Dual-Modality Neural Network Integrating Sequence and Structural Information for Drug-Target Interaction Prediction

Dongjie Chen, Yuemeng Liu, Binfeng Liu, Zican Wang, Hangxing Lin, and Zhihui Yang*

Foshan University, Foshan, China
zhy@whu.edu.cn

Abstract. Accurately predicting drug-target interactions (DTI) is central to accelerating computational drug discovery; however, existing methods still face challenges in effectively constructing multimodal features, fusing heterogeneous features, and capturing complex non-linear relationships. This paper proposes a novel dual-modal neural network framework, MGK-DTI, designed to deeply integrate the sequence and structural information of drugs and proteins to address these challenges. First, we designed Mambaformer in the sequence branch, utilizing Mamba to capture long-range dependencies while leveraging the advantages of the multi-head attention mechanism in global context; subsequently, we designed a hybrid Gated Graph Convolutional Network in the structural branch, which uses an adaptive gating mechanism to dynamically balance intra-modal topological updates with cross-modal feature injection, achieving bidirectional fine-grained interaction between protein structural features and drug molecular graph structures; finally, after unifying the dual-path feature space through the Latent Alignment Module, the Kolmogorov-Arnold Network is introduced to strengthen nonlinear interaction modeling. Experimental results on four benchmark datasets demonstrate that MGK-DTI outperforms state-of-the-art methods. Ablation studies validate the contribution of each core module, and the case study on Histamine H1 receptor further confirms the immense potential of the model in virtual screening. Code is available at https://github.com/HellojiePro/MGK-DTI.

Keywords: Drug-target interaction · Mamba · KAN · Deep learning.

1 Introduction

Accurate identification of drug-target interactions (DTI) is beneficial for drug discovery and repositioning [2, 19, 23]. However, DTI prediction methods based on wet-lab experiments are expensive, laborious, and time-consuming [4]. In recent years, with the explosive growth of bioinformatics data and the evolution

* Corresponding author: Zhihui Yang (Email: zhy@whu.edu.cn)

of deep learning technologies, deep learning-based DTI prediction methods have become an indispensable auxiliary tool in the field of drug research and development [1, 18].

These methods typically adopt specific biological sequences or molecular graph data as input, extracting features and effectively modeling non-linear relationships through sequence models (such as CNN, RNN, and Transformer) or GNN [15]. For instance, DeepConv-DTI [14] utilizes 1D-CNN to extract protein features and ECFP to encode drug structures, followed by prediction via fully connected networks; MCANet [16] captures local drug-protein correlation features under different receptive fields through multi-scale convolution and cross-attention mechanisms; HyperAttentionDTI [24] combines hyper-attention mechanisms to enhance feature representation; TransformerCPI [3] utilizes GNN to encode drugs and integrates a Transformer decoder to capture cross-attention between atoms and amino acids; the IIFDTI [5] method combines CNN and GATs within an encoder-decoder framework for drug-protein representation learning. Despite significant progress, existing methods are still limited to single-modality feature extraction, ignoring the significant complementarity between different modalities and failing to fully utilize the multi-dimensional information of drugs and targets. The rise of biological omics has driven the development of DTI research, leading to methods that utilize multi-modal data. UnitedDTA [8] integrates 1D, 2D, and 3D embeddings of drugs and proteins to obtain a unified and discriminative representation aligned across modalities. MMDG-DTI [9] merges generalized text and structural features through domain adversarial training and contrastive learning, significantly enhancing model capabilities. TriMulDTI [6] integrates three modalities—drug graphs, sequences, and protein structures—to improve prediction performance using multi-modal fusion. Currently, DTI models not only need to fully utilize multi-modal features but also require the design of efficient fusion strategies to deeply mine the synergistic and complementary relationships between different modalities.

