Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 1, January 2021: 528-541

An Efficient LSTM and SVM based Fall Event Detection System for Elderly People

S. Vijaya Kumar

Research Scholar, [Reg No.:17231172161017]
Department of Computer Science,
Rani Anna Government College for Women, Tirunelveli
svjaykumar@gmail.com

Dr. R. Shenbagavalli

Associate Professor, Department of Computer Science,
Rani Anna Government College for Women, Tirunelveli.
rtillam3000@yahoo.com

Abstract

In recent days, there is a lack of time and energy for the children to care for the elderly people, thus the elderly people monitoring system is developed to take care of them. This system is developed especially for monitoring the living conditions of elderly people in real-time. In the future, with this proposed system one can monitor the elderly by using the alarm function. Here mainly the falling detection is done, which is one of the major issues in the elderly monitoring systems. The objective is to identify the fall event as normal and non-fall events as abnormal for monitoring elderly people. In this research paper, a deep learning mechanism is used to deduct the fall and build a fall identification system that is efficient. The Long ShortTerm Memory (LSTM) recurrent neural networks are one of the deep learning models used for feature extraction that is used for automatically identifying an elderly person's unusual behaviors. The proposed method is tested on UR Fall, Weizmann and KTH datasets.Its performance is evaluated using Precision, Recall, F-score and the accuracy of classifications. Finally, the result is compared with other deep learning algorithms. The investigational result of proposed work proved that its supremacy by achieve 100% accuracy for the UR Fall dataset, 95% for the Weizmann dataset and 96% for the KTH dataset which is more efficient than the other existing algorithms.

Keywords: LST, SVM, UR Fall, Classification, Sensing, Video Abstraction.

1. Introduction

Recognizing activities provide data about the level of the patient's activity and it will help them to become more active and to live longer with good health. There are many studies regarding this activity recognition. One among them is Smartphone accelerometer data of elder people. Physical activity is very serious for aged people to reduce their risk of increasing comorbidities, and for extend their quality of life. An activity classification technique without any misclassification and high accuracy is the key for providing a proper response to elderly people. There are many classification algorithms, but most of the algorithms help younger adults only. As there are huge differences in features among the elder and younger adults, those algorithms developed for younger adults may not be accurate when applied to the elderly [1]. Therefore, there is a need for developing the best classification algorithm for elderly people.

For fall detection, many kinds of systems were designed and developed [2] [3]. The objective of the fall finding system is to differentiate fall events as abnormal events and the activities of daily life as normal events. The performance of these systems is analyzed using the following metrics Precision, Recall, F-measure and Accuracy. Fall detection systems are categorized into context-aware systems, devices for wearable, and mobile phone-based systems. In context-aware systems for identifying fall events sensors are positioned in the environment. The wide-angle cameras [4] [5], and depth cameras [6-19] are the most commonly used sensors. Other types of sensors such as microphones [10] and pressure sensors [11] are also used. Radars are one among them, which is used in context-aware systems. Here, sensing is done based on motions on the wavelet transform [12]. Wearable devices are miniature electronic sensor devices that are worn as clothes. An accelerometer is one of the wearable devices [13] [14] [15].

Generally, the fall-detection system mainly focuses on extracting the suitable features that can be able to detect fall events. In this paper Deep Learning model based on the LSTM is used to the overall spatiotemporal feature pattern for predicting the fall event. This approach focuses on using a less number of video frames to automatically learn the spatiotemporal features involved in the surveillance video frames and also to find the best classification algorithm to identify the fall events without any misclassification.

2. Related work

Now a day there is nothing without the Internet. The Internet provides an easy means for communicating with each other with very little human involvement. So there is a need to develop a better world for humans. Things around us are based on some awareness of the circumstance which can respond to some of the basic needs. Elegant houses are one of the major applications of the Internet of Things that are receiving special notice from young researchers [16]. Smart homes give a secure and sheltered environment for needy humans. It has the ability for tracking the activities of the residents without disturbing their daily life, and for tracking the behaviors of the residents and also to observe the health of the residents by using different sensors that are surrounded in the livelihood spaces of the people [17]. The information collected from these smart homes is deeply examined and studied, for extracting valuable information about the daily routines of the residents, and especially the activities of human beings.

