Vol 6 No 2 (2024) E-ISSN: 2682-860X

International Journal on Robotics, Automation and Sciences

Human Fall Motion Prediction – A Review

Raphon Galuh Candraningtyas, Andi Prademon Yunus* and Yit Hong Choo

***Abstract* – In predicting human fall motion, focused on enhancing safety and quality of life for the elderly and individuals at risk of falls. By highlighting the critical role of Human Pose Estimation, advancements in human motion forecasting, and fall prediction. It explores the continuous efforts to improve fall detection systems using innovative technologies, such as wearable sensors and IoT devices to implement deep learning models and analyze human poses and gestures. Various methods show promise in accurately predicting human fall motion by capturing complex patterns and relationships in the data. For instance, self-attention mechanisms can revolutionize human motion prediction by effectively capturing these intricate patterns, leading to more accurate predictions. Future research directions should focus on enhancing model accuracy, exploring new techniques for capturing complex patterns, and enabling real-time implementation in wearable devices or smart environments. By addressing these areas, fall detection systems can be significantly improved, benefiting individuals and healthcare systems worldwide.**

Keywords—Human Fall Motion Prediction, Human Motion, Fall Detection System, IOT Devices, Wearable Sensors.

I. Introduction

Falling is an event that happens under some circumstances unintentionally in our daily life which can cause severe injuries or even death. Globally, it is estimated that as many as 646,000 people die each year due to accidental falls, of which more than 80% occur in low- and middle-income countries [1]. The risk of falling is also experienced by the elderly, where falls in the elderly can cause morbidity, mortality, reduced function, and the potential need for early admission into nursing homes. This is particularly concerning given that in some studies, the proportion of patients who fell during hospitalization was as high as 12.2% with a total of 69 falls. The study evaluated the Morse Fall Scale (MFS) as a tool to assess the risk of falls in hospitalized patients, particularly in a Swiss hospital setting. The MFS was tested across various patient populations, and its effectiveness was measured using different cut-off scores. The study found that a cut-off score of 55 points was optimal, providing a balance between sensitivity and specificity, with an accuracy of 66.8%. However, the study also highlighted limitations, such as high false positive rates, which suggests that while the MFS is a useful tool, it may require further validation in different clinical settings. The paragraph concludes by emphasizing the importance of a real-time prevention system to accurately assess and prevent falls through continuous monitoring of patient behavior and motions, enabling early intervention and potentially reducing the incidence of falls [2].

The prevention system can prevent falling and reduce unwanted accidents. By realizing the human falling prediction system to predict the falling human motion, we could send mechanical tools such as airbags to prevent severe collisions before the event and reduce the death rate caused by falling.

*Corresponding Author, email: andidemon@ittelkom-pwt.ac.id ORCID: 0000-0002-1185-5763

Raphon Galuh Candraningtyas is with Faculty of Informatics, Institut Teknologi Telkom Purwokerto, Banyumas, Indonesia (e-mail: 21110034@ittelkom-pwt.ac.id).

Andi Prademon Yunus is with Faculty of Informatics, Institut Teknologi Telkom Purwokerto, Banyumas, Indonesia (e-mail: andidemon@ittelkom-pwt.ac.id).

Yit Hong Choo is with Institute for Intelligent System Research and Innovation, Deakin University, Australia (e-mail: y.choo@deakin.edu.au).

International Journal on Robotics, Automation and Sciences (2024) 6, 2:52-58
https://doi.org/10.33093/ijoras.2024.6.2.8
Manuscript received: 15 Apr 2024 | Revised: 20 June 2024 | Accepted: 12 July 2024 | Published: : 30 Sep 2024
Published by MMU PRESS. URL: http://journals.mmupress.com/ijoras

Several techniques to detect human falling down event such as utilizing accelerometer sensor and gyroscope to K-Nearest Neighbors algorithm and Decision Tree for deciding the states [3]. However, this sensor-embedded system is impractical since it must be attached to the human body. Other studies used research on an image-based fall detection system where they developed image-based fall detection to detect falls while walking, forward, sitting, and standing. The research used YOLOv3 for data preprocessing, posture detection, and fall detection with the Image-based Fall Detection System (IFADS) [4]. The system uses the surveillance camera's input and recognizes humans by pose using pose estimation techniques. Then deep learning method is employed to decide the states. These techniques highly depend on the accuracy of human pose estimation which leads the system to be vulnerable. Although, Human Pose Estimation (HPE) has been extensively studied in computer vision literature, involves estimating the configuration of human body parts from input data captured by sensors, specifically images and videos [5]. HPE methods trained on existing standard datasets may not generalize well across different domains. The recent trend to alleviate the domain gap is utilizing GAN-based learning approaches. Nevertheless, how to effectively transfer the human pose knowledge to bridge domain gaps remains unaddressed [5]. However, these techniques have some advantages in utility, practicality, and future development utilization.

