

Singapore Management University

## Institutional Knowledge at Singapore Management University

---

Research Collection School Of Computing and Information Systems

School of Computing and Information Systems

---

8-2024

### Contrastive general graph matching with adaptive augmentation sampling

Jianyuan BO

Singapore Management University, jybo.2020@phdcs.smu.edu.sg

Yuan FANG

Singapore Management University, yfang@smu.edu.sg

Follow this and additional works at: [https://ink.library.smu.edu.sg/sis\\_research](https://ink.library.smu.edu.sg/sis_research)



Part of the [Artificial Intelligence and Robotics Commons](#)

---

#### Citation

BO, Jianyuan and FANG, Yuan. Contrastive general graph matching with adaptive augmentation sampling. (2024). *Proceedings of the International Joint Conference on Artificial Intelligence 2024 : Jeju, South Korea, August 3-9*. 3724-3732.

Available at: [https://ink.library.smu.edu.sg/sis\\_research/9537](https://ink.library.smu.edu.sg/sis_research/9537)

This Conference Proceeding Article is brought to you for free and open access by the School of Computing and Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Computing and Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email [cherylids@smu.edu.sg](mailto:cherylids@smu.edu.sg).

# Contrastive General Graph Matching with Adaptive Augmentation Sampling

Jianyuan Bo , Yuan Fang

Singapore Management University, Singapore

{jybo.2020, yfang}@smu.edu.sg

## Abstract

Graph matching has important applications in pattern recognition and beyond. Current approaches predominantly adopt supervised learning, demanding extensive labeled data which can be limited or costly. Meanwhile, self-supervised learning methods for graph matching often require additional side information such as extra categorical information and input features, limiting their application to the general case. Moreover, designing the optimal graph augmentations for self-supervised graph matching presents another challenge to ensure robustness and efficacy. To address these issues, we introduce a novel Graph-centric Contrastive framework for Graph Matching (GCGM), capitalizing on a vast pool of graph augmentations for contrastive learning, yet without needing any side information. Given the variety of augmentation choices, we further introduce a Boosting-inspired Adaptive Augmentation Sampler (BiAS), which adaptively selects more challenging augmentations tailored for graph matching. Through various experiments, our GCGM surpasses state-of-the-art self-supervised methods across various datasets, marking a significant step toward more effective, efficient and general graph matching.

## 1 Introduction

Graph Matching (GM) refers to the process of *establishing a correspondence between the nodes of two graphs* based on their structure and node attributes [Yan *et al.*, 2016; Zanfir and Sminchisescu, 2018], and entails various applications. In pattern recognition [Shokoufandeh *et al.*, 2012], graph matching facilitates the alignment of visual patterns within images [Wang *et al.*, 2018]. In bioinformatics [Krissinel and Henrick, 2004], it is crucial for understanding protein-protein interaction networks [Zaslavskiy *et al.*, 2009]. In social networks, aligning user profiles and their connections is essential for understanding social behaviors [Zhang *et al.*, 2018].

Despite the widespread applications of graph matching, prevailing methods such as BBGM [Rolínek *et al.*, 2020] and NGM [Wang *et al.*, 2021] heavily depend on the supervised learning paradigm, which requires a large number of label to

annotate the node correspondence across graphs. To avoid the costly labeling effort, self-supervised learning (SSL) on graphs [Peng *et al.*, 2020; Zeng and Xie, 2021] becomes a practical alternative. Different from supervised learning, SSL exploits label-free graphs to model intrinsic graph properties general to an application domain, which guides the learning of node representations amenable to graph matching. While label-free graphs are often abundant or easy to obtain, existing SSL-based approaches for graph matching often assume the availability of additional side information besides graphs. On the one hand, approaches like GANN [Wang *et al.*, 2023] and IA-SSGM [Guo *et al.*, 2023] still require graph pairs that are potentially matchable as training data, although they do not need exact annotations on node correspondence. The potentially matchable graph pairs are typically sampled from the same category, under the assumption that graph pairs from distinct categories involve totally different kinds of nodes and cannot be matched. Hence, such approaches require extra categorical information, which can be considered as a form of weak supervision. On the other hand, SCGM [Liu *et al.*, 2022] requires raw image as input in order to perform augmentation, leveraging additional image features in visual applications, and thus cannot be generalized to graph matching in non-visual domains. Additionally, image features may not be available for privacy or commercial considerations.

Thus, toward more general graph matching, an immediate challenge arises: *Is it feasible to employ SSL for graph matching without relying on any side information?* To address this, we develop a Graph-centric Contrastive framework for Graph Matching (GCGM), using a *large pool* of graph augmentations to compensate for the lack of any side information. As a form of SSL, contrastive learning (CL) [Wu *et al.*, 2018; Chen *et al.*, 2020] regards corresponding/non-corresponding nodes from two different views of the same graph as positive/negative pairs, where a graph can be augmented in different ways to generate different views [Hassani and Khasahmadi, 2020]. For robustness, we consider both structure- and feature-space augmentations. First, the structure-space augmentations include edge removal, node dropping, and node replacement, etc., which alter the structure of the graph, thereby improving matching against structure variations. Second, the feature-space augmentations include feature masking and scaling, which manipulate nodes features, thereby improving matching against feature variations. Generally speak-

ing, each type of augmentation focuses on a specific kind of variation across graphs to enhance a specific aspect of matching capability. Moreover, each type of augmentation (e.g., node dropping) in fact entails a family of augmentations defined by some hyperparameters (e.g., probability of dropping a node) to vary the difficulty levels within each type. Together, we form a large, comprehensive pool of augmentations to improve the robustness of graph matching.

However, given a comprehensive pool of graph augmentations, a second challenge follows: *How do we select the most effective augmentations for graph matching?* The pool of graph augmentations includes various types of structure- and feature-space augmentations, and each type can instantiate a family of augmented views by tuning the hyperparameters of that type. In existing graph contrastive learning, hyperparameter tuning for each type of augmentation is the prevailing solution [You *et al.*, 2020], which can be expensive and tends to overfit. A few automated methods exist in selecting the right type of augmentations. JOAO [You *et al.*, 2021] always exploits the most challenging augmentation w.r.t. the current loss, while recent research [Cai *et al.*, 2020; Robinson *et al.*, 2021; Wang and Liu, 2021] finds that the hardest instances are not necessarily the best for learning. Meanwhile, GCA [Zhu *et al.*, 2021] adaptively augments each graph through edge removal based on node centrality and masking unimportant node features, which requires custom design for each type of augmentation and does not generalize to other types. To counter this challenge, we present a **Boosting-inspired Adaptive Augmentation Sampler (BiAS)** that automatically and adaptively samples graph augmentations beyond naïve uniform sampling, providing a general solution to the otherwise expensive task of selecting and tuning the optimal augmentations.

