多层次双向融合分子特征的药物–靶标结合亲和力预测

曾文亮，张龙信*

湖南工业大学计算机学院，湖南株洲 412007

摘要： 传统药物开发需要经历一个漫长的过程。通过计算机准确预测药物与靶标蛋白质的结合亲和力（Drug-target binding affinity, DTA）能够极大地加快药物开发进程。准确预测 DTA 的关键在于如何准确地挖掘出药物和靶标的潜在特征。为解决此问题，本文提出了一个基于多通道结构融合双向靶标特征的 DTA 预测模型 MBDTA（Multi-layer Bidirectional Drug-Target Binding Affinity），该模型利用药物与靶标的深层特征表达预测 DTA。MBDTA 分为三个步骤：首先，通过标签编码得到药物分子和靶标的初始特征；其次将初始特征分别送入图神经网络模块和递归神经网络模块学习其中的潜在特征；最后，将两组潜在特征整合后利用全连接层预测 DTA。在 KIBA 数据集的实验结果表明，MBDTA 在三个指标（*MSE*、*CI* 和 r_m^2）上的性能比当前最先进的 DTA 预测模型平均提升了 32.42%、3.84%和 23.64%。

关键词： 药物-靶标结合亲和力；图神经网络；深度学习；多阶邻域特征

DOI: 10.57237/j.cst.2023.04.009

Drug-target Binding Affinity Prediction by Multi-layer Bidirectional Fusion Molecular Feature

Zeng Wenliang, Zhang Longxin*

College of Computer Science, Hunan University of Technology, Zhuzhou 412007, China

Abstract: Traditional drug development requires a long process. Accurate prediction of drug-target binding affinity (DTA) by computer can greatly accelerate the drug development process. The key to predicting DTA is how to accurately mine the potential features of drugs and targets. To solve this problem, this paper proposes a DTA prediction model based on a multi-layer structure fused with bidirectional target features (MBDTA). DTA is predicted by this model through exploiting the representation of deep features of the drug-target pairs. MBDTA is split into three steps: Firstly, the initial features are obtained by encoding the drug molecules and targets through label encoding; Secondly, the initial features are fed into the graph neural network module and the recurrent neural network module to learn the potential features in them, respectively; Finally, the two sets of potential features are integrated and then the DTA is predicted exploiting a fully connected layer. Experimental results on the KIBA dataset show that MBDTA improves performance by an average of 32.42%, 3.84%, and 23.64% on the three metrics, *MSE*, *CI*, and r_m^2, compared to the current state-of-the-art DTA prediction model.

基金项目：湖南省自然科学基金 (2023JJ50204)；湖南省教育厅科学研究项目 (23B0560).

*通信作者：张龙信, longxin.zhang@163.com

收稿日期: 2023-11-15；接受日期: 2023-12-13；在线出版日期: 2023-12-27 http://www.computscitech.com

Keywords: Drug-target Binding Affinity; Graph Neural Networks; Deep Learning; Multi-order Neighborhood Features

1 引言

传统药物开发是一个漫长的过程，需要耗费大量的时间进行药物实验并不断测试。据统计，一种药物的开发过程平均需要花费 13 年左右，成本投入在 6.18-26 亿美元之间[1]。大多数药物都是通过与特定靶标分子如酶、核受体、G-蛋白偶联受体和离子通道[2]的体内相互作用达到治疗效果。因此，准确预测药物-靶标相互作用（Drug-Target Interactions, DTI）能够大大缩短药物开发过程耗费的时间。DTI 通常以两者之间的结合亲和力值作为度量，因此 DTI 预测也称为药物-靶标结合亲和力（Drug-Target Binding Affinity, DTA）预测。由于一款新药研发过程代价昂贵，且失败概率极高[3]。因此，大量的研究人员希望从现有药物中发现潜在的 DTI，从而加速药物的研发过程，即药物重定位。药物重定位是为了从现有药物中识别新药理适应症。它可以将新开发药物用于治疗除药物的预期治疗用途以外的疾病，同时也能够为已知药物建立新的治疗用途[4]。例如，阿司匹林在研发之初被用于治疗各种疼痛和炎症。它通过抑制血小板（抗血小板药）的正常功能来抑制血液凝结，形成血凝块。通过大量临床实验进行不断试验，如今的阿司匹林已经成为预防心血管疾病的基础药物之一，并且能够用于治疗心脏病和中风。这说明药物重定位是药物开发的一种有效手段。

分子对接和机器学习模型长期以来一直是药物重定位的重要工具[5]。分子对接方法通过对药物与靶标蛋白质的三维结构进行建模，预测两者之间的相互作用关系。由于部分生物大分子的三维结构数据难以获取[6]，此类方法无法处理大批量的药物组学数据。传统的机器学习模型广泛应用于定量构效活性关系（Quantitative Structure Activity Relationship, QSAR）以及蛋白质化学计量学等领域。这些领域中的数据集存在正负样本高度不平衡的问题，因此预测精度并不理想。尽管研究人员尝试使用卷积神经网络（Convolutional Neural Network, CNN）[7]提高 DTI 预测的准确性，但是 CNN 的平移不变性无法处理药物分子图和药物-靶标相互作用网络这类非欧式空间数据。因此，预测准确率不如传统机器学习模型[8]。

近年来，图神经网络（Graph Neural Network, GNN）因其在非欧氏空间数据上具有强大的学习能力而受到广泛关注。因此，研究人员逐渐将 GNN 模型应用在药物-靶标相互作用关系的预测任务中。Shao 等人[9]将 DTI 预测任务视为链路预测问题，并提出了基于异质信息的图神经网络方法 DTI-HETA。尽管 GNN 能够解决非欧式结构数据难以处理的问题[10]，但在实际应用过程中，GNN 还存在以下三个缺点：(1) 当节点位于图中不同位置并且节点邻域中的节点拓扑结构相似或相同时，GNN 聚合邻域节点特征并更新节点信息后，将得到相同的节点嵌入表达。尽管部分图数据可以采用堆叠网络层数的方式区分拓扑结构相似或相同的子图，但是对于药物分子图而言，即使网络层数迭代多次也无法区分这类子图。(2) 随着网络层数堆叠过深，模型会出现梯度消失、过拟合或过度平滑等问题[11]。(3) 当 GNN 层数过少时，模型无法有效地提取到分子图的局部拓扑结构特征。当前已有部分工作对基于 GNN 的 DTA 预测模型模型对以上三个确定进行改进。例如，DeepNC [12]采用可自行学习聚合参数的图神经网络模型对药物分子图的潜在特征学习，在 DTA 预测任务中达到了一定的效果。然而，模型的预测精度不够理想。

针对上述问题，本文提出了一种基于多通道结构融合双向靶标特征的 DTA 预测模型 MBDTA（Multi-channel Bi-directional Drug-Target Binding Affinity）。MBDTA 的核心模块为多层通用汇聚网络（Multi-layer General Aggregation Network, MLGEN）模块、双向长短周期记忆（Bidirectional Long Short-Term Memory, BiLSTM）[13]模块和多层感知机（Multilayer Perceptron, MLP）模块。MLGEN 模块利用多层次特征融合机制整合原子多阶邻居节点的特征表达。BiLSTM 模块分别从靶标序列的前向与后向对序列进行建模，充分利用长文本序列数据中的依赖关系，提取靶标序列的潜在特征表达。MLP 模块将 MLGEN 模块和 BiLSTM 模块捕获的特征拼接后对 DTA 进行预测。本文的主要贡献如下：

(1). 提出一种基于多阶邻域的节点特征提取技术。

为了提高 DTA 预测任务中结合亲和力的预测精度，本文基于药物分子的原子节点的多阶邻域构建 MLGEN 模块，缓解在图卷积过程中由于子图的拓扑结构相同或相似引起的预测精度下降问题；

(2). 使用 BiLSTM [13]对模型进行构建与优化。BiLSTM 对靶标氨基酸序列的前向和后向分别进行建模，综合考虑序列的时序特征，获取靶标的深层次特征信息；

(3). 大量的实验证明，在 MSE 标准下，MBDTA 在 Davis 和 KIBA 数据集上比与当前最先进的 DTA 预测模型在性能上平均提高了 15.83%和 32.42%。

本文第 2 节对国内外在 DTA 预测任务的相关工作进行总结；第 3 节给出 MBDTA 的模型结构，并对其进行详细介绍；第 4 节为实验结果与对比分析；第 5 节总结全文并对未来的研究工作进行展望。

2 相关工作

DTA 预测已经成为药物重定位中的一个重要问题。在过去十几年的研究当中，众多学者基于机器学习方法提出多种 DTA 预测模型。例如，PUCPI [14]将蛋白质结构域和化合物子结构分别作为蛋白质特征和化合物特征，然后计算两组特征的张量积作为化合物-蛋白质对特征，使用偏置支持向量机对其进行分类训练。Tapio 等人[15]基于随机森林模型提出了 KronRLS 方法。KronRLS 通过 Kronecker 正则最小二乘算法预测药物-靶标对结合亲和力。为了弥补 KronRLS 中存在的线性依赖问题，SimBoost [16]采用梯度增强方法构建药物和靶标之间的相似性特征网络，以提高预测性能。尽管机器学习算法在 DTA 预测任务中表现出合适的性能，但是这类算法往往使用精心设计的手工特征，而这些手工特征通常需要特别的专业知识与经验将其组合在一起[17]。

