Personal Assistance Sytem in Healthcare: Elderly Fall Detection

Aleena Roy
Department of AI & DS
SJCET Palai
Kottayam, India
aleenadeshaami@gmail.com

Namitha Pramod
Department of AI & DS
SJCET Palai
Kottayam, India
namithajispra@gmail.com

Saira Susan John
Department of AI & DS
SJCET Palai
Kottayam, India
sairasusanjohn@gmail.com

Sandra Ann Thomas
Department of AI & DS
SJCET Palai
Kottayam, India
sandraannthomas001@gmail.com

Dheeraj N
Department of AI & DS
SJCET Palai
Kottayam, India
dhraj.n@gmail.com

***Abstract*—Developments in computer vision have sparked creative approaches to eldercare, especially in the area where fall is detected, which is very essential in guaranteeing security and welfare of the elderly. For improving effectiveness of the fall detecting systems in actual surveillance setups devoted to senior care, this study investigates the mutually beneficial interaction between Human Pose Estimation (HPE) approaches and Synthetic Data Generation (SDG). Convolutional neural networks (CNNs) and key point localization are two techniques used by HPE algorithms to accurately interpret visual input and determine the configurations and movements of the human body. In addition to HPE, SDG approaches overcome the drawbacks of inadequately annotated datasets by simulating realistic but artificially manufactured data. Accuracy, adaptability, and robustness of fall detection systems could be enhanced by integrating HPE and SDG. The study explores techniques like OpenPose and GAN-based data augmentation and talks about how they are used in medical environments. To determine whether humans were present in the environment, a real-time detection mechanism was put into place in addition to the model evaluation. In order to detect whether human subjects are present or not, this system continuously analyzes live data streams. By enabling quick reaction to possible fall events, this real-time detection feature improves the whole fall detection system's adaptability and usefulness. By adding real-time detection, the system's usefulness and responsiveness are further increased, making it an invaluable instrument in healthcare environments where the goal is to ensure the safety and health of the elderly citizens.**

***Keywords*—Human Pose Estimation (HPE), Synthetic Data Generation (SDG) , Convolutional Neural Networks and Generative Adversarial Networks.**

I. Introduction

The study highlights the growing need for customised healthcare solutions in the context of an ageing populace, especially when it comes to reducing the likelihood of falls among senior citizens. As a result, the project aims to develop a personalised assistance system in healthcare that is especially designed to identify falls in older persons.

The Personal Assistance System uses cutting-edge machine learning algorithms to provide proactive monitoring and prompt intervention in the event of a fall. The project compares and assesses the effectiveness of several machine learning models in correctly classifying falls using datasets that include both fall and non-fall instances. Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), K-Nearest Neighbours (KNN), Support Vector Machines (SVM), and Logistic Regression are the various models used for comparing the accuracy.

The CNN model is the most promising, according to the findings, as it has better accuracy in differentiating between non-fall actions. Further improvements and adjustments strengthen the CNN model's performance and adaptability in practical situations.

The effort intends to give senior citizens an extra layer of security and comfort by smoothly incorporating the Personal Assistance System into their everyday activities, enabling them to preserve their independence and general well-being. This programme not only meets the urgent healthcare requirements of the elderly but also establishes the foundation for a time when everyone will have access to proactive care and individualised support. Fall detection systems for healthcare have advanced significantly with the addition of real-time detection technology, especially in terms of meeting the requirements of the elderly. Through continuous analysis of real-time data streams, this system can quickly and reliably identify whether humans are present in the surrounding area. This real-time feature not only makes fall detection more successful overall, but it also makes it possible to respond quickly to possible falls, which lowers the chance of harm and improves patient outcomes. In this introduction, we explore the application and effects of real-time detection technology in healthcare environments, highlighting its critical function in ensuring the safety of the elderly.

II. Literature Survey

The literature review in the field of fall detection and healthcare technology have yielded essential information about the creation and implementation of cutting-edge systems designed to improve the security and welfare of senior citizens. After conducting thorough examinations of several machine learning models and techniques, scientists have shown that Convolutional Neural Networks (CNNs) are especially good at correctly identifying falls in practical situations. Furthermore, adding real-time detection capabilities has been shown to be a viable way to improve the usefulness and responsiveness of fall detection systems, enabling prompt interventions and reducing any fall-related dangers. These results highlight how crucial it is to use state-of-the-art technology to meet the urgent healthcare demands of the aging population, which will eventually improve security, welfare, and general quality of life.

