



# OnionNet-2: A Convolutional Neural Network Model for Predicting Protein-Ligand Binding Affinity Based on Residue-Atom Contacting Shells

Zechen Wang<sup>1</sup>, Liangzhen Zheng<sup>2</sup>, Yang Liu<sup>1</sup>, Yuanyuan Qu<sup>1</sup>, Yong-Qiang Li<sup>1</sup>, Mingwen Zhao<sup>1</sup>, Yuguang Mu<sup>3\*</sup> and Weifeng Li<sup>1\*</sup>

<sup>1</sup>School of Physics, Shandong University, Jinan, China, <sup>2</sup>Tencent AI Lab, Shenzhen, China, <sup>3</sup>School of Biological Sciences, Nanyang Technological University, Singapore

## OPEN ACCESS

### Edited by:

Yong Wang,  
Ningbo University, China

### Reviewed by:

Martin Zacharias,  
Technical University of Munich,  
Germany  
Jian Zhang,  
Nanjing University, China

### \*Correspondence:

Yuguang Mu  
ygm@ntu.edu.sg  
Weifeng Li  
lwf@sdu.edu.cn

### Specialty section:

This article was submitted to  
Theoretical and Computational  
Chemistry,

a section of the journal  
Frontiers in Chemistry

**Received:** 04 August 2021

**Accepted:** 06 October 2021

**Published:** 27 October 2021

### Citation:

Wang Z, Zheng L, Liu Y, Qu Y, Li Y-Q, Zhao M, Mu Y and Li W (2021) OnionNet-2: A Convolutional Neural Network Model for Predicting Protein-Ligand Binding Affinity Based on Residue-Atom Contacting Shells. *Front. Chem.* 9:753002. doi: 10.3389/fchem.2021.753002

One key task in virtual screening is to accurately predict the binding affinity ( $\Delta G$ ) of protein-ligand complexes. Recently, deep learning (DL) has significantly increased the predicting accuracy of scoring functions due to the extraordinary ability of DL to extract useful features from raw data. Nevertheless, more efforts still need to be paid in many aspects, for the aim of increasing prediction accuracy and decreasing computational cost. In this study, we proposed a simple scoring function (called OnionNet-2) based on convolutional neural network to predict  $\Delta G$ . The protein-ligand interactions are characterized by the number of contacts between protein residues and ligand atoms in multiple distance shells. Compared to published models, the efficacy of OnionNet-2 is demonstrated to be the best for two widely used datasets CASF-2016 and CASF-2013 benchmarks. The OnionNet-2 model was further verified by non-experimental decoy structures from docking program and the CSAR NRC-HIQ data set (a high-quality data set provided by CSAR), which showed great success. Thus, our study provides a simple but efficient scoring function for predicting protein-ligand binding free energy.

**Keywords:** protein-ligand binding, deep learning, onionnet, residue-atom distance, structure-based affinity prediction

## 1 INTRODUCTION

Protein-ligand binding is the basic of almost all processes in living organisms Du et al. (2016) thus predicting binding affinity ( $\Delta G$ ) of protein-ligand complex becomes the research focus of bioinformatics and drug design Guedes et al. (2018); Guvench and MacKerell (2009); Ellingson et al. (2020). Theoretically, molecular dynamics (MD) simulations and free energy calculations (for instance, thermal integration method and free energy perturbation can provide accurate predictions of  $\Delta G$  relying on extensive configurational sampling and calculation, leading to a large demand in computational cost Ellingson et al. (2020); Michel and Essex (2010); Gilson and Zhou (2007); Hansen and Van Gunsteren (2014). Therefore, developing simple, accurate and efficient scoring methods to predict protein-ligand binding will greatly accelerate the drug design process Liu and Wang (2015). To achieve this, several theoretical methods (scoring functions) have been proposed. Typically, the scoring functions are based on calculations of interactions between protein and ligand atoms Du et al. (2016); Liu and Wang (2015); Huang et al. (2010); Grinter and Zou (2014). This includes quantum mechanics calculations, molecular dynamics simulations (electrostatic interaction, van der Waals

interaction, hydrogen-bond and etc.), empirical-based interacting models. Du et al. (2016); Huang et al. (2010); Grinter and Zou (2014); Huang et al. (2006).

In recent years, approaches based on machine learning (ML) have been applied in scoring functions and demonstrated great success Lo et al. (2018); Vamathevan et al. (2019); Shen et al. (2020a); Lavecchia (2015). For instance, RF-Score Ballester and Mitchell (2010) and NNScore are two pioneering ML-based scoring functions Durrant and McCammon (2010). Compared with classical approaches, these ML-based methods allow higher flexibility in selecting configurational representations or features for protein and ligand. More importantly, these methods have been demonstrated to perform better and more effective than classical approaches Ain et al. (2015); S Heck et al. (2017). Recently, the deep learning (DL) approaches have provided alternative solution. Compared with ML, the DL models perform better at learning features from the raw data to extract the relationship between these features and labels. Wang and Gao (2019); Gawehn et al. (2016); Chen et al. (2018) Thus, DL algorithms have been introduced to model the structure-activity relationships Yang et al. (2019); Winter et al. (2019); Ghasemi et al. (2018). One of the most popular methods of DL is the convolutional neural network (CNN), which is a multi-layer perceptron inspired by the neural network of living organisms Lavecchia (2019).

Inspired by the great success of DL and CNN techniques, several models applying CNN to predict protein-ligand interaction have been reported Jiménez et al. (2017); Gomes et al. (2017); Öztürk et al. (2018); Stepniewska-Dziubinska et al. (2018); Jiménez et al. (2018); Torng and Altman (2019a); Zheng et al. (2019); Morrone et al. (2020); Torng and Altman (2019b). For example, Öztürk and co-workers reported a DeepDTA model based on one-dimensional (1D) convolution, which took protein sequences and simplified molecular input line entry specification (SMILES) codes of ligand as inputs to predict drug-target  $\Delta G$  Öztürk et al. (2018). Using 3D CNN model, two independent groups developed scoring functions, named Pafnucy Stepniewska-Dziubinska et al. (2018) and  $K_{deep}$  Jiménez et al. (2018), to model the complex in a cubic box centered on the geometric center of the ligand to predict the  $\Delta G$  of protein-ligand complex. More interestingly, Russ et al. employed Graph-CNNs to automatically extract features from protein pocket and 2D ligand graphs, and demonstrated that the Graph-CNN framework can achieve superior performance without relying on protein-ligand complexes Torng and Altman (2019a). Our group has proposed a 2D convolution-based predictor, called OnionNet, based on element-pair-specific contacts between ligands and protein atoms Zheng et al. (2019). As is shown, these DL and CNN based approaches, achieved higher accuracy in  $\Delta G$  prediction than most traditional scoring functions, such as Auto Dock Morris et al. (1998); Huey et al. (2007), X-Score. Wang et al. (2002) and KScore. Zhao et al. (2008).

Physically, the dominating factors for overall binding affinity involve electrostatic interactions, van der Waals interactions, hydrogen bonds, hydration/de-hydration during complexation. Instead, for DL scoring functions, how to treat

with the high-dimensional structural information encoded in the 3D structures and convert to the low-dimensional features for ML (or DL) training is critical. For most structure-based ML/DL models, the features are usually derived from the atomic information of proteins and ligands, such as the element type and spatial coordinates of the atom and even other atomic properties Stepniewska-Dziubinska et al. (2018); Jiménez et al. (2018). In our OnionNet model, we characterized the protein-ligand interactions by the number of element-pair-specific contacts in multiple distance shells. Zheng et al. (2019); Song et al. (2020).

