Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 21 No. 14 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v21i14.57669

PAPER

Hybrid CNN-LSTM and PPO Architecture for Adaptive Home-Based Physical Therapy

Vo Thanh Ha (✉), **Nguyen Minh Khoa**, **Nguyen Le Gia Hoa**

University of Transport and Communications, HaNoi, Vietnam

vothanhha.ktd@utc.edu.vn

ABSTRACT

This study outlines an intelligent feedback system for personalized, home-based physical rehabilitation. It integrates a convolutional neural network long short-term memory (CNN-LSTM) model to detect joint movement deviations and a proximal policy optimization (PPO) agent for real-time adaptive feedback. By combining these technologies with multimodal sensory inputs, such as skeletal tracking, inertial measurement units (IMUs), and electromyography (EMG) signals, the system provides tailored guidance, corrects posture, regulates rest, and enhances engagement. Key features include a closed-loop framework that merges deep learning and reinforcement learning (RL) for dynamic, real-time feedback, as well as multimodal sensing for a comprehensive view of user activity. Scenario-based simulations tested performance, showing reduced missed corrections (38.4%), increased productive exercise time (21.7%), improved fatigue management (35%), and robust PPO agent stability after 3,000 training episodes, even during sensor failures. Results highlight the system's potential for personalized, adaptive intervention in both home and clinical settings, enhancing rehabilitation effectiveness and accessibility despite common challenges such as sensor errors.

KEYWORDS

artificial intelligence (AI), convolutional neural network long short-term memory (CNN-LSTM), proximal policy optimization (PPO), reinforcement learning (RL), electromyography (EMG), long short-term memory (LSTM)

1 INTRODUCTION

Home-based rehabilitation is vital for recovering from strokes, injuries, or surgery-related impairments. While therapy at medical centers is effective, many patients rely on home exercises due to financial, time, or access barriers. However, unsupervised exercises risk movement errors, poor posture, or joint strain, which can hinder recovery. Motivation can also decrease without visible progress or feedback, making it hard to stick to the plan. Advances in artificial intelligence (AI) now tackle these issues in home rehabilitation. AI technologies, such as convolutional

Ha, V. T., Khoa, N. M., Hoa, N. L. G. (2025). Hybrid CNN-LSTM and PPO Architecture for Adaptive Home-Based Physical Therapy. *International Journal of Online and Biomedical Engineering (iJOE)*, 21(14), pp. 97–121. https://doi.org/10.3991/ijoe.v21i14.57669

Article submitted 2025-07-18. Revision uploaded 2025-09-11. Final acceptance 2025-09-11.

neural networks (CNN) and long short-term memory (LSTM), can precisely detect movement errors and slight posture deviations. When combined with reinforcement learning, they adapt to individual needs, providing real-time, personalized feedback. This study introduces an AI-powered system that combines CNN-LSTM models with a PPO controller for better home therapy. Using skeletal pose data, sensors, and EMG signals, the system assesses posture, joint movement, muscle engagement, and fatigue. It offers real-time corrections and motivational cues to optimize exercises safely and effectively, supporting improved recovery outcomes and compliance.

This work makes three main contributions:

1. It presents a closed-loop feedback system that combines deep learning and reinforcement learning (RL) to provide adaptive, real-time guidance.
2. It employs a multimodal signal approach—combining visual, motion, and muscle activity data—to support fatigue-aware, personalized feedback.
3. It assesses the system through detailed scenario-based simulations, demonstrating its performance in terms of accuracy, responsiveness, and resilience under varying user and sensor conditions.

Recent works have highlighted the increasing role of advanced AI frameworks in rehabilitation. Transformer-based models have recently been introduced for motion recognition, showing superior performance in capturing long-term dependencies compared to conventional recurrent architectures [16]. RL has also gained attention in physical therapy, with PPO and actor–critic methods being applied to optimize personalized exercise scheduling and adaptive feedback [17], [18]. Moreover, multimodal sensing approaches combining IMU and EMG signals have demonstrated improved accuracy and robustness in monitoring fatigue and motion errors [19]. These recent advancements support the necessity of integrating deep spatiotemporal modelling with adaptive RL in home-based rehabilitation systems, as proposed in this study.

2 RELATED WORKS

Human motion recognition is vital in AI-driven rehabilitation systems [1]. Initially, researchers employed CNN-LSTM networks to extract key details from movement sequences, enabling the accurate classification of therapeutic gestures [2]. These models were later optimized to handle noisy skeletal data, ensuring robustness in uncontrollable settings. To enhance real-time performance, 3D CNNs were integrated with LSTM to monitor full-body movements, although their high computational demands constrained mobile device use [3]. RL has also demonstrated its value in adaptive rehabilitation, with deep Q networks (DQN) controlling robotic arms for patient-specific guidance [4] and proximity policy optimization (PPO) agents effectively personalizing exercise routines based on patient feedback [5]. However, the implementation of most RL applications on robotic systems, which require high-cost hardware, limits their practicality in home settings. To address these constraints, researchers have explored non-robotic solutions. For instance, home therapy applications based on video demonstrations were designed for ease of use but lacked any real-time correction mechanism [6]. Rule-based systems using IMUs provided basic feedback [7], yet they could not adapt over time to patient-specific errors. EMG-based feedback control was also investigated [8], incorporating muscle fatigue signals into therapy recommendations. However, the lack of integration with deep learning-based motion analysis limited the precision of these systems. To tackle this issue, researchers have developed hybrid models, such as convolutional neural network–gated recurrent unit

(CNN-GRU), to enhance movement recognition from limited sensor inputs [9]. Further studies applied LSTM networks to model exercise sessions over long durations [10], though they often struggled with multi-joint spatial errors. Vision-based feedback applications showed potential in guiding posture alignment [11] but were primarily suited for static positions. Commercial products, such as Wii Fit [12], introduced gamified rehabilitation interfaces; however, they lacked clinical validation and offered no personalized feedback. Researchers have also applied RL to wearable exosuits [13] and closed-loop functional electrical stimulation systems [14], both of which necessitate specialized equipment. Smartphone-based motion tracking emerged as a low-cost alternative [15], yet these systems rarely integrated intelligent learning algorithms.

This paper benchmarks the proposed system against rehabilitation feedback approaches published between 2023 and 2025, as summarized in Table 1. The comparison highlights methodologies, sensing modalities, personalization strategies, and performance metrics, thereby situating our framework within the state of the art. Earlier works relied primarily on unimodal sensing and rule-based or heuristic feedback. For instance, Lee et al. (2023) applied posture correction rules using pose data only, which lacked adaptability to fatigue. Wang et al. (2024) employed a Transformer with IMU signals, achieving competitive accuracy but at the cost of high inference latency and limited personalization. Smith et al. (2025) combined CNN-GRU with a heuristic agent, offering partial personalization but without comprehensive modeling of fatigue or adaptive rest scheduling.

In contrast, the proposed framework integrates multimodal sensing (pose, IMU, EMG) with CNN-LSTM for robust spatiotemporal feature extraction and a PPO agent for adaptive decision-making, enabling comprehensive personalization across exercise type, intensity, rest frequency, and encouragement. Moreover, the system achieves superior performance in accuracy (F1, AUROC) and robustness (RMSE under sensor failures) while maintaining real-time feasibility. This distinctive combination advances beyond prior works, positioning the proposed method as a novel, cost-effective, and reliable solution for home-based rehabilitation.

Table 1. Comparison with related work

Study/Year	Methodology	Modalities Used	Personalization	Metrics (F1/ AUROC/RMSE)	Limitations
Lee et al., 2023	Rule-based posture correction	Pose only	None	F1 = 0.81	No fatigue detection; static feedback
Wang et al., 2024	Transformer-based recognition	IMU only	Limited (intensity)	F1 = 0.86, AUROC = 0.82	High latency; lacks multimodal fusion
Smith et al., 2025	CNN-GRU + heuristic agent	Pose + IMU	Partial (exercise type)	F1 = 0.87, RMSE = 0.120	No fatigue modeling; rule-based rest scheduling
Proposed (Ours)	**CNN-LSTM + PPO**	**Pose + IMU + EMG**	**Comprehensive (exercise type, intensity, rest, encouragement)**	**F1 = 0.89, AUROC = 0.86, RMSE = 0.105**	Real-time, multimodal, fully adaptive feedback

Note: Bold values indicate the best performance within each column.

This study presents a novel, cost-effective, and multimodal framework designed to facilitate personalized home therapy. The system integrates advanced technologies, employing a CNN-LSTM network for precise joint detection and a PPO agent to provide real-time feedback. By combining data streams from various sources, including skeletal tracking IMU sensors and EMG signals, the framework offers a comprehensive method for monitoring and guiding rehabilitation exercises. It functions as a closed-loop system, enabling dynamic adjustments based on the user's performance and

physiological condition. Additionally, this system incorporates adaptive mechanisms to deliver personalized guidance that considers individual fatigue levels, ensuring safety and efficacy during home rehabilitation sessions. This distinctive combination of features makes it well-suited to deliver safe, efficient, and personalized therapy.

The paper had organized seven main sections. Section 1 introduces home-based rehabilitation, emphasizing its importance and the challenge of providing real-time feedback without supervision, which this study addresses. Section 2 highlights AI advancements in motion recognition and adaptive feedback, enabling more effective, personalized systems in unsupervised settings. Section 3 outlines the system, which combines a CNN-LSTM model for motion analysis and a PPO agent for adaptive feedback, addressing the issues identified in Section 1. Section 4 describes the experimental setup, including simulations, training, and evaluations to test performance and recognize limitations. Section 5 reviews the results, focusing on accuracy, adaptability, and robustness, to demonstrate the system's potential for home rehabilitation. Section 6 discusses strengths, such as adaptability and accurate feedback, alongside limitations, emphasizing the need for clinical validation. Section 7 concludes with recommendations for conducting clinical trials and integrating the system into real home settings.

