e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.3 ,Mar 2022
  UGC CARE Listed (Group -I) Journal

AN OPTIMIZATION OF ELDERLY FALL DETECTION SYSTEM USING OPENPOSE SKELETON MODEL

S.Vijaya Kumar [1], Dr.R. Shenbagavalli [2]

[1]*Research Scholar,Part Time Extenal* [Reg No. :17231172161017] *Dept. of Computer Science , Rani Anna Govt. college for women, (Affiliated to Manonmaniam Sundaranar University), Tirunelveli, India. svjaykumar@gmail.com*

[2]*Assistant Professor, Department of Computer Science, Rani Anna Govt. college for women, (Affiliated to Manonmaniam Sundaranar University), Tirunelveli, India. Rtillam3000@yahoo.com*

Abstract

For the elderly, falls pose a serious public health danger on a global scale. If a fall is not prevented in time, it may significantly reduce an elderly person's mobility, independence, and quality of life. Battery life, user discomfort, expense, complexity of installation, furniture occlusion, and computational demands are only some of the issues with today's detection technologies. In this chapter, we offer a sensor-free method of detecting falls with the use of computer vision and applied machine learning techniques. We used OpenPose real-time multi-person 2D pose estimation to track an object's motion across two datasets totaling 570×30 frames captured in five rooms from eight vantage points[3,4]. A total of 13 joint points in the human leg are retrieved by the device, and any shifts in those points indicate human movement. The system is accurate in its identification of human joints and is able to filter out background noise to do so. By focusing on joint points rather than images, we may reduce training time and avoid the drawbacks of conventional image-based methods such as motion blur, lighting, and shadows. The usage of single-view images in this work helps to cut down on costly machinery. To study the dynamic changes in human joint points over time, we tried out time series recurrent neural network (RNN), long- and short-term memory (LSTM), and convolution neural network (CNN) models. The experimental findings demonstrate that the suggested model has superior fall detection accuracy to the state-of-the-art methods.

INTRODUCTION

As the world's ageing population continues to expand at an alarming rate in the twenty-first century, multidisciplinary development in fields such as healthcare and tracking systems has accelerated. Falls are a major source of injury and death in the aged population. More than one-third of seniors who live at home and two-thirds of those in residential care fall at least once a year. More than two-thirds of persons who have fallen are at risk of falling again within a year after the occurrence[5,6]. The psychological implications frequently result in reduced mobility and independence among the elderly. Fall-related injury is one of the top 20 most expensive medical diseases[1] among the elderly who live in the community. Most senior persons are unable to get up on their own after a fall, and it has been observed that half of those who spent a lengthy period of time laying on the floor (>1 h) died within six months of the occurrence, even if they had no physical injuries. It is defined as "An occurrence that causes a person to come to rest unintentionally on the ground or other lower level". A person's posture and shape change during a fall, which can happen in a fraction of a second but usually takes between 0.45 and 0.85 seconds. When it comes to spotting a fall, these alterations are critical.

PROBLEM STATEMENT

The deployment of sensory devices such as cameras, microphones, infrared, and pressure sensors to follow the movement of individuals in restricted locations is referred to as a context-aware system. These systems are prone to difficulties such as restricted coverage, expensive installation costs, a large number of false alerts caused by other mobile entities, and privacy. Even though wearable-based FDS have some advantages over video-based systems, they do have significant drawbacks. Existing wearable-based FDS depended on the user being awake and attentive to press the button and inform emergency responders when they fell. It was found that the pushbuttons were useless if the patient was unable to think or was disoriented due to panic, or if the button was accidentally pressed, and the device was not worn by the user when he or she fell into the water during a fall. Furthermore, the installation of more sensors makes the user uncomfortable and inconvenient. If the user is using a battery-powered sensor, they must stop using it while the batteries are being recharged or replaced. These are some of the problems that users encounter when utilizing wearable FDS. Hence, in this research, we investigate about the fall detection framework using open pose skeleton and LSTM/CNN models.

