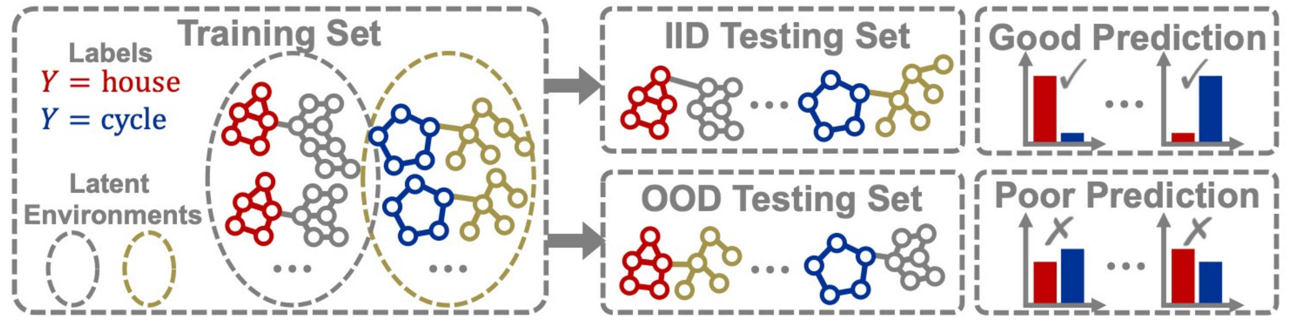
Learning Invariant Graph Representations for Out-of-Distribution Generalization

Haoyang Li, Ziwei Zhang, Xin Wang, Wenwu Zhu
Tsinghua University

Motivation

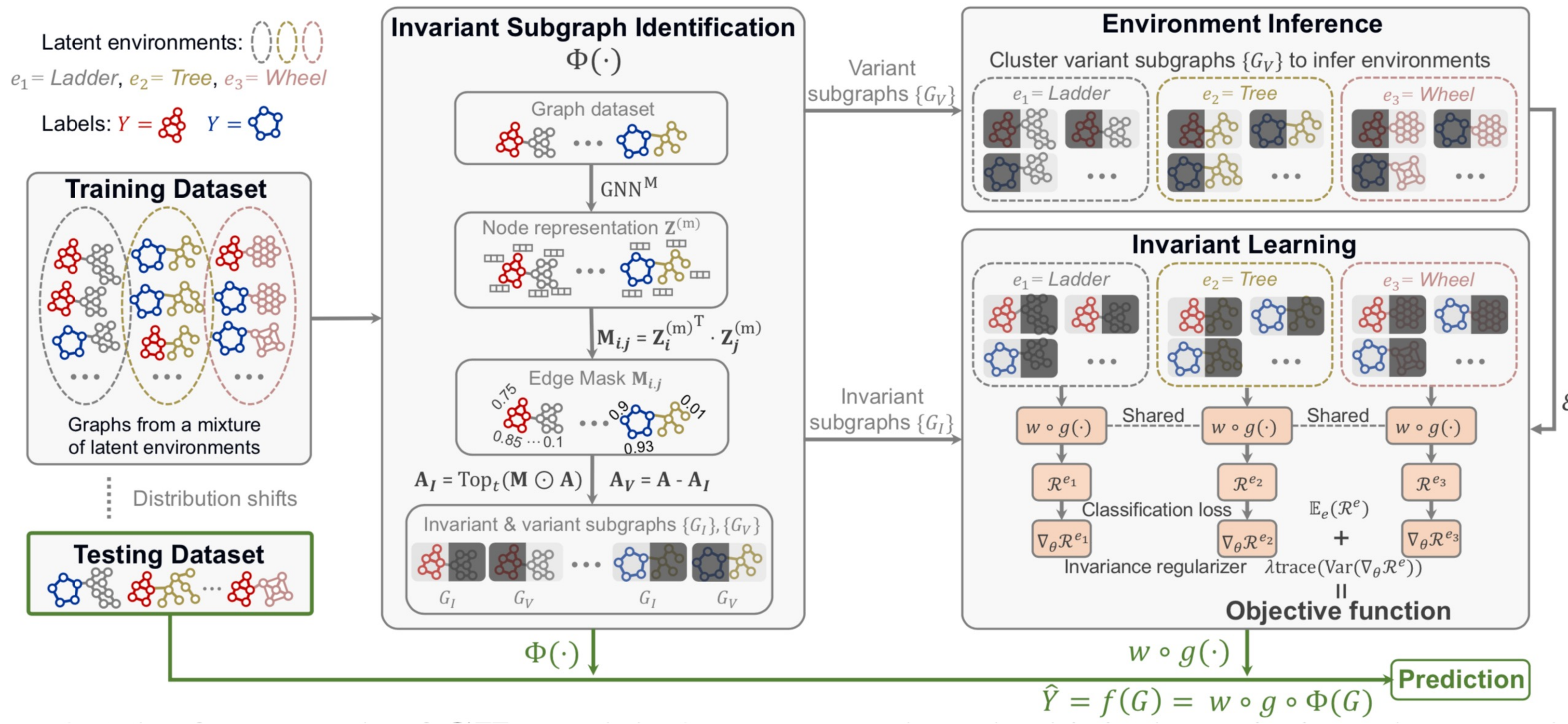
- Graph structured data is ubiquitous in the real world.
- GNNs have shown profound successes in graph representation learning.
- Most GNNs perform well when testing and training data come from ***identical distribution***.
- However, in real world, the testing distribution may suffer ***unobserved or uncontrolled shifts (out-of-distribution)*** compared with the training distribution.



