

·综述·

孙连伟,连雪全,杨佳林,李建伟

(河北工业大学人工智能与数据科学学院,天津 300401)

摘要: 药物-靶标预测是现代药物研发中至关重要的一个环节,它克服了传统药物研发中发现和验证药物-靶标对的高成本和周期长等问题。药物-靶标相互作用(DTI)预测和药物-靶标亲和力(DTA)预测都是药物-靶标预测的任务。预测药物-靶标相互作用,筛选出具有潜在相互作用的药物-靶标对,对药物重定位以及新药研发具有重要意义。预测药物-靶标亲和力对药物的设计和优化可起到辅助性作用。图神经网络(GNN)作为一种用于处理图结构数据的深度学习模型,它可有效地对同构网络或异构网络中的节点和边进行表示学习,现已被广泛应用于各种药物-靶标预测任务中。本文对图神经网络在药物-靶标预测领域的研究进展作一综述,以期图神经网络在药物研发领域的进一步发展提供参考。

关键词: 药物-靶标预测;药物-靶标相互作用;药物-靶标亲和力;图神经网络;异构网络

中图分类号: Q811

文献标识码: A

DOI: 10.3969/j.issn.1006-1959.2025.08.036

文章编号: 1006-1959(2025)08-0167-05

Research Progress of Graph Neural Network in the Field of Drug-target Prediction

SUN Lianwei, LIAN Xuequan, YANG Jialin, LI Jianwei

(School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China)

Abstract: Drug-target prediction is a crucial part of modern drug research and development. It overcomes the problems of high cost and long cycle in the discovery and verification of drug-target pairs in traditional drug research and development. Drug-target interaction (DTI) prediction and drug-target affinity (DTA) prediction are both tasks of drug-target prediction. Predicting drug-target interactions and screening drug-target pairs with potential interactions are of great significance for drug repositioning and new drug development. Predicting drug-target affinity can play an auxiliary role in drug design and optimization. As a deep learning model for processing graph-structured data, graph neural network (GNN) can effectively represent nodes and edges in homogeneous or heterogeneous networks, and has been widely used in various drug-target prediction tasks. This paper reviews the research progress of graph neural network in the field of drug-target prediction, in order to provide reference for the further development of graph neural network in the field of drug research and development.

Key words: Drug-target prediction; Drug-target interaction; Drug-target affinity; Graph neural network; Heterogeneous network

药物研发是一个漫长且复杂的过程,涉及到药物发现、设计、优化和临床试验等多个环节^[1]。传统的药物研发工作费时费力且成本高昂。为了加速药物研发进程,人们不断努力寻找新的技术和方法。药物靶标在药物开发中具有重要意义,它们是药物与生物体内特定分子相互作用的目标。药物靶标可以是蛋白质、酶、受体、离子通道等生物分子,这些分子在疾病的发生和发展进程中扮演着重要的角色。因此,药物-靶标预测已成为药物研发中的一个重要研究方向^[2]。药物-靶标预测(drug-target prediction)是一种基于大量的已知药物-靶标相互作用数据、药物信息和蛋白质靶标信息,通过机器学习、深度学习等方法,对药物分子和蛋白质靶标之间的相互作用和亲和力进行预测的研究方法^[3]。在分析与

药物和靶标相关的数据基础上,药物-靶标预测通过构建药物分子与蛋白质靶标之间的关联模型,预测新的药物分子与潜在靶标之间的相互作用和亲和力值,指导药物研发的方向选择和优化,减少研发过程中的不确定性^[4]。图神经网络(graph neural network, GNN)是一类用于处理图结构数据的深度学习模型。它的核心思想是通过聚合节点的特征信息和邻居节点的拓扑信息,获得节点的综合表示,实现对图结构的全局理解和预测^[5]。在药物-靶标预测领域中,GNN模型已被广泛应用,可更好地理解药物与靶标之间的相互作用机制,发现新的药物靶点,加速新药的研发。本文对图神经网络应用于药物-靶标预测的研究进展进行综述,方便研究者能够快速了解图神经网络在药物-靶标预测领域的研究现状,以期图神经网络与药物研发相结合的进一步发展提供帮助。

1 GNN 简介

GNN 的框架结构通常由输入层、多个卷积层、非线性激活函数和输出层组成^[6],见图 1。在输入层

基金项目:国家自然科学基金项目(编号:62072154)

