Indian Journal of Modern Research and Reviews

This Journal is a member of the *'Committee on Publication Ethics'*

Online ISSN:2584-184X

Research Article

Deep Learning Strategies for Predicting Drug–Target Interactions: Advances, Challenges, and Future Perspectives

Avinash Bajpai [1], Sachin Sharma [2*], Birender Singh [3], KM Nisha [4]

[1-2] Assistant Professor, ITS College of Pharmacy, Muradnagar, Ghaziabad, Uttar Pradesh, India

[3] MM College of Pharmacy, Maharishi Markandeshwar (Deemed to be University), Mullana, Haryana, India

[4] Assistant Professor, ITS College of Pharmacy, Muradnagar, Ghaziabad, Uttar Pradesh, India

Corresponding Author: * Sachin Sharma

DOI: https://doi.org/10.5281/zenodo.20344084

Abstract

Drug–target interaction (DTI) prediction lies at the heart of modern drug discovery, determining whether a candidate small molecule will bind to and modulate a biological macromolecule of therapeutic relevance. Traditional experimental high-throughput screening is expensive, time-consuming, and constrained by library size, while classical computational approaches—docking, pharmacophore modelling, and quantitative structure–activity relationship (QSAR) modelling suffer from limitations in scalability and generalizability. The emergence of deep learning (DL) has fundamentally transformed the field, enabling end-to-end learning of molecular representations and interaction patterns from heterogeneous, large-scale biomedical data.

This review provides a comprehensive synthesis of DL-based DTI prediction methodologies, covering convolutional neural networks (CNNs), recurrent neural networks (RNNs), graph neural networks (GNNs), transformer-based architectures, autoencoders, and multi-modal fusion frameworks. We discuss the critical role of molecular representation—from one-dimensional SMILES strings and circular fingerprints to three-dimensional molecular graphs and protein contact maps. Benchmark datasets including Davis, KIBA, BindingDB, ChEMBL, and PDBbind are reviewed with respect to their composition, metric conventions, and appropriate use. We critically examine key challenges: data scarcity and imbalance, negative-sample bias, interpretability deficits, cold-start generalization, and the limited availability of experimental three-dimensional protein structures. Emerging solutions—pre-trained chemical language models, AlphaFold3 integration, federated learning, knowledge graph-augmented GNNs, and causal interpretability methods—are discussed as future directions. This review aims to serve as an authoritative reference for computational chemists, bioinformaticians, and medicinal chemists seeking to leverage DL for accelerated, cost-effective drug discovery.

Manuscript Information

- **ISSN No:** 2584-184X
- **Received:** 04-04-2026
- **Accepted:** 17-05-2026
- **Published:** 22-05-2026
- **MRR:**4(5); 2026: 249-259

- **Plagiarism Checked:** Yes
- **Peer Review Process:** Yes

How to Cite this Article

Bajpai A, Sharma S, Singh B, K M N. Deep Learning Strategies for Predicting Drug–Target Interactions: Advances, Challenges, and Future Perspectives. Indian J Mod Res Rev. 2026;4(5):249-259.

Access this Article Online

www.mrrjournal.in

KEYWORDS: Drug–Target Interaction; Deep Learning; Graph Neural Networks; Molecular Representation; Binding Affinity; Drug Discovery; Transformer; Alphafold

1. INTRODUCTION

The process of bringing a new drug from initial discovery to clinical approval remains one of the most resource-intensive endeavours in science, typically requiring 10–15 years and exceeding US$2.5 billion in direct costs. Central to early-phase drug discovery is the identification of specific drug–target interactions (DTIs): the binding of a small-molecule compound to a protein, enzyme, receptor, ion channel, or nucleic acid structure that mediates a disease phenotype. Accurate prediction of DTIs can dramatically reduce the search space, prioritize experimental validation efforts, and reveal off-target liabilities early in development.

Historically, DTI discovery relied on serendipity, phenotypic screening, and later on rational structure-based design. The advent of genomics and proteomics in the 1990s–2000s expanded the druggable proteome to several thousand putative targets, far exceeding the capacity of experimental screening. Computational approaches molecular docking, homology modelling, pharmacophore mapping, and machine learning-based QSAR provided partial solutions but struggled with flexibility, scalability, and the need for high-quality three-dimensional structures.

Deep learning, a subfield of machine learning that uses hierarchical feature representations learned from raw data, has achieved breakthrough performance in image recognition, natural language processing, and game-playing. Its application to biomolecular problems, ignited by AlphaFold's revolution in protein structure prediction, has catalysed equally transformative advances in DTI prediction. DL models can simultaneously learn drug representations, protein representations, and their cross-domain interaction signals without manual feature engineering, operating at a scale and accuracy unattainable by previous methods.

The purpose of this review is threefold: (i) to systematically describe the major DL architectures applied to DTI prediction, with attention to their theoretical underpinnings and published implementations; (ii) to survey benchmark datasets, evaluation metrics, and reproducibility considerations; and (iii) to identify open challenges and delineate promising future research directions. We anticipate that this review will serve as a comprehensive resource for both newcomers and experts navigating the rapidly evolving landscape of AI-driven drug discovery.

2. Background and Theoretical Foundations

2.1 Drug–Target Interaction: Definitions and Importance

A DTI is operationally defined as a physico-chemical association between a ligand (small molecule, peptide, or macromolecule) and a macromolecular target (protein, RNA, DNA) that alters target function with pharmacological consequence. DTIs are quantified through binding affinity metrics including equilibrium dissociation constant (Kd), inhibition constant (Ki), half-maximal inhibitory concentration (IC50), and KIBA scores—composite metrics combining multiple affinity readouts. The ultimate goal of DTI prediction is to accurately estimate these values or classify pairs as interacting/non-interacting across the chemical and proteomic space.

Network pharmacology has further revealed that most drugs act on multiple targets simultaneously (polypharmacology), and that target pathways are embedded within complex biological networks. This systems-level perspective demands DTI models that can handle multi-target, multi-pathway interactions at scale a requirement ideally suited to DL approaches operating over knowledge graphs.

