

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2024 年 12 月 12 日 (12.12.2024)



(10) 国际公布号
WO 2024/250183 A1

(51) 国际专利分类号:
G16C 20/30 (2019.01)

(21) 国际申请号: PCT/CN2023/098718

(22) 国际申请日: 2023 年 6 月 6 日 (06.06.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 之江实验室(ZHEJIANG LAB) [CN/CN]; 中国浙江省杭州市余杭区中泰街道科创大道之江实验室, Zhejiang 311121 (CN)。

(72) 发明人: 郑鹏飞(ZHENG, Pengfei); 中国浙江省杭州市余杭区中泰街道科创大道之江实验室, Zhejiang 311121 (CN)。 劳传祺(LAO, Chuanqi); 中国浙江省杭州市余杭区中泰街道科创大道之江实验室, Zhejiang 311121 (CN)。 陈红阳(CHEN,

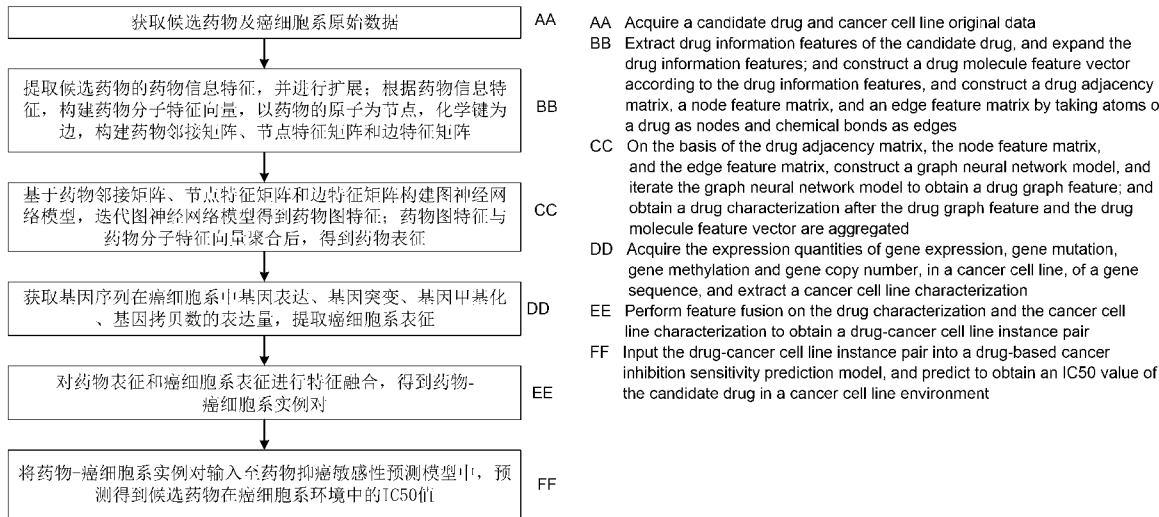
Hongyang); 中国浙江省杭州市余杭区中泰街道科创大道之江实验室, Zhejiang 311121 (CN)。

(74) 代理人: 杭州求是专利事务所有限公司(HANGZHOU QIUSHI PATENT OFFICE CO.,LTD.); 中国浙江省杭州市西湖区古墩路701号绿城紫金国际广场C座1506室, Zhejiang 310030 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD,

(54) Title: DRUG-BASED CANCER INHIBITION SENSITIVITY PREDICTION METHOD AND DEVICE BASED ON GRAPH NEURAL NETWORK

(54) 发明名称: 一种基于图神经网络的药物抑癌敏感性预测方法与装置



(57) Abstract: Disclosed in the present invention are a drug-based cancer inhibition sensitivity prediction method and device based on a graph neural network. The method comprises: acquiring a candidate drug and cancer cell line original data; extracting drug information features of the candidate drug, and expanding the drug information features; constructing a drug molecule feature vector according to the drug information features, and constructing a drug adjacency matrix, a node feature matrix, and an edge feature matrix by taking atoms of a drug as nodes and chemical bonds as edges; constructing and iterating a graph neural network model to obtain a drug graph feature; obtaining a drug characterization after the drug graph feature and the drug molecule feature vector are aggregated; acquiring and storing the expression quantities of gene expression, gene mutation, gene methylation and gene copy number, in a cancer cell line, of a gene sequence, and extracting a cancer cell line characterization; and performing feature fusion on the drug characterization and

SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚
(AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE,
BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR,
HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO,
PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN,
TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

the cancer cell line characterization to obtain a drug-cancer cell line instance pair, inputting the instance pair into a drug-based cancer inhibition sensitivity prediction model, and predicting to obtain an IC50 value of the candidate drug in a cancer cell line environment.

(57) 摘要: 本发明公开了一种基于图神经网络的药物抑癌敏感性预测方法与装置, 包括: 获取候选药物及癌细胞系原始数据; 提取候选药物的药物信息特征, 并进行扩展; 根据药物信息特征, 构建药物分子特征向量, 以药物的原子为节点, 化学键为边, 构建药物邻接矩阵、节点特征矩阵和边特征矩阵; 构建并迭代图神经网络模型, 得到药物图特征; 药物图特征与药物分子特征向量聚合后, 得到药物表征; 获取存储基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量, 提取癌细胞系表征; 对药物表征和癌细胞系表征进行特征融合, 得到药物-癌细胞系实例对, 将实例对输入至药物抑癌敏感性预测模型中, 预测得到候选药物在癌细胞系环境中的IC50值。

一种基于图神经网络的药物抑癌敏感性预测方法与装置

技术领域

本发明主要涉及人工智能药物研发领域，具体涉及一种基于图神经网络的药物抑癌敏感性预测方法与装置。

背景技术

得益于深度学习算法的快速发展及大量癌细胞系及其全基因组信息的出现，以数据为驱动的癌症药物敏感性预测是加速癌症药物研发的必然技术趋势，能有效地缓解传统癌症药物研发周期长、成本高、效率低等问题，能为医生的临床用药提供关键指导，有助于癌症的个性化治疗。基于深度学习来进行药物抑癌敏感性预测是当下的热点，该类方法将药物和癌细胞系数据输入到深度学习模型当中，组成“药物-癌细胞系”实例对，并通过深度学习算法来进行特征提取，挖掘实例对内在联系，从而进行药物抑癌敏感性预测。然而，当下的深度学习方法进行药物抑癌敏感性存在以下的局限性：

(1) 药物数据获取麻烦、药物特征信息不全、药物特征稀疏，使得药物特征提取不够丰富。

(2) 癌细胞系表征困难，仅由少量癌细胞系组学数据表示会导致癌细胞系特征提取不够充分。

因此，亟需提出一种药物抑癌敏感性预测方法，来获取更丰富的药物和癌细胞系信息，提取更准确的“药物—癌细胞系”表征。

发明内容

针对现有技术的不足，本发明提出了一种基于图神经网络的药物抑癌敏感性预测方法与装置。

根据本发明实施例的第一方面，提供了一种基于图神经网络的药物抑癌敏感性预测方法，所述方法包括：

获取候选药物及癌细胞系原始数据；

提取候选药物的药物信息特征，并进行扩展；根据药物信息特征，构建药物分子特征向量，以药物的原子为节点，化学键为边，构建药物邻接矩阵、节点特征矩阵和边特征矩阵；

基于药物邻接矩阵、节点特征矩阵和边特征矩阵构建图神经网络模型，迭代图神经网络模型得到药物图特征；药物图特征与药物分子特征向量聚合后，得到药物表征；

获取基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量，

提取癌细胞系表征；

对药物表征和癌细胞系表征进行特征融合，得到药物-癌细胞系实例对；

将药物-癌细胞系实例对输入至药物抑癌敏感性预测模型中，预测得到候选药物在癌细胞系环境中的 IC50 值。

根据本发明实施例的第二方面，提供了一种基于图神经网络的药物抑癌敏感性预测装置，所述装置包括：多源数据整合模块、存储模块和处理模块；

多源数据整合模块，用于获取候选药物及癌细胞系原始数据；提取候选药物的药物信息特征，并进行扩展；根据药物信息特征，构建药物分子特征向量，以药物的原子为节点，化学键为边，构建药物邻接矩阵、节点特征矩阵和边特征矩阵；

