Journal of Engineering Science and Technology Review 18 (1) (2025) 227-234

Research Article

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

GraphRAG-based Sleep Staging Method Using Multi-Strategy Adaptive Reinforcement Reward Learning and Multilevel Distillation Pruning

Wei Zhao[1,*] and Rongsheng Zhao[2]

[1]*Hebei Minzu Normal University, Hebei 067000, China*
[2]*PetroChina Company Limited, Jilin 138000, China*

Received 11 November 2024; Accepted 23 February 2025

Abstract

Traditional deep learning methods have been widely used in automatic feature extraction and sleep discrimination. However, these methods need excessive computing resources in time and space, which severely limits their direct deployment and application in resource-constrained edge devices. In order to significantly reduce the size and complexity of the model, this study proposed an innovative sleep discrimination method-based graph retrieval-augmented generation (GraphRAG) using multi-strategy adaptive reinforcement reward learning and multilevel distillation pruning (i.e., RDLG). First, sleep-related data were constructed into a graph structure with GraphRAG technology based on prior knowledge or automatic learning between different data signals. Second, the adaptive strategy of reinforcement reward learning was used to approximate the true value. Third, the importance of model nodes or edges was determined, and a multilevel distillation pruning operation was performed on the sleep model. Lastly, detailed experimental tests were conducted on the polysomnography data set to verify the effectiveness of the RDLG model. Results demonstrate that, (1) the model shows significant performance advantages in sleep discrimination tasks. For the prediction results of the experiment for the group aged 20–50, the accuracy of this model is about 0.8, which is higher than that of other models at 0.6. (2) This model can effectively reduce complexity. The parameter of the model is in the order of 10^3, which is much lower than those of other models. (3) The model's adaptive multi-strategy reinforcement reward learning converges faster. This study is of practical significance for improving the accuracy of sleep recognition, greatly reducing model size, computational complexity, and resource consumption, and promoting the widespread application of sleep recognition technology on edge devices.

Keywords: RDLG, Sleep, Distillation, Reinforcement learning

1. Introduction

As a common health problem worldwide, sleep disorders cover a wide range of fields, including insomnia, hypersomnia, and sleep apnea syndrome. These disorders not only severely affect patients' quality of life but may also trigger a range of other health problems, such as cardiovascular disease, reduced immune function, and memory loss. Therefore, accurate diagnosis of sleep disorders is particularly critical. With the advancement of medical technology and bioinformatics, more studies have been focusing on how to use advanced algorithms and technologies to improve the diagnostic accuracy of sleep disorders. Traditional sleep identification methods are mainly based on polysomnography (PSG) data and use artificial feature extraction and classifiers (e.g., decision trees and SVM) to identify sleep states. Silva et al. [1] developed a model based on recurrent neural network (RNN) to extract sleep-related features by combining data from accelerometers and photoplethysmography sensors, achieving real-time monitoring of healthy people and sleep apnea patients.

However, with the development of sleep disorder identification technology, sleep models have been evolving toward multimodality and low complexity, and the accuracy and computational efficiency have increased. While the performance in accurate discrimination of sleep models is improving, their computing resource consumption is also increasing. Deep learning methods are often used in the design process. Their large number of parameters and graph structure data design factors make the model highly complex, which brings huge challenges to the study of sleep discrimination models.

Scholars have conducted numerous studies of sleep discrimination methods [2-4]. However, problems such as high model complexity and low robustness remain. Therefore, how to achieve accurate feature extraction and model optimization, improve the interpretability and generalization ability of the model while reducing the burden on computing resources, and achieve efficient and robust sleep discrimination is a key issue that needs to be solved urgently.

Therefore, this study proposed an innovative sleep discrimination method-based graph retrieval-augmented generation (GraphRAG) using multi-strategy adaptive reinforcement reward learning and multilevel distillation pruning (i.e., RDLG). This method aims to improve the accuracy and efficiency of sleep discrimination tasks. Multilevel distillation pruning technology is used to reduce model redundancy and improve generalization capabilities by constructing the RDLG model structure, combined with the multi-strategy adaptive reinforcement reward learning mechanism to optimize the model decision-making process.

