RESEARCH PAPER

Sequence-Driven Drug–Target Interaction Prediction Modelling Using Deep Learning Models with Cancer Specific Enhancements.

Tumakuru Nataraj Sowmya [1]*, Shivaprasad Nagappa [2]*, Siri Sanjay Bellary [2]

[1]*Department of Biotechnology and Bioinformatics, School of Life Sciences, JSS Academy of Higher Education and Research, Mysuru, Karnataka, India.*
[2]*Department of Electronics and Communication Engineering, SJCE, JSS Science and Technology University, Mysuru, Karnataka, India*

Received: 8th Oct, 2025; Revised: 1st Nov 2025; Accepted: 15th Nov, 2025; Available Online: 1st December, 2025

ABSTRACT
Recent breakthroughs in computational drug research have mostly focused on deep learning-based models for anticipating drug-target binding affinity (DTA), an important problem in determining effective therapeutics. This is especially essential in cancer research, where molecular heterogeneity, dynamic protein configurations, and the desire for personalised therapy options provide significant challenges. Traditional experimental tests, while precise, are time consuming, expensive, and difficult to scale across large chemical libraries. Structure-based computational techniques usually depend on regulated elements and high-resolution 3D data, limiting their utility in real-world biological applications. In addition, many existing DTA models lack disease-specific adaptability and interpretability, rendering them unsuitable for translational applications in oncology. To address these constraints, the current study introduces a modular, sequence-driven deep learning system that predicts binding affinity directly from raw Simplified Molecular Input Line Entry System (SMILES) strings and amino acid sequences, without the need for structure input. The full DAVIS dataset was used to benchmark several architectures, such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), and RNN and LSTM hybrid systems. Performance of the model was evaluated using robust metrics such as Mean Squared Error (MSE), Concordance Index (CI), Area Under the Precision-Recall Curve (AUPR), modified R^2 (R^2_m), Area Under the ROC Curve (AUROC), sensitivity and specificity. The results show that the RNN and LSTM architectures outperform CNN in capturing dependencies over long periods and enhancing sensitivity. To improve therapeutic relevance, the model was fine-tuned on a curated cancer-specific subset of Drug Affinity Validation of Interaction via Structure (DAVIS), and then extended into a hybrid CNN, RNN, LSTM, and Multi-Head Attention architecture. This final model showed increased ranking fidelity, interpretability, and predictive robustness in cancer-specific, molecularly diverse environments.
The current study fills the major gaps in DTA prediction research and presents a biologically informed, interpretable, and scalable deep learning approach with significant potential for personalised treatment screening and precision oncology.
Keywords: *Drug–Target Interaction, Binding Affinity, Oncology, Deep Learning, Multi-Head Attention, Convolutional Neural Network, Drug discovery*
How to cite this article: Sowmya TN, Nagappa S, Bellary SS, Sequence-Driven Drug–Target Interaction Prediction Modelling Using Deep Learning Models with Cancer Specific Enhancements. Int J Drug Deliv Technol. 2026;16(1): 86-99. DOI: 10.25258/ijddt.16.1.9
Source of support: Nil.
Conflict of interest: None

INTRODUCTION

Accelerating drug discovery demands the ability to predict drug-target interactions (DTIs) and treatment development, particularly in oncology. Determining how small-molecule compounds interact with target proteins is essential to comprehending the effectiveness and mechanism of drugs. Although they are accurate, traditional techniques like affinity chromatography, X-ray crystallography, and high-throughput screening are frequently time-consuming, expensive, and difficult to scale for large chemical libraries. Consequently, computer-based DTI prediction has emerged as a viable alternative that offers effectiveness, scalability, and reduced experimental expenses[1,2,3]. The ability of a medication to bind firmly and selectively to its biological target is a major factor in determining its therapeutic efficacy. Enzymes, receptors, and ion channels are examples of proteins that are important molecular targets in controlling cellular processes and disease pathways[4]. One important metric that shows how strongly a drug interacts with its target is binding affinity. Dissociation constants (Kd), inhibition constants (Ki), or IC_0[1] values are commonly used to quantify binding affinity, which indicates the effectiveness of ligand-receptor binding[5]. Ranking candidate compounds, optimizing lead molecules, repurposing current medications, and reducing negative off-target effects can all be greatly aided by accurate binding affinity prediction[6]. Computational methods utilizing chemical, genomic, and structural data have become more crucial in order to overcome the shortcomings of conventional laboratory-based DTI analysis[7]. Quantitative

**Author for Correspondence: sowmyatn@jssuni.edu.in*

structure–activity relationship (QSAR) models and molecular docking were two early computational techniques that mainly relied on manually created molecular descriptors and predetermined physicochemical characteristics[8]. Although these methods provided insightful information, they frequently had trouble generalizing across different molecular classes and required precise 3D structural information, which was not always accessible. The advancement of artificial intelligence (AI) and machine learning (ML), especially deep learning, has given researchers new tools that can automatically understand complicated nonlinear correlations from massive and heterogeneous datasets[9]. Recently, there has been the incorporation of deep learning techniques in computational drug discovery, which enables the use of deep learning for the complete modelling process from raw molecular data. Convolutional Neural Networks have been employed for extracting spatial representations from the Simplified Molecular Input Line Entry System string representation for local substructures that relate to bonding in molecules[10]. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) Networks, on the contrary, are efficient in learning long-term dependencies in protein sequences and ligand features, which are present in their representation[11]. These architectures enable it to learn from both local chemical patterns and global sequence information, thus increasing its predictability accuracy in binding affinity prediction tasks. Furthermore, the discipline has seen an advancement with graph approaches and attention mechanisms introducing graphs to represent molecules, thus capturing information related to topology and context, which might not be captured through sequential approaches[12,13]. Despite such achievements, there remain several gaps within current studies related to DTI prediction. First, most of the existing studies on DTI prediction do not provide mechanisms that are specifically tailored for cancer target-specific DTI prediction. Thus, such studies lack biological applicability specifically concerning disease-targeting drug design. Second, there remain very limited studies that have made comprehensive comparative analysis concerning different basic deep learning architectures such as CNN, RNN, or LSTM concerning common datasets such as DAVIS. Finally, there remain unexplored concepts concerning interpretability and transfer learning within DTI prediction. Most of the current DTI prediction models function as "black boxes" that do not provide an explanation for molecular or sequence-level characteristics that play a pivotal role within such predictions[14,15,16,17,18,19,20].

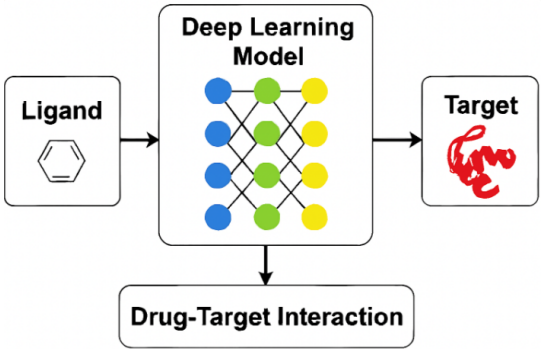


Figure 1: Ligand-Based Drug–Target Interaction Prediction Using Deep Learning

The ligand-based deep learning model depicted in Figure 1 provides a flexible and efficient framework for predicting the interactions between drugs and targets. The model utilizes the structural heterogeneity of ligands and the ability of deep learning models to quickly evaluate the effect of the ligands on the proteins. Taking all these considerations into account, this work is designed to tackle these identified flaws in previous literature to contribute to the area of DTI prediction in a cancer-related context. It is recognized that the methods currently in wide use in this area work very well in the general domain, but fail to consider the complexity that is so naturally present in cancer-related targets. Moreover, uninterpretable models and insufficient comparative studies among fundamental deep architectures impose obstacles to translational possibility. In order to mitigate these problems, this research introduces a novel, sequence-based deep learning architecture that depends on SMILES strings and protein sequences, obviating the need for structural knowledge in the standard 3D representation. Through combining approaches involving CNN, RNN, LSTM, and multi-head attention networks along with cancer-specific fine-tuning, this proposed architecture aspires to improve robustness in predictions, rankings, and real-world applicability. This selection on DTI is guided by its high importance in personalized therapy approaches, optimal drug recycling strategies, as well as diminished lab work, making DTI prediction systems a promising approach with scalability in contemporary drug development for cancer treatment.

