A Benchmark of Human Body Movements for Physical Rehabilitation Exercises

Amal Bouallegui †, Mohamed Nidhal Krifa *§, Abdessalem Ben Abdelali ‡§
* Higher Institute of Applied Science and Technology of Kairouan, University of Kairouan, Kairouan, Tunisia;
§ Laboratory of Electronics and Microelectronics (EμE), Faculty of Sciences of Monastir,
University of Monastir, Monastir 5000, Tunisia;
‡ Higher Institute of Computer Science and Mathematics, University of Monastir, Monastir 5000, Tunisia
† Higher Institute of Applied Science and Technology of Kasserine, University of Kairouan, Kairouan, Tunisia;
Emails: amalbouallegue56@gmail.com, kmnidhal@yahoo.fr and abdessalem.benabdelali@enim.rnu.tn

***Abstract*—The analysis of human body movements plays a crucial role in physical rehabilitation, enabling accurate assessment and feedback for patients undergoing therapy. However, the lack of standardized datasets and benchmarks limits the development of robust AI-based evaluation models. In this paper, we propose a benchmark for human body movements in the context of physical rehabilitation exercises. A key contribution is the creation of a custom dataset comprising 400 annotated images, providing a valuable resource for evaluating motion analysis models. After extracting body landmarks, we compare multiple architectures, including a simple artificial neural network (ANN), convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory (LSTM), gated recurrent units (GRU), long-term recurrent convolutional networks (LRCN), and bidirectional LSTM (BiLSTM). The experimental results demonstrate the superiority of the BiLSTM model with an accuracy equal to 91%, highlighting its potential for improving the automatic assessment of rehabilitation exercises. This benchmark serves as a foundation for future research in AI-driven rehabilitation monitoring.**

***Index Terms*—Deep learning, LSTM, BLSTM, Human Body Movements, Physical Rehabilitation Exercises, Movement assessment, Kinect, Vicon.**

I. Introduction

With the rise of telemedicine and digital health, healthcare in cyberspace has become an increasingly important aspect of modern medicine. In particular, the integration of technology with physical rehabilitation movements has shown great promise. Remote monitoring and coaching through cameras, wearable devices and mobile apps can allow healthcare professionals to track patients' progress and provide guidance outside of traditional clinical settings.

A home exercise solution can make the performance of physiotherapy and rehabilitation exercises easier for patients who are unable to move. Indeed, achieving a recommended set of physical exercises in a home environment is often an important component of a patient's rehabilitation treatment. According to the literature, over 90% of rehabilitation sessions are conducted in a home-based environment [1].

There are ongoing efforts to develop telerehabilitation programs that can be carried out in patient homes, and these programs may become a viable option for certain patients in the future [2]. The objectives of this work are to design a home-based tele-rehabilitation protocol for patients. We need to design a solution that can accurately capture body movements and assess exercise quality in real-time, in a home environment, without direct supervision from a therapist. Challenges include managing different patient morphologies, varying lighting conditions, the need to differentiate between very similar postures, and accounting for temporal aspects of movements.

We aim to capture body movements during therapy sessions and automatically assess patient performance and adherence to recommended exercises.

To accomplish this goal, we have:

- Compiled a new dataset consisting of 400 JPG annotated images representing four rehabilitation exercises.
- Conducted an in-depth analysis of human postures by identifying precise anatomical landmarks. Based on these key points, joint connections were mapped to compute angular measurements relative to their spatial positions. The extracted angular data were structured into a dataframe.
- Investigated various architectures for data analysis, including CNN, RNN, LSTM, BiLSTM, GRU, and LRCN. These models incorporate both temporal dimensions (tracking posture evolution over time) and morphological variations (body differences) to assess postures effectively.

Results point out that the BLSTM architecture outperformed the other architectures, achieving a precision rate of 91.03% on the proposed database.

The GRU model achieves a precision of 87%, the precision of the RNN is 85%, and that of the LSTM is 81%. The simple neural networks and the CNN show lower precisions, equal to 73% and 67% respectively. The LRCN, achieves an accuracy of only 71%.

