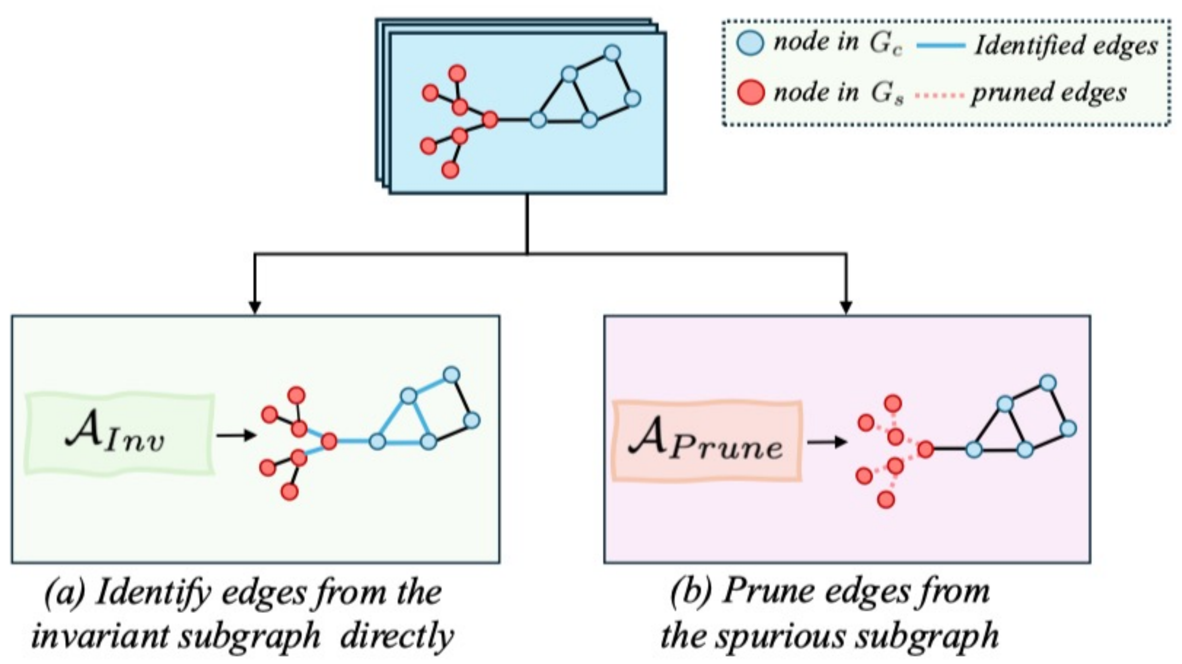
Pruning Spurious Subgraphs for Graph Out-of-Distribution Generalization

Tianjun Yao, Haoxuan Li, Yongqiang Chen, Tongliang Liu, Le Song, Eric Xing, Zhiqiang Shen

