IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 29 June 2024, accepted 8 August 2024, date of publication 13 August 2024, date of current version 26 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3442931

RESEARCH ARTICLE

A Multi-Feature and Dual-Attribute Interaction Aggregation Model for Predicting Drug-Target Interactions

YUANDONG LIU[1], HAOQIN YANG[2], LONGBO ZHANG[1], HONGZHEN CAI[3], MAOZU GUO[4], AND LINLIN XING[1]

[1]Department of Computer Science and Technology, Shandong University of Technology, Zibo, Shandong 255000, China
[2]Department of Mechanical Engineering, Shandong University of Technology, Zibo, Shandong 255000, China
[3]Department of Agricultural Engineering and Food Science, Shandong University of Technology, Zibo, Shandong 255000, China
[4]Department of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

Corresponding author: Linlin Xing (xinglinlin@sdut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62002206 and Grant 62202373.

ABSTRACT Discovering potential drug-target interactions is crucial for advancing pharmacology. In recent years, the development of large-scale DTI datasets has propelled advancements in DTI prediction computational methods. Various deep learning approaches for interaction prediction often rely on sequence data or structural complexity, yet the synergistic integration of diverse bioinformatics and binding site data remains underexploited, constraining prediction precision. Therefore, a novel approach to integrate available data is required to enhance DTI prediction performance. In this paper, we present a novel aggregation prediction model named MDiDTI, designed to facilitate multi-attribute dual interaction learning. The multi-head self-attention interaction network extracts substructure information of drug molecules and pocket information of targets from biomedical data, enabling spatial-level learning of structural attributes. Meanwhile, the dual-weight mapping network aggregates the chemical semantic features of drug-target pairs, facilitating semantic attribute learning at the sequence level. Lastly, the model combines structural and semantic attributes to compute the interaction values for DTI tasks. Performance evaluation metrics were conducted on three mainstream datasets: BioSNAP, BindingDB, and Human. Experimental results indicate that MDiDTI outperforms existing methods and serves as a reliable and highly generalizable tool for DTI prediction.

INDEX TERMS Drug target interaction, self-attention network, dual-weight mapping network, multi-feature, joint attention.

I. INTRODUCTION

The drug development process is time-consuming and expensive. Drug-target interactions (DTIs) are a key part of the drug discovery and development process, describing the interactions between a drug molecule and the target on which it acts. The study of DTIs is crucial for the repurposing of existing drugs as well as for the discovery of novel drugs. With the rapid advancement of proteomics and drug molecule research, the quantity of drug databases measured by wet experiments is also increasing. Although traditional manual experiments and clinical experiments are reliable methods for predicting DTI, this approach is not practical [1]. Therefore, it has become crucial to explore the interaction mechanism of DTIs, assist wet laboratory techniques, and develop more effective computational methods [2].

Developing a method that can predict DTIs with both efficiency and accuracy represents a substantial challenge [3]. Molecular docking simulation (MDS) is one of the early computing methods [4]. Although this method has been initially applied to drug-target interaction prediction, its application is limited due to its complex operation and

The associate editor coordinating the review of this manuscript and approving it for publication was Jinhua Sheng.

high computational cost [5]. The emergence of data-driven methods, has successfully overcome some challenges in the drug design and development process and has been widely studied and applied [6]. For example, Yamanishi et al. [7] proposed to apply bipartite graph technology to DTI prediction. They transformed the task of predicting interactions into a supervised learning problem of bipartite graphs and proposed a statistical learning method that can combine chemical structure information and genomic information of drugs. The results show that this method exhibits certain robustness in four categories of drug target sets. Laarhoven et al. [8] utilized a kernel least squares-based classifier for interaction prediction, and the input to the model included only the drug-target interaction network. Although traditional machine learning methods have achieved good performance, these methods cannot abstract and automatically extract features.

Deep learning algorithms can automatically extract important features and learn complex nonlinear interactions between drug targets [9]. They can be divided into two categories according to input representation in DTI prediction. One class of models uses input representations based on drug and target protein sequences. The advantage of these models is that they can handle variable-length inputs and consider contextual information [10]. However, their limitation lies in their inability to effectively capture the structural information of molecules, resulting in reduced prediction performance [11]. The shortcomings of this type of model prompted researchers to consider using molecular representations that are more consistent with experimental settings and then developed the second type of deep learning models, which are input representations based on molecular graphs [12]. Where atoms and chemical bonds correspond to nodes and edges, respectively. All these graph-based models represent proteins using amino acid sequences, which are unable to capture the three-dimensional structural features, which are key factors in predicting DTIs. Obtaining high-resolution 3D structures of proteins is a challenging task, in addition to the fact that proteins contain a large number of atoms, which requires large-scale sparse matrices to capture the entire structure. To alleviate this problem, an alternative strategy is used in which proteins are represented by a 2D contact (or distance) map that shows protein residue pair interactions in matrix form. However, the contact (or distance) map, which is usually the output of protein structure prediction, is heuristic-based and provides only an approximate abstraction of the true structure of the protein. Considering that the binding of proteins to many molecules occurs in different binding pockets rather than the whole protein, network-based approaches cannot consider both the semantic and structural properties of drug targets; previous models only learn a global representation of the drug target without explicitly considering local interactions between drug target substructures. We believe that combining sequence and structural properties is a more accurate and effective approach. Therefore, in this work, we propose a multi-attribute dual interaction DTI prediction model called MDiDTI. The model incorporates multi-level attributes of semantics and structure. One branch learns the properties at the structural level. Another branch aims at explicitly learning local interaction properties at the semantic level. The contribution of semantic attributes and structural attributes to the model is comprehensively considered to perform DTI prediction more effectively. The proposed framework views predicting drug-target interactions as a binary classification task. The input is the drug-target pair *(d, p)*, and the output $y \in (0, 1)$, where $y = 1$ means that the drug interacts with the target and $y = 0$ means that it does not.

This paper has the following key contributes:

- The fine structural information of protein binding sites and drug molecule substructures are deeply mined and effectively utilized, and effective integration sequence and strustructural are achieved.
- GraphSage networks and scalable convolutional networks are used to jointly extract valuable feature representations of proteins and drugs, effectively capturing the topological and chemical information of molecules.
- For multi-feature information inputs, a fusion of a multi-head attention network and dual weight mapping network is proposed for dual attribute interactions at the structure and sequence levels of proteins and drugs. It achieves a multi-level fusion of protein-drug and significantly improves the effectiveness of drug-target interaction prediction.

