G-Mixup: Graph Data Augmentation for Graph Classification

Xiaotian Han, Zhimeng Jiang1, Ninghao Liu, Xia Hu

International Conference on Machine Learning 2022

Presenter: Tanya Djavaherpour
CAS 747 –Winter 2024

Overview

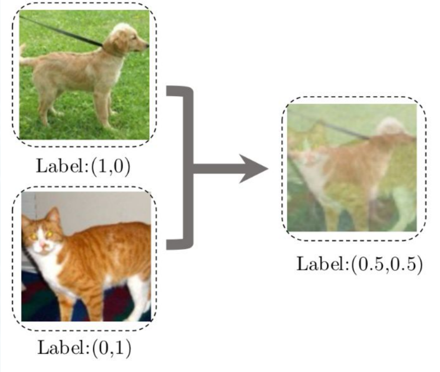
- Background and Motivation
- Methodology
- Experiments
- Conclusion

Background and Motivation

Mixup

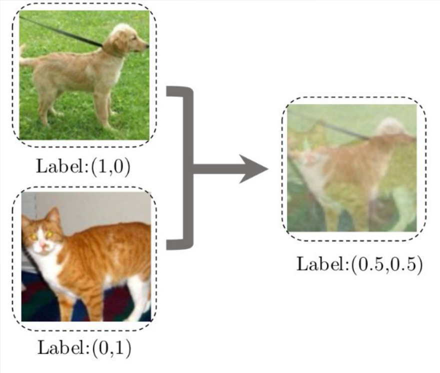
Mixup is a cross-instance data augmentation method, which linearly interpolates random sample pair to generate more synthetic training data.

Mixup have been empirically and theoretically shown to improve the generalization and robustness of deep neural networks.



$$\mathbf{x}_{new} = \lambda \mathbf{x}_i + (1 - \lambda)\mathbf{x}_j,$$
$$\mathbf{y}_{new} = \lambda \mathbf{y}_i + (1 - \lambda)\mathbf{y}_j,$$

Challenges for Graph Mixup



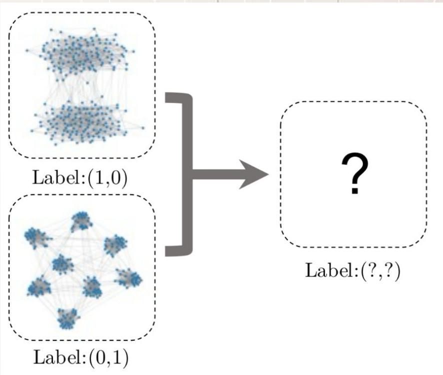


Image data is regular (image can be represented as matrix)

Image data is well-aligned (pixel to pixel correspondence)

Graph data is irregular (the number of nodes)

Graph data is not well-aligned (nodes not naturally ordered)

What are Graphons?

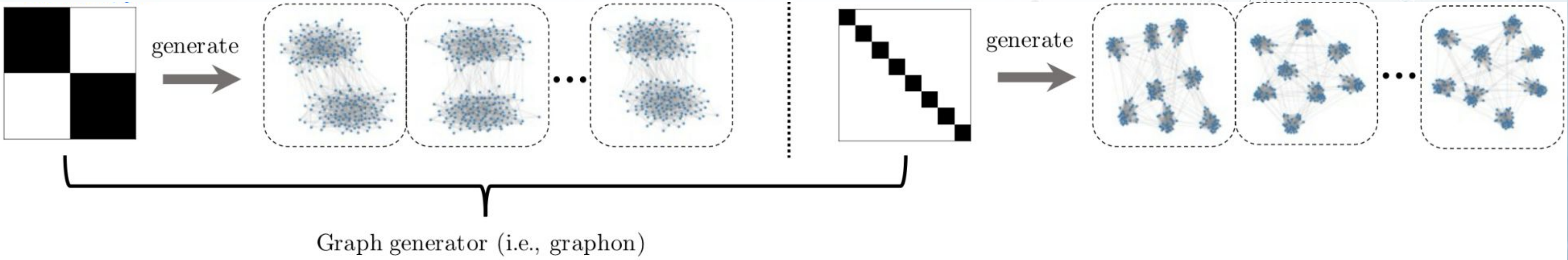
In this study Graphons are used to achieve the input graph mixup.