2.2 Classical Computational Approaches and Their Limitations

Molecular docking simulates the binding geometry of a ligand within a protein binding site, estimating binding free energy through scoring functions. While conceptually powerful, docking accuracy is limited by receptor flexibility, scoring function errors, and the requirement for high-quality crystal structures. QSAR models link chemical descriptors to biological activity but are constrained by descriptor quality, applicability domain, and interpretability. Similarity-based methods (chemical or biological) assume activity consistency with structural similarity a principle that fails for activity cliffs. These limitations collectively motivate data-driven deep learning approaches.

2.3 Overview of Deep Learning

Deep learning comprises neural network architectures with multiple layers of non-linear transformations that extract progressively abstract representations from raw input data. The key components are: (i) an input representation layer; (ii) a stack of parameterized feature extraction layers (convolutional, recurrent, attentive, or graph-based); (iii) an aggregation or interaction module; and (iv) a task-specific output head. Models are trained by minimizing a loss function—mean squared error (MSE) for regression, binary cross-entropy (BCE) for classification—using gradient-based optimization via backpropagation. Regularization strategies (dropout, batch normalization, weight decay) combat overfitting, while data augmentation, transfer learning, and semi-supervised methods address data scarcity.

3. MOLECULAR REPRESENTATION METHODS

3.1 Drug Representations

3.1.1 SMILES Strings and Sequence-Based Encoding

The Simplified Molecular Input Line Entry System (SMILES) encodes molecular connectivity as a string of ASCII characters. CNNs and RNNs treat SMILES as biological sequences, learning local motifs and global dependencies. Extended Connectivity Fingerprints (ECFP/Morgan) encode circular neighbourhood environments around each atom into bit vectors, providing a compact, fixed-length representation amenable to dense neural network layers. Despite computational efficiency, SMILES-based representations are non-unique (multiple valid SMILES per molecule) and may mislead models through tokenization artifacts.

3.1.2 Molecular Graph Representations

A molecule is naturally a graph: atoms are nodes with feature vectors (atomic number, charge, hybridization, aromaticity) and bonds are edges with features (bond order, stereochemistry). GNNs operate directly on this structure, propagating information through message-passing operations and learning permutation-invariant molecular embeddings. Graph representations preserve topology, avoid SMILES non-uniqueness, and enable end-to-end learning of chemical intuition. Libraries such as RDKit and DeepChem standardize graph construction from SMILES input.

3.1.3 Three-Dimensional Representations

Three-dimensional representations capture spatial arrangement: Cartesian coordinates, Coulomb matrices, or point clouds of atoms. Models such as SchNet, DimeNet, and equivariant GNNs (E(3)-equivariant neural networks) incorporate distance and angle information, achieving high accuracy on quantum chemistry tasks. The primary limitation is conformer generation cost and the existence of multiple low-energy conformers for flexible molecules.

3.1.4 Pre-trained Chemical Embeddings

Self-supervised pre-training on large chemical databases (ChEMBL, ZINC) using SMILES-based transformers (e.g., ChemBERTa, MolBERT) or graph-based strategies (e.g., GROVER, Pre-GNN) produces rich, transferable drug embeddings. These substantially improve DTI prediction in data-scarce settings, analogous to the role of BERT embeddings in NLP.

3.2 Protein Representations

3.2.1 Sequence-Based Encoding

Protein sequences, encoded as amino acid character strings, are processed by CNNs (sliding window motif extraction) or bidirectional LSTMs. Position-specific scoring matrices (PSSMs) from multiple sequence alignments add evolutionary information. Pre-trained protein language models ESM-2, ProtTrans learn contextual amino acid embeddings at scale, capturing structural and functional information without three-dimensional coordinates.

3.2.2 Structure-Based Encoding

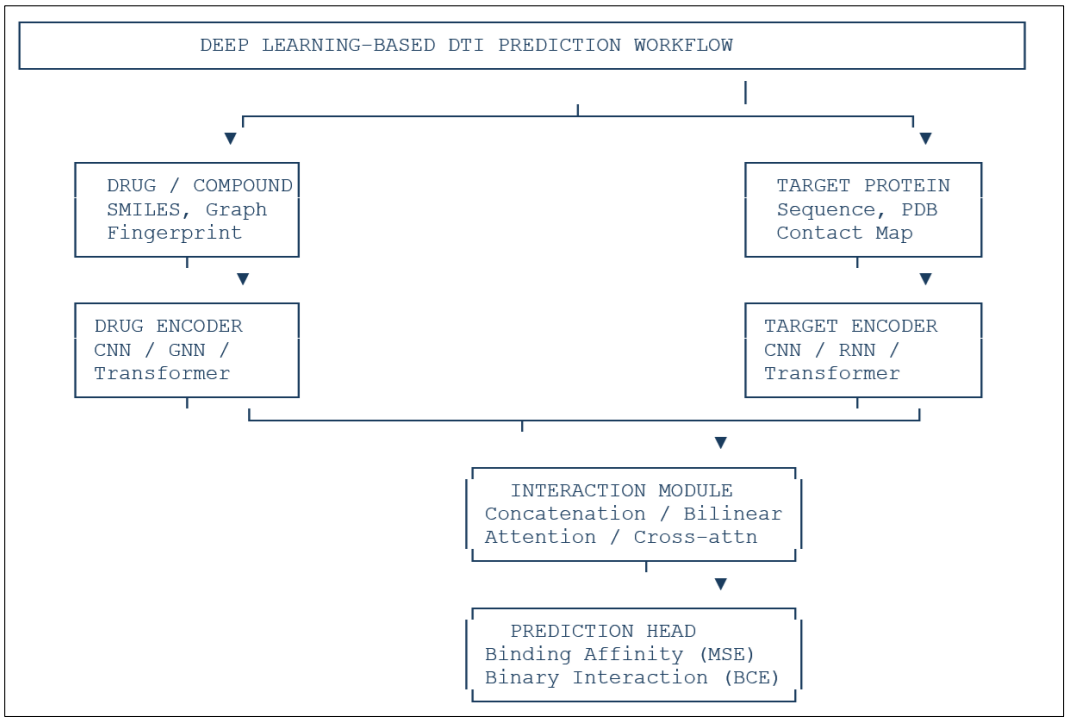
Protein contact maps and distance matrices derived from crystallographic or AlphaFold-predicted structures are processed by 2D CNNs or graph networks. Residue-level graphs with edges defined by spatial proximity enable structure-aware GNNs (e.g., GVP-GNN) to encode protein binding pockets with high fidelity. The widespread availability of AlphaFold2-predicted structures for nearly the entire human proteome has dramatically expanded the scope of structure-based DTI prediction.