PROPOSED METHODOLOGY

The proposed approach retrieves skeletal joint points through preprocessing using an existing UR Fall detection dataset. Missing data may be reconstructed from the processed joint points data. They are then separated into training and testing sets. The data processing, training, and testing method block diagram is shown in Figure 1 The videos in the dataset are first broken down into continuous picture sequences inside the framework. It utilizes OpenPose to get the coordinates of 13 points on the human leg skeleton from a series of images. We suggested a minimum and maximum normalization technique to preprocess the joint point data. The processed joint point data were then entered into the LSTM model in order to undergo training and testing. The data were then separated into a test set and a training set.

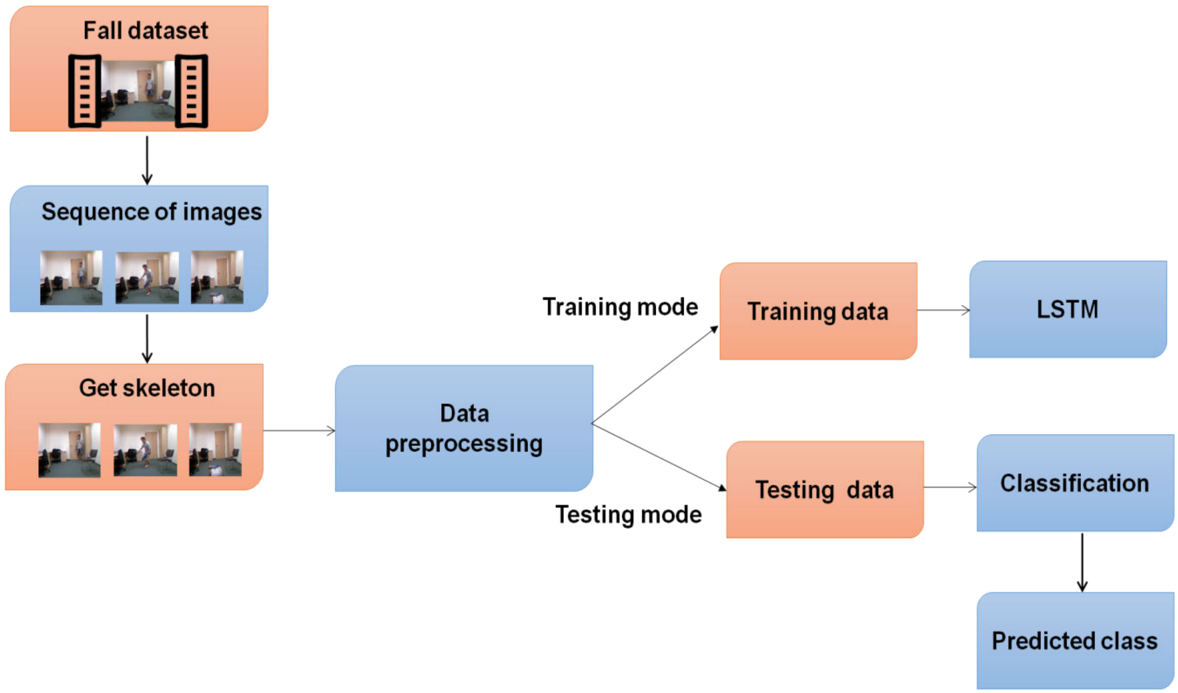


Figure 1 Flow of the proposed methodology

Retrieval of human skeleton using openpose skeleton method

For 2D pose estimation in real time with several subjects, Human body motions, face emotions, and finger motions may all be used to estimate the user's posture[2]. It has a high recognition rate and works well in both single- and multi-user environments.

The human skeleton is determined through the OpenPose algorithm's use of Part Affinity Fields (PAFs) and a Confidence Map. A series of 2D Confidence Maps make up the first branch, which is used to forecast joint parts, while PAFs make up the second branch, which is used to forecast limb parts. In order to create 2D skeletons for the characters in the figure, a greedy approach is used to examine the confidence map and PAFs. There are 13 estimated joint locations in the human leg according to OpenPose, as shown in Figure 2

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.3 ,Mar 2022
  UGC CARE Listed (Group -I) Journal

