

# A geometric XAI approach to protein pocket detection.

Giovanni Bocchi<sup>1,\*</sup>, Patrizio Frosini<sup>2</sup>, Alessandra Micheletti<sup>1</sup>, Alessandro Pedretti<sup>3</sup>, Gianluca Palermo<sup>4</sup>, Davide Gadioli<sup>4</sup>, Carmen Gratteri<sup>5</sup>, Filippo Lunghini<sup>6</sup>, Andrea Rosario Beccari<sup>6</sup>, Anna Fava<sup>6</sup> and Carmine Talarico<sup>6</sup>

<sup>1</sup>Department of Environmental Science and Policy, University of Milan, Via Saldini 50, Milano, 20133, Italy

<sup>2</sup>Department of Mathematics, University of Bologna, Piazza di Porta S. Donato 5, Bologna, 40126, Italy

<sup>3</sup>Department of Pharmaceutical Sciences, University of Milan, Via Mangiagalli 25, Milano, 20133, Italy

<sup>4</sup>Department of Electronics, Information and Bioengineering, Politecnico di Milano, Via Ponzio 34/5, 20133, Italy

<sup>5</sup>Department of Health Sciences, University Magna Græcia di Catanzaro, Viale Europa, Catanzaro, 88100, Italy

<sup>6</sup>Dompé Farmaceutici S.p.A., Via Tommaso de Amicis 95, Napoli, 80123, Italy

## Abstract

Protein pocket detection is an essential step in structure-based virtual screening methods for identifying potential drug targets. To facilitate efficient molecular docking, an accurate determination of target binding sites is indispensable. In this study, we present GENEOnet, an innovative machine learning model based on Group Equivariant Non-Expansive Operators (GENEOs) for protein pocket detection. Our proposed method sets itself apart from other artificial intelligence techniques in the domain due to its reduced number of parameters, increased transparency, and integration of prior knowledge. The experimental assessment validates GENEOnet's efficacy with a limited training dataset, surpassing several established state-of-the-art methods based on multiple critical performance indicators computed using extensive public datasets of ligand-protein complexes. GENEOnet, the result of an ongoing collaborative effort between Italian universities and the pharmaceutical company Dompé Farmaceutici S.p.A., is accessible as a web service at <https://geneonet.exscalate.eu> to enable the scientific community to evaluate the pre-trained model for pocket detection.

## Keywords

Molecular Docking, Protein Pocket Detection, Equivariance, GENEOnet

## 1. Introduction

Molecular docking, a critical research area in both pharmaceuticals and academia, is a simulation technique utilized to predict the binding mode of small molecules with respect to specific target proteins based on their three-dimensional structures [1]. This method plays a crucial

---

*Late-breaking work, Demos and Doctoral Consortium, colocated with The 2nd World Conference on eXplainable Artificial Intelligence: July 17–19, 2024, Valletta, Malta*

\*Corresponding author.

✉ giovanni.bocchi1@unimi.it (G. Bocchi); patrizio.frosini@unibo.it (P. Frosini); alessandra.micheletti@unimi.it (A. Micheletti); alessandro.pedretti@unimi.it (A. Pedretti); gianluca.palermo@polimi.it (G. Palermo); davide.gadioli@polimi.it (D. Gadioli); carmen.gratteri@studenti.unicz.it (C. Gratteri); filippo.lunghini@dompe.com (F. Lunghini); andrea.beccari@dompe.com (A. R. Beccari); anna.fava@dompe.com (A. Fava); carmine.talarico@dompe.com (C. Talarico)

ORCID 0000-0002-4017-8212 (G. Bocchi); 0000-0001-6711-1926 (P. Frosini); 0000-0002-5369-5657 (A. Micheletti); 0000-0001-5916-2029 (A. Pedretti); 0000-0001-7955-8012 (G. Palermo); 0000-0002-0143-0737 (D. Gadioli); 0009-0007-3527-2080 (C. Gratteri)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

role in elucidating molecular interactions and has significant applications, particularly in drug repurposing [2]. Repurposing existing and approved drugs involves identifying new therapeutic uses and, bypassing the time-consuming and costly drug development process, enables researchers to expedite the delivery of novel therapies to patients in an efficient and affordable manner.

