DOI: https://doi.org/10.31351/vol35iss2pp209-226

AI-Based Approach To Screen Opioid Receptor Ligands Via Grind Fingerprints and Convolutional Neural Network

Mohammed N. Al-Qattan*,[1]

[1]Department of Pharmaceutical Chemistry, College of Pharmacy, Ninevah University, Mosul, Iraq.
***Corresponding author**
Received 3/6/2025, Accepted 2/1/2026, Published 24/6/2026

Abstract

Opioid receptors belong to the G-Protein Coupled Receptor (GPCR) family and are divided into 3 subtypes Mu, Kappa, and Delta. Ligands for opioid receptors vary in size, chemical scaffold, and binding (orthosteric/allosteric) sites. Prediction of ligand activities using neural networks involves molecular descriptor encoding. In this research, grid-based alignment-independent molecular fingerprint (GRIND) and convolutional neural network (CNN) models are used to classify a database of ligands of experimentally measured opioid activities. The GRIND-CNN model outperforms The Extended-Connectivity Fingerprints with deep neural network (ECFP-DNN) in terms of activity-relevant feature recognition, while being almost equivalent to the PubChem fingerprint with deep neural network (PubChem-DNN). By combining the predictions of both GRIND-CNN and PubChem-DNN models, the prediction performance is further improved. Saliency maps of GRIND fingerprints for models trained to differentiate agonists vs antagonists, and differentiate ligands vs non-ligands for different opioid receptor subtypes showed statically significant differences at the correlogram bins (pixels) level. These differences are consistent with the message-address theory of opioid ligands. The models are freely available on https://github.com/mohammednooraldeen/GRIND-CNN.

Keywords: Alignment-independent GRIND fingerprint, Opioids virtual screening, Convolutional, Neural network, Artificial intelligence

Introduction

The GPCRs are flexible trans-membranal receptors that mediate multiple intracellular signaling pathways [(1)]. Although GPCRs represent the primary target of 34% of FDA-approved drugs, more than 220 non-olfactory GPCRs have disease associations that are yet to be explored in clinical research [(2, 3)].

Opioid receptors (Mu, Kappa, and Delta) are G-protein-coupled receptors (GPCRs) that play a vital role in pain modulation, mood regulation, and numerous physiological processes. The discovery of opioid receptor agonists and antagonists has significantly impacted medicine, particularly in pain management, addiction treatment, and neuropharmacology [(4)].

The prediction of ligands' activities against opioid receptors is useful in virtual screening for agonists and antagonists for potential therapeutic uses. The use of a ligand-based QSAR approach is a promising alternative to molecular docking concerning GPCRs [(5-7)], since the former technique overcomes the obstacles of receptor flexibility, selection between orthosteric and/or allosteric pockets, internal cooperativity, and multimerization [(8)]. The main scope of traditional QSAR is to formulate multiple linear regression terms that describe the correlation between molecular descriptors (encodings) and biological activities.

In QSAR models, molecular structures are encoded into 1D, 2D, 3D, or higher dimensionality fingerprints [(9)]. The common 2D molecular fingerprints include PubChem and ECFB formats. The PubChem fingerprint is a binary vector of 881 bits which encodes the presence/absence of element count, a type of ring system, atom pairing, atom environment (nearest neighbors), etc. (ftp://ftp.ncbi.nlm.nih.gov/pubchem/specifications/). The Extended-Connectivity Fingerprint (ECFP) is a circular topological fingerprint that involves iterated circular analysis of the molecular substructures and connectivity. Unlike PubChem, the ECFP bits representations are not predefined and can represent an essentially infinite number of different molecular features (including stereochemical information) [(10)]. The Grid-Independent Molecular Descriptor (GRIND) is considered a 3D fingerprint and provides a powerful approach for representing molecules with diverse scaffolds irrespective of molecular substructures and conformational alignment [(11, 12)]. GRIND encodes molecular features by summarizing the interaction fields of various chemical probes into discrete

hotspots. Pairwise distances between relevant probe otspots are then calculated and compiled into frequency distributions, capturing key spatial relationships independent of molecular alignment.

Recently, artificial neural networks have been found to outperform multiple linear regression in correlating molecular fingerprints to activities, especially for GPCRs [13-15]. Convolutional Neural Networks (CNNs) are a class of deep NN (DNN) that has emerged as a powerful tool in the field of QSAR and virtual screening for medicinal molecules. Similar to DNN, CNN uses molecular fingerprints and activities for training. However, unlike DNN, CNN can capture spatial correlations that are most relevant to activities even if it is shifted within 2D or 3D fingerprints. CNN is different from DNN in using convolutional layers which can extract specific features that can be related to the measurement of interest and thus reduce noise dependence compared to DNN [16]. Therefore, CNN can provide better performance in 1D-[17], 2D- [18, 19] and 3D-QSAR studies [20] and potentially outperforms traditional docking methods in virtual screening [21].

The GRIND fingerprint is frequently correlated to activity using PCA [22, 23] or PLS analysis [24-26], and is shown to be suitable for virtual screening [11] . Few reports are available on using MIFs as inputs for CNN models [27], however, no report for using GRIND yet. In this research, GRIND is reshaped as a 2D molecular fingerprint that is correlated with agonistic/antagonistic activity against Mu, Kappa, and Delta opioid receptors using CNN (GRIND-CNN). Other models of, PubChem-DNN and ECFP-DNN were also developed and used for comparison.

Materials and Methods

All scripts used are written using Python 3.9.10 in Pycharm Professional Edition 2022.3. All scripts and data are available on https://github.com/mohammednooraldeen/GRIND-CNN.

1. Datasets

The dataset used in the research was obtained from (https://tripod.nih.gov/tox/oprpred/data.zip) which includes 2905 molecules from NCATS Pharmaceutical Collection (NPC) of approved and investigational drugs. The activities were provided as 3 classes: inactive, inconclusive, and active, which assigned values of 0, 1, and 2, respectively. All compounds of inconclusive activity were excluded in this study.

2. GRIND molecular fingerprints

The molecules were generated from smiles code and conformationally optimized using MMFF94 force field in RDKit [28]. The alignment-independent molecular fingerprint (GRIND) was generated using Pentacle software (version 1.06) [11]. The probes used were DRY, O, N1 and TIP over grid step of 0.5 Å with default energy cutoff values of -0.5, -2.6, -4.2 and -0.75 Kcal/Mol. The Amanda discretization and MACC2 encoding methods were used [11]. The generated finger print is composed of 1870 floating values distributed as 187 bins frequencies for each of the 10 probe-probe MIFs distances (i.e. Dry-Dry, O-O, N1-N1, TIP-TIP, Dry-O, Dry-N1, Dry-TIP, O-N1, O-TIP and N1-TIP).

3. Setup and training of neural networks

The CNN is composed of an input layer with the number of neurons equal to the size of the descriptor vector, two sequential convolution layers of 3 kernels, each with the size of 10x10 neurons, a dense layer of 10 neurons, and an output layer of a single neuron. The activation functions used are leaky_relu for the convolutional and dense layers, while the output layer uses the sigmoidal function. Both convolutional layers and deep layers used a dropout rate of 0.15 and kernel L2 regularizer of 0.1 and 0.3, respectively. Both convolutional layers use kernel initializer of he_normal and end with max pooling layer of (1,4). The batch normalization layer is used between convolutional and deep layers.

The DNN network is composed of an input layer equal to the size of a fingerprint, three dense layers of 250, 90, and 20 neurons, respectively, and an output layer of one neuron. The dense layers use the leaky_relu activation function and kernel L2 regularizer of 0.03, and dropout rate of 0.25. Batch normalization layers are used between dense layers. The network architectures are shown in (Figure 1) while trainable parameters are listed in Table 1 and Table 2.

The networks are trained using Adam optimizer, and binary cross entropy as loss function. The training is done using a batch of 5 ligands per epoch and continued for 200 epochs unless the loss function score is not improved for 20 successive epochs in such case best model weights are restored.

4. Evaluation of models

Since the dataset is biased toward inactive ligands, the models are trained and validated using an under-sampling stratified five-fold cross-validation technique [29, 30]. The technique involves selecting a balanced subset of the dataset by randomly selecting samples to achieve an active-to-inactive ligands ratio of 1:2. This subset is then partitioned into five folds, each maintaining the same class ratio. In each training process, four folds are combined and used for training, while the remaining fold is reserved for validation. This process is repeated five times, with each fold serving once as the validation set, and then the averaged performance is measured. The five-fold cross-validation is iterated 20 times to cover almost all active and inactive ligands in the dataset.
The measurements used to estimate model performance include mean ± standard deviation for area under the curve of receiver operating characteristic (AUC-ROC) as well as Balanced

Accuracy (BA) and Matthews Correlation Coefficient (MCC) calculated from True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) predictions [30, 31] as shown in following equations:

$$sensitivity = \frac{TP}{TP + FN}$$

$$specificity = \frac{TN}{TN + FP}$$

$$BA = \frac{sensitivity + specificity}{2}$$

$$MCC = \frac{(TP \times TN) - (FP - FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

To exclude the possibility of correlation development by noise dependence, the Y-randomization technique is used. The technique involves the use of randomly shuffled activities for molecular fingerprints. The performance measurements of trained the models are compared with those of models trained on original data [32].

5. Saliency maps for GRIND fingerprint

The relevant hot regions within the GRIND fingerprint during the training of CNN models are shown as saliency maps. The saliency map was computed by taking the gradient of the trained model's output with respect to the input fingerprint pixels. During 5-fold cross-validation training, the saliency map was calculated at the end of each run. The mean ± standard deviation for accumulated 100 saliency maps (for 20 iterations) are then calculated. The calculation was done for six models trained on agonists and antagonists with respect to each of the Mu, Kappa, and Delta receptor subtypes. Similar calculations were also done for models trained on individual receptor ligands (i.e. combined agonists and antagonists) and for models trained on general agonists and antagonists (i.e. combined for all receptors). The t-test is used to allocate pixels that show statistically significant differences of potential relevancy to the address-message theory of opioid ligands.

Results and Discussion

The dataset used to train the models includes molecules that are experimentally classified as active, inconclusive (i.e. intermediate), and inactive when tested for agonistic or antagonistic activities against the Mu, Kappa, and Delta opioid receptors. The total molecules used in this study include 2723 molecules out of 2905 molecules in the original dataset, due to incompatibility with Pentacle software. The dataset is biased toward inactive ligands as shown in Figure 2. The ratio of active: inactive molecules for each of Mu, Kappa, and Delta agonists is about 0.19, 0.06, and 0.03, respectively, while for antagonists is about 0.06, 0.15, and 0.08, respectively. According to molecular weight distribution, most of the active molecules are in 200 to 400 Daltons. Some molecules show cross-reactivity among opioid receptors as shown in Figure 3. With respect to multi-receptor agonistic activity, 27 molecules are active and 1997 are inactive across all opioid receptors. While, for multi-receptor antagonistic activity, 81 molecules are active, while 1975 are inactive. Concerning di-receptor activity, 99 molecules are active as agonists, while 104 are active as antagonists on two opioid receptors. A total of 102 molecules show both agonistic and antagonistic activities on different opioid receptors.

