[1]Dr. Rahul Patil

[2]Dr. Prashant Ahire

[3]Dr. G. B. Sambare

[4]Dr. Yashwant Dongre

[5]Mrs. Aarti S. Gaikwad

[6]Nishant Ashtekar

[7]Sushant Chaudhari

[8]Shailendra Chaudhari

[9]Abhijit Gadhave

Comparative Analysis of Human Pose Estimation Methods with Fall Detection and Smart Alarming System

Abstract: - The population of old age peoples are increasing day by day and it will grow up by 31% of today's population till 2031. In many cases the old age people are prone to fall. So, old age peoples can help in various aspect by using advanced technology. We already know that fall discovery is very much critical part of senior care as fall pose a threat to their health. To this challenge, this discussion and study proposes a fall detection model comprising three crucial phases: Pose Estimation, Fall Detection, and Smart Alerting System. Using computer vision model similar as OpenPose, PoseNet, MoveNet and YOLOv7. By the help of this models, we directly track the act of senior individualities in real time and identifies fall incidents grounded by abnormal changes in posture. When fall detect, the WhatsApp Alert message with fall detected frame and phone call will be received by Caregiver or their family members to take urgent actions. In case of fall, one can prevent from severe injuries which could be caused due to delay in mediation. These proposed models will enhance the safety of old age population.

Keywords: Pose Estimation, Fall Detection, and Smart Alerting System.

I. INTRODUCTION

Nowadays Senior Citizens and disabled people can overcome their disabilities in carrying out daily tasks in many facilities. But they have trouble being independent and sometimes they fall, and due to delay in late treatment after fall it increases the chances of severe injuries. Hence keeping this in mind an idea to develop a fall detection system emerged. This report mainly focuses on the problem we discussed. In recent years, artificial intelligence techniques have shown a remarkable detection of human fall in comparison to traditional techniques. Human Fall detection systems aim to reduce both dependency and care costs in the elderly community.

The proposed system consists of three stages namely pose estimation, fall detection, and alarming system. This system exceptionally proves itself better than existing solutions because no physical devices will be used. For eg. Smart Watches, Wrist Bands. Also, other available systems monitor only a single person in frame whether this system works for more than one person too. In the second stage i.e. fall detection a comparison between different pose detection models was made to find and provide the best suited solution for the problem. In the proposed

[1]Asst. Prof., Dept. Of Computer Engg., Pimpri Chinchwad College of Engineering, Pune., Email:patil.rahul3068@gmail.com

[2] Asst. Prof., Dept. Of Computer Engg., Pimpri Chinchwad College of Engineering, Pune., Email:prasshantahire@gmail.com

[3] Asso. Prof., Dept. Of Computer Engg., Pimpri Chinchwad College of Engineering, Pune., Email:santosh.sambare@pccoepune.org

[4] Asst. Prof., Dept. Of Computer Engg., Vishwakarma Institute of Information Technology, Pune., Email:yashwant.dongre@gmail.com

[5]Asst. Prof., Dept. Of Information Technology, D. Y. Patil College of Engg., Email:asgaikwad@dypcoeakurdi.ac.in

[6]Dept of Computer Engg. PimpriChinchwad College of Engineering, Pune.

[7]Dept of Computer Engg. PimpriChinchwad College of Engineering, Pune.

[8]Dept of Computer Engg. PimpriChinchwad College of Engineering, Pune.

[9]Dept of Computer Engg. PimpriChinchwad College of Engineering, Pune.

system Smart Alarming System plays the final and most essential role as it will to convey about the detected fall to the user's relatives or caregivers to provide the treatment.

II. RELATED WORK

A. State-of-the-arts

The authors of [1] have in-depth knowledge of sensor and low cost of the Kinect camera which can be used as a perfect option for detection of fall in home. To analyze the Kinect video streams, authors used computer vision algorithms. The computer vision method to detect falls from ordinary falls includes identifying a person's bones, and then applying machine learning algorithms. The possible accuracy of Kinect-based systems ranges from 80 to 95%, their disadvantages include a restricted range.

