Article

Patterns

Exploring complex and heterogeneous correlations on hypergraph for the prediction of drug-target interactions

Highlights

- A hypergraph framework to model high-order correlations in heterogenous biological network
- An embedding learning method for drugs and targets using hypergraphs
- High-order correlation between drugs and targets can contribute to DTI predictions

Authors

Ding Ruan, Shuyi Ji, Chenggang Yan, ..., Yue Gao, Changqing Zou, Qionghai Dai

Correspondence

gaoyue@tsinghua.edu.cn (Y.G.), aaronzou1125@gmail.com (C.Z.), qhdai@tsinghua.edu.cn (Q.D.)

In brief

Ruan et al. propose a new method for predicting drug-target interactions (DTIs). They pay attention to the high-order correlations in heterogeneous biological networks and use the hypergraph to model them. Their experimental results indicate that the high-order correlations among drugs and targets contribute significantly to DTIs predictions, and other associations besides DTIs are also useful in this task.

Ruan et al., 2021, Patterns 2, 100390
December 10, 2021
https://doi.org/10.1016/j.patter.2021.100390

CellPress

Patterns

CellPress
OPEN ACCESS

Article

Exploring complex and heterogeneous correlations on hypergraph for the prediction of drug-target interactions

Ding Ruan,[1,7] Shuyi Ji,[2,3,7] Chenggang Yan,[1,7] Junjie Zhu,[2] Xibin Zhao,[2] Yuedong Yang,[4] Yue Gao,[2,3,8,*] Changqing Zou,[5,*] and Qionghai Dai[3,6,*]

[1]School of Automation, Hangzhou Dianzi University, Hangzhou, China
[2]School of Software, KLISS, BNRist, Tsinghua University, Beijing, China
[3]Institute for Brain and Cognitive Sciences, Tsinghua University, Beijing, China
[4]School of Computer Science, Sun Yat-sen University, Guangzhou, China
[5]Huawei Vancouver Research Center, Huawei Canada Technologies, Vancouver, Canada
[6]Department of Automation, Tsinghua University, Beijing, China
[7]These authors contributed equally
[8]Lead contact
*Correspondence: gaoyue@tsinghua.edu.cn (Y.G.), aaronzou1125@gmail.com (C.Z.), qhdai@tsinghua.edu.cn (Q.D.)
https://doi.org/10.1016/j.patter.2021.100390

THE BIGGER PICTURE The prediction of drug-target interactions (DTIs) plays a crucial role in drug discovery. In this work, we discover that the high-order correlations in heterogeneous biological networks are essential for DTI predictions. The hypergraph structure is ultilized to model the high-order correlations in the biological networks, then the embeddings are generated for the drugs and targets, respectively. Finally, the interaction between them can be predicted according to the similarity of the embeddings. Our proposed method has been evaluated on multiple public datasets and the improved performance demonstrates that the high-order correlations among drugs and targets contribute significantly on DTI predictions, and other associations besides DTIs are also useful in this task.
Our method can also be used in other scenarios containing complex correlations.

1 2 3 4 5 **Proof-of-Concept:** Data science output has been formulated, implemented, and tested for one domain/problem

SUMMARY

The continuous emergence of drug-target interaction data provides an opportunity to construct a biological network for systematically discovering unknown interactions. However, this is challenging due to complex and heterogeneous correlations between drug and target. Here, we describe a heterogeneous hypergraph-based framework for drug-target interaction (HHDTI) predictions by modeling biological networks through a hypergraph, where each vertex represents a drug or a target and a hyperedge indicates existing similar interactions or associations between the connected vertices. The hypergraph is then trained to generate suitably structured embeddings for discovering unknown interactions. Comprehensive experiments performed on four public datasets demonstrate that HHDTI achieves significant and consistently improved predictions compared with state-of-the-art methods. Our analysis indicates that this superior performance is due to the ability to integrate heterogeneous high-order information from the hypergraph learning. These results suggest that HHDTI is a scalable and practical tool for uncovering novel drug-target interactions.

INTRODUCTION

The prediction of drug-target interactions (DTIs) plays a crucial role in drug discovery.[5] However, the biochemical experimental approaches widely used in wet laboratories are expensive and time consuming,[6] thus slowing down the progress of drug discovery. The ever-growing demand for inexpensive, effective, and rapid prediction methods has driven the development of

computational approaches, which provide a cheaper and faster way to predict potential interactions between drugs and targets. Conventional computational approaches tend to begin with the inherent properties of drugs and targets, such as the chemical structure of drugs and the three-dimensional (3D) structure of proteins. Molecular docking,[7] an important tool in structural molecular biology and computer-assisted drug design, is used to predict the binding mode(s) of a ligand with a protein of known 3D structure. Keiser et al.[8] use a complementary technique based on the chemical similarity of ligands to quantitatively group and relate proteins and discover unexpected ligand-target links. However, molecular docking predictions cannot be successful without a known and accurate 3D protein structure, and ligand-based methods require several known binding ligands.

Recently, machine learning methods[9] have attracted more attention and shown greater promise in drug discovery. Unlike the aforementioned methods, one key idea of current machine learning-based approaches is that similar drugs may share similar targets and vice versa.[1] Typical computational approaches adopt machine learning methods to catalog the similarities of drugs and targets based on biological features and then predict DTIs.[10–12] Yamanishi et al.[13] made the first attempt to predict DTIs based on biological feature information, such as the similarity between drug chemical structure and target protein sequence, unifying the chemical and genomic spaces of known drugs and targets into pharmacological spaces. Yu et al.[14] integrated features from chemical and genomic space for large-scale drug discovery using random forest and support vector machine algorithms. Gao et al.[15] used low-level representations such as Gene Ontology annotations, amino acids sequences, and chemical structural graphs as inputs to the neural network, generating embeddings for the targets and drugs, respectively, and then calculating the similarity between the embeddings to predict the interaction. This type of approach adequately extracts information from inherent properties, but problems arise when sufficient and reliable information is not available.

In addition to the inherent properties of drugs and targets, there is increasing interest in exploring the correlations among drugs, targets, and other biological entities in the data structure of a heterogeneous biological network. Compared with biological feature-based methods, network topology information-based methods make predictions based on the topology information of the network.[16,17] Several recent attempts have explored topological structures of model DTIs, with biological entities such as drugs, targets, side effects, and diseases denoting vertices in the biological graph and the interactions or associations indicating edges among them. Campillos et al.[18] constructed a network of 1,018 side effect-driven drug-drug relations and validated 13 implied drug-target relations. Cheng et al.[19] compared network-based inference with drug-based similarity inference and target-based similarity inference , showing that the former achieved higher-quality results. Chen et al.[20] integrated and annotated data from public datasets to build a semantic-linked network. They developed a statistical model to assess the association of drug-target pairs and observed that drugs from the same disease area will cluster together. They noted that this mode of clustering is difficult to infer based on inherent properties alone. We hypothesized that correlation among various biological entities can provide useful information that cannot be obtained from inherent properties. Some recent methods formulate DTI prediction tasks as ''link predictions'' in complex networks.[17,21,3] TriModel[3] represents heterogeneous topological correlations in the form of a knowledge graph and generates embeddings to predict whether there is a link between a drug and a target (supplementary note). Furthermore, similarities based on both inherent properties and topological correlations can be used to predict DTIs. DTINet[1] integrates diverse inherent properties and topological correlations through a network diffusion process. It generates representations for drugs and targets, containing the similarities of vertices in the biological network, and then performs predictions using these representations (supplementary note). DeepDTnet[2] is another network-based method that integrates information based on the inherent properties of drugs and targets (supplementary note). NeoDTI[22] also integrates information from heterogeneous network data and predicts DTIs by learning the topological preservation representations of drugs and targets.

In summary, previous methods have performed DTI predictions by extracting the similarities between drugs and targets. However, they describe the interactions between drugs and targets in a low-order manner where only pairwise correlations are taken into consideration, i.e., one-drug, one-target paradigms. However, the connections among biomedical entities can be far more intricate than merely pairwise links. For example, a single drug may be connected to a number of targets (so-called multi-target drugs, which can target various complex diseases as they are ubiquitous and effective[23]), and these targets may share subtle but important pharmacological characteristics that contribute to the interactions. When further considering more connections, such as drug-disease associations and target-disease associations, the overall heterogeneous biological network becomes even more complex and emerges in a many-to-many pattern. Under such circumstances, it is important to formulate and explore the underlying higher-order topological correlations for drug discovery, which is beyond the capability of the pairwise correlation-based methods. To tackle this issue, we adopted a heterogeneous hypergraph-based model to explore complex and heterogeneous correlations for drug-target interaction prediction (HHDTI) (see section ''experimental procedures'' for more details).

