A network representation approach for COVID-19 drug recommendation

Haifeng Liu [a], Hongfei Lin [a,*], Chen Shen [a], Liang Yang [a], Yuan Lin [b], Bo Xu [a], Zhihao Yang [a], Jian Wang [a], Yuanyuan Sun [a]

[a] Department of Computer Science, Dalian University of Technology, LiaoNing, China
[b] Faculty of Humanities and Social Sciences, Dalian University of Technology, LiaoNing, China

ARTICLE INFO

ABSTRACT

The coronavirus disease 2019 (COVID-19) has outbreak since early December 2019, and COVID-19 has caused over 100 million cases and 2 million deaths around the world. After one year of the COVID-19 outbreak, there is no certain and approve medicine against it. Drug repositioning has become one line of scientific research that is being pursued to develop an effective drug. However, due to the lack of COVID-19 data, there is still no specific drug repositioning targeting the COVID-19. In this paper, we propose a framework for COVID-19 drug repositioning. This framework has several advantages that can be exploited: one is that a local graph aggregating representation is used across a heterogeneous network to address the data sparsity problem; another is the multi-hop neighbors of the heterogeneous graph are aggregated to recall as many COVID-19 potential drugs as possible. Our experimental results show that our COVDR framework performs significantly better than baseline methods, and the docking simulation verifies that our three potential drugs have the ability to against COVID-19 disease.

1. Introduction

SARS-CoV-2 is the virus that causes COVID-19, the respiratory illness responsible for the SARS-CoV-2 pandemic. As of February 2021, the SARS-CoV-2 has affected over 190 countries, leading to over 100 million global cases with over 2 million deaths (fatality rate of 2.2%). The COVID-19 pandemic has far-reaching economic consequences beyond the spread of the disease, which leads to global economic recession with an estimated over 500 billion dollars and 25 million jobs [1]. The severe impact of the COVID-19 pandemic results in people needing a treatment drug for COVID-19. However, there is still no specific medicine for COVID-19. Even if the vaccine can prevent people from contracting COVID-19, the confirmed patients cannot be treated with the vaccine, we need to explore an effective medicine for curing SARS-CoV-2. To accelerate drug development, drug repositionings have received great attention in recent years.

With the wide application of web-based clinical information systems and clinical disease data to increase rapidly, it is possible to find hidden relationships between diseases and existing drugs with drug repositioning. Meanwhile, many drugs have more than one target, which is intended in the drug design stage. The targets that are not originally intended are called "off-target". Such off-target interactions, though may cause adverse drug reactions (ADRs), can provide opportunities to seek new use of existing drugs in drug discovery.

Most drug repositionings have focused on machine learning approach, network analysis approach, and text mining of semantic inference approach [2], such as [3–5]. However, most methods are inapplicable for COVID-19. The above methods have two main challenges for SARS-CoV-2 drug repositioning: One is that the most current methods often focus on only one aspect of drug repositionings and lacks systematic consideration of multiple interactive networks, such as drug-target interaction prediction [1] or drug-disease association prediction [6]. How to integrate the SARS-COV-2 heterogeneous network to explore hidden drugs is a challenge. The other is the SARS-CoV-2 is a novel virus, and there is currently no effective treatment for SARS-CoV-2. Fig. 1 displays the 3D structure of the spike protein bound with the host receptor angiotensin-converting enznyme2 (ACE2) in SARS-CoV-2 (PDB ID:6CS2). Recent virus research demonstrates that SARS-CoV-2 utilizes ACE2 as an entry receptor in ACE2-expressing cells [7], suggesting potential drug targets for therapeutic development. It is possible to search SARS-CoV-2 potential drugs from ACE2 protein docking small molecules.

In this paper, we propose a practical and effective framework (named COVDR) to methodically infer new drug-disease interaction for SARS-

Fig. 1. The 3D structure of the spike protein bound with the host receptor ACE2 in SARS-CoV-2.

CoV-2 drug repositioning. The target of COVDR is to fuse diverse information from a heterogeneous network and infer potential drugs for SARS-CoV-2. Since there is no published drug repositioning dataset for SARS-CoV-2 when we started to research, we construct a drug repositioning dataset based on DrugBank, repoDB, News reports, and literatures. Then, we construct a heterogeneous graph of COVID-19 to integrate diversified information of COVID-19, such as the drug-drug network, drug-disease network, and drug-target network. At the same time, we utilized the graph convolution function to aggregate local network information in a heterogeneous network, which promotes the drug, disease, and target non-linear feature interaction and improve predict accuracy. Finally, we use the drug docking poses simulation to evaluate the Top-N recommended drugs, and the simulation demonstrated that our potential drugs can protect ACE2 targets from SARS-CoV-2 infection.

In summary, our major contributions are threefold:

Fig. 2. Our framework for SARS-CoV-2 drug repositioning.

