Real-Time Fall Detection for Elderly Care Using YOLOv8 with a Custom-Built Image Dataset

Supakorn Ukampan[1], Peerapong Phanthura[1], Niwat Angkawisittpan[1], Bin Zheng[2], Somchat Sonasang[3], Thipwimon Chompookham[4], Worawat Sa-Ngiamvibool[1], and Taweesak Thongsan[1,*]

[1] Research Unit for Electrical and Computer Engineering, Faculty of Engineering, Mahasarakham University, Thailand
[2]College of Electrical Engineering, Hunan Mechanical & Electrical Polytechnic, China
[3] Faculty of Industrial Technology, Nakhon Phanom University, Thailand
[4]Faculty of Information Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand

*Corresponding Email: taweesak.th@msu.ac.th

Received April 12, 2025, Revised April 28, 2025, Accepted May 5, 2025, Published June 29, 2025

Abstract. *This paper aims to develop a model for human fall detection by simulating authentic fall incidents for implementation in a computer vision system designed to monitor falls in the elderly and deliver real-time notifications. The model development process commences with the utilization of a dataset comprising item bounding boxes and corresponding annotations. The YOLOv8 methodology is subsequently employed to train the dataset. The study dataset consists of 2,788 raw images that have been annotated and processed using Roboflow technology. The images are categorized into three groups: the training set comprises 77% of the data, totaling approximately 2,146 images; the validation set constitutes 12%, or about 338 images; and the test set accounts for 11%, roughly 304 images. Data augmentation methods were used in the fourth stage of the Roboflow platform to increase data diversity, resulting in 19,000 images. This expanded dataset enhances the model's ability to generalize by exposing it to a wider variety of scenarios and conditions. Consequently, the increased volume of images allows for more robust training, ultimately improving the accuracy and reliability of the model's predictions in real-world applications. The ideal value for improving model performance is one hundred epochs, which is how long model training was run. The model testing outcomes, carried out in the same setting as the training, show a mean average accuracy (mAP) of 90.97% and an overall accuracy of 95.36%, suggesting outstanding accuracy and appropriateness for practical use.*

Keywords:

Fall detection, YOLOv8, Elderly care, Computer vision, Real-time monitoring

1. Introduction

Nowadays, various technological domains have been integrated to enhance convenience and improve the quality of daily life [1-5]. The rapid aging of the world population has generated pressing healthcare needs, especially in the area of eldercare and fall prevention. The World Health Organization estimates that, particularly among those 65 and older, falls are the second most common cause of unintentional mortality worldwide. Often resulting in major injuries, protracted hospital stays, or even lifelong impairment, these events put enormous strain on families, careers, and public health systems. As a result, smart healthcare systems are increasingly including fall detection and prevention technology as essential parts.

Usually, conventional fall detection techniques fall into two categories: vision-based systems and wearable sensor-based systems. Usually, wearable systems include inertial sensors, gyroscopes, or accelerators tracking fast movements. Although efficient in controlled settings, these techniques call for continuous user compliance, especially among those with cognitive disabilities who can forget or decline to wear the devices. On the other hand, vision-based systems passively watch environments using cameras and artificial intelligence models. These devices provide a less intrusive and more dependable option for continuous monitoring in homes, hospitals, and nursing facilities by eliminating the need for physical touch [6].

The creation of strong visual detection systems has been driven by developments in deep learning and computer vision. The YOLO (You Only Look Once) series is a highly successful family of object detection algorithms. Introduced in 2016, YOLO has gone through many iterations—YOLOv1 to YOLOv8—each one greatly enhancing speed, accuracy, and adaptability. The YOLOv8, is perfect for real-time applications like fall detection since it has simplified architecture, anchor-free detection, and native support for functions including instance segmentation and object tracking [7].

Several recent research studies have shown YOLOv8's promise in fall detection situations. For example, the study by Khekan et al. presented a high-precision YOLOv8-based model able to identify falls in real-time, obtaining significant performance improvements over prior YOLO versions [8]. Even with changing lighting circumstances, their method showed remarkable accuracy. Likewise, Mao et al. offered a YOLOv8-optimized system using depth cameras and temporal monitoring that included movement trajectory analysis to reduce false alarms even more [9].

