CTU Journal of Innovation and Sustainable Development
ISSN 2588-1418 | e-ISSN 2815-6412

DOI:10.22144/ctujoisd.2026.050

3D-ReSE-VN: Advanced Deep Learning for Protein Ligand Affinity Prediction in Cancer-Associated Targets

Muhammad Hussain[1], Maisam Abbas[1], Qaisar Ali[1], Muhammad Hassan[3], Hussain Abbas[4], and Quang Thai Ho[1,2*]

[1]*Department of Computer Science and Engineering, Yuan Ze University, Taiwan*

[2]*College of Information and Communication Technology, Can Tho University, Viet Nam*

[3]*Department of Computer Science and Engineering, University of Udine, Italy*

[4]*Graduate Institute of Artificial Intelligence and Big Data in Healthcare, National Taipei University of Nursing and Health Sciences, Taiwan*

**Corresponding author (hqthai@saturn.yzu.edu.tw)*

Article info.

Received 20 Jun 2026
Revised 8 Sep 2026
Accepted 4 Oct 2026

Keywords

Anticancer drug discovery, attention, molecular docking, protein-ligand binding affinity, residual learning

ABSTRACT

Accurate protein-ligand binding affinity prediction is important for structure-based anticancer drug discovery. This study presents 3D-ReSE-VN, a three-dimensional residual attention network that combines voxelized docking-derived complexes with residual learning, squeeze-and-excitation channel attention, and spatial attention. The model was evaluated on 4,620 protein-ligand complexes using random five-fold cross-validation and a predefined train/validation/test split. In cross-validation, 3D-ReSE-VN achieved an RMSE of 1.9679 ± 0.8242, an MAE of 0.6952 ± 0.1079, a Pearson correlation of 0.8169 ± 0.1095, and an R^2 of 0.6724 ± 0.1755. On the fixed hold-out test set of 693 complexes, the model achieved an RMSE of 1.3251, an MAE of 0.6420, a Pearson correlation of 0.9098, and an R^2 of 0.8271. The model showed competitive performance against the reported architectures, although performance varied across data partitions and affinity ranges. These findings support further evaluation on larger, chemically diverse, and disjoint protein-ligand datasets.

1. INTRODUCTION

Cancer remains a major global health challenge with substantial clinical, economic, and societal consequences (Adeshakin et al., 2021; Smith et al., 2022; Zhang et al., 2024). Its development involves molecular alterations that affect signaling, proliferation, apoptosis, metabolism, immune regulation, and the tumor microenvironment (Adeshakin et al., 2021; Wang et al., 2024a; Zhang et al., 2024). Because many of these processes involve dysregulated proteins, identifying small molecules that effectively bind disease-associated targets is central to targeted drug discovery and precision therapeutics (Kim et al., 2024; Wang et al., 2024a). Accurate protein-ligand binding affinity prediction supports virtual screening, compound prioritization, and lead optimization before experimental validation (Rezaei et al., 2022; Abdelkader & Kim, 2024; Liu et al., 2024b). Molecular docking is widely used to generate plausible binding poses and estimate protein-ligand interactions, with tools such as AutoDock Vina enabling efficient exploration of ligand conformations (Eberhardt et al., 2021; Mursal et al., 2024). However, conventional scoring functions use simplified energy approximations and may

