 doi:10.20944/preprints202209.0226.v1

A Survey of GNN in Bioinformation Data

ZHU Zhenyi
The Hong Kong University of Science and Technology
zzhubh@connect.ust.hk

Abstract

With the development of data science, more and more machine learning technologies have been designed to solve complicated and challenging real-world tasks containing a large volume of data. And many significant real-world datasets contain data in the form of networks or graphs. Graph Neural Networks is one of the powerful machine learning tools, which could provide a perfect solution to processing a large amount of non-Euclidean data. And because most bio information data in bioinformatics is in the non-Euclidean domain, Graph Neural Networks could then directly be applied to solve problems in bioinformatics. Much research has been done in the field of GNN, and there are also some surveys (Zhang et al., 2021a) related to GNN and its applications. However, there has been little research focusing on GNN in bioinformatics, and we think in the future we could better utilize GNN in the field of biology, so we would like to write a literature review to help take a comprehensive look at GNN and their applications in the field of bioinformatics. In this paper, we would first introduce SOTA models in Graph Neural Networks, and second, introduce their applications in bio information. And then we would provide future directions of Graph Neural Networks in bioinformatics.

Keywords : Graph Neural Networks, Non-Euclidean Data, Bioinformation

1 Introduction

In this digital era, a huge amount of data emerges every day, and there has been a significant proportion of the data comes from non-Euclidean space. Examples are the World Wide Web, knowledge graphs, social networks, and protein-interaction networks, etc. To analyze these data, we need to use GNN as a powerful tool.

GNN is a very important helper for processing graph-based data. It has evolved for several years and has lots of SOTA models. For instance, there have been RecGNN, GCN, GAT, etc. And these models could be applied in lots of cases, from social network analysis to chemistry analysis. And with the significant amounts of bio information data such as omics, genes, images, and signal data accumulated, machine learning has become a strong tool to tackle tasks in bioinformatics. Many Machine learning tools have been applied in the bioinformatic area, such as convolutional neural networks, deep neural networks, recurrent neural networks, etc.

Nowadays, many technology companies have developed a platform for analyzing bio information. For instance, IBM has developed an artificial intelligence system named Watson to analyze Oncology(Ferrucci et al., 2010). And Google has developed the DeepMind AI system, which could predict the 3D structure of the protein (Silver et al., 2016). In this literature review, we would like to focus on analyzing how GNN acts in bioinformatics and how to use GNN to process bio information data.

With the innovation in data science emerging every day, there are generating more and more techniques in GNN. However, there is few research about providing a comprehensive understanding of GNN and their application in bio information. So, we would like to write the literature review to conclude these breakthroughs in GNN and then give people more insights into its applications. In this literature review, we provide a detailed introduction to SOTA GNN, their current applications in bioinformatics, and their future direction of them.

2 Graph Neural Networks

We know that deep learning has become one of the most powerful tools in solving a wide range of machine learning tasks, ranging from image classification and object detection to speech recognition and natural language processing. The data in the above tasks can be represented in a Euclidean space.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0226.v1

However, there exists a large number of data generated from non-Euclidean domains such as graph data (Wu et al., 2020b). A graph does not exist in Euclidian space and cannot be mapped to a 2-D or 3-D Euclidian space. And a graph usually does not keep a fixed form. So, graph data is difficult to be analyzed by traditional graph analysis methods and algorithms. Some clustering methods such as the K-means method and Highly Connected Components method, searching algorithms such as Breadth First Search and Depth First Search, spanning-tree algorithms such as Prim's algorithm and shortest path algorithms such as Nearest Neighbor algorithm and Dijkstra's algorithm could not perform graph level machine learning task such as classification. Then a neural network that could be directly applied to graph-level tasks was proposed by Gori(Gori et al., 2005) and then further elaborated by Scarselli (Scarselli et al., 2008).

2.1 Simple Introduction to GNN

First, we would provide some basic concepts and definitions of Graph Neural Networks (GNN). The basic concepts in GNN come from Graph theory, where we need to consider vertices and edges in the graph. And the graph is a mathematical structure to help analyze the pair-wise relationship between objects and entities. And we usually define a graph as G (V, E), where V is the set of vertices and E is the set of edges. The idea of a graph could better help us to deal with some abstract concepts like interactions and relationships, and then assist us by simplifying the problems into much simpler representations. So, we need to introduce a new method to process these graphs.

The basic working structure of GNN is shown in Figure 1, which comes from THOMAS KIPF. The input data is in the form of graphs, and then use the activation function to process the graphs, finally, we could get the output which is usually a generated graph modeling the relationships between different objects.

2.2 Applications of GNN

Before discussing GNN in bioinformatics, we would like to provide several applications of GNN in other cases. GNN could roughly solve three problems: node classification, link prediction, and graph classification. And in the node classification case, the task is usually trained in a semi-supervised way. A typical application is like Facebook friends' relationships analysis, citation networks analysis, etc. In the task of link prediction, GNN is usually applied in building a recommendation system, where we need to constantly tune the system and track the users' preferences. And in the graph classification case, GNN could be used in many different areas such as the classification of molecular structure in the field of chemistry, and the classification of gene structure in the field of biology.

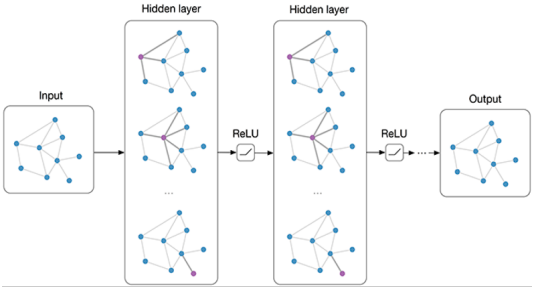

Figure 1: Basic structure of GNN (**?**).

