Biomedical Signal Processing and Control 92 (2024) 106110

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

GDilatedDTA: Graph dilation convolution strategy for drug target binding affinity prediction

Longxin Zhang [a,*], Wenliang Zeng [a], Jingsheng Chen [a], Jianguo Chen [b,*], Keqin Li [c]

[a] College of Computer Science, Hunan University of Technology, Zhuzhou 412007, China
[b] School of Software Engineering, Sun Yat-Sen University, Zhuhai 519082, Guangdong, China
[c] Department of Computer Science, State University of New York, New Paltz, NY 12561, USA

ARTICLE INFO

Keywords:
Deep learning
DTA
Drug repositioning
Graph neural network

ABSTRACT

Traditional drug development methods require a huge amount of time. The utilization of deep learning methods for predicting drug–target binding affinity (DTA) has the potential to expedite the drug screening process, thereby accelerating drug development. The key to accurately predicting DTA depends on effectively mining the potential features of drugs and targets. To address this problem, a DTA prediction model (GDilatedDTA) based on graph dilated convolution strategy, which predicts DTA by learning the potential feature representations of drug–target pairs, is proposed in this study. GDilatedDTA includes three steps. First, the feature encoding module acquires the initial features of the drug molecule and the target through integer encoding. Second, the initial encoded features are output to a subsequent deep learning model to mine their potential features. Third, the potential features are combined to predict DTA. A large number of experimental results show that the mean squared error (*MSE*) of GDilatedDTA on the KIBA and Davis datasets is reduced by 35.71% and 15.57%, on the average, respectively, compared with the state-of-the-art (SOTA) DTA prediction models. Under the drug cold-start setting, the GDilatedDTA's *MSE* score on the Davis dataset is 26.58% lower than the previous SOTA model, more robust than the baselines, and more effective for unseen drugs. In addition, the visualization of drug–target interaction can be beneficial to predict the binding region of both datasets in the structure-based DTA prediction, thereby improving the interpretability of GDilatedDTA.

1. Introduction

Developing new drugs is a time-consuming process that requires extensive drug experiments and ongoing clinical trials. According to statistics, the average development period of a drug is approximately 13 years, and the cost is between $0.618 billion and $2.6 billion [1]. The therapeutic effects of most drugs are realized through in vivo interactions with specific target molecules, such as enzymes, G-protein coupled receptors, and ion channels [2]. Therefore, accurate prediction of drug–target interaction (DTI) can significantly reduce the time spent in the drug development process. The strength of DTI is usually measured by the binding affinity value between the drug and target; thus, DTI prediction is also called drug–target binding affinity (DTA) prediction. A substantial amount of researchers hope to discover potential DTI from existing drugs to accelerate the drug development process, namely, drug repositioning, given the high cost of developing a new drug and the high probability of failure [3]. Drug repositioning seeks to discover novel pharmacological indications for existing drugs, allowing them to be repurposed for treating conditions beyond their initially intended therapeutic use. This approach also aims to unveil new therapeutic applications for well-established drugs [4]. Aspirin, for example, was initially developed to treat various types of pain and inflammation. Through numerous clinical trials, Aspirin is currently one of the basic drugs used to prevent cardiovascular disease and treat heart attacks and strokes. This finding indicates that drug repositioning is effective for drug development.

Molecular docking and machine learning models have been pivotal tools in the field of drug repositioning for an extended period [5]. Molecular docking predicts DTI by modeling the 3D structure of the drug and the target protein. Nevertheless, this approach faces limitations in managing extensive pharmacogenomics data, primarily due to challenges in acquiring 3D structural data for certain biomacromolecules [6]. Traditional machine learning models are widely used in the fields of quantitative structure–activity relationship (QSAR) and

* Corresponding authors.
E-mail addresses: longxinzhang@hut.edu.cn (L. Zhang), wenliang_zeng416@163.com (W. Zeng), jingshengchen01@foxmail.com (J. Chen), chenjg33@mail.sysu.edu.cn (J. Chen), lik@newpaltz.edu (K. Li).

https://doi.org/10.1016/j.bspc.2024.106110
Received 25 May 2023; Received in revised form 3 February 2024; Accepted 9 February 2024
Available online 13 February 2024

proteochemometric [7]. The datasets in these fields are highly imbalanced in terms of positive and negative samples; thus, the prediction accuracy is unsatisfactory. Despite the efforts of researchers to enhance the accuracy of DTI predictions using convolutional neural networks (CNNs) and other intelligent techniques that have been highly successful in the field of computer vision [8–10], these methods still struggle to effectively extract features from non-Euclidean spatial data, such as drug molecule maps and DTI networks. Therefore, the DTI prediction accuracy of CNN models is not as good as that of the conventional machine learning models [11].

In the previous decade, many scholars have proposed various DTA prediction models based on machine learning methods. For example, the compound–protein interaction prediction model based on positive and negative sample learning [12] takes the protein structural domain and compound substructure as protein and compound features, respectively. Then, the tensor product of the two sets of features is calculated as the compound–protein pair feature, which is trained for classification using a biased support vector machine. Pashikala et al. [13] presented the Kronecker regularized least squares (KronRLS) model, which defines the similarity score of a drug–target pair as the Kronecker product of the similarity matrix. To address the linear dependence problem in KronRLS, SimBoost [14] used a gradient boosting approach to construct similarity feature networks between drugs and targets to improve the prediction performance in a nonlinear manner. Although machine learning models show acceptable performance in DTA prediction problem, such algorithms often employ well-designed manual features, which usually require special professional knowledge and experience to combine them [15].

The application of deep learning techniques to DTI prediction problem has received increased attention, owing to the great success of these approaches in diverse fields. Most existing deep learning models predict DTI based on topological similarity. For example, Yuan et al. [16] developed a novel end-to-end deep collaborative learning model named EDC-DTI, designed for predicting DTIs. This model leverages the combined features of drugs and targets derived from both homogeneous and heterogeneous networks. Additionally, it extracts features from the chemical structural attributes of biomolecules and the topological properties of relational networks. Wan et al. [17] developed a nonlinear end-to-end deep learning model named neural integration of neighbor information for DTI prediction (NeoDTI) to improve predictive performance by aggregating information from heterogeneous network data and automatically learning the topological representation of drugs and targets. As research has advanced, numerous scholars have come to the consensus that the molecular structure of drugs constitutes non-Euclidean spatial data. Moreover, there is a prevailing belief that graph neural networks (GNNs) exhibit robust learning capabilities when applied to such data. Therefore, researchers gradually apply the GNN model to the drug repositioning problem. Existing DTA prediction models can be divided into two categories according to different input data: interaction network-based and structure-based prediction models [2]. The interaction network-based DTA prediction model treats the prediction of DTA as a binary classification problem, and predicts the unknown DTI based on the multiple biological networks of known drugs and targets; it outputs 1 when DTI exists, and 0 otherwise. Peng et al. [18] constructed a novel heterogeneous network by analyzing and combining multiple heterogeneous network graphs, suggested a heterogeneous GNN-based approach to learn the feature representations among them, and predicted DTI in an end-to-end approach. The structure-based DTA prediction model considers DTA prediction as a regression problem, predicts DTA based on drug molecule structure and target amino acid sequence, and outputs the interaction strength value of the drug–target pair, i.e., the binding affinity value. Nguyen et al. [19] reported a graph-based model named GraphDTA to predict DTA, which uses GNN and CNN models to extract drug features and target sequence features, respectively. Nonetheless, GNN's difficulty in distinguishing subgraphs with similar topologies has always limited the performance of the above models in DTA prediction. Welling et al. [20] attempted to use one-hot coding to calculate the distance between nodes. However, this approach is highly dependent on the structural information of the entire graph and cannot perform migration learning on graph structure data with new nodes. While GNN effectively address the challenge of handling non-Euclidean structural data [21], it is worth noting that GNN-based DTA prediction models may encounter difficulties in effectively distinguishing subgraphs with highly similar or identical topologies within drug molecular graphs in practical applications. Researchers have attempted to distinguish such subgraphs by stacking the network layers, while the DTA prediction model suffers from gradient disappearance, overfitting, or over-smoothing because the GNN layers are stacked deeply [22]. Therefore, the current GNN-based DTA prediction model is not highly accurate.

The existing DTA prediction models still have the following problems: (1) traditional machine learning models rely overly on manual features and cannot accurately predict DTA without the guidance of professional knowledge; (2) the CNN-based DTA prediction model can automatically extract the features of drug–target pair, but the feature extraction performance of CNN for graph data is always limited; (3) the GNN-based DTA prediction model overcomes the difficulty in processing the graph structure data to a certain extent, but the prediction accuracy should be further improved, especially when facing drug molecular graphs with similar or the same topological structure subgraphs. To address these problems, a DTA prediction model, named graph dilated convolution network drug–target binding affinity (GDilatedDTA) based on graph dilated convolution strategy, is proposed in this study. The fundamental components of GDilatedDTA encompass the multilayer residual connection network (MLRCN) module, the multichannel general aggregation network (McGEN) module, and the bidirectional long short-term memory (BiLSTM) [23] module. MLRCN builds a multilayer residual network to extract local chemical information features in the SMILES sequence. McGEN utilizes multichannel feature fusion mechanism and combines with dilated convolution strategy to integrate node feature representations within the indirect neighborhood of atoms. BiLSTM separately models from different directions of the target sequence to extract potential features of the target sequence. GDilatedDTA improves the prediction accuracy of binding affinity by extracting effective drug–target pair features from the perspective of practical applications. The main contributions of this study are as follows:

- A novel graph dilated convolution strategy based on graph data is proposed and used to construct the McGEN module, which improves the prediction accuracy of the DTA model.
- The MLRCN module is designed to extract the local chemical features of drugs and is exploited to extract the potential structural features in SMILE sequences to enrich the feature information of drugs.
- Extensive experiments have demonstrated that the mean squared error (*MSE*) of GDilatedDTA on Davis and KIBA datasets shows an average performance improvement of 16.10% and 30.68%, respectively, compared with the current state-of-the-art (SOTA) DTA prediction models.

The remainder of the study is organized as follows. Section 2 briefly discusses related works in the field of DTA prediction. Section 3 presents the model structure of GDilatedDTA in detail. Section 4 introduces the experimental settings. Section 5 provides and analyzes the experimental results. Section 6 concludes the entire study.

