

基于图同构网络与自注意力机制的药物-药物相互作用预测

陈 诚^{1,2}, 王 文^{1,2}, 夏迎春^{1,2}, 唐良贵^{1,2}, 许文俊^{1,2}, 王庆勇^{1,2}, 辜丽川^{1,2}

1. 安徽农业大学 信息与人工智能学院, 合肥 230036

2. 智慧农业技术与装备安徽省重点实验室, 合肥 230036

摘要:目的 药物-药物相互作用 (Drug-Drug Interaction, DDI) 预测能够辅助确定有效和安全的药物联合治疗。但以往的预测方法通常围绕药物的二维分子建模, 没有考虑药物的三维几何结构; 并且忽略了药物作用子结构的多样性, 单一的特征拟合路径可能会限制模型的预测能力。方法 为此提出了基于图同构网络与注意力机制的 DDI 预测模型, 它基于图同构网络构建分子图特征生成模块用以生成三维坐标系上的药物分子图特征; 并在此基础上使用自注意力机制构建子结构特征提取模块, 从多个维度提取药物的空间子结构; 最后使用 DDI 三元组预测模块来识别潜在的 DDI。整个模型采用双层相互独立的网络, 以此提取不同作用的子结构。结果 实验结果证明: 该模型在 DrugBank 数据集上相对以往方法取得了不错的优势, 特别是在预测新药-新药相互作用的实验中, 它在准确率 ACC、受试者工作特征曲线下面积 AUC、 F_1 分数上相比最好的基线分别提升了 7.59%、9.45%、12.95%, 展现了良好的泛化性能。结论 消融实验证明: 子结构特征提取模块能够有效提取空间子结构, 双层网络使模型获得了更好的预测性能。

关键词: 药物-药物相互作用预测; 图同构网络; 自注意力机制; 空间子结构; 双层网络

中图分类号: TP181 **文献标识码:** A **doi:** 10.16055/j.issn.1672-058X.2026.0002.005

Drug-drug Interaction Prediction Based on Graph Isomorphic Network and Self-attention Mechanism

CHEN Cheng^{1,2}, WANG Wen^{1,2}, XIA Yingchun^{1,2}, TANG Lianggui^{1,2}, XU Wenjun^{1,2}, WANG Qingyong^{1,2}, GU Lichuan^{1,2}

1. School of Information and Artificial Intelligence, Anhui Agricultural University, Hefei 230036, China

2. Key Laboratory of Intelligent Agricultural Technology and Equipment of Anhui Province, Hefei 230036, China

Abstract: Objective Drug-drug interaction (DDI) prediction can assist in determining effective and safe combinations of drugs for treatment. However, previous prediction methods typically model drugs based on their two-dimensional molecular structures, without considering the three-dimensional geometric structure of drugs. Moreover, they overlook the diversity of drug functional substructures, and a single feature-fitting path may limit the predictive ability of the model.

Methods Therefore, a DDI prediction model based on a graph isomorphic network and attention mechanism was proposed. The model constructed a molecular graph feature generation module using the graph isomorphic network to generate drug molecular graph features in a three-dimensional coordinate system. On this basis, a self-attention mechanism was used to build a substructure feature extraction module to extract spatial substructures of drugs from multiple dimensions. Finally, a DDI triplet prediction module was used to identify potential DDIs. The entire model adopted a dual-layer, mutually independent network to extract different acting substructures. **Results** Experimental results

收稿日期: 2024-03-05 修回日期: 2024-06-17 文章编号: 1672-058X(2026)02-0035-08

基金项目: 国家自然科学基金(62301006); 国家重点研发计划项目(2023YFD1802200)资助。

作者简介: 陈诚(1999—), 男, 湖北黄石人, 硕士研究生, 从事生物信息学研究。

通信作者: 辜丽川(1974—), 男, 合肥蜀山区人, 博士, 教授, 从事人工智能与大数据研究。Email: glc@ahau.edu.cn

引用格式: 陈诚, 王文, 夏迎春, 等. 基于图同构网络与自注意力机制的药物-药物相互作用预测[J]. 重庆工商大学学报(自然科学版), 2026, 43(2): 35-42.

CHEN Cheng, WANG Wen, XIA Yingchun, et al. Drug-drug interaction prediction based on graph isomorphic network and self-attention mechanism[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2026, 43(2): 35-42.

demonstrated that the model achieved significant advantages over previous methods on the DrugBank dataset, particularly in predicting interactions between new drugs. It achieved improvements of 7.59% in accuracy (ACC), 9.45% in area under the receiver operating characteristic curve (AUC), and 12.95% in F_1 score compared with the best baseline, demonstrating excellent generalization performance. **Conclusion** Ablation experiments prove that the substructure feature extraction module can effectively extract spatial substructures, and the dual-layer network makes the model obtain better prediction performance.

