Briefings in Bioinformatics, 23(1), 2022, 1–16

https://doi.org/10.1093/bib/bbab340
Review

Graph representation learning in bioinformatics: trends, methods and applications

Hai-Cheng Yi, Zhu-Hong You, De-Shuang Huang and Chee Keong Kwoh

Corresponding author: Zhu-Hong You, School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China. Tel.: +8618160622862; E-mail: zhuhongyou@nwpu.edu.cn

Abstract

Graph is a natural data structure for describing complex systems, which contains a set of objects and relationships. Ubiquitous real-life biomedical problems can be modeled as graph analytics tasks. Machine learning, especially deep learning, succeeds in vast bioinformatics scenarios with data represented in Euclidean domain. However, rich relational information between biological elements is retained in the non-Euclidean biomedical graphs, which is not learning friendly to classic machine learning methods. Graph representation learning aims to embed graph into a low-dimensional space while preserving graph topology and node properties. It bridges biomedical graphs and modern machine learning methods and has recently raised widespread interest in both machine learning and bioinformatics communities. In this work, we summarize the advances of graph representation learning and its representative applications in bioinformatics. To provide a comprehensive and structured analysis and perspective, we first categorize and analyze both graph embedding methods (homogeneous graph embedding, heterogeneous graph embedding, attribute graph embedding) and graph neural networks. Furthermore, we summarize their representative applications from molecular level to genomics, pharmaceutical and healthcare systems level. Moreover, we provide open resource platforms and libraries for implementing these graph representation learning methods and discuss the challenges and opportunities of graph representation learning in bioinformatics. This work provides a comprehensive survey of emerging graph representation learning algorithms and their applications in bioinformatics. It is anticipated that it could bring valuable insights for researchers to contribute their knowledge to graph representation learning and future-oriented bioinformatics studies.

Key words: graph representation learning; deep learning; graph neural network; graph embedding; knowledge graph; healthcare

Introduction

Graph is a natural data structure that contains a set of objects and a collection of pairwise relationships between objects. It is a universal language for describing and modeling ubiquitous real-life complex systems, such as social network [1], academic citation network [2] and word co-occurrence networks [3]. From molecular structure to healthcare systems, biomedical graphs are pervasive in the field of biomedicine and life sciences, for instance, gene regulatory networks [4], protein–protein interaction (PPI) network [5], human brain connectomes [6] and biomedical knowledge graphs. Graphs are increasingly

Hai-Cheng Yi is a PhD candidate at the University of Chinese Academy of Sciences and the Xinjiang Technical Institute of Physics and Chemistry, Chinese Academy of Sciences. He works on representation learning, network analysis and bioinformatics.
Zhu-Hong You is a professor at Northwestern Polytechnical University. His research interests include big data analysis, data mining, pattern recognition and bioinformatics.
De-Shuang Huang is the director of the Institute of Machines Learning and Systems Biology, Tongji University. He is an IAPR Fellow, IEEE Fellow.
Chee Keong Kwoh received PhD Degree from Imperial College of Science, Technology and Medicine, University of London, in 1995. He has been with the School of Computer Engineering, Nanyang Technological University (NTU), since 1993. He is the Program Director of the MSc in Bioinformatics program at NTU. His research interests include data mining, soft computing and graph-based inference, and applications areas include bioinformatics and biomedical engineering.
Submitted: 24 June 2021; **Received (in revised form):** 18 July 2021

becoming the major paradigm for modeling, learning and reasoning in biomedical systems.

The rapid development of molecular biology, physiology and omics has promoted the understanding of how biomolecules, cells, organs cooperate to carry out the important biochemical or physiological activities. Representing biological components as nodes and interactions between nodes as edges, complicated biological systems can be modeled naturally as graphs. This austere concept is gradually being accepted and expanded by researchers. The trend of modeling and analyzing biological graphs to solve biological problems can be categorized into three phases: bipartite biological graphs, multirelational biological graphs and multimodal biomedical knowledge graphs. We present brief introductions of them as follow: (a) bipartite biological graph. It contains two kind of biological objects and the associations between them [7]. It has already applied to a great deal of important biological tasks [8], such as protein functions annotations based on PPIs graph [9, 10], inferring new indications of drug from drug–target interactions graph [11–13], miRNA–disease associations prediction [14, 15], lncRNA–disease associations prediction [16] and circRNA–disease associations detection [17]; (b) multirelational biological graph. This is a more complicate multilayer heterogeneous network for describing complex synergy between multiple biological elements. Emerging studies confirm that there is mutual regulation and competition between molecules, i.e. competing endogenous RNA hypothesis [18]. And for drug discovery and disease treatment, it is necessary to comprehensively examine drug–target, drug–disease, drug–gene, disease–gene and drug–drug interactions. These complex systems can be great formed as heterogeneous multirelational biological graphs, e.g. lncRNA–mRNA–miRNA graph [19], drug–target–disease graph [13, 20–22], miRNA–gene–disease tripartite graph [23], chemical–gene–disease graph [24] and miRNA–gene–lncRNA–disease graph for miRNA–disease associations prediction [25]; (c) biomedical knowledge graph. A knowledge graph has many names in history, e.g. semantic networks, knowledge base or ontology. It mines 'knowledge' from a large volume of information in a wealth of scattered documents and databases and links these entity relationships as graph. Each piece of knowledge is represented as a Subject–Predicate–Object triplet [26]. Knowledge graph is regarded as an fundamental infrastructure for the next generation of artificial intelligence and has many cutting-edge applications in the field of bioinformatics, including healthcare knowledge graph for clinical decision support [27], comprehensive molecular associations graph [28] and biomedical knowledge graphs (e.g. PharmGKB [29], DrugBank [30], Gene Ontology (GO) [31], Disease Ontology [32] and KEGG [33]) for disease treatment.

To transform rapidly accumulated biomedical big data into valuable knowledge, machine learning, especially deep learning, succeeds in broad scenario of bioinformatics, such as sequence analysis, structure prediction, and biomedical image processing and diagnosis. The data in these tasks are directly represented in Euclidean domain, e.g. sequences (1-D), biomedical images (2-D) and structure (3-D). Deep learning models are designed to handle these regular Euclidean data, which has been well reviewed by previous works [34–36]. However, there are clear challenges between non-Euclidean biological graphs and typical deep learning models. For instance, the nodes in graphs have diverse connections, arbitrary neighbor size, complex topological structure and no fixed node ordering. To address these need, graph representation learning bridges rich valuable biological graphs and advanced machine learning techniques, including shallow graph embedding methods and emerging graph neural networks (GNNs). The main paradigms of graph representation learning are shown in Figure 1. Among them, graph embedding aims to learn low-dimensional representations of nodes, links or subgraphs while maximally preserving graph topology and inherent attributes that are fit for off-the-shelf machine learning methods for downstream graph analytics tasks, such as node classification, link prediction, community detection and visualization. However, GNNs can not only learn embeddings that retain graph topology and node attribute through a series of message aggregation and propagation, but also directly complete the tasks on graphs end-to-end (as shown in Figure 2). As suggested by previous research of graph embedding techniques [37–41], graph embedding approaches can be classified into homogeneous graph embedding, heterogeneous graph embedding and attributed graph embedding methods. Although the GNNs can be summarized into graph recurrent networks (GRNs), graph convolutional networks (GCNs), graph autoencoders (GAEs) and graph generative adversarial networks (GGANs) based on their model architectures and training strategies [42–45].

Although there have been some beneficial attempts to apply graph representation learning to improve biological graph analysis tasks, some valuable previous works have reviewed the individual types of graph representation learning methods and the traditional applications in several scenarios. For example, Barabasi *et al*. first reviewed many network-based methods that can accurately discover biomarkers and effectively identify targets for drug development and outlined the network medicine of disease modules, pathways and molecular relations with distinct phenotypes [46, 47]. The contribution of Barabasi's series of work is not only in methodology (the method is outdated from the current perspective), but in attracting and disseminating network-based systematic views. Furthermore, Pavlopoulos *et al*. [48] provided a comprehensive survey of bipartite graph and related applications in network biology and medicine. Su *et al*. [49] introduced the works of applying graph embedding to accelerate the downstream biomedical tasks. Ding *et al*. [50] presented an overview of heterogeneous network and its application in drug development and human interactome. Yue *et al*. [51] conducted a comparison of typical graph embedding methods for link prediction (drug–drug, protein–protein, drug–disease interactions prediction) and node classification (medical term classification and protein function annotation) tasks on bipartite biological graphs. Muzio *et al*.'s work [52] briefly summarized deep learning on biological networks, specifically, the application of GNNs in proteomics, drug development, metabolic and gene regulatory networks. And Nelson *et al*.'s [53] work mainly focuses on the traditional network embedding methods and their applications in network biology. However, due to the complexity and diversity of biomedical graphs and the rapid development of graph representation learning, these works are not comprehensive and systematic. There is an urgent need to summarize the types of biological graphs and emerging graph representation learning algorithms systematically and clearly.

In this work, we provide a comprehensive review of graph representation learning and its brilliant applications in bioinformatics. We first conducted a detailed summary and discussion of the shallow graph embedding algorithms and emerging GNNs in graph representation learning. Then, the representative applications of graph representation learning in board bioinformatics problems are introduced. Moreover, we discuss the challenges and opportunities of graph representation learning methods in bioinformatics. In addition, we summarize the open resources platforms and libraries for graph computing and graph representation learning and provide the implementation of the