By addressing the challenges, we make three major contributions to graph matching. (1) We propose GCGM, a graph-centric contrastive framework for general graph matching, based on a large pool of graph augmentations without the need of any labeled data or side information. (2) We introduce BiAS, an adaptive augmentation selection strategy for GCGM, to optimize graph matching and eliminate manual tuning effort over the pool of augmentations. (3) We conduct experiments on both real-world and synthetic datasets, demonstrating that our approach consistently outperforms state-of-the-art self-supervised baselines.

## 2 Related Work

**Deep Graph Matching** In supervised graph matching, GMN [Zanfir and Sminchisescu, 2018] has pioneered the use of deep learning for GM. Subsequent research, like PCA [Wang *et al.*, 2020], has advanced this by learning node, edge, and affinity features end-to-end. Another notable approach is the channel-independent embedding with Hungarian attention in deep graph matching [Yu *et al.*, 2019].

Furthermore, NGM [Wang *et al.*, 2021] extends GM to hypergraphs and multi-graph matching (MGM). Additionally, BBGM [Rolínek *et al.*, 2020] addresses GM using differentiable operations on combinatorial solvers. Besides, addressing outliers and noise is crucial for obtaining accu-

rate matches. Notably, RGM [Liu *et al.*, 2023] introduced a reinforcement learning approach with a revocable action scheme to counteract the impact of outliers, using a pre-trained BBGM as a feature extractor that still requires labels.

While supervised graph matching methods have progressed using large labeled datasets, unsupervised or self-supervised approaches have emerged to tackle label scarcity. GANN utilizes graduated assignment mechanisms that allows for joint MGM and clustering. SCGM introduces a two-stage contrastive learning framework where image augmentation is essential, necessitating visual access and limiting its use in other domains. Given the advantage of self-supervised learning, our research primarily benchmarks against these approaches in the evolving landscape of graph matching.

**Graph Contrastive Learning** Graph learning has advanced significantly with the advent of graph neural networks (GNNs) [Veličković *et al.*, 2017] and the introduction of contrastive learning methods [Chen *et al.*, 2020], leading to the novel concept of graph contrastive learning (GCL) [You *et al.*, 2020]. As a major branch of self-supervised learning, GCL enhances the representation of nodes [Tan *et al.*, 2022], edges [Li *et al.*, 2023], and entire graphs [Peng *et al.*, 2020] by contrasting corrupted views augmented by different graph augmentations [Xu *et al.*, 2022], providing superior performance to traditional graph learning techniques. However, these approaches face the challenge of designing effective augmentations, which may involve custom designs or extensive hyperparameter tuning [You *et al.*, 2020; Zhao *et al.*, 2021], posing a significant overhead.

## 3 Proposed Method

In this section, we introduce our GCGM framework, starting with the overall framework and preliminaries, followed by our augmentation designs and adaptive sampling strategy.

### 3.1 Overall Framework

**Pre-training** Fig. 1(a) depicts the self-supervised pre-training phase. The illustrated augmentation process is applied across all graphs in the entire training set.

Specifically, for each graph  $\mathcal{G} = \{\mathbf{X}, \mathbf{A}\}$  in a batch during pre-training, defined by node features  $\mathbf{X} \in \mathbb{R}^{N \times F}$  and the adjacency matrix  $\mathbf{A}$ , we sample an augmentation pair  $(\tau_j, \tau_k)$  based on their weights w.r.t. the current mini-batch, from a randomly initialized augmentation pairs pool as shown in Fig. 1(a)(i). Through graph augmentation, we produce two distinct views of each input graph:  $\tilde{\mathcal{G}}^A$  and  $\tilde{\mathcal{G}}^B$ , with  $N_A$  and  $N_B$  nodes respectively, which will then be processed by an graph encoder, capturing both local and global graph patterns. Node embeddings are then refined through the node-level contrastive strategy, optimizing the embedding by matching congruent nodes across views and differentiating non-correspondence nodes. Finally, the encoded embeddings will be processed by the subsequent affinity learning module and GM solver to produce the predicted matching matrix  $\hat{\mathbf{G}} \in \mathbb{R}^{N_A \times N_B}$ .

Enhancing this workflow, as illustrated in Fig. 1(a)(ii), our BiAS technique dynamically updates the weights of the

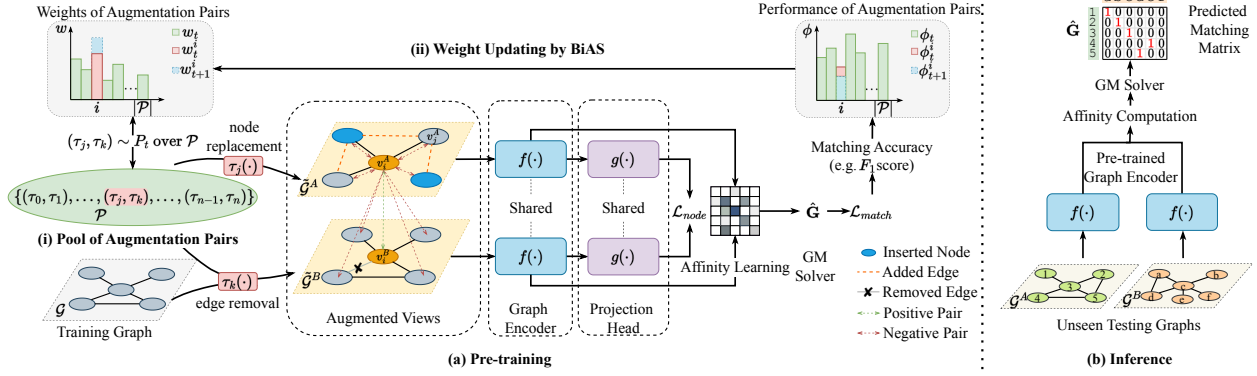


Figure 1: Framework of the proposed GCGM with BiAS. (a) Pre-training: Each graph  $\mathcal{G}$  in the training set is augmented by sampling augmentation pairs from a large pool, guided by the BiAS strategy. (To be clear, we only show one training graph here; the same augmentation process is applied to every graph in the training set.) (b) Inference: The pre-trained model is frozen and applied to unseen graph pairs.

augmentation pair according to the matching accuracy between the two corresponding augmented views. This adjustment influences the sampling probability for subsequent mini-batches. By doing so, BiAS adaptively samples challenging augmentations tailored for GM, mitigating the necessity for exhaustive hyperparameter tuning and augmentation pool selection.

