Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 3, July 2021: 5237-5248

Research Article

Domestic - Monitoring System For TheElderly Using Deep Learning

G. Deepa[1], V. Suresh Kumar[2], Hari Baskar C[3], Hari Prakathish T[4], Deepak Kumar J[5]

Abstract

Taking care of elderly people is a significant duty. Elderly people are prone to fall-related fatalities including deaths. Fall affects them physically as well as mentally. Even though old adults are monitored manually there are cases when the caretaker is away for a moment and the old adults are at high risk of indoor accidents one such is falling. The time gap between the incident of fall and proper medical treatment is crucial in enhancing the chances of recovery. Thusa digitally assisted approach is required when manual monitoring is unavailable and to ease the same. In this project, deep learning-based digital monitoring is proposed for elderly people in indoor environments to detect falls. The proposed neural network architecture is computationally light, owing to lesser trainable parameters and a less complex pre-processing pipeline involving an optical-flow algorithm. Optical flow has the dual advantage of ensuring the privacy of the individualand requirement of inexpensive hardware which is a common RGBcamera.

Keywords— *Optical-flow, CNN, Transfer Learning*

[1]Assistant Professor (Sr.G) Department of Electronics and Communication Engineering Kongu Engineering College, Perundurai, Tamilnadu, India.
[2]Assistant Professor (Sr.G) Department of Electrical and Electronics Engineering K.Ramakrishnan college of Technology, Samayapuram, Trichy, Tamilnadu, India.
[3, 4, 5]Bachelor of Engineering Department of Electronics and Communication Engineering Kongu Engineering College, Perundurai, Tamilnadu, India.
Email id: g.deepa@kongu.ac.in

I. INTRODUCTION

Falls are more likely in case of physical weakness. The elderly suffer falls detrimental to their health due to age-related weakness. Falls are among the prime reasons for mortality in elderly people as briefed by Ambrose et al. [1]. This fact is justified with high frequency of falls: 33.33% of adults fall at least once in a year over the age of 65. The impact is a solicitude for human health and welfare systems as old adults require attention on par with infants. Falls lead to both physical and mental illness. The medical expenditures as a consequence is humongous[2]. An expenditure of US$55 billion by the end of 2020.

The most common strategy for fall detection is gathering data from necessary sensors and computing with the data to interpret result[3]. Advances have led to smart assisting environments for old adults, previously limited to home settings [4].

This work is a vision relevant approach for fall detection. The common vision-based hardware component isthe camera. Its data is very much valuable provided we have the necessary algorithm to interpret the data into classifying target of interest classes. Moreover, cameras are used for surveillance in public places as well as elder-care homes andother indoor environments. This helps to deploy reliable vision-based fall detection methods in healthcare and accident detection strategies

The impact of deep learning has improved accuracy and reliability in object localization, and segmentation [5]. This proposed CNN detects falls from an optical-flow image stack. In the beginning step, the proposed model is trainedon the Imagenet dataset [6] to enhance chances of detecting relevant boundaries and features. In the next step, as the approach [7], the CNN is trained on the optical flow images generated from individual samples of the UCF101 action dataset [8]. This is to generalize the CNN to differentiate various everyday human actions. The third step, transfer learning is employed, trained on Fall Datasets.

The main contributions of the work carried out are (i)Transfer learning for fall detection from the action recognition domain . This is to leverage the limited fall samples in datasets available. (ii) Optical-flow data as input to neural network to suppress background noise. Optical- flow only contains data of relative motion between any two successive video frames suppressing the subject's physical features such as colour, brightness, and contrast. Hence, the privacy of the intended individual is safeguarded.

