Fall Detection using OpenPose

Divya R
Asst. Professor
Dept. of Computer Science and Engineering
SCET , Thrissur

Riya T B , Rona Johns P , Sreelakshmi T J , Theres Davies
UG Students
Dept. Of Computer Science and Engineering
SCET , Thrissur

Abstract:- Falls are a fatal threat to the elderly peoples health. It is a notable cause of morbidity and mortality in elders. Falls can even lead to serious injuries and death of the person , if they are not given proper attention. Above 30% of persons aged 65 years or above , fall each year and they mostly are reoccurring. The severity of such falls are due to the increasing age, cognitive impairment and sensory deficits. A multidisciplinary approach should be developed to prevent future falls. This paper emphasizes the need and development of an advanced fall detection system using Machine Learning and Artificial Intelligence technologies. The fall detection systems are currently categorized into wearable and non-wearable devices existing in the market. These wearable devices use sensors which may not be accurate always and it would be difficult for the elderly person to wear it around their body all the time. The architecture that is proposed in this paper uses open source libraries such as OpenPose for a much better detection and alert system, among non-wearable devices. The system retrieves the locations of 18 joint points of the human body and detects human movement through detecting its location changes. The system is able to effectively identify the various joints of the human body as well as eliminating environmental noise for an improved accuracy. This results in improved effective training time as well as eliminating blurriness, light, and shadows. The developed approach falls within the scope of computer vision-based human activity recognition and has attracted a lot of interest.

Keywords:- *OpenPose ; OpenCV ; Fall detection ; Artificial Intelligence ; Human Action Recognition ; Convolutional Neural Networks ; LSTM ; Image preprocessing ; Recurrent Neural Network.*

I. INTRODUCTION

The ageing of the population is caused by a decrease in the birth rate and a longer life span. According to studies, the senior population will grow substantially in the future, with the share of senior people in the global population reaching 28% in 2050. Human function declines as people age, increasing the chance of falling. Fractures and other long-term diseases are the result of falls. It can result in disability, loss of freedom, and a psychological fear of re-falling. Falls not only hurt the elderly, but they also put a mental and financial strain on the person and their family. As a result, a reliable fall detection system and emergency aid are required.

Human Activity Recognition (HAR) is a vast topic of research that tries to categorise statistical activities. Normal bodily movements such as standing up, sitting down, leaping, walking, and so on can be included in these exercises. When a person does an action, it normally takes a few seconds. Such photos are not classified into an activity class by image classification techniques. To recognise the movements, existing systems rely on sensor data acquired by accelerometers, smart phones, or similar harnessing devices. The collection of such data is difficult and expensive, requiring several sensors, specialised technology, and software.

A. PROBLEM DOMAIN

Existing systems uses image processing techniques for human action and pose detection. In earlier methods, they capture human images and process the image using techniques like Artificial Neural Networks to get the output. The human actions are happening simultaneously in a fraction of seconds. There fore existing method is time consuming and less effective. Some of the existing systems uses hardware devices like movement sensors. These sensors can be fit onto the human body and can look for shakes. When a particular level of shake is detected, then it is guessed as a fall detection. This is not accurate as every shakes may not be falls.

In the marketplace, there are a variety of products to choose from. The main drawback of such goods is that they necessitate large hardware components and would be prohibitively expensive. Some of the systems use base stations and Radio Frequency, which are hardwired to a phone line and then call a call centre for assistance. The problem is that they all require an intermediary call service, which costs a lot of money each month and is confined to the range of a house because they rely on a central base station for outside communication.

The Human Activity Recognition system continuously monitors human actions, which could be useful in senior surveillance. It aids in health-care surveillance, aberrant behaviour detection, identity, mental state awareness, and geriatric care. Despite the progress made in this field, substantial dependencies remain, such as data intake from wearable sensors and their installation, accuracy, and annoyance. Sensors are placed throughout the body to collect input data. Smart phones and smart watches take data from accelerometers, gyroscopes, and other sensors.

Over a period of time, the Patterns are examined. Sleeping activities are frequently discovered by interpreting minute motions. The person is identified, a unique identity is created, and the activity is checked out by the video surveillance system. Due to background noise, subject occlusions, low light circumstances, and the difficulty in tracking a person, developing a low-error system is tough. The suggested classification system for human activities looks promising and addresses the issues raised by the previous systems. The data is collected in real time with the help of connected cameras. This eliminates the need for the wearable sensors to be connected. Openpose derives keypoint properties from the skeleton form and handles situations such as ground noise and distorted light. As a result, the suggested algorithm is robust. As a result, a new fall detection method is necessary.

