Human Fall Down Recognition using Coordinates Skeleton

Wisam hamed abd
Mustansiriyah University, Iraq
Baghdad, Iraq
wisamhamedabd@gmail.com

Ahmed T. Sadiq
Technology University, Iraq
Baghdad, Iraq
Ahmed.t.sadiq@uotechnology.edu.iq

Khalid Ali Hussein
Mustansiriyah University, Iraq
Baghdad, Iraq
dr.khalid.ali68@gmail.com

Abstract: Falls are a significant hazard to human safety and may have devastating consequences in a matter of seconds. This is especially true for elderly, since falls are the primary cause of hospitalization and injury-related mortality in this age group. In this paper, a deep learning based model is proposed for analyzing video detected pictures in real-time monitoring camera. Two conditions were tested in order to determine the falling severity. This paper proposes a vision-based method for falling detection. It analyzes an extracted skeleton to identify human posture. Open Pose network is utilized to obtain skeleton information regarding the human body. The falling can be identified using two parameters: the first is the calculation of the angle formed between the shoulder, hip or any other points (knee, ankle or the heel), the second parameter is the distance calculated during the horizontal falling between the shoulder, hip or any other points such as knee, ankle, heal or the toe. The test results approved that the proposed method outperformed the other state-of-the-art methods in terms of, sensitivity (99.38%), specificity (96%), and accuracy (98%).

Keywords: fall down, skeleton, tracking.

I. INTRODUCTION

In computer vision realm, the understanding and detection of the human behaviors is still one of the challenging and crucial tasks. Its beneficial in various research cases, which include the intelligent surveillance cameras, elder surveillance, video information analysis, and human-computer interactions, [1]-[3]. It may be divided into 2 groups, according to the aims of identification and classification of the normal human behaviors, in addition to detections and warning of abnormal ones [4]-[6]. The elderly individuals have a higher likelihood of being influenced by the falls that can lead to dangerous injuries or even death. Most widespread negative impacts of falls include long-term illnesses and fractures that lead to the independence losses, incapacity, and psychological anxiety concerning the possibility to fall again. Falling leads to putting elderly people at risks for moderate to severe injuries, in addition to burdening them and their families psychologically and financially [7] [8]. In such case, it is quite critical to identify elderly’s falls effectively and immediately for the purpose of providing urgent assistance. To summarize, the individuals who fall and are not capable of requesting help, have to be identified and treated in a very timely manner [9] [10].

Both vision- and sensor-based systems of fall detection are available. Sensor-based systems [3] allow building inexpensive systems with the use of simple sensors; none-the-less, those systems have to be connected to the patient’s body. Manufacturing environment requires having vision-based systems which require no connection to the body [11] [12]. Those devices could be portable and have a micro-controller attached to camera for the real-time processing and analyses, or they can simply be represented by cameras sending video streams to some centralized devices for the operation of the processing. The process of the moving object detection remains one of the difficult tasks. A number of the conventional approaches have accomplished the tasks of detection and classifications. Amongst the present methods, DL with the NNs has been found quite prospective in terms of capability. However, it is rather demanding concerning the required processing time [13]-[15]. This process performs the processing of each one of frames. It utilizes 2-D Pose Estimation with the use of Part Affinity Fields skeleton extraction method for obtaining file of the skeletal data of the individuals on a screen, and that has been captured from surveillance video. The horizontal and vertical coordinate values are a representation of every one of the nodes in system of the planar coordinates. Conditions determine behaviors of the fall and if an individual is able to stand alone after they fall, through the examining of coordinates of head center-to-center of shoulder line between the points of the shoulders, distance between head center and ground, and angle of human body center-line with ground [16] [17].

Employees in the healthcare sector can use the findings of the present study in order to quickly and accurately recognize actions in real-time manner from massive video data sizes from a system of surveillance to swiftly and effectively respond in the case of the emergency. Additionally, it has the ability of improving the logical awareness of the subject of the public safety. Which is why, reliable detection of the human fall in the public situations has social as well as economic impacts. The main aim of this paper is to identify falling down, inside buildings in this paper we propose a novel technique that uses the Open Pose methodology for analyzing video-detected pictures in a real-time monitoring camera, a couple of criteria were tested to determine the falling severity.

The rest of the present study has been arranged as: section II contains the related works. Section III explains the models...
of the system, whereas section IV discusses experimental results and section V includes the conclusions.

