Molecular Diversity
https://doi.org/10.1007/s11030-025-11303-6

REVIEW

A comprehensive review of neural network-based approaches for drug–target interaction prediction

Fatemeh Panahandeh[1] · **Najme Mansouri**[1]

Received: 7 May 2025 / Accepted: 16 July 2025

Abstract
Predicting Drug–Target Interactions (DTI) is vital for accelerating drug discovery and repurposing. This review assesses the efficacy of neural network-based methods, including Convolutional Neural Networks (CNNs), Graph Neural Networks (GNNs), and Transformer architectures utilizing self-attention mechanisms, in predicting DTI. Leveraging recent studies and benchmark datasets, such as DrugBank, neural networks outperform traditional methods, with hybrid architectures achieving an Area Under the Receiver-Operating Characteristic Curve (AUROC) of 0.979. These models enhance precision and reduce errors by integrating multimodal data, such as molecular graphs and protein sequences. However, challenges like high computational costs, limited interpretability, and data scarcity persist. Future advances may combine multimodal data with explainable Artificial Intelligence (AI) to improve DTI analysis and unlock new therapeutic opportunities. This study highlights the transformative potential of neural networks in computational drug discovery.

Keywords Drug–Target Interaction (DTI) · Neural networks · Deep neural networks · CNN · GNN · Transformers · Explainable AI

Introduction

Predicting Drug–Target Interactions (DTIs) involves identifying and analyzing relationships between drug molecules and biological targets, such as proteins, DNA, and RNA. These interactions are often responsible for the pathophysiology of diseases and are essential for cellular processes. The accuracy of DTI predictions facilitates novel therapeutic discoveries and drug repurposing, reducing both time and cost compared to experimental methods. Early phases of drug discovery have been shortened using computational approaches, as seen in efforts targeting cancer pathways.

For DTI prediction, molecular docking and molecular dynamics simulations have been widely used for a long time. In molecular docking, optimal binding positions and energies are evaluated, whereas in molecular dynamics, structural changes and stability are examined. Although these methods are exact, they require high-quality structural data and incur high computational costs, limiting their scalability. DTI prediction has been transformed by Machine Learning (ML) and Neural Networks (NN) in the past decade. By handling vast and heterogeneous datasets, these methods uncover non-linear relationships that are difficult to model using traditional approaches. The use of Graph Neural Networks (GNNs) can improve the prediction of molecular structures. At the same time, Transformer architectures—a specific class of neural networks utilizing self-attention mechanisms—have proven effective in sequence-based predictions, such as modeling protein or SMILES sequences [1].

There is growing evidence that neural networks can address these challenges. Several innovative models have emerged to enhance prediction performance and generalizability, including OverfitDTI [2] (which eliminates overfitting), FRnet [3] (which improves feature extraction), and HGDTI [4] (which enhances accuracy through heterogeneous graphs). Despite these advancements, comprehensive reviews are scarce, leaving gaps in our understanding of the field. Using neural network-based DTI prediction methods, this study aims to fill this gap.

✉ Najme Mansouri
n.mansouri@uk.ac.ir

Fatemeh Panahandeh
F.panah@math.uk.ac.ir

[1] Department of Computer Science, Shahid Bahonar University of Kerman, Box No. 76135-133, Kerman, Iran

Published online: 26 July 2025

In Fig. 1, the article's flow is shown from the introduction to future directions. The "Introduction" section discusses the current methods and neural network-based techniques for predicting Drug–Target Interactions (DTIs), followed by a Background (the "Background" section) of the article. The "Related survey" section provides a literature review, and the "Research methodology" section describes the research method, covering data selection, search processes, and study selection criteria. Several neural network models are presented for DTI prediction in the "Review of neural network models in drug-target interaction prediction" section, including the Hybrid, CNN, Transformers, GNN, and Deep Neural Network (DNN) models. In the "Analysis and comparisons" section, the models are compared and analyzed, followed by Conclusions (the "Future directions and open issues" section) and Future Directions and Open Issues ("Conclusions" section). In addition to providing insight into future developments, the logical flow offers a detailed review of neural network-based DTI prediction methods.

Background

Table 1 shows the standard abbreviations of this article.

Overview of existing methods for DTI prediction

DTI prediction plays a crucial role in drug discovery, drug repurposing, and molecular understanding of biological processes. With the exponential growth of biological data, such as protein sequences, molecular structures, high-throughput compound libraries, and multi-omics profiles, computational DTI prediction has become increasingly necessary. In addition to identifying potential drug candidates, these methods also enable researchers to repurpose existing drugs for new therapeutic indications and understand how biological interactions work. Traditional DTI methods rely on structural and chemical data. In contrast, machine learning-based methods utilize statistical patterns in large datasets, while neural network-based methods model complex, non-linear relationships. Hybrid methods, on the other hand, overcome the limitations of the approaches above. A hierarchical overview

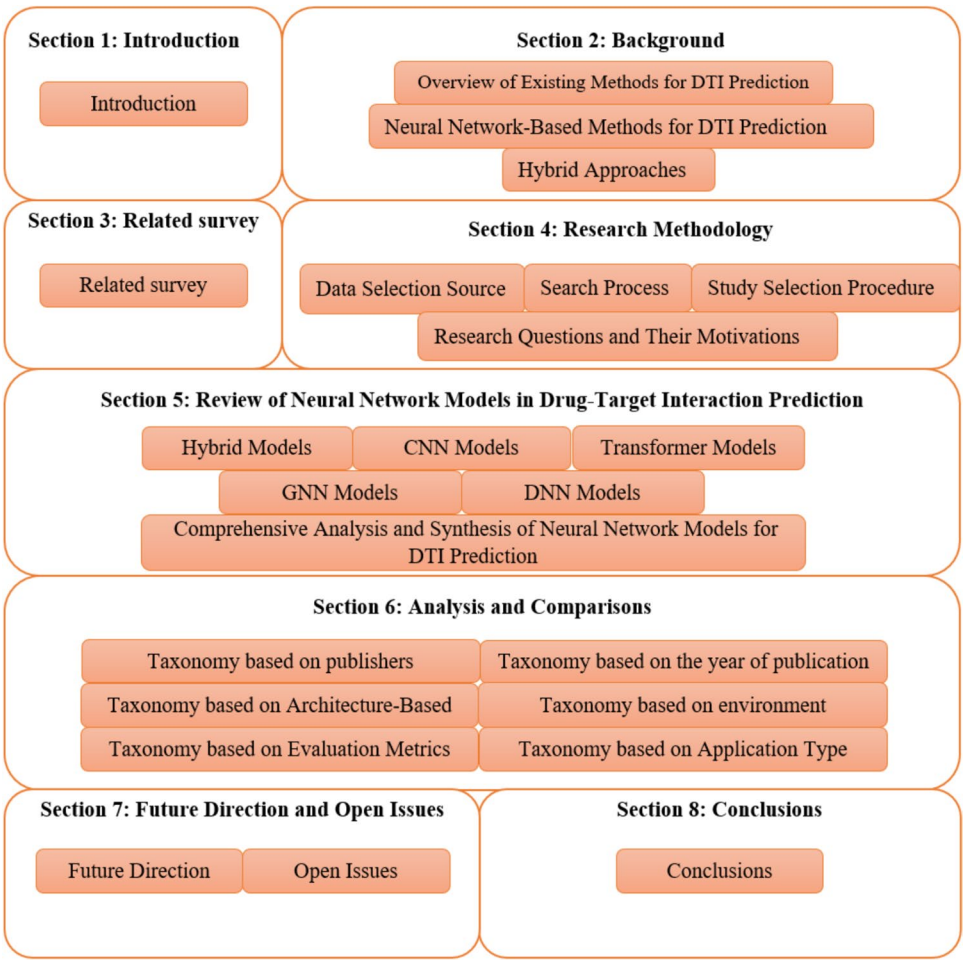


Fig. 1 Article structure

Table 1 Abbreviation

Abbreviation	Definition	Abbreviation	Definition
DTI	Drug–target interaction	DDI	Drug–Drug Interaction
CNN	Convolutional neural network	CADD	Computer-Aided Drug Design
GNN	Graph neural network	RF	Random Forest
RNN	Recurrent neural network	k-NN	k-Nearest Neighbors
LSTM	Long short-term memory	GCN	Graph Convolutional Network Graph Convolutional Network
DNN	Deep neural network	IFSM	Improved Frequent Subsequence Mining
MLP	Multi-layer perceptron	OMBO	Opposition-based Monarch Butterfly Optimization
MD	Molecular dynamics	TBAN	Text CNN-BiLSTM-Attention Network
ML	Machine learning	MSA	Multi-Head Self-Attention
AI	Artificial intelligence	MCA	Multi-head Cross Attention
XAI	Explainable artificial intelligence	MHRW	Metropolis–Hastings Random Walk
AUROC	Area under the receiver-operating characteristic curve	MKL	Multi-Kernel Learning
AUPR	Area under the precision–recall curve	NT-Xent	Normalized Temperature-scaled Cross Entropy Loss
SMILES	Simplified molecular input line entry system	GRU	Gated Recurrent Unit
QSAR	Quantitative structure–activity relationship	ReLU	Rectified Linear Unit
GPCR	G-protein-coupled receptor	ECFP	Extended-Connectivity Fingerprints
PPI	Protein–protein interaction	NMR	Nuclear Magnetic Resonance
GAEMSDNN	Graph auto-encoder and multi-subspace deep neural networks	NN	Neural Networks
VR	Virtual reality	(GAN)	Generative Adversarial Networks
EHR	Electronic health records		

of traditional and hybrid methods is shown in Fig. 2. In the "Traditional methods" to "Hybrid approaches" sections, we explore each category in detail, discussing its principles, applications, strengths, and limitations [1, 5].

Traditional methods

In traditional DTI prediction methods, structural and chemical data are used to simulate drug–target interactions. The value of these methods is especially apparent during the early stages of drug discovery, when mechanistic insights into binding are crucial for advancing the process. There are two main subgroups of these methods: structure-based and ligand-based.

Structure-Based Methods: In these methods, three-dimensional (3D) molecular models are used to simulate drug–target interactions at the atomic level, thereby identifying binding sites and predicting interaction dynamics. Structure-based methods use molecular docking to predict the optimal binding positions and energies for small molecules (ligands) by placing them into the active sites of biological targets, such as enzymes and receptors. Tools like AutoDock, GOLD, and Glide evaluate interactions based on factors like hydrogen bonds, van der Waals forces, hydrophobic effects, and steric complementarity. During the COVID-19 pandemic, molecular docking played a crucial role in the design of nirmatrelvir, an inhibitor of the SARS-CoV-2 main protease. Additionally, docking has been essential in oncology for identifying inhibitors of cancer-related proteins, such as EGFR, which has paved the way for precision medicine treatments. However, docking does not fully capture the dynamic nature of biological systems, since it offers a static view of interactions. Molecular Dynamics simulations (MD) provide a complementary approach by modeling drug–target interactions in a time-dependent manner. Through MD simulations, atomic movements in a binding pocket are tracked over time, conformational changes are assessed, and solvent molecules play a role. In studies of kinase inhibitors used for treating chronic myeloid leukemia, MD simulations have revealed how structural changes in the target protein can significantly change binding affinity, providing clues about drug resistance. Force-field models, such as GROMACS and AMBER, enable nanosecond-to-microsecond simulations. In MD simulations of HIV protease inhibitors, structural adaptations in the binding pocket enhanced drug efficacy by stabilizing the drug–target complex [1, 6].

- *Advantages:* X-ray crystallography, Nuclear Magnetic Resonance (NMR), and cryo-electron microscopy provide high-resolution structural data, making structure-based methods highly effective. Scientists can use these techniques to investigate interactions at the atomic level,

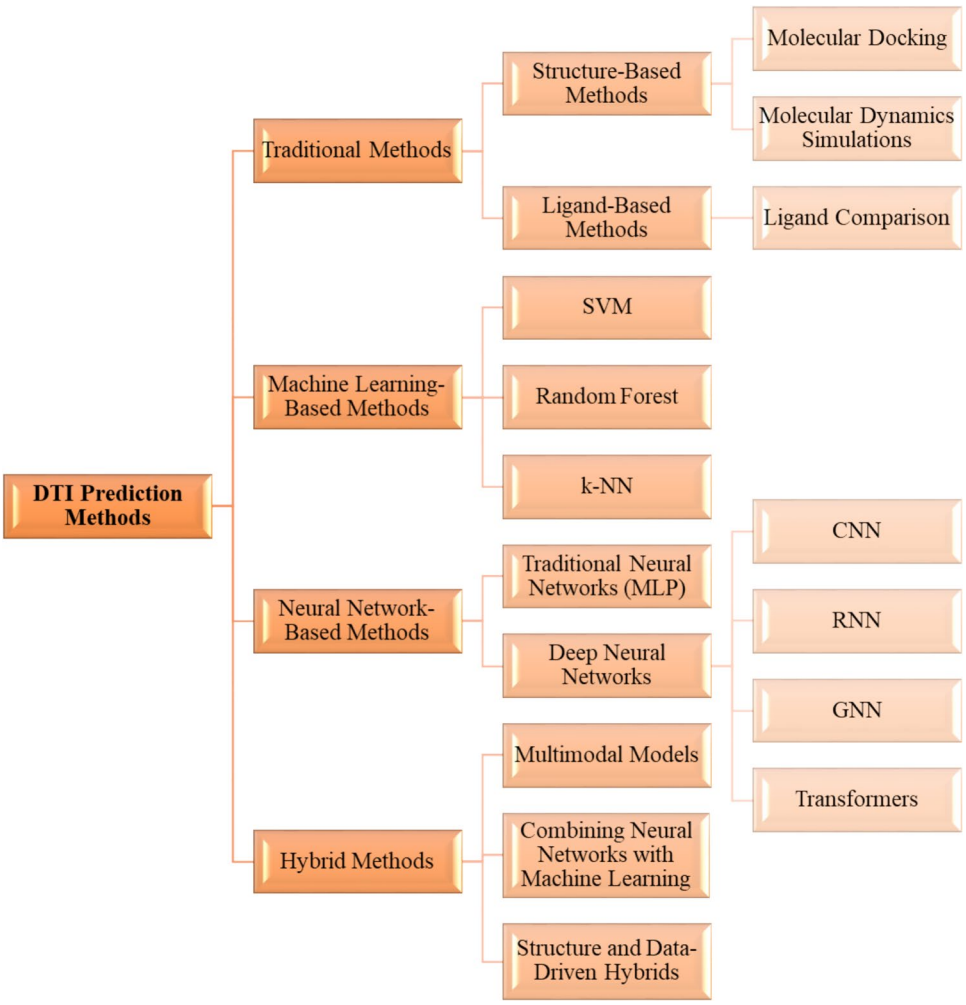


Fig. 2 Structure of DTI methods

test hypotheses, and rationally design targeted therapies. In the case of HIV protease, X-ray crystallography enabled the development of potent antiviral drugs, such as saquinavir, by elucidating its three-dimensional structure. They also provide valuable insights into drug–target interactions based on mechanistic insights.

- *Disadvantages:* A significant limitation. Due to their insoluble nature and structural flexibility, membrane proteins, such as G-protein-coupled receptors (GPCRs) and ion channels, are notoriously difficult to crystallize. These methods are further restricted in their application to novel targets due to experimental challenges in crystallization and solubilization. The cost of docking and MD is also high. Even with high-performance computing clusters, it can take quite some time to complete a single MD simulation for large compound libraries with millions of molecules, making these methods unsuitable for high-throughput applications.

Ligand-Based Methods: Compared to structure-based methods, ligand-based approaches rely on chemical similarity between known ligands to predict bioactivity. Molecular structures with similar biological activities tend to exhibit similar properties. This category of study is dominated by Quantitative Structure–Activity Relationships (QSAR), which uses mathematical models to link chemical properties (such as molecular weight, polarity, hydrophobicity, and electronic properties) to biological activity. Another key method is pharmacophore mapping, which identifies shared structural features (e.g., hydrogen-bond donors, acceptors, or hydrophobic regions) among active ligands. Descriptors like Extended-Connectivity Fingerprints (ECFPs) and MACCS keys capture molecular topology and functional groups in these methods. Ligand-based approaches are beneficial for discovering new drugs targeting GPCRs and ion channels, where structural data may be limited. In antidepressant development, QSAR has identified structural similarities among serotonin reuptake inhibitors, leading to the creation of improved compounds with enhanced efficacy. Due to its chemical similarity to known active compounds, thalidomide was repurposed for the treatment of multiple myeloma using ligand-based

methods. Additionally, these methods have enabled early stage discovery of targets for neglected tropical diseases, where existing information is scarce [1, 7].

Ligand-based approaches enhance bioactivity prediction by relying on chemical similarities between known ligands, rather than requiring target structure information. Techniques such as Quantitative Structure–Activity Relationship (QSAR) and pharmacophore mapping link structural and physicochemical features (e.g., hydrogen-bond donors, acceptors, and hydrophobic regions) to biological activity. These features are encoded using descriptors like Extended-Connectivity Fingerprints (ECFPs) and MACCS keys, which capture molecular topology. Ligand-based approaches are particularly valuable in early stage screening, especially for targets such as GPCRs and ion channels, where structural data are limited. They have been applied in antidepressant development and drug repurposing—such as identifying thalidomide as a treatment for multiple myeloma—and in discovering treatments for neglected tropical diseases where available data are sparse [1, 7].

- *Advantages:* Ligand-based approaches are less computationally intensive than structure-based approaches, making them ideal for rapid screening of large compound libraries in the early stages of drug discovery. Additionally, their relative simplicity makes them accessible to research groups with modest computational resources, allowing them to be applied in resource-constrained settings.
- *Disadvantages:* For ligand-based methods to perform effectively, existing ligand datasets must be readily available and of high quality. Data are often sparse and of low quality for poorly studied targets or rare diseases (e.g., orphan diseases). Additionally, these methods struggle to extrapolate beyond known chemical spaces, which limits their ability to predict novel interactions for targets with limited prior data. It makes them less effective for discovering drug–target pairs outside of well-characterized biological systems.

Machine learning-based methods

In DTI prediction, Machine Learning (ML) has replaced labor-intensive, structure-dependent methods with scalable, data-driven strategies. With machine learning, large and diverse datasets are leveraged to achieve fast and accurate outcomes. Often, features are extracted automatically, which enhances the model's efficiency. It is typically possible for these algorithms to generalize to new, previously unknown combinations of drugs and targets, as they are trained on labeled datasets of known drug–target pairs. For high-throughput applications, machine learning has become a cornerstone of modern drug discovery [5, 7].

Common algorithms for DTI prediction *Support Vector Machines (SVMs): ML methods, such as SVMs, were among the first to predict DTI by constructing hyperplanes that distinguish between drug–target pairs that interact and those that do not*. For drugs, SVM models often utilize molecular descriptors (e.g., logP, molecular weight, and polar surface area), while for targets, amino acid compositions are employed. It has been demonstrated that SVMs are highly effective in virtual screening, particularly when combined with data from databases such as ChEMBL. In GPCRs, a prominent drug target family, SVM-driven models have been successful in identifying novel ligands by distinguishing between active and inactive compounds. Their ability to handle noisy data makes them particularly useful in early stage discovery, where datasets may be incomplete or contain errors [5, 7].

- *Advantages:* DTI prediction tasks frequently involve high-dimensional and sparse datasets, which SVMs excel at. Due to their mathematical formulation, the boundaries of decisions are clearly defined through optimal hyperplanes, resulting in high classification accuracy. Furthermore, SVMs are highly flexible, because they can incorporate different kernels (e.g., linear, radial basis functions) without requiring extensive parameter tuning. Due to this flexibility and strong generalization ability, SVMs are well suited for early stage virtual screening with limited labeled data.
- *Disadvantages*: Although SVMs have strengths, they often face challenges in terms of scalability and computational efficiency, particularly with large datasets. This is because training time increases dramatically as data size increases. Their interpretability is also limited, especially when non-linear kernels are used, making it difficult to understand the biological basis of predictions. Moreover, SVMs are sensitive to parameter selection (e.g., the regularization parameter C, kernel type, and gamma), which can have a significant impact on performance. Additionally, SVMs are not inherently capable of providing probabilistic outputs, which limits their application in drug discovery scenarios [6].

Random Forest (RF): By constructing multiple decision trees based on random subsets of features—such as Morgan fingerprints for drugs or protein domain profiles for targets—RF improves accuracy and prevents overfitting by combining their predictions by majority voting or averaging. In kinase inhibitor screening, RF has been beneficial for ranking lead compounds in large compound libraries, which significantly reduces experimentation costs. In the development of kinase-targeted therapies, RF models prioritized high-potential candidates for experimental validation, thereby streamlining the drug discovery process [5, 7].

- *Advantages:* Its ensemble nature makes it robust to data variability and overfitting, reducing sensitivity to both. Its strengths include the ability to assess feature importance, which helps identify which molecular or biological features are most influential in DTI prediction (e.g., specific functional groups or protein domains). Real-world biomedical datasets with noise or incomplete information can also benefit from RF's ability to handle missing or incomplete data without requiring extensive processing.
- *Disadvantages:* There is a possibility that RF may be computationally intensive, especially for massive datasets or when computing resources are limited. Its complex internal structure—comprising many decision trees—makes it difficult to understand, which is critical in applications requiring biological reasoning, such as clinical validation. Additionally, hyperparameter choices, such as the number of trees or maximum depth, can impact RF performance [6].

k-Nearest Neighbors (k-NN): This is a similarity-based method that compares the chemical structures and numerical features of target and drug pairs with their nearest neighbors in a feature space to predict interactions. Drug repurposing, which transfers bioactivity from known targets to new ones based on chemical similarity, has been particularly beneficial. To identify candidates for antimalarial properties in the existing drugs, k-NN was used to compare their chemical profiles with those of known antimalarial compounds. In this way, researchers have been able to repurpose drugs for new indications with minimal experimental effort [1, 5].

- *Advantages*: It is a non-parametric and intuitive algorithm that does not make assumptions about data distribution, making it flexible for situations in which the relationships between variables are complex or poorly understood. Adaptable to different data types, it has become an integral part of early stage drug discovery platforms due to its similarity metrics. In addition to Tanimoto and Euclidean distance functions, researchers can select from a variety of data types.
- *Disadvantages:* The k-NN approach is computationally inefficient during inference due to the need to calculate distances to every instance in the dataset. Furthermore, it is susceptible to noisy or outlier data. Moreover, it is challenging to select the optimal value of k, as a small value may result in instability due to noise, while a considerable value may dilute local patterns.

