Pattern Recognition 176 (2026) 113157

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Class-imbalanced graph contrastive clustering for sleep apnea prediction in mental health

Xin Liu [a], Xinke Wang [b,c,*], Jinyang Huang [a], Dan Guo [a], Meng Wang [a]

[a] School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, 230009, Anhui, China
[b] School of Artificial Intelligence, Anhui University, Hefei, 230601, Anhui, China
[c] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, 230088, Anhui, China

ARTICLE INFO

Keywords:
Obstructive sleep apnea
Mental health
Class imbalance
Graph contrastive learning
Graph clustering

ABSTRACT

Obstructive Sleep Apnea (OSA), as a prevalent sleep disorder, has impacts that extend far beyond the physiological level. It disrupts the neural mechanisms underlying affective processing and cognitive control, leading to emotional dysregulation, anxiety, and depressive symptoms, which pose significant risks to mental health. Research in this area provides key insights into emotional dynamics and mental health assessment. In this paper, we consider class imbalance in OSA data and rethink the homophilous graph assumption, observing that tail-class samples can form identifiable subgraphs and the message passing mechanism enhances their representational capacity. Hence, we construct a homophilous graph based on respiratory samples using dynamic time warping. Building on this, we propose a Class-Aware Graph Contrastive Clustering (GA-GCC) framework to address class imbalance. Specifically, at first, a locally weighted graph contrastive learning module is utilized to mine same-class samples within low-order neighborhoods. Second, to further emphasize tail class, we design a frequency-weighted graph semantic clustering module that generates frequency-based weights using pseudo-labels derived from semantic clustering. Finally, a globally class-aware weight optimization module is develop to integrate local and frequency-based weights, enabling GA-GCC to adaptively balances class distributions and mitigates intra-class redundancy. Since the lack of public non-contact OSA datasets limits multimodal affective research, we release ROSA, a radar-based dataset for three-class sleep apnea prediction. Extensive experiments on ROSA and other long-tailed datasets demonstrate the superior performance of GA-GCC, highlighting its potential for clinical and mental health applications.

1. Introduction

Obstructive Sleep Apnea (OSA) is a common sleep disorder characterized by recurrent interruptions of breathing during sleep [1,2]. It poses significant risks to mental health, particularly for emotional and affective functioning. In terms of emotional impact, obstructive sleep apnea can induce mood instability, emotional blunting, and deficits in emotional recognition and processing [3,4]. From an affective perspective, patients with severe OSA often exhibit abnormal affective responses and behavioral disturbances [5,6], including anxiety, depression, and cognitive-emotional dysregulation. These affective impairments are primarily attributed to nocturnal hypoxemia and fragmented sleep architecture, particularly the loss of restorative deep sleep. Therefore, early detection and continuous monitoring of OSA are essential for preserving good emotion and mental health.

Recent advances [7,8] in affective computing highlight that emotion and sentiment analysis (ESA) is a cornerstone of artificial emotional intelligence. In particular, a recent comprehensive survey [9] on label-efficient ESA systematically formulates a hierarchical taxonomy covering seven learning paradigms and provides a unified computational perspective for emotion recognition under limited supervision, offering valuable methodological guidance for affective modeling in data-scarce and noisy real-world scenarios. Motivated by these insights, integrating obstructive sleep apnea research with affective computing can enhance the understanding of emotional perception, physiological health [10,11]. Hence, accurate and timely prediction of obstructive sleep apnea not only facilitates the exploration of physiological responses associated with emotional states, but also provides a foundation for more precise assessment of affective dynamics. This integration offers novel opportunities to develop personalized interventions that target mental health and emotional regulation, bridging the gap between sleep medicine and affective computing applications.

Despite the significance of OSA, the current diagnostic standard, polysomnography (PSG), remains costly, labor-intensive, and

* Corresponding author.
E-mail address: wa22101009@stu.ahu.edu.cn (X. Wang).

https://doi.org/10.1016/j.patcog.2026.113157
Received 24 October 2025; Received in revised form 7 January 2026; Accepted 23 January 2026
Available online 30 January 2026

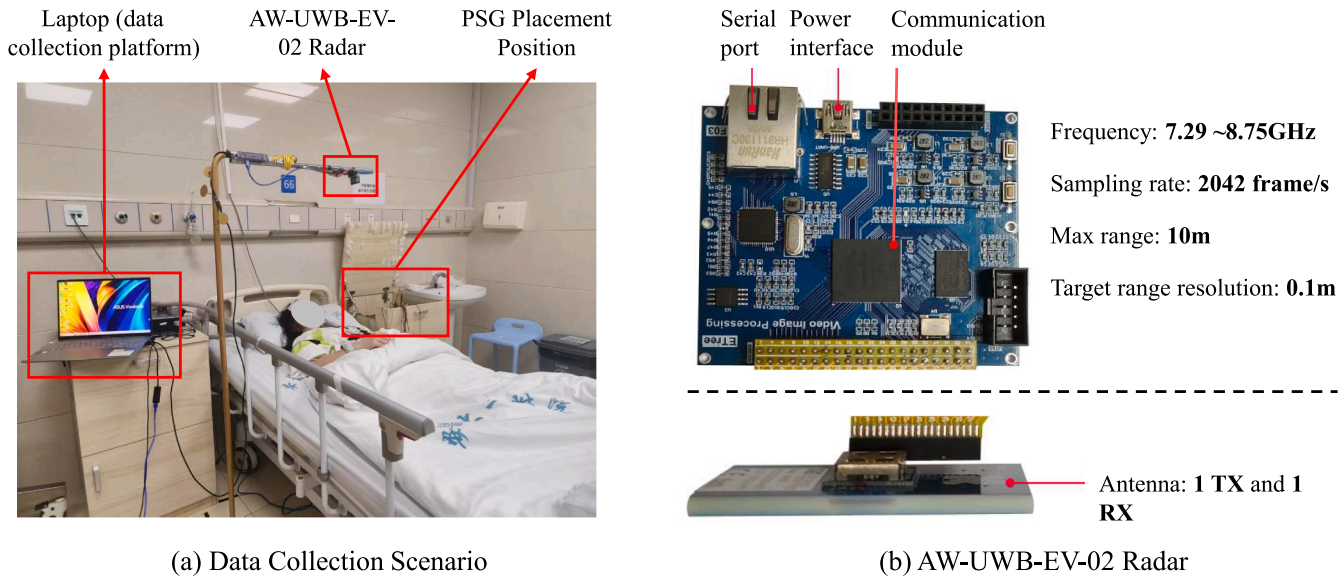

Fig. 1. (a) Clinical setting for data collection using radar. (b) The parameters of AW-UWB-EV-02 radar employed in the study.

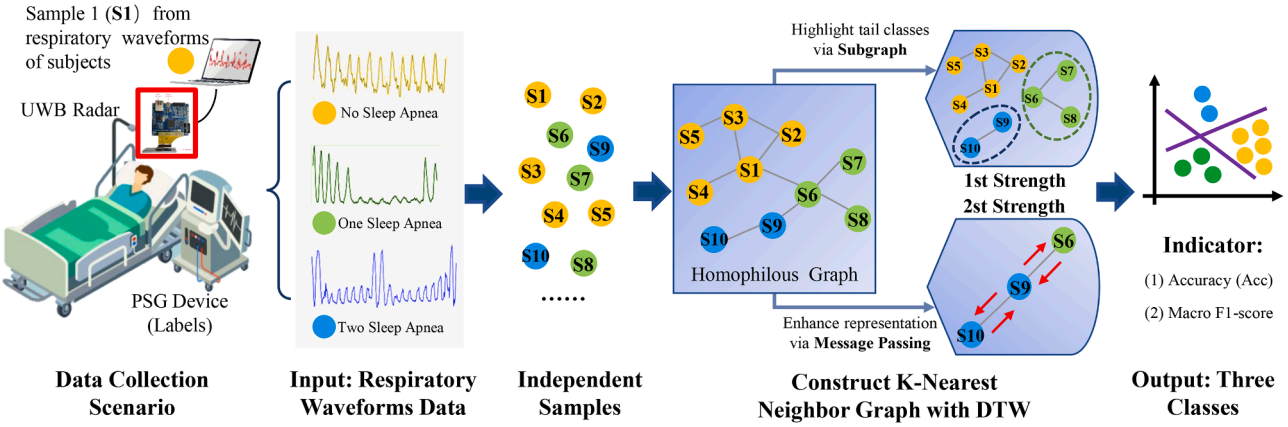

Fig. 2. Overall Process of radar-based sleep apnea prediction via homophilous graph: Based on radar-based respiratory waveform data, a homophilous temporal graph is constructed using DTW and K-Nearest Neighbors to predict sleep apnea class. In the homophilous graph, nodes of the same class can form clusters that are easier to identify, while the message passing mechanism allows nodes of the tail class to enhance their feature representations through connections with neighbors, thereby alleviating class imbalance.

uncomfortable for patients. To address these limitations, we employ a non-contact ultra-wideband radar (AW-UWB-EV-02) to collect sleep data, as shown in Fig. 1, recording overnight data from 17 subjects. These recordings are then partitioned into 6179 one-minute respiratory samples with fine-grained annotations, forming the radar-based obstructive sleep apnea (ROSA) dataset. Since sleep apnea events occur only during a portion of the total sleep period, severe class imbalance in ROSA and poses a significant challenge for analysis.

Some studies have shown that supervised contrastive learning can alleviate the problem of class imbalance to a certain extent, therefore, most methods prefer to utilize the contrastive learning paradigm to design the model, to solve the long-tailed distribution of independent sample data. For example, balanced contrastive learning (BCL) is proposed in [12] by using gradient contributions to ensure that all classes are present in each mini-batch. From the perspective of data distribution, Hou et al. [13] develop a subclass-balancing contrastive learning method to divide classes into subclasses, thereby alleviating the impact of imbalance at a finer granularity. More methods for exploring class imbalance using contrastive learning are discussed in [14,15].

