Computers in Biology and Medicine 173 (2024) 108376

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

G-K BertDTA: A graph representation learning and semantic embedding-based framework for drug-target affinity prediction

Xihe Qiu [a,1], Haoyu Wang [a,1], Xiaoyu Tan [b,1], Zhijun Fang [c,*]

[a] School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai, China
[b] INF Technology (Shanghai) Co., Ltd., Shanghai, China
[c] School of Computer Science and Technology, Donghua University, Shanghai, China

ARTICLE INFO

Keywords:
Drug-target affinity
Molecular semantics
Drug properties mining
Graph attention networks
DTA prediction

ABSTRACT

Developing new drugs is costly, time-consuming, and risky. Drug-target affinity (DTA), indicating the binding capability between drugs and target proteins, is a crucial indicator for drug development. Accurately predicting interaction strength between new drug-target pairs by analyzing previous experiments aids in screening potential drug molecules, repurposing them, and developing safe and effective medicines. Existing computational models for DTA prediction rely on strings or single-graph neural networks, lacking consideration of protein structure and molecular semantic information, leading to limited accuracy. Our experiments demonstrate that string-based methods may overlook protein conformations, causing a high root mean square error (RMSE) of 3.584 in affinity due to a lack of spatial context. Single graph networks also underperform on topology features, with a 6% lower confidence interval (CI) for activity classification. Absent semantic information also limits generalization across diverse compounds, resulting in 18% increment in RMSE and 5% in misclassifications within quantifications study, restricting potential drug discovery. To address these limitations, we propose **G-K BertDTA**, a novel framework for accurate DTA prediction incorporating protein features, molecular semantic features, and molecular structural information. In this proposed model, we represent drugs as graphs, with a GIN employed to learn the molecular topological information. For the extraction of protein structural features, we utilize a DenseNet architecture. A knowledge-based BERT semantic model is incorporated to obtain rich pre-trained semantic embeddings, thereby enhancing the feature information. We extensively evaluated our proposed approach on the publicly available benchmark datasets (i.e., KIBA and Davis), and experimental results demonstrate the promising performance of our method, which consistently outperforms previous state-of-the-art approaches. Code is available at https://github.com/AmbitYuki/G-K-BertDTA.

1. Introduction

Drug-target affinity (DTA) refers to the binding ability between drugs and their targeted proteins, which reflects the degree of binding between drugs and their target molecules [1]. It is one of the important indicators for drug efficacy and new drug development [2]. The higher the DTA, the tighter the binding between drugs and targets, and the stronger the therapeutic effect of the drug [3]. Therefore, accurate prediction of DTA can assist in screening more potential drug molecules, providing new targets for disease treatment, and assisting in repurposing new drugs [4,5]. This can also guide drug dosage, administration methods, and treatment time, thereby improving the efficacy and safety of drugs [6].

Currently, numerous computational models have been developed to forecast drug-target affinity [7,8]. One common approach is to calculate the interaction forces between drug molecules and target molecules [9,10] to predict the affinity between them. An alternative strategy involves analyzing the tridimensional structure of target molecules and predicting drug molecule conformation [11] to establish their association with the target [12,13]. Quantum chemistry calculations can also be used to analyze the quantum mechanical interactions between drug targets and protein compositions. For example, density functional theory (DFT) can simulate and quantify the intermolecular forces dictating binding affinity, such as hydrogen bonds, charge transfers, and π-stacking [14]. However, the computational complexity of DFT limits throughput, with typical predictions taking upwards of five hours per drug-target complex [15]. Compared to data-driven deep learning techniques that can evaluate thousands of candidates within minutes with better accuracy (MSE of 0.21), this approach

* Corresponding author.
E-mail address: zjfang@dhu.edu.cn (Z. Fang).
[1] Equal contribution.

https://doi.org/10.1016/j.compbiomed.2024.108376
Received 21 December 2023; Received in revised form 21 March 2024; Accepted 24 March 2024
Available online 25 March 2024

hinders high-volume screening. Additionally, large amounts of drug-target affinity data can be collected and used as input strings for binary classification [16] or regression problem [17] analysis. Deep learning models have achieved notable advancements in predicting drug-target affinity [18,19], leading to the development of various algorithms based on drug-target interactions (DTI). They are computationally efficient and accurate in prediction [20], and are playing an increasingly important role in drug discovery [21,22]. However, most of the above models are based on the structural information of protein strings and ignore the spatial structural information of the target molecule [23,24]. More recently, graph-based DTA models have been proposed [25], which consider both the structural information of proteins and drugs. The simplified molecular input line entry system (SMILES) is a standardized method for representing molecules as strings of characters. The SMILES molecules contain multidimensional feature relationships of target information, which can better calculate molecular properties and molecular docking information [26], and facilitate the model to search for and learn molecular structure features. The aforementioned models use the SMILES molecule information as input for the graph structure in the graph neural network (GNN) and leverage convolutional neural network (CNN) to extract higher-dimensional molecular features of proteins [27]. Unlike models that primarily rely on string-based structural information, graph-based approaches are capable of concurrently processing both SMILES data and protein structural information, thereby achieving more accurate predictions.

Existing methods mainly rely on graph neural networks (e.g., GNN, GCN, and GAT) to process SMILES formulas [28]. However, these models can result in the loss of vital information [29] and cannot consider the spatial order information [30] in the SMILES sequence [31], hindering the exploration of dependence relationships between various chemical bonds and chiral molecules [32]. Furthermore, the graph networks utilized in these methods cannot fully utilize the topology information of drug molecules [33]. To overcome these challenges, we utilized the Graph Isomorphism Network (GIN) to acquire features from SMILES molecule information [34]. In comparison to GCN and GAT, which only consider local node features [35–37], GIN focuses on the global feature representation of the graph, enabling the learning of comprehensive spatial order, chiral, and topology information in SMILES formulas [38].

The state-of-the-art (SOTA) drug-target affinity (DTA) prediction approach ignores the semantic information between molecules [39]. This leads to the model's inability to fully extract the complex features of the molecule [40]. Integrating semantic information can aid in detecting the correlation between drugs and targets [41]. For instance, certain specific functional groups in a drug molecule's chemical structure can predict its interaction with specific targets [42]. Additionally, it can improve generalization ability [43] and enable better handling of diverse SMILES molecule structures [44,45].

To enhance the accuracy of DTA prediction, we introduce the **G-K BertDTA** framework. This model comprehensively integrates protein feature information, molecular semantic features, and structural data of molecules. In this framework, we utilize a GIN model to learn the topological data of SMILES molecules and derive features for molecular structure analysis. Additionally, we propose a novel DenseSENet architecture tailored for protein feature extraction that contains DenseNet and squeeze-and-excitation (SE) blocks. DenseNet enables representational reuse through dense-layer interconnections. After that, We integrate SE blocks to re-calibrate channel-wise feature importance. We also leverage knowledge-based BERT (KB-BERT), a pre-trained language model based on bidirectional encoder representations from transformers, to extract semantic features of SMILES molecules. These features collectively enhance the completeness and accuracy of our predictions. We then concatenate these embeddings and input them into the output layer to calculate affinity. This approach effectively integrates protein feature information, molecular semantic features, and molecular structural information to improve DTA prediction.

The key contributions of our work are summarized as follows.

- To address the high dimensionality and spatial complexity of protein features, we have redesigned and improved the DenseNet network, called DenseSENet. Leveraging the advantages of information flow and feature reuse in DenseNet, we also introduce the SE blocks to adaptively learn channel weights and adjust the activation level of feature maps. The SE blocks evaluate and select the importance of feature channels, enhancing the salient features for classification tasks. Our DenseSENet effectively utilizes features from all previous layers and captures important features from protein sequences.
- To enhance high-dimensional feature extraction from molecular structures using SMILES formulas and address the issue of missing node labels, we employ an improved graph isomorphism network (GIN) to capture high-dimensional relational features between isomorphic SMILES structures. By learning the topological representation of molecules, the GIN network can further improve the understanding of SMILE's structural characteristics and adaptively learn the features of the graph. Furthermore, we employ CNNs for additional feature dimensionality reduction, enabling a more targeted focus on critical features.
- Through the KB-BERT semantic model, we learn the semantic features of SMILES representation and consider important features that were previously ignored in predicting molecules. This enhances the accuracy and robustness of our model, enabling it to better predict the essential characteristics of molecules. To the best of our knowledge, no prior model network has been specifically designed for extracting semantic information from SMILES molecular formulas.
- We design a specific framework based on graph representation learning and pre-trained semantic embeddings for precise prediction of drug-target affinity. We extensively evaluate our framework and compare it with SOTA DTA methods. Our method demonstrates a significant improvement, with a confidence interval (CI) increase of 0.019 and a decrease in mean squared error (MSE) to 0.12, compared to the baseline method, respectively. These results indicate that our method will contribute to the advancement of medical research and the development of safe and effective drugs. Our GraphDTA model has general applicability beyond drug-target affinity prediction, as it can be applied to any similar problem that involves graph-based data input with semantic features.

The rest of the paper is structured as follows. Preliminaries and the proposed framework are provided in Section 2. A brief overview of the related work is provided in Section 3, then the experimental findings and ablation investigations are presented in Section 4. Finally, we draw a conclusion in Section 5, and the overall diagram of our proposed framework is in Fig. 1.

2. Materials and methodology

2.1. Data collection

We utilized DAVIS [46] and KIBA [47] to evaluate the performance of our proposed model. Both datasets are commonly used benchmarks in drug-target interaction prediction and are widely recognized in the field. The DAVIS dataset consists of 68 drug-protein pairs and 442 SMILES targets [48]. The KIBA dataset comprises a combination of different biological activity data for kinase inhibitors, such as Ki, Kd, and IC50 [49]. We filter the dataset for drugs and targets that have at least ten interactions, resulting in 229 unique proteins and 2111 unique drugs [50]. We provide an overview of the datasets in Table 1.

Furthermore, the predictive capabilities of our G-K BertDTA approach demonstrate strong generalizability across BindingDB [51], Therapeutic Target [52] [53], and DrugMAP [54] datasets. BindingDB is a publicly accessible database of measured binding affinities focusing on the interactions between proteins considered to be drug

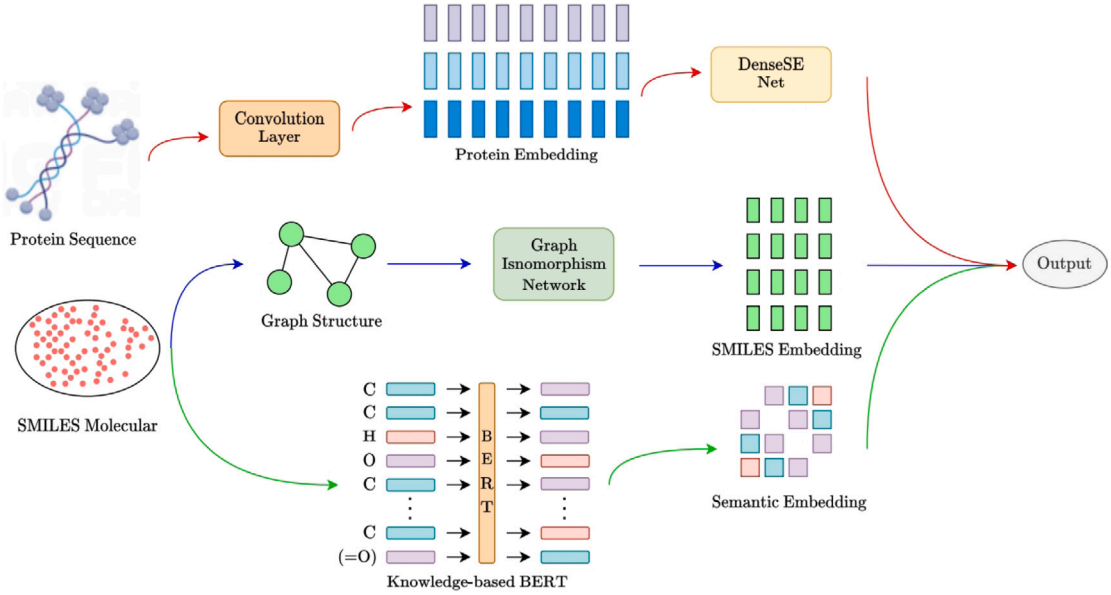

Fig. 1. Framework overview of the proposed G-K BertDTA model. It comprises (1) DenseSENet for protein feature extraction, (2) GIN to learn topological representations of molecular graphs, (3) KB-BERT that acquires semantic knowledge of SMILES, and (4) feature interaction and model training to predict drug-target binding affinity.