2. Related work

The evolution of cloud technology not only offers crucial computing power support but also ensures high reliability for deep learning applications [24,25]. Scientists have made significant strides in drug

repositioning by reassessing approved medications, uncovering their potential effectiveness against new diseases [4]. A notable example is Azvudine Tablets, originally intended for HIV treatment, but later discovered to be beneficial in combating COVID-19 [26]. Since these drugs are already in use for humans, their safety profiles have undergone thorough validation. Advanced techniques in virtual screening enable the molecular-level assessment of approved drugs, predicting their potential therapeutic impact on new pathogens or diseases [27]. This approach significantly streamlines the screening process, allowing researchers to swiftly and efficiently pinpoint potential candidates for drug repositioning. The prediction of DTA plays a pivotal role in virtual screening. Within this context, DTA prediction results aid researchers in identifying potential drug–target pairs, offering insights into existing drugs that may hold therapeutic promise for specific diseases [28]. Addressing the DTA prediction challenge, deep learning methods discover potential features within numerous known drug–target pairs by acquiring information about them. These features, often challenging to capture through traditional experimental methods, can be automatically learned and represented by the multi-layer neural network structure inherent in deep learning models [29]. The distinctions among various deep learning methods lie in their model architectures and the representation of input data. Deep DTA prediction (DeepDTA) [30] designs two CNN modules to learn features in the simplified molecular input line entry system (SMILES) sequences of drug molecules and amino acid sequences of target proteins, and then combines them with a deep neural network to predict DTA. DeepMGT-DTI [31] integrates the structural information of drug molecules by using the molecular complementary graph CNN to combine the feature information of multilayer graphs. Moreover, the transformer architecture [32] is used to fuse multilayer graph features to enrich the potential characteristics of drugs, and good results have been achieved in the dataset extracted from DrugBank database [33]. The Transformer is an innovative deep learning model that leverages text-based context-aware technology and has gained significant traction across diverse fields. To illustrate, Cx-CP-ABE [34] employs context-aware technology to integrate hybrid cryptography with blockchain technology, thereby enhancing the security strategy for accessing sensitive data and ensuring data security. Adel et al. utilized context-aware technology to build an agent-based autonomous semantic context-aware health monitoring and disease detection platform [35]. By incorporating context-aware technologies, these platforms not only enhance the adaptability and versatility of health monitoring and disease surveillance systems, but also offer users more accurate and tailored services, considering their diverse data and resource constraints. Different from context-aware technology, deep neural computation (DeepNC) [36] is utilized to investigate drug characteristics. DeepNC combines three distinct GNN models, which are then fed into a fully connected layer to predict DTA. This approach addresses the issue of single GNN models being unable to effectively aggregate neighboring node features, by integrating target potential features into the prediction process.

Drug molecules can be represented in 1D space as sequences, but proteins are macromolecular substances with complex interactions that are difficult to represent in simply 1D space. Despite the limited availability of 3D structure datasets for proteins, a number of researchers have recently begun to use them in their research processes. Ragoza et al. [37] introduced CNN scoring functions designed to extract high-level features from protein–ligand structural data. They posited that CNNs are adept at extracting crucial information from spatial data of this nature. The CNN scoring functions take the 3D representation of protein–ligand interactions as input and enhance interaction predictions by autonomously learning essential features pertinent to binding. Another deep neural network, Pafnucy [38], utilizes 3D grid representations of molecular complexes, factoring in similarities between proteins and ligands. Employing a regularization technique, Pafnucy captures overarching interaction properties, thereby elevating the accuracy of DTA predictions. Nevertheless, acquiring high-quality experimental protein 3D structures poses a significant challenge. Numerous models leveraging protein 3D structures rely on CNNs, yet these CNN-based approaches exhibit sensitivity to various orientations within 3D structures, often demanding substantial computational resources. To address these challenges, recent studies have fused protein 3D structures with GNN models to forecast DTA. For instance, GraphBAR [39] develops a GCN model tailored for DTA prediction using 3D structural insights. This model effectively mitigates computational demands compared to CNN approaches. It represents protein–ligand complexes as graphs, typically comprising multiple adjacency matrices with parameters derived from atomic molecular properties and the identity matrix. Additionally, the authors enhance data through a docking simulation method, further elevating model performance. In a similar vein, Lim et al. [40] introduced a novel deep learning approach for DTI prediction. Their method involves a distance-aware graphical attention algorithm based on GNN, enabling discrimination between various interaction types among molecules. This algorithm extracts graphical features directly from the 3D structural information within the protein–ligand binding site. Consequently, this approach enhances the precision of DTI prediction, moving beyond mere pattern memorization of ligand molecules. Graph-CNN, introduced by Torng et al. [41], operates as an unsupervised learning model designed to learn fixed-size representations from a collection of representative pocket graphs and 2D ligand graphs within protein–ligand complexes. Initially, the model constructs a graphical autoencoder to derive these representations. Subsequently, two Graph-CNNs are trained to extract distinctive features from both the pocket and 2D ligand graphs. This approach showcases the Graph-CNN model's efficiency in autonomously capturing DTIs regardless of the intricacies within protein–ligand complexes. Contrarily, Zheng et al. [42] contested the inefficiency of training models directly using 3D protein structures. They proposed an end-to-end deep learning methodology to predict interactions between proteins and drugs. This innovative approach adopts a visual question-and-answer pattern: proteins are represented through 2D distance maps derived from single structures, while drugs are depicted using molecular linear symbols. This strategy aims to streamline and optimize the prediction process for protein–drug interactions.

3. Materials and methods

3.1. Overview of GDilatedDTA

We develop a novel deep learning model GDilatedDTA for DTA prediction, whose overall framework is shown in Fig. 1. GDilatedDTA views DTA prediction as a regression problem, using drug–target pair as inputs to predict the binding affinity value between the two. Different from the existing methods, GDilatedDTA extracts the direct neighbor node and indirect neighbor node features of atomic nodes, and then combines with the local chemical information learned by MLRCN to obtain drug features. GDilatedDTA consists of three modules, including the feature encoding module (FEM), the representation learning module (RLM), and the DTA prediction module (DTAPM). FEM encodes the initial feature information of the drug–target pair. RLM employs different components to obtain potential feature representations of drugs and targets. DTAPM merges these features and predicts the binding affinity between them.

3.2. Feature encoding module

FEM encodes the initial feature information of a drug–target pair using three separate components, namely, local chemical feature encoding (LCFE), drug structure encoding (DSE), and protein structure encoding (PSE).

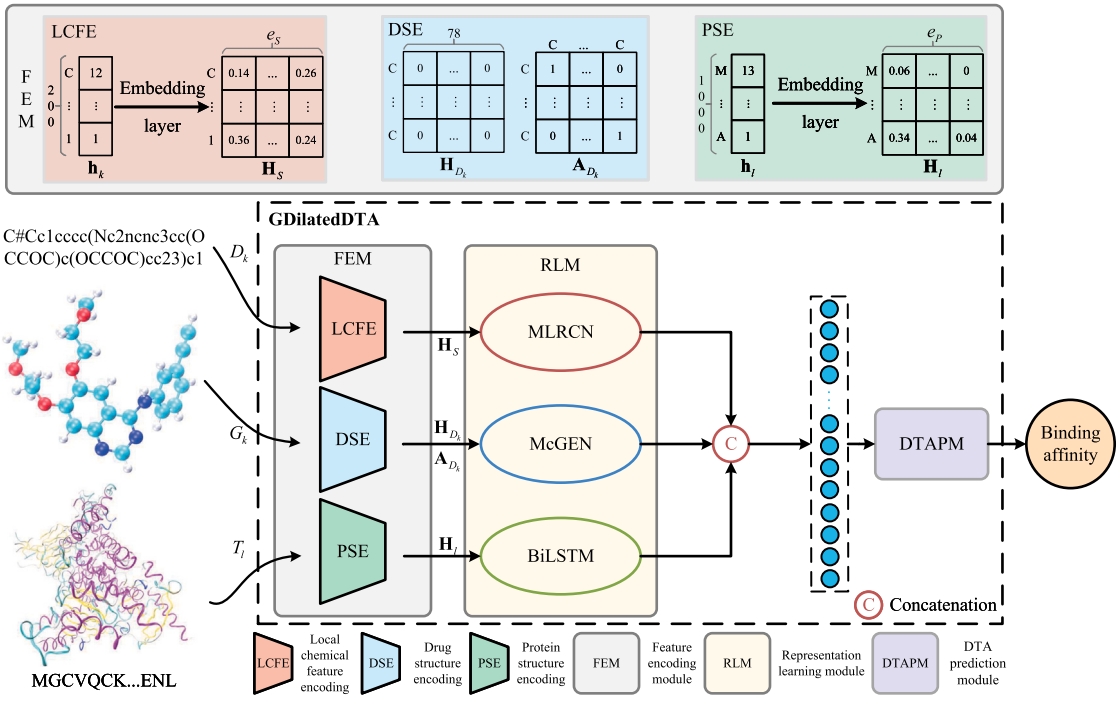

Fig. 1. Overall framework of GDilatedDTA.

3.2.1. Local chemical feature encoding

Drug molecules are commonly depicted using SMILES sequences, utilizing alphanumeric characters to describe the 3D structure of the molecule (for instance, 'S' denotes sulfur atoms, '=' represents carbon–carbon double bonds, and '#' signifies carbon–carbon triple bonds). Drug molecules consist of atoms linked by diverse valence bonds, creating distinct functional groups. This detailed chemical information is considered the local chemical information of the drug. To establish the initial features of the drug, LCFE constructs a mapping relationship from characters to digits (for instance, '1' corresponding to 1 and 'S' corresponding to 32). As illustrated in Fig. 1, the SMILES sequence "C#C...oc1" of drug D_K undergoes the process of character-to-number mapping, resulting in the generation of the integer-encoded feature $\mathbf{h}_k$. Given the varying sequence lengths across different drugs, LCFE establishes a constant initial feature dimension for the drug, specifically set at 1×200. In cases where the sequence length falls short of 200, the remaining dimensions are padded with zeros, and if the length exceeds 200, it is truncated accordingly. Therefore, the dimension of the vector $\mathbf{h}_k$ in Fig. 1 is 1×200. The encoded integer features, represented as $\mathbf{h}_k$ for D_k, are then input into an embedding layer to derive the initial features. The process can be formalized as follows [30]:

$$\boldsymbol{H}_s = Embedding(\boldsymbol{h}_k, e_S),\ \boldsymbol{H}_s \in \mathbb{R}^{200\times e_S}, \tag{1}$$

where $Embedding(\cdot)$ is a random initiation function, and e_S is the dimension of the feature vector for each label value.

3.2.2. Drug structure encoding

In addition to the local chemical information, GDilatedDTA models the drug molecular graph structure to obtain its initial feature representation. DSE uses the open source tool RDKit [43] to transform D_k into the corresponding molecular graph $G_k = (V_k, E_k)$, where $v_i \in V_k (i = 1, 2, \ldots, N^k_{\text{atom}})$ denotes the ith atomic node of D_k, and $e_{i,j} \in E_k (j \in \{1, 2, \ldots, N_{\text{nei}}\})$ denotes the valence bond edge between nodes v_i and v_j. N^k_{atom} and N_{nei} denote the number of atoms and neighbor nodes of v_i, respectively. In the drug molecule graph, GDilatedDTA does not consider the strength of the chemical bonds; thus, G_k is an undirected graph. GDilatedDTA employs the DeepChem method [44] to represent each node, utilizing a 78-dimensional binary feature vector. This approach enables the capture and description of node features in the model. The feature vector contains five types of information, including atom signature $\mathbf{h}_{\text{sym}}$, number of neighboring atom $\mathbf{h}_{\text{nei}}$, number of neighboring hydrogen atom $\mathbf{h}_{\text{H}}$, atomic implied valence $\mathbf{h}_{\text{hid}}$, and whether the atom belongs to the aromatic hydrocarbon structure $\mathbf{h}_{\text{isaro}}$. The node features and their respective encoding details are presented in Table 1. In this process, DSE transforms each node of the drug D_k into a 78-dimensional vector, represented as $\mathbf{H}_{D_k}$ in Fig. 1. The precise calculation formula is elucidated in Eq. (2)

$$\mathbf{h}^k_{v_i} = \left[\mathbf{h}^k_{\text{sym}} \,\middle\|\, \mathbf{h}^k_{\text{nei}} \,\middle\|\, \mathbf{h}^k_{\text{H}} \,\middle\|\, \mathbf{h}^k_{\text{hid}} \,\middle\|\, \mathbf{h}^k_{\text{isaro}}\right],\ \mathbf{h}^k_{v_i} \in \mathbb{R}^{1\times N_{\text{fea}}}, \tag{2}$$

where $\|$ denotes the concatenation of vectors, and N_{fea} represents the dimension of a feature vector of size 78. In this case, the initial feature $\boldsymbol{H}_{D_k}$ of the drug molecule D_k can be expressed as follows:

$$\mathbf{H}_{D_k} = \left[\mathbf{h}^k_{v_1}, \mathbf{h}^k_{v_2}, \ldots, \mathbf{h}^k_{N^k_{\text{atom}}}\right]^T,\ \mathbf{H}_{D_k} \in \mathbb{R}^{N^k_{\text{atom}}\times N_{\text{fea}}}. \tag{3}$$

where $\mathbf{h}^k_{v_1}$ is the feature vector of the first atomic node of D_k. The adjacency matrix $\mathbf{H}_{\text{dir}} = \mathbf{A}_{D_k} \in \mathbb{R}^{N^k_{\text{atom}}\times N^k_{\text{atom}}}$ represents the association features of nodes. $\mathbf{A}_{D_k}(i,j) = \mathbf{A}_{D_k}(j,i) = 1$ when a valence bond connection exists between atomic nodes v_i and v_j; otherwise, it is 0.