The GRIND fingerprint was calculated for the dataset using 4 GRID probes: DRY (which represents hydrophobic interactions), O (sp2 carbonyl oxygen, representing an H bond acceptor), N1 (neutral flat NH like in amide, an H bond donor) and the TIP probe (molecular shape descriptor). A total of 10 auto- and cross-correlograms are calculated to generate a fingerprint vector of 1870 floating values i.e. 187 bins for each probe-probe correlogram. Each bin has a value representing the frequency of that particular distance (Figure 4). The first few bins for any given probe-probe correlogram are for short-distance correlations. Such bins describe localized patches on the molecular surface. On the other hand, bins for long distances are related to molecular shape and functional group distribution. The bins of distances more than 56 Å were excluded from model training in order to reduce noise and computational cost. The long-distance bins are only relevant for infrequent large-sized molecules and are not highly influential on models' performance (as shown in supplementary material tables 3S and 4S).

The 1D GRIND fingerprint is converted to a 2D array where each row represents one of the ten probe-probe correlograms at the input layer of CNN. The first convolutional layer of CNN includes 3 kernels, where each kernel is composed of 10 rows and 10 columns of trainable floating parameters (i.e. filters). Each kernel's window spans all correlograms and at a width of 10 Å across bins. The kernels are iteratively multiplied by a 2D input matrix with only a horizontal stride by 1 column (i.e. 1 Å) at a time. During training, the first convolutional kernels learn to discriminate relevant probe-probe correlograms with less focus on individual bins. This can be observed as rows of almost similar values across multiple columns

(Figure 5). On the other hand, the second-layer kernels are trained to be more bin- specific.

In order to establish comparison with fingerprints used in the previous report [30], the PubChem substructure fingerprint (https://ftp.ncbi.nlm.nih.gov/pubchem/specifications/pubchem_fingerprints.pdf) and Extended-Connectivity Fingerprint (ECFP) [10] of binary vectors of 881 and 1024-bit length, respectively, were used. While PubChem fingerprints represent structural keys, where each bit corresponds to a "pre-defined" structural feature (e.g., substructure or fragment), the ECFP fingerprint is considered hashed where there is no one-to-one correspondence between fragments and fingerprint bits. The DNN is trained using either of the above fingerprints. Since DNN is not designed to capture activity-relevant features, a layer of higher drop-out ratio (compared to CNN) was used in order to avoid overfitting data.

1. Models evaluation

The dataset used in training was imbalanced toward inactive ligands with an overall active: inactive ligands ratio of 1:11. During training, the ratio was adjusted to 1:2 using under-sampling techniques [30, 33]. The performance was evaluated from five-fold cross-validation training that is iterated 20 times (i.e. 100 training runs). The three models of GRIND-CNN, PubChem-DNN, and ECFP-DNN were trained using real and randomized activities. Finally, the averaged AUC-ROC, BA, MCC, TP, TN, FP, and FN values from every 100 runs were calculated [30, 31, 34, 35]. The results are summarized in

Figure 6.

The GRIND-CNN model exhibited promising classification performance across all six opioid receptor activity groups (

Table **3**), with AUC-ROC values ranging from 0.70 to 0.83, BA from 0.58 to 0.71, and MCC from 0.21 to 0.46. Notably, the Mu agonist model achieved the highest performance, with AUC-ROC = 0.83 ± 0.03, BA = 0.71 ± 0.04, and MCC = 0.46 ± 0.07, indicating strong predictive ability and good balance between sensitivity and specificity. Conversely, the Delta antagonist model had the lowest overall performance with AUC-ROC of 0.70 ± 0.05, BA of 0.58 ± 0.05, and MCC of 0.21 ± 0.12. The higher performance of the Mu agonist model could be attributed to the highest ratio of active/inactive ligands in the dataset (0.19) compared to the delta antagonist dataset (0.08), in addition, to possibly a clearer pattern of Mu agonist compared to others like kappa antagonist dataset which has actives: inactive ligands ratio of 0.15. The performance of the PubChem-DNN model (AUC-ROC between 0.74 and 0.90, BA between 0.63 and 0.80, and MCC between 0.29 and 0.61) is comparable to that of the ECFP-DNN model (AUC-ROC between 0.74 to 0.91, BA between 0.66 to 0.83 and MCC between 0.34 to 0.67) as shown in

Table **4** and

Table **6**, respectively.

Upon using randomly shuffled activities (i.e. Y-randomized inputs), the performance dropped for all models. The drop in mean AUC-ROC values for GRIND-CNN (between 0.57 and 0.6) and PubChem-DNN models (between 0.53 and 0.61) approaches random change (0.5), thus indicating no meaningful predictive power. Similarly, BA values were dropped (~0.5) and MCC values approached zero for both GRIND-CNN and PubChem-DNN models, which indicates a lack of any real signal in the randomized models. The results suggest a lower potential for the models to overfit to noise. On the other hand, the ECFP-DNN showed higher noise dependence and failure to capture activity-relevant features for Mu agonists and Kappa antagonists as shown by higher AUC-ROC (between 0.55 and 0.72), MCC (up to 0.3) and BA (between 0.53 and 0.64). This could be attributed to limitations in ECFP where a single fingerprint bit (i.e., hashed feature) can represent multiple, structurally different substructures [10]. Accordingly, the ECFP-DNN model is less efficient in capturing activity-relevant features compared to other models.

The prediction performance is further improved when combining the outputs of the GRIND-CNN model with either of PubChem-DNN model (Table 5) or the ECFP-DNN model (

Table **7**). However, the GRIND-CNN/PubChem-DNN model provides relatively lower noise dependence (i.e. lower AUC-ROC, BA, and MCC values upon using Y-randomized values). Therefore, the later combined model can potentially be used to screen for opioid ligands irrespective of the molecular scaffold.

2. Hot regions within GRIND fingerprint of potential correspondence to "address" and "message" of opioid ligand

Saliency (hot) regions within the GRIND fingerprint that are relevant to opioid activities can be identified by using the gradient of the trained model's output with respect to the input fingerprint pixel. The analysis of saliency maps accumulated during 20 iterations of 5-fold cross-validation training revealed a notable degree of consistency in the regions contributing most to model predictions (

Figure **7** and

Figure **8**). The correlograms of Dry-Dry, TIP-TIP, and Dry-N1 are most relevant to activities for all opioid ligands (including agonists and antagonists). The Dry-Dry correlogram describes distances between hydrophobic patches on the molecule, while the TIP-TIP correlogram is related

to molecular size, and the Dry-N1 correlogram for distances between hydrophobic patches and hydrogen bond acceptor patches [12, 36].

In order to discriminate the "message" and "address" hot correlogram pixels within the 2D GRIND fingerprints, the statistical difference in saliency maps was used. The GRIND-CNN model was trained twice: once using the combined dataset of agonist activities and once using the combined dataset of antagonist activities, each spanning all three opioid receptor subtypes (Mu, Kappa, and Delta). This approach allowed for capturing the patterns distinguishing agonist versus antagonist ("message pixels"). While in order to discriminate the receptor-specific "address pixels", the model was trained three times; each time exclusively on combined active compounds (agonists and antagonists) specific to one receptor subtype (Mu, Kappa, or Delta). The performance of these models is shown in supplementary material (Tables 5 and 6).

The accumulated saliency maps for agonists versus antagonists across all receptors (Figure 9) show hot correlogram pixels (distances or bins) that are significantly different (Figure 10). Antagonist hot correlograms occur relatively at larger inter-feature distances compared to those of agonists. Notably, antagonists show pronounced hot hydrophobic-related correlograms; particularly in DRY-DRY (33–35 Å), DRY-O (40–42 Å), and DRY-TIP (50–53 Å) regions. This suggests a significant effect of protruding hydrophobic fields for antagonist messages [37]. In contrast, agonist messages are significantly related to small localized hydrophobic patches which are observed as DRY-DRY (10–27 Å) and DRY-N1 (15–25 Å). Consequently, the significant differences between antagonist and agonist messages are related to hydrophobic-related correlograms.

The accumulated saliency maps for ligands selective to either Mu, Kapp, or Delta receptors (Figure 11) also show regions of significant differences that are potentially relevant to the addresses (Figure 12). Mu-selective hot correlograms are predominantly located at shorter distances (<30 Å), primarily involving DRY-DRY (15–30 Å), DRY-N1 (15–20 Å), and TIP-TIP (4–8 Å). In contrast, Kappa-selective hot correlograms are fewer and span DRY-DRY (18–43 Å) and TIP-TIP regions at approximately 25 Å (relative to Mu) and 53–54 Å (relative to Delta). Accordingly, Kappa's selective address depends mainly on excluding Mu and Delta selective hot correlograms rather than dependence on its hot correlograms. Since, the largest distance TIP-TIP correlogram is related to molecular size [36], the Kappa address looks to be distinct from the Delta address (but not the Mu address) by dependence on the difference in size [38]. Delta-selective hot correlograms are more abundant and look to be distinct from Mu and Kappa. The saliency maps difference of Mu-Delta and Kappa-Delta are partially overlapped indicating dependence of Delta address on equivalent hot correlograms against the other receptors. Accordingly, the Mu address potentially depends on the presence of localized structural features to be distinct from the Kappa address and on the exclusion of local and distal structural features to be distinct from Delta. While Kappa address depends mainly on the exclusion of features to be distinct from Mu and Delta addresses as well as on difference in size with respect to Delta address. The Delta address looks the most distinct with respect to features that span the whole 2D fingerprint, such features are absent from both Mu and Kappa addresses.

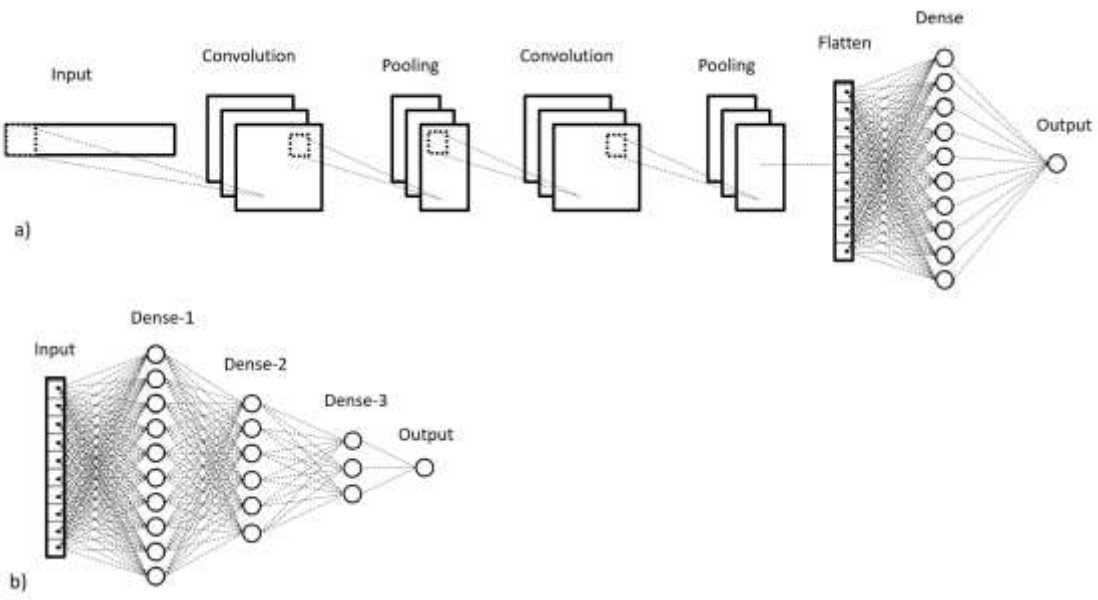


Figure 1. The architectures of a) convolutional neural network (CNN), and b) dense neural network (DNN).

