

Article

Detecting Human Falls in Poor Lighting: Object Detection and Tracking Approach for Indoor Safety

Xing Zi, Kunal Chaturvedi, Ali Braytee, Jun Li and Mukesh Prasad * 

School of Computer Science, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney 2007, Australia

* Correspondence: mukesh.prasad@uts.edu.au

Abstract: Falls are one the leading causes of accidental death for all people, but the elderly are at particularly high risk. Falls are severe issue in the care of those elderly people who live alone and have limited access to health aides and skilled nursing care. Conventional vision-based systems for fall detection are prone to failure in conditions with low illumination. Therefore, an automated system that detects falls in low-light conditions has become an urgent need for protecting vulnerable people. This paper proposes a novel vision-based fall detection system that uses object tracking and image enhancement techniques. The proposed approach is divided into two parts. First, the captured frames are optimized using a dual illumination estimation algorithm. Next, a deep-learning-based tracking framework that includes detection by YOLOv7 and tracking by the Deep SORT algorithm is proposed to perform fall detection. On the Le2i fall and UR fall detection (URFD) datasets, we evaluate the proposed method and demonstrate the effectiveness of fall detection in dark night environments with obstacles.

Keywords: fall detection; YOLOv7; dual illumination estimation; CNN; Deep SORT



Citation: Zi, X.; Chaturvedi, K.; Braytee, A.; Li, J.; Prasad, M. Detecting Human Falls in Poor Lighting: Object Detection and Tracking Approach for Indoor Safety. *Electronics* **2023**, *12*, 1259. <https://doi.org/10.3390/electronics12051259>

Academic Editors: Yuji Iwahori, Aili Wang and Haibin Wu

Received: 23 December 2022

Revised: 28 February 2023

Accepted: 5 March 2023

Published: 6 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The World Health Organization states that falling has been ranked as the second leading cause of accidental death [1]. A fall is described as a rapid change from a normal state to a reclined or extended position of the whole body. It can be caused by discomfort or unsteadiness while standing [2]. Recent studies indicate that the death rate among the elderly due to falls is nearly three times higher than that of younger age groups [3,4]. Most older people, especially those living independently without a carer, fear severe injuries or even fatalities from falling because of delayed assistance. This demonstrates the importance of early warning and convenient management of falls. It is foreseeable that with practical and timely detection and warning mechanisms, the rate of severe injuries and even fatalities in falls, especially among the elderly, will be significantly reduced [5,6].

Most of the current fall detection and warning systems employ sensors such as barometers, accelerometers, gyroscopes, and inertial sensors [7,8]. These sensors are usually used in wearable devices, including a smartwatch, shoes, and necklace, to monitor users' body parameters to detect a fall. However, they have the drawback of requiring the devices to be put on the user's body, which makes it uncomfortable and infeasible to be worn constantly. Furthermore, sensor-based fall detection systems are expensive and have data privacy and security concerns.

With the advent of vision-based systems, computer vision has witnessed a growing trend in its applications. Notably, significant amounts of research are focused on vision-based systems to monitor any anomalies in body movement for fall detection [5,9–13]. Such methods can be divided into two categories based on the data used, namely, (1) RGB-based detection and (2) depth-based detection [14]. In this paper, we perform fall detection on an RGB-based dataset using a deep learning framework. Deep learning networks, in

particular, convolutional neural networks (CNNs), have emerged as a popular computation framework in computer vision [15–19]. Many deeper and more complicated networks are being developed to enable CNNs to deliver near-human accuracy in many applications, such as classification [20,21], detection [22–24], and segmentation [25,26]. A CNN-based deep learning approach on RGB-based datasets has been established as state-of-the-art for fall detection [12,13,27,28]. Although these methods enable flexible automatic feature extraction from images, they use only one keyframe, which may not be sufficient to adequately classify falls in videos. For example, a fall can be summarized using distinctive keyframes based on variations in position from standing or sitting to lying down [29]. Specifying the location of such keyframes for falls would be too tedious and prone to errors. Therefore, there is a need for a frame-to-frame object tracking mechanism to characterize the current state of the body.

This paper presents a novel deep-learning-based tracking-by-detection approach to detect falls in videos. Tracking-by-detection methods consist of: (1) an object detection algorithm that gives the object's bounding box coordinates for each frame, and (2) a tracking algorithm that decides if the newly detected object can be connected to the predicted position of existing tracks. The quality of both the detection algorithm and the data association algorithm heavily influences the effectiveness of object tracking [30]. Specifically, the proposed method integrates Simple Online and Realtime Tracking with a Deep association metric (Deep SORT) algorithm [31] with YOLOv7 [32] for fall detection. Performance is highly dependent on the light in the environment. In the real-world setting, fall detection can be challenging due to light changes and poor illumination. To improve the overall efficiency, we combine the deep-learning-based tracking-by-detection with an image enhancement algorithm. Our proposed hybrid method significantly reduces the false positive rates in the fall detection task.

The main contributions of this paper are as follows:

- A novel deep-learning-based approach for vision-based fall detection is proposed, integrating YOLOv7 for object detection and the Deep SORT algorithm for tracking and trajectory analysis.
- The proposed method incorporates dual illumination estimation, utilizing a Retinex-based image enhancement algorithm, to effectively tackle the issue of inconsistent lighting conditions and exposure levels.
- The effectiveness and superiority of the proposed approach over current state-of-the-art methods are extensively demonstrated, offering a robust solution for vision-based fall detection.

2. Related Works

Elderly care requires continuous monitoring by health care staff, which is costly, time-consuming, and even considered an intrusive process that disrupts individuals' routines and privacy [33]. Thus, an automatic fall detection system is required in healthcare services to provide cost-effective, efficient, and timely care to the person affected. Researchers have exhaustively explored many methods of fall detection to reduce the number of injuries caused by falls. These systems are broadly classified based on the device used to detect the fall, including wearable sensors, environment sensors, and vision-based sensors. In wearable devices, one or more sensors are put on various body parts of the person to detect and identify falls from activities of daily living (ADL). It detects any rapid increase in negative acceleration when the individual suddenly changes its position from standing to lying down [34,35]. For environment-based sensors, signals from acoustic [36], infrared [37], vibration [38] and pressure [39] are collected and analyzed to detect falls. These devices are less intrusive than wearable sensor-based devices and require less interaction with the individual, which reduces privacy and security concerns. However, environment-based sensors are prone to false alarms due to changes in the external environment. For example, different noises or other falling items in the room can impact the sensors' performance [40].