由于深度学习在多个领域取得巨大成功，越来越多的人把注意力转移到将深度学习技术应用在预测 DTA 的问题上。现有的多数深度学习模型都是基于拓扑相似性对结合亲和力值进行预测。例如，DeepDTIs [18]通过堆叠受限玻尔兹曼机器构造了一个深度信任网络模型预测药物-靶标对的相互作用关系。Wan 等人[19]提出了一个非线性的端对端深度学习模型 NeoDTI。NeoDTI 通过聚合异构网络数据信息，自动学习药物和靶标的拓扑结构表示提高模型的预测性能。由于 CNN 在计算机视觉领域取得的巨大成功[20, 21]，部分研究人员将 CNN 应用于 DTA 预测问题。DeepDTA [22]使用两个 CNN 构件，分别对药物分子的 SMILES 序列和蛋白质的氨基酸序列进行学习，并与深度神经网络结合预测 DTA。随着 GNN 近年来在图结构数据任务上展现出其优越的性能，研究人员也开始利用 GNN 模型对 DTA 进行预测。现有的 GNN 模型主要分为两种：基于交互网络的 DTA 预测模型和基于结构的 DTA 预测模型。前者将 DTA 预测任务视为二分类任务，基于药物与靶标的多种相互作用网络对两者之间是否存在相互作用关系进行预测，存在相互作用关系时输出为 1，否则为 0。例如，Peng 等人[23]通过分析并组合多种异构网络图构建一个新的异构网络，提出一种基于异构图卷积神经网络的方法学习其中的特征表示，采用端对端的方式对 DTI 进行预测。后者将其视为回归任务，基于药物分子结构和靶标氨基酸序列对 DTA 进行预测，输出药物-靶标对的相互作用强度值，即结合亲和力值。例如，GraphDTA [24]使用 RDKit [25]构建分子图并提取原子特征，并通过 DeepChem 方法[26]描述节点特征，如原子符号、氢原子总数和原子的隐含值等。GraphDTA 采用 GNN 模型提取药物分子特征，使用 CNN 方法获取靶标序列特征。DeepMGT-DTI [27]通过使用 MCGCN 模型整合药物分子的结构信息，并通过 Transformer 网络[28]融合多层图特征丰富药物的潜在特征，在 DrugBank 数据库[29]中提取的数据集上取得了一定的成效。DeepNC [30]通过组合三种不同的 GNN 算法学习药物特征，随后与靶标潜在特征结合后送入全连接层预测 DTA。然而，GNN 难以区分具有相似拓扑结构的子图的问题始终限制着上述模型在 DTA 预测任务中的性能。Welling [31]等人尝试使用 One-Hot 编码计算节点之间的距离，但是这种方式高度依赖全图的结构信息，无法在具有新节点的图结构数据上进行迁移学习。

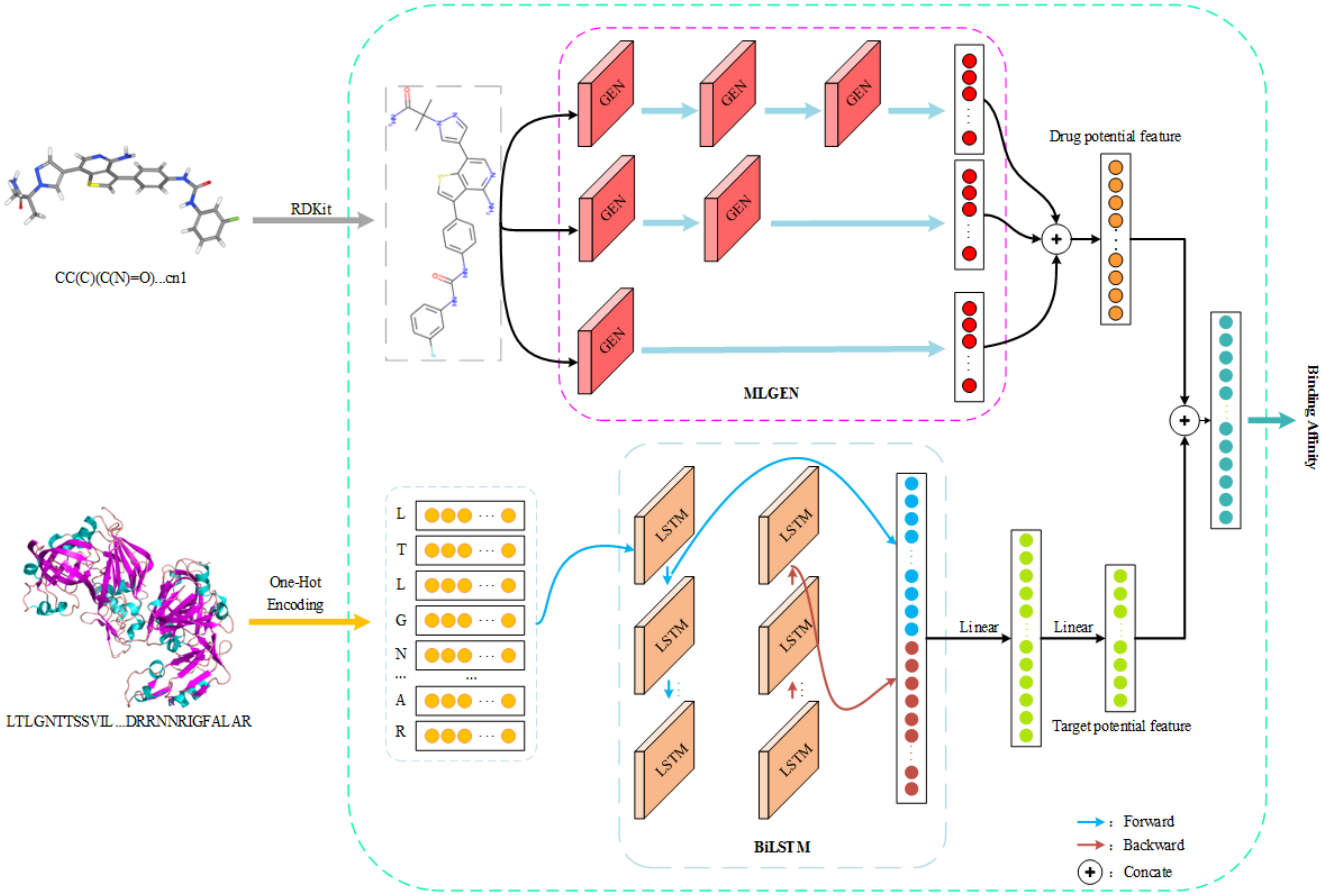


图 1 MBDTA 整体框架

综上，现有的 DTA 预测模型仍存在以下问题：(1) 传统机器学习模型依赖于手工特征，在没有专业知识的引导下，机器学习模型无法准确预测药物-靶标对的相互作用强度；(2) 基于 CNN 的 DTA 预测模型能自动提取药物-靶标对的特征，但是对于图结构数据，CNN 的特征提取性能始终受限；(3) 基于 GNN 的 DTA 预测模型在一定程度上克服了图结构数据难以处理的问题，但预测准确度还需继续提升，尤其是面对具有相似或相同拓扑结构的子图的图数据时。针对以上问题，本文从实际应用的角度出发，提出了高效的特征提取模型，并提高了结合亲和力的预测准确率。

3 MBDTA

MBDTA 整体框架如图 1 所示。MBDTA 以药物分子的 SMILES 序列和靶标蛋白质的氨基酸序列作为输入，对两者之间的结合亲和力进行预测并输出。MBDTA 包含三个模块：特征编码模块、潜在特征提取模块以及 DTA 预测模块。

3.1 特征编码模块

MBDTA 模型将分别对药物分子和靶标蛋白质的特征进行编码，其中药物特征编码模块的输入是药物分子的 SMILES 序列，靶标特征编码模块的输入是靶标的氨基酸序列，这两个模块的输出分别为药物和靶标的初始特征矩阵。

3.1.1 药物特征编码

药物分子通常以具有价键和原子的结构图表示。在 DTA 预测的计算过程中，处理药物分子的 3D 结构图需要耗费大量的计算资源。因此，MBDTA 采用 SMILES 序列表示药物分子结构。SMILES 序列采用 ASCII 字符将分子的 3-D 化学结构转换为一串字符，使其可以被计算机读取。本文使用 $S_{drug}=\{D_1,D_2,...D_{N_d}\}$ 表示药物分子 SMILES 字符串集合，其中 D_k 表示第 k 个药物分子的 SMILES 序列表达，N_d 表示药物分子总数。通过使用开源工具 RDKit，

药物特征编码模块会将 D_k 转换为相应的分子图 $G_k=(V_k,E_k)$，其中 $v_i \in V_k(i=1,2,\ldots N_{atom}$ 表示分子的第 i 个原子节点，$e_{i,j} \in E_k(j \in \{1,2,...,N_N\})$ 表示原子节点 v_i 与 v_j 之间存在的价键。在药物分子图中，MBDTA 不考虑化学键的强度，因此 G_k 为无向图。N_{atom} 和 N_N 分别代表原子数量和节点 v_i 的一阶邻居节点数量。为了描述节点特性，MBDTA 采用 DeepChem 方法将每个节点表示为一个 78 维的二进制特征向量。特征向量共包含五条信息：原子符号 $\mathbf{h}_{sym}$、相邻原子数量 $\mathbf{h}_{nei}$、相邻氢原子数量 $\mathbf{h}_H$、原子隐含价态 $\mathbf{h}_{hid}$ 以及原子是否属于芳香烃结构 $\mathbf{h}_{isaro}$。通过对 v_i 采用独热编码可以得到第 k 个药物分子中每个原子节点的初始化特征表示，计算过程如公式(1)：

$$\mathbf{h}_{v_i}=[\mathbf{h}_{sym} \| \mathbf{h}_{nei} \| \mathbf{h}_H \| \mathbf{h}_{hid} \| \mathbf{h}_{isaro}],\ \mathbf{h}_{v_i} \in \mathbb{R}^{1\times N_{fea}} \quad (1)$$

