Perception-Edge-Cloud Architecture for Post-Stroke Rehabilitation: A Biosensor-Based Intelligent System for Movement Recognition and Compliance Assessment

Tianhui Li
Heilongjiang Open University, Harbin City, Heilongjiang Province, 150000, China
E-mail: litianhui0722@163.com

Keywords: Post-stroke rehabilitation, intelligent evaluation system, nursing compliance, multimodal biosensing, action recognition

Received: 25 January 2026

To address the problem of distorted NC (Nursing Compliance) assessment for stroke patients during home rehabilitation, this paper proposes a four-level collaborative intelligent assessment and guidance system architecture: "Perception-Edge-Cloud-Feedback." First, upper limb movement and electromyography signals are synchronously collected using IMU (Inertial Measurement Unit) and sEMG (Surface Electromyography) sensors. Signal preprocessing, combined with wavelet denoising and SW (Sliding Window) segmentation, improves interference resistance in unstructured home environments. Second, multidimensional features are extracted from the time-frequency-spatial domain, and the input space is optimized. Finally, an attention-enhanced LSTM (Long Short-Term Memory)-Gated Recurrent Unit (GRU) hybrid architecture is constructed to adaptively capture key spatiotemporal features. Finally, a mapping mechanism between action semantics and rehabilitation prescriptions was constructed, along with a dynamically weighted compliance scoring model. The methodology integrates Daubechies wavelet denoising, sliding window segmentation, and an attention-enhanced LSTM-GRU network to process synchronous IMU and sEMG signals. Experimental validation involving 32 patients demonstrates that this system achieves an average recognition accuracy of 93.6%, an Intraclass Correlation Coefficient (ICC) of 0.89 between automatic and manual compliance scoring, and a significant improvement in Fugl-Meyer Assessment scores ($p < 0.001$) for the high-adherence group.

Povzetek: Članek predstavi štiristopenjski sistem »Percepcija–Edge–Oblak–Povratna zanka«, ki z valčnim razšumljanjem in drsečimi okni obdeluje sinhrone IMU+sEMG signale ter z pozornostno okrepljenim hibridom LSTM–GRU samodejno ocenjuje in usmerja rehabilitacijsko skladnost po možganski kapi.

1 Introduction

As the incidence of stroke continues to rise, a large number of patients enter the long-term rehabilitation stage. Home rehabilitation has become an important model for reducing medical burdens and improving quality of life. However, current home rehabilitation is generally plagued by problems such as non-standard training, inconsistent implementation, and difficulty in quantifying compliance, which restrict the effective improvement of rehabilitation effects [1-3]. While intelligent sensing and AI technologies offer new avenues for motion recognition, unstructured home environments present significant challenges, including substantial individual variability and environmental interference. These factors often compromise the robustness of motion recognition algorithms, leading to distorted Nursing Compliance (NC) evaluations [4-6]. Therefore, there is an urgent need to build an intelligent system that can accurately identify rehabilitation movements and achieve objective and dynamic compliance assessment. The primary objectives of this study are to construct a robust multimodal motion recognition model capable of operating in unstructured environments and to establish a clinically interpretable compliance scoring system that correlates quantitatively with functional recovery outcomes.

This paper focuses on patients with upper limb motor dysfunction after stroke. Their rehabilitation compliance directly affects the process of neural plasticity reconstruction and functional recovery. Gebreheat G [7] pointed out that digital home rehabilitation solutions for stroke patients have the potential to replace traditional training in home or clinic environments; Olk K [8] emphasized that the most important elements of a rehabilitation program are careful guidance, focused practice, and repeated practice. Repeated use of the damaged limb promotes brain plasticity and helps reduce the severity of the injury. Rocha C D [9] pointed out that stroke survivors usually develop compensatory strategies to cope with upper limb dysfunction. These compensatory behaviors must be detected and corrected during rehabilitation to promote long-term functional rehabilitation. Peznola S [10] used objective compliance

measurements provided by rehabilitation games to help therapists identify early which patients scheduled behaviors indicate poor compliance. However, its core indicators are still training frequency and duration, and it fails to deeply evaluate the standardization of movement execution, a key quality dimension that affects rehabilitation effectiveness. Zhou C [11] designed a hemiplegic gait dataset with three severity levels based on medical prior knowledge, specifically for evaluating stroke rehabilitation. Wen X [12] research shows that personalized feedback, psychological support, and time effect are the key to improving rehabilitation compliance in elderly stroke patients. Therefore, accurate recognition and execution quality assessment of typical upper limb rehabilitation movements are crucial for effective compliance monitoring.

In recent years, researchers have attempted to leverage wearable sensors and machine learning to improve rehabilitation movement recognition capabilities. Maskeliūnas R [13] developed a rehabilitation system based on Virtual Reality (VR), which combines a VR physical training monitoring environment with upper limb rehabilitation technology to achieve precise interaction and improve patients' participation in rehabilitation training. Jin F [14] constructed a fusion tensor based on the data fusion mechanism of multiple sensors and used a bidirectional long short-term memory network to improve the performance of activity classification based on wearable sensors. Bijalwan V [15] used a hybrid deep learning method to identify upper limb physical therapy exercises for post-stroke patients with high accuracy and robustness. Hamid I [16] used electromyography-based machine learning to improve the accuracy of upper limb exoskeleton movements for the rehabilitation of stroke patients; Xie J [17] used a heterogeneous classifier based on TinyML (Tiny Machine Learning) to monitor upper limb rehabilitation and promote stroke rehabilitation, achieving a training accuracy of 96.1%, a test accuracy of 95.09%, and a deployment accuracy of 88.01%. Tang C [18] built a multimodal smart home platform to provide continuous home rehabilitation for post-stroke patients. Its plantar pressure insoles equipped with machine learning processes can classify patients according to the stage of motor recovery with an accuracy of up to 94%. Existing methods are mostly limited to action category recognition and lack deep semantic analysis of the relationship between action quality and rehabilitation goals, making it difficult to support clinical-level intelligent compliance evaluation. A comparative analysis reveals that while VR-based systems focus on interaction immersion and TinyML approaches prioritize deployment efficiency, they often sacrifice the depth of physiological signal analysis required for precise compensatory movement detection. The proposed perception-edge-cloud architecture addresses these gaps by integrating multimodal sEMG-IMU fusion with clinical rule-based compliance quantification. Recent advancements in distributed computing and secure transmission in high-mobility networks offer valuable references for medical IoT systems. Studies on DDoS mitigation using Chebyshev polynomials in 5G vehicular networks [19] and fog computing-based authentication schemes [20-21] demonstrate robust strategies for maintaining data integrity in edge-cloud interactions. These strategies parallel the requirements for protecting sensitive patient rehabilitation data against replay and tampering attacks, as highlighted in recent reviews of vehicular ad hoc networks [22] and lattice-based privacy-preserving schemes [23], [24]. Intelligent control frameworks in cyber-physical systems often employ differential flatness-based approaches for trajectory tracking in dual-arm manipulators to ensure real-time adaptability under uncertain conditions. The proposed hybrid LSTM-GRU architecture diverges from these model-based control strategies by adopting a data-driven attention mechanism to handle environmental uncertainty without the high computational cost associated with solving complex differential equations in real-time control loops, thereby optimizing latency for consumer-grade edge devices.

This study proposes a biosensor-based method for motion recognition and intelligent evaluation of NC in post-stroke home rehabilitation. The system utilizes a four-tiered architecture: "perception-edge-cloud-feedback." Multimodal wearable sensors (IMU and sEMG) deployed at key locations on the upper limb collect motion and physiological signals. The IMU precisely captures joint kinematics, whereas the sEMG directly quantifies muscle activation patterns and neural drive, facilitating the differentiation between target rehabilitation movements and compensatory behaviors. The combination of the two effectively distinguishes target movements from compensatory movements. These signals are then uploaded to the cloud after time synchronization and preprocessing at the edge. At the edge layer, wavelet denoising and SW segmentation are used to improve interference resistance and recognition robustness in unstructured home environments. Subsequently, in the cloud, multi-dimensional feature extraction is performed based on the preprocessed data, and an attention-enhanced LSTM-GRU hybrid model is constructed to achieve high-precision recognition of six categories of upper limb rehabilitation movements. Furthermore, through semantic parsing of movements, the recognition results are mapped to individualized rehabilitation prescriptions. A four-dimensional compliance scoring model is constructed, combining completion rate, standardization, frequency consistency, and duration compliance, generating graded feedback. This method innovatively proposes a deep network that integrates multimodal fusion and attention mechanisms to improve the robustness of motion recognition. It introduces a DTW (Dynamic Time Warping)-based quantification of motion norms, using the calculated joint angle trajectory deviation as a core metric. This metric is strongly correlated with the Fugl-Meyer scale's scoring dimensions for movement fluency and coordination, providing a clinically interpretable basis for automated scoring. This paper establishes a "perception-edge-cloud-feedback" closed-loop system that not only achieves a high recognition accuracy of 93.6%, but its edge-cloud collaborative architecture also controls end-to-end latency to less than 203.5ms, ensuring real-time feedback.