inadequately represent protein flexibility, solvent and entropic effects, and complex intermolecular interactions (Liu et al., 2024c; Mursal et al., 2024). These limitations have encouraged artificial intelligence approaches that learn interaction patterns directly from molecular and structural data (Rezaei et al., 2022; Jiang et al., 2023; Abdelkader & Kim, 2024). Deep learning methods employ diverse representations for affinity prediction. GraphDTA uses molecular graphs and graph neural networks (Nguyen et al., 2021), whereas MolTrans learns interactions between molecular substructures and protein sequences (Huang et al., 2021). Graph-based interaction modeling and active-site sequence representations have further demonstrated the value of interaction-aware features (Jiang et al., 2021; Born et al., 2022), while recent sequence- and attention-based methods have extended these representations (Wang et al., 2022; Liu et al., 2024d). Nevertheless, sequence and two-dimensional graph representations do not explicitly encode complete three-dimensional binding geometry. Structure-based deep learning addresses this limitation by incorporating the spatial organization of protein-ligand complexes. Spatial attention, structural deep learning, and structure-aware graph models have demonstrated the value of learning binding geometry and atomic interactions (Wang et al., 2021; Wang et al., 2022; Guo, 2024). HAC-Net further combines three-dimensional convolution, channel-wise attention, and graph convolution, illustrating the complementary value of voxel, attention, and graph representations (Kyro et al., 2023). More recent approaches have explored electron-density representations, structure-based graph ensembles, interaction-aware models, and learned protein-ligand scoring functions (Guo, 2024; Isert et al., 2024; Lai et al., 2024; Liang et al., 2024a). These developments reflect a broader shift toward geometry-aware and interaction-aware affinity prediction. Model reliability and generalizability remain important challenges. Performance can depend strongly on dataset construction, molecular representation, preprocessing, and evaluation strategy (Abdelkader & Kim, 2024; Wang et al., 2024b; Zhou & DiMaio, 2025). Chemical or structural similarity between training and evaluation samples can make random partitions less stringent than evaluations involving unseen ligands or targets (Kyro et al., 2023; Mastropietro et al., 2023). Uncertainty-aware approaches have therefore been investigated to identify potentially unreliable affinity predictions (Rayka & Naghavi, 2025). Recent developments in 2025 and 2026 further emphasize structure-aware learning. Integrated structure prediction, docking, and affinity frameworks have connected multiple stages of molecular recognition (Wu et al., 2025), while evaluations of deep learning docking methods continue to identify challenges involving novel structural contexts (Morehead et al., 2026). Vision transformers have been applied to three-dimensional affinity prediction to capture long-range spatial dependencies (Poziemski & Siedlecki, 2026). E-CloudBind combines electron-density point clouds with molecular graphs to address structural uncertainty and out-of-distribution structures (Liu et al., 2026), whereas PBCNet2.0 uses atomic-level geometric representations to learn protein-ligand recognition and relative affinity relationships (Yu et al., 2026). Despite these advances, three-dimensional affinity prediction remains sensitive to structural quality, representation choice, dataset composition, and chemical diversity, while voxel representations can impose substantial computational requirements (Abdelkader & Kim, 2024; Isert et al., 2024; Wang et al., 2024b; Rayka & Naghavi, 2025; Liu et al., 2026). These limitations motivate efficient architectures that capture local physicochemical interactions and broader spatial dependencies while emphasizing informative feature channels and interaction-rich regions.

Accordingly, this study proposes 3D-ReSE-VN, a three-dimensional residual attention network for protein-ligand binding affinity prediction in cancer-associated targets. AutoDock Vina (Eberhardt et al., 2021) generates binding poses that are transformed into voxelized three-dimensional representations. 3D-ReSE-VN combines residual learning, squeeze-and-excitation channel attention, and spatial attention and is evaluated using a fixed hold-out test set and separate random five-fold cross-validation. Given the restricted chemical diversity of the dataset, performance is interpreted within the represented chemical space rather than as evidence of unrestricted generalization to unseen chemotypes.

2. MATERIALS AND METHODS

2.1. Data Collection

The dataset comprised 21 cancer-associated protein targets obtained from the Protein Data Bank (PDB) (Burley et al., 2025): BCR-ABL1, BRCA1, CDK2, EGFR, HER2, MET, PARP1, ROS1, AKT1, androgen receptor, BRAF, CDK4, CDK6, DNMT1, EZH2, KRAS, mTOR, PIK3CA, STAT3, TP53, and

VEGFR2. Structurally suitable entries with biologically relevant binding conformations were prioritized. Fourteen anticancer ligands were obtained from ChEMBL (Zdrazil et al., 2024) and were standardized, protonated, optimized, and prepared for docking. After affinity-label matching and preprocessing, 4,620 complexes were partitioned into 3,234 training, 693 validation, and 693 test samples, corresponding to a 70:15:15 split. Protein and ligand identifiers and associated structural information are provided in the supplementary material. Table 1 summarizes the dataset.

Table 1. Composition and partitioning of the protein-ligand dataset

Targets	Ligands	Complexes	Train	Validation	Test
21	14	4,620	3,234	693	693

2.2. Structure-Based Docking Simulations

Protein structures were prepared by removing water molecules, buffer ions, unrelated heteroatoms, and co-crystallized ligands after their coordinates had been recorded for binding-site definition. Polar hydrogen atoms and partial charges were added, and the prepared protein structures were converted to PDBQT format. Ligands were standardized, protonated, geometrically optimized, assigned Gasteiger charges, configured with rotatable bonds, and converted to PDBQT format. AutoDock Vina (Eberhardt et al., 2021) was used to dock each protein-ligand pair. The search grid was centered on the co-crystallized ligand when available; otherwise, it was defined using experimentally reported active-site residues. Multiple binding poses were generated for each protein-ligand pair, and the lowest-energy pose was retained for subsequent analysis. Docking scores and three-dimensional atomic coordinates were recorded, and the selected complexes were converted into fixed-size voxel grids encoding atomic, spatial, and physicochemical information for input to 3D-ReSE-VN.

2.3. Three-Dimensional Residual Squeeze-and-Excitation Voxel Network

The proposed 3D-ReSE-VN predicts protein-ligand binding affinity from voxelized three-dimensional complexes. As shown in Figure *1*, the model combines 3D convolution, residual learning, squeeze-and-excitation (SE) channel attention, spatial attention, global average pooling, and regression.

