Artificial Intelligence-driven Remote Monitoring Model for Physical Rehabilitation

Mohamed Jleli[1,*] , Bessem Samet[1] and Ashit Kumar Dutta[2,*]

[1]Department of Mathematics, College of Science, King Saud University, Riyadh 11451, Saudi Arabia
[2]Department of Computer Science and Information Systems, College of Applied Sciences, AlMaarefa University, Ad Diriyah, Riyadh 13713, Saudi Arabia

Correspondence to:
Mohamed Jleli*, e-mail: jleli@ksu.edu.sa
Ashit Kumar Dutta*, e-mail: adotta@mcst.edu.sa

ABSTRACT

In recent times, the quality of life of several individuals has been affected by chronic diseases. Traditional forms of rehabilitation occasionally involve face-to-face sessions, which restricts accessibility and presents challenges for real-time monitoring. Lack of comprehensive understanding of the aspects impacts long-term patient engagement and adherence to remote rehabilitation programs. Individuals and healthcare systems incur a significant portion of the costs associated with rehabilitation treatment. A home-based rehabilitation program reduces the rehabilitation cost. However, the clinicians' absence may affect the effectiveness of rehabilitation programs. There is a demand for an artificial intelligence-based remote monitoring model for evaluating the physical movements of individuals. Therefore, the study proposes a framework for generating scores for physical rehabilitation exercises. It supports the home-based rehabilitation program by assessing the individual's movements. The authors employ the You Only Look Once V5–ShuffleNet V2-based image processor for generating scores using the variations between the joints. In addition, they build bidirectional long short-term memory networks for delivering a score for each exercise. Finally, the two outcomes are compared using the modulated rank averaging method for presenting the final score. The authors evaluate the performance of the proposed model using the KiMoRe dataset. The comparative analysis outcome suggested that the proposed model obtained an exceptional mean absolute deviation, mean absolute percentage error, and root mean square error of 0.425, 1.120, and 0.985, respectively. It highlighted the significance of the proposed framework in assessing the individual's physical movement. Further studies will improve the performance of the proposed framework.

KEYWORDS

rehabilitation, Bi-LSTM, ShuffleNet V2, assessment score, quality of life

INTRODUCTION

Rehabilitation is frequently recommended for individuals with physical limitations or who need to regain functional ability resulting from an accident or surgery (Capecci et al., 2019). In the literature, several studies emphasize the significance of physical rehabilitation for better patient outcomes and the robust relationship between exercise intensity and rehabilitation success rates (Zhao et al., 2017; Vakanski et al., 2018; Zhang et al., 2019). In clinical rehabilitation programs, clinicians teach and supervise patients' activities (Li et al., 2018). This sort of rehabilitation treatment requires patients' schedules and is limited by clinician availability. Thus, home-based rehabilitation is utilized as a complement to clinic-based programs to provide additional program adaptability. In home-based rehabilitation, the physician prepares a unique rehabilitation program for each patient by suggesting a series of activities (Allahbakhsh et al., 2020).

Patients follow the instructions, register their daily progress, and attend the clinic for progress assessments. Recent studies report that over 90% of rehabilitation treatments are carried out at home (Sarsfield et al., 2019).

In order to regain muscular strength along with enhanced balance, individuals with a variety of physical impairments benefit greatly from participating in home-based rehabilitation programs and performing the recommended physical exercises (Lin et al., 2018). Without a medical specialist, individuals in these programs cannot evaluate their action performance. In the field of activity monitoring, vision-based sensors have become more prevalent in recent years. These devices may obtain highly accurate measurements of skeletons. Recent advances in computers, robotics, machine learning, connectivity, and sensor downsizing provide compassionate intelligent gadgets and technology-embedded

environments (Petersen et al., 2018; Prabhu et al., 2020; Ferreira et al., 2021). The majority of intelligent systems are for manufacturing, military, space exploration, and entertainment. Improving health-related quality of life is an emerging topic in artificial intelligence (AI)-based applications. Assistive technology has become compromised in the shuffle between breakthroughs in medical and intelligent system technologies. In order to build intelligent devices and systems that assist and interact with humans, a fundamental knowledge of how they relate to human functions is essential.

In recent years, the quality of life of stroke survivors has been the focus of significant research and development efforts. These efforts have focused on utilizing a diverse array of technological tools (Nambi et al., 2017; Derungs et al., 2018; Soro et al., 2019; Ishii et al., 2020). However, relatively few medical institutions make use of computer-based rehabilitation solutions. According to the studies, more than 90% of rehabilitation takes place at home. To achieve the best possible outcome in terms of safety and efficacy, a rehabilitation program is typically delivered in clinical facilities or at home by a physiotherapist who continuously provides feedback on gesture accuracy, in terms of goal, motion, and posture (Ebert et al., 2017; Sardari et al., 2020; Raihan et al., 2021; Kanade et al., 2022; Rahman et al., 2022; Li et al., 2023). Motor learning and retention are enhanced by the feedback and information on gesture accuracy provided between and following a physical movement (Bevilacqua et al., 2020; Liao et al., 2020; Deb et al., 2022; Abedi et al., 2023). To provide safe and successful rehabilitation, the supervision of the gesture is essential to achieve an effective outcome.

In order for the patients to keep track of their progress in the rehabilitation program, they are required to either self-monitor or seek assistance from family members. The voluntary nature of home-based rehabilitation programs might lead to low patient adherence, prolonging posthospitalization recovery (Chowdhury et al., 2021; Guo and Khan, 2021; Mottaghi and Akbarzadeh-T, 2022). Home-based rehabilitation lacks corrective input on movement quality and accuracy. Robotic-assistive devices, virtual reality and gaming interfaces, and Kinect-based support are a few of the technology tools accessible to patients receiving home-based rehabilitation (Mourchid and Slama, 2023).