II. LITERATURE REVIEW

A. *Human Pose Estimation (HPE)*

Methods for identifying the 2D or 3D motion of the human body in frames are gaining popularity due to their necessity in completing tracking automation. Most of the methods created for human motion analysis utilize models, such as volumetric, planar, or kinematic models, to identify different human body parts in provided images [6].

You Only Look Once (YOLO) is a cutting-edge object detection software that has been popular in recent years and is known for its efficiency and accuracy in detecting objects in images or videos [7].YOLO's approach of processing the entire image in a single pass through a convolutional neural network allows for real-time object detection. One of the main advantages of YOLO is its user-friendly interface, making it accessible to developers and researchers with different levels of expertise. The software's high object detection precision has made it the tool of choice for tasks such as surveillance, autonomous vehicles, and image analysis [4].

On the other hand, OpenPose as a human pose estimation method offers an open-source system used for real-time human pose detection developed by Carnegie Mellon University (CMU) based on object detection precision has made it the tool of choice for tasks such as surveillance, autonomous vehicles, and image [8].Now future development utilization based on human pose estimation could lead to predicting the long-term human body motion ahead consecutively. A study of human motion forecasting obtained a good prediction with 64.7 mm by mean per joint position error [9]. Thus this technique opens the possibility of falling down human motion forecasting. Some techniques and methods are applicable to that are described in Section 2.

As we can see the YOLO-Pose approach combines object detection and pose estimation forecasting as demonstrated in a study where they used the YOLO-Pose approach, which integrates the YOLOv5 framework for object detection with a novel heatmap-free method for 2D multi-person pose estimation [10]. Similarly, a study utilized the YOLO-Pose approach for object detection [11]. The difference is that the study used YOLOv3 for human body detection. Additionally, research on real-time detection of safety harnesses used YOLO-Pose for its object detection, whereas it used the YOLOv5 algorithm to enhance object detection by adding focus structure and adaptive image scaling, and OpenPose for its human posture estimation [12].

The YOLO-Pose approach efficiently merges object detection with precise human pose estimation. YOLO's ability to process images rapidly and accurately in real-time complements Open Pose's capability to track human body parts, including faces, limbs, and fingers, in a variety of scenarios. This combination represents a promising direction in computer vision, providing a versatile and robust framework for a wide range of applications that require both object detection and human pose estimation. Convolutional Neural Network (CNN) [13]. OpenPose is mostly used for skeleton sequence extraction. This model can realize the tracking of human faces, limbs, and even fingers. OpenPose contributed as one of the baseline pose estimation methods for detecting a single person and also for multiple people at once [14].

A research utilized HPE techniques to forecast human locomotion in an egocentric view by disentangling pedestrian dynamics into global and local components [15]. HPE can also be utilized to develop an end-to-end trainable approach for diverse and controllable human motion predictions, where the accuracy of human pose allows for the generation of realistic and diverse future poses [16]. Additionally, HPE can be integrated with self-localization within a 3D scene using wearable sensors, leading to the development of the Human Positioning System (HPS) [17]. This system estimates human motion prediction while localizing the person within a pre-scanned large 3D scene using wearable sensors.

Furthermore, HPE has been utilized to advance action recognition, and human pose estimation through benchmarking state-of-the-art methods on the IKEA ASM dataset. This enables accurate tracking and understanding of human movements during tasks like furniture assembly [18].

These studies demonstrate HPE's crucial role in various applications related to human pose estimation analysis. HPE has significantly advanced action recognition, and human motion forecasting tasks, as the accuracy and versatility of HPE make it a valuable tool for understanding and analyzing human movements in diverse scenarios. By providing precise and versatile tracking capabilities, HPE methods contribute to the development of more effective systems for predicting and preventing falls, thereby enhancing safety and quality of life.