Table 1: Comparison of Molecular Representation Methods for Drug–Target Interaction Prediction

Representation	Type	Dimensionality	Advantages / Limitations
Morgan Fingerprint	Bit vector	1024–4096 bits	Fast; loses stereochemistry
SMILES string	Sequence	Variable length	Human-readable; non-unique
Molecular graph	Graph	N atoms + E bonds	Topology-preserving; GNN-ready
3D conformer	Point cloud	3N coordinates	Captures geometry; expensive
Coulomb matrix	Matrix	N×N	Quantum-informed; size-dependent
ECFP (circular FP)	Hashed vector	512–2048 bits	Efficient; no bond-order info
Pre-trained embeddings	Dense vector	128–768 dims	Transfer learning; black-box

4. Deep Learning Architectures for DTI Prediction

Multiple DL architectures have been adapted and purpose-built for DTI prediction. Figure 1 depicts the general workflow shared by most frameworks. Table 1 compares the major architecture families.



Figu 1: Schematic representation of a generic deep learning-based drug–target interaction prediction workflow. Drug and target encoders independently learn representations, which are fused in an interaction module to predict binding affinity or binary interaction.

4.1 Convolutional Neural Networks

CNNs apply learned filters across SMILES or sequence inputs to extract local chemical or structural motifs. DeepDTA (Öztürk *et al*., 2018) [4] independently applies CNN blocks to drug SMILES and protein sequences, concatenates the resulting feature vectors, and feeds them into fully connected layers for binding affinity regression. WideDTA (Pahikkala *et al*., 2019) augments this by incorporating protein domains and drug maximum common substructures as additional inputs, improving generalization. DeepConvDTI employs convolutional filters of multiple sizes in parallel (inception-style) to capture multi-scale motifs in protein sequences.

CNN-based models are computationally efficient and easy to train but lack explicit modelling of molecular topology. The fixed context window of convolutional filters limits capture of long-range dependencies, a limitation addressed by attention mechanisms or recurrent layers.

4.2 Recurrent Neural Networks and LSTM

RNNs process sequential data with hidden states that carry context through time steps, making them natural for SMILES and protein sequence modelling. Long short-term memory networks (Hochreiter and Schmidhuber, 1997) [40] mitigate the vanishing gradient problem through gated mechanisms. Bidirectional LSTMs (BiLSTMs) process sequences in both directions, capturing upstream and downstream context. In DTI prediction, BiLSTMs encode protein sequences while CNNs encode drugs in hybrid architectures, combining the complementary strengths of both.

4.3 Graph Neural Networks

GNNs operate on molecular graphs through message-passing neural networks (MPNN), where each atom aggregates feature messages from its neighbours, updates its hidden state, and after L layers generates a global molecular embedding via a readout function. Figure 2 illustrates the core message-passing operation. GraphDTA (Nguyen *et al*., 2021) [5] applies four GNN variants GCN, GAT, GIN, and GAT-GCN to molecular graphs while encoding protein sequences with CNNs, achieving state-of-the-art performance on Davis and KIBA benchmarks.

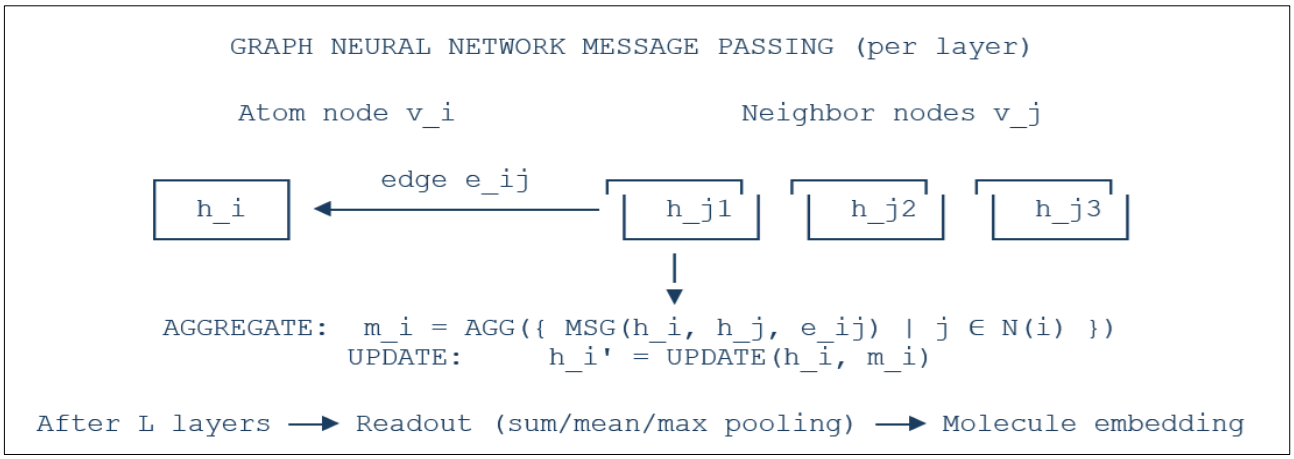


Fig 2: Illustration of message passing in a Graph Neural Network (GNN) for molecular encoding. Each atom node aggregates feature messages from its neighbours, and a global readout produces a fixed-size molecular embedding.

MGraphDTA (Yang *et al*., 2022) [17] introduces multi-scale graph convolutions that simultaneously capture local atomic environments and global molecular topology, achieving the highest reported concordance index on standard benchmarks at the time. GAT-based models (graph attention networks) assign learnable attention weights to neighbouring atoms, providing a built-in explainability mechanism by revealing which atomic neighborhoods drive predictions. Protein-side GNNs build residue-level graphs from contact maps or predicted distance matrices. Each residue is a node with features derived from its amino acid physicochemical properties, secondary structure state, and solvent accessibility. Edges connect residues within a spatial threshold, enabling the model to learn binding pocket geometry.

