OXFORD

Briefings in Bioinformatics, 2022, **23(4)**, 1–12
https://doi.org/10.1093/bib/bbac209
Advance access publication date: 7 June 2022
Problem Solving Protocol

STNN-DDI: a Substructure-aware Tensor Neural Network to predict Drug–Drug Interactions

Hui Yu, ShiYu Zhao and JianYu Shi

Corresponding authors: Hui Yu, School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China. Tel.: +862988431537; E-mail: huiyu@nwpu.edu.cn; Jian-yu Shi, School of Life Sciences, Northwestern Polytechnical University, Xi'an 710072, China. Tel.: +862988460332; E-mail: jianyushi@nwpu.edu.cn

Abstract

Computational prediction of multiple-type drug–drug interaction (DDI) helps reduce unexpected side effects in poly-drug treatments. Although existing computational approaches achieve inspiring results, they ignore to study which local structures of drugs cause DDIs, and their interpretability is still weak. In this paper, by supposing that the interactions between two given drugs are caused by their local chemical structures (substructures) and their DDI types are determined by the linkages between different substructure sets, we design a novel Substructure-aware Tensor Neural Network model for DDI prediction (STNN-DDI). The proposed model learns a 3-D tensor of ⟨ *substructure, substructure, interaction type* ⟩ triplets, which characterizes a substructure–substructure interaction (SSI) space. According to a list of predefined substructures with specific chemical meanings, the mapping of drugs into this SSI space enables STNN-DDI to perform the multiple-type DDI prediction in both transductive and inductive scenarios in a unified form with an explicable manner. The comparison with deep learning-based state-of-the-art baselines demonstrates the superiority of STNN-DDI with the significant improvement of AUC, AUPR, Accuracy and Precision. More importantly, case studies illustrate its interpretability by both revealing an important substructure pair across drugs regarding a DDI type of interest and uncovering interaction type-specific substructure pairs in a given DDI. In summary, STNN-DDI provides an effective approach to predicting DDIs as well as explaining the interaction mechanisms among drugs. Source code is freely available at https://github.com/zsy-9/STNN-DDI.

Keywords: Tensor Neural Network, drug–drug interactions, substructure–substructure interactions, multi-type interactions

Introduction

In the treatment of human complex disease, co-prescription of drugs has become a common therapy [1] because individual drugs often fail to meet clinical needs. However, two or more drugs are taken together would induce pharmacokinetic or pharmacodynamic changes [2], referred to as drug–drug interactions (DDIs), which may trigger unexpected adverse drug events [3]. DDI-triggered adverse effects are harmful even dangerous to patients [4]. Therefore, the detection of potential DDIs before making co-prescriptions has been gaining attention from biologists, pharmacologists and clinicians. Due to the enormous combinational number of drugs, traditional methods (e.g. in-vitro approaches and clinical trials) are time-consuming, low-efficiency and costly [5]. In recent years, computational methods have been demonstrated their ability to infer potential DDIs rapidly and cheaply [6, 7].

Most computational methods for DDI prediction are based on machine learning, which can be roughly categorized into four classes: feature-based, similarity-based, network-based and matrix decomposition-based methods. Feature-based methods generally represent drugs into feature vectors derived from various drug properties, such as chemical structure [8, 9], side-effect [10], etc. For example, Zhang *et al.* [11] collected a variety of drug features, i.e. pathway, indication, transporter, etc., and build prediction models, respectively, by the neighbor recommender method, the random walk method and the matrix perturbation method. Chu *et al.* [12] constructed an autoencoder to predict potential DDIs. Similarity-based methods assume that two similar drugs tend to have common interactions. They usually calculate the similarity between drugs and represent drugs by similarity vectors [13, 14], then predict DDIs via different prediction models. For instance, Gottlieb *et al.* [7] used 7 types of drug features to generate similarity vectors and constructed a DDI prediction model by logistic-based regression. Ferdousi *et al.* [15] utilized structure similarity vectors as the input of a deep neural network to predict potential DDIs. By constructing the knowledge graph of DDIs, network-based methods generally obtain drug embeddings by their local or global topological information in the graph to infer potential DDIs. For example,

Hui Yu received the Master's and PhD degrees from Northwestern Polytechnical University, Xi'an, China, where he is currently an Associate Professor. His research interest includes bioinformatics, granular computing, three-way decisions and data mining.
ShiYu Zhao is currently pursuing her Master's degree in computer science at Northwestern Polytechnical University, Xi'an, China. She is interested in graph representation learning and applications.
JianYu Shi received his PhD degree from Northwestern Polytechnical University, Xi'an, China, where he is currently a Professor. His research interests include bioinformatics, cheminformatics and artificial intelligence.
Received: February 22, 2022. **Revised:** April 25, 2022. **Accepted:** May 5, 2022

Wang *et al.* [16] generated drug embedding representation on the drug structure graph by graph convolution neural network (GCN). Huang *et al.* [17] utilized GCN to generate the first- and second-order neighbors to represent drugs and predicted DDIs by a coefficient matrix. Takeda *et al.* [18] represented drugs by the position similarity in the DDI knowledge graph and predicted DDIs by the logistic regression model. Matrix factorization-based methods decompose the adjacency matrix of DDIs into several factor matrices and reconstruct the adjacency matrix to identify potential DDIs [19–21]. Yu *et al.* [22] predicted potential DDIs through non-negative matrix factorization. Rohani *et al.* [23] integrated a nonlinear multi-similarity fusion with matrix factorization to identify potential DDIs.

Although existing methods have achieved inspiring prediction, some issues are still challenging as follows.