Activity recognition [18] is a significant feature of the smart home. It consists of data categorization that is recorded by the differently incorporated environment and by wearable devices, which turns to proper and well-known movements. So there is a need to detect the abnormal activities which are gaining importance for dependent humans to ensure that behaviors are carried out any error [19]. It is used to ensure the safety of the well-being.

Anomaly detection of daily living (ADL) of a human being is to be carried out for detecting the nonconformities from the usual ADL model. Machine learning models complete the process [20]. Online telecare needs a system with extremely high accuracy, minimum computation instance and minimum involvement of the user. This is because of the increase and complexity in the data level [21]. The deep learning model presents a way for automatically extracting the functional and important spatial and sequential features of the unrefined data which doesn't require data classification as it leads to a complex and error-prone process. Therefore deep learning architecture is generalization in various contexts. LSTM is a dominant deep learning algorithm that is used for identifying sequences and anomalies in sequential data [22]. LSTM is used to take out the sequential features with long-time gaps. These possessions of LSTM gain more significance in smart houses and ADL for understanding the behaviors of humans, those changes over time.

The dataset described that real [23] or computer-generated [6] fall and non-fall events, which are composed and record using different devices to process and classify them. To classify the actions like falls or non-falls and various techniques like as decision trees [7], Support Vector Machine [4], Kalman Filtering [5], Thresh holding Techniques [6], Nearest-neighbor Rule [6], Mixture Gaussian Model [24], Rule-based methods, Multi-frame Gaussian Classifier [26], Gaussian Distribution of Clustered Knowledge [27], Bayesian Filtering [28], Hidden Markov Models [29], and Fuzzy Logic [30] are used.

The classification methodology selection depends on the features that are obtained for mapping input to output. Fall detection event systems are based on computer vision approaches. These features are simple because height and width ratio of the bounding box are surrounding the human [31], and it is very difficult as the distance points in a person point obscure to the floor [6]. Other important features are pattern extraction concern about the person curve [9], or orientation of human silhouette [32] at some point in a go down event. It also explained the way of predict, fall events, with the smallest amount of misclassification error.

3. Proposed Methodology

This detailed architecture of the fall event and the non-fall event classification system was explained in this section. Fig.1. shows the overall architecture of the proposed system. The proposed architecture consists of four phase's sensing, video abstraction, video analysis and communication. The sensing phase comprises of video capturing device details, and also the dataset description. The video abstraction phase includes sequence reduction. The video analysis phase comprises preprocessing, feature extraction and classification. Finally, in the communication phase, notification is sent through SMS, MMS and E-mail through the mobile phone.