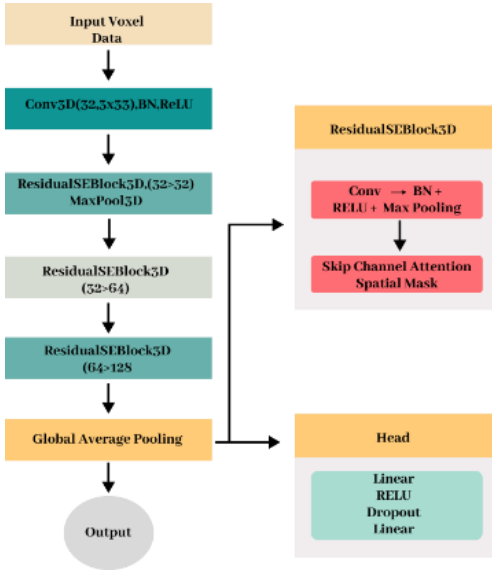


Figure 1. Architecture of 3D-ReSE-VN

(Note: Voxelized protein-ligand complexes are processed by 3D convolutional residual-attention blocks, followed by global average pooling and affinity regression)

2.3.1. Voxelized Input and Feature Extraction

Each complex was represented as a fixed-size tensor

$$\mathbf{X} \in \mathbb{R}^{D\times H\times W\times C_{\text{in}}},$$

where D, H, and W are the voxel-grid dimensions and C_{in} is the number of feature channels encoding structural and physicochemical information. Initial features were extracted using 32 $3\times3\times3$ convolutional filters followed by batch normalization and ReLU activation.

2.3.2. Residual and Channel-Attention Learning

Residual SE blocks progressively extracted features with channel dimensions of 32, 64, and 128, while max-pooling reduced spatial resolution. Residual learning was defined as

$$\mathbf{F}_{l+1} = \mathbf{F}_l + \mathcal{T}_l(\mathbf{F}_l),$$

where $\mathcal{T}_l(\cdot)$ denotes the learned residual transformation. Within each block, global average pooling generated channel descriptors that were transformed through two fully connected layers and a sigmoid function. The resulting SE weights recalibrated the feature channels, emphasizing informative structural and physicochemical features.

2.3.3. Spatial Attention and Affinity Regression

Spatial attention further emphasized informative binding regions. Channel-wise average and maximum descriptors were concatenated and processed by a 3D convolution and sigmoid function:

$$\mathbf{M}_{\text{spatial}} = \sigma(\text{Conv3D}[\text{AvgPool}_c(\mathbf{F}); \text{MaxPool}_c(\mathbf{F})]).$$

The resulting mask reweighed the feature tensor. Global average pooling then produced a compact feature vector, which was passed through a fully connected layer, ReLU activation, dropout, and a final linear layer:

$$\hat{y} = \text{FC}_{\text{out}}\Big(\text{Dropout}\Big(\text{ReLU}\big(\text{FC}(\mathbf{v})\big)\Big)\Big).$$

Together, residual learning, SE channel attention, and spatial attention enable hierarchical learning of structural and physicochemical features for protein-ligand affinity prediction.

2.4. Performance Evaluation Metrics

The predictive performance of 3D-ReSE-VN was evaluated using root mean square error (RMSE), mean absolute error (MAE), the coefficient of determination (R^2), and the Pearson correlation coefficient (r). Let n denote the number of samples, y_i the reference affinity, $\hat{y}_i$ the predicted affinity, $\bar{y}$ the mean reference value, and $\bar{\hat{y}}$ the mean predicted value.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}.$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|.$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}.$$

$$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})\left(\hat{y}_i - \bar{\hat{y}}\right)}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}\left(\hat{y}_i - \bar{\hat{y}}\right)^2}}.$$

Lower RMSE and MAE values indicate smaller prediction errors, while higher R^2 and r values indicate stronger predictive agreement.

3. RESULTS AND DISCUSSION

3.1. Cross-Validated Performance Comparison with Baseline Architectures

To evaluate performance across different random partitions, 3D-ReSE-VN was assessed using five-fold cross-validation on all 4,620 protein-ligand complexes and compared with Simple3DCNN, Re3DNet, and VoxFormer. Table 2 summarizes the mean and standard deviation of RMSE, MAE, Pearson correlation, and R^2 across the five folds.

3D-ReSE-VN achieved the best mean values among the reported architectures, with an RMSE of 1.9679 ± 0.8242, an MAE of 0.6952 ± 0.1079, a Pearson correlation of 0.8169 ± 0.1095, and an R^2 of 0.6724 ± 0.1755. These results indicate improved average predictive performance compared with the selected baselines. However, the relatively large variability in RMSE and R^2 shows that performance depends on the composition of the random folds. Therefore, the results should be interpreted as robustness across random data partitions rather than evidence of ligand- or target-disjoint generalization.