- Most GNNs do not consider the out-of-distribution generalization ability, so that their performances can ***drop substantially*** on out-of-distribution testing graphs.

Method

- We propose a **G**raph **I**nvariant **L**earning model (**GIL**) for ***Graph OOD generalization***.
- 1) **Invariant Subgraph Identification**: a GNN-based subgraph generator identifies potentially invariant subgraphs for the input graphs.
- 2) **Environment Inference**: infer environment labels by clustering the environment-discriminative features of variant subgraphs.
- 3) **Invariant Learning**: optimize the maximal invariant subgraph generator criterion given the identified invariant subgraphs and inferred environments to generate representations.



Theoretical Analysis

- **OOD generalization on graphs → finding a maximal invariant subgraph generator**

Theorem 4.1. *Let Φ^* be the optimal invariant subgraph generator in Assumption 3.1 and denote the complement as $G \backslash \Phi^*(G)$, i.e., the corresponding variant subgraph. Then, we can obtain the optimal predictor under distribution shifts, i.e., the solution to Problem 1, as follows:*

$$\arg\min_{w,g} w \circ g \circ \Phi^*(G) = \arg\min_{f} \sup_{e \in \mathrm{supp}(\mathcal{E})} \mathcal{R}(f|e), \qquad (10)$$

- **Our method satisfies permutation invariance**

Theorem 4.2. *Our proposed **GIL** model is permutation-invariant if $\mathrm{GNN}^{\mathrm{M}}$ and $\mathrm{GNN}^{\mathrm{I}}$ are permutation-equivariant and $\mathrm{READOUT}^{\mathrm{I}}$ is permutation-invariant.*

