



太原理工大学学报  
*Journal of Taiyuan University of Technology*  
ISSN 1007-9432, CN 14-1220/N

## 《太原理工大学学报》网络首发论文

题目： 基于注意力感知和模态融合的药物-靶点相互作用预测方法  
作者： 彭杨, 朱小飞, 胡冬冬  
网络首发日期： 2024-03-14  
引用格式： 彭杨, 朱小飞, 胡冬冬. 基于注意力感知和模态融合的药物-靶点相互作用预测方法[J/OL]. 太原理工大学学报.  
<https://link.cnki.net/urlid/14.1220.N.20240313.1646.004>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 基于注意力感知和模态融合的药物-靶点相互作用预测方法

彭杨, 朱小飞, 胡冬冬

(重庆理工大学 计算机科学与工程学院, 重庆 400054)

**摘要:** 预测药物-靶点相互作用是药物发现的关键步骤。现有的基于图神经网络的预测方法取得了较好的效果, 但是仍然存在两方面挑战: 一是如何更高效地提取药物和靶点中深层次的特征和丰富的语义信息, 二是如何明确地建模并学习药物和靶点之间的相互作用, 以便更好地预测和解释。为了解决以上问题, 本文提出了基于注意力感知和模态融合的药物-靶点相互作用预测方法。对于药物分支, 利用 Graph Transformer 和图卷积神经网络来联合表征药物分子的全局结构和生化信息。对于蛋白质分支, 采用了一个独立的编码模块来提取蛋白质的全局和局部特征。为了更准确的建模和学习药物和靶点之间的作用关系, 设计了一个基于 Transformer 的模块来融合药物和靶点特征。最后, 在三个公开数据集上的实验结果表明, 在主要评价指标上均有所提升, 该方法能够更加准确的预测药物和靶点的相互作用。

**关键词:** 药物发现; 图神经网络; 模态融合; 注意力机制

**中图分类号:** TP 391

**文献标志码:** A

## A Drug-target Interaction Prediction Method Based on Attention Perception and Modality Fusion

PENG Yang, ZHU Xiaofei, HU Dongdong

(College of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China)

**Abstract:** Predicting drug-target interactions is a critical step in drug discovery. The existing prediction methods based on graph neural networks have achieved good results, but there are still two challenges: first, how to extract the deep features and rich semantic information in drugs and targets more efficiently, and second, how to explicitly model and learn the interactions between drugs and targets for better prediction and interpretation. In order to solve the above problems, this paper proposes a drug-target interaction prediction method based on attention perception and modality fusion. For drug branches, Graph Transformer and Graph Convolutional Neural Network were used to jointly characterize the global structure and biochemical information of drug molecules. For protein branches, a separate encoding module was employed to extract the global and local features of the protein. In order to more accurately model the relationship between drugs and targets, a Transformer-based module was designed to fuse drug and target features. Finally, the experimental results on the three public datasets show that the main evaluation indicators have been improved, and the method can more accurately

predict the interaction between drugs and targets.

**Keywords:** drug discovery; graph neural networks; modality fusion; attention mechanisms

药物-靶点相互作用 (Drug Target Interaction, DTI) [1] 表示药物与靶点 (蛋白质) 之间的结合作用关系, 识别药物-靶点相互作用是药物发现过程 [1] 中的重要步骤。虽然传统的生物医学体外实验测量筛选是可靠的, 但是该方法具有昂贵的开发成本和耗时的开发周期, 这阻碍了其在大规模数据 [2] 上的应用。相比之下, 通过计算机模拟方法识别药物-靶点对可以大大缩小化合物候选物的搜索范围, 并提供药物组合中潜在副作用原因的解释。因此, 计算机模拟方法在近几年里得到了越来越多的关注, 并取得了较大的进展。

最近, 基于深度学习 (Deep Learning, DL) 的方法在各个领域的成功, 在 DTI 预测方面也取得了较快的发展 [3]。目前大部分研究都是从化学基因组学的角度构建的, 它将化学空间、基因组空间和相互作用信息集成到一个统一的端到端的框架。由于具有可用三维结构的生物靶点的数量有限, 许多基于 DL 的模型将药物和蛋白质的线性或二维结构信息作为输入。将 DTI 预测视为二分类任务, 并通过将输入信息馈送到不同的深度编码和解码模块来进行预测, 例如深度神经网络 (Deep Neural Networks, DNN) [4, 5]、图神经网络 (Graph Neural Networks, GNN) [6, 7] 和 Transformer [8] 架构。随着深度学习技术的进步, 这类模型可以从大规模 DTI 数据中自动学习药物和蛋白质的数据驱动 [11] 表示, 而不仅仅是使用预定义的标签和描述符。

现有的基于 DL 的方法仍然存在两方面挑战, 一是如何更高效地提取药物和靶点中深层次的特征和丰富的语义信息。图神经网络已经成为从图结构数据中进行表征学习 [11] 的最有效的方法之一, 图卷积神经网络 (Graph Convolutional Networks, GCN) [12] 通过聚合来自每个节点的邻域的特征直接在图上定义卷积, 这有利于泛化到未知的节点。图注意力网络 (Graph Attention Networks, GAT) [13] 通过在特征聚合步骤分配具有不同重要性的相邻节点, 将注意力机制推广到图数据中。药物分子的表征过程如图 1 所示, 基于药物分子的图表示学习为每个药物分子定义了一个图结构, 能够更完整、稳定地表示原子特征和原子间的化学键, 有助于模型准确预测分子的生化性质。然而, 现有的大多数基于 GNN 的