Unlike traditional graphs that model pairwise correlations, the hypergraph can model higher-order correlations and is thus more flexible and powerful, with the ability to incorporate complex correlations. There are precedents for modeling biological networks using hypergraphs, but they have not been used to predict DTIs. Vaida et al.[24] modeled relations between pairs of drugs as a hypergraph and used a two-layer graph convolution neural network as an encoder to predict drug interactions. Niu et al.[25] used diseases as hyperedges, connected microbes associated with them, and developed a hypergraph-based random walk model for microbe-disease association prediction.

Hypergraphs are indeed suitable for modeling drug-target interaction networks. When a drug-target hypergraph is constructed, targets are denoted by vertices, and the interactions between a specific drug and a certain number of targets can be modeled by a hyperedge. In this hypergraph, all targets interacting with the same drug are connected by a hyperedge; therefore, all the target vertices connected by one hyperedge can be regarded as a set. Rather than a graph edge in a

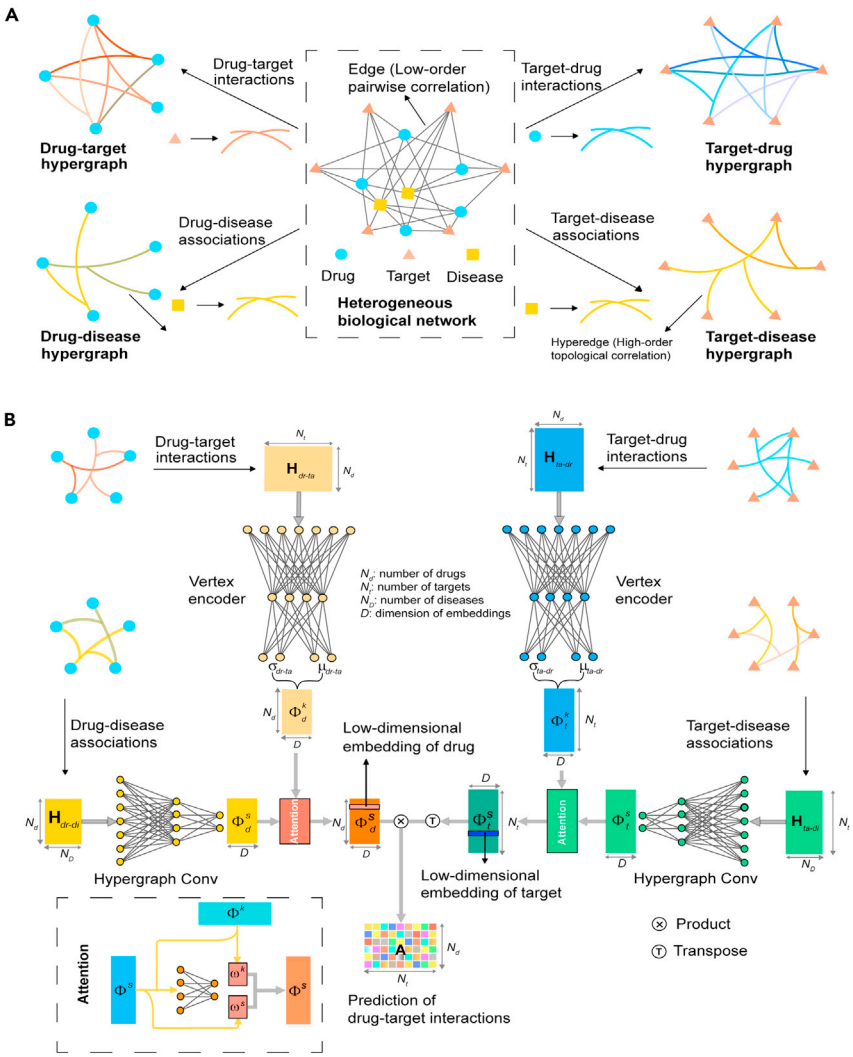

Figure 1. Schematic flowchart of the HHDTI pipeline

(A) Illustration of the hypergraph construction.

(B) Given the heterogeneous biological network in (A), four distinct types of sub-hypergraphs (drug-target, drug-disease, target-drug, and target-disease) can be built. Taking the target-drug interactions as an example, we used a hyperedge to connect all targets that interact with the same drug, i.e., a hyperedge in the heterogeneous biological hypergraph represents a drug. These hypergraphs provide the input for the key and side embedding learning in (B). The incidence matrix **H** represents the sub-hypergraph and serves as the input of the model, and Φ^k, Φ^s, and Φ^S represent the key embeddings, side embeddings, and structural embeddings, respectively. μ and σ, respectively, refer to the means and variances obtained by the variational autoencoder when generating the key embeddings. "Attention" means bi-embedding attention fusion module.

heterogeneous biological graph representing a two-order pairwise correlation (i.e., indicating direct DTIs), a hyperedge in a heterogeneous biological hypergraph instead models high-order multilateral (i.e., many-to-many) correlations between targets and drugs. Moreover, to provide a thorough understanding of DTIs, we comprehensively integrated several types of connections among various vertices (e.g., drug-target, target-disease, and drug-disease connections) in the heterogeneous biological networks. A representation modeled on higher-order correlations can significantly improve the predication performance of DTIs.

Specifically, HHDTI infers candidate DTIs by fusing two types of embeddings: key and side embeddings. Key embeddings provide initial and major vector representations for all drugs and targets, which are learned using the direct drug-target interaction information. By contrast, side embeddings offer complementary representations learned by leveraging disease-relevant information. Structural drug-target embeddings are achieved by fusing the key embeddings with the side embeddings, with HHDTI estimating drug-target similarity to perform DTI predictions. We have demonstrated that, based on this embedding learning process, HHDTI consistently achieved higher-quality prediction results when analyzing several popular datasets compared with alternative state-of-the-art methods. Comprehensive evaluations have determined that the proposed HHDTI is a promising and powerful tool for drug discovery.

RESULTS

Overview of HHDTI

We propose a computational framework for DTI prediction, called HHDTI, which captures implicit high-order topological correlations in heterogeneous biological networks. HHDTI first uses a generative model to construct key embeddings from drug-target and target-drug interactions (Figure 1). It then extracts drug-disease correlations and target-disease correlations to generate side embeddings using hypergraph neural networks (HGNNs).[26] Ultimately, HHDTI fuses the key embeddings and side embeddings and obtains structural embeddings to perform DTI prediction. Integrating diverse information from heterogeneous biological data can assist in determining higher-order topological correlations among different vertices. HHDTI then can infer potential DTIs by computing and ranking the prediction scores of all candidate interactions. In summary, embeddings encode both topological properties and association information, resulting in a low-dimensional vector space where the distance between drug-target pairs correlates with their likelihood of interaction. More details of the HHDTI framework can be found in the section "experimental procedures."

Better DTI prediction performance by HHDTI

We initially evaluated the overall prediction performance of HHDTI using a 10-fold cross-validation procedure. We conducted these experiments on three public datasets (DTINet_17,[1] deepDTnet_20,[2] and KEGG_MED[3]) and compared HHDTI with four