Then we would like to show some more concrete applications of GNN. For instance, (Wu et al., 2021a) and (Wei et al., 2020) found the GNNs such as RecGNN could be utilized in Natural Language Processing (Wu et al., 2021b). And the GNN could also be in the field of Computer Vision. For instance, (Shi and Rajkumar, 2020) and (Luo et al., 2020) found that GNN can be used in object detection. (Coupeau et al., 2022) proposed that GNN can also do semantic segmentation tasks. And the most interesting application for us is that GNN could perform well in the zero-shot learning (ZSL). (Gao and Xu, 2020) (Wang et al., 2021) (Chen et al., 2022) And in zero-shot task, GNN can help model relationships such as relationships between text descriptions well. And there are still other domains that GNN could be applied to such as traffic control (Hu et al., 2020), logic reasoning (Zhang et al., 2020) (Zhang et al., 2019), adversarial attack prevention (Chen et al., 2020a), social influence prediction (Song et al., 2021), etc. We could easily see that the development of GNN has a great influence in many real-world fields.

3 Methodologies in Graph Neural Networks

In this section, we would like to introduce several SOTA models and their models in graph neural networks.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0226.v1

3.1 Deep Learning in Bioinformation Data

Before introducing benchmarks in Graph Neural Networks, we would like to first introduce deep learning methods in bioinformatics. Deep learning has been widely used to solve many realistic tasks. And with the development of bioinformatics, deep learning becomes the tool to develop new treatments, identify disease susceptibility genes, discover drugs, do forensic analysis, improve crops, and manage biodiversity (Agbachi, 2017).

Many deep learning methods could be applied to bioinformation data (Min et al., 2017). For instance, we could use deep neural networks to process Omics, Biomedical signals, and images. Then we could solve tasks such as predicting protein structure, regulating gene expression, segmentation, anomaly detection, and brain coding ((Min et al., 2017)). Figure 2 shows the working flow. We could also use convolutional neural networks

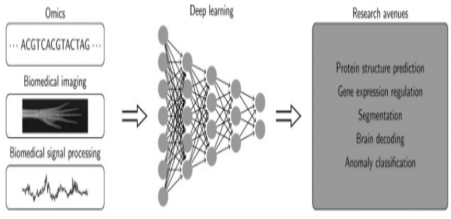

Figure 2: Application of deep learning methods in bioinformatics (Min et al., 2017)

to process X-rays for finger joint detection (Lee et al., 2015). And recurrent neural networks could be used to process biomedical signals such as EEG signals to perform lapse detection (Davidson et al., 2007).

3.2 Benchmarks in GNN

There have been several benchmarks and techniques in GNN during recent years. Compared to traditional methods, these benchmarks could be stronger tools in bioinformatics. In this section, we would like to introduce some SOTA models such as GCN, RecGNN, Graph Attention Networks, etc.

3.2.1 Recurrent Graph Neural Network (RecGNN)

To better analyze the link problem the Recurrent Graph Neural Network was proposed Recurrent Graph Neural Networks could be the pioneer version of Graph Neural Networks, which is mainly based on an assumption of the Banach Fixed-Point Theorem. And RecGNN defines a parameterized function f_w as follows:

$$x_n = f_w(l_n, l_{co[n]}, x_{ne[n]}, l_{ne[n]})$$

And l_n, $l_{co[n]}$, $x_{ne[n]}$, $l_{ne[n]}$ respectively represent the feature of the current node [n], the edges of the node [n], the state of the neighboring nodes, and the features of the neighboring nodes. Finally, after several iterations, we could produce an output to make a decision for each node by using the state of the final node. And the final output function can be defined as:

$$O_n = g_w(x_n, l_n)$$

Figure 3 shows the basic structure of RecGNN. RecGNN could have many applications in different

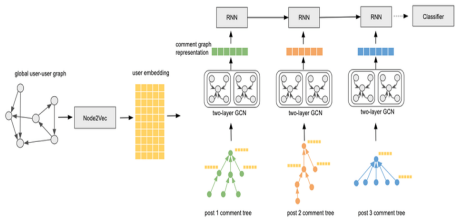

Figure 3: Structure of the Recurrent GNN (Huang et al., 2021). Features can be extracted from the left model and then fed to RNN cells, then we could model the temporal propagation.

cases. Andrei (Nicolicioiu et al., 2019) proposed a Recurrent Space-time Graph Neural Network to better process data in both space and time, and then help to capture interactions. And in the field of bioinformatics, (Elbasani et al., 2021) proposed a RecGNN named GCRNN to help predict the compound-protein interaction.

3.2.2 Graph Convolutional Networks (GCN)

Graph Convolutional Networks is purposed to extract the spatial features of the topology. And the two main methods to complete the purpose: the first method lies in the spatial domain (also called vertex domain), which needs to figure out all neighbors of each vertex; the second method lies in the domain of edges. And different from input and output in normal neural networks, our GCN could process a graph directly, which means the input to neural networks in GCN will be a graph. And GCN could infer values and make predictions for every node in the graph by utilizing the neighbor nodes in the calculation. Figure 4 shows simple comparison of the two networks' structures. Figure 5 shows one of the practical working structures of GCN under the background of forecasting retweet (Vijayan

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0226.v1

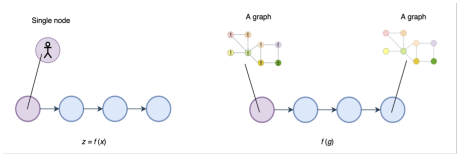

Figure 4: Simple comparison of NN and GCN.

and Mohler, 2018), where we could conclude that the basic working structure of GCN is like CNN, and we just need to set the whole graph as the input signal. Application GCN is a common strategy for

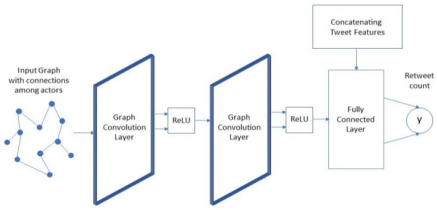

Figure 5: Structure of GCN. The input is usually graphs and then it passes one convolutional layer and activation function.

many tasks such as the prediction of biomedical interactions (Kishan et al., 2021) (Nguyen et al., 2021) and associations (Wang et al., 2020).