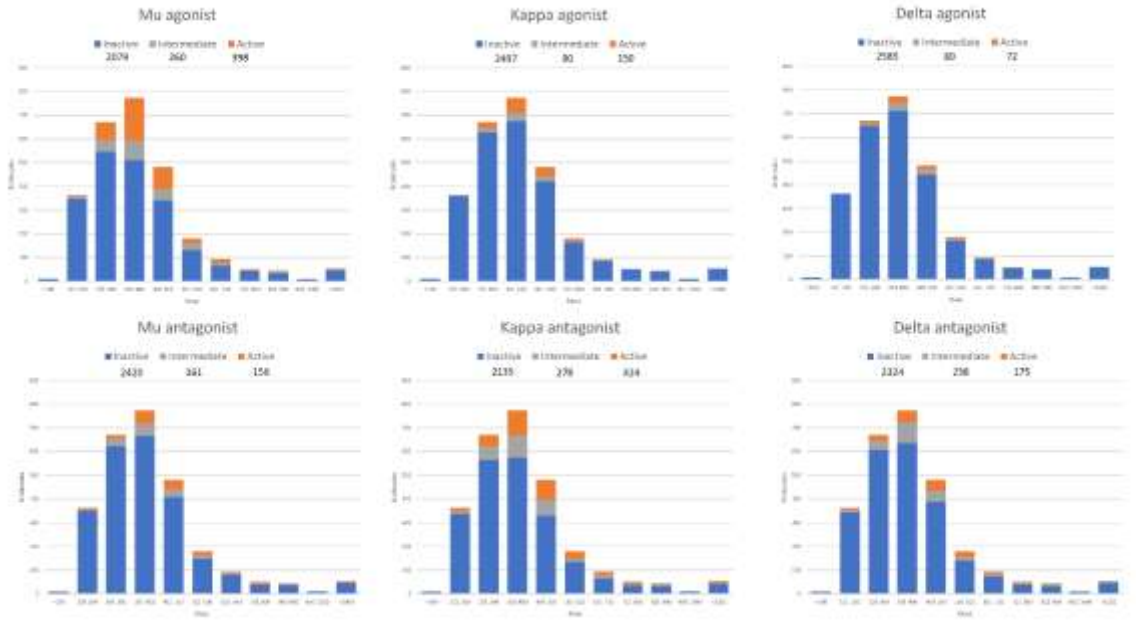
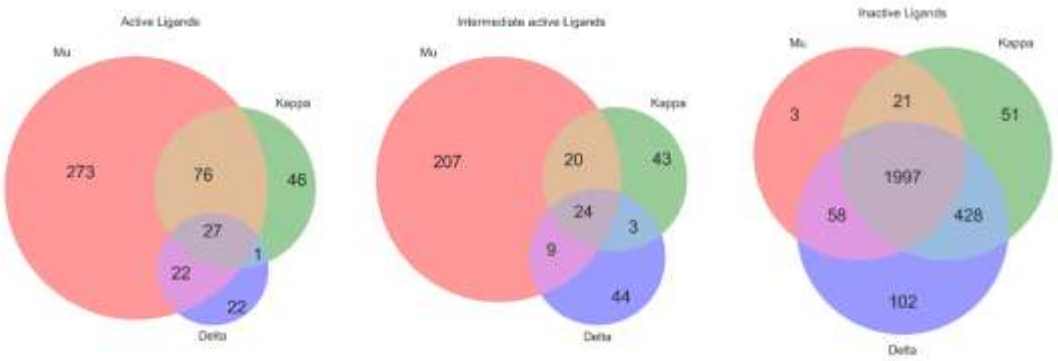


Figure 2. Analysis for training set molecules showing histograms for molecular weight distribution, total number of inactive, intermediate and active molecules.



a) Agonist activity overlap across opioid receptors

b) Antagonist activity overlap across opioid receptors

Figure 3.The distribution of ligands in the training dataset with respect to (a) agonistic and (b) antagonistic activities across opioid receptors.

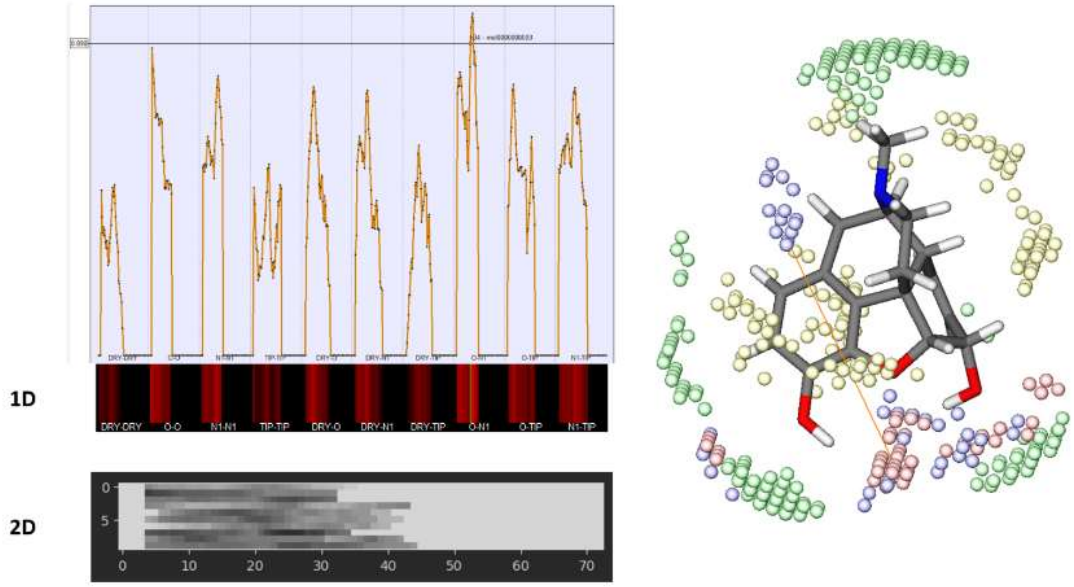
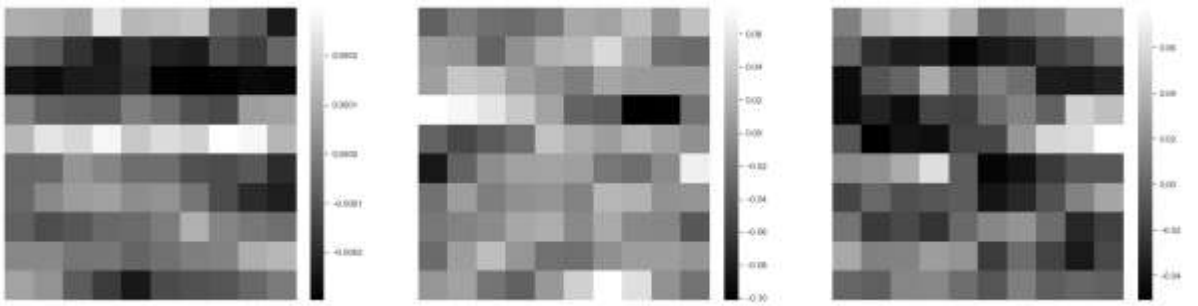
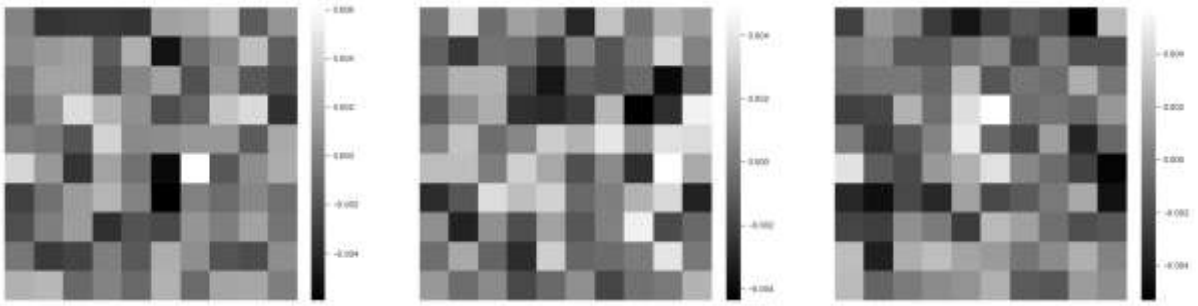


Figure 4.The construction of 2D finger print from 1D frequency of distances distribution for probe-probe interaction fields for morphine molecule. Bin of distances up to 72 Å are shown.



a)



b)

Figure 5 .The kernels of a) first and b) second convolutional layer of CNN trained on Mu agonists. The first-layer kernels exhibiting horizontal bands of negative values that diminish the relevancy of particular correlograms on final prediction, while the second-layer kernels is more bins' specific.