Keywords: drug-drug interaction prediction; graph isomorphism network; self-attention mechanism; spatial substructure; dual-layer network

医生为病人提供的治疗策略通常会包含药物的组合使用,它在提升疗效的同时可能会因为药物之间理化性质不相容而引起副作用^[1-6]。因此,药物-药物相互作用是药物研究中重要的一环,但药物的组合空间巨大,无效的药物组合会给研究带来资源上的浪费。为了解决这个问题,目前已经有大量的 DDI 预测方法被提出,并证明了其有效性。

早期的 DDI 预测方法大多遵循一个假设,即特征相似的药物可能会发生相似的相互作用^[7-8]。从这个假设出发,后续的研究人员为了找到药物尽可能多的相似特征,从最开始的手动搜集,到后来利用开源的大型医学知识图谱来获取药物的外部信息。随着图神经网络^[9-11]的发展,研究人员围绕知识图谱构建 DDI 预测模型^[12-14]成为主流,特征的多样性以及算法的进步^[15-16]使得 DDI 预测有了很大进展。但是,其研究忽视了药物相互作用发生的主要机制,即药物与药物的化学子结构之间发生了相互作用^[17-18]。同时在新药的研发中,通常只有分子结构作为原始特征,围绕药物外部信息构建的预测模型不能被有效地利用。

为了解决上述问题,有许多聚焦于药物分子结构的 DDI 预测方法被提出,基于药物的分子图构建模型,在关于新药的 DDI 预测中取得了不错的效果。Xu 等^[19]提出了 MR-GNN,它在图神经网络的基础上引入了双图状态 LSTM(Long short-term memory, LSTM),为 DDI 预测提供了多样的子图结构特征;Nyamabo 等^[20]提出了 SSI-DDI,与传统方法不同,它不只考虑药物的整个化学结构,还注意到了发生相互作用的子结构,并将其作为预测的基础,这种方法也为预测带来了不错的效果;Nyamabo 等^[21]后来又提出了一种名为门控消息传递神经网络(GMPNN)的 DDI 预测方法。它利用药物的分子图学习不同大小和形状的化学子结构,药物对的子结构特征会进行成对交互,并根据不同权重对最终的预测分数造成不同影响。事实上,GMPNN 的做法跟 MR-GNN 有些类似,都是在药物分子图上提取多样的子图以供预测,但 GMPNN 的优势在于子图权重的引入,无效子结构与作用子结构的区分会让预测更加准确;Yang 等^[22]提出了 SA-DDI,它采用一种新颖的

子结构注意机制,这种机制实际上是另一种子结构权重分配方法,通过分子内的消息传递给子结构并提供不同的注意力,以此来找到作用子结构;Li 等^[23]提出了一种名为 DSN-DDI 的双视图药物表示学习网络,能够同时从单个药物和药物对中学习化学子结构,药物对分子之间的原子交互学习能使模型进一步确定作用子结构,从而提升预测性能。

虽然上述工作大多数都考虑了药物的子结构交互,但这些研究都局限于在药物的二维分子上建模,而对于一个药物来说,只考虑它的二维结构显然是不充分的,并且它与不同药物发生相互作用的子结构并不固定,这些基本信息的缺失会对模型的预测性能造成影响。针对上述局限性,本文设计了一种基于 GIN 的双层组合网络模型,从药物的 3D 几何结构出发,对三维坐标系每一维度的分子平面进行建模,并通过图同构网络^[24](Graph Isomorphism Network, GIN)拟合该平面上的分子图特征。为了突出药物分子的作用子结构,在图同构层之间引入了自注意力机制^[25]来构建子结构特征提取模块。除此之外,由于药物与不同药物发生相互作用的子结构并不唯一,所以为了增加子结构特征的多样性,本文使用双层组合网络来实现药物分子特征的最终提取。本文提出的方法主要贡献如下:

(1) 设计了一种基于 GIN 架构的 DDI 预测模型,通过对药物分子的三维结构进行建模,使用自注意力机制提取药物的空间作用子结构特征,利用双层网络来提升子结构特征的多样性,从而提升 DDI 的预测性能。

(2) 为了验证模型的有效性,分别在 DrugBank 公共数据集上进行直推式实验与归纳实验。实验结果表明:本文提出的模型在多个预测任务中都展现出了良好的性能。特别是在新药-新药相互作用预测中,相比于最好的基线在性能指标 ACC、AUC、 F_1 分数上分别提升了 7.59%、9.45%、12.95%,展现出了强大的泛化能力。

1 相关理论概述

1.1 图同构网络

Keyulu 等在图神经网络的基础上提出了图同构网络,这是一种能有效提取图表示的架构。它可以在节点特征传递和聚合的过程中保留图的同构性信息,从

而提高图数据的建模能力和表征学习效果,并在理论上能够逼近图同构测试(Weisfeiler-Lehman, WL)。

GIN 模型的网络结构通常包含多个图神经网络层。每层都由两个关键步骤组成:节点特征传递和节点特征聚合。节点特征传递阶段,每个节点将其特征向量与邻居节点的特征向量进行聚合操作,得到更新后的节点特征。节点特征聚合阶段,对所有节点的特征向量进行池化,得到整个图的表示向量。GIN 强大的表征能力为本文在药物分子图上提取代表性特征提供了可能,也为后续作用子结构的提取打下了基础。

1.2 自注意力机制

Transformer 的出现带来了自注意力机制,从一开始应用于自然语言处理,逐步扩展到计算机视觉、生物信息等领域。自注意力机制的核心是让原本的特征通过三个全连接层得到相应的查询矩阵 Q 、键矩阵 K 、值矩阵 V ,这三种特征尽管不完全一样,但仍保留了原始特征。接下来就是通过计算 Q 和 K 之间的注意力,得到不同向量之间的注意力矩阵,最终进行标准化后与值矩阵融合得到最终的输出。