Table 1
The datasets details used for the experiment.

Dataset	Drugs	Targets	Interactions	Train	Validation	Test
Davis	72	442	30,056	21,039	3006	6011
KIBA	2116	229	118,254	82,778	11,825	23,651
BindingDB	770,124	7352	1,735,582	1,219,907	173,987	341,688
TTD	38,583	3527	130,524	91,367	13,052	26,105
DrugMAP	31,284	5489	201,942	141,360	20,194	40,388

targets and small, drug-like molecules that bind to them [55]. It contains 1,735,582 binding data entries for 7352 protein targets and 770,124 small molecules.[2] By compiling a vast amount of quantitative data on protein-ligand binding affinities, BindingDB enables better understanding and prediction of interactions important for drug discovery.

Therapeutic Target Database (TTD) [56] is a drug target database that provides information about known and explored therapeutic protein and nucleic acid targets, targeted diseases, pathway data, and the corresponding approved, clinical trial and experimental drugs.[3] As of 2022, TTD contains over 3500 drug targets and 38,000 drug molecules. For each target, TTD integrates data including the target's sequence, pathway, disease, structure information, associated drugs, and clinical trials.

DrugMAP [54] is a new dataset that describes the molecular portraits and drug information. It provides a complete list of interacting molecules for over 30,000 drugs/drug candidates and gives the differential expression patterns of over 5000 interacting molecules in various disease sites,[4] ADME (Absorption, Distribution, Metabolism, and Excretion) related organs, and physiological tissues. Furthermore, an extensive and accurate network containing over 200,000 interactions between drugs and molecules has been compiled.

2.2. Experiment setup and implementation

The experiment is conducted on an NVIDIA GeForce RTX 3090 GPU with an 8-core CPU. The training for each validation iteration lasted for 4 h. Table 2 provides a detailed description of the hyperparameter settings. Our code is available at the link.[5]

Table 2
Detailed description of computational resources and hyperparameters.

Hardware	Configuration
GPU	NVIDIA GeForce RTX 3090 24GB GDDR6X
GPU Memory Bandwidth	936 GB/s
GPU Boost Clock	1.7 GHz
GPU CUDA Cores	10 496
CPU	AMD Ryzen 9 5950X (16 Cores, 32 Threads)
CPU Base Clock	3.4 GHz
CPU Boost Clock	4.9 GHz
System Memory	64GB DDR4 3600MHz
Storage	1TB NVMe SSD
Network	10 Gigabit Ethernet
Software Environment	TensorFlow, PyTorch, CUDA 11.2, cuDNN 8.1
Average Training Time	4 h/iteration
Peak Training Time	6 h/iteration
Batch Size	512
Learning Rate	0.0005
Optimizer	Adam
Loss Function	Mean Squared Error
Protein Sequence Length	256 amino acids
SMILES Sequence Length	128 tokens
Dense Blocks	3
Convolution Layers per Block	3
Dropout Rate	0.1
Epochs	300

The consistency index (CI) is an essential metric in evaluating the predictive performance of DTA [23,25]. It measures the concordance between the estimated binding affinity scores and the true values for drug-target pairs. We use paired t-tests with a 95% CI to handle statistical significance, and CI is widely employed to measure the degree of prediction of binding strength values in protein-ligand interactions relative to true values. The formula used to compute CI is given below.

$$CI = \frac{1}{Z} \sum_{\delta_i > \delta_j} h\left(b_i - b_j\right) \tag{1}$$

[2] https://www.bindingdb.org/rwd/bind
[3] https://db.idrblab.net/ttd
[4] http://drugmap.idrblab.net
[5] https://github.com/AmbitYuki/G-K-BertDTA

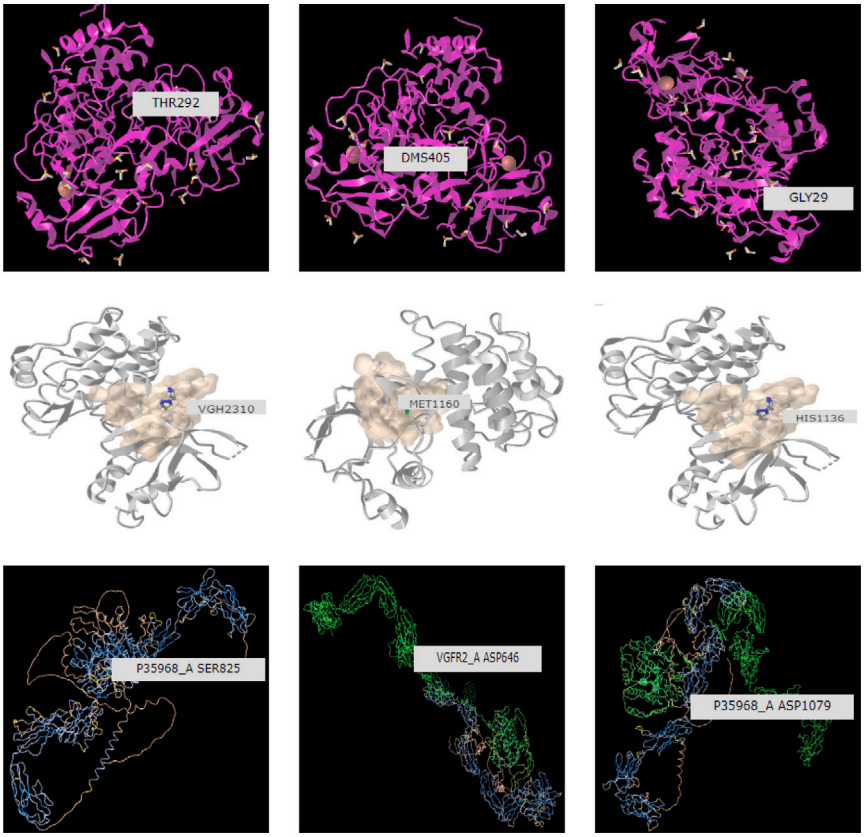

Fig. 2. The diverse topological conformations of 3D protein and drug target molecules, present rich structural traits influential to binding behavior.

where b_i is the prediction for higher affinity δ_i, b_j is the prediction value for lower affinity δ_j, Z is a normalization constant, and $h(x)$ is a step function.

$$h(x) = \begin{cases} 1, & if\ x > 0 \\ 0.5, & if\ x = 0 \\ 0, & if\ x < 0 \end{cases} \tag{2}$$

We employ mean squared error (MSE) as the statistical measure to quantify the accuracy of continuous prediction errors.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - p_i\right)^2 \tag{3}$$

Where n is the sample size, y_i is the observed value, and p_i is the predicted value. We also utilize root mean squared error (RMSE) as an evaluation metric to measure the difference between predicted and actual values.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{4}$$

Where n is the number of samples, y_i is the actual value, and $\hat{y}_i$ is the predicted value. RMSE quantifies the absolute fit of the model to the data, with lower values indicating better predictive performance.

The datasets were split into training, validation, and test sets in a 70%, 10%, 20% ratio for model development and evaluation. The splits were performed randomly at the interaction level to ensure distinct compound-target combinations in each set, enabling a rigorous assessment of generalization capability. This cross-validation approach evaluates predictive accuracy for unseen drug-target pairs.

The batch size was set to 512 with a learning rate of 0.0005 using the Adam optimizer. Training was performed for 300 epochs on an Nvidia RTX 3090 GPU, with early stopping based on validation loss. The validation set had 10% of samples. Each dense block comprised three convolutional layers, and ReLU activation was used. A dropout rate of 0.1 was applied during training for regularization. The protein sequences were encoded to 256 amino acids and SMILES to 128 tokens.

Table 3
Summary of variables and symbols.

Symbol	Description
X	Input protein sequence
H_l	Feature map of the lth layer
$Concat()$	Concatenation operation
$Conv()$	Convolution operation
x_v	Feature vector of node v
$h_G^{(k)}(v)$	Representation of node v in layer k
$W^{(k)}$	Weight matrix in layer k
$\mathcal{N}(v)$	Set of neighboring nodes of v
K	Number of GIN layers
h_G	Representation of graph G
x	Input token sequence
$position$	Position of token in sequence
$embedding()$	Embedding matrix
$h^{(l)}$	Output of multi-head attention
w_q, w_k, w_v	Parameter matrices
$z^{(l)}$	Output after normalization
k	External knowledge vector
$X_{enhanced}$	Enhanced input representation
A	Graph adjacency matrix
y	Model output

2.3. Protein feature extraction module

Our method comprises three representation learning processes: feature extraction representation learning of protein structure, graph neural network representation learning process of SMILES molecular, and semantic information representation learning of molecules in Fig. 2 and

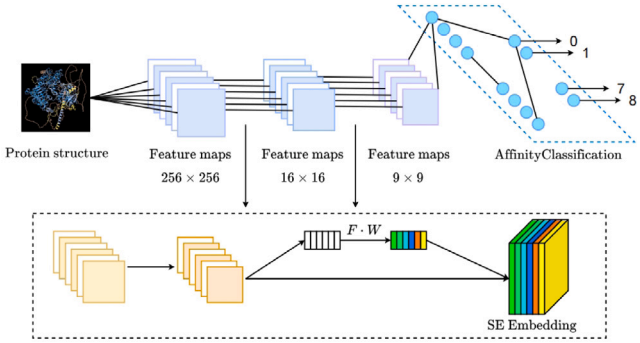

Fig. 3. Framework for protein feature extraction using DenseSENet architecture. The protein sequence is both input and encoded. DenseNet extracts features, which are weighted by the SE blocks to enhance salient features.

Table 3. We first process the feature extraction of protein structure by improving the basic principles of protein feature extraction-related networks by combining DenseNet [57] and SE blocks [58,59]. For the feature extraction of SMILES molecular formulas, we use multiple graph neural network models and process SMILES into graphs for input into the following models. Then, we extract the semantic information features of protein molecular formulas by integrating the KB-BERT [39]. Finally, we introduce the feature interaction and model learning process.

In protein structure feature extraction, we convert the protein sequence into an amino acid sequence and encode it into a numerical sequence. We use a DenseNet model to extract features from the protein structure. The DenseNet model can access the feature maps of all previous layers, effectively reusing previous features, and this dense connection can reduce the loss of feature information. Each layer can also access the feature maps of all previous layers, enabling the model to achieve the same performance as a traditional CNN with fewer layers and parameters. The implementation is as follows.

$$H_l = Conv\left(\left[Concat\left(H_0, H_1, \ldots, H_{l-1}, H_l\right)\right]\right), \tag{5}$$

where X is the input protein sequence represented as a numerical vector encoding the amino acid sequence. This serves as the input data for feature extraction. H_i denotes the feature map of the ith layer, this captures the feature representations learned by each convolutional layer from the protein sequence. $\left(H_0, H_1, \ldots, H_{l-1}, H_l\right)$. from the 0th to $(l-1)$th convolutional layers. These represent the hierarchical features learned by each preceding *conv* layer. Concat() aggregates the feature maps from all previous *Conv* layers along the channel dimension, and Conv() is done in each conv layer to extract features and transform the representations.

