Pakistan Journal of Bioinformatics (PJB)
Volume 1, Issue 1, 2026 — ISSN: 2222-7628

Predicting Drug–Target Interactions from Drug Structure and Protein Sequence Using Novel Convolutional Neural Networks: A Review

Laiba Riaz, Anam Saleem
Department of Computer Science, University of Agriculture, Faisalabad
laibariaz9512@gmail.com

Abstract

Drug–target interaction prediction is a key area of computational drug discovery, thereby making it possible to reveal presumptive targets and speed up medication repositioning. Experimental identification of DTIs is costly, time-consuming, and in feasible at large scale. Recent deep learning advances, especially convolutional neural Networks, or CNNs, have considerably enhanced DTI prediction by learning complicated hierarchical patterns directly from drug molecular structures and protein sequences. This Review provides a comprehensive overview of CNN-based prediction methods for DTI, covering representation learning strategies, network architectures, benchmark datasets, evaluation metrics, comparative performance analysis, challenges, and future research directions. This study highlights the superiority of CNN models over traditional machine learning approaches and discusses emerging trends such as attention mechanisms, graph neural networks, and multi-modal data integration.

Keywords: Bioinformatics; Computational Biology; Data Science

1 Introduction

The identification of interactions between chemical compounds and biological targets provides an play a fundamental role in modern drug discovery and development. A successful drug must bind effectively toward a certain protein target in order to modulate its biological function.

However, experimentally testing every possible drug–target pair is not practical because of the huge combinatorial space. In fact, with over 10,000 approved drugs and almost 20,000 human proteins, the potential. The interaction space exceeds 200 million combinations.

Traditional experimental methods include high-throughput screening and binding assays.They are expensive as well as time-consuming. Therefore, computational methods have become tools that reduce the number of candidate interactions before experimental validation.

Early Computational strategies relied heavily on molecular docking, ligand-based similarity searches, a chemogenomic approaches. Although powerful in certain contexts, these methods necessitate detailed structural information or prior biological knowledge, limiting their scalability and generalizability.

Machine learning methods introduced a paradigm shift by learning predictive patterns from existing interaction data. However, traditional ML models using Support Vector Ma chines and Random Forests rely on hand-crafted features, which can be inadequate to depict Complex non-linear relationships. Deep learning, especially convolutional neural networks.CNNs learn the representation of the input through the application of convolutional filters. This allows the CNN-based DTI model to learn the intricate relationships between the chemical substructures and protein patterns that correspond to binding affinity or probability. These limitations are addressed by CNNs, which automatically learn hierarchical feature representations from raw input data.(1, 2).

2 Traditional Approaches for DTI Prediction

2.1 Molecular Docking

Molecular docking simulates the physical interaction between a drug molecule and a protein binding site. Although docking provides structural insights, it is computationally expensive and highly dependent on accurate protein structures. Molecular docking is a computational tool used to predict how a small molecule-ligand (like a drug)-binds to its larger molecule-receptor (like a protein)-to form a stable complex, a step intrinsic to drug discovery. This technique simulates molecular ”handshakes” in order to find out the best fit, or pose, and binding strength, or score, of potential ligands to their receptors. It involves sampling potential ligand orientations (poses) within the protein’s binding site and scoring these poses to identify the most favorable interactions-like hydrogen bonds and van der Waals forces-that accelerate the quest for new drugs and the explanation of biochemical processes.

2.2 Ligand-Based Methods

Ligand-based approaches assume that chemically similar compounds bind to similar targets. These methods perform poorly when novel compounds lack known analogs.Ligand-based methods in computational chemistry and drug discovery use information from known active molecules (ligands) to find new drug candidates, especially when the 3D structure of the protein target is unknown.

2.3 Chemogenomic Approaches

Chemogenomics combines chemical and genomic information, but it is often based on incomplete biological information and curated databases. Chemogenomic approaches combine chemistry with genomics in mapping interactions between chemical compounds and biological targets (like proteins), using large libraries of diverse molecules along with genomic information to understand disease pathways, discover new drugs, validate targets, and even allow personalized medicine that links patient genetics to drug response.