- We build a COVID-19 based drug repositioning dataset for SARS-CoV-2 potential drug prediction.
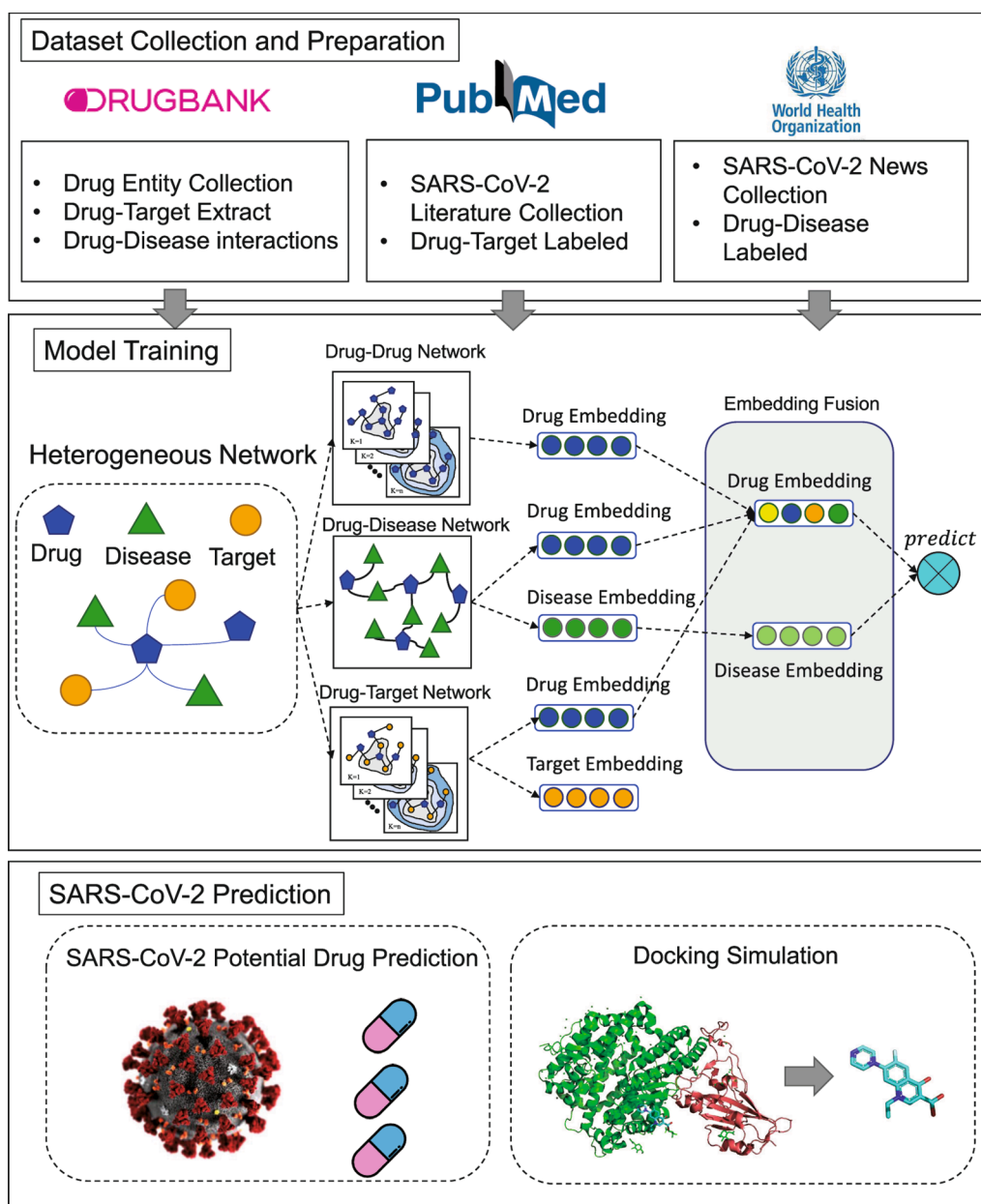- We propose a novel drug repositioning framework, named COVDR, which use a heterogeneous graph convolutional network to infer potential drug for COVID-19.
- The extensive experiments and the drug docking poses simulation demonstrated that COVDR can effectively predict potential drugs for COVID-19.

2. Related work

Generally, drug repositioning can be divided into three groups: The machine learning approach, the network analysis approach, and the text mining of semantic inference approach [2]. In recent years, many artificial intelligence methods are widely applied for drug repositioning, For instance, Yang et al. [3] proposed a probabilistic matrix factorization approach to infer drug-disease associations for drug repositioning. As the drug-disease interaction relations is a kind of network, some researchers trends to used network representation learning methods to predict drug-disease interaction. Menden et al. [8] proposed network representation models to predict the response of cancer cell lines to drug treatment through IC_{50} values. Zeng et al. [9] developed a networks-based deep learning approach to integrate heterogeneous networks for human disease. Luo et al. [4] proposed a random walk model with drug-disease-protein to predict potential drugs for diseases. Some heterogeneous network-based methods are proposed for drug repositioning. For example, Yang et al. propose a heterogeneous graph inference with matrix completion (HGIMC) method to predict potential indications for approved and novel drugs [10]. Wang et al. [11] proposed a computational framework based on a heterogeneous network model and applied the approach on drug repositioning by using existing omics data about diseases, drugs, and drug targets. Different from network-based methods, text mining of semantic inference approach mined and retrieved possible drug information via the amount of literature. Chen et al. [12] developed a statistical model to assess drug-target associations from a semantic network. Dong et al. [5] proposed a literature mining approach to extract hidden disease-drug associations from literature. Lee et al. [13] proposed a deep convolutional neural network to training PubMed abstracts and News articles for drug repositioning. However, the above methods only consider part of the drug-disease association, the drug-drug networks often were ignored. Meanwhile, the SARS-CoV-2 has only limited data and cannot use the above methods for SARS-CoV-2 drug repositioning.

3. Materials and methods

To prioritize candidate drugs for SARS-CoV-2, we propose a novel drug repositioning framework, named COVDR. The whole framework is showed in Fig. 2. The top part is dataset collection and preparation, which collected the SARS-CoV-2 relevant data from the DrugBank, PubMed, and the World Health Organization (WHO). We first give descriptions of the used dataset and construct a heterogeneous network by integrating multi-source data. The middle part is the model training, we propose a prediction method based on a graph convolutional network model, which is developed to learn low-dimension feature representation for nodes. The bottom part is the SARS-CoV-2 prediction, in this part, we predict potential drugs for SARS-CoV-2 based on feature representation and validate the predicted drugs with docking simulation and literature reference.

3.1. Dataset

Since there is no published drug repositioning dataset for SARS-CoV-

Table 1
Statistics of the dataset used in this study.