- Inductive prediction. Practical applications usually require two inductive scenarios of DDIs prediction [24]: to deduce interactions between new drugs and known drugs or interactions among new drugs. However, many methods, such as network-based and matrix decomposition-based methods, cannot meet the demands in the inductive scenario [25].
- Multi-label interaction prediction. Our observation shows that two drugs may trigger one or more types of drug interactions (referred to as multi-label interaction in this article); however, most existing methods only predict whether two drugs trigger an interaction or trigger a specific type of interaction (referred to as ordinary multi-type interactions in this article).
- Interpretability. A predictive model with good interpretability surely helps clinicians to uncover the underlying mechanism of DDIs. Nevertheless, the interpretability of many existing methods is weak yet. Especially, the methods are derived from knowledge-graph or deep learning.

Based on the medicinal chemistry knowledge [26–28], a drug is an entity which composed by different functional groups/chemical substructures who determine all of its pharmacokinetic and pharmacodynamic properties, and ultimately all of its interactions. Therefore, we assume that the interactions between two drugs are determined by the occurrence of their substructures and their DDI types are determined by the corresponding type of linkages between their substructure sets. As illustrated in Figure 1, two drugs, denoted as d_a and d_b, are represented as a set of substructures with specific chemical meanings, respectively. They trigger four types of interactions (i.e. multiple-label interactions, such as 'blood calcium increased', 'narcolepsy', 'pyoderma', 'apoplexy' and so on.), colored by yellow, blue, orange and green from top to bottom (Figure 1a). Each type of their interactions is caused by specific associations between two substructure sets derived from the two drugs, respectively. For example, the yellow interaction is mainly caused by two pairs of substructures (the pair of s_1^a in d_a and s_3^b in d_b as well as the pair of s_n^a and s_1^b). The blue interaction is mainly caused by only one pair of substructures, s_3^a in d_a and s_2^b in d_b (Fig. 1b). Overall, the multiple-to-multiple associations between their substructures determine their interactions. Also, a substructure may attend multiple-type interactions. In the context of DDI, we term these associations between substructures as **s**ubstructure–**s**ubstructure **i**nteractions (SSIs).

Under this assumption, we propose a Substructure-aware Tensor Neural Network for DDI Prediction, referred to as STNN-DDI. The proposed model learns a *substructure* × *substructure* × *interaction* tensor (named ***ST***), which characterizes a SSI space, expanded by a series of rank-one tensors [29]. According to a list of predefined substructures with specific chemical meanings (e.g. PubChem fingerprint), two given drugs are embedded into this SSI space to discriminate what types of interactions they trigger and how likely they trigger a specific type of interaction. By leveraging the substructure tensor ***ST***, STNN-DDI has the following advantages.

- It provides a unified form for the DDI prediction in both transductive and inductive scenarios, where the latter ac-counts for both predicting interactions between new drugs and known drugs and predicting interactions among new drugs.
- It also provides an integrated solution for the case that two drugs trigger single interaction (ordinary multi-type interactions) and the case that they trigger multiple-type interactions simultaneously (multi-label interactions).
- More importantly, its explicit interpretability enables to both reveal an important substructure pair across drugs regarding a DDI type of interest and uncover major interaction type-specific substructure pairs in a given DDI.

The remaining sections are organized as follows. Section 2 briefly introduces the preliminary knowledge of tensors. Section 3 formulates the problem of DDI prediction and gives the detailed modeling and solution of STNN-DDI. Section 4 designs corresponding experiments to demonstrate the superiority of STNN-DDI in different scenarios and illustrate its interpretability. Section 5 concludes our contributions.

Preliminary

To illustrate STNN-DDI clearly, in this section, we shall briefly review several basic definitions, including tensor and its operations. A tensor is the generalization of a matrix to >2 dimensions and can consequently be treated as multidimensional fields [29]. Its decompositions originally appeared in 1927 [30] but have remained untouched by the computer science community until the late 20th century [31]. Several core terms about tensor are used in this paper.

Tensor: An *N*-way or *N*th-order tensor is an element of the tensor product of *N* vector spaces, each of which has its coordinate system.

Rank-one tensors: An *N*-order tensor $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined as rank-one tensor when it can be strictly decomposed into the outer product of *N* vectors. $\boldsymbol{\mathcal{Y}}$ is given as

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{a}^{(1)} \circ \boldsymbol{a}^{(2)} \circ \cdots \circ \boldsymbol{a}^{(N)}, \tag{1}$$

where $(\circ)$ is the vector outer product and $\boldsymbol{a}^{(i)}$ is a vector. Each element of tensor $\boldsymbol{\mathcal{Y}}$ is the product of the corresponding vector elements, the operation can be written as

$$y_{i_1 i_2 \ldots i_N} = {a_{i_1}}^{(1)} {a_{i_2}}^{(2)} \ldots {a_{i_N}}^{(N)}, \tag{2}$$

where $a_{i_j}^{(j)}$ is the i_jth element of $a^{(j)}$. For a third-order rank-one tensor $\boldsymbol{\mathcal{Y}} = \boldsymbol{a} \circ \boldsymbol{b} \circ \boldsymbol{c}$, the (i, j, k) element of $\boldsymbol{\mathcal{Y}}$ is given by $y_{ijk} = a_i b_j c_k$.