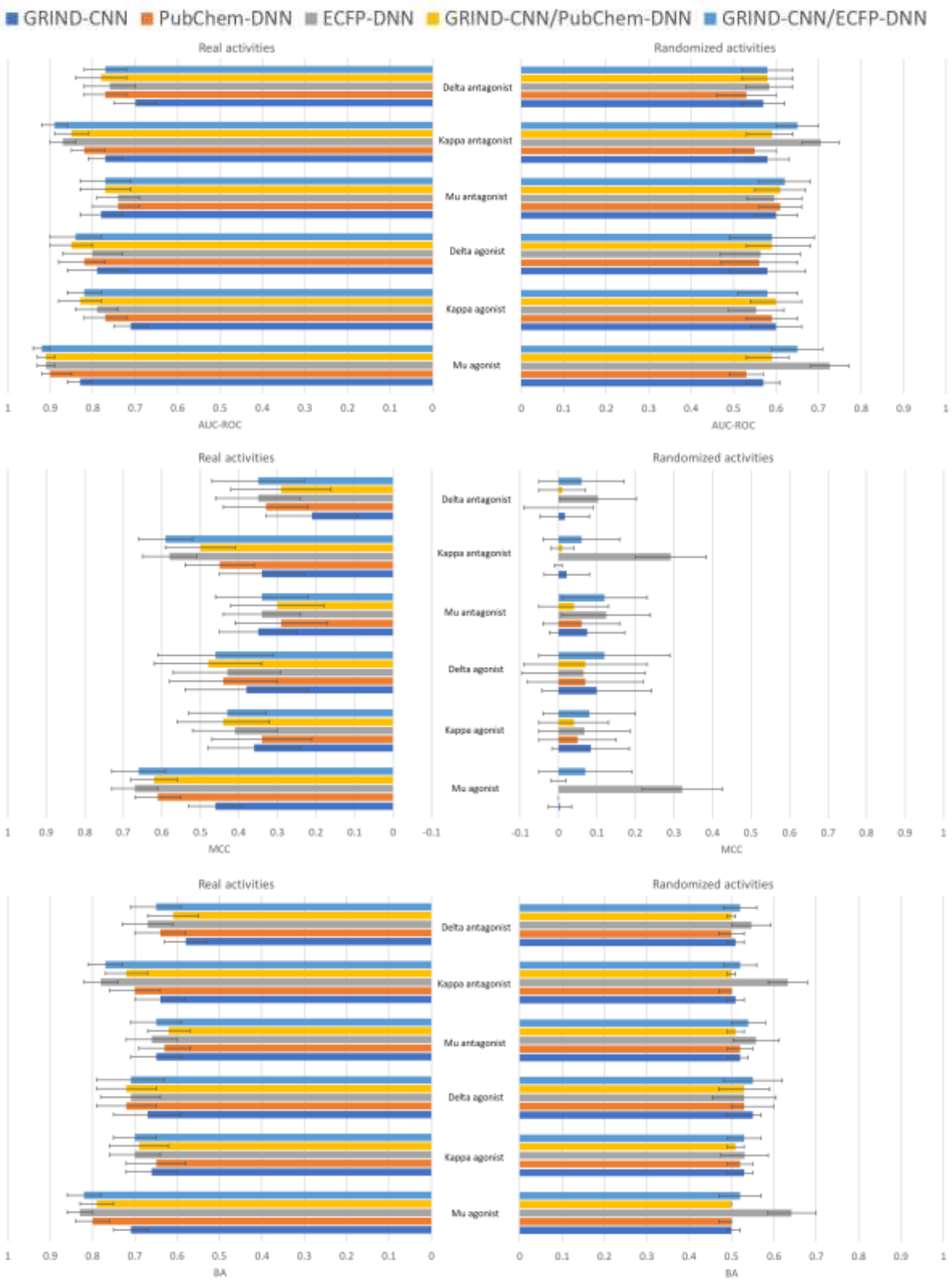
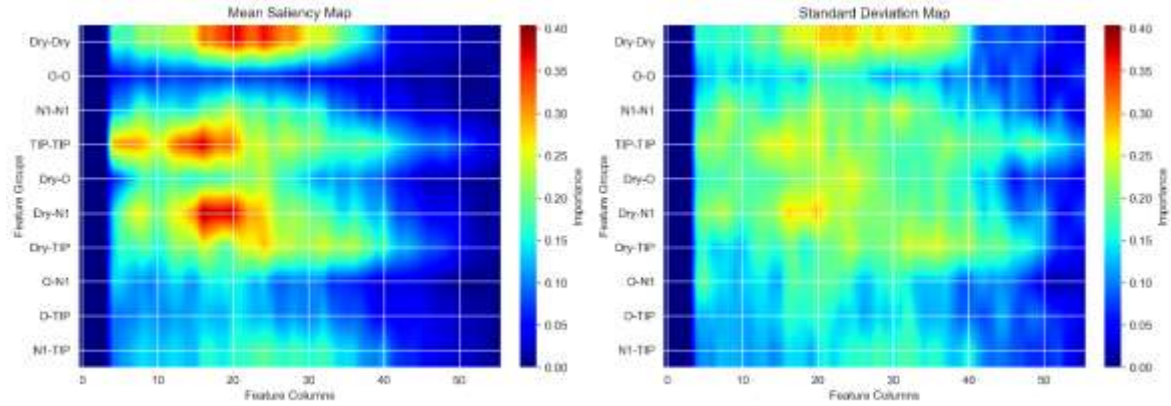
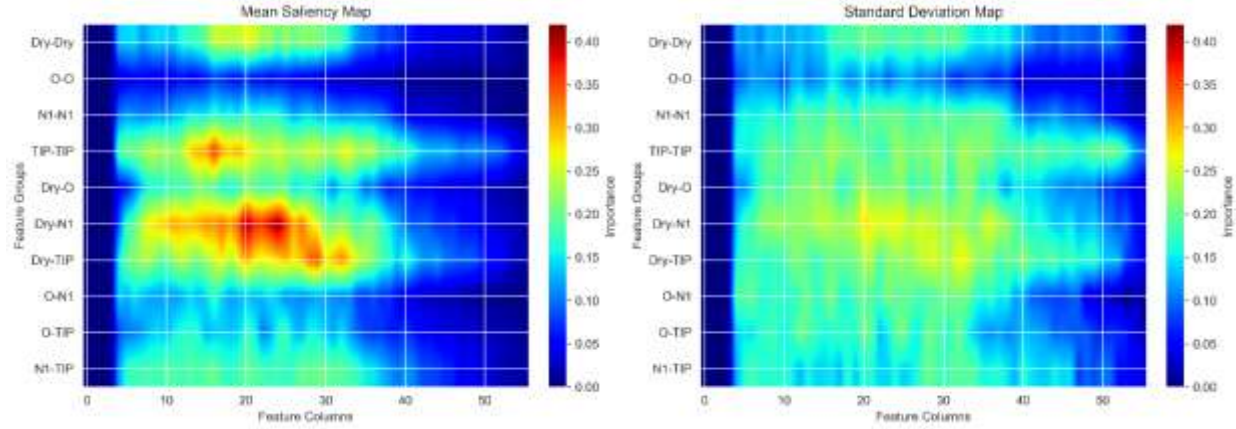


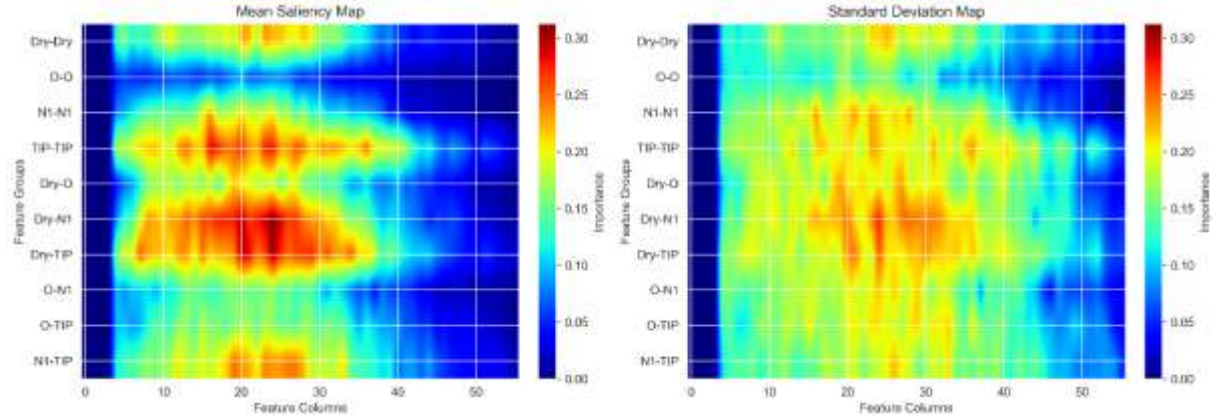
Figure 6 The mean±standard deviation values of AUC-ROC, MCC and BA from 5-fold cross validation for GRIND-CNN, PubChem-DNN, ECFP-DNN and GRIND-CNN/PubChem-DNN, and GRIND-CNN/ECFP-DNN models using real and randomized activities.



Mu agonist

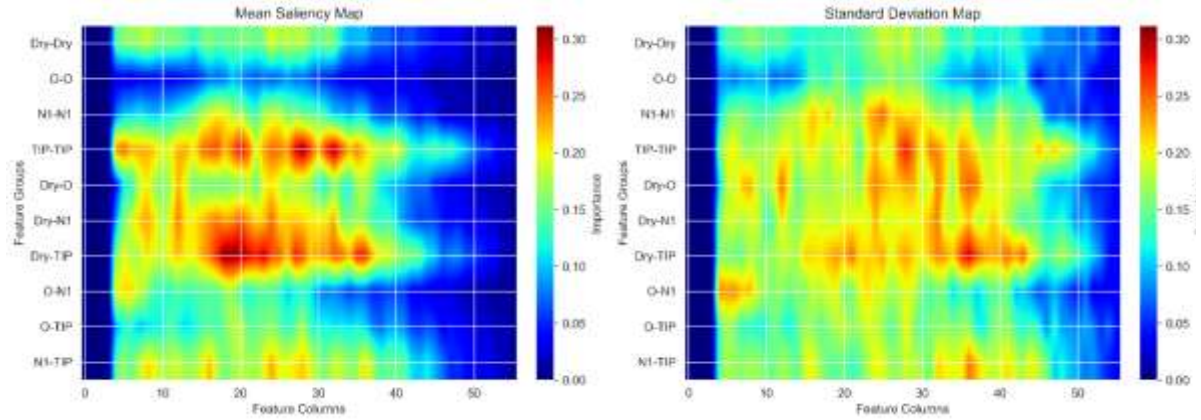


Kappa agonist

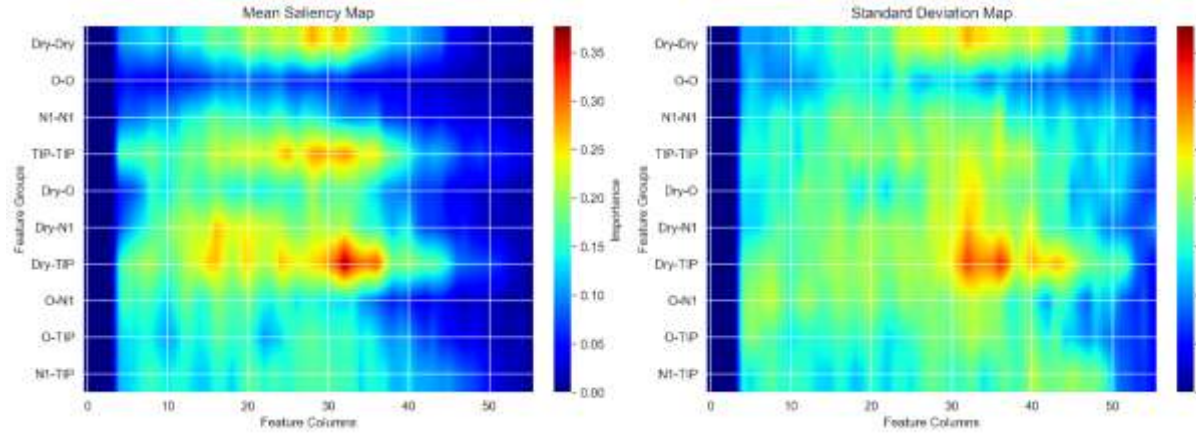


Delta agonist

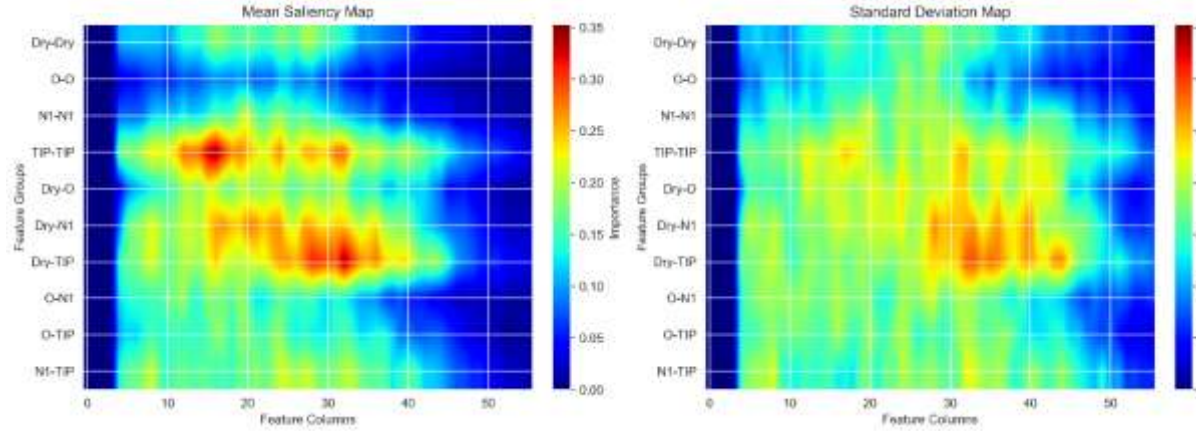
Figure 7 the mean and standard deviation saliency map (shown separately) for GRIND-CNN models trained to predict agonists of each of Mu, Kappa and Delta receptors.



