

Article

Fall Detection Based on Key Points of Human-Skeleton Using OpenPose

Weiming Chen , Zijie Jiang , Hailin Guo and Xiaoyang Ni *

Faculty of Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China; chenweiming@cug.edu.cn (W.C.); jzj0803@foxmail.com (Z.J.); hl_guo2004@163.com (H.G.)

* Correspondence: nixiaoyang@cug.edu.cn or xy_ni@163.com

Received: 25 March 2020; Accepted: 16 April 2020; Published: 5 May 2020



Abstract: According to statistics, falls are the primary cause of injury or death for the elderly over 65 years old. About 30% of the elderly over 65 years old fall every year. Along with the increase in the elderly fall accidents each year, it is urgent to find a fast and effective fall detection method to help the elderly fall. The reason for falling is that the center of gravity of the human body is not stable or symmetry breaking, and the body cannot keep balance. To solve the above problem, in this paper, we propose an approach for reorganization of accidental falls based on the symmetry principle. We extract the skeleton information of the human body by OpenPose and identify the fall through three critical parameters: speed of descent at the center of the hip joint, the human body centerline angle with the ground, and width-to-height ratio of the human body external rectangular. Unlike previous studies that have just investigated falling behavior, we consider the standing up of people after falls. This method has 97% success rate to recognize the fall down behavior.

Keywords: fall detection; openpose; skeleton extraction

1. Introduction

The decline of birth rate and the prolongation of life span lead to the aging of the population, which has become a worldwide problem [1]. According to the research [2], the elderly population will increase dramatically in the future, and the proportion of the elderly in the world population will continue to grow, which is expected to reach 28% in 2050. Aging is accompanied by a decline in human function, which increases the risk of falls. According to statistics, falls are the primary cause of injury or death for the elderly over 65 years old. About 30% of the elderly over 65 years old fall every year [3]. In 2015, there were 29 million elderly falls in the United States, of which 37.5% required medical treatment or restricted activities for 1 day or more, and about 33,000 people died [4]. The most common immediate consequences of falls are fractures and other long-term ailments, which can lead to disability and loss of independence and psychological fear of falling again [5]. Falls not only make the elderly suffer moderate or severe injuries, but also bring a mental burden and economic pressure to the elderly and their relatives [6]. Faced with this situation, it is particularly important to quickly and effectively detect the fall of the elderly and provide emergency assistance. In a word, it is extremely important for those who fall and cannot call for help to be found in time and to be treated.

This paper proposes a new detection method for falling. This method processes every frame captured by monitoring, which is to use the OpenPose skeleton extraction algorithm to obtain the skeleton data of people on the screen. In the plane coordinate system, the horizontal and vertical coordinates are used to represent each node. According to the speed of descent at the center of the hip joint, the human body centerline angle with the ground, and the width-to-height ratio of the human body external rectangular, these determine the conditions to identify falling behavior and if the person can stand on his/her own after a fall.

The remainder of this paper is organized as follows: Section 2 reviews the current methods of skeleton extraction and fall detection. Section 3 details the approach (e.g., skeleton extraction and behavior recognition). Section 4 presents the results of an experiment to validate the effectiveness and feasibility of our proposed approach. Section 5 discusses the limitations of the study and potential future work.

2. Related Work

In this paper, the skeleton estimation method is used to extract human joint points for fall detection. Skeleton extraction and fall detection are discussed in the literature review.

2.1. Skeleton Estimation

The skeletal structure of the human body determines the geometric structure of the human movement. In the field of computer vision, the skeleton is defined as a human torso and head and limb position diagram of the model [7]. In other words, the relative position of the joints in the skeleton determines the posture of the body.

With the coming of the 21st century, people have entered the information age; artificial algorithm have been widely used in all kind of fields including Public Opinion Polarization Process [8,9], Emergency Logistics Network [10], network structure optimization [11], mixed traffic flow model [12], Road Network [13] and so on. People have been eager to use computers instead of humans to do repetitive and simple work, hoping that computers can learn like humans. In this case, computer vision was born, which is used in most skeleton estimations. The methods for skeleton estimation from pictures or videos can be divided into two categories: depth images and normal RGB images.

Since the depth image contains the position information of human joint points, we can convert the position information of joint points into human skeletons to infer human behavior. There are two methods to obtain depth images: passive range sensor and active depth sensor. The most commonly used method of the passive range sensor is binocular stereo vision [14], which obtains two images of the same scene at the same time by two cameras with a certain distance apart, finds the corresponding pixel points in two images by stereo algorithm, and then calculates the time difference information according to the triangle principle; the time difference information can represent the depth information of the object in the scene by conversion. Based on the stereo matching algorithm [15], the depth image of the same scene can also be obtained by photographing a group of images with different angles in the same scene. Compared with the passive range sensor, the active depth sensor has the most important feature: the equipment itself needs to transmit energy to complete the depth information collection. The methods of active depth sensor mainly include TOF (time of flight) [16] and Kinect [17].

The principle of normal RGB images skeleton estimation is human skeleton key points detection. The detection of human skeleton key points is also known as a pose estimation. It mainly detects some key points of the human body, such as joints and five senses, and describes human skeleton information through key points. There are two main directions of the 2D human skeleton key point detection algorithm: top-down detection method [18] and bottom-up detection method [19]. The OpenPose used in this paper is a bottom-up detection method. It is a new way to use OpenPose to get the data of human skeleton to study the fall. Sungil et al. [20] introduced a fall-detection method based on skeleton data by OpenPose and constructed a fall detection system using LSTM. Experiments show that the method of fall detection compared with untreated raw frame data is more effective. Xu et al. [21] used OpenPose to get the data set of a human skeleton map and trained to get a new model that can predict the fall.

2.2. Fall Detection

In the current fall detection research, both articles [22,23] divide fall systems into three categories: vision-based sensors, wearable device-based sensors, and ambient sensors. Later, Ren et al. [24] proposed a more comprehensive classification scheme on fall detection from the

sensor apparatus. According to the sensing equipment used in the existing fall detection system, as shown in Figure 1, fall detection is divided into four categories: inertial sensor-based, context-based, RF-based, and sensor fusion-based.

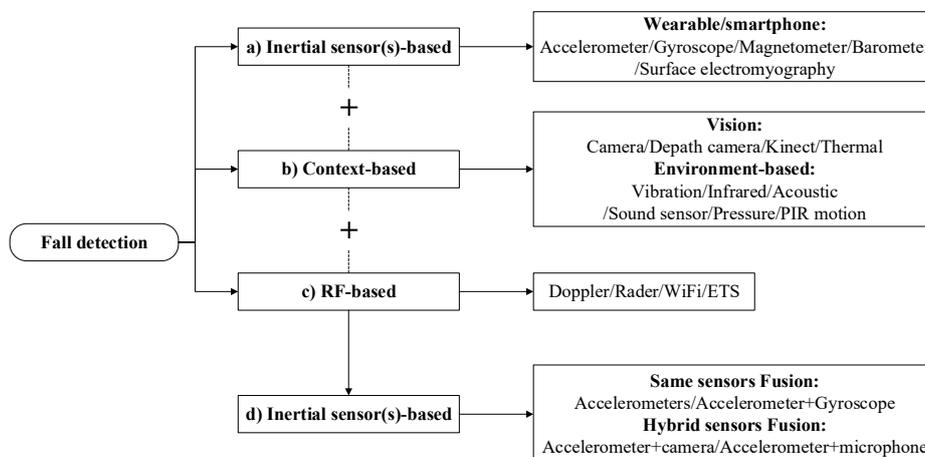


Figure 1. Taxonomy for fall detection from sensor apparatus aspect.