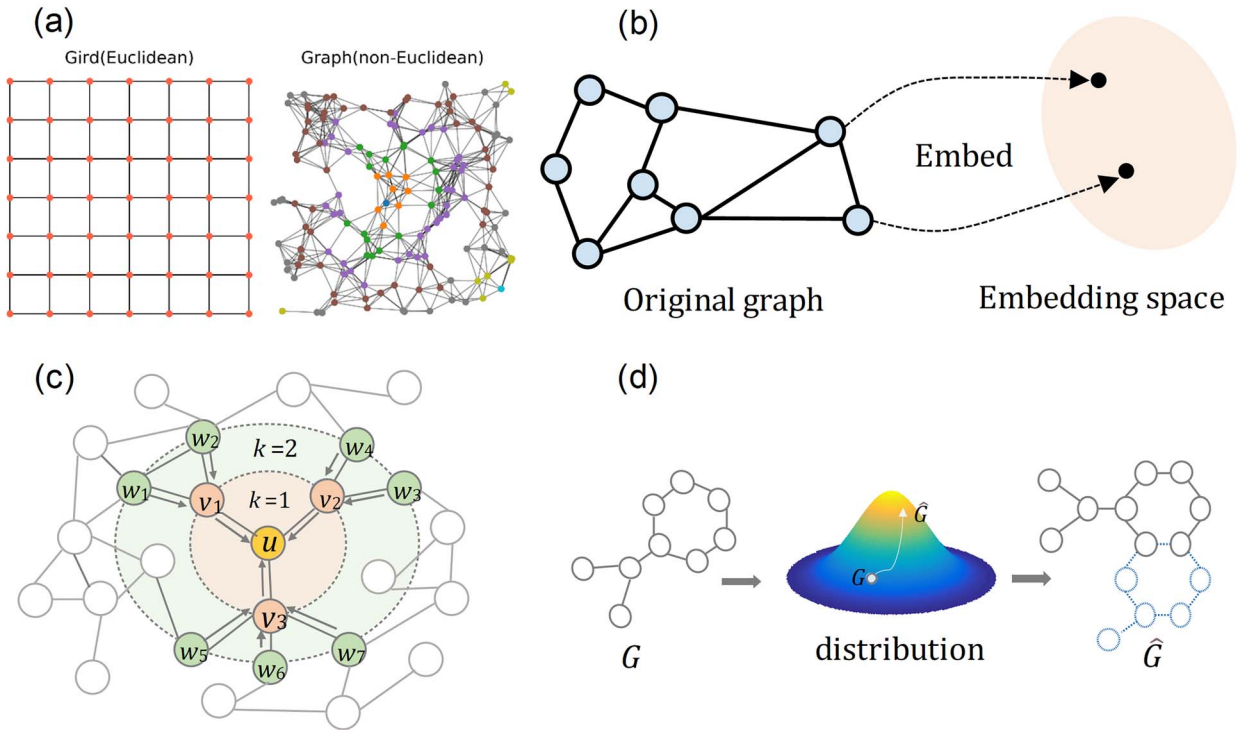

Figure 1. Graph representation learning paradigms. Graph representation learning bridges non-Euclidean graph data and modern machine learning techniques. Here, we summarize both graph embedding methods and graph neural networks. **(A)** An illustration of gird-like Euclidean data versus non-Euclidean graphs (not learning friendly). **(B)** Graph embedding methods generate node representations through directly projecting nodes in original graph to low-dimensional representation spaces. **(C)** Graph neural networks learn graph representations through diverse message aggregation and propagation. **(D)** Graph generative models learn the distribution of input samples to generate molecular graph with desired properties.

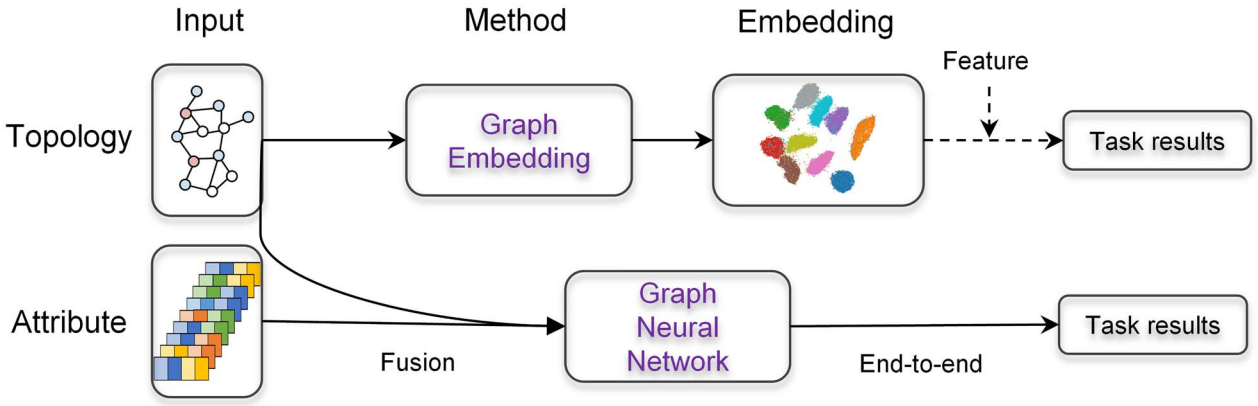

Figure 2. Comparison of graph embedding methods and graph neural networks. Graph embedding methods generate node representations that can be combined with machine learning models to preform downstream tasks, whereas graph neural networks fuse graph topology and attributes to perform end-to-end graph tasks.

graph embedding and GNN models reviewed in this work. This work provides comprehensive survey of emerging graph representation learning and its applications in bioinformatics, which aims to serve as a useful guide for researchers to apply graph representation learning approaches in bioinformatics studies.

Graph representation learning: a brief overview

In this section, we provide a brief overview of graph representation learning methods. Graph representation learning aims to encode the nodes in the graph into low-dimensional vector representations, which maximize the preservation of graph topology and node attribute information. The notations and definitions regrading graphs and proximities are described first. And then we outline the key types of both shallow graph embedding methods (including homogeneous graph embedding, heterogeneous embedding and attributed graph embedding) and GNNs. The hierarchical relationship between the different methods reviewed in this section is summarized in Figure 3.

Basic definitions

Many real-world systems can be abstractly represented as different levels of information graphs, which focus on the components and the associations among these components. Graph representation learning method aims to solve the generalized graph

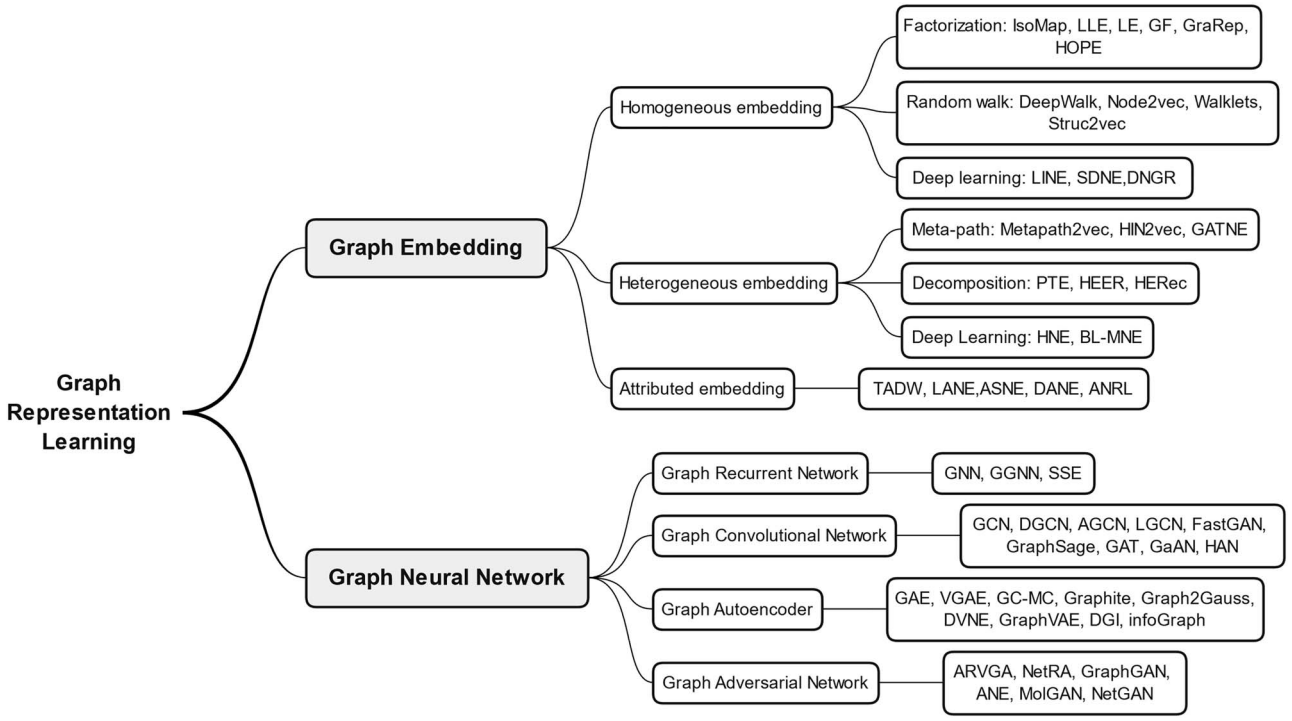

Figure 3. Taxonomy and representative methods of graph representation learning.

embedding problem. In this section, we first define important concepts related to graph representation learning, including different type of graphs, and graph proximities that different graph embedding algorithms rely on. For ease of presentation and analysis, formal definitions of the notations are first introduced.

Suppose $G = (V, E)$ indicates a graph, consists of a set of vertices (a.k.a. nodes) $V = \{v_1, v_2, \ldots, v_{|V|}\}$, where $|V|$ represents the number of vertices and a set of links (a.k.a. edges) $E = \{e_{i,j}\} \in \mathbb{R}^{V \times V}$. The adjacency matrix W of graph G retains nonnegative weights associated with each edge, if v_i is linked with v_j, $w_{i,j} > 0$, if there is no link between v_i and v_j, $w_{i,j} = 0$. For undirected graph, the adjacency matrix is symmetric, $w_{i,j} = w_{j,i}, \forall i, j \in [V]$. There is also a node type mapping function $\phi : V \to T$, and a link type mapping function $\psi : E \to R$. T and R are the sets of predefined node types and link types.

Definition 1. Homogenous and Heterogeneous graph. Given an information graph G, based on its graph topology and attributed property (with or without node attributes), it can be categorized into different type of graph. If the node types $|T| > 1$ or link types $|R| > 1$, that is $|T| + |R| > 2$, the graph is heterogeneous graph. Otherwise, it is a homogenous graph ($|T| = 1$ and $|R| = 1$) [54–56]. The homogenous graph has only one type of node and a unique link type, whereas the heterogeneous graph contains multityped, interconnected objects, such as drug–target–disease graph. At the same time, multiplex graph is a special type of heterogeneous graph. Multiplex graph is also known as multiview graph [57, 58] or multidimensional graph [59, 60], which has only one type of node but multiple types of edges. It can be treated as a special type of heterogeneous graph with $|T| = 1$ and $|R| > 1$ [61, 62].

Definition 2. Attributed graph. Abstract vertices in an information graph usually have their inherent properties. An attributed graph can be formally defined as $G = (V, E, A)$, where A denotes an attribute representation matrix. For each node $v_i \in V$, there is a feature vector $a_i \in A$ is affiliated with it, where $A = \{a_i | v_i \in V\}$ is the set of node attribute features for all nodes. a_i is the ith row of the attribute matrix that belongs to node v_i [41, 63].

Definition 3. Meta-path. For heterogeneous graph, a meta-path $P = T_1 \xrightarrow{R_1} T_2 \xrightarrow{R_2} T_3 \to \ldots \xrightarrow{R_l} T_{l+1}$ is defined on a network schema $\tau(G) = (T, R)$, which consists of a composite relation $R = R_1 \circ R_2 \circ R_3 \circ \cdots \circ R_l$ between node types T_1 and T_{l+1}, where l denotes the length of path ($l \geq 1$), and $\circ$ indicates the composition operator on relations [64]. Meta-path can handle semantic information effectively, for example a path $\text{Drug}_a \xrightarrow{\text{target}} \text{protein}_b \xrightarrow{\text{interact}} \text{disease}_c$ notes a treatment mechanism of a disease in a biomedical graph.

