Artificial Intelligence Advances
https://ojs.bilpublishing.com/index.php/aia

ARTICLE

Elderly Fall Detection by Sensitive Features Based on Image Processing and Machine Learning

Mohammad Hasan Olyaei[1*] **Ali Olyaei**[2] **Sumaya Hamidi**[3]

1. Faculty of Electrical Engineering, Sadjad University of Technology, Mashhad, Iran
2. Department of Computer Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
3. ARIVET Project, Mashhad, Iran

ARTICLE INFO

Article history
Received: 29 January 2022
Revised: 31 March 2022
Accepted: 2 April 2022
Published: 12 April 2022

Keywords:
Human fall detection
Machine learning
Computer vision
Elderly

ABSTRACT

The world's elderly population is growing every year. It is easy to say that the fall is one of the major dangers that threaten them. This paper offers a Trained Model for fall detection to help the older people live comfortably and alone at home. The purpose of this paper is to investigate appropriate methods for diagnosing falls by analyzing the motion and shape characteristics of the human body. Several machine learning technologies have been proposed for automatic fall detection. The proposed research reported in this paper detects a moving object by using a background subtraction algorithm with a single camera. The next step is to extract the features that are very important and generally describe the human shape and show the difference between the human falls from the daily activities. These features are based on motion, changes in human shape, and oval diameters around the human and temporal head position. The features extracted from the human mask are eventually fed in to various machine learning classifiers for fall detection. Experimental results showed the efficiency and reliability of the proposed method with a fall detection rate of 81% that have been tested with UR Fall Detection dataset.

1. Introduction

The number of elderly people (people with the age of ≥ 65 years) is increasing rapidly [1]. Based on the UK Office for National Statistics (ONS), more than 8.6 million people aged 65 or higher will exist in the next 5 decades [2]. The population of the elderly people in Iran according to the statistics of the United Nations in 2019 is 5272000 and it is predicted that this population will reach 8849000 in 2030 [3]. Existing health care systems face many shortcomings in meeting the care needs of the elderly and disabled. An independent life supported by self-monitoring is a suitable solution to the growing population of the elderly with limited budgets to care for them. Current technologies such as environmental intelligence can help the elderly and disabled to live longer in their homes. Falls

*Corresponding Author:
Mohammad Hasan Olyaei,
Faculty of Electrical Engineering, Sadjad University of Technology, Mashhad, Iran;
Email: mh.olyaei123@sadjad.ac.ir

DOI: https://doi.org/10.30564/aia.v4i1.4419

are a relatively common occurrence among the elderly, which can have significant health consequences. Studies have shown that 28% to 34% of the elderly fall at least once a year [4] and falls are the second leading reason for accidental death in more than 87% of the elderly [5]. Falling is one of the main causes of various types of injuries, such as fractures, muscle bruises or brain injuries in the elderly [6]. It is possible that the injured person will remain lying on the ground for a significant period of time before receiving assistance after the fall. Fall detection and alerting systems are unable to prevent falls.

However, these systems can ensure that the person who falls in quickly receives the help quickly, and this can reduce the effects of the accident [7]. Therefore, significant attention should be paid to develop the automatic human fall detection systems [8].

Currently, some of the advanced technologies for alarming of a fall accident are manually, inactive and based on devices that a person is required to push a button during an accident and wears it as a pendant or bracelet. This device is recommended for the elderly who live alone, as well as vulnerable people who have a history of falls and need to be worn constantly throughout the day. In the event of a fall, the victim is forced to push a button to generate a signal that is sent to crisis telecare center. But this approach depends on the level of awareness of the person at the time of the accident, it is necessary to have full mental awareness and cognizant of the required response, otherwise this device is simply decorative. Automatic human fall detection systems are needed to monitor vulnerable people susceptible to falling. A number of new methods of fall detection were established over the past few years to enable seniors to live in their own homes independently while also having acceptable monitoring tools. And this plan requires large-scale monitoring strategy. Ideally, a comprehensive detection system can do this while conserving the individuals' privacy being monitored. Numerous sensors are used to detect falls, such as cameras, infrared sensors, sound and pressure sensors, accelerometers and others. The development of these sensors has also helped improve fall detection algorithms, especially by reducing the price of cameras. And the development of novel image acquisition like cameras (Xtion, Kinect, etc.). Also, increasing the scope of computer vision has facilitated the deployment and development of computer visual programs. The use of in-depth data due to having an additional dimension has significantly increased the level of reliability and accuracy (in fact it has 2.5 dimensions). The high-level data provided by the OpenNI and Kinect SDK also simplify the division of the human body. With the arrival of Kinect technology in 2011, new research into fall detection has been utilizes using the features of depth sensors like 3D analysis with privacy, thus, more acceptable solutions were provided using deep data (Kinect) in fall detection system.