方法直接使用节点信息进行建模, 忽略了分子内部的化学键。事实上, 化学药物分子中的大多数原子具有不同的标记或特征, 一个分子中化学键的类型是不同的。尽管现有的研究取得了不错的进展, 但在现有的基于图神经网络的预测方法中仍然存在一些挑战。蛋白质由多个氨基酸组成, 基于卷积神经网络 (Convolutional Neural Network, CNN) [14] 和递归神经网络 (Recurrent Neural Network, RNN) [15] 的编码器被广泛用于氨基酸序列建模。虽然 CNN 具有  $O(n)$  的时间和空间复杂度, 并且可以处理长序列数据, 但其主要用于提取序列的局部特征, 其卷积核的单个操作只能用于序列中的小范围计算。与句子相比, 氨基酸序列特征具有更难处理的模式, 以及更长的范围。精确地捕获序列特征需要更全面的学习架构来从整个序列中提炼序列信息, 由于 CNN 的感受野在序列的中间扩展得较快, 而在两侧扩展得更慢, CNN 不能公平地利用所有位置的信息, 导致序列位置对卷积的输出产生影响。

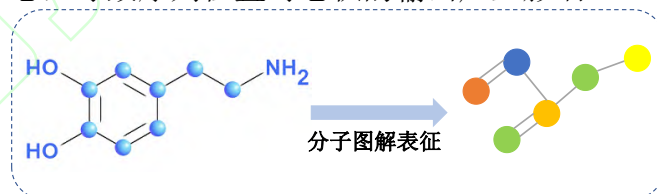


图 1 药物分子的图解表征过程

Fig.1 Graphical characterization of drug molecules

第二个挑战是明确学习药物和蛋白质的局部结构之间的相互作用, DTI 基本上由药物化合物中的重要分子子结构和氨基酸序列中结合位点之间的相互作用决定的。然而, 许多先前的研究在其分别编码后直接学习全局表示, 而没有明确地学习局部作用点位之间的交互, 并且仅在黑盒解码模块中隐含地学习互信息。药物和靶点之间的关系与其子结构特别相关, 因此单独的全局表示学习往往会限制建模能力和预测性能。此外, 如果没有对局部相互作用的明确学习, 预测结果很难解释。

针对以上问题, 本文提出了一个基于注意力感知和模态融合的神经网络 (Attention Perception and Modality Fusion, APMF)。主要贡献可以总结为以下几个方面:

(1) 对于药物分子的表示, 引入原子间的共价键来编码分子内化学键的信息, 利用 Graph

Transformer 和 GCN 来联合节点信息和分子内化学键对药物分子进行表征。

(2) 针对蛋白质靶点的表示, 使用了一个独立的蛋白质特征提取模块, 该模块利用了强大的局部表征和上下文信息提取能力, 提高了模型

表达氨基酸中复杂依赖性的能力。

(3) 针对药物和靶点两个模态, 设计了一个基于 Transformer 的融合模块, 将两个分支的输出以多模态融合的方式进行连接, 最后解码输出预测结果。

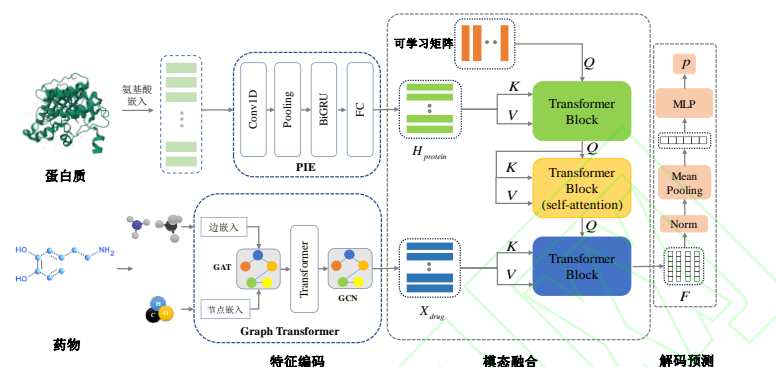


图 2 APMF 整体模型框架

Fig.2 Overall framework of APMF model

## 1 相关工作

DTI 被描述为药物分子与靶点 (通常是蛋白质) 的结合, 其中药物抑制或激活靶点以治疗疾病。DTI 的准确鉴定可以大大加速药物发现过程。大多数工作仅采用化学和基因组数据进行 DTI 预测, 通过以字符串、分子图或复合物的形式建模药物和靶点。然而, 相对有限的知识之间的复杂关系的化学空间和基因组空间限制了 DTI 预测。

随着基因组学、蛋白质组学、转录组学、代谢组学等生物技术的快速发展, 庞大的生物数据库资源可以公开获取, 这为 DTI 预测提供了丰富的信息。基于这些数据, 许多研究人员进行了 DTI 预测任务的研究。DrugVQA<sup>[16]</sup> 采用了距离结构来表示蛋白质的特征, 借助语言模型的思想进行序列建模预测; 对于特征建模, DeepDTA<sup>[17]</sup> 和 DrugVQA 以字符为基本单位为序列和靶点序列建立字典, 每个字符对应一个数字, 然后从该字典中构建每个向量表示, 并分别使用卷积神经网络 (CNN) 和双向长短时记忆网络 (Bi-directional Long Short-Term Memory ,

