4th International Conference on Electronics, Engineering, Computer Science and Applied Development (EESD 2026)

Article

Personalized Rehabilitation Assistive Device Parameter Design by Integrating Human Motion-Capture Data and Artificial Intelligence

Jianwei Zhang[1,*]

[1] School of Nursing, Jilin University, Changchun, China
* Correspondence: Jianwei Zhang, School of Nursing, Jilin University, Changchun, China

Abstract: Personalized rehabilitation assistive devices need to set parameters based on an individual's joint mobility ability, movement performance, and compensatory behaviors. Human motion capture data can provide objective kinematic information, but current artificial intelligence methods are mostly used for motion recognition and motion quality assessment, and are less used for determining device parameters. To address this issue, this study proposes PMAD-Net, a personalized motion assistance design network that combines three-dimensional skeletal trajectories and inertial measurement data. The network first aligns the two types of data onto the same time axis, then uses a spatiotemporal encoder to analyze the coordination relationships between joints, and combines individual movement characteristics for fusion. The model can be set specifically according to the movement conditions of the subjects, including the range of motion of the shoulders and elbows, auxiliary gain, virtual stiffness, the starting time and duration of the movement. The predicted results also need to meet the biomechanical constraints to avoid unreasonable settings. PMAD-Net can adjust the device settings based on the subject's movement performance, covering shoulder and elbow range of motion, auxiliary gain, virtual stiffness, as well as the start and duration of the movement. To ensure that the settings are in line with the actual movement conditions, the model's output will also undergo biomechanical constraint checks. The experiment was conducted based on the public rehabilitation movement dataset, and ten independent subjects were selected for testing. The parameters of PMAD-Net have an MAE of 0.078 ± 0.006, a motion tracking RMSE of 5.73 ± 0.39, and a constraint violation rate of 2.9 ± 0.5%. Compared to the time transformer, these three indicators have decreased by 14.3%, 10.6%, and 39.6% respectively. This method can provide equipment setting suggestions for professionals and assist in the personalized adjustment of rehabilitation equipment.

Keywords: Human motion capture; Personalized rehabilitation; Assistive-device parameter design; Spatiotemporal graph learning; Biomechanical safety constraints

Received: 01 July 2026
Revised: 19 August 2026
Accepted: 01 September 2026
Published: 05 September 2026

1. Introduction

Rehabilitation assistive devices such as adjustable orthoses, robot trainers and wearable support systems are used in rehabilitation training to help patients regain their motor functions [1]. The actual effectiveness of the equipment is closely related to the parameter settings. The range of motion, assistive force, virtual stiffness and driving timing all need to be adjusted according to the specific conditions of the patients [2]. Currently, this task is mainly carried out by therapists. This process is relatively dependent on personal experience and is also time-consuming. When patients have joint movement limitations, abnormal movement speeds, asymmetry in left and right movements or compensatory movements, parameter settings are often more difficult [3].

The human motion capture technology provides a data basis for the equipment parameters. Through the three-dimensional skeletal data, the position changes of joints during the movement process and the coordination between different joints can be observed; the inertial measurement unit can record acceleration and angular velocity, and can still supplement the motion information even when some visual information is missing [4]. Artificial intelligence models can also capture temporal variations and inter-joint connections that are difficult to reflect in range of motion measurements from these data [5]. However, existing artificial intelligence research in the rehabilitation field mostly focuses on action classification, correctness judgment, and motion quality scoring [6], and rarely further determines device parameters based on movement performance.

Currently, there are three main issues in this field. There are differences in sampling rate, coordinate form, and missing data between skeletal data and inertial data, and direct fusion may affect the results. However, most of the existing motion assessment models only stop at judging the movement performance, and rarely further answer how the equipment parameters should be adjusted. Moreover, if the prediction process has no corresponding constraints, the model may also give parameters beyond the actual movement range, or produce inappropriate combinations in terms of assistance force, stiffness and driving timing. These factors still restrict the practical application of motion data in the personalized settings of rehabilitation equipment.

To address the aforementioned issues, this study proposes PMAD-Net, which is used to assist in determining the parameters of rehabilitation equipment based on an individual's movement situation. The research first synchronizes and scales the bone and inertia data, and retains the information of sensor absence to reduce the differences between different data sources. Subsequently, the model analyzes joint movements by combining the movement characteristics of the subject and predicts six equipment parameters. The effectiveness of the model is verified by comparing it with four baseline methods and through experiments without individual features. For the problem of unreasonable parameters, the study adds biomechanical constraints in the output stage and uses a feedback mechanism to adjust the parameters. The relevant experiments were mainly designed to test the performance of the model when data is missing, as well as to determine whether the security module can reduce the output of unreasonable parameters.

PMAD-Net starts from motion data and provides personalized parameter suggestions for rehabilitation equipment. The model incorporates safety constraints during the process of handling data and generating parameters to reduce unreasonable results. It should be noted that the system output is for professional reference only and is not used for independently formulating clinical prescriptions. The study also considered the possible data missing issues in actual use and tested the stability and rationality of the model results. In practical applications, it is still necessary to combine anonymized data processing, setting of safety ranges, and professional review.