其中 $\|$ 表示向量级联，N_{fea} 表示特征向量维度，大小为 78。此时，药物分子 D_k 的初始特征 $\mathbf{H}_{D_k}$ 表示为公式(2)：

$$H_{D_k}=[h_{v_1},h_{v_2},...,h_{N_{atom}}]^T.\ H_{D_k} \in \mathbb{R}^{N_{atom}\times N_{fea}} \quad (2)$$

对原子节点关联结构特征，MBDTA 采用邻接矩阵 $A_{D_k} \in \mathbb{R}^{N_{atom}\times N_{atom}}$ 表示。当原子节点 v_i 与 v_j 之间存在价键连接时，$A_{D_k}(i,j)=A_{D_k}(j,i)=1$，否则为 0。为了准确预测 DTA，了解药物分子图中不同节点之间的相互作用关系是十分重要的。此时的邻接矩阵 $\mathbf{A}_{D_k}$ 仅表示节点与其一阶邻居节点的连通性。然而，由于药物分子通常是由多个原子和不同价键组成的复杂结构，仅聚合一阶邻域信息无法深入挖掘原子局部结构特征。因此，特征编码模块通过对节点的多阶邻域特征进行编码，得到二阶邻接矩阵和三阶邻接矩阵，计算公式如式(3)、(4)所示。

$$\mathbf{A}^2_{D_k}=\mathbf{A}_{D_k}\times\mathbf{A}_{D_k}, \quad (3)$$

$$\mathbf{A}^3_{D_k}=\mathbf{A}_{D_k}\times\mathbf{A}_{D_k}\times\mathbf{A}_{D_k}. \quad (4)$$

$\mathbf{A}^n_{D_k}(i,j)$ 表示分子图中节点 v_i 到 v_j 的距离为 n 的边的数量。通过这种方式，MBDTA 能够聚合原子节点的多阶邻域特征，捕获节点之间的多重连通性关系。

3.1.2 靶标特征编码

蛋白质是由多个氨基酸通过肽键相连而成的具有一定结构的生物大分子物质。由于使用分子图结构表示蛋白质分子比较困难，可靠性较低。因此，蛋白质分子通常以多个 ASCII 字符组成的序列进行表示，序列中的每个字符代表一个氨基酸（比如 A 代表丙氨酸，R 代表精氨酸）。本节使用 $S_{pro}=\{T_1,T_2,...T_{N_t}\}$ 靶标序列集合，其中 T_l 表示第 l 个靶标的氨基酸字符序列，N_t 表示靶标分子总数。组成生命体中蛋白质的氨基酸总共有二十种，每种氨基酸都对应字母表中的一个字母。MBDTA 将靶标的特征矩阵维度设置为 1000。当靶标序列长度不足 1000 时，则用 0 补齐；序列长度超过 1000 则进行截断。靶标特征编码模块使用字母顺序对 T_l 进行标签编码得到靶标的初始分子特征表达 $\mathbf{h}_l=[x_1,x_2,...,x_{1000}]$。随后通过公式(5)将初始分子特征表达转换为靶标初始特征：

$$\mathbf{H}_l=Embedding(\mathbf{h}_l), \quad (5)$$

$\mathbf{H}_l \in \mathbb{R}^{1\times 1000\times 128}$ 表示第 l 个靶标分子的初始特征。每个靶标通过初始化编码函数 $Embedding(\cdot)$ 被映射到一个 128 维的向量空间中。

3.2 特征提取模块

MBDTA 的特征提取模块分为两个部分：MLGEN 模块和 BiLSTM 模块。

3.2.1 MLGEN 模块

由于传统的 GNN 模型随着网络的不断加深会产生梯度消失、过度平滑以及过拟合问题。MLGEN 对节点的多阶邻域进行编码从而减少 GNN 层数。同时，MLGEN 使用能够训练更深层网络的通用聚合网络（General Aggregation Network, GEN）组成提取不同邻域特征的三个通道，以多通道方式提取多阶邻域中的节点特征，加深分子网络图层次。

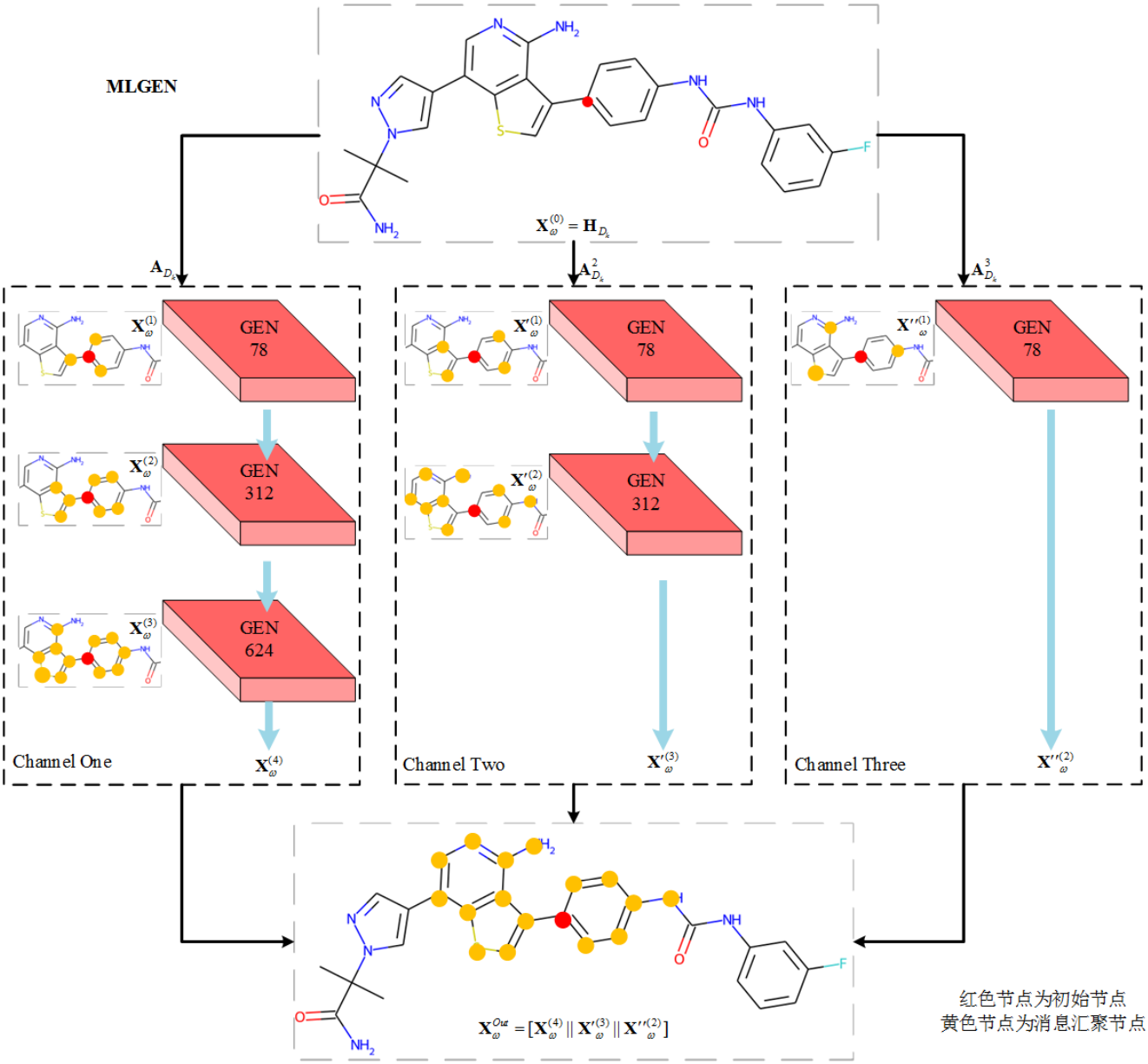


图 2 MLGEN 模块图

GEN 由 Li 等人[32]在 2020 年提出，它对基于谱域的图卷积网络（Graph Convolutional Networks, GCN）的聚合函数进行改进并提出了图跳跃连接和图归一化层增强 GCN 的特征提取能力。GCN 模型对 l 层的节点进行消息传递的过程如公式(6-8)所示：

$$\mathbf{m}_{vu}^{(l)} = \alpha^{(l)}(\mathbf{h}_v^{(l)}, \mathbf{h}_u^{(l)}, \mathbf{h}_{e_{vu}}^{(l)}),\ u \in N(v), \tag{6}$$

$$\mathbf{m}_v^{(l)} = \beta^{(l)}(\{\mathbf{m}_{vu}^{(l)} \mid u \in N(v)\}), \tag{7}$$

$$\mathbf{h}_v^{(l+1)} = \delta^{(l)}(\mathbf{h}_v^{(l)}, \mathbf{m}_v^{(l)}), \tag{8}$$