3.2.3. Protein structure encoding

Target proteins are biological macromolecules composed of multiple amino acids linked by peptide bonds. Target proteins are usually represented as sequences made up of multiple letters of the alphabet due to the difficulty and low reliability of using 3D structures to represent proteins. Each letter in the sequence stands for an amino acid (i.e., 'A' for alanine and 'R' for arginine). Let $S_{\text{pro}} = \{T_1, T_2, \ldots, T_{N_t}\}$ be the target sequence set, where T_l is the amino acid sequence of the lth target, and N_t is the total number of targets. Similar to the SMILES sequences the amino acid sequences may vary in length from protein to protein. GDilatedDTA sets the dimension size of the target to 1×1000 and processes the same as SMILES. In the PSE module displayed in Fig. 1, protein T_l undergoes encoding in this phase, represented as an integer matrix with a length of 1000. PSE encodes T_l in alphabetical

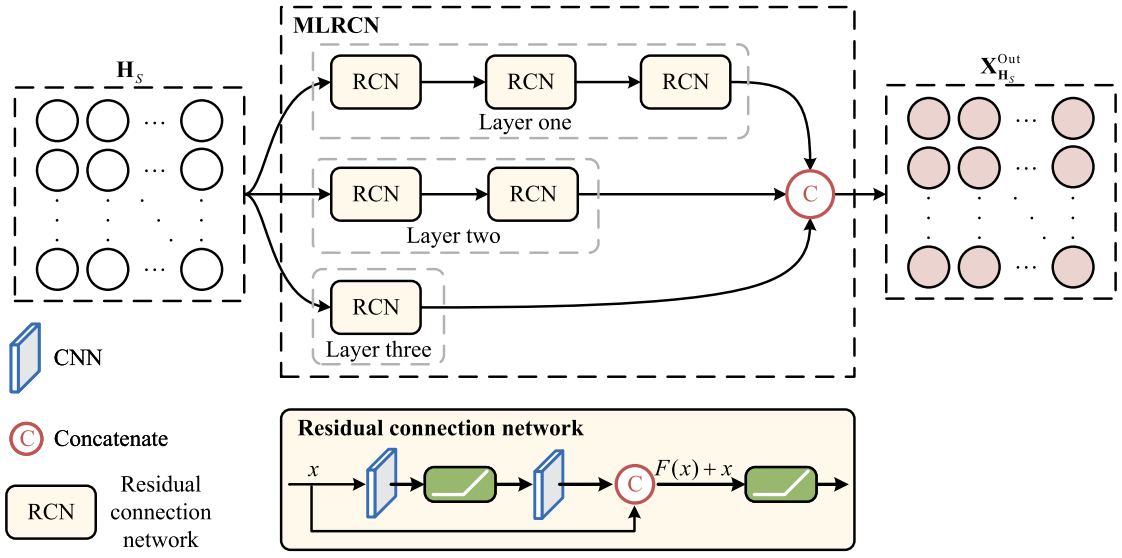

Fig. 2. MLRCN module.

Table 1
Atomic features on the Human and C.elegans datasets.

Name	Description	Dimension
$\mathbf{h}_{\mathrm{sym}}$	[C, N, O, S, F, Si, P, Cl, Br, Mg, Na, Ca, Fe, As, Al, I, B, V, K, Tl, Yb, Sb, Sn, Ag, Pd, Co, Se, Ti, Zn, H, Li, Ge, Cu, Au, Ni, Cd, In, Mn, Zr, Cr, Pt, Hg, Pb, other] (one-hot)	44
$\mathbf{h}_{\mathrm{nei}}$	Number of covalent bonds [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10] (one-hot)	11
$\mathbf{h}_{\mathrm{H}}$	Number of connected hydrogens [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10] (one-hot)	11
$\mathbf{h}_{\mathrm{hid}}$	Implicit valence of the atom [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10] (one-hot)	11
$\mathbf{h}_{\mathrm{isaro}}$	Whether the atom is part of an aromatic system [0/1] (binary)	1

order to obtain the target label $\mathbf{h}_l = [x_1, x_2, \ldots, x_{1000}]^T$. Subsequently, the initial structural feature $\mathbf{H}_l$ of the target is obtained by Eq. (4) [30], as follows:

$$\boldsymbol{H}_l = Embedding(\boldsymbol{h}_l, e_P),\ \boldsymbol{H}_l \in \mathbb{R}^{1000\times e_P}, \tag{4}$$

$\mathbf{H}_l \in \mathbb{R}^{1000\times e_P}$ is the initial structural feature of protein T_l, and e_P represents the dimension of the feature vector for each amino acid.

3.3. Representation learning module

RLM contains three components, including the MLRCN module, the McGEN module, and the BiLSTM module.

3.3.1. Multilayer residual connection network

In GDilatedDTA, the MLRCN module is specifically designed to extract local chemical feature information from SMILES sequences. This is achieved through the utilization of the Residual Connectivity Network (RCN) [45]. RCN was originally used to address the problems of gradient disappearance and network degradation caused by the deepening of deep neural network layers in the field of computer vision. During the DTA prediction process, the traditional CNN cannot effectively extract the local chemical information in the SMILES sequences, whereas the RCN can mine the potential local chemical information in the SMILES sequence while preserving the original features. The structure of MLRCN is shown in Fig. 2.

All parameters within the MLRCN undergo selection via the grid search method, encompassing factors such as the number of RCN layers and the count of nodes in the CNN. Fig. 2 illustrates the RCN module, comprising two CNN layers within each RCN, followed by a ReLU activation function. This design facilitates the RCNs in capturing potential drug features while preserving the inherent characteristics. To mitigate computational overload, each layer within the MLRCN employs fewer than three RCN layers. The RCN layer grid comprises [**(3, 2, 1)**, (3, 3, 3), (2, 2, 2), (1, 1, 1)], with the bold values representing the parameters yielding the best prediction performance. For instance, (3, 2, 1) signifies that the MLRCN consists of three layers, wherein the first layer has 3 RCN layers, the second layer has 2 RCN layers, and the third layer contains 1 RCN layer. The configuration of CNN nodes within the RCN is meticulously designed to ensure the MLRCN can effectively learn significant drug-related features. The grid for the number of CNN nodes is [(512, 256), (256, 128), (128, 96), (96, 64), (64, 32)]. Initially, an identical count of CNN nodes is used for all RCNs, but this approach restricts the model from learning features of varying dimensions in the drug, leading to suboptimal prediction performance. Through the RCN grid analysis, it is observed that the most effective CNN node parameters for (3, 3, 3), (2, 2, 2), and (1, 1, 1) are (128, 96), (96, 64), and (64, 32) respectively. Hence, the MLRCN comprises 128 and 96 CNN nodes per RCN in the first layer, 96 and 64 in the second layer, and 64 and 32 in the third layer. The input to the MLRCN consists of the local chemical information feature $\mathbf{H}_s$, while the final output yields the potential feature representation of the drug's local chemical information, $\mathbf{X}_{\mathbf{H}_s}^{\mathrm{Out}}$.

3.3.2. Multichannel general aggregation network

McGEN utilizes the general aggregation network (GEN) [21] to compose the two-channel extraction of the global structural features of the drug molecular graph.

(1) GEN

GEN was discussed by Li et al. [22] in 2020, it improves the aggregation function of the spectral-based graph convolutional networks (GCN) [20] and describes that the feature extraction capability of the GCN is enhanced by graph residual connections and message normalization. The GCN model can formalize the process of message passing to the nodes of the lth layer, as follows:

$$\mathbf{m}_{vu}^{(l)} = \alpha^{(l)}\left(\mathbf{h}_v^{(l)}, \mathbf{h}_u^{(l)}, \mathbf{h}_{e_{vu}}^{(l)}\right),\quad u \in N(v), \tag{5}$$

$$\mathbf{m}_v^{(l)} = \beta^{(l)}\left(\left\{\mathbf{m}_{vu}^{(l)} \middle| u \in N(v)\right\}\right), \tag{6}$$

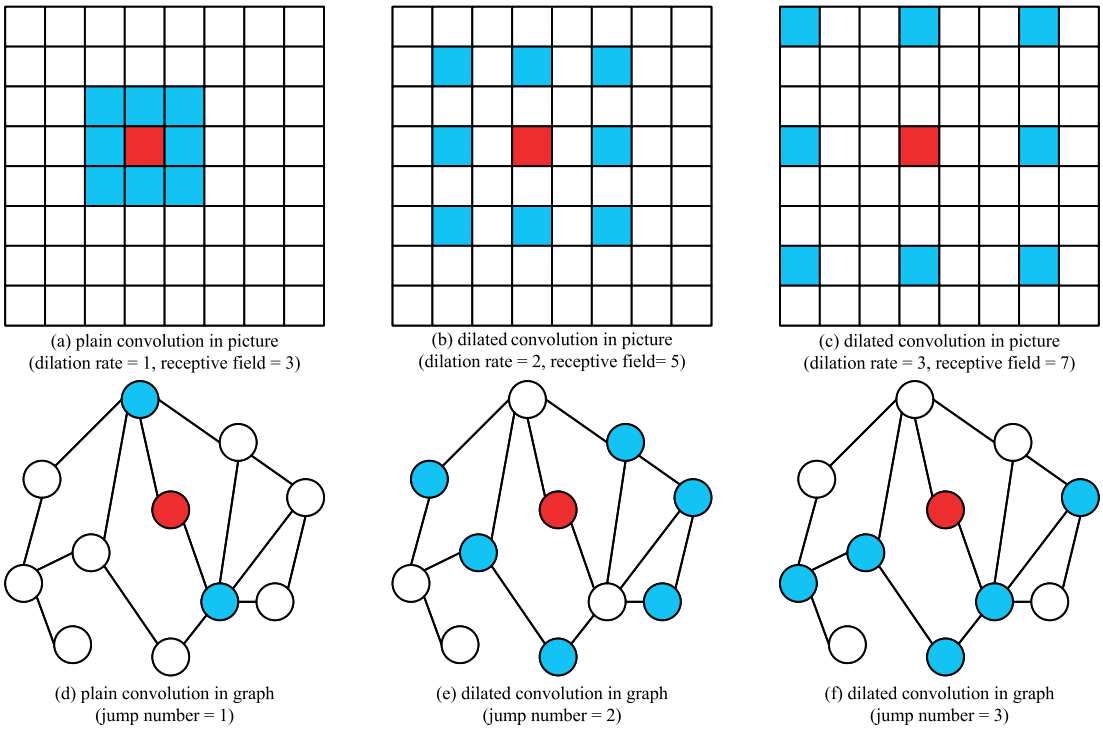

Fig. 3. Dilated convolutions strategy.