在模型的训练下,原始特征里的重要特征会被放

大,这会让预测变得更加准确。而这正是本文需要的,自注意力机制能让分子图中的作用子结构被放大,从而使得模型能够准确地识别 DDI。

2 模型框架设计

如图 1 所示,该模型由两层相同的网络组成,每层网络主要包含两部分,分别是分子图特征生成模块以及子结构特征提取模块,并且它们在输出后共用一个 DDI 预测模块。分子图特征生成模块一共有 6 层,第 1 层属于初级编码器,负责对原子特征以及边特征进行初始化;后 5 层都由 GIN 构成,它将已经编码过的三维原子特征与边特征按照聚合模式提取为相对应的分子图特征,并将更新后的节点特征输入子结构特征提取模块和下一层进行更深层次的分子图特征生成。子结构特征提取模块同样有 5 层,它通过分子图特征对自身的注意力分数获取关键的作用子结构特征,其中每一层的作用子结构特征都会参与最终的输出,这会进一步完善作用子结构的整体。DDI 预测模块会将上述两个模块的所有输出进行累加得到药物的最终特征,然后按照药物-关系-药物的方式进行矩阵相乘求和并激活,从而得到最终的预测分数。



图 1 模型框架图

Fig. 1 Model framework

2.1 分子图特征生成模块

本文在分子图特征生成模块使用了 GIN 实现药物分子图特征的提取,在此之前,每个药物分子都会进入初级编码器中进行原子特征与原子之间边特征的随机初始化。对于原子特征来说,它的初始化发生在模型初始化之前,因为原子特征初始化的范围是整个药物空间。具体来说,不同药物所包含的相同原子应该具有相同的初始特征。但边特征并非如此,因为即便是两个相同的原子之间也不一定生成相同的化学键,所以本文将边特征设置为随机参数,与模型一起初始化。

为了引入药物的 3D 几何结构,本文将每个原子的原始特征设为 $\mathbf{a} \in \mathbf{R}^{1 \times 3 \times n}$, 由一个三维矩阵表示, n 代表向量维度, 向量维度为超参数, 它会决定药物特征向量的最终维度。而药物的节点特征矩阵可以表示为 $\mathbf{A}^{(k)} \in \mathbf{R}^{N \times 3 \times n}$, N 表示原子节点的总数。那么药物中的原子都可以初始化为

$$\boldsymbol{\alpha} = (x, y, z) \cdot \mathbf{a}, \boldsymbol{\alpha} \in \mathbf{R}^{1 \times 3 \times n} \quad (1)$$

其中, (x, y, z) 表示原子的三维坐标, “ \cdot ” 表示三维坐标与原子特征内的三个向量进行对应乘积, 最后生成一个同样大小的三维表示。边特征进行随机初始化得到 $\mathbf{e} \in \mathbf{R}^{3n}$, 它不参与更新。在所有的初始化完成后, 整个分子图会被送入 GIN 层。在 GIN 层中, 原子特征会转换为向量, 即 $\boldsymbol{\alpha} \in \mathbf{R}^{3n}$, 这样每一维度都会与边特征一起参与整个分子图的更新。分子图中原子节点特征的更新函数如下:

$$\boldsymbol{\alpha}_i^{(k)} = \text{MLP}^{(k)} \left(\sum_{j \in N(i)} \boldsymbol{\alpha}_i^{(k-1)} + \mathbf{e}_{ij} \right) \quad (2)$$

其中, i 表示当前需要更新的原子节点, k 表示当前的层数, $N(i)$ 表示当前原子邻接节点的集合。这个函数意味着节点特征的更新由上一层输入的原子特征与其所有邻接的边特征求和, 并输入一个多层感知机 (MLP) 中完成。值得注意的是, 同一个分子图上的所有节点特征会经过同一个 MLP, 而处在不同层的分子图上的节点特征进入的 MLP 是不同的。每层最终的分子图特征由所有的节点特征进行平均池化得到, 具体的函数如下:

$$\mathbf{g}'^{(k)} = \text{AveragePool}(\{ \boldsymbol{\alpha}_i^{(k)} \mid \boldsymbol{\alpha}_i^{(k)} \in \mathbf{A}^{(k)}, i \in N \}) \quad (3)$$

其中, i 表示原子节点的总数, $\mathbf{g}'^{(k)} \in \mathbf{R}^{3 \times n}$ 表示分子图特征向量, 它将被合并为多维向量 $\mathbf{g}^{(k)} \in \mathbf{R}^{3n}$ 作为最终的分子图表示。然后, 本文将每层的分子图特征进行累加得到最终的分子图特征, 它将参与药物特征的最终合成。最后, 本文将更新后的节点特征矩阵 $\mathbf{A}^{(k)}$ 的大小转换为二维矩阵 $3N \times n$, 并输入对应的自注意力层, 以求得作用子结构的完整提取。