Limitations of traditional and ML methods

In the context of modern drug discovery, traditional and ML-based approaches to drug discovery face significant limitations. Due to challenges in crystallization and solubilization, traditional structure-based methods, such as molecular docking and MD simulations, are heavily dependent on accurate 3D experimental structures. They are not suitable for screening large libraries of compounds at high throughput due to their high computational costs. It is often challenging to predict novel interactions outside the known chemical space using ligand-based methods, since they do not require target structures. In contrast to traditional approaches, ML methods, such as SVM, RF, and k-NN, offer improved scalability and speed but rely on hand-engineered features (e.g., molecular fingerprints, physicochemical properties). It is often hard to identify allosteric effects or complex molecular graph relationships in DTI data using these human-defined features. Furthermore, ML methods require extensive preprocessing to deal with unstructured and high-dimensional data, which can introduce bias. Modern datasets are complex, encompassing millions of compounds from high-throughput screening and multi-omics profiling, exceeding the ability of traditional and ML methods. In "Neural Network-Based Methods for DTI Prediction" section, we discuss how neural network-based methods excel at extracting complex patterns from raw, diverse data [1, 5].

Neural network-based methods for DTI prediction

A neural network is a computer model based on the structure and function of the human brain. It consists of layers of nodes (neurons) connected, with each node taking inputs, processing them, and passing on the result to the next layer through weighted links. Neural networks learn from examples by adjusting these weights during training.

Neural networks represent a broad family of architectures, ranging from shallow architectures, such as Multi-Layer Perceptrons (MLPs), to Deep Neural Networks (DNNs), including specialized architectures like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Graph Neural Networks (GNNs), and Transformer architectures. Although shallow networks typically rely on comparatively small numbers of hidden layers and are suitable for structured inputs, deep neural networks can automatically learn hierarchical, high-dimensional features from unprocessed data. Such a capability renders them extremely potent in predicting Drug–Target Interactions (DTIs). Here, we present an overview of both standard (shallow) and deep neural network-based approaches for DTI modeling [8, 9].

Traditional neural networks (MLPs)

Multi-Layer Perceptrons (MLPs) are the simplest of neural networks (NNs), consisting of several layers of interlinked neurons. The MLP model processes inputs, such as molecular weight, logP, and protein descriptors through hidden layers by applying activation functions (e.g., Rectified Linear

Unit (ReLU) and sigmoid), and predicts drug–target interactions [9].

- Applications: MLPs were initially used in initial research to forecast DTI for simple, well-defined tasks. To show how NNs could be used to forecast kinase inhibitor-target interactions, they utilized precalculated ligand descriptors. Small experiments or controlled conditions with well-structured data can utilize them computationally at low cost.
- Example: Initially, MLPs were used to make interaction predictions for kinase inhibitors based on precomputed ligand descriptors. More advanced NN architectures will need to be created that can scale to larger, more diverse datasets.

Deep neural networks (DNNs)

Deep Neural Networks (DNNs) extend Multi-Layer Perceptrons (MLPs) by incorporating multiple hidden layers, enabling them to learn hierarchical features directly from raw data. Modern Drug–Target Interaction (DTI) prediction heavily relies on DNNs, with specialized architectures designed for specific data types, as they can extract complex patterns without requiring manual feature engineering [8, 9].

Convolutional Neural Networks (CNNs): The convolutional network perceives grid-based data, such as 2D molecular images or 1D protein sequences, through learning local patterns. A CNN is particularly suited to recognizing chemical substructures in drug representations (e.g., functional groups) and conserved motifs in protein sequences (e.g., binding domains). These are the essence of predicting interactions in DTIs. Due to their capacity for dimensionality reduction through pooling across layers, they are well suited for large-scale screening. CNNs have been used to predict drug–target binding affinities, due to the utilization of Simplified Molecular Input Line Entry System (SMILES) strings (textual representations of molecule structures) [10].

Recurrent Neural Networks (RNNs): The usage of RNNs, in their more recent forms such as Long Short-Term Memory (LSTM) units, is best suited for sequence data, including protein sequences or SMILES strings. The algorithms process sequentially, maintaining long-range relationships and time correlations between amino acids or molecule fragments. They are, therefore, highly effective in modeling dynamic drug–target interactions where the position in the sequence is critical to binding affinity. For predicting protein–ligand interactions, RNNs are used to describe the sequence of amino acid residues in a target protein and identify patterns that define binding specificity. Sequential processing of extremely long sequences is computationally expensive and limits their scalability [6, 8, 11].

Graph Neural Networks (GNNs): In a molecular graph, atoms are nodes and bonds are edges. GNNs are designed to process graph-structured data. To acquire structural interaction between the target and the drug, GNNs integrate topological and spatial information. To predict interactions, they are capable of modeling molecular graphs directly from raw data. Because they learn embeddings of molecular graphs that capture local and global structural information, GNNs are superior to traditional approaches in binding affinity prediction [12, 13].

Transformers: A family of deep neural networks based on self-attention, Transformers are primarily famous for capturing long-range dependencies in graph-structured and sequential data. They outperform RNNs as they handle data in parallel and, therefore, are more suitable for large datasets. Transformers can encode both molecular graphs and protein sequences simultaneously, achieving better DTI prediction performance by learning sophisticated cross-modal relations. With its joint domain adaptation and cross-attention, CAT-DTI [14] achieves an AUROC of 0.983 and an Area Under the Precision–Recall Curve (AUPR) of 0.976, performing well in and across domains [15].

Application of DNNs DNN application to DTI prediction is widespread. OverfitDTI [2] is a DNN-based framework designed to overfit training sets and capture the chemical and biological drug–target space. It integrates a Graph Auto-Encoder with Multi-Subspace Deep Neural Networks (GAEMSDNN) and achieves a concordance index of 0.9480 [11]. DNNs are flexible in managing intricate prediction scenarios, ranging from dynamic interactions to cold-start issues [9].

Table 2 synthesizes the major strengths and limitations of the preceding neural network architectures. The emphasis here is placed on the trade-off between model complexity, computational cost, interpretability, and accuracy with different types of DTI data.

Hybrid approaches

To address the weaknesses of traditional, ML, and NN models, new hybrid approaches have been suggested that combine their advantages into an even stronger and more general system. In computational drug discovery, various architectures, algorithms, and data sources are combined to enhance predictive performance and meet the dual requirements of accuracy and interpretability [6, 8].

Multimodal Models: Multimodal models integrate various sources of data to enhance the accuracy of prediction. Bidirectional multi-head attention, for instance, is employed by the MCL-DTI [16] algorithm to integrate molecular images and chemical features. Multimodal fusion is well suited for MCL-DTI by utilizing CNNs for visual patterns

Table 2 Advantages and disadvantages of neural network models

Model type	Advantages	Disadvantages
MLP	Low-cost and straightforward structure; suitable for structured features and resource-limited settings [8, 9]	Highly dependent on input feature quality; weak for complex or raw data like sequences and graphs [8, 9]
CNN	Strong local feature extraction (e.g., functional groups or protective motifs); faster and scalable with pooling [3, 6, 10]	Poor at modeling long-range dependencies and non-local interactions; requires careful preprocessing; low interpretability [3, 6, 10]
RNN	Models sequential relationships and retains long-term information, effective for sequence and SMILES data [6, 8, 11]	Computationally expensive and slow; vanishing/exploding gradient issues; sensitive to noisy/incomplete data [6, 8, 11]
GNN	Direct modeling of molecular graph structures; suitable for complex molecule–protein interactions [4, 6, 12, 13]	Requires precise structural data, high computational cost, limited interpretability, and complex parameter tuning [4, 6, 12, 13]
Transformers	Learns long-range and non-local relationships via attention, parallel processing; integrates sequence and structural data [6, 14, 15]	Needs large datasets; complex design; prone to overfitting with small data; computationally expensive and low interpretability [6, 14, 15]

and neural networks for integrating chemical data. Drug–target interactions are more comprehensively analyzed through multiple representations of data (e.g., structural, sequential, and chemical) [6].

Combining Neural Networks with Machine Learning: In some of the hybrid models, NNs and ML approaches are combined to leverage their strengths. For instance, HGDTI [4] integrates CNNs with GNNs within a heterogeneous graph framework, achieving an AUROC of 0.979 and an AUPR of 0.961. In addition to improving robustness and generalizability across datasets like DrugBank, HGDTI [4] promotes improved generalizability by learning node embeddings from protein and molecular data. As such, the model can identify both local patterns (via CNNs) and topological relations (via GNNs). Meanwhile, ML can also be utilized to post-process predictions or carry out tasks such as uncertainty estimation or feature selection [6].

Structure and Data-Driven Hybrids: Hybrid approaches leverage both precise structural information and high-scale patterns by combining traditional structure-based methods with data-driven NNs. Molecular docking, for example, calculates binding energies and structural descriptors that serve as input to an NN for pattern matching. In this approach, docking is combined with the predictive power of NNs to yield more interpretable predictions with higher accuracy. NN-based predictions have been employed to study protein–ligand interactions in drug discovery [1, 6]. Table 3 summarizes the advantages and disadvantages of hybrid models.

Advantages of hybrid approaches

There are several key benefits to hybrid approaches. Due to its integrated architecture, HGDTI [4] outperforms standalone GNNs, achieving higher accuracy. Additionally, they can adapt to a wide range of data types (e.g., sequences, graphs, and images) and support various applications, including drug discovery and repurposing. Additionally, multimodal models provide a more balanced approach to DTI prediction by combining knowledge-based (e.g., structural data) and data-driven (e.g., NN-based) methods [4, 6].

Challenges of hybrid approaches

Hybrid approaches present several challenges. In addition to computational complexity, integrating multiple architectures increases resource demands. Due to its complex graph-based framework, HGDTI [4] requires advanced hardware and extended training periods. Additionally, the complexity of the design necessitates significant expertise and time to select and tune components (e.g., CNNs, GNNs, and attention mechanisms). Additionally, interpretability remains problematic; although some models, such as FragXsiteDTI

Table 3 Advantages and disadvantages of hybrid models

Model type	Advantages	Disadvantages
Multimodal models	Integrates diverse data types (images, chemical features, sequences) for higher accuracy; achieves AUROC 0.987; effective multimodal fusion with CNN and deep neural network architectures [6, 16]	High computational cost due to processing multiple data types; requires high-quality, consistent data; limited interpretability due to black-box nature [6]
Combining neural networks with machine learning	Combines CNN and GNN to capture local and topological features; improves robustness and generalizability; achieves AUROC 0.979, AUPR 0.961 [4, 6]	Computationally intensive; requires advanced hardware; sensitive to noisy/incomplete data; needs expertise for model design and parameter tuning [4, 6]
Structure and data-driven hybrids	Combines molecular docking with NN for interpretable and accurate predictions; integrates mechanistic insights with data-driven learning [1, 6]	Limited availability of high-quality structural data; increased computational cost; complex integration design; limited interpretability due to black-box nature [6]

[17], emphasize drug fragments and protein pockets, most hybrid approaches retain the black-box nature of NNs, making it difficult to understand the biological rationale behind predictions. Lack of transparency limits their adoption in clinical settings, where interpretability is crucial for validation [6].

Related survey

The purpose of this section is to provide an overview of recent surveys and studies related to computational methods for predicting Drug–Target Interaction (DTI), with a focus on neural network-based approaches. It identifies key trends, contextualizes advancements in DTI prediction, and highlights gaps that are subsequently addressed. Tables 4, 5, and 6 present a comparative analysis of the studies.

Zuo et al. [1] noted the importance of identifying Drug–Target Interactions (DTI) for drug discovery and repositioning; however, high experimental validation costs limited their effectiveness. They found computational methods to be cost-effective. Feature encodings and databases for DTI prediction were summarized in this review. In addition to highlighting their performance across diverse datasets, it also evaluated the strengths and weaknesses of prediction models from the past 5 years. To improve prediction accuracy and efficiency in future DTI research, advanced computational technologies, such as quantum computing and multi-omics data integration, must be utilized.

Shi et al. [5] highlighted the challenges of traditional experimental methods and the importance of predicting drug–target interactions. Moreover, they categorized a variety of machine learning-based approaches using a new taxonomy for deep neural networks. Furthermore, these models are often challenging to interpret due to the presence of imbalanced datasets. To enhance understanding and effectiveness in DTI prediction, the authors advocate for the use of semi-supervised learning, transfer learning, and interpretable models.

Alizadehsani et al. [6] discussed how Artificial Intelligence (AI) and Machine Learning (ML) transformed drug discovery, with transparency and interpretability being essential. Explainable Artificial Intelligence (XAI) was developed to enhance the understanding of machine learning models and their predictions. XAI techniques received increased attention in drug discovery over the past few years, and this paper presented a review of XAI in the context of drug discovery. Additionally, they discussed the application of XAI in target identification, compound design, and toxicity prediction in the context of drug discovery.

Atasever [7] noted that this systematic review focused on virtual screening and machine learning. ML, molecular docking, bioinformatics, and cheminformatics improved

Table 4 Recent DTI studies

Article	Year	Study area	Main focus	Author	Title	Keywords
[1]	2024	Drug–target interaction prediction	Reviewing recent advancements in computational DTI prediction methods, databases, and potential directions in the last 5 years	Zuo et al.	Research Progress on Drug–Target Interactions in the Last 5 Years	Drug–target interaction, Machine learning, Network representation learning, Graph neural networks
[5]	2024	Drug–target interaction prediction	A survey of machine learning-based approaches for DTI prediction, with an emphasis on data representation and model taxonomy	Shi et al.	A survey of machine learning-based approaches for drug–target interaction prediction	Drug–target Interaction, Machine learning, Data representation, Deep neural network models
[6]	2024	Explainable AI for drug discovery and development	Presenting an overall overview of the applications, challenges, and opportunities of Explainable AI (XAI) in drug discovery and drug development	Alizadehsani et al.	Explainable Artificial Intelligence for Drug Discovery and Development: A Comprehensive Survey	Explainable artificial intelligence, Drug discovery, Machine learning, Big data
[7]	2024	Drug discovery	Reviewing ML-based in silico methods in drug discovery, focusing on virtual screening and computational tools	Atasever	In silico drug discovery: a machine learning-driven systematic review	Machine Learning, Drug Discovery, In Silico Methods, Systematic Review, PRISMA
[8]	2023	Drug discovery	Reviewing the applications of deep neural networks in drug discovery, including drug design, recommendation, and response prediction	Li et al.	A compact review of the progress and prospects of deep neural networks in drug discovery	Deep Neural Network, Drug Discovery, Drug Repositioning, Drug Design, Drug Reaction
[12]	2023	Drug–Target interaction prediction	Reviewing methods for DTI and DTA prediction using Graph Neural Networks (GNNs)	Zhang et al.	A survey of drug–target interaction and affinity prediction methods via graph neural networks	Drug–target Interaction, Drug–target affinity, Graph neural networks, deep neural network
[18]	2024	Drug–drug interactions prediction using deep neural network and knowledge graph	Deep neural network and knowledge graph-based methods for DDI prediction: review, approach classification, and challenges discussion	Luo et al.	Deep neural network and knowledge graph-based drug–drug interactions prediction: A review	Drug-drug interactions, Deep neural network, Knowledge graph
[19]	2024	Computer-aided drug design and machine learning in anesthetic drug discovery	Summary of the computational techniques in anesthetic development, including the uses of ML and DNN in anesthetic drug safety improvement and efficacy improvement	Liu et al.	Computer-aided drug design and machine learning in anesthetic drug discovery	Computer-aided drug design, Machine learning, Anesthetic, Drug discovery
[20]	2022	Drug–target interaction	Reviewing AI-based techniques and data formats for predicting Drug–Target Interactions (DTI) and ligand–protein interactions (LPI)	Liyaqat & Ahmad	A brief review on Artificial Intelligence-based Drug Target Interaction Prediction	Artificial Intelligence, Drug Discovery, Drug–Target Interactions

Table 5 DTI study comparison

Article	Highest order evaluation metric	Comparative analysis	Taxonomy	Presenting a comparative chart	Set up environment details	Datasets analysis	Motivation or open issues
[1]	-	-	-	-	×	-	-
[5]	-				×		
[6]	-				×		
[7]	-				×		
[8]	-				×		
[12]	-				-		
[18]	-				×		
[19]	-				×		
[20]	-				×		

drug development efficiency. Tools such as ChEMBL and the random forest algorithm were introduced. It was found that ML could accelerate drug discovery, and AI-based approaches could transform therapeutic drug development.

Li et al. [8] discussed the role of deep neural networks in enhancing the efficiency of drug discovery. They explored the limitations of traditional methods, such as virtual screening, due to the complexity of the data, which limits their effectiveness. Transfer learning helped mitigate challenges from data scarcity in drug target discovery and new drug design. For drug discovery to be effective, deep neural networks—particularly RNNs, CNNs, and GNNs—must be appropriately selected in terms of architecture and feature representation.

Zhang et al. [12] recognized Graph Neural Networks (GNNs) as a crucial tool for predicting Drug–Target Interactions (DTIs) and affinity (DTAs). In these approaches, molecular graphs and protein sequences were modeled to extract structural and functional information. This approach used contact maps to represent proteins and model multimodal features of drugs. DTI and DTA predictions enhanced drug development efficiency despite challenges related to managing large datasets and selecting optimal network structures.

Huimin et al. [18] discussed the potential harmful effects of drug–drug interactions. Time-consuming and costly traditional detection methods necessitated the development of computational approaches. Drug–Drug Interaction (DDI) prediction methods utilizing deep neural networks and knowledge graphs are systematically categorized into three main classes: deep neural network-based, knowledge graph-based, and hybrid methods. Furthermore, asymmetric and high-order interactions were discussed as challenges in DDI prediction in the review.

Liu et al. [19] reviewed Computer-Aided Drug Design (CADD), which can reduce costs and time during the drug discovery process. The integration of machine learning and deep neural networks enhanced CADD's potential. In this article, computational methods are explored in the development of novel anesthetics, along with their advantages and limitations. This research was conducted to develop safer and more effective anesthetic drugs.

Liyaqat and Ahmad [20] examined ligand–protein interaction (LPI) prediction and the importance of technological advances in the drug discovery process. As a result of this prediction, millions of compounds could be screened, thereby speeding up drug discovery. In addition to addressing data representation limitations, continuous-valued parameters were also discussed as challenges associated with accurate LPI prediction. Additionally, they explored methods to enhance the accuracy of LPI predictions by utilizing effective representation techniques and deep neural network approaches.

Table 4 summarizes these studies, while Tables 5 and 6 provide comparative analyses of evaluation metrics, taxonomy, experimental setup, dataset analysis, open research issues, and additional aspects, such as publisher analysis.

Table 6 DTI evaluation aspects

Article	Allocation problem	Publisher analysis	SI algorithms analysis	Environment analysis	Evaluation tool analysis
[1]	×	×	-	-	-
[5]	-	×	×	×	×
[6]	-	-	-	×	×
[7]	×	-	-	×	×
[8]	-	-	-	-	-
[12]	-	×	-	-	-
[18]	-	×	×	×	×
[19]	×	-	×	×	×
[20]	-	×	-	×	×

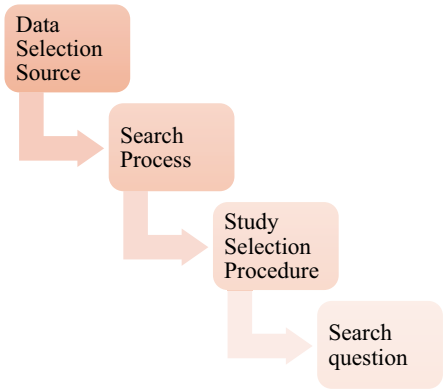


Fig. 3 Flowchart of research methodology

Table 7 Database sources

No	Source	URL
1	ScienceDirect	https://www.sciencedirect.com/
2	Springer Link	https://link.springer.com/
3	IEEEXplore	https://ieeexplore.ieee.org/

Research methodology

This research aims to study and evaluate neural network approaches for predicting Drug–Target Interactions (DTIs). Figure 3 illustrates the methodology, including data selection, search processes, and study selection. This systematic approach ensures the analysis of high-quality studies.

Data selection source

This study selected articles from three widely recognized and reliable scientific databases: ScienceDirect, IEEE Xplore, and SpringerLink. These databases were chosen for their extensive peer-reviewed content and relevance to DTI and neural network research, ensuring access to valid and current studies that form a robust foundation for analysis. Table 7 lists the URLs of these databases.

Search process

For comprehensive coverage of current research, a step-by-step approach was formulated to find suitable articles. As a first step, a list of keywords related to the research subject was compiled. The keywords were DTI, Drug–Target Interaction, Neural Network, Deep Neural Network, CNN, RNN, GNN, and Transformers. The terms were used to query selected scientific databases, including ScienceDirect, IEEE Xplore, and SpringerLink. The search was not time-limited, because the subject is relatively recent and has a small body of previous research. To narrow the results to the most critical, extraneous papers and non-English research were excluded. With the help of inclusion and exclusion criteria, a vast number of results were narrowed down. Additionally, the abstracts and keywords of the shortlisted articles were reviewed to select the most suitable studies.

Study selection procedure

Figure 4 illustrates a rigorous, structured process used to identify high-quality, relevant articles. The steps involved in this process were as follows:

- Initial screening:
- DTI-related and neural network-related titles and abstracts were excluded from the databases.
- Application of inclusion criteria:
- Articles were selected for further review if they met the following criteria:
- A clear focus on Drug–Target Interaction (DTI).
- Utilization of neural network methods.
- Published in reputable, peer-reviewed scientific journals.
- Application of exclusion criteria:
- Articles were excluded if they:
- Lacked sufficient methodological detail.
- Were duplicates of other articles.

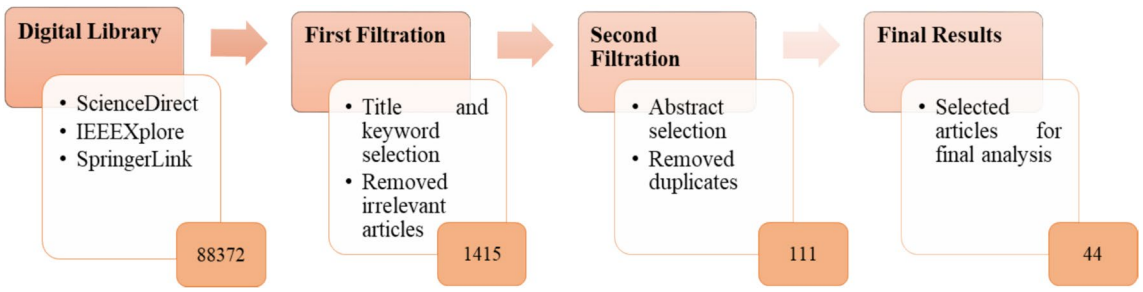


Fig. 4 Article selection process

- Had no full-text availability.
- Final review:
- Articles meeting all criteria were thoroughly reviewed and selected for analysis.