As we all know, the same elements in different residues have quite different physical and chemical properties, which might greatly affect the protein-ligand binding event. Therefore, it may be insufficient enough to characterize the intrinsic physical and chemical properties of proteins by dividing the protein into eight types of atoms. Considering that the twenty types of amino acids can be treated as intrinsic classification of protein compounds which involve lay features of them, like polar, apolar, aromatics, and etc. It may be more reasonable to characterize the physicochemical properties of proteins through individual residues. Especially the residues in the binding pocket, they directly participate in the construction of the binding environment, which plays a decisive role in ligand binding events. It is undeniable that using residues as the basic unit actually “coarse-grained” the protein, which will lose part of the structural information, but this may also help to improve the generalization ability of the model itself. In view of these, we anticipate that it may be more beneficial to encode protein as residues instead of atoms in developing DL scoring functions.

In this work, we proposed a simple OnionNet-2—a 2D CNN based regression model to predict protein-ligand  $\Delta G$ , which adopts the rotation-free residue-atom-specific contacts in multiple distance shells to describe the protein (residues) - ligand (atoms) interactions. The model was trained on the PDBbind database Li et al. (2014a) and tested by the comparative assessment of scoring function (CASF) benchmarks, where CASF-2016 Su et al. (2018) is employed as the primary benchmark. When the total number of shells was 62, OnionNet-2 achieved the best performance with the Pearson correlation coefficient ( $R$ ) reaching 0.864 and a root-mean-squared error (RMSE) of 1.164. When applied to the earlier version of CASF-2013, Li et al. (2014a), Li et al. (2014b), Li et al. (2018) our present model achieved  $R$  of 0.821 and RMSE of 1.357. An additional high-quality data set, the CSAR NRC-HiQ data set (consisting of two subsets) Dunbar et al. (2011) was also used to verify OnionNet-2. Our model achieved  $R$  of 0.89 for NRC-HiQ subset 1 (55 complexes) and 0.87 for NRC-HiQ subset 2 (49 complexes). The performance is indeed higher than two previous ML scoring models, RF-Score ( $R = 0.78$  and  $0.75$ , respectively) and  $K_{deep}$  ( $R = 0.72$  and  $0.65$ , respectively). We demonstrated that, our present method can significantly improve the prediction power by about 3.7% than previous models, thus providing an efficient and accurate approach for predicting protein-ligand interactions and uncover a new trend of using DL technique for massive biological structures training for drug design.

## 2 METHODS

### 2.1 Evaluation Metrics

We used Pearson correlation coefficient  $R$ , RMSE and Standard Deviation (SD) to evaluate the scoring power of the model which are defined as:

$$R = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2} \sqrt{\sum_i^n (y_i - \bar{y})^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2} \quad (2)$$

$$SD = \sqrt{\frac{\sum_i^n [y_i - (a + bx_i)]^2}{n - 1}} \quad (3)$$

where  $x_i$  is the predicted  $pK_d$  for  $i$ th complex;  $y_i$  is the experimental  $pK_d$  of this complex;  $\bar{x}$  and  $\bar{y}$  are the averages of all predicted values and experimental values;  $a$  and  $b$  are the intercept and the slope of the regression line, respectively Su et al. (2018).

### 2.2 Preparation of Dataset

We mainly used the protein-ligand complexes of PDBbind database v.2019 (<http://www.pdbbind-cn.org/>) for training. This database consists of two overlapping subsets, the general set and the refined set. The general set includes all available complexes and the refined set comprises protein-ligand complexes with high-quality structure and binding information selected from the general set. For each structure of the protein-ligand complex, the corresponding binding affinity is represented by the negative logarithms ( $pK_d$ ) of the dissociation constants ( $K_d$ ), inhibition constants ( $K_i$ ) or half inhibition concentrations ( $IC_{50}$ ). In order to evaluate the predictive ability and compare with other scoring functions, OnionNet-2 was evaluated on the CASF-2016 test set (core set v.2016) Su et al. (2018) and CASF-2013 test set (core set v.2013) Li et al. (2014a), Li et al. (2014b), Li et al. (2018). It should be noted that the CASF-2016 test set is the latest update of CASF-2016, which contains 285 high-quality complexes. While for core set v.2013, it is a subset of the PDBbind database v.2013, consisting of 195 protein-ligand complexes classified in 65 clusters with binding constants spanning nearly 10 orders of magnitude. Besides, a data set called CSAR NRC-HiQ, consisting of two subsets containing 176 and 167 complexes respectively, Dunbar et al. (2011) was employed as a third test set. For the previous models of Kdeep and RF-score, 55 and 49 complexes in two subsets were used as test data Jiménez et al. (2018). To provide a direct comparison with  $K_{deep}$  and RF-score, we adopted the same data for the OnionNet-2 test.

In order to perform normal training and testing, it is necessary to redistribute remaining complexes in PDBbind database v.2019. First, we excluded the complexes contained in three test sets from PDBbind database v.2019 (general set and refined set). Then, as a common practice (Reference: Pafnucy Stepniewska-Dziubinska et al. (2018) and OnionNet Zheng et al. (2019)), 1,000 complexes were randomly sampled from v.2019 refined set (after filtering all complexes used in the test sets described previously) as the

validating set. Finally, the remaining complexes (that is, the complexes that are not included in the three test sets and validating set) were adopted for the training set. This ensures that there is no overlapping protein-ligand complex in the training set, validating set and test sets.

### 2.3 Descriptors

The features employed are the pair numbers of the specific residue (protein)-atom (ligand) combination in multiple distance shells. The minimum distances between any atom in the ligand and any residue of protein are treated as the representative distances. First, around each atom in the ligand, we defined  $N$  continuously packed shells. The thickness of each shell is  $\delta$ , except that the first shell is a sphere with a radius of  $d_0$ . The boundary  $K_i$  of the  $i$ th shell is as follows

$$\begin{aligned} 0 < K_i < d_0, \quad i = 1 \\ d_0 + (i - 2)\delta \leq K_i < d_0 + (i - 1)\delta, \quad i \geq 2 \end{aligned}$$

Meanwhile, we classified atoms in the ligand into eight types, namely C, H, O, N, P, S, HAL, and DU, where HAL represents the halogen elements (F, Cl, Br, and I), and DU represents the element types excluded in these seven types.

$$T_e \in \{C, H, O, N, P, S, HAL, DU\}$$

When pre-processing the structure file, water and ions were treated explicitly because crystal water molecules and ions could affect the protein-ligand binding García-Sosa (2013); Spyraakis et al. (2017). In addition to the twenty standard residues, we added an expanded type named "OTH" to represent water, ions and any other non-standard residues.