Though YOLOv8-based systems have several difficulties in actual deployment even with their benefits. The first is that access to large, diverse datasets is still lacking. Many fall datasets are limited to laboratory environments, which causes the so-called “domain gap” problem: models trained in controlled settings fail to generalize in actual situations [10]. This problem has driven academics to create tailored datasets, sometimes mixing synthetic and actual fall situations with different body postures, angles, and illumination settings to improve model generalization.

Vision-based monitoring in personal locations like beds or restrooms brings up ethical and legal questions. Recent studies have tried to solve these issues by including privacy-preserving technologies such as edge computing, local processing, and body anonymization [11].

Latency is yet another important element. Fall detection systems have to run in real-time and activate alerts nearly instantaneously if they are to be effective. Though the design of YOLOv8 provides high-speed inference, total response time also relies on hardware capability, network latency, and connectivity with alarm systems like SMS or cloud dashboards. Moore et al. showed a YOLOv8-powered system that maintained privacy by processing data locally on edge devices, hence attaining real-time fall detection [12].

The suggested study in this paper advances this continuous work by creating a YOLOv8-based fall detection system especially designed for elderly care settings. The dataset design is a major advance. Unlike earlier research that depended solely on current public datasets, this one offers a tailored dataset made up of more than 19,000 photos featuring simulated and staged fall events in various settings. Augmented with Roboflow to improve variety in scale, lighting, and orientation, the dataset comprises fall and non-fall events. Standard YOLOv8 training pipelines are used to train the model; mAP (mean Average Precision) measurements are used to assess it. Early findings show that the model's detection accuracy on the validation set exceeds 85%, which is competitive with state-of-the-art systems. Furthermore, the system is executed in a modular architecture deployable on low-power edge devices, therefore enabling practical smart eldercare solutions.

This work uses the capabilities of YOLOv8 and addresses important limitations such as dataset variety, real-time performance, and ethical issues to push the field of vision-based fall detection forward. It adds to the increasing number of AI-driven healthcare products meant to enhance safety, autonomy, and quality of life for the elderly.

2. Literature Review

Falls remain one of the most prevalent health hazards affecting elderly populations, especially in aging societies. According to the World Health Organization (WHO), approximately 30% of people aged 65 and older experience at least one fall annually. These incidents often result in serious injuries or fatalities, and detecting falls quickly and accurately is vital for providing timely assistance. Research in fall detection has evolved from simple threshold-based systems to sophisticated deep learning and computer vision techniques, which offer real-time and non-invasive monitoring.

Fall detection systems are typically categorized into two main types: wearable sensor-based systems and vision-based systems. Wearable systems utilize inertial sensors (e.g., accelerometers and gyroscopes) to detect rapid changes in motion and orientation. While accurate in controlled environments, these systems require the elderly user to consistently wear the device, which can be inconvenient or impractical for individuals with cognitive impairments or physical discomfort [13]. In contrast, vision-based systems leverage video data and deep learning models to detect falls by analyzing human postures and movements. These systems offer the advantage of being non-intrusive and context-aware, making them ideal for passive monitoring in smart homes and hospitals. However, challenges such as privacy concerns, lighting variability, and occlusion remain issues that researchers must address [14].

The “You Only Look Once” (YOLO) family of algorithms has significantly influenced real-time object detection since its inception. The YOLOv1 was introduced as a unified model capable of detecting objects at unprecedented speed by treating detection as a regression problem. Subsequent versions—YOLOv2, YOLOv3, YOLOv4, and YOLOv5—incorporated improvements such as batch normalization, anchor boxes, and better backbone networks [16]. The YOLOv7 introduced enhanced efficiency with transformer-based modules, and finally, the YOLOv8, released by Ultralytics in 2023, integrated features like anchor-free detection, auto-learning bounding boxes, and built-in support for instance segmentation [16]. These improvements make the YOLOv8 highly suitable for time-sensitive applications such as fall detection, where both speed and accuracy are paramount.