4.4 Transformer and Attention-Based Models

The transformer architecture (Vaswani *et al*., 2017) [39], based on scaled dot-product self-attention, captures long-range dependencies without sequential processing, enabling parallel training over entire sequences or graphs. MolTrans (Huang *et al*., 2020) [16] designs a sub-structure interaction module where drug and protein sub-sequences interact through bilinear attention, achieving interpretable interaction maps. DrugBAN (Bai *et al*., 2023) [19] employs a bilinear attention network with domain adaptation to improve generalization to novel drug–target pairs. BACPI (Wang *et al*., 2022) [18] introduces bi-directional attention between drug molecular graphs and protein sequences, enabling fine-grained cross-modal interaction modelling.

Protein foundation models (ProtTrans, ESM-2) pre-trained on hundreds of millions of sequences produce contextual residue embeddings that serve as powerful protein encoders for DTI, replacing sequence-level CNNs and dramatically reducing the need for labelled DTI data.

4.5 Autoencoder and Variational Autoencoder Approaches

Autoencoders learn compact latent representations by reconstructing inputs through encoder-decoder bottlenecks. Variational autoencoders (VAEs) impose a probabilistic prior on the latent space, enabling generative sampling of novel molecules. In DTI, autoencoders are used to learn drug and protein embeddings in a unified latent space where proximity implies interaction. MONN (Li *et al*., 2020) uses a non-covalent interaction network to simultaneously predict drug–protein interaction maps and binding affinity, with an autoencoder-derived protein pocket representation.

4.6 Generative Models and Data Augmentation

Generative models GANs and diffusion models are increasingly applied to generate novel drug-like molecules or augment training data for rare drug–target pairs. MolGAN (De Cao and Kipf, 2018) generates small molecular graphs using a GAN framework trained with reinforcement learning reward signals for drug-likeness and target affinity. ORGAN (Guimaraes *et al*., 2018) uses a recurrent GAN on SMILES strings with pharmacological objectives. Generated molecules can augment training sets for underrepresented target classes, partially alleviating data imbalance.

4.7 Multi-Modal Fusion Frameworks

Real-world DTI prediction benefits from integrating multiple heterogeneous information sources: chemical fingerprints, sequence embeddings, three-dimensional coordinates, gene expression profiles, and knowledge graph embeddings. DeepPurpose (Huang *et al*., 2020) [16] provides a modular library supporting 15 drug and protein encoders in combination, enabling rapid benchmarking of architecture choices. Multi-modal fusion strategies early fusion (concatenation of raw features), intermediate fusion (concatenation of latent embeddings), and late fusion (ensemble of predictions) each have different inductive biases. Attention-based cross-modal fusion, as in DrugBAN, achieves superior performance by explicitly modelling drug–protein co-representation.

Table 1: Comparison of Deep Learning Architectures for Drug–Target Interaction Prediction

Architecture	Key Features	Input Representation	Notable Models	Limitations
Convolutional Neural Network (CNN)	Local feature extraction; parameter sharing	SMILES sequences; protein sequences	DeepDTI, WideDTA, GraphDTA (conv branch)	Limited global context; fixed kernel size
Recurrent Neural Network (RNN/LSTM)	Sequential dependency modeling	SMILES strings; amino acid sequences	DeepConvDTI (LSTM branch)	Vanishing gradients; slow training
Graph Neural Network (GNN)	Molecular topology; permutation invariant	Molecular graphs; protein contact maps	GraphDTA, MGraphDTA, DGAT	Scalability to large graphs; over-smoothing
Transformer / Attention	Long-range dependency; self-attention	Tokenized SMILES; ESM embeddings	MolTrans, DrugBAN, BACPI	High compute; requires large data
Autoencoder (VAE/AE)	Latent space learning; generative	Fingerprints; embeddings	MONN, DrugVQA	Reconstruction loss vs. task loss
Generative Adversarial Network (GAN)	Data augmentation; de novo design	Latent molecular vectors	MolGAN, ORGAN	Mode collapse; training instability
Multi-modal Fusion	Combines heterogeneous data sources	Sequence + graph + 3D structure	DeepPurpose, MINN, AttentionDTA	Data integration complexity
Capsule Network	Part-whole relationships; equivariance	Molecular graphs	CapsNet-DTI	Computationally expensive; nascent

5. Benchmark Datasets and Evaluation Metrics

5.1 Key Datasets

Rigorous evaluation requires standardized benchmark datasets with well-defined positive and negative interaction sets. Table 2 summarizes the most widely used DTI databases. The Davis dataset comprises binding affinities (Kd) for 68 kinase inhibitors across 442 kinase domains, making it the gold standard for affinity regression. The KIBA dataset integrates Ki, Kd, and IC50 values through a computational weighting scheme, offering a larger and more heterogeneous benchmark. BindingDB and ChEMBL provide millions of bioactivity measurements across diverse chemical and proteomic spaces, while PDBbind includes experimentally determined three-dimensional co-crystal structures.