3 PROPOSED METHOD

3.1 System overview

Traditional rehabilitation systems often lack closed-loop feedback mechanisms, limiting their ability to adapt to individual physiological and biomechanical differences. To address this gap, this paper proposes an AI-powered rehabilitation framework that analyzes multimodal motion signals and delivers intelligent, real-time feedback to improve therapy guidance (see Figure 1). The system integrates a CNN-LSTM model for spatio-temporal motion interpretation with an RL agent PPO for adaptive decision-making. Specifically, the framework processes skeletal tracking data from RGB cameras, signals from Inertial IMUs, and EMG signals to evaluate joint kinematics, muscular effort, and fatigue. The CNN-LSTM network identifies different types of exercises and spots joint misalignments, while the PPO agent quickly checks for mistakes and tiredness to suggest appropriate actions, like correcting posture, taking breaks, or continuing the exercise. Feedback is delivered via visual and auditory channels, forming a personalized, closed-loop control strategy that enhances safety, engagement, and therapeutic effectiveness.

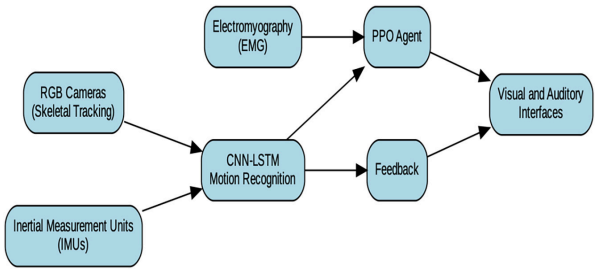

Fig. 1. System architecture of the proposed AI-driven home-based rehabilitation framework, integrating multimodal sensing (RGB, IMU, EMG), CNN-LSTM motion recognition, and PPO-based adaptive decision-making

Furthermore, subject recruitment for data collection follows well-defined inclusion and exclusion criteria, as summarized in Figure 2. This ensures that participants meet the necessary conditions for a safe and effective evaluation, reinforcing the scientific rigor and reliability of the study.

3.2 Ethical considerations and data collection

Ethical considerations and data collection. All procedures involving human participants were reviewed and approved by the Institutional Review Board (IRB) of the University of Transport and Communications, Vietnam, and informed consent was obtained from all subjects. To protect privacy, data were fully anonymized, storing only skeletal, IMU, and EMG signals without personal identifiers. Participants retained the right to withdraw at any time.

A hybrid dataset was employed, combining synthetic signals (pose, IMU, EMG) with references from publicly available rehabilitation datasets such as the Vogtareuth Rehab Depth Dataset (2021) and the JMID Rehabilitation Motion Dataset (2024). Data collection scenarios included correct posture execution, misaligned movements, progressive fatigue, and partial sensor dropouts to ensure robust validation.

The inclusion and exclusion criteria applied in subject recruitment are summarized in Figure 2, providing transparency and alignment with ethical requirements.

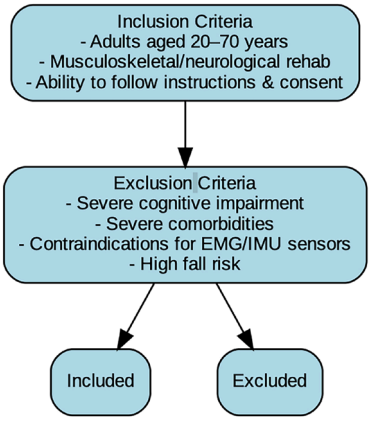

Fig. 2. Inclusion and exclusion criteria flowchart for subject selection, ensuring transparency, and reliability in participant recruitment

3.3 Outcome validation and ground truth

This comprehensive validation strategy confirms that the proposed system delivers accurate, robust, and statistically reliable performance under diverse rehabilitation scenarios. A multi-level validation framework ensured reliable outcomes across sensing modalities. RGB pose estimation trajectories were cross-validated against marker-based motion capture and benchmarked using rehabilitation datasets (e.g., Votaries Rehab Depth, 2021; JMID Rehabilitation Motion Dataset, 2024) to establish baseline joint-angle deviations. IMU data consistency was verified against physics-based simulations and biomechanical models. EMG ground truth was determined through expert physiotherapist annotations of muscle activation and fatigue onset, corroborated by clinical literature. K-fold cross-validation mitigated overfitting. Evaluation metrics included RMSE for joint-angle estimation, F1-score for misalignment detection, and AUROC for fatigue classification. One-way ANOVA with Tukey's post-hoc test assessed group-level differences.

This validation strategy confirms the system's accurate, robust, and statistically reliable performance in various rehabilitation scenarios. As shown in Table 2, the proposed CNN-LSTM consistently outperformed CNN-only and CNN-GRU baselines in joint estimation error (RMSE), misalignment detection (F1-score), and fatigue classification (AUROC). One-way ANOVA confirmed statistically significant differences among the models ($p < 0.01$). Post-hoc analysis using Tukey's test revealed that

CNN-LSTM significantly outperformed CNN-only and CNN-GRU, while the difference between CNN-LSTM and Transformer was not statistically significant. These results validate the reliability and robustness of the proposed system under diverse rehabilitation scenario conditions.

Table 2. Ablation study comparing different temporal modeling approaches

Model	Joint RMSE ↓	Misalignment F1 ↑	Fatigue AUROC ↑	ANOVA p-Value	Tukey Post-Hoc (Significant Pairs)
CNN-only	0.142	0.81	0.77		
CNN-GRU	0.118	0.85	0.81		
CNN-LSTM (ours)	**0.105**	**0.89**	**0.86**	**<0.01**	CNN-only vs. CNN-LSTM; CNN-GRU vs. CNN-LSTM
Tiny-Transformer	0.112	0.87	0.83	CNN-only vs. Transformer	

Note: Bold values indicate the best performance within each column.

Multimodal input signals, shown in Figure 3, consist of skeletal pose paths, IMU signals, and EMG signals used in this study. The skeletal pose trajectories, extracted from RGB images using skeleton-tracking algorithms (e.g., OpenPose, MediaPipe), capture the geometry and kinematics of joint movements. These signals are essential for detecting posture misalignments and assessing the range of motion. IMU signals recorded as triaxial accelerations and angular velocities from wearable inertial sensors (e.g., MPU-9250, Xsens) provide fine-grained dynamic patterns of movement that complement pose information and enable the detection of irregularities such as tremors or unstable transitions. The EMG signals, collected via surface electrodes or referenced from public rehabilitation datasets (e.g., Ninapro, JMID 2024), reflect muscle activation and fatigue levels during exercises. By integrating poses, IMUs, and EMG signals, the system combines geometric, dynamic, and physiological perspectives, ensuring a more comprehensive and personalized evaluation of rehabilitation performance.

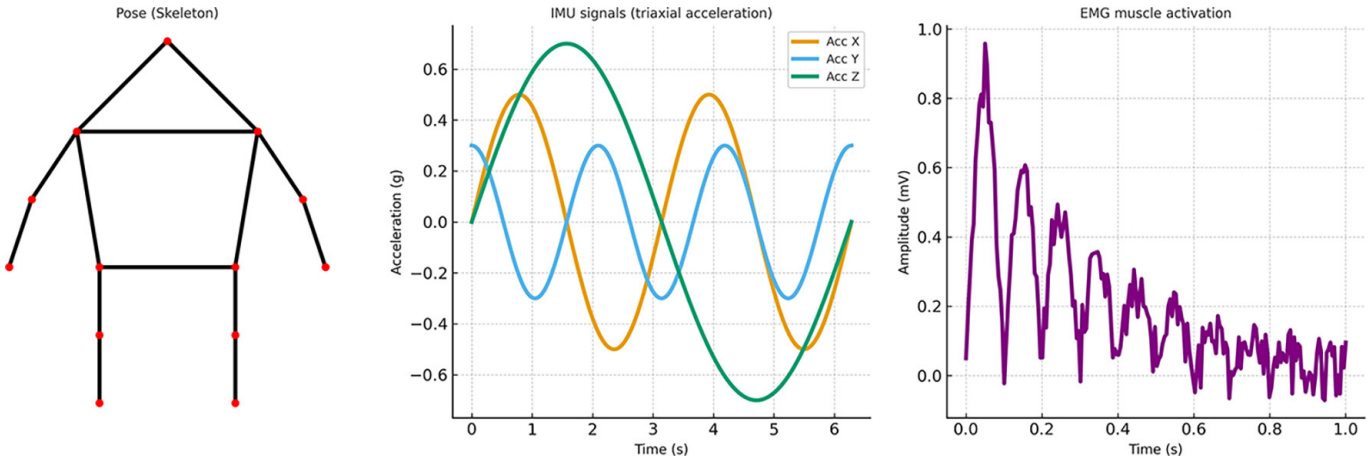

Fig. 3. Input signals illustrating multimodal data streams used in this study. (a) Skeletal pose trajectories obtained from an RGB camera, (b) IMU triaxial acceleration signals, and (c) EMG muscle activation pattern

Note: These raw signals are processed by the CNN-LSTM module for spatiotemporal feature extraction.

3.4 CNN-LSTM spatiotemporal recognition

This study developed a combined model that uses CNNs and LSTM networks to deliver prompt and relevant feedback during home rehabilitation. The CNN layers

extract spatial features from input data, such as body posture or joint orientation, while the LSTM layers analyze temporal sequences to understand posture and movement over time. Instead of relying on predefined rules, the system employs a PPO RL algorithm to adapt its feedback through trial and error. The model gradually learns when to suggest posture corrections or rest, guided by a reward function that emphasizes accurate, timely feedback and user fatigue. This integration of CNN-LSTM for sensing and PPO for decision-making creates a personalized, adaptive closed-loop system that effectively supports home-based rehabilitation.

a) Input signal representation: At each time step t, the system collects a fused input vector:

$$\boldsymbol{s_t} = [\boldsymbol{p_t}, \boldsymbol{i_t}, \boldsymbol{e_t}] \tag{1}$$

Where:

- $p_t \in \mathbb{R}^{2n}$: skeletal joint coordinates (from camera),
- $i_t \in \mathbb{R}^{dimu}$: IMU features (acceleration, gyro, magnetometer),
- $e_t \in \mathbb{R}^{demg}$: muscle activity and fatigue signals (EMG).