To address the aforementioned challenges, this study proposes a novel drug-target interaction prediction framework (MGK-DTI). This framework integrates the strengths of Mamba in capturing long-range dependencies with the advantages of multi-head attention mechanisms in focusing on short-range dependencies and global attention features, achieving complementary advantages to significantly improve the accuracy and efficiency of DTI prediction. Simultaneously, we designed a Hybrid Gated Graph Convolutional Network (HGGCN) to enhance model performance at the structural modality level by extracting features reflecting the structural relationships between drug atoms and protein amino acids. Finally, multi-modal features are unified in space via the Latent Alignment Module (LAM) and undergo deep fusion and non-linear modeling by the Kolmogorov-Arnold Network (KAN) to complete the final interaction prediction. The main contributions are as follows: (1) introducing the Mambaformer architecture in the sequence branch to effectively capture long-range and short-range dependencies within sequences; (2) establishing a structural branch based on HGGCN to supplement topological features; (3) introducing the KAN to sig-

nificantly enhance the modeling capability for complex non-linear interactions; (4) comprehensive experiments demonstrate that MGK-DTI outperforms state-of-the-art methods, validating its effectiveness in DTI prediction.

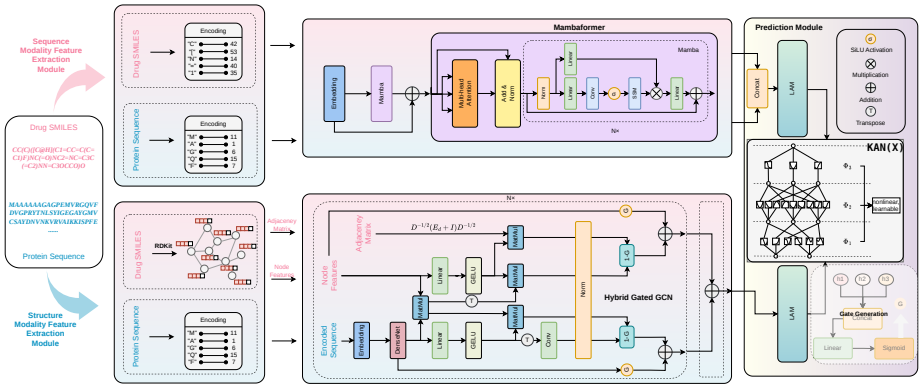

Fig. 1: Overview of the MGK-DTI framework.

2 Methodology

The overall workflow of the MGK-DTI framework is illustrated in Fig. 1. This process mainly includes three consecutive stages: first, in the feature construction stage, the raw drug molecules and target proteins are converted into input representations processable by the model; subsequently, the Sequence Modality Feature Extraction Branch and the Structure Modality Feature Extraction Branch are respectively utilized to extract the global semantics of 1D sequences and the topological features of 2D graph structures; finally, it enters the Prediction Module, where the extracted dual-stream high-level features are aligned and deeply fused, thereby outputting the final interaction prediction result.

2.1 Sequence Modality Feature Extraction Branch

The sequence branch of this research adopts the Mambaformer architecture: after the mapping of tokens in the embedding layer, the first Mamba layer captures initial temporal information. Since Mamba is capable of implicitly capturing the temporal information within the sequence, the model can learn the positional information of the sequence without the need for additional positional encodings. The calculation process is as follows:

$$H_x = \text{Mamba}(E_x) \in \mathbb{R}^{L_x \times D_{seq}}, \quad x \in \{d, p\} \tag{1}$$

where E_x denotes the output features of the embedding layer. Subsequently, we adopt a standard multi-head self-attention mechanism to capture various

dependencies between different positions in the sequence, thereby extracting global context information $H_x^{att} \in \mathbb{R}^{L_x \times D_{seq}}$. By processing multiple attention heads in parallel, each head can focus on different patterns of the sequence, effectively compensating for the limitations of local modeling.

