



多源描述符融合的药物-靶标相互作用预测框架

成志兴¹, 丁彦蕊^{1,2}

(1. 江南大学 江苏省媒体设计与软件技术重点实验室, 江苏 无锡 214122; 2. 江南大学 理学院, 江苏 无锡 214122)

摘要 识别药物靶标相互作用(DTI)是药物发现的一项关键任务,在虚拟筛选、药物重定位和识别药物副作用等领域都发挥了至关重要的作用。通过传统的生物实验方法识别DTI通常昂贵且耗时,随着深度学习在自然语言处理等领域取得的巨大成功,已有许多研究利用深度学习预测潜在的DTI。然而,先前的许多研究仅利用药物和蛋白的单视角特征预测DTI,忽视了多源描述符的潜在价值。该文提出了一种融合多源描述符的药物靶标相互作用预测框架(DFDTI),充分利用药物和蛋白不同视角的结构信息。首先,通过全连接(FC)层为多源描述符生成低维表示;然后,考虑到不同类型的描述符对DTI预测的贡献度不同,利用通道注意力机制给予不同的描述符权重;此外,使用单层Transformer编码器增强描述符的特征表示;最后,拼接药物和蛋白的增强特征表示,并输入到深度神经网络(DNN)中以预测DTI。实验结果表明,DFDTI能够有效融合不同类型的描述符,在3类评价指标上均优于该文中所有的基线方法。

关键词 药物-靶标相互作用;特征融合;注意力机制;深度学习

中图分类号: TP399 **DOI**: 10.16152/j.cnki.xdxbzr.2023-05-010

Prediction framework for drug-target interactions with multi-source descriptors fusion

CHENG Zhixing¹, DING Yanrui^{1,2}

(1. Jiangsu Key Laboratory of Media Design and Software Technology, Jiangnan University, Wuxi 214122, China;

2. School of Science, Jiangnan University, Wuxi 214122, China)

Abstract Identifying drug-target interactions (DTIs) is a key task in drug discovery, and it plays an important role in the fields of virtual screening, drug repurposing and identification of potential drug side effects. Traditional biological experimental methods of identifying DTIs are often expensive and time-consuming. With the great success of deep learning in natural language processing and other fields, many studies have used deep learning to predict DTIs. However, many previous studies only utilized single-view features of drugs and proteins to predict DTIs, ignoring the potential value of multi-source descriptors. A drug-target interaction prediction framework with multi-source descriptors fusion (DFDTI) is proposed in this paper to make full use of the structural information of drugs and proteins from different view. First, low dimensional representations are generated for the multi-source descriptors through fully connected (FC) layers. Then, considering the different contribution of different types of descriptors, the channel attention mechanism is used to give weights to de-

收稿日期: 2022-11-29

基金项目: 国家自然科学基金(62161050)

第一作者: 成志兴, 男, 从事机器学习和生物信息学研究, 6191611021@stu.jiangnan.edu.cn。

通信作者: 丁彦蕊, 女, 教授, 从事机器学习和生物信息学研究, yr_ding@jiangnan.edu.cn。

scriptors. In addition, one layer Transformer encoder is used to enhance the feature representation of descriptors. Finally, the enhanced features representations of drugs and proteins are concatenated and input into the deep neural network (DNN) to predict DTIs. Experimental results show that the DFDTI can effectively fuse different types of descriptors and outperform all baseline methods in this paper on all three types of evaluation metrics.

Keywords drug-target interaction; feature fusion; attention mechanism; deep learning

药物的发现和开发过程漫长而复杂^[1],尽管在过去几十年中药物研发投入显著增加,但每年批准的新药数量仍然很低。因此,探索有效的方法来提高药物开发的成功率至关重要^[2]。

在药物发现和开发过程中,药物靶标相互作用(drug-target interaction, DTI)的识别非常关键^[3]。由于通过传统的生物实验确定DTI通常成本高、耗时长^[4],为有效减少时间和成本,近年来基于计算机的方法得到了迅速发展。DTI预测计算方法大致可以分为3类:基于结构的方法^[5]、基于配体的方法^[6]和基于机器学习的方法^[7]。

基于结构的方法即分子对接模拟,该方法主要利用蛋白质的三维结构信息来预测DTI。由于该方法非常耗时且需要蛋白的三维结构,因此其性能受到了限制。基于配体的方法利用相似的配体往往具有相似的生物学特性的理论,将靶标蛋白的结合配体与候选药物进行比对和评分,从而为靶标蛋白推荐候选药物。然而,该方法非常依赖靶标蛋白的结合配体信息,当所选靶标蛋白只有少量已知结合配体时,效果不佳^[8]。

为了探索更高效的计算方法,近年来许多研究致力于利用机器学习预测DTI。基于机器学习的方法通常将预测DTI的问题视为二分类任务或回归任务,每个药物蛋白对都有一个标签来指示药物与蛋白之间是否存在相互作用关系。具体地,基于机器学习的方法大致可以分为基于相似性的方法、基于网络的方法和基于特征的方法。

基于相似性的方法需要预先计算药物或蛋白的多种相似性,然后将这些相似性值输入到各种机器学习方法中。Perlman等人计算了5种药物相似性与3种蛋白相似性,并利用逻辑回归分类器自动组合这些特征以产生最终的关联得分^[9]。Olayan等人计算了药物和蛋白的多种相似性,通过启发式过程选择了一组信息丰富且冗余较少的相似性组合,并利用相似性网络融合算法SNF融合了多个相似性值;然后,基于融合后的相似性值与已知DTI构建了异构网络以提取基于路径的特

征,并训练随机森林模型来预测DTI^[10]。然而,基于相似性的方法需要预先计算多种相似性值,因此,非常依赖药物与蛋白特征的注释丰富程度,难以应用于无法计算这些相似性值的药物或蛋白。

基于网络的方法通常会构建一个包括药物与蛋白的网络,通过构建的网络来挖掘药物与蛋白之间的潜在关联信息。Luo等人构建了一个包含药物、蛋白、疾病以及副作用的异构网络,通过重启随机游走算法和扩散分量分析相结合来整合异构网络中的各种信息并生成低维特征表示,最后,利用矩阵完成算法预测DTI^[11]。基于Luo等人构建的网络^[11],Wan等人采用邻域信息聚合操作更新节点的特征,通过重构异构网络学习拓扑特征以预测DTI^[12]。与基于相似性的方法类似,基于网络的方法非常依赖已知的关联信息,如果药物或蛋白没有已知的关联或已知关联非常少,则预测性能不佳。

