Deep learning-assisted pipeline for Virtual Screening of ligand compound databases: Application on inhibiting the entry of SARS-CoV-2 into human cells

Stelios Mylonas*, Apostolos Axenopoulos*, Sotiris Katsamakas†, Ioannis Gkekas‡, Kostas Stamatopoulos‡,
Spyros Petrakis‡, and Petros Daras*
*Information Technologies Institute, Centre for Research and Technology Hellas,Thessaloniki, Greece
†Institute of Chemical Biology, National Hellenic Research Foundation, Athens, Greece
‡Institute of Applied Biosciences, Centre for Research and Technology Hellas,Thessaloniki, Greece

Emails: {smylonas,axenop,daras}@iti.gr, sotikats@eie.gr, {gkekasioannis,kostas.stamatopoulos,spetrak}@certh.gr

Abstract—**Drug discovery involves extremely costly and time consuming procedures and can be significantly benefited by computational approaches, such as virtual screening (VS). Structure-based VS relies on scoring functions which aim to evaluate the binding of a candidate compound (ligand) on a protein target. Over the last few years, the advancement of the deep learning field has led to the development of novel scoring functions based on convolutional neural networks (CNN), which have achieved state-of-the-art results. In this paper, we present an integrated end-to-end VS pipeline for application on real-world drug discovery scenarios. It combines multiple conformations of the ligand with a new CNN scoring function based on the ResNet architecture, called ResNetVS, which incorporates also the docking output score in its evaluation. After experiments on the DUD-E dataset, it has shown notable performance, especially in early enrichment, where it overcomes current benchmarks. The proposed pipeline is finally applied on the emerging case of COVID-19 pandemic, in a struggle to discover inhibitors for the viral spike protein-ACE2 interaction.**

Index Terms—**virtual screening, deep learning, ResNetVS, scoring function, covid-19**

I. INTRODUCTION

Virtual screening (VS) is a computational technique used to predict potentially bioactive chemical compounds (ligands) that bind to a specific target, usually a protein receptor or enzyme [1]. VS can substantially speed up the drug discovery process and reduce its significant costs, by reducing the time invested in expensive experimental assays and improving the hit-rate of experimental verification. It consists of narrowing down an initial large library of chemical compounds to a considerably smaller final set of hit compounds, which are then tested experimentally *in vitro* to confirm their biological activity.

Structure-based VS involves docking of candidate ligands into the protein target followed by a scoring function which estimates the likelihood that the ligand will bind to the protein with high affinity [2]. Docking is the computational task used to predict the position and orientation of a ligand when it bounds to the protein receptor. Traditional VS approaches are based solely on the docking scoring functions, which are usually physically inspired and aim to predict the binding affinity values [3]–[5]. These techniques, although interpretable, are inherently limited in their ability to capture complex interactions due to the use of fixed functional forms. With the employment of machine learning techniques, the formation of generalised data-driven scoring functions was enabled, providing more flexibility in the representation of the underlying binding mechanisms [6], [7]. However, the use of descriptors without any spatial arrangement between them can lead to loss of the precise spatial relationships existed in the raw structural data. Recent works tried to mitigate this issue by adopting a graph-based representation [8], [9].

The increasing availability of large amount of data and the advancement of the deep learning (DL) field in computer vision inspired recently the development of novel structure-based scoring schemes, which are based on 3D convolutional neural networks (CNN) [10]–[13]. The VS task is treated in this case as a binary classification problem, where the CNN is trained to discriminate between "binding" (active) and "non-binding" (decoy) complexes. Specifically, after discretizing the protein-ligand complex into a 3D grid, it is imported to a 3D-CNN which automatically extracts the most suitable features required for the discrimination. While AtomNet [10] was the first attempt to employ a shallow 3D-CNN for the VS task, the effectiveness of the DL approach was mainly displayed through an extensive study in [11], where the proposed CNN-based scoring surpassed considerably both empirical and ML-based scoring functions. The works of Bindscope [12] and Imrie *et al.* [13] further improved the obtained results by
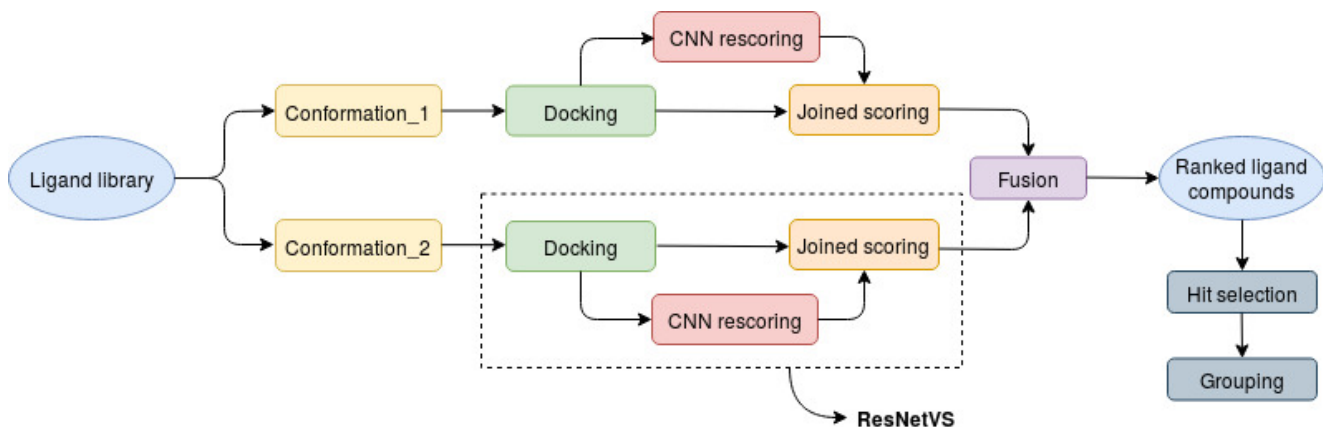
Fig. 1. Flowchart of the proposed virtual screening pipeline.