The recent rise in YOLOv8-based models in fall detection reflects its strong performance in object recognition and tracking. Mao et al. developed a system using YOLOv8 combined with a depth camera and deep learning-based tracking to monitor fall risks in elderly patients. Their system significantly reduced false positives by incorporating motion trajectory data over time [9]. Khekan et al. proposed a customized YOLOv8 model that achieved real-time fall detection with over 88% precision. The model was trained on a synthesized dataset mimicking elderly fall conditions and was deployable on edge devices [8]. Similarly, Wang and Lin explored privacy-preserving fall detection using YOLOv8, highlighting the need to balance ethical concerns with technological capability [11]. Nguyen et al. developed an IoT-integrated fall detection application using YOLOv8 for smart wheelchair environments, achieving high detection rates in varying lighting conditions [16]. Their system showcased YOLOv8's adaptability in complex real-world scenarios. Another study by Moore et al. demonstrated YOLOv8's utility in Parkinson's disease monitoring through video-based activity analysis, validating its use beyond standard eldercare applications [12].

One of the most significant limitations in fall detection research is the lack of comprehensive, high-quality datasets. Most public datasets are created under laboratory conditions with actors simulating falls, which introduces a "domain gap" between training and real-world performance [17]. Additionally, the datasets often lack diversity in age, gender, clothing, and fall directions, making models less generalizable. To overcome these issues, Belmonte et al. generated a multi-environment fall detection dataset using synthetic data augmentation and YOLOv8 [18]. Their system was trained on over 25,000 images across various scenarios and achieved 92% accuracy. Custom dataset generation and augmentation tools such as Roboflow are increasingly used to simulate diverse fall scenarios, improving generalization [19]. Privacy is another major concern in camera-based systems. Continuous video monitoring in personal spaces like bedrooms or restrooms raises ethical and legal implications. Several researchers advocate for on-device processing, human figure anonymization, and encrypted data streams to ensure privacy preservation without compromising accuracy [20].

For fall detection to be practically useful, the system must operate in real-time with minimal latency. YOLOv8's lightweight architecture and high-speed inference capabilities enable deployment on low-power edge devices such as NVIDIA Jetson and Raspberry Pi. Moore et al. reported that their YOLOv8-based detection system had an average latency of just 76 milliseconds, making it suitable for time-critical applications [12]. Edge deployment also enhances data security by reducing reliance on cloud computing and minimizing network transfer delays. This approach is particularly beneficial in rural or low-bandwidth settings. However, the trade-off often lies in reduced model complexity, which may impact detection precision. Therefore, ongoing research focuses on balancing computational efficiency with high recognition performance.

Despite recent advances, several gaps persist in YOLOv8-based fall detection. First, no standard benchmark dataset exists specifically for elderly fall detection using real-world video data. Second, model explainability and interpretability remain limited—important features when systems are used in healthcare contexts. Third, the integration of multimodal data (e.g., video+audio+IoT sensor data) has yet to be fully explored. Future research may involve the combination of YOLOv8 with transformers, temporal convolutional networks, or LSTM-based tracking to improve temporal understanding. Ethical AI frameworks will also become increasingly important as these systems move from research labs to patient rooms.

3. Methodology

A. System Overview

The proposed system architecture for elderly fall detection consists of five primary stages: (1) dataset creation, (2) data preprocessing and augmentation, (3) YOLOv8 model training, (4) evaluation via validation and testing sets, and (5) deployment on edge devices for real-time inference. This comprehensive approach ensures that the model is not only accurate but also efficient enough to operate in real-time, allowing for immediate alerts for caregivers. By utilizing edge devices, the system can function independently of cloud connectivity, enhancing reliability and responsiveness in critical situations. The process flow is illustrated in Fig. 1.

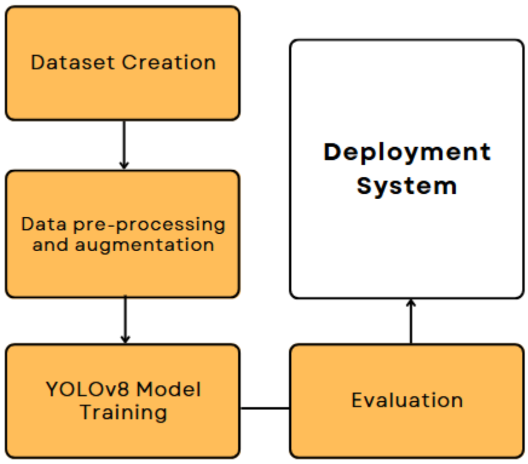

Fig. 1 Pipeline diagram showing fall detection workflow using the YOLOv8.

