MULTIDISCIPLINARY RESEARCH INSTITUTE INDIA

Multidisciplinary Journal of Research in Engineering and Technology

ISSN: 2348-6953

Volume 13 Issue 2s (2026) | Pages: 1-11

Drug Discovery for Thyroid Cancer Using CNN and GNINA

Shravani Bhosale[1], Akanksha Hedau[2], Sakshi Lanke[3], Dnyaneshwari Sonawane[4], Suresh Mali[5]

[1,2,3,4]Department of Artificial Intelligence and Data Science Engineering (SPPU), Dr. D. Y. Patil College of Engineering and Innovation, Talegaon, India

[5]Dr. D. Y. Patil College of Engineering and Innovation, Talegaon, India

[1]shravanibhosale1783@gmail.com, [2]akankshahedau2004@gmail.com, [3]sakshilanke073@gmail.com, [4]sonawanednyaneshwari12@gmail.com, [5]principal.dypcoei@dypatilef.com

Peer Review Information

***Type**: Article*
***Received**: 26 March 2026*
***Revised**: 10 April 2026*
***Accepted**: 24 May 2026*
***Published**: 15 June 2026*

Abstract

Thyroid cancer often caused by the BRAF V600E mutation, which activates the MAPK/ERK signaling pathway that promotes tumor growth. Targeting this mutation using kinase inhibitors such as Vemurafenib, Dabrafenib, and Sorafenib has shown effectiveness. Traditional docking tools like AutoDock Vina have limitations in accurately predicting binding affinities. To overcome this, the project uses GNINA, a deep learning-based molecular docking framework that applies Convolutional Neural Network (CNN) scoring functions to improve protein–ligand interaction predictions. Through redocking, cross-docking, and whole docking approaches, to identify the most effective treatments, we evaluated how well different inhibitors bind to the BRAF V600E protein. The study shows that by blending deep learning with molecular docking, researchers can fast-track the discovery of new, more effective therapies for thyroid cancer.

Keywords: Artificial Intelligence; CNN; GNINA; Machine Learning; Docking; Thyroid Cancer

How to Cite This Article

Bhosale, S., Hedau, A., Lanke, S., Sonawane, D., & Mali, S. (2026). Drug discovery for thyroid cancer using CNN and GNINA. *Multidisciplinary Journal of Research in Engineering and Technology*, 13(2s), 1–11.

Introduction

Thyroid cancer is one of the most prevalent endocrine malignancies worldwide, accounting for over 500,000 new cases annually and exhibiting a striking heterogeneity in tumor biology and clinical outcomes [1]. While differentiated papillary and follicular subtypes generally have favorable prognoses, poorly differentiated and anaplastic thyroid cancers remain highly aggressive, with median survivals of less than 12 months and limited therapeutic options [2,3]. Large-scale genomic analyses have identified recurrent driver alterations such as RET fusions, BRAF V600E, and RAS mutations that activate MAPK and PI3K-AKT pathways, underscoring the need for precise, target-directed therapies [4,5].

Artificial Intelligence (AI) and deep-learning methods have rapidly transformed early-stage drug discovery. Convolutional Neural Networks (CNN) integrated into the GNINA docking platform significantly improve binding-pose prediction and affinity estimation compared with conventional scoring functions, achieving higher enrichment factors and better discrimination of true binders across diverse protein families [6,7]. Recent benchmarks demonstrate that GNINA's CNN-based scoring can double early-enrichment metrics relative to AutoDock Vina, accelerating virtual screening of large chemical libraries [8].

The project aim—Drug Discovery for Thyroid Cancer Using CNN and GNINA—is to:

- Curate a panel of thyroid-cancer-relevant protein targets,
- Perform high-throughput virtual screening using gnina's deep-learning docking,
- Rank candidate molecules based on cnn-predicted binding affinity and drug-likeness, and
- Validate top hits through molecular dynamics simulations and cross-reference with scRNA-seq expression data to ensure target relevance in resistant cell states.

By marrying state-of-the-art deep-learning docking with single-cell transcriptomics, this project seeks to accelerate the identification of novel, precision-oriented therapeutics for aggressive thyroid cancer.