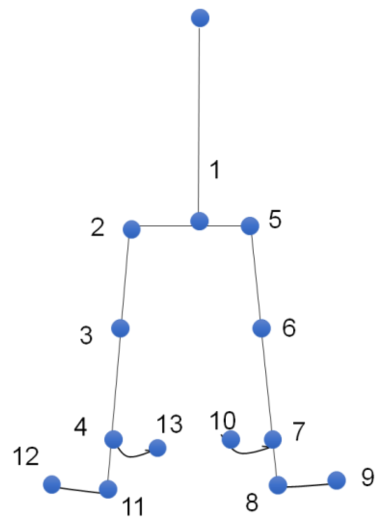
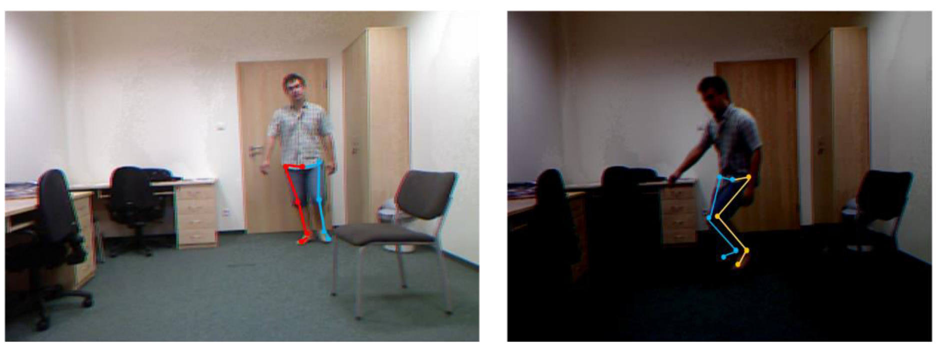


Figure 2 Joint locations in the human leg

OpenPose's captured joint points are shown in Figure 3 It demonstrates that the approach is able to locate human joints in a variety of settings, including low-light conditions (b). Each frame produces an $M_3 \times_{25}$ matrix that contains the (x, y) coordinates of the 25 joint points together with the prediction score for that joint point. A higher score indicates a more precisely located joint.



(a) (b)

Figure 3 Joint points captured by OpenPose (a) Bright ambient light (b) Dim ambient light

RESULTS AND DISCUSSION

In this section, we discuss in detail about the findings of the fall detection. We use three different models, namely RNN, LSTM, and CNN for experiments and compare their performance in fall detection.

These are divided into four prediction situations: (True Positive (TP)), (True Negative (TN)), (False Positive (FP)) and (False Negative (FN)). TP is the main target of correct prediction, TN is the secondary target of correct prediction, FP is the main target of wrong prediction, and FN is the secondary target of wrong prediction.

- **SENSITIVITY**

Sensitivity refers to the proportion of all samples that are actually the main classification target (positive), which is predicted to be the main classification target. If the true event is a fall (positive), the model predicts a fall as well. Sensitivity is defined as the ratio of the two is

$$Sensitivity = \frac{TP}{TP+FN} \times 100\% \quad (1)$$

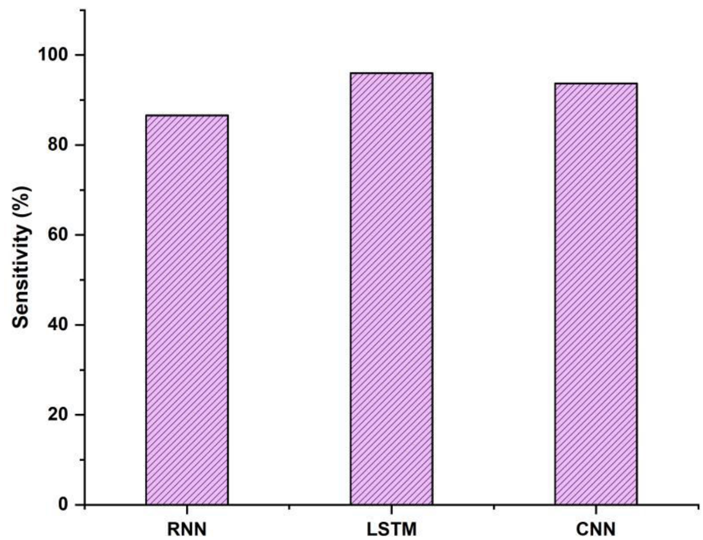


Figure 4 Comparison of sensitivity

Figure 4 shows the sensitivity comparison of the proposed method with certain conventional methods. As shown in figure 4 we can see that the sensitivity is 86% for RNN, 95% for LSTM and 92% for CNN. From the above comparison, we can clearly see that LSTM has highest sensitivity.