The LCR-Net++ approach was introduced by Rogez, Weinzaepfel, and Schmid as a way to recognize multi-person poses in natural photographs. This technique simultaneously recognizes 2D and 3D positions, offering important new perspectives on how people move. The IEEE Transactions on Pattern Analysis and Machine Intelligence have released their paper, which advances computer vision, particularly in the field of pose estimation. [1].Yang, Wang, Dantcheva, Garattoni, Francesca, and Bremond introduced UNIK, a unified framework for skeleton-based action recognition in real-world scenarios. Their paradigm, which was published as an arXiv preprint, tackles the difficulties in interpreting human skeletal movements, which are especially important for comprehending intricate actions like falls. This work has significance for fall detection applications and advances action recognition systems [2]. Iqbal, Milan, and Gall presented PoseTrack, a method for collaborative multi-person position estimation and tracking, at the IEEE Conference on Computer Vision and Pattern Recognition. Their method makes it possible to continuously monitor human positions throughout time, which is essential for identifying falls and examining movement patterns that precede fall events. Our knowledge of human position estimation and its uses in fall detection systems is improved by this research [3]. In their arXiv preprint, Redmon and Farhadi presented YOLOv3, an object detection method renowned for its speed and accuracy. The ideas behind YOLOv3 can be applied to fall detection systems in order to effectively identify people in photos or videos, even though they are not directly related to pose or fall detection. This work advances the field of object detection, which has ramifications for fall detection among other applications [4]. Serpa, Nogueira, Neto, and Rodrigues evaluated posture estimation as a possible fall detection solution at the IEEE International Conference on Serious Games and Applications for Health. Their study probably sheds light on approaches and conclusions related to fall detection systems that use pose estimation techniques. This research helps to increase the precision and dependability of fall detection systems by assessing the usefulness of pose estimation for fall detection [5]. Collectively, these advancements contribute to the development of more precise and reliable pose and fall detection systems, ultimately enhancing patient safety and well-being in healthcare settings.

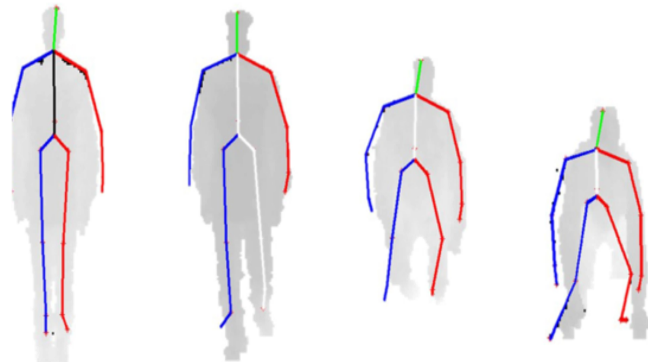

Fig. 1: Keypoint Identification

III. Methodology

The research aims to improve fall detection systems' accuracy and effectiveness by utilising deep learning techniques, specifically Transformers and Long Short Term Memory (LSTM). The dataset was annotated into discrete classes using the YOLOv5 framework, highlighting the dataset's complexity to guarantee the best possible model performance. Five distinct models were used during training, with the CNN model exhibiting the highest accuracy of 96 percentage. Refer to Fig.1 for various keypoint identification. Throughout the training phase, a focus was on making sure the model learned and generalised well using the fall detection data that was supplied. YOLOv5 was used to perform real-time fall detection, which allowed the system to recognise a wide variety of human actions, including walking, standing, sitting, lying down, and other common movements. When a fall is detected, the user-friendly interface delivers timely and informative alerts, enabling fast response and intervention as needed.

A. *Dataset and Data Preparation*

A JPEG-formatted dataset was assembled for this study from many web sources, guaranteeing a wide and thorough assortment of photos. Every image in the dataset was painstakingly annotated using the Roboflow platform to aid with model training. In order to successfully train the model, this annotating procedure require manually annotating the important elements within each photos, with a focus on capturing every complex human poses. The annotation process was led by many factors, which ensured a comprehensive analysis of human postures from all angles. The focus was on getting a variety of poses and motions that were pertinent to the goals of the research. This includes bending, stretching, sitting, standing, and other positions that are frequently seen in everyday life.