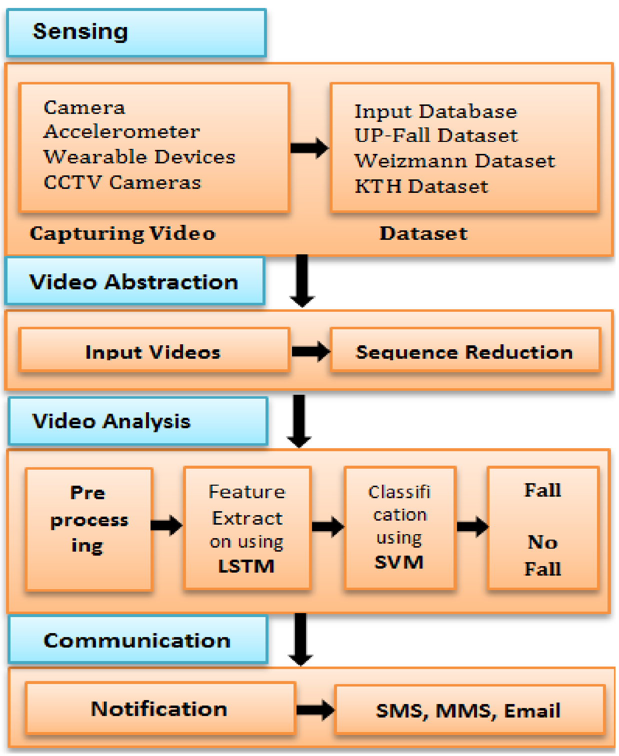

Fig. 1. Overall Architecture of Proposed System

3.1 Sensing

UR Fall detection dataset, Weizmann dataset and KTH dataset were used in this system to analyze the performance of the system. The dataset contains videos of human activities obtained from various devices of sensors, video devices and surveillance cameras.

3.1.1 UR Fall Dataset

The UR Fall discovery dataset is very vast. It comprises of 70 activities. Among these 30 fall events and 40 non-fall events. Humans (Subjects) perform daily activities of humans and fall events. Adults without any impairment are used for performing these activities. The dataset are publicly available in

http://fenix.univ.rzeszow.pl/~mkepski/ds/uf.html.

Fig. 2. Sample videos from UR Fall dataset

3.1.2 Weizmann Actions

The Weizmann dataset comprises ten action classes. The dataset includes different human actions. Fig. 3. shows the sample input from the Weizmann dataset.

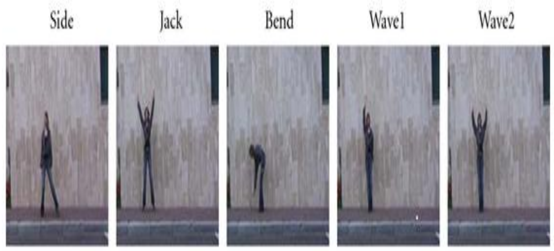

Fig. 3. Sample videos from Weizmann Dataset

3.1.3 KTH Actions

The KTH dataset comprise six human being action classes like walk, jog, run, boxing, waving, and clap. Twenty-five subjects performed each act many times. The sample input from the KTH dataset is illustrated in Fig. 4. There are totally 2391 video samples.

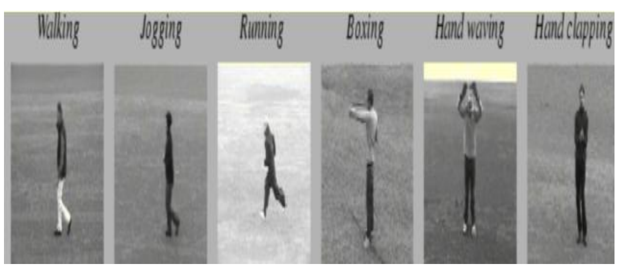

Fig. 4. Sample videos from KTH dataset

3.2 Video Abstraction

Video abstraction is the process of creating a short video summary from a long video. It is a sequence of stationary images or moving images. In terms of identifying the fall event, efficiently the video abstract is needed. Because within a specific time constraint it provides detailed content about the video sequence.Here is the goal of the system is to identify the fall events. So sequence reduction is done. Sequence reduction is done to get the abstract video which is sufficient for identifying the fall events.

3.2.1 Sequence Reduction

The input video sequences consist of videos of varying sizes. In general, the dataset contains very few sequences of longer length. This can be verified using a histogram of the sequence length. More padding can harmfully impact the accuracy of classifications. The longest sequences are removed to improve the accuracy.

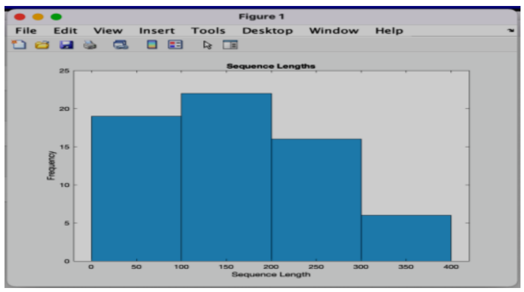

Fig. 5. Sample Output for Sequence Reduction

3.3 Video Analysis

Video analysis is the process of without human intervention to analyze the video and detect the spatial and temporal actions involved in the video frames. This video analysis stage includes preprocessing of frames, extraction of features and classification methods. In pre - screening stage center cropping is done, the feature extraction stage LSTM algorithm is used to extract temporal features and a multi-class SVM classifier for classifying the attributes.