**Inference** In the inference phase, as shown in Fig. 1(b), we apply the pre-trained, now-frozen graph encoder and affinity learning module to unseen graph pairs. The GM solver predicts the matching matrix, which determines the node-to-node alignment for each input graph pair.

We will next introduce some technical preliminaries (Sect. 3.2), before discussing our main contributions on the comprehensive pool of graph augmentations (Sect. 3.3) and our adaptive sampler BiAS (Sect. 3.4), respectively.

### 3.2 Preliminaries

**Graph Encoder** During training, our proposed graph encoder takes an augmented graph  $\tilde{\mathcal{G}}$  as input to produce both node and graph-level representations as output. Given the comprehensive graph augmentations we carry out, it is essential to leverage information at multiple granularities to preserve both local and global graph patterns. To achieve this, we have opted for the commonly used GraphSAGE framework [Hamilton *et al.*, 2017] as our base graph encoder, denoted as  $f(\cdot)$  in Fig. 1. This encoder captures both local node-level embedding  $\mathbf{h}_v$ , using GraphSAGE layers with skip connections, and global graph-level embedding  $\tilde{\mathbf{h}}_{\tilde{\mathcal{G}}}$  via a readout function. The node- and graph-level embeddings are then utilized for the affinity learning and GM solver modules, to be elaborated next. For detailed specifications of the encoding function  $f(\cdot)$ , refer to Appendix A.

**Affinity Learning and GM Solver** On one hand, the affinity learning module accepts these embeddings from the two views (graphs) as input and generates an affinity matrix, containing the pairwise affinity scores between the nodes from the two views. Subsequently, the GM solver leverages the affinity matrix to determine the node correspondence across

the two graphs. We will use the affinity learning module and GM solver following previous work [Rolínek *et al.*, 2020; Wang *et al.*, 2021].

**Contrastive Loss** After processing the two augmented views of the graph through our encoder, as depicted in Fig. 1(a), we employ a contrastive framework to learn node embeddings. Adapting from GRACE [Zhu *et al.*, 2020], we employ both intra-view and inter-view contrastive objectives to enhance the comparison and alignment of nodes between distinct graph views.

First, the intra-view contrastive objective aims to differentiate anchor nodes from other nodes within the same view by maximizing their dissimilarity. Specifically, in view  $\tilde{\mathcal{G}}^A$  with  $N_A$  nodes, the dissimilarity between the anchor node  $v_i^A$  with feature  $\tilde{\mathbf{h}}_i^A$  and the negative nodes is

$$\ell_{\text{intra}}^A(v_i) = \sum_{j=1; j \neq i}^{N_A} \exp\left(\text{sim}\left(g(\tilde{\mathbf{h}}_i^A), g(\tilde{\mathbf{h}}_j^A)\right) / T\right), \quad (1)$$

where  $\text{sim}(\cdot, \cdot)$  denotes cosine similarity,  $g(\cdot) : \mathbb{R}^{d_h} \rightarrow \mathbb{R}^{d'_h}$  is a non-linear projection head, and  $T$  is the temperature hyperparameter. This approach ensures that distinct nodes within the same view become dissimilar.

Second, our inter-view contrastive objective, operating across distinct augmented views  $\tilde{\mathcal{G}}^A$  and  $\tilde{\mathcal{G}}^B$ , maximizes the similarity of corresponding nodes across views while minimizing that of an anchor node and all other nodes in the opposite view. We first sum up the similarities across all inter-view pairs, centered around the anchor node in view  $\tilde{\mathcal{G}}^A$ :

$$\ell_{\text{inter}}^A(v_i) = \sum_{j=1}^{N_B} \exp\left(\text{sim}\left(g(\tilde{\mathbf{h}}_i^A), g(\tilde{\mathbf{h}}_j^B)\right) / T\right). \quad (2)$$

Integrating both intra- and inter-view objectives based on the NT-Xent loss [Chen *et al.*, 2020], the contrastive loss for an anchor node is

$$\ell_i^A = -\log \frac{\exp(\text{sim}(g(\tilde{\mathbf{h}}_i^A), g(\tilde{\mathbf{h}}_i^B)) / T)}{\ell_{\text{intra}}^A(v_i) + \ell_{\text{inter}}^A(v_i)}. \quad (3)$$

The total contrastive loss averages across all anchor nodes in the two views:

$$\mathcal{L}_{\text{node}} = \frac{1}{N_A + N_B} \left( \sum_{i=1}^{N_A} \ell_i^A + \sum_{i=1}^{N_B} \ell_i^B \right). \quad (4)$$

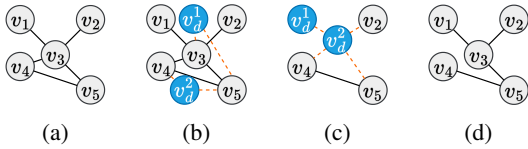


Figure 2: Graph augmentation: (a) input graph; (b) node insertion; (c) node replacement; (d) edge removal. The blue node represents the inserted node, and the dotted edge indicates the added edge.

In essence, our contrastive strategy attempts to align corresponding nodes in the two views to achieve graph matching.

**Graph Matching Loss** While our contrastive loss refines individual node representations, our approach recognizes the importance of holistic graph matching. To this end, we incorporate a graph matching loss  $\mathcal{L}_{\text{match}}$ . This loss bridges the gap between our predicted matching matrix,  $\hat{\mathbf{G}}$ , and a self-labeled ground-truth matching matrix  $\mathbf{G}^{\text{self}}$ .  $\mathbf{G}^{\text{self}}$  can be readily determined without supervision, as  $\tilde{\mathcal{G}}^A$  and  $\tilde{\mathcal{G}}^B$  originate from the same input graph.

$$\mathcal{L}_{\text{match}} = \mathcal{L}(\hat{\mathbf{G}}, \mathbf{G}^{\text{self}}) \quad (5)$$

The specific GM solver used determines the form of the loss: a permutation loss for NGM [Wang *et al.*, 2021], or a Hamming loss for BBGM [Rolínek *et al.*, 2020].

Our overall loss  $\mathcal{L} = \mathcal{L}_{\text{node}} + \mathcal{L}_{\text{match}}$  combines contrastive loss and graph matching loss, ensuring that the overall graph representations are optimized for graph matching tasks.

### 3.3 Comprehensive Graph Augmentations

Appropriate graph augmentations are crucial for the success of contrastive learning toward general graph matching. As illustrated in Fig. 1(a), we propose to select from a comprehensive pool of augmentation pairs to generate two views for each input graph. While existing methods often adopt a small number of standard augmentations, we argue that, in the absence of side information for graph matching, we need to formulate a comprehensive pool of augmentations for robustness. In particular, we not only focus on a diverse range of augmentation types to improve different aspects of matching, but also expand each augmentation type into a family of augmentations with varying difficulty levels. In the following, we first introduce different types of augmentations; then, we discuss the construction of the pool of paired augmentations for contrastive learning.

