



Article Knowledge-Graph-Based Drug Repositioning against COVID-19 by Graph Convolutional Network with Attention Mechanism

Mingxuan Che¹, Kui Yao², Chao Che^{2,*}, Zhangwei Cao³ and Fanchen Kong³

- ¹ Department of Information Engineering, Dalian University, Dalian 116622, China; chemingxuan@s.dlu.edu.cn
- ² Key Laboratory of Advanced Design and Intelligent Computing, Ministry of Education, Dalian University, Dalian 116622, China; yaokui@s.dlu.edu.cn
- ³ Department of Software Engineering, Dalian University, Dalian 116622, China; caozhangwei@s.dlu.edu.cn (Z.C.); kongfanchen@s.dlu.edu.cn (F.K.)
- Correspondence: chechao@dlu.edu.cn; Tel.: +86-0411-8740-2046

Abstract: The current global crisis caused by COVID-19 almost halted normal life in most parts of the world. Due to the long development cycle for new drugs, drug repositioning becomes an effective method of screening drugs for COVID-19. To find suitable drugs for COVID-19, we add COVID-19-related information into our medical knowledge graph and utilize a knowledge-graph-based drug repositioning method to screen potential therapeutic drugs for COVID-19. Specific steps are as follows. Firstly, the information about COVID-19 is collected from the latest published literature, and gene targets of COVID-19 are added to the knowledge graph. Then, the information of COVID-19 of the knowledge graph is extracted and a drug–disease interaction prediction model based on Graph Convolutional Network with Attention (Att-GCN) is established. Att-GCN is used to extract features from the knowledge graph and the prediction matrix reconstructed through matrix operation. We evaluate the model by predicting drugs for both ordinary diseases and COVID-19. The model can achieve area under curve (AUC) of 0.954 and area under the precise recall area curve (AUPR) of 0.851 for ordinary diseases. On the drug repositioning experiment for COVID-19, five drugs predicted by the models have proved effective in clinical treatment. The experimental results confirm that the model can predict drug–disease interaction effectively for both normal diseases and COVID-19.

Keywords: COVID-19; drug–disease interaction prediction; knowledge graph; graph convolutional network

1. Introduction

Coronavirus disease 2019 (COVID-19) has been listed as an international public health emergency by WHO [1], and on March 11, it was defined as a global "pandemic". As of December 16, more than 20.58 million people were infected worldwide. As the number of COVID-19 patients is dramatically increasing worldwide, treatment in intensive care units (ICUs) has also become a major challenge [2]. Under the current circumstance of the absence of specific vaccines and medicines against COVID-19, it is urgent to discover effective therapies, especially drugs, to treat COVID-19. Considering that it usually takes 10-15 years to develop a new drug, probably the best strategy for the treatment of COVID-19 is drug repositioning. Drug repositioning, also known as "new use of old drugs" and "re-examination of old drugs", refers to the discovery of new indications or new uses for drugs already on the market, including repositioning, repurposed, and repurposed drugs that are in the clinical research stage or approved for marketing evaluation, reorientation of treatment direction, etc. Under normal circumstances, it takes 10-15 years for new drug development from the determination of the idea to the drug market, and there are uncertainties such as safety and pharmacokinetics so that R&D costs and risks of drug development are significant. However, since drugs used for drug repositioning studies usually have passed



Citation: Che, M.; Yao, K.; Che, C.; Cao, Z.; Kong, F. Knowledge-Graph-Based Drug Repositioning against COVID-19 by Graph Convolutional Network with Attention Mechanism. *Future Internet* **2021**, *13*, 13. https:// doi.org/10.3390/fi13010013

Received: 15 November 2020 Accepted: 28 December 2020 Published: 7 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/). several stages of clinical trials or are already on the market, their risks are lower compared with strategies such as developing from scratch and obtaining patent licenses. In addition, compared with obtaining patent licenses and restructuring strategies, it has the advantages of a short time to market and a greater possibility of discovering differences in drug effects, so higher returns are expected. Therefore, drug repositioning is one of the strategies with the best risk/benefit ratio among currently known drug development strategies.

