Improving Subgraph Representation Learning via Multi-View Augmentation

Yili Shen [1 2] **Jiaxu Yan** [1 3] **Cheng-Wei Ju** [1 4] **Jun Yi** [1] **Zhou Lin** [1] **Hui Guan** [5]

Abstract

Subgraph representation learning based on Graph Neural Network (GNN) has exhibited broad applications in scientific advancements, such as predictions of molecular structure–property relationships and collective cellular function. In particular, graph augmentation techniques have shown promising results in improving graph-based and node-based classification tasks. Still, they have rarely been explored in the existing GNN-based subgraph representation learning studies. In this study, we develop a novel multi-view augmentation mechanism to improve subgraph representation learning models and thus the accuracy of downstream prediction tasks. Our augmentation technique creates multiple variants of subgraphs and embeds these variants into the original graph to achieve highly improved training efficiency, scalability, and accuracy. Benchmark experiments on several real-world biological and physiological datasets demonstrate the superiority of our proposed multi-view augmentation techniques in subgraph representation learning.

1. Introduction

Subgraph representation learning using Graph Neural Networks (GNNs) has been widely applied to essential scientific areas, including chemistry, biology, pharmaceutics, and materials science. For example, a PPI (Protein–Protein Interaction) network (Zitnik et al., 2018) represents proteins as nodes and their interactions as edges. Typically, a set of proteins and their interactions collaboratively determine biological processes or cellular functions. This set of proteins and their interactions correspond to subgraphs in the PPI networks. As another example, in a large molecular aggregate, single atoms and chemical bonds can be viewed as nodes and edges, respectively, and any single molecule in the aggregate can be treated as a subgraph of the system.

Although several subgraph-based GNNs have been proposed to learn subgraph representations for subgraph-related tasks (Alsentzer et al., 2020; Kim & Oh, 2022; Wang & Zhang, 2021), none of them have implemented graph augmentation techniques to improve their accuracy. Specifically, graph augmentation approaches have shown great potential in promoting the expressive power of GNNs on graph-level and node-level tasks. To name a few, GraphCL (Graph Contrastive Learning) (You et al., 2020) drops nodes to build perturbed graphs for contrastive representation learning. MPNN (Message Passing Neural Networks) (Gilmer et al., 2017) promotes the graph message passing by inserting a supernode into the existing graph. DropEdge (Rong et al., 2020) drops random edges to avoid over-fitting and relieve over-smoothing (You et al., 2020; 2021; Zhu et al., 2020). NeuralSparse (Zheng et al., 2020) introduces a graph sparsification strategy to augment graph data. To the best of our knowledge, existing graph augmentation approaches are designed to improve representation learning for nodes or graphs instead of subgraphs.

This work addresses a fundamental question: *can graph augmentation strategies boost subgraph representation learning?* To answer the question, we apply general graph augmentation techniques such as dropping nodes to existing subgraph-based GNNs and observe surprisingly significant accuracy degradation. We see that the micro-F1 performance of GLASS (GNN with LAbeling trickS for Subgraph) (Wang & Zhang, 2021) on HPO-METAB (Human Phenotype Ontology-Metabolism) (Splinter et al., 2018; Hartley et al., 2020) drops by 1.2% when applying the DropEdge on the whole graph and by 12.5% when applying the DropEdge directly on subgraphs. This motivates us to develop subgraph-specific augmentation techniques to improve the subgraph representation learning model.

To this end, we present a novel multi-view approach to augment subgraph representation learning. We hypothesis that the accuracy drop from DropEdge is because DropEdge

[1]Department of Chemistry, University of Massachusetts Amherst, Amherst, MA 01003, United States [2]School of Software Engineering, Tongji University, Shanghai 200070, China [3]Department of Mechanical and Automation Enginering, Shanghai University, Shanghai 200444, China [4]Pritzker School of Molecular Engineering, The University of Chicago, Chicago, IL 60637, United States [5]College of Information and Computer Sciences, University of Massachusetts Amherst, Amherst, MA 01003, United States. Correspondence to: Zhou Lin <zhoulin@umass.edu>, Hui Guan <huiguan@umass.edu>.