Blind docking involves the tentative to accommodate the ligand trying anywhere on the protein surface. However, blind docking comes with challenges like false positives, where ligands may score well but have no experimental binding affinity to the target due to limitations in scoring functions.

To address these concerns, “informed” docking [3] is an alternative approach that restricts the search only to promising areas of a protein’s surface, known as pockets. This method reduces false positives and computational time needed for extensive virtual screening. Protein pocket detection, which is the problem of identifying such pockets on a target protein, in recent years has been addressed through machine learning and AI methods. However, explainability and trustworthiness of these methods have been neglected.

In this paper, we present some recent findings relative to GENEOnet [4] a simple and thus quite explainable machine learning method for protein pocket detection which is built using the mathematical theory of Group Equivariant Non-Expansive Operators.

## 2. Methods

### 2.1. GENEOnets

In this study, we further investigate GENEOnet a shallow network composed of Group Equivariant Non-Expansive Operator units, whose architecture is depicted in Figure 1a. For a comprehensive understanding of GENEOnets, we refer the reader to [5, 6]. GENEOnets possess good mathematical properties that make them interesting for machine learning applications. Two primary features of GENEOnets are equivariance and non-expansivity.

Equivariance [7] refers to the method’s ability to behave coherently with respect to the action of specific groups of geometrical transformations of the data. The simplest form of equivariance is called invariance: for instance, a rotation-invariant method must assign the same label to an image of an elephant facing upwards as it does to the same image rotated 90 degrees left independently on the correctness of the label.

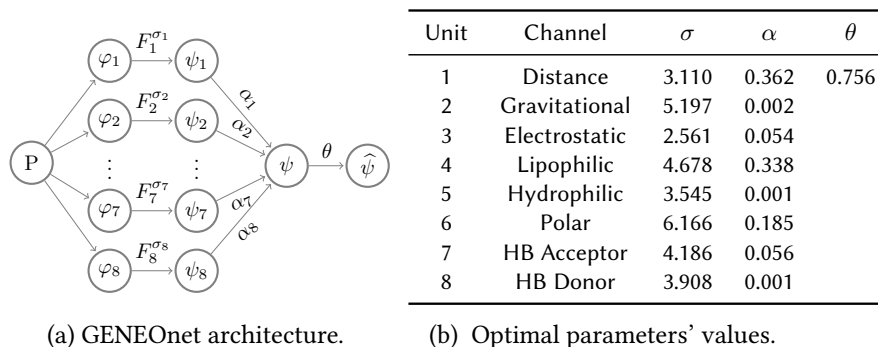
Non-expansivity means that the outputs of a GENEOnet should not be further apart than their respective inputs with respect to an appropriate distance metric. This property can be read as a form of theoretical stability, ensuring that small input perturbations result in minimal output changes.

In this regard, adversarial attacks [8] aim to assess deep learning methods’ robustness against perturbations, such as rotations and noise perturbations of the inputs, and many others. Regrettably, black box models have often been shown to be susceptible to these forms of deception, which could hardly mislead human observers. Instead, by definition, GENEOnets can be made resilient to the two attacks mentioned above, making them an attractive option when seeking robust and trustworthy results.

## 2.2. GENEOnet

GENEOnet, with a representative output displayed in Figure 2a, is broken down into six steps:

1. Data preprocessing: Given an input protein  $P$ , GENEOnet first computes a voxelization of the space enclosed within a bounding box surrounding the protein surface. Subsequently, eight potential functions are computed on the resulting voxel grid, each one modeling essential aspects of the protein structure from geometrical, physical and chemical viewpoints.
2. GENE layer: A convolutional rigid motion equivariant operator (mostly based on gaussian kernels), depending on a shape parameter  $\sigma$ , is applied to each potential function. The resulting function is normalized between 0 and 1.
3. Convex combination: The eight GENE outputs are combined via convex combination with weights  $\alpha$ . The resulting combination output, denoted as  $\psi$ , is normalized between 0 and 1 and encodes the probability that a voxel belongs to a pocket.
4. Thresholding: The final output is obtained by considering the connected components of the set of voxels where the value of  $\psi$  is above a threshold parameter  $\theta$ .
5. Evaluation: If the ground truth (i.e., the co-crystallized ligand) is accessible, the output can be compared against it using a volumetric accuracy function.
6. Scoring: Each predicted pocket is assigned a score, computed as a weighted average of  $\psi$  within the spatial region identified by a pocket, enabling ranking of pockets from the most promising to the least promising.



