Applying TCN-Based Rhythmic Electroencephalogram Data to Assist in the Construction of Autism Intervention Pathways in Music Therapy

Anshuang Zhang[1], Wei Sun[2], Boyuan Song[3]
[1]Music Therapy Teaching and Research Office, Changzhi Medical College, Changzhi 046000, Shanxi, China
[2]Department of Fine Arts, Jinzhong University, Jinzhong 030619, Shanxi, China
[3]Nanjing normal university of special education, Nanjing 210038, Jiangsu, China

Corresponding author: Boyuan Song (e-mail: xinfunihao1199@163.com)

ABSTRACT This study addresses the challenge of personalizing music therapy pathways for autism spectrum disorder (ASD) by developing a rhythmic electroencephalogram (EEG) decoding framework based on Temporal Convolutional Networks (TCN). Considering EEG as a representative bioelectromagnetic signal, the proposed approach provides an effective means for extracting neural dynamics relevant to individualized intervention assessment. Pre- and post-intervention EEG data were collected from 55 children with ASD during a 12-week music therapy program. The TCN model achieved an accuracy of 0.861 (F1 = 0.843) in classifying treatment responders, outperforming baseline models including Transformer (accuracy = 0.816). For regression, it attained an R2 of 0.692 (RMSE = 6.52) in predicting changes in the Social Responsiveness Scale (ΔSRS). Responders exhibited significantly higher Beta/Gamma relative power and Theta/ Alpha-weighted Phase Lag Index (wPLI). A Neural Response Index (NRI), derived from TCN-decoded features, showed a strong correlation with ΔSRS ($r = 0.712$, $p < 0.001$). Empirical application of NRI-based decision rules yielded an 88.9% response rate in the high-NRI pathway. These findings demonstrate that TCN can effectively decode bioelectromagnetic neural patterns associated with therapeutic outcomes and that NRI provides a quantitative, data-driven indicator for personalized music therapy planning, offering methodological insights for advanced electrophysiological signal analysis and intelligent biomedical sensing applications.

INDEX TERMS rhythmic EEG data, temporal convolutional network, bioelectromagnetic signal analysis, music therapy, autism intervention pathway

I. INTRODUCTION

AUTISM spectrum disorder (ASD) is a common neurodevelopmental disorder characterized by social communication deficits and the presence of repetitive, stereotyped behaviors, which leads to a significant burden for the individual, their families, and society overall. Although behavioral and pharmacological interventions are currently the most common methods of intervention, they also have limitations, including inconsistent effectiveness and large individual variability [1-2]. Music therapy, an emerging neuromodulatory intervention with potential for improving emotional regulation and social interaction in ASD patients due to its structured and rhythmic benefits[3-4]. Unfortunately, these advancements in neuromodulation therapies have long suffered from a core challenge: efficacy assessments have relied primarily on subjective behavioral scales and lacked objective, quantitative neurophysiological indicators of efficacy.

At the computational level, machine learning and deep learning models have been attempted for the assisted diagnosis and biomarker mining of ASD. Early studies mostly used traditional models such as support vector machines, but their performance heavily depended on the quality of manual feature extraction [5]. Subsequently, recurrent neural networks and convolutional neural networks are applied to automatically learn EEG features [6-7]. In summary, while existing research has made progress, it generally suffers from three shortcomings: First, the models fail to optimally balance long-term dependency capture, computational efficiency, and training stability; second, research objectives are mostly focused on static classification or diagnosis, rarely combined with dynamic intervention processes; third, the model outputs have failed to be transformed into quantitative and actionable indicators that can guide clinical intervention decisions.

To systematically address these limitations, this study

employs a Temporal Convolutional Network (TCN), which effectively models long-range dependencies in temporal data while maintaining computational efficiency and training stability. Building on this, we propose an integrated framework spanning from neural signal decoding to intervention pathway construction. The contributions of this work are threefold. First, methodologically, we demonstrate the superiority of TCN in identifying ASD music therapy responders and extracting meaningful neural features, offering a viable solution to balance temporal modeling depth with computational tractability. Second, we introduce a Neural Response Index (NRI) derived from TCN-decoded features, translating multidimensional EEG patterns into a single, interpretable metric that bridges neural dynamics and clinical decision-making. Third, we establish a set of data-driven, dynamically adaptable decision rules for music therapy intervention pathways based on NRI, providing empirical support for moving toward more individualized, neuro-informed intervention strategies.

II. METHODS

DATA ACQUISITION AND PREPROCESSING

EEG DATA ACQUISITION

This study employs high-density EEG to acquire neurophysiological data in both resting and task states [8-9]. All data are acquired using a 30- lead electrode system with matching electrode caps. The electrode layout strictly adheres to the international 10-20 standard system, covering major brain regions including Fp1, Fp2, F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T3, C3, Cz, C4, T4, TP7, CP3, CPz, CP4, TP8, T5, P3, Pz, P4, T6, O1, Oz, and O2 to ensure comprehensive coverage of cerebral cortical activity. During acquisition, the left mastoid process is used as an online reference, and the data is converted to an average reference offline after acquisition. Vertical electrooculography (VEOG) and horizontal electrooculography (HEOG) are collected concurrently for future identification and correction of eye movement artifacts. Data sampling occurs at a rate of 1000 Hz to cover thinner details of neural oscillations at and above the Gamma band. Prior to the commencement of the experiment, the electrical impedance between participant scalp and where electrodes are placed is decreased to below 5 kΩ to ensure signal quality and signal-to-noise ratio. All recordings take place in a professional electromagnetically shielded environment, and participants are instructed to stay relaxed and minimize unnecessary movements [10-11].

DATA PREPROCESSING WORKFLOW

To assure trustworthiness of future analyses, the raw EEG data goes through a set of thorough preprocessing procedures to remove noise caused by various physiological and non-physiological artifacts. This workflow combines the most recognized utilized approaches in the discipline, with the intent to retain EEG signals that are meaningful for driving neurophysiological hypothesis testing as much as possible, while at the same time removing noise, thus providing high-quality, clean inputs of data for future temporal convolutional network models [12-13]. The EEG data preprocessing workflow is presented in Figure 1.

Figure 1 provides an overview of the workflow, beginning with the raw continuous EEG data. Initially, frequency domain filtering is employed to remove irrelevant frequency components: a 1Hz high-pass filter is used to remove slow baseline drift due to sweating, breathing etc., and a 100hz low-pass filter is used to attenuate high-frequency electromyographic noise and environmental noise. A notch filter at 50Hz is also applied simultaneously to remove mains power (50Hz) and harmonic interference.

After removing bad leads and segments, independent component analysis (ICA) was used for further artifact correction. Specifically, the Infomax algorithm was used to perform ICA decomposition on the data, yielding 30 independent components. Subsequently, the following multimodal information was used to identify and remove artifact components: (1) the time series and topology of VEOG and HEOG signals, used to identify components highly correlated with vertical and horizontal eye movements; (2) the time series spectral characteristics of the components (e.g., high-frequency energy prominence may indicate EMG artifacts); and (3) the topological distribution of the components (e.g., concentration in the forehead may indicate eye movement artifacts, while concentration in the occipital region may indicate ECG artifacts). After automatic identification, components identified as artifacts were manually checked and removed, and the remaining components were reconstructed into clean EEG signals. This process effectively removed major physiological artifacts such as eye movement, ECG, and EMG, providing high-quality neural oscillation data for subsequent analysis. After frequency filtering, the operator identified and marked “severe artifact segments” to be excluded through a combination of visual inspection and automated algorithms. Significant signal drift or saturation due to large head or limb movements, complete signal loss due to momentary electrode detachment, and persistent high-frequency noise caused by strong environmental electromagnetic interference are all serious artifacts. These artifacts disrupt the continuity and basic structure of the data and therefore must be removed directly before component analysis. Conversely, physiological artifacts such as blinking and chewing usually manifest as specific spatiotemporal patterns superimposed on the signal. Their waveforms are relatively stable and have separable independent component characteristics, so the data for that time period is retained.