II. LITERATURE REVIEW

The writing on fall identification is split between vision-based and sensor-based methodologies. The sensors put together recognition has ordinarily transferred with reference to the use of accelerometers, which give appropriate speed increase estimates like vertical speed increase. On account of falls, these actions are totally different contrasted with day by day exercises or jumbling occasions, (for example, twisting around or crouching), permitting us to acknowledge betwixt them. The knowledge of a 3-hub accelerometer (speed increase esteems in x-, y-, and z-pivot) was put forward by Vallejo et al. [9] and Sengtoet.al [10] taking care of a Multilayer Perceptron (MLP). An Inertial Measurement Unit (IMU) joined with the profundity maps got from a Kinect camera was applied by Kwolek and Kepski [11]. By taking care of it the knowledge from the IMU and therefore the Kinect, they additionally utilized a Support Vector Machine (SVM) classifier. Approaches just like the latter and [12] consolidated sensors with vision methods. Nonetheless, they utilized vision relevant arrangements just to work out the forecast of the sensorrelevant methodology.

The simply vision-put together methodologies center with reference to the sides of recordings to differentiate falls.Through PC vision procedures, significant highlights like outlines or jumping boxes are separated from the casings to figure with recognition. a couple of arrangements utilizethose highlights as contribution for a classifier (e.g., GMM, SVM, and MLP) to naturally distinguish fall event. Positioning frameworks is additionally much reached out; Lee et.al[17] deployed following methods during a nearby climate to acknowledge falls. They proposed utilizing associated parts marking to process the outline of private and removing highlights just like the spatial direction of the focus of the outline or its mathematical direction. Consolidating this data they will distinguish positions and furthermore fall. Rougier et al. [18] proposed utilizing outlines also, which may be a typical procedure within the writing. Applying a coordinating with framework along the video to follow the deformity of the outline, they brokedown the state of the body lastly acquired an outcome with aGMM. The individual’s head was followed by Mubashir et al. [3] to enhance their base outcomes utilizing a multi-outline Gaussian classifier, which was taken care of with the bearing of the chief segment and therefore the change proportion of the outline. Another regular procedure comprises of processing the bouncing boxes of the things to make a decision whether or not they contain a private and after-ward identify the autumn through highlights extricated from it (see, as an example, [19, 20]). Following a comparative methodology, to work the attitude proportion, level and vertical inclinations of a piece of writing, and fall point Vishwakarma et.al[21] has worked with bounding boxes. GMM classifier as the end of pipeline. Numerous arrangements depend upon regulated realizing, that is, removing tons of highlights from crude pictures and utilizinga classifier to require during a choice from named information. This is often things, for instance, I. Charfi et al.

[16] removed 14 highlights (change in first and second subordinates, the Fourier change, and therefore the Wavelet change) SVM to the last stage. Zerrouki et al. [22]

figured inhabitance regions round the body’s gravity local area, separated their points, and took care of them into different classifiers, being the SVM the one which got the simplest outcomes. In 2017, to exhibit the various body presents [13] same maker expanded his past work by adding Curvelet coefficients as additional highlights and applying a Hidden Markov Model (HMM). Harrou et al.[23] utilized a less ceaseless methodology and also applied Multivariate Exponentially Weighted Moving Average (MEWMA) graphs. In any case, they couldn’t recognize falls and puzzling occasions, which may be a significant issue that's considered within the proposed arrangement. Truth be told, not having the choice to separate between such circumstances delivers a unprecedented number of bogus alerts.

Vision-based fall recognition frameworks are 3D spatial constructions. Aloof frameworks or dynamic profundity cameras from basis of this system. Kinect camera is employed to assemble a 3D outline to then investigate the quantity dispersion along vertical pivot [24] then extricate 3D highlights and afterward applying a worldwide positioning framework to differentiate the falls, a similar work was proposed by Gasparrini et al. [25]. Kinect programming locates locomotory joints of body, which were utilized by R.Planinc et al. [26]. G.Diraco et al. [27] utilized profundity guides to register 3D highlights. Another straightforward on the other hand fascinating methodology was given by Mastorakis.et.al [28], they applied 3D bouncing boxes. All other previously mentioned strategies exploited the 3D data given by their camera frameworks. The disadvantages of such methodologies are identified with framework organization: they have either numerous synchronized cameras zeroed in on similar region or dynamic profundity cameras, limited frame of coverage a restricted profundity. This way is consistent with the attitude of framework sending, 2D inactive frameworks are generally a superior choice, given their reasonable cost. It’s likewise imperative to feature that cameras are now introduced in numerous public spots, like air terminals, shops, and older consideration habitats. These reasons make 2-Dimensionaluninvolved camera-based fall detection.