In [2] cameras and sensors like Kinect sensors are used for detection of fall. They examine several techniques of video data analysis, such as body identification, motion tracking, and machine learning to distinguish falls from everyday activities. The accuracy rates of Kinect-based sensors are 80–95%.

The authors of [3] research proposes an Internet of Things (IoT) system to track the movements and detect falls of senior people who live alone. The movement is tracked using cameras or sensors like motion detectors and then machine

In [4] the research investigates the main use of deep learning techniques to infer human body posture from pictures and videos. While showing the benefits of deep learning for posture estimation, the research also notes several difficulties, such as handling busy circumstances. It ends by encouraging further research into these problems and new possibilities for pose estimation based on deep learning.

The authors of [5] investigates the detection of fall which focus on data accessibility. It emphasizes how challenging the falls can be and how difficult it would be to record it for research. The authors emphasize that to increase the accuracy and effectiveness of fall detection systems, more additional data is required.

In [6] the study proposes a process on data on a cheap android box in which fall is detected using an eight-camera fall detection system. Rather than using wearable sensors, the techniques use machine learning models to identify fall patterns in the video streams. This method has several benefits, which includes the ability to see beyond transparent barriers, which can cover a huge distance (up to 60 meters), and can work in different lighting situations.

The authors of [7] researches how pose estimation can be used to identify falls. Fall detection systems depends heavily on pose estimation algorithms, which identify a person's joints inside a video frame. The study talks about how these methods provides useful information on how a person's body posture changes after a fall. It also draws attention to problems that might affect the accuracy, such as object interference and camera angles.

B. Research Gap Analysis

In [1] research gaps are found in the survey on Kinect cameras for detection of fall. The analysis reveals the limitations in the datasets that were employed, which may not accurately represent in real-world situations. Additionally, some user groups might not be able to utilize the current systems.

A potential gap or lacking found in [2] could be that there would be privacy issues even though using cameras for fall detection is a advantage. Understanding user preferences and the use of various systems: such as wearable and non-wearable devices are the main areas of attention for this gap.

The paper [3] has some possible weaknesses in the present technology which are studied on the application of Internet of Things (IoT) for detection of fall. Security and privacy of data is another problem. The study has highlighted the necessity of strong measures to protect private data of patients that is gathered by Internet of Things (IoT) devices.

In [4] the authors discuss about how pose estimation using deep learning algorithms still faces several challenges. In addition to that, accurate 3D reconstruction and 2D pose estimation are not the same. Lastly, additional study is required due to privacy concerns with camera-based systems.

A potential gap or lacking found in [5] mostly concentrates on the limited data. This study mainly focuses on the fact that current datasets might not include a range of fall kinds, such as user characteristics (age, physical ability), and contexts (at home vs. outside).

In [6] research on a machine learning-based eight-camera fall detection system presents a method, where might be some areas which may require additional investigation. Even though, the system can cover huge environment with eight cameras, still it may not be scalable in situations, such as home or any other facility. The paper [7] has some of the significant weakness in pose estimate for fall detection. Many rooms may have dim illumination, occlusions (items obscuring the body), or intricate backgrounds, because of these possible things pose estimation can be less accurate. The development of pose estimation methods for detection of fall which are generalizable is required.

C. Contributions

1. We have compared the various Pose Estimation models and selected the best suitable algorithm for the Human Pose Estimation. Our major contribution was to compare MoveNet Lightning, MoveNet Thunder, OpenPose, PoseNet and YOLOv7 based on factors such as multi-persion detection, Keypoints, accuracy, Response Time, speed, etc. We came to conclusion that to select such a model which is quick and can detect multiple persons.

2. We used Movenet Lightning for Pose Estimation which is fast and helpful in our project to detect the human pose based on centre heatmap so that it can detect human pose quickly and fall can be detected as early as possible.

3.We have also integrated our system with the smart alerting system. When there is any such case of fall the caregiver will be notified with the WhatsApp message with the frame as how the has fallen and with the point of impact and also, he will be alerted with the phone calls where he will receive the voice message.