EEG RHYTHMS AND FUNCTIONAL CONNECTIVITY INDICATORS

To provide a valid comparison with end-to-end TCN models, this study also extracts classic rhythmic features based on frequency domain and functional connectivity. These features aim to quantify the intensity of local brain activity across different frequency bands and the collaborative work-

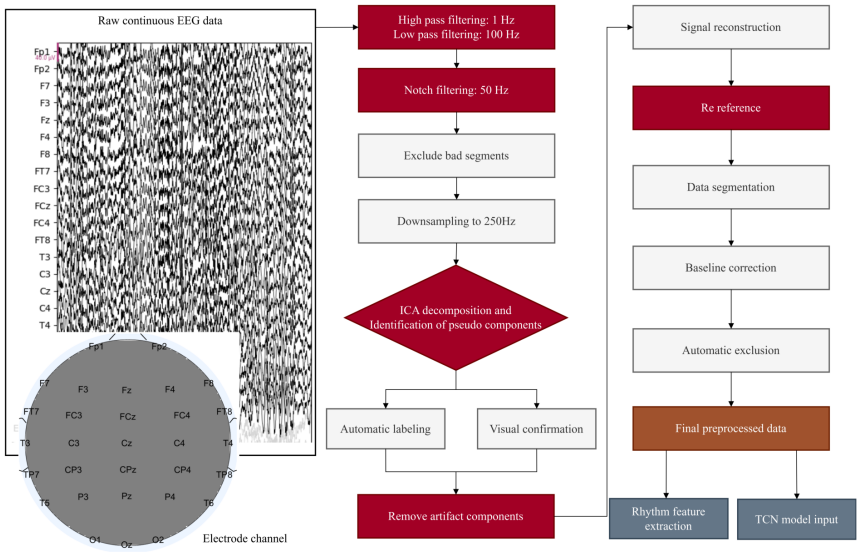

FIGURE 1. EEG data preprocessing workflow

ing patterns across brain regions. The preprocessed clean EEG data is divided into five standard neural oscillation bands by a finite impulse response bandpass filter [14-15]: Delta (δ): 1 - 4 Hz; Theta (θ): 4 - 8 Hz; Alpha (α): 8 - 13 Hz; Beta (β): 13 - 30

Hz; Gamma (γ): 30 - 45 Hz (the upper limit is set to 45 Hz to avoid power frequency noise).

The Welch method is used to estimate the power spectral density of the signal within each frequency band. Continuous data is segmented (using a 2-second Hanning window with 50% overlap in this study), and the averaged values are obtained after Fourier transform to reduce variance. For the i-th frequency band, its absolute power $P_{\text{abs}}(i)$ is obtained by integrating the power spectral density (PSD) over that band. Relative power is an effective indicator for eliminating differences in overall power levels among individuals and can better reflect the balance of activity in different frequency bands. The formula for calculating the relative power $P_{\text{rel}}(i)$ of the i-th frequency band is:

$$P_{\text{rel}}(i) = \frac{P_{\text{abs}}(i)}{\sum_{j=1}^{5} P_{\text{abs}}(j)} \tag{1}$$

In Formula 1, the denominator is the sum of the absolute powers of all five frequency bands (Delta, Theta, Alpha, Beta, Gamma). Finally, five relative power characteristics are calculated for each channel and each trial.

Functional connectivity is used to measure the statistical dependence of neural activity over time in different brain regions. This study calculates two phase-based connectivity indices that are insensitive to volumetric conduction effects.

PLV (Phase Locking Value) measures the stability of the phase difference sequence of two signals at a specific frequency, with a value range of [0, 1] [16-17]. The closer the value is to 1, the stronger the phase synchronization. A Hilbert transform is performed on the signal of each channel to extract the instantaneous phase $\phi(t)$; for a pair of channels m and n, the instantaneous phase difference is calculated: $\Delta\phi_{mn}(t) = \phi_m(t) - \phi_n(t)$. The formula for calculating PLV over a period of time is:

$$\text{PLV}_{mn} = \left| \frac{1}{N} \sum_{t=1}^{N} e^{j\Delta\phi_{mn}(t)} \right| \tag{2}$$

In Formula 2, N is the total number of time points, and j is the imaginary unit. wPLI (Weighted Phase Lag Index) reduces the impact of zero-hysteresis synchronization by weighting the amplitude of the instantaneous phase difference dot product, thus exhibiting stronger robustness against noise and volumetric conduction [18-19]. The calculation formula is:

$$\text{wPLI}_{mn} = \frac{|E\{|\mathcal{I}\{X_{mn}(t)\}| \cdot \text{sign}(\mathcal{I}\{X_{mn}(t)\})\}|}{E\{|\mathcal{I}\{X_{mn}(t)\}|\}} = \frac{|E\{\mathcal{I}\{X_{mn}(t)\}\}|}{E\{|\mathcal{I}\{X_{mn}(t)\}|\}} \tag{3}$$

In Formula 3: I{X_mn(t)} is the imaginary part of the cross spectrum of the signals m and n. E represents the expected value (average over time). The range of wPLI is also [0, 1], with a larger value indicating stronger true phase synchronization.

TCN MODEL ARCHITECTURE AND TRAINING

MODEL INPUT

In this study, the input to the TCN model is preprocessed and segmented single-trial EEG data [20-21]. The input data format is defined as a two-dimensional tensor: number of channels $\times$ time points. Specifically, for the 30-channel EEG system and the segmented 2-second duration data in this study (downsampled to 250Hz, with a single trial length of 500 time points), the input dimension of the model is 30 $\times$ 500. This tensor does not require traditional frequency domain or manual feature extraction; it is directly input into the network as

the original time domain signal or its slightly filtered form, enabling end-to-end feature learning and pattern recognition.

TCN CORE STRUCTURE

Temporal convolutional networks (TCNs) allow for effective modeling of long-sequence temporal data with several designed components. Their strength lies in using dilated causal convolutions to create a large receptive field and maintaining residual connections for training sufficiently deep networks, while also being able to learn complex temporal-dependent dynamics in local and global electrical brain activity from EEGs. The architecture of the TCN model is shown in Figure 2.