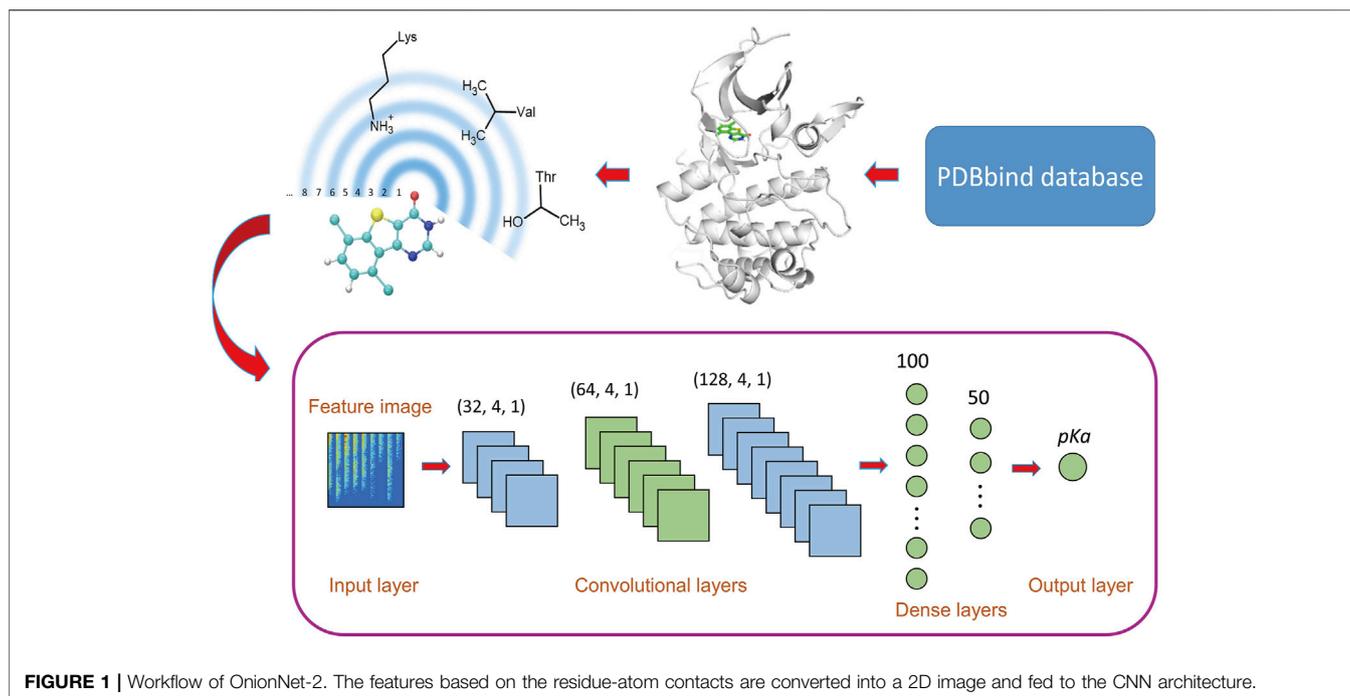
$$T_r \in \{GLY, ALA, VAL, LEU, ILE, PRO, PHE, TYR, TRP, SER, THR, CYS, MET, ASN, GLN, ASP, GLU, LYS, ARG, HIS, OTH\}$$

It is worth mention that the residue-atom distance is defined as the distance between the atom in the ligand and the nearest heavy atom in the residue. A 2D visual representation is depicted at the upper left of **Figure 1**. For any shell, the number of contacts for each residue-atom pair is calculated and used as a feature. Each shell has  $8 \times 21 = 168$  residue-atom combinations, which means that there are 168 features for a shell. Thus, if the total number of shells is  $N$ ,  $168 \times N$  features will be generated.

$$n_{T_r, T_e} = \sum_{r=1}^R \sum_{e=1}^E c_{r,e}^i \quad (4)$$

$$c_{r,e}^i = \begin{cases} 1, & (i - 2)\delta + d_0 \leq d_{r,e} < (i - 1)\delta + d_0 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

Here,  $R$  is the total number of residues in the protein, and  $E$  is the total number of atoms in the ligand. The  $d_{r,e}$  is the minimum distance between the residue  $r$  in the protein and the atom  $e$  in the ligand, and  $n_{T_r, T_e}^i$  is the number of contacts of the specific residue-element combination in the  $i$ th shell. The  $c_{r,e}^i$  is 1 when  $(i - 2)\delta + d_0 \leq d_{r,e} < (i - 1)\delta + d_0$ , otherwise  $c_{r,e}^i$  is 0. Following our previous study, Zheng et al. (2019) we used  $d_0 = 1 \text{ \AA}$  and  $\delta = 0.5 \text{ \AA}$ .



Interestingly such shell-like, or radial, representations of protein environments, have been demonstrated to be superior features in protein function prediction Torng and Altman (2019b). The source code of OnionNet-2 is available at <https://github.com/zchwang/OnionNet-2/>.

## 2.4 Architecture

We adopted a CNN model based on 2D convolution to learn the relationship between the contact features and the  $\Delta G$ . The model was constructed using the Keras package in tensorflow Abadi et al. (2016). The workflow architecture is shown in **Figure 1**.

The raw data is pre-processed before input into the CNN model. Here, the features are standardized through the scikit-learn package Pedregosa et al. (2011), and the processed features confirmed the standard normal distribution. Considering the big success of CNN model using 2D convolution in image recognition, Pak and Kim (2017) the 1D vector containing the protein-ligand interactions was converted into a 2D matrix to mimic the feature images, which was used as the input of CNN model.

CNNs usually consist of multiple layers with different functions, and the convolution layer is the key part of CNN models, which is to extract different features from input data Lavecchia (2019). The filter, also called the convolution kernel, is the core part of the convolutional layer, and the local features of the input “picture” are extracted through the sliding of the filter Gawehn et al. (2016); Angermueller et al. (2016). In the OnionNet-2 model, we used three convolutional layers, with 32, 64, and 128 filters respectively and the filter sizes were all set as 4, with strides as 1. The results of the last convolutional layer need to be flattened before being passed to the fully

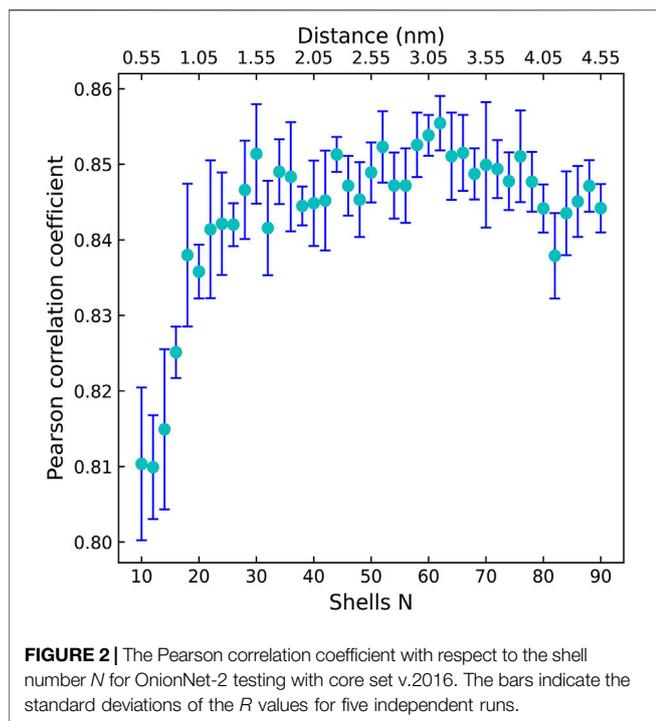
connected layers. The fully connected layers can integrate the local features extracted by the convolutional layers to give the prediction of  $pK_d$  value. Increasing the width and length of the fully connected layer can improve the complexity and nonlinear expression ability of the model, but in practice, this may lead to unexpected overfitting. In addition, increasing parameters will significantly increase the computational cost. After preliminary tests, two fully connected layers with 100 and 50 neurons are used before the output layer, which is capable of capturing the nonlinear relationship between the features and the  $pK_d$  values.

To further increase the nonlinear ability of the model, a rectified linear unit (RELU) layer was added after each convolutional layer and fully connected layer. Also, a batch normalization layer was used after the fully connected layer. The stochastic gradient descent (SGD) optimizer was adopted and the learning rate was set as 0.001. To reduce overfitting, L2 regularization with weight decay  $\lambda = 0.01$  was used after each fully connected layer. The number of samples processed per batch is 64.

To evaluate the performance of the OnionNet-2, we adopted the loss function defined in the previous work Zheng et al. (2019).

$$loss = \alpha(1 - R) + (1 - \alpha)RMSE \quad (6)$$

where  $R$  and  $RMSE$  represent Pearson correlation coefficient (**Eq. 1**) and root-mean-squared error (**Eq. 2**), respectively. The  $\alpha(0 \leq \alpha \leq 1)$  value is an adjustable factor for adjusting the weight with  $R$  and  $RMSE$ , which was finally set to 0.7. For each independent training task, we adopted early stopping (patience = 20, that is, if the change of the loss value in the validating set is less than 0.001 after 20 epochs, the training is terminated) and save the model that performed best on the validating set. For the



prediction in each case, five independently trainings were conducted to obtain the predicted mean value.