Despite the effectiveness of these methods, they do not focus on the interaction between independent samples. Inspired by graph representation learning [16], we summarize the two advantages of graph structures compared to independent sample data when facing class imbalance: (1) The homophilous graph can connect the samples of the tail class to each other, forming a subgraph that is easier to discover; (2) The message passing mechanism for graph-structured data enables the nodes in the tail class to "see" more nodes by connecting with other nodes, thereby enhancing their representation, as illustrated in Fig. 2. Hence, we construct a homophilous graph from radar-based respiratory waveform samples to perform sleep apnea prediction.

The principle of homophily requires that two samples connected by an edge have a high similarity. Since cosine similarity depends on the quality of sample representations, accurate similarity measurements usually require representations learned through a pre-trained model. However, a good pre-trained model is not a trivial matter. Hence, we adopt dynamic time warping (DTW) [17,18] to measure similarity between samples, which enables non-linear alignment to capture similar patterns occurring at different temporal positions, while maintaining robustness to noise and local distortions, thus providing a more reliable measure of similarity between temporal signals. Building on this, we establish a tail-class-friendly homophilous graph, as shown in Fig. 2.

Nevertheless, the imbalance node classification is still a critical problem, many existing methods consider designing models based on ideas such as reweighting and resampling [19]. The key insight of these methods is to increase the attention to the tail class while reducing the attention to the head class. Inspired by our LocWGCL [20], negative pairs in graph contrastive learning can reflect the attention of the anchor to negative samples. Essentially, seeking false hard negatives (samples that are of the same class as the anchor under the true label and have a high similarity) is a strategy to reduce the mutual attention between similar samples, implying that head and tail classes no longer attend to themselves.

When we apply LocWGCL to the obviously imbalance graph data, there exist two apparent limitations. On the one hand, LocWGCL only focuses on samples within a 2-hop subgraph, because most of the nodes of the same class as the anchor in the homophilous graph exist in this area. On the other hand, LocWGCL cannot measure the weight relationship between the head and tail classes. To address these problems, we propose graph clustering as an effective strategy to capture high-order neighborhood information and distinguish head and tail classes using pseudo-labels.

How to use graph clustering strategies to obtain accurate pseudo-labels is a crucial but challenging issue. There is no doubt that using only kmeans to address this problem is not sufficient for the graph data. The key reason is that most graph contrastive clustering methods are centered on node representation, while ignoring the structure of graph. Therefore, developing a clustering loss that combines representation and graph structure is essential. Inspired by NS4GC [21], we argue that the node similarity matrix needs to satisfy the inherent semantic relationships among nodes, and design a clustering loss for the high-order nodes, enabling the designed clustering loss to work synergistically with the locally weighted loss of LocWGCL. Then, based on the pseudo-labels generated from the clustering layer, we design a frequency-weighted strategy to further address the two limitations of LocWGCL. Here, frequency-based weights can be assigned to negative pairs outside the role of LocWGCL in graph contrastive learning, whose key insight is to increase the weight of negative pairs from different classes, especially from the head and tail classes, respectively, and vice versa. Therefore, the frequency-weighted strategy highlights the focus on tail classes.

By integrating frequency-based weights with local weights of LocWGCL, we design a new global weight strategy to increase the focus on tail classes while eliminating the adverse impact of false negatives. Specifically, we can assign the weights of first-order nodes using the local weight strategy of LocWGCL. In terms of the second-order neighborhood, we can assign weights of second-order nodes based on both the message similarity and the representation similarity of LocWGCL. As for remaining second-order and high-order nodes, normalized cosine similarity and frequency-weighted strategy are employed to weight negative pairs of the same and different classes, respectively. Hence, all nodes of the head and tail classes are balanced with adaptive weights.

In this paper, we propose the Class-Aware Graph Contrastive Clustering (GA-GCC) framework, which constructs a homophilous graph from radar-based respiratory samples and design a globally class-aware weight optimization module to balance head and tail classes, yielding expressive node representations for accurate prediction of three-class sleep apnea. We evaluate GA-GCC on the ROSA dataset and public datasets with long-tail distributions in other fields, demonstrating its effectiveness and generalization in mitigating class imbalance. The main contributions are summarized as follows:

- We release the ROSA dataset, the first publicly available radar-based sleep apnea dataset, providing a benchmark for non-contact sleep apnea prediction.
- The class imbalance issue is mitigated by constructing a DTW-based homophilous graph, where tail-class samples form identifiable subgraphs and their representations are enhanced through message passing.
- Building on LocWGCL, a frequency-weighted graph semantic clustering module is proposed, which generates frequency-based weights to enhance tail-class representation. These frequency-based weights are combined with the local weights of LocWGCL to obtain globally balanced class-aware weights.
- A Class-Aware Graph Contrastive Clustering framework is proposed that balances head and tail classes to alleviate class imbalance, producing expressive node representations to support accurate sleep apnea detection. The extended experiments reveal the outstanding performance of GA-GCC.

2. Related work

2.1. Sleep apnea prediction

Obstructive sleep apnea (OSA) is a common sleep disorder, which not only has physiological characteristics but also profoundly impacts emotional regulation and mental health. Mounting evidence suggests it is associated with impaired mood and affective functioning, including disruptions to emotional stability, emotional processing, and affective regulation, often manifesting as affective blunting, anxiety, depression, and cognitive-emotional dissonance. OSA analysis can provide a deeper understanding of the connectedness between sleep disorders and emotional states, paving the way for affective computing and mental health monitoring [22,23]. In early studies, polysomnography (PSG) serves as the gold standard, achieving diagnosis through multi-channel signal recording including electroencephalography (EEG), electrocardiography (ECG), and blood oxygen saturation, yet face limitations of cumbersome equipment and high costs [24]. With the advancement of deep learning, Chaw et al. [25] employ deep convolutional neural networks to process peripheral oxygen saturation signals to achieve sleep apnea prediction. Subsequently, Bitners and Arens [24] address pediatric OSA by proposing a comprehensive evaluation framework, including clinical history, polysomnography scoring, and surgical interventions. Innovations in hardware devices have propelled the development of home-based detection. Edouard et al. [26] analyze respiratory patterns and cardiac activities through mattress pressure sensors and deep learning algorithms. In recent years, researchers have increasingly focused on algorithm efficiency and non-contact monitoring. For example, Li et al. [27] utilize ultra-wideband (UWB) signals with attention-enhanced models to monitor respiratory rate and diagnose sleep apnea. Although existing work has achieved radar-based sleep apnea prediction, the inherent class imbalance problem in sleep apnea data has not been explored. This paper focuses on the class imbalance problem from the perspectives of graph structure and class-aware weights.

2.2. Class imbalance

Class Imbalance is a universal phenomenon in practical scenarios, especially for mental health data. Recent studies have shown that contrastive learning, especially supervised contrastive learning, holds significant promise in addressing the class imbalance problem [12,13,28]. For example, to explore the advantages of supervised contrastive learning in solving class imbalance, Hou et al. [13] propose a subclass-balancing contrastive learning method, which refines the class structure by dividing classes into subclasses, to more finely mitigate the impact of imbalance. There are also some methods dedicated to optimizing the training process or modeling long-tailed distribution data, which are discussed in [14,15,29]. From contrastive learning to graph contrastive learning (GCL), some approaches study GCL to alleviate the node class imbalance. To explore the role of GCL, Qian et al. [30] propose co-modality graph contrastive learning (CM-GCL), including inter-modality GCL and intra-modality GCL. Recently, a dynamic graph mixed supervised contrastive learning (DGMSCL) method is designed [31], which

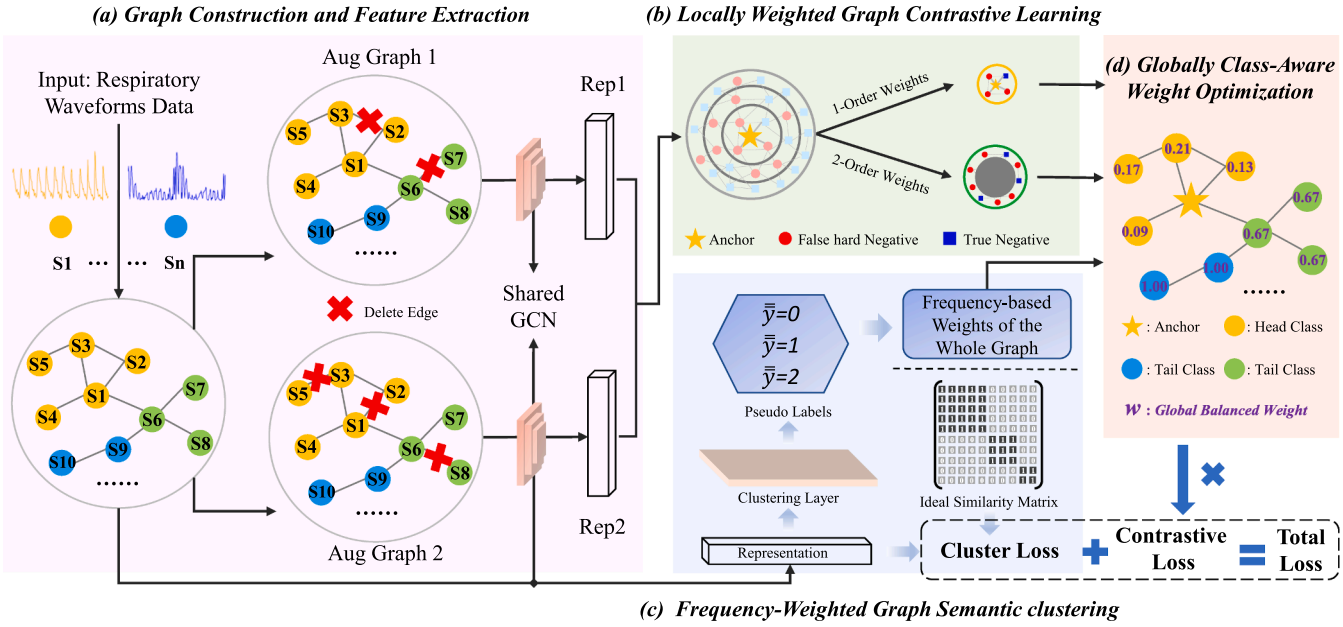

Fig. 3. The overview of the proposed GA-GCC Framework. There are four main steps: (a) The respiratory waveform data are constructed as a homophilous graph, and node representations are learned using graph contrast learning paradigm. (b) The locally weighted graph contrastive learning (LocWGCL) module is introduced to generate local weights for low-order neighborhoods. (c) The frequency-weighted graph semantic clustering (FreGSC) module is designed to produce frequency-based weights focusing on tail classes. (d) Based on the local and frequency-based weights, the globally class-aware weight optimization module (GloCWO) is applied to obtain globally balanced class weights for graph contrastive loss.

integrates dynamic graph characteristics and mixed supervision mechanisms, further advancing performance on imbalanced graph data. Despite handling class imbalance well, GCL struggles with graph structure mining. Therefore, in this paper, we emphasize the perception of semantic relationships between nodes.