2. Related Works

2.1. Human Motion Capture for Rehabilitation Assessment

Human motion capture can provide relatively objective motion data for rehabilitation training, which is used to record joint movements and their coordination situations [7]. The common collection methods currently include optical systems, depth cameras, and inertial measurement units. The optical system can provide relatively accurate spatial information, but it has high requirements for the experimental environment and equipment calibration. The deployment of depth cameras is relatively simple, but during actual collection, they are prone to interference from obstructions and changes in viewing angle. The applicable scenarios for different collection methods vary. The spatial measurement of the optical system is relatively accurate, but it usually requires a stable experimental environment and prior calibration. The depth camera has relatively flexible requirements for the shooting position, but body obstruction or changes in the shooting angle may affect the bone data. The inertial measurement unit is more convenient for movement and wearing, but changes in the sensor position and signal drift can introduce certain errors [8]. Therefore, in actual research, it is necessary to select the appropriate collection method based on the specific scenario.

Different sensors vary in sampling frequency, time delay, and data missing situations. Some existing studies process various types of data separately before fusing them, while others directly assume that the data is already synchronized, which may ignore short-term deviations during the movement process [9]. Therefore, before fusing multi-source motion data, it is necessary to first handle the differences in time and scale.

2.2. Artificial Intelligence for Movement Evaluation

Artificial intelligence methods have been widely used in rehabilitation motion analysis. Recurrent neural networks, time convolution networks, and graph models can learn the changes in actions and the relationships between joints from continuous motion data, and the attention mechanism can also help the model identify the motion information that has a significant impact on the results [10]. Compared w

However, most existing models are used to determine the type of action, correctness, or movement quality, and are less likely to further guide the parameter adjustment of assistive devices [11]. The same motion quality problems may come from different movement restrictions, so the corresponding assistive methods may also be different. Moreover, many models mainly learn from group data and do not consider the individual's physical condition and basic movement ability [12]. This makes it difficult to directly apply the model's evaluation results to personalized device parameter settings.

2.3. Personalized Rehabilitation Assistive-Device Design

Personalized rehabilitation equipment can adjust the auxiliary parameters according to the user's movement situation. Existing research mainly adopts rule-based setting, optimization search or data-driven methods [13,14]. The rule-based method is easy to understand but relies on manual experience; the optimization method usually requires multiple trials and adjustments, and the calibration process is relatively cumbersome [15]; the data-driven method can automatically adjust the degree of assistance based on the movement data, but it is prone to be limited by the equipment and the usage environment. In addition, the existing methods pay insufficient attention to the safety of the prediction results, and the model may provide parameters beyond the individual's activity range. The performance of the system when the sensor data is incomplete and the parameter prediction is not

Currently, motion analysis and equipment control are two relatively independent directions. Motion analysis is usually used to judge the quality of the movement, while equipment control relies more on the equipment data and repeated trials to determine the parameters. How to directly obtain suitable equipment settings for individuals using human motion data, while ensuring that the results are consistent with actual movement and equipment limitations, still lacks a relatively complete solution.

2.4. Research Gap and Positioning of This Study

PMAD-Net combines skeletal and inertial data, predicts equipment parameters based on the movement characteristics of the subject, and incorporates biomechanical constraints in the output stage. The model also utilizes motion information from different sources to reduce the output of unreasonable parameters. The system provides parameter suggestions, which are reviewed and adjusted by professionals, and are not used for independently formulating clinical prescriptions. The system outputs parameter suggestions for professionals to refer to. These suggestions need to be reviewed and adjusted manually and cannot be directly used for clinical prescriptions. The model's performance is compared through baseline models, ablation experiments, and hold-out set tests, and its performance is also examined in cases of data missing and constraint violations.

The model processes the bone and inertia data on a unified time scale and combines individual body scale to adjust the data representation, reducing the deviation between the two types of data and the impact of partial data missing.The subject-specific spatial-temporal encoder is used to extract joint motion information and combine with individual movement ability to predict six equipment parameters. After parameter prediction, the feasible set projection of biomechanics is added, and the output is adjusted according to the individual activity range and equipment operation restrictions. The feedback mechanism is used for subsequent parameter updates to improve the stability of the results.

3. Methodology

3.1. Overall Architecture

PMAD-Net is a personalized motion assistance design network, designed to predict the parameters of rehabilitation equipment using motion capture data from various sources. The model links motion analysis with parameter prediction and adjusts based on individual motion characteristics. The preliminary prediction results are subject to biomechanical constraints and corrected by incorporating motion feedback, making the parameters more in line with actual movement conditions.

As shown in Figure 1, PMAD-Net consists of five stages. Firstly, the three-dimensional skeletal trajectories and inertial measurement unit signals are synchronized and normalized according to the body scale. Subsequently, the model extracts motion features from both skeletal data and inertial data, with the former used to analyze the relationships between joints and the latter used to capture changes in acceleration and angular velocity. After fusion of these two types of features, they are combined with the individual's own motion characteristics for parameter prediction. The model outputs six equipment parameters, including the range of motion of the shoulder and elbow, assistance force, virtual stiffness, action start time, and duration. Finally, the predicted results are adjusted according to the individual's range of motion and equipment limitations, and the parameters are further corrected using motion feedback.