Research on vision-based monitoring systems to understand human behavior has steadily grown recently [9]. Computer-based image processing systems have provided promising new solutions to help older people stay at home [10]. One of the applications based on video surveillance is the analysis of human behavior and the detection of their unusual behavior.

According to the latest report of the World Health Organization (WHO), every year about 646,000 people die due to falls in the world. The fall is the second leading cause of death due to injury worldwide [11] and at least 1/3 of the elderly fall one or more times a year [12].

There are different types of falls, depending on the direction of the body when falling, and each of them has a different movement. Hence, a detection algorithm is needed to identify different types and minimize false detections. Also, considering privacy issues may negatively affect the choice of a vision-based fall detection system.

In this paper, we focus on the automatic human fall detection from data generated by the camera (or Webcam). This will be based on a monitoring system with low-cost video cameras to track elderly people who living independently.

2. Fall Detection Methods Based on Image Processing

Computer vision-based techniques can provide information about falls and other daily life behaviors. Fall events can be found using an automatic fall detection algorithm via intelligent monitoring systems [13]. Much work was done in the field of visual sensors in fall detection utilizing single, multiple and omni-directional cameras [14-16]. Presently, in-depth cameras such as Microsoft Kinect 2.0 and ASUS Xtion Pro Live was also utilized to detect falls. Kinect is a motion sensing technology that uses an RGB camera and a depth sensor to track a moving object in 3D [17-19]. Depth cameras like the Microsoft Kinect have been used to detect falls. Depth cameras are used to identify human activities. For example, the authors propose a spatial approach based on the skeleton to describe spatiotemporal aspects of a human activity using three-dimensional in-depth images [20]. Two features are calculated from the 3D joint position concluded the spatial and the temporal aspects of the activity. These features are used as input to the random tree algorithm to teach the model of human activity. The trained classified model achieved an accuracy of 80.92%.

The accuracy of a fall detection system must be such that it can be trusted in a real environment. Activities such as lying down may be misdiagnosed and lead to unnecessary alert generation. These movements are described as activities of daily living (ADL) or not falling. Therefore, algorithms must minimize the impact of such events, to increase system reliability, and to reduce false positive (FP).

There are different types of falls, depending on the direction of the body when falling, and each has a different movement. Therefore, a detection algorithm is needed to detect different types and minimize false detections. Also, considering privacy issues may affect negatively the choice of a vision-based fall detection system. A surveillance system should be able to hide some recognizable features, such as people's faces if needed.

Existing research on fall detection systems is normally based on limited amounts of human fall data (which are not real events, and are generated in a tuned or trained manner).

Recording fall events in real life is a complex process that requires costly infrastructure. Several research groups have documented falls in hospitals and nursing homes, but such data is not limited and available to the public due to copyright and privacy restrictions. An alternative approach is used, such as video recordings of human-simulated falls, in which participants try to guide researchers. Data collection is very common in data science, and collecting fall detection videos faces the following issues:

1) Demographics: For example, fall data and daily activity data, people should be of different ages. Likewise, these individuals with various physical features (e.g. weight, height, delivery) should be examined. Finally, the behavioral characteristics of these individuals, such as gait patterns, should be recorded and examined.

2) The quality, quantity and availability of the simulated samples are very important and should be in different places of the house.

3) Scene conditions: Visual obstruction is another issue that needs to be addressed. Dining tables, chairs, sofas and other accessories can act as a barrier in the home scene.

So far, previous methods have often achieved good accuracy by connecting gadgets and devices to the elderly, such as necklaces or mobile phone. But most seniors are dissatisfied with connecting extra gadgets or forget to always have them. For this reason, vision-based methods and the use of a camera or webcam can be a better alternative for these people.

In image processing, feature selection is one of the most important parameters to increase accuracy. It must also have a high processing speed for real-time execution. In our proposed algorithm, the important features that are selected and combined show that it has both speed and accuracy.