Table 2: Summary of Benchmark Datasets for Drug–Target Interaction Prediction

Dataset	Drugs	Targets	Interactions	Data Type	Key Reference
Davis	68	442	30,056	Kd values	Davis *et al*., 2011 [31]
KIBA	2,111	229	118,254	KIBA scores	Tang *et al*., 2014 [32]
BindingDB	>900,000	>8,000	>2.9 M	Ki, IC50, Kd	Liu *et al*., 2007
ChEMBL	>2.4 M	>15,000	>17 M	Bioactivity	Mendez *et al*., 2019 [34]
DrugBank	14,051	4,900+	Varied	Approved/exp.	Wishart *et al*., 2022 [35]
PDBbind	~22,000	~22,000	~22,000	Kd/Ki (3D)	Wang *et al*., 2004
DTC	4,276	1,050+	4.5 M+	Ki, IC50	Tanoli *et al*., 2018 [52]
STITCH	~0.5 M	~9.6 M	~15 M	Combined score	Szklarczyk *et al*., 2016 [37]
UniProt/Swiss-Prot	N/A	~570,000	—	Protein sequences	The UniProt Consortium, 2023 [51]
PubChem BioAssay	>1 M	>10,000	>250 M	Bioactivity	Kim *et al*., 2021 [36]

5.2 Evaluation Metrics

The concordance index (CI) measures the probability that a model ranks a higher-affinity pair above a lower-affinity pair, analogous to the area under the ROC curve (AUROC) for ranking. Mean squared error (MSE) and Pearson correlation coefficient (R) quantify regression accuracy. For binary classification, AUROC, area under the precision-recall curve (AUPR), F1 score, and Matthew's correlation coefficient (MCC) are reported. The choice of dataset splitting strategy random, scaffold-based (for drug generalization), target-based (for protein generalization), or temporal critically affects reported performance and should match the intended deployment scenario.

5.3 Performance Comparison

Table 3 presents the concordance index of selected DL models on the Davis and KIBA benchmarks, illustrating the progressive improvement from CNN-only architectures (DeepDTA, 2018) to GNN-transformer hybrids (DrugBAN, ICAN, 2022–2023). Notably, performance gains have been accompanied by increasing model complexity, raising questions about practical deployability and overfitting to benchmark-specific artifacts.

Table 3: Performance Comparison of Deep Learning Models on Davis and KIBA Benchmarks (Concordance Index, CI)

Model	Architecture	Davis CI	KIBA CI	Year
DeepDTA	CNN (seq only)	0.878	0.863	2018
WideDTA	CNN + domains	0.886	0.875	2019
GraphDTA	GNN + CNN	0.893	0.891	2021
AttentionDTA	CNN + Attention	0.892	0.882	2021
MGraphDTA	Multi-scale GNN	0.900	0.904	2022
BACPI	Bilinear Attention	0.901	0.898	2022
DrugBAN	GNN + Transformer	0.907	0.912	2022
ColdDTA	Meta-learning	0.897	0.895	2023
ICAN	Interaction-aware GNN	0.910	0.915	2023

6. Challenges in Deep Learning-Based DTI Prediction

6.1 Data Scarcity, Imbalance, and Negative Sample Bias

Publicly available DTI datasets are heavily imbalanced: known interacting pairs are vastly outnumbered by non-interacting pairs, yet true negatives are rarely experimentally confirmed most apparent non-interactions are simply 'unknown.' This positive-unlabelled (PU) learning problem biases classifiers towards positive predictions and inflates apparent accuracy. Strategies including random negative sampling, semi-supervised PU learning, and contrastive self-supervised learning have been proposed, but no consensus solution exists.

6.2 Cold-Start Generalization

Standard random-split evaluations significantly overestimate real-world performance because training and test sets share structurally similar drugs or sequence-similar proteins. Scaffold-split, cluster-split, and full cold-start (novel drug AND novel target) evaluations reveal far lower generalization, particularly for CNNs. Meta-learning frameworks (MAML, ProtoNet) and pre-trained foundation models show promise for cold-start DTI prediction but remain significantly below warm-start performance.

6.3 Interpretability and Explainability

Clinical deployment of DTI models demands mechanistic insights: which atoms, residues, or interactions drive predictions? Attention weights in transformer and GAT models provide correlative but not causal explanations. Post-hoc methods—SHAP (SHapley Additive exPlanations), GNNExplainer, Integrated Gradients identify influential features but may not faithfully represent model reasoning. Developing intrinsically interpretable DL architectures for DTI remains an open and pressing challenge.

6.4 Three-Dimensional Structure Integration

Despite the centrality of three-dimensional binding geometry to DTI, most DL models operate on sequences or two-dimensional graphs due to limited experimental structural data. AlphaFold2 and ESMFold have dramatically expanded the availability of predicted protein structures, but predicted structure quality varies, and accounting for conformational flexibility (induced fit, allostery) remains unresolved. Equivariant neural networks that respect the rotational and translational symmetries of three-dimensional molecular geometry are an active research frontier.

6.5 Data Heterogeneity and Reproducibility

Reported performance metrics vary substantially across studies due to differences in train-test splitting, negative sampling, hyperparameter optimization, and evaluation metric selection. The absence of standardized benchmarking pipelines analogous to the GLUE benchmark in NLP hinders fair comparison. Initiatives such as the Therapeutics Data Commons (TDC) and the PEER benchmark aim to address this by providing canonical splits and leaderboards.

7. Emerging Approaches and Future Perspectives

7.1 Foundation Models for Drug Discovery

Large language models trained on vast chemical and biological corpora analogous to GPT-4 in NLP are emerging as universal molecular intelligence platforms. Models such as Galactica, ChemGPT, and BioMedGPT learn joint representations of small molecules, proteins, genes, and clinical phenotypes, enabling zero-shot and few-shot DTI prediction without task-specific fine-tuning. These foundation models represent a paradigm shift from task-specific architectures to general-purpose scientific AI.

7.2 AlphaFold3 and Co-Folding Integration

AlphaFold3 (Abramson *et al*., 2024) [48] extends structure prediction to ligand-protein complexes, nucleic acids, and post-translational modifications, providing predicted binding pose and interaction geometry. Integrating AlphaFold3-predicted complex structures into DTI prediction pipelines—using three-dimensional GNNs or physics-informed neural networks—is expected to substantially improve accuracy for challenging targets without experimental structures, including intrinsically disordered proteins and membrane receptors.

7.3 Knowledge Graph-Augmented Models

Biological knowledge graphs integrating drug, protein, disease, gene, pathway, and phenotype nodes provide rich relational context for DTI prediction. Graph attention networks operating on knowledge graphs (e.g., DrugKGC, HetioNet) can infer DTIs through multi-hop reasoning: a drug may interact with a target because they share a molecular pathway linked through intermediate entities. Combining GNN-based molecular encoders with knowledge graph embeddings represents a powerful direction for mechanistically grounded DTI prediction.