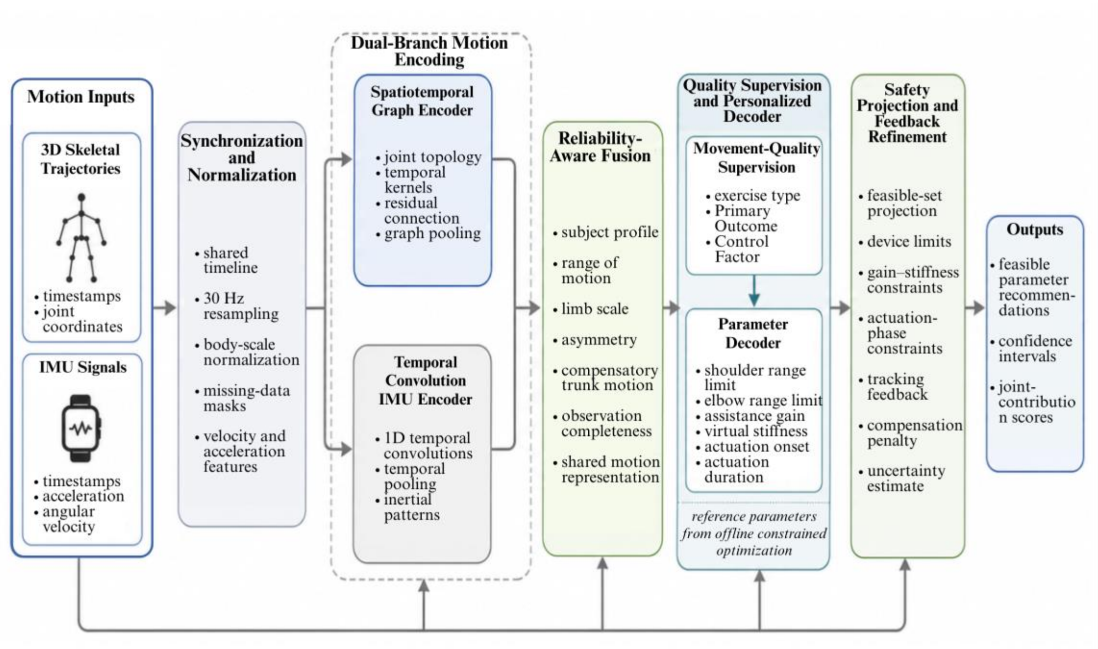

Figure 1. Overall Architecture of PMAD-Net for Personalized Rehabilitation Assistive-Device Parameter Design.

3.2. Cross-Sensor Synchronization and Normalization

For sample i, the skeletal data is denoted as $X_i^s \in \mathbb{R}^{T_s \times J \times 3}$, where T_s represents the number of sampling frames and J is the number of joints. The inertial data is denoted as $X_i^u \in \mathbb{R}^{T_u \times C}$, where T_u indicates the number of sampling points and C represents the number of channels of accelerometers and gyroscopes. After the two types of data are temporally aligned, they are uniformly mapped onto the timeline $\{\tau_t\}_{t=1}^{T}$:

$$\tilde{X}_{i,t}^m = M_{i,t}^m \odot \mathcal{J}_m(X_i^m, \tau_t), m \in \{s, u\}, \quad (1)$$

In the formula, $\mathcal{J}_m(\cdot)$ represents interpolation operation, $M_{i,t}^m$ is used to indicate whether there is an observation value at the corresponding position, and $\odot$ represents element-wise multiplication. Both types of signals are ultimately resampled to 30 Hz. When there are less than 5 consecutive missing frames, interpolation is used to complete the data; when the missing frames reach 5 or more, the corresponding mask is retained and no filling is performed.

The skeletal data is based on the spine base as the reference point, and the body scale is adjusted according to the shoulder width or the length of the upper arm. The joint speed and acceleration are further calculated based on the processed bone positions. Feature normalization only uses the statistics of the training set, avoiding the information of the test data from entering the training process.

3.3. Spatiotemporal Joint Encoding

The skeleton is represented as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ represents the joints and E represents the anatomical connections and movement relationships between the joints. Let $H_i^{(l)}$ be the feature representation of the l-th layer, and its update formula is:

$$H_i^{(l+1)} = \sigma\left[\sum_{k=1}^{K} \widetilde{A}_k H_i^{(l)} W_k^{(l)} + TCN_l\left(H_i^{(l)}\right)\right], \quad (2)$$

$\widetilde{A}_k$ represents the normalized adjacency matrix, and $W_k^{(l)}$ are the trainable parameters in the model. $TCN_l(\cdot)$ is used for time convolution, and $\sigma(\cdot)$ adopts the GELU activation function.

The skeletal encoder consists of two layers of graph convolutions with channel numbers of 64 and 128 respectively. Each layer uses a time convolution kernel of five frames, and residual connections, layer normalization, and 0.2 Dropout are added to capture the changes between joints and during the movement process.

3.4. Reliability-Aware Multimodal Fusion

The IMU branch consists of three one-dimensional time convolutions with channel numbers of 64, 96, and 128 respectively. After time pooling and graph pooling, the inertial representation h_i^u and the skeletal representation h_i^s are obtained.

The individual feature vector a_i is used to describe the movement situation of the subject itself, including normalized limb length, observable range of motion of the shoulder and elbow, movement speed, bilateral movement differences, and trunk compensation. The fusion process of multimodal features is defined as:

$$\begin{aligned} \alpha_i &= \sigma(W_\alpha[h_i^s \parallel h_i^u \parallel a_i \parallel \rho_i] + b_\alpha), \\ z_i &= \alpha_i \odot h_i^s + (1 - \alpha_i) \odot h_i^u, \end{aligned} \quad (3)$$

Here, ρ_i represents the observation proportion of each modality, $\parallel$ denotes feature concatenation, and z_i is the fused representation. This approach can reduce the influence of the modality with a large amount of missing data on the final result.