作者简介:孙连伟(1999.2-),男,河北唐山人,硕士研究生,主要从事药物相关的生物信息学研究

通讯作者:李建伟(1974.11-),男,河北唐山人,博士,教授,主要从事生物信息学研究

中,图数据被表示为节点特征矩阵和邻接矩阵的形式。其中,节点特征矩阵用来存储每个节点的属性信息,例如节点的数值特征或者向量表示;邻接矩阵则用来描述图中节点之间的拓扑关系,记录节点之间的边及其权重。卷积层是 GNN 的核心组件,通过多次迭代来更新节点的特征表示。在每一次迭代中,卷积层会聚合节点的邻居信息,并将其传递给目标

节点。因此,节点的特征表示可以逐渐从局部信息扩展到全局信息,获得更深层次且具有表征能力的特征表示^[7]。非线性激活函数在 GNN 中起到引入非线性变换的作用,将节点的特征映射到非线性的空间中,增加模型的表达能力。常见的非线性激活函数有 ReLU、Sigmoid 和 Tanh 等。输出层则将节点的特征表示映射到预测的标签或分类结果。

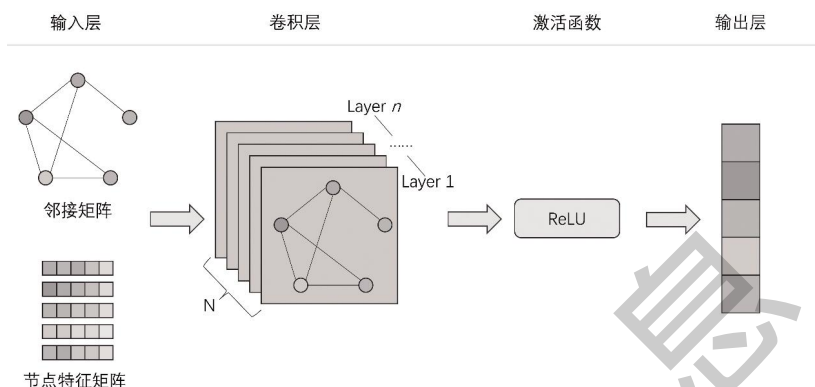


图 1 GNN 的框架图

近年来,图神经网络在许多领域中都得到了广泛应用,例如:社交网络分析、推荐系统、生物信息网络数据挖掘等。药物和蛋白质靶标可自然表示为带有化学元素或氨基酸节点的图形结构,GNN 方法可有效对药物-靶标相互作用进行预测。在具体研究中,研究者们提出了许多不同类型的 GNN 模型。图卷积神经网络(graph convolutional network, GCN)是最早、最经典的 GNN 模型之一,它通过对节点和其邻居节点的特征进行加权平均来更新节点表示^[8]。GraphSAGE 通过从节点和其邻居中采样一小部分节点,并使用聚合函数(如平均池化、最大池化等)来聚合邻居节点的特征^[9]。图注意力网络(graph attention network, GAT)是一种使用注意力机制的 GNN 模型,它通过计算节点与邻居节点之间的注意力权重,来决定邻居节点对目标节点的贡献程度^[10]。图同构网络(graph isomorphism network, GIN)是一种基于图同构性的 GNN 模型,它通过对节点的特征进行非线性变换和聚合,来获得节点的特征表示^[11]。GIN 模型对节点的邻居节点的顺序不敏感,常用来完成对图结构变化不敏感的任务。

2 药物-靶标预测简介

2.1 药物-靶标相互作用预测 药物-靶标相互作用预测是指通过计算和模型构建来预测药物分子与蛋白质靶标之间是否存在相互作用的方法。传统药物

研发中,筛选和验证药物的靶标通常需要大量的实验和时间,成本高昂且效率低下。药物-靶标相互作用预测的基本原理是将药物和靶标的特征进行编码和表示,通过训练和优化模型,预测药物与靶标之间的相互作用概率^[12]。预测结果可帮助研究人员筛选出存在潜在相互作用的药物-靶标对,降低药物发现过程中的实验成本。药物-靶标相互作用预测在药物研发中具有重要实际意义,可有效加速药物发现过程、降低实验成本和时间,为药物研发提供重要指导和决策。