存储模块包括：药物信息数据库和癌细胞系信息数据库；

药物信息数据库，用于存储药物分子特征向量、药物邻接矩阵、节点特征矩阵和边特征矩阵；

癌细胞系信息数据库，用于存储基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量；

处理模块包括：药物表征提取单元、癌细胞系表征提取单元、药物-癌细胞系实例对获取单元、药物抑癌敏感性预测单元；

药物表征提取单元，基于药物邻接矩阵、节点特征矩阵和边特征矩阵构建图神经网络模型，迭代图神经网络模型得到药物图特征；药物图特征与药物分子特征向量聚合后，得到药物表征；

癌细胞系表征提取单元，通过获取基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量，提取癌细胞系表征；

药物-癌细胞系实例对获取单元，用于对药物表征和癌细胞系表征进行特征融合，得到药物-癌细胞系实例对；

药物抑癌敏感性预测单元，通过将药物-癌细胞系实例对输入至药物抑癌敏感性预测模型中，预测得到候选药物在癌细胞系环境中的 IC50 值。

根据本发明实施例的第三方面，提供了一种电子设备，包括一个或多个处理器，用于实现上述的基于图神经网络的药物抑癌敏感性预测方法。

根据本发明实施例的第四方面，提供了一种计算机可读存储介质，其上存储有程序，该程序被处理器执行时，用于实现上述的基于图神经网络的药物抑癌敏感性预测方法。

本发明的有益效果是：本发明提供了一种基于图神经网络的药物抑癌敏感性预测方法与装置，针对药物数据获取麻烦、药物特征信息不全、药物特征稀疏，提取候选药物的药物信

息特征，并进行扩展，融合了各大药物信息数据源，打破了单一数据源药物信息特征少、部分药物信息特征缺失等限制瓶颈，充分捕获药物特征信息，增加药物信息特征维度；针对癌细胞系表征困难，通过基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量来精准的进行表示癌细胞系，并通过 Transformer 网络和前馈神经网络对其进行编码，提取得到癌细胞系表征。本发明对药物表征和癌细胞系表征进行特征融合，得到药物-癌细胞系实例对，通过更准确、更丰富的药物-癌细胞系实例对，显著提高了药物抑癌敏感性的预测准确率。

附图说明

为了更清楚地说明本发明实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 是本发明实施例提供的基于图神经网络的药物抑癌敏感性预测方法的流程图；

图 2 是本发明实施例提供的获取药物表征的流程性框图；

图3是本发明实施例提供的获取癌细胞系表征的流程性框图；

图4是本发明实施例提供的基于图神经网络的药物抑癌敏感性预测装置的示意图；

图 5 是本发明实施例提供的一种电子设备的示意图。

具体实施方式

下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

需要说明的是，在不冲突的情况下，下述的实施例及实施方式中的特征可以相互组合。

为了解决药物数据获取麻烦、药物特征信息不全、药物特征稀疏和癌细胞系表征困难的问题，实施例提供了一种基于图神经网络的药物抑癌敏感性预测方法与装置，融合了各大药物信息数据源，打破了单一数据源药物信息特征少、部分药物信息特征缺失等限制瓶颈，充分捕获药物特征信息，增加药物信息特征维度；通过基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量来精准的进行表示癌细胞系，并通过 Transformer 和前馈神经网络对其进行编码。本发明通过更准确、更丰富的“药物—癌细胞系”表征的提取，显著提高了药物抑癌敏感性的预测结果。

图 1 示出了本发明实施例提供的基于图神经网络的药物抑癌敏感性预测方法的流程图，所述方法具体包括以下步骤：

步骤 S1，获取候选药物及癌细胞系原始数据。

步骤 S2，提取候选药物的药物信息特征，并进行扩展；根据药物信息特征，构建药物分子特征向量，以药物的原子为节点，化学键为边，构建药物邻接矩阵、节点特征矩阵和边特征矩阵。

具体地，所述步骤 S2 具体包括以下子步骤：

步骤 S201，以分子的 SMILES 序列为输入标准，从各药物信息数据源(DeepChem、RDKit、PubChem 等)中获取药物信息特征，通过阅读各药物信息数据源官方文档，明确药物信息特征中每一维度的具体含义。

其中，可提取的药物信息特征包括药物的原子特征(原子类型、形式电荷、杂化轨道等)、药物的化学键特征(键类型、同环性、共轭性等)、药物分子特征(分子量、价电子数等)。

以化合物环丁烷为例，其 SMILES 序列为“C1CCC1”，通过“MolGraphConvFeaturizer”可以得到第一个“C”原子的原子特征为“[1,0]”，特征维度为 30 维，其中 1-10 维表示原子类型，11 维表示形式电荷数，12-14 维表示杂化轨道，15-16 维表示氢键供体或受体，17 维表示原子是否属于芳香环，18-24 维表示原子的度，25-30 维表示与原子相连的氢的数量。通过“MolGraphConvFeaturizer”可以得到第一个“C”原子和第二个“C”原子之间的连边的边特征为“[1,0,0,0,1,0,1,0,0,0,0]”，特征维度为 11 维，其中 1-4 维表示连边的键类型，5 维表示边所连的两个原子是否在同一个环里，6 维表示连边是否是共轭的，7-11 维表示连边的键的立体构型。通过“rdkit”可以得到“C1CCC1”的分子特征如分子量为 56.0626，忽略氢的分子平均分子量为 48.0440，油水分配系数 LogP 为 1.5604 等。

步骤 S202，撰写可提取的药物信息特征标准文档，对步骤 S201 中得到的药物信息特征进行去重、补全、融合。

步骤 S203，根据步骤 S202 中得到的药物信息，以药物的原子为节点，化学键为边，构建药物邻接矩阵A、节点特征矩阵V、边特征矩阵E和药物分子特征向量Molecule_Feature。

由此，药物邻接矩阵A由维度为(shape, shape)的矩阵形式给出，其中“shape”代表该药物的节点数量(即药物的原子数量)。

节点特征矩阵V由维度为(shape, dim1)的矩阵形式给出，其中“dim1”代表该药物每个节点的特征数量(即药物每个原子的特征数量)。

边特征矩阵E由维度为(shape, shape, dim2)的邻接矩阵形式给出，其中“dim2”代表该药物每条边的特征数量(即药物化学键的特征数量)。进一步地，当两节点间不存在边时，边特征矩阵中相应的坐标元素用“0”填充。

