Jurnal Teknik Informatika Vol. 17 No. 2, Oktober 2024 (171-178)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v17i2.41027

JURNAL TEKNIK INFORMATIKA

Homepage : http://journal.uinjkt.ac.id/index.php/ti

Human Fall Motion Prediction: Fall Motion Forecasting and Detection with GRU

Andi Prademon Yunus[1], Amalia Beladinna Arifa[2], Yit Hong Choo[3]

[1,2]Department of Informatics, Telkom University
[3]Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University
[1,2]Jl. DI Panjaitan No 128, Purwokerto Selatan, Kab. Banyumas, Jawa Tengah, Indonesia
[3]75 Pigdons Rd, Waurn Ponds VIC 3216, Australia

Article:
Accepted: August 26, 2024
Revised: May 27, 2024
Issued: October 29, 2024

***Correspondence Address:**
andiay@telkomuniversity.ac.id

ABSTRACT

The human fall motion prediction system is a preventive tool aimed at reducing the risk of falls. In our research, we developed a deep learning model that utilizes pose estimation to track human body posture and integrated this with a Gated Recurrent Unit (GRU) to forecast human motion and predict falls. GRU, an enhancement of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) models offers improved memorization and more efficient memory usage and performance. Our study presents the human fall motion prediction, which combines the forecasting and classification of potential falls.The CAUCAFall dataset is used as the benchmark of this study, which contains the image sequences of single human motion with ten actions conducted by ten actors. We employed the YOLOv8 Pose model to track the 2D human body pose as the input in our system. A thorough evaluation of the CAUCAFall dataset highlights the effectiveness of our proposed system. Evaluation using the CAUCAFall dataset demonstrates that the model achieved a Mean Per Joint Position Error (MPJPE) of 4.65 pixels from the ground truth, with a 70% accuracy rate in fall prediction. However, the model also exhibited a Mean Relative Error (MRE) of 0.3, indicating that 30% of the predictions were incorrect. These findings underscore the potential of the GRU-based system in fall prevention.

Keywords : *human body motion prediction; fall motion prediction; deep learning; fall;*

Jurnal Teknik Informatika Vol. 17 No. 2, Oktober 2024 (171-178)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v17i2.41027

1. INTRODUCTION

Human activities in some cases can lead to injuries, such as falling or even road traffic accidents. Walking or running is an example of a very basic human activity. Even though the activity is simple, an incident like an ankle sprain could occur in some circumstances. Studies about the sprain ankle mechanism and frequency in sports have been done by examining 2,840 participants in 14 sports, as a result, 1,176 participants sustained injuries and 14% of all injuries involved the ankle [1]. Not only in sports activities, but injuries can also occur in general cases like walking down the stairs, walking or running fast without care movement, and more. In general terms, falls are the second leading cause of unintentional injury death. It is estimated that 684,000 individuals die each year from fall accidents globally, with 80% of the cases occurring in low- and middle-income countries [2]. WHO also reports that 37.3 million falls severely which requires medical attention occur each year [2]. In medical terms, patients fall are the most common type of in-hospital accidents [3].

Preventing human falling caused by various activities is one major goal in many studies. Research conducted by Sourav Kumar Bhoi et. al. utilized accelerometer and gyroscope sensors to classify several activities which include falling for elderly healthcare fall detection system [4]. They employed K-Nearest Neighbors (KNN) and Decision Tree to analyze the data obtained by the sensors. As a result, they achieved 98.75% accuracy with KNN for classification labels: sleeping, sitting, and falling. However, the IoT-based approach has a general limitation that the required devices should be attached to the subject to obtain the data. Therefore, the system linearly costs the subject number, and the device installation needs technician knowledge.

Furthermore, another study used image-based input with only a standard video camera to detect falls and recognition of human activities [5]. To carry out their study, the UP-FALL public dataset was used, employing Random Forest, Support Vector Machine, Multi-Layer Perceptron, and K-Nearest Neighbor. As a result, random forest obtained the best accuracy for fall detection with 99.34 in standard deviation of 0.03. With this approach, the system can easily be installed in any place with only a video camera needed, even though the machine learning approach can be expensive in terms of computational needs. To tackle these computational needs, the traditional calculation is one other approach to detecting the human falling down movement. A study using the human body's middle point as the key feature to calculate the falling action named as Human Torso Motion Model (HTMM). Assuming by tracking the changing rate of torso angle and centroid height, the human falls can be detected by a simple threshold. The calculation is purely based on a mathematical formula, this approach can be optimally processed by any computer.

