SCIENCE CHINA
Information Sciences

• PERSPECTIVE •

February 2022, Vol. 65 126101:1–126101:3
https://doi.org/10.1007/s11432-020-3210-2

Geometric deep learning: progress, applications and challenges

Wenming CAO, Canta ZHENG, Zhiyue YAN & Weixin XIE*

College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China

Received 9 September 2020/Revised 2 December 2020/Accepted 4 March 2021/Published online 14 January 2022

Citation Cao W M, Zheng C T, Yan Z Y, et al. Geometric deep learning: progress, applications and challenges. Sci China Inf Sci, 2022, 65(2): 126101, https://doi.org/10.1007/s11432-020-3210-2

Deep learning technology, like convolutional neural networks, has been widely used in machine learning applications such as image detection and classification, speech processing, natural language processing and cross-modal retrieval [1]. Although deep learning technology has achieved great success in various fields, it has been unable to match non-Euclidean data effectively. In this context, the emergence of geometric deep learning has filled the above-mentioned technical gaps and realized an effective combination of deep learning technology and non-Euclidean data, consisting of manifolds and graphs. Non-Euclidean data is very common, so it is important to study non-Euclidean data with deep learning.

We will introduce geometric deep learning from several aspects. First is an introduction about related progress, in which we briefly introduce basic methods and general frameworks for graph learning and manifold learning in recent years. Second, there are some popular geometric deep learning applications. These applications are usually combined with an expanded graph learning or manifold learning, providing a comprehensive understanding of various fields. The last part discusses some existing challenges, which proposes some future work for researchers in this field. Instead of covering various methods of geometric deep learning in other reviews, this perspective focuses on the above three aspects, which can help new readers have a fast grasp of development trends in geometric deep learning.

Progress. The rapid development of neural networks has greatly promoted the development of geometric deep learning. Graph convolutional neural networks have been developing and making progress in recent years. The proposed model GCN [2] in 2016 has transformed frequency domain graph convolution into spatial domain graph convolution, which greatly improves the computational efficiency of graph convolution models. Subsequently, great achievements were made in graph convolution models. Many variants of graph convolutional neural networks have been proposed, such as GraphSAGE [3], which changes the GCN from two aspects. On the one hand, through the strategy of sampling neighbors, it transformed GCN from a full-batch training method to a node-center mini-batch training method, which makes the distributed training of large-scale graph data available. On the other hand, the method has been expanded into several new ways of aggregating neighbors, replacing GCN operation. Besides, the graph attention network (GAT) [4], which uses the attention mechanism to perform aggregation operations on neighbor nodes, realizes the adaptive distribution of different neighbor weights, thereby greatly improving the expressive ability of the graph neural networks. Graph data in real life is often heterogeneous. Taking the different relationships between nodes into consideration, some methods add a dimension of relationship based on the operation of aggregating neighbors, making the aggregation operation of nodes a double aggregation process, which extends graph convolutional neural network to the heterogeneous information network. Table 1 shows the predicted performance of four typical GNNs in four common datasets. With the initial residual connection and identity mapping, GCNII [5] makes the best performance in the four datasets.

Table 1 Performances in four datasets cited from [4, 5]

Method	Cora (%)	Citeseer (%)	Pubmed (%)	PPI (%)
GCN	81.4	70.9	79.0	–
GraphSAGE	–	–	–	76.8
GAT	83.0	72.5	79.0	97.3
GCNII	85.5	73.4	80.2	99.53

As for manifold deep learning, it has still been combined with the traditional neural network or machine learning algorithm. There are many proposed models for manifold deep learning which are voxel-based, multiview-based, point-based, and peometric algebra-based. Voxel data is a data representation form of 3D data. Each voxel in the 3D image corresponds to each pixel in the 2D image, and the voxel is the smallest data unit in the 3D space. The disadvantages of this method are high memory consumption and slow training. There are several ways to improve it. The main idea of the multi-view neural network [6] is to use multiple two-dimensional images to extract the surface features of the three-dimensional shape from different angles,

* Corresponding author (email: wxxie@szu.edu.cn)

and to segment the three-dimensional shape. Using two-dimensional CNN technology to process the corresponding two-dimensional images directly to identify and detect three-dimensional objects. The point cloud-based method directly processes the original 3D surface with 3D coordinates, and must not be affected by the number and arrangement of data.

