

# Summary of Graph Based Deep Learning Architectures

Jarred M. Kvamme

*University of Idaho*

*Department of Bioinformatics and Computation Biology*

September 22, 2023

# Contents

<b>LIST OF TABLES</b>	<b>1</b>
<b>LIST OF FIGURES</b>	<b>1</b>
<b>1 Graph Learning</b>	<b>2</b>
1.1 Generic Graph Neural Networks (GNN) . . . . .	2
1.2 Graph Convolutional Neural Network (GCN) . . . . .	2
1.3 Graph Attention Neural Network (GAT) . . . . .	3
<b>2 Geometric Autoencoder Architectures</b>	<b>3</b>
2.1 Graph AutoEncoder (GAE) . . . . .	3
2.2 Graph Attention Autoencoder (GATE) . . . . .	4
2.3 Community Detection Based on Unsupervised Attributed Network Em- bedding (CDBNE) . . . . .	5

## List of Figures

1	Graph AutoEncoder (GAE) Architecture . . . . .	4
2	Graph Attention Autoencoder (GATE) Architecture . . . . .	5
3	CDBNE Architecture . . . . .	6

# 1 Graph Learning

Graph structured data is found in many real-world applications, such as biological networks. Recent advancements in machine learning have allowed neural network models to take advantage of the rich relational information present in graph structured data. Geometric models are generically referred to as graph neural networks (GNNs) in the deep learning paradigm. These models consist of "message-passing" layers - layers which consist of a differentiable aggregation function which iteratively updates node features based on local structure, and then passes the aggregated "message" through dense linear layer and nonlinear activation [1, 2].

Recently, [3] proposed graph convolutional neural networks (GCNs), as a generalization of convolutional neural networks to graph structured data. GCN performs convolutions on a graph using a 1-step spectral filter which aggregates node and neighbor information together by smoothing node features over a local neighborhood in the graph. However, the learned spectral filters of GCN depend on the eigendecomposition of the spectral graph. Thus the filters are specific to the input graph structure and cannot be directly applied to a graph of with different structure.

Another type of layer is the graph attention network (GAT) [4]. GAT integrates Graph Neural Networks (GNNs) with the traditional self-attention mechanism [5]. When applied to graph-structured data, the attention coefficients serve as a valuable tool for quantifying the significance of a node's neighbors, enabling the model to prioritize important relationships among nodes.

## 1.1 Generic Graph Neural Networks (GNN)

A generic graph neural network consists of message passing layers in which structural information is used to determine which nodes allowed to share feature information based on local linkages. A simple GNN model takes as input an attributed graph  $\mathcal{G}(E, V, \mathbf{X})$  with  $V = V_{i=1}^N$  vertices,  $e_{i,j} \in E$  edges between vertices, and node attributes  $\mathbf{X}$ . The following is the generic architecture of a message-passing layer.

$$\mathbf{h}_u = \psi \left( \mathbf{x}_u \cdot \text{AGGREGATE}_{v \in \mathbb{N}}(\phi(\mathbf{x}_u, \mathbf{x}_v), e_{u,v}) \right) \quad (1)$$

where  $\mathbf{h}_u$  is the updated representation of node  $u$ ,  $\mathbf{x}_u$  is the feature vector for node  $u$ ,  $\mathbf{x}_v$  is the feature vector of node  $v$ , and  $e_{u,v} = (u, v)$  is the edge between nodes  $u$  and  $v$ , and  $\phi$  and  $\psi$  are differentiable activation activation functions.

## 1.2 Graph Convolutional Neural Network (GCN)

[3] presented a spectral view of graph convolutions as the 1-D signal Fourier transform in the spectral graph domain [6]. In the GCN architecture, the graph structure is encoded by the normalized Laplacian  $\tilde{\mathbf{L}}$  representation of the adjacency matrix  $\mathbf{A}$ .

$$H = \phi(\tilde{\mathbf{L}}\mathbf{X}\mathbf{W})$$

where  $\tilde{\mathbf{L}}$  is the normalized graph Laplacian with added self loops such that  $\tilde{\mathbf{L}} = \mathbf{D}^{1/2}\mathbf{I} + \mathbf{A}\mathbf{D}^{1/2}$