AI and other forms of cutting-edge technology have advanced rehabilitation to an entirely new level (Mourchid and Slama, 2023). It is possible to train AI algorithms to analyze enormous amounts of patient data, such as medical history, vital signs, and lifestyle variables, in order to generate a tailored treatment plan for the patient. More precise and rapid diagnosis might result in better health outcomes for patients (Mourchid and Slama, 2023). AI enables physicians to closely monitor rehabilitation progress and adapt treatment plans to speed patient recovery. Deep learning (DL) is a widely applied AI technique for classifying medical images (Albert et al., 2021). It can assist therapists in estimating the rehabilitation period. In addition, it determines whether their patients are ready for the subsequent phase in their recovery process. Rehabilitation solutions driven by AI may lower the financial and logistical barriers that prevent more

individuals from benefiting from physical treatment (Albert et al., 2021). If rehabilitation is more efficient and successful, patients will need fewer sessions, reducing healthcare expenses (Maradani and Levkowitz, 2017).

A significant number of individuals globally necessitate physical rehabilitation; nonetheless, the accessibility of rehabilitation services may be restricted due to variables such as geographical location, mobility impairments, or limitations in healthcare resources. Traditional treatment involves regularly attending sessions in person, resulting in a lack of monitoring and intervention in the time between sessions. The research motivation for building an AI-driven remote monitoring model for physical rehabilitation is to enhance accessibility, individualization, continual tracking, and the overall efficacy of rehabilitation services. Implementing AI-based technology can enhance rehabilitation, yielding more significant outcomes for individuals undergoing physical rehabilitation.

In Vision 2030, the Kingdom established the health sector transformation program. The initiative aims to provide universal healthcare, improve regional fairness, and promote e-health utilization (Zahra et al., 2022). The program prioritizes health innovation in order to enhance health care, particularly rehabilitation. However, there is a demand for automating the process of monitoring rehabilitation programs to provide an effective service for disabled individuals. Therefore, the study intends to generate an AI-driven movement assessment score for monitoring the home-based rehabilitation program.

The contributions of this study are:
- An effective rehabilitation monitoring framework using bidirectional long short-term memory (Bi-LSTM) and You Only Look Once (YOLO) V5–ShuffleNet V2 models.
- Evaluation of the proposed framework using the benchmark dataset and the baseline models.

The study is organized as follows: The Literature Review section presents the features and limitations of the existing literature. The research methodology is presented in the Materials and Methods section. The Results section presents the experimental outcomes of the proposed study. The study findings are discussed in the Discussions section. Finally, the Conclusion section concludes the proposed study.

LITERATURE REVIEW

In AI, human behavior analysis is considered a significant and formidable challenge (Capecci et al., 2019). It analyzes human body movements by evaluating joint, bone, and muscle motions. The modeling and analysis of human movements using DL techniques have recently gained popularity due to their exceptional performance (Zhao et al., 2017; Li et al., 2018; Lin et al., 2018; Vakanski et al., 2018; Sarsfield et al., 2019; Zhang et al., 2019; Allahbakhsh et al., 2020). It has been widely used for motion classification, gesture recognition, and action localization. DL techniques, including convolutional neural networks (CNNs), LSTM, and gated recurrent units, are employed to assess the physical movements in

the rehabilitation program (Lin et al., 2018; Petersen et al., 2018; Prabhu et al., 2020; Ferreira et al., 2021). These movements can be interpreted as various gestures, human–human interactions, group activities, and behaviors depending on the level of complexity. These activities can reveal persons' personalities, physiological and psychological states, and potential goals and intents. Automatic human behavior analysis systems are gaining popularity for assisting professionals in healthcare, public monitoring, and driverless systems (Nambi et al., 2017; Derungs et al., 2018; Soro et al., 2019; Ishii et al., 2020; Li et al., 2023). Nevertheless, the effectiveness of motion tracking, data preprocessing, representation learning, and assessment approaches heavily influences the development of a human activity analysis system.

In order to develop effective tools and systems for home-based rehabilitation, it is crucial to quantify the degree of accuracy in performing recommended activities. Existing studies typically analyze the movement by contrasting a patient's performance on a certain activity with the optimal performance of healthy individuals. Several previously published research used machine learning techniques to categorize workout repetitions as correct or incorrect (Ebert et al., 2017). Researchers applied the Adaboost classifier, k-nearest neighbors, and CNN models to assess the individual's performance quality. These approaches were unable to identify the movement quality or incremental patient performance improvements throughout rehabilitation (Rahman et al., 2022).

It is common practice for clinicians to provide their patients with a series of exercises combined with a recommended number of repetitions (Sardari et al., 2020). Evaluating exercise performance involves objective criteria such as following set and repetition guidelines, maintaining adequate technique, quality of movements, and right posture. Thus, temporal segmentation is the initial step in an AI-driven workout assessment pipeline. Segmentation can be used to calculate exercise repetition counts, or it may be conducted independently. Multiple data sources, such as the inertial measurement unit, sensor, video, and skeletal position data, were utilized for repetitive segmentation and counting.

The existing datasets cover a small set of physical movements. The limited number of the participants may reduce the performance of the CNN model in generating the assessment score. The KiMoRe dataset offers video, images, and skeletal body joints (BJ) data (Capecci et al., 2019). The number of healthy and unhealthy individuals participating in the dataset is higher than the existing datasets. In addition, recent studies are widely applying this dataset for developing physical rehabilitation assessment models. The studies (Ebert et al., 2017; Derungs et al., 2018; Bevilacqua et al., 2020; Ishii et al., 2020; Liao et al., 2020; Sardari et al., 2020; Chowdhury et al., 2021; Guo and Khan, 2021; Raihan et al., 2021; Deb et al., 2022; Kanade et al., 2022; Rahman et al., 2022; Abedi et al., 2023; Li et al., 2023) utilized the KiMoRe dataset to their assessment frameworks. Sardari et al. (2020) proposed a Vi-net-based movement assessment model. They employed images to evaluate the individual's movements. Kanade et al. (2022) built a DL-based framework for assessing movement quality. The skeletal BJ data were used to generate a score for each individual's rehabilitation exercise.

Raihan et al. (2021) developed an exercise assessment model using a genetic algorithm (GA)-optimized CNN. They generated scores for each exercise using a local binary pattern (LBP) mechanism.