2.2.1. Inertial Sensor(s)-Based Fall Detection

There are severe changes during the falling, such as collisions, changes in body orientation, or severe tilts. These features can be measured by sensors such as accelerometers, barometers, gyroscopes, magnetometers, etc. Shahzad et al. [25] developed a pervasive fall detection system on smartphones, in which the system uses accelerometer signals embedded in the smartphone and a proposed two-step algorithm to detect falls. Fino et al. [26] proposed two novel methods and combined them to achieve remote monitoring of turning behavior using three uniaxial gyroscopes, and found the relationship between rotation frequency and fall. The principle of the pressure sensor is to detect and track the pressure based on the weight of the object. Light et al. [27] have built pressure sensors into smart shoes that detect falls by measuring whether a person's gait is consistent. Han et al. [28] used the bidirectional EMG (electromyographic) sensor network model to realize simple communication between the user and the nursing staff and finally proved that the method could detect the fall events more flexibly and effectively. Sun et al. [29] used a plantar inclinometer sensor to obtain the angle change information in the process of walking and the angle status after falling. Select the threshold in four directions from the plantar angle of the fall state: forward, backward, left, and right. They conducted 100 tests on falls under different circumstances, and the detection rate was 92%. The advantages of the inertial sensor(s)-based fall detection method are: portability, easy to implement, good real-time, few privacy issues, and high accuracy. However, this method also has corresponding shortcomings. The most obvious is that people need to wear the corresponding device on their bodies, which is undoubtedly an intrusion for users.

2.2.2. Context-Based Fall Detection

Context-based fall detection can be divided into two categories: ambient-based and vision-based. The common ground of these two methods is that they detect falls by detecting external environmental information to track human behavior. The ambient-based system mainly uses sensors to collect vibration, acoustic and pressure signals to track the human body. Droghini et al. [30] used a floor acoustic sensor to acquire sound waves passing through the floor and established a human fall classification system. Using the new sensing technology of piezoresistive pressure sensors, Chaccour et al. [31] designed an intelligent carpet to detect falls. Infrared array sensors are also used in fall detection systems. The difference is that Fan et al. [32] used a variety of deep learning methods to improve them, making their fall detection systems have more obvious advantages. Compared with the inertial-based system,

the biggest advantage of ambient-based devices is that they basically do not interfere with people. The obscure and minimal interaction with people also determines that the ambient-based fall detection system rarely involves security and privacy issues. However, these methods have a limited detection range. Besides, ambient sensors are easier affected by the external environment.

As the camera is widely used in our daily life, the camera is also gradually used to obtain relevant information, which is also considered as vision-based fall detection systems. Many studies have used depth camera(s) (Kinect), RGB camera(s), thermal sensor(s), or even a combination of cameras to track changes in body shape, the trajectory of the head, or to monitor the body posture of the subject to detect or prevent falls. Fan et al. [33] proposed a novel fall detection method based on vision; this method is used to describe the human body posture to analyze extraction. Based on Kinect, Liu et al. [34] developed a novel fall recognition algorithm, which can quickly and effectively recognize human falls after experimental verification. Kong et al. [35] got the outline of the binary image by the canny filter and a depth camera. Then, the output outline image was used for falling detection. Rafferty et al. [36] combined the computer vision processes with the thermal vision sensor installed on the ceiling for fall detection. This novel method overcomes the shortcomings of traditional methods. However, there are some disadvantages in vision-based fall detection, such as considerable computing and storage capacity to run the real-time algorithm, privacy issues, and limited capture space can be monitored.

2.2.3. RF-Based Fall Detection

The study found that violent body movements can cause abnormal changes in the RF signal. This feature provides a new idea for fall detection, which is to detect falls through the fluctuation of the RF signal. RF-based fall detection systems are mainly divided into two categories, including radar frequency-based and wireless channel-based system. Tang et al. [37] proposed a fall prevention system based on FMCW radar, which predicts a fall by continuously measuring the distance between the radar and the surrounding environment and analyzing the relationship between human motion and radar frequency. The fall-related system based on wireless channel state information can quickly estimate changes in wireless signals caused by different human activities, which can be WiFi or Bluetooth. Wang et al. [38] designed an indoor fall detection system using ordinary WiFi equipment, which has many advantages such as being real-time, non-contact, low-cost, and accurate. As the frequency signal of radar is ubiquitous, its biggest advantage is that it can detect the fall event conveniently without being intrusive to the user. However, RF-based technology also has its limitations. Most wireless networks are deployed in houses within a limited range, and there are problems with their comprehensive coverage.

2.2.4. Sensor Fusion-Based Fall Detection

The problem of low accuracy or high false positives is widespread in the fall detection system with a single sensor, which means that other information is needed to improve the accuracy of the system. For example, Lu et al. [39] used a combination of infrared sensors and pressure sensors to detect fall events. Quadros et al. [40] used an accelerometer, gyroscope and magnetometer to obtain a variety of information such as acceleration, velocity, displacement, and direction components and then integrated them. Using the fusion information, they proposed a fall detection method based on the combination of threshold value and machine learning. Kepski et al. [41] proposed an efficient fall detection algorithm, which uses information derived from wireless inertial sensors and depth images. Ramezani et al. [42] detected falls according to ground vibration. Different from traditional methods, ground vibration signals were acquired by combining Wi-Fi Channel State Information (CSI) with the ground-mounted accelerometer.

The sensor fusion system can provide more human activity information. On the one hand, the increase in information has significantly improved the performance of the fall detection system, but at the same time, a large amount of information also brings many disadvantages, such as poor performance of information fusion methods and redundant information and robust fusion algorithm.

The advantages and disadvantages of various classifications are shown in Table 1. The method proposed in this paper belongs to vision-based. It can make full use of the cameras around our lives, which is convenient and has high accuracy, low cost, and is easy to implement.

Table 1. Advantages and shortcomings of each classification.

	Inertial Sensor(s)-Based	Context-Based		RF-Based	Sensors Fusion-Based
		Ambient-Based	Vision-Based		
Advantages	Easy to implement Few privacies issue High accuracy Real time	Least intrusive Few privacy and security issues	Convenient Accurate	Real-time Contactless Low-cost Nonintrusive	Accurate Significant performance
Shortcoming	Intrusive	Limited detection range Easier affected by the external environment	Considerable computing Privacy issue Limited capture space	Coverage issue Limited range	Information redundancy Robust fusion algorithm

3. Methods

Our proposed approach consists of five key steps: (1) OpenPose gets the skeleton information of the human body; (2) Decision condition one (The angle between the centerline of the body and the ground); (3) Decision condition two (The angle between the centerline of the human and the ground); (4) Decision condition three (The width to height ratio of the human body external rectangular); and (5) The procedure of implementation of our proposed approach is as shown in Figure 2.