### 1.3 Graph Attention Neural Network (GAT)

The GAT model is similar to GNN/GCN in that the contribution of source nodes to a given target is a weighted average. However, in the GAT model, the weight applied to each node is a function of the source and targets nodes embeddings. This differs from classical GNN/GCN in which the weights are the node degrees. This additional flexibility allows GAT to capture more complex relationships between nodes such as their features or local structure. The GAT layer is defined as

## 2 Geometric Autoencoder Architectures

Autoencoders are generative models which consist of two main components (i) an encoder which learns a lower dimension representation of the data which are embedded at a bottleneck layer and (ii) a decoder which reconstructs the embeddings back to their original input. In geometric learning, the autoencoder framework has been routinely employed for a number of deep learning task including community detection [7, 8, 9, 10]. However, these methods are typically restricted to reconstructing either the graph or the node attributes but rarely both.

### 2.1 Graph AutoEncoder (GAE)

GAE is the non-probabilistic version of the variational autoencoder first proposed by [7] for node classification. GAE takes as input an attributed graph and passes the graph and node-features through two GCN encoder layers which learn a representation  $\mathbf{Z}$  which can be used to do a number of tasks. The "decoder" of GAE consists of a simple non-trainable layer they call the "dot-product decoder" as it takes the outer product of  $\mathbf{Z}$  followed by logistic activation to reconstruct the input graph. The encoder layers are given by

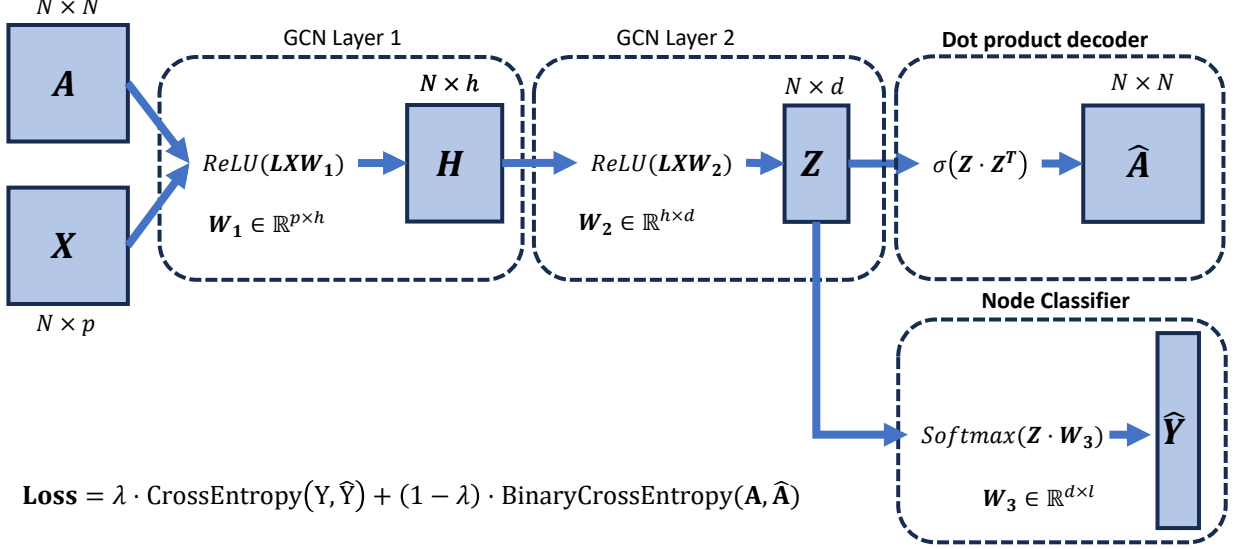
$$h_u = ReLU \left( \sqrt{\frac{1}{d_u}} \sum_{v \in \mathcal{N}(u)} \sqrt{\frac{1}{d_v}} \mathbf{W} h_v \right)$$

The decoder layer is given as

$$\hat{\mathbf{A}} = \sigma(\mathbf{Z}\mathbf{Z}^T)$$

Figure 1: Graph AutoEncoder (GAE) Architecture

**Graph AutoEncoder (GAE) - Kipf and Welling, 2016**  
**Original Task: Node Classification**



Kipf, Thomas N., and Max Welling. "Variational graph auto-encoders." *arXiv preprint arXiv:1611.07308* (2016).