These days, utilization of profound neural organizations is filling in numerous difficult areas, together with fall recognition using data from regular cameras. A PCAnet[28] based method isproposed in [14] removing highlights from shading pictures and after-ward tried an SVM to identify falls. This methodology is like [32] yet rather than a PCAnet, they utilize a changed VGG16 design [30] that permits us to handle different casingsto consider movement. Another examination work [15], consolidated HOG (Histograms of Oriented Gradients), LBP (Local Binary Pattern), and Caffe [31] neural organization synthesized highlights to perceive an outline. An SVM classifier finally.

III. RELATED WORKS

Adrián .et.al**[32]** propose a VGG16**[30]** based architecture as a neural network to be trained. The training phase is similar to the proposed methodology. The difference lies in how the optical-flow stacks are constructed. The parameter '*L'* in their work accounts for the number of optical-flow stacks precisely '*2 X L'*. The optical flow stack, in this case, is a *four*-dimensional stack where a particular optical-flow image in i^{th} position of one row is repeated in the $(i+1)^{th}$ position of the consecutive row, forming a series of images depicting motion across each frame of the original video clip. However, this work is bulkier in terms of trainable weights (138 million) and a delayed pre-processing pipeline due to bulkier 4- dimensional input (224 **X** 224 **X** $2*L$ **X** $(N - L+1)$), where N is the number of frames captured from a video clip.

Gupta R .et.al**[33]** developed a modified Inflated 3D (I3D) architecture. The prepared dataset consisted of compressed spatio-temporal data result of compressive sensing framework, synthesized from human action video- clips. The video compression framework is said to address privacy issues in the case of RGB images from a regular camera. Ten no-fall action classes were selected from Kinetics-400 after compressive sensing was used to train the3D ConvNet. This methodology takes 3 channel inputs (RGB), which is not an essential parameter to the solution. According to K. de Miguel et. al**[34]**, their algorithm aims tolocalize subjects in fall states. The algorithm initially determines the subject's current state from the scene. This data acquisition is a step-wise process: The first step subtracts the subject from the background, progressively learning the altering environment of the subject, and detecting uninteresting objects, localizing the subject through the scene, and scanning for subjects occluded by objects. Second step Kalman filter to reduce noisy data and accumulate the periodic changes common to different human actions. Finally, the KNN algorithm **[35]** processes the obtained data classifying the current state of the subject. Their work also includes occlusion detection but fails in case of varying light conditions for which they insist on using optical flow in the future.

A highly quantitative method is proposed by X. Kong **[36]** is to get binary image and its outline by a depth camera followed by canny filter respectively. The outline image is a contour of white pixels, for which the tangent vector angleis calculated and divided into 15° groups. The fall is classified when most of the calculated tangent angles are less than 45°. This requires a depth camera which is expensive and fails in cases where only the initiation to fall occurs and the fall never occurred in the near future.

Lezzar. et.al**[37]** use object recognition as the initialstage to pick out the person in a video frame and then compute for fall detection. The first step is extracting a frame from a video stream. Person detection by the YOLOV3 algorithm in the extracted frame features extraction from a bounding box which is the result of a YOLO detector. YOLO is a time consuming object detection and localization algorithm to train and bulkier when it comes to detecting multiple bounding boxes. This approach is misleading when the person is simply laying on the ground and not the result of a fall. Thus, the training methodology proposed by Adrián.et.al[32] can be modified to be trained with only a lesser dimensional input stack. The architecture in Gupta R.et.al**[33]** has inception modules for fall detection that reduce the number of trainable parameters,thus a backbone InceptionV3 architecture would result in a compact model without compromising accuracy.

