Can machine learning-based docking approaches outperform or complement classical docking approaches in virtual screening?

Thi Ngoc Lan Vu[1,2,3], Hosein Fooladi[1,2,3] and Johannes Kirchmair[1,2]

[1]*Christian Doppler Laboratory for Molecular Informatics in the Biosciences, Department for Pharmaceutical Sciences, University of Vienna, 1090 Vienna, Austria*
[2]*Department of Pharmaceutical Sciences, Division of Pharmaceutical Chemistry, University of Vienna, Josef-Holaubek-Platz 2, 1090 Vienna, Austria*
[3]*Vienna Doctoral School of Pharmaceutical, Nutritional and Sport Sciences (PhaNuSpo), University of Vienna, 1090 Vienna, Austria*

In recent years, machine learning (ML) based docking approaches have become available. Unlike the classical search-based approach, ML-based docking programs sample ligand poses using regression or generative modeling[1]. They have been shown to yield competitive performance in pose prediction[1] (compared to existing, force field-based approaches), but their performance in virtual screening (VS) remains open for investigation. A recent study[2] takes root mean square deviation (RMSD) between poses generated by an ML-based docking approach and two classical docking programs into a consensus scoring scheme. The reported improvement promises potential benefits of ML-based docking approaches in VS.

The present study proposes a new protocol for adopting ML-based docking approaches for VS. Their performance is investigated on a conventional VS benchmarking dataset. Furthermore, their generalizability for VS is evaluated on a newly curated dataset with dissimilar protein structures. The results suggest the potential application of ML-based docking approaches for VS.

Bibliography **:**

[1] Corso, G.; Jing, B.; Barzilay, R.; Jaakkola, T. Diffdock: Diffusion steps, twists, and turns for molecular docking. ICLR (2023).
[2] Nelen, J.; Carmena-Bargueño, M.; Martínez-Cortés, C.; Rodríguez-Martínez, A.; Villalgordo-Soto, J. M.; Pérez-Sánchez, H. ESSENCE-Dock: A Consensus-Based Approach to Enhance Virtual Screening Enrichment in Drug Discovery. J. Chem. Inf. Model. 64, 5 (2024) 1605–1614.