



PDF Download
3718734.pdf
11 January 2026
Total Citations: 13
Total Downloads:
3174

 Latest updates: <https://dl.acm.org/doi/10.1145/3718734>

SURVEY

Class-Imbalanced Learning on Graphs: A Survey

YIHONG MA, University of Notre Dame, Notre Dame, IN, United States

YIJUN TIAN, University of Notre Dame, Notre Dame, IN, United States

NUNO MONIZ, University of Notre Dame, Notre Dame, IN, United States

NITESH V. CHAWLA, University of Notre Dame, Notre Dame, IN, United States

Published: 22 March 2025
Online AM: 19 February 2025
Accepted: 03 February 2025
Revised: 20 January 2025
Received: 13 April 2023

[Citation in BibTeX format](#)

Open Access Support provided by:

University of Notre Dame

Class-Imbalanced Learning on Graphs: A Survey

YIHONG MA, University of Notre Dame, Notre Dame, United States

YIJUN TIAN, University of Notre Dame, Notre Dame, United States

NUNO MONIZ, University of Notre Dame, Notre Dame, United States

NITESH V. CHAWLA, University of Notre Dame, Notre Dame, United States

Rapid advancement in machine learning is increasing the demand for effective graph data analysis. However, real-world graph data often exhibits class imbalance, leading to poor performance of standard machine learning models on underrepresented classes. To address this, Class-Imbalanced Learning on Graphs (CILG) has emerged as a promising solution that combines graph representation learning and class-imbalanced learning. This survey provides a comprehensive understanding of CILG's current state-of-the-art, establishing the first systematic taxonomy of existing work and its connections to traditional imbalanced learning. We critically analyze recent advances and discuss key open problems. A continuously updated reading list of relevant articles and code implementations is available at <https://github.com/yihongma/CILG-Papers>.

CCS Concepts: • **General and reference** → **Surveys and overviews**; • **Computing methodologies** → **Machine learning**; • **Mathematics of computing** → **Graph algorithms**;

Additional Key Words and Phrases: Class imbalance, graph representation learning, graph neural networks

ACM Reference Format:

Yihong Ma, Yijun Tian, Nuno Moniz, and Nitesh V. Chawla. 2025. Class-Imbalanced Learning on Graphs: A Survey. *ACM Comput. Surv.* 57, 8, Article 207 (March 2025), 16 pages. <https://doi.org/10.1145/3718734>

1 Introduction

Graphs are a powerful and prevalent data structure for representing complex relational systems, such as social networks, citation networks, and knowledge graphs. In these systems, nodes represent entities, while edges represent their relationships. In recent years, graph representation learning techniques have proven effective in discovering meaningful vector representations of nodes or the entire graph, resulting in successful applications across a wide range of downstream tasks [49, 51, 68]. Despite these advances, graph data often presents a significant challenge in the form of class imbalance, where instances of one class significantly outnumber those of other classes. When left unaddressed, such class imbalance can significantly degrade model performance, particularly for minority classes that often represent the most critical cases in real-world applications.

Class-Imbalanced Learning on Graphs (CILG) is an emerging research area addressing class imbalance in graph data, where traditional methods for non-graph data might be unsuitable or ineffective for several reasons. Firstly, the unique, irregular, and non-Euclidean structure of graph

Authors' Contact Information: Yihong Ma, University of Notre Dame, Notre Dame, Indiana, United States; e-mail: yma5@nd.edu; Yijun Tian, University of Notre Dame, Notre Dame, Indiana, United States; e-mail: yijun.tian@nd.edu; Nuno Moniz, University of Notre Dame, Notre Dame, Indiana, United States; e-mail: nuno.moniz@nd.edu; Nitesh V. Chawla, University of Notre Dame, Notre Dame, Indiana, United States; e-mail: nchawla@nd.edu.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2025 Copyright held by the owner/author(s).

ACM 0360-0300/2025/03-ART207

<https://doi.org/10.1145/3718734>

data complicates traditional class-imbalanced techniques designed for Euclidean data [76]. Furthermore, graph data often contain rich relational information, requiring specialized techniques to preserve and leverage such information during the learning process [53]. Finally, node dependencies and interactions in a graph make class rebalancing techniques complex, as naive oversampling or undersampling methods may disrupt the graph's structure and thus lead to poor performance [42].

Although class-imbalanced learning has already attracted significant attention in areas such as computer vision [73] and natural language processing [21], its application to graph data remains relatively underexplored. This survey addresses this gap by providing the first comprehensive overview of CILG. We introduce the background of CILG, systematically categorize existing CILG techniques under our proposed taxonomies, and present a thorough evaluation framework that includes benchmark datasets and a critical assessment of current performance metrics.

In addition to reviewing existing work, this survey identifies three critical future research directions to advance CILG. First, we explore opportunities beyond node classification, examining both other node-level tasks and graph-level tasks where class imbalance affects the entire graph structure. Second, we discuss the application of CILG to more complex graph types beyond traditional homogeneous and homophilic settings, where existing methods may fall short. Third, we investigate topology imbalance as a fundamental challenge distinct from class imbalance, considering how factors such as positional distributions of nodes, subclass-level topological motifs, and hierarchical structures in graphs can impact graph learning. These research directions highlight fundamental challenges that must be addressed to develop more comprehensive CILG solutions.

The article is organized as follows: Section 2 introduces the notations for CILG and provides a background on graph representation learning, class-imbalanced learning, and their relationship with CILG. Section 3 details recent CILG methods, including the data level (Section 3.1) and the algorithm level (Section 3.2). Section 4 discusses benchmark datasets, train-test split strategies, and evaluation metrics for CILG. Finally, Section 5 discusses three important future research directions.

In summary, the main contributions of this article include:

- We present the first comprehensive survey on CILG, analyzing recent advances in addressing class imbalance challenges unique to graph data;
- We propose a systematic taxonomy of CILG methods, organizing them into two primary categories with six distinct sub-categories, providing a clear framework for existing solutions;
- We identify three future directions in CILG: (1) exploring tasks beyond node classification at both node and graph levels, (2) extending to heterogeneous or heterophilic graphs, and (3) investigating topology imbalance beyond traditional class imbalance;
- We maintain a continuously updated reading list of relevant articles and code implementations at <https://github.com/yihongma/CILG-Papers>.

2 Background

2.1 Notations

Consider an attributed graph $G = (\mathcal{V}, \mathcal{E})$ with a set of n nodes $\mathcal{V} = \{v_1, \dots, v_n\}$ and a set of m edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ denote the adjacency matrix where $A_{ij} = 1$ if $(v_i, v_j) \in \mathcal{E}$ and 0 otherwise. The node feature matrix $\mathbf{X} \in \mathbb{R}^{n \times d}$ has each row \mathbf{x}_i representing the node feature vector of node v_i with a feature dimension of d .

For node classification, let $\mathbf{y} = \{y_1, \dots, y_n\}$ denote the set of node labels, where $y_i \in \{1, \dots, K\}$ corresponds to the label of node v_i with K classes. The node class distribution $P_k = \frac{|C_k|}{\sum_{i=1}^K |C_i|}$ represents the proportion of nodes in class k , where C_k is the set of nodes in class k .

2.2 Graph Representation Learning

Graph representation learning [74] focuses on discovering meaningful representations of nodes or the entire graph for various downstream tasks. Current methods fall into three categories:

- (1) Network embedding methods [8] that learn node representations by preserving structural proximities. These methods (e.g., DeepWalk [43] and node2vec [17]) can only capture purely structural information, making them less suitable for attributed graphs.
- (2) **Graph neural networks (GNNs)** [63] that learn node representations through iterative neighborhood aggregation operations. By jointly leveraging both graph structure and node features, GNNs have become particularly effective for attributed graphs, leading to their dominance in modern graph learning.
- (3) Knowledge graph embedding methods [56] that learn representations by optimizing scoring functions over fact triplets (head entity, relation, and tail entity). These methods are specifically designed for multi-relational graphs where edges have different types or meanings.