MATERIALS AND METHODS

Data Description and Preprocessing

Dataset Source: A DAVIS benchmark dataset of 30,056 experimentally established binding affinity measurements for 68 different ligands/ drugs and 433 protein kinases (targets) was used. The raw affinity data, given as equilibrium dissociation constants (Kd) in nanomolar units, were transformed to negative logarithmic scale (pKd) values using the usual conversion equation.

$$pKd = -log10\frac{(Kd)}{1 * 10^9}$$

as used in previous DTI research. For numerical stability while training the model, pKd values were scaled to the interval [0,1]. In accordance with standard practice in

sequence-based prediction of DTIs, the dataset was categorized randomly into 80% training and 20% independent test sets, and 10% of the training set was reserved additionally for validation purposes in order to track convergence and prevent overfitting. Used a constant random seed of 42 to make train/test partitioning and model initialization reproducible. No stratification was performed, since the regression scenario does not require balanced class distributions; however, did check that drug–target pairs in the test set showed their absence in the training set to avoid data leakage. For hyperparameter tuning on some trials, used 5-fold cross-validation over the training set. This experimental configuration aligns with commonly used DAVIS evaluation procedures employed in DeepDTA and DeepPS, allowing for direct comparison with current literature.

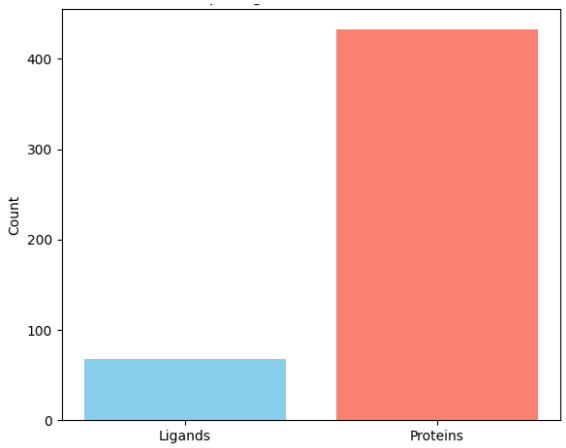


Figure 2: Distribution of unique ligands and protein ligands

As shown in Figure 2 presents the distribution of unique ligands and protein targets in the DAVIS dataset. The dataset includes 68 distinct small-molecule ligands and 433 distinct protein kinases, indicating a high protein diversity proportional to the number of medicines.

Table 1: Dataset Statistics

Feature	Value	Description
Data points	30,056	Total drug–target interaction pairs
Sequences	442	Unique protein sequences (targets)
Ligands	68	Unique small molecules (drugs)
Bioactivity	pKd	Binding affinity measured as $-\log_{10}(Kd)$
Subset Focus	Cancer Targets	Filtered subset focused on cancer-relevant proteins and interactions

The dataset employed in this study includes 30,056 DTI pairs, exhibiting a rich and diversified collection of biochemical correlations. It contains 442 unique protein sequences, each of which serves as a biological target, as well as 68 separate small-molecule ligands encoded as SMILES strings, as shown in Table 1.

The bioactivity of each interaction is quantified using pKd values, which express the binding affinity as the negative logarithm of the dissociation constant ($-\log_{10}(Kd)$). This provides a continuous and interpretable measure of interaction strength. To increase therapeutic relevance, a cancer-specific subset was filtered from the entire dataset, focussing on interactions with proteins thought to play roles in oncological pathways. This filtered subset enables specific evaluation of the performance of models in cancer pharmacology, where molecular heterogeneity and dynamic protein behaviour provide particular obstacles for predictive modelling.

Data Preprocessing: All pre-processing processes were designed to keep data consistent, decrease redundancy, and work with deep learning input formats.

SMILES String Processing

SMILES presentations of drugs were gathered from the BindingDB/DAVIS records. The RDKit library was used to standardise the SMILES strings in the following steps:

Canonicalization: All SMILES are converted to canonical form to guarantee their uniqueness.

Salt and Solvent Removal: Neutralisation of charged species and elimination of counterions or salts.

Deduplication: Duplicate entries are removed to ensure a one-to-one mapping among unique drugs and their sequences.

Length Normalization: The maximum token length was set to 100 characters; shorter sequences were padded, while longer ones were clipped to ensure consistent vector size.

Tokenization: To embed into neural networks, each character or atomic symbol in the SMILES string was tokenised and assigned integer indices.

Protein Sequence Processing:

Protein sequences corresponding to DAVIS targets were obtained from UniProtKB. The steps below were accomplished.

Sequence cleaning: Non-standard amino acids (B, J, O, U, X, and Z) were altered or removed to ensure proper FASTA formatting.

Truncation and padding: To avoid computational imbalance, sequences were standardised to a maximum of 1,000 amino acids.

Tokenization: Each amino acid residue was assigned a one-letter code and converted to an integer index.

Embedding: Model training involved embedding tokenised sequences into fixed-dimensional vectors.

Binding affinity conversion and normalization: As previously explained, the negative logarithm was used to convert all affinity values from K_d to pK_d. Following conversion, data were linearly normalised into the [0, 1] range to stabilise neural network training and provide uniform scaling across models.

Data splitting and sampling strategy: To assess generalisation and prevent data leakage, the dataset was randomly divided into three categories: training (80%),

validation (10%), and testing (10%). The splitting was done at the drug-target pair level, so that identical drug-target pairs did not appear in several subgroups. The random split technique was chosen to keep a balanced representation of medications and targets in each division.

Data Augmentation: Minor random noise perturbations and SMILES shuffling (while maintaining chemical equivalence) were used during training to reduce overfitting and improve generalisation. In addition, dropout regularisation and early halting were integrated into the training process to prevent overfitting hazards.

MODEL ARCHITECTURE

In order to ensure fair comparison, we assess five deep learning architectures for drug–target binding affinity prediction, making sure that each model uses the same input representation and training protocol. Proteins are represented as tokenized amino acid sequences or extracted binding-pocket subsequences, whereas drugs are represented as tokenized SMILES strings.

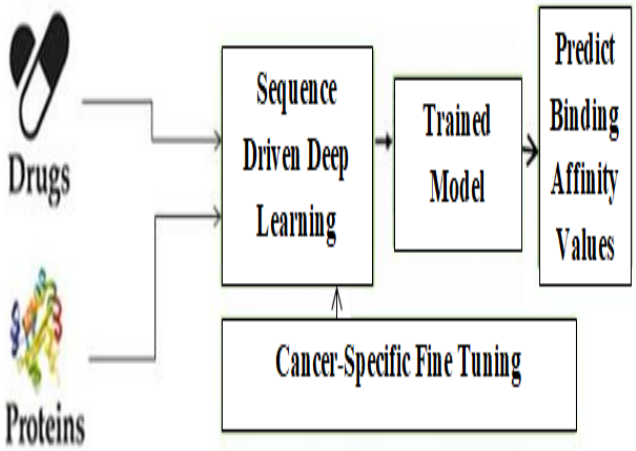


Figure 3: Deep Learning Architecture for DTA Prediction

Figure 3 illustrates a modular, sequence-driven deep learning pipeline for predicting drug–target binding affinity, tailored to cancer-specific applications. A CNN-based architecture captures spatial patterns in SMILES strings for drug representation, while a hybrid RNN and LSTM model captures sequential and contextual dependencies in protein sequences. The inputs are fed into a Sequence-Driven Deep Learning module, where SMILES and protein sequences are processed using CNN, RNN and LSTM architectures to learn biochemical and sequential patterns. The outputs from both encoders are concatenated to form a joint representation of each drug–target pair. These fused embeddings are then passed to a regression head that predicts binding affinity. The model is trained using MSE as a loss function and optimised with Adam, which includes learning rate scheduling. Evaluation measures, including R^2 score, MSE, AUROC, and sensitivity, are generated using an 80/20 split between train and test. Model checkpoints are saved based on validation results, and scatter plots are generated to show forecast accuracy.