**Types of Augmentation** We introduce four major graph augmentation types to robustly handle real-world matching scenarios, most notably the complexities in managing outliers, noises [Cho *et al.*, 2014] and high-order structures. The first three types operate in the structure space, while the last one operates in the feature space.

**Node Insertion (NI).** Recognizing outliers, or nodes that are exclusive to only one of the graph in a candidate matching pair, is a crucial aspect for graph matching. Inspired by node addition in contrastive learning for graph classification [Zeng and Xie, 2021], we propose a node insertion strategy to simulate outliers. As shown in Fig. 2(b), given a randomly

selected subset of nodes from the original input graph, we aggregate their features to construct a dummy node (e.g.,  $v_d^1$ ) and randomly link it to randomly chosen nodes not in the given subset. The augmented graph view with the dummy nodes can provide contrastive samples to learn outliers.

**Node Replacement (NR).** This is a variant of node insertion, to simulate the matching scenario where there can still be outliers even though the total number of nodes are the same in a candidate pair. As illustrated in Fig. 2(c), we first drop some nodes randomly from the original graph (e.g.,  $v_1, v_3$ ) along with their incident edges, and then insert an equal number of dummy nodes (e.g.,  $v_d^1, v_d^2$ ) following the node insertion strategy above.

**Edge Removal (ER).** In many real-world graphs, first-order connections (i.e., edges to one-hop neighbors) are often noisy or sparse. This limitation can be alleviated by considering high-order structural connectivity or node features. To recognize node correspondences without overly relying on first-order connections, we adopt an edge removal strategy [Hasani and Khasahmadi, 2020]. As shown in Fig. 2(d), a subset of edges (e.g.,  $v_3-v_4$ ) is randomly removed from the graph, which encourages the model to focus more on the high-order connections (e.g.,  $v_3-v_5-v_4$ ) or invariant features.

**Feature Scaling (FS).** Node features in graphs can exhibit significant variability and contain considerable noises. To address this, we present two feature scaling techniques [Laskin *et al.*, 2020]. First, node features are scaled using a univariate random variable, such that all dimensions of the feature vector of the same node are scaled by the same factor. Second, we apply a multivariate random variable, such that different scaling factors can be applied to different dimensions.

In addition to above four major types of augmentations designed to deal with different matching scenarios, we further adopt a series of standard graph augmentations aimed to capture general graph properties that are widely used in previous work. We summarize the four major of augmentations used in our approach in Tab. 1; the additional types of augmentation are in Appendix B. Note that each type of augmentation is also associated with one or more hyperparameters, adjusting which can instantiate different augmentations with variable difficulty. Additionally, we introduce a special type  $\tau_{\emptyset}$  that does nothing, which enables us to generate a view using the the original graph.

Formally, let  $\mathcal{T}$  denote the set of augmentation types including  $\tau_{\emptyset}$ . Further define  $\mathcal{I}_{\tau}$  as the set of augmentations that can be instantiated from type  $\tau \in \mathcal{T}$ , by sampling different hyperparameter values associated with  $\tau$ . Hence, the full set of augmentations under consideration is  $\mathcal{I} = \cup_{\tau \in \mathcal{T}} \mathcal{I}_{\tau}$ .

**Pool of Paired Augmentations** As illustrated in Fig. 1(a)(i), given an input graph, we generate two augmented views of the graph by applying a pair of augmentations. Existing graph contrastive learning approaches mostly employ well-tuned hyperparameters for every augmentation type. However, given a large number of augmentation types and associated hyperparameters, extensive tuning can be combinatorially costly and tends to overfit. Moreover, recent research shows [Cai *et al.*, 2020] that training can often benefit from samples of variable difficulty level. In our

Type	Matching scenario	Hyperparameters
NI	node outlier (unequal node count in two views)	$p_{nr} \in [0.1, 0.9]$ : fraction of nodes inserted; $k_{ni} \geq 2$ : size of subset; $agg_{r_{ni}} \in \{\text{mean, max}\}$ : aggregation function; $e_{ni} \geq 1$ : number of edges inserted
NR	node outlier (equal node count in two views)	$p_{nr} \in [0.1, 0.9]$ : fraction of nodes replaced; $k_{nr} \geq 2$ : size of subset; $agg_{r_{nr}} \in \{\text{mean, max}\}$ : aggregation function; $e_{nr} \geq 1$ : number of edges inserted
ER	sparse/noisy first-order connections	$p_{er} \in [0.1, 0.9]$ : probability of each edge being removed
FS	feature variations & noises	$\alpha \in [0.2, 0.8]$ : lower bound of uniform distribution; $\beta \in [1.2, 1.8]$ : upper bound of uniform distribution

Table 1: Details of the four major types of graph augmentation.

case, as the difficulty of augmentation is controlled by its associated hyperparameters, it is not ideal to only rely on one set of tuned hyperparameters.

In our approach, given an input graph, we propose to apply a pair of augmentations sampled from  $\mathcal{P}$ , the pool of all possible pairs. Most generally, we can define  $\mathcal{P} = \mathcal{I}^2$ , i.e., the Cartesian product of  $\mathcal{I}$  with itself, where  $\mathcal{I}$  is the set of all augmentations. However, we exclude two special cases: (1) Both augmentations in the pair belong to the  $\tau_{\emptyset}$  type, since both views being the original graph does not form a meaningful contrastive example; (2) both belong to the mixup type following earlier work [Liu *et al.*, 2022] which tends to give poor empirical performance.

### 3.4 Sampling Augmentation Pairs via BiAS

Given a pool of augmentation pairs  $\mathcal{P}$ , we sample a pair to be applied to each input graph to form the contrastive example for training. Adopting a uniform sampler is the most straightforward approach, but its disadvantage is twofold. First, it does not distinguish the importance of different augmentations; second, it is static and task irrelevant, which does not adapt to the fluctuations in training.

In our GCGM framework, we propose BiAS, a **Boosting-inspired Adaptive Augmentation Sampler**. As the name suggests, BiAS is adaptive to training fluctuations, inspired by the boosting technique [Schapire, 1999] in ensemble learning that iteratively refines weak learners by focusing on more difficult examples. Specifically, augmentation pairs that result in subpar matching performance in a mini-batch during training are marked as difficult, suggesting that the model needs further adaptation to data variations introduced by such augmentations. These challenging augmentation pairs subsequently receive more frequent sampling by increasing their probability. Conversely, pairs that can be solved effortlessly with high accuracy are deemed easy and are less frequently sampled in the following mini-batches. Essentially, this strategy biases toward certain augmentations dynamically, addressing the drawbacks of a uniform sampler.

