

# Research on Cognitive Augmented Real-Time Decision-Making Technology Based on Cross-Modal Biological Signal Fusion

Jiayi Shen, Luyan Shen, Sitong Ruan, Zhihui Xu, Jingyi Xu\*

*Artificial Intelligence College, Zhejiang Dongfang Polytechnic, Wenzhou, Zhejiang, China*

*\*Corresponding Author*

**Abstract:** Aiming at the industrial pain points in the current rehabilitation medical field, including the shortage of therapists, the lack of real-time professional guidance for patients during independent home training, non-standard training movements, and inconsistent rehabilitation effects, this paper proposes a cognitive augmented real-time decision-making technology based on cross-modal biological signal fusion. Relying on VR intelligent rehabilitation equipment, this technology integrates three heterogeneous biological and motion perception signals, namely IMU inertial sensing, eye tracking, and brain-computer interface signals, to construct a multi-dimensional and high-precision patient rehabilitation training environment and state perception system. Based on massive standardized patient rehabilitation training datasets provided by hospitals, special fine-tuning and training of AI large models are completed, endowing the models with full-process intelligent capabilities including rehabilitation movement recognition, error deviation judgment, real-time correction guidance, and training effect evaluation, thereby forming a mature real-time intelligent feedback, error correction and evaluation function. Experimental and application results show that the proposed technology can accurately identify movement deviations, state abnormalities and insufficient cognitive cooperation of patients during rehabilitation training without on-site guidance from professional therapists, and output standardized correction decision instructions in real time. It significantly improves the accuracy and standardization of patients' independent training, effectively solves the problem of unbalanced supply and demand of rehabilitation medical resources, and provides a new technical support for intelligent, inclusive and normalized home-

based rehabilitation and out-of-hospital continuous rehabilitation.

**Keywords:** Cross-Modal Signal Fusion; Biological Signal Perception; AI Large Model; Real-Time Intelligent Error Correction; Rehabilitation Training; Cognitive Augmentation

## 1. Introduction

### 1.1 Research Background

With the aggravation of population aging and the continuous growth of patients with chronic diseases and sports injuries, the social demand for rehabilitation medical services has shown an explosive growth trend. As a core method for patients with functional disorders to recover limb, cognitive and motor abilities, the standardization, sustainability and accuracy of rehabilitation training directly determine the prognosis of rehabilitation. The traditional rehabilitation model is highly dependent on one-to-one on-site guidance, movement correction and effect evaluation by professional therapists, and the training process is restricted by multiple factors such as human resources, venue and time. At present, the rehabilitation medical industry in China generally faces core problems including a serious shortage of professional therapists, excessive number of patients served per capita, and a lack of out-of-hospital rehabilitation guidance. A large number of patients suffer from deformed training movements, incorrect force application modes and inappropriate training intensity due to the lack of real-time and professional movement correction and state guidance during independent home training without therapist supervision. These problems not only greatly reduce the efficiency of rehabilitation training, but also may cause secondary injuries due to incorrect training, seriously restricting the improvement of overall rehabilitation medical

quality and the inclusive development of rehabilitation services.

Meanwhile, the rapid iteration of Virtual Reality (VR) perception technology, brain-computer interface technology, multi-modal sensor fusion technology and general artificial intelligence large models has provided a new path for technological innovation in the field of intelligent rehabilitation. The traditional single-sensor monitoring mode can only collect patients' limb movement posture or single physiological signal, and cannot fully cover patients' limb movement state, visual fixation state and brain cognitive activation state during training. It has defects such as single perception dimension, one-sided data and large decision deviation, making it difficult to achieve high-precision training error correction and intelligent evaluation. In contrast, cross-modal biological signal fusion technology can integrate multi-source heterogeneous perception data, make up for the information shortcomings of single sensors, and realize full-process intelligent perception, analysis and feedback of rehabilitation training combined with the powerful feature learning and decision reasoning capabilities of AI large models, providing core technical support for solving the pain points of traditional rehabilitation.