Dataset	Drugs	Diseases	Targets
	7,365	1,230	4,596
	Drug-Drug Interactions	Drug-Disease Interactions	Drug-Target Interactions
	290,836	6,680	25,012

2 when we started research, we build our dataset through collect multiple open sources, including DrugBank[1], repoDB[2], News reports[3] and some literature[4]. As a whole, we construct three different networks: (1) drug-disease interactions network. The interactions between drugs and diseases are constructed of two parts. One part of drug-disease interactions derives from the repoDB[2] dataset, which releases some "Approved" drug-disease dataset for drug repositioning by Drug Central and ClinicalTrials.gov. Another part of drug-disease derives from the manual annotation. But until now there has been no special drug to treat COVID-19 disease. Based on the News report[3] of the WHO, we manually labeled three potential drugs for COVID-19, including Umifenovir (DB13609), Remdesivir (DB14761), and Favipiravir (DB12466), all of them are still used in COVID-19 treatment.[5] (2) drug-target interactions network. The drug-target interactions comprise three parts, the first part drug-target interactions collected from DrugBank[1] website, which is a biomedical database with detailed drug data and includes associated drug evidence and direct or indirect gene targets. The second part of the drug-target dataset are collected from literature [9], which released 9,744 pairs of drug-target interactions. The third part of the dataset derives from manual labeling. Since there is currently no drug-target interaction data on SARS-CoV-2, we analyze SARS-CoV-2 related literature[4] and labeled 70 pairs of drug-target information about SARS-CoV-2 for drug repositioning. 3) drug-drug interactions. The interactions between drugs and drugs are collected from literature [9], which provided 290,836 clinically reported drug-drug interactions, and we select the clinically reported drug-drug interactions to experiment. The detailed statistic of the dataset is shown in Table 1.

3.2. Model training on the SARS-CoV-2 heterogeneous network

Based on the COVID-19 drug repositioning dataset, COVDR integrates the process of the graph convolutional network on the heterogeneous network to search candidate drugs for the SARS-CoV-2 disease, as shown in Fig. 2 model training part. Our model is primarily divided into three components, including the embedding layer, heterogeneous network modeling, and prediction component. The first component is the embedding layer, which is proposed to initialize representation embedding for nodes in the heterogeneous network. The second component is heterogeneous network modeling, which aims to learn the low-dimension feature representation of drugs, diseases, and targets from different perspectives. The third component is the prediction component, which aims to achieve the drugs and disease representation for prediction. Next, we will introduce each model component.

3.2.1. Embedding layer

First, we introduce the embedding layer. We embedded disease, drug, and target into representation embedding with initializer, that denoted as $p \in \mathbb{R}^{L \times D}$ for diseases, $q \in \mathbb{R}^{M \times D}$ for drugs, $r \in \mathbb{R}^{N \times D}$ for targets, where D means embedding dimensions, the L, M, N mean numbers of disease, drugs and targets. We define p_a means the embedding of

---

[1] https://www.drugbank.ca/
[2] http://apps.chiragjpgroup.org/repoDB/
[3] https://www.who.int/emergencies/diseases/novel-coronavirus-2019
[4] https://bigd.big.ac.cn/ncov/publication
[5] https://www.fda.gov/media/137574/download

disease a, $\boldsymbol{q_b}$ means the embedding of drug b, $\boldsymbol{r_c}$ means the embedding of target c. Then we update the initialized vector with a multi-hop graph convolutional network based on the heterogeneous network.

3.2.2. Heterogeneous network modeling

In addition to the disease-drug graph, the drug-drug graph and drug-target graph supply possibilities to learn disease and drug representations from different aspects. The network-based representation learning methods employ graph neural networks to aggregate graph features of neighboring nodes, which makes the representation embedding more powerful information. However, network-based drug repositioning methods [14] only design for a single graph or two graphs, real-world drug-disease-target interactions scenarios often contains multiple interactive graphs. Different from other works, we propose to enhance the graph neural network with a heterogeneous graph convolutional network for COVID-19 drug repositioning.

Disease-Drug Aggregation.
To enhance the representations of diseases in the disease-drug graph, we propose to utilize the graph convolutional networks to accumulate local neighbors information as follows,

$$\boldsymbol{p_a^g} = Agg_{drug2disease}\{\frac{1}{|R_a|}\sum_{b \in R_a} \boldsymbol{q_b}\} \tag{1}$$

where R_a means disease a associations drug b set, and $Agg\{*\}$ is a mean-based aggregating operation of GNNs [15].

Drug-Drug Aggregation. As suggested by DDI theories [16], drug prediction is more likely to be influenced by drug-drug interactions. It is important to integrate drug-drug interactions information into drug representations. Moreover, tie strengths among drugs can differently influence drugs' treatment ability. To recognize heterogeneous strengths of drug-drug interactions, we propose to aggregate differentiate drugs' local neighbors during aggregating operation in graph neural networks as follows,

$$\boldsymbol{q_b^g} = \sigma(\boldsymbol{W} \cdot Agg_{drug2drug}\{\sum_{d \in DDI_b} w_{d,b}(\boldsymbol{q_d} \odot \boldsymbol{q_b})\}) \tag{2}$$

where DDI_b means drug b interactions set of drug d in drug-drug network, the σ means relu activation function, and the $\odot$ denote the element-wise product, $\boldsymbol{W}$ is a trainable parametric matrix. Equivalently, $w_{d,b}$ is the important weight between drug d and drug b, here we used GCN laplacian matrix [15] as weight matrix.

On the other side, the directly connected nodes are beneficial for enhancing representation. Since the features can be passed throughout the DDI network, and drugs might be affected by the k-hop neighbor drugs. Accordingly, we propose to accumulate DDI information through k-layer aggregation as follows,

$$\boldsymbol{q_b^{g^{k+1}}} = \sigma(\boldsymbol{W} \cdot Agg_{drug2drug}\{\sum_{d \in DDI_b} w_{d,b}(\boldsymbol{q_d^{g^k}} \odot \boldsymbol{q_b^{g^k}})\}) \tag{3}$$

where $\boldsymbol{q_b^{g^k}}$ denotes the drug b representation after k-layer aggregating operation. The drug initial representation $\boldsymbol{q_b^{g^k}}$ is equal to $\boldsymbol{q_b^g}$ when $k = 0$.