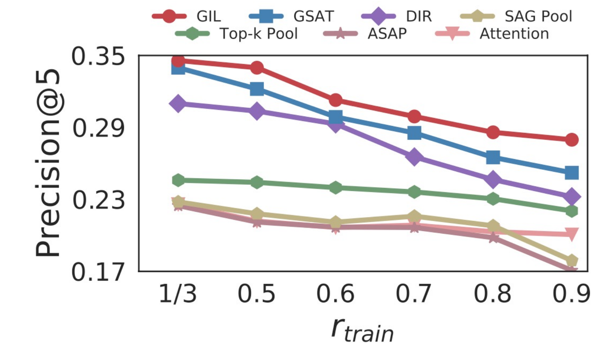
Experiment

- **Best OOD classification performance**

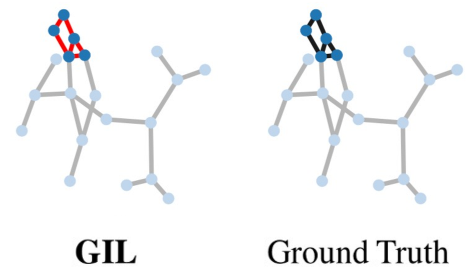
	Scenario 1: $r_{test} = 1/3$						Scenario 2: $r_{test} = 0.2$					
r_{train}	$r = 1/3$	$r = 0.5$	$r = 0.6$	$r = 0.7$	$r = 0.8$	$r = 0.9$	$r = 1/3$	$r = 0.5$	$r = 0.6$	$r = 0.7$	$r = 0.8$	$r = 0.9$
ERM	53.60±3.79	51.24±4.13	47.04±7.01	38.80±3.72	37.84±3.01	37.44±2.15	48.48±4.53	41.72±4.81	36.92±6.93	35.72±8,33	28.80±3.91	19.60±1.66
Attention	54.31±3.98	53.24±3.56	42.52±6.20	35.20±1.05	34.48±1.18	33.88±1.01	44.04±4.33	31.64±0.67	25.72±5.34	24.80±4.06	23.20±3.60	18.04±2.88
Top-k Pool	54.68±2.71	53.12±5.58	44.56±4.57	37.44±2.04	35.24±2.28	34.28±4.11	45.68±5.16	34.20±4.34	31.00±2.89	30.64±3.59	29.16±2.18	27.56±3.91
SAG Pool	54.08±3.66	52.60±3.52	44.68±5.25	37.68±4.03	34.28±1.82	32.72±1.83	44.36±6.09	38.64±3.02	31.36±4.40	32.84±1.86	28.72±3.11	26.60±5.37
ASAP	54.00±4.21	51.92±3.81	45.12±1.98	36.28±0.86	34.24±2.02	34.40±3.15	49.88±4.90	34.52±4.35	27.00±2.61	27.20±2.53	27.96±3.89	22.88±4.33
GroupDRO	53.20±4.91	51.40±4.35	48.32±5.35	39.12±4.27	38.40±2.76	37.64±1.69	52.68±4.04	43.68±4.05	31.92±6.84	34.36±8.41	28.88±5.14	20.32±1.64
IRM	52.00±2.34	50.60±3.54	47.84±6.95	38.80±3.72	39.84±3.21	39.00±3.98	50.24±6.73	41.60±4.75	35.24±5.35	34.92±8.03	29.44±5.47	21.84±3.57
V-REx	53.16±3.25	46.04±6.11	45.36±3.66	40.24±3.86	39.48±3.00	39.12±3.48	50.56±2.83	37.16±6.24	34.52±3.00	29.72±4.58	27.32±3.18	24.04±6.08
DIR	52.96±5.06	52.08±1.93	50.12±2.76	49.84±2.46	45.20±1.11	41.24±4.73	50.68±5.20	49.96±1.75	45.44±6.00	40.56±2.36	39.92±4.53	32.52±4.59
GSAT	53.67±3.65	53.34±4.08	51.54±3.78	50.12±3.29	45.83±4.01	44.22±5.57	51.36±4.21	50.48±3.98	46.93±5.03	43.55±3.67	40.35±4.21	33.87±5.19
GIL	**55.44**±3.11	**54.56**±3.02	**53.60**±4.82	**53.12**±2.18	**51.24**±3.88	**46.04**±3.51	**54.80**±3.93	**52.48**±4.41	**50.08**±5.47	**47.44**±2.87	**46.36**±3.80	**35.80**±5.03

