Extended Abstract Track

Augmented Geometric Multi Resolution Analysis for Graph features representation

Editors: List of editors' names

Abstract

We present an Augmented Geometric Multi-Resolution Analysis (AGMRA) framework for graph feature representation learning. Our approach extends the classical GMRA framework to graph domains, enabling adaptive multi-scale representation feature for graph structures. We overcome the limitation of original GMRA framework by using dense extraction of wavelet features, which we call AGMRA. By incorporating graph-specific geometric properties and augmenting the analysis with domain-specific features, we demonstrate the effectiveness of our method on real-world network datasets, showing improvements in performance compared to traditional graph representation learning.

Keywords: geometric representation, graph learning, multi-resolution analysis, wavelets

1. Introduction

1.1. Graph Feature Learning

Graph feature learning underpins applications with complex relational structure—e.g., social networks and protein graphs—where signals span local neighborhoods up to meso-/macro-scale communities. Classical graph methods often miss these hierarchical, multi-scale patterns, which is especially limiting in dynamic threat networks. Broadly, approaches split into (i) transductive methods (e.g., DeepWalk (Perozzi et al. (2014)), node2vec (Grover and Leskovec (2016))) that learn embeddings for a fixed graph but struggle to generalize to new nodes/graphs, and (ii) inductive methods (e.g., GNNs/GCNs/GraphSAGE (Kipf and Welling (2017))) that generalize via node features and topology but may under-capture rich multi-scale structure in highly nonstationary graphs.

We address these gaps by bringing Geometric Multi-Resolution Analysis (GMRA, Allard et al. (2012)) to graph feature learning. GMRA yields a data-adaptive, multiscale decomposition with efficient coarse-to-fine representations. While effective in vision and signal domains (Wang et al. (2016)), its use on graphs is underexplored. We extend GMRA to graph-derived features, producing geometry-aware, multiscale embeddings that align local and global structure and are amenable to downstream learning on evolving graphs.

1.2. Geometric Multi-Resolution Analysis

In many graph-learning problems, data live in high-dimensional, non-Euclidean spaces with complex global structure but locally low-dimensional regularity (Bronstein et al. (2017)). We adopt Geometric Multi-Resolution Analysis (GMRA) to organize graph-derived features across scales: GMRA builds data-adaptive, multiscale dictionaries that approximate local affine structure, yielding geometry-aware representations that capture both coarse and fine variations. Practically, we partition the feature space into a dyadic tree (via spatial or

Algorithm 1: Augmented Forward GMRA

Input: GMRA tree $\mathcal{T}$ with nodes, each node v associated with scale and cell index j, k and corresponding $\{(\Phi_{j,k}, \Psi_{j,k}, c_{j,k})\} = \{(\Phi_v, \Psi_v, c_v)\}$;
Data point $x \in \mathbb{R}^D$; Leaf set $\mathcal{L}$

Output: Augmented GMRA feature vector $q_x \in \mathbb{R}^M$

$q_x \leftarrow \emptyset$;
foreach *leaf node* $\ell \in \mathcal{L}$ *(in fixed order)* **do**
  $x_\ell \leftarrow x$;
  $v \leftarrow \ell$; `// current node; parent(v) returns parent or ∅`
  $p_J = \Phi_\ell^* (x_\ell - c_\ell)$
  $q_J = \Phi_v^* \Phi_v \, p_J$;
  $j \leftarrow \text{level}(v)$; `// also finest scale J`
  while *parent*$(v) \neq \varnothing$ **do**
    $v \leftarrow \text{parent}(v)$; `// moves to scale j-1`
    $j \leftarrow j - 1$;
    $p_{j-1} \leftarrow \Phi_v^* \Phi_\ell p_J + \Phi_v^* (c_\ell - c_v)$;
    $q_{j-1} \leftarrow \Psi_v^* \Phi_v p_{j-1}$;
  end
  Append $(q_0, q_1, \ldots, q_J)$ to q_x;
end
return q_x

graph-based clustering), fit local PCA models in each cell, and use the forward GMRA transform to encode each point by a coarse projection plus geometric wavelet refinements; see (Allard et al. (2012)) for details and guarantees.