In recent years, the development of vision-based sensors has led to several computer vision-based approaches being proposed for fall detection. Vision-based approaches extract data from sensors which can be RGB-based, depth-based, thermal-based, or even a combination of these. The extracted data is then utilized by computer vision algorithms to detect any unusual changes related to the body trajectories, postures, or shapes to identify falls. The most common approach in computer vision is based on hand-crafted feature extraction [9,11]. Sehairi et al. [11] use a finite state machine (FSM) on a human silhouette extracted by background subtraction technique to determine if a fall happened. Albawendi et al. [9] perform fall detection based on hand-crafted features extracted from projection histogram, motion information, and human shape variation. Recently, deep learning methods have gained popularity over hand-crafted feature-based methods due to their ability to extract important features in high dimensional data independently.

Deep learning has been successfully applied in fall detection, demonstrating high performance. Lu et al. [12] incorporate a combination of a three-dimensional convolutional neural network (3D CNN) and a long short-term memory (LSTM) in their fall prevention method. Han et al. [13] improve the fall detection speed using a two-stream approach with the lightweight MobileVGG. Recently, human skeleton joint coordinates extracted from RGB data have also been used to detect falls [10,28]. The authors [10] propose a spatiotemporal network based on CNNs, GRUs, and fully connected layers to classify falls and ADL. The existing methods for fall detection involve either a combination of various tasks, such as foreground and background separation or human skeleton joint coordinates. In contrast, our proposed method uses spatio-temporal features from RGB data using a deep-learning-based tracking-by-detection approach.

A considerable amount of research has been conducted in the field of tracking-by-detection approach [41–44]. These methods use spatiotemporal data to extract relevant features and then use them as inputs to distinguish between the identified objects. Finally, a tracker follows the object during the video flow. Despite the promising results obtained by tracking-by-detection methods, these methods can still suffer from colour bias or complex underexposed conditions [45]. Therefore, this paper presents a two-stage fall detection framework. First, we perform exposure correction of the video frames using a dual illumination estimation method [46]. The method adaptively corrects the frames based on exposure conditions, including overexposed, underexposed, and partially over- and underexposed conditions. Second, we use the YOLOv7 algorithm [26] to detect the subject's activity and combine the appearance information from object-detection CNNs with a fast, improved Deep SORT [25] based multiple object tracking (MOT) method to extract motion features and compute trajectories.

3. Methodology

3.1. Experimental Data

In this paper, we use two publicly available datasets, i.e., the Le2i Fall Detection dataset [47] and the UR Fall Detection dataset (URFD) [48]. The Le2i dataset is composed of falls captured by narrow angle cameras. The dataset includes videos of various actors falling and not falling in different illumination scenarios. The videos vividly depict real-world scenes of people falling while performing their daily activities. These scenes are captured in a variety of environments, including homes, workplaces, and pantries; it makes for more accurately simulated realistic video sequences of falls during everyday activities, as shown in Figure 1. The UR Fall Detection dataset [48] contains 60 videos taken by two cameras placed at different angles. Figure 1a,b shows some example video frames from the Le2i and URFD dataset, respectively.

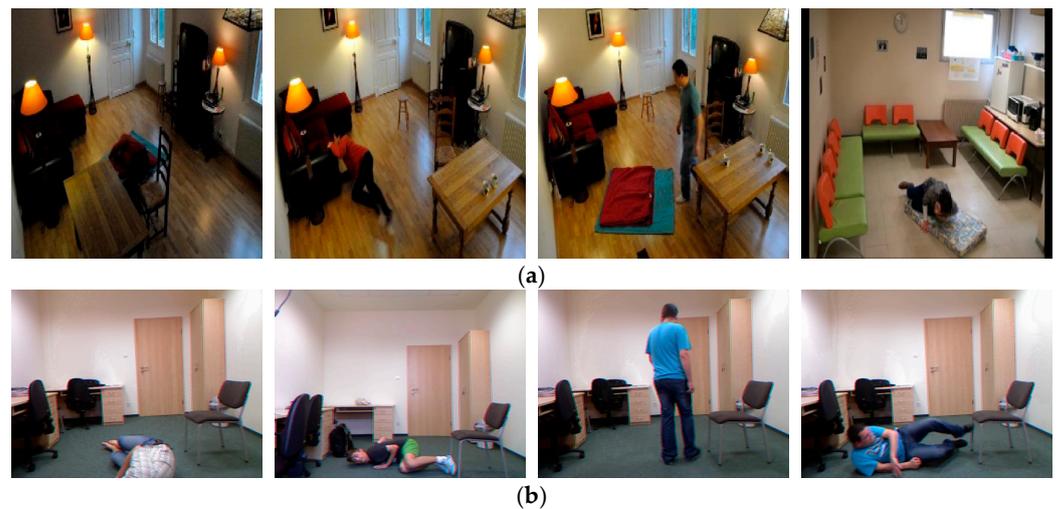


Figure 1. Some example video frames from the publicly available datasets (a) Le2i and (b) URFD dataset.

3.2. Annotation

Our annotate method detects the person class in each frame using a pre-trained YOLOv5 model, and then stores the bounding box coordinates in a CSV table. In the next step, we examine each frame and correct the annotation results for fall and upright, as shown in Table 1. Figure 2 shows the detection results of YOLOv5 and Manual post-correction.

Table 1. Example bounding box coordinates and annotations.

| File Name | Xmin | Ymin | Xmax | Ymax | YOLOv5 Annotations | Manually Corrected Annotations |
|------------|------|------|------|------|--------------------|--------------------------------|
| 000018.jpg | 120 | 101 | 187 | 200 | person | upright |
| 000019.jpg | 109 | 113 | 194 | 198 | person | upright |
| 000020.jpg | 108 | 116 | 195 | 197 | person | fall |
| 000021.jpg | 107 | 125 | 198 | 200 | person | fall |
| 000022.jpg | 107 | 125 | 198 | 200 | person | fall |
| 000023.jpg | 107 | 125 | 198 | 200 | person | fall |

3.3. Proposed Framework

The proposed framework is shown in Figure 3, which employs a two-stage framework to detect falls in low-light indoor environments. The object detection model is used as the main component of the fall detection system to further elaborate on the proposed technique. In the first stage, the captured video frames are pre-processed to improve the visual quality of the footage. This is accomplished using a dual illumination estimation-based exposure correction technique field [40]. The goal of this step is to adjust the brightness, contrast, and detail of the video frames so that later stages can detect falls more accurately. On the enhanced video frames, the second stage implements a deep-learning-based approach for person tracking and detection. This is accomplished using a technique known as people tracking by detection, which involves using an object detection model to track the movements of people in video footage. The algorithm can detect when a person falls by analysing the motion patterns of the tracked objects. To detect falls in low-light indoor environments, the proposed fall detection system utilizes a sophisticated combination of exposure correction and deep-learning-based people tracking by detection algorithms. By improving the visual quality of the video footage and leveraging advanced object detection techniques, the system can improve the accuracy and reliability of fall detection.