Canonical Polyadic (CP) Decomposition: An *N*-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ can be approximately expressed by the sum of rank-one tensors, the operation is defined as

$$\boldsymbol{\mathcal{X}} \approx \sum_{r=1}^{R} \lambda_r \boldsymbol{\mathcal{Y}}_r, \tag{3}$$

where $\boldsymbol{\mathcal{Y}}_r \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is a rank-one tensor, λ_r is the weight of $\boldsymbol{\mathcal{Y}}_r$. The CP Decomposition of a 3-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I \times J \times K}$ can be expressed as

$$\boldsymbol{\mathcal{X}} \approx \sum_{r=1}^{R} \lambda_r (\boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r) = \langle \boldsymbol{\lambda}; \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \rangle, \tag{4}$$

where R is the rank of tensor $\boldsymbol{\mathcal{X}}$, λ_r is the weight of the rth rank-one tensor, and $\boldsymbol{a}_r \in \mathbb{R}^I$, $\boldsymbol{b}_r \in \mathbb{R}^J$, $\boldsymbol{c}_r \in \mathbb{R}^K$ are the vectors decomposed by $\boldsymbol{\mathcal{X}}$ in three directions, respectively. Vector $\boldsymbol{\lambda} \in \mathbb{R}^R$ is composed by λ_r. Factor matrices $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots \boldsymbol{a}_R]$, $\boldsymbol{B} = [\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_R]$, $\boldsymbol{C} = [\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_R]$ are the combination of vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ in columns, where $\boldsymbol{A} \in \mathbb{R}^{I \times R}$, $\boldsymbol{B} \in \mathbb{R}^{J \times R}$ and $\boldsymbol{C} \in \mathbb{R}^{K \times R}$, respectively. These matrices are also called the 2-dimensional representation of $\boldsymbol{\mathcal{X}}$.

STNN-DDI modeling

In this section, we will describe STNN-DDI technically, including problem definition, notations, mathematical derivation and the construction of the neural network.

Problem definition

This paper aims to address three tasks of DDIs simultaneously as follows (Figure 2):

- C1: to infer potential pairwise interactions (marked by dashed lines) between known drugs,
- C2: to infer potential interactions between known drugs having interactions and new drugs,
- C3: to infer potential pairwise interactions among new drugs,

where we refer to known drugs as the drugs which have reported interactions and new drugs as the drugs which have no known interaction yet. During the construction of a machine learning model, both known drugs and their partial interactions attend the training, whereas neither new drugs nor their interaction attends the training. New drugs are only used to test the performance of the trained model in the three abovementioned tasks of DDI prediction.

Notations of STNN-DDI

Let $S = \{s_i\}_{i=1}^n$ be the set of all n substructures or functional groups in a list of predefined substructures with explicit chemical meanings (e.g. PubChem fingerprint), $D = \{d_i\}_{i=1}^m$ be the set of m drugs, $L = \{l_i\}_{i=1}^f$ be the set of all f relations between drug pairs. Here, l_k is encoded by a one-hot vector $\boldsymbol{v}_k = (v_i^k)_{i=1}^f$ and d_p is encoded by Pubchem fingerprint, namely $\boldsymbol{e}_p = (e_i^p)_{i=1}^n$. Table 1 lists the notations of STNN-DDI used in the following sections.

Tensor construction in STNN-DDI model

The underlying assumption of STNN-DDI is that the interaction between two drugs is determined by the occurrence of their substructures and their DDI types are determined by the linkages between different substructure sets. Thus, we model DDI from the perspective of SSI by P_{pqk}^d, which is defined as

$$P_{pqk}^d = \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ijk}^s e_i^p e_j^q. \tag{5}$$

It represents the occurring probability of the triplet $\langle d_p, d_q, l_k \rangle$ as the sum of the occurring probability of $\langle s_i, s_j, l_k \rangle$, in which s_i and s_j are the substructures included in drug pair (d_p, d_q), and the types of SSI are defined as the same as the corresponding types of DDI.

Following the above definition, by learning the probability of each SSI under a list of predefined chemical substructures, potential DDIs can be obtained. Thus, a *substructure* × *substructure* × *interaction* tensor, named $\boldsymbol{ST}$, is designed. The construction and application of $\boldsymbol{ST} \in \mathbb{R}^{n \times n \times f}$ is given in Figure 3, where coordinate axis x or y enumerates the set of predefined substructures, coordinate axis z enumerates the set of all interaction types in DDI set, which is referred to as the SSI space. Because of the probability $P_{ijk}^s = ST_{ijk}$, as Figure 3, the calculation of P_{pqk}^d by $\boldsymbol{ST}$ can be rewritten as

$$P_{pqk}^d = \sum_{i=1}^{n} \sum_{j=1}^{n} ST_{ijk} e_i^p e_j^q. \tag{6}$$

This form allows STNN-DDI to predict DDIs in both transductive and inductive scenarios in a unified form

Table 1. Notations of STNN-DDI.

n	n=881/167, which is the number of all the chemical substructures in the PubChem/MACCS fingerprint.
m	m=555, which is the number of drugs in the dataset.
f	f=1318, which is the number of all types of interactions in the dataset.
s_i	The ith chemical substructure or functional group in S
d_p	The pth drug in D.
$\boldsymbol{e}_p$	A vector, fingerprint of drug d_p.
e_i^p	The ith element of e_p, if d_p contains s_i, $e_i^p = 1$, else $e_i^p = 0$.
l_k	The kth relation (interaction type) in L.
$\boldsymbol{v}_k$	A one-hot vector, the embedding of the relation l_k.
v_i^k	The ith element of $\boldsymbol{v}_k$, if $i = k$, v_i^k=1, else v_i^k=0
DDI_{pqk}	Represents the triplet $\langle d_p, d_q, l_k\rangle$, which has the relation l_k in drug pair (d_p, d_q).
P_{pqk}^d	Probability of DDI_{pqk}.
SSI_{ijk}	Represents the triplet $\langle s_i, s_j, l_k\rangle$, which is the relation l_k in (s_i, s_j).
P_{ijk}^s	Probability of SSI_{ijk}.
$\boldsymbol{ST}$	Substructure–substructure–interaction Tensor.
ST_{ijk}	The value of the element of the ith row, jth column and kth dimension in $\boldsymbol{ST}$.
$\widehat{P_{pqk}^d}$	Estimated value of P_{pqk}^d calculated by STNN-DDI.