$$\mathbf{h}_v^{(l+1)} = \delta^{(l)}\left(\mathbf{h}_v^{(l)}, \mathbf{m}_v^{(l)}\right), \tag{7}$$

where $\alpha^{(l)}$, $\beta^{(l)}$, and $\delta^{(l)}$ represent the message constructor, message aggregation, and vertex update functions of the lth layer, respectively. For simplicity, we only analyze the update of vertex features at each layer. $\alpha^{(l)}$ constructs a separate node information $\mathbf{m}_{vu}^{(l)}$ for each neighbor node $u \in N(v)$ through the vertex feature $\mathbf{h}_v^{(l)}$ of node v at the lth layer, the vertex features $\mathbf{h}_u^{(l)}$ of neighbor node u, and the corresponding edge attribute features $\mathbf{h}_{e_{vu}}^{(l)}$. $\mathbf{h}_{e_{vu}}^{(l)}$ is not considered the drug molecular graph, but an undirected graph. Subsequently, $\beta^{(l)}$ aggregates the neighbor node information to obtain the neighbor feature $\mathbf{m}_v^{(l)}$ of node v. Finally, $\delta^{(l)}$ updates the vertex features $\mathbf{h}_v^{(l+1)}$ of the next layer by synthesizing vertex features and aggregating neighbor features. A single aggregation function cannot effectively aggregate the node features of all neighbor nodes because the number of neighbor nodes is inconsistent. Therefore, GEN adopts the differentiable general aggregation function $SoftMax\_Agg_{\mu}(\cdot)$ to aggregate the neighbor node features, whose expressions are calculated as follows:

$$SoftMax\_Agg_{\mu}(\mathbf{m}_{vu}) = \sum_{u \in N(v)} \frac{\exp(\mu \mathbf{m}_{vu})}{\sum_{i \in N(v)} \exp(\mu \mathbf{m}_{vi})} \cdot \mathbf{m}_{vu}, \tag{8}$$

where μ is a learnable continuous variable with an initial value of 1. With the use of μ, $\beta^{(l)}$ can find the best aggregation function for graphs with different structures and obtain more effective node features by aggregation.

(2) Graph dilated convolution strategy

Dilated convolution network (DCN) [46] was originally used to solve the problem of image segmentation. In the field of image segmentation, dilated convolution can enlarge the receptive field while maintaining the consistent size of the feature image. Figs. 3(a)–(c) show that the receptive field of the image expands as the dilation rate gradually increases. This approach improves the accuracy of the image segmentation algorithm to a certain extent. During the process of aggregating neighboring nodes in the GNN model, the graph structures of the nodes at different locations may be similar or identical, thereby easily producing error node feature representation. Therefore, the graph dilated convolution strategy constructs a new node association matrix, namely, indirect neighbor node association matrix $\mathbf{H}_{\text{indir}}$, using different jump numbers and considering the neighbor node association features. It is calculated by $\mathbf{H}_{\text{indir}} = \mathbf{A}_{D_k}^{Jn}, Jn > 1$, where Jn denotes the number of jumps and the power of matrix $\mathbf{H}_{D_k}$, and $\mathbf{H}_{\text{indir}}(i,j)$ denotes the number of trips from node v_i to node v_j with Jn jumps. Figs. 3(d)–(f) show that the subgraph structure varies as Jn increases. This strategy alleviates the problem of declining prediction accuracy caused by subgraphs with similar or identical structures, thereby extracting rich structural features.

The McGEN module is shown in Fig. 4. McGEN comprises two channels, each of which contains three GENs. GEN(78, 312) indicates that the number of feature channels for input and output are 78 and 312, respectively.

Channel one of McGEN is used to aggregate the features of the direct neighbor nodes of the vertexes. The input of channel one is the initial feature $\mathbf{X}_{D_k}^{(0)}$ of the drug molecule D_k and direct neighbor node association feature $\mathbf{H}_{\text{dir}}$. Then, after the cascade of three layers of GEN, channel one aggregates the neighbor node feature information and updates it to obtain the vertex feature representation $\mathbf{X}_{D_k}^{(3)}$. Channel two exploits the indirect neighbor node with $Jn = 3$ to construct $\mathbf{H}_{\text{indir}}$, thereby aggregating the features of the indirect neighbor node of the current vertex. Similarly, the feature representation of the vertex is obtained after processing by three GEN modules. Finally, McGEN concatenates the output of the two channels to obtain the potential feature representation $\mathbf{X}_{D_k}^{\text{Out}} \in \mathbb{R}^{1\times 1248}$ of D_k. Fig. 4 shows that channel two with the graph dilated convolution strategy evidently aggregates more node feature information, thereby enriching the structural information of the drug molecule.

3.3.3. BiLSTM

The BiLSTM module includes two long short-term memory networks (LSTM). LSTM is used to solve the long-term dependency problem in the recurrent neural network [47]. The input vector of the target sequence is generally denoted as a set of discrete biological words. Therefore, BiLSTM uses LSTM to extract the sequence potential features from the

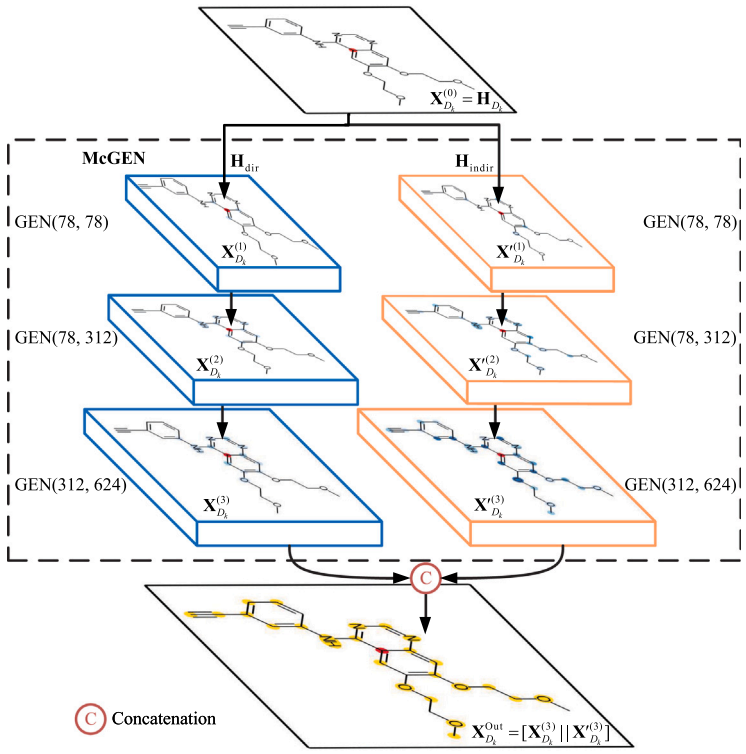

Fig. 4. McGEN module. The red node is the vertex, the blue nodes in the dashed box are the message aggregation nodes of each layer, and the yellow nodes represent the nodes aggregated by McGEN. The darker blue indicates more aggregation.

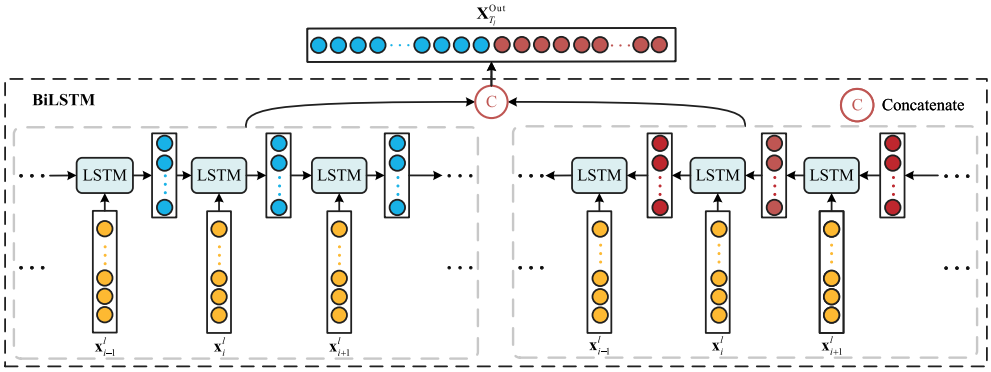

Fig. 5. BiLSTM module.

forward and backward directions of the target sequence. As shown in Fig. 5, the initial features $\mathbf{H}_l$ of the target T_l is considered an example to introduce BiLSTM.

In Fig. 5, $\mathbf{x}_i^l \in \mathbb{R}^{1\times e_P}$ denotes the amino acid vector representation at position i in the target sequence. Amino acid sequence features are input into the LSTM module from both directions, and then BiLSTM combines two sets of features and outputs the embedded feature representation of the target. The process can be formalized, as follows:

$$\mathbf{X}_{T_l}^{\text{Out}} = \left[\overrightarrow{LSTM}(\boldsymbol{H}_l) \middle\| \overleftarrow{LSTM}(\boldsymbol{H}_l)\right], \mathbf{X}_{T_l}^{\text{Out}} \in \mathbb{R}^{1000\times e_{\text{tar}}}, \tag{9}$$

where $\overrightarrow{LSTM}(\mathbf{H}_l)$ denotes the LSTM module that extracts the features from the left end to the right end of the sequence, $e_{\text{tar}} = 2 \times e_P$.

3.4. DTA prediction module

GDilatedDTA combines $\mathbf{X}_{D_k}^{\text{Out}}$, $\mathbf{X}_{\mathbf{H}_s}^{\text{Out}}$, and $\mathbf{X}_{T_l}^{\text{Out}}$ of the RLM output to obtain the drug–target pair features $\mathbf{X}^{\text{Out}} = \left[\mathbf{X}_{D_k}^{\text{Out}} \middle\| \mathbf{X}_{\mathbf{H}_s}^{\text{Out}} \middle\| \mathbf{X}_{T_l}^{\text{Out}}\right]$, which are input into DTAPM to predict DTA. DTAPM is shown in Fig. 6.

$\mathbf{X}^{\text{Out}}$ is the input for DTAPM, and the binding affinity value is calculated after forward propagation. Subsequently, GDilatedDTA updates the network layer weights using back propagation by establishing a loss function for the predicted and true binding affinity values, and these network weights can be used to visualize the region of DTI. The loss function is expressed as follows:

$$\mathcal{L}_{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2, \tag{10}$$

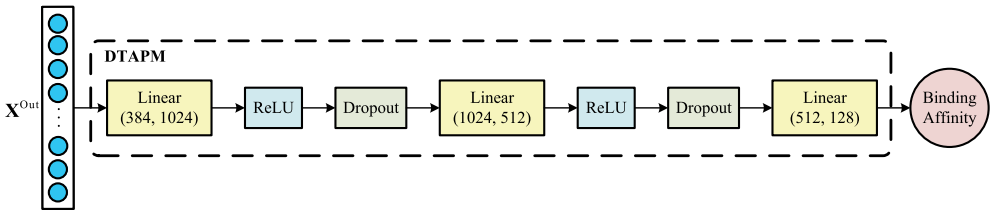

Fig. 6. DTA prediction module.

where Y_i and $\hat{Y}_i$ are the true and predicted binding affinity values, respectively, and n is the number of drug–target pairs. $\mathcal{L}_{MSE}$ represents the *MSE* value between the predicted and true binding affinity values.

4. Experiments

4.1. Baselines

Six current SOTA models are selected as baselines to verify the efficiency of GDilatedDTA.