3. Our Proposed Algorithm

The proposed fall detection system in this paper is divided into four stages: data collection, foreground segmentation (or moving Human detection), feature extraction and fall detection.

Foreground segmentation is performed to detect moving objects [21]. The steps of the model training algorithm are shown in Figure 1 and the model test in Figure 1-B.

As shown in Figure 1, the videos are first read from the database and then the next steps are performed for all frames of all videos. In each frame, the mask of the person is first identified using the background estimation and Frame Difference methods (BMaker_oly.py) and the background subtraction method. The subtraction method is mainly aimed to convert frame images to a gray scale at any time and then separated from the background image at the pixel level for producing a binary image such as Equation (1) [22].

$$REk(X, Y) = FRk(X, Y) - BA(X, Y). \quad (1)$$

REK is the absolute difference between background and current images. FRK represents the current image and BA denotes the background image. The main idea of the

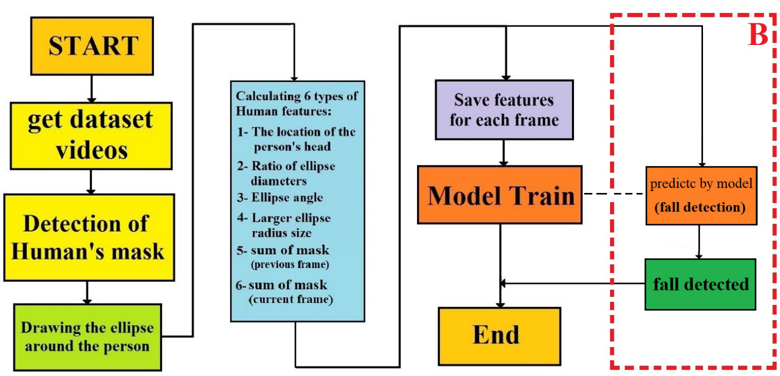

Figure 1. Our proposed algorithm - train and test (B section)

background estimation method (BMaker_oly.py) is that 12 frames of video are randomly selected and images are averaged between them. This averaging causes the moving objects to fade or even be removed and the background to remain. This background can be used to detect a moving object. Figure 2 shows a sample background estimate output for a video.

After this step, an ellipse is estimated around the human mask as shown in Figure 1 using OpenCV library and python language. From the detected and elliptical mask, 6 special features are extracted. These features are head position, ellipse diameter ratio, ellipse angle, larger ellipse radius, and sum of mask, respectively. The proposed method of this detection is the extraction of motion and shape features from previous observations of human falls compared to other daily activities, including severe changes in body shape and form, as well as the study of sudden and unexpected movements.

Many usual fall detection algorithms are oriented by the higher acceleration of fall events have than a daily life activity. Ultimately, the extracted properties are utilized as input vectors for existing classifiers in machine learning for categorizing fall and non-fall events.

Figure 3 shows the sum of mask_final changes of person pixels and the main label for videos of fall-1 and fall-4. It is clear that the sum of motion in frames of 110 and 40 has changed suddenly.

In these two diagrams, it is quite clear that in sum of motion mask, at the moment of the fall, creates a clear peak. This means that when we have a peak in sum of motion mask, it means that we have most likely fallen, because in this frame, there was a sudden movement. This proves that the sum of mask can be used as a very good feature.