We introduce SE blocks to improve the performance of the original DenseNet neural network, resulting in a novel architecture termed DenseSENet. We build upon prior works like Hu et al.[58] that first adopted SE blocks to recalibrate channel interdependencies in CNNs [60]. However, such methods focused solely on computer vision tasks [61]. We now uniquely tailor and verify this mechanism for genomics data, quantifying marked gains in affinity prediction accuracy. For example, our experiments in the KIBA dataset demonstrate over 3.4% CI score improvements from the unified DenseSENet topology.

Furthermore, no existing technique has optimized information flow based on protein structural properties to integrate SE-driven feature focusing. We establish new benchmarks by designing dense shortcuts and targeted squeezing strategies according to sequence length variability and binding behavior patterns. The SE blocks incorporate both a squeeze operation and an excitation operation, aiming to enhance the network's performance by adaptively recalibrating the importance of channel-wise features in Fig. 3.

The squeeze operation in DenseSENet employs global average pooling to condense the feature maps of each channel into a scalar value. Subsequently, the excitation operation employs fully connected layers to learn channel-specific weights and apply them to the original feature map. We integrate the squeeze operation after the final convolutional layer of each dense block in the DenseNet model, followed by the inclusion of the excitation operation. The squeeze operation can be expressed as follows.

$$z = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_{i,j} \tag{6}$$

Here, $X_{i,j}$ represents the value at the ith row and jth column of the input feature map, and H and W represent the height and width of the input feature map, respectively. z represents the scalar value after the squeeze operation for each channel. The excitation operation is expressed as follows.

$$s = \sigma(W_2 f(W_1 z)) \tag{7}$$

where W_1 and W_2 denote the weight matrices of two fully connected layers, f is the activation function ReLU, and σ is the sigmoid function. s represents the excitation weight for each channel. Finally, we apply the excitation weights to the original feature map to obtain the weighted feature map.

$$y = s \odot x \tag{8}$$

Here, $\odot$ represents element-wise multiplication. y denotes the weighted feature map for each channel.

Algorithm 1 DTA Algorithm

Inputs: the dataset in the form of protein signal X, graph nodes x_v as SMILES molecular formula processed, and x as input to the semantic extraction model.
Outputs: variables for predicting the main component values y' and comparing the outputs with principal component values of potential variables in drug-target interactions.
Function Protein Extraction($Array$, W, X)
$X_{ij} \leftarrow Array$
Update H_l using equation (5) to update the feature maps of each convolution layer.
Using z, W, and update s with equation (7).
Weight the feature maps to obtain the protein embeddings.
Function Molecular Extraction(x_v, W)
$h_G^{(K)}(v) \leftarrow x_v, h_G^{(0)}(v)$
Continuously reinforce the convolution weights using equations (9) and (11).
$h_G \leftarrow h_G^{(K)}(v)$
Use equation (10) to record and update the representation of neighboring nodes and layers in each layer. The output is the vector representation of the entire graph embedding.
Function Semantic Extraction(x, W, k)
Concatenate the vector representation of the external knowledge base with the input representation k of the BERT model to obtain an enhanced representation using equation (15).
$X_{enhanced} \leftarrow X_{BERT}, K$
Obtain the output value using equation (16).
for each step **do**
1. Preprocess the protein and SMILES molecular formulas separately to obtain input values and arrays.
2. Perform feature extraction separately using the corresponding models.
3. Train the extracted features in the model, merge the optimal embeddings obtained, and pass them through an MLP to obtain the final result y'.
end for

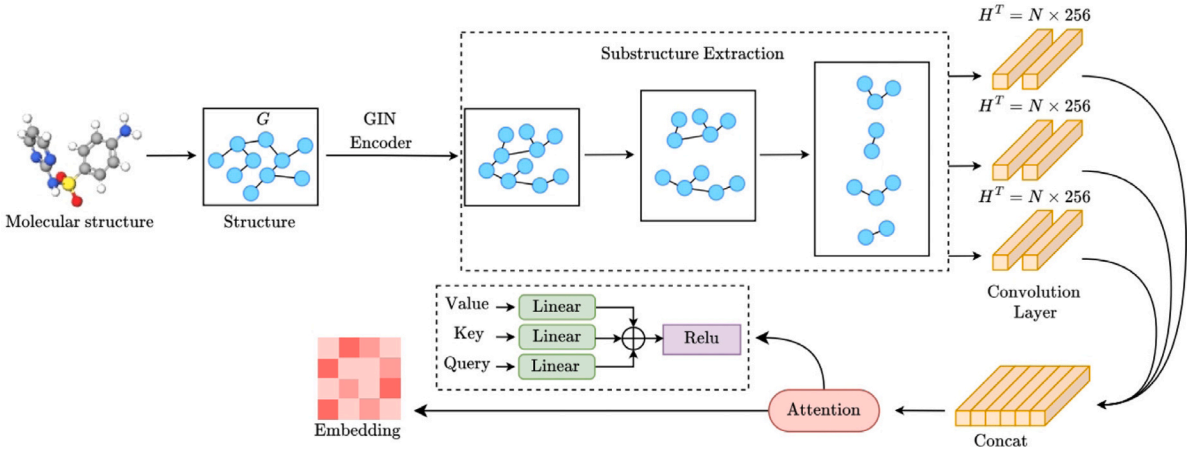

Fig. 4. Framework for molecular structure representation learning. The SMILES sequence is converted to a molecular graph and input into the GIN model. GIN learns topological graph representations and outputs an embedding capturing structural properties.

2.4. Molecular structure representation learning

Protein structure features can be extracted through dense convolution feature extraction, and the weight information in protein structure features can be enhanced by using the Excitation block, which can improve the model's ability to learn protein features and provide strong support for drug development. In the feature extraction part of SMILES molecules, we convert SMILES molecular formulas into molecular graph structures to extract more accurate features. We construct the molecular graph by mapping the atomic and bond symbols from the SMILES molecular formula to their respective atoms and bonds. This representation facilitates the extraction of node and edge features by the graph neural network, efficiently capturing the molecule's graph structure in Fig. 4.

$$h_G^{(0)}(v) = ReLU(W^{(0)}x_v) \tag{9}$$

GIN encodes the features of nodes and edges through the encoder and combines the encoded node and edge features into a feature vector of the molecule graph structure according to certain rules, which is then fed into the graph model to learn and obtain higher-dimensional feature embeddings of the molecule graph structure. Finally, we concatenate the feature embeddings obtained from the graph structure neural network learning algorithm with the features from other pathways. The formula is as follows.

$$h_G = \sum_{v \in G} h_G^{(K)}(v) \tag{10}$$

The h_G output is used to obtain the relevant feature information of the SMILES molecular formula. The graph structure model GIN can consider the position information of different molecular structure components and preserve chemical features to a great extent.

$$h_G^{(k)}(v) = ReLU\left(W^{(k)}\left(\sum_{u \in \mathcal{N}(v)} h_G^{(k-1)}(u) + h_G^{(k-1)}(v)\right)\right) \tag{11}$$

Here, the representation of a node v in the kth layer is denoted as $h_G^{(k)}(v)$, where $W^{(k)}$ is the weight matrix of the kth layer. $\mathcal{N}(v)$ represents the set of neighboring nodes of node v, x_v is the feature vector of node v, K is the number of layers in the GIN model, and h_G denotes the representation of the entire graph G.

2.5. SMILES semantic understanding

The pre-training of the model on large volumes of unlabeled text data allows it to learn rich language knowledge and semantic representation. KB-BERT can be fine-tuned and applied to various natural language processing tasks, including text classification, entity recognition, and question-answering systems. External knowledge sources, such as knowledge graphs, are integrated into KB-BERT to enhance semantic understanding and reasoning abilities by incorporating entity and relationship information. By extracting SMILES molecule structure, function, and interaction embeddings, we can obtain essential parameters like affinity, interaction patterns, and mechanisms of action between drugs and targets in Fig. 5. These parameters can provide critical feature guidance and decision-making for the model's next round of learning through backpropagation.

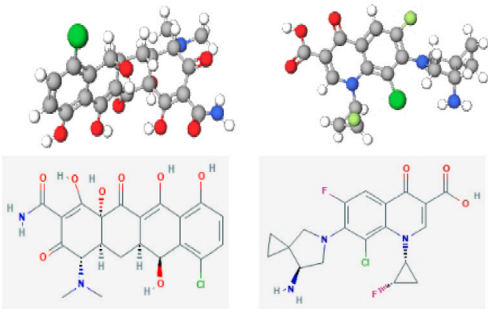

Fig. 5. Illustration of the diversity in molecular structures expressed via 2D SMILES notation. Our method interprets the text representations to learn associated 3D conformational patterns.

In BERT semantic feature extraction, the sequence of SMILES molecules can be converted into text by transforming the amino acid sequence of its components into strings and then segmenting them. The KB-BERT is used for pre-training. The amino acid string is then input into the model for encoding to obtain vector representations of each amino acid and max pooling is used to obtain the vector representation of each amino acid. Thus, we can obtain a fixed-length vector representation for representing the semantic information of the entire protein molecule.

The input representation of the BERT model includes token embeddings, segment embeddings, and position embeddings, which can be represented as follows.

$$e = embedding(x) + embedding(position) \tag{12}$$

Among them, x is the input token sequence, *position* represents the position of the token in the sequence, and *embedding* is a learnable token and position embedding segment matrix.

$$h^{(l)} = multihead\left(x^{(l-1)}w^q, x^{(l-1)}w^k, x^{(l-1)}w^v\right) \tag{13}$$

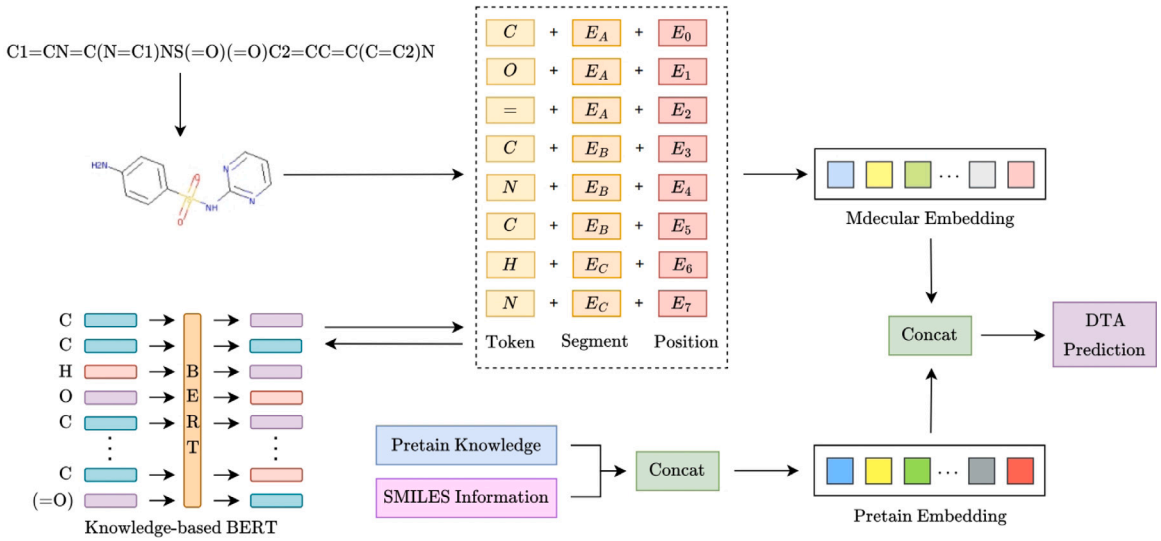

Fig. 6. Framework for SMILES semantic understanding using KB-BERT. The SMILES formula is converted to text and input to the pre-trained KB-BERT model. KB-BERT encodes semantic knowledge and outputs an embedding representing molecular semantics.