Mu antagonist

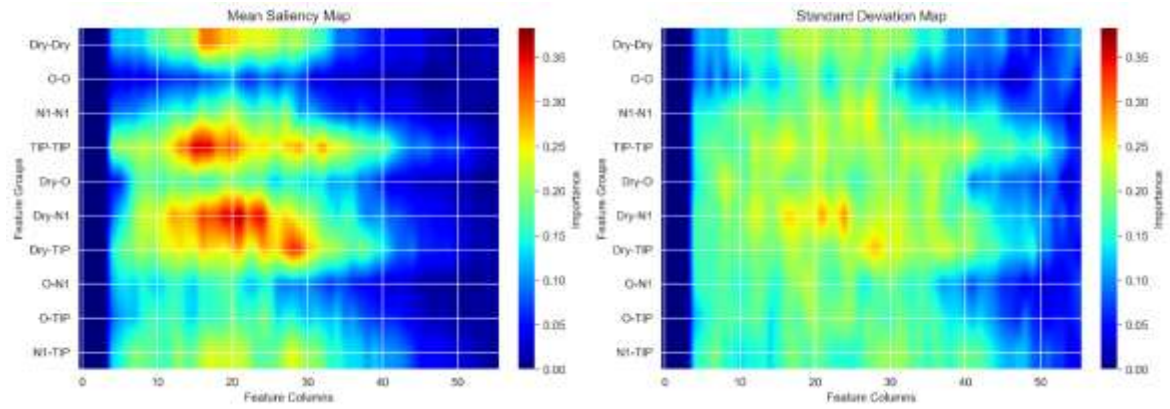


Kappa antagonist

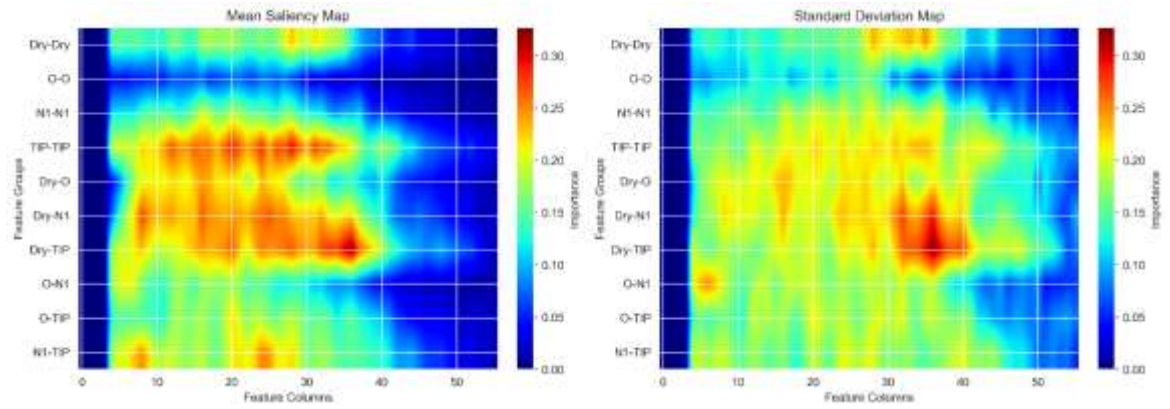


Delta antagonist

Figure 8. the mean and standard deviation saliency map (shown separately) for GRIND-CNN models trained to predict antagonists of each of Mu, Kappa and Delta receptors.



a) Agonists



b) Antagonists

Figure 9 .The mean and standard deviation saliency maps (shown separately) for GRIND-CNN models trained to predict (a) agonists and (b) antagonists irrespective of receptor type.

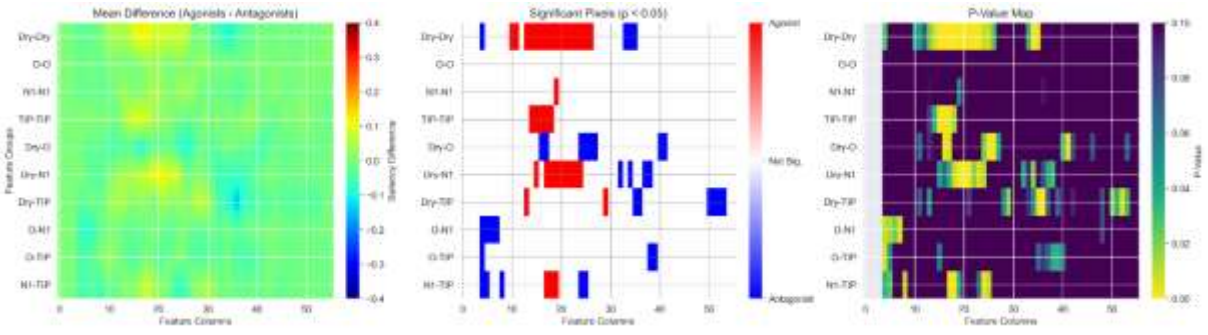
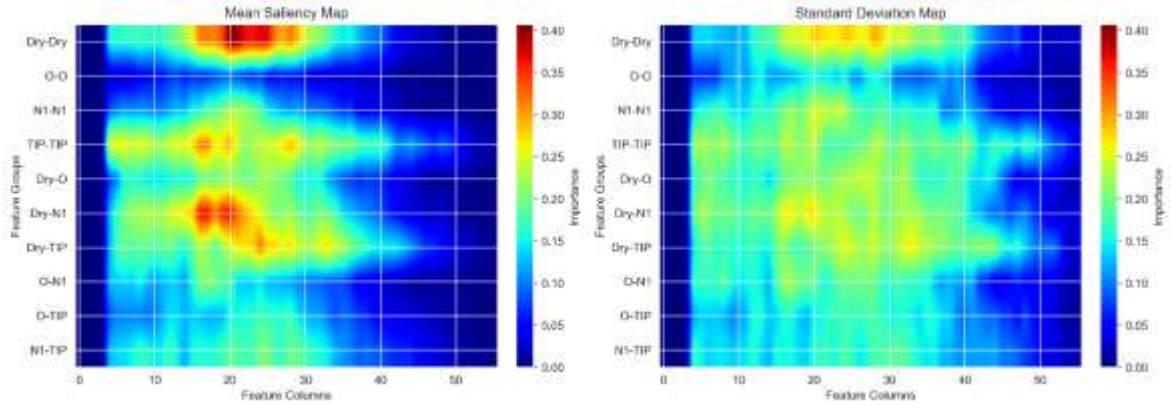
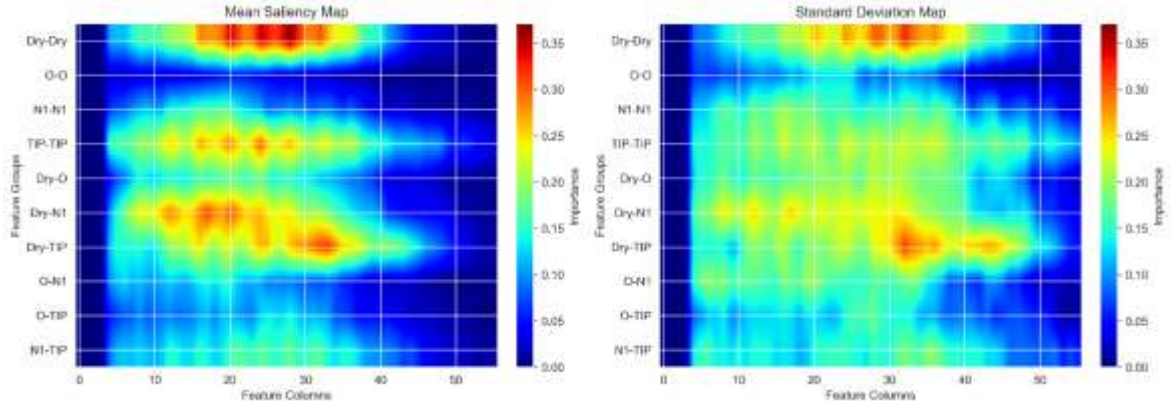


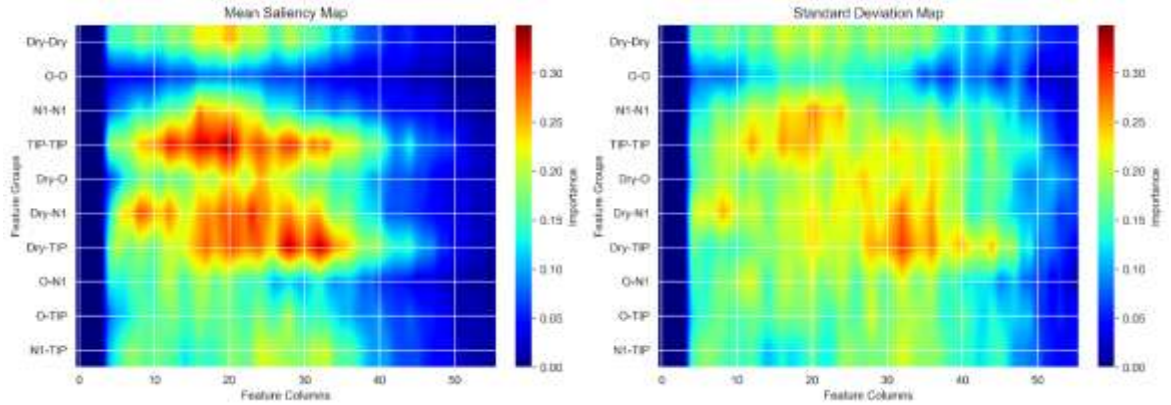
Figure 10. The significant difference between accumulated saliency maps relevant to opioid's message (agonist vs antagonist).



