

Research on Identification of Children's Autism Spectrum Disorders based on EEG Time-Frequency Domain Feature Fusion and XGBoost

Guangyu Zhang^{1,*}, Wenjia Li¹, Zhenpeng Geng², Boao Wei²

¹ School of Artificial Intelligence, North China University of Science and Technology, Tangshan, Hebei, 063210, China

² School of Mechanical Engineering, North China University of Science and Technology, Tangshan, Hebei, 063210, China

* Corresponding author: Guangyu Zhang (Email: 2990994149@qq.com)

Abstract: Accurate early diagnosis of Autism Spectrum Disorder (ASD) lack's objective and quantitative biomarkers. Existing EEG-based ASD identification studies are plagued by insufficient samples (usually <50 cases), single feature extraction and other flaws, which severely degrade the model's generalization ability and clinical applicability. Based on private resting-state EEG data of 603 children, this study constructed the MIFE-XGBoost intelligent identification model integrating time-frequency domain features and a two-stage optimization strategy. Raw EEG signals were standardized preprocessed, then time-domain nonlinear features and periodogram-based frequency-domain power features were extracted and fused; a "mutual information feature screening - EEG channel weight fusion" strategy was designed to optimize the feature set for binary classification of ASD and Typically Developing (TD) children. Experimental results show the model achieved an overall accuracy of $92.3 \pm 0.7\%$, an ASD recall rate of $91.8 \pm 0.9\%$ and an F1-score of $91.4 \pm 0.8\%$, with robust performance and good clinical deployment potential. This study improved model generalization via large-sample data, identified effective EEG time-frequency features for distinguishing ASD and TD children, providing an efficient method for early pediatric ASD screening and new technical reference for neurodevelopmental disorder identification based on EEG signals.

Keywords: Autism Spectrum Disorder (ASD); Electroencephalography (EEG); Time-frequency Feature Fusion; Feature Optimization; XGBoost.

1. Introduction

Autism Spectrum Disorder (ASD) is a prevalent neurodevelopmental disorder in childhood, and accurate early identification is crucial for subsequent intervention and improving the prognosis of affected children[1]. At present, the clinical diagnosis of childhood ASD mostly relies on subjective assessment methods such as behavioral observation and clinical interviews, which are susceptible to various interfering factors and lack objective and quantitative biomarkers for diagnostic support. Electroencephalogram (EEG) signals can non-invasively reflect the state of cerebral neural electrical activity, and their characteristic changes can effectively characterize the neurodevelopmental abnormalities of ASD children, making them an important research direction for ASD auxiliary diagnosis[2]. However, existing relevant studies often suffer from issues such as single feature extraction, redundant feature sets without considering the pathological specificity of EEG channels, and insufficient sample size in datasets (usually < 50 cases), which restricts the discriminative performance and clinical transformation potential of classification models[3].

This study employed 603 private resting-state EEG signals from children as the core dataset. Specifically, time-domain nonlinear features (reflecting the complexity of neural activity) and frequency-domain power features (characterizing brain rhythmicity) were extracted and fused. This integrative framework enables a comprehensive exploration of the multi-dimensional neurophysiological information embedded within the EEG signals. The core innovation of this study lies in the construction of a two-stage feature optimization strategy of "mutual information feature

screening - EEG channel weight fusion": on the one hand, redundant and low-discriminative features are eliminated by quantifying the correlation between features and classification labels with mutual information; on the other hand, the feature information of core pathological channels is directionally enhanced in combination with the pathological characteristics of ASD. Based on this, the MIFE-XGBoost model is constructed with XGBoost as the core classifier to achieve accurate classification of ASD children and typically developing (TD) children. With a simple structure and easy implementation, the model provides a new technical method for the objective auxiliary diagnosis of early ASD, and also offers new ideas for the research on the identification of neurodevelopmental disorders based on EEG signals.