As many drugs treat disease by acting on related targets, most drug repositioning research studies predict new drug-disease interaction (DDI) by discovering new drugtarget interaction (DTI). However, for a new disease, because the corresponding target of the disease has not been fully discovered, if drug repositioning is achieved only through the discovery of new DTI, the results may be limited. In addition, the existing data of known drugs and related information, such as proteins and genes, are huge, so it is particularly important to use appropriate databases and preliminary screening of data when redirecting drugs to an emerging disease. Wearable devices and internet medical service platforms generate a large number of high-dimensional medical data. Traditional machine learning methods cannot be used to process high-dimensional medical data from different sources. Deep learning methods are widely used in feature extraction [3] and disease prediction [4] for medical data. In the drug redirection problem, the drug and target naturally form a graph structure, and the most used deep learning model is the Graph Convolutional Network (GCN).

In order to find drug candidates against COVID-19, we construct a knowledge graph (KG) for COVID-19 and propose a new model called Graph Convolutional Network with Attentional mechanism for Drug-Disease Interaction (Att-GCN-DDI) to predict potential therapeutic drugs for COVID-19. We first collect information about COVID-19 in the latest published literature and add the gene targets of COVID-19 to our drug KG. Then, we screen the nodes and relationships associated with COVID-19 to build the KG. Borrowing the idea from Neo-DTI [5] and DTI-NET [6], the Att-GCN-DDI model first employs GCN with attentional mechanism to extract features from KG and then performs matrix factorization for DDI prediction. The tests on two scenarios in DDI prediction have demonstrated that Att-GCN-DDI can significantly outperform several baseline prediction methods. Att-GCN-DDI also has good performance on the DDI prediction against COVID-19. Five drug candidates predicted by Att-GCN-DDI have proved effective in the clinical treatment.

The main contributions of our work lie in:

- We gather and add the target of COVID-19 to our KG. Then, we select the related (1)knowledge to construct a KG for COVID-19, which is applied to find the potential therapeutic drugs against COVID-19.
- (2)We propose a GCN-based model for drug repositioning on KG. The model can learn the topology around the disease effectively, which is utilized to predict new drugs for the disease.
- (3) We evaluate our method by predicting drugs for both ordinary diseases and COVID-19. Att-GCN-DDI finds five effective drugs against COVID-19, which have been proved in clinical treatment and outperforms five other baseline models in the drug repositioning for ordinary diseases. The experimental results confirm the strong predictive power of Att-GCN-DDI.

2. Related Work

The most important step in drug repositioning is to find novel drug–disease or drug– target interactions. In order to achieve this goal, various methods have been developed, including computational approaches, biological experimental approaches, and mixed approaches [7]. In recent years, researchers mainly use computational methods to realize drug repositioning because biological experimental methods cost a lot of money and time [8]. Now the most widely used computing methods include the following four categories: network-based methods, knowledge graph embedding-based method, text mining methods and biological feature-based methods [9].

Text mining methods extract useful information from known literature to find the novel interactions between drugs and diseases. "MAM" [10], "PharmGKB" [11] and "Chem2bio2rdf" [12] are all based on semantic similarity to measure the relationship between drug and diseases. Recently, researchers began to use machine learning methods to achieve this goal. Fu et al. [13] proposed a semantic similarity framework using random forest (RF) and support vector machine (SVM) methods. However, the diversity of language expression and the contradictory information found in the literature limit the performance of the text mining method [14].

The biological feature-based method realizes drug repositioning by using machine learning approaches to extract the biological feature of drugs and targets [15]. These methods usually include two key parts: feature extraction and relationship prediction. "SimBoost" [16] trains a gradient-enhanced machine model to learn the similarity between drugs and proteins to understand their binding affinity. "NRLMF" [17] uses the similarity between drugs and proteins to simulate the probability of drug–target interaction through logical matrix decomposition [18]. These methods improve the accuracy of DTI prediction to a certain extent. However, these methods do not take drug–drug or protein–protein interactions into account [9].