B. Human Motion Prediction

Human motion prediction involves predicting future movements based on current or past actions by analyzing human poses and gestures captured in video data. This predictive capability is crucial in applications such as fall detection, where anticipating and detecting instances where a person is likely to fall based on their movements and posture can prevent injuries.

Human motion prediction enables biometric identification through the analysis of distinct gait patterns, providing a secure method for individual authentication. Additionally, the ability to predict human motion facilitates the generation of lifelike gestures, which can be valuable in virtual reality, animation, and human-robot interaction applications. Moreover, in the realm of healthcare, human motion prediction aids in designing personalized rehabilitation programs tailored to patients with movement disorders, ultimately enhancing their motor skills and overall quality of life [19]. Implementing motion prediction models in healthcare can enhance patient care, improve treatment outcomes, and optimize rehabilitation processes by leveraging technology to analyze and predict human movements effectively [20].

Several studies illustrate the application of human motion prediction. A study proposes an approach for human motion prediction that utilizes a Skeleton-joint Co-Attention mechanism to capture spatial coherence and temporal evolution, which identify unstable or unusual movements indicative of a falling state [21]. Another study utilizes convolution-based networks with the High-Resolution Spatio-Temporal Attention Network (HR-STAN) architecture and attention modules to improve long-term predictions, which is relevant for monitoring movements overextended [21]. Integrating self-attention mechanisms, such as the T-transformer module, to capture long-range dependencies in human motion sequences that had been done by, are effective in modeling temporal dependencies in sequential data, making them suitable for predicting based on historical motion patterns [22]. Furthermore, unsupervised learning techniques for human motion prediction also contribute to fall detection. A study has focused on unsupervised action representation learning, enabling models to learn motion representation from unlabeled data. This enhances the understanding of human motion for action recognition [23].

Human motion prediction is a pivotal technology that extends its benefits across various domains, with significant implications for fall detection. By analyzing human poses and gestures captured in video data, this potential can effectively anticipate and detect falls. The integration of advanced techniques such as CNNs, RNNs, and unsupervised learning enhances the ability to capture complex motion patterns, improving the accuracy and reliability of fall detection systems. As research and technological advancements continue, the applications of human motion prediction are expected to further enhance safety, improve healthcare outcomes, and provide much valuable insights for diverse applications, that ultimately will benefit many individuals.

C. Human Fall Motion Prediction

Human motion forecasting involves predicting the future movements of a person based on their current or past actions, achieved through analysis of human poses and gestures captured in video data.

This knowledge is valuable in various applications, one of them is fall detection. Predicting human fall motion involves anticipating and detecting instances where a person is likely to fall based on their movements and posture. Human Fall Motion Prediction focuses on preventing injuries by analyzing body poses and acceleration patterns in human motion data [24]. A study developed the Human Torso Motion Model (HTMM), where the model will compare the changing rates of torso angle and centroid height with specific thresholds to determine if a fall has occurred. This method was found to be highly accurate in discriminating falls when compared to other fall detection approaches [25].

In the healthcare fields predicting the fall motion can help by implementing proactive measures, to ensure timely assistance and reduce the risk of serious consequences. One study centers on developing a smart fall detection mechanism for healthcare by utilizing a fuzzy adaptive thresholds algorithm with a triaxial accelerometer in a smartwatch for fall detection and indoor positioning, which aims to improve the quality of life for the aging population while reducing labor costs and resource consumption [26]. Another study aimed to develop a cost-effective fall-detection system for the elderly, using a Pyroelectric Infrared (PIR) sensor to collect data on normal and falling events for detecting the elderly [27]. In a different approach, a study proposed developing a Fall Detection System based on the Internet of Things (IoT) using accelerometer and gyroscope sensors to classify various activities, including falling [3]. Another approach focused on a sensor-based fall detection system using deep learning models, utilizing wearable sensor data for daily activities and fall behavior [3]. Furthermore, developed a fall detection system that used multiple sensors on different body locations and integrated eXplainable Artificial Intelligence (XAI) for model interpretability [28].