2.2 药物-靶标亲和力预测 药物-靶标亲和力预测是指预测药物与蛋白质靶标之间的亲和力或结合能力。药物与靶标之间的亲和力反映了它们之间相互作用的强度,亲和力高表示药物与靶标结合紧密^[13]。预测药物-靶标亲和力的方法较多,包括基于分子结构和物理化学性质的计算方法、机器学习和深度学习等方法^[14]。这些方法通过对药物和靶标的特征进行编码和表示,构建预测模型,从而预测药物与靶标之间的亲和力。预测结果可以帮助研究人员评估药物与靶标之间的结合能力,筛选出具有高亲和力的化合物。药物-靶标亲和力预测有助于理解药物的作用机制,指导药物设计和优化,提高新药物的研发成功率。

3 图神经网络在药物-靶标预测领域中的应用

3.1 图神经网络应用于药物-靶标相互作用预测 药物-靶标相互作用预测基于现有的海量组学数据,预测出未知的且具有潜在相互作用的药物-靶标对。将图神经网络应用于药物-靶标相互作用领域,采用深度学习模型进行药物-靶标相互作用预测,克服了传统药物研发过程中耗时长、成本高等问题,提高筛选潜在候选药物和研发新药的效率^[15]。

Zhao T 等^[16]提出了一种基于图卷积网络和深度神经网络的药物-靶标相互作用预测模型“GCN-DTI”。首先,该模型基于已知的药物-靶标相互作用关系构建一个药物-靶标对网络,每个节点表示一个药物-靶标对,而节点之间的边则表示为两个药物-靶标对之间的连接强度。其次,使用 GCN 方法从药物-靶标对网络中提取每个药物-靶标对的特征表示。最后,将特征表示输入到深度神经网络来预测每个药物-靶标对之间存在相互作用的可能性。Li Y 等^[17]提出了一种基于多通道图神经网络的预测模型“DTI-MGNN”。该模型首先从药物、靶标蛋白、副作用和疾病之间相互作用的异质网络中获得药物和靶标蛋白在其独立空间中的表示。通过药物和蛋白的特征向量构建药物-蛋白对的特征向量,进一步构造药物-蛋白对的拓扑图和特征图。对于每一个药物-蛋白对,DTI-MGNN 模型使用两个独立的图注意力网络学习它的拓扑和特征表示,并与 GCN 得到的通用表示组合。最后,将组合好的药物-蛋白对的表示向量输入到多层感知机 (multilayer perceptron, MLP) 中,以此判定药物和蛋白之间是否存在相互作用关系。Jiang L 等^[18]提出了一种结合跨模态相似性的异构图注意力网络的药物-靶标预测模型“DTIH-NC”。该模型首先通过去噪自编码器获取药物、靶标、疾病和药物副作用的低维特征向量,并构建一个包含药物、靶标、疾病和药物副作用四类节点的异构网络。然后,模型采用异构图注意力网络模型来获取各个节点的特征向量表示。对于每个节点,模型根据它的一阶邻居节点和多阶邻居节点信息来更新该节点的特征信息,从四类节点之间的跨尺度关系中计算了药物和靶标的跨模态相似性,将异构图注意力网络得到的节点特征和跨模态相似性通过多层卷积神经网络进行深度整合,最终对药物和靶标之间相互作用进行评估。Zhang P 等^[19]提出了一种包含多层信息的 Transformer 网络的药物-靶标相互作用