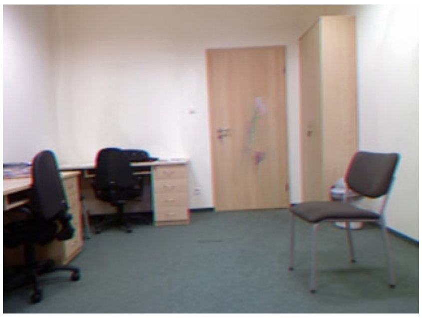

Figure 2. Sample output of background estimation

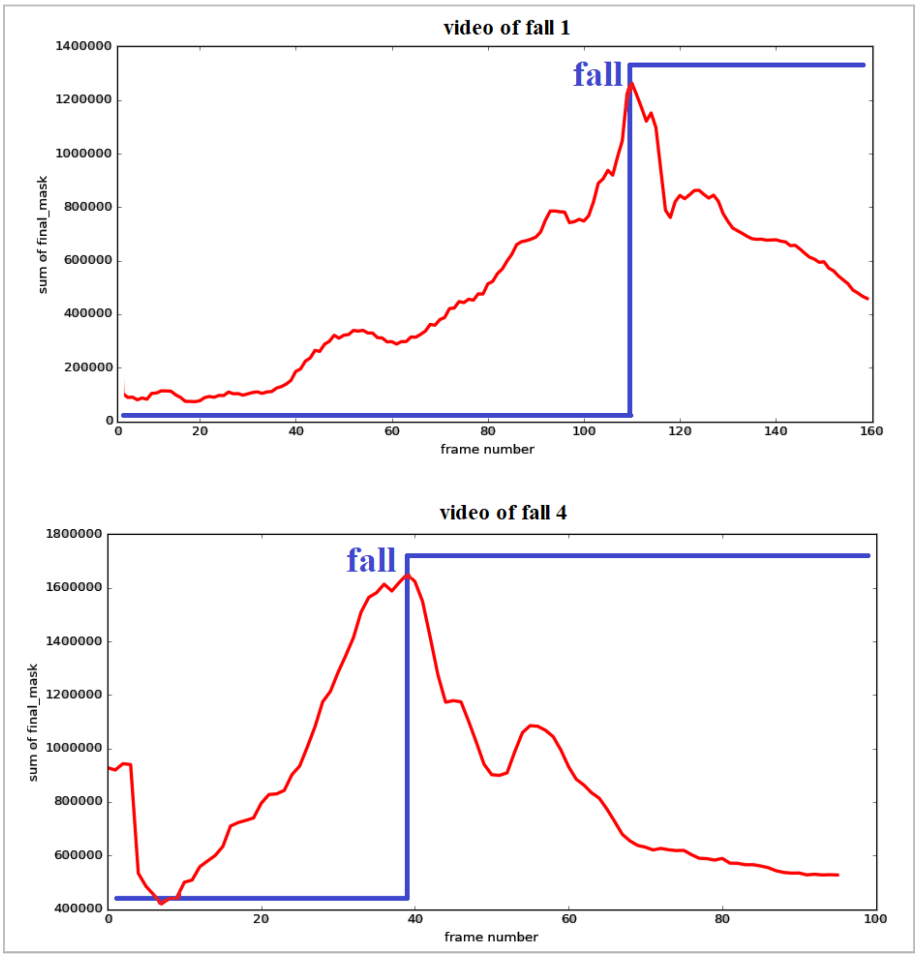

Figure 3. Sum of motion and label

4. Simulation Results

In this paper, we have used UR dataset [23] and the powerful Python language.

Seventy (30 falls + 40 activities of daily living) sequences are included in this dataset. Using 2 Microsoft Kinect cameras and correspondent accelerometric data, fall events are recorded. Only one device (camera 0) and accelerometer were used to record ADL events. To collect sensor data, x-IMU (256Hz) and PS Move (60Hz) devices were used.

The UR Fall Detection dataset includes video frames from the front and top of the head taken by two Kinect cameras. The first camera is installed at a height of 1 meter from the floor and the second camera is located at the ceiling at a height of 3 meters. There are 30 fall scenarios recorded by two Kinect sensors and an accelerometer. Daily activities include walking, sitting, bending over, lifting an object off the floor, and lying down. Normal activities (30 scenarios) using a Kinect sensor that is parallel to the floor and collects 10 sequences of activities such as falling, such as lying on the floor fast and lying on the bed. Five participants were asked to perform daily activities and fall simulations. The dataset was recorded at 30 frames per second. In video sequences containing fall events, the number of images is 3,000, but the number of images in video sequences with ADLs is 10,000. Falling in different directions is simulated according to the camera view. Various types of fall accidents have been recorded, including falling backward, forward, and sideways. Participants performed the fall while standing and sitting on a chair. Figure 4 shows examples of frames taken from the UR fall detection dataset. Only RGB camera video data was used for the suggested fall detection approach.

The dataset is organized as follows. Sequence of depth and RGB images for camera 0 and camera 1 are contained in each row, that we only used its camera 0 videos and RGB section of each frame.

In general, three labels are recorded for each frame. If the person is normal, the label is –1. If a person starts to fall, the label=0. Finally, after a complete fall, the label is 1. You can see an example of these three situations in the Figure 4.

Machine learning-based methods compared in this paper include support vector machine (SVM), k-nearest neighbors (KNN), Adaptive Boosting (Adaboost), Multiple-layer perceptron neural network (MLPNN) and Elman neural network (ENN).