However, vanilla GMRA is unsuitable for learning pipelines: (i) representation dimensionality varies across samples, (ii) encoding is slow to batch because bases are node-specific, and (iii) coefficients are not globally aligned, so identical indices may represent different subspaces. We address these issues with Augmented GMRA, which produces fixed-dimensional, globally aligned, and computationally efficient multiscale embeddings (Section 2.1).

2. Methods

2.1. Augmented Geometric Multi-Resolution Analysis

Given a trained GMRA—i.e., a dyadic tree with scaling bases $\{\Phi_{j,k}\}$ and wavelet bases $\{\Psi_{j,k}\}$ (citegmra)—we modify the forward transform to produce fixed-dimensional, aligned, and batch-friendly features. **Augmented GMRA** encodes every point via the same leaf-centric traversal: for each leaf ℓ, follow its path to the root and project the point onto the affine subspaces along that path (optionally using the nearest leaf to the point). Repeating this for all leaves and concatenating the per-leaf coefficients yields a globally consistent multiscale embedding, removing variable dimensionality, enabling efficient batching, and preserving GMRA's geometric interpretability. See Algorithm 1.

Table 1: MNIST classification results comparing GMRA, PCA, and Raw Pixels features.

	Accuracy			F1 Score		
Classifier	AGMRA	PCA	Raw Pixels	AGMRA	PCA	Raw Pixels
KNN	0.8800	0.8225	0.9000	0.8803	0.8185	0.8998
Naive Bayes	0.7525	0.5925	0.5775	0.7509	0.5687	0.5485
Neural Network	0.9125	0.8750	0.9225	0.9120	0.8745	0.9224
Random Forest	0.8575	0.8550	0.9200	0.8569	0.8522	0.9195
SVM	0.8975	0.8400	0.9275	0.8970	0.8510	0.9268

Table 2: CIFAR-10 classification results comparing GMRA, PCA, and Raw Pixels features.

	Accuracy			F1 Score		
Classifier	AGMRA	PCA	Raw Pixels	AGMRA	PCA	Raw Pixels
KNN	0.2600	0.2950	0.2750	0.2358	0.2708	0.2458
Naive Bayes	0.3150	0.2100	0.2850	0.2964	0.1884	0.2598
Neural Network	0.3250	0.2800	0.3000	0.3271	0.2793	0.3008
Random Forest	0.3850	0.2300	0.3500	0.3832	0.2187	0.3439
SVM	0.3450	0.4000	0.4000	0.3282	0.3826	0.3884

2.2. Graph Neural Networks Integration

Because GMRA/AGMRA operate on Euclidean point clouds, we apply the transform to node or graph-level embeddings and integrate it with GNN workflows. Concretely, we stack AGMRA on a transductive baseline (Node2vec): (i) run Node2vec to obtain fixed-dimensional node embeddings; (ii) build a GMRA tree on these embeddings and apply *Augmented GMRA* (Sec. 2.1) to produce fixed-dimensional, globally aligned multiscale features; (iii) concatenate AGMRA features with the original Node2vec vectors and train a simple classifier (e.g., MLP) for node classification or link prediction. This retains Node2vec's local connectivity signal while injecting multiscale structure captured by AGMRA.

3. Experiments

3.1. AGMRA on MNIST and CIFAR10 classification

In order to verify the correctness of our AGMRA implementation, we empirically test it on MNIST and CIFAR10 classification tasks. We first train a GMRA structure on the training set of MNIST and CIFAR10, then we apply the AGMRA transform to the training and test sets to obtain the augmented features. Finally, we train a simple MLP classifier on top of the AGMRA features and evaluate its performance on the test set. The results are shown in Table 1 and 2. Features extracted by AGMRA outperform PCA features (with roughly same number of dimensions) and are competitive with raw pixel features (no dimension reduction).

Table 3: Node classification accuracy (validation).