Drug-Target Aggregation. A drug target is a molecule in the body, usually a protein, that is intrinsically associated with a particular disease process and that could be addressed by a drug to produce a desired therapeutic effect. It is very important to profile drugs from the related target. Hence, we introduce the drug-target aggregating operation to enhance the representation embedding of drugs. We introduce to aggregate drug target interaction information as follows,

$$\boldsymbol{q_b^t} = \sigma(\boldsymbol{W} \cdot Agg_{target2drug}\{\sum_{c \in DTI_b} w_{c,b}(\boldsymbol{q_b} \odot \boldsymbol{r_c})\}) \tag{4}$$

where DIT_b indicates neighbor targets c of drug b in drug-target graph, and $Agg_{target2drug}$ is the aggregating function. $\boldsymbol{r_c}$ means the initialized

vector of target c.

In addition, we propose to aggregate target information through k-layer aggregating operation as follows,

$$\boldsymbol{q_b^{t^{k+1}}} = \sigma(\boldsymbol{W} \cdot Agg_{target2drug}\{\sum_{c \in DTI_b} w_{c,b}(\boldsymbol{q_b^{t^k}} \odot \boldsymbol{r_c^{t^k}})\}) \tag{5}$$

where $\boldsymbol{q_b^{t^k}}$ denotes the drug b representation after k-layer aggregating operation. The drug initial representation $\boldsymbol{q_b^{t^k}}$ is equal to $\boldsymbol{q_b^t}$ when $k = 0$.

3.3. Prediction component

The last component in our proposed model is the prediction component, which aims to finalize the disease and drug representation for prediction. Since the drug-drug graph and drug-target graph provide important signals to understand drugs interactions information, we propose to integrate the final drug representation $\boldsymbol{q_b^f}$ as follows:

$$\boldsymbol{q_b^f} = \boldsymbol{q_b^{g^{k+1}}} \oplus \boldsymbol{q_b^{t^{k+1}}} \tag{6}$$

where $\oplus$ indicates summation operation.

With the disease and drug representation (e.g., $\boldsymbol{p_a^g}$ and $\boldsymbol{q_b^f}$), we calculate the prediction score via vector inner product as follows,

$$r_{a,b} = \boldsymbol{p_a^{gT}} \boldsymbol{q_b^f} \tag{7}$$

3.4. Model training

To training our model's parameters, we utilize the pair-wise loss [17] as an objective function for the SARS-CoV-2 drug repositioning task, the detail loss function is as follows,

$$\min_\theta Loss = \sum_{a=1}^{L} \sum_{(b,d) \in R_a} -\sigma(r_{a,b} - r_{a,d}) + \lambda \|\theta\|^2 \tag{8}$$

where $\sigma(\cdot)$ is a sigmoid activation function. L indicate the number of disease for training. θ means all trainable model parameters in our COVDR framework. λ is a regularization parameter that controls the complexity of disease and drug graph representation. R_a denotes disease a interactive drugs set. By optimizing the loss function, all parameters are updated via backward propagation.

4. Results

4.1. Baseline of COVDR

Since most related studies investigate drug-target prediction or drug-disease associations, few models were designed for the triple associations on SARS-CoV-2. Thus we compare our method to some pairwise association drug prediction methods, such as deepDR [9] and MLP [18] methods, which design neural networks to construct node representation for heterogeneous work and predict drug-disease interactions. In addition to the above two deep-based models, we further investigate the effects of using different additional information. We designed some variants to verify the COVDR performance, including COVDR + DDI, COVDR + DTI, and COVDR + No, the detailed introduction is as follows.

- **deepDR** [9]: deepDR model is one of the state-of-the-art drug repositioning methods, constructed nine different networks, including drug-drug interactions, drug-target interactions, drug-side-effect associations, drug-target interactions, drug-side effect interactions, chemical similarities and so on. Since there is currently no SARS-CoV-2 drug side effect association, we choose the drug-disease interactions to compare with our model.
- **MLP** [18]: A multilayer perceptron (MLP) is a class of neural network models, which consists of three or more layers of nodes. The
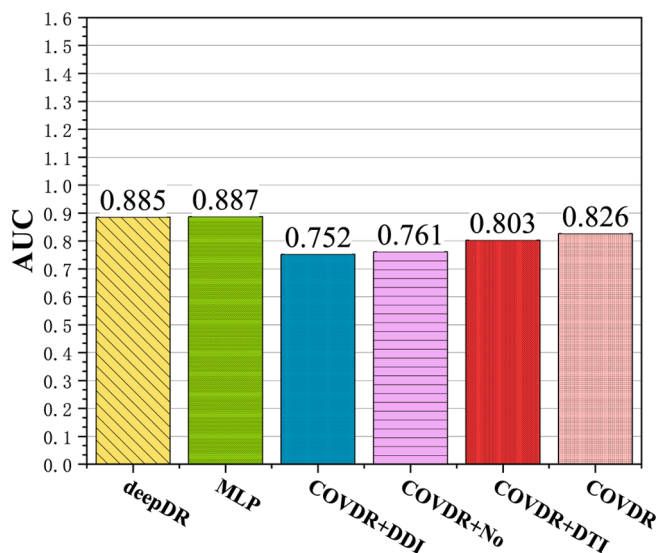
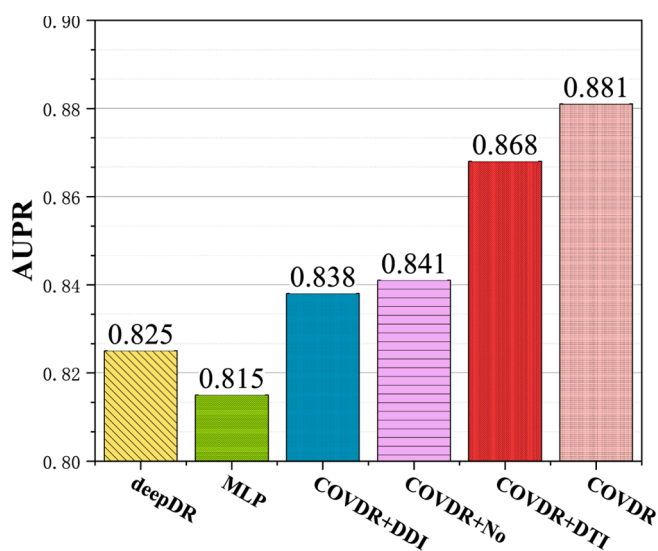**Fig. 3.** Performance of experiment dataset on AUC.