7.4 Federated Learning for Privacy-Preserving DTI

Pharmaceutical companies possess large proprietary DTI datasets that cannot be shared due to intellectual property constraints. Federated learning enables collaborative model training across institutional silos without data sharing, preserving privacy while aggregating global knowledge. Preliminary studies demonstrate that federated DTI models approach the performance of centrally trained counterparts, opening the possibility of pre-competitive data collaboration at scale.

7.5 Multi-Omics Integration

DTI prediction increasingly benefits from multi-omics context: gene expression profiles, proteomics, metabolomics, and phenomics data capture the cellular environment in which drug–target binding occurs. Integrating these modalities through multi-view learning or heterogeneous graph networks produces pharmacogenomically aware DTI models, enabling personalized medicine applications where predictions account for patient-specific target expression and genetic variation.

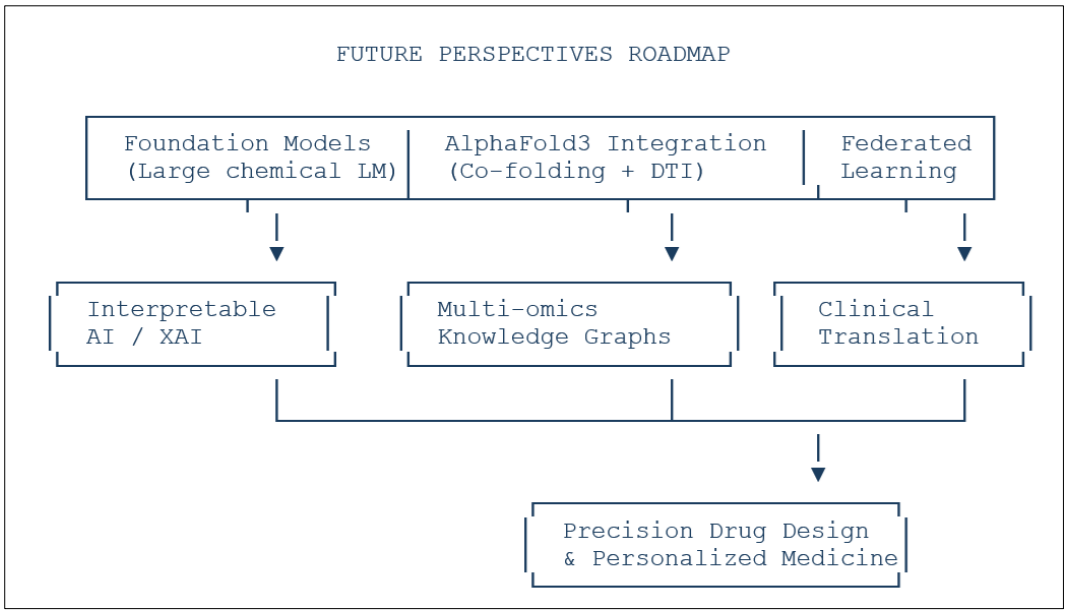


Fig 4: Roadmap of future research directions in deep learning-based drug–target interaction prediction, highlighting foundation models, structural prediction integration, interpretability, federated learning, and personalized medicine.

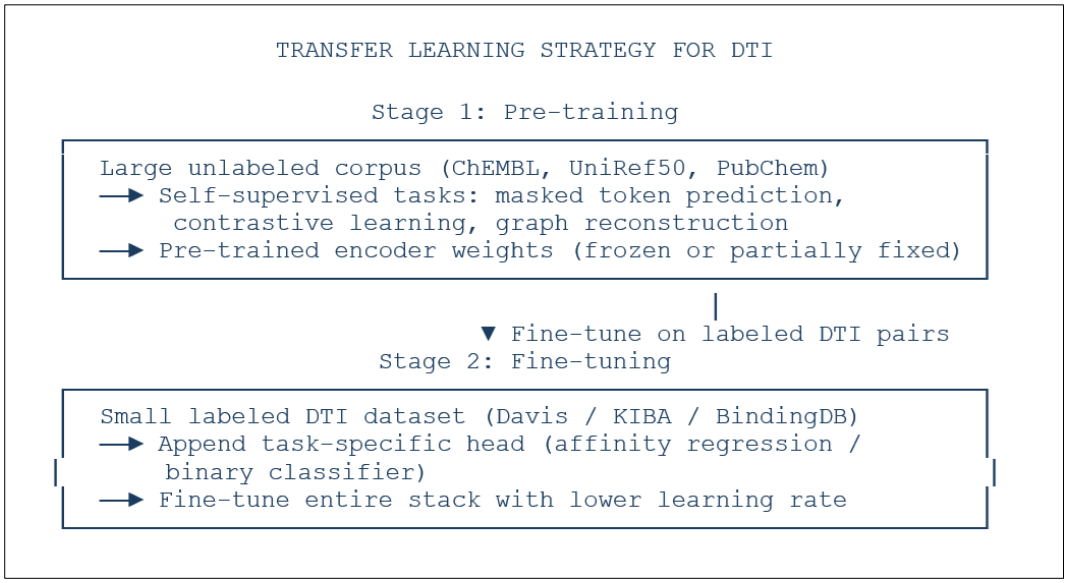


Fig 3: Two-stage transfer learning strategy for DTI prediction. Stage 1: self-supervised pre-training on large unlabeled chemical and protein corpora. Stage 2: fine-tuning on small labelled DTI datasets with a task-specific prediction head.