	MNIST-75sp	Graph-SST2	MOLSIDER	MOLHIV
ERM	14.94±3.27	81.44±0.59	57.57±1.56	76.20±1.14
Attention	16.44±3.78	81.57±0.71	56.99±0.54	75.84±1.33
Top-k Pool	15.02±3.08	79.78±1.35	60.63±1.52	73.01±1.65
SAG Pool	19.34±1.73	80.24±1.72	61.29±1.31	73.26±0.84
ASAP	15.14±3.58	81.57±0.84	55.77±1.34	73.81±1.17
GroupDRO	15.72±4.35	81.29±1.44	56.31±1.15	75.44±2.70
IRM	18.74±2.43	81.01±1.13	57.10±0.92	74.46±2.74
V-REx	18.40±1.12	81.76±0.08	57.76±0.78	75.62±0.79
DIR	17.38±3.52	83.29±0.53	57.74±1.63	77.05±0.57
GSAT	20.12±1.35	82.95±0.58	60.82±1.36	76.47±1.53
GIL	**21.94**±0.38	**83.44**±0.37	**63.50**±0.57	**79.08**±0.54

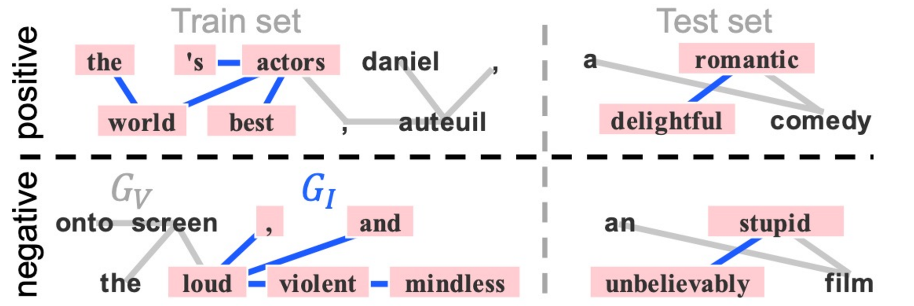
- **The learned invariant subgraphs are accurate**



Precision@5 of discovering the ground-truth invariant subgraphs on SP-Motif.

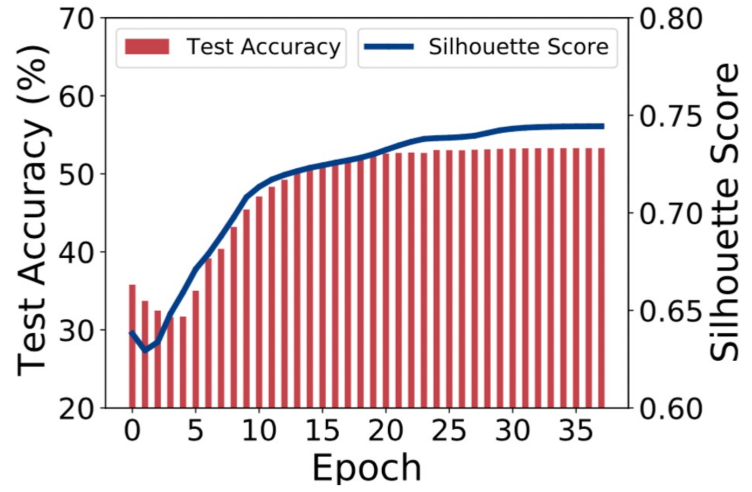


Showcases on synthetic SP-Motif.

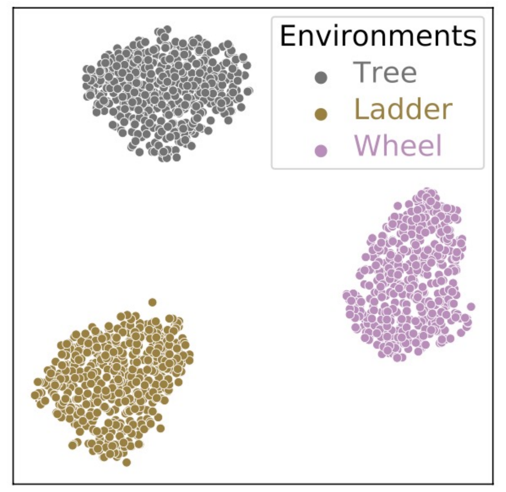


Showcases on real-world Graph-SST2.

- **Deeper Analysis of Environment Inference**



The environment inference and invariant learning can mutually promote each other.



The inferred environment is also accurate.

Contact: Haoyang Li, Email: lihy18@mails.tsinghua.edu.cn