The remainder of this paper is organized as follows. The next section presents the literature work on deep learning strategies applied to assess physical rehabilitation exercises. Section III details the experimental protocol, the database collection, the obtained results, and their discussion. Section IV presents the main conclusion of this work and some future work.

II. RELATED WORK

Mathematical modeling and representation of human movements typically fall into two main categories: top-down approaches, which involve introducing hidden states to describe the temporal dynamics of the movements, and bottom-up approaches, which utilize local features to represent the movements. The first category commonly uses methods such as Kalman filters [7], hidden Markov models [8], [9], and Gaussian mixture models [10].

The recent advancements in artificial neural networks (NNs) have generated considerable interest in their potential for modeling and analyzing human motions [11]. Long Short-Term Memory (LSTM) and Bidirectional LSTM (BLSTM) are two popular types of recurrent neural networks (RNNs) that have been widely used in various fields of research. LSTM and BLSTM are effective in modeling sequential data, such as handwriting recognition [12], speech recognition [13], fingerprint recognition [14], keystroke dynamics recognition [15] due to their ability to capture long-term dependencies and handle vanishing and exploding gradients. Furthermore, this technology has been researched extensively for various real-life applications, including driver action recognition [16], safe intelligent transportation systems (ITS) [17], video text recognition [18], and many others.

LSTM was first introduced by Hochreiter and Schmidhuber in 1997 [19] as a solution to the vanishing gradient problem in traditional RNNs. LSTM uses a memory cell, input gate, forget gate, and output gate to control the flow of information through the network. The memory cell allows the network to store information over time, while the gates regulate the flow of information into and out of the cell.

The recognition of human physical rehabilitation movements using machine learning algorithms, especially LSTM and BLSTM has shown promising results in recent research [20], [21].

LSTM and BLSTM architectures have been utilized to model the complex temporal dependencies present in the rehabilitation movement data. The input to the network consists of time-series data of joint angles, accelerations [22], positions, and other kinematic parameters collected from sensors attached to the patient's body [23]–[25] or from cameras installed in front of the patient.

However, challenges remain, particularly in adapting to interindividual morphological variations, minimizing latency for truly real-time feedback, and generalizing these systems to unstructured environments (variable lighting, occlusions). The integration of these technologies into mainstream clinical applications also requires validating their robustness in uncontrolled settings. This is why selecting the most suitable database remains a significant hurdle.

Indeed, several databases, such as IRDS [3], KIMORE [4], and UI-PRMD [5], have been created and applied within the rehabilitation field. However, these datasets present major limitations for the real-time implementation of deep learning algorithms in practical contexts:

- Controlled Lighting Conditions: Collected in laboratory settings under stable lighting, these datasets struggle to generalize models to real-world environments where illumination varies unpredictably.
- Rigid Camera Configurations: The fixed camera positions, optimized for experimental captures, do not reflect dynamic scenarios (e.g., varying angles, non-standard movements) encountered in clinical or home settings.
- Dependence on Specialized Equipment: The use of expensive sensors (e.g., Kinect V2, Vicon systems) limits compatibility with accessible devices (webcams, smartphones), restricting large-scale adoption.

A new dataset is being developed to:

- Incorporate realistic lighting variations, enhancing model

DATASET	Target group	Population	Sensors	Physical activity	Collected Modalities	Limitations
IRDS [3]	General	29	Kinect V1	Several repetitions of nine general rehabilitation exercises	Skeletal data, depth images	Limited number of modalities, limited number of subjects, discrete labels suited only for HAR research, imbalanced data.
KIMORE [4]	Back pain, Stroke, Parkinson's disease	78	Kinect V2	5 repetitions of 5 exercises for back pain	Skeleton data, depth images, RGB (non-public)	Specific target population, Specific physical activities, Limited number of actions.
UI-PRMD [5]	General	10	Kinect and VICON	10 repetitions of 10 general rehabilitation exercises.	Skeleton data	Limited number of modalities, limited number of subjects.
AHA-3D [6]	Assessment of lower body fitness levels	21	Kinect V2 RGB camera	79 sequences of 4 actions (lower limb)	Skeleton data, depth images, RGB images	Specific physical activities, limited to a few limbs

TABLE I: analysis of some available databases.

robustness.