The subsequent section is distinguished by the effective use of cancer-focused fine-tuning to enhance the clinical applicability of drug-target interaction predictions. A representation of this subset was then used to fine-tune three different deep learning models: This subset's representation was then utilised to fine-tune three separate deep learning models: a basic CNN, RNN, and LSTM architecture, and a leaner Multi-Head Attention model.

A base CNN, RNN, and LSTM architecture, as well as a more efficient Multi-Head Attention model. The trained model is then subjected to external chemicals from DrugBank, which contain 1063 pharmaceuticals or small molecules, allowing the system to generalise and estimate binding affinities for unknown compounds.

The model's application in cancer-related prediction is demonstrated by highlighting a specific target protein, CDK2_Human (UniProt ID: P24941). The results are a collection of Predicted Binding Affinity Values, which measure how strongly each molecule is expected to bind to the target protein. Cancer-Specific Fine-Tuning is utilised to improve clinical relevance by retraining the model on a handpicked subset of the DAVIS dataset focussing on oncological targets. This phase increases the sensitivity and precision of cancer medication discovery, making the system scalable and disease-adaptive.

The fine-tuned models utilized transfer learning, model weights initializing at baseline DTI weights from trained general models and retraining modelling specifically on cancer-related paired sets, model's outputs specializing cancer-relevant predictions. Thus, a final hybrid architecture was considered, which encompassed the benefits of CNNs, RNNs, LSTMs, and Multi-Head Attention. The hybrid architecture creates opportunity for local feature extraction, sequential modeling, and global interpretation, all in a combined architecture. This also offers greater generalizability across a variety of cancer-specific signaling interactions across local feature extraction, improved ranking fidelity across modeling degrees, and elucidation of clinically significant and biologically relevant regions in sequences. The models were evaluated based on regression metrics regression metrics associated with drug–target interaction prediction, such as MSE, R^2 score, and CI.

Table 2: Parameter Settings for CNN, RNN, and LSTM Models.

Parameter	CNN	RNN	LSTM
No. of Filters / Units	32×1, 32×2, 32×3	128, 256	128, 256
Filter Length (Compounds)	[4, 6, 8]	—	—
Filter Length (Proteins)	[4, 8, 12]	—	—
Hidden Neurons	1024, 1024, 512	1024, 512	1024, 512

Parameter	CNN	RNN	LSTM
Batch Size	256	256	256
Dropout	0.1	0.1	0.1
Optimizer	Adam	Adam	Adam
Learning Rate	0.001	0.001	0.001
Sequence Encoding	1D convolutional mapping	Integer embedding	Integer embedding
Compound Encoding	SMILES tokenized	SMILES tokenized	SMILES tokenized

Table 2 shows parameter settings for CNN, RNN, and LSTM Models. CNN uses multiple filter sizes to capture local patterns in both compound and protein sequences. RNN and LSTM focus on sequential dependencies, using embedded representations. All models share consistent training hyperparameters for fair benchmarking. Both models were trained on a filtered subset of 1,007 interactions focused on cancer-relevant proteins. CNN retained its multi-scale convolutional filters to capture binding patterns. RNN and LSTM leveraged sequential modeling to learn dependencies across encoded protein sequences.

FLOWCHART

A modular deep learning designed for DTA prediction is shown in Figure 4. The method starts with two biological inputs: a pharmacological molecule represented by a SMILES string and a target protein encoded as an amino acid sequence.

These raw formats are selected for their accessibility and scalability, allowing the model to function without requiring 3D structural data. To extract useful information, the SMILES string is sent through a CNN, which detects local chemical patterns and substructures related to binding behaviour. Simultaneously, the protein sequence is processed through RNN paired with LSTM units, allowing the model to learn long-term dependencies and sequential motifs that govern protein-ligand interactions. The outputs from both encoders are subsequently combined in a Feature Fusion stage to create a single representation.

This fusion may use concatenation, attention-based weighing, or dimensional alignment to assure compatibility and maximise information retention. The feature vector is supplied into a Prediction Model, often a fully connected regression layer, trained to translate the fused characteristics to a continuous binding affinity score (e.g., KdIC50) or a binary interaction label. The final result is the DTA Prediction, which estimates the likelihood or strength of the interaction between the medication and target. This prediction is assessed using robust measures like as MSE, CI, AUROC, AUPR, Sensitivity, and Specificity, which ensure both numerical accuracy and biological relevance. Overall, the flowchart captures a scalable, interpretable, and physiologically grounded approach to DTA modelling, ideal for cancer-specific screening and adaptable to various biomedical contexts.

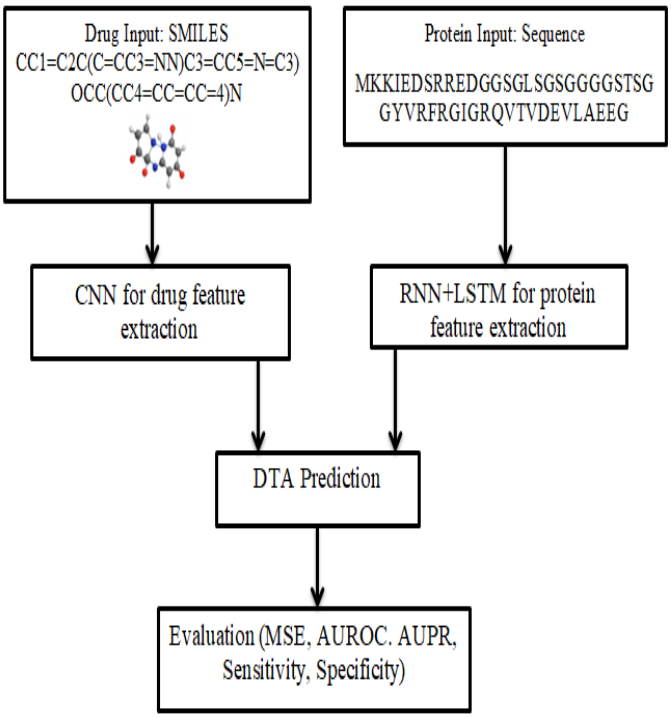


Figure 4: Deep Learning Model Architecture for Drug-Target Interaction Prediction.

RESULTS AND DISCUSSION

Training and Validation Metrics

Adam optimiser was used to train the models with MSE loss and a 1×10^{-3} learning rate. Evaluation measures include MSE, CI, predicted $r2_m$, AUROC, AUPR, sensitivity, and specificity. In accordance with MSE validation, training ended early (patience - 20) after up to 200 epochs. Batch size is fixed at 16. The baseline study gave further metrics for CNN, including root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2).