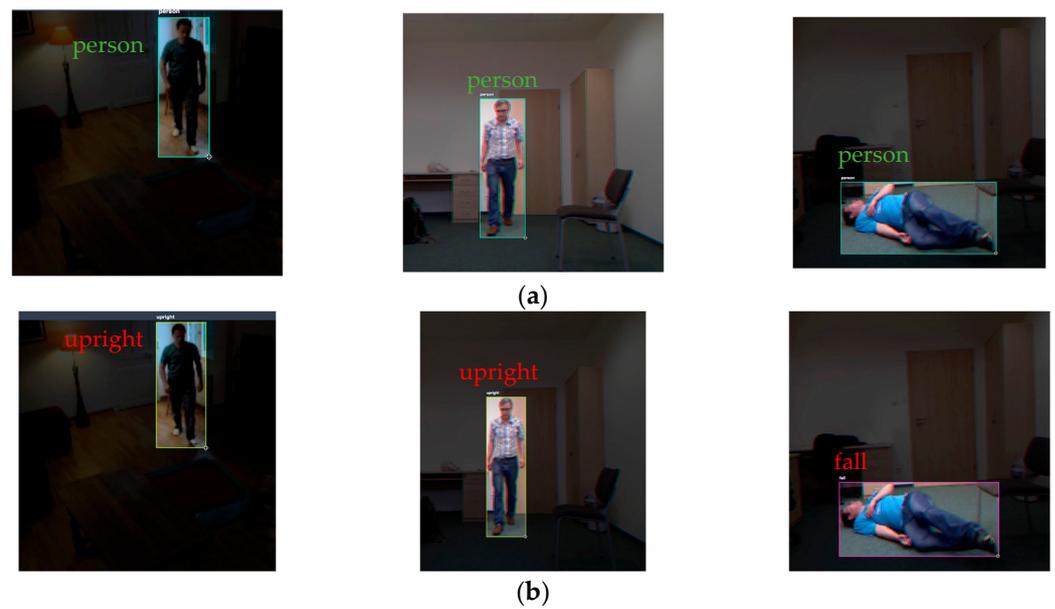


Figure 2. The detection results of YOLOv5 and Manual post-correction. (a) YOLOv5 pre-trained model. (b) Manual post-correction.

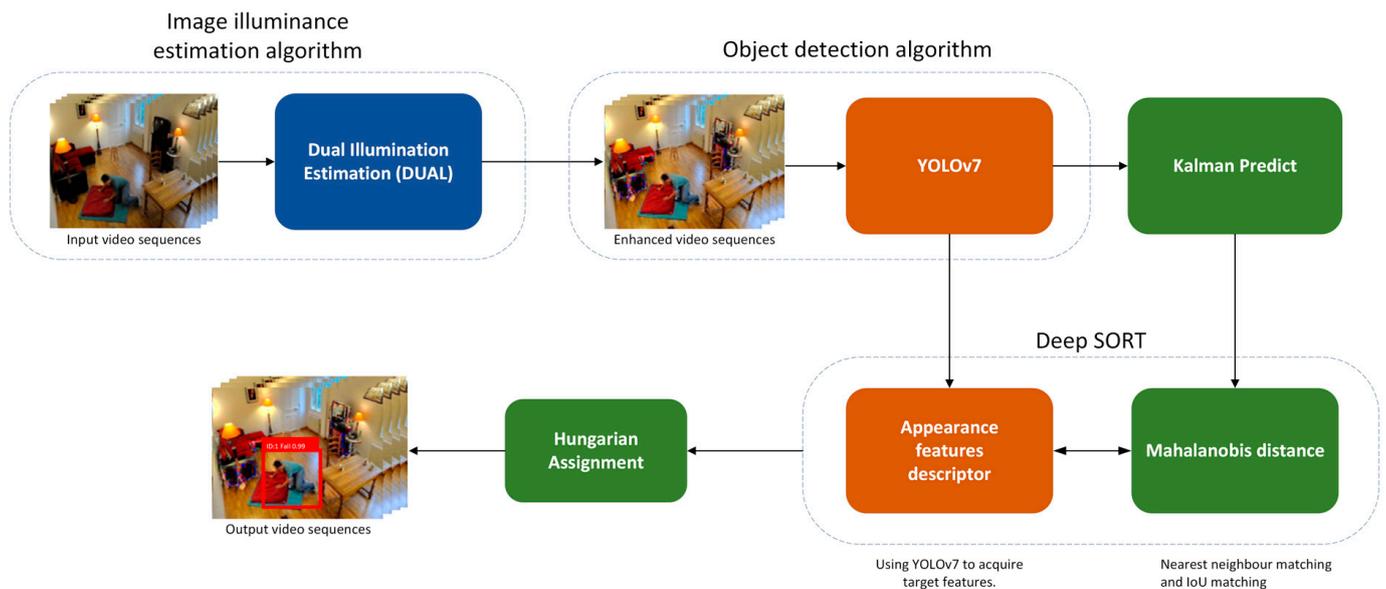


Figure 3. Schematic representation of the proposed framework.

3.3.1. Object Detection

RetinaNet was proposed by Lin et al. [49] in 2018, which is based on three parts: the ResNet backbone, the feature pyramid network (FPN), and the object classification and bounding box regression subnetwork. Deeper neural networks capture more information, and the ResNet backbone uses identity mappings to ensure that performance does not degrade as the network’s depth increases. The FPN brings the benefits of multi-scale fusion by capturing coarse to fine semantic information, allowing for more feature details to be obtained for small object detection tasks and thus improving accuracy. The RetinaNet model runs the classification and regression sub-networks in parallel before attaching them to the FPN module. A sigmoid activation function is used to predict the classification at each location by the classification module. The regression sub-network outputs coordinate values of shape 4 via a fully connected layer. In object detection tasks [50–52], an imbalance between positive and negative samples is a major cause of classification difficulties. The

Focal Loss method solves this problem by adjusting the weights of the difficult and easy-to-classify samples.