相比基于相似性的方法和基于网络的方法,基于特征的方法因通常仅需要药物的结构信息以及蛋白的序列信息,因此适用范围更广。这类方法通常分别对药物和蛋白进行编码,然后将药物与蛋白的特征组合作为最终特征输入到机器学习方法中。随着深度学习在多个领域中的成功应用,许多研究致力于利用深度学习预测潜在的DTI^[13]。Öztürk等人提出了一种端到端的DTI预测模型DeepDTA^[14],该模型利用卷积神经网络(convolutional neural networks, CNN)从药物的SMILES(simplified molecular input line entry system)字符串^[15]和蛋白的氨基酸序列中提取特征,然后,通过全连接(Fully Connected, FC)层预测DTI得分。Lee等人提出了DTI预测模型DeepConv-DTI^[16],该模型计算了药物直径为4的扩展连通性指纹(extended connectivity fingerprints with a diameter of 4, ECFP4)^[17]作为药物特征,并利用多尺度一维CNN提取蛋白序列上的特征。Nguyen人提出了一个端到端的DTI预测模型GraphDTA^[18],将药

物视为分子图,利用图神经网络(graph neural networks, GNN)学习药物的表示,并利用 CNN 提取蛋白的低维表示。受自然语言处理领域中算法的启发,Wan 等人将自然语言处理领域中的潜在语义分析和 Word2vec 与深度学习相结合,提出了一个通用可扩展的 DTI 预测框架 DeepCPI^[19]。

鉴于注意力机制^[20]能够在训练过程中自动关注任务相关的信息而忽略不相关的信息,一些研究致力于将注意力机制融入到 DTI 预测模型中。受 Transformer^[21]可以挖掘两个序列之间的特征的启发,Chen 等人将药物和蛋白视为两个序列,提出了一种基于 Transformer 的模型来预测 DTI^[22]。基于 DeepDTA 的模型框架,Zhao 等人加入了特征级的注意力机制以模拟小分子药物中的原子与蛋白中的氨基酸之间复杂的相互作用,提出了预测模型 HyperAttentionDTI^[23]。

虽然上述基于特征的方法取得了不错的预测性能,然而,大多仅使用单个描述符作为特征或从原始的药物 SMILES 字符串和蛋白序列中自动提取特征,忽略了含有具体意义的多源药物结构描述符以及蛋白结构描述符的潜在价值。为了有效融合来自不同视角的多源描述符以预测 DTI,本

文提出了一个名为 DFDTI 的预测模型。首先,DFDTI 利用嵌入层计算药物和蛋白描述符的低维表示;然后,利用通道注意力机制给予不同描述符类型以不同的权重;其次,利用 Transformer 编码器融合药物和蛋白相关的多种类型的描述符的特征表示;最后,拼接药物和蛋白的特征,并通过深度神经网络(deep neural networks, DNN)预测 DTI 得分。实验结果表明,DFDTI 在 3 类评价指标上均优于所有的基线方法,证明了融合多源描述符以预测 DTI 的潜在价值以及 DFDTI 的优越性。

1 材料和方法

本文提出了一个基于描述符融合的深度学习方法 DFDTI 预测潜在的 DTI,模型的框架如图 1 所示。DFDTI 共包括 4 个模块:①生成药物和蛋白描述符的低维表示的嵌入层;②基于通道注意力的特征加权模块;③基于 Transformer 编码器的特征增强模块;④基于 DNN 的药物蛋白相互作用预测模块。

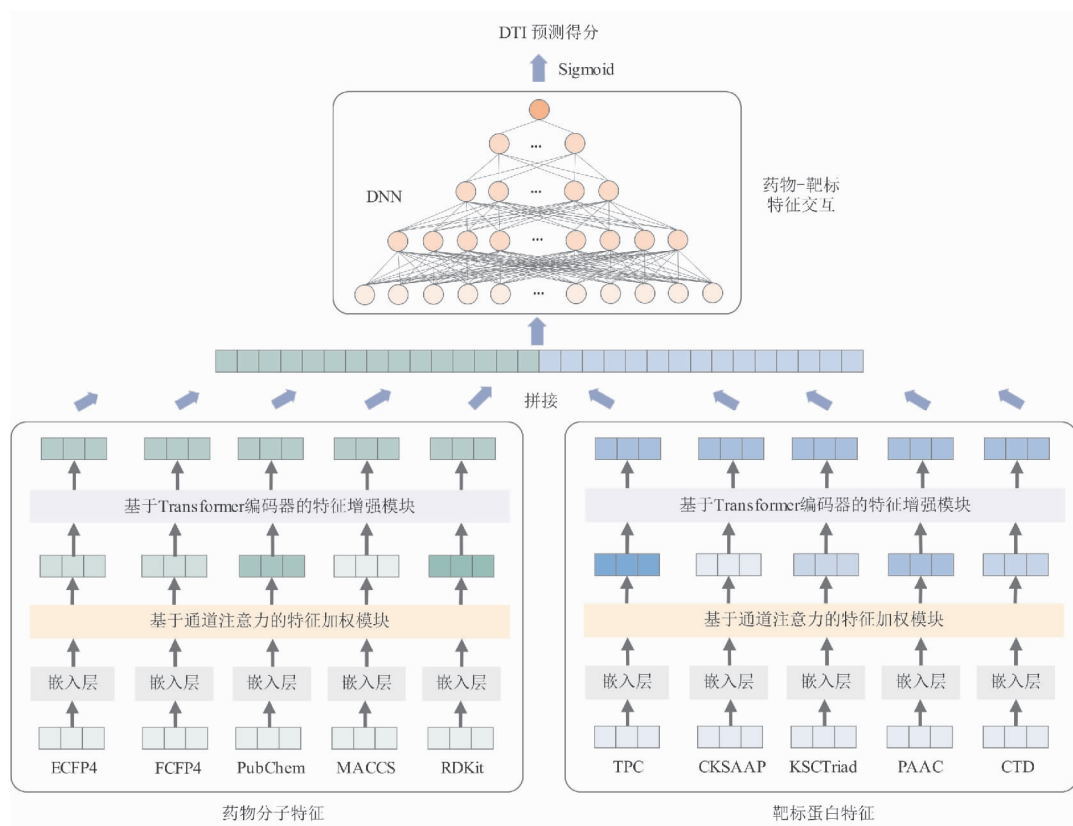


图 1 DFDTI 模型的框架
Fig. 1 The framework of the DFDTI model

1.1 数据集

研究中所用的药物蛋白相互作用数据是从 DrugBank 数据库^[24]中提取的,包括 1 409 个来自 DrugBank 数据库的 FDA(food and drug administration)已批准的小分子药物与 1 648 个来自 Uniprot 数据库^[25]的已审查智人蛋白之间的 8 020 个药物蛋白相互作用。此外,从 DrugBank 数据库中提取了药物的 SMILES 结构信息,从 Uniprot 数据库中提取了蛋白的氨基酸序列。

1.2 多源描述符计算

在本节中,计算了多种描述符作为药物与蛋白的初始特征,并将其作为 DFDTI 的输入。使用的药物描述符共包含 5 种药物分子指纹,蛋白描述符共包含 5 种蛋白序列描述符。

1.2.1 药物描述符

分子指纹是编码分子结构特征的一种方式,在虚拟筛选、基于相似性的化合物搜索、靶标分子排名以及其他药物发现过程中都发挥着重要作用^[26]。常见的分子指纹包括基于子结构的指纹、基于拓扑或路径的指纹和圆形指纹。本节计算的药物描述符包括以下 3 种。

