

METHODOLOGY

Open Access



CNN-DDI: a learning-based method for predicting drug–drug interactions using convolution neural networks

Chengcheng Zhang^{1†}, Yao Lu^{2†} and Tianyi Zang^{1*} 

From Biological Ontologies and Knowledge bases workshop 2020 Virtual. 16 December 2020

*Correspondence:

tianyi.zang@hit.edu.cn

[†]Chengcheng Zhang and Yao Lu have contributed equally to this work and should be considered co-first authors

¹ Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
Full list of author information is available at the end of the article

Abstract

Background: Drug–drug interactions (DDIs) are the reactions between drugs. They are compartmentalized into three types: synergistic, antagonistic and no reaction. As a rapidly developing technology, predicting DDIs-associated events is getting more and more attention and application in drug development and disease diagnosis fields. In this work, we study not only whether the two drugs interact, but also specific interaction types. And we propose a learning-based method using convolution neural networks to learn feature representations and predict DDIs.

Results: In this paper, we proposed a novel algorithm using a CNN architecture, named CNN-DDI, to predict drug–drug interactions. First, we extract feature interactions from drug categories, targets, pathways and enzymes as feature vectors and employ the Jaccard similarity as the measurement of drugs similarity. Then, based on the representation of features, we build a new convolution neural network as the DDIs' predictor.

Conclusion: The experimental results indicate that drug categories is effective as a new feature type applied to CNN-DDI method. And using multiple features is more informative and more effective than single feature. It can be concluded that CNN-DDI has more superiority than other existing algorithms on task of predicting DDIs.

Keywords: Drug–drug interactions, Drug categories, Convolutional neural network, Multiple features combination

Background

Drug–drug interactions (DDIs) mean the reactions between drugs. They are compartmentalized into three types: synergistic, antagonistic and no reaction [1–3]. The DDIs play a significant role in drug development and disease diagnosis fields, which still consumes manpower, substance sources and time [4].

Powered by advanced machine learning technology, methods of DDIs' prediction have been evolved from traditional methods [5–7], including text mining methods and



statistical methods, to machine learning methods. Furthermore, more and more studies use deep learning methods in the field of bio-informatics [8–15].

The task of predicting DDIs is vitally interrelated with similarities between drugs. The fundamental hypothesis of this task is that if drug A and drug B interact each other, causing a specific biological impact, drugs have similarity to drug A (or drug B) are possible to interact with drug B (or drug A) and causes same effect [16].

Cami et al. [17] utilized a logistic regression model to solve the DDIs' problem. On this basis, Gottfried et al. [18] exploited more different drug–drug similarities and proposed another logistic regression model. Two similarity-based models based on drug interaction profile fingerprints were proposed [16, 19] and a heterogeneous network-assisted inference framework was introduced by Cheng et al. [20]. Some other algorithms were extended on the task of DDIs' prediction. For instance, TMFUF [12] is based on the triple matrix factorization, DDINMF [21] is based on the semi-nonnegative matrix factorization. Three algorithms were proposed in [22], including neighbor recommender algorithm, random walk algorithm, and the matrix perturbation algorithm. Further, they proposed a novel algorithms named 'Manifold Regularized Matrix Factorization'. In 2019, SFLLN was proposed in [23] based on linear neighborhood regularization using four types of drug features. It is a sparse feature learning ensemble method.

DeepDDI was proposed [10] to classify the DDIs' events from DrugBank [24]. DeepDDI calculates features' similarity and reduces features' dimension by principal component analysis (PCA). Lee et al. [25] concentrated on concrete types of two drugs, not simply whether they interact or not. DDIMDL [11] is a multimodal deep neural network algorithm, which combines diverse drug features that predicting 65 types of DDI events.

Convolutional neural network (CNN) is a typical artificial neural network based on supervised learning, which has good performance on computer vision filed [26]. And it develops more network structures from CNN. They have been used extensively in bio-informatics [27, 28]. Many studies apply deep learning method in the task of DDIs' prediction, and most of them choose deep neural network (DNN). But compared with deep neural network, CNN performs better in feature learning and can alleviate the degree of over-fitting effectively. Considering features selected contain noise and advantages of CNN, we decide to use CNN to solve the problem of DDIs' prediction.