*E-mail address: she-shiv@163.com
ISSN: 1791-2377
doi:10.25103/jestr.181.23

2. State of the art

Traditional sleep discrimination methods are mainly based on PSG data, which use deep learning to extract artificial features and identify sleep states. However, the excessive model parameters and complex training processes cause challenges in sleep model prediction. Signal preprocessing and feature extraction are key steps in traditional sleep discrimination methods. The half-wave method piecewise linear data reduction technique proposed by Yash [5] reduced signal complexity by simplifying the electroencephalogram (EEG) signal into a piecewise linear form while retaining key features of sleep stages. However, this method lost some subtle but important signal features, especially when processing high-frequency or low-amplitude signals, which led to incomplete feature information. In addition, this method had shortcomings in processing graph-structured data and was difficult to be directly applied to resource-constrained edge devices. Erdem Tuncer et al. presented a wavelet transform and feature extraction algorithm. In the feature extraction stage, dynamic time warping and medi an frequency features were obtained from electrocardiogram (ECG) data through wavelet transform. This method could effectively extract the frequency and time characteristics of the signal, but the basis function and parameter selection of the wavelet transform were complex. Moreover, the feature extraction process was time consuming, affecting the real-time performance of the algorithm. Xiao Shuyuan presented an improved K-means clustering algorithm [7]. This algorithm aimed to solve the problem that the original K-means clustering algorithm was sensitive to the initial cluster center and outliers. It combined the density idea to optimize the selection of the initial center and updated the center through the “3σ rule.” This method performed well in feature extraction and clustering stages and could effectively improve the accuracy and robustness of sleep staging. However, K-means clustering had strong assumptions about the data distribution shape and could not work well for nonspherical distributed data. The convergence speed and results of the algorithm also depended on the choice of initial parameters. Liu Zhiyong [8] designed a sequence connectivity analysis feature parameter extraction algorithm. Feature parameters such as the slope of the connectivity distribution and the mean value of the connectivity distance were extracted, and the least squares method was used for training and learning. From the perspective of signal connectivity, this algorithm provided a new viewpoint for feature extraction of EEG signals and was suitable for feature analysis of complex signals. However, this algorithm extracted considerable feature parameters, which led to feature redundancy and increased computational complexity, and the least squares method could not be flexible enough when processing nonlinear data, affecting the generalization ability of the model.

In recent years, with the advancement of artificial intelligence technology, scholars have tried to identify sleep on the basis of deep learning algorithms. Garcia-Vicente developed the Sleep ECG-Net interpretable deep learning method [9], which combined CNN and RNN to train ECG signals to directly assess obstructive sleep apnea (OSA) severity in high-risk children. Although this method could handle time series data, the interpretability of deep learning models was still a challenge, the training of the model required a large amount of annotated data, and the training process could be time consuming. Jiménez-García et al. [10] constructed a CNN-based deep learning architecture. AF and SpO_2 signals were analyzed using CNN to assess the severity of OSA in children. This algorithm simplified the diagnosis process through deep learning and improved the accuracy and efficiency of diagnosis. However, CNN had strict requirements on the size and format of input data, the model’s generalization ability was limited by data diversity, and its adaptability to small-sample data sets was poor. Shao Hengyi [11] presented an unsupervised domain adaptation algorithm that integrated class rebalancing and semi-supervised learning. A deep learning automatic sleep staging algorithm based on electroencephalography introduced a balanced loss function to alleviate the data imbalance problem. This algorithm improved the performance of the model on imbalanced data sets through semi-supervised learning and class-rebalancing strategies. However, the performance of semi-supervised learning was highly dependent on the quality of a small amount of annotated data, and the class-rebalancing strategy could not work well on some extreme imbalanced data sets. Tian Yunzhi [12] established a stochastic deep residual network (TL-SDResNet) based on transfer learning, which used single-channel EEG signals, transfer learning, and stochastic deep residual networks for sleep staging, combined with Butterworth filtering and continuous wavelet transform for preprocessing. This algorithm improved the generalization ability of the model through transfer learning, but the performance of transfer learning depended on the similarity between the source and target domains. The training process of the random deep residual network was also relatively complex and could lead to overfitting. Jin Zheng designed a hybrid attention sequential network [13], which replaced the traditional CNN with RNN. The intra-segment temporal attention and channel attention mechanisms were combined to achieve the fusion of intra-segment and channel correlation features of the signal. This method could effectively extract the temporal features of PSG, but it had high complexity, long training and inference time, and limited interpretability of the attention mechanism, making it difficult to intuitively understand the decision-making process of the model. Silva proposed an RNN-based model [1], which extracted sleep-input features from accelerometer and photoplethysmography sensor data. The model was used for comparative monitoring of healthy people and sleep apnea people to predict sleep stages in intervals. Although this model could handle time series data, the training process of RNNs was complicated and prone to overfitting.

In addition, scholars have adopted algorithms based on machine learning. Zeinab et al. [14] suggested the MLP algorithm, which used a multilayer perception neural network combined with a back propagation algorithm to classify sleep apnea. This method improved the accuracy of the network based on optimization algorithms, but it was prone to falling into local optima and sensitive to the selection of hyper parameters and required high computational costs. Uddin et al. developed an airflow (AF) signal peak amplitude encoding algorithm [15]. Apnea events were detected by encoding each sample of the peak amplitude of the AF signal. This algorithm used a simple signal processing method to achieve efficient event detection, but it relied on the peak amplitude of the signal and was sensitive to noise interference. Furthermore, the encoding process might lose some important timing information, affecting the diagnosis accuracy. Ye established an XGBoost diagnostic model [16]. The OSA diagnostic model established based on the XGBoost algorithm verified the performance of the model through various classification ability evaluation indicators. This method took advantage of the efficiency and scalability of XGBoost to provide a reliable machine learning solution for OSA diagnosis. However, XGBoost had high requirements for data preprocessing and was sensitive to outliers, which could affect the robustness of the model. Sheta established an efficient classification framework [17]. The performance of various machine learning classifiers was evaluated using feature selection schemes based on metaheuristic algorithms and fixed and adaptive learning methods to identify the most suitable classifier for the collected data. This method improved the accuracy of the model by optimizing feature selection and classifier performance, but it had high computational complexity and required excessive data preprocessing. Wang Qi proposed a semi-supervised sleep staging algorithm [18]. On the basis of EEG signals, an improved convolutional encoder–decoder and a GAN were used to construct a shallow feature extraction network. With the hard swish activation function, model convergence was accelerated, and a weighted cross-entropy loss function was used to improve classification accuracy. This algorithm provided an efficient and accurate solution for sleep staging through semi-supervised learning and the improved network architecture. However, the performance of semi-supervised learning was highly dependent on the quality of a small amount of annotated data, and the weighted cross-entropy loss function could not work well on some extremely imbalanced data sets.