Table 5: Summary of Key Challenges and Proposed Solutions in Deep Learning-Based DTI Prediction

Challenge	Current Approaches	Future Directions
Data scarcity & imbalance	Transfer learning; SMOTE; data augmentation	Few-shot learning; generative augmentation (GANs, diffusion models)
Negative sample bias	Semi-supervised learning; PU learning	Contrastive self-supervised learning on unlabeled pairs
Model interpretability	Attention heatmaps; SHAP values	Graph-based explainability (GNNExplainer); causal inference
Cold-start (unseen entities)	Meta-learning (MAML); zero-shot	Foundation models pretrained on chemical universe
3D structure availability	AlphaFold-predicted structures	End-to-end co-folding + DTI (AlphaFold3 paradigm)
Reproducibility gaps	Open benchmarks; fixed splits	Standardized leaderboards (PEER, Therapeutics Data Commons)
Multi-target polypharmacology	Multi-task learning	Network pharmacology + GNNs; drug-disease-target KG

8. Applications in Drug Discovery

8.1 Virtual Screening and Hit Identification

DL-based DTI models enable ultra-large-scale virtual screening of billion-compound libraries against disease-relevant targets in silico, prioritizing compounds for experimental validation at a fraction of the cost of physical high-throughput screening. Stokes *et al*. (2020) demonstrated this paradigm by using a message-passing GNN to screen 107 million compounds against *E. coli* growth inhibition, identifying halicin a structurally novel antibiotic subsequently validated experimentally.

8.2 Drug Repurposing

Drug repurposing leverages existing safety profiles by repositioning approved drugs for new indications. DL-based DTI prediction identifies previously unknown interactions between approved drugs and disease-relevant targets, as exemplified by computational predictions of COVID-19 drug candidates during the pandemic. Network-based repurposing models integrate DTI predictions with disease gene networks to rank approved drugs by mechanistic relevance to new indications.

8.3 Polypharmacology and Off-Target Prediction

Most approved drugs exhibit measurable activity at multiple targets beyond their primary indication. DL models trained on large bioactivity databases can predict the full target interaction profile of a drug, identifying potential off-target toxicity liabilities and revealing unexpected therapeutic opportunities. This polypharmacology-aware DTI prediction is particularly valuable in CNS drug discovery, where pleiotropy may be beneficial.

8.4 De Novo Drug Design

Generative DL models VAEs, GANs, diffusion models, and flow-matching networks can design novel molecular structures optimized for target affinity, selectivity, and drug-likeness. These approaches are conditioned on DTI prediction models that serve as reward functions, guiding generation towards the desired target binding profile. Reinforcement learning from human feedback (RLHF)-inspired frameworks are emerging as powerful tools for iterative molecular optimization.

9. CONCLUSION

Deep learning has emerged as the dominant paradigm for computational drug–target interaction prediction, surpassing classical methods in accuracy, scalability, and generalizability across diverse chemical and proteomic spaces. The trajectory from early CNN-based sequence models (2018) to multi-modal GNN-transformer hybrids (2022–2023) and nascent foundation models (2024–) reflects the extraordinary pace of innovation. Critical advances include the adoption of graph neural networks for topology-aware molecular encoding, pre-trained protein language models for rich protein representations, and attention mechanisms for interpretable cross-modal interaction modelling.

Nevertheless, substantial challenges persist. Data imbalance and negative sample uncertainty remain fundamental statistical problems. Cold-start generalization predicting interactions for entirely novel drugs and targets is far from solved. Interpretability lags behind accuracy, limiting clinical translation. Integration of three-dimensional structural information, particularly for flexible and membrane-associated targets, remains an area of active development.

Looking forward, the convergence of foundation models, AlphaFold3-enabled structural predictions, knowledge graph reasoning, federated learning, and multi-omics integration promises to elevate DTI prediction from a research benchmark exercise to a genuine engine of clinical drug discovery. We anticipate that DL-based DTI models, combined with active learning and automated experimental validation platforms, will compress the drug discovery timeline and enable the development of therapeutics for diseases that have historically resisted intervention.

10. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of MM College of Pharmacy, Maharishi Markendeswar (Deemed to be University), Mullana, and Gateway College of Pharmacy, Sonipat, Haryana. The authors declare no conflict of interest.

REFERENCES

1. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–444. https://doi.org/10.1038/nature14539
2. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today. 2018;23(6):1241–1250.
3. Wen M, Zhang Z, Niu S, *et al*. Deep-learning-based drug-target interaction prediction. J Proteome Res. 2017;16(4):1–10.
4. Öztürk H, Özgür A, Ozkirimli E. DeepDTA: deep drug-target binding affinity prediction. Bioinformatics. 2018;34(17):i821–i829.
5. Nguyen T, Le H, Quinn TP, Nguyen T, Le TD, Venkatesh S. GraphDTA: predicting drug-target binding affinity with graph neural networks. Bioinformatics. 2021;37(8):1140–1147.
6. Tian K, Shao M, Wang Y, *et al*. Boosting compound-protein interaction prediction by deep learning. Methods. 2016;110:64–72.
7. Abbasi K, Raza A, Hassanpour H, *et al*. DeepCDA: deep cross-domain compound-protein affinity prediction. Bioinformatics. 2020;36(22–23):5461–5467.
8. Wang K, Zhou R, Li Y, Li M. DeepDTAF: a deep learning method to predict protein-ligand binding affinity. Brief Bioinform. 2021;22(5):bbab072.
9. Zhao Q, Duan G, Yang M, *et al*. AttentionDTA: prediction of drug-target binding affinity using attention model. IEEE Int Conf Bioinformatics Biomed. 2019;64–69.
10. Peng J, Li J, Shang X. A learning-based method for drug-target interaction prediction based on feature representation