The input data goes through initial channel-dimensional feature mapping with a 1x1 convolutional layer. After initial mapping, the core is comprised of several stacked TCN blocks, where each block includes two dilated causal convolutions. The dilation coefficient establishes a depthwise exponential increase, meaning that neurons at a greater depth have receptive fields that cover the entire input sequence. Weight normalization is performed after each convolution to accelerate convergence, and a ReLU (Rectified Linear Unit) activation function is applied to apply non-linearity. Causal convolution ensures that the model does not "peek" into future information when processing time series [22]. This is achieved by padding forward on top of forward convolution. For one-dimensional convolution, zeros are padded to the left of the input sequence so that the receptive field of each point in the output sequence is completely located at the corresponding time point in the input sequence and the time points before it. The output x_t^l of the layer l at time t is:

$$x_t^l = f\left(\sum_{i=0}^{k-1} w_i \cdot x_{t-d\cdot i}^{l-1} + b\right) \tag{4}$$

In Formula 4, k is the kernel size; d is the dilation coefficient; w and b are the weights and biases; f is the activation function.

Dilated convolution expands the receptive field exponentially by inserting holes between kernel elements without increasing the number of parameters [23]. The dilation coefficient d defines the sampling interval of the kernel when processing input data. For a one-dimensional input sequence X, the output of the dilated convolution operation at position t is:

$$F(t) = \sum_{i=0}^{k-1} w_i \cdot X_{t-d\cdot i} \tag{5}$$

When d=1, it is a standard convolution. As the number of layers n increases, d grows exponentially. For a TCN with N stacked layers and kernel size k, the receptive field of the top neuron is calculated as:

$$RF = 1 + 2 \cdot (k-1) \cdot \sum_{i=0}^{N-1} d_i \tag{6}$$

Residual learning addresses the vanishing and degradation problems in deep networks by applying a shortcut path that directly adds the input to the output of the block [24]. A TCN residual block can be formally represented as:

$$\text{Output} = \text{Activation}(\text{Input} + F(\text{Input})) \tag{7}$$

In Formula 7, F represents a series of transformations within the residual block. If the input and output dimensions of the main path F are inconsistent, a linear projection of the Input is required:

$$\text{Output} = \text{Activation}(W_{\text{proj}} \cdot \text{Input} + F(\text{Input})) \tag{8}$$

In Formula 8, W_{proj} represents the weights of the 1x1 convolution.

MODEL TASK SETTING

To comprehensively evaluate the potential of the TCN model in assisting clinical decision-making, this study sets up two tasks: a classification task for early screening and a regression task for refined prediction.

Task 1: Classification (responder vs. non-responder) - based on EEG data from patients in the early stages of music therapy, whether they can eventually become therapy responders is predicted.

Based on the changes in the Social Responsiveness Scale (SRS) before and after the intervention, responders are objectively defined. If a patient's total SRS score decreases by $\geq 30\%$ from baseline after the intervention, they are defined as a "responder" and marked with 1; otherwise, they are defined as a "non-responder" and marked with 0.

This task is a binary classification problem. Therefore, the final output layer of the model uses the Sigmoid activation function to compress the output to the range (0, 1), representing the probability that the input sample is predicted as a "responder". The formula is:

$$\hat{y}_{\text{cls}} = \sigma(W_{\text{cls}} \cdot h + b_{\text{cls}}) \tag{9}$$

In Formula 9, h is the feature vector after global average pooling; W_{cls} and b_{cls} are the weights and biases of the classification output layer; σ is the Sigmoid function.

Task 2: Regression (predicting changes in SRS score) - this task aims to more precisely predict the exact change in a patient's total SRS score, thereby quantifying the degree of symptom improvement. The label for this task is a continuous value, specifically the change in the total SRS score (ΔSRS). The calculation formula is:

$$\Delta\text{SRS} = \text{SRS}_{\text{baseline}} - \text{SRS}_{\text{post-treatment}} \tag{10}$$

Therefore, a positive ΔSRS indicates symptom relief, and a larger value indicates more significant improvement. The output layer uses a linear activation function to directly predict the value of ΔSRS. The formula is:

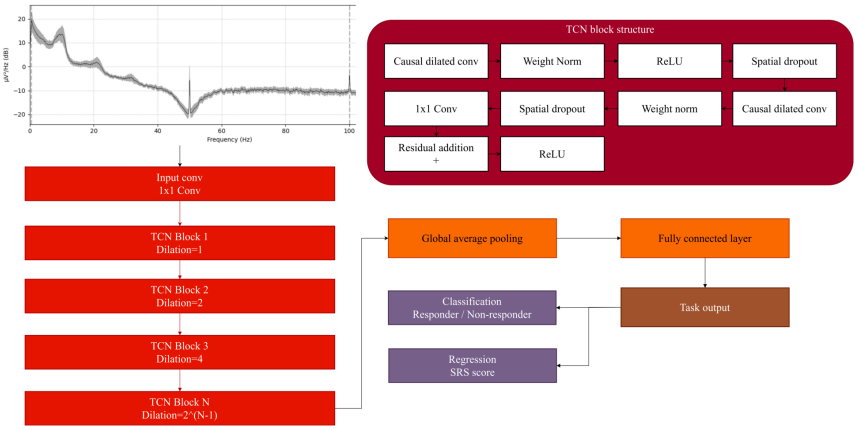

FIGURE 2. TCN model architecture

$$\hat{y}_{\text{reg}} = W_{\text{reg}} \cdot h + b_{\text{reg}} \tag{11}$$

In Formula 11, W_{reg} and b_{reg} are the weights and biases of the regression output layer.

Since the two tasks share the same backbone feature extractor, a multi-task learning paradigm is adopted. During training, a weighted summation loss function is used to jointly optimize the classification and regression objectives.

MODEL TRAINING DETAILS

The model employs a multi-task loss function, which is a weighted sum of the losses from the classification and regression tasks. The classification task uses a binary cross-entropy loss.

$$\mathcal{L}_{\text{cls}} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \cdot \log(\hat{y}_i) + (1-y_i)\cdot \log(1-\hat{y}_i)\right] \tag{12}$$

The regression task uses smoothed L1 loss, the formula of which is:

$$\mathcal{L}_{\text{reg}} = \frac{1}{N}\sum_{i=1}^{N}\begin{cases} 0.5\cdot(\hat{y}_i - y_i)^2/\beta, & |\hat{y}_i - y_i| < \beta, \\ |\hat{y}_i - y_i| - 0.5\cdot\beta, & \text{otherwise} \end{cases} \tag{13}$$

The total loss is:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{cls}}\mathcal{L}_{\text{cls}} + \lambda_{\text{reg}}\mathcal{L}_{\text{reg}} \tag{14}$$

In Formula 14, the weights λ_{cls} and λ_{reg} are adjusted to 1.0 and 0.5 based on the performance of the validation set.

AdamW optimizer, with a weight decay of 1×10^{-4}, effectively prevents overfitting. An initial learning rate of 5×10^{-4} is used, employing a cosine annealing learning rate scheduling strategy to gradually reduce the learning rate after each training epoch, helping the model converge to a better local minimum. The model is trained using mini-batch gradient descent. To obtain unbiased model performance estimates and fully utilize the limited data, hierarchical K-fold cross-validation is used, with K=5. This ensures that the ratio of responders to non-responders remains consistent with the overall dataset in each fold. The model hyperparameter settings are shown in Table 1.