GNNs have emerged as the dominant backbone for CILG methods due to their ability to: (1) capture both structural and feature information, (2) propagate information across the graph through message passing, and (3) naturally incorporate task-specific downstream objectives. GNNs adopt a message-passing mechanism [15], which updates a node's representation by aggregating information from its neighboring nodes and edges. After k iterations of aggregation, the representation of node v_i captures the structural information within its k -hop neighborhood:

$$\mathbf{h}_i^{(k)} = \text{UPDATE} \left(\mathbf{h}_i^{(k-1)}, \text{AGGREGATE}(\{\mathbf{h}_j^{(k-1)} \mid \forall v_j \in \mathcal{N}(v_i)\}) \right), \quad (1)$$

where $\mathbf{h}_i^{(k)}$ is the representation vector of node v_i in the k th GNN layer; $\mathcal{N}(v_i)$ represents the set of neighbors of node v_i . Common choices for $\text{AGGREGATE}(\cdot)$ include mean, max, or sum operations, while $\text{UPDATE}(\cdot)$ typically involves neural networks that combine the aggregated neighborhood information with the node's current representation. The initial representation $\mathbf{h}_i^{(0)}$ is initialized with the node feature vector \mathbf{x}_i , allowing the GNNs to leverage node attributes as well.

2.3 Class-Imbalanced Learning

Class-imbalanced learning [20, 23] addresses the challenge of learning from skewed class distributions, where certain classes (majority classes) have significantly more training instances than others (minority classes). The severity of class imbalance is quantified by an imbalance ratio ρ :

$$\rho = \frac{\max_k |C_k|}{\min_k |C_k|}, \quad (2)$$

where $|C_k|$ denotes the number of samples in class k .

Traditional class-imbalanced learning methods can be categorized into two main approaches: (1) data-level methods that modify the distribution of training data by oversampling minority classes [4], undersampling majority classes [34], or hybrid approaches [47] that combine both; and (2) algorithm-level methods that modify the learning algorithm through cost-sensitive learning [80], ensemble methods [5], or loss function engineering [3, 9, 10, 30].

2.4 CLIG

CILG focuses primarily on node classification tasks where the distribution of node classes is skewed. While CILG can also encompass graph classification [59, 65], this survey focuses on node classification, which has received the majority of research attention. We formally define the node classification task under the class-imbalanced setting:

Table 1. A Comprehensive Summary of Representative CILG Models, Arranged in Chronological Order, where “D.I.” Stands for Data Interpolation (Section 3.1.1), “A.G.” Stands for Adversarial Generation (Section 3.1.2), “P.L.” Stands for Pseudo-Labeling (Section 3.1.3), “M.T.R.” Stands for Model and Training Refinement (Section 3.2.1), “L.E.” Stands for Loss Function Engineering (Section 3.2.2), and “P.A.” Stands for Post-hoc Adjustment (Section 3.2.3)

| Model | Year | Venue | Data-Level | | | Algorithm-Level | | | Key Components | Code |
|-----------------|------|-------------|------------|------|------|-----------------|------|------|--|----------------------|
| | | | D.I. | A.G. | P.L. | M.T.R. | L.E. | P.A. | | |
| RSDNE [60] | 2018 | AAAI | | | ■ | ■ | □ | | DeepWalk + Intra/Inter-class similarity | link |
| SPARC [78] | 2018 | KDD | | | ■ | □ | □ | | Label propagation + Curriculum learning | link |
| ImVerde [61] | 2018 | Big Data | | | | ■ | | | DeepWalk + Balanced mini-batch sampling | link |
| DR-GCN [52] | 2020 | IJCAI | | ■ | | □ | □ | | GAN + Distribution alignment (labeled/unlabeled) | link |
| GraphSMOTE [76] | 2021 | WSDM | ■ | | | □ | □ | | SMOTE + Pre-trained edge generator | link |
| GraphENS [42] | 2021 | ICLR | ■ | | | □ | | | Ego network Mixup + Saliency-based feature filtering | link |
| ImGAGN [46] | 2021 | KDD | | ■ | | | | | GAN + Class-conditional generation | link |
| ReNode [6] | 2021 | NeurIPS | | | | | ■ | | Topology-aware reweighting | link |
| SET-GNN [24] | 2021 | ICONIP | | | ■ | | | | Label propagation + Noise-aware self-training | – |
| TAM [53] | 2022 | ICLR | | | | | ■ | | Topology-aware margin adjustment | link |
| ALLIE [7] | 2022 | WWW | | | | ■ | □ | □ | Active learning + RL-based sampling | – |
| GATSMOTE [37] | 2022 | Mathematics | ■ | | | □ | □ | | SMOTE + Homophily-guided attention edge generator | – |
| ACS-GNN [39] | 2022 | ICNSC | | | | ■ | □ | | Attention message passing + Class-aware reweighting | – |
| EGCN [55] | 2022 | ICNSC | | | | ■ | | | Class-aware message passing | – |
| KINC-GCN [1] | 2022 | ICNSC | | | | ■ | □ | | Kernel-based message passing + Node clustering | – |
| FACS-GCN [50] | 2022 | IJCNN | | □ | | □ | ■ | | Class-aware reweighting + GAN | link |
| DPGNN [58] | 2022 | MLG | | | ■ | □ | | | Prototype label propagation + Distance metric learning | link |
| LTE4G [69] | 2022 | CIKM | | | | ■ | | □ | Knowledge distillation + Class prototype inference | link |
| GraphMixup [62] | 2022 | ECML PKDD | ■ | | | □ | □ | | Semantic-space Mixup + RL-based sampling | link |
| SORAG [11] | 2022 | ECML PKDD | | ■ | | | □ | | Dual-generator synthesis (labeled/unlabeled) | – |
| CM-GCL [45] | 2022 | NeurIPS | | | | ■ | □ | | Contrastive learning + Topology-aware pruning | link |
| GraphSR [79] | 2023 | AAAI | | | ■ | □ | | | Label propagation + RL-based sampling | – |
| ImGCL [70] | 2023 | AAAI | | | | ■ | □ | | Contrastive learning + Topology-aware sampling | – |
| BNE [81] | 2023 | Info. Sci. | | | | ■ | | | Balanced neighbor influence exploration | – |
| INS-GNN [25] | 2023 | Info. Sci. | | | ■ | □ | □ | | Self-supervised pre-training + Self-training | link |
| GraphSANN [33] | 2023 | IJCAI | ■ | | | □ | □ | | Homophilic/Heterophilic feature Mixup | – |
| DataDec [71] | 2023 | ICML | | | | ■ | □ | | Contrastive learning + Dynamic sparse training | link |
| GraphSHA [27] | 2023 | KDD | ■ | | | □ | | | Mixup + Neighbor class sampling | link |
| SNS [14] | 2023 | CIKM | ■ | | | □ | □ | | Semantic-aware sampling + Graph augmentation | link |
| ReVar [66] | 2023 | NeurIPS | | | | □ | ■ | | Variance reduction + Graph augmentation | link |
| GNN-CL [28] | 2024 | Neurocomp. | ■ | | | □ | | | SMOTE + Curriculum learning | – |

■: Primary Category, □: Secondary Category.

Definition 1 (Class-Imbalanced Node Classification). Given a graph G and a set of labeled nodes $\mathcal{V}^\ell \subseteq \mathcal{V}$ with imbalanced class distribution P_k , learn a classifier $f: \mathcal{V} \rightarrow \{1, \dots, K\}$ that accurately predicts labels for unlabeled nodes $\mathcal{V}^u = \mathcal{V} \setminus \mathcal{V}^\ell$ across all classes.