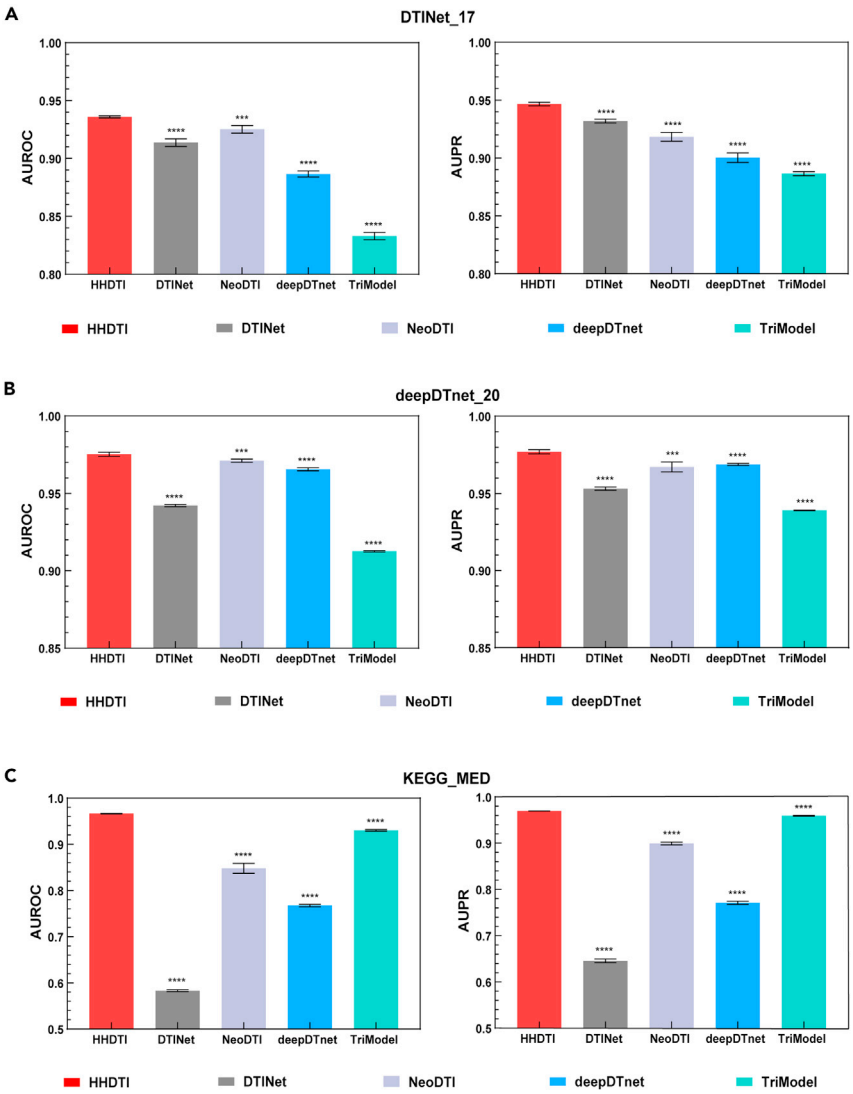

Figure 2. HHDTI outperforms other models when used on all three datasets

(A–C) Experimental results as measured by AUROC and AUPR. 10-fold cross-validations were performed on (A) DTINet_17, (B) deepDTnet_20, and (C) KEGG_MED databases to compare the prediction ability of HHDTI with DTINet, NeoDTI, deepDTnet, and TriModel (supplementary note). The results of five trials for each method are expressed as mean ± SD; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

state-of-the-art network-based drug discovery methods: DTINet, NeoDTI, deepDTnet, and TriModel. Under the experimental setting, 10% of the known drug-target interaction pairs and non-interaction pairs were randomly chosen as the positive and negative samples, respectively, for testing. The remaining 90% were used for training. Two widely used metrics, the area under the receiver operating characteristic (AUROC)[27] curve and the area under the precision-recall (AUPR) curve, were calculated to comprehensively compare the performance of different methods. We conducted separate experiments on these three datasets and found that there was no data overlap between the training and test sets within each dataset. The four methods were consistent with the results provided in the original papers for their corresponding datasets (Figure 2). However, HHDTI outperformed each of these competitive baselines, consistently achieving the highest prediction results for all three datasets. All four methods are network-based methods, each with minor differences. DTINet, deepDTnet, and NeoDTI blend the inherent properties of drugs and targets and the topological correlations among biological entities. For this reason, both methods perform poorly on the KEGG_MED dataset, which does not include any information related to inherent properties such as the chemical structures of drugs and the primary sequences of proteins. Although these baseline methods attempt to fuse diverse information in heterogeneous biological networks, they are still limited in terms of data modeling as they can only capture low-order pairwise correlations between vertices rather than high-order correlations.

The superior performance of the prediction methods might result from the easy predictions of homologous proteins or similar drugs in the dataset. To investigate this issue, we refer to the work of Luo et al.[6] and performed an additional test on the DTINet_17 dataset without the DTIs involving homologous proteins (sequence identity scores >40%). In this test, the removal of homologous proteins can reduce the potential redundancy in the DTIs that may lead to an inflated performance evaluation. The test results were robust even after removing homologous proteins from the training data, suggesting that HHDTI capturing high-order correlation information can still achieve good performance and outperform other prediction methods even in the absence of similar targets (Figure S1.).

Additional association information for DTI prediction

We further investigated how the quantity of potential isolated data influences DTI prediction results. We extracted all known drug-target interaction pairs of three different amounts of drugs (20%, 50%, and 80%) within the datasets as positive samples and the same number of non-interaction pairs as negative samples to generate the test sets (i.e., there are no known drug-target interaction pairs in the training data for these drugs). This experimental setting simulated the so-called cold-start problem by artificially creating isolated vertices, resulting in extremely difficult DTI predictions. Our analysis showed that the side embeddings generated from the association information (i.e., drug-disease and target-disease associations) can help improve DTI predictions to some extent, despite the absence of any known drug-target interaction pairs within the training sets (Figure 3). These studies also showed that additional association information can be captured by the

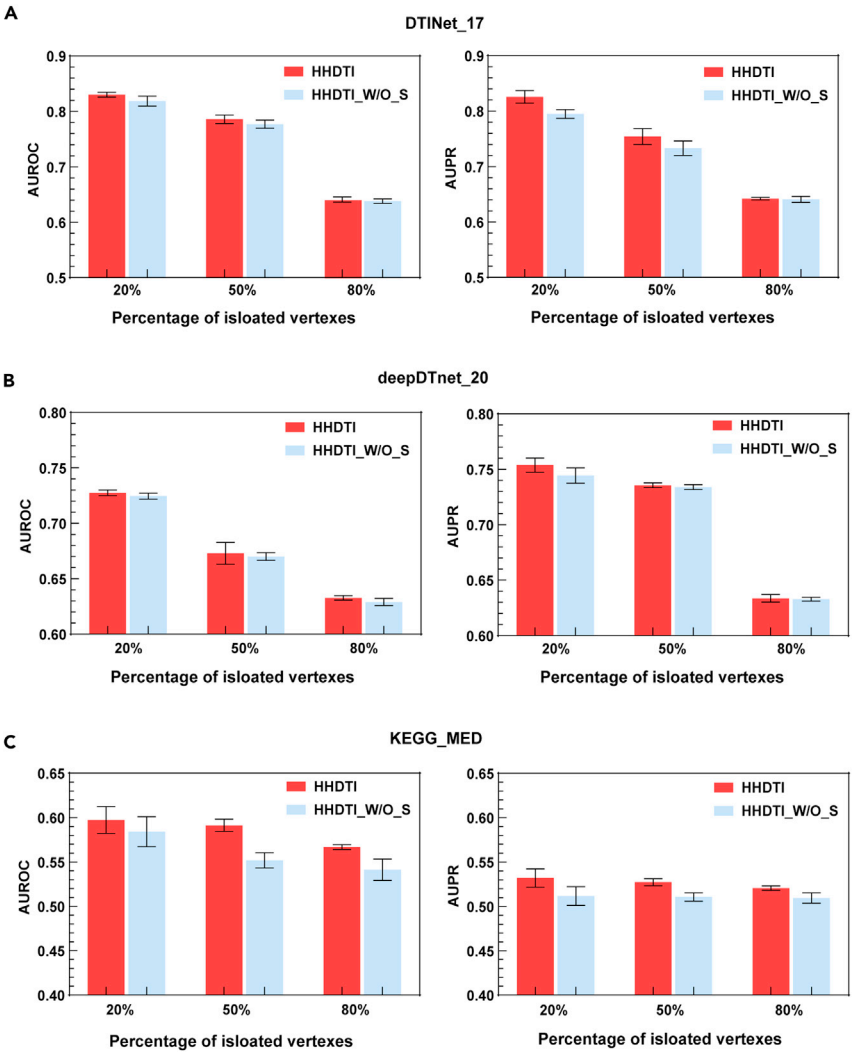

Figure 3. HHDTI evaluated under cold-start conditions

(A–C) All known interactions of three different amounts of drugs (20%, 50%, and 80%) in the datasets (A) DTINet_17, (B) deepDTnet_20, and (C) KEGG_MED and the same number of negative samples form the test sets. Specifically, in the first experiment, 20% of the drug vertexes in the training set are isolated vertexes; in the second experiment, 50% of the drug vertexes in the training set are isolated vertexes; and in the third experiment, 80% of the drug vertexes in the training set are isolated vertexes. HHDTI_W/O_S means HHDTI does not use side embeddings for DTI prediction. The results summarize five trials and are expressed as mean ± SD.

proposed HHDTI to enhance DTI predictions, which may provide new insights into understanding interaction mechanisms among drugs, targets, and diseases.

