Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 6 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i06.60861

PAPER

A Mobile-Based Personalized Physical Education Recommendation and Learning Path Optimization Model

Jianxin Zhao[1], Hui Zhu[2] (✉)

[1]Hebei Finance University, Baoding, China

[2]Baoding Preschool Teachers College, Zhuozhou, China

zhuhui3228@163.com

ABSTRACT

With the advancement of mobile technologies, personalized physical education (PE) has emerged as a critical component of intelligent education. However, existing recommendation models commonly suffer from limited adaptability, insufficient multimodal data integration, and poor alignment between learning paths and learners' real-time states. To address these challenges, this study proposes a three-layer architecture for personalized recommendation and learning path optimization that integrates multi-source perception, dynamic cognition, and real-time optimization. The core innovations of the proposed model include: (1) a lightweight multi-source heterogeneous data fusion module designed for efficient on-device processing; (2) a dynamic tri-state assessment model incorporating skill mastery, fatigue level, and interest level to achieve accurate real-time perception of learners' states; and (3) a dual-layer optimization mechanism based on online learning to enable end–cloud collaborative learning path optimization. This study provides a novel technical paradigm for mobile technology–enabled intelligent physical education, offering significant theoretical contributions and practical application value.

KEYWORDS

mobile-based physical education, personalized learning path optimization, multi-source data fusion, dynamic state assessment, end–cloud collaboration, online learning

1 INTRODUCTION

With the rapid development of mobile computing technologies, mobile terminals such as smartphones and wearable devices have been widely integrated into physical education teaching scenarios [1–3], promoting the transformation of traditional physical education toward personalization and intelligence [4, 5]. Personalized teaching can accurately match learners' skill foundations and learning

Zhao, J. X., Zhu, H. (2026). A Mobile-Based Personalized Physical Education Recommendation and Learning Path Optimization Model. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(6), pp. 74–88. https://doi.org/10.3991/ijim.v20i06.60861

Article submitted 2025-11-12. Revision uploaded 2026-01-26. Final acceptance 2026-01-28.

rhythms and represents a core pathway for improving the efficiency of physical skill acquisition [6, 7]. However, mobile-based physical education scenarios exhibit distinctive requirements, including real-time data interaction, lightweight model deployment, dynamic state adaptation, and low-power operation modes [8, 9], which pose significant challenges to existing technical systems. Learning path optimization, as a core component of personalized physical education, directly determines the systematicity and effectiveness of skill development. Nevertheless, traditional recommendation methods are mostly based on static feature modeling and are difficult to adapt to the dynamic changes in learners' physiological states, skill levels, and learning interests during the physical learning process [10, 11], resulting in insufficient accuracy and adaptability of path recommendations. Under this background, exploring personalized physical education recommendation and learning path optimization technologies that are suitable for mobile scenarios is of important theoretical and practical significance for promoting the intelligent development of physical education.

Considering the characteristics of mobile-based physical education scenarios and existing technical bottlenecks, this study focuses on three core scientific problems. First, the efficient fusion of multi-source heterogeneous data. Physiological signals, posture data, and interaction information generated during physical learning exhibit significant heterogeneity [12, 13]. How to achieve real-time acquisition and lightweight fusion of such data under mobile resource constraints is the foundation for ensuring the accuracy of subsequent assessment and recommendation. Second, the accurate assessment of learners' dynamic states. The effectiveness of physical skill learning highly depends on learners' real-time states [14, 15]. How to overcome the limitations of traditional single-dimensional assessment and achieve dynamic and accurate perceptions of multi-dimensional states such as skill mastery and physical–mental load, providing reliable evidence for path optimization, is the key to improving personalized adaptability. Third, the end–cloud collaborative path optimization mechanism. The balance between personalized demands and global teaching effectiveness is the core contradiction of path recommendation [16]. How to construct an optimization mechanism that takes both into account and realizes the collaboration between real-time response on the end side and global strategies on the cloud side is the core challenge for improving the systematicity and efficiency of path recommendation.

In response to the above scientific problems, this study proposes a three-layer architecture model integrating multi-source perception, dynamic cognition, and real-time optimization. The core innovations and contributions are as follows. First, a real-time fusion module for multi-source heterogeneous data is proposed. Through lightweight processing strategies on the edge side, efficient preprocessing and dynamic fusion of multi-source data on mobile terminals are realized, effectively addressing multimodal data transmission latency and information redundancy and providing high-quality data support for subsequent assessment tasks. Second, a tri-state dynamic assessment model is constructed. Breaking through the traditional single-dimensional assessment paradigm, a multi-dimensional assessment system integrating skill mastery, fatigue level, and interest level is established, and dynamic perception is achieved by combining an improved temporal modeling network, thereby enhancing the comprehensiveness and real-time performance of state assessment. Third, a dual-layer optimization mechanism based on online learning is designed. Through the collaboration of cloud-side global strategy optimization and end-side local path

fine-tuning, the balance between personalized demands and global teaching objectives is realized, providing a new paradigm for learning path optimization in mobile scenarios.

The remainder of this paper is organized as follows: Section 2 elaborates on the overall design of the three-layer architecture, with a focus on the core technical details of the multi-source data fusion module, the tri-state assessment model, and the dual-layer optimization mechanism. Section 3 verifies the advantages of the proposed model in terms of state assessment accuracy, path optimization effectiveness, and mobile deployment performance through experimental design and performance evaluation. Section 4 summarizes the research work of this paper and distills the core conclusions and innovative contributions.