3.3.1 Preprocessing

In this work, pre-trained deep network models are used for feature extraction. Hence the size of the input video sequence should match the previous one. This phase involves cropping the center portion of the previous network to remove the unwanted background. The size of the frame in each video sequence will be 224 x 224. The cropped video sequence is given for feature extraction.

3.3.2 Feature extraction using LSTM

Feature extraction plays an important phase in classification purposes. If the extracted features are sufficient and if the obtained features are useful then only the events can be classified correctly without any misclassification errors. Here the temporal features are extracted using LSTM.

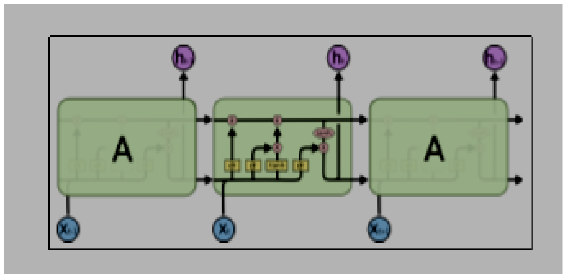

Fig. 6. Architecture of LSTM cell.

To break the difficulty of losing gradients, Long Short-Term Memory architecture was proposed. An LSTM cell is used for learning the long-term dependencies. LSTM comprises of hidden state and a cell state. They are maintained by structures called gates. The gate is the sigmoid function of the input. The architecture of an LSTM cell is shown in Fig. 6.

The input gate contains the new data that are stored in the cell state. This is represented in "Eq. (1) and (2)". The existing information in the cell state that is to be thrown away is decided by the forget gate. This is represented in "Eq. (3)". Finally, the output gate is used for filtering the output and decides the final cell output. This is represented in "Eq. (4) and (5)". o_t is the output-gate layer output, and h_t is the resultant unseen state for the given input.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (1)$$

$$c_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2)$$

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (3)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (5)$$

Now the frames are converted into feature vectors using LSTM to produce temporal features.

3.3.3 Support Vector Machine for Classification

SVM is one of the supervised learning approaches used for binary classification. The SVM method is used to find the best surface for making unification between optimistic and pessimistic training features that depends on training and testing set error reduction. The hyper-plane differentiates the vectorized data into two classes. The decision taken is dependent on this support-vector.

Let 'N' is the linearly separable training set with feature vector 'x' of 'd' dimension. For dual optimization,
Where, $\alpha \in R^N$ and $y \in \{1, -1\}$
Then the result of SVMs can be discussed as follows:

$$\vec{\alpha}^* = argmin\left\{-\sum_{i=1}^{n} \alpha_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \langle \vec{x_i}, \vec{x_j} \rangle\right\} \quad (6)$$

Where,

$$\sum_{i=1}^{n} \alpha_i y_i = 0; \qquad 0 \leq \alpha_i \leq C \quad (7)$$

The linear dataset is in SVM classification is separated with a single hyper-plane that is used for separating the two classes as a fall and the non-fall event with a given feature subset. The main problem in the SVM approach is to find the separate hyper-plane those classes are efficiently depends on the maximal margin. The SVM classifier algorithm takes the optimized features as input, classifies them, and applies the SVM with the RBF kernel to specify the hyper-plane. Then it checks the obtained accuracy and validity.

4. Experimental Results

The goal is to identify the fall eventand non-fall event. The performance of the proposed method is analyzed, and it is tested using UR Fall dataset, Weizmann dataset and the KTH dataset. Then to analyze the accuracy of the proposed work a comparison is made with existing techniques.

The deep features are extracted using the proposed LSTM approach based on an SVM classifier. Experiments are conducted in the standard database. The performance evaluation of the proposed machine learning-based SVM classifier is compared with the existing methods. To analyze the performance of the classifier, many performance evaluation attributes are available. Among them, Precision Rate, Recall Rate, F-Measure and Accuracy are used as performance metrics for this work.