3 Methods

This section reviews existing CILG approaches, categorizing them into two main groups: (1) data-level methods and (2) algorithm-level methods. Data-level methods are further divided into (i) data interpolation, (ii) adversarial generation, and (iii) pseudo-labeling. Algorithm-level methods are further classified into (i) model and training refinement, (ii) loss function engineering, and (iii) post-hoc adjustment. Table 1 presents a comprehensive summary of representative CILG models.

3.1 Data-Level Methods

Data-level techniques are crucial for addressing class imbalance problems by modifying training data in feature or label spaces to achieve a more balanced learning environment. Basic data-level approaches include oversampling, which increases the number of minority class samples, and undersampling, which decreases the majority class samples. However, new techniques are required to augment training data within the graph structural space due to node dependencies and inter-connections in graph data. In the context of CILG, we further classify data-level methods into data interpolation, adversarial generation, and pseudo-labeling. Each method differs in its approach to generating features, structures, or labels for synthetically created minority data instances.

3.1.1 Data Interpolation. Data interpolation generates synthetic training samples for minority classes through linear combinations of existing samples. The foundational approach is the **Synthetic Minority Over-sampling Technique (SMOTE)** [4], which creates minority training samples $\tilde{\mathbf{x}}$ by interpolating between two training instances from the same minority class:

$$\tilde{\mathbf{x}} = \mathbf{x}_i + \lambda(\mathbf{x}_j - \mathbf{x}_i), \quad (3)$$

where \mathbf{x}_i is a random minority instance, \mathbf{x}_j is one of its k -nearest neighbors from the same minority class, and $\lambda \in [0, 1]$. However, directly applying SMOTE to graphs presents unique challenges, as the interpolation process must also consider the topological structure of graph data. Specifically, after synthesizing new nodes through interpolation, we must determine how these synthetic nodes connect to existing nodes while preserving a meaningful graph structure.

To address this challenge, several graph-specific SMOTE adaptations have been proposed. GraphSMOTE [76] is the first to adapt SMOTE for graphs by performing interpolation in the node embedding space and using a pre-trained edge predictor to determine the connectivity between synthetic and real nodes. GATSMOTE [37] improves this approach by employing a neighborhood-constrained attention mechanism and homophily principles to generate edges that preserve local structures. GNN-CL [28] further advances the field by combining adaptive oversampling with curriculum learning to gradually expose the model to increasingly complex synthetic samples.

Another significant development is the adaptation of Mixup [72] to graphs. Unlike SMOTE's focus on minority classes only, Mixup interpolates across all classes in both feature and label spaces:

$$\tilde{\mathbf{x}} = \lambda \mathbf{x}_i + (1 - \lambda) \mathbf{x}_j, \quad \tilde{y} = \lambda y_i + (1 - \lambda) y_j, \quad (4)$$

where (\mathbf{x}_i, y_i) and (\mathbf{x}_j, y_j) are two random instances, and $\lambda \in [0, 1]$.

Recent works have proposed various innovations in graph-based Mixup. GraphENS [42] addresses the limited neighbor views of minority nodes by synthesizing their ego networks through interpolation between the ego networks of the minority node and a random node. It also uses gradient-based feature saliency to filter out class-specific attributes during mixing. GraphMixup [62] performs Mixup in the semantic space to prevent the generation of out-of-domain samples, incorporates self-supervised objectives to preserve node connections, and uses reinforcement learning to adaptively determine oversampling ratios for minority classes. GraphSHA [27] expands minority class boundaries through Mixup between hard minority samples and neighbor class nodes with selective edge sampling only from the minority node's subgraph. GraphSANN [33] extends Mixup to heterophilic graphs by enabling both homophilic and heterophilic interpolation. For heterogeneous graphs, SNS [14] addresses the semantic imbalance through meta-path based graph decomposition and personalized PageRank to identify influential neighbors for minority nodes.

3.1.2 Adversarial Generation. Adversarial generation methods utilize **generative adversarial networks (GANs)** to create synthetic minority nodes. GANs [16] are a class of neural networks comprising a generator and a discriminator. The generator G aims at creating fake data that closely resemble real data, while the goal of the discriminator D is to differentiate between the real and generated data. GANs are trained using a minimax objective function:

$$\min_G \max_D \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_z} [\log (1 - D(G(\mathbf{z})))], \quad (5)$$

where p_{data} is the distribution of real data \mathbf{x} , and p_z is the distribution of random noise \mathbf{z} .

DR-GCN [52] is the first work to leverage adversarial learning in CILG, using a GAN to generate synthetic nodes and incorporating KL-divergence to align labeled and unlabeled node representations. ImGAGN [46] expands this by introducing a GAN that synthesizes both minority nodes and their connections to the graph, with a shared GCN [26] serving as both a discriminator

to distinguish real nodes from synthetic nodes and a classifier to identify minority class membership. However, ImGAGN is restricted to binary classification, as extending it to multi-class settings would require non-trivial modifications. FACS-GCN [50] integrates adversarial learning with class-adaptive reweighting, where the adversarial component helps generate balanced feature representations while the reweighting mechanism dynamically adjusts the importance of different classes during training. SORAG [11] introduces a dual-generator architecture combining a standard GAN and a conditional GAN [40]: the former generates unlabeled nodes to prevent over-aggregation of majority features in minority regions, while the latter creates labeled synthetic nodes in regions where minorities form local majorities. The method incorporates both global and local minority perspectives in seed node selection and determines connections between synthetic and real nodes using an edge predictor, ensuring structural coherence in the augmented graph.

3.1.3 Pseudo-Labeling. Pseudo-labeling has emerged as a powerful approach for handling class imbalance in partially labeled graphs. Unlike data interpolation or adversarial generation methods that create synthetic data, pseudo-labeling exploits the abundant unlabeled nodes naturally present in graphs while preserving the original graph structure and node distribution.

The development of graph-based pseudo-labeling methods has been built on the basis of label propagation [64], a foundational work in semi-supervised learning. SPARC [78] enhances this foundation by combining label propagation with self-paced curriculum learning to enable gradual learning from easy to hard samples when enriching minority classes. SET-GNN [24] further advances label propagation by reducing noise in self-training through iterative refinement, while promoting the propagation of minority classes through reweighting. DPGNN [58] enhances label propagation with class prototypes and distance metric learning to better capture the node relationships with class prototypes. GraphSR [79] improves the reliability of pseudo-labels through a two-stage selection process: first filtering potential candidates based on similarity and then using reinforcement learning to adaptively select reliable unlabeled nodes for minority class augmentation. INS-GNN [25] advances this direction through a model-agnostic framework. The framework first employs pre-training with self-supervised objectives that naturally mitigates class imbalance, and then generates pseudo-labels using both node and structural information, while performing targeted edge augmentation for minority nodes to increase their engagement in message passing.

3.2 Algorithm-Level Methods

Algorithm-level methods aim at modifying learning algorithms to address class imbalance. These approaches adjust the model's training process to mitigate problems arising from class imbalance. We classify CILG algorithm-level methods into three categories: model and training refinement, loss function engineering, and post-hoc adjustment, each targeting distinct components of the learning algorithm. Each category represents a distinct strategy to address the class imbalance.

3.2.1 Model and Training Refinement. Model and training refinement approaches handle class imbalance by modifying model architectures or training pipelines. Rather than addressing imbalance through pre-processing or post-processing steps, these methods integrate robustness to class imbalance directly into how graph representation learning methods process and learn from data.