Definition 4. First-order proximity. The first-order proximity captures the local pairwise similarity between two direct neighbor nodes. If there is a link between two vertices, these two nodes are similar, otherwise, they are dissimilar. Formally, the first-order proximity of two nodes v_m and v_n is measured by $S_{m,n}$. If the node pair $(v_m, v_n) \in E$, $S_{m,n} > 0$; if $(v_m, v_n) \notin E$, $S_{m,n} = 0$ [56, 58].

Definition 5. High-order proximity. The high-order proximity captured the k-hops ($k \geq 2$) neighborships between nodes. And the second-order proximity is a special case of high-order proximity ($k = 2$), which is determined by the number of neighbor nodes connected by intermediate nodes. The high-order proximity of two nodes v_m and v_n is measured by the k-hops transition probability from v_m to v_n, that is $S_{m,n} = \hat{E} + \hat{E}^2 + \hat{E}^3 + \cdots + \hat{E}^k$, where $\hat{E}$ denotes the first-hop transition probability. The high-order proximity captured the global proximity [65].

Definition 6. Semantic proximity. The semantic proximity of two node v_m and v_n is obtained by the similarity of their attribute feature vector a_m and a_n [66, 67]. And commonly used similarity measures include Cosine similarity, Pearson correlation

coefficient, Jaccard similarity coefficient and Gaussian interaction profile (GIP) kernel similarity.

Homogeneous graph embedding

The first category of graph embeddings is homogeneous graph embedding, which is also known as network embedding or nonattributed graph embedding. And it is the plainest graph representation learning method, which was first developed. Homogeneous graph embedding methods generally aim to preserve graph topology when learning low-dimensional representations of vertices. Based on their technical details, we categorize these homogeneous graph embedding methods into three major categories: matrix factorization-based methods, random walk-based methods and deep learning-based methods.

Matrix factorization-based methods

Matrix factorization aims to factorize a matrix into low-dimensional matrix while still maintaining the latent manifold structure and topological properties in the original matrix. The pioneer efforts, such as IsoMap [68], Locally Linear Embedding [69], Laplacian Eigenmaps [70] and graph factorization [71], represent the relationships between nodes as graph adjacency matrix, Laplacian matrix or similarity matrix, and then adopt matrix factorization to obtain the embeddings. The difference between these methods is that they are based on different first-order matrices that capture the structure of graph, and they usually obtain a shallow embedding of nodes.

More recently, matrix factorization-based graph embedding methods that can preserve high-order proximity also have been developed. For example, the GraRep [65] and HOPE [72] consider the high-order proximity by factorizing k-hop transition probability matrices and similarity matrix (measured by Kate index, Rooted Page Rank, Common Neighbors and Adamic–Adar score), respectively. Although its effectiveness has been proven, scalability is still a key bottleneck for matrix factorization-based methods because of the huge memory overhead and extremely high computational cost.

Random walk-based methods

Inspired by word2vec [73], researchers extended the embedding method in natural language processing from word sequences to graph node sequences, i.e. paths. Random walk is applied to generate node sequences in graph to capture structural relationships of nodes. Specifically, for a given graph and a starting node, Random walk randomly selects its neighbor node and then moves to this node. By repeating this process, a graph is transformed into node sequences. Then, probability model like skip-gram model can be employed to learn the node embedding based on the generated node sequences, which preserve structural proximity of the graph. The initial work of this category is Deepwalk [74]. Similarly, node2vec [75] improves a flexible biased random walk, smoothly combining breadth-first sampling and depth-first sampling to obtain node sequences. Therefore, both local and global proximities are preserved. Furthermore, two variants Walklets [76] and stuc2vec [77] are proposed by biasing and modifying the random walks. Walklets modified the random walk strategy by skipping over some nodes. Rather than neighborhood node information, struc2vec define random walks based on structural similarity of nodes.

Deep learning-based methods

The impressive representation learning capabilities of deep learning techniques are also demonstrated in the field of graph embedding learning. The widely used embedding method LINE [58], which can be regard as using a multilayer perceptron to approximate the first-order proximity and second-order proximity to learn node embedding. Based on deep autoencoder architecture, SDNE [78] preserves both global and local graph structure by modeling both first-order proximity (measured based on Laplacian Eigenmaps) and second-order proximity of nodes. To capture higher order proximity, DNGR [79] learns deep low-dimensional node embedding by applying stacking denoising autoencoders on the positive pointwise mutual information matrix. Deep learning-based embedding methods can learn nonlinearity in graphs, but their calculation cost is generally high. And most importantly, these deep learning-based methods are still very primitive. They can only generate embeddings for nodes that have appeared in training phase. To alleviate these address, more sophisticated GNNs have been rapidly developed recently. And we will discuss them specifically in the 'Graph neural networks' section.

Heterogeneous graph embedding

Heterogeneous graphs are more naturally related to real-world scenarios with multitype of objects and associations, whereas the homogeneous graph embedding methods mentioned above cannot be directly work on them. Heterogeneous graph embedding emerged few years ago and rapidly become a booming research domain. Existing heterogeneous graph embedding methods can also be roughly divided into three types: meta-path-based, decomposition-based and deep learning-based methods.

Due to the heterogeneity of structure and contents, the random walk is difficult to find an effective walking strategy to capture the rich semantics contained in the entire graph. The meta-path restricts the direction of random walk to reduce the complexity of traversal in the heterogeneous graph. Metapath2vec [61] formalizes the meta-path-guided random walk to generate heterogeneous neighborhoods of a node and then leverages a heterogeneous skip-gram model to conduct node embeddings. As an extension, HIN2vec [62] also uses meta-path-based random walk and proposed a neural network to capture the graph heterogeneity. The differences are that it uses the generated meta-paths as objects to directly learn the representation of both meta-paths and nodes. Later on, GATNE [41] extended this strategy to multiplex heterogeneous networks. However, the appropriate settings of number of walks and walk length are crucial for this type of model. Otherwise, these methods cannot completely preserve the whole structure of a graph.

To ease the complexity of heterogeneous graph, another strategy for solving heterogeneous graph embedding is like divide and conquer, which divides an input heterogeneous graph into several small homogeneous or bipartite graphs. For instance, PTE [80] decomposes a heterogeneous graph into multiple bipartite graphs according to edge types, and then utilizes LINE on each bipartite graph to learn the shared node embeddings. HEER [81] extended the PTE by considering the typed closeness of node pair atop their edge embedding. According to defined meta-paths, the HERec [82] also projects a heterogeneous graph into different dimensions, and then employed the metapath2vec within each subgraph to learn vertex embeddings.

Despite the importance of this problem, few efforts have been made on heterogeneous graph embedding based on deep learning techniques. For example, HNE [83] learns representation for topological information of graphs and contents through a

deep neural network and a convolutional neural network, and then jointly projects them into a common embedding space. And the BL-MNE [84] model introduced a deep aligned autoencoder-based embedding method for heterogeneous graph. Recently, diverse GNNs have also been applied to heterogeneous graph embedding. To better introduce them below, they are grouped together and sorted out.

Attributed graph embedding

In addition to graph topology information used by most of graph embedding algorithms, some works have demonstrated that the abundant node attributes should also be fully exploited to assist graph representation learning. Attribute graph embedding can also have an intersection with the homogeneous or heterogeneous graph embedding, for example attributed homogenous graph embedding or attributed heterogeneous graph embedding [41]. The TADW [67] first adopted an inductive matrix factorization to fuse node textual attributes and graph structure. LANE [85] smoothly applies graph Laplacian technique combined with label information to jointly learn embedding from node attributes and topological structure. ASNE [86] develops a social graphs embedding framework by integrating both node proximity and structural proximity. DANE [87] captures the high nonlinearity of various topological structure and node attributes proximities. Liu *et al*. proposes a unified attributed heterogeneous graph embedding model AHNG [88] by fusing graph structure, semantic information and attributes with a Gaussian distribution. ANRL [89] designed a neighbor enhancement autoencoder model to merge node attributes affinity and structural proximity into low-dimensional embedding spaces.

GNNs

Deep learning has gained substantial progress in various fields. However, original deep learning methods are designed for Euclidean data, such as texts, sequences, images. To apply deep learning on non-Euclidean graphs, GNNs have recently been rapidly developed and widely applied. In addition to learning graph embedding, GNNs can usually capture both graph structural information and node properties through a variety of local message aggregation and propagation steps and can directly address graph-related tasks in an end-to-end manner. Suggested by previous work [42, 43], in this section, we summarize the latest GNNs into four categories based on model architectures and training strategies: GRNs, GCNs, GAEs and GGANs. And some variants can combine and overlap these architectures based on different training strategies, such as reinforcement learning and contrastive learning.

GRNs

Graph recurrent neural networks are the pioneer works of GNNs, which encode high-dimensional node representations by applying the same parameters recurrently over nodes in a graph. Based on an information diffusion mechanism, Gori *et al*. presented the GNN [90, 91], which modified the original recurrent neural network model to process graph data, where node aggregates their neighbors information until a stable equilibrium is reached. As a notable improvement, Li *et al*. developed the gated GNNs [92] by employing a gated recurrent unit [93] as a mapping function with shared parameters and adopted the back-propagation through time to train it. Stochastic Steady-state Embedding [94] is also presented to improve the scalability of GNN model, which updates hidden states of node recurrently and is more effective for large-scale graphs.

GCNs

GCNs extend the convolution operation from structured data such as images to graph data. The main idea is to learn a mapping function $f(\cdot)$ to produce a node's embedding by aggregating its own feature and its neighbor's features. Graph convolutions can be divided into spectral and spatial method. Kipf *et al*. introduce the first GCN [95], which applied a renormalization trick to address the gradient exploding or vanishing problem. Then, Zhuang *et al*. present the dual GNN [96], which jointly consider the local and global consistency on graphs with two convolutional layers, and the adjacency matrix is replaced by positive pointwise mutual information matrix. However, these methods require storing the entire adjacency or Laplacian matrix in memory, which will result in expensive calculations. They are also a lot of variants of GCNs (for instance, the AGCN [97], LGCN [98], FastGCN [99]). An important improvement of GCN is GraphSAGE [100]. It provides a general inductive learning framework that can generate embedding for unseen nodes by sampling and aggregating local neighbor's features. Attention mechanisms [101] also can be employed to improve GCNs, graph attention network (GAT) [102] introduces self-attention into propagation step, and multihead attention is further considered to enhance the capacity and stability of model. Gated attention networks [103] improves the multihead attention mechanism by learning different weights for different attention heads. HAN [55] learns node embedding for heterogeneous graph by performing a hierarchical attention on both node level and semantic level, while the node-level attention is used to learn different weights to aggregating meta-path-based neighbors, and the semantic-level attention lean the importance of different meta-paths.