As shown in Fig. 4, the digital library initially contained 88,372 articles, which were reduced to 1,415 after the first filtration (removal of irrelevant articles), then to 111 after the second (removal of duplicate articles), and finally to 44 for detailed analysis. The systematic approach ensured that the collection of studies was balanced and representative of the overall research.

Research questions and their motivations

In this study, neural network methods will be employed to investigate the prediction of Drug–Target Interactions (DTIs). Table 8 presents key research questions that address different aspects of these challenges, problems, and opportunities.

Review of neural network models in drug–target interaction prediction

Specifically, this section addresses key research questions from "Research Questions and Their Motivations" section (e.g., Questions 1, 7, and 10 regarding popular architectures, performance comparisons, and limitations) regarding neural network-based methods for Drug–Target Interaction (DTI) prediction. As shown in Fig. 2, the review includes hybrid, Convolutional Neural Networks (CNN), Transformer architectures, Graph Neural Networks (GNN), and Deep Neural Networks (DNN) approaches. The models are evaluated based on performance metrics, application scope, and their ability to address challenges, such as dataset diversity, computational costs, and interpretability. Detailed analyses are supported by Tables 9, 10, 11, 12, 13, 14, 15, 16, 17, and 18, which summarize methods, databases, and comparative insights, and identify future research opportunities.

Hybrid models

Yu et al. [4] demonstrated that in silico DTI prediction can significantly reduce the search space of potential drugs. Powerful algorithms can be used to uncover bioinformatics networks that include drugs, proteins, and related information. HGDTI combines node embedding learning and DTI classification phases within a heterogeneous graph neural network. After extracting initial node features from drug molecular fingerprints and protein pseudo-amino acid compositions, a Bi-LSTM was employed to aggregate heterogeneous neighbors using attention mechanisms. Among the key novelties of HGDTI is the integration of heterogeneous network embeddings with Bi-LSTM and attention mechanisms, which improves prediction accuracy and robustness in complex networks. HGDTI performed better than state-of-the-art DTI models based on negative sampling in comparison. In addition, the authors demonstrated the stability of HGDTI in heterogeneous networks and validated its reasonableness through further experiments. HGDTI has the potential to advance drug development via drug and target embeddings learning. HGDTI was composed of feature learning neural networks and label prediction neural networks. End-to-end optimization of HGDTI parameters could enable the former algorithm to learn reliable features and the latter to predict labels more precisely. When more heterogeneous networks were integrated, the model became more accurate. All these enhancements notwithstanding, negative sampling remains a limitation that continues to present challenges in some datasets, especially when there are imbalanced data. The model can be enhanced by solving this issue. HGDTI is applicable in both drug repositioning and discovery. The HGDTI method's limitations, on the other hand, manifest when negative sampling can restrict its generalizability to datasets of varying characteristics.

Drug–target Interaction prediction is crucial for new drug discovery, as stated by Ye et al. [11]. Deep Neural Network (DNN) prediction for DTI has garnered significant attention over the past few years. However, limitations in DTI data and its dimensionality may hinder its full potential. Ye et al. used graph autoencoders and multi-subspace deep neural networks to cross this barrier. In GAEMSDNN, with the graph auto-encoder, subspace layer, and ensemble layer, this model reinforces learning. It preserves reconstruction information and learns subsets of useful features through the subspace layer. The ensemble layer efficiently trains the DNN by extracting more features from the input data through the ensemble of these subsets. Experiments confirmed the superiority of GAEMSDNN over previous methods. One of the model's most significant innovations is the subspace layer, which enhances data feature extraction to facilitate the learning of more robust and diverse feature subsets, thereby significantly improving DTI prediction. The ensemble layer also enhances feature diversity and learning accuracy. The difficulty, however, is the complexity of combining several feature subspaces, which must be carefully optimized. This complexity also prescribes greater computational demands, which can become a limiting factor in practical applications. Furthermore, the model's use of random functions for subspace creation can be a source of instability in its performance on various DTI datasets.

Qian et al. [16] said that most deep neural networks focus on unimodal representations, i.e., SMILES sequences, molecular graphs, or drug molecular images. Most methods do not integrate information from multiple sources,

Table 8 DTI research questions

No	Research question	Motivation
1	Which neural network architectures are most commonly used for predicting drug–target interactions, and why are they preferred?	Understanding the popularity and advantages of specific architectures (like CNN, RNN, and GNN) can highlight trends and best practices in the field
2	What challenges do researchers face when applying neural networks to DTI prediction?	Identifying these challenges can help pinpoint limitations and suggest areas where improvements are needed
3	How can combining different neural network models enhance the accuracy and reliability of DTI predictions?	Exploring hybrid models reveals how blending approaches can overcome the limitations of individual methods
4	How can the scarcity of labeled data in DTI prediction be addressed effectively?	Addressing this issue can involve showcasing creative solutions, such as transfer learning or generating synthetic datasets, to enhance the performance of machine learning models
5	What role do Transformer architectures play in improving DTI predictions, especially when dealing with complex datasets?	Investigating Transformer architectures illustrates how cutting-edge architectures can effectively handle intricate data relationships
6	How can attention mechanisms make DTI prediction models more interpretable and transparent?	Attention mechanisms can help researchers and practitioners better understand how models arrive at their predictions by highlighting key features that contribute to these predictions
7	How do CNN, RNN, and GNN compare in terms of performance for DTI prediction tasks?	Comparing these architectures helps determine which model works best under different circumstances
8	What impact do biological data, such as protein sequences, have on the performance of DTI prediction models?	Examining this relationship can help researchers select the most impactful data sources
9	How can unsupervised learning be used to discover new drug–target interactions that aren't yet labeled in existing datasets?	This approach could unlock hidden patterns and enable the discovery of novel interactions
10	What are the key limitations of current neural network models in DTI prediction?	Understanding these limitations provides a roadmap for improving the next generation of models
11	How do existing DTI prediction methods handle noisy or incomplete data?	Identifying these strategies helps improve model robustness and reliability
12	What types of biological and chemical data are most effective for improving DTI prediction accuracy?	Exploring this question facilitates the integration of diverse and relevant datasets to achieve better outcomes
13	How can hybrid approaches leverage limited data to enhance DTI prediction performance?	Combining methods can enhance the use of limited data and improve prediction accuracy
14	Can transfer learning help DTI prediction models perform better when working with limited or sparse data?	Leveraging knowledge from pre-trained models can improve outcomes in data-scarce scenarios
15	What are the significant research gaps in applying neural networks to DTI prediction?	Highlighting these gaps can inspire innovative research and guide the field toward solving critical issues

Table 9 Overview of hybrid models

Article	Year	Method	Model type	Description	Used database	Feature extractor	Metrics
[4]	2022	HGDTI	Hybrid	Heterogeneous graph neural network model for DTI prediction, using node embeddings and aggregation with an attention mechanism	DrugBank,	Heterogeneous Graph Neural Network, Bi-LSTM	AUROC, AUPR
[11]	2023	GAEMSDNN	Hybrid	Graph auto-encoder and multi-subspace deep neural network for improving DTI prediction accuracy	Gold Standard Dataset, DrugBank	Graph Auto-encoder, Multi-subspace Neural Network	AUC, AUPR, Accuracy (ACC), MCC
[16]	2023	MCL-DTI	Hybrid	Uses multimodal information (molecular image and chemical text) with bidirectional cross-attention for improved DTI prediction	Human, C. elegans, Davis	Multi-head self-attention, Bidirectional cross-attention	AUROC, AUPR, Accuracy
[21]	2024	CCL-ASPS	Hybrid	The CCL-ASPS model utilizes contrastive learning and self-paced sampling to enhance DTI prediction by combining multiple networks and selecting informative negative samples	DrugBank, UniProt, RCSB Protein Data Bank (PDB),	GCN, Cooperative Contrastive Learning, SMILES Embeddings, Amino Acid Sequence Embeddings	ACC, AUROC, AUPR, MCC, F1
[22]	2024	Embedding aggregation	Hybrid	Comparison of embedding aggregation strategies for DTI prediction with MLP, dot product, and tensor product methods	Davis, KIBA	MLP, Dot Product, Tensor Product	MSE, R^2, CI
[23]	2024	MSH-DTI	Hybrid	Multimodal self-supervised learning with heterogeneous aggregation for DTI prediction	DrugBank	Graph Neural Network (GCN), Attention	AUROC, AUPR
[24]	2024	CCL-DTI	Hybrid	Uses contrastive loss learning to improve DTI prediction, combining multimodal data with attention-based fusion	Wang et al., Luo, Davis, KIBA	Attention-based feature fusion	ACC, AUROC, AUPR
[25]	2023	GIFDTI	Hybrid	Uses CNNFormer and intermolecular interaction modeling for DTI prediction	DrugBank, Davis, KIBA, BindingDB	CNNFormer, Global Molecular Feature Extractor	AUROC, AUPR, Accuracy, Precision, F1 Score, Recall
[26]	2024	DeepMHAttGRU-DTI	Hybrid	Uses random walk embeddings and GRU with multi-head attention for DTI prediction	GPCR, Ion Channels	Random Walk Graph Embedding, GRU with Multi-head Attention	ROC_AUC, AUPR
[27]	2024	MSF-DTI	Hybrid	Multi-scale feature fusion neural network for DTI prediction and binding affinity prediction	Human, C. elegans	Multi-scale Feature Fusion, Attention	AUROC, AUPR, Accuracy

Table 9 (continued)

Article	Year	Method	Model type	Description	Used database	Feature extractor	Metrics
[28]	2021	HGSDTI	Hybrid	Uses a heterogeneous graph convolutional network integrating multimodal similarities for DTI prediction	DTI dataset, SIDER, CTD dataset	Graph Convolutional Network (GCN), Convolutional Neural Network (CNN)	AUROC, Precision, Recall
[29]	2023	CSDTI	Hybrid	Uses cross-attention with GNN-based aggregation for interpretable DTI prediction	DrugBank, Human,	Cross-attention, GNN-based Drug Molecule Aggregation	AUROC, Precision, Recall, MSE, CI LDS-CNN
[30]	2020	MMIDTI	Hybrid	Explores multi-level mutual information with a Graph Convolutional Network (GCN) for type-aware and meta-path augmented DTI prediction	Drug-Protein Interaction, Drug–Drug Interaction, Protein–Protein Interaction, Drug-Disease Association, Protein-Disease Association, Drug-Side-Effect Association	Graph Convolutional Network (GCN), Mutual Information (MI)	AUROC, RMSE, MAE, AUPR
[31]	2020	MRCH-GCAE	Hybrid	Uses a multi-resolution collaborative heterogeneous graph convolutional auto-encoder for predicting DTIs	DTI, DDI, PPI	Graph Convolutional Auto-Encoder (GCAE), Collaborative Aggregation	AUROC, AUPR, Precision, Recall
[32]	2021	MHRW2Vec-TBAN	Hybrid	TBAN	DTI (Enzyme, Ion Channel, GPCR, Nuclear Receptor), DDI (DrugBank, KEGG, PharmGKB)	MHRW2Vec, TextCNN, Bi-LSTM, Attention Network	Accuracy, Precision, Recall, F1-Score
[33]	2021	GDNet-DTI	Hybrid	GDNet-DTI uses a GCN and DeepWalk for DTI prediction, combining drug molecular graphs and protein target features for improved accuracy, especially for COVID-19 data	DrugBank, KEGG, BRENDA	Graph Convolutional Network (GCN), DeepWalk, GIN	AUC, AUPR
[34]	2022	ConformerDTI	Hybrid	ConformerDTI combines CNN and Transformer models to enhance DTI prediction by integrating local and global features from drug and protein sequences. It outperforms other methods in multiple datasets	DrugBank, Biosnap, Davis	CNN, Transformer, Convolutional Interaction Network	AUROC, AUPR, RMSE, R^2
[35]	2024	SCTDTI	Hybrid	Collaborative prediction using sequence-based CNN and Transformer architectures for DTI	Davis, KIBA	CNN, Transformer, Attention Mechanism	AUROC, AUPR, Accuracy

Table 9 (continued)

Article	Year	Method	Model type	Description	Used database	Feature extractor	Metrics
[36]	2022	Transformer-CNN-OMBO	Hybrid	Utilizes Transformer architecture and CNN layers, combined with OMBO optimization, for predicting DTIs, focusing on molecular sub-structural patterns	BindingDB	Transformer, CNN, Opposition-based Monarch Butterfly Optimization (OMBO)	accuracy, sensitivity, specificity
[37]	2020	DTIGCCN	Hybrid	Combines GCN and CNN for predicting drug–target interactions, refining the features from both drug and target expression profiles	A375, A549, HA1E, HCC515,	GCN, CNN	AUROC, Precision, Recall, F1-Score
[38]	2023	CHADTI	Hybrid	Sequence-based model using an attention mechanism for extracting complex interactions between drugs and proteins	DrugBank, Davis	1D-CNN, Pre-trained model (ESM), Attention Mechanism	AUROC, AUPR, Accuracy, Precision
[39]	2022	DTIPred	Hybrid	Random walk and CNN for DTI prediction, integrating heterogeneous data	DrugBank, Side Effect Database	Random Walk, CNN	AUROC, Precision, Recall, F1-Score
[40]	2022	MHSADTI	Hybrid	Multi-head self-attention and Graph Attention Network (GAT) for predicting DTIs, focusing on long-dependent contextual relationships	Human, C. elegans, DUD-E, DrugBank	Multi-Head Self-Attention, Graph Attention Network (GAT)	AUC, Precision, Recall, AUPR, F1-score
[41]	2024	SMGCN	Hybrid	Multiple Similarity and Kernel Fusion-Based Graph Convolutional Network for DTI prediction, integrating Random Walk with Restart and MKL	Gold standard datasets	Graph Convolutional Neural Network (GCN), Multiple Kernel Learning (MKL)	AUC, AUPR
[42]	2021	GanDTI	Hybrid	Multi-task neural network employing GNN and attention for DTI prediction, capable of both interaction classification and affinity prediction	DUD-E, Human, BindingDB	Graph Neural Network (GNN), Attention, Multi-layer Perceptron (MLP)	Accuracy, Precision, F1-Score, RMSE

Table 10 Advantages and disadvantages of hybrid models

Article	Advantage	Disadvantage
[4]	-High DTI accuracy with heterogeneous graphs - Uses attention for improved predictions - Accessible as a web server	- Needs diverse datasets - High computational cost
[11]	- Employs Graph Auto-Encoder and multi-subspace deep neural networks to extract features and strengthen gradient signals - Capable of learning diverse, comprehensive features from high-dimensional data - Utilizes multi-subspace infrastructure for strong feature extraction and unified optimization	- Requires substantial training data due to the use of Graph Auto-Encoder and multi-subspace network - Complex design and implementation with a need for precise tuning - Model performance may decline without large datasets
[16]	- Uses multimodal drug info (images + text) - Effective cross-attention mechanism - Good for other tasks (e.g., DDI)	- High resource demand - Requires precise tuning
[21]	- Uses Collaborative Contrastive Learning (CCL) for consistent drug–target representations - Adaptive self-paced sampling for better adverse sample selection - Demonstrates improved performance over baselines	- Complexity of model design and implementation - Requires extensive computational resources for training
[22]	- Compares different aggregation strategies (MLP, dot product, tensor product) - Provides insights on the effectiveness of each method for DTI prediction - Useful for selecting aggregation methods in a deep neural network	- Limited to theoretical and empirical comparison, without proposing a novel model - Aggregation strategy recommendations may not generalize across all DTI datasets
[23]	- Incorporates multiple graph types and self-supervised learning for feature extraction - Uses an attention mechanism for focusing on key features - Achieves high AUROC and AUPR scores on the DTINet dataset	- High computational cost due to self-supervised learning methods - Limited applicability if pre-trained models are not available
[24]	- Uses multiple contrastive loss functions to enhance prediction accuracy - Combines multimodal knowledge with protein and drug data - Employs attention-based networks to focus on key features of drugs and proteins	- High complexity and requires careful parameter tuning due to multiple contrastive functions - High computational resources needed for training, especially with large datasets - May require adaptation for various data types
[25]	- Uses CNNFormer to combine CNN and Transformer architecture for local and long-range feature extraction - Improves model performance in drug–target interactions by extracting global and intermolecular features - Useful for low-cost drug screening in early predictions	- High computational demands due to combining the CNN and Transformer architecture - Limited applicability for small or low-diversity datasets - Model may need comprehensive, diverse data for optimal accuracy
[26]	- Combines knowledge graph embeddings with GRU and multi-head attention for improved feature extraction - Effective for integrating heterogeneous data sources for DTI prediction - Monte Carlo random walk with Metropolis–Hastings for better feature embeddings	- Complexity in implementing and training due to multiple embedding techniques - May face challenges in scalability with huge graphs
[27]	- Uses a multi-scale feature fusion network to improve protein feature representation - Incorporates binding affinity prediction, not just binary DTI classification - Outperforms previous models in accuracy for both DTI and binding affinity tasks	- High complexity due to multi-scale and multi-stage fusion - Limited interpretability in the fusion process, making biological insights less accessible
[28]	- Integrates multimodal similarities in a heterogeneous network (drugs, proteins, diseases, side effects) - Utilizes attention-based GCN for capturing relational dependencies - Enhanced robustness with a denoising auto-encoder for high-quality embeddings	- Limited interpretability due to complex multimodal integration - High memory and computational requirements for handling heterogeneous data

Table 10 (continued)

Article	Advantage	Disadvantage
[29]	- Employs a cross-attention network for improved interaction interpretation between drugs and targets - Aggregates high-order dependencies for global molecular representation - Provides enhanced biological insights through attention visualization	- Complexity in interpretability due to the cross-attention network - May struggle with scalability on larger datasets due to model depth
[30]	- Incorporates multi-level mutual information to capture both local and global information for better DTI prediction - Utilizes a heterogeneous network structure, integrating drug–protein, drug-disease, and drug-side effect data - Outperforms baseline methods in prediction accuracy on benchmark datasets	- Complexity in model structure, requiring extensive tuning - High computational cost due to the heterogeneous network and multi-level mutual information approach
[31]	- Uses a multiresolutional collaborative heterogeneous GCN with an auto-encoder for more interpretable embeddings - Effective for predicting DTIs for novel drugs and proteins due to multi-level aggregation - Demonstrates high accuracy and interpretability	- Complex implementation, with a need for multiple convolution kernels - High memory requirements due to multi-resolution and multi-type data integration
[32]	- Combines knowledge graph with improved neural network techniques for DTI and Drug–Drug Interaction (DDI) prediction - Utilizes MHRW2Vec for enhanced feature vectors, capturing long-distance correlations - Flexible and adaptable for predicting various drug–target and drug–drug pairs	- Complex training process due to the combination of a knowledge graph and a neural network - May struggle with interpretability due to the complex knowledge graph embedding approach
[33]	- Uses a GCN with DeepWalk to capture local and global topological information in a molecular graph - Demonstrates high predictive accuracy on both DrugBank and benchmark datasets - Case study on COVID-19 data shows potential for real-world DTI applications	- Limited scalability due to GCN and DeepWalk integration, which requires extensive computational resources - Complexity in feature extraction may require fine-tuning for different types of drug and target data
[34]	- Combines CNN and Transformer architecture for enhanced local–global feature extraction - Cross-attention mechanism improves the prediction of Drug–Target Interactions (DTIs) - Demonstrated high accuracy across multiple datasets	- High computational demands due to complex architecture - Model tuning is challenging due to the interplay between the CNN and the Transformer architecture
[35]	- Collaborative CNN-Transformer approach enhances feature interaction for DTI prediction - Incorporates an attention mechanism for improved intermolecular feature extraction - High accuracy across multiple benchmark datasets	- High model complexity, making it resource-intensive - Potentially challenging to tune due to the collaboration between the CNN and the Transformer architecture
[36]	- Transformer-CNN-OMBO model uses sub-structural pattern discovery for DTI - Effective at capturing molecular substructures for improved prediction accuracy - OMBO optimizes CNN hyperparameters for better performance	- High complexity in implementing OMBO and substructure discovery - Requires large, labeled datasets for optimal results
[37]	- DTIGCCN model combines GCN and CNN to improve DTI prediction accuracy - Spectral-based GCN for effective feature extraction from drug and target profiles - Superior performance compared to traditional feature extraction methods	- High computational demands due to the combination of GCN and CNN - Complexity in implementation and training due to the spectral-based approach
[38]	- CHADTI uses a sequence-based model with attention to capture drug–protein interactions - Pre-trained ESM model enhances protein embedding accuracy - Achieves state-of-the-art performance on benchmark datasets	- High computational complexity due to the attention mechanism - Requires extensive pre-training for optimal protein feature extraction
[39]	- DTIPred uses a random walk and CNN for capturing local and global features in heterogeneous networks - Effective in retrieving true interactions within top predictions - Demonstrated effectiveness in real-world case studies	- High complexity due to random walk integration - Requires substantial computational resources for training

Table 10 (continued)

Article	Advantage	Disadvantage
[40]	- MHSADTI model uses multi-head self-attention and GAT for DTI prediction - Captures long-term dependencies effectively in amino acid sequences - Outperforms previous models on various datasets	- High computational cost due to multi-head self-attention - Complex architecture requiring fine-tuning for optimal results
[41]	- SMGCN combines multiple similarity measures with GCN for enhanced feature extraction - Multiple kernel fusion improves model robustness and performance - High accuracy on standard DTI datasets (AUC, AUPR)	- High resource demand due to multiple similarities and kernel fusion - Complexity in integrating multiple kernels into the GCN
[42]	- GanDTI model utilizes GNN with an attention mechanism for multi-task DTI prediction - Effective for both interaction classification and binding affinity prediction - Outperforms state-of-the-art on multiple benchmark datasets	- Limited interpretability due to the use of GNN - High computational demand for processing large datasets

resulting in an incomplete representation of drugs and their targets. They utilized molecular photos, along with chemical properties, to obtain a more comprehensive understanding of drug properties. As a result, the drug images contained structural and spatial information, while the chemical information provided complementary functions and properties, enhancing their representation. The performance of DTI was improved through the addition of a bidirectional multi-head attention mechanism. One of the main novelties of the MCL-DTI model is the integration of multimodal data, including structural, spatial, and functional data, thereby providing a more comprehensive and integrative representation of drug–target interactions. The bidirectional multi-head attention mechanism enables the model to learn complex interactions more effectively, thereby significantly improving the performance of DTI prediction. However, the integration of multimodal data also presents challenges; for example, it may not always be feasible to obtain and process molecular images and chemical properties due to the availability and quality of such multimodal data. This could limit its general applicability, especially in environments where data diversity is not present.