Model architectural modifications focus on re-designing the core components of network embedding models or GNNs. For network embedding methods, several approaches improve upon DeepWalk [43]: RSDNE [60] introduces auxiliary components to balance intra-class similarities and inter-class dissimilarities in the embedding space, while ImVerde [61] adjusts the random walk process by modifying the transition probabilities and implemented balanced mini-batch sampling. Recent GNN-based approaches focus on modifications to message passing: ACS-GNN [39]

incorporates an attention mechanism for class-aware neighbor weighting during message passing, EGCN [55] implements class-proportional message passing to balance information flow between nodes of majority and minority classes, and KINC-GCN [1] introduces kernel-based message passing combined with node clustering to capture both local and global structural patterns.

Meanwhile, adjustments in the training pipeline modify how models learn from imbalanced data while maintaining the underlying model architecture. ALLIE [7] employs active learning that samples instances from both majority and minority classes based on reinforcement learning, guided by an imbalance-aware reward function. LTE4G [69] leverages knowledge distillation by training class-specific student models guided by teacher GNNs that operate on node subsets with balanced class distributions and node degrees. BNE [81] achieves efficient balanced training by identifying nodes that are heavily influenced by minority classes during message passing.

3.2.2 Loss Function Engineering. Loss function engineering adapts training objectives to address class imbalance in machine learning models. Traditional approaches either reweight losses to emphasize the minority class samples [9, 30] or modify decision boundaries to increase the separation between the majority and minority classes [3, 10].

Early CILG methods used loss function engineering as an auxiliary component. ReNode [6] later made it central to the approach through topology-aware reweighting. ReNode adjusts the influence of labeled nodes based on their positions relative to class boundaries in the graph structure:

$$\mathcal{L}_{\text{ReNode}} = \frac{1}{|\mathcal{V}^l|} \sum_{v \in \mathcal{V}^l} w_v \frac{|\bar{\mathcal{C}}|}{|C_{y_v}|} \mathcal{L}_{\text{CE}}(\mathbf{l}_v, y_v), \quad (6)$$

where $\mathcal{L}_{\text{CE}}(\cdot, \cdot)$ is the cross-entropy loss, $\mathbf{l}_v \in \mathbb{R}^K$ is the logit of node v , and w_v reweights nodes based on their proximity to topological class centers. $|\bar{\mathcal{C}}|$ denotes the average number of samples per class, while $|C_{y_v}|$ is the number of samples in node v 's class.

TAM [53] focuses on decision boundary expansion. Recognizing that minority nodes with high connectivity to majority classes are more likely to be misclassified, TAM adjusts decision boundaries through topology-aware margin calibration:

$$\mathcal{L}_{\text{TAM}} = \frac{1}{|\mathcal{V}^l|} \sum_{v \in \mathcal{V}^l} \mathcal{L}_{\text{CE}}(\mathbf{l}_v + \alpha \mathbf{m}_v^{\text{ACM}} + \beta \mathbf{m}_v^{\text{ADM}}, y_v), \quad (7)$$

where $\mathbf{m}_v^{\text{ACM}} \in \mathbb{R}^K$ reduces class margins when a node shows anomalously high neighbor density for specific classes and $\mathbf{m}_v^{\text{ADM}} \in \mathbb{R}^K$ further adjusts margins based on the deviation between the node's neighbor distribution and class-wise average statistics.

On the other hand, ReVar [66] introduces a fundamentally different theoretical perspective by connecting class-imbalanced node classification with bias-variance decomposition theory. Through theoretical analysis, ReVar demonstrates that class imbalance leads to poor prediction accuracy due to increased model variance. This insight led to a composite loss function:

$$\mathcal{L}_{\text{ReVar}} = \lambda_1 \mathcal{L}_{\text{VR}} + \lambda_2 \mathcal{L}_{\text{IR}} + \mathcal{L}_{\text{CE}} \quad (8)$$

where \mathcal{L}_{VR} implements confidence-based consistency regularization to mitigate confirmation bias in pseudo-labeling for unlabeled nodes and \mathcal{L}_{IR} introduces intra-class aggregation to maintain representation invariance within classes. ReVar's key innovation lies in using graph data augmentation to model different training set distributions, enabling practical estimation of model variance.

Another line of research uses self-supervised learning to address class imbalance. Pretraining a feature encoder through self-supervised objectives helps capture inherent data structures independent of class labels, naturally leading to a more balanced representation space [22, 67]. Among

self-supervised approaches, contrastive learning is dominant in graph learning [36]. It learns node representations by preserving similarity between structurally related nodes using an anchor node u , a positive sample v^+ , and a set of negative samples \mathcal{V}^- , combined in the InfoNCE loss [41]:

$$\mathcal{L}_{\text{InfoNCE}} = -\log \frac{\exp(\text{sim}(\mathbf{e}_u, \mathbf{e}_{v^+})/\tau)}{\exp(\text{sim}(\mathbf{e}_u, \mathbf{e}_{v^+})/\tau) + \sum_{v^- \in \mathcal{V}^-} \exp(\text{sim}(\mathbf{e}_u, \mathbf{e}_{v^-})/\tau)}, \quad (9)$$

where \mathbf{e}_v represents node embeddings, $\text{sim}(\cdot, \cdot)$ measures embedding similarity, and τ is a temperature parameter that controls distribution concentration.

Recent work has improved contrastive learning for class imbalance. CM-GCL [45] combines multi-view contrastive learning with topology-aware network pruning to capture diverse node relationships, ImGCL [70] develops sophisticated sampling strategies that are topology-aware when selecting positive and negative samples, and DataDec [71] combines dynamic sparse training with contrastive learning to improve representation learning quality.

3.2.3 Post-hoc Adjustment. Post-hoc adjustments modify model predictions during inference to mitigate class imbalance. Unlike methods that modify the training process, these approaches recalibrate model outputs after training, offering a computationally efficient solution for large-scale graphs. Despite this advantage, these approaches remain notably underexplored in CILG.

Current approaches in this category focus primarily on embedding-based adjustments. LTE4G [69] introduces class prototype-based inference. It constructs class prototypes by averaging the learned node embeddings of each class to create reference points in the embedding space. Test nodes are then classified based on their similarities to these prototypes rather than using a classifier, helping to mitigate bias toward majority classes. ALLIE [7] incorporates post-hoc evaluation to guide its active learning process. ALLIE evaluates performance on a validation set to generate imbalance-aware reward signals. These signals, which treat minority and majority predictions differently, are then used to optimize the node selection policy by maximizing expected rewards.

The limited exploration of post-hoc adjustments in CILG opens up promising research directions. Future work could investigate graph-specific threshold adjustment techniques that account for local neighborhood structure and feature similarity when modifying decision boundaries for individual nodes [31]. Another direction is to develop topology-aware calibration methods that consider both node relationships and structural roles when recalibrating predictions [18, 44].

4 Evaluation

The following sections provide a critical overview of the evaluation settings for CILG, including benchmark datasets for node classification, class-imbalanced train-test split strategies, and commonly used performance metrics in CILG research.

4.1 Datasets

Table 2 summarizes commonly used real-world benchmark datasets for graph node classification. These datasets span diverse application domains, from citation networks where nodes represent articles to be classified by research area to co-purchase networks where nodes represent products to be categorized. While not specifically designed for studying class imbalance, many of these datasets naturally exhibit imbalanced class distributions, with imbalance ratios ranging from 1.92 to 17.73. The Amazon-Computers dataset [51] represents the most extreme case with an imbalance ratio of 17.73:1, making it particularly challenging for standard classification approaches.