Fig. 4. Performance of experiment dataset on AUPR.

authors used MLP to distinguish data and predict new drugs for the disease. We choose the MLP model and set two hidden layers as the compare model.

- **COVDR + DDI:** To analysis the influence of drug-drug interactions, we design the COVDR + DDI, which is a variant model of COVDR, and only used drug-drug interactions and drug-disease interactions for SARS-CoV-2 drug repositioning.
- **COVDR + DTI:** To analysis the influence of drug-target interactions, we design the COVDR + DTI, which is a variant model of COVDR, and only used drug-target interactions and drug-disease interactions for SARS-CoV-2 drug repositioning.
- **COVDR + No:** To analysis our model ability of drug prediction, we design the COVDR + No, which only used drug-disease interactions without any other interactions for SARS-CoV-2 drug repositioning.
- **COVDR:** We combine three interaction networks, such as drug-drug interactions, drug-target interactions, and drug-disease interactions, and construct a heterogeneous drug-disease-protein network to predict potential drugs for SARS-CoV-2.
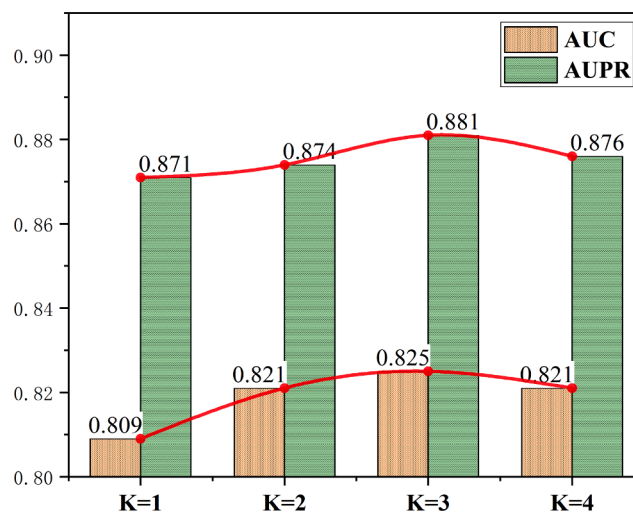

Fig. 5. Effect of multi-hop size k on experiment datasets.

4.2. Performance of COVDR on the dataset

To evaluate the performance of COVDR, we use 5-fold cross-validation to test the performance of triplet association prediction. Similar to other works [9,18], we choose AUC and AUPR to evaluate the drug repositioning performance. We apply the randomly selected 20% subset as the test, and the remaining 80% drug-disease pairs applied to train model parameters. The Fig. 3 and Fig. 4. reports the performance comparison results about AUC and AUPR. We have the following observations.

As shown in Fig. 3 and Fig. 4, we can find that: (1) The network-based model achieves better drug prediction performance than MLP, which demonstrates that MLP hard to learn heterogeneous network information without considering network structure, and the network local neighbor nodes can improve representation embedding. (2) The MLP achieves poor performance on AUPR than deepDR. This indicates that network-based representation learning is sufficient to capture the complex relationship between drugs and items, DeepDR consistently outperforms MLP across all cases, demonstrating the importance of graph interactions between drugs and diseases embeddings. (3) Comparing to deepDR and MLP, the performance of COVDR achieves the best performance on AUPR and accomplishes comparable improvement on AUC. Comparing with the COVDR + DDI, COVDR + DTI, and COVDR + No, the COVDR improves the performance on AUPR by approximately 4.3%. The main reason is that COVDR builds the representation of the drug-disease interaction via integrating drug-drug interactions and drug-target interactions. The design of the model can consider protein interaction among the heterogeneous network. Meanwhile, our COVDR achieves slightly inferior to deepDR in AUC indicator, we analysis the dataset and found that the reason is unbalanced data distribution. The ineffective drugs for diseases are far more than the number of effective drugs, and the AUC indicator is often invalid on the imbalanced data, but the AUPR indicator is more convincing.

4.3. Hyper-parameters sensitivity

In this subsection, we design the different k values for the model k-hop aggregating component by comparing AUC and AUPR to validate the effectiveness of our proposed COVDR model. The detail is shown in Fig. 5.

As shown in Fig. 5, when $k = 1$, the model only considers first-order neighbors, and our model gets the worst performance than other sizes k, which implies that nodes' higher-order neighbors' information is beneficial to enhance the drug repositioning. Meanwhile, we can find that the size $k = 3$ with multi-hop achieves a better performance than k

Table 2
Top 10 predicted drug for SARS-CoV-2.