And there are two kinds of GCN: spectral-based GCN and spatial-based GCN. And the spectral one is derived from the angle of graph signal processing. The spectral-based GCN could exploit the principle of Laplacian and Fourier transform to map the irregular structure of a graph to a regular Euclidean space for convolution operation. The spatial-based GCN could directly utilize the information dissemination mechanism on the graph to complete the convolution operation, and its propagation method is similar to the original GNN (Zhang et al., 2021a).

3.2.3 Graph Spectral Convolutional Network

Graph Convolutional Networks is recognized as spectral graph convolution. And it is a localized first-order approximation of spectral graph convolutions with the propagation rule below, where A is the adjacency matrix, D is the degree matrix, W is the weight matrix, sigma is the activation function and H is the output vector. We have the following propagation rule:

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)})$$

where

$$\tilde{A} = A + I_n$$
$$\tilde{D}_{ii} = \Sigma_j + \tilde{A}_{ij}$$

The first version of Graph was proposed by (Bruna et al., 2013), which contains multiple convolutional layers. (Defferrard et al., 2016) solved the high computational complexity of the model by proposing the ChebNets model, which uses a K-degree polynomial filter in the convolutional layer. And (Kipf and Welling, 2016) simplified the above ChebNets by truncating the Chebyshev polynomial to one time.

And the model could be widely used in bioinformatics. Federico (Baldassarre et al., 2021)use it to assess protein model quality. And [38] found many of its applications in biological data.

3.2.4 Graph Spatial Convolutional Network

The intuition of Spatial Convolution Networks is similar to Convolutional Neural Networks. In short, the idea of convolution on an image is to sum the neighboring pixels around a center pixel, specified by a filter with parameterized size and learnable weight. Spatial Convolutional Network adopts the same idea by aggregating the features of neighboring nodes into the center node. Figure 6 shows the draft about convolution on a regular graph and arbitrary graph. While spectral-based GNN is based on

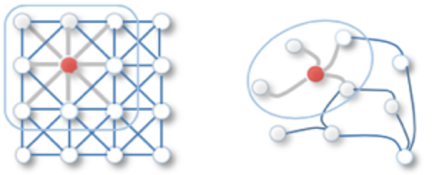

Figure 6: Left subgraph: The convolution on the regular graph. Right subgraph: The convolution on the arbitrary graph. Figure from (Wu et al., 2020b)

the convolution theorem, the spatial model starts from the node domain and aggregates each central node. (Atwood and Towsley, 2016) proposed that the convolution is a diffusion process (DCNN) between nodes and the structure of layer m is defined as follows:

$$H^{m+1} = f(Wp^k H^m)$$

Where p^k is the k-hop reachability probability between nodes in a random walk, W is a parameter reached by learning from the model. However, DCNN couldn't be applied to large-scale graphs well because the complexity of computation is too high. (Hamilton et al., 2017) proposed GraphSage in 2017 to adapt the model to large graph networks by doing random sampling on neighboring nodes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0226.v1

And the Graph Spatial Convolutional Network could be used to do metric learning on some bioinformation datasets such as the brain networks dataset (Ktena et al., 2017a) and drug dataset [33].

3.2.5 Graph Attention Network

The Graph Attentions Network (GAT) was purposed by (Velickovic et al., 2017), which is a non-spectral graph network utilizing node's spatial information for graph learning. The basic element of GAT is the Graph Attention layer. And the principle and process of GAT contain about five steps: linear transformation, attention coefficients computation, attention coefficients normalization, output feature computation, and multiple attention mechanisms computations. So, GAT could be regarded as an advanced GCN. And the key steps are shown in the following Figure 7.

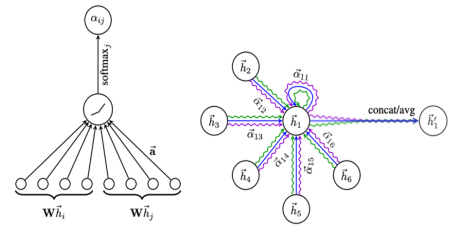

Figure 7: left subfigure shows the model employs and parametrizes the mechanism a(whi,whj) with a weight vector a. and the right subfigure shows the multi-head attention by node 1 on its neighborhood(Nicolicioiu et al., 2019).

Graph Attention Network has a large number of practical applications. For instance, one of it is applications in bioinformatics is utilizing graph attention mechanisms to a gene contact network to distinguish the significance among various neighboring genes of each gene and then predict co-expression in supervised predation to distill biological insights by learning the representation of gene (Zhang et al., 2022).

3.2.6 Graph Auto-encoder

The autoencoder (AE) is a popular tool in the field of unsupervised learning. And Graph Auto-encoder is derived from a sparse autoencoder (SAE) proposed by (Tian et al., 2014), where L2 reconstruction loss was used. And based on SAE, (Wang et al., 2016) proposed a structured deep network embedding (SDNE), which improved the L2 reconstruction used in SAE. And then [5] and [46] showed a GAE for network embedding, where the encoder uses graph convolutional layers to get a network for each node. And the decoder could compute the pair-wise distance with the network embeddings. And the network is trained by minimizing the difference between the real adjacency matrix and the reconstructed adjacency matrix. The basic structure of Graph Auto-encoder (GAE) is shown in Figure 8. And the GAE has lots of applications in bioinformatics. (Wu et al., 2021c), (Li et al., 2021), (Ding et al., 2021) respectively show that GAE could be used to predict the associations in IncRNA-disease and miRNA-disease. GAE could also predict the unobserved node features on biological networks (Hasibi and Michoel, 2021). And it could denoise the protein-protein interaction network (Yao et al., 2020).