High-order topological correlations for DTI prediction

We conducted ablation experiments on the DTINet_17, deepDTnet_20, and KEGG_MED datasets, respectively, to study the advantages and disadvantages of high-order topological correlations relative to low-order pairwise correlations. To this end, we replaced the hypergraph representation in HHDTI with plain graph representations and used this as the comparative method (specifically, we constructed standard plain graphs on these three datasets and performed a similar key-side embedding learning procedure as HHDTI for DTI prediction). The experimental results showed that HHDTI consistently outperformed the low-order correlations-based comparative method when used on either of the three datasets (Figure 4).

Practical drug discovery

Our goal was to study HHDTI's capability as a practical tool for unknown DTI discovery. We chose Target Drug-UniProt Links (approved) of the DrugBank database[4] in version 5.1.0 for the evaluation, as it contains detailed and complete interaction information for targets and drugs. DeepDTnet[2] was chosen as the comparative method because it achieved the highest quantitative prediction among the baselines. Since there is no disease association information in this dataset, we compared HHDTI (no disease) with DeepDTnet. We trained the two methods using all the data in Target Drug-UniProt Links (approved) and produced a top-10 target prediction list for each drug using each of the two methods (Table S1). Data S1 and S2 are the lists of DTIs predicted by HHDTI (no disease) and deepDTnet, respectively, and validated by the literature. In the lists predicted by both methods, aside from the known targets in the training set, we observed that there was a subset of new predicted DTIs that were unknown in the training set but had been reported in the literature. Statistical analysis showed that HHDTI successfully predicted 17.9% more DTIs than deepDTnet. To further compare HHDTI (no disease) and deepDTnet, we used "recall @ top-10" as the evaluation metric,[28,29] which is defined as the fraction of true interacting targets retrieved in the list of top-10 predictions for a drug. With this evaluation metric, the average recall at top-10 of HHDTI (no disease) and deepDTNet were 0.0590 and 0.0573, respectively. This indicates that both methods can successfully discover targets that interact with a given drug and that HHDTI (no disease) is more powerful than deepDTNet.

Figure 5 illustrates specific practical drug discovery results produced by HHDTI (no disease) and deepDTNet. The data in the training set show that the anti-epileptic drug phenytoin acts on nuclear receptor subfamily 1, group I, member 2 (NR1I2) and several targets from the sodium channel family (SCN1A, SCN3A, and SCN5A). The drug brivaracetam, which is commonly used in the treatment of partial-onset seizures, is a ligand of synaptic vesicle protein 2A (SV2A) and inhibits voltage-gated sodium channels. Existing low-order correlation-

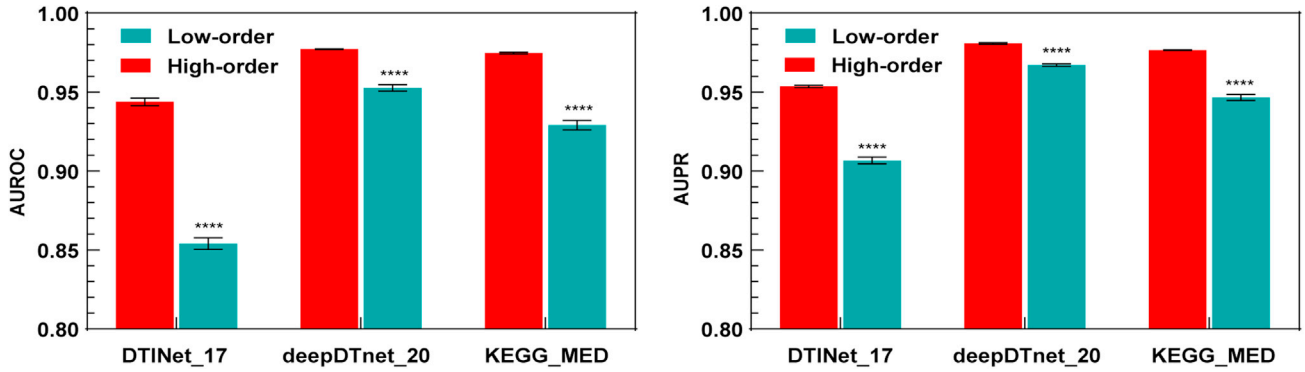

Figure 4. Ablation experiments determine the contribution of high-order topological correlation to HHDTI

We performed ablation experiments using the DTINet_17, deepDTnet_20, and KEGG_MED datasets to evaluate the superiority of high-order correlations. The results summarize five trials and are expressed as mean ± SD; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

based methods, including deepDTNet, make DTI inferences based on the "guilt-by-association" assumption that similar drugs may share similar targets and vice versa. Since both brivaracetam and phenytoin act on similar targets, deepDTNet predicted that phenytoin acts on a member of sodium channel family SCN8A. However, deepDTNet failed to predict the interaction between phenytoin and KCNH2, which is not similar to NR1I2 or the sodium channel family. The experimental results reveal that the problem with these methods is that they are only able to predict targets that are similar to known targets. In contrast, HHDTI (no disease) successfully predicted that phenytoin acts on KCNH2. As shown in Figure 5, the training data reveal the similarity between NR1I2 and KCNH2 because both NR1I2 and KCNH2 have interactions with the same drug, ketoconazole. The two targets NR1I2 and KCNH2 are thus linked by a hyperedge and are regarded as a set. We first train the model to find a certain similarity between the targets in the set and project it into a low-dimensional common feature space as the embedding of the drug. In the same way, we can obtain the embedding of the target. The drug embedding and the target embedding with known interactions are then positioned close to each other (i.e., the embedding of ketoconazole and the embedding of KCNH2 are close in the low-dimensional feature space). Since phenytoin and ketoconazole also act on SCN5A, their embeddings will also be near each other in the feature space. Due to the transfer of similarity, HHDTI successfully predicted the interaction of phenytoin with KCNH2. The interaction of propafenone with SCN5A and KCNH2 can also help predict the interaction between phenytoin and KCNH2. Furthermore, SCN5A and KCNH2 belong to the voltage-gated ion channel superfamily, suggesting that our method finds some similarity between these two proteins and facilitates us to further explore the role and structure of the proteins. The high-order topological correlation allows HHDTI to take full advantage of known interaction information in the heterogeneous biological network and recall more potential DTIs in a top-N prediction list.

We conducted additional rigorous testing. We downloaded the earliest available release (v4.6.0, released on 20 April 2016) from the DrugBank database.[30] Using all the data in Target Drug-UniProt Links (approved) from this release, we obtained some results that prove the validity of HHDTI. As shown in Table S2, these results have been validated in the literature and the publication time of these literatures is later than April 2016. For example, the interactions related to the drug celiprolol (DB04846) in the training set were first documented in the literature in 2007.[31] HHDTI predicts that the drug also interacts with beta-3 adrenergic receptor (ADRB3, P13945) and alpha-2A adrenergic receptor (ADRA2A, P08913), and these results were proved by the literature in 2017.[32]

DISCUSSION

The HHDTI method presented here is a computational approach based on hypergraph networks and deep neural networks. Based on known DTIs, HHDTI extracts the intrinsic characteristics of drugs and targets, models these correlations with a hypergraph capable of higher-order modeling, and then enhances these correlations with complementary information to generate structural embeddings for both drugs and targets. The major advantage of the proposed method lies in its powerful capability of modeling high-order correlations among various entities and its flexible framework capable of integrating several types of complementary information. Our study found it can discover more DTIs that have been previously validated by the literature than other state-of-the-art computational approaches. It can therefore identify potential DTI candidates to efficiently guide validation experiments in the wet laboratory. In the future, we plan to perform wet experimental validation as a method of cross-validation through cooperation with drug discovery industry partners, which will help us further improve the framework in return.