药物分子特征向量Molecule_Feature由维度为(1, dim3)的特征向量形式给出，其中“的

特征 3”表示该药物的分子特征维度。

对于药物邻接矩阵A，由于其是固定的，因此可以直接通过基准数据源得到。

对于节点特征矩阵V的构建与边特征矩阵E，由于两者的构建方法相同，以节点特征矩阵的构建为例进行阐述。

为表达方便，本实例只取两个数据源，并将数据源 A 和数据源 B 可以获得的节点特征矩阵 V_A ， V_B 分别表示成以下形式：

$$V_A = \begin{pmatrix} & F_{A1} & \cdots & F_{An} \\ \text{Atom}_{A1} & [1,1,0,1] & [\dots] & [\dots] \\ \text{Atom}_{A2} & [1,0,0,1] & [\dots] & [\dots] \\ \text{Atom}_{A3} & [1,1,0,0] & [\dots] & [\dots] \\ \vdots & [\dots] & [\dots] & [\dots] \\ \text{other} & [0,0,0,0] & [\dots] & [\dots] \end{pmatrix}, \quad V_B = \begin{pmatrix} & F_{B1} & \cdots & F_{Bm} \\ \text{Atom}_{B1} & [1,0,0,1] & [\dots] & [\dots] \\ \text{Atom}_{B2} & [1,1,0,1] & [\dots] & [\dots] \\ \text{Atom}_{B3} & [1,1,0,0] & [\dots] & [\dots] \\ \vdots & [\dots] & [\dots] & [\dots] \\ \text{other} & [0,0,0,0] & [\dots] & [\dots] \end{pmatrix}$$

若特征 F_{B1} 并未出现在数据源 A 中，本实例将特征 F_{B1} 直接添加到 V_A 末尾，即可表示成以下形式：

$$V = \begin{pmatrix} & F_{A1} & \cdots & F_{An} & F_{B1} \\ \text{Atom}_1 & [1,1,0,1] & [\dots] & [\dots] & [1,0,0,1] \\ \text{Atom}_2 & [1,0,0,1] & [\dots] & [\dots] & [1,1,0,1] \\ \text{Atom}_3 & [1,1,0,0] & [\dots] & [\dots] & [1,1,0,0] \\ \vdots & [\dots] & [\dots] & [\dots] & [\dots] \\ \text{other} & [0,0,0,0] & [\dots] & [\dots] & [0,0,0,0] \end{pmatrix}$$

若特征 F_{B1} 与数据源 A 中的 F_{A1} 是同个性质的表征，当特征 F_{A1} 与 F_{B1} 是离散化数据时，本实例中取 F_{A1} 与 F_{B1} 中表示更多的那个特征作为该性质的表征。如原子的“Hybridization”属性，MolGraphConvFeaturizer 提取器只将其表示成“sp”，“sp2”，“sp3”，而 DMPNNFeaturizer 提取器将其表示成“sp”，“sp2”，“sp3”，“sp3d”，“sp3d2”，因此对于“Hybridization”属性，本实例采用 DMPNNFeaturizer 提取器中提取到的属性作为最终属性；当特征 F_{A1} 与 F_{B1} 是非离散化数据时，本实例将 F_{A1} 与 F_{B1} 直接相连作为该性质的表征，即可表示成以下形式：

$$V = \begin{pmatrix} & F_{AB} & \cdots \\ \text{Atom}_1 & [1,1,0,1,1,0,0,1] & [\dots] \\ \text{Atom}_2 & [1,0,0,1,1,1,0,1] & [\dots] \\ \text{Atom}_3 & [1,1,0,0,1,1,0,0] & [\dots] \\ \vdots & [\dots] & [\dots] \\ \text{other} & [0,0,0,0,0,0,0,0] & [\dots] \end{pmatrix}$$

对于药物分子特征向量Molecule_Feature，可由不同数据源得到不同的分子特征，并通过“one-hot”等编码方式，将这些分子特征编码成药物分子特征向量Molecule_Feature。

步骤 S204，将药物邻接矩阵A，节点特征矩阵V，边特征矩阵E以及药物分子特征向量Molecule_Feature传入到药物信息数据库中进行存储。

如图 2 所示, 步骤 S3, 基于药物邻接矩阵、节点特征矩阵和边特征矩阵构建图神经网络模型, 迭代图神经网络模型得到药物图特征; 药物图特征与药物分子特征向量聚合后, 得到药物表征。

需要说明的是, 在本实例中可以从从药物信息数据库中获取物信息特征及相应的药物邻接矩阵A, 节点特征矩阵V, 边特征矩阵E等。搭建并迭代图神经网络模型, 以节点更新、边更新的方式进行消息传递和聚合, 得到药物图特征 (Graph_Feature)。与药物信息数据库中的药物分子特征向量 (Molecule_Feature) 聚合后, 得到最终的药物表征 (Whole_Feature)。

具体地, 所述步骤 S3 具体包括以下子步骤:

步骤 S301, 以搭建并迭代图神经网络模型, 以节点更新、边更新的方式进行消息传递和聚合, 得到药物图特征 (Graph_Feature)。

其中, 以节点更新的方式更新图神经网络模型, 在图神经网络模型第 l 层进行节点特征更新, 表达式如下:

$$V^{(l+1)} = \text{Update}(V^l, \sigma(E_i^{(l)} V^l W_i^{(l+1)} + b_i^{(l+1)}) | i \in (1, \dots, \text{dim2}))$$

式中, 边特征矩阵为维度为 (shape, shape, dim2) 的矩阵, 其中“dim2”代表药物每条边的特征数量, 即药物化学键的特征数量; 在节点更新过程中, 将边特征矩阵拆分成 “dim2” 个 (shape, shape, 1) 的边特征矩阵; $V^{(l+1)}$ 表示第 $l+1$ 层的节点特征矩阵, Update 表示更新函数, $V^{(l)}$ 表示第 l 层的节点特征子矩阵, σ 为激活函数, $E_{(i)}^{(l)}$ 表示第 l 层上的边特征矩阵经过拆分后的第 i 个边特征子矩阵, $W_i^{(l+1)}$ 和 $b_i^{(l+1)}$ 表示图神经网络模型参数。

以边更新的方式更新图神经网络模型, 在图神经网络模型第 l 层进行边特征更新, 表达式如下:

$$e_{ij}^{(l+1)} = \sigma((V_i^{(l+1)} || V_j^{(l+1)}) W_{ij}^{(l+1)} + b_{ij}^{(l+1)})$$

$$E_{ij}^{(l+1)} = \text{Update}(E_{ij}^l, e_{ij}^{(l+1)})$$

式中, $e_{ij}^{(l+1)}$ 表示第 $l+1$ 层的节点 i 与节点 j 之间的连边 e_{ij} 的更新变量, σ 为激活函数, $V_i^{(l+1)}$ 表示节点 i 在第 $l+1$ 层的特征向量, $V_j^{(l+1)}$ 表示节点 j 在第 $l+1$ 层的特征向量, $W_{ij}^{(l+1)}$ 和 $b_{ij}^{(l+1)}$ 表示图神经网络模型参数, $E_{ij}^{(l+1)}$ 表示第 $l+1$ 层的节点 i 与节点 j 之间的边特征矩阵, Update 表示更新函数, $E_{ij}^{(l)}$ 表示第 l 层的节点 i 与节点 j 之间的边特征矩阵。