Formally, we introduce a weight updating scheme to dynamically adjust the sampling probability of each augmentation pair from the pool  $\mathcal{P}$ , as illustrated in Fig. 1(a)(ii). Specifically, each augmentation pair  $i \in \mathcal{P}$  is associated with a weight  $w_t^i$  in the  $t$ -th mini-batch. The weight is updated in

the next mini-batch as follows.

$$w_{t+1}^i = \lambda w_t^i + (1 - \lambda)e^{\alpha(1 - \phi_t^i)}, \quad (6)$$

where  $\phi_t^i$  indicates the mean performance score for all matchings where augmentation pair  $i$  was previously applied up to the current mini-batch  $t$ . In our implementation, we adopt the  $F_1$  score as the performance score (see Appendix E), which measures the accuracy of the predicted matching against the contrastive self-supervision generated from the two graph views. The hyperparameter  $\alpha \geq 1$  regulates the magnitude of weight adjustment, while  $\lambda \in [0, 1]$  serves the purpose of momentum update to constrain the potentially exponential increase in weights of those very challenging pairs yielding low  $\phi_t^i$  repeatedly. The weights in the  $t$ -th mini-batch are then normalized into a probability distribution over the pool of augmentation pairs  $\mathcal{P}$ , given by  $P_t(i) = w_t^i / \sum_{j \in \mathcal{P}} w_t^j$ . The initial weight  $w_0^i$  is set to  $e^\alpha$  for all augmentation pairs, i.e., we start with a uniform sampler.

Finally, given an input graph in the  $t$ -th mini-batch, we generate two views using an augmentation pair  $i \sim P_t(i)$  drawn from the distribution over the pool  $\mathcal{P}$ . Pseudocode of our method can be found in Appendix A.

In summary, BiAS offers an adaptive strategy to choose the right augmentations. Compared to the uniform sampling strategy, BiAS dynamically selects the most useful augmentation pairs in each mini-batch during training. Compared to the conventional approach of tuning the hyperparameters of each augmentation type, BiAS is much cheaper and can benefit from a wider range of augmentations.

## 4 Experiment

In this section, we empirically evaluate the proposed model GCGM and BiAS.

### 4.1 Experiment Setup

**Datasets** We tested three real-world datasets. (1) Pascal VOC [Bourdev and Malik, 2009; Everingham *et al.*, 2010] includes images from 20 classes; (2) Willow [Cho *et al.*, 2013] offers 256 images over five classes; (3) SPair-71k [Min *et al.*, 2019] has 70,958 image pairs across 18 classes. Besides, we followed a recent work [Liu *et al.*, 2023] to generate a synthetic dataset from random 2D node coordinates for the general non-visual domain. All graphs are constructed based on Delaunay triangulation. More dataset details are presented in Appendix C.

**Baselines** We assessed the performance of GCGM against diverse baselines, including supervised, learning-free, and unsupervised methods. Our emphasis is benchmarking against graph-centric self-supervised methods, given our exclusion of labeled data or side information. Supervised methods included cutting-edge techniques like CIE [Yu *et al.*, 2019], BBGM [Rolínek *et al.*, 2020], and NGMv2 [Wang *et al.*, 2021]. Learning-free methods included RRWM [Cho *et al.*, 2010], IFPF [Leordeanu *et al.*, 2009], and SM [Leordeanu and Hebert, 2005]. For self-supervised baselines, we compared against GANN-GM [Wang *et al.*, 2023] and SCGM [Liu *et al.*, 2022].

Methods	Pascal VOC		Willow Intsec	SPair-71k	
	Intsec	Unfilt		Intsec	Unfilt
CIE (SUP)	66.8±0.4	-	82.6±0.2	69.3±0.3	-
BBGM (SUP)	77.3±0.1	55.0±0.1	96.2±0.1	77.7±0.2	48.4±0.2
NGMv2 (SUP)	76.8±0.1	56.7±0.1	94.5±0.3	76.6±0.2	49.8±0.08
IPFP	45.8±0.02	31.5±0.002	80.1±0.06	57.0±0.04	31.7±0.01
RRWM	47.2±0.02	31.7±0.001	83.4±0.09	58.6±0.05	32.3±0.01
SM	46.2±0.03	30.4±0.002	81.3±0.08	57.7±0.04	30.3±0.01
GANN-GM <sup>ˆ</sup>	34.5±0.3	23.4±0.2	89.3±0.1	34.7±0.4	19.4±0.3
SCGM+BBGM	54.8±0.05	<u>36.6±0.04</u>	93.1±0.08	60.2±0.05	34.1±0.01
SCGM+NGMv2	50.8±0.1	32.9±0.03	84.2±0.1	59.8±0.1	30.5±0.3
GCGM+BBGM	<u>56.8±0.02</u>	36.2±0.01	94.4±0.3	<u>60.6±0.1</u>	<u>35.9±0.07</u>
GCGM+NGMv2	<u>57.3±0.11</u>	<u>37.4±0.07</u>	<u>95.0±0.1</u>	<u>62.6±0.02</u>	<u>35.4±0.07</u>

Table 2: Performance (%) on real-world datasets. Supervised methods are marked with ‘SUP’, included for reference only. -: running exception in handling the matching scenario. <sup>ˆ</sup>: self-supervised methods that require categorical information. Bold/underlined: best/runner-up results (excluding supervised methods).

**Settings and Hyperparameters** GCGM with BiAS is a generic framework that can be applied to any GM solver. To demonstrate this, we paired them with both BBGM [Rolínek *et al.*, 2020] and NGMv2 [Wang *et al.*, 2021]. As GCGM is graph-centric, we froze the image backbone on the visual datasets, and treated each input image solely as a graph. Our implementation is based on ThinkMatch [Wang *et al.*, 2021], and we kept the original configurations for both solvers. Our reported results for supervised methods and SCGM might be slightly lower than their original papers due to our 80:20 train-validation split from the original training set, as the original splits lack a validation set. We repeated the splits five times using varied random seeds. We tuned the hyperparameters of our encoder using Tree-structured Parzen Estimator (TPE) [Bergstra *et al.*, 2011] via Optuna [Akiba *et al.*, 2019] based on the validation set. For BiAS, we set  $\lambda = 0.8$ ,  $\alpha = 3$ ,  $|\mathcal{P}| = 512$ . However, for the Willow dataset, due to its smaller size, we adjust  $|\mathcal{P}|$  to 128. Early stopping was applied if performance improvements were below the threshold  $\epsilon = 0.001$ . Detailed model and parameter configurations can be found in Appendix A.