1) 基于子结构的指纹 MACCS^[27] 和 PubChem^[28]。基于子结构通常根据给定的子结构或特征是否存在而设置,其中,MACCS 包含 166 个子结构,而 PubChem 包含 881 个子结构。

2) 基于拓扑的指纹 RDKit^[29]。基于拓扑的指纹根据从一个原子出发指定长度的所有路径产生指纹,RDKit 根据给定最小长度和最大长度之间的所有分子路径计算分子指纹,由 2 048 个比特位组成。

3) 圆形指纹 ECFP4 以及直径为 4 的功能基指纹(functional-class fingerprints with a diameter of 4,FCFP4)^[30]。圆形指纹会根据从原子出发指定半径范围的子结构产生分子指纹,其中,最常用的指纹类型为 ECFP4。与 ECFP 不同的是,FCFP 中具有相同功能或相似功能的原子没有区别。圆形指纹的长度通常有 1 024 和 2 048 两种,本节中使用的是 1 024 位的圆形指纹。

指纹 MACCS、RDKit、ECFP4 和 FCFP4 由开源化学信息 Python 包 RDKit^[29] 计算,指纹 PubChem 由 R 包 rcdk^[31] 计算。

1.2.2 蛋白描述符

基于序列的分析和预测是基本的生物信息学任务,有助于理解蛋白的结构功能^[32]。为了从蛋

白序列中提取不同的特征以生成数字向量,计算了以下 5 种蛋白序列描述符作为蛋白描述符。

1) 三肽组成(tripeptides composition, TPC)^[33]。每一位都代表了蛋白序列中每种类型三肽的出现频率,向量维数为 8 000 维。

2) k 间隔氨基酸对组成(composition of k -spaced amino acid pairs, CKSAAP)^[34]。通过计算间隔从 0 到 k 的氨基酸对在蛋白序列中的频率提取序列特征。本节中 k 取值为 3,向量维数为 1 600 维。

3) k 间隔联合三元组(k -spaced conjoint triad, KSCTriad)^[35]。通过计算间隔从 0 到 k 的三元组(3 个连续的氨基酸组成的单元)在蛋白序列中的频率提取特征。本节中 k 取值为 3,向量维数为 1 372 维。

4) 伪氨基酸组成(pseudo-amino acid composition, PAAC)^[36]。前 20 位代表了每种类型的氨基酸在蛋白序列中的频率,其他代表各种包含序列顺序信息的伪成分。本节中使用的 PAAC 向量维数为 22 维。

5) 组成、转化和分布(composition, transition, and distribution, CTD)^[37]。将氨基酸序列转化成具有某些结构或理化性质的残基序列,然后为给定的结构或理化性质分别计算组成、转化和分布 3 种描述符,最终向量维数为 273 维。

所有类型的蛋白序列描述符均由开源平台 iLearnPlus^[32] 计算。

1.3 药物、蛋白的低维特征表示

由于描述符向量初始维度高,且不同类型的描述符向量维度相差较大,因此,首先通过多个结构相同但大小不同的嵌入层为每种类型的药物描述符和蛋白质描述符生成对应的维度相同的低维表示。对于任意类型的药物描述符 $s \in \{\text{MACCS}, \text{PubChem}, \text{RDKit}, \text{ECFP4}, \text{FCFP4}\}$,通过一层 FC 层生成其低维表示。

$$\mathbf{h}_{\text{dr}}^s = \sigma(\mathbf{W}_{\text{dr}}^s \mathbf{h}_{\text{dr}}^s + \mathbf{b}_{\text{dr}}^s) \quad (1)$$

式中: $\mathbf{W}_{\text{dr}}^s$ 和 $\mathbf{b}_{\text{dr}}^s$ 表示类型为 s 的药物描述符的嵌入层的可训练权重和偏差; $\mathbf{h}_{\text{dr}}^s$ 表示类型为 s 的药物描述符向量; σ 是非线性激活函数 Relu。

类似地,对于任意类型的蛋白描述符 $t \in \{\text{TPC}, \text{CKSAAP}, \text{KSCTriad}, \text{PAAC}, \text{CTD}\}$,其低维表示生成过程为

$$\mathbf{h}_{\text{p}}^t = \sigma(\mathbf{W}_{\text{p}}^t \mathbf{h}_{\text{p}}^t + \mathbf{b}_{\text{p}}^t) \quad (2)$$

式中: $\mathbf{W}_{\text{p}}^t$ 和 $\mathbf{b}_{\text{p}}^t$ 表示类型为 t 的蛋白描述符的嵌入

层的可训练权重和偏差; $\mathbf{h}'_p$ 表示类型为 t 的蛋白描述符向量; σ 是非线性激活函数 Relu。

1.4 基于通道注意力机制的特征加权模块

由于不同类型的描述符对 DTI 预测的贡献度是不同的,因此,需要对不同类型的描述符加权以区分其贡献度,使更重要的描述符发挥更大的作用。而在输入到预测模型之前,其贡献度是未知的,因此,受 Hu 等人^[38]启发,本文通过通道注意力模块自动学习不同描述符的权重,模块框架如图 2 所示。

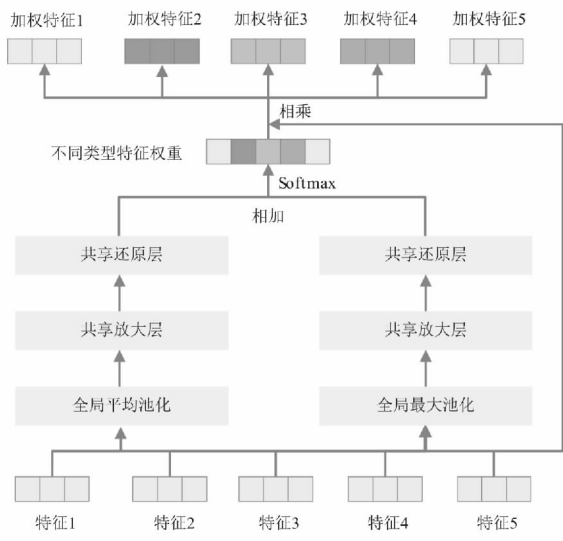


图 2 基于通道注意力的特征加权框架

Fig. 2 The framework of feature weighting based on channel attention

该模块通过将不同类型的特征视为特征中不同的通道,然后,利用通道注意力给与特征不同的权重。以药物为例,在获得药物多种描述符的低维表示后,将各个描述符的低维表示视为不同通道,将综合特征视为药物特征 $\mathbf{h}'_{dr}$ 。然后分别对其进行全局平均池化和全局最大池化,获取其不同视角下的全局特征,计算公式如下,

$$\mathbf{g}_{dr}^1 = \text{average}(\mathbf{h}'_{dr}) \quad (3)$$

$$\mathbf{g}_{dr}^2 = \max(\mathbf{h}'_{dr}) \quad (4)$$

式中:average 代表全局平均池化操作;max 代表全局最大池化操作。得到药物的 2 种全局表示后,将其作为输入得到其通道注意力权重,

$$\mathbf{a}_{dr}^1 = \mathbf{W}_2 \sigma(\mathbf{W}_1 \mathbf{g}_{dr}^1 + \mathbf{b}_1) + \mathbf{b}_2 \quad (5)$$

$$\mathbf{a}_{dr}^2 = \mathbf{W}_2 \sigma(\mathbf{W}_1 \mathbf{g}_{dr}^2 + \mathbf{b}_1) + \mathbf{b}_2 \quad (6)$$

$$\mathbf{a}_{dr} = \text{Softmax}(\mathbf{a}_{dr}^1 + \mathbf{a}_{dr}^2) \quad (7)$$