其中 $\alpha^{(l)}$、$\beta^{(l)}$ 和 $\delta^{(l)}$ 分别为 l 层节点的消息构造函数、消息聚合函数以及顶点更新函数。为简单起见，本节仅分析每一层顶点特征更新的情况。消息构造函数 $\alpha^{(l)}$ 通过 l 层节点 v 的顶点特征 $\mathbf{h}_v^{(l)}$、邻居节点 u 的顶点特征 $\mathbf{h}_u^{(l)}$ 以及对应的边属性特征 $\mathbf{h}_{e_{vu}}^{(l)}$ 针对每个邻居节点 $u \in N(v)$ 构造单独的节点信息 $\mathbf{m}_{vu}^{(l)}$。随后，消息聚合函数会对节点信息进行聚合，从而得到 v 的聚合邻居特征 $\mathbf{m}_v^{(l)}$。最后，顶点更新函数通过综合顶点特征以及聚合邻居特征更新下一层的顶点特征 $\mathbf{h}_v^{(l+1)}$。由于不同节

点的邻居节点数量不一致，采用单一的聚合函数无法综合考虑所有邻居节点的节点特征。因此，GEN 提出了两个可微的通用聚合函数 $SoftMax\_Agg_{\mu}(\cdot)$ 和 $PoweMean\_Agg_{\theta}(\cdot)$，其表达式分别如式(9)、(10)所示：

$$SoftMax\_Agg_{\mu}(\cdot)=\sum_{u\in N(v)}\frac{\exp(\mu\mathbf{m}_{vu})}{\sum_{i\in N(v)}\exp(\mu\mathbf{m}_{vi})}\cdot\mathbf{m}_{vu}, \quad (9)$$

$$PowerMean\_Agg_{\theta}(\cdot)=(\frac{1}{|N(v)|}\sum_{u\in N(v)}\mathbf{m}_{vu}^{\theta})^{1/\theta}, \ \theta\neq 0, \quad (10)$$

其中 μ 和 θ 均为可学习的连续变量，初始值为 1。通过引入可学习参数，聚合函数能够针对节点的每个邻居节点寻找到最佳聚合函数，从而聚合得到更有效的节点特征。

MLGEN 模块如图 2 所示，主要由三个通道组成，每个通道分别由不同数量的 GEN 模块构成。GEN 模块的特征通道数分别为 78、312 以及 624。每个通道分别汇聚节点不同邻域中的节点特征，捕获节点之间的连通性关系。本节将以药物 D_k 的特征矩阵 $\mathbf{H}_{D_k}$ 和特征矩阵 $\mathbf{A}_{D_k}$ 对 MLGEN 模块进行介绍。

MLGEN 的通道一由三层 GEN 模块构成，主要是用于聚合原子节点的一阶邻域特征。通道一的输入药物分子 D_k 的初始特征 $\mathbf{X}_{D_k}^{(0)}$ 和一阶邻接矩阵 $\mathbf{A}_{D_k}$。随后，通道一将对初始节点的一阶邻居节点、二阶邻居节点和三阶邻居节点的特征信息进行聚合得到节点特征信息 $\mathbf{X}_{D_k}^{(4)}$。MLGEN 的通道二和通道三分别由两层 GEN 模块和一层 GEN 模块组成。同样，MLGEN 的通道二和通道三通过将二阶邻接矩阵 $\mathbf{A}_{D_k}^{2}$ 和三阶邻接矩阵 $\mathbf{A}_{D_k}^{3}$ 作为输入分别得到了 $\mathbf{X}'^{(3)}_{D_k}$ 和 $\mathbf{X}''^{(2)}_{D_k}$。最后，MLGEN 将三个通道的输出进行级联得到 D_k 的潜在特征表达 $\mathbf{X}_{D_k}^{Out}\in\mathbb{R}^{1\times 1014}$。

3.2.2 BiLSTM 模块

BiLSTM 模块由两个长短期记忆网络（Long Short-Term Memory, LSTM）模块组成，分别从靶标序列的前向和后向对序列中的氨基酸时序特征进行提取。本节以靶标 T_l 的初始特征 $\mathbf{H}_l$ 对 BiLSTM 模块进行介绍。BiLSTM 模块如图 3 所示。在图 3 中，x_i 表示靶标序列中第 i 个位置的氨基酸。靶标的氨基酸序列特征将从两个方向分别送入 LSTM 模块，BiLSTM 模块会将两组特征整合后输出靶标的嵌入特征表达。

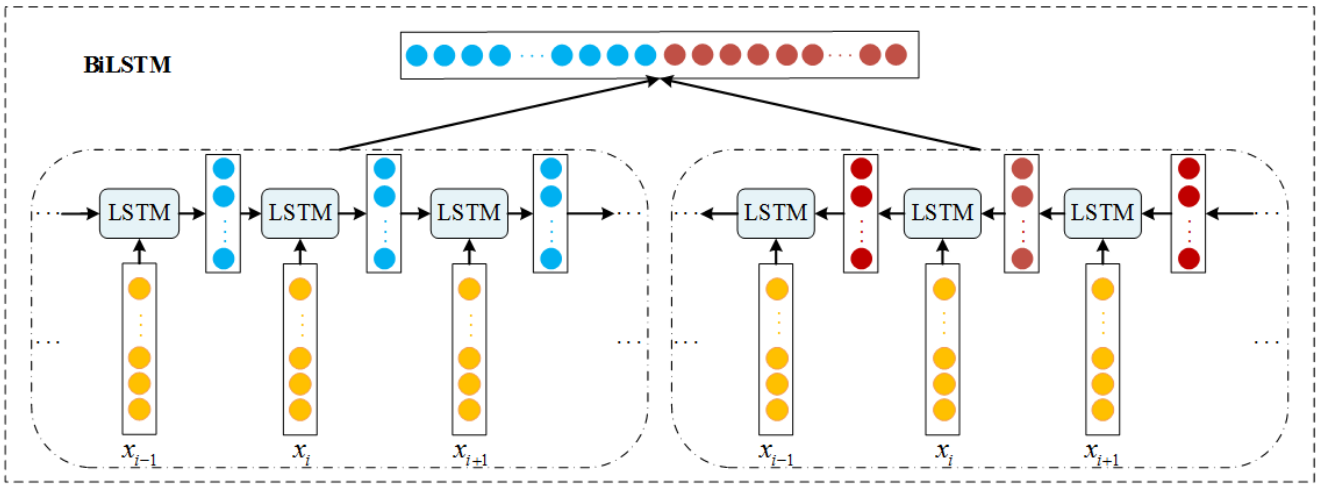


图 3 BiLSTM 模块

3.3 DTA 预测

MBDTA 将药物与靶标之间的关系预测视为回归任务，通过对药物与靶标之间的结合亲和力值进行预测从而得到两者之间的关系强度。MBDTA 将特征提取模块获取到的 $X_{D_k}^{Out}$ 和 $X_{T_l}^{Out}$ 进行整合后送入 MLP 网络进行 DTA 预测，整合后药物-靶标对特征为 $\mathbf{X}^{Out}=[\mathbf{X}_{D_k}^{Out}\,|\,\mathbf{X}_{T_l}^{Out}$。DTA 预测模块如图 4 所示。

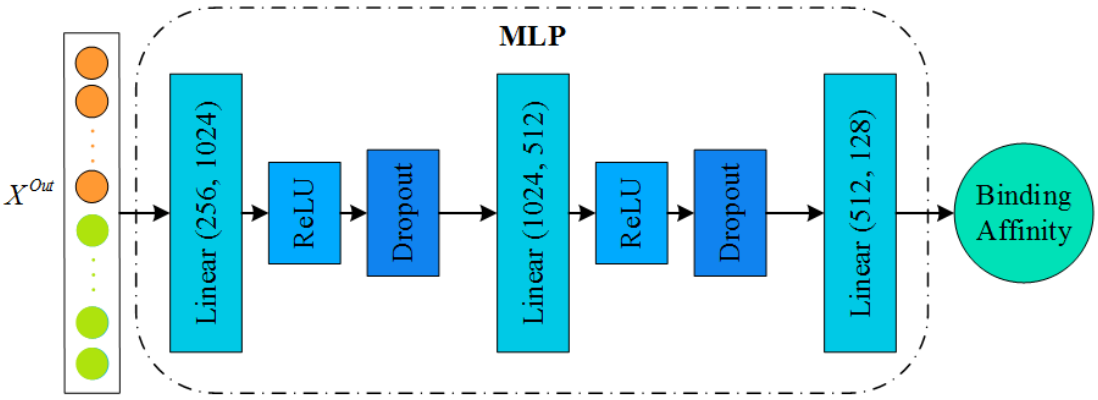


图 4 DTA 预测模块

药物-靶标对特征作为 MLP 的输入，经过前向传播后计算得到结合亲和力值作为 MLP 的输出。随后，MLP 通过对预测得到的结合亲和力值与真实结合亲和力值建立损失函数，使用反向传播对隐藏层进行更新。MLP 使用的损失函数如式(11)所示：

$$\mathcal{L}_{MSE}=\frac{1}{n}\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2, \tag{11}$$