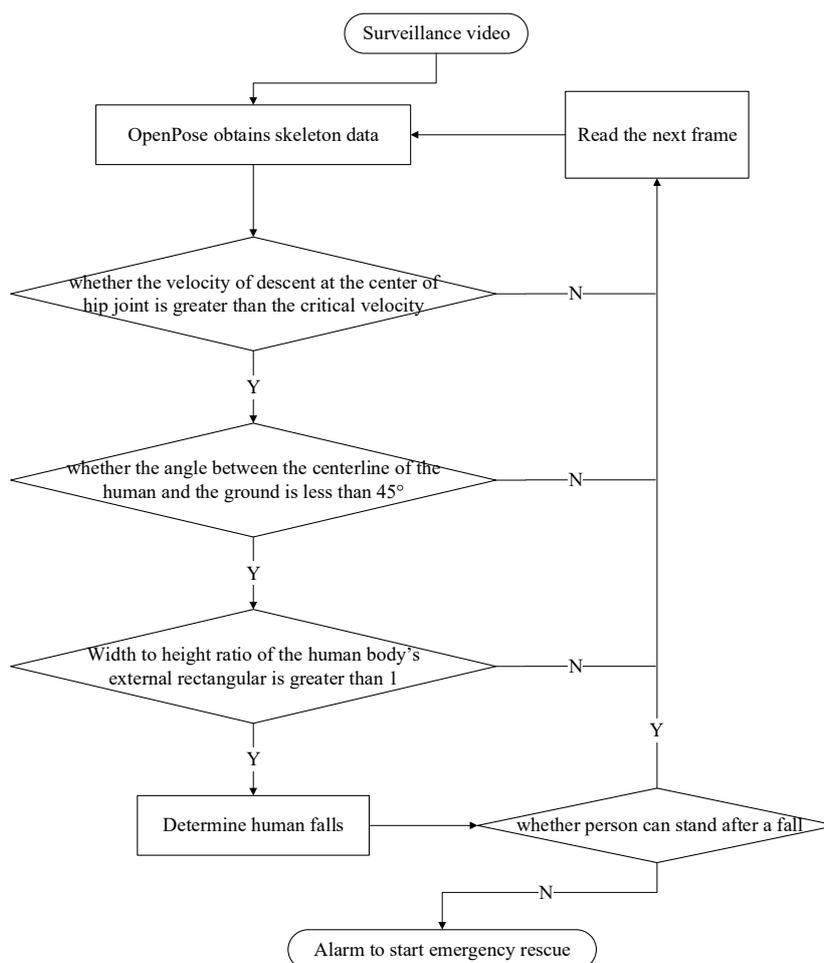


Figure 2. The workflow of our proposed approach.

3.1. OpenPose Gets the Skeleton Information of the Human Body

The OpenPose human gesture recognition project is an open-source library developed by Carnegie Mellon University (CMU) based on convolutional neural network and supervised learning and based on Caffe (Convolutional Architecture for Fast Feature Embedding) [43]. In 2017, researchers from Carnegie Mellon University released the source code of the human skeleton recognition system of OpenPose to realize real-time tracking of targets under the monitoring of video. It can capture the COCO (Common Objects in Context) human skeleton information in the color video and provide joints information in the scene. OpenPose human key node recognition system can realize real-time detection of multi-person skeleton information. It adopts the top-down human body attitude estimation algorithm to detect the position of key points of the human body and then uses the feature vector affinity parameter to determine the hot spot map of human key nodes. OpenPose can realize human movement, facial expression, finger movement, and other posture estimation. It is suitable for a single person and many people with excellent robustness.

As shown in Figure 3, the screen taken by the surveillance camera uses OpenPose to obtain the information of human key nodes. The surveillance video is divided into a series of frames, each showing the skeleton of a person.

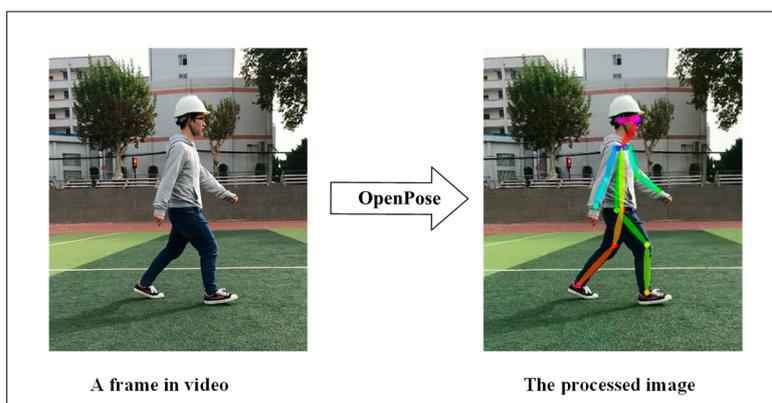


Figure 3. OpenPose gets the skeleton information of the human body.

As shown in Table 2, the position information of each joint point is represented by the horizontal and vertical coordinate values, and the accuracy of each joint point is provided. For some joints, the accuracy of their coordinate position is not very ideal. This problem is mainly due to the defects of OpenPose algorithm itself, but the deviation of some key points has little effect on the recognition of the whole fall action. The specific joint points corresponding to each joint point number in the table are shown in Figure 4.

Table 2. Joint point data obtained through OpenPose.

Joint Number	Accuracy	X-Coordinate	Y-Coordinate
0	0.97517216	333	93
1	0.86759883	355	113
2	0.86547723	351	113
3	0.70167869	343	154
4	0.89546448	331	184
5	0.92826366	360	112
6	0.91062319	372	148
7	0.95447254	376	180
8	0.73973876	355	190
9	0.90461838	377	238
10	0.92706913	398	284
11	0.80891138	355	189
12	0.92123324	337	239

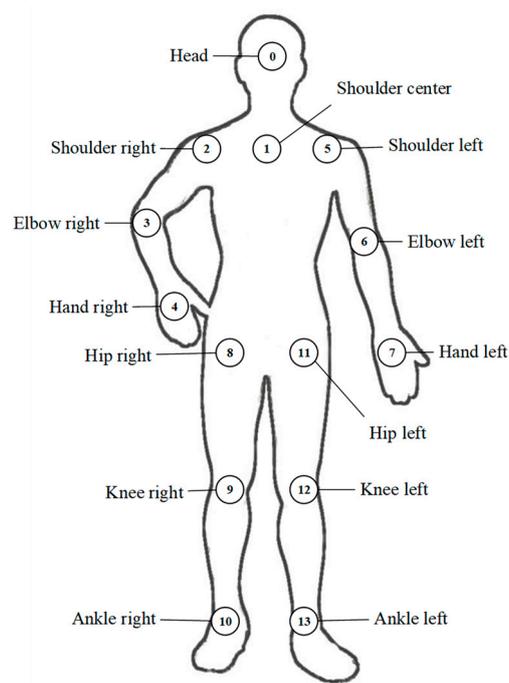


Figure 4. Human node model diagram.