2nd AI4Science Workshop at the 39^{th} International Conference on Machine Learning (ICML), 2022.

changes the original connections in subgraphs and the representation learned on augmented subgraphs only is less useful for downstream tasks. Our basic idea is to not only create perturbations of the original subgraphs (i.e., augmented subgraphs), but also preserve the original view of the subgraphs (i.e., original subgraphs) during training. Generating separate graphs to include different views of the same subgraph is not scalable because the size of the augmented graph will grow linearly with the number of subgraphs and the number of views. We address the scalability issue by embedding augmented subgraphs in the original graph, significantly decreasing the demand for GPU resources. We obtain the embeddings of both the original and augmented subgraphs by feeding them to subgraph-based GNNs and generate subgraph embeddings by applying a pooling function.

In summary, this work makes the following contributions:

- This work proposes a novel multi-view augmentation strategy to improve the accuracy of subgraph-based learning tasks. This study is the first to explore the benefits of graph augmentation techniques in subgraph representation learning.
- The proposed multi-view augmentation strategy dynamically binds augmented subgraph views to the whole graph to drop exaggerated GPU resource consumption in order to achieve highly-improved training efficiency and task accuracy.
- Empirical evaluations on three subgraph datasets demonstrate that our augmentation approach can improve existing subgraph representation learning by 0.1%–1.6% in accuracy, which is 4.6% in average higher than general graph augmentation techniques DropEdge and DropNode.

2. Related Works

Subgraph Representation Learning Subgraph representation learning using GNNs has gained substantial attention these years (Meng et al., 2018) due to its broad applications in scientific domains. Outstanding examples include SubGNN (SubGraph Neural Network) (Alsentzer et al., 2020), which routes messages for internal and border properties within sub-channels of each channel, including neighborhood, structure, and position. After that, the anchor patch is sampled and the features of the anchor patch are aggregated to the connected components of the subgraph through six sub-channels. GLASS (Wang & Zhang, 2021) employs a labeling trick (Zhang et al., 2021) and labels nodes belonging to any subgraph to boost plain GNNs on subgraph tasks. S2N (Subgraph-To-Node) (Kim & Oh, 2022) translates subgraphs into nodes and thus reduces the scale of the input graph. These approaches focus on developing novel subgraph-based GNNs to improve task accuracy, but they have never implemented graph augmentation techniques.

Graph Augmentation Many general graph augmentation techniques have been proposed to improve task accuracy recently. DGI (Deep Graph Infomax) (Veličković et al., 2019) perturbs the nodes by performing a row-wise swap of the input feature matrix while the adjacency matrix remains unchanged, generating negative samples for comparison learning and maximizing the mutual information of input and output. GAug-M (Zhao et al., 2021) generates and removes edges of the graph by training an edge predictor to finally achieve the effect of high connectivity between nodes within the same class and low connectivity between nodes from different classes. NeuralSparse (Neural Sparsification) (Zheng et al., 2020) proposes a supervised graph sparsification technique that improves generalization by learning to remove potentially task-irrelevant edges from the input graph. GraphCL (You et al., 2020) points out that different data augmentation techniques introduce different advantages in graph learning tasks in different domains. For example, edge perturbation can enhance learning in social network graphs, but can be counterproductive in compound graphs learning by destroying the original information. SUBG-CON (SUBGraph CONtrast) (Jiao et al., 2020) samples a series of subgraphs containing regional neighbors from the original graph as training data to serve as an augmented node representation. Although these methods show promising results for augmenting graphs for node- and graph-based downstream tasks, they are not designed for augmenting subgraphs for subgraph-based tasks.

Multi-view Graph Learning Multi-view representation learning on graphs has attracted significant attention because they capture different properties on the same graph. Hassani et al. (Hassani & Khasahmadi, 2020) introduce a multi-view graph learning manner to perform contrastive learning. O2MAC (One2Multi graph AutoenCoder) (Fan et al., 2020) proposes a multi-view-based auto-encoder to promote self-supervised learning. MV-GNN (Multi-View Graph Neural Network) (Ma et al., 2020) utilizes two MPNNs to encode atom and bond information respectively via multi-view graph construction. They construct multi-view graphs to express different levels of information in a graph, which is an intuitive and efficient way of building augmented graphs. Our work also leverages multi-view–based augmentation but focuses on subgraph-based tasks.