Redmon et al. [53] proposed You Only Look Once (YOLO), which combines object detection and regression into a single task. The input image is segmented into multiple grid cells to accomplish this. The case where the object's centre point is located on the grid is seen as the anchor box responsible for detecting the object class. The YOLOv5 [54] and YOLOv7 [32] models used in this paper are optimized from YOLOv3 [55]. The reparameterization convolution method is used in YOLOv7. This method fuses the convolutional layers and batch normalization into a single convolutional module, which greatly improving the model inference speed.

The main advantage of one-stage models used in the proposed method is the high computational efficiency and better inference speeds over two-stage models. In real-time, vision-based systems for fall detection need to be faster and more efficient. Furthermore, the one-stage models address the class imbalance problem in fall detection to improve the detection accuracy.

3.3.2. Object Tracking

The Simple Online and Realtime Tracking (SORT) algorithm [30], proposed in 2016, is a straightforward and quick method for multiple object tracking (MOT). The correlation between the preceding and following frames is processed using Kalman filtering and then measured using the Hungarian algorithm [56]. The Deep SORT algorithm is an extension of the SORT algorithm. The Deep SORT algorithm's neural network weights training on the pedestrian dataset as best suited for human fall detection. By matching the extracted features to the object's nearest neighbors, it is well adapted for improving object tracking [57] and detection in obstacle-type environments [31]. The feature extraction network can significantly improve the Deep SORT algorithm's robustness to obstacles and object loss.

Using object-tracking methods in conjunction with fall detection systems is a promising approach, especially given the limitations of current visual method studies, which frequently overlook the possibility of increased false-negative detection rates. Deep SORT, fortunately, can help address this issue by predicting the potential location of the next frame and calculating correlations, providing an early warning if a person becomes obscured from view. It is possible to improve the accuracy and reliability of fall detection systems by incorporating this approach, thereby improving the safety and well-being of individuals at risk of falling. It is notable to combine object tracking methods with fall detection systems. Most current visual method studies do not account for the potential increase in false-negative detection rates. Deep SORT provides early warning if a person is obscured by predicting the potential location of the next frame and calculating correlations.

3.3.3. Dual Illumination Estimation (DUAL)

Zhang et al. [40] proposed the dual illumination estimation algorithm for dark light image enhancement. This method is based on the core concept of Retinex. The colour of an object is determined by its ability to reflect light of three different wavelengths. The support for this concept is based on colour constancy. Dual illumination enhancement demonstrates that uneven lighting conditions do not affect the colour of the object. When there is underexposure or overexposure, the colour of the object does not change. The illumination is estimated in both the forward and reverse directions to recover a properly exposed image.

The objective of this project is to create the image I by using the image light mapping method L to multiply each pixel value in the existing image I' . When an overexposed image is reversed, it produces an underexposed image. This can be interpreted as locating the

overexposed portion of the current image by entering a forward and reverse image. The generated image I is defined in Equation (1) below,

$$I = I' \times L \quad (1)$$

Here, I' is calculated according to Equation (2), where $I_{inv} = 1 - I$ is the formula for the inverted image. At this point, the illuminance map L_{inv} for the current state is estimated. As a result of calculating the underexposure correction, the image I_{inv} can be obtained.

$$I' = 1 - (I_{inv} \times I_{inv}^{-1}) \quad (2)$$

$$L'_p = \max I_p^c, \forall c \in \{r, g, b\} \quad (3)$$

where c is the colour channel; p is the pixel and I_p^c is expressed as the colour channel c in pixel point p . Before estimating the illumination of the image, it needs to extract the maximum channel value for each pixel in the RGB three-channel image. It is possible to compose an initialized illuminance L'_p by iteratively obtaining each maximum value as shown in Equation in (3). The desired illumination map L is obtained using the objective defined in Equation (4).

$$\operatorname{argmin}_L \sum_p \left((L_p - L'_p)^2 + \lambda (w_{x,y} (\partial_x L)_p^2 + w_{y,p} (\partial_y L)_p^2) \right) \quad (4)$$

In real-world conditions, the performance of deep learning models suffers due to videos captured under suboptimal lighting conditions caused by dim or uneven light. For example, in an indoor setting, a video captured at night has dark or under-exposed regions due to insufficient illumination. This may result in the proposed systems failing to monitor individuals' activities, resulting in high false-negative rates. Figure 4 shows the results from Le2i FDD obtained by using a dual estimation algorithm.

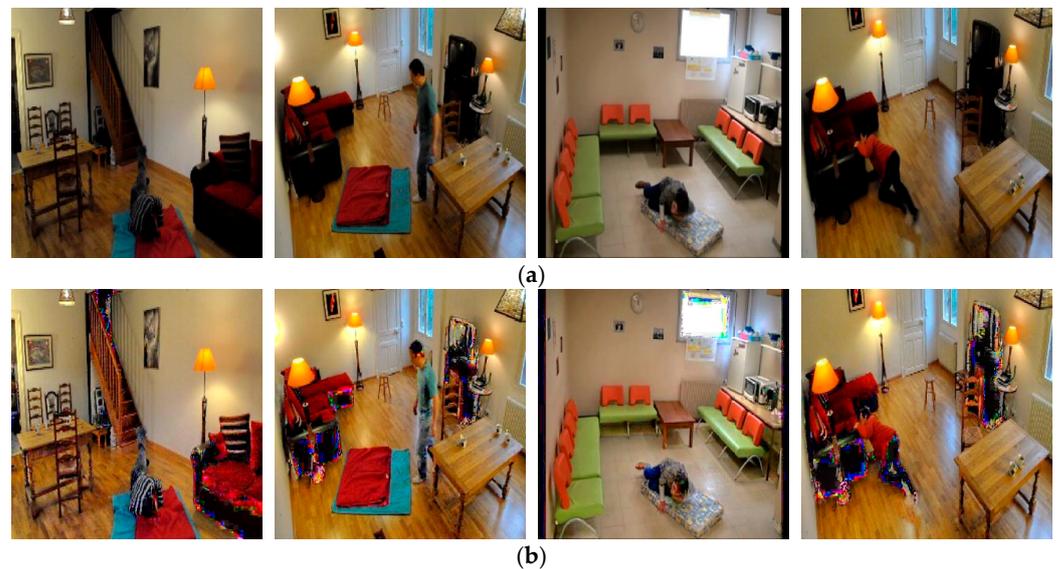


Figure 4. Examples of images from Le2i Fall Detection dataset. (a) The original images. (b) The images after processing by the DUAL illumination estimation.