Aiming the prevention system for humans from falling cannot be done instantly. For example, when the system detects the subject (human) is falling, or has already fallen, the system only has a few milliseconds or even no time to send a device or tools to catch the subject. The previous related research developed models to detect humans falling. However, these models are not meant to prevent humans from falling as there is no time for it. For this matter, we aim to develop a human motion fall prediction, which can forecast where the subject will move and will this person fall in the future. This study adapts to the previous related works which is the most optimal input of the system is video camera based. The input is a sequence of frames with an expected output of another sequence of frames in the next determined steps. A study predicting the next sequence of frames of human pose is called human motion forecasting or human motion prediction. Human motion forecasting studies have been conducted over time. The study differs based on the input and expected output which are 2D human pose and 3D human pose.

For 3D human motion forecasting, a baseline study proposed the Encoder-Recurrent-Decoder (ERD) model[6]. In their research, they used the Human3.6M dataset as a benchmark in human motion forecasting[7]. As a result, they obtained short-term predictions up to 560ms ahead with the best result on LSTM-3LR by motion prediction error as the evaluation metrics. Following the baseline of 3D human motion forecasting, an approach employed the Recurrent Neural Network series like LSTM and GRU with the residual network

to help the model on the previous knowledge before the vanishing gradient [8]. As a result, they obtained better performance than the baseline. The recent works on human motion forecasting with the Human3.6M dataset benchmark are conducted mostly with Mean Per Joint Position Error (MPJPE) as the evaluation metric. Recent studies conducted 3D human motion forecasting with short- and long-term prediction. One study proposed a GCNext, an optimized universal graph convolutional (UniGC) model framework. Their proposed method obtained the best score with 64.7 MPJPE on 1000ms prediction. With GCNext performing better than Separable Temporal Spatial Graph Convolutional Network (STS-GCN) [9], siMLPe [10], PGBIG [11], and MotionMixer [12].

On the other hand, 2D human motion forecasting neglects one dimension in the real-world situation. Some cases like monitoring 2D location for humans require only 2D input and output. It is easier to obtain the data and not expensive for computational needs. A study about 2D human motion forecasting has been done using the same benchmark as the 3D input, which is the Human3.6M dataset on the MPJPE evaluation metric. Research has been conducted by employing the Self-Attention-based method for 2D human motion forecasting[13]. They proposed a Time-Series Self-Attention (TSSA) model for short- and long-term 2D human motion forecasting. As a result, they obtained 10.30 pixels MPJPE for long-term 1000ms prediction. Another experiment developed Temporal-Spatial Time-Series Self-Attention (TS-TSSA) following the TSSA for the same task[14]. However, the TS-TSSA obtained was not better than TSSA with 15.17 pixels MPJPE. As a comparison in the 3D human motion forecasting, TS-TSSA obtained slightly better than STS-GCN with 73.2 MPJPE, but not better than MotionMixer.

Human motion forecasting is a fundamental core tool for human interaction in the real world. Realizing the prevention system to avoid unexpected accidents that happen to humans requires the function of this fundamental work. In this research, we combine human motion forecasting with human fall motion detection to realize a solid fall prevention system. Nevertheless, the previous related works are conducted with annotated datasets. In the real-world case, the human body pose can't be estimated 100% correct. Devices like RGB-D cameras are necessary to capture the human body pose. This brings another problem as the availability of RGB-D cameras is not quite common. Another option is to use the pose estimation method to capture the human body pose from the RGB camera like a standard video camera. This option is more low cost and more available. However, this option also sacrifices the correctness of the dataset since sometimes the pose estimation method does not correctly predict the key point of the human limbs which makes the dataset noisy. For this reason, we conduct this research using pose estimation to capture the human body pose in the dataset. We believe this study strongly contributes as a baseline for human fall motion prediction.

Carnegie Mellon University (CMU) has developed a supervised convolutional neural network called OpenPose, based on Caffe, for real-time multi-person 2D pose estimation [15]. This system accurately estimates human body movements, facial expressions, and finger movements. It is suitable for both single-user and multi-user settings, providing excellent recognition capabilities and fast processing speeds.

Diller et al. 2022 proposed a probabilistic approach to model the potential variability in the distribution of likely characteristic poses to predict them [16].