For other supervised and SSL baselines, we also applied the same early stopping criterion. On visual datasets, we allowed them to fine-tune their visual backbone and employ visual augmentations (if any) per their original paper. More detailed baseline settings can be found in Appendix D.

**Evaluation Metric** We used the  $F_1$  score for evaluation (see Appendix E), combining precision and recall for a balanced measure, which is vital for matching graphs with outliers. We repeated the experiments on the five different splits described above, and report their average and standard deviation in  $F_1$ .

## 4.2 Performance Evaluation

In our experiments, we adopted the two pairwise GM settings outlined by SCGM. The *Intersection* (Intsec) setting confines keypoints (i.e., nodes) to only those shared between the input graph pair. Meanwhile, the *Unfiltered* (Unfilt) setting includes all keypoints, accounting for potential outliers and variations in keypoint counts between graphs. Notably, the

Methods	Synthetic	
	Intsec	Unfilt
GANN-GM <sup>ˆ</sup>	11.2±0.04	10.2±0.03
SCGM + BBGM	33.5±2.0	24.3±1.2
SCGM + NGMv2	35.2±0.6	25.0±0.4
GCGM	<b>58.1±0.5</b>	<b>39.9±0.4</b>

Table 3: Performance (%) of SSL methods on the synthetic dataset.

*Unfiltered* setting is more challenging because of the presence of the outliers. Note that in the case of the Willow dataset, the *Intersection* and *Unfiltered* scenarios are equivalent due to each graph class containing 10 shared keypoints (i.e., nodes) without outliers. Consequently, we only report the *Intersection* results. For training, we used the original graph without any filtering, retaining all the keypoints.

**Real-world Datasets** In Tab. 2, we present the performance evaluation on the visual datasets (see detailed per-class results in Appendix F). We make several observations.

First, in terms of overall performance, GCGM consistently surpasses other SSL baselines and learning-free methods. On the Willow dataset, GCGM outperforms even some supervised methods. This result highlights the effectiveness of our contrastive approach, especially given that the baselines are allowed to fine-tune the visual backbone and apply image augmentations (if applicable), whereas our GCGM is graph-centric and does not rely on these visual elements.

Second, when paired with the same GM solver, GCGM consistently outperforms the state-of-the-art SCGM. This finding highlights the flexibility and robustness of GCGM when working with different GM solvers. Specifically, when pairing with NGMv2, a GNN-based solver, our approach tends to perform the best. For brevity, in subsequent results, we default to reporting only the performance of GCGM using the NGMv2 solver.

Third, GCGM demonstrates solid performance in both *Intersection* and *Unfiltered* settings. In particular, the *Unfiltered* setting is often more challenging given the presence of outliers. Hence, all methods give lower performance in this setting, while GCGM continues to maintain a clear advantage.

**Synthetic Dataset** The performance of various SSL methods on the synthetic dataset is presented in Tab. 3, providing a focused and fair comparison since our GCGM is self-supervised. Note that supervised methods (CIE, BBGM and NGMv2) and learning-free methods (IPFP, RRWM and SM) as shown in Tab. 2 are for reference only; we report their results on the synthetic dataset in Appendix F.

In this dataset, GCGM excels over other SSL methods in both *Intersection* and *Unfiltered* settings. It is worth noting that the performances of SCGM and GANN-GM significantly trail behind GCGM, widening the gap observed on the visual datasets. The reason is that the synthetic dataset does not require visual backbone processing or visual input, making SCGM and GANN-GM unable to exploit the visual aspects in their original design. Hence, they tend to perform poorly in general graph matching outside the visual domain. In contrast, our GCGM is graph-centric and does not rely on vi-

Augmentation Set	Pascal VOC		Willow	SPair-71k		Synthetic	
	Intsec	Unfilt	Intsec	Intsec	Unfilt	Intsec	Unfilt
$\mathcal{T} \setminus \text{NI}$	56.9	36.6	94.8	61.8	34.9	57.9	<b>40.5</b>
$\mathcal{T} \setminus \text{NR}$	56.5	36.5	<b>95.1</b>	61.8	34.4	57.8	40.0
$\mathcal{T} \setminus \text{ER}$	57.3	37.2	95.0	59.8	32.5	57.9	40.0
$\mathcal{T} \setminus \text{FS}$	<b>57.5</b>	37.2	95.0	62.1	35.1	57.8	40.3
$\mathcal{T}$	57.3	<b>37.4</b>	95.0	<b>62.6</b>	<b>35.4</b>	<b>58.1</b>	39.9

Table 4: Ablation study on graph augmentations.

sual information, leading to its effectiveness and robustness in general graph matching.

### 4.3 Ablation and Model Analyses

**Effect of Graph Augmentations** We explored the contribution of different types of graph augmentation, by excluding each type from the augmentation set  $\mathcal{T}$ . As observed in Tab. 4, Node Insertion (NI) and Node Replacement (NR) generally have a significant impact on the performance, particularly on Pascal VOC and SPair-71k. For the synthetic dataset, the two augmentation types appear not useful. A potential reason is that the node features are generated randomly, which means the feature differences between outliers and inliers are less significant. Thus, NI and NR can help less on the synthetic dataset. On the other hand, excluding Edge Removal (ER) and Feature Scaling (FS) generally results in a slight performance dip. In particular, ER appears to be the most useful to SPair-71k, due to the significant variability in viewpoint and scale in this dataset, as compared to other datasets [Min *et al.*, 2019]. This variability leads to diverse geometric configurations of keypoints across images, which impacts and alters the graph structures formed. By randomly removing edges, ER encourages the model to focus on invariant features over variable graph structures, which is important on SPair-71k.

**Effect of BiAS and Augmentation Pool** Tab. 5 compares the performance of GCGM under different strategies for selecting augmentations. In the first category including ‘Random’, ‘Tuning’ and ‘Tuning + BiAS’, we only adopt one hyperparameter setting for each augmentation type. In ‘Random’, the hyperparameters of an augmentation type are initialized randomly, whereas in ‘Tuning’ or ‘Tuning + BiAS’, they are selected via 100 trials of TPE based on validation performance. Furthermore, ‘Random’ and ‘Tuning’ applies a uniform sampler when selecting the augmentation pair in each mini-batch, whereas ‘Tuning + BiAS’ applies our BiAS sampler. In the second category including ‘Uniform’ and ‘BiAS’, we do not rely on just one set of hyperparameters for each augmentation type. As discussed in Sect. 3.3, we instantiate a family of augmentations from each type with different hyperparameter settings. In each mini-batch, we select an augmentation pair from all pairs of instantiations using either the uniform or BiAS sampler.