In this paper, we propose a novel algorithm based on CNN, named CNN-DDI, to learn the best combination of drug features and predict DDI-associated events. CNN-DDI method contains two parts. One part is a feature selection framework. We utilize drug categories as another feature, and choose the best combination form of drug features. The other part is a CNN-based DDI's predictor. We utilize a new CNN to predict DDI-associated events based on features pairs selected from feature selection framework.

Results and discussion

Evaluation criteria

Predicting DDI' events can be regarded as a multi-label classification problem. Therefore, the prediction results are divided into four kinds, true positive (TP), false positive (FP), true negative (TN) and false negative (FN). In addition, precision and recall criteria are common used evaluation criteria, which can evaluate the accuracy of results. Precision means in the classified positive samples, the proportion of TP samples. And recall

means in all positive samples, the proportion of correct samples classified. The expressions are as follows:

$$precision = \frac{TP}{TP + FP} \quad (1)$$

$$recall = \frac{TP}{TP + FN} \quad (2)$$

Based on precision and recall, Accuracy, F1-score, area under the precision-recall curve (AUPR) and area under the ROC curve (AUC) are utilized to evaluate the performance of the algorithm.

In the study, we adopt Accuracy, F1-score, micro-averaged AUPR and micro-averaged AUC as the evaluation metrics. Micro-averaged metrics means metrics are averaged after getting the results of all classes.

Performance

To analyze the effect of different similarity algorithms on performance of CNN-DDI, we utilize cosine similarity, Jaccard similarity and Gaussian similarity to calculate features' similarities. Table 1 shows the experimental results of our method on three similarity measures. It can be seen that using different similarity measures exhibits similar properties. CNN-DDI is robust to these three similarity measures, so Jaccard similarity measure is used in the experiments.

To demonstrate the superiority of drug categories and influence of different combination forms, we further test the performance of CNN-DDI model with different features' types. The experimental results are shown in Table 2. As for one feature, CNN-DDI using drug categories as the feature performs best, the AUPR score using drug categories is 0.9139, which is quite higher than the second highest score produced by drug targets (the value is 0.8470). Similarly, using drug categories achieves the highest scores of other five evaluation metrics. So the drug category is effective as a new feature type applied to CNN-DDI method. On the whole, using multiple features is informative and helps CNN-DDI perform better than single feature. The combination of four features has the highest AUPR score (the value is 0.9251) in all combinations. Thus it can be proved that every feature improves the performance of CNN-DDI to a certain extent.

Comparison experiments

We evaluate the effectiveness of our algorithm and four state-of-art algorithms. The four algorithms are random forest (RF), gradient boosting decision tree (GBDT), logistic regression (LR) and K-nearest neighbor (KNN). We measure feature

Table 1 The experimental results of CNN-DDI on three similarity measures

Similarity	ACC	AUPR	AUC	F1	Precision	Recall
Jaccard	0.8871	0.9251	0.9980	0.7496	0.8556	0.7220
Cosine	0.8871	0.9251	0.9979	0.7492	0.8855	0.7721
Gaussian	0.8870	0.9248	0.9979	0.7489	0.8859	0.7720

Table 2 Results of CNN-DDI using different features

Feature	ACC	AUPR	AUC	F1	Precision	Recall
T	0.7915	0.8470	0.9953	0.6099	0.6932	0.5716
P	0.7820	0.8381	0.9952	0.5805	0.6822	0.5364
E	0.6580	0.7098	0.9897	0.3344	0.4419	0.2957
C	0.8702	0.9139	0.9966	0.7421	0.7994	0.7125
T+P	0.8227	0.8898	0.9969	0.6778	0.7589	0.6375
T+E	0.8242	0.8712	0.9956	0.6360	0.7373	0.5849
T+C	0.8792	0.9185	0.9960	0.7627	0.8167	0.7405
P+E	0.8255	0.8747	0.9958	0.6227	0.7130	0.5781
P+C	0.8796	0.9179	0.9961	0.7440	0.7955	0.7485
E+C	0.8496	0.8895	0.9948	0.6928	0.7726	0.6488
T+P+E	0.8243	0.8690	0.9947	0.6489	0.7332	0.6063
T+P+C	0.8797	0.9199	0.9960	0.7490	0.8164	0.7232
T+E+C	0.8539	0.8899	0.9933	0.6938	0.7726	0.6539
P+E+C	0.8559	0.8919	0.9939	0.6845	0.7575	0.6485
T+P+E+C	0.8871	0.9251	0.9980	0.7496	0.8556	0.7220