### 1.2 Research Status

Current domestic and foreign research on intelligent rehabilitation mainly focuses on single-sensor monitoring or simple data statistical analysis. Most systems only collect patients' limb movement trajectories through inertial sensors to complete basic movement compliance judgment, lacking synchronous perception of patients' cognitive state and visual cooperation state, and unable to realize full-dimensional state evaluation of "limb movement-cognitive participation-visual cooperation". Some studies have introduced eye tracking or brain-computer sensing technology, but most of them adopt independent single-modal application without in-depth fusion and correlation analysis of multi-source signals, resulting in low data utilization and insufficient intelligent decision accuracy [1].

In terms of artificial intelligence application, most existing intelligent rehabilitation systems adopt traditional machine learning models or lightweight algorithms that rely on manual feature extraction and have poor generalization

ability. They can only adapt to fixed and standardized rehabilitation movements, failing to cope with individual differences of patients and diverse scenarios of movement deviations, and difficult to realize refined and personalized real-time error correction. At the same time, most models are not trained based on large-scale clinical patient data, lacking medical professionalism and clinical adaptability. There are large deviations between intelligent evaluation results and manual evaluation standards of professional therapists, which limits their application in actual clinical rehabilitation scenarios. In general, existing technologies are plagued by problems such as single perception dimension, weak cross-modal fusion capability, insufficient professionalism of intelligent decision-making and delayed real-time feedback, which cannot meet the high-precision independent rehabilitation training requirements without therapist guidance.

### 1.3 Research Contents and Innovations

Focusing on the core goals of real-time intelligent error correction and accuracy improvement of independent rehabilitation training, this paper carries out research on cognitive augmented real-time decision-making technology based on cross-modal biological signal fusion. The core research contents include: constructing a VR multi-sensor integrated perception system to realize synchronous collection and preprocessing of IMU motion data, eye tracking data and brain-computer biological signals; designing a multi-modal signal in-depth fusion algorithm to solve the problems of heterogeneous data dimension differences, noise interference and time series misalignment; completing special training and fine-tuning of AI large models based on massive clinical patient training data from hospitals to build a rehabilitation-specific intelligent decision model; developing a real-time intelligent feedback, error correction and evaluation function to realize full-process automation of training movement deviation identification, error grading, real-time correction guidance and quantitative evaluation of training effects [2].

The core innovations of this paper are as follows. First, it innovatively integrates three types of heterogeneous perception signals, breaks through the limitations of single-modal perception, and realizes full-dimensional and

accurate perception of patients' training motion state, cognitive state and visual cooperation state. Second, the AI large model is trained based on real clinical patient data, abandoning the limitations of manual rules. The decision logic of the model conforms to the professional diagnosis and treatment standards of therapists, with strong clinical professionalism and scenario generalization. Third, it realizes cognitive augmented real-time decision-making [3], incorporating patients' cognitive state into the training evaluation system, which is different from the traditional evaluation mode focusing only on limb movements and improves the comprehensiveness and scientificity of rehabilitation evaluation. Fourth, it supports the implementation of independent and normalized intelligent rehabilitation scenarios, effectively makes up for the shortage of therapist manpower, and realizes the inclusive extension of rehabilitation medical resources [4].

## 2. System Overall Architecture and Technical Principles

### 2.1 System Overall Architecture

The cognitive augmented real-time decision system designed in this paper consists of four layers: multi-modal perception and collection layer, data preprocessing and fusion layer, AI model training and decision layer, and real-time feedback error correction application layer. All layers cooperate to realize a closed loop of full-process intelligent perception, analysis, decision-making and feedback for rehabilitation training. The overall architecture is lightweight with high real-time performance and can be deployed locally on VR terminals [5].

With VR helmets as the core carrier, the multi-modal perception and collection layer integrates high-precision IMU inertial sensors, eye tracking sensors and brain-computer interface sensors. It synchronously collects multi-dimensional data such as limb posture, motion trajectory, motion speed, visual fixation points, fixation duration, brain cognitive activation intensity and nerve response signals during patients' rehabilitation training, realizing all-round dynamic perception of training scenarios and patient states, and providing complete data support for subsequent intelligent decision-making.

Aiming at the problems of heterogeneity, time series asynchrony and noise interference of the

three types of signals, the data preprocessing and fusion layer successively completes data cleaning, denoising, time series alignment and dimension normalization. It adopts a cross-modal attention fusion algorithm to mine correlation features of multi-source data, eliminate invalid and redundant information, retain core rehabilitation state features, and form standardized and structured fused feature datasets for input into AI large models for training and inference [6].