其中，其中 Y_i 为真实结合亲和力值，$\hat{Y}_i$ 为预测结合亲和力值，n 为药物-靶标对数量。

4 实验分析

4.1 数据集

根据已有工作的实验设置，本文在实验部分选取了两个专用于 DTA 预测的数据集，分别为 Davis 和 KIBA 数据集。

(1) Davis

Davis 数据集来自 Davis 等人[33]对激酶蛋白质进行的选择性实验数据。Davis 数据集包含 442 种蛋白质和 68 种药物分子，并包含了这些蛋白质和药物全部的 30056 个相互作用对。药物与蛋白质的相互作用程度由解离常数 K_d 表示，K_d 反映的是化合物对靶标的亲和力大小，值越小时亲和力越强。由于 K_d 数值分布区间过大，He 等人[16]将其转换到对数空间，最终用于表示 Davis 药物-靶标对结合亲和力的 pK_d 数值分布在[5, 10.8]的区间内，转换公式如式(12)：

$$pK_d=\log_{10}(\frac{K_d}{1e^9}). \tag{12}$$

(2) KIBA

KIBA 数据集是由 Tang 等人[34]对药物靶标相互作用数据的不同来源统一整合而得到的数据集。初始的 KIBA 数据集包含 467 个蛋白质靶标，52498 个化学小分子，共计 246088 个药物-靶点相互作用对。由于该数据集中包含大量稀疏的相互作用关系，因此去除结合亲和力值低于 10 的药物和蛋白质，形成过滤后的数据集。KIBA 数据集中的蛋白质与药物的亲和力由 KIBA score 度量，其中 KIBA score 是由 IC_{50}（半抑制浓度），K_i（抑制常数）和 K_d 等信息整合得到的。

在训练过程中，MBDTA 将 Davis 数据集和 KIBA 数据集随机划分为六个不同的子集，并选取其中五个作为训练集，其余作为测试集。

4.2 评价指标

在训练过程中，MBDTA 将 Davis 数据集和 KIBA 数据集随机划分为六个不同的子集，并选取其中五个作为训练集，其余作为测试集。

本文采用回归任务中常用的三个指标对 MBDTA 模型的预测性能进行评估。它们包括均方误差（Mean-square Error, MSE）、一致性指数（Concordance Index, CI）和 QSAR 模型外部预测性能指标 r_m^2。

(1) *MSE*

均方误差为预测值和真实值的匹配程度，当匹配程度越大，*MSE* 值越小，模型预测效果越好。其计算公式为：

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2, \tag{13}$$

其中 Y_i 为真实值，$\hat{Y}_i$ 为预测值，n 为药物-靶标对

数量。

(2) *CI*

一致性指数用于衡量预测的结合亲和力值与真实值对应的药物-靶标对中的顺序排列是否一致，一致性指数越大说明模型性能更好。计算公式为：

$$CI = \frac{1}{N}\sum_{y_i > y_j} \xi(f_i - f_j), \tag{14}$$

其中 N 为归一化常数，数值上等于结合亲和力值不同的药物-靶标对数量。 f_i，f_j 分别代表结合亲和力指数 y_i 和 y_j（$y_i > y_j$）对应的预测得分。$\xi(x)$ 是一个跃迁函数，函数式如下：

$$\xi(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0.5, & \text{if } x = 0 \\ 0, & \text{if } x < 0. \end{cases} \tag{15}$$

(3) r_m^2

r_m^2 用于评估 QSAR 模型的外部预测性能。当且仅当 $r_m^2 \geq 0.5$ 时，才为可接受模型。r_m^2 越大，模型的预测性能越好。计算公式为：

$$r_m^2 = r^2 \times (1 - \sqrt{r^2 - r_0^2}), \tag{16}$$

其中 r^2 和 r_0^2 分别表示真实值与预测值之间有无截距的平方相关系数值。

4.3 基准模型

为验证 MBDTA 在 DTA 预测任务中的性能，本文将与以下六种模型进行比较分析。

(1). KronRLS [15]：KronRLS 基于正则化最小二乘模型对药物-靶标对的结合亲和力值进行预测，旨在最小化目标函数：

$$\mathcal{J}(f) = \sum_{i=1}^{m} (y_i - f(x_i))^2 + \lambda \| f \|_k^2, \tag{17}$$

其中 $x_i (i = 1, 2, ..., m)$ 是一组药物-靶标对特征，f 是非线性函数，y_i 表示它们对应的真实结合亲和力值。λ 为预定义的正则化参数并且 $\lambda > 0$，$\| f \|_k^2$ 表示符合核函数 k 的 f 范数。

(2). SimBoost [16]：SimBoost 通过定义三种类型的特征分别描述药物、靶标和药物-靶标对的性质，并训练了梯度提升树模型对药物与靶标之间的结合亲和力值进行预测。

(3). DeepDTA [22]：DeepDTA 使用药物分子和靶标的标签编码分别训练两个 3 层的 CNN 模型进行 DTA 预测。两个 CNN 模型分别从药物分子的 SMILES 字符串和靶标氨基酸序列学习其中的潜在特征，然后将两者进行连接后传递到全连接层预测药物-靶标对的结合亲和力值。

(4). GraphDTA [24]：GraphDTA 属于 DeepDTA 的拓展。它通过综合多种 GNN 模型取代 DeepDTA 中提取药物特征的 CNN 模型，将每张药物分子图视为图结构并对其中的潜在特征进行挖掘，最后与 CNN 模型学习到的靶标隐藏特征进行整合后预测药物-靶标对的结合亲和力指数。

(5). DeepGLSTM [35]：针对靶标氨基酸序列，DeepGLSTM 采用了在长文本序列中性能更优的 BiLSTM 模型；对药物分子图，DeepGLSTM 使用 GCN 方法充分挖掘分子图中的原子连通性关系。最后，通过将两组特征进行连接后预测 DTA。

(6). DeepGS [36]：DeepGS 由三个模块构成，分别为 CNN 模块、BiGRU 模块和 GAT 模块。CNN 模块和 BiGRU 模块分别对提取靶标和药物的一维字符序列特征，GAT 模块对药物的全局拓扑结构特征进行学习。

4.4 实验设置

实验基于 Ubuntu 系统进行训练，利用 4 张 GTX 1080 Ti 10 GB 显卡进行加速。基于深度学习框架 Pytorch 1.6.0 + CUDA 10.1，对 BiLSTM 网络和 MLP 神经网络进行构建。基于图深度学习库 Pytorch geometric 2.0.4 [37]对 GNN 模型进行构建和训练，利用 RDkit 2022.3.4 软件包提取数据集中的药物分子拓扑结构图。

MBDTA 采用 MLGEN 模块与 BiLSTM 模块分别提取药品与靶标的特征。MLGEN 模块中三个通道的 GEN 层数分别为 3、2、1，GEN 层中的聚合函数均使用可微聚合函数 softmax_Agg 对聚合参数进行学习。每一个 GEN 层后都使用 ReLu 激活函数去

除多余特征。MLP 层的隐藏层大小为 512，输入层和隐藏层后均应用 ReLU 激活函数和 Dropout 方法，Dropout Rate 为 0.2。在训练过程中，MBDTA 设置的 Batch_Size 大小为 128，并利用 Adam 优化器对模型参数进行学习，学习率为 0.0005。MBDTA 模型参数如表 1 所示。

表 1 MBDTA 模型参数

参数	设置
GEN 层数	[3, 2, 1]
GEN 聚合函数	softmax_Agg
优化器	Adam
学习率	0.0005
Epoch	100
Dropout Rate	0.2
MLP 隐藏层大小	512
Batch_Size	128
激活函数	ReLU

4.5 对比实验与分析

为了检验 MBDTA 在 DTA 预测任务上的高效性，本节将 MBDTA 与当前最先进的 DTA 预测模型（机器学习模型与深度学习模型）进行比较分析。表 2 中显示了在过滤后的 KIBA 数据集上不同模型的最佳 *MSE*、*CI* 和 r_m^2 得分，其中 KronRLS、SimBoost、DeepDTA 和 DeepGS 模型的结果均取自文献[36]。GraphDTA 模型仅考虑最优结果，即 GraphDTA (GIN)。

表 2 KIBA 数据集上不同模型的最佳 *MSE*，*CI*，r_m^2

Model	*MSE*	*CI*	r_m^2
KronRLS	0.411	0.782	0.342
SimBoost	0.222	0.836	0.629
DeepDTA	0.194	0.863	0.673
GraphDTA	0.251	0.808	0.631
DeepGLSTM	0.185	0.855	0.705
DeepGS	0.193	0.860	0.684
MBDTA	0.164	0.866	0.755

从表 2 中可以看出，深度学习模型的平均 *MSE* 得分、平均 *CI* 得分和平均 r_m^2 得分比机器学习模型分别提升了 37.63%、5.12%和 42.04%。这说明深度学习模型在 DTA 预测任务上总体来说是优于机器学习模型的。产生这一现象的原因在于：机器学习模型在 DTA 预测任务中很大程度上依赖手工生成的特征和药物与靶标的相似性矩阵。当特征完备性不足时，预测精度随之下降。然而，基于深度学习的模型能主动学习药物-靶标对的特征信息。在五种深度学习算法中 MBDTA 模型的 *MSE*、*CI* 和 r_m^2 得分别为：0.164、0.866 和 0.755，均优于其余四个模型。相比于其余深度学习模型，MBDTA 在 *MSE* 得分上平均降低了 20.29%，在 *CI* 和 r_m^2 得分上分别提高了 2.30%和 12.14%。在 *MSE* 得分上，MBDTA 相比于当前性能最好的深度学习模型 DeepGLSTM（0.185）降低了 11.35%；在 *CI* 得分上，比当前性能最好的模型 DeepDTA（0.863）提高了 0.03%；在 r_m^2 得分上，比当前性能最好的模型 DeepGLSTM（0.705）提高了 7.09%。原因在于：(1) DeepDTA、GraphDTA 和 DeepGS 仅对药物原子节点的单阶邻域特征进行聚合。在药物特征聚合过程中，模型难以处理具有相似图结构节点。对于靶标氨基酸序列，卷积神经网络无法有效获取这类时序数据的深层语义信息。(2) 尽管 DeepGLSTM 同样对节点的多阶邻域特征进行聚合，但是模型无法处理聚合过程中出现的大量重复节点信息。MBDTA 使用能够自动学习聚合参数的 GEN 构建多通道汇聚药物原子节点多阶邻域特征，并结合 BiLSTM 对氨基酸序列数据进行特征提取，综合考虑了药物和靶标的全局拓扑结构和局部化学信息，提高了 DTA 的预测准确率。