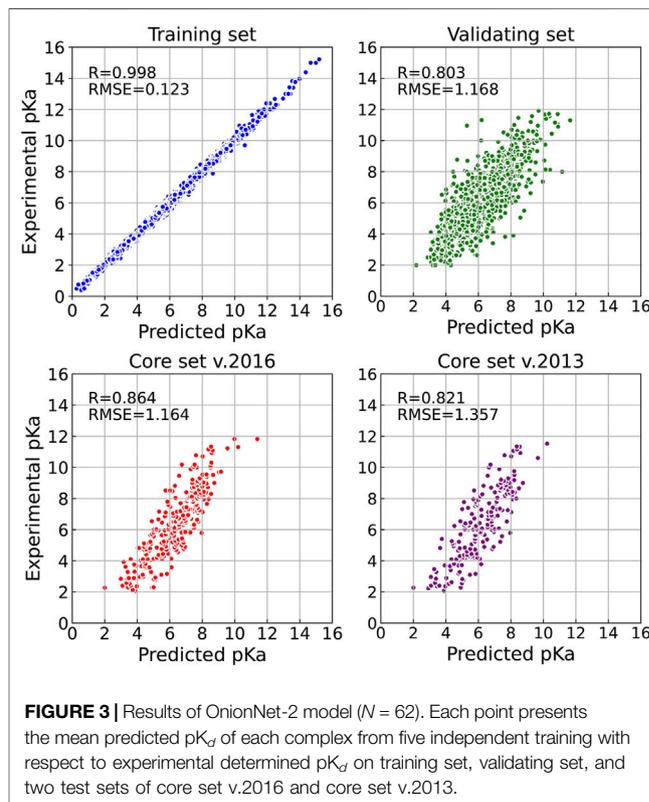
### 3 RESULTS AND DISCUSSIONS

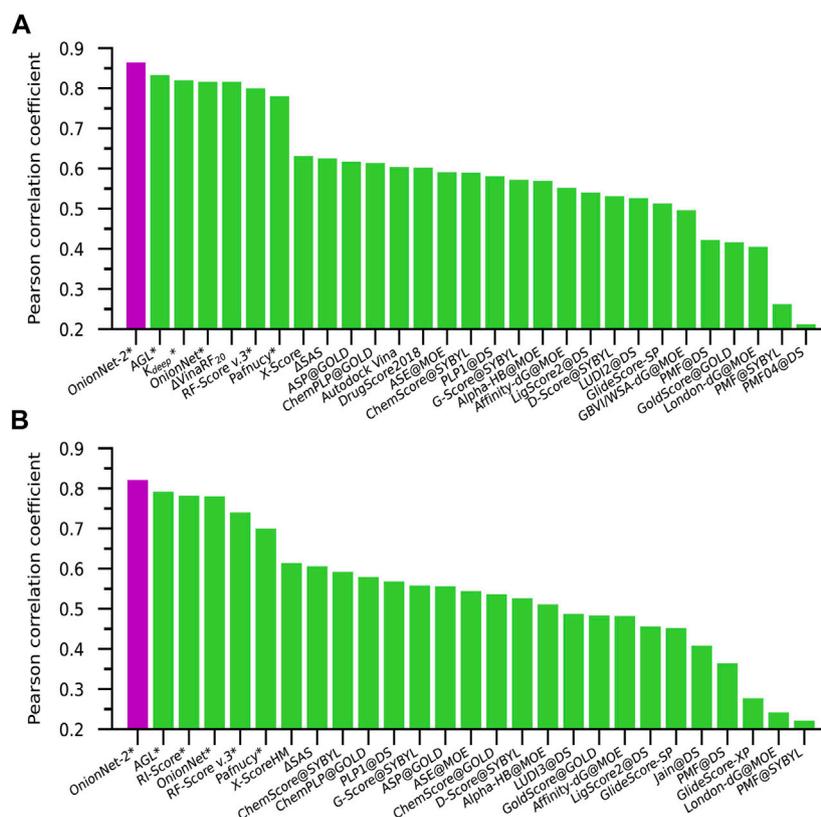
#### 3.1 The Predictive Power of OnionNet-2

Firstly, we explored the effect of shell number  $N$  on the predictive capability of the OnionNet-2 model. A range of the total shell number  $10 \leq N \leq 90$  was tested with interval of 2. According to our definitions of distance shell, this covers a separation between the residue and the atom from 0.55 to 4.55 nm. **Figure 2** depicts the trend of the  $R$  value to shell number  $N$  testing with core set v.2016. For  $N$  from 10 to 20, the  $R$  quickly increases as the total number of shells increases. This is expected because as the number of shells increases, the interactions between ligand and protein were gradually captured by the model. The  $R$  value reached the first peak for  $N$  is 30. This means that OnionNet-2 can achieve high prediction accuracy at a relatively low computational cost. Then,  $R$  fluctuates in a range of 0.01 until reaches the global maximum value when  $N = 62$ . **Figure 3** summarized the mean predicted value of each complex from five independent training, with respect to experimental value, using  $N = 62$ , on the training set, validating set and two testing sets, core set v.2016 and core set v.2013. It shows that the predicted  $pK_d$  and experimental  $pK_d$  are highly correlated for the two testing sets and validating set. After this point,  $R$  decreases when  $N$  increase. We attribute this to the enormous data that leads to the introduction of noise in the training. Unless otherwise specified, we adopted  $N$  of 62 in the following discussions. In addition, we also re-trained the

model with two elder versions (v.2016 and v.2018) of the PDBbind database, and the  $R$  values of our re-trained models are almost the same (**Supplementary Figure S1** and **Supplementary Table S1**).

The performance of some published scoring functions and OnionNet-2 tested on CASF-2016 and CASF-2013 are showed in **Figures 4A,B**, respectively. The corresponding  $R$  and RMSE (or SD) achieved by these representative scoring functions can be found in **Supplementary Table S2**. Firstly, our OnionNet-2 model achieved highest  $R$  of 0.864 and RMSE of 1.164 with the core set v.2016, and  $R = 0.821$  and RMSE = 1.357 with the core set v.2013. These were significantly higher than other scoring functions. The 2<sup>nd</sup> best scoring function was AGL, which adopted the gradient boosting trees (GBTs) algorithm, focusing on multiscale weighted labeled algebraic subgraphs to characterize protein-ligand interactions Nguyen and Wei (2019). For two 3D convolution-based scoring functions  $K_{deep}$  Jiménez et al. (2018) and Pafnucy Stepniewska-Dziubinska et al. (2018), they adopted 3D voxel representation to model the protein-ligand complex and explicitly treated with physical properties of atoms such as hydrophobic, hydrogen-bond donor or acceptor and aromatic etc. into consideration. It is interesting to find that although we only employed the residue-atom contact to mimic the interactions between the protein and the ligand in OnionNet-2, the predicting power is higher. This further reveals that the selected features have a great impact on the predictive power of the CNN-based scoring functions. Secondly, as is expected, the introduction of ML/DL





**FIGURE 4 |** Pearson correlation coefficient of different scoring functions on **(A)** CASF-2016 and **(B)** CASF-2013 benchmarks. The scoring functions marked with an asterisk are based on ML/DL models. The purple column is the performance of our proposed model.