IV. PROPOSED APPROACH

A. DATASET

The datasets used belong to correspond to the categories general human action detection and fall detection. UFC101 dataset **[8]** was used for training the proposed model initially. Then transfer learning was employed by reusing the weights to train on fall detection datasets URFD **[41]** and FDD **[32]**.

B. PRE-PROCESSING

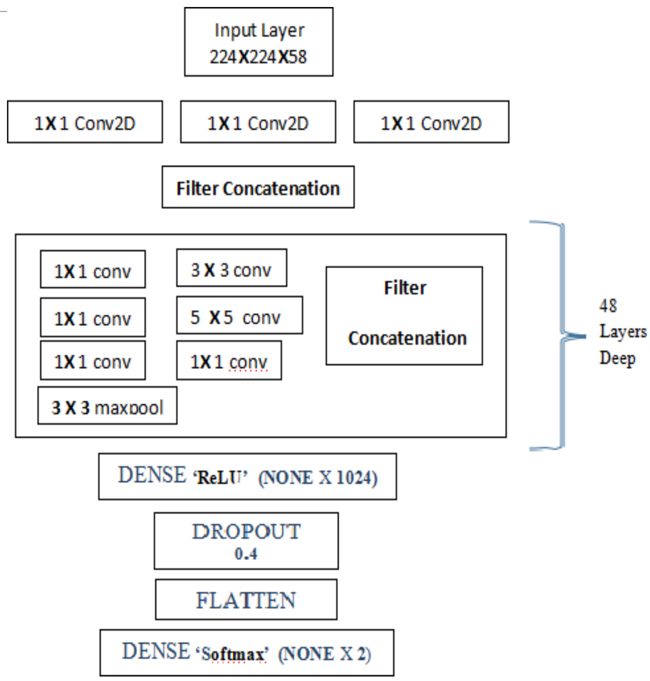
The video clips in the dataset were individually reduced to short clips of 1.2-second temporal duration. At frame rate of 25 frames per second, 1.2 seconds correspond to 30 frames as shown in ***equation 1*** of the 3 channel image data stack. This temporal duration is based on the fact that fall incidents typically last 1.2 - 1.3 seconds.

$$F = F_{rate} \times t_{fall} \quad \rightarrow (1)$$

$$N_{stack} = (F - 1) \times 2 \quad \rightarrow (2)$$

Optical flow displacement images are generated with the TVL-1 algorithm **[40]**. It takes the input of two consecutive frames and computes optical flow between them. The result is two arrays consisting of displacement of pixels along the horizontal and vertical dimensions. Hence 29 pairs of optical flow images are generated which comprise a single stack for training shown by ***equation 2***. The dimension of an individual stack is (224 X 224 X 58). A sample output of the greyscale image stack produced by the TVL-1 algorithm is shown.

Figure: 1 Optical-flow images generated from theresult of TVL-1 dense-flow algorithm



C. TRAINING PHASE

Training consists of two phases.

i. Training with optical flow stacks of action detection dataset (UFC101) which has a million samples belonging to 101 different action categories.

ii. Transfer learning and training with optical flow stacks of action detection datasets. The benchmark URFDand FDD datasets were used.

D. PROPOSED ARCHITECTURE

The proposed architecture is a backbone of Inception V3 **[39]** architecture. Inception V3 is immune to vanishing gradient problems and it has far lesser trainable parameters (24 million) as opposed to similar implementation with VGG - 16 **[32]** architecture (138 million trainable parameters).