Although the Transformer model performs excellently in capturing global dependencies, the quadratic time complexity of its attention mechanism limits its application in capturing long-range dependencies, making it suitable only for modeling short-range dependencies. To this end, we employ stacked Mamba blocks to ensure that important correlations and long-range dependencies throughout the entire process are effectively preserved. The specific calculation process is as follows:

$$X_x = \sigma(\text{Conv1D}(\omega(H_x^{att}))), \quad Z_x = \sigma(\omega(H_x^{att})), \quad x \in \{d, p\} \tag{2}$$

$$H'_x = \text{LayerNorm}(\omega(\text{SSM}(X_x) \cdot Z_x) + H_x^{att}) \in \mathbb{R}^{L_x \times D_{seq}}, \quad x \in \{d, p\} \tag{3}$$

where $\omega(\cdot)$ denotes the linear layer, $\sigma(\cdot)$ denotes the SiLU activation function, Conv1D$(\cdot)$ denotes the 1D convolution operation, and SSM$(\cdot)$ represents the fundamental calculation of the State Space Model. Finally, the high-order sequence representation $H'_{seq} = [H'_d \oplus H'_p] \in \mathbb{R}^{(L_d+L_p) \times D_{seq}}$ is obtained by concatenating the features of the drug and the target. For more details about Mamba, readers may refer to the literature [7].

2.2 Structure Modality Feature Extraction Branch

To incorporate the structural topological correlation between drugs and target proteins into the DTI model, this research designs a structural feature extractor based on the HGGCN.

Regarding the design logic of this branch, drug molecules can be constructed as topological graphs with atoms as nodes through graph representation; in contrast, protein sequences cannot directly form a graph structural relationship between amino acids. Although spatial structures can be predicted through deep learning, high computational overhead and prediction errors often limit model performance. To this end, this paper first utilizes an embedding layer and a DenseNet [12] encoder to extract the base protein features $H_p \in \mathbb{R}^{L_p \times D_{str}}$. Subsequently, we treat adjacent amino acids as connected nodes and utilize 1D Depthwise Separable Convolution (DSConv) [9] to process H_p, thereby efficiently extracting the internal structural topological information of the protein.

For the drug modality, drug molecules are parsed via RDKit to obtain the initial node matrix $H_d \in \mathbb{R}^{L_d \times D_{str}}$ and adjacency matrix $E_d \in \{0, 1\}^{L_d \times L_d}$. Subsequently, symmetric normalization is performed on the drug adjacency matrix:

$$\tilde{A} = D^{-1/2}(E_d + I)D^{-1/2} \tag{4}$$

where I is the identity matrix and D is the degree matrix with self-loops. In the core HGGCN layers, features are dynamically updated through cascaded interaction blocks. Within each interaction layer, drug features H_d and protein

features H_p are first transformed into a to-be-aggregated state via a non-linear projection, followed by intra-modal graph convolutions to capture topological information and cross-modal correlation feature capture to extract interaction features. The specific process is as follows:

$$H_d^{intra} = BN(\tilde{A} \cdot \psi(H_d)), \quad H_d^{inter} = BN(\text{softmax}(H_d H_p^T) \cdot \psi(H_p)) \tag{5}$$

$$H_p^{intra} = BN(DSConv(\psi(H_p))), \quad H_p^{inter} = BN(\text{softmax}(H_p H_d^T) \cdot \psi(H_d)) \tag{6}$$

where $\psi(\cdot)$ denotes the linear layer and GELU activation function, $H^{intra} \in \mathbb{R}^{L_x \times D_{str}}$ denotes the intra-modal topological features, and $H^{inter} \in \mathbb{R}^{L_x \times D_{str}}$ denotes the inter-modal interaction features extracted through cross-modal correlation mapping. To effectively integrate the original information with the newly extracted multi-source features, this study introduces a Gated Feature Fusion mechanism. This mechanism calculates a dynamic gating coefficient G_x by concatenating the original features, intra-graph features, and interaction features:

$$G_x = \sigma(\omega(H_x \oplus H_x^{intra} \oplus H_x^{inter})), \quad x \in \{d, p\} \tag{7}$$

$$H'_x = (1 - G_x) \odot H_x + G_x \odot \tanh(\omega(H_x^{intra} \oplus H_x^{inter})) \in \mathbb{R}^{L_x \times D_{str}}, \quad x \in \{d, p\} \tag{8}$$

where $\omega(\cdot)$ denotes the linear projection, σ is the Sigmoid activation function, $\oplus$ denotes feature concatenation, and $\odot$ is the element-wise multiplication. Finally, after N layers of stacked processing, the updated protein features H'_p and drug features H'_d are concatenated to obtain the high-order structural modality representation $H'_{str} \in \mathbb{R}^{(L_d + L_p) \times D_{str}}$.