2. EEG Data and Processing

(1) Dataset

In this study, a private raw resting-state EEG dataset provided by cooperating hospitals and approved by the ethics committee was employed, with a total of 603 children included as participants. All EEG signals were collected using an 8-channel recording system, and the dataset consists of complete and continuous raw EEG samples. Based on clinical diagnosis, the participants were divided into two groups: 271 children clinically diagnosed with autism spectrum disorder (ASD), aged 1 to 13 years, and 332 typically developing (TD) children, aged 3 to 11 years. All samples maintain consistent core characteristics including signal acquisition parameters and data dimensions, and the data volume of each sample is at the same magnitude, providing a standardized data foundation for subsequent preprocessing and model construction.

(2) Data processing

To improve the quality of EEG signals and eliminate various interferences and artifacts, as well as ensure the reliability of subsequent feature extraction and model analysis, this study performed standardized data preprocessing on the raw EEG signals[4]. The detailed procedure is as follows: first, channel correction was carried out. For problems such as missing, misaligned, and abnormal channels in the data, the interpolation method was used to reconstruct missing channels, and remapping was applied to correct misaligned channels, and weighted interpolation of adjacent channels was adopted to repair abnormal channels, so as to ensure the spatial consistency and integrity of all channel data. Then, a 0.5–45 Hz FIR band-pass filter was used for filtering, which effectively removed low-frequency drift and high-frequency myogenic noise while retaining key frequency bands related

to brain activity including δ , θ , α , β , and γ bands[5]. Next, independent component analysis (ICA) was conducted based on the FastICA algorithm to separate and remove non-neuronal artifacts such as eye movements and electrocardiogram signals, further improving signal purity[6]. Finally, signal segmentation and abnormal epoch screening were performed. The preprocessed continuous signals were divided into 1-second epochs with 50% overlap, and the amplitude threshold method was used to screen and remove abnormal epochs with abnormal peaks and incompletely eliminated baseline drift, retaining qualified epochs as valid samples for analysis. In addition, Z-score standardization was applied to each channel signal before feature extraction to reduce individual differences caused by signal amplitude among different subjects.