Method	Accuracy
Only node2vec	0.74
Only GMRA	0.75
node2vec + GMRA (ours)	**0.77**

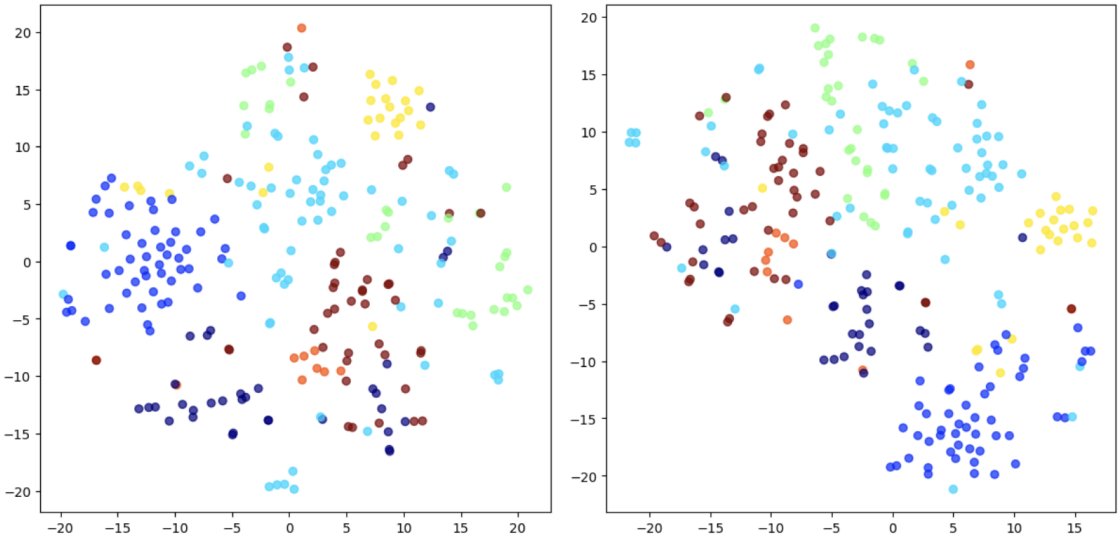

Figure 1: TSNE visualization of node embeddings before (left) and after (right) applying AGMRA. Notice the disperse of the light blue and brown clusters are reduced.

3.2. AGMRA on Graph Node Classification

We evaluate the performance of our AGMRA-enhanced graph neural network (GNN) on Cora dataset (McCallum (2024)) for node classification task. We use 10% for train and 90% for test set. We use Covertree to build GMRA tree with finest scale $J = 3$ and prune any cell with less than 15 points. We compare the results of our method with baseline GNN models without AGMRA. The results are shown in Table 3. We also provide a TSNE visualization of the node embeddings before and after applying AGMRA in Figure 1.

4. Conclusion

We presented Augmented Geometric Multi-Resolution Analysis (AGMRA), a novel extension of GMRA for graph feature representation learning. By enforcing fixed-dimensional, globally aligned multiscale embeddings, AGMRA overcomes key limitations of vanilla GMRA and integrates seamlessly with graph neural networks. Empirical results on MNIST, CIFAR10, and Cora node classification demonstrate AGMRA's effectiveness in capturing rich geometric structure and improving downstream performance. Future work includes exploring AGMRA in inductive graph learning methods.

References

William K. Allard, Guangliang Chen, and Mauro Maggioni. Multi-scale geometric methods for data sets ii: Geometric multi-resolution analysis. *Applied and Computational Harmonic Analysis*, 32(3):435–462, 2012. ISSN 1063-5203. doi: https://doi.org/10.1016/j.acha.2011.08.001. URL https://www.sciencedirect.com/science/article/pii/S1063520311000868.

Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: Going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, July 2017. ISSN 1558-0792. doi: 10.1109/msp.2017.2693418. URL http://dx.doi.org/10.1109/MSP.2017.2693418.

Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 855–864, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450342322. doi: 10.1145/2939672.2939754. URL https://doi.org/10.1145/2939672.2939754.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*, 2017. URL https://openreview.net/forum?id=SJU4ayYgl.

Andrew McCallum. Cora, 2024. URL https://dx.doi.org/10.21227/jsg4-wp31.

Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '14, page 701–710. ACM, August 2014. doi: 10.1145/2623330.2623732. URL http://dx.doi.org/10.1145/2623330.2623732.

Yi Wang, Guangliang Chen, and Mauro Maggioni. High-dimensional data modeling techniques for detection of chemical plumes and anomalies in hyperspectral images and movies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(9):4316–4324, 2016. doi: 10.1109/JSTARS.2016.2539968.