III. PROPOSED SYSTEM

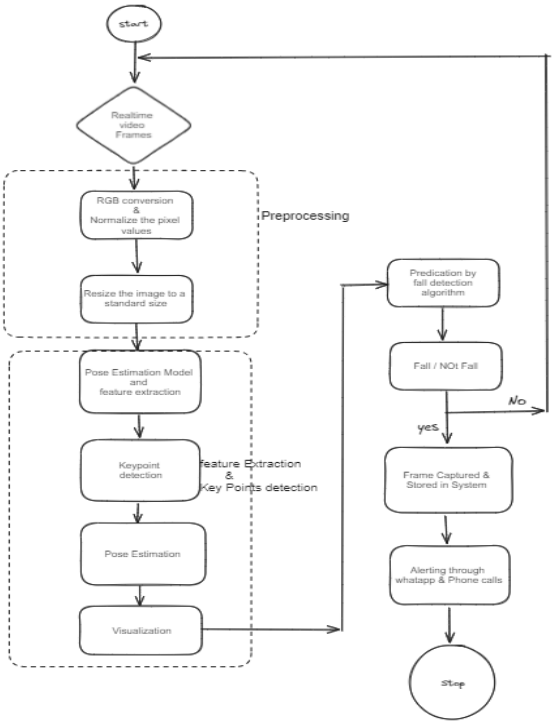

Figure. 1 Project Flow Diagram

Working:

(1) Real-time pose Estimation:

- Initializing the MoveNet model by loading a TensorFlow Lite and allocating tensors.
- Process the video frames captured by web cam, preprocessing it by normalizing and predicting key Points using MoveNet model.
- Draw keypoints and connecting the frames based on confidence scores.
- Show the real-time annoted video displaying the keypoints and their connection with the movement of body.

(2) Fall Detection Logic:

- Detecting keypoints and bounding boxes for each person in the frame.
- Determine if a person likely ot fall or not based on estimating the pose and certain threshold value).
- It the fall is detected, Save the frame, and break the loop.

(3) Emergency Notification System:

- Make an automated voice calls to the emergency number which is registered
- Send a WhatsApp message with image frame which saved after fall to detect the fall incident.

IV. PROPOSED METHODOLOGY

To identify the critical spots in the human body, we employed pose estimation techniques in the proposed approach. We utilized four distinct models for this: MoveNet, PoseNet, OpenPose, and Yolov7. Our comprehensive study of these models' operation yielded keypoint coordinates within an input frame as an output. For usage in applications, these pretrained models are stored on TensorFlow-Hub and are freely accessible to everybody.

Pose estimation in computer vision is an approach for predicting keypoints or joints in the human body. One way to anticipate bodily joints or keypoints is to identify keypoints such as the ear, nose, shoulder, knees, wrists, and so on. By first estimating the human posture in the frame and then identifying the fall by analyzing the coordinates of keypoints provided by the model, this technique increases accuracy and speed while reducing effort by doing away with the typical method of training the model using a big dataset.
The provided problem statement has been split up into three sections: Posture Estimation, Fall Detection, and Smart Alarm System.

A. Posture Estimation

For detecting objects and people in images or videos, pose estimation solution in computer vision is used. For example, it can find keypoints such as elbow, face, shoulder, etc in an image of a person. Gesture control, augmented reality, and action detection are the wide-range of application that uses the Pose estimation approaches. We deployed four different models, which are all available on Tensorflow-Hub for pose estimation. MoveNet, OpenPose, PoseNet and YOLOv7 are the models we compared for analysis.

1. MoveNet:

MoveNet is a pose estimation model available on TF Hub. There are two version MoveNet Thunder and MoveNet Lightning which are readily available to everyone and are available as pretrained model on TF Hub. These models return 17 keypoints, in which lightning model is used for latency-critical application where speed is important factor. Thunder is used where accuracy is important concern and is applied where precision is more essential than quickness. MoveNet uses bottom-up approach and parallel process for heatmap generation, keypoint regression, and background suppression which makes it faster. MobileNet is used in MoveNet as pre-trained feature extractor

which increases its efficiency. It helps in capturing complex human movements with less computational power, which makes it best suitable for use.

