CS460 PROJECT

ON THE EXPRESSIVE POWER OF GEOMETRIC GNN

Soumya Dasgupta, Akankshya Nayak

National Institute of Science Education and Research

March 10, 2023

Introduction

- Graphical Neural Networks (GNNs) can be simply defined as topologies of neural networks that operate on graphs
- Research into the expressive capabilities of GNNs has primarily relied on the Weisfeiler-Lehman (WL) test [Leman 2018]
- WL test is not suitable for analyzing geometric graphs that are embedded in Euclidean space
- Geometric WL (GWL) test explores the expressive power of geometric GNNs while accounting for physical symmetries [Joshi et al. 2023]
  Invariant layers have limited expressivity as they fail to distinguish one-hop identical geometric graphs
  Equivariant layers distinguish a larger class of graphs by propagating geometric information beyond local neighbourhoods
  Higher order tensors and scalarization enable maximally powerful geometric GNNs
  GWL's discrimination-based perspective is equivalent to universal approximation

INTRODUCTION

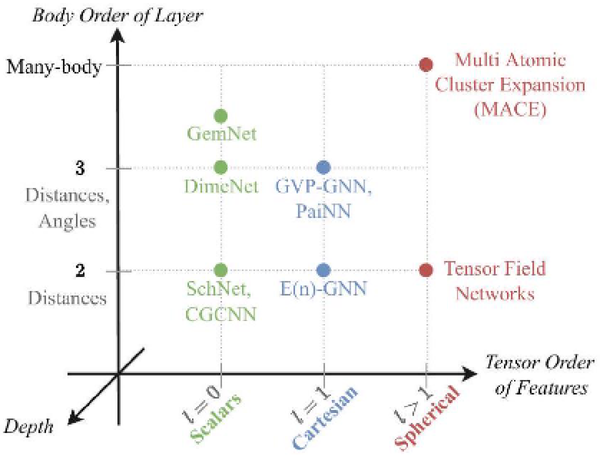

Figure. Axes of geometric GNN expressivity: (1) Scalarisation body order: increasing body order of scalarisation builds expressive local neighbourhood descriptors; (2) Tensor order: higher order spherical tensors determine the relative orientation of neighbourhoods; and (3) Depth: deep equivariant layers propagate geometric information beyond local neighbourhoods [Joshi et al. 2023]

Related Work

- Standard GNNs are at most expressive as the WL algorithm [Leman 2018, Xu et al. 2018, Morris et al. 2018]
- k-WL hierarchy: generalizing the WL algorithm for classifying k-tuples of vertices [Grohe 2017]
- N-WL hierarchy: built on high-order subgraphs within neighborhood aggregation [Wang et al. 2023]
- Architectures such as TFN, GemNet, and GVP-GNN can serve as universal approximators of continuous, $\mathfrak{G}$-equivariant, or $\mathfrak{G}$-invariant multiset functions on point clouds [Dym and Maron 2020, Villar et al. 2021, Klicpera, Becker, and Gunnemann 2021, Jing et al. 2020]

Baseline

WL test

- 0th iteration: WL assigns a colour $c_i^{(0)} \in C$ where C is a countable space of colours. Nodes having the same features are given the same colour.
- In the subsequent iterations the colours of each node are updated in the following manner:

$$\boldsymbol{c}_i^{(t)} := \mathrm{HASH}\left(\boldsymbol{c}_i^{(t-1)}, \left\{\left\{\boldsymbol{c}_j^{(t-1)} \mid j \in \mathcal{N}_i\right\}\right\}\right)$$

  The test terminates when the partition of the nodes induced by the colours becomes stable.
- Given two graphs $\mathcal{G}$ and $\mathcal{H}$, if there exists some iteration t for which $\left\{\left\{c_i^{(t)} \mid i \in \mathcal{V}(\mathcal{G})\right\}\right\} \left\{\left\{c_i^{(t)} \mid i \in \mathcal{V}(\mathcal{H})\right\}\right\}$, then the graphs are not isomorphic. Otherwise inconclusive.