Code Available

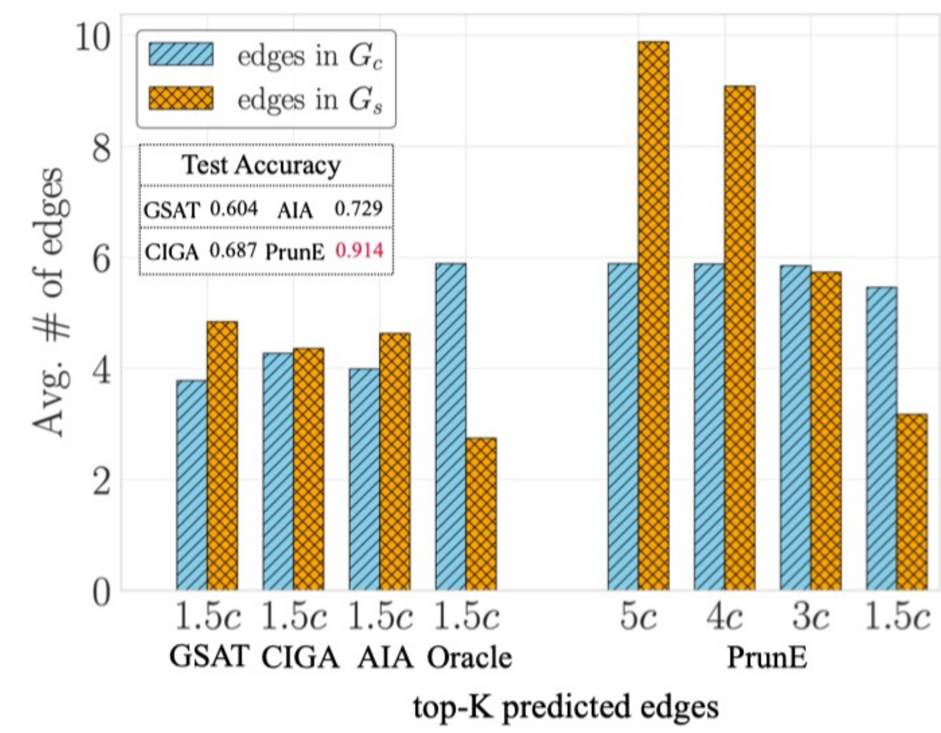
We show that pruning spurious (uninformative) edges can be more effective than directly identifying invariant subgraphs!



(a) Identify edges from the invariant subgraph directly

(b) Prune edges from the spurious subgraph

Illustration of Two Learning Pradigms



Pruning spurious edges lead to better preservation of invariant subgraph

Two Simple Regularizers

Overall goal: regularize the model to remove spurious edges

- Graph-size constraint: encourage the model to remove uninformative edges.

$$\mathcal{L}_e = \mathbb{E}_{\mathcal{G}} \left(\frac{\sum_{(i,j)\in\mathcal{E}} \widetilde{\mathbf{A}}_{ij}}{|\mathcal{E}|} - \eta \right)^2$$

- ϵ-probability Alignment: Align each edge probability to close to zero

$$\mathcal{L}_s = \mathbb{E}_{\mathcal{G}} \frac{1}{|\mathcal{E}_s|} \sum_{e_{ij}\in\mathcal{E}_s} |p_{ij} - \epsilon|$$

- Final Objective: ERM loss tend to learn all useful features, while the two regularizers prunes uninformative edges, and retain invariant substructure

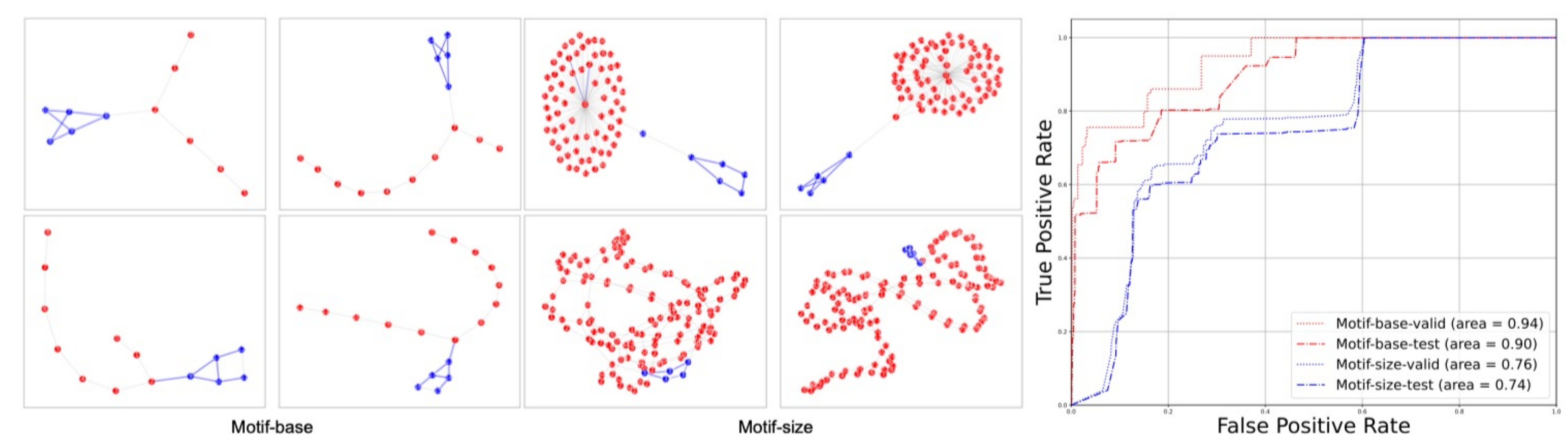
$$\mathcal{L} = \mathcal{L}_{GT} + \lambda_1 \mathcal{L}_e + \lambda_2 \mathcal{L}_s$$