GAEs

The encoder–decoder architectures have also been widely applied in both graph embedding and graph generation tasks. GAE [104] first extends this architecture into graph embedding, which employs GCNs as the encoder to encode the structural and node feature information and uses the decoder to reconstruct the adjacency matrix. And they also proposed a variational graph autoencoder (VGAE) by training the GAE in a variational manner [104]. By adopting GCN as the encoder and using a simple bilinear function as the decoder, GC-MC is proposed and demonstrated on recommendation tasks [105]. Furthermore, iterative generative modeling of graphs (Graphite) [106] extends them by designing a more complicated decoder, which iterates between paired graph convolutions and decoding functions. In addition to encode nodes into low-dimensional embedding, Graph2Gauss (G2G) [107] captures the uncertainties of nodes by learning each node's Gaussian distribution. Inspired by SDNE and G2G, DVNE [108] also represents each node as Gaussian distribution using a variational autoencoder (VAE), and the Wasserstein distance is adopted to remain the transitivity of the node similarities. Assuming a Gaussian prior distribution, GraphVAE [109] uses a GCN as the encoder and a simple multilayer perception as the decoder for graph generation tasks. Based on GAEs, contrastive learning is another way for unsupervised graph embedding, which is first introduced in Deep Graph Infomax (DGI) [110]. DGI captures graph global topological information by maximizing mutual information between node embedding and graph representation. Similarly, InfoGraph [111] learns graph representations by maximizing mutual information between graph-level representation and subgraph-level representations of different scales.

GGANs

The generative adversarial networks (GANs) [112] can also be extended to graph domains. The main idea behind GAN is adversarial training. The generator aims to generate fake samples to fool the discriminator, whereas the discriminator is designed to correctly discriminate between real and generated samples. In the end, both models will benefit from the joint training of this minimax game. Adversarially regularized graph autoencoder [113] adopts the adversarial training principle to regularize GCN-based GAE for learning robust node embeddings. Network representations with adversarially regularized autoencoders (NetRA) [114] presents a encoder–decoder framework to learn network representations, whereas the input of NetRA is random walks rooted each node, and the learned embedding is regularized within a prior distribution through adversarial training. GraphGAN [115] enhances the inference abilities of node graph embedding by using a GAN. By employing the adversarial training strategy, adversarial network embedding [116] enhances existing graph embedding methods, such as Deepwalk, by taking a prior distribution as real data and treating the embedding vectors as the generated samples, where a GAN is adopted as an additional regularization term. Meanwhile, some brilliant works demonstrated the adversarial training can also improve the generalization capability of GNNs. To address the molecular graph generation problem, molecular GAN [117] combines GCNs, GANs and reinforcement learning scheme to generate molecular graphs with desired properties. NetGAN [118] treats the graph generation task as to learn the biased random walk's distribution. The generator produces conceivable random walks using an LSTM network, and the discriminator try to determine the fake ones.

In addition, some recent studies on pretraining and acceleration of GNN and very deep GNN models are worth noting, e.g. GPT-GNN [119], Graph-Bert [120], RevGNN-Deep (deepest GNN with >1000 layers) [121] and Graph-MLP [122], a new framework for graph learning without message passing.

Open resources

To facilitate researchers to implement and develop graph representation learning algorithms, we first provide graph representation learning and graph computing platforms and libraries in Table 1. These platforms and libraries assist researchers to quickly benchmark graph representation learning algorithms and develop their own models. And we also summarize the implementations of graph embedding and GNNs reviewed in the paper in Table 2, most of which are official implementations.

Applications in bioinformatics

From molecular level to healthcare level, graph is widely applied to represent and model multimodal biological and medical systems. An illustration of biomedical graphs at different scales is shown in Figure 4. Although several enlightened works have applied graph representation learning techniques in biomedical tasks, e.g. molecular generation, drug repurposing, interaction predication, the use of graph representation learning in biomedical tasks has not been thoroughly explored. In this section, we introduce a series of representative applications of graph representation learning in molecular graph analysis, multi-omics graph analysis, pharmaceutical and healthcare graph analysis.

Graph representation learning for molecules

The structure of molecules such as proteins and chemical compounds can be regarded as molecular graphs composed of atoms and bonds. The nodes are atoms or amino acids, and the edges are chemical bonds or peptide bonds. Graph representation learning for molecules aims to generate novel molecules efficiently and automatically with optimized properties.

Molecule representation learning

Learning efficient representations of molecules plays a fundamental role in many downstream tasks, such as protein function prediction, molecule property prediction and drug discovery. In addition to string-based representations, graph representation learning provides more flexible and better representations of molecules that are optimal for special tasks. For example, Duvenaud *et al.* [123] proposed an end-to-end framework to learn a differentiable molecular fingerprints by using GNNs. Based on geometric deep learning, Gainza *et al.* [124] presented molecular surface interaction fingerprinting to capture fingerprints that are optimized for specific biomolecular interactions under the hypothesis that proteins involved in similar interactions may share common fingerprints. Recently, Li *et al.* [125] proposed the 3DMol-Net to learn molecule representation using adaptive GCN and considered both the topology and rotation invariance of the 3D molecular structure.

Molecule properties prediction

Accurate prediction of molecular properties is crucial for compound design and drug discovery. Gilmer *et al.* [126] proposed a unified framework message passing neural networks (MPNNs) and demonstrated superior performance on molecular properties prediction benchmark. To remain the spatial connection information on molecules, a convolution spatial graph embedding layer (C-SGEL) is introduced by Wang *et al.* [127] to predict molecular property utilizing molecular graph data. Multilayer C-SGEL is integrated to form a convolution spatial graph embedding model and molecular fingerprints are fused to predict molecular properties. And Wieder *et al.* [128] provided a compact review of GNNs with different architectures for predicting molecular properties.

Molecule graph generation

To design or generate molecules with desired properties is a challenge problem with applications in drug discovery and development. Existing graph generative models aim to model the joint distribution directly. Jin *et al.* [129] presented a VAEs-based method to generate a junction tree-structured chemical substructure and then combined them into a graph using a graph massage passing network. Shi *et al.* [130] proposed a flow-based autoregressive model for generating molecular graph. They formulate graph generation as a sequential decision process, generate a new atom in each step, and then determine the bonds between generated atoms and existing atoms. Zang *et al.* [131] proposed an invertible flow-based generative model for molecular graph generation and achieved the state of the art performance in molecular graph generation and reconstruction, property optimization, etc. Recently, Mahmood *et al.* [132] developed a masked graph model for molecule generation, which learns the conditional distribution of masked graph components, given the rest of graph, using simple MPNN GNNs.

Graph representation learning for multi-omics

The integrated analysis of multi-omics data has become a new direction for exploring life mechanisms. Graph representation learning is a valuable tool to accelerate relational multi-omics data analysis, including genomics, proteomics and transcriptomics.

Table 1. Platforms and libraries for graph computing and graph representation learning

Package	URL	Framework	Contributor	Models included/Overview
OpenNE	https://github.com/thunlp/OpenNE/tree/pytorch	TensorFlow/PyTorch	Tsinghua University	LLE, GF, GraRep, Deepwalk, node2vec, LINE, TADW, SDNE, HOPE, GCN, GAE, VGAE
CogDL [161]	https://github.com/THUDM/cogdl	PyTorch/TensorFlow	Tsinghua University	Extensive task-oriented graph embedding methods and graph neural networks
PyTorch Geometric (PyG) [162]	https://github.com/rusty1s/pytorch_geometric	PyTorch	Dortmund University of Technology	Geometric deep learning extension library for PyTorch. A variety of methods for deep learning on graphs from published papers
Deep Graph Library (DGL) [163]	https://github.com/dmlc/dgl	PyTorch/MXNet/TensorFlow	New York University and Amazon Web Services (AWS)	Easy-to-use, high performance and scalable Python package for deep learning on graphs
Dive into Graphs (DIG) [164]	https://github.com/divelab/DIG	PyTorch/ PyTorch Geometric/RDKit	Texas A&M University	Unified testbed for higher level, research-oriented graph deep learning tasks, such as graph generation, self-supervised learning, explainability and 3D graphs
Graphvite [165]	https://github.com/deepgraphlearning/graphvite	Linux/Python	Mila-Quebec AI Institute	General graph embedding engine, dedicated to high-speed and large-scale embedding learning in various applications.
Graph-Learn [166]	https://github.com/alibaba/graph-learn	Python	Alibaba	Industrial graph neural network framework
Paddle Graph Learning (PGL)	https://github.com/PaddlePaddle/PGL	PaddlePaddle/Cython	Baidu	Efficient and flexible graph learning framework based on PaddlePaddle

Genomics graph analysis

Li *et al.* [133] proposed a single-cell representation learning method based on LINE to learn meaningful representations for single-cell high-throughput RNA sequencing (scRNA-Seq) data by considering gene–gene associations from gene expression data and pathway priors. Li *et al.* merged a variety of genomic and phenotype graphs into a heterogenous multigraph and developed a random walk-based method for disease gene identification [134]. GCN-MF [135] combined the GCN and matrix factorization to discover gene–disease associations. By using a subset of gene expression matrices, Yang *et al.* [136] proposed a unified model of graph variational generative adversarial nets (CONDGEN) integrated GCN, VAE and GAN framework for graph generation. Rhee *et al.* [137] combined the gene expression data into PPI graph and uses it as the GCN's input, and they defined a relation network to given priority to the edges weighted by graph convolutional layer, representing the associated gene sets.

Proteomics graph analysis

Protein is the direct bearer of life activities, and proteomics plays an important role in elucidating the molecular mechanisms of life activities and complex diseases. You *et al.* [138] presented to use IsoMap-based embedding method to encode protein nodes in PPI network. And they measure the similarity between proteins in the embedding space to predict PPIs. An attributed network embedding method, Graph2Go [139] fused the attribute feature and graph embedding of protein and adopted VGAE and GCN to infer protein functions and GO. Yao *et al.* [140] used stacked GCN to construct more reliable PPI network by removing less credible PPIs for protein complex detection.