The above studies have made significant progress in the field of sleep discrimination, but the existing methods still have shortcomings in many aspects. First, although algorithms based on signal preprocessing and feature extraction can effectively reduce the amount of data, they are prone to losing key signal features and have limitations in processing graph-structured data and resource-limited edge devices. Second, although algorithms based on deep learning improve the accuracy and efficiency of diagnosis, the model has poor interpretability, complex training process, and high demand for annotated data, making it difficult to adapt to small-sample data sets. In addition, although machine learning-based algorithms optimize classification performance, they are prone to falling into local optima, sensitive to hyper parameters, and not robust enough to noise and outliers. The proposed GraphRAG-based sleep discrimination method using multi-strategy adaptive reinforcement reward learning and multilevel distillation pruning effectively solves the shortcomings of existing methods by optimizing feature extraction, model training, and model structure. This method not only improves the interpretability and generalization ability of the model but also reduces the spatiotemporal overhead of computing resources, making it highly applicable to resource-constrained edge devices. Moreover, through graph structure data processing optimization, the RDLG model can efficiently handle complex data and improve the adaptability and robustness of the model. These improvements provide new ideas and methods for accurate diagnosis of sleep disorders.

The remainder of this study is structured as follows. Section 3 presents the GraphRAG sleep discrimination method based on multi-strategy adaptive reinforcement reward learning and multilevel distillation pruning (i.e., RDLG). Section 4 validates the effectiveness of RDLG technology and conducts detailed experimental tests on different models and data sets. Section 5 summarizes this study. RDLG shows obvious advantages in performance optimization and recognition accuracy. It can significantly reduce model complexity and the amount of calculation parameters while ensuring high recognition accuracy.

3. Methodology

3.1 Model architecture

(1) RDLG graph construction

Sleep-related data are constructed into a graph structure

$$G=(V,E) \tag{1}$$

Where the node set V contains multiple types of nodes, such as EEG signal nodes, ECG signal nodes, electromyogram (EMG) signal nodes, and nodes representing sleep stages in different frequency bands. The edge set E represents the relationships between nodes. These relationships can be automatically learned on the basis of prior knowledge, such as the correlation between physiological signals and the connection between different signals and sleep stages, or data-driven methods.

(2) Multi-strategy adaptive reinforcement reward learning module

State space: State space S , $s \in S$, represents the internal state of the model at a certain moment in the process of identifying the sleep state. It includes the current feature representation of the graph structure data and processed node information.

Action space: Action space A contains the operations that the model can take on the graph, such as selecting the next node to be processed, adjusting the connection weights between nodes, and updating node characteristics.

Reward function: The reward function $R(s,a)$ is defined in accordance with the effect of the model's action a on the sleep discrimination result in the state. If the action a can improve the accuracy of sleep discrimination (e.g., make the sleep stage predicted closer to the true value), a positive reward will be given; otherwise, a negative reward will be given. A simple form of the reward function is given below:

$$R(s,a)=\begin{cases} r_{pos}, & \text{if accuracy}_{\text{new}} > \text{accuracy}_{\text{old}} \\ r_{neg}, & \text{othewise} \end{cases} \tag{2}$$

Where r_{pos} and r_{neg} are the values of positive and negative rewards, respectively. $\text{accuracy}_{\text{new}}$ and $\text{accuracy}_{\text{old}}$ are the sleep discrimination accuracy before and after executing the action a , respectively.

(3) Adaptive strategy

The adaptive strategy selects actions a on the basis of the current state s . This strategy adopts a multi-strategy hybrid approach, such as combining a value function-based strategy and a policy gradient-based strategy. The value function-based strategy selects actions on the basis of the state-value function and aims to maximize long-term rewards. The gradient-based strategy directly optimizes the policy function $\pi(s)$. During the training process, the adaptive strategy is continuously adjusted to adapt to different sleep data distributions. Value function update: The update formula of the value function adopts the time difference learning method, i.e.,

$$V(s) \leftarrow V(s) + \alpha[R(s,a) + \gamma V(s') - V(s)] \quad (3)$$

Where α is the learning rate, γ is the discount factor, and s' is the next state after executing the action a.

(4) Multilevel distillation pruning module

Teacher–student model structure: A teacher model and a student model are constructed. The teacher model is a relatively complex RDLG model with high accuracy but possible redundancy. The student model is a simplified RDLG model designed to improve performance and reduce model complexity by distilling the knowledge of the teacher model.