2.3. Graph contrastive and clustering learning

Graph Contrastive Learning (GCL) has emerged as a prominent paradigm in graph self-supervised learning, where node-level GCL has attracted significant attention due to its fine-grained representation capabilities. Early and influential methods is Deep Graph Infomax (DGI) [32] maximizes mutual information between node-level representations and the global summary of the graph. Following this idea, GRACE [33] adopts the SimCLR [34] framework and designs random augmentation functions for the graph to enhance node representations. Its subsequent variants, such as GCA [35] and directed-GCL [36], further refine augmentation strategies or incorporate structural priors. Since both random augmentations may distort semantic structures. To address this issue, recent works [37,38] tend to explore learnable and adaptive augmentations. Graph clustering, as a fundamental task in graph data mining, has witnessed significant developments with the advancement of self-supervised learning. The rise of contrastive learning brings new paradigms to graph clustering. For efficiency, Liu et al. [39] introduce a simple graph contrastive clustering, which simplifies the network architecture using low-pass denoising and parameter-unshared siamese MLPs. Recent works focus on low-order interaction modeling and personalized feature selection. Liu et al. [21] develop a reliable node similarity matrix guided graph contrastive clustering (NS4GC). Xie et al. [40] propose the feature personalized graph clustering to select cluster-relevant features for each node. Although the above graph contrastive clustering models achieve competitive performance, they may conflict with LocWGCL loss, which hinders the model from reaching ideal performance. Inspired by the ideal of NS4GC, we propose a new semantic clustering loss.

3. Methodology

3.1. Overview of our framework

In this paper, we focus on the prediction of three-class sleep apnea in mental health, with the aim of addressing the class imbalance problem inherent in this task. As illustrated in Fig. 3, we propose the Class-Aware Graph Contrastive Clustering (GA-GCC) framework, which enhances node representations while balancing head and tail classes. The implementation of GA-GCC consists of four modules: (a) graph construction and contrastive learning (Section 3.2), which builds a homophilous graph using DTW from ROSA samples and obtains node representations by shared graph convolutional neural networks (GCN); (b) locally weighted graph contrastive learning (LocWGCL, Section 3.3), which mines same-class samples within low-order neighborhoods; (c) frequency-weighted graph semantic clustering (FreGSC, Section 3.4), which generates frequency-based weights using pseudo-labels derived from semantic clustering layer to enhance tail-class representations; and (d) globally class-aware weight optimization (GloCWO, Section 3.5), which integrates local and frequency-based weights to obtain globally balanced class weights. Together, these modules define the overall optimization function of GA-GCC, as detailed in Section 3.6. Finally, based on the learned GA-GCC framework, we obtain discriminative representations of respiratory samples, which are then fed into a classifier to produce prediction of three-class sleep apnea.

3.2. Graph construction and contrastive learning

Inspired by the homophily and message passing mechanism of the graph, we summarize the advantages of graph structures compared to independent sample data under class imbalance. On the one hand, the homophilous graph naturally connects the samples from the tail class to each other, forming a subgraph that is easier to discover in the whole graph. On the other hand, the message passing mechanism of graph representation learning enables the nodes in the tail class to receive more messages from other nodes, which enhances the representation

ability. Therefore, in this paper, we construct a homophilous graph with independent sleep apnea data.

Cosine similarity is a typical method for measuring the similarity of nodes, since the L2 norm of cosine similarity can project the node representation onto the hypersphere. However, an accurate similarity measurement is generally achieved through the learned representation. For simplicity, we introduce dynamic time warping (DTW) to measure the similarity between samples instead of the pre-trained model. DTW calculates the similarity between two time series by dynamically "stretching" or "compressing" the time axis to identify the best possible alignment, which suits for respiratory waveforms with different breathing frequencies.

DTW uses dynamic programming to find a path with the least cost to achieve optimal alignment and constructs a distance matrix. The minimum total difference between two sequences under optimal alignment measures the similarity between the sequences. Given N sleep apnea samples, we can obtain a similarity matrix $S_{dtw} \in \mathbb{R}^{N\times N}$ based on DTW, where each element represents the DTW-based similarity between two samples. To construct the homophilous graph, the K-nearest neighbor graph is a reasonable solution, ensuring that nodes in the tail class can receive messages from other nodes.

Let $\mathcal{X} = \{x_1, x_2, \ldots, x_N\}$ be the set of sleep apnea samples, S_{dtw} is the similarity matrix. For each node x_i, the first-order neighborhood is defined as $\mathcal{N}_1^k(x_i) = \{x_j \in \mathcal{X}\backslash\{x_i\}\}$, where x_j is among the top-k most similar to x_i according to S_{dtw}. Therefore, the adjacency matrix A is defined as:

$$A_{ij}^{\text{dir}} = \begin{cases} 1, & \text{if } x_j \in \mathcal{N}_1^k(x_i), \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Then, we transform this directed k-nearest neighbor graph into an undirected graph:

$$A_{ij} = \begin{cases} 1, & \text{if } x_j \in \mathcal{N}_1^k(x_i) \text{ or } x_i \in \mathcal{N}_1^k(x_j), \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The adjacency matrix can be converted to the edge set $\mathcal{E}$. Currently, we have constructed an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{X})$ based on independent sleep apnea samples.

To learn the node representation via graph contrastive learning, we employ random augmentation functions, including masking node features and removing edges, to obtain two augmentation views $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$ of the original graph $\mathcal{G}$. Generally, we assume that an anchor node v_i^1 is from $\tilde{\mathcal{G}}_1$, thus, the corresponding node v_i^2 in $\tilde{\mathcal{G}}_2$ is the positive sample for the anchor. In parallel, other nodes of $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$ are treated as negative samples, forming the intra-view and inter-view negative pairs. Therefore, node representations of the original graph and two augmentation graphs can be learned using the GCN [16] encoder in the graph contrastive learning paradigm.

3.3. Locally weighted graph contrastive learning

As definitions in locally weighted graph contrastive learning (LocWGCL), false negative is a negative sample of the same class as the anchor, and hard negative is a negative sample with a high similarity to the anchor. LocWGCL first studies the distribution of false hard negatives. Guided by the homophily prior, we investigate the issue through representation smoothness, capturing similarity variations across multi-hop subgraphs. Higher smoothness suggests more false hard negatives. The smoothness is defined as:

$$\boldsymbol{e}^T L \boldsymbol{e} = \frac{1}{2} \sum_{(i,j)\in\mathcal{E}} A_{ij}(\boldsymbol{e}_i - \boldsymbol{e}_j)^2, \tag{3}$$

where $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ are the representation of node v_i and v_j, A is adjacency matrix, L is the Laplacian matrix, defined as $L = D - A$ (D denotes the degree matrix. A lower value of Eq. (3) encourages higher subgraph smoothness. According to the subgraph smoothness, false hard negatives are distributed primarily over the local neighborhoods of the anchor, especially within its first- and second-order neighborhoods.

Specifically, for the first-order nodes of the anchor, they are more likely to be of the same class as the anchor, implying that the message passing is direct. Therefore, adopting only representation similarity is sufficient to identify false hard negatives, and representation similarity is also utilized to assign the weight of first-order negative pairs. Furthermore, for the second-order nodes of the anchor, LocWGCL first filters out nodes that are unlikely to be regarded as hard negatives via the average representation similarity between all second-order neighbors and the anchor. Then, the random walk with restart is employed to measure the passed message between the anchor and the second nodes through their common first-order nodes. Upon representation similarity and message similarity, an adaptive strategy is designed to weight the probability that the negatives belong to the false hard negatives in the second-order neighborhood.

Let node v_i be the anchor, given the union $\mathcal{N}_1^U(v_i)$ of first-order neighbor and the intersection $\mathcal{N}_2^I(v_i)$ of second-order neighbor from the two augmentation views $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$, where the intersection $\mathcal{N}_2^I(v_i)$ are initially filtered via the average representation similarity. Based on $\mathcal{N}_1^U(v_i)$ and $\mathcal{N}_2^I(v_i)$, we calculate the weights $w_1(i,j)$ and $w_2(i,j)$ for the first- and second-order neighbors:

$$\begin{aligned} w_1(i,j) &= 1 - \frac{s_{rep1}(i,j) + s_{rep2}(i,j)}{2}, j \in \mathcal{N}_1^U(v_i), \\ w_2(i,j) &= 1 - (\alpha_j \frac{s_{rep1}(i,j) + s_{rep2}(i,j)}{2} + \\ &\qquad (1-\alpha_j \frac{s_{mes1}(i,j) + s_{mes2}(i,j)}{2}, j \in \mathcal{N}_2^I(v_i), \end{aligned} \tag{4}$$

where s_{rep1} and s_{rep2} denote the normalized representation similarity from $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$, respectively. Similarly, s_{mes1} and s_{mes2} are message similarity. α_j is a regulator factor, which is calculated as follows:

$$\alpha_j = \frac{\sigma(s_{rep}(i,j))}{\sigma(s_{rep}(i,j)) + \sigma(s_{mes}(i,j))}, \tag{5}$$

where $\sigma(\cdot)$ is the sigmoid function for normalization. For detailed calculation process of message similarity, please refer to LocWGCL [20].