Table 2. A Comprehensive Overview of Benchmark Datasets for Real-world Node Classification on Graphs

| Domain | Dataset | Properties | | | Class Distribution (%) | | | | | | | | | |
|-------------|-------------------|-----------------|-----------|--------|------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | | $ \mathcal{V} $ | \bar{k} | ρ | C_0 | C_1 | C_2 | C_3 | C_4 | C_5 | C_6 | C_7 | C_8 | C_9 |
| Citation | Cora [68] | 2,708 | 1.95 | 4.54 | 30.21 | 15.73 | 15.44 | 12.96 | 11.00 | 8.01 | 6.65 | — | — | — |
| | CiteSeer [68] | 3,327 | 1.37 | 2.66 | 21.07 | 20.08 | 17.91 | 17.73 | 15.27 | 7.94 | — | — | — | — |
| | PubMed [68] | 19,717 | 2.25 | 1.92 | 39.94 | 39.25 | 20.81 | — | — | — | — | — | — | — |
| Social | GitHub [49] | 37,700 | 7.67 | 2.87 | 74.17 | 25.83 | — | — | — | — | — | — | — | — |
| | Facebook [49] | 22,470 | 7.61 | 2.07 | 30.62 | 28.91 | 25.67 | 14.81 | — | — | — | — | — | — |
| Co-purchase | Amazon-Photo [51] | 7,650 | 15.57 | 5.86 | 25.37 | 22.04 | 11.96 | 11.53 | 10.76 | 9.19 | 4.82 | 4.33 | — | — |
| | Amazon-Comp. [51] | 13,752 | 17.88 | 17.73 | 37.51 | 15.68 | 15.58 | 10.28 | 5.95 | 3.94 | 3.54 | 3.17 | 2.24 | 2.12 |

Notes: $|\mathcal{V}|$: number of nodes; \bar{k} : average degree; ρ : imbalance ratio (largest/smallest class); C_i : percentage (%) of class i .

4.2 Class-Imbalanced Train-Test Split

Although real-world datasets provide natural examples of class imbalance, the evaluation of CILG models often requires more extreme scenarios to thoroughly assess their robustness. To create such controlled scenarios, researchers have developed specialized train-test split strategies. The process first constructs a class-balanced test set by randomly selecting an equal number of nodes from each class. The remaining nodes are then assigned to the training set and sampled using one of two strategies: (1) long-tailed sampling, which creates a smooth exponential decline in class sizes suitable for supervised learning, or (2) step sampling, which creates a stark division between majority and minority classes to simulate semi-supervised learning with limited labeled data.

4.2.1 Long-Tailed Sampling. Long-tailed sampling, characterized by an imbalance ratio ρ , adapts techniques originally developed for computer vision datasets such as CIFAR [9] to the graph domain. This method carefully reduces the number of nodes per class while maintaining graph connectivity [42]. The sampling follows an exponential decay pattern across classes, where each successive class contains $\rho^{1/(K-1)}$ times fewer samples than the previous one, ultimately achieving a head-to-tail class ratio of ρ . This approach creates a smooth transition from majority to minority classes, mimicking natural power-law distributions often observed in real-world networks.

4.2.2 Step Sampling. An imbalance ratio ρ and a parameter μ define the step sampling method, following a well-established semi-supervised node classification setting [26, 68]. A small number n of training samples (e.g., 20) are selected for each class. The method randomly chooses μK classes as minority classes and $(1 - \mu)K$ classes as majority classes, where K is the number of class labels. From the minority classes, it randomly selects ρn nodes to form a class-imbalanced training set.

4.3 Performance Metrics

Performance evaluation in class-imbalanced settings requires metrics that effectively capture model behavior across all classes. While we present these metrics for binary classification, they generalize to multiclass scenarios through micro- or macro-averaging. These metrics are computed with **true positives (TP)**, **true negatives (TN)**, **false positives (FP)**, and **false negatives (FN)**.

4.3.1 Accuracy: A Misleading Metric for Class Imbalance. **Acc** (accuracy) is the ratio of the number of correct predictions to the total number of predictions:

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \quad (10)$$

Although commonly used, accuracy is inappropriate for imbalanced learning. It exhibits a strong bias towards majority classes, as a naive classifier that always predicts the majority class can achieve deceptively high accuracy while failing entirely in minority classes [38].

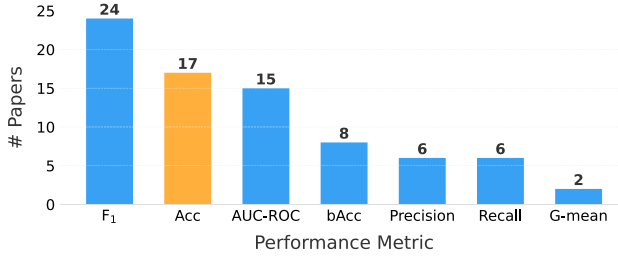


Fig. 1. Frequency analysis of performance metrics in CILG literature, with accuracy (shown in orange) being an improper metric for imbalanced learning scenarios.

4.3.2 Class-Balanced Evaluation Metrics. To address this limitation and provide a more comprehensive assessment of model performance across all classes, we recommend the following metrics:

- **F₁** is the harmonic average of **Precision** (positive predicted values) and **Recall** (sensitivity):

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}, \quad F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (11)$$

- **bAcc** (balanced accuracy) equally weights the performance of both classes:

$$\text{bAcc} = \frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{FP + TN} \right). \quad (12)$$

- **G-mean** (geometric mean) captures performance balance through geometric averaging:

$$\text{G-mean} = \sqrt{\frac{TP}{TP + FN} \cdot \frac{TN}{FP + TN}}. \quad (13)$$

- **AUC-ROC** (area under the ROC curve) measures the model's ability to distinguish between classes across different classification thresholds by calculating the area under the **receiver operating characteristic (ROC)** curve.

As is shown in Figure 1, our analysis reveals concerning evaluation practices. Accuracy, despite its known limitations in imbalanced settings, remains the main metric in 55% of the articles (17/31). Furthermore, one article reports only precision and two reports only recall, providing incomplete performance assessment. For meaningful evaluation in imbalanced scenarios, we strongly advocate adopting the recommended metrics above, which collectively provide a comprehensive view of model performance across all classes. Continuing to use accuracy can mask critical failures in minority class prediction and lead to over-optimistic performance assessments.

5 Future Directions

Most existing research on CILG mainly focuses on node classification in simple homogeneous and homophilic graphs. However, imbalance issues appear in various graph mining tasks and complex graph types. Beyond quantity imbalances in node classes, the field also needs to address topology imbalances and other scenarios that affect graph-based applications in real-world situations.

5.1 Beyond Node Classification

While node classification has been well-studied in CILG, other graph mining tasks at both node and graph levels present unique challenges in handling imbalance.

5.1.1 Other Node-Level Tasks.

Link Prediction. In real-world graphs, the number of non-existent edges typically far exceeds the number of existing edges, creating a natural imbalance that makes link prediction particularly challenging [29]. Similar to node classification, this extreme ratio between positive (existing) and negative (non-existing) edges can lead to biased predictions favoring the majority class. The local structures and features characterizing true connections might be overwhelmed by the abundance of non-connections during training, affecting the model's ability to identify meaningful potential links. This challenge becomes especially significant in evolving networks where new nodes or communities have limited connection patterns, creating scenarios where certain edge types effectively become minority classes that are difficult to predict accurately.

Node Regression. In node regression, the objective is to predict underrepresented numerical target values of nodes. Unlike classification where class distributions provide clear metrics, defining and measuring imbalance in continuous target spaces remains challenging [48]. Models tend to be biased toward mean values and struggle to accurately predict extreme values in the tails of the target distribution, leading to systematic underestimation of high values and overestimation of low values. Additionally, traditional regression metrics may not adequately reflect model performance on rare extreme values, as they are dominated by errors in common cases [2].