式中: $\mathbf{W}_1 \in \mathbf{R}^{C \times r}$ 和 $\mathbf{W}_2 \in \mathbf{R}^{r \times C}$ 代表可训练的权重; C 代表通道数; r 代表比例; $\mathbf{b}_1$ 与 $\mathbf{b}_2$ 代表可训练

的偏差; Softmax 代表归一化函数; σ 是激活函数 Relu。本节中 C 取值为 5, r 取值为 16。

此时,通过该通道注意力权重 $\mathbf{a}_{dr}$ 可以区分不同描述符特征的贡献度,使得对最终 DTI 预测贡献更大的特征获得更大的权重值。在得到药物不同通道的注意力得分后,将其与原特征相乘得到加权后的药物特征表示,

$$\mathbf{h}''_{dr} = \mathbf{a}_{dr} \times \mathbf{h}'_{dr} \quad (8)$$

同理,通过上述过程,可以得到加权后的蛋白特征表示 $\mathbf{h}''_p$ 。

1.5 基于 Transformer 编码器的特征增强模块

虽然不同描述符之间是相互独立的,但是同一药物的不同描述符或同一蛋白的不同描述符之间可能存在相关性。受 Transformer 能够适应多模态数据以挖掘多类型特征之间的相关性及互补性的启发,本文将药物与蛋白的不同类型的描述符看作药物与蛋白的多模态数据,利用单层 Transformer 编码器增强药物与蛋白的特征表示。

由于使用的特征之间没有序列关联,本节中的 Transformer 编码器去除了位置编码模块,由单头自注意力层、残差连接和层归一化、前馈层 3 部分组成,模块框架如图 3 所示。最终可以获得药物的增强特征和蛋白的增强特征。

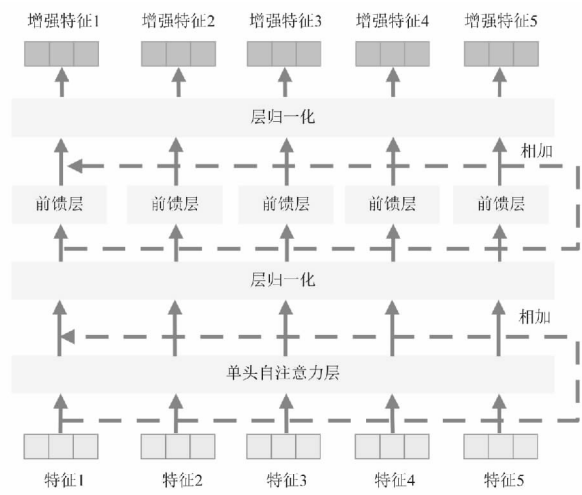


图 3 基于 Transformer 编码器的特征增强框架

Fig. 3 The framework of feature enhancement based on Transformer encoder

1) 单头自注意力层。自注意力层是 Transformer 中的关键技术,可以捕捉不同特征向量之间的相关性。具体地,对于每种类型的特征输入,自注意力层为每个输入生成 3 个不同的向量,分别命名为查询 Q 、键 K 和值 V 。给定 Q ,自注意力层会计算该 Q 与每个 K 的注意力得分,然后,将该

注意力得分乘以每个 $\mathbf{K}$ 相应的 $\mathbf{V}$ 。注意力得分的具体计算过程如下,

$$a_{\text{self}}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \tag{9}$$

式中: d_k 是取决于图层大小的比例因子,即 $\mathbf{K}$ 向量维度的平方根。自注意力机制可以减少对外部信息的依赖,专注于捕获不同类型特征的内部相关性。

2)残差连接和层归一化。残差连接用于改善信息流,避免因网络过度深化而导致梯度消失和退化的问题。残差连接表示将一层的输出添加到前一层的输出。层归一化旨在在一定程度上避免过拟合,即对样本中特定类型特征的特征向量进行归一化。

3)前馈层。前馈层由 2 层 FC 层组成,目的是将自注意力层得到的向量投影到一个空间中,以便更容易地提取所需的信息。

1.6 基于 DNN 的相互作用预测模块

在获得所有类型的增强特征后,将药物和蛋白质的所有增强特征拼接作为药物蛋白对的最终特征表示 $\mathbf{h}_{\text{dr-p}}$,然后将 $\mathbf{h}_{\text{dr-p}}$ 输入到 DNN 中,预测药物蛋白相互作用。DNN 可以表示为

$$\mathbf{y}' = \text{DNN}(\mathbf{h}_{\text{dr-p}}) \tag{10}$$

本节中构建的 DNN 共由 3 个隐藏层和 1 个输出层组成。DNN 中第 l 层隐藏层更新特征的过程可表示为

$$\mathbf{y}^{(l)} = \sigma(\text{BN}(\mathbf{W}_h^{(l)} \mathbf{y}^{(l-1)} + \mathbf{b}_h^{(l)})) \tag{11}$$

式中: $\mathbf{W}_h^{(l)}$ 和 $\mathbf{b}_h^{(l)}$ 代表第 l 层隐藏层的可训练权重和偏差; $\mathbf{y}^{(l-1)}$ 代表第 $l-1$ 层隐藏层的输出;BN 代表用于防止过拟合的批次归一化; σ 是非线性激活函数 Relu。DNN 中的输出层可以表示为

$$\mathbf{y}' = \text{Sigmoid}(\mathbf{W}_o \mathbf{y}^h + \mathbf{b}_o) \tag{12}$$

式中: $\mathbf{W}_o$ 和 $\mathbf{b}_o$ 代表输出层的权重和偏差; $\mathbf{y}^h$ 代表最后一层隐藏层的输出向量;Sigmoid 代表归一化函数,用于将输出值映射到 0 到 1 之间的交互得分。

1.7 损失函数

损失函数定义为药物蛋白对的真实标签 y 与预测得分 y' 之间的二分类交叉熵损失函数。对于包含 N 个样本的批次数据,损失函数可以表示为

$$l = -\frac{1}{N} \sum_{i=1}^N y_i \ln y'_i + (1 - y_i) \ln(1 - y'_i) \tag{13}$$

1.8 动态学习率衰减

学习率是神经网络训练时的重要参数,如果过大则会导致不收敛,过小则会导致收敛速度太慢。因此,本文使用动态学习率衰减策略优化模型的训练过程。当验证集的 AUC 连续 20 次迭代均没有升高时,说明此时模型的学习遇到了瓶颈,学习率降低为当前学习率的一半。