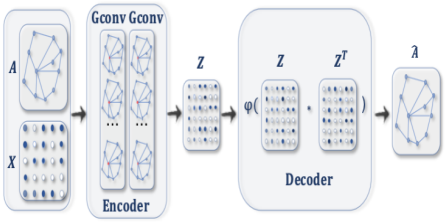

Figure 8: Basic structure of GAE (Source:(Wu et al., 2020b)). A and X are inputted into the encoder, and then with the application of a non-linear activation function, the decoder could reconstruct the adjacency matrix of the graph.

3.2.7 Graph Reinforcement Learning

Reinforcement Learning is a machine learning training method that depends on the reward strategy and has been utilized in more and more cases to help us better perceive and interpret the environment. Nowadays, the ideas in Reinforcement Learning gradually creep into the field of graph learning.

The graph convolution would adapt to the changes and dynamics of the underlying graph of the multi-agent environment. And the relation kernels could capture the interplay between an agent with the multi-agent environment. (Jiang et al., 2018) In this task, we usually would provide many multi-agent environments. And because the convolution could increase the receptive field of an agent, so considering only neighboring agents are effective and efficient. Figure 9 shows the learning curve comparison in routing. And there has been a large number of applications of Graph Reinforcement Learning. for instance, it could be used for traffic signal control (Peng et al., 2021) (Nishi et al., 2018), and financial portfolio man-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0226.v1

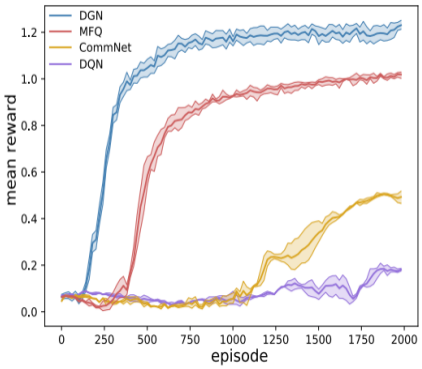

Figure 9: Learning curve comparison in routing. (Jiang et al., 2018)

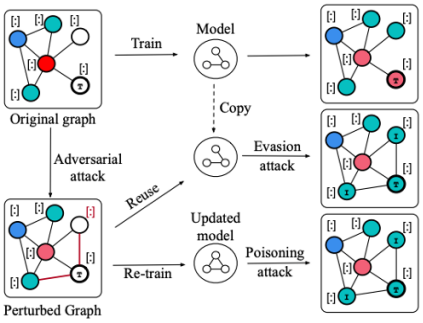

Figure 10: Example of the evasion attack and the poisoning attack where input is the graphs (Chen et al., 2020b).

agement(Soleymani and Paquet, 2021). And in the field of bioinformatics, it can perform medicine prediction (Wang et al., 2019a).

3.2.8 Graph Adversarial Methods

In recent years, deep learning on graphs has achieved great success in different graph analysis tasks such as node classification, link prediction, and graph clustering (Chen et al., 2020b). However, these models face the problem that the uncertainty and unreliability are against well-designed inputs. So, we need to develop some graph adversarial methods to solve the defeat. There has been a surge of research on studying the vulnerability and security of graph models. (Zügner et al., 2018) proposed the first study about adversarial learning, where node features and graph structure have little perturbation and the target classifiers are easily fooled nodes. (Wang et al., 2019b) later propose an improved GCN to better its robustness. There are mainly three types of attacks: white-box attack, gray-box attack, and black-box attack. Figure 10 shows an example of an evasion attack and poisoning attack. Graph Adversarial Methods also have many applications. It could be utilized to improve social recommendation (Yu et al., 2020). And in the field of bioinformatics, for instance, (Isallari and Rekik, 2021) found that it could be a powerful tool in functional brain connectivity tasks.

4 GNN in Bioinformation Data

4.1 Introduction to Bioinformatics

The main task of bioinformatics focuses on the collection, classification, storage, and analysis of bioinformation data by using computer science. And the bioinformation data is a treasure in analyzing real-world tasks in bioinformatics. There have many different types of data in the field of bioinformatics: DNA sequences of genes, three-dimensional structures of proteins, amino acid sequences of proteins, genetics of diseases, protein-nucleic acid complexes, etc. And there exist many databases such as Genbank, PDB, Pfam, Motif, etc. (Benson et al., 2012) And in this section, we would like to introduce how could CNN be applied to solve these questions.

4.2 Applications of GNN in bioinformatics

In this part, we would like to introduce several applications of GNN in bioinformation. In this part, we would like to mainly focus on three typical applications: GNN in drug discovery, GNN in medical images, and GNN in disease association prediction and classification.

4.2.1 GNN in Drug Discovery

Drug discovery is a significant task in bioinformatics. Drug development task contains drug target and candidate drug determination, lead compound discovery (Vohora and Singh, 2018). However, many of nowadays machine learning methods can't directly process structural information well. So, many methods in GNN were developed to solve this limitation.

For the specific task of prediction of molecular properties, (Duvenaud et al., 2015) proposed the initial application of molecular properties learning. and (Liu et al., 2019) combined GCN with MT-DNNs to further improve prediction accuracy.

For the specific task about prediction of protein structure, (Zamora-Resendiz and Crivelli, 2019) proposed a methodology for large data sets, which could make the application more transferable by using natural spatial representation substituting for previous 2D or 3D models. And (Ingraham et al.,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0226.v1

2019) proposed a protein design framework based on one graph attention method to construct a conditional generation model. And there have been two types of methods to predict protein function. One is based on protein structure (Ioannidis et al., 2019), and the other is based on PPI. (Gligorijević et al., 2021) modeled the structure of the protein to the graph.