BiLSTM) 进行特征学习和表示; TransformerCPI<sup>[18]</sup> 和 Moltrans<sup>[10]</sup> 则是采用 Transformer 架构对 DTI 任务进行建模, 两者分别对药物和靶点以字符级别和子结构级别进行特征建模, 并结合了生物学领域知识, 将原子的多种特征合并为序列嵌入, 前者使用 word2vec 获得蛋白质的序列特征, 后者将药物和靶点子结构看作同一空间的特征, 但两种方法都采用相同的编码器进行特征提取; 基于 GCN 或 GAT 的模型从多个药物和靶点相关网络中提取药物和靶点表示, GraphDTA<sup>[7]</sup> 和 HGAN<sup>[8]</sup> 利用分子的分子图形式, 使用卷积神经网络和图卷积神经网络提取药物特征并预测药物-靶点相互作用; 对于靶点特征, GraphDTA 使用一维序列嵌入并利用 CNN 提取蛋白质的潜在特征; HGAN 提出了使用异构图注意网络来捕捉生物异构图中复杂的结构和丰富的语义, 以用于 DTI 预测; DrugBAN<sup>[19]</sup> 是一个深度双线性注意力网络框架, 具有域适应性, 用于学习药物和靶点之间的成对局部相互作用。深度学习能够进行非线性的变换, 捕捉隐性特征, 挖掘药物-靶点之间的隐性关联, 还可以降低对特征选择的依赖, 弥补传

统机器学习方法的不足。

## 2 方法与模型

### 2.1 问题定义

在 DTI 预测中,任务是确定药物化合物和靶点是否会相互作用。对于靶点,将每个蛋白质表示为  $\mathcal{P}=(a_1, \dots, a_n)$ , 其中每个标记  $a_i$  代表 23 种氨基酸中的一种。对于药物化合物,现有的基于深度学习的方法用简化分子输入行输入系统 (Simplified Molecular Input Line Entry System, SMILES) [20] 表示。SMILES 是描述药物分子中的化学原子和边信息的一维序列。具体来说,一个药物分子图定义为  $\mathcal{G}=(\mathcal{V}, \mathcal{E})$ , 其中  $\mathcal{V}$  是顶点 (原子) 的集合,  $\mathcal{E}$  是边 (化学键) 的集合。给定一个序列  $\mathcal{P}$  和一个药物分子图  $\mathcal{G}$ , DTI 预测旨在学习一个模型  $\mathcal{M}$ , 将特征联合表示  $\mathcal{P} \cdot \mathcal{G}$  映射到相互作用概率得分  $p \in [0, 1]$ 。

### 2.2 APMF模型框架

本文提出了一个基于注意力感知模态融合的深度模型来预测药物靶点相互作用,由于任务的特异性,存在药物和蛋白质特征编码、特征相互作用建模和解码预测三个模块。为了提高模型的可扩展性,设计了基于药物和蛋白质两个独立的特征表示分支。具体来说,利用带有边嵌入模块的 Graph Transformer 提取药物分子图的高阶语义,利用蛋白质独立编码模块提取蛋白质的氨基酸序列组信息,然后将两类特征融合,最后通过解码模块预测相互作用概率,模型如图 2 所示。

### 2.3 药物编码器

先前的工作联合 GCN 和 GAT 用于 DTI 预测 [21], 基于该方法,对于药物化合物,将每个 SMILES 字符串转换为二维的分子图  $\mathcal{G}$ 。初始化药物分子 SMILES 字符串的图以获得邻接矩阵  $A \in \mathbb{R}^{n \times n}$ , 以及原子节点的特征矩阵  $X \in \mathbb{R}^{n \times d}$ ,  $n$  是图中的节点数,也称为药物分子中的原子数。使用  $W_{edge} \in \mathbb{R}^{m \times m}$  通过  $E_{Edge} \in \mathbb{R}^{m \times d}$  将图的边映射成  $E_{edge} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m\}$ ,  $\vec{e}_i \in \mathbb{R}^d$ 。

为了提取分子图的全局信息,引入了多头注意力编码器层来编码每个节点表示。在本文中,多头注意力编码器层是由几个自注意力来提取图的信息:

$$Attention(X) = \text{soft max} \left( \frac{(XW_i^Q)(XW_i^K)}{\sqrt{d}} \right) XW_i^V \quad (1)$$

等式(1)中  $W_i^Q \in \mathbb{R}^{d \times d}$ ,  $W_i^K \in \mathbb{R}^{d \times d}$ ,  $W_i^V \in \mathbb{R}^{d \times d}$ 。  $i$  是多头的索引,在进行多头注意力提取之后,获得节点矩阵的表示  $X_1 \in \mathbb{R}^{n \times d}$ 。

然后,采用 GAT 在模型中进行第一次消息传递。将  $X_1$  作为 GAT 的节点输入,并使用边特征  $E_{Edge}$  表示邻居结点的依赖关系,两个节点  $i$  和  $j$  之间注意力得分计算如下:

$$e_{ij} = \sigma(a^T [W_G \vec{x}_i \parallel W_G \vec{x}_j]) \quad (2)$$

其中  $a$  是单层前馈神经网络,  $\sigma$  使用 LeakyReLU 作为激活函数,  $a \in \mathbb{R}^{2d}$ 。使用边特征  $E_{Edge}$  作为权重共享  $W_G \in \mathbb{R}^{n \times n}$  来表示相邻节点的依赖关系,  $e_{ij}$  表示节点  $j$  和  $i$  之间的重要性,然后使用 softmax 函数归一化节点  $i, j$  的相关系数:

$$\alpha_{ij} = \text{softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{j \in \mathcal{N}_i} \exp(e_{ik})} \quad (3)$$