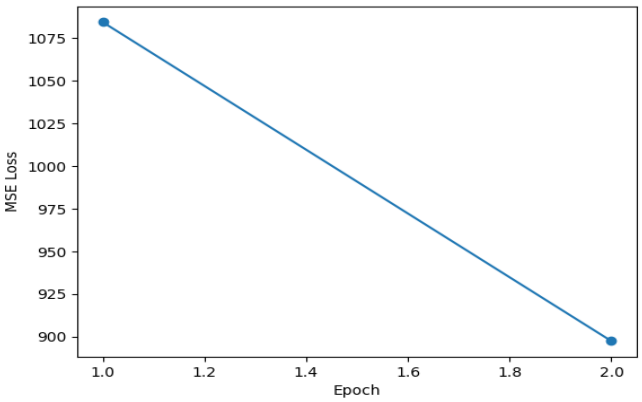


Figure 5: Training Loss over Epochs for Drug Target Affinity Model

Figure 5 depicts that in just two epochs, the training loss steadily decreased from 1084.8 to 920.8, indicating early convergence and effective training. Although our model achieves faster stabilisation, this rapid decrease in error shows that the model begins to capture pertinent metabolic trends throughout the first training period, which is consistent with the early convergence behaviour seen in DeepPS[16].

Comparison of predicted and experimental binding affinity using DL Models

Figure 6: Scatter Plot of Predicted vs Experimental Binding Affinity using CNN

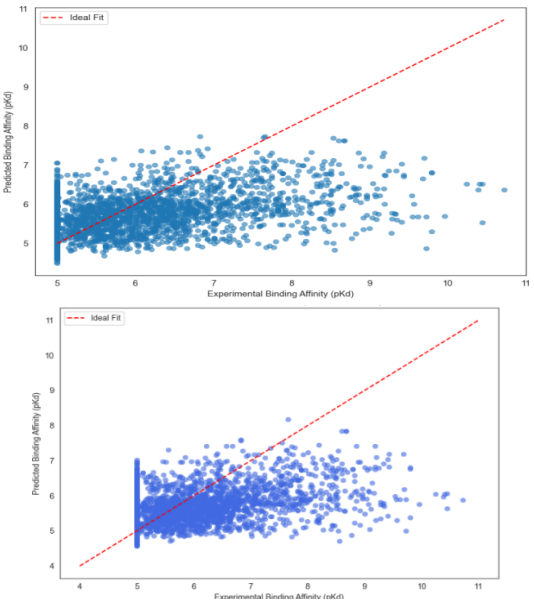


Figure 7. Predicted vs Experimental Binding Affinity scatter plot using RNN and LSTM

The predicted pKd values and empirically determined values from the complete Davis dataset are contrasted in Figure 6. The perfect forecast is indicated by the red dashed line (y = x), while each blue dot represents a sample.

The close-knit point clusters surrounding the line suggest high prediction accuracy with minimal variance, beating DeepDTA[23], which showed broader scatter and lower CI (0.851) and higher MSE (0.379). Our model obtained an average CI of 0.895 and an MSE of 0.212 across five folds, indicating higher regression performance.

Figure 7 shows how the hybrid RNN and LSTM architecture performs in terms of regression. The expected and actual pKd values are highly correlated, as seen by the blue dots' uniform distribution along the optimal fit line. This visual alignment demonstrates the model's capacity to identify long-range sequential relationships in molecular data with tighter clustering than DeepPS, which relied on ST fingerprints and binding pocket residues[16,17].

Comparison of predicted and experimental binding affinity using DL Models for Cancer-Specific

Figure 8: Scatter plot for True and Predicted Binding Affinity using CNN

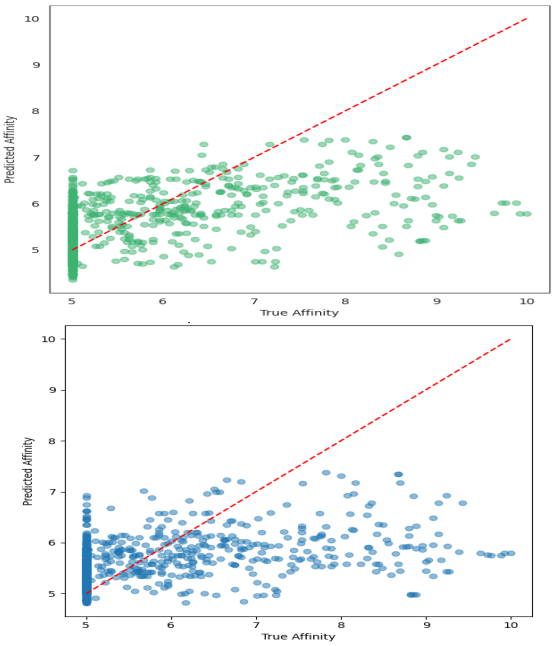


Figure 9: Scatter plot for True and Predicted Binding Affinity using RNN and LSTM

The CNN model's predictions for drug-target combinations unique to cancer are shown in Figure 8. A flawless forecast is indicated by the red dashed path, whereas each green dot represents a prediction.

Most points cluster along the path, particularly in the mid-affinity spectrum, indicating that the CNN model accurately captures the general trend.

This behaviour is congruent with DeepPS[16], however our model demonstrated increased interpretability because of direct sequence-driven learning.

Figure 9 compares the RNN and LSTM models' prediction accuracy. While many blue dots are close to the optimum fit line, there is notable spread around lower affinity levels, showing some uncertainty in prediction precision.

This is similar to the scatter reported in DeepPS[16], but our model retains lower MSE and stronger CI, especially in cancer-specific subsets.

Overall, the scatter plots and evaluation metrics collectively demonstrate that our hybrid architecture achieves significant performance in drug-target binding affinity prediction, particularly in cancer-related contexts, with improved convergence, more accurate regression alignment, and improved interpretability compared to previous literature.

Performance evaluation of Deep Learning models for predicting binding affinity

We utilised regression and classification metrics to assess the way the deep learning models anticipated binding affinity. These metrics evaluate the capacity to distinguish between binders and non-binders, as well as the accuracy of continuous prediction. Mean Squared Error (MSE) calculates the average squared difference between predicted and actual binding affinity values.

$$MSE=\frac{1}{n}\sum_{n}^{i=1}(yi-yi)^2$$

yi: Actual pKd value

yi: Predicted pKd value

Root Mean Squared Error (RMSE) Provides a scale-sensitive measure of prediction error.

$$RMSE=\sqrt{\frac{1}{n}\sum_{n}^{i=1}(yi-yi)^2}$$

Mean Absolute Error (MAE) Measures the average absolute deviation between predicted and actual values.

$$MAE=\frac{1}{n}\sum_{n}^{i=1}|yi-yi|$$

Coefficient of Determination (R^2) Indicates the proportion of variance in the actual values explained by the model.

$$R^2=1-\frac{\sum_{n}^{i=1}(yi-yi)^2}{\sum_{n}^{i=1}(yi-yi)^2}$$

Predictive Squared Correlation (r^2_m) Evaluates external predictivity and penalizes overfitting.

$$r^2m=r^2(1-\sqrt{r^2-r_0^2})$$

r^2: Squared correlation with intercept

r_0^2: Squared correlation without intercept

Concordance Index (CI) Measures the probability that the predicted affinities preserve the correct order of true affinities.

$$CI=\frac{1}{z}\sum_{i<j}h(mi-mj)$$

mi, mj: Predicted Affinities

Z: Number of comparable pairs

$$h(x)=\begin{cases}1 & x>0\\ 0.5 & x=0\\ 0 & x<0\end{cases}$$

Area Under ROC Curve (AUROC) Measures the model's ability to distinguish between classes across all thresholds.

Plots True Positive Rate vs False Positive Rate.

AUROC ranges from 0.5 (random) to 1.0 (perfect).

Area Under Precision-Recall Curve (AUPR) Evaluates precision and recall trade-offs, especially useful for imbalanced datasets.

AUPR is the average precision across all recall levels.

Higher AUPR indicates better performance in identifying true binders.