Relying on massive real clinical rehabilitation training data of patients accumulated by cooperative hospitals, the AI model training and decision layer conducts special fine-tuning training on general AI large models. It enables the models to learn professional rehabilitation movement standards, common error types, deviation thresholds, and correlation relationships between cognitive states and training quality, so as to construct an intelligent decision model with professional rehabilitation cognition, realizing automatic movement error identification, quantitative deviation degree judgment, intelligent generation of correction strategies and quantitative evaluation of training effects [7].

Based on model inference results, the real-time feedback error correction application layer builds a real-time intelligent feedback, error correction and evaluation function. It identifies wrong movements at the millisecond level during patient training, outputs real-time correction instructions through VR visual prompts, voice guidance and posture correction prompts, and automatically generates quantitative evaluation reports after training, completing full-process intelligent auxiliary rehabilitation [8].

### 2.2 Technical Principles of Multi-Modal Signal Perception

The three core perception signals adopted by the system are complementary, jointly constructing a full-dimensional rehabilitation state perception system. Among them, the VR helmet IMU inertial sensor serves as the core of motion perception. It captures kinematic parameters such as patients' head and limb motion trajectories, posture offsets, movement amplitudes and motion rates in real time through three-axis acceleration, three-axis angular velocity and three-axis posture angle data, accurately judging the completion form

and standard degree of patients' training movements and solving the problem of limb motion state perception.

The eye tracking sensor is the core of visual cooperation perception. By capturing data such as patients' eyeball rotation trajectories, fixation areas, gaze duration and blink frequency, it analyzes the concentration of attention and visual cooperation degree during training, judges whether patients accurately follow training instructions to complete movement cooperation, and identifies negative states such as scattered attention and visual offset that affect training quality [9].

The brain-computer sensor is the core of cognitive state perception. It collects electroencephalogram signals of patients during training in real time, extracts cognitive features such as cognitive activation, attention investment and nerve response intensity, and quantifies patients' subjective cognitive participation during training. Different from the traditional pure limb movement evaluation mode, it realizes dual state evaluation of "movement execution + cognitive participation" and achieves the core goal of cognitive augmentation.

### 2.3 Cross-Modal Signal Fusion Algorithm

IMU motion signals, eye movement signals and brain-computer signals differ greatly in data dimension, sampling frequency and physical characteristics, and direct fusion is prone to feature confusion and loss of effective information. This paper adopts a cross-modal attention weighted fusion algorithm to realize in-depth fusion of multi-source signals. Firstly, independent feature extraction is performed on the three types of signals to screen motion features, visual features and cognitive features respectively. Secondly, an attention mechanism is introduced to dynamically allocate weights of each modal feature according to different training scenarios, strengthen weights of core features affecting training quality and weaken weights of redundant noise features. Finally, feature splicing and dimension fusion are completed to generate unified cross-modal fused feature vectors. This algorithm can effectively solve the heterogeneous data fusion problem, improve the accuracy of the model in identifying patients' training states, movement errors and cognitive deviations, and provide core data support for high-precision intelligent

decision-making.

### 2.4 Special Training Mechanism of AI Large Model

To ensure the professionalism and accuracy of intelligent decision-making, this study abandons the direct inference mode of general models and adopts a special fine-tuning training scheme based on clinical data. The training dataset is derived from real rehabilitation training data of patients in cooperative hospitals, covering standard training movements, common wrong movements, cognitive state changes, manual correction records and effect evaluation data of therapists for patients of different ages, diseases and rehabilitation stages. The dataset has a large sample size, comprehensive scenario coverage and high clinical professionalism [10]. In the training process, the preprocessed cross-modal fused feature data are paired and trained with professional labeled data of therapists, enabling the AI large model to deeply learn professional standards in the rehabilitation field, including the threshold range of standard movements, feature differences of various movement errors, error severity grading, corresponding standardized correction schemes, and correlation laws between cognitive states and training effects. Through multiple iterative training, parameter optimization and model verification, the final model can accurately replace manual work to complete professional tasks such as movement correction, state evaluation and effect evaluation, with stable and reliable clinical-level intelligent decision-making capability.