B. PROPOSED SYSTEM

Human Pose Estimation is a fascinating area in Artificial Intelligence. The practise of locating human joints from photographs or movies is known as pose estimation. It's also known as the process of looking for a certain posture among all articulated positions. The method considers human joints, which are frequently referred to as key points, and estimates the pose from a live video.

Openpose is utilised in the proposed method to determine the exact human position. OpenPose is a popular bottom-up methodology for estimating multi-person human poses. It's an open source, real-time project that uses a single photograph to recognise the human body, including hands, face, facial expressions, and legs. OpenPose, like many other bottom-up techniques, first discovers keypoints that belong to each person in the image, then assigns keypoints to separate individuals. The suggested system takes the user's video and sends it to the openpose estimater for analysis. The COCO dataset is used with the openpose estimator, which is a pre-trained model. (The coco dataset is made up of 18 different human activities that were recorded under rigorous conditions.) When the user tries to wake up, the openpose estimater, which is running in the background, collects these behaviours and generates an alarm.

A graphical representation of the Human Pose Skeleton is available. The set of coordinates is linked to describe the person's stance. A joint is the name given to each co-ordinate in the skeleton (also known as keypoint). A pair is a legitimate connection between two joints (or a limb). Not every part combination would be a good match. Action recognition, animation, gaming, and other applications are among them. The first method approximated a single person's stance in an image. These methods work by first identifying the individual keypoints, then connecting them to construct the pose.

Using OpenPose and Long Short-Term Memory Networks, this research describes a systematic way for recognising human behaviours in real time. This method is based on real-time images collected by the camera. Using a sliding window technique, the output of body features would be separated into sub-sequences called windows. The LSTM (Long Short Term Memory) is a type of recurrent neural network (RNN). Feed forward neural networks have feedback connections, allowing them to handle both single data points and sequences of data. The LSTM algorithm is appropriate since it quickly learns the essential point attributes and returns an activity class. Running, sitting, standing, waving one hand, waving two hands, jumping, and clapping are all detected in real time by this system.

This approach examines each frame acquired by a camera, extracting skeletal data from humans using the OpenPose skeleton extraction algorithm and displaying it on the screen. Each node in the plane coordinate system is represented by its horizontal and vertical coordinates. The velocity of descent at the centre of the hip joints, the human body centre line angle with the ground, and the width-to-height ratio of the human body exterior rectangle are used to identify falling behaviour. The following is how the paper is structured: The second section examines the system's flow and architecture. The working explanation is detailed in Section 3. The software is described in Section 4. Section 5 discusses the conclusion.

II. SYSTEM ARCHITECTURE

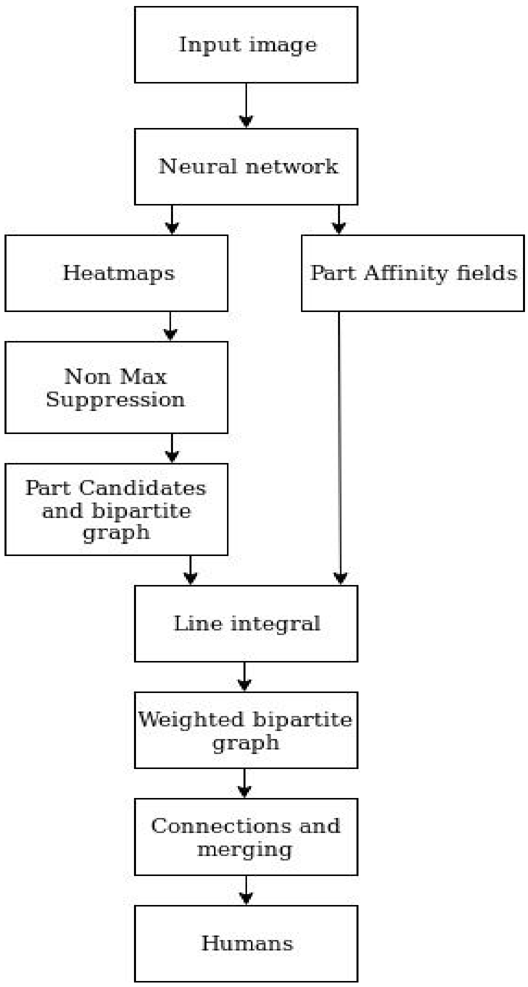

Fig 1: Flow diagram

III. WORKING EXPLANATION

A. PSEUDOCODE

1. Identifying two distinct activities Fall action (FA) and current action (CA)
2. Begin recording video using the system camera.
3. Launch the OpenPose library.
4. Using the video captured, draw the essential points of the human body.
5. The keypoints are linked together, resulting in CA:
   if CA==FA,
      a fall has been detected, and an email has been issued;
   else,
      video recording will continue.
6. End.