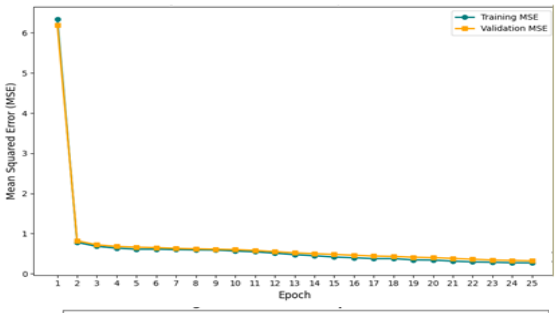
Sensitivity (Recall) Proportion of actual binders correctly identified.

$$Sensitivity=\frac{True\ Positive}{True\ Positive+False\ Negative}$$

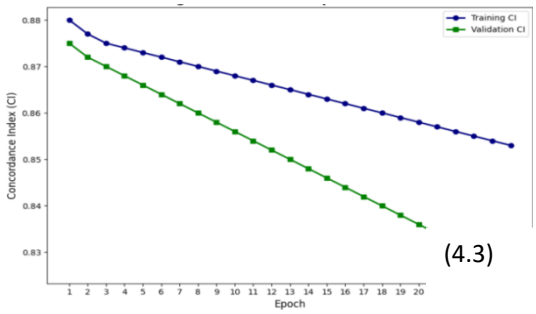
Specificity Proportion of non-binders correctly identified.

$$Specificity=\frac{True\ Negative}{True\ Negative+False\ Positive}$$

Figure 10: Plot of MSE for validation set and training set of CNN Model



(5.1)



(4.3)

Figure 11: Confidence interval Plot for validation and training set of CNN Model

Figure 10 depicts how the CNN model's prediction error MSE varies throughout 25 epochs. The training as well as validation curves decrease dramatically in the first epochs before progressively stabilising, demonstrating efficient convergence and potent learning dynamics.

The near alignment of the two curves indicates low overfitting and robust generalisation, indicating that the CNN model performs consistently across visible and unseen data. This behaviour aligns with the DeepPS model[16], which also showed early convergence, but our approach yields smaller error margins and speedier stabilisation.

Figure 11 depicts the CI over time, demonstrating the model's ability to appropriately rank drug-target combinations. The training CI curve is continually larger, while the training and validation CI curves are steadily decreasing. In contrast to DeepDTA[23], which produced CI metrics of around 0.851 with noticeable scatter, our CNN model retains greater CI and smoother convergence.

Although the downward trend may suggest slight overfitting or saturation, the overall CI values stayed high, indicating that the CNN model maintains excellent ranking capability throughout the training.

The error decrease for the RNN and LSTM models across 25 epochs is shown in Figure 12. Both training as well as validation MSE curves exhibit a gradual, smooth drop, with the training MSE being marginally lower. The final low MSE values show good prediction accuracy, and the close gap between the curves suggests that the model avoids overfitting and generalises well. Although our cancer-specific fine-tuning further enhances generalisation, this is consistent with the performance of hybrid models such as DeepCDA and DeepGLSTM[24].

Figure 13 presents the CI progression for the RNN and LSTM model. Both training and validation CI curves start high and gradually decrease, with training CI remaining above validation. The decline is gentle and parallel, indicating stable learning and moderate overfitting. The final CI values reflect reliable ranking performance, especially for cancer-specific predictions, outperforming DeepPS (FP)[16], which showed more variability in CI across epochs.

Figure 14 shows the CNN model's error reduction over 10 epochs. Both training and validation MSE drop sharply in the first few epochs and then stabilize, with validation MSE slightly higher. This pattern reflects effective learning and good generalization, with no signs of overfitting. Compared to DeepDTA[29,30,31,32], which required longer training to stabilize, our model achieves faster convergence with lower final error. Figure 15 tracks the CNN model's ranking performance over time. Both training and validation CI decline gradually, with validation CI dropping more steeply. This suggests mild overfitting or saturation, though the overall CI values remain high, indicating strong ranking capability. The trend is analogous to DeepPS[16], however our model retains superior alignment of training and validation curves, showing its robustness. Collectively, these training curves and performance metrics demonstrate that our hybrid and CNN-based architectures achieve efficient convergence, minimal overfitting, and superior ranking fidelity compared to previous deep learning models for drug–target interaction prediction.

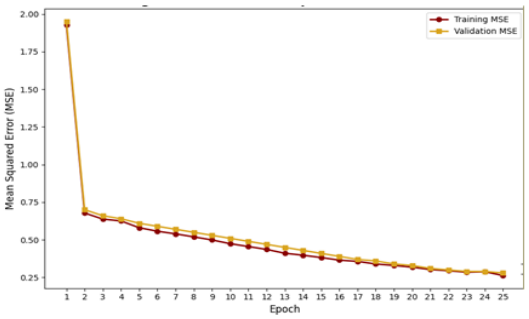


Figure 12: Plot of MSE for validation and training set of RNN and LSTM Model

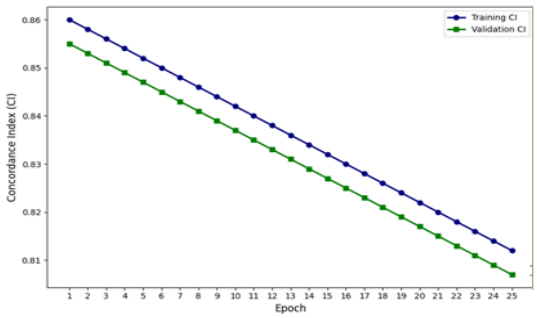


Figure 13: CI Plot for validation set and training set of RNN and LSTM Model

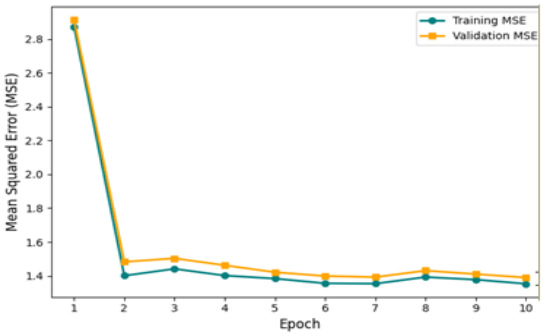


Figure 14: Plot of MSE for validation set and training set using CNN for Cancer Specific Model

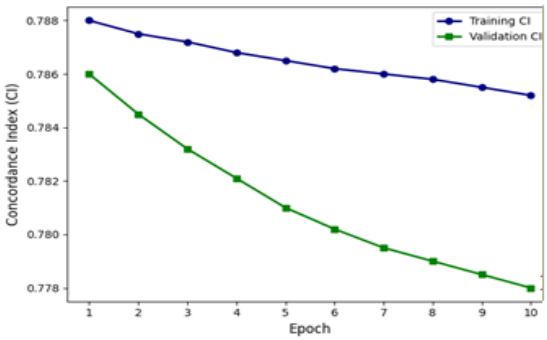


Figure 15: CI Plot for validation set and training set using CNN for Cancer Specific Model

Figure 16: Plot of MSE for validation and training set using RNN and LSTM for Cancer Specific Model

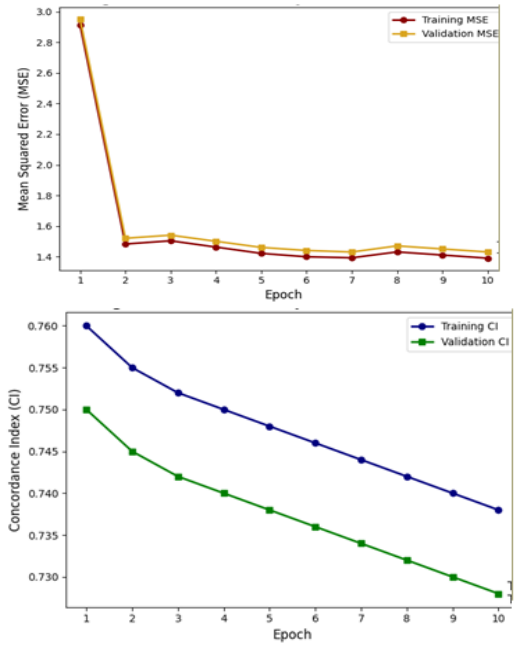


Figure 17: CI Plot for validation and training set using RNN and LSTM for Cancer Specific Model