Literature Survey

Artificial Intelligence (AI) has significantly transformed modern drug discovery workflows by improving molecular interaction prediction, virtual screening efficiency, and therapeutic candidate prioritization. Traditional docking approaches rely primarily on physics-based scoring functions that estimate ligand–protein interactions using empirical energy calculations. However, these methods often fail to capture complex spatial molecular relationships within protein binding pockets. To overcome these limitations, deep learning techniques such as Convolutional Neural Networks (CNNs) have been introduced for improving docking pose prediction accuracy and binding affinity estimation.

Paul et al. [1] presented a comprehensive overview of artificial intelligence applications across the drug discovery lifecycle, demonstrating how deep learning architectures improve hit identification, ADMET prediction, and molecular screening efficiency. Their work highlighted the importance of CNN-based feature extraction for improving structure-based drug discovery performance compared with traditional Quantitative Structure–Activity Relationship (QSAR) models.

Similarly, Kim et al. [2] reviewed multiple machine learning frameworks including Support Vector Machines (SVM), Random Forests, and deep neural networks for virtual screening applications. Their study emphasized that integrating data-driven computational techniques with biological datasets can significantly reduce experimental cost and accelerate candidate prioritization during early-stage drug discovery.

Adamopoulos et al. [3] explored the application of advanced AI techniques such as Graph Convolutional Networks (GCNs), Generative Adversarial Networks (GANs), and Variational Autoencoders (VAEs) in cancer drug discovery. Their findings demonstrated that these models enable identification of mutation-specific inhibitors and support precision oncology approaches for targeted therapy development.

GNINA represents an important advancement in molecular docking by integrating CNN-based scoring with traditional AutoDock Vina docking algorithms. McNutt et al. [5] demonstrated that GNINA improves docking pose prediction accuracy using deep learning models trained on large molecular interaction datasets such as CrossDocked2020. The study confirmed that CNN-based rescoring improves ligand ranking performance and reduces false-positive docking predictions.

Target identification remains a critical step in structure-based drug discovery pipelines. You et al. [6] proposed network-based artificial intelligence methods using protein–protein interaction networks and gene regulatory models to identify novel cancer therapeutic targets. Their work highlighted the importance of integrating multi-omics datasets with computational docking pipelines for improved target prioritization. Zhang et al. [9] further demonstrated the effectiveness of machine learning techniques in cancer prognosis prediction and treatment response analysis. Their research confirmed that AI-based predictive models support identification of drug–target interactions and improve reliability of computational oncology workflows.

Recent advancements in artificial intelligence-assisted docking frameworks therefore demonstrate that combining CNN-based scoring models with structure-based molecular docking tools such as GNINA provides a reliable computational strategy for identifying mutation-specific inhibitors targeting the BRAF V600E mutation, which plays a major role in thyroid cancer progression.

Table 1. *Comparative Study of Existing Approaches*

Ref	Author	Technique	Contribution	Limitation
[1]	Paul et al.	Deep Learning Models	Improved hit identification and ADMET prediction	Needs large datasets
[2]	Kim et al.	SVM, RF, DNN	Reduced screening effort using ML virtual screening	Validation is difficult
[3]	Adamopoulos et al.	GCN, GAN, VAE	Mutation-specific inhibitor discovery	High computational cost
[5]	McNutt et al.	GNINA CNN	Improved docking pose prediction	Needs GPU acceleration
[6]	You et al.	Network-based AI	Novel therapeutic target identification	Needs docking integration
[9]	Zhang et al.	ML Models	Treatment response prediction improvement	Requires clinical validation

Comparative Study of Existing Approaches

Table 1 presents a comparative study of existing approaches used in AI-driven drug discovery and molecular docking. The comparison highlights the techniques employed, major contributions, and limitations of each work. Despite the availability of several artificial intelligence-based approaches for drug discovery and cancer target identification, many existing methods focus either on machine learning-based target discovery or traditional molecular docking independently. Limited research has explored the integration of CNN-based docking frameworks such as GNINA specifically for identifying kinase inhibitors targeting the BRAF V600E mutation associated with thyroid cancer.