For the specific task of drug response prediction, the combination of genomics data and drug information promoted the development of personalized medicine. (Huang et al., 2020) proposed a GCN to predict the association between miRNA expression profile and drug structure fingerprint information.

4.2.2 GNN in Medical Images

Medical images are a significant data source in the field of bioinformatics, especially for disease diagnosis and treatment. And the main task in it is similar to that of a normal computer vision task. To be more specific, GNN could be used to visualize histological images (Levy et al., 2020), analyze surgical images (Zhang et al., 2018), predict disease ((Parisot et al., 2017) (Anirudh and Thiagarajan, 2019) (Zhang and Bellec, 2020) (Zhang et al., 2021b)), perform image segmentation ((Gopinath et al., 2019),(Wang et al., 2019c)), do research on brain connectivity (Ktena et al., 2017b), etc.

For the specific task of image segmentation, we know that there have existed many CNN-based image segmentation technology. To segment 3D images and predict points' movement, (Ma et al., 2018) proposed a gated graph neural network for segmentation. And (Avelar et al., 2020) utilized the Graph Attention Networks to perform superpixel image classification. (Yun et al., 2021) proposed a GNN-based method for online handwritten diagram recognition.

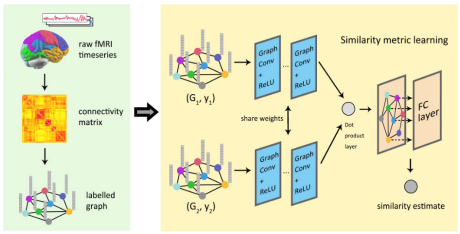

Figure 11: Similarity metric learning of GNN in brain graphs (Zhang et al., 2021a).

For the specific task of brain connectivity research, we know that regular images are the default data in most machine learning models. And (Ktena et al., 2018) introduced the initial research on brain connection research by proposing depth graph networks. Figure 11 shows the process for analysis of functional magnetic resonance imaging (fMRI) in the above study.

4.2.3 GNN in Disease Association Prediction and Classification

Discovering disease association is a very important task in the field of bioinformatics. And many methodologies in machine learning and matrix decomposition have been utilized to complete the task. For instance, Luo, J proposed to use transduction learning to predict the disease-associate miRNAs in 2016. However, these methods have some limitations because they usually ignore the representation of topological relationships between entities.

For the specific task of disease-gene association, GNN could infer the interaction between cells(Rao et al., 2021), predict disease state (Ravindra et al., 2020), predict cancer (Iglehart and Silver, 2009), and analyze the mechanism of cancer (Schulte-Sasse et al., 2019).

For the specific task RNA-Disease association, we know that the task is important because if we could identify these associations, then we could explore the pathogenesis of complicated diseases. (Pan and Shen, 2019) proposed a semi-supervised model named DimiG, which could predict the association between miRNAs and diseases. And the following Figure 12 shows the working flow of DimiG. (Wu et al., 2020a) used the autoencoder method to reconstruct node features.

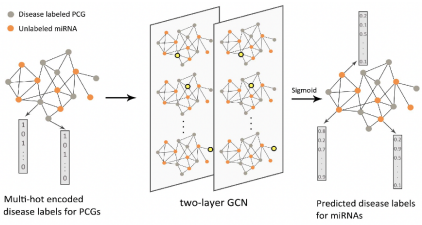

Figure 12: DimiG Working Flow (Pan and Shen, 2019).

5 Future Directions

5.1 Interpretability of GNN in Biominforatics

In the field of machine learning, many theories couldn't be interpreted well, so we need to figure out the basic theoretical explanation behind machine learning. And in GNN, there are still many

applications unexplainable to us. However, the entities and relationships in GNNs are always related to different types of objects existing in the real world, so we could better get an interpretable result from GNNs (Selsam et al., 2018).

5.2 Exploration of Deep Structure

We know that in the field of deep learning, most networks have many layers. For instance, VGG16 (Qassim et al., 2018) has 16 layers and Resnet (He et al., 2016) has 152 layers, etc. However, in the field of GNNs, the number of layers is always small (Zhou et al., 2020). And we know that if we increase the complexity of the model, we could better train the model and predict more precisely. So, we need to focus on increasing the number of layers in the GNN structure in the next step.

6 Conclusion

In this paper, we introduce a lot of SOTA models and methodologies in GNN and their applications in the field of bioinformatics. GNN is a strong tool to help us solve real-world tasks, especially in biology, bioinformatics, and biological medicine. And with the development of new skills nowadays, more and more strong GNN models emerge every day. However, there are still many difficulties in the application of GNN models. We think in the future GNN could be more widely in solving real-world tasks, especially in processing bioinformation datasets.

References

CPE Agbachi. 2017. Pathways in bioinformatics: A window in computer science. *Int J Comput Trends Technol*, 49(2):83–90.

Rushil Anirudh and Jayaraman J Thiagarajan. 2019. Bootstrapping graph convolutional neural networks for autism spectrum disorder classification. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3197–3201. IEEE.

James Atwood and Don Towsley. 2016. Diffusion-convolutional neural networks. *Advances in neural information processing systems*, 29.

Pedro HC Avelar, Anderson R Tavares, Thiago LT da Silveira, Clíudio R Jung, and Luís C Lamb. 2020. Superpixel image classification with graph attention networks. In *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 203–209. IEEE.

Federico Baldassarre, David Menéndez Hurtado, Arne Elofsson, and Hossein Azizpour. 2021. Graphqa: protein model quality assessment using graph convolutional networks. *Bioinformatics*, 37(3):360–366.