(1) **KronRLS** [13]: KronRLS forecasts DTA based on the Kronecker regularized least-squares approach. It aims at minimizing the objective function, as follows:

$$\mathcal{J}(f) = \sum_{i=1}^{m} \left(y_i - f(x_i)\right)^2 + \lambda \left|\left|f\right|\right|_k^2, \tag{11}$$

where $x_i (i = 1, 2, \ldots, m)$ is a set of drug–target pair features, $f(\cdot)$ is a nonlinear function, and y_i denotes the true binding affinity value. λ is a regularization parameter, and $\lambda > 0$. $\left|\left|f\right|\right|_k^2$ represents the f-norm consistent with kernel function k.

(2) **SimBoost** [14]: SimBoost describes the attributes of drug, target, and drug–target pair by defining three types of features, respectively, and trains a gradient boost tree model for DTA prediction.

(3) **DeepDTA** [30]: DeepDTA separately trains a three-layer CNN model for the drug and target, and then exploits CNN to learn potential features in SMILES string and target sequence, respectively. Finally, they are connected and passed to the fully connected layer to predict DTA.

(4) **GraphDTA** [19]: GraphDTA is an extension of DeepDTA. It replaces the CNN model for extracting drug features in DeepDTA by applying multiple GNN models, views each drug molecule graph as a graph structure, and mines its potential features. Finally, DTA is predicted via integrating GraphDTA with hidden features of the target learned by CNN models.

(5) **DeepGLSTM** [48]: For target amino acid sequences, the BiLSTM model is applied by DeepGLSTM to extract the potential features. For drug molecular graph, DeepGLSTM exploits the GCN model to mine the atomic features. Finally, the DTA is predicted by concatenating the two sets of features.

(6) **DeepGS** [49]: DeepGS is composed of three modules, namely, CNN, BiGRU, and GAT modules. The CNN module and BiGRU module extract 1D sequence features of targets and drugs, respectively, and the GAT module is used to learn the global topological structure features of drugs.

4.2. Datasets

The two public datasets dedicated to DTA prediction are used in the experiments, as follows:

Davis [50] dataset contains 442 proteins, 68 drug molecules, and 30,056 interaction pairs made up of these proteins and drugs. The binding affinity between drug and protein in this dataset is measured by the dissociation constant K_d. The smaller K_d value indicates stronger affinity. He et al. [14] utilized Eq. (12) to transform it to logarithmic space given that the distribution interval of K_d value is extremely large, and finally, the range of pK_d used to characterize the binding affinity is [5, 10.8]. The expression for pK_d is as follows:

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

KIBA [51] dataset contains 229 protein targets, 2,111 chemical small molecules, and 118,254 drug–target interaction pairs. The DTA in this dataset is measured by KIBA score, which is calculated from a combination of the semi-inhibitory concentration IC_{50} and the inhibition constant K_i and K_d.

4.3. Metrics

GDilatedDTA adopts the same dataset partition ratio (training set to test set ratio is 5:1) and performance evaluation metrics as the other models. The metrics include the MSE, concordance index (CI), and QSAR model external prediction performance metric r_m^2.

(1) MSE

MSE is used to measure the difference between the predicted and true values. The smaller MSE value indicate better model prediction effect. The expression of MSE can be computed by the following equation:

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

where Y_i indicates true value, $\hat{Y}_i$ indicates predicted value, and n is the number of drug–target pairs.

(2) CI

CI is adopted to assess whether the predicted binding affinity value is consistent with the true value in the sequential alignment of drug–target pairs. A larger CI indicates better model performance. CI can be estimated as follows:

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

where N is the normalization constant, which is numerically equal to the number of drug–target pairs with different binding affinity values. f_i and f_j represent the prediction scores corresponding to the binding affinity index y_i and y_j ($y_i > y_j$), respectively. $\xi(x)$ is a transition function, which can be calculated by the following:

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

(3) r_m^2

r_m^2 is used to evaluate the external prediction performance of the QSAR model. It is an acceptable model if and only if $r_m^2 \geq 0.5$. The larger value of r_m^2 indicates better predictive performance of the model. r_m^2 is expressed as follows:

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

where r^2 and r_0^2 denote the squared correlation coefficients with and without intercept between the true and predicted values, respectively.

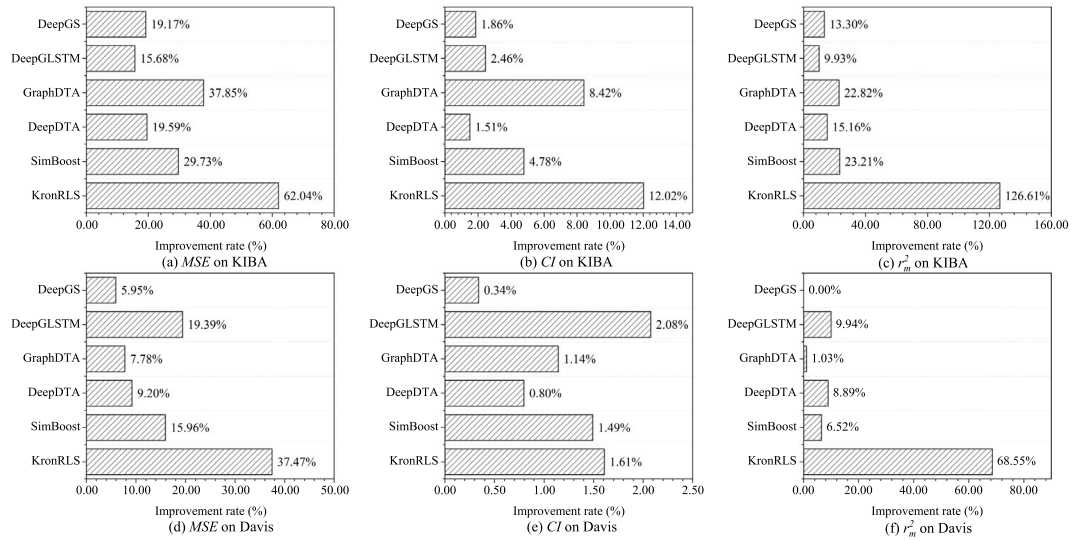

Fig. 7. Improvement rate on KIBA and Davis.

Table 2
Hyperparameter grids of GDilatedDTA in the experiments.

Hyperparameters	Setting
MLRCN layer	[**[3, 2, 1]**, [1, 1, 1], [2, 2, 2], [3, 3, 3]]
CNN nodes	[(512, 256), (256, 128), **(128, 96)**, **(96, 64)**, **(64, 32)**]
e_S	[512, 256, **128**, 64]
e_P	[512, 256, **128**, 64]
GEN layer	[1, **3**, 5, 7, 9]
Optimizer	[**Adam**, Sgd]
Learning rate	[0.0001, **0.0005**, 0.001, 0.002]
Dropout rate	[0.1, **0.2**, 0.3, 0.4, 0.5]
Batch Size	[32, 64, **128**, 256, 512]

4.4. Implementation details

The experimental hardware includes 12-core Intel Core i7-6850K processors and four GTX 1080 Ti graphics cards with memory size of 10 GB. Python 3.6, CUDA 11.4, and Pytorch 1.6.0 deep learning frameworks are built in the Ubuntu 20.04.2 system as running environments. The efficiency of GDilatedDTA is verified on Davis and KIBA datasets. The training epoch of the model is 100, and all hyperparameters in the training process are determined by grid search. The hyperparameter grids are shown in Table 2, where bold values are the hyperparameters used in these experiments.

5. Results and discussion

5.1. Comparison results

Table 3 compares the experimental comparison results of GDilatedDTA with the baselines on the KIBA dataset, where the results of KronRLS, SimBoost, DeepDTA, and DeepGS are obtained from Ref. [49].

Table 3 indicates that the GDilatedDTA model outperforms the other five deep learning algorithms, showcasing MSE, CI, and r_m^2 scores of 0.156, 0.876, and 0.775, respectively. These scores surpass those of the remaining models. Specifically, compared to the other deep learning models, GDilatedDTA demonstrates an average reduction in MSE score by 23.07% while enhancing CI and r_m^2 scores by 3.56% and 15.30%, respectively. Compared to the current best-performing deep learning model, DeepGLSTM, GDilatedDTA delivers a 15.68% decrease in MSE score (0.185 for DeepGLSTM). In terms of the CI score, it boasts a 2.46% enhancement compared to the current leader, DeepDTA, which scored 0.863. Moreover, concerning the r_m^2 score, GDilatedDTA presents a 9.93% improvement over the current best-performing model, DeepGLSTM (DeepGLSTM scored 0.705).

Table 4 presents the experimental results of different models on the Davis dataset.

Table 4 shows that GDilatedDTA outperforms the baselines in all evaluation metrics. The GDilatedDTA model shows an average score reduction of 10.90% compared to the other four deep learning models, accompanied by improvements of 1.09% in CI score and 5.77% in r_m^2 score. This emphasizes the efficacy of GDilatedDTA in addressing the DTA prediction problem. In comparison to DeepGS, the current leading deep learning model, GDilatedDTA demonstrates a 5.95% reduction in the MSE score while enhancing the CI score by 0.003, maintaining an equivalent r_m^2 score. Compared to the DeepGLSTM model, which also utilizes a BiLSTM model for feature extraction from the target, the GDilatedDTA model demonstrates a 19.39% reduction in MSE score, coupled with improvements of 2.08% and 9.94% in CI and r_m^2 scores, respectively. These advancements can be attributed to the integration of McGEN and the well-crafted MLRCN module. McGEN, capable of dynamically adjusting model parameters based on training effects, along with the thoughtfully designed MLRCN module, significantly enhance the efficacy of feature mining in drugs. This synergy contributes to achieving superior prediction outcomes.

To further illustrate the competitiveness of GDilatedDTA in the DTA prediction problem, Eq. (17) is used to calculate the relative improvement rate V_{ir}, as follows:

$$V_{\text{ir}} = (A_{\text{new}} - A_{\text{old}})/A_{\text{old}} \times 100\%, \tag{17}$$

where A_{new} and A_{old} refer to the evaluation metric values for GDilatedDTA and baselines, respectively. The improvement rates of GDilatedDTA over the baselines on the KIBA and Davis datasets are shown in Fig. 7. Fig. 7(a) shows that on the KIBA dataset, the MSE of GDilatedDTA is decreased by 45.89%, on the average, compared with KronRLS and SimBoost. GDilatedDTA carries out an average reduction of 24.18% in terms of MSE compared with the other deep learning models. The *MSE* of GDilatedDTA is reduced by 37.85%, 15.68%, and 19.17% on the KIBA dataset and by 7.78%, 19.39%, and 5.95% on the Davis dataset compared to the GraphDTA, DeepGLSTM, and DeepGS models without the graph dilated convolution strategy. The results show that GDilatedDTA with graph dilated convolution strategy can effectively extract the deep information of drug molecules and target amino acid sequences to achieve accurate prediction of DTA.

Table 3
Comparison results of GDilatedDTA and baselines on the KIBA dataset.

Model		Drug (graph)	Drug (SMILES)	Target	MSE	CI	r_m^2
Machine learning model	KronRLS	PubChem	–	S-W	0.411	0.782	0.342
	SimBoost	PubChem	–	S-W	0.222	0.836	0.629
Deep learning model	DeepDTA	CNN	–	CNN	0.194	0.863	0.673
	GraphDTA	GIN	–	CNN	0.251	0.808	0.631
	DeepGLSTM	GCN	–	BiLSTM	0.185	0.855	0.705
	DeepGS	GAT	BiGRU	CNN	0.193	0.86	0.684
Deep learning model (proposed method)	GDilatedDTA	McGEN	MLRCN	BiLSTM	**0.156**	**0.876**	**0.775**

The values in bold indicate the optimal results achieved by the detection model with respect to the current evaluation metric.