Although network-based methods have been applied,[1,2] the correlation modeling based on one-to-one correspondence may not produce the essential features reflecting a single drug acting on multiple targets or multiple drugs acting on the same target. Integrating network biology and polypharmacology promises an expanded opportunity for druggable targets,[33] which cannot be achieved without effective high-order correlation modeling. Capturing the high-order topological correlations among various vertices in a heterogeneous biological network can achieve more accurate and robust prediction performance, which is worthy of more attention for further study. Although

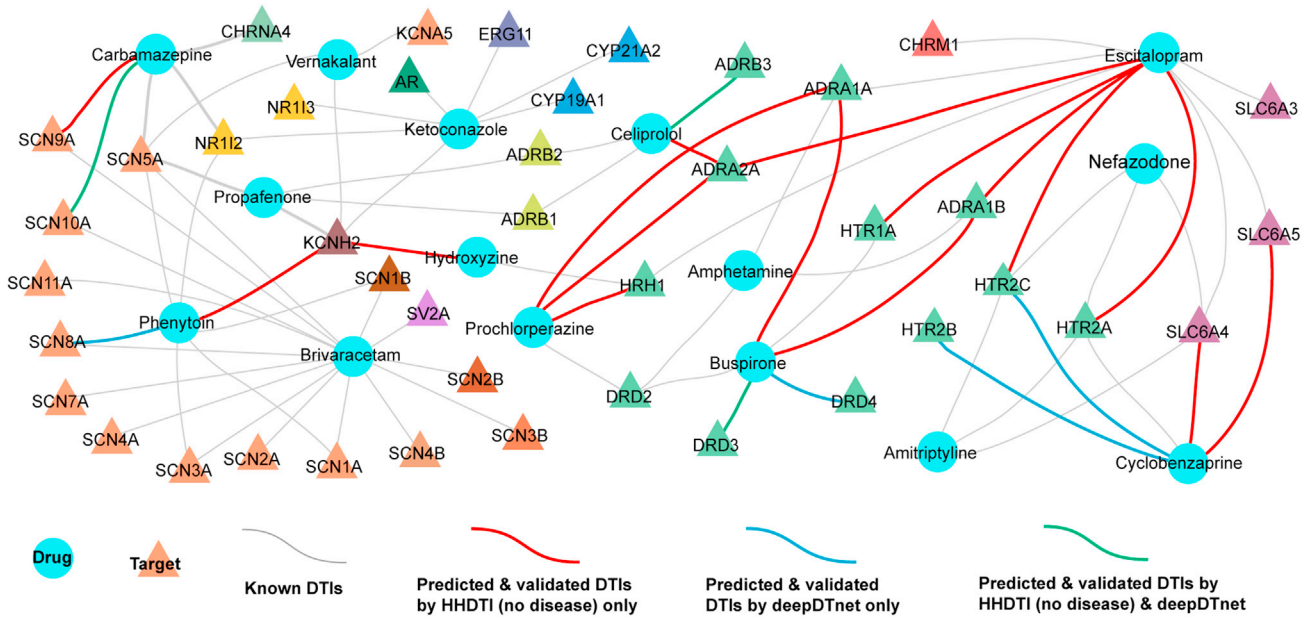

Figure 5. Predicted and validated DTI examples visualized in a heterogeneous biological network

Predicted and validated DTIs refer to the predicted DTIs that can also be confirmed by known experimental or clinical evidence in the literature. Targets of the same color belong to the same protein family. HHDTI can discover more interaction targets that are not close to the known interaction targets in terms of protein family proximity for drugs than the state-of-art network-based method deepDTnet.

computational approaches have achieved decent results after years of development, there are still many under-resolved problems. The biological data used in this study are considered large-scale datasets, but the number of drug vertices, target vertices, and DTIs included in each dataset is quite limited.[1,3,34,35] For example, the approved Target Drug-UniProt Links in DrugBank database (version 5.1.0)[4] only contains 2,020 drugs, 2,669 targets, and 9,796 DTIs. To construct a large-scale comprehensive heterogeneous biological network, more types of vertices in addition to drugs and targets should be provided to obtain complex relationships at different levels.[36] It is not easy to accomplish this task using a single dataset. Fortunately, we may integrate complementary information from different public databases. For instance, we can integrate the known drug-disease associations from Drug Central,[37] clinically reported drug side effects from the Comparative Toxicogenomics Database (CTD),[38] protein-protein interactions data from the Human Protein Resource Database (HPRD)[39] and the HuRI,[40] and clinically reported drug-drug interactions data from the DrugBank database. Even with plenty of data, coping with the noise from multiple databases is a challenging problem for data integration. The sample imbalance problem may also be raised by collecting only positive sample information and ignoring information for non-interaction pairs. Furthermore, even an evaluated DTI may be rejected in the future.[4] We believe that a high-quality, large-scale dataset that integrates various classes of information will significantly progress the development of computational approaches.

By convention, the HHDTI selects drug-target pairs with no known interactions as negative samples. These negative samples are potentially positive, making it difficult to select genuine no-interaction drug-target pairs.

The proposed HHDTI method can be further expanded to incorporate more topological information (e.g., drug side effect associations) and other types of information. For example, the similarity computed from the inherent property information of drugs and targets, such as drug chemical similarity and protein sequence similarity, can also be modeled in the form of hypergraphs to explore the high-order correlations in this respect, which will be considered in our future research. Importantly, although HHDTI was developed for DTI predictions, it can also be used as a general framework to address link prediction-related problems in other fields (e.g., drug interactions).

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests for code and data should be directed to and will be fulfilled by the lead contact, Yue Gao (gaoyue@tsinghua.edu.cn).

Materials availability

This study did not generate any physical materials.

Data and code availability

The four datasets used in the experiments can be found in DTINet,[1] deepDTnet,[2] TriModel,[3] and DrugBank database https://doi.org/10.1093/nar/gkx1037.[4] HHDTI source code can be downloaded from https://github.com/iMoonLab/HHDTI.

The framework of the HHDTI

The framework of the proposed HHDTI is shown in Figure 1. Taking the biological hypergraphs as input, HHDTI can achieve prediction performance that outperforms other state-of-the-art methods by simultaneously optimizing both the high-order association capture process and the DTI prediction model in an end-to-end manner. We first construct hypergraphs to model the biological network and then employ a structural embedding learning framework to capture the

high-order correlation and generate structural embeddings for both targets and drugs. The interaction likelihood between a given drug and target is predicted by estimating the similarity of their structural embeddings. Specifically, for drug i and target j, the DTI score can be computed as $Sigmoid((\mathbf{\Phi}_d^{\mathbf{S}})_i((\mathbf{\Phi}_t^{\mathbf{S}})_j)^{\mathbf{T}})$, where $\mathbf{\Phi}_d^{\mathbf{S}}$ and $\mathbf{\Phi}_t^{\mathbf{S}}$ denote the drug structural embeddings and target structural embeddings, respectively. These low-dimensional structural embeddings, $\mathbf{\Phi}_d^{\mathbf{S}}$ or $\mathbf{\Phi}_t^{\mathbf{S}}$, are generated by fusing key and side embeddings by a biembedding attention fusion module; drug (target) structural embeddings $\mathbf{\Phi}_d^{\mathbf{S}}(\mathbf{\Phi}_t^{\mathbf{S}})$ are generated by fusing the key drug embeddings $\mathbf{\Phi}_d^{k}(\mathbf{\Phi}_t^{k})$ and side drug embeddings $\mathbf{\Phi}_d^{s}(\mathbf{\Phi}_t^{s})$.

Heterogeneous hypergraph modeling of biological networks

Biological networks in this work present both direct and indirect relationships between drugs and targets. A heterogeneous biological network $\boldsymbol{G_h} = \{\boldsymbol{V_h}, \boldsymbol{E_h}\}$ refers to a biological network containing multiple types of vertices and edges, where $\boldsymbol{V_h}$ represents the set of vertices and $\boldsymbol{E_h}$ represents the set of edges. In our biological network, the sets of vertex types $\boldsymbol{O}$ include {drug, target, disease}, the sets of correlation types $\boldsymbol{R}$ include {drug-target interaction, target-drug interaction, drug-disease association, target-disease association}. Given different types of correlations, a heterogeneous multiple hypergraph $\boldsymbol{G} = \{\boldsymbol{V}_r = \{v_1, ..., v_{M_r}\}, \boldsymbol{E}_r = \{e_1, ..., e_{N_r}\}\}$ with M_r vertices and N_r hyperedges is constructed to model the biological networks, where r represents different types of correlations and $r = 1, 2, 3, 4$. In this work, the heterogeneous hypergraph modeling of the biological networks is illustrated in Figure 1A. For each correlation, we achieve an individual sub-hypergraph. We achieve four types of sub-hypergraph in total. The heterogeneous hypergraph modeling results are four incidence matrices, which can be represented by $\mathbf{H} \in \mathbb{R}^{M \times N}$, where $\mathbf{H}_{i,j} = 1$ if vertex i has connected with hyperedge j; otherwise, $\mathbf{H}_{i,j} = 0$. We obtain four types of incidence matrices ($\mathbf{H}_{dr\text{-}ta}$, $\mathbf{H}_{ta\text{-}dr}$, $\mathbf{H}_{dr\text{-}di}$, $\mathbf{H}_{ta\text{-}di}$) based on $\boldsymbol{R}$. Both drugs and targets employ the same structural embedding learning framework to generate the structural embeddings. For conciseness, we next present how drug structural embeddings are generated from this structural embedding learning framework.