$j \in \mathcal{N}_i$  是图中邻居节点的集合,来自每个头的聚合特征被拼接以获得:

$$\vec{x}_i^u = \sigma \left( \sum_{k \in \mathcal{N}_j} \alpha_{ij}^k W^k \vec{x}_j \right) \quad (4)$$

其中  $\sigma(\cdot)$  是 LeakyReLU 非线性激活函数,  $k \in (1, K)$  是多头的索引,  $K$  是多头的总数。

从 GAT 输出的节点特征表示为  $X_{GAT} = \{\vec{x}_1^u, \vec{x}_2^u, \dots, \vec{x}_n^u\}$ , 再使用等式(1)来获得节点的表示  $X_2 = \{\vec{z}_1^u, \vec{z}_2^u, \dots, \vec{z}_n^u\}$ , 图的节点表示  $X_2$  的消息聚合由 GCN 输出的节点得到最终的药物表示  $X_{Drug} = \{\vec{z}_1, \vec{z}_2, \dots, \vec{z}_n\}$ :

$$\vec{z}_i = \sigma \left( \sum_{j \in \mathcal{N}_i} W \vec{z}_j^u \right) \quad (5)$$

其中  $W \in \mathbb{R}^{h \times h}$  是线性变换矩阵,  $\sigma(\cdot)$  也是非线性激活函数。

### 2.4 蛋白质编码器

Chen 等人 [22] 的工作已经证明,基于氨基酸序列的蛋白质特征表示有利于提高蛋白质相互作用预测任务的性能,受此启发,本文采用蛋白质独立编码器 (Protein Independent Encoding, PIE) 来捕获蛋白质的潜在语义特征。PIE 利用全局序列信息和局部特征。该深度神经编码器堆叠两个计算模块的多个实例,即具有池化和双向残差门控递归单元的卷积层。

PIE 由连续的一维卷积层组成,其将输入氨

氨基酸序列转换为潜在特征空间中的矩阵表示。矩阵的每一行表示蛋白质中的氨基酸的表示。借鉴词嵌入的概念, 首先将所有氨基酸初始化为一个可学习的嵌入矩阵  $E_p \in \mathbb{R}^{23 \times h}$ , 其中 23 是所有氨基酸类型的数量,  $h$  是潜在空间维度。通过查询  $E_p$ , 每个蛋白质可以被初始化为对应的特征矩阵  $H_p \in \mathbb{R}^{\Theta \times h}$ , 其中  $\Theta = 735$  是序列的最大允许长度, 其被设置为比对不同的氨基酸长度并进行批量训练。氨基酸序列的超过最大允许长度被切断, 具有较小的长度用零填充。蛋白质编码器描述如下编码器描述如下:

$$H_p^{(l+1)} = \text{Conv}(W_c^{(l)}, b_c^{(l)}, H_p^{(l)}) \quad (6)$$

$W_c^{(l)}$  和  $b_c^{(l)}$  是第  $l$  层中的可学习权重矩阵和偏置向量,  $H_p^{(l)}$  是蛋白质的第  $l$  层隐藏层表示, 经过池化后得到更高级的蛋白质嵌入  $H^v$ , 双向 GRU 层在两个方向上表征氨基酸序列信息。包含前向编码过程  $\overrightarrow{GRU}$ , 用来读取输入向量序列  $H^v = \{\overrightarrow{h_1^v}, \overrightarrow{h_2^v}, \dots, \overrightarrow{h_t^v}\}$  的  $\overrightarrow{h_1^v}$  到  $\overrightarrow{h_t^v}$ , 以及反向编码过程  $\overleftarrow{GRU}$ , 其读取方向相反。两个过程的编码结果拼接每个输入项  $\overrightarrow{h_t^v}$ , 计算如下所示:

$$\overrightarrow{h_t^v} = [\overrightarrow{GRU}(\overrightarrow{h_t^v}) + \overrightarrow{h_t^v} \parallel \overleftarrow{GRU}(\overrightarrow{h_t^v}) + \overrightarrow{h_t^v}] \quad (7)$$

PIE 交替地堆叠以上两个中间神经网络组件。卷积层用作第一编码层以从输入序列提取局部特征。残差 GRU 层接收保留的局部特征, 其输出被传递到全连接层。在网络结构中重复这两个组件对氨基酸序列进行多粒度特征聚合过程, 同时保留所选特征的每个粒度上的顺序和上下文化信息。最后再次使用全连接层应用于先前的序列来获得最终的蛋白质特征表示  $H_{protein} = \{\overrightarrow{h_1}, \overrightarrow{h_2}, \dots, \overrightarrow{h_t}\}$ 。

## 2.5 模态融合预测

在视觉问答(Visual Question Answering, VQA)<sup>[16]</sup>任务中。给定图像和相关自然语言问题, VQA 任务旨在提供文本-图像匹配答案。与 VQA 问题类似, DTI 任务可以被视为多模态任务以适应多模态学习, 受 Transformer 架构<sup>[10]</sup>和 Perceiver IO<sup>[23]</sup>模型启发, 提出了一种基于 Transformer 架构的模块来融合药物和靶点特征, 如图 3 所示, 模型考虑每对多模态输入通道, 即药物分子和氨基酸序列的交互学习表示。