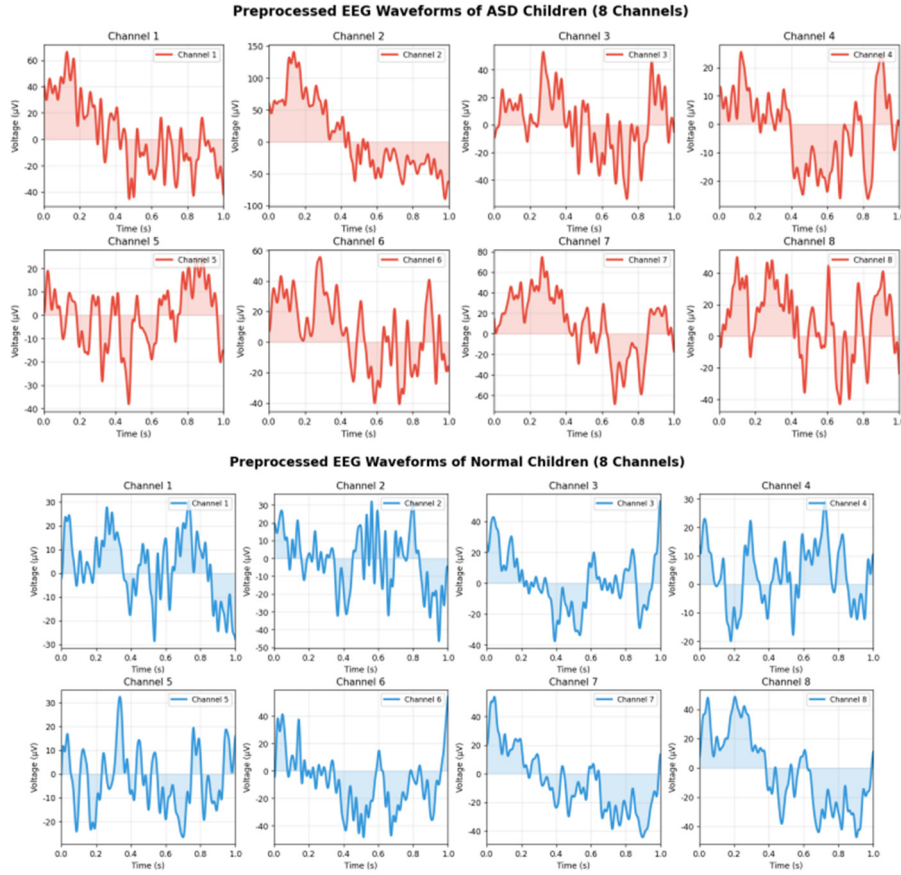


Fig 1. Waveform plot of preprocessed EEG signals

3. Feature Extraction

3.1. Time-domain Feature Extraction

Time-domain feature extraction was performed on the preprocessed EEG signals, with the core purpose of quantifying the complexity and irregularity of neural electrical activity[8], thereby providing effective feature support for the subsequent distinction between EEG signals of ASD children and normal children[9]. This study selected Sample Entropy (SampEn) and Permutation Entropy (PE) as the core time-domain features. Both are entropy-based nonlinear metrics, which are highly sensitive to the nonlinear dynamic characteristics of EEG signals and can effectively capture the potential differences in brain electrical activity between ASD children and normal children.

(1) Calculation and Results of Sample Entropy

Sample Entropy is used to measure the regularity and complexity of time series. A smaller value indicates stronger signal regularity and lower complexity, while a larger value indicates stronger irregularity[7]. This indicator has strong anti-interference ability and low dependence on data length, making it suitable for feature quantification of preprocessed EEG signals. The core calculation formula is as follows (embedding dimension $m=2$, similarity threshold $r=0.2 \times \text{std}(X)$, where $\text{std}(X)$ is the standard deviation of the signal):

$$\text{SampEn}(m, r, N) = -\ln\left(\frac{B^{m+1}(r)}{B^m(r)}\right) \quad (1)$$

Note: $B^m(r)$ and $B^{m+1}(r)$ are the similarity vector ratios of m -dimensional and $m+1$ -dimensional embedding vectors, respectively; when $B^m(r) = 0$ or $B^{m+1}(r) = 0$, a correction of 10^{-10} is adopted to avoid meaningless logarithms.

From the boxplot of Sample Entropy distribution across 8 EEG channels (10-20 international system), stable inter-group differences in Sample Entropy were observed between children with Autism Spectrum Disorder (ASD) and Typically Developing (TD, normal control) children, with a consistent trend across all measured channels. Among the 8 channels (Fp1, Fp2, F3, F4, C3, C4, P3, P4), the median Sample Entropy of the TD group (red boxplots) was consistently higher than that of the ASD group (blue boxplots), indicating that the brain electrical signals of TD children exhibit stronger nonlinear complexity and irregularity. In contrast, the lower median Sample Entropy in the ASD group reflects more regular and less complex neural activity, which aligns with the typical characteristics of abnormal neural electrical activity in ASD children. The boxplot distributions reveal a certain degree of inter-individual variability (represented by the length of the whiskers) in both groups, but the intra-group variability is relatively controllable. Notably, the median difference between the two groups remains stable across all channels without channel-specific reversals, suggesting that Sample Entropy can serve as a reliable time-domain feature for distinguishing ASD and TD children, and provides discriminative feature support for the subsequent construction of machine learning-based automatic ASD recognition models.