The bold values indicate the result of CNN-DDI with four types of features. So it can be concluded that the drug category is effective as a new feature type and multiple features can improve the performance of CNN-DDI

Table 3 Results of CNN-DDI and other state-of-art models

Algorithm	ACC	AUPR	AUC	F1	Precision	Recall
CNN-DDI	0.8871	0.9251	0.9980	0.7496	0.8556	0.7220
GBDT	0.8327	0.8828	0.9970	0.6730	0.7817	0.6133
RF	0.7837	0.8446	0.9959	0.5167	0.6973	0.4444
KNN	0.7581	0.8166	0.9881	0.6250	0.7562	0.5596
LR	0.7558	0.8087	0.9950	0.3894	0.5617	0.3331

similarities in the same manner. In the experiment, we set the decision tree number of RF to be 100 and the neighbor number of KNN to be 4.

Table 3 shows CNN-DDI algorithm has better performance than other four methods in these 6 accuracy assessments. The score of ACC is 0.8871, it is better than the score of GBDT, RF, KNN and LR (0.8327, 0.7837, 0.7581 and 0.7558 respectively). And other evaluation metrics achieved by CNN-DDI are 0.9251, 0.9980, 0.7496, 0.8556 and 0.7220, respectively, which are significantly higher than the scores of other methods. The LR algorithm gets the worst performance, whose scores are 0.7558, 0.8087, 0.9950, 0.3894, 0.5617 and 0.3331, respectively. Compared with GBDT, which gets the second best performance, the ACC score is 0.8871, increased by 6.53%. And the score of AUPR is 0.9251, increased by 4.79%, all of other evaluation metrics have been improved in varying degrees.

And we compare our algorithm with DDIMDL. Considering DDIMDL using different features, we retrain DDIMDL model with features selected by CNN-DDI. As shown in Table 4, DDIMDL represents the original algorithm proposed by original paper [11]. DDIMDL* represents DDIMDL with features selected by CNN-DDI. It can be concluded that the drug category is effective as a new feature type, and CNN-DDI still performs better than DDIMDL in the case of using the same features.

Table 4 Comparison of CNN-DDI with DDIMDL

Algorithm	ACC	AUPR	AUC	F1	Precision	Recall
CNN-DDI	0.8871	0.9251	0.9980	0.7496	0.8556	0.7220
DDIMDL	0.8852	0.9208	0.9976	0.7585	0.8471	0.7182
DDIMDL*	0.8865	0.9230	0.9976	0.7559	0.8513	0.7204

The single asterisk represent DDIMDL with features selected by our method. It can be concluded that the drug category is effective as a new feature type

Conclusions

In the work, we proposed a novel semi-supervised algorithm using a CNN architecture, named CNN-DDI, to predict drug–drug interactions. First, we extract feature interactions from drug categories, targets, pathways and enzymes as feature vectors. Then, based on the representation of feature, we proposed a new convolution neural network as the predictor of DDIs-associated events. The predictor consists of five convolutional layers, two full-connected layers and a softmax layer based on CNN.

To demonstrate the performance of our method, we compare it with other start-of-the-art methods. The evaluation shows our method, CNN-DDI, has better performance than other existing state-of-art measures. Meanwhile, we discuss the contribution of combinational features and each single feature. Overall, CNN-DDI has more advantages on predicting DDIs' events. In consideration of consuming longer time, we will try to improve the efficiency of CNN-DDI in the future.

Methods

We propose a novel method called CNN-DDI to predict DDI-associated events. The method mainly contain two parts, combinational features selection module and CNN-based prediction module. As shown in Fig. 1, we combine four drug features and obtain a low dimensional as the CNN model inputs. Then a deep CNN model is built to calculate the probability of DDIs' types. In this section, we will thoroughly expound the structure and principle of CNN-DDI.

Data collection

DDIMDL proposed a data set that classifying DDIs' events into 65 types, not simply focusing on whether they interact or not. The data set includes 572 drugs and 74,528 DDIs-associated events collected from DrugBank. Which is a manually collected data source that provides drugs comprehensive information and unified syntax in describing DDIs.

To extend the information of DDIMDL, we extract drugs categories from DrugBank. 572 drugs have 1622 types of categories in DDIMDL.

In our paper, cross validation is utilized to demonstrate the effectiveness of our method. We set the fold number of cross validation is 5. In our experiments, we randomly divide the data set into five subsets, choose four subsets as the train set and another one as the test set. We test on the data set five times following the above steps, and the final result is the average of multiple results.