3 Deep Learning for Drug-Target Interaction Prediction

Deep Learning (DL) revolutionizes Drug-Target Interaction (DTI) prediction by learning complex patterns from drug (SMILES, graphs) and protein (sequences, structures) data, overcoming traditional limitations, with models like GNNs, Transformers, CNNs, and RNNs identifying potential drug-target pairs for faster discovery, though challenges in data scarcity, interpretability, and uncertainty quantification persist, leading to advanced methods like evidential learning.

3.1 Why CNNs for DTI?

CNNs excel at capturing spatial and sequential dependencies, making them suitable for modeling chemical substructures and protein motifs. Unlike traditional ML models, CNNs eliminate the need for manual feature engineering.CNNs are used for Diffusion Tensor Imaging (DTI) because they excel at learning spatial patterns in image data, allowing them to accurately estimate diffusion tensors from fewer measurements (reducing scan time), denoise images, classify brain conditions like Alzheimer's, and extract relevant features from raw data like protein sequences, outperforming traditional methods in speed and accuracy for complex DTI analysis.

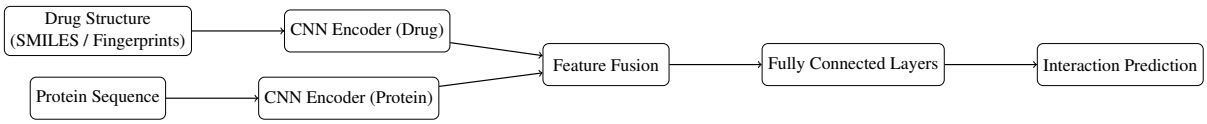

Figure 1: CNN-based DTI prediction workflow

4 Representation Learning

4.1 Drug Representation

The molecules or drugs are represented using molecules descriptors or fingerprints such as Morgan fingerprints, Extended Connectivity Fingerprints (ECFP), or SMILE representations. The representations include chemical structures or topologies and physicochemical characteristics. The SMILE string representations of molecules can be processed into either vector or image representations amenable to CNNs .

Example: A SMILES expression describing the molecular structure of a drug compound might be represented as a 2D matrix where each row denotes the atomic properties and the columns the bonding relationships as an 'image' input for the convolutional layers.

4.2 Protein Representation

Proteins are usually modeled by a set of sequence features, which are

Amino acid composition

Frequencies associated with dipeptides

Pseudo-amino acid

Physicochemical property vectors (hydrophobicity

Such features can then be rearranged into 2D matrices or aggregated as embeddings that can work as the input for CNNs.

4.3 Joint Representation

Joint embedding of drug and protein features enables CNNs to learn cross-domain interactions efficiently. Joint representation means merging diverse drug (structure/SMILES) and protein (sequence/structure) features into a unified input, often by converting them to image-like matrices or graph embeddings, to let the Convolutional Neural Network learn combined patterns, capturing complex relationships for better prediction than using single-modality features alone. Models use techniques like converting SMILES/sequences to images (MPS2IT),

graph-based features (GNNs), or concatenated vectors, allowing the CNN to extract both local atomic/residue details and global structural context for accurate interaction prediction.

5 CNN Architectures for DTI Prediction

CNN-based architectures consist of convolutional layers, pooling layers, and fully connected layers. Advanced architectures incorporate dropout, batch normalization, residual connections, and attention mechanisms.In general, CNN models for predicting DTI could consist of:

1. Input layers for drugs and proteins

2.Convolutional Layers for detecting local patterns in chemical and sequence features

3. Pooling layers to reduce the dimensionality and extract dominant features

4.Full connected layers for combination of feature and output interaction probability. Other improvements that can be added are the addition of dropout layers that address the overfitting problem, residual connections that enable deeper designs, and attention mechanisms that target important residues or substructures (3).