a) Mu active ligands



b) Kappa active ligands



c) Delta active ligands

Figure 11. the mean and standard deviation saliency map (shown separately) for GRIND-CNN models trained to predict activity (irrespective whether agonist or antagonist) on each of a) Mu, b) Kappa and c) Delta receptors

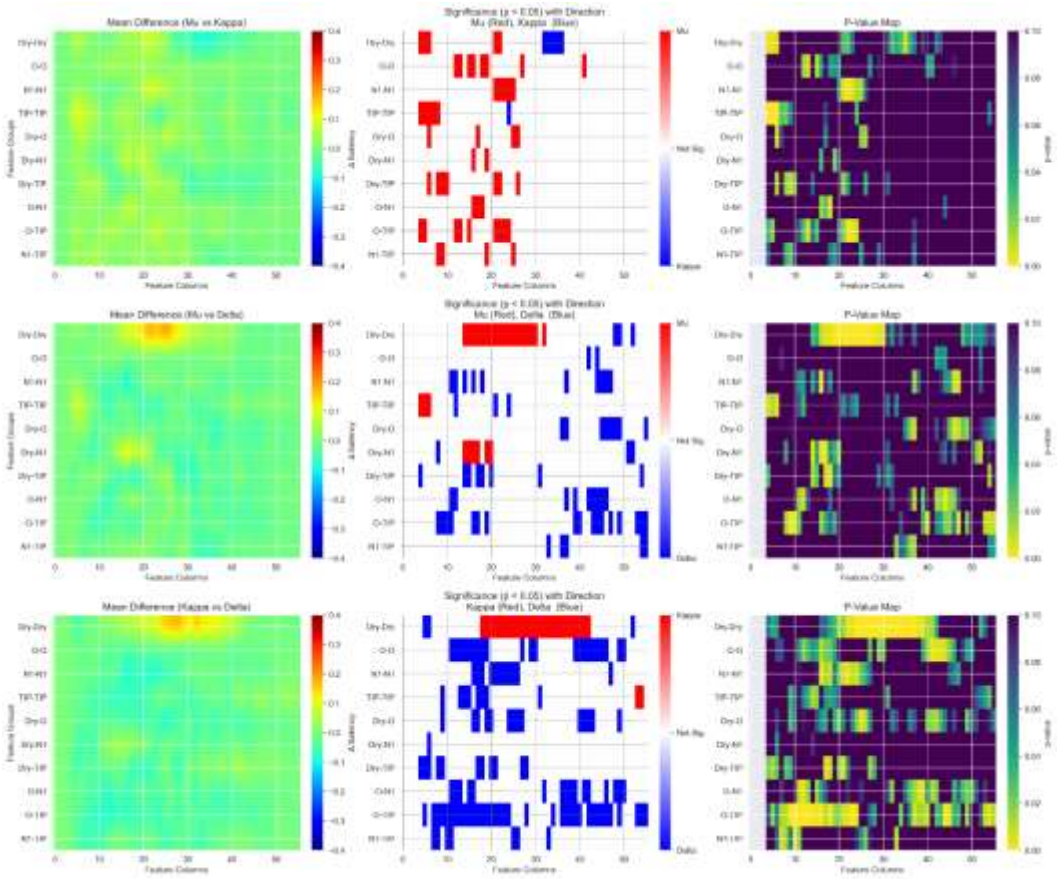


Figure 12. The pairwise significant difference between accumulated saliency maps relevant to opioids' addresses (Mu, Kappa and Delta selectivity).

Table 1. Architecture of convolutional neural network (CNN)

	Layer (type)	Output Shape	Params
Input layer	Reshape	(None, 10, 56, 1)	0
	Masking	(None, 10, 56, 1)	0
1st convolution layer	Conv2D	(None, 10, 56, 3)	303
	Batch Normalization	(None, 10, 56, 3)	12
	Activation function (Leaky Relu)	(None, 10, 56, 3)	0
	Dropout (0.15)	(None, 10, 56, 3)	0
	Max Pooling (1,4)	(None, 10, 14, 3)	0
2nd convolution layer	Conv2D	(None, 10, 14, 3)	903
	Batch Normalization	(None, 10, 14, 3)	12
	Activation function (Leaky Relu)	(None, 10, 14, 3)	0
	Dropout (0.15)	(None, 10, 14, 3)	0
	Max Pooling (1,4)	(None, 10, 3, 3)	0
	Flatten	(None, 90)	0
Dense layer	Dense	(None, 10)	910
	Batch Normalization	(None, 10)	40
	Activation function (Leaky Relu)	(None, 10)	0
	Dropout (0.15)	(None, 10)	0
Output layer	Dense	(None, 1)	11
Total params:			**2,191**
Trainable params:			**2,159**
Non-trainable params:			**32**

Table 2 . Architecture of deep neural network (DNN)

	Layer (type)	Output Shape	Params
Input layer	Reshape	(None, 881)*	0
1st Dense Layer	Dense0	(None, 250)	220500
	Batch Normalization	(None, 250)	1000
	Activation function (Leaky Relu)	(None, 250)	0
	Dropout (0.15)	(None, 250)	0
2nd Dense Layer	Dense	(None, 90)	22590
	Batch Normalization	(None, 90)	360
	Activation function (Leaky Relu)	(None, 90)	0
	Dropout (0.15)	(None, 90)	0
3rd Dense Layer	Dense	(None, 20)	1820
	Batch Normalization	(None, 20)	80
	Activation function (Leaky Relu)	(None, 20)	0
	Dropout (0.15)	(None, 20)	0
Output layer	Dense	(None, 1)	21
Total params:			**246,371**
Trainable params:			**245,651**
Non-trainable params:			**720**

* for ECFP-DNN the reshape layer is (None,1024) and dense layer contains 256,250 parameters, thus all trainable parameters are 281,401

Table 3. Performance measurements of GRIND-CNN model using real and randomized activities

	Average	Mu agonist	Kappa agonist	Delta agonist	Mu antagonist	Kappa antagonist	Delta antagonist
Real activities	**AUC-ROC**	**0.83 ± 0.03**	**0.71 ± 0.04**	**0.79 ± 0.07**	**0.78 ± 0.05**	**0.77 ± 0.04**	**0.70 ± 0.05**
	BA	0.71 ± 0.04	0.66 ± 0.06	0.67 ± 0.08	0.65 ± 0.06	0.64 ± 0.06	0.58 ± 0.05
	MCC	0.46 ± 0.07	0.36 ± 0.12	0.38 ± 0.16	0.35 ± 0.10	0.34 ± 0.11	0.21± 0.12
	TP	42.66 ± 8.03	13.17 ± 4.38	7.16 ± 2.46	25.51 ± 9.21	24.70 ± 8.85	9.02 ± 5.19
	TN	141.44 ± 6.52	52.64 ± 3.19	24.50 ± 2.46	116.77 ± 6.05	117.36 ± 5.88	63.10 ± 4.74
	FP	17.76 ± 6.49	7.36 ± 3.19	4.30 ± 2.34	12.83 ± 6.03	12.24 ± 5.82	6.90 ± 4.74
	FN	36.94 ± 7.91	16.83 ± 4.38	7.24± 2.33	39.29 ± 9.19	40.10 ± 8.85	25.98 ± 5.19
Randomized Activities	**AUC-ROC**	**0.57 ± 0.04**	**0.60 ± 0.06**	**0.58 ± 0.09**	**0.60 ± 0.05**	**0.58 ± 0.05**	**0.57 ± 0.05**
	BA	0.50 ± 0.01	0.53 ± 0.04	0.55 ± 0.06	0.52 ± 0.03	0.51 ± 0.02	0.51 ± 0.02
	MCC	0.004 ± 0.03	0.084 ± 0.10	0.100± 0.142	0.075 ± 0.098	0.022 ± 0.059	0.017 ± 0.064
	TP	0.35 ± 1.71	3.54 ± 3.39	4.44 ± 2.31	2.91 ± 3.63	1.18 ± 3.67	0.93 ± 2.14
	TN	158.83 ± 1.62	56.08± 3.89	22.53 ± 3.62	59.50 ± 3.83	128.70 ± 3.04	68.89 ± 2.40
	FP	0.37 ± 1.60	3.92 ± 3.89	6.27 ± 3.61	2.90 ± 3.79	0.90 ± 3.00	1.11 ± 2.40
	FN	79.25 ± 1.79	26.46 ± 3.39	9.96 ± 2.29	28.29 ± 3.64	63.62 ± 3.65	34.07 ± 2.14

Table 4. Performance measurements of PubChem-DNN model using real and randomized activities

	Average	Mu Agonist	Kappa Agonist	Delta Agonist	Mu Antagonist	Kappa Antagonist	Delta Antagonist
Real activities	**AUC-ROC**	0.90 ± 0.02	0.77 ± 0.05	0.82 ± 0.06	0.74 ± 0.06	0.82 ± 0.03	0.77 ± 0.05
	BA	0.80 ± 0.04	0.65 ± 0.07	0.72 ± 0.07	0.63 ± 0.06	0.70 ± 0.06	0.64 ± 0.06
	MCC	0.61 ± 0.06	0.34 ± 0.13	0.44 ± 0.14	0.29 ± 0.12	0.45 ± 0.09	0.33 ± 0.11
	TP	57.25 ± 6.86	15.46 ± 5.45	8.78 ± 2.08	12.02 ± 4.77	33.01 ± 9.22	14.38 ± 5.40
	TN	140.84 ± 6.19	60.53 ± 4.32	23.68 ± 2.21	54.04 ± 4.20	116.16 ± 6.42	61.20 ± 4.74
	FP	18.36 ± 6.10	9.47 ± 4.32	5.12 ± 2.14	8.36 ± 4.23	13.44 ± 6.48	8.80 ± 4.74
	FN	22.35 ± 6.75	19.54 ± 5.45	5.62 ± 2.04	19.18 ± 4.75	31.79 ± 9.23	20.62 ± 5.40
Randomized Activities	**AUC-ROC**	0.53 ± 0.04	0.59 ± 0.06	0.56 ± 0.09	0.61 ± 0.05	0.55 ± 0.05	0.53 ± 0.07
	BA	0.50 ± 0.00	0.52 ± 0.03	0.53 ± 0.07	0.52 ± 0.03	0.50 ± 0.00	0.50 ± 0.03
	MCC	0.00 ± 0.00	0.05 ± 0.10	0.07 ± 0.15	0.06 ± 0.10	0.00 ± 0.01	0.00 ± 0.09
	TP	0.00 ± 0.00	2.87 ± 3.04	4.42 ± 2.10	3.79 ± 3.61	0.01 ± 0.10	0.00 ± 3.27
	TN	159.20 ± 0.40	56.38 ± 4.11	21.86 ± 3.34	57.56 ± 4.82	129.60 ± 0.49	159.20 ± 3.18
	FP	0.00 ± 0.00	3.62 ± 4.11	6.94 ± 3.28	4.84 ± 4.72	0.00 ± 0.00	0.00 ± 3.18
	FN	79.60 ± 0.49	27.13 ± 3.04	9.98 ± 2.13	27.41 ± 3.61	64.79 ± 0.43	79.60 ± 3.27

Table 5. Performance measurements of GRIND-CNN/PubChem-DNN models where predictions are merged by 0.5 coefficient.