3.5. Movement-Quality Supervision and Training Labels

The model training uses information such as exercise type, continuous exercise quality score, individual exercise characteristics, and reference device parameters. The main outcome indicators and control factor scores in the dataset are normalized based on the statistics of the training set and are used for the quality regression part. For samples with incomplete score records, the model still retains their exercise characteristics, but the missing scores are not included in the regression calculation.

The model simultaneously employs supervised contrastive learning and quality regression:

$$\mathcal{L}_{\mathrm{rep}} = -\frac{1}{N}\sum_{i=1}^{N}\frac{1}{|P(i)|}\sum_{p\in P(i)}\log\frac{\exp(e_i^{T}e_p/\tau)}{\sum_{a\neq i}\exp(e_i^{T}e_a/\tau)} + \frac{\lambda_q}{N}\sum_{i=1}^{N} m_i\,(\hat{q}_i - q_i)^2, \quad (4)$$

Among them, N represents the batch size, e_i is the normalized projection of z_i , and P(i) indicates the set of samples belonging to the same motion type and having the same normalized quality score within the training set interval. τ is the temperature coefficient, q_i and $\hat{q}_i$ represent the actual score and predicted score respectively, and m_i is used to indicate whether the score is available.

The mod $\sum_{i=1}^{N} m_i$ s movement performance through contrastive learning, and then uses quality regression learning to predict the changes in scores.

Reference parameters are generated through offline constrained optimization:

$$\theta_i^{\mathrm{ref}} = \arg\min_{\theta\in\Theta_i}\left[\frac{1}{T}\sum_{t=1}^{T}\parallel y_{i,t} - \hat{y}_{i,t}(\theta)\parallel_2^2 + \gamma_c C_i(\theta) + \gamma_e\sum_{i=1}^{N} m_i\,(\theta)\right], \quad (5)$$

Among them, $y_{i,t}$ represents the target joint state, and $\hat{y}_{i,t}(\theta)$ represents the joint state obtained based on the parameters θ. C_i is used to measure compensatory movements, and E_i is used to limit excessive assistance. γ_c and γ_e are the corresponding penalty weights. The optimized parameters are used as reference values during model training. Among them, F represents the target joint state, and $\hat{y}_{i,t}(\theta)$ represents the joint state obtained based on the parameters θ. C_i used to measure compensatory movements, and E_i is used to limit excessive assistance. γ_c and γ_e are the corresponding penalty weights. The optimized parameters are used as reference values during model training.

3.6. Personalized Parameter Decoder

The parameter vector is defined as:

$$\theta_i = [r_i^{sh}, r_i^{el}, g_i, k_i, \phi_i^{on}, \phi_i^{dur}], \quad (6)$$

Among them, r_i^{sh} and r_i^{el} respectively represent the range of motion restrictions for the shoulder and the elbow, g_i represents the auxiliary gain, k_i represents the virtual stiffness, and ϕ_i^{on} and ϕ_i^{dur} respectively represent the auxiliary start time and duration.

The output parameters of the decoder are:

$$\hat{\theta}_i = \theta_{\min} + (\theta_{\max} - \theta_{\min}) \odot \sigma(f_d[z_i \parallel \hat{q}_i \parallel a_i]), \quad (7)$$

Among them, $f_d(\cdot)$ represents a multi-layer perceptron with two hidden layers, and the number of hidden units in each layer is 128 and 64 respectively. $\theta_{\min}$ and $\theta_{\max}$ represent the upper and lower limits of the device parameters, and these ranges will be used to constrain the prediction results in the subsequent steps.

3.7. Biomechanical Safety Projection

The feasible set Θ_i is determined based on the observable movement range of the individual, equipment limitations, as well as the combinations of allowable gains and stiffness. The predicted results are projected as:

$$\theta_i^{*} = \arg\min_{\theta\in\Theta_i}\parallel\theta - \hat{\theta}_i\parallel_2^2, \Theta_i = \{\theta: g_r(\theta, a_i) \leq 0,\ r = 1, \dots, R\}, \quad (8)$$

where $g_r(\cdot)$ represents the r-th constraint condition, and R is the total number of constraint conditions.

After projection, the parameters that satisfy the constraints remain unchanged, while only the parts beyond the range are adjusted. This can prevent situations such as exceeding the joint range of motion, unreasonable combination of gain and stiffness, and deviation of auxiliary movements from the main movement phase.

3.8. Feedback Update and Joint Optimization

The model updates the prediction parameters based on the simulation results or the feedback from the equipment, and rechecks their feasibility after each update:

$$\begin{aligned} \theta_{i,t+1}^{*} &= \Pi_{\Theta_i}\left[\theta_{i,t}^{*} - \eta_t\nabla_\theta\left(E_{i,t} + \beta C_{i,t}\right)\right], \\ \mathcal{L} &= \mathcal{L}_{\mathrm{rep}} + \lambda_\theta \parallel \hat{\theta}_i - \theta_i^{\mathrm{ref}} \parallel_1 + \lambda_y \mathrm{RMSE}(\hat{y}_i, y_i) \\ &\quad + \lambda_s \sum_{r=1}^{R}[g_r(\hat{\theta}_i, a_i)]_+^2 + \lambda_u \mathcal{L}_{\mathrm{unc}}, \end{aligned} \quad (9)$$

Among them, Π_{Θ_i} represents the feasible set projection, $E_{i,t}$ represents the tracking error, $C_{i,t}$ represents the penalty for compensating motion, η_t is the update rate, and β is used to control the degree of compensation. λ_θ , λ_y , λ_s , and λ_u respectively control the weights of parameters, reconstruction, safety, and uncertainty loss, $[v]_+ = \max(0, v)$.