Table 2. Five-fold cross-validation performance (mean ± standard deviation)

Model	RMSE	MAE	Pearson r	R^2
Simple3DCNN	2.4938 ± 0.5749	1.3295 ± 0.1095	0.6983 ± 0.0797	0.4850 ± 0.1139
Re3DNet	2.1684 ± 0.5725	0.8566 ± 0.0356	0.7856 ± 0.0749	0.6109 ± 0.1094
VoxFormer	2.6675 ± 0.5469	1.2632 ± 0.0818	0.6482 ± 0.0818	0.4095 ± 0.1068
3D-ReSE-VN	$\mathbf{1.9679 \pm 0.8242}$	$\mathbf{0.6952 \pm 0.1079}$	$\mathbf{0.8169 \pm 0.1095}$	$\mathbf{0.6724 \pm 0.1755}$

(Notes: Best values are shown in bold)

3.2. Performance on the Fixed Hold-Out Test Set

The architectures were also evaluated on the predefined fixed hold-out test set containing 693 complexes. This test partition was not used for training, early stopping, or checkpoint selection and was evaluated separately from the cross-validation experiment. Table 3 reports the resulting performance.

Table 3. Performance on the fixed hold-out test set ($N = 693$)

Model	RMSE	MAE	Pearson r	R^2
Simple3DCNN	1.3648	0.6601	0.9043	0.8166
Re3DNet	1.3289	**0.5990**	**0.9104**	0.8261
VoxFormer	2.2448	1.3509	0.7105	0.5039
3D-ReSE-VN	**1.3251**	0.6420	0.9098	**0.8271**

(Notes: Best values are shown in bold)

On the fixed hold-out test set, 3D-ReSE-VN achieved an RMSE of 1.3251, an MAE of 0.6420, a Pearson correlation of 0.9098, and an R^2 of 0.8271. The model obtained the lowest RMSE and highest R^2 among the reported architectures, whereas Re3DNet achieved the lowest MAE (0.5990) and a marginally higher Pearson correlation (0.9104). The results therefore demonstrate competitive rather than uniformly superior performance. Cross-validation and fixed hold-out results are reported separately because they represent different evaluation protocols and provide complementary information about predictive accuracy and partition sensitivity.

3.3. Prediction Accuracy and Residual Error Analysis

Prediction behavior on the fixed hold-out test set was examined using reference-versus-predicted affinity values. As shown in Figure *2*, most observations were concentrated near the identity line, consistent with the strong overall test-set performance. However, one sparsely represented high-affinity observation was substantially underpredicted. Such large individual errors have a greater effect on RMSE than on MAE and can therefore influence the overall squared-error estimate.

Signed residuals were calculated as predicted minus reference affinity. Figure *3* shows a statistically significant moderate negative relationship between reference affinity and signed error, with $r = -0.4146$ and $p = 3.6 \times 10^{-30}$. Thus, predictions tended to become more negative relative to the reference values as affinity increased. The moderate magnitude of the correlation, however, indicates that this behavior was not uniform across the entire test set and was partly influenced by sparsely represented high-affinity observations.

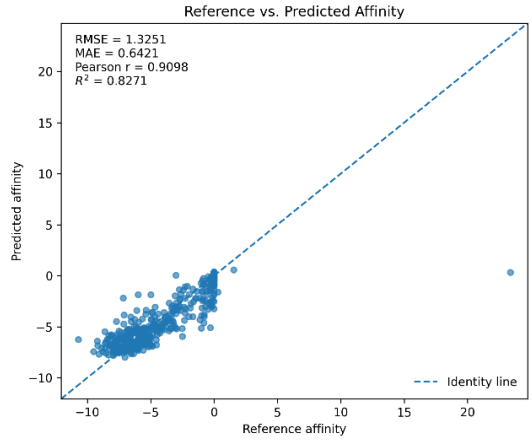


Figure 2. Reference versus predicted affinity values for 3D-ReSE-VN on the fixed hold-out test set

(Notes: The dashed diagonal line represents ideal agreement.)

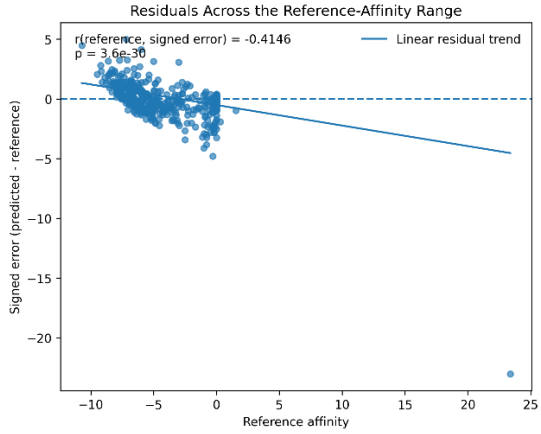


Figure 3. Signed prediction error across the reference-affinity range

(Notes: The dashed horizontal line represents zero error, and the solid line represents the fitted residual trend.)