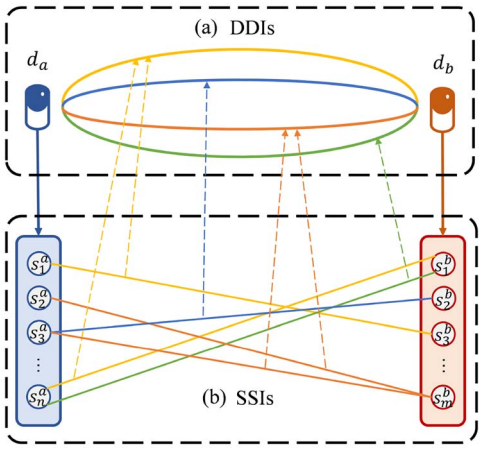


Figure 1. SSIs and DDIs. (d_a, d_b) is a drug pair, s_i^a is the substructure of drug d_a, s_i^b is the substructure of drug d_b, different colored edges represent the type of different interaction.

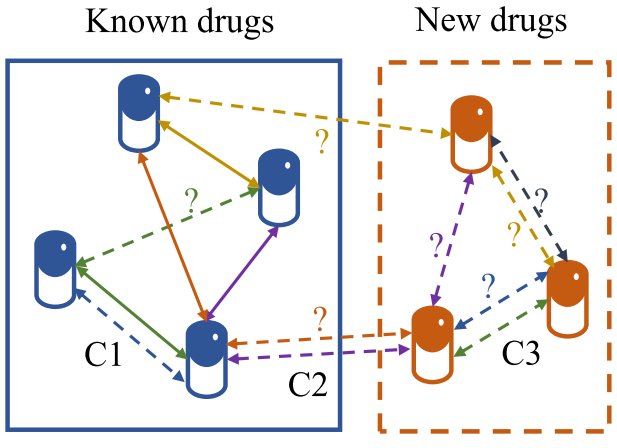


Figure 2. Problem definition. Solid lines represent approved DDIs, and dashed lines represent interactions to be predicted. Different colors of lines represent multiple types of DDIs. The left panel illustrates a graph containing the interactions between known drugs, while the right panel shows a graph containing new drugs. Three prediction tasks are marked by C1, C2 and C3, respectively.

because both known drugs and new drugs are mapped into a common SSI space no matter whether a drug has an interaction or not.

The DDI prediction by tensor *ST*

To perform the DDI prediction via tensor $\boldsymbol{ST}$, we implement Eq. 6 by the n-mode product for tensor and vectors. Given an N-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ and a vector $\boldsymbol{v} \in \mathbb{R}^{I_n}$, the n-mode product for $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{v}$ is denoted as

$$(\boldsymbol{\mathcal{X}} \overline{\times}_n \boldsymbol{v})_{i_1 \ldots i_{n-1} i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 \ldots i_N} \boldsymbol{v}_{i_n}. \tag{7}$$

Therefore, with the n-mode product mathematical principle, P_{pqk}^d can be calculated by $\boldsymbol{ST}$ mode product with $\boldsymbol{v}_k$, $\boldsymbol{e}_p$, $\boldsymbol{e}_q$ individually, this operation can be expressed as follows:

$$P_{pqk}^d = \boldsymbol{ST} \overline{\times}_3 \boldsymbol{v}_k \overline{\times}_2 \boldsymbol{e}_p \overline{\times}_1 \boldsymbol{e}_q. \tag{8}$$

The mathematical principles are given as follows and shown in Figure 4.

First, $\boldsymbol{ST}_{::k} = \boldsymbol{ST} \overline{\times}_3 \boldsymbol{v}_k$, where the frontal slices $\boldsymbol{ST}_{::k}$ can be described as a *substructure* × *substructure* matrix for interaction l_k.

Then, $\boldsymbol{st}_{(p,k)} = \boldsymbol{ST}_{::k} \overline{\times}_2 \boldsymbol{e}_p$, the jth element of the vector $st_j^{(p,k)} = \sum_{i=1}^n ST_{ijk} e_i^p$ can be described as the probability of interaction l_k happens caused by s_j under the action of d_p.

Finally, $P_{pqk}^d = \boldsymbol{st}_{(p,k)} \overline{\times}_1 \boldsymbol{e}_q = \sum_{j=1}^n st_j^{(p,k)} e_j^q$, which can also be written as $P_{pqk}^d = \sum_{i=1}^n \sum_{j=1}^n ST_{ijk} e_i^p e_j^q$, P_{pqk}^d can be calculated according to the structure of d_q.

The factor matrices of *ST*

According to Eq. 4, a tensor can be approximated by the factor matrices of CP decomposition. Therefore, the