## 2.2 Graph Attention Autoencoder (GATE)

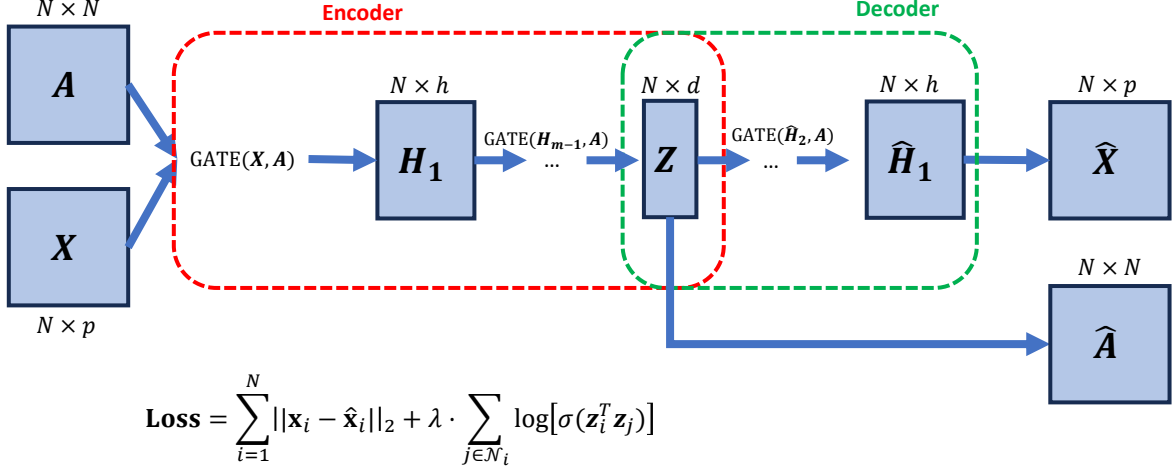
The authors [11] propose graph attention autoencoder (GATE) which uses conventional GAT layers in an autoencoder framework to reconstruct both the node attributes and graph structure for both supervised and unsupervised graph learning tasks **Figure 2** for node classification and link prediction. The reconstructed node attributes are taken as the last layer of the decoder model and the reconstructed graph is found using dot-product decoder used in GAE. The attention mechanism is computed as follows

$$\alpha_{ij} = \frac{\sigma(\exp(\mathbf{a}_s^T \phi[\mathbf{W}h_i] + \mathbf{a}_v^T \phi[\mathbf{W}h_j]))}{\sum_{k \in \mathcal{N}(i)} \sigma(\exp(\mathbf{a}_s^T \phi[\mathbf{W}h_i] + \mathbf{a}_v^T \phi[\mathbf{W}h_k]))}$$

where  $\mathbf{W}$ ,  $\mathbf{a}_s^T$ , and  $\mathbf{a}_v^T$  are learnable parameters,  $\phi$  denotes an activation function, and  $\sigma$  denotes the logistic sigmoid function. In the original work of [4]  $\phi$  was given as the LeakyReLU activation. The attention mechanism adopted by [11] is a slight variation of the original self-attention proposed [4] which uses separate attention weights for nodes  $i$  and  $j$

Figure 2: Graph Attention Autoencoder (GATE) Architecture

**Graph Attention AutoEncoder (GATE) – Salehi and Davulcu, 2019**  
**Original Task: Node Classification**



Salehi, Amin, and Hasan Davulcu. "Graph attention auto-encoders." *arXiv preprint arXiv:1905.10715* (2019).