3.4. Affinity-Range Error Distribution and Prediction Bias

To examine whether prediction accuracy was consistent across the affinity range, the fixed hold-out set was divided into five quantile bins according to reference affinity. Figure *4* shows a non-monotonic error pattern: prediction error was appreciable in the lower-affinity region, decreased across parts of the intermediate range, and increased again for RMSE in the highest-affinity quantile.

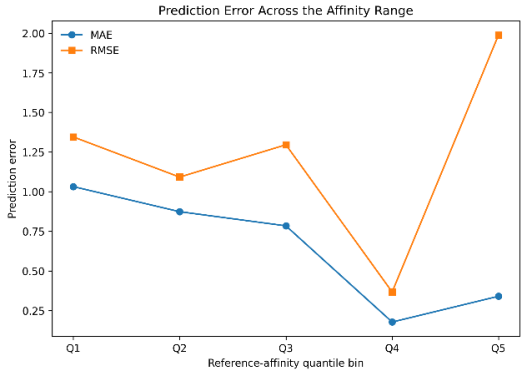


Figure 4. MAE and RMSE across five quantiles of the reference-affinity distribution on the fixed hold-out test set

In the highest-affinity quantile, MAE remained comparatively modest, whereas RMSE increased to approximately 2.0. This difference suggests that a limited number of large prediction errors, rather than uniformly poor predictions, dominated the squared error in this region. The lower-affinity bins also showed noticeable error, indicating that reduced accuracy was not restricted only to high-affinity samples. Overall, prediction reliability varied across the affinity distribution.

3.5. Agreement Between Reference and Predicted Affinity Values

Agreement between reference and predicted affinity values was further examined using Bland-Altman analysis. The mean signed difference was -0.081, while the approximate 95% limits of agreement ranged from -2.675 to 2.513, as shown in Figure *5*.

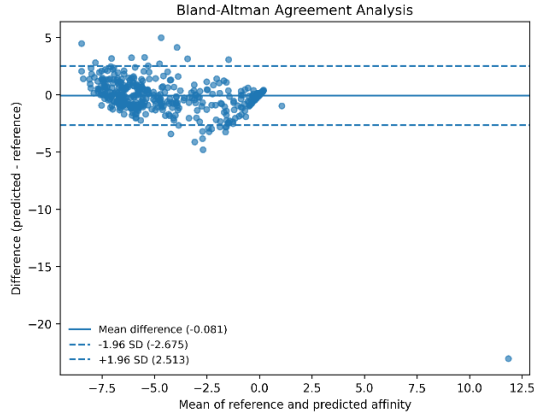


Figure 5. Bland-Altman analysis on the fixed hold-out test set

(Notes: The solid line shows the mean signed difference (-0.081), and the dashed lines show the approximate 95% limits of agreement (-2.675, 2.513))

Most observations were located within the limits of agreement, indicating a small average bias for most samples. However, one extreme observation fell substantially below the lower limit and corresponds to the large underprediction identified in the preceding analyses. Therefore, the small mean difference reflects good average agreement but does not imply uniform accuracy across all samples.

3.6. Component Ablation and Statistical Analysis

A component-removal analysis was performed on the same fixed hold-out test set to assess the contributions of SE channel attention, spatial attention, and residual connections. Four variants removed SE, spatial attention, both attention mechanisms, or residual connections. Table *4* summarizes the results.

Table 4. Component-removal ablation results on the fixed hold-out test set ($N = 693$)

Configuration	RMSE	MAE	Pearson r	R^2
Full 3D-ReSE-VN	1.3251	0.6420	0.9098	0.8271
No SE	1.3526	0.6651	0.9105	0.8199
No Spatial	1.3400	0.6436	0.9107	0.8232
No SE/Spatial	**1.2857**	**0.6310**	**0.9162**	**0.8373**
No Residual	1.3121	0.6445	0.9126	0.8305

(Notes: Best values are shown in bold)

Removing SE alone increased RMSE and MAE, while removing spatial attention produced performance close to that of the full model. In contrast, removing both attention mechanisms yielded the numerically best aggregate performance, with an RMSE of 1.2857, MAE of 0.6310, Pearson $r = 0.9162$, and $R^2 = 0.8373$. The No-Residual variant also slightly improved RMSE and R^2 but produced a marginally higher MAE. These findings indicate that the architectural components do not provide uniformly additive improvements under the present evaluation setting. Paired Wilcoxon signed-rank tests on absolute prediction errors were therefore performed, followed by Holm correction for multiple comparisons. Table *5* reports statistical results. A positive ΔMAE indicates a higher mean absolute error for the full model relative to the corresponding ablated model.