2 Algorithm design

2.1 System Architecture and multimodal data collection

This system utilizes a four-level distributed architecture: "perception-edge-cloud-feedback," enabling a complete process from data collection, intelligent analysis, to real-time intervention. The overall system architecture is shown in Fig. 1. The perception layer consists of three wearable sensor nodes deployed at key locations of the patient's upper limbs (shoulder joint, elbow joint, and distal forearm). Each node integrates a six-axis IMU and a dual-channel sEMG sensor. The sampling frequency is uniformly configured as f_s=200Hz to meet the Nyquist sampling theorem's coverage requirements for the main frequency band of upper limb rehabilitation movements [25-26]. Assume that the original physical quantities collected by the ith sensor nodes are as shown in Formulas (1) and (2):

Acceleration vector:

$$\mathbf{a}_i(k)=[a_{x,i}(k),a_{y,i}(k),a_{z,i}(k)]^T\in\mathrm{R}^3 \quad (1)$$

$\mathbf{a}_i(k)$ is the acceleration vector of the node i, angular velocity vector:

$$\boldsymbol{\omega}_i(k)=[\omega_{x,i}(k),\omega_{y,i}(k),\omega_{z,i}(k)]^T\in\mathrm{R}^3 \quad (2)$$

$\boldsymbol{\omega}_i(k)$ is the angular velocity vector of the i node. After the sEMG signal is full-wave rectified and low-pass filtered, the complete multimodal observation vector of the node is defined as follows:

$$\mathbf{x}_i(k)=[\mathbf{a}_i(k)^T,\boldsymbol{\omega}_i(k)^T,v_{\mathrm{emg}_1,i}(k),v_{\mathrm{emg}_2,i}(k)]^T\in\mathrm{R}^8 \quad (3)$$

$v_{\mathrm{emg}_1,i}(k)$ and $v_{\mathrm{emg}_2,i}(k)$ are the surface electromyographic signal amplitudes of the i node. All nodes transmit data packets to the edge computing unit in a broadcast-subscribe mode via the Bluetooth 5.0 protocol. A soft synchronization mechanism based on a high-precision timer is introduced, and the local timestamp of each node is mapped to the global time axis as shown in Formula (4):

$$t(k)=t_0+\frac{k}{f_s} \quad (4)$$

t_0 is the session start time. When the edge receives an asynchronous data packet, a linear time alignment algorithm is used to reconstruct the synchronization sequence. Assuming that the data of the node i at the global time $t(k)$ is not directly sampled, the data is interpolated through its neighboring sampling points $t(j)$ and $t(j+1)$ as shown in Formula (5):

$$\hat{\mathbf{x}}_i(k)=\mathbf{x}_i(j)+\frac{t(k)-t(j)}{t(j+1)-t(j)}(\mathbf{x}_i(j+1)-\mathbf{x}_i(j)),t(j)\le t(k)<t(j+1) \quad (5)$$

$\mathbf{x}_i(j)$ and $\mathbf{x}_i(j+1)$ are the original observations at the corresponding moments, and $\hat{\mathbf{x}}_i(k)$ is the reconstructed value. To enhance the system's adaptability in unstructured home environments, two key mechanisms are introduced at the edge: Ambient Noise Adaptive Filtering: The edge calculates the spectral entropy of each IMU channel signal in real time as a quantitative indicator of the ambient electromagnetic noise level. Sensor Baseline Self-Calibration: The patient is guided to perform a standardized resting posture for 5 seconds. The edge collects IMU and sEMG signals during this phase, calculates and stores new zero bias and baseline values for dynamic correction of subsequent data, eliminating measurement drift caused by sensor displacement or environmental changes. This adaptive calibration strategy parallels robust control frameworks used in MIMO systems, where maintaining system stability against parameter variations is critical. The baseline correction mechanism functions similarly to finite-time fuzzy synchronization in chaotic systems, mitigating the impact of sensor placement inconsistencies and signal noise to ensure the reliability of the compliance scoring model in variable real-world home environments.The synchronized multi-channel signals are encapsulated into structured data frames and uploaded to the cloud inference service via the MQTT (Message Queuing Telemetry Transport) protocol. To ensure the security of large-scale rehabilitation data transmission, future iterations of this protocol will incorporate post-quantum cryptographic techniques. Batch authentication approaches similar to those used in VANETs, such as lattice-based conditional privacy-preserving schemes and certificateless authentication, provide a theoretical basis for securing high-frequency biosensor data streams against replay and tampering attacks. In the event of a network interruption, the local microSD (Secure Digital) card is used for circular caching, and the data is automatically resumed after recovery. The cloud-side end-to-end latency is defined as Equation (6):

$$\tau_{\text{end-to-end}}=\tau_{\text{sampling}}+\tau_{\text{transmission}}+\tau_{\text{alignment}}+\tau_{\text{inference}} \quad (6)$$

After identifying and evaluating the patient's current movements, the cloud sends the results to the feedback layer. This engine receives the cloud-based results and, based on a pre-defined clinical rule base and personalized rehabilitation prescriptions, generates immediate, specific, and easy-to-understand feedback instructions. The feedback engine supports multimodal output: for voice feedback, the system calls the cloud-based TTS (Text-to-Speech) service to convert structured text instructions into natural speech and play it in real time through Bluetooth headphones; for visual feedback, the system pushes correct and incorrect demonstrations of key actions in the form of 3D animations to the patient's mobile phone APP, achieving immersive guidance.

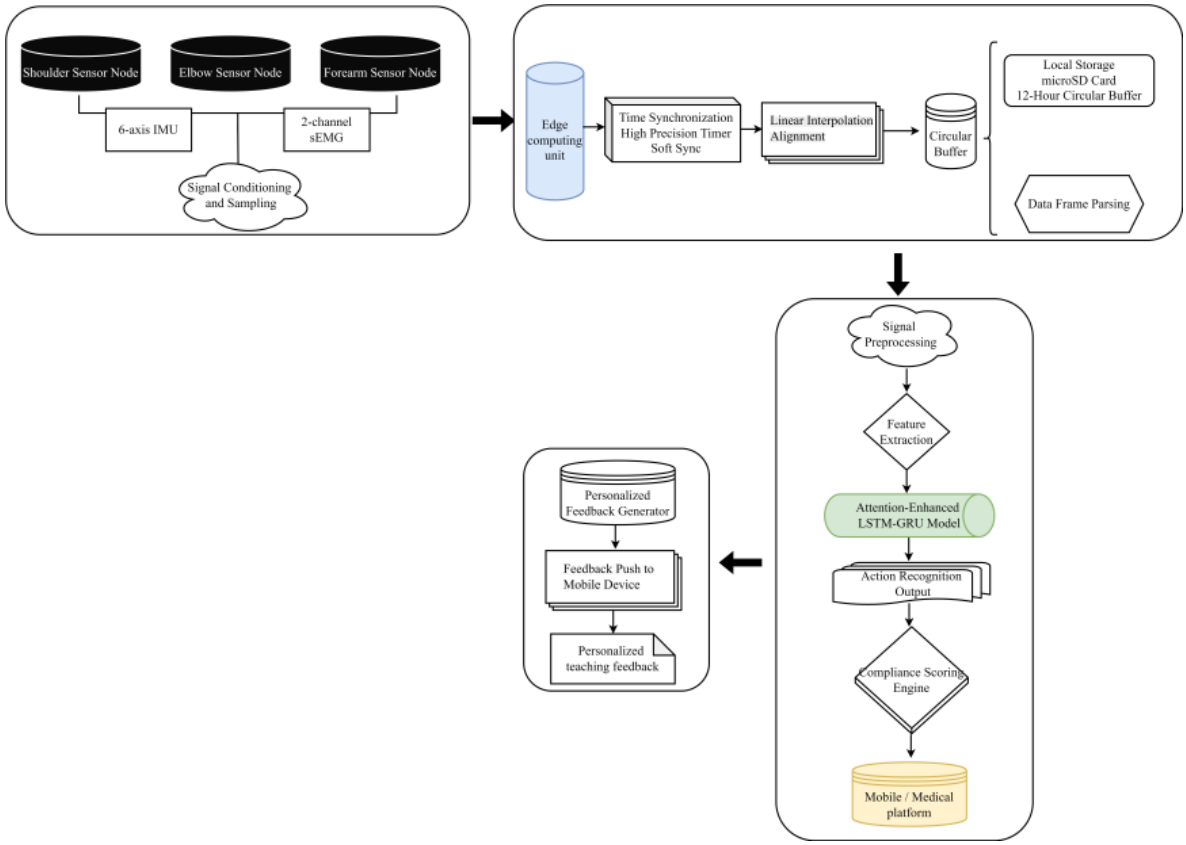

Figure 1: Overall framework

Fig. 1 illustrates the architecture of the biosensor-based post-stroke home rehabilitation motion recognition and NC intelligent assessment system. The perception layer deploys wearable nodes with integrated IMUs and sEMG at key locations on the patient's upper limbs to collect motion and electromyographic signals in real time and transmit the data to the edge layer via Bluetooth 5.0. The edge layer performs multi-node time synchronization, linear interpolation alignment, and local preprocessing, and uses microSD card caching to ensure data integrity when the network is disconnected. The data is then uploaded to the cloud via the MQTT protocol for signal reconstruction, feature extraction, attention-enhanced motion recognition, and intelligent compliance scoring. The results are ultimately pushed to mobile devices or healthcare platforms. Finally, the feedback layer generates personalized instructional feedback based on the scoring results, guiding patients to correct incorrect movements. This creates a closed-loop management system that includes data collection and intelligent feedback, supporting long-term, continuous, and remote home rehabilitation monitoring.

2.2 Signal preprocessing and motion segmentation

This section performs step-by-step preprocessing and precise segmentation of the multimodal raw signals. Assume that the original acceleration, angular velocity, and sEMG signals collected by the ith sensor node in the time series are $\mathbf{a}_i(k)$, $\boldsymbol{\omega}_i(k)$, $\boldsymbol{e}_i(k)$, respectively. First, the motion signals $\mathbf{a}_i(k)$ and $\boldsymbol{\omega}_i(k)$ are subjected to a discrete wavelet transform, and the Daubechies wavelet basis ψ_{db6} is selected to perform a 5-layer multi-resolution decomposition [27-28], as shown in Formula (7):

$$c_j(l),d_j(l)=\langle x,\phi_{j,l}\rangle,\langle x,\psi_{j,l}\rangle,j=1,\ldots,5 \tag{7}$$

In Formula (7), c_j is the approximation coefficient, d_j is the detail coefficient, $\phi_{j,l}(t)=2^{-j/2}\phi(2^j t-l)$ is the scaling function, and $\psi_{j,l}(t)=2^{-j/2}\psi(2^j t-l)$ is the wavelet basis function. The soft threshold method is used to denoise the high-frequency detail coefficients as shown in Formula (8):

$$\tilde{d}_j(l)=\text{sgn}(d_j(l))\cdot max(|d_j(l)|-\lambda_j,0),\lambda_j=\sigma_j\sqrt{2\log\ N_j} \tag{8}$$