步骤 S303, 计算药物图特征 (Graph_Feature), 表达式如下:

$$\text{Graph_Feature} = \frac{1}{|V|} \sum_{i \in V} V_i^{(L)}$$

其中, $|V|$ 表示节点数量, $V_i^{(L)}$ 表示节点 i 在第 L 层的特征向量, L 为图神经网络模型的层数。

步骤 S304, 药物图特征与药物分子特征向量进行聚合, 得到药物表征, 表达式如下:

$$\text{Whole_Feature} = \text{Concat}(\text{Graph_Feature}, \text{Molecule_Feature})$$

式中, Whole_Feature 表示药物表征, Graph_Feature 表示药物图特征, Molecule_Feature 表示药物分子特征。

如图 3 所示, 步骤 S4, 获取基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量, 提取癌细胞系表征。

具体地, 所述步骤 S4 具体包括以下子步骤:

步骤 S401, 获取基因序列在癌细胞系中的各表达量的相关信息, 对于癌细胞系 C_1 , 其表达序列 X_{c1} 可以表达成以下形式:

$$X_{c1} = \{X_{c1,g1}, X_{c1,g2}, \dots, X_{c1,gn}\} \in \mathbb{R}^{n \times 4}$$

$$X_{c1,gi} = [\text{express}_{c1,gi}, \text{mutat}_{c1,gi}, \text{methy}_{c1,gi}, \text{copy}_{c1,gi}] \in \mathbb{R}^4, i \in (1, \dots, n)$$

其中, n 表示为癌细胞系中致病基因数量, $X_{c1,gi}$ 表示第 i 个致病基因在癌细胞系中基因表达($\text{express}_{c1,gi}$)、基因突变($\text{mutat}_{c1,gi}$)、基因甲基化($\text{methy}_{c1,gi}$)、基因复制水平 $\text{copy}_{c1,gi}$ 的表达量。

步骤 S402, 将癌细胞系 C_1 的表达序列 X_{c1} 输入至 Transformer 网络中, 输出表征向量 Z_{c1} , 表达式如下:

$$Z_{c1}^i = \text{Softmax}\left(\frac{(X_{c1}W_q^i)(X_{c1}W_k^i)^T}{\sqrt{d}}\right)(X_{c1}W_v^i) \in \mathbb{R}^{n \times d}$$

$$Z_{c1} = \text{Concat}(Z_{c1}^1, Z_{c1}^2, \dots, Z_{c1}^K)$$

特别地, 本实例在 Transformer 网络中采用多头注意力机制。其中, Z_{c1}^i 表示癌细胞系 C_1 的表达序列 X_{c1} 在经过第 i 个注意力头后得到的表征, Softmax 为激活函数, W_q^i , W_k^i , W_v^i 分别为 Transformer 网络的Q, K, V矩阵, $d = \frac{4}{K}$, Z_{c1} 为 X_{c1} 经过 Transformer 网络后的输出表征。

步骤 S403, 将表征向量 Z_{c1} 输入至前馈神经网络中, 得到癌细胞系表征 Cell_Feature , 表达式如下:

$$\text{Cell_Feature} = \sigma(Z_{c1}W + b)$$

式中, σ 为激活函数, W 和 b 表示前馈神经网络模型参数

步骤 S5, 对药物表征和癌细胞系表征进行特征融合, 得到药物-癌细胞系实例对, 表达式如下:

$$\text{Mol_Cell_Feature} = \text{Whole_Feature} || \text{Cell_Feature}$$

式中, Mol_Cell_Feature 表示药物-癌细胞系实例对, Whole_Feature 表示药物表征, Cell_Feature 表示癌细胞系表征。

步骤 S6, 将药物-癌细胞系实例对输入至药物抑癌敏感性预测模型中, 预测得到候选药物在癌细胞系环境中的 IC50 值。

进一步地, 药物抑癌敏感性预测模型为一分类器; 将药物-癌细胞系实例对输入至分类器中, 进行药物抑癌敏感性预测, 预测得到药物在癌细胞系环境中的 IC50 值, 由此量化药物对癌细胞系的抑制能力大小, 表达式如下:

$$\text{Pred} = \text{Softmax}(\text{Mol_Cell_Feature} * W + b)$$

式中, Mol_Cell_Feature 表示药物-癌细胞系实例对, W 和 b 为药物抑癌敏感性预测模型参数。

综上所述, 本发明提供了一种基于图神经网络的药物抑癌敏感性预测方法, 针对药物数据获取麻烦、药物特征信息不全、药物特征稀疏, 提取候选药物的药物信息特征, 并进行扩展, 融合了各大药物信息数据源, 打破了单一数据源药物信息特征少、部分药物信息特征缺失等限制瓶颈, 充分捕获药物特征信息, 增加药物信息特征维度; 针对癌细胞系表征困难, 通过基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量来精准的表示癌细胞系, 并通过 Transformer 网络和前馈神经网络对其进行编码, 提取得到癌细胞系表征。本发明对药物表征和癌细胞系表征进行特征融合, 得到药物-癌细胞系实例对, 通过更准确、更丰富的药物-癌细胞系实例对, 显著提高了药物抑癌敏感性的预测准确率。

本发明实施例还提供了一种基于图神经网络的药物抑癌敏感性预测装置, 如图 4 所示, 所述装置包括: 多源数据整合模块、存储模块和处理模块;

多源数据整合模块, 用于获取候选药物及癌细胞系原始数据; 提取候选药物的药物信息特征, 并进行扩展; 根据药物信息特征, 构建药物分子特征向量, 以药物的原子为节点, 化学键为边, 构建药物邻接矩阵、节点特征矩阵和边特征矩阵;

存储模块包括: 药物信息数据库和癌细胞系信息数据库;

药物信息数据库, 用于存储药物分子特征向量、药物邻接矩阵、节点特征矩阵和边特征矩阵;

癌细胞系信息数据库, 用于存储基因序列在癌细胞系中基因表达、基因突变、基因甲基

化、基因拷贝数的表达量；

处理模块包括：药物表征提取单元、癌细胞系表征提取单元、药物-癌细胞系实例对获取单元、药物抑癌敏感性预测单元；

药物表征提取单元，基于药物邻接矩阵、节点特征矩阵和边特征矩阵构建图神经网络模型，迭代图神经网络模型得到药物图特征；药物图特征与药物分子特征向量聚合后，得到药物表征；

癌细胞系表征提取单元，通过获取基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量，提取癌细胞系表征；

药物-癌细胞系实例对获取单元，用于对药物表征和癌细胞系表征进行特征融合，得到药物-癌细胞系实例对；

药物抑癌敏感性预测单元，通过将药物-癌细胞系实例对输入至药物抑癌敏感性预测模型中，预测得到候选药物在癌细胞系环境中的 IC50 值。

需要说明的是，所述基于图神经网络的药物抑癌敏感性预测装置在处理模块的输入端还设置有输入控制接口，用于控制输入数据；处理模块的输出端还设置有输出控制接口，用于控制输出数据；输出控制接口与外部接口相连，其输出表现形式除 IC50 值外，还可进行药物特征可视化输出、癌细胞系特征可视化输出等。

关于上述实施例中的装置，其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

对于装置实施例而言，由于其基本对应于方法实施例，所以相关之处参见方法实施例的部分说明即可。以上所描述的系统实施例仅仅是示意性的，其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本申请方案的目的。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