For the convenience of representation, $S = \{s_0, s_1, \dots, s_{13}\}$ represents the joint position set. We define the Joint Coordinates (JC): Define the position of the node j at time t as $s_j(t) = (x_{tj}, y_{tj})$, $j \in \{0, 1, \dots, 13\}$.

3.2. Decision Condition One (the Speed of Descent at the Center of the Hip Joint)

As shown in Figure 5, in the process of sudden fall, the center of gravity of the human body will change in the vertical direction. The central point of the human hip joint can represent the center of gravity of the human body and reflect this feature. By processing the joint point data obtained from the OpenPose, the longitudinal coordinates of the hip joint center point of each frame of the image are obtained. Because it is a very short process from standing posture to falling posture, and the time used is also very short, it is detected once every five adjacent frames, with a time interval of 0.25 s. The coordinates of the hips are $s_8(t) = (x_{t8}, y_{t8})$ and $s_{11}(t) = (x_{t11}, y_{t11})$. Assume that the y-coordinate of the center of the human hip joint at time t_1 is $y_{t_1} = \frac{y_{t_18} + y_{t_111}}{2}$ and the y-coordinate at time t_2 is $y_{t_2} = \frac{y_{t_28} + y_{t_211}}{2}$. According to these, the descent velocity of the hip joint center can be obtained.

$$\Delta t = t_2 - t_1 \quad (1)$$

$$v = \frac{|y_{t_2} - y_{t_1}|}{\Delta t} \quad (2)$$

where v is greater than or equal to the critical speed \bar{v} , the fall feature is considered to be detected. According to the experimental results, this paper chooses 0.009 m/s as the threshold of the falling speed of the hip joint center.

$$M_1 = \begin{cases} 0; & v < \bar{v} \\ 1; & v \geq \bar{v} \end{cases} \quad (3)$$

when $v \geq \bar{v}$ $M_1 = 1$, it can be considered to satisfy the decision condition one.



Figure 5. The falling process.

3.3. Decision Condition Two (the Angle between the Centerline of the Human and the Ground)

In the process of falling, the most obvious feature of the human body is the body tilt, and tilt degree will continue to increase. In order to reflect the characteristics of the body’s continuous tilt in the process of human fall, a human centerline L is defined in this paper (Let the midpoint of joint s_{12} and joint point s_{13} be \bar{s} , and the connection of midpoint \bar{s} and joint s_0 is the centerline L of the human body).

As shown in Figure 6, θ is the angle between the centerline of the human and the ground. Through OpenPose, the data of joint points 0, 10 and 13 are $s_0(t) = (x_{t0}, y_{t0})$, $s_{10}(t) = (x_{t10}, y_{t10})$ and $s_{13}(t) = (x_{t13}, y_{t13})$ respectively. So $\bar{s} = \frac{s_{10}+s_{13}}{2}$, $\bar{s}(t) = (\bar{x}_t, \bar{y}_t)$. At time t , the angle between the centerline of human body and the ground is $\theta_t = \arctan\left|\frac{y_{t0}-\bar{y}_t}{x_{t0}-\bar{x}_t}\right|$.

$$M_2 = \begin{cases} 0; & \theta \geq \theta_0 \\ 1; & \theta < \theta_0 \end{cases} \tag{4}$$

when $\theta < \theta_0$ ($\theta_0 = 45^\circ$) $M_2 = 1$, it can be considered as satisfying the decision condition two for the occurrence of the fall event.

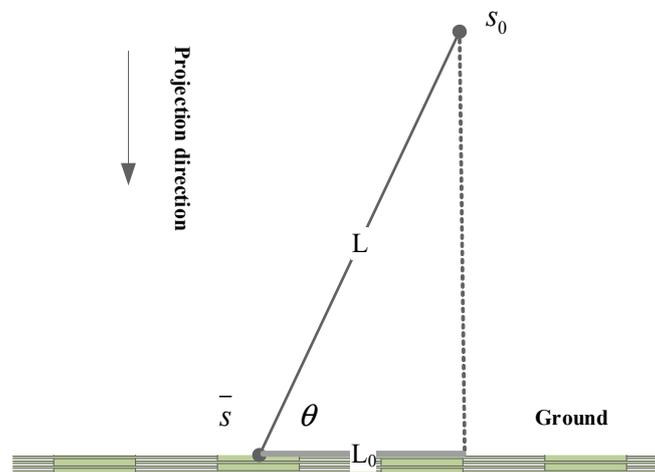


Figure 6. The angle between the centerline of the body and the ground.

3.4. Decision Condition Three (the Width to Height Ratio of the Human Body External Rectangular)

When a fall is detected, the most intuitive feature is a change in the contours of the body. If we simply compare the length and height of the moving target, both the length and height of the moving target will change due to the distance from or near the camera, while their ratio will not exist. We will detect the falling behavior through the change of the length and height ratio of the target contour rectangle.

As shown in Figure 7, the ratio of width to the height of the outer rectangle of the human body is $P = \text{Width}/\text{Height}$. When the human body falls, the outer rectangle of the target will also change; the most significant manifestation is the change of the length–height ratio.

$$M_3 = \begin{cases} 1; & P \geq T \\ 0; & P < T \end{cases} \quad (5)$$

where T is the threshold. According to the actual situation, when a human body normally walks, the width-to-height ratio P is less than 1, while the width-to-height ratio for falling is greater than 1. When $P \geq T$, $M_3 = 1$, it can be considered as satisfying decision condition three of the occurrence of the fall event.