2 实验

2.1 实验环境

本文所提方法 DFDTI 的平台配置为:Inter® Core™ i7-9750H@2.60 GHz,内存 32 GiB,GPU 类型为 NVIDIA GeForce RTX 2060,操作系统为 Windows 10,开发环境为 Python 3.8 + CUDA 10.2 + cuDNN 7.6.5 + PyTorch 1.10.0。

2.2 超参数设置

DFDTI 的超参数设置如表 1 所示。

表 1 DFDTI 的超参数
Tab. 1 Hyperparameter of DFDTI

超参数	取值
学习率	0.000 1
权值衰减	0.000 1
批次大小	64
迭代次数	200
DNN 隐藏层神经元个数	256、128、64
DNN 输出层神经元个数	1
低维表示大小	128
Transformer 编码器前馈层神经元个数	128
Transformer 编码器 Dropout 概率	0.1
优化器	Adam

2.3 性能评估

由于 DTI 预测为分类任务,因此,使用准确率 (accuracy rate, ACC, 式中简记 R_{ACC}) 作为衡量模型的评价指标之一,分类阈值设为 0.5。

$$R_{\text{ACC}} = \frac{N_{\text{TP}} + N_{\text{TN}}}{N_{\text{P}} + N_{\text{N}}} \tag{14}$$

式中: N_{TP} 和 N_{TN} 分别代表真阳性和真阴性的数量; N_{P} 和 N_{N} 分别代表阳性样本和阴性样本的数量。

考虑到人工设定分类阈值可能难以准确衡量不同方法的预测性能,为了避免阈值选择的主观

性,本文还使用接收机工作特性(receiver operating characteristic, ROC)曲线下面积(area under curve, AUC)和精确召回率(precision recall, PR)曲线下面积(area under precision recall curve, AU-PR)作为衡量模型性能的主要指标。ROC 曲线绘制了不同阈值下的真阳性率(true positive rate, TPR, 式中简记 R_{TP})与假阳性率(false positive rate, FPR, 式中简记 R_{FP}), PR 曲线绘制了不同阈值下的精准率(precision, 式中简记 R_{pre})与召回率(recall, 式中简记 R_{rec})。TPR 和 FPR 的定义为

$$R_{TP} = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (15)$$

$$R_{FP} = \frac{N_{FP}}{N_{TN} + N_{FP}} \quad (16)$$

式中: N_{FP} 和 N_{FN} 分别代表假阳性和假阴性的数量。

精准率和召回率定义为

$$R_{pre} = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (17)$$

$$R_{rec} = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (18)$$

2.4 数据集划分

在研究中,已知的药物蛋白相互作用被视为正样本,未知关联的药物蛋白对被视为负样本。由于未知药物蛋白对的数量远远大于已知药物蛋白相互作用的数量,随机从未标记药物蛋白对中选择一定数量的药物蛋白对作为负样本。本文构建了两类数据集:均衡数据集和非均衡数据集。

1) 均衡数据集。从未标记药物蛋白对中随机选择与正样本数量相等的药物蛋白对作为负样本,最终数据集中共包含 8 020 个正样本和 8 020 个负样本。考虑到手动选择不同负样本时实验结果不同,为了在一定程度上避免单次随机选择负样本时可能带来的实验误差,并公平比较不同的 DTI 预测方法,从未标记的药物蛋白对中随机选取 3 次负样本,并将其构成的数据集分别命名为 DTI-1、DTI-2 与 DTI-3。

2) 非均衡数据集。考虑到实际情况中负样本的数量大于正样本的数量,构建了 2 个负样本多于正样本的数据集,以验证模型的预测性能。分别从未标记药物蛋白对中随机选择数量为正样本数量 3 倍和 5 倍的药物蛋白对作为负样本构建数据集(即 24 060 个负样本和 40 100 个负样本),并将这 2 个数据集分别命名为 DTI-3fold 和

DTI-5fold。

为了公平验证预测方法的性能,所有数据集训练集、验证集和测试集按 8: 1: 1 的比例随机划分 5 次,并以 5 次实验的平均结果作为最终结果。为了保证所有方法使用相同的训练集、验证集和测试集,数据集的划分在输入预测方法之前完成。

2.5 基线方法

为了证明模型性能的优越性,将 DFDTI 的性能与 6 个先进的基于深度学习的 DTI 预测方法进行了比较。

1) DeepDTA^[14]。分别利用 CNN 从药物的 SMILES 字符串和蛋白序列中提取特征,然后将药物与蛋白的特征拼接输入到 DNN 中。

2) DeepConv-DTI^[16]。首先,计算药物的 ECFP4 指纹作为药物的初始特征。然后,通过 FC 层生成药物的低维表示。此外,利用 CNN 从蛋白序列中提取特征。最后,拼接药物与蛋白的特征输入到 DNN 中。

3) GraphDTA^[18]。首先,将药物的 SMILES 字符串转化为以原子为节点的分子图,并通过 GNN 生成药物的低维表示。然后,利用 CNN 提取蛋白特征,并拼接药物与蛋白的特征输入到 DNN 中。

4) DeepCPI^[19]。利用直径为 2 的 ECFP 指纹生成药物的子结构,并通过潜在语义分析生成药物的特征。对于蛋白,将每个蛋白的氨基酸序列视为“句子”,将每 3 个不重叠的氨基酸视为“单词”,然后,利用 Word2vec 生成蛋白质的特征。随后,分别通过两层 FC 层提取药物和蛋白的低维表示,并将其拼接输入到 DNN 中。

5) TransformerCPI^[22]。通过图卷积神经网络(graph convolutional network, GCN)学习药物的特征,利用 Word2vec 将蛋白序列转换为蛋白特征。此外,该方法修改了 Transformer 编码器,使其更适合处理基于序列的 DTI 预测任务。随后,将药物特征作为解码器的输入,将蛋白特征作为编码器的输入,最后,将解码器输出的相互作用向量输入到 FC 层中,得到预测得分。

6) HyperAttentionDTI^[23]。与 DeepDTA 类似,不同的是该方法利用特征矩阵上的注意力机制为每个原子和氨基酸分配一个注意力向量。

各基线方法的框架如表 2 所示。

表 2 基线的框架

Tab. 2 The framework of baselines

方法	药物输入	蛋白输入	药物特征提取	蛋白特征提取	DTI 预测
DeepDTA	SMILES	氨基酸序列	CNN	CNN	DNN
DeepConv-DTI	ECFP4	氨基酸序列	FC	CNN	DNN
GraphDTA	SMILES	氨基酸序列	GNN	CNN	DNN
DeepCPI	低维表示 (ECFP2 +	低维表示	FC	FC	DNN
	潜在语义分析)	(Word2vec)			
TransformerCPI	SMILES	低维表示	GCN + Transformer	修改的 Transformer	FC
		(Word2vec)	解码器	编码器	
HyperAttentionDTI	SMILES	氨基酸序列	CNN + 注意力机制	CNN + 注意力机制	DNN