As noted by Tian et al. [21], most previous methods have failed to fully leverage the complementary properties of biological networks. Additionally, the process of adversarial sampling selection has a significant influence on contrastive learning-based frameworks. An Adaptive Self-Paced Sampling model was proposed for drug–target interaction prediction, such as Collaborative Contrastive Learning (CCL). Various networks were employed in CCL-ASPS to fuse drug embeddings and targets for consistency. In contrastive learning, ASPS selected more informative negative sample pairs. In comparison with current state-of-the-art methods, CCL-ASPS achieved dramatic gains. The key novelty of CCL-ASPS lies in the use of self-paced sampling and collaborative contrastive learning, which ensures more informative negative sample pairs and improves overall DTI prediction performance. However, relying on self-paced learning and contrastive learning remains challenging. Negative sample optimization is crucial, and it includes specializing the model to specific datasets. CCL-ASPS's complexity and dependence on contrastive methods may limit its scalability to other datasets of varying natures, making it more task-dependent.

Iliadis et al. [22] noted that most deep neural networks introduced over the past decade share a typical two-branch architecture, differing only in their feature representations and branches. The method for fusing the output of the branches (embeddings) has been the same. Recommender systems utilize the same overall architecture, yet the actual merging techniques are not specified. Drug–target Interaction (DTI) prediction was tested with three embedding fusion techniques. These approaches were specifically defined and were universal mathematical approximators. The experiments compared different DTI prediction approaches on benchmark datasets to demonstrate under what conditions some approaches were preferable. The main novelty of the work of Iliadis et al. is the evaluation of different aggregation approaches for the outputs of branches in DTI prediction, which were universal approximators. However, the need for universal approximators for proteins and compounds is a drawback, because these networks are not universally applicable without special configurations. This can limit the model's flexibility on different protein and compound datasets. The model's ability to represent a wide variety of functions is, nonetheless, a valuable contribution to DTI prediction.

Zhang et al. [23] stated that current DTI prediction models are based on small datasets for extracting drug and target features, which may limit their scope and robustness. DTI prediction was carried out using several networks; however, the aggregation and attention mechanisms used

Table 11 Overview of CNN models

Article	Year	Method	Model type	Description	Used database	Feature extractor	Metrics
[3]	2020	FRnet-DTI	CNN	Uses two deep convolutional neural networks for feature manipulation (FRnet-Encode) and classification (FRnet-Predict) of DTIs	Gold Standard Datasets	Convolutional Neural Network (CNN)	AUROC, AUPR
[10]	2020	DEEPScreen	CNN	DEEPScreen is a deep neural network-based system that predicts Drug–Target Interactions (DTIs) using 2D compound structure images. It outperforms traditional methods, aiding in early stage drug discovery and drug repurposing	Random-Split Dataset	Convolutional Neural Network (CNN)	ACC, Precision, Recall
[43]	2023	LDS-CNN	CNN	Uses large-scale drug screening with unified encoding for predicting interactions between small-molecule drugs and protein targets	DrugBank, STITCH	Convolutional Neural Network (CNN), Unified Encoding	AUC, AUPRC, ACC
[44]	2020	LDCNN-DTI	CNN	A light deep convolutional neural network for DTI predictions with fewer protein descriptors	DrugBank	Light Deep Convolutional Neural Network (LDCNN)	50% fewer neurons, AUC loss, AUPR
[45]	2023	DeepCNN-DTI	CNN	Uses Convolutional Neural Networks (CNN) to transform protein and drug data into features for DTI classification, achieving high accuracy	DrugBank	Convolutional Neural Network (CNN), Legendre Moments (LMs)	Accuracy, Precision, Sensitivity, Specificity, F1-Score
[46]	2023	CNN-DDI	CNN	Uses CNN for predicting Drug–Drug Interactions (DDI) with high validation accuracy	DDI-Corpus	Convolutional Neural Network (CNN)	F1-Score, Accuracy
[47]	2022	CNN-OSBO	CNN	CNN encoder–decoder architecture with Opposition-based Satin Bowerbird Optimizer (OSBO) for predicting DTI in COVID-19 targets	Custom Dataset, DrugBank	CNN, OSBO	Accuracy, Precision, Recall
[48]	2021	CNN	CNN	It uses Random Walk and CNN on heterogeneous networks to predict DTI, integrating multiple similarities and associations of drugs and proteins	DrugBank	Random Walk, Convolutional Neural Network (CNN)	Accuracy, AUC, Precision

Table 12 Advantages and disadvantages of CNN models

Article	Advantage	Disadvantage
[3]	Generates 4,096 features for improved accuracy -Introduces 20 new drug–target pairs -Improved AUROC and AUPR scores	-High Computational Demand may struggle with small datasets -Complexity may lead to overfitting
[10]	- Utilizes 2D images of drug structures to extract richer features - High accuracy in interaction predictions due to Convolutional Neural Networks (CNN) - Suitable for early stage identification of drugs and chemical compounds	- Strong reliance on 2D data, which may limit accuracy in complex interactions - Challenges in model generalization to unfamiliar data - Performance may depend on large, comprehensive datasets
[43]	- Utilizes a large-scale CNN for Drug–Target Interaction (DTI) prediction, which achieves high accuracy with an AUC of 0.96 - Employs unified encoding for handling various data formats, enhancing compatibility - Effective for large datasets with different formats	- High computational demands due to the extensive dataset processing - May not generalize well to smaller datasets due to encoding complexity
[44]	- A lightweight DTI prediction model that reduces the number of neurons and filters, achieving efficiency - Uses convolution on amino acid sequences to extract protein features effectively - Requires 50% fewer resources compared to other models with minimal accuracy loss	- May have slightly reduced accuracy (1.3% AUC and 4% AUPR loss compared to other models) - Simplicity may limit its application to more complex DTI scenarios
[45]	- DeepCNN model effectively captures complex drug–target relationships - Achieves high performance in accuracy, precision, and other metrics - A comparatively simple architecture with strong feature extraction	- Limited generalizability to unseen data due to a lack of long-range feature extraction - Relies heavily on CNN, which may limit performance on highly complex interactions
[46]	- CNN-based model shows state-of-the-art performance in DTI prediction - Simple architecture allows for efficient training and testing - High accuracy and F1 score on DDI-Corpus dataset	- Limited interpretability due to a simple CNN approach - May not handle long-range dependencies as effectively as Transformer-based architectures
[47]	- CNN-OSBO encoder–decoder model focuses on COVID-19 target proteins - Uses OSBO for hyperparameter optimization, improving model accuracy - Demonstrated accuracy improvement over previous COVID-19 DTI prediction models	- Limited to COVID-19 targets, which may limit generalizability - OSBO optimization adds computational overhead
[48]	- CNN-based model for DTI discrimination with high accuracy (AUC 0.9527 on DrugBank dataset) - Novel approach for generating credible negative instances improves robustness - Demonstrated generalization on multiple datasets	- Negative instance generation adds computational overhead - Limited interpretability due to CNN structure

were simplistic, restricting their capabilities. In this work, MSH-DTI is proposed for predicting drug–target interactions. Drug and target structure features were extracted using self-supervised learning. With heterogeneous interaction-augmented Feature Fusion Modules, node features were learned by convolutional neural networks. The model could predict multiple features via an attention mechanism. The MSH-DTI model achieved superior performance compared to other models on the DTINet dataset with an AUROC and AUPR of 0.9605. The key novelty of MSH-DTI lies in its self-supervised learning approach for extracting structural features, combined with attention mechanisms for feature prediction. Yet, its reliance on straightforward aggregation and attention mechanisms could limit its scalability to extremely complicated or imbalanced datasets that would require more sophisticated feature extraction methods.

Dehghan et al. [24] illustrated that deep neural network-based models have gained increasing attention in recent years. Categorical cross-entropy and mean squared error were often used to train the models. Contrastive-based loss functions improve machine learning feature space discrimination. For information fusion, an attention-based fusion method was developed based on multimodal knowledge. The model was further empowered by investigating four specific contrastive loss functions: Normalized Temperature-scaled Cross Entropy Loss (NT-Xent), max-margin contrastive loss, triplet loss, and multi-class N-pair loss. Their multimodal network outperformed other methods that had

Table 13 Review of transformer models

Article	Year	Method	Model type	Description	Used database	Feature extractor	Metrics
[14]	2024	CAT-DTI	Transformer	Cross-attention Transformer-based architecture with domain adaptation for DTI prediction, leveraging Transformer and CNN for protein and drug encoding	Human, BindingDB	Transformer, Convolutional Neural Network (CNN)	AUROC, AUPR, Accuracy, Sensitivity & Specificity,
[15]	2024	MocFormer	Transformer	Two-stage pre-training-driven Transformer architecture for DTI prediction, with molecule and protein models in the first stage	DrugBank, Epigenetic-regulators	Transformer, Pre-training with transfer learning	AUROC, AUPR, Recall, Accuracy
[17]	2024	FragXsiteDTI	Transformer	Transformer-based architecture using drug molecule fragments and protein pockets for DTI prediction, inspired by the Perceiver IO framework	Human, C. elegans, DrugBank	Cross-attention Transformer architecture, Self-attention Transformer architecture	AUROC, AUPR, Accuracy, RMSE
[49]	2022	DTI-GTN	Transformer	Uses a graph transformer architecture to model and predict DTI by transforming interactions into line graph nodes	Peng Dataset	Graph Transformer, Line Graph	AUROC, AUPR
[50]	2023	MGDTI	Transformer	Graph transformer architecture with meta-learning to address the cold-start problem in DTI predictions, using drug–drug and target–target similarity as additional info	DrugBank, Biosnap, KEGG	Graph Transformer, Meta-learning	AUC, AUPR
[51]	2024	BERT-DTI	Transformer	Uses BERT for protein and drug SMILES sequences, concatenated and fed to a random forest classifier for predicting drug–target interactions	DrugBank, KEGG, PubChem	BERT for Proteins and Drugs, Random Forest	Precision, Recall, F1-Score, AUC

Table 14 Advantages and disadvantages of transformer models

Article	Advantage	Disadvantage
[14]	- Strong cross-domain performance (CDAN) - Combines Transformer architecture and CNN for feature learning - Captures complex interactions	- Complex, requires tuning - May need data adaptation
[15]	- Utilizes pre-trained Transformer architecture for drug–target interactions - Reduces overfitting with a two-stage framework - Outperforms state-of-the-art methods in accuracy and AUROC	- Heavy reliance on pre-training, which is time-consuming and costly - Limited interpretability due to complex model architecture
[17]	- FragXsiteDTI leverages Transformer-based interpretation for DTI - Uses drug fragments and protein pockets, improving model interpretability - Achieves high accuracy with an explainable AI approach	- High model complexity due to Transformer-based interpretation - Requires comprehensive data preparation for accurate protein pockets
[49]	- Uses line graphs to capture complex drug–protein relationships - Employs Graph Transformer Network (GTN) for node interactions - High performance on Peng et al. dataset (AUROC, AUPR)	- Limited effectiveness for datasets without high-quality interaction data - GTN-based models can be resource-intensive
[50]	- Addresses the cold-start problem in DTI prediction using meta-learning and graph transformer - Uses similarity matrices for drugs and targets to improve prediction for new entities - Validated with effective results on benchmark datasets	- Complexity in implementation due to meta-learning and transformer integration - Requires extensive training time, especially with large datasets
[51]	- Utilizes BERT for protein and SMILES representations, improving feature richness - Effective for feature extraction from diverse datasets without the need for traditional descriptors - High validation accuracy on benchmark datasets	- High computational resources needed for BERT-based feature extraction - May face scalability issues with huge datasets

utilized protein sequences, drug molecules, protein–protein interaction networks, and drug–drug interaction networks. The most noticeable innovation of Dehghan et al.'s model is the meticulous combination of attention-based fusion with contrastive loss functions to augment feature extraction and interaction prediction. The challenge, nevertheless, remains the computational cost of jointly optimizing multiple contrastive loss functions, which can introduce scalability issues for large-scale datasets or real-time systems.

Zhao et al. [25] explained that among the challenges in building deep neural network-based models was the proper representation of drugs and proteins, as well as local chemical environments and long-range information in proteins. Drug–Target Interactions (DTIs) are primarily based on intermolecular interactions between drugs and their corresponding protein targets. They proposed a novel model, GIFDTI, which comprises three primary components: a sequence feature extractor (CNNFormer), a molecular feature extractor (GF), and an intermolecular interaction model (IIF). The innovation of GIFDTI lies in the integration of CNN and Transformer architectures, enabling the model to capture both local motifs and long-range dependencies among tokens (atoms or amino acids). GF captures molecular features and intermolecular interaction features by IIF. Six realistic experiments were constructed to validate the model's effectiveness. Different drugs and proteins were utilized to demonstrate GIFDTI's powerful and robust predictive ability. Affordable drugs and personalized medicine are attainable based on the DTI model. Its stability was attributed to CNNFormer's powerful feature extraction ability as well as the complementarity of global and intermolecular interactions. However, the model's complexity in incorporating different architectures will impede its application in resource-limited environments.

Yu et al. [26] presented the model DeepMHAttGRU-DTI to predict drug–target interactions. Drug–target interactions are crucial to identify in drug discovery and drug repurposing. Here, a knowledge graph was utilized to integrate biological data from multiple sources, thereby reducing experimental costs and the time required for developing new medications. After that, the feature vectors were extracted from the graph using a random walk-based graph embedding algorithm. To train a deep neural network model, the preprocessed data were mapped to universal datasets, and successful mappings were used to build a binary classification dataset. To predict drug–target interactions, the proposed model employs a Gated Recurrent Unit (GRU) neural network enhanced with a multi-head attention mechanism. MHRW-based Monte Carlo Random Walks provided better node feature vectors in experiments. Multi-head attention

Table 15 Review of GNN models

Article	Year	Method	Model type	Description	Used database	Feature extractor	Metrics
[52]	2024	HiGraphDTI	GNN	Hierarchical graph representation learning to enhance DTI prediction using atom, motif, and molecule-level information	Human, C. elegans, BindingDB, GPCR	Hierarchical Graph Learning, Hierarchical Attention	AUC, AUPR, Accuracy
[53]	2024	DSG-DTI	GNN	Dual-Stream Graph Neural Network combining heterogeneous graph auto-encoder and heterogeneous attention-based matrix completion	DTI-HN	Heterogeneous Graph Encoder, Attention-based Matrix Completion	Precision, Recall,
[54]	2023	SubgraphDTI	GNN	Uses drug subgraph fingerprint extraction and attention mechanism for DTI prediction	DUD-E Dataset, Human Dataset, BindingDB Dataset	Subgraph Fingerprint Extraction, Subgraph Attention	AUC, Precision, Recall
[55]	2021	GCN-DTI	GNN	Graph Convolutional Neural Network for DTI prediction using drug–target and drug–drug interaction networks	DTI-Net	Graph Convolutional Neural Network (GCN), Decagon	AUC, AUPR

also contributed to the performance boost of the GRU model. One of the key new contributions of this model is the integration of knowledge graphs with Gated Recurrent Units (GRU) and multi-head attention, enabling more accurate and effective drug–target predictions. However, based on Monte Carlo random walks, integrating big biological data networks can be computationally costly, particularly with huge and heterogeneous datasets.

As noted by Yang et al. [27], sequence-based deep neural network models had some disadvantages. As a result, amino acid sequences were not sufficient for feature extraction of protein targets, and drugs and targets were not effectively integrated. Many methods treated DTI merely as a binary classification problem, without predicting binding affinity, which indicated the strength of drug–target interactions. To predict drug–target interactions, they developed a multi-scale feature fusion neural network (MSF-DTI) that integrates the semantic features of amino acid sequences, protein feature representations, and drug and target features within a feature fusion module. Both experiments on DTI classification and binding affinity prediction demonstrated that MSF-DTI outperformed state-of-the-art methods. The main innovation of MSF-DTI is its ability to combine multiple forms of features from various biological data sources to achieve more robust and precise predictions. However, integrating the features remains a challenge, because it requires the accurate integration of disparate data types, which can be difficult when handling noisy or missing data. Additionally, the model's reliance on a binary classification task may overlook nuanced variations in the strength of interactions, potentially leading to a less precise understanding of drug–target affinity.

Jiang et al. [28] proposed HGSDTI as a new method. They initially employed a denoising auto-encoder to learn low-dimensional representations of drugs, proteins, diseases, and their associated side effects. Then they used a three-layer Graph Convolutional Network (GCN) to integrate low-dimensional features and learn neighbor topology information from the heterogeneous network. Their next step was to calculate multimodal drug similarities and protein similarities based on multi-scale relationships. Finally, the neighborhood topology of the neighborhood was combined with drug similarity information using the Convolutional Neural Network (CNN). Experiments revealed that this method outperformed state-of-the-art methods. The most significant contribution of HGSDTI is the application of denoising autoencoders in feature extraction to reduce noise and improve the model's ability to handle missing data. A limitation of this method is that it relies on multimodal data, which is not consistently available or of sufficient quality, thereby limiting its generalizability across various datasets.

Pan et al. [29] reported that deep neural network-based DTI methods use Graph Neural Networks (GNNs) to

Table 16 Advantages and disadvantages of GNN models

Article	Advantage	Disadvantage
[52]	- Utilizes a hierarchical graph representation learning approach to leverage chemical information at atomic, motif, and molecular levels - Employs attentional feature fusion to improve the representation of protein targets - Offers interpretability in identifying crucial molecular segments	- Complexity in implementing hierarchical graph representation and attention fusion - High computational demands for processing detailed graph structures
[53]	- DSG-DTI model uses dual-stream GNN with matrix completion for accurate DTI prediction - Handles heterogeneous data with an attention-based graph approach - Shows strong generalization on public benchmark datasets	- Limited interpretability due to complex graph structure - Computationally expensive, especially on large datasets
[54]	- Leverages attention mechanism to highlight necessary substructures in drugs - High accuracy in complex interactions due to focus on critical drug components - Suitable for detailed protein and drug data processing	- Complexity in implementation and requires fine-tuning for different drug substructures - Potential errors if drug substructures do not match expectations - May need diverse data to improve generalization
[55]	- The GCN model for DTI prediction utilizes graph structures for rich feature extraction - Effective in capturing topological information from heterogeneous networks - Achieves high AUC and AUPR on DTI-Net data	- Limited to COVID-19 targets, which may limit generalizability - OSBO optimization adds computational overhead

Table 17 Review of DNN models

Article	Year	Method	Model type	Description	Used database	Feature extractor	Metrics
[2]	2023	OverfitDTI	DNN	Deep neural network trained to overfit Drug–Target Interaction (DTI) data, focusing on reconstructing the complex non-linear relationships of DTI	DTC	Various DNNs (CNN, GNN, VAE)	MSE, Concordance Index (CI)

represent drug molecules and attention mechanisms to represent drug–target interactions. It is challenging to represent global 3D structure and edge information in GNNs, as they primarily operate on local neighbors. Drug–target interactions are poorly represented in current attention-based models. They introduced CSDTI as a solution for them. The cross-attention mechanism keeps drug–target interactions. Drug molecular graph aggregators keep high-order dependencies. Subsequent downstream tasks are executed in parallel. DTI prediction tasks on CSDTI surpass state-of-the-art baselines, demonstrated by rigorous experiments. On a range of benchmark datasets, CSDTI competes well with existing models. The AUC and recall for CSDTI were both improved by 3.5% and 3.7% over the second-best algorithm. In addition, attention weights demonstrate the biological relevance of the proposed approach. CSDTI's unique aspect lies in its cross-attention mechanism and high-order dependency modeling that enables a richer representation of complex drug–target relationships. However, the model's structural information is limited in its ability to represent it effectively through its application to protein representations. The tertiary structure of proteins predominantly determines their biological function, and failure to accurately represent this information limits the model's applicability.

Chen et al. [30] introduced MMIDTI, a DTI predictive model derived from heterogeneous networks (e.g., drug–protein or drug–drug interactions). MMIDTI employed an encoder–decoder strategy, where the encoder learned type-aware and meta-path-enriched node representations. Graph convolutional networks were used for decoding the original heterogeneous network. Mutual information was employed at two scales in MMIDTI: (1) mining local mutual information to limit node representations, and (2) optimizing global mutual information to limit node representations. During DTI prediction, MMIDTI learns node discriminative representations (e.g., drugs, protein targets) from mutual information. MMIDTI adopts enhanced aggregation to encode heterogeneous

Table 18 Advantages and disadvantages of DNN models

Article	Advantage	Disadvantage
[2]	-Utilization of All Data: OverfitDTI uses all available data for learning, unlike standard methods that use partial data to avoid overfitting -Prediction of New Compounds: The model successfully predicted new compounds as TEK inhibitors that were previously unknown	-High Computational Resource Requirement: OverfitDTI requires significant computational resources, especially for model tuning and feature selection -Need for Experimental Validation: Additional experimental tests are required to confirm the model's predictions in real-world scenarios

graph information. Mutual information maximization modules were also added to the model. MMIDTI outperforms baseline methods on real DTI prediction tasks. Perhaps, the most significant innovation of MMIDTI is the use of mutual information to enhance node representations, thereby making more informative and stable predictions. One limitation of this method is that it is computationally intensive, particularly in the use of mutual information to restrict node representations, which can be computationally costly and challenging to scale up to massive datasets.

Jin et al. [31] have generally assumed that the simultaneous application of Drug–Drug Interactions (DDI) and Protein–Protein Interactions (PPI) networks could yield a more precise prediction. Nevertheless, existing methods have mainly ignored complementary information concealed in various types of interactions between edges and nodes. For predicting DTI, this paper introduces the Multi-Resolution Collaborative Heterogeneous Graph Convolutional Auto-Encoder (MRCH-GCAE), which aggregates embeddings from heterogeneous drug–target networks via different types of links, thereby generating more interpretable embeddings. It accurately predicts DTIs and recommends new target proteins for developing new drugs. With its unique structure, this method leverages the local and global topological properties of heterogeneous drug–target networks. It also included multiresolutional neighboring data representation embeddings. The originality of MRCH-GCAE lies in its use of multi-resolution collaborative learning to collect embeddings from heterogeneous networks, thereby producing more interpretable and accurate predictions. Although the model's complexity—particularly in handling multi-resolution embeddings and heterogeneous information—may limit its efficiency when used in real-time or with extensive networks, Zhang et al. [32] describe a hybrid approach that combines graph representation learning with neural networks. KG-DTI and KG-DDI knowledge graphs are constructed in the initial step using MHRW2Vec-TBAN. They then built feature vectors from network structure information using MHRW2Vec, a state-of-the-art graph representation learning model. Such feature vector analysis led to the creation of a novel neural network model called Text CNN-BiLSTM-Attention Network (TBAN). Drugs and their potential neighborhoods were predicted more precisely with this approach. Moreover, COVID-19 DTI can be predicted by the model. MHRW2Vec-TBAN effectively learnt local and global knowledge graph representations. The proposed methodology is novel in its combination of graph representation learning and TextCNN-BiLSTM-Attention models, which enhance the predictability of drug–target interactions, particularly for COVID-19. However, application of knowledge graphs to feature extraction may render the model rigid, because the quality and availability of

these graphs are not always guaranteed, which could affect its performance in other areas.