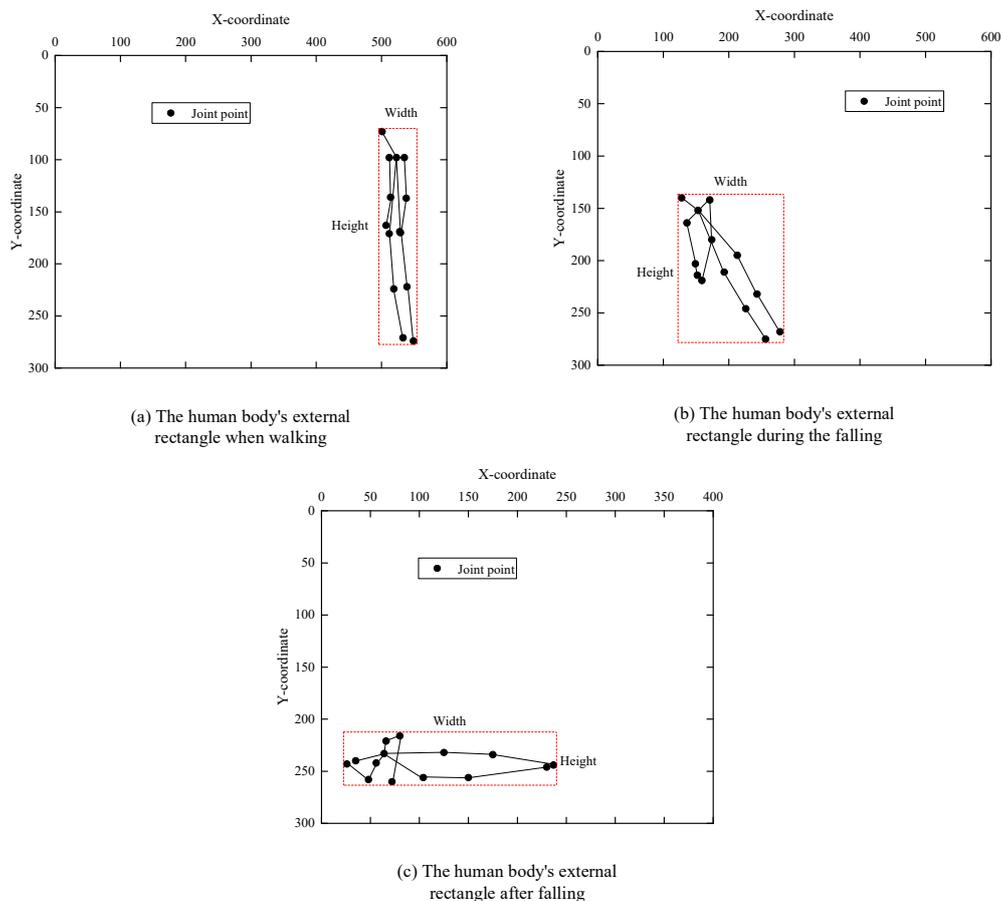


Figure 7. Human body external rectangular.

3.5. Determine Whether a Person Can Stand after a Fall

If a person can stand on his own within a period after falling, no alarm is required. Nowadays, most of the fall detection focuses on the analysis of the fall process, rarely considering that people stand on their own within a short time after falling. As shown in Figure 8, standing up after a fall can be regarded as an inverse process of a fall. The only difference is that the whole process is slower than a fall. According to the analysis of this paper, if the ratio of height to width of the external rectangle of the human body is less than 1 and the inclination angle of the central line is greater than 45° in a period of time after a fall, it can be concluded that the person has stood up. The point of judging whether people can stand up on their own after a fall is to reduce unnecessary alarms because sometimes falls do not cause serious injury to the human body.



Figure 8. The process of standing up after a fall.

4. Experimental Results

4.1. Experiment Data and Test

In order to verify the effectiveness of the proposed method, the fall event is tested. Because this experiment has certain risks, the experimental site is chosen in the laboratory. We randomly select 10 experimenters who made falls or non-falls during the test. As shown in Table 3, the actions collected in the experiment are divided into three categories, namely falling actions (fall, stand up after a fall), similar falling actions (squat, stoop), and daily actions (walk, sit down). A total of 100 actions are collected, including 60 falling actions and 40 non-falling actions, each lasting about 5–11 s. From each video, 100–350 valid video frames can be extracted as samples.

Table 3. Collection action classification.

Action	Action Description	Sample Size
Falling actions	Fall	40
	Stand up after a fall	20
Similar falling actions	Squat/Stoop	20
Daily actions	Walk/Sit down	20

In order to ensure the universality of the system in the test experiment, 10 different types of experimental subjects are randomly selected. The height and weight data of 10 experimenters are shown in Table 4. In the experiment, each person performed 10 actions, including six falls and four non-falls, with a total of 100 action samples.

Table 4. The height and weight data of 10 experimenters.

Experiment Number	1	2	3	4	5	6	7	8	9	10
Height (cm)	172	170	175	163	172	175	177	178	183	177
Weight (kg)	65	64	72.5	67	67.5	67	68	57	67	60

In the test of falling, there are four possible cases: In the first case, a fall event occurs and the algorithm correctly detects the fall; in the second case, the fall did not happen but the algorithm misidentified it as a fall; in the third case, a fall occurs but the algorithm judges that it did not fall; in the fourth case, the fall did not happen and the algorithm did not detect the fall. The above four cases are defined as TP, FP, TN, and FN respectively.

True positive (TP): a fall occurs, the device detects it.

False positive (FP): the device announces a fall, but it did not occur.

True negative (TN): a normal (no fall) movement is performed, the device does not declare a fall.

False negative (FN): a fall occurs but the device does not detect it.

To evaluate the response to these four situations, two criteria are proposed:

Sensitivity is the capacity to detect a fall:

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

Specificity is the capacity to detect only a fall:

$$Specificity = \frac{TN}{TN + FP} \quad (7)$$

Accuracy is the capacity to correctly detect fall and no fall:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)$$

4.2. Analysis of the Experimental Results

Before the final experimental judgment, we analyze the feasibility of the three conditions and the final conditions of standing up after falling.

When detecting the descending speed of the hip joint center point, the speed of change of each action is shown in Figure 9 below. We can see that the speed of fall and squat can exceed the critical value (0.09 m/s). In other words, only falling and squatting down meet the conditions by decision condition one (the speed of descent at the center of the hip joint).

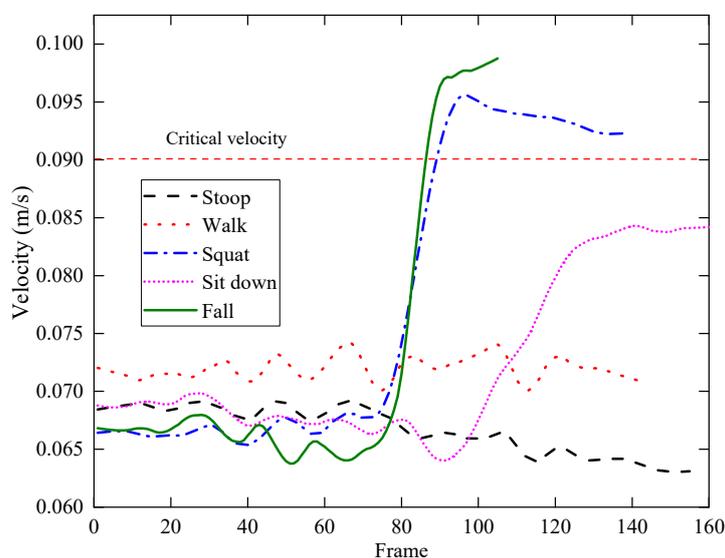


Figure 9. Speed change of each action.

As shown in Figure 10: When walking and sitting down, the inclination angle of the human body fluctuates less; when squatting down, the inclination angle of the human body fluctuates, but the whole body is relatively stable; only when stooping and falling, the inclination angle of the human body fluctuates greatly, and the inclination angle is less than the critical angle 45° . We can exclude walking, sitting down and squatting from the decision condition two (the angle between the centerline of the body and the ground).