相应的，本申请还提供一种电子设备，包括：一个或多个处理器；存储器，用于存储一个或多个程序；当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如上述的基于图神经网络的药物抑癌敏感性预测方法。如图 5 所示，为本发明实施例提供的基于图神经网络的药物抑癌敏感性预测方法所在任意具备数据处理能力的设备的一种硬件结构图，除了图 5 所示的处理器、内存以及网络接口之外，实施例中装置所在的任意具备数据处理能力的设备通常根据该任意具备数据处理能力的设备的实际功能，还可以包括其他硬件，对此不再赘述。

相应的，本申请还提供一种计算机可读存储介质，其上存储有计算机指令，该指令被处理器执行时实现如上述的基于图神经网络的药物抑癌敏感性预测方法。所述计算机可读存储介质可以是前述任一实施例所述的任意具备数据处理能力的设备的内部存储单元，例如硬盘或内存。所述计算机可读存储介质也可以是外部存储设备，例如所述设备上配备的插接式硬盘、智能存储卡（Smart Media Card，SMC）、SD卡、闪存卡（Flash Card）等。进一步的，所述计算机可读存储介还可以既包括任意具备数据处理能力的设备的内部存储单元也包括外部存储设备。所述计算机可读存储介质用于存储所述计算机程序以及所述任意具备数据处理能力的设备所需的其他程序和数据，还可以用于暂时地存储已经输出或者将要输出的数据。

以上实施例仅用于说明本发明的设计思想和特点，其目的在于使本领域内的技术人员能够了解本发明的内容并据以实施，本发明的保护范围不限于上述实施例。所以，凡依据本发明所揭示的原理、设计思路所作的等同变化或修饰，均在本发明的保护范围之内。

权 利 要 求 书

1、一种基于图神经网络的药物抑癌敏感性预测方法，其特征在于，所述方法包括：

获取候选药物及癌细胞系原始数据；

提取候选药物的药物信息特征，并进行扩展；根据药物信息特征，构建药物分子特征向量，以药物的原子为节点，化学键为边，构建药物邻接矩阵、节点特征矩阵和边特征矩阵；

基于药物邻接矩阵、节点特征矩阵和边特征矩阵构建图神经网络模型，迭代图神经网络模型得到药物图特征；药物图特征与药物分子特征向量聚合后，得到药物表征；

获取基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量，提取癌细胞系表征；

对药物表征和癌细胞系表征进行特征融合，得到药物-癌细胞系实例对；

将药物-癌细胞系实例对输入至药物抑癌敏感性预测模型中，预测得到候选药物在癌细胞系环境中的 IC50 值。

2、根据权利要求 1 所述的基于图神经网络的药物抑癌敏感性预测方法，其特征在于，药物信息特征包括药物的原子特征、药物的化学键特征、药物分子特征；

药物邻接矩阵为维度为 (shape, shape) 的矩阵，其中“shape”代表药物的节点数量，即药物的原子数量；

节点特征矩阵为维度为 (shape, dim1) 的矩阵，其中“dim1”代表药物每个节点的特征数量，即药物每个原子的特征数量；

边特征矩阵为维度为 (shape, shape, dim2) 的矩阵，其中“dim2”代表药物每条边的特征数量，即药物化学键的特征数量；

药物分子特征向量为维度为 (1, dim3) 的特征向量，其中“dim3”代表药物的分子特征维度。

3、根据权利要求 1 或 2 所述的基于图神经网络的药物抑癌敏感性预测方法，其特征在于，节点特征矩阵与边特征矩阵的构建方法相同；

其中，节点特征矩阵的构建过程包括：

获取数据源 A 和数据源 B，数据源 A 和数据源 B 对应的第一节点特征矩阵 V_A 、第二节点特征矩阵 V_B 的表达式如下：

$$V_A = \begin{pmatrix} & F_{A1} & \cdots & F_{An} \\ \text{Atom}_{A1} & [1,1,0,1] & [\dots] & [\dots] \\ \text{Atom}_{A2} & [1,0,0,1] & [\dots] & [\dots] \\ \text{Atom}_{A3} & [1,1,0,0] & [\dots] & [\dots] \\ \vdots & [\dots] & [\dots] & [\dots] \\ \text{other} & [0,0,0,0] & [\dots] & [\dots] \end{pmatrix}, \quad V_B = \begin{pmatrix} & F_{B1} & \cdots & F_{Bm} \\ \text{Atom}_{B1} & [1,0,0,1] & [\dots] & [\dots] \\ \text{Atom}_{B2} & [1,1,0,1] & [\dots] & [\dots] \\ \text{Atom}_{B3} & [1,1,0,0] & [\dots] & [\dots] \\ \vdots & [\dots] & [\dots] & [\dots] \\ \text{other} & [0,0,0,0] & [\dots] & [\dots] \end{pmatrix}$$

当特征 F_{B1} 未出现在数据源 A 中，将特征 F_{B1} 直接添加到第一节点特征矩阵 V_A 末尾，得到节点特征矩阵 V ，表达式如下：

$$V = \begin{pmatrix} & F_{A1} & \cdots & F_{An} & F_{B1} \\ \text{Atom}_1 & [1,1,0,1] & [\dots] & [\dots] & [1,0,0,1] \\ \text{Atom}_2 & [1,0,0,1] & [\dots] & [\dots] & [1,1,0,1] \\ \text{Atom}_3 & [1,1,0,0] & [\dots] & [\dots] & [1,1,0,0] \\ \vdots & [\dots] & [\dots] & [\dots] & [\dots] \\ \text{other} & [0,0,0,0] & [\dots] & [\dots] & [0,0,0,0] \end{pmatrix}$$

当特征 F_{B1} 与数据源 A 中的 F_{A1} 是同一个性质的表征，且特征 F_{A1} 与 F_{B1} 是离散化数据时，取 F_{A1} 与 F_{B1} 中表示更多的特征作为该性质的表征；

当特征 F_{A1} 与 F_{B1} 是非离散化数据时，将 F_{A1} 与 F_{B1} 直接相连作为该性质的表征，节点特征矩阵 V 的表达式为：

$$V = \begin{pmatrix} & F_{AB} & \cdots \\ \text{Atom}_1 & [1,1,0,1,1,0,0,1] & [\dots] \\ \text{Atom}_2 & [1,0,0,1,1,1,0,1] & [\dots] \\ \text{Atom}_3 & [1,1,0,0,1,1,0,0] & [\dots] \\ \vdots & [\dots] & [\dots] \\ \text{other} & [0,0,0,0,0,0,0,0] & [\dots] \end{pmatrix}$$

4、根据权利要求 1 所述的基于图神经网络的药物抑癌敏感性预测方法，其特征在于，迭代图神经网络模型得到药物图特征包括：

以节点更新的方式更新图神经网络模型，其中在图神经网络模型第 l 层进行节点特征更新，表达式如下：

$$V^{(l+1)} = \text{Update}\left(V^l, \sigma(E_i^{(l)} V^l W_i^{(l+1)} + b_i^{(l+1)}) | i \in (1, \dots, \text{dim2})\right)$$

式中，边特征矩阵为维度为 (shape, shape, dim2) 的矩阵，其中“dim2”代表药物每条边的特征数量，即药物化学键的特征数量；在节点更新过程中，将边特征矩阵拆分成“dim2”个 (shape, shape, 1) 的边特征矩阵； $V^{(l+1)}$ 表示第 $l+1$ 层的节点特征矩阵， Update 表示更新函数， $V^{(l)}$ 表示第 l 层的节点特征子矩阵， σ 为激活函数， $E_{(i)}^{(l)}$ 表示第 l 层上的边特征矩阵经过拆分后的第 i 个边特征子矩阵， $W_i^{(l+1)}$ 和 $b_i^{(l+1)}$ 表示图神经网络模型参数。