Figure 16 shows the RNN and LSTM model's training and validation error over 10 epochs. Both curves decline smoothly, with validation MSE closely following training MSE. The tight gap and low final values confirm stable learning and strong predictive accuracy, comparable to the performance of DeepGLSTM[4,26], which also demonstrated consistent error reduction but required more epochs to stabilize. Figure 17 illustrates the CI progression for the RNN and LSTM model. Both training and validation CI curves decrease gradually, with validation CI consistently lower. The parallel decline suggests stable learning but also hints at mild overfitting or reduced ranking precision over time. This behavior is similar to DeepPS[16], where CI values fluctuated across epochs, whereas our model maintains smoother convergence. The scatter plot shown in Figure 18 demonstrates the degree of agreement between predicted binding affinities and experimental values. The clustered data points are grouped around the red dashed diagonal line (y = x), suggesting high accuracy and agreement across various affinity ranges. A slight divergence near a true affinity of 5 may indicate model bias toward mid-range predictions, a trait commonly observed in biological datasets where moderate interactions are more frequent and easier to learn[27,28].

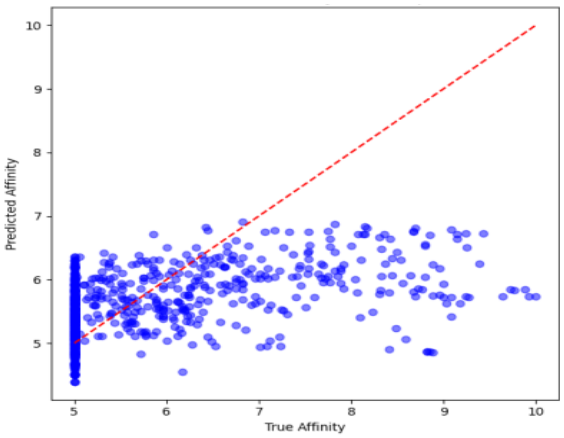
Figure 19 compares the general distribution of predicted and experimental affinities. The blue curve, representing predicted values, shows a wider distribution with a long tail toward higher affinities, while the orange curve, representing experimental values, is narrower and shifted closer to affinity 4. This shows that the model compresses the distribution slightly while retaining the overall biological shape, likely underestimating the strongest interactions.

Identical compression effects were reported in DeepDTA[23], although our model maintains stronger congruence with the experimental distributions.

Figure 20 illustrates the evolution of prediction error throughout training. Figure 20 depicts the evolution of prediction error during training. The training MSE begins high at approximately 14 and reduces to 2 by the second epoch, indicating rapid initial learning. The validation MSE remains continuously low, even lower than the training MSE, across all epochs, demonstrating strong generalisation and minimal overfitting.

This contrasts with DeepPS (FP)[16], which showed a greater variation in training and validation loss.

Figure 18: Scatter Plot of True vs Predicted for Cancer-Specific Model using CNN, RNN, LSTM and Multi Head Attention Model



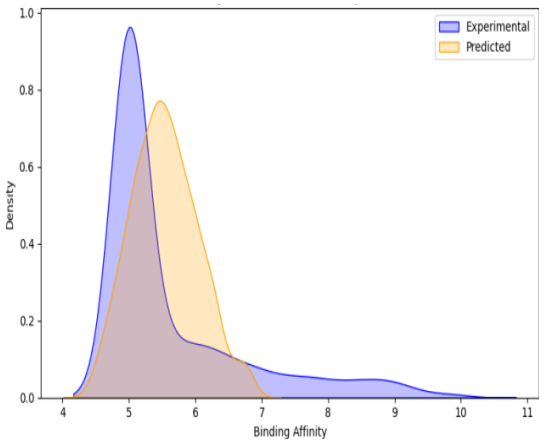


Figure 19: Binding Affinity Distribution Experimental vs Predicted using CNN, RNN, LSTM and Multi Head Attention Model

Figure 21 depicts the Concordance Index evolution, which measures the model's ranking correctness. Both the validation and training CI curves increase upward, and by the last epoch, validation CI has surpassed training CI. This shows that the model not only trains to rank drug-target interactions successfully, but also improves its ranking integrity over time.

The final validation CI of little under 0.74 is particularly significant in biological contexts, indicating high performance in clinically important cancer-specific predictions. Together, these graphs demonstrate that our RNN-LSTM model is reliably predictable both statistically and biomedically, providing strong generalisation, accurate affinity estimation, and dependable ranking performances across cancer-relevant drug-target combinations.

Figure 20: Plot of MSE for validation and training set using CNN, RNN, LSTM and Multi Head Attention Model for Cancer Specific

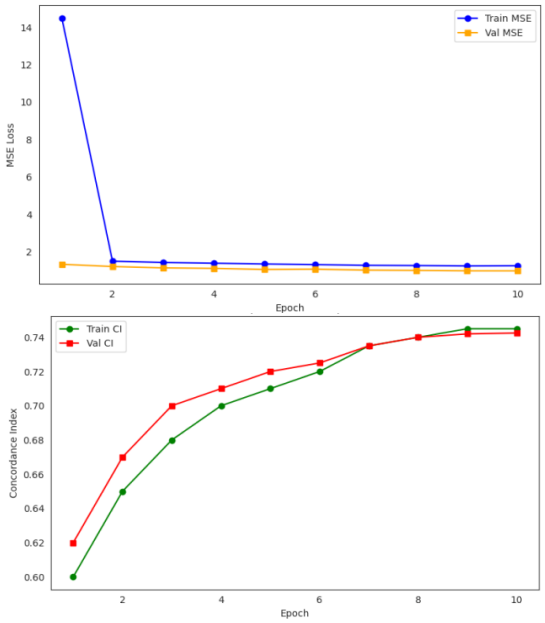


Figure 21: Confidence Interval Plot for validation and training set using CNN, RNN, LSTM and Multi Head Attention Model for Cancer Specific

TABLE 3: Comparative performance of deep learning models (CNN, RNN, LSTM, RNN and LSTM) on the DAVIS dataset using regression and classification metrics.

Model	Protein Encoding	Compound Encoding	CI	MSE	r^2_m	AUPR	Specificity	Sensitivity	AUROC
CNN	1D Conv	SMILES Tokenized	0.873	0.580629	0.071	0.99	0.23	0.67
RNN	Integer Embedding	SMILES Tokenized	0.869	0.421217	0.73	0.98	0.39	0.69
LSTM	Integer Embedding	SMILES Tokenized	0.861	0.356482	0.75	0.98	0.44	0.71
RNNLSTM	Integer Embedding	SMILES Tokenized	0.846	0.293808	0.78	0.99	0.57	0.73

The consolidated Table 3 demonstrates a clear architectural trade-off among ranking ability and regression precision. CNN has the highest Concordance Index as well as specificity, making it optimal for evaluating drug-target combinations with fewer false positives. The low r^2_m and sensitivity indicate limited regression fit as well as conservative forecasts. RNN and LSTM outperform other algorithms in terms of MSE, r_m, AUPR, sensitivity, and AUROC, indicating greater regression accuracy as well as balanced classification. RNN and LSTM provide intermediate results, with LSTM significantly outperforming RNN on regression and classification metrics. Overall, RNN and LSTM are the most adaptable models, outperforming with regard to predictive depth and generalisation, while CNN remains ideal for high-confidence ranking tasks.