employing deeper DenseNet architectures, while the latter one achieved state-of-the-art performance by additionally creating protein-family specific models.

In this work, we propose an integrated pipeline for application on real-world VS tasks. More specifically, a novel scoring scheme is introduced, called ResNetVS, which enhances the score directly computed by the docking algorithm through combination with a 3D-CNN scoring function. In order to increase the robustness of the proposed framework, the results of multiple conformations for each ligand are considered and fused appropriately. Experiments conducted in benchmark dataset demonstrate the effectiveness of the proposed approach. Moreover, in an attempt to contribute to the endeavours against the COVID-19 pandemic, the framework is currently being applied on a real-case scenario. Specifically, a database of thousands of molecules is screened against the viral spike protein-ACE2 interface which is responsible for the entry of SARS-CoV-2 into human cells.

II. PROPOSED PIPELINE

A flowchart of our proposed VS pipeline is shown in Fig. 1. Initially, a large library of possible hit compounds is selected and is subjected by the experts to a succession of case-specific filters, in order to reduce the number of ligands that will proceed to docking.

One factor that has large impact on the docking performance is the initial conformation of the ligand compounds, i.e. their 3D structure. Since in most cases input compounds are provided in two-dimensional (2D) format, the conversion into their three-dimensional (3D) form is performed computationally. This may introduce great ambiguity in the VS pipeline, since an erroneously predicted structure can significantly mislead the geometry-based docking procedure. To alleviate this issue, multiple conformations of the ligands can be employed [1]. A parallel scheme of multiple ligand conformations was taken into account in [14], which proved more versatile and robust in handling unknown compounds, and led to higher screening performance compared to that of single conformations. In our pipeline, we adopt a similar approach with two

ligand conformations, generated with OpenBabel's toolbox [15] by employing the default genetic algorithm-based method with MMFF94 energies scoring and the weighted rotor search method for lowest energy conformer, respectively.

For the docking procedure, we used Smina [16], a fork of AutoDock Vina [5], which implements an improved empirical scoring function and minimization than its predecessor. Although previous works have discovered new ligands by performing docking screen [17]–[19], solely relying on docking scoring functions can suffer from many weaknesses [20], which can be overcome either by rescoring the obtained docked poses [21] or by applying a consensus scheme of multiple scoring functions [14].

Following the above rationales, we trained a deep learning-based method for rescoring the docked poses generated by Smina and integrated it, along with the Smina scores, into a joined scoring function, called ResNetVS. Despite the proven effectiveness of CNNs in improving the docking scores, we propose here a joined scoring, in order to handle with possible failures of the CNN model. More specifically, when employed on real-world applications, CNN-based scoring may suffer from poor generalization ability due to the inadequate size or possible biases existed in its training set. Returned values by the individual scoring methods are of different context and range, since CNN returns probability values in the range of [0,1], while Smina returns affinity values measured in kcal/mol. For this reason the following sigmoid function is applied on Smina affinities in order to obtain a score in range [0,1]:

$$Sm^{score} = \frac{1}{1 + e^{1.465*(a_f+6)}}. \tag{1}$$

For affinity values of $a_f = -6$ kcal/mol, a moderate score of 0.5 is obtained, while a larger affinity value of $a_f = -8$ kcal/mol receives a higher score of 0.95. The parameters of the sigmoid were determined after inspecting the distribution of affinity scores returned by Smina on our training set. The joined score is finally obtained through:

$$ResNetVS^{score} = Sm^{score} * CNN^{score}. \qquad (2)$$

Multiplication operates here as an intersection operator, promoting solutions with high scores in both single metrics.

A single score for each ligand should be extracted in the next step, by combining the scores of the various poses of this ligand. According to previous works, three different scoring tactics (single-pose, multi-pose by considering the best pose, multi-pose by averaging the top poses) were examined and evaluated in our experimental study in Section III. Finally, in order to mitigate the possible structural variability of the different conformations, an intersection operator, notably the minimum operator, is employed to fuse the different conformation results. Specifically, the final score of a candidate compound is the minimum of the scores from the two different conformations, forcing both 3D structures of the ligand to achieve high scores. It can also imply the existence of possible physicochemical properties that guide the binding of the ligand on the receptor. The ligand compounds are finally sorted according to the obtained multi-conformation score and the most prominent solutions, usually a top-k%, are selected for further examination.