- Represent dynamic camera positions, including simulated movements and variable angles.
- Ensure compatibility with standard hardware, facilitating real-world usability.

The proposed dataset aims to bridge the gap between controlled environments and clinical needs, providing a cost-effective solution tailored to the challenges of real-time rehabilitation.

III. Proposed Benchmark for Motion Recognition System

A. Data Collection

Building upon the UI-PRMD dataset [5] , we developed a dataset specifically tailored to our domain, prioritizing accessibility and practicality. Unlike many existing approaches that rely on specialized hardware and motion capture sensors, often making implementation costly, our solution leverages readily available resources. The posture data used in our study are sourced from publicly accessible online sources, providing a cost-effective alternative. Additionally, most existing datasets are proprietary or require payment, restricting their usability. In contrast, our approach ensures greater accessibility by enabling real-time detection using only a standard camera. Furthermore, our dataset addresses common challenges related to lighting conditions, camera angles, and patient positioning, enhancing its robustness for real-world applications.

This dataset consists of 400 JPG images, organized into four distinct folders, each corresponding to a specific rehabilitation exercise posture:

- Deep squat,
- Hurdle step,
- Inline lunge, and
- Side lunge.

These folders are grouped within a main directory. To ensure diversity and relevance in the targeted postures, all images were sourced from Google Images.

B. Feature Extraction

In our approach, we utilized MediaPipe library [26] to detect and visualize human body landmarks on a given image. When key points are successfully identified, they are overlaid on the image, each represented by a small red circle as depicted in Fig.1. The position of each landmark is determined using normalized coordinates (x, y), which are then scaled by the image dimensions to obtain their exact positions on the screen.

Extracting angles from landmarks detected in images is a critical step in pose analysis, especially for applications such as posture correction or activity recognition. Using tools like MediaPipe Pose, key points of the body (e.g., joints such as shoulders, elbows, hips, and knees) are identified in the image.

Fig. 1: Detection of human body landmarks

These key points are then used to calculate joint angles using trigonometric principles.

The angle formed by three points (two limb extremities and the central joint) is calculated using the arctangent function. The formula applied is as follows:

$$\text{Angle} = |\arctan 2(C_y - B_y, C_x - B_x) -$$
$$\arctan 2(A_y - B_y, A_x - B_x)(1)$$

Where:

B represents the central joint (e.g., elbow or knee), A and C are the adjacent points.

To ensure that the angle remains between 0° and 180°, it is normalized as follows:

$$\text{angle} = \begin{cases} \text{angle}, & \text{if angle} \leq 180^\circ \\ 360^\circ - \text{angle}, & \text{if angle} > 180^\circ \end{cases} \quad (2)$$

This normalization ensures measurement consistency and provides an accurate representation of body posture. The extracted angles serve as key features for pose analysis, enabling deviation detection and facilitating appropriate corrections.

Angles are calculated for various joints based on the key point numbering provided by MediaPipe Pose. The primary analyzed joints include: the upper body, represented by the shoulders and elbows; the lower body, represented by the hips and knees; and the overall torso alignment. These angles provide a comprehensive representation of body posture, enabling detailed and precise analysis.

The angles calculated for each image are stored in a list, thus forming a set of features. The associated labels, extracted from the folder name corresponding to the pose category, are added to another list. Once all the images in a folder are processed, these lists are combined to create a Pandas DataFrame.