预测模型“DeepMGT-DTI”。首先,该模型通过包含多层图信息的 Transformer 网络来获取药物分子结构的特征向量表示。其次,通过卷积神经网络来捕获靶标序列中局部残基信息,进而提取出靶标的特征向量表示。最后,通过提取的药物分子特征向量和靶标序列的特征向量来计算药物和靶标之间存在相互作用的可能性。Yu L 等^[20]提出了一种基于异构图神经网络的药物-靶标相互作用预测模型“HGDTI”。该模型首先获取了药物的分子指纹特征、靶标蛋白的伪氨基酸组成 (pseudo-amino acid composition, PseAAC) 特征、疾病的随机游走特征和药物副作用的随机游走特征。然后,将四种特征分别通过 Bi-LSTM 对每个节点特征进行编码。同时,利用注意力机制对异质邻居节点进行聚合以获得药物和靶标的嵌入特征。最终,将获得的药物和靶标的特征输入到两层的神经网络对药物-靶标相互作用的结果进行预测。Shao K 等^[21]提出了一种基于图卷积神经网络和图注意力网络的适用于异构图的预测模型“DTI-HETA”。该模型首先根据药物-药物相似度矩阵、靶标-靶标相似度矩阵和药物-靶标相互作用矩阵构建异构图。然后,通过图卷积神经网络和图注意力网络获得药物和靶标的节点嵌入特征。最后,该模型应用内积解码器来预测药物和靶标之间的相互作用。Wang W 等^[22]通过在异构网络上引入图卷积技术提出了“GCHN-DTI”模型。该模型整合了药物-靶标相互作用、药物-药物相互作用、药物相似性、靶标-靶标相互作用和靶标相似性的网络信息。在五种网络构成的异构网络中使用图卷积运算,获得药物和靶标的节点嵌入特征。该模型在图卷积层之间引入了注意力机制来融合每一层得到的节点特征。最后,根据融合后的药物和靶标节点嵌入特征来对预测药物和靶标之间是否存在相互作用。Li M 等^[23]将元路径聚合异构图神经网络应用在药物-靶标相互作用预测领域,提出了“MHGNN”模型。该模型通过元路径对异构图中的高阶关系进行建模。其次,该模型通过构建以药物-靶标对为节点的拓扑结构图,极大地丰富了药物-靶标对之间的高阶相关性,解决了目前模型中只能聚合图结构中直接邻居节点信息的弊端,增强了异构图中的高阶语义学习。Veleiro U 等^[24]提出了一种基于图神经网络的药物-靶标相互作用推理方法“GeNNius”。该模型由一个图神经网络编码器和一个基于神经网络的分类器组成,通过图神

神经网络编码器将药物、靶标的矩阵分别转换成指定嵌入维度的药物特征矩阵和靶标特征矩阵。将获得的对应药物特征与靶标特征同时输入到分类器中,预测最终的相互作用结果。

3.2 图神经网络应用于药物-靶标亲和力预测 将图神经网络融入到药物-靶标亲和力预测领域,可加深对药物与靶标相互作用强度的深入理解,指导药物研发者对药物结构和剂量进行优化,提高药物的活性、安全性和耐受性,提升药物的疗效并减少药物的不良反应^[25]。

Tian Q 等^[26]提出了一种基于循环神经网络和图神经网络的药物-靶标亲和力预测模型。该模型将 GraphDTA 模型^[27]升级为三通道模型。对于药物,该模型使用了 GraphDTA 模型中使用四个变异网络提取药物拓扑特征的方法,使用 BiGRU^[28]提取了药物序列中部分原子的局部化学背景。对于靶标,该模型将靶标蛋白的序列解释为时间序列并用 LSTM 网络提取其特征表示。最后,将药物和靶标向量输入到两层的全连接网络中获得可供参考的药物-靶标亲和力值。Zhao Q 等^[29]将注意力机制应用到药物-靶标相互作用预测中,提出了“AttentionDTA”模型。它使用两个独立的一维卷积神经网络从药物的 SMILES 序列和靶标蛋白的氨基酸序列中分别提取各自的语义特征信息,还嵌入了一种双侧多头注意力机制以此来探索药物特征信息与靶标特征信息之间的关系。该模型将靶标和药物的特征向量输入到全连接层,得到药物和靶标之间的亲和力得分。Tran HNT 等^[30]提出了“DeepNC”模型。该模型使用了 GENConv、GCNConv 和 HypergraphConv 三种图神经网络算法构建框架,并在两个基准数据集(Davis、Kiba)和一个独立提出的数据集(Allergy)上对模型进行性能评估。该模型使用 GNN 和一维卷积网络层来分别学习药物和靶标的特征向量信息,将药物特征向量和靶标特征向量输入到全连接层中获得结合亲和力值。Jiang M 等^[31]提出了一种基于序列的药物-靶标亲和力预测模型“WGNN-DTA”。该模型首先提取加权蛋白质图和药物分子图。然后,利用 GCN 和 GAT 来提取靶标蛋白和药物分子图的潜在特征向量。最后,将提取到的靶标潜在向量和药物潜在向量输入到全连接层中实现亲和力值的预测。Wang S 等^[32]使用图神经网络和多尺度拓扑特征融合技术,提出了“MSGNN-DTA”模型。该模型将门控