2.2 子结构特征提取模块

在这个模块中, 分别使用五层与 GIN 相对应的自

注意力层来进行作用子结构特征的提取, 并将输入设为第 k 层的节点特征矩阵 $\mathbf{A}^{(k)} \in \mathbf{R}^{3N \times n}$ 。可以发现, 尽管上个模块中初始的原子节点特征经过聚合从三维矩阵变换成了表示分子图特征的向量, 但是坐标 (x, y, z) 维度上的分子图信息仍然独立, 没有很好地融合, 而本文引入的自注意力机制可以很好地解决这一问题。对于节点特征矩阵来说, $\mathbf{A}^{(k)}$ 代表了三个坐标维度上的分子图特征, 在获取自注意力的过程中, 每个维度的节点特征都会与另两个维度上的分子图特征进行交互, 得到相应的注意力, 这些注意力将会与最终的值矩阵融合, 使得模块能从空间中提取作用子结构。

在每层自注意力层中, 输入的节点特征矩阵 $\mathbf{A}^{(k)}$ 会被输送进三个不同的全连接层中, 三个全连接层的参数分别为 $\mathbf{W}_Q^{(k)}$ 、 $\mathbf{W}_K^{(k)}$ 、 $\mathbf{W}_V^{(k)}$, 生成相应的 $\mathbf{Q}^{(k)}$ 、 $\mathbf{K}^{(k)}$ 、 $\mathbf{V}^{(k)}$, 随后按照如下函数生成表示:

$$\mathbf{Q}^{(k)} = \mathbf{A}^{(k)} \mathbf{W}_Q^{(k)} \quad (4)$$

$$\mathbf{K}^{(k)} = \mathbf{A}^{(k)} \mathbf{W}_K^{(k)} \quad (5)$$

$$\mathbf{V}^{(k)} = \mathbf{A}^{(k)} \mathbf{W}_V^{(k)} \quad (6)$$

$$\text{Attention}(\mathbf{Q}^{(k)}, \mathbf{K}^{(k)}, \mathbf{V}^{(k)}) = \text{Softmax} \left(\frac{\mathbf{Q}^{(k)} \mathbf{K}^{(k)}}{\sqrt{N}} \right) \mathbf{V}^{(k)} \quad (7)$$

其中, 矩阵相连代表着两个矩阵之间进行了点积, 最终生成的作用子结构特征矩阵大小与节点特征矩阵 $\mathbf{A}^{(k)}$ 大小相同, 不同的是, 此时的特征矩阵已经强调了节点特征矩阵中异常关键的原子特征, 也就是作用子结构特征。同时, 三个维度的节点特征也进行了注意力交互, 它们从空间的角度进一步完善了作用子结构的特征提取。但是, 注意力矩阵的大小与分子图特征的大小并不相同, 本文需要将其按照维度进行拼接, 最终得到输出矩阵 $\mathbf{B}^{(k)} \in \mathbf{R}^{N \times 3n}$ 。那么作用子结构特征向量 $\mathbf{h}^{(k)} \in \mathbf{R}^{3n}$ 同样可以通过平均池化得到:

$$\mathbf{h}^{(k)} = \text{AveragePool}(\{ \mathbf{b}_i^{(k)} \mid \mathbf{b}_i^{(k)} \in \mathbf{B}^{(k)}, i \in N \}) \quad (8)$$

值得注意的是, 单层的自注意力机制并不能很好地提取作用子结构, 它依赖于分子图特征的完整性, 也就是看输入的节点特征矩阵能否代表一个完整的分子图, 因而本文采用自注意力层与 GIN 层一一对应的方式来提取作用子结构。

2.3 DDI 三元组预测模块

这个模块解决的是直推式任务和归纳任务。直推式任务是指在测试集的药物出现在训练集的情况下完成模型的训练, 从而在测试集中推理出药物对是否发生了某种具体的相互作用。简单来说, 就是预测已知药物之间相互作用是否发生。但归纳任务中测试集里的一部分或所有药物都从未在训练集里出现过, 也就是需要模型完成已知药物与新药、新药与新药的相互

作用预测。

在这个 DDI 三元组预测模块中,本文将每一种 DDI 类型定义为关系矩阵 $\mathbf{P}^{(u)} \in \mathbf{R}^{n \times n}$,其中 u 代表着相互作用的具体类型。 $\mathbf{P}^{(u)}$ 实际上也是参数矩阵,它与模型一同初始化并参与训练,这意味着 DDI 类型的具体特征由训练集中出现的药物对来确定。在得到前两个模块来完成药物特征的提取,过程如下:

$$\mathbf{d} = \sum_k^{l=5} (\mathbf{g}^{(k)} + \mathbf{h}^{(k)}) \quad (9)$$

式(9)中, l 表示层数, $\mathbf{d} \in \mathbf{R}^{3n}$ 是一层网络的输出,为了增加作用子结构的多样性,本文用两层网络来进行训练,并将两层网络的输出增加一个维度进行拼接得到最终的药物特征矩阵 $\mathbf{D} \in \mathbf{R}^{2 \times 3n}$ 。本文将最终的得分函数设为

$$z_{ij} = \text{SUM}(\mathbf{D}_i \mathbf{P}^{(u)} \mathbf{D}_j^T) \quad (10)$$

其中 i, j 分别表示发生相互作用 u 的两种药物,函数里相邻的矩阵按照从左到右的顺序发生了点积,而点积后的结果会按照所有维度相加,生成相应的得分。在进行测试集的预测时,本文使用 Sigmoid 函数作为激活函数,并选择 0.5 作为划分正负样本的阈值。那么最终的预测分数可以表示为

$$y_{ij} = \text{Sigmoid}(z_{ij}) \quad (11)$$

在激活函数为 Sigmoid 时,要使正样本的预测分数大于等于 0.5,那么就要使正样本的得分 z_{ij} 大于等于 0。反之,本文希望负样本的得分能够小于 0。因此,本