Table 4
Comparison results of GDilatedDTA and baselines on the Davis dataset.

Model		Drug (graph)	Drug (SMILES)	Target	MSE	CI	r_m^2
Machine learning model	KronRLS	PubChem	–	S-W	0.379	0.871	0.407
	SimBoost	PubChem	–	S-W	0.282	0.872	0.644
Deep learning model	DeepDTA	CNN	–	CNN	0.261	0.878	0.630
	GraphDTA	GIN	–	CNN	0.257	0.875	0.679
	DeepGLSTM	GCN	–	BiLSTM	0.294	0.867	0.624
	DeepGS	GAT	BiGRU	CNN	0.252	0.882	**0.686**
Deep learning model (proposed method)	GDilatedDTA	McGEN	MLRCN	BiLSTM	**0.237**	**0.885**	**0.686**

5.2. Cold-start setting

The cold-start setting refers to the input of data that has not been seen during the model training phase. To solve practical issues, researchers often predict drugs or targets that has never seen before. The challenge of the DTA prediction problem in the cold-start setting is whether a model that has excellent performance on a specific dataset can also perform well on unknown data. This experiment assesses the robustness of the model in the face of new environments (e.g., mutated proteins) by testing the predictive performance of GDilatedDTA under the cold-start setting.

In the cold-start settings, we conduct a comparative analysis involving GDilatedDTA and four other deep learning models, namely GraphDTA [19], GLFA [52], GEFA [52], and FusionDTA [53]. GLFA and GEFA are novel DTA prediction models proposed in Ref. [52], which use a novel graph-in-graph neural network with an attention mechanism to solve the problem of changing target representations caused by binding effects. FusionDTA model was introduced in Ref. [53] as a DTA prediction model based on a multi-headed linear attention mechanism, which aggregates global information according to attention weights to achieve accurate DTA predictions.

Table 5 reports the performance of models under the cold-start settings of drug, target, and drug–target on the Davis dataset, corresponding to unseen drugs, proteins, and drug–target pairs, respectively. As can be seen from Table 5, among GLFA and GEFA, GEFA demonstrates superior effectiveness. In three distinct cold-start settings, the MSE score of GDilatedDTA experiences a reduction of 40.90%, 11.98%, and 27.40% compared to GEFA. This improvement can be attributed to the utilization of a graph-in-graph neural network with an attention mechanism in the GEFA model, which concentrates attention on the binding effect. Nevertheless, the binding interaction between a drug and its target is inherently determined by their respective characteristics. In this regard, the GDilatedDTA model delves deeply into the inherent features of both drugs and targets, yielding more accurate predictive outcomes. Notably, when compared to GraphDTA, the improvements brought about by GDilatedDTA are more pronounced. Across three distinct environments, the MSE score sees an average reduction of 35.74%, while the CI score exhibits an average improvement of 9.48%. It can be seen that the *MSE* score of GDilatedDTA in the drug cold-start setting is 0.500, which is reduced by 26.58% compared to FusionDTA. GDilatedDTA is 1.87% higher than FusionDTA in terms of *CI* score. In the cold-start settings of target and drug–target, FusionDTA shows better prediction performance than GDilatedDTA, with lower *MSE* scores of 0.051 and 0.002, and higher *CI* scores of 0.016 and 0.052 than GDilatedDTA, respectively. Such differences are due to the fact that GDilatedDTA focuses most of its attention on drugs, while BiLSTM exhibits only limited performance in extracting target features. FusionDTA extracts more effective target features by using the transformer framework to construct global features of targets. Although the prediction performance of GDilatedDTA in the cold-start scenarios of target and drug–target does not surpass the existing best model (FusionDTA), GDilatedDTA shows a remarkable advantage in the drug cold-start setting. The experiments fully demonstrate the enhanced robustness of GDilatedDTA in predicting unseen drugs under the cold-start setting constraint, which also shows appropriate performance in predicting unseen targets or unseen drug–target pairs.

Table 5
The performance of cold-start setting of drug, protein, and drug–protein on the Davis dataset.

Model	Drug cold-start		Protein cold-start		Drug–protein cold-start	
	MSE	*CI*	*MSE*	*CI*	*MSE*	*CI*
GraphDTA [19]	0.920	0.675	0.510	0.706	1.130	0.627
GLFA [52]	0.861	0.670	0.453	0.780	1.144	0.636
GEFA [52]	0.846	0.709	0.434	0.795	0.989	0.639
FusionDTA [53]	0.681	0.747	**0.331**	**0.826**	**0.716**	**0.685**
GDilatedDTA	**0.500**	**0.761**	0.382	0.810	0.718	0.633

5.3. Ablation study

The core strategies of GDilatedDTA are to construct subgraphs of different structures under the graph dilated convolution strategy and to employ MLRCN to mine the local chemical structure features of drug molecules to improve the accuracy of the model in the DTA prediction problem. Four variants of GDilatedDTA are implemented in our study, namely, GDilatedDTA_J1, GDilatedDTA_J2, GDilatedDTA_J4, and GDilatedDTA_NS to validate the efficiency of graph dilated convolution strategy and MLRCN. The difference between the four variant models lies in the selection of Jn size and whether the MLRCN module is used. The specific setting is shown in Table 6. The variant models in the experiment adopt the same hyperparameter settings as GDilatedDTA. This section compares the values of three evaluation metrics of different variants on two datasets, and the comparison results are presented in Fig. 8.

Fig. 8 clearly illustrates that the models with the graph dilated convolution strategy ($Jn > 1$) generally outperform GDilatedDTA_J1 ($Jn = 1$) on the KIBA and Davis datasets, where the MSE of GDilatedDTA on the KIBA is 16.96% lower than that of GDilatedDTA_J1.

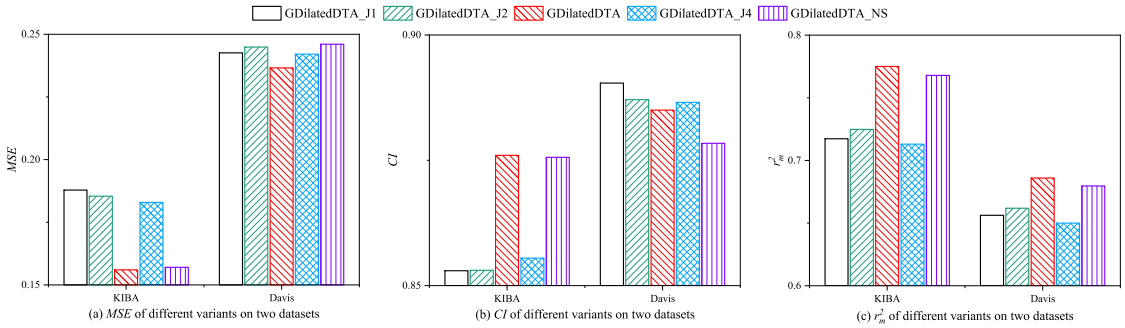

Fig. 8. Metrics for evaluation of different variants on two datasets.

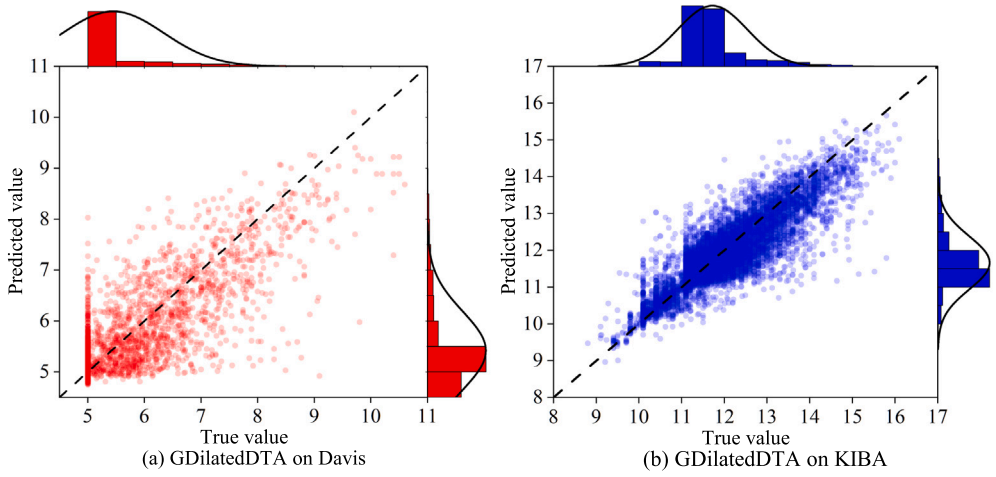

Fig. 9. True affinity value against the predicted value on Davis and KIBA datasets.

Table 6
GDilatedDTA model and four variants of GDilatedDTA.

Model	Jn	MLRCN
GDilatedDTA_J1	1	✓
GDilatedDTA_J2	2	✓
GDilatedDTA	3	✓
GDilatedDTA_J4	4	✓
GDilatedDTA_NS	3	–

Meanwhile, the prediction performance of the model gradually improves with the increase in Jn. Fig. 8 shows that the model achieves the best prediction performance when Jn is 3. When Jn is greater than 3, the prediction performance of the model decreases. GDilatedDTA and GDilatedDTA_NS show roughly the same performance on both datasets, whereas the former performs better. On the KIBA dataset, the MSE, CI and r_m^2 of GDilatedDTA_NS are 0.157, 0.876, and 0.768, respectively. On the contrast, the MSE index of GDilatedDTA is 0.65% lower, and the index is 0.92% higher. The reason for this finding is that the drug SMILES sequences of the KIBA dataset may lose some functional group information in the feature encoding process, resulting in no significant effect of the MLRCN module. Note here that the CI score of GDilatedDTA on the Davis dataset is 0.005, 0.002, and 0.002 lower than those of GDilatedDTA_J1, GDilatedDTA_J2, and GDilatedDTA_J4, respectively. This phenomenon is related to the size of the dataset. The CI score of GDilatedDTA surpasses GDilatedDTA_J1, GDilatedDTA_J2, and GDilatedDTA_J4 by 0.023, 0.023, and 0.021, respectively, on the KIBA dataset with a larger amount of data.

Evidently, the graph dilated convolution strategy can effectively extract the potential features of drug molecules, and MLRCN can learn local chemical information to further enrich the potential features of drugs and finally realize accurate DTA prediction.

5.4. Model analysis

This section further validates the prediction performance of GDilatedDTA on the test sets of Davis and KIBA. The x-axis in Fig. 9 denotes the true value T, and the y-axis represents the predicted value P. An excellent model's predicted value should be close to the true value, that is, $P \rightarrow T$. Hence, the predicted value should fall on or near the black dashed line.

Fig. 9 shows the difference in the predicted value density between the KIBA and Davis datasets. The darker color in the figure indicates greater density. The majority of the predicted values of GDilatedDTA on the Davis dataset in Fig. 9(a) is in the range of [5, 6], whereas the dense region on the KIBA dataset in Fig. 9(b) is within [10, 14]. Fig. 9(a) shows that the drug–target pairs with binding affinity values within [5, 6] on the Davis dataset account for more than 50% of the entire dataset (i.e., 24,495/30,056). Similarly, the intensive area of KIBA's predicted value in Fig. 9(b) falls in [10, 14]. This group of experiments confirm that GDilatedDTA has promising predictive performance.