Distillation loss function: The distillation loss function is defined as

$$L_{distill} = KL(p_{teacher}(y \mid x), p_{student}(y \mid x)) \quad (4)$$

Where KL is the KL divergence, $p_{teacher}(y \mid x)$ is the probability distribution of the teacher model's input x - predicted output y, and $p_{student}$ is the probability distribution of the student model for input x. Pruning operation: Pruning operations are performed during the distillation process. In accordance with the importance of nodes or edges to the sleep discrimination results, the nodes or edges with low importance are removed. The measurement of importance can be based on methods such as gradient information or information gain. For example, for node v, its importance $I(v)$ can be measured by calculating the gradient of the node features relative to the change in the discriminant loss, i.e.,

$$I(v) = \left\| \frac{\partial L}{\partial \text{feature}(v)} \right\| \quad (5)$$

Where L is the sleep discrimination loss, and $\text{feature}(v)$ is the feature vector of node v.

3.2 Model training

(1) Pretrained teacher model

The teacher model is pretrained on a large-scale sleep data set. The supervised learning method is used during the pretraining process to minimize the cross-entropy loss between the predicted and real sleep stages, i.e.,

$$L_{ce}(v) = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) \quad (6)$$

Where N is the number of samples, y_i is the real sleep stage label of the i th sample, and $\hat{y}_i$ is the sleep stage probability distribution of the i th sample predicted by the teacher model.

(2) Distillation pruning student model training

After the teacher model pretraining, the student model is trained. The loss function of the student model consists of two parts: supervised learning loss and distillation loss. The supervised learning loss aims to enable the student model to learn real labels directly, and the distillation loss enables the student model to learn the knowledge of the teacher model. The total loss function is

$$L = \beta L_{ce} + (1+\beta) L_{distill} \quad (7)$$

Where β is the weight coefficient that balances the supervised learning loss and the distillation loss. During the training process, pruning operations are performed on the basis of the importance of nodes or edges to reduce model complexity.

4. Result Analysis and Discussion

4.1 Experimental data

The publicly available PSG data set, containing sleep data from subjects of different ages, genders, and health conditions, is used. The data include various physiological signals, such as EEG, ECG, and EMG signals, as well as corresponding expert-labeled sleep stages (awake, light sleep, deep sleep, rapid eye movement sleep, etc.).

Data preprocessing: Preprocessing operations, such as filtering and denoising, are performed on physiological signals. For example, a band-pass filter is used on the EEG signal to remove low- and high-frequency noise, and all physiological signals are normalized so that the value range is between [0,1] to improve the training effect of the model.

4.2 Experimental setup

For the RDLG model, the number of layers of the graph neural network is set to 3, and the hidden dimension of the nodes in each layer is 128. The learning rate in reinforcement learning is α=0.001, and the discount factor is γ=0.9. The weight coefficient is β=0.5 in the distillation process. In the pruning operation, the importance threshold is set to 0.1, i.e., nodes or edges with an importance lower than 0.1 are pruned. For comparison models, the traditional sleep discrimination model based on SVM, the simple GNN model (without reinforcement learning and distillation pruning), and the LSTM model based on deep learning are selected.

4.3 Experimental results

Fig. 1 shows the overall sleep discrimination accuracy of different age groups. The abscissa represents the three age groups of 20–30 years old, 31–40 years old, and 41–50 years old, and the ordinate is the accuracy rate. The accuracy of the RDLG model in the 20–30 age group is about 0.8, and the accuracy in the 31–40 and 41–50 age groups is also close to 0.8. The accuracy of the SVM, GNN, and LSTM models in all age groups is lower than that of the RDLG model, mostly around 0.6. That is, the RDLG model has good adaptability to the sleep state discrimination of people of different age groups. It can effectively extract the characteristics of sleep data of different age groups for accurate judgment and is less affected by age factors.

Fig. 2 shows the overall sleep discrimination accuracy for different genders, where the abscissa indicates the male and female gender, and the ordinate is the accuracy. The RDLG model has an accuracy of around 0.8 in men and women, and that of the SVM, GNN, and LSTM models is around 0.6. This finding demonstrates that the RDLG model has stable performance in different-gender sleep discrimination, i.e., gender differences have minimal effect on its performance.