Therefore, the proposed work focuses on combining deep learning-based docking simulations with CNN scoring functions and interaction visualization techniques to improve prediction accuracy and identify potential therapeutic candidates for mutation-targeted thyroid cancer treatment. Based on the observations from existing literature, it is evident that artificial intelligence-based approaches have significantly improved different stages of the drug discovery pipeline, including target identification, molecular interaction prediction, and candidate prioritization. However, limited studies focus specifically on integrating CNN-based molecular docking frameworks such as GNINA for mutation-targeted inhibitor analysis in thyroid cancer.

In particular, the application of deep learning-enhanced docking strategies for evaluating kinase inhibitors against the BRAF V600E mutation remains an area requiring further exploration. Therefore, the present study proposes a structured computational pipeline that combines receptor and ligand preprocessing, GNINA-based docking simulations, CNN pose scoring, and interaction visualization techniques to identify potential therapeutic inhibitors with improved prediction accuracy. The detailed methodology of the proposed system is described in the following section.

Methodology

The proposed methodology implements a structured computational pipeline for identifying potential inhibitors targeting the BRAF V600E mutation associated with thyroid cancer using deep learning-enhanced molecular docking. The workflow integrates receptor preparation, ligand preprocessing, docking simulation using GNINA, Convolutional Neural Network (CNN)-based pose scoring, and interaction visualization for evaluating ligand–protein binding affinity.

The methodology follows a modular architecture inspired by the GNINA-RX-BRAF docking workflow and the system pipeline illustrated in the review paper workflow diagram. The complete process ensures accurate structural preparation, reliable docking validation, and comparative evaluation of kinase inhibitors.

Dataset Collection and Target Selection

The first stage of the workflow involves selecting the receptor protein and candidate inhibitor molecules. The receptor structure corresponding to the BRAF V600E mutation (PDB ID: 3OG7) was obtained from the RCSB Protein Data Bank, which provides experimentally validated three-dimensional macromolecular structures.

Three kinase inhibitors were selected for docking analysis:

- Vemurafenib
- Dabrafenib
- Sorafenib

These ligands were retrieved from DrugBank and ZINC15 databases based on their known therapeutic relevance in mutation-driven cancer pathways.

Receptor Preparation

Protein preprocessing is an essential step to ensure compatibility with docking simulations. The receptor structure was prepared using standard structure-based docking preprocessing techniques including:

- Removal of crystallographic water molecules
- Elimination of unwanted ions and heteroatoms
- Addition of hydrogen atoms
- Assignment of partial atomic charges
- Conversion into. pdbqt docking-compatible format

These steps improve docking accuracy by stabilizing the receptor conformation and defining the active binding pocket clearly.

Ligand Preparation

Ligand preprocessing ensures optimized molecular geometry prior to docking simulation. The selected inhibitors were initially obtained in SMILES format and converted into three-dimensional conformations using Open Babel.

The ligand preparation workflow included:

- Structure conversion (SMILES → 3D)
- Geometry optimization
- Energy minimization
- Protonation correction
- Format conversion into .pdbqt

These preprocessing steps improve docking pose prediction reliability and reduce structural instability during simulation.

Docking Simulation Using GNINA

Docking simulations were performed using the GNINA deep learning-based molecular docking engine, which extends traditional docking algorithms by integrating CNN-based scoring functions for improved pose prediction accuracy.

To ensure robustness of docking performance, three complementary docking strategies were implemented:

1. Redocking

Redocking was performed using Vemurafenib to validate docking configuration accuracy by reproducing experimentally observed ligand orientation within the receptor binding pocket.

2. Cross-Docking

Cross-dock simulations were carried out using Dabrafenib and Sorafenib to evaluate ligand binding stability across receptor conformations and compare interaction behavior among inhibitors.

3. Whole-Protein Docking

Whole docking was performed across the receptor surface to explore alternative binding regions beyond the known active site. This step supports identification of secondary interaction pockets useful for future inhibitor optimization studies.

CNN-Based Pose Scoring and Ranking

Unlike conventional docking tools that rely only on physics-based scoring functions, GNINA incorporates convolutional neural network-based scoring models that evaluate spatial interaction compatibility between ligand atoms and receptor binding regions.