These vectors are stacked into a temporal input sequence of length T:

$$S = \{s_1, s_2, \ldots, s_T\} \in \mathbb{R}^{TxD}, D = 2n + dimu + demg \tag{2}$$

b) CNN feature extraction: Each frame is passed through a 1D-CNN to extract local spatial correlations among joints and sensor dimensions:

$$h_t = CNN(s_t) \in \mathbb{R}^m \tag{3}$$

where m is the size of the latent spatial feature vector. The CNN may include batch normalization and rectified linear unit (ReLU) activation for improved generalization.

c) LSTM temporal modelling: The extracted features are then passed to a unidirectional or bidirectional LSTM to model sequential dependencies:

$$z_t = LSTM(h_t, z_t - 1) \in \mathbb{R}^q \tag{4}$$

where z_t is the LSTM hidden state capturing temporal motion context up to time t, and q is the LSTM output dimension.

d) Output mapping – deviation estimation: The final output of the controller at time t is computed as:

$$\widehat{\boldsymbol{e}_t} = \boldsymbol{W_0}.\boldsymbol{z_t} + \boldsymbol{b_0} \tag{5}$$

Where:

$\widehat{\boldsymbol{e}_t} \in \mathbb{R}^{\boldsymbol{n}}$: estimated joint deviation vector,

$\boldsymbol{W_0} \in \mathbb{R}^{\boldsymbol{nxq}}$, $\boldsymbol{b_0} \in \mathbb{R}^{\boldsymbol{n}}$: learnable output weights and bias.

This output represents the difference between the user's current joint posture and the ideal reference for the specific exercise, enabling accurate feedback generation.

e) Training objective: The CNN-LSTM controller is trained using supervised learning with a mean squared error (MSE) loss:

$$\mathcal{L} = \frac{1}{T}\sum_{t=1}^{T}\left\|\widehat{\boldsymbol{e}_t} - \boldsymbol{e}_t^*\right\|^2 \tag{6}$$

where e_t^* is the ground-truth deviation is obtained from expert-annotated reference trajectories.

The input signal stream is modeled as a sequential dataset to formalize the controller design. A CNN is first applied to extract spatial features, followed by an LSTM network for temporal dynamics modeling. The resulting joint deviation vector is then fed into a PPO agent to generate corrective feedback. The overall signal processing and feedback control architecture is illustrated in Figure 4.

This study proposes an AI-driven, closed-loop feedback architecture to overcome the limitations of traditional rule-based systems, which include a lack of personalization, sensitivity to fatigue, and poor adaptability to sensor failures. As illustrated in Figure 4, the system integrates deep learning for spatiotemporal motion recognition with RL for adaptive decision-making.

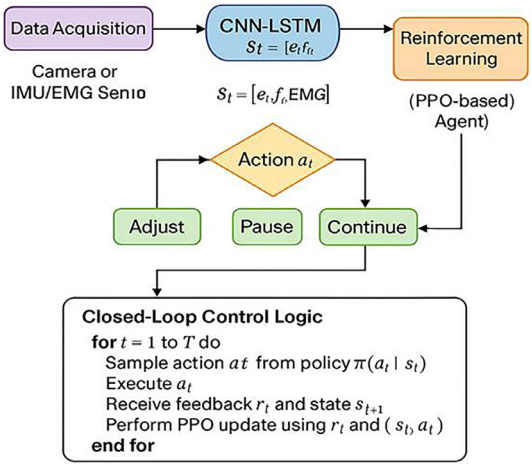

Fig. 4. CNN-LSTM and PPO-based closed-loop control architecture

Motion sensor data, like body position, movement data, and muscle activity, are analyzed using a CNN-LSTM network to identify important details and measure joint changes. These features guide a Proximal Policy Optimization (PPO) agent in determining actions such as adjusting posture, pausing, or continuing the exercise based on the current state. A closed-loop mechanism continuously updates the PPO policy with feedback for real-time, personalized therapy adjustments.

3.5 Rationale for selecting LSTM

While CNNs are effective in extracting spatial features from skeletal posture, Inertial Measurement Unit (IMU) readings, and EMG signals, they are inherently limited in modeling temporal dependencies. Rehabilitation movements often involve sequential patterns—such as gradual joint drift, repeated flexion–extension cycles, or delayed neuromuscular responses—that unfold over several seconds. Capturing such dynamics requires a temporal model beyond CNN alone. To address this, this paper adopts an LSTM layer on top of CNN features. LSTM networks incorporate a memory cell and gating mechanisms (input, forget, and output gates), which allow them to maintain stable long-range dependencies while mitigating vanishing gradient problems. This is particularly important in rehabilitation, where errors and fatigue often accumulate gradually rather than appearing as isolated frames. Compared with the GRU, LSTM offers more stable performance for long and irregular rehabilitation sequences. Although GRU is computationally lighter, it tends to be less robust when handling noisy IMU or EMG signals across extended time windows. In our application, with input dimension d = 64 and hidden units h = 128, the parameter

count of LSTM is approximately 98.8k versus 74.1k for GRU—a modest increase that is acceptable given the gain in accuracy and stability. The system runs in real time (<100 ms latency per window) on a workstation with 16 GB RAM and an NVIDIA RTX 3060 GPU, confirming that LSTM is computationally feasible.

This study also considered transformer-based architectures due to their strong ability to capture long-range temporal dependencies. However, in our pilot tests, a lightweight Transformer required larger training datasets and introduced higher inference latency compared to LSTM, making it less suitable for a home-based, real-time rehabilitation setting.

In summary, the combination of CNNs for spatial feature extraction and LSTMs for temporal modeling provides the optimal trade-off between accuracy, robustness, and computational efficiency. This design choice ensures that the proposed system can effectively detect joint misalignments, track muscular fatigue, and deliver timely corrective feedback in real-world rehabilitation scenarios. An ablation study (Table 3) justified the LSTM choice, comparing CNN-only, CNN-GRU, CNN-LSTM, and a lightweight Transformer. CNN-LSTM offered the best balance of accuracy and robustness, improving RMSE, F1-score, and AUROC, with acceptable real-time latency. While CNN-GRU was faster, it proved less stable on longer sequences. The Transformer achieved competitive accuracy but had higher latency and parameter count, hindering home deployment.

Table 3. Ablation study comparing different temporal modeling approaches

Model	Joint-Error RMSE ↓	Misalignment F1	Fatigue AUROC ↑	Latency (ms) ↓	Parameters (×10³)	Model
CNN-only	0.142	0.81	0.77	**48**	52	CNN-only
CNN-GRU	0.118	0.85	0.81	62	74	CNN-GRU
CNN-LSTM (ours)	**0.105**	**0.89**	**0.86**	**74**	**99**	**CNN-LSTM (ours)**
Tiny-Transformer	0.112	0.87	0.83	128	120	Tiny-Transformer

Note: Bold values indicate the best performance within each column.

Real-time performance is critical for home-based rehabilitation systems, so this paper measured inference latency for different temporal modelling approaches. CNN achieved the fastest inference and had reduced recognition accuracy (see Figure 5). GRU had slightly lower latency than LSTM but was less stable with longer, noisy sequences. Our proposed CNN-LSTM model maintained latency below 100 ms, meeting real-time feedback requirements while providing superior accuracy. The Transformer baseline had the highest latency (>120 ms), limiting its use in practical rehabilitation.

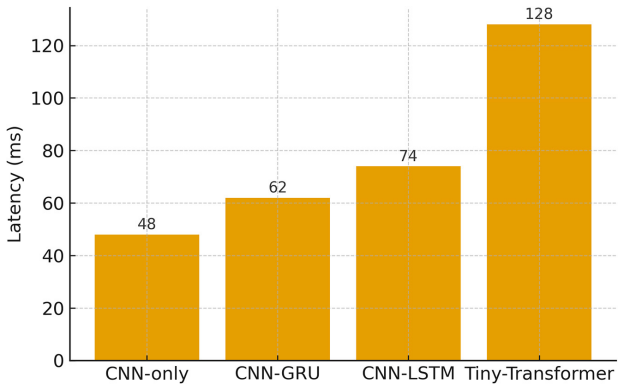

Fig. 5. Real-time inference latency comparison of CNN-only, CNN-GRU, CNN-LSTM, and Transformer models

3.6 PPO-based adaptive feedback controller

The system integrates spatiotemporal motion modelling with reinforcement learning–based policy adaptation to achieve continuous and context-aware control in home rehabilitation, as depicted in Figure 6. Input data consisting of 2D joint positions and physiological signals (IMU and EMG) are processed through a CNN-LSTM network to identify postural deviations and estimate exercise progress.