- **SPECIFICITY**

Specificity refers to the proportion of all samples that are actually secondary classification targets (negative) judged to be secondary classification targets. If the true event is non-fall (negative), the model prediction is also non-fall, and the ratio of the two is

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.3 ,Mar 2022
Research Paper UGC CARE Listed (Group -I) Journal

$$Specificity = \frac{TN}{TN+FP} \times 100\% \qquad (2)$$

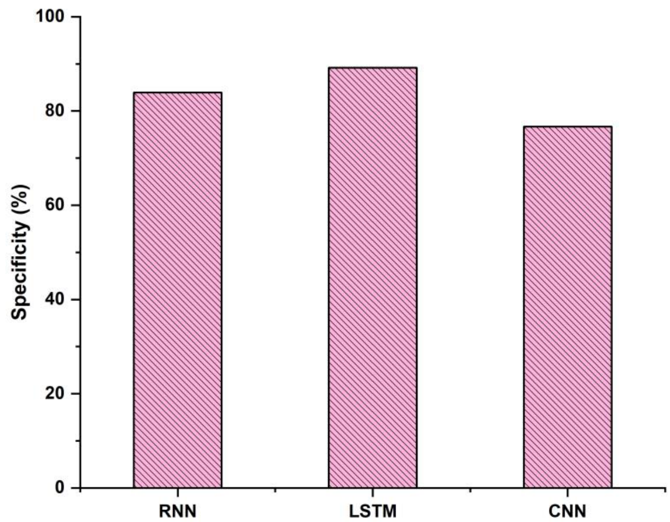


Figure 5 Comparison of specificity

Figure 4.10 shows the comparison of the proposed method's specificity to certain conventional methods. As shown in figure 4.10, we can see that the specificity is 83% for RNN, 89% for LSTM and 77% for CNN. From the above comparison, we can clearly see that LSTM has highest specificity.

- **ACCURACY**

Accuracy refers to the proportion of the primary and secondary classification targets that are correctly judged in all classified samples. Like all events, the model correctly predicts the proportion of falls and non-falls

$$Accuracy = \frac{TN+TP}{TP+TN+FP+FN} \times 100\% \qquad (3)$$

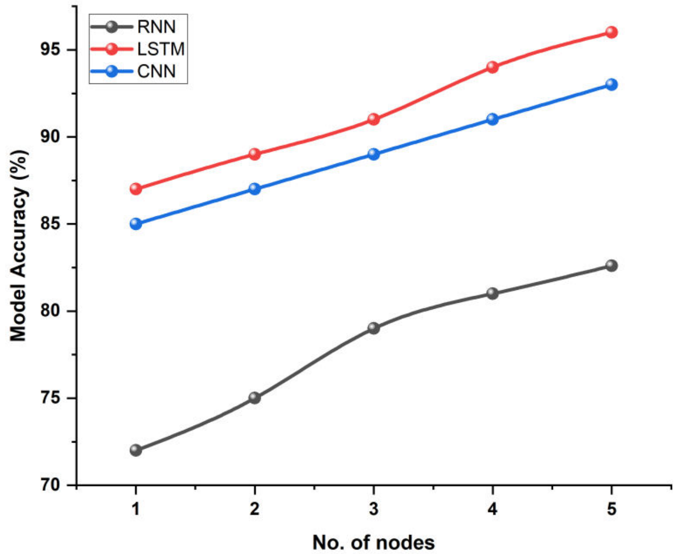


Figure 6 Model accuracy vs. number of nodes

As shown in figure 6, we can see that the accuracy is 82% for RNN. LSTM, with an accuracy of 96%. We hypothesize that this may be due to the complex structure of the LSTM model, and when too many Nodes are used it may over-describe the event, thus causing a decrease in accuracy. We observe the same trend for CNN and LSTM. The highest accuracy of CNN is 93%.