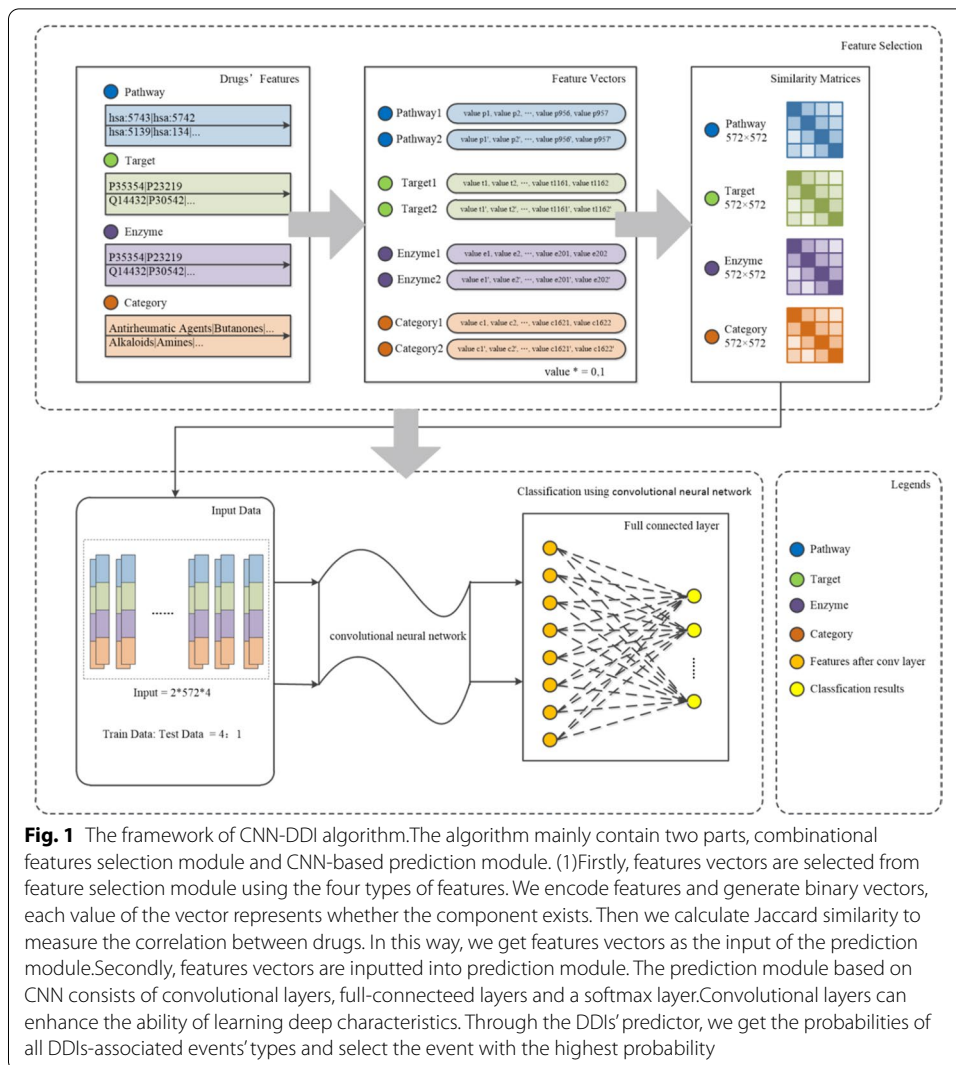


Fig. 1 The framework of CNN-DDI algorithm. The algorithm mainly contains two parts, a combinational features selection module and a CNN-based prediction module. (1) Firstly, feature vectors are selected from the feature selection module using the four types of features. We encode features and generate binary vectors, where each value of the vector represents whether the component exists. Then we calculate Jaccard similarity to measure the correlation between drugs. In this way, we get feature vectors as the input of the prediction module. Secondly, feature vectors are inputted into the prediction module. The prediction module based on CNN consists of convolutional layers, fully connected layers, and a softmax layer. Convolutional layers can enhance the ability of learning deep characteristics. Through the DDI predictor, we get the probabilities of all DDI-associated event types and select the event with the highest probability

CNN-DDI algorithm

Drug-drug similarity

There are three common similarity measures, Jaccard similarity, cosine similarity, and Gaussian similarity. To better measure the drug feature vectors' similarity, we analyze the difference of measures' results. Jaccard similarity calculates the intersection of components and the union. Gaussian similarity utilizes the Gaussian kernel function. And cosine similarity is used to calculate the cosine between two vectors in an inner product space [29].