Xu et al. [33] proposed the GDNet-DTI predictor as a variant of GCN and DeepWalk. They first constructed a molecular map from the SMILES sequence of drugs to understand the complex interaction between atoms. This was followed by feature extraction from the map using GIN to learn about the map's features and uncover more complex interactions between atoms. Protein sequences were represented as word vectors, enabling the extraction of information at multiple levels. They then created a DTI graph, where drugs and targets are nodes, and interactions between them are represented as edges. They then utilized GDNet-DTI to derive the topology and neighborhood information of the DTI graph. In contrast to other state-of-the-art models, the method demonstrated better accuracy and efficiency, achieving superior performance based on DrugBank and four benchmark data sets, and outperforming other state-of-the-art models. Additionally, a COVID-19-based case study demonstrated that this method was capable of predicting actual drug–target interactions and aiding in the development of new drugs. What is novel in GDNet-DTI is that it can integrate SMILES-based molecular maps with GCN and DeepWalk methods to gain a deeper understanding of drug–target interactions. However, the complexity of handling multi-source molecular data might limit its application in settings of heterogeneous or incomplete data.

Wang et al. [34] stated that current methods lack local and global characteristics of drug molecules and protein sequences. This was resolved using Convolutional Neural Networks (CNN) and Transformer designs. For pharmaceuticals, SMILES strings and amino acids were used as inputs to the Transformer and CNN models. Feature globalization was coupled with localization. The Transformer models were rendered extremely efficient by processing both CNN's local feature extraction and global feature extraction concurrently. ConformerDTI also utilizes dynamic filters derived from drug–target interactions for modeling these interactions. Their model outperformed the best deep neural network methods on three varied datasets. The primary innovation of this model is to combine CNN and Transformer architectures to achieve maximum efficiency in global and local feature extraction, ultimately enhancing drug–target interaction prediction. However, combining CNN and Transformer may increase processing time and computational cost requirements.

With Transformer and CNN models, Li et al. [35] proposed a sequence-based collaborative prediction method named SCTDTI. SCTDTI encodes local and global molecular properties as well as intermolecular interactions among drug and protein SMILES sequences, utilizing the local feature extraction capabilities of CNNs and global feature modeling capabilities of Transformers. SCTDTI outperformed the conventional DTI prediction methods on two challenging sets on which it was trained and evaluated. The novelty of SCTDTI lies in its incorporation of CNN and Transformer to represent both local and global molecular properties within a single robust framework, thereby facilitating enhanced predictions through the establishment of deeper relationships between the features. Nonetheless, the approach's dependency on extensive training and computational cost can limit its effectiveness in real-time scenarios and when processing large datasets.

Nandhini and Mayilvahanan [36] introduced the Transformer-CNN-OMBO method for predicting DTI interactions. Protein and input drug data sets were transformed into explicit substructures using Improved Frequent Subsequence Mining (IFSM). After the transformer binding process, contextual bindings were established for every structure according to the substructures. With the addition of an optimal CNN layer to these interactions, they were able to make predictions regarding neighborhood interactions. For their CNN hyperparameter optimization, they employed Opposition-based Monarch Butterfly Optimization (OMBO). Transformers and MPNN, as well as CNN models, were tested on the BindingDB dataset. This model improves accuracy, sensitivity, and specificity. The main innovation of this approach is the use of an ensemble of a Transformer and a CNN, coupled with OMBO optimization, which increases the accuracy of DTI interaction prediction. However, optimizing CNN hyperparameters can be challenging, particularly in larger or more complex datasets where fine-tuning is necessary.

Shao et al. [37] introduced a DTIGCCN prediction model. This model abstracts latent interactions from the expression profiles of drugs and targets using GCNs and CNNs. To build an effective classifier, the abstracted features were concatenated together. DTIGCCN optimized the abstracted features, and drug–target correlation was employed entirely. For feature extraction in this study, spectral-based GCNs and CNNs were used from drug and target expression profiles. Extracted features were passed through a classifier. One of the novelties of DTIGCCN is that it utilizes GCN and CNN to extract various types of features from drug and target expression profiles, which can lead to more accurate predictions of drug–target interactions. However, model structure is an essential determinant of its output, and the need for high-quality graphs derived from Gaussian kernels for feature extraction would limit its performance when graph quality is low.

Chen and Peng [38] also introduced a sequence-based equivalent DTI prediction model called CHADTI. Drug features were extracted by stacking 1D-CNN layers, and proteins were represented using a pre-trained model (ESM). They introduced a new attention mechanism to learn drug–protein binding. Their experiments verified that CHADTI achieves state-of-the-art results on two

benchmark datasets. They also introduced a new attention mechanism, Cross-HyperAttention, to understand drug–target interactions. Cross-HyperAttention interactive features captured the semantic interdependency between drugs and proteins. The innovation in CHADTI is the inclusion of Cross-HyperAttention, which more effectively captures drug–protein interactions, thereby enhancing prediction performance. However, the use of precise and complete data for model training can impact model performance in the presence of noisy or missing data.

Xu et al. [39] combined drug interactions, side effects, and similarities into a heterogeneous network. To make predictions of DTIs, convolutional neural networks and random walks are employed. Furthermore, DTIPred also utilizes the topology of the heterogeneous networks along with several original features. For the prediction model, there were two sides for each drug–protein pair. The topological vectors of the nodes were learned by training a random walk with restart. Topological representations were constructed using convolutional neural networks to incorporate drug–protein interactions and similarities, enabling comprehension of the initial representation on the right-hand side. The innovation in DTIPred lies in its ability to integrate the topology of heterogeneous networks, along with drug–protein interaction features, thereby strengthening predictions. However, the model's complexity, particularly with the use of random walks and multiple categories of features, may make it computationally intensive, which hinders its ability to scale with large datasets.

Cheng et al. [40] proposed an end-to-end deep neural network approach (referred to as MHSADTI) for DTI prediction from a graph attention network and a multi-head self-attention mechanism. They initially decoupled protein properties from drug properties in the first step, using a graph attention network. For predicting drug interaction, they utilized attention scores. DTI was predicted using fully connected layers after computing feature vectors for proteins and drugs. MHSADTI enhanced the interpretability of DTIs with the self-attention mechanism. Graph attention networks also produced more efficient molecular features. Various cross-validation tests evaluated MHSADTI. Four datasets were used—human, C. elegans, DUD-E, and DrugBank—where their method outperformed existing methods in AUC, Precision, Recall, AUPR, and F1-scores. Their case studies also illustrated the interpretation of results from predictions in light of known biological information. A primary innovation of MHSADTI is its ability to combine graph attention networks with multi-head self-attention, thereby enhancing both interpretability and prediction performance. However, the ease of employing both graph attention and self-attention mechanisms may lead to an undue computational burden, especially in the context of large datasets.

Wang et al. [41] introduced a new method for Graph Convolutional Networks (GCNs) that incorporates multiple similarity and kernel fusion. The similarity fusion matrices are combined with Random Walk with Restart (RWR) and cosine similarity to extract the structural features of the network. They gained multilayer embeddings through GCN after learning low-dimensional embeddings. MKL replaced traditional GCN methods. The final step consisted of predicting DTIs using combinatorial kernels with Dual Laplace Regularized Least Squares. For gold-standard datasets such as E, IC, GPCR, and NR, they used tenfold cross-validation to validate the performance of their proposed model in predicting DTIs. Their method was also compared with other complex state-of-the-art approaches, and experiments proved that it outperformed others. For example, their method outperformed BICTR, eBICT, NRLMF, FLapRLS, MDMF2A, and FRnet on the IC dataset by 16.5%, 20.5%, 19.8%, 10.6%, 5%, and 50.6% in AUPR, respectively. Their method achieved higher AUCs and AUPRs, demonstrating its effectiveness compared to existing approaches. One of its limitations, however, is that it relies on similarity fusion matrices and multiple kernel learning, which can lead to higher computational complexity and longer training times.

GanDTI is a deep neural network model as developed by Wang et al. [42] for classification and binding affinity prediction. It used graphs and protein sequences. The model consisted of only a graph neural network (GNN), an attention module, and a Multi-Layer Perceptron (MLP). However, it outperformed state-of-the-art methods in the tasks of binding affinity prediction and interaction classification. They could provide a novel solution to improve performance by optimizing their model. The model consumed compound fingerprint data through a residual graph neural network and produced a vector for product-based attention projection onto protein sequences. In MLP processing, the two halves of the data were concatenated. Traditional approaches, such as LSTM and CNN, were not used. It performed better than state-of-the-art deep neural network methods on multiple metrics when tested on three public benchmarking datasets. The key innovation of GanDTI lies in the use of residual graph neural networks for processing compound fingerprint data, enabling more efficient learning of molecular properties. The attention mechanism also contributes to the model's robustness, particularly in enhancing the prediction of protein–drug interactions based on practical features. It outnumbers traditional models, such as LSTM and CNN, and opens up a new research direction for DTI prediction that can potentially enhance efficiency in both training and inference. However, one limitation of the approach is that it relies on residual graph neural networks and attention mechanisms, which, although extremely powerful, may require substantial computational resources, especially when handling large datasets or more complex protein–protein interaction networks. The

model's reliance on compound fingerprint data might also limit its generalizability in the absence of high-quality, comprehensive fingerprint data. The performance of GanDTI can be further influenced by the quality of protein sequence data, which may restrict its applicability to datasets with varying levels of data completeness.

CNN models

CNNs can predict DTI, because they can process structured data such as 2D molecular images and protein sequences. With an AUROC of 0.88, models like DeepDTI [10] are reviewed in Table 11, while Table 12 highlights their advantages (e.g., high accuracy with 2D data) and disadvantages (e.g., overfitting on small datasets). For generalization, CNN-based models such as DTIPred [39] and DTIGCCN [37] utilize large datasets to address local patterns.

Rayhan et al. [3] used autoencoders and convolutional neural networks to manipulate features. The FRnet-Encode and FRnet-Predict are two convolutional neural networks. For each instance, FRnet-Predict utilized 4,096 features generated by FRnet-Encode. They tested their method on four gold-standard datasets. In three out of four datasets, FRnet-DTI outperformed state-of-the-art methods in terms of ROC and PR areas. Additionally, 20 new drug targets were identified.

Sureyya Rifaioglu et al. [10] proposed a large-scale DTI prediction system called DEEPScreen based on deep convolutional neural networks for early drug discovery. DEEPScreen exhibited considerable superiority to conventional descriptors through the use of easily obtainable 2D structural representations as input. These 2D representations learn DEEPScreen complex features, producing highly accurate predictions. DEEPScreen was trained on 704 target proteins following exhaustive hyperparameter optimization experiments. They evaluated DEEPScreen in comparison to state-of-the-art methods on several benchmark datasets by molecular docking analysis and literature-based verification. DEEPScreen predicted JAK proteins as new targets of the approved drug cladribine in vitro through STAT3 phosphorylation. The system can be extended to drug discovery and repurposing for the creation of new DTIs that can be tested experimentally in the chemogenomic space. The innovation in DEEPScreen lies in the use of 2D structural representations, i.e., more straightforward yet highly effective methods for encoding complex molecular features. The drawback of the technique, nonetheless, is that it may not capture 3D structural information, which may be critical in successfully modeling complex interactions and predicting more diverse DTI scenarios.

Wang et al. [43] state that Large-scale Drug target Screening Convolutional Neural Networks (LDS-CNNs) can predict small-molecule–protein target interactions. They employed a unified encoding model to simplify the computation of different data formats within the model, facilitating feature abstraction and potential object prediction. In the evaluation of the proposed method, 898,412 interaction data points out of 8.8 billion records were analyzed. Results showed an Area Under the Curve (AUC) of 0.96, an Area Under the Precision–Recall Curve (AUPRC) of 0.95, and an accuracy of 90.13%. The experimental results demonstrated high precision in predicting drug–target interactions using the proposed method. LDS-CNN was able to predict various large datasets and data types, overcoming the challenges of encoding large-scale data. One of the innovations of LDS-CNN is that it can efficiently handle large-scale and multi-format data, simplifying the prediction process and reducing experimental time and cost. However, the model's reliance on large training sets and the complexity in handling various types of data could increase computational demands, making it less suitable for small-scale scenarios or environments with limited data available become a trending topic in predicting DTI, but with the challenge of designing lightweight learning frameworks that utilize drug discovery and drug repositioning. As DTI data have accumulated exponentially in recent years, deep neural network technology has become a trending topic in predicting DTI, but with the challenge of designing lightweight learning frameworks that utilize fewer protein descriptors. This study presents a Lightweight Deep Convolutional Neural Network (LDCNN) for DTI prediction, which utilizes amino acid sequences of varying lengths to generate a limited number of protein descriptors. In comparison to DeepConv, LDCNN reduced the number of neurons in convolutional layers and filters by 50%, resulting in a 1.3% loss in AUC and a 4% loss in AUPR. The innovation of LDCNN lies in its ability to reduce model complexity while maintaining a significant improvement in prediction performance. However, LDCNN's drawback is the trade-off between model size and prediction accuracy, which can compromise the model's performance for more complex or varied datasets.

Ghozzi et al. [45] proposed a DTI prediction model based on known DTIs, drug molecule substructure fingerprints, and target protein sequences. Drug–target feature vectors were extracted by concatenating Legendre moments and fingerprint binary vectors. CNN classifiers ultimately predicted DTIs. When compared with state-of-the-art classifiers in this field, the model was discovered to be superior. When combined with AutoML techniques, the proposed approach can be further enhanced. Ensemble techniques enhance the prediction accuracy by combining several independent classifiers. The innovation here is the coupling of Legendre moments and substructure fingerprints to create more potent drug–target feature vectors, thereby enhancing the model's predictive power. Its drawback is that it relies on high-quality substructure fingerprints, which may not be

readily available for all drug molecules; thus, its application may be limited to less-characterized molecules.

Yaseen et al. [46] proposed a multi-drug combination classification model using Convolutional Neural Networks (CNNs). Clinical reports and experimental data showed increasing inconsistencies when compared. There were also some outliers in every class, suggesting unusual pharmacological pathways or drug–drug interactions. This study employed machine learning and deep neural networks to develop a CNN-based system for predicting drug–target interactions. With an F1 score of 0.82 for the single model and validation accuracy of 96.72%, the CNN performed well on the DDI-Corpus dataset. Empirical evidence and clinical reports were disparate, suggesting atypical pharmacological mechanisms or drug–drug interactions. Furthermore, this study recognized the issue of class imbalance in the dataset and proposed resampling techniques to enhance the model's performance. The major novelty of this model is the use of resampling techniques to rectify class imbalance, which significantly improved its performance. However, the model's reliance on the DDI-Corpus dataset, which may fail to account for all types of drugs' pharmacological pathways or interactions, may limit its generalizability across all drug types.

Nandhini and Thailambal [47] found that the Coronavirus disease 2019 (COVID-19) resulted in clusters of fatal pneumonia clinically resembling SARS-CoV. Due to its varied forms and structures, the clinical implications of COVID-19 are more challenging to diagnose accurately. It was therefore crucial to predict the interaction between various drugs and the SARS-CoV target protein at that time, as it would have led to the discovery of new drugs for the disease. DTI prediction has now been extended to Deep Neural Network (DNN) techniques. Since CNN was a DNN model that could be used to generate predictive feature vectors or embeddings, they presented a CNN–OSBO encoder–decoder system to predict COVID-19 targets. To achieve the best output, they individually fed the drug input data and COVID-19 target input data through a Convolutional Neural Network (CNN) that features an Opposition-based Satin Bowerbird Optimizer (OSBO) encoder module. For CNN layers, OSBO was used to tune hyperparameters (HPs). To model a binding module, both encoded data were embedded. At last, the CNN decoder module predicted drug interactions against COVID-19 targets. The innovation here is the use of CNN with OSBO as a hyperparameter optimization method, which provides maximum prediction accuracy. The limitation is the reliance on specific SARS-CoV data, and the model's generalizability to other diseases or targets with less data could be limited.

Hu et al. [48] proposed a deep neural network-based prediction system, which creates new negative instances for detecting DTIs. They proposed a CNN model that distinguishes between potential drug–target interactions involved in drug repositioning. Concatenated descriptors of drugs and target proteins were mapped to a lower-dimensional subspace randomly. Their model achieved reasonable performance on three benchmark datasets. On the same benchmark dataset, CNN grasped more subtle and detailed information. Deep neural networks have proven to be a successful approach in most biological fields, including the prediction of drug synergistic effects, non-coding RNAs, and drug targets. They also introduced a reasonable strategy for selecting negative examples, which reduced the possibility of false negatives. The innovation in this model is the generation of novel negative examples for the effective prediction of DTI, addressing an open issue in DTI research. Yet, the limitation is that it performs a random projection of drug–target descriptors, which can oversimplify complex interactions and limit the model's performance in more complicated scenarios.

Transformer models

CAT-DTI [14] achieved an AUROC of 0.983 due to the existence of long-range dependencies in DTI prediction. Table 14 summarizes the advantages (e.g., handling sophisticated data) and disadvantages (e.g., requiring extensive tuning) of this approach. MocFormer [15] and FragXsiteDTI [17] added interpretability and flexibility.

Zeng et al. [14] stated that despite the recent progress being made in Drug–Target Interaction (DTI) prediction, numerous issues remain. DTI prediction begins with training on protein features, drugs, and drug–protein interactions. It is also necessary to generalize the model more effectively in practical scenarios. These were addressed by CAT-DTI, a cross-attention and Transformer model capable of performing domain adaptation. Aside from out-of-distribution data, CAT-DTI was able to learn drug–target interactions. A Convolutional Neural Network (CNN) hybrid with a Transformer-based neural model was used to represent amino acid distance in protein sequences. Concurrently, a cross-attention module was used to mine drug–target interaction features. The CDAN was used to transfer DTI representations between different distributions to novel DTI scenarios. CAT-DTI outperformed state-of-the-art methods in terms of DTI prediction performance in both in-domain and cross-domain scenarios. A Graph Convolutional Network (GCN) was used to extract drug features, while CNN-Transformer models were used to extract protein targets. To extract more crucial DTI features, the cross-attention module enhances the bidirectional feature interactions between drugs and proteins, thereby improving the model's accuracy. CAT-DTI demonstrated strong adaptability and prediction capabilities in cross-domain tasks using CDAN. Through experiments, CAT-DTI outperformed both state-of-the-art and conventional machine learning models in both in-domain

and cross-domain settings. The strength of CAT-DTI lies in its ability to capture complex dependencies and interactions using cross-attention and Transformer-based architectures, thereby making it more accurate and generalizable across various datasets. However, its shortcoming is its high computational cost and fine-tuning complexity, which may pose an issue for real-time use or when resources are limited.

Zhang et al. [15] noted that general deep neural network strategies suffer from low-confidence predictions due to the complexity of proteins and drugs, as well as bias and sparsity in labeled data. For drug–target interaction prediction, they proposed a two-stage pre-trained model. This platform addressed drug diversity by reducing prejudice and enhancing prediction performance using pre-trained models of molecules and proteins. During the second phase, a Transformer-based neural network with bilinear pooling and a fully connected layer permitted feature vector-based prediction. To evaluate the efficacy of the framework, DrugBank and epigenetic regulator datasets were utilized. Apart from accuracy, the proposed framework outperformed state-of-the-art methods in terms of the area under the receiver-operating characteristic curve, recall, and the area under the precision–recall curve. The architecture effectively processed drug and protein diversity, as well as data bias. Quantitative and qualitative analyses of DrugBank and epigenetic regulator databases were satisfactory. Two stages of processing were employed, and optimization was required at each stage in turn. The innovation here is the two-stage pre-trained scheme, which reduces data biases and enhances the accuracy of predictions, especially for complex molecular systems. The disadvantage of this scheme is the risk of inefficiency when handling enormous datasets, because each stage requires unique optimization, which could increase the total computational cost.

Khodabandeh Yalabadi et al. [17] concluded that model interpretability and performance optimization remain challenging. To address these challenges, FragXsiteDTI was proposed, a deep neural network based on the Transformer. For the first time, FragXsiteDTI employed parallel drug molecule fragments and protein pockets. Protein and drug interactions were studied by their information-rich representations. This perceiver IO (Transformer-based neural architecture) model features a learnable latent array that initially interacts with protein-binding site embeddings via cross-attention, and then undergoes fine-tuning with self-attention to locate drug fragments through a cross-attention block within the Transformer architecture. By free-flowing translation of information, the learnable query array preserved crucial nuances in drug–protein interactions. On three benchmarking sets, the model outperformed some state-of-the-art models, demonstrating better predictive power. Moreover, they showed the explainability of their model for the most essential elements of the target proteins as well as drug molecules. One of the standout features of their method was its ability to determine which drug fragment targeted what region on the protein. Global maps of interaction hotspots provided researchers with an overall map to guide drug optimization and modifications. Attention modules also offered insight into the model's decision-making process and a superior mechanism for performance. Pharmacologically active fragmental scaffolds were identified by attention scores. Model-based data enabled drug designers to make more informed decisions and potentially sidestep some trial-and-error attempts. The advantage of FragXsiteDTI is that it can utilize both drug fragments and protein pockets simultaneously, providing an in-depth understanding of drug–protein interactions. Its only disadvantage is that the model's reliance on the learnable latent array and cross-attention modules renders it computationally expensive and potentially difficult to scale for large datasets.

Wang et al. [49] recognized that most recent drug–target interaction research has been conducted on individual nodes of drugs or targets in isolation, without considering their connectivity. This article proposes a way that captures drug–target relationships as multi-level relational information as opposed to regular node-based representations. Seven levels of relationships were employed in building drug–target interactions. This interaction was then represented in a line graph. The second part involved designing a graph Transformer-based deep neural network to predict drug–target interactions. Based on Jaccard similarity and random walks with restarts, the authors initially created features from seven levels of drug-to-target relationships. Subsequently, a line graph was used to convert drug–target interactions into links. For prediction, a graph Transformer-based deep neural network was designed. For comparison purposes, AUROC and AUPR values were employed. During the experiment, the DTI network revealed a novel pattern for understanding drug–target interactions. The key innovation in this approach lies in utilizing multi-level relational data to describe drug–target interactions, thereby enabling a deeper understanding of the interaction landscape. Yet, the drawback is that the process might be computationally intensive to handle large-scale data with multiplex relational levels, making it less favorable for real-time processing.