In terms of performance, we observe that ‘Tuning’ generally outperforms ‘Random’, implying that hyperparameter settings are important to each augmentation type. Moreover, ‘Tuning + BiAS’ generally improves over ‘Tuning’, and

Settings	$\mathcal{P}$	Pascal VOC			Willow		SPair-71k		
		Intsec	Unfilt	Time/h	Intsec	Time/h	Intsec	Unfilt	Time/h
Random	$\times$	55.0	35.9	0.26	93.8	0.04	61.4	35.2	0.38
Tuning	$\times$	55.9	36.8	23.76	94.8	4.54	61.5	35.6	31.96
Tuning + BiAS	$\times$	55.8	37.0	23.95	<b>95.4</b>	4.54	61.9	<b>36.0</b>	31.88
Uniform	$\checkmark$	56.9	36.7	0.32	94.7	0.05	62.0	34.8	0.35
BiAS	$\checkmark$	<b>57.3</b>	<b>37.4</b>	0.39	95.0	0.05	<b>62.6</b>	35.4	0.34

Table 5: Performance of different initialization of augmentations and the use of augmentation pool. Time (hour) represents the total wall clock time spent on tuning the augmentations (for ‘Tuning’ methods) and training the model. The ‘ $\mathcal{P}$ ’ column indicates if a pool of augmentation pairs is used.

‘BiAS’ consistently improves over ‘Uniform’, highlighting the effectiveness of the BiAS sampler. Furthermore, using a large pool of augmentation pairs can compensate for the lack of hyperparameter tuning, as evident from the improvements made by ‘Uniform’ and ‘BiAS’ w.r.t. ‘Random’.

In terms of time cost, ‘Tuning’ and ‘Tuning + BiAS’ fare poorly due to the need of hyperparameter tuning. Yet, their performance rarely surpasses BiAS. On the other hand, the time cost of BiAS is two orders of magnitude smaller than tuning-based methods, and simultaneously achieves the best or near-best performance. It is also worth-noting that the weight updating strategy in BiAS incurs a negligible overhead, taking similar training time as ‘Random’ or ‘Uniform’ while achieving better results.

In summary, BiAS offers a robust solution for selecting optimal augmentations, yet without the significant time overhead required by the tuning-based methods.

**Additional Model Analyses** Additionally, we experimented our framework with varying sizes of the augmentation pool and different combinations of the hyperparameter  $\lambda$  and  $\alpha$ . We also investigated the effect of the individual design elements in BiAS, namely, momentum update and performance criterion. Finally, we compared GCGM with SCGM when the level of input information varies, to show the robustness of our approach given minimum input. Due to space constraint, we omit these results here, and present them in Appendix F.

## 5 Conclusion

We introduced the Graph-centric Contrastive framework for Graph Matching (GCGM), a novel framework toward general graph matching without the need of any side information. GCGM utilizes a comprehensive pool of graph augmentations for self-supervised contrastive learning, which enhances matching robustness in the absence of side information. It is further complemented with a Boosting-inspired Adaptive Augmentation Sampler (BiAS), which dynamically selects the challenging augmentations for optimal results without tuning the hyperparameters associated with the augmentations. Together, through our experiments GCGM with BiAS achieves superior performance in graph matching, surpassing state-of-the-art self-supervised methods across various domains. At the same time, it is significantly more efficient than conventional augmentations with hyperparameter tuning.



## Acknowledgements

This research project is supported by the Ministry of Education, Singapore, under its Academic Research Fund Tier 2 (Proposal ID: T2EP20122-0041). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the Ministry of Education, Singapore.

## References

- [Akiba *et al.*, 2019] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.
- [Bergstra *et al.*, 2011] James Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for hyper-parameter optimization. In *Advances in Neural Information Processing Systems*, volume 24, 2011.
- [Bourdev and Malik, 2009] L. Bourdev and J. Malik. Poselets: Body part detectors trained using 3d human pose annotations. In *International Conference on Computer Vision*, pages 1365–1372, 2009.
- [Cai *et al.*, 2020] Tiffany Tianhui Cai, Jonathan Frankle, David J Schwab, and Ari S Morcos. Are all negatives created equal in contrastive instance discrimination? *arXiv preprint arXiv:2010.06682*, 2020.
- [Chen *et al.*, 2020] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International Conference on Machine Learning*, pages 1597–1607, 2020.
- [Cho *et al.*, 2010] Minsu Cho, Jungmin Lee, and Kyoung Mu Lee. Reweighted random walks for graph matching. In *European Conference on Computer Vision, Part V*, pages 492–505, 2010.
- [Cho *et al.*, 2013] Minsu Cho, Karteek Alahari, and Jean Ponce. Learning graphs to match. In *International Conference on Computer Vision*, pages 25–32, 2013.
- [Cho *et al.*, 2014] Minsu Cho, Jian Sun, Olivier Duchenne, and Jean Ponce. Finding matches in a haystack: A max-pooling strategy for graph matching in the presence of outliers. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2083–2090, 2014.
- [Everingham *et al.*, 2010] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88:303–338, 2010.
- [Guo *et al.*, 2023] Wenqi Guo, Lin Zhang, Shikui Tu, and Lei Xu. Self-supervised bidirectional learning for graph matching. In *AAAI Conference on Artificial Intelligence*, volume 37, pages 7784–7792, 2023.
- [Hamilton *et al.*, 2017] Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems*, volume 30, 2017.
- [Hassani and Khasahmadi, 2020] Kaveh Hassani and Amir Hosein Khasahmadi. Contrastive multi-view representation learning on graphs. In *International Conference on Machine Learning*, pages 4116–4126, 2020.
- [Krissinel and Henrick, 2004] Eugene Krissinel and Kim Henrick. Secondary-structure matching (ssm), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica Section D: Biological Crystallography*, 60(12):2256–2268, 2004.
- [Laskin *et al.*, 2020] Misha Laskin, Kimin Lee, Adam Stooke, Lerrel Pinto, Pieter Abbeel, and Aravind Srivas. Reinforcement learning with augmented data. In *Advances in Neural Information Processing Systems*, volume 33, pages 19884–19895, 2020.
- [Leordeanu and Hebert, 2005] Marius Leordeanu and Martial Hebert. A spectral technique for correspondence problems using pairwise constraints. In *IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 2, pages 1482–1489, 2005.
- [Leordeanu *et al.*, 2009] Marius Leordeanu, Martial Hebert, and Rahul Sukthankar. An integer projected fixed point method for graph matching and map inference. In *Advances in Neural Information Processing Systems*, volume 22, 2009.
- [Li *et al.*, 2023] Jintang Li, Ruofan Wu, Wangbin Sun, Liang Chen, Sheng Tian, Liang Zhu, Changhua Meng, Zibin Zheng, and Weiqiang Wang. What’s behind the mask: Understanding masked graph modeling for graph autoencoders. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 1268–1279, 2023.
- [Liu *et al.*, 2022] Chang Liu, Shaofeng Zhang, Xiaokang Yang, and Junchi Yan. Self-supervised learning of visual graph matching. In *European Conference on Computer Vision*, pages 370–388, 2022.
- [Liu *et al.*, 2023] Chang Liu, Zetian Jiang, Runzhong Wang, Lingxiao Huang, Pinyan Lu, and Junchi Yan. Revocable deep reinforcement learning with affinity regularization for outlier-robust graph matching. In *International Conference on Learning Representations*, 2023.
- [Min *et al.*, 2019] Juhong Min, Jongmin Lee, Jean Ponce, and Minsu Cho. Spair-71k: A large-scale benchmark for semantic correspondence. *arXiv preprint arXiv:1908.10543*, 2019.
- [Peng *et al.*, 2020] Zhen Peng, Wenbing Huang, Minnan Luo, Qinghua Zheng, Yu Rong, Tingyang Xu, and Junzhou Huang. Graph representation learning via graphical mutual information maximization. In *ACM The Web Conference*, pages 259–270, 2020.
- [Robinson *et al.*, 2021] Joshua Robinson, Ching-Yao Chuang, Suvrit Sra, and Stefanie Jegelka. Contrastive learning with hard negative samples. In *International Conference on Learning Representations*, 2021.