Jaccard similarity can be calculated as follows:

$$\begin{aligned}
 Sim_f(x_i, x_j) &= Sim_f(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \\
 &= \frac{M_{11}}{M_{01} + M_{10} + M_{11}}
 \end{aligned}
 \tag{3}$$

where x_i and x_j are feature vectors of two drugs, X and Y are the vector sets respectively. $|X \cup Y|$ represents the union of X and Y , $|X \cap Y|$ represents the intersection. Further, M represents the number of elements. Subscript 11 means the elements where x_i and x_j are 1, 01 means elements where x_i is 0 and x_j is 1, 10 means elements where x_i is 1 and x_j is 0.

Cosine similarity can be calculated as follows:

$$Sim_c(x_i, x_j) = \frac{x_i \cdot x_j}{\|x_i\| \|x_j\|} \quad (4)$$

where $\|\cdot\|$ represents the Euclidean norm.

Gaussian similarity can be calculated as follows:

$$Sim_G(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (5)$$

where γ represents hyper parameters. And $\gamma = 1 / \left(\sum_{i=1}^n |x_i| / n \right)$.

Feature selection module

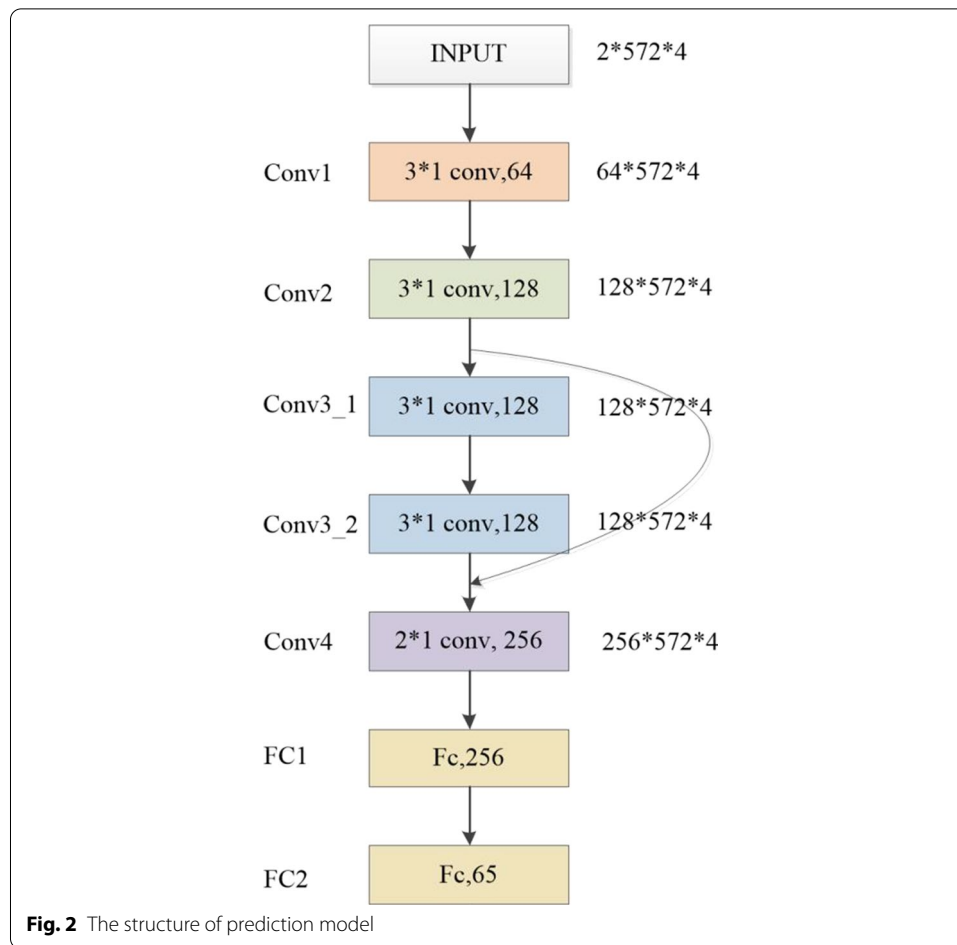
Firstly, we evaluate the similarity between two drugs. The feature selection includes two steps: (1) calculating the similarity scores to evaluate correlation between drugs. (2) Generating feature vectors as the input to the prediction module.