II. RELATED WORK

Deep learning based methods for solving DTI prediction problems have achieved remarkable success. The main difference between deep learning methods is their architecture and the representation of the input data. For example, Huang et al. [13] created a large corpus to segment the original sequences and then used Transformer to encode the segmented sequences directly. While these sequence representations contain atoms and continuously learn semantic relationships between atoms, none of the sequence representations cover the spatial structure of the molecule. The loss of spatial structure information may weaken the predictive power of the model as well as the functional relevance of the learned latent space. Zheng et al. [14] proposed an end-to-end deep learning framework to represent proteins with 2D distance maps and follow a visual question-answering paradigm to predict interactions between drug targets. The model combines a dynamic attentional CNN capable of learning fixed-size representations, and a self-attentional sequence model capable of automatically extracting semantic features from linear symbols. Through the learned attention weights, the model can provide interpretability of the interaction contribution. However, the method is based on the conversion of protein sequences into a heuristic protein distance map, which lacks dynamic information and loses structural information. Bai et al. [15] proposed a DrugBAN

prediction model based on a three-layer GCN encoding drug sequences and a CNN network encoding protein sequences. To better learn the local interaction features between proteins and drugs, the model incorporated an attention mechanism to generate a new joint feature, and a bilinear attention network was able to better learn the link between the two subfeatures. Finally, the integration of the CDAN module into the modeling process enhances the generalization ability of the model. However, the characterization of targets and molecules only contains sequence information. The binding process of target and drug is a dynamic fitting process, and relying on a fixed 2D molecular map to make predictions is not sufficient.

As mentioned earlier, small molecules of drugs can be easily and efficiently represented in one dimension, but proteins are much larger molecules with complex interactions for which a one-dimensional representation may be insufficient. While datasets containing the 3D structures of proteins are limited, some deep learning-based work has used them. AtomNet [16] is the first network that applies local convolutional filters to structural target information and uses a binary classifier to predict DTI. Ragoza et al. [17] describe a CNN scoring function that automatically learns important protein-ligand interaction features relevant to binding. In addition, DataDTA [18] proposes a method to predict pockets from the three-dimensional structure. This model considers vertical and horizontal features to achieve the fusion and capture of multi-scale interactive information.

III. MATERIALS AND METHOD

A. THE OVERALL ARCHITECTURE OF MDIDTI

This paper proposes a multi-attribute dual interaction learning aggregation model MDiDTI to predict DTIs, as shown in Fig 1. The model includes four modules: sequence embedding and graph construction module, multi-feature encoding module, dual-attribute interaction module, and decoding prediction module. In the sequence embedding and graph construction module, the SMILES sequence and protein sequence are deeply embedded, and the protein pocket binding site graph and ligand subgraph are constructed, as shown in Fig 1.A. In the multi-feature encoding module, the sequence embeddings are input to the scalable convolutional network layer to obtain the sequence encoding. At the same time, the GraphSage network encodes the drug molecule subgraph and target site graph to obtain the Structural encoding, as shown in Fig 1.B. Next, structure coding and sequence coding interact in the dual-attribute interaction module respectively. The multi-head self-attention interaction network combines the structural coding to calculate the attention weight matrix to reflect the contribution of the binding site in drug-target interaction, and outputs structural attributes; the sequence coding is input into the dual-weight mapping network, and the influence of the interaction of protein residues pair is reflected through the mapping of the drug target sequence coding to the shared matrix, and semantic attributes are output. Finally, the decoding prediction module obtains the structural attribute representation and the semantic attribute representation and combines these two attributes with a learnable hyperparameter α to perform DTI prediction.

B. SEQUENCE EMBEDDING AND GRAPH CONSTRUCTION MODULE

In DTI prediction, the model is tasked with binary classification of whether an interaction occurs between a compound and a target. The input representation of the protein consists of sequences and pockets. Algorithm [19] calculates the boundary coordinates of each binding pocket and obtains the binding site in the coordinates, they are represented as separate graphs [20], Helps reduce model complexity. A one-shot encoding of the atom type, atomicity, and atom implicit price is used to calculate the eigenvector for each atom, resulting in a vector of size 31 for each node. The process of binding sites graph construction is as in Algorithm 1.

Algorithm 1 binding Sites Graph Construction Procedure

Input: PDB file
Output: site graphs

```
P: Pdb file
m: Molecules read from PDB files
am: Neighbourhood matrix of molecules
ami: Submatrix of am
pockets: Find pocket from pdb
d2: Position of the conformational atom
for pocket in pockets do
    x, y, z = Getpocketboundarycoordinates(pocket);
    bindingPartsAtoms = []
    for atom in d2 do
        bindingPartsAtoms ← atom
    end for
     H = get Atom Feature(bindingPartsAtoms)
    g = Creating a graph from an adjacency matrix(ami)
    graph = Create DGL graph (g, H)
    graphs ← graph
end for
```

The protein sequence embeddings and site maps are then input into a multi-feature encoding layer, which maps protein representations into latent feature space. Drug compounds are represented using the SMILES, which is a one-dimensional sequence that describes chemical atoms and chemical bonds in drug molecules. Although many classic deep learning frameworks use the SMILES format to encode drug information, sequences are not natural representations of molecules and some structural information may be lost [21], thus affecting DTI prediction performance. Therefore, the model first converts the input SMILES into the corresponding two-dimensional molecular graph. The drug molecule graph is defined as $G=(V,E)$. Then, in molecules, the number and types of chemical bonds in atomic nuclei are relatively small,