Several methods have been explored for fall human detection, with various approaches to the subject. For instance, developed a fall detection method using dual-channel feature integration, defining falling-state and fallen-state perspectives, and implementing a dual-channel sliding window model for feature extraction [29]. Another study utilized a 360-degree camera to develop a fall detection system, addressing limitations of field of view in existing solutions [30]. The research by utilizing a human model for body posture recognition, and height changes to detect fall motion [4]. A study that employs the Motion History Image and C-Motion method to quantify motion, distinguishing between normal and unusual activities to develop a human fall detection system using video surveillance [31] Research of fall detection using skeleton trajectories and positive semidefinite matrices to analyze fall events done by employing the Dynamic Time Warping (DTW) algorithm which predicts fall motion based on similarity scores between sequences [32].

Advancements in human motion forecasting and fall motion prediction are vital for enhancing safety and quality of life, particularly for the elderly and those at risk of falls. The variety of approaches and techniques discussed in the research demonstrates the continuous effort to improve fall detection systems, utilizing innovative technologies such as wearable sensors, IoT devices, and deep learning models. As study in this field progresses, it is likely to lead to further improvements in fall detection systems, ultimately benefiting individuals and healthcare systems worldwide.

III. Prospective Method

A. Kalman Filter

Kalman Filter (KF), also known as Linear Quadratic Estimation, predicts the future state of a system based on previous data and the current state. KF is expressed by an equation, but it is separated into two steps including prediction and update. In the prediction step, estimates are obtained for the current state by utilizing a series of state estimates at previous intervals. These predicted estimates are prior knowledge related to the previous estimates and there are no observations for that system at the current state. In the update step, previous estimates are blended with current observations to present an estimate of the current and subsequent state settings of the system. These steps are usually repeated alternately, i.e., predictions are made until the next observation, and then updates are made using the current observation. If there are no observations in an interval, multiple prediction updates will be performed until the next observation. Similarly, if several independent observations are made periodically, several updates with different matrices will be obtained based on each observation [33].

One study on fall detection used KF to preprocess raw data from MEMS-based inertial sensors. This process helps reduce noise in tri-axial accelerations and angular velocities, thereby improving the accuracy of monitoring falls and daily activities. By enhancing the data quality, KF contributes significantly to the system's ability to distinguish between falls and normal activities with high precision [30] Another study utilized the Kalman Filter for human motion prediction, which it reduced noise and uncertainty in measurements, making it a valuable tool for enhancing the accuracy of human motion prediction algorithms [34].

The approach of KF for predicting the future state based on past data and current observations makes it a powerful tool for human fall detection. The iterative process of the KF enables continual refinement of its predictions, resulting in more accurate and reliable estimates of the system's current and future states. KF's precise and reliable state estimation capabilities enable the development of advanced fall detection systems that can significantly improve the safety and quality of life for individuals at risk of falls.

B. Recurrent Neural Network (RNN)

Recurrent Neural Network is a class of neural networks where the connections in the units create a shared structure with temporal order. RNNs have an internal memory to process a series of input data. The computational units in RNN have time-varying real-valued activations and adjustable weights. RNNs are created by recursively applying equal weights on a graph-like structure. The learned model in RNN has the right input size since it transitions from one state to another [34]. Long Short-Term Memory (LSTM) can be stored based on the RNN's internal memory that stores weights and computation data. However, RNNs cannot retain such a set of data in the long term for prediction. Here, LSTM acts as an extended form of RNN, which contains additional memory based on structure. Hochreiter and Schmidhuber invented the LSTM in 1997, which works and can handle signals mixed with low and high-frequency components [34].

The RNN algorithm is widely used for detecting human motion prediction, as demonstrated which utilized RNNs to process acceleration measurements and detect falls based on temporal patterns in the data [24]. A study that has been done, used an LSTM-based sequence model, which is a type of RNN algorithm, for fall detection, leveraging data from five different sensors [28]. Similarly, a study also used LSTM network architecture and Bidirectional LSTM to address the vanishing gradient issue in long-term dependency learning for their fall detection model [28]. Other studies also use the RNN algorithm, such as Bidirectional RNN and LSTM, to model skeleton data as a time sequence to capture temporal dependencies between consecutive frames. Furthermore, an attention-enhanced graph convolutional LSTM network was employed for skeleton-based action recognition from a study [23].