Related applications. Geometric deep learning has been widely explored in various fields. The recommendation system uses social networks or buying behaviors to provide users with accurate and fast services, such as friends, merchandise items or personalized content for each user to improve their experience and create opportunities for attracting new users. Besides, community detection and fraud detection in social network are also popular in the field of graph neural networks. Community detection is a common task in social network analysis, with the purpose of clustering nodes into multiple groups called communities. Graph neural networks can achieve both overlapping and nonoverlapping community detection.

Except for the above, geometric deep learning is also applied in many other fields such as computer vision [7], chemistry, biology, and knowledge graph. There are also many interesting types of research in image analysis. Related methods are presented in Appendix A.

Challenges. Research on geometric deep learning, especially in graph data, has made great success in structural data in the past years. But there are still some problems and challenges which researchers should face.

(1) Computation cost. There are numerous parameters in the deep neural network that bring great computation cost to researchers. So does the geometric deep neural network. For one thing, nodes in graph data have a complicated connection and unpredictable attributes, which makes it more difficult to train the neural network model. The information in manifold entities has lots of high-dimensional information. Therefore, the data with a number of entities we need to train our model is quite difficult to process and train. For another thing, the current deep learning frameworks, such as Tensorflow and Pytorch, integrated many common models like CNN, LSTM, which are used to directly deal with Euclidean data like image and text, with the powerful calculating ability of modern GPUs to accelerate the model's training process. Geometric data is so irregular in most situations that different methods are needed for efficient and complex calculations. It has a long way to go for more powerful frameworks and abilities for calculating.

(2) Deeper architecture. As we know, it is a lot effective for most kinds of tasks to increase the number of neural network layers in the deep learning field. The traditional deep neural networks can stack hundreds of layers to obtain better performance, because the deep neural network has more structural parameters, which greatly improves its expression ability. However, graph neural networks are often shallow, most of which do not exceed three layers. Overlaying multiple GCN layers will result in over smoothing; that is, all nodes will converge to a similar value. Some methods are proposed to alleviate over smoothing and achieve a certain effect. The comparison of different deep GCNs is presented in Appendix B.

However, the proposed methods for deep architecture are effective just in some certain graph data. It has still been the biggest limitation of GNN. Designing a true deep GNN is an exciting challenge for future research and will make a huge contribution to understanding GNN.

(3) Dynamic graph. A big challenge is how to deal with dynamic graphs. Static graphs are stable, so feasibility modeling can be done on them, while dynamic graphs have changing structures. New users may join in, and the old user's exit and connections between users will change at any time. When edges and nodes appear or disappear, static GNN cannot change adaptively. Although the graph is a common way to model complex systems, this abstraction is usually too simple because the real-world system is dynamic and changes over time. Sometimes time-continuous behavior provides the key information about the system. How to process the dynamic graph effectively in various types of fields is still a big challenge for the researchers. Dynamic GNN is actively studied, but research about dynamics needs deeper study before it becomes mature. We think it is a milestone in overall GNN stability and adaptability.

(4) Heterogeneity. For graph data, the current methods aim to process homogeneous information networks, less to process heterogeneous information networks that are those with more than one type of nodes or more than one type of connection. For example, users on Twitter network can retweet, reply, and mention other users. Therefore, there are three or more user relationship types for a Twitter network, and the network data contains a large amount of multimedia information such as image and text. This is an example of a typical heterogeneous information network. In [8], there proposed a heterogeneous information network analysis to mine heterogeneous graphs with richer and more comprehensive structural information and attribute information. It still needs deeper researches.

(5) Scalability. In the real world, there are many so large graph networks that it is hard to research on with graph neural networks. Large graphs consist of local information and global information that are important for graph analysis. The current graph neural network focuses only on local information, neglecting global information which is important but hard to obtain. It is difficult to calculate the Laplacian matrix of a large graph for GCN model. Besides, every node in a large graph has its local structure, making it hard to train the model with mini-batch method. Some methods have been proposed to solve this problem, such as rapid sampling and subgraph training but they are not effective enough. Besides, high-order structures are also very important in complex networks, such as describing protein-protein interactions in biological applications. However, most graph neural networks are limited to nodes and edges. Incorporating such a structure into a message passing mechanism can bring more expressive power to models based on graph neural networks.