In this paper, to more accurately detect the human mask, the mask from the background estimate and frame difference and the motion mask from the subtractor model (changes in a person's movement) are added together logically. For example, look at Figure 5.

Figure 5A is the result of detecting and drawing an ellipse, 5B is the result of motion changes, 5C is the result of background estimation, and finally Figure 5D is the result of the logical adding of two masks.

Figure 6 shows an example of the implementation of the steps in Figure 1.

In the background subtraction method, if the amount of brightness due to the pixel difference is less than the threshold value, the object is considered as the background pixel and the value 0 in the binary image is assigned to it. Otherwise, the pixel is taken as the background and assigned a value of 1 in Equation (2) [24].

$$D_K(X,Y) = \begin{cases} 0 & background \quad R_K(X,Y) > T \\ 1 & target \quad R_K(X,Y) < T \end{cases} \tag{2}$$

By applying the proposed algorithm and using 5 machine learning methods, the results of performance evaluation of the models are shown in Table 1. By observing and comparing the methods with each other, it can be concluded that the highest accuracy is for Adaboost method and it shows in Figure 7.

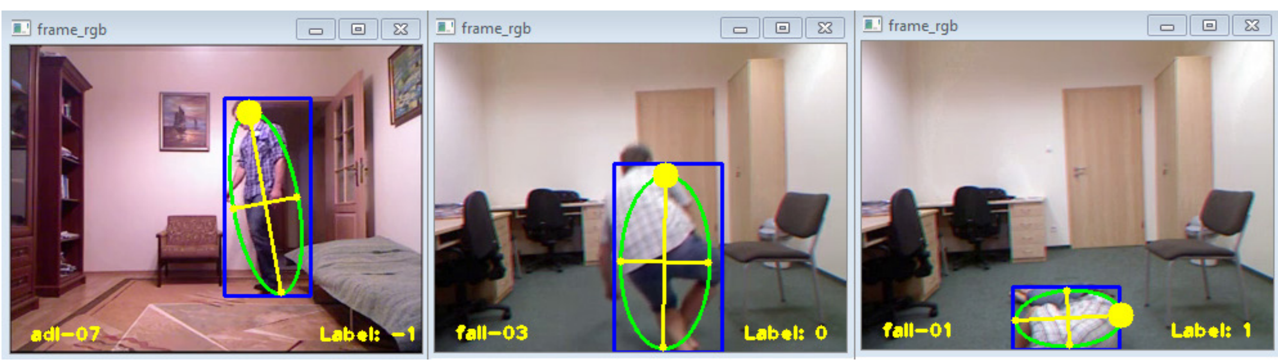

Figure 4. sample of three classes, "–1" means person is not lying, "0" is temporary pose, "1" means person is lying on the ground