RNNs represent a significant advancement in neural network architecture, enabling the modeling of temporal dependencies in sequential data. RNNs, characterized by their internal memory and ability to process the input data in a series, have found extensive use in various applications, including human motion prediction and fall detection. The LSTM networks have also further enhanced RNN capabilities, allowing for the retention of important information over long time intervals. Studies have shown that LSTM-based models, including variants like Bidirectional LSTM and attention-enhanced graph convolutional LSTM networks, excel in capturing complex temporal patterns, making them valuable tools for tasks requiring sophisticated sequential data analysis. The RNNs and LSTM can model the temporal dependencies in these sequences of poses, predicting potential falls by analyzing patterns over time.

C. Self-Attention

The self-attention mechanism within the transformer model is a mechanism that connects different positions in a single sequence to compute the representation of that sequence. In computer vision, the attention mechanism models can be classified into soft attention and hard attention. Soft attention has been used in various computer vision fields such as classification, detection, segmentation, modeling, and video processing. Some categories of the soft attention mechanism include spatial attention, channel attention, mixed attention, and self-attention. Self-attention is a mechanism that allows each pixel in the feature map to be considered a random variable, and the pixel's predicted value can be enhanced or diminished based on its similarity to other pixels in the image.

Currently, no research uses self-attention mechanisms for fall motion prediction, but there are several research that use self-attention mechanisms for human motion prediction. One such study aimed to maintain both high and low-resolution features in human motion prediction to preserve motion details while allowing the network to focus on large-scale features simultaneously. They developed a method for 3D human motion prediction using the high-resolution spatiotemporal attention network (HR-STAN), which combines spatiotemporal convolution (STConv) and spatiotemporal attention to efficiently capture spatial-temporal dependencies, resulting in highly accurate short-term motion predictions [35].

Another study proposed an enhancement to human motion prediction based on fusion strategies, where current prediction frames can be combined with previous frames to minimize prediction errors effectively and improve prediction continuity. They developed the spatiotemporal transformer graph convolutional network (STTG-Net) to minimize prediction errors and enhance prediction sequence smoothness, thus reducing the problem of error accumulation in prediction [22]. Additionally, researchers aimed to address the limitations of conventional encoder models by focusing more on previous motion information that better reflects the current motion context. They developed an attention-based motion prediction method that dynamically adapts to previous motion with the current context. This approach utilized a Graph Convolutional Network to learn spatial and temporal dependencies in human motion data, a Motion Attention Model to find similar sub-sequences from the past, and a Discrete Cosine Transform-based representation to accurately extract motion patterns from historical data [20].

Lastly, another study employed an end-to-end approach for human motion prediction, allowing the system to predict various motions without the need for multiple mappings to achieve motion diversity. They developed a unified framework for human motion prediction to achieve diversity and controlled human motion predictions [16].

Self-attention mechanisms play a crucial role in connecting different positions in a sequence, enabling the computation of sequence representation. While self-attention has not been directly applied to human fall prediction, it has been extensively utilized in several human motion prediction research that demonstrated the effectiveness of self-attention in capturing spatial-temporal dependencies and enhancing prediction accuracy. These advancements in human motion prediction could potentially be applied to fall detection systems.

IV. CONCLUSION

Studies have been conducted to realize the system for securing people from unexpected accidents and severe injuries by performing the human pose estimation, predicting the falling and fallen state, predicting the human movement, as well as predicting the falling by using sensors. A system that predicts the falling state by attaching the sensor to the human body is one of the best options to realize. However, the real-life implementation is not cost-effective and impractical. Another option is to utilize the camera and computer vision for detecting the fallen state. However, the device can only detect the state without a prevention system before the collision happens. Thus, human motion prediction takes part as the main technique to be utilized in the prevention system. Combining human motion prediction and falling down detection on one model can advance the implementation of the safety surveillance system to prevent falling and severe accidents. Novel method is required to be developed to process the video input and expect the future motion with falling detection output which can be the combination of RNN-based method, Self-Attention-based method, or unsupervised learning-based method such as Kalman Filter.

Human fall motion prediction and the significant advancements in human motion analysis and prediction can contribute to the development of human motion prediction technology. As research in this field continues to evolve, it is expected will lead to more reliable and effective systems for preventing falls and enhancing the overall well-being of individuals.