2.3 Prediction Module and Loss Function

After completing the dual-branch feature extraction, since the resulting features originate from different semantic spaces, we utilize LAM to unify the dual-path feature spaces:

$$H_m^{l+1} = \phi(Conv1D(H_m^l)) + H_m^l \tag{9}$$

$$v_m = MLP(GAP(H_m^L)) \in \mathbb{R}^{D_{latent}}, \quad m \in \{seq, str\} \tag{10}$$

where $H_m^0 = H'_m$. The model projects the features into a shared latent space through residual 1D convolution, Global Average Pooling (GAP), and Multi-Layer Perceptron (MLP). Subsequently, we utilize KAN [17] to replace the traditional MLP to enhance the ability to model complex non-linear interactions, and the generated final prediction output $\hat{y}$ is:

$$\hat{y} = KAN([v_{seq} \oplus v_{str}]) \in [0, 1] \tag{11}$$

The model is trained using the Binary Cross-Entropy (BCE) loss, and the total loss is composed of a weighted combination of the main prediction loss and the auxiliary supervisory losses of the branches:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{main} + \lambda_2 \mathcal{L}_{seq} + \lambda_3 \mathcal{L}_{str} \tag{12}$$

where λ_i are hyperparameters used to balance the contribution of each loss term.

Table 1: Statistics of the benchmarking datasets.

Dataset	Protein	Drug	Interaction	Positive	Negative
Human	852	1,052	6,212	3,369	2,843
C.elegans	2,504	1,434	7,511	4,000	3,511
BindingDB	811	49,567	60,780	33,777	27,003
DrugBank	4,254	6,645	35,022	17,511	17,511

Table 2: Hyper-parameter settings.

Parameter	Description	Value
L_d, L_p	Length	100, 1000
D_{seq}, D_{str}	Dimension	256, 75
D_{latent}	Dimension	128
LR, Epoch	Training	1×10^{-5}, 100
Dropout	Training	0.05
Batch size	Batch size	16 / 64

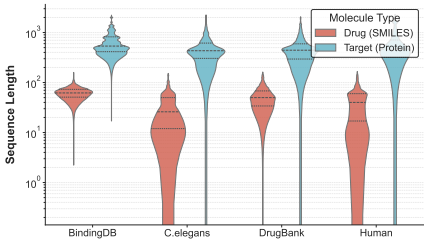

Fig. 2: Distribution of sequence.

3 Experiments and Results

Dataset. We utilized four benchmark DTI datasets: Human [10], C.elegans [10], DrugBank [24], and BindingDB [10]. These datasets are structured in text format, comprising drug SMILES strings, protein amino acid sequences, and binary interaction labels. Detailed statistical summaries for each dataset are provided in Table 1.

Implementation Details. As illustrated in Fig. 2, to strike a balance between computational efficiency and information integrity, we limited the maximum lengths of drug SMILES and protein sequences to 100 and 1,000, respectively. This configuration covers the vast majority of samples in the datasets: while insufficient lengths risk the loss of critical functional domains or chemical structures due to truncation, excessive lengths would introduce redundant padding noise and increase computational overhead. Consequently, we normalized the sequence lengths through truncation or zero-padding to ensure uniform input dimensions for the model.Other hyperparameters and experimental settings are summarized in Table 2.

Evaluation Metrics and Experimental Environment Area Under the ROC Curve (AUC-ROC) and Area Under the Precision-Recall Curve (AUC-PR) were utilized as the primary evaluation metrics. All experiments were conducted on a system equipped with an Intel Xeon Gold 6348 CPU (2.60 GHz) and an NVIDIA A800-SXM4-80GB GPU (80 GB).