The drugs' feature can be represented as a binary vector, the value is 1 or 0. Value 1 means presence of components, value 0 means absence. For instance, the data set has 1622 types of categories. So the categories can be expressed as a 1622-dimensional bit vector, the value means that the drug belongs to the category or not. Similarly, we can extract four binary feature vectors from one drug corresponding four features. Then we calculate the similarity between two drugs' feature vectors by similarity measures. By this means, similarity matrices are generated as $S = (s_{ij})$, where the value of s_{ij} is from 0 to 1. The closer the value is to 1, the higher the similar degree of drugs.

Prediction module using convolutional neural network

As shown in Fig. 1, CNN-based prediction module is the important part to predict DDIs' events. Features selected from selection module are input vectors into the prediction module. Considering features selected contain noise and advantages of CNN, we decide to use CNN in the prediction module.

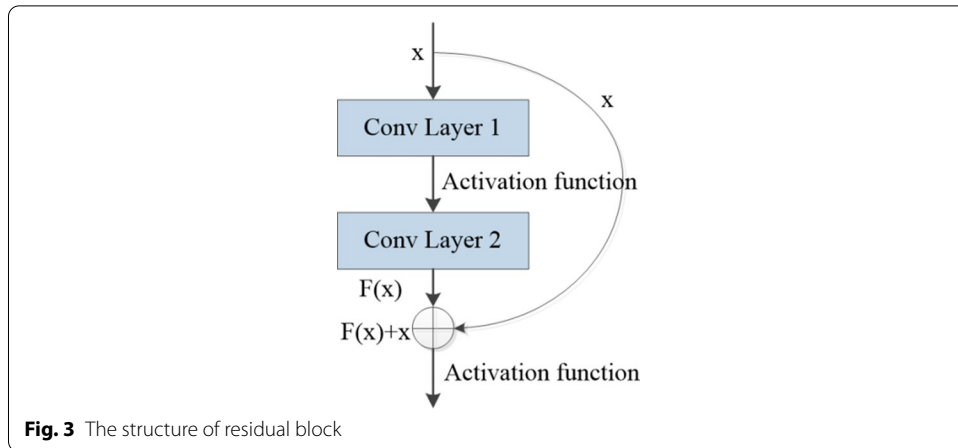
CNN is widely used and performs well on computer vision, like image classification, image detection and image segmentation. And powered by advanced deep learning technology, more and more studies have explored its application in bio-informatics field [30]. Compared with the pure deep neural network, CNN has the following advantages: (1) the convolutional layer has less parameters by using connections' sparsity and parameters sharing. (2) The convolutional layer extracts information from global features and local features. On the task of DDIs' prediction, Results of classification are strongly related to not only global drug features but also part of features combination. So it can enhance the capability of features learn. Consequently, in this article, we apply CNN as the supervised model for distilling integrated features information to predict DDIs.

**Table 5** The convolution layers of CNN-DDI

Layer name	number of filters	Kernel size	Output shape
Conv1	64	3×1	(64, 572, 4)
Conv2	128	3×1	(128, 572, 4)
Conv3_1	128	3×1	(128, 572, 4)
Conv3_2	128	3×1	(128, 572, 4)
Conv4	256	3×1	(256, 572, 4)

The structure of prediction model is shown in Fig. 2. The prediction model based on CNN includes five convolutional layers, two full-connected layers and a softmax layer. Among them, convolutional layers are mainly responsible for subspace feature extraction from the input vectors. Table 5 shows the specific configuration. The kernel size of each convolutional layer is same (3×1), and the filters' number is increasing layer-by-layer.

In addition, we add a residual block [31] to build one short connection between two layers. Figure 3 shows the structure of residual block. The output of residual block is expressed as follows:



$$y = F(x) + x = W_2[\sigma_1(W_1x + b_1)] + b_2 + x \tag{6}$$

where x is the input vectors, y is the output vectors. W_1, W_2 are the weight vectors of two layers, b_1, b_2 are the biases, and σ_1 is the activation function of first layer.

The residual block strengthen the correlation of multi-layer features. The short connection’s input vectors and output vectors must have the same dimensions, and the stacked convolutional layers’ output vectors are added together. It should be noted that no additional parameters are added in the residual block.