In Formula (8), σ_j is the estimated value of the noise standard deviation of the jth layer, and N_j is the coefficient length of this layer. The threshold determination relies on the Stein's Unbiased Risk Estimate (SURE) principle to adaptively minimize signal distortion while effectively suppressing white noise. For the sliding window segmentation, a window length of 2 seconds with a 50% overlap is configured to capture complete motion cycles, ensuring that the transition phases between repetition are adequately represented for the subsequent feature extraction. The reconstructed signal is as shown in Formula (9):

$$x_{\text{denoised}}=\sum_l \tilde{c}_5(l)\phi_{5,l}+\sum_{j=1}^{5}\sum_l \tilde{d}_j(l)\psi_{j,l} \tag{9}$$

The sEMG signal $\boldsymbol{e}_i(k)$ is first full-wave rectified, and then the envelope is extracted through a second-order

zero-phase Butterworth low-pass filter. The differential equation is as follows:

$$y[n]=b_0x[n]+b_1x[n\text{-}1]+b_2x[n\text{-}2]-a_1y[n\text{-}1]-a_2y[n\text{-}2] \quad (10)$$

The filtering is performed once in the forward direction and once in the reverse direction to achieve zero phase delay, as shown in Formula (11):

$$e_{\text{env}}=\text{filtfilt}(b,a,e_{\text{rect}}) \quad (11)$$

The denoised acceleration and angular velocity signals are further subjected to bandpass zero-phase filtering to retain the effective frequency band of rehabilitation movements. A second-order Butterworth bandpass filter is designed, and its transfer function is shown in Formula (12):

$$H(z)=\frac{\sum_{m=0}^{2} b_m z^{-m}}{1+\sum_{n=1}^{2} a_n z^{-n}} \quad (12)$$

After filtering is completed, overlapping SW segmentation is performed on the continuous signal stream. Let the current signal segment be $\mathbf{s}(k)\in\mathrm{R}^8$, and the signal segment of the mth window is as shown in Formula (13):

$$\mathbf{S}_m=\{\mathbf{s}(k)|k\in[m\cdot S,m\cdot S+L\text{-}1]\},m=0,1,2,\ldots \quad (13)$$

This strategy ensures that t_{onset} is covered by at least one complete window at the start or end of any action, satisfying $t_{\text{onset}}\in[\frac{mS}{f_s},\frac{mS+L}{f_s})$, thereby improving the sensitivity of the subsequent classification model to action boundaries and the robustness of recognition.

2.3 Multi-dimensional spatiotemporal feature extraction and selection

This section extracts three types of features from the time domain, frequency domain, and spatial domain based on the signal segments after SW segmentation, constructs a high-dimensional feature space, and performs feature selection to optimize the model input [29-30]. Time domain feature extraction: The following statistics are calculated for each channel signal $x(k)$:

The mean is as shown in Formula (14):

$$\mu_x=\frac{1}{L}\sum_{k=1}^{L}x(k) \quad (14)$$

The standard deviation is as shown in Formula (15):

$$\sigma_x=\sqrt{\frac{1}{L\text{-}1}\sum_{k=1}^{L}(x(k)\text{-}\mu_x)^2} \quad (15)$$

The peak-to-peak value is as shown in Formula (16):

$$R_x=\max_k x(k)-\min_k x(k) \quad (16)$$

The zero-crossing rate (ZCR) is as follows:

$$Z_x=\frac{1}{L\text{-}1}\sum_{k=1}^{L\text{-}1}\mathrm{I}\,(x(k)\cdot x(k+1)<0) \quad (17)$$

In Formula (17), $\mathrm{I}\,(\cdot)$ is the indicator function. Frequency domain feature extraction: Perform fast Fourier transform on each channel signal as shown in Formula (18):

$$X(f_n)=\sum_{k=0}^{L\text{-}1}x(k)e^{-j2\pi nk/L},f_n=\frac{nf_s}{L},n=0,1,\ldots,\frac{L}{2} \quad (18)$$

Calculate the power spectral density (PSD) as shown in Formula (19):

$$P(f_n)=\frac{2|X(f_n)|^2}{f_sL} \quad (19)$$

Extract frequency domain indicators: Dominant Frequency as shown in Formula (20):

$$f_{\text{dom}}=\arg\max_{f_n\in[1,50]}P(f_n) \quad (20)$$

The energy proportion of each frequency band is calculated using Formula (21):

$$E_b=\frac{\sum_{f_n\in B_b}P(f_n)\Delta f}{\sum_{f_n=1}^{50}P(f_n)\Delta f},b=1,\ldots,5 \quad (21)$$

Spatial kinematic feature extraction: Based on the quaternion attitude solution results of the shoulder and elbow dual IMU, the joint angle sequence $\theta(k)$ is obtained. The angular velocity $\dot{\theta}(k)$ and angular acceleration $\ddot{\theta}(k)$ are defined by central difference as shown in Formula (22):

$$\dot{\theta}(k)=\frac{\theta(k+1)\text{-}\theta(k\text{-}1)}{2\Delta t},\ddot{\theta}(k)=\frac{\theta(k+1)\text{-}2\theta(\mathrm{k})+\theta(\mathrm{k}\text{-}1)}{(\Delta t)^2} \quad (22)$$

Calculate the spatial feature: Range of Motion (ROM) as shown in Formula (23):

$$\mathrm{ROM}=\max_k\theta(k)-\min_k\theta(k) \quad (23)$$

The trajectory curvature is as shown in Formula (24):

$$\kappa(k)=\frac{|\dot{\theta}(k)\ddot{\theta}(k)|}{(1+\dot{\theta}(k)^2)^{3/2}},\bar{\kappa}=\frac{1}{L}\sum_{k=1}^{L}\kappa(k) \quad (24)$$

The trajectory smoothness is as shown in Formula (25):

$$j(k)=\frac{d^3\theta}{dt^3}\approx\frac{\ddot{\theta}(k+1)\text{-}\ddot{\theta}(k\text{-}1)}{2\Delta t},\mathrm{JI}=\sqrt{\frac{1}{L}\sum_{k=1}^{L}j(k)^2} \quad (25)$$

To enhance the clinical interpretability of the features and directly capture compensatory behavior, two key features based on medical prior knowledge are introduced: the co-contraction index (CCI), as shown in Formula (26):

$$\mathrm{CCI}=\frac{2}{T}\int_0^1\frac{\min(\mathrm{EMG}_{\text{agonist}}(t),\mathrm{EMG}_{\text{antagonist}}(t))}{\mathrm{EMG}_{\text{agonist}}(t)+\mathrm{EMG}_{\text{antagonist}}(t)}\mathrm{d}t \quad (26)$$

The symmetry of joint range of motion is as shown in Formula (27):

$$\mathrm{ROMS}=\frac{\max(|\omega_{\text{affected}}|)}{\max(|\omega_{\text{unaffected}}|)} \quad (27)$$

Merge the three types of features to form a high-dimensional vector and perform Z-score normalization as shown in Formula (28):

$$\hat{f}_i=\frac{f_i\text{-}\mu_i}{\sigma_i} \quad (28)$$

The recursive feature elimination algorithm is combined with a linear support vector machine (SVM) to

perform feature sorting [31-32]. Initialize the weight vector and train the SVM:

$$min\frac{1}{2}\|\mathbf{w}\|^2+C\sum_{j=1}^{M}\xi_j \quad (29)$$

$$y_j(\mathbf{w}^T\hat{\mathbf{f}}_j+b)\geq 1-\xi_j,\xi_j\geq 0 \quad (30)$$

The absolute value of each feature weight is calculated, the smallest weight is removed, and retraining is repeated until the remaining 64-dimensional features are obtained, forming the optimal feature subset that serves as the input for the subsequent deep model.

2.4 Construction of an attention-enhanced hybrid deep model

This paper constructs an attention-enhanced LSTM-GRU hybrid deep neural network. The network architecture is shown in Fig. 2. This architecture fully combines the advantages of LSTM in modeling long-term dependencies with the rapid response of GRU to dynamic changes, and introduces a temporal attention mechanism to enhance the semantic representation of key action segments [33-34]. Assume that the input is a feature sequence consisting of T time windows after feature selection and is mapped to a high-dimensional latent space through a fully connected layer, as shown in Formula (31):

$$\mathbf{H}^{(0)}=\tanh\ (\mathbf{W}_e\mathbf{X}+\mathbf{b}_e) \quad (31)$$

The bidirectional LSTM layer extracts sequence features in the forward and backward propagation directions, respectively. The forward hidden state is denoted as $\mathbf{H}_{\text{fwd}}$ and the backward hidden state as $\mathbf{H}_{\text{bwd}}$. The final hidden state representation $\mathbf{H}^{(1)}=[\mathbf{H}_{\text{fwd}},\mathbf{H}_{\text{bwd}}]$ is obtained through a concatenation operation. A temporal attention mechanism is introduced at the LSTM output to enhance the representation capability of key action frames [35-36]. The attention weight for each time step is calculated as shown in Formula (32):

$$e_t=\mathbf{u}_c^T\tanh\ (\mathbf{W}_a\mathbf{h}_t^{(1)}),\alpha_t=\frac{\exp\ (e_t)}{\sum_{t'=1}^{T}\exp\ (e_{t'})} \quad (32)$$

The context vector obtained by weighted fusion is as shown in Formula (33):

$$\mathbf{r}=\sum_{t=1}^{T}\alpha_t\mathbf{h}_t^{(1)} \quad (33)$$

This vector serves as the initial memory input of the GRU layer, guiding the model to focus on key stages such as the start and peak of the action. The context vector, generated by the weighted summation of hidden states, acts as a semantic anchor that highlights the temporal segments with the highest diagnostic value, thereby preventing the subsequent GRU layers from being diluted by irrelevant resting or transition phases.The Bidirectional-GRU layer further captures dynamic changes, and its update gate $\mathbf{z}_t$ and reset gate $\mathbf{r}_t$ are defined as follows:

$$\mathbf{z}_t=\sigma(\mathbf{W}_z[\mathbf{h}_{t-1},\mathbf{r}]+\mathbf{b}_z),\mathbf{r}_t=\sigma(\mathbf{W}_r[\mathbf{h}_{t-1},\mathbf{r}]+\mathbf{b}_r) \quad (34)$$