This DataFrame contains columns representing the names of the calculated angles (e.g., leftelbow, rightelbow, leftknee, etc.), as well as an additional column, label, which indicates the corresponding pose category. Finally, the DataFrame is saved as a CSV file titled pose-data.csv. Each row in this file represents an individual image, including its extracted angles

Category	Articulation	Description
Elbows	Left Elbows	Angle between left shoulder, left elbow and left wrist
	Right Elbows	Angle between right shoulder, right elbow and right wrist
Shoulders	Left Shoulders	Angle between left hip, left shoulder and left elbow
	Right Shoulders	Angle between right hip, right shoulder and right elbow
Knees	Left Knees	Angle between left hip, left knee and left ankle
	Right Knees	Angle between right hip, right knee and right ankle
Hips	Left Hips	Angle between left shoulder, left hip and left knee
	Right Hips	Angle between right shoulder, right hip and right knee
Torso	Left Torso	Angle between left shoulder, left hip and left ankle
	Right Torso	Angle between right shoulder, right hip and right ankle

TABLE II: Description of joint angles for postural analysis

and associated pose category. This structure allows for efficient data organization and facilitates its use for analysis or machine learning tasks.

C. Model Evaluation

1) Data management : Splitting data into training and testing sets is a crucial step in the machine learning model development process. It allows evaluating the model's performance on previously unseen data. 20% of the data is reserved for the test set, while the remaining 80% is used for model training. We ensured the reproducibility of the results by setting a constant random state. This split ensures that the model is trained on a subset of the data and evaluated on a separate set, which allows for efficient measurement of its performance and reduces the risk of overfitting.

D. Obtained results

The proposed architecture demonstrated promising results as it was able to recognize the four rehabilitation exercises based on the collected images.

The effectiveness of the architectures tested was assessed in the test set of the proposed database, with accuracy serving as the primary evaluation metric. Comparing the results obtained in Figure 2, we notice a considerable improvement in the classification accuracy when dealing with the BLSTM architecture. In fact, BLSTM can extract more comprehensive features from the data. It can capture different patterns and representations from the forward and backward directions, enabling the model to capture complex relationships and nuances in the data more effectively.

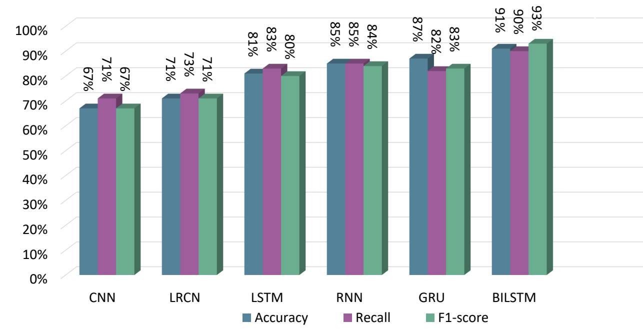


Fig. 2: Obtained performances for all tested models

The confusion matrix for human body movement classification in physical rehabilitation exercises provides an insightful analysis of model performance in different exercises as shown in Figure 3. Each row represents the actual exercise performed, while each column indicates the predicted exercise, allowing for a clear visualization of correct and misclassified instances. The results highlight that all exercises are generally well recognized, with high classification accuracy across most movements. Notably, exercise number 3 exhibits the highest performance, showing the least misclassification and the highest true positive rate. This suggests that the model is particularly effective in distinguishing this exercise from the others, possibly due to its distinct motion patterns or well-represented features in the dataset.

E. Comaprison

The table III provides a comparative analysis of various research approaches in the field of physical rehabilitation exercise recognition and assessment, emphasizing their methodologies, datasets, and achieved accuracies. Notably, my approach distinguishes itself through the use of the BILSTM architecture applied to the collected data, achieving an impressive accuracy of 91.03%, surpassing all other approaches.