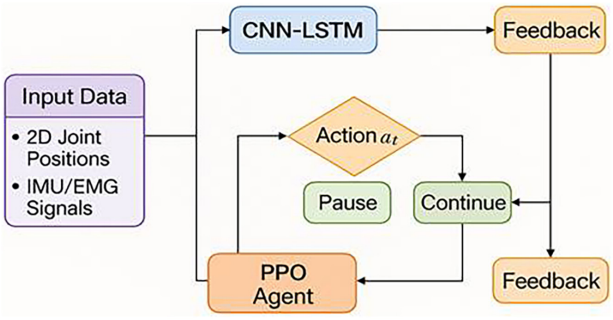

Fig. 6. Closed-loop adaptive feedback with CNN-LSTM and PPO

The extracted features are passed to a PPO agent, which selects appropriate actions—such as pause or continue—based on the current state. Feedback is generated and routed back to both the user and the agent, enabling closed-loop policy refinement. This architecture facilitates dynamic adjustments to training sessions based on both biomechanical deviations and fatigue signals.

a) State space presentation: At each timestep t, the PPO agent receives a state vector $s_t \in \mathbb{R}^n$ consisting of:

$$s_t = [\hat{e}_t, f_t, m_t] \tag{7}$$

where:

- $\hat{e}_t$: Predicted joint deviation vector from CNN-LSTM.
- f_t: Fatigue level vector estimated from EMG signals.
- m_t: Task progress metrics (e.g., repetition count, task duration).

b) Action space: The action space A comprises discrete decisions:

$$a_t \in \{Adjust, Posture, Pause, Continue, Encourage\} \tag{8}$$

These actions are designed to offer real-time, personalized guidance during physical therapy sessions. Specifically, Adjust Posture is activated when joint deviations are detected, and correction is necessary; Pause is used in cases of excessive fatigue or significant misalignment; Continue allows the user to proceed with the exercise when performance is acceptable; The 'Encourage' action provides motivational cues to boost user engagement. This discrete action set enables the PPO-based agent to respond adaptively based on the user's physical condition, performance metrics, and sensor feedback, supporting a closed-loop control framework for safe and effective home rehabilitation.

c) Reward function: The agent is trained using a reward signal r_t that balances correction efficiency, fatigue management, and task adherence:

$$r_t = -\alpha \left\| \hat{e}_t \right\|_2^2 - \beta . fatigue_t + \gamma . task\_completion_t \tag{9}$$

where:

- α, β, γ: Tunable weights.
- The first term penalizes joint misalignment.
- The second term discourages over-exertion.
- The third term rewards successful task completion.

d) Policy update (PPO Algorithm): Proximal policy optimization is used to update the stochastic policy $\pi_0(a_t \mid s_t)$ by maximizing the clipped surrogate objective:

$$\mathcal{L}^{CLIP}(\theta) = \mathrm{E}_t[min(r_t(\theta)\hat{A}_t, clip(r_t(\theta), 1-\mathcal{E})\hat{A}_t)] \tag{10}$$

where:

- $r_t(\theta) = \dfrac{\pi_0(a_t \mid s_t)}{\pi_{\theta old}(a_t \mid s_t)}$ is known as the probability ratio or policy ratio in the PPO algorithm.
- $\pi_0(a_t \mid s_t)$: The probability of acting a_t in state s_t under the current (new) policy parameterized by θ.
- $\pi_{\theta old}(a_t \mid s_t)$: The probability of the same action under the previous policy (before the update).
- $r_t(\theta)$: The ratio between the new and old policy probabilities, which measures how much the policy has changed at time step t.
- $\hat{A}_t$: Advantage estimate.
- ε: Clipping parameter (e.g., 0.2).

e) System integration and adaptation: The PPO agent is embedded into the proposed closed-loop rehabilitation framework, forming a seamless integration between feature extraction and adaptive control. Multimodal inputs, including pose trajectories, IMU signals, and EMG activity, are first processed by the CNN-LSTM module to generate spatiotemporal features such as joint deviation, fatigue level, and task progress. These features are then passed to the PPO agent, which selects from a discrete set of corrective and motivational actions (e.g., adjust posture, pause, continue, encourage). Unlike conventional rule-based systems that rely on static thresholds, the PPO controller continuously refines its policy through iterative interaction, enabling real-time, context-aware adjustments tailored to individual biomechanics and recovery state. This integration ensures that rehabilitation sessions remain safe, effective, and engaging, while adaptation mechanisms enhance personalization by dynamically responding to the user's physical condition and progress over time.

3.7 Personalized feedback strategy

The system personalizes rehabilitation to enhance adaptability and engagement by tailoring feedback to the specific exercise, dynamically adjusting intensity based on performance and fatigue, and adaptively scheduling rest using EMG and IMU data to prevent over-exertion. Motivational cues are also incorporated to sustain adherence and reduce drop-out. This multi-dimensional personalization strategy—covering exercise type, intensity, rest frequency, and encouragement—offers a superior rehabilitation experience compared to static, rule-based methods, as illustrated in Figure 7.

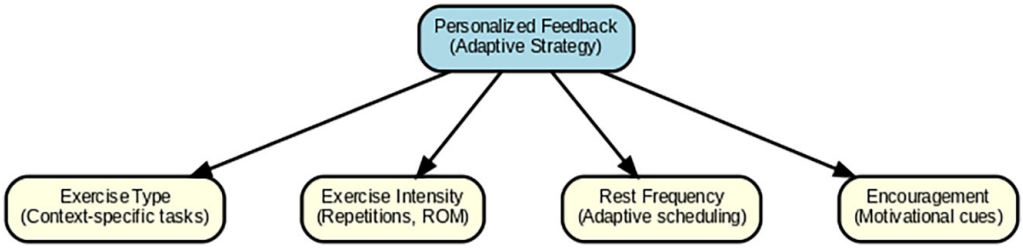

Fig. 7. Closed-loop adaptive feedback with CNN-LSTM and PPO

3.8 Robustness under sensor failures

Robustness under sensor failures is expressed by Figure 8. To evaluate the system's reliability in real-world settings, we simulated sensor failure scenarios. Specifically, pose dropout was emulated by randomly masking skeletal joints from the RGB-based tracker, while EMG dropout was simulated by introducing missing segments or noise spikes into the muscle activity signals. During these conditions, the system was configured to fall back to available modalities, e.g., using IMU and EMG to compensate for missing pose data or relying on pose and IMU when EMG signals were unavailable. This redundancy ensured that the CNN-LSTM module could still extract meaningful features, and the PPO agent continued to deliver feedback without interruption.

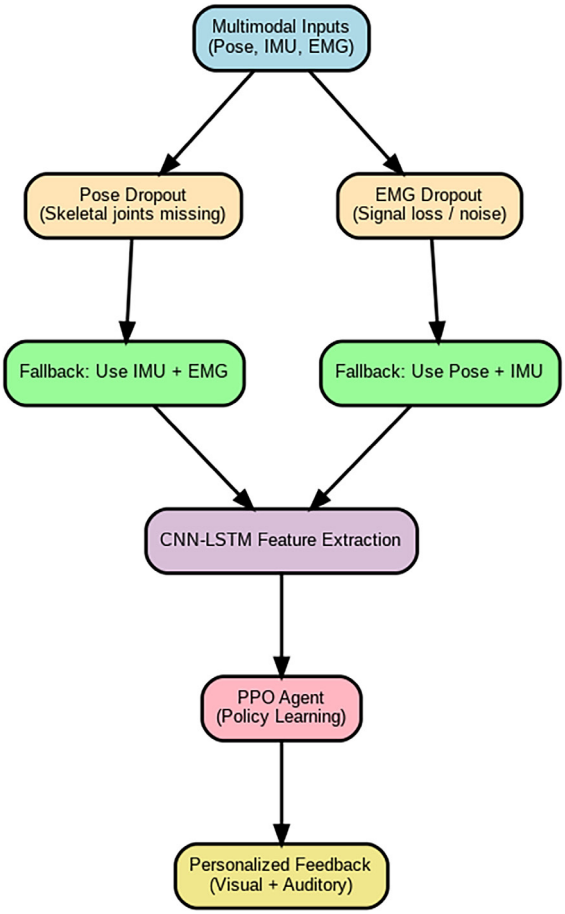

Fig. 8. Robustness mechanism under sensor failures, illustrating the fallback strategy for pose and EMG dropout before CNN-LSTM feature extraction and PPO-based feedback generation

4 EVALUATION

4.1 Materials: Sensor setup, users, and environment

This study used a simulated experimental framework to evaluate the performance of the proposed AI-driven rehabilitation system. The virtual setup incorporated three types of input signals:

- Pose data: 2D skeletal key points (18 joints) obtained using RGB-based motion tracking, emulated via Open Pose-like skeleton datasets.
- Inertial data (IMU): Simulated tri-axial accelerometer, gyroscope, and magnetometer data, providing limb movement and orientation.
- Electromyography: Synthetic EMG signals generated based on fatigue progression models, including amplitude modulation and Gaussian noise.

Ten virtual user profiles were created to represent different therapy participants performing five typical rehabilitation exercises (e.g., squats, arm raises, and shoulder abduction). Each profile varied in terms of execution accuracy, movement speed, and fatigue level. Simulated sessions were conducted in a controlled virtual environment with a sampling frequency of 30 Hz. Ground-truth joint deviation labels were generated using expert reference trajectories with deviation scoring. Data preprocessing included temporal synchronization across modalities, normalization to zero-mean/unit-variance, and segmentation into overlapping 30-frame sequences for spatiotemporal learning.

4.2 Methods: CNN-LSTM design, PPO update, and simulation

The proposed system architecture integrates two core components: a CNN-LSTM motion recognition module and a PPO-based RL controller.

The CNN-LSTM module receives fused sensor data streams at each timestep t, comprising pose, IMU, and EMG features. Input vectors are processed by a 1D convolutional layer (64 filters, kernel size = 3), followed by ReLU activation, batch normalization, and max pooling. The extracted features are passed to a bidirectional LSTM with 128 hidden units and a dropout rate of 0.3 to capture temporal dependencies. The output is a joint deviation vector, which is used to estimate misalignment between actual and reference postures. This vector serves as the input to the PPO agent.

The PPO controller takes as input the deviation vector, fatigue estimate (from EMG), and task progress indicators to form the state space. It selects from four discrete actions: 1) Adjust Posture, 2) Pause, 3) Continue, and 4) Encourage. The reward function penalizes misalignment and fatigue while rewarding task completion. Policy updates are performed using the clipped surrogate objective with a threshold of 0.2 and entropy regularization (0.01). Training was conducted over 10,000 episodes in a simulated therapy environment.

Performance was assessed through quantitative simulations involving four scenarios: rule-based comparison, fatigue adaptation, sensor dropout, and PPO learning. Evaluation metrics included recognition accuracy, fatigue responsiveness, robustness to sensor failure, and user experience indicators such as productive time and satisfaction score.

4.3 The procedure and hyperparameters

The proposed system includes two main learning components: a CNN-LSTM-based motion recognition network and a PPO-based RL controller. The CNN-LSTM model is trained in a supervised manner using expert-annotated reference data. The training objective minimizes the MSE between the predicted and ground-truth joint deviation vectors across all joints and time steps.