II. RELATED WORK

Due to the fact that the cameras are very widespread in the today’s society, they’re used increasingly for the purpose of acquiring helpful information. They’re common tools in the vision-based systems of fall recognition. For the identification or avoiding of the falls, numerous researches utilized the thermal sensors of the depth, cameras (Kinect), RGB cameras, or even a combination of cameras for the monitoring of the head trajectory, body size, or body position variations. J. Zhang et al. [7] Have suggested a human body position fall recognition model that has been referred to as “5-point inverted pendulum model,” which performs the extraction and construction of human posture’s pendulum structure in the complicated natural scenes with the use of improved 2-branch multistage convolutional neural network (M-CNN) model. W. Chen et al. [18] utilized the OpenPose for the purpose of obtaining the skeleton information from human body and for the detection of the fall with the use of 3 fundamental parameters, which are: descent velocity at the hip joint core, externally rectangular ratios of width-to-height and angles of human body centroid with floor. This method has the ability to recognize fall-down action with 97% rate of success. Fann et al. [19] Have proposed an innovative vision-based approach for the detection of the falls through the analysis of the features of the extraction utilized for defining the position of the body. Characteristics have been put in a directed acyclic graph SVM for distinguishing 4 highly associated human postures, which are: sitting, standing, lying and crouching. For fall detection, it performs the counting of number of the occurrences of the postures of laying in limited intervals of the time. Following the majority vote, the verification of immobility determines an event of the fall. Based on experimental results, the recognition of the 4 postures has a general 97.10% accuracy, with 1% only of postures being classified wrongly as lying down position. A system of fall detection could achieve as high as 95.20% accuracy of fall detection on public fall data-set.

Liu et al. [20] have based their research on Kinect. Following an extensive testing, researchers have developed an innovative system of fall identification, which can accurately and quickly detect the human fall. This approach includes 3 parts, which are: depth of the movement targets of the capture of the image, processing of the depth images, and identification of the target movement behaviors. The approach of the detection utilizes sequence of the depth map that has been created with the use of Kinect. The data on the bending and the fall has been obtained then compared in the study. Anti-noise performances of Otsu approach had been utilized for the processing of depth map. It leads to simplifying physical outline extraction. After removing contours, a corrosion approach has been deployed for securing edges. The aspect ratio of the outer rectangle of human, the center of the gravity of the human body, and the degree of the inclination are retrieved thereafter.

Kong et al. [21] have suggested a system for the protection of the elderly, their research had proposed an algorithm for the recognition of the hazardous cases in living rooms. Depth camera in addition to the Canny filter obtained binary image contours. Outline image which resulted has been utilized after that for the identification of the falling objects. It gathers every white pixel in the image of the outline, computes their angles of the tangent vector, and divides them to 15 classes. In a case where most tangent angle values are <45°, a decline has been found. Rafferty et al. The algorithms of computer vision have been combined with sensor of thermal vision that is mounted on ceiling for the detection of the falls. The conventional methods have weaknesses, which new solution in this approach address. However, vision-based fall detection is related to a number of the drawbacks as well, which include the high computational and storage requirements to run real-time algorithm, issues that are related to the privacy, and restricted capture area which may be observed. An initial research of that approach had resulted in encouraging the findings with 68% rate of accuracy. However, it raised concerns as well about the false positive cases [22]. When it comes to the video recognition systems it can be observed that there are some academic works that applied some deep learning techniques such as [26]-[29]. To track the human body with high degrees of accuracy, it could be essential to modify the marker technology and acquisition system. Due to the fact that they are utilized in controlled situations, marker-based tracking systems are technically easy to implement. For instance, lighting and field-of-view conditions can be precisely controlled. There has been a lot of effort in the scientific community to study and implement systems that don't use any marker. In fact, in the field of indoor detection, recent works describe applications where human tracking is performed without the use of markers [30]. Outdoor pedestrian detection is typically more challenging compared to indoor pedestrian detection due to the fact that there are various variables in the outside world and scenes that are to be captured are utterly unpredictable. The latest researches on the marker-based human tracking were conducted, the majority of which deploy the drones [31]. Yet, it's possible to detect and follow a human without using any external markers [32].

III. METHODOLOGY

The proposed method can be divided into 4 steps: (1) get the human body skeleton information using Open Pose, (2) trace the human body with the skeletal information, (3) the angel calculation is conducted as a first decision criterion where the angel among any triplet three points are calculated relative to the central point which is the hip point (4) the second decision criterion is considered by calculating the distance between any two points of the body in both sides relative to the hip and the shoulder referencing points.