5、根据权利要求 1 所述的基于图神经网络的药物抑癌敏感性预测方法，其特征在于，迭

代图神经网络模型得到药物图特征包括：

以边更新的方式更新图神经网络模型，其中在图神经网络模型第 l 层进行边特征更新，表达式如下：

$$e_{ij}^{(l+1)} = \sigma\left((V_i^{(l+1)} || V_j^{(l+1)})W_{ij}^{(l+1)} + b_{ij}^{(l+1)}\right)$$

$$E_{ij}^{(l+1)} = \text{Update}\left(E_{ij}^l, e_{ij}^{(l+1)}\right)$$

式中， $e_{ij}^{(l+1)}$ 表示第 $l+1$ 层的节点 i 与节点 j 之间的连边 e_{ij} 的更新变量， σ 为激活函数， $V_i^{(l+1)}$ 表示节点 i 在第 $l+1$ 层的特征向量， $V_j^{(l+1)}$ 表示节点 j 在第 $l+1$ 层的特征向量， $W_{ij}^{(l+1)}$ 和 $b_{ij}^{(l+1)}$ 表示图神经网络模型参数， $E_{ij}^{(l+1)}$ 表示第 $l+1$ 层的节点 i 与节点 j 之间的边特征矩阵， Update 表示更新函数， E_{ij}^l 表示第 l 层的节点 i 与节点 j 之间的边特征矩阵。

6、根据权利要求1所述的基于图神经网络的药物抑癌敏感性预测方法，其特征在于，获取存储基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量，提取癌细胞系表征包括：

癌细胞系 C_1 的表达序列 X_{c1} 的表达式如下：

$$X_{c1} = \{X_{c1,g1}, X_{c1,g2}, \dots, X_{c1,gn}\} \in \mathbb{R}^{n \times 4}$$

$$X_{c1,gi} = [\text{express}_{c1,gi}, \text{mutat}_{c1,gi}, \text{methy}_{c1,gi}, \text{copy}_{c1,gi}] \in \mathbb{R}^4, i \in (1, \dots, n)$$

式中， n 表示癌细胞系中致病基因数量， $X_{c1,gi}$ 表示第 i 个致病基因在癌细胞系中的表达量， $\text{express}_{c1,gi}$ 表示第 i 个致病基因在癌细胞系中的基因表达的表达量， $\text{mutat}_{c1,gi}$ 表示第 i 个致病基因在癌细胞系中的基因突变的表达量， $\text{methy}_{c1,gi}$ 表示第 i 个致病基因在癌细胞系中的基因甲基化的表达量， $\text{copy}_{c1,gi}$ 表示第 i 个致病基因在癌细胞系中的基因拷贝数的表达量；

将癌细胞系 C_1 的表达序列 X_{c1} 输入至Transformer网络中，输出表征向量 Z_{c1} ；

将表征向量 Z_{c1} 输入至前馈神经网络中，得到癌细胞系表征。

7、根据权利要求1所述的基于图神经网络的药物抑癌敏感性预测方法，其特征在于，将药物-癌细胞系实例对输入至药物抑癌敏感性预测模型中，预测得到候选药物在癌细胞系环境中的IC50值包括：

药物抑癌敏感性预测模型为一分类器；将药物-癌细胞系实例对输入至分类器中，进行药物抑癌敏感性预测，预测得到药物在癌细胞系环境中的IC50值，表达式如下：

$$\text{Pred} = \text{Softmax}(\text{Mol_Cell_Feature} * W + b)$$

式中， Mol_Cell_Feature 表示药物-癌细胞系实例对， W 和 b 为药物抑癌敏感性预测模型参

数。

8、一种基于图神经网络的药物抑癌敏感性预测装置，其特征在于，所述装置包括：多源数据整合模块、存储模块和处理模块；

多源数据整合模块，用于获取候选药物及癌细胞系原始数据；提取候选药物的药物信息特征，并进行扩展；根据药物信息特征，构建药物分子特征向量，以药物的原子为节点，化学键为边，构建药物邻接矩阵、节点特征矩阵和边特征矩阵；

存储模块包括：药物信息数据库和癌细胞系信息数据库；

药物信息数据库，用于存储药物分子特征向量、药物邻接矩阵、节点特征矩阵和边特征矩阵；

癌细胞系信息数据库，用于存储基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量；

处理模块包括：药物表征提取单元、癌细胞系表征提取单元、药物-癌细胞系实例对获取单元、药物抑癌敏感性预测单元；

药物表征提取单元，基于药物邻接矩阵、节点特征矩阵和边特征矩阵构建图神经网络模型，迭代图神经网络模型得到药物图特征；药物图特征与药物分子特征向量聚合后，得到药物表征；