5.1.2 Graph-Level Tasks.

Graph Classification. Class-imbalanced graph classification aims to predict labels for entire graphs while dealing with challenges caused by multiple forms of imbalance. Unlike node classification, where minority nodes can leverage neighboring information, graphs are independent entities, making it difficult to share information between samples in minority classes [59]. This independence prevents the direct application of traditional neighborhood-based techniques for handling imbalance, particularly in applications such as chemical compound classification, where a vast majority of compounds are labeled as inactive while only a small portion is labeled as active. Moreover, minority class graphs may also exhibit distinct characteristics in terms of their overall structure [65]. This combination of sample scarcity and unique topological properties in minority classes requires sophisticated approaches that can address both aspects simultaneously.

Graph Regression. Graph regression tasks, such as molecular property prediction [19], face imbalance challenges when annotated data concentrate on popular property ranges while leaving other areas of the property space underrepresented. Unlike node regression, addressing this imbalance requires considering how entire graph structures correlate with property values, while expensive experimental costs often limit data collection for underrepresented ranges. Recent work has shown promising results in adapting pseudo-labeling methods mentioned in Section 3.1.3 originally developed for imbalanced node classification to imbalanced graph regression tasks [32].

5.2 Beyond Homogeneous and Homophilic Graphs

Although a significant portion of existing research in graph mining has focused on simple graphs that are homogeneous and homophilic, it is crucial to explore the potential for addressing class imbalance in more complex graph types. In particular, heterogeneous graphs with multiple node and edge types, and heterophilic graphs with diverse connectivity patterns, present unique challenges in handling class imbalance that extend beyond those found in simple graphs.

5.2.1 Heterogeneous Graphs. Heterogeneous graphs [57] contain multiple types of nodes and edges, reflecting complex relationships in various real-world scenarios. While most CILG research

focuses on homogeneous graphs, addressing class imbalance in heterogeneous graphs presents unique challenges beyond those in homogeneous graphs. Different node types may exhibit varying degrees of imbalance, and the semantic relationships between node types can affect how the imbalance issue propagates through the graph [14]. For example, in financial networks, the scarcity of bankrupt companies (minority class) affects not only the company nodes but also their associated transaction patterns and relationships with other entity types [35]. The interplay between different node types and their relationships requires careful consideration of how imbalance mitigation strategies might affect the overall semantic structure of the heterogeneous graph.

5.2.2 Heterophilic Graphs. Heterophilic graphs [77] exhibit variations in the connectivity patterns and attributes of nodes, leading to diverse subpopulations within the graph. In the context of class imbalance, heterophilic graphs present unique challenges that extend beyond simple quantity imbalance. The heterophilic nature of connections means minority class nodes often receive messages primarily from majority class nodes during feature propagation, which can lead to their characteristics being overwhelmed or assimilated by the majority class [13, 33]. This interaction between heterophily and class imbalance creates a compounded challenge: not only are minority class instances rare, but their feature representations may also be distorted due to the dominance of different-class neighbors in the message passing process. These challenges require careful consideration of node similarity metrics and neighbor selection strategies that can preserve minority class characteristics while still leveraging the graph structure.

5.3 Beyond Class Imbalance

In this section, we investigate *topology imbalance*, which focuses on how the structural positions and patterns of the labeled nodes affect learning, rather than just their quantity. While class imbalance focuses on the number of samples per class, topology imbalance spans multiple dimensions: (1) the asymmetric distribution of nodes in the graph topology [6], (2) the uneven representation of subclass-level topological motifs in training data [75], and (3) the hierarchical roles of nodes in the graph structure [12]. Each dimension of topology imbalance presents distinct learning challenges. In positional distribution, nodes near topological class boundaries can disproportionately influence decision boundaries [6]. At the subclass level, underrepresented topological motifs can lead models to overfit to majority patterns [75]. In hierarchical structures, the distribution of labeled nodes across different levels can significantly impact model learning, causing severe decision boundary shifts despite having balanced quantities and positions in the graph [12].

The flow of supervision information through the graph reveals additional challenges. Two key phenomena emerge [54]: under-reaching, where nodes distant from labeled examples receive insufficient supervision, and over-squashing, where supervision information becomes diluted when traversing narrow paths between different regions of the graph. These effects intensify in hierarchical structures, where cross-hierarchy connectivity creates information bottlenecks [12].

6 Conclusion

In conclusion, this article presents the first comprehensive survey on CILG, addressing the significant challenges of class imbalance in graph data. By systematically categorizing existing CILG methods, discussing pressing research questions, and offering a roadmap for future investigations, the article aims to raise awareness within both the graph learning and imbalanced learning communities and to encourage further research in this emerging area. As the applicability of class-imbalanced learning techniques expands across various graph-based domains, researchers and practitioners will be better equipped to develop effective and efficient solutions for handling class imbalance problems, benefiting a wide range of applications.

Acknowledgments

This work was supported in part by National Science Foundation through the NSF Center for Computer-Assisted Synthesis (C-CAS), under grant number CHE-2202693.

References

- [1] Xiang-En Bai, Jing An, Zi-Bo Yu, Han-Qiu Bao, and Ke-Fan Wang. 2022. A kernel propagation-based graph convolutional network imbalanced node classification model on graph data. In *Proceedings of the ICNSC*. IEEE, 1–6.
- [2] Paula Branco, Luis Torgo, and Rita P. Ribeiro. 2016. A survey of predictive modeling on imbalanced domains. *ACM Computing Surveys* 49, 2 (2016), 1–50.
- [3] Kaidi Cao, Colin Wei, Adrien Gaidon, Nikos Arechiga, and Tengyu Ma. 2019. Learning imbalanced datasets with label-distribution-aware margin loss. In *Proceedings of the NeurIPS*. 1567–1578.
- [4] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* 16 (2002), 321–357.
- [5] Nitesh V. Chawla, Aleksandar Lazarevic, Lawrence O. Hall, and Kevin W. Bowyer. 2003. SMOTEBoost: Improving prediction of the minority class in boosting. In *Proceedings of the PKDD*. Springer, 107–119.
- [6] Deli Chen, Yankai Lin, Guangxiang Zhao, Xuancheng Ren, Peng Li, Jie Zhou, and Xu Sun. 2021. Topology-imbalance learning for semi-supervised node classification. In *Proceedings of the NeurIPS*. 29885–29897.
- [7] Limeng Cui, Xianfeng Tang, Sumeet Katariya, Nikhil Rao, Pallav Agrawal, Karthik Subbian, and Dongwon Lee. 2022. Allie: Active learning on large-scale imbalanced graphs. In *Proceedings of the WWW*. 690–698.
- [8] Peng Cui, Xiao Wang, Jian Pei, and Wenwu Zhu. 2018. A survey on network embedding. *IEEE Transactions on Knowledge and Data Engineering* 31, 5 (2018), 833–852.
- [9] Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, and Serge Belongie. 2019. Class-balanced loss based on effective number of samples. In *Proceedings of the CVPR*. 9268–9277.
- [10] Qi Dong, Shaogang Gong, and Xiatian Zhu. 2018. Imbalanced deep learning by minority class incremental rectification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41, 6 (2018), 1367–1381.
- [11] Yijun Duan, Xin Liu, Adam Jatowt, Hai-tao Yu, Steven Lynden, Kyoung-Sook Kim, and Akiyoshi Matono. 2022. Anonymity can help minority: A novel synthetic data over-sampling strategy on multi-label graphs. In *Proceedings of the ECML PKDD*. Springer, 20–36.
- [12] Xingcheng Fu, Yuecen Wei, Qingyun Sun, Haonan Yuan, Jia Wu, Hao Peng, and Jianxin Li. 2023. Hyperbolic geometric graph representation learning for hierarchy-imbalance node classification. In *Proceedings of the WWW*. 460–468.
- [13] Jiayi Gao, Jiaxing Li, Ke Zhang, and Youyong Kong. 2023. Topology uncertainty modeling for imbalanced node classification on graphs. In *Proceedings of the ICASSP*. IEEE, 1–5.
- [14] Xinyi Gao, Wentao Zhang, Tong Chen, Junliang Yu, Hung Quoc Viet Nguyen, and Hongzhi Yin. 2023. Semantic-aware node synthesis for imbalanced heterogeneous information networks. In *Proceedings of the CIKM*. 545–555.
- [15] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. 2017. Neural message passing for quantum chemistry. In *Proceedings of the ICML*. PMLR, 1263–1272.
- [16] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative adversarial networks. *Communications of the ACM* 63, 11 (2020), 139–144.
- [17] Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable feature learning for networks. In *Proceedings of the KDD*. 855–864.
- [18] Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. 2017. On calibration of modern neural networks. In *Proceedings of the ICML*. PMLR, 1321–1330.
- [19] Zhichun Guo, Kehan Guo, Bozhao Nan, Yijun Tian, Roshni G. Iyer, Yihong Ma, Olaf Wiest, Xiangliang Zhang, Wei Wang, Chuxu Zhang, et al. 2023. Graph-based molecular representation learning. In *Proceedings of the IJCAI*. 6638–6646.
- [20] Guo Haixiang, Li Yijing, Jennifer Shang, Gu Mingyun, Huang Yuanyue, and Gong Bing. 2017. Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications* 73 (2017), 220–239.
- [21] Sophie Henning, William Beluch, Alexander Fraser, and Annemarie Friedrich. 2023. A survey of methods for addressing class imbalance in deep-learning based natural language processing. In *Proceedings of the EACL*. 523–540.
- [22] Ziyu Jiang, Tianlong Chen, Ting Chen, and Zhangyang Wang. 2021. Improving contrastive learning on imbalanced data via open-world sampling. In *Proceedings of the NeurIPS*. 5997–6009.
- [23] Justin M. Johnson and Taghi M. Khoshgohfar. 2019. Survey on deep learning with class imbalance. *Journal of Big Data* 6, 1 (2019), 1–54.
- [24] Xin Juan, Meixin Peng, and Xin Wang. 2021. Exploring self-training for imbalanced node classification. In *Proceedings of the ICONIP*. Springer, 28–36.