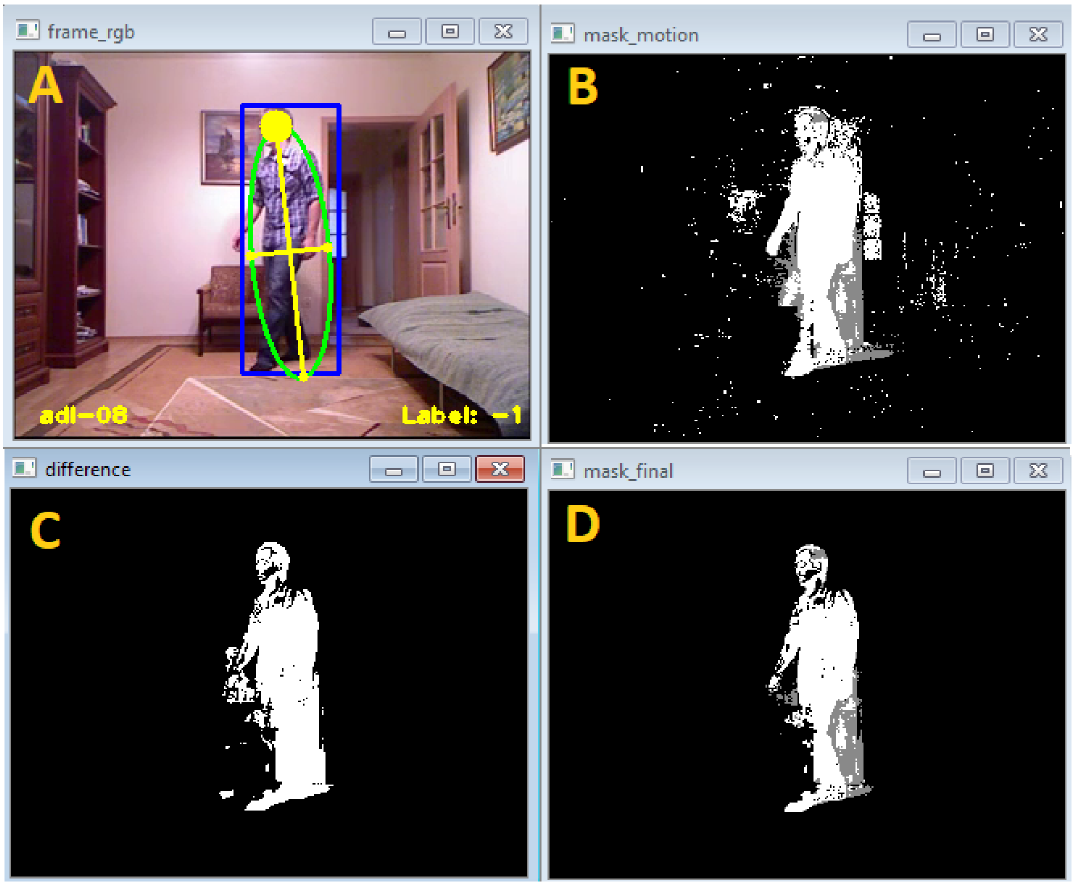

Figure 5. A) output image, B) motion mask, C) Human mask by Frame difference, D) final Human mask