**Figure 1:** Model architecture and optimal parameters.

## 2.3. Training, model selection and comparison

Our study employed PDBbind, a publicly accessible database of curated protein-ligand complexes for binding affinity analysis, as the primary data source (available at <http://pdbind.org.cn/>). We retrieved 12330 complexes from the 2020 version for evaluation and analysis.

Steps 1 through 5 were considered during the training process, with backpropagation used to optimize parameters to optimize accuracy (introduced in Step 5). GENEOnet was trained on

a set of 200 molecules extracted from the retrieved complexes. However, this method did not address the challenge of identifying true pockets within high-ranked predictions (as per Step 6).

To tackle this issue, we generated multiple models by repeating the optimization process while maintaining the same initial parameter settings but varying the training sets (each comprising of 200 complexes). The final model was selected based on maximizing its ability to identify true pockets at the highest rank using a Validation set consisting of nearly 3000 complexes.

Moreover, we compared GENEOnet’s performance against other state-of-the-art methods (see [4] for results) in terms of scoring precision, volumetric matching, and shape similarity on a Test set containing approximately 9000 complexes. In all experiments, GENEOnet outperformed the others, which was unexpected due to its unique transparency, explainability, and robustness features not shared by the other methods. We will further explore some of these properties in the following section.

### 3. Analysis

In this section, our aim is to provide an additional in-depth analysis that underscores the explainability and robustness characteristics of GENEOnet, many of which can be attributed to the adoption of GENEOnets as building blocks.

#### 3.1. Prior knowledge

In the context of developing GENEOnet-based models, the selection and design of an appropriate pool of operators can be considered as a crucial step of information engineering for the problem at hand.

In the case of GENEOnet, experts in molecular docking identified lipophilicity, hydrophilicity, hydrogen bond acceptor or donor atoms, etc. (see Table 1b) as essential features to identify potential pockets. For example, regarding the lipophilicity potential, it was argued that a favorable pocket candidate should exhibit an average high value of this quantity. Thus, a GENEOnet from the family of convolutional operators with a Gaussian kernel was chosen to process this potential. The tunable shape parameter in this case is linked to the shape of the Gaussian. Analogous considerations were used to select the remaining seven GENEOnets for the other potentials.

#### 3.2. Parameters investigation

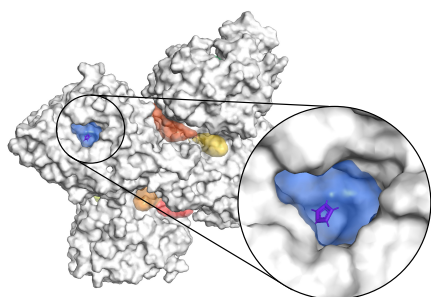
GENEOnet, as described in Section 2.2, includes a total of seventeen trainable parameters: eight shape coefficients for operator kernels, eight coefficients for the convex combination, and one threshold coefficient.

We believe that Equivariance is a primary factor contributing to GENEOnet’s relatively few trainable parameters versus other CNN-based methods like DeepSite[9] (844529) and DeepPocket[10] (665122). Indeed the requirement for equivariance necessitated the adoption of rotation-invariant kernels, limiting the kernel search to radial functions chosen to depend on only one shape parameter. This leads to a greater understanding and interpretability of GENEOnet’s kernel parameters as opposed to standard CNNs. Moreover, while not directly

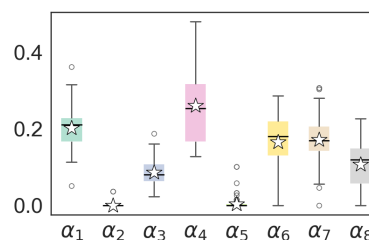
responsible for reducing the number of parameters, the convex constraint promotes sparsity in the combination coefficients, akin to Lasso regularisation. This may result in some biologically relevant GENEOnets having small coefficients due to their correlation with other operators.