The future directions in human fall Motion Prediction will aim to enhance the model accuracy by exploring new techniques for capturing complex patterns in data, enabling real-time implementation in wearable devices or smart environments. By integrating self-attention-based models with other technologies like IoT sensors and AI algorithms, it can create a more robust fall detection system. Advancements in HPE technologies will further improve its accuracy and versatility, while validation studies and clinical trials will evaluate system effectiveness in real-world settings. Developing user-friendly interfaces, addressing concerns, and exploring cross-domain applications will also be crucial. Additionally, long-term monitoring methods and education programs for healthcare professionals and caregivers will ensure effective implementation and utilization of these systems, ultimately improving safety and quality of life for individuals at risk of falls.

ACKNOWLEDGMENT

I would like to express my deepest gratitude to my university, Institut Teknologi Telkom Purwokerto, for their guidance and encouragement which have been invaluable for me to write this literature review. I would also like to acknowledge that there is no financial support from agencies funding this research work. My sincere appreciation goes to all the researchers and authors whose work and insights have contributed to the development of this literature review.

AUTHOR CONTRIBUTIONS

Raphon Galuh Candraningtyas: Writing – Original Draft Preparation;

Andi Prademon Yunus: Writing – Review & Editing;

Yit Hong Choo: Writing – Review & Editing;

CONFLICT OF INTERESTS

No conflict of interests were disclosed.

ETHICS STATEMENTS

Our publication ethics follow The Committee of Publication Ethics (COPE) guideline. https://publicationethics.org/

REFERENCES

[1] M.S. Reddy and V. Vankayalapati, "Need of global student safety and insurance day observance: a suggestion," *International Journal of Community Medicine and Public Health*,vol.7,no.4,pp.1587, 2020.
DOI: https://doi.org/10.18203/2394-6040.ijcmph20201434

[2] R. Schwendimann, S. De Geest, and K. Milisen, "Evaluation of the Morse Fall Scale in hospitalised patients," *Age Ageing*, vol. 35, no. 3, pp. 311–313, 2006.
DOI: https://doi.org/10.1093/ageing/afj066

[3] S.K. Bhoi *et al.*, "FallDS-IoT: A Fall Detection System for Elderly Healthcare Based on IoT Data Analytics," *Proceedings of the 2018 International Conference on Information Technology (ICIT)*, pp. 155–160, 2018.
DOI: https://doi.org/10.1109/ICIT.2018.00041

[4] K.L. Lu and E.T.H. Chu, "An image-based fall detection system for the elderly," *Applied Sciences*, vol. 8, no. 10, pp. 1995, 2018.
DOI: http://dx.doi.org/10.3390/app8101995

[5] C. Zheng, W. Wu, C. Chen, T. Yang, S. Zhu, J. Shen, N. Kehtarnavaz and M. Shah, "Deep Learning-based Human Pose Estimation: A Survey," *ACM Computing Surveys*, vol. 56, no. 1, Art. no. 11, pp. 1-37, 2024.
DOI: https://doi.org/10.1145/3603618

[6] L. Gril, P. Wedenig, C. Torkar and U. Kleb, "A Tensor-based Regression Approach for Human Motion Prediction," *Quality and Reliability Engineering International*, vol. 39, no. 2, pp. 481–499, 2023.
DOI: https://doi.org/10.1002/qre.3153

[7] U. Sirisha, S.P. Praveen, P.N. Srinivasu, P. Barsocchi and A. K. Bhoi, "Statistical Analysis of Design Aspects of Various YOLO-Based Deep Learning Models for Object Detection," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, pp. 1–29, 2023.
DOI: https://doi.org/10.1007/s44196-023-00302-w

[8] J. Lee and K.I. Hwang, "YOLO with adaptive frame control for real-time object detection applications," *Multimedia Tools and Applications*, vol. 81, no. 25, pp. 36375–36396, 2022.
DOI: https://doi.org/10.1007/s11042-021-11480-0

[9] C. Xing, W. Mao and M. Liu, "Scene-aware Human Motion Forecasting via Mutual Distance Prediction," *Computing Research Repository*, vol. abs/2310.00615, pp. 1-20, 2023.
DOI: https://doi.org/10.48550/arXiv.2310.00615