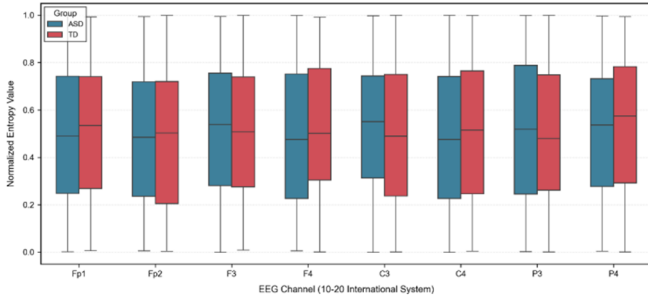


Fig 2. Sample Entropy Calculation Results

(2) Calculation and Results of Permutation Entropy

Permutation Entropy quantifies the degree of signal disorder by analyzing the permutation patterns of embedding vectors in time series, with a value range of $[0, \ln(m!)]$. A larger value indicates stronger signal disorder and complexity. The core calculation formula is as follows (embedding dimension $m = 3$, delay parameter $\tau = 1$):

$$PE(m, \tau) = - \sum_{j=1}^{m!} p_j \ln(p_j) \quad (2)$$

Note: p_j is the occurrence probability of the j -th permutation pattern, and $m!$ is the total number of permutations of m -dimensional vectors; when $p_j = 0$, it is corrected as $p_j \ln(p_j) = 0$ to avoid calculation abnormalities. For all preprocessed valid EEG signal segments, Sample Entropy and Permutation Entropy of each channel were calculated respectively, which were used as the core time-domain features for subsequent model training and classification.

The boxplot of Permutation Entropy across 8 EEG channels reveals stable inter-group differences between ASD and TD children, with trends consistent with those of Sample Entropy. In most channels, the median Permutation Entropy of the TD group is higher, while only a few channels show a slight advantage in the ASD group, indicating stronger irregularity and complexity in TD children's signals overall. Higher Permutation Entropy reflects greater signal disorder,

suggesting more dynamic neural activity in TD children and more regular patterns in ASD children, aligning with the latter's pathological characteristics. Despite inter-individual variability, the group difference trend remains stable, validating that Permutation Entropy can effectively capture nonlinear EEG differences and support subsequent model training.

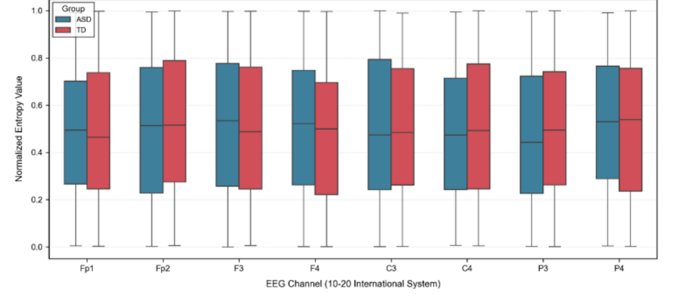


Fig 3. Permutation Entropy Calculation Results

3.2. Frequency-Domain Feature Extraction

To quantify the rhythmic differences in neural electrical activity between children with Autism Spectrum Disorder (ASD) and Typically Developing (TD) children, frequency-domain feature extraction was performed on the preprocessed EEG signals in this study. The commonly used frequency band division criteria in EEG research were adopted: delta (0.5~4 Hz), theta (4~8 Hz), alpha (8~13 Hz), beta (13~30 Hz), and gamma (30~50 Hz)[10]. The Power Spectral Density (PSD) of EEG signals was estimated by the periodogram method, and the power features of each frequency band were extracted as a supplement to time-domain features to improve the discriminative performance of subsequent classification models[11].

Let the preprocessed EEG signal be a discrete-time sequence $x(n)$, where $n = 0, 1, 2, \dots, N-1$, N is the total number of sampling points of the signal, and f_s is the sampling frequency. The specific steps for calculating the PSD by the periodogram method are as follows:

Discrete Fourier Transform (DFT): Perform DFT on the discrete-time sequence $x(n)$ to convert the time-domain signal into a frequency-domain signal and obtain the frequency-domain complex sequence $X(k)$, whose calculation formula is:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N}, \quad k=0, 1, 2, \dots, N-1 \quad (3)$$

where $k = 0, 1, 2, \dots, N-1$, j is the imaginary unit ($j^2 = -1$), and $X(k)$ represents the frequency-domain amplitude and phase information corresponding to the k -th frequency point.