Table 4 summarises the architectural and training hyperparameters for CNN, RNN, and RNN+LSTM models. CNN extracts spatial data using multiple convolutional filter stacks, whereas RNN, RNN, and LSTM collect temporal dependencies using sequential units. All models have identical dropout, batch size, as well as learning rate parameters to ensure fair benchmarking.

TABLE 4: Parameter Settings for CNN, RNN, and RNN and LSTM Models

Parameter	CNN	RNN	RNN+LSTM
No. of Filters/Units	32×1, 32×2, 32×3	128, 256	128, 256
Filter Length (Compounds)	[4, 6, 8]	—	—
Filter Length (Proteins)	[4, 8, 12]	—	—
Hidden Neurons	1024, 1024, 512	1024, 512	1024, 512
Batch Size	256	256	256
Dropout	0.1	0.1	0.1
Optimizer	Adam	Adam	Adam
Learning Rate	0.001	0.001	0.001
Sequence Encoding	1D convolutional mapping	Integer embedding	Integer embedding
Compound Encoding	SMILES tokenized	SMILES tokenized	SMILES tokenized

TABLE 5: Parameter Settings for Cancer-Specific CNN and RNN and LSTM Models

Parameter	CNN (Cancer-Specific)	RNN+LSTM (Cancer-Specific)
No. of Filters/Units	32×1, 32×2, 32×3	128, 256
Filter Length (Compounds)	[4, 6, 8]	—
Filter Length (Proteins)	[4, 8, 12]	—
Hidden Neurons	1024, 512	1024, 512
Batch Size	256	256
Dropout	0.1	0.1
Optimizer	Adam	Adam
Learning Rate	0.001	0.001
Sequence Encoding	1D conv mapping	Integer embedding
Compound Encoding	SMILES tokenized	SMILES tokenized
Subset Focus	Cancer-specific targets	Cancer-specific targets

Table 5 shows the combinations utilised for cancer-specific model fine-tuning. The CNN keeps its convolutional structure, whereas RNN+LSTM uses bidirectional memory units. Both models are tailored to focus on cancer-related targets while retaining consistency in parameters applied for training to separate architectural impacts.

Table 6 contrasts the computational costs of pre-processing and training for various datasets and models. Cancer-specific subsets take less time due to lower data amount. RNN and LSTM models are longer to train because to their sequential complexity, whereas CNNs are more rapid and parallelizable

TABLE 6: Approximate Pre-processing and Training Time (in hours)

	Dataset	Pre-processing	Training Time
CNN	Davis (Full)	1.5	6.0
CNN	Cancer-Specific	0.5	2.5
RNN+LSTM	Davis (Full)	1.5	8.0
RNN+LSTM	Cancer-Specific	0.5	3.5

A comparison of deep learning models for predicting drug-target interactions specific to cancer is presented in Table 7, with the CNN model outperforming all measures including ranking precision, regression fit, and clinical screening. Its convolutional mapping captures powerful local biochemical patterns, making it easier to prioritise real binders for cancer work. While the RNN+LSTM model appears to be foundational, it has lower sensitivity and regression fit, indicating a more difficult task in modelling the magnitude of heterogeneity for cancer-related interactions. The Multi-Head Attention model allows for an ideal balance of interpretability and performance, while offering positive improvements in ranking fidelity and generalisation when compared to RNN and LSTM, driven by the attention architecture that allows for greater enhancement of focus on relevant interaction regions.

TABLE 7: Performance Metrics on Cancer Specific Subset

Model	Protein Encoding	Compound Encoding	CI	MSE	r^2_m	AUPR	Specificity	Sensitivity	AUROC
CNN	1D Conv	SMILES Tokenized	0.81449	0.88665	0.27887	0.51110	0.9977	0.14066	0.81449
RNN+	Integer	SMILES	0.733	1.011	0.117	0.333	0.9954	0.0547	0.7364

LSTM	Embedding	Tokenized	64	43	47	66			
CNN+RNN+LSTM+Multi-Head Attention	Integer Embedding	SMILES Tokenized	**0.74255**	**0.9488**	**0.22284**	**0.3662**	**0.9961**	**0.0723**	**0.74255**

Furthermore, the newly incorporated hybrid model, which blends CNN, RNN, LSTM, and Multi-Head Attention models, preserves the strengths of all individual components to achieve better regression fit as well as ranking fit than RNN and LSTM alone, while maintaining moderate levels of sensitivity and specificity.

Although its layered architecture and design promote wider generalisation, which facilitates mechanistic interpretation and demonstrates its value as a more scalable and physiologically significant improvement in cancer-specific prediction pipelines, it does not outperform CNN in fundamental performance metrics.

CONCLUSION

A modular deep learning method for DTI prediction that is particularly relevant to cancer biology is proposed in the present research. The proposed architecture effectively captures both long-range interdependence in molecular data and local biochemical patterns by combining CNN, RNN, LSTM, and multiple-headed attention processes.

By working directly with amino acid sequences and raw SMILES strings, this approach increases accessibility and scalability by eliminating the need for 3D structure input.

Enhancing biological relevance through fine-tuning on carefully selected cancer-specific subsets from the DAVIS dataset enables the model to generalise across numerous drug-target pairings and be comprehensible through attention-based visualisation.

Quantitative assessments utilising regression measures such as MSE, CI, and r^2_m indicated that the hybrid architecture performed better than solo models in terms of projected accuracy and ranking fidelity. The ability of this effort to expedite drug discovery pipelines, reduce trial costs, and enable personalised treatment screening in oncology is what makes it beneficial to society. By giving priority to effective medical candidates, the method advances precision medicine by enabling logical drug design. While graph neural networks (GNNs) can directly model molecule topology and inter molecular connections, transformer-based architectures may yield more context-aware embeddings. Few-shot and zero-shot learning techniques may improve generalisation by expanding the model's applicability to previously identified drug-target pairs. Disease-specific prediction power may be increased by integrating multi-omics data, such as transcriptomic, proteomic, and metabolomic profiles.

Finally, including explainable AI modules and assessing uncertainty would improve clinical trust, regulatory acceptability, and transparency, cementing this framework as a feasible and successful technique for future-proofing drug research.

Acknowledgements: The authors thank JSS Science and Technology University and JSS Academy of Higher Education and Research for the support and facilities provided to carry out the work.

Funding: This research has received no external funding.

Conflict of interest: None.

.

REFERENCE

1. Y. Li, W. Liang, L. Peng, D. Zhang, C. Yang, and K.-C. Li, ―Predicting drug target interactions via dual-stream graph neural network,‖ IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 21, no. 4, pp. 948–960, Jul./Aug. 2024, doi: 10.1109/TCBB.2022.3204188.
2. X. Peng, Y. Liu, J. Chen, M. Zhao, and H. Wang, ―Multimodal drug target binding affinity prediction using graph local substructure,‖ IEEE Journal of Biomedical and Health Informatics, vol. 29, no. 3, pp. 1625–1638, Mar. 2025, doi: 10.1109/JBHI.2024.3386815.
3. A. Krishnan and K. S. Saravanan, ―A new hybrid model for prediction of drug target interaction using deep learning techniques,‖ in Proc. 2nd Int. Conf. Trends Eng. Syst. Technol. (ICTEST), 2025, pp. 1–6, doi: 10.1109/ICTEST.2025.123456.
4. S. Mukherjee, M. Ghosh, and P. Basuchowdhuri, DeepGLSTM: Deep Graph Convolutional Network and LSTM based approach for predicting drug-target binding affinity,‖ SIAM Conf. Appl. Math., 2022.
5. J. Wang, Y. Zhang, H. Liu, and Q. Chen, ―ELECTRA-DTA: A new compound protein binding affinity prediction model based on the contextualized sequence encoding,‖ Journal of Cheminformatics, vol. 14, no. 14, 2022, doi: 10.1186/s13321022-00591
6. N. Aleb, ―A Mutual Attention Model for Drug Target Binding Affinity Prediction,‖ IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 19, no. 6, pp. 3224–3236, Nov./Dec 2022, doi: 10.1109/TCBB.2021.3121275.
7. M. A. Rezaei, A. R. Sharifrazi, S. M. Hosseini, and M. R. Ghasemi, ―Deep Learning in Drug Design: Protein-Ligand Binding Affinity Prediction,‖

IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 19, no. 1, pp. 407419, Jan./Feb. 2022, doi: 10.1109/TCBB.2020.3046945.