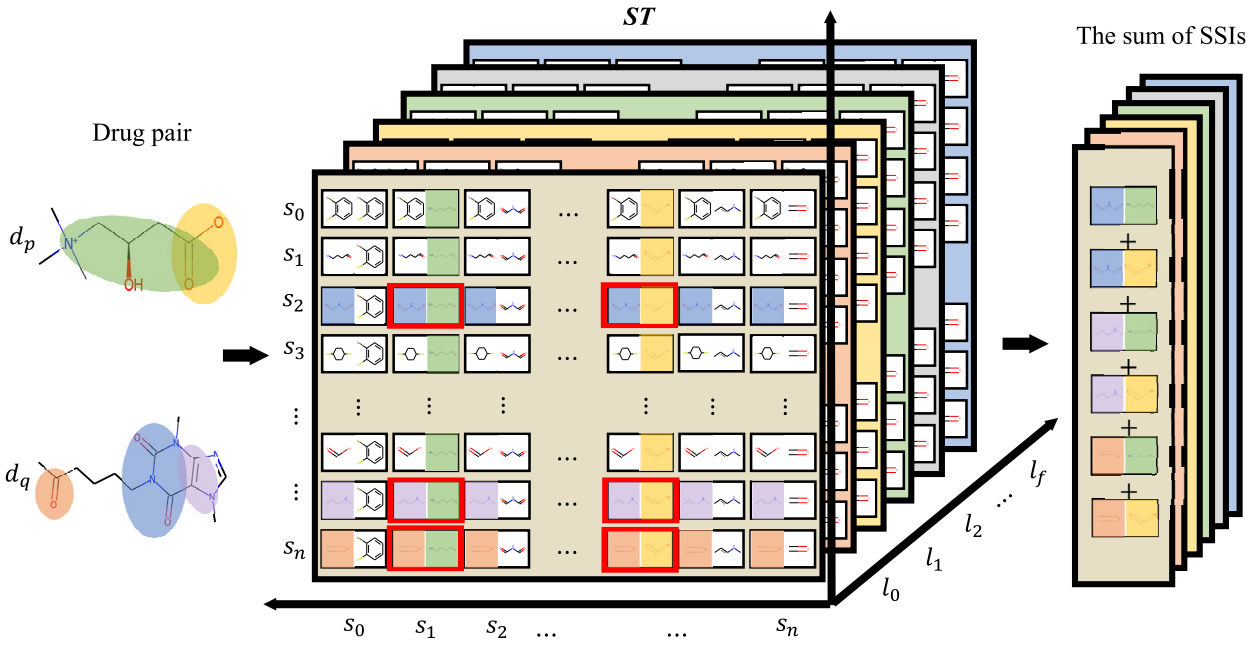


Figure 3. Construction of tensor ***ST***. Axis x and y are substructures and z represents the interactions, substructures in the drug pair (d_p, d_q) are highlighted by semitransparent colorful ellipses and the probability P^d_{pqk} of $\langle d_p, d_q, l_k \rangle$ is the sum of the occurring probability of $\langle s_i, s_j, l_k \rangle$.

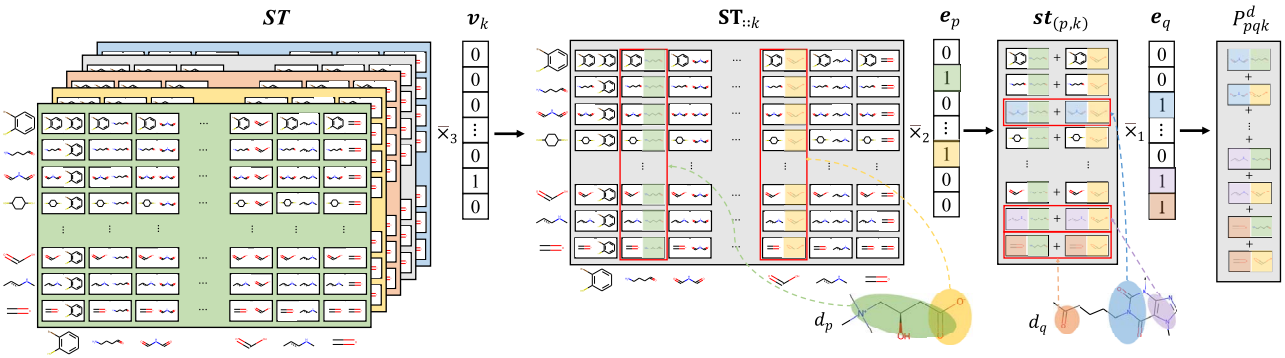


Figure 4. ***ST*** mode product with $\boldsymbol{v}_k$, $\boldsymbol{e}_p$, $\boldsymbol{e}_q$ to calculate P^d_{pqk}, substructures in the drug pair (d_p, d_q) are highlighted by semitransparent colorful ellipses.

unknown ***ST*** can be obtained by studying the matrices, that is ***ST*** can be represented as factor matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n \times R}$ and $\boldsymbol{C} \in \mathbb{R}^{f \times R}$, where R represents the number of rank-one tensors decomposed by ***ST*** and R also is a manually parameter in STNN-DDI.

In CP decomposition, each factor matrix collects the latent information of a specific dimension. $\boldsymbol{A}$ collects the latent information of axis x, $\boldsymbol{B}$ collects the latent information of axis y and C collects the latent information of axis z. Therefore, each row of matrix $\boldsymbol{A} \in \mathbb{R}^{n \times R}$ and $\boldsymbol{B} \in \mathbb{R}^{n \times R}$ can be regard as the embedding of each substructure. Hence, there is $\boldsymbol{A} = \boldsymbol{B}$. While matrix $\boldsymbol{C} \in \mathbb{R}^{f \times R}$ can be regard as the embedding of the relations. Then, ***ST*** can be expressed as

$$\boldsymbol{ST} \approx \sum_{r=1}^{R} \lambda_r (\boldsymbol{a}_r \circ \boldsymbol{a}_r \circ \boldsymbol{c}_r) = \langle \boldsymbol{\lambda}; \boldsymbol{A}, \boldsymbol{A}, \boldsymbol{C} \rangle. \tag{9}$$

As mentioned in section 3.4, P^d_{pqk} can be calculated by n-mode product for ***ST*** and $\boldsymbol{v}_k$, $\boldsymbol{e}_p$, $\boldsymbol{e}_q$, and ***ST*** can be approximated by the factor matrices $\boldsymbol{A}$ and $\boldsymbol{C}$.