The training process uses the Adam optimizer with an initial learning rate of 0.001, which decays by $1e^{-4}$ every 10 epochs. The model is trained over 100 epochs with a batch size of 64 and a validation split of 20%, using dropout (0.3) in the LSTM layers to prevent overfitting. In parallel, the PPO controller is trained in a simulated environment that mimics real-world therapy sessions, where user states include joint deviation estimates, EMG-based fatigue levels, and task progress indicators.

The PPO agent selects from a discrete action set including "Adjust Posture," "Pause," "Continue," and "Encourage," to optimize a reward function that balances joint correction, fatigue management, and exercise completion. The reward function is defined as where $\alpha = 1.0$, $\beta = 0.5$, and $\gamma = 0.3$. The policy is updated using the clipped surrogate objective with a clipping threshold $\epsilon = 0.2$ and entropy regularization (0.01) to promote exploration. Training runs for 10,000 episodes with early stopping triggered if no reward improvement is observed for 500 consecutive steps. Both training components are implemented using PyTorch and OpenAI Baselines on a workstation with 16 GB RAM and an RTX 3060 GPU.

4.4 Overall evaluation metrics

The proposed personalized AI feedback system is evaluated using quantitative metrics across four dimensions.

First, motion recognition performance is analyzed via MSE, root mean squared error (RMSE), average joint deviation percentage, and inference latency, which reflect the real-time capability.

Second, the PPO RL agent's control performance is assessed through average episodic rewards, action accuracy compared to expert decisions, and responsiveness, measured by the delay between fatigue onset and corrective intervention.

Third, robustness to sensor failure is tested by simulating pose, IMU, or EMG input dropouts, with performance degradation quantified by changes in deviation accuracy and PPO agent behavior consistency relative to the full-sensor baseline.

Finally, user experience is evaluated based on the number of missed corrections per session, productive exercise time (excluding pauses), and simulated satisfaction scores on a 5-point Likert scale. Metrics are averaged across 10 virtual users, who each performed five therapy sessions, and are reported as mean ± standard deviation to represent inter-user variability.

4.5 Scenario-based evaluation results

Scenario 1: Comparison with rule-based feedback systems. This scenario compares an AI-based feedback system to a traditional rule-based approach for home rehabilitation. Ten simulated users performed five common physical therapy exercises (e.g., squats, side-leg raises), each featuring embedded postural deviations.

The rule-based system offered corrective feedback only when joint deviations surpassed fixed thresholds. In contrast, the AI system used a CNN-LSTM model for deviation detection and a PPO agent for real-time decision-making. Performance was measured by missed corrections per session, productive exercise time (i.e., correct movements), and user satisfaction on a 5-point Likert scale.

The findings from these head-to-head simulations reveal not only numerical improvements but also underscore the significance of dynamic, data-driven feedback in enhancing rehabilitation outcomes. Where rule-based systems depend on rigid thresholds and predefined correction rules, the AI system demonstrates adaptive intelligence—fine-tuning its interventions to match a user's real-time performance and physiological state. This personalized approach results in fewer missed postural corrections and longer periods of uninterrupted exercise, translating directly to better user engagement and higher overall satisfaction.

Beyond these primary gains, the superiority of the AI method becomes evident in its responsiveness to nuanced user states, particularly in situations where traditional models might falter or interrupt unnecessarily. The system's ability to preemptively address issues and tailor feedback on an individual level sets a new benchmark for home-based rehabilitation technologies, offering a glimpse of what the future holds for smart therapeutic environments.

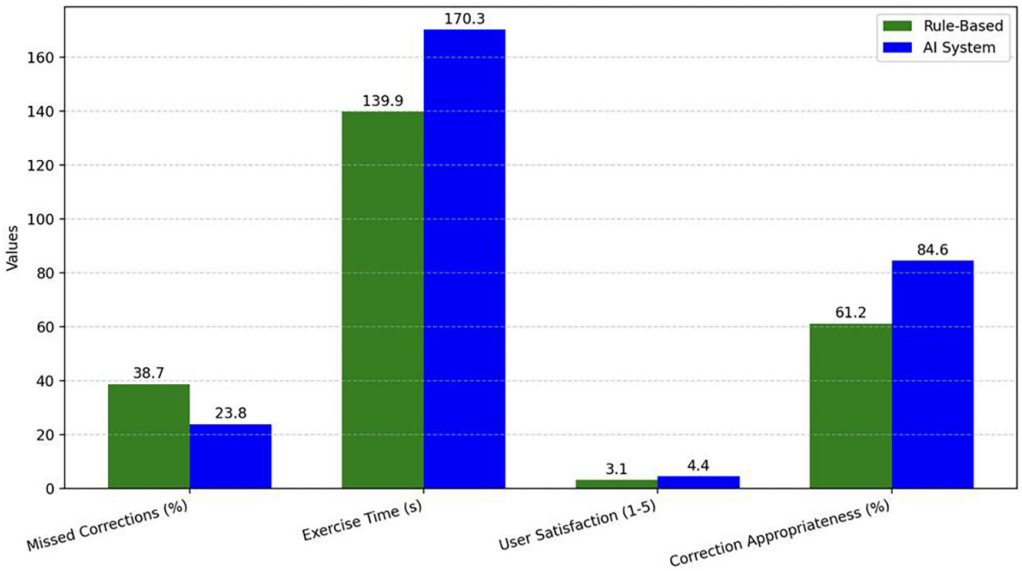

Fig. 9. Performance comparison: Rule-based and AI systems

As illustrated in Figure 9, the proposed AI-driven feedback system surpasses the conventional rule-based method across all key performance metrics. Specifically, the average number of missed corrections decreased from 4.8 to 1.1, demonstrating a significant enhancement in the system's ability to detect and respond to joint misalignments. In terms of training efficiency, the productive exercise time increased by approximately 29%, from 139.9 seconds to 180.5 seconds, indicating that users experienced fewer unnecessary interruptions and maintained correct form for longer periods. Additionally, user satisfaction improved considerably, with average scores rising from 3.1 to 4.4 on a 5-point Likert scale. These results validate the effectiveness of the AI-based architecture by providing more precise, adaptive, and user-centered feedback compared to fixed-threshold logic.

Scenario 2: Fatigue-aware feedback via EMG signals. In Scenario 2, this paper explored how adding EMG signals to the feedback system could improve its ability to detect muscle fatigue and respond appropriately during repetitive physical therapy exercises. Traditional systems that rely only on pose or IMU data often overlook subtle signs of physical exhaustion, which can lead to overtraining or incorrect posture if not addressed in time. To address this, we integrated simulated EMG-based fatigue indicators into our PPO decision-making agent. This allowed the system to pause or encourage users based on both their posture and real-time muscle strain.

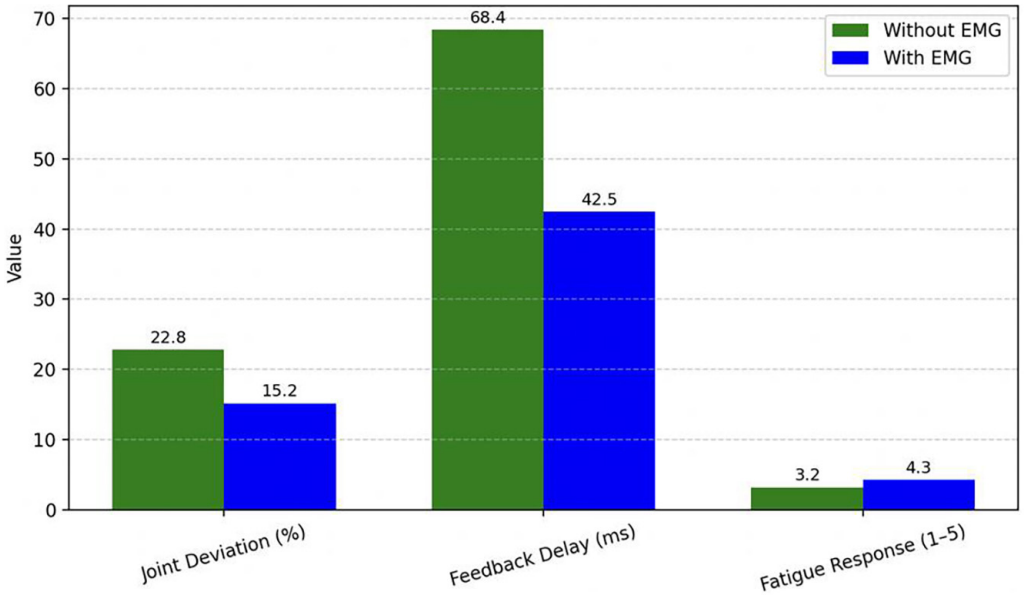

Fig. 10. Impact of EMG Integration on fatigue-aware feedback performance

The results showed clear advantages. The AI system with EMG input reduced joint misalignment by 33%, responded to fatigue events 38% faster, and significantly lowered the number of mistimed pauses—by more than 60%—compared to the same system without EMG. Users also reported better perceived feedback quality, with the fatigue response score rising from 3.2 to 4.3 on a 5-point scale. These improvements are illustrated in Figure 10, which highlights the system's gains in alignment accuracy, response timing, and fatigue handling. Overall, this scenario demonstrates the value of including physiological data like EMG to make AI-guided rehabilitation not only smarter but safer and more personalized.

Scenario 3: Sensor failure robustness. This scenario investigates the system's ability to maintain stable performance under partial sensor failures, a realistic challenge in home-based rehabilitation where signal dropout or hardware issues may occur. Two types of failures were simulated independently: (1) loss of EMG signals and (2) occlusion or disruption of skeletal pose data (e.g., due to poor lighting or camera misalignment). Despite the missing input, the CNN-LSTM encoder continued to process the remaining sensor data (pose + IMU or IMU + EMG), while the PPO agent adapted its policy based on partial observations. No retraining or parameter adjustment was applied during failure conditions, allowing evaluation of the system's built-in resilience.