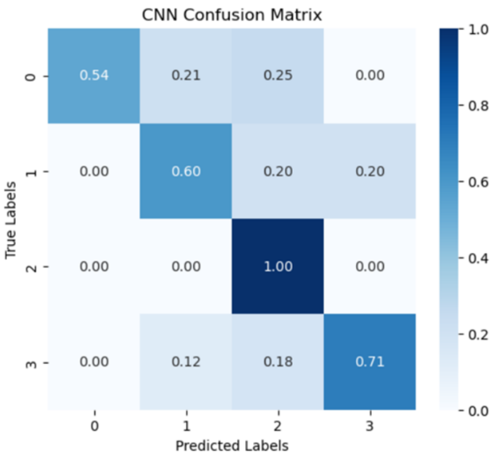
Article	Method	Accuracy
Proposed approach	BILSTM	91.03%
[27]	CNN(YOLO)	84%
[28]	CNN	82%
[21]	LSTM	83%

TABLE III: Comparison

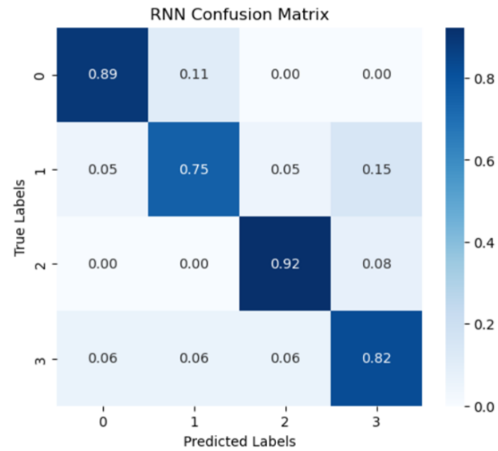
IV. Conclusion

In this paper, we propose a system that allows the assessment of human posture during physical rehabilitation exercises by Deep Learning.

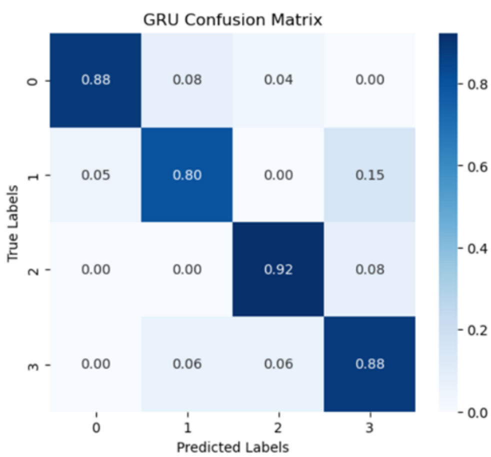
A key contribution of this work is the development of a custom dataset comprising 400 annotated images of 4 different rehabilitation exercises. Following landmark extraction, multiple architectures were evaluated, including CNN, RNN,