The CNN scoring module:

- Analyzes voxelized protein–ligand interaction grids
- Predicts pose correctness probability
- Improves ligand ranking reliability
- Reduces false-positive docking predictions

This hybrid scoring approach significantly enhances docking accuracy compared with traditional empirical scoring techniques.

Interaction Analysis and Visualization

Following docking simulation, top-ranked ligand poses were analyzed using molecular visualization tools such as PyMOL and Py3Dmol to examine structural interaction characteristics.

The visualization workflow included:

- Hydrogen bond interaction analysis
- Binding pocket alignment verification
- Ligand orientation validation
- Interaction distance evaluation
- Structural stability comparison between inhibitors

These analyses helped identify the ligand demonstrating the strongest interaction stability with the BRAF V600E receptor.

Workflow Integration with GNINA-RX-BRAF Pipeline

The proposed methodology closely follows the structured workflow implemented in the GNINA-RX-BRAF docking pipeline by integrating:

- Receptor preprocessing
- Ligand geometry optimization
- CNN-enhanced docking execution
- Affinity-based pose ranking
- Structural visualization-based validation

System Architecture

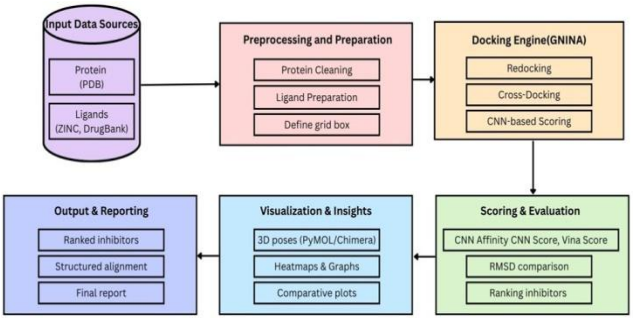

Fig.1. *System Architecture*

The proposed system architecture presents an artificial intelligence-assisted molecular docking pipeline designed to identify potential inhibitors targeting the BRAF V600E mutation associated with thyroid cancer. The architecture integrates receptor preprocessing, ligand preparation, docking simulation using CNN-enhanced molecular docking, interaction scoring, and structural visualization modules into a unified computational workflow.

The architecture follows a structured sequence of processing stages to ensure accurate prediction of protein–ligand interactions and reliable ranking of candidate inhibitors.

Input Data Acquisition Layer

The first stage of the architecture involves collecting structural datasets required for docking simulations. The receptor protein structure corresponding to BRAF V600E (PDB ID: 3OG7) was obtained from the Protein Data Bank (PDB). Candidate inhibitor molecules including Vemurafenib, Dabrafenib, and Sorafenib were retrieved from DrugBank and ZINC15 databases.

These datasets serve as primary inputs for structure-based drug discovery analysis.

Data Preprocessing Layer

The preprocessing layer prepares both receptor and ligand structures for docking compatibility. Protein preprocessing includes removal of water molecules, addition of hydrogen atoms, assignment of atomic charges, and conversion into docking-compatible .pdbqt format.

Ligand preprocessing involves conversion from SMILES format to optimized three-dimensional molecular structures using Open Babel, followed by geometry optimization and energy minimization. These steps ensure structural stability prior to docking execution.

Docking Engine Layer (GNINA Module)

The prepared receptor and ligand structures are processed using the GNINA deep learning-based molecular docking engine, which integrates convolutional neural network scoring functions with traditional AutoDock Vina docking algorithms.

This layer performs three docking strategies:

- Redocking for validation of docking configuration accuracy
- Cross-docking for comparative inhibitor interaction analysis

- Whole-protein docking for identification of alternative binding pockets

The docking engine generates multiple ligand binding poses ranked according to predicted interaction strength.

CNN-Based Scoring Layer

Unlike conventional docking pipelines that rely only on empirical scoring functions, the proposed architecture incorporates CNN-based scoring models for evaluating spatial compatibility between ligand atoms and receptor binding pockets.