The final hidden state is updated through a gating mechanism. The bidirectional LSTM layer is configured with 128 hidden units per direction to capture long-term dependencies, while the subsequent bidirectional GRU layer utilizes 64 hidden units to refine local dynamic features. To prevent overfitting, a dropout rate of 0.5 is applied after the recurrent layers, and the model is optimized using the Adam optimizer with an initial learning rate of 0.001 and a batch size of 32. The classification layer uses a time-distributed fully connected network, and the Softmax outputs the probability distribution of the six types of rehabilitation actions as shown in Formula (35):

$$\mathbf{P}_t=\text{Softmax}(\mathbf{W}_o\mathbf{h}_t^{(2)}+\mathbf{b}_o)\in R^6 \quad (35)$$

The model uses cross-entropy as the loss function, as shown in Formula (36):

$$L=-\frac{1}{T}\sum_{t=1}^{T}\sum_{c=1}^{6}y_{t,c}\log\ \ p_{t,c} \quad (36)$$

Fig. 2 illustrates the attention-enhanced LSTM-GRU hybrid network architecture proposed in this paper. The input is a 64-dimensional temporal feature sequence after feature selection. This is first mapped to a high-dimensional latent space through a fully connected layer. A bidirectional LSTM layer then captures the forward and backward temporal dependencies of the action sequence, concatenating them to form a latent state sequence containing contextual information. Building on this foundation, a temporal attention mechanism is introduced. By calculating the attention weight at each time step, it performs weighted focus on keyframes, generating a context vector that serves as the initial memory for the GRU layer. A bidirectional GRU layer then further extracts dynamic features, enhancing responsiveness to transient motion patterns. Finally, a temporally distributed fully connected layer and a softmax classifier output the probability distribution of six rehabilitation motion categories, which are then trained end-to-end using a cross-entropy loss function.

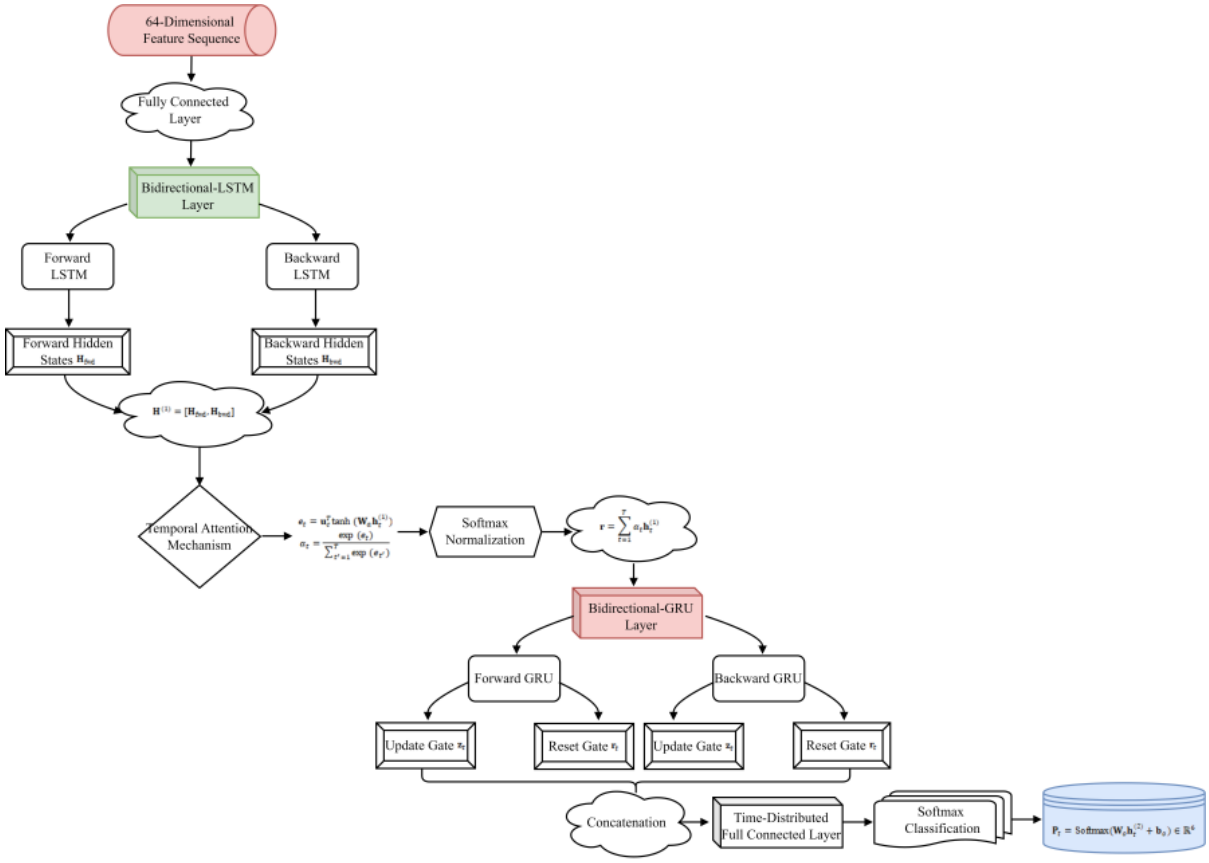

Figure 2: Network architecture

2.5 Rehabilitation movement semantic parsing, intelligent scoring, and personalized feedback

This paper constructs an intelligent scoring and feedback model based on action semantic parsing. This model aligns the underlying Action Recognition (AR) results with clinical rehabilitation prescriptions and integrates multidimensional behavioral indicators to establish a quantifiable dynamic scoring system, which subsequently generates personalized, real-time instructional feedback [37-38]. Assume that the system identifies the action sequence $A=\{a_1,a_2,\ldots,a_N\}$ within the monitoring period $T=[t_{\text{start}},t_{\text{end}}]$. The doctor's preset rehabilitation prescription P is defined as a triple: $P_c=(c,G_c,F_c,D_c),c\in C$, G_c is the target number of daily sets, F_c is the target number of times per set, and D_c is the standard duration of a single action. Perform semantic parsing on the recognition results, extract the subsequence $A_c=\{a_n|a_n=c\}$ belonging to category c in the actual execution sequence, and divide the training groups based on time interval clustering. Let the time interval between adjacent actions be $\Delta t_n=t_{n+1}-t_n$. If $\Delta t_n>T_{\text{rest}}$, it is considered as the boundary between groups. The actual number of completed groups G_c^{actual} and the total number of N_c^{actual} are respectively as shown in Formula (37):

$$G_c^{\text{actual}}=|\{\text{cluster}_k(A_c)\}|,N_c^{\text{actual}}=|A_c| \tag{37}$$

Define the four-dimensional compliance indicators: completion rate as Formula (38):

$$R_c=\frac{N_c^{\text{actual}}}{G_c\cdot F_c} \tag{38}$$

Normalization: The similarity between the action trajectory and the standard template is calculated based on the DTW distance and normalized as shown in Formula (39):

$$N_c^{\text{norm}}=\frac{1}{1+\exp\ (\gamma(d_{\text{dtw}}-d_0))},d_{\text{dtw}}=\text{DTW}_{\text{avg}}(\mathbf{q}^{(i)},\mathbf{q}_{\text{std}}) \tag{39}$$

The design of this standardization index aims to establish a correlation with clinical functional recovery. Frequency consistency: reflects the regularity of daily training time distribution as shown in Formula (40):

$$F_c^{\text{cons}}=\frac{1}{1+\alpha\cdot\sigma(T_{\text{daily}})} \tag{40}$$

The duration compliance rate is as shown in Formula (41):

$$D_c^{\text{rate}}=\frac{1}{N_c^{\text{actual}}}\sum_{i=1}^{N_c^{\text{actual}}}\text{I}\,(\Delta t_i\geq 0.8D_c) \tag{41}$$

Two optimization mechanisms are proposed to achieve an accurate evaluation that dynamically matches the patient's rehabilitation process. 1) Stage-adaptive dynamic weighting: The scoring weight is dynamically adjusted according to the patient's rehabilitation stage; 2) Trend reinforcement based on forgetting factor: The exponentially weighted moving average and forgetting factor are introduced to smooth the compliance indicators

for multiple consecutive days, reduce the impact of early historical data, and enhance sensitivity to recent training trends. The four-dimensional indicators after dynamic weighting and trend reinforcement are integrated to construct an adaptive compliance scoring model as shown in Formula (42):

$$S_c=100\cdot(w_1R_c+w_2N_c^{\text{norm}}+w_3F_c^{\text{cons}}+w_4D_c^{\text{rate}}) \quad (42)$$

Scoring results are categorized into three levels: $S_c<60$ indicates low compliance, triggering an alert and generating corrective teaching feedback; $60\leq S_c<80$ indicates moderate compliance, generating encouraging reminders and improvement suggestions; and $S_c\geq 80$ indicates high compliance, recording incentives, and providing further guidance. The system pushes both the structured score and personalized feedback instructions to mobile devices, supporting a closed-loop remote nursing intervention. With the theoretical framework and algorithmic design fully established, the subsequent section will validate the system's practical performance through rigorous experiments, focusing on recognition accuracy, compliance scoring consistency, and system latency under varying network conditions.

3 Experimentation and validation

3.1 Experimental design

Six standard rehabilitation exercises were performed in a home environment: elbow flexion, elbow extension, shoulder flexion, shoulder abduction, forearm pronation, and grip training. Dual-channel sEMG and three-axis IMU signals were collected simultaneously, and a dual-view HD camera was set up to record video as the basis for action ground truth annotation. All data was de-identified and then constructed into a dedicated dataset, which was divided into training, validation, and test sets in an 8:1:1 ratio. The dataset comprises data collected from 32 post-stroke patients (18 males, 14 females; mean age 62.4 ± 5.8 years) recruited from the rehabilitation department, all meeting the inclusion criteria of Brunnstrom recovery stages III to IV. Sensors were consistently placed on the muscle belly of the biceps brachii, triceps brachii, and the dorsal aspect of the forearm to ensure high-fidelity signal acquisition during the six prescribed exercises.