4.1 Evaluate the Precision Rate

Precision is the fraction of retrieved fall and non-fall events that are relevant. Precision is calculated utilizing “Eq. (8)”.

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

Where TP = True Positive (the fall event that is identified as fall event)
FP = False Positive (the non-fall events that are identified as fall events)

4.2 Recall Rate

The recall is the fraction of relevant fall and non-fall instances that are retrieved. The recall is calculated utilizing “Eq. (9)”

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

Where TP = True Positive (the fall event that is identified as fall event)
FN = False Negative (fall events that are identified as the non-fall events)

4.3 F-Measure

It uses both precision and recall and it is called the harmonic mean of precision and recall. The balanced F-score is calculated utilizing “Eq. (10)”

$$F_m = 2\,\frac{Precision * Recall}{(Precision + Recall)} \quad (10)$$

4.4 Accuracy

The Accuracy measure is the accurate prediction of the fall and the non-fall events. The Accuracy is calculated utilizing “Eq. (11)”

$$Accuracy = \frac{TP+TN}{TP+Tn+FN+FP} \quad (11)$$

Performance Metrics	Weizmann Dataset	KTH Dataset	UR Fall Dataset
No. of Actions	99	100	80
Precision	94.3	96.5	100
Recall	93.5	95.4	100
Accuracy	95.7	96.3	100
F-measure	94.8	96.9	100

Table 1. Performance Analysis of Proposed Work

The Performance of the proposed LSTM feature extraction method with an SVM classifier for identifying the fall and Non fall events are analyzed by three different datasets. It is observed from Table 1 that the proposed work recognizes events with higher accuracy.

Actions		Non Fall Events	Fall Events
Non Fall Events	Walk	100	0
	Run	100	0
	Jump	100	0
	Pjump	100	0
	Jack	95	5
	Bend	92	8
	Wave1	94	6
	Wave2	94	6
	Skip	96	4
	Side	95	5

Table 2. Results on Weizmann dataset

The above Table 2. shows that which of the classes are confused with the fall events collected from the UR Fall dataset. As observed from Table 2 the proposed method suffers some failures in differentiating Bend, Jack, Wave1, Wave2, Skip and Side with the fall event. Actions other than these provide 100% accuracy. The overall accuracy of the Weizmann dataset is 95%.

Actions		Non Fall Events	Fall Events
Non Fall Events	Walk	100	0
	Run	100	0
	Jog	91	9
	Hwave	96	4
	Hclap	90	10
	Box	98	2

Table 3: Results on KTH dataset

As observed from the above Table 3, for the KTH dataset out of 6 actions only 4 actions, namely Hwave, Hclap, jog and box are confusing with fall events.The KTH dataset provides an overall accuracy of 96 percent.

Actions	Non Fall Events	Fall Events
Fall-01-cam0	0	100
Fall-02-cam0	0	100
Fall-03-cam0	0	100

Fall-04-cam0	0	100
Fall-05-cam0	0	100
Adl-01-cam0	100	0
Adl-02-cam0	100	0
Adl-03-cam0	100	0
Adl-04-cam0	100	0
Adl-05-cam0	100	0

Table 4: Result on UR Fall dataset

During testing on the UR Fall dataset there is no misclassification error between fall and non-fall events. In the UR Fall dataset, all events are correctly classified. The UR Fall dataset provides an overall accuracy of 100%.

5. Conclusion

In worldwide, the number of elderly people living alone increases day by day. In this work, a fall detection system for controlled environments based on the user needs is proposed. LSTM based feature extraction is done to extract temporal features. The proposed framework shows better execution when recognizing falls, featuring among them the model LSTM. The proposed framework has a few disservices, for example, being inclined to vulnerability on encompassing temperature and the presence of articles in the space of coverage. In the end, the proposed framework shows that a profound learning-based LSTM based component extraction algorithm is reasonable for perceiving fall-occasion. It is seen that the profound learning-based example portrayals help to build the exactness for fall-detection systems when contrasted with customary example portrayal strategies. The worldly elements were automatically educated and extricated utilizing LSTM neural organization. The classification is done using an SVM classifier. Then finally a notification is sent as an alarm through SMS, MMS via E-mail. The experimental results reveal that the presented approach improved the overall accuracy of the system. The UR-Fall detection dataset provides 100% accuracy, the Weizmann dataset provides 95% accuracy and the KTH dataset provides 96% accuracy. This system can also be employed under real-life conditions.