4. Experimental Evaluation

4.1. Experimental Setup

This paper uses RetinaNet, YOLOv5, and YOLOv7 object detection algorithms based on a PyTorch framework to evaluate the proposed approach. Experimental results are conducted on a Tesla P100 GPU with 16,280 Mb video memory. The initial learning rate

for YOLO series models is set to 0.1, gradually decreasing to 0.01 during training, whereas RetinaNet uses a learning rate of 0.00025. As shown in Table 2, the batch size for YOLOv7 and RetinaNet is set to 8, and for YOLOv5, the batch size is set to 32. All the models are trained for 100 epochs each. For the YOLOv7/w DUAL, we use the same model parameters as YOLOv7. McNemar’s significance test is used to statistically validate the performance of the object detection models. Next, the Deep SORT tracking algorithm is applied to the best performing object detection model. Specifically, we use the performance metrics such as accuracy, 0.5 mAP, and precision score for evaluation.

Table 2. Details of the parameters of the model trained.

| Parameters | The Proposed Method | YOLOv7 | YOLOv5 | RetinaNet |
|------------------|---------------------|----------|------------|-----------|
| Learning Rate | 0.01–0.1 | 0.01–0.1 | 0.001–0.01 | 0.00025 |
| Batch Size | 8 | 8 | 32 | 8 |
| Epochs | 100 | 100 | 100 | 100 |
| Data Enhancement | Yes | - | - | - |

4.2. Results

Table 3 shows the results of McNemar’s significance test (p values indicate the results obtained by the best performing model) on the Le2i Fall Detection dataset where YOLOv7 is statistically different than the results by RetinaNet and YOLOv5. Table 4 shows the results from different object detection algorithms on the Le2i Fall Detection dataset (Le2i FDD). YOLOv7 outperforms YOLOv5 and the RetinaNet with an accuracy of 90.5%, and mAP of 0.966. Here, the RetinaNet model performs the worst with an accuracy of 59.02%, and mAP of 0.842. Then, we apply Deep SORT to the best performing object detection model, i.e., YOLOv7, which is trained and tested on enhanced video frames obtained by performing dual illumination estimation. Here, the proposed method (YOLOv7 + Deep SORT/w DUAL) gives an accuracy of 94.5% and mAP of 0.986, which is a significant improvement over the object detection methods. Moreover, YOLOv7 + Deep SORT/w DUAL is compared to the existing state-of-the-art (SOTA) methods on Le2i FDD. Table 4 shows that the proposed approach achieves the highest accuracy, outperforming the SOTA methods by Poonsri et al. [58] and Chamle et al. [59], which have the fall accuracy of 91.38% and 79.41%, respectively. Poonsri et al. [58] and Chamle et al. [59] have used background-based subtraction techniques. However, due to insufficient illumination conditions in videos of Le2i FDD, background subtraction techniques used in these methods incorrectly extract other objects as a human silhouette. This results in high false-positive rates for fall detection. The proposed method is compared to the traditional method proposed by Poonsri et al. [48] and Chamle et al. [49] based on their annotated images and the result presented. Furthermore, the visual results using the YOLOv7 and the proposed method are shown in Figure 5. When the environment light is not insufficiently illuminated, as shown in the first row of Figure 6, the YOLOv7 + Deep SORT misclassifies the fall as upright. However, the proposed method correctly classifies it as a fall, as shown in the second row of Figure 5.

Table 3. p -Values of the McNemar’s significance test on Le2i dataset. Here, $p < 0.05$ is statistically significant.

| Method | RetinaNet | YOLOv5 | YOLOv7 |
|--------|-----------|--------|--------|
| YOLOv7 | <0.001 | 0.0095 | 1 |

Table 4. Performance of the different state-of-the-art methods on Le2i dataset.

| Method | Accuracy (%) | 0.5 mAP | Precision of Fall |
|---------------------|--------------|---------|-------------------|
| Poonsri et al. [58] | 91.38 | - | 0.886 |
| Chamle et al. [59] | 79.31 | - | 0.794 |
| RetinaNet [49] | 59.02 | 0.842 | 0.775 |
| YOLOv5 [54] | 86.0 | 0.947 | 0.896 |
| YOLOv7 [32] | 90.5 | 0.966 | 0.935 |
| The proposed method | 94.5 | 0.986 | 0.970 |

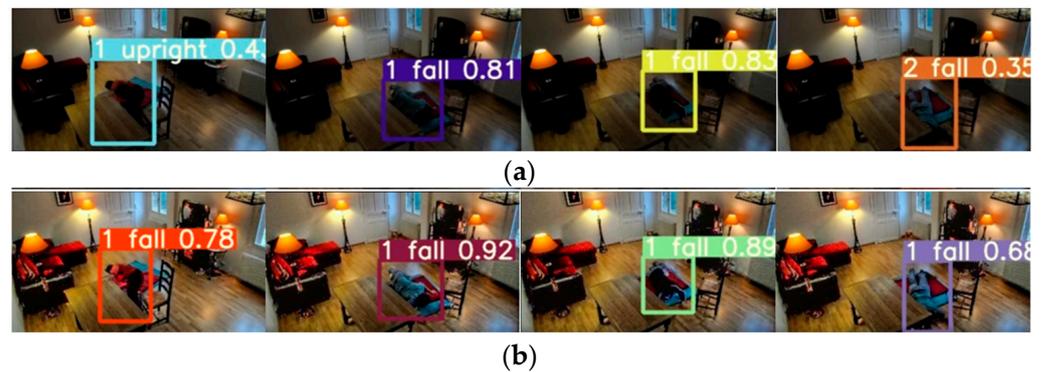


Figure 5. Visual results of object tracking on Le2i Fall Detection dataset. (a) YOLOv7 + Deep SORT method. (b) The proposed method.