3.2 Action recognition accuracy test

The AR accuracy was calculated frame by frame based on the test set as shown in Formula (43):

$$\text{Accuracy}=\frac{\sum_{t=1}^{T}\text{I}(\hat{y}_t=y_t)}{T} \quad (43)$$

y_t is the ground truth value of the video, and $\hat{y}_t$ is the predicted label output by the model. The confusion matrix is used to statistically analyze the recognition performance of various actions and calculate the overall accuracy.

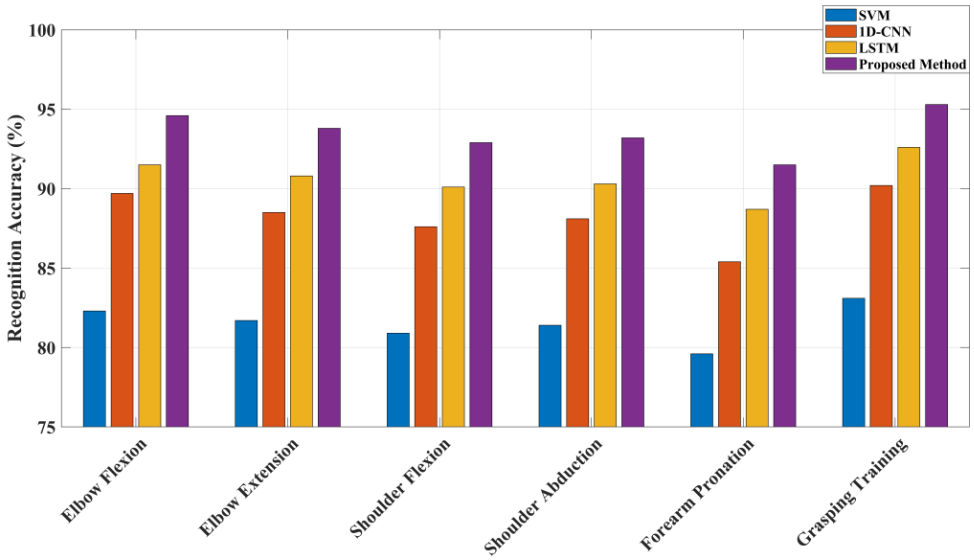

Figure 3: AR accuracy

Fig. 3 compares the recognition accuracy of six upper limb rehabilitation exercises using SVM, 1D-CNN (1 Dimensional-Convolutional Neural Network), LSTM, and the PM (Proposed Method). The horizontal axis represents six typical upper limb rehabilitation exercises: elbow flexion, elbow extension, shoulder flexion, shoulder abduction, forearm pronation, and grip training. The vertical axis represents recognition accuracy. PM achieved the highest accuracy across all action categories, achieving 94.6%, 93.8%, 92.9%, 93.2%, 91.5%, and 95.3%, respectively, with an overall average of 93.6%. This significantly outperformed SVM (79.6-83.1%), 1D-CNN (85.4-90.2%), and LSTM (88.7-92.6%). In the dynamic, complex, and compensatory forearm pronation movement, PM achieved a 2.8% improvement over LSTM and an 11.9% improvement over SVM, demonstrating greater robustness. The performance advantage of the proposed method stems from the synergistic integration of multimodal biosignals, where sEMG data provides critical information on muscle activation patterns that purely kinematic methods miss. The temporal attention mechanism further differentiates this solution from

standard LSTM or CNN models by dynamically assigning higher weights to key action frames, thereby effectively suppressing environmental noise and capturing subtle compensatory behaviors that distinguish high-quality rehabilitation movements.

This performance improvement stems from the synergistic effect of the proposed multimodal fusion and attention-enhancing mechanisms. The dual IMU and sEMG signal inputs provide complementary information on kinematics and physiological activation, enabling the system to more accurately distinguish between real movements and compensatory actions. This method utilizes a temporal attention mechanism to assign higher weights to key frames of action, effectively suppressing noise interference and enhancing the ability to capture dynamic features of the movement. A lower DTW distance clinically correlates with reduced pathological synergy and smoother joint coordination, serving as a quantitative indicator of improved motor control quality compared to the standard kinematic templates of healthy subjects. This metric effectively translates complex kinematic deviations into a unified standardization score that parallels the qualitative assessment criteria of the Fugl-Meyer Assessment, enabling the automated system to detect functional improvements that correspond to neural plasticity recovery. LSTMs lack a selective focus mechanism and are susceptible to redundant frames; CNNs are limited to a local receptive field and struggle to model long-term dependencies. This paper introduces motion trajectory curvature and DTW-related indicators at the feature level to further enhance the perception of movement quality. This provides high-fidelity input for subsequent intelligent compliance evaluation, with only non-judgment categories.

3.3 Quantitative evaluation of movement standardization

For each type of movement, the joint angle trajectory is extracted, and the DTW distance is calculated compared with the mean trajectory of the healthy control group, as shown in Formula (44):

$$\mathrm{DTW}(\theta,\theta_{\mathrm{std}})=\min_{\pi}\sum_{(i,j)\in\pi}\|\theta(i)-\theta_{\mathrm{std}}(j)\|^2 \tag{44}$$

Path π is the optimal alignment sequence. The optimization objective function ensures the minimization of cumulative Euclidean distances along the warping path π, which satisfies the boundary, continuity, and monotonicity constraints. This mathematically guarantees that the calculated distance represents the optimal global alignment between the patient's variable-speed motion sequence and the standard clinical template, providing a rigorous metric for non-linear temporal comparison.

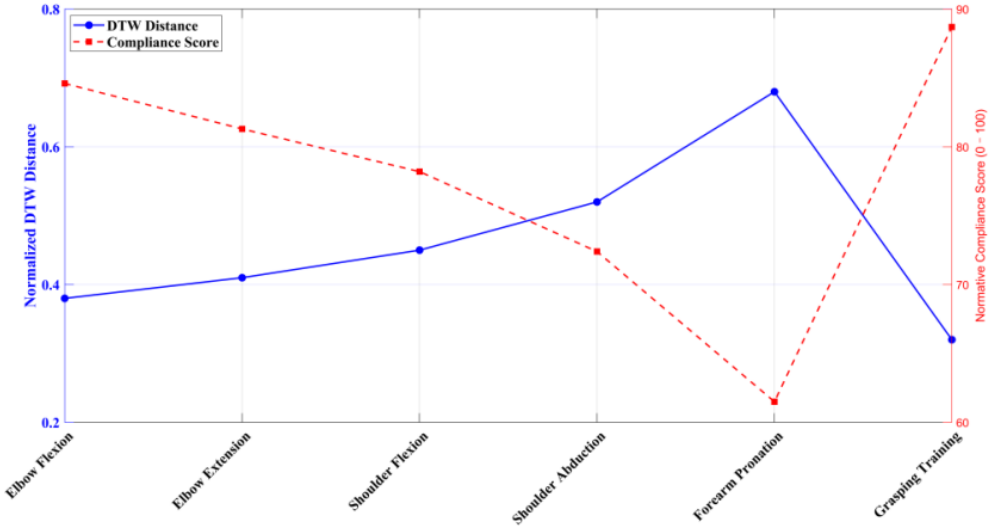

Figure 4: Quantitative assessment of movement standardization

Fig. 4 shows the relationship between normalized DTW distance and standardization score for six upper limb rehabilitation movement types. The horizontal axis represents elbow flexion, elbow extension, shoulder flexion, shoulder abduction, forearm pronation, and grip training. The left vertical axis represents normalized DTW distance, and the right vertical axis represents standardization score (0–100). The DTW distance was significantly negatively correlated with the standardization score, with lower DTW values associated with higher standardization scores. The gripping exercise had the smallest DTW distance (0.32) and the highest standardization score (88.7), indicating that this movement was most closely executed. The forearm pronation exercise had the highest DTW distance of 0.68, corresponding to a score of only 61.5, indicating the poorest execution quality. Shoulder abduction and flexion exercises also exhibited high DTW values (0.52 and 0.45) and low scores (72.4 and 78.2), reflecting the common difficulty controlling shoulder movements in stroke patients. The overall trend demonstrates that the system is capable of quantifying motion deviations using DTW and mapping them into interpretable clinical scores.

This assessment capability stems from the PM's deep integration of multimodal biosignals and kinematic modeling. The joint angle trajectories derived from IMU posture analysis have high temporal resolution. Combined with the DTW algorithm's flexible alignment capabilities

for non-isometric and asynchronous movements, this effectively overcomes matching challenges associated with slow movements and irregular starts and stops. During fine-control movements like forearm pronation, stroke patients often rely on shoulder-arm compensation due to insufficient activation of the pronator muscles, resulting in significant trajectory deviation and a significant increase in DTW distance. Specifically, in a documented case of a patient with moderate spasticity, the system detected a compensatory trunk lean of 15 degrees during the pronation task. This subtle deviation was successfully captured by the IMU's kinematic chain analysis, resulting in a distinct DTW spike that triggered an immediate corrective voice prompt, demonstrating the system's capacity to handle complex compensatory patterns that simple threshold-based methods miss. Grip training exercises are frequently used daily, retain better neuroplasticity, and have movement patterns closer to healthy models, resulting in lower DTW distances and higher scores. The standardization score utilizes a nonlinear mapping function that imposes a greater penalty on the mid- to high-DTW range, highlighting serious non-standard behaviors of clinical concern. This enhances the score's sensitivity and practicality, providing deeper quality support for adherence assessment beyond completion.