Deb et al. (2022) extracted features from the red, green, and blue (RGB) videos for presenting the quality assessment score. Bevilacqua et al. (2020) developed an LSTM and a boosting aggregation exercise assessment model. They employed accelerometer and gyroscopic data. Liao et al. (2020) proposed an assessment framework for supervising physical rehabilitation exercises. They employed a log likelihood of a Gaussian mixture model for the score generation. In addition, a deep encoder network was used to encode the low-dimensional data. Abedi et al. (2023) developed a set of DL models for exercise repetition segmentation using the skeletal BJ data. They applied multiple sequential neural networks for producing an outcome. Chowdhury et al. (2021) utilized depth sensor data and developed an assessment model for supervising the rehabilitation exercises. Guo and Khan (2021) employed a feature extraction method for evaluating physical rehabilitation exercises. The RGB videos of the KiMoRe dataset are used for building the model. Mottaghi and Akbarzadeh-T (2022) developed deep mixture density neural networks for the automated evaluation of rehabilitation exercises. Lastly, Albert et al. (2021) used generative adversarial networks for extracting the features from the RGB videos. The existing literature focused on implementing a score generator using the RGB videos, images, and skeletal BJ data. However, the current models demand a substantial computational cost and high training time. There is a demand for an effective feature extraction technique to support the score generator model.

There is a lack of significant clinical validation and evidence about the effectiveness of AI-driven remote monitoring models in various rehabilitation scenarios; lack of collaboration between AI researchers and rehabilitation specialists; insufficient and comprehensive ethical and legal frameworks for implementing AI in rehabilitation; and inadequate validation of AI-driven models across heterogeneous patient populations encompassing various age cohorts, cultural contexts, and medical conditions.

MATERIALS AND METHODS

In this study, the authors developed an integrated framework for assessing the movements of disabled individuals. Table 1 presents the features and limitations of the current literature. Figure 1 highlights the proposed assessment framework (PAF). It includes a preprocessing method, YOLO V5 (Ge et al., 2021)–ShuffleNet V2 (Ma et al., 2018) model, and Bi-LSTM (Huang et al., 2015) model. The preprocessing technique is employed in order to improve the image quality and remove irrelevant data from the dataset. The authors apply YOLO V5 to extract features from the images. In addition, the ShuffleNet V2 model classifies the images and generates a score using the movement variations. In addition, Bi-LSTM is used to process the BJ data. The modulated rank averaging (MRA) method (De and Chowdhury,

Table 1: Characteristics of the existing literature.

Authors	Methods	Dataset	Features	Limitations
Sardari et al. (2020)	Two-dimensional CNN	KiMoRe	Used the Vi-net model to process the images and generate a score	The skeletal BJ data are not considered in the study.
Kanade et al. (2022)	Transformer-based attention mechanism	KiMoRe	Applied the transformer network and attention mechanism for data imputation	The study's outcome was limited to the skeletal BJ data.
Raihan et al. (2021)	CNN	KiMoRe	Employed GA-based CNN and local binary mechanism for score generation	GA requires a huge computation cost for optimizing the CNN model.
Deb et al. (2022)	Graph convolutional network	KiMoRe	Applied graph convolutional network (GCN)-based quality assessment score generator	GCN requires an additional computational cost.
Bevilacqua et al. (2020)	LSTM model	KiMoRe	Employed LSTM and boosting aggregation method for processing the skeletal BJ data	The LSTM model passes the information in one direction. The model may face challenges in dealing with data overfitting.
Liao et al. (2020)	Gaussian mixture model	University of Idaho- Physical Rehabilitation Movement Dataset (UI-PRMD)	Employed log likelihood to compute a score	UI-PRMD contains 10 exercises without ground truth labels.
Abedi et al. (2023)	DL models	KiMoRe, UI-PRMD, and IntelliRehabDS	Applied a set of DL techniques	Irregularities in the exercises may affect the model's efficiency.
Chowdhury et al. (2021)	DL model	KiMoRe	Employed depth sensor data and graph convolutional network	Graph convolutional network demands an additional computational time.
Guo and Khan (2021)	DL-based feature extraction model	KiMoRe	Utilized the RGB videos and extracted features	The model requires a huge computational cost for processing the videos.
Mottaghi and Akbarzadeh-T (2022)	DL model	KiMoRe	Deep mixture density neural network	The model's performance is based on the low-quality sensor data.
Albert et al. (2021)	CNN	KiMoRe	Employed GAN for the feature extraction	The model presented the features rather than generating score for each exercise.

Abbreviations: BJ, body joints; CNN, convolutional neural network; DL, deep learning; GA, genetic algorithm; GAN, generative adversarial network; LSTM, long short-term memory; RGB, red, green, and blue.
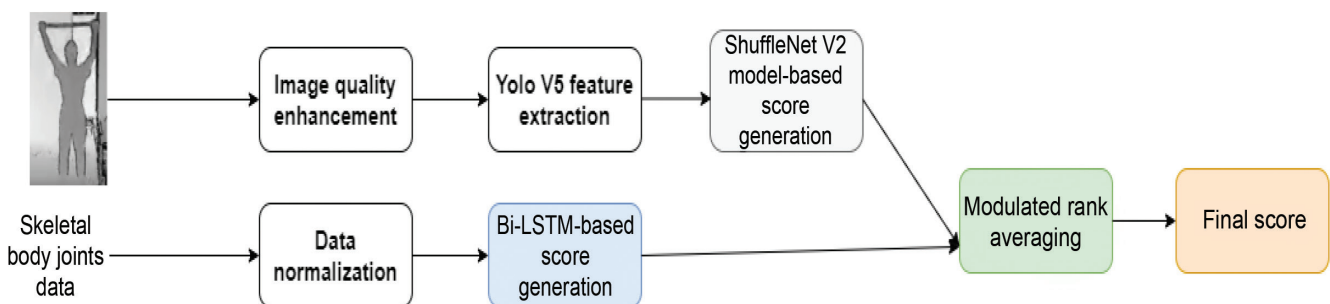
Figure 1: The proposed framework.