- [Rolínek *et al.*, 2020] Michal Rolínek, Paul Swoboda, Dominik Zietlow, Anselm Paulus, Vít Musil, and Georg Martius. Deep graph matching via blackbox differentiation of combinatorial solvers. In *European Conference on Computer Vision, Part XXVIII*, pages 407–424, 2020.
- [Schapire, 1999] Robert E. Schapire. A brief introduction to boosting. In *International Joint Conference on Artificial Intelligence*, page 1401–1406, San Francisco, CA, USA, 1999. Morgan Kaufmann Publishers Inc.
- [Shokoufandeh *et al.*, 2012] A Shokoufandeh, Y Keselman, M Fatih Demirci, D Macrini, and Sven Dickinson. Many-to-many feature matching in object recognition: a review of three approaches. *IET Computer Vision*, 6(6):500–513, 2012.
- [Tan *et al.*, 2022] Zhen Tan, Kaize Ding, Ruocheng Guo, and Huan Liu. Supervised graph contrastive learning for few-shot node classification. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 394–411, 2022.
- [Veličković *et al.*, 2017] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.
- [Wang and Liu, 2021] Feng Wang and Huaping Liu. Understanding the behaviour of contrastive loss. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2495–2504, 2021.
- [Wang *et al.*, 2018] Qianqian Wang, Xiaowei Zhou, and Kostas Daniilidis. Multi-image semantic matching by mining consistent features. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 685–694, 2018.
- [Wang *et al.*, 2020] Runzhong Wang, Junchi Yan, and Xiaokang Yang. Combinatorial learning of robust deep graph matching: an embedding based approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(6):6984–7000, 2020.
- [Wang *et al.*, 2021] Runzhong Wang, Junchi Yan, and Xiaokang Yang. Neural graph matching network: Learning lawler’s quadratic assignment problem with extension to hypergraph and multiple-graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9):5261–5279, 2021.
- [Wang *et al.*, 2023] Runzhong Wang, Junchi Yan, and Xiaokang Yang. Unsupervised learning of graph matching with mixture of modes via discrepancy minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8):10500–10518, 2023.
- [Wu *et al.*, 2018] Zhirong Wu, Yuanjun Xiong, Stella X Yu, and Dahua Lin. Unsupervised feature learning via non-parametric instance discrimination. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3733–3742, 2018.
- [Xu *et al.*, 2022] Xovee Xu, Fan Zhou, Kunpeng Zhang, and Siyuan Liu. Ccgl: Contrastive cascade graph learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(5):4539–4554, 2022.
- [Yan *et al.*, 2016] Junchi Yan, Xu-Cheng Yin, Weiyao Lin, Cheng Deng, Hongyuan Zha, and Xiaokang Yang. A short survey of recent advances in graph matching. In *ACM on International Conference on Multimedia Retrieval*, pages 167–174, 2016.
- [You *et al.*, 2020] Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, and Yang Shen. Graph contrastive learning with augmentations. In *Advances in Neural Information Processing Systems*, volume 33, pages 5812–5823, 2020.
- [You *et al.*, 2021] Yuning You, Tianlong Chen, Yang Shen, and Zhangyang Wang. Graph contrastive learning automated. In *International Conference on Machine Learning*, pages 12121–12132, 2021.
- [Yu *et al.*, 2019] Tianshu Yu, Runzhong Wang, Junchi Yan, and Baoxin Li. Learning deep graph matching with channel-independent embedding and hungarian attention. In *International Conference on Learning Representations*, 2019.
- [Zanfir and Sminchisescu, 2018] Andrei Zanfir and Cristian Sminchisescu. Deep Learning of Graph Matching. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2684–2693, 2018.
- [Zaslavskiy *et al.*, 2009] Mikhail Zaslavskiy, Francis Bach, and Jean-Philippe Vert. Global alignment of protein-protein interaction networks by graph matching methods. *Bioinformatics*, 25(12):i259–i267, 2009.
- [Zeng and Xie, 2021] Jiaqi Zeng and Pengtao Xie. Contrastive self-supervised learning for graph classification. In *AAAI Conference on Artificial Intelligence*, volume 35, pages 10824–10832, 2021.
- [Zhang *et al.*, 2018] Jing Zhang, Bo Chen, Xianming Wang, Hong Chen, Cuiping Li, Fengmei Jin, Guojie Song, and Yutao Zhang. Mego2vec: Embedding matched ego networks for user alignment across social networks. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 327–336, 2018.
- [Zhao *et al.*, 2021] Tong Zhao, Yozen Liu, Leonardo Neves, Oliver Woodford, Meng Jiang, and Neil Shah. Data augmentation for graph neural networks. In *AAAI Conference on Artificial Intelligence*, volume 35, pages 11015–11023, 2021.
- [Zhu *et al.*, 2020] Yanqiao Zhu, Yichen Xu, Feng Yu, Qiang Liu, Shu Wu, and Liang Wang. Deep graph contrastive representation learning. *arXiv preprint arXiv:2006.04131*, 2020.
- [Zhu *et al.*, 2021] Yanqiao Zhu, Yichen Xu, Feng Yu, Qiang Liu, Shu Wu, and Liang Wang. Graph Contrastive Learning with Adaptive Augmentation. In *ACM The Web Conference*, pages 2069–2080, April 2021.