DrugBank ID	Drug Name	Target UniProt ID	References
DB01212	Ceftriaxone	P0A3M6, Q9NSA0, Q4U2R8, Q8TCC7, P46059	[19,20]
DB00537	Ciprofloxacin	P43702, P43700, P11388, Q12809	[21,22]
DB00319	Piperacillin	Q75Y35, P0A3M6, Q7CRA4, Q8DNB6	[23,24]
DB00681	Amphotericin B	Q9UKR5	[25–27]
DB00254	Doxycycline	Q63X76	[28,29]
DB01112	Cefuroxime	Q8XJ01	[30–32]
DB00199	Erythromycin	O43193, Q12809	[33]
DB01017	Minocycline	P0A7X3, P0A7V8, P01584, P09917, P14780, P15692	[28]
DB00595	Oxytetracycline	P0A7X3, P0A7V8	[34,35]
DB00759	Tetracycline	P02359, P0AG59, P0A7W7, P0A7U3, P0A7V3	[28]

= 4 on AUC, and AUPR. The experiment result verifies that incorporated too many high-order neighbor nodes will lead to noise and reduce the accuracy of drug repositioning. The intuitive explanation is that only limited drugs can treat a disease.

4.4. SARS-CoV-2 potential drug prediction

We check the top 100 drug predictions and show the statistics of their projected top 10 pairwise associations in Table 2. The detailed literature and publications about are listed in references, and the target UniProt ID about each drug are showed Table 2.

As shown in Table 2, we find some evidence to confirm ten therapeutic associations, For example, One of the patients was treated with Ceftriaxone, and his cough, fever, and fatigue were improved gradually, he completed the 5–day course of hydroxychloroquine and azithromycin, 7 days of ceftriaxone [20,19]. As reported in [21,22], ciprofloxacin and glycyrrhizic acid were selected based on their reported anti-viral activity, safety, availability, affordability and subjected for Molecular Dynamics (MD) simulation. The MD simulation results indicate that ciprofloxacin may be repurposed against SARS-CoV-2. In patients in the ICU, piperacillin was the most commonly prescribed antibiotic, and the piperacillin was used for the treatment of SARS-CoV-2 in the study [22]. The studies [25–27] demonstrated that the amphotericin B and caspofungin were added, assuming a potential synergistic effect with liposomal amphotericin B for treat SARS-CoV-2. Several preclinical studies [28,29] showed that doxycycline (DOX) and minocycline (MIN), semi-synthetic tetracyclines frequently used in clinical practice against a variety of infective agents and well-tolerated,

are also effective against some RNA viruses. As such, DOX/MIN may be effective also against SARS-CoV-2 infection. Cefuroxime is a second-generation cephalosporin antibiotic, and the literature [30–32] shown that cefuroxime is a potent inhibitor of 3 key SARS-CoV-2 proteins, and cefuroxime axetil exhibited the possibility to halt the pocket of SARS-CoV-2 Mpro by forming stable interaction through covalent bonding and hydrogen bonding. The study [37] believes that Tetracycline is one of the potential therapeutic candidates, which demonstrates tetracycline antibiotic therapy corresponded to significant reductions in the duration of mechanical ventilation and ICU stay in ARDS patients. Oxytetracycline is one of the tetracyclines used for the treatment of infections caused by a variety of Gram-positive and Gram-negative microorganisms, and the studies [34,35] demonstrate that oxytetracycline is a relatively well-tolerated antibiotic that shows one of the highest docking scores. Erythromycin is used to treat certain infections caused by bacteria, and the study [33] demonstrate that Erythromycin could be an important combined treatment. Nonetheless, erythromycin has inhibited SARS-CoV-2 activity and may combine with hydroxychloroquine to reduce infection.

4.5. SARS-CoV-2 potential drug docking poses simulation

To validate the ability of our method predicting for SARS-CoV-2 drug-disease interactions, we check the top 100 predictions and find some potential drug-target interactions (Table 2. We select 10 potential drugs, which have supported references and 3D molecular structure for docking simulation. We perform computational docking for all of the top 10 predictions using Autodock Vina [38].

Our docking studies showed that the three drugs (i.e., Doxycycline, Cefuroxime, and Erythromycin) can dock to the structures of SARS-CoV-2-ACE2 (PDB ID:6M0J), and displayed different binding patterns (Fig. 6). In particular, all three drugs are fitted into the active sites of ACE2. More specifically, the doxycycline binding to ACE2 by forming hydrogen bonds with residues ASN546 with a bond length of 2.1Å in Fig. 6a, the cefuroxime binding to ACE2 by forming hydrogen bonds with residues LYS534, GLU536 with a bond length of 1.7Å, 2.0Å in Fig. 6b and the erythromycin binding to ACE2 by forming hydrogen bonds with residues LYS416 with a bond length of 2.2Å in Fig. 6c. These docking results may provide important hints for understanding the structural basis of the predicted SARS-CoV-2 potential drugs and thus help reveal the underlying molecular mechanisms of drug action.

5. Conclusion

To search for a specific drug for COVID-19, we propose a drug-

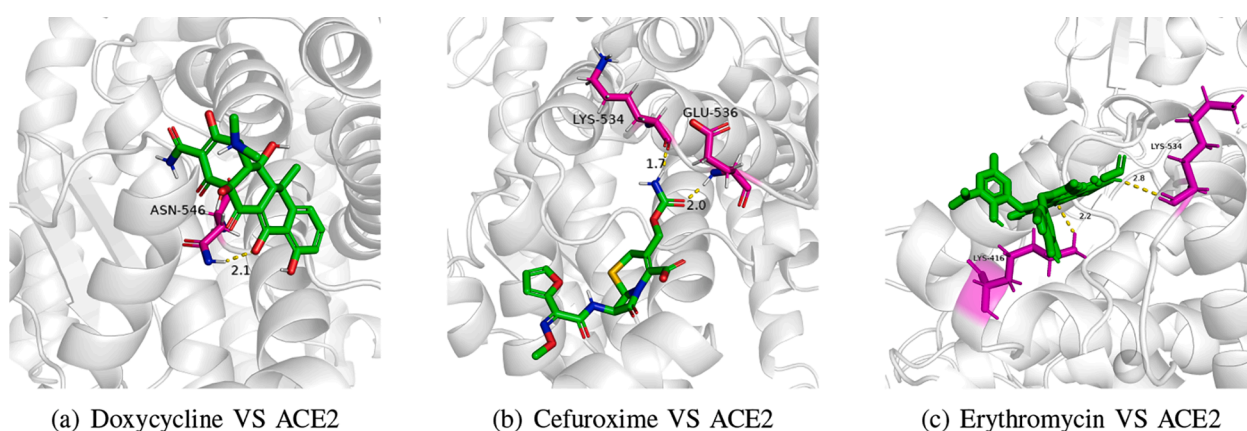(a) Doxycycline VS ACE2   (b) Cefuroxime VS ACE2   (c) Erythromycin VS ACE2