Vendrow et al. (2022) introduced the Social Motion Transformer (SoMoFormer) for multi-person 3D pose prediction [17]. This transformer architecture uniquely represents human motion input as a joint sequence instead of a time sequence. This allows it to perform attention over individual joints while predicting an entire future motion sequence for each joint simultaneously. SoMoFormer can naturally be applied to multi-person scenes by using the joints of all individuals in a scene as input queries. By leveraging learned embeddings to indicate the type of joint, person identity, and global position, the model can learn the connections between joints and individuals, prioritizing joints from the same or nearby people.

Forecasting human poses from historical pose frames has many important applications, particularly in smart home safety. Deep learning has improved computer vision-based pose forecasting, but there are still two unresolved

Jurnal Teknik Informatika Vol. 17 No. 2, Oktober 2024 (171-178)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v17i2.41027

issues for practical implementation in IoT edge environments. In Li and Li (2021), mentioned that existing methods struggle to accurately predict long-term poses due to inadequate modeling of connected human joint information. Second, pre-trained prediction models may struggle in deployment environments due to visual domain shifts. Hence, they proposed a hybrid cloud-edge system called GPFS (Graph-based Human Pose Forecasting System) to solve those issues [18]. Specifically, a novel graph convolutional neural network (GCN)-based sequence-to-sequence learning method that enriches the sequence encoder by incorporating a graph to represent the spatial and temporal connections of human joints in input frames is presented.

In this research, we strongly believe that this study contributes:

1. A novel idea on human motion fall prediction.
2. A baseline for the human fall motion prediction method.

A novel framework for conducting human motion forecasting and fall detection at the same time

2. METHODS

In this study, we propose human fall motion prediction, combining the forecasting and classification of falling or not falling in the future. Figure 1 shows the overall general flow.

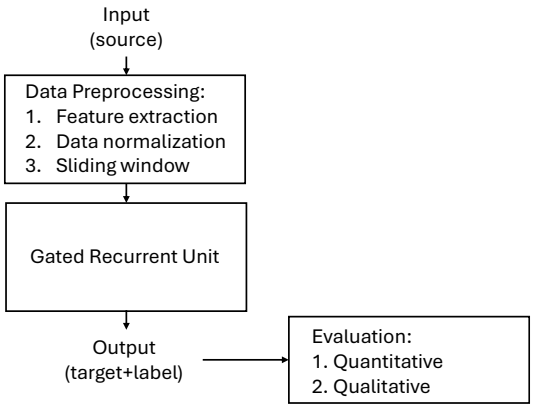


Figure 1. *Overview system flow*

2.1. Data Preprocessing

The feature extraction is done by capturing the human body pose in the sequence of images by utilizing the YOLOv8 pose model. The captured body pose is normalized into the range of 0 to 1, and then a sliding window is employed to make a subset T_p from the length of the whole dataset T. The detail of each step in data preprocessing is described in Section II.1.1, II.1.2, and II.1.3 respectively.

2.1.1. Pose Estimation by YOLOv8 Pose

YOLO is a state-of-the-art object detection model. In general, the YOLO model is built to detect common objects like chairs, animals, tables, etc. Since the model improved by the performance of each version. YOLO can be used in detecting the key points in human body pose or is mostly known for pose estimation. In this research, YOLO version 8 specific to the pose estimation also known as YOLOv8 pose is utilized. YOLOv8 pose can detect multi-person in the frame estimating 17 key points for each person.

2.1.2. Data Normalization

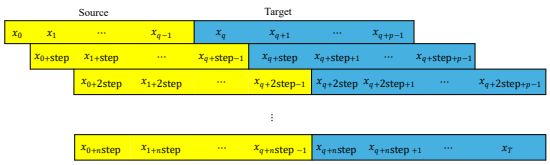
Data normalization is done to simplify the input as well as the give the model a limitation to predict not up to a certain point. With the shape of the image, the frame is 448 times 640 with a range from 0 to 448 or 640, the normalization follows the shape dimension of the data as described in Eq. 1.

$$x_i = \frac{x_i - \min(\text{dim})}{\max(\text{dim}) - \min(\text{dim})}, \qquad (1)$$

where the x_i represents the data in the i-th sequential index, dimension dim is varied to the cartesian coordinate. When the x_i is the x or y coordinate of the key point the maximum value of dimension dim is 448, and 640 for the y coordinate. While the minimum value of the dimension is fixed at 0. With Eq. 1, the data with a range of 448 and 640 could be changed from 0 to 1.