	Average	Mu Agonist	Kappa Agonist	Delta Agonist	Mu Antagonist	Kappa Antagonist	Delta Antagonist
Real activities	AUC-ROC	0.91 ± 0.02	0.83 ± 0.05	0.85 ± 0.05	0.77 ± 0.06	0.85 ± 0.04	0.78 ± 0.06
	BA	0.79 ± 0.04	0.69 ± 0.07	0.72 ± 0.07	0.62 ± 0.05	0.72 ± 0.05	0.61 ± 0.06
	MCC	0.62 ± 0.06	0.44 ± 0.12	0.48 ± 0.14	0.30 ± 0.12	0.50 ± 0.09	0.29 ± 0.13
	TP	53.35 ± 6.36	14.19 ± 4.22	7.97 ± 2.30	10.00 ± 4.19	32.81 ± 8.13	9.92 ± 5.18
	TN	146.41 ± 4.42	54.54 ± 2.57	25.64 ± 1.89	57.13 ± 3.32	120.04 ± 4.45	65.43 ± 3.28
	FP	12.79 ± 4.32	5.46 ± 2.57	3.16 ± 1.76	5.27 ± 3.27	9.56 ± 4.44	4.57 ± 3.28
	FN	26.25 ± 6.32	15.81 ± 4.22	6.43 ± 2.25	21.20 ± 4.20	31.99 ± 8.13	25.08 ± 5.18
Randomized Activities	AUC-ROC	0.59 ± 0.04	0.60 ± 0.06	0.59 ± 0.09	0.61 ± 0.06	0.59 ± 0.05	0.58 ± 0.06
	BA	0.50 ± 0.00	0.51 ± 0.02	0.53 ± 0.06	0.51 ± 0.02	0.50 ± 0.01	0.50 ± 0.01
	MCC	0.00 ± 0.02	0.04 ± 0.09	0.07 ± 0.16	0.04 ± 0.09	0.01 ± 0.03	0.01 ± 0.06
	TP	0.05 ± 0.30	1.13 ± 1.68	2.76 ± 2.09	1.08 ± 2.18	0.16 ± 0.94	0.33 ± 1.05
	TN	159.11 ± 0.53	58.98 ± 1.48	25.04 ± 2.51	61.49 ± 1.78	129.54 ± 0.61	69.77 ± 0.91
	FP	0.09 ± 0.32	1.02 ± 1.48	3.76 ± 2.40	0.91 ± 1.75	0.06 ± 0.34	0.23 ± 0.91
	FN	79.55 ± 0.61	28.87 ± 1.68	11.64 ± 2.10	30.12 ± 2.15	64.64 ± 1.00	34.67 ± 1.05

Table 6.Performance measurements of ECFP-DNN model using real and randomized activities

	Average	Mu Agonist	Kappa Agonist	Delta Agonist	Mu Antagonist	Kappa Antagonist	Delta Antagonist
Real activities	AUC-ROC	0.91 ± 0.02	0.79 ± 0.05	0.80 ± 0.07	0.74 ± 0.05	0.87 ± 0.03	0.76 ± 0.06
	BA	0.83 ± 0.03	0.70 ± 0.06	0.71 ± 0.07	0.66 ± 0.06	0.78 ± 0.04	0.67 ± 0.06
	MCC	0.67 ± 0.06	0.41 ± 0.11	0.43 ± 0.14	0.34 ± 0.10	0.58 ± 0.07	0.35 ± 0.11
	TP	61.08 ± 5.64	17.76 ± 3.71	8.88 ± 1.93	16.74 ± 4.22	44.16 ± 5.73	18.79 ± 4.51
	TN	142.30 ± 6.78	48.83 ± 4.04	23.36 ± 2.34	49.47 ± 4.63	114.32 ± 6.20	56.49 ± 4.50
	FP	16.90 ± 6.71	11.17 ± 4.04	5.44 ± 2.30	12.93 ± 4.57	15.28 ± 6.19	13.51 ± 4.50
	FN	18.52 ± 5.64	12.24 ± 3.71	5.52 ± 1.90	14.46 ± 4.22	20.64 ± 5.66	16.21 ± 4.51
Randomized Activities	AUC-ROC	0.727 ± 0.045	0.553 ± 0.065	0.563 ± 0.094	0.596 ± 0.065	0.705 ± 0.043	0.584 ± 0.055
	BA	0.642 ± 0.056	0.531 ± 0.056	0.530 ± 0.075	0.558 ± 0.054	0.633 ± 0.046	0.547 ± 0.046
	MCC	0.322 ± 0.104	0.068 ± 0.118	0.065 ± 0.161	0.124 ± 0.115	0.292 ± 0.092	0.103 ± 0.099
	TP	32.37 ± 11.03	9.46 ± 3.87	4.70 ± 2.26	11.49 ± 3.97	27.73 ± 6.98	11.34 ± 4.02
	TN	139.51 ± 8.12	44.77 ± 6.11	21.14 ± 3.74	46.63 ± 6.78	108.67 ± 7.21	53.94 ± 6.12
	FP	19.69 ± 8.08	15.23 ± 6.11	7.66 ± 3.72	15.77 ± 6.76	20.93 ± 7.23	16.06 ± 6.12
	FN	47.23 ± 11.07	20.54 ± 3.87	9.70 ± 2.23	19.71 ± 3.95	37.07 ± 6.97	23.66 ± 4.02

Table 7.Performance measurements of GRIND-CNN/ECFP-DNN models where predictions are merged by 0.5 coefficient.

	Average	Mu Agonist	Kappa Agonist	Delta Agonist	Mu Antagonist	Kappa Antagonist	Delta Antagonist
Real activities	AUC-ROC	0.92 ± 0.02	0.82 ± 0.04	0.84 ± 0.06	0.77 ± 0.06	0.89 ± 0.03	0.77 ± 0.05
	BA	0.82 ± 0.04	0.70 ± 0.05	0.71 ± 0.08	0.65 ± 0.06	0.77 ± 0.04	0.65 ± 0.06
	MCC	0.66 ± 0.07	0.43 ± 0.10	0.46 ± 0.15	0.34 ± 0.12	0.59 ± 0.07	0.35 ± 0.12
	TP	56.45 ± 6.66	15.49 ± 3.58	7.81 ± 2.20	13.23 ± 3.98	39.29 ± 5.86	14.53 ± 4.68
	TN	147.28 ± 3.86	52.78 ± 3.22	25.45 ± 1.97	54.62 ± 3.51	120.54 ± 3.47	62.08 ± 3.19
	FP	11.92 ± 3.84	7.22 ± 3.22	3.35 ± 1.86	7.78 ± 3.43	9.06 ± 3.36	7.92 ± 3.19
	FN	23.15 ± 6.63	14.51 ± 3.58	6.59 ± 2.17	17.97 ± 4.03	25.51 ± 5.88	20.47 ± 4.68
Randomized Activities	AUC-ROC	0.65 ± 0.06	0.58 ± 0.07	0.59 ± 0.10	0.62 ± 0.06	0.65 ± 0.05	0.58 ± 0.06
	BA	0.52 ± 0.05	0.53 ± 0.04	0.55 ± 0.07	0.54 ± 0.04	0.52 ± 0.04	0.52 ± 0.04
	MCC	0.07 ± 0.12	0.08 ± 0.12	0.12 ± 0.17	0.12 ± 0.11	0.06 ± 0.10	0.06 ± 0.11
	TP	4.14 ± 8.71	3.66 ± 3.29	3.77 ± 2.24	5.02 ± 3.39	3.42 ± 6.35	2.78 ± 3.53
	TN	157.51 ± 3.66	55.73 ± 3.98	24.20 ± 2.72	57.60 ± 3.44	127.03 ± 4.57	67.12 ± 3.21
	FP	1.69 ± 3.63	4.27 ± 3.98	4.60 ± 2.69	4.80 ± 3.44	2.57 ± 4.58	2.88 ± 3.21
	FN	75.46 ± 8.73	26.34 ± 3.29	10.63 ± 2.27	26.18 ± 3.41	61.38 ± 6.33	32.22 ± 3.53

Conclusion

Three types of models were developed to predict opioid receptor activities for agonists and antagonists. Each model was trained to predict the activities of agonists and antagonists for the three opioid receptor subtypes of Mu, Kappa, and Delta giving 6 trained models. The GRIND-CNN model demonstrates comparable performance to

PubChem-DNN and ECFP-DNN models. However, the ECFP-DNN model showed the lowest activity-related feature recognition. The combination of GRIND-CNN and PubChem-DNN predictions provides improved performance and activity-related feature recognition compared to separated models. The combined models can be used to screen for opioid ligands irrespective of molecular scaffold. The statistically significant difference in saliency maps for trained GRIND-CNN models revealed hot correlogram bins of potential relevancy to the message and address theory of opioid ligands

Acknowledgment

Thanks to information technology school at Nineveh university for assistance in accessing supercomputing power.

Conflicts of Interest

Author declares no conflicts of interest.

Funding

Research is not funded by any institutions.

Ethics Statements

No ethical approval is needed.

Author Contribution

The author confirm contribution to the paper as follows: study conception and design, data collection and processing, analysis and interpretation of results, drafting manuscript and revision are all done by Mohammed Noorladeen Al-Qattan.

References

1. Che T, Roth BL. Molecular basis of opioid receptor signaling. Cell. 2023;186(24):5203-19.

2. Hauser AS, Attwood MM, Rask-Andersen M, Schiöth HB, Gloriam DE. Trends in GPCR drug discovery: new agents, targets and indications. Nature reviews Drug discovery. 2017;16(12):829-42.

3. Sriram K, Insel PA. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? Molecular pharmacology. 2018;93(4):251-8.

4. Stein C. Opioid Receptors. Annual Review of Medicine. 2016;67(Volume 67, 2016):433-51.

5. Zhang L, Zhang J-T, Hang L, Liu T. Mu Opioid Receptor Heterodimers Emerge as Novel Therapeutic Targets: Recent Progress and Future Perspective. Frontiers in Pharmacology. 2020;11.

6. Alananzeh WA, Al-Qattan MN, Ayipo YO, Mordi MN. N-substituted tetrahydro-beta-carboline as mu-opioid receptors ligands: in silico study; molecular docking, ADMET and molecular dynamics approach. Molecular diversity. 2024;28(3):1273-89.

7. Evers A, Hessler G, Matter H, Klabunde T. Virtual Screening of Biogenic Amine-Binding G-Protein Coupled Receptors: Comparative Evaluation of Protein- and Ligand-Based Virtual Screening Protocols. Journal of Medicinal Chemistry. 2005;48(17):5448-65.

8. Mahmod Al-Qattan MN, Mordi MN. Molecular Basis of Modulating Adenosine Receptors Activities. Current pharmaceutical design. 2019;25(7):817-31.

9. Bajusz D, Rácz A, Héberger K. 3.14 - Chemical Data Formats, Fingerprints, and Other Molecular Descriptions for Database Analysis and Searching. In: Chackalamannil S, Rotella D, Ward SE, editors. Comprehensive Medicinal Chemistry III. Oxford: Elsevier; 2017. p. 329-78.