文将训练中的损失函数设为交叉熵损失函数:

$$L = -y^{(u)} \log \sigma(z_{ij}) - (1 - y^{(u)}) \log \sigma(-z_{ij}) \quad (12)$$

其中, $y^{(u)}$ 表示真实标签, σ 表示 Sigmoid 函数。该函数在输出激活之前将负样本得分取负,这有利于负样本得分在训练中倾向负数。

3 实验与结果分析

3.1 实验数据集

本文在一个广泛使用的数据集 DrugBank 上评估模型。DrugBank 包含 191 808 条 DDI 三元组,共计 1 706 种药物,86 种相互作用类型,并且一条药物对只存在一种相互作用,其中每个药物的三维结构分别通过 DrugBank 数据库以及 PubChem 数据库进行下载,负样本按照基线 DSN-DDI 中方法生成。本文按照 DSN-DDI 中相同的数据划分方案,对 DrugBank 数据集中的训练集、验证集和测试集以 6:2:2 的比例在直推式任务中进行划分,并进行五折交叉实验。对于归纳设置,本文也遵循了 DSN-DDI 等相同的数据分割方案,即随机选取 20% 的药物作为未知药物,并将其余的药物视为已知药物。然后,本文将只包含已知药物的药物对作为训练集,包含已知药物跟未知药物的药物对跟只包含未知药物的药物对分别划为归纳任务的两个测试集。为了公平地对比,本文同样对归纳任务重复 3 次(表 1,表 2,表 3)。

表 1 与其他先进模型在直推式设置下对比

Table 1 Comparison with other advanced models for the transductive setting

模 型	ACC	AUC	AP	F_1
MHCADDI	83.80 ± 0.27	91.16 ± 0.31	89.26 ± 0.37	85.06 ± 0.31
GAT	89.81 ± 1.00	95.21 ± 0.70	93.56 ± 0.90	90.18 ± 0.74
GMPNN-CS	95.30 ± 0.05	98.46 ± 0.01	97.94 ± 0.02	95.39 ± 0.05
MR-GNN	96.04 ± 0.05	98.87 ± 0.04	98.57 ± 0.06	96.10 ± 0.05
SA-DDI	96.23 ± 0.01	98.80 ± 0.02	98.36 ± 0.04	96.29 ± 0.09
SSI-DDI	96.33 ± 0.09	98.95 ± 0.08	98.57 ± 0.14	96.38 ± 0.09
DSN-DDI	96.94 ± 0.02	99.47 ± 0.01	99.37 ± 0.02	96.93 ± 0.02
Ours	97.87 ± 0.07	99.62 ± 0.02	96.26 ± 0.19	97.91 ± 0.07

每个值都是平均值 (%) ± 标准差的形式,粗体表示最优性能

表 2 与其他先进模型在归纳设置(新药,已知药物)下对比

Table 2 Comparison with other advanced models for the inductive setting (new drug, existing drug)

模 型	ACC	AUC	AP	F_1
MHCADDI	70.58 ± 0.94	77.84 ± 1.08	76.16 ± 1.45	72.74 ± 0.65
GAT	69.83 ± 1.41	77.29 ± 1.63	75.79 ± 1.95	73.01 ± 0.85
GMPNN-CS	77.72 ± 0.30	84.84 ± 0.15	84.87 ± 0.40	78.29 ± 0.16
MR-GNN	74.67 ± 0.33	83.15 ± 0.60	83.81 ± 0.69	69.88 ± 0.86
SA-DDI	75.55 ± 1.12	82.95 ± 1.05	84.11 ± 0.92	71.94 ± 1.57
SSI-DDI	76.38 ± 0.92	84.23 ± 1.05	84.94 ± 0.76	73.54 ± 1.50
DSN-DDI	81.92 ± 1.20	91.01 ± 0.76	91.09 ± 0.93	80.18 ± 1.49
Ours	81.47 ± 0.20	92.72 ± 0.39	73.70 ± 0.13	84.00 ± 1.18

每个值都是平均值 (%) ± 标准差的形式,粗体表示最优性能

表 3 与其他先进模型在归纳设置(新药,新药)下对比
Table 3 Comparison with other advanced models for the inductive setting (new drug, new drug)

模 型	ACC	AUC	AP	F_1
MHCADDI	66.50 ± 0.62	72.53 ± 0.92	71.06 ± 1.61	67.21 ± 0.59
GAT	66.31 ± 0.61	72.75 ± 0.78	71.61 ± 1.00	68.68 ± 0.60
GMPNN-CS	68.57 ± 0.30	74.96 ± 0.40	75.44 ± 0.50	65.32 ± 0.23
MR-GNN	62.63 ± 0.77	70.92 ± 0.84	73.01 ± 1.23	45.81 ± 2.51
SA-DDI	67.15 ± 0.88	73.62 ± 1.25	73.39 ± 1.40	63.40 ± 1.53
SSI-DDI	65.40 ± 1.30	73.43 ± 1.81	75.03 ± 1.42	54.12 ± 3.46
DSN-DDI	73.42 ± 1.29	81.79 ± 1.12	81.82 ± 1.48	70.34 ± 0.98
Ours	81.01 ± 2.06	91.24 ± 1.16	73.74 ± 3.31	83.29 ± 1.07