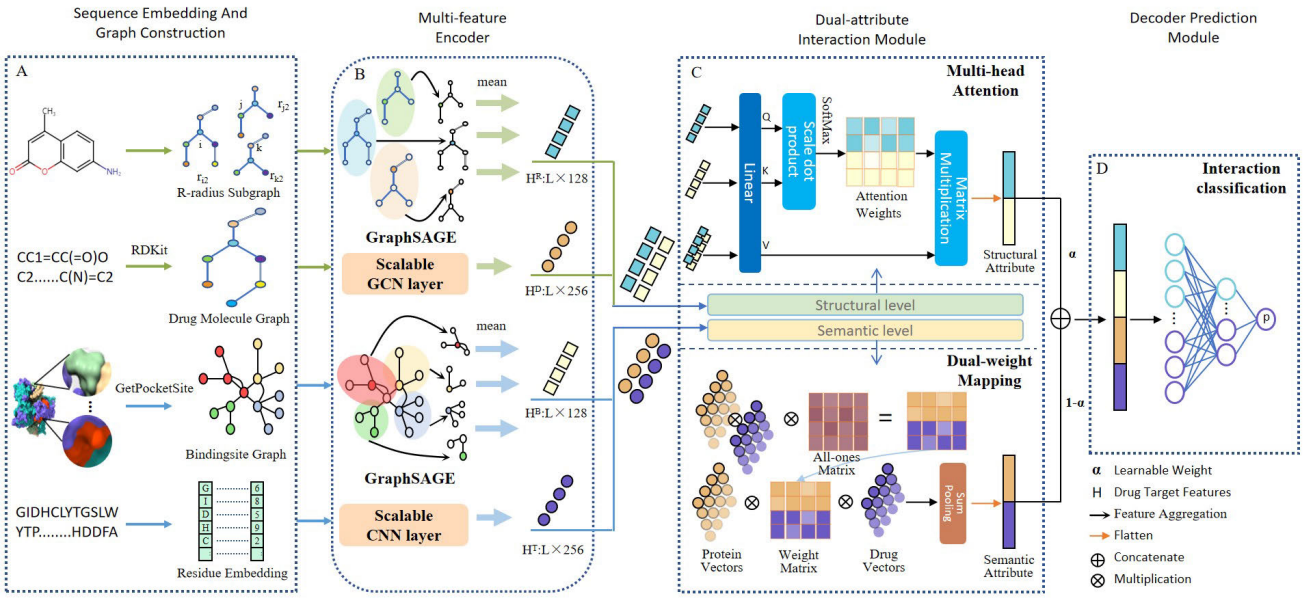

FIGURE 1. The MDiDTI model includes four modules, (A) Sequence Embedding and Graph Construction module, which learns the embedding representation of sequences through the deep representation learning module, and constructs graph representations of drug molecule subgraphs and protein binding sites; (B) Multi-feature Encoder module, constructs scalable convolutional networks for drugs and proteins, and obtains structural feature encoding and semantic feature encoding of drugs and proteins respectively; (C) Dual-attribute dual interaction module based on the attention mechanism, obtains the interaction results respectively structural attribute representation and semantic attribute representation. (D) Decoder Prediction module for DTI prediction.

resulting in correspondingly fewer parameters that can be learned in the model, which in turn leads to insufficient representation learning. To overcome this problem, we adopt the R-radius subgraph algorithm to represent subgraphs of molecular graphs induced by the r-radii of vertices adjacent to vertices and edges, where the r-value is set to start from a vertex hop count. Flowing the previous work setting [22], for a graph *G=(V, E)*, the set of adjacent vertices within the radius r of the *i-th* vertex is denoted as $N_{i,j}$. Here, $N_{i,0} = i$. The r-radius subgraph of vertex v_i is then defined as:

$$v_i^{(r)} = \left(V_i^{(r)}, E_i^{(r)}\right) \tag{1}$$

where:

$$V_i^{(r)} = \left\{v_j \mid j \in N_{(i,r)}\right\}$$
$$E_i^{(r)} = \left\{e_{mn} \in E \mid (m,n) \in N_{(i,r)} \times N_{(i,r-1)}\right\}$$

thereby overcoming the limitation of insufficient learning parameters.

C. MULTI-FEATURE ENCODER

This module uses graph neural networks to learn sequence feature encoding and structural feature encoding of drug targets respectively.

1) GRAPHSAGE FOR STRUCTURAL FEATURES ENCODE OF DRUG AND TARGET

Drug molecules are represented as SMILES strings. MDiDTI extracts atomic and bond information from each SMILES string, which contains connectivity and structural details. The latent information within the sequence exists in non-euclidean space, making it possible to utilize GNN for extracting high-level node representation through information propagation. In graph G, $v_i \in V$ represents the *i-th* atom, and $e_{ij} \in E$ is the chemical bond between the *i-th* atom and the *j-th* atom. Fewer learnable parameters in drug molecular graphs, we employ the r-radius subgraph approach to represent the composite graph.

GraphSage [23], as a spatial-based convolutional GNN, adeptly processing large-scale graph data and facilitates efficient computation. Employing sampling and aggregation techniques, it learns node representations effectively, capturing both local neighborhood features and global graph structural characteristics to generate graph encodings. According to the definition of a subgraph, $V_i^{(r)}$ is represented as V, and the features of the subgraph are denoted as X_V. The hidden state of vertex V at a time step of t is represented as h_v^t, and the corresponding hidden state of its adjacent subgraphs is denoted as $h_{N(V)}^t$. The initial hidden state of h_v^0 is initialized as X_v, and the calculation process is as follows:

$$h_{N(V)}^t = mean\left(\left\{\sigma\left(w_p h_{ui}^{t-1}\right), \forall u_i \in N_{(V)}\right\}\right) \tag{2}$$
$$h_v^t = \sigma\left(w_v^t \cdot CONCAT\left(h_v^{t-1}, h_{N(V)}^t\right)\right) \tag{3}$$

Among them, each layer utilizes mean aggregation, σ represents the ReLU nonlinear activation function, w_p and w_v^t are the weight matrices, $t \in \{1 \ldots k\}$. $N_{(v)}$ denotes its neighbor node-set. Neighbor subgraph information is aggregated using mean aggregation in conjunction with hidden

state information from the previous time step. Following the acquisition of neighbor information, GraphSage updates the hidden state by linking the hidden states of the previous time step h_v^{t-1} and $h_{N(V)}^t$ in preparation for the subsequent time step update. The final hidden state H^K is fed into a fully connected layer to derive the ultimate drug structure encoding $x_{ds} \in R^{n\times d}$, serving as input for the subsequent synergy layer, where n represents the number of subgraphs and d = 75.

The number of binding sites of a protein molecule is huge and contains a large number of graph nodes. Utilizing GraphSage to encode protein binding site graph is fitting. Define the binding site graph $G' = \left(V', E'\right)$. For a specific node v', its neighbor node set is $N\left(v'\right)$, the node v' is denoted by hv', and the neighbor nodes of node v' are expressed as $\left\{h_{ui'} \mid ui' \in N(v')\right\}$. The hidden state of this node is updated as follows:

$$h_{N(V')}^t = mean\left(\left\{\sigma\left(w_p h_{ui'}^{t-1}\right), \forall ui' \in N(v')\right\}\right) \tag{4}$$

$$h_{v'}^t = GlobalAttPool\left(\sigma\left(w_{v'}^t \cdot CONCAT\left(h_{v'}^{t-1}, h_{N(V')}^t\right)\right)\right) \tag{5}$$

After mean aggregation and global attention pooling operations, GraphSage effectively captures the contextual relationships of the protein node graph. To handle varying sizes of the final pooling layer, zero padding or truncation operations are applied. The final protein structure code $X_{ps} \in R^{m\times d}$ is then generated, where m represents the protein length and d=128. Subsequently, the structural code is fed into the multi-head self-attention interaction network.

2) SCALABLE GCN FOR SMILES FEATURES ENCODE OF DRUG

In the drug sequence feature encoding module, we transform each SMILES into a 2D molecular graph G. Firstly, initialize each atomic node based on its chemical properties and use the DGL-life [24] tool to represent each atom as a 128 dimensional integer vector. This vector contains eight types of information about the atom. The drug sequence is standardized to a fixed length. For molecules with fewer nodes, zero-filled virtual nodes are introduced, while longer sequences undergo truncation. The node feature of molecular graph is defined as $X_{dg} \in R^{d\times 128}$, where d is the number of nodes in the graph, and then through linear layer transformation, the dense matrix representation of drug sequence is obtained and input into GCN layer. Scalable convolutional networks gradually expand the perceptual domain of the network to be able to extract multi-scale information [25], [26]. The scalable GCN extends the convolution operator to the irregular domain and updates the atomic eigenvector by aggregating the neighbor nodes connected by chemical bonds. This broadcast pathway can automatically capture the structural information of molecules. Finally, the drug node is represented as:

$$h_d^{(m+1)} = \sigma\left(GCN\left(w_s^{(m)}, b_s^{(m)}, h_s^{(m)}\right)\right) \tag{6}$$

where $w_s^{(m)}$ and $b_s^{(m)}$ are the weight matrices and bias vectors that can be learned in each layer of GCN, and $h_d^{(m)}$ is the hidden representation of the *m-th* layer node. When m=0, $h_d^0 = X_{dg}$.