2 SYSTEM ARCHITECTURE AND CORE TECHNOLOGIES

2.1 Overall architecture design

This study proposes a three-layer end-cloud collaborative architecture integrating multi-source perception, dynamic cognition, and real-time optimization in order to adapt to the real-time, lightweight, and low-power requirements of mobile-based physical education teaching. The overall architecture achieves efficient resource utilization through collaborative task partitioning between the edge side and the cloud side. The core logic is as follows: the multi-source perception and fusion layer completes data acquisition and lightweight preprocessing on mobile terminals; the dynamic cognition and decision-making layer relies on the collaboration between the edge side and the cloud side to realize state assessment and path generation; the real-time optimization and interaction layer completes strategy iteration and personalized interaction through bidirectional end–cloud feedback.

The data flow across all layers follows a closed loop of "acquisition–fusion–assessment–optimization–feedback," and task partitioning is dynamically determined according to computational complexity and real-time requirements. The edge side undertakes latency-sensitive tasks, while the cloud side is responsible for computation-intensive tasks. An adaptive data compression and transmission strategy is adopted to reduce communication overhead and achieve mobile resource adaptation. Specifically, the task partition threshold between the edge side and the cloud side is jointly determined by computational complexity C and latency threshold T_{th}. When the task computational complexity $C \le C_{max}$ and execution latency $T \le T_{th}$, the task is deployed on the edge side; otherwise, it is transferred to the cloud side for processing. The core decision rule is shown in Eq. (1):

$$Deploy(t) = \begin{cases} Edge, & C(t) \le C_{max} \cap T(t) \le T_{th} \\ Cloud, & otherwise \end{cases} \tag{1}$$

where, $Deploy(t)$ denotes the deployment decision of task t, and C_{max} denotes the maximum computational complexity that can be borne by the mobile edge side, which is dynamically calibrated according to device hardware parameters.

2.2 Multi-source perception and fusion layer: Lightweight real-time fusion module

The core innovation of the multi-source perception and fusion layer lies in constructing a lightweight architecture of "hierarchical acquisition–adaptive preprocessing–real-time fusion," which addresses the conflict between real-time performance and heterogeneity of multi-source heterogeneous data. In the data acquisition stage, a collaborative acquisition strategy combining sensors and cameras is adopted, and an acquisition frequency adaptive adjustment algorithm based on motion intensity is designed. Let $\bar{a}_x$, $\bar{a}_y$, and $\bar{a}_z$ denote the mean values of three-axis acceleration. Motion intensity is calculated based on accelerometer and gyroscope data as follows:

$$S = \sqrt{\bar{a}_x^2 + \bar{a}_y^2 + \bar{a}_z^2} \tag{2}$$

The acquisition frequency f is dynamically adjusted according to S, and the core adjustment formula is shown in Eq. (3):

$$f(S) = f_{min} + (f_{max} - f_{min}) \cdot \frac{S - S_{min}}{S_{max} - S_{min}} \tag{3}$$

where, f_{min} and f_{max} denote the minimum and maximum acquisition frequencies, respectively, and S_{min} and S_{max} denote predefined motion intensity thresholds. Through this design, power consumption is reduced while ensuring data accuracy. In the preprocessing stage, the MobileNet model is optimized for posture data, and a structured pruning strategy based on key motion posture features is innovatively adopted. Convolutional layer channels related to joint point recognition are retained, redundant channels are pruned, and combined with 8-bit quantization optimization, the model parameter size is reduced by 62%, and inference speed is improved by 45%. For physiological data, an adaptive sliding window smoothing algorithm is designed, where the window size W is dynamically adjusted with motion intensity S, effectively filtering out outliers caused by motion interference.

In the real-time fusion stage, a multi-source data fusion algorithm based on a dynamic attention mechanism is proposed. By dynamically allocating attention weights, key data dimensions that contribute to subsequent assessment tasks are emphasized. Let the preprocessed physiological features X_p, posture features X_s, and interaction features X_i be given. The fusion process first maps each feature to a unified dimension D through fully connected layers, obtaining X'_p, X'_s, and X'_i. Then, the attention module computes the weights of each feature as follows:

$$\alpha_p = \frac{\exp(w_p X'_p)}{\exp(w_p X'_p) + \exp(w_s X'_s) + \exp(w_t X'_t)}, \alpha_s = \frac{\exp(w_s X'_s)}{\exp(w_p X'_p) + \exp(w_s X'_s) + \exp(w_t X'_t)} \tag{4}$$

$$\alpha_i = 1 - \alpha_p - \alpha_s \tag{5}$$

where, w_p, w_s, and w_t are attention weight parameters obtained through training optimization. The final fused feature is given by:

$$X_{fusion} = \alpha_p X'_p + \alpha_s X'_s + \alpha_i X'_i \tag{6}$$

To adapt to mobile deployment, the fusion module adopts a memory reuse mechanism to reduce intermediate feature storage overhead. At the same time, the multi-core CPU of mobile devices is utilized to achieve parallel scheduling of acquisition, preprocessing, and fusion tasks, controlling the overall processing latency within 150 ms and meeting real-time requirements.