- Object detection and extraction from live video: The live video was collected using the system web camera, and the object (human body) was detected and extracted from the backdrop using OpenCV algorithms.
- Image Pre-processing: From the live video, the collected image (16 frames per second) is converted to RGB format and prepared for processing. For processing, the image is resized to 250x250 pixels.
- OpenPose is a programme that determines the exact human posture. OpenPose is a bottom-up method for estimating multi-person human poses. OpenPose finds the keypoints, which are mostly the joints of the human body skeleton belonging to each individual in the retrieved image. Get crucial features including thigh deflection angles, calf deflection angles, spine deflection angle, and spine ratio using OpenPose. These features are calculated, and the body position is determined depending on the features.
- As demonstrated in Fig 2, the OpenPose estimater, which is a pre-trained model, works with the COCO dataset to extract 18 essential points of the human anatomy from a 2D image. The attributes collected from the collected live video are then compared to the specified action to determine whether or not the current action is a fall. If the system detects a fall, it will emit a buzzer sound and send an alert notification to the provided email address; otherwise, it will repeat the capture procedure.

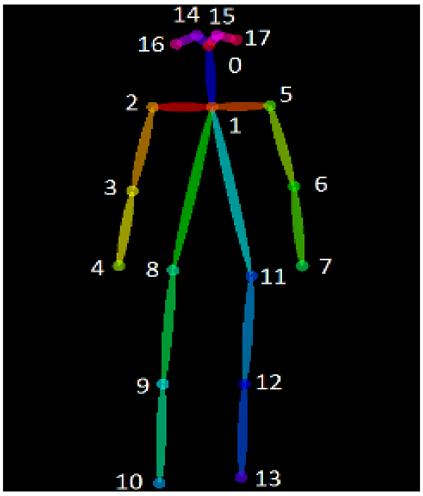

Fig 2: Keypoints

IV. HARDWARE REQUIREMENTS

- Processor : i5
- RAM : 4 GB
- Hard Disk : 500 GB

V. SOFTWARE REQUIREMENTS

- Front end : Python

Python is a general-purpose programming language that may be used for a variety of tasks, including back-end development, software development, and writing system scripts. Python 3.7 was used in this project.

- Back end : MySql

Python programmes can use MySQL connector/Python to connect to a MySQL database.

- IDE : Spyder

It's a cross-platform open source IDE. The Python Spyder IDE is entirely written in Python. Spyder is best used with Jupyter notebooks or other scientific computing tools such as Anaconda, rather than as a general Python development environment. It was created by scientists and is primarily composed of scientists, data analysts, and engineers.

- Tools : OpenPose

Openpose is a technique for determining a person's exact position. OpenPose is a popular bottom-up methodology for estimating multi-person human poses. It's an open source, real-time project that uses a single photograph to recognise the human body, including hands, face, facial expressions, and legs. The COCO dataset was used to train the Openpose model to extract 18 body key points.

VI. CONCLUSION

Falling and being seriously injured are two of the most important public health issues in an ageing society. It is required to determine the features of the fall movement in order to detect a fall. This study presents a method for detecting human body keypoints from collected frames using the TensorFlow OpenPose module. This approach extracts the human skeleton and detects falls using skeleton data, i.e. the changing trajectory of joint points is related to human movements. The suggested system employs computer vision technologies in conjunction with artificial intelligence libraries such as OpenPose to detect human falls and provide timely alerts to users. Because OpenPose extracts the body main point features in skeleton form, which can handle instances like background noise and low light distorted situations, this method has been proved to operate well in a complex space environment with lower equipment costs.

In some postures and motions, the joint points can become forgotten, causing the model to coagulate throughout training. To circumvent this, set the relative position of the joint points and interpolate a portion of the data to reduce excessive noise and unnormalized events. Recursive neural networks can aid in the learning of time changes in human joint points while also shortening training time. The fall detection model proposed in this paper learns the changes in human joint points over time. The falling motion is then recognised by setting three conditions: (a) the speed of descent at the centre of the hip joint, (b) the angle between the human body's centerline and the ground, and (c) the exterior rectangular width-to-height ratio. Based on the awareness of falls and consideration of the condition of persons rising up after falls, the process of standing up after falls is seen as an inverse process of falling. As a result, the suggested algorithm is robust. Experiments have proven that the procedure works and that it produces the desired result, as well as a proper alert.

ACKNOWLEDGMENT

We want to offer our heartfelt gratitude and heartfelt thanks to everyone who helped us make this initiative a huge success. We thank the almighty god for all the benefits he has bestowed upon us. Prof. Divya R, our guide, deserves a huge thanking . We want to thank our Principal, Dr. Nixon Kuruvila, and our Department Head, Dr. M Rajeswari, for their unwavering support throughout the project.