- **LEARNING RATE COMPARISON**

In this research, we test when the neuron is 512 under the learning rate of 0.1, 0.01 and 0.001 to find the most suitable learning parameters. From figure 7, we can see that RNN has the highest accuracy of 89% when LR = 0.1, while LSTM and CNN achieves higher accuracy when LR = 0.01. The accuracy is 98.3% and 96.6%, respectively

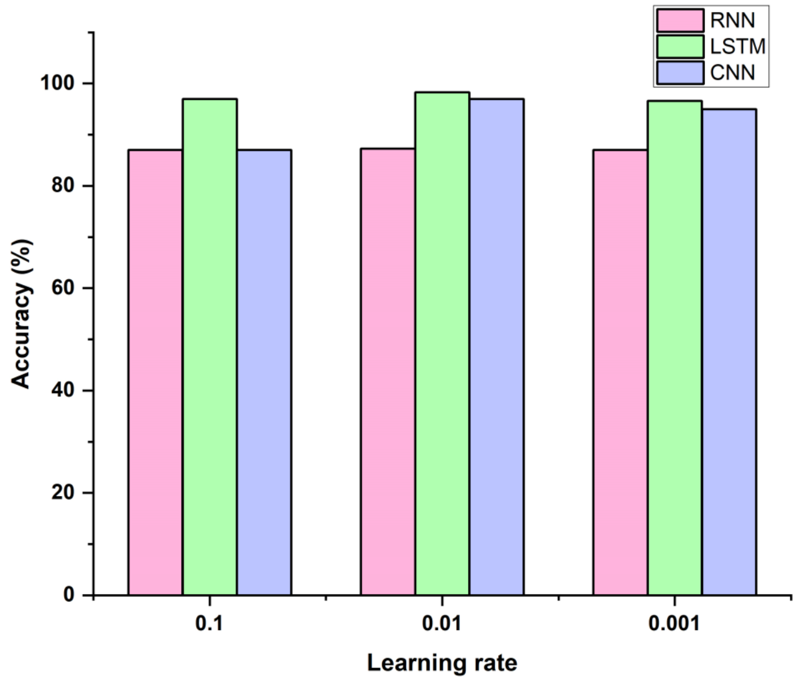


Figure 7 Comparison of average accuracy of model learning rate (RP)

CONCLUSION

This research proposes a fall detection framework using OpenPose to extract human skeleton from captured frames and RNN, LSTM and CNN to detect fall based on skeleton data. Due to the correlation of the human skeleton, we believe that the change trajectory of the joint points is related to the human movement. This method has been shown to work well in a complex space environment with lower equipment costs. However, the joint points will be lost in some postures and actions, causing the model to have coagulation during training. Therefore, we will move the joint points to the set relative position and interpolate part of the data to improve the excessive noise and unnormalized happening. Recursive neural networks can effectively help learn the time changes in human joint points and reduce training time. Therefore, this paper proposes a fall detection model that learns the changes in human joint points in continuous time. The experimental results show that the fall detection accuracy of the proposed model is far better than conventional techniques.

REFERENCES

[1].Romeo, L.; Marani, R.; Lorusso, N.; Angelillo, M.T.; Cicirelli, G. Vision-based Assessment of Balance Control in Elderly People.In Proceedings of the 2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Bari, Italy,1–3 July 2020; pp. 1–6.

[2].Cao, Z.; Simon, T.; Wei, S.; Sheikh, Y. Realtime multi-person 2D Pose Estimation Using Part Affinity Fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1302–1310.

[3].Kwolek, B.; Kepski, M. Human Fall Detection on Embedded Platform Using Depth Maps and Wireless Accelerometer. Computer. Methods Programs Biomed. 2014, 117, 489–501. [CrossRef] [PubMed]

[4]. Adhikari, K.; Bouchachia, H.; Nait-Charif, H. Activity Recognition for Indoor Fall Detection Using Convolutional Neural Network.

[5].The global burden of falls: global, regional and national estimates of morbidity and mortality from the Global Burden of Disease Study 2017 Injury prevention, 26(Suppl 2), pp.i3-i11.

[6].Journal of the American Medical Directors Association.Zhang, L., Ding, Z., Qiu, L., and Li, A., 2019 Falls and risk factors of falls for urban and rural community-dwelling older adults in ChinaBMC Geriatrics, 19(1), pp.1-17.

Authors Profile

S.Vijayakumar is a Ph.D Research Scholar Part Time Extenal 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.