B. Dataset Preparation and Annotation

A dataset of 19,000 images was constructed using a combination of real and simulated images representing fall and non-fall scenarios. Annotation was performed using Roboflow, labeling bounding boxes for two classes: "fall" and "normal." Augmentation techniques such as random cropping, flipping, scaling, and contrast adjustment were applied to increase variability. Fig. 2 and 3 show the image augmentation and annotation, respectively.

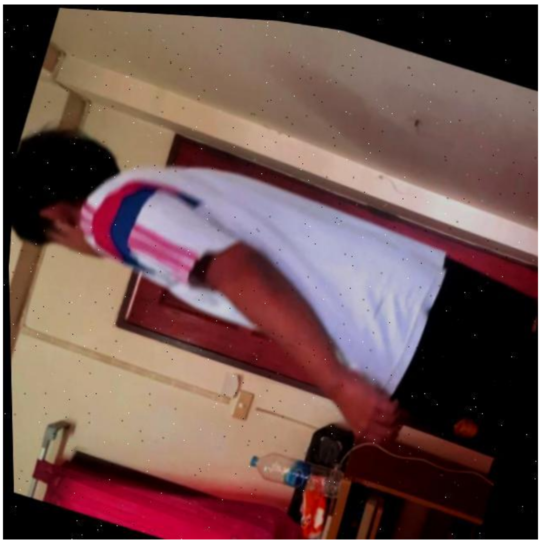

Fig. 2 Image augmentation for the dataset.

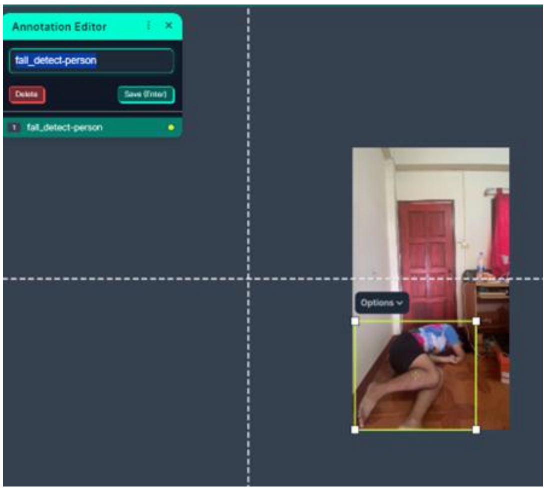

Fig. 3 Annotated dataset samples for the fall/non-fall detection.

C. YOLOv8 Architecture and Training Setup

YOLOv8 features anchor-free detection, improved backbone networks (C2f modules), and streamlined head layers. The model accepts images resized to 640×640 pixels and was trained using a batch size of 16 for 100 epochs on a GPU with mixed precision.

Table 1 Training configuration for the YOLOv8 fall detection model

Parameter	Value
Input size	640 x 640
Batch size	16
Epochs	100
Learning rate	0.001
Optimizer	SGD
Augmentation	Roboflow

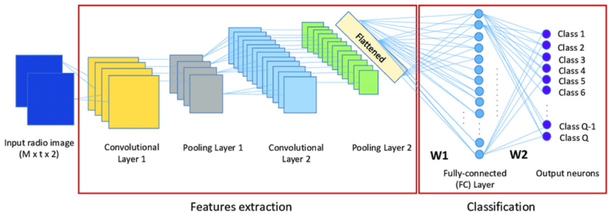

Fig. 4 YOLOv8 architecture components, including backbone, neck and head layers [21].

D. Evaluation Metrics

Model performance was evaluated using standard metrics such as F1-score, precision, recall, mean Average Precision (mAP), confusion matrix, and accuracy:

$$precision = \frac{TP}{TP+FP} \tag{1}$$

$$recall = \frac{TP}{TP+FN} \tag{2}$$

$$mAP=mean(AP1,\ AP2,\ldots,APn) \tag{3}$$

$$F1score= \frac{2TP}{2TP+FP+FN} \tag{4}$$

where TP, FP, and FN refer to true positives, false positives, and false negatives, respectively. The model achieved a mAP of 92.10% on the validation set and demonstrated high generalization under varied lighting and orientation conditions.