For efficiency reasons, before proceeding to the *in vitro* examination of the activity of the top hit compounds, a further sampling is usually required. In order to obtain a representative sample, a suitable criterion, like a shape-based one, should be applied. Two molecules with a similar shape are likely to fit in the same binding pocket and thereby exhibit similar biological activity. Therefore, it is important to ensure structural diversity by selecting compounds that are different from one another, and this can be achieved by grouping the hit compounds according to their shape using clustering. Hit compounds that belong to different clusters can finally be selected for *in vitro* experiments.

A. CNN rescoring

We follow here the approach described in [11], and adopted also in [13], where a 3D CNN network is employed in order to automatically extract relevant features from a protein-ligand complex and assign to it a score of binding probability. Specifically, a complex of a protein receptor with a ligand docked pose is locally discretized into a, centered on the ligand, 3D grid of $24 \times 24 \times 24$ voxels with resolution 1 Å and is, then, imported as input to the 3D CNN. Each voxel holds in 28 channels (see Table I) the information for the existence or not of the various heavy atom types, each of them corresponding to either a protein (12) or ligand (16) smina atom type [16]. Compared to the 34 channels initially employed in both [11] and [13], we excluded those that appear more rare in the training dataset. As a representation scheme, we chose a simple boolean one which, despite its simplicity, exhibited in [11] comparable results to a most sophisticated Gaussian representation. In the boolean scheme, a voxel's channel is assigned a value of one, if it overlaps with the atom sphere of an atom of the corresponding type.

TABLE I
SMINA ATOM TYPES EMPLOYED IN RESNETVS

Type	Receptor	Ligand
AliphaticCarbonXSHydrophobe	✓	✓
AliphaticCarbonXSNonHydrophobe	✓	✓
AromaticCarbonXSHydrophobe	✓	✓
AromaticCarbonXSNonHydrophobe	✓	✓
Nitrogen	✓	✓
NitrogenXSDonor	✓	✓
NitrogenXSDonorAcceptor	✓	✓
NitrogenXSAcceptor	✓	✓
OxygenXSDonorAcceptor	✓	✓
OxygenXSAcceptor	✓	✓
Sulfur	✓	✓
Phosphorus		✓
Fluorine		✓
Chlorine		✓
Bromine		✓
Iodine		✓
GenericMetal	✓	

The only information embedded in this representation is the spatial arrangement of the atoms, their atom types and their distinction between protein or ligand.

The main difference of ResNetVS rescoring to the previous related CNN-based scoring schemes is the network architecture employed. Contrary to the 3-layered CNN used in [11] and to the DenseNet used in [13], we employ here a deeper ResNet architecture [22] of 18 layers with the exact structure being shown in the original work. Belonging to the residual networks family, the main attribute of ResNet is the existence of skip connections between adjacent layers, so as to avoid the vanishing gradient problem. Although it has shown considerable results in many computer vision problems, its application in structural bioinformatics is still limited [23].

ResNetVS was implemented on Python using the Tensorflow framework. Regarding the training process, L2 regularization was applied on the weights of all convolutional layers ($\lambda = 10^{-4}$), while batch normalization was applied with its default parameters. All models were trained for 25 epochs, with batch size of 32 samples, and were optimized by the Adam optimizer [24] with a learning rate of 10^{-3}.

B. Grouping of hit compounds

In the final step, the hit compounds are grouped into clusters based on shape similarity, i.e. compounds of similar shape should belong to the same cluster. This is achieved by utilizing an intuitive molecular shape descriptor that was introduced in [25]. The process for extracting this hand-crafted feature is as follows: taking as input the 3D structure of the molecule, the Solvent Excluded Surface (SES) is computed. Then, a set of keypoints is uniformly sampled on the surface and an Augmented Local Descriptor (ALD) is computed for each local surface patch centered at the corresponding keypoint. ALD describes the local shape of the patch as a 4-dimensional joint histogram of different geometric features. Finally, for each compound, a global descriptor of $n = 1000$ values is generated from the partial ALDs using Bag-of-Words.

In order to measure the similarity between two molecules, the chi-square metric for histogram comparison is used. Specifically, if $x = \{x_1, x_2, ..., x_n\}$ and $y = \{y_1, y_2, ..., y_n\}$ are two descriptor vectors, the chi-square distance between them is computed from:

$$d_{xy} = \frac{1}{2} \sum_{i=1}^{n} \frac{(x_i - y_i)^2}{x_i + y_i}. \tag{3}$$

These distances are then embedded in a hierarchical clustering process, which finally forms the clusters of the hit compounds. The number of clusters is controlled by a user-specified similarity threshold that defines the maximum intracluster distance.