3 结果与分析

3.1 均衡数据集上与基线方法的比较

为了证明模型 DFDTI 的优越性,在 3 个数据集上将 DFDTI 与基线方法进行了比较,实验结果如表 3 所示。可以看出,相比于基线方法,DFDTI 在 3 个数据集上的 3 类指标中均取得了最优性能,而 HyperAttentionDTI 均取得了次优性能。相比于 HyperAttentionDTI,3 个数据集上的 ACC 分别高出了 0.016 6、0.013 9 和 0.022 7,AUC 分别高出了 0.011 7、0.007 2 和 0.011 5,而 AUPR 分别高出了 0.014 5、0.008 0 和 0.012 9。

从均衡数据集上的 AUC 来看,所有方法的预测结果从高到低分别是 DFDTI、HyperAttentionDTI、DeepDTA、DeepConv-DTI、DeepCPI、Transform-

erCPI 和 GraphDTA。其中,所有基线方法均是从某个视角出发计算药物与蛋白的特征,然后,通过单视角特征预测 DTI,而 DFDTI 融合了多个视角下的描述符特征,因此,DFDTI 可以提取更全面的结构信息。此外,排名靠后的 GraphDTA 和 TransformerCPI 均利用了 GNN 从分子图中提取药物的结构信息,这样的方法虽然可以从复杂的分子图中提取更详细的信息,但是由于数据集大小有限,无法完全学习到复杂的分子图中的拓扑特征,甚至不如与其特征提取过程类似的 ECFP 指纹。

从表 3 中还可以看出,除 GraphDTA 的结果有一定程度上的变化以外,其他基线方法在 3 个数据集上的结果差异均不明显,证明了取 5 次结果平均值作为最终结果衡量方法的性能具有一定的稳定性,有利于公平比较不同的方法性能。

表 3 均衡数据集上 DFDTI 和基线的平均性能

Tab. 3 Average performance of DFDTI and baselines on balanced datasets

方法	ACC			AUC			AUPR		
	DTI-1	DTI-2	DTI-3	DTI-1	DTI-2	DTI-3	DTI-1	DTI-2	DTI-3
DeepDTA	0.872 3	0.869 8	0.867 8	0.939 4	0.936 8	0.938 9	0.935 9	0.931 6	0.937 2
DeepConv-DTI	0.854 0	0.849 1	0.860 4	0.926 7	0.924 1	0.929 6	0.921 7	0.913 7	0.927 9
GraphDTA	0.846 0	0.807 5	0.808 9	0.906 1	0.865 1	0.872 2	0.905 2	0.862 7	0.881 4
DeepCPI	0.841 5	0.836 0	0.844 3	0.917 1	0.914 6	0.919 7	0.913 5	0.909 0	0.919 5
TransformerCPI	0.816 5	0.813 3	0.824 0	0.893 5	0.888 4	0.897 6	0.889 9	0.882 2	0.893 1
HyperAttentionDTI	0.874 6	0.873 9	0.871 3	0.941 3	0.941 4	0.941 9	0.939 3	0.940 1	0.940 8
DFDTI	0.891 2	0.887 8	0.894 0	0.953 0	0.948 6	0.953 4	0.953 8	0.948 1	0.953 7

3.2 非均衡数据集上与基线方法的比较

为了进一步证明模型 DFDTI 的优越性,在 2 个非均衡数据集上与基线方法进行了比较,实验结果如表 4 所示。可以看出,相比于基线方法,DFDTI 在 2 个数据集上的 3 类指标中均取得了最优性能。相比于次优方法,2 个数据集上的 ACC 分别高出了 0.011 8 和 0.006 8,AUC 分别高出了

0.001 1 和 0.000 2,而 AUPR 分别高出了 0.012 1 和 0.009 9。

结合表 3 和表 4 的结果可以看出,随着数据集中负样本数量的增加,大多数方法的 ACC 和 AUC 都有一定程度的增加,而 AUPR 有一定程度的降低。这样的结果与常识是符合的,即在非均衡数据集中 AUPR 能提供更有价值的性能评估。

从非均衡数据集上的 AUPR 来看,所有方法的预测结果从高到低分别是 DFDTI、HyperAttentionDTI、DeepDTA、DeepConv-DTI、DeepCPI、TransformerCPI 和 GraphDTA。

表 4 非均衡数据集上 DFDTI 和基线的平均性能

Tab. 4 Average performance of DFDTI and baselines on imbalanced datasets

方法	ACC		AUC		AUPR	
	DTI-3fold	DTI-5fold	DTI-3fold	DTI-5fold	DTI-3fold	DTI-5fold
DeepDTA	0.914 8	0.935 3	0.954 1	0.952 9	0.901 4	0.864 2
DeepConv-DTI	0.911 8	0.935 5	0.948 1	0.949 7	0.893 9	0.863 3
GraphDTA	0.885 0	0.911 4	0.893 2	0.897 1	0.815 8	0.765 5
DeepCPI	0.888 7	0.892 4	0.950 9	0.953 8	0.891 3	0.851 5
TransformerCPI	0.875 6	0.904 4	0.919 7	0.920 1	0.823 7	0.769 6
HyperAttentionDTI	0.915 8	0.935 2	0.955 8	0.957 0	0.908 0	0.875 1
DFDTI	0.927 6	0.942 3	0.956 9	0.957 2	0.920 1	0.885 0

3.3 与最优基线方法的比较

从表 3 和表 4 的结果可以看出,性能最优的基线方法为 HyperAttentionDTI。本节从以下几个方面对 HyperAttentionDTI 与 DFDTI 进行详细的对比。

1) 输入。HyperAttentionDTI 直接以药物的 SMILES 字符串和蛋白的氨基酸序列作为输入,而 DFDTI 以药物的 5 种分子指纹和 5 种蛋白序列描述符作为输入。HyperAttentionDTI 通过 one-hot 编码对 SMILES 字符串的 64 种字符和氨基酸序列的 20 种氨基酸进行编码,并通过设定最大长度将变长数据转换成定长数据(若小于最大长度则用 0 补齐,超过最大长度则截断,药物字符串和蛋白字符串的最大长度设为 100 个字符)。显然,这样的处理方法不适用于处理长度特别长或特别短的药物字符串和蛋白序列。DFDTI 的输入为定长的描述符,描述符中每一位都代表了某种特定的含义。

2) 嵌入层。HyperAttentionDTI 利用嵌入层将稀疏的 one-hot 编码向量转化为密集向量,输出为二维向量。DFDTI 利用嵌入层将不同长度的描述符向量转化为统一大小的密集特征,每个嵌入层的输出为一维向量。

3) 特征提取算法。HyperAttentionDTI 利用 3 层 CNN 提取药物与蛋白的结构特征,并利用基于特征的注意力机制为每一个原子和每一个氨基酸训练了一个注意力向量,以突出部分原子或氨基酸。DFDTI 利用通道注意力机制对不同类型的描述符特征进行加权,然后,利用 Transformer 编码器增强药物与蛋白的特征表示。