3. Method

3.1. Preliminaries

Let $G = (\mathcal{V}, \mathcal{E}, \mathbf{X})$ denote a graph, where $\mathcal{V} = \{1, 2, .., N\}$ represents the node set, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represents the edge

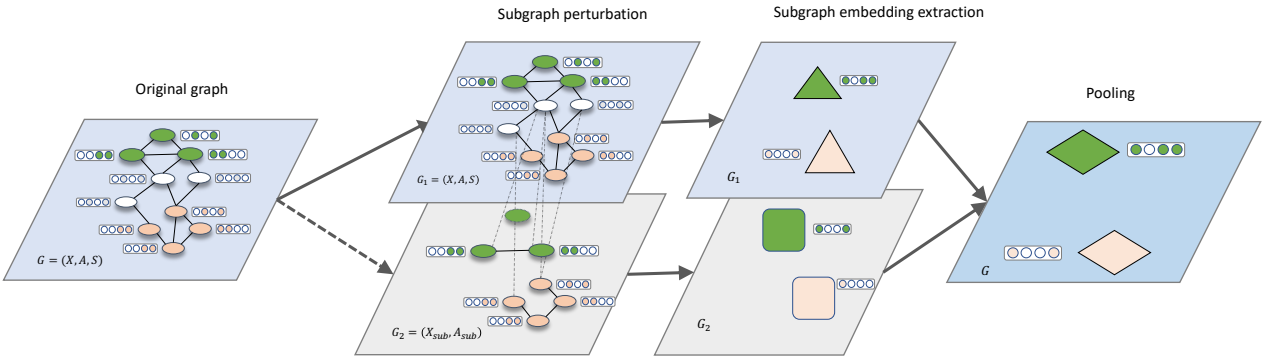

Figure 1. Overview of our proposed subgraph augmentation approach. The two subgraphs in the original graph are colored in gree and orange. We first generate multi-subgraph views via stochastic augmentation. Following that we connect the augmented subgraph to the remaining part of the original graph, by adding edges that link the augmented subgraph and the whole graph. After feeding forward the whole graph into subgraph-specific GNNs, we extract the subgraph embeddings of different views, respectively (triangles and squares). Ultimately, we fuse the embeddings of different views by a pooling function and obtain the augmented subgraph embeddings (diamonds).

sets, and $\mathbf{X}$ represents the corresponding node feature. X_i, the i^{th} row of $\mathbf{X}$, represents the features associated with the i^{th} node and the adjacency matrix $\mathbf{A} \in \{0,1\}^{N\times N}$, where $a_{ij} = 1$ denotes that $(v_i, v_j) \in \mathcal{E}$.

$G_S = (\mathcal{V}_S, \mathcal{E}_S, \mathbf{X}_S)$ denotes a subgraph of G, where $\mathcal{V}_S \subseteq \mathcal{V}$, $\mathcal{E}_S \subseteq \mathcal{E} \cap (\mathcal{V}_S \times \mathcal{V})_S$, and $\mathbf{X_S}$ stacks the rows of $\mathbf{X}$ belonging to $\mathcal{V}_S$. The adjacency matrix of a subgraph G_S is $\mathbf{A}_S$.

Given the subgraphs $\mathcal{S} = (G_{S_1}, G_{S_2}, .., G_{S_n})$ and their labels $\mathbf{T} = \{t_{S_1}, t_{S_2}, ..., t_{S_n}\}$, the goal of *subgraph property prediction* is to learn a representation embedding h_{S_i} for each subgraph G_{S_i} to predict the corresponding t_{S_i}.

3.2. Proposed Multi-View Augmentation

This section presents our proposed multi-view augmentation approach (Figure 1). At each forward step, we first generate augmented views of subgraphs by perturbing original subgraphs randomly with dropping nodes. After that, we add the augmented subgraphs into the original graph and feed the new graph into a subgraph-specific neural network. Here, we obtain subgraph embeddings of both the original subgraph and the augmented subgraph. These embeddings are fed into a pooling function to generate a single subgraph embedding for each subgraph, which is used for downstream subgraph-based tasks.

3.2.1. DropNode

DropNode is a graph data pertubation strategy that randomly drops nodes in a graph (You et al., 2020). We apply DropNode for each subgraph to generate an augmented subgraph. For each subgraph, we generate a stochastic boolean mask $\mathbf{M}_p \in \mathbb{R}^{m\times m}$, where m is the number of nodes in the subgraph and p represents the rate of dropping nodes. We use the boolean mask to set the corresponding value of the adjacency matrix of the subgraph to zero if $\mathbf{M}_p \neq 1$. The new adjacency matrix becomes

$$\mathbf{A}'_S = \mathbf{A}_S - \mathbf{M}_p \odot \mathbf{A}_S \tag{1}$$

where $\odot$ means element-wise product.