III. EXPERIMENTS ON BENCHMARK

In order to validate the performance of ResNetVS, we conducted experiments on the benchmark dataset of DUD-E (Database of Useful Decoys: Enhanced) [26]. DUD-E contains 102 protein targets with a group of active molecules for each target (224 ligands on average) and decoys (50 decoys per active ligand), leading to more than 20.000 active and one million decoy molecules, respectively. The utilized docked poses are those provided by [11]. Following [11] and [13], the proposed scoring scheme was evaluated by 3-fold cross validation. Specifically, the set of 102 proteins was separated into three folds, based on a sequence-based clustering, which ensured that similar proteins with high sequence identity are retained in the same fold. This procedure prevents training and testing on overly similar targets and resembles more to the actual screening process of a novel target.

Due to the heavy imbalance between the active and decoy molecules in DUD-E, we used undersampling on the decoys set during training, by forcing the same number of positive and negative examples at each epoch. Especially, at each epoch a different set of negative examples is randomly chosen from the decoys set. Data augmentation was also employed by randomly rotating and translating (up to 3 Å) the input structures.

For the evaluation of the method, two well-known metrics were employed, the area under the curve (AUC) of the receiver operating characteristic (ROC) curve, which plots the true positive rate (TPR) against the false positive rate (FPR) and the AUC of the precision recall curve (PRC). According to [27], while the two metrics are correlated, AUC PRC is more informative in cases of high class imbalance, as is the case of virtual screening. Both previous metrics are global metrics, measuring the overall performance of the method on the whole set of ligand compounds. However, in practice, only a percentage of the top scoring compounds are kept for further experimental evaluation. For this reason, the enrichment factor (EF) of the ROC curve may be a most relevant metric for VS, since it evaluates the performance on the top ranked compounds by measuring the ratio of TPR to FPR at low FPR values (0.5%, 1%, 2%, 5%).

After aggregating the partial results of the three folds, Table II holds the overall performance of ResNetVS on DUD-E using different scoring schemes. i.e. the single-pose, the

	Single-pose		Multi-pose (Best pose)		Multi-pose (Avg of top 9)	
	CNN	*Joined*	*CNN*	*Joined*	*CNN*	*Joined*
AUC ROC	0.873	0.875	0.896	0.897	0.908	0.908
AUC PRC	0.33	0.35	0.375	0.385	0.438	0.447
EF (0.5%)	58.63	62.81	66.19	68.36	79.39	81.78
EF (1%)	36.4	38.39	41.36	42.36	47.27	48.42
EF (2%)	22.14	22.97	25	25.75	27.79	28.32
EF (5%)	11.28	11.44	12.53	12.71	13.37	13.48

	Smina	Ragoza *et al.* [11]	Imrie *et al.* [13]		ResNetVS
			DenseU	*DenseFS*	
AUC ROC	0.725	0.862	0.904	0.917	0.908
AUC PRC	0.1	0.263	0.368	0.443	0.447
EF (0.5%)	16.09	44.52	64.89	79.32	81.78
EF (1%)	11.37	30.65	40.92	47.99	48.42
EF (2%)	8.02	19.72	25.08	28.41	28.32
EF (5%)	5.25	10.6	12.7	13.74	13.48

multi-pose where only the best pose is considered and the multi-pose where an average of the top poses is evaluated. In the last case, and according to [13], we computed the average on the top 9 poses. In compliance to previous findings in [13], averaging the top ranked poses outperforms the other two scoring schemes in all metrics. Regarding now the incorporation of Smina output into the joined scoring, we notice a performance gain in early enrichment, which is increased as the EF thresholds are getting smaller. As previously stated, this is an important aspect in real-case VS scenarios, where we are particularly interested in the performance on the top compounds. Although in some metrics the additive contribution of the joined scoring is negligible, if we take into account that it can computed without additive cost, its inclusion in the final scoring scheme could be considered as a good practice.

Our highest scoring alternative is compared to Smina docking and to three competing 3D-CNN rescoring methods in Table III. DenseFS is the finally proposed scheme by [13], while DenseU corresponds to an intermediate model, without the employment of protein family-specific models. Our first remark is that all 3D-CNN methods outperform considerably the single Smina output, confirming the need for pose rescoring. Among the CNN rescoring methods, ResNetVS and Imrie *et al.*, that employ deeper network architectures and apply the multi-pose average scoring, achieve higher results than [11]. Comparing now our proposed scoring to DenseU, we can see that ResNetVS performs better by all metrics, especially in early enrichment factors and in PRC. Although the comparison to DenseFS is not totally fair, since it is enhanced with family-specific models, we can see that ResNetVS achieves comparable and in some cases higher accuracies (as in 0.5% EF).

IV. APPLICATION ON COVID-19

A. Problem description

In December 2019, coronavirus disease 2019 (COVID-19) emerged in Wuhan, Hubei province of China. This disease is caused by a new, highly pathogenic severe-acute-respiratory-syndrome coronavirus 2 (SARS-CoV-2) [28]. Similar to other coronaviruses, SARS-CoV-2 affects the lower respiratory tract, gastrointestinal system, heart, central nervous system and kidney [29], [30]. On the 30th of January 2020, the World Health Organization (WHO) declared the SARS-CoV-2 epidemic a public health emergency of international concern. More than 6.4 million confirmed patients were reported globally until the 2nd of June 2020, resulting in more than 380.000 related deaths.