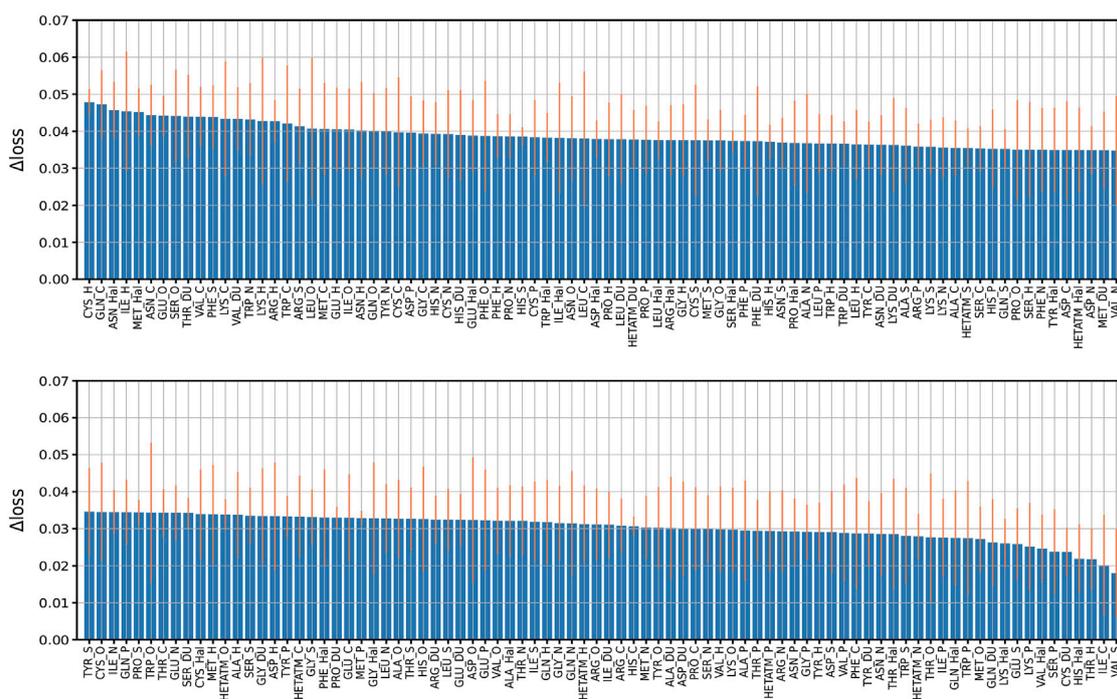
techniques into models has systematically enhanced the predicting accuracy.

In order to explore the feature importance of different residue-atom combinations to the total performance of the model, we valued the importance of combination features following our previous strategy. In detail, we re-trained the model with missing a certain residue-atom combination, then calculated the performance change in loss ( $\Delta$ loss). Here, the  $\Delta$ loss is defined as the difference between the loss of a model with missing a certain feature and the loss of the best model. Therefore, larger  $\Delta$ loss represents that this feature has higher importance. The results are summarized in **Figure 5**. The most important combination is “CYS\_H”, which is mainly due to the high occurrences of hydrogen bond between CYS and H atom of ligand molecules. Also, “CYS\_N” also showed relatively high importance because N atom of ligand molecule acts as the donor of hydrogen bonds. Besides, “ASN\_Hal” is also recognized to be an important feature to the protein-ligand binding affinity prediction, which may be attributed to the formation of the halogen bond between the halogen atom in the ligand and the O atom in the ASN Cavallo et al. (2016). Generally, it is worth mention that although we identified the different importance of the combinations, missing of any residue-atom combination indeed does not cause clear decreases in the overall performance of the model. For instance, most  $\Delta$ loss are in the

range of 0.02–0.04 except that missing of the first 23 combinations caused a  $\Delta$ loss near 0.05, indicating the high robustness of this model.

### 3.2 Evaluation of the Generalization Ability of the Model on Different Test Sets

Generally, DL models display a good generalization behavior in practical applications Neyshabur et al. (2017). To verify the generalization ability of the OnionNet-2, the CSAR NRC-HiQ data set provided by CSAR. Dunbar et al. (2011) was used as an additional test set in this study. This data set contains two subsets which contain 176 and 167 protein-ligand complexes, respectively. For the two previous ML models,  $K_{deep}$  and RF-Score, the researchers used 55 and 49 complexes in two subsets respectively as test data Jiménez et al. (2018). To provide a direct comparison with them, we adopted the same data for the OnionNet-2 test. It is worth mention that the two test subsets from the CSAR NRC-HiQ only have two common complexes with core set v.2013, namely 2jdy and 2qmj, and does not overlap with the training set, validation set and core set v.2016. The performance of  $K_{deep}$ , RF-Score and OnionNet-2 on these two subsets are shown in **Table 1**, and the scatter plots of the  $pK_d$  predicted by OnionNet-2 with respect to experimental  $pK_d$  can be found in **Supplementary Figure S1**.



**FIGURE 5 |** Performance change ( $\Delta$ loss) due to missing features. Performance change ( $\Delta$ loss) when a certain residue-atom combination is removed. The  $\Delta$ loss is defined as the difference between the loss of a model with missing a certain feature and the loss of the best model. The bars indicate the standard deviations of the  $\Delta$ loss for five independent runs.

**TABLE 1 |** The performance of OnionNet-2,  $K_{deep}$  and RF-Score achieved on subsets from CSAR NRC-HIQ data set.

	Subset 1		Subset 2	
	R	RMSE	R	RMSE
OnionNet-2	0.89	1.50	0.87	1.21
$K_{deep}$ Jiménez et al. (2018)	0.72	2.09	0.65	1.92
RF-Score Jiménez et al. (2018)	0.78	1.99	0.75	1.66

As expected, our model achieved a higher performance than  $K_{deep}$  and RF-score. For subset 1, the present OnionNet-2 achieved  $R$  of 0.89, which is considerably higher than that of  $K_{deep}$  (0.72) and RF-Score (0.78). This is also true for subset 2. Especially that, the  $R$  value of  $K_{deep}$  model is only 0.65 for subset 2, indicating weak predicting capability on these data. These results effectively demonstrated that OnionNet-2 has a good generalization ability.

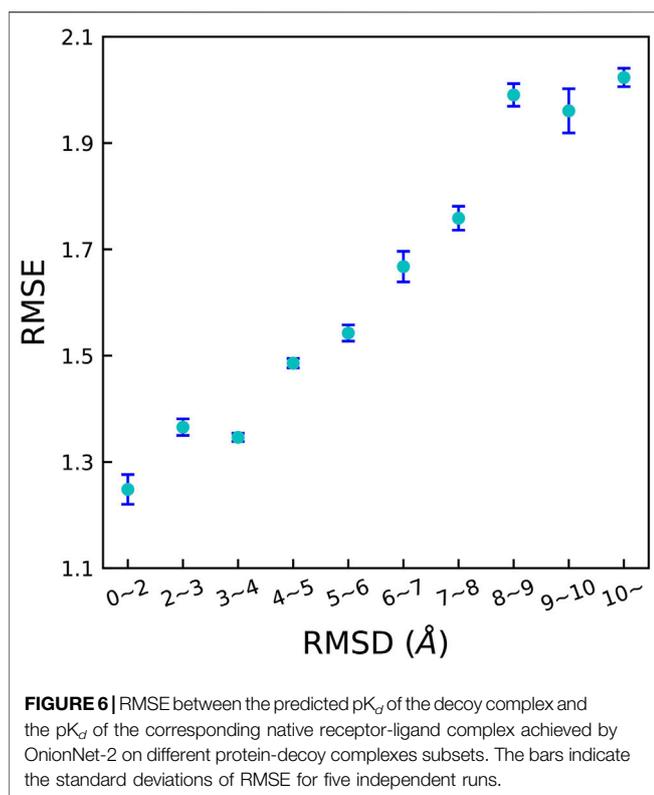
### 3.3 Evaluations on Subsets of Non-Experimental Decoy Structures