Experimental results

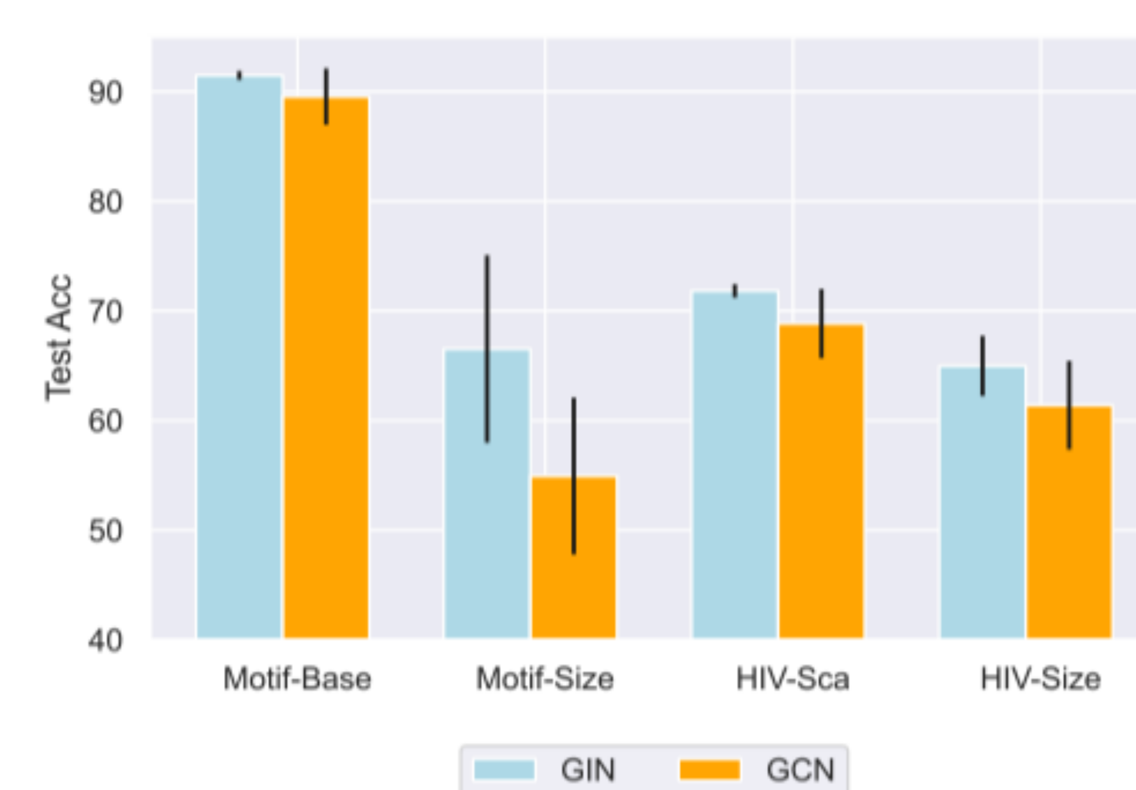
- PrunE achieves SOTA performance across various domains