3.4 Compliance score consistency analysis

Two rehabilitation therapists independently manually scored the patients' 7-day training behavior based on video playback (maximum 100 points). The scoring criteria included the number of completions, movement quality, and execution regularity. The ICC between the system's automatic score and the mean of the manual score was calculated as shown in Formula (45):

$$\mathrm{ICC}(2,1)=\frac{\mathrm{MS}_R-\mathrm{MS}_W}{\mathrm{MS}_R+(k-1)\mathrm{MS}_W} \tag{45}$$

Table 1: Analysis of consistency of compliance scores

Assessment Method	Mean Score (± SD)	Inter-Rater ICC (Therapist A vs. B)	System vs. Mean Human ICC
Therapist A	76.3 ± 14.2	—	—
Therapist B	77.1 ± 13.8	0.87	—
System (Automated)	75.9 ± 15.1	—	0.89

Table 1 shows the results of the consistency analysis between the system's automatic scores and the manual scores of two rehabilitation therapists, using the ICC for quantitative evaluation. As can be seen from Table 1, the consistency between the two therapists' scores was 0.87, indicating high inter-expert agreement. The ICC between the system's automatic scores and the mean of the manual scores reached 0.89, slightly higher than the inter-therapist agreement level. The 95% confidence interval for the ICC was calculated to be $[0.82,0.93]$, indicating a tight agreement range. Furthermore, a Bland-Altman plot analysis revealed that 94% of the difference points fell within the 1.96 standard deviation limits of agreement, confirming that the automated scoring system possesses no significant systematic bias compared to manual expert evaluation. The average scores for the three groups were 76.3, 77.1, and 75.9, respectively, with standard deviations ranging from 13.8 to 15.1. This indicates that the system scores are not only numerically close to manual assessments, but also have controllable variability. This result confirms that the system has comparable stability and reliability to clinical experts in the adherence assessment task, and can objectively reflect the actual implementation of patients' home rehabilitation, providing a statistical basis for replacing or supplementing manual follow-up.

Unlike traditional coarse-grained statistics that rely solely on the number of movements, this method uses DTW to quantify movement standardization, combining completion rate, execution frequency, and duration, creating an evaluation framework that is highly consistent with rehabilitation therapists' scoring logic. The attention-enhanced LSTM-GRU network architecture accurately identifies movement start and end points and key phases, reducing misjudgments. The action-prescription mapping library ensures that scoring criteria strictly align with individualized rehabilitation plans, enhancing clinical interpretability. The system avoids the fatigue and subjective bias often associated with manual scoring, maintains stable output over seven consecutive days of monitoring, and achieves slightly better than expert-to-expert agreement, providing a new paradigm for scalable and reproducible intelligent nursing assessment for post-stroke home rehabilitation.

3.5 System response latency measurement

The end-to-end latency was recorded from the time the sensor acquires the first frame of data to the time the motion recognition result is returned to the cloud. Tests were conducted 500 times each over Wi-Fi and 4G networks, using high-precision timestamps to measure the duration of each stage: sensor acquisition, Bluetooth transmission, edge preprocessing, and cloud inference.

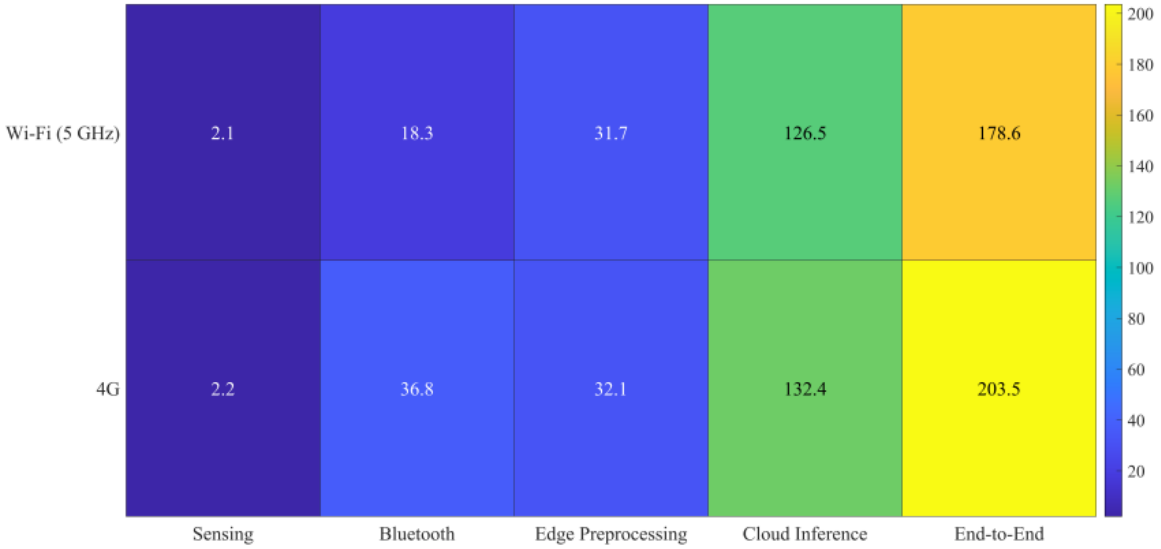

Figure 5: System response latency measurement

Fig. 5 shows the system response latency distribution at each stage, from signal acquisition to result return, in the form of a heat map under Wi-Fi and 4G network conditions. The horizontal axis represents the five stages: sensor acquisition, Bluetooth transmission, edge preprocessing, cloud inference, and end-to-end total latency, while the vertical axis corresponds to the two network environments. Wi-Fi enabled low latency across all stages, with a total end-to-end latency of 178.6ms, meeting the time window required for real-time closed-loop feedback. However, 4G enabled latency increased to 203.5ms, primarily due to the significant increase in Bluetooth transmission and cloud-based inference. Fig. 5 clearly illustrates the critical impact of network bandwidth on the remote inference link. During data upload, 4G communication became a performance bottleneck, causing some tasks to exceed real-time requirements.

Although signal preprocessing and feature extraction are completed at the edge, significantly reducing the amount of uploaded data, AR still relies on complex cloud-based models for high-precision inference, making it highly sensitive to network stability. In 4G environments, high transmission jitter and congestion control mechanisms extend the Bluetooth-cloud communication cycle, which is more pronounced in situations with multiple users or signal fluctuations. Wi-Fi offers lower latency and higher throughput, ensuring efficient collaboration between the edge and cloud. Cloud-based inference itself accounts for approximately 70.8% of total latency (126.5%/178.6% on Wi-Fi), indicating that there is still room for optimization in model lightweighting. In future deployments, edge inference offloading strategies can be combined to achieve end-cloud collaboration while ensuring accuracy, further improving the robustness and practicality of the system in diverse network environments. To specifically mitigate high latency in 4G environments, future iterations will implement model quantization techniques to deploy lightweight inference engines directly on edge nodes. This approach reduces reliance on continuous cloud uplink transmission, thereby maintaining real-time feedback capabilities even under constrained network bandwidth.

3.6 Multidimensional compliance index trend analysis

The actual training completion rate for each patient for 7 consecutive days was calculated, and the intra-week CV (Coefficient of Variation) was calculated as shown in Formula (46):

$$\mathrm{CV}=\frac{\sigma(R_c^{(1)},\ldots,R_c^{(7)})}{\mu(R_c^{(1)},\ldots,R_c^{(7)})}\times 100\% \tag{46}$$

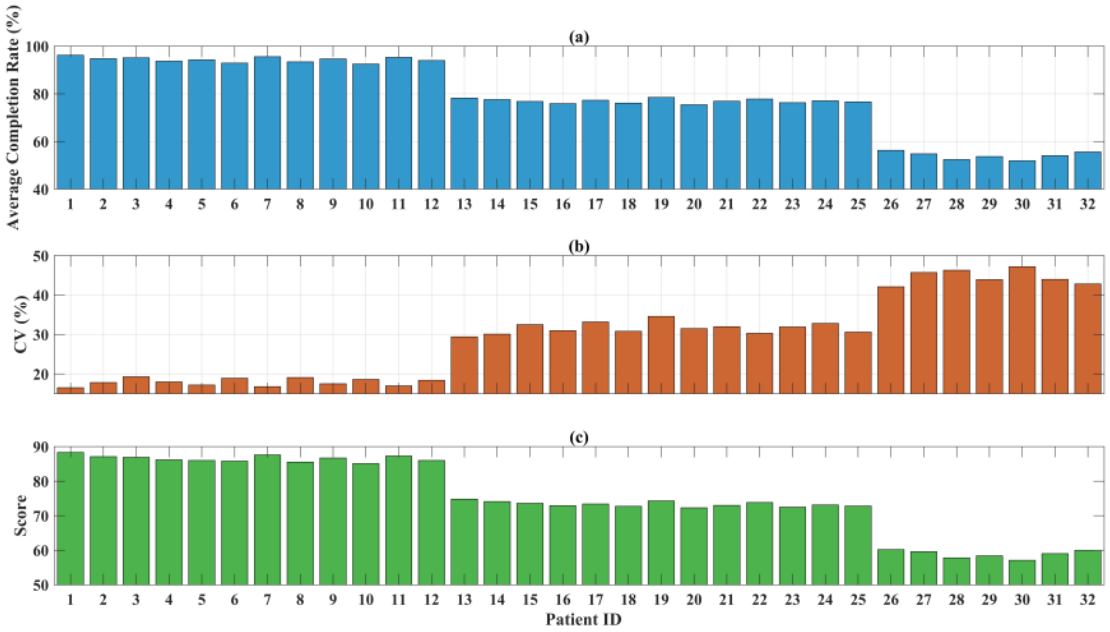

Figure 6: Multidimensional adherence trend analysis; (a): average completion rate, (b): CV, (c): adherence score

Fig. 6 shows the multi-day adherence trends of stroke patients during 7 consecutive days of home rehabilitation. The horizontal axis represents the patient number, and the vertical axis represents the mean completion rate, the intra-week CV, and the adherence score. The first 12 patients (1–12) had a completion rate generally above 92%, CVs mainly ranging from 16% to 19%, and scores exceeding 85, indicating high adherence. The middle 13 patients (13–25) had a completion rate between 75% and 79%, CVs ranging from 29% to 35%, and scores ranging from 72% to 75%, indicating a moderate level of adherence. The last 7 patients (26–32) had a completion rate below 57%, CVs as high as 42% to 48%, and scores below 60.2, indicating highly irregular behavior. The overall trends in Figs 6(a), 6(b), and 6(c) demonstrate a significant negative correlation between completion rate and CV, which jointly influence the final score, revealing significant inter-individual variability in adherence.