Zhao et al. [50] noted that fewer techniques have attempted to address the cold-start issue, as most existing techniques only represent existing interactions, not those unique to new drugs or targets with few interactions. They proposed the MGDTI graph Transformer-based deep neural network method, a meta-learning-based method. To mitigate the shortage of interaction, drug–drug and target–target similarity were employed as side information. Additionally, long-range dependencies were also addressed by a graph Transformer-based deep neural network, thereby avoiding oversmoothing. Experimental results on a benchmark dataset

indicated that their proposed MGDTI was effective in DTI prediction. The novelty in MGDTI lies in its use of meta-learning to address the cold-start problem and its ability to incorporate similarity-based data for enhanced predictions. However, the limitation is that the model cannot process datasets that lack rich drug–drug or target–target similarity data, which will impact prediction precision in such scenarios.

Simon & Bankapur [51] developed BERT to discover protein–drug interactions. Additionally, they emphasized the importance of data transformation and data preprocessing to make machine learning algorithms operational. This approach generates features from protein sequences and SMILES. Random forests can be used for classifying drug–target interaction pairs with these features. It relies on deep representations (proteins and drug SMILES) as opposed to molecular descriptors for predicting DTI. To predict drug–target interactions, this model did not require molecular descriptors based on wet lab experiments. Integrating feature models and classification models can increase the precision of the prediction. Bidirectional encoder representations of the BERT model from this work were applied to supplement DTI prediction by protein SMILES data. The BERT model could predict drug–protein interactions without a descriptor dataset. The innovation in BERT for DTI prediction lies in its ability to predict drug–target interactions without relying on molecular descriptors from wet lab experiments, but rather using only deep representations of proteins and SMILES. However, its accuracy will also be highly dependent on the quality and comprehensiveness of SMILES data. It may render this data of limited use where it is of poor or incomplete quality.

GNN models

According to Table 15, GNNs effectively model graphical structures, with HiGraphDTI [52] reaching an AUC of 0.910. Table 16 shows advantages and challenges (e.g., capturing topological features). CSDTI [29] and DSG-DTI [53] supported unsupervised learning of DTI ("Research Questions and Their Motivations" section, Question 9).

According to Liu et al. [52], deep neural network models perform better in DTI prediction, since they can extract expressive and strong features from the chemical structures of targets and drugs. Deep neural network methods do not use chemical properties that are transferable by motifs. Molecular representation at the atom–drug level and motif level cannot utilize structural information. Additionally, sequential model-based target feature extraction is computationally costly and biased toward limited contextual information. Atoms, molecules, and motifs are combined in HiGraphDTI to describe chemical information at three distinct levels. Expressive target features are learned using attentional feature fusion in this module. To provide a complementary understanding of interactions, the hierarchical attention mechanism was used. HiGraphDTI performed better than state-of-the-art approaches in explaining interactions and predicting new DTIs. Four datasets showed that the model outperformed six baselines in terms of AUC and AUPR. HiGraphDTI can also enable new drug discovery from attention-weighted visualizations. The breakthrough of HiGraphDTI lies in its hierarchical attention mechanism, which incorporates multi-level chemical information (atoms, molecules, and motifs), allowing it to capture more complex drug–target interactions. However, the disadvantage of HiGraphDTI is the computationally intensive process of processing multi-level chemical data, which would make it challenging to implement on large scales or in real-time systems.

According to Li et al. [53], although deep neural networks are capable of predicting drug–target interactions, some difficulties remain to be addressed: (1) drug and protein interactions are complex; (2) intermediate nodes are not calibrated. DSG-DTI addressed such problems. To complete matrices, DSG-DTI utilizes heterogeneous graph autoencoding. There exist high-dimensional spaces of drugs, targets, side effects, and diseases. Additionally, attention-based heterogeneous graph-based matrix completion has been shown to identify long-range dependencies. The first step is to auto-learn a latent node representation from entity-relationship graphs. Second, in GNODE, they proposed a heterogeneous graph convolution layer to combat oversmoothing. For end-to-end training of the model, integration of GNODE into the optimization framework was required. A public benchmark verified their model. Out of two publicly available benchmark apps, the proposed scheme showed better performance compared to other baselines. In summary, it slightly lowered performance for new targets and registered drugs. Drug–target interactions are a critical process in drug repurposing and discovery. Big data and deep neural networks were combined to establish the HGMC model. It outperformed other baselines in experiments. Graph Neural Ordinary Differential Equations (GNODE) coupled with heterogeneous graph attention networks improved performance. The breakthrough in DSG-DTI lies in its use of heterogeneous graph autoencoding to address the complexity of drug–target interactions and the issue of uncalibrated intermediate nodes. The limitation is the decrease in performance on newly registered targets and drugs, indicating that the model may struggle to make predictions for new or underrepresented data points.

Wang et al. [54] clarified that previous virtual screens only considered the sequence or graph structure of the drug but not its subgraph structure. Drug subgraph fingerprints are derived from drug subgraphs in this manner. Graph neural networks utilized fingerprint vector data to represent the spatial structure of drug subgraphs. In the final step, the subgraph attention mechanism updated protein target

knowledge, allowing the model to extract more information. Experimentally, the model outperformed the state-of-the-art model in every respect. The model also predicted affinity regressions and classifications. Target sequence information was extracted through semantic segmentation. Protein target information matrix weights were redistributed through subgraph attention. Next, the drug and target feature vectors were contrasted. Using two classification prediction sets and one affinity prediction set as the dataset, the model demonstrated generalizability. The final prediction performed well both at classification and affinity regression. The novelty of this approach lies in its application of drug subgraph fingerprints and subgraph attention mechanisms to learn more informative features from drugs, thereby improving the accuracy of DTI prediction. The drawback, however, is reliance on high-quality subgraph representations, which in some cases may not be readily available for all drugs, potentially affecting the model's performance when subgraph information is noisy or incomplete.

Graph Convolutional Networks (GCN) is a strong deep neural network method for complex networks, as stated by Wang et al. [55]. The approach processes node features and structure data simultaneously by extending the convolution operation from Euclidean space to non-Euclidean space. For end-to-end learning, nodes in a mapping graph combine their features with those of neighboring nodes using a mapping function. In this paper, a method named Decagon for GCN link prediction is proposed. Experiments were performed with DTI-net. The DTI-net provides drug–drug interactions, target–target interactions, and drug–target interactions. DTI heterogeneous networks were constructed using drug features and target features in the attribute property form. This paper compares multiple-parameter selection and optimization experiment predictions to enhance the performance of drug–target relationship prediction. With the ability to efficiently extract features from heterogeneous networks, the GCN algorithm is suitable for drug–target interaction prediction. Decagon's innovation lies in its ability to push GCNs beyond Euclidean space, thereby enabling it to handle complex interactions and relationships involved in drug–target prediction more effectively. A disadvantage of this approach is that the model's reliance on non-Euclidean space and higher-order graph structures can lead to computationally expensive operations, particularly with large-scale data and complex topologies.

DNN models

DNNs extract high-level features, as reviewed in Table 17, with OverfitDTI [2] achieving a CI of 0.9480. Table 18 notes advantages (e.g., utilizing all data) and disadvantages (e.g., computational demands). This flexibility supports hybrid integration ("Research Questions and Their Motivations" section, Question 3).

Xiaolin et al. [2] proposed OverfitDTI, a novel framework for resisting overfitting in deep neural network models. Deep Neural Networks (DNN) in OverfitDTI are deliberately overfit to learn the chemical space of drugs and the biological space of targets. Trained DNN weights implicitly represented drugs and targets. Three public datasets demonstrate that OverfitDTI effectively captures this non-linear relationship using overfit DNN models. Fifteen compounds against TEK, a receptor tyrosine kinase involved in vascular homeostasis, with AT9283 and dorsomorphin experimentally confirmed as TEK inhibitors. OverfitDTI outperformed state-of-the-art baselines and enhanced the model's generalization ability even in cold starts. The innovation of OverfitDTI lies in its deliberate overfitting approach, which is utilized to learn more complex, non-linear drug–target relationships that other models do not capture. This new approach enables the model to generalize even in cases where sparse data or cold-start problems are present. The limitation of this model is that, although strategic overfitting is being leveraged, it may still be challenging to scale up to much larger, more diverse datasets, especially when there are more complex drug–target interactions or extremely high-dimensional chemical spaces.

Comprehensive analysis and synthesis of neural network models for DTI prediction

This subsection provides a detailed analysis and comparison of the neural network models being reviewed for predicting Drug–Target Interactions (DTIs). The overview covers Hybrid models ("Hybrid Models" section), Convolutional Neural Networks (CNNs, "CNN Models" section), Transformer models ("Transformer Models" section), Graph Neural Networks (GNNs, "GNN Models" section), and Deep Neural Networks (DNNs, "DNN Models" section). The intention here is to provide the performance of the models and highlight their general strengths and weaknesses under various circumstances.

Comparative analysis of model performance

The models being compared employ diverse architectures to address the difficulties in DTI prediction, each with its distinct merits. Based on Tables 9, 10, 11, 12, 13, 14, 15, 16, 17, and 18, hybrid models, such as HGDTI [4], GAEMS-DNN [11], and MSH-DTI [23], integrate multimodal data (e.g., molecular graphs, SMILES sequences, and protein interactions) to achieve high AUROC and AUPR values (typically > 0.95) on databases such as DrugBank, DTINet,

and BindingDB. These models excel in complex environments due to their ability to handle heterogeneous data sources.

CNN-based models, such as DEEPScreen [10] and LDS-CNN [43], excel in handling two-dimensional (2D) structural data (e.g., molecular images) and achieve AUCs ranging from 0.88 to 0.96. These models are computationally lightweight and can be used in environments with limited resource availability, but are hindered by their inability to handle three-dimensional (3D) information or long-range dependencies. Transformer models, such as CAT-DTI [14] with an AUROC value of 0.983 and FragXsiteDTI [17], utilize attention mechanisms to capture long-range dependencies and complex interactions more effectively, thereby offering significant biological interpretability.

GNN models, such as HiGraphDTI [52] with an AUC value of 0.91 and DSG-DTI [53], are best suited for modeling graph structures and topological interactions in heterogeneous data environments. Finally, DNN models, such as OverfitDTI [2], with a Concordance Index (CI) value of 0.9480, emphasize non-linear relationships and work effectively in cold-start situations; however, they require large datasets and substantial computational power.

A comparison of performance indicates that hybrid and transformer models outperform CNN and DNN models for most measurements (e.g., AUPR, AUROC, and accuracy), as they can handle multimodal data and learn intricate dependencies. However, CNN models, due to their simplicity and computational efficiency, are more suitable for scalable use or when limited data is available. GNN models perform better when a large amount of graphical data are available.

Common strengths

Advanced feature extraction: All models employ advanced techniques, including attention mechanisms (e.g., CAT-DTI [14], HGDTI [4]), graph-based learning (e.g., HiGraph-DTI [52], DSG-DTI [53]), and multimodal encoding (e.g., MCL-DTI [16], GAEMSDNN [11]), to extract robust features from drugs and proteins, thereby enhancing prediction accuracy.

Data flexibility: Hybrid, GNN, and transformer models are very flexible in handling heterogeneous data (e.g., drug–drug interactions, protein–protein interactions, and multimodal data), which promotes robustness in different databases.

Biological interpretability: FragXsiteDTI [17] and HiGraphDTI [52] models offer insightful biological understandings of drug–target relationships through attention-based visualizations, facilitating targeted drug design.

Handling complex scenarios: Transformer and DNN models, as well as CAT-DTI [14] and OverfitDTI [2], excel in cold-start situations or with scarce data, thereby enhancing generalizability.

Common weaknesses

High computational resource requirement: The majority of models, including hybrid (e.g., GIFDTI [25], MSH-DTI [23]), transformer (e.g., CAT-DTI [14], MocFormer [15]), and GNN models (e.g., HiGraphDTI [52], DSG-DTI [53]), require high computational resources due to complex architectures and the requirement for fine-tuning, which is difficult in real-time or low-resource settings.

High-quality data reliance: Models that rely on multimodal or graph data (e.g., MCL-DTI [16], HGSDTI [28], and SubgraphDTI [54]) may work suboptimally if high-quality or complete data (e.g., molecular images, drug–drug similarities, or graph structures) is lacking.

Generalization problems: Certain models, such as DSG-DTI [53] and OverfitDTI [2], struggle to predict results for new drugs or targets (cold-start settings) or when dealing with imbalanced data, highlighting limitations in generalizability.

Implementation difficulty: Complex models, such as FragXsiteDTI [17] and MRCH-GCAE [31], which utilize sophisticated architectures (e.g., multilayer transformers or heterogeneous graphs), require precise tuning and specialized expertise for implementation.

Analysis and comparisons

In this section, 44 articles are analyzed using neural network-based techniques to predict drug–target interactions, categorizing them across multiple dimensions to identify trends, preferences, and gaps. This analysis addresses research questions from "Research Questions and Their Motivations" section (e.g., Questions 1, 7, 12) and provides a basis for future directions in "Conclusions" section.

Taxonomy based on publishers

Figure 5 presents the publisher distribution of articles, where IEEE Xplore leads with 54.55%, followed by Springer at 36.36%, ScienceDirect at 6.82%, and RSC at 2.27%. This distribution can most likely be attributed to the focus of IEEE Xplore on bioinformatics and computational biology conferences, particularly ones like BIBM, which are predominantly responsible for the interdisciplinary nature of Drug–Target Interaction (DTI) prediction. This multidisciplinary focus in IEEE Xplore has contributed to its dominance in publishing articles related to DTI. In contrast, Springer also contains an enormous proportion of published articles, perhaps because of the more specialized focus and

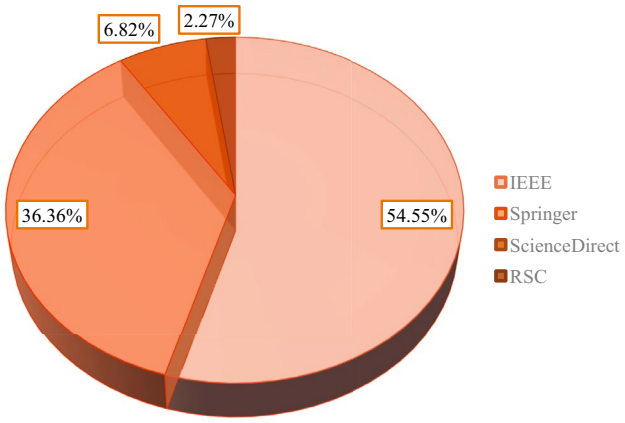


Fig. 5 Distribution of articles by publisher

stricter peer-review requirements of its journals. In particular, BMC Bioinformatics (Table 19), one of the leading outlets for DTI research, accounts for a large proportion of Springer's publications. From Tables 19, 20, the field favors a combination of two publication approaches: on the one hand, articles are published in peer-reviewed journals that offer a high scientific reputation, and on the other hand, rapid publication through conference proceedings also plays a significant role. The 50–50 split between journals and conferences reflects this dual strategy, whereby scholars prefer to seek a balance between rapid publication and maintaining high scientific quality.

Table 19 The most relevant journals

Journal	Number of papers	Article numbers
BMC Bioinformatics	7	[4, 14, 16, 22–24, 49]
IEEE/ACM Transactions on Computational Biology and Bioinformatics	7	[11, 25, 39–41, 48, 53]
iScience	1	[2]
Heliyon	1	[3]
BMC Biology	1	[21]
International Journal of Computational Intelligence Systems	1	[15]
Chemical Science	1	[10]
Applied Intelligence	1	[29]
Health Information Science and Systems	1	[43]
Computational Biology and Chemistry	1	[42]

Table 20 The most relevant conference

Conference	Number of papers	Article numbers
IEEE International Conference on Bioinformatics and Biomedicine (BIBM)	8	[28, 30–34, 44, 50]
Advanced Data Mining and Applications (ADMA)	1	[54]
Machine Learning and Knowledge Discovery in Databases (ECML PKDD)	1	[52]
Intelligent Computing Theories and Applications (ICIC)	1	[26]
Neural Information Processing (ICONIP)	1	[27]
International Conference on Control, Decision and Information Technologies (CoDIT)	1	[45]
COMSNETS	1	[51]
International Conference on Computer-Supported Cooperative Work in Design (CSCWD)	1	[35]
International Congress on Human–Computer Interaction, Optimization and Robotic Applications (HORA)	1	[46]
IEEE International Conference on Bioinformatics and Computational Biology (ICBCB)	1	[55]
International Conference on Electronics, Communication and Aerospace Technology (ICECA)	1	[47]
International Conference on Trends in Electronics and Informatics (ICOEI)	1	[36]
IEEE International Conference on Tools with Artificial Intelligence (ICTAI)	1	[37]
IEEE International Conference on Systems, Man, and Cybernetics (SMC)	1	[38]
RECOMB	1	[17]

Taxonomy based on the year of publication

Figure 6 depicts the time-series distribution of DTI prediction articles from 2020 through 2024. The plot reveals an exponential rise in the use of neural network-based methods in the recent 5 years, clearly addressing Research Question 1 (the best-performing architectures). Six articles were published in 2020—three CNN-based and three hybrid-based. This led to the increased use of CNN on structured data, such as 2D molecular images, discussed in "CNN Models" section. Due to the novelty of such data structures for DTI prediction, models including Transformer-based, GNN-based, and DNN-based had not yet been utilized in 2020. The number of publications rose to six (four hybrid, one CNN, and one GNN) in 2021, marking the beginning of a boom in GNNs for molecular graph modeling. The rise in GNN usage reflected the shift toward models capable of handling complex relationships between molecular components. Seven papers were published in 2022 (five hybrid, one CNN, and one Transformer), reflecting the growing popularity of hybrid models and the application of Transformers, which are well suited to handle long-range dependencies in molecular interactions. In the year 2023, the number of papers increased to 11 (5 hybrid, 3 CNN, 1 Transformer, 1 GNN, and 1 DNN). This is indicative of the broad embrace of various models as researchers sought more flexible alternatives in DTI prediction. Fourteen papers (eight hybrid, four Transformer, and two GNN) in 2024 indicated that the prevalence of hybrid models persisted, as did the popularity of Transformer and GNN models capable of processing complex, multidimensional data. Hybrid models, which represented 57% of the papers in 2024, demonstrate a strategic emphasis on integrative and adaptive design approaches. For instance, models like HGDTI (AUROC 0.979) outperform earlier DNN-based models, such as DeepDTI (AUROC 0.88), as indicated in Table 20. The recent trend of hybrid and Transformer-based models will continue to evolve, incorporating advances in Explainable AI, as outlined in "Conclusions" section.

From Fig. 7, it is clear that the reference list of papers is equally split between academic conferences and highly regarded academic journals, each at 50%. This balanced split suggests that both journal articles and conference proceedings are equally popular means for publishing studies on Drug–Target Interactions (DTIs). Such a pattern highlights the importance of both peer-reviewed publications based on rigorous research and the rapid dissemination of results through conference proceedings. Tables 19 and 20 give a complete list of the conferences and journals included in the study. Some of the key journals, such as BMC Bioinformatics, IEEE/ACM Transactions on Computational Biology and Bioinformatics, and Chemical Science, are often quoted. Some of the key international conference proceedings, such

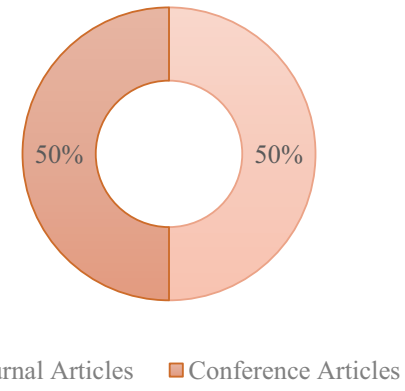
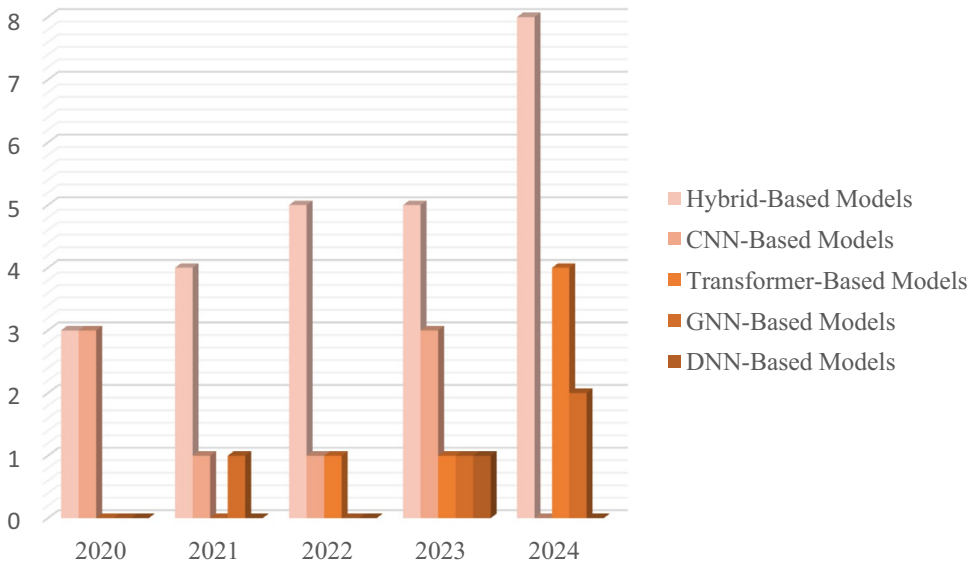


Fig. 7 Distribution of articles across journals and conferences

Fig. 6 Yearly distribution



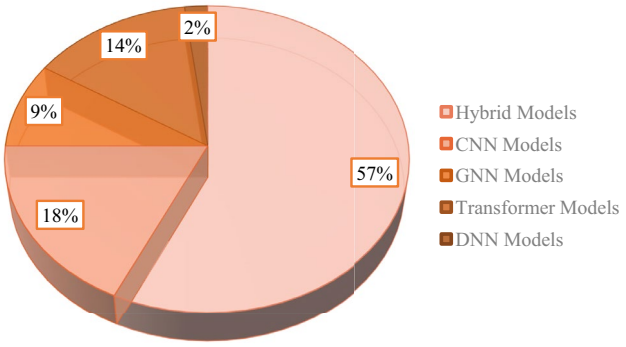


Fig. 8 Distribution of articles by neural network architecture

as the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), are also significant sources of DTI-related research.