Drug structural embedding learning

We introduce a Bayesian deep generative model that is a framework for unsupervised learning on a hypergraph-structured data-based variational auto-encoder[41] to learn drug key embeddings from $\mathbf{H}_{dr\text{-}ta}$ and employ the HGNN[26] model to generate the drug side embeddings from $\mathbf{H}_{dr\text{-}di}$. For the drug-target interaction hypergraph $\mathbf{H}_{dr\text{-}ta}$, this Bayesian generative model is instantiated as a vertex encoder, which models the similarity and correlations of the drugs interacting with the same target. The vertex encoder (Figure 1B, vertex encoder) performs a nonlinear mapping from the observed space $\mathbf{H}_{dr\text{-}ta}$ to the common latent space $\mathbf{\Phi}'_{dr-ta}$ by

$$\mathbf{\Phi}'_{dr-ta} = f(\mathbf{H}_{dr-ta}\mathbf{W}_{dr-ta} + \mathbf{b}_{dr-ta}) \quad \text{(Equation 1)}$$

where $f(\cdot)$ is a nonlinear activation function to enable our model to approximate a nonlinear function.[42] Based on our experiments (Figure S2), we adopted the hyperbolic tangent $\tanh(x) = (\exp(x) - \exp(-x))/(\exp(x) + \exp(-x))$ for the activation function due to its simplicity and superiority of performance. $\mathbf{W}_{dr-ta} \in \mathbb{R}^{D_{in} \times D_{out}}$ and $\mathbf{b}_{dr-ta} \in \mathbb{R}^{D_{out}}$ are the weight and bias learned by the encoder, and D_{in} and D_{out} are the dimensionalities of $\mathbf{H}_{dr\text{-}ta}$ and $\mathbf{\Phi}'_{dr-ta}$, respectively. After obtaining $\mathbf{\Phi}'_{dr-ta}$, two individual fully connected layers are used to estimate the means $\boldsymbol{\mu}_{dr\text{-}ta}$ and variances $\boldsymbol{\sigma}_{dr\text{-}ta}$:

$$\boldsymbol{\mu}_{dr-ta} = f(\mathbf{\Phi}'_{dr-ta}\mathbf{W}^{\boldsymbol{\mu}}_{dr-ta} + \mathbf{b}^{\boldsymbol{\mu}}_{dr-ta}) \quad \text{(Equation 2)}$$

$$\boldsymbol{\sigma}_{dr-ta} = f(\mathbf{\Phi}'_{dr-ta}\mathbf{W}^{\boldsymbol{\sigma}}_{dr-ta} + \mathbf{b}^{\boldsymbol{\sigma}}_{dr-ta}) \quad \text{(Equation 3)}$$

where $\mathbf{W}^{\boldsymbol{\mu}}_{dr-ta}, \mathbf{W}^{\boldsymbol{\sigma}}_{dr-ta} \in \mathbb{R}^{D_{out} \times D}$ and $\mathbf{b}^{\boldsymbol{\mu}}_{dr-ta}, \mathbf{b}^{\boldsymbol{\sigma}}_{dr-ta} \in \mathbb{R}^{D}$ are the learnable weights and biases, respectively. The dimensionality of the drug key embedding $\mathbf{\Phi}_d^k$ is D, and we sample this by

$$\mathbf{\Phi}_d^k = \boldsymbol{\mu}_{dr-ta} + \boldsymbol{\sigma}_{dr-ta} \odot \boldsymbol{\varepsilon} \quad \text{(Equation 4)}$$

where $\varepsilon \sim \mathbf{N}(0, \mathbf{I})$, and $\odot$ stands for the element-wise product.

The key embeddings characterize the high-order topological correlations from the direct relationships between targets and drugs. Recent studies have found that integrating multiple types of information can improve prediction accuracy.[43] For example, drug side effects are observable phenotypic effects resulting from drugs acting on genetic off-targets in human bodies.[44] Phenotypic side effect similarity can be used to infer whether two drugs share a target.[18] Hu et al.[45] found that targets can be used as bridges to link drugs and diseases. Inspired by these studies, we integrated additional types of association correlations in HHDTI to provide complementary information so that the method can predict correctly even in the case of extreme challenges like the cold-start problem.

As shown in Figure 1B, we learn drug side embeddings from the drug-disease incidence matrices ($\mathbf{H}_{dr\text{-}di}$) to provide complementary information for the drug key embeddings. This is achieved by the HGNN[26] model (Figure 1B, hypergraph convolutional layers). HGNN consists of hypergraph convolutional layers that encode high-order correlations:

$$\mathrm{Convh}(\mathbf{H}, \mathbf{X}|\mathbf{W}) = f\left((\mathbf{D}^v)^{-\frac{1}{2}}\mathbf{H}(\mathbf{D}^e)^{-1}\mathbf{H}^{\mathrm{T}}(\mathbf{D}^v)^{-\frac{1}{2}}\mathbf{X}\mathbf{W}\right) \quad \text{(Equation 5)}$$

where $\mathbf{D}^v$ and $\mathbf{D}^e$ are the diagonal degree matrices of the vertex and hyperedge respectively, with $(\mathbf{D}^v)_{k,k} = \sum_{j=1}^{L}\mathbf{H}^{k,j}$ being the degree of vertex and $(\mathbf{D}^e)_{j,j} = \sum_{k=1}^{N}\mathbf{H}^{k,j}$ being the degree of hyperedge. $\mathbf{X}$ denotes the vertex features, $\mathbf{W}$ is the learnable weight matrix, and $(\cdot)^{\mathrm{T}}$ is the transposition operator.

The output of the HGNN model is the side embeddings, which represent high-order correlations. The adopted HGNN has two hypergraph convolutional layers. Taking the drug side embedding learning on $\mathrm{H}_{dr\text{-}di}$ as an example, each layer can be formulated as

$$\mathbf{\Phi}_d^{s\,(l)} = \mathrm{Convh}\left(\mathbf{H}_{dr-di}, \mathbf{\Phi}_d^{s\,(l-1)}\middle|\mathbf{W}^{(l-1)}\right) \quad \text{(Equation 6)}$$

where $\mathbf{\Phi}_d^{s(l-1)}$, $\mathbf{\Phi}_d^{s(l)}$, and $\mathbf{W}^{(l-1)}$ are the input, output, and trainable weight matrix of the (l-1)-th layer, respectively. The vertex feature X is the inherent properties of the drugs, and we replaced with an identity matrix for $\mathbf{\Phi}_d^{s(0)} = \mathbf{X} = \mathbf{I}$. Then, we employ attention modules to fuse the key and side embeddings into a shared vector space to construct low-dimensional structural embeddings. We propose the bi-embedding attention fusion (Figure 1B, attention) to compute the coefficients ω^i to give different weights to the key embeddings and side embeddings:

$$\omega^i = \frac{\exp\left(f\left(\mathbf{\Phi}^i\mathbf{W}^i + \mathbf{b}^i\right)\cdot\mathbf{P}^i\right)}{\sum_{j\in k,s}\exp\left(f\left(\mathbf{\Phi}^j\mathbf{W}^j + \mathbf{b}^j\right)\cdot\mathbf{P}^j\right)} \quad \text{(Equation 7)}$$

where $\mathbf{\Phi}^i (i \in k, s)$ stands for key embeddings or side embeddings and $\mathbf{W}^i \in \mathbb{R}^{D \times D'}$, $\mathbf{b}^i \in \mathbb{R}^{D'}$, and $\mathbf{P}^i \in \mathbb{R}^{D' \times 1}$ are trainable parameters for embeddings $\mathbf{\Phi}^i$, respectively. D' is dimensionality of the trainable parameters. The overall structural embeddings $\mathbf{\Phi}^{\mathbf{S}}$ can be achieved by

$$\mathbf{\Phi}^{\mathbf{S}} = \omega^k\mathbf{\Phi}^k + \omega^s\mathbf{\Phi}^s \quad \text{(Equation 8)}$$

where $\mathbf{\Phi}^k$ and $\mathbf{\Phi}^s$ are the key and side embeddings, respectively.

Target structural embedding learning

By contrast, the target structural embedding learning uses the target-drug interaction hypergraph and the target-disease association hypergraph as inputs. It models the similarity and correlations of the targets interacting with the same drug to generate the target key embeddings $\mathbf{\Phi}_t^k$ through a vertex encoder (with the same structure as the vertex encoder in drug structural embedding learning). It also uses the HGNN[26] model to generate the target side embeddings $\mathbf{\Phi}_t^s$ from $\mathbf{H}_{ta\text{-}di}$ and fuses the target key embeddings and target side embeddings by biembedding attention fusion to obtain target structural embeddings $\mathbf{\Phi}_t^{\mathbf{S}}$.