Indeed, GENEOnet's limited trainable parameters facilitate precise interpretation of each of them: shape parameters influence operator kernels; convex combination coefficients serve as feature importances for potentials (Table 1b); and the threshold coefficient determines the significance of voxel activation. The convex combination coefficients warrant further exploration. The lipophilic potential, essential for pockets due to their need to exhibit lipophilic properties, receives the highest score. Surprisingly, potentials linked to hydrogen bonds, despite being biologically significant, have very low scores. As already noticed, this may be a result of the convex constraint: indeed, these potentials tend to be more concentrated in small regions surrounding acceptor or donor atoms, and thus other potentials, such as the polar potential, which might be strongly correlated with hydrogen bonds, are preferred. This could be the case with the selected model where the polar operator has a convex combination weight which is the third highest.

During the model selection process, parameter sensitivity analyses were performed as byproducts when generating 200 models from different training sets (Figure 2b shows boxplots of the convex combination coefficients). The findings largely confirm the conclusions of the previous paragraphs: on average, hydrogen bonds play a role but are outweighed by other operators, such as lipophilicity, which exhibit a coefficients distribution with a box containing higher values ( $\alpha_4$ ). If this analysis confirms the importance of potentials like those related to hydrogen bonds, it also reveals that the Gravitational and Hydrophilic potentials are given very little importance, a result that was somewhat anticipated given the nature of these potentials in connection to pocket detection.



(a) The input protein is depicted in white, while predicted pockets with colors. A zoom shows the true pocket containing the ligand.



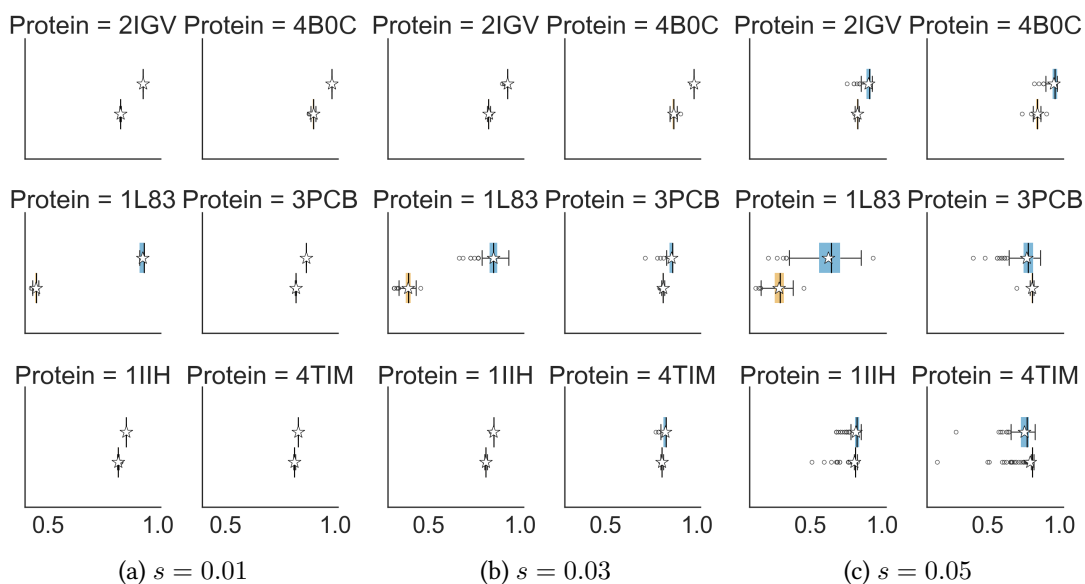
(b) Sensitivity analysis of convex combination coefficients. Lines indicate median values, whereas stars represent means. The order is the same as in Table 1b.