### 3. Core Function Design: Real-Time Intelligent Feedback, Error Correction and Evaluation Function

Based on the achievements of cross-modal signal fusion and AI large model training, this study develops the core function—real-time intelligent feedback, error correction and evaluation function. As the core carrier to solve the shortage of rehabilitation manpower and realize unmanned high-precision rehabilitation training, this function mainly includes four core modules: real-time movement error identification, dynamic intelligent correction feedback, quantitative training effect evaluation and personalized training optimization.

Relying on fused multi-modal data and the trained AI model, the real-time movement error

identification module compares patients' real-time training movements with clinical standard movements at the millisecond level, and accurately identifies limb problems such as insufficient movement amplitude, posture offset, incorrect force application and rhythm disorder. Meanwhile, combined with eye movement and brain-computer signals, it identifies implicit training problems such as inattention and insufficient cognitive participation, realizing dual identification of explicit movement errors and implicit cognitive deviations.

According to the error type and deviation degree judged by the model, the dynamic intelligent correction feedback module matches the corresponding professional correction strategies, and outputs real-time visual correction prompts and voice guidance instructions through VR equipment to guide patients to adjust movement postures and correct wrong force application modes in real time. It forms a closed loop of "error occurrence-real-time identification-instant correction", avoiding the accumulation of wrong training and greatly improving the accuracy of independent training.

After a single training session, the quantitative training effect evaluation module integrates the whole-process motion data, visual cooperation data and cognitive state data, and generates quantitative indicators such as training standardization score, movement accuracy rate, cognitive participation score and training compliance rate automatically combined with the AI model's professional evaluation system. It synchronously generates professional training evaluation reports that accurately align with therapists' manual evaluation standards, providing objective and comprehensive training effect references for patients and medical staff.

The personalized training optimization module iteratively adapts to individual training characteristics of patients based on their historical training data and error correction records, dynamically adjusts error correction sensitivity and training difficulty, and strengthens guidance for patients' high-frequency wrong movements. It realizes personalized intelligent rehabilitation guidance tailored to each patient and solves the problem of poor adaptability of traditional standardized training.

#### **4. Technical Application Value and**

#### **Advantage Analysis**

##### **4.1 Solving Core Industry Pain Points**

The core contradiction in the current rehabilitation medical field is the mismatch between the growing rehabilitation demand and scarce professional therapist resources. In the traditional model, patients have no professional guidance during out-of-hospital independent training, resulting in unguaranteed training quality. Rehabilitation resources are concentrated in hospitals with insufficient inclusiveness. By replacing manual work with AI intelligent decision-making to complete real-time error correction and effect evaluation, this technology breaks the human resource dependence of rehabilitation training on therapists, effectively reduces the work pressure of therapists, fills the gap of out-of-hospital rehabilitation guidance, enables patients to enjoy professional-level rehabilitation guidance without on-site accompaniment, and greatly improves the coverage and efficiency of rehabilitation medical services.

##### **4.2 Core Technical Advantages**

Compared with traditional single-sensor intelligent rehabilitation technology and traditional manual rehabilitation mode, this technology has multiple core advantages. First, it has a more comprehensive perception dimension, integrating three types of cross-modal biological signals of motion, vision and cognition, breaking through the limitations of single data perception and realizing full-dimensional rehabilitation state evaluation. Second, it has higher decision-making professionalism. Trained based on massive clinical patient data, the model's decision-making conforms to medical professional standards and is superior to traditional manual rule algorithms. Third, it has higher real-time feedback performance, realizing millisecond-level identification and correction, dynamic intervention in the training process and elimination of wrong training accumulation. Fourth, it has wider adaptability, applicable to patients with different diseases and rehabilitation stages, and supporting multiple scenarios such as in-hospital auxiliary training, home independent training and remote rehabilitation management. Fifth, it realizes cognitive augmented rehabilitation, incorporating patients' cognitive participation

state into the evaluation system, taking into account both limb function recovery and cognitive ability improvement, and conforming to the modern concept of precision rehabilitation.

#### 4.3 Application and Promotion Value

This technology can be widely applied in hospital rehabilitation departments, community rehabilitation centers, home rehabilitation, remote rehabilitation medical treatment and other scenarios. It can not only serve as an auxiliary tool for therapists to improve the efficiency and standardization of in-hospital rehabilitation training, but also act as an independent intelligent rehabilitation terminal to provide normalized and accurate intelligent training guidance for home patients. The implementation of this technology can effectively reduce the human cost of rehabilitation medical treatment, improve the quality of national rehabilitation training, promote the intelligent, digital and inclusive transformation of rehabilitation medical treatment, and has high clinical application value and social promotion value.