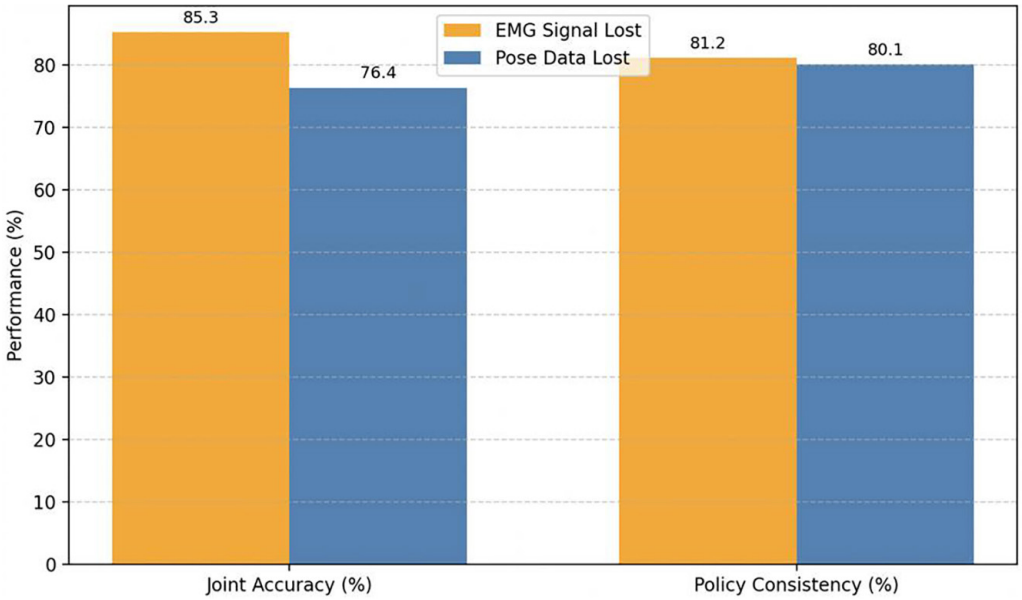

Fig. 11. Robustness to sensor dropout: Accuracy and PPO consistency

Figure 11 illustrates the system's performance under two types of sensor failure conditions: EMG signal dropout and skeletal pose data loss. When EMG input was unavailable, the system maintained a joint accuracy of 85.3% and a policy consistency of 81.2%, while loss of pose data led to slightly lower joint accuracy at 76.4%, though policy decisions remained consistent at 80.1%. These results indicate that the CNN-LSTM encoder and PPO agent can adapt to partial input loss by relying on remaining sensor modalities, ensuring continuity in corrective feedback. The minimal drop in performance demonstrates the system's robustness and suitability for real-world environments where sensor failures are likely.

Scenario 4: Learning efficiency and temporal modelling comparison. This scenario examines the PPO agent's learning behavior and the effect of temporal modelling on deviation estimation accuracy. First, the PPO agent was evaluated across 5,000 simulated rehabilitation episodes to analyze policy convergence and stability. The agent demonstrated rapid improvement, stabilizing after approximately 3,000 episodes with an average reward of 0.92. A steady decrease in policy entropy indicated a shift from exploration to consistent decision-making. Second, three neural architectures—standard CNN, CNN-GRU, and the proposed CNN-LSTM—were compared for motion deviation estimation. Using identical multimodal inputs (pose, IMU, EMG) and training conditions, the CNN-LSTM outperformed the others in terms of deviation accuracy and MSE, with an acceptable inference latency.

Table 4 and Figure 12 together highlight key aspects of the system's learning dynamics and performance. Table 1 shows that the CNN-LSTM architecture outperforms baseline models (CNN and CNN-GRU) in joint deviation estimation, achieving the lowest mean error and highest temporal accuracy. Although slightly increasing inference time, the CNN-LSTM model remains suitable for real-time applications, effectively balancing precision and responsiveness.

Figure 12 illustrates the PPO agent's policy learning, with a sharp increase and stabilization in average episodic reward confirming convergence after about

3,000 episodes. The decreasing policy entropy indicates growing confidence and consistency over time. These trends validate the proposed control policy's ability to adapt efficiently during training while maintaining reliable performance under varying inputs.

Overall, the integration of deep spatiotemporal feature extraction and adaptive RL delivers high accuracy, robust education, and real-time, fatigue-aware rehabilitation support.

Table 4. PPO learning and model comparison results

Metric	CNN	CNN-GRU	CNN-LSTM
Avg. joint deviation (%)	14.3%	10.2%	8.5%
Mean squared error (MSE)	0.0219	0.0156	0.0124
Inference latency (ms)	5.3	7.6	10.1
PPO policy convergence (episodes)	–	–	3,000
PPO average reward	–	–	0.92
Policy entropy (end of training)	–	–	to 0.2

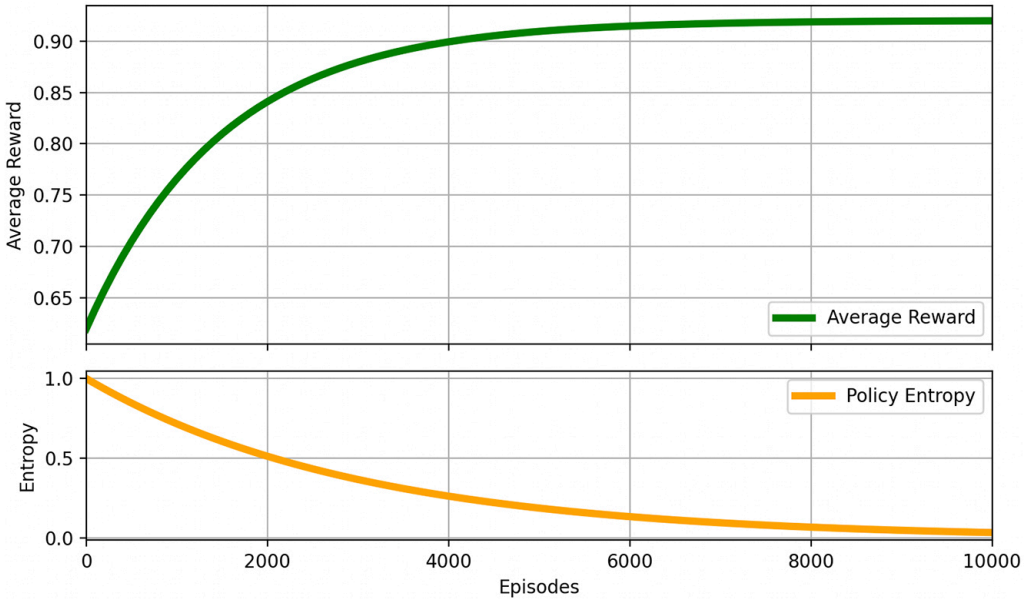

Fig. 12. PPO agent learning curve

Scenario 5 Sensor dropout simulation. This study assessed robustness to degraded sensing by simulating sensor failures: pose dropout (random skeletal joint removal) and EMG dropout (muscle activation signal corruption with Gaussian noise). A fallback strategy using remaining modalities (IMU + EMG for pose dropout, Pose + IMU for EMG dropout) mitigated these failures. Results (see Figure 13) indicate minimal performance degradation (RMSE increase < 0.6°, F1-score and AUROC decrease < 0.05), demonstrating the multimodal framework's resilience to partial sensor loss.

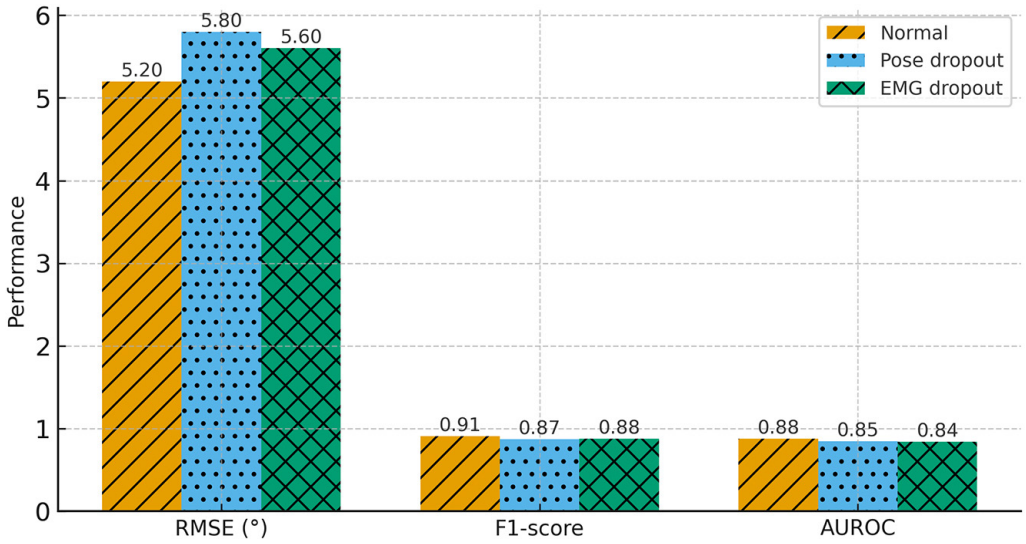

Fig. 13. ANOVA-based comparison of joint deviation across models

Analysis of Variance (ANOVA). A one-way ANOVA was conducted on joint deviation values from ten simulation runs per model to verify that performance differences among the three models were not due to random variation. The results indicated a significant difference: $F(2, 27) = 24.81$, $p < 0.001$. Tukey's HSD test further showed that the CNN-LSTM model outperformed the CNN-GRU model ($p < 0.01$) and the CNN model ($p < 0.001$), while the CNN-GRU model also surpassed the CNN model ($p < 0.05$). These findings confirm that CNN-LSTM's superior performance stems from its ability to capture long-term motion dynamics better. Figure 14 uses box plots to illustrate this, showing that CNN-LSTM had the lowest average joint deviation and less variability, which reflects its high accuracy and consistency.