模态融合的过程包含三个阶段: 第一阶段采用高斯分布采样初始化一个可学习矩阵, 该

矩阵用来捕获蛋白质作用点位的特征表示, 通过注意力机制关注蛋白质作用点位, 得到可学习的蛋白质结合位点的加权表示。在第二阶段, 来自第一阶段的加权表示通过自注意进行进一步的细化。经过处理的 query 包含了蛋白质结合点位的细粒度特征, 为与药物分子的相互作用建模做准备。在第三阶段, 来自前一步骤的已处理的 query 充当 Transformer Block 中的  $Q$ , 药物分子表示同时充当  $K$  和  $V$ 。

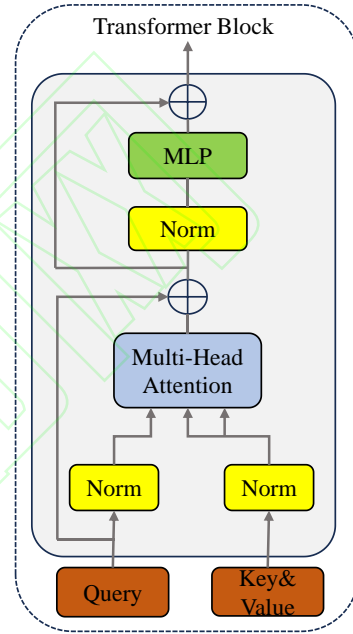


图 3 Transformer Block 框架

Fig.3 Overall framework of Transformer Block

这种设置使模型能够专注于发现与蛋白质结合位点最相关的药物子结构, 从而得到药物-蛋白质相互作用的联合表示  $F$ :

$$\text{Attention}(Q, K, V) = \text{softmax} \left( \frac{QK^T}{\sqrt{d_k}} \right) V \quad (8)$$

其中, 使用线性变换矩阵  $WQ, WK, WV$  计算得到  $Q, K, V$ 。经过层归一化和池化过后, 得到联合向量的表示  $\vec{f}$ , 最后经过多层感知机 (Multi-Layer perceptron, MLP) 分类后得到最后的预测结果  $p$ , 经过分类后的交叉熵损失, 其定义为:

$$\mathcal{L}_{CE} = - \sum_{i=1} y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \quad (9)$$

其中,  $y_i$  表示第  $i$  个药物-靶点对的事实标签,  $p_i$  是模型的输出概率。最后, 通过反向传播联合优化所有可学习的参数。

### 3 实验

#### 3.1 实验设置

##### 3.1.1 数据集

表 1 实验数据集

Table 1 Experimental dataset statistics

数据集	#药物	#蛋白质	#作用对
BindingDB	14,643	2,623	49,199
BioSNAP	4,510	2,181	27,464
Human	2,726	2,001	6,728

本文在 3 个公开的 DTI 数据集：BindingDB<sup>[24]</sup>，BioSNAP<sup>[25]</sup>和 Human<sup>[26]</sup>上评估了 APMF 和 6 个基线模型的性能。BindingDB 数据集是一个可通过网络访问的数据库，包含了实验验证的结合亲和力，主要关注药物小分子和蛋白质的相互作用，BioSNAP 数据集由 Huang<sup>[10]</sup>等人从 DrugBank 数据库创建，它是一个平衡的

表 2 BindingDB 和 BioSNAP 数据集的性能比较

Table 2 performance comparison on the BindingDB and BioSNAP datasets

Dataset	Model	AUROC	AUPRC	Accuracy	Sensitivity	Specificity
BindingDB	SVM	0.939	0.928	0.825	0.781	0.886
	RF	0.942	0.921	0.880	0.875	0.892
	DeepConv-DTI	0.945	0.925	0.882	0.873	0.894
	GraphDTA	0.951	0.934	0.888	0.882	0.897
	MolTrans	0.952	0.936	0.887	0.877	0.902
	DrugBAN	0.960	0.948	0.904	<b>0.900</b>	0.908
	APMF	<b>0.964</b>	<b>0.954</b>	<b>0.905</b>	0.895	<b>0.919</b>
BioSNAP	SVM	0.862	0.864	0.777	0.711	0.841
	RF	0.860	0.886	0.804	0.823	0.786
	DeepConv-DTI	0.886	0.890	0.805	0.760	0.851
	GraphDTA	0.887	0.890	0.800	0.745	0.854
	MolTrans	0.895	0.897	0.825	0.818	0.831
	DrugBAN	0.903	0.902	0.834	0.820	0.847
	APMF	<b>0.909</b>	<b>0.915</b>	<b>0.843</b>	<b>0.830</b>	<b>0.855</b>

##### 3.1.3 基线模型

APMF 和六个基线模型上在 DTI 预测性能上进行了比较，两种浅层机器学习方法：(1) 支持向量机 (Support Vector Machine, SVM)<sup>[27]</sup>；(2) 随机森林 (Random Forest, RF)<sup>[28]</sup>来编码药物分子特征；(3) DeepConv-DTI<sup>[5]</sup>，其使用 CNN 和一个全局最大池化层来提取蛋白质序列中的局部特征，并使用全连接网络来编码药物特征；(4) GraphDTA<sup>[7]</sup>使用图神经网络对 DTI 进行建模，以编码药物分子图，并使用 CNN 对蛋白质序列进行编码。将药物表示向量和蛋白质表示向量简单的级联组合，然后进行预测；(5) MolTrans<sup>[10]</sup>，一个 Transformer 架构的深度学习

数据集，包括经过验证的相互作用的正样本和从未知的药物-靶点对中随机获得的相等数量的负样本。Human 数据集由 Liu 等人<sup>[26]</sup>创建，通过计算机筛选方法获得高可信度的阴性样品，表 1 是实验数据集统计。