4) 预测算法。HyperAttentionDTI 和 DFDTI 均在拼接药物与蛋白的特征后使用 DNN 进行 DTI

预测,不同点在于 HyperAttentionDTI 使用 Dropout 增强泛化性能,DFDTI 通过批次归一化增强泛化性能。

5) 复杂度。相比于 HyperAttentionDTI,DFDTI 嵌入层的输出向量维度更低,模型的参数量更少,训练时间也更短。

3.4 单一描述符性能

为了证明特征融合的有效性,即特征融合后的结果优于任意单一描述符作为特征的结果,本节在 DTI-1 数据集上,将所有类型的单一药物描述符与单一蛋白描述符两两交叉输入到 DFDTI 中的 DNN 模块中预测 DTI。不同组合的描述符作为特征的 AUC 结果如图 4 所示。

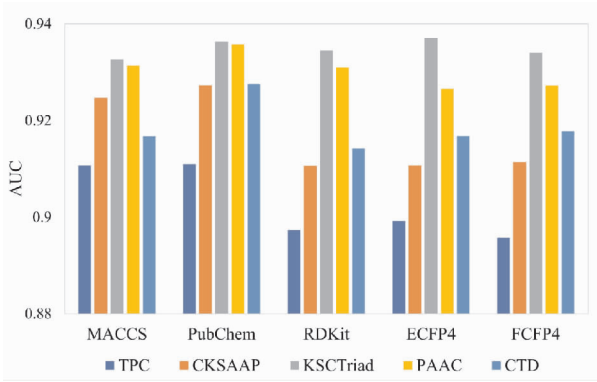


图 4 单一描述符特征的 AUC 结果

Fig. 4 The AUC results of single descriptor feature

从图 4 可以看出,单一描述符特征组合 AUC 最高时(AUC = 0.937 1),药物特征为圆形指纹 ECFP4,蛋白特征为包含了三元组信息的 KSC-Triad。DFDTI 模型在 DTI-1 数据集上的 AUC 为 0.953 0,比最优单一特征组合的情况高出了 0.015 9,充分证明了 DFDTI 融合多源描述符特征的方法可以提高预测 DTI 的性能。

此外,当药物特征为 ECFP4 时,不同蛋白描述符作为特征时的 AUC 结果从高到低为 KSCTriad、PAAC、CTD、CKSAAP 和 TPC。当蛋白特征为 KSCTriad 时,不同药物分子指纹作为特征时的 AUC 结果从高到低为 ECFP4、PubChem、RDKit、FCFP4 和 MACCS。

3.5 DFDTI 中不同组件的影响

为了证明 DFDTI 模型中不同组成部分的有效性,即所提特征融合方法的有效性,在多种设置下进行了实验与 DFDTI 比较,包括:①DFDTI 中去除所有的注意力模块,仅保留描述符嵌入模块以及基于 DNN 的 DTI 预测模块(Only-DNN);②DFDTI 中去除 Transformer 编码器模块,即注意力模块中仅保留基于通道注意力的特征加权模块(Only-Channel);③DFDTI 中去除通道注意力模块,注意力模块中仅保留基于 Transformer 编码器的特征增强模块(Only-Transformer)。

不同设置下的实验在 DTI-1 数据集上的 AUC 结果如图 5 所示。可以看出,去除了所有注意力模块后的 AUC 为 0.939 8,相比于 DFDTI 降低了 0.013 2,证明了提出的注意力模块的有效性。从 3.4 节的预测结果可知,单个描述符组合的最优 AUC 为 0.937 1,而将多个描述符特征的低维表示拼接输入到 DNN 中的 AUC 结果为 0.939 8,仅比单个描述符组合的最优结果高 0.002 7。这样的结果证明了直接拼接多个描述符的低维表示以融合特征的能力是非常有限的。

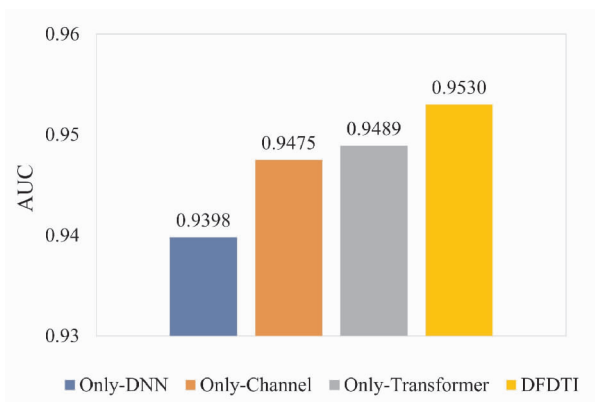


图5 DFDTI 中不同组件的影响

Fig. 5 The effect of different components in DFDTI

从图 5 可以看出,Only-Channel 的 AUC 比 Only-DNN 高出了 0.007 7,证明了基于通道注意力的特征加权模块的有效性。相比于 Only-DNN,Only-Channel 添加的通道注意力自动学习了不同描述符的权重,使得更有价值的描述符在进行预

测时有更高的重要性,从而提高了最终的预测性能。如图 5 所示,Only-Transformer 的 AUC 比 Only-DNN 高出了 0.009 1,证明了 Transformer 编码器模块利用不同描述符之间的相关性提取特征的有效性。相比于 Only-DNN,Only-Transformer 添加的 Transformer 编码器更有效地融合了多个自相关的描述符特征。

此外,相比于 Only-Channel 和 Only-Transformer,DFDTI 的 AUC 分别高出了 0.005 5 和 0.004 1,证明了将基于通道注意力的特征加权模块与基于 Transformer 编码器的特征增强模块进行组合预测 DTI 的有效性。综上所述,这 2 个模块不仅可以独立地发挥作用,组合在一起更可以增强 DTI 预测的性能。

4 结语

本文提出了一种基于深度学习的药物-靶标相互作用预测模型 DFDTI。考虑到单一的描述符仅能代表某个角度下的特征,设计了一种基于特征融合的 DTI 预测框架,以融合来源于不同视角的药物描述符和蛋白描述符。为了证明模型的优越性,将 DFDTI 与 6 种最新的基线方法进行了比较。实验结果表明,与基线方法相比,DFDTI 在 ACC、AUC 和 AUPR 这 3 种评价指标中均获得了最优结果。此外,进行了一系列实验证明了 DFDTI 融合描述符特征模块的有效性。

后续将从 2 个方面继续进行研究以提升模型的预测性能。①加入更多类型的描述符并采用特征选择方法选择最优的描述符组合;②药物特征与蛋白特征之间采用更合理、更具可解释性的交互方法取代拼接操作。

参考文献

- [1] VAMATHEVAN J, CLARK D, CZODROWSKI P, et al. Applications of machine learning in drug discovery and development[J]. Nature Reviews Drug Discovery, 2019, 18(6): 463-477.
- [2] LUO H M, LI M, YANG M Y, et al. Biomedical data and computational models for drug repositioning: A comprehensive review[J]. Briefings in Bioinformatics, 2021, 22(2): 1604-1619.
- [3] ZHAO Q C, YANG M Y, CHENG Z J, et al. Biomedical data and deep learning computational models for predicting compound-protein relations [J]. IEEE/