Input is the optical-flow stack which represents a turret of frames depicting actual motion. The stack consists of all grey-scale images only depicting displacement. The inception architecture strictly requires three-channel input, thus there is a need for channel reduction without loss of data. Two-dimensional convolution layers serve this processreducing the channel size to fit the Inception V3 input layer.

The InceptionV3 architecture is initially loaded with the 'Imagenet' dataset. This is done to better identify boundary features in the images. This worked arguably well,Wang et al. **[7]** prove in his work that all-inclusive appearance features obtained on training in 'Imagenet' **[6]** initialize the network weights to acquire optical-flowfeatures.

The output from the InceptionV3 is connected to a Dense layer of dim (**None, 1024**) and activation = '**ReLu**', Dropout layer of dropout fraction = **0.4**, Flatten layer, lastly Dense layer of dim (**None, 2**) and activation = '**Softmax'.**

Figure 2: Architecture of proposed InceptionV3backbone model

V. RESULTS AND DISCUSSION

A. EVALUATION METHODOLOGY

The proposed work is a supervised learning, binary classification model. The input to the classifier is specific set of optical-flow frames composed from a video-clip, which the classifier has to classify as fall or non-fall. Common evaluation metrics to evaluate this type of classifier include: Sensitivity (probability of true positives), and specificity (probability of true negatives). Sensitivity in this context is the accuracy of predicting falls, specificity is the accuracy of "no falls". The evaluation metrics are defined as:

Sensitivity: → **(4)** $TP \div (TP + FP)$

Specificity: → **(5)** $TN \div (TN + FN)$

Accuracy: → **(6)** $(TP + TN) \div (TP + TN + FP + FN)$

TP - True Positives TN - True NegativesFP - False Positives FN - False Negatives

B. ACHIEVED EVALUATION PARAMETERS:

The VGG-16 architecture proposed by Adrián Núñez-Marcos et.al **[32]** shows remarkable values of sensitivity, specificity, and accuracy are shown in table 1.

Table: 1 Evaluation parameters of replicated VGG16model Evaluated on URFD dataset.

Dataset	Sensitivity %	Specificity %	Accuracy %
URFD	95.45%	97.53%	95.33%

The training and validation set accuracy versus the number of epochs for the proposed model trained with the benchmark URFD dataset shows the plummeting accuracy over the number of epochs is seen for the range 0 to 30 epochs after which accuracy values oscillate between within

0.9 and 0.96. The difference between training and validationaccuracy is less than 1% evident from the graph in thefigure: 2, thus the model is optimal.

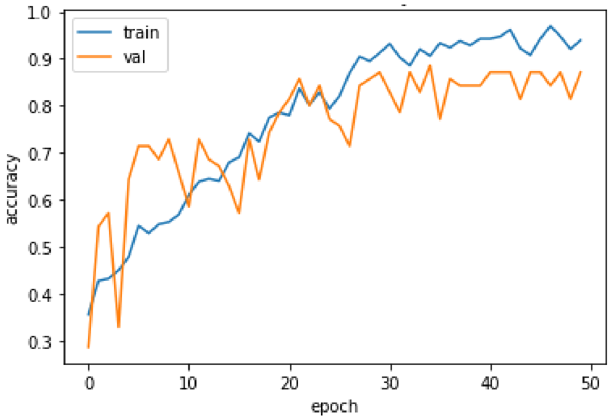


Figure: 3Training and validation accuracy versus epochsfor a proposed model trained on URFD dataset

The training set accuracy value for benchmark URFD dataset for the proposed model and the sensitivity, specificity values which are more than that obtained for the above-mentioned VGG16 model.

Table: 2 Evaluation parameters of replicated VGG16model Evaluated on FDD dataset.

Dataset	Sensitivity %	Specificity %	Accuracy %
URFD	95.45%	97.53%	96.03%

Figure 4 depicts the variation of training and validation set accuracy with the number of epochs is plotted for benchmark FDD dataset trained for 100 epochs after which accuracy did not improve. The result is obtained for the proposed model.