[10] D. Maji, S. Nagori, M. Mathew and D. Poddar, "YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation Using Object Keypoint Similarity Loss," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2022-June, pp. 2636–2645, 2022.
DOI: https://doi.org/10.1109/CVPRW56347.2022.00297

[11] M.F.R. Lee, Y.C. Chen and C.Y. Tsai, "Deep Learning-Based Human Body Posture Recognition and Tracking for Unmanned Aerial Vehicles," *Processes*, vol. 10, no. 11, pp. 1-22, 2022.
DOI: http://dx.doi.org/10.3390/pr10112295

[12] C. Fang, H. Xiang, C. Leng, J. Chen and Q. Yu, "Research on Real-Time Detection of Safety Harness Wearing of Workshop Personnel Based on YOLOv5 and OpenPose," *Sustainability*, vol. 14, no. 10, pp. 5872, 2022..
DOI: https://doi.org/10.3390/su14105872

[13] M. Gao, J. Li, D. Zhou, Y. Zhi, M. Zhang and B. Li, "Fall detection based on OpenPose and MobileNetV2 network," *IET Image Process.*, vol. 17, no. 3, pp. 722–732, 2023.
DOI: http://dx.doi.org/10.1049/ipr2.12667

[14] X. Zhang, Q. Xie, W. Sun, Y. Ren and M. Mukherjee, "Dense Spatial-Temporal Graph Convolutional Network Based on Lightweight OpenPose for Detecting Falls," *Computational Materials Continua*, vol. 77, no. 1, pp. 47–61, 2023.
DOI: https://doi.org/10.32604/cmc.2023.042561

[15] K. Mangalam, E. Adeli, K.H. Lee, A. Gaidon and J.C. Niebles, "Disentangling human dynamics for pedestrian locomotion forecasting with noisy supervision," *Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision*, pp. 2773–2782, 2020.
DOI: https://doi.org/10.1109/WACV45572.2020.9093350

[16] W. Mao, M. Liu and M. Salzmann, "Generating Smooth Pose Sequences for Diverse Human Motion Prediction," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 13289-13298, 2021.
DOI: https://doi.org/10.1109/ICCV48922.2021.01306

[17] V. Guzov, A. Mir, T. Sattler, and G. Pons-Moll, "Human POSEitioning System (HPS): 3D Human Pose Estimation and Self-localization in Large Scenes from Body-Mounted Sensors," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 4316–4327, 2021.
DOI: https://doi.org/10.1109/CVPR46437.2021.00430

[18] Y. Ben-Shabat, X. Yu, F. Saleh, D. Campbell, C. Rodríguez, H. Li and S. Gould, "The IKEA ASM Dataset: Understanding People Assembling Furniture through Actions, Objects and Pose," *Proceedings of the IEEE/CVF Winter Conference on Applications of Vision*, pp. 846-858, 2021.
DOI: https://doi.org/10.1109/WACV48630.2021.00089

[19] M.A. Khatun, M.A. Yousuf, S. Ahmed, M.Z. Uddin, S.A. Alyami, S. Al-Ashhab, H.F. Akhdar, A. Khan, A. Azad and M. A. Moni, "Deep CNN-LSTM With Self-Attention Model for Human Activity Recognition Using Wearable Sensor," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 10, pp. 2700316, 2022.
DOI: http://dx.doi.org/10.1109/JTEHM.2022.3177710

[20] W. Mao, M. Liu, M. Salzmann and H. Li, "Multi-level Motion Attention for Human Motion Prediction," *International Journal of Computer Vision*, vol. 129, no. 9, pp.2513–2535, 2021.
DOI: http://dx.doi.org/10.1007/s11263-021-01483-7

[21] X. Shu, L. Zhang, G.J. Qi, W. Liu, and J. Tang, "Spatiotemporal Co-Attention Recurrent Neural Networks for Human-Skeleton Motion Prediction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 6, pp. 3300–3315, 2022.
DOI: https://doi.org/10.1109/TPAMI.2021.3050918

[22] L. Chen, R. Liu, X. Yang, D. Zhou, Q. Zhang and X. Wei, "STTG-net: a Spatio-temporal network for human motion prediction based on transformer and graph convolution network," *Visual Computing for Industry, Biomedicine, and Art*, vol. 5, no. 1, pp. 19, 2022.
DOI: https://doi.org/10.1186/s42492-022-00112-5