2021) fuses the outcomes of the YOLO V5–ShuffleNet V2 and Bi-LSTM models and presents a final score. Finally, the authors generalized the PAF using the KiMoRe dataset (KiMoRe dataset, 2023). They employed benchmark metrics that are used to evaluate the performance of the proposed framework. Figure 1 outlines the PAF.

joint orientations of 78 individuals. The dataset was collected from 44 healthy and 34 unhealthy individuals. A total of five exercises were conducted and respective images, videos, and features of exercises were recorded. Table 2 outlines the characteristics of the dataset. Figure 2 shows the sample images of exercises 4 and 5 of the KiMoRe dataset.

Data acquisition phase

The authors utilized the KiMoRe to train the PAF. The dataset is publicly available in the repository (KiMoRe dataset, 2023). It includes RGB images, videos, BJ positions, and

Preprocessing phase

The images were extracted from the videos. It may contain noises, which may affect the image quality. Thus, the authors apply the linear filtering technique for removing the noises.

Table 2: KiMoRe dataset characteristics.

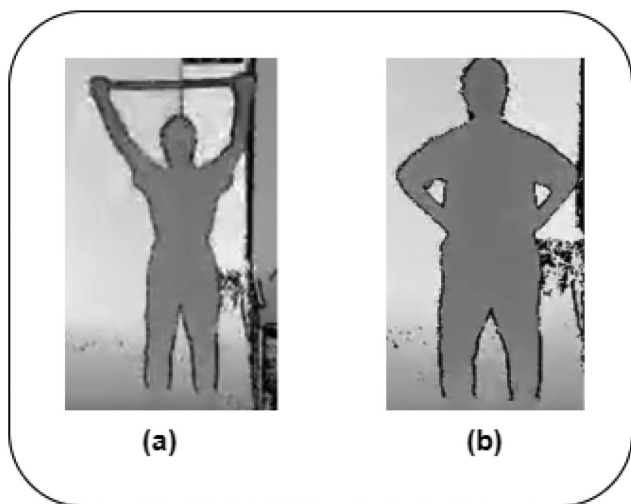| Exercises	Features	Number of individuals with standard deviation and average age
1	7	1. 44 healthy individuals, 36.5, and 35 years.
2	9	
3	9	
4	5	2. 34 unhealthy, 60.44, and 60 years
5	9	

Figure 2: (a) Exercise 5. (b) Exercise 4.

In addition, average and median filtering techniques are employed to overcome the salt and pepper noises. In addition, BJ data are normalized to support the Bi-LSTM model.

YOLO V5–ShuffleNet V2-based score generation

The YOLO V5 model utilizes a deep neural network as its foundation to extract hierarchical features from the images. Figure 3 offers the structure of the proposed model. These characteristics can be employed for the purpose of identifying and locating objects. The system utilizes transfer learning by leveraging pretrained weights from extensive datasets such as ImageNet. This enables the model to utilize acquired characteristics from a wide range of images prior to refining its performance object detection assignments. The system includes feature pyramid networks or comparable structures. This facilitates the capture of characteristics at several scales, enabling the model to effectively detect objects of varying sizes. The YOLO V5 model employs the residual and dense blocks for forwarding the information to the deepest layers to overcome the challenges in generating the feature maps. Let I be the image; b be the bounding box; a, d, h, and w are the locations of the bounding box; and σ be the image variation in time (t). Eqs. 1-4 express the computation of the bounding box.

$$b_a = (2X\sigma(t_a) - 0.5) + C_a \tag{1}$$

$$b_d = (2X\sigma(t_b) - 0.5) + C_b \tag{2}$$

$$b_w = P_w X\left((2X\sigma(t_w))^2\right) \tag{3}$$

$$b_h = P_h X\left((2X\sigma(t_h))^2\right) \tag{4}$$

where c is the fixed constant and P is the specified location in the bounding box. Using the YOLO V5 model, the features are extracted and corresponding feature maps are generated. The authors use the ShuffleNet V2 model for processing the feature maps and produce a score for the individual's performance in the specific exercise.

ShuffleNet V2 is a lightweight CNN model that demands fewer computing resources for classifying complex images. It employs a limited number of feature channels to reduce the number of floating point operations (FLOPs). A channel shuffle is used for communicating the information between the channel groups. Furthermore, the feature channels are divided into branches. Each branch includes three convolutions with similar input and output channels. Element-wise operations, including depth-wise convolution, channel shuffle, and channel split, are merged into a single element-wise operation. It identifies the BJ data and generates the score according to the variations. For instance, individuals' postures during the exercises are analyzed and variations are calculated based on the specific postures of the exercises. The authors employ the Adam optimizer to fine-tune the performance of the ShuffleNet V2 model. Based on the fine-tuned parameters, the additional rectified linear unit (ReLu), fully connected CNNs, and dropout layers are integrated into
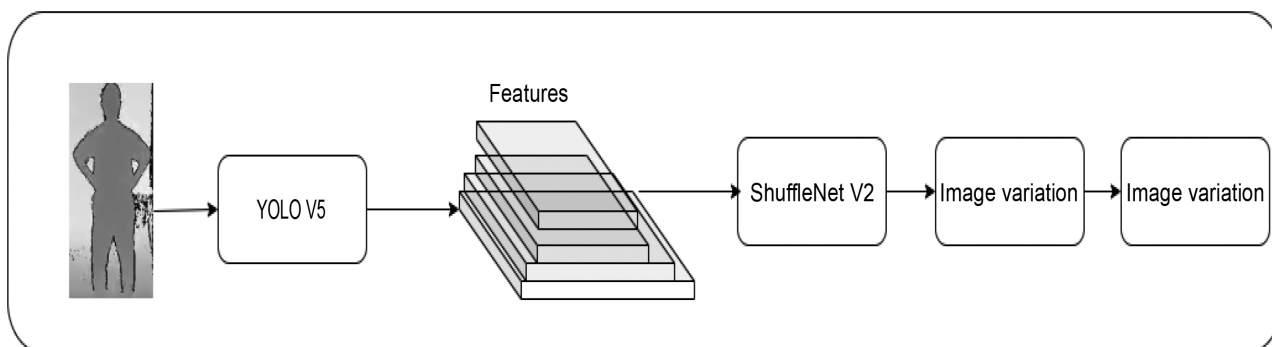

Figure 3: The structure of YOLO V5 and ShuffleNet V2 models.

the primary ShuffleNet V2 model. Figure 2 outlines the proposed feature extraction and image-based score generation.