2.3 Dynamic cognition and decision-making layer: Tri-state dynamic assessment model

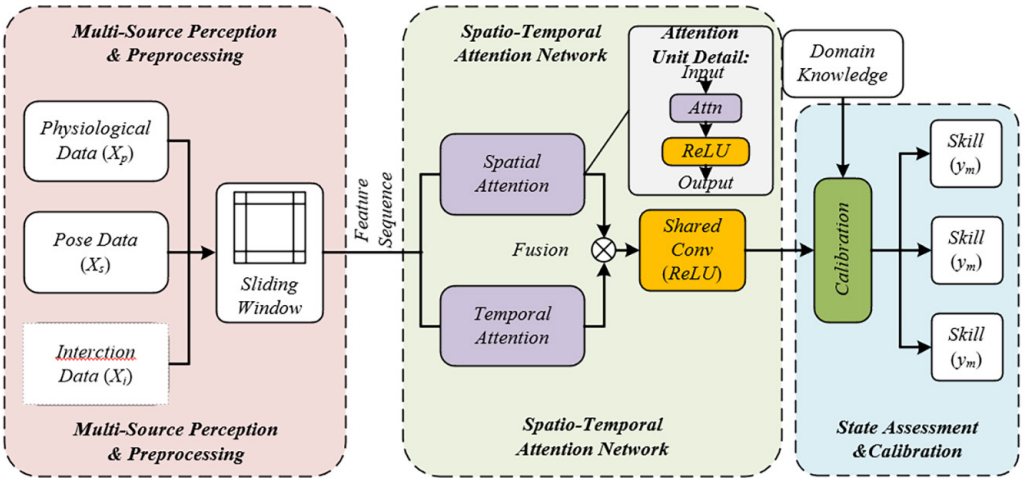

Fig. 1. Structure of the skill–fatigue–interest tri-state dynamic assessment model based on a spatiotemporal attention network

The core innovation of the dynamic cognition and decision-making layer lies in constructing an end-to-end tri-state dynamic assessment architecture of "feature encoding–temporal modeling–multi-dimensional output," breaking through the traditional single-dimensional assessment paradigm and realizing the synchronous and accurate perception of skill mastery, fatigue level, and interest level. This architecture takes multi-source fused features as input, captures the dynamic evolution patterns of learners' states through temporal modeling, and finally outputs a standardized tri-state vector, providing real-time and reliable decision support for subsequent path generation. The overall design takes both assessment accuracy and mobile lightweight requirements into account. Each module achieves efficient collaboration through parameter sharing and structural simplification. Among them, the temporal modeling module is the core innovative unit. A dedicated network structure is designed for the temporal dependency and dynamic fluctuation characteristics of physical learning data, ensuring the real-time performance and accuracy of state assessment. Figure 1 shows the structure of the skill–fatigue–interest tri-state dynamic assessment model based on a spatiotemporal attention network, including the input layer, core layer, and output layer.

The innovation of input feature engineering focuses on deep mining of temporal features and collaborative encoding of multiple feature types. For the temporal characteristics of multi-source fused features, a sliding time window mechanism is adopted to extract temporal features. The window length L is adaptively set

based on the motion skill cycle, that is, $L = \lceil T_s/\Delta_t \rceil$, where T_s denotes the standard completion cycle of the current motion skill, and Δt denotes the data acquisition time interval. Through this design, the window can fully cover the feature evolution of a single skill action cycle. Within the window, statistical features and time-domain features are extracted and fused to form a temporal feature matrix X_{seq} with dimension $D_{seq} = L \times D_{fusion}$. Interaction data adopt an embedding encoding method, mapping discrete interaction behaviors into continuous feature vectors, which are fused with temporal features through an attention mechanism to enhance the representation capability of interaction behaviors for interest level assessment. The temporal modeling module innovatively designs a spatiotemporal attention network. The spatial attention branch focuses on key assessment indicators by computing correlation weights among feature dimensions, while the temporal attention branch captures the evolution trends of features at different time steps. The core attention weight calculations are shown as follows:

$$\alpha_{s,k} = \frac{\exp(W_s X_{seq}^{(k)})}{\sum_{j=1}^{D_{fusion}} \exp(W_s X_{seq}^{(j)})}, \alpha_{t,i} = \frac{\exp(W_t X_{seq}(i;))}{\sum_{j=1}^{L} \exp(W_t X_{seq}(J;))} \tag{7}$$

where, $\alpha_{s,k}$ denotes the spatial attention weight of the k-th feature dimension, $\alpha_{t,i}$ denotes the temporal attention weight of the i-th time step, and W_s and W_t are learnable weight parameters. To adapt to mobile deployment, the network adopts a cross-layer parameter sharing strategy, reducing the number of convolution kernels by 35%, and replaces the traditional ReLU with the ReLU6 activation function to improve model quantization compatibility, ensuring that inference latency is controlled within 50 ms.