10. Rogers D, Hahn M. Extended-connectivity fingerprints. J Chem Inf Model. 2010;50(5):742-54.

11. Durán Á, Zamora I, Pastor M. Suitability of GRIND-Based Principal Properties for the Description of Molecular Similarity and Ligand-Based Virtual Screening. Journal of Chemical Information and Modeling. 2009;49(9):2129-38.

12. Pastor M, Cruciani G, McLay I, Pickett S, Clementi S. GRid-INdependent Descriptors (GRIND): A Novel Class of Alignment-Independent Three-Dimensional Molecular Descriptors. Journal of Medicinal Chemistry. 2000;43(17):3233-43.

13. Tsou LK, Yeh SH, Ueng SH, Chang CP, Song JS, Wu MH, et al. Comparative study between deep learning and QSAR classifications for TNBC inhibitors and novel GPCR agonist discovery. Sci Rep. 2020;10(1):16771.

14. Mouhibi R, Zahouily M, El Akri K. Using multiple linear regression and artificial neural network techniques for predicting CCR5 binding affinity of substituted 1-(3, 3-Diphenylpropyl)-Piperidinyl amides and ureas. 2013.

15. Nguyen ATN, Nguyen DTN, Koh HY, Toskov J, MacLean W, Xu A, et al. The application of artificial intelligence to accelerate G protein-coupled receptor drug discovery. British Journal of Pharmacology. 2024;181(14):2371-84.

16. Liu S, Liu T, Gao L, Li H, Hu Q, Zhao J, et al. Convolutional Neural Network and Guided Filtering for SAR Image Denoising. Remote Sensing. 2019;11(6):702.

17. Kwon S, Bae H, Jo J, Yoon S. Comprehensive ensemble in QSAR prediction for drug discovery. BMC Bioinformatics. 2019;20(1):521.

18. Goh GB, Siegel C, Vishnu A, Hodas NO, Baker N. Chemception: a deep neural network with minimal chemistry knowledge matches the performance of expert-developed QSAR/QSPR models. arXiv preprint arXiv:170606689. 2017.

19. Hentabli H, Bengherbia B, Saeed F, Salim N, Nafea I, Toubal A, et al. Convolutional Neural Network Model Based on 2D Fingerprint for

Bioactivity Prediction. International journal of molecular sciences. 2022;23(21).
20. Huo X, Xu J, Xu M, Chen H. An improved 3D quantitative structure-activity relationships (QSAR) of molecules with CNN-based partial least squares model. Artificial Intelligence in the Life Sciences. 2023;3:100065.
21. Shan W, Li X, Yao H, Lin K. Convolutional neural network-based virtual screening. Current Medicinal Chemistry. 2021;28(10):2033-47.
22. Munawar S, Windley MJ, Tse EG, Todd MH, Hill AP, Vandenberg JI, et al. Experimentally Validated Pharmacoinformatics Approach to Predict hERG Inhibition Potential of New Chemical Entities. Front Pharmacol. 2018;9:1035.
23. Ismatullah H, Jabeen I. Combined pharmacophore and grid-independent molecular descriptors (GRIND) analysis to probe 3D features of inositol 1, 4, 5-trisphosphate receptor (IP3R) inhibitors in cancer. International journal of molecular sciences. 2021;22(23):12993.
24. Khan HA, Jabeen I. Combined Machine Learning and GRID-Independent Molecular Descriptor (GRIND) Models to Probe the Activity Profiles of 5-Lipoxygenase Activating Protein Inhibitors. Frontiers in Pharmacology. 2022;13.
25. Shiri F, Pirhadi S, Ghasemi JB. Alignment independent 3D-QSAR, quantum calculations and molecular docking of Mer specific tyrosine kinase inhibitors as anticancer drugs. Saudi pharmaceutical journal : SPJ : the official publication of the Saudi Pharmaceutical Society. 2016;24(2):197-212.
26. Manouchehrizadeh E, Mostoufi A, Tahanpesar E, Fereidoonnezhad M. Alignment-independent 3D-QSAR and molecular docking studies of tacrine-4-oxo-4H-Chromene hybrids as anti-Alzheimer's agents. Computational biology and chemistry. 2019;80:463-71.
27. Moriwaki H, Tian YS, Kawashita N, Takagi T. Three-Dimensional Classification Structure-Activity Relationship Analysis Using Convolutional Neural Network. Chemical & pharmaceutical bulletin. 2019;67(5):426-32.
28. Bento AP, Hersey A, Félix E, Landrum G, Gaulton A, Atkinson F, et al. An open source chemical structure curation pipeline using RDKit. Journal of Cheminformatics. 2020;12(1):51.
29. Liu S, Alnammi M, Ericksen SS, Voter AF, Ananiev GE, Keck JL, et al. Practical Model Selection for Prospective Virtual Screening. Journal of Chemical Information and Modeling. 2019;59(1):282-93.
30. Sakamuru S, Zhao J, Xia M, Hong H, Simeonov A, Vaisman I, et al. Predictive Models to Identify Small Molecule Activators and Inhibitors of Opioid Receptors. Journal of Chemical Information and Modeling. 2021;61(6):2675-85.
31. Richardson E, Trevizani R, Greenbaum JA, Carter H, Nielsen M, Peters B. The receiver operating characteristic curve accurately assesses imbalanced datasets. Patterns (New York, NY). 2024;5(6):100994.
32. Qian X, Dai X, Luo L, Lin M, Xu Y, Zhao Y, et al. An Interpretable Multitask Framework BiLAT Enables Accurate Prediction of Cyclin-Dependent Protein Kinase Inhibitors. Journal of Chemical Information and Modeling. 2023;63(11):3350-68.
33. 33. Lopez-del Rio A, Picart-Armada S, Perera-Lluna A. Balancing Data on Deep Learning-Based Proteochemometric Activity Classification. Journal of Chemical Information and Modeling. 2021;61(4):1657-69.
34. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters. 2006;27(8):861-74.
35. Chicco D, Tötsch N, Jurman G. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. BioData Mining. 2021;14(1):13.
36. Fontaine F, Pastor M, Sanz F. Incorporating Molecular Shape into the Alignment-free GRid-INdependent Descriptors. Journal of Medicinal Chemistry. 2004;47(11):2805-15.
37. Shim J, Coop A, MacKerell AD, Jr. Molecular details of the activation of the μ opioid receptor. The journal of physical chemistry B. 2013;117(26):7907-17.
38. Kajino K, Tokuda A, Saitoh T. Morphinan Evolution: The Impact of Advances in Biochemistry and Molecular Biology. The Journal of Biochemistry. 2024;175(4):337-55.

استعمال الذكاء الاصطناعي لترشيح المركبات الفعالة تجاه مستقبلات الأفيون عبر البصمات الجزيئية والشبكة العصبية التلافيفية

محمد نورالدين محمود القطان[١]

[١]فرع الكيمياء الصيدلانية، كلية الصيدلة، جامعة نينوى، الموصل، العراق.

الخلاصة

تنتمي مستقبلات الأفيونات إلى عائلة المستقبلات المقترنة بالبروتين ، وتنقسم إلى ثلاثة أنواع فرعية هي: ميو(Mu) ، وكابا(Kappa) ، ودلتا (Delta). وتُستهدف هذه المستقبلات بواسطة ناهضات (Agonists) ومناهضات (Antagonists) لعلاج مجموعة متنوعة من الحالات الطبية. وعلى الرغم من أن الروابط الذاتية الرئيسية لهذه المستقبلات هي ببتيدات داخلية المنشأ، إلا أن مجموعة متنوعة من الجزيئات غير الببتيدية يمكنها أيضًا الارتباط بالمواقع الأرثوستيرية والألّوستيرية في المستقبل. إلا أن المسح الافتراضي للناهضات والمناهضات الفعالة باستخدام الالتحام الجزيئي القائم على البنية الجزيئية يواجه تحديات تتعلق بمرونة المستقبل، وتعدّد الوحدات، وتنسيق المواقع الأرثوستيرية/الألّوستيرية. بالمقابل، فإن المسح الافتراضي القائم على خصائص المركب الرقمية يكون أكثر موثوقية. في هذا البحث، تم استخدام بصمات جزيئية غير معتمدة على المحاذاة (GRIND) والشبكة العصبية الالتفافية (CNN) لتصنيف قاعدة بيانات من مركبات تم قياس نشاط اتحادها مع مستقبلات الافيون تجريبيًا. تم تدريب ستة نماذج GRIND-CNN للتعرف على الميزات داخل البصمة الجزيئية التي ترتبط بالنشاطات الناهضة والمناهضة لكل من مستقبلات ميو، وكابا، ودلتا. وقد تم تقييم أداء النماذج باستخدام المساحة تحت منحنى المميز لاداء المستقبل (AUC-ROC) ، ومعامل ارتباط ماثيوز (MCC)، والدقة المتوازنة (BA) . أظهرت نماذج GRIND-CNN متوسط قيم AUC-ROC تراوح بين (٠,٧٠ إلى ٠,٨٣)، وهي قيم مقاربة لما تم الحصول عليه باستعمال البصمات الجزيئية من نوع PubChem و Extended-Connectivityالمستخدمة في نماذج الشبكات العصبية العميقة PubChem-DNN) و ECFP-DNNعلى التوالي .إلا أن نماذج GRIND-CNN و PubChem-DNN تفوقت على نموذج ECFP-DNN في التعرف على الميزات المرتبطة بالنشاط داخل البصمات، حيث كانت متوسطات قيم MCC للشبكات المدربة باستعمال معامل عشوائي Y-randomizedعبر جميع مجموعات المركبات هي ٠,٠٥ و٠,٠٣ و٠,١٦ على التوالي .وعند استخدام تنبؤات مجمعة من كلا النموذجين GRIND-CNN و PubChem-DNN بمعامل ٠,٥ لكل منهما، تحسنت قيمة AUC-ROC (من ٠,٧٧ إلى ٠,٩١)، وانخفضت قيمة MCC في تدريب Y-randomized إلى ٠,٠٢، مما يشير إلى تحسن في التعرف على الميزات. كما أظهرت خرائط الأهمية لبصمات GRIND، في النماذج المدربة على التمييز بين الناهضات والمناهضات، وبين المركبات القادرة على الاتحاد وغير الغير قادرة على الاتحاد بالمستقبلات الأفيونات المختلفة، وجود فروق إحصائية معنوية على مستوى وحدات المخطط الارتباطي (البكسلات). وترتبط هذه الفروق بنظرية "الرسالة-العنوان" في تفسير آلية ارتباط الروابط الأفيونية. يوفر هذا البحث أدوات للمسح الافتراضي للروابط الأفيونية بناءً على التعرف على ميزات مرتبطة بالنشاط من خلال البصمة الجزيئية، بغض النظر عن الهيكل الجزيئي أو الحجم، بطريقة تشبه كيفية استشعار اليد لشكل جسم ما عبر أصابعها المختلفة. ان البرامج والموديلات المطورة في هذا البحث هي متوفرة للباحثين على الرابط:
https://github.com/mohammednooraldeen/GRIND-CNN.