References

[1] Musci M, De Martini D, Blago N, Facchinetti T, Piastra M,'Online fall detection using recurrent neural networks', 2018.

[2] Igual, R., Medrano, C., and Plaza, I. 'Challenges, issues and trends in fall detection systems', Biomedical engineering, online 12, 1 2013.

[3] Mubashir, M., Shao, L., and Seed, L, 'A survey on fall detection: Principles and approaches'. Neurocomputing, 2013, pp:144–152.

[4] Bosch-Jorge, M., Sanchez-Salmer, A.-J., Valera, ´A., and Ricolfe-Viala, C,'Fall detection based on the gravity vector using a wide angle camera', Expert Systems with Applications 41, 17 (2014),pp: 7980–7986.

[5] Rezaee, K., Haddadnia, J., and Delbari, A, 'Modeling abnormal walking of the elderly to predict risk of the falls using kalman filter and motion estimation approach', Computers & Electrical Engineering, 2015,pp: 471–486.

[6] Kwolek, B., and Kepski, M,'Improving fall detection by the use of depth sensor and accelerometer',Neurocomputing,vol.168, 2015,pp: 637–645.

[7] Choi, Y., Ralhan, A., and Ko, S, 'A study on machine learning algorithms for fall detection and movement classification','Information Science and Applications (ICISA), 2011, IEEE, pp.1–8.

[8] Akagunduz, E., Aslan, M., Sengur, A., Wang, H., and Ince, M, 'Silhouette orientation volumes for efficient fall detection in depth videos', IEEE journal of biomedical and health informatics, 2016.

[9] Ma, X., Wang, H., Xue, B., Zhou, M., Ji, B., and Li, Y,'Depth-based human fall detection via shape features and improved extreme learning machine, IEEE journal of biomedical and health informatics 18, 6 (2014), 1915–1922.

[10] Li, Y., Ho, K., and Popescu, M, 'A microphone array system for automatic fall detection', IEEE Transactions on Biomedical Engineering ,2012, pp:1291–1301.

[11] Tzeng, H.-W., Chen, M.-Y., and Chen, J.-Y, 'Design of fall detection system with floor pressure and infrared image', International Conference on System Science and Engineering ,2010, IEEE, pp. 131–135.

[12] Su, B. Y., Ho, K., Rantz, M. J., and Skubic, M, 'Doppler radar fall activity detection using the wavelet transform', IEEE Transactions on Biomedical Engineering, 2015, pp:865–875.

[13] Tong, L., Song, Q., Ge, Y., and Liu, M., 'Hmm-based human fall detection and prediction method using tri-axial accelerometer',. IEEE Sensors Journal, 2013, pp: 1849–1856.

[14] Choi, Y., Ralhan, A., and Ko, S,'A study on machine learning algorithms for fall detection and movement classification', Information Science and Applications (ICISA), 2011, pp. 1–8.

[15] Li, H., and Yang, Y.-L, 'Research of elderly fall detection based on dynamic time warping algorithm', Control Conference (CCC), 2016, pp. 5190–5194.

[16] Nef T, Urwyler P, Büchler M, Tarnanas I, Stucki R, Cazzoli D, Müri R, Mosimann U, 'Evaluation of Three State-of-the-Art Classifiers for Recognition of Activities of Daily Living from Smart Home Ambient Data', Sensors.

2015; 15(5):11725-11740. https://doi.org/10.3390/s150511725

[17] Ransing RS, Rajput M, ‘Smart home for elderly care, based on Wireless Sensor Network.’, International Conference on Nascent Technologies in the Engineering Field (ICNTE) 2015 Jan 9 (pp. 1-5).