The parameter loss is used to constrain the difference between the predicted result and the offline reference value. The reconstruction loss is used to measure the motion trajectory corresponding to the predicted parameters. The safety loss penalizes the results that violate the constraints before projection. The uncertainty loss is used to estimate the confidence range of the parameter prediction. All weights are determined based on the validation set and remain unchanged in subsequent experiments.

[[TABLE_001]]

4. Experiment and Results

4.1. Experimental Settings

The experiment utilized an open public dataset for stroke rehabilitation exercises, which was collected using 3D depth sensors and IMU sensors, and included data from 128 participants. There were 631 Kinect skeletal files and 128 dual IMU files, covering five types of rehabilitation exercises. The data also included basic participant information, exercise types, main outcome metric scores, and control factor scores. The skeletal data was collected by Microsoft Kinect v2, and the IMU data contained signals from three-axis accelerometers, gyroscopes, and magnetometers. The dataset can be obtained from Mendeley Data and is licensed under the CC BY 4.0 license.

PMAD-Net is mainly used for predicting upper limb assist parameters, so actions such as arm lifting, arm extension with trunk tilt and trunk rotation were selected. These actions can reflect shoulder-elbow coordination and trunk compensation. Sequences with incomplete participant information or where the skeleton and IMU could not be correctly aligned were excluded, but each participant was retained at least one valid sequence. The main outcome metrics and control factor scores were normalized based on the statistics of the training set and used as continuous motion quality targets, without setting binary correctness labels. Reference parameters were generated based on the optimization results of the constraints in Section 3.5 and served as training targets.

The skeleton coordinates were based on the reference point of the base joint of the spine, and were normalized according to shoulder width or upper arm length. Joint positions, velocities, and accelerations were used as node features. The IMU signals were uniformly resampled to 30 Hz and aligned using timestamps and motion energy cross-correlation. All sequences were uniformly processed into 128 frames, and shorter sequences were completed through interpolation or masking. Interpolation was performed for continuous gaps of no more than 5 frames, and the observation mask was retained for gaps of more than 5 frames. The statistical quantities required for normalization were calculated from the training set.

The data is divided by participants. In each round of the experiment, 90 participants are used for training, 19 for validation, and 19 for testing. The experiment is conducted with 10 independent divisions to ensure no overlapping of participants among different data sets. Weighted sampling is only used for the training set, while the validation set and test set maintain the original data distribution.

The experiment was executed in the Python 3.11 and PyTorch 2.3 environment, with GPU acceleration using CUDA 12.1. The hardware contains an Intel Core i7-12700K, 32GB of memory, and an NVIDIA RTX 3060. During model training, the batch size was set to 32, the learning rate started at 1×10^{-3}, the weight decay was 1×10^{-4}, and AdamW was used for optimization. The training was carried out for a maximum of 120 epochs. If the validation set did not improve for 15 consecutive epochs, the training was prematurely terminated. The temperature coefficient and Dropout rate in contrastive learning were set to 0.1 and 0.2.

All four baseline models adopted the same data partitioning and training settings as PMAD-Net, namely DTW-kNN parameter search, BiLSTM regression model, ST-GCN regression model, and time transformer regression model. They used different modeling approaches for predicting device parameters, covering sample retrieval, recurrent networks, graph models, and attention mechanisms.

The model performance is evaluated from three aspects: parameter prediction, motion reconstruction, and safety. The parameter prediction error is measured by the normalized parameter MAE, involving six parameters. The RMSE of motion tracking reflects the error when the model reconstructs the joint angles. The violation rate of constraints is the proportion of parameters in the prediction results that exceed the individual or equipment limits. A lower value indicates better safety.

Each result is calculated based on 10 independent runs and presented as the mean ± standard deviation (n = 10). The results of different methods were subjected to a two-sided paired t-test, and multiple comparisons were corrected using the Holm method. The significance level is set at $\alpha = 0.05$. The error range in the figure corresponds to the 95% confidence interval.

4.2. Performance Comparison

Table 1 compares the performance of PMAD-Net with four baseline methods. The experiment uses 10 independent groups of participants (n = 10), and the results are presented as mean ± standard deviation. A two-sided paired t-test is used between PMAD-Net and the best baseline model, and the Holm method is employed for correction.

Table 1. Performance Comparison on the Subject-Independent Test Sets.