3.2.2. Augmenting the Original Graph

After subgraph augmentation, we obtain an augmented subgraph $G'_S = (\mathcal{V}'_S, \mathcal{E}'_S, \mathbf{X}'_S)$. We enrich the original graph to include both the augmented subgraph and the original subgraph. The enriched graph is thus called a *Multi-View Graph*.

Mathematically, the multi-view graph $G' = (\mathcal{V}', \mathcal{E}', \mathbf{X}')$ where $\mathcal{V}' = \mathcal{V} \cup \mathcal{V}'_S$. The consequent adjacency matrix becomes

$$\mathbf{A}' = \begin{bmatrix} \mathbf{A} & \mathbf{A}[:, \mathcal{V}'_S] \\ \mathbf{A}[\mathcal{V}'_S, :] & \mathbf{A}_S \end{bmatrix}. \tag{2}$$

Feeding forward the multi-view graph into subgraph-specific neural networks, we can get the embeddings of both the augmented subgraph and the original subgraph. We fuse different subgraph embeddings into one embedding by applying a pooling function (e.g., MaxPool or AvgPool):

$$h_S = \text{Pool}(\text{GNN}(G')). \tag{3}$$

With the learned subgraph embeddings, we can predict the subgraph properties by applying a Multi-Layer Perception (MLP):

$$t_S = \text{MLP}\,(h_S). \tag{4}$$

Table 1. Statistics of three real-world datasets.

DATASET	# NODES	# EDGES	# SUBGRAPHS
PPI-BP	17,080	316,951	1,591
HPO-METAB	14,587	3,238,174	2,400
EM-USER	57,333	4,573,417	324

4. Experiments

In this section, we evaluate the efficacy of our proposed augmentation technique by comparing it with several baselines from real-world datasets.

4.1. Experiment Settings

Datasets Table 1 summarizes the subgraph structure related statistics of the datasets obtained from SubGNN (Alsentzer et al., 2020). Specifically, PPI-BP (Zitnik et al., 2018) aims to predict the collective cellular function of a given set of genes known to be associated with specific biological processes in common. The graph shows the correlation of the human PPI (protein-protein interaction) network where nodes represent proteins and edges represent the interaction between proteins. A subgraph is defined by the collaboration of proteins and labeled according to cellular functions from six categories (metabolism, development, signal transduction, stress/death, cell organization, and transport).

HPO-METAB (Splinter et al., 2018; Hartley et al., 2020) simulates rare disease diagnosis with the task of predicting subcategories of metabolic disorders that are the most consistent with these phenotypes. The graph is a knowledge graph containing phenotypic and genotypic information for rare diseases. A subgraph consists of a collection of phenotypes associated with rare monogenic diseases. HPO-METAB subgraphs are labeled according to the type of metabolic disease.

The task of the EM-USER dataset (Ni et al., 2019) is to predict the characteristics of a user in social fitness networks. In the network, nodes represent workouts and edges exist between workouts completed by multiple users. Each subgraph is represented by a sub-network that constitutes a user's workout history and is labeled with features about the user.

Models The proposed augmentation technique is compatible with many subgraph-based GNNs and can be integrated on top of them because it does not alter the GNN model. In the evaluation, we select GLASS (Wang & Zhang, 2021), the current state-of-the-art subgraph representation learning model to integrate our subgraph augmentation approach. We refer to our approach as **GLASS w/ Multi-View**.

For baselines, we compare our model with four different approaches that apply DropNode and DropEdge to GLASS, **GLASS w/ DropNode**, **GLASS w/ DropNode sub**, **GLASS w/ DropEdge**, **GLASS w/ DropEdge sub**, several other subgraph-based GNN models (**Sub2Vec** (Adhikari et al., 2018) and **SubGNN** (Alsentzer et al., 2020)), and general GNN baselines (**MLP** (Multi-Layer Perceptron) and **GBDT** (Gradient Boosting Decision Tree) (Chen & Guestrin, 2016)). GLASS w/ DropNode and GLASS w/ DropEdge apply DropNode (You et al., 2020) and DropEdge (Rong et al., 2020), respectively, to the entire graph while GLASS w/ DropNode sub and GLASS w/ DropEdge sub apply DropNode and DropEdge to subgraphs only, respectively.

Implementation Details The hyperparameters in GLASS w/ Multi-View include the probability of dropping nodes and the number of augmented views. We use grid search to optimize the hyperparameters. The dropping node probability is 0.15 for PPI-BP and HPO-METAB and 0.3 for EM-USER. We find one augmented view is sufficient and adding more augmented views do not increase task accuracy, which echos the findings by Hassani et al. (Hassani & Khasahmadi, 2020).