The method based on the knowledge graph embedding (KGE) maps the entities and relationships in the knowledge graph to a low-dimensional continuous vector space, which can retain the inherent characteristics of the knowledge graph and alleviate the feature sparse problem that may be faced in the application of the knowledge graph. The training process is divided into multiple stages. First, the KGE model uses random noise to initialize the embedding vector. Then, the loss error is calculated by the score function, and training is performed. GrEDeL [19] uses TransE to learn the embedding vector based on a biomedical knowledge graph by exploiting the relations extracted from biomedical abstracts. Then, the Long Short-Term Memory (LSTM) Networks model is trained to discover candidate drugs for diseases of interest from biomedical literature. TriModel [20] is an extension of DistMultand and ComplEx models, using three embedding vectors to represent each entity and relationship. These methods have high requirements on the quality of knowledge graphs and are suitable for finding new associations between drugs and targets that have been fully studied; however, they are not suitable for drug repositioning for emerging diseases due to incomplete disease-related information in the knowledge graph.

In recent years, the network-based method has been widely used. This method mainly includes three steps: network construction, feature extraction and relationship prediction [9]. The network-based method calculates the similarity between drugs and targets based on the network topology, and the purpose of this method is to predict unknown interactions based on known interactions [21]. The basic principle is that drugs tend to combine with similar targets or diseases. DDR [22] constructed drug-drug interaction network and protein-protein interaction network based on the similarity between proteins and drugs [18]. They then used the RF method to predict the combination of drugs and proteins. DTI-NET [6], MSCMF(Multiple Similarities Collaborative Matrix Factorization) [23] and HNM(Heterogeneous Network Model) [24] can further improve the accuracy of DTI prediction by integrating information from heterogeneous data sources and improving the relationship prediction method. Neo-DTI [5] uses a new feature extraction scheme based on DTI-NET [6] to enhance the accuracy. However, the existing research on network-based method is limited to the prediction of drug-target interactions, so, the mining of heterogeneous network is not thorough enough. In addition, there is still a lot of data loss in the process of feature extraction.

3. Data Acquisition and Processing

3.1. The Drug KG

We adopted the drug KG that was built in our previous study, which integrated six pharmaceutical knowledge bases: DrugBank [25], KEGG DRUG [26], TTD [27], DID [28], PharmGKB [29] and SIDER [30]. We first analyze the original data in the knowledge

base to extract the triples, and then insert the data according to the graphical data model integration data triples to obtain the knowledge graph [31]. The KG contains five types of entities, including drugs, genes, diseases, channels, side effects, and nine relationships among them. The data schema of our drug KG is shown in Figure 1.



Figure 1. Data schema of our drug knowledge graph (KG).

3.2. Acquisition of COVID-19 Information

Nowadays, drugs against COVID-19 are divided into two categories according to their targets (genes) [32]. The first is to act on the immune cells of the human body to enhance the immune function of human. The second is acting on the COVID-19 itself, binding receptors, and the enzymes needed for its replication. Through gathering the information of COVID-19, we have screened out four targets (genes) of COVID-19 with clear function and high reliability. They are RNA dependent RNA polymer (RdRp) [32], ACE2 [32], pp1ab [33], human immunity virus type 1 protection (pol) [34]. Therefore, we link the COVID-19 entity to KG through four drug–gene interactions: COVID-19-RdRp, COVID-19-ACE2, COVID-19-pp1ab and COVID-19-pol.

3.3. Construction of the KG for COVID-19

Our drug KG contains more than 100,000 entities and more than 670,000 relationships. It is extremely difficult to perform a computation consuming model GCN on such a largescale KG. In addition, the information of the drugs and proteins that are not related to COVID-19 in the KG will also interfere with drug repositioning. To reduce the calculation amount and improve the accuracy of the DDI prediction, we need to construct a KG for COVID-19 by extracting the related knowledge from the drug KG.

In our drug KG, if a drug can treat a disease, there is usually a path with a distance less than 4 between them beside the direct connection. The path between them contains the information of why the drug can treat the disease. In addition, Att-GCN-DDI predicts DDIs through the similarity of topological structure, and diseases with similar topological structure are usually connected by the path with distance less than 4, such as disease–drug–disease, disease–gene–disease and disease–gene–disease.