Method	N-MAE ↓	Tracking RMSE (°) ↓	CVR (%) ↓
DTW-kNN	0.132 ± 0.010	8.46 ± 0.58	7.9 ± 1.1
BiLSTM	0.109 ± 0.008	7.31 ± 0.49	6.4 ± 0.9
ST-GCN	0.095 ± 0.007	6.58 ± 0.44	5.1 ± 0.8
Temporal Transformer	0.091 ± 0.006	6.41 ± 0.42	4.8 ± 0.7
PMAD-Net	0.078 ± 0.006	5.73 ± 0.39	2.9 ± 0.5

From Table 1, it can be seen that the results of PMAD-Net are the lowest in all three indicators. Compared with the temporal transformer, its N-MAE decreased by 14.3%, the tracking RMSE decreased by 10.6%, and the CVR decreased by 39.6%. After Holm correction, the three differences still reached statistical significance. The p-values for N-MAE, tracking RMSE, and CVR are 0.004, 0.018, and 0.002, respectively.

The performance of DTW-kNN is relatively poor, which may be related to the lack of a direct correspondence between motion similarity and device requirements. BiLSTM can utilize the temporal sequence of actions, but it does not consider the connection between joints. ST-GCN incorporates information about human joint structure, while Transformer is better at capturing movement changes over a longer time range. However, neither of these two types of methods further considers individual movement capabilities, nor does it set clear safety limits for the prediction results.

The improvement of PMAD-Net in CVR is greater than that of the ordinary prediction error, indicating that feasible set projection has a more obvious effect on reducing unreasonable parameters. The model adjusts results that exceed the joint activity range or do not meet the gain, stiffness, and execution timing restrictions. Therefore, N-MAE reflects the accuracy of parameter prediction, while CVR focuses more on whether the prediction results can meet the limitations in actual use.

4.3. Ablation and Mechanism Verification

Table 2 mainly examines the effects of synchronous processing, IMU branch, personalized settings, and safety mechanisms. Apart from the tested parts, all groups of experiments maintained the same data division, decoder settings, training rounds, and reference parameters.

Table 2. Ablation Results on the Subject-Independent Test Sets.

Variant	N-MAE ↓	Tracking RMSE (°) ↓	CVR (%) ↓
w/o synchronization	0.088 ± 0.007	6.26 ± 0.45	3.7 ± 0.6
w/o IMU branch	0.084 ± 0.007	6.03 ± 0.43	3.4 ± 0.6
w/o personalization	0.091 ± 0.008	6.41 ± 0.47	4.0 ± 0.7
w/o safety projection	0.081 ± 0.006	5.79 ± 0.41	8.6 ± 1.2
Full PMAD-Net	0.078 ± 0.006	5.73 ± 0.39	2.9 ± 0.5

The influence of individual information on the model is the most obvious. After removing limb scale, observable joint range of motion, bilateral differences, and compensatory movements, the performance of the model in parameter prediction and motion tracking has significantly changed. The corresponding p-values for N-MAE and tracking RMSE are 0.003 and 0.009 respectively. This indicates that it is difficult to distinguish the inherent movement limitations of an individual from the accidental fluctuations in a single action based on the changes in the action itself.

Removing synchronization affects the model's performance. When there is a time deviation between the bone data and the IMU data, all three indicators decrease. This shows that the two types of signals need to maintain a good time correspondence. In contrast, the IMU branch brings relatively smaller changes when the data is complete. After the coordinates of the shoulder or elbow become missing, the impact of removing the IMU branch becomes significantly greater. This indicates that when visual information is insufficient, IMU data can supplement some of the motion information.

The safety mechanism has a small impact on the prediction error, but is more significant in reducing unreasonable parameters. After removing this mechanism, N-MAE increased from 0.078 to 0.081, tracking RMSE increased from 5.73 to 5.79, and CVR increased from 2.9% to 8.6% ($p < 0.001$). This indicates that the safety mechanism is mainly used to control whether the prediction results meet the constraints.

4.4. Convergence Analysis

As shown in Figure 2, the validation loss of PMAD-Net reached its minimum value at the 73 ± 8th training cycle, which was 0.184 ± 0.011. The temporal transformer reached its minimum point around the 100 ± 9th cycle, with a loss of 0.217 ± 0.014. There was a significant difference in the minimum validation loss between the two models ($p = 0.006$).

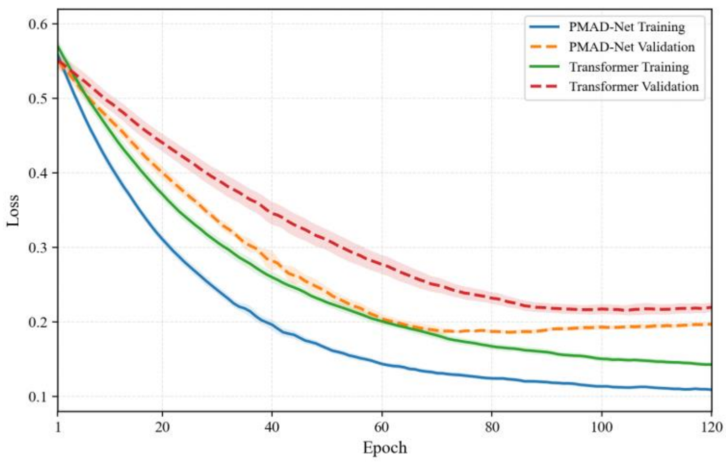

Figure 2. Training and Validation Losses of PMAD-Net and the Temporal Transformer over Ten Runs; Shaded Areas Represent 95% Confidence Intervals.

PMAD-Net showed a rapid decrease in the early training stage, with the most significant changes occurring in the first 30 cycles. There were some fluctuations from the 35th to the 45th cycle, which were related to the gradual increase in the safety loss weight. After that, the curve gradually stabilized and reached its minimum point around the 73rd cycle. After the 90th cycle, the training loss continued to decrease, but the validation loss began to rise, which also indicates that stopping training early is reasonable.