DTI prediction

The DTI predictions are produced from the reconstruction space **A**, which is achieved by computing the likelihood of the drug and target structural embeddings.

$$\mathbf{A} = \mathrm{Sigmoid}\left(\mathbf{\Phi}_d^{\mathbf{S}}\left(\mathbf{\Phi}_t^{\mathbf{S}}\right)^{\mathbf{T}}\right) \quad \text{(Equation 9)}$$

where sigmoid(·) is the sigmoid activation function. We optimize the variational lower bound L:

$$\mathcal{L} = \mathbb{E}_q\left[\log p\left(\mathbf{A}\middle|\mathbf{\Phi}_d^{\mathbf{S}}, \mathbf{\Phi}_t^{\mathbf{S}}\right)\right] - \beta\left(\mathrm{KL}\left(q\left(\mathbf{\Phi}_d^k\middle|\mathbf{A}\right)\middle|\middle|p\left(\mathbf{\Phi}_d^k\right)\right) + \mathrm{KL}\left[q\left(\mathbf{\Phi}_t^k\middle|\mathbf{A}\right)\middle|\middle|p\left(\mathbf{\Phi}_t^k\right)\right]\right) \quad \text{(Equation 10)}$$

where KL[$q(\cdot)\|p(\cdot)$] is the Kullback-Leibler divergence between $q(\cdot)$ and $p(\cdot)$. Varying β encourages different learned representations by changing the degree of applied learning pressure during training. Referring to the work of the variational autoencoder, we further take Gaussian priors $p\left(\mathbf{\Phi}_d^k\right) = \prod_i p\left(\varphi_i^d\right) = \prod_i \mathcal{N}\left(\varphi_i^d\middle|0, \mathbf{I}\right)$ and $p\left(\mathbf{\Phi}_t^k\right) = \prod_j p\left(\varphi_j^t\right) = \prod_j \mathcal{N}\left(\varphi_j^t\middle|0, \mathbf{I}\right)$. $\mathrm{E}_q[\log p(\cdot|\cdot)]$ is the likelihood of reconstruction space **A** learned by HHDTI.

Model evaluation metrics

We introduced two evaluation metrics, the AUROC curve and the AUPR curve, to evaluate prediction performance. A confusion matrix is shown in Figure S3. In the receiver operating characteristic (ROC) space, the ROC curve gives a pair of x and y values where x is the false-positive rate (FPR) and y is the true-positive rate (TPR). We connected all points obtained by changing the cutoff to create the ROC curve.

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad \text{(Equation 11)}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{TN} + \mathrm{FP}} \quad \text{(Equation 12)}$$

where true-positives (TPs) and false-positives (FPs) are positive samples correctly predicted as positive and negative samples incorrectly predicted as positive, respectively. True-negatives (TNs) are negatives correctly identified as negative. False-negatives (FNs) correspond to positives incorrectly predicted as negative.

The precision-recall curve is plotted in a comparable way to the ROC curve but with the x axis being recall and the y axis being precision:

$$\mathrm{recall} = \frac{TP}{TP + FN} \quad \text{(Equation 13)}$$

$$\mathrm{precision} = \frac{TP}{TP + FN} \quad \text{(Equation 14)}$$

As discussed in previous work,[46,47] AUPR can provide a better assessment when the data for testing are highly skewed (supplementary note).

Datasets

The three public datasets proposed in DTINet,[1] deepDTnet,[2] and TriModel[3] (named DTINet_17, deepDTnet_20, and KEGG_MED, respectively) as well as the Target Drug-UniProt Links (approved) from the DrugBank database (version 5.1.0)[4] were used for evaluation.

The data in DTINet_17 were collected from public databases. Drug vertices, protein vertices, and disease vertices were obtained from the DrugBank database (version 3.0),[48] the HPRD database (release 9),[39] and CTD,[38] respectively. The known DTIs were imported from the DrugBank database (version 3.0),[48] and the drug-disease and target-disease associations were extracted from the CTD.[38]

The deepDTnet_20 dataset was also derived from the integration of information in multiple databases. The DTIs were collected from the DrugBank database (version 4.3),[30] the Therapeutic Target Database,[49] and the PharmGKB database.[50] The drug-disease association information came from the DrugBank database (version 4.3),[30] Drug Central,[37] and repoDB.[51] The drug-disease association data were integrated from the bioinformatics data sources CTD[38] and HuGe navigator.[52]

The KEGG_MED dataset was larger than the above two datasets and was extracted from multiple databases, including KEGG,[53] DrugBank database,[54] InterPro,[55] and UniProt.[56]

The Target Drug-UniProt Links (approved) dataset was extracted from the DrugBank database (version 5.1.0).[4]

More specific information regarding the four datasets is shown in Table S3. For more information about the datasets, please refer to the works of DTINet, deepDTnet, TriModel, and DrugBank database (version 5.1.0).

Statistical analysis

All statistical analyses were performed using GraphPad Prism software (version 8.0.2). The data shown in the study were obtained from at least five independent experiments. Values in different experimental groups are expressed as the mean ± standard deviation. $p < 0.05$ was considered statistically significant.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.patter.2021.100390.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China under grant 2020YFB1406604 and National Natural Science Foundation of China (62088102, U1701262, U1801263).

AUTHOR CONTRIBUTIONS

D.R., C.Z., Y.G., and Q.D. conceived the research project. D.R., S.J., C.Y., Y.G., and J.Z. designed the methodology. D.R., S.J., C.Y., J.Z., and C.Z. conducted experiments. D.R., X.Z., Y.Y., C.Z., Q.D., and Y.G. analyzed the results. All authors wrote the paper and contributed to the revision of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 11, 2021
Revised: July 23, 2021
Accepted: October 21, 2021
Published: November 16, 2021

REFERENCES

1. Luo, Y., Zhao, X., Zhou, J., Yang, J., Zhang, Y., Kuang, W., Peng, J., Chen, L., and Zeng, J. (2017). A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. Nat. Commun. *8*, 1–13.
2. Zeng, X., Zhu, S., Lu, W., Liu, Z., Huang, J., Zhou, Y., Fang, J., Huang, Y., Guo, H., and Li, L. (2020). Target identification among known drugs by deep learning from heterogeneous networks. Chem. Sci. *11*, 1775–1797.
3. Mohamed, S.K., Nováček, V., and Nounu, A. (2020). Discovering protein drug targets using knowledge graph embeddings. Bioinformatics *36*, 603–610.
4. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2017). DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. *46*, D1074–D1082.
5. Barabási, A.-L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. Nat. Rev. Genet. *12*, 56–68.
6. Whitebread, S., Hamon, J., Bojanic, D., and Urban, L. (2005). Keynote review: in vitro safety pharmacology profiling: an essential tool for successful drug development. Drug Discov. Today *10*, 1421–1433.
7. Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., and Olson, A.J. (2009). AutoDock4 and AutoDockTools4:

Automated docking with selective receptor flexibility. J. Comput. Chem. *30*, 2785–2791.