Transcriptomics graph analysis

The transcriptome of an organism contains a large amount of noncoding RNAs, including miRNAs, lncRNAs, circRNAs, etc., which play an important role in gene expression, cell development and diverse life activities, and are closely related to complex human diseases. MMGCN [141] developed a multiview multichannel attention-based GCN to predict miRNA–disease associations. Sheng *et al.* [142] constructed a triple-layer heterogeneous graph to integrated similarities and associations between miRNAs, lncRNAs and diseases. They also proposed heterogeneous attributed embedding methods VADLP combined random walk, convolutional autoencoder and VAE using attention mechanism to learn node feature for predicting lncRNA–disease associations. Wang *et al.* [17] proposed a FastGCN-based

Table 2. A summary of open-source implementations of graph representation learning algorithms

Category	Algorithm	Source code	Description	Last update	References
Homogeneous graph embedding	GraRep	https://github.com/ShelsonCao/GraRep	SVD; proximity matrix	2017	[65]
	HOPE	http://git.thumedia.org/embedding/HOPE	Matrix factorization	2016	[72]
	Deepwalk	https://github.com/phanein/deepwalk	Random walk; Skip-gram	2020	[74]
	node2vec	https://github.com/aditya-grover/node2vec	Biased Random walk	2016	[75]
	Walklets	https://github.com/benedekrozemberczki/walklets	Python/C++	2019	[76]
	struc2vec	https://github.com/leoribeiro/struc2vec	Python; structural identify	2018	[77]
	LINE	https://github.com/tangjianpku/LINE	C++/python	2018	[58]
	SDNE	https://github.com/suanrong/SDNE	Autoencoder; Python	2018	[78]
	DNGR	https://github.com/ShelsonCao/DNGR	MATLAB/Keras	2016	[79]
Heterogeneous graph embedding	metapath2vec	https://ericdongyx.github.io/metapath2vec/m2v.html	Meta-path	2017	[61]
	HIN2vec	https://github.com/csiesheep/hin2vec	Meta-path	2019	[62]
	GATNE	https://github.com/THUDM/GATNE	TensorFlow/PyTorch	2021	[41]
	PTE	https://github.com/mnqu/PTE	C++	2017	[80]
	HEER	https://github.com/GentleZhu/HEER	PyTorch	2019	[81]
	HERec	https://github.com/librahu/HERec	Factorization	2019	[82]
Attributed graph embedding	TADW	https://github.com/thunlp/tadw	Text information	2015	[67]
	LANE	https://github.com/xhuang31/LANE	MATLAB	2018	[85]
	ASNE	https://github.com/lizi-git/ASNE	TensorFlow	2018	[86]
	DANE	https://github.com/gaoghc/DANE	Python	2018	[87]
	ANRL	https://github.com/cszhangzhen/ANRL	TensorFlow	2020	[89]
Graph recurrent networks	GGNN	https://github.com/yujiali/ggnn	Lua; Torch	2016	[92]
	SSE	https://github.com/Hanjun-Dai/steady_state_embedding	C++	2018	[94]
Graph convolutional networks	GCN	https://github.com/tkipf/gcn	TensorFlow	2020	[95]
	DGCN	https://github.com/ZhuangCY/DGCN	Theano; semi-supervised	2018	[96]
	AGCN	https://github.com/yimutianyang/AGCN	TensorFlow; attribute inference	2020	[97]
	LGCN	https://github.com/divelab/lgcn	Large-scale; Python	2020	[98]
	FastGCN	https://github.com/matenure/FastGCN	TensorFlow	2019	[99]
	GraphSAGE	https://github.com/williamleif/GraphSAGE	Inductive learning; TensorFlow/PyTorch	2018	[100]
	GAT	https://github.com/PetarV-/GAT	Attention mechanism	2021	[102]
	GaAN	https://github.com/jennyzhang0215/GaAN	MXNet; GRU	2019	[103]
	HAN	https://github.com/Jhy1993/HAN	Heterogeneous; GAT-based	2020	[55]
Graph autoencoders	VGAE	https://github.com/tkipf/gae	TensorFlow	2020	[104]
	GC-MC	https://github.com/riannevdberg/gc-mc	Tensorflow; matrix completion	2018	[105]
	Graphite	https://github.com/ermongroup/graphite	TensorFlow; generative	2019	[106]
	Graph2Gauss	https://github.com/abojchevski/graph2gauss	Inductive learning	2019	[107]

(Continued)

Table 2. Continued

Category	Algorithm	Source code	Description	Last update	References
	GraphVAE	https://github.com/snap-stanford/GraphRNN/tree/master/baselines/graphvae	PyTorch; unofficial implementation	2018	[109]
	DGI	https://github.com/PetarV-/DGI	contrastive learning; PyTorch	2020	[110]
	infoGraph	https://github.com/fanyun-sun/InfoGraph	PyTorch	2021	[111]
Graph adversarial networks	ARGA	https://github.com/GRAND-Lab/ARGA	TensorFlow	2018	[113]
	NetRA	https://github.com/chengw07/NetRA	PyTorch	2019	[114]
	GraphGAN	https://github.com/hwwang55/GraphGAN	Adversarial training	2019	[115]
	MolGAN	https://github.com/nicola-decao/MolGAN	Molecular generation	2020	[117]
	NetGAN	https://github.com/danielzuegner/netgan	TensorFlow	2020	[118]

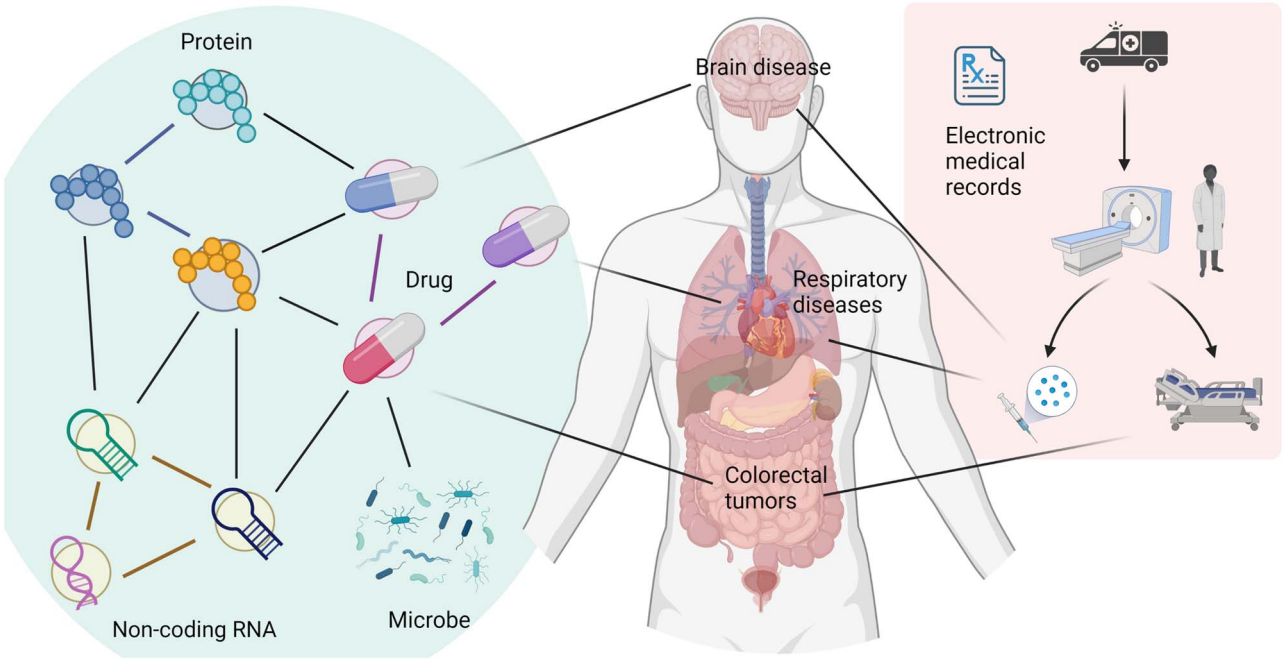

Figure 4. An overview of graphs in biomedicine at different scales. From molecules to healthcare systems, graphs are ubiquitous in the biomedical field with multiple types of intraclass and interclass relationships. The structures and functions of protein and pharmaceutical compounds can be regarded as molecular graphs. Comprehensive associations between proteins and noncoding transcripts (including miRNAs, lncRNAs, circRNAs, etc.) modeled the multi-omics graphs. Drugs, protein targets, ncRNAs, microbes, disease indications and their interactions constituted pharmaceutical graphs. And electronic medical records, personalized omics and other data can be further integrated into the healthcare systems-level knowledge graphs. These interconnected multimodal graphs can be systematically integrated and fully understood based on a holistic perspective.

method GCNCDA to predict circRNA–disease associations by merging disease semantic similarity information and GIP of circRNA. In our previous work [143], we proposed a molecular association network that systematically integrated comprehensive associations among miRNAs, lncRNAs, circRNAs, mRNAs, proteins, microbes, drugs and diseases, and proposed SDNE- and node2vec-based methods to learn node embedding. The node embedding and node attributes are fused to predict intermolecular relations such as lncRNA–protein interaction and miRNA–disease association.

Graph representation learning for pharmaceutical

Modern pharmaceuticals have huge investments, long cycles and high risks of failure. Graph representation learning can accelerate the drug discovery and drug repositioning by integrating compounds chemical information, target interactions and clinical data such as side-effects and drug combination information.

Drug–target interaction prediction

Zong *et al.* [20] applied Deepwalk on drug–target–disease tripartite graph to predict drug–target interactions. Zhao *et al.* [144] incorporated the associations between drug–target interactions and drug–protein pairs and proposed a GCN-based method to encode the feature of drug–protein pairs for inferring drug–target interactions. Based on heterogeneous graph representation learning, Peng *et al.* [145] proposed an end-to-end learning method to predict drug–target interactions based on GCN. And

they considered the association between drug, protein, disease and side effect. Thafar *et al.* [146] use node2vec to learn representation for drugs and targets and calculated similarities between drug–drug, drug–target and target–target. Nguyen *et al.* [147] introduced GraphDTA that encode the drugs as graphs and applied the GNN to predict the drug–target binding affinity.

Drug–disease association prediction

Accurate prediction of drug–disease associations can find new indications for existing drugs or promote new drug for diseases, which so called drug repositioning. Zhang *et al.* [148] proposed a similarity constrained graph factorization-based approach to predict drug–disease associations by using known drug–disease associations, drug features and disease semantic information. By using graph embedding methods such as Deepwalk, LINE, SDNE and HOPE, Guo *et al.* [149] represented the Medical Subject Headings as drug–disease graph to learn the embedding of drugs and diseases. Yu *et al.* [150] constructed a heterogeneous graph consists of known drug–disease associations, drug–drug similarities and disease–disease similarities, and proposed layer attention GCN to learn embedding of drugs and diseases for predicting drug–disease associations.

Drug–drug interaction prediction

Drug–drug interactions can affect the effects of different drug combinations and even lead to serious adverse effects. Effective drug–drug interaction prediction is critical for patients and reduction of drug development cost. Based on multiple data sources, Karim *et al.* [151] combined various knowledge graph embedding approaches with convolutional LSTM and classic machine learning classifiers for drug–drug interaction prediction. They formed a knowledge graph consists of drug features from DrugBank, PharmGKB and KEGG. Park *et al.* [152] developed an attention-based GCNs for extracting drug–drug interaction from the biomedical literature. To enhance the scalability and robustness of existing drug–drug interaction prediction methods, Chen *et al.* [153] explored a graph representation learning-based method for more accurate drug–drug interaction prediction. Celebi *et al.* [154] compared and evaluated different knowledge graph embedding methods for predicting drug–drug interaction and tested the drug–drug interaction prediction task under disjoint cross-validation.

Graph representation learning for healthcare

Recent graph representation learning-based computational methods were also used to integrate and exploit multimodal healthcare system data, such as biomedical knowledge graphs, electronic health records (EHRs), electronic medical records (EMRs) and biomedical images, to better enable personalized medicine. EHRs or EMRs are usually indicated by International Classification of Disease codes with hierarchical structure, which can naturally represent as comprehensive medical knowledge graphs. And the disease symptoms, molecules information, drug interactions and side-effects information can also be involved.