The temporal transformer did not incorporate a planned safety regularization, so the training curve was smoother but with a slower convergence speed. After the 60th cycle, the confidence interval of PMAD-Net also gradually narrowed. At this point, the personalized information and constraint learning became stable, and the model's result differences for different participants decreased.

4.5. Stability, Generalization, and Trustworthiness

Figure 3 further examined the performance of the model under data missing, time displacement, and task changes. Without adding perturbation, the N-MAE of PMAD-Net was 0.078 ± 0.006. After increasing the joint Dropout ratio to 10%, 20%, and 30%, the N-MAE was 0.086 ± 0.007, 0.098 ± 0.009, and 0.116 ± 0.011, respectively. The loss of 20% of IMU data packets increased the N-MAE to 0.089 ± 0.008, and when the time offset was 4 frames, it reached 0.101 ± 0.009.

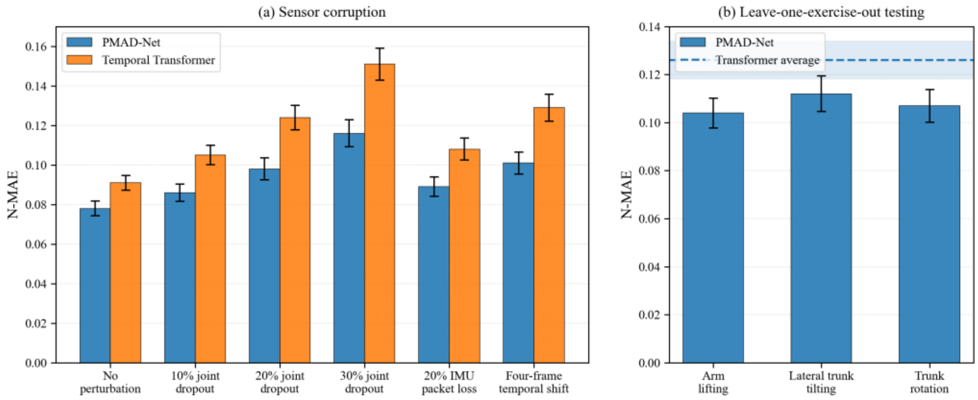

Figure 3. Robustness and Cross-Exercise Generalization of PMAD-Net: (A) N-MAE under Joint Dropout, IMU Packet Loss, and Temporal Misalignment; (B) Leave-One-Exercise-Out N-MAE for Arm Lifting, Lateral Trunk Tilting, and Trunk Rotation. Error Bars and Shaded Regions Represent 95% Confidence Intervals over Ten Runs.

The corresponding Transformer results were 0.091±0.006, 0.105±0.008, 0.124±0.010, 0.151±0.013, 0.108±0.009, and 0.129±0.011 respectively. In comparison, PMAD-Net performed better in three scenarios of 20% joint loss ($p=0.009$), 30% joint loss ($p=0.003$), and time alignment error ($p=0.007$). The model will distinguish between the true zero coordinates and the missing data, and adjust its influence based on the completeness of the data of each modality.

However, even with 30% of the joint data missing, the N-MAE still increased by 48.7%. This indicates that when there is insufficient effective observation for the shoulder and elbow joints, the model performance will significantly decline. Severe data loss remains a limitation for the model.

The N-MAE obtained from the Leave-One-Out test were 0.104 ± 0.010 (for arm lifting), 0.112 ± 0.012 (for trunk tilt), and 0.107 ± 0.011 (for trunk rotation). Among them, trunk tilt was the most difficult to predict, possibly because the trunk compensation in the movement was more obvious. The average error was still lower than that of the Transformer (0.126 ± 0.013).

The analysis results show that the model's judgment of different parameters will refer to different movement changes. The movements of the shoulders and elbows are mainly related to the range of motion, while angular velocity is more used to determine the auxiliary gain and start time. The movements of the trunk are related to stiffness and compensation. When there are many missing data or abnormal duration of the action, the uncertainty of the model will increase. The model uses anonymized skeletons and IMU data, and does not require the collection of facial images. The final results still need to be reviewed by professionals before being used.

5. Conclusion

This study investigated whether various motion capture data could be used to generate rehabilitation equipment parameter recommendations that are more suitable for individual conditions and meet safety constraints. The experimental results were largely consistent with the research goals proposed in the previous studies.

Firstly, after synchronizing and scaling the body size of the IMU data with the skeletal data, the model's predictions became more stable. Without the synchronization processing, the N-MAE increased from 0.078 ± 0.006 to 0.088 ± 0.007. Tests with time misalignment and sensor data missing also indicated that marking the missing data and adjusting the influence of different information based on the completeness of the data could alleviate the impact of incomplete data. Secondly, after incorporating individual movement information, the model's prediction of equipment parameters became more accurate. The model would combine information such as joint range of motion, limb length, left-right difference, and compensatory movement. The N-MAE of PMAD-Net was 0.078 ± 0.006, and the motion tracking RMSE was 5.73 ± 0.39, which were 14.3% and 10.6% lower than those of the Temporal Transformer, respectively. Removing individual information led to the most significant decline in model performance, indicating that the movement characteristics of different users do have a significant impact on parameter prediction. Finally, the projection of the biomechanical feasible set reduced the violation rate to 2.9 ± 0.5%, while removing this step increased it to 8.6 ± 1.2%. This shows that merely looking at the prediction error is not sufficient to determine whether the parameters are suitable for practical use.