The entry of SARS-CoV-2 into human cells is mediated by its transmembrane homotrimeric spike (S) glycoprotein (each monomer comprises an S1 and an S2 subunit) [31]. The virus interacts with the angiotensin-converting enzyme II (ACE2) receptor of the host cells through the receptor-binding domain (RBD) of its S1 subunit. Then, the S2 subunit fuses the host and viral membranes, enabling the entry of SARS-CoV-2 genome into human cells [32]. Inside the cell, the viral genome is replicated and its coding proteins are produced, the virion is assembled and the newly synthesized SARS-CoV-2 particles are released, surrounded by the membrane of the host cells [33]. Any of the above discrete steps can be pharmacologically targeted to prevent infection or spread of the virus.

Several in-silico-based approaches have been previously utilized to identify potential drugs which would specifically prevent from SARS-CoV-2 infection. Some research groups target viral enzymes which are crucial for the life cycle of the virus, such as its main protease [34], [35], some others target the N-protein that binds to the viral RNA genome [36], while others target the spike S1 subunit-ACE2 protein interaction, either by using models [37], [38] or crystal structures [39] of each protein or simulations of the interaction [40].

In this work, we focus on the interaction between the viral spike protein S1 subunit and the human ACE2 enzyme and apply the proposed VS pipeline to computationally identify novel inhibitors of this interaction. All previous relative studies have performed virtual screening using drug-repurposing libraries to identify clinically-approved molecules that would block this interaction. Although drug repurposing is of first concern, due to the avoidance of the time consuming clinical trial processes, its efficiency is doubtful. These drugs have been designed on specific substrates (targets) different than the one considered here, hence finding a drug that could specifically target spike-ACE2 interaction is unlikely. Indeed, no effective inhibitors for the entry of SARS-CoV-2 into host cells have been reported to date [41]. In our approach, we target the specified interaction with a broader set of small molecule libraries provided from ChemBridge Corporation [42].
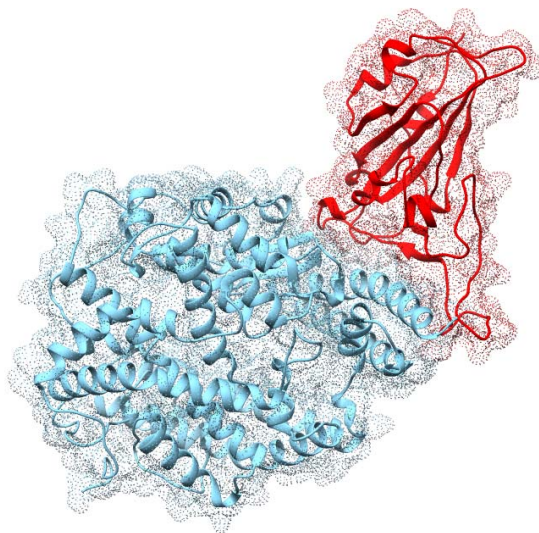

Fig. 2. Crystal structure of the complex between the S1 subunit of SARS-CoV-2 spike protein (red) and ACE2 human enzyme (blue).

B. Screening and preliminary results

The crystal structure of spike protein S1 subunit bound to the extracellular domain of ACE2 receptor (PDB ID: 6M0J) [43] was solved by X-ray crystallography at a 2.45 Å resolution and is displayed here in Fig. 2. From this complex, we isolated the coronavirus's spike protein part which have been the target in our screening process. The protein structure was accordingly protonated and charged with OpenBabel before entering the docking procedure.

Commercially available libraries provided from Chem-Bridge Corporation [42] of more than 1 million drug-like molecules have been filtered accordingly based on their pharmacokinetic (PK) properties to identify hit compounds able to disrupt this interaction. Firstly, the libraries were filtered with the help of OpenBabel [15] to clean all salt forms included within in order to have a clear calculation of compounds PK properties to the later stages. Following the cleaning of our drug-like subset we utilized FAF-drugs4 [44] filtering to eliminate pan-assay interfering compounds and unwanted metabolite moieties and, afterwards, we applied the Eli Lilly MedChem set of rules [45] along with PPI profiling [46]. The remaining compound in our datasets after those filters were applied is a median of 9.5% in total numbers.

The described screening process is in progress, since a 30% of the initial library has been processed so far, and has resulted after filtering to 26000 candidate compounds. The herein proposed pipeline was then employed to screen the filtered compounds on the targeted S-protein. Docking was constrained on the interaction site of the S-protein with the ACE2, while 50 docked poses were generated for each conformation of the ligands.