8. Keiser, M.J., Roth, B.L., Armbruster, B.N., Ernsberger, P., Irwin, J.J., and Shoichet, B.K. (2007). Relating protein pharmacology by ligand chemistry. Nat. Biotechnol. *25*, 197–206.
9. Bagherian, M., Sabeti, E., Wang, K., Sartor, M.A., Nikolovska-Coleska, Z., and Najarian, K. (2020). Machine learning approaches and databases for prediction of drug-target interaction: a survey paper. Brief. Bioinform. *22*, 247–269.
10. Lavecchia, A. (2015). Machine-learning approaches in drug discovery: methods and applications. Drug Discov. Today *20*, 318–331.
11. Mayr, A., Klambauer, G., Unterthiner, T., Steijaert, M., Wegner, J.K., Ceulemans, H., Clevert, D.-A., and Hochreiter, S. (2018). Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. Chem. Sci. *9*, 5441–5451.
12. Lo, Y.-C., Rensi, S.E., Torng, W., and Altman, R.B. (2018). Machine learning in chemoinformatics and drug discovery. Drug Discov. Today *23*, 1538–1546.
13. Yamanishi, Y., Araki, M., Gutteridge, A., Honda, W., and Kanehisa, M. (2008). Prediction of drug-target interaction networks from the integration of chemical and genomic spaces. Bioinformatics *24*, i232–i240.
14. Yu, H., Chen, J., Xu, X., Li, Y., Zhao, H., Fang, Y., Li, X., Zhou, W., Wang, W., and Wang, Y. (2012). A systematic prediction of multiple drug-target interactions from chemical, genomic, and pharmacological data. PLoS one *7*, e37608.
15. Gao, K.Y., Fokoue, A., Luo, H., Iyengar, A., Dey, S., and Zhang, P. (2018). Interpretable Drug Target Prediction Using Deep Neural Representation (IJCAI).
16. Nascimento, A.C., Prudêncio, R.B., and Costa, I.G. (2016). A multiple kernel learning algorithm for drug-target interaction prediction. BMC Bioinformatics *17*, 46.
17. Zong, N., Kim, H., Ngo, V., and Harismendy, O. (2017). Deep mining heterogeneous networks of biomedical linked data to predict novel drug-target associations. Bioinformatics *33*, 2337–2344.
18. Campillos, M., Kuhn, M., Gavin, A.-C., Jensen, L.J., and Bork, P. (2008). Drug target identification using side-effect similarity. Science *321*, 263–266.
19. Cheng, F., Liu, C., Jiang, J., Lu, W., Li, W., Liu, G., Zhou, W., Huang, J., and Tang, Y. (2012). Prediction of drug-target interactions and drug repositioning via network-based inference. PLoS Comput. Biol. *8*, e1002503.
20. Chen, B., Ding, Y., and Wild, D.J. (2012). Assessing drug target association using semantic linked data. PLoS Comput. Biol. *8*, e1002574.
21. Lü, L., and Zhou, T. (2011). Link prediction in complex networks: A survey. Physica A *390*, 1150–1170.
22. Wan, F., Hong, L., Xiao, A., Jiang, T., and Zeng, J. (2019). NeoDTI: neural integration of neighbor information from a heterogeneous network for discovering new drug–target interactions. Bioinformatics *35*, 104–111.
23. Pei, J., Yin, N., Ma, X., and Lai, L. (2014). Systems biology brings new dimensions for structure-based drug design. J. Am. Chem. Soc. *136*, 11556–11565.
24. Vaida, M., and Purcell, K. (2019). Hypergraph link prediction: learning drug interaction networks embeddings. In Paper Presented at: 2019 18th IEEE International Conference on Machine Learning and Applications (ICMLA) (IEEE), pp. 1860–1865. https://doi: 10.1109/ICMLA.2019.00299.
25. Niu, Y.W., Qu, C.Q., Wang, G.H., and Yan, G.Y. (2019). RWHMDA: random walk on hypergraph for microbe-disease association prediction. Front. Microbiol. *10*, 1578.
26. Feng, Y., You, H., Zhang, Z., Ji, R., and Gao, Y. (2019). Hypergraph neural networks. In Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI), pp. 3558–3565.
27. Powers, D.M. (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. J. Mach. Learn. Tech. *2*, 37–63.
28. Natarajan, N., and Dhillon, I.S. (2014). Inductive matrix completion for predicting gene–disease associations. Bioinformatics *30*, i60–i68.
29. Singh-Blom, U.M., Natarajan, N., Tewari, A., Woods, J.O., Dhillon, I.S., and Marcotte, E.M. (2013). Prediction and validation of gene-disease associations using methods inspired by social network analyses. PLoS One *8*, e58977.
30. Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A.C., Liu, Y., Maciejewski, A., Arndt, D., Wilson, M., Neveu, V., et al. (2013). DrugBank 4.0: shedding new light on drug metabolism. Nucleic Acids Res. *42*, D1091–D1097.
31. Hayashi, T., Juliet, P.A., Miyazaki-Akita, A., Funami, J., Matsui-Hirai, H., Fukatsu, A., and Iguchi, A. (2007). beta1 antagonist and beta2 agonist, celiprolol, restores the impaired endothelial dependent and independent responses and decreased TNFalpha in rat with type II diabetes. Life Sci. *80*, 592–599.
32. Nawarskas, J.J., Cheng-Lai, A., and Frishman, W.H. (2017). Celiprolol: a unique selective adrenoceptor modulator. Cardiol. Rev. *25*, 247–253.
33. Hopkins, A.L. (2008). Network pharmacology: the next paradigm in drug discovery. Nat. Chem. Biol. *4*, 682–690.
34. Lim, H., Gray, P., Xie, L., and Poleksic, A. (2016). Improved genome-scale multi-target virtual screening via a novel collaborative filtering approach to cold-start problem. Sci. Rep. *6*, 1–11.
35. Yao, Y., Tong, H., Yan, G., Xu, F., Zhang, X., Szymanski, B.K., and Lu, J. (2014). Dual-regularized one-class collaborative filtering. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management (ACM), pp. 759–768.
36. Sun, M., Zhao, S., Gilvary, C., Elemento, O., Zhou, J., and Wang, F. (2019). Graph convolutional networks for computational drug development and discovery. Brief. Bioinform. *21*, 919–935.
37. Ursu, O., Holmes, J., Knockel, J., Bologa, C.G., Yang, J.J., Mathias, S.L., Nelson, S.J., and Oprea, T.I. (2016). DrugCentral: online drug compendium. Nucleic Acids Res. *45*, D932–D939.
38. Davis, A.P., Grondin, C.J., Johnson, R.J., Sciaky, D., McMorran, R., Wiegers, J., Wiegers, T.C., and Mattingly, C.J. (2018). The Comparative Toxicogenomics Database: update 2019. Nucleic Acids Res. *47*, D948–D954.
39. Keshava Prasad, T.S., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., Telikicherla, D., Raju, R., Shafreen, B., Venugopal, A., et al. (2009). Human protein reference database—2009 update. Nucleic Acids Res. *37*, D767–D772.
40. Luck, K., Kim, D.-K., Lambourne, L., Spirohn, K., Begg, B.E., Bian, W., Brignall, R., Cafarelli, T., Campos-Laborie, F.J., Charloteaux, B., et al. (2020). A reference map of the human binary protein interactome. Nature *580*, 402–408.
41. Kingma, D.P., and Welling, M. (2013). Auto-Encoding Variational Bayes (arXiv), arXiv:1312.6114.
42. Gardner, M.W., and Dorling, S.R. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmos. Environ. *32*, 2627–2636.
43. Madhukar, N.S., Khade, P.K., Huang, L., Gayvert, K., Galletti, G., Stogniew, M., Allen, J.E., Giannakakou, P., and Elemento, O. (2019). A Bayesian machine learning approach for drug target identification using diverse data types. Nat. Commun. *10*, 1–14.
44. Zhou, M., Chen, Y., and Xu, R. (2019). A drug-side effect context-sensitive network approach for drug target prediction. Bioinformatics *35*, 2100–2107.
45. Hu, Q.N., Deng, Z., Tu, W., Yang, X., Meng, Z.B., Deng, Z.X., and Liu, J. (2014). VNP: interactive visual network pharmacology of diseases, targets, and drugs. CPT Pharmacometrics Syst. Pharmacol. *3*, e105.
46. van Laarhoven, T., Nabuurs, S.B., and Marchiori, E. (2011). Gaussian interaction profile kernels for predicting drug–target interaction. Bioinformatics *27*, 3036–3043.

47. Davis, J., and Goadrich, M. (2006). The relationship between precision-recall and ROC curves. In Proceedings of the 23rd International Conference on Machine Learning (ACM), pp. 233–240.

48. Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., et al. (2010). DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Nucleic Acids Res. *39*, D1035–D1041.

49. Yang, H., Qin, C., Li, Y.H., Tao, L., Zhou, J., Yu, C.Y., Xu, F., Chen, Z., Zhu, F., and Chen, Y.Z. (2015). Therapeutic target database update 2016: enriched resource for bench to clinical drug target and targeted pathway information. Nucleic Acids Res. *44*, D1069–D1074.

50. Sangkuhl, K., Berlin, D.S., Altman, R.B., and Klein, T.E. (2008). PharmGKB: understanding the effects of individual genetic variants. Drug Metab. Rev. *40*, 539.

51. Brown, A.S., and Patel, C.J. (2017). A standard database for drug repositioning. Sci. Data *4*, 170029.

52. Yu, W., Gwinn, M., Clyne, M., Yesupriya, A., and Khoury, M.J. (2008). A navigator for human genome epidemiology. Nat. Genet. *40*, 124–125.

53. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. *45*, D353–D361.

54. Wishart, D.S., Knox, C., Guo, A.C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., and Woolsey, J. (2006). DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res. *34*, D668–D672.

55. Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.-Y., El-Gebali, S., Fraser, M.I., et al. (2018). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res. *47*, D351–D360.

56. Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., et al. (2004). UniProt: the universal protein knowledgebase. Nucleic Acids Res. *32*, D115–D119.