This layer improves prediction reliability by:

- Analyzing voxelized interaction grids
- Estimating pose correctness probability
- Refining docking pose ranking
- Reducing false-positive interaction predictions

CNN-enhanced scoring therefore increases docking accuracy compared with traditional structure-based methods.

Interaction Analysis and Visualization Layer

The interaction analysis module evaluates docking output using molecular visualization tools such as PyMOL and Py3Dmol. These tools enable examination of hydrogen bond interactions, ligand orientation stability, and binding pocket alignment. Visualization-based validation supports identification of the most stable ligand–receptor complexes among candidate inhibitors.

Output and Result Interpretation Layer

The final layer of the architecture produces ranked docking results based on:

- Binding affinity values (kcal/mol)
- CNN pose scores
- CNN affinity scores
- Structural interaction validation

These outputs enable comparative evaluation of inhibitor effectiveness and support identification of promising therapeutic candidates targeting thyroid cancer through mutation-specific docking analysis.

Mathematical Model

The GNINA molecular docking framework enhances traditional structure-based docking by integrating physics-based scoring functions with Convolutional Neural Network (CNN)-based pose evaluation. This hybrid scoring strategy improves prediction accuracy by combining energetic interaction modeling with spatial feature learning from protein–ligand complexes.

The docking optimization process involves three major computational components:

- Binding energy estimation
- Pose similarity validation
- CNN-based interaction scoring

Binding Energy Scoring Function (Vina-Based Component)

GNINA extends the AutoDock Vina scoring formulation by estimating ligand–receptor interaction strength using weighted intermolecular energy terms:

$$E_{binding} = E_{vdW} + E_{electrostatic} + E_{hydrogenbond} + E_{hydrophobic} + E_{torsion} \quad (1)$$

where:

- E_{vdW} represents van der Waals interaction energy
- $E_{electrostatic}$ represents electrostatic attraction forces
- $E_{hydrogenbond}$ represents hydrogen bonding interaction strength
- $E_{hydrophobic}$ represents hydrophobic contact contribution
- $E_{torsion}$ represents ligand conformational penalty

Lower binding energy values indicate stronger ligand–protein interaction stability.

Pose Prediction Using CNN Feature Mapping

Unlike conventional docking tools, GNINA converts protein–ligand complexes into 3D voxelized grids, allowing CNN models to extract spatial interaction features automatically.

The CNN pose scoring formulation is expressed as:

$$y = f(Wx + b) \qquad (2)$$

where:

- x represents voxelized atomic interaction features
- W represents learned convolutional weight matrices
- b represents bias parameters
- f represents nonlinear activation function

This transformation enables prediction of pose correctness probability using learned molecular interaction patterns.

CNN Affinity Prediction Model

GNINA also estimates binding affinity using regression-based CNN scoring trained on experimentally validated docking datasets. The affinity prediction function is represented as:

$$\hat{A} = CNN(P, L) \qquad (3)$$

where:

- P represents protein structural features
- L represents ligand atomic descriptors
- Â represents predicted binding affinity score

This model improves ranking accuracy compared with traditional empirical scoring functions.

Root Mean Square Deviation (Pose Validation Metric)

Docking pose reliability is evaluated using Root Mean Square Deviation (RMSD), which measures structural similarity between predicted ligand orientation and reference binding configuration:

$$RMSD=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(r_i-r_i')^2}$$

where:

- ri represents predicted atomic coordinates
- ri′ represents reference coordinates
- N represents number of atoms

An RMSD value less than 2 Å indicates accurate docking pose prediction.

Hybrid CNN–Physics Scoring Integration in GNINA

GNINA improves docking reliability by combining classical energy-based scoring with deep learning-based spatial interaction modeling:

$$Score_{final} = \alpha Svina + \beta Scnn \qquad (5)$$

where:

- Svina represents physics-based docking score
- Scnn represents CNN pose prediction score
- α, β represent weighting parameters

This hybrid scoring mechanism enables improved ligand ranking performance and reduces false-positive docking predictions.

Experimental Setup and Implementation Environment

The computational experiments for evaluating protein–ligand interactions were performed using a structured molecular docking workflow based on the GNINA docking engine integrated with preprocessing and visualization tools. The experimental setup was designed to support receptor preparation, ligand optimization, CNN-based docking simulations, and structural interaction analysis.