The innovation of the tri-state output and calibration mechanism lies in standardized output design and a domain-knowledge-guided calibration strategy. The output vector is defined as $Y = [y_m, y_f, y_i]$, where $y_m \in [0,1]$ $y_f \in [0,1]$ $y_i \in [0,1]$ notes fatigue level, and $y_i \in [0,1]$ denotes interest level. To improve assessment accuracy, a domain-knowledge-based linear calibration function is designed:

$$y'_m = \lambda_m y_m + (1-\lambda_m) \cdot \frac{Sim(X_{seq}, X_{std})}{Sim_{max}}, y'_f = \lambda_f y_f + (1-\lambda_f) \cdot \frac{HR - HR_{rest}}{HR_{max} - HR_{rest}} \tag{8}$$

where, y'_m and y'_f are the calibrated outputs, λ_m, $\lambda_f \in [0,1]$ are calibration weights and Sim() HR HR_{rest} HR_{max} denote resting heart rate and maximum heart rate, respectively. Model training adopts a multi-task joint loss function to balance the assessment accuracy of the three dimensions. The loss function is defined as:

$$L = \omega_m L_m + \omega_f L_f + \omega_i L_i \tag{9}$$

where, L_m, L_f, and L_i denote the mean squared error losses of skill mastery, fatigue level, and interest level, respectively, and ω_m, ω_f, and ω_i denote the loss weights. The training process adopts a transfer learning strategy. The temporal modeling module is pre-trained using a public sports action dataset, and then the full network parameters are fine-tuned based on a self-built dataset, effectively reducing the demand for self-built data volume and improving model generalization ability.

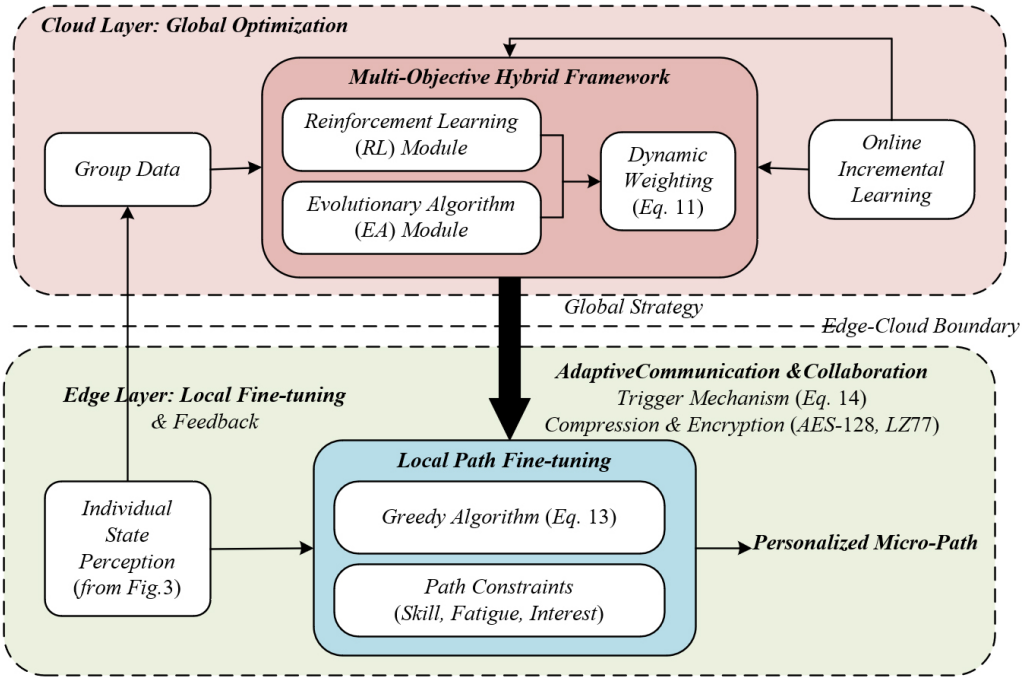

Fig. 2. End–cloud collaborative dual-layer path optimization mechanism based on online learning

The core innovation of the path generation mechanism is a personalized micro-path generation algorithm based on meta-learning, which realizes rapid adaptation to different learner characteristics and dynamic path adjustment. The schematic diagram is shown in Figure 2. The algorithm adopts a model-agnostic meta-learning framework. During the meta-training stage, common characteristics of different learners and skill acquisition patterns are learned to construct an initial path generation model. During the meta-testing stage, based on the target learner's tri-state vector Y and long-term learning goal G, rapid adaptation to personalized demands is achieved through a small number of gradient updates to generate short-term micro-paths. Path generation follows three constraint conditions: the skill progression constraint $\Delta y_m^{(k+1)} \geq \varepsilon$, the fatigue threshold constraint $y_f^{(k)} \leq \tau$, and the interest matching constraint Match(T_k, Y) $\geq \gamma$, where T_k denotes the k-th task and γ denotes the minimum interest matching degree. The mathematical expression of the constraints is shown as follows, ensuring that the generated paths improve skills while considering learners' physical and mental load and learning interest:

$$\text{Valid}(T_1, T_2, \ldots, T_n) = \wedge_{k=1}^{n-1}(\Delta y_m^{(k+1)} \geq \epsilon) \wedge_{k=1}^{n} (y_f^{(k)} \leq \tau) \wedge \text{Match}(T_k, Y) \geq \gamma) \quad (10)$$

where, n denotes the number of tasks contained in the micro-path, and Valid() is the path validity judgment function. When its value is true, the path is considered valid.