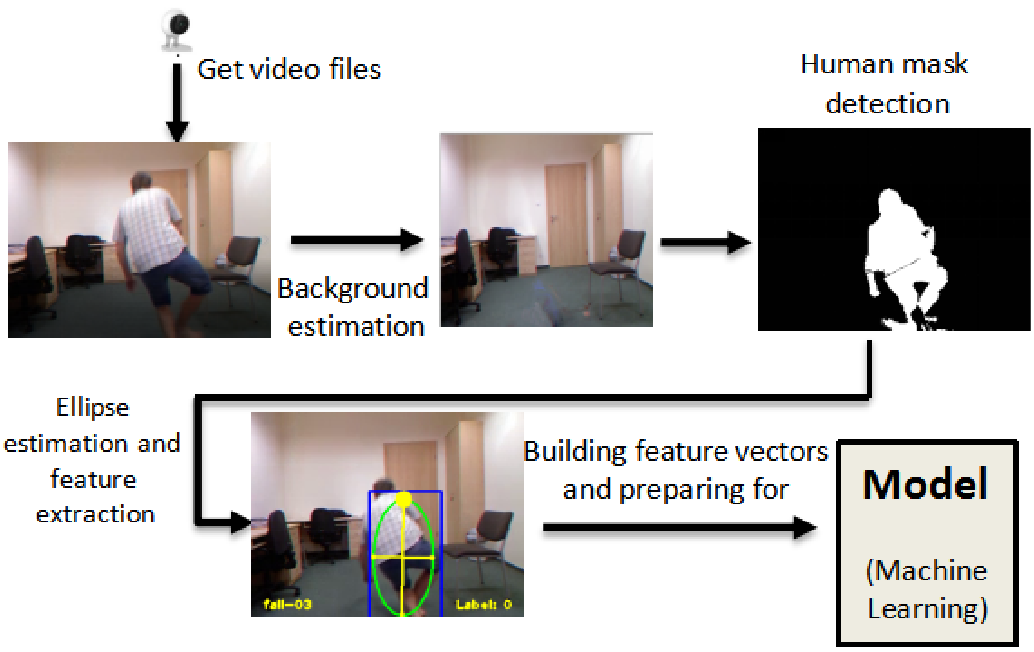

Figure 6. An example of the output of this algorithm

Table 1. Detection results of the proposed algorithm using UR dataset

Method	SVM	KNN	MLP	ENN	Adaboost
Accuracy	70%	65%	69%	54%	81%

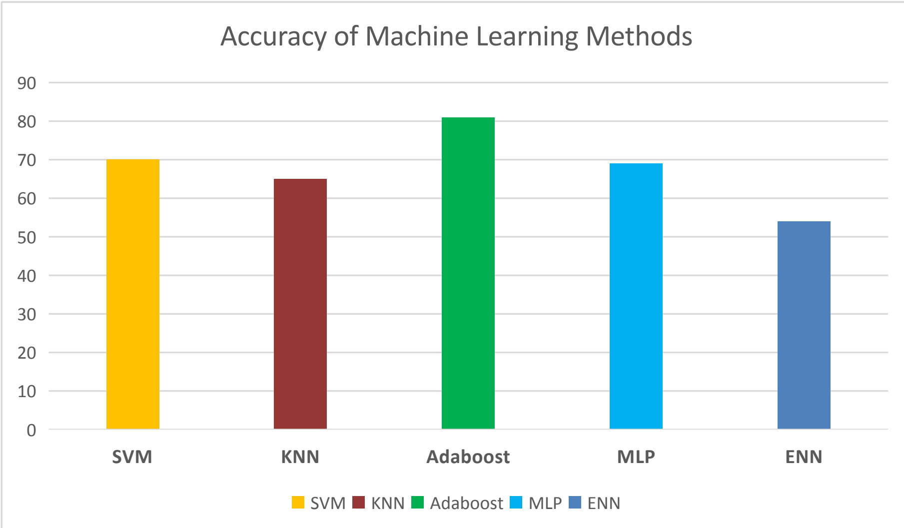

Figure 7. Accuracy of machine learning methods with proposed algorithm for fall detection using UR dataset

5. Conclusions

The algorithm introduced in this paper is a new purpose to fall detection automatically to support a better life for the elderly. The features of the mask extracted from the human play an essential duty in the robustness and effectiveness of human fall detection. Our main focus and purpose in this paper is to combine the features extracted from moving human, which are suitable for real-time fall detection only with images received from the webcam. By combining features such as the coordinates of the human head, various parameters of the estimated ellipse around the human, and the motion features of the human, we have created a very suitable feature vector that demonstrates high accuracy with machine learning algorithms. This type of extraction and combination of features and the use of machine learning models create a new algorithm. Experimental results in Table 1 show that the proposed fall detection algorithm is one of the best fall detection methods in the literature. The advantage of this method over the previous methods is the high speed in real-time image processing, and the reason is the use of easy steps in feature extraction and processing. In addition, there is no need to connect gadgets and devices to the elderly, and it can be used only by placing a camera or webcam in the place of movement. It is considered that the main reason for the success of the method is the strong and effective selection features to distinguish between fall and non-fall activities. The highest accuracy with the proposed algorithm for detecting and combining features and using the Adaboost method is 81%. The results of Table 1 show that this method can also be used in practice. This suggests that the proposed fall detection method can effectively predict the fall.

Conflict of Interest

There is no conflict of interest.

References

[1] Wu, G., 2000. Distinguishing fall activities from normal activities by velocity characteristics. Journal of Biomechanics. 33(11), 1497-1500.
DOI: https://doi.org/10.1016/s0021-9290(00)00117-2

[2] Nash, A., 2019. National population projections: 2018-based. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections

[3] McNicoll, G., 2002. World Population Ageing 1950-2050. Population and development Review, 28(4), 814-816. https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf

[4] Leone, A., Diraco, G., Distante, C., et al., 2008. multi-sensor approach for people fall detection in home environment. Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications-M2SFA2. https://hal.inria.fr/inria-00326739