As all the training and validating sets are composed of well-validated native structures in previous studies, it is largely unknown whether the DL method is capable to distinguish “bad data” that are incorporated in these integrated data sets, for instance, non-native binding poses. To verify this, we tested OnionNet-2 to deal with non-experimental structures (generated

by docking programs). Technically, non-native binding poses (called decoys) were generated based on core set v.2016 complexes by AutoDock Vina Trott and Olson (2010); Forli et al. (2016). The detailed information of the generation of decoys can be found in the Supplementary Information. In CASF-2016 benchmark, the similarity between two binding poses is measured by the root-mean-square deviation (RMSD) value. Following previous studies by Allen et al. (2014), we adopted the Hungarian algorithm to calculate RMSD between decoy ligand and native structure which is implemented in spyrmsd Meli and Biggin (2020). The treatment of decoy was as following:

- 1) For each receptor, up to 20 decoy ligands were generated by AutoDock Vina. The actual number may be less than 20 because of limited size and shape of the binding pocket in the target protein. For each decoy, the RMSD with respect to native structure was calculated.
- 2) We used 10 RMSD intervals, [0 Å, 2 Å], [2 Å, 3 Å], [3 Å, 4 Å], ..., [9 Å, 10 Å] and [10 Å:].
- 3) For all ligands in every interval, we selected the decoy with the smallest RMSD value to put into the corresponding subsets. 4. Ten test subsets containing non-experimental complexes were used for OnionNet-2 training.

The predicting accuracy was evaluated by calculating the RMSE between the predicted  $pK_d$  of the decoy complex and the  $pK_d$  of the corresponding native receptor-ligand complex which is shown in **Figure 6**. It is clear that, the RMSE quickly



increased with increasing RMSD. This is expected because decoys with larger RMSD result in more severe change of  $\Delta G$ . These results reveal that OnionNet-2 can accurately respond to changes of the ligand binding poses and distinguish the native structure.

### 3.4 Effects of Hydrophobic Scale, Buried Solvent-Accessible Area and Excluded Volume Inside the Binding Pockets on the Prediction Accuracy

Principally, the physical interactions between protein and ligand determine the  $\Delta G$ . The dominating factors for overall  $\Delta G$  involve electrostatic interactions, van der Waals interactions, hydrogen bonds, hydration/de-hydration during complexation. However, such mechanistic interactions were not directly input into DL features. At molecular level, these involves the size and shape of the binding pocket, and the nature of residues around the binding pocket which determine its physicochemical characteristics Stank et al. (2016). However, whether DL models can accurately represent the structural specificity of the binding pocket is poorly documented.

The entire CASF-2016 test set can be divided into three subsets by each of three descriptors according to physical classifications of the binding pocket on the target protein Su et al. (2018). The three descriptors include H-scale (hydrophobic scale of the binding pocket),  $\Delta SAS$  (buried percentage of the solvent-accessible area of the ligand after binding) and  $\Delta VOL$  (excluded volume inside the binding pocket after ligand binding). Protein-ligand complexes in CASF-2016 were

grouped into 57 clusters, and the authors sorted all 57 clusters in ascending order by each descriptor. Then, these complex clusters were divided into three subsets according to the chosen descriptor, labeled as H1, H2, and H3 or S1, S2, and S3 or V1, V2, and V3. These subsets were used as validations of our OnionNet-2 model. As comparison, previous scoring functions were also tested on these three sets of subsets by Su et al. (2018), and the results are summarized in **Table 2**.

As can be seen in **Table 2**, OnionNet-2 achieved higher prediction accuracy compared with other scoring functions when tested on H, S, and V-series subsets. This indicates that the feature based on the contact number of residue-atom pairs in multiple shell is capable of capturing the hydrophobic scale of the binding pocket. The number of contacts in different shells (specifically the shells within the binding pocket) may be able to reflect the buried solvent-accessible surface area and the excluded volume of the ligand.

We noticed that, compared to other subsets, the R value of OnionNet-2 on V2 subset is clearly lower than other subsets (nevertheless it is still high than other scoring functions). This may indicate that our model is less sensitive to medium-sized binding pockets. Thus it may be still challenging for current scoring functions to recognize the size and shape of the binding pocket.

Furthermore, we plotted the detailed scatter plots of predicted  $pK_d$  and experimental  $pK_d$  in **Supplementary Figure S3** according to the specific H, S and V range. It is interesting to find almost no dependence of  $pK_d$  with the values of H, S, or V. Thus we speculate that a more realistic descriptor for the ligand characteristic in the binding pocket is essential to guide the protein-ligand  $\Delta G$  prediction.

### 3.5 Discussions of Screening Power

From the above results, we demonstrated that OnionNet-2 has high efficiency in treating with protein-ligand binding affinity prediction with simple calculations of structural features. However, we noticed a work from Hou group reporting that mostly developed ML models (including the OnionNet) performed poorly in the virtual screening tasks Shen et al. (2020b). In such screening power examinations, theoretical models have to pick true binders from a lot of false “decoy” molecules. It is not surprising that all ML models which were trained solely on true binders (for example, PDBbind databases) were not taught to distinguish decoy molecules from the true binders. In order to improve the screening power of the ML/DL based scoring functions, decoy molecules should be included in the training sets. Such work is now being undertaken by us.

## 4 CONCLUSION

To summarize, a 2D convolution-based CNN model, OnionNet-2, is proposed for prediction of the protein-ligand binding free energy. The contacting pair numbers between the protein residues and the ligand atoms were used as features for DL training. Using CASF-2013 and CASF-2016 as benchmarks, our model achieved the highest accuracy to predict  $\Delta G$  than

**TABLE 2** | Pearson correlation coefficients achieved by OnionNet-2 and previous scoring functions on three series of subsets.

	Subset H			Subset S			Subset V		
	H1	H2	H3	S1	S2	S3	V1	V2	V3
OnionNet-2	0.872	0.866	0.856	0.868	0.839	0.869	0.856	0.774	0.866
$\Delta_{vina}RF_{20}$	0.820	0.832	0.804	0.843	0.765	0.823	0.727	0.760	0.818
X-Score	0.698	0.570	0.661	0.743	0.536	0.572	0.437	0.622	0.579
X-Score <sup>HS</sup>	0.711	0.565	0.647	0.748	0.557	0.546	0.433	0.630	0.559
$\Delta$ SAS	0.641	0.643	0.589	0.746	0.572	0.494	0.480	0.669	0.541
X-Score <sup>HP</sup>	0.669	0.558	0.672	0.740	0.510	0.575	0.450	0.575	0.580
AutoDock Vina	0.626	0.586	0.641	0.745	0.521	0.507	0.484	0.564	0.486

Results of the last six rows taken from Su et al. (2018).

previous scoring functions. In addition, when employing different versions of PDBbind database for training, the performance of OnionNet-2 is nearly the same. We also evaluated the generalization ability of the model through testing on the CSAR NRC-HiQ data set and the decoys structures. Our result also indicates that OnionNet-2 has the capability to recognize these physical natures (in detail, hydrophobic scale of the binding pocket, buried percentage of the solvent-accessible area of the ligand upon binding and excluded volume inside the binding pocket upon ligand binding) of the ligand-binding pocket interaction. This systematic study also verified our initial hypothesis that is using the intrinsic types residues as the basic units can better characterize the physicochemical properties of the protein, which will be beneficial to improve the performance of the protein-ligand binding affinity prediction.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