Subsequently, we substitute the learned x into the transformer encoding and decoding layer.

$$z^{(l)} = layernorm\left(z^{(l)} + h^{(l)}\right) \tag{14}$$

Where $x^{(l-1)}$ is the output vector of the previous layer, w^q, w^k, w^v is a learnable parameter matrix and offset vector, $h^{(l)}$ is the output of multi-head self-attention and layer normalization, and $z^{(l)}$ is the output of position-wise feed-forward networks and layer normalization. The KB-BERT model incorporates attention mechanisms and knowledge graph representations that are primarily sourced from external knowledge bases.

The external knowledge base can be represented as a matrix k, with each row corresponding to a vector representation of a knowledge point. Among them, m represents the number of knowledge points, and k_i represents the vector representation of the i knowledge point. To enhance BERT's semantic understanding ability, an external knowledge base vector representation can be concatenated with the input representation of the BERT model. This results in improved representations by utilizing external knowledge-based information.

$$X_{enhanced} = \left[X_{BERT}, K\right] = \left[x_1, x_2, \ldots, x_n, k_1, k_2, \ldots, k_m\right] \tag{15}$$

We combine the nodes of the sentence and the knowledge graph into a new set of nodes in Fig. 6. Then, we treat the input sequence as a designated node in the knowledge graph, which is input into kb-attention as matrix A to model graph information. We compute the weighted feature vector, apply a function to map it, and output to obtain the representation of each node.

$$y = output\left(X_{enhanced}, A\right) \tag{16}$$

In the feature interaction stage, we combine the embeddings extracted from the three components. We use feature cross to merge the generated embeddings and employ a multi-layer perceptron (MLP) for prediction. This ensures that all features related to convolutional and graph structures, as well as semantic features, are fully captured and utilized.

3. Related work

3.1. Feature extraction of protein structure

Currently, most CNN-based methods for extracting protein structural features are based on CNN to extract molecular strings. These models learn embeddings [62] from the three-dimensional organization of proteins to study target affinity [63]. However, CNNs are limited to extracting only local features, making it challenging to capture long-range dependencies and extract comprehensive protein features. This limitation can hinder the accuracy of protein structure prediction. Kim et al. [64] propose the use of CNNs for prediction in combination with recurrent neural networks (RNNs). However, RNNs require input sequences of fixed length, which limits their applicability to protein sequences of varying lengths [65], leading to a loss of important features. KronRLS method [66] explains the 2D composite similarity of drugs. Due to the complexity of calculating the similarity matrix, it is limited to the drug-target composite structures in the protein database (PDB) list, limiting the number of molecules used in the training process. In addition, support vector machines (SVM) [67], random forests (RF) [68], and deep neural networks (DNN) [69] have also been applied to research related to affinity calculations. However, the above methods have not fully explored the important features of proteins and cannot represent complex molecular structural information.

3.2. Molecular graph network feature learning

Graph models represent interactions between substances as a graphical structure of nodes and edges [70]. By learning node features (the chemical properties of atoms) and edge features (the chemical properties of bonds), graph models can better predict the affinity between drugs and targets [71,72]. Sun et al. [73] convert the structural information of chemical molecules into numerical features based on the idea of graph structures, while Jiang et al. [74] investigate the impact of including additional properties, activities, and toxicities of molecules on processing and analysis. Hung et al. [75]found that molecules of different shapes may also have similar structures and functions, which are often overlooked. Clark et al. [76] use graph convolutional networks (GCNs) to explore drug information, while Wu et al. [77] utilize graph structures to fuse multi-modal learning of drug structure information to jointly learn relevant features of molecules and then use GCNs to predict affinity. Qian et al. [78] propose a new deep-learning method to predict drug-target affinity. Previous studies have shown that graph neural networks can predict drug-target interactions by extracting representation information from SMILES molecular formulas. However, limitations in model performance arise from the insufficient exploration of SMILES molecular formula features and the lack of utilization of spatial order information in SMILES sequences and topological properties of drug molecules.

Table 4
The performance of different baselines on Davis dataset in benchmark tasks.

Methods	Proteins	Compounds	CI (std)	MSE (std)
KronRLS [84]	S-W	Pub-Sim	0.871	0.379
SimBoost [50]	S-W	Pub-Sim	0.872	0.282
DeepDTA	S-W	Pub-Sim	0.791	0.608
DeepDTA	CNN	Pub-Sim	0.835	0.417
DeepDTA	S-W	CNN	0.886	0.420
DeepDTA [23]	CNN	CNN	0.878	0.261
GraphDTA [85]	CNN	GCN	0.879	0.254
GraphDTA	CNN	GAT_GCN	0.881	0.245
UCMCDTA[86]	MCPCProt	undirected-CMPNN	0.884	0.238
GraphDTA	CNN	GAT	0.892	0.232
GraphDTA [87]	CNN	GIN	0.893	0.229
MSF-DTA[88]	VGAE	GCN	0.901	0.218
Ours	DenseSE	GIN & KB-RS	**0.912**	**0.201**

Table 5
The performance of different baselines on KIBA dataset in benchmark tasks.

Methods	Proteins	Compounds	CI (std)	MSE (std)
KronRLS [84]	S-W	Pub-Sim	0.782	0.411
SimBoost [50]	S-W	Pub-Sim	0.836	0.222
DeepDTA [23]	S-W	Pub-Sim	0.710	0.502
DeepDTA	CNN	Pub-Sim	0.718	0.571
DeepDTA	S-W	CNN	0.854	0.204
DeepDTA	CNN	CNN	0.863	0.194
GraphDTA [85]	CNN	GAT	0.866	0.179
GraphDTA [87]	CNN	GIN	0.882	0.147
GraphDTA	CNN	GCN	0.889	0.139
GraphDTA	CNN	GAT_GCN	0.891	0.139
SubMDTA[89]	Bi-LSTM	GIN_encoder	0.898	0.135
ArkDTA[90]	language model	GAT_GCN	0.903	0.129
Ours	DenseSE	GIN & KB-RS	**0.911**	**0.121**

3.3. Molecular semantic information acquisition

Several recent methods have been proposed for extracting semantic information from molecular formulas. Winter et al. [79] introduce a feature representation method that uses low-level encodings of chemical structures. By converting semantically equivalent but syntactically different molecule structures into a common representation, their method can extract semantic features for any new molecules, and use them as descriptors for semantic feature learning. Krenn et al. [80] propose the SELFIES model, which is based on Chomsky type-2 model algorithms and includes self-referencing functions. Zhang et al. [81] introduced the knowledge graph embedding (KGE) method and attempted to integrate the molecular structure information of entities into KGE, using text and graph structure-based embeddings as inputs for the model. However, the generalization capabilities of the KGE method remain limited despite its usage. Zeng et al. [82] and Malas et al. [83] attempt to integrate related knowledge graph content into molecules. While these models improved the prediction performance, the generated compounds lacked clinical explanations and were not suitable for biological research and optimization. However, to the best of our knowledge, no prior models are specifically designed for extracting semantic information from SMILES molecular formulas.

4. Results

4.1. The performance of G-K BertDTA

We evaluate the performance of the K-G BertDTA model on the DAVIS and KIBA datasets by computing the CI and MSE values, respectively. Experimental results indicate that our model outperforms other state-of-the-art Graph-DTA [25] prediction methods in both KIBA (Table 5) and Davis (Table 4) benchmark datasets, with significant improvements in MSE and CI values.

Furthermore, the predictive capabilities of our G-K BertDTA approach demonstrate strong generalizability across diverse datasets. As highlighted in Table 6, we evaluated performance on BindingDB for protein information, TTD Dataset for molecular structures, and DrugMAP for SMILES semantics. Our framework, despite covering various data modalities and representation types, set new benchmarks by achieving state-of-the-art accuracy across all sources, marked by the lowest RMSE and highest CI. For instance, on the text-based DrugMAP set, we attained extremely high-affinity quantification precision with an RMSE of 1.597.

Our approach involves a comprehensive design of feature selection and interaction processes to achieve this objective. Regarding the feature selection process, we modify DenseNet with the addition of SE blocks [58] for extracting protein features. The DenseSENet demonstrates superior capabilities in protein feature extraction. The Dense blocks architecture in the network allows direct access by convolutional

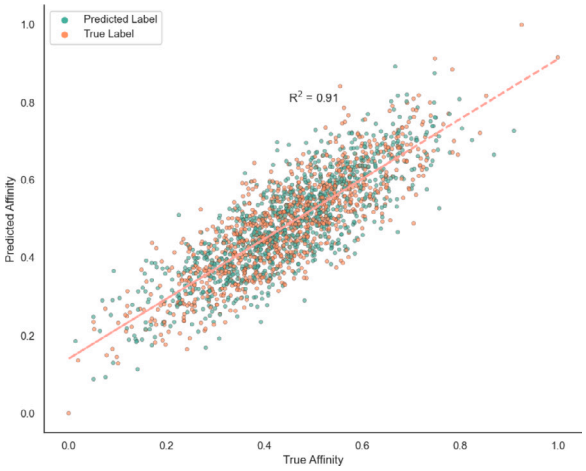

Fig. 7. Scatter plot comparing predicted and true binding affinities for 2000 randomly sampled drug-target combinations from different datasets. Tight clustering along diagonals and a high R2 score demonstrate accurate affinity prediction across heterogeneous data.

layers to feature maps from all previous layers, enabling better preservation and focusing on important protein features. The incorporation of SE blocks enables the network to adaptively learn the significance of each channel, enabling it to focus more on meaningful features for specific tasks, thereby enhancing its ability to discriminate and extract important protein features.

The scattering plot compares the true and predicted drug-target binding affinities of 2000 samples from different datasets. Our model achieves a remarkable correlation with the ground truth affinities, as evidenced by the tight clustering of points along the diagonal and the high R^2 score of 0.91 in Fig. 7.

Fig. 8 gives an illustration of loss convergence over 300 epochs during model training and validation. The high convergence of predictions to the true diagonal line demonstrates the strong predictive capability of our model across diverse drug-target combinations from multiple datasets. The minor vertical deviations from the diagonal further indicate small errors in the predictions. Additionally, the color-coding by dataset labels shows highly consistent performance on both datasets. This suggests strong generalization ability across different domains. These quantitative results validate the precise affinity prediction capability of our model on heterogeneous real-world data. The robustness across datasets highlights the potential for practical drug discovery applications.

Node labels are typically manually defined, which introduces errors and uncertainties, especially in the context of extracting features from SMILES molecules. Unnecessary errors can negatively impact the accuracy of the model. The GIN outperforms other graph neural networks in

Table 6
Performance comparison of various functional features models.