跳跃连接机制引入到了药物和靶标特征学习阶段,得到信息丰富的药物和靶标蛋白的特征表示。对于每种药物,该模型同时构建了原子拓扑图和基序拓扑图。该模型使用了蛋白质结构预测模型来获取该蛋白的残基接触图,同时构建了加权的蛋白拓扑图。之后,将三种拓扑图分别输入到 GNN 中以获取药物和蛋白质的多尺度拓扑特征表示,通过注意力机制将药物的两种拓扑特征和蛋白质的一种拓扑特征进行融合得到联合特征。最后,该模型将联合特征输入到多个全连接层中以预测药物和靶标之间的亲和力的值。

4 总结与展望

随着计算机技术的不断发展和生物学数据库的不断扩大,越来越多的图神经网络模型被应用到药物-靶标预测领域中,以期加快药物研发的进程,提高药物研发的成功率。虽然大多数模型在某个方面或者某些方面存在着一定优势,但也存在着一些局限,需要不断优化升级。随着数据库的不断扩充及算法的不断优化,有关药物-靶标预测的图神经网络模型的预测结果也会越来越精确,效率也会逐渐提高。

将图神经网络模型应用于药物-靶标预测领域具有重要的应用价值和广阔的应用前景。随着大规模药物-靶标数据的积累和图神经网络模型的不断发展,图神经网络模型将进一步提高药物-靶标预测的准确性和可解释性,加深对药物的作用机制和靶标的生物功能的理解。此外,图神经网络还将加速药物研发的进程。总之,通过对药物-靶标相互作用和亲和力预测,图神经网络模型必将成为指导药物筛选、设计和优化的重要辅助工具之一,为新药的研发以及疾病的治疗提供强大的支持和帮助。

参考文献:

- [1]Gupta R,Srivastava D,Sahu M,et al.Artificial intelligence to deep learning: machine intelligence approach for drug discovery [J].Mol Divers,2021,25(3):1315-1360.
- [2]Ezzat A,Wu M,Li XL,et al.Computational prediction of drug-target interactions using chemogenomic approaches: an empirical survey[J].Brief Bioinform,2019,20(4):1337-1357.
- [3]Abbasi K,Razzaghi P,Poso A,et al.Deep Learning in Drug Target Interaction Prediction: Current and Future Perspectives [J].Curr Med Chem,2021,28(11):2100-2113.
- [4]Ru X,Ye X,Sakurai T,et al.Current status and future prospects of drug-target interaction prediction [J].Brief Funct