The output of each convolutional layer is passed through an activation function that enhances positive vectors and inhibits negative vectors from previous layer. In the paper, the activation function we use is Leaky ReLU. Compare with other activations, ReLU can increase feature sparsity and decrease the possibility of vanishing gradient. The expression is as follows:

$$LeakyRelu(x) = \begin{cases} x, & x \geq 0 \\ ax, & x < 0 \end{cases} \tag{7}$$

where a represents hyper-parameters, a is set 0.2.

There are two full-connected layers after convolutional layers. The first full-connected layer has 267 hidden units and the second has 65 hidden units. Considering predicting DDI’s events is a classification task, softmax function is used as the activation of the last full-connected layer. So the loss function of the prediction module is as follows:

$$Loss = - \sum_{i=1}^K y_i \log(p_i) \tag{8}$$

where K represents the number of events’ types, y_i represents the true value, 0 or 1.

The CNN-DDI algorithm

The algorithm mainly contain two parts, combinational features selection module and CNN-based prediction module. The pseudocode of CNN-DDI is shown in Algorithm 1.

Algorithm 1: The CNN-DDI algorithm

Input: All features of drugs F_i ($i = 1, 2, \dots, n$) including categories, targets, pathway, pathways and enzymes;

Output: Prediction results.

Run feature selection module:

1. For all $i \in N$ do
2. Calculate Jaccard similarity of drugs' feature vectors to measure drugs similarity;
3. Optimize data from high dimension to low dimension;
4. End for

Run prediction module

5. Put feature vectors of the train set as the predictor input;
6. Train CNN-DDI predictor to optimize Eq.(8);
7. Put feature vectors of the test set as the predictor input;
8. Through forward propagation, the model output the classification scores of all possible DDIs' events.

Return: The DDIs' event having the highest classification score as the prediction results.

Abbreviations

DDIs: Drug–drug interactions; CNN: Convolutional neural network; DNN: Deep neural network; AUPR: Area under the precision-recall curve; AUC: Area under the ROC curve; RF: Random forest; GBDT: Gradient boosting decision tree; LR: Logistic regression; KNN: K-nearest neighbor; PCA: Principal component analysis.

Acknowledgements

The authors thank the anonymous referees for their many useful suggestions.

About this supplement

This article has been published as part of BMC Bioinformatics Volume 23 Supplement 1, 2022: Selected articles from the Biological Ontologies and Knowledge bases workshop 2020. The full contents of the supplement are available online at <https://bmcbioinformatics.biomedcentral.com/articles/supplements/volume-23-supplement-1>.

Authors' contributions

TYZ helped revise this paper. CCZ and YL did the study's experiments and wrote the paper. CCZ and YL contributed equally to this work. All authors have read and approved the final manuscript.

Funding

Publication costs are funded by the National Key Research and Development Program of China (2017YFC0907503). The funder had no role in the design of the work, data collection, data analysis, and writing the manuscript.

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the drugbank repository and DDIMDL repository. <https://go.drugbank.com/> <https://github.com/YifanDengWHU/DDIMDL>.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China. ²General Hospital of Heilongjiang Province Land Reclamation Bureau, Harbin, China.