Therefore, in this paper, we focus on the COVID-19 node, and select the drugs and disease nodes whose shortest path distance from the COVID-19 node is less than 4. After that, we use these drugs and disease nodes as the initial data to supplement the related gene, side effect and pathway nodes. Finally, the drug–disease drug–gene, gene–pathway, drug–drug, drug–side effect, gene–gene, and disease–gene relationships among nodes were supplemented too. Tables 1 and 2 show the number of entities and relationship of the KG for COVID-19.

Table 1. The number of entities of each type in the KG for COVID-19.

Entity Type	Drug	Disease	Gene	Side Effect	Pathway
Number	1470	752	1741	274	53

Relation Type	Drug- Disease	Drug-Gene	Gene– Pathway	Drug–Drug	Drug–Side Effect	Gene-Gene	Disease- Gene
Number	1659	1898	62	921	1432	263	877

Table 2. The number of relations of each type in the KG for COVID-19.

4. Method

We designed a model called Att-GCN-DDI to discover unknown DDIs based on the drug KG. The workflow of Att-GCN-DDI is shown in Figure 2. Att-GCN-DDI mainly includes three main steps: (a) node embedding based on Att-GCN; (b) topology-preserving learning of the node embedding; (c) reconstruction of DDI matrix. Through step (a), the topological features of each node in the KG are extracted into an F-dimension vector, and the feature vectors of all drugs and diseases constitute the drug feature matrix and disease feature matrix, respectively (where X is the drug feature matrix, where each row represents the feature vector of a drug, and Y is the disease feature matrix, where each row represents the feature vector of a disease). Through step (b), we attempt to find an optimal projection from the drug space to the protein space by supervised learning so that the mapped drug feature vectors geometrically approach the diseases of their known interactions. The projection matrix Z is supervised by known drug-disease interactions and learns to minimize the difference between the known interaction matrix P and XZY^T. Then, Att-GCN-DDI performs matrix operations on the projection matrix Z and feature matrix by step (c), and finally reconstructs DDIs matrix. Then, we can get the novel DDIs based on the reconstructed DDI matrix.



Figure 2. The workflow of the Graph Convolutional Network with Attentional mechanism for Drug–Disease Interaction (Att-GCN-DDI).

Next, we will introduce the mathematical formulations of these three main steps.

The given KG is defined as an undirected graph $G = (\mathbf{V}, \mathbf{E}) \mathbf{V} = \{v_1, v_2, ..., v_n\}$ is the set of nodes and $\mathbf{E} = \{e_1, e_2, ..., e_m\}$ is the set of edges; where *m* is the number of nodes, *n* is the number of edges and $\mathbf{E} \in \mathbf{V} \times \mathbf{V}$. The adjacency matrix **A** is usually represented in binary, 1 means that there is a connection between nodes, 0 means that there is no connection between nodes.

The key step of topological feature extraction using Att-GCN is to construct the Laplace matrix, and in our model, the Laplacian matrix should be:

$$L = D - A = I_n - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$$
(1)

where I_n is an identity matrix and D is the inverse degree matrix [9].

Finally, the topological feature of each node in the heterogeneous network can be extracted using the following formula:

$$\mathbf{M} = \operatorname{Re}Lu(\mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}\mathbf{X})$$
(2)

where **M** is a feature vector of the input entity, and **X** is the characteristic of each node itself.

The extracted features are used to form a drug feature matrix F_{drug} and a disease feature matrix $F_{disease}$ in which each row represents a drug or disease feature vector **M**. Then, we use supervised learning to find the most appropriate projection matrix **Z**. The learning objectives are as follows:

$$\min_{w} \left\| \mathbf{P} - \mathbf{F}_{drug} \mathbf{Z}_{w} \mathbf{F}_{disease}^{T} \right\|^{2} + \lambda \left\| w \right\|_{1}$$
(3)

where **P** is the known DDI matrix. Note that the same matrix construction method is also used in [1,2] to solve the problem of relationship prediction. In order to achieve quick convergence, we use $\lambda ||w||_{1}$ as the regularizer.