##### 3.1.2 实验设计

APMF 在 Python3.8 和 Pytorch 1.7.1 中实验，batch\_size 大小设置为 64，Adam 优化器的学习率设置为 5e-5。允许模型在所有数据集上运行最多 100 个 epoch。选择在验证集上给出最佳 AUROC 分数的模型，然后使用该模型来评估测试集上的最终性能。训练模型使用的设备信息是：Intel(R) Core(TM) i7-9700K，GeForce RTX 2080 Ti。

模型，用于编码药物和蛋白质信息，以及一个基于 CNN 的交互式模块，用于学习子结构相互作用；(6) DrugBAN<sup>[19]</sup>，一个深度双线性注意力网络框架，具有域自适应功能，可显式学习药物和靶点之间的成对局部相互作用。对于上述深度 DTI 模型，在实验中遵循其论文中推荐的模型超参数设置。

##### 3.1.4 评价策略和衡量标准

本文分别在 3 个公共数据集上研究分类性能，在测试集中保持药物-靶点对的标签“未知”以进行评估。每个实验数据集以 7:1:2 的比例随机划分为训练集、验证集和测试集。每次数据集划分使用不同的随机种子，重复进行五次独立的

实验。每次选择在验证集上具有最佳 AUROC 的模型, 然后在测试集上评估选定的模型, 最后取平均值以报告性能度量。

实验采用了受试者工作特征曲线下的面积 (area under the receiver operating characteristic curve, AUROC) 和精确度-调用曲线下的面积 (area under the precision-call curve, AUPRC) 被用作评估模型分类性能的主要度量。此外, 为了全面的评估模型的性能和泛化能力, 还报告了最佳 F1 评分阈值的准确性 (Accuracy)、灵敏度 (Sensitivity) 和特异性 (Specificity)。

## 3.2 实验结果与分析

### 3.2.1 总体实验性能

表 2 是 APMF 在 BindingDB 和 BioSNAP 数据集上的性能比较。模型在 AUROC、AUPRC 和准确性等各个评价指标上始终优于基线模型, 而且在灵敏度和特异性方面的表现也具有较好的竞争力。结果表明, 在 DTI 预测中, 数据驱动的表达学习可以捕获比预定义的描述符特征更重要的信息。此外, APMF 可以通过其成对模态融合作用模块捕获局部结构之间的相互作用模式, 进一步提高预测性能。

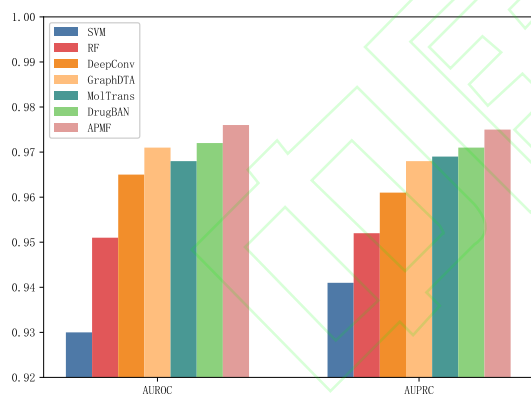


图 4 Human 数据集上的性能比较

Fig.4 performance on the Human dataset

图 4 是 Human 数据集上的结果。在随机划分下, 基于深度学习的模型都实现了较好的性能 (AUROC>0.96)。然而, Liu 等人<sup>[29]</sup>指出, Human 数据集有一些隐藏的配体偏差, 导致正确的预测是仅基于药物特征而不是相互作用模式。高准确性可能是由于偏差和过拟合, 而不是模型在预测方面的真实表现。但是, 从实验结果可以看到 APMF 相比于其他的深度学习基线, 仍然实现了较好的性能。

### 3.2.2 消融实验

本研究进行了消融实验, 以研究注意力感知和模态融合的神经网络的效果。结果在表 3 中,

为了验证深度特征提取模块的有效性, 研究了 APMF 的 3 种变体: 不使用模态融合 (w/o fusion), 对池化之后的药物和蛋白质表示的向量进行拼接后直接预测; 单侧药物特征表示网络 (w/o protein) 和单侧蛋白质特征表示网络 (w/o drug), 单侧特征表示是 Tsubaki 等<sup>[29]</sup>提出的神经网络机制, 将 APMF 中的双侧特征表示替换为单侧特征表示以生成两个变体。变体之间在药物和蛋白质之间的联合表示计算方面有所不同。

表 3 BindingDB 和 BioSNAP 数据集上的 AUROC 消融研究

Table 3 Ablation study in AUROC on the BindingDB and BioSNAP datasets

Model	BindingDB	BioSNAP
w/o fusion	0.941	0.883
w/o drug	0.950	0.890
w/o protein	0.953	0.892
APMF	<b>0.964</b>	<b>0.909</b>

如表 3 所示, 实验发现, 虽然双侧特征表示的结构表明两个分支的特征在融合之前没有耦合, 但结果表明它们之间存在明显的相互作用, 使用单侧表征蛋白质序列的有效性仅在另一分支中使用 Graph Transformer 时才表现出来, 这意味着双侧深度特征提取是用于 DTI 相互作用预测的最有效方法。从模型中移除模态融合模块时。模型中两个数据集上的性能均下降明显, 由此可知, 提出的模态融合模块有助于提高预测的性能。