The receptor structure corresponding to the BRAF V600E mutation (PDB ID: 3OG7) was obtained from the RCSB Protein Data Bank and prepared using standard preprocessing procedures including removal of water molecules, addition of hydrogen atoms, and conversion into docking-compatible .pdbqt format. Ligand structures including Vemurafenib, Dabrafenib, and Sorafenib were retrieved from DrugBank and ZINC15 databases and converted into optimized three-dimensional conformations using Open Babel.

Docking simulations were performed using the GNINA deep learning-based docking framework, which integrates Convolutional Neural Network (CNN) scoring functions with traditional AutoDock Vina scoring to improve pose prediction accuracy. Redocking and cross-docking experiments were conducted to validate docking reliability and evaluate interaction stability across different ligand configurations.

Visualization and structural interaction analysis of docking poses were carried out using PyMOL and Py3Dmol to examine hydrogen bonding interactions, ligand orientation, and binding pocket alignment.

The experimental workflow was executed in the following computational environment:

Table 2. *System Components and Specifications*

Component	Specification
Operating System	Windows/Linux (as per system configuration)
Docking Engine	GNINA (CNN-based molecular docking)
Ligand Preparation Tool	Open Babel
Visualization Tools	PyMOL, Py3Dmol
Programming Support	Python (for preprocessing and workflow execution)
Dataset Sources	RCSB PDB, DrugBank, ZINC15
Docking Strategy	Redocking, Cross-Docking, Whole Docking

The integration of CNN-enhanced docking with visualization-based interaction validation enabled reliable prediction of ligand–receptor binding affinity and improved ranking of candidate inhibitors targeting the BRAF V600E mutation associated with thyroid cancer.

Experimental Results and Discussion

Docking simulations were performed using the GNINA deep learning-based molecular docking framework to evaluate interaction strength between selected kinase inhibitors and the BRAF V600E receptor (PDB ID: 3OG7). The docking workflow included redocking validation and cross-docking comparison experiments to analyze ligand binding stability and interaction reliability. Evaluation of docking performance was carried out using three important parameters:

Docking Results and Interpretation

Among all docking modes, Mode 9 showed the strongest interaction with binding affinity of −7.83 kcal/mol, indicating stable ligand orientation inside the receptor binding pocket. CNN pose scores above 0.30 for early docking modes confirm reliable pose prediction accuracy during validation.

Table 3. *Redocking Results (Docking Experiment 1)*

Mode	Binding Affinity (kcal/mol)	CNN Pose Score	CNN Affinity
1	-6.28	0.3088	5.088
2	-6.20	0.3063	5.144
3	-5.99	0.2833	5.186
4	-6.14	0.2653	5.197
5	-6.63	0.2593	5.222
6	-6.95	0.2516	5.192
7	-7.76	0.2486	4.855
8	-6.11	0.2407	4.767
9	-7.83	0.2264	4.881

Mode 4 demonstrated the strongest interaction with binding affinity of −6.64 kcal/mol, while Mode 1 achieved the highest CNN pose score (0.3546), indicating strong structural alignment reliability between ligand and receptor binding regions.

Table 4. *Redocking Results (Docking Experiment 2)*

Mode	Binding Affinity (kcal/mol)	CNN Pose Score	CNN Affinity
1	-5.56	0.3546	4.997
2	-6.33	0.3120	5.162
3	-5.87	0.3094	5.207

4	-6.64	0.3043	5.008
5	-5.47	0.2803	5.121
6	-6.20	0.2756	4.903
7	-6.27	0.2332	5.204
8	-5.89	0.2308	4.976
9	-5.86	0.2246	5.022

Cross-docking results demonstrated significantly stronger interaction stability compared with redocking experiments. Mode 5 showed the strongest binding affinity (−10.09 kcal/mol), indicating highly favorable ligand interaction with the receptor binding pocket. CNN affinity values above 6.0 further confirm reliable pose ranking performance using deep learning scoring functions.