## REFERENCES

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). "Tensorflow: A System for Large-Scale Machine Learning," in 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16), 265–283.
- Ain, Q. U., Aleksandrova, A., Roessler, F. D., and Ballester, P. J. (2015). Machine-learning Scoring Functions to Improve Structure-Based Binding Affinity Prediction and Virtual Screening. *Wires Comput. Mol. Sci.* 5, 405–424. doi:10.1002/wcms.1225
- Allen, W. J., and Rizzo, R. C. (2014). Implementation of the Hungarian Algorithm to Account for Ligand Symmetry and Similarity in Structure-Based Design. *J. Chem. Inf. Model.* 54, 518–529. doi:10.1021/ci400534h
- Angermueller, C., Pärnamaa, T., Parts, L., and Stegle, O. (2016). Deep Learning for Computational Biology. *Mol. Syst. Biol.* 12, 878. doi:10.15252/msb.20156651
- Ballester, P. J., and Mitchell, J. B. O. (2010). A Machine Learning Approach to Predicting Protein-Ligand Binding Affinity with Applications to Molecular Docking. *Bioinformatics* 26, 1169–1175. doi:10.1093/bioinformatics/btq112
- Cavallo, G., Metrangolo, P., Milani, R., Pilati, T., Priimagi, A., Resnati, G., et al. (2016). The Halogen Bond. *Chem. Rev.* 116, 2478–2601. doi:10.1021/acs.chemrev.5b00484

## AUTHOR CONTRIBUTIONS

ZW performed experiments, analyzed data and wrote the article. LZ, WL, and YM designed the experiment and analyzed data. YL, YQ, Y-QL, and MZ analyzed data. ZW, LZ, WL, and YM wrote the article.

## FUNDING

This work is supported by the Natural Science Foundation of Shandong Province (ZR2020JQ04), National Natural Science Foundation of China (11874238) and Singapore MOE Tier 1 Grant RG146/17. This work is also supported by the Ministry of Education, Singapore, under its Academic Research Fund Tier 2, MOE-T2EP30120-0007.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2021.753002/full#supplementary-material>

- Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., and Blaschke, T. (2018). The Rise of Deep Learning in Drug Discovery. *Drug Discov. Today* 23, 1241–1250. doi:10.1016/j.drudis.2018.01.039
- Du, X., Li, Y., Xia, Y.-L., Ai, S.-M., Liang, J., Sang, P., et al. (2016). Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Ijms* 17, 144. doi:10.3390/ijms17020144
- Dunbar, J. B., Jr, Smith, R. D., Yang, C.-Y., Ung, P. M.-U., Lexa, K. W., Khazanov, N. A., et al. (2011). CSAR Benchmark Exercise of 2010: Selection of the Protein-Ligand Complexes. *J. Chem. Inf. Model.* 51, 2036–2046. doi:10.1021/ci200082t
- Durrant, J. D., and McCammon, J. A. (2010). NNScore: A Neural-Network-Based Scoring Function for the Characterization of Protein-Ligand Complexes. *J. Chem. Inf. Model.* 50, 1865–1871. doi:10.1021/ci100244v
- Ellingson, S. R., Davis, B., and Allen, J. (2020). Machine Learning and Ligand Binding Predictions: a Review of Data, Methods, and Obstacles. *Biochim. Biophys. Acta (Bba) - Gen. Subjects* 1864, 129545. doi:10.1016/j.bbagen.2020.129545
- Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., and Olson, A. J. (2016). Computational Protein-Ligand Docking and Virtual Drug Screening with the AutoDock Suite. *Nat. Protoc.* 11, 905–919. doi:10.1038/nprot.2016.051
- García-Sosa, A. T. (2013). Hydration Properties of Ligands and Drugs in Protein Binding Sites: Tightly-Bound, Bridging Water Molecules and Their Effects and Consequences on Molecular Design Strategies. *J. Chem. Inf. Model.* 53, 1388–1405. doi:10.1021/ci3005786