[18] Chun Zhu & Weihua Sheng 2011, ‘Motion-and location-based online human daily activity recognition’, Pervasive and Mobile Computing, Vol.7, No.2, pp.256–269.

[19] Palaniappan, A., Bhargavi, R., Vaidehi, V, ‘Abnormal human activity recognition using SVM based approach’,International Conference on Recent Trends Information Technology, ICRTIT 2012, pp. 97–102.

[20] Bouchachia, A, ‘Activity recognition and abnormal behaviour detection with recurrent neural networks’. Proc. Comput. Sci. 110, 86–93 (2017).

[21] Zerkouk M, Chikhaoui B,‘Spatio-temporal abnormal behavior prediction in elderly persons using deep learning models’, Sensors, 2020.

[22] Zerkouk M, Chikhaoui B,‘Long short term memory based model for abnormal behavior prediction in elderly persons’, International Conference on Smart Homes and Health Telematics,2019 ,pp. 36-45, Springer.

[23] Stone, E. E., and Skubic, M,‘Fall detection in homes of older adults using the Microsoft kinect’, IEEE journal of biomedical and health informatics,vol.19,No.1, 2015, pp: 290–301.

[24] Rougier, C., Meunier, J., St-Arnaud, A., and Rousseau, J. Robust,‘video surveillance for fall detection based on human shape deformation’, IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no.5 ,2011,pp: 611–622.

[25] Vishwakarma, V., Mandal, C., and Sural, S,‘Automatic detection of human fall in video’, International conference on pattern recognition and machine intelligence, 2007, Springer, pp. 616–623.

[26] Hazelhoff, L., Han, J., ‘Video-based fall detection in the homeusing principal component analysis’, International Conference on Advanced Concepts for Intelligent Vision Systems 2008, Springer, pp:298–309.

[27] Yuwono, M., Moulton, B. D., Su, S. W., Celler, B. G., and Nguyen, H. T, ‘Unsupervised machine-learning method for improving the performance of ambulatory fall-detection systems’, Biomedical engineering online, vol.11,no. 1,2012.

[28] Rimminen, H., Lindstr¨om, J., Linnavuo, M., and Sepponen, R, ‘Detection of falls among the elderly by a floor sensor using the electric near field’, IEEE transactions on information technology in biomedicine, a publication of

the IEEE Engineering in Medicine and Biology Society, vol.14,no.6 , 2010, pp:1475–1476.
[29] Cucchiara, R., Prati, A., and Vezzani, R, 'A multi-camera vision system for fall detection and alarm generation', Expert Systems, vol.24, no.5 ,2007,pp: 334–345.
[30] Planinc, R., and Kampel, M,'Robust fall detection by combining 3d data and fuzzy logic' In Asian Conference on Computer Vision ,2012, Springer, pp. 121–132.
[31] Rougier, C., Meunier, J., St-Arnaud, A., and Rousseau, J,'Robust video surveillance for fall detection based on human shape deformation', IEEE Transactions on Circuits and Systems for Video Technology, vol. 21,no. 5 ,2011, pp: 611–622.
[32] Akagunduz, E., Aslan, M., Sengur, A., Wang, H., and Ince, M, 'Silhouette orientation volumes for efficient fall detection in depth videos', IEEE journal of biomedical and healthinformatics, 2016.

Authors Profile

S.Vijayakumar is a Ph.D Research Scholar in Rani Anna Government Arts College for Women, Manonmaniam Sundaranar University,Tirunelveli. He has more than 10 years of Teaching Experience in Computer Science. His area of interest is IoT, Deep Learning and Video image processing in health care system.

Dr.R.Shenbagavalli. She is currently working as an Associate Professor, Department of Computer Science, Rani Anna Government College for Women, Tirunelveli. Her current research interests include Feature Extraction of soil images for retrieval based on Statistics, Satellite image edge detection using Fuzzy Logic and its applications video image processing and health care sectors.