However, the applicability of the model is still limited by certain conditions. When a large amount of joint data is missing, the model's performance will significantly decline. After missing 30% of the joints, the N-MAE increased by 48.7%. The results of the leave-one-out test were also lower than those of the conventional independent subject test, with relatively weaker performance in trunk tilt. Additionally, this study only used an open dataset, and the reference parameters were obtained from offline constraint optimization, not through repeated adjustments by clinical personnel on actual equipment. Compared to simple regression models, the current graph structure and multimodal design require more computational resources.

Future work can verify these parameter recommendations on real adjustable orthoses or rehabilitation robots, and include parameters obtained from actual clinical adjustments as references. At the same time, tests should be conducted in different institutions, with different sensor equipment, and for different patient groups. Electromyography, force, and fatigue information can also be used as subsequent supplements. The model itself can be further simplified to reduce the computational load during actual use. When data is severely missing or the uncertainty given by the model is too high, clear judgment criteria need to be established to decide when to abandon or re-evaluate the prediction results.

References

1. U. Lugrís, M. Pérez-Soto, F. Michaud, and J. Cuadrado, "Human motion capture, reconstruction, and musculoskeletal analysis in real time," *Multibody System Dynamics*, vol. 60, no. 1, pp. 3–25, 2024.
2. R. I. Hamilton, Z. Glavcheva-Laleva, M. I. H. Milon, Y. Anil, J. Williams, P. Bishop, and C. Holt, "Comparison of computational pose estimation models for joint angles with 3D motion capture," *Journal of Bodywork and Movement Therapies*, vol. 40, pp. 315–319, 2024.
3. J. Li, J. Xue, R. Cao, X. Du, S. Mo, K. Ran, and Z. Zhang, "Finerehab: A multi-modality and multi-task dataset for rehabilitation analysis," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2024, pp. 3184–3193.
4. Z. Niu, K. Lu, J. Xue, X. Qin, J. Wang, and L. Shao, "From methods to applications: A review of deep 3D human motion capture," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 11, pp. 11340–11359, 2024.
5. M. Devanne, O. R. Neris, M. Lempereur, and A. Thepaut, "A medical low-back pain physical rehabilitation database for human body movement analysis," in *2024 Int. Joint Conf. Neural Netw. (IJCNN)*, 2024, pp. 1–8.
6. I. Kourbane, P. Papadakis, and M. Andries, "SSL-Rehab: Assessment of physical rehabilitation exercises through self-supervised learning of 3D skeleton representations," *Comput. Vis. Image Underst.*, vol. 251, p. 104275, 2025.

7. N. Bhagat, P. Raghavan, and V. Kapila, "Role of joint interactions in upper limb joint movements: a disability simulation study using wearable inertial sensors for 3D motion capture," *J. NeuroEng. Rehabil.*, vol. 21, no. 1, p. 197, 2024.
8. O. Yuhai, A. Choi, Y. Cho, H. Kim, and J. H. Mun, "Deep-learning-based recovery of missing optical marker trajectories in 3D motion capture systems," *Bioengineering*, vol. 11, no. 6, p. 560, 2024.
9. S. Scataglini, E. Abts, C. Van Bocxlaer, M. Van den Bussche, S. Meletani, and S. Truijen, "Accuracy, validity, and reliability of markerless camera-based 3D motion capture systems versus marker-based 3D motion capture systems in gait analysis: a systematic review and meta-analysis," *Sensors*, vol. 24, no. 11, p. 3686, 2024.
10. W. Jia, H. Wang, Q. Chen, T. Bao, and Y. Sun, "Analysis of Kinect-Based Human Motion Capture Accuracy Using Skeletal Cosine Similarity Metrics," *Sensors*, vol. 25, no. 4, p. 1047, 2025.
11. I. Kourbane, P. Papadakis, and M. Andries, "Optimized assessment of physical rehabilitation exercises using spatiotemporal, sequential graph-convolutional networks," *Comput. Biol. Med.*, vol. 186, p. 109578, 2025.
12. X. B. Bruce, Y. Liu, K. C. Chan, and C. W. Chen, "Egcn++: A new fusion strategy for ensemble learning in skeleton-based rehabilitation exercise assessment," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 9, pp. 6471–6485, 2024.
13. R. Hu, Y. Diao, Y. Wang, G. Li, R. He, Y. Ning, and G. Zhao, "Effective evaluation of HGcnMLP method for markerless 3D pose estimation of musculoskeletal diseases patients based on smartphone monocular video," *Front. Bioeng. Biotechnol.*, vol. 11, p. 1335251, 2024.
14. L. A. Espitia-Mora, M. A. Vélez-Guerrero, and M. Callejas-Cuervo, "Development of a low-cost markerless optical motion capture system for gait analysis and anthropometric parameter quantification," *Sensors*, vol. 24, no. 11, p. 3371, 2024.
15. S. Miao, Z. Liu, D. Wang, X. Shen, and N. Shen, "Applying hybrid deep learning models to assess upper limb rehabilitation," *IEEE Access*, vol. 12, pp. 154337–154348, 2024.