As can be shown in Fig.1, at the beginning we capture the images from the human movement video through the motion
detection of the body. Afterwards, the captured images will be processed via the open pose module in order to obtain the skeleton data of the human being where only the main points are considered which are 32 points Fig. 2. The proposed system aims are determining the human body and to discover his/her movement in order to identify if there is any falling down case or not. Via the use of the deep learning i.e. the Open Pose algorithm and the mathematical functions. The system starts working through the capturing of a video clip, a live video or real-time video and recognizing it if it was a human or not through the use of the artificial intelligence and transform this video into a sequence of frames where the images will be processed before selecting the active points of the human body using the Open Pose technology. After the selection of the points, each point is separated into a x and y coordinates. These coordinates are input into a variable in order to deal with them for a later use. In fact, 32 active points of the body will be considered as main points for recognition. At this moment, each angle will be calculated from three points. These points are determined according to the tuning of the design.

![Fig. 2 Active points on the skeleton after capture capturing using open pose](image)

**A. Angel Calculation**

Evaluate the first decision criteria where the angel is calculated in both body sides. Each angle is computed based on three points. It is noteworthy to mention that not only the triplet points from one side are calculated. However, the combinations of the triplet points from both sides can be formed. For example, it might be possible that the falling situation can involve the shoulder and the hip from the right side while at the same time it can also involve a third point situated on the opposite side such as the knee point on the other side. All possible other combinations are taken into account. Indeed, the hip point is considered the central point in each side where all combinations must involve at least one hip point from any side of the body. In Fig.3 the angle combinations are illustrated.

Where in Fig.3, the SR is the right shoulder, H(c) is the hip center, k is the knee, H is heel, and the A is the angle. These are all the points of the right side of the body. The figure shows that there is one angle between the right shoulder, right hip and the right knee. Similarly, there are two angles between the right shoulder, the hip and the right ankle. This is true for all other points at the opposite side.

And the SL is the left shoulder, H(c) is the hip center, k is the knee, H is heel, and the A is the ankle. These are all the points of the left side of the body. The figure shows that there is one angle between the k left shoulder, left hip and the left knee. Similarly, there are two angles between the left shoulder, the hip and the left ankle. This is true for all other points at the opposite side.

![Fig. 1: proposed model and rules](image)
It can be done via the pose estimation and tracking. The majority of these points will be taken into account to calculate the angle among three points where there is a spatial and temporal relationship between these points to determine the status of the falling situation. In this case, the falling will be detected according to specific threshold value. This value was determined as a result of several experimentations and testing more than one database in a direct and real-time manner. There is a time interval that measures the wait time of the falling down. As for the distance, they are measured using the same technique of calculating the angle.

IV. EXPERIMENTAL RESULTS

A. Data of Test Description

The test results depend on the images captured indoor using a web camera. The camera used is the Panasonic HC-MDH2 AVCHD Shoulder Mount Camcorder shoots SD and HD video matched with PAL criteria. It is supporting Full HD 1080 p 50, utilizing AVCHD version 2 compressions at 28 Mb/s. Furthermore, it shoots 1080i 50 records at 24 Mb/s, and higher compression rate enables longer record durations. The camera could further be recording SD footages. Thus, it combines into an SD or HD workflows. This would make it a good candidate to capture images in our proposed system. The samples of the captured images are shown in the Fig.4.

B. Evaluation Measures

The falling tests show the next expected outcomes. The falling event occurs in 1st instance, and the suggested method appropriately discovered the falling among the other postures which is called the True Positive (TP). In 2nd instance, the falling didn’t occur. Nonetheless, the method erroneously diagnosed it as falling case which is called the False Positive.
(FP). However, the method didn’t recognize it as falling in the 3rd instance which is called the True Negative (TN). As for the last case, the falling didn’t happen, and the method didn’t diagnose the falling which is called the False Negative (FN) [26]-[29].

True positive (TP): a falling happened while it is recognized. A false positive (FP) a falling did not happen while, the instrument detects it as a falling. True negative (TN): a falling did not happen while the instrument did not consider it as a falling.

A false negative (FN) a falling happens while is not recognized by the device.

For evaluating the reactions of those cases, couple standards have been shown.

The capability of detecting a falling case is named as sensitivity:

$$SE = \frac{TP}{(TP+FN)}$$  \hspace{1cm} (1)

The specificity is the capability of the system of identifying merely a falling:

$$SP = \frac{TN}{(TN+FP)}$$  \hspace{1cm} (2)

Accuracy is referent to the capability of distinguishing the falling from the non-falling cases.

$$AC = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$  \hspace{1cm} (3)

Based on the aforementioned criteria, the next test results were gained for this paper are, sensitivity (99.38%), specificity (96%), and last is accuracy (98%).