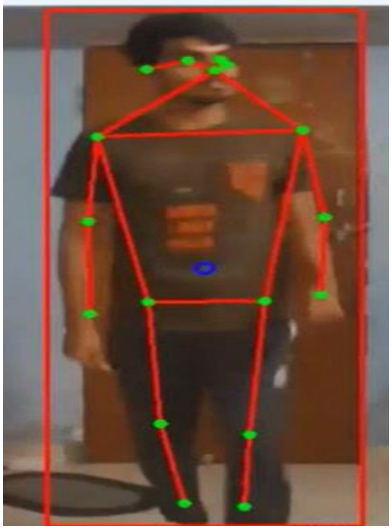

Figure 2. Pose Estimation using MoveNet Lightning

2. OpenPose:

OpenPose is the multi-person system which can find the joints and extremities in the body. Bottom-up approach is used in OpenPose. It is known for its versatility and broad applicability with various application. In a single input image, it can recognize 135 body keypoints. It uses a multi-stage architecture with several interconnected stages. As it finds more keypoints it is computationally expensive and slow as compare to MoveNet.

3. PoseNet:

PoseNet is built upon Convolutional Neural Network (CNN) from which postures can be identified from a single RGB image. Also, real-time systems have 5ms/frame speed. It finds keypoints on human body with a skeletal representation on it. For feature extraction and keypoint detection it includes multiple convolutional layers. It is more efficient and light weight as compared to OpensPose. PoseNet does not offer multi-person keypoints prediction capability.

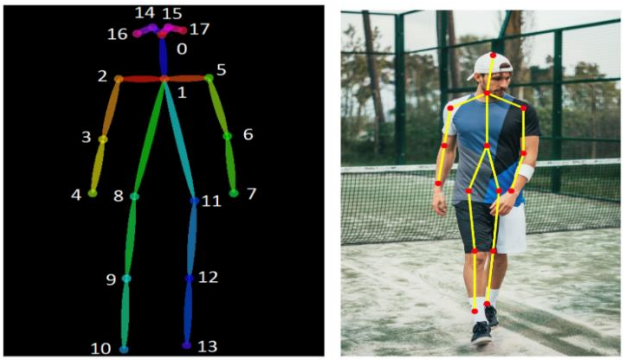

Figure 3. Pose Detection using PoseNet[12]

4. YOLOv7:

YOLOv7 is an object detection model focused on classifying and identifying objects within an image/video. Bounding Boxes and class probabilities are predicted directly from the grid cells by dividing an input image into grid. YOLOv7 uses deep neural network architecture with additional features of CSPDarknet53 and PANet. The primary focus of YOLOv7 is object detection and not pose estimation. It does not give the accurate results for pose estimation as compare to other Pose Estimation algorithms.

Figure 4 YOLOv7 detection [10]

Model	Key Points	Estimation Method	Speed	Multiple Person Detection	Accuracy
Movenet	17	Bottom-Up	Fastest	Yes	Moderate
OpenPose	137	Bottom-Up	Slower	Yes	High
PoseNet	17	Top-down	Moderate	No	Moderate
YOLOv7	17	Top Down	High for COCO and MPII benchmarks but slow for other cases	Yes	Not specified

Figure 5 Comparison table of Models

From above comparison table, we can conclude that Movenet Lightning is quick with good accuracy. As OpenPose has high accuracy but it will be not suitable for our application as its speed is very less. And PoseNet can be used for single person detection. Therefore, we selected MoveNet model for Pose estimation.The algorithmic steps for MoveNet models for Pose Estimations are:

Algorithm for Estimating Pose.
Input: Video/ Image frame **Output:** Returned the estimated keypoints locations.
1. import required libraries like tensorflow, matplotlib, sklearn, cv2, etc. #Loading movenet_lightning model 2. model_name = "movenet_lightning" #Initialize the TFLite interpreter 3. interpreter = tf.lite.Interpreter(model_path="model.tflite") #Architecture of Movenet Model #Extracting the feature of the image 4. features=features_extractor(input_image) #Allocating Scores to the Pixels to ignore background and focus on humans 5. center_heatmap= person_center_heatmap(input_image) # Using vector regression for the to estimate initial location for each keypoint from the predicted centre point 6. keypoint_regression= keypoint_regression_fields(features)

```
# Based onn above vectors keypoints are deployed on the
image
7. keypoint_heatmap= actual_keypoint_heatmap(features)
# Model defines the keypoint offset to make it more precise
and accurate
8. keypoint_offset= keypoint_offset_field(features)
```