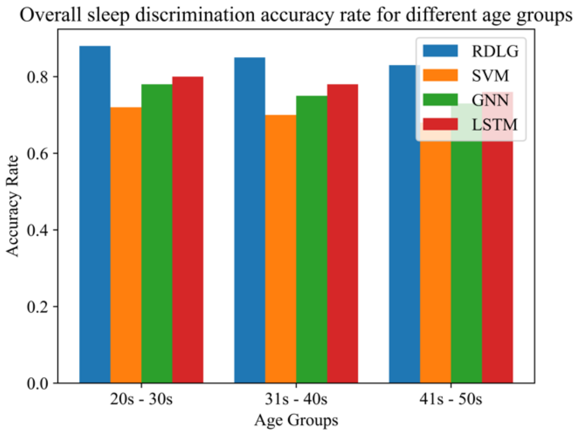

Fig. 1. Overall sleep discrimination accuracy rates for different age groups

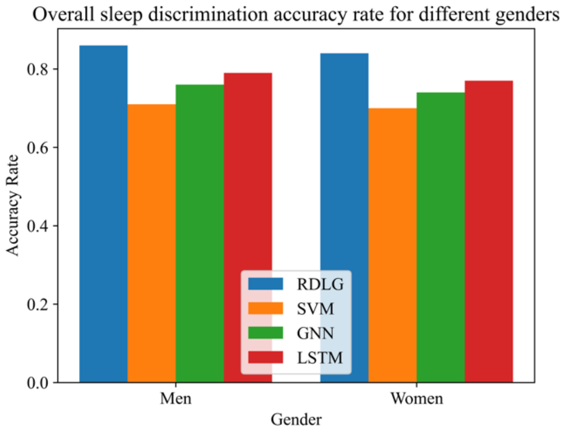

Fig. 2. Overall sleep discrimination accuracy rates for different genders

Fig.3 shows the accuracy of different models under different data set sizes, where the abscissa denotes the small data set (100 samples), the medium data set (500 samples), and the large data set (1000 samples), and the ordinate is the accuracy. The accuracy of the RDLG model is about 0.6 on small data sets, about 0.8 on medium data sets, and about 0.8 on large data sets. The accuracy of SVM, GNN, and LSTM models is about 0.4 on small data sets and about 0.6 on medium data sets, and the improvement is smaller on large data sets. This result reflects the strong adaptability of the RDLG model to data of different sizes. As the data set increases, its advantages strengthen. It can use more data to learn more abundant features, thereby improving the discrimination accuracy, whereas other models have a relatively weak response to the increase in data volume.

Fig. 4 compares the parameter amounts of different models, indicating 10^3 level for SVM, 10^4 level for GNN, 10^4 level for LSTM, and 10^3 level for RDLG. Although the proposed RDLG model has a high accuracy, the number of parameters is equivalent to that of the SVM model and much lower than that of the simple GNN and LSTM models. This finding shows that the multilevel distillation pruning technology effectively reduces the model redundancy and complexity.

Fig.5 shows the convergence speed of the reinforcement learning strategy in the RDLG model. The abscissa represents the three strategy types, namely, value function-based strategy, policy gradient-based strategy, and multi-strategy hybrid approach, and the ordinate is the convergence number. The convergence number of the multi-strategy hybrid approach is lowest, reaching about 40, the convergence number of the value function-based strategy is about 80, and the convergence number of the strategy gradient-based strategy is about 100. This result shows that the multi-strategy hybrid approach can enable the model to find a better decision-making strategy faster, speed up the convergence of the training process, and improve the training efficiency of the model.

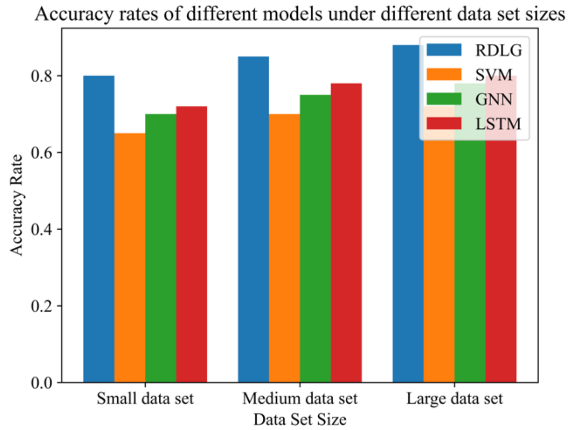

Fig. 3. Accuracy rates of different models under varying data set sizes

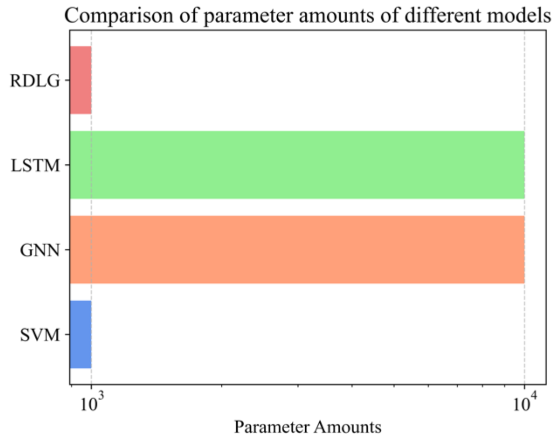

Fig. 4. Comparison of parameters of different models

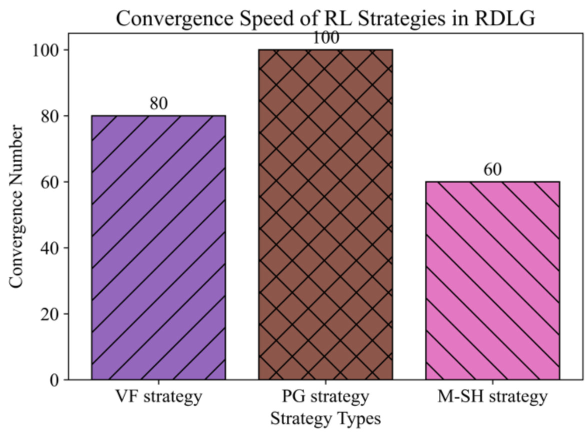

Fig. 5. Convergence speed of the reinforcement learning strategy in the RDLG model