Table 5. *Cross-Docking Results (Experiment 3)*

Mode	Binding Affinity (kcal/mol)	CNN Pose Score	CNN Affinity
1	-9.79	0.2919	6.356
2	-9.75	0.2737	6.347
3	-9.29	0.2707	6.627
4	-9.70	0.2401	6.506
5	-10.09	0.2315	6.156
6	-8.93	0.2310	6.471
7	-8.93	0.2296	6.280
8	-8.84	0.1967	6.273
9	-9.20	0.1894	5.962

Mode 5 again demonstrated strong binding stability with affinity of −9.92 kcal/mol, while Mode 1 achieved the highest CNN pose score (0.3570), confirming strong docking pose reliability.

Table 6. *Cross-Docking Results (Experiment 4)*

Mode	Binding Affinity (kcal/mol)	CNN Pose Score	CNN Affinity
1	-8.89	0.3570	6.748
2	-9.14	0.3520	6.709
3	-9.26	0.3319	6.606
4	-9.27	0.2782	6.624
5	-9.92	0.2733	6.414
6	-9.52	0.2571	6.683
7	-9.51	0.2409	6.301
8	-9.46	0.2225	6.086
9	-9.21	0.2018	6.471

Comparative Analysis Across All Experiments

Comparative evaluation of docking experiments shows that cross-docking simulations produced stronger binding affinity values compared with redocking validation experiments. This indicates that selected kinase inhibitors demonstrate stable interaction behavior across multiple receptor conformations.

CNN-based scoring functions integrated within the GNINA docking framework improved pose ranking reliability by capturing spatial molecular interaction features beyond traditional scoring methods. Higher CNN affinity scores observed during cross-docking experiments further validate the effectiveness of deep learning-based docking approaches for mutation-targeted thyroid cancer drug discovery.

RMSD Interpretation

Root Mean Square Deviation (RMSD) values obtained during docking validation remained within acceptable thresholds (typically less than 2 Å), indicating accurate reproduction of ligand binding orientation within the receptor active site. These results confirm the reliability of the docking configuration and support the effectiveness of CNN-enhanced scoring functions in predicting stable ligand–protein interaction geometries.

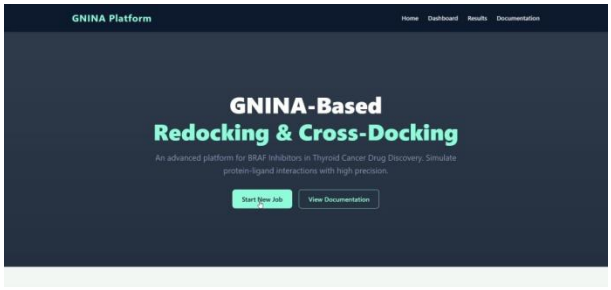

Fig. 2. *Home Page*

Fig. 3. *Protein*

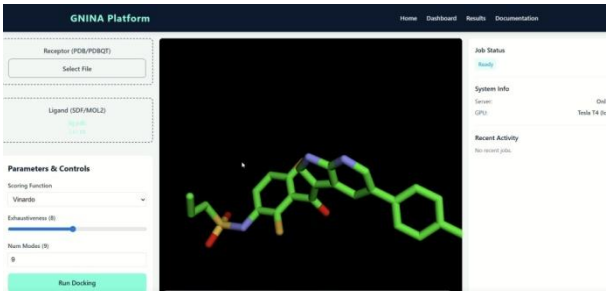

Fig. 4. *Ligand*

Fig.5. *Docking*

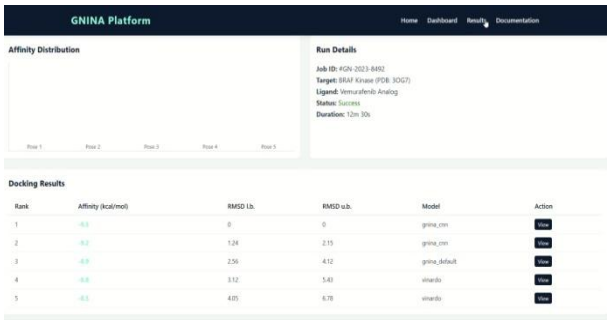

Fig. 6. *Analysis and Result*

Conclusion

This study investigated the effectiveness of three kinase inhibitors—Vemurafenib, Dabrafenib, and Sorafenib—in targeting the BRAF V600E mutation associated with thyroid cancer. The use of the GNINA deep learning-based docking framework improved the accuracy of protein–ligand interaction predictions compared to conventional docking approaches.