癌细胞系表征提取单元，通过获取基因序列在癌细胞系中基因表达、基因突变、基因甲基化、基因拷贝数的表达量，提取癌细胞系表征；

药物-癌细胞系实例对获取单元，用于对药物表征和癌细胞系表征进行特征融合，得到药物-癌细胞系实例对；

药物抑癌敏感性预测单元，通过将药物-癌细胞系实例对输入至药物抑癌敏感性预测模型中，预测得到候选药物在癌细胞系环境中的 IC50 值。

9、一种电子设备，其特征在于，包括一个或多个处理器，用于实现权利要求 1-7 中任一项所述的基于图神经网络的药物抑癌敏感性预测方法。

10、一种计算机可读存储介质，其上存储有程序，其特征在于，该程序被处理器执行时，用于实现权利要求 1-7 中任一项所述的基于图神经网络的药物抑癌敏感性预测方法。

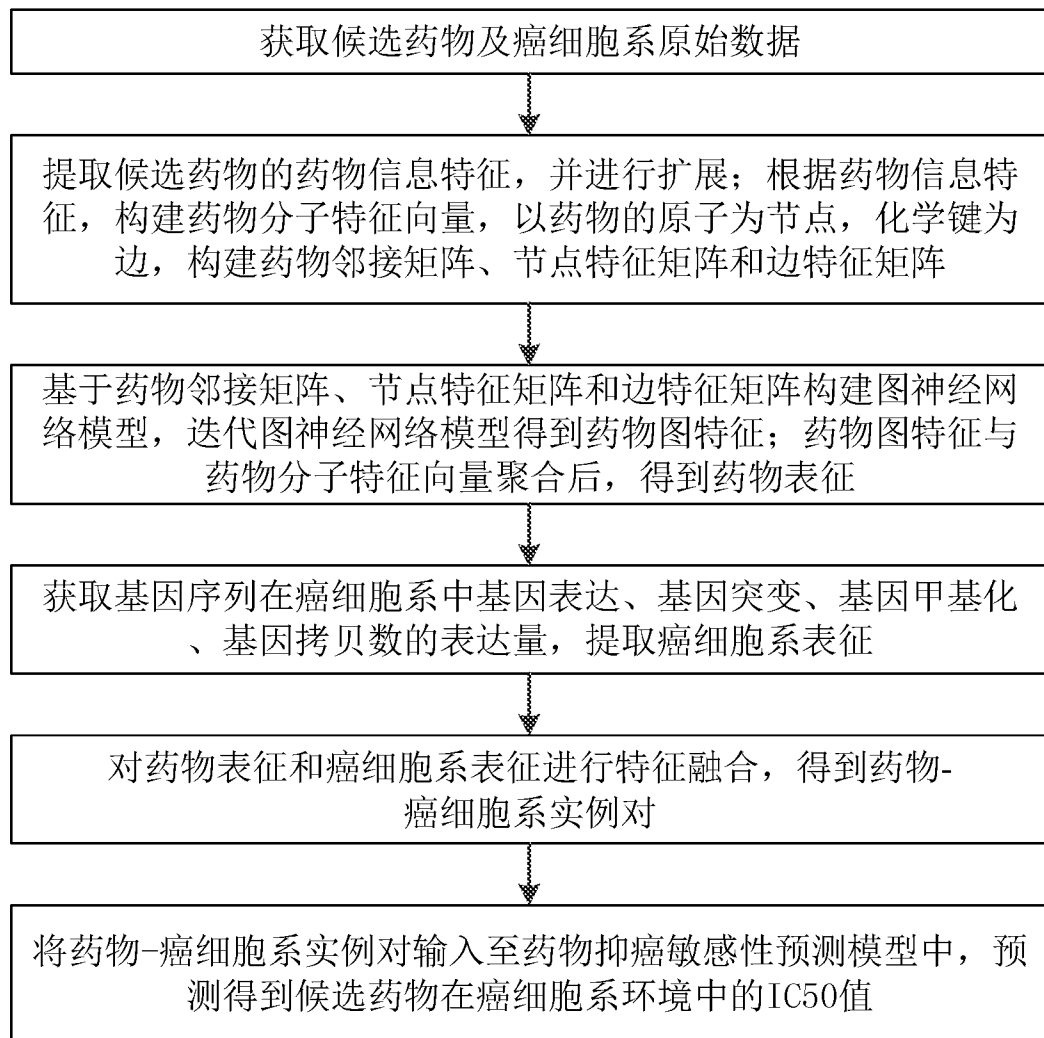


图 1

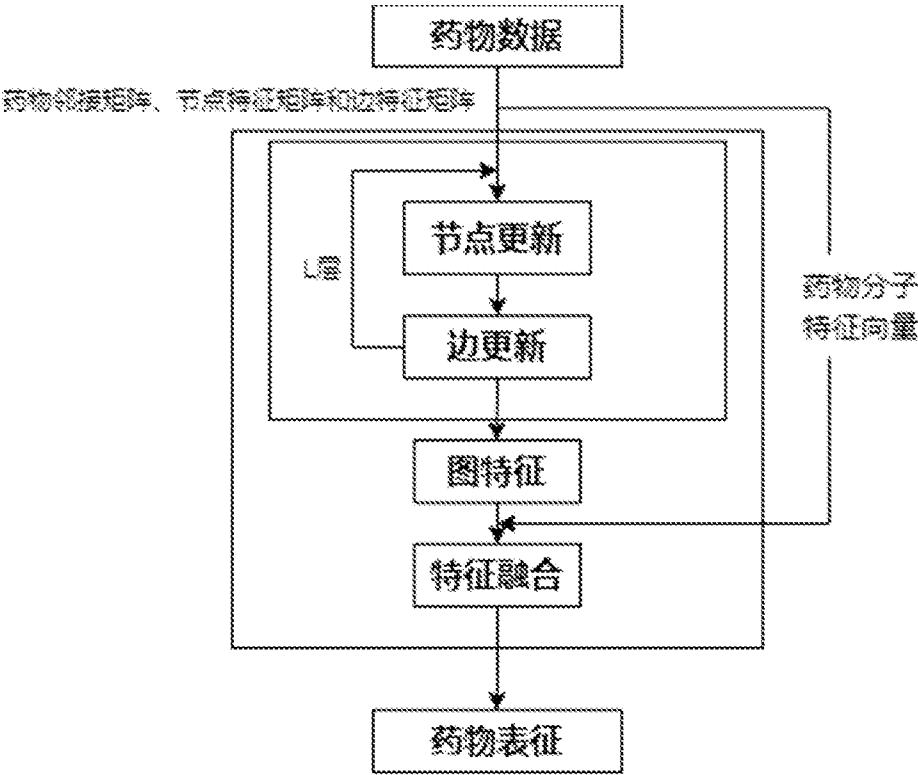


图 2

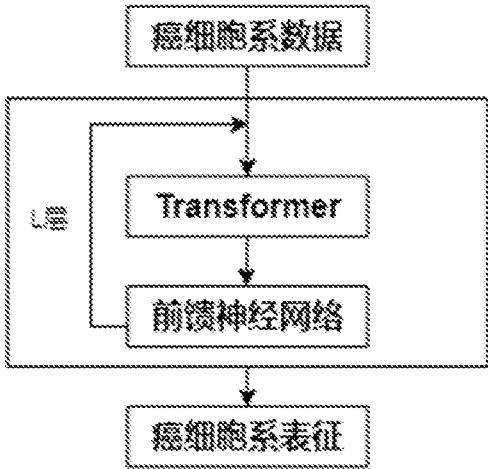


图 3

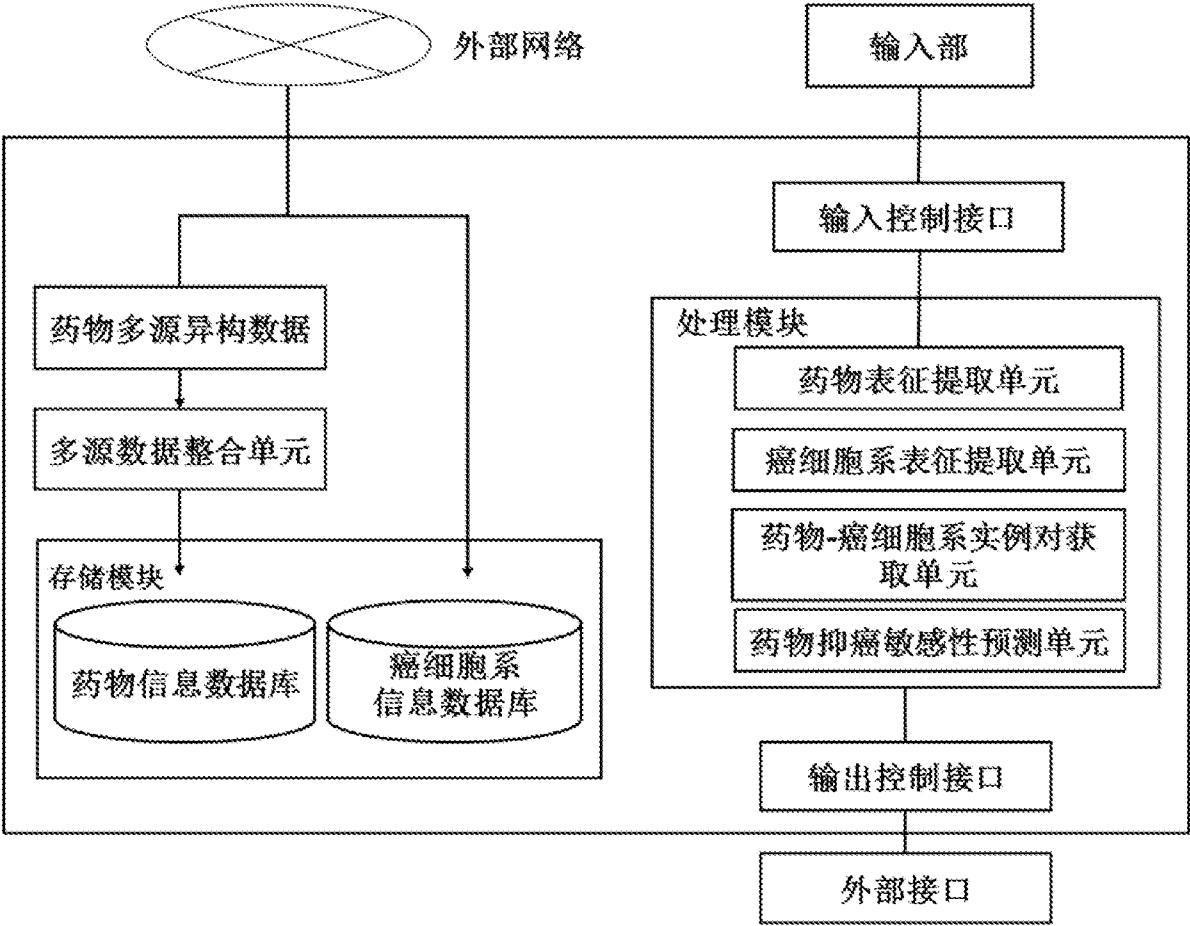


图 4

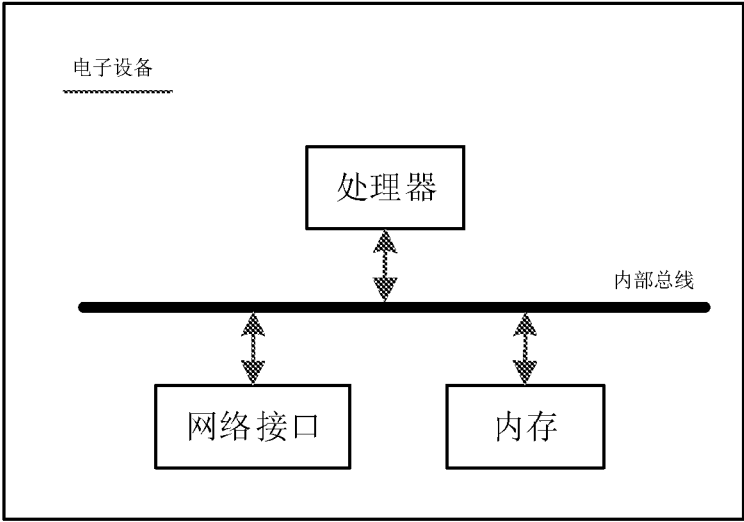


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/098718

A. CLASSIFICATION OF SUBJECT MATTER

G16C20/30(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC:G16C,G16H,G16B,G06F,G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT, ENTXT, ENTXTC, CNKI, DWPI: 药, 癌, 抑癌, 细胞, 特征, 向量, 分子, 原子, 点, 边, 化学键, 矩阵, 图, 神经网络, 基因, 聚合, 融合, 实例对, 预测, drug, cancer, suppression, cell, feature, vector, molecule, atom, chemical bond, matrix, graph, neural network, gene, aggregation, fusion, example pair, predict+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 114974610 A (HUNAN UNIVERSITY) 30 August 2022 (2022-08-30) entire document	1-10
A	CN 114496303 A (HUNAN UNIVERSITY) 13 May 2022 (2022-05-13) entire document	1-10
A	CN 115274136 A (SHANGHAI JIAO TONG UNIVERSITY) 01 November 2022 (2022-11-01) entire document	1-10
A	CN 108830040 A (CENTRAL SOUTH UNIVERSITY) 16 November 2018 (2018-11-16) entire document	1-10
A	CN 113782089 A (ZHEJIANG UNIVERSITY) 10 December 2021 (2021-12-10) entire document	1-10
A	CN 116110509 A (ZHEJIANG UNIVERSITY) 12 May 2023 (2023-05-12) entire document	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

06 February 2024

Date of mailing of the international search report

08 February 2024

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/098718**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2016246919 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 25 August 2016 (2016-08-25) entire document	1-10
A	WO 2022214036 A1 (SHANGHAI ORIGIMED CO., LTD. et al.) 13 October 2022 (2022-10-13) entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/098718

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	114974610	A	30 August 2022	None			