表 3 Davis 数据集上不同模型的最佳 *MSE*，*CI*，r_m^2

Model	*MSE*	*CI*	r_m^2
KronRLS	0.379	0.871	0.407
SimBoost	0.282	0.872	0.644
DeepDTA	0.261	0.878	0.630
GraphDTA	0.257	0.875	0.679
DeepGLSTM	0.294	0.867	0.624
DeepGS	0.252	0.882	0.686
MBDTA	0.242	0.879	0.697

为进一步验证 MBDTA 模型的高效性，本文在数据规模较小的 Davis 数据集上进行了同样的对比试验。表 3 展示了不同模型在 Davis 数据集上的结果，其中 KronRLS、SimBoost、DeepDTA 和 DeepGS 模型的结果同样取自文献[36]。

从表 3 中可以看出，在 Davis 数据集上，MBDTA 的 *MSE*、*CI* 和 r_m^2 得分分别为 0.242、0.879 和 0.697。

深度学习模型相比于机器学习模型而言，*MSE* 得分平均降低了 20.97%，*CI* 和 r_m^2 得分分别提升了 0.54%和 26.20%。这进一步证明了深度学习模型更适用于 DTA 预测任务。MBDTA 在 *MSE* 得分相较于当前性能最好的深度学习模型 DeepGS（0.252）降低了 3.97%，在 r_m^2 得分上比 DeepGS（0.686）提高了 1.60%。由于训练数据的减少，MBDTA 的 *CI* 得分比 DeepGS（0.882）少 0.003。这一点的原因在于：DeepGS 使用的编码方式在编码前已经在大型生物语料库中进行了训练。这使得 DeepGS 在数据量较小的情况下，实验结果的一致性指标也能够获得更优的性能。但是对于预测准确率和预测性能而言，MBDTA 相比于其他深度学习模型平均提升了 9.02%和 0.40%。实验结果表明：在数据量不足的情况下，MBDTA 的预测性能可能会受到一定的限制。随着训练数据量的增大，MBDTA 的 MLGEN 组件逐渐发挥出其聚合多阶领域特征的优势。

为进一步证明了 MBDTA 模型在 DTA 预测任务中的竞争力，本文根据公式(18)计算相对提升率 *UP* 。

$$UP = (A_{new} - A_{old}) / A_{old} \times 100\%, \tag{18}$$

其中 A_{new} 和 A_{new} 分别表示 MBDTA 和其他模型的评价指标值。MBDTA 在两个数据集上相对于其他模型的提升率如 5 所示。

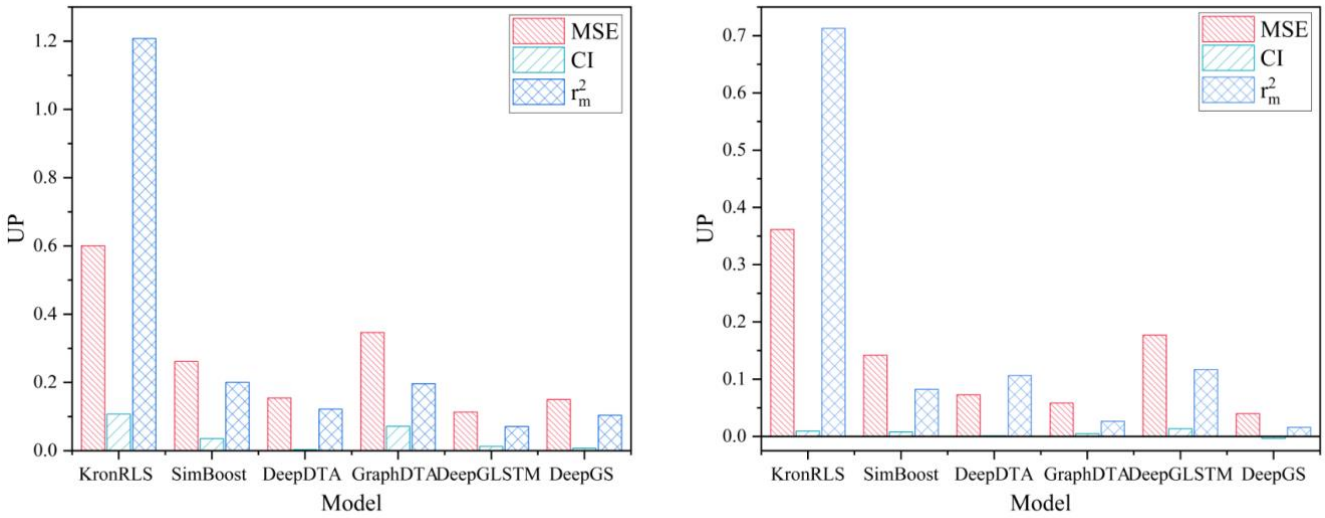


图 5 相对提升率

从图 5 中能够明显的看出 MBDTA 在两个数据集上相对于其他模型的提升效果。在 Davis 数据集上，MBDTA 的 *MSE* 得分相较于其它模型降低了 3.97%-36.15%；在 r_m^2 指标上，MBDTA 提升了 1.06%-71.25%。尽管在 *CI* 得分上比 DeepGS 模型略低，但是在数据量更大的 KIBA 数据集上，MBDTA 与其它模型相比，*CI* 得分提高了 0.35%-10.74%。而在另外两个评价指标上，MBDTA 在 *MSE* 指标上比其它模型降低了 11.35%-60.09%，在 r_m^2 指标上比其他模型提升了 7.09%-120.08%。实验结果表明 MBDTA 模型能有效提取药物分子和靶标氨基酸序列的深层次信息，实现对 DTA 的准确预测。

4.6 消融实验

MBDTA 的核心策略在于通过自动汇聚多阶邻域节点特征充分挖掘分子图中的全局拓扑结构和局部化学信息，并以不同方向获取靶标序列的上下文信息，从而更好地学习药物与靶标的潜在特征。为了验证以上策略的有效性，本文实现了 MBDTA 的四种变体：MBDTA_A1、MBDTA_A2、MBDTA_A3 和 MBDTA_LSTM 并将其在 Davis 数据集上的训练过程进行分析比较。为了保证网络层次对模型预测结果的干扰，本文将 MLGEN 模块中的通道二和通道三融合为一个网络，作为 MBDTA_A2 和 MBDTA_A3 的多阶邻域特征提取网络。变体模型的差异在于多阶邻域以及

靶标特征提取网络的选取（A1、A2 和 A3 分别表示节点的一阶、二阶以及三阶邻域），详细模块设置如表 4 所示。在以下实验中，所有模型均采用与表 1 相同的实验设置，比较结果如图 6 所示。

表 4 MBDTA 模型及四个变体主要模块说明

模型	多阶邻域	靶标特征提取网络
MBDTA	A1、A2、A3	BiLSTM
MBDTA_A1	A1	BiLSTM
MBDTA_A2	A1、A2	BiLSTM
MBDTA_A3	A1、A3	BiLSTM
MBDTA_LSTM	A1、A2、A3	LSTM

从图 7(a)、(e)、(i)中可以看出，尽管变体模型 MBDTA_A1 使用了六个 GEN 层对一阶邻居节点特征信息进行聚合，但是在训练过程中，MBDTA 的得分在每个轮次都低于 MBDTA_A1，和得分均高于 MBDTA_A1。MBDTA_A1 的最低得分为 0.258。通过添加多阶邻域特征，MBDTA_A2、MBDTA_A3 和 MBDTA 的得分分别降低了 5.71%、3.78%和 6.63%。在指标上，MBDTA_A2 和 MBDTA_A3 相比于 MBDTA_A1 分别提高了 3.20%和 2.09%，而 MBDTA 取得了五个模型中的最高得分（0.697）。MBDTA_A1、MBDTA_A2 和 MBDTA_A3 在得分相比于 MBDTA 分别提高了 0.004、0.005 和 0.005。这些现象的成因有以下几点：(1) 网络层次的增加使得 MBDTA_A1 汇聚了更多的原子节点信息。尽管 MBDTA_A1 无法有效区分相似的分子图结构导致增大，但是在网络层叠加下，得分得到了 0.47%的提升。(2) MBDTA 中对原子节点的二阶邻域和三阶邻域分别使用两层和单层 GEN，无法有效整合数据集中的部分药物分子的特征，导致得分相比于 MBDTA_A2 和 MBDTA_A3 分别降低了 0.60%和 0.61%。实验结果说明：综合考虑多个多阶邻域特征能够有效提高 DTA 预测任务的准确度，得到更加精确的药物-靶标对结合亲和力值。以单向 LSTM 模块对靶标氨基酸序列进行特征提取的 MBDTA_LSTM 模型在三个指标上的得分分别为：0.245、0.878 和 0.656。相比于 MBDTA 在上增长了 0.002，在和得分上分别降低了 0.001 和 0.041。实验结果说明靶标氨基酸序列在不同的方向具有一定的潜在特征。综合以上三个指标的分析，整合多个多阶邻域特征并提取氨基酸序列双向特征的 MBDTA 能够更有效地提取药物-靶标对的潜在特征，准确预测药物-靶标对的结合亲和力值。