每个值都是平均值(%)±标准差的形式,粗体表示最优性能

3.2 对比算法

MR-GNN:利用图卷积网络捕获药物的子结构表示用于 DDI 预测。

MHCADDI^[26]:利用共同注意机制在药物分子结构上学习药物表示。

GAT-DDI^[27]:利用图注意力网络在药物分子上学习药物表示并进行 DDI 预测。

SA-DDI:设计了一种新的子结构注意机制来学习 DDI 预测药物。

GMPNN-CS:通过门控神经网络学习不同的化学子结构,用以 DDI 预测。

SSI-DDI:将每个原子节点的隐藏表示作为子结构,然后计算这些子结构之间的交互来预测 DDI。

DSN-DDI:通过双视图学习药物原子之间的交互,并生成药物表示进行 DDI 预测。

3.3 评价指标与参数设置

本文使用准确率 ACC、受试者工作特征曲线下面积 AUC、平均精确率 AP、 F_1 分数作为评价指标,以获得更公平的评价。

在直推式任务的设置中,本文将学习率设为 0.001,特征向量的维度 n 设为 300,训练批大小设为 2 048,训练周期设为 1 500 轮。在归纳任务的设置中,学习率与特征向量的维度不变,不同的是本文将批大小与训练周期分别调整为 128、24 000 轮。

3.4 实验结果分析

表 1 表明了各个模型在设置的 Drugbank 数据集上的 ACC、AUC、AP 分数和 F_1 。其中,基线的评价结果来自文献,因为在数据集以及划分方案一致的情况下,本文直接采用文献中给出的性能数据进行对比。与所有基线相比,本文的方法在除了 AP 以外的评价指标都获得了最好的评价。这说明本文的方法能根据经验有效地从三维空间中提取作用子结构特征,并将其应用于已知药物的预测。表 2 展示了所有模型在已知药物-新药测试集上的性能。可以发现本文的方法在 AUC、 F_1 分数上取得了最好的评价,并且 ACC 仅以微弱的劣势输给了最好的基线。它意味着本文除了有很好的直

推式推理能力外还有很好的泛化能力。事实上,这正是本文有效提取作用子结构的进一步佐证。

表 3 中的归纳任务是预测新药-新药之间的相互作用。可以很直观地发现,它与表 1 的情况相似,模型在 ACC、AUC、 F_1 分数上相比最好的基线 DSN-DDI 分别提升了 7.59%、9.45%、12.95%。同时本文可以观察到,大部分模型在新药-已知药物相互作用预测的任务中表现不错,但是在新药-新药相互作用的任务中性能表现不佳。而本文的模型在这两个任务中的变化很少,这说明分子图特征生成模块与子结构特征提取模块真实有效地在药物三维几何结构中提取出了药物的代表性特征。与其他模型相比,本文的优势在于三维信息与自注意力机制的纳入,使得模型能够让药物分子图特征在三个维度进行交互,从而注意到空间中的作用子结构。同时本文不能忽视的是图同构网络对药物分子图特征生成的贡献,它能够准确地拟合分子图,并为不同药物中相似的结构分配预测所需的相似性。

3.5 消融实验

为了验证子结构特征提取模块的有效性,本文分别进行了 3 组对照实验。结果如图 2、图 3、图 4 所示,本文模型去掉子结构特征提取模块后,在直推式推理中的性能只是略有下降。但在归纳学习的两个实验中性能下降严重,这说明子结构特征提取模块对整个模型的泛化能力起着很重要的作用。

作用子结构的多样性同样是本文关心的重点,因此本文使用单一网络来进行新药-新药相互作用的实验。结果如图 5 所示,使用双层网络的模型性能相比单一网络在各个指标上都有所提升。需要注意的是,模型进行的是泛化能力测试,这对双层网络能否发掘药物子结构是一个很大的挑战。并且第一层网络与第二层网络都是独立且互不影响,因此每层提取出的作用子结构也是独立的。所以,在引入双层网络后模型性能有限的提升证明了药物子结构特征单一拟合路径的局限性,双层网络在一定程度上拟合了药物的不同作用子结构。

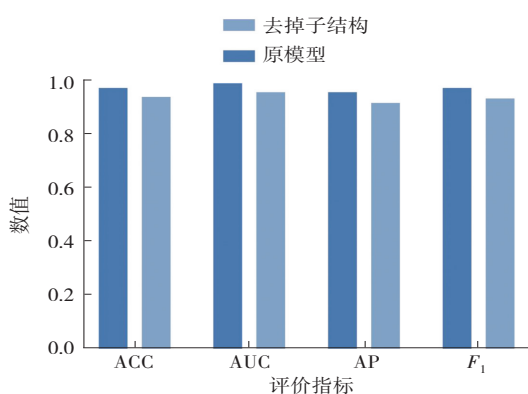


图 2 直推任务的消融研究

Fig. 2 Ablation study for transductive task

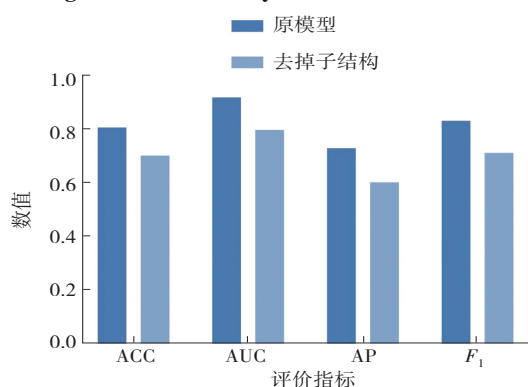


图 3 归纳任务(已知-新药)的消融研究

Fig. 3 Ablation study for the inductive task (known-new drug)