Bi-LSTM-based score generation

Bi-LSTM is used to study the BJ data for computing the score. It offers the PAF to read backward and forward the information at each step. The hidden state preserves the information of the directions. The initial layer of the Bi-LSTM model is the vectorization layer. It encodes the BJ data into a sequence of tokens. The tokens are processed and assigned to a trainable vector. The vector can adjust themselves using the neighbor values. Lastly, Bi-LSTM processes the vectors and calculates an outcome for each exercise. Table 3 presents the notations and description of BJ.

Eqs. 5 and 6 show the computation of exercise 1.

$$E1_s = \sum_{i=1}^{3} JP_i + (\alpha/r) \tag{5}$$

where $E1_s$ is the side view of exercise 1.

$$E1_f = \sum_{i=1}^{13} JP_i + (\varnothing/r) + (\gamma/r) + (\psi/r) + d_a + d_h \tag{6}$$

where $E1_s$ is the frontal view of exercise 1.

The frontal views of exercise 2 are covered in Eqns. 7 and 8.

$$E2_{F1} = \sum_{i=1}^{4} JP_i + (\beta/r) \tag{7}$$

$$E2_{F2} = \sum_{i=1}^{13} JP_i + d_s + (\gamma/r) + d_{hip} + (Z_h/r, X_h/r) \\ + (\psi/r) + (\varnothing/r) \tag{8}$$

where $E2_{F1}$ and $E2_{F2}$ are the frontal views of exercise 2.

Eqs. 9 and 10 highlight the expression of the exercise 3 top and frontal views, respectively.

Table 3: KiMoRe dataset—notations and its description.

Notations	Description
(a/r)	Sagittal plane
(γ/r)	Elbow angles
$(\varnothing/r)$	Knee angles
(ψ/r)	Hip angles
(β/r)	Vertical axis of the hip and shoulders
A_t	Torso area
d_a	Ankle distance
d_h	Hands distances
d_k	Knee distance
d_{hip}	Hip distance
d_x	Horizontal distance between the elbows
(h/r)	Distance between wrists and shoulders
$(Z_h/r, X_h/r)$	Transverse plane coordinates of the hip
(η/r)	Shoulder extension angle
d_s	Shoulder distance
(Z_h/r)	Depth coordinates of the hip
(Z_s/r)	Depth coordinates of the shoulder
$(Z_s/r, X_s/r)$	Transverse plane coordinates of the hip
(d_s/r)	Distance between hand and shoulder
(x,z)	Transverse plan
JP	Joint positions

$$E3_T = \sum_{i=1}^{2} JP_i + d_x \tag{9}$$

$$E3_F = \sum_{i=1}^{13} JP_i + d_s + d_{hip} + d_h + (\gamma/r) + (\eta/r) + (h/r) \\ + (Z_h/r) + (\psi/r) + (\varnothing/r) \tag{10}$$

where $E3_T$ and $E3_F$ are the top and frontal views of exercise 3, respectively.

Eqs. 11 and 12 outline the combinations of angles and variations related to exercise 4.

$$E4_{F1} = JP_1 + x + z \tag{11}$$

$$E4_{F2} = \sum_{i=1}^{13} JP_i + (\gamma/r) + d_s + (Z_s/r) + d_{hip} + (\varnothing/r) \tag{12}$$

where $E4_{F1}$ and $E4_{F2}$ are the frontal views of exercise 4.

Finally, Eqs. 13 and 14 represent the expression of exercise 5.

$$E5_s = \sum_{i=1}^{6} JP_i + (\alpha/r) \tag{13}$$

$$E5_F = \sum_{i=1}^{13} JP_i + A_t + d_{hip} + d_k + d_s + d_a + (Z_s/r, X_s/r) + d_h \tag{14}$$

where $E5_s$ and $E5_F$ are the side and frontal views of exercise 5, respectively.

Score ranking model

In order to rank the scores of Bi-LSTM and YOLO V5–ShuffleNet V2 models, the authors follow the study. They apply the MRA method to fuse the scores and generate an outcome. The loss function outcome of Bi-LSTM and YOLO V5–ShuffleNet V2 models is used as a weight and combined with the generated scores. Based on these values, the accuracy is computed using the MRA method.

Evaluation metrics

In order to evaluate the performance of the PAF, the authors follow the benchmark evaluation metrics. They employ mean absolute deviation (MAD), mean absolute percentage error (MAPE), and root mean square error (RMSE) for the performance evaluation. MAD is used to derive the average distance between each BJ data with the mean. MAPE is the loss function that measures the prediction accuracy of the framework. RMSE is used to compute the difference between the predicted and ground truth values. In addition, the authors computed the accuracy using the MRA method. Eqs. 15-17 show the expression for computing MAD, MAPE, and RMSE.

$$MAD = \frac{1}{m} \sum_{i=1}^{m} |x - \hat{x}| \tag{15}$$

$$MAPE = \frac{1}{m} \sum_{i=1}^{m} \left| \frac{x - \hat{x}}{x} \right| \times 100 \tag{16}$$

Table 4: Computation strategies of the movement assessment framework.

Methods/strategies	Learning rate	Parameters (M)	FLOPs (G)	GPU speed (batches/sec)
PAF	1×10^{-4}	40	2.3	412
Deb et al.	1×10^{-3}	38	3.4	503
Sardari et al.	1×10^{-4}	25	4.7	586
Raihan et al.	1×10^{-3}	40	5.1	490
Kanade et al.	1×10^{-2}	42	4.8	535
Abedi et al.	2×10^{-4}	51	3.9	621

Abbreviations: FLOP, floating point operations; G, giga; GPU, graphics processing unit; M, millions; PAF, proposed assessment framework.

$$\text{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(x-\hat{x})^2} \qquad (17)$$

where m is the size of the dataset, and x and $\hat{x}$ are the real-time and predicted values, respectively.