B. Model Architecture

The same annotated datasets were used to train five different models in a rigorous evaluation process. Interestingly, the CNN model was adopted for real-time detection jobs since it was the most accurate model. The YOLOv5 model was selected for fall detection in particular because of its effectiveness and simplified complexity, which makes use of several Convolutional Neural Network layers for reliable performance. YOLOv5's architecture stands out for being simple and effective, with a focus on real-time object identification tasks. Deep convolutional neural networks (CNNs) combined Cross-Stage Partial Networks (CSPNet) architecture provide the basis of YOLOv5. The powerful feature extractor function of this backbone network allows it to extract fine information from input photos at different spatial resolutions. Additionally, YOLOv5 has a Feature Pyramid Network (FPN), which facilitates the extraction of multi-scale features required for the identification of objects with varying sizes and degrees of complexity.

C. Model Training

Before starting the training procedure, the dataset was meticulously divided into training and testing partitions. Various classes were carefully created to capture a broad range of situations. The Convolutional Neural Networks model, Recurrent Neural Network model, K-Nearest Neighbour model, Support Vector Machines model, and Logistic Regression model. Annotated datasets are used to train the models for particular tasks, such as fall detection and real-time detection, by giving the models labeled samples to work with. In this instance, five distinct models using a range of architectures and we trained the methodologies using the same annotated datasets for all. These models were put through a rigorous evaluation procedure to see how well they performed in terms of speed, accuracy, and efficiency. The Convolutional Neural Network model was the most accurate of the models that were examined; it showed greater skills to discover and recognize the patterns in the data. Fig 2 refers to the fall detection system's sequential workflow is depicted in the block diagram. To improve the quality of the data, datasets are first vetted and then preprocessed using techniques like blur removal and annotation. Then, relevant parameters are extracted from the photos using feature extraction algorithms, which makes it easier to find significant patterns. The Convolutional Neural Network (CNN) algorithm, which is the fundamental part for evaluating and categorising fall and non-fall occurrences, is then given the processed data. Based on the retrieved features, the CNN algorithm uses its deep learning skills to distinguish between fall and non-fall events with accuracy. All things considered, this methodical technique makes it possible to accurately and efficiently identify falls in the photographs that are supplied, which advances safety protocols in the medical field and other connected fields.

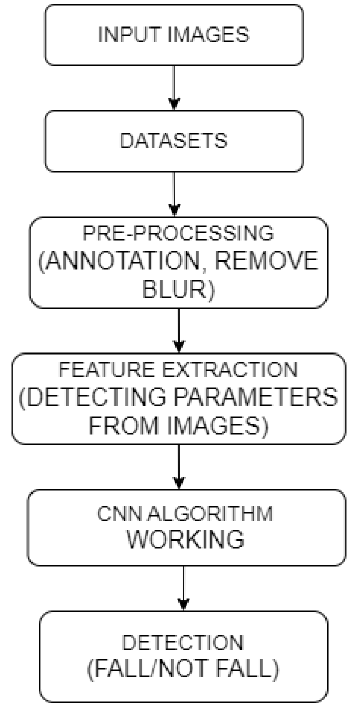

Fig. 2: Block Diagram Depicting the Process

D. Object Detection

Elements of the images were carefully retrieved and labelled, paying special attention to the study of human posture. Techniques for person cropping were used to guarantee accurate placement analysis. To accurately classify various human postures, labels were assigned, including standing, about to fall, sitting, and asleep.The picture components were meticulously extracted and labeled, with particular focus on the analysis of human posture.The person-cropping technique was applied to ensure precise placement evaluation.People are isolated from a visual backdrop,by person cropping techniques, which enable accurate placement analysis. This crucial stage lays the foundation for precise posture-categorization by narrowing the emphasis to just human participants. After cropping, every image is carefully tagged with descriptive captions that depict a variety of human positions. Every posture is classified to offer thorough annotations for the model training, ranging from the sitting and standing positions to more complex movements like bending or stretching.

E. Output and Utilisation

Performance measures of all five models, accuracy measurements, were painstakingly collected and documented. The detection process was carried out with accuracy and precision, guaranteeing consistent and dependable results for every model. Users were immediately notified of the events which are identified through the user interface, allowing for any rapid actions and intervention when necessary. To enable smooth detection of features and user engagement, an easy interface was created which offers a user-friendly way to monitor and react to detected events in real-time. Users may efficiently exploit the system's detection capabilities and take prompt action based on observed events by using this interface, which is a vital tool.