Received: 10 February 2022 Accepted: 14 February 2022

Published online: 07 March 2022

References

1. Liu S, Tang B, Chen Q, et al. Drug–drug interaction extraction via convolutional neural networks. *Comput Math Methods Med*. 2016;56:1–8.
2. Hiroyuki K. How far should we go? Perspective of drug–drug interaction studies in drug development. *Drug Metab Pharmacokinet*. 2014;29(3):227–8.
3. Percha B, Altman RB. Informatics confronts drug–drug interactions. *Trends Pharmacol Sci*. 2013;34(3):178–84.
4. Fang H, Chen X, Pei X, Grant S, Tan M. Experimental design and statistical analysis for three-drug combination studies. *Stat Methods Med Res*. 2015;26(3):1261–80.
5. Isabel S, Paloma M, César S. Extracting drug–drug interactions from biomedical texts. *BMC Bioinform*. 2010;11(5):9.
6. Yan S, Jiang X, Chen Y. Text mining driven drug–drug interaction detection. In: 2013 IEEE international conference on bioinformatics and biomedicine. IEEE. 2013;349–54.
7. Tari L, Anwar S, Liang S, Cai J, Baral C. Discovering drug–drug interactions: a text-mining and reasoning approach based on properties of drug metabolism. *Bioinformatics*. 2010;26(18):547–53.
8. Zhao T, Hu Y, Peng J, Cheng L. DeepLGP: a novel deep learning method for prioritizing lncRNA target genes. *Bioinformatics*. 2020;36(16):4466–72.
9. Zhao T, Hu Y, Cheng L. Deep-DRM: a computational method for identifying disease-related metabolites based on graph deep learning approaches. *Brief Bioinform*. 2020;36(16):4466–72.
10. Ryu JY, Kim HU, Sang YL. Deep learning improves prediction of drug–drug and drug–food interactions. *Proc Natl Acad Sci*. 2018;115(18):201803294.
11. Deng Y, Xu X, Qiu Y, Xia J, Liu S. A multimodal deep learning framework for predicting drug–drug interaction events. In: 2020 15th IEEE international conference on automatic face and gesture. 2020.
12. Shi J, Huang H, Lin J, et al. Tmfuf: a triple matrix factorization-based unified framework for predicting comprehensive drug–drug interactions of new drugs. *BMC Bioinform*. 2018;19(Suppl 14):411.
13. Peng J, Guan J, Hui W, et al. A novel subnetwork representation learning method for uncovering disease–disease relationships. *Methods*. 2020.
14. Zhao T, Liu J, Zeng X, et al. Prediction and collection of protein–metabolite interactions. *Brief Bioinform*. 2021.
15. Peng J, Xue H, Wei Z, Tuncali I, Hao J, Shang X. Integrating multi-network topology for gene function prediction using deep neural networks. *Brief Bioinform*. 2021;22(2):2096–105.
16. Vilar S, Harpaz R, et al. Drug–drug interaction through molecular structure similarity analysis. *J Am Med Inform Assoc*. 2012;19(6):1066–74.
17. Cami A, Manzi S, Arnold A, Reis BY, Medina MA. Pharmacointeraction network models predict unknown drug–drug interactions. *PLoS ONE*. 2013;8(4):e61468.
18. Assarf G, Gideon S, Oran Y, et al. Indi: a computational framework for inferring drug interactions and their associated recommendations. *Mol Syst Biol*. 2012;17(8):592.
19. Vilar S, Uriarte E, Santana L, et al. Similarity-based modeling in large-scale prediction of drug–drug interactions. *Nat Protoc*. 2014;9(9):2147–63.
20. Cheng F, Zhao Z. Machine learning-based prediction of drug–drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *J Am Med Inform Assoc*. 2014;21(e2):e278–86.
21. Zhang W, Chen Y, Liu F, Luo F, Tian G, Li X. Predicting potential drug–drug interactions by integrating chemical, biological, phenotypic and network data. *BMC Bioinform*. 2017;18(1):18.
22. Yu H, Mao K, Shi J, et al. Predicting and understanding comprehensive drug–drug interactions via semi-nonnegative matrix factorization. *BMC Syst Biol*. 2018;12(S1):14.
23. Zhang WA, Jing KC, Huang FB. SFLLN: A sparse feature learning ensemble method with linear neighborhood regularization for predicting drug–drug interactions. *Inf Sci*. 2019;497(23):189–201.
24. Wiwshart D, et al. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018;46(D1):D1074–82.
25. Lee G, Park C, Ahn J. Novel deep learning model for more accurate prediction of drug–drug interaction effects. *BMC Bioinform*. 2019;20(1):415.
26. Alex K, Llya S, et al. Imagenet classification with deep convolutional neural network. *Commun ACM*. 2017;6(6):94–90.
27. Zhao T, Yang H, et al. Identifying drug–target interactions based on graph convolutional network and deep neural network. *Brief Bioinform*. 2020;22(2):2141–50.
28. Peng J, Wang Y, Guan J, Li J, Han R, Hao J, Wei Z, Shang X. An end-to-end heterogeneous graph representation learning-based framework for drug–target interaction prediction. *Brief Bioinform*. 2021.
29. Deepika S, Geetha TV. Drug side effect prediction through linear neighborhoods and multiple data source integration. *J Biomed Inform*. 2019;84:136–47.
30. Zhuang Z, Pan W, Shen X. A simple convolutional neural network for prediction of enhancer–promoter interactions with dna sequence data. *Bioinformatics*. 2019;35(17):2899–906.
31. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. In: Conference on computer vision and pattern recognition (CVPR). 2016.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.