Then, we introduced the attention mechanism to assign weights to the feature matrix. The calculation formula of attention is as follows:

Softmax
$$(\mathbf{QF}_{\mathbf{drug}}^{\mathbf{I}})\mathbf{F}_{\mathbf{drug}}$$
 (4)

where **Q** is the embedding representation of drug features.

After the supervised learning of the projection matrix, the process of reconstructing the DDI matrix is as follows:

$$\mathbf{F}_{\mathrm{ddi_reconstruct}} = \mathbf{F}_{\mathrm{drug}} \mathbf{Z} \mathbf{F}_{\mathrm{disease}}^{\mathrm{T}}$$
(5)

5. Experiments

Our model is tested on two experiments. First, we employ the Att-GCN-DDI model to predict drug candidates for COVID-19 and conduct a case study based on predicted drugs. Then we compare Att-GCN-DDI with 5 baseline models on the drug repositioning experiment for ordinary diseases.

5.1. DDI Prediction for COVID-19

5.1.1. Results

In the experiment, we used the Att-GCN-DDI model to predict drug candidates against COVID-19. Here, we take COVID-19 KG as the input. Since the KG does not include the interactions between COVID-19 and drugs, we can use all known DDIs as the training set to train our model and finally reconstruct the prediction matrix.

Then, we extracted the predictive scores for COVID-19 and all drugs from the reconstructed matrix and ranked the drugs by the scores. The score here reflects the strength of the underlying interaction between a specific drug and COVID-19. Here, we extracted the top 30 drugs as drug candidates against COVID-19. They are Efavirenz, Lamivudine, Stavudine, Abacavir, Nevirapine, Tipranavir, Roquinimex, Zalcitabine, Delavirdine, Emtricitabine, Didanosine, Tenofovir, Lopinavir, Amprenavir, Zidovudine, Saquinavir, Darunavir, Ritonavir, Atazanavir, Indinavir, Moexipril, Rilpivirine, Cefroxadine, Brecanavir, Lisinopril, Ribavirin, Pentanal, Alfaxalone, 5-[(5-fluoro-3-methyl-1H-indazol-4-yl)oxy]benzene-1,3dicarbonitrile, SPP1148.

After analysis and search, 5 drugs among them have been clinically proven to be viable for COVID-19 treatment. Their information is shown in Table 3. These results indicate that the drug candidates against COVID-19 predicted by Att-GCN-DDI are basically reliable.

DrugBank Id	Drug Name	Source of Clinical Feasibility
DB00300	Tenofovir	Published medical literature [35]
DB01601	Lopinavir	Experimental drugs for COVID-19 in DrugBank [34]
DB01264	Darunavir	Experimental drugs for COVID-19 in DrugBank [34]
DB00503	Ritonavir	Published Treatment Protocol in China [36]
DB00811	Ribavirin	Published Treatment Protocol in China [36]

Table 3. Clinically proven drugs in our drug candidates.

5.1.2. Case Study

We analyzed the path between COVID-19 and our drug candidates in the KG to understand why these drugs are more likely to treat COVID-19 than others. The path can be divided into two types. The first is to directly connect COVID-19 through genes. In this case, we can think of drugs acting on COVID-19 related genes to treat COVID-19. The second is to link diseases directly without genes. Take drug Tipranavir as an example. The paths between this drug and COVID-19 in KG are shown in Figure 3. It can be found that although this drug does not directly act on COVID-19-related genes, drugs related to Tipranavir directly act on COVID-19-related genes, ACE2 and pol, respectively. Therefore, we believe that this drug is indeed more likely to have a therapeutic relationship with COVID-19 than other drugs.



Figure 3. Path between Tipranavir and COVID-19.

5.2. DDI Prediction for Other Diseases

At present, only a small number of medicines have proven to have a therapeutic effect on COVID-19. Therefore, only performing DDI experiments on COVID-19 cannot fully verify the effectiveness of the Att-GCN-DDI model Therefore, we also test our model on DDI prediction for other diseases.