Regarding the ResNetVS rescoring, the applied CNN model was trained on the whole DUD-E dataset, following the same
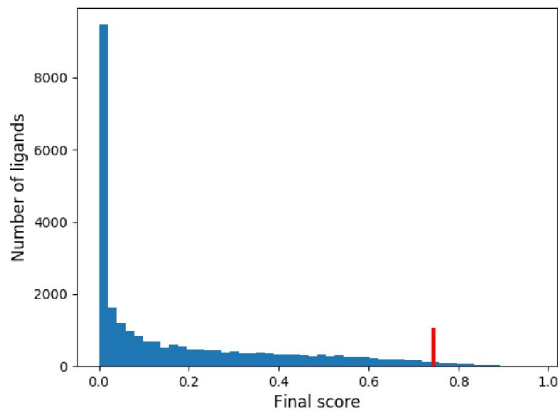
Fig. 3. Histogram of the final ligand scores obtained by ResNetVS. The red line corresponds to the selection threshold of the top-2% compounds.

TABLE IV
FINAL SCORES OF THE SO FAR TOP-20 RANKED COMPOUNDS AND THEIR
ASSIGNMENT TO GROUPS

Rank	Score	Group
1	0.972	1
2	0.951	2
3	0.931	3
4	0.924	4
5	0.918	5
6	0.916	6
7	0.915	6
8	0.905	4
9	0.901	7
10	0.900	8
11	0.899	9
12	0.897	5
13	0.896	10
14	0.895	6
15	0.895	11
16	0.894	9
17	0.893	3
18	0.893	9
19	0.891	12
20	0.889	11

Fig. 4. Heatmap of the normalized shape distances between the top-20 ranked compounds

principles as described in Section II-A. The histogram of the obtained scores is shown in Fig. 3. As we can see, our proposed scoring scheme assigned to the vast majority of the ligands very low (near zero) scores, and only few of them were distinguished. Ligands with score higher than 0.744 (red line) were selected by the top-2% rule for further examination.

These 520 selected compounds were then grouped into 122 clusters according to their shape. The scores of the top-20 ranked compounds, along with their group assignment, are indicatively presented in Table IV. Due to the incomplete experimentation, we avoided to include the compound names in our results. The separation of these compounds into clusters is illustrated when inspecting the normalized shape distances between them, which are shown in Fig. 4. Compound pairs with smaller distances between them (blue color), and therefore most similarly-shaped, were assigned to the same group.

After completing the screening process on the entire compound library, the created groups will be subject to a secondary filtering where possible interaction patterns between the ligands and the ACE2 receptor will be investigated through SMART search. As a last step, the final hit compounds will be further validated *in vitro*. Specifically, they will be tested whether they block the interaction between spike S1 subunit and ACE2 receptor in a high-throughput cell-based assay, as previously described in [47].

V. CONCLUSION

We have presented an end-to-end pipeline for performing virtual screening on large ligand databases. The proposed VS pipeline combines multiple ligand conformations with ResNetVS, an improved CNN-based scoring function, which incorporates the docking output score in its computation. The fusion of multiple ligand conformations provide robustness in cases of inaccurate ligand 3D structures, while the consideration of the docking score into the ResNetVS alleviates possible failures of the CNN scoring. After experimentation on the DUD-E database, ResNetVS achieved performance comparable to the current state of the art and in the case of early enrichment, even better. The pipeline is currently applied on an effort to confront with the COVID-19 pandemic. Using a combination of computational and experimental methods, we expect to identify compounds which may inhibit the entry of SARS-CoV-2 into human cells, by screening large compound libraries against the interaction of the viral spike S-protein with the human ACE2 enzyme.

REFERENCES

[1] A. Gimeno, M. J. Ojeda-Montes, S. Tomás-Hernández, A. Cereto-Massagué, R. Beltrán-Debón, M. Mulero, G. Pujadas, and S. Garcia-Vallvé, "The light and dark sides of virtual screening: what is there to know?" *International journal of molecular sciences*, vol. 20, no. 6, p. 1375, 2019.

[2] R. T. Kroemer, "Structure-based drug design: docking and scoring," *Current protein and peptide science*, vol. 8, no. 4, pp. 312–328, 2007.

[3] D. T. Moustakas, P. T. Lang, S. Pegg, E. Pettersen, I. D. Kuntz, N. Brooijmans, and R. C. Rizzo, "Development and validation of a modular, extensible docking program: Dock 5," *Journal of computer-aided molecular design*, vol. 20, no. 10-11, pp. 601–619, 2006.

[4] R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, and D. T. Mainz, "Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein- ligand complexes," *Journal of medicinal chemistry*, vol. 49, no. 21, pp. 6177–6196, 2006.

[5] O. Trott and A. J. Olson, "Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *Journal of computational chemistry*, vol. 31, no. 2, pp. 455–461, 2010.

[6] P. J. Ballester and J. B. Mitchell, "A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking," *Bioinformatics*, vol. 26, no. 9, pp. 1169–1175, 2010.

[7] J. D. Durrant and J. A. McCammon, "Nnscore 2.0: a neural-network receptor–ligand scoring function," *Journal of chemical information and modeling*, vol. 51, no. 11, pp. 2897–2903, 2011.

[8] J. C. Pereira, E. R. Caffarena, and C. N. dos Santos, "Boosting docking-based virtual screening with deep learning," *Journal of chemical information and modeling*, vol. 56, no. 12, pp. 2495–2506, 2016.