We first train the model in an unsupervised manner as is performed in GLASS (Wang & Zhang, 2021), and then use supervision from downstream tasks to fine-tune the model parameters. We perform 10 different training and validation processes with 10 distinct random seeds.

4.2. Results

The empirical performance is summarized in Table 2. Our proposed subgraph augmentation improves task accuracy across all three datasets. It consistently performs better than all the baseline approaches, mainly because it inhibits over-smoothing and over-fitting. Specifically, our approach improves the Micro-F1 scores by 0.1%–1.6%, 2.4%–8.8%, and 12.5%–23.5% compared to plain GLASS, SubGNN, and Sub2Vec, respectively, which are state-of-the-art approaches for subgraph representation learning. The results also demonstrate the superiority of subgraph-based GNNs over general graph representation learning approaches.

We also observe a significant accuracy drop after applying DropEdge to GLASS. The accuracy drop of GLASS w/ DropEdge sub is more significant than GLASS w/ DropEdge, which applies DropEdge to the entire graph instead of subgraphs. It implies that the accuracy drop from DropEdge comes from the fact that DropEdge removes some of the internal connections of a subgraph. The subgraph representation using only the augmented subgraph is thus less useful than integrating the original subgraph. This result demonstrates the necessity to keep at least one view of the original subgraph structure.

Table 2. Mean Micro-F1 scores with standard deviations of the mean on three real-world datasets. Results are provided from runs with 10 random seeds.

METHOD	PPI-BP	HPO-METAB	EM-USER
GBDT	0.446 ± 0.000	0.404 ± 0.000	0.694 ± 0.000
MLP	0.445 ± 0.003	0.386 ± 0.011	0.524 ± 0.019
SUB2VEC	0.388 ± 0.001	0.472 ± 0.010	0.779 ± 0.003
SUBGNN	0.599 ± 0.008	0.537 ± 0.008	0.816 ± 0.013
GLASS W/ DROPEDGE	0.593 ± 0.007	0.602 ± 0.090	0.814 ± 0.032
GLASS W/ DROPEDGE SUB	0.575 ± 0.007	0.490 ± 0.020	0.808 ± 0.026
GLASS W/ DROPNODE	0.601 ± 0.009	0.612 ± 0.011	0.796 ± 0.055
GLASS W/ DROPNODE SUB	0.577 ± 0.016	0.453 ± 0.025	0.816 ± 0.228
GLASS	0.613 ± 0.007	0.614 ± 0.005	0.888 ± 0.006
GLASS w/ Multi-View (this work)	$\mathbf{0.623 \pm 0.008}$	$\mathbf{0.615 \pm 0.010}$	$\mathbf{0.904 \pm 0.006}$

5. Conclusion

This paper presents a novel multi-view augmentation technique to improve subgraph representation learning. The augmentation randomly perturbs the subgraph components in an input graph and generates multi-views of this subgraph. Existing subgraph-based GNNs can be used to produce embeddings for each view of the subgraph. The embeddings are fused by a pooling function to create a final representation of the subgraph. We evaluate the augmentation technique on top of GLASS, a state-of-the-art subgraph-based GNN, and the experimental results on three subgraph datasets demonstrate that the augmentation can effectively improve the accuracy of the GNN.

Acknowledgment

This work is financially supported by UMass ADVANCE Collaborative Research Seed Fund and UMass Start-up Funds offered to Z.L. and H.G. We would like to thank helpful discussions from Dr. Lixue Cheng, Dr. Chenru Duan, Dr. Zhenkun Huang, Dr. Kun Yao, and Mr. Shuo Zhang.

References

Adhikari, B., Zhang, Y., Ramakrishnan, N., and Prakash, B. A. Sub2Vec: Feature learning for subgraphs. In Phung, D., Tseng, V. S., Webb, G. I., Ho, B., Ganji, M., and Rashidi, L. (eds.), *Advances in Knowledge Discovery and Data Mining*, pp. 170–182, 2018.

Alsentzer, E., Finlayson, S. G., Li, M. M., and Zitnik, M. Subgraph neural networks. In *NeurIPS*, 2020.

Chen, T. and Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pp. 785–794, 2016.