IV. Implementation and Result

In the initial phase of execution, the research team procured datasets including illustrations of both falls and non-fall incidents from reputable sources like Roboflow and Kaggle.
The data was then extensively preprocessed to create distinct autumn and non-fall groups. Techniques such as one-hot encoding were employed to ensure the optimal performance of the machine learning algorithms. The generated datasets were then utilized for training and evaluation of many models, such as Support Vector Machines, Convolutional Neural Networks , Recurrent Neural Networks, K-Nearest Neighbors, and Logistic Regression. Following extensive testing and analysis, it was found that the CNN model was the most successful in accurately recognizing fall occurrences. This underscores the necessity for further optimization and refining work.

During the implementation stage, reliable sources such as Roboflow and Kaggle provided datasets with both fall and non-fall instances. The data was then heavily preprocessed using methods like one-hot encoding to maximise the effectiveness of machine learning algorithms in order to separate them into separate fall and non-fall groups. The produced datasets were used for training and evaluating a variety of models, such as Support Vector Machines (SVM), Recurrent Neural Networks (RNN), Convolutional Neural Networks (CNN), K-Nearest Neighbours (KNN), and Logistic Regression. After extensive testing and analysis, the CNN model showed an amazing accuracy score of 97.03%, making it the most successful in correctly recognising fall incidents.Furthermore, the system was able to incorporate real-time detection capabilities with ease, which significantly improved the system's practicality and responsiveness by enabling the early recognition of human presence in the surrounding environment.

A. *Performance Evaluation*

Table I: Accuracy Of Various Models

	Score	Model
0	0.970298	CNN
1	0.950009	RNN
2	0.690000	Logistic Regression
3	0.670000	Support Vector Machine
4	0.660000	K Nearest NeighhbourClassifiier

Table 1 depicts an outstanding accuracy score of 97.03%, the Convolutional Neural Network (CNN) was determined to be the best performer in the project's performance evaluation. The Recurrent Neural Network (RNN), with an accuracy of 95.00%, trailed closely behind. In the meantime, Logistic Regression showed a moderate accuracy of 69.00%, while K-Nearest Neighbours (KNN) and Support Vector Machine (SVM) showed lower accuracies of 66.00% and 67.00%, respectively.

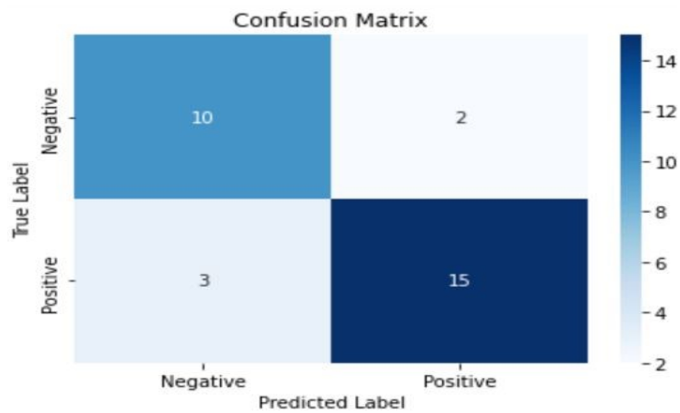

Fig. 3: Confusion Matrix Of CNN Model

These findings highlight how well deep learning models—in particular, CNN and RNN—perform when it comes to correctly recognising falls in the elderly when compared to traditional machine learning algorithms. The remarkable precision of the CNN model indicates how well it performs in actual fall detection situations and suggests its possible application in senior-focused healthcare systems. During the results phase, a thorough examination of several machine learning models showed notable differences in their performance indicators. Notably, when it came to reliably and properly differentiating between fall and non-fall behaviours, the CNN model outperformed other algorithms. This demonstrates how convolutional neural networks function well in fall detection applications. Subsequent enhancements reinforced the CNN model's effectiveness and resilience, confirming its appropriateness for incorporation into personal assistance systems for the elderly. These conclusions are supported by the CNN confusion matrix shown in Figure 3, which also shows the accuracy breakdown of the various models, with CNN obtaining the best accuracy. All things considered, these findings highlight the encouraging possibility of using state-of-the-art technology to meet the pressing healthcare requirements of the senior citizenry, providing real solutions to improve their safety and well-being.