3) SCALABLE CCN FOR SEQUENCE FEATURES ENCODE OF PROTEIN

First, the protein sequence was transformed into an integer sequence, and each residue was encoded as 1-25, representing 25 residue types. According to the length distribution of protein sequence, the maximum length of protein sequence is set to 1200, the longer sequence is truncated, and the shorter sequence is zero filled. Next, the input integer sequence is transformed into a matrix representation in the potential space, and each row in the matrix represents the feature representation of the subsequence. In this way, the protein sequence is initialized into a learnable embedding matrix $X_{pg} \in R^{n\times 128}$, where n is the sequence length and 128 is the embedding dimension. The feature representation of the protein is obtained by three consecutive one-dimensional convolution layers. The convolution kernel size of the first layer is 3×3, and then each layer continues to expand the receptive field to learn the local abstract features of protein sequences. The coding process is as follows:

$$h_p^{m+1} = \sigma\left(CNN\left(w_p^{(m)}, b_p^{(m)}, h_p^{(m)}\right)\right) \tag{7}$$

where $w_p^{(m)}$ and $b_p^{(m)}$ are the learnable weight matrix and bias vector of CNN layer, $h_p^{(m)}$ is the hidden protein expression of layer m, and when m=0, $h_p^{(0)} = X_{pg}$. $\sigma(\cdot)$ is the ReLU activation function.

D. DUAL-ATTRIBUTE INTERACTION MODULE

After acquiring the drug and protein feature codes, the multi-head self-attention interactive network and the dual-weight mapping network are utilized to fuse sequence features and structural features interactively. This captures attribute information of the drug and protein from both structural and semantic perspectives. Drawing inspiration from the successful Transformer [27] and the multi-head attention model Bert [28] in natural language processing, the structural attributes of drugs and proteins are obtained through the multi-head self-attention interaction network. The semantic attributes are acquired through the dual-weight mapping network. This simultaneous learning at structural and semantic levels allows for a more comprehensive understanding of the relationship between drugs and proteins.

1) MULTI-HEAD SELF-ATTENTION INTERACTION NETWORK

In this module, this article constructs a multi-head self-attention interaction network, using the attention mechanism to integrate the structural feature encoding of proteins and drugs, as shown in Fig 1.C Multi-head attention. In previous studies, the input of multi-head attention was a separate drug code or protein code [29], which separated the correlation between drugs and proteins and failed to consider the

interaction. This module first fuses drug and protein structure codes to create a fusion structure matrix. The drug structure feature encoding, protein structure feature encoding, and fusion structure matrix are respectively projected into the query (Q), key (K), and value (V) matrices through the linear layer, where the matrix Q and the matrix K represent the query and key of the drug and protein, while matrix V represents the value of whether the drug interacts with the protein and the values are calculated taking into account the relationship between the Q and K. Subsequently, the attention score of the drug target is obtained by multiplying the query vector matrix by the transpose of the key vector matrix, by factor normalizing each element by the square root of the key dimension. Finally, the interaction matrix V is assigned through the attention score matrix, and the implementation process can be expressed as:

$$Q, K, V = linear_Q(X_p), linear_k(X_d), linear_v(X_{p+d}) \tag{8}$$

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d}}\right)V \tag{9}$$

where $X_p = h_{v'}$, $X_d = h_{v'}$, $\sqrt{d}$ is the square root factor of the key dimension, and the attention matrix is transformed into a standard normal distribution.

$$MHA\,(Q, K, V) = CONCAT\,(head_1, head_2, \ldots, head_h)\,W \tag{10}$$

where:

$$head_i = Attention\left(QW_i^Q, KW_i^W, VW_i^V\right) \tag{11}$$

$W_i^Q, W_i^K, W_i^V \in R^{m\times d_n}$ is the corresponding *i-th* head projection matrix, $W \in R^{d_n\times m}$, and d_n is the dimension of each head output. The output structure attribute of the multi-head self-attention interaction module represents y_1:

$$y_1 = FC\,(MHA\,(Q, K, V)) \tag{12}$$

2) DUAL-WEIGHT MAPPING NETWORK

To capture the pairwise local interactions between drugs and protein sequences, this paper designs a dual-weight mapping network. The dual-weight mapping network is a variant of the bilinear attention network, which is improved based on the algorithm in literature [15]. Due to excessive attention capture in the bilinear attention network, the contribution weight of the drug target to the substructure is not clear enough. Therefore, we remove the respective weight matrices of the drug and the target and upgrade the all-one vector to a learnable attention weight matrix that is used uniformly by the drug and the target. The dual weight mapping network is shown in Fig 1.C, including (1) the dual network interaction layer; (2) the pooling layer. Among them, the network action layer involves two network interactions, which are used to initialize the attention weight map and capture the paired attention weights of drug proteins; the pooling layer extracts the semantic attribute representation of drug proteins y_2.

Define the coding representation of drugs and targets $h_d = \left\{h_d^{'1}, h_d^{'2}, \ldots, h_d^{'j}, \right\}$, $h_p = \left\{h_p^{'1}, h_p^{'2}, \ldots, h_p^{'k}, \right\}$, respectively for the output of scalableGCN encoder and scalableCNN encoder. j and k represent the count of coding atoms in the drug molecule and the count of residue substructures in the protein, respectively. First, initialize the all-one matrix. The drug and protein coding are mapped from high-dimensional to low-dimensional space to the all-one matrix. The initial weight matrix A is obtained and belongs to $R^{j\times k}$:

$$A = \left(1\cdot q^T\right) \circ \sigma\left(h_d, h_p\right) \tag{13}$$

where one is a fixed all-one matrix, $q \in R^m$ is a learnable weight matrix, $\circ$ represents the Hadamard product of element units, and the value in A represents the element-level interaction strength in the drug-protein pair. For a certain substructure pair of the drug target, it can be expressed as:

$$A_{m,n} = q^T \circ \sigma\left(h_d^m, h_p^n\right) \tag{14}$$

where h_d^m represents the *m-th* column of h_d, and h_p^n represents the *n-th* column of h_p, which correspond to the representations of the *m-th* and *n-th* substructure of the drug-protein, respectively. Subsequently, the representations of drugs and proteins are concurrently projected into a feature space A with weighted representation. The drug target enhances interpretability by showcasing the contribution of the substructure to the prediction outcomes via its interaction with the feature space A. This layer can be formulated as:

$$y'_2 = \sigma\left(h_d\cdot A\cdot h_p\right) = \sum_{m=1}^{j}\sum_{n=1}^{k} A_{m,n}h_d^m h_p^n \tag{15}$$

