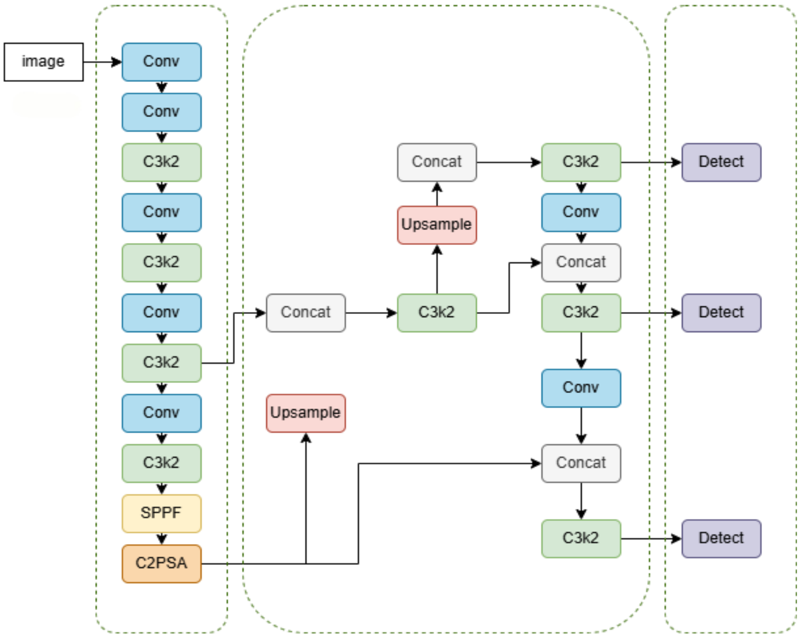

Figure 2. YOLOv10 network architecture

4.2. Fall Dataset Construction and Preprocessing

To improve generalization and robustness, diverse data augmentation strategies are used, including random rotation, cropping, horizontal flipping, and color jitter. Mosaic augmentation (training on a 4-image composite) increases per-batch diversity and helps with small-object detection [6]. MixUp further enriches training samples by blending image pairs and labels.

Annotations follow the YOLO format. Tools such as LabelImg are used to draw precise bounding boxes for human targets, with manual review ensuring accuracy and consistency. The dataset is split into training, validation, and test sets in a typical 7:1:2 ratio, ensuring representative coverage of scenes, lighting, and poses for training, hyperparameter tuning, and final evaluation.

4.3. YOLOv10 Training and Optimization

Transfer learning is adopted: a pretrained YOLOv10 model is fine-tuned on the constructed fall dataset, shortening training time and enhancing task-specific performance. A composite loss function—combining classification, regression, and objectness terms—follows the official YOLOv10 implementation with appropriate adjustments.

Optimizers such as AdamW and SGD are compared, with learning-rate schedules like cosine annealing or step decay used to dynamically adjust learning rates, avoid local minima, and accelerate convergence. Key hyperparameters (batch size, initial learning rate, epochs) are tuned experimentally.

Evaluation metrics used during training and on the test set include Precision, Recall, F1-score at various IoU thresholds, and FPS (frames per second). These metrics jointly reflect detection accuracy, sensitivity to fall events, and real-time performance.

4.4. Fall-Detection Workflow Implementation

The core software workflow proceeds as follows. First, OpenCV captures real-time video and splits it into frames. Each frame is preprocessed (resize, normalization, channel ordering) to meet YOLOv10 input requirements. The frame is then fed into YOLOv10 for inference, producing bounding boxes, class confidences, and objectness scores. Based on detections of the "person" class, the system analyzes geometrical features and inter-frame dynamics to infer a fall—for example, a significant increase in the bounding-box aspect ratio beyond a preset threshold (indicating a vertical-to-horizontal posture change) that persists across multiple consecutive frames, optionally combined with a rapid downward shift of the box center, triggers an initial fall decision. Although YOLOv10 employs NMS-free strategies during training, a lightweight post-processing NMS may still be applied at inference to handle edge cases and remove redundant, highly overlapping boxes.

4.5. UI Design and Backend Services

The backend is built on Spring Boot, with business logic organized around three modules: user center, monitoring management, and video-record management—handling user identity, alerts, and video-stream operations while interacting with the data layer. The data layer employs stored procedures for complex operations, transactions for atomicity, and a custom permission model aligned with business-layer controls. Redis caches frequent data with expiration policies; synchronization ensures consistency between MySQL and cache. File I/O is supported, and RabbitMQ enables asynchronous messaging. An end-to-end permission framework safeguards security and efficiency from business logic down to data.