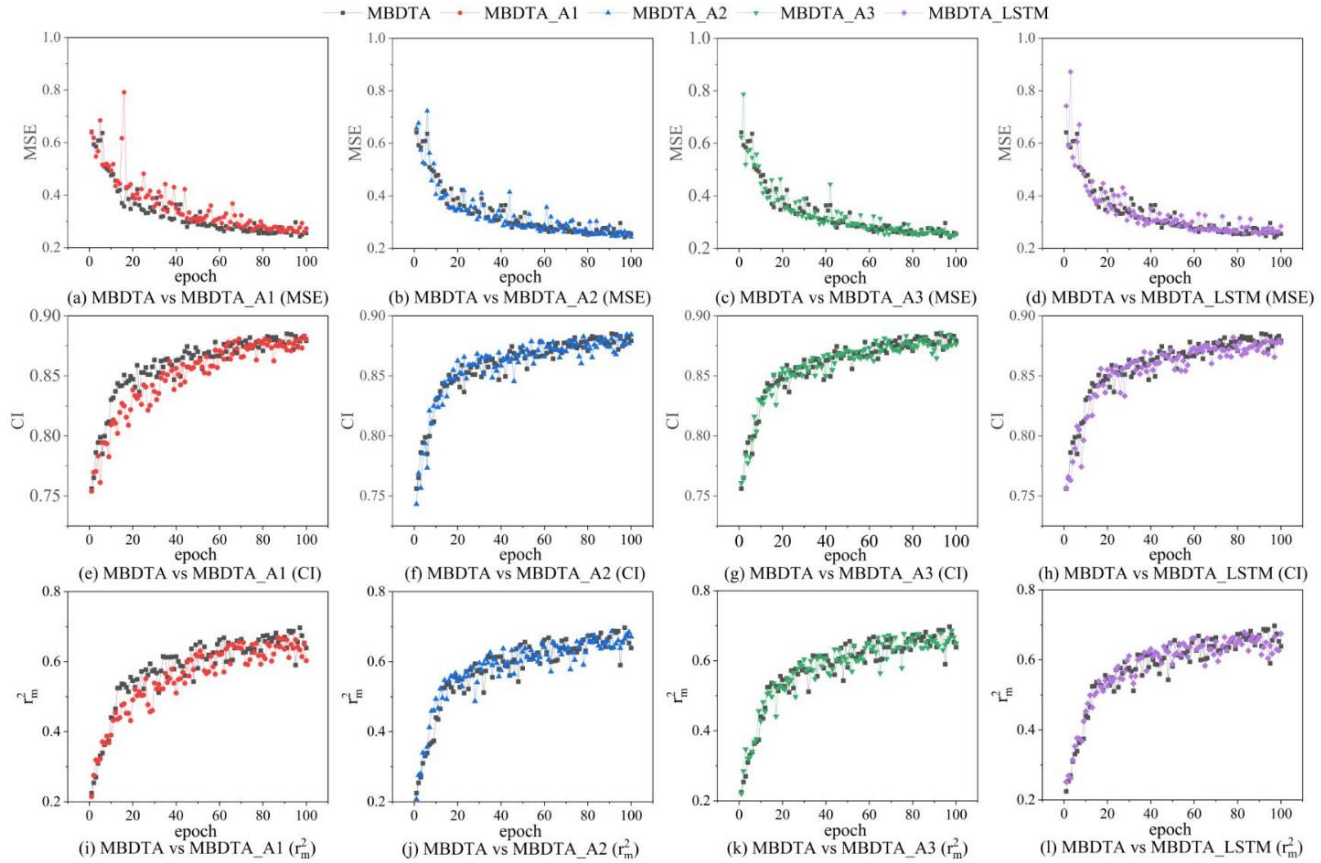
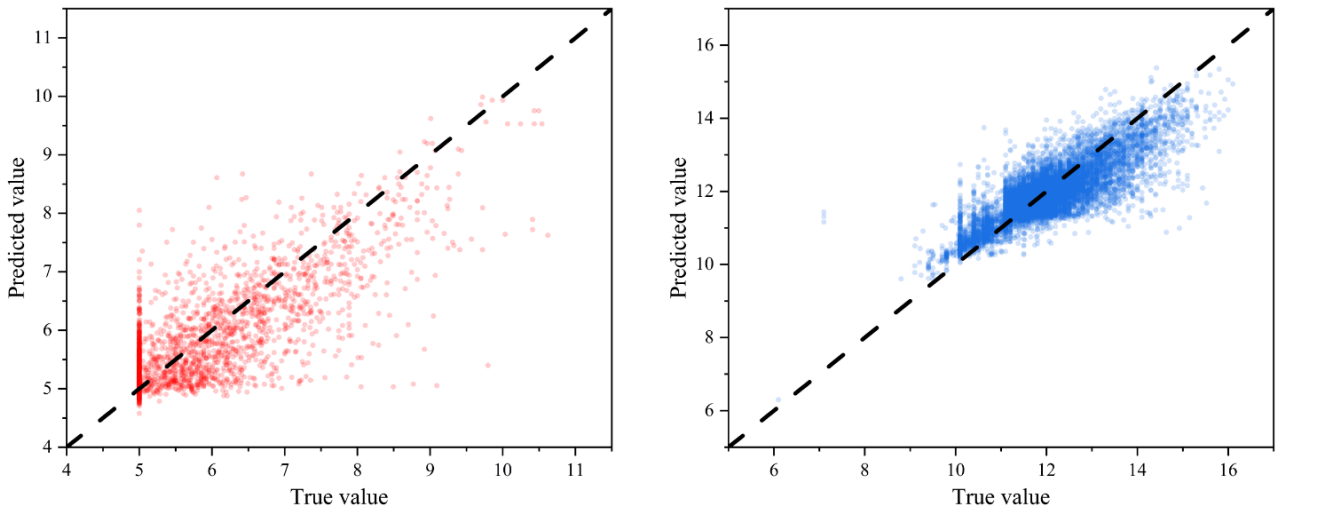


图 6 MBDTA 及其四个变体在 Davis 数据集上的评价指标变化过程

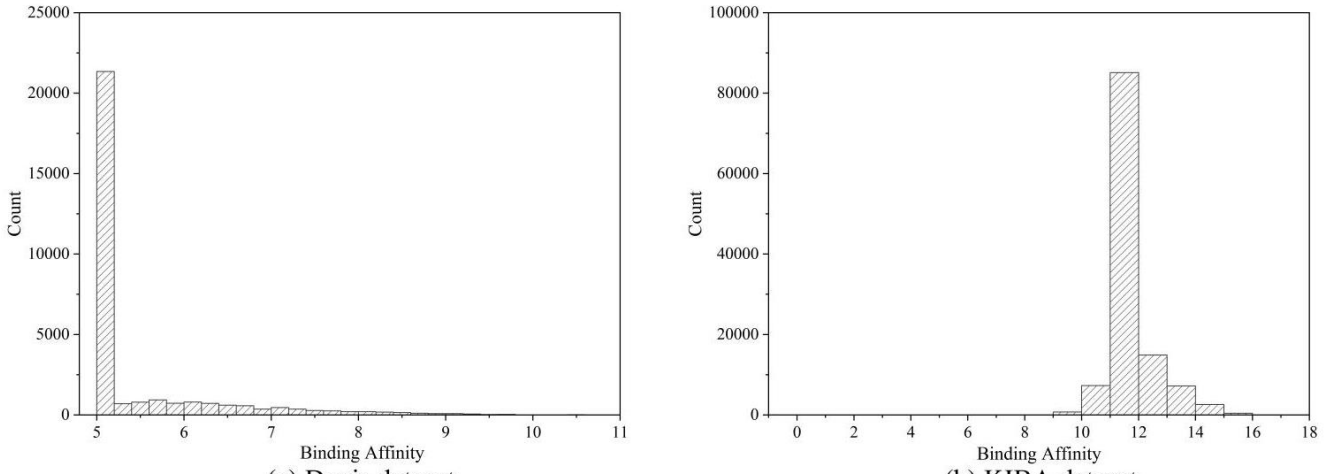


(a) Scatter plot of true and predicted values (Davis) (b) Scatter plot of true and predicted values (KIBA)

图 7 MBDTA 模型 T-P 散点图

4.7 模型分析

本节基于 Davis 和 KIBA 数据集的测试集对 MBDTA 的预测性能进行进一步验证。图 7 展示了 MBDTA 在两个测试集上预测的 DTA 散点图。图中横轴表示真实值 T，纵轴代表预测值 P。一个优秀的模型的预测值应接近于真实值，即 $P \to T$。因此预测值应该落在黑色虚线上或者靠近黑色虚线。