Finally, after pooling and dimensionality reduction, a compact feature representation of the semantic attributes of the drug target in this module is obtained, where s is the step size.

$$y_2 = Avgpool\left(y'_2, s\right) \tag{16}$$

E. DECODER PREDICTION MODULE

The decoding module obtains the interactive structural attribute representation and semantic attribute representation, and sets an adaptive learnable hyperparameter α to allocate appropriate weights, thereby obtaining a multi-attribute joint feature representation for DTI prediction. In this study, the extracted attribute features are mapped through three fully connected layers, and the last layer uses the Sigmoid activation function to output the predicted value in the form of probability.

$$p = Sigmoid\,(FC\,(CONCAT\,(\alpha\times y_1, (1-\alpha)\times y_2))) \tag{17}$$

Finally, the model is trained via backpropagation to optimize all learnable parameters. The training goal is to minimize the cross-entropy loss:

$$L_{(\theta)} = -\sum_{i=1}^{n}\left(y_i \log p_i + (1-y_i)\log\left(1-p_i\right)\right) + \frac{\lambda}{2}\parallel\theta\parallel_2^2 \tag{18}$$

where θ is the set of all weight matrices and bias vectors above, n is the total number of drug-protein pairs in the training data set, y_i is the true label of the *i-th* drug target pair, p_i is the predicted value, and λ is a hyperparameter for L2 regularization.

IV. EXPERIMENTS

A. DATASETS

The model presented in this paper is compared with several state-of-the-art methods on three benchmark datasets. These datasets provide the essential 3D structural information of targets. Sample information of the dataset is provided in Table 1.

The BioSNAP dataset was created by Huang et al. [13]. from the DrugBank database [30] and consists of 4510 drugs and 2181 proteins. Following the prior works by Bai et al [15]. we used part of the data set and obtained a balanced data set with the same positive and negative samples.

BindingDB focuses on interactions between small drug molecules and proteins. It is used as real-world data to evaluate models. Reference dataset [15], target data that could provide pdb information were retained, resulting in a data set with a total of 7689 interaction information.

The Human data set was created by Liu et al. [31]. Following previous research [32], we used a balanced version of the dataset containing nearly the same number of negative and positive samples. The handling approach of this dataset is identical to that of the previous two datasets.

B. EXPERIMENTAL STRATEGIES AND EVALUATION INDICATORS

The model was implemented in PyTorch 2.1.1, with the batch size set to 120, the learning rate is set to 5e-5. Models were run for up to 50 epochs. Select the parameter model with the highest AUROC score on the validation set for testing on the test set. Our model is experimented on NVIDIA A100 80G.

We selected a total of four indicators to evaluate the model's performance, namely AUROC, AUPRC, Accuracy, and F1-score. The larger the values of AUROC and AUPRC, the better the model performance. They are the main indicators to evaluate the classification performance of the model. In addition, we also added the F1 score and accuracy index. F1 comprehensively considers the precision and recall of the model. The calculation formula is as follows:

$$Precision = \frac{TP}{(TP + FP)} \tag{19}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{20}$$

$$F1 - score = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \tag{21}$$

Accuracy calculation formula is:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{22}$$

TABLE 1. The three benchmark datasets are randomly divided into training, validation and testing sets in the ratio of 7:2:1.

Datasets	Drugs	3D-Proteins	DT pairs	Active	Inactive
BioSNAP	1506	999	2028	1059	969
BindingDB	3474	1722	7689	3891	3798
Human	2508	1553	5844	3023	2821

C. COMPARISON WITH EXISTING MODELS

To verify the effectiveness and generalization ability of proposed model, we compare its performance with existing methods on three datasets. Each baseline model is run three times using three sets of random seeds to reduce errors caused by randomness. Based on the best performance on the validation set, it is evaluated on the test set, and the average of the three test results is used as the final result. We compared the proposed model with a machine learning method (SVM) and three state-of-the-art deep DTI prediction tools (DrugBAN, Drugvqa(seq), MolTrans). For the above model, we follow the settings of the hyperparameters described by the author.

(1) The prediction of DTI is carried out on a shallow machine learning method such as SVM. The sequence features are spliced and then fed into the SVM. The kernel function is selected to be linear. The output result is the probability value of the drug target interaction prediction, and the gamma parameter is set to 0.02.

(2) Drugvqa(seq) uses dynamic convolution with sequential attention to extract protein feature representation. The size of the convolution kernel remains consistent with the original text. The difference is that the input of the protein is a sequence. The bidirectional LSTM with multi-head self-attention obtains feature representations from the molecular sequence, sums all attention vectors performs weight normalization, and finally inputs the feature splicing into the classifier for prediction.

(3) MolTrans uses a substructure mining algorithm to obtain the substructure sequence of the drug target, uses an enhanced transformer architecture to encode the semantic relationship between the substructures of the drug and the protein, and models the high-order interactions of the substructures based on the CNN layer.

(4) DrugBAN uses graph convolutional networks and one-dimensional CNN to encode sequences respectively and then inputs the encoding features into a bilinear attention network to learn the interactions of drug target pairs. The experimental results are achieved within a single domain.

1) PERFORMANCE ON THE BIOSNAP DATASET

On the BioSNAP data set, the comparison results of model MDiDTI on the main indicators AUROC and AUPRC are shown in Fig 2. Compared with existing models, model MDiDTI performs excellently on all evaluation indicators. Especially in terms of AUROC and AUPRC indicators, compared with the current highest-scoring models DrugBAN and MolTrans, our model improved by 7.06% and 3.42% respectively. Comparing these two types of models, the

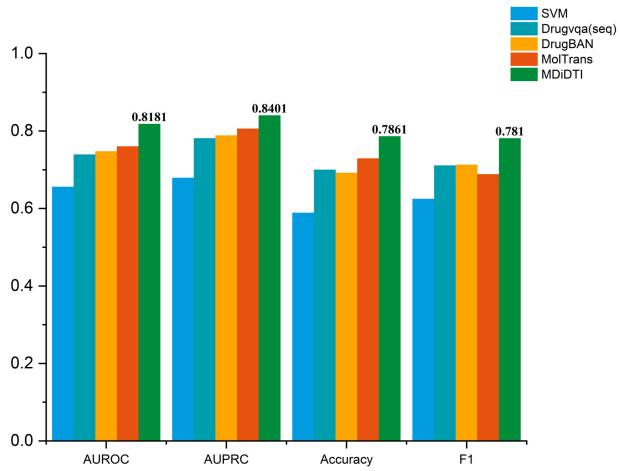

FIGURE 2. Results on the BioSNAP dataset.

TABLE 2. Results on the BindingDB dataset (best, second best).