E. Deployment and Real-Time Inference

The trained model was exported and deployed on NVIDIA Jetson Nano for real-time execution. With an average inference time of under 100 milliseconds per frame, the system supports live monitoring while preserving user privacy. Edge deployment ensures privacy by avoiding cloud data transmission and enables offline operation.

4. Results and Discussion

The YOLOv8 model was trained on a custom dataset comprising over 19,000 images labeled as “fall,” “non-fall,” and “background”. The training process spanned 100 epochs with a learning rate of 0.001. The evaluation was conducted on a test set comprising 11% of the dataset. Tables 2, 3, and 4 present the performance metrics of YOLOv8 on the test set. The model achieved a mean Average Precision (mAP) of 90.97% and an overall accuracy of 95.36%.

Experimental results validate that the proposed YOLOv8-based approach particularly addresses significant issues of real-time fall detection in elderly care facilities. With an F1-score of 95%, the system indicated fair balance between precision (98%) and recall (93%), implying ongoing dependability in distinguishing fall events from fall-free activities. Those numbers much exceed those for older YOLOv5-based techniques shown in other studies [8-9], hence suggesting the advancement the design of YOLOv8 offers for object identification activities with complex human motions.

In practical uses where fall events could happen in rather varied and unstructured settings, that level of success is most important. The results of this study are reconcilable and demonstrate the model's great relevance in fall detection settings when compared to earlier research such as Khekan et al. and Nguyen et al., which also reported high mAP results using YOLOv8 [6],[8]. The model was evaluated with an entirely different dataset designed and customized for actual seniors' fall circumstances. By closely mimicking actual deployment circumstances unlike in previous studies that often used limited or false datasets, such targeted augmentation evolved the model toward robustness. The system ran inferences under 100 milliseconds per frame on the Jetson Nano platform, hence suggesting its potential for real-time deployment in edge settings in terms of computational convincingness. Such efficiency is essential for applications like home care and healthcare monitoring since significant contributors are latency and on-device processing constraints. Though the model produced human-interpretable visual displays with well-defined bounding boxes and tags, there were occasional misclassifications—most notably in partial occlusions and poor-lighting situations. Failures like this suggest potential paths for study in the diversity of datasets and data augmentation strategies, including tactics such as synthetic occlusion, contrast adjustments, and adversarial augmentation. This study now offers solid proof to confirm the real-time fall detection feasibility of YOLOv8, therefore offering both quick processing and outstanding accuracy. Future studies could investigate model pruning or quantization to further reduce inference time and/or the incorporation of the LSTM or Transformer module with temporal analysis for improving the identification of unusual motion patterns.

Table 2 The confusion matrix of YOLOv8 on the test set

	Predicted: Fall	Predicted: Non-Fall	Predicted: Background
Class: Fall	**0.93**	**0.04**	**0.03**
Class: Non-Fall	**0.02**	**0.96**	**0.02**
Class: Background	**0.01**	**0.02**	**0.97**

Table 3 The precision, recall and F1-score of YOLOv8 on the test set

	Precision	Recall	F1-score
Class: Fall	**0.98**	**0.93**	**0.95**
Class: Non-Fall	**0.94**	**0.95**	**0.94**
Class: Background	**0.94**	**0.97**	**0.95**

Table 4 The mAP and accuracy of YOLOv8 on the test set

Accuracy (%)	mAP (%)
95.36	**90.97**

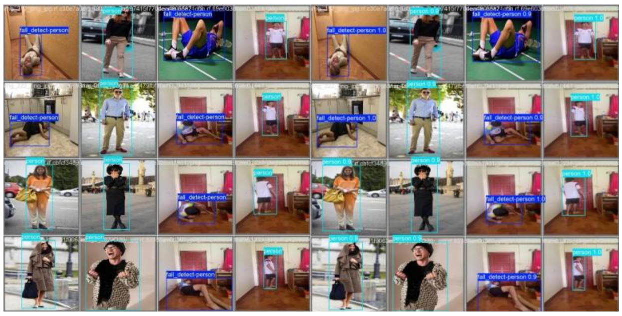

Fig. 5 Sample prediction outputs of YOLOv8 showing fall and non-fall cases from the dataset.