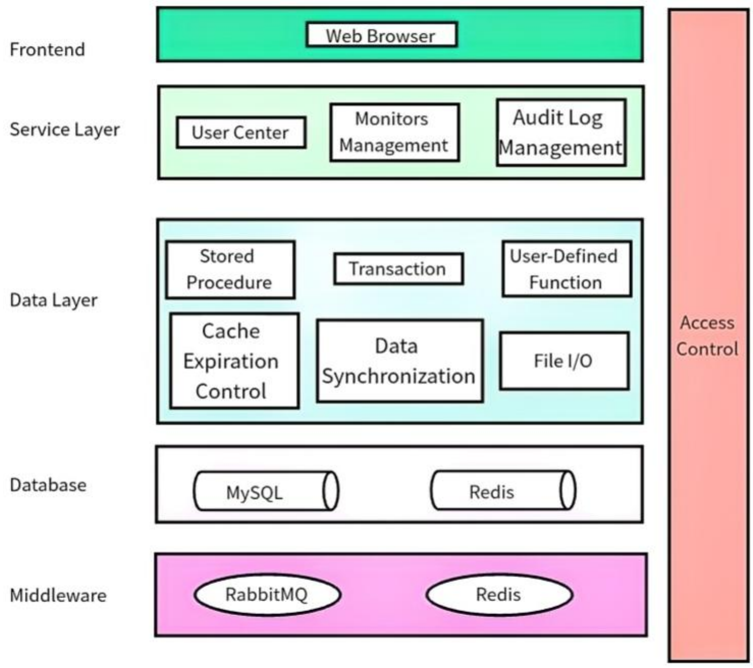

Figure 3. Backend architecture

The frontend uses React to deliver a responsive UI featuring real-time monitoring, alert-record queries, and system parameter configuration. The interface emphasizes usability with a clean layout and intuitive interactions, and includes accessibility-minded adjustments to element sizes, typography, and interaction flows to serve older users well.