3.4. Frequency-weighted graph semantic clustering

LocWGCL identifies false hard negatives among the first- and second-order neighborhoods, helping to determine which nodes are potentially of the same class as the anchor. Nevertheless, higher-order neighborhoods remain unexplored. In addition, it is insufficient to focus solely on nodes of the same class, and attention must also be given to the tail classes. To realize this goal, we tend to introduce clustering to provide the pseudo-label. However, applying K-means clustering directly to the learned representations often fails to generate reliable pseudo-labels, which potentially leads to conflicts with the LocWGCL weight. To improve the quality of pseudo-labels, exploring the semantic relationships among nodes is essential.

Inspired by NS4GC [21], a semantic clustering loss is proposed to alleviate conflicts with LocWGCL loss, whose key insight lies in adding edges between disconnected node pairs when they exhibit relatively high similarity. To avoid conflicts with locwgcl in second-order neighborhoods, semantic cluster loss only applies to nodes outside of LocWGCL identification, which is formulated as follows:

$$\ell_{clu} = \sum_{A_{ij}=0,\ j\neq i,\ j\notin\mathcal{N}_1^U(v_i),\ j\notin\mathcal{N}_2^I(v_i)} \sigma\left(\frac{S_{ij} - r}{\tau}\right), \tag{6}$$

where σ is the sigmoid function, $\mathcal{N}_1^U(v_i)$ and $\mathcal{N}_2^I(v_i)$ denote the sets of first- and second-order nodes identified by LocWGCL, and A and S represent the adjacency matrix and the cosine similarity matrix from the original graph $\mathcal{G}$, respectively. The similarity threshold r is used to decide whether an edge connection should be formed between the nodes. τ is the temperature parameter, which controls the smoothness of the

similarity-based penalty. A smaller τ sharpens the discrimination, enforcing a harder separation, while a larger τ allows for some fuzziness.

The clustering loss can improve the reliability of pseudo-labels via the exploration of both node representation and graph structure. Pseudo-labels are assigned through K-means, the number of nodes in each class can be obtained. The number of clusters in K-means is prior, which is set to 3 for the ROSA dataset. Since the proposed model is designed for sleep apnea prediction, as long as the data partitioning is consistent, the features of this domain are relatively stable and reproducible patterns of respiratory events. Therefore, once the number of clusters in this domain is determined, the same prior can be reused on subsequent sleep apnea data without any modifications.

Therefore, the number of nodes for each class can be expressed as $m_c, c \in [1, 3]$, where m_c is sorted from head class to tail class. Upon this, we design a relative frequency difference weight for negative pairs to enhance attention to tail classes, which can be formulated as:

$$w^{fre} = \begin{cases} 0, & \text{if } c_i = c_j, \\ \dfrac{\left| m_{c_i} - m_{c_j} \right|}{\max\left(m_{c_i},\ m_{c_j} \right)}, & \text{if } c_i \neq c_j, \end{cases} \tag{7}$$

where c indicates the class of the node. To match the weights of LocWGCL, we normalize w_{fre} with the min-max strategy as follows:

$$w^{FRE} = \frac{w^{fre} - \min(w^{fre})}{\max(w^{fre}) - \min(w^{fre})}. \tag{8}$$

Until now, we have finished the frequency-weighted graph semantic clustering (FreGSC) module, which use frequency-based weights w^{FRE} to quantify the weighting relationship between head and tail classes. This module assigns greater weights to negative pairs stemming from different classes, particularly emphasizing pairs involving head and tail classes, and vice versa.

3.5. Globally class-aware weight optimization

LocWGCL seeks false (hard) negatives via representation similarity and message similarity from a local perspective, while the FreGSC module identifies false negatives and tail classes based on graph clustering. To integrate LocWGCL and LocWGCL, we design a globally class-aware weight opetimization (GloCWO) module to alleviate the adverse impact of class imbalance by adaptive weights.

Specifically, given two augmentation views $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$. For the anchor node v_i, we can obtain the union $\mathcal{N}_1^U(v_i)$ of the first-order neighborhood. The weights of the first-order nodes are calculated as follows:

$$w_1(i, j) = 1 - s(i, j), \tag{9}$$

where $s(\cdot, \cdot)$ denotes the normalized cosine similarity and $j \in \mathcal{N}_1^U(v_i)$. According to the assumption of homophilous graph, all nodes within $\mathcal{N}_1^U(v_i)$ are the same class as the anchor and estimate their likelihood of belonging to the same class with representational similarity. Even if some first-order nodes are not of the same class as the anchor, they can still be measured by the normalized cosine similarity.

In terms of the second-order neighborhood, for the anchor node v_i, we can get the intersection $\mathcal{N}_2^I(v_i)$ of second-order nodes from the two augmentation views, while $\mathcal{N}_2^{I^{sel}}(v_i)$ represents the intersection of the sets of second-order false hard negatives selected from the two augmented views, as described in Section 3.3. Let $j \in \mathcal{N}_2^U(v_i)$, the weight of the second-order neighborhood is defined as follows:

$$w_2(i, j) = \begin{cases} w_2^{\text{LocWGCL}}(i, j), & \text{if } j \in \mathcal{N}_2^{I_{sel}}(v_i), \\ 1 - s(i, j), & \text{if } j \in \mathcal{N}_2^{I}(v_i) \setminus \mathcal{N}_2^{I_{sel}}(v_i) \text{ and } \hat{y}_i = \hat{y}_j, \\ w_2^{\text{FRE}}(i, j), & \text{if } j \in \mathcal{N}_2^{I}(v_i) \setminus \mathcal{N}_2^{I_{sel}}(v_i) \text{ and } \hat{y}_i \neq \hat{y}_j, \end{cases} \tag{10}$$

where $s(\cdot, \cdot)$ denotes the normalized cosine similarity, w_2^{LocWGCL} denotes the weights assigned to the selected second-order false hard negatives based on the LocWGCL method and w_2^{FRE} represents the frequency-based weights assigned to second-order nodes via the FreGSC module. In LcoWGCL, due to the significant structural differences of augmentation views $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$, the nodes within the intersection $\mathcal{N}_2^{I_{sel}}(v_i)$ constitute only a small fraction of the total second-order nodes. Besides, these nodes are jointly identified based on representational similarity and message similarity, which introduces both node features and graph structural properties. Therefore, the identified nodes are more reliable. For other second-order nodes, the cosine similarity and frequency-based weights are employed to weight negative pairs from the same and different classes, respectively.

As for other nodes, the set can be defined as $\mathcal{N}_{other}(v_i)$ of the anchor node v_i. Similar to $w_2(i, j)$, we also utilize the pseudo-labels to judge the class of nodes, and cosine similarity and frequency-based weights are adopted to weight negative pairs. When $j \in \mathcal{N}_{other}(v_i)$, the $w_{other}(i, j)$ is formulated as follows:

$$w_{other}(i, j) = \begin{cases} 1 - s(i, j), & \text{if } \hat{y}_i = \hat{y}_j, \\ w_{other}^{\text{FRE}}(i, j), & \text{if } \hat{y}_i \neq \hat{y}_j. \end{cases} \tag{11}$$

Finally, $w_1(i, j)$, $w_2(i, j)$, and $w_{other}(i, j)$ are integrated into $w(i, j)$ for negative pairs consisting of the anchor v_i and the node v_j, written as follows:

$$w(i, j) = w_1(i, j) + w_2(i, j) + w_{other}(i, j). \tag{12}$$

3.6. Overall loss function for GA-GCC

Based on Eq. (12), we can calculate the global weights for all negative pairs. Hence, the global weights are applied to the GCL loss to guide model training. The loss function of GCL with global weights can be written as follows:

$$\ell_w(\boldsymbol{e}_i^1, \boldsymbol{e}_i^2) = \log \frac{\exp\left(\frac{\theta(\boldsymbol{e}_i^1, \boldsymbol{e}_i^2)}{\tau} \right)}{\exp\left(\frac{\theta(\boldsymbol{e}_i^1, \boldsymbol{e}_i^2)}{\tau} \right) + \sum_{j \neq i} w(i, j) \exp\left(\frac{\theta(\boldsymbol{e}_i^1, \boldsymbol{e}_j^2)}{\tau} \right) + \sum_{j \neq i} w(i, j) \exp\left(\frac{\theta(\boldsymbol{e}_i^1, \boldsymbol{e}_j^2)}{\tau} \right)}, \tag{13}$$

$$\ell_{wgcl} = -\frac{1}{2N} \sum_{i=1}^{N} (\ell_w(\boldsymbol{e}_i^1, \boldsymbol{e}_i^2) + \ell_w(\boldsymbol{e}_i^2, \boldsymbol{e}_i^1)), \tag{14}$$

where τ is the temperature parameter, also used in Eq. (6). The main function of temperature parameter τ in (Eq. (13)) can be summarized as adjusting the sharpness of the similarity distribution, thereby affecting the model's ability to distinguish between positive and negative samples. When τ is small, the distribution becomes sharp, amplifying subtle differences in similarity. When τ is large, the distribution is smoother and the differences between positive and negative samples are blurred. It can be observed that the temperature parameter τ plays a similar role in Eqs. (6) and (13). Upon Eqs. (6) and (13), the final loss function can be formulated as:

$$L = \lambda \ell_{wgcl} + (1 - \lambda) \ell_{clu}, \tag{15}$$

where λ is the hyperparameter to balance the contributions of ℓ_{wgcl} and ℓ_{clu}.