TABLE 1. Model hyperparameter settings

Parameter name	Value	Notes/Explanation
Batch size	32	Balancing training efficiency and stability
Hidden layer dimension	64	Number of channels in feature maps
Convolution kernel size	3	Width of 1D convolution kernels
Number of residual blocks	4	Number of stacked TCN core blocks
Dilation coefficient sequence	[1, 2, 4, 8]	Exponential growth with number of layers
Dropout rate	0.1	Ratio of spatial dropout
Optimizer	AdamW	Adam with decoupled weight decay
Initial learning rate	5×10^{-4}	Decrement using cosine annealing strategy
Weight decay	1×10^{-3}	Used for L2 regularization
Training epochs	200	Maximum number of epochs with early stopping
Early stopping patience value	15	Wait period after validation loss shows no improvement

CONSTRUCTION OF INTERVENTION PATHWAYS BASED ON NEURAL RESPONSES

RESPONDER DEFINITION

To establish objective and quantifiable assessment criteria, this study operationally defines "treatment responder" based on the SRS scale [25-26]. This definition references improvement thresholds commonly used in clinical trials of mental disorders to measure clinical significance. In this study, a patient with autism spectrum disorder is defined as a "treatment responder" if, after completing a pre-planned music therapy program, their total SRS score decreases by 30% or more compared to their baseline assessment score. The calculation formula is:

$$\text{SRS}_{\text{change rate}} = \frac{\text{SRS}_B - \text{SRS}_T}{\text{SRS}_B} \times 100\% \tag{15}$$

In Formula 15, a patient is labeled a responder when SRS change rate$\geq$30%. This threshold ensures that individuals classified as "responders" experience symptom improvement with clear clinical significance, rather than minor fluctuations that may have no clinical value.

CONSTRUCTION AND CALCULATION OF NEURAL RESPONSE INDEX

Using a trained TCN classification model, forward propagation is performed on EEG data collected from each patient

during the first four music therapy sessions. Gradient-weighted class activation mapping is employed to analyze the interpretability of the model's decisions. This technique identifies the spatiotemporal regions of the input signals that contribute most to the classification decision (responder vs. non-responder). The two

most critical features are identified as the Gamma band power in the prefrontal cortex and the Theta band functional connectivity (wPLI) in the temporoparietal cortex.

Interpretability analysis of the trained TCN classification model was performed using gradient-weighted class activation mapping, identifying two key neural features that contribute most to distinguishing between responders and non-responders: (1) the relative power change rate ($\Delta P_{\gamma\mathrm{PFC}}$) in the Gamma band (30-45 Hz) of the prefrontal cortex (F3, F4, Fz electrodes); and (2) the weighted phase lag exponential change rate ($\Delta C_{\theta\mathrm{TP}}$) in the Theta band (4-8 Hz) of the temporoparietal cortex (represented by the CP3 and CP4 electrode pairs). The NRI was constructed strictly based on these two features.

For the prefrontal cortex Gamma power, there is:

$$\Delta P_{\gamma\mathrm{PFC}} = \frac{\bar{P}_{\gamma\mathrm{PFC}}^{\mathrm{early}} - \bar{P}_{\gamma\mathrm{PFC}}^{\mathrm{baseline}}}{\bar{P}_{\gamma\mathrm{PFC}}^{\mathrm{baseline}}} \tag{16}$$

Similarly, the change rate $\Delta C_{\theta\mathrm{TP}}$ of the Theta connection in the temporoparietal lobe is calculated. NRI is a weighted linear combination of the rates of change of key characteristics. The weights (w_1, w_2) are determined through logistic regression or linear discriminant analysis to maximize the separation between responders and non-responders. The calculation formula is:

$$\mathrm{NRI} = w_1 \cdot \Delta P_{\gamma\mathrm{PFC}} + w_2 \cdot \Delta C_{\theta\mathrm{TP}} + b \tag{17}$$

In Formula 17, b represents the bias term. The final calculated NRI is a continuous value; the higher the value, the stronger the tendency for the patient's brain neural activity to evolve towards a "responder" pattern.

Specifically, the NRI is constructed as follows: First, the EEG data collected during the first four music therapy sessions for each patient are forward-propagated using a trained TCN classification model, and key neural features are extracted using gradient-weighted class activation mapping, namely the relative power change rate of the Gamma band in the prefrontal cortex and the weighted phase lag exponent change rate of the Theta band in the temporoparietal cortex. Both are calculated according to formula (16) and similar forms, respectively. Subsequently, linear discriminant analysis is used to determine the weights and biases of each feature on the training set to maximize the separation between responders and non-responders. Finally, the NRI is calculated according to formula (17). This process is performed independently in each fold of the five-fold cross-validation to ensure that the determination of weights and thresholds does not depend on a single data partition, thereby supporting the generalization performance of the index in new samples.

ESTABLISHMENT OF DECISION-MAKING RULES

Based on the calculated NRI, a data-driven decision support system is established to provide clinicians with suggestions for adjusting the intervention pathway in the early stages of treatment (after the fourth treatment). The thresholds for the decision rules are determined by analyzing the correspondence between the distribution of NRI in the training set data and the final treatment outcomes. The intervention pathway decision rules are shown in Table 2.

The decision thresholds (θ_H, θ_L) are not determined solely based on the distribution fit of the training set, but rather through validation set performance optimization within a cross-validation framework. Specifically, in each fold cross-validation, the model is trained using the training set and the NRI is calculated. Subsequently, the response rate differences of the three intervention paths under different threshold divisions are evaluated on the validation set, and the threshold combination that maximizes the response rate gradient between paths and optimizes the overall classification performance of the validation set is selected.

TABLE 2. NRI-based intervention path decision rules

Intervention Pathway	NRI threshold range	Clinical Interpretation and Decision-Making Recommendations
Pathway A: maintenance path	$\mathrm{NRI} \geq \theta_H$	The patient's brain demonstrated a strong, positive neuroplastic response to the current music therapy program. Recommendation: Maintain the type and intensity of the current treatment program, focusing on generalization and consolidation of behavioral skills.
Pathway B: adjustment path	$\theta_L \leq \mathrm{NRI} < \theta_H$	The patient demonstrated some neural response, but the intensity was insufficient. Recommendation: Adapt the current treatment program, such as adding targeted social interaction elements, fine-tuning the music tempo to better align with her circadian rhythm, or increasing the frequency of treatment as appropriate, and reassess the NRI in 2–4 weeks.
Pathway C: early conversion pathway	$\mathrm{NRI} < \theta_L$	The patient's brain did not produce the expected neural response pattern to the current music therapy program. Recommendation: Rather than waiting for behavioral changes, initiate a clinical decision-making meeting as soon as possible and actively consider switching to other evidence-based interventions (e.g., behavioral interventions, other neuromodulation modalities) to avoid delaying the optimal intervention period.

III. EXPERIMENTAL SETUP AND EVALUATION

PARTICIPANTS AND EXPERIMENTAL DESIGN

This study recruits 80 participants aged 6-12 years. The experimental group consists of 55 children diagnosed with autism spectrum disorder (ASD) by a senior psychiatrist according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. The control group consisted of 25 age- and sex-matched children with typical development (TD). All participants have a nonverbal IQ (Intelligence Quotient) above 70 to ensure they possess the basic cognitive abilities to complete the experimental tasks. Informed consent is obtained from all participants with ASD or their guardians. All eligible participants signed an informed

consent form. The basic information of the participants is shown in Table 3.