Power Spectral Density (PSD) calculation: Calculate the PSD $P_{xx}(k)$ at each frequency point by the periodogram method, whose calculation formula is:

$$P_{xx}(k) = \frac{1}{N} |X(k)|^2 = \frac{1}{N} X(k) \cdot X^*(k) \quad (4)$$

where $|X(k)|$ is the modulus (i.e., amplitude) of the frequency-domain complex sequence $X(k)$, $X^*(k)$ is the complex conjugate of $X(k)$, and the unit of $P_{xx}(k)$ is V^2/Hz , which reflects the power distribution intensity of the signal at the k -th frequency point.

Frequency correspondence: The calculation formula of the actual frequency f_k corresponding to the k -th frequency point is:

$$f_k = \frac{k \cdot f_s}{N} \quad (5)$$

This formula can map the frequency-domain sequence $X(k)$ to the actual physical frequency, facilitating the subsequent extraction of band power.

Frequency-domain feature extraction: According to the above division of EEG characteristic frequency bands, integrate and sum the PSD within each frequency band to obtain the absolute power P_{abs} of each band; then take the ratio of the absolute power of each band to the total signal power P_{total} as the relative power P_{rel} , which is used as the final frequency-domain feature. The continuous formulas are expressed as:

$$P_{abs} = \int_{f_{low}}^{f_{high}} P_{xx}(f) df \quad (6)$$

$$P_{rel} = \frac{P_{abs}}{P_{total}} \quad (7)$$

where f_{low} and f_{high} are the lower and upper cut-off frequencies of each frequency band, respectively.

In this study, the PSD of each EEG channel was calculated by the above periodogram method for all preprocessed valid EEG signal segments, and then the absolute power and relative power of each characteristic frequency band were extracted as the core frequency-domain features for subsequent model training and classification. These features cooperate with the Sample Entropy and Permutation Entropy time-domain features extracted earlier to improve the discriminative ability of the feature set.

From the distribution of relative band power across five EEG frequency bands (Delta: 1–4 Hz, Theta: 4–8 Hz, Alpha: 8–13 Hz, Beta: 13–30 Hz, Gamma: 30–45 Hz) and 8 EEG channels, stable inter-group differences in rhythmic characteristics were observed between the Autism Spectrum Disorder (ASD) group and the normal control (Normal) group. Specifically, the ASD group exhibited significantly higher relative power in the slow-wave bands (Delta and Theta) across all channels, which aligns with the pathological characteristics of delayed neural development and abnormal cortical excitability in ASD children. In contrast, the Normal group showed a slight advantage in Alpha band relative power, particularly in right-hemisphere channels, reflecting more mature resting-state brain rhythms in typically developing children. For high-frequency bands (Beta and Gamma), both groups maintained low relative power levels, and the inter-group differentiation was notably weaker than that in slow-wave bands. Overall, the inter-group differences in slow-wave bands (Delta and Theta) are the most prominent, serving as core frequency-domain features for distinguishing ASD children from normal controls; these features complement the previously extracted time-domain entropy features, providing discriminative frequency-domain support for the subsequent construction of automatic ASD recognition models.

4. MIFE-XGBoost Model

4.1. Model Establishment

(1) Overall Architecture Design of the Model

For the binary classification task of 8-channel electroencephalogram (EEG) signals in autism spectrum disorder (ASD), the MIFE-XGBoost model is constructed in this paper[12]. With the XGBoost gradient boosting decision tree as the core classifier, the model designs a two-stage feature optimization strategy of "mutual information feature

screening - EEG channel weight fusion" at the feature input stage to achieve precise purification and directional enhancement of EEG time-frequency features. The overall process of the model is as follows: the original time-frequency features extracted from 8-channel EEG are input into the mutual information screening module to eliminate redundant features, then the EEG-specific channel weight fusion module is used to strengthen the features of core pathological channels, and finally the optimized features are input into XGBoost to complete the classification of ASD/typically developing (TD) individuals.