Table 5. Paired Wilcoxon analysis of the full model and ablated variants ($N = 693$), with Holm correction

Comparison	ΔMAE	W	Raw p	Holm p
Full vs. No SE	0.0230	100317	0.000158	**0.000632**
Full vs. No Spatial	0.0016	114187	0.251267	0.502533
Full vs. No SE/Spatial	-0.0111	115781	0.398151	0.502533
Full vs. No Residual	0.0024	103429	0.001433	**0.004300**

After Holm correction at $\alpha = 0.05$, significant differences were observed between the full model and the No-SE variant ($p = 0.000632$) and between the full model and the No-Residual variant ($p = 0.004300$). No significant differences were observed for No Spatial or No SE/Spatial (both $p = 0.502533$). Importantly, although No SE/Spatial produced the best aggregate metrics, its paired absolute-error distribution was not significantly different from that of the full model. The ablation results therefore suggest interactions among residual learning, channel attention, and spatial attention rather than independent benefits from every component.

3.7. Limitations

The dataset contains 4,620 protein-ligand complexes across 21 cancer-associated targets but only 14 unique ligands, which limit chemical diversity and broader conclusions about unseen chemotypes. In addition, random partitioning is less stringent than ligand- or target-disjoint evaluation because related samples may appear across different partitions. The cross-validation results should therefore be interpreted as robustness across random partitions rather than direct evidence of generalization to unseen ligand or target identities. Prediction accuracy was also non-uniform across the affinity distribution, as indicated by the residual trend, elevated RMSE in the highest-affinity quantile, and the extreme Bland-Altman observation. Furthermore, the ablation study showed that the full architecture did not outperform every simplified variant, and the numerically strongest No SE/Spatial configuration was not significantly different from the full model after Holm correction.

3.8. Future Work

Future work should evaluate 3D-ReSE-VN on larger and more chemically diverse datasets using explicit ligand- and target-disjoint protocols. Increasing representation at affinity extremes would also help determine whether the observed residual trend and large prediction errors persist. Sampling strategies, weight loss, and related approaches may be explored to improve learning in underrepresented affinity regions. Further studies should also examine interactions among residual learning, SE channel attention, and spatial attention through repeated ablation across multiple datasets and chemically distinct evaluation settings. Such analysis would clarify whether the complete architecture provides reproducible advantages or whether simpler configurations are preferable under data regimes.

4. CONCLUSION

This study presented 3D-ReSE-VN, a three-dimensional residual attention network that combines voxelized docking-derived protein-ligand complexes with residual learning, squeeze-and-excitation channel attention, and spatial attention. On the fixed hold-out test set, the model achieved an RMSE of 1.3251, an MAE of 0.6420, a Pearson correlation of 0.9098, and an R^2 of 0.8271, demonstrating competitive performance among the reported architectures. Random five-fold cross-validation produced an RMSE of 1.9679 ± 0.8242, an MAE of 0.6952 ± 0.1079, a Pearson correlation of 0.8169 ± 0.1095, and an R^2 of 0.6724 ± 0.1755, indicating sensitivity to data partitioning.

The limited diversity of 14 ligands further constrains conclusions about generalization to unseen chemical space. Future work should therefore evaluate the framework on larger, chemically diverse datasets using ligand- and target-disjoint validation.

ACKNOWLEDGMENTS

The authors would like to thank Yuan Ze University, Taiwan, the University of Udine, Italy, the National Taipei University of Nursing and Health Sciences, Taiwan, and Can Tho University, Viet Nam, for their support and research facilities. No external funding was received for this study.

CONFLICT OF INTEREST

The authors declare that they have no known financial or non-financial conflicts of interest that could have influenced the research presented in this manuscript.

REFERENCES

Abdelkader, G. A., & Kim, J.-D. (2024). Advances in protein-ligand binding affinity prediction via deep learning: A comprehensive study of datasets, data preprocessing techniques, and model architectures. *Current Drug Targets*, *25*(15), 1041–1065. *https://doi.org/10.2174/0113894501330963240905083020*

Adeshakin, F. O., Adeshakin, A. O., Afolabi, L. O., Yan, D., Zhang, G., & Wan, X. (2021). Mechanisms for modulating anoikis resistance in cancer and the relevance of metabolic reprogramming. *Frontiers in Oncology*, *11*, 626577. *https://doi.org/10.3389/fonc.2021.626577*

Born, J., Huynh, T., Stroobants, A., Cornell, W. D., & Manica, M. (2022). Active site sequence representations of human kinases outperform full sequence representations for affinity prediction and inhibitor generation: 3D effects in a 1D model. *Journal of Chemical Information and Modeling*, *62*(2), 240–257. *https://doi.org/10.1021/acs.jcim.1c00889*

Burley, S. K., Bhatt, R., Bhikadiya, C., Bi, C., Biester, A., Biswas, P., Bittrich, S., Blaumann, S., Brown, R., Chao, H., … & Zardecki, C. (2025). Updated resources for exploring experimentally-determined PDB structures and computed structure models at the RCSB Protein Data Bank. *Nucleic Acids Research*, *53*(D1), D564–D574. *https://doi.org/10.1093/nar/gkae1091*

Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. (2021). AutoDock Vina 1.2.0: New docking methods, expanded force field, and Python bindings. *Journal of Chemical Information and Modeling*, *61*(8), 3891–3898. *https://doi.org/10.1021/acs.jcim.1c00203*

Guo, J. (2024). Improving structure-based protein-ligand affinity prediction by graph representation learning and ensemble learning. *PLOS ONE*, *19*(1), e0296676. *https://doi.org/10.1371/journal.pone.0296676*

Huang, K., Xiao, C., Glass, L. M., & Sun, J. (2021). MolTrans: Molecular interaction transformer for drug-target interaction prediction. *Bioinformatics*, *37*(6), 830–836. *https://doi.org/10.1093/bioinformatics/btaa880*

Isert, C., Atz, K., Riniker, S., & Schneider, G. (2024). Exploring protein-ligand binding affinity prediction with electron density-based geometric deep learning. *RSC Advances*, *14*, 4492–4502. *https://doi.org/10.1039/D3RA08650J*

Jiang, D., Hsieh, C.-Y., Wu, Z., Kang, Y., Wang, J., Wang, E., Liao, B., Shen, C., Xu, L., Wu, J., Cao, D., & Hou, T. (2021). InteractionGraphNet: A novel and efficient deep graph representation learning framework for accurate protein-ligand interaction predictions. *Journal of Medicinal Chemistry*, *64*(24), 18209–18232. *https://doi.org/10.1021/acs.jmedchem.1c01830*

Jiang, M., Shao, Y., Zhang, Y., Zhou, W., & Pang, S. (2023). A deep learning method for drug-target affinity prediction based on sequence interaction information mining. *PeerJ*, *11*, e16625. *https://doi.org/10.7717/peerj.16625*

Kim, H., Taslakjian, B., Kim, S., Tirrell, M. V., & Guler, M. O. (2024). Therapeutic peptides, proteins and their nanostructures for drug delivery and precision medicine. *ChemBioChem*, *25*(8), e202300831. *https://doi.org/10.1002/cbic.202300831*

Kyro, G. W., Brent, R. I., & Batista, V. S. (2023). HAC-Net: A hybrid attention-based convolutional neural network for highly accurate protein-ligand binding affinity prediction. *Journal of Chemical Information and Modeling*, *63*(7), 1947–1960. *https://doi.org/10.1021/acs.jcim.3c00251*

Lai, H., Wang, L., Qian, R., Huang, J., Zhou, P., Ye, G., Wu, F., Wu, F., Zeng, X., & Liu, W. (2024). Interformer: An interaction-aware model for protein-ligand docking and affinity prediction. *Nature Communications*, *15*, 10223. *https://doi.org/10.1038/s41467-024-54440-6*

Liang, L., Duan, Y., Zeng, C., Wan, B., Yao, H., Liu, H., Lu, T., Zhang, Y., Chen, Y., & Shen, J. (2024a). CPIScore: A deep learning approach for rapid scoring and interpretation of protein-ligand binding interactions. *Journal of Chemical Information and Modeling*, *64*(23), 8809–8823. *https://doi.org/10.1021/acs.jcim.4c01175*

Liu, H., Chen, P., Zhai, X., Huo, K.-G., Zhou, S., Han, L., & Fan, G. (2024b). PPB-Affinity: Protein-protein

binding affinity dataset for AI-based protein drug discovery. *Scientific Data*, *11*, 1316. *https://doi.org/10.1038/s41597-024-03997-4*

Liu, H., Hu, B., Chen, P., Wang, X., Wang, H., Wang, S., Wang, J., Lin, B., & Cheng, M. (2024c). Docking Score ML: Target-specific machine learning models improving docking-based virtual screening in 155 targets. *Journal of Chemical Information and Modeling*, *64*(14), 5413–5426. *https://doi.org/10.1021/acs.jcim.4c00072*

Liu, M., Meng, X., Mao, Y., Li, H., & Liu, J. (2024d). ReduMixDTI: Prediction of drug-target interaction with feature redundancy reduction and interpretable attention mechanism. *Journal of Chemical Information and Modeling*, *64*(23), 8952–8962. *https://doi.org/10.1021/acs.jcim.4c01554*

Liu, Y., Wang, Y., Wang, Q., Peng, M., Chen, Y., Lin, Y., Shen, D., Liu, X., Xu, S., & Liu, B. (2026). An electron-density point-cloud framework for robust protein-ligand interaction prediction. *Nature Communications*, *17*, 7424. *https://doi.org/10.1038/s41467-026-74196-5*

Mastropietro, A., Pasculli, G., & Bajorath, J. (2023). Learning characteristics of graph neural networks predicting protein-ligand affinities. *Nature Machine Intelligence*, *5*, 1427–1436. *https://doi.org/10.1038/s42256-023-00756-9*

Morehead, A., Giri, N., Liu, J., Neupane, P., & Cheng, J. (2026). Assessing the potential of deep learning for protein-ligand docking. *Nature Machine Intelligence*, *8*, 32–41. *https://doi.org/10.1038/s42256-025-01160-1*