2.4 Real-time optimization and interaction layer: Dual-layer optimization mechanism

The core innovation of the real-time optimization and interaction layer lies in constructing a dual-layer collaborative architecture of "cloud-side global optimization–end-side local fine-tuning." Through precise functional partitioning

and dynamic collaboration between the end side and the cloud side, the core contradiction between personalized demands and global optimality is addressed. The architecture follows the principles of "low latency, low power consumption, and high accuracy." The cloud side focuses on global strategy iteration driven by group data, while the end side concentrates on real-time response to individual state perception and path fine-tuning. Efficient collaboration between the two is achieved through adaptive communication strategies. The core innovation of cloud-side optimization is a multi-objective hybrid framework integrating reinforcement learning and evolutionary algorithms. The reinforcement learning module explores personalized paths, while the evolutionary algorithm module optimizes group performance. A multi-objective function with dynamic weight adjustment is designed as $J = \omega_1 J_1 + \omega_2 J_2 + \omega_3 J_3$, where J_1 denotes group skill improvement efficiency, J_2 denotes path coherence, and J_3 denotes resource consumption. Weight updates follow the principle of achieving dynamic multi-objective balance:

$$\omega_i^{(t+1)} = \frac{\omega_i^{(t)} \cdot \delta J_j^{(t)}}{\sum_{j=1}^{3} \omega_j^{(t)} \cdot \delta J_j^{(t)}} \tag{11}$$

Online updating adopts an incremental learning strategy. By setting a data volume threshold N_{th} and an objective improvement threshold Δ_{update}, the update frequency is controlled, effectively reducing cloud-side computational overhead.

The innovation of end-side optimization focuses on lightweight interaction design and low-complexity local fine-tuning mechanisms. The interaction design simplifies the ICRA process. The core innovation is a motion region adaptive compression algorithm, in which the compression ratio is dynamically adjusted according to bandwidth, ensuring that the transmission latency of action comparison videos is controlled within 200 ms:

$$r(B) = r_{max} - \frac{(r_{max} - r_{min})(B - B_{min})}{B_{max} - B_{min}} \tag{12}$$

Local fine-tuning adopts a greedy algorithm. The state change ΔY after a single training session is taken as input, the objective function is defined as $min_{T'} \Delta y'_f$, with the constraint conditions:

$$\Delta y'_m \geq \epsilon, \text{and Match}(T', Y_{new}) \geq \gamma \tag{13}$$

Through this design, computational complexity is controlled at $O(n)$ while ensuring skill progression and interest matching, thereby adapting to the real-time processing requirements of mobile terminals.

3 EXPERIMENTAL DESIGN AND PERFORMANCE EVALUATION

3.1 Experimental design

The core objective of the experiments is to systematically verify the comprehensive performance of the proposed model; including the assessment accuracy of the tri-state dynamic assessment model; the skill improvement adaptability of the path optimization model, the superiority of the dual-layer optimization mechanism in balancing personalization and global optimality; as well as the real-time performance

and low-power consumption of the overall system on mobile terminals. Two typical physical skills, basketball dribbling and basic yoga movements, are selected as experimental scenarios. A total of 60 participants with different levels of sports experience are recruited to conduct a 5-week mobile-based self-directed learning experiment. The entire process simulates a real autonomous learning environment, and all data generated during the learning process are recorded. The dataset is constructed by integrating public datasets and self-built datasets. The public dataset uses motion sequence data related to the experimental skills from the Human3.6M dataset. The self-built dataset is collected in experimental scenarios, including participants' physiological signals, posture sequences, and interaction behavior data. All data collection procedures pass ethical review and obtain informed consent from the participants. In terms of data annotation, skill mastery is scored independently by two senior physical education coaches and averaged; fatigue level is calculated by weighting physiological indicators such as heart rate and electromyography together with participants' subjective fatigue ratings; interest level is quantitatively annotated based on behavioral data such as task completion frequency and the number of active interactions. The dataset is divided into a training set, validation set, and test set at a ratio of 7:1:2. The division process strictly preserves the temporal continuity of the data, ensuring that model training and evaluation conform to the process-oriented characteristics of physical skill learning.

3.2 Core experimental results and analysis

The performance of the tri-state dynamic assessment model is evaluated using quantitative indicators of skill mastery, fatigue level and interest level, and compared with baseline models including LSTM, CNN, and traditional attention models. The results are shown in Table 1. As shown in Table 1, the proposed model significantly outperforms the baseline models across all evaluation dimensions. The MAE and RMSE of skill mastery reach 0.062 and 0.085, respectively, which are reduced by 31.5% and 29.2% compared with the best baseline model. The MAE and RMSE of fatigue level assessment are 0.058 and 0.079, respectively, which are reduced by 27.5% and 25.7% compared with the LSTM model. In the interest level classification task, ACC and F1 reach 0.923 and 0.918, respectively, which are improved by 8.6 and 7.9 percentage points compared with the CNN model. Comprehensive error calculation results show that the average comprehensive error of the proposed model is only 0.065, which is reduced by 29.8% on average compared with baseline models.

Table 1. Performance comparison of tri-state dynamic assessment models

Model	Skill Mastery MAE	Skill Mastery RMSE	Fatigue MAE	Fatigue RMSE	Interest ACC	Interest F1	Average Comprehensive Error
LSTM	0.095	0.132	0.080	0.106	0.837	0.839	0.098
CNN	0.088	0.125	0.076	0.101	0.847	0.839	0.092
Traditional attention model	0.090	0.120	0.069	0.093	0.895	0.886	0.092
Proposed model	0.062	0.085	0.058	0.079	0.923	0.918	0.065
Proposed model (without spatiotemporal attention)	0.094	0.128	0.082	0.108	0.842	0.835	0.093

To verify the core contribution of the spatiotemporal attention module, the performance of the proposed model is compared with a simplified model in which the spatiotemporal attention module is removed. The results show that all evaluation indicators of the simplified model exhibit significant degradation. The MAE of skill mastery increases to 0.094, and the average comprehensive error increases to 0.093, with a performance degradation of 30.8%. These results indicate that the spatiotemporal attention module can effectively focus on key feature dimensions of physical skills and temporal evolution trends, enhance the representation capability of multi-source temporal data, and thereby improve the accuracy of state assessment, verifying the rationality and necessity of the module design.