TABLE 3. Participant basic information

Demographic Variables	ASD group (n=55)	TD group (n=25)	Statistical test value	p value
Age (years)				
Mean (SD)	8.7 (1.8)	8.9 (1.6)	$t = -0.49$	0.625
Range	6–12	6–12		
Gender (n)				
Male/Female	44 / 11	19 / 6	$\chi^2 = 0.24$	0.624
Nonverbal intelligence quotient				
Mean (SD)	98.5 (12.3)	105.2 (10.1)	$t = -2.45$	0.017
Baseline SRS total score				
Mean (SD)	82.4 (15.7)	45.2 (8.3)	$t = 11.82$	< 0.001

Note: SD (Standard Deviation).

There are no significant differences in age and sex between the two groups, meeting the matching requirements. Although the nonverbal IQ of the ASD group is significantly lower than that of the TD group, it is still within the normal range (>70). The baseline SRS total score of the ASD group is significantly higher than that of the TD group, consistent with its core symptom manifestations.

MODEL COMPARISON AND EVALUATION INDICATORS

To comprehensively and rigorously evaluate the performance of the proposed TCN model, this paper compares it with a series of representative state-of-the-art baseline models. These baseline models cover feature-engineered machine learning, classic deep learning temporal models, and state-of-the-art Transformer architectures. The baseline models compared include: SVM (Support Vector Machine), XGBoost, LSTM (Long Short-Term Memory), Bi-LSTM, CNN-LSTM, Transformer, Informer, and Pyraformer.

To ensure comprehensiveness of the evaluation, a complementary set of evaluation metrics is used for both classification and regression tasks. The formula for accuracy is:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{18}$$

The formula for precision is:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{19}$$

The formula for recall rate is:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{20}$$

The formula for the F1 score is:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{21}$$

The RMSE (Root Mean Square Error) formula is:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{22}$$

The coefficient of determination reflects the proportion of the variance of the predicted value that is explained by the actual value, and the formula is:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{23}$$

MUSIC THERAPY INTERVENTION PROGRAM

The music therapy protocol used in this study is based on an improvisational neuromusic therapy protocol, which aims to specifically modulate the neural circuits and core symptoms of autism spectrum disorder through musical experiences. The treatment protocol is administered by a registered music therapist with more than 5 years of clinical experience in ASD.

The total intervention period is 12 weeks. Individualized treatment is administered twice weekly, with each session lasting 45 minutes. Each treatment follows a structured procedure, but the content is personalized based on the patient's immediate response.

Welcome and Emotional Validation (5 minutes): a therapeutic framework is established using a fixed "greeting song", and the patient's current emotional state is assessed and validated through simple instrumental performance.

Rhythm Synchronization and Sensorimotor Integration (15 minutes): this session involves rhythmic imitation, alternation, and synchronization exercises using percussion instruments such as drums and maracas. The aim is to enhance rhythmic entrainment in the brainstem and sensorimotor cortex, improving sensory processing and multisensory integration abilities.

Musical Improvisation and Emotional Integration (15 minutes): therapist and patient engage in free improvisation using melodic instruments. Therapist follows and expands upon the patient's musical expression through techniques such as mirroring, matching, and integration, creating a non-verbal social interaction platform. This session aims to activate and exercise the "social brain" network, promoting emotional expression and interactive regulation.

Relaxation and structured closing (10 minutes): listening to soothing music, guided visualization, or humming can help patients calm their physiological arousal levels and end the session clearly with a "goodbye song", providing predictability and a sense of security.

STATISTICAL ANALYSIS METHODS

All statistical analyses are performed using SPSS 26.0 and Python (Scipy, Statsmodels library). For continuous variables such as age and nonverbal IQ that conform to a normal distribution (verified by Shapiro-Wilk test), independent samples t-tests are used to compare the ASD and TD groups.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{24}$$

In Formula 24, $\bar{X}$ is the mean; s^2 is the variance; n is the sample size.

For categorical variables such as gender, the chi-square test is used for comparison. For variables that do not conform to a normal distribution or are ordinal data, the Mann-Whitney U test is used.

To explore the relationship between NRI and behavioral improvement (ΔSRS), Pearson product-moment correlation analysis is used, with the formula as follows:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (25)$$

In Formula 25, X_i and Y_i are the observed values of NRI and ΔSRS, respectively, and $\bar{X}$ and $\bar{Y}$ are their means. The range of r is [-1, 1], and the larger its absolute value, the stronger the linear correlation.

IV. RESULTS

BEHAVIORAL RESULTS

To assess the overall effectiveness of music therapy, this study quantitatively analyzes the core behavioral symptoms of the ASD group before and after intervention, and the results are shown in Table 4.

TABLE 4. Changes in behavioral scale scores before and after intervention (Mean $\pm$ SD)

Groups	Assessment time point	n	SRS total score	ABC total score
ASD group	Baseline	55	82.4 $\pm$ 15.7	58.3 $\pm$ 20.1
	Post-treatment	55	65.1 $\pm$ 18.9	45.6 $\pm$ 17.5
	Intra-group comparison p value		< 0.001	0.003
TD group	Baseline	25	45.2 $\pm$ 8.3	22.1 $\pm$ 6.5

Behavioral data shows that after 12 weeks of music therapy intervention, the ASD group exhibits statistically significant improvement in core symptoms of social response. In particular, a large significant reduction (82.4 points at baseline, 65.1 points after treatment; $p < 0.001$) in SRS total score indicates a decrease in overall social communication difficulties among these patients. In addition, similarly defined more specific behavioral

symptoms show a large significant reduction (58.3 points to 45.6 points; $p = 0.003$) in ABC total score, representing decreased frequency and intensity of abnormal behaviors in the sensory, social, and motor domains within the group.

MODEL PERFORMANCE EVALUATION

To verify the superiority of the proposed TCN model, it is systematically compared with several state-of-the-art baseline models on a classification task. The classification performance evaluation results are shown in

The TCN model is the best at classification tasks. In Figure 3(a), the TCN model outperforms every baseline model in all four evaluation metrics of accuracy, precision, recall, and F1 score, with an accuracy, precision, recall, and F1 score of 0.861, 0.847, 0.839, and 0.843, respectively. This fully exemplifies the unique advantages of the TCN architecture when processing temporal EEG signals. Not very surprisingly, the TCN outperforms the equally-capable Transformer model and its computationally-efficient variants (Informer, Pyraformer). This is mainly because the TCN causal convolutional structure is better aligned with the physical mechanisms in which EEG signals are generated. Furthermore, the receptive field of TCNs grows exponentially with the depth of the model, allowing the TCN to more effectively capture long-range neural oscillation patterns associated with the participants' responses to the music therapy intervention. The training process of the TCN is also much more stable compared to RNN-type models, such as LSTM and Bi-LSTM.

To further quantify the accuracy of the models in predicting symptom improvement, the performance of each model is compared on the task of predicting changes in SRS scores (ΔSRS). The results are shown in Figure 4.

The accuracy of the TCN model at the individual patient level is shown in Figure 4(a). The trends seen in predicted values (red dashed line) and actual observed values (blue solid line) show agreement, with the majority of sampled cases predicting ΔSRS with small prediction error (short gray dashed lines). The TCN model is also capable of tracking the changes even for one sample where a sizeable change in ΔSRS takes place; this suggests that the TCN model performs well and robustly across patients when the degree of improvement varies. For a few samples, this moderate bias may indicate a more complex mechanism for symptom improvement, and/or confounding factors not in the model. In summary, the TCN model provides reliable predictions of efficacy at the level of the population, and at the individual patient level provides an acceptable level of accuracy.