Mursal, M., Hussain, S., Ahmad, M., & Khan, M. F. (2024). *Navigating the computational seas: A comprehensive overview of molecular docking software in drug discovery*. IntechOpen. *https://doi.org/10.5772/intechopen.1004802*

Nguyen, T., Le, H., Quinn, T. P., Nguyen, T., Le, T. D., & Venkatesh, S. (2021). GraphDTA: Predicting drug-target binding affinity with graph neural networks. *Bioinformatics*, *37*(8), 1140–1147. *https://doi.org/10.1093/bioinformatics/btaa921*

Poziemski, J., & Siedlecki, P. (2026). Application of vision transformers to protein-ligand affinity prediction. *Scientific Reports*. *https://doi.org/10.1038/s41598-026-62430-5*

Rayka, M., & Naghavi, S. S. (2025). Uncertainty quantification enables reliable deep learning for protein-ligand binding affinity prediction. *Scientific Reports*, *15*, 43156. *https://doi.org/10.1038/s41598-025-27167-7*

Rezaei, M. A., Li, X., Wu, D., Li, C., & Li, Y. (2022). Deep learning in drug design: Protein-ligand binding affinity prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *19*(1), 407–417. *https://doi.org/10.1109/TCBB.2020.3046945*

Smith, G. L., Banegas, M. P., Acquati, C., Chang, S., Chino, F., Conti, R. M., Greenup, R. A., Kroll, J. L., Liang, M. I., Pisu, M., Primm, K. M., Roth, M. E., Shankaran, V., & Yabroff, K. R. (2022). Navigating financial toxicity in patients with cancer: A multidisciplinary management approach. *CA: A Cancer Journal for Clinicians*, *72*(5), 437–453. *https://doi.org/10.3322/caac.21730*

Wang, C., Wang, W., Wang, M., Deng, J., Sun, C., Hu, Y., & Luo, S. (2024a). Different evasion strategies in multiple myeloma. *Frontiers in Immunology*, *15*, 1346211. *https://doi.org/10.3389/fimmu.2024.1346211*

Wang, D. D., Wu, W., & Wang, R. (2024b). Structure-based, deep-learning models for protein-ligand binding affinity prediction. *Journal of Cheminformatics*, *16*, 2. *https://doi.org/10.1186/s13321-023-00795-9*

Wang, H., Liu, H., Ning, S., Zeng, C., & Zhao, Y. (2022). DLSSAffinity: Protein-ligand binding affinity prediction via a deep learning model. *Physical Chemistry Chemical Physics*, *24*(17), 10124–10133. *https://doi.org/10.1039/D1CP05558E*

Wang, Y., Qiu, Z., Jiao, Q., Chen, C., Meng, Z., & Cui, X. (2021). Structure-based protein-drug affinity prediction with spatial attention mechanisms. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 92–97). IEEE. *https://doi.org/10.1109/BIBM52615.2021.9669781*

Wu, M.-H., Xie, Z., & Zhi, D. (2025). A Folding-Docking-Affinity framework for protein-ligand binding affinity prediction. *Communications Chemistry*, *8*, 108. *https://doi.org/10.1038/s42004-025-01506-1*

Yu, J., Sheng, X., Fan, Z., Wang, Z., Cao, D., Tang, Y., Hao, Y., Zhang, Y., Shao, P., Ma, H., Cao, T., Zeng, C., Rao, J., Chen, M., Chen, K., Li, X., Teng, D., Luo, X., Wang, M., Zhang, S., & Zheng, M. (2026). Atomic-level protein-ligand recognition with PBCNet2.0 for probe discovery. *Nature Chemical Biology*, *22*, 1299–1310. *https://doi.org/10.1038/s41589-026-02241-x*

Zdrazil, B., Felix, E., Hunter, F., Manners, E. J., Blackshaw, J., Corbett, S., de Veij, M., Ioannidis, H., Mendez Lopez, D., Mosquera, J. F., Magarinos, M. P., Bosc, N., Arcila, R., Kizilören, T., Gaulton, A., Bento, A. P., Adasme, M. F., Monecke, P., Landrum, G. A., & Leach, A. R. (2024). The ChEMBL database in 2023: A drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Research*, *52*(D1), D1180–D1192. *https://doi.org/10.1093/nar/gkad1004*

Zhang, X., Ma, H., Gao, Y., Liang, Y., Du, Y., Hao, S., & Ni, T. (2024). The tumor microenvironment: Signal transduction. *Biomolecules*, *14*(4), 438. *https://doi.org/10.3390/biom14040438*

Zhou, G., & DiMaio, F. (2025). Protein ligand structure prediction: From empirical to deep learning approaches. *Current Opinion in Structural Biology*, *91*, 102998. *https://doi.org/10.1016/j.sbi.2025.102998*