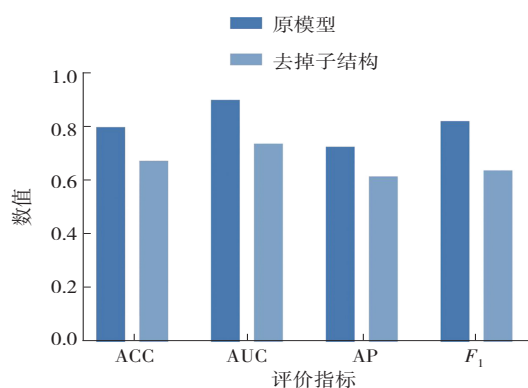


图 4 归纳任务(新药-新药)的消融研究

Fig. 4 Ablation study for the inductive task (new drug-new drug)

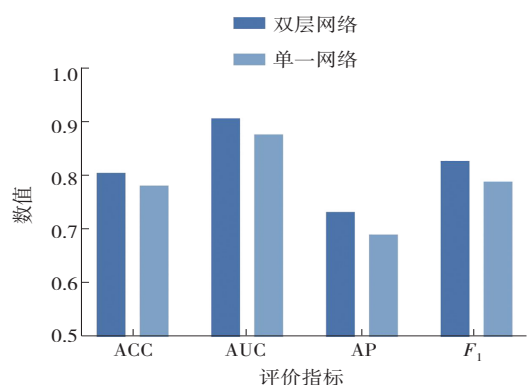


图 5 单一网络和双层网络的对比

Fig. 5 Comparison of single network and dual-layer network

4 结论

在图同构网络的基础上引入自注意力机制是拟合药物作用子结构的关键。图同构网络尽可能地将一个药物分子图量化,而在引入 3D 信息后,自注意力机制让多个维度的分子图节点相互注意,使模型从空间中学习到了预测所需的作用子结构。在 DrugBank 数据集上进行的三个实验都证明了这一点,无论是在直推式学习中,还是在归纳学习中,本文方法相比过去的预测模型,取得了不错的效果。特别是在新药与新药的相互作用预测中,本文方法具备突出的预测能力,可以为药物研发过程中 DDI 的研究提供帮助,具备一定的实际价值。

在未来的工作中,进一步关注药物三维信息的融合方法,寻求更加有效的方式拟合药物的空间作用子结构,以此提升模型的预测能力;同时根据注意力矩阵探究药物分子内部空间子结构特征,具体的拟合过程可以很好地帮助模型获得可解释性,从而为 DDI 的研究提供服务。

参考文献(References):

- [1] JIA J, ZHU F, MA X, et al. Mechanisms of drug combinations: interaction and network perspectives[J]. *Nature Reviews Drug Discovery*, 2009, 8(2): 111-128.
- [2] 王斯, 张国浩, 陈义安. 基于 GWO-KELM 与 GBDT 的抗乳腺癌药物性质预测[J]. *重庆工商大学学报(自然科学版)*, 2023, 40(6): 93-104.
WANG Si, ZHANG Guo-hao, CHEN Yi-an. Prediction of anti-breast cancer drug properties based on GWO-KELM and GBDT[J]. *Journal of Chongqing Technology and Business University (Natural Science Edition)*, 2023, 40(6): 93-104.
- [3] SINDELAR M, MCCABE D, CARROLL E. Tacrolimus drug-drug interaction with nirmatrelvir/ritonavir (paxlovid™) managed with phenytoin[J]. *Journal of Medical Toxicology*, 2023, 19(1): 45-48.
- [4] PRESKORN S H. Comparative pharmacology of the 3 marketed dual orexin antagonists-daridorexant, lemborexant, and suvorexant: Part 2. principal drug metabolizing enzyme, drug-drug interactions, and effects of liver and renal impairment on metabolism[J]. *Journal of Psychiatric Practice*, 2023, 29(1): 38-41.
- [5] CHAN W J J, ADIWIDJAJA J, MCLACHLAN A J, et al. Interactions between natural products and cancer treatments: Underlying mechanisms and clinical importance[J]. *Cancer Chemotherapy and Pharmacology*, 2023, 91(2): 103-119.
- [6] DALLE FRATTE C, GAGNO S, RONCATO R, et al. CYP2D6 and CYP2C8 pharmacogenetics and pharmacological