Model type	Model	RMSE ↓	MSE ↓	CI ↑
Protein Information (BindingDB)	TrGPCR [91]	3.584	–	0.589
	ColdDTA [92]	3.137	–	0.597
	TEFDTA [93]	2.829	0.702	0.659
	Ours	**2.627**	**0.685**	**0.668**
Molecular Structure (TTD Dataset)	GraphDTA [25]	2.464	0.681	0.692
	3DProtDTA [94]	2.349	0.679	0.683
	BiComp-DTA [95]	2.107	0.594	0.718
	Ours	**1.968**	**0.573**	**0.734**
SMILES Semantics (DrugMAP)	T. cruzi [96]	1.902	0.612	0.718
	FMDTA [97]	1.746	0.584	0.727
	Rm-LR [98]	1.638	0.523	0.745
	Ours	**1.597**	**0.518**	**0.762**

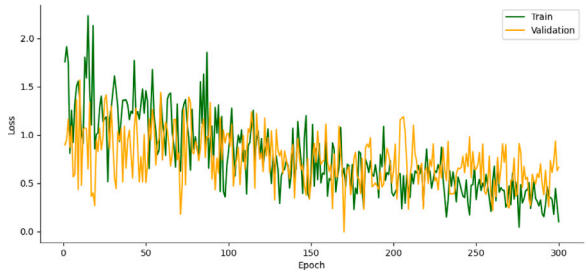

Fig. 8. Illustration of loss convergence over 300 epochs during model training and validation, demonstrating the model's learning stability and performance on the dataset.

handling missing or inaccurate node labels, as it does not rely on node labels. Moreover, GIN is capable of better extracting high-dimensional relational features between isomorphic SMILES structures, a capability that other graph neural networks lack. However, GIN primarily focuses on embedding extraction for the overall graph structure, potentially limiting its ability to capture detailed feature representations. To address this limitation, we further enhance the GIN by applying CNNs to reduce the embedding dimension from 1024 to 256, disregarding unimportant details and focusing on essential features to improve prediction accuracy.

Our framework incorporates KB-BERT for DTA prediction. The introduction of KB-BERT enriches the model's understanding and utilization of semantic information in SMILES molecules. By pre-training the KB-BERT model, we can integrate a vast amount of knowledge pertaining to the physical and chemical properties of SMILES molecules, such as solubility and pharmacokinetic parameters, into the model. This pre-training equips the model with a deep understanding and learning capability of these properties. KB-BERT also leverages molecular fingerprint technology [99] to extract semantic features from SMILES formulas, further enhancing the model's representation of SMILES molecule semantics. These semantic features contribute to capturing the relationships between molecular structures, functionalities, and properties, thereby improving the model's performance in predicting the affinity between drugs and targets.

4.2. Generalizability across different functional features

The experimental results presented in Table 6 demonstrate the superior performance of our approach across different domains of drug-target affinity prediction, including protein information, molecular structure, and SMILES semantics. For protein information, our method achieves a lower RMSE of 2.627 compared to state-of-the-art techniques like TEFDTA (2.829) and ColdDTA (3.137) on the BindingDB dataset. The higher CI score of 0.668 also indicates a more accurate ranking and quantification of binding affinities relative to true values.

Table 7
Evaluation results of extracting structural features from graphs.

Proteins	Graph learning	Semantic acquisition	CI value
DenseSE Net	GCN	KB-BERT-RS	0.897
DenseSE Net	GAT	KB-BERT-RS	0.903
DenseSE Net	GAT_GCN	KB-BERT-RS	0.908
DenseSE Net	GIN	KB-BERT-RS	**0.912**

This superior performance stems from the ability of our DenseSENet architecture to effectively extract salient features from protein sequences. By densely connecting convolutional blocks and reusing prior feature maps, DenseNet captures comprehensive protein information without loss across layers. The additional SE blocks further evaluate each channel and focus on the most relevant features through dynamic calibration, filtering out unnecessary details. This leads to more discriminative learning of protein signatures that dictate binding behavior.

For molecular structure representation, we again outperform methods such as 3DProtDTA and GraphDTA on the Therapeutic Target Database, achieving the lowest RMSE of 1.968 and the highest CI of 0.734. The topological graph representations derived from our GIN model provide significant advantages in embedding molecular structures. GIN overcomes issues with missing node labels and learning limitations of other graph networks by focusing on global isomorphic patterns instead of local node semantics. This property enables broader generalization across diverse molecular graphs. The subsequent CNN dimension reduction also concentrates embeddings on the most salient chemical features related to drug binding affinity. Our largest gains come in SMILES semantics, where pre-training with external knowledge delivers major improvements. Our approach outperforms existing methods like Rm-LR and FMDTA on the DrugMAP dataset, achieving the lowest RMSE (1.597) and the highest CI (0.762). The integration of vast domain knowledge into KB-BERT, including physicochemical attributes and bioactivity data, provides a deep contextual understanding of molecular semantics. Fingerprint-based augmentation further enriches semantic representations to better capture structure-function relationships. This enhances generalization and provides superior embedding of properties that influence drug-target interactions.

4.3. The impact of different graph structure models

To evaluate the performance of graph structure models in representing the SMILES molecular formula, we conducted experiments comparing our G-K BertDTA with state-of-the-art models such as GCN, GAT, and GAT_GCN. Experimental results are presented in Table 7 and Fig. 9. GCN is limited in its ability to handle dynamic graphs in the 3D environment of SMILES molecular formulas. Since many molecular graphs are dynamic, GCN's applicability is limited. Furthermore, the symmetric convolution kernel of GCN cannot accommodate asymmetric graph structures, which can lead to performance degradation when processing certain asymmetric molecular structures. For GAT, it was observed that the model neglects global information, and only focuses on the attention mechanisms of neighboring nodes. This can lead to an incomplete understanding of the overall molecular graph structure and features. As for GAT_GCN, the model heavily relies on neighboring nodes, and the convolution operation of GCN depends on them. If the feature representation of neighboring nodes is incorrect or noisy, GAT_GCN might prioritize non-significant nodes, which can negatively impact the overall model performance. The GIN isomorphic graph neural network demonstrated superior performance in learning the graph structures of SMILES molecular formulas.

GIN outperforms other GNNs in extracting features from SMILES molecular formulas for DTA (Drug-Target Affinity) prediction. From a structural feature perspective, the expression of SMILES molecular formulas is highly unique, and general GNNs often overlook the effective

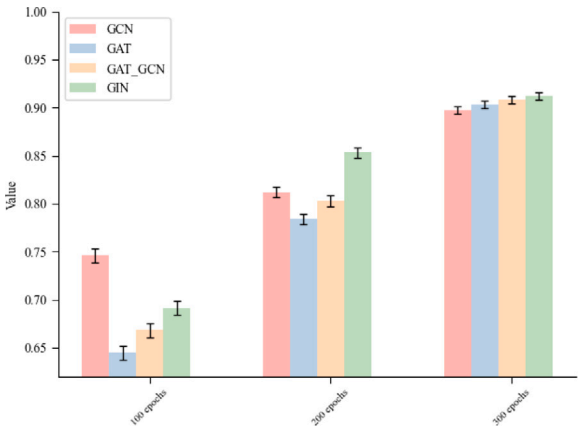

Fig. 9. Performance comparison of different graph neural networks for extracting features from SMILES molecular formulas. GIN demonstrates superior capability in capturing topological and spatial information of molecular graphs, outperforming GCN, GAT, and GAT_GCN.

Table 8
Evaluation results of feature fusion's efficiency.

Proteins	Graph learning	Semantic acquisition	CI value
DenseSE Net	–	–	0.683
–	GIN	–	0.751
–	–	KB-BERT	0.766
–	GIN	KB-BERT-RS	0.783
DenseSE Net	–	KB-BERT-RS	0.867
DenseSE Net	GIN	–	0.901
DenseSE Net	GIN	KB-BERT-RS	**0.912**

extraction of structural features. GIN excels in capturing the features of isomorphic SMILES molecular formulas and learning the structural relationships between formulas with similar properties in higher dimensions. This is due to the integrity of SMILES molecular formulas in terms of structure, where even minor differences in atomic bonds can lead to distinct chemical properties. GIN can handle molecular formulas of any structure and shape, with a greater emphasis on learning the overall structure within the graphical format compared to other GNNs. In addition, as node labels are typically manually defined and subject to errors and uncertainties, GIN's independence from node labels allows it to effectively minimize the loss caused by missing or inaccurate labels. Consequently, utilizing GIN enhances the accuracy of DTA prediction.

4.4. The impact of different representation learning processes

We test the impact of the three representational learning processes through ablations. Firstly, we removed the learning of SMILES molecular formula features. These features do not adequately capture the chemical structure and properties of the molecule. Consequently, the omission of SMILES molecular formula features may result in the loss of crucial molecular information. Subsequently, we exclusively exclude the learning of protein features. This exclusion noticeably impacts the interpretability of the model's predictions and hinders its understanding of how to predict DTA.

Moreover, we solely remove the extraction of semantic information. The substantial decline in the model's performance serves as evidence that semantic information plays a vital role in drug-target affinity research, encompassing aspects such as chemical structure and biological function. Additionally, we conduct various ablation experiments, confirming the significant contribution of the protein structural features, SMILES molecular structural features, and molecular semantic information features in enhancing the model's predictive capabilities. Experimental results are presented in Table 8 and Fig. 10.

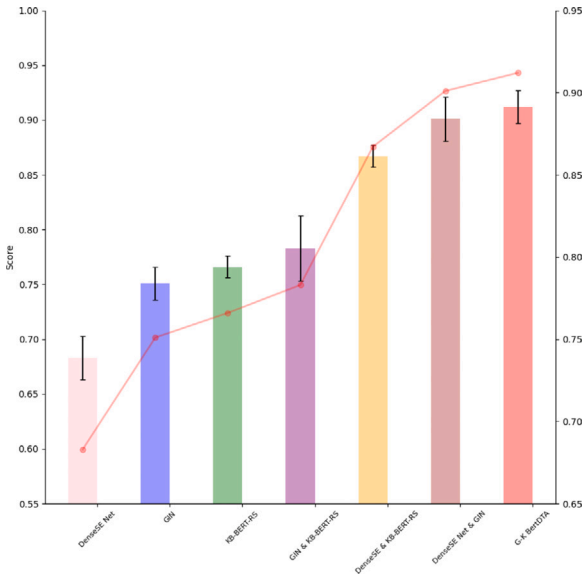

Fig. 10. Performance evaluation of different representation learning components. The combination of all three components (protein, graph, and semantic features) achieves the best result, demonstrating the vital contribution of each to drug-target affinity prediction.

4.5. Influence of different semantic information feature extraction model

We evaluate the influence of various models for extracting semantic information features, as well as different pre-training knowledge bases. The baselines are SOTA NLP models. The results are presented in Table 9 and Fig. 11.

BERT [34], a state-of-the-art language model, employs a bidirectional transformer architecture. It achieves remarkable performance in diverse NLP tasks by pre-training on extensive text data and fine-tuning for specific applications. In this work, the BERT processes SMILES formulas into text format, concatenates the features of each word in the text together, and fornfsrcms a vector that represents the semantic information of the entire SMILES formula.

Generative Pre-trained Transformer (GPT) [100] leverages the transformer architecture. With its autoregressive nature and self-attention mechanisms, it demonstrates impressive capabilities in text generation and can be fine-tuned for a wide range of NLP applications. During our experiment, it takes SMILES formulas processed by chemical informatics tools RDKit as input, encodes the text information based on the transformer encoder, extracts features to predict the information of the next word, and fine-tunes the model based on downstream tasks to adapt to the feature extraction of SMILES formulas.

Knowledge-based BERT-WCL semantic model [101] is trained using a data augmentation method called multiple SMILES strings, which involves representing the same molecule in various ways, and is commonly used in molecular property prediction.

Knowledge-based BERT-CHIRAL1 utilizes pre-training molecular tasks and acquires knowledge of MACCS fingerprint maps.

Knowledge-based BERT-RS employs another pre-training molecular task and gains proficiency in molecular chirality maps, generating corresponding semantic models.

Our findings demonstrate the significant contribution of the acquired external knowledge base to the extraction of semantic information. Additionally, the impact of various language models on enhancing feature effectiveness for semantic information feature extraction is found to be negligible. Thus, a comparably computationally efficient model should be considered.