In Fig. 6, the abscissa is the number of graph neural network layers (1.00–3.00), and the ordinate is the accuracy. When the graph neural network has three layers, the model accuracy is about 0.84. Within a certain range, as the number of layers increases, the accuracy shows an upward trend. However, after a certain number of layers, overfitting or performance degradation may occur. Thus, the number of graph neural network layers has a significant effect on model performance. The three-layer setting can better balance model complexity and performance under the current experimental conditions and provide a reference for determining the optimal number of graph neural network layers in the model.

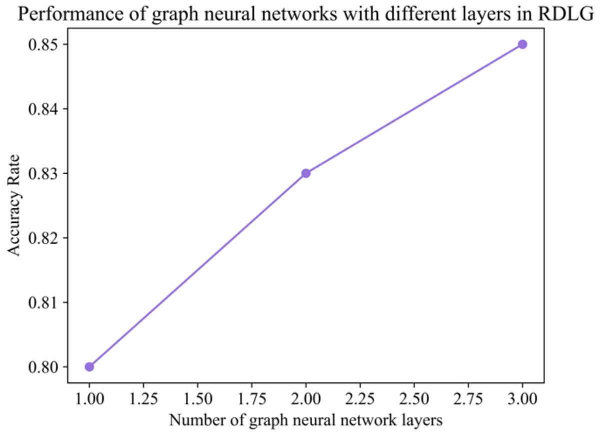

Fig. 6. Performance of graph neural networks with different layers in the RDLG model

In Fig. 7, the abscissa is the reward value (−4–4), and the ordinate is the frequency or frequency correlation. Reward values are distributed within a certain range, either positive or negative. For example, reward values are mainly spread in the range 1–3. The increase in positive rewards means that the model actions are more conducive to improving the accuracy of sleep discrimination. Through the distribution of reward values, we can understand the effect of the model's actions on the sleep discrimination results during the training process and evaluate the effectiveness of the reinforcement learning strategy and the learning status of the model.

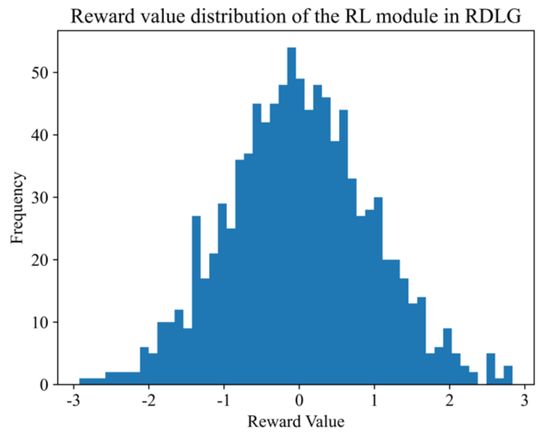

Fig. 7. Reward value distribution of the reinforcement learning module in the RDLG model

In Fig. 8, the abscissa is the state before and after pruning, and the ordinate is the number of model parameters. The number of model parameters before pruning is about 2000, which then decreases to about 1400 after pruning. That is, the multilevel distillation pruning technology effectively reduces the complexity of the model and redundant information in the model. Based on other charts, while the complexity is reduced, the accuracy of the model is not significantly negatively affected; instead, it improves, indicating that this technology enhances not only the efficiency of model operation but also the generalization ability of the model.

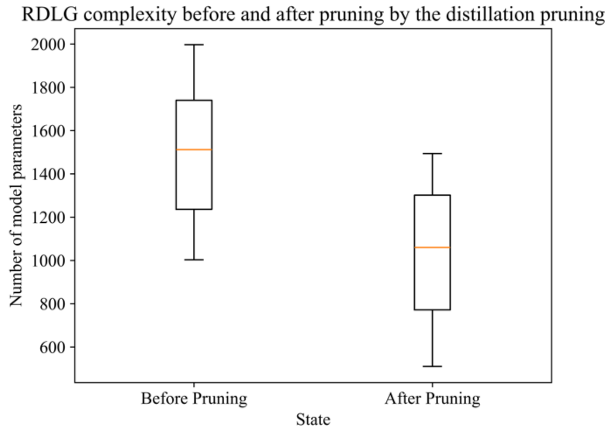

Fig. 8. Complexity of the RDLG model before and after pruning by the distillation pruning module

In Fig. 9, the abscissa is the learning rate (0.001, 0.005, 0.01), and the ordinate is the accuracy. When the learning rate is 0.001, the model accuracy is about 0.85; when the learning rate is 0.005, the accuracy is about 0.86; when the learning rate is 0.01, the accuracy is about 0.84. Accordingly, hyper parameters exert considerable effect on model performance. Different learning rate settings lead to changes in model accuracy. Reasonable selection of hyper parameters can optimize the effect of the model, providing an experimental basis for further adjusting model hyper parameters and improving model performance.

