**Convolutional Neural Networks for Protein-Ligand Scoring**

Matt Ragoza1,2, Elisa Idrobo1,6, Joshua Hochuli2,4, Jocelyn Sunseri5 and David Koes5

1Department of Neuroscience, 2Department of Computer Science, 3TECBio REU @ Pitt, 4Department of Biological Sciences, 5Dept. of Computational and Systems Biology, University of Pittsburgh, PA 15260,  
6The College of New Jersey, Ewing, NJ 08648

**Abstract**

Computational approaches to drug discovery reduce the time and cost associated with experimental assays and enable the screening of novel chemical space. Structure-based drug design methods rely on scoring functions to rank and predict binding affinities and poses. The ever-expanding amount of protein-ligand binding and structural data enables deep machine learning techniques for protein-ligand scoring.

We describe a convolutional neural network (CNN) scoring function that takes as input a comprehensive 3D representation of a protein-ligand interaction. A CNN scoring function automatically learns the key features of protein-ligand interactions that determine binding. We train and optimize our CNN scoring functions to discriminate between correct and incorrect binding poses and known binders and non-binders.

We find that our CNN scoring function outperforms the AutoDock Vina scoring function when ranking poses both for pose prediction and virtual screening.

**Methods**

**Protein-Ligand Scoring**

Protein-ligand scoring provides a metric of binding strength between small molecules and target proteins. This has a wide array of uses such as within virtual libraries that filters large databases of candidate molecules for potential hits, and docking, which predicts the binding pose of a bound.

**Machine Learning**

Machine learning strategies have been used to treat scoring as a classification problem between “good” and “bad” binding states, though this often requires manually selecting properties that the model uses for discrimination, for example pairwise interactions and global counts of typical chemical interactions. However, other machine learning models can learn the most important features directly from data.

**Neural Networks**

Neural networks are a supervised machine learning algorithm inspired by the nervous system. A basic neural network consists of an input layer, one or more hidden layers, and an output layer of interconnected nodes. Each hidden node computes a feature that is a function of the weighted input it receives from the nodes of the previous layer.

**Input Data**

Input data are fed through the network, and a prediction is output by the last layer. A neural network is trained by iteratively adjusting the weights by minimization of an objective function, for example, the mean squared deviation between predictions and their ground truth labels.

**GPU Accelerated**

Within the last decade convolutional neural networks have become the state-of-the-art in image classification. Convolutional layers only have connection weights to small spatial subsets of the previous layer, and apply these weights to features across the entire input to produce feature maps.

The fact that convolutional layers learn local features and apply them across the entire input space leads to faster training and improved accuracy on data with a spatial structure.

The rise of GPU computing in combination with other advancements has made training networks with many more layers feasible, leading to the surge of research in deep learning.

Each successive layer in a deep neural network extracts features at a higher level of abstraction.

**D U D - E**

**Model Evaluation**

The CNN model was trained on residue-level features (80% split) and validated on residue-level features (20% split) across the entire input to produce feature maps.

The evaluation of the CNN model can be achieved by training and testing on data with a spatial structure.

**Conclusion**

We have shown that a convolutional neural network with a well-optimized model and appropriate training dataset has great potential in aiding with drug discovery. Creating variety in the poses through rotation, translation and shuffling during training are important in training the model. Other parameters such as network depth and width can also reduce overfitting. Visualizations highlight the pose sensitivity of the CNN model and can emphasize regions of interest in the protein-ligand complex.

Our best model performs better than AutoDock Vina at pose selection when evaluated for pose prediction performance (CCIA) and virtual screening performance (DUD-E), although the nature of the training data greatly influences the results.

---

**Visualizations**

**Ligand binding prediction with DUD-E**

**Optimization**

**Future Work**

**Evaluation of alternative network topologies, such as residual neural networks**

**Conclusion**

We have shown that a convolutional neural network with a well-optimized model and appropriate training dataset has great potential in aiding with drug discovery. Creating variety in the poses through rotation, translation and shuffling during training are important in training the model. Other parameters such as network depth and width can also reduce overfitting. Visualizations highlight the pose sensitivity of the CNN model and can emphasize regions of interest in the protein-ligand complex.

Our best model performs better than AutoDock Vina at pose selection when evaluated for pose prediction performance (CCIA) and virtual screening performance (DUD-E), although the nature of the training data greatly influences the results.