if available

THE FIRST INTERNATIONAL CONGRESS ON HEALTH SERVICES TRANSFORMATION

اولین کنگره بین المللی تحول خدمات سلامت

۲۷ و ۲۸ مهر ماه

Code: 1357



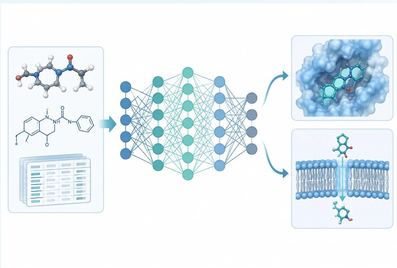


Systematic Review of Machine Learning and Deep Learning Approaches for Target Binding and Drug Permeability Prediction

Maryam shahin[1], Alireza Zarei*[2]

(1) B.Sc. Student in Biomedical Engineering, Apadana Institute of Higher Education, Shiraz, Iran

(2) MSc in Biomedical Engineering, Department of Biomedical Engineering, Imam Reza International University, Mashhad, Iran



Backgrounds

- **The Drug Discovery Bottleneck:**High attrition rates in pharmaceutical pipelines are primarily driven by late-stage pharmacokinetic (**PK**) and toxicokinetic (**TK**) failures.
- **Pivotal Determinants:**Target Binding Affinity (**TBA**) and Cell Membrane Permeability (e.g., **BBB**, **Caco-2**) are critical ADMET determinants of bioavailability and therapeutic efficacy.
- **Limitations of Classical Assays:**Traditional *in vitro* / *in vivo* empirical assays and rigid molecular docking are cost-prohibitive, low-throughput, and time-intensive.
- **The AI Paradigm Shift:**Machine Learning (**ML**) and Deep Learning (**DL**) offer robust, data-driven *in silico* screening alternatives.
- **Study Objective:**Systematic evaluation of state-of-the-art architectures (2025–2026), moving from 1D representations to **2D GNNs**, **3D embeddings**, **Multi-Task Learning**, and **Explainable AI (XAI)**.

Methods

1. **Sources:** PubMed, Scopus, IEEE Xplore, Web of Science.
2. **Scope:** Studies published 2025–2026.
3. **Selection Flow:** 600 screened →90 assessed →**23 included**.
4. **Quality Tool:** MINORS criteria.
5. **Synthesis:** Narrative analysis of architectural and dataset trends.

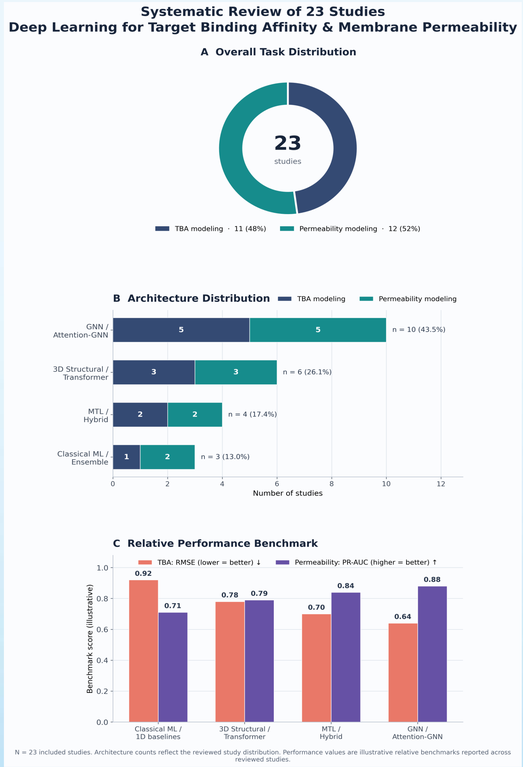


Findings

Target Domain	Dominant Architectures	Key Advantages	Typical Benchmark Gains
Target Binding Affinity (TBA)	Attention-based GNNs, 3D Transformers	Captures spatial ligand-protein interaction hotspots	15–22% reduction in RMSE vs. 1D descriptors
Membrane Permeability	Multi-Task Learning (MTL), Graph ConvNets	Efficient cross-learning; mitigates low-data scarcity	10–12% increase in AUC-ROC & PR-AUC
Model Explainability (XAI)	Integrated Gradients, Subgraph Attribution	Visualizes pharmacophores; enables rational design	Bridges the gap between "Black-box" predictions and experimental validation

Comparison of Molecular and Target Representation Techniques in Predictive Modeling

Technique Category	Primary Modalities	Key Advantages	Limitations
1D Sequential Descriptors	SMILES (Ligand), FASTA (Protein)	Computationally lightweight; highly compatible with NLP architectures (RNNs, 1D-CNNs).	Fails to capture 3D conformational changes and folding dynamics; treats molecules as text.
2D Graph Representations	Molecular Graphs (Nodes=Atoms, Edges=Bonds)	Explicitly captures molecular topology and substructure connectivity.	Ignores spptial geometry and protein pocket cavities critical for binding.
3D Structural Embeddings	Point Clouds, Voxel Grids, Distance Matrices	State-of-the-art representation of spatial geometry, physical constraints, and stereochemistry.	Highly computationally intensive; heavily reliant on high-quality crystal structures (e.g., PDB).



Conclusion

- **Dominance of Graph-Based Architectures:**Graph Neural Networks (GNNs) and Attention-augmented variants consistently achieve superior predictive performance across both Target Binding Affinity (TBA) and membrane permeability benchmarks, primarily due to their natural representation of 2D/3D molecular topologies.
- **Synergy in Multi-Task Learning (MTL):**Jointly optimizing binding affinity and pharmacokinetic properties (e.g., Caco-2 / PAMPA permeability) through MTL and hybrid frameworks effectively mitigates label scarcity and improves cross-domain generalizability.
- **Persistent Translational Bottlenecks:**Despite high in silico benchmark metrics, real-world drug discovery pipelines remain constrained by out-of-distribution (OOD) generalization failures and limited model interpretability.
- **Future Directions**
- **Multimodal Foundation Models:** Integration of molecular graphs, 3D structural pockets, and text-based biomedical knowledge graphs.
- **Uncertainty Quantification:** Establishing robust confidence metrics prior to wet-lab synthesis and biological assay validation.

References

1. Larsson S, Johansson E, Nilsson A. Multitask deep learning for solubility, permeability, protein binding, and clearance prediction. Pharmacophore. 2025;16(5):1.
2. Ma Z, Niu M, Song Y, Zhang L, Wang W, Wang T, et al. An interpretable, thermodynamics-based deep learning framework for predicting and optimizing drug membrane permeability. J Med Chem. 2026;69(8):9489-506.
3. Chakraborty A, Gholap AD, Khuspe PR, Sundaram G, Webster TJ, Khalid M, et al. Advancing drug discovery with AI: Machine and deep learning strategies for target identification and precision nanomedicine. Int J Nanomedicine. 2026;21:600651.