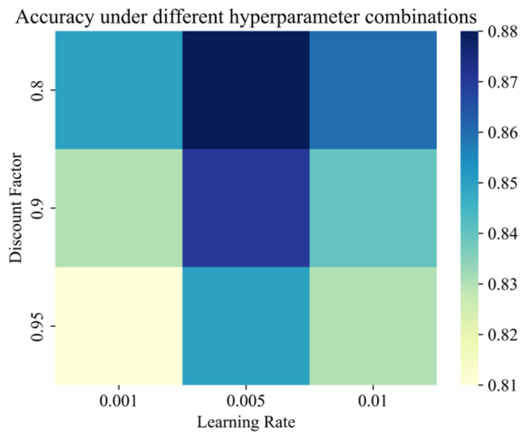

Fig. 9. Accuracy rate of the RDLG model under different hyper parameter combinations

As shown in Fig. 10, the RDLG model has a high accuracy, and only two samples are misclassified (category 0 is predicted as category 1, and category 1 is predicted as category 0). This result indicates that the RDLG model performs better in handling this classification task, with high precision and recall.

The accuracy of other models is relatively low, with four samples being misclassified (categories 0 and 1 are mis-predicted twice each). That is, other models perform worse than the RDLG model in this classification task and have higher error rates.

The superiority of the RDLG model is due to the multi-strategy adaptive reinforcement reward learning and multilevel distillation pruning technology it adopts, which helps the model better learn the characteristics of the data, thereby improving classification performance.

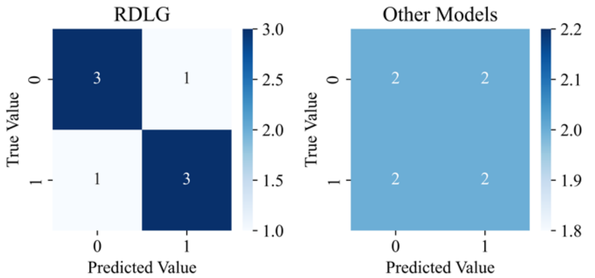

Fig. 10. Comparison of model performance

Fig. 11 shows the correlation between physiological indicators such as heart rate and EEG θ frequency band. The correlation between the EEG θ frequency band and a certain indicator 1 may reach about 0.6, whereas the correlation between heart rate and some indicators is weak, about 0.3. Through the correlation matrix, we can understand the correlation degree of different physiological indicators in the light sleep stage, which provides a basis for selecting appropriate physiological characteristics for sleep discrimination models and helps understand the relationship between sleep physiological mechanisms and model discrimination basis, thereby optimizing the model's ability to discriminate light sleep stages.

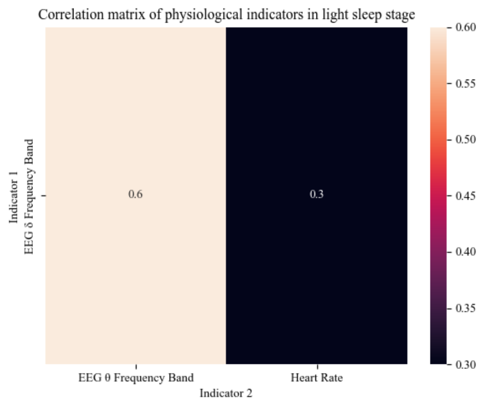

Fig. 11. Correlation matrix of physiological indicators in the light sleep stage

5. Conclusions

To explore compression and optimization methods for sleep recognition models and significantly reduce the size and complexity of the models, this study combined model improvement comparison and experimental research while retaining the core performance of the original model. The GraphRAG-based sleep discrimination method using multi-strategy adaptive reinforcement reward learning and multilevel distillation pruning was analyzed. The following conclusions could be drawn:

(1) The RDLG model shows significant performance advantages in sleep discrimination tasks. It is superior to traditional methods and other deep learning models in terms of accuracy. The RDLG model has an accuracy of about 0.8 in the 20–30 age group, and the accuracy in the 31–40 and 41–50 age groups is also close to 0.8. The SVM, GNN, and LSTM models have an accuracy of about 0.6, lower than that of the RDLG model in all age groups.

(2) The RDLG model can maintain or enhance the recognition accuracy while reducing the number of parameters, effectively reduce the complexity of the model, and improve the generalization ability. In terms of parameter amount, given 10^3 level for SVM, 10^4 level for GNN, 10^4 level for LSTM, and 10^3 level for RDLG, the RDLG model is equivalent to the SVM model, far lower than the simple GNN and LSTM models.

(3) The adaptive multi-strategy reinforcement reward learning of the RDLG model indicates high convergence speed and training efficiency.