(a) Davis dataset (b) KIBA dataset

图 8 药物-靶标对结合亲和力分布直方图

从图 7 可以看出，两个数据集的预测值密集区域有所不同（颜色越深越密集）。在图 7(a)中，MBDTA 在 Davis 数据集上的预测点大部分都聚集在[5, 6]的区间内，而图 7(b)中 KIBA 数据集上的预测点密集区域在[10, 14]内。为分析此种现象，本文通过对数据集的结合亲和力值绘制直方图进行分析，如图 8 所示。从图 8 可以看出，Davis 数据集中结合亲和力值 pK_d 在[5, 6]区间内的药物-靶标对占整个数据集的 50%以上（即 24,495/30,056）。图 7(b)中 KIBA 的预测值密集区域出现在[10, 14]内的原因与 Davis 数据集相似。这同时证明了 MBDTA 模型具有不错的预测性能。

5 结论

当下，利用深度学习模型挖掘原始数据中潜在特征信息的方法已经得到广泛的应用和认可。本文提出了一种基于深度学习的药物-靶标结合亲和力预测模型

MBDTA。MBDTA 通过综合药物分子图的多阶邻域特征和靶标序列的双向特征信息，利用图神经网络模型对其中的潜在特征信息进行挖掘，实现对 DTA 的准确预测。本文通过在 Davis 和 KIBA 两个数据集上进行大量实验，并与当前最先进的 DTA 预测模型进行对比研究。实验结果表明，在 KIBA 数据集上，MBDTA 的三个评价指标得分（ *MSE* 、 *CI* 和 r_m^2 ）比当前最先进的模型的性能平均提升了 32.42%、3.84%和 23.64%。这证明 MBDTA 模型在 DTA 预测任务中的高效性。MBDTA 将表示药物分子的 SMILES 序列进行独热编码，随后送入 MLGEN 模块提取药物分子的整体特征。这种方式可能会丢失 SMILES 序列中用于表示药物分子图级结构的符号特征，虽然每个节点加入了部分结构特征信息，但是这种方式可能并不是最佳方案。因此，未来的研究应当继续探索如何综合考虑药物分子图的图级结构特征对 DTA 进行预测。

参考文献

[1] PRASAD V, MAILANKODY S. Research and Development Spending to Bring a Single Cancer Drug to Market and Revenues After Approval [J]. JAMA internal medicine, 2017, 177(11): 1569-75.

[2] LANDRY Y, GIES J P. Drugs and their molecular targets: an updated overview [J]. Fundamental and clinical pharmacology, 2008, 22(1): 1-18.

[3] 王可鉴, 贺林, 杨仑. 生物信息学在药物研究和开发中的应用 [J]. 中国药理学与毒理学杂志, 2014, 28(1): 118-25.

[4] PUSHPAKOM S, IORIO F, EYERS P, et al. Drug repurposing: progress, challenges and recommendations [J]. Nature reviews Drug discovery, 2019, 18(1): 41-58.

[5] D'SOUZA S, PREMA K, BALAJI S. Machine learning models for drug–target interactions: current knowledge and future directions [J]. Drug Discovery Today, 2020, 25(4): 748-56.

[6] EZZAT A, WU M, LI X, et al. Computational prediction of drug–target interactions using chemogenomic approaches: an empirical survey [J]. Briefings in Bioinformatics, 2019, 20(4): 1337-57.

[7] ZHANG L, CHEN J, CHEN J, et al. LDD-Net: Lightweight printed circuit board defect detection network fusing multi-scale features [J]. Engineering Applications of Artificial Intelligence, 2024, 129: 107628.

[8] XU B, CEN K, HUANG J. A review of graph convolutional neural networks [J]. Chinese Journal of Computers, 2020, 43(5): 755-80.

[9] SHAO K, ZHANG Y, WEN Y, et al. DTI-HETA: prediction of drug–target interactions based on GCN and GAT on heterogeneous graph [J]. Briefings in Bioinformatics, 2022, 23(3): bbac109.

[10] HU W, LIU B, GOMES J, et al. Strategies For Pre-training Graph Neural Networks; proceedings of the International Conference on Learning Representations (ICLR), New Orleans, Louisiana, United States, F, 2020 [C].

[11] LI G, MULLER M, THABET A, et al. Deepgcns: Can gcns go as deep as cnns?; proceedings of the Proceedings of the IEEE/CVF international conference on computer vision, Seoul, Korea (South), F, 2019 [C].

[12] TRAN H N T, THOMAS J J, MALIM N H A H J P. DeepNC: a framework for drug-target interaction prediction with graph neural networks [J]. 2022, 10: e13163.

[13] PASCANU R, MIKOLOV T, BENGIO Y. On the difficulty of training recurrent neural networks; proceedings of the International conference on machine learning, Miami, Florida, USA, F, 2013 [C]. PMLR.

[14] CHENG Z, ZHOU S, WANG Y, et al. Effectively Identifying Compound-Protein Interactions by Learning from Positive and Unlabeled Examples [J]. IEEE/ACM transactions on computational biology and bioinformatics, 2016, 15(6): 1832-43.

[15] CICHONSKA A, PAHIKKALA T, SZEDMAK S, et al. Learning with multiple pairwise kernels for drug bioactivity prediction [J]. Bioinformatics, 2018, 34(13): i509-i18.

[16] HE T, HEIDEMEYER M, BAN F, et al. SimBoost: a read-across approach for predicting drug–target binding affinities using gradient boosting machines [J]. Journal of cheminformatics, 2017, 9(1): 1-14.

[17] WU Y, GAO M, ZENG M, et al. BridgeDPI: a novel Graph Neural Network for predicting drug–protein interactions [J]. Bioinformatics, 2022, 38(9): 2571-8.

[18] WEN M, ZHANG Z, NIU S, et al. Deep-learning-based drug–target interaction prediction [J]. Journal of proteome research, 2017, 16(4): 1401-9.

[19] WAN F, HONG L, XIAO A, et al. NeoDTI: neural integration of neighbor information from a heterogeneous network for discovering new drug–target interactions [J]. Bioinformatics, 2019, 35(1): 104-11.

[20] ZHANG L, HU Y, CHEN J, et al. MSSIF-Net: an efficient CNN automatic detection method for freight train images [J]. Neural Computing and Applications, 2023, 35(9): 6767-85.

[21] ZHANG L, CHEN J, ZENG W, et al. FDNet: Lightweight train image fault detection network in edge computing environments [J]. IEEE Sensors Journal, 2023, 23(20): 25105-25115.

[22] ÖZTÜRK H, ÖZGÜR A, OZKIRIMLI E. DeepDTA: deep drug–target binding affinity prediction [J]. Bioinformatics, 2018, 34(17): i821-i9.

[23] PENG J, WANG Y, GUAN J, et al. An end-to-end heterogeneous graph representation learning-based framework for drug-target interaction prediction [J]. Briefings in Bioinformatics, 2021, 22(5): bbaa430.

[24] NGUYEN T, LE H, QUINN T P, et al. GraphDTA: Predicting drug–target binding affinity with graph neural networks [J]. Bioinformatics, 2021, 37(8): 1140-7.

[25] LANDRUM G. RDKit: A software suite for cheminformatics, computational chemistry, and predictive modeling [J]. Greg Landrum, 2013.

[26] RAMSUNDAR B. Molecular machine learning with DeepChem [D]; Stanford University, 2018.

[27] ZHANG P, WEI Z, CHE C, et al. DeepMGT-DTI: Transformer network incorporating multilayer graph information for Drug-Target interaction prediction [J]. Computers in biology medicine, 2022, 142: 105214.

[28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in neural information processing systems, 2017, 30: 6000-10.

[29] WISHART D S, FEUNANG Y D, GUO A C, et al. DrugBank 5.0: a major update to the DrugBank database for 2018 [J]. Nucleic acids research, 2018, 46(D1): D1074-D82.

[30] TRAN H N T, THOMAS J J, AHAMED HASSAIN MALIM N. DeepNC: a framework for drug-target interaction prediction with graph neural networks [J]. PeerJ, 2022, 10: e13163.

[31] WELLING M, KIPF T N. Semi-supervised classification with graph convolutional networks; proceedings of the International Conference on Learning Representations (ICLR 2017), Palais des Congrès Neptune, Toulon, France, F, 2016 [C].

[32] LI G, XIONG C, ALI T, et al. Deepergcn: All you need to train deeper gcns [J]. arXiv preprint arXiv: 200607739, 2020: 1-16.

[33] DAVIS M I, HUNT J P, HERRGARD S, et al. Comprehensive analysis of kinase inhibitor selectivity [J]. Nature biotechnology, 2011, 29(11): 1046-51.

[34] TANG J, SZWAJDA A, SHAKYAWAR S, et al. Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis [J]. Journal of Chemical Information Modeling, 2014, 54(3): 735-43.

[35] MUKHERJEE S, GHOSH M, BASUCHOWDHURI P. DeepGLSTM: Deep Graph Convolutional Network and LSTM based approach for predicting drug-target binding affinity; proceedings of the Proceedings of the 2022 SIAM International Conference on Data Mining (SDM), Zurich, Switzerland, F, 2022 [C]. SIAM.

[36] LIN X. DeepGS: Deep Representation Learning of Graphs and Sequences for Drug-Target Binding Affinity Prediction; proceedings of the The 24th European Conference on Artifical Intelligence (ECAI), Santiago de Compostela, Spain, F, 2020 [C]. {IOS} Press.

[37] FEY M, LENSSEN J E. Fast graph representation learning with PyTorch Geometric [J]. ICLR Workshop on Representation Learning on Graphs and Manifolds, 2019: 1-9.