- Gawehn, E., Hiss, J. A., and Schneider, G. (2016). Deep Learning in Drug Discovery. *Mol. Inf.* 35, 3–14. doi:10.1002/minf.201501008
- Ghasemi, F., Mehridehnavi, A., Pérez-Garrido, A., and Pérez-Sánchez, H. (2018). Neural Network and Deep-Learning Algorithms Used in Qsar Studies: Merits and Drawbacks. *Drug Discov. Today* 23, 1784–1790. doi:10.1016/j.drudis.2018.06.016
- Gilson, M. K., and Zhou, H.-X. (2007). Calculation of Protein-Ligand Binding Affinities. *Annu. Rev. Biophys. Biomol. Struct.* 36, 21–42. doi:10.1146/annurev.biophys.36.040306.132550
- Gomes, J., Ramsundar, B., Feinberg, E. N., and Pande, V. S. (2017). *Atomic Convolutional Networks for Predicting Protein-Ligand Binding Affinity*. arXiv preprint arXiv:1703.10603.
- Grinter, S., and Zou, X. (2014). Challenges, Applications, and Recent Advances of Protein-Ligand Docking in Structure-Based Drug Design. *Molecules* 19, 10150–10176. doi:10.3390/molecules190710150
- Guedes, I. A., Pereira, F. S. S., and Dardenne, L. E. (2018). Empirical Scoring Functions for Structure-Based Virtual Screening: Applications, Critical Aspects, and Challenges. *Front. Pharmacol.* 9, 1089. doi:10.3389/fphar.2018.01089
- Guvench, O., and MacKerell, A. D., Jr (2009). Computational Evaluation of Protein-Small Molecule Binding. *Curr. Opin. Struct. Biol.* 19, 56–61. doi:10.1016/j.sbi.2008.11.009
- Hansen, N., and Van Gunsteren, W. F. (2014). Practical Aspects of Free-Energy Calculations: a Review. *J. Chem. Theor. Comput.* 10, 2632–2647. doi:10.1021/ct500161f
- Heck, G. S., Pintro, V. O., Pereira, R. R., de Ávila, M. B., Levin, N. M. B., and de Azevedo, W. F. (2017). Supervised Machine Learning Methods Applied to Predict Ligand- Binding Affinity. *Curr. Med. Chem.* 24, 2459–2470. doi:10.2174/0929867324666170623092503
- Huang, N., Kalyanaraman, C., Irwin, J. J., and Jacobson, M. P. (2006). Physics-Based Scoring of Protein-Ligand Complexes: Enrichment of Known Inhibitors in Large-Scale Virtual Screening. *J. Chem. Inf. Model.* 46, 243–253. doi:10.1021/ci0502855
- Huang, S.-Y., Grinter, S. Z., and Zou, X. (2010). Scoring Functions and Their Evaluation Methods for Protein-Ligand Docking: Recent Advances and Future Directions. *Phys. Chem. Chem. Phys.* 12, 12899–12908. doi:10.1039/c0cp00151a
- Huey, R., Morris, G. M., Olson, A. J., and Goodsell, D. S. (2007). A Semiempirical Free Energy Force Field with Charge-Based Desolvation. *J. Comput. Chem.* 28, 1145–1152. doi:10.1002/jcc.20634
- Jiménez, J., Doerr, S., Martínez-Rosell, G., Rose, A. S., and De Fabritiis, G. (2017). DeepSite: Protein-Binding Site Predictor Using 3d-Convolutional Neural Networks. *Bioinformatics* 33, 3036–3042. doi:10.1093/bioinformatics/btx350
- Jiménez, J., Škalič, M., Martínez-Rosell, G., and De Fabritiis, G. (2018). KDEEP: Protein-Ligand Absolute Binding Affinity Prediction via 3D-Convolutional Neural Networks. *J. Chem. Inf. Model.* 58, 287–296. doi:10.1021/acs.jcim.7b00650
- Lavecchia, A. (2019). Deep Learning in Drug Discovery: Opportunities, Challenges and Future Prospects. *Drug Discov. Today* 24, 2017–2032. doi:10.1016/j.drudis.2019.07.006
- Lavecchia, A. (2015). Machine-learning Approaches in Drug Discovery: Methods and Applications. *Drug Discov. Today* 20, 318–331. doi:10.1016/j.drudis.2014.10.012
- Li, Y., Han, L., Liu, Z., and Wang, R. (2014a). Comparative Assessment of Scoring Functions on an Updated Benchmark: 2. Evaluation Methods and General Results. *J. Chem. Inf. Model.* 54, 1717–1736. doi:10.1021/ci500081m
- Li, Y., Liu, Z., Li, J., Han, L., Liu, J., Zhao, Z., et al. (2014b). Comparative Assessment of Scoring Functions on an Updated Benchmark: 1. Compilation of the Test Set. *J. Chem. Inf. Model.* 54, 1700–1716. doi:10.1021/ci500080q
- Li, Y., Su, M., Liu, Z., Li, J., Liu, J., Han, L., et al. (2018). Assessing Protein-Ligand Interaction Scoring Functions with the CASF-2013 Benchmark. *Nat. Protoc.* 13, 666–680. doi:10.1038/nprot.2017.114
- Liu, J., and Wang, R. (2015). Classification of Current Scoring Functions. *J. Chem. Inf. Model.* 55, 475–482. doi:10.1021/ci500731a
- Lo, Y.-C., Rensi, S. E., Torng, W., and Altman, R. B. (2018). Machine Learning in Chemoinformatics and Drug Discovery. *Drug Discov. Today* 23, 1538–1546. doi:10.1016/j.drudis.2018.05.010
- Meli, R., and Biggin, P. C. (2020). Spyrmsd: Symmetry-Corrected Rmsd Calculations in python. *J. Cheminform.* 12, 1–7. doi:10.1186/s13321-020-00455-2
- Michel, J., and Essex, J. W. (2010). Prediction of Protein-Ligand Binding Affinity by Free Energy Simulations: Assumptions, Pitfalls and Expectations. *J. Comput. Aided Mol. Des.* 24, 639–658. doi:10.1007/s10822-010-9363-3
- Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., et al. (1998). Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* 19, 1639–1662. doi:10.1002/(sici)1096-987x(19981115)19:14<1639:aid-jcc10>3.0.co;2-b
- Morrone, J. A., Weber, J. K., Huynh, T., Luo, H., and Cornell, W. D. (2020). Combining Docking Pose Rank and Structure with Deep Learning Improves Protein-Ligand Binding Mode Prediction over a Baseline Docking Approach. *J. Chem. Inf. Model.* 60, 4170–4179. doi:10.1021/acs.jcim.9b00927
- Neyshabur, B., Bhojanapalli, S., McAllester, D., and Srebro, N. (2017). *Exploring Generalization in Deep Learning*. arXiv preprint arXiv:1706.08947.
- Nguyen, D. D., and Wei, G.-W. (2019). AGL-score: Algebraic Graph Learning Score for Protein-Ligand Binding Scoring, Ranking, Docking, and Screening. *J. Chem. Inf. Model.* 59, 3291–3304. doi:10.1021/acs.jcim.9b00334
- Öztürk, H., Özgür, A., and Ozkirimli, E. (2018). DeepDTA: Deep Drug-Target Binding Affinity Prediction. *Bioinformatics* 34, i821–i829. doi:10.1093/bioinformatics/bty593
- Pak, M., and Kim, S. (2017). “A Review of Deep Learning in Image Recognition,” in 2017 4th international conference on computer applications and information processing technology (CAIPT) (IEEE), 1–3. doi:10.1109/caipt.2017.8320684
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in python. *J. Machine Learn. Res.* 12, 2825–2830.
- Shen, C., Ding, J., Wang, Z., Cao, D., Ding, X., and Hou, T. (2020a). From Machine Learning to Deep Learning: Advances in Scoring Functions for Protein-Ligand Docking. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 10, e1429. doi:10.1002/wcms.1429
- Shen, C., Hu, Y., Wang, Z., Zhang, X., Pang, J., Wang, G., et al. (2020b). Beware of the Generic Machine Learning-Based Scoring Functions in Structure-Based Virtual Screening. *Brief. Bioinform.* 22, bbaa070. doi:10.1093/bib/bbaa070
- Song, T., Wang, S., Liu, D., Ding, M., Du, Z., Zhong, Y., et al. (2020). Se-onionnet: A Convolution Neural Network for Protein-Ligand Binding Affinity Prediction. *Front. Genet.* 11, 1805.
- Spyrakis, F., Ahmed, M. H., Bayden, A. S., Cozzini, P., Mozzarelli, A., and Kellogg, G. E. (2017). The Roles of Water in the Protein Matrix: a Largely Untapped Resource for Drug Discovery. *J. Med. Chem.* 60, 6781–6827. doi:10.1021/acs.jmedchem.7b00057
- Stank, A., Kokh, D. B., Fuller, J. C., and Wade, R. C. (2016). Protein Binding Pocket Dynamics. *Acc. Chem. Res.* 49, 809–815. doi:10.1021/acs.accounts.5b00516
- Stepniewska-Dziubinska, M. M., Zielenkiewicz, P., and Siedlecki, P. (2018). Development and Evaluation of a Deep Learning Model for Protein-Ligand Binding Affinity Prediction. *Bioinformatics* 34, 3666–3674. doi:10.1093/bioinformatics/bty374
- Su, M., Yang, Q., Du, Y., Feng, G., Liu, Z., Li, Y., et al. (2018). Comparative Assessment of Scoring Functions: the Casf-2016 Update. *J. Chem. Inf. Model.* 59, 895–913. doi:10.1021/acs.jcim.8b00545
- Torng, W., and Altman, R. B. (2019a). Graph Convolutional Neural Networks for Predicting Drug-Target Interactions. *J. Chem. Inf. Model.* 59, 4131–4149. doi:10.1021/acs.jcim.9b00628
- Torng, W., and Altman, R. B. (2019b). High Precision Protein Functional Site Detection Using 3d Convolutional Neural Networks. *Bioinformatics* 35, 1503–1512. doi:10.1093/bioinformatics/bty813
- Trott, O., and Olson, A. J. (2010). Autodock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* 31, 455–461. doi:10.1002/jcc.21334
- Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., et al. (2019). Applications of Machine Learning in Drug Discovery and Development. *Nat. Rev. Drug Discov.* 18, 463–477. doi:10.1038/s41573-019-0024-5
- Wang, R., Lai, L., and Wang, S. (2002). Further Development and Validation of Empirical Scoring Functions for Structure-Based Binding Affinity Prediction. *J. Computer-Aided Mol. Des.* 16, 11–26. doi:10.1023/a:1016357811882
- Wang, W., and Gao, X. (2019). *Deep Learning in Bioinformatics*.
- Winter, R., Montanari, F., Noé, F., and Clevert, D. A. (2019). Learning Continuous and Data-Driven Molecular Descriptors by Translating Equivalent Chemical Representations. *Chem. Sci.* 10, 1692–1701. doi:10.1039/c8sc04175j

- Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., et al. (2019). Analyzing Learned Molecular Representations for Property Prediction. *J. Chem. Inf. Model.* 59, 3370–3388. doi:10.1021/acs.jcim.9b00237
- Zhao, X., Liu, X., Wang, Y., Chen, Z., Kang, L., Zhang, H., et al. (2008). An Improved Pmf Scoring Function for Universally Predicting the Interactions of a Ligand with Protein, Dna, and Rna. *J. Chem. Inf. Model.* 48, 1438–1447. doi:10.1021/ci7004719
- Zheng, L., Fan, J., and Mu, Y. (2019). OnionNet: a Multiple-Layer Intermolecular-Contact-Based Convolutional Neural Network for Protein-Ligand Binding Affinity Prediction. *ACS Omega* 4, 15956–15965. doi:10.1021/acsomega.9b01997

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wang, Zheng, Liu, Qu, Li, Zhao, Mu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.