2.1.3. Sliding Window

The sequence of images in the dataset contains a long set of $x_i \in \mathbb{R}^{2K}$ consisting of the x and y coordinates of K keypoints. The sliding window technique is a common technique to make a subset from a sequence of frames. In this research, we utilize the sliding window with source and target labeling schemes for forecasting and classification tasks. Figs. 2 and 3 show the scheme of the sliding window in our research.



Jurnal Teknik Informatika Vol. 17 No. 2, Oktober 2024 (171-178)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v17i2.41027

Figure 2. *Sliding window scheme. Source is the input in the model and target is the expected output of the model. Each row represents a window containing source with length of query q and prediction p. The next window is shifted by the step variable. Data is all on the subset of window till the last index of dataset which is represented by the length of T.*

Figure 3. *Sliding window with label L is concatenated to the target.*

Using a sliding window, the data are adjusted into three parts: source, target, and label. One window contains three of them—these three parts of data are used for different cases. Source will be fed to the model, while the model is expected to predict target and label simultaneously.

2.2. Gated Recurrent Unit

Gated Recurrent Unit (GRU) is one of the improvements of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) with better memorization and more efficiency in terms of memory usage and performance[19]. GRU is developed to solve the vanishing gradient problem that comes with standard RNN. GRU works with two gates which are the update gate and reset gate[19]. In this research, we employed GRU as the model to predict both forecasting and classification tasks.

The GRU model is expected to predict the next ten frames in the shape of (windows, frames, and 1D keypoints of 34) for the forecasting task. In addition, we concatenated the output of the forecasting task with the expected output of binary class 0 or 1 respected to the frames for the fall classification task. As a result, the expected output from the model contains 1D tensors of 35 respected to the frames and windows as shown in Fig. 4.

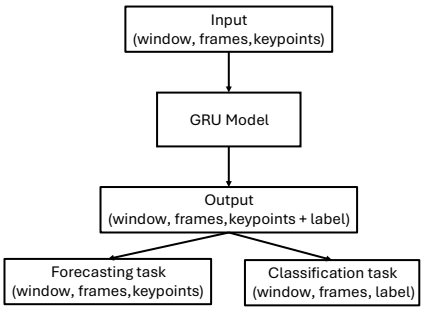


Figure 4. *Forecasting and classification task expected output from GRU model*

2.3. Loss Function

The prediction result from the model training is computed by comparing the distance of the prediction and the ground truth. The prediction result is separated into two parts, which are the forecasting and classes. We calculate the loss function for the forecasting task with the RMSE and classification task with L1 loss as described in the Eq. 2 and 3.

$$\mathcal{L}_{\text{RMSE}} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\left(P_i - \hat{P}_i\right)^2}, \tag{2}$$

$$\mathcal{L}_{L1} = \frac{1}{N}\sum_{i=1}^{N}\left|L_i - \hat{L}_i\right|, \tag{3}$$

where P_i represents the ground truth forecasting data in the index i-th, $\hat{P}_i$ is the prediction result for the forecasting task, N is the total number of the data, L_i is the ground truth label, and $\hat{L}_i$ is the prediction result for falling or not falling down classification task.

This loss then is combined as described in Eq. 4 for the backpropagation loss score in the next iteration or epoch.

$$\mathcal{L} = \mathcal{L}_{\text{RMSE}} + \mathcal{L}_{L1} \tag{4}$$

2.4. Evaluation Method

Following the benchmark of human motion forecasting, this research adopted the MPJPE as evaluation metrics for human motion forecasting as described in Eq. 5, the classification task is evaluated by the Mean Rounded Error (MRE) to show the accuracy of prediction and ground truth as described in Eq. 6.

$$\text{MPJPE} = \frac{1}{N}\left\|\boldsymbol{P} - \hat{\boldsymbol{P}}\right\|, \tag{5}$$

$$\text{MRE} = \frac{1}{N}\left|\boldsymbol{L} - \left\lfloor\hat{\boldsymbol{L}}\right\rceil\right|, \tag{6}$$

where N represents the total number of testing data, $\boldsymbol{P}$ is the vector containing all ground truth for the forecasting task, $\hat{\boldsymbol{P}}$ is the vector containing the prediction result of the forecasting task. $\boldsymbol{L}$ is the vector of all ground truth for the classification task, and $\left\lfloor\hat{\boldsymbol{L}}\right\rceil$ is the rounded prediction result for the classification task.