B. Fall Detection:

We have performed fall detection after detecting keypoints with ratings obtained from our pose estimation methods. As discussed, we used outputs from the MoveNet Lightning model because this version may be very quick. We want short effects from the pose estimation models to provide instant hospital treatment to the users. We were given outputs in form of keypoints, for the MoveNet lightning we were given x and y coordinates of that 17 keypoints with the confidence score of that keypoint. That keypoints could be specially used to hit upon the autumn. We have used keypoints from the shoulder and nose to come across the drastic change within the coordinates of the shoulder and nostril.

For detection technique is used to detect the fall using various equipment such as cameras and sensors. As elderly people are very much prone to fall. The fall detection technique can be helpful to them. It will help in critical situation; the caregivers could be immediately notified regarding the falls and the immediate actions would be taken to be prevent from severe injuries to elderly people after the fall.

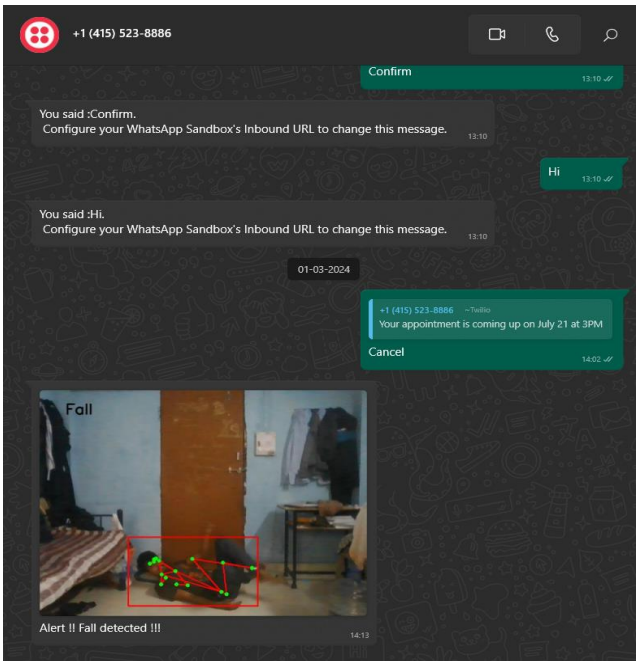

Figure 6 Fall Detected Alert Message to Caretakers

C. Smart Alarming System

In this part, we integrated our system with features to phone calls and WhatApp message to the caregivers or family members. Twilio is software to exchange call and send or receive messages. It is affordable software API. We used some twilio features such as whatApp messaging device and contact name machine to create alert signals. The care givers will receive realtime frame detected during the fall and help to find the point of impact.

Twilio can even play a computer-generated audio message after receiving the call to allow the person recognize about the fall detected. We have also helped fallen person to communicate with caregivers if he/she is fine or not.

V. EXPERIMENTAL RESULTS

We have used four kind of pose estimation models MoveNet, Openpose, Posenet and YOLOV7. We have examined that the MoveNet Models is the quickest pose estimator and especially the MoveNet Lightning version could be very fast as compared to the alternative fashions. The MoveNet thunder model offers better accuracy than the Lightning version but it's miles slow in comparison to the Lightning version. We have done more than one man or woman pose estimation using the MoveNet, Openpose and Yolov7 fashions and determined that these models can flawlessly come across a couple of character inside the given input frame. Openpose can stumble on multiple people in a body however it's far slower than the MoveNet Lightning. They have compared these models on each input pix in addition to on real time camera inputs [9]. The yolov7 model is used for item detection and presents a new model to carry out pose estimation which offers keypoints.

We have compared this model based on various factors such as accuracy, speed, Response-time, and various other factor which helps to identify which model will be the best and suitable for our Fall detection system.