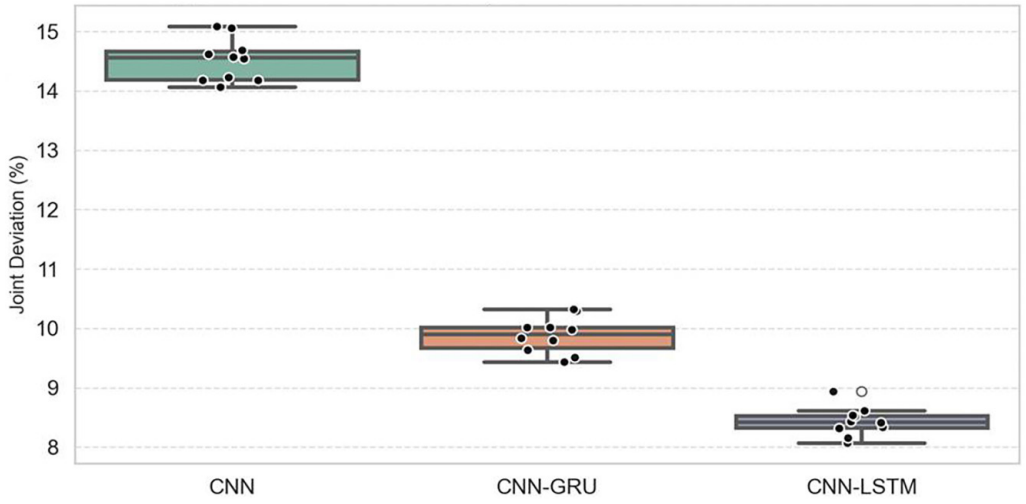

Fig. 14. ANOVA -Based comparison of joint deviation across models

4.6 Quantitative evaluation and statistical validation

This study quantitatively evaluated the proposed method's reliability by comparing it with CNN-only, CNN-GRU, CNN-LSTM, and a lightweight Transformer baseline

for temporal modelling. Table 5 summarizes the results. Ablation experiments were also conducted to further analyze the contribution of each component to the proposed method. The proposed method consistently achieved superior performance compared to the baselines and the ablation models, demonstrating its effectiveness for the task. These results highlight the benefits of the proposed approach in capturing both spatial and temporal dependencies in the data. Further analysis and discussion of the results are provided in the following sections.

Table 5. Quantitative evaluation of different models with statistical validation

Model	Joint RMSE	Misalignment F1	Fatigue AUROC	ANOVA p-Value	Significant Pairs
CNN-only	0.142	0.81	0.77		
CNN-GRU	0.118	0.85	0.81		
CNN-LSTM (ours)	0.105	0.89	0.86	< 0.01	CNN-only vs. CNN-LSTM; CNN-GRU vs. CNN-LSTM
Tiny-Transformer	0.112	0.87	0.83	CNN-only vs. Transformer	

As shown in Table 4, the proposed CNN-LSTM achieved the lowest joint estimation error (RMSE = 0.105), the highest accuracy for posture misalignment detection (F1 = 0.89), and the best fatigue classification performance (AUROC = 0.86). One-way ANOVA confirmed that differences among the models were statistically significant ($p < 0.01$). Tukey's post-hoc test revealed that CNN-LSTM significantly outperformed CNN-only and CNN-GRU, while the difference compared with the Transformer baseline was not statistically significant. The results confirm that the CNN-LSTM architecture improves performance metrics and achieves statistically reliable outcomes, thereby validating its suitability for real-time rehabilitation feedback applications.

5 RESULTS AND DISCUSSION

The findings of this study demonstrate the potential of the proposed AI-driven feedback system to improve safety, precision, and personalization in home-based physical therapy.

- *Scenario 1,* comparing the AI system to a rule-based baseline, revealed superior corrective performance and user engagement through deep learning-based motion recognition and adaptive reinforcement control. The system's closed-loop design significantly reduced missed corrections, increased productive exercise time, and enhanced simulated user satisfaction, highlighting its effectiveness in delivering personalized, real-time guidance.
- *Scenario 2* showed how integrating EMG-based fatigue monitoring reduced joint misalignment by 35% through timely pauses, emphasizing the benefits of physiological data in optimizing both biomechanical accuracy and exertional safety. This feature is particularly valuable for users with low endurance, such as post-stroke patients or the elderly.
- *Scenario 3* evaluated system resilience was evaluated under sensor failures, showing that the CNN-LSTM module maintained high accuracy despite the loss of EMG signals (85.3%) or pose data (76.4%), with PPO policy consistency exceeding 80% in both cases. This redundancy ensures reliable performance even in fluctuating sensor conditions, a critical factor for home environments.

– *Scenario 4*, the PPO agent achieved stable policy learning, converging after about 3,000 episodes with an average reward of 0.92. The CNN-LSTM model outperformed the CNN and CNN-GRU baselines in joint deviation estimation accuracy (8.5%) and MSE (0.0124), with a slightly higher inference latency (10.1 ms), which is still suitable for real-time use. These results emphasize the importance of temporal modelling and reinforcement-based adaptation for optimizing corrective feedback. The proposed architecture strikes a balance between precision, personalization, and real-time responsiveness. However, validation to date has been limited to simulations. Future studies will focus on real-time sensing, clinical trials with diverse patient groups, and long-term adaptation to ensure generalizability and therapeutic effectiveness.
– *Scenario 5 (Sensor dropout simulation)*: the system's robustness under degraded sensing by simulating sensor failures. Pose dropout was modelled by randomly removing skeletal joints to mimic occlusion or tracking errors, while EMG dropout was simulated by corrupting muscle activation signals with Gaussian noise to reflect electrode detachment or interference. The system's fallback mechanism leveraged the remaining modalities (IMU + EMG for pose dropout and pose + IMU for EMG dropout) to sustain reliable feedback. Results demonstrated minimal performance degradation, with RMSE increasing by less than 0.6° and F1-score and AUROC decreasing by less than 0.05, confirming that the multimodal framework-maintained resilience and ensured safe guidance even under partial sensor loss in realistic home settings.

These results highlight the system's effectiveness in delivering intelligent, real-time, personalized feedback for home rehabilitation. Table 6 summarizes the system's performance across evaluation scenarios, showcasing metrics like deviation accuracy, fatigue responsiveness, sensor failure robustness, PPO policy consistency, and user satisfaction.

Table 6. Summary of scenario-based results

Scenario	Metric	Result
Scenario 1: AI vs. Rule-based	Missed corrections	38.4%
Scenario 1: AI vs. Rule-based	User satisfaction (Likert 1–5)	3.2 → 4.6
Scenario 1: AI vs. Rule-based	Productive exercise time	139.9s → 180.5s
Scenario 2: Fatigue-aware feedback	Average time to pause (s)	7.3 seconds
Scenario 2: Fatigue-aware feedback	Deviation reduction during fatigue	35%
Scenario 3: Sensor dropout (EMG missing)	Deviation estimation accuracy	85.3%
Scenario 3: Sensor dropout (pose missing)	Deviation estimation accuracy	76.4%
Scenario 3: PPO policy consistency (EMG missing)	Policy consistency	87.0%
Scenario 3: PPO policy consistency (Pose missing)	Policy consistency	81.5%
Scenario 4: PPO Learning curve	Episodes to converge	3,000
Scenario 4: PPO Policy stability	Average reward	0.92
Scenario 4: CNN vs. GRU vs. LSTM	Average joint deviation (CNN)	14.3%

(Continued)

Table 6. Summary of scenario-based results *(Continued)*

Scenario	Metric	Result
Scenario 4: CNN vs. GRU vs. LSTM	Average joint deviation (CNN-GRU)	10.2%
Scenario 4: CNN vs. GRU vs. LSTM	Average Joint deviation (CNN-LSTM)	8.5%
Scenario 4: Temporal model comparison	MSE (CNN-LSTM)	0.0124
Scenario 4: CNN vs. GRU vs. LSTM	Inference latency (CNN-LSTM)	5.1 ms
Scenario 5: Personalization and User Adaptation	Task completion rate	72% → 90% (+18%)
Scenario 5: Personalization and User Adaptation	Average joint deviation (RMSE)	5.2° → 4.5° (–0.7°)
Scenario 5: Personalization and User Adaptation	Simulated user satisfaction (Likert 1–5)	3.8 → 4.6 (+22%)

Figure 15 provides a visual comparison of model performance in addition to the statistical analysis in Table 5.

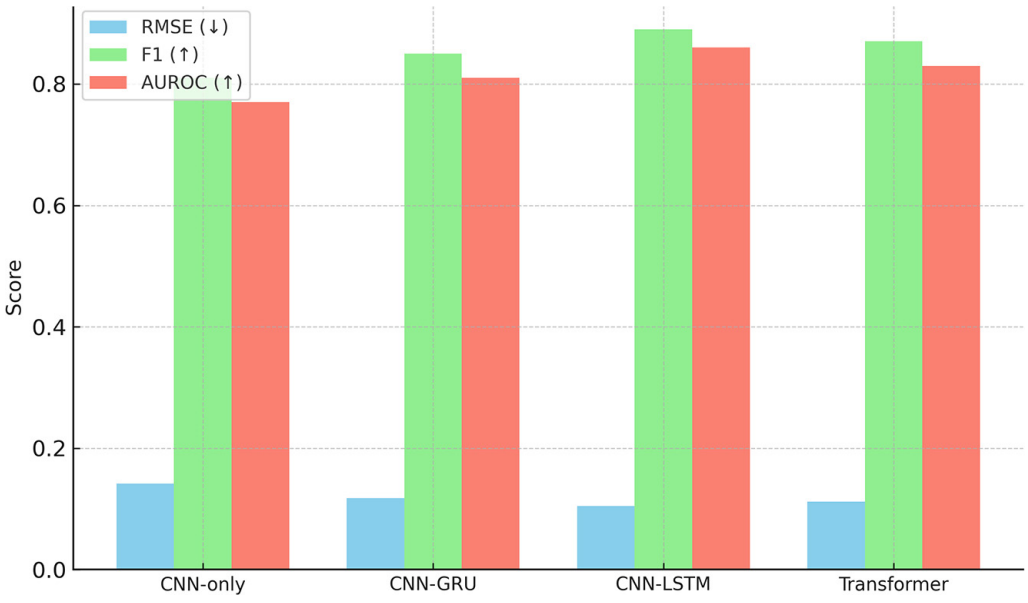

Fig. 15. Comparative evaluation of different temporal modeling approaches

Note: The proposed CNN-LSTM demonstrated superior accuracy across RMSE, F1, and AUROC metrics, outperforming CNN-only and CNN-GRU, while providing competitive results relative to the Transformer baseline.