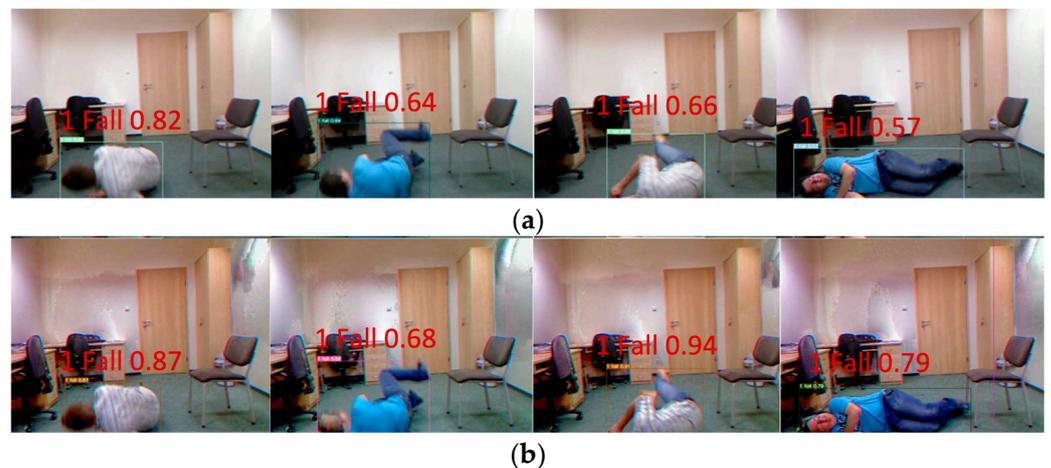


Figure 6. Visual results of object tracking on UR Fall Detection dataset (a) YOLOv7 + Deep SORT method (b) the proposed method.

Table 5 shows the results of McNemar’s significance test on UR-Fall dataset where the performance of YOLOv7 is statistically significant ($p < 0.05$) over RetinaNet, but it is not statistically different from that of YOLOv5. Table 6 shows the different object detection algorithms on the UR Fall detection dataset. YOLO models achieve high results on metrics such as accuracy and mAP, with YOLOv7 performing the best among all the models. However, RetinaNet did not perform well on the UR-Fall dataset, reporting the lowest accuracy of 40.9% as well as the lowest mAP of 0.464. Similarly, to the experiments on Le2i FDD, YOLOv7 is integrated with the Deep SORT tracking algorithm to report the results on the UR-Fall dataset. The proposed method outperforms all the methods in terms of accuracy, mAP, and precision. As shown in the second row of Figure 6, the visual results on the UR-Fall dataset using the proposed method highlights the improved performance on the enhanced video frames as compared to the YOLOv7 + Deep SORT.

Table 5. *p*-Values of the McNemar’s significance test on UR-Fall dataset.

| Method | RetinaNet | YOLOv5 | YOLOv7 |
|--------|-----------|--------|--------|
| YOLOv7 | <0.001 | 0.085 | 1 |

Table 6. Testing Performance of different models in UR Fall detection dataset.

| Method | Accuracy (%) | 0.5 mAP | Precision of Fall |
|---------------------|--------------|---------|-------------------|
| RetinaNet [49] | 40.9 | 0.464 | 0.818 |
| YOLOv5 [54] | 89.8 | 0.925 | 0.881 |
| YOLOv7 [32] | 92.4 | 0.944 | 0.893 |
| The proposed method | 93.2 | 0.960 | 0.920 |

5. Conclusions

This paper proposes a vision-based fall detection system with an improved deep-learning-based tracking-by-detection method. The proposed method integrates dual illumination estimation to the YOLOv7 + Deep SORT tracking algorithm to enhance fall detection performance under suboptimal lighting conditions caused by dim or uneven light. The proposed method also incorporates exposure correction for fall detection in videos. The performance of the proposed method is validated on two fall detection datasets, namely, Le2i FDD and UR-Fall datasets. For future experiments, we aim to implement a self-learning framework that automatically adapts to false alarms and adds correct results to help the current fall detection systems to improve their performance.

Author Contributions: Conceptualization, X.Z. and M.P.; methodology, X.Z. and K.C.; software, X.Z. and K.C.; validation, X.Z., K.C., A.B., J.L. and M.P.; formal analysis, X.Z., K.C. and A.B.; investigation, X.Z. and K.C.; resources, A.B., J.L. and M.P.; data curation, X.Z. and K.C.; writing—original draft preparation, X.Z.; writing—review and editing, X.Z., K.C., A.B., J.L. and M.P.; visualization, K.C., A.B. and J.L.; supervision, J.L. and M.P.; project administration, A.B. and M.P.; funding acquisition, J.L. and M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- World Health Organization (WHO). Falls. 2021. Available online: <https://www.who.int/news-room/fact-sheets/detail/falls> (accessed on 17 July 2022).
- Alam, E.; Sufian, A.; Dutta, P.; Leo, M. Vision-based human fall detection systems using deep learning: A review. *Comput. Biol. Med.* **2022**, *146*, 105626. [CrossRef] [PubMed]
- Burns, E.; Kakara, R. Deaths from Falls Among Persons Aged ≥ 65 Years—United States, 2007–2016. *MMWR. Morb. Mortal. Wkly. Rep.* **2018**, *67*, 509–514. [CrossRef]
- Kelsey, J.L.; Procter-Gray, E.; Hannan, M.T.; Li, W. Heterogeneity of Falls Among Older Adults: Implications for Public Health Prevention. *Am. J. Public Health* **2012**, *102*, 2149–2156. [CrossRef]
- Vishnu, C.; Datla, R.; Roy, D.; Babu, S.; Mohan, C.K. Human Fall Detection in Surveillance Videos Using Fall Motion Vector Modeling. *IEEE Sensors J.* **2021**, *21*, 17162–17170. [CrossRef]
- Mubashir, M.; Shao, L.; Seed, L. A survey on fall detection: Principles and approaches. *Neurocomputing* **2013**, *100*, 144–152. [CrossRef]
- Saleh, M.; Jeannes, R.L.B. Elderly Fall Detection Using Wearable Sensors: A Low Cost Highly Accurate Algorithm. *IEEE Sensors J.* **2019**, *19*, 3156–3164. [CrossRef]
- Wang, X.; Ellul, J.; Azzopardi, G. Elderly Fall Detection Systems: A Literature Survey. *Front. Robot. AI* **2020**, *7*. [CrossRef]
- Albawendi, S.; Lotfi, A.; Powell, H.; Appiah, K. Video Based Fall Detection using Features of Motion, Shape and Histogram. In Proceedings of the 11th Pervasive Technologies Related to Assistive Environments Conference, Corfu, Greece, 26–29 June 2018; pp. 529–536. [CrossRef]
- Yadav, S.K.; Luthra, A.; Tiwari, K.; Pandey, H.M.; Akbar, S.A. ARFDNet: An efficient activity recognition & fall detection system using latent feature pooling. *Knowl. Based Syst.* **2022**, *239*, 107948. [CrossRef]