Then, to predict DDIs by these factor matrices, Multilinear Tensor Transformation is introduced [32]. A tensor $\boldsymbol{\mathcal{X}}$ can be transformed on multiple dimensions by hitting each vector producing $\boldsymbol{\mathcal{X}}$ with a transformation vector (or matrix) from the left side. In the 3-dimensional case, the multilinear tensor transformation using vectors $\boldsymbol{v}_i$ is defined as

$$\begin{aligned} &\boldsymbol{\mathcal{X}} \bar{\times}_3 \boldsymbol{v}_3 \bar{\times}_2 \boldsymbol{v}_2 \bar{\times}_1 \boldsymbol{v}_1 \\ &\approx \sum_{r=1}^{R} \lambda_r (\boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r) \bar{\times}_3 \boldsymbol{v}_3 \bar{\times}_2 \boldsymbol{v}_2 \bar{\times}_1 \boldsymbol{v}_1 \\ &= \sum_{r=1}^{R} \lambda_r (\boldsymbol{v}_1^T \cdot \boldsymbol{a}_r)(\boldsymbol{v}_2^T \cdot \boldsymbol{b}_r)(\boldsymbol{v}_3^T \cdot \boldsymbol{c}_r), \end{aligned} \tag{10}$$

where $(\cdot)$ is the dot product between vectors. Therefore, P^d_{pqk} can be calculated as

$$\begin{aligned} P^d_{pqk} &\approx \sum_{r=1}^{R} \lambda_r (\boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r) \overline{\times}_3 \boldsymbol{v}_k \overline{\times}_2 \boldsymbol{e}_p \overline{\times}_1 \boldsymbol{e}_q \\ &= \sum_{r=1}^{R} \lambda_r (\boldsymbol{e}_q^T \cdot \boldsymbol{a}_r)(\boldsymbol{e}_p^T \cdot \boldsymbol{b}_r)(\boldsymbol{v}_k^T \cdot \boldsymbol{c}_r) \\ &= [(\boldsymbol{e}_q^T \cdot \boldsymbol{A}) \odot (\boldsymbol{e}_p^T \cdot \boldsymbol{A}) \odot (\boldsymbol{v}_k^T \cdot \boldsymbol{C})] \cdot \boldsymbol{\lambda}, \end{aligned} \tag{11}$$

where $\odot$ is the Hadama product. The formula derivation can be found in Supplementary Material.

Estimate P^d_{pqk} by neural network

With the above mathematical derivation, P^d_{pqk} can be approximately calculated by the factor matrices of $\boldsymbol{ST}$. Thus, it is reasonable for STNN-DDI to obtain P^d_{pqk} by a neural network. The modeling of STNN-DDI can be described as follows.

Firstly, we initialize R-dimension embeddings for every chemical substructure and interaction randomly, as mentioned in section 3.4, $\boldsymbol{ST}$ can be represented by factor matrices $\boldsymbol{A}$ and $\boldsymbol{C}$, where $\boldsymbol{A} \in \mathbb{R}^{n \times R}$ is constituted by the embeddings of all chemical substructures, $\boldsymbol{C} \in \mathbb{R}^{n \times R}$ is constituted by the embeddings of all interactions. According to Eq. 11, we can get

$$\widehat{P^d_{pqk}} = [(\boldsymbol{e}_q^T \cdot \boldsymbol{A}) \odot (\boldsymbol{e}_p^T \cdot \boldsymbol{A}) \odot (\boldsymbol{v}_k^T \cdot \boldsymbol{C})] \cdot \boldsymbol{W}_\lambda + \boldsymbol{bias}, \tag{12}$$

where $\boldsymbol{W}_\lambda \in \mathbb{R}^R$ is the weight of each rank-one tensor in tensor reconstruction, $\boldsymbol{bias} \in \mathbb{R}^R$ is added for enhancing the robustness of STNN-DDI. Naturally, we can design a fully connected neural network model based on Eq. 12. Figure 5 shows the workflow of the learning process by using a neural network. It is an end-to-end learning model.

Since STNN-DDI has trainable parameters, including $\boldsymbol{A}$, $\boldsymbol{C}$ and $\boldsymbol{W}_\lambda$, we calculate the loss in all classes over all the training subjects by

$$loss = \sum_{DDI_{pqk} \in DDI_{train}} (P^d_{pqk} - \widehat{P^d_{pqk}})^2, \tag{13}$$

where P^d_{pqk} indicates DDIs if DDI_{pqk} occurs, $P^d_{pqk} = 1$, otherwise 0.

Experiment

The dataset of DDI used in this paper was taken from [33], which collected 555 drugs and their 3 576 513 pairwise interactions involving 1318 interaction types from TWOSIDES [6]. Each drug in the dataset has at least one DDI and each interaction occurs at least once. From our statistics, there are 1∼541 interactions between a pair of drugs. The original data of TWOSIDES were collected from the source of FARES [34] and MedEffect [35] respectively. More details can be found in TWOSIDES.

We use two kinds of molecular fingerprints: PubChem and MACCS, to test the performance of the proposed model. The number of substructures of MACCS fingerprint is 167, covering most of the chemical characteristics in drug discovery and virtual screening. While PubChem fingerprint contains 881 substructures, covering a wide range of different substructures and functional groups. For easy distinction, we define STNN-DDI formed by Pubchem fingerprint as STNN-DDI(P) and formed by MACCS fingerprint as STNN-DDI(M). Then, we adopt 10-fold cross-validation to estimate the performance of STNN-DDI and compare it with six state-of-the-art methods under three scenarios.

For task C1, all the DDIs were randomly divided into 10 parts of approximately equal sizes in each round, where 90% DDIs are sampled as the training set and the remaining 10% DDIs are used for testing. For task C2, all the drugs were randomly divided into 10 parts, where 90% of drugs were sampled as known drugs for training, and the remaining 10% of drugs were treated as new drugs for testing. Note that none of the interactions of new drugs attend in the training. C2 aims to deduce potential interactions between the training drugs and the testing drugs.