As shown in Figure 11, in all the actions, only the width–height ratio of the external rectangle of the human body in the falling action is greater than 1. By decision condition three (the width to height ratio of the human body external rectangular), we can find that only the falling action meets the requirement.

As shown in Figure 12, it shows that the common feature of falling action is that the inclination angle of the human body must fall below 45° and the aspect ratio of the external rectangle of the human body will be greater than 1 at a certain time. For the action of standing up after a fall, the fall process can be judged according to the judgment conditions of the fall. In the subsequent rise process, it can be found that the inclination angle of the human body will gradually increase to above 45° , and the width–height ratio of the external rectangle of the human body is also less than 1.

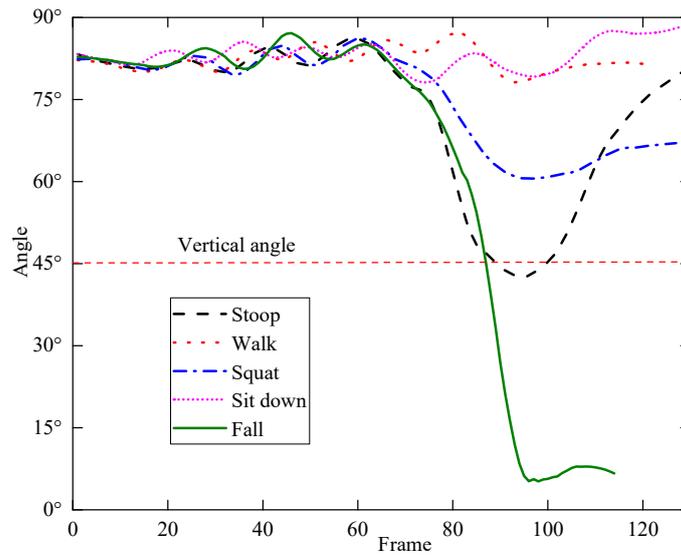


Figure 10. The change of inclination angle of each action.

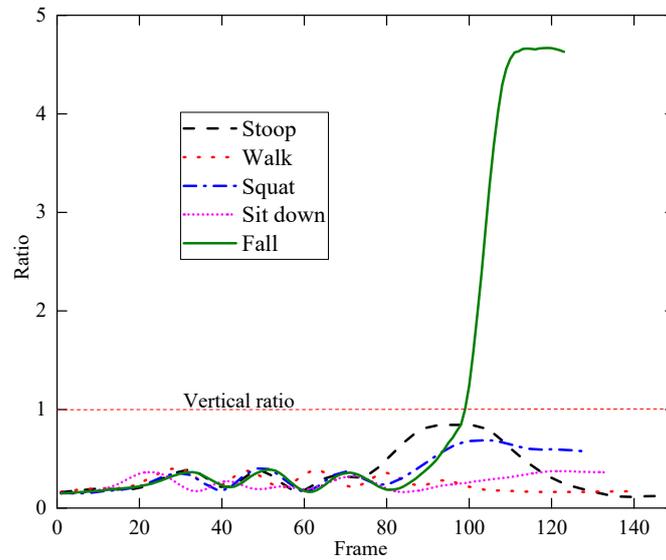


Figure 11. The change of the aspect ratio of the outer rectangle for each action.

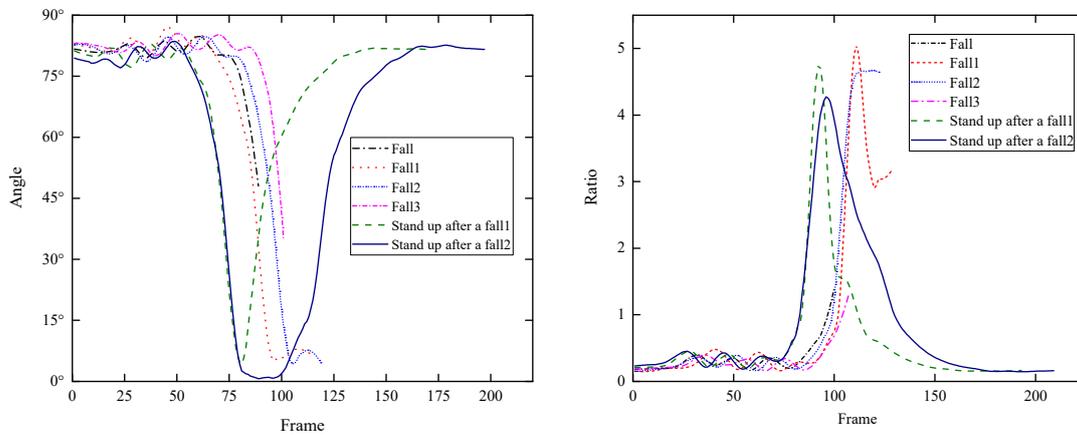


Figure 12. The characteristic of standing up after a fall.

Through the analysis of a total of 100 experimental actions, the specific situation is shown in the Table 5 below. In the table, ✓ indicates that the action is correctly identified, × indicates that the action is incorrectly identified. It can be seen that No.1 and No.3 experiments' stooping actions in the non-falling actions are wrongly identified as falling, and only one time in the falling actions is wrongly identified as non-falling.

Table 5. Test results of 100 actions.

Number	Action									
	Stoop	Squat	Walk	Sit Down	Fall	Fall1	Fall2	Fall3	Stand Up after a Fall1	Stand Up after a Fall2
1	×	✓	✓	✓	✓	✓	✓	✓	✓	✓
2	✓	✓	✓	×	✓	✓	✓	✓	✓	✓
3	×	✓	✓	✓	✓	✓	✓	✓	✓	✓
4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
5	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
6	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
7	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
8	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
9	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
10	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

According to the calculation formula proposed in Section 4.2, the sensitivity, specificity and accuracy are 98.3%, 95% and 97% in Table 6. There are the following reasons for wrong discrimination: (a) The lack of joint points in skeleton estimation results in incomplete data, which affects the final recognition. (b) The three thresholds selected in the experiment are not necessarily optimal. (c) During the experiment, due to the self-protection consciousness of the experimenter, there are still differences between the recorded falls and the real falls.

Table 6. The experimental calculation results.