### 3.2.3 超参数设置和分析

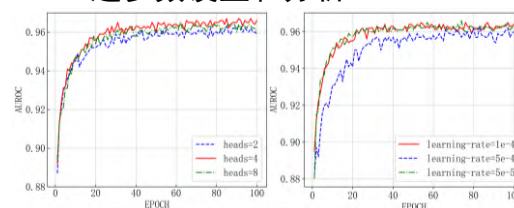


图 5 BindingDB 数据集上不同超参数的学习率曲线

Fig.5 Learning curves with the different choices of hyperparameters on the BindingDB

在所有实验数据集上使用相同的超参数, 图 5 是 BindingDB 数据集上选择不同超参数下的 AUROC 的性能, 包括注意力头数和学习率, 通过实验结果表明它们的性能差异不大, 通常在 30 到 40 个 epoch 之间收敛。这是由于在较低维度特征空间中映射序列更容易学习丰富的特征信息, 而且不增加时间复杂度。同时也表明多头机制可以显著降低特征提取的难度, 从实验中可知, 当多头数目设置为 4 时, 预测性能最好, 因为较少的注意头可能更有助于模型提取更通用



和抽象的特征,而过多的注意头可能导致模型过度关注输入中的细节和噪声,而忽略了更重要的特征和及作用点位信息。合适的注意头可以更好地捕捉输入之间的相关性,帮助模型更好地泛化到新的样本。

### 3.2.4 对不可见药物/靶点的性能比较

为了研究 APMF 和其他深度学习基线在不可见的药物/靶点上的性能表现,在 BindingDB 和 BioSNAP 上进行了额外的实验。对于每个数据集,实验随机选择 20% 的药物/靶点蛋白。然后评估与这些药物/靶蛋白相关的所有 DTI 对的预测性能(70%作为评估的测试集,30%作为验证集),剩余药物靶点对作为用于模型优化的训练集。每个实验设置有五次独立的运行实验,由表 4 可知,测试集的 AUROC 结果分别包含随机划分(Random Split, RS)、未见药物(Unseen Drug, UD)、未见靶点(Unseen Target, UT)的结果,在两个数据集性能提升明显。实验结果表明 APMF 在 6 个实验设置中最优,实现了较好性能。

表 4 在随机划分、未知药物和未知靶点设置下

BindingDB 和 BioSNAP 数据集的性能

Table 4 Performance comparison on the BindingDB and BioSNAP datasets with random split

Setting	DeepConv -DTI	Graph -DTA	MolTrans	DrugBAN	APMF
BindingDB					
RS	0.945	0.951	0.952	0.960	<b>0.964</b>
UD	0.943	0.950	0.945	0.959	<b>0.960</b>
UT	0.627	0.670	0.661	0.692	<b>0.701</b>
BioSNAP					
RS	0.886	0.887	0.895	0.903	<b>0.909</b>
UD	0.856	0.858	0.856	0.886	<b>0.892</b>
UT	0.692	0.704	0.714	0.710	<b>0.723</b>

### 3.2.5 缺失数据下的性能比较

为了验证所提出的模型在 BindingDB 上执行大量缺失数据下的性能表现。根据 MolTrans 中的缺失数据设置,使用每个数据集缺失 5%, 10%, 20% 和 30% 来训练 APMF 和其他深度学习基线,并评估其余数据的预测性能(90%作为测试集,10%作为验证集,用于确定实验的提前停止运行)。

图 6 呈现了所获得的结果,显示 APMF 在

所有设置中具有最佳性能。特别是在更大的数据集(BindingDB)上的提升更大,这是由于大数据集提供了更多的变化和多样性,这使得模型能够学习到更丰富的特征表示。通过观察更多样本之间的差异,模型可以学习到更具判别性的特征表示,从而更好地区分不同的类别。通过提供更多的数据样本,模型可以更好地学习数据的统计规律,减少对特定样本的过度拟合,从而进一步提高模型的泛化能力。

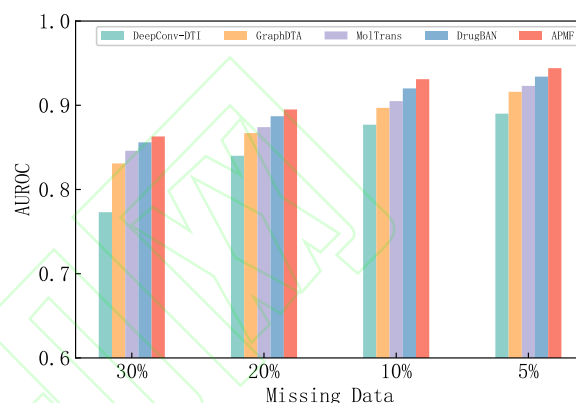


图 6 缺失数据下 BindingDB 数据集上的性能比较

Fig.6 Performance comparison on the BindingDB datasets with high fraction of missing data

## 4 结束语

本文提出了一种新的药物-靶点相互作用预测的方法。使用深度特征提取表征蛋白质序列和药物化合物,使用模态融合的方式对药物-靶点相互作用进行预测,模型可以为解释药物-靶点相互作用的性质提供生物学解释。本文提出的方法相对于基线模型在三个评测数据集上均取得了较好的性能。所提出的思想具有普遍性,可以推广到其他相互作用预测问题,如药物相互作用预测和蛋白质相互作用的预测。本研究的重点是使用一维蛋白质序列和二维分子图作为输入进行建模,由于高精度度的三维结构蛋白质已经出现,部分蛋白质三维结构预测工作取得了巨大进展,但是这项工作模型暂时没有考虑使用这种结构信息进行建模。我们相信,在复杂的三维结构上进一步扩展本文的思想,在未来的工作中也可以带来更好的性能。

## 参考文献:

- [1] 张然,王学志,汪嘉霞,等. 药物-靶点相互作用预测的计算方法综述[J]. 计算机工程与应用, 2023, 59(12):1-13.DOI:10.3778/j.issn.1002-8331.2210-0108.