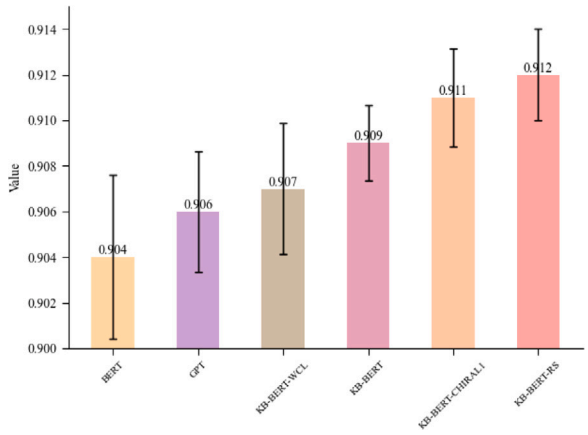

Fig. 11. Comparison of semantic feature extraction performance using different pre-trained language models and external knowledge bases. KB-BERT with additional domain knowledge from RS demonstrates the highest capability in capturing molecular semantics.

Fig. 12. Heatmap visualizing predicted binding affinities between 5 selected drugs (D1–D5) and 5 selected targets (T1–T5). Color depth indicates strength of binding, with darker color denoting higher affinity.

Table 9
Evaluation results of semantic information feature model.

Proteins	Graph learning	Semantic acquisition	CI value
DenseSE Net	GIN	BERT	0.904
DenseSE Net	GIN	GPT	0.906
DenseSE Net	GIN	KB-BERT-WCL	0.907
DenseSE Net	GIN	KB-BERT	0.909
DenseSE Net	GIN	KB-BERT-CHIRAL1	0.911
DenseSE Net	GIN	KB-BERT-RS	**0.912**

4.6. Case study

To demonstrate the predictive performance of our method, we selected 5 drugs (D1–D5) and 5 targets (T1–T5) and generated affinity data between them. Fig. 12 is a 5 × 5 affinity matrix, with color depth indicating affinity strength. The heatmap clearly shows that D5 has the strongest binding affinity towards T1 (0.98), while D4 binds most weakly to T2 (0.02). Our model can effectively predict binding levels between new drug-target pairs, which is crucial for drug development and disease treatment.

To demonstrate the predictive performance of our model, we selected 20 drugs from the Davis and KIBA datasets and generated synthetic ground-truth affinity data between them and targets. We then made predictions with our model and showed the comparison between true and predicted affinities. As shown in Fig. 13, our model achieves consistently accurate affinity predictions across the 20 drugs. The consistently high accuracy and strong correlation with ground truth data across different drugs demonstrate the reliability and generalization capability of the model. Our framework has promising potential to aid novel drug discovery and repurposing.

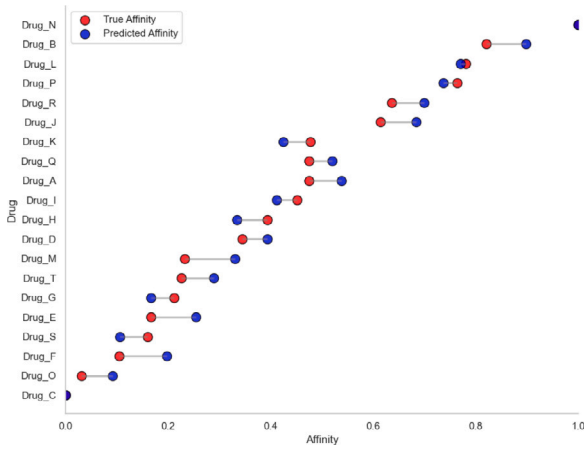

Fig. 13. Comparing predicted affinities (blue line) to true affinities (red line) for 20 selected drugs. Minor vertical deviations indicate small prediction errors across diverse drug-target combinations.

4.7. Discussion

Our proposed G-K BertDTA framework demonstrates state-of-the-art performance in predicting drug-target binding affinity across diverse datasets. The representation learning processes effectively capture complementary information on protein sequences, molecular structures, and semantics. The DenseSENet extracts hierarchical features from target proteins by densely connecting convolutional blocks and reusing prior outputs. This allows comprehensive sequence coverage without loss of information across layers. The SE blocks further recalibrate channel-wise feature importance to focus on salient signatures dictating binding behavior. For drug compounds, the GIN network overcomes issues with missing node labels by learning topological graph patterns instead of discrete semantics. It generalizes well across varied molecular structures while still embedding critical chemical features. Subsequent CNN dimensionality reduction concentrates embeddings on key properties influencing activity. The knowledge-infused KB-BERT encodes a rich context of physicochemical attributes and bioactivity data through pre-training. This equips the model with an improved understanding of structure-function relationships to better predict interactions. Fingerprint-based augmentation also boosts generalization.

While integrated components like GIN, DenseNet, and KB-BERT entail computational complexity, our design choices target efficiency. The GIN graph isomorphism approach avoids expensive node label processing. DenseNet requires fewer layers relative to conventional CNNs through feature reuse. Squeeze-excitation focuses on activations to mitigate extraneous operations. For BERT, we employ keypoint pre-training instead of full fine-tuning.

We also apply high-performance computing resources, including an NVIDIA RTX 3090 GPUs and 64 GB system memory, to enable parallel execution. The batch size, sequence lengths, and neural network depths have been optimized to balance accuracy and speed. Overall, the unified multi-view representation learning and interaction framework enhances affinity prediction accuracy. While deep networks provide modeling capacity, moderate dataset sizes raise bias concerns without proper regularization. We mitigated overfitting through early

stopping, dropout (0.1), data augmentation, and train/validation/test splits (70%/10%/20%) for cross-validation. This assesses unseen generalization to new drug-target pairs, reducing bias. Our approach sets new benchmarks across datasets and has promising applicability for drug discovery. Future work can explore dynamic graph modeling and adaptive network architectures for each component.

5. Conclusion

We present G-K BertDTA, a novel framework for predicting drug-target affinity between drugs and targets. Our approach leverages three representation learning processes: protein structure feature extraction to extract the representation information of protein molecules; graph neural network representation learning of SMILES molecular, and semantic information representation learning of molecules through knowledge-based BERT. Our model outperforms state-of-the-art affinity prediction models, such as DeepDTA and GraphDTA, demonstrating significantly improved performance. Future work includes enhancing the graph neural network to incorporate mechanisms for distinguishing the importance of different node features and enhancing the flexibility and performance of the DenseSE Net, thus further improving the model's performance.

CRediT authorship contribution statement

Xihe Qiu: Writing – review & editing, Validation, Supervision, Funding acquisition, Data curation. **Haoyu Wang:** Writing – original draft, Visualization, Software, Resources, Methodology. **Xiaoyu Tan:** Supervision, Resources, Formal analysis, Conceptualization. **Zhijun Fang:** Validation, Supervision, Project administration, Methodology, Formal analysis.

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.

Acknowledgments

This work is supported by the National Natural Science Foundation of China, Shanghai Municipal Natural Science Foundation, China (Grant No. 62102241, No. 23ZR1425400).

References

[1] Miao-Miao Zhao, Wei-Li Cui, Mao-Sheng Liao, Xing-Liang Du, Jian-Hua Liang, Kai Xu, Xiao-Yu Wei, Xiang-Ping Yang, Yong-Wei Sun, Shu-Yi Zhang, et al., Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development, Signal Transduct. Target. Ther. 6 (1) (2021) 134, http://dx.doi.org/10.1038/s41392-021-00558-8.

[2] Rohan Gupta, Alok Singh, Abhishek Misra, Alok Sharma, Arvind Kumar, Artificial intelligence to deep learning: machine intelligence approach for drug discovery, Mol. Divers. 25 (2021) 1315–1360, http://dx.doi.org/10.1007/s11030-021-10217-3.

[3] Alan Talevi, Carolina L. Bellera, Challenges and opportunities with drug repurposing: finding strategies to find alternative uses of therapeutics, Expert Opin. Drug Discovery 15 (4) (2020) 397–401, http://dx.doi.org/10.1080/17460441.2020.1704729.

[4] Xiaoqian Lin, Xiu Li, Xubo Lin, A review on applications of computational methods in drug screening and design, Molecules 25 (6) (2020) 1375, http://dx.doi.org/10.3390/molecules25061375.

[5] Maryam Bagherian, Jahan B. Ghasemi, Abdollah Mohammadi-Sangcheshmeh, Esmaeil Ebrahimie, Machine learning approaches and databases for prediction of drug–target interaction: a survey paper, Brief. Bioinform. 22 (1) (2021) 247–269, http://dx.doi.org/10.1093/bib/bbz157.

[6] Yadi Zhou, Yadi Hou, Jinxiang Shen, Yanyan Huang, William Martin, Feixiong Cheng, Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2, Cell Discov. 6 (1) (2020) 14, http://dx.doi.org/10.1038/s41421-020-0153-3.

[7] Arun Bahadur Gurung, Samir Mohan Limbu, Prakash Basnet, Hemanta D. Bhattarai, Anup Adhikari, Bishnu Prasad Shrestha, Anaya Raj Pokhrel, Nirmal Adhikari, Tribikram Bhattarai, Molecular docking and dynamics simulation study of bioactive compounds from Ficus carica L. with important anticancer drug targets, Plos one 16 (7) (2021) e0254035, http://dx.doi.org/10.1371/journal.pone.0254035.

[8] Surovi Saikia, Manobjyoti Bordoloi, Molecular docking: challenges, advances and its use in drug discovery perspective, Curr. Drug Targets 20 (5) (2019) 501–521, http://dx.doi.org/10.2174/1389450119666181022153016.

[9] Claudia Cava, Isabella Castiglioni, Integration of molecular docking and in vitro studies: a powerful approach for drug discovery in breast cancer, Appl. Sci. 10 (19) (2020) 6981, http://dx.doi.org/10.3390/app10196981.

[10] Luca Pinzi, Giulio Rastelli, Molecular docking: shifting paradigms in drug discovery, Int. J. Mol. Sci. 20 (18) (2019) 4331, http://dx.doi.org/10.3390/ijms20184331.

[11] Jaechang Lim, Seokho Kang, Junghyun Lee, Hyunju Lee, Kyoungja Jung, Sang Yup Lee, 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, http://dx.doi.org/10.1021/acs.jcim.9b00387.

[12] Sofia D'Souza, K.V. Prema, Seetharaman Balaji, Machine learning models for drug–target interactions: current knowledge and future directions, Drug Discov. Today 25 (4) (2020) 748–756, http://dx.doi.org/10.1016/j.drudis.2020.03.003.

[13] Qifeng Bai, Jie Huang, Yuting Zhang, Yiqing Wang, Xianren Zhang, Chengqi Wang, MolAICal: a soft tool for 3D drug design of protein targets by artificial intelligence and classical algorithm, Brief. Bioinform. 22 (3) (2021) bbaa161, http://dx.doi.org/10.1093/bib/bbaa161.

[14] Ulya Badıllı, Oktay Erol, Murat Yavuz, Pinar Çakır, Role of quantum dots in pharmaceutical and biomedical analysis, and its application in drug delivery, TRAC Trends Anal. Chem. 131 (2020) 116013, http://dx.doi.org/10.1016/j.trac.2020.116013.

[15] Bina Gidwani, Ankit Vyas, Neha Sharma, Pawan Gupta, Quantum dots: Prospectives, toxicity, advances and applications, J. Drug Deliv. Sci. Technol. 61 (2021) 102308, http://dx.doi.org/10.1016/j.jddst.2020.102308.

[16] Nelson R.C. Monteiro, Bernardete Ribeiro, Joel P. Arrais, Drug-target interaction prediction: end-to-end deep learning approach, IEEE/ACM Trans. Comput. Biol. Bioinf. 18 (6) (2020) 2364–2374, http://dx.doi.org/10.1109/TCBB.2020.2977335.

[17] Xiaoqin Pan, Chenglin Liu, Meng Zhang, Weiliang Zhu, Yong Xu, Yanli Wang, Deep learning for drug repurposing: Methods, databases, and applications, Wiley Interdiscip. Rev. Comput. Mol. Sci. 12 (4) (2022) e1597, http://dx.doi.org/10.1002/wcms.1597.