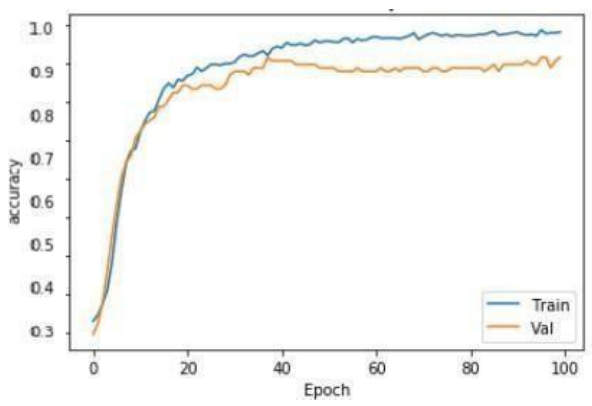


Figure: 4 Training and validation accuracy versusepochs for a proposed model trained on FDD dataset

The sensitivity and specificity values for the proposed model trained and validated on the FDD dataset depict the accuracy of true positives and percentage of true negatives respectively, evident from table:5.2c. The overall training accuracy is high compared to the replicated VGG16 model.

Table: 3 Evaluation parameters of proposed modelEvaluated on FDD dataset.

Dataset	Sensitivity %	Specificity %	Accuracy %
URFD	95.45%	97.53%	95.33%

The overall performance of the proposed model is compared with that of the well-performing models. The proposed model surpasses previous models in actual accuracy and latency observed in table 4.

Table: 4 Comparisons of various standard models withour proposed model.

MODEL	Number of trainable parameters	Latency	Accuracy in URFD dataset
VGG16	138 million	27	95.33%
Proposed backbone InceptionV3	26.2 million	20	96.03%
I3D with inception modules	<1 million	22	94.67%

VI. CONCLUSION AND FUTURE SCOPE

Modeling sequence data processing into a discrete data problem using a convolutional neural network is necessary, reducing the number of trainable parameters. The result is lesser latency for the model to predict and the main objective of reducing the time between an accidentand its detection. This is advantageous in timely detection and alerting fall accidents, enhancing the chances of recovery of elderly people. The proposed neural network model works with the output from a regular camera module, thus reducing cost by not requiring expensive motion sensing devices for input.

In the future micro-controllers dedicated to computer vision will be used to further analyze the model based on the optimum point of positions. These controllers are inexpensive and allow for independent decentralized data processing. This will increase the area of monitoring and compensate for any missed detections accounted for by the monitoring node. Further other video compression standards that perform well in varying lighting conditions will be experimented with to reduce latency.