### 5. Conclusion and Prospect

#### 5.1 Research Conclusion

The cognitive augmented real-time decision-making technology based on cross-modal biological signal fusion proposed in this paper realizes real-time intelligent feedback error correction and quantitative evaluation in rehabilitation training scenarios by integrating VR equipment's IMU, eye tracking and brain-computer multi-source perception signals and combining with clinical big data-driven AI large model training. The cross-modal signal fusion technology effectively solves the problems of one-sided single perception dimension and insufficient accuracy of state identification. The cognitive augmentation mechanism optimizes the traditional rehabilitation evaluation system that only focuses on limb movements. The AI model trained with clinical data has professional-level intelligent decision-making capability and can completely replace manual work to complete rehabilitation training error correction and evaluation in non-guided scenarios. This technology fundamentally solves the industrial problems such as shortage of rehabilitation

therapists, low accuracy of patients' independent training and lack of out-of-hospital rehabilitation guidance, significantly improves the standardization, accuracy and effectiveness of rehabilitation training, and provides a brand-new technical scheme and implementation path for the intelligent rehabilitation field.

#### 5.2 Future Prospect

Future research will further optimize the cross-modal fusion algorithm to improve the signal adaptability and real-time decision-making speed in complex training scenarios; continuously expand the clinical training dataset to cover more disease types and rehabilitation scenarios, and enhance the model's generalization ability and personalized adaptation accuracy. At the same time, combined with the Internet of Things and telemedicine technology, we will build a cloud intelligent rehabilitation management platform to realize cloud storage of patient training data, remote supervision by medical staff and dynamic iterative upgrading of rehabilitation schemes, and construct a full-process intelligent rehabilitation ecosystem of "perception-analysis-decision-feedback-iteration", so as to further promote the large-scale implementation and application of precision rehabilitation and smart medical care.

#### References

- [1] Liu C L, Jin Y R, Wang M X. Research on Intelligent Health Perception Technology Based on Multi-Modal Physiological Signal Fusion. *Information Fusion*, 2026(02):112-120.
- [2] Wu K, et al. BrainFusion: Standardization Framework of Multi-Modal Biological Signals and Its Application in Intelligent Rehabilitation. *Advanced Science*, 2025(05):89-97.
- [3] Zhang M, Li R. Multi-Modal Physiological Signal Feature Fusion Algorithm Based on Attention Mechanism. *Computer Engineering and Applications*, 2025(18):134-141.
- [4] Wang H Y, Chen X. Research Progress of AI-Driven Real-Time Feedback Technology for Intelligent Rehabilitation. *Chinese Medical Equipment Journal*, 2024(09):45-50.
- [5] Li Y T, et al. Review of Cross-Modal Biosensing and AI Decision-Making in Rehabilitation Medical Applications.

- Journal of Biomedical Engineering, 2026(01):210-218.
- [6] Zhang Yangzhou, Jiang Ziyi, Shi Na. Research on the Long-term Mechanism of the “Five Integrations” Textbook Construction in Higher Vocational Colleges in the New Era. *Journal of Zhejiang Industry & Trade Vocational College*, 2023(03): 47-51.
- [7] Liu Fen, Li Yajing. Current Situation, Direction and Strategies of the Construction of New “Paper-Digital Integration” Textbooks in Vocational Education. *Journal of Hubei Open Vocational College*, 2024, 37(22): 52-54.
- [8] Ji Zhaohua. Exploration and Practice of Digital Textbook Development for Computer Specialty Courses in Vocational Education. *E-Learning*, 2024, (12): 72-74.
- [9] Kong Ling, Yuan Lihui. Research on the Construction of Digital Textbooks for Computer Networks in Application-oriented Undergraduate Education. *High-Technology & Industrialization*, 2024, 30(09): 124-125.
- [10] Liu Chao, Huang Ronghuai, Wang Hongyu. Exploration on the Construction and Application Path of New Textbooks Based on Knowledge Graph. *China University Teaching*, 2023(8): 10-16.