3.1 Comparative experiments

To evaluate the performance of MGK-DTI, we conduct comparative experiments with seven mainstream methods, including TransformerCPI [3], HADTI [24], GanDTI [21], MCANet [16], TriMulDTI [6], MFR-DTA [11], and PSC-CPI [22].

Table 3 summarizes the performance comparison between MGK-DTI and various baseline models across four public datasets. The experimental results show that MGK-DTI achieves the best performance across all evaluation metrics. Specifically, on the Human and C.elegans datasets, MGK-DTI outperforms other methods, fully demonstrating its superior performance on small-scale datasets. On the larger-scale DrugBank and BindingDB datasets, MGK-DTI maintains a significant leading advantage, further verifying the powerful generalization capability and superiority of the method across different dataset scales.

Table 3: Comprehensive Performance Comparison Across All Datasets

	Human		C.elegans		DrugBank		BindingDB	
Method	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR
TransformerCPI	0.973	0.979	0.988	0.987	0.837	0.836	0.951	0.949
HADTI	0.978	0.980	0.985	0.987	0.889	0.897	0.956	0.951
GanDTI	0.984	0.973	0.985	0.988	0.855	0.854	0.985	0.987
MCANet	0.983	0.989	0.989	0.990	0.897	0.902	0.961	0.963
TriMulDTI	0.985	0.986	0.986	0.988	0.895	0.897	0.963	0.967
MFR-DTA	0.986	0.983	0.988	0.985	0.893	0.891	0.962	0.959
PSC-CPI	0.988	0.993	0.993	0.992	0.895	0.897	0.969	0.972
MGK-DTI	**0.991**	**0.993**	**0.996**	**0.997**	**0.902**	**0.907**	**0.988**	**0.988**

3.2 Ablation Study

To evaluate the contribution of different functional modules to MGK-DTI, we conduct ablation studies on the smaller C.elegans and larger BindingDB datasets, comparing the following five model variants: (i) **w/o MHA**: removing the multi-head attention mechanism; (ii) **w/o Gated**: removing the gated fusion mechanism in the structural branch; (iii) **w/o LAM**: removing LAM and directly concatenating features; (iv) **w/o KAN**: replacing KAN with a standard MLP; (v) **w/o Seq**: removing the sequence modality branch; (vi) **w/o Str**: removing the structural modality branch. The experimental results in Table 4 demonstrate that each constituent module of MGK-DTI plays an indispensable role in enhancing model efficacy, collectively validating the rationality and necessity of the model design.

3.3 Case Study

The Histamine H1 receptor is a pivotal member of the G-protein coupled receptor (GPCR) family, playing a crucial role in mediating allergic reactions, inflammatory processes, and the regulation of sleep-wake cycles, making it a central target

Table 4: Performance metrics for ablation studies (AUC-ROC) evaluated on C.elegans and BindingDB.

Variant	C.elegans	BindingDB
w/o MHA	0.994	0.984
w/o Gated	0.992	0.985
w/o LAM	0.990	0.983
w/o KAN	0.993	0.982
w/o Seq	0.989	0.981
w/o Str	0.991	0.981
MGK-DTI	**0.996**	**0.988**

Table 5: Top 10 DTI predictions.

Drug Name	Rank	Evidence
DB07227	1	N/A
DB07563	2	N/A
DB00719	3	DrugBank, TTD
DB00920	4	DrugBank, TTD
DB06153	5	DrugBank
DB08685	6	N/A
DB03507	7	N/A
DB06678	8	DrugBank
DB06766	9	DrugBank, TTD
DB08802	10	DrugBank