To promote clinical decision support systems in medicine and healthcare, Rotmensch *et al.* [27] developed an automated approach for mining and constructing high-quality medical knowledge graph connecting diseases and symptoms from EMRs. Ruiz *et al.* [155] proposed a powerful method to explain disease treatment, they integrated multiple disease-perturbed proteins, drug targets and biological functions into a multiscale interactome and developed a random walk-based method to capture how drugs effects through PPIs and biological functions. The multiscale interactome predicts disease treatment-related drug–disease associations, proteins and biological functions, and predicts genes that affect treatment effects and adverse reactions. Based on brain magnetic resonance imaging images, Song *et al.* [156] proposed a GCN-based method to classify Alzheimer's disease. Wu *et al.* [157] proposed ME2Vec to learn continuous low-dimensional embeddings for general entities in EHRs, and the medical services, doctors and patients are embedded by word2vec, GAT and LINE, respectively. To avoid the limitation of manually labeled EMR data, Sun *et al.* [158] introduced a GNN-based model for disease prediction by using external knowledge bases to augment the insufficient EMR data to learn effective representation for diseases, symptoms and patients based on patients record graph and medical concept graphs. And they further explored GAT and graph isomorphic network [159] aggregators for comparison. Furthermore, Choi *et al.* [160] proposed a graph convolutional transformer to learn hidden structure of EHR rather than treating EHR data as a flat-structured bag-of-features.

Challenges and opportunities

Although graph representation learning has demonstrated promising results in diverse biomedical tasks, multi-omics data integration will promote biological and medical research. However, current graph representation learning on biomedical graphs is not good enough to provide fabulous solutions for any biological and medical graph in any condition. There are some challenges and opportunities for future directions.

Data quality

Compared with clean and well-organized data in other fields, biomedical graphs are usually sparse, noisy and incomplete. At the same time, collecting original and reliable data usually requires time consuming and laborious wet experiments, and the data has the problem of high false negative and false positive rates. Meanwhile, biomedical data is scattered and accumulated quickly, lacking a good structure. In view of the sparse and incomplete of biomedical data, it is a challenging problem to better integrate multisource high-quality data and develop targeted graph representation learning approaches.

Complex graph structure

Graphs structures are flexible and complicated in biomedical and healthcare applications. Various efforts are made to handle homogeneous graph, and several works considered complex graph structures, for example heterogeneous graphs and spatial and temporal dynamic graphs. In real-life biomedical scenarios, the nodes and linkages may appear and disappear, and graphs are dynamically changed time by time. Moreover, scRNA-seq data and domain knowledge-associated data, which also offers promising opportunities but with complex graph structure. How to deal with complex biomedical graphs for downstream applications is a promising issue.

Interpretability and robustness

The risk-sensitive scenarios of biomedicine put forward higher requirements for the interpretability and robustness of graph representation learning methods, whereas neural

network-based GNNs are still black-boxes and lack of explanations. Therefore, the ability to interpret the results of deep graph learning is crucial in decision-making applications. Also, as a lot of models based on deep learning, like the domains of computer vision and natural language processing, these graph representation learning methods are vulnerable to adversarial attacks. More robust and interpretable graph representation learning methods are important to apply graph representation learning on biomedical problems with trusted explanations and credible defenses.

Conclusion

Graph representation learning bridges comprehensive graph-structured biomedical data and advanced machine learning methods, which promotes biomedical research from molecules to healthcare systems. In this work, we conducted a comprehensive and structured survey of graph representation learning and its applications in bioinformatics. Both graph embedding methods, including homogeneous graph embedding, heterogeneous graph embedding, attribute network embedding, and emerging GNNs such as GRNs, GCNs, GAEs and GGANs are summarized. And we analyzed representative applications of graph representation learning for molecules, genomics, pharmaceutical and healthcare. In addition, open resource platforms and libraries for graph representation learning are also provided. It is anticipated that this work could promote graph representation learning and biomedical studies.

Key Points

- From molecules to healthcare systems, graphs are ubiquitous to effectively integrate and model multi-source multimodal biomedical relational data, which is rapidly generated and accumulated by biomedical research.
- Graph representation learning bridges non-Euclidean biomedical graphs and machine learning techniques to promotes drug discovery, molecular mechanisms exploration, complex diseases diagnosis and treatment, and healthcare.
- This article provides a clear and comprehensive summary of graph representation learning and its brilliant applications in the biomedicine field, which will benefit interdisciplinary researchers as useful guidance.
- Challenges and opportunities are discussed for future research, and graph computing platforms and implementations are provided to accelerate benchmark research and development of methods suitable for bioinformatics problems.

Data Availability Statement

No new data were generated or analyzed in support of this research.

Funding

This work was supported by the National Natural Science Foundation of China (under grants 61873212, 61732012), in part by the National Outstanding Youth Science Foundation of National Natural Science Foundation of China (under grant 61722212).

References

1. Freeman LC. Visualizing social networks. *J Soc Struct* 2000;**1**:4.
2. Kajikawa Y, Ohno J, Takeda Y, *et al*. Creating an academic landscape of sustainability science: an analysis of the citation network. *Sustain Sci* 2007;**2**:221.
3. RFI C, Solé RV. The small world of human language. *Proc R Soc Lond Series B Biol Sci* 2001;**268**:2261–5.
4. Thattai M, van Oudenaarden A. Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci* 2001;**98**:8614–9.
5. Theocharidis A, van Dongen S, Enright AJ, *et al*. Network visualization and analysis of gene expression data using BioLayout Express3D. *Nat Protoc* 2009;**4**:1535–50.
6. Sporns O, Tononi G, Kötter R. The human connectome: a structural description of the human brain. *PLoS Comput Biol* 2005;**1**:e42.
7. Venkatesan K, Rual J-F, Vazquez A, *et al*. An empirical framework for binary interactome mapping. *Nat Methods* 2009;**6**:83–90.
8. Zhang Z-C, Zhang X-F, Wu M, *et al*. A graph regularized generalized matrix factorization model for predicting links in biomedical bipartite networks. *Bioinformatics* 2020;**36**:3474–81.
9. Stelzl U, Worm U, Lalowski M, *et al*. A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 2005;**122**:957–68.
10. Vascon S, Frasca M, Tripodi R, *et al*. Protein function prediction as a graph-transduction game. *Pattern Recognit Lett* 2020;**134**:96–105.
11. Cheng F, Liu C, Jiang J, *et al*. Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput Biol* 2012;**8**:e1002503.
12. Manoochehri HE, Nourani M. Drug-target interaction prediction using semi-bipartite graph model and deep learning. *BMC Bioinformatics* 2020;**21**:1–16.
13. Daminelli S, Haupt VJ, Reimann M, *et al*. Drug repositioning through incomplete bi-cliques in an integrated drug–target–disease network. *Integr Biol* 2012;**4**:778–88.
14. Ding Y, Tian L-P, Lei X, *et al*. Variational graph auto-encoders for miRNA-disease association prediction. *Methods* 2021;**192**:25–34.
15. Li J, Zhang S, Liu T, *et al*. Neural inductive matrix completion with graph convolutional networks for miRNA-disease association prediction. *Bioinformatics* 2020;**36**:2538–46.
16. Tan H, Sun Q, Li G, *et al*. Multiview consensus graph learning for lncRNA–disease association prediction. *Front Genet* 2020;**11**:89.
17. Wang L, You Z-H, Li Y-M, *et al*. GCNCDA: a new method for predicting circRNA-disease associations based on graph convolutional network algorithm. *PLoS Comput Biol* 2020;**16**:e1007568.
18. Salmena L, Poliseno L, Tay Y, *et al*. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 2011;**146**:353–8.
19. Tran NC, Gao JX. Network representation of large-scale heterogeneous RNA sequences with integration of diverse multi-omics, interactions, and annotations data. *Pac Symp Biocomput* 2020;**25**:499–510.
20. Zong N, Kim H, Ngo V, *et al*. Deep mining heterogeneous networks of biomedical linked data to predict novel drug–target associations. *Bioinformatics* 2017;**33**:2337–44.
21. Zong N, Wong RSN, Yu Y, *et al*. Drug–target prediction utilizing heterogeneous bio-linked network embeddings. *Brief Bioinform* 2021;**22**:568–80.