Fig. 6 Sample prediction outputs of YOLOv8 showing fall and non-fall cases from the webcam.

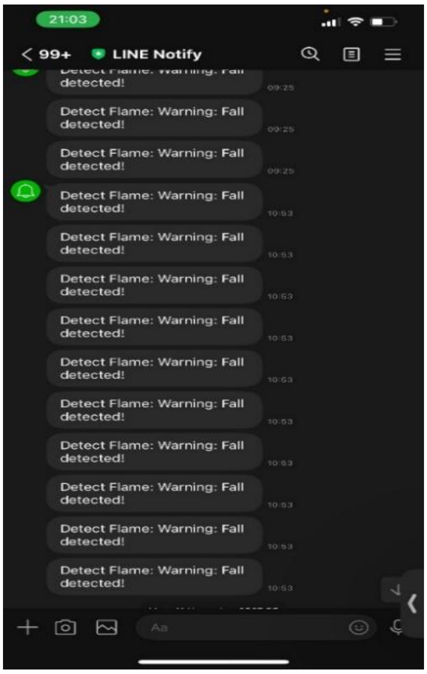

Fig. 7 Social media notification with Line Application.

5. Conclusions

This work proposes an extensive YOLOv8 vision system designed for fall detection in elderly living environments with a focus on accuracy, efficacy, and real-time implementation. Utilizing a specially created dataset consisting of more than 19,000 labeled images, the system showed high performance metrics with 98% precision and 93% recall for fall events. Results validate that YOLOv8 is very effective in the detection of fall incidents with diverse illumination, camera positions, and real-world situations. Besides better detection performance, the system was shown to be operationally viable on edge devices like NVIDIA Jetson Nano, enabling decentralized, privacy-preserving surveillance. Robust model generalization and resilience were further aided by the application of Roboflow to data augmentation and preprocessing. Additionally, the visual interpretation of predictions and smooth model convergence validated the architecture's suitability for real-time deployment. Traditional wearable- and vision-based detection models were outperformed by the approach regarding detection speed and contextual correctness. However, some degradation in performance under partial occlusion and low-contrast scenarios was noted. Models in the future could be explored with hybrid models that link vision data with inertial sensors or with depth data. Additionally, adding more real-world falling videos to the dataset, increasing model interpretability, and adding lightweight attention modules could enhance model performance. In conclusion, this study highlights the viability of YOLOv8 to act as a foundational technology in smart eldercare fall detection systems to promote more independent living and a sense of security among the elderly population.

Acknowledgements

This research project was financially supported by Mahasarakham University, Thailand.

References

[1] X. Fu and N. Angkawisittpan, "Detecting surface defects of heritage buildings based on deep learning," in *Journal of Intelligent Systems*, vol. 13, no. 1, 20230048, 2024, doi: 10.1515/jisys-2023-0048.

[2] S. Sonasang, S. Srisawat, R. Phromloungsri, W. Rattanangam and N. Angkawisittpan, "Liquid Level Measurement Using Sensors with Microstrip Parallel Coupled Lines," *2019 IEEE 2nd International Conference on Power and Energy Applications* (ICPEA2019), Singapore, 2019, pp. 106-109, doi: 10.1109/ICPEA.2019.8818540.

[3] S. Balasubramaniam, J. R. Park, T. Mistry, N. Angkawisittpan, A. Akyurtlu, T. Rao and R. Nagarajan, "Conformal passive sensors for wireless structural health monitoring," *Materials Research Society Symposium Proceedings*, Boston, Massachusetts, USA, 2009, vol. 1129, pp. 341-348, doi: 10.1557/PROC-1129-V04-19.

[4] M. Sopa and N. Angkawisittpan, "An Application of Cuckoo Search Algorithm for Series System with Cost and Multiple Choices Constraints," in *Procedia Computer Science (iEECON2016)*, Chiang Mai, Thailand, vol. 86, pp. 453-456, 2016, doi: 10.1016/j.procs.2016.05.079.

[5] N. Angkawisittpan and A. Siritaratiwat, "A Dual Frequency Monopole Antenna with Double Spurlines for PCS and Bluetooth Applications," in *Applied Computational Electromagnetics Society Journal*, vol. 31, no. 8, pp. 976-981, 2016.