REFERENCES

1. A. F. Ambrose, G. Paul, and J. M. Hausdorff (2013), "Risk factors for falls among older adults: A review of the literature", Maturitas, vol. 75, no. 1, pp. 51–61.
2. J. C. Davis, M. C. Robertson, M. C. Ashe, T. Liu-Ambrose, K. M. Khan, and C. A. Marra (2010), "International comparison of cost of falls in older adults living in the community: A systematic review", Osteoporosis International, vol. 21, no. 8, pp. 1295–1306.
3. M. Mubashir, L. Shao, and L. Seed (2013), "Asurvey onfall detection: principles and approaches", Neurocomputing, vol. 100, pp. 144– 152.
4. L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu (2012), "Sensor-based activity recognition", IEEE Transactions on Systems, Man, and Cybernetics, Part C:Applications and Reviews, vol. 42, no. 6, pp. 790–808.
5. Y. LeCun, Y. Bengio, and G. Hinton (2015), "Deep learning", Nature, vol. 521, no. 7553, pp. 436–444.
6. J. Deng, W. Dong, and R. Socher (2009), "ImageNet: a large-scale hierarchical image database", in Proceedingsof the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 248–255.
7. L. Wang and et al (2015), "Towards good practices for very deep two-stream ConvNets".
8. S. Khurram, A. R. Zamir, and M. Shah, Amir Roshan Zamir, and Mubarak Shah (2012), UCF101: A dataset of 101 human actions classes from videos in the wild.
9. M. Vallejo, C. V. Isaza, and J. D. Lopez (2013), "Artificial Neural Networks as an alternative to traditional fall detection methods", in Proceedings ofthe 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,EMBC 2013, pp. 1648–1651.
10. A. Sengto and T. Leauhatong (2012), "Human falling detection algorithm using backpropagation neural network", in Proceedings of the 5th 2012 Biomedical Engineering International Conference, BMEiCON 2012.
11. B. Kwolek and M. Kepski (2014), "Human fall detection on embedded platform using depth maps and wireless accelerometer", Computer Methods and Programs in Biomedicine, vol. 117, no. 3, pp. 489–501.
12. F. Harrou, N. Zerrouki, Y. Sun, and A. Houacine (2016), "Statistical control chart and neural network classification for improving human fall detection," in Proceedings of the 8th International Conference on Modelling, Identification and Control, ICMIC 2016, pp. 1060–1064.
13. N. Zerrouki and A. Houacine (2017), "Combined curvelets and hidden Markov models for human fall detection", Multimedia Tools and Applications, pp. 1– 20.
14. S. Wang, L. Chen, Z. Zhou, X. Sun, and J. Dong (2015), "Human fall detection in surveillance video based on PCANet", Multimedia Tools and Applications,vol. 75, no. 19, pp. 11603–11613.
15. K. Wang, G. Cao, D. Meng (2016), W. Chen, and W. Cao, "Automatic fall detection of human in video using combination of features", in Proceedings of the 2016 IEEE

International Conference on Bioinformatics and Biomedicine, BIBM 2016, pp. 1228–1233.

16. I. Charfi, J. Miteran, J. Dubois, M. Atri, and R. Tourki (2012), “Definition and performance evaluation of a robust SVM based fall detection solution”, in Proceedings of the 8th International Conference on Signal Image Technology and Internet Based Systems, SITIS 2012, pp. 218–224.
17. T. Lee and A. Mihailidis (2005), “An intelligent emergency response system: Preliminary development and testing of automated fall detection”, Journal of Telemedicine and Telecare, vol. 11, no. 4, pp. 194–198.
18. C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau (2011), “Robust video surveillance for fall detection based on human shape deformation”, IEEETransactions on Circuits and Systems for Video Technology, vol. 21, no. 5, pp. 611–622.
19. S.-G. Miaou, P.-H. Sung, and C.-Y. Huang (2006), “A customized human fall detection system using omni- camera images and personal information”, in proceedings of the 1st Transdisciplinary Conference on Distributed Diagnosis and Home Healthcare, D2H2 2006, pp. 39–42.
20. C.-L. Liu, C.-H. Lee and P.-M. Lin (2010), “A fall detection system using k-nearest neighbor classifier”, Expert Systems with Applications, vol. 37, no. 10, pp. 7174–7181.
21. V. Vishwakarma, C. Mandal, and S. Sural (2007), in Pattern Recognition and Machine Intelligence, Automatic detection of human fall in video, Ed., pp. 616–623.
22. N. Zerrouki, F. Harrou, A. Houacine, and Y. Sun (2016), “Fall detection using supervised machine learning algorithms: A comparative study”, in Proceedings of the 8th International Conference on Modelling, Identification and Control, ICMIC 2016, pp. 665–670.
23. F. Harrou, N. Zerrouki, Y. Sun, and A. Houacine (2016), “A simple strategy for fall events detection”, in Proceedings of the 14th IEEE International Conference on Industrial Informatics, INDIN 2016, pp. 332–336.