22. Zheng Y, Peng H, Zhang X, *et al*. Predicting adverse drug reactions of combined medication from heterogeneous pharmacologic databases. *BMC Bioinformatics* 2018; **19**:49–59.
23. Peng H, Lan C, Zheng Y, *et al*. Cross disease analysis of co-functional microRNA pairs on a reconstructed network of disease-gene-microRNA tripartite. *BMC Bioinformatics* 2017;**18**:193.
24. Davis AP, Murphy CG, Saraceni-Richards CA, *et al*. Comparative Toxicogenomics Database: a knowledgebase and discovery tool for chemical–gene–disease networks. *Nucleic Acids Res* 2008;**37**:D786–92.
25. Liu Y, Zeng X, He Z, *et al*. Inferring microRNA-disease associations by random walk on a heterogeneous network with multiple data sources. *IEEE/ACM Trans Comput Biol Bioinform* 2017;**14**:905–15.
26. Nickel M, Murphy K, Tresp V, *et al*. A review of relational machine learning for knowledge graphs. *Proc IEEE* 2015;**104**:11–33.
27. Rotmensch M, Halpern Y, Tlimat A, *et al*. Learning a health knowledge graph from electronic medical records. *Sci Rep* 2017;**7**:5994.
28. Yi H-C, You Z-H, Huang D-S, *et al*. Learning representations to predict intermolecular interactions on large-scale heterogeneous molecular association network. *iScience* 2020;**23**:101261.
29. Hewett M, Oliver DE, Rubin DL, *et al*. PharmGKB: the pharmacogenetics knowledge base. *Nucleic Acids Res* 2002; **30**:163–5.
30. Wishart DS, Knox C, Guo AC, *et al*. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res* 2008;**36**:901–6.
31. Ashburner M, Ball CA, Blake JA, *et al*. Gene ontology: tool for the unification of biology. *Nat Genet* 2000;**25**:25–9.
32. Kibbe WA, Arze C, Felix V, *et al*. Disease Ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic Acids Res* 2015;**43**:D1071–8.
33. Kanehisa M, Araki M, Goto S, *et al*. KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 2007;**36**: D480–4.
34. Min S, Lee B, Yoon S. Deep learning in bioinformatics. *Brief Bioinform* 2017;**18**:851–69.
35. Angermueller C, Pärnamaa T, Parts L, *et al*. Deep learning for computational biology. *Mol Syst Biol* 2016;**12**:878.
36. Leung MK, Delong A, Alipanahi B, *et al*. Machine learning in genomic medicine: a review of computational problems and data sets. *Proc IEEE* 2015;**104**:176–97.
37. Cai H, Zheng VW, Chang KC-C. A comprehensive survey of graph embedding: problems, techniques, and applications. *IEEE Trans Knowl Data Eng* 2018;**30**:1616–37.
38. Wang Q, Mao Z, Wang B, *et al*. Knowledge graph embedding: a survey of approaches and applications. *IEEE Trans Knowl Data Eng* 2017;**29**:2724–43.
39. Cui P, Wang X, Pei J, *et al*. A survey on network embedding. *IEEE Trans Knowl Data Eng* 2018;**31**:833–52.
40. Goyal P, Ferrara E. Graph embedding techniques, applications, and performance: a survey. *Knowl Based Syst* 2018; **151**:78–94.
41. Cen Y, Zou X, Zhang J, *et al*. Representation learning for attributed multiplex heterogeneous network. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, Anchorage, AK, USA, 2019, p. 1358–68.
42. Wu Z, Pan S, Chen F, *et al*. A comprehensive survey on graph neural networks. *IEEE Trans Neural Netw Learn Syst* 2020;**32**:4–24.
43. Zhou J, Cui G, Hu S, *et al*. Graph neural networks: a review of methods and applications. *AI Open* 2020;**1**:57–81.
44. Zhang Z, Cui P, Zhu W. Deep learning on graphs: a survey. *IEEE Trans Knowl Data Eng* 2020. doi: 10.1109/TKDE.2020.2981333.
45. Li MM, Huang K, Zitnik M. Representation learning for networks in biology and medicine: advancements, challenges, and opportunities. *arXiv* preprint arXiv:2104.04883, 2021.
46. Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* 2011;**12**:56–68.
47. Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature* 2009;**461**:218–23.
48. Pavlopoulos GA, Kontou PI, Pavlopoulou A, *et al*. Bipartite graphs in systems biology and medicine: a survey of methods and applications. *GigaScience* 2018;**7**:giy014.
49. Su C, Tong J, Zhu Y, *et al*. Network embedding in biomedical data science. *Brief Bioinform* 2020;**21**:182–97.
50. Ding P, Ouyang W, Luo J, *et al*. Heterogeneous information network and its application to human health and disease. *Brief Bioinform* 2020;**21**:1327–46.
51. Yue X, Wang Z, Huang J, *et al*. Graph embedding on biomedical networks: methods, applications and evaluations. *Bioinformatics* 2020;**36**:1241–51.
52. Muzio G, O'Bray L, Borgwardt K. Biological network analysis with deep learning. *Brief Bioinform* 2020;**22**:1515–30.
53. Nelson W, Zitnik M, Wang B, *et al*. To embed or not: network embedding as a paradigm in computational biology. *Front Genet* 2019;**10**:381.
54. Sun Y, Han J. Mining heterogeneous information networks: a structural analysis approach. *ACM SIGKDD Explor Newsletter* 2013;**14**:20–8.
55. Wang X, Ji H, Shi C, *et al*. Heterogeneous graph attention network. In: *The World Wide Web Conference*. Association for Computing Machinery, San Francisco, CA, USA, 2019, p. 2022–32.
56. Zhang D, Yin J, Zhu X, *et al*. Network representation learning: a survey. *IEEE Trans Big Data* 2020;**6**:3–28.
57. Shi Y, Han F, He X, *et al*. mvn2vec: preservation and collaboration in multi-view network embedding. *arXiv* preprint arXiv:1801.06597, 2018.
58. Tang J, Qu M, Wang M, *et al*. LINE: large-scale information network embedding. In: *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, Florence, Italy, 2015, p. 1067–77.
59. Ma Y, Ren Z, Jiang Z, *et al*. Multi-dimensional network embedding with hierarchical structure. In: *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. Association for Computing Machinery, Marina Del Rey, CA, USA, 2018, p. 387–95.
60. Ma Y, Wang S, Aggarwal CC, *et al*. Multi-dimensional graph convolutional networks. In: *Proceedings of the 2019 SIAM International Conference on Data Mining*. SIAM, 2019, p. 657–65.
61. Dong Y, Chawla NV, Swami A. metapath2vec: scalable representation learning for heterogeneous networks. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, Halifax, NS, Canada, 2017, p. 135–44.
62. Fu T-Y, Lee W-C, Lei Z. Hin2Vec: explore meta-paths in heterogeneous information networks for representation

learning. In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. Association for Computing Machinery, 2017, p. 1797–806.
63. Park C, Han J, Yu H. Deep multiplex graph infomax: attentive multiplex network embedding using global information. *Knowl Based Syst* 2020;**197**:105861.
64. Sun Y, Han J, Yan X, *et al*. Pathsim: meta path-based top-k similarity search in heterogeneous information networks. *Proc VLDB Endow* 2011;**4**:992–1003.
65. Cao S, Lu W, Xu Q. GraRep: learning graph representations with global structural information. In: *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. Association for Computing Machinery, Melbourne, Australia, 2015, p. 891–900.
66. Zhang D, Yin J, Zhu X, *et al*. Homophily, structure, and content augmented network representation learning. In: *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE, 2016, p. 609–18.
67. Yang C, Liu Z, Zhao D, *et al*. Network representation learning with rich text information. In: *Proceedings of the 24th International Conference on Artificial Intelligence*. AAAI Press, Buenos Aires, Argentina, 2015, p. 2111–7.
68. Tenenbaum JB, De Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science* 2000;**290**:2319–23.
69. Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science* 2000;**290**:2323–6.
70. Belkin M, Niyogi P. Laplacian Eigenmaps for dimensionality reduction and data representation. *Neural Comput* 2003;**15**:1373–96.
71. Ahmed A, Shervashidze N, Narayanamurthy S, *et al*. Distributed large-scale natural graph factorization. In: *Proceedings of the 22nd International Conference on World Wide Web*. Association for Computing Machinery, Rio de Janeiro, Brazil, 2013, p. 37–48.
72. Ou M, Cui P, Pei J, *et al*. Asymmetric transitivity preserving graph embedding. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, San Francisco, California, USA, 2016, p. 1105–14.
73. Mikolov T, Sutskever I, Chen K, *et al*. Distributed representations of words and phrases and their compositionality. *arXiv preprint arXiv* preprint arXiv:1310.4546, 2013.
74. Perozzi B, Al-Rfou R, Skiena S. DeepWalk: online learning of social representations. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, New York, New York, USA, 2014, p. 701–10.
75. Grover A, Leskovec J. node2vec: scalable feature learning for networks. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, San Francisco, California, USA, 2016, p. 855–64.
76. Perozzi B, Kulkarni V, Chen H, *et al*. Don't Walk, Skip! Online learning of multi-scale network embeddings. In: *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*. Association for Computing Machinery, Sydney, Australia, 2017, p. 258–65.
77. Ribeiro LFR, Saverese PHP, Figueiredo DR. *struc2vec*: learning node representations from structural identity. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, Halifax, NS, Canada, 2017, p. 385–94.
78. Wang D, Cui P, Zhu W. Structural deep network embedding. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, San Francisco, California, USA, 2016, p. 1225–34.
79. Cao S, Lu W, Xu Q. Deep neural networks for learning graph representations. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, 2016;**30**:1145–52.
80. Tang J, Qu M, Mei Q. PTE: predictive text embedding through large-scale heterogeneous text networks. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, 2015, p. 1165–74.
81. Shi Y, Zhu Q, Guo F, *et al*. Easing embedding learning by comprehensive transcription of heterogeneous information networks. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, London, United Kingdom, 2018, p. 2190–9.
82. Shi C, Hu B, Zhao WX, *et al*. Heterogeneous information network embedding for recommendation. *IEEE Trans Knowl Data Eng* 2018;**31**:357–70.
83. Chang S, Han W, Tang J, *et al*. Heterogeneous network embedding via deep architectures. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, Sydney, NSW, Australia, 2015, p. 119–28.
84. Zhang J, Xia C, Zhang C, *et al*. BL-MNE: emerging heterogeneous social network embedding through broad learning with aligned autoencoder. In: 2017 IEEE International Conference on Data Mining (ICDM). IEEE, 2017, 605–14.
85. Huang X, Li J, Hu X. Label informed attributed network embedding. In: *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. Association for Computing Machinery, Cambridge, United Kingdom, 2017, p. 731–9.
86. Liao L, He X, Zhang H, *et al*. Attributed social network embedding. *IEEE Trans Knowl Data Eng* 2018;**30**: 2257–70.
87. Gao H, Huang H. Deep attributed network embedding. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. AAAI Press, Stockholm, Sweden, 2018, p. 3364–70.
88. Liu M, Liu J, Chen Y, *et al*. AHNG: representation learning on attributed heterogeneous network. *Inf Fusion* 2019; **50**:221–30.
89. Zhang Z, Yang H, Bu J, *et al*. ANRL: attributed network representation learning via deep neural networks. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. AAAI Press, Stockholm, Sweden, 2018, 3155–61.
90. Gori M, Monfardini G, Scarselli F. A new model for learning in graph domains. In: *Proceedings. 2005 IEEE International Joint Conference on Neural Networks*. IEEE, 2005, p. 729–34.
91. Scarselli F, Gori M, Tsoi AC, *et al*. The graph neural network model. *IEEE Trans Neural Netw* 2008;**20**:61–80.
92. Li Y, Tarlow D, Brockschmidt M, *et al*. Gated graph sequence neural networks. *arXiv* preprint arXiv:1511.05493, 2015.
93. Cho K, Van Merriënboer B, Gulcehre C, *et al*. (eds). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* preprint arXiv:1406.1078, 2014.