Graphon serves as a tool in graph theory for approximating large-scale network structures. It works based on the probability of an edge existing between each two nodes.

The graphons of different graphs are regular and well-aligned.

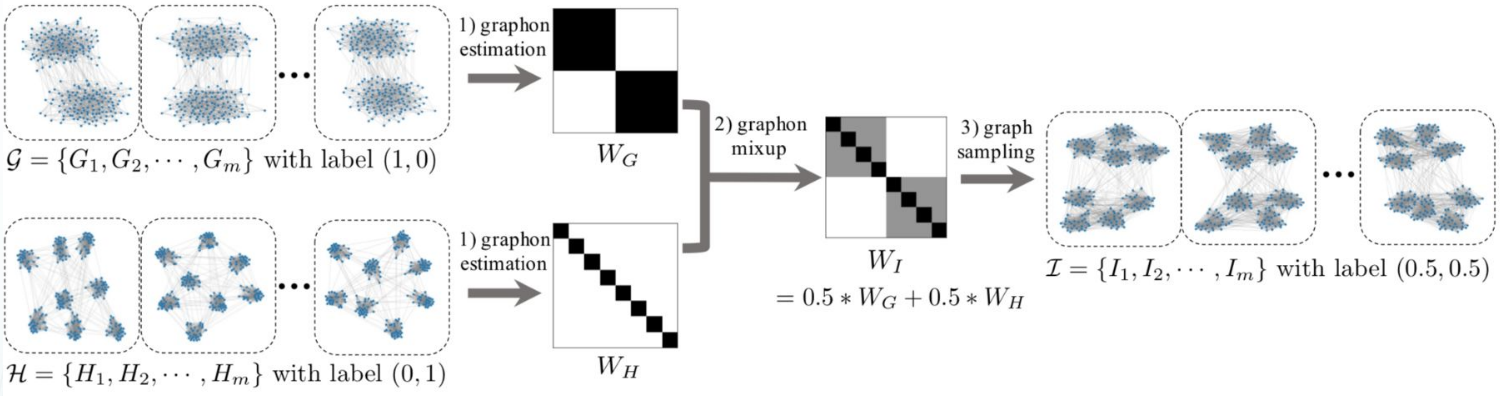
Graph Generator: Graphon

The real-world graphs can be regarded as generated from generator (i.e., graphon). For example,



Methodology

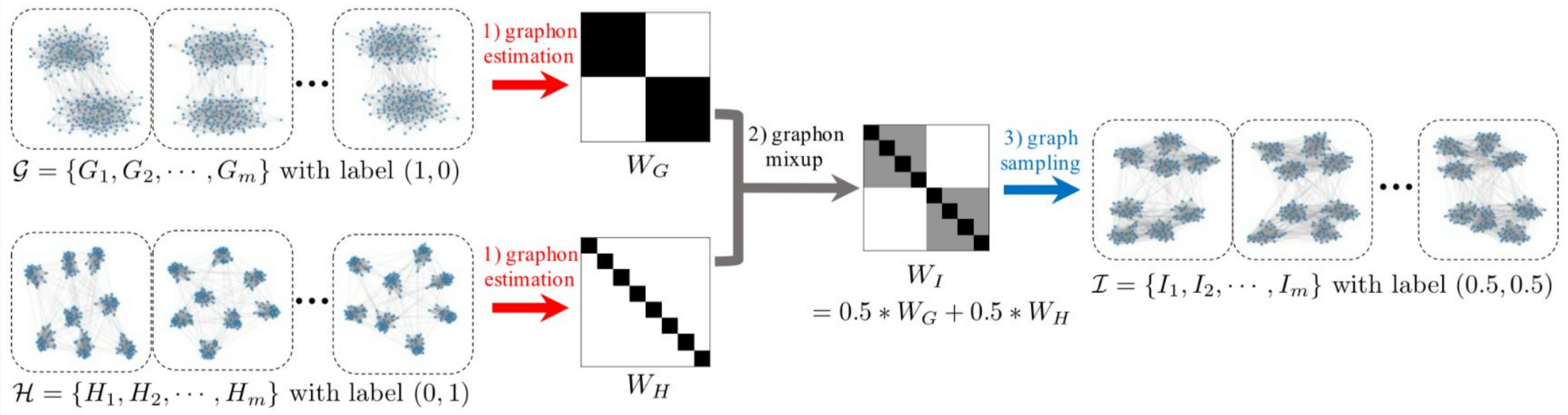
G-Mixup



The formal mathematical expression are as follows:

$$\begin{aligned}
&\text{Graphon Estimation:} && \mathcal{G} \to W_{\mathcal{G}}, \mathcal{H} \to W_{\mathcal{H}} \\
&\text{Graphon Mixup:} && W_{\mathcal{I}} = \lambda W_{\mathcal{G}} + (1-\lambda) W_{\mathcal{H}} \\
&\text{Graph Generation:} && \{I_1, I_2, \cdots, I_m\} \overset{\text{i.i.d}}{\sim} \mathbb{G}(K, W_{\mathcal{I}}) \\
&\text{Label Mixup:} && \mathbf{y}_{\mathcal{I}} = \lambda \mathbf{y}_{\mathcal{G}} + (1-\lambda) \mathbf{y}_{\mathcal{H}}
\end{aligned}$$

Implementation

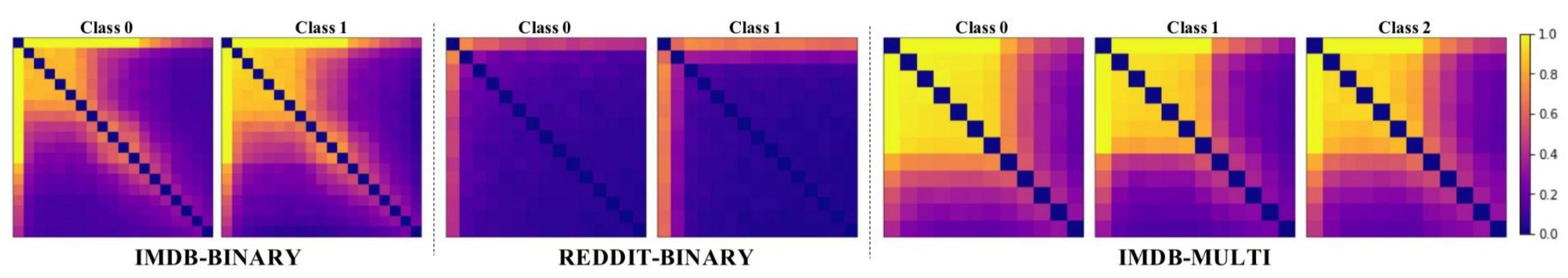


1. Graphon Estimation Using Graphons
2. Synthetic Graphs Generation Using Bernoulli Distributions

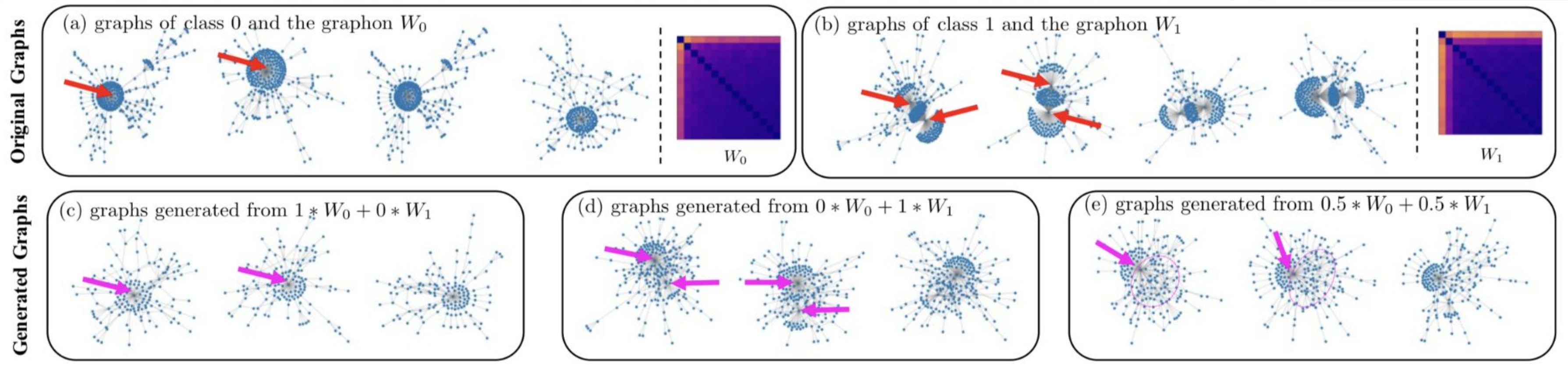
Experiments

Do different classes of graphs have different graphons?