Dennis A Benson, Mark Cavanaugh, Karen Clark, Ilene Karsch-Mizrachi, David J Lipman, James Ostell, and Eric W Sayers. 2012. Genbank. *Nucleic acids research*, 41(D1):D36–D42.

Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. 2013. Spectral networks and locally connected networks on graphs. *arXiv preprint arXiv:1312.6203*.

Jinyin Chen, Xiang Lin, Hui Xiong, Yangyang Wu, Haibin Zheng, and Qi Xuan. 2020a. Smoothing adversarial training for gnn. *IEEE Transactions on Computational Social Systems*, 8(3):618–629.

Liang Chen, Jintang Li, Jiaying Peng, Tao Xie, Zengxu Cao, Kun Xu, Xiangnan He, and Zibin Zheng. 2020b. A survey of adversarial learning on graphs. *arXiv preprint arXiv:2003.05730*.

Shiming Chen, Ziming Hong, Guosen Xie, Qinmu Peng, Xinge You, Weiping Ding, and Ling Shao. 2022. Gndan: Graph navigated dual attention network for zero-shot learning. *IEEE Transactions on Neural Networks and Learning Systems*.

P Coupeau, J-B Fasquel, and M Dinomais. 2022. On the relevance of edge-conditioned convolution for gnn-based semantic image segmentation using spatial relationships. In *2022 Eleventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 1–6. IEEE.

Paul R Davidson, Richard D Jones, and Malik TR Peiris. 2007. Eeg-based lapse detection with high temporal resolution. *IEEE Transactions on Biomedical Engineering*, 54(5):832–839.

Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. 2016. Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems*, 29.

Yulian Ding, Xiujuan Lei, Bo Liao, and Fang-Xiang Wu. 2021. Predicting mirna-disease associations based on multi-view variational graph auto-encoder with matrix factorization. *IEEE journal of biomedical and health informatics*, 26(1):446–457.

David K Duvenaud, Dougal Maclaurin, Jorge Iparraguirre, Rafael Bombarell, Timothy Hirzel, Alán Aspuru-Guzik, and Ryan P Adams. 2015. Convolutional networks on graphs for learning molecular fingerprints. *Advances in neural information processing systems*, 28.

Ermal Elbasani, Soualihou Ngnamsie Njimbouom, Tae-Jin Oh, Eung-Hee Kim, Hyun Lee, and Jeong-Dong Kim. 2021. Gcrnn: graph convolutional recurrent neural network for compound–protein interaction prediction. *BMC bioinformatics*, 22(5):1–14.

David Ferrucci, Eric Brown, Jennifer Chu-Carroll, James Fan, David Gondek, Aditya A Kalyanpur, Adam Lally, J William Murdock, Eric Nyberg, John Prager, et al. 2010. Building watson: An overview of the deepqa project. *AI magazine*, 31(3):59–79.

Junyu Gao and Changsheng Xu. 2020. Ci-gnn: Building a category-instance graph for zero-shot video classification. *IEEE Transactions on Multimedia*, 22(12):3088–3100.

Vladimir Gligorijević, P Douglas Renfrew, Tomasz Kosciolek, Julia Koehler Leman, Daniel Berenberg, Tommi Vatanen, Chris Chandler, Bryn C Taylor, Ian M Fisk, Hera Vlamakis, et al. 2021. Structure-based protein function prediction using graph convolutional networks. *Nature communications*, 12(1):1–14.

Karthik Gopinath, Christian Desrosiers, and Herve Lombaert. 2019. Graph convolutions on spectral embeddings for cortical surface parcellation. *Medical image analysis*, 54:297–305.

Marco Gori, Gabriele Monfardini, and Franco Scarselli. 2005. A new model for learning in graph domains. In *Proceedings. 2005 IEEE international joint conference on neural networks*, volume 2, pages 729–734.

Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30.

Ramin Hasibi and Tom Michoel. 2021. A graph feature auto-encoder for the prediction of unobserved node features on biological networks. *BMC bioinformatics*, 22(1):1–17.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

Xiaorong Hu, Chenguang Zhao, and Gang Wang. 2020. A traffic light dynamic control algorithm with deep reinforcement learning based on gnn prediction. *arXiv preprint arXiv:2009.14627*.

Di Huang, Jacob Bartel, and John Palowitch. 2021. Recurrent graph neural networks for rumor detection in online forums. *arXiv preprint arXiv:2108.03548*.

Yu-an Huang, Pengwei Hu, Keith CC Chan, and Zhu-Hong You. 2020. Graph convolution for predicting associations between mirna and drug resistance. *Bioinformatics*, 36(3):851–858.

J Dirk Iglehart and Daniel P Silver. 2009. Synthetic lethality—a new direction in cancer-drug development.

John Ingraham, Vikas Garg, Regina Barzilay, and Tommi Jaakkola. 2019. Generative models for graph-based protein design. *Advances in neural information processing systems*, 32.

Vassilis N Ioannidis, Antonio G Marques, and Georgios B Giannakis. 2019. Graph neural networks for predicting protein functions. In *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 221–225. IEEE.

Megi Isallari and Islem Rekik. 2021. Brain graph super-resolution using adversarial graph neural network with application to functional brain connectivity. *Medical Image Analysis*, 71:102084.

Jiechuan Jiang, Chen Dun, Tiejun Huang, and Zongqing Lu. 2018. Graph convolutional reinforcement learning. *arXiv preprint arXiv:1810.09202*.

Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

KC Kishan, Rui Li, Feng Cui, and Anne R Haake. 2021. Predicting biomedical interactions with higher-order graph convolutional networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 19(2):676–687.

Sofia Ira Ktena, Sarah Parisot, Enzo Ferrante, Martin Rajchl, Matthew Lee, Ben Glocker, and Daniel Rueckert. 2017a. Distance metric learning using graph convolutional networks: Application to functional brain networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 469–477. Springer.