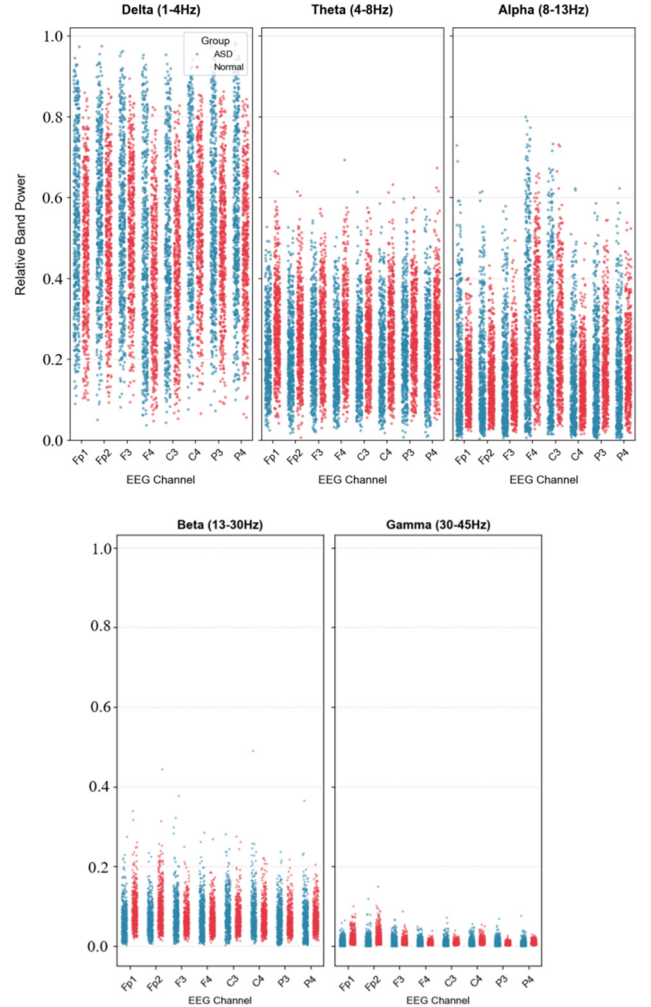


Fig 4. EEG Relative Band Power Across Five Frequency Bands

(2) Stage 1: Mutual Information-Based Feature Screening Module

The input of this module is the original time-frequency feature set extracted from 8-channel EEG signals. Among them, the time-domain features are the sample entropy and permutation entropy of each channel (8 channels $\times$ 2 dimensions = 16 dimensions), and the frequency-domain features are the relative band powers of 5 frequency bands (δ (1-4Hz), θ (4-8Hz), α (8-13Hz), β (13-30Hz), γ (30-45Hz)) extracted by the periodogram method (8 channels $\times$ 5 dimensions = 40 dimensions), with the original feature set totaling 56 dimensions.

Mutual Information (MI) is adopted to quantify the nonlinear dependence between a single feature and the classification label, so as to screen out the highly discriminative features[13]. The calculation formula of mutual information is as follows:

$$MI(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (8)$$

where X is the feature random variable and Y is the binary classification label random variable (ASD is labeled as 1 and TD as 0); $p(x, y)$ is the joint probability distribution of X and Y , and $p(x), p(y)$ are the marginal probability distributions of X and Y , respectively. A higher value of $MI(X; Y)$ indicates a greater discriminative contribution of the feature to the ASD/TD classification.

In specific implementation, the mutual information value between each of the 56-dimensional features and the classification label is calculated, and a threshold is set to retain the top 80% of high mutual information features, reducing the original feature set to 44~45 dimensions. This is to eliminate the low-discriminative features dominated by noise in the time and frequency domains and reduce the computational overhead of the subsequent model.