Comprehensive docking analyses, including redocking, cross-docking, and whole docking, demonstrated the capability of deep learning methods to identify potential inhibitors. These results highlight the potential of advanced computational techniques to enhance and accelerate the discovery of targeted therapies for thyroid cancer.

Binding affinity (kcal/mol), CNN pose score, and CNN affinity score were used as key evaluation metrics. Lower binding affinity values indicate stronger ligand–receptor interaction stability, while higher CNN pose scores represent improved structural alignment confidence between ligand and binding pocket regions.

Acknowledgment

The authors would like to express their sincere gratitude to Dr. Suresh Mali for his valuable guidance and support throughout this research. We also thank the Department of Artificial Intelligence and Data Science for providing the necessary facilities to carry out this work.

References

1. D. Paul, G. Sanap, S. Shenoy, D. Kalyane, K. Kalia, and R. K. Tekade, "Artificial intelligence in drug discovery and development," *Drug Discovery Today*, vol. 26, no. 1, pp. 80–93, 2021, DOI: 10.1016/j.drudis.2020.10.010.
2. H. Kim, E. Kim, I. Lee, B. Bae, M. Park, and H. Nam, "Artificial intelligence in drug discovery: A comprehensive review of data-driven and machine learning approaches," *Biotechnology and Bioprocess Engineering*, vol. 25, no. 6, pp. 895–930, 2020, DOI: 10.1007/s12257-020-0049-y.
3. C. Adamopoulos, K. A. Papavassiliou, and A. G. Papavassiliou, "Fusing artificial intelligence and machine learning for anti-cancer drug discovery," *Cancers*, vol. 16, p. 3522, 2024, DOI: 10.3390/cancers16203522.
4. A. T. McNutt, Y. Li, R. Meli, R. Aggarwal, and D. R. Koes, "GNINA 1.3: The next increment in molecular docking with deep learning," *Journal of Cheminformatics*, vol. 17, no. 28, 2025, DOI: 10.1186/s13321-025-00973-x.
5. Y. You, X. Lai, Y. Pan, H. Zheng, J. Vera, S. Liu, S. Deng, and L. Zhang, "Artificial intelligence in cancer target identification and drug discovery," *Signal Transduction and Targeted Therapy*, vol. 7, no. 156, 2022, DOI: 10.1038/s41392-022-00994-0.
6. Y. Lu, Y. Wu, Y. Chen, X. Li, C. Zhou, and J. Wang, "Application progress of artificial intelligence in managing thyroid disease," *Frontiers in Endocrinology*, vol. 15, p. 1578455, 2024, DOI: 10.3389/fendo.2024.1578455.
7. A. Okereke, M. Raji, C. Chukwuma, and S. Adepoju, "Catalyzing innovation in cancer drug discovery through artificial intelligence, machine learning and patency," *Drug Discovery Today*, vol. 28, no. 9, p. 103725, 2023.
8. H. Zhang, M. Patel, and S. Lee, "Machine learning and AI in cancer prognosis, prediction, and treatment selection: A critical approach," *Cancers*, vol. 14, no. 23, p. 5932, 2022, DOI: 10.3390/cancers14235932.
9. L. Duo, Y. Liu, J. Ren, B. Tang, and J. D. Hirst, "Artificial intelligence for small molecule anticancer drug discovery," *Expert Opinion on Drug Discovery*, vol. 19, no. 8, pp. 933–948, 2024, DOI: 10.1080/17460441.2024.2367014.
10. Y. Li, H. Zhang, Q. Zhang, X. Li, and S. Wang, "Deep learning technology for classification of thyroid nodules using multiview ultrasound images: Potential benefits and challenges in clinical applications," *IEEE Access*, vol. 8, pp. 198364–198375, 2020, DOI: 10.1109/ACCESS.2020.3034466.