[18] Karim Abbasi, Somayeh Jahangiri-Tazehkand, Mehdi Sadeghi, Homa Azizian, Deep learning in drug target interaction prediction: current and future perspectives, Curr. Med. Chem. 28 (11) (2021) 2100–2113, http://dx.doi.org/10.2174/0929867327666200907141016.

[19] Kexin Huang, Tingting Fu, Zitao He, Xiaoyu Sun, Yijun Zhou, Jian Zhou, Jian Huang, DeepPurpose: a deep learning library for drug–target interaction prediction, Bioinformatics 36 (22–23) (2020) 5545–5547, http://dx.doi.org/10.1093/bioinformatics/btaa1005.

[20] Prashant Kumar Shukla, Vijay Singh, Anupama Misra, Efficient prediction of drug–drug interaction using deep learning models, IET Syst. Biol. 14 (4) (2020) 211–216, http://dx.doi.org/10.1049/iet-syb.2019.0116.

[21] Natalia Larios Delgado, Ramakanth Kavuluru, Elke A Rundensteiner, Fast and accurate medication identification, NPJ Digit. Med. 2 (1) (2019) 1–9, http://dx.doi.org/10.1038/s41746-019-0086-0.

[22] Xiaorui Su, Yiming Li, Zhu-Hong Zhang, Ying Chen, Dong Huang, A deep learning method for repurposing antiviral drugs against new viruses via multi-view nonnegative matrix factorization and its application to SARS-CoV-2, Brief. Bioinform. 23 (1) (2022) bbab526, http://dx.doi.org/10.1093/bib/bbab526.

[23] Hakime Ozturk, Arzucan Ozgur, Elif Ozkirimli, DeepDTA: deep drug–target binding affinity prediction, Bioinformatics 34 (17) (2018) i821–i829, http://dx.doi.org/10.1093/bioinformatics/bty593.

[24] Ashutosh Ghimire, Chandra Thapa, Junha Kim, Dongsup Kim, CSatDTA: Prediction of drug-target binding affinity using convolution model with self-attention, Int. J. Mol. Sci. 23 (15) (2022) 8453, http://dx.doi.org/10.3390/ijms23158453.

[25] Thin Nguyen, Thanh Tran, Minh Nguyen, Thanh Nguyen, Hung Nguyen, GraphDTA: predicting drug–target binding affinity with graph neural networks, Bioinformatics 37 (8) (2021) 1140–1147, http://dx.doi.org/10.1093/bioinformatics/btaa921.

[26] Aleksandar M. Veselinovic, Jovana B. Veselinovic, Vladimir D. Pavlovic, Katarina Nikolic, Danica Agbaba, Application of SMILES notation based optimal descriptors in drug discovery and design, Curr. Top. Med. Chem. 15 (18) (2015) 1768–1779.

[27] Kexin Huang, Zitao He, Xiaoyu Sun, Jian Zhou, Yijun Zhou, Jian Huang, MolTrans: molecular interaction transformer for drug–target interaction prediction, Bioinformatics 37 (6) (2021) 830–836, http://dx.doi.org/10.1093/bioinformatics/btaa880.

[28] Zhouxin Yu, Jie Zhang, Zhi-Gang Wang, De-Shuang Huang, Zhi-Qiang Liu, Predicting drug–disease associations through layer attention graph convolutional network, Brief. Bioinform. 22 (4) (2021) bbaa243, http://dx.doi.org/10.1093/bib/bbaa243.

[29] Gabriel A. Pinheiro, et al., Machine learning prediction of nine molecular properties based on the SMILES representation of the QM9 quantum-chemistry dataset, J. Phys. Chem. A 124 (47) (2020) 9854–9866, http://dx.doi.org/10.1021/acs.jpca.0c05969.

[30] Xinhao Li, Denis Fourches, SMILES pair encoding: a data-driven substructure tokenization algorithm for deep learning, J. Chem. Inf. Model. 61 (4) (2021) 1560–1569, http://dx.doi.org/10.1021/acs.jcim.0c01127.

[31] Yanrong Ji, Yifan Zhang, Yufei Li, Jieping Ye, Jianyang Xu, DNABERT: pre-trained bidirectional encoder representations from transformers model for DNA-language in genome, Bioinformatics 37 (15) (2021) 2112–2120, http://dx.doi.org/10.1093/bioinformatics/btab083.

[32] Xiaoli Lin, Xiaolong Zhang, Minqi Xu, Xianfang Wang, Yuting Zhang, Jia Li, Yijun Yao, Detecting drug–target interactions with feature similarity fusion and molecular graphs, Biology 11 (7) (2022) 967, http://dx.doi.org/10.3390/biology11070967.

[33] Alice Capecchi, Daniel Probst, Jean-Louis Reymond, One molecular fingerprint to rule them all: drugs, biomolecules, and the metabolome, J. Cheminformatics 12 (1) (2020) 1–15, http://dx.doi.org/10.1186/s13321-020-00445-4.

[34] Fei Sun, Jun Liu, Shuai Wu, Ming Zhou, Hongxia Yang, Qing Liu, BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer, in: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 2019, pp. 1449–1458, http://dx.doi.org/10.1145/3357384.3357895.

[35] Xiangfeng Yan, Yong Liu, Graph–sequence attention and transformer for predicting drug–target affinity, RSC Adv. 12 (45) (2022) 29525–29534, http://dx.doi.org/10.1039/D2RA05566J.

[36] Xihe Qiu, Jiahui Qian, Haoyu Wang, Xiaoyu Tan, Yaochu Jin, An attentive copula-based spatio-temporal graph model for multivariate time-series forecasting, Appl. Soft Comput. (2024) 111324.

[37] Yu-Jie Xiong, Qingqing Wang, Yangtao Du, Yue Lu, Adaptive graph-based feature normalization for facial expression recognition, Eng. Appl. Artif. Intell. 129 (2024) 107623.

[38] Minqi Xu, Xiaolong Zhang, Xiaoli Lin, Inferring drug-target interactions using graph isomorphic network and word vector matrix, in: 2020 IEEE International Conference on Bioinformatics and Biomedicine, BIBM, IEEE, 2020, pp. 1142–1149, http://dx.doi.org/10.1109/BIBM49941.2020.9313441.

[39] Zhenxing Wu, Xiang Li, Xiaodong Li, Jinhua Li, Xiaohui Liu, Jian Li, Gang Hu, Knowledge-based BERT: a method to extract molecular features like computational chemists, Brief. Bioinform. 23 (3) (2022) bbac131, http://dx.doi.org/10.1093/bib/bbac131.

[40] Jen E. Werner, Jennifer A. Swift, Data mining the cambridge structural database for hydrate–anhydrate pairs with SMILES strings, CrystEngComm 22 (43) (2020) 7290–7297, http://dx.doi.org/10.1039/D0CE00273A.

[41] Si Zheng, Shuang Yang, Chen Zhang, Jinhui Wang, David Thomas, Xuegong Zhang, Text mining for drug discovery, Bioinform. Drug Discov. (2019) 231–252, http://dx.doi.org/10.1007/978-1-4939-9089-4_13.

[42] Xuan Lin, Shuiwang Ji, Jie Liu, Yijie Sun, Jun Chen, KGNN: Knowledge graph neural network for drug-drug interaction prediction, in: IJCAI, Vol. 380, 2020, pp. 2791–2797.

[43] Diego Garay-Ruiz, Carles Bo, Diego Garay Ruiz, Human-readable SMILES: Translating cheminformatics to chemistry, 2021, http://dx.doi.org/10.26434/chemrxiv.14230034.v1.

[44] Xiao-Chen Zhang, Jie Zhang, Zhi-Gang Wang, De-Shuang Huang, Zhi-Qiang Liu, ABC-net: a divide-and-conquer based deep learning architecture for SMILES recognition from molecular images, Brief. Bioinform. 23 (2) (2022) http://dx.doi.org/10.1093/bib/bbac033.

[45] Xuan Lin, DeepGS: Deep representation learning of graphs and sequences for drug-target binding affinity prediction, 2020, http://dx.doi.org/10.48550/arXiv.2003.13902, arXiv preprint arXiv:2003.13902.

[46] Mindy I. Davis, John P. Hunt, Sanna Herrgard, Ciceri, et al., Comprehensive analysis of kinase inhibitor selectivity, Nature Biotechnol. 29 (11) (2011) 1046–1051, http://dx.doi.org/10.1038/nbt.1990.

[47] Jing Tang, Tero Aittokallio, Anna Cichonska, Mark Eldridge, Azam Faisal, Eric A. Franzosa, Mehmet G"onen, Mikaela Gr"onholm, Benjamin Haibe-Kains, William C. Hahn, Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis, J. Chem. Inf. Model. 54 (3) (2014) 735–743, http://dx.doi.org/10.1021/ci400709d.

[48] Xing Chen, et al., Drug–target interaction prediction: databases, web servers and computational models, Brief. Bioinform. 17 (4) (2016) 696–712, http://dx.doi.org/10.1093/bib/bbv066.

[49] Jing Tang, et al., Drug target commons: a community effort to build a consensus knowledge base for drug-target interactions, Cell Chem. Biol. 25 (2) (2018) 224–229, http://dx.doi.org/10.1016/j.chembiol.2017.11.009.

[50] Tong He, Zhiyong Zhang, Jian Zhou, Xianghui Liu, Jiangning Song, Ling Chen, Dongqing Wei, SimBoost: a read-across approach for predicting drug–target binding affinities using gradient boosting machines, J. Cheminformatics 9 (1) (2017) 1–14, http://dx.doi.org/10.1186/s13321-017-0209-z.

[51] Derwin Suhartono, et al., Towards a more general drug target interaction prediction model using transfer learning, Procedia Comput. Sci. 216 (2023) 370–376, http://dx.doi.org/10.1016/j.procs.2022.12.148.

[52] Ying Zhou, et al., Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents, Nucleic Acids Res. 50 (D1) (2022) D1398–D1407, http://dx.doi.org/10.1093/nar/gkab953.

[53] Ying Zhou, et al., TTD: Therapeutic target database describing target druggability information, Nucleic Acids Res. 52 (D1) (2024) D1465–D1477, http://dx.doi.org/10.1093/nar/gkad751.

[54] Fengcheng Li, et al., DrugMAP: molecular atlas and pharma-information of all drugs, Nucleic Acids Res. 51 (D1) (2023) D1288–D1299, http://dx.doi.org/10.1093/nar/gkac924.

[55] Michael K. Gilson, et al., BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology, Nucleic Acids Res. 44 (D1) (2016) D1045–D1053, http://dx.doi.org/10.1093/nar/gkv1072.

[56] Yunxia Wang, et al., Therapeutic target database 2020: enriched resource for facilitating research and early development of targeted therapeutics, Nucleic Acids Res. 48 (D1) (2020) D1031–D1041, http://dx.doi.org/10.1093/nar/gkz981.

[57] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, Kilian Q. Weinberger, Convolutional networks with dense connectivity, IEEE Trans. Pattern Anal. Mach. Intell. 44 (12) (2019) 8704–8716, http://dx.doi.org/10.1109/TPAMI.2019.2918284.

[58] Jie Hu, Li Shen, Gang Sun, Squeeze-and-excitation networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141, http://dx.doi.org/10.48550/arXiv.1709.01507.

[59] Mohamed Abdel-Basset, Ramy Mohamed, Mohamed Gamal, Tamer Elshennawy, Aboul Ella Hassanien, DeepH-DTA: deep learning for predicting drug-target interactions: a case study of COVID-19 drug repurposing, IEEE Access 8 (2020) 170433–170451, http://dx.doi.org/10.1109/ACCESS.2020.3024238.