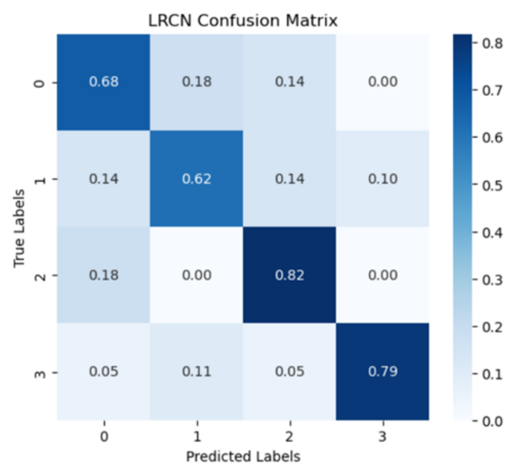
(a) CNN architecture



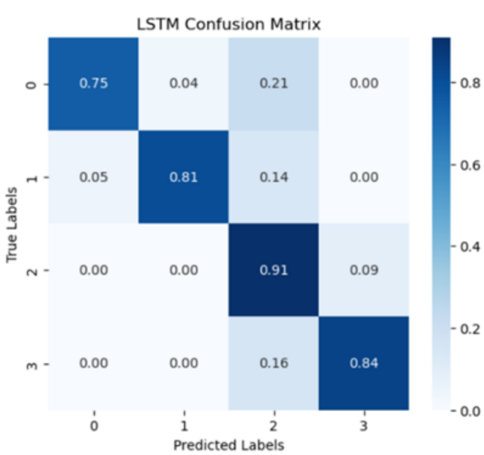
(b) RNN architecture



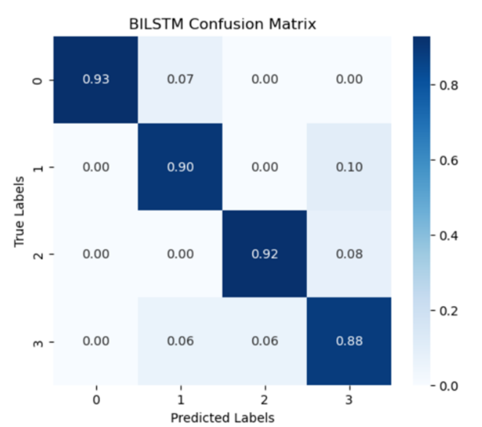
(c) GRU architecture



(d) LRCN architecture



(e) LSTM architecture



(f) BLSTM architecture

Fig. 3: Confusion matrices of the implemented architectures

LSTM, GRU, LRCN, and BiLSTM. The results demonstrated the superior accuracy of the BiLSTM model.

This paper put forward a novel framework for the assessment of home-based rehabilitation that investigates a deep neural network model designed to handle spatial and temporal variability in human movements.

The obtained results are promising and can be further enhanced by incorporating additional rehabilitation exercises into the dataset. Furthermore, we plan to implement this approach on an FPGA board to assess its performance in a real-world application.

References

1 Komatireddy, R., Chokshi, A., Basnett, J., Casale, M., Goble, D., and Shubert, T., "Quality and quantity of rehabilitation exercises delivered by a 3-d motion controlled camera: A pilot study," *International journal of physical medicine & rehabilitation*, vol. 2, no. 4, 2014.

2 Chen, Y., Abel, K. T., Janecek, J. T., Chen, Y., Zheng, K., and Cramer, S. C., "Home-based technologies for stroke rehabilitation: A systematic review," *International journal of medical informatics*, vol. 123, pp. 11–22, 2019.

3 Miron, A., Sadawi, N., Ismail, W., Hussain, H., and Grosan, C., "Intellirehabds (irds)—a dataset of physical rehabilitation movements," *Data*, vol. 6, no. 5, p. 46, 2021.

4 Capecci, M., Ceravolo, M. G., Ferracuti, F., Iarlori, S., Monteriu, A., Romeo, L., and Verdini, F., "The kimore dataset: Kinematic assessment of movement and clinical scores for remote monitoring of physical rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 7, pp. 1436–1448, 2019.

5 Vakanski, A., Jun, H.-p., Paul, D., and Baker, R., "A data set of human body movements for physical rehabilitation exercises," *Data*, vol. 3, no. 1, p. 2, 2018.

6 Antunes, J., Bernardino, A., Smailagic, A., and Siewiorek, D. P., "Aha-3d: A labelled dataset for senior fitness exercise recognition and segmentation from 3d skeletal data." in *BMVC*, 2018, p. 332.

7 Yun, X. and Bachmann, E. R., "Design, implementation, and experimental results of a quaternion-based kalman filter for human body motion tracking," *IEEE transactions on Robotics*, vol. 22, no. 6, pp. 1216–1227, 2006.

8 Panahandeh, G., Mohammadiha, N., Leijon, A., and Händel, P., "Continuous hidden markov model for pedestrian activity classification and gait analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 5, pp. 1073–1083, 2013.

9 Deters, J. K. and Rybarczyk, Y., "Hidden markov model approach for the assessment of tele-rehabilitation exercises," *International Journal of Artificial Intelligence*, vol. 16, no. 1, pp. 1–19, 2018.

10 Huang, Y., Englehart, K. B., Hudgins, B., and Chan, A. D., "A gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 11, pp. 1801–1811, 2005.

11 Vakanski, A., Ferguson, J., and Lee, S., "Mathematical modeling and evaluation of human motions in physical therapy using mixture density neural networks," *Journal of physiotherapy & physical rehabilitation*, vol. 1, no. 4, 2016.