2.5. Experiment

2.5.1. Dataset

CAUCAFall dataset is used as the main dataset in this research. CAUCAFall is a database created to recognize human falls [20]. This dataset contains 10 actions in an uncontrolled home environment performed by

Jurnal Teknik Informatika Vol. 17 No. 2, Oktober 2024 (171-178)
ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v17i2.41027

10 subjects. There are 5 falling down actions which include forward falls, backward falls, falls to the left, falls to the right, and falls arising from the sitting. Another 5 activities of daily living include walking, hopping, picking up an object, sitting, and kneeling. The dataset also includes the "fall" or "no fall" annotation, which makes this dataset a good fit for our research. Fig. 5 shows one sample of the image illustrating falling backward on subject 1.

Figure 5. *CAUCAFall dataset sample of falling down action*

2.5.2. Experimental Setup

We set up our experiment in a Python environment, using the PyTorch framework to build the deep learning model. The sliding window is performed with the source length q equal to 10, target length p equal to 10, and the step is set to 5. The input size dimension is flattened from $x_i \in \mathbb{R}^{2K}$ to $x_i \in \mathbb{R}^{34}$. The training is run with 1000 epochs as shown in Fig.6. The hidden dimension is set to 128 and the dropout layer with 0.5. We used Stochastic Gradient Descent (SGD) as the optimizer with a learning rate of 0.0001. The subject for training includes subjects 1, 2, 3, 4, 5, and 6. While subject 7 is used for validation, and subjects 8, 9, and 10 are used for testing.

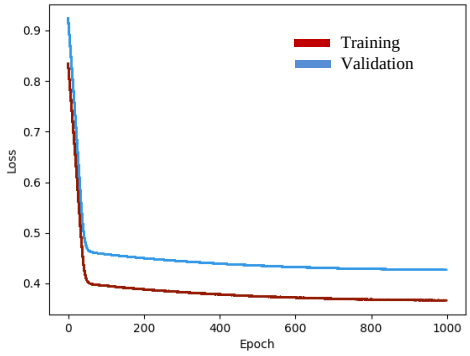


Figure 6. *Training GRU model with 1000 epochs*

3. RESULTS AND DISCUSSION

3.1. Quantitative Evaluation

Fig. 5 shows that the training and validation on the experiment can learn quite well with the significant decreasing loss score around epoch 50. After epochs 60 to 70, the GRU model did not show a significant decrease in the loss score on both training and validation. The evaluation of the testing set using the MPJPE and MRE is shown in Table 1. GRU can predict well with differences of 4.65 pixels away from the ground truth. This indicates GRU quantitatively can produce good forecasting in human motion in the next 10 frames or 400ms in 25 frames per second video. On the other hand, the GRU loss of 0.3 MRE or 30% of the prediction is not correct. This indicates that GRU can only obtain around 70% correct prediction or 0.7 accuracy of human motion fall prediction.

Table 1. *Experiment result*

Evaluation Metric	Score
MPJPE	4.65 (pixels) or 0.0093 (scaled)
MRE	0.30

3.2. Qualitative Evaluation

We randomly illustrate the prediction result compared to the ground truth to show the differences in the human body pose. As a result, Fig. 7 shows the qualitative evaluation in one frame. In a frame, the prediction is quite close to the location of the ground truth. Even though, the human body pose does not mimic the ground truth. This shows that GRU can predict the location of the human body in a frame but fails to predict the human body pose.

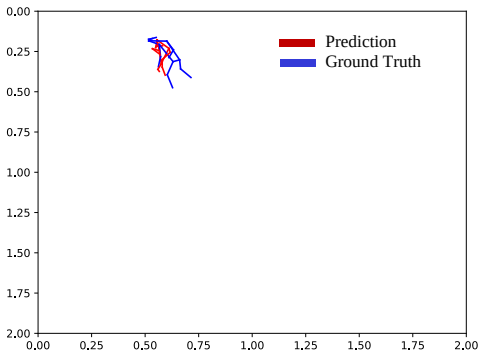


Figure 7. *Qualitative comparison in a frame*

Figure 8 shows the differences in the sequence of human motion forecasting. Showing qualitatively slight differences in human body location and GRU can mimic a bit

ISSN: p-ISSN 1979-9160 (Print)| e-ISSN 2549-7901 (Online)
DOI: https://doi.org/10.15408/jti.v17i2.41027

of the human body pose. However, the prediction results and ground truth data showed different sizes of pose. It indicates the GRU can predict the motion and changes of poses but is still not quite good at mimicking pose forecasting.