Taxonomy based on architecture

The architecture-based classification of the 44 articles, illustrated in Fig. 8 (a pie chart), categorizes neural network models into five distinct categories, addressing Research Question 1 (top architectures) and Question 7 (evaluation of performance). As one can see from the figure, the hybrid models are the most frequent among the distribution, with 57% of the papers. Such superiority reflects the growing practice of combining different architectures, for instance, CNNs, GNNs, and Transformer models, to cope with the burgeoning complexity of DTI prediction. Hybrid models, such as HGDTI with an AUROC of 0.979, work efficiently, because they can leverage the strengths of different architectures in encoding complex molecular interactions. CNN-based models, which constitute 18% of the papers, are highly suitable for structured data, such as 2D molecular images. However, even though they are robust in image-related tasks, their performance is inferior compared to hybrid models, such as DeepDTI, with an AUROC of 0.88. GNN-based models, which constitute 9% of the papers, are suitable for graph-structured data, such as molecular networks, and are reviewed in "GNN Models" section. For example, HiGraph-DTI (AUC of 0.90) is particularly good at capturing the graph-level relationships among molecules. Transformer-based models, which constitute 14%, are compelling in capturing long-range dependencies of molecular structure. These models, for example, CAT-DTI (0.983 AUROC), demonstrate a promising ability to handle complex interactions that are difficult to capture using traditional methods. DNN-based models, which contribute only 2% of the articles, are less common due to their more straightforward structure, which is less capable of capturing the complexity of DTI interactions as advanced models are. This distribution highlights the increasing prominence of advanced, integrative architectures for DTI prediction, with a strong trend toward hybrid and Transformer-based models, which are best suited for more complex and high-dimensional data.

Taxonomy based on environment

Researchers rely on a variety of sources based on their analysis of the articles surveyed. Classifying models based on the databases they utilize reveals the most commonly used databases. This classification is presented in Table 21. The DrugBank database is utilized in 21 out of 44 articles, likely due to its comprehensive and accessible data. Nine articles utilize the Human database, which is used for its biological data on humans. Davis is cited in 7 articles as a means to validate models. BindingDB is also referenced in 6 articles, and KEGG in 4. DUD-E, KIBA, and the Gold Standard Datasets are referenced in fewer articles, with all of them being cited in only three articles, indicating that they are used for case-specific or testing purposes. Some of these articles have utilized a combination of these databases in varying proportions, indicating a preference for using more diverse and comprehensive data to achieve higher prediction accuracy. For the practical training and assessment of DTI prediction models, reliable and varied databases are preferable.

Taxonomy based on evaluation metrics

In "Taxonomy based on Evaluation Metrics" section, the evaluation metrics for Drug–Target Interaction (DTI) prediction models across 44 articles are analyzed; Table 22 indicates that AUROC (used in 37 articles, e.g., [49] with 0.9973), AUPR (used in 31 articles, e.g., [16] with 0.989), and Accuracy (used in 24 articles, e.g., [41] with 0.999) are the most common metrics. The distribution is

Table 21 Number of articles per database

Database	Number of articles
DrugBank	21
Human	9
Davis	7
BindingDB	6
KEGG	4
KIBA	4
C. elegans	3
DUD-E	3
Gold standard	3
Other databases	1 or 2

Table 22 Metrics for neural network-based

Article	Dataset	Accuracy	AUROC	AUPR
[2]	DTC	–	–	–
[3]	Enzymes	0.95 <	0.9754	0.70
[4]	DrugBank		0.979	0.961
[10]	Random-Split Dataset	0.87	0.88	0.85
[11]	DrugBank, NR (Nuclear Receptors)	0.8935 0.8819	0.8672 0.8556	0.7125 0.6308
[14]	Human BindingDB	0.942	0.983	0.976
[15]	DrugBank, Epigenetic-regulators	0.834 0.596	0.912 0.661	0.910 0.645
[16]	Human, C. elegans	–	0.987 0.992	0.989 0.994
[17]	Human, C. elegans, DrugBank	–	0.991	–
[21]	DrugBank	0.8955	0.9548	0.9644
[22]	Davis KIBA	–	–	–
[23]	DrugBank	–	0.9620	0.9605
[24]	Wang et al. dataset	0.878	0.978	0.964
[25]	DrugBank, Davis, KIBA, BindingDB	0.823	0.905	0.912
[26]	GPCR, Ion Channels	–	0.9798 0.9896	0.9689 0.9889
[27]	Human, C. elegans	–	0.98 0.97	0.94 0.91
[28]	DTI, SIDER, CTD	–	0.9623	0.9674
[29]	DrugBank, Human	–	0.902 0.982	–
[30]	Drug-Protein Interaction, Drug–Drug Interaction, Protein–Protein Interaction, Drug-Disease Association, Protein-Disease Association, Drug-Side-Effect Association	–	0.963	0.884
[31]	DTI, DDI, PPI	–	0.9805	0.8416
[32]	Enzyme, GPCR	–	0.9864 0.9632	0.9757 0.9614
[33]	DrugBank, KEGG, BRENDA	0.983	0.99	0.994
[34]	DrugBank, Biosnap, Davis	0.865	0.917	0.897
[35]	Davis, KIBA	0.8739	0.9225	0.8395
[36]	BindngDB	0.81	–	–
[37]	A375	0.9113	0.9151	–
[38]	DrugBank, Davis	0.836	0.905	0.908
[39]	DrugBank, Side Effect Database	–	0.968	0.462
[40]	Human, DrugBank	–	0.9822 0.8628	0.9568 0.8351
[41]	enzymes	0.999	0.995	0.968
[42]	DUD-E, Human	–	0.997 0.983	–
[43]	DrugBank, STITCH	0.9013	0.962	0.959
[44]	DrugBank	0.79	0.85	0.80
[45]	DrugBank	0.824	–	–
[46]	DDI-Corpus	0.9672	–	–
[47]	Custom Dataset, DrugBank	0.87	–	–
[48]	DrugBank	0.9800	0.9965	
[49]	Peng Dataset	–	0.9973	0.9976

Table 22 (continued)

Article	Dataset	Accuracy	AUROC	AUPR
[50]	DrugBank, Biosnap, KEGG	–	0.936 (cold-target task at mask-rate 0.5)	0.940 (cold-target task at mask-rate 0.5)
[51]	DrugBank, KEGG, PubChem	0.75	–	–
[52]	Human, BindingDB	–	0.985 0.954	0.988 0.955
[53]	DTI-HN	–	0.9568	0.8629
[54]	DUD-E, Human, BindingDB	Human: 0.964	DUD-E: 0.999 Human: 0.985 BindingDB: 0.980	–
[55]	DTI-net	0.8464	0.9186	0.9221

confirmed by Fig. 9, emphasizing discrimination precision and handling imbalanced data; hybrid models (e.g., [4] with AUROC 0.979) and Transformer architecture-based models (e.g., [14] with AUROC 0.983) in Table 22 outperform CNNs (e.g., [10] with AUROC 0.88) and GNNs (e.g., [52] with AUROC 0.910). Although Table 23 details secondary metrics, such as F1 (in 18 articles, e.g., [48], with a value of 0.9798) and Recall (in 21 articles), the limited reporting of CI and MCC (in 2 and 6 articles, respectively) highlights the need for standardization. Although Tables 22 and 23 present a broad overview of model performance, the metrics are reported under varied experimental conditions and on different datasets, limiting the fairness of direct comparisons.

Standardized classification of metrics by dataset

In this section, a uniform categorization of evaluation metrics for Drug–Target Interaction (DTI) prediction models based on the datasets used in the reviewed studies is introduced. Since some articles have used multiple datasets, the most commonly used and widely applied datasets in these studies were selected. The three datasets, DrugBank, Human, and Davis, are the most frequently used in the reviewed articles; therefore, these three datasets were used for categorizing the evaluation metrics. As shown in Tables 24, 25, and 26, performance metrics, including Accuracy, AUROC, and AUPR, are categorized based on the datasets used in the experiments. This categorization aims to enhance fairness and comparability of model performance under similar experimental conditions. Table 24 provides a summary of model evaluation scores on the

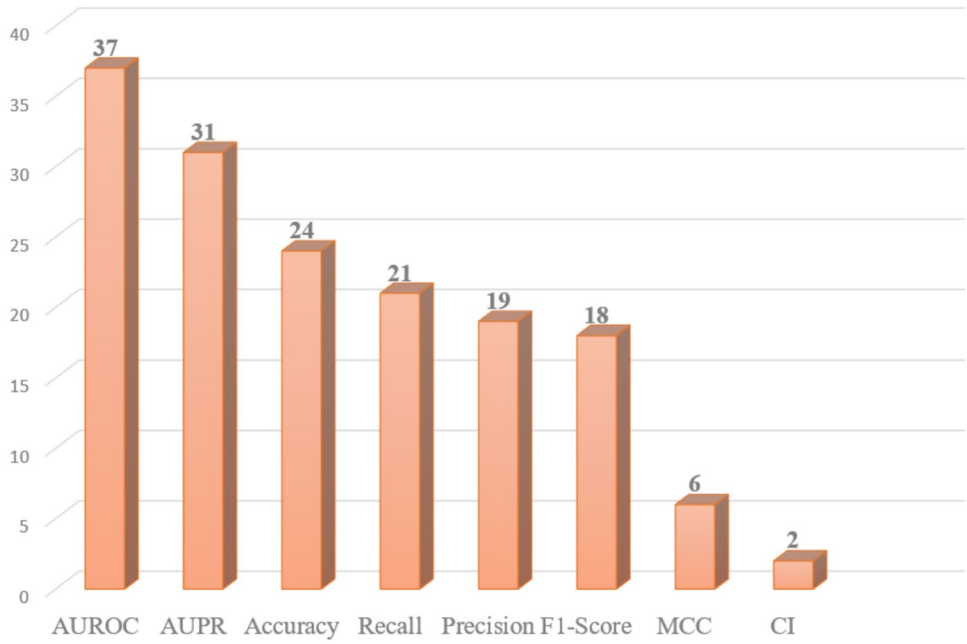


Fig. 9 DTI metrics distribution

Table 23 Detailed secondary performance metrics

Article	CI	F1 score	MCC	Recall	Precision
[2]	0.9480	–	–	–	–
[10]	–	0.87	0.74	–	–
[11]	–	–	0.90	–	–
[14]	–	0.944	–	–	–
[15]	–	–	–	0.832	–
[16]	–	–	–	0.961	–
[17]	–	0.964	–	0.952	0.977
[21]	–	0.8932	0.8013	–	–
[22]	DAVIS:0.8702 KIBA:0.73	–	–	–	–
[24]	–	0.8932	0.8013	–	–
[25]	–	0.830	–	0.86	0.801
[26]	–	–	–	0.80	0.83
[27]	–	0.83	–	0.84	0.88
[28]	–	0.82	0.78	–	0.79
[29]	–	0.86	0.84	0.87	0.89
[32]	–	0.9501	–	0.9135	0.9614
[34]	–	–	–	0.848	0.823
[35]	–	–	–	0.7814	0.7787
[36]	–	0.91	–	0.93	0.89
[37]	–	–	–	–	–
[38]	–	–	–	0849	0.836
[39]	–	–	–	(top 30) = 0.849 (top 90) = 0.957 (0.981) = 0.981	–
[40]	–	0.9346	–	0.9365	0.9472
[41]	–	0.929	–	0.925	–
[42]	–	–	–	0.933	0.90
[45]	–	0.9023	–	0.9186	0.8866
[46]	–	0.82	–	–	–
[47]	–	0.65	–	0.92	0.46
[48]	–	0.9798	–	–	0.9887
[51]	–	0.72	–	0.75	0.80
[52]	–	–	–	0.86	0.84
[54]	–	–	–	Human:0.964	Human:0.961

DrugBank dataset, Table 25 includes evaluations on the Human dataset, and Table 26 presents evaluation scores on the Davis dataset.

However, it is worth noting that these results were not obtained under identical experimental conditions, and there may be differences in the experimental setup or parameters. Therefore, for a more accurate evaluation and a better understanding of how the use of these datasets impacts the results, it is recommended to refer to the original articles. This uniform categorization facilitates easier comparison between models evaluated on the same datasets, enabling a clearer understanding of how dataset selection affects the overall performance of DTI prediction models.

Taxonomy based on application type

This section categorizes 44 reviewed articles based on the primary application types of neural network-based Drug–Target Interaction (DTI) prediction models, addressing Research Question 12 on how biological and chemical data can enhance DTI prediction accuracy and its practical implications. These models are applied in various areas, as illustrated in Fig. 10, including drug discovery, drug repurposing, prediction of binding affinity, and toxicity analysis. Taxonomy provides insight into the practical utility of DTI prediction and its alignment with therapeutic and pharmacological priorities. According to Fig. 10, 54.55% of the articles (24 articles) are devoted to drug discovery, reflecting

Table 24 Evaluation metrics for DrugBank

Article	Accuracy	AUROC	AUPR
[4]	–	0.979	0.961
[11]	0.8935 0.8819	0.8672 0.8556	0.7125 0.6308
[21]	0.8955	0.9548	0.9644
[23]		0.9620	0.9605
[25]	0.823	0.905	0.912
[29]	–	0.902 0.982	–
[33]	0.983	0.99	0.994
[38]	0.836	0.905	0.908
[39]	–	0.968	0.462
[43]	0.9013	0.962	0.959
[44]	0.79	0.85	0.80
[46]	0.9672	–	–
[47]	0.87	–	–
[48]	0.9800	0.9965	–

Table 25 Evaluation metrics for Human

Article	Accuracy	AUROC	AUPR
[14]	0.942	0.983	0.976
[16]	–	0.987	0.989
[17]	–	0.991	–
[27]	–	0.98	0.94
[29]	–	0.902	–
[40]	–	0.9822	0.9568
[41]	0.999	0.995	0.968
[42]	–	0.983	–
[52]	–	0.985	0.988

Table 26 Evaluation metrics for Davis

Article	Accuracy	AUROC	AUPR
[25]	0.823	0.905	0.912
[34]	0.865	0.917	0.897
[35]	0.8739	0.9225	0.8395
[38]	0.836	0.905	0.908

the field's focus on identifying novel drug–target pairs to accelerate the development of new therapeutics, as discussed in "Future Directions and Open Issues" section. This trend is exemplified by models such as HGDTI [4] and CAT-DTI [14], which utilize hybrid and Transformer architectures to discover new interactions, thereby supporting early stage drug development. The use of DTI prediction to identify new uses for existing drugs follows second with 22.73% (10 articles). Especially relevant in urgent scenarios, such as

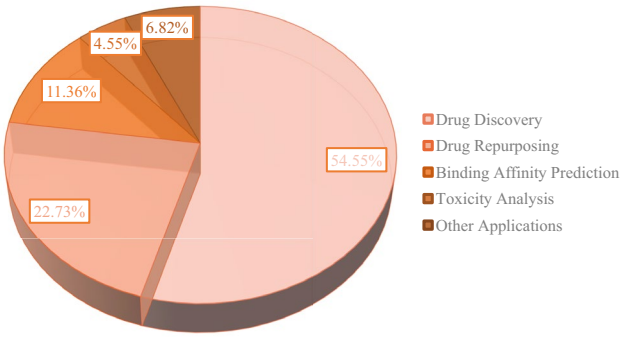


Fig. 10 Application distribution

pandemics (e.g., [32] for COVID-19), is this application. MSF-DTI [27] incorporates multi-scale features to predict interaction strengths, thereby enhancing the precision of drug design by 11.36% (based on five articles). Often, accurate affinity data are available, with a focus on databases such as DrugBank (36 articles). The toxicity analysis, represented by 6.82% (3 articles), addresses safety concerns with approaches such as those that utilize Explainable AI to assess possible adverse effects, aligning with future directions for interpretability. The remaining 4.55% (2 articles) illustrate the versatility of these models in broader pharmacological contexts. Drug discovery (54.55%) exhibits a strong correlation with comprehensive databases, such as DrugBank, and metrics, including AUROC (42 articles), which ensure the robust validation of new interactions. Adaptability to existing chemical spaces is enhanced by the significant share of drug repurposing (22.73%). There is a growing interest in binding affinity prediction (11.36%) and toxicity analysis (6.82%), both of which indicate a shift toward precision and safety, which are likely to continue as multimodal learning and multi-omics data advance. The findings suggest that while drug discovery remains central, the field is diversifying to address repurposing, affinity, and safety needs, which sets the stage for future innovations.

Future directions and open issues

The application of neural network-based methods for Drug–Target Interactions (DTI) prediction has been a game-changing development in drug discovery and repurposing. Using advanced algorithms and big data, it is now feasible to predict novel interactions between biological targets and chemical molecules. Nevertheless, several complex challenges remain in the field, which must be further investigated and novel solutions developed to address them. This chapter broadly outlines the main future directions and open issues,

providing pragmatic solutions and explanatory examples to illuminate the path forward for future research.

Integration of multimodal and multi-omics data

One of the most significant upcoming developments in DTI prediction is multimodal data integration (for example, protein sequence, three-dimensional molecular structure, chemical graph, molecular image, genomic, proteomic, metabolomic, and transcriptomic data) and clinical data (for instance, medical records and drug response from patients). This approach has the potential to significantly enhance the accuracy and coverage of DTI predictions by providing a complete picture of biological interactions. For instance, the HGDTI [4] model, which is based on heterogeneous graph networks, combines drug structural data and protein sequence information with an AUROC value of 0.979. As such, MCL-DTI [16] integrates two-dimensional molecular images and chemical descriptors to achieve an AUROC of 0.987, highlighting the strength of multimodal data to offer complementary structural and spatial information.

However, despite all the achievements, the heterogeneity of data (differences in scale, format, and quality), incomplete data (particularly for less well-studied diseases or targets), and the need for advanced tools for data alignment and integration remain significant issues. For example, metabolomic measurements can be on different scales than genomic measurements, requiring proper normalization. Additionally, in conditions like Alzheimer's or the case of rare cancers, clinical measurements can be incomplete or sparse due to limited patient populations.

Proposed Solutions:

Utilize self-supervised learning techniques to extract latent features from missing and unlabeled data.
Use matrix completion or tensor-based techniques to impute missing data.
Develop multimodal transformers that can automatically ingest and combine heterogeneous data. For example, a transformer model can combine pharmacokinetic data (e.g., drug absorption, distribution, and elimination) with genomic data to enable more personalized predictions.
Employ transcriptomic data (e.g., RNA-seq) to locate indirect or allosteric interactions not evident from structural data alone. This is particularly critical for drug target discovery in multifactorial diseases, such as cancer, where metabolic and genomic effects play key roles.

Practical example: Take the example of designing a drug for breast cancer. By combining gene expression data (RNA-seq) from patients, metabolomic data (the levels of metabolites in cancer cells), and the chemical structures of drugs, it is possible to identify which drugs bind to target proteins in cancer cells with minimal side effects.

Enhancing interpretability with explainable AI (XAI)

Deep neural networks, such as Graph Neural Networks (GNNs) and transformers, also lack the transparency they require, as they are "black-box" models and therefore not transparent. This limits their application in healthcare scenarios where decision-making requires regulation, approval, and high trust. The FragXsiteDTI [17] model, which utilizes attention modules, was able to identify significant drug fragments and protein-binding sites, generating attention maps that indicate which molecular segments interact with specific regions of the protein. Such maps have enabled the design of new drugs with increased effectiveness.

However, for broader clinical applications, more advanced XAI techniques are needed. Techniques like SHAP (Shapley Additive exPlanations) are capable of quantifying the significance of each input feature (e.g., functional groups on a drug molecule or specific protein domains) to the final prediction. For instance, SHAP can inform us about the degree to which the presence of a hydroxyl group on a molecule influences the binding affinity. Moreover, sophisticated visualizations, e.g., three-dimensional heatmaps or Virtual Reality (VR) interaction displays, can render complex interactions more intuitive and usable for clinicians and researchers.

Proposed solutions:

Develop hybrid models that perform both prediction and explanation simultaneously, such as the conjunction of GNNs with SHAP or the utilization of multi-level attention mechanisms for hierarchical explanations (molecule-to-biology).
Create easy-to-use visualization interfaces, such as interactive dashboards that allow clinicians to click on a molecule and observe its main interactions.
Train clinical personnel to read XAI outputs, thereby increasing their confidence in these tools.

Challenges:

The computational complexity of XAI approaches, such as SHAP, can be significant, potentially requiring a considerable amount of time.
Complexity of results for non-expert users (e.g., clinicians who lack in-depth machine learning expertise).

Practical example: During a clinical trial, FragXsiteDTI [17] utilized attention maps to demonstrate that one amine group in a potential drug exhibits a significant interaction with a hydrophobic pocket of the target protein. Such

information guided the optimization of the drug's chemical structure to enhance its effectiveness.

Addressing data scarcity and cold-start scenarios

Insufficient data, particularly for new targets, less-known drugs, or rare diseases (cold-start conditions), remain a primary impediment to developing DTI models. The MGDTI [50] model, based on meta-learning and drug–drug and target–target similarity information, has shown stunning performance in low-data conditions (AUPR 0.92). However, its reliance on auxiliary data impedes its performance where such data are unavailable.

Proposed solutions:

Apply transfer learning: Models can be pre-trained on extensive, inclusive databases, such as DrugBank or ChEMBL and subsequently fine-tuned for target-specific or orphan disease prediction. For instance, OverfitDTI [2] with a CI of 0.9480 proved that this method improves generalizability in cold-start applications.
Use Generative Adversarial Networks (GANs): These can be used to create realistic synthetic data (such as molecular structures or simulated interactions) for enriching training datasets.
Frequent biochemical simulations: Techniques such as molecular dynamics or quantum calculations can simulate drug–target interactions at the atomic level, generating new data for model training and refinement.

Challenges:

The synthetic data produced by GANs may not be as good as laboratory data.
Biochemical simulations require significant computational resources and time.

Practical example: For an orphan disease like cystic fibrosis, for which there are limited experimental data, GANs can generate molecular structures similar to existing drugs, and molecular dynamics simulations can screen their interactions with the CFTR protein.

Reducing computational complexity

Complex models, such as HGDTI [4] and CAT-DTI [14], constructed on intricate architectures like GNNs and transformers, require substantial computational power. For instance, HGDTI [4] achieves an AUROC of 0.979; however, training on high-end GPUs may take days. Such complexity does not allow utilization by research groups with limited resources or production environments that require speedy prediction.

Proposed solutions:

Apply knowledge distillation: A large, complex model (teacher) transfers its knowledge to a smaller, more lightweight model (student), retaining accuracy while reducing runtime.
Use model pruning: Pruning away unnecessary or low-impact parameters can reduce computational load. For example, LDCNN [44] reduced complexity by pruning but maintained satisfactory performance.
Leverage quantum computing: Quantum algorithms can accelerate certain complex computations, such as graph optimization.
Implement parallelization and utilize advanced hardware, such as multiple GPUs or TPUs, to leverage distributed algorithms and reduce training time.

Practical example: A low-budget research team can distill the HGDTI [4] model into a leaner variant that runs on a regular laptop with only a 5% loss in accuracy.

Standardizing evaluation metrics

Although model performance comparisons are attempted between papers, experiment results within a paper vary significantly, and variations in these settings interfere with meaningful and apples-to-apples model-to-model comparisons. For the majority of documents, model settings, datasets, and preprocessing methods are various, and experimental conditions directly influence the performance reported. To be specific, model comparisons in these papers are made under different and non-uniform conditions, rendering the final comparison questionable. With the absence of a single criterion for evaluation, model comparisons are significantly restricted and unfair, affecting both equality and accuracy. For instance, the reporting of performance metrics, such as AUROC and AUPR, is customary in most papers, whereas crucial metrics, such as CI and MCC, are rarely reported. This not only diminishes the comparability of the tests but also casts doubt on the validity of the results.