- ZHANG RAN, WANG XUEZHI, WANG JIAJIA, MENG ZHEN. Survey on Computational Approaches for Drug-Target Interaction Prediction[J]. *Computer Engineering and Applications*, 2023, 59(12): 1-13.
- [2] LUO Y, ZHAO X, ZHOU J, et al. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information[J]. *Nature communications*, 2017, 8(1): 573.
- [3] ZITNIK M, NGUYEN F, WANG B, et al. Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities[J]. *Information Fusion*, 2019, 50: 71-91.
- [4] GAO K Y, FOKOUE A, LUO H, et al. Interpretable drug target prediction using deep neural representation[C]//IJCAI. 2018, 2018: 3371-3377.
- [5] LEE I, KEUM J, NAM H. DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences[J]. *PLoS computational biology*, 2019, 15(6): e1007129.
- [6] HINNERICHS T, HOEHNDORF R. DTI-Voodoo: machine learning over interaction networks and ontology-based background knowledge predicts drug-target interactions[J]. *Bioinformatics*, 2021, 37(24): 4835-4843.
- [7] NGUYEN T, LE H, QUINN T P, et al. GraphDTA: predicting drug-target binding affinity with graph neural networks[J]. *Bioinformatics*, 2021, 37(8): 1140-1147.
- [8] LI M, CAI X, LI L, et al. Heterogeneous Graph Attention Network for Drug-Target Interaction Prediction[C]// *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 2022: 1166-1176.
- [9] YU Z, GAO H. Molecular Graph Representation Learning via Heterogeneous Motif Graph Construction[J]. 2022.DOI:10.48550/arXiv.2202.00529.
- [10] HUANG K, XIAO C, GLASS L M, et al. MolTrans: molecular interaction transformer for drug-target interaction prediction[J]. *Bioinformatics*, 2021, 37(6): 830-836.
- [11] BENGIO Y, COURVILLE A, VINCENT P. Representation learning: A review and new perspectives[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(8): 1798-1828.
- [12] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[J]. *arXiv preprint arXiv:1609.02907*, 2016.
- [13] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[J]. *arXiv preprint arXiv:1710.10903*, 2017.
- [14] YIN W, SCHÜTZE H. Convolutional neural network for paraphrase identification[C]//*Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2015: 901-911.
- [15] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural network regularization[J]. *arXiv preprint arXiv:1409.2329*, 2014.
- [16] ZHENG S, LI Y, CHEN S, et al. Predicting drug-protein interaction using visual question answering system[J]. *Nature Machine Intelligence*, 2020, 2(2): 134-140.
- [17] ÖZTÜRK H, ÖZGÜR A, OZKIRIMLI E. DeepDTA: deep drug-target binding affinity prediction[J]. *Bioinformatics*, 2018, 34(17): i821-i829.
- [18] TSUBAKI M, TOMII K, SESE J. Compound-protein interaction prediction with end-to-end learning of neural networks for graphs and sequences[J]. *Bioinformatics*, 2019, 35(2): 309-318.
- [19] BAI P, MILJKOVIĆ F, JOHN B, et al. Interpretable bilinear attention network with domain adaptation improves drug-target prediction[J]. *Nature Machine Intelligence*, 2023, 5(2): 126-136.
- [20] WEININGER D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules[J]. *Journal of chemical information and computer sciences*, 1988, 28(1): 31-36.
- [21] KANGHAO S, YUNHAO Z, YUQI W, et al. DTI-HETA: prediction of drug-target interactions based on GCN and GAT on heterogeneous graph[J]. *Briefings in Bioinformatics*(3):3[2024-03-03].DOI:10.1093/bib/bbac109.
- [22] CHEN M, JU J T, ZHOU G, et al. Multifaceted protein-protein interaction prediction based on Siamese residual RCNN[J]. *Bioinformatics*, 2019, 35(14): i305-i314. DOI:10.1093/bioinformatics/btz328.
- [23] JAEGLE A, BORGEAUD S, ALAYRAC J B, et al. Perceiver IO: A General Architecture for Structured Inputs &

- Outputs[J]. 2021.DOI:10.48550/arXiv.2107.14795.
- [24] LIU T, LIN Y, WEN X, et al. BindingDB: a web-accessible database of experimentally determined protein ligand binding affinities[J]. *Nucleic acids research*, 2007, 35(suppl\_1): D198-D201.
- [25] ZITNIK M, SOSI R, MAHESHWARI S, et al. Stanford biomedical network dataset collection[J]. *Biosn. Datasets Stanford Biomed. Netw. Dataset Collect*, 2018.
- [26] LIU H, SUN J, GUAN J, et al. Improving compound–protein interaction prediction by building up highly credible negative samples[J]. *Bioinformatics*, 2015, 31(12): i221-i229.
- [27] CORTES C, VAPNIK V. Support-vector networks[J]. *Machine learning*, 1995, 20: 273-297.
- [28] Ho T K. Random decision forests[C]//*Proceedings of 3rd international conference on document analysis and recognition. IEEE*, 1995, 1: 278-282.
- [29] NICOLA G, LIU T, GILSON M K. Public domain databases for medicinal chemistry[J]. *Journal of medicinal chemistry*, 2012, 55(16): 6987-7002.