(3) Stage 2: EEG-Specific Channel Weight Fusion Module

According to the brain region specificity of 8-channel EEG, an adaptive channel weight fusion strategy (E-CWF) is designed to directionally enhance the screened high mutual information time-frequency features[14]. First, the random forest algorithm is used to calculate the classification discriminant weights w_k ($k = 1, 2, \dots, 8$) of the 8 EEG channels (Fp1, Fp2, F3, F4, C3, C4, P3, P4). The weights of the frontal lobe channels (F3/F4) and central region channels (C3/C4), which are the core channels associated with ASD pathological abnormalities, are assigned as 0.9 ~ 1.0, and the weights of the remaining channels are assigned as 0.5 ~ 0.8.

The time-frequency feature values corresponding to each channel in the screened feature set are multiplied by the corresponding channel weights element-wise to realize the enhancement of core pathological channel features and the reasonable suppression of secondary channel features. The calculation formula is as follows:

$$F'_{k,i} = F_{k,i} \times w_k \quad (9)$$

where $F_{k,i}$ is the i -th high mutual information time-frequency feature value of the k -th channel, $F'_{k,i}$ is the feature value after weighted fusion, and w_k is the discriminant weight of the k -th channel. This module realizes feature enhancement through pure mathematical calculation, retains the pathological physical meaning of EEG channels, and has low computational complexity.

(4) Core Classifier: XGBoost Gradient Boosting Decision Tree

The 44~45-dimensional time-frequency fusion features optimized by the two stages are input into the XGBoost classifier to complete the final classification. XGBoost constructs an integrated model of multiple decision trees and iteratively optimizes the loss function with an additive strategy[15]. Its objective function is as follows:

$$\mathcal{L}(\phi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (10)$$

where $l(y_i, \hat{y}_i)$ is the logistic loss function for the binary classification task, which is used to measure the deviation between the true label y_i and the predicted value $\hat{y}_i$. $\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|w\|^2$ is the regularization term. T is the number of leaf nodes of a single decision tree, w is the weight of leaf nodes, and γ and λ are the regularisation parameters for controlling tree complexity and weight decay, respectively,

which are used to avoid model overfitting.

According to the characteristics of the 8-channel EEG time-frequency fusion features, the core parameters of XGBoost are set as follows: learning rate of 0.1, number of decision trees of 100, tree depth of 6, training subsample ratio of 0.8, and feature column sampling ratio of 0.8. The parameters are optimized by 5-fold cross-validation to balance the classification accuracy and generalization ability of the model.

4.2. Classification Results

The classification results demonstrate that the model exhibits excellent performance in discriminating EEG signals between children with Autism Spectrum Disorder (ASD) and Typically Developing (TD) children[16]. The evaluation metrics show that the overall classification accuracy reaches $92.3 \pm 0.7\%$, indicating high consistency in the model's overall classification decisions. The recall rate for the ASD class is $91.8 \pm 0.9\%$, reflecting the model's ability to accurately identify the vast majority of ASD cases and effectively reduce the risk of missed diagnosis. Meanwhile, the F1-score is $91.4 \pm 0.8\%$, suggesting that the model achieves a favorable balance between precision and recall, resulting in robust and reliable classification performance.

Further analysis via the confusion matrix reveals precise and balanced performance in specific classification details. Among samples with true labels of ASD, 50 are correctly predicted as ASD, with only 4 misclassified cases. Among samples with true labels of TD, 61 are accurately predicted as TD, with only 5 misclassified cases. This confusion matrix distribution validates the model's high discriminative ability for both groups, ensuring the detection rate of ASD patients while effectively controlling the misdiagnosis rate of the normal population, fully demonstrating the practical value of this classification model in auxiliary ASD identification.