Fig. 6. The docked poses for the predicted interaction between three drugs (i.e., Doxycycline, Cefuroxime, and Erythromycin) and proteins (i.e, ACE2). The protein structures of ACE2 were downloaded from the Protein Data Bank (PDB ID:6M0J). The structures of the small molecules were obtained from the ZINC. The docking program Autodock was used for docking modeling. Hydrogen bonds were computed by PyMOL [36] and represented by the yellow dashe.d lines in ACE2.

disease-target triplet association prediction method (COVDR) based on the graph convolutional network. Since the SARS-CoV-2 is still in the research stage and there is no relevant drug repositioning dataset for SARS-CoV-2, we first collect a SARS-CoV-2 based drug repositioning dataset from the existing data (e.g., drugbank, repoDB). Secondly, apart from the gold standard drug-disease association network, we integrate one drug-drug interaction network and one drug-target association network to construct a complicated heterogeneous network. The heterogeneous network contains diverse information and a multi-perspective view for predicting SARS-CoV-2 potential drugs. Furthermore, our COVDR uses graph convolutional neural networks to learn heterogeneous networks to obtain low-dimensional feature representations of drugs, diseases, and targets. Then we predict SARS-CoV-2 potential drugs based on the feature representation. COVDR can learn the high-order nonlinear interactive information and improve the potential drugs recall rate via adopting a multi-hop mechanism. Comparing with baselines, COVDR achieves better performance than other methods. Moreover, we have validated most of the prediction drugs by literature reference. We have checked the top 10 predicted drugs and validated the three most relevant drugs by computational docking, the molecular docking results proved that Doxycycline, Cefuroxime, and Erythromycin have the potential to against the SARS-CoV-2.

CRediT authorship contribution statement

Haifeng Liu: Conceptualization, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Hongfei Lin:** Supervision. **Chen Shen:** Writing – review & editing. **Liang Yang:** Writing – review & editing. **Yuan Lin:** Writing – review & editing. **Bo Xu:** Writing – review & editing. **Zhihao Yang:** Writing – review & editing. **Jian Wang:** Writing – review & editing. **Yuanyuan Sun:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] Y. Zhou, Y. Hou, J. Shen, Y. Huang, W. Martin, F. Cheng, Network-based drug repurposing for novel coronavirus 2019-ncov/sars-cov-2, Cell Discovery 6 (1) (2020) 14.

[2] J. Li, S. Zheng, B. Chen, A.J. Butte, S.J. Swamidass, Z. Lu, A survey of current trends in computational drug repositioning, Briefings Bioinformatics 17 (1) (2016) 2–12.

[3] J. Yang, Z. Li, X. Fan, and Y. Cheng, "Drug–disease association and drug-repositioning predictions in complex diseases using causal inference–probabilistic matrix factorization," Journal of Chemical Information and Modeling, vol. 54, no. 9, pp. 2562–2569, 2014, pMID: 25116798.

[4] H. Luo, J. Wang, M. Li, J. Luo, P. Ni, K. Zhao, F.-X. Wu, Y. Pan, Computational drug repositioning with random walk on a heterogeneous network, IEEE/ACM Trans. Comput. Biology Bioinformatics 16 (6) (2018) 1890–1900.

[5] G. Dong, P. Zhang, J. Yang, D. Zhang, and J. Peng, "A systematic framework for drug repurposing based on literature mining," IEEE International Conference on Bioinformatics and Biomedicine, 2019.

[6] D.E. Gordon, G.M. Jang, M. Bouhaddou, J.X., et al., A sars-cov-2 protein interaction map reveals targets for drug repurposing, Nature 583 (7816) (2020) 459–468.

[7] P. Zhou, X. Yang, X. Wang, B. Hu, L. Zhang, W. Zhang, H. Si, Y. Zhu, B. Li, C. Huang, et al., A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (7798) (2020) 270–273.

[8] M.P. Menden, F. Iorio, M.J. Garnett, U. Mcdermott, C.H. Benes, P.J. Ballester, and J. Saezrodriguez, "Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties," PLOS ONE, vol. 8, no. 4, 2013.

[9] X. Zeng, S. Zhu, X. Liu, Y. Zhou, R. Nussinov, F. Cheng, deepdr: a network-based deep learning approach to in silico drug repositioning, Bioinformatics 35 (24) (2019) 5191–5198.

[10] M. Yang, L. Huang, Y. Xu, C. Lu, J. Wang, Heterogeneous graph inference with matrix completion for computational drug repositioning, Bioinform. 36 (22–23) (2021) 5456–5464.

[11] W. Wang, S. Yang, X. Zhang, J. Li, Drug repositioning by integrating target information through a heterogeneous network model, Bioinform. 30 (20) (2014) 2923–2930.

[12] B. Chen, Y. Ding, and D.J. Wild, "Assessing drug target association using semantic linked data," PLOS Computational Biology, vol. 8, no. 7, 2012.

[13] K. Lee, B. Kim, Y. Choi, S. Kim, W. Shin, S. Lee, S. Park, S. Kim, A.C. Tan, J. Kang, Deep learning of mutation-gene-drug relations from the literature, BMC Bioinformatics 19 (1) (2018) 21.

[14] J. Lim, S. Ryu, K. Park, Y.J. Choe, J. Ham, W.Y. Kim, Predicting drug–target interaction using a novel graph neural network with 3d structure-embedded graph representation, J. Chem. Inf. Model. 59 (9) (2019) 3981–3988.