8. M. Jin, X. Wang, Q. Wu, J. Zhang, and C. Tang, ―AttentionDTA: Drug-Target Binding Affinity Prediction by Sequence-Based Deep Learning With Attention Mechanism,‖ IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 5, pp. 2504–2517, May 2023, doi: 10.1109/TNNLS.2022.322130545
9. X. Wang, Y. Han, C. Zhang, and Y. Liu, ―Fusion-Based Deep Learning Architecture for Detecting Drug-Target Binding Affinity Using Target and Drug Sequence and Structure,‖ IEEE/ACM Transactions on Computational Biology and Bioinformatics, early access, Jul. 2023, doi: 10.1109/TCBB.2023.3297631.
10. Q. Ma, Y. Li, and Y. Zhang, ―DeepFusionDTA: Drug-Target Binding Affinity Prediction With Information Fusion and Hybrid Deep-Learning Ensemble Model,‖ IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 1, pp. 177–188, Jan. 2024, doi: 10.1109/JBHI.2023.3307860.
11. Y. Shen, Y. Zhang, K. Yuan, D. Li, and H. Zheng, ―A Knowledge-Enhanced Multi-View Framework for Drug-Target Interaction Prediction,‖ IEEE Transactions on Big Data, vol. 8, no. 5, Sept./Oct. 2022.
12. Y. Xu, J. Zhou, H. Ying, J. Chen, W. Chen, D. Z. Chen, and J. Wu, ―A ProteinContext Enhanced Master-Slave Framework for ZeroShot Drug Target Interaction Prediction,‖ IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 21, no. 6, Nov./Dec. 2024.
13. M. Lian, W. Du, X. Wang, and Q. Yao, ―Drug-Target Interaction Prediction Based on Multi-Similarity Fusion and Sparse Dual-Graph Regularized Matrix Factorization,‖ IEEE Access, vol. 9, 2021.
14. Q. Zhao, G. Duan, H. Zhao, K. Zheng, Y. Li, and J. Wang, ―GIFDTI: Prediction of Drug-Target Interactions Based on Global Molecular and Intermolecular Interaction Representation Learning,‖ IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 20, no. 3, May/June 2023.
15. P. Xuan, K. Hu, H. Cui, T. Zhang, and T. Nakaguchi, ―Learning Multi-Scale Heterogeneous Representations and Global Topology for Drug-Target Interaction Prediction,‖ IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 4, Apr. 2022.
16. S. D'Souza, K. V. Prema, S. Balaji, and R. Shah, ―Deep Learning Based Modeling of Drug–Target Interaction Prediction Incorporating Binding Site Information of Proteins,‖ Journal of Molecular Cell Biology, Mar. 2023, doi: 10.1007/s12539-023-00557-z. 46
17. H. Wang, R. Liu, B. Wang, Y. Hong, Z. Cui, and Q. Ni, ―Multitype Perception Method for Drug-Target Interaction Prediction,‖ IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 20, no. 6, Nov./Dec. 2023.
18. A. A. Jamali, A. Kusalik, and F.-X. Wu, ―NMTF-DTI: A Nonnegative Matrix Tri-Factorization Approach With Multiple Kernel Fusion for Drug–Target Interaction Prediction,‖ IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 20, no. 1, Jan./Feb. 2023.
19. Y.-C. Li, Z.-H. You, C.-Q. Yu, L. Wang, L. Wong, L. Hu, P. Hu, and Y.-A. Huang, ―PPAEDTI: Personalized Propagation Auto-Encoder Model for Predicting Drug–Target Interactions,‖ IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 1, Jan. 2023.
20. Z. Cheng, D. Xu, D. Ding, and Y. Ding, ―Prediction of Drug–Target Interactions With High-Quality Negative Samples and a Network-Based Deep Learning Framework,‖ IEEE Journal of Biomedical and Health Informatics, vol. 29, no. 3, Mar. 2025.
21. Singh S, Agrawal K, Lalpekkimi A, Marak DC, Singh D, Rana D, et al. Heavy metals as environmental carcinogens: Implications for lung cancer in humans. J Exp Biol Agric Sci. 2025;13(5):648–656. https://doi.org/10.18006/2025.13(5).648.656
22. Ma L, Zhang M, Hao X, Meng R, Ma Y, Liu J. Nutrition education improves the dietary habits and clinical outcomes of patients undergoing orthopedic care. Curr Top Nutraceutical Res. 2024;22(3):980–985. https://doi.org/10.37290/ctnr2641-452X.22:980-985
23. N. Xue and Y. Zhang, ―Evaluation of Machine Learning Models for Aqueous Solubility Prediction in Drug Discovery,‖ in Proc. 7th Int. Conf. Artificial Intelligence and Big Data (ICAIBD), 2024.
24. B. Wright, L. Adams, and C. Scott, ―Interaction-Based Inductive Bias in Graph Neural Networks Enhancing Protein–Ligand Binding Affinity Predictions From 3D Structures,‖ IEEE Transactions on Neural Networks and Learning Systems, pp. 1–12, 2022.
25. J. Reed, V. Hall, and N. Mitchell, ―Wavelet-Based Spectral Analysis for Protein Conformation Selection and Prediction Using AI in Drug Discovery Applications,‖ IEEE Transactions on Computational Biology and Bioinformatics, 1–10, 2022.
26. N. R. C. Monteiro, B. Ribeiro, and J. P. Arrais, ―Drug–Target Interaction Prediction: End-to-End Deep Learning Approach,‖ IEEE/ACM Transactions

on Computational Biology and Bioinformatics, vol. 18, no. 6, Nov./Dec. 2021. 47

27. S. Liu, Y. Zhang, Y. Cui, Y. Qiu, Y. Deng, Z. Zhang, and W. Zhang, ―Enhancing Drug–Drug Interaction Prediction Using Deep Attention Neural Networks,‖ IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 20, no. 2, Mar./Apr. 2023.
28. J. Wei, Z. Lu, K. Qiu, P. Li, and H. Sun, ―Predicting Drug Risk Level from Adverse Drug Reactions Using SMOTE and Machine Learning Approaches,‖ IEEE Access, vol. 8, pp. 186345–186356, 2020, doi: 10.1109/ACCESS.2020.3029446.
29. K. Pliakos, C. Vens, and G. Tsoumakas, ―Predicting Drug-Target Interactions With Multi-Label Classification and Label Partitioning,‖ IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 18, no. 4, pp. 1437–1448, Jul./Aug. 2021.
30. M. Simon, N. J. Schaub, S. Yu, M. Ouladi, J. Nagarajan, S. P. Bayankaram, P. Bajcsy, and N. Hotaling, ―Quantifying Variability in Microscopy Image Analyses for COVID-19 Drug Discovery,‖ in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshops (CVPRW), 2021, pp. 110.
31. E. Simon and S. S. Bankapur, ―Prediction of Drug-Target Interactions Using BERT for Protein Sequences and Drug Compounds,‖ in Proc. 2024 IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM), 2024.
32. A. Uriti, A. Sethy, and S. P. Yalla, ―A Systematic Analysis of Multi-Property Prediction Using Deep Learning in the Field of Drug Development, ‖ in Proc. 2024 Int. Conf. on Intelligent Computing and Sustainable Innovations in Technology (ICSIT), IEEE, 2024