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]. Lie, W.; Le, A.T.; Lin, G. Human Fall-Down Event Detection Based on 2D Skeletons and Deep Learning Approach. In Proceedings of the 2018 International Workshop on Advanced Image Technology (IWAIT), Chiang Mai, Thailand, 7–10 January 2018; pp. 1–4.

[3]. Lin, C.; Wang, S.; Hong, J.; Kang, L.; Huang, C. Vision-Based Fall Detection through Shape Features. In Proceedings of the IEEE Second International Conference on Multimedia Big Data (BigMM), Taipei, Taiwan, 20–22 April 2016; pp. 237–240.

[4]. M. Zadghorban and M. Nahvi, (2018), An algorithm on sign words extraction and recognition of continuous Persian sign language based on motion and shape features of hands Pattern Anal. Appl., vol. 21, no. 2, pp. 323–335.

[5]. Z. Wu, Y. Huang, L. Wang, X. Wang, and T. Tan, (2017) A comprehensive study on cross-view gait based human identification with deep CNNsIEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 2, pp. 209–226.

[6]. P. Ji, C. Wu, H. Li, and X. Xu}, (2016) Vision-based posture recognition using an ensemble classifier and a vote filter Proc. SPIE, vol. 157, pp. 101571J-101584J.

[7]. J. C. Núñez, R. Cabido, J. J. Pantrigo, A. S. Montemayor, and J. F. Vélez, (2018) Convolutional neural networks and long short-term memory for skeleton-based human activity and hand gesture recognition Pattern Recognit.vol. 76, pp. 80–94.

[8]. Q. Chen, Y. H. Chen, and W. G. Jiang, (2015) The change detection of high spatial resolution remotely sensed imagery based on OB-HMAD algorithm and spectral features Guang Pu Xue Yu Guang Pu Fen Xi= Guang Pu, vol. 35, no. 6, pp. 1709–1714.

[9]. Y. Li, Q. Miao, K. Tian, Y. Fan, X. Xu, Z. Ma, and J. Song,}, (2019), Y. Li, Q. Miao, K. Tian, Y. Fan, X. Xu, Z. Ma, and J. Song. Lett., vol. 119, pp. 187–194.

[10]. B. Xie, X. He, and Y. Li, (2018), RGB-D static gesture recognition based on convolutional neural network J. Eng., vol. 2018, no. 16, pp. 1515–1520.

[11]. M. Arsalan and A. Santra, (2019), Character recognition in air-writing based on network of radars for human-machine interfaceIEEE Sensors J., vol. 19, no. 19, pp. 8855–8864.

[12]. S. Huo, T. Hu, and C. Li, (2017) Improved collaborative representation classifier based on l2-regularized for human action recognition J. Electr. Comput. Eng., vol. 2017, pp. 1–6.

[13]. J. Wang, T. Zheng, and P. Lei, (2018), Hand gesture recognition method by radar based on convolutional neural network J. Bjing Univ. Aeronaut. Astronaut., vol. 44, no. 6, pp. 1117–1123.

[14]. E. Tsironi, P. Barros, C. Weber, and S. Wermter, (2017), An analysis of convolutional long short-term memory recurrent neural networks for gesture recognition Neurocomputing, vol. 268, pp. 76–86.

[15]. S. Y. Kim, H. G. Han, J. W. Kim, S. Lee, and T. W. Kim, (2017), A hand gesture recognition sensor using reflected impulses IEEE Sensors J., vol. 17, no. 10, pp. 2975–2976.

[16]. C. D. Santos, J. L. A. Samatelo, and R. F. Vassallo (2020), Dynamic gesture recognition by using CNNs and star RGB: A temporal information condensation Neuro computing, vol. 400, pp. 238–254.

[17]. L. Chen, J. Fu, H. Li, B. Zheng, and Y. Wu (2020) Hand gesture recognition using compact CNN via surface electromyography signalsSensors, vol. 20, no. 3, pp. 672–683.

[18]. Tsai, T.; Hsu, C. Implementation of Fall Detection System Based on 3D Skeleton for Deep Learning Technique. IEEE Access 2019, 7, 153049–153059.

[19]. S. K. Leem, F. Khan, and S. H. Cho, (2020), Detecting mid-air gestures for digit writing with radio sensors and a CNN IEEE Trans. Instrum. Meas vol. 69, no. 4, pp. 1066–1081.

[20]. X. Zhang and X. Li, (2019), Dynamic gesture recognition based onFuture Internet, vol. 11, no. 4, pp. 91–102.