Method	GOODMotif		GOODHIV		EC50			OGBG-Molbbbp	
	base	size	scaffold	size	scaffold	size	assay	scaffold	size
ERM	68.66±4.25	51.74±2.88	69.58±2.51	59.94±2.37	62.77±2.14	61.03±1.88	64.93±6.25	68.10±1.68	78.29±3.76
IRM	70.65±4.17	51.41±3.78	67.97±1.84	59.00±2.92	63.96±3.21	62.47±1.15	72.27±3.41	67.22±1.15	77.56±2.48
GroupDRO	68.24±8.92	51.95±5.86	70.64±2.57	58.98±2.16	64.13±1.81	59.06±1.50	70.52±3.38	66.47±2.39	79.27±2.43
VREx	71.47±6.69	52.67±5.54	70.77±2.84	58.53±2.88	64.23±1.76	63.54±1.03	68.23±3.19	68.74±1.03	78.76±2.37
DropEdge	45.08±4.46	45.63±4.61	70.78±1.38	58.53±1.26	63.91±2.56	61.93±1.41	73.79±4.06	66.49±1.55	78.32±3.44
G-Mixup	59.66±7.03	52.81±6.73	70.01±2.52	59.34±2.43	61.90±2.08	61.06±1.74	69.28±1.36	67.44±1.62	78.55±4.16
FLAG	61.12±5.39	51.66±4.14	68.45±2.30	60.59±2.95	64.98±0.87	64.28±0.54	74.91±1.18	67.69±2.36	79.26±2.26
LiSA	54.59±4.81	53.46±3.41	70.38±1.45	52.36±3.73	62.60±3.62	60.96±1.07	69.73±0.62	68.11±0.52	78.62±3.74
DIR	62.07±8.75	52.27±4.56	68.07±2.29	58.08±2.31	63.91±2.92	61.91±3.92	66.13±3.01	66.86±2.25	76.40±4.43
DisC	51.08±3.08	50.39±1.15	68.07±1.75	58.76±0.91	59.10±5.69	57.64±1.57	61.94±7.76	67.12±2.11	56.59±10.09
CAL	65.63±4.29	51.18±5.60	67.37±3.61	57.95±2.24	65.03±1.12	60.92±2.02	74.93±5.12	68.06±2.60	79.50±4.81
GREA	56.74±9.23	54.13±10.02	67.79±2.56	60.71±2.20	64.67±1.43	62.17±1.78	71.12±1.87	69.72±1.66	77.34±3.52
GSAT	60.42±9.32	53.20±8.35	68.66±1.35	58.06±1.98	65.12±1.07	61.90±2.12	74.77±4.31	66.78±1.45	75.63±3.83
CIGA	68.71±10.9	49.14±8.34	69.40±2.39	59.55±2.56	65.42±1.53	64.47±0.73	74.94±1.91	64.92±2.09	65.98±3.31
AIA	72.91±5.62	55.85±7.98	71.15±1.81	61.64±3.37	64.71±0.50	63.43±1.35	76.01±1.18	**70.79**±1.53	81.03±5.15
PrunE	**91.48***±0.40	**66.53***±8.55	**71.84***±0.61	**64.99***±1.63	**67.56***±0.34	**65.46***±0.88	**78.01***±0.42	70.32±1.73	**81.59**±5.35

- PrunE provides **explainability** for graph-data under distribution shift



- Perform better with more **expressive** GNN



- Robust to Concept Shift

Method	GOODHIV	GOODMotif
	size	base
ERM	63.26±2.47	81.44±0.45
IRM	59.90±3.15	80.71±0.46
VRex	60.23±1.70	81.56±0.35
GSAT	56.76±7.16	76.07±3.48
GREA	60.07±5.40	78.27±4.29
CIGA	73.62±0.86	81.68±3.01
AIA	74.21±1.81	82.51±2.81
PrunE	**79.50**±1.57	**90.28**±1.72