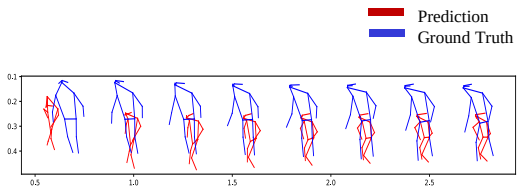


Figure 8. *Qualitative comparison in a sequence*

CONCLUSION

Human fall motion prediction is a preventive prediction system that can be implemented as a core to save humans from falling. In this research, we developed a model to forecast human motion and predict whether motion shortly would be falling or not. Based on the experiment result, a sequential method like GRU fits well with the task. As a result, GRU obtained an MPJPE of 4.65 pixels away from the ground truth on the forecasting task, and 70% of the falling prediction was correct. However, the MRE of 0.3 shows that 30% of the falling prediction is incorrect, the model is far from good enough to be implemented in a real-world case. More research and experiments are needed to develop human fall motion prediction.

REFERENCES

[1] J. G. Garrick, "The frequency of injury, mechanism of injury, and epidemiology of ankle sprains*," Am J Sports Med, vol. 5, no. 6, pp. 241–242, Nov. 1977, doi: 10.1177/036354657700500606.

[2] World Health Organization, "Falls," https://www.who.int/news-room/fact-sheets/detail/falls.

[3] D. C. Anderson, T. S. Postler, and T.-T. Dam, "Epidemiology of Hospital System Patient Falls," American Journal of Medical Quality, vol. 31, no. 5, pp. 423–428, Sep. 2016, doi: 10.1177/1062860615581199.

[4] S. K. Bhoi et al., "FallDS-IoT: A Fall Detection System for Elderly Healthcare Based on IoT Data Analytics," in 2018 International Conference on Information Technology (ICIT), IEEE, Dec. 2018, pp. 155–160. doi: 10.1109/ICIT.2018.00041.

[5] 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: 10.1109/ACCESS.2021.3061626.

[6] K. Fragkiadaki, S. Levine, P. Felsen, and J. Malik, "Recurrent Network Models for Human Dynamics."

[7] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments," IEEE Trans Pattern Anal Mach Intell, vol. 36, no. 7, pp. 1325–1339, Jul. 2014, doi: 10.1109/TPAMI.2013.248.

[8] J. Martinez, M. J. Black, and J. Romero, "On human motion prediction using recurrent neural networks." [Online]. Available: https://github.com/una-dinosauria/

[9] T. Sofianos, A. Sampieri, L. Franco, and F. Galasso, "Space-Time-Separable Graph Convolutional Network for Pose Forecasting." [Online]. Available: https://github.com/FraLuca/STSGCN

[10] W. Guo, Y. Du, X. Shen, V. Lepetit, X. Alameda-Pineda, and F. Moreno-Noguer, "Back to MLP: A Simple Baseline for Human Motion Prediction," Jul. 2022.

[11] T. Ma, Y. Nie, C. Long, Q. Zhang, and G. Li, "Progressively Generating Better Initial Guesses Towards Next Stages for High-Quality Human Motion Prediction," Mar. 2022.

[12] A. Bouazizi, A. Holzbock, U. Kressel, K. Dietmayer, and V. Belagiannis, "MotionMixer: MLP-based 3D Human Body Pose Forecasting," Jul. 2022.

[13] 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, May 2023, doi: 10.20965/jaciii.2023.p0445.

[14] A. P. Yunus, K. Morita, N. C. Shirai, and T. Wakabayashi, "Temporal-Spatial Time Series Self-Attention 2D & 3D Human Motion Forecasting," in 2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), IEEE, Jul. 2023, pp. 66–72. doi: 10.1109/IAICT59002.2023.10205596.

[15] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," Dec. 2018.

[16] C. Diller, T. Funkhouser, and A. Dai, "Forecasting Characteristic 3D Poses of Human Actions," Nov. 2020.

[17] E. Vendrow, S. Kumar, E. Adeli, and H. Rezatofighi, "SoMoFormer: Multi-Person Pose Forecasting with Transformers," Aug. 2022.

[18] X. Li and D. Li, "GPFS: A graph-based human pose forecasting system for smart home with online learning," ACM Trans Sens Netw, vol. 17, no. 3, Aug. 2021, doi: 10.1145/3460199.

[19] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," Dec. 2014.

[20] J. C. E. Guerrero, E. M. España, M. M. Añasco, and J. E. P. Lopera, "Dataset for human fall recognition in an uncontrolled environment," Data Brief, vol. 45, p. 108610, Dec. 2022, doi: 10.1016/j.dib.2022.108610.