- [25] Xin Juan, Fengfeng Zhou, Wentao Wang, Wei Jin, Jiliang Tang, and Xin Wang. 2023. INS-GNN: Improving graph imbalance learning with self-supervision. *Information Sciences* 637 (2023), 118935.
- [26] Thomas N. Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. In *Proceedings of the ICLR*.
- [27] Wen-Zhi Li, Chang-Dong Wang, Hui Xiong, and Jian-Huang Lai. 2023. Graphsha: Synthesizing harder samples for class-imbalanced node classification. In *Proceedings of the KDD*. 1328–1340.
- [28] Xiaohe Li, Zide Fan, Feilong Huang, Xuming Hu, Yawen Deng, Lei Wang, and Xinyu Zhao. 2024. Graph neural network with curriculum learning for imbalanced node classification. *Neurocomputing* 574 (2024), 127229.
- [29] Ryan N. Lichtenwalter, Jake T. Lussier, and Nitesh V. Chawla. 2010. New perspectives and methods in link prediction. In *Proceedings of the KDD*. 243–252.
- [30] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. 2017. Focal loss for dense object detection. In *Proceedings of the ICCV*. 2980–2988.
- [31] Charles X. Ling and Victor S. Sheng. 2008. Cost-sensitive learning and the class imbalance problem. *Encyclopedia of Machine Learning* 2011 (2008), 231–235.
- [32] Gang Liu, Tong Zhao, Eric Inae, Tengfei Luo, and Meng Jiang. 2023. Semi-supervised graph imbalanced regression. In *Proceedings of the KDD*. 1453–1465.
- [33] Jie Liu, Mengting He, Guangtao Wang, Quoc Viet Hung Nguyen, Xuequn Shang, and Hongzhi Yin. 2023. Imbalanced node classification beyond homophilic assumption. In *Proceedings of the IJCAI*. 7206–7214.
- [34] Xu-Ying Liu, Jianxin Wu, and Zhi-Hua Zhou. 2008. Exploratory undersampling for class-imbalance learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39, 2 (2008), 539–550.
- [35] Yucheng Liu, Zipeng Gao, Xiangyang Liu, Pengfei Luo, Yang Yang, and Hui Xiong. 2023. QTIAH-GNN: Quantity and topology imbalance-aware heterogeneous graph neural network for bankruptcy prediction. In *Proceedings of the KDD*. 1572–1582.
- [36] Yixin Liu, Ming Jin, Shirui Pan, Chuan Zhou, Yu Zheng, Feng Xia, and S. Yu Philip. 2022. Graph self-supervised learning: A survey. *IEEE Transactions on Knowledge and Data Engineering* 35, 6 (2022), 5879–5900.
- [37] Yongxu Liu, Zhi Zhang, Yan Liu, and Yao Zhu. 2022. Gatsmote: Improving imbalanced node classification on graphs via attention and homophily. *Mathematics* 10, 11 (2022), 1799.
- [38] Victoria López, Alberto Fernández, Salvador García, Vasile Palade, and Francisco Herrera. 2013. An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information Sciences* 250 (2013), 113–141.
- [39] Chao Ma, Jing An, Xiang-En Bai, and Han-Qiu Bao. 2022. Attention and cost-sensitive graph neural network for imbalanced node classification. In *Proceedings of the ICNSC*. IEEE, 1–6.
- [40] Mehdi Mirza. 2014. Conditional generative adversarial nets. arXiv:1411.1784. Retrieved from <https://arxiv.org/abs/1411.1784>
- [41] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. 2018. Representation learning with contrastive predictive coding. arXiv:1807.03748. Retrieved from <https://arxiv.org/abs/1807.03748>
- [42] Joonhyung Park, Jaeyun Song, and Eunho Yang. 2021. Graphens: Neighbor-aware ego network synthesis for class-imbalanced node classification. In *Proceedings of the ICLR*.
- [43] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. Deepwalk: Online learning of social representations. In *Proceedings of the KDD*. 701–710.
- [44] John Platt. 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in Large Margin Classifiers* 10, 3 (1999), 61–74.
- [45] Yiyue Qian, Chunhui Zhang, Yiming Zhang, Qianlong Wen, Yanfang Ye, and Chuxu Zhang. 2022. Co-modality graph contrastive learning for imbalanced node classification. In *Proceedings of the NeurIPS*. 15862–15874.
- [46] Liang Qu, Huaisheng Zhu, Ruiqi Zheng, Yuhui Shi, and Hongzhi Yin. 2021. Imagn: Imbalanced network embedding via generative adversarial graph networks. In *Proceedings of the KDD*. 1390–1398.
- [47] Enislay Ramentol, Yailé Caballero, Rafael Bello, and Francisco Herrera. 2012. Smote-rs b*: A hybrid preprocessing approach based on oversampling and undersampling for high imbalanced data-sets using smote and rough sets theory. *Knowledge and Information Systems* 33 (2012), 245–265.
- [48] Rita P. Ribeiro and Nuno Moniz. 2020. Imbalanced regression and extreme value prediction. *Machine Learning* 109 (2020), 1803–1835.
- [49] Benedek Rozemberczki, Carl Allen, and Rik Sarkar. 2021. Multi-scale attributed node embedding. *Journal of Complex Networks* 9, 2 (2021), cnab014.
- [50] Francisco Santos, Junke Ye, Farzan Masrour, Pang-Ning Tan, and Abdol-Hossein Esfahanian. 2022. FACS-GCN: Fairness-aware cost-sensitive boosting of graph convolutional networks. In *Proceedings of the IJCNN*. IEEE, 1–8.
- [51] Oleksandr Shchur, Maximilian Mumme, Aleksandar Bojchevski, and Stephan Günnemann. 2018. Pitfalls of graph neural network evaluation. arXiv:1811.05868. Retrieved from <https://arxiv.org/abs/1811.05868>