	AUROC	AUPRC	ACCURACY	F1
SVM [33]	0.7475	0.7823	0.6743	0.6680
Drugvaq(seq) [14]	0.7547	0.7876	0.6986	0.6967
MolTrans [13]	0.82	0.8303	0.7553	0.7633
DrugBAN [15]	0.8312	0.846	0.7482	0.7661
MDiDIT	**0.8363**	**0.8515**	**0.756**	**0.7701**

bilinear attention network of the DrugBAN model effectively learns the interaction between drug targets, but does not consider the three-dimensional structure of the target, resulting in a decrease in performance; the MolTrans model uses subsequences mining algorithm obtains subsequences of drug targets, enriches semantic and structural information, and provides rich learnable parameters. However, there is still potential to enhance the effective combination of feature information, and it is certain to be competitive in terms of accuracy. Taken together, our model performs reliably and has outstanding performance on the BioSNAP data set.

2) PERFORMANCE ON THE BINDINGDB DATASET

In this section, the MDiDTI model is compared with existing prediction methods on the BindingDB dataset. We compare it with SVM, Drugvqa(seq), Moltrans, and in-domain versions of DrugBAN, and retrain all models. Parameter settings remain consistent with the original paper. On this data set, MDiDTI conducted three sets of random experiments with other models and calculated and recorded the average AUROC, AUPRC, accuracy, and F1 score.

The test process is shown in Fig 3, Fig 4, and the results are shown in Table 2. The results show that our model performs well on all indicators and is always ahead of other models. In particular, it is better than the MolTrans model in terms of accuracy, and the accuracy of the MolTrans model is higher than the DrugBAN model. By comparing the differences between these three models, we found that the model that considers the subsequence structure and the model that only considers the complete sequence has

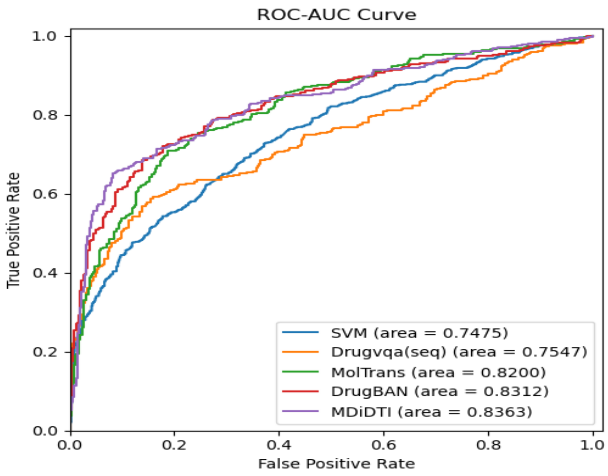

FIGURE 3. AUROC of different methods on the BindingDB dataset.

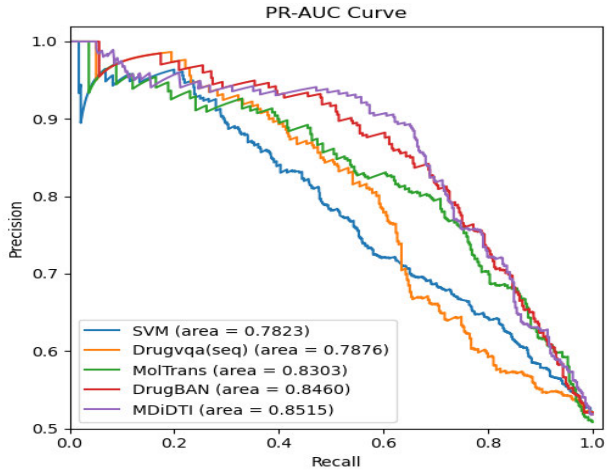

FIGURE 4. AUPRC of different methods on the BindingDB dataset.

TABLE 3. Results on Human dataset (best, second best).

	AUROC	AUPRC	ACCURACY	F1
SVM [33]	0.9368	0.9415	0.8631	0.8642
Drugvaq(seq) [14]	0.8289	0.8482	0.7644	0.7370
MolTrans [13]	0.9787	**0.9784**	0.9409	0.9410
DrugBAN [15]	0.977	0.9750	0.9427	0.9427
MDiDIT	**0.9794**	0.9765	**0.9452**	**0.9446**

excellent performance in evaluation indicators, highlighting the importance of substructure features as supplementary information for drug targets. This is further confirmed by our model.

3) PERFORMANCE ON THE HUMAN DATASET

On this data set, we still use AUROC and AUPRC as indicators to evaluate our model. the results are shown in Table 3. The MDiDTI model is better than the MolTrans model on the main indicator AUROC and is slightly lower than this model on the AUPRC indicator. Our analysis believes that due to the partially hidden target bias in the

TABLE 4. Ablation results on the BindingDB dataset.

Model	AUROC	AUPRC	ACCURACY	F1
Without pockets	0.8258	0.8367	0.7469	0.7677
Without R-radius Subgraph	0.8257	0.8329	0.7486	0.7615
Without multi-head attention	0.8299	0.8359	0.7553	0.7698
Without Dual-weigh mapping	0.8177	0.8258	0.7431	0.7692
MDiDIT	0.8363	0.8515	0.756	0.7701

Human data set [34], this will, to a certain extent, lead to the model making correct predictions based only on shallow features rather than on the true interaction, while the MolTrans model due to its The interaction module is simpler and therefore achieves slightly higher performance. On the SVM model, this is even more obvious due to its learning of shallow features. Despite this, the MDiDTI model still achieved the best results on the other three indicators and showed high accuracy and stability. Considering the overall effect, our model still shows excellent performance.

D. ABLATION STUDY

To assess the effectiveness of each module of the proposed model at various stages and its impact on the final performance, we conducted an ablation study and performed experiments on the BindingDB dataset, and designed four model configurations for comparison: (1) remove protein pockets; (2) remove R-radius subgraph; (3) remove the multi-head self-attention interaction module; (4) remove the dual weight mapping network module. The research results for different configurations are shown in Table 4.

1) THE EFFECTS OF POCKETS

To evaluate the impact of protein pockets on experimental results, we conducted experiments that omitted pocket information. As can be seen from Table 4, the model variant without considering pocket information performed lower than MDiDTI. Specifically, the model without pocket information achieves 0.8258 and 0.8367 in AUROC and AUPRC respectively. Pocket information is beneficial to DTI prediction and provides valuable protein-related information for DTI prediction.

2) THE EFFECTS OF R-RADIUS SUBGRAPH

Molecular subgraph features not only enrich the structural information of drug molecules but also provide more learnable parameters, which are helpful for model learning. Studies have shown that few studies focus on molecular structure graphs, and research on molecular subgraphs is even rarer [35]. Therefore, we explored the impact of molecular structure information in DTI prediction. The AUROC and AUPRC effects of the model that does not consider subgraphs on the test set are 0.8257 and 0.8329, which are both lower than the AUROC and AUPRC values of MDiDTI (0.8363 and 0.8515 respectively). At the same time, Accuracy is 0.7486 and F1 is 0.7615, both lower than the Accuracy and F1 values of MDiDTI. In summary, substructure information is effective as a structural input for drugs and can be used as supplementary drug information in DTI prediction.