Remarkably, since the augmentation graphs are regenerated in each epoch, the neighborhood of the anchor changes in real time. Therefore, we dynamically update the weights according to a certain training interval to capture the structural changes of the augmentation views for more reliable node representation and pseudo-labels.

4. Experiments

4.1. Datasets

4.1.1. ROSA: Radar-based obstructive sleep apnea dataset

To construct the radar-based obstructive sleep apnea (ROSA) dataset, the ultra-wideband radar AW-UWB-EV-02 is employed to collect the

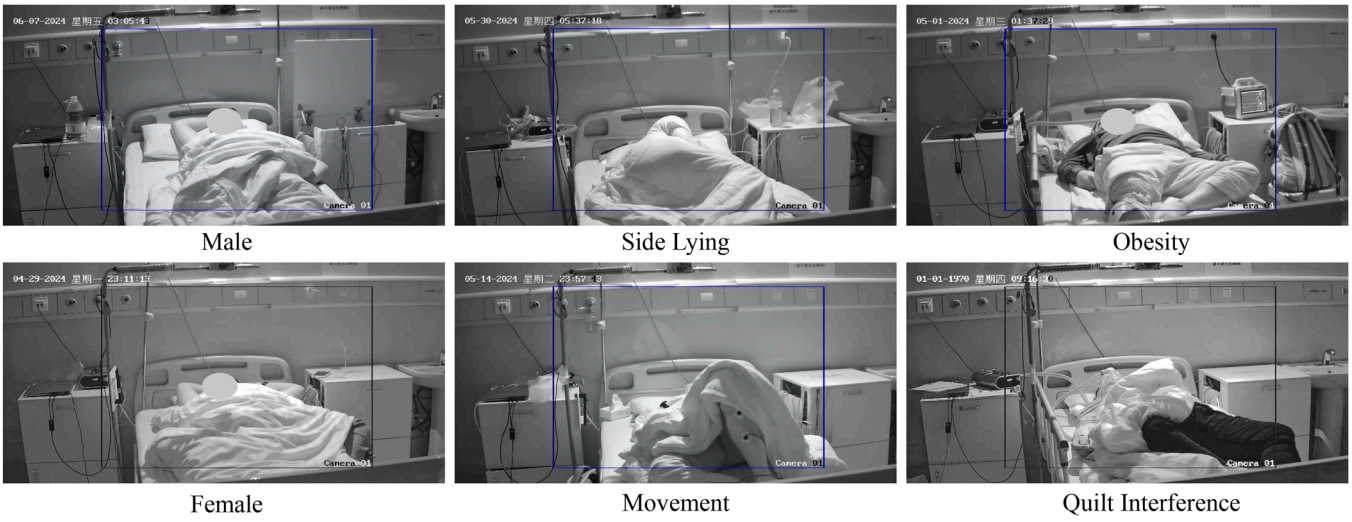

Fig. 4. Real-world data collected from subjects of different genders and sleeping postures.

sleep apnea data. To ensure the naturalness of the clinical data and better simulate real-world scenarios, no constraints are imposed on the environment and subjects, such as covering, gender, and sleeping posture, as shown in Fig. 4.

Since the sleep apnea dataset is initially established for sleep apnea classification tasks, the radar data are first divided into different samples. Taking into account that obstructive sleep apnea usually lasts more than 10 s, we adopt a one-minute sliding window and a one-minute step size to segment radar data based on PSG. Unfortunately, continuous segmentation operation often causes a one-time sleep apnea to be cut by two consecutive segmented samples. To avoid this issue, we dynamically move forward or backward one-time sleep apnea sampling frames that are split between two samples to ensure that one-time sleep apnea sampling frames can be completely included in one segmented sample. In addition, since the sampling rate of AW-UWB-EV-02 is too high (2042 frames/s), we downsample the sample to 5 frames/s to reduce redundant information.

The ultra-wideband radar senses respiratory movement and expresses it through complex signals. Specifically, radar data are composed of the in-phase component (I) and the quadrature component (Q), which is collected by the ultra-wideband pulse radar to describe the complex form of the received radio wave. Therefore, the IQ component is used to recover the amplitude and phase sequences that reflect the respiratory movement. To accurately sense the respiratory movement, we calculate the signal-to-noise ratio (SNR) of the two sequences, respectively, and then select the sequence with the larger SNR as the respiratory waveform. Specifically, based on empirical physiological knowledge, the respiratory frequency of both healthy subjects and patients with sleep apnea is typically within the range of 0-4 Hz. Upon this, we calculate the signal-to-noise ratio (SNR) of both the amplitude sequence and the phase sequence, where the in-band energy (0-4 Hz) is considered as signal power and the out-of-band energy is considered as noise power. We then compare the SNR values of two sequences and select the one with higher SNR as the final respiratory waveform for subsequent analysis.

Finally, we totally cut 6179 one-minute respiratory samples from 17 subjects to form the ROSA dataset, where the number of samples in the head class is 7.83 times that of the tail class. Based on the ROSA, we construct the graph, termed the ROSA-Graph, where each sample is regarded as a node, with edges defined according to the similarity of waveform features. The node attributes correspond to the radar-captured respiratory signals, which reflect characteristic patterns of sleep apnea data.

4.1.2. Other publicly available datasets

Since there are currently no publicly available datasets for sleep apnea, we further verify the performance of GA-GCC on other publicly available datasets with inherent class imbalance, including Amazon-Photo and Amazon-Computers [41]. These datasets exhibit varying degrees of long-tailed distribution. Specifically, the number of samples in the head class of the Amazon-Photo and Amazon-Computers datasets exceeds that in the tail class by 5.86 and 17.73 times, respectively. These datasets also allow us to evaluate the generalization ability of the proposed method.

4.2. Experimental setup

During the training process, a GCN encoder is trained via the self-supervised contrastive learning paradigm. For the public dataset, we adopt the public split criteria, that is, the data are randomly split into 10%, 10% and 80%, serving as training, validation, and test sets, respectively. In terms of the ROSA dataset, due to an obvious long-tailed distribution of the apnea dataset, random partition could lead to significant performance differences. Therefore, we employ five-fold cross-validation to ensure a more reliable performance evaluation. For public datasets, a ℓ_2-regularized logistic classifier is trained on the training set. We repeat the training for 20 runs and take the mean and variance. Following LocWGCL, we train the encoder with Adam SGD optimizer for all datasets. The weight decay factor is set to $1e\text{-}5$ while the dropout rate is zero. For all three datasets, the learning rate, the temperature hyperparameter τ, and the weight update frequency are set to 0.005, 0.2, and 400, respectively. To highlight the advantages of GA-GCC on class imbalance problem, the Macro F1-score is utilized to measure the performance. In addition, the Amazon-Photo and Amazon-Computers datasets can be used for transductive node classification tasks, therefore, the classification accuracy is also adopted. All experiments are conducted with PyTorch on the NVIDIA GeForce RTX 3090 and Intel(R) Xeon(R) W-2133 CPU @ 3.60GHz.

4.3. Results and analysis on ROSA

4.3.1. Experimental comparison results

To verify the effectiveness of GA-GCC, we compare it with state-of-the-art graph contrastive learning (GCL) methods on the ROSA dataset, including GRACE [33], GCA [35], ProGCL [42], HNGCL [43], and LocWGCL [20]. Since the ROSA dataset is a self-built radar-based

Table 1
Comparison of node classification accuracy (%) and Macro F1-score (%) on the ROSA dataset. "Gain" represents the improvement of GA-GCC over each baseline method and is not applicable to GA-GCC itself. '-' indicates performance degradation relative to GA-GCC. The best results are reported with **boldface** for the accuracy and Macro F1-score.

Method	Acc (%) ↑	Macro F1 (%) ↑	Acc-Gain	F1-Gain
GRACE [33]	94.52±0.91	89.84±1.30	−1.98	−4.75
GCA [35]	94.97±1.02	91.63±1.14	−1.53	−2.96
ProGCL-Weight [42]	95.04±1.02	91.62±1.69	−1.46	−2.97
ProGCL-Mix [42]	95.13±1.15	91.84±2.03	−1.37	−2.75
HNGCL [43]	96.17±1.73	93.96±1.36	−0.33	−0.63
LocWGCL [20]	96.15±1.77	94.05±3.35	−0.35	−0.54
GA-GCC (Ours)	**96.50±1.92**	**94.59±3.92**	*	*

Table 2
Comparison of node classification accuracy (%) and Macro F1-score (%) on the public imbalanced graph datasets: Amazon-Photo and Amazon-Computers. The best results are reported with **boldface**.

Method	Amazon-Photo		Amazon-Computers	
	Acc (%) ↑	Macro F1 (%) ↑	Acc (%) ↑	Macro F1 (%) ↑
GRACE [33]	92.40±0.11	91.62±0.24	87.34±0.18	85.13±0.21
GCA [35]	92.55±0.03	91.68±0.15	87.82±0.11	85.77±0.18
ProGCL-Weight [42]	93.30±0.09	91.79±0.23	89.28±0.15	86.30±0.24
ProGCL-Mix [42]	93.64±0.13	91.94±0.16	89.55±0.16	86.75±0.19
HNGCL [43]	93.87±0.07	92.21±0.11	89.70±0.06	87.01±0.14
LocWGCL [20]	93.95±0.06	92.60±0.17	**89.97±0.09**	88.44±0.17
GA-GCC (Ours)	**94.04±0.12**	**93.01±0.19**	89.91±0.14	**88.82±0.22**

obstructive sleep apnea dataset, existing GCL methods have not reported results on ROSA. For a fair comparison, we reproduce these methods under the same experimental environment as GA-GCC and report the experimental results in Table 1. Then, we summarize the following conclusions:

(1) From the perspective of the dataset, we observe that all methods achieve over 94% on Acc and 89% on Macro F1-score. This phenomenon evidences that the ROSA dataset is of high quality and holds significant value for the study of sleep apnea.