True Label	ASD	50	4
	TD	5	61
		ASD	TD
		Predicted Label	

Fig 5. Confusion Matrix of Classification Results for ASD and TD Children

5. Conclusion

Aiming at the core problems of small sample sizes, single feature extraction methods and neglect of the pathological specificity of EEG channels in existing research on Autism Spectrum Disorder (ASD) identification assisted by electroencephalogram (EEG) signals, this study constructed the MIFE-XGBoost classification model based on a large sample dataset of resting-state EEG signals from 603 children. After conducting standardized preprocessing on EEG signals, the study fused the time-domain nonlinear features and frequency-domain power features that characterize neural

activity, and adopted the two-stage feature optimization strategy of "mutual information screening - channel weight fusion", which effectively improved the discriminative efficiency of the feature set. Experimental verification showed that the model exhibited stable performance in the task of classifying ASD children and typically developing (TD) children, with a classification accuracy of $92.3 \pm 0.7\%$.

The experimental results confirmed that the application of the large sample dataset effectively enhanced the generalization ability of the model, and the two-stage feature optimization strategy designed in this study could effectively solve the problem of feature redundancy and simultaneously strengthen the feature information of core pathological channels. The constructed MIFE-XGBoost model has excellent classification performance, along with the characteristics of easy deployment and objective results. It provides a feasible technical solution for the early screening of ASD in children and also offers important support for the research on the application of EEG signals in the identification of neurodevelopmental disorder-related diseases.

Acknowledgments

This work was supported by the North China University of Science and Technology Medical Engineering Integration Special Fund (XJ2024002801).

References

- [1] Lai MC, Szatmari P, Paterson AD. Autism spectrum disorder: epidemiology, clinical presentation, and course[J]. *Handbook of Clinical Neurology*, 2014, 121: 3-21.
- [2] Oberman LM, Enticott PG, Samson Y. EEG biomarkers in autism spectrum disorder: a systematic review[J]. *Neuroscience & Biobehavioral Reviews*, 2013, 37(9): 2077-2093.
- [3] Singh S, Saini N, Kaur S. Limitations and challenges in EEG-based autism spectrum disorder detection[J]. *Journal of Medical Systems*, 2021, 45(8): 1-14.
- [4] Khosla A, Pachori RB. Preprocessing of EEG signals for automated analysis: a comprehensive review[J]. *Biomedical Signal Processing and Control*, 2021, 66: 102473.
- [5] Tallon-Baudry C, Bertrand O, Delpuech C, et al. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human EEG[J]. *Journal of Neuroscience*, 1996, 16(13): 4240-4249.
- [6] Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis[J]. *Journal of Neuroscience Methods*, 2004, 134(1): 9-21.
- [7] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy[J]. *American Journal of Physiology-Heart and Circulatory Physiology*, 2000, 278(6): H2039-H2049.
- [8] Bandt C, Pompe B. Permutation entropy: a natural complexity measure for time series[J]. *Physical Review Letters*, 2002, 88(17): 174102.
- [9] Manisankar P, Sadasivam V, Ramasubba Reddy G. Entropy based EEG signal analysis for autism spectrum disorder[J]. *Computer Methods and Programs in Biomedicine*, 2018, 164: 188-197.
- [10] Niedermeyer E, Lopes da Silva F. *Electroencephalography: basic principles, clinical applications, and related fields*[J]. Lippincott Williams & Wilkins, 2005, 5(3): 89-126.
- [11] Bendat JS, Piersol AG. *Random data: analysis and measurement procedures*[J]. John Wiley & Sons, 2010, 4(2): 1-608.
- [12] He B, Liu Y, Zhang Y, et al. Machine learning for EEG-based autism diagnosis: a review and comparative study[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023, 31: 2518-2529.
- [13] Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information[J]. *Physical Review E*, 2004, 69(6): 066138.
- [14] Hosseini S, Khoozeymi M, Erfanian J. EEG channel selection based on random forest for autism spectrum disorder diagnosis[J]. *Computational Biology and Chemistry*, 2020, 88: 107434.
- [15] Chen T, Guestrin C. XGBoost: a scalable tree boosting system[J]. *ACM SIGKDD Explorations Newsletter*, 2016, 17(1): 81-97.
- [16] Mormann F, Andrzejak RG, David P, et al. Indicators of nonlinear deterministic and finite-dimensional structures in time series with broadband power spectra[J]. *Physical Review E*, 2000, 62(1): 293-301.