	Sensitivity	Specificity	Accuracy
Result	98.3%	95%	97%

5. Conclusions and Future Work

Conclusions

At present, because there are no suitable public datasets of falls, we cannot directly compare our results with previous results in detail. As shown in Table 7, we list the algorithms, classifications, features, and final accuracy of other fall detection technologies. Droghini et al. [30] detected falls by capturing sound waves transmitted on the floor. The accuracy of the experimental results is high, but the experiment uses a puppet to imitate falls, which is still very different from the real human fall. In addition, its detection method is extremely susceptible to interference from external noise, and the available environment is limited. Shahzad et al. [25] make good use of the sensors in smartphones and improves the power consumption of the algorithm, but the phone can always also cause false positives and requires the user to wear the phone. Kepski et al. [44] proposed a fall recognition system based on microwave doppler sensor, which can not only distinguish fall and fall-like movements accurately, but also does not infringe on the human body. The only disadvantage of this method is that the detection range is too small. Quadros et al. [40], the threshold method and machine learning are used to fuse multiple signals to identify falls, which undoubtedly improves the reliability of the recognition results. However, the user needs to wear the device for a long time, and the endurance of the device should also be considered. The method of OpenPose [20,21] can be used to identify the images captured by the camera, which is convenient and fast, and has a broad prospect in video-based methods. Compared with other methods, vision-based is more convenient. OpenPose gets the skeleton

information of the human body, which is convenient and accurate. To some degree, our method not only has high accuracy but also is simple and low cost.

Table 7. Comparison of our proposed algorithm with other fall detection approaches.

Algorithm	Classification	Features	Accuracy (%)
Mel-Frequency Cepstral Coefficients + SVM [30]	Ambient-based	Acoustic waves	99.14–100%
Threshold + SVM [25]	Inertial sensor (s)-based	Acceleration, Magnitude variation, Max peak	91.7–97.8%
Microwave Doppler + Markov model [44]	RF-based	Velocity Frequency	95%
Threshold + Madgwick's decomposition [40]	Fusion-based	Acceleration, Velocity Displacement	91.1%
OpenPose + LSTM [20]	Vision-based	Coordinate, Speed	98.7%
OpenPose + Convolutional neural network [21]	Vision-based	skeleton map	91.7%
Our proposed OpenPose + three thresholds	Vision-based	Velocity, Angle, Ratio	97%

According to statistics, the elderly population will continue to increase in the future, and falling is one of the major public health problems in an aging society. It is necessary to find out the characteristics of the fall movement for fall detection. In this paper, we introduce a novel method for this problem. Using OpenPose algorithm to process video captured by surveillance, the data of human joint points are obtained. Then, the falling motion is recognized by setting three conditions: the speed of descent at the center of the hip joint, the angle between the centerline of the human body and the ground, and the width-to-height ratio of the human body external rectangular. Based on the recognition of falls, considering the situation of people standing up after falls, the process of standing up after falls is regarded as an inverse process of falling. The method is verified by experiments and achieved the ideal result. The sensitivity is 98.3%, the specificity is 95%, and the accuracy is 97%.

With the popularity of the camera and the clearer quality of the captured image, the vision-based fall detection method has a broader space. In the future, we can carry out the following work:

- The environment of daily life is complex, there may be situations in which peoples' actions cannot be completely captured by surveillance. In the future, we can study the estimation and prediction of peoples' behavior and actions in the presence of partial occlusion.
- In this paper, the action is identified from the side, and the other directions are not considered. Future research can start with multiple directions recognition and then comprehensively judge whether to fall.
- Building a fall alarm system for people. In the event of a fall, the scene, time, location, and other detailed information shall be timely notified to the rescuer, to speed up the response speed of emergency rescue.

Author Contributions: W.C. contributed to the conception of the study. Z.J. performed the experiment; W.C., Z.J. performed the data analyses and wrote the manuscript; H.G., X.N. helped perform the analysis with constructive discussions. All authors read and approved the manuscript.

Funding: This research was funded by the Open Fund of Teaching Laboratory of China University of Geosciences (Wuhan) grant number SKJ2019095.

Acknowledgments: Thanks to everyone who helped with the experiment. We are also very thankful for the editors and anonymous reviewers.

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

References

1. WHO. Number of People over 60 Years Set to Double by 2050; Major Societal Changes Required. Available online: <https://www.who.int/mediacentre/news/releases/2015/older-persons-day/en/> (accessed on 17 March 2020).
2. Lapiere, N.; Neubauer, N.; Miguel-Cruz, A.; Rincon, A.R.; Liu, L.; Rousseau, J. The state of knowledge on technologies and their use for fall detection: A scoping review. *Int. J. Med. Inform.* **2018**, *111*, 58–71. [[CrossRef](#)] [[PubMed](#)]
3. Christiansen, T.L.; Lipsitz, S.; Scanlan, M.; Yu, S.P.; Lindros, M.E.; Leung, W.Y.; Adelman, J.; Bates, D.W.; Dykes, P.C. Patient activation related to fall prevention: A multisite study. *Jt. Comm. J. Qual. Patient Saf.* **2020**, *46*, 129–135. [[CrossRef](#)] [[PubMed](#)]
4. Grossman, D.C.; Curry, S.J.; Owens, D.K.; Barry, M.J.; Caughey, A.B.; Davidson, K.W.; Doubeni, C.A.; Epling, J.W.; Kemper, A.R.; Krist, A.H. Interventions to prevent falls in community-dwelling older adults: US Preventive Services Task Force recommendation statement. *JAMA* **2018**, *319*, 1696–1704. [[PubMed](#)]
5. Gates, S.; Fisher, J.; Cooke, M.; Carter, Y.; Lamb, S. Multifactorial assessment and targeted intervention for preventing falls and injuries among older people in community and emergency care settings: Systematic review and meta-analysis. *BMJ* **2008**, *336*, 130–133. [[CrossRef](#)] [[PubMed](#)]
6. Faes, M.C.; Reelick, M.F.; Joosten-Weyn Banningh, L.W.; Gier, M.D.; Esselink, R.A.; Olde Rikkert, M.G. Qualitative study on the impact of falling in frail older persons and family caregivers: Foundations for an intervention to prevent falls. *Aging Ment. Health* **2010**, *14*, 834–842. [[CrossRef](#)] [[PubMed](#)]
7. Johansson, G. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* **1973**, *14*, 201–211. [[CrossRef](#)]
8. Chen, T.; Li, Q.; Fu, P.; Yang, J.; Xu, C.; Cong, G.; Li, G. Public opinion polarization by individual revenue from the social preference theory. *Int. J. Environ. Res. Public Health* **2020**, *17*, 946. [[CrossRef](#)]
9. Chen, T.; Li, Q.; Yang, J.; Cong, G.; Li, G. Modeling of the public opinion polarization process with the considerations of individual heterogeneity and dynamic conformity. *Mathematics* **2019**, *7*, 917. [[CrossRef](#)]
10. Chen, T.; Wu, S.; Yang, J.; Cong, G. Risk Propagation Model and Its Simulation of Emergency Logistics Network Based on Material Reliability. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4677. [[CrossRef](#)]
11. Chen, T.; Shi, J.; Yang, J.; Li, G. Enhancing network cluster synchronization capability based on artificial immune algorithm. *Hum. Cent. Comput. Inf. Sci.* **2019**, *9*, 3. [[CrossRef](#)]
12. Jiang, C.; Chen, T.; Li, R.; Li, L.; Li, G.; Xu, C.; Li, S. Construction of extended ant colony labor division model for traffic signal timing and its application in mixed traffic flow model of single intersection. *Concurr. Comput. Pract. Exp.* **2020**, *32*, e5592. [[CrossRef](#)]
13. Chen, T.; Wu, S.; Yang, J.; Cong, G.; Li, G. Modeling of emergency supply scheduling problem based on reliability and its solution algorithm under variable road network after sudden-onset disasters. *Complexity* **2020**, *2020*. [[CrossRef](#)]
14. Ye, Q.; Dong, J.; Zhang, Y. 3D Human behavior recognition based on binocular vision and face–hand feature. *Optik* **2015**, *126*, 4712–4717. [[CrossRef](#)]
15. Alagoz, B.B. Obtaining depth maps from color images by region based stereo matching algorithms. *arXiv* **2008**, arXiv:0812.1340.
16. Foix, S.; Alenya, G.; Torras, C. Lock-in time-of-flight (ToF) cameras: A survey. *IEEE Sens. J.* **2011**, *11*, 1917–1926. [[CrossRef](#)]
17. Zhang, Z. Microsoft kinect sensor and its effect. *IEEE Multimed.* **2012**, *19*, 4–10. [[CrossRef](#)]
18. Newell, A.; Yang, K.; Deng, J. Stacked hourglass networks for human pose estimation. In Proceedings of the Computer Vision—14th European Conference, Amsterdam, The Netherlands, 18 October 2016; pp. 483–499.
19. Insafutdinov, E.; Pishchulin, L.; Andres, B.; Andriluka, M.; Schiele, B. Deepcruc: A deeper, stronger, and faster multi-person pose estimation model. In Proceedings of the European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; pp. 34–50.
20. Jeong, S.; Kang, S.; Chun, I. Human-skeleton based Fall-Detection Method using LSTM for Manufacturing Industries. In Proceedings of the 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), Jeju Shinhwa World, Korea, 23–26 June 2019; pp. 1–4.
21. Xu, Q.; Huang, G.; Yu, M.; Guo, Y. Fall prediction based on key points of human bones. *Phys. A Stat. Mech. Its Appl.* **2020**, *540*, 123205. [[CrossRef](#)]