94. Dai H, Kozareva Z, Dai B, *et al*. Learning steady-states of iterative algorithms over graphs. In: *International Conference on Machine Learning*. PMLR, 2018, p. 1106–14.
95. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. *arXiv* preprint arXiv: 1609.02907, 2016.
96. Zhuang C, Ma Q. Dual graph convolutional networks for graph-based semi-supervised classification. In: *Proceedings of the 2018 World Wide Web Conference*. International World Wide Web Conferences Steering Committee, Lyon, France, 2018, p. 499–508.
97. Li R, Wang S, Zhu F, *et al*. Adaptive graph convolutional neural networks. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018;**32**:3546–53.
98. Gao H, Wang Z, Ji S. Large-scale learnable graph convolutional networks. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, London, United Kingdom, 2018, p. 1416–24.
99. Chen J, Ma T, Xiao C. FastGCN: fast learning with graph convolutional networks via importance sampling. *arXiv* preprint arXiv:1801.10247, 2018.
100. Hamilton WL, Ying R, Leskovec J. Inductive representation learning on large graphs. *arXiv* preprint arXiv:1706.02216, 2017.
101. Vaswani A, Shazeer N, Parmar N, *et al*. Attention is all you need. *Adv Neural Inf Process Syst* 2017;5998–6008.
102. Veličković P, Cucurull G, Casanova A, *et al*. Graph attention networks. *arXiv* preprint arXiv:1710.10903, 2017.
103. Zhang J, Shi X, Xie J, *et al*. GaAN: gated attention networks for learning on large and spatiotemporal graphs. *arXiv* preprint arXiv:1803.07294, 2018.
104. Kipf TN, Welling M. Variational graph auto-encoders. *arXiv* preprint arXiv:1611.07308, 2016.
105. Rvd B, Kipf TN, Welling M. Graph convolutional matrix completion. *arXiv* preprint arXiv:1706.02263, 2017.
106. Grover A, Zweig A, Ermon S. Graphite: iterative generative modeling of graphs. In: *International Conference on Machine Learning*. PMLR, 2019, p. 2434–44.
107. Bojchevski A, Günnemann S. Deep Gaussian embedding of graphs: unsupervised inductive learning via ranking. In: *International Conference on Learning Representations*. 2018.
108. Zhu D, Cui P, Wang D, *et al*. Deep variational network embedding in Wasserstein space. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, London, UK, 2018, p. 2827–36.
109. Simonovsky M, Komodakis N. GraphVAE: towards generation of small graphs using variational autoencoders. In: *International Conference on Artificial Neural Networks*. Springer, 2018, p. 412–422.
110. Velickovic P, Fedus W, Hamilton WL, *et al*. Deep Graph Infomax. In: *ICLR (Poster)*, 2019.
111. Sun F-Y, Hoffmann J, Verma V, *et al*. Infograph: unsupervised and semi-supervised graph-level representation learning via mutual information maximization. *arXiv* preprint arXiv:1908.01000, 2019.
112. Goodfellow I, Pouget-Abadie J, Mirza M, *et al*. Generative adversarial networks. *Commun ACM* 2020;**63**:139–44.
113. Pan S, Hu R, Fung SF, *et al*. Learning graph embedding with adversarial training methods. *IEEE Trans Cybernet* 2020;**50**:2475–87.
114. Yu W, Zheng C, Cheng W, *et al*. Learning deep network representations with adversarially regularized autoencoders. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, London, UK, 2018, p. 2663–71.
115. Wang H, Wang J, Wang J, *et al*. GraphGAN: graph representation learning with generative adversarial nets. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018;**32**:2508–15.
116. Dai Q, Li Q, Tang J, *et al*. Adversarial network embedding, Proc AAAI Conf Artif Intell. 2018;**32**:2167–74.
117. De Cao N, Kipf T. MolGAN: an implicit generative model for small molecular graphs. *arXiv* preprint arXiv:1805.11973, 2018.
118. Bojchevski A, Shchur O, Zügner D, *et al*. Netgan: generating graphs via random walks. In: *International Conference on Machine Learning*. PMLR, 2018, p. 610–9.
119. Hu Z, Dong Y, Wang K, *et al*. GPT-GNN: generative pre-training of graph neural networks. In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, 2020, p. 1857–67.
120. Zhang J, Zhang H, Xia C, *et al*. Graph-Bert: only attention is needed for learning graph representations. *arXiv* preprint arXiv:2001.05140, 2020.
121. Li G, Müller M, Ghanem B, *et al*. Training graph neural networks with 1000 layers. *arXiv* preprint arXiv:2106.07476, 2021.
122. Hu Y, You H, Wang Z, *et al*. Graph-MLP: node classification without message passing in graph. *arXiv* preprint arXiv:2106.040, 2021.
123. Duvenaud D, Maclaurin D, Aguilera-Iparraguirre J, *et al*. Convolutional networks on graphs for learning molecular fingerprints. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*. MIT Press, Montreal, Canada, 2015, p. 2224–32.
124. Gainza P, Sverrisson F, Monti F, *et al*. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nat Methods* 2020;**17**:184–92.
125. Li C, Wei W, Li J, *et al*. 3DMol-Net: learn 3D molecular representation using adaptive graph convolutional network based on rotation invariance. *IEEE J Biomed Health Inform* 2021;1–1.
126. Gilmer J, Schoenholz SS, Riley PF, *et al*. Neural message passing for quantum chemistry. In: *International Conference on Machine Learning*. PMLR, 2017, p. 1263–1272.
127. Wang X, Li Z, Jiang M, *et al*. Molecule property prediction based on spatial graph embedding. *J Chem Inf Model* 2019;**59**:3817–28.
128. Wieder O, Kohlbacher S, Kuenemann M, *et al*. A compact review of molecular property prediction with graph neural networks. *Drug Discov Today Technol* 2020.
129. Jin W, Barzilay R, Jaakkola T. Junction tree variational autoencoder for molecular graph generation. In: *International Conference on Machine Learning*. PMLR, 2018, p. 2323–32.
130. Shi C, Xu M, Zhu Z, *et al*. GraphAF: a flow-based autoregressive model for molecular graph generation. *arXiv* preprint arXiv:2001.09382, 2020.
131. Zang C, Wang F. MoFlow: an invertible flow model for generating molecular graphs. In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Virtual Event*. Association for Computing Machinery, CA, USA, 2020, p. 617–26.
132. Mahmood O, Mansimov E, Bonneau R, *et al*. Masked graph modeling for molecule generation. *Nat Commun* 2021;**12**:3156.

133. Li X, Chen W, Chen Y, *et al*. Network embedding-based representation learning for single cell RNA-seq data. *Nucleic Acids Res* 2017;**45**:e166–6.
134. Li Y, Li J. Disease gene identification by random walk on multigraphs merging heterogeneous genomic and phenotype data. *BMC Genomics* 2012;**13**:S27.
135. Han P, Yang P, Zhao P, *et al*. GCN-MF: disease-gene association identification by graph convolutional networks and matrix factorization. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, Anchorage, AK, USA, 2019, p. 705–13.
136. Yang C, Zhuang P, Shi W, *et al*. Conditional structure generation through graph variational generative adversarial nets. In: *Proceedings of the 33rd International Conference on Neural Information Processing Systems*. Curran Associates Inc., 2019, p. 1338–49.
137. Rhee S, Seo S, Kim S. Hybrid approach of relation network and localized graph convolutional filtering for breast cancer subtype classification. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. AAAI Press, Stockholm, Sweden, 2018, p. 3527–34.
138. You Z-H, Lei Y-K, Gui J, *et al*. Using manifold embedding for assessing and predicting protein interactions from high-throughput experimental data. *Bioinformatics* 2010;**26**:2744–51.
139. Fan K, Guan Y, Zhang Y. Graph2GO: a multi-modal attributed network embedding method for inferring protein functions. *GigaScience* 2020;**9**:giaa081.
140. Yao H, Guan J, Liu T. Denoising protein–protein interaction network via variational graph auto-encoder for protein complex detection. *J Bioinform Comput Biol* 2020;**18**: 2040010.
141. Tang X, Luo J, Shen C, *et al*. Multi-view multichannel attention graph convolutional network for miRNA–disease association prediction. *Brief Bioinform* 2021;bbab174.
142. Sheng N, Cui H, Zhang T, *et al*. Attentional multi-level representation encoding based on convolutional and variance autoencoders for lncRNA–disease association prediction. *Brief Bioinform* 2021;**22**:bbaa067.
143. Yi HC, You ZH, Guo ZH, *et al*. Learning representation of molecules in association network for predicting intermolecular associations. *IEEE/ACM Trans Comput Biol Bioinform* 2020;1.
144. Zhao T, Hu Y, Valsdottir LR, *et al*. Identifying drug–target interactions based on graph convolutional network and deep neural network. *Brief Bioinform* 2020;**22**:2141–50.
145. Peng J, Wang Y, Guan J, *et al*. An end-to-end heterogeneous graph representation learning-based framework for drug–target interaction prediction. *Brief Bioinform* 2021; bbaa430.
146. Thafar MA, Olayan RS, Ashoor H, *et al*. DTiGEMS+: drug–target interaction prediction using graph embedding, graph mining, and similarity-based techniques. *J Chem* 2020;**12**: 1–17.
147. Nguyen T, Le H, Quinn TP, *et al*. GraphDTA: predicting drug–target binding affinity with graph neural networks. *Bioinformatics* 2021;**37**:1140–7.
148. Zhang W, Yue X, Lin W, *et al*. Predicting drug-disease associations by using similarity constrained matrix factorization. *BMC Bioinformatics* 2018;**19**:233.
149. Guo Z-H, You Z-H, Huang D-S, *et al*. MeSHHeading2vec: a new method for representing MeSH headings as vectors based on graph embedding algorithm. *Brief Bioinform* 2021;**22**:2085–95.
150. Sun M, Zhao S, Gilvary C, *et al*. Graph convolutional networks for computational drug development and discovery. *Brief Bioinform* 2019;**21**:919–35.
151. Karim MR, Cochez M, Jares JB, *et al*. Drug-drug interaction prediction based on knowledge graph embeddings and convolutional-LSTM network. In: *Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*. Association for Computing Machinery, Niagara Falls, NY, USA, 2019, p. 113–23.
152. Park C, Park J, Park S. AGCN: attention-based graph convolutional networks for drug-drug interaction extraction. *Expert Syst Appl* 2020;**159**:113538.
153. Chen X, Liu X, Wu J. Drug-drug interaction prediction with graph representation learning. In: *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2019, p. 354–61.
154. Celebi R, Uyar H, Yasar E, *et al*. Evaluation of knowledge graph embedding approaches for drug-drug interaction prediction in realistic settings. *BMC Bioinformatics* 2019;**20**:726.
155. Ruiz C, Zitnik M, Leskovec J. Identification of disease treatment mechanisms through the multiscale interactome. *Nat Commun* 2021;**12**:1796.
156. Song TA, Chowdhury SR, Yang F, *et al*. Graph convolutional neural networks for Alzheimer's disease classification. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019, p. 414–17.
157. Wu T, Wang Y, Wang Y, *et al*. Leveraging graph-based hierarchical medical entity embedding for healthcare applications. *Sci Rep* 2021;**11**:5858.
158. Sun Z, Yin H, Chen H, *et al*. Disease prediction via graph neural networks. *IEEE J Biomed Health Inform* 2021;**25**:818–26.
159. Xu K, Hu W, Leskovec J, *et al*. How powerful are graph neural networks? *arXiv* preprint arXiv:1810.00826, 2018.
160. Choi E, Xu Z, Li Y, *et al*. Learning the graphical structure of electronic health records with graph convolutional transformer. *Proc AAAI Conf Artif Intell* 2020;**34**:606–13.
161. Cen Y, Hou Z, Wang Y, *et al*. CogDL: an extensive toolkit for deep learning on graphs. *arXiv* preprint arXiv:2103.00959, 2021.
162. Fey M, Lenssen JE. Fast graph representation learning with PyTorch geometric. *arXiv* preprint arXiv:1903.02428, 2019.
163. Wang M, Zheng D, Ye Z, *et al*. Deep graph library: a graph-centric, highly-performant package for graph neural networks. *arXiv* preprint arXiv:1909.0131, 2019.
164. Liu M, Luo Y, Wang L, *et al*. DIG: a turnkey library for diving into graph deep learning research. *arXiv* preprint arXiv:2103.12608, 2021.
165. Zhu Z, Xu S, Tang J, *et al*. GraphVite: a high-performance CPU-GPU hybrid system for node embedding. In: *The World Wide Web Conference*. Association for Computing Machinery, San Francisco, CA, USA, 2019, p. 2494–504.
166. Yang H. AliGraph: a comprehensive graph neural network platform. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Association for Computing Machinery, Anchorage, AK, USA, 2019, p. 3165–6.