Sofia Ira Ktena, Sarah Parisot, Enzo Ferrante, Martin Rajchl, Matthew Lee, Ben Glocker, and Daniel Rueckert. 2017b. Distance metric learning using graph convolutional networks: Application to functional brain networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 469–477. Springer.

Sofia Ira Ktena, Sarah Parisot, Enzo Ferrante, Martin Rajchl, Matthew Lee, Ben Glocker, and Daniel Rueckert. 2018. Metric learning with spectral graph convolutions on brain connectivity networks. *NeuroImage*, 169:431–442.

Sungmin Lee, Minsuk Choi, Hyun-soo Choi, Moon Seok Park, and Sungroh Yoon. 2015. Fingernet: Deep learning-based robust finger joint detection from radiographs. In *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–4. IEEE.

Joshua Levy, Christian Haudenschild, Clark Barwick, Brock Christensen, and Louis Vaickus. 2020. Topological feature extraction and visualization of whole slide images using graph neural networks. In *BIOCOMPUTING 2021: Proceedings of the Pacific Symposium*, pages 285–296. World Scientific.

Zhengwei Li, Jiashu Li, Ru Nie, Zhu-Hong You, and Wenzheng Bao. 2021. A graph auto-encoder model for mirna-disease associations prediction. *Briefings in Bioinformatics*, 22(4).

Ke Liu, Xiangyan Sun, Lei Jia, Jun Ma, Haoming Xing, Junqiu Wu, Hua Gao, Yax Sun, Florian Boulnois, and Jie Fan. 2019. Chemi-net: a molecular graph convolutional network for accurate drug property prediction. *International journal of molecular sciences*, 20(14):3389.

Ao Luo, Xin Li, Fan Yang, Zhicheng Jiao, Hong Cheng, and Siwei Lyu. 2020. Cascade graph neural networks for rgb-d salient object detection. In *European Conference on Computer Vision*, pages 346–364. Springer.

Tengfei Ma, Cao Xiao, Jiayu Zhou, and Fei Wang. 2018. Drug similarity integration through attentive multi-view graph auto-encoders. *arXiv preprint arXiv:1804.10850*.

Seonwoo Min, Byunghan Lee, and Sungroh Yoon. 2017. Deep learning in bioinformatics. *Briefings in bioinformatics*, 18(5):851–869.

Tuan Nguyen, Giang TT Nguyen, Thin Nguyen, and Duc-Hau Le. 2021. Graph convolutional networks for drug response prediction. *IEEE/ACM transactions on computational biology and bioinformatics*, 19(1):146–154.

Andrei Nicolicioiu, Iulia Duta, and Marius Leordeanu. 2019. Recurrent space-time graph neural networks. *Advances in neural information processing systems*, 32.

Tomoki Nishi, Keisuke Otaki, Keiichiro Hayakawa, and Takayoshi Yoshimura. 2018. Traffic signal control based on reinforcement learning with graph convolutional neural nets. In *2018 21st International conference on intelligent transportation systems (ITSC)*, pages 877–883. IEEE.

Xiaoyong Pan and Hong-Bin Shen. 2019. Inferring disease-associated micrornas using semi-supervised multi-label graph convolutional networks. *Iscience*, 20:265–277.

Sarah Parisot, Sofia Ira Ktena, Enzo Ferrante, Matthew Lee, Ricardo Guerrero Moreno, Ben Glocker, and Daniel Rueckert. 2017. Spectral graph convolutions for population-based disease prediction. In *International conference on medical image computing and computer-assisted intervention*, pages 177–185. Springer.

Hao Peng, Bowen Du, Mingsheng Liu, Mingzhe Liu, Shumei Ji, Senzhang Wang, Xu Zhang, and Lifang He. 2021. Dynamic graph convolutional network for long-term traffic flow prediction with reinforcement learning. *Information Sciences*, 578:401–416.

Hussam Qassim, Abhishek Verma, and David Feinzimer. 2018. Compressed residual-vgg16 cnn model for big data places image recognition. In *2018 IEEE 8th annual computing and communication workshop and conference (CCWC)*, pages 169–175. IEEE.

Jiahua Rao, Xiang Zhou, Yutong Lu, Huiying Zhao, and Yuedong Yang. 2021. Imputing single-cell rna-seq data by combining graph convolution and autoencoder neural networks. *Iscience*, 24(5):102393.

Neal Ravindra, Arijit Sehanobish, Jenna L Pappalardo, David A Hafler, and David van Dijk. 2020. Disease state prediction from single-cell data using graph attention networks. In *Proceedings of the ACM conference on health, inference, and learning*, pages 121–130.

Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2008. The graph neural network model. *IEEE transactions on neural networks*, 20(1):61–80.

Roman Schulte-Sasse, Stefan Budach, Denes Hnisz, and Annalisa Marsico. 2019. Graph convolutional networks improve the prediction of cancer driver genes. In *International Conference on Artificial Neural Networks*, pages 658–668. Springer.

Daniel Selsam, Matthew Lamm, Benedikt Bünz, Percy Liang, Leonardo de Moura, and David L Dill. 2018. Learning a sat solver from single-bit supervision. *arXiv preprint arXiv:1802.03685*.

Weijing Shi and Raj Rajkumar. 2020. Point-gnn: Graph neural network for 3d object detection in a point cloud. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1711–1719.

David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. 2016. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489.

Farzan Soleymani and Eric Paquet. 2021. Deep graph convolutional reinforcement learning for financial portfolio management–deeppocket. *Expert Systems with Applications*, 182:115127.

Changhao Song, Bo Wang, Qinxue Jiang, Yehua Zhang, Ruifang He, and Yuexian Hou. 2021. Social recommendation with implicit social influence. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1788–1792.