[5] Diraco, G., Leone, A., Siciliano, P., 2010. An active vision system for fall detection and posture recognition in elderly healthcare. 2010 Design, Automation & Test in Europe Conference & Exhibition(DATE 2010). pp. 1536-1541. DOI: https://doi.org/10.1109/DATE.2010.5457055
[6] Perner, P., Imiya, A., 2011. Machine learning and data mining in pattern recognition. Springer-Verlag Berlin Heidelberg. pp. 1-264. https://link.springer.com/book/10.1007/978-3-642-23199-5
[7] Chen, Y.T., Lin, Y.Ch., Fang, W.H., 2010. A hybrid human fall detection scheme. 2010 IEEE International Conference on Image Processing. pp. 3485-3488. DOI: https://doi.org/10.1109/ICIP.2010.5650127
[8] Xiang, Y.J., Arora, J.S., Abdel-Malek, K., 2010. Physics-based modeling and simulation of human walking: a review of optimization-based and other approaches. Structural and Multidisciplinary Optimization. 42(1), 1-23. DOI: http://dx.doi.org/10.1007/s00158-010-0496-8
[9] Gjoreski, H., Lustrek, M., Gams, M. 2011. Accelerometer placement for posture recognition and fall detection. Seventh International Conference on Intelligent Environments. pp. 47-54. DOI: https://doi.org/10.1109/IE.2011.11
[10] Bergen, G., Stevens, M. R., Burns, E. R. 2016. Falls and fall injuries among adults aged 65 years United States. MMWR Morbidity and mortality weekly report. 65(37), 993-998. DOI: https://doi.org/10.15585/mmwr.mm6537a2
[11] Ma, X., Wang, H.B., Xue, B.X., et al., 2014. Depth-based human fall detection via shape features and improved extreme learning machine. IEEE journal of biomedical and health informatics. 18(6), 1915-1922. DOI: https://doi.org/10.1109/JBHI.2014.2304357
[12] Jansen, B., Deklerck, R., 2006. Context aware inactivity recognition for visual fall detection. Pervasive Health Conference and Workshops. pp. 1-4. DOI: https://doi.org/10.1109/PCTHEALTH.2006.361657
[13] Nait-Charif, H., McKenna, S.J., 2004. Activity summarisation and fall detection in a supportive home environment. Proceedings of the 17th International Conference on Pattern Recognition. 4(1), 323-326. DOI: https://doi.org/10.1109/ICPR.2004.1333768
[14] Dai, J.P., Bai, X.L., Yang, Zh.M., et al., 2010. Mobile phone-based pervasive fall detection. Personal and Ubiquitous Computing. 14(7), 633-643. DOI: http://dx.doi.org/10.1007/s00779-010-0292-x
[15] Dubitzky, W., Granzow, M., Berrar, D., 2007. Fundamentals of data mining in genomics and proteomics. Springer Science & Business Media. https://link.springer.com/book/10.1007/978-0-387-47509-7
[16] Hazelhoff, L., Han, J.G., 2008. Video-based fall detection in the home using principal component analysis. International Conference on Advanced Concepts for Intelligent Vision Systems. pp. 298-309. DOI: http://dx.doi.org/10.1007/978-3-540-88458-3_27
[17] Leclercq, S., 2000. In-company same-and low-level falls: From an understanding of such accidents to their prevention. International journal of industrial ergonomics. 25(1), 59-67. DOI: https://doi.org/10.1016/S0169-8141(98)00095-X
[18] Xiang, Y.J., Arora, J.S., Abdel-Malek, K., 2010. Physics-based modeling and simulation of human walking: a review of optimization-based and other approaches. Structural and Multidisciplinary Optimization. 42(1), 1-23. DOI: https://doi.org/10.1007/s00158-010-0496-8
[19] Pai, Y.Ch., Patton, J., 1997. Center of mass velocity-position predictions for balance control. Journal of Biomechanics. 30(4), 347-354. DOI: https://doi.org/10.1016/S0021-9290(96)00165-0
[20] Popescu, M., Li, Y., Skubic, M., et al., 2008. An acoustic fall detector system that uses sound height information to reduce the false alarm rate. 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 4628-4631. DOI: https://doi.org/10.1109/IEMBS.2008.4650244
[21] Hartikainen, S., Lönnroos, E., Louhivuori, K., 2007. Medication as a risk factor for falls: critical systematic review. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 62(10), 1172-1181. DOI: https://doi.org/10.1093/gerona/62.10.1172
[22] Yazdi, M., Bouwmans, T., 2018. New trends on moving object detection in video images captured by a moving camera: A survey. Computer Science Review. 28(1), 157-177. DOI: https://doi.org/10.1016/j.cosrev.2018.03.001
[23] Kwolek, B., Kepski, M., 2014. Human fall detection on embedded platform using depth maps and wireless accelerometer. Computer Methods and Programs in Biomedicine. 117(3), 489-501. DOI: https://doi.org/10.1016/j.cmpb.2014.09.005
[24] Jadhav, M.J.J., 2014. Moving object detection and tracking for video survelliance. International Journal of Engineering Research and General Science. 2(4), 372-378. https://ijergs.org.managewebsiteportal.com/files/documents/MOVING-46.pdf