3) THE EFFECTS OF MULTI-HEAD SELF-ATTENTION INTERACTION NETWORK

To evaluate the impact of our proposed multi-head self-attention interaction module for drug-target interaction on the overall model performance, in the interaction part, we only retained the dual-weight mapping network and the interaction of the structure was simply performed on the structural feature attributes splicing. The evaluation results in Table 4. The AUROC value and AUPRC value of the model that does not include the multi-head self-attention interaction module are 0.8299 and 0.8359 respectively, which are both lower than MDiDTI's 0.8363 and 0.8515. In addition, judging from the Accuracy value and F1 value, the MDiDTI also boasts excellent performance. Therefore, this interaction network we proposed plays an obvious role in DTIs. This also verifies the importance of structural information in drug-protein interaction prediction experiments and cannot be ignored.

4) THE EFFECTS OF DUAL-WEIGHT MAPPING NETWORK

By adjusting the dual-weight mapping network, the interaction between drug and protein sequences overlooked the processing of the dual-weight mapping network, opting solely for vector summation of the feature representations. Experimental results demonstrate that in models lacking the dual-weight mapping network, the AUROC value is 0.8177, and the AUPRC value is 0.8258, with MDiDTI exhibiting improvements of 2.27% and 3.11%, respectively. From other metrics, the Accuracy value of 0.7431 and the F1 score of 0.7692 are both inferior to those of MDiDTI. These results highlight the effectiveness of the dual-weight mapping network in capturing information regarding drug-target sequence interactions, significantly enhancing the overall performance of the MDiDTI.

V. DISCUSSION

We propose an approach grounded in graph neural networks and attention mechanisms for discerning drug-protein interactions. Addressing the limitations of extant methods, which often overlook the three-dimensional protein structure and lack interpretability in drug-target interactions, we present a solution termed drug-target multi-feature encoding dual-attribute interactive learning. This framework comprehensively integrates structural and semantic attributes and incorporates an Attention mechanism module to enhance interpretability. We analyzed that there are several factors for the model to achieve excellent performance performance. First, the input representation module, where the input representation significantly affects the predictive performance of the model. The use of more advanced and comprehensive input feature representations, such as structure maps and sequences, can effectively capture the chemical information

and topological relationships of drug molecules and targets, further improving the performance. Second, the prediction module. Traditional ML-based techniques rely heavily on the quality of hand-crafted features that are unable to learn complex nonlinear relationships, while self-attentive mechanisms provide powerful automatic feature extraction to learn higher-order nonlinear relationships. Despite the considerable advancements in protein three-dimensional structure prediction, the following limitations still exist:

- Limitations of datasets. Not every drug-target interaction pair in the various current datasets has its corresponding protein PDB file confirmed by wet experiments. Thus limiting the use of datasets in the study.
- Limitations of structure prediction. We used a structure prediction method based on 3D bounding box coordinates to predict protein binding sites [19], which achieved more than 70% accuracy in both the dataset with 86 elements and the dataset with 130 elements, leaving some room for optimization.

Nevertheless, our proposed method exhibits notable efficacy. Future endeavors will prioritize refining pocket prediction techniques for enhanced accuracy and leveraging more extensive datasets to optimize and bolster the stability of the model.

VI. CONCLUSION

In this study, we propose an aggregation model named MDiDTI, which effectively integrates multi-attribute features of drug targets through multi-feature encoding dual-attribute interactive learning. In this work, four different inputs of the drug target are deeply represented, including the sequence information, the target pocket, and the drug molecular graph. Subsequently, multi-feature encoders for drugs and targets are designed to encode the embedding matrix. Finally, a strategy is designed to combine the multi-head self-attention interaction network and the dual-weight mapping network to achieve the fusion and capture of structural attributes and semantic attributes at the structural level and semantic level. It outperforms existing methods in prediction DTI experiments and exhibits a high degree of generalization ability on different datasets. Looking forward, we believe that with the continuous development of protein structure prediction and binding site identification technologies, the integration of protein three-dimensional structure into DTI prediction and other related prediction tasks holds promising prospects. We aspire for the MDiDTI model proposed in this paper to serve as an effective virtual screening tool in the field of drug research and development.

REFERENCES

[1] M. Ma and X. Lei, "A dual graph neural network for drug–drug interactions prediction based on molecular structure and interactions," *PLOS Comput. Biol.*, vol. 19, no. 1, Jan. 2023, Art. no. e1010812.
[2] X. Su, P. Hu, H. Yi, Z. You, and L. Hu, "Predicting drug-target interactions over heterogeneous information network," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 1, pp. 562–572, Jan. 2023.
[3] R. Chen, F. Xia, B. Hu, S. Jin, and X. Liu, "Drug–target interactions prediction via deep collaborative filtering with multiembeddings," *Briefings Bioinf.*, vol. 23, no. 2, 2022, Art. no. bbab520.
[4] S. M. Kandathil, J. G. Greener, A. M. Lau, and D. T. Jones, "Ultrafast end-to-end protein structure prediction enables high-throughput exploration of uncharacterized proteins," *Proc. Nat. Acad. Sci. USA*, vol. 119, no. 4, 2022, Art. no. e2113348119.
[5] V. Kairys, L. Baranauskiene, M. Kazlauskiene, D. Matulis, and E. Kazlauskas, "Binding affinity in drug design: Experimental and computational techniques," *Expert Opinion Drug Discovery*, vol. 14, no. 8, pp. 755–768, 2019.
[6] Y. Cao, T. A. Geddes, J. Y. H. Yang, and P. Yang, "Ensemble deep learning in bioinformatics," *Nature Mach. Intell.*, vol. 2, no. 9, pp. 500–508, 2020.
[7] Y. Yamanishi, M. Araki, A. Gutteridge, W. Honda, and M. Kanehisa, "Prediction of drug–target interaction networks from the integration of chemical and genomic spaces," *Bioinformatics*, vol. 24, no. 13, pp. i232–i240, Jul. 2008.
[8] T. van Laarhoven, S. B. Nabuurs, and E. Marchiori, "Gaussian interaction profile kernels for predicting drug–target interaction," *Bioinformatics*, vol. 27, no. 21, pp. 3036–3043, Nov. 2011.
[9] A. S. Rifaioglu, E. Nalbat, V. Atalay, M. J. Martin, R. Cetin-Atalay, and T. Doğan, "DEEPScreen: High performance drug–target interaction prediction with convolutional neural networks using 2-D structural compound representations," *Chem. Sci.*, vol. 11, no. 9, pp. 2531–2557, 2020.
[10] Y. Luo, X. Zhao, J. Zhou, J. Yang, Y. Zhang, W. Kuang, J. Peng, L. Chen, and J. Zeng, "A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information," *Nature Commun.*, vol. 8, no. 1, p. 573, Sep. 2017.
[11] G. Li, P. Zhang, W. Sun, C. Ren, and L. Wang, "Bridging-BPs: A novel approach to predict potential drug–target interactions based on a bridging heterogeneous graph and BPs2vec," *Briefings Bioinf.*, vol. 23, no. 2, 2022, Art. no. bbab557.
[12] Y. Gan, X. Huang, W. Guo, C. Yan, and G. Zou, "Predicting synergistic anticancer drug combination based on low-rank global attention mechanism and bilinear predictor," *Bioinformatics*, vol. 39, no. 10, 2023, Art. no. btad607.
[13] K. Huang, C. Xiao, L. M. Glass, and J. Sun, "MolTrans: Molecular interaction transformer for drug–target interaction prediction," *Bioinformatics*, vol. 37, no. 6, pp. 830–836, May 2021.
[14] S. Zheng, Y. Li, S. Chen, J. Xu, and Y. Yang, "Predicting drug–protein interaction using quasi-visual question answering system," *Nature Mach. Intell.*, vol. 2, no. 2, pp. 134–140, 2020.
[15] P. Bai, F. Miljković, B. John, and H. Lu, "Interpretable bilinear attention network with domain adaptation improves drug–target prediction," *Nature Mach. Intell.*, vol. 5, no. 2, pp. 126–136, Feb. 2023.
[16] I. Wallach, M. Dzamba, and A. Heifets, "AtomNet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery," 2015, *arXiv:1510.02855*.
[17] M. Ragoza, J. Hochuli, E. Idrobo, J. Sunseri, and D. R. Koes, "Protein–ligand scoring with convolutional neural networks," *J. Chem. Inf. Model.*, vol. 57, no. 4, pp. 942–957, 2017.
[18] Y. Zhu, L. Zhao, N. Wen, J. Wang, and C. Wang, "DataDTA: A multi-feature and dual-interaction aggregation framework for drug–target binding affinity prediction," *Bioinformatics*, vol. 39, no. 9, 2023, Art. no. btad560.
[19] S. M. Saberi Fathi and J. A. Tuszynski, "A simple method for finding a protein's ligand-binding pockets," *BMC Struct. Biol.*, vol. 14, p. 18, Jul. 2014.
[20] M. Yazdani-Jahromi, N. Yousefi, A. Tayebi, E. Kolanthai, C. J. Neal, S. Seal, and O. O. Garibay, "AttentionSiteDTI: An interpretable graph-based model for drug-target interaction prediction using NLP sentence-level relation classification," *Briefings Bioinf.*, vol. 23, no. 4, 2022, Art. no. bbac272.
[21] Y. Chen, T. Ma, X. Yang, J. Wang, B. Song, and X. Zeng, "MUFFIN: Multi-scale feature fusion for drug–drug interaction prediction," *Bioinformatics*, vol. 37, no. 17, pp. 2651–2658, Sep. 2021.
[22] M. Tsubaki, K. Tomii, and J. Sese, "Compound–protein interaction prediction with end-to-end learning of neural networks for graphs and sequences," *Bioinformatics*, vol. 35, no. 2, pp. 309–318, Jan. 2019.
[23] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.