The protein sequence and target sequence are truncated, that is, the feature matrices of D_k and T_l are changed, and then the predictive performance of the model is tested on the KIBA dataset. FEM sets the lengths of the SMILES sequence and the target amino acid sequence to [50, 100, 150, **200**, 250] and [100, 500, **1000**, 1500 2000], respectively, where the bold indicates the default value.

Let L_S and L_P be the length of SMILES sequences and target sequences, respectively. Fig. 10 shows the prediction results of GDilatedDTA when L_S is set to 50, 100, 150, 200, and 250. Fig. 10 shows that GDilatedDTA reaches the best performance when L_S = 200. To analyze this phenomenon, the statistical results of the length distribution of the SMILES sequences and target sequences on KIBA are shown in Fig. 12, where the average length of SMILES sequence and

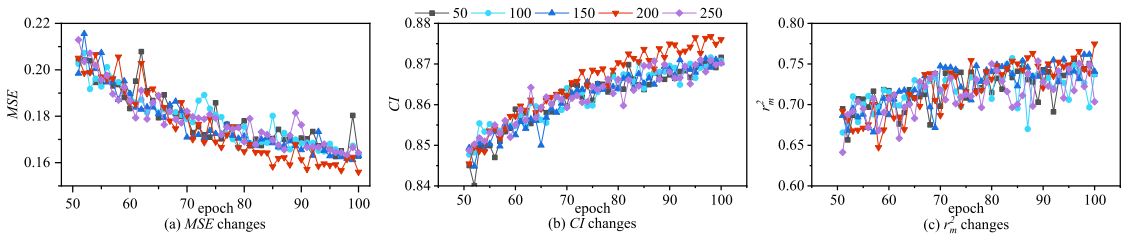

Fig. 10. Comparison of MSE, CI and r_m^2 of GDilatedDTA with different L_S on the KIBA dataset.

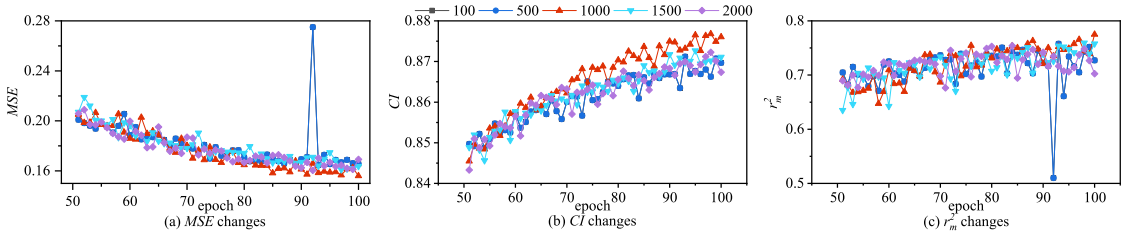

Fig. 11. Comparison of MSE, CI and r_m^2 of GDilatedDTA with different L_P on the KIBA dataset.

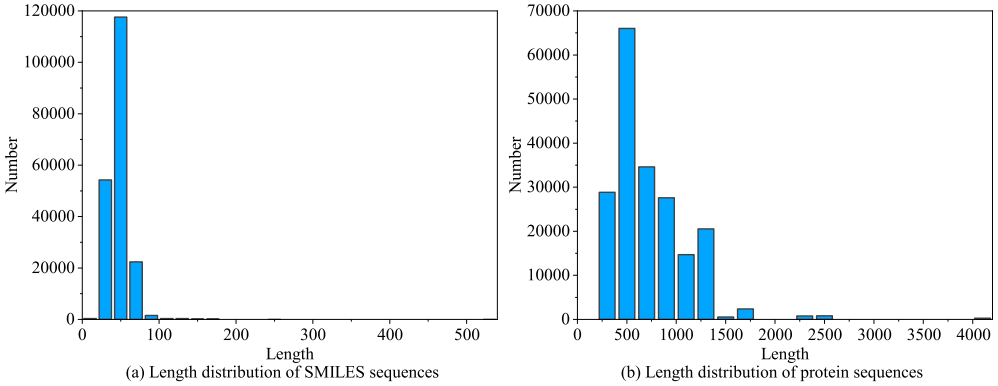

Fig. 12. Length distribution of SMILES and protein sequences on the KIBA dataset.

target sequence are 85 and 752, respectively. Fig. 12(a) shows that the length of the SMILES sequence of most drugs is in the range of [50, 200], whereas that for a minority of drugs is greater than 250. This finding is a good explanation for why this phenomenon occurs. LCFE can obtain effective information about the local chemical structure of the drug only by cutting part of the SMILES sequence or filling in a small number of zero values to when $L_S = 200$.

Fig. 11 shows the prediction results of GDilatedDTA when L_P is set to 100, 500, 1000, 1500, and 2000 on the KIBA dataset. Evidently, GDilatedDTA has the best prediction effect when $L_P = 1000$. Fig. 12(b) illustrates that most targets have amino acid sequence ranging in length from 500 to 1500, whereas a few targets have amino acid sequence greater than 2000. We infer that the underlying reason is that PSE retains most of the available information of the target when $L_P = 1000$, thereby producing accurate feature representation and improving the model prediction accuracy of GDilatedDTA.

5.5. Visualization with network weights

The weight matrices of different network layers in GDilatedDTA can be used to analyze which part of the drug–target interaction plays a key role in the binding pocket. The weight matrix can be used to calculate the partial critical regions of DTI. In this experiment, the amino acid sequence of the target and SMILES features of the drug's small molecule is initially calculated, and then the weights of the amino acid and atom node are computed by back propagation, thereby visualizing the key interaction regions. Fig. 13 shows an example of weight visualization for GDilatedDTA. PI3Kdelta (PDB ID: 5I6U) and CHEMBL3806195 are selected for interaction visualization analysis. The region with the weight value of more than 1.07E−2 in the protein and all the atomic nodes in the drug molecule are colored to highlight the interaction region. In Fig. 13, the blue color in the protein structure and the orange color in the drug molecule map denote the locations in the binding pocket, where the target interacts with drug's small molecule. The darker orange color in the drug molecule map indicates higher weight value, and the darker blue color indicates smaller weight value. Most weight values of amino acids in PI3Kdelta are concentrated from 6.54E−4 to 5.76E−3. The major amino acid regions captured by the GDilatedDTA model are residues 812–919, where residues 907–914 are relatively heavy, and the peak is at ASP-911, which exactly falls into the binding pocket. Fig. 13 shows that the drug atom node binding to residue ASP-911 occupies a large weight value as well. Weight visualization operations can help researchers reveal the manner of interaction between drugs and targets, understand the interaction patterns of drug atoms and residues, and ultimately facilitate the drug repositioning process. The results of weight visualization analysis demonstrate that the proposed GDilatedDTA model can accurately predict the binding location of DTI, while improving the interpretability of GDilatedDTA.

6. Conclusion and future work

The deep learning model has been widely accepted to mine the potential feature information in raw data. We develop a novel DTA

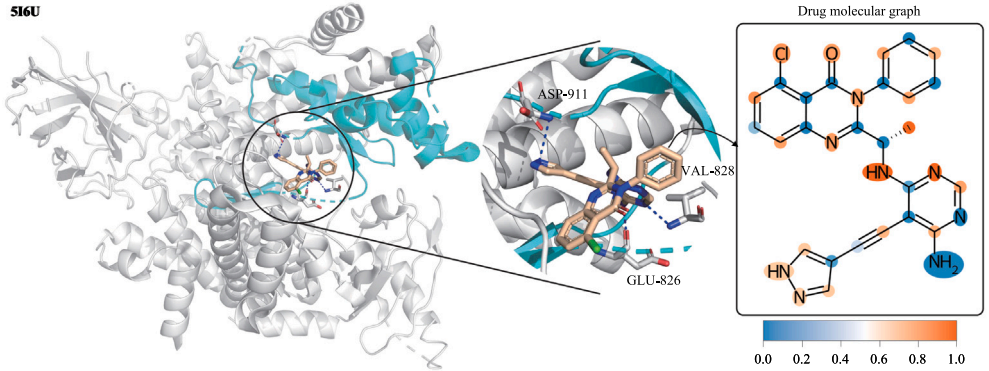

Fig. 13. Example of weight visualization of GDilatedDTA model. PI3Kdelta (PDB ID: 5I6U) protein is expressed in 3D structure, whereas the drug molecule CHEMBL3806195 is expressed in stick form and molecular graph.

prediction model named GDilatedDTA based on deep learning. GDilatedDTA designs a graph dilated convolution strategy to construct atomic node association feature matrices to aggregate feature information of nodes in indirect neighborhoods. It also combines drug local chemical information and target sequence time-series feature information to realize accurate prediction of DTA. Extensive experiments are conducted on two public datasets, namely, Davis and KIBA, and the experimental results are compared with KronRLS, SimBoost, DeepDTA, GraphDTA, DeepGLSTM, and DeepGS. The experimental results show that the MSE, CI, and r_m^2 metrics of GDilatedDTA on the KIBA dataset achieve an average improvement of 30.68%, 5.17%, and 35.17% over the current SOTA models. This finding demonstrates the high efficiency of GDilatedDTA in the DTA prediction problem. Furthermore, GDilatedDTA demonstrates greater robustness than current SOTA DTA prediction models in the drug cold-start setting. In addition, GDilatedDTA explores the explainability of the DTA models by mapping network layer weights to drug atoms and amino acids, further demonstrating the reliability of GDilatedDTA.

GDilatedDTA exclusively takes into account the 2D structure of the target. As advancements in protein structure prediction methodologies continue, obtaining 3D structures of proteins is expected to become more accessible in the future. Consequently, our forthcoming research aims to investigate the DTA prediction in the context of 3D protein structures.

CRediT authorship contribution statement

Longxin Zhang: Writing – review & editing, Writing – original draft, Software, Methodology, Funding acquisition, Conceptualization. **Wenliang Zeng:** Writing – original draft, Investigation, Data curation. **Jingsheng Chen:** Validation, Software. **Jianguo Chen:** Writing – review & editing, Funding acquisition. **Keqin Li:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank two anonymous reviewers for their suggestions to improve the manuscript. This work was partially funded by the Natural Science Foundation of Hunan Province, China (Grant Nos. 2023JJ50204), the Scientific Research Foundation of Hunan Provincial Education Department, China (Grant No. 23B0560), the National Key R&D Program of China (Grant No. 2018YFB1003401), and the National Natural Science Foundation of China (Grant Nos. 61702178, 62072172, 62002110, 62372486),