RESULTS

The authors deployed the proposed model using Python 3.8.2, Nvidia Geforce rtx 3050 ti, Intel I7 processor, and 10 GB RAM. They used PyTorch and Tensorflow libraries for constructing the model. They trained the model using the batch size of 24 and Epochs of 242. Table 4 highlights the strategies of movement monitoring frameworks for generating the outcome. Compared to the state-of-the-art frameworks, the PAF required a limited learning rate and FLOPs. The ShuffleNet V2 model assisted the framework to identify the image variation and produce an optimal outcome with a limited computation cost.

Table 5 presents the performance of the PAF. The PAF generalization using the KiMoRe dataset highlighted its significance in assessing the rehabilitation exercises. Bi-LSTM and YOLO V5–ShuffleNet V2 models processed the key patterns of the images and streamlined the process of generating an outcome. Figure 4 reveals the findings of the performance analysis.

Table 6 reveals the findings of the comparative analysis. It is evident that the PAF outperforms the existing movements monitoring framework. The modified assessment framework (MAF) method presented a significant improvement in the PAF's performance. It employed the weights of Bi-LSTM and YOLO V5–ShuffleNet V2 models and generated the final score effectively. The lack of skeletal BJ data or images reduced the effectiveness of the existing frameworks. Figure 5 represents the outcome of the comparative analysis.

Table 5: Outcome of the performance analysis.

Exercises/metrics	MAD	MAPE	RMSE	Accuracy
Ex 1	0.482	1.112	1.026	86
Ex 2	0.521	1.205	1.102	91
Ex 3	0.389	1.098	1.206	78
Ex 4	0.478	0.981	1.054	89
Ex 5	0.489	1.108	0.996	91
Average	0.4718	1.1008	1.0768	87

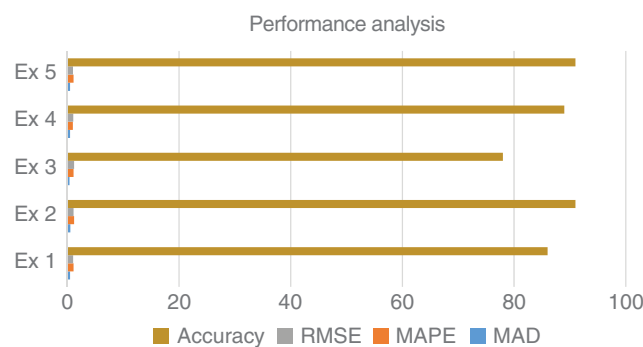Abbreviations: MAD, mean absolute deviation; MAPE, mean absolute percentage error; RMSE, root mean square error.

Figure 4: Performance analysis outcome.

Table 6: Outcome of the comparative analysis.

Frameworks/metrics	MAD	MAPE	RMSE	Accuracy (%)
PAF	0.4718	1.1008	1.0768	87
Deb et al.	0.581	1.305	1.2682	85
Sardari et al.	0.674	1.523	1.325	76
Raihan et al.	0.812	1.751	1.524	81
Kanade et al.	0.759	2.0123	1.458	77
Abedi et al.	0.536	1.268	1.362	82

Abbreviations: MAD, mean absolute deviation; MAPE, mean absolute percentage error; PAF, proposed assessment framework; RMSE, root mean square error.
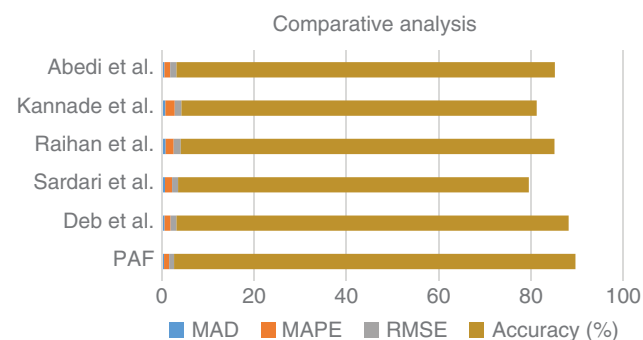

Figure 5: Comparative analysis outcome.

DISCUSSIONS

The proposed study presented an AI-based framework for monitoring the rehabilitation processes. It generates a score for each exercise using the skeleton's position data and images. The authors employed YOLO V5–ShuffleNet V2

models and the Bi-LSTM model to build the proposed framework. The KiMoRe dataset was utilized to evaluate the performance of the PAF. The dataset covers videos and images of five exercises. It provides stroke patient (SP) data for each exercise. Initially, the image quality is enhanced using the contrast-limited adaptive histogram equalization technique. YOLO V5 is used to extract features from images. The extracted features are processed using ShuffleNet V2 to generate scores using the skeleton joints. In addition, the skeleton's positions are used by Bi-LSTM to generate a score. Finally, the MRA method is used to compare the scores of Bi-LSTM and ShuffleNet V2 models using the weights and their scores to deliver the final score.

The PAF has major implications and benefits for patients, healthcare professionals, and the healthcare system. Personalizing treatment programs corresponding to each patient's unique requirements maximizes the efficacy of therapies while decreasing the time patients engage in rehabilitation. Using the suggested model makes it possible to promptly identify any problems, complications, or deviations from the planned path of recovery. This allows for appropriate interventions and minimizes the likelihood of setbacks. Integrating the model can improve patient involvement, resulting in increased adherence rates and improved rehabilitation outcomes. In addition, the proposed model provides healthcare clinicians with vital insights, facilitating evidence-based decision-making, individualized modifications to treatment plans, and enhanced overall patient care.

The PAF addressed the study's intention to develop an automated movement assessment framework. It offers an opportunity for clinicians and physiotherapists to provide effective services for disabled individuals. The recent developments in Internet of Things (IoT) devices can support the PAF in presenting an effective home-based rehabilitation monitoring environment. IoT-based cameras can be integrated with the PAF to supervise the individuals' movements. The lightweight application can be deployed in remote locations and assist physiotherapists in rendering services to aged and unhealthy individuals.