V. Conclusion

In conclusion, with the most recent advancements in pose and fall detection documented in the literature review, the utilization of computer vision and machine learning for healthcare applications has improved dramatically. Promising techniques for accurately tracking human motions and

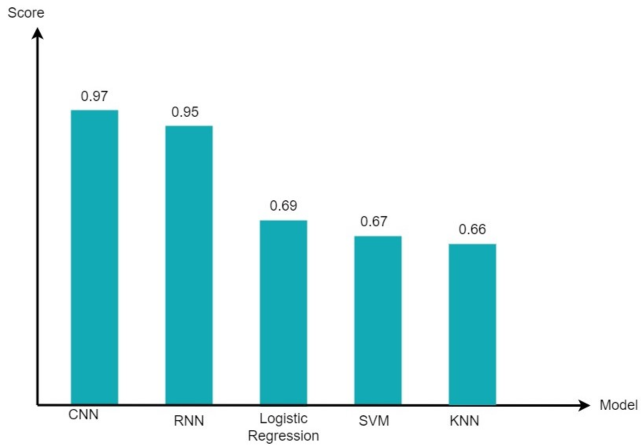

Fig. 4: Bar Graph Depicting Accuracy Of Models

detecting human poses include PoseTrack, UNIK, and LCR-Net++. This information is useful for understanding complex activities such as falls. Researchers can develop more dependable systems that can recognize individuals and follow their motions in real-world scenarios by combining these techniques with efficient object detection algorithms like YOLOv3.

The extensive examination of the literature offers priceless insights into the most recent developments in pose and fall detection technologies. The description of YOLOv3, research works like LCR-Net++, UNIK, PoseTrack, and reviews of position estimate algorithms all make a substantial contribution to the development of reliable fall detection systems, especially for older populations. Convolutional Neural Networks (CNNs) show promise as a viable option for precise fall detection, highlighting the need of utilising deep learning methods in medical applications. Additionally, the incorporation of real-time detecting capabilities improves fall detection systems' practicality and reactivity, which in turn improves patient outcomes and also the quality of life.

In the future, fall detection systems could pursue a number of areas for investigation and improvement. Fall detection may be more accurate and dependable if multimodal sensor data—from depth sensors or wearables—is integrated. Additionally, fall prevention measures may be more effective if customised health recommendations based on personal data and activity patterns are created. Enhancing real-time detection algorithms and incorporating them into wearable or Internet of Things sensors may also make it possible to monitor continuously and take timely action in real-world situations. Fall detection technology could be revolutionised by continued study and innovation in this area, improving senior people's security and well-being and improving healthcare outcomes. Ultimately, these advancements create new opportunities for the development of customized support systems in the healthcare sector, particularly for senior care, where early fall detection can significantly improve patient outcomes and quality of life. Further research and innovation in computer vision and machine learning for healthcare applications is necessary to improve patient safety and well-being in healthcare settings, hence collaboration and exploration in these fields are essential.

References

[1] Rogez, G., Weinzaepfel, P. and Schmid, C., 2019. Lcr-net++: Multi-person 2d and 3d pose detection in natural images. IEEE transactions on pattern analysis and machine intelligence, 42(5), pp.1146-1161.

[2] Yang, D., Wang, Y., Dantcheva, A., Garattoni, L., Francesca, G. and Brémond, F., 2021. Unik: A unified framework for real-world skeleton-based action recognition. arXiv preprint arXiv:2107.08580.

[3] Iqbal, U., Milan, A. and Gall, J., 2017. Posetrack: Joint multi-person pose estimation and tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2011-2020). 2011–2020.

[4] Redmon, J. and Farhadi, A., 2018. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

[5] Serpa, Y.R., Nogueira, M.B., Neto, P.P.M. and Rodrigues, M.A.F., 2020, August. Evaluating pose estimation as a solution to the fall detection problem. In 2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH) (pp. 1-7). IEEE.

[6] Zherdev, D., Zherdeva, L., Agapov, S., Sapozhnikov, A., Nikonorov, A. and Chaplygin, S., 2021. Producing synthetic dataset for human fall detection in ar/vr environments. Applied Sciences, 11(24), p.11938.