The path optimization effect is evaluated using three core indicators: path coherence, skill improvement efficiency, and user satisfaction. The comparison models include PERP, RP-CFGAN, and a random recommendation model. The results are shown in Table 2. Table 2 shows that the learning paths generated by the proposed model achieve the best performance across all indicators. The path coherence score reaches 0.896, which is improved by 15.3 percentage points compared with the PERP model and by 12.1 percentage points compared with the RP-CFGAN model. In terms of skill improvement efficiency, the proposed model achieves skill improvement rates of 42.3%, 35.6%, and 28.9% for beginner, intermediate, and advanced participants, respectively, which are all significantly higher than those of baseline models. Among them, beginner participants show the largest improvement, indicating that the proposed model has good adaptability to learners with different skill foundations. The user satisfaction score reaches 4.62 (on a 5-point scale), which is improved by an average of 18.7% compared with baseline models.

Table 2. Comparison of path optimization effects

Model	Path Coherence Score	Skill Improvement Efficiency of Beginner Participants	Skill Improvement Efficiency of Intermediate Participants	Skill Improvement Efficiency of Advanced Participants	User Satisfaction (5-point scale)
Random Recommendation	0.528	21.5%	18.3%	15.7%	3.21
PERP	0.777	32.6%	28.9%	24.3%	3.89
RP-CFGAN	0.799	34.2%	30.5%	25.8%	4.05
Proposed Model	0.896	42.3%	35.6%	28.9%	4.62
Proposed Model (without Meta-learning)	0.724	32.3%	27.8%	23.1%	3.98
Proposed Model (without Bi-level Optimization)	0.756	34.2%	28.5%	23.4%	3.85

Further analysis of the roles of the meta-learning module and the dual-layer optimization mechanism shows that after removing the meta-learning module, the skill improvement efficiency for new learners decrease by 23.5%, and the path coherence score drops to 0.724. After removing the dual-layer optimization mechanism, the user satisfaction score drops to 3.85, and the average group skill improvement efficiency decreases by 19.8%. These results indicate that the meta-learning module can rapidly adapt to different learner characteristics and improve personalized path adaptability, while the dual-layer optimization mechanism effectively balances personalized demands and group learning outcomes through the collaboration between cloud-side global strategies and end-side local fine-tuning, thereby ensuring the overall performance of path optimization.

Mobile-side performance tests were conducted on three typical types of mobile devices, including high-end, mid-range, and entry-level devices. The test metrics included data processing latency, model inference latency, power consumption per unit time, and memory usage. The comparison model was a baseline model without lightweight optimization. The results are shown in Table 3. As shown in Table 3, the proposed model exhibits excellent lightweight performance across devices with different configurations. On high-end devices, the total latency is only 198 ms, which is reduced by 52.3% compared with the baseline model. On mid-range devices, the total latency is 265 ms, meeting the real-time response requirements of mobile applications. On entry-level devices, the total latency is 342 ms, which is reduced by 48.7% compared with the baseline model. In terms of power consumption, the proposed model achieves power consumption per unit time of 2.3 W, 2.1 W, and 1.9 W on high-end, mid-range, and entry-level devices, respectively, with an average reduction of 41.5% compared with the baseline model. In terms of memory usage, the proposed model achieves an average memory footprint of 186 MB, which is reduced by 56.8% compared with the baseline model.

Table 3. Results of mobile-side performance tests

Device Type	Model	Data Processing Latency (ms)	Model Inference Latency (ms)	Total Latency (ms)	Power Consumption per Unit Time (W)	Memory Usage (MB)
High-end	Baseline Model	156	248	404	3.9	431
High-end	Proposed Model	82	116	198	2.3	186
Mid-range	Baseline Model	189	312	501	3.6	428
Mid-range	Proposed Model	105	160	265	2.1	184
Entry-level	Baseline Model	245	438	683	3.2	425
Entry-level	Proposed Model	142	200	342	1.9	182

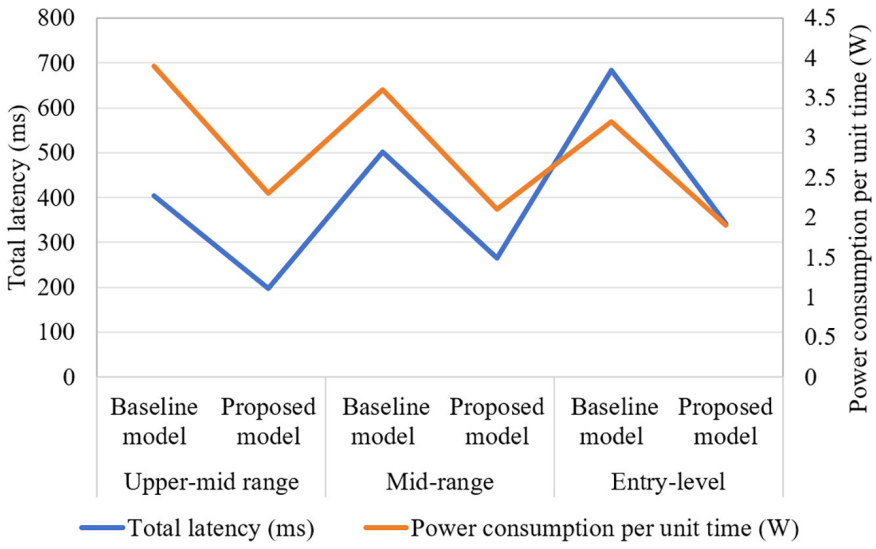

Fig. 3. Mobile-side deployment performance: trade-off analysis between latency and power consumption