The experimental outcome revealed that the proposed framework outperformed the existing AI-based assessment framework. The model by Sardari et al. (2020) obtained a remarkable outcome using the images. However, the skeleton's position data can be applied to support disabled individuals in a real-time environment. The proposed framework integrated images and SP data for assessing the individual's movement. Raihan et al. (2021) extracted features from the SP data using LBP. The GA-based CNN demands additional training time for generating the score. Similarly, Deb et al. (2022) employed a graph convolutional network to evaluate the exercises. It obtained a MAD, MAPE, and RMSE of 0.581, 1.305, and 1.2682. In contrast, the PAF delivered a superior outcome. The frameworks of Kanade et al. (2022) and Abedi et al. (2023) required a high computation cost for generating the outcome. On the other hand, the PAF achieved a superior result with a limited computation cost.

The PAF contains some limitations. The authors validated the study using the KiMoRe dataset. The irregularities of the dataset may reduce the effectiveness of the PAF. In addition,

the structure of the ShuffleNet V2 model may affect the PAF's performance in the real-time environment, dependency on the availability of high-quality and diverse datasets for training the proposed model, challenges in precisely monitoring and evaluating the effectiveness of intricate rehabilitation exercises using remote monitoring, the dynamic and changing regulatory environment, and the lack of assurance in the approval processes for healthcare solutions driven by AI technologies. In the future, the authors will address the limitations of the PAF. They will focus on creating a large dataset using IoT and Kinect sensors. In addition, the muscle activity measurement will be considered for generating the score, integrating AI-driven monitoring, interactive activities, teleconsultations, and secure communication channels into comprehensive telerehabilitation platforms for seamless remote medical care. Edge computing is becoming more popular for processing AI algorithms in real-time, improving distant monitoring systems' responsiveness and decreasing latency. The fusion of AI algorithms with robotic rehabilitation devices enables the development of intelligent and adaptable systems that can offer dynamic and individualized support during physical therapy. Utilizing blockchain technology can strengthen data security, privacy, and integrity by addressing apprehensions related to preserving and exchanging confidential patient data.

CONCLUSION

In this study, the authors addressed the limitations of the existing movement monitoring framework by developing an integrated framework. They implemented the PAF by using Bi-LSTM and YOLO V5–ShuffleNet V2 models. The images of healthy and unhealthy individuals during the rehabilitation exercises and the skeletal BJ data were used to train the framework. The low-quality images were improved using the image preprocessing model. The YOLO V5–ShuffleNet V2 model generated a score based on the images. In addition, Bi-LSTM produced a score using the skeletal BJ data. Finally, the MRA method fused the scores and produced the final score for each rehabilitation exercise. The authors generalized the PAF using the KiMoRe dataset. The performance analysis outcome suggested that the PAF generated an optimal outcome for five exercises. Furthermore, the comparative analysis findings highlighted the significance of the PAF in supervising the rehabilitation exercises. The PAF yielded a superior outcome compared to the existing frameworks. It can be implemented in healthcare centers across the Kingdom of Saudi Arabia (KSA). It provides an effective environment for physiotherapists to treat disabled individuals. In addition, IoT cameras can be integrated with the PAF in order to present a home-based rehabilitation exercise monitoring system. The findings suggested that the proposed model facilitates the delivery of additional care beyond the designated treatment periods, consequently fostering continuous rehabilitation, mitigating the occurrence of relapses, and assisting individuals in effectively managing long-term diseases. The proposed model has the ability for scalability, enabling them to reach a broader population and deliver rehabilitation treatments to a more significant number of persons concurrently.

An AI-powered remote monitoring model enhances the robustness of healthcare systems, providing a practical solution for rehabilitation in times of pandemics or emergencies. However, the PAF demands additional training time for generating scores from complex real-time images. Future studies will consider improving the PAF's performance to deliver highly accurate results.

ACKNOWLEDGEMENTS

REFERENCES

Abedi A., Bisht P., Chatterjee R., Agrawal R., Sharma V., Jayagopi D.B., et al. (2023). Rehabilitation exercise repetition segmentation and counting using skeletal body joints. arXiv preprint arXiv:2304.09735.

Albert J., Glöckner P., Pfitzner B. and Arnrich B. (2021). Data augmentation of kinematic time-series from rehabilitation exercises using GANs. In: *Proceedings of the 2021 IEEE International Conference on Omni-Layer Intelligent Systems (COINS)*, IEEE, Barcelona, Spain, 23-25 August 2021; pp. 1-6.

Allahbakhsh M., Amintoosi H., Behkamal B., Kanhere S.S. and Bertino E. (2020). AQA: an adaptive quality assessment framework for online review systems. *IEEE Trans. Serv. Comput.*, 15(3), 1486-1497.

Bevilacqua A., Ciampi G., Argent R., Caulfield B. and Kechadi T. (2020). Combining real-time segmentation and classification of rehabilitation exercises with LSTM networks and pointwise boosting. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, New York, USA, 7-12 February 2020; (Vol. 34, No. 08); pp. 13229-13234.

Capecci M., Ceravolo M.G., Ferracuti F., Iarlori S., Monteriu A., Romeo L., et al. (2019). The KIMORE dataset: KInematic assessment of MOvement and clinical scores for remote monitoring of physical REhabilitation. *IEEE Trans. Neural. Syst. Rehabil. Eng.*, 27(7), 1436-1448.

Chowdhury S.H., Al Amin M., Rahman A.M., Amin M.A. and Ali A.A. (2021). Assessment of rehabilitation exercises from depth sensor data. In: *Proceedings of the 2021 24th International Conference on Computer and Information Technology (ICCIT)*, IEEE, Dhaka, Bangladesh, 8-20 December 2021; pp. 1-7.

De A. and Chowdhury A.S. (2021). DTI based Alzheimer's disease classification with rank modulated fusion of CNNs and random forest. *Expert Syst. Appl.*, 169, 114338.

Deb S., Islam M.F., Rahman S. and Rahman S. (2022). Graph convolutional networks for assessment of physical rehabilitation exercises. *IEEE Trans. Neural. Syst. Rehabil. Eng.*, 30, 410-419.

Derungs A., Schuster-Amft C. and Amft O. (2018). Physical activity comparison between body sides in hemiparetic patients using wearable motion sensors in free-living and therapy: a case series. *Front. Bioeng. Biotechnol.*, 6, 136.