They have used the coco and MPI datasheet and group the images in three parts where each part contains 1000 images. [9] We analyzed that all the models that are OpenPose, MoveNet lightning, moveNet Thunder, PoseNet and drawn some results.

Group1 – Image with single person

Group 2 – Image with Multiple person

Group 3 – Image without person

A. Total time and Standard Deviation to Estimate all groups (1000 images)

$$T_r^p = \Sigma \sum_{k=0}^{n-1} E(I_k^p)$$

T_r^p – Total time to estimate image., n – number of images in each part, and E(I_r^p) – Time to estimate image I_r^p in each part

$$C^p = \frac{e^p}{n^p}$$

e^p- number estimated images in part 1, n^p- number of images in part 1, And C^p – accuracy of each part.

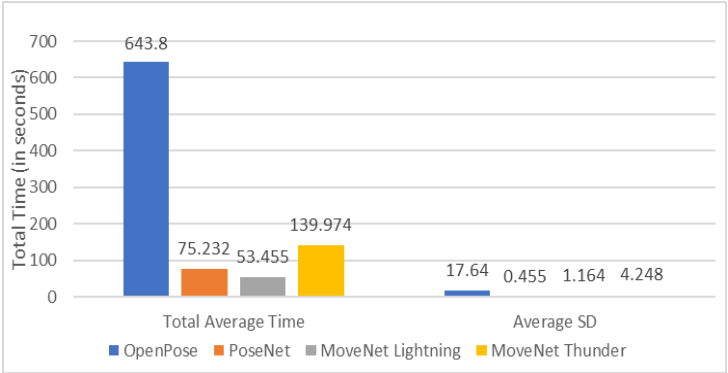

Figure 7 Average total time and Standard Deviation for all models [9]

From fig. (7), we can draw conclusions that total average time for OpenPose is much higher than other models. As per our system requirement, MoveNet Lightning would be the suitable model if we compare it with other models. PoseNet gives the better result with respect to standard Deviation means it is more consistent but it cannot be used for multi-person detection [9].

B. Accuracy of models on Group 1 and Group 2 images

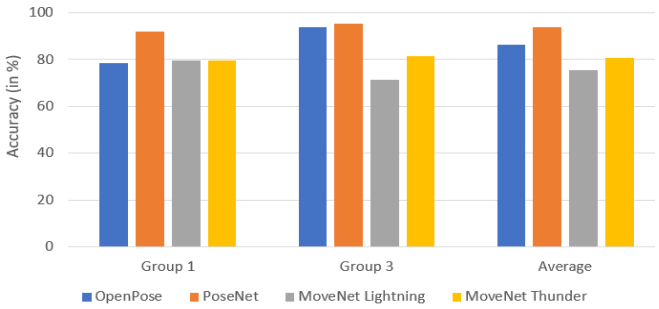

Figure .8 Pose Estimation Accuracies of models and their average [9]

They compared the models for Pose Estimation, PoseNet is best among four in Pose Estimation but it is single person. MoveNet thunder has good accuracy than MoveNet Lightning but it slower while OpenPose is also better with accuracy but Estimate Pose in very slower rate. MoveNet even with slightly lesser accuracy suites for the Fall Detection [9].

VI. CONCLUSION

Our Project demonstrates the best suitable model among OpenPose, PoseNet, MoveNet Lightning, MoveNet Thunder and YOLOv7 for Fall Detection. We come to conclusion that MoveNet Lightning is best among all other model to estimate the fall. Apart from that, we also created the system where we used this model to estimate the pose and detect the fall using different fall detecting algorithm. MoveNet Lightning has quick response and multi-person detection with moderate accuracy.

The Fall detection system has many applications such as monitoring system in health sectors which ultimately will makes elderly peoples and caregivers life easy. By studying, analysing and optimizing this can be improved by enhancing its efficiency and applicability.

Conflicts of interest

The authors declare no conflicts of interest

REFERENCES

[1] Jarraya, Salma Kammoun, Computer Vision-Based Fall Detection Methods Using the Kinect Camera: A Survey (December 10, 2018). International Journal of Computer Science & Information Technology (IJCSIT) Vol 10, No 5, October 2018.