References

[1] V. Prasad, S. Mailankody, Research and development spending to bring a single cancer drug to market and revenues after approval, JAMA Internal Med. 177 (11) (2017) 1569–1575.
[2] Z. Zhang, L. Chen, F. Zhong, D. Wang, J. Jiang, S. Zhang, H. Jiang, M. Zheng, X. Li, Graph neural network approaches for drug-target interactions, Curr. Opin. Struct. Biol. 73 (2022) 102327.
[3] X. Yang, Z. Niu, Y. Liu, B. Song, W. Lu, L. Zeng, X. Zeng, Modality-DTA: Multimodality fusion strategy for drug-target affinity prediction, IEEE/ACM Trans. Comput. Biol. Bioinform. 20 (2) (2023) 1200–1210.
[4] Y. Hua, X. Dai, Y. Xu, G. Xing, H. Liu, T. Lu, Y. Chen, Y. Zhang, Drug repositioning: Progress and challenges in drug discovery for various diseases, Eur. J. Med. Chem. 234 (2022) 114239.
[5] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R.K. Ambasta, P. Kumar, Artificial intelligence to deep learning: Machine intelligence approach for drug discovery, Mol. Diversity 25 (2021) 1315–1360.
[6] Z. Hu, W. Liu, C. Zhang, J. Huang, S. Zhang, H. Yu, Y. Xiong, H. Liu, S. Ke, L. Hong, SAM-DTA: A sequence-agnostic model for drug–target binding affinity prediction, Brief. Bioinform. 24 (1) (2023) bbac533.
[7] Z. Wu, M. Zhu, Y. Kang, E.L.-H. Leung, T. Lei, C. Shen, D. Jiang, Z. Wang, D. Cao, T. Hou, Do we need different machine learning algorithms for QSAR modeling? A comprehensive assessment of 16 machine learning algorithms on 14 QSAR data sets, Brief. Bioinform. 22 (4) (2021) bbaa321.
[8] L. Zhang, J. Chen, J. Chen, Z. Wen, X. Zhou, LDD-Net: Lightweight printed circuit board defect detection network fusing multi-scale features, Eng. Appl. Artif. Intell. 129 (2024) 107628.
[9] A. Soualmi, L. Laouamer, A. Adel, A blind watermarking approach based on hybrid imperialistic competitive algorithm and SURF points for color images' authentication, Biomed. Signal Process. Control 84 (2023) 105007.
[10] L. Zhang, Y. Hu, J. Chen, C. Li, K. Li, MSSIF-Net: An efficient CNN automatic detection method for freight train images, Neural Comput. Appl. 35 (9) (2023) 6767–6785.
[11] B. Xu, K. Cen, J. Huang, H. Shen, X. Cheng, A review of graph convolutional neural networks, J. Comput. 43 (05) (2020) 755–780.
[12] X. Lin, Z. Quan, Z.-J. Wang, Y. Guo, X. Zeng, S.Y. Philip, Effectively identifying compound-protein interaction using graph neural representation, IEEE/ACM Trans. Comput. Biol. Bioinform. 20 (2) (2022) 932–943.
[13] T. Pahikkala, A. Airola, S. Pietilä, S. Shakyawar, A. Szwajda, J. Tang, T. Aittokallio, Toward more realistic drug–target interaction predictions, Brief. Bioinform. 16 (2) (2015) 325–337.
[14] T. He, M. Heidemeyer, F. Ban, A. Cherkasov, M. Ester, SimBoost: A read-across approach for predicting drug–target binding affinities using gradient boosting machines, J. Chem. 9 (1) (2017) 1–14.

[15] Y. Wu, M. Gao, M. Zeng, J. Zhang, M. Li, BridgeDPI: A novel graph neural network for predicting drug–protein interactions, Bioinformatics 38 (9) (2022) 2571–2578.
[16] Y. Yuan, Y. Zhang, X. Meng, Z. Liu, B. Wang, R. Miao, R. Zhang, W. Su, L. Liu, EDC-DTI: An end-to-end deep collaborative learning model based on multiple information for drug-target interactions prediction, J. Mol. Graph. Model. 122 (2023) 108498.
[17] F. Wan, L. Hong, A. Xiao, T. Jiang, J. Zeng, NeoDTI: Neural integration of neighbor information from a heterogeneous network for discovering new drug–target interactions, Bioinformatics 35 (1) (2019) 104–111.
[18] J. Peng, Y. Wang, J. Guan, J. Li, R. Han, J. Hao, Z. Wei, X. Shang, An end-to-end heterogeneous graph representation learning-based framework for drug–target interaction prediction, Brief. Bioinform. 22 (5) (2021) bbaa430.
[19] T. Nguyen, H. Le, T.P. Quinn, T. Nguyen, T.D. Le, S. Venkatesh, GraphDTA: Predicting drug–target binding affinity with graph neural networks, Bioinformatics 37 (8) (2021) 1140–1147.
[20] T.N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, 2016, arXiv preprint arXiv:1609.02907.
[21] G. Bouritsas, F. Frasca, S. Zafeiriou, M.M. Bronstein, Improving graph neural network expressivity via subgraph isomorphism counting, IEEE Trans. Pattern Anal. Mach. Intell. 45 (1) (2022) 657–668.
[22] G. Li, C. Xiong, G. Qian, A. Thabet, B. Ghanem, DeeperGCN: Training deeper GCNs with generalized aggregation functions, IEEE Trans. Pattern Anal. Mach. Intell. 45 (11) (2023) 13024–13034.
[23] Z. Wei, Z. Dai, BiLSTM with novel feature matrix predicts the binding affinity between MHC-I and peptides, in: Proceedings of the 5th International Conference on Big Data Technologies, ICBDT '22, Association for Computing Machinery, New York, NY, USA, 2022, pp. 351–356.
[24] L. Zhang, M. Ai, K. Liu, J. Chen, K. Li, Reliability enhancement strategies for workflow scheduling under energy consumption constraints in clouds, IEEE Trans. Sustain. Comput. (2023) 1–14.
[25] L. Zhang, M. Ai, R. Tan, J. Man, X. Deng, K. Li, Efficient prediction of makespan matrix workflow scheduling algorithm for heterogeneous cloud environments, J. Grid Comput. 21 (4) (2023) 75.
[26] X. Ji, X. Meng, X. Zhu, Q. He, Y. Cui, Research and development of Chinese anti-COVID-19 drugs, Acta Pharm. Sinica B 12 (12) (2022) 4271–4286.
[27] S. Singh, K. Kumar, M. Panda, A. Srivastava, A. Mishra, V.K. Prajapati, High-throughput virtual screening of small-molecule inhibitors targeting immune cell checkpoints to discover new immunotherapeutics for human diseases, Mol. Diversity 27 (2) (2023) 729–751.
[28] Q. Pei, L. Wu, J. Zhu, Y. Xia, S. Xie, T. Qin, H. Liu, T.-Y. Liu, R. Yan, Breaking the barriers of data scarcity in drug–target affinity prediction, Brief. Bioinform. 24 (6) (2023) bbad386.
[29] M. Kalemati, M. Zamani Emani, S. Koohi, Bicomp-DTA: Drug-target binding affinity prediction through complementary biological-related and compression-based featurization approach, PLoS Comput. Biol. 19 (3) (2023) e1011036.
[30] H. Öztürk, A. Özgür, E. Ozkirimli, DeepDTA: Deep drug–target binding affinity prediction, Bioinformatics 34 (17) (2018) i821–i829.
[31] P. Zhang, Z. Wei, C. Che, B. Jin, DeepMGT-DTI: Transformer network incorporating multilayer graph information for drug–target interaction prediction, Comput. Biol. Med. 142 (2022) 105214.
[32] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L.u. Kaiser, I. Polosukhin, Attention is all you need, in: I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), in: Advances in Neural Information Processing Systems, vol. 30, Curran Associates, Inc., 2017.
[33] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, et al., DrugBank 5.0: a major update to the DrugBank database for 2018, Nucleic acids research 46 (D1) (2018) D1074–D1082.
[34] B. Annane, A. Alti, L. Laouamer, H. Reffad, Cx-CP-ABE: Context-aware attribute-based access control schema and blockchain technology to ensure scalable and efficient health data privacy, Secur. Privacy 5 (5) (2022) e249.
[35] A. Alti, L. Laouamer, Agent-based autonomic semantic context-aware platform for smart health monitoring and disease detection, Comput. J. 65 (3) (2022) 736–755.
[36] H.N.T. Tran, J.J. Thomas, N.H.A.H. Malim, Deepnc: A framework for drug-target interaction prediction with graph neural networks, PeerJ 10 (2022) e13163.
[37] M. Ragoza, J. Hochuli, E. Idrobo, J. Sunseri, D.R. Koes, Protein–ligand scoring with convolutional neural networks, J. Chem. Inf. Model. 57 (4) (2017) 942–957.
[38] M.M. Stepniewska-Dziubinska, P. Zielenkiewicz, P. Siedlecki, Development and evaluation of a deep learning model for protein–ligand binding affinity prediction, Bioinformatics 34 (21) (2018) 3666–3674.
[39] J. Son, D. Kim, Development of a graph convolutional neural network model for efficient prediction of protein-ligand binding affinities, PLoS One 16 (4) (2021) e0249404.
[40] J. Lim, S. Ryu, K. Park, Y.J. Choe, J. Ham, W.Y. Kim, Predicting drug–target interaction using a novel graph neural network with 3D structure-embedded graph representation, J. Chem. Inf. Model. 59 (9) (2019) 3981–3988.
[41] W. Torng, R.B. Altman, Graph convolutional neural networks for predicting drug-target interactions, J. Chem. Inf. Model. 59 (10) (2019) 4131–4149.
[42] S. Zheng, Y. Li, S. Chen, J. Xu, Y. Yang, Predicting drug–protein interaction using quasi-visual question answering system, Nat. Mach. Intell. 2 (2) (2020) 134–140.
[43] M.A. Thafar, M. Alshahrani, S. Albaradei, T. Gojobori, M. Essack, X. Gao, Affinity2Vec: Drug-target binding affinity prediction through representation learning, graph mining, and machine learning, Sci. Rep. 12 (1) (2022) 4751.
[44] M. Korshunova, B. Ginsburg, A. Tropsha, O. Isayev, OpenChem: A deep learning toolkit for computational chemistry and drug design, J. Chem. Inf. Model. 61 (1) (2021) 7–13.
[45] L. Zhang, J. Chen, W. Zeng, Z. Tong, Z. Wen, X. Zhang, FDNet: Lightweight train image fault detection network in edge computing environments, IEEE Sens. J. 23 (20) (2023) 25105–25115.
[46] F. Yu, V. Koltun, T. Funkhouser, Dilated residual networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2017, pp. 472–480.
[47] G. Sharma, A. Parashar, A.M. Joshi, DepHNN: A novel hybrid neural network for electroencephalogram (EEG)-based screening of depression, 66, 2021, 102393,
[48] S. Mukherjee, M. Ghosh, P. Basuchowdhuri, DeepGLSTM: Deep graph convolutional network and LSTM based approach for predicting drug-target binding affinity, in: Proceedings of the 2022 SIAM International Conference on Data Mining, SDM, SIAM, 2022, pp. 729–737.
[49] X. Lin, K. Zhao, T. Xiao, Z. Quan, Z.-J. Wang, P.S. Yu, DeepGS: Deep representation learning of graphs and sequences for drug-target binding affinity prediction, in: ECAI 2020, IOS Press, 2020, pp. 1301–1308.
[50] M.I. Davis, J.P. Hunt, S. Herrgard, P. Ciceri, L.M. Wodicka, G. Pallares, M. Hocker, D.K. Treiber, P.P. Zarrinkar, Comprehensive analysis of kinase inhibitor selectivity, Nature Biotechnol. 29 (11) (2011) 1046–1051.
[51] J. Tang, A. Szwajda, S. Shakyawar, T. Xu, P. Hintsanen, K. Wennerberg, T. Aittokallio, Making sense of large-scale kinase inhibitor bioactivity data sets: A comparative and integrative analysis, J. Chem. Inf. Model. 54 (3) (2014) 735–743.
[52] T.M. Nguyen, T. Nguyen, T.M. Le, T. Tran, Gefa: Early fusion approach in drug-target affinity prediction, IEEE/ACM Trans. Comput. Biol. Bioinform. 19 (2) (2021) 718–728.
[53] W. Yuan, G. Chen, C.Y.-C. Chen, FusionDTA: Attention-based feature polymerizer and knowledge distillation for drug-target binding affinity prediction, Brief. Bioinform. 23 (1) (2022) bbab506.