[9] A. Gonczarek, J. M. Tomczak, S. Zareba, J. Kaczmar, P. Dabrowski, and M. J. Walczak, "Interaction prediction in structure-based virtual screening using deep learning," *Computers in biology and medicine*, vol. 100, pp. 253–258, 2018.

[10] I. Wallach, M. Dzamba, and A. Heifets, "Atomnet: a deep convolutional neural network for bioactivity prediction in structure-based drug discovery," *arXiv preprint arXiv:1510.02855*, 2015.

[11] M. Ragoza, J. Hochuli, E. Idrobo, J. Sunseri, and D. R. Koes, "Protein–ligand scoring with convolutional neural networks," *Journal of chemical information and modeling*, vol. 57, no. 4, pp. 942–957, 2017.

[12] M. Skalic, G. Martínez-Rosell, J. Jiménez, and G. De Fabritiis, "Play-molecule bindscope: large scale cnn-based virtual screening on the web," *Bioinformatics*, vol. 35, no. 7, pp. 1237–1238, 2019.

[13] F. Imrie, A. R. Bradley, M. van der Schaar, and C. M. Deane, "Protein family-specific models using deep neural networks and transfer learning improve virtual screening and highlight the need for more data," *Journal of chemical information and modeling*, vol. 58, no. 11, pp. 2319–2330, 2018.

[14] M. Schneider, J.-L. Pons, W. Bourguet, and G. Labesse, "Towards accurate high-throughput ligand affinity prediction by exploiting structural ensembles, docking metrics and ligand similarity," *Bioinformatics*, vol. 36, no. 1, pp. 160–168, 2020.

[15] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open babel: An open chemical toolbox," *Journal of cheminformatics*, vol. 3, no. 1, p. 33, 2011.

[16] D. R. Koes, M. P. Baumgartner, and C. J. Camacho, "Lessons learned in empirical scoring with smina from the csar 2011 benchmarking exercise," *Journal of chemical information and modeling*, vol. 53, no. 8, pp. 1893–1904, 2013.

[17] N. London, R. M. Miller, S. Krishnan, K. Uchida, J. J. Irwin, O. Eidam, L. Gibold, P. Cimermančič, R. Bonnet, B. K. Shoichet *et al.*, "Covalent docking of large libraries for the discovery of chemical probes," *Nature chemical biology*, vol. 10, no. 12, p. 1066, 2014.

[18] C. Colas, C. Grewer, N. J. Otte, A. Gameiro, T. Albers, K. Singh, H. Shere, M. Bonomi, J. Holst, and A. Schlessinger, "Ligand discovery for the alanine-serine-cysteine transporter (asct2, slc1a5) from homology modeling and virtual screening," *PLoS computational biology*, vol. 11, no. 10, 2015.

[19] V. N. Stone, H. I. Parikh, F. El-rami, X. Ge, W. Chen, Y. Zhang, G. E. Kellogg, and P. Xu, "Identification of small-molecule inhibitors against meso-2, 6-diaminopimelate dehydrogenase from porphyromonas gingivalis," *PloS one*, vol. 10, no. 11, 2015.

[20] J. J. Irwin and B. K. Shoichet, "Docking screens for novel ligands conferring new biology: Miniperspective," *Journal of medicinal chemistry*, vol. 59, no. 9, pp. 4103–4120, 2016.

[21] X. Zhang, S. E. Wong, and F. C. Lightstone, "Toward fully automated high performance computing drug discovery: a massively parallel virtual screening pipeline for docking and molecular mechanics/generalized born surface area rescoring to improve enrichment," pp. 324–337, 2014.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[23] S. K. Mylonas, A. Axenopoulos, and P. Daras, "Deepsurf: A surface-based deep learning approach for the prediction of ligand binding sites on proteins," *arXiv preprint arXiv:2002.05643*, 2020.

[24] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[25] A. Axenopoulos, D. Rafailidis, G. Papadopoulos, E. N. Houstis, and P. Daras, "Similarity search of flexible 3d molecules combining local and global shape descriptors," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 13, no. 5, pp. 954–970, 2015.

[26] M. M. Mysinger, M. Carchia, J. J. Irwin, and B. K. Shoichet, "Directory of useful decoys, enhanced (dud-e): better ligands and decoys for better benchmarking," *Journal of medicinal chemistry*, vol. 55, no. 14, pp. 6582–6594, 2012.

[27] Z. Wu, B. Ramsundar, E. N. Feinberg, J. Gomes, C. Geniesse, A. S. Pappu, K. Leswing, and V. Pande, "Moleculenet: a benchmark for molecular machine learning," *Chemical science*, vol. 9, no. 2, pp. 513–530, 2018.

[28] X. Xu, P. Chen, J. Wang, J. Feng, H. Zhou, X. Li, W. Zhong, and P. Hao, "Evolution of the novel coronavirus from the ongoing wuhan outbreak and modeling of its spike protein for risk of human transmission," *Science China Life Sciences*, vol. 63, no. 3, pp. 457–460, 2020.