Fei Tian, Bin Gao, Qing Cui, Enhong Chen, and Tie-Yan Liu. 2014. Learning deep representations for graph clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28.

Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *stat*, 1050:20.

Raghavendran Vijayan and George Mohler. 2018. Forecasting retweet count during elections using graph convolution neural networks. In *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 256–262. IEEE.

Divya Vohora and Gursharan Singh. 2018. *Pharmaceutical medicine and translational clinical research*. Academic Press.

Daixin Wang, Peng Cui, and Wenwu Zhu. 2016. Structural deep network embedding. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1225–1234.

Lei Wang, Zhu-Hong You, Yang-Ming Li, Kai Zheng, and Yu-An Huang. 2020. Gcncda: a new method for predicting circrna-disease associations based on graph convolutional network algorithm. *PLOS Computational Biology*, 16(5):e1007568.

Qunbo Wang, Wenjun Wu, Yongchi Zhao, and Yuzhang Zhuang. 2021. Graph active learning for gcn-based zero-shot classification. *Neurocomputing*, 435:15–25.

Shanshan Wang, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Jun Ma, and Maarten de Rijke. 2019a. Order-free medicine combination prediction with graph convolutional reinforcement learning. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 1623–1632.

Shen Wang, Zhengzhang Chen, Jingchao Ni, Xiao Yu, Zhichun Li, Haifeng Chen, and Philip S Yu. 2019b. Adversarial defense framework for graph neural network. *arXiv preprint arXiv:1905.03679*.

Xiaosong Wang, Ling Zhang, Holger Roth, Daguang Xu, and Ziyue Xu. 2019c. Interactive 3d segmentation editing and refinement via gated graph neural networks. In *International Workshop on Graph Learning in Medical Imaging*, pages 9–17. Springer.

Xinde Wei, Hai Huang, Longxuan Ma, Ze Yang, and Liutong Xu. 2020. Recurrent graph neural networks for text classification. In *2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS)*, pages 91–97. IEEE.

Lingfei Wu, Yu Chen, Heng Ji, and Bang Liu. 2021a. Deep learning on graphs for natural language processing. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2651–2653.

Lingfei Wu, Yu Chen, Kai Shen, Xiaojie Guo, Hanning Gao, Shucheng Li, Jian Pei, and Bo Long. 2021b. Graph neural networks for natural language processing: A survey. *arXiv preprint arXiv:2106.06090*.

Qing-Wen Wu, Jun-Feng Xia, Jian-Cheng Ni, and Chun-Hou Zheng. 2021c. Gaerf: predicting lncrna-disease associations by graph auto-encoder and random forest. *Briefings in bioinformatics*, 22(5):bbaa391.

Ximin Wu, Wei Lan, Qingfeng Chen, Yi Dong, Jin Liu, and Wei Peng. 2020a. Inferring lncrna-disease associations based on graph autoencoder matrix completion. *Computational Biology and Chemistry*, 87:107282.

Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. 2020b. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1):4–24.

Heng Yao, Jihong Guan, and Tianying Liu. 2020. Denoising protein–protein interaction network via variational graph auto-encoder for protein complex detection. *Journal of Bioinformatics and Computational Biology*, 18(03):2040010.

Junliang Yu, Hongzhi Yin, Jundong Li, Min Gao, Zi Huang, and Lizhen Cui. 2020. Enhance social recommendation with adversarial graph convolutional networks. *IEEE Transactions on Knowledge and Data Engineering*.

Xiao-Long Yun, Yan-Ming Zhang, Fei Yin, and Cheng-Lin Liu. 2021. Instance gnn: a learning framework for joint symbol segmentation and recognition in online handwritten diagrams. *IEEE Transactions on Multimedia*, 24:2580–2594.

Rafael Zamora-Resendiz and Silvia Crivelli. 2019. Structural learning of proteins using graph convolutional neural networks. *bioRxiv*, page 610444.

Ke Zhang, Chenxi Wang, Liping Sun, and Jie Zheng. 2022. Prediction of gene co-expression from chromatin contacts with graph attention network. *Bioinformatics*.

Xi Zhang, Lifang He, Kun Chen, Yuan Luo, Jiayu Zhou, and Fei Wang. 2018. Multi-view graph convolutional network and its applications on neuroimage analysis for parkinson's disease. In *AMIA Annual Symposium Proceedings*, volume 2018, page 1147. American Medical Informatics Association.

Xiao-Meng Zhang, Li Liang, Lin Liu, and Ming-Jing Tang. 2021a. Graph neural networks and their current applications in bioinformatics. *Frontiers in genetics*, 12.

Yu Zhang and Pierre Bellec. 2020. Transferability of brain decoding using graph convolutional networks. *BioRxiv*.

Yu Zhang, Loïc Tetrel, Bertrand Thirion, and Pierre Bellec. 2021b. Functional annotation of human cognitive states using deep graph convolution. *NeuroImage*, 231:117847.

Yuyu Zhang, Xinshi Chen, Yuan Yang, Arun Ramamurthy, Bo Li, Yuan Qi, and Le Song. 2019. Can graph neural networks help logic reasoning? *arXiv preprint arXiv:1906.02111*.

Yuyu Zhang, Xinshi Chen, Yuan Yang, Arun Ramamurthy, Bo Li, Yuan Qi, and Le Song. 2020. Efficient probabilistic logic reasoning with graph neural networks. *arXiv preprint arXiv:2001.11850*.

 doi:10.20944/preprints202209.0226.v1

Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. 2020. Graph neural networks: A review of methods and applications. ai open 1 (2020), 57–81.

Daniel Zügner, Amir Akbarnejad, and Stephan Günnemann. 2018. Adversarial attacks on neural networks for graph data. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2847–2856.