[6] P. H. Ng, A. Mai and H. Nguyen, "Building an AI-powered IoT App for Fall Detection using YOLOv8 Approach," *Intelligence of Things: Technologies and Applications (ICIT 2023)*, Ho Chi Minh City, Vietnam, 2023, pp. 65-74, doi: 10.1007/978-3-031-46749-3_7.

[7] M. G. Ragab, S. J. Abdulkadir, A. Muneer, A. Alqushaibi, E. H. Sumiea, R. Qureshi, S. M. Al-Selwi and H. Alhussian, "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," in *IEEE Access*, vol. 12, pp. 57815-57836, 2024, doi: 10.1109/ACCESS.2024.3386826.

[8] A. R. Khekan, H. S. Aghdasi and P. Salehpour, "Fast and High-Precision Human Fall Detection Using Improved YOLOv8 Model," in *IEEE Access*, vol. 13, pp. 5271-5283, 2025, doi: 10.1109/ACCESS.2024.3470319.

[9] Y. J. Mao, A. Y. C. Tam, Q. T. K. Shea, Y. P. Zheng and J. C. W. Cheung, "eNightTrack: Restraint-Free Depth-Camera-Based Surveillance and Alarm System for Fall Prevention Using Deep Learning Tracking," in *Algorithms*, vol. 16, no. 10, 477, 2023, doi: 10.3390/a16100477.

[10] X. Huang, X. Li, L. Yuan, Z. Jiang, H. Jin, W. Wu, R. Cai, M. Zheng and H. Bai, "SDES-YOLO: A high-precision and lightweight model for fall detection in complex environments," in *Scientific Reports*, vol. 15, 2026, 2276, 2025, doi: 10.1038/s41598-025-86593-9.

[11] C. Y. Wang and F. S. Lin, "AI-driven privacy in elderly care: Developing a comprehensive solution for camera-based monitoring of older adults," in *Applied Sciences*, vol. 14, no. 10, 4150, 2024, doi: 10.3390/app14104150.

[12] J. Moore, Y. Celik, S. Stuart, P. McMeekin, R. Walker, V. Hetherington and A. Godfrey, "Using Video Technology and AI within Parkinson's Disease Free-Living Fall Risk Assessment," in *Sensors*, vol. 24, no. 15, 4914, 2024, doi: 10.3390/s24154914.

[13] A. T. Özdemir and B. Barshan, "Detecting Falls with Wearable Sensors Using Machine Learning Techniques," in *Sensors*, vol. 14, no. 6, pp. 10691-10708, 2014, doi: 10.3390/s140610691.

[14] J. Gutiérrez, V. Rodríguez and S. Martin, "Comprehensive Review of Vision-Based Fall Detection Systems," in *Sensors*, vol. 21, no. 3, 947, 2021, doi: 10.3390/s21030947.

[15] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 779-788, 2016, doi: 10.1109/CVPR.2016.91.

[16] N. Ladi, "Object Detection with a pre-trained Ultralytics YOLOv8 model," 2024, Available: https://www.ultralytics.com/blog/object-detection-with-a-pre-trained-ultralytics-yolov8-model

[17] F. X. Gaya-Morey, C. Manresa-Yee and J. M. Buades-Rubio, "Deep learning for computer vision based activity recognition and fall detection of the elderly: a systematic review," in *Applied Intelligence*, 54, pp. 8982–9007, 2024, doi: 10.1007/s10489-024-05645-1.

[18] A. Bustamante, L. M. Belmonte, R. Morales, A. Pereira and A. F. Caballero, "Bridging the Appearance Domain Gap in Elderly Posture Recognition with YOLOv9," in *Applied Sciences*, vol. 14, no. 21, 9695, 2024, doi: 10.3390/app14219695.

[19] Roboflow, "Data augmentation and preprocessing tools," Available: https://roboflow.com

[20] M. Lupión, V. González-Ruiz, J. F. Sanjuan and P. M. Ortigosa, "Privacy-aware fall detection and alert management in smart environments using multimodal device," in *Internet of Things*, vol. 30, 101526, 2025, doi: 10.1016/j.iot.2025.101526.

[21] Ultralytics, "Explore Ultralytics YOLOv8," 2023, Available: https://docs.ultralytics.com/models/yolov8/