learning and deep neural network. BMC Bioinformatics. 2020;21(Suppl 13):394.
11. Feng Q, Dueva E, Cherkasov A, Ester M. PADME: a deep learning-based framework for drug-target interaction prediction. arXiv. 2018;1807.09741.
12. Tsubaki M, Tomii K, Sese J. Compound-protein interaction prediction with end-to-end learning of neural networks for graphs and sequences. Bioinformatics. 2019;35(2):309–318.
13. Lee I, Keum J, Nam H. DeepConvDTI: prediction of drug-target interactions via deep learning with convolution on protein sequences. PLOS Comput Biol. 2019;15(6):e1007129.
14. Lim J, Ryu S, Park K, *et al*. Predicting drug-target interaction using a novel graph neural network with 3D information-embedded representation. J Chem Inf Model. 2019;59(9):3981–3988.
15. Sun M, Zhao S, Gilvary C, *et al*. Graph convolutional networks for computational drug development and discovery. Brief Bioinform. 2020;21(3):919–935.
16. Huang K, Xiao C, Hoang T, *et al*. DeepPurpose: a deep learning library for drug-target interaction prediction. Bioinformatics. 2020;36(22–23):5545–5547.
17. Yang Z, Zhong W, Zhao L, Chen CYC. MGraphDTA: deep multiscale graph neural network for explainable drug-target binding affinity prediction. Chem Sci. 2022;13(3):816–833.
18. Wang Y, Jiang M, Li Z, *et al*. BACPI: a bi-directional attention neural network for compound-protein interaction and binding affinity prediction. Bioinformatics. 2022;38(7):1995–2002.
19. Bai P, Miljkovic F, John B, Lu H. Interpretable bilinear attention network with domain adaptation improves drug-target prediction. Nat Mach Intell. 2023;5:126–136.
20. Jin Z, Wei Y, Xu P, Zhang K. ColdDTA: utilizing data augmentation and feature optimization for drug-target binding affinity prediction. Comput Biol Med. 2023;152:106336.
21. Liu B, Pliakos K, Vens C, Tsoumakas G. Drug-target interaction prediction via an ensemble of multitask networks. Artif Intell Med. 2021;121:102179.
22. Bagherian M, Sabeti E, Wang K, *et al*. Machine learning approaches and databases for prediction of drug-target interaction: a survey paper. Brief Bioinform. 2021;22(1):247–269.
23. Sachdev K, Gupta MK. A comprehensive review of feature based methods for drug target interaction prediction. J Biomed Inform. 2019;93:103159.
24. Thafar MA, Olayan RS, Ashoor H, *et al*. DTiGEMS+: drug-target interaction prediction using graph embedding, graph mining, and similarity-based techniques. J Cheminform. 2020;12(1):44.
25. Ye Q, Zhang C, Ju CJT, Wang W. Unified rational protein engineering with sequence-based deep representation learning. Nat Methods. 2021;18:1546–1555.
26. Elnaggar A, Heinzinger M, Dallago C, *et al*. ProtTrans: towards cracking the language of life's code through self-supervised deep learning and high performance computing. IEEE Trans Pattern Anal Mach Intell. 2022;44(10):7112–7127.
27. Rives A, Meier J, Sercu T, *et al*. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. Proc Natl Acad Sci USA. 2021;118(15):e2016239118.
28. Stokes JM, Yang K, Swanson K, *et al*. A deep learning approach to antibiotic discovery. Cell. 2020;180(4):688–702.
29. Zheng S, Li Y, Chen S, Xu J, Yang Y. Predicting drug-protein interaction using quasi-visual question answering system. Nat Mach Intell. 2020;2:134–140.
30. Liu T, Altman RB. Using multiple microarray datasets to study compound-target interactions. Bioinformatics. 2009;25(21):2791–2797.
31. Davis MI, Hunt JP, Herrgard S, *et al*. Comprehensive analysis of kinase inhibitor selectivity. Nat Biotechnol. 2011;29(11):1046–1051.
32. Tang J, Szwajda A, Shakyawar S, *et al*. Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. J Chem Inf Model. 2014;54(3):735–743.
33. Liu Z, Li Y, Han L, *et al*. PDB-wide collection of binding data: current status of the PDBbind database. Bioinformatics. 2015;31(3):405–412.
34. Mendez D, Gaulton A, Bento AP, *et al*. ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Res. 2019;47(D1):D930–D940.
35. Wishart DS, Feunang YD, Guo AC, *et al*. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 2018;46(D1):D1074–D1082.
36. Kim S, Chen J, Cheng T, *et al*. PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Res. 2021;49(D1):D1388–D1395.
37. Szklarczyk D, Santos A, von Mering C, *et al*. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. Nucleic Acids Res. 2016;44(D1):D380–D384.
38. Gilmer J, Schütt AT, Brockherde F, *et al*. Neural message passing for quantum chemistry. Proc Int Conf Mach Learn. 2017;70:1263–1272.
39. Vaswani A, Shazeer N, Parmar N, *et al*. Attention is all you need. Adv Neural Inf Process Syst. 2017;30:5998–6008.
40. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735–1780.
41. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv. 2017;1609.02907.
42. Xu K, Hu W, Leskovec J, Jegelka S. How powerful are graph neural networks? arXiv. 2019;1810.00826.
43. Goodfellow I, Pouget-Abadie J, Mirza M, *et al*. Generative adversarial nets. Adv Neural Inf Process Syst. 2014;27.
44. Kingma DP, Welling M. Auto-encoding variational bayes. arXiv. 2014;1312.6114.
45. Yang K, Swanson K, Jin W, *et al*. Analyzing learned molecular representations for property prediction. J Chem Inf Model. 2019;59(8):3370–3388.

46. Hu W, Liu B, Gomes J, *et al*. Strategies for pre-training graph neural networks. arXiv. 2020;1905.12265.
47. Jumper J, Evans R, Pritzel A, *et al*. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–589.
48. Abramson J, Adler J, Dunger J, *et al*. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630:493–500.
49. Réau M, Renaud N, Xue LC, Bonvin AMJJ. DeepRank-GNN: a graph neural network framework to learn patterns in protein-protein interfaces. Bioinformatics. 2023;39(1):btac759.
50. Kexin H, Cao X, Zitnik M. Few-shot graph learning for molecular property prediction. Proc Web Conf. 2021;2559–2567.
51. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2023. Nucleic Acids Res. 2023;51(D1):D523–D531.
52. Tanoli Z, Alam Z, Vähä-Koskela M, *et al*. Drug Target Commons 2.0: a community platform for systematic analysis of drug-target interaction profiles. Database. 2018;2018:bay083.
53. Chen L, Tan X, Wang D, *et al*. TransformerCPI: improving compound-protein interaction prediction by sequence-based deep learning with self-attention mechanism and label reversal experiments. Bioinformatics. 2020;36(16):4406–4414.

About the Corresponding Author

Sachin Sharma is an Assistant Professor at ITS College of Pharmacy. He is actively engaged in pharmaceutical education and academic research, contributing to the advancement of pharmacy studies through teaching, student mentorship, and participation in scholarly and professional activities in the field of pharmaceutical sciences.