Fan, S., Wang, X., Shi, C., Lu, E., Lin, K., and Wang, B. One2Multi graph autoencoder for multi-view graph clustering. In *Proceedings of The Web Conference 2020 (WWW '20)*, 2020.

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *ICML'17*, pp. 1263–1272, 2017.

Hartley, T., Lemire, G., Kernohan, K. D., Howley, H. E., Adams, D. R., and Boycott, K. M. New diagnostic approaches for undiagnosed rare genetic diseases. *Annual Review of Genomics and Human Genetics*, 21(1):351–372, 2020.

Hassani, K. and Khasahmadi, A. H. Contrastive multi-view representation learning on graphs. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20, pp. 4116–4126, 2020.

Jiao, Y., Xiong, Y., Zhang, J., Zhang, Y., Zhang, T., and Zhu, Y. Sub-graph contrast for scalable self-supervised graph representation learning. In *2020 IEEE International Conference on Data Mining (ICDM)*, pp. 222–231, 2020.

Kim, D. and Oh, A. Efficient representation learning of subgraphs by subgraph-to-node translation. In *ICLR 2022 Workshop on Geometrical and Topological Representation Learning*, 2022.

Ma, H., Bian, Y., Rong, Y., Huang, W., Xu, T., Xie, W., Ye, G., and Huang, J. Dual message passing neural network for molecular property prediction. *CoRR*, abs/2005.13607, 2020.

Meng, C., Mouli, S. C., Ribeiro, B., and Neville, J. Subgraph pattern neural networks for high-order graph evolution prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Ni, J., Muhlstein, L., and McAuley, J. Modeling heart rate and activity data for personalized fitness recommendation. In *The World Wide Web Conference*, WWW '19, pp. 1343–1353, 2019.

Rong, Y., Huang, W., Xu, T., and Huang, J. Dropedge: Towards deep graph convolutional networks on node classification. In *International Conference on Learning Representations*, 2020.

Splinter, K., Adams, D. R., Bacino, C. A., Bellen, H. J., Bernstein, J. A., Cheatle-Jarvela, A. M., Eng, C. M., Esteves, C., Gahl, W. A., Hamid, R., Jacob, H. J., Kikani, B., Koeller, D. M., Kohane, I. S., Lee, B. H., Loscalzo, J., Luo, X., McCray, A. T., Metz, T. O., Mulvihill, J. J., Nelson, S. F., Palmer, C. G., Phillips, J. A., Pick, L., Postlethwait, J. H., Reuter, C., Shashi, V., Sweetser, D. A., Tifft, C. J., Walley, N. M., Wangler, M. F., Westerfield, M., Wheeler, M. T., Wise, A. L., Worthey, E. A., Yamamoto, S., and Ashley, E. A. Effect of genetic diagnosis on patients with previously undiagnosed disease. *New England Journal of Medicine*, 379(22):2131–2139, 2018.

Veličković, P., Fedus, W., Hamilton, W. L., Liò, P., Bengio, Y., and Hjelm, R. D. Deep graph infomax. In *International Conference on Learning Representations*, 2019.

Wang, X. and Zhang, M. GLASS: GNN with labeling tricks for subgraph representation learning. In *International Conference on Learning Representations*, 2021.

You, Y., Chen, T., Sui, Y., Chen, T., Wang, Z., and Shen, Y. Graph contrastive learning with augmentations. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H. (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 5812–5823, 2020.

You, Y., Chen, T., Shen, Y., and Wang, Z. Graph contrastive learning automated. In *International Conference on Machine Learning*, pp. 12121–12132, 2021.

Zhang, M., Li, P., Xia, Y., Wang, K., and Jin, L. Labeling trick: A theory of using graph neural networks for multi-node representation learning. In Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, 2021.

Zhao, T., Liu, Y., Neves, L., Woodford, O., Jiang, M., and Shah, N. Data augmentation for graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 11015–11023, 2021.

Zheng, C., Zong, B., Cheng, W., Song, D., Ni, J., Yu, W., Chen, H., and Wang, W. Robust graph representation learning via neural sparsification. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pp. 11458–11468, 2020.

Zhu, Y., Xu, Y., Yu, F., Liu, Q., Wu, S., and Wang, L. Deep graph contrastive representation learning. *arXiv preprint arXiv:2006.04131*, 2020.

Zitnik, M., Sosič, R., Maheshwari, S., and Leskovec, J. BioSNAP Datasets: Stanford biomedical network dataset collection. http://snap.stanford.edu/biodata, 2018.