11. Sehairi, K.; Chouireb, F.; Meunier, J. Elderly fall detection system based on multiple shape features and motion analysis. In Proceedings of the 2018 IEEE International Conference on Intelligent Systems and Computer Vision (ISCV), Fez, Morocco, 2–4 April 2018; pp. 1–8. [CrossRef]
12. Lu, N.; Wu, Y.; Feng, L.; Song, J. Deep Learning for Fall Detection: Three-Dimensional CNN Combined with LSTM on Video Kinematic Data. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 314–323. [CrossRef]
13. Han, Q.; Zhao, H.; Min, W.; Cui, H.; Zhou, X.; Zuo, K.; Liu, R. A Two-Stream Approach to Fall Detection with MobileVGG. *IEEE Access* **2020**, *8*, 17556–17566. [CrossRef]
14. Khraief, C.; Benzarti, F.; Amiri, H. Elderly fall detection based on multi-stream deep convolutional networks. *Multimed. Tools Appl.* **2020**, *79*, 19537–19560. [CrossRef]
15. Li, J.; Han, L.; Zhang, C.; Li, Q.; Liu, Z. Spherical Convolution Empowered Viewport Prediction in 360 Video Multicast with Limited FoV Feedback. *ACM Trans. Multimed. Comput. Commun. Appl.* **2023**, *19*, 1–23. [CrossRef]
16. Feng, Q.; Feng, Z.; Su, X. Design and Simulation of Human Resource Allocation Model Based on Double-Cycle Neural Network. *Comput. Intell. Neurosci.* **2021**, *2021*, 7149631. [CrossRef] [PubMed]
17. Liu, H.; Liu, M.; Li, D.; Zheng, W.; Yin, L.; Wang, R. Recent Advances in Pulse-Coupled Neural Networks with Applications in Image Processing. *Electronics* **2022**, *11*, 3264. [CrossRef]
18. Qin, X.; Liu, Z.; Liu, Y.; Liu, S.; Yang, B.; Yin, L.; Liu, M.; Zheng, W. User OCEAN Personality Model Construction Method Using a BP Neural Network. *Electronics* **2022**, *11*, 3022. [CrossRef]
19. Lu, H.; Zhu, Y.; Yin, M.; Yin, G.; Xie, L. Multimodal Fusion Convolutional Neural Network with Cross-Attention Mechanism for Internal Defect Detection of Magnetic Tile. *IEEE Access* **2022**, *10*, 60876–60886. [CrossRef]
20. Zhou, W.; Wang, H.; Wan, Z. Ore Image Classification Based on Improved CNN. *Comput. Electr. Eng.* **2022**, *99*, 107819. [CrossRef]
21. Huang, C.-Q.; Jiang, F.; Huang, Q.-H.; Wang, X.-Z.; Han, Z.-M.; Huang, W.-Y. Dual-Graph Attention Convolution Network for 3-D Point Cloud Classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, 1–13. [CrossRef]
22. Xu, S.; He, Q.; Tao, S.; Chen, H.; Chai, Y.; Zheng, W. Pig Face Recognition Based on Trapezoid Normalized Pixel Difference Feature and Trimmed Mean Attention Mechanism. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3500713. [CrossRef]
23. Zhou, X.; Zhang, L. SA-FPN: An effective feature pyramid network for crowded human detection. *Appl. Intell.* **2022**, *52*, 12556–12568. [CrossRef]
24. Shi, Y.; Xu, X.; Xi, J.; Hu, X.; Hu, D.; Xu, K. Learning to Detect 3D Symmetry from Single-View RGB-D Images With Weak Supervision. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, 1–15. [CrossRef] [PubMed]
25. Yang, D.; Zhu, T.; Wang, S.; Wang, S.; Xiong, Z. LFRSNet: A robust light field semantic segmentation network combining contextual and geometric features. *Front. Environ. Sci.* **2022**, *10*, 1443. [CrossRef]
26. Sheng, H.; Cong, R.; Yang, D.; Chen, R.; Wang, S.; Cui, Z. UrbanLF: A Comprehensive Light Field Dataset for Semantic Segmentation of Urban Scenes. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 7880–7893. [CrossRef]
27. Galvão, Y.M.; Ferreira, J.; Albuquerque, V.A.; Barros, P.; Fernandes, B.J. A multimodal approach using deep learning for fall detection. *Expert Syst. Appl.* **2020**, *168*, 114226. [CrossRef]
28. Ramirez, H.; Velastin, S.A.; Meza, I.; Fabregas, E.; Makris, D.; Farias, G. Fall Detection and Activity Recognition Using Human Skeleton Features. *IEEE Access* **2021**, *9*, 33532–33542. [CrossRef]
29. Cheng, S.; Liu, J.; Li, Z.; Zhang, P.; Chen, J.; Yang, H. 3D error calibration of spatial spots based on dual position-sensitive detectors. *Appl. Opt.* **2023**, *62*, 933–943. [CrossRef] [PubMed]
30. Bewley, A.; Ge, Z.; Ott, L.; Ramos, F.; Upcroft, B. Simple online and realtime tracking. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 3464–3468. [CrossRef]
31. Wojke, N.; Bewley, A.; Paulus, D. Simple online and realtime tracking with a deep association metric. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 3645–3649. [CrossRef]
32. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. 2022. Available online: <http://arxiv.org/abs/2207.02696> (accessed on 14 December 2022).
33. Abbate, S.; Avvenuti, M.; Bonatesta, F.; Cola, G.; Corsini, P.; Vecchio, A. A smartphone-based fall detection system. *Pervasive Mob. Comput.* **2012**, *8*, 883–899. [CrossRef]
34. Palmerini, L.; Klenk, J.; Becker, C.; Chiari, L. Accelerometer-Based Fall Detection Using Machine Learning: Training and Testing on Real-World Falls. *Sensors* **2020**, *20*, 6479. [CrossRef] [PubMed]
35. Bagalà, F.; Becker, C.; Cappello, A.; Chiari, L.; Aminian, K.; Hausdorff, J.M.; Zijlstra, W.; Klenk, J. Evaluation of Accelerometer-Based Fall Detection Algorithms on Real-World Falls. *PLoS ONE* **2012**, *7*, e37062. [CrossRef] [PubMed]
36. Irtaza, A.; Adnan, S.M.; Aziz, S.; Javed, A.; Ullah, M.O.; Mahmood, M.T. A framework for fall detection of elderly people by analyzing environmental sounds through acoustic local ternary patterns. In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2017; pp. 1558–1563. [CrossRef]
37. Fan, X.; Zhang, H.; Leung, C.; Shen, Z. Robust unobtrusive fall detection using infrared array sensors. In Proceedings of the 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Daegu, Republic of Korea, 16–18 November 2017; pp. 194–199. [CrossRef]
38. Muheidat, F.; Tawalbeh, L.A.; Tyrer, H. Context-Aware, Accurate, and Real Time Fall Detection System for Elderly People. In Proceedings of the 2018 IEEE 12th International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 31 January–2 February 2018; pp. 329–333. [CrossRef]