C. Comparison

Table I lists a compared analysis between different methods used for the falling detection. The Open Pose [12] [24] [18] technique was utilized for the recognition of images captured by camera; visions-based techniques were more appropriate than others methods. The skeleton information of human body is obtained using Open Pose, which is both accurate and easy. The proposed technique was better than other state-of-the-art methods as it is easy to perform arithmetic operations. The optimal accuracy was obtained, and the tests approved this as it would not be depending on a trained dataset. Through this comparison, it is expected to approve that the suggested method outperformed the other methods on the real-time datasets.

In this section the test results of the proposed falling down detection system has been given in details. In table II the angle test using different values.

| Method                          | Accuracies (%) | SE  | SP  |
|--------------------------------|----------------|-----|-----|
| Sungil Jeong [12]              | 98.7           | none| none|
| Qingzhen Xu [24]               | 91.7           | none| none|
| Weiming Chen [18]              | 97             | none| none|
| [25] Ali, Mohammed Abduljabbar| 97.8           | none| none|
| Proposed                       | 98%            | 99% | 96% |

Table II has the selected angle number and its relationship with the precision value. As can be seen in the table that several experimentations have been conducted using different angles, in which we started from the angle 110 and we tested the human body falling down. After that we reduced the angle’s values gradually until we found the best accuracy in the angle >90.

V. CONCLUSIONS

Falling down is one of the worst consequences that might occur to elderly or people who suffer from some physical or mental health problems. Therefore, technology can give active solutions to such a severe problem. In this paper, we presented a technique that used Open Pose methodology for analyzing video detected pictures in real-time monitoring camera. Two parameters were tested in order to determine the falling severity. The first is to calculate the angels in particular points of the body i.e. the shoulder, hip, knee, and the ankle while the second parameter depends on calculating the distances between these points including the toe points from each body side. It was observed from the test results in this paper that not all angles calculated have the significance to determine the fall. For instance, the shoulder-hip-knee is considered as the most determinant angle that can clearly describe this movement. It was also observed that when the ankle in
invisible by the camera that would make it uneasy to determine the fall. Therefore, there is about 2% the fall cases can be misclassified. In general, the test results approved that the proposed method outperformed the other state-of-the-art methods in terms, sensitivity (99.38%), specificity (96%), and last is accuracy (98%).

REFERENCES

[1] W. M. Salih Abedi, I. Nadher, and A. T. Sadiq, “Modification of deep learning technique for face expressions and body postures recognitions,” Int. J. Adv. Sci. Technol., vol. 29, no. 3 Special Issue, pp. 313–320, 2020.

[2] A. T. Saaqeq, E. K. Jabbar, and N. J. Ibrahim, “Human detection and Recognition System,” vol. 5, no. 1, pp. 28–51, 2015.

[3] W. M. S. Abedi, I. Nadher, and A. T. Sadiq, “Modified deep learning method for body postures recognition,” Int. J. Adv. Sci. Technol., vol. 29, no. 2, pp. 3830–3841, 2020.

[4] Learning a Deep Model for Human Action Recognition from Novel Viewpoints, Rahmani et al. - IEEE Transactions on Pattern Analysis and Machine Intelligence - 2018, doi: 10.1109/TPAMI.2017.2691768.

[5] Action recognition based on motion of oriented magnitude patterns and feature selection,Phan et al. - IET Computer Vision - 2018, doi: 10.1049/iet-cvi.2017.0282.

[6] Sparse Representation and Weighted Clustering Based Abnormal Behavior Detection,Jin et al. - 2018 24th International Conference on Pattern Recognition (ICPR) - 2018, doi: 10.1109/ICPR.2018.8545794.

[7] Human Fall Detection Based on Body Posture Spatio-Temporal Evolution,Zhang et al. - Sensors - 2020, doi: 10.3390/s20030946.

[8] Design and Implementation of Smart City Applications Based on the Internet of Things,Alikabi & Als Jasim - International Journal of Interactive Mobile Technologies (IJIM) - 2021, doi: 10.3991/ijim.v15i13.22331.

[9] Pose Estimation of Objects Using Digital Image Processing for Pick-and-Place Applications of Robotic Arms,Hameed et al. - Engineering and Technology Journal - 2020, doi: 10.30684/etj.v38i5a.518.

[10] Anticipating Atrial Fibrillation Signal Using Efficient Algorithm,Al-dabag et al. - International Journal of Onine and Biomedical Engineering (IJOE) - 2021, doi: 10.3991/ijoec.v17i02.19183.