**Figure 2: GENEOnet example (Protein ID 6WIY) and sensitivity analysis.**

### 3.3. Robustness test

Here we tested robustness to perturbations. In our context, robustness is strictly related to the Non-Expansivity property of GENEOnets, and it is very important for pocket identification

since proteins may exhibit extremely dynamic shapes, while in the data, they are frozen until crystallizing to capture an image of the structure. To test robustness, six proteins were selected: two belonging to the set of those whose true pocket was matched by the top-ranked prediction (proteins 2IGV, 4B0C), another two from the set of pockets matched by the second-ranked prediction (proteins 1L83, 3PCB), and the remaining two from those matched by the third-ranked predicted pocket (proteins 1IIH, 4TIM). White noise was introduced by summing i.i.d. Gaussian r.v.  $\mathcal{N}(0, s^2)$  with zero mean and variance  $s^2$  to every voxel. Although i.i.d. noise is not realistic, as perturbations in the protein conformation are likely correlated because of atom interactions, it serves as a global test for robustness. We conducted multiple experiments by varying the standard deviation  $s$  in the set  $\{0.01, 0.03, 0.05\}$ , each time generating 200 different perturbations. Figure 3 shows that overlaps and scores distributions of perturbations exhibit great robustness for all values of standard deviation  $s$ . Figure 4 indicates greater robustness for rankings of top and second-ranked pockets while a mildly lower degree of stability is showed for third-ranked pockets.

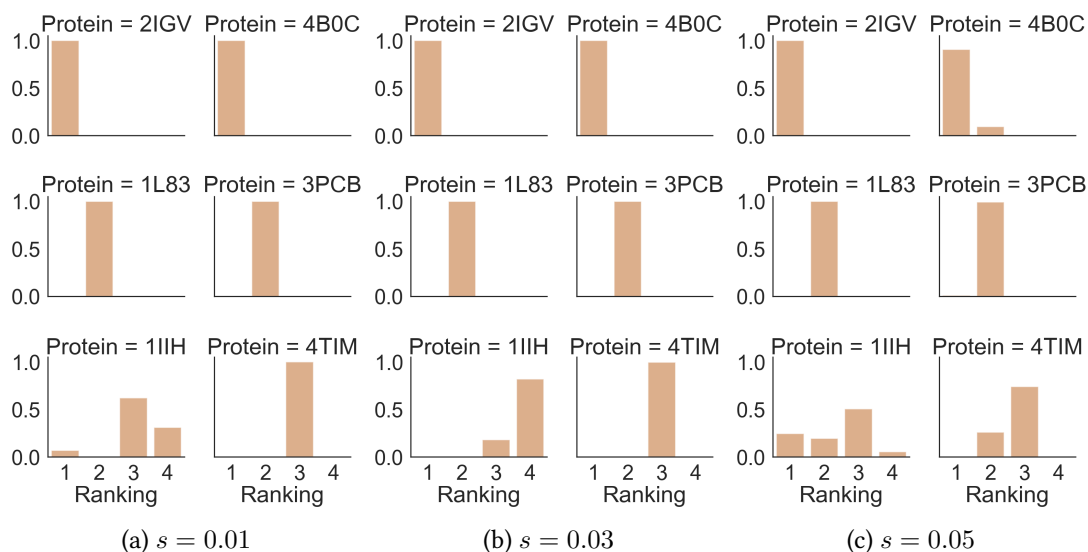


**Figure 3: Box Plots of Overlaps and Scores:** For each protein, the upper boxplot illustrates the distribution of overlaps between the predicted pocket having the highest overlap to the true pocket and the true pocket itself, while the lower boxplot depicts the corresponding distribution of scores.

## 4. Discussion

The analysis presented in the preceding section provides an initial understanding of GENEOnet’s explainability and trustworthiness, summarised as follows:

- GENEOnet can incorporate prior knowledge not only during feature selection but also in operator pool selection. Although feature selection is common in machine learning



**Figure 4: Bar Plots of Rank Positions:** Each bar chart depicts the distribution of rankings assigned to the predicted pockets that correspond to the true ones across the perturbed sample set.

algorithms, this second step is less frequently considered, making the choice of operators more complex and less stable for methods like Convolutional Neural Networks.