Proposed solutions:

Develop a consolidated evaluation framework with benchmark datasets (e.g., DrugBank, BindingDB, and ChEMBL), uniform data splitting protocols (e.g., 80–20 split), and comprehensive metrics (AUROC, AUPR, CI, MCC, and F1-Score).
Develop an open-source platform for testing models under identical conditions and reporting results in standardized comparison tables.
Challenge researchers to report all relevant metrics, provide code availability, and publish datasets to enable transparency and reproducibility.

Challenges:

No single evaluation criterion for studies.
Variability in experimental conditions, such as datasets and preprocessing methods.
Inadequate reporting of key measures like CI and MCC.

Practical example: An online platform could test various DTI models on *BindingDB* with an 80–20 data split and display results in an interactive dashboard, enabling fair and transparent comparisons.

Applications in personalized medicine

DTI prediction can potentially become the core of individualized medicine, particularly by leveraging patients' genomic data (e.g., genome-wide association studies, whole-genome sequencing, or RNA sequencing). It is possible to forecast which medications are more effective or safer in patients with specific genetic profiles. In cancer treatment, for example, where drug response is heavily affected by genetic variability, DTI models can identify best-in-class drugs for each patient.

Proposed solutions:

Integrate real-world clinical data [e.g., Electronic Health Records (EHR)] with DTI models for personalized prediction.
Construct drug recommendation systems based on DTI models that take patients' genetic data as input.
Utilize multi-task learning to learn drug efficacy and side effects from genomic data simultaneously.

Practical example: For a patient with lung cancer and an EGFR mutation, a DTI model may suggest a medicine that directly engages with the mutated protein, resulting in fewer side effects compared to conventional drugs.

Predicting side effects and toxicity

Because only 6.82% of the papers examined addressed toxicity prediction, this is still an under-explored area. DTI models can be extended to predict efficacy in addition to toxicity by integrating toxicity information (e.g., LD50, SIDER data, or adverse event reports) and employing multi-task learning.

Challenges:

Limited availability of comprehensive toxicity information for certain drugs.
Difficulty in predicting off-target interactions.

Proposed solutions:

Develop multi-task models that provide dual outputs for efficacy and toxicity.
Utilize post-market reports and clinical data to identify any unexpected side effects.
Model off-target interactions using molecular docking or molecular dynamics.

Practical example: A DTI model might predict that a new anti-inflammatory drug interacts with a liver enzyme other than its intended target, thereby increasing the risk of hepatotoxicity, which can be corrected in the early stages of drug discovery.

Ethical considerations and bias mitigation

Data bias can damage the performance of DTI models for some populations or disorders. For example, suppose DrugBank data consist only of drugs prevalent in developed countries. In that case, models can perform poorly when predicting interactions for local or traditional drugs (e.g., Asian or African herbal medicines). Data bias can lead to inequities in access to treatment.

Proposed solutions:

Apply fairness techniques, such as reweighing or Adversarial Debiasing, to combat data bias.
Obtain more diverse data from global sources, including those from developing countries.
Establish ethical standards for data privacy (e.g., genomic data) protection and transparency when using models.

Practical example: A model of DTI trained using diverse data from African, Asian, and European patients would recommend a malaria drug that is effective and safe in all genetic populations.

Interdisciplinary collaboration and emerging technologies

DTI prediction innovations require close collaboration between engineers, chemists, data scientists, and biologists. New technologies, such as quantum computing, federated learning, and blockchain, also help address the existing challenges.

Proposed solutions:

Use quantum computing to accelerate molecular interaction simulations.
Use federated learning for model training on decentralized datasets (e.g., hospital records) without compromising privacy.
Use blockchain for transparent record-keeping of model development and findings.

Practical example: An international project could utilize federated learning to train a DTI model on hospital information worldwide, while keeping data in local locations to preserve privacy.

Conclusions

The objective of this study was to review neural network-based approaches for predicting Drug–Target Interactions (DTIs), with a focus on their transformative role in drug discovery and repurposing. The analysis of 44 articles allowed us to identify key trends, performance metrics, and challenges in the field. Neural network-based methods, comprising CNNs, GNNs, Transformer architectures, and hybrid models, have surpassed machine learning-based and conventional methods with AUROCs ranging from 0.88 to 0.99. This signifies the increasing contribution of deep neural network techniques to the accuracy and robustness of DTI prediction. Hybrid models, which fuse different architectures and multimodal data, achieved the highest accuracy. Models like HGDTI [4] and MCL-DTI [16] have demonstrated how incorporating multiple network structures can enhance predictive power, particularly when dealing with complex molecular interactions. At the same time, Transformer-based models like CAT-DTI [14] have been shown to perform effectively with long-range dependencies and complex data structures, further boosting their potential in DTI prediction. These findings highlight the value of model flexibility and underscore the need for approaches that integrate multiple methods in addressing challenges in DTI prediction. In taxonomic research, hybrid models were the majority, accounting for 57% of the papers. Their prevalence is a testament to their flexibility and ability to integrate multiple methods for improved performance. Moreover, the focus on drug discovery (54.55%) and drug repurposing (22.73%) reflects the field's ongoing interest in both new therapy development and the repurposing of existing drugs. The DrugBank database, employed in 36 papers due to its comprehensiveness, is especially significant for model training and testing, as well as standard performance metrics such as AUROC and AUPR. Despite these advancements, new targets face challenges of computational cost, lack of interpretability, and data scarcity. While models like FragXsiteDTI [17] have helped address interpretability by ranking drug fragments and protein pockets, further work on Explainable AI (XAI) is needed to ensure transparency, especially in clinical applications. For greater clinical utility, future research is required to render these models more interpretable, integrate multi-omics data, and reduce their computational complexity. This may enable them to better align with personalized medicine initiatives. Overall, neural network-based approaches hold promise for accelerating drug discovery and enhancing therapeutic efficiency. By transcending existing limitations, such as computational complexity and a lack of interpretability, and leveraging emerging technologies, including quantum computing and XAI, computational pharmacology can contribute to the development of safer and more effective therapies. Future research should continue to explore the integration of multiple data sources and refine advanced models to maximize the impact of these methodologies on real-world clinical practices.

Author contributions Fatemeh Panahandeh: design the study, gathering data for the article, and verification. Najme Mansouri: investigation, and writing—original draft preparation.

Funding Not applicable.

Data availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

Consent for publication Not applicable.

Ethical approval Not applicable.

References

1. Zuo Y, Wu X, Ge F, Yan H, Fei S, Liang J, Deng Z (2025) Research progress on drug-target interactions in the last five years. Anal Biochem 697:115691. https://doi.org/10.1016/j.ab.2024.115691
2. Xiaolin X et al (2023) Overfit deep neural network for predicting drug-target interactions. iScience 26(9):107646. https://doi.org/10.1016/j.isci.2023.107646
3. Rayhan F, Ahmed S, Mousavian Z, Farid DM, Shatabda S (2020) FRnet-DTI: Deep convolutional neural network for drug-target interaction prediction. Heliyon 6(3):e03444. https://doi.org/10.1016/j.heliyon.2020.e03444
4. Yu L, Qiu W, Lin W, Cheng X, Xiao X, Dai J (2022) HGDTI: predicting drug–target interaction by using information aggregation based on heterogeneous graph neural network. BMC Bioinformatics 23(1):126. https://doi.org/10.1186/s12859-022-04655-5
5. Shi W, Yang H, Xie L, Yin X-X, Zhang Y (2024) A review of machine learning-based methods for predicting drug–target interactions. Health Inf Sci Syst 12(1):30. https://doi.org/10.1007/s13755-024-00287-6
6. Alizadehsani R, Oyelere SS, Hussain S, Jagatheesaperumal SK, Calixto RR, Rahouti M, Roshanzamir M, de Albuquerque VHC (2024) Explainable artificial intelligence for drug discovery and development: a comprehensive survey. IEEE Access 12:35796–35812. https://doi.org/10.1109/ACCESS.2024.3373195
7. Atasever S (2024) In silico drug discovery: a machine learning-driven systematic review. Med Chem Res 33(9):1465–1490. https://doi.org/10.1007/s00044-024-03260-w
8. Li H, Zou L, Kowah JAH, He D, Liu Z, Ding X, Wen H, Wang L, Yuan M, Liu X (2023) A compact review of progress and prospects of deep learning in drug discovery. J Mol Model 29(4):117. https://doi.org/10.1007/s00894-023-05492-w

9. Peng J, Li J, Shang X (2020) A learning-based method for drug-target interaction prediction based on feature representation learning and deep neural network. BMC Bioinformatics 21(S13):394. https://doi.org/10.1186/s12859-020-03677-1
10. Rifaioglu AS, Nalbat E, Atalay V, Martin MJ, Cetin-Atalay R, Doğan T (2020) DEEPScreen: high performance drug–target interaction prediction with convolutional neural networks using 2-D structural compound representations. Chem Sci 11(9):2531–2557. https://doi.org/10.1039/C9SC03414E
11. Ye Q, Zhang X, Lin X (2023) Drug-target interaction prediction via graph auto-encoder and multi-subspace deep neural networks. IEEE ACM Trans Comput Biol Bioinform 20(5):2647–2658. https://doi.org/10.1109/TCBB.2022.3206907
12. Zhang Y, Hu Y, Han N, Yang A, Liu X, Cai H (2023) A survey of drug-target interaction and affinity prediction methods via graph neural networks. Comput Biol Med 163:107136. https://doi.org/10.1016/j.compbiomed.2023.107136
13. Zhang Z, Lifan C, Feisheng Z, Dingyan W, Jiaxin J, Sulin Z, Hualiang J, Mingyue Z, Xutong L (2022) Graph neural network approaches for drug-target interactions. Curr Opin Struct Biol 73:102327. https://doi.org/10.1016/j.sbi.2021.102327
14. Zeng X, Chen W, Lei B (2024) Cat-dti: cross-attention and transformer network with domain adaptation for drug-target interaction prediction. BMC Bioinformatics 25(1):141. https://doi.org/10.1186/s12859-024-05753-2
15. Zhang Y-L, Wang W-T, Guan J-H, Jain DK, Wang T-Y, Roy SK (2024) Mocformer: a two-stage pre-training-driven transformer for drug-target interactions prediction. Int J Comput Intell Syst 17(1):165. https://doi.org/10.1007/s44196-024-00561-1
16. Qian Y, Li X, Wu J, Zhang Q (2023) MCL-DTI: using drug multimodal information and bi-directional cross-attention learning method for predicting drug–target interaction. BMC Bioinformatics 24(1):323. https://doi.org/10.1186/s12859-023-05447-1
17. Khodabandeh Yalabadi A, Yazdani-Jahromi M, Yousefi N, Tayebi A, Abdidizaji S, Garibay OO (2024) FragXsiteDTI: revealing responsible segments in drug–target interaction with transformer-driven interpretation. In: Ma J (ed) Research in computational molecular biology. Springer Nature Switzerland, Cham, pp 68–85
18. Luo H, Yin W, Wang J, Zhang G, Liang W, Luo J, Yan C (2024) Drug-drug interactions prediction based on deep learning and knowledge graph: a review. iScience 27(3):109148. https://doi.org/10.1016/j.isci.2024.109148
19. Liu X, Xue Z, Luo M, Ke B, Lv J (2024) Anesthetic drug discovery with computer-aided drug design and machine learning. APS 2(1):7. https://doi.org/10.1007/s44254-023-00047-x
20. Liyaqat T, Ahmad T (2022) A brief review on Artificial Intelligence based Drug Target Interaction Prediction. In 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India: IEEE, 544–549. https://doi.org/10.1109/ICAC3N56670.2022.10074321
21. Tian Z, Yu Y, Ni F, Zou Q (2024) Drug-target interaction prediction with collaborative contrastive learning and adaptive self-paced sampling strategy. BMC Biol 22(1):216. https://doi.org/10.1186/s12915-024-02012-x
22. Iliadis D, De Baets B, Pahikkala T, Waegeman W (2024) A comparison of embedding aggregation strategies in drug–target interaction prediction. BMC Bioinformatics 25(1):59. https://doi.org/10.1186/s12859-024-05684-y
23. Zhang B, Niu D, Zhang L, Zhang Q, Li Z (2024) MSH-DTI: Multi-graph convolution with self-supervised embedding and heterogeneous aggregation for drug-target interaction prediction. BMC Bioinformatics 25(1):275. https://doi.org/10.1186/s12859-024-05904-5
24. Dehghan A, Abbasi K, Razzaghi P, Banadkuki H, Gharaghani S (2024) CCL-DTI: contributing the contrastive loss in drug–target interaction prediction. BMC Bioinformatics 25(1):48. https://doi.org/10.1186/s12859-024-05671-3
25. Zhao Q, Duan G, Zhao H, Zheng K, Li Y, Wang J (2023) GIFDTI: prediction of drug-target interactions based on global molecular and intermolecular interaction representation learning. IEEE ACM Trans Comput Biol Bioinform 20(3):1943–1952. https://doi.org/10.1109/TCBB.2022.3225423
26. Yu W, Yu H (2024) DeepMHAttGRU-DTI: Prediction of Drug–Target Interactions Based on Knowledge Graph Random Walk Embeddings and GRU Neural Network. In Advanced Intelligent Computing in Bioinformatics, vol. 14882, D.-S. Huang, Y. Pan, and Q. Zhang, Eds., in Lecture Notes in Computer Science, vol. 14882. Singapore: Springer Nature Singapore, 96–107. https://doi.org/10.1007/978-981-97-5692-6_9
27. Yang Z, Bai B, Long J, Wei P, Li J (2024) Multi-scale Feature Fusion Neural Network for Accurate Prediction of Drug–Target Interactions. In Neural Information Processing, vol. 1964, B. Luo, L. Cheng, Z.-G. Wu, H. Li, and C. Li, Eds., in Communications in Computer and Information Science, vol. 1964., Singapore: Springer Nature Singapore, 176–188. https://doi.org/10.1007/978-981-99-8141-0_14
28. Jiang L, Sun J, Wang Y, Ning Q, Luo N, Yin M. Heterogeneous Graph Convolutional Network integrates Multi-modal Similarities for Drug–Target Interaction Prediction. In 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Houston, TX, USA: IEEE, Dec. 2021, 137–140. https://doi.org/10.1109/BIBM52615.2021.9669468
29. Pan Y, Zhang Y, Zhang J, Lu M (2023) CSDTI: an interpretable cross-attention network with GNN-based drug molecule aggregation for drug-target interaction prediction. Appl Intell 53(22):27177–27190. https://doi.org/10.1007/s10489-023-04977-8
30. Chen J, Zhang L, Cheng K, Jin B, Lu X, Che C, Liu Y (2020) Exploring Multi-level Mutual Information for Drug–target Interaction Prediction. In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Seoul, Korea (South): IEEE, 251–256. https://doi.org/10.1109/BIBM49941.2020.9313395
31. Jin X, Liu M, Wang L, He W, Huang Y, Xie M (2020) Multi-Resolutional Collaborative Heterogeneous Graph Convolutional Auto-Encoder for Drug–Target Interaction Prediction. In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Seoul, Korea (South): IEEE, 150–153. https://doi.org/10.1109/BIBM49941.2020.9313489
32. Zhang S, Lin X, Zhang X (2021) Discovering DTI and DDI by Knowledge Graph with MHRW and Improved Neural Network. In 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Houston, TX, USA: IEEE, 588–593. https://doi.org/10.1109/BIBM52615.2021.9669347
33. Xu S, Lin X, Yu H (2021) Prediction of Drug–Target Interactions Using Molecular Graph and GDNet-DTI Model. In 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Houston, TX, USA: IEEE, 3747–3754. https://doi.org/10.1109/BIBM52615.2021.9669652
34. Wang T, Yang W, Chen J, Tian Y, Wei D-Q (2022) ConformerDTI: Local Features Coupling Global Representations for Drug–Target Interaction Prediction. In 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA: IEEE, 1227–1234. https://doi.org/10.1109/BIBM55620.2022.9995344
35. Li H, Zhang W, Shang Y, Lu W (2024) Collaborative Prediction of Drug–target Interaction using Sequence-based CNN and Transformer. In 2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Tianjin, China: IEEE, 2276–2281. https://doi.org/10.1109/CSCWD61410.2024.10580505

36. Nandhini K, Mayilvahanan P (2022) Molecular Sub-structural Pattern Discovery using Transformer-CNN-OMBO Algorithm for Drug–target interaction (DTI) Prediction. In 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India: IEEE, 241–247. https://doi.org/10.1109/ICOEI53556.2022.9776666
37. Shao K, Zhang Z, He S, Bo X (2020) DTIGCCN: Prediction of drug–target interactions based on GCN and CNN. In 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI), Baltimore, MD, USA: IEEE, 337–342. https://doi.org/10.1109/ICTAI50040.2020.00060
38. Chen Z, Peng J (2023) CHADTI: A Sequence-Based Model with Attention Mechanism for Drug–Target Interaction Prediction. In 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Honolulu, Oahu, HI, USA: IEEE, 107–113. https://doi.org/10.1109/SMC53992.2023.10394079
39. Xu X, Xuan P, Zhang T, Chen B, Sheng N (2022) Inferring drug-target interactions based on random walk and convolutional neural network. IEEE ACM Trans Comput Biol Bioinform 19(4):2294–2304. https://doi.org/10.1109/TCBB.2021.3066813
40. Cheng Z, Yan C, Wu F-X, Wang J (2022) Drug-target interaction prediction using multi-head self-attention and graph attention network. IEEE ACM Trans Comput Biol Bioinform 19(4):2208–2218. https://doi.org/10.1109/TCBB.2021.3077905
41. Wang W, Yu M, Sun B, Li J, Liu D, Zhang H, Wang X, Zhou Y (2024) SMGCN: multiple similarity and multiple kernel fusion based graph convolutional neural network for drug-target interactions prediction. IEEE ACM Trans Comput Biol Bioinform 21(1):143–154. https://doi.org/10.1109/TCBB.2023.3339645
42. Wang S, Shan P, Zhao Y, Zuo L (2021) GanDTI: a multi-task neural network for drug-target interaction prediction. Comput Biol Chem 92:107476. https://doi.org/10.1016/j.compbiolchem.2021.107476
43. Wang Y, Zhang Z, Piao C, Huang Y, Zhang Y, Zhang C (2023) LDS-CNN: a deep learning framework for drug-target interactions prediction based on large-scale drug screening. Health Inf Sci Syst 11(1):42. https://doi.org/10.1007/s13755-023-00243-w
44. Wang S, Du Z, Ding M, Zhao R, Rodriguez-Paton A, Song T (2020) LDCNN-DTI: A Novel Light Deep Convolutional Neural Network for Drug–Target Interaction Predictions. In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Seoul, Korea (South): IEEE, 1132–1136. https://doi.org/10.1109/BIBM49941.2020.9313585
45. Ghozzi WB, Chaabani A, Kodia Z, Said LB (2023) DeepCNN-DTI: A Deep Learning Model for Detecting Drug–Target Interactions. In 2023 9th International Conference on Control, Decision and Information Technologies (CoDIT), Rome, Italy: IEEE, 1677–1682. https://doi.org/10.1109/CoDIT58514.2023.10284177
46. Yaseen BT (2023) Drug Target Interaction Prediction Using Convolutional Neural Network (CNN). In 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Istanbul, Turkiye: IEEE, 1–5. https://doi.org/10.1109/HORA58378.2023.10156717
47. Nandhini K, Thailambal G (2022) CNN-OSBO Encoder-Decoder Architecture for Drug–Target Interaction (DTI) Prediction of Covid-19 Targets. In 2022 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India: IEEE, 1199–1206. https://doi.org/10.1109/ICECA55336.2022.10009182
48. Hu S, Xia D, Su B, Chen P, Wang B, Li J (2021) A convolutional neural network system to discriminate drug-target interactions. IEEE ACM Trans Comput Biol Bioinform 18(4):1315–1324. https://doi.org/10.1109/TCBB.2019.2940187
49. Wang H, Guo F, Du M, Wang G, Cao C (2022) A novel method for drug-target interaction prediction based on graph transformers model. BMC Bioinformatics 23(1):459. https://doi.org/10.1186/s12859-022-04812-w
50. Zhao Z, He C, Qu Y, Zheng H, Duan L, Zuo J (2023) MGDTI: Graph Transformer with Meta-Learning for Drug–Target Interaction Prediction. In 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Istanbul, Turkiye: IEEE, 801–806. https://doi.org/10.1109/BIBM58861.2023.10385671
51. Simon E, Bankapur SS (2024) Prediction of Drug–Target Interactions Using BERT for Protein Sequences and Drug Compounds. In 2024 16th International Conference on Communication Systems & Networks (COMSNETS), Bengaluru, India: IEEE, 436–438. https://doi.org/10.1109/COMSNETS59351.2024.10427536
52. Liu B, Wu S, Wang J, Deng X, Zhou A (2024) HiGraphDTI: Hierarchical Graph Representation Learning for Drug–Target Interaction Prediction," in Machine Learning and Knowledge Discovery in Databases. Research Track, vol. 14946, A. Bifet, J. Davis, T. Krilavičius, M. Kull, E. Ntoutsi, and I. Žliobaitė, Eds., in Lecture Notes in Computer Science, vol. 14946. Cham: Springer Nature Switzerland, 354–370. https://doi.org/10.1007/978-3-031-70365-2_21
53. Li Y, Liang W, Peng L, Zhang D, Yang C, Li K-C (2024) Predicting drug-target interactions via dual-stream graph neural network. IEEE ACM Trans Comput Biol Bioinform 21(4):948–958. https://doi.org/10.1109/TCBB.2022.3204188
54. Wang L, Zhang X, Lin X, Hu J (2023) Drug–Target Interaction Prediction Based on Drug Subgraph Fingerprint Extraction Strategy and Subgraph Attention Mechanism," in Advanced Data Mining and Applications, vol. 14178, X. Yang, H. Suhartanto, G. Wang, B. Wang, J. Jiang, B. Li, H. Zhu, and N. Cui, Eds., in Lecture Notes in Computer Science, vol. 14178. , Cham: Springer Nature Switzerland, 3–17. https://doi.org/10.1007/978-3-031-46671-7_1
55. Wang X, Wang J, Wang Z (2021) A Drug–Target Interaction Prediction Based on GCN Learning. In 2021 IEEE 9th International Conference on Bioinformatics and Computational Biology (ICBCB), Taiyuan, China: IEEE, 42–47. https://doi.org/10.1109/ICBCB52223.2021.9459231