[11] Sleeping posture recognition using fuzzy c-means algorithm Hsiao et al. - BioMedical Engineering OnLine - 2018, doi: 10.1186/s12938-018-0584-3.

[12] Human-skeleton based Fall-Detection Method using LSTM for Manufacturing Industries,Jeong et al. - 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC) - 2019, doi: 10.1109/ITC-CSCC.2019.8793342.

[13] A Video based Vehicle Detection, Counting and Classification System,Memon et al. - International Journal of Image, Graphics and Signal Processing - 2018, doi: 10.5815/ijigps.2018.09.05.

[14] A. Raychaudhuri, S. Maitry, A. Chakrabarti, and D. Bhattacharjee, Detection of Moving Objects in Video Using Block-Based Approach. Springer Singapore, 2019.

[15] An Efficient Approach for Detecting and Classifying Moving Vehicles in a Video Based Monitoring System,Mahmood & Saud - Engineering and Technology Journal - 2020, doi: 10.30684/etj.v38i6a.438.

[16] Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields,Cao et al. - 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - 2017, doi: 10.1109/CVPR.2017.143.

[17] Hand Gesture Recognition of Static Letters American Sign Language (ASL) Using Deep Learning,Abdulhussein & Raheem - Engineering and Technology Journal - 2020, doi: 10.30684/etj.v38i6a.535.

[18] Fall Detection Based on Key Points of Human-Skeleton Using OpenPose,Chen et al. - Symmetry - 2020, doi: 10.3390/sym1205074.

[19] Fall detection via human posture representation and support vector machine,Fan et al. - International Journal of Distributed Sensor Networks - 2017, doi: 10.1177/1550147717707418.

[20] Human body fall detection based on the Kinect sensor,Liu et al. - 9th International Conference on Image and Signal Processing (CISP) - 2015, doi: 10.1109/CISP.2015.7407906.

[21] Fall detection for elderly persons using a depth camera,Kong et al. - 2017 International Conference on Advanced Mechatronic Systems (ICAMechS) - 2017, doi: 10.1109/ICAMechS.2017.8316483.

[22] Fall Detection Through Thermal Vision Sensing,Rafferty et al. - Ubiquitous Computing and Ambient Intelligence - 2016, doi: 10.1007/978-3-319-48799-1_10.

[23] K.A. Hussein, and Z. Tariq Al-Ani. "Iraqi License Plate Recognition Based on Neural Network Technique." Journal of Physics: Conference Series. Vol. 2322. No. 1. IOP Publishing, 2022.

[24] R.F. Ghani, et al. "Key Frames Extraction Using Spline Curve Fitting for Online Video Summarization." 2019 11th Computer Science and Electronic Engineering (CEEC). IEEE, 2019.

[25] A.W.S. Ibrahim, and S. Mhamed Taher. "Inserting virtual static object with geometry consistency into real video." Journal of Physics: Conference Series. Vol. 1530. No. 1. IOP Publishing, 2020.

[26] S.A. Mahmood, A. Monsheid Abid, and W. Abdul Khuder Naser. "Contextual Anomaly Detection Based Video Surveillance System." 2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE). IEEE, 2021.

[27] A. M. Jalil, F.I Sahib Hasan, and A.A.Hesham "Speaker identification using convolutional neural network for clean and nois speech samples." 2019 first international conference of computer and applied sciences (CAS). IEEE, 2019.

[28] M.B.Tuieb, A. Shakir Mahmood, and Farah Neamah Abbas. "Fusion of DWT as a Novel Approach for Efficient, Secured and Reversible Video Steganography," Journal of Physics: Conference Series. Vol. 2322. No. 1. IOP Publishing, 2022.

[29] S.K. Salah, A. Othman Khalaf, and Waleed Rasheed. "Safe Medical Radiography Compressive Sensing based on Discrete-Rational-Random-Transform." Journal of Green Engineering 10 (2020): 6815-683

[30] Collecting pedestrian trajectories,Boltes & Seyfried - Neurocomputing - 2013, doi: 10.1016/j.neucom.2012.01.036.

[31] M. Wani, F. Bhat, S. Afzal, and A. Khan, "Advances in Deep Learning, Book." Springer, (2019),10.1007/978-981-13-6794-6.

[32] Computer vision and deep learning techniques for pedestrian detection and tracking: A survey,Brunetti et al. - Neurocomputing - 2018, doi: 10.1016/j.neucom.2018.01.092.

237