RDLG shows obvious advantages in performance optimization, recognition accuracy, computing time, and memory usage. While ensuring low model complexity and parameter volume, it can significantly reduce the time and space cost of calculation, which makes it highly competitive in sleep discrimination systems and suitable for running in resource-constrained environments. It also improves the overall efficiency and performance of the system. Moreover, the adaptive multi-strategy reinforcement and reward learning hybrid method speeds up the convergence speed of the training process and provides a reference for further improving the training efficiency of the model. Future study can further optimize the reinforcement learning strategy and parameters of the model, explore more effective distillation pruning methods, and apply the model to more sleep-related fields, such as early warning of sleep disorders and personalized sleep management.

References

[1] F. B. Silva *et al.*, "Sleep staging algorithm based on smartwatch sensors for healthy and sleep apnea populations," *Sleep Med.*, vol. 119, pp. 535-548, May 2024.

[2] P. Ladosz, L. Weng, M. Kim, and H. Oh, "Exploration in deep reinforcement learning: A survey," *Inform. Fusion*, vol. 85, pp. 1-22, Oct. 2022.

[3] J. Gou, B. Yu, S. J. Maybank, and D. Tao, "Knowledge distillation: A survey," *Int. J. Comput. Vision*, vol. 129, pp. 1789-1819, Mar. 2021.

[4] C. K. Joshi, F. Liu, X. Xun, J. Lin, and C. S. Foo, "On representation knowledge distillation for graph neural networks," *IEEE T. Neur. Net. Lear.*, vol. 35, no. 4, pp. 4656-4667, April 2024.

[5] Y. Paul, R. Singh, S. Sharma, R. Singh, and I.H. Ra, "Efficient Sleep Stage Identification Using Piecewise Linear EEG Signal Reduction: A Novel Algorithm for Sleep Disorder Diagnosis," *Sensors-Basel*, vol. 24, no. 16, Aug. 2024, Art. no. 5265.

[6] E. Tuncer, "Development of New Machine Learning Based Algorithm for the Diagnosis of Obstructive Sleep Apnea from ECG Data," *J. Comput. Sci. Res.*, vol. 5, no. 3, pp. 15–21, July 2023.

[7] S. Xiao, B. Wang, and J. Zhang, "Automatic Sleep Stage Classification Based on an Improved K-means Clustering Algorithm," *J. Biomed. Eng.*, vol. 33, no. 5, pp. 847-854, Oct. 2016.

[8] Z. Liu and J. Sun, "Sleep Staging from the Visibility Graph Algorithm of Series," *Acta Electron. Sin.*, vol. 45, no. 1, pp. 225-231, Jan. 2017.

[9] C. Garcia-Vicente, G. C. Gutierrez-Tobal, F. Vaquerizo-Villar, A. Martin-Montero, D. Gozal, and R. Hornero, “SleepECG-Net: explainable deep learning approach with ECG for pediatric sleep apnea diagnosis,” *IEEE J. Biomed. Health Inform.*, vol. 29, no. 2, pp. 1021-1034, Feb. 2025.

[10] J. Jiménez-García *et al.*, “A 2D convolutional neural network to detect sleep apnea in children using airflow and oximetry,” *Comput. Biol. Med.*, vol. 147, Aug. 2022, Art. no. 105784.

[11] H. Shao, Y. Zheng, Y. Luo, L. Li, and L. Zhang, “Sleep staging algorithm based on class rebalancing unsupervised domain adaptation,” *J. Beijing Univ. Posts Telecommun.*, vol. 46, no. 6, pp. 8-14, Dec. 2023.

[12] Y. Tian, Q. Zhou, and W. Li, “Automatic sleep staging algorithm for stochastic depth residual networks based on transfer learning,” *J. Biomed. Eng.*, vol. 40, no. 2, pp. 286-294, 2023.

[13] Z. Jin, K. Jia and Y. Yuan, “A hybrid attention temporal sequential network for sleep stage classification,” *J. Biomed. Eng.*, vol. 38, no. 2, pp. 241-248, Mar. 2021.

[14] Z. Kohzadi, R. Safdari, and K. S. Haghighi, “Evaluation of the PSO Metaheuristic Algorithm in Different Types of Sleep Apnea Diagnosis Using RR Intervals,” *J. Biomed. Phys. Eng.*, vol. 13, no. 2, pp. 147-156, Oct. 2023.

[15] M. B. Uddin, C. M. Chow, and S. H. Ling, “A generalized algorithm for the automatic diagnosis of sleep apnea from per-sample encoding of airflow and oximetry,” *Physiol. Meas.*, vol. 43, no. 6, June 2022, Art. no. 065004.

[16] P. Ye *et al.*, “Diagnosis of obstructive sleep apnea in children based on the XGBoost algorithm using nocturnal heart rate and blood oxygen feature,” *Am. J. Otolaryng.*, vol. 44, no. 2, Mar. 2023, Art. no.103714.

[17] A. Sheta *et al.*, “Diagnosis of obstructive sleep apnea using feature selection, classification methods, and data grouping based age, sex, and race,” *Diagnostics*, vol. 13, no. 14, July 2023, Art. no. 2417.

[18] Q. Wang and S. Tong, “Semi-supervised EEG Sleep Staging Algorithm Based on Hybrid Neural Network,” *Radio Eng.*, vol. 53, no. 4, pp. 925-935, Mar. 2023.