[29] D. Wu, T. Wu, Q. Liu, and Z. Yang, "The sars-cov-2 outbreak: what we know," *International Journal of Infectious Diseases*, 2020.

[30] Y. Zhang, X. Geng, Y. Tan, Q. Li, C. Xu, J. Xu, L. Hao, Z. Zeng, X. Luo, F. Liu *et al.*, "New understanding of the damage of sars-cov-2 infection outside the respiratory system," *Biomedicine & Pharmacotherapy*, p. 110195, 2020.

[31] J. Shang, G. Ye, K. Shi, Y. Wan, C. Luo, H. Aihara, Q. Geng, A. Auerbach, and F. Li, "Structural basis of receptor recognition by sars-cov-2," *Nature*, pp. 1–4, 2020.

[32] X. Ou, Y. Liu, X. Lei, P. Li, D. Mi, L. Ren, L. Guo, R. Guo, T. Chen, J. Hu *et al.*, "Characterization of spike glycoprotein of sars-cov-2 on virus entry and its immune cross-reactivity with sars-cov," *Nature communications*, vol. 11, no. 1, pp. 1–12, 2020.

[33] M. A. Shereen, S. Khan, A. Kazmi, N. Bashir, and R. Siddique, "Covid-19 infection: Origin, transmission, and characteristics of human coronaviruses," *Journal of Advanced Research*, 2020.

[34] M. Kandeel and M. Al-Nazawi, "Virtual screening and repurposing of fda approved drugs against covid-19 main protease," *Life sciences*, p. 117627, 2020.

[35] R. Pokhrel, P. Chapagain, and J. Siltberg-Liberles, "Potential rna-dependent rna polymerase inhibitors as prospective therapeutics against sars-cov-2," *Journal of Medical Microbiology*, p. jmm001203, 2020.

[36] P. Sarma, N. Shekhar, M. Prajapat, P. Avti, H. Kaur, S. Kumar, S. Singh, H. Kumar, A. Prakash, D. P. Dhibar *et al.*, "In-silico homology assisted identification of inhibitor of rna binding against 2019-ncov n-protein (n terminal domain)," *Journal of Biomolecular Structure and Dynamics*, pp. 1–9, 2020.

[37] S. Choudhary, Y. S. Malik, S. Tomar, and S. Tomar, "Identification of sars-cov-2 cell entry inhibitors by drug repurposing using in silico structure-based virtual screening approach," *Chemrxiv*, 2020.

[38] O. V. de Oliveira, G. B. Rocha, A. S. Paluch, and L. T. Costa, "Repurposing approved drugs as inhibitors of sars-cov-2 s-protein from molecular modeling and virtual screening," *Journal of Biomolecular Structure and Dynamics*, no. just-accepted, pp. 1–14, 2020.

[39] S. Durdagi, B. Aksoydan, B. Dogan, K. Sahin, A. Shahraki, and N. Birgül-İyison, "Screening of clinically approved and investigation drugs as potential inhibitors of sars-cov-2 main protease and spike receptor-binding domain bound with ace2 covid19 target proteins: A virtual drug repurposing study," *ChemRxiv*, 2020.

[40] M. Smith and J. C. Smith, "Repurposing therapeutics for covid-19: supercomputer-based docking to the sars-cov-2 viral spike protein and viral spike protein-human ace2 interface," *ChemRxiv*, 2020.

[41] H. Ledford, "Dozens of coronavirus drugs are in development-what happens next?" *Nature*, 2020.

[42] C. . U. ChemBridge Corporation, San Diego, "Chembridge screening libraries." [Online]. Available: https://www.chembridge.com/screening_libraries/

[43] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang *et al.*, "Structure of the sars-cov-2 spike receptor-binding domain bound to the ace2 receptor," *Nature*, pp. 1–6, 2020.

[44] D. Lagorce, L. Bouslama, J. Becot, M. A. Miteva, and B. O. Villoutreix, "Faf-drugs4: free adme-tox filtering computations for chemical biology and early stages drug discovery," *Bioinformatics*, vol. 33, no. 22, pp. 3658–3660, 2017.

[45] R. F. Bruns and I. A. Watson, "Rules for identifying potentially reactive or promiscuous compounds," *Journal of medicinal chemistry*, vol. 55, no. 22, pp. 9763–9772, 2012.

[46] O. Sperandio, C. H. Reynès, A.-C. Camproux, and B. O. Villoutreix, "Rationalizing the chemical space of protein–protein interaction inhibitors," *Drug discovery today*, vol. 15, no. 5-6, pp. 220–229, 2010.

[47] S. Petrakis, T. Raskó, J. Russ, R. P. Friedrich, M. Stroedicke, S.-P. Riechers, K. Muehlenberg, A. Möller, A. Reinhardt, A. Vinayagam *et al.*, "Identification of human proteins that modify misfolding and proteotoxicity of pathogenic ataxin-1," *PLoS Genet*, vol. 8, no. 8, p. e1002897, 2012.