The above results verify the effectiveness of the lightweight design of the proposed model. Specifically, structured pruning and quantization optimization in the multi-source fusion module, parameter sharing strategies in the temporal modeling

module, and parallel scheduling design for on-device tasks significantly reduce computational complexity and resource consumption. These optimizations ensure efficient and stable operation of the model on mobile devices with different configurations, satisfying the low-latency and low-power requirements of mobile sports teaching scenarios. Figure 3 visualizes the data in Table 3. The x-axis represents device types, the left y-axis represents total latency, and the right y-axis represents power consumption. Different colored lines or bars are used to compare the "baseline model" and the "proposed model," intuitively demonstrating the advantages of lightweight design.

Table 4. Ablation experiment results

Model Version	Tri-Dimensional State Evaluation Accuracy	Path Optimization Effect	Mobile-Side Performance	Comprehensive Performance Score	Performance Degradation	Module Contribution
Complete Model	0.912	0.896	0.820	0.876	–	–
Without Multi-source Fusion Module	0.725	0.683	0.586	0.652	25.6%	25.6%
Without Tri-dimensional Evaluation Dimensions	0.758	0.712	0.605	0.689	21.3%	21.3%
Without Bi-level Optimization Mechanism	0.786	0.756	0.632	0.724	17.4%	17.4%
Without Spatiotemporal Attention Module	0.765	0.724	0.605	0.698	20.3%	20.3%

To verify the necessity of each core module, ablation experiments were designed by sequentially removing the multi-source fusion module, tri-dimensional evaluation dimensions, bi-level optimization mechanism, and spatiotemporal attention module. The performance differences between each ablated version and the complete model were compared to quantify the contribution of each module. The results are shown in Table 4. As shown in Table 4, the complete model achieves a comprehensive performance score of 0.876, while all ablated versions exhibit varying degrees of performance degradation. After removing the multi-source fusion module, the comprehensive performance score drops to 0.652, with a contribution rate of 25.6%, indicating that effective fusion of multi-source heterogeneous data is the foundation for ensuring evaluation and optimization accuracy. After removing the tri-dimensional evaluation dimensions, the comprehensive performance score drops to 0.689, with a contribution rate of 21.3%, verifying the critical role of multi-dimensional state evaluation in personalized path adaptation. After removing the bi-level optimization mechanism, the comprehensive performance score drops to 0.724, with a contribution rate of 17.4%, highlighting the importance of edge–cloud collaborative optimization in balancing personalization and global optimality. After removing the spatiotemporal attention module, the comprehensive performance score drops to 0.698, with a contribution rate of 20.3%, further confirming the core role of this module in enhancing temporal feature representation capability.

The quantitative results of module contributions indicate that the multi-source fusion module and tri-dimensional evaluation dimensions are the core foundational modules of the proposed model, while the spatiotemporal attention module and bi-level optimization mechanism are key innovative modules for performance improvement. The synergistic effect of all modules ensures the comprehensive

advantages of the model and validates the rationality and completeness of the proposed architectural design.

4 CONCLUSION

This study addressed the core problems of insufficient dynamic adaptability of recommendation models, inadequate multimodal data fusion, and poor path adaptability in mobile-side personalized sports teaching. A three-layer edge–cloud collaborative architecture integrating multi-source perception, dynamic cognition, and real-time optimization was proposed, and three core innovative modules were designed: a lightweight real-time multi-source heterogeneous data fusion module, a tri-dimensional dynamic state evaluation model of skill mastery–fatigue–interest, and a bi-level optimization mechanism based on online learning. Experimental results demonstrated that the proposed model achieved superior performance in tri-dimensional state evaluation tasks, with average evaluation errors across dimensions reduced by 29.8% compared with baseline models. The introduction of the spatiotemporal attention module improved performance by 30.8%, effectively ensuring evaluation accuracy. In terms of path optimization, the generated learning paths achieved a coherence score of 0.896 and a user satisfaction score of 4.62. The meta-learning and bi-level optimization mechanisms ensured personalized adaptation for learners with different skill levels, with skill improvement efficiency reaching 42.3% for beginner participants. Mobile-side performance tests show that the model achieved efficient operation on high-end, mid-range, and entry-level devices, with a minimum total latency of 198 ms, an average reduction in power consumption of 41.5%, and a reduction in memory usage of 56.8%, demonstrating the effectiveness of the lightweight design. Ablation experiments further confirmed the necessity of each core module, among which the multi-source fusion module showed the highest contribution, and the synergistic effects of all modules ensure the comprehensive advantages of the model.

5 ACKNOWLEDGEMENTS

This paper was funded by the Research and Practice Project on Education and Teaching Reform of Hebei Finance University (Grant No.: 2025TML08), titled "Digital and Intelligent Transformation of Physical Education in Finance and Economics Colleges: Construction of a Dual-driven Path of 'Digital and Intelligent Empowerment – Finance and Economics Integration' and Educational Reform Practice."