IDENTIFICATION OF KEY NEURAL OSCILLATION PATTERNS

To gain a deeper understanding of the decision-making mechanism of the TCN model, this study uses gradient-weighted class activation mapping to visualize and analyze the key neural features on which the model's

Through inverse visualization analysis of the TCN model's decision-making process, it is found that when distinguishing between treatment responders and non-responders, the model's attention is significantly focused on the dorsal attentional network consisting of the prefrontal-central-parietal lobe. Specifically, the brain regions with the highest feature importance are concentrated in the dorsolateral prefrontal lobe, as evidenced by sustained high activation of electrodes F3, F4, and Fz, reflecting the model's emphasis on executive function and cognitive control networks.

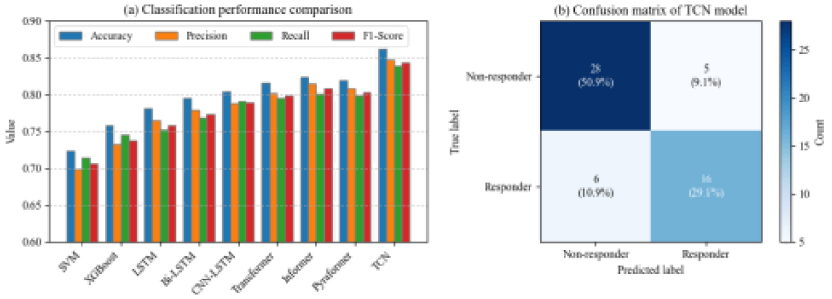

FIGURE 3. Classification performance evaluation: (a) Comparison of classification indicators; (b) Classification confusion matrix

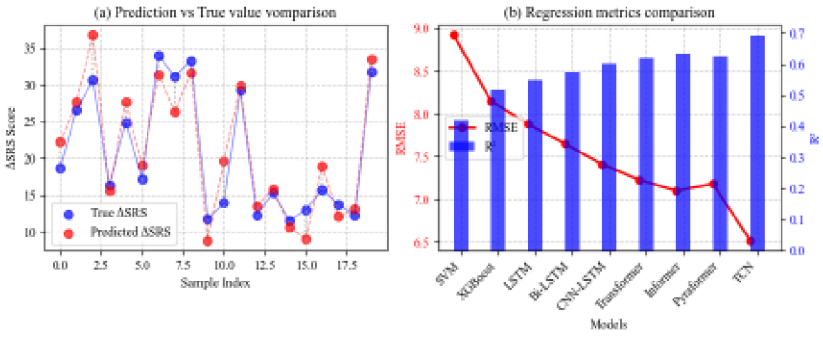

FIGURE 4. Regression performance analysis results: (a) Comparison of predicted and actual values; (b) Regression index analysis

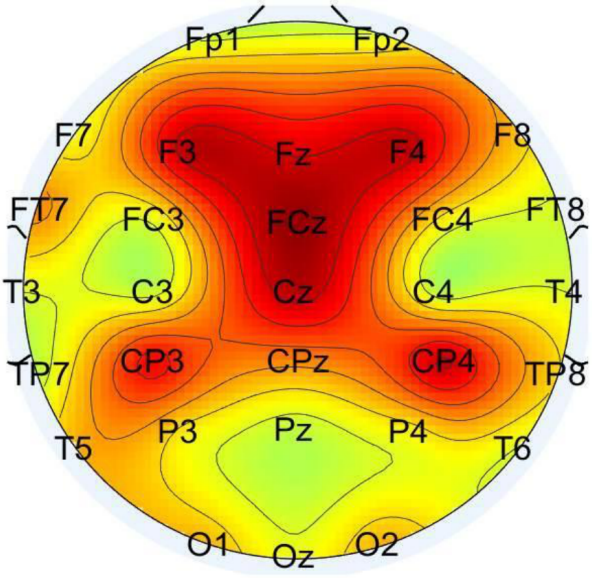

FIGURE 5. Brain topography

Simultaneously, electrodes FCz and Cz corresponding to the supplementary motor area and the anterior cingulate cortex also show high importance, indicating that neural activity related to motor planning and motivation is a key discriminative feature. Notably, electrodes CP3 and CP4 in the temporoparietal association area also show significant activation, suggesting the model's dependence on sensory integration and social cognitive neural circuits.

To systematically evaluate the neurophysiological differences between the two groups of patients, statistical analysis is performed across the entire frequency band and at the group level, based on three dimensions: relative power, PLV, and wPLI. The results are shown in Figure 6.

Inter-group comparisons reveal multi-level, frequency-specific differences in neural oscillations between responders and non-responders. In Figure 6(a), the responder group shows significantly higher relative power in the Beta and Gamma bands ($p<0.05$). Enhanced Beta rhythms are asso-

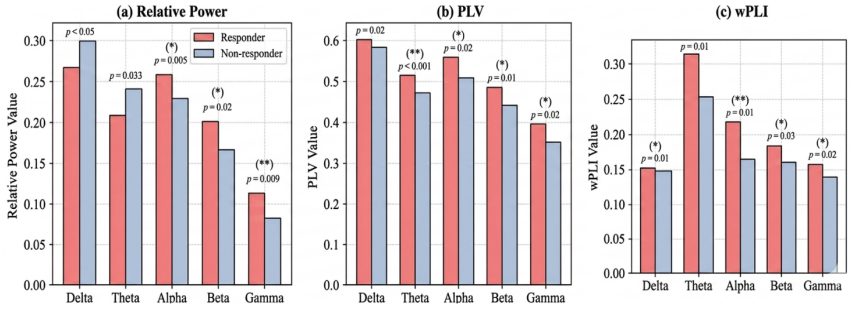

FIGURE 6. Comparison of key neural features between responders and non-responders: (a) Relative power; (b) PLV; (c) wPLI

ciated with sensorimotor integration and sustained attention, while enhanced prefrontal Gamma activity is closely related to higher-order cognitive processing, such as social information processing and cognitive flexibility, indicating that responders possess more efficient local neural network information processing capabilities. At the level of brain network connectivity, the responder group exhibits a more significant advantage in long-range functional connectivity. In terms of phase synchronicity (PLV in Figure 6(b)), responders show stronger connectivity in the Alpha and Gamma bands, reflecting more efficient coordination between resting-state networks and task-related networks. Crucially, on the wPLI index (Figure 6(c)), which is insensitive to volumetric conduction, the responder group shows significantly enhanced connectivity in the Theta and Alpha bands ($p<0.05$).

ASSOCIATION BETWEEN NEUROLOGICAL CHARACTERISTICS AND BEHAVIORAL IMPROVEMENT

To explore the quantitative relationship between TCN-based neurological indicators and clinical efficacy, Pearson correlation coefficients are calculated between NRI and key neurological features and behavioral improvement (ΔSRS). The results are shown in Table 5.

TABLE 5. Correlation results

Neural profile	Correlation coefficient (r) with ΔSRS	p value
NRI	0.712	$< 0.001^{***}$
Prefrontal Gamma Relative Power	0.583	0.002^{**}
Central Beta Relative Power	0.445	0.021^{*}
Prefrontal Alpha wPLI	0.521	0.007^{**}
Midline Theta wPLI	0.386	0.045^{*}
Whole-Brain Gamma PLV	0.402	0.037^{*}

Table 5 lists the neural features used to construct the NRI (prefrontal gamma relative power, temporoparietal Theta wPLI) and other features that showed significance in the correlation analysis (such as prefrontal alpha wPLI, midline Theta wPLI, etc.). Although the latter was not included in the NRI calculation, its correlation with ΔSRS further supports the universal importance of Theta/Alpha band functional connectivity in the neural response of music therapy and indirectly verifies the rationality of the model feature selection.