- interactions to predict imatinib plasmatic exposure in GIST patients[J]. *British Journal of Clinical Pharmacology*, 2023, 89(3): 1089–1098.
- [7] VILAR S, HARPAZ R, URIARTE E, et al. Drug-drug interaction through molecular structure similarity analysis[J]. *Journal of the American Medical Informatics Association*, 2012, 19(6): 1066–1074.
- [8] GUO L, LEI X, CHEN M, et al. MSResG: Using GAE and residual GCN to predict drug-drug interactions based on multi-source drug features[J]. *Interdisciplinary Sciences, Computational Life Sciences*, 2023, 15(2): 171–188.
- [9] BALAMURALIDHAR N, SURENDRAN P, SINGH G, et al. Prediction of drug interactions using graph-topological features and gnn[C]//IFIP International Conference on Artificial Intelligence Applications and Innovations. Cham: Springer Nature Switzerland, 2023: 135–144.
- [10] 侯磊, 刘金环, 于旭, 等. 图神经网络研究综述[J/OL]. *计算机科学*, 2023: 1–25. <https://kns.cnki.net/kcms/detail/50.1075.TP.20231010.1541.022.html>.
HOU Lei, LIU Jin-huan, YU Xu, et al. Review of graph neural networks[J/OL]. *Computer Science*, 2023: 1–25. <https://kns.cnki.net/kcms/detail/50.1075.TP.20231010.1541.022.html>.
- [11] 甘宇雨, 余彦佳, 刘勇. CBSG-PPI: 基于图神经网络的蛋白质-蛋白质相互作用的预测算法[J/OL]. *基因组学与应用生物学*, 2024: 1–12. <https://kns.cnki.net/kcms/detail/45.1369.Q.20240108.0925.002.html>.
GAN Yu-yu, YU Yan-jia, LIU Yong. CBSG-PPI: A prediction algorithm for protein-protein interaction based on graph neural networks[J/OL]. *Genomics and applied biology*, 2024: 1–12. <https://kns.cnki.net/kcms/detail/45.1369.Q.20240108.0925.002.html>.
- [12] LIN X, QUAN Z, WANG Z J, et al. Kgnn: Knowledge graph neural network for drug-drug interaction prediction[C]//IJCAI. 2020, 380: 2739–2745.
- [13] HONG Y, LUO P, JIN S, et al. LaGAT: link-aware graph attention network for drug-drug interaction prediction[J]. *Bioinformatics*, 2022, 38(24): 5406–5412.
- [14] ASADA M, MIWA M, SASAKI Y. Integrating heterogeneous knowledge graphs into drug-drug interaction extraction from the literature[J]. *Bioinformatics*, 2023, 39(1): 754–754.
- [15] 孙泽旭, 赵楠, 解染, 等. 利托那韦生理药代动力学模型的建立及其介导的药物相互作用预测[J]. *中国临床药理学杂志*, 2024, 40(2): 259–263.
SUN Ze-xu, ZHAO Nan, XIE Ran, et al. Establishment of physiological based pharmacokinetic model of ritonavir and prediction of its drug-drug interactions[J]. *The Chinese Journal of Clinical Pharmacology*, 2024, 40(2): 259–263.
- [16] 杨双萌, 于江, 侯文彬, 等. 人工智能算法用于药物研发的研究进展[J]. *现代药物与临床*, 2023, 38(12): 3150–3160.
YANG Shuang-meng, YU Jiang, HOU Wen-bin, et al. Research progress on artificial intelligence algorithms for drug discovery[J]. *Drugs & Clinic*, 2023, 38(12): 3150–3160.
- [17] SILVERMAN R B, HOLLADAY M W. *The organic chemistry of drug design and drug action*[M]. USA: Academic press, 2014.
- [18] YU H, WANG J, ZHAO S Y, et al. GGI-DDI: Identification for key molecular substructures by granule learning to interpret predicted drug-drug interactions[J]. *Expert Systems with Applications*, 2024, 240: 122500.
- [19] XU N, WANG P, CHEN L, et al. MR-GNN: Multi-resolution and dual graph neural network for predicting structured entity interactions[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. California: International Joint Conferences on Artificial Intelligence Organization, 2019: 3968–3974.
- [20] NYAMABO A K, YU H, SHI J Y. SSI-DDI: Substructure-substructure interactions for drug-drug interaction prediction[J]. *Briefings in Bioinformatics*, 2021, 22(6): 133–145.
- [21] NYAMABO A K, YU H, LIU Z, et al. Drug-drug interaction prediction with learnable size-adaptive molecular substructures[J]. *Briefings in Bioinformatics*, 2022, 23(1): 441–453.
- [22] YANG Z, ZHONG W, LV Q, et al. Learning size-adaptive molecular substructures for explainable drug-drug interaction prediction by substructure-aware graph neural network[J]. *Chemical Science*, 2022, 13(29): 8693–8703.
- [23] LI Z, ZHU S, SHAO B, et al. DSN-DDI: An accurate and generalized framework for drug-drug interaction prediction by dual-view representation learning[J]. *Briefings in Bioinformatics*, 2023, 24(1): 591–597.
- [24] XU K, HU W, LESKOVEC J, et al. How powerful are graph neural networks? [C]//International Conference on Learning Representations, 2018: 102–122.
- [25] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//NIPS. 2017: 5998–6008.
- [26] ANDREEA D, YU-HSIANG H, PETAR V, et al. Drug-drug adverse effect prediction with graph co-attention [C]//ICML Workshop on Computational Biology, New Orleans, Louisiana, United States: ICLR. 2019: 47–55.
- [27] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks [EB/OL]. 2017: 1710.10903. <http://arxiv.org/abs/1710.10903v3>.

责任编辑:代小红