Visualization the estimated graphons on IMDB-BINARY, REDDIT-BINARY, and IMDB-MULTI:



What is G-Mixup doing? A case study



The class 0 has one high-degree node while class 1 have two (a)(b).

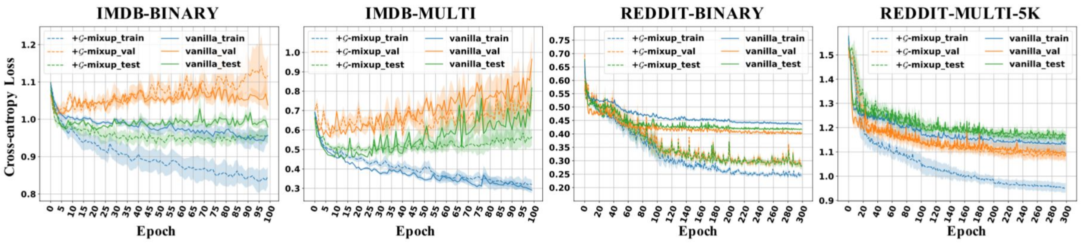
The generated graphs based on

- $(1 * W_0 + 0 * W_1)$ have one high-degree node (c).
- $(0 * W_0 + 1 * W_1)$ have two high-degree nodes (d).
- $(0.5 * W_0 + 0.5 * W_1)$ have a high-degree node and a dense subgraph (e).

Graphs generated by $\mathcal{G}$-Mixup are the mixture of original graphs.

Can G-Mixup improve the performance of GNNs?

	Dataset	IMDB-B	IMDB-M	REDD-B	REDD-M5	REDD-M12
	#graphs	1000	1500	2000	4999	11929
	#classes	2	3	2	5	11
	#avg.nodes	19.77	13.00	429.63	508.52	391.41
	#avg.edges	96.53	65.94	497.75	594.87	456.89
GCN	vanilla	72.18	48.79	78.82	45.07	46.90
	w/ Dropedge	72.50	49.08	81.25	51.35	47.08
	w/ DropNode	72.00	48.58	79.25	49.35	47.93
	w/ Subgraph	68.50	49.58	74.33	48.70	47.49
	w/ M-Mixup	72.83	49.50	75.75	49.82	46.92
	w/ $\mathcal{G}$-Mixup	**72.87**	**51.30**	**89.81**	**51.51**	**48.06**
GIN	vanilla	71.55	48.83	92.59	55.19	50.23
	w/ Dropedge	**72.20**	48.83	92.00	55.10	49.77
	w/ DropNode	72.16	48.33	90.25	53.26	49.95
	w/ Subgraph	68.50	47.25	90.33	54.60	49.67
	w/ M-Mixup	70.83	49.88	90.75	54.95	49.81
	w/ $\mathcal{G}$-Mixup	71.94	**50.46**	**92.90**	**55.49**	**50.50**



The loss curves of G-Mixup are lower than the vanilla model.

G-Mixup can improve the generalization of graph neural networks.

Conclusion

Conclusion

G-Mixup Introduction: A new method for graph data augmentation.

Graphon Utilization: Mixing graphons from different classes to generate new graphs.

Improved GNN Performance: Enhanced performance, generalization, and noise robustness.

Innovation in Graph Data Handling: A novel approach to address graph data augmentation challenges.

References

Cordaro, D., Cox, S., Ren, Y., & Yu, T. (2023). On the Reproducibility of “G-Mixup:Graph Data Augmentation for Graph Classification”. ML Reproducibility Challenge 2022.

Lovász, L. (2012). Large networks and graph limits (Vol. 60). American Mathematical Soc.

Xu, H., Luo, D., Carin, L., & Zha, H. (2021). Learning graphons via structured gromov-wasserstein barycenters. Proceedings of the AAAI Conference on Artificial Intelligence, 10505–10513.

Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (2017). Mixup: Beyond empirical risk minimization. International Conference on Learning Representations.

Zhang, L., Deng, Z., Kawaguchi, K., Ghorbani, A., & Zou, J. (2021). How does mixup help with robustness and generalization? International Conference on Learning Representations.

Thanks for your attention!