The DDI prediction is a binary classification problem, in which known DDIs are considered as positive examples and unknown DDIs are considered as negative examples. In the experiment, we used COVID-19 KG and first ran a 10-fold cross-validation test on all positive examples and a set of randomly sampled negative examples, which were 10 times as many as positive samples. This scenario basically mimicked the practical situation of drug repositioning [5]. For each fold, 90% of randomly chosen positive and negative examples were used as the training set, and the remaining 10% of positive and negative examples were regarded as the test set.

We then compared the performance of Att-GCN-DDI with GCN-DDI, Neo-DTI [5], DTI-NET [6] and HNM [24] in predicting DDIs. GCN-DDI indicates that the GCN model in our framework does not use attention mechanism. The area under the precise recall area curve (AUPR) and the area under receiver operating characteristic curve (AUROC) were used to evaluate the predictive performance of all prediction methods. The evaluation results are shown in Figure 4. The results show that Att-GCN-DDI is superior to other

methods in AUROC and AUPR. Att-GCN-DDI adopts the Att-GCN model to extract features by node embedding, so the extracted features can better retain the topology structure of nodes. This makes the prediction more accurate.



Figure 4. Comparison results of Att-GCN-DDI and previous methods in the first scenario.

Next, we tested Att-GCN-DDI in another scenario by including all positive and negative examples in the 10-fold cross-validation procedure. The evaluation results are shown in Figure 5. We find that the performance of Att-GCN-DDI in this scenario is also superior to other methods. However, the AUPR of GCN-DDI is very close to GCN-DDI. The main reason for this situation is that with the expansion of data volume, the demand of the model for accurate topology feature extraction decreases. Therefore, compared with other methods, the application of GCN for feature extraction has no obvious advantages. From Figures 4 and 5, we can see that the Att-GCN-DDI using attention mechanism performs obviously better than GCN-DDI in two scenarios. This is because the attention mechanism can flexibly capture the relationship between global information and local information. This is very important for extracting feature information of topology structure in the knowledge graph.



Figure 5. Comparison results of Att-GCN-DDI and previous methods in the second scenario.

6. Discussion

This paper collects COVID-19 information and inserts it into the existing medical KG to build a KG of COVID-19. Based on the KG, the GCN-based drug repositioning model is used to predict potential therapeutic drugs for COVID-19. We conduct drug repositioning experiments on COVID-19 and other diseases, respectively. Our model ultimately identified 30 potential drugs for COVID-19 treatment, of which five have proven to be effective clinically. On the DDI prediction experiments for other diseases, our model outperforms other baseline methods. Our work provides help for the preliminary screening

of drugs in the face of new diseases and helps medical staff to screen out potential drugs for new diseases in the shortest time. However, our research also has some shortcomings. For example, the GCN model can effectively learn the structural information and the relation between nodes in the knowledge map, but it cannot learn the representation of the relation, let alone the directivity of the relation. Therefore, in the future, we will learn from the structure of GraphSAGE and the GAT(Graph Attention Network) model to improve the ability of our model in relation representation.