Ebert A., Beck M.T., Mattausch A., Belzner L. and Linnhoff-Popien C. (2017). Qualitative assessment of recurrent human motion. In: *Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO)*, IEEE, Kos Island, Greece, 28 August-2 September 2017; pp. 306-310.

Ferreira B., Ferreira P.M., Pinheiro G., Figueiredo N., Carvalho F., Menezes P., et al. (2021). Deep learning approaches for workout repetition counting and validation. *Pattern Recognit. Lett.*, 151, 259-266.

Ge Z., Liu S., Wang F., Li Z. and Sun J. (2021). Yolox: exceeding Yolo series in 2021. arXiv preprint arXiv:2107.08430.

Guo Q. and Khan S.S. (2021). Exercise-specific feature extraction approach for assessing physical rehabilitation. In: *Proceedings of the 4th IJCAI Workshop on AI for Aging, Rehabilitation and Intelligent Assisted Living*. IJCAI, Montreal, Canada, 21-26 August 2021.

Huang Z., Xu W. and Yu K. (2015). Bidirectional LSTM-CRF models for sequence tagging. arXiv preprint arXiv:1508.01991.

Ishii S., Yokokubo A., Luimula M. and Lopez G. (2020). ExerSense: physical exercise recognition and counting algorithm from wearables robust to positioning. *Sensors*, 21(1), 91.

Kanade A., Sharma M. and Muniyandi M. (2022). A robust and scalable attention guided deep learning framework for movement quality assessment. arXiv preprint arXiv:2204.07840.

KiMoRe dataset. https://vrai.dii.univpm.it/content/KiMoRe-dataset, accessed on January 15, 2023.

Li C., Zhong Q., Xie D. and Pu S. (2018). Co-occurrence feature learning rom skeleton data for action recognition and detection with hierarchical aggregation. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, Stockholm, Sweden, 13-19 July 2018; pp. 786-792.

Li C., Shao M., Yang Q. and Xia S. (2023). High-precision skeleton-based human repetitive action counting. *IET Computer Vision*, 3, 700-709.

Liao Y., Vakanski A. and Xian M. (2020). A deep learning framework for assessing physical rehabilitation exercises. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 28(2), 468-477.

Lin J.F.S., Joukov V. and Kulić D. (2018). Classification-based segmentation for rehabilitation exercise monitoring. *J. Rehabil. Assist. Technol. Eng.*, 5, 2055668318761523.

Ma N., Zhang X., Zheng H.T. and Sun J. (2018). ShuffleNet v2: practical guidelines for efficient CNN architecture design. In: *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, 8-14 September 2018; pp. 116-131.

Maradani B. and Levkowitz H. (2017). The role of visualization in tele-rehabilitation: a case study. In: *Proceedings of the 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence*, IEEE, Setubal, Portugal, 24-26 April 2017; pp. 643-648.

Mottaghi E. and Akbarzadeh-T M.R. (2022). Automatic evaluation of motor rehabilitation exercises based on deep mixture density neural networks. *J. Biomed. Inform.*, 130, 104077.

Mourchid Y. and Slama R. (2023). D-STGCNT: a dense spatio-temporal graph Conv-GRU network based on transformer for assessment of patient physical rehabilitation. *Comput. Biol. Med.*, 165, 107420.

Nambi S.N.A.U., Gonzalez L. and Prasad RV. (2017). CoachMe: activity recognition using wearable devices for human augmentation. In: *Proceedings of the 2017 International Conference on Embedded Wireless Systems and Networks*, Uppsala, Sweden, 20-22 February 2017; pp. 174-179.

Petersen C.L., Wechsler E.V., Halter R.J., Boateng G.G., Proctor P.O., Kotz D.F., et al. (2018). Detection and monitoring of repetitions using an mHealth-enabled resistance band. In: *Proceedings of the 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies*, New York, USA, 26-28 September 2018; pp. 22-24.

Prabhu G., O'connor N.E. and Moran K. (2020). Recognition and repetition counting for local muscular endurance exercises in exercise-based rehabilitation: a comparative study using artificial intelligence models. *Sensors*, 20(17), 4791.

Rahman S., Sarker S., Haque A.N., Uttsha M.M., Islam M.F. and Deb S. (2022). AI-driven stroke rehabilitation systems and assessment: a systematic review. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 31, 192-207.

Raihan M.J., Ahad M.A.R. and Nahid A.A. (2021). Automated rehabilitation exercise assessment by genetic algorithm-optimized CNN. In: *Proceedings of the 2021 Joint 10th International Conference on Informatics, Electronics & Vision (ICIEV) and 2021 5th International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, IEEE, Kitakyushu, Japan, 16-21 August 2021; pp. 1-6.

Sardari F., Paiement A., Hannuna S. and Mirmehdi M. (2020). Vi-net—view-invariant quality of human movement assessment. *Sensors*, 20(18), 5258.

Sarsfield J., Brown D., Sherkat N., Langensiepen C., Lewis J., Taheri M., et al. (2019). Segmentation of exercise repetitions enabling real-time patient analysis and feedback using a single exemplar. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 27(5), 1004-1019.

Soro A., Brunner G., Tanner S. and Wattenhofer R. (2019). Recognition and repetition counting for complex physical exercises with deep learning. *Sensors*, 19(3), 714.

Vakanski A., Jun H., Paul D. and Baker R. (2018). A data set of human body movements for physical rehabilitation exercises. *Data*, 3(1), 2.

Zahra A., Hassan M.S., Park J.H., Hassan S.U.N. and Parveen N. (2022). Role of environmental quality of life in physical activity status of individuals with and without physical disabilities in Saudi Arabia. *Int. J. Environ. Res. Public Health*, 19(7), 4228.

Zhang D., Dai X. and Wang Y.-F. (2019). Dynamic temporal pyramid network: a closer look at multi-scale modeling for activity detection. In: *Computer Vision – ACCV 2018* (Jawahar C., Li H., Mori G. and Schindler K., eds.) pp. 712-728, Springer, Cham.

Zhao YS., Xiong Y., Wang L., Wu Z., Lin D. and Tang X. (2017). Temporal action detection with structured segment networks. In: *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 22-29 October 2017; pp. 2933-2942.