Figure 4. Frontend concept

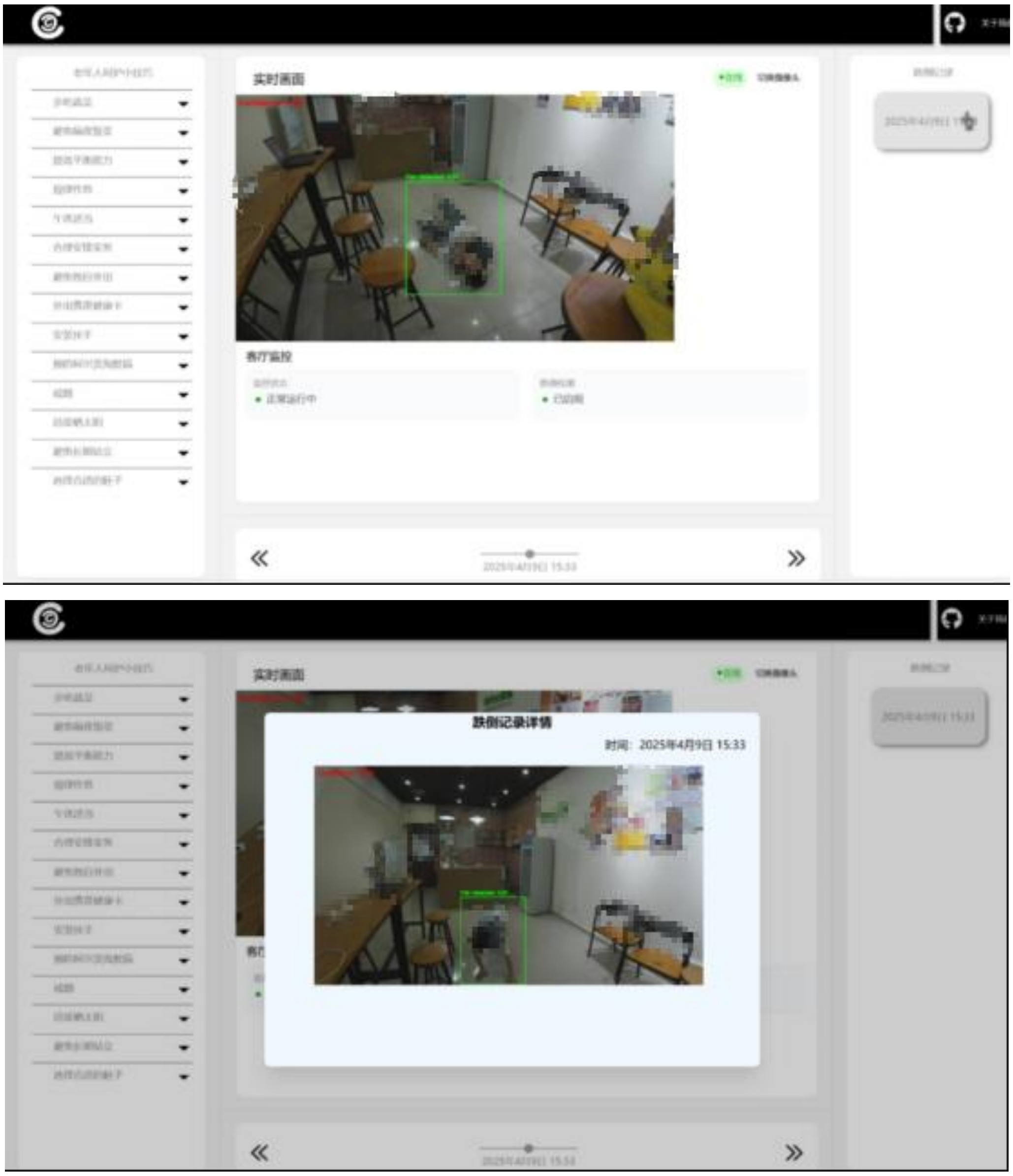

Figure 5. Example system interface

5. CONCLUSION

This paper addresses the pressing problem of falls among older adults in aging societies by designing and implementing an intelligent fall detection and alert system based on the YOLOv10 framework. We successfully apply an advanced real-time object detector to precise recognition of falls and integrate it into a full alerting solution. By constructing a high-quality fall dataset, thoroughly training and optimizing YOLOv10, and introducing a temporal confirmation mechanism for events, the system achieves higher detection accuracy and resilience to false alarms. Experiments indicate strong performance in accuracy, real-time operation, and robustness, reaching an average 32 FPS and F1-score above 90%, validating YOLOv10's substantial potential for elderly fall detection. Limitations remain in extremely complex scenes. Future work includes exploring lighter, more robust YOLOv10 variants; applying quantization and pruning for efficient edge deployment; and integrating with smart-home platforms to embed fall alerts into broader home IoT ecosystems.

ACKNOWLEDGEMENTS

This work has been supported by the key project of extra-curricular open experiment of Southwest Petroleum University (grant no. 2023KSZ61009).

REFERENCES

[1] Wan, H., Chen, L., Pan, L., et al. (2013). Research on the Fall Detection Algorithm for the Elderly Based on Intelligent Video Surveillance. Journal of Taiyuan University of Science and Technology, 34(04), 245-249.

[2] Ye, L., Qi, Y., Shi, J., et al. (2020). A Review of Human Fall Detection Technologies. Electronic Testing, (02), 50-51+65. DOI: 10.16520/j.cnki.1000-8519.2020.02.017.

[3] Jin, X., Zhang, B., Song, J. (2013). Design and Development of a Tag Recognition System Based on RFID Technology. Journal of Communication University of China (Natural Science Edition), 20(01), 35-39. DOI: 10.16196/j.cnki.issn.1673-4793.2013.01.011.

[4] Xin, W., Hao, H., Bu, M., et al. (2021). Real-Time Static Gesture Recognition Method Based on ShuffleNetv2-YOLOv3 Model. Journal of Zhejiang University (Engineering Edition), 55(10), 1815-1824+1846.

[5] Liu, Y., Sui, J., Wei, X., et al. (2023). Real-Time Small Object Detection Based on Lightweight YOLOv4. Progress in Laser and Optoelectronics, 60(06), 107-114.

[6] Chen, S., Yuan, Y. (2023). Improved YOLOv5 Sign Language Alphabet Recognition Algorithm. Mini and Micro Computer Systems, 44(04), 838-844. DOI: 10.20009/j.cnki.21-1106/TP.2021-0664.