6 REFERENCES

[1] M. Peng, "Evaluating the effectiveness of mobile interactive technology in university physical education courses," *International Journal of Interactive Mobile Technologies*, vol. 18, no. 20, pp. 85–98, 2024. https://doi.org/10.3991/ijim.v18i20.51883

[2] A. Martín-Rodríguez and R. Madrigal-Cerezo, "Technology-enhanced pedagogy in physical education: Bridging engagement, learning, and lifelong activity," *Education Sciences*, vol. 15, no. 4, p. 409, 2025. https://doi.org/10.3390/educsci15040409

[3] S. Wu and Y. Zhan, "Development and optimization of a physiological data-driven mobile interactive feedback system for university physical education," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 22, pp. 154–168, 2025. https://doi.org/10.3991/ijim.v19i22.59061

[4] N. Mulato, F. Hidayatulloh, S. K. Purnama, and R. Syaifullah, "Optimization of learning physical education in digital era: A systematic review," *Retos*, vol. 54, no. 1, pp. 844–849, 2024. https://doi.org/10.47197/retos.v54.105211

[5] A. Rodríguez-Cayetano, S. Pérez-Muñoz, D. Neila-Simón, and P. T. Morales-Campo, "Effect of the use of new technologies on mental health in physical education students: A systematic review," *Education Sciences*, vol. 15, no. 10, p. 1282, 2025. https://doi.org/10.3390/educsci15101282

[6] C. Kwok, P. Y. Leung, K. Y. Poon, and X. C. Fung, "The effects of internet gaming and social media use on physical activity, sleep, quality of life, and academic performance among university students in Hong Kong: A preliminary study," *Asian Journal of Social Health and Behavior*, vol. 4, no. 1, pp. 36–44, 2021. https://doi.org/10.4103/shb.shb_81_20

[7] J. O'loughlin, D. N. Chróinín, and D. O'Grady, "Digital video: The impact on children's learning experiences in primary physical education," *European Physical Education Review*, vol. 19, no. 2, pp. 165–182, 2013. https://doi.org/10.1177/1356336X13486050

[8] C. W. Wu *et al.*, "Enhancing fan engagement in a 5G stadium with AI-based technologies and live streaming," *IEEE Systems Journal*, vol. 16, no. 4, pp. 6590–6601, 2022. https://doi.org/10.1109/JSYST.2022.3169553

[9] S. Paine, A. Ghosh, A. Choudhury, S. Akbar, A. Ray, and K. Mukherjee, "Smartphone-enabled biomedical sensors for healthcare applications: progress and perspective," *Measurement*, vol. 256, no. 1, p. 118466, 2025. https://doi.org/10.1016/j.measurement.2025.118466

[10] Y. Li, A. Fayazi, and H. G. Zadeh, "Federated knowledge graph-based hybrid personalized curriculum recommendation for privacy-preserving physical education teaching systems," *Journal of Cloud Computing*, vol. 15, no. 1, pp. 1–15, 2025. https://doi.org/10.1186/s13677-025-00812-2

[11] N. J. Kirk-Sanchez, J. G. Moore, G. W. Hartley, and M. Wong, "The use of movement scripts for clinical reasoning in physical therapist education and practice," *Physical Therapy*, vol. 102, no. 7, p. pzac061, 2022. https://doi.org/10.1093/ptj/pzac061

[12] V. Camomilla, E. Bergamini, S. Fantozzi, and G. Vannozzi, "Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: A systematic review," *Sensors*, vol. 18, no. 3, p. 873, 2018. https://doi.org/10.3390/s18030873

[13] S. J. Preece, J. Y. Goulermas, L. P. Kenney, D. Howard, K. Meijer, and R. Crompton, "Activity identification using body-mounted sensors—a review of classification techniques," *Physiological Measurement*, vol. 30, no. 4, p. R1, 2009. https://doi.org/10.1088/0967-3334/30/4/R01

[14] F. M. Clemente, M. S. Couceiro, F. M. Martins, R. Mendes, and A. J. Figueiredo, "Measuring tactical behaviour using technological metrics: Case study of a football game," *International Journal of Sports Science & Coaching*, vol. 8, no. 4, pp. 723–739, 2013. https://doi.org/10.1260/1747-9541.8.4.723

[15] R. T. Thorpe, G. Atkinson, B. Drust, and W. Gregson, "Monitoring fatigue status in elite team-sport athletes: Implications for practice," *International Journal of Sports Physiology and Performance*, vol. 12, no. s2, pp. S2–27, 2017. https://doi.org/10.1123/ijspp.2016-0434

[16] J. K. Tarus, Z. Niu, and D. Kalui, "A hybrid recommender system for e-learning based on context awareness and sequential pattern mining," *Soft Computing*, vol. 22, no. 8, pp. 2449–2461, 2018. https://doi.org/10.1007/s00500-017-2720-6

7 AUTHORS

Jianxin Zhao obtained her bachelor's degree from the Faculty of Physical Education at Shanxi Normal University. Then she obtained her master's degree at Xi'an Physical Education University and currently works at Hebei Finance University. She has published seven papers, two of which are indexed by SCI. Her research interests include physical education teaching and sports training and social sports (E-mail: zhaojianxin@hbfu.edu.cn).

Hui Zhu obtained her master's degree at Hebei Normal University and currently works at Baoding Preschool Teachers College. Her research interests include physical education teaching and sports training and social sports (E-mail: zhuhui3228@163.com).