- ACM Transactions on Computational Biology and Bioinformatics, 2022, 19(4): 2092-2110.
- [4] ZHENG S J, LI Y J, CHEN S, et al. Predicting drug-protein interaction using quasi-visual question answering system[J]. Nature Machine Intelligence, 2020, 2(2): 134-140.
- [5] FORLI S, HUEY R, PIQUE M E, et al. Computational proteinligand docking and virtual drug screening with the AutoDock suite[J]. Nature Protocols, 2016, 11(5): 905-919.
- [6] KEISER M J, ROTH B L, ARMBRUSTER B N, et al. Relating protein pharmacology by ligand chemistry[J]. Nature Biotechnology, 2007, 25(2): 197-206.
- [7] 刘皓森, 杨志伟, 王力卓, 等. 基于机器学习的药物-靶标相互作用预测[J]. 中国生物工程杂志, 2022, 42(4): 40-48.
- LIU H M, YANG Z Y, WANG L Z, et al. Research progress of drug target interaction prediction based on machine learning[J]. China Biotechnology, 2022, 42(4): 40-48.
- [8] 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.
- [9] PERMAN L, GOTTLIEB A, ATIAS N, et al. Combining drug and gene similarity measures for drug-target elucidation[J]. Journal of Computational Biology, 2011, 18(2): 133-145.
- [10] OLAYAN R S, ASHOOR H, BAJIC V B. DDR: Efficient computational method to predict drugtarget interactions using graph mining and machine learning approaches[J]. Bioinformatics, 2018, 34(7): 1164-1173.
- [11] LUO Y N, ZHAO X B, ZHOU J T, et al. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information[J]. Nature Communications, 2017, 8: 573.
- [12] WAN F P, HONG L X, XIAO A, et al. NeoDTI: Neural integration of neighbor information from a heterogeneous network for discovering new drugtarget interactions[J]. Bioinformatics, 2019, 35(1): 104-111.
- [13] 戴青青, 余俊霖, 李国波. 深度学习辅助药物发现的研究进展[J]. 药学进展, 2022, 46(1): 60-70.
- DAI Q Q, YU J L, LI G B. Recent advances in deep learning aided drug discovery[J]. Progress in Pharmaceutical Sciences, 2022, 46(1): 60-70.
- [14] ÖZTÜRK H, ÖZGÜR A, OZKIRIMLI E. DeepDTA: Deep drug-target binding affinity prediction[J]. Bioinformatics, 2018, 34(17): i821-i829.
- [15] 王波, 高映新, 袁煦. SMILES 编码技术在化学物质数据库中的应用[J]. 中国石油和化工, 2010(9): 41-42.
- [16] 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.
- [17] ROGERS D, HAHN M. Extended-connectivity fingerprints[J]. Journal of Chemical Information and Modeling, 2010, 50(5): 742-54.
- [18] NGUYEN T, LE H, QUINN T P, et al. GraphDTA: Predicting drugtarget binding affinity with graph neural networks[J]. Bioinformatics, 2021, 37(8): 1140-7.
- [19] WAN F, ZHU Y, HU H, et al. DeepCPI: A deep learning-based framework for large-scale in silico drug screening[J]. Genomics, Proteomics & Bioinformatics, 2019, 17(5): 478-95.
- [20] 任欢, 王旭光. 注意力机制综述[J]. 计算机应用, 2021, 41(S1): 1-6.
- REN H, WANG X G. Review of attention mechanism[J]. Journal of Computer Applications, 2021, 41(S1): 1-6.
- [21] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all You need[EB/OL]. (2017-06-12)[2022-08-20]. <https://arxiv.org/abs/1706.03762>.
- [22] CHEN L F, TAN X Q, WANG D Y, et al. TransformerCPI: Improving compound—protein interaction prediction by sequence-based deep learning with self-attention mechanism and label reversal experiments[J]. Bioinformatics, 2020, 36(16): 4406-4414.
- [23] ZHAO Q C, ZHAO H C, ZHENG K, et al. HyperAttentionDTI: Improving drugprotein interaction prediction by sequence-based deep learning with attention mechanism[J]. Bioinformatics, 2022, 38(3): 655-662.
- [24] 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-D1082.
- [25] CONSORTIUM T U, BATEMAN A, MARTIN M J, et al. UniProt: The universal protein knowledgebase in 2021[J]. Nucleic Acids Research, 2021, 49(D1): D480-D489.
- [26] GAO K F, NGUYEN D D, SRESHT V, et al. Are 2D fingerprints still valuable for drug discovery? [J]. Physical Chemistry Chemical Physics: PCCP, 2020,

- 22(16): 8373-8390.
- [27] DURANT J L, LELAND B A, HENRY D R, et al. Reoptimization of MDL keys for use in drug discovery [J]. Journal of Chemical Information and Computer Sciences, 2002, 42(6): 1273-1280.
- [28] KIM S, CHEN J, CHENG T J, et al. PubChem in 2021: New data content and improved web interfaces [J]. Nucleic Acids Research, 2020, 49 (D1): D1388-D1395.
- [29] LANDRUM G. RDKit: Open-source cheminformatics [EB/OL]. (2021-09-13) [2022-09-01]. <https://www.rdkit.org>.
- [30] ROGERS D, BROWN R D, HAHN M. Using extended-connectivity fingerprints with laplacian-modified Bayesian analysis in high-throughput screening follow-up[J]. SLAS Discovery, 2005, 10(7): 682-686.
- [31] GUHA R. Chemical informatics functionality in R[J]. Journal of Statistical Software, 2007, 18(5):1-16.
- [32] CHEN Z, ZHAO P, LI C, et al. iLearnPlus: A comprehensive and automated machine-learning platform for nucleic acid and protein sequence analysis, prediction and visualization[J]. Nucleic Acids Research, 2021, 49(10): e60.
- [33] LUAI B, JIANG J B, ZENG Q G, et al. Predicting apoptosis protein subcellular location with PseAAC by incorporating tripeptide composition [J]. Protein & Peptide Letters, 2011, 18(11): 1086-1092.
- [34] CHEN K, KURGAN L A, RUAN J S. Prediction of flexible/rigid regions from protein sequences using k-spaced amino acid pairs[J]. BMC Structural Biology, 2007, 7:25.
- [35] SHEN J W, ZHANG J, LUO X M, et al. Predicting protein-protein interactions based only on sequences information[J]. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104 (11): 4337-4341.
- [36] CHOU K C. Prediction of protein cellular attributes using pseudo-amino acid composition [J]. Proteins: Structure, Function, and Bioinformatics, 2001, 43 (3): 246-255.
- [37] DUBCHAK I, MUCHNIK I, HOLBROOK S R, et al. Prediction of protein folding class using global description of amino acid sequence[J]. Proceedings of the National Academy of Sciences of the United States of America, 1995, 92(19): 8700-8704.
- [38] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City:IEEE, 2018:7132-7141.

(编 辑 李 静)