22. Koshmak, G.; Loutfi, A.; Linden, M. Challenges and issues in multisensor fusion approach for fall detection. *J. Sens.* **2016**, *2016*. [[CrossRef](#)]
23. Mubashir, M.; Shao, L.; Seed, L. A survey on fall detection: Principles and approaches. *Neurocomputing* **2013**, *100*, 144–152. [[CrossRef](#)]
24. Ren, L.; Peng, Y. Research of fall detection and fall prevention technologies: A systematic review. *IEEE Access* **2019**, *7*, 77702–77722. [[CrossRef](#)]
25. Shahzad, A.; Kim, K. FallDroid: An automated smart-phone-based fall detection system using multiple kernel learning. *IEEE Trans. Ind. Inform.* **2018**, *15*, 35–44. [[CrossRef](#)]
26. Fino, P.C.; Frames, C.W.; Lockhart, T.E. Classifying step and spin turns using wireless gyroscopes and implications for fall risk assessments. *Sensors* **2015**, *15*, 10676–10685. [[CrossRef](#)] [[PubMed](#)]
27. Light, J.; Cha, S.; Chowdhury, M. Optimizing pressure sensor array data for a smart-shoe fall monitoring system. In Proceedings of the 2015 IEEE SENSORS, Busan, Korea, 1–4 November 2015; pp. 1–4.
28. Han, H.; Ma, X.; Oyama, K. Flexible detection of fall events using bidirectional EMG sensor. *Stud. Health Technol. Inform.* **2017**, *245*, 1225. [[PubMed](#)]
29. Sun, J.; Wang, Z.; Pei, B.; Tao, S.; Chen, L. Fall detection using plantar inclinometer sensor. In Proceedings of the 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom), Beijing, China, 10–14 August 2015; pp. 1692–1697.
30. Droghini, D.; Principi, E.; Squartini, S.; Olivetti, P.; Piazza, F. Human fall detection by using an innovative floor acoustic sensor. In *Multidisciplinary Approaches to Neural Computing*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 97–107.
31. 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.
32. 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, Korea, 16–18 November 2017; pp. 194–199.
33. Fan, K.; Wang, P.; Hu, Y.; Dou, B. Fall detection via human posture representation and support vector machine. *Int. J. Distrib. Sens. Netw.* **2017**, *13*, 1550147717707418. [[CrossRef](#)]
34. Liu, Y.; Wang, N.; Lv, C.; Cui, J. Human body fall detection based on the Kinect sensor. In Proceedings of the 2015 8th International Congress on Image and Signal Processing (CISP), Shenyang, China, 14–16 October 2015; pp. 367–371.
35. Kong, X.; Meng, L.; Tomiyama, H. Fall detection for elderly persons using a depth camera. In Proceedings of the 2017 International Conference on Advanced Mechatronic Systems (ICAMEchS), Xiamen, China, 6–9 December 2017; pp. 269–273.
36. Rafferty, J.; Synnott, J.; Nugent, C.; Morrison, G.; Tamburini, E. Fall detection through thermal vision sensing. In *Ubiquitous Computing and Ambient Intelligence*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 84–90.
37. Tang, Y.; Peng, Z.; Ran, L.; Li, C. iPrevent: A novel wearable radio frequency range detector for fall prevention. In Proceedings of the 2016 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT), Taipei, Taiwan, 24–26 August 2016; pp. 1–3.
38. Wang, H.; Zhang, D.; Wang, Y.; Ma, J.; Wang, Y.; Li, S. RT-Fall: A real-time and contactless fall detection system with commodity WiFi devices. *IEEE Trans. Mob. Comput.* **2016**, *16*, 511–526. [[CrossRef](#)]
39. Lu, C.; Huang, J.; Lan, Z.; Wang, Q. Bed exiting monitoring system with fall detection for the elderly living alone. In Proceedings of the 2016 International Conference on Advanced Robotics and Mechatronics (ICARM), Macau, China, 18–20 August 2016; pp. 59–64.
40. De Quadros, T.; Lazzaretti, A.E.; Schneider, F.K. A movement decomposition and machine learning-based fall detection system using wrist wearable device. *IEEE Sens. J.* **2018**, *18*, 5082–5089. [[CrossRef](#)]
41. Kepski, M.; Kwolek, B. Event-driven system for fall detection using body-worn accelerometer and depth sensor. *IET Comput. Vis.* **2017**, *12*, 48–58. [[CrossRef](#)]

42. Ramezani, R.; Xiao, Y.; Naeim, A. Sensing-Fi: Wi-Fi CSI and accelerometer fusion system for fall detection. In Proceedings of the 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), Las Vegas, NV, USA, 4–7 March 2018; pp. 402–405.
43. Cao, Z.; Simon, T.; Wei, S.-E.; Sheikh, Y. Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7291–7299.
44. Shiba, K.; Kaburagi, T.; Kurihara, Y. Fall detection utilizing frequency distribution trajectory by microwave Doppler sensor. *IEEE Sens. J.* **2017**, *17*, 7561–7568. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).