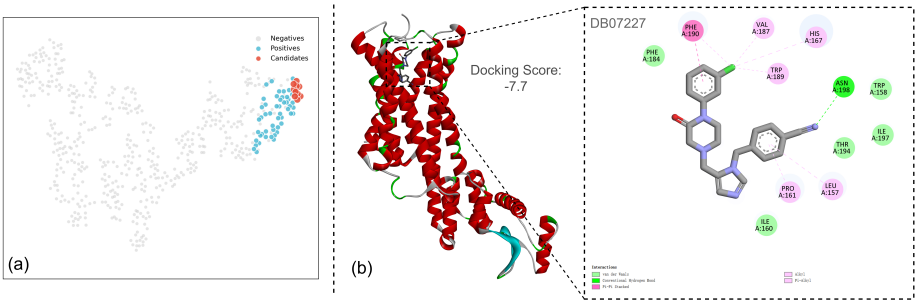

Fig. 3: Fig. 3: Case study of P35367. (a) t-SNE visualization of the latent space. (b) Molecular docking poses and key 2D interactions.

for antihistamine drug development. To further validate the practical effectiveness of the MGK-DTI model, we conducted a case study on the target protein P35367 (Histamine H1 receptor). As shown in Table 5, the majority of the top 10 candidate drugs predicted by the model have been confirmed in external databases such as DrugBank [13] and TTD [25]. Furthermore, as illustrated in Fig. 3(a), we generated a t-SNE visualization of the latent features based on the top 10 candidate drugs, 50 randomly selected positive samples, and 500 negative samples. The results indicate that the model can effectively map the candidate drugs and known positive samples into a similar feature space, demonstrating good discrimination and clustering capabilities. The molecular docking [20] results with the top-ranked candidate drug in Fig. 3(b) further reveal its stable binding with key residues (such as PHE A:190 and TRP A:189). These results demonstrate the immense potential of the model in virtual screening.

4 Conclusion

In this work, we propose MGK-DTI, a novel dual-modality framework that deeply integrates sequence and structural information for DTI prediction. It employs Mambaformer to extract the global and long-range dependencies of sequences, and captures structural and interaction features through a HGGCN. Furthermore, the model unifies the feature space via a LAM and utilizes the KAN network to achieve highly non-linear interaction prediction. Experimental results and case studies fully demonstrate the superior performance of MGK-DTI, validating its effectiveness in DTI prediction and virtual screening.

Acknowledgments. This study was funded by the Guangdong Basic and Applied Basic Research Foundation (No. 2024A1515110165), the Guangdong Medical Science and Technology Research Fund (No. B2026201), the Foshan Science and Technology Innovation Project (No. 2420001004566), and the Foshan Emergency Science and Technology Program for Chikungunya Fever Control (No. JYSZ2025001).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ahmad, B., Ouahada, K., Hamam, H.: Machine learning for drug-target interaction prediction: A comprehensive review of models, challenges, and computational strategies. Computational and Structural Biotechnology Journal (2026)
2. Aragh, A.H., Amirani, R.M., Givehchian, P., Eskafi, M., Bajgiran, Y.A., Eslahchi, C.: Mirage-dti: a novel approach for drug–target interaction prediction by integrating drug and target similarity metrics. Computers in Biology and Medicine **192**, 110249 (2025)
3. Chen, L., Tan, X., Wang, D., Zhong, F., Liu, X., Yang, T., Luo, X., Chen, K., Jiang, H., Zheng, M.: Transformercpi: improving compound–protein interaction prediction by sequence-based deep learning with self-attention mechanism and label reversal experiments. Bioinformatics **36**(16), 4406–4414 (2020)
4. Chen, R., Liu, X., Jin, S., Lin, J., Liu, J.: Machine learning for drug-target interaction prediction. Molecules **23**(9), 2208 (2018)
5. Cheng, Z., Zhao, Q., Li, Y., Wang, J.: Iifdti: predicting drug–target interactions through interactive and independent features based on attention mechanism. Bioinformatics **38**(17), 4153–4161 (2022)
6. Dehghan, A., Razzaghi, P., Abbasi, K., Gharaghani, S.: Tripletmultidti: multimodal representation learning in drug-target interaction prediction with triplet loss function. Expert Systems with Applications **232**, 120754 (2023)
7. Gu, A., Dao, T.: Mamba: Linear-time sequence modeling with selective state spaces. In: First conference on language modeling (2024)
8. Hu, J., Liu, Y., Zeng, X., Zou, Q., Su, R., Wei, L.: Multi-modal deep representation learning accurately identifies and interprets drug-target interactions. IEEE Journal of Biomedical and Health Informatics (2025)
9. Hua, Y., Feng, Z., Song, X., Wu, X.J., Kittler, J.: Mmdg-dti: Drug–target interaction prediction via multimodal feature fusion and domain generalization. Pattern Recognition **157**, 110887 (2025)