- GENEOnet has a smaller number of parameters than other protein pocket detection methods based on deep learning architectures. This attribute allows us to perform sensitivity analysis and provides interpretability *by design* rather than using *post hoc* explanation techniques.
- GENEOnet’s equivariance with respect to rigid motions allows us to overlook the input protein’s spatial pose and helps to reduce the number of parameters.
- GENEOnet’s non-expansivity provides the model with a reasonable degree of robustness against small physiological modifications in the input protein’s spatial conformation. The model shows great robustness for lower noise standard deviations when analysing the identifications made by the first or second predicted pockets. However, larger input changes may influence the final prediction. We leave this study to subsequent work.
- The reduced data requirements for training GENEOnet (200 complexes) compared to deep learning methods help decrease computational costs and processing times. This attribute is not directly related to explainability but makes the model more accessible as fewer resources are needed to reproduce its results or delve deeper into its analysis.

In conclusion, GENEOnet holds promise for building explainable and robust models such as GENEOnet. Future work includes creating more instances of GENEOnet-based models [11] and further evaluation of GENEOnet. A promising advancement could be the use of GENEOnet for analyzing molecular dynamics data. In molecular dynamics simulations, tracking the progression of GENEOnet scores for pockets jointly with techniques that identify breakpoints (see e.g. [12]) may help pinpoint optimal frames for bond formation.



## Acknowledgments

This research was financially supported by Dompé Farmaceutici S.p.A. Additional scientific support is acknowledged to the Italian GNAMPA - INDAM group. Computational resources were partially provided by the INDACO core facility for HPC at Università degli Studi di Milano.

**Availability:** GENEOnet can be tested at the following URL: <https://geneonet.exscalate.eu>

## References

- [1] A. C. Anderson, The process of structure-based drug design, *Chem. Biol.* 10 (2003) 787–797. doi:10.1016/j.chembiol.2003.09.002.
- [2] W. Shi, X. Chen, et al., A review of recent developments and progress in computational drug repositioning, *Curr. Pharm. Design* 26 (2020) 3059–3068. doi:10.2174/1381612826666200116145559.
- [3] D. Ghersi, R. Sanchez, Improving accuracy and efficiency of blind protein-ligand docking by focusing on predicted binding sites, *Proteins* 74 (2009) 417–424. doi:10.1002/prot.22154.
- [4] G. Bocchi, P. Frosini, et al., GENEOnet: A new machine learning paradigm based on group equivariant non-expansive operators. an application to protein pocket detection, 2022. arXiv:2202.00451.
- [5] M. G. Bergomi, P. Frosini, et al., Towards a topological-geometrical theory of group equivariant non-expansive operators for data analysis and machine learning, *Nat. Mach. Intell.* 1 (2019) 423–433. doi:10.1038/s42256-019-0087-3.
- [6] G. Bocchi, S. Botteghi, et al., On the finite representation of linear group equivariant operators via permutant measures, *Ann. Math. Artif. Intell.* 91 (2023) 465–487. doi:10.1007/s10472-022-09830-1.
- [7] T. S. Cohen, M. Welling, Group equivariant convolutional networks, in: *ICML 2016*, volume 48, 2016, pp. 2990–2999.
- [8] M. A. Alcorn, Q. Li, et al., Strike (with) a pose: Neural networks are easily fooled by strange poses of familiar objects, in: *CVPR 2019*, 2019, pp. 4840–4849. doi:10.1109/CVPR.2019.00498.
- [9] J. Jimenez, S. Doerr, et al., Deepsite: protein-binding site predictor using 3d-convolutional neural networks, *Bioinformatics* 33 (2017) 3036–3042. doi:10.1093/bioinformatics/btx350.
- [10] R. Aggarwal, A. Gupta, et al., Deep-pocket: Ligand binding site detection and segmentation using 3d convolutional neural networks, *J. Chem Inf. Model.* 62 (2022) 5069–5079. doi:10.1021/acs.jcim.1c00799.
- [11] D. Lavado, C. Soares, et al., Low-resource white-box semantic segmentation of supporting towers on 3d point clouds via signature shape identification, 2023. arXiv:2306.07809.
- [12] A. Micheletti, G. Aletti, et al., A weighted  $\chi^2$  test to detect the presence of a major change point in non-stationary markov chains, *Stat. Method. Appl.* 29 (2020) 899 – 912. doi:10.1007/s10260-020-00510-0.