The division of drugs in task C3 is similar to that in task C2. Remarkably, task C3 aims to infer potential interactions among the testing drugs by the model learned from the interactions of the training drugs.

In addition, since the number of negative samples is larger than that of positive samples, the class imbalance would impact the predicting performance of the model. Therefore, we use all the positive samples and randomly select the same number of negative samples during the training. Four evaluating metrics were adopted to measure the models' performance, including AUC, AUPR, Accuracy (Acc) and Precision (Pre).

Hyper-parameter tuning

The only hyper-parameter of STNN-DDI is R, which is the number of rank-one tensors decomposed by $\boldsymbol{ST}$ and indicates the embedding dimension of the SSI space. Let R_P be the hyper-parameter of $\boldsymbol{ST}$ in STNN-DDI(P), and R_M be the hyper-parameter of $\boldsymbol{ST}$ in STNN-DDI(M). To investigate how it influences the performance of STNN-DDI, we tuned the value of R_P for the list of {10, 20, 50, 100, 200, 300, 350, 400, 450, 500}, and the value of R_M for the list of {10, 20, 50, 80, 100, 150} under three prediction tasks. The results show that with the increment of R, its value gradually approaches the rank of $\boldsymbol{ST}$, which brings the better performance of STNN-DDI. STNN-DDI achieves the best performance at $R_P = 400$ (Figure 6) and $R_M = 80$ (Figure 7) regarding all four metrics. After that, with the increase of R, meaningless information will be introduced into the tensor, and the effect of the model becomes worse. In addition, it is observed that the small

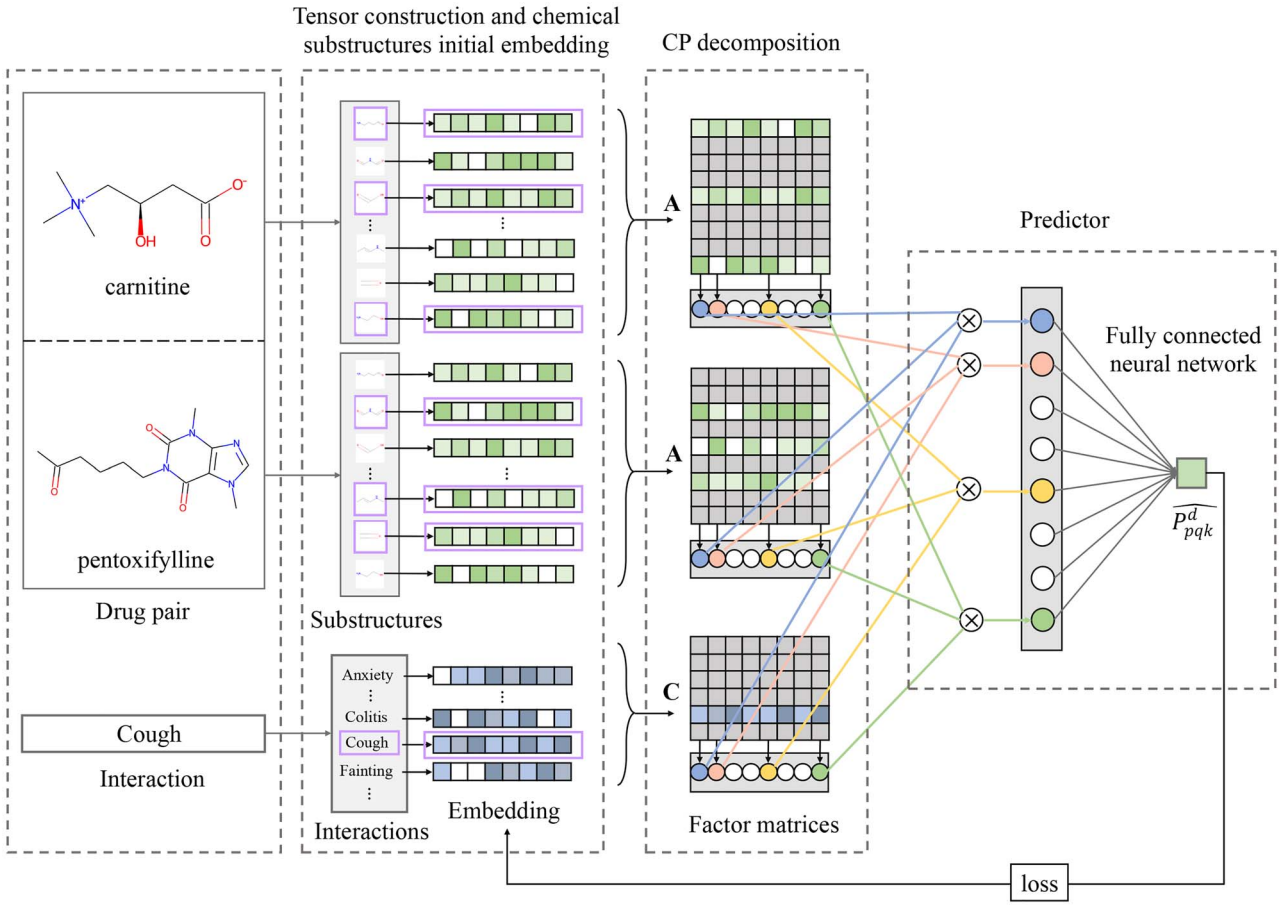


Figure 5. The end-to-end learning workflow of STNN-DDI. (1) Input triplet; (2) initialize R-dimensional embedding for all substructures and interactions randomly; (3) the embeddings of substructure and interaction constitute **A** and **C** in rows, respectively; (4) construct a fully connected neural network according to Eqs.12 and 13 to calculate P^d_{pqk}.

value of R still has an acceptable performance. Thus, STNN-DDI is a robust model for the hyper-parameter. In all the subsequent experiments, we fixed $R_P = 400$ and $R_M = 80$ to run STNN-DDI.