Author Contributions: Conceptualization, M.C. and C.C.; methodology, M.C. and K.Y.; validation, M.C., K.Y. and Z.C.; data curation, M.C. and Z.C.; writing—original draft preparation, M.C. and K.Y.; writing—review and editing, M.C. and C.C.; visualization, F.K.; project administration, C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No.62076045) and the Guidance Program of Liaoning Natural Science Foundation (No.2019-ZD-0569).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/FangpingWan/NeoDTI.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Velavan, T.P.; Meyer, C.G. The COVID-19 epidemic. Trop. Med. Int. Health 2020, 25, 278–280. [CrossRef] [PubMed]
- Velavan, T.P.; Meyer, C.G. Mild versus severe COVID-19: Laboratory markers. Int. J. Infect. Dis. 2020, 95, 304–307. [CrossRef] [PubMed]
- 3. Ali, F.; El-Sappagh, S.; Islam, S.M.R.; Ali, A.; Attique, M.; Imran, M.; Kwak, K.-S. An intelligent healthcare monitoring framework using wearable sensors and social networking data. *Futur. Gener. Comput. Syst.* **2021**, *114*, 23–43. [CrossRef]
- Ali, F.; El-Sappagh, S.H.A.; Islam, S.M.R.; Kwak, D.; Ali, A.; Imran, M.; Kwak, K.S. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Inf. Fusion* 2020, 63, 208–222. [CrossRef]
- 5. Wan, F.; Hong, L.; Xiao, A.; Jiang, T.; Zeng, J. NeoDTI: Neural integration of neighbor information from a heterogeneous network for discovering new drug-target interactions. *Bioinformatics* **2018**, *35*, 104–111. [CrossRef] [PubMed]
- Luo, Y.; Zhao, X.; Zhou, J.; Yang, J.; Zhang, Y.; Kuang, W.; Peng, J.; Chen, L.; Zeng, J. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nat. Commun.* 2017, *8*, 1–13. [CrossRef]
- Xue, H.; Li, J.; Xie, H.; Wang, Y. Review of Drug Repositioning Approaches and Resources. Int. J. Biol. Sci. 2018, 14, 1232–1244. [CrossRef]
- Mathur, A.; Loskill, P.; Shao, K.; Huebsch, N.; Hong, S.; Marcus, S.G.; Marks, N.C.; A Mandegar, M.; Conklin, B.R.; Lee, L.P.; et al. Human iPSC-based Cardiac Microphysiological System For Drug Screening Applications. *Sci. Rep.* 2015, *5*, srep08883. [CrossRef]
- 9. Zhao, T.; Hu, Y.; Valsdottir, L.R.; Zang, T.; Peng, J. Identifying drug–target interactions based on graph convolutional network and deep neural network. *Briefings Bioinform.* 2020, bbaa044. [CrossRef]
- 10. Zhu, S.; Okuno, Y.; Tsujimoto, G.; Mamitsuka, H. A probabilistic model for mining implicit 'chemical compound-gene' relations from literature. *Bioinformatics* **2005**, *21*, ii245–ii251. [CrossRef]
- 11. Hewett, M.; Oliver, D.E.; Rubin, D.L.; Easton, K.L.; Stuart, J.M.; Altman, R.B.; Klein, T.E. PharmGKB: The Pharmacogenetics Knowledge Base. *Nucleic Acids Res.* 2002, *30*, 163–165. [CrossRef] [PubMed]
- 12. Chen, B.; Dong, X.; Jiao, D.; Wang, H.; Zhu, Q.; Ding, Y.; Wild, D.J. Chem2Bio2RDF: A semantic framework for linking and data mining chemogenomic and systems chemical biology data. *BMC Bioinform.* **2010**, *11*, 255. [CrossRef] [PubMed]
- 13. Fu, G.; Ding, Y.; Seal, A.; Chen, B.; Sun, Y.; Bolton, E. Predicting drug target interactions using meta-path-based semantic network analysis. *BMC Bioinform.* **2016**, *17*, 1–10. [CrossRef]
- 14. Fotis, C.; Antoranz, A.; Hatziavramidis, D.; Sakellaropoulos, T.; Alexopoulos, L. Network-based technologies for early drug discovery. *Drug Discov. Today* **2018**, *23*, 626–635. [CrossRef]
- 15. Lavecchia, A. Machine-learning approaches in drug discovery: Methods and applications. *Drug Discov. Today* **2015**, *20*, 318–331. [CrossRef] [PubMed]
- 16. He, T.; Heidemeyer, M.; Ban, F.; Cherkasov, A.; Ester, M. SimBoost: A read-across approach for predicting drug–target binding affinities using gradient boosting machines. *J. Chemin.* **2017**, *9*, 1–14. [CrossRef] [PubMed]
- 17. Liu, Y.; Wu, M.; Miao, C.; Zhao, P.; Li, X.-L. Neighborhood Regularized Logistic Matrix Factorization for Drug-Target Interaction Prediction. *PLoS Comput. Biol.* 2016, 12, e1004760. [CrossRef] [PubMed]