This hierarchical model reflects the system's ability to deeply capture the dynamic characteristics of rehabilitation behavior. Completion rate reflects whether training is completed, while CV reflects whether it is performed regularly. This method calculates CV based on daily fluctuations in completion rate, revealing stable differences in patients' self-management abilities. Adherence scoring integrates multiple metrics to avoid bias from a single dimension. This allows patients with similar completion rates but higher CVs to receive lower scores, demonstrating the sensitivity and refinement of the assessment. A dynamic modeling mechanism based on long-term trends enables the system to not only identify the extent of completion but also determine whether it is consistently and correctly performed. This provides data-driven support for personalized intervention strategies and promotes the evolution of home rehabilitation from an outcome-oriented approach to process management. The identification of distinct adherence patterns allows clinicians to intervene early with patients exhibiting high variance in training frequency, shifting the focus from passive monitoring to active behavioral modification. For the low-adherence group, the system suggests simplifying the rehabilitation prescription or increasing the frequency of gamified incentives to prevent the abandonment of therapy before functional decline occurs.

3.7 Verification of cross-user generalization

Leave-one-out cross-validation is used to evaluate the generalization performance of the model on unseen individuals. This validation strategy simulates the cold-start scenario where the system must recognize the movements of a new patient without prior calibration. The consistent performance metrics indicate that the model learns universal motor features rather than memorizing subject-specific patterns, ensuring the system's rapid deployability in clinical settings without requiring extensive personal data collection for every new user. The F1-score, precision, and recall of each test set are calculated as follows:

$$F1=2\cdot\frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}},\text{Precision}=\frac{TP}{TP+FP},\text{Recall}=\frac{TP}{TP+FN} \quad (47)$$

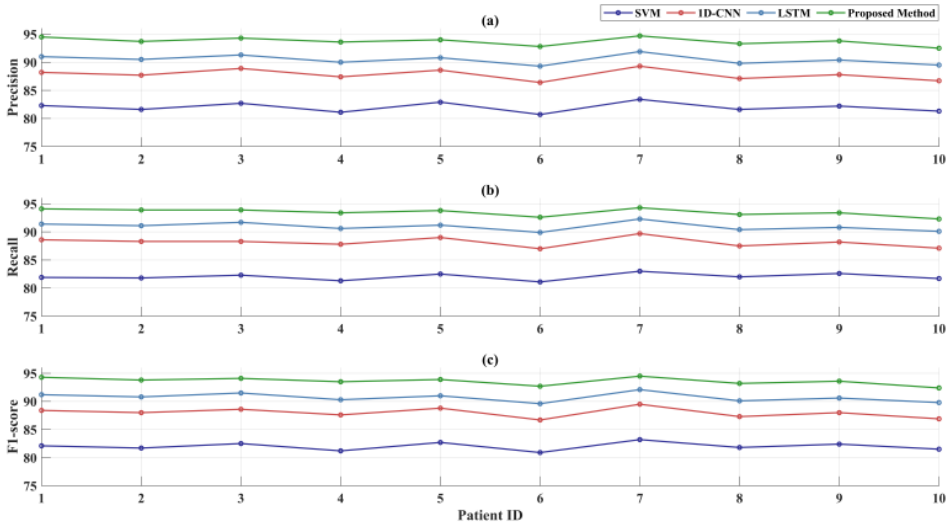

Figure 7: Cross-user generalization; (a): precision, (b): recall, (c): F1-score

Fig. 7 shows a comparison of the cross-user generalization performance of SVM, 1D-CNN, LSTM, and PM. The horizontal axis represents patient number, and the vertical axes represent precision, recall, and F1-score (%), respectively. PM significantly outperformed across all patients and all metrics: its precision remained consistently between 92.5% and 94.7%, its recall remained between 92.3% and 94.3%, and its F1-score consistently exceeded 92.4%, reaching a maximum of 94.5%. The SVM approach exhibited significant fluctuations, with a maximum F1-score of only 83.2%, representing the lowest overall performance. The 1D-CNN and LSTM approaches performed in the middle, but still showed significant declines when accounting for individual motion pattern variations. PM achieved the highest performance, demonstrating excellent robustness and consistency. Comparison with existing single-modality approaches highlights that the proposed multimodal fusion strategy maintains superior F1-scores across all subjects by leveraging the complementary nature of kinematic and physiological data. This resilience to inter-subject variability contrasts with the significant performance degradation observed in SVM and 1D-CNN models when processing data from patients with varying degrees of hemiplegia.

Multimodal biosensing, integrating kinematic and physiological information, provides the model with a deeper understanding beyond a single motion trajectory. In cross-user scenarios, different patients' motion trajectories often exhibit significant variations due to varying degrees of hemiplegia or compensation strategies. When performing the same target movement, core muscle activation patterns often exhibit greater physiological commonality. This characteristic enables the model to capture the relatively stable movement intent behind movement morphology. By utilizing sEMG signals as a key clue, the system can more accurately determine rehabilitation intentions when identifying new users' movements, significantly enhancing the model's cross-user generalization capabilities. This stable recognition of intentions ensures that the compliance assessment model is not invalidated by individual differences in execution, providing a reliable technical foundation for personalized, automated rehabilitation monitoring for a wide range of stroke patients.

3.8 Clinical validation

To verify the clinical validity of this system's adherence score and its role in promoting functional recovery, a 4-week prospective intervention study was conducted. Participants were assessed for functional ability using the FMA-UE (Fugl-Meyer Assessment - Upper Extremity). ANCOVA was used to analyze the effect of different adherence levels on the rate of functional recovery. The study followed a prospective cohort design involving 32 participants stratified by their system-assessed adherence levels over a 4-week intervention period. To isolate the effect of adherence on recovery, the Analysis of Covariance (ANCOVA) treated the baseline FMA-UE score as a covariate, ensuring that the observed improvements in the Week-4 assessment were statistically attributable to the training compliance rather than initial functional differences.

Table 2: Clinical validation

Adherence Level	Baseline FMA-UE	Week-4 FMA-UE	FMA-UE Improvement
High Adherence	42.3	56.5	14.2
Medium Adherence	43.1	51.8	8.7
Low Adherence	41.9	47.1	5.2
Statistical Test	F-value	p-value	Partial η^2
	32.47	< 0.001	0.68

Table 2 clearly demonstrates the significant differences in functional recovery outcomes among patients with different adherence levels during the 4-week rehabilitation intervention. The baseline functional levels of the three groups were similar (41.9-43.1 points), ensuring relative fairness. At the end of the intervention, the FMA-UE score (56.5 points) in the high-adherence group was significantly higher than that in the medium-adherence group (51.8 points) and the low-adherence group (47.1 points). The magnitude of improvement (14.2 points) also far exceeded that in the medium-adherence group (8.7 points) and the low-adherence group (5.2 points), demonstrating a clear gradient effect. The ANCOVA analysis results (F-value 32.47, $p < 0.001$, partial η^2 0.68) confirmed the statistical significance and large effect size of this difference, demonstrating that the system's automatically assigned adherence level is a strong predictor of patients' functional recovery.

This significant difference in functional recovery is primarily due to the system's use of multimodal IMU and sEMG sensing, which not only determines whether the patient has performed the movement but also quantifies its accuracy using the DTW algorithm, accurately identifying and correcting ineffective or incorrect movements such as compensatory shrugs. Patients in the high-adherence group received this high-quality feedback continuously, and their training process was more aligned with the requirements for neuroplasticity reconstruction. Each movement more effectively promoted the remodeling of functional brain areas. Patients in the low-adherence group, perhaps due to a lack of effective supervision, frequently resorted to compensatory strategies, resulting in poor training quality and significantly reduced neuroplasticity efficiency.

4 Conclusions

This study proposed a biosensor-based intelligent method for movement recognition and NC assessment in post-stroke home rehabilitation. The system collects upper limb movement and electromyographic signals through multimodal wearable sensors. After edge synchronization preprocessing, the attention-enhanced LSTM-GRU hybrid deep model is used to achieve accurate recognition of six types of rehabilitation movements. It also innovatively integrates completion rate, DTW-based standardization, frequency consistency, and duration compliance rate to construct a multi-dimensional dynamic compliance scoring system. Experimental results demonstrate that this method achieves an average motion recognition accuracy of 93.6%, and the ICC between the system's automatic scoring and manual scoring by rehabilitation therapists reaches 0.89, achieving refined, objective, and intelligent evaluation. This study provides a new paradigm for scalable, reproducible, high-precision, and robust remote monitoring of stroke patients, potentially enabling personalized, closed-loop rehabilitation management in the home setting, improving rehabilitation outcomes, and reducing medical costs. Beyond individual home use, this architecture can be deployed in community health centers to facilitate remote triage, allowing therapists to manage multiple patients simultaneously through the cloud platform. The standardized scoring mechanism also supports the integration of this system into telemedicine insurance reimbursement policies by providing objective evidence of rehabilitation adherence. Future research will focus on integrating 5G-enabled edge computing modules to further reduce end-to-end latency for tactile feedback applications. Additionally, we plan to incorporate the aforementioned lattice-based cryptographic protocols to enhance data privacy and conduct large-scale, multi-center clinical trials to verify the system's long-term impact on neural plasticity recovery.

Authorship contribution statement

Tianhui Li: Writing-Original draft preparation, Conceptualization, Supervision, Project administration.

Data availability

On Request

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Author Statement

The manuscript has been read and approved by all the authors, the requirements for authorship, as stated earlier in this document, have been met, and each author believes that the manuscript represents honest work.

Ethical approval

All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.