Baseline

Isomorphism in Geometric graphs, GWL, IGWL

Isomorphism: Two geometric graphs $\mathcal{G}$ and $\mathcal{H}$ are geometrically isomorphic if there exists an attributed graph isomorphism b such that the geometric attributes are equivalent, up to global group actions $Q_{\mathfrak{g}} \in \mathfrak{G}$ and $\vec{\boldsymbol{t}} \in T(d)$

$$\left(\boldsymbol{s}_i^{(\mathcal{G})}, \vec{\boldsymbol{v}}_i^{(\mathcal{G})}, \vec{\boldsymbol{x}}_i^{(\mathcal{G})}\right) = \left(\boldsymbol{s}_{b(i)}^{(\mathcal{H})}, \boldsymbol{Q}_{\mathfrak{g}} \vec{\boldsymbol{v}}_{b(i)}^{(\mathcal{H})}, \boldsymbol{Q}_{\mathfrak{g}} \left(\vec{\boldsymbol{x}}_{b(i)}^{(\mathcal{H})} + \vec{\boldsymbol{t}}\right)\right) \quad \text{for all } i \in \mathcal{V}(\mathcal{G})$$

In simple words, two geomtric graphs are isomorphic if the two graphs super-impose on each other after some rotations and translations.

GWL

- 0th iteration

$$c_i^{(0)} := \text{HASH}\left(\boldsymbol{s}_i\right), \quad \boldsymbol{g}_i^{(0)} := \left(c_i^{(0)}, \vec{\boldsymbol{v}}_i\right)$$

- tth iteration

$$\boldsymbol{g}_i^{(t)} := \left(\left(c_i^{(t-1)}, \boldsymbol{g}_i^{(t-1)}\right), \left\{\left\{\left(c_j^{(t-1)}, \boldsymbol{g}_j^{(t-1)}, \vec{\boldsymbol{x}}_{ij}\right) \mid j \in \mathcal{N}_i\right\}\right\}\right)$$

$$c_i^{(0)} := \text{I-HASH}^t(\boldsymbol{g}_i^{(t)})$$

IGWL

$$c_i^{(t)} := \text{I} - \text{HASH}\left(\left(c_i^{(t-1)}, \vec{\boldsymbol{v}}_i\right), \left\{\left\{\left(c_j^{(t-1)}, \vec{\boldsymbol{v}}_j, \vec{\boldsymbol{x}}_{ij}\right) \mid j \in \mathcal{N}_i\right\}\right\}\right).$$

Baseline

Charactersing the expressive power of gGNN

Theorem Any pair of geometric graphs distinguishable by a $\mathfrak{G}$-equivariant GNN is also distinguishable by GWL.
Proposition: $\mathfrak{G}$-equivariant GNNs have the same expressive power as GWL if the following conditions hold: (1) The aggregation AGG is an injective, $\mathfrak{G}$-equivariant multiset function. (2) The scalar part of the update UPD_s is a $\mathfrak{G}$-orbit injective, $\mathfrak{G}$-invariant multiset function. (3) The vector part of the update UPD_v is an injective, $\mathfrak{G}$-equivariant multiset function. (4) The graph-level readout f is an injective multiset function.

Experiments

Depth and non local properties

Here we train $\mathfrak{G}$-equivariant and $\mathfrak{G}$-invariant GNNs on two k chains graph in each run. Each pair of k chain graphs contain $k+2$ nodes with k nodes arranged in a line and are different by the orientation of the two end points. In every run the number of layers are increased. Since the k chain graphs are $\left(\left\lfloor \frac{k}{2} \right\rfloor + 1\right)$ hop different so theoretically GWL requires only $\left(\left\lfloor \frac{k}{2} \right\rfloor + 1\right)$ iterations to distinguish them