Scalability is one of the key factors limiting industrial applications, and industrial applications usually have to deal with very large graphs and low latency constraints. Until recently, the academic research community has almost ignored this aspect, and many models described in the literature are completely unsuitable for large-scale scenarios. Also, the perfect combination of graphics hardware (GPU) and classic deep learning architecture is one of the main driving forces for mutual success, but is not necessarily suitable to graph structure data. In the long run, specialized graphics hardware may be needed.

(6) Causal reasoning and interpretability. The Turing award winner Judea Pearl pointed out that graph neural network has the potential to deal with problems on causal reasoning that cannot be solved by a classical neural network. Geometric deep learning is the expansion of various

neural networking methods. With the powerful end-to-end learning ability and relational inductive reasoning, graph neural networks are expected to solve the relational reasoning problems that cannot be dealt with by classical neural networks, providing structured information and structured behaviors with a direct interface. Ref. [9] was the first to focus on relational reasoning with a general framework of GCNs and GNNs.

Since graphs are usually related to other disciplines, deep learning models that interpret graphs are essential to decision-making problems. For example, in medical or disease-related problems, interpretability is very important to transform computer experiments into clinical applications. Human cognition makes a strong assumption that the world is composed of objects and relationships. As graph networks make similar assumptions, their behavior is often easier to explain. However, the interpretability of graph-based deep learning is more challenging than other black-box models, because there is a high degree of interconnection between the nodes and edges in the graph. The entities and relationships usually correspond to things that humans understand, thereby supporting more interpretable analysis and visualization. Further exploration of interpretability is an interesting direction for future work on geometric deep learning.

Conclusion. Thanks to advances in powerful computing, model flexibility and training algorithms, as well as the rapid development of open datasets, deep learning techniques for geometric data have achieved great success in both theoretical research and practical applications. We give a general introduction to geometric deep learning methods. Methods are divided into two parts including graph-based and manifold-based. As is pointed out above, there are still many studies to do in the future as follows.

• Expanding modeling. Complex networks are diverse and not only topology may vary. It is a lot vital to design effective universal frameworks for various homogeneous and heterogeneous data.

• Scalability and dynamic graph. Data in the real world is usually large-scale and dynamic. How to model detailed dynamic time information in large-scale ground-truth data is a popular research direction. Moreover, it still has a long way to go.

• Adoption in related fields. Advances in geometric deep learning directly support the development of related applications based on geometric deep learning. Besides, it has been a trend to apply graph learning to related researches with connection information.

In summary, geometric deep learning provides a new framework in which non-Euclidean data can be effectively addressed in order to solve problems in various related fields. From a long-term perspective, geometric deep learning will set off a big wave in the field of artificial intelligence.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 61771322, 61871186, 61971290) and Fundamental Research Foundation of Shenzhen (Grant No. JCYJ20190808160815125).

Supporting information Appendixes A and B. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Lin Q, Wenming C, He Z, et al. Mask cross-modal hashing networks. IEEE Trans Multimedia, 2020, 23: 550–558

2 Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks. 2016. ArXiv:1609.02907

3 Hamilton W L, Ying R, Leskovec J. Inductive representation learning on large graphs. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017. 1025–1035

4 Veličković P, Cucurull G, Casanova A, et al. Graph attention networks. 2017. ArXiv:1710.10903

5 Chen M, Wei Z, Huang Z, et al. Simple and deep graph convolutional networks. 2020. ArXiv:2007.02133

6 Su H, Maji S, Kalogerakis E, et al. Multi-view convolutional neural networks for 3D shape recognition. In: Proceedings of the IEEE International Conference on Computer Vision, 2015. 945–953

7 Wang R, Shen M M, Wang X Y, et al. RGA-CNNs: convolutional neural networks based on reduced geometric algebra. Sci China Inf Sci, 2021, 64: 129101

8 Shi C, Li Y, Zhang J, et al. A survey of heterogeneous information network analysis. IEEE Trans Knowl Data Eng, 2017, 29: 17–37

9 Battaglia P W, Hamrick J B, Bapst V, et al. Relational inductive biases, deep learning, and graph networks. 2018. ArXiv:1806.01261