Correlation analysis strongly confirms a significant and well-defined quantitative relationship between TCN-based neural indicators and clinical behavioral improvement. The most prominent finding is the strongest positive correlation between NRI and ΔSRS ($r = 0.712$, $p < 0.001$), validating the effectiveness of NRI as a predictive biomarker for treatment efficacy and demonstrating that the neural features integrated by the TCN model robustly reflect the degree of symptom improvement. At the specific neural feature level, the prefrontal cortex Gamma relative power ($r = 0.583$) and prefrontal cortex Alpha band functional connectivity (wPLI, $r = 0.521$) show the most significant correlations with behavioral improvement.

EMPIRICAL VERIFICATION AND CASES OF INTERVENTION PATHWAYS

To verify the clinical effectiveness of the constructed intervention pathway, NRI-based decision rules are applied, and typical cases are selected to systematically demonstrate the complete process from neural response prediction to clinical decision-making. The results are shown in Figure 7.

The empirical results strongly demonstrate the effectiveness and practicality of the NRI-based intervention pathway decision-making system in real-world clinical scenarios. Figure 7 (a) shows that patients are successfully stratified into three differentiated pathways, with pathway B having the most patients (n=22), consistent with the general trend that the majority of patients in the clinical population exhibit moderate responses to treatment. The most crucial evidence comes from Figure 7 (b), where the proportion of actual responders in the three pathways shows a perfect gradient difference: the response rate of pathway A (NRI $\geq$ 0.4) is as high as 88.9%, higher than that of pathway B (45.5%) and pathway C (13.3%), proving that NRI can accurately distinguish patient groups with different prognoses in the early stages of treatment (week 4). Figure 7 (c) vividly illustrates this decision-making process: Patient #23′s NRI rises rapidly during the first four weeks of treatment, from 0.15 to 0.61, far exceeding the threshold of Pathway A (0.4), thus recommending the "maintenance path" in the fourth week. Subsequent follow-up confirms the correctness of this recommendation; the patient's SRS score significantly improves from 85 to 52 (ΔSRS=33) at the end of treatment, making them a typical treatment responder.

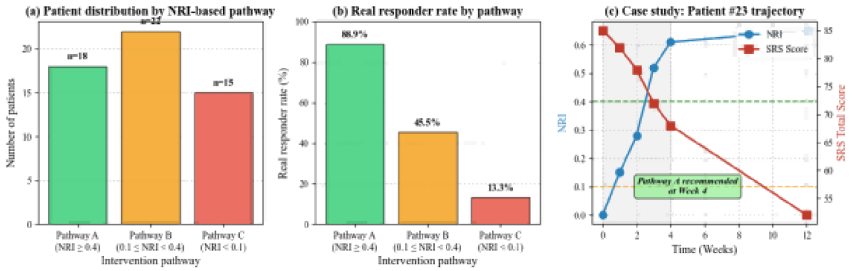

FIGURE 7. Intervention path decision effect: (a) Patient distribution by NRI-based pathway; (b) Real responder rate by pathway; (c) Case study: Patient #23 trajectory

The superiority of TCN models in classification and regression tasks stems not only from their efficient modeling capabilities for long sequences but also from their inherent compatibility with the temporal dynamics of neural oscillations. Long-range dependencies in EEG signals often manifest as cross-frequency and cross-brain region phase synchronization and power modulation. These patterns frequently exhibit cross-timescale evolution during music intervention (e.g., coupling between slow oscillations in the Theta-Alpha band and fast oscillations in the Gamma band). TCN expands the receptive field exponentially through causal convolutions, enabling it to capture neural response trajectories ranging from hundreds of milliseconds to several seconds, matching the timescale of music rhythm-induced neural plasticity. Simultaneously, residual connections ensure the stability of deep network training, allowing the model to effectively learn hierarchical temporal features—from local event-related potentials to global brain network coordinative changes. Therefore, TCN is not only a computationally efficient architecture but also a neurophysiologically interpretable tool capable of analyzing the dynamics of neural responses in music therapy, providing a new computational perspective for understanding the "brain-music" interaction mechanism.

V. CONCLUSIONS

This study systematically constructs a complete research framework from "neural signal decoding" to "clinical pathway construction". This paper applies TCN to EEG data analysis in music therapy for autism, validating its advantages over traditional and modern deep learning models in handling long-term temporal dependencies. Secondly, leveraging the interpretability of the TCN model, key features with clear neuroscientific significance are extracted from complex EEG signals, and a quantitative NRI is constructed based on this. Finally, based on the NRI, a data-driven, three-pathway intervention decision rule is established, transforming abstract neural features into concrete clinical action guidelines. This study confirms the superior performance of TCN in decoding ASD EEG temporal patterns, providing a new solution to the contradiction between long-term dependencies and computational efficiency. The innovative NRI concept realizes the transformation from multidimensional neural features to a single, actionable clinical indicator, filling the gap between neuroscience and clinical decision-making. A dynamically adjusted intervention pathway for music therapy for autism based on neural responses is established, providing empirical evidence for achieving personalized and precise intervention.

AUTHOR CONTRIBUTIONS

Conceptualization –Anshuang Zhang, Wei Sun and Boyuan Song; methodology – Anshuang Zhang, Wei Sun and Boyuan Song; investigation – Anshuang Zhang and Boyuan Song; writing-original draft preparation – Anshuang Zhang, Wei Sun and Boyuan Song. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

HUMAN RESEARCH SUBJECTS

All eligible participants signed an informed consent form.

FUNDING

This research was supported by,

1. Changzhi Medical College 2024 Doctoral Research Startup Fund Project: A Study on the Role of Art Therapy under the Healthy China Initiative - A Case Study of College Students (2024BS16)

2. This research was supported by the 2025 Shanxi Provincial Educational Science 14th Five-Year Plan Project: "Research on Educational Pathways for Enhancing Mental Health Literacy among Rural Elderly in Shanxi Province through Music Therapy Interventions" (GH-25017)

3. This research was supported by the 2025 Special Project on "Taihang Spirit and Medical Talent Cultivation in the New Era" of Changzhi Medical College: "Research on the Pathway of Integrating Music Interventions into the 'Great Ideological and Political Course' under the Guidance of the Taihang Spirit" (2025SZ15).

4. Teaching Methodology Reform Research Project of Nanjing Normal University of Special Education: Innovative Research on Practice-Oriented Teaching Models for Music Therapy Ethics (2024XJJG38). One of the achievements of the 'Key (Construction) Laboratory for Special Children's Disorders and Intervention Technologies' in Jiangsu Province (2024.1).

ACKNOWLEDGEMENTS
Not applicable.

REFERENCES

[1] S. Majhi, S. Kumar, and L. Singh, "A review on autism spectrum disorder: pathogenesis, biomarkers, pharmacological and non-pharmacological interventions," CNS & Neurological Disorders-Drug Targets, vol. 22, no. 5, pp. 659-677, 2023, doi: 10.2174/1871527321666220428134802.