References

[1] Y.-C. Chen, W. Chou, R.-B. Hong, J.-H. Lee, and J.-H. Chang (2023). Home-based rehabilitation versus hospital-based rehabilitation for stroke patients in post-acute care stage: Comparison on the quality of life, *Journal of the Formosan Medical Association*, 122(9), pp. 862–871, https://doi.org/10.1016/j.jfma.2023.05.007

[2] P. Langhorne, F. Coupar, and A. Pollock (2009). Motor recovery after stroke: a systematic review, The Lancet Neurology, 8(8), pp. 741–754,

[3] P. Babbar, K. V. Kumar, P. T. A. Joshua, M. Chakrapani, and Z. K. Misri (2021). Adherence to home-based neuro-rehabilitation exercise program in stroke survivors, *Bangladesh Journal of Medical Science*, 20(1), pp. 145–153, https://doi.org/10.3329/bjms.v20i1.50361

[4] I. Senadheera *et al.* (2024). AI applications in adult stroke recovery and rehabilitation: a scoping review using AI, *Sensors (Basel)*, 24(20), p. 6585, https://doi.org/10.3390/s24206585

[5] S. Rahman, S. Sarker, A. K. M. N. Haque, M. M. Uttsha, M. F. Islam, and S. Deb (2022). AI-driven stroke rehabilitation systems and assessment: A systematic review, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31, pp. 192–207, https://doi.org/10.1109/TNSRE.2022.3219085

[6] R. Proffitt, M. Ma, and M. Skubic (2023). Development and testing of a daily activity recognition system for post-stroke rehabilitation, *Sensors*, 23(18), p. 7872, https://doi.org/10.3390/s23187872

[7] G. Gebreheat, A. Goman, and A. Porter-Armstrong (2024). The use of home-based digital technology to support post-stroke upper limb rehabilitation: A scoping review, *Clin Rehabil*, 38(1), pp. 60–71, https://doi.org/10.1177/02692155231189257

[8] K. Olk and S. Persaud (2021). Repeat After Me: Stroke rehab technologies Leverage repetition and natural strength to optimize neuroplasticity., *Rehab Manag*, 34(2), pp. 12–14,

[9] C. D. Rocha, I. Carneiro, M. Torres, H. P. Oliveira, E. J. S. Pires, and M. F. Silva (2025). Post-stroke upper limb rehabilitation: clinical practices, compensatory movements, assessment, and trends, *Progress in Biomedical Engineering*, 7(4), p. 042001, DOI: 10.1088/2516-1091/adeb1e

[10] S. Peznola, L. V Gauthier, M. C. Claypool, B. Roop, and A. C. Lammert (2023). Compliance with in-home self-managed rehabilitation post-stroke is largely independent of scheduling approach, *Arch Phys Med Rehabil*, 104(4), pp. 554–561, https://doi.org/10.1016/j.apmr.2022.10.007

[11] C. Zhou, D. Feng, S. Chen, N. Ban, and J. Pan (2024). Portable vision-based gait assessment for post-stroke rehabilitation using an attention-based lightweight CNN, *Expert Syst Appl*, 238, p. 122074, https://doi.org/10.1016/j.eswa.2023.122074

[12] X. Wen *et al.* (2025). Enhancing long-term adherence in elderly stroke rehabilitation through a digital health approach based on multimodal feedback and personalized intervention, *Sci Rep*, 15(1), p. 14190, https://doi.org/10.1038/s41598-025-95726-z

[13] R. Maskeliūnas, R. Damaševičius, T. Blažauskas, C. Canbulut, A. Adomavičienė, and J. Griškevičius (2023). BiomacVR: A virtual reality-based system for precise human posture and motion analysis in rehabilitation exercises using depth sensors, *Electronics (Basel)*, 12(2), p. 339, https://doi.org/10.3390/electronics12020339

[14] F. Jin, M. Zou, X. Peng, H. Lei, and Y. Ren (2023). Deep learning-enhanced internet of things for activity recognition in post-stroke rehabilitation, *IEEE J Biomed Health Inform*, 28(7), pp. 3851–3859, https://doi.org/10.1109/JBHI.2023.3332735

[15] V. Bijalwan, V. B. Semwal, G. Singh, and T. K. Mandal (2023). HDL-PSR: Modelling spatio-temporal features using hybrid deep learning approach for post-stroke rehabilitation, *Neural Process Lett*, 55(1), pp. 279–298, https://doi.org/10.1007/s11063-022-10744-6

[16] I. Hamid and B. Utomo (2025). Improving the Accuracy of Upper Limb Exoskeleton Movements Using Machine Learning Based on EMG for Stroke Patient Rehabilitation, *Jurnal Teknokes*, 18(1), pp. 18–31, https://doi.org/10.35882/76cntw33

[17] J. Xie, Q. Wu, N. Dey, F. Shi, R. S. Sherratt, and Y. Kuang (2025). Empowering stroke recovery with upper limb rehabilitation monitoring using TinyML based heterogeneous classifiers, *Sci Rep*, 15(1), p. 18090, https://doi.org/10.1038/s41598-025-01710-y

[18] C. Tang *et al.* (2024). An AI-driven multimodal smart home platform for continuous monitoring and intelligent assistance in post-stroke patients, *arXiv preprint arXiv:2411.19000*, https://doi.org/10.48550/arXiv.2411.19000

[19] M. A. Khan, H. U. Rahman, and M. Al-Ghamdi (2024). A Novel DDoS Mitigation Strategy in 5G-Based Vehicular Networks Using Chebyshev Polynomials, IEEE Transactions on Intelligent Transportation Systems, 25(1), pp. 112–125, https://doi.org/10.1007/s13369-023-08535-9

[20] S. U. Rehman, M. A. Khan, T. A. Zia, and T. Xie (2023). FCA-VBN: Fog computing-based authentication scheme for 5G-assisted vehicular blockchain network, IEEE Internet of Things Journal, 10(19), pp. 17234–17246, https://doi.org/10.1016/j.iot.2024.101096

[21] M. A. Khan, S. R. A. Jafri, and A. Al-Dhaqm (2024). Chebyshev Polynomial Based Emergency Conditions With Authentication Scheme for 5G-Assisted Vehicular Fog Computing, IEEE Transactions on Vehicular Technology, 73(2), pp. 2451–2463, https://doi.org/10.1109/TDSC.2025.3553868

[22] A. R. Rajput, Q. Li, M. T. Ahvanooey, and I. Masood (2023). Review of Prevention schemes for Replay Attack in Vehicular Ad hoc Networks (VANETs), IEEE Access, 11, pp. 1450–1465, https://doi.org/10.1109/ICICSP50920.2020.9232047

[23] J. Zhang, T. Xu, and Y. Zhang (2024). L-CPPA: Lattice-based conditional privacy-preserving authentication scheme for fog computing with 5G-enabled vehicular system, IEEE Systems Journal, 18(1), pp. 560–571, https://doi.org/10.1371/journal.pone.0292690

[24] M. A. Khan, S. Kumar, and S. A. Chaudhry (2023). ECA-VFog: An efficient certificateless authentication scheme for 5G-assisted vehicular fog computing, Ad Hoc Networks, 140, p. 103052, https://doi.org/10.1371/journal.pone.0287291

[25] Z. Zhao *et al.* (2023). Multimodal sensing in stroke motor rehabilitation, *Advanced Sensor Research*, 2(9), p. 2200055, https://doi.org/10.1002/adsr.202200055

[26] K. Xia *et al.* (2025). Exploratory development of human–machine interaction strategies for post-stroke upper-limb rehabilitation, *J Neuroeng Rehabil*, 22(1), p. 144, https://doi.org/10.1186/s12984-025-01680-2

[27] C.-H. Shih, P.-J. Lin, Y.-L. Chen, and S.-L. Chen (2024). A post-stroke rehabilitation system with compensatory movement detection using virtual reality and electroencephalogram technologies, *IEEE*

Access, 12, pp. 61418–61432, https://doi.org/10.1109/ACCESS.2024.3392513

[28] J. Chen *et al.* (2023). fNIRS-EEG BCIs for motor rehabilitation: a review, *Bioengineering*, 10(12), p. 1393, https://doi.org/10.3390/bioengineering10121393

[29] A. Noor *et al.* (2021). Decoding of ankle joint movements in stroke patients using surface electromyography, *Sensors*, 21(5), p. 1575, https://doi.org/10.3390/s21051575

[30] M. Y. T. Sulistyono, E. S. Pane, E. M. Yuniarno, and M. H. Purnomo (2024). Correlation Analysis Approach Between Features and Motor Movement Stimulus for Stroke Severity Classification of EEG Signal Based on Time Domain, Frequency Domain, and Signal Decomposition Domain, *Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI*, 13(3), pp. 597–611, https://doi.org/10.23887/janapati.v13i3.85550

[31] J. Sung *et al.* (2022). Classification of stroke severity using clinically relevant symmetric gait features based on recursive feature elimination with cross-validation, *Ieee Access*, 10, pp. 119437–119447, https://doi.org/10.1109/ACCESS.2022.3218118

[32] K. Apostolidis *et al.* (2024). Machine Learning Algorithms for the Prediction of Language and Cognition Rehabilitation Outcomes of Post-stroke Patients: A Scoping Review, *Human-Centric Intelligent Systems*, 4(1), pp. 147–160, https://doi.org/10.1007/s44230-023-00051-1

[33] H. Cao, Q. Chen, and S. Wang (2025). A novel spatio-temporal attention-based graph model for patient rehabilitation behavior recognition, *Signal Image Video Process*, 19(9), p. 764, https://doi.org/10.1007/s11760-025-04366-3

[34] F. Yang, S. Li, C. Sun, X. Li, and Z. Xiao (2024). Action recognition in rehabilitation: combining 3D convolution and LSTM with spatiotemporal attention, *Front Physiol*, 15, p. 1472380, https://doi.org/10.3389/fphys.2024.1472380

[35] S. Rani, S. Masood, and D. R. Rizvi (2025). A Bi-LSTM based approach for Compensation detection during robotic stroke rehabilitation therapy, *Procedia Comput Sci*, 258, pp. 3145–3154, https://doi.org/10.1016/j.procs.2025.04.572

[36] J. Bai, Z. Wang, X. Lu, and X. Wen (2023). Improved spatial–temporal graph convolutional networks for upper limb rehabilitation assessment based on precise posture measurement, *Front Neurosci*, 17, p. 1219556, https://doi.org/10.3389/fnins.2023.1219556

[37] C. Wang, L. Peng, Z.-G. Hou, P. Zhang, and P. Fang (2023). An easy-to-use assessment system for spasticity severity quantification in post-stroke rehabilitation, *IEEE Trans Cogn Dev Syst*, 16(3), pp. 828–839, https://doi.org/10.1109/TCDS.2023.3304352

[38] X. Wang, J. Zhang, S. Q. Xie, C. Shi, J. Li, and Z.-Q. Zhang (2024). Quantitative upper limb impairment assessment for stroke rehabilitation: A review, *IEEE Sens J*, 24(6), pp. 7432–7447, https://doi.org/10.1109/JSEN.2024.3359811