(2) From the perspective of the graph structure, all methods that take ROSA-Graph as input exhibit strong performance. The lowest accuracy (94.52%) and Macro F1-score (89.84%) are both at high levels, indicating that the constructed graph based on ROSA is effective and enables these methods to learn accurate node representations.

(3) From the perspective of performance, GA-GCC achieves the best overall performance. Especially for the Macro F1-score, GA-GCC improves by at most 4.75%, which highlights the advantages of GA-GCC in solving the class imbalance problem. We suggest that both the structure of homophilous graph and the globally balanced weight strategy can enhance the focus on tail classes to improve the performance.

4.3.2. *Visualization analysis on ROSA*

We further visualize the constructed graph to demonstrate the essential role of the homophilous graph in alleviating the class imbalance. As illustrated in Figs. 5–7, we observe that in the 1-hop subgraph, most first-order nodes are the same class as the anchor, especially for the tail class (class 2). Therefore, in Fig. 7, the subgraph consisting of most tail class samples is easier to identify, and the message passing of this subgraph can naturally enhance the representation of these samples.

Furthermore, we also mark the prediction class of each node (respiratory sample) in Figs. 5–7. Even though there are multiple classes in the subgraph, as shown in Figs. 6 and 7, GA-GCC can still accurately predict the class of tail samples. Especially in Fig. 6, the first-order neighborhood of anchor 349 indeed contains a small number of samples from different classes. This phenomenon is a normal and expected characteristic of homophilous graphs, especially in real-world data where noise and overlapping physiological states may exist. Although nodes tend to connect to others with the same label in a homophilous graph, this tendency is statistical rather than absolute. Importantly, most neighboring nodes still share the same class as the anchor, which preserves the overall homophilous property of the constructed graph. Even in this case, the model is still able to accurately identify the class of first-order nodes. This not only highlights the rationality of exploring class imbalance from the perspective of data structure, but also demonstrates the effectiveness of the globally balanced weight strategy of GA-GCC.

4.4. *Results on other publicly available datasets*

In order to improve the generalization of GA-GCC, we further evaluate GA-GCC on two public graph datasets with long-tail distributions: Amazon-Photo and Amazon-Computers. We compare GA-GCC against the GCL methods in Section 4.3.1. These methods are also designed for transductive node classification tasks, for a fair comparison, we directly adopt their reported classification accuracy from the original papers. However, as the Macro F1-score is not provided, we reproduce these algorithms to report.

As results reported in Table 2, GA-GCC achieves the highest Macro F1-score, demonstrating its effectiveness in handling class imbalance. For the Amazon Computers dataset, GA-GCC does not achieve the best node classification accuracyc, which is attributed to the fact that our model mainly emphasizes solving the class imbalance problem. Despite this, these results are sufficient to prove the generalization.

In particular, methods targeting false negatives (ProGCL, HNGCL, LocWGCL) exhibit superior performance compared to other baselines. In contrastive learning, they reduce the adverse impact of false negatives that are actually close to the anchor under the ground truth. Although these methods are designed for node classification, they also help mitigate class imbalance.

Table 3
Ablation study on FreGSC, LocWGCL, and GloCWO modules for node classification accuracy (%) and Macro F1-score (%) indicators.

Case	Metric	ROSA	Amazon Photo	Amazon Computers
GA-GCC	Acc ↑	96.50	94.04	89.91
	Macro F1-score ↑	94.59	93.01	88.82
w/o LocWGCL	Acc ↑	96.09	93.89	89.52
	Macro F1-score ↑	93.64	92.85	88.47
w/o FreGSC	Acc ↑	96.12	93.77	89.82
	Macro F1-score ↑	93.73	92.64	88.42
w/o GloCWO	Acc ↑	95.93	93.81	89.70
	Macro F1-score ↑	93.56	92.82	88.36

4.5. *Ablation study*

Impact of LocWGCL Module: We study the impact of the LocWGCL weight strategy in GA-GCC. As shown in Table 3, when the LocWGCL module is eliminated, GA-GCC *w/o* LocWGCL has a lower performance than GA-GCC (93.64% vs. 94.59% for Macro F1-score on the ROSA dataset). This implies that LocWGCL can accurately identify false negatives in local neighborhoods and assign lower weights to the corresponding negative pairs.

Impact of FreGSC Module: The key insight of the frequency-weighted graph semantic clustering (FreGSC) module lies in the focus on samples of tail classes in the negative sample pairs, while weakening the adverse effect of false negatives. Hence, FreGSC is essential to alleviate class imbalance from a global perspective. To prove this, we remove the FreGSC module and name it GA-GCC *w/o* FreGSC. The results are reported in Table 3. We observe that the performance of GA-GCC *w/o* FreGSC on the three datasets has different degrees of degeneration. This trend high-

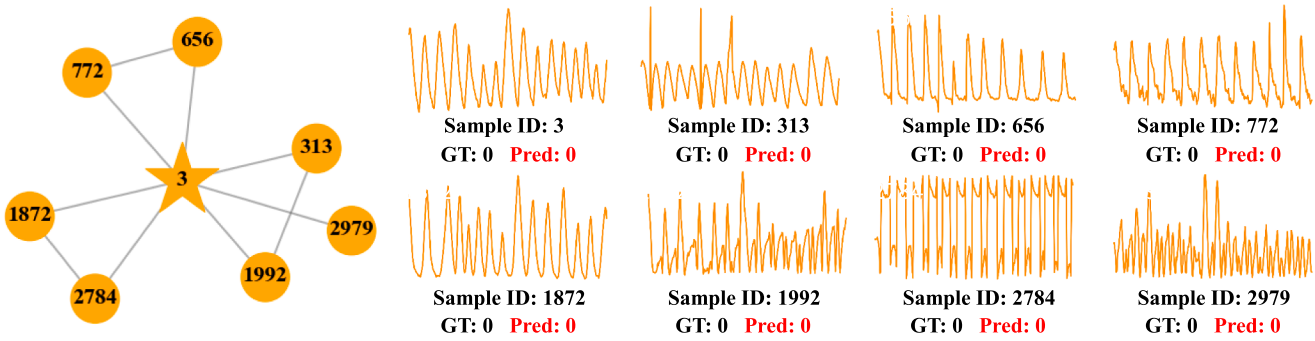

Fig. 5. Visualization of the 1-hop subgraph of anchor '3'. ★: the anchor of class 0. ●: the node of class 0. **GT**: ground truth. **Pred**: predicted class.

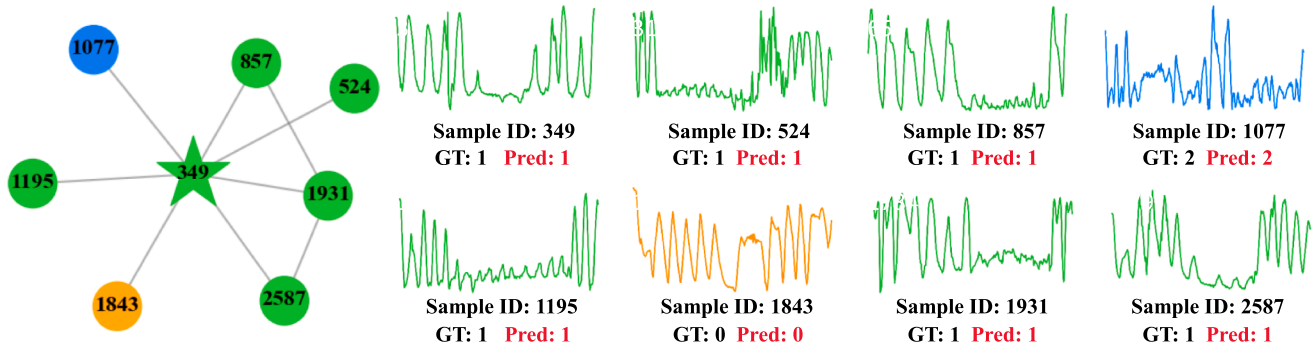

Fig. 6. Visualization of the 1-hop subgraph of anchor '349'. ★: the anchor of class 1. ●: the node of class 0. ●: the node of class 1. ●: the node of class 2. **GT**: ground truth. **Pred**: predicted class.

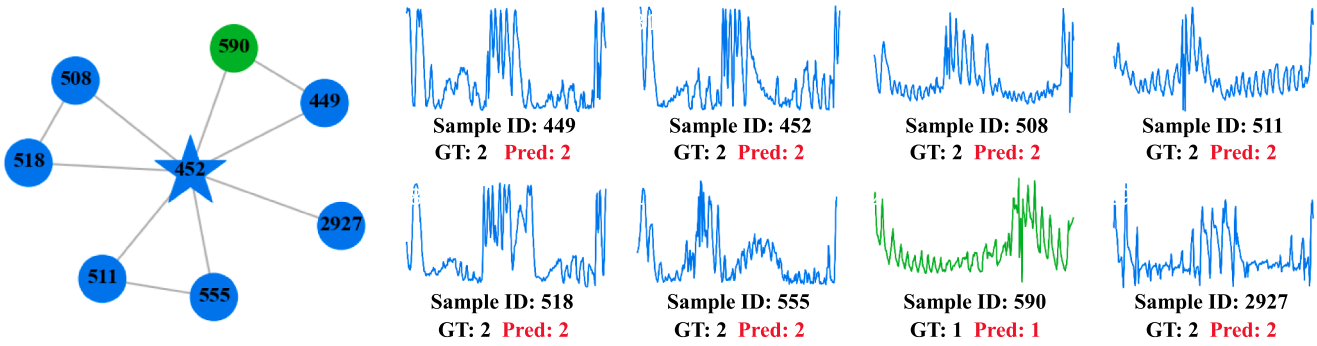

Fig. 7. Visualization of the 1-hop subgraph of anchor '452'. ★: the anchor of class 2. ●: the node of class 1. ●: the node of class 2. **GT**: ground truth. **Pred**: predicted class.

lights the essential role of the FreGSC, which focuses on higher-order nodes that are more numerous and not of the same class as anchors.