10. Hua, Y., Song, X., Feng, Z., Wu, X.J., Kittler, J., Yu, D.J.: Cpinformer for efficient and robust compound-protein interaction prediction. IEEE/ACM transactions on computational biology and bioinformatics **20**(1), 285–296 (2022)
11. Hua, Y., Song, X., Feng, Z., Wu, X.: Mfr-dta: a multi-functional and robust model for predicting drug–target binding affinity and region. Bioinformatics **39**(2), btad056 (2023)
12. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708 (2017)
13. Knox, C., Wilson, M., Klinger, C.M., Franklin, M., Oler, E., Wilson, A., Pon, A., Cox, J., Chin, N.E., Strawbridge, S.A., et al.: Drugbank 6.0: the drugbank knowledgebase for 2024. Nucleic acids research **52**(D1), D1265–D1275 (2024)
14. Lee, I., Keum, J., Nam, H.: Deepconv-dti: Prediction of drug-target interactions via deep learning with convolution on protein sequences. PLoS computational biology **15**(6), e1007129 (2019)
15. Li, X., Su, R., Liu, L.: M3st-dti: A multi-task learning model for drug-target interactions based on multi-modal features and multi-stage alignment. arXiv preprint arXiv:2510.12445 (2025)
16. Li, X., Zhang, G., Cui, H., Hou, S., Wang, S., Li, X., Chen, Y., Li, Z., Zhang, L.: Mcanet: A joint semantic segmentation framework of optical and sar images for land use classification. International Journal of Applied Earth Observation and Geoinformation **106**, 102638 (2022)
17. Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T.Y., Tegmark, M.: Kan: Kolmogorov-arnold networks. arXiv preprint arXiv:2404.19756 (2024)
18. Subramanian, M.K., Spencer, J.: Deep neural networks for in silico prediction of drug-target interactions: Advancements and future directions. Frontiers in Emerging Artificial Intelligence and Machine Learning **2**(03), 1–12 (2025)
19. Suruliandi, A., Idhaya, T., Raja, S.: Drug target interaction prediction using machine learning techniques–a review (2024)
20. Trott, O., Olson, A.J.: Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry **31**(2), 455–461 (2010)
21. Wang, S., Shan, P., Zhao, Y., Zuo, L.: Gandti: A multi-task neural network for drug-target interaction prediction. Computational biology and chemistry **92**, 107476 (2021)
22. Wu, L., Huang, Y., Tan, C., Gao, Z., Hu, B., Lin, H., Liu, Z., Li, S.Z.: Psc-cpi: Multi-scale protein sequence-structure contrasting for efficient and generalizable compound-protein interaction prediction. In: Proceedings of the AAAI conference on artificial intelligence. vol. 38, pp. 310–319 (2024)
23. Zhang, Z., Chen, L., Zhong, F., Wang, D., Jiang, J., Zhang, S., Jiang, H., Zheng, M., Li, X.: Graph neural network approaches for drug-target interactions. Current Opinion in Structural Biology **73**, 102327 (2022)
24. Zhao, Q., Zhao, H., Zheng, K., Wang, J.: Hyperattentiondti: improving drug–protein interaction prediction by sequence-based deep learning with attention mechanism. Bioinformatics **38**(3), 655–662 (2022)
25. Zhou, Y., Zhang, Y., Zhao, D., Yu, X., Shen, X., Zhou, Y., Wang, S., Qiu, Y., Chen, Y., Zhu, F.: Ttd: Therapeutic target database describing target druggability information. Nucleic acids research **52**(D1), D1465–D1477 (2024)