39. Chaccour, K.; Darazi, R.; el Hassans, A.H.; Andres, E. Smart carpet using differential piezoresistive pressure sensors for elderly fall detection. In Proceedings of the 2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Abu Dhabi, United Arab Emirates, 19–21 October 2015; pp. 225–229. [CrossRef]
40. Ren, L.; Peng, Y. Research of Fall Detection and Fall Prevention Technologies: A Systematic Review. *IEEE Access* **2019**, *7*, 77702–77722. [CrossRef]
41. Wang, S.; Sheng, H.; Yang, D.; Zhang, Y.; Wu, Y.; Wang, S. Extendable Multiple Nodes Recurrent Tracking Framework With RTU++. *IEEE Trans. Image Process.* **2022**, *31*, 5257–5271. [CrossRef] [PubMed]
42. Wu, Y.; Sheng, H.; Zhang, Y.; Wang, S.; Xiong, Z.; Ke, W. Hybrid Motion Model for Multiple Object Tracking in Mobile Devices. *IEEE Internet Things J.* **2022**, *1*. [CrossRef]
43. Xiong, S.; Li, B.; Zhu, S. DCGNN: A single-stage 3D object detection network based on density clustering and graph neural network. *Complex Intell. Syst.* **2022**, 1–10. [CrossRef]
44. Lu, S.; Liu, S.; Hou, P.; Yang, B.; Liu, M.; Yin, L.; Zheng, W. Soft Tissue Feature Tracking Based on Deep Matching Network. *Comput. Model. Eng. Sci.* **2023**, *136*, 363–379. [CrossRef]
45. Zhao, L.; Lu, S.-P.; Chen, T.; Yang, Z.; Shamir, A. Deep Symmetric Network for Underexposed Image Enhancement with Recurrent Attentional Learning. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, BC, Canada, 11–17 October 2021; pp. 12055–12064. [CrossRef]
46. Zhang, Q.; Nie, Y.; Zheng, W. Dual Illumination Estimation for Robust Exposure Correction. *Comput. Graph. Forum* **2019**, *38*, 243–252. Available online: <http://arxiv.org/abs/1910.13688> (accessed on 21 July 2022). [CrossRef]
47. Charfi, I.; Miteran, J.; Dubois, J.; Atri, M.; Tourki, R. Optimized spatio-temporal descriptors for real-time fall detection: Comparison of support vector machine and Adaboost-based classification. *J. Electron. Imaging* **2013**, *22*, 41106. [CrossRef]
48. Kwolek, B.; Kepski, M. Human fall detection on embedded platform using depth maps and wireless accelerometer. *Comput. Methods Programs Biomed.* **2014**, *117*, 489–501. [CrossRef]
49. Lin, T.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2999–3007. [CrossRef]
50. Cheng, E.J.; Prasad, M.; Yang, J.; Khanna, P.; Chen, B.-H.; Tao, X.; Young, K.-Y.; Lin, C.-T. A fast fused part-based model with new deep feature for pedestrian detection and security monitoring. *Measurement* **2019**, *151*, 107081. [CrossRef]
51. Hong, G.-J.; Li, D.-L.; Pare, S.; Saxena, A.; Prasad, M.; Lin, C.-T. Adaptive Decision Support System for On-Line Multi-Class Learning and Object Detection. *Appl. Sci.* **2021**, *11*, 11268. [CrossRef]
52. Cheng, E.J.; Prasad, M.; Yang, J.; Zheng, D.R.; Tao, X.; Mery, D.; Young, K.Y.; Lin, C.T. A novel online self-learning system with automatic object detection model for multimedia applications. *Multimed. Tools Appl.* **2020**, *80*, 16659–16681. [CrossRef]
53. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788. [CrossRef]
54. Jocher, G.; Stoken, A.; Chaurasia, A.; Borovec, J.; Kwon, Y.; Michael, K.; Liu, C.; Fang, J.; Abhiram, V.; Skalski, S.P. Ultralytics/yolov5: V6.0—YOLOv5n ‘Nano’ models, Roboflow integration, TensorFlow export, OpenCV DNN support. *Zenodo Tech. Rep.* **2021**. [CrossRef]
55. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767. Available online: <http://arxiv.org/abs/1804.02767> (accessed on 20 January 2023).
56. Huang, C.-J. Integrate the Hungarian Method and Genetic Algorithm to Solve the Shortest Distance Problem. In Proceedings of the 2012 Third International Conference on Digital Manufacturing & Automation, Guilin, China, 31 July–2 August 2012; pp. 496–499. [CrossRef]
57. Chang, L.C.; Pare, S.; Meena, M.S.; Jain, D.; Li, D.L.; Saxena, A.; Prasad, M.; Lin, C.T. An Intelligent Automatic Human Detection and Tracking System Based on Weighted Resampling Particle Filtering. *Big Data Cogn. Comput.* **2020**, *4*, 27. [CrossRef]
58. Poonsri, A.; Chiracharit, W. Fall detection using Gaussian mixture model and principle component analysis. In Proceedings of the 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE), Phuket, Thailand, 12–13 October 2017; pp. 1–4. [CrossRef]
59. Chamle, M.; Gunale, K.G.; Warhade, K.K. Automated unusual event detection in video surveillance. In Proceedings of the 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–27 August 2016; pp. 1–4. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.