Impact of GloCWO Module: To validate the effectiveness of globally class-aware weight optimization module (GloCWO), we perform the ablation study by removing clustering loss ℓ_{clu}. As observed in Table 3, there is the largest performance gap between GA-GCC and GA-GCC *w/o* GloCWO (94.59% vs. 93.56% for Macro F1-score) on the ROSA dataset. This proves that our GloCWO can effectively model the semantic relationship between indirect connected nodes, thus increasing the representation similarity of samples in the same cluster and improving the accuracy of clustering.

Table 4
Node classification performance on the ROSA dataset when selecting different Top-K neighbors to construct K-nearest neighbor graphs. Metrics include classification accuracy (%) and Macro F1-score (%).

Top-K Settings when Constructing Graph	Average Degree	Acc (%) ↑	Macro F1-score (%) ↑
Top-5	7.87	96.50 ± 1.92	94.59 ± 3.92
Top-10	15.44	95.83 ± 1.12	92.89 ± 2.40
Top-15	22.92	95.23 ± 1.03	91.44 ± 2.00
Top-20	30.36	94.24 ± 1.03	90.09 ± 1.54
Top-25	37.77	94.09 ± 1.13	89.69 ± 1.92
Top-30	45.12	93.99 ± 0.68	89.46 ± 1.95

4.6. Hyperparameter analysis

Impact of the Average Degree of Graph: To study the impact of the graph structure, we also analyze how the average degree that reflects the density of the graph affects the model performance. Generally, when constructing the K-nearest neighbor graph, the larger the value of K, the higher the average degree of the constructed graph. The results of different graphs are reported in Table 4. We observe a clear downward

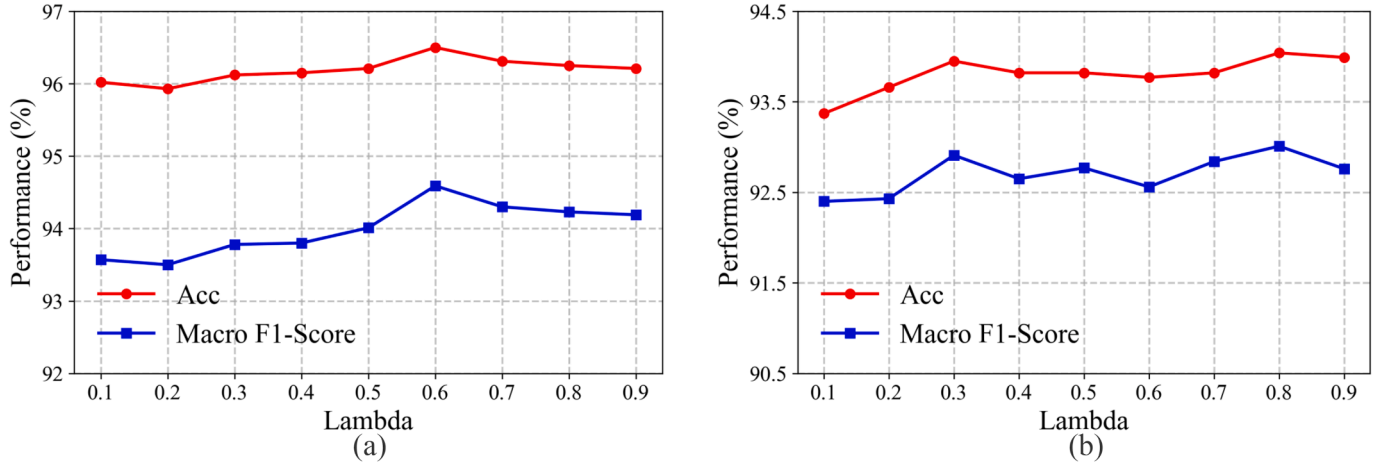

Fig. 8. The hyperparameter analysis of λ on ROSA (a) and Amazon-Photo (b) datasets.

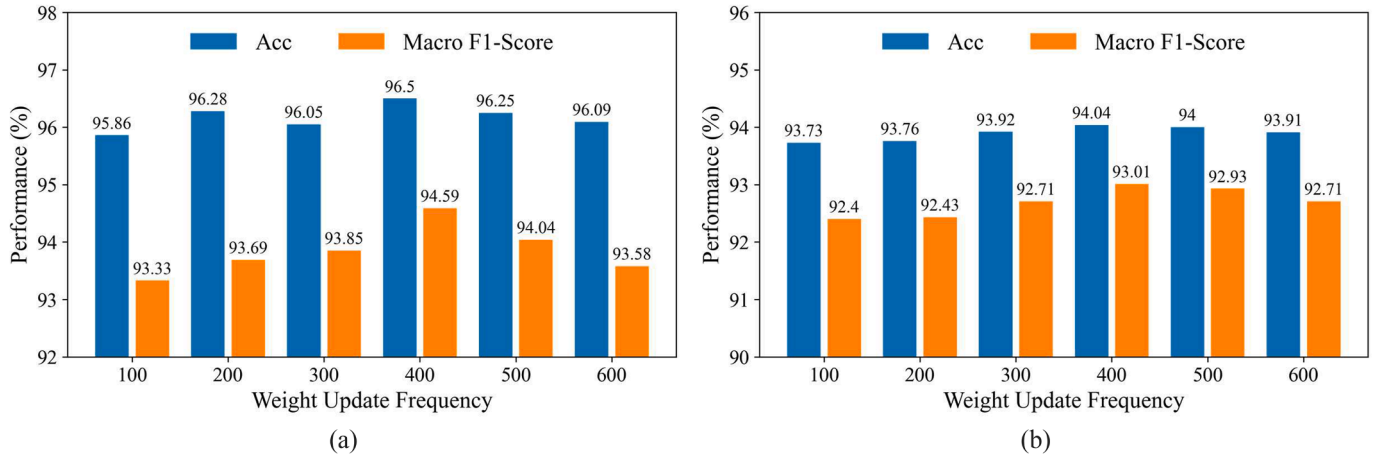

Fig. 9. The hyperparameter analysis of weight update frequency on Apnea (a) and Amazon-Photo (b) datasets.

trend in performance as the average degree increases. This degradation can be attributed to the fact that excessively high degrees can lead to spurious structural relationships in the graph, which is particularly problematic for sleep apnea data that contain only three classes and exhibit a severely long-tailed distribution.

Impact of the Tradeoff Hyperparameter: We study the impact of the tradeoff hyperparameter λ, which is adopted to balance the importance of ℓ_{wgcl} and ℓ_{clu}. Specifically, as shown in Fig. 8, we vary λ in the range of 0.1 to 0.9 on the ROSA and Amazon-Photo datasets. It is worth noting that the fluctuation of performance is relatively remarkable with the change of λ, which highlight the significant role of the tradeoff hyperparameter. In particular, the best Acc and Macro F1-score are obtained when λ is set to 0.6 and 0.8 on the ROSA and Amazon-Photo datasets, respectively.

Finally, the impact of weight update frequency is discussed. Since the augmented views are regenerated for each epoch, the weight update frequency is essential to dynamically identify the samples of tail classes and to emphasize attention to these samples. To this end, the weight update frequency is set as $\{100, 200, 300, 400, 500, 600\}$. As illustrated in Fig. 9, the best performance occurs when the update frequency is 400. We suggest that too large a frequency is not conducive to accurately seeking the tail category, while too small a frequency does not allow the learned weight to fully play its role within a certain number of epochs.

5. Conclusion

In this paper, we release the first radar-based obstructive sleep apnea (ROSA) dataset, composed of 6179 respiratory samples, to boost multimodal affective research. To address the class imbalance problem inherent in sleep apnea prediction, we propose a Class-Aware Graph Contrastive Clustering (GA-GCC) method via adaptive weights. We first construct a homophilous graph with dynamic time warping (DTW) to make the tail classes easier to discover and to enhance the representation of their nodes. Then, locally weighted graph contrastive learning (LocWGCL) is introduced to identify false negatives on lower-order neighborhoods. Next, we develop a clustering loss from the perspective of semantic relationships among nodes, improving the quality of pseudo-labels. Based on pseudo-labels, a frequency-weighted strategy is designed to highlight tail classes globally. Finally, we integrate frequency-based weights and local weights to balance the head and tail classes and eliminate the adverse impact of false negatives. Extensive experiments on sleep apnea dataset ROSA and other public datasets verify the effectiveness and of generalization GA-GCC.

Combining OSA research with affective computing offers a promising approach to jointly modeling physiological signals and emotional responses. Accurate OSA predictions can not only analyze sleep physiological indicators related to emotional states but also support a refined assessment of affective dynamics. This interdisciplinary integration opens

new opportunities for personalized mental health interventions, bridging sleep medicine and affective computing. In the future, we will study more robust model to improve the reliability of non-contact sleep apnea prediction.

CRediT authorship contribution statement

Xin Liu: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Xinke Wang:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Jinyang Huang:** Writing – review & editing, Supervision, Methodology, Formal analysis; **Dan Guo:** Writing – review & editing, Validation, Supervision, Resources, Investigation, Funding acquisition, Conceptualization; **Meng Wang:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is supported by National Natural Science Foundation of China (72188101, 62272144, 62020106007), National Key R&D Program of China (NO. 2024YFB3311600), the Major Project of Anhui Province (202203a05020011), the Anhui Provincial Natural Science Foundation (2408085J040), and the Fundamental Research Funds for the Central Universities (JZ2023YQTD0072, JZ2024HGTG0309, JZ2024AHST0337), and the New Cornerstone Science Foundation through the XPLORER PRIZE.