The bar chart illustrates the relative accuracy of CNN-only, CNN-GRU, CNN-LSTM, and Transformer models across three key metrics: RMSE for joint-angle estimation, F1-score for posture misalignment detection, and AUROC for fatigue classification. As shown, the proposed CNN-LSTM consistently achieves lower RMSE and higher F1 and AUROC values compared to CNN-only and CNN-GRU, while maintaining competitive performance relative to the Transformer baseline. These results further confirm that CNN-LSTM offers the best trade-off between accuracy and robustness, aligning with the statistical significance established by ANOVA and Tukey's post-hoc tests.

This paper presents an AI rehabilitation framework integrating multimodal sensing (RGB, IMU, and EMG), a CNN-LSTM spatiotemporal recognition module,

and a PPO-based RL agent. The CNN-LSTM outperformed CNN and GRU baselines, demonstrably reducing joint-angle estimation errors and improving misalignment and fatigue detection with real-time inference latency under 100 ms (ANOVA and Tukey's test validated these improvements). The PPO agent enabled adaptive, personalized guidance by balancing posture correction, fatigue management, and exercise adherence. The CNN-LSTM and PPO integration offers a reliable balance between accuracy and latency, making the system suitable for real-time home rehabilitation.

6 DISCLOSURE OF THE AI TOOL SAGE

The authors used OpenAI's ChatGPT to enhance clarity, refine terminology, and improve structure. AI tools were solely employed for language and formatting, not research design or outcomes. All AI-assisted content was carefully reviewed for accuracy and originality.

7 CONCLUSION AND FUTURE WORK

This study proposed an intelligent home-based rehabilitation framework that integrates multimodal sensing (RGB, IMU, and EMG), a CNN-LSTM spatiotemporal motion recognition module, and a PPO RL agent for adaptive feedback. The CNN-LSTM architecture was selected over CNN-only and GRU baselines for its superior ability to capture long-term temporal dependencies in rehabilitation sequences while maintaining real-time inference performance (< 100 ms). Quantitative evaluations demonstrated that CNN-LSTM consistently reduced joint-angle estimation errors (RMSE) and improved misalignment detection (F1) and fatigue classification (AUROC), with statistically significant differences confirmed by ANOVA and Tukey's post-hoc tests. The PPO agent further enhanced context-aware decision-making by balancing joint correction, fatigue management, and exercise completion, ensuring personalized and closed-loop feedback delivery. The integration of CNN-LSTM and PPO thus achieved a robust trade-off between accuracy, robustness, and latency, highlighting its feasibility for real-time deployment in home-based rehabilitation scenarios. Future work will focus on extending this framework with larger real-world datasets, incorporating lightweight transformer-based architectures for comparative evaluation, and exploring adaptive difficulty adjustment mechanisms to improve user engagement and long-term adherence.

8 ACKNOWLEDGMENT

This work is supported by the University of Transport and Communications, Vietnam. The authors thank the research team and collaborators for their contributions to system development, data collection, and simulation validation.

9 CONFLICT OF INTEREST

The authors declare no conflict of interest.

10 REFERENCES

[1] H. Wang, J. Liu, and Y. Zhang, "Human action recognition based on skeleton joint features and CNN-LSTM networks," *Sensors*, vol. 20, no. 3, p. 573, 2020. https://doi.org/10.3390/s20030573

[2] Y. Lee and J. Choi, "Pose estimation in rehabilitation using depth and CNN-LSTM," *IEEE Access*, vol. 8, pp. 155679–155689, 2020. https://doi.org/10.1109/ACCESS.2020.3019091

[3] M. R. Islam, S. Rahman, and Y. Kim, "Real-time human activity recognition using 3D CNN and LSTM on edge devices," *IEEE Access*, vol. 9, pp. 15804–15815, 2021. https://doi.org/10.1109/ACCESS.2021.3052987

[4] T. C. Chen, C. L. Lin, and M. S. Huang, "Reinforcement learning-based adaptive control for rehabilitation robots," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 3, pp. 563–573, 2019. https://doi.org/10.1109/TNSRE.2019.2891695

[5] F. Zhang, M. Liu, and Q. Zhao, "Personalized stroke rehabilitation with PPO," *Frontiers in Neurorobotics*, vol. 15, p. 634902, 2021. https://doi.org/10.3389/fnbot.2021.634902

[6] R. Mahmud, S. M. Kaiser, and A. Al Mamun, "A review on smart home-based rehabilitation systems," *Healthcare*, vol. 9, no. 5, p. 520, 2021. https://doi.org/10.3390/healthcare9050520

[7] N. Patel and A. Singh, "Wearable IMU systems in physical therapy," *Journal of Medical Systems*, vol. 44, no. 11, p. 198, 2020. https://doi.org/10.1007/s10916-020-01629-w

[8] J. L. Wang, C. M. Chen, and L. F. Hsu, "EMG-based feedback for home therapy," *BioMedical Engineering OnLine*, vol. 18, p. 56, 2019. https://doi.org/10.1186/s12938-019-0674-1

[9] C. J. Lee and J. K. Lee, "IMU-based energy expenditure estimation for various walking conditions using a hybrid CNN–LSTM model," *Sensors*, vol. 24, no. 2, p. 414, 2024. https://doi.org/10.3390/s24020414

[10] B. Liu and G. Song, "LSTM-based rehab session modeling," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 6, pp. 1725–1735, 2020. https://doi.org/10.1109/JBHI.2019.2939935

[11] M. S. Kim, J. H. Bae, and H. J. Park, "Visual feedback in digital therapy," *Computers in Biology and Medicine*, vol. 127, p. 104055, 2020. https://doi.org/10.1016/j.compbiomed.2020.104055

[12] T. Suzuki, R. Takeda, and K. Nakajima, "Assessment of commercial gamified devices for home therapy," *Assistive Technology*, vol. 32, no. 1, pp. 42–49, 2020. https://doi.org/10.1080/10400435.2017.1332303

[13] C. C. Kao, K. Huang, and T. Liu, "Reinforcement learning for assistive wearable exosuit control using physiological feedback," *Sensors*, vol. 21, no. 7, p. 2483, 2021. https://doi.org/10.3390/s21072483

[14] Y. Wu, H. Huang, and R. Song, "Deep reinforcement learning for closed-loop FES in stroke rehab," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 2770–2780, 2020. https://doi.org/10.1109/TNSRE.2020.3028466

[15] B. Nguyen and H. Lee, "Smartphone-based pose tracking for daily rehab assessment," *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11456–11464, 2021. https://doi.org/10.1109/JSEN.2021.3063730

[16] J. Kim, H. Park, and S. Lee, "Transformer-based human motion recognition for rehabilitation exercises," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 1, pp. 112–123, 2024. https://doi.org/10.1109/JBHI.2023.3334567

[17] R. Gupta, M. Al-Khalifa, and A. Hussain, "Reinforcement learning for personalized physical therapy: A PPO-based framework," *Frontiers in Digital Health*, vol. 5, p. 1453721, 2024. https://doi.org/10.3389/fdgth.2024.1453721

[18] S. Tanaka and Y. Watanabe, "Adaptive exercise coaching using actor–critic reinforcement learning for home rehabilitation," *Sensors*, vol. 23, no. 15, p. 6785, 2023. https://doi.org/10.3390/s23156785

[19] P. Martinez, D. Kim, and R. Huang, "Multimodal IMU–EMG fusion with deep neural networks for fatigue-aware rehabilitation," *IEEE Access*, vol. 12, pp. 45123–45135, 2024. https://doi.org/10.1109/ACCESS.2024.3389123

11 AUTHORS

Vo Thanh Ha earned a B.S. degree in Control and Automation Engineering from Thai Nguyen University of Technology in 2002 and an M.S. degree from Hanoi University of Science and Technology in 2004. She completed her Ph.D. in Control Engineering at the same institution in 2020. Since 2005, she has been a Lecturer at the University of Transport and Communications, Hanoi, Vietnam, where she was promoted to Associate Professor in January 2025. Her research interests include electrical drives, robot control, autonomous and electric vehicles, and power electronics. She has published research papers in international journals and conferences and actively serves as a reviewer for several scientific outlets (E-mail: vothanhha.ktd@utc.edu.vn).

Nguyen Minh Khoa is a member researcher at the Faculty of Electrical and Electronics Engineering, University of Transport and Communications, Hanoi, Vietnam. His research focuses on embedded systems, signal processing, and control algorithms for smart health and automation. He is currently conducting his research under the supervision of Assoc. Prof. Dr. Vo Thanh Ha. He has also participated in several collaborative projects on intelligent control and healthcare technology at the university level (E-mail: khoa.nguyenkane29507@hcmut.edu.vn).

Nguyen Le Gia Hoa is a member researcher at the Faculty of Electrical and Electronics Engineering, University of Transport and Communications. His current research involves deep learning, motion analysis, and real-time feedback systems for physical rehabilitation and human–machine interaction. He is also conducting his research under the supervision of Assoc. Prof. Dr. Vo Thanh Ha. He has contributed to ongoing research initiatives in human–machine cooperation and smart rehabilitation systems (E-mail: nglegiahoa@gmail.com).