- [52] Min Shi, Yufei Tang, Xingquan Zhu, David Wilson, and Jianxun Liu. 2020. Multi-class imbalanced graph convolutional network learning. In *Proceedings of the IJCAI*. 2879–2885.
- [53] Jaeyun Song, Joonhyung Park, and Eunho Yang. 2022. TAM: Topology-aware margin loss for class-imbalanced node classification. In *Proceedings of the ICML*. PMLR, 20369–20383.
- [54] Qingyun Sun, Jianxin Li, Haonan Yuan, Xingcheng Fu, Hao Peng, Cheng Ji, Qian Li, and Philip S. Yu. 2022. Position-aware structure learning for graph topology-imbalance by relieving under-reaching and over-squashing. In *Proceedings of the CIKM*. 1848–1857.
- [55] Kefan Wang, Jing An, and Qi Kang. 2022. Effective-aggregation graph convolutional network for imbalanced classification. In *Proceedings of the ICNSC*. IEEE, 1–5.
- [56] Quan Wang, Zhendong Mao, Bin Wang, and Li Guo. 2017. Knowledge graph embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering* 29, 12 (2017), 2724–2743.
- [57] Xiao Wang, Deyu Bo, Chuan Shi, Shaohua Fan, Yanfang Ye, and S. Yu Philip. 2022. A survey on heterogeneous graph embedding: Methods, techniques, applications and sources. *IEEE Transactions on Big Data* 9, 2 (2022), 415–436.
- [58] Yu Wang, Charu Aggarwal, and Tyler Derr. 2022. Distance-wise prototypical graph neural network for imbalanced node classification. In *Proceedings of the 17th International Workshop on Mining and Learning with Graphs (MLG)*.
- [59] Yu Wang, Yuying Zhao, Neil Shah, and Tyler Derr. 2022. Imbalanced graph classification via graph-of-graph neural networks. In *Proceedings of the CIKM*. 2067–2076.
- [60] Zheng Wang, Xiaojun Ye, Chaokun Wang, Yuexin Wu, Changping Wang, and Kaiwen Liang. 2018. RSDNE: Exploring relaxed similarity and dissimilarity from completely-imbalanced labels for network embedding. In *Proceedings of the AAAI*. 475–482.
- [61] Jun Wu, Jingrui He, and Yongming Liu. 2018. Imverde: Vertex-diminished random walk for learning imbalanced network representation. In *Proceedings of the Big Data*. IEEE, 871–880.
- [62] Lirong Wu, Jun Xia, Zhangyang Gao, Haitao Lin, Cheng Tan, and Stan Z. Li. 2022. Graphmixup: Improving class-imbalanced node classification by reinforcement mixup and self-supervised context prediction. In *Proceedings of the ECML PKDD*. Springer, 519–535.
- [63] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S. Yu Philip. 2020. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems* 32, 1 (2020), 4–24.
- [64] Zhu Xiaojin and Ghahramani Zoubin. 2002. *Learning from Labeled and Unlabeled Data with Label Propagation*. Technical Report CMU-CALD-02–107. Carnegie Mellon University.
- [65] Wei Xu, Pengkun Wang, Zhe Zhao, Binwu Wang, Xu Wang, and Yang Wang. 2024. When imbalance meets imbalance: Structure-driven learning for imbalanced graph classification. In *Proceedings of the WWW*. 905–913.
- [66] Divin Yan, Gengchen Wei, Chen Yang, Shengzhong Zhang, and Zengfeng Huang. 2024. Rethinking semi-supervised imbalanced node classification from bias-variance decomposition. In *Proceedings of the NeurIPS*. 29174–29200.
- [67] Yuzhe Yang and Zhi Xu. 2020. Rethinking the value of labels for improving class-imbalanced learning. In *Proceedings of the NeurIPS*. 19290–19301.
- [68] Zhilin Yang, William Cohen, and Ruslan Salakhudinov. 2016. Revisiting semi-supervised learning with graph embeddings. In *Proceedings of the ICML*. PMLR, 40–48.
- [69] Sukwon Yun, Kibum Kim, Kanghoon Yoon, and Chanyoung Park. 2022. Lte4g: Long-tail experts for graph neural networks. In *Proceedings of the CIKM*. 2434–2443.
- [70] Liang Zeng, Lanqing Li, Ziqi Gao, Peilin Zhao, and Jian Li. 2023. Imgcl: Revisiting graph contrastive learning on imbalanced node classification. In *Proceedings of the AAAI*. 11138–11146.
- [71] Chunhui Zhang, Chao Huang, Yijun Tian, Qianlong Wen, Zhongyu Ouyang, Youhuan Li, Yanfang Ye, and Chuxu Zhang. 2023. When sparsity meets contrastive models: Less graph data can bring better class-balanced representations. In *Proceedings of the ICML*. PMLR, 41133–41150.
- [72] Hongyi Zhang, Moustopha Cisse, Yann N Dauphin, and David Lopez-Paz. 2018. mixup: Beyond empirical risk minimization. In *Proceedings of the ICLR*.
- [73] Yifan Zhang, Bingyi Kang, Bryan Hooi, Shuicheng Yan, and Jiashi Feng. 2023. Deep long-tailed learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 9 (2023), 10795–10816.
- [74] Ziwei Zhang, Peng Cui, and Wenwu Zhu. 2020. Deep learning on graphs: A survey. *IEEE Transactions on Knowledge and Data Engineering* 34, 1 (2020), 249–270.
- [75] Tianxiang Zhao, Dongsheng Luo, Xiang Zhang, and Suhang Wang. 2022. Topoimb: Toward topology-level imbalance in learning from graphs. In *Proceedings of the Learning on Graphs Conference (LoG)*. PMLR, 37:1–37:18.
- [76] Tianxiang Zhao, Xiang Zhang, and Suhang Wang. 2021. Graphsmote: Imbalanced node classification on graphs with graph neural networks. In *Proceedings of the WSDM*. 833–841.
- [77] Xin Zheng, Yi Wang, Yixin Liu, Ming Li, Miao Zhang, Di Jin, Philip S Yu, and Shirui Pan. 2022. Graph neural networks for graphs with heterophily: A survey. arXiv:2202.07082. Retrieved from <https://arxiv.org/abs/2202.07082>

- [78] Dawei Zhou, Jingrui He, Hongxia Yang, and Wei Fan. 2018. Sparc: Self-paced network representation for few-shot rare category characterization. In *Proceedings of the KDD*. 2807–2816.
- [79] Mengting Zhou and Zhiguo Gong. 2023. GraphSR: A data augmentation algorithm for imbalanced node classification. In *Proceedings of the AAAI*. 4954–4962.
- [80] Zhi-Hua Zhou and Xu-Ying Liu. 2005. Training cost-sensitive neural networks with methods addressing the class imbalance problem. *IEEE Transactions on Knowledge and Data Engineering* 18, 1 (2005), 63–77.
- [81] Zonghai Zhu, Huanlai Xing, and Yuge Xu. 2023. Balanced neighbor exploration for semi-supervised node classification on imbalanced graph data. *Information Sciences* 631 (2023), 31–44.

Received 13 April 2023; revised 20 January 2025; accepted 3 February 2025