12 Carbune, V., Gonnet, P., Deselaers, T., Rowley, H. A., Daryin, A., Calvo, M., Wang, L.-L., Keysers, D., Feuz, S., and Gervais, P., "Fast multi-language lstm-based online handwriting recognition," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 23, no. 2, pp. 89–102, 2020.

13 Bhaskar, S. and Thasleema, T., "Lstm model for visual speech recognition through facial expressions," *Multimedia Tools and Applications*, vol. 82, no. 4, pp. 5455–5472, 2023.

14 Jia, B., Qiao, W., Zong, Z., Liu, S., Hijji, M., Del Ser, J., and Muhammad, K., "A fingerprint-based localization algorithm based on lstm and data expansion method for sparse samples," *Future Generation Computer Systems*, vol. 137, pp. 380–393, 2022.

15 Mhenni, A., Rosenberger, C., and Essoukri Ben Amara, N., "Keystroke dynamics classification based on lstm and blstm models," in *2021 International Conference on Cyberworlds (CW)*. IEEE, 2021, pp. 295–298.

16 Jegham, I., Khalifa, A. B., Alouani, I., and Mahjoub, M. A., "Soft spatial attention-based multimodal driver action recognition using deep learning," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1918–1925, 2021.

17 Chougule, A., Shah, J., Chamola, V., and Kanhere, S., "Enabling safe its: Eeg-based microsleep detection in vanets," *IEEE Transactions on Intelligent Transportation Systems*, 2022.

18 Muthumani, I., Malmurugan, N., and Ganesan, L., "Resnet cnn with lstm based tamil text detection from video frames." *Intelligent Automation & Soft Computing*, vol. 31, no. 2, 2022.

19 Schmidhuber, J., Hochreiter, S. *et al.*, "Long short-term memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, 1997.

20 Guerra, B. M. V., Ramat, S., Beltrami, G., and Schmid, M., "Recurrent network solutions for human posture recognition based on kinect skeletal data," *Sensors*, vol. 23, no. 11, p. 5260, 2023.

21 Tsakanikas, V. D., Gatsios, D., Dimopoulos, D., Pardalis, A., Pavlou, M., Liston, M. B., and Fotiadis, D. I., "Evaluating the performance of balance physiotherapy exercises using a sensory platform: The basis for a persuasive balance rehabilitation virtual coaching system," *Frontiers in digital health*, vol. 2, p. 545885, 2020.

22 Postolache, O., Hemanth, D. J., Alexandre, R., Gupta, D., Geman, O., and Khanna, A., "Remote monitoring of physical rehabilitation of stroke patients using iot and virtual reality," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 2, pp. 562–573, 2020.

23 Alemayoh, T. T., Lee, J. H., and Okamoto, S., "Leg-joint angle estimation from a single inertial sensor attached to various lower-body links during walking motion," *Applied Sciences*, vol. 13, no. 8, p. 4794, 2023.

24 Bijalwan, V., Semwal, V. B., and Gupta, V., "Wearable sensor-based pattern mining for human activity recognition: Deep learning approach," *Industrial Robot: the international journal of robotics research and application*, vol. 49, no. 1, pp. 21–33, 2022.

25 Kristoffersson, A. and Lindén, M., "A systematic review of wearable sensors for monitoring physical activity," *Sensors*, vol. 22, no. 2, p. 573, 2022.

26 Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M., Lee, J. *et al.*, "Mediapipe: A framework for perceiving and processing reality," in *Third workshop on computer vision for AR/VR at IEEE computer vision and pattern recognition (CVPR)*, vol. 2019, 2019.

27 Izadmehr, Y., Satizábal, H. F., Aminian, K., and Perez-Uribe, A., "Depth estimation for egocentric rehabilitation monitoring using deep learning algorithms," *Applied Sciences*, vol. 12, no. 13, p. 6578, 2022.

28 Cejnog, L. W. X., de Campos, T., Elui, V. M. C., and Cesar Jr, R. M., "A framework for automatic hand range of motion evaluation of rheumatoid arthritis patients," *Informatics in Medicine Unlocked*, vol. 23, p. 100544, 2021.