References

[1] X. Qiu, Y. Wei, X. Tan, W. Xu, H. Wang, J. Ma, J. Huang, Z. Fang, MIMAR-OSA: enhancing obstructive sleep apnea diagnosis through multimodal data integration and missing modality reconstruction, Pattern Recognit. 169 (2026) 111917.
[2] H.-L. Chan, H.-M. Yuen, C.-T. Au, K.C.-C. Chan, A.M. Li, L.-M. Lui, Classification of childhood obstructive sleep apnea based on X-ray images analysis by quasi-conformal geometry, Pattern Recognit. 152 (2024) 110454.
[3] J. Gu, K. Li, F. Wang, Y. Wei, Z. Wu, H. Fan, M. Wang, Motion matters: motion-guided modulation network for skeleton-based micro-action recognition, in: Proceedings of the 33rd ACM International Conference on Multimedia, 2025, pp. 5461–5470.
[4] K. Li, P. Liu, D. Guo, F. Wang, Z. Wu, H. Fan, M. Wang, MMAD: multi-label micro-action detection in videos, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2025, pp. 13225–13236.
[5] Y. Hong, C. Pei, L. Hao, K. Xu, F. Liu, Z. Ding, The study of the relationship between moderate to severe sleep obstructive apnea and cognitive impairment, anxiety, and depression, Front. Neurol. 15 (2024) 1363005.
[6] Y. Hai, W. Kou, Z. Gu, C. Zhang, Q. Zou, F. Wang, H. Yao, P. Wei, Obstructive sleep apnea affects the psychological and behavioural development of children–a case–control study, J. Sleep Res. 33 (1) (2024) e13924.
[7] S. Zhao, X. Yao, J. Yang, G. Jia, G. Ding, T.-S. Chua, B.W. Schuller, K. Keutzer, Affective image content analysis: two decades review and new perspectives, IEEE Trans. Pattern Anal. Mach. Intell. 44 (10) (2021) 6729–6751.
[8] K. Li, D. Guo, G. Chen, C. Fan, J. Xu, Z. Wu, H. Fan, M. Wang, Prototypical calibrating ambiguous samples for micro-action recognition, in: Proceedings of the AAAI Conference on Artificial Intelligence, 39, 2025, pp. 4815–4823.
[9] S. Zhao, X. Hong, J. Yang, Y. Zhao, G. Ding, Toward label-efficient emotion and sentiment analysis, Proc. IEEE 111 (10) (2023) 1159–1197.
[10] S. Zhao, H. Yao, Y. Gao, R. Ji, G. Ding, Continuous probability distribution prediction of image emotions via multitask shared sparse regression, IEEE Trans. Multimedia 19 (3) (2016) 632–645.
[11] J. Zhao, F. Wang, K. Li, Y. Wei, S. Tang, S. Zhao, X. Sun, Temporal-frequency state space duality: an efficient paradigm for speech emotion recognition, in: ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 2025, pp. 1–5.
[12] J. Zhu, Z. Wang, J. Chen, Y.-P.P. Chen, Y.-G. Jiang, Balanced contrastive learning for long-tailed visual recognition, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 6908–6917.
[13] C. Hou, J. Zhang, H. Wang, T. Zhou, Subclass-balancing contrastive learning for long-tailed recognition, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 5395–5407.
[14] Z. Ning, C. Guo, D. Zhang, Meta-distribution-based ensemble sampler for imbalanced semi-supervised learning, Pattern Recognit. 164 (2025) 111552.
[15] Z. Huang, Z. Chen, B. Dong, C. Liang, E. Zhou, W. Zuo, IMWA: Iterative model weight averaging benefits class-imbalanced learning, Pattern Recognit. 161 (2025) 111293.
[16] T.N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, in: International Conference on Learning Representations, 2017.
[17] X. Chang, F. Tung, G. Mori, Learning discriminative prototypes with dynamic time warping, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 8395–8404.
[18] H. Deng, W. Chen, Q. Shen, A.J. Ma, P.C. Yuen, G. Feng, Invariant subspace learning for time series data based on dynamic time warping distance, Pattern Recognit. 102 (2020) 107210.
[19] Y. Zhang, J. Chen, K. Wang, F. Xie, ECL: class-enhancement contrastive learning for long-tailed skin lesion classification, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2023, pp. 244–254.
[20] X. Liu, B. Qian, H. Liu, D. Guo, Y. Wang, M. Wang, Seeking false hard negatives for graph contrastive learning, IEEE Trans. Circuits Syst. Video Technol. 34 (8) (2024) 7454–7466.
[21] Y. Liu, X. Gao, T. He, T. Zheng, J. Zhao, H. Yin, Reliable node similarity matrix guided contrastive graph clustering, IEEE Trans. Knowl. Data Eng. 36 (12) (2024) 9123–9135.
[22] A.S. BaHammam, A.R. Pirzada, S.R. Pandi-Perumal, Neurocognitive, mood changes, and sleepiness in patients with REM-predominant obstructive sleep apnea, Sleep Breathing 27 (1) (2023) 57–66.
[23] W. Qian, K. Li, D. Guo, B. Hu, M. Wang, Cluster-phys: facial clues clustering towards efficient remote physiological measurement, in: Proceedings of the 32nd ACM International Conference on Multimedia, 2024, pp. 330–339.
[24] A.C. Bitners, R. Arens, Evaluation and management of children with obstructive sleep apnea syndrome, Lung 198 (2) (2020) 257–270.
[25] H.T. Chaw, S. Kamolphiwong, K. Wongsritrang, Sleep apnea detection using deep learning, Tehnički Glasnik 13 (4) (2019) 261–266.
[26] P. Edouard, D. Campo, P. Bartet, R.-Y. Yang, M. Bruyneel, G. Roisman, P. Escourrou, Validation of the withings sleep analyzer, an under-the-mattress device for the detection of moderate-severe sleep apnea syndrome, Journal of Clinical Sleep Medicine 17 (6) (2021) 1217–1227.
[27] S. Li, B. Jin, Z. Wang, F. Zhang, X. Ren, H. Liu, Leveraging attention-reinforced UWB signals to monitor respiration during sleep, ACM Trans. Sens. Netw. 20 (5) (2024) 1–28.
[28] Y. Marrakchi, O. Makansi, T. Brox, Fighting class imbalance with contrastive learning, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2021, pp. 466–476.
[29] D. Guo, K. Li, B. Hu, Y. Zhang, M. Wang, Benchmarking micro-action recognition: dataset, methods, and applications, IEEE Trans. Circuits Syst. Video Technol. 34 (7) (2024) 6238–6252.
[30] Y. Qian, C. Zhang, Y. Zhang, Q. Wen, Y. Ye, C. Zhang, Co-modality graph contrastive learning for imbalanced node classification, Adv. Neural Inf. Process. Syst. 35 (2022) 15862–15874.
[31] L. Qian, Q. Zuo, D. Li, H. Zhu, DGMSCL: A dynamic graph mixed supervised contrastive learning approach for class imbalanced multivariate time series classification, Neural Netw. 185 (2025) 107131.
[32] P. Veličković, W. Fedus, W.L. Hamilton, P. Liò, Y. Bengio, R.D. Hjelm, Deep graph Infomax, in: International Conference on Learning Representations, 2019.
[33] Y. Zhu, Y. Xu, F. Yu, Q. Liu, S. Wu, L. Wang, Deep graph contrastive representation learning,(2020). arXiv preprint arXiv:2006.04131
[34] T. Chen, S. Kornblith, M. Norouzi, G. Hinton, A simple framework for contrastive learning of visual representations, in: International Conference on Machine Learning, 119, PMLR, 2020, pp. 1597–1607.
[35] Y. Zhu, Y. Xu, F. Yu, Q. Liu, S. Wu, L. Wang, Graph contrastive learning with adaptive augmentation, in: Proceedings of the Web Conference 2021, 2021, pp. 2069–2080.
[36] Z. Tong, Y. Liang, H. Ding, Y. Dai, X. Li, C. Wang, Directed graph contrastive learning, Adv. Neural Inf. Process. Syst. 34 (2021) 19580–19593.
[37] J. Wu, X. Chen, B. Shi, S. Li, K. Xu, SEGA: structural entropy guided anchor view for graph contrastive learning, in: International Conference on Machine Learning, 202, PMLR, 2023, pp. 37293–37312.
[38] C. Wei, Y. Wang, B. Bai, K. Ni, D. Brady, L. Fang, Boosting graph contrastive learning via graph contrastive saliency, in: International Conference on Machine Learning, 202, PMLR, 2023, pp. 36839–36855.
[39] Y. Liu, X. Yang, S. Zhou, X. Liu, S. Wang, K. Liang, W. Tu, L. Li, Simple contrastive graph clustering, IEEE Trans. Neural Netw. Learn. Syst. 35 (10) (2023) 13789–13800.

[40] X. Xie, B. Li, E. Pan, Z. Guo, Z. Kang, W. Chen, One node one model: featuring the missing-half for graph clustering, in: Proceedings of the AAAI Conference on Artificial Intelligence, 39, 2025, pp. 21688–21696.
[41] O. Shchur, M. Mumme, A. Bojchevski, S. Günnemann, Pitfalls of graph neural network evaluation, (2018). arXiv preprint arXiv:1811.05868
[42] J. Xia, L. Wu, G. Wang, S.Z. Li, ProGCL: rethinking hard negative mining in graph contrastive learning, in: International Conference on Machine Learning, 162, PMLR, 2022, pp. 24332–24346.
[43] S. Wang, H. Yan, J. Du, J. Yin, J. Zhu, C. Li, J. Wang, Adversarial hard negative generation for complementary graph contrastive learning, in: Proceedings of the 2023 SIAM International Conference on Data Mining (SDM), SIAM, 2023, pp. 163–171.