(k=4 chains) GNN Layer	Number of Layers				
	[k/2]	**[k/2]+1=3**	**[k/2]+2**	**[k/2]+3**	**[k/2]+4**
IGWL	**50%**	**50%**	**50%**	**50%**	**50%**
SchNet	50.00±0.00	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
DimeNet	50.00±0.00	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
GWL	**50%**	**100%**	**100%**	**100%**	**100%**
E-GNN	50.00±0.00	50.00±0.0	50.00±0.0	50.00±0.0	100.00±0.0
GVP-GNN	50.00±0.00	100.00±0.0	100.00±0.0	100.00±0.0	100.00±0.0
TFN	50.00±0.00	50.00±0.0	50.00±0.0	80.0±24.5	85.0±22.9
MACE	50.00±0.00	90.0±20.0	90.0±20.0	95.0±15.0	95.0±15.0

Table. Results for experiment on depth

Experiments

Higher order tensor and rotational symmetry

An L-fold symmetric structure does not change when rotated by an angle $\frac{2\pi}{L}$ about a point in 2D and an axis in 3D. Two distinct rotated versions of each L-fold symmetric structure are taken and single layer $\mathfrak{G}$-equivariant GNNs are trained on them to classify them.

GNN Layer	Rotational Symmetry			
	2-fold	**3-fold**	**5-fold**	**10-fold**
E-GNN (L=1)	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
GVP-GNN (L=1)	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
TFN/MACE (L=1)	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
TFN/MACE (L=2)	100.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
TFN/MACE (L=3)	100.00±0.0	100.00±0.0	50.00±0.0	50.00±0.0
TFN/MACE (L=5)	100.00±0.0	100.00±0.0	100.00±0.0	50.00±0.0
TFN/MACE (L=10)	100.00±0.0	100.00±0.0	100.00±0.0	100.00±0.0

Table. Result for experiment on tensor order. L denotes the tensor order taken

References I

- Dym, Nadav and Haggai Maron (2020). *On the Universality of Rotation Equivariant Point Cloud Networks*. DOI: 10.48550/ARXIV.2010.02449. URL: https://arxiv.org/abs/2010.02449.
- Grohe, Martin (Aug. 2017). *Descriptive Complexity, Canonisation, and Definable Graph Structure Theory*. Cambridge University Press. DOI: 10.1017/9781139028868. URL: https://doi.org/10.1017/9781139028868.
- Jing, Bowen et al. (2020). *Learning from Protein Structure with Geometric Vector Perceptrons*. DOI: 10.48550/ARXIV.2009.01411. URL: https://arxiv.org/abs/2009.01411.
- Joshi, Chaitanya K. et al. (2023). *On the Expressive Power of Geometric Graph Neural Networks*. DOI: 10.48550/ARXIV.2301.09308. URL: https://arxiv.org/abs/2301.09308.
- Klicpera, Johannes, Florian Becker, and Stephan Gunnemann (2021). "GemNet: Universal Directional Graph Neural Networks for Molecules". In: *ArXiv* abs/2106.08903.
- Leman, Adrien (2018). "THE REDUCTION OF A GRAPH TO CANONICAL FORM AND THE ALGEBRA WHICH APPEARS THEREIN". In.
- Morris, Christopher et al. (2018). *Weisfeiler and Leman Go Neural: Higher-order Graph Neural Networks*. DOI: 10.48550/ARXIV.1810.02244. URL: https://arxiv.org/abs/1810.02244.

References II

Villar, Soledad et al. (2021). "Scalars are universal: Equivariant machine learning, structured like classical physics". In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato et al. Vol. 34. Curran Associates, Inc., pp. 28848–28863. URL: https://proceedings.neurips.cc/paper/2021/file/f1b0775946bc0329b35b823b86eeb5f5-Paper.pdf.

Wang, Qing et al. (2023). "$\mathscr{N}$-WL: A New Hierarchy of Expressivity for Graph Neural Networks". In: *The Eleventh International Conference on Learning Representations*. URL: https://openreview.net/forum?id=5cAI0qXxyv.

Xu, Keyulu et al. (2018). *How Powerful are Graph Neural Networks?* DOI: 10.48550/ARXIV.1810.00826. URL: https://arxiv.org/abs/1810.00826.