[2] Nikita Vidua | Prof. Avinash Sharma **"Analysis of Fall Detection Systems: A Review**" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-4 | Issue-1, December 2019, pp.103-108

[3] Prof. Deepali K. Shende | Mr. Sidheshwar Madrewar | Mr. Shivam Bhongade | Mr. Shivam Dugade **"Dementia Patient Activity Monitoring and Fall Detection using IoT for Elderly**" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-4, June 2019, pp.363-367

[4] Ce Zheng, Wenhan Wu, Chen Chen, Taojiannan Yang, Sijie Zhu, Ju Shen, Nasser Kehtarnavaz, and Mubarak Shah. 2018. Deep Learning-Based Human Pose Estimation: A Survey. J. ACM 37, 4, Article 111 (August 2018), 37 pages

[5] Khan, Shehroz & Hoey, Jesse. (2017). Review of Fall Detection Techniques: A Data Availability Perspective. Medical Engineering & Physics. 39. 10.1016/j.medengphy.2016.10.014.

[6] Shu, F., Shu, J. An eight-camera fall detection system using human fall pattern recognition via machine learning by a low-cost android box. Sci Rep 11, 2471 (2021). https://doi.org/10.1038/s41598-021-81115-9

[7] Rastogi, Shikha & Singh, Jaspreet. (2023). A Comprehensive Study of Pose Estimation in Human Fall Detection. 10.1007/978-981-19-1142-2_31.

[8] Iazzi, Abderrazak & Rziza, Mohammed & Rachid, Oulad haj thami. (2018). Fall detection based on posture analysis and support vector machine. 1-6. 10.1109/ATSIP.2018.8364462.

[9] B. Jo and S. Kim, “Comparative Analysis of OpenPose, PoseNet, and MoveNet Models for Pose Estimation in Mobile Devices,” Feb. 2022, doi: 10.18280/ts.390111.

[10] Z. Zhang, C. Conly, and V. Athitsos, ‘‘A survey on vision-based fall detection,’’ in Proc. 8th ACM Int. Conf. PErvasive Technol. Related Assistive Environ., New York, NY, USA, Jul. 2015, p. 46

[11] Ren, Lingmei & Peng, Yanjun. (2019). Research of Fall Detection and Fall Prevention Technologies: A Systematic Review. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2922708.

[12] Romaissa, Beddiar & Oussalah, Mourad & Nini, Brahim. (2021). Fall detection using body geometry and human pose estimation in video sequences. Journal of Visual Communication and Image Representation. 82. 103407. 10.1016/j.jvcir.2021.103407.

[13] Toshpulatov, M., Lee, W., Lee, S. et al. Human pose, hand and mesh estimation using deep learning: a survey. J Supercomput 78, 7616–7654 (2022). https://doi.org/10.1007/s11227-021-04184-7

[14] Patil, R. A. ., and P. D. . Patil. “Skewed Evolving Data Streams Classification With Actionable Knowledge Extraction Using Data Approximation and Adaptive Classification Framework”. International Journal on Recent and Innovation Trends in Computing and Communication, vol. 11, no. 1, Feb. 2023, pp. 38-52, doi:10.17762/ijritcc.v11i1.5985.

[15] Patil, R.A., Patil, P.D. Efficient approximation and privacy preservation algorithms for real time online evolving data streams. World Wide Web 27, 5 (2024). https://doi.org/10.1007/s11280-024-01244-9

[16] Ahire, P.G., Patil, P.D. CAC-WOA: context aware clustering with whale optimization algorithm for knowledge discovery from multidimensional space in electricity application. Cluster Computing 27, 499–513 (2024). https://doi.org/10.1007/s10586-023-03965-4

[17] Ahire, P. G., & Patil, P. D. (2023). Context-Aware Clustering and the Optimized Whale Optimization Algorithm: An Effective Predictive Model for the Smart Grid. International Journal on Recent and Innovation Trends in Computing and Communication, 11(1), 62–76. https://doi.org/10.17762/ijritcc.v11i1.5987