[2] S. Zhou M, M. Nasir, C. Farhat L, et al., "Meta-analysis: pharmacologic treatment of restricted and repetitive behaviors in autism spectrum disorders," Journal of the American Academy of Child & Adolescent Psychiatry, vol. 60, no. 1, pp. 35-45, 2021, doi: 10.1016/j.jaac.2020.03.007.

[3] V. Marquez-Garcia A, J. Magnuson, J. Morris, et al., "Music therapy in autism spectrum disorder: A systematic review," Review Journal of Autism and Developmental Disorders, vol. 9, no. 1, pp. 91-107, 2022, doi: 10.1007/s40489-021-00246-x.

[4] L. Gassner, M. Geretsegger, and J. Mayer-Ferbas, "Effectiveness of music therapy for autism spectrum disorder, dementia, depression, insomnia and schizophrenia: update of systematic reviews," European journal of public health, vol. 32, no. 1, pp. 27-34, 2022, doi: 10.1093/eurpub/ckab042.

[5] M. Rahman M, L. Usman O, C. Muniyandi R, et al., "A review of machine learning methods of feature selection and classification for autism spectrum disorder," Brain sciences, vol. 10, no. 12, pp. 949, 2020, doi: 10.3390/brainsci10120949.

[6] A. Manoharan T and M. Radhakrishnan, "Region-Wise brain response classification of ASD children using EEG and BiLSTM RNN," Clinical EEG and neuroscience, vol. 54, no. 5, pp. 461-471, 2023, doi: 10.1177/15500594211054990.

[7] A. Ardakani H, M. Taghizadeh, and F. Shayegh, "Diagnosis of autism disorder based on deep network trained by augmented EEG signals," International journal of neural systems, vol. 32, no. 11, Art. no. 2250046, 2022, doi: 10.1142/S0129065722500460.

[8] M. Hankus, P. Ochman-Pasierbek, M. Brzozowska, et al., "Electroencephalography in autism spectrum disorder," Journal of Clinical Medicine, vol. 14, no. 6, pp. 1882, 2025, doi: 10.3390/jcm14061882.

[9] M. Melinda, D. Purnamasari P, F. Fahmi, et al., "A comprehensive EEG dataset and performance assessment for Autism Spectrum Disorder," Scientific Reports, vol. 15, no. 1, Art. no. 34981, 2025, doi: 10.1038/s41598-025-18934-7.

[10] E. Santarone M, S. Zambrano, N. Zanotta, et al., "EEG features in autism spectrum disorder: a retrospective analysis in a cohort of preschool children," Brain Sciences, vol. 13, no. 2, pp. 345, 2023, doi: 10.3390/brainsci13020345.

[11] S. Pierce, G. Kadlaskar, A. Edmondson D, et al., "Associations between sensory processing and electrophysiological and neurochemical measures in children with ASD: an EEG-MRS study," Journal of neurodevelopmental disorders, vol. 13, no. 1, pp. 5, 2021, doi: 10.1186/s11689-020-09351-0.

[12] S. Neo W, D. Foti, B. Keehn, et al., "Resting-state EEG power differences in autism spectrum disorder: a systematic review and meta-analysis," Translational Psychiatry, vol. 13, no. 1, pp. 389, 2023, doi: 10.1038/s41398-023-02681-2.

[13] V. Sysoeva O, "Neurophysiological markers that link genes and behavior in humans: examples from rare genetic syndromes associated with autism spectrum disorders," Genes & Cells, vol. 18, no. 4, pp. 297-307, 2023, doi: 10.23868/gc567774.

[14] A. Alturki F, K. AlSharabi, M. Abdurraqeeb A, et al., "EEG signal analysis for diagnosing neurological disorders using discrete wavelet transform and intelligent techniques," Sensors, vol. 20, no. 9, pp. 2505, 2020, doi: 10.3390/s20092505.

[15] A. Tawhid M N, S. Siuly, and H. Wang, "Diagnosis of autism spectrum disorder from EEG using a time-frequency spectrogram image-based approach," Electronics Letters, vol. 56, no. 25, pp. 1372-1375, 2020, doi: 10.1049/el.2020.2646.

[16] N. Alotaibi and K. Maharatna, "Classification of autism spectrum disorder from EEG-based functional brain connectivity analysis," Neural computation, vol. 33, no. 7, pp. 1914-1941, 2021, doi: 10.1162/neco_a_01394.

[17] K. Kang, S. Orlandi, J. Leung, et al., "Electroencephalographic interbrain synchronization in children with disabilities, their parents, and neurologic music therapists," European Journal of Neuroscience, vol. 58, no. 1, pp. 2367-2383, 2023, doi: 10.1111/ejn.16036.

[18] P. Garces, S. Baumeister, L. Mason, et al., "Resting state EEG power spectrum and functional connectivity in autism: a cross-sectional analysis," Molecular autism, vol. 13, no. 1, pp. 22, 2022, doi: 10.1186/s13229-022-00500-x.

[19] C. O'Reilly, S. Huberty, S. van Noordt, et al., "EEG functional connectivity in infants at elevated familial likelihood for autism spectrum disorder," Molecular Autism, vol. 14, no. 1, pp. 37, 2023, doi: 10.1186/s13229-023-00570-5.

[20] L. Huang, K. Zhou, S. Chen, et al., "Automatic detection of epilepsy from EEGs using a temporal convolutional network with a self-attention layer," BioMedical Engineering OnLine, vol. 23, no. 1, pp. 50, 2024, doi: 10.1186/s12938-024-01244-w.

[21] Y. Ding, S. Zhang, C. Tang, et al., "Masa-tcn: Multi-anchor space-aware temporal convolutional neural networks for continuous and discrete eeg emotion recognition," IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 7, pp. 3953-3964, 2024, doi: 10.1109/JBHI.2024.3392564.

[22] Y. Lim Z, H. Pang Y, Y. Ooi S, et al., "MLTCN-EEG: metric learning-based temporal convolutional network for seizure EEG classification," Neural Computing and Applications, vol. 37, no. 4, pp. 2849-2875, 2025, doi: 10.1007/s00521-024-10783-1.

[23] D. Ji, L. He, X. Dong, et al., "Epileptic seizure prediction using spatiotemporal feature fusion on EEG," International journal of neural systems, vol. 34, no. 08, Art. no. 2450041, 2024, doi: 10.1142/S0129065724500412.

[24] Q. Abbas, R. Baig A, and A. Hussain, "Classification of post-covid-19 emotions with residual-based separable convolution networks and eeg signals," Sustainability, vol. 15, no. 2, pp. 1293, 2023, doi: 10.3390/su15021293.

[25] K. Gergoudis, A. Weinberg, J. Templin, et al., "Psychometric study of the social responsiveness scale in Phelan-McDermid syndrome," Autism research, vol. 13, no. 8, pp. 1383-1396, 2020, doi: 10.1002/aur.2299.

[26] L. Bezemer M, A. Blijd-Hoogewys E M, and M. Meek-Heekelaar, "The predictive value of the AQ and the SRS-A in the diagnosis of ASD in adults in clinical practice," Journal of Autism and Developmental Disorders, vol. 51, no. 7, pp. 2402-2415, 2021, doi: 10.1007/s10803-020-04699-7.