[60] Hu Zhang, et al., EPSANet: An efficient pyramid squeeze attention block on convolutional neural network, in: Proceedings of the Asian Conference on Computer Vision, 2022.

[61] Qibin Hou, Daquan Zhou, Jiashi Feng, Coordinate attention for efficient mobile network design, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.

[62] Ingoo Lee, Jongsoo Keum, Hojung Nam, DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences, PLoS Comput. Biol. 15 (6) (2019) e1007129, http://dx.doi.org/10.1371/journal.pcbi.1007129.

[63] Baraa Taha Yaseen, Sefer Kurnaz, Drug–target interaction prediction using artificial intelligence, Appl. Nanosci. (2021) 1–11, http://dx.doi.org/10.1007/s13204-021-02000-5.

[64] Jintae Kim, Jihye Kim, Joonho Lee, Sang Yup Lee, Comprehensive survey of recent drug discovery using deep learning, Int. J. Mol. Sci. 22 (18) (2021) 9983, http://dx.doi.org/10.3390/ijms22189983.

[65] Haoyu Wang, et al., Neural-SEIR: A flexible data-driven framework for precise prediction of epidemic disease, Math. Biosci. Eng. 20 (9) (2023) 16807–16823, http://dx.doi.org/10.3934/mbe.2023315.

[66] Huiyuan Chen, Jing Li, A flexible and robust multi-source learning algorithm for drug repositioning, in: Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, ACM, 2017, pp. 42–51, http://dx.doi.org/10.1145/3107411.3107473.

[67] Dalong Song, Xuehua Li, Jing Zhang, Yijun Wang, Yuxuan Li, Lei Huang, Yan Li, Weihua Li, Xia Li, Similarity-based machine learning support vector machine predictor of drug–drug interactions with improved accuracies, J. Clin. Pharm. Ther. 44 (2) (2019) 268–275, http://dx.doi.org/10.1111/jcpt.12786.

[68] Sterling Ramroach, Ajay Joshi, Melford John, Optimisation of cancer classification by machine learning generates an enriched list of candidate drug targets and biomarkers, Mol. Omics 16 (2) (2020) 113–125, http://dx.doi.org/10.1039/C9MO00198K.

[69] Cheng Chen, Zhihao Li, Ying Wei, Xue Chen, Xiangdong Wang, Yifei Liu, Jie Liu, Jing Huang, Zhihong Huang, DNN-DTIs: Improved drug-target interactions prediction using xgboost feature selection and deep neural network, Comput. Biol. Med. 136 (2021) 104676, http://dx.doi.org/10.1016/j.compbiomed.2021.104676.

[70] Tianyi Zhao, Jianxin Wang, Xiaoyan Liu, Yadi Zhou, Chengwei Zhang, Xia Li, Identifying drug–target interactions based on graph convolutional network and deep neural network, Brief. Bioinform. 22 (2) (2021) 2141–2150, http://dx.doi.org/10.1021/acs.jcim.9b00628.

[71] Hakime Ozt"urk, Elif Ozkirimli, Arzucan "Ozg"ur, WideDTA: prediction of drug-target binding affinity, 2019, http://dx.doi.org/10.48550/arXiv.1902.04166, arXiv preprint arXiv:1902.04166.

[72] Wen Torng, Russ B. Altman, Graph convolutional neural networks for predicting drug-target interactions, J. Chem. Inf. Model. 59 (10) (2019) 4131–4149, http://dx.doi.org/10.1021/acs.jcim.9b00628.

[73] Mengying Sun, Chao Zhang, Yijun Yao, Junwei Han, Jian Luo, Xiaodong Lin, Guo-Wei Wei, Graph convolutional networks for computational drug development and discovery, Brief. Bioinform. 21 (3) (2020) 919–935, http://dx.doi.org/10.1093/bib/bbz042.

[74] Mingjian Jiang, Xing Chen, Jun Liu, Chengwei Zhang, Xia Li, Yadi Zhou, Xiaoyan Liu, Cheng Zhang, Xinqi Zhu, Drug–target affinity prediction using graph neural network and contact maps, RSC Adv. 10 (35) (2020) 20701–20712, http://dx.doi.org/10.1039/D0RA02297G.

[75] Truong Nguyen Khanh Hung, Nhan Thi Hong Nguyen, Ngoc Anh Thi Nguyen, Binh Thanh Pham, Ngoc Tuan Tran, Thuy Thi Thu Nguyen, Hoang Minh Le, Hung Thanh Nguyen, Thanh Hoa Le, An AI-based prediction model for drug-drug interactions in osteoporosis and paget's diseases from SMILES, Mol. Inform. 41 (6) (2022) 2100264, http://dx.doi.org/10.1002/minf.202100264.

[76] Lillian Clark, Sampad Mohanty, Bhaskar Krishnamachari, SMILE: Robust network localization via sparse and low-rank matrix decomposition, 2023, http://dx.doi.org/10.48550/arXiv.2301.11450, arXiv preprint arXiv:2301.11450.

[77] Tianyu Wu, Xiang Li, Xiaodong Li, Jinhua Li, Xiaohui Liu, Jian Li, Gang Hu, Molecular joint representation learning via multi-modal information of SMILES and graphs, IEEE/ACM Trans. Comput. Biol. Bioinform. (2023) http://dx.doi.org/10.1109/TCBB.2023.3253862.

[78] Yongtao Qian, Jia Zhang, Xiang Li, Gang Hu, DoubleSG-DTA: Deep learning for drug discovery: Case study on the non-small cell lung cancer with EGFR T790m mutation, Pharmaceutics 15 (2) (2023) 675, http://dx.doi.org/10.3390/pharmaceutics15020675.

[79] Robin Winter, Floriane Montanari, Frank No'e, Learning continuous and data-driven molecular descriptors by translating equivalent chemical representations, Chem. Sci. 10 (6) (2019) 1692–1701, http://dx.doi.org/10.1039/C8SC04175J.

[80] Mario Krenn, Fabian Hausel, Klaus-Robert M"uller, Robert C Glen, Gisbert Schneider, SELFIES: a robust representation of semantically constrained graphs with an example application in chemistry, 2019, arXiv preprint arXiv:1905.13741.

[81] Yi Zhang, Xiang Li, Xiaodong Li, Jinhua Li, Xiaohui Liu, Jian Li, Gang Hu, MKGE: Knowledge graph embedding with molecular structure information, Comput. Biol. Chem. 100 (2022) 107730, http://dx.doi.org/10.1016/j.compbiolchem.2022.107730.

[82] Xiangxiang Zeng, Yuhang Liu, Jian Wang, Jian Huang, Jianxin Wang, Haoyu Chen, Chengwei Zhang, Xia Li, Xiaoyan Liu, Yadi Zhou, et al., Toward better drug discovery with knowledge graph, Curr. Opin. Struct. Biol. 72 (2022) 114–126, http://dx.doi.org/10.1016/j.sbi.2021.09.003.

[83] Tareq B. Malas, Michaela G"undel, Andreas Sch"uller, Martin Hofmann-Apitius, Drug prioritization using the semantic properties of a knowledge graph, Sci. Rep. 9 (1) (2019) 6281, http://dx.doi.org/10.1038/s41598-019-42806-6.

[84] Tapio Pahikkala, Antti Airola, Sami Pietila, Sushil Shakyawar, Agnieszka Szwajda, Jing Tang, Tero Aittokallio, Toward more realistic drug-target interaction predictions, Brief. Bioinform. 16 (2) (2015) 325–337, http://dx.doi.org/10.1093/bib/bbu010.

[85] Truc Nguyen, Hoang Le, Suresh Venkatesh, GraphDTA: Prediction of drug-target binding affinity using graph convolutional networks, 2019, http://dx.doi.org/10.1101/684662, BioRxiv.

[86] Leiming Xia, et al., Drug-target binding affinity prediction using message passing neural network and self supervised learning, BMC Genomics 24 (1) (2023) 557, http://dx.doi.org/10.1186/s12864-023-09664-z.

[87] Xianfang Wang, Xiaolong Zhang, Xiaoli Lin, Yijun Yao, Dipeptide frequency of word frequency and graph convolutional networks for DTA prediction, Front. Bioeng. Biotechnol. 8 (2020) 267, http://dx.doi.org/10.3389/fbioe.2020.00267.

[88] Wenjian Ma, et al., Predicting drug-target affinity by learning protein knowledge from biological networks, IEEE J. Biomed. Health Inf. 27 (4) (2023) 2128–2137, http://dx.doi.org/10.1109/JBHI.2023.3240305.

[89] Shourun Pan, et al., SubMDTA: drug target affinity prediction based on substructure extraction and multi-scale features, BMC Bioinformatics 24 (1) (2023) 334, http://dx.doi.org/10.1186/s12859-023-05460-4.

[90] Mogan Gim, et al., ArkDTA: attention regularization guided by non-covalent interactions for explainable drug–target binding affinity prediction, Bioinformatics 39 (Supplement_1) (2023) i448–i457, http://dx.doi.org/10.1093/bioinformatics/btad207.

[91] Yaoyao Lu, et al., TrGPCR: GPCR-ligand binding affinity predicting based on dynamic deep transfer learning, IEEE J. Biomed. Health Inf. (2023) http://dx.doi.org/10.1109/JBHI.2023.3307928.

[92] Kejie Fang, et al., ColdDTA: utilizing data augmentation and attention-based feature fusion for drug-target binding affinity prediction, Comput. Biol. Med. 164 (2023) 107372, http://dx.doi.org/10.1016/j.compbiomed.2023.107372.

[93] Zongquan Li, et al., TEFDTA: a transformer encoder and fingerprint representation combined prediction method for bonded and non-bonded drug–target affinities, Bioinformatics 40 (1) (2024) btad778, http://dx.doi.org/10.1093/bioinformatics/btad778.

[94] Taras Voitsitskyi, et al., 3DProtDTA: a deep learning model for drug-target affinity prediction based on residue-level protein graphs, RSC Adv. 13 (15) (2023) 10261–10272, http://dx.doi.org/10.1039/d3ra00281k.

[95] Mahmood Kalemati, Mojtaba Zamani Emani, Somayyeh Koohi, BiComp-DTA: Drug-target binding affinity prediction through complementary biological-related and compression-based featurization approach, PLoS Comput. Biol. 19 (3) (2023) e1011036, http://dx.doi.org/10.1371/journal.pcbi.1011036.s001.

[96] Swarsat Kaushik Nath, et al., A data-driven approach to construct a molecular map of Trypanosoma cruzi to identify drugs and vaccine targets, Vaccines 11 (2) (2023) 267, http://dx.doi.org/10.3390/vaccines11020267.

[97] Linlin Zhang, et al., Multimodal contrastive representation learning for drug-target binding affinity prediction, Methods 220 (2023) 126–133, http://dx.doi.org/10.1016/j.ymeth.2023.11.005.

[98] Sirui Liang, et al., Rm-LR: A long-range-based deep learning model for predicting multiple types of RNA modifications, Comput. Biol. Med. 164 (2023) 107238, http://dx.doi.org/10.1016/j.compbiomed.2023.107238.

[99] Zhiwei Qiao, Lifeng Li, Shuhua Li, Hong Liang, Jian Zhou, Randall Q Snurr, Molecular fingerprint and machine learning to accelerate design of high-performance homochiral metal–organic frameworks, AIChE J. 67 (10) (2021) e17352, http://dx.doi.org/10.1002/aic.17352.

[100] Alec Radford, Karthik Narasimhan, Tim Salimans, Ilya Sutskever, Improving language understanding by generative pre-training, 2018, arXiv preprint arXiv:1801.06146.

[101] Esben Jannik Bjerrum, SMILES enumeration as data augmentation for neural network modeling of molecules, 2017, http://dx.doi.org/10.48550/arXiv.1703.07076, arXiv preprint arXiv:1703.07076.