[15] T. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," International conference on learning representations, 2017.

[16] R.S. Obach, R.L. Walsky, K. Venkatakrishnan, Mechanism-based inactivation of human cytochrome p450 enzymes and the prediction of drug-drug interactions, Drug Metab. Dispos. 35 (2) (2007) 246–255.

[17] S. Rendle, C. Freudenthaler, Z. Gantner, L. Schmidt-Thieme, Bpr: Bayesian personalized ranking from implicit feedback, in: UAI '09 Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, 2009, pp. 452–461.

[18] A. Aliper, S. Plis, A. Artemov, A. Ulloa, P. Mamoshina, A. Zhavoronkov, Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data, Mol. Pharm. 13 (7) (2016) 2524–2530.

[19] M. Khalid, T. Dernaika, L. Jacob, P. Annamaraju, and A.K. Guddati, "Weather the cytokine storm: A report of successful management of a colon cancer survivor and a critically ill patient with covid-19," Case Reports in Oncology, pp. 754–759, 2020.

[20] E.M. Saricaoglu, I. Hasanoglu, R. Guner, The first reactive arthritis case associated with covid-19, J. Med. Virol. (2020).

[21] W. Wang, J. Yu, J. He, Z. Wang, F. Li, Ambroxol inhibits mucoid conversion of pseudomonas aeruginosa and contributes to the bactericidal activity of ciprofloxacin against mucoid p. aeruginosa biofilms, Apmis 124 (7) (2016) 611–618.

[22] B. Br, H. Damle, S. Ganju, and L. Damle, "In silico screening of known small molecules to bind ace2 specific rbd on spike glycoprotein of sars-cov-2 for repurposing against covid-19," F1000Research, vol. 9, p. 663, 2020.

[23] B. Beović, M. Doušak, J. Ferreira-Coimbra, K. Nadrah, F. Rubulotta, M. Belliato, J. Berger-Estilita, F. Ayoade, J. Rello, H. Erdem, Antibiotic use in patients with covid-19: a 'snapshot' infectious diseases international research initiative (id-iri) survey, J. Antimicrob. Chemother. (2020).

[24] F. Khamis, I. Al-Zakwani, H. Al Naamani, S. Al Lawati, N. Pandak, M.B. Omar, M. Al Bahrani, Z.A. Bulushi, H. Al Khalili, I. Al Salmi, et al., Clinical characteristics and outcomes of the first 63 adult patients hospitalized with covid-19: An experience from oman, J. Infection Public Health (2020).

[25] C. Poignon, M. Blaize, C. Vezinet, A. Lampros, A. Monsel, A. Fekkar, Invasive pulmonary fusariosis in an immunocompetent critically ill patient with severe covid-19, Clinical Microbiol. Infection (2020).

[26] S. Arikan, M. Lozano-Chiu, V. Paetznick, J.H. Rex, In vitro synergy of caspofungin and amphotericin b against aspergillus and fusarium spp, Antimicrobial Agents Chemotherapy 46 (1) (2002) 245–247.

[27] J.-M. Vagace, C. Sanz-Rodriguez, M.-S. Casado, N. Alonso, M. Garcia-Dominguez, F.G. de la Llana, L. Zarallo, M. Fajardo, R. Bajo, Resolution of disseminated fusariosis in a child with acute leukemia treated with combined antifungal therapy: a case report, BMC Infect. Dis. 7 (1) (2007) 1–4.

[28] E. Francini, S.T. Miano, A.I. Fiaschi, G. Francini, Doxycycline or minocycline may be a viable treatment option against sars-cov-2, Med. Hypotheses 144 (2020), 110054.

[29] C. Sachdeva, A. Wadhwa, A. Kumari, F. Hussain, P. Jha, and N.K. Kaushik, "In silico potential of approved antimalarial drugs for repurposing against covid-19," OMICS: A Journal of Integrative Biology, 2020.

[30] A.B. Durojaiye, J.-R.D. Clarke, G.A. Stamatiades, C. Wang, Repurposing cefuroxime for treatment of covid-19: a scoping review of in silico studies, J. Biomol. Struct. Dyn. (2020) 1–8.

[31] K. Al-Khafaji, D. AL-DuhaidahawiL, and T. Taskin Tok, "Using integrated computational approaches to identify safe and rapid treatment for sars-cov-2," Journal of Biomolecular Structure and Dynamics, no. just-accepted, pp. 1–11, 2020.

[32] A.A. Elfiky, Sars-cov-2 rna dependent rna polymerase (rdrp) targeting: An in silico perspective, J. Biomol. Struct. Dyn. (2020) 1–9.

[33] Z. Yuan, M.A. Pavel, H. Wang, and S. Hansen, "Hydroxychloroquine: mechanism of action inhibiting sars-cov2 entry." bioRxiv, 2020.

[34] F.-J. Meyer-Almes, Repurposing approved drugs as potential inhibitors of 3cl-protease of sars-cov-2: Virtual screening and structure based drug design, Comput. Biol. Chem. (2020), 107351.

[35] S.K. Maurya, A.K. Maurya, N. Mishra, H.R. Siddique, Virtual screening, adme/t, and binding free energy analysis of anti-viral, anti-protease, and anti-infectious compounds against nsp10/nsp16 methyltransferase and main protease of sars cov-2, J. Recept. Signal Transduction (2020) 1–8.

[36] W.L. DeLano, "The pymol molecular graphics system," http://www. pymol. org, 2002.

[37] J.D. Byrne, R. Shakur, J. Collins, S.L. Becker, C.C. Young, H. Boyce, and C. Traverso, "Prophylaxis with tetracyclines in ards: Potential therapy for covid-19-induced ards?" medRxiv, 2020.

[38] O. Trott, A.J. Olson, Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2) (2010) 455–461.