[23] Y. Yang, G. Liu and X. Gao, "Motion Guided Attention Learning for Self-Supervised 3D Human Action Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 12, pp. 8623–8634, 2022.
DOI: https://doi.org/10.1109/TCSVT.2022.3194350

[24] H. Ramirez, S. A. Velastin, I. Meza, E. Fabregas, D. Makris and G. Farias, "Fall Detection and Activity Recognition Using Human Skeleton Features," *IEEE Access*, vol. 9, pp. 33532–33542, 2021.
DOI: https://doi.org/10.1109/ACCESS.2021.3061626

[25] L. Yao, W. Min and K. Lu, "A new approach to fall detection based on the human torso motion model," *Applied Sciences*, vol. 7, no. 10, pp. 993, 2017.
DOI: https://doi.org/10.3390/app7100993

[26] G.J. Horng and K.H. Chen, "The Smart Fall Detection Mechanism for Healthcare Under Free-Living Conditions," *Wireless Personal Communications*, vol. 118, no. 1, pp. 715-

753, 2021.
DOI: https://doi.org/10.1007/s11277-020-08040-4

[27] C.A.U. Hassan, F.K. Karim, A. Abbas, J. Iqbal, H. Elmannai, S. Hussain, S.S. Ullah and M.S. Khan, "A Cost-Effective Fall-Detection Framework for the Elderly Using Sensor-Based Technologies," *Sustainability*, vol. 15,no.5, pp. 3982, 2023. DOI: https://doi.org/10.3390/su15053982

[28] H. Mankodiya, D. Jadav, R. Gupta, S. Tanwar, A. Alharbi, A. Tolba, B.-C. Neagu, and M. S. Raboaca, "XAI-Fall: Explainable AI for Fall Detection on Wearable Devices Using Sequence Models and XAI Techniques," *Mathematics*, vol. 10, no. 12, pp. 1990, 2022.
DOI: https://doi.org/10.3390/math10121990

[29] B.H. Wang, J. Yu, K. Wang, X.Y. Bao and K.M. Mao, "Fall Detection Based on Dual-Channel Feature Integration," *IEEE Access*, vol. 8, pp. 103443–103453, 2020.
DOI: https://doi.org/10.1109/ACCESS.2020.2999503

[30] S. Saurav, R. Saini and S. Singh, "A dual-stream fused neural network for fall detection in multi-camera and 360° videos," *Neural Computing and Applications*, vol. 34, no. 2, pp. 1455–1482, 2022.
DOI: https://doi.org/10.1007/s00521-021-06495-5

[31] N. Worrakulpanit and P. Samanpiboon, "Human Fall Detection Using Standard Deviation of C-Motion Method," *Journal of Automation and Control Engineering*, vol. 2, no. 4, pp. 388–391,2014,
DOI: https://doi.org/10.12720/joace.2.4.388-391

[32] A. Youssfi Alaoui, Y. Tabii, R. Oulad Haj Thami, M. Daoudi, S. Berretti and P. Pala, "Fall Detection of Elderly People Using the Manifold of Positive Semidefinite Matrices," *Journal of Imaging*, vol. 7, no. 7, pp. 109, 2021.
DOI: https://doi.org/10.3390/jimaging7070109

[33] M. Khodarahmi and V. Maihami, "A Review on Kalman Filter Models," *Archives of Computational Methods in Engineering*, vol. 30, no. 1, pp. 727–747, 2023.
DOI: https://doi.org/10.1007/s11831-022-09815-7

[34] A. P. Yunus, K. Morita, N. C. Shirai and T. Wakabayashi, "Time Series Self-Attention Approach for Human Motion Forecasting: A Baseline 2D Pose Forecasting," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 27, no. 3, pp. 445–457, 2023.
DOI: https://doi.org/10.20965/jaciii.2023.p0445

[35] O. Medjaouri and K. Desai, "HR-STAN: High-Resolution Spatio-Temporal Attention Network for 3D Human Motion Prediction," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 2539–2548, 2022.
DOI: https://doi.org/10.1109/CVPRW56347.2022.00286