[24] M. Li, J. Zhou, J. Hu, W. Fan, Y. Zhang, Y. Gu, and G. Karypis, "DGL-LifeSci: An open-source toolkit for deep learning on graphs in life science," *ACS Omega*, vol. 6, no. 41, pp. 27233–27238, 2021.
[25] S. Zhang, X. Zhang, S. Wan, W. Ren, L. Zhao, and L. Shen, "Generative adversarial and self-supervised dehazing network," *IEEE Trans. Ind. Informat.*, vol. 20, no. 3, pp. 4187–4197, Mar. 2024.
[26] Y. Liu, Z. Yan, J. Tan, and Y. Li, "Multi-purpose oriented single nighttime image haze removal based on unified variational retinex model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 4, pp. 1643–1657, Apr. 2023.
[27] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.
[28] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.
[29] L. Huang, J. Lin, R. Liu, Z. Zheng, L. Meng, X. Chen, X. Li, and K.-C. Wong, "CoaDTI: Multi-modal co-attention based framework for drug–target interaction annotation," *Briefings Bioinf.*, vol. 23, no. 6, 2022, Art. no. bbac446.
[30] D. S. Wishart, C. Knox, A. C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B. Gautam, and M. Hassanali, "DrugBank: A knowledgebase for drugs, drug actions and drug targets," *Nucleic Acids Res.*, vol. 36, pp. D901–D906, Jan. 2008.
[31] H. Liu, J. Sun, J. Guan, J. Zheng, and S. Zhou, "Improving compound–protein interaction prediction by building up highly credible negative samples," *Bioinformatics*, vol. 31, no. 12, pp. i221–i229, Jun. 2015.
[32] F. Li, Z. Zhang, J. Guan, and S. Zhou, "Effective drug–target interaction prediction with mutual interaction neural network," *Bioinformatics*, vol. 38, no. 14, pp. 3582–3589, Jul. 2022.
[33] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, Sep. 1995.
[34] L. Chen, X. Tan, D. Wang, F. Zhong, X. Liu, T. Yang, X. Luo, K. Chen, H. Jiang, and M. Zheng, "TransformerCPI: Improving compound–protein interaction prediction by sequence-based deep learning with self-attention mechanism and label reversal experiments," *Bioinformatics*, vol. 36, no. 16, pp. 4406–4414, 2020.
[35] Z. Yang, W. Zhong, L. Zhao, and C. Y.-C. Chen, "ML-DTI: Mutual learning mechanism for interpretable drug–target interaction prediction," *J. Phys. Chem. Lett.*, vol. 12, no. 17, pp. 4247–4261, 2021.

YUANDONG LIU was born in Shandong, China, and received his Bachelor of Engineering degree from Shandong University of Technology in 2020, where he is currently pursuing his Master's degree. His research interests include bioinformatics and deep learning.

HAOQIN YANG was born in Shandong, China. He received the B.S. degree in computer science from Shandong University of Technology, in 2022, where he is currently pursuing the M.S. degree. His research interests include bioinformatics and deep learning.

LONGBO ZHANG was born in Shandong, China, in 1968. He received the Ph.D. degree, in 2008. He has presided over and participated in more than 30 National Key Research and Development projects, National Natural Science Foundation of China projects, and Social Science Foundation projects. His research interests include databases and big data analysis and mining.

HONGZHEN CAI received the Ph.D. degree from Shenyang Agricultural University, Shenyang, China, in 2015. She is currently a Master's Tutor with Shandong University of Technology. Her research interest includes bio-based materials.

MAOZU GUO received the Ph.D. degree from Harbin Institute of Technology, Harbin, China, in 1997. He is currently a Ph.D. Supervisor with Beijing University of Civil Engineering and Architecture. He is the Key Project Leader of the National Natural Science Foundation of China and the National Key Research and Development Task. He has published more than 200 articles. His research interests include machine learning, artificial intelligence, intelligent construction, smart cities, and bioinformatics. He received the Natural Science Award of the Ministry of Education and the Provincial Natural Science Award.

LINLIN XING received the Ph.D. degree from Harbin Institute of Technology, Harbin, China, in 2018. He is currently a Master's Tutor with Shandong University of Technology. His research interests include bioinformatics and machine learning.