[7] Hwang, H., Jang, C., Park, G., Cho, J. and Kim, I.J., 2021. Eldersim: A synthetic data generation platform for human action recognition in eldercare applications. IEEE Access, 11, pp.9279-9294.

[8] Papandreou, G., Zhu, T., Kanazawa, N., Toshev, A., Tompson, J., Bregler, C. and Murphy, K., 2017. Towards accurate multi-person pose estimation in the wild. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4903-4911).

[9] Khodabandeh, M., Reza Vaezi Joze, H., Zharkov, I. and Pradeep, V., 2018. DIY human action dataset generation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 1448-1458).

[10] Mubashir, M., Shao, L. and Seed, L., 2013. A survey on fall detection: Principles and approaches. Neurocomputing, 100, pp.144-152.

[11] Dai, J., Bai, X., Yang, Z., Shen, Z. and Xuan, D., 2010, March. PerFallD: A pervasive fall detection system using mobile phones. In 2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops) (pp. 292-297). IEEE.

[12] Montesinos, L., Castaldo, R. and Pecchia, L., 2018. Wearable inertial sensors for fall risk assessment and prediction in older adults: A systematic review and meta-analysis. IEEE transactions on neural systems and rehabilitation engineering, 26(3), pp.573-582.

[13] Ludl, D., Gulde, T. and Curio, C., 2020. Enhancing data-driven algorithms for human pose estimation and action recognition through simulation. IEEE transactions on intelligent transportation systems, 21(9), pp.3990-3999.

[14] Briana Moreland, M.P.H., Ramakrishna Kakara, M.P.H. and Ankita Henry, M.P.H., 2018. Trends in Nonfatal Falls and Fall-Related Injuries Among Adults Aged 65 Years—United States, 2012–2018. Increases in Health-Related Workplace Absenteeism Among Workers in Essential Critical Infrastructure Occupations During the COVID-19 Pandemic—United States, March–April 2020, 2001, p.875.

[15] Mohammed Sharook, K., Premkumar, A., Aishwaryaa, R., Amrutha, J.M. and Deepthi, L.R., 2022. Fall Detection Using Transformer Model. In ICT Infrastructure and Computing: Proceedings of ICT4SD 2022 (pp. 29-37). Singapore: Springer Nature Singapore.

[16] Noury, N., Fleury, A., Rumeau, P., Bourke, A.K., Laighin, G.O., Rialle, V. and Lundy, J.E., 2007, August. Fall detection-principles and methods. In 2007 29th annual international conference of the IEEE engineering in medicine and biology society (pp. 1663-1666). IEEE.

[17] Xu, T., Zhou, Y. and Zhu, J., 2018. New advances and challenges of fall detection systems: A survey. Applied Sciences, 8(3), p.418.

[18] Abbate, S., Avvenuti, M., Bonatesta, F., Cola, G., Corsini, P. and Vecchio, A., 2012. A smartphone-based fall detection system. Pervasive and Mobile Computing, 8(6), pp.883-899.

[19] Özdemir, A.T., 2016. An analysis on sensor locations of the human body for wearable fall detection devices: Principles and practice. Sensors, 16(8), p.1161.

[20] Martínez-Villaseñor, L., Ponce, H., Brieva, J., Moya-Albor, E., Núñez-Martínez, J. and Peñafort-Asturiano, C., 2019. UP-fall detection dataset: A multimodal approach. Sensors, 19(9), p.1988.

[21] Planinc, R. and Kampel, M., 2013. Introducing the use of depth data for fall detection. Personal and ubiquitous computing, 17, pp.1063-1072.

[22] Harrou, F., Zerrouki, N., Sun, Y. and Houacine, A., 2017. Vision-based fall detection system for improving safety of elderly people. IEEE Instrumentation Measurement Magazine, 20(6), pp.49-55.

[23] Yacchirema, D., De Puga, J.S., Palau, C. and Esteve, M., 2018. Fall detection system for elderly people using IoT and big data. Procedia computer science, 130, pp.603-610.

[24] Thome, N., Miguet, S. and Ambellouis, S., 2008. A real-time, multi-view fall detection system: A LHMM-based approach. IEEE transactions on circuits and systems for video technology, 18(11), pp.1522-1532.

[25] Mastorakis, G. and Makris, D., 2014. Fall detection system using Kinect's infrared sensor. Journal of Real-Time Image Processing, 9, pp.635-646.