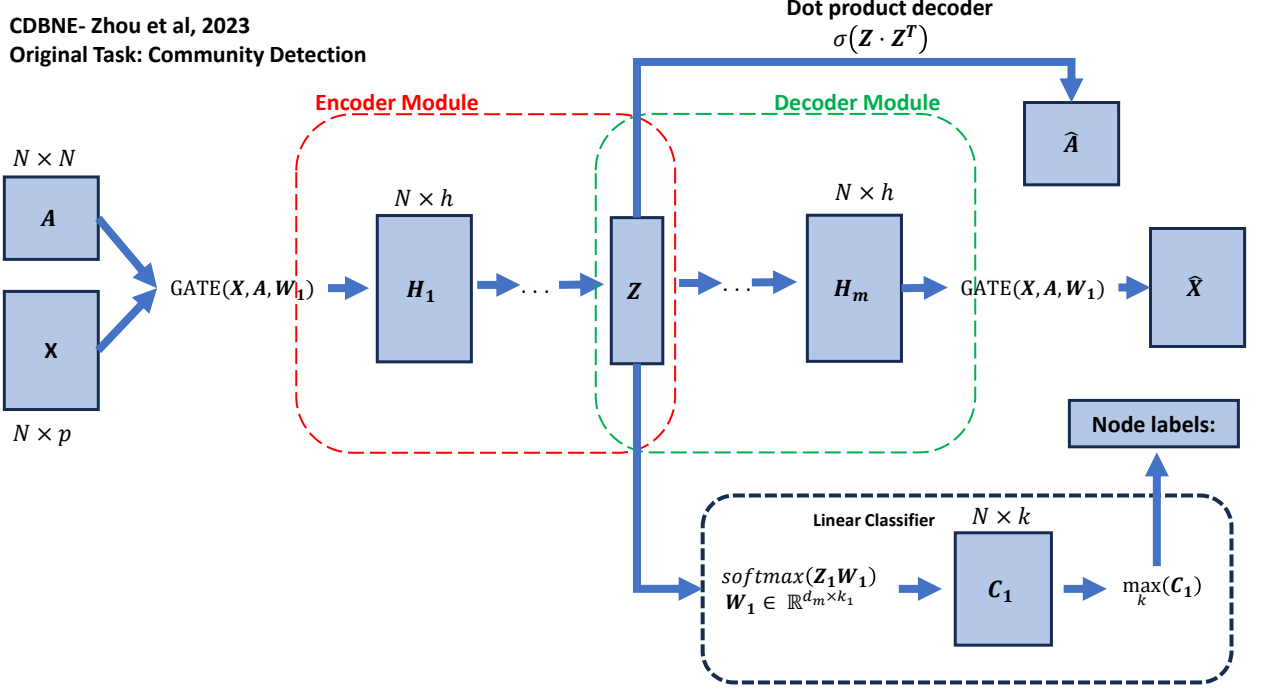
### 2.3 Community Detection Based on Unsupervised Attributed Network Embedding (CDBNE)

CDBNE is a model proposed by [9] which adopts GATE under a different loss function for community detection on graphs. Specifically, the CDBNE uses three loss components: (i) A modularity loss  $L_M$  computed as the modularity of inferred graph communities, (ii) the reconstruction loss  $L_R$  which consists of the reconstruction of the node attributes  $L_X$  and the reconstruction of the adjacency matrix  $L_A$  (iii) a clustering loss  $L_C$  based on the KL divergence between the inferred communities and some set of "soft" initial predictions of the communities. The final objective function

$$L_{\text{total}} = L_R + \delta L_C - \gamma L_M$$

which has two tuning parameters  $\delta$  and  $\gamma$ .

Figure 3: CDBNE Architecture



## References

- [1] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE transactions on neural networks*, 20(1):61–80, 2008.
- [2] Marco Gori, Gabriele Monfardini, and Franco Scarselli. A new model for learning in graph domains. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, volume 2, pages 729–734. IEEE, 2005.
- [3] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
- [4] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.
- [5] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.
- [6] Si Zhang, Hanghang Tong, Jiejun Xu, and Ross Maciejewski. Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6(1):1–23, 2019.

- [7] Thomas N Kipf and Max Welling. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308*, 2016.
- [8] Chun Wang, Shirui Pan, Guodong Long, Xingquan Zhu, and Jing Jiang. Mgae: Marginalized graph autoencoder for graph clustering. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 889–898, 2017.
- [9] Xinchuang Zhou, Lingtao Su, Xiangju Li, Zhongying Zhao, and Chao Li. Community detection based on unsupervised attributed network embedding. *Expert Systems with Applications*, 213:118937, 2023.
- [10] Juexin Wang, Anjun Ma, Yuzhou Chang, Jianting Gong, Yuexu Jiang, Ren Qi, Cankun Wang, Hongjun Fu, Qin Ma, and Dong Xu. scgcn is a novel graph neural network framework for single-cell rna-seq analyses. *Nature communications*, 12(1):1882, 2021.
- [11] Amin Salehi and Hasan Davulcu. Graph attention auto-encoders. *arXiv preprint arXiv:1905.10715*, 2019.