CN	114496303	A	13 May 2022	None			
CN	115274136	A	01 November 2022	None			
CN	108830040	A	16 November 2018	CN	108830040	B	15 June 2021
CN	113782089	A	10 December 2021	CN	113782089	B	18 February 2022
CN	116110509	A	12 May 2023	CN	116110509	B	04 August 2023
US	2016246919	A1	25 August 2016	WO	2015054266	A1	16 April 2015
WO	2022214036	A1	13 October 2022	CN	112768089	A	07 May 2021
				CN	112768089	B	22 June 2021

A. 主题的分类

G16C20/30(2019.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC:G16C,G16H,G16B,G06F,G06N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNTEXT,ENTXT,ENTXTC,CNKI,DWPI,药, 癌, 抑癌, 细胞, 特征, 向量, 分子, 原子, 点, 边, 化学键, 矩阵, 图, 神经网络, 基因, 聚合, 融合, 实例对, 预测, drug, cancer, suppression, cell, feature, vector, molecule, atom, chemical bond, matrix, graph, neural network, gene, aggregation, fusion, example pair, predict+

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 114974610 A (湖南大学) 2022年8月30日 (2022 - 08 - 30) 全文	1-10
A	CN 114496303 A (湖南大学) 2022年5月13日 (2022 - 05 - 13) 全文	1-10
A	CN 115274136 A (上海交通大学) 2022年11月1日 (2022 - 11 - 01) 全文	1-10
A	CN 108830040 A (中南大学) 2018年11月16日 (2018 - 11 - 16) 全文	1-10
A	CN 113782089 A (浙江大学) 2021年12月10日 (2021 - 12 - 10) 全文	1-10
A	CN 116110509 A (浙江大学) 2023年5月12日 (2023 - 05 - 12) 全文	1-10
A	US 2016246919 A1 (UNIV CALIFORNIA) 2016年8月25日 (2016 - 08 - 25) 全文	1-10

☒ 其余文件在C栏的续页中列出。

☒ 见同族专利附件。

★ 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“p” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2024年2月6日

国际检索报告邮寄日期

2024年2月8日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

史江峰

电话号码 (+86) 62089926

PCT/ISA/210 表(第2页) (2022年7月)

C. 相关文件		
类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	WO 2022214036 A1 (SHANGHAI ORIGIMED CO LTD 等) 2022年10月13日 (2022 - 10 - 13) 全文	1-10

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/098718

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	114974610	A	2022年8月30日	无	
CN	114496303	A	2022年5月13日	无	
CN	115274136	A	2022年11月1日	无	
CN	108830040	A	2018年11月16日	CN	108830040 B 2021年6月15日
CN	113782089	A	2021年12月10日	CN	113782089 B 2022年2月18日
CN	116110509	A	2023年5月12日	CN	116110509 B 2023年8月4日
US	2016246919	A1	2016年8月25日	WO	2015054266 A1 2015年4月16日
WO	2022214036	A1	2022年10月13日	CN	112768089 A 2021年5月7日
				CN	112768089 B 2021年6月22日