- Mei, J.-P.; Kwoh, C.-K.; Yang, P.; Li, X.-L.; Zheng, J. Drug–target interaction prediction by learning from local information and neighbors. *Bioinformatics* 2013, 29, 238–245. [CrossRef]
- 19. Sang, S.; Yang, Z.; Liu, X.; Wang, L.; Lin, H.F.; Wang, J.; Dumontier, M. GrEDeL: A Knowledge Graph Embedding Based Method for Drug Discovery from Biomedical Literatures. *IEEE Access* **2019**, *7*, 8404–8415. [CrossRef]
- Mohamed, S.K.; Nováček, V.; Nounu, A. Discovering Protein Drug Targets Using Knowledge Graph Embeddings. *Bioinformatics* 2019, 36, 603–610. [CrossRef]
- 21. Pei, J.; Yin, N.; Ma, X.; Lai, L. Systems Biology Brings New Dimensions for Structure-Based Drug Design. J. Am. Chem. Soc. 2014, 136, 11556–11565. [CrossRef] [PubMed]
- 22. Olayan, R.S.; Ashoor, H.; Bajic, V.B. DDR: Efficient computational method to predict drug-target interactions using graph mining and machine learning approaches. *Bioinformatics* **2018**, *34*, 1164–1173. [CrossRef]
- Zheng, X.; Ding, H.; Mamitsuka, H.; Zhu, S. Collaborative matrix factorization with multiple similarities for predicting drug-target interactions. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; ACM Inc.: New York, NY, USA, 2013; pp. 1025–1033. [CrossRef]
- 24. Wang, W.; Yang, S.; Zhang, X.; Li, J. Drug repositioning by integrating target information through a heterogeneous network model. *Bioinformatics* **2014**, *30*, 2923–2930. [CrossRef] [PubMed]
- 25. Whirl-Carrillo, M.; McDonagh, E.M.; Hebert, J.M.; Gong, L.; Sangkuhl, K.; Thorn, C.F.; Altman, R.B.; Klein, T.E. Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* **2012**, *92*, 414–417. [CrossRef] [PubMed]
- Yang, H.; Qin, C.; Li, Y.H.; Tao, L.; Zhou, J.; Yu, C.Y.; Xu, F.; Chen, Z.; Zhu, F.; Chen, Y.Z. Therapeutic target database update 2016: Enriched resource for bench to clinical drug target and targeted pathway information. *Nucleic Acids Res.* 2016, 44, D1069–D1074. [CrossRef]
- 27. Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **2008**, *36*, D480–D484. [CrossRef]
- 28. Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A.C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V.; et al. Drugbank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Res.* **2014**, *42*, D1091–D1097. [CrossRef]
- 29. Kuhn, M.; Letunic, I.; Jensen, L.J.; Bork, P. The SIDER database of drugs and side effects. *Nucleic Acids Res.* 2016, 44, D1075–D1079. [CrossRef]
- 30. Sharp, M.E. Toward a comprehensive drug ontology: Extraction of drug-indication relations from diverse information sources. *J. Biomed Semant.* **2017**, *8*, 2. [CrossRef]
- 31. Zhu, Y.; Che, C.; Jin, B.; Zhang, N.; Su, C.; Wang, F. Knowledge-driven drug repurposing using a comprehensive drug knowledge graph. *Health Inform. J.* 2020. [CrossRef]
- Wu, C.; Liu, Y.; Yang, Y.; Zhang, P.; Zhong, W.; Wang, Y.; Wang, Q.; Xu, Y.; Li, M.; Li, X.; et al. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm. Sin. B* 2020, *10*, 766–788. [CrossRef] [PubMed]
- 33. Liu, Q.; Wang, X. Strategies for the development of drugs targeting novel coronavirus 2019-nCoV. *Acta Pharm. Sin. B* 2020, 55, 181–188. [CrossRef]
- 34. COVID-19 Dashboard | DrugBank Online. Available online: https://www.drugbank.ca/\protect\unhbox\voidb@x\ hbox{COVID-19} (accessed on 14 November 2020).
- 35. Zhang, C.; Chen, S.; Zhang, J.; Guo, Y. Analysis of chemical drugs applied for clinical trial for the treatment of COVID-19. *Acta Pharm. Sin. B* 2020, *55*, 355–365. [CrossRef]
- Wei, P. Iagnosis and Treatment Protocol for Novel Coronavirus Pneumonia (Trial Version 7). *Chin. Med. J.* 2020, 133, 1087–1095. [CrossRef]