24. E. Auvinet, F. Multon, A. Saint-Arnaud, J. Rousseau, and J. Meunier (2011), “Fall detection with multiple cameras: An occlusion-resistant method based on 3-D silhouette vertical distribution”, IEEE Transactions on Information Technology in Biomedicine, vol. 15, no. 2, pp. 290–300.
25. S. Gasparrini, E. Cippitelli, S. Spinsante, and E. Gambi (2014), “A depth-based fall detection system using a Kinect® sensor”, Sensors, vol. 14, no. 2, pp. 2756–2775.
26. R. Planinc and M. Kampel (2013), “Introducing the use of depth data for fall detection”, Personal and Ubiquitous Computing, vol. 17, no. 6, pp. 1063– 1072.
27. G. Diraco, A. Leone, and P. Siciliano (2010), “An active vision system for fall detection and posture recognition in elderly healthcare”, in Proceedings of theDesign, Automation and Test in Europe Conference andExhibition, DATE 2010, pp. 1536–1541.
28. G. Mastorakis and D. Makris (2012), “Fall detection system using Kinect’s infrared sensor”, Journal of Real-Time Image Processing, vol. 9, no. 4, pp. 635– 646.
29. T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma (2015), “PCANet: A simple Deep Learning baseline forImage Classification?” IEEE Transactions on ImageProcessing, vol. 24, no. 12, pp. 5017–5032.
30. K. Simonyan and A. Zisserman (2014), “Very deep convolutional networks for large-scale image recognition”.
31. Y. Jia, E. Shelhamer, J. Donahue et al. (2014), “Caffe: convolutional architecture for fast feature embedding”, in Proceedings of the ACM International Conference on Multimedia, pp. 675– 678.
32. Adrián Núñez-Marcos, Gorka Azkune, Ignacio Arganda-Carreras (2017), "Vision-Based Fall Detection with Convolutional Neural Networks", Wireless Communications and Mobile Computing, vol.2017, Article ID 9474806, 16 pages.
33. Gupta R., Anand P., Chaudhury S., Lall B., Singh S. (2020), “Compressive Sensing Based Privacy for FallDetection”. Computer Vision, Pattern Recognition, Image Processing, and Graphics. NCVPRIPG 2019, Communications in Computer and Information Science, vol 1249.
34. K. de Miguel, A. Brunete, M. Hernando, and E. Gambao (2017), “Home Camera-Based Fall Detection System for the Elderly”, Sensors, vol. 17, no. 12, p. 2864.
35. Mucherino A., Papajorgji P.J., Pardalos P.M. (2009), “k-Nearest Neighbor Classification”. In: Data Mining inAgriculture. Springer Optimization and Its Applications,vol 34.
36. X. Kong, L. Meng and H. Tomiyama (2017), “Fall detection for elderly persons using a depth camera”, 2017 International Conference on Advanced Mechatronic Systems (ICAMechS), Xiamen, China, 2017, pp. 269-273.
37. Lezzar, Fouzi & Benmerzoug, Djamel & Kitouni, Ilham(2020), “Camera-Based Fall Detection System for the Elderly with Occlusion Recognition”. Applied Medical Informatics. Vol.42, no. 3, pp. 169-179.
38. F. A. Hamprecht, C. Schnörr, and B. Jähne (2007), “A duality based approach for realtime TV-L 1 optical flow”, in Pattern Recognition, pp. 214–223.
39. Xialong Xia,Cui Xu,Bing Nan (2017), “Inception- v3for Flower Classification”, IEEE 2nd International Conference on Image Vision and Computing(ICIVC).
40. F. A. Hamprecht, C. Schnörr, and B. Jähne (2007), “A duality based approach for realtime TV-L 1 optical flow”, in Pattern Recognition, pp. 214–223.
41. Bogdan Kwolek, Michal Kepski (2014), “Human fall detection on embedded platform using depth maps and wireless accelerometer”, Computer Methods andPrograms in Biomedicine, Volume 117, no. 3, pp 489- 501.