- Genomics,2021,20(5):312–322.
- [5]Zhang XM,Liang L,Liu L,et al.Graph Neural Networks and Their Current Applications in Bioinformatics [J].Front Genet, 2021,12:690049.
- [6]Wu Z,Pan S,Chen F,et al.A Comprehensive Survey on Graph Neural Networks[J].IEEE Trans Neural Netw Learn Syst, 2021,32(1):4–24.
- [7]Lin X,Quan Z,Wang ZJ,et al.Effectively Identifying Compound–Protein Interaction Using Graph Neural Representation [J].IEEE/ACM Trans Comput Biol Bioinform,2023,20(2):932–943.
- [8]Kipf TN,Welling M.Semi–Supervised Classification with Graph Convolutional Networks[C]//ICLR.2016.
- [9]Zhang Y,Lei X,Pan Y,et al.Drug Repositioning with GraphSAGE and Clustering Constraints Based on Drug and Disease Networks[J].Front Pharmacol,2022,13:872785.
- [10]Velikovi P,Cucurull G,Casanova A,et al.Graph Attention Networks [C]//Proceedings of the International Conference on Learning Representations,2018.
- [11]Xu K,Hu W,Leskevci J,et al.How Powerful are Graph Neural Networks?[C]//ICLR.2018.
- [12]Liu S,An J,Zhao J,et al.Drug–Target Interaction Prediction Based on Multisource Information Weighted Fusion[J].Contrast Media Mol Imaging,2021,2021:6044256.
- [13]Zeng Y,Chen X,Luo Y,et al.Deep drug–target binding affinity prediction with multiple attention blocks [J].Brief Bioinform,2021,22(5):bbab117.
- [14]Ren YS,Li HL,Piao XH,et al.Drug affinity responsive target stability (DARTS) accelerated small molecules target discovery: Principles and application [J].Biochem Pharmacol,2021,194: 114798.
- [15]Monteiro NRC,Ribeiro B,Arrais JP.Drug–Target Interaction Prediction: End–to–End Deep Learning Approach [J].IEEE/ACM Trans Comput Biol Bioinform,2021,18 (6):2364–2374.
- [16]Zhao T,Hu Y,Valsdottir LR,et al.Identifying drug–target interactions based on graph convolutional network and deep neural network[J].Brief Bioinform,2021,22(2):2141–2150.
- [17]Li Y,Qiao G,Wang K,et al.Drug–target interaction prediction via multi–channel graph neural networks [J].Brief Bioinform,2022,23(1):bbab346.
- [18]Jiang L,Sun J,Wang Y,et al.Identifying drug–target interactions via heterogeneous graph attention networks combined with cross–modal similarities [J].Brief Bioinform,2022,23 (2): bbac016.
- [19]Zhang P,Wei Z,Che C,et al.DeepMGT–DTI: Transformer network incorporating multilayer graph information for Drug–Target interaction prediction [J].Comput Biol Med,2022,142: 105214.
- [20]Yu L,Qiu W,Lin W,et al.HGDTI: predicting drug–target interaction by using information aggregation based on heterogeneous graph neural network [J].BMC Bioinformatics,2022,23(1): 126.
- [21]Shao K,Zhang Y,Wen Y,et al.DTI–HETA: prediction of drug–target interactions based on GCN and GAT on heterogeneous graph[J].Brief Bioinform,2022,23(3):bbac109.
- [22]Wang W,Liang S,Yu M,et al.GCHN–DTI: Predicting drug–target interactions by graph convolution on heterogeneous networks[J].Methods,2022,206:101–107.
- [23]Li M,Cai X, Xu S,et al.Metapath–aggregated heterogeneous graph neural network for drug–target interaction prediction [J]. Brief Bioinform,2023,24(1):bbac578.
- [24]Veleiro U,de la Fuente J,Serrano G,et al.GeNNius: an ultrafast drug–target interaction inference method based on graph neural networks[J].Bioinformatics,2024,40(1):btad774.
- [25]Lin S,Shi C,Chen J.GeneralizedDTA: combining pre–training and multi–task learning to predict drug–target binding affinity for unknown drug discovery [J].BMC Bioinformatics,2022,23 (1):367.
- [26]Tian Q,Ding M,Yang H,et al.Predicting Drug–Target Affinity Based on Recurrent Neural Networks and Graph Convolutional Neural Networks [J].Comb Chem High Throughput Screen,2022,25(4):634–641.
- [27]Nguyen T,Le H,Quinn TP,et al.GraphDTA: predicting drug–target binding affinity with graph neural networks [J]. Bioinformatics,2021,37(8):1140–1147.
- [28]Liu J,Lei X,Zhang Y,et al.The prediction of molecular toxicity based on BiGRU and GraphSAGE [J].Comput Biol Med, 2023,153:106524.
- [29]Zhao Q,Duan G,Yang M,et al.AttentionDTA: Drug–Target Binding Affinity Prediction by Sequence–Based Deep Learning With Attention Mechanism [J].IEEE/ACM Trans Comput Biol Bioinform,2023,20(2):852–863.
- [30]Tran HNT,Thomas JJ,Ahamed Hassain Malim NH.DeepNC: a framework for drug–target interaction prediction with graph neural networks[J].PeerJ,2022,10:e13163.
- [31]Jiang M,Wang S,Zhang S,et al.Sequence–based drug–target affinity prediction using weighted graph neural networks [J]. BMC Genomics,2022,23(1):449.
- [32]Wang S,Song X,Zhang Y,et al.MSGNN–DTA: Multi–Scale Topological Feature Fusion Based on Graph Neural Networks for Drug–Target Binding Affinity Prediction[J].Int J Mol Sci,2023,24(9):8326.

收稿日期:2024–02–25;修回日期:2024–03–06

编辑/成森