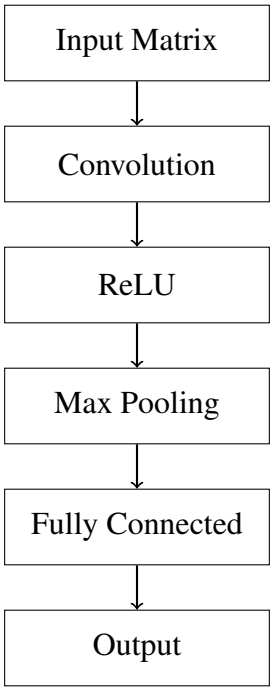

Figure 2: CNN architecture for DTI prediction

6 Datasets Used for DTI Prediction

CNN-based models of DTIs are trained on known interactions in these datasets:

Enzyme, GPCRs, Ion Channels, Nuclear Receptor

Positive samples: drug-target pairs that interact

Negative samples: randomly selected drug-target pairs for which the interaction is unknown

Dataset balance is very important because interactions in the real world are sparser compared to the total combinations of drugs and targets. (4).

7 Evaluation Metrics and Case Studies

Common metrics include: accuracy, precision, recall, F1-score, AUC-ROC, MCC.

CNN models consistently outperform traditional ML methods(1)

Research indicates that CNN models perform better than traditional machine learning algorithms in predicting DTI. For example:

DeepConv-DTI reported an AUC value¿ 0.90 on enzyme, GPCR, and ion channel datasets (1)

DeepDTA and DeepACTION demonstrated a better generalization on novel compounds and proteins (2)

CNNs capture nonlinear, hierarchical relationships between drug substructures and protein motifs, providing better predictive power than feature-engineered methods.(3)

Table 1: Performance comparison

Model	Dataset	Acc	Prec	Rec	F1	AUC
DeepConv-DTI	Enzyme	0.92	0.90	0.91	0.91	0.93
DeepConv-DTI	GPCR	0.90	0.88	0.89	0.89	0.91
DeepDTA	Ion Channel	0.89	0.87	0.88	0.88	0.90
SVM	Enzyme	0.81	0.79	0.80	0.80	0.82

8 Challenges and Limitations

Although it performs well, CNN-based DTI prediction also encounters several issues:

1.Scarcity of data or imbalance: fewer known interactions compared to the search space of possible combinations.

2. Cold-start problem: It is difficult to predict interactions for novel compounds and proteins.

3.Interpretability: CNNs are often CNNs are often criticized for their limited interpretability, as the complex hierarchical feature representations make it challenging to explain model decisions.

4.Missing 3D structural information: sequence-based representations can leave out crucial interactions based on the three-dimensional folding of proteins, active-site geometry, and spatial orientation of ligands.

To deal with these problems, there are several approaches, such as techniques based on explainable AI, graph structures, and transfer learning (1, 4) .

9 Future Directions

Future enhancements for CNN-modelled DTI prediction could:

Graph Neural Networks for representation of molecules and protein structure.

Attention mechanisms for the key residues or moieties Multi-modal integration with omics data, docking, or network pharmacology.

Transfer learning and meta-learning for better predictions of new drugs.

Applications to emerging diseases (such as COVID-19) and Personalized Medicine.(2)

10 Conclusion

CNN-based DTI predictions offer a computationally efficient and effective solution to explore the possibility of binding interactions between drugs and targets. By incorporating information from chemical and protein sequences, CNNs have shown superiority over other ML models. However, future studies involving graph models, attention models, and multiple models would help in improving predictions and would open up hopes for speedy breakthroughs in drugs and their effective repurposing.(1, 4)

References

[1] Hu,S., Zhang, C., Chen, P., Gu, P., Zhang, J. and Wang, B., 2019. Predicting drug-target interactions from drug structure and protein sequence using novel convolutional neural networks. *BMC bioinformatics* , 20(Suppl 25), p.689.

[2] Lee, I., Keum, J., Nam, H. (2019). DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences. *PLoS computational biology*, 15(6), e1007129

[3] Öztürk, H., Özgür, A., Ozkirimli, E. (2018). DeepDTA: deep drug–target binding affinity prediction. *Bioinformatics*, 34(17), i821–i829.

[4] Gao, K., et al. (2020). Interpretable drug-target interaction prediction using deep learning. *Briefings in Bioinformatics*, 21(4), 1382–1397.

[5] 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(Suppl 13), 394.

[6] Meng, F. R., You, Z. H., Chen, X., Zhou, Y., An, J. Y. (2017). Prediction of drug–target interaction networks from the integration of protein sequences and drug chemical structures. *Molecules* , 22(7), 1119.