Baselines and comparison

To test the performance of STNN-DDI, we compared it with four state-of-the-art baselines in our dataset, which are briefly introduced as follows.

- DeepDDI [13]: DeepDDI uses the structural similarity between drugs to represent drugs and predicts DDI through deep learning.
- DDIMDL [14]: DDIMDL generates drug embeddings through multiple drug propertie similarities and trains the model through auto-encoder and neural network.
- DNN [14, 36]: It used a deep learning model to predict DDIs and Deng *et al.* [14] modified it which takes drug molecular fingerprints as input.
- MDNN [37]: MDNN designs a two-pathway framework based on drug knowledge graph and heterogeneous feature then predicts DDIs by multimodal fusion neural layer.
- GoGNN [16]: GoGNN extracts drug features in drug structure graphs as well as a DDI knowledge graph by a multi-resolution architecture and trains the model by a dual attention mechanism.
- Rescal [38]: We also implement a model based on Rescal, which is an efficient tensor decomposition model for link prediction. In this model, we organized all the DDIs into a tensor, which was decomposed by Rescal into a core tensor and drug latent representations, and predict potential interactions by reconstructing the tensor.

In addition, since neither GoGNN nor Rescal can only be applied in the inductive scenario, we did not run them in C2 and C3.

The comparison results are shown in Table 2. In a nutshell, STNN-DDI(P) significantly outperforms all the baselines across all the four metrics in task C1. As to the task of C2, STNN-DDI(M) achieves the best except for Acc. While in the case of C3, STNN-DDI(M) achieves the best in terms of AUPR and Precision, though it cannot beat the baselines in terms of AUC and Acc. Our extra investigation of the underlying reason reveals that some substructures occurring in the testing drugs do not occur in the training drugs especially in STNN-DDI(P) model. The SSI space trained without these substructures failed to characterize new drugs and resulted in the worse performance of STNN-DDI in C2 or C3, while

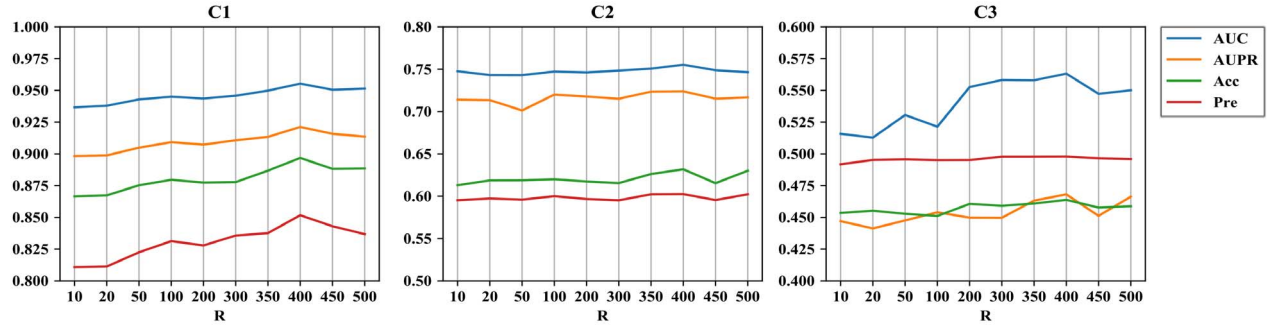


Figure 6. Experimental results of different prediction tasks under different R_P.

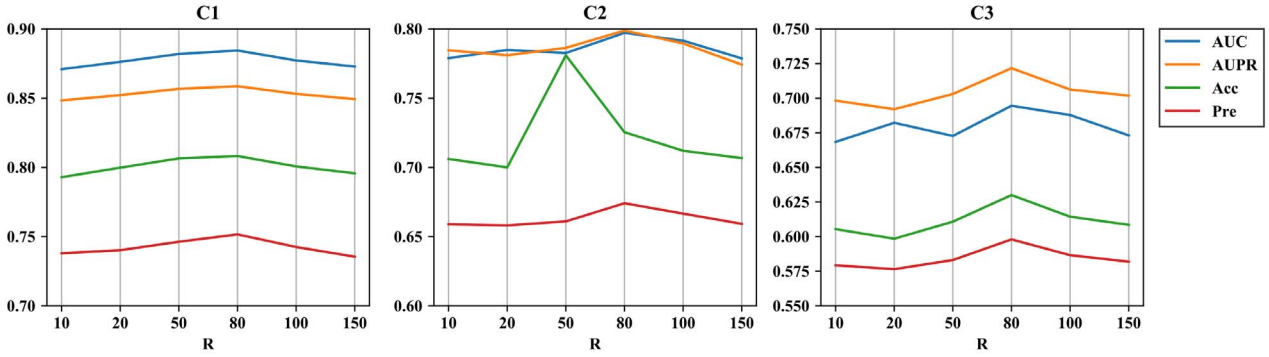


Figure 7. Experimental results of different prediction tasks under different R_M.

STNN-DDI(P) beat all the baselines in C1 due to the SSI space was trained by all the occurring substructures in the drugs. We believe that STNN-DDI would achieve a better performance in C2 and C3 if all the substructure pairs can attend its training. In summary, the comparison with baselines validates the superiority of STNN-DDI for multi-type DDI prediction in both transductive and inductive scenarios.

Interpretability illustration

We elaborated on two experiments to illustrate the explicit interpretability of STNN-DDI. The first experiment illustrated how STNN-DDI can reveal an important substructure pair across drugs regarding a DDI type of interest, while the second one illustrated how it can uncover major interaction type-specific substructure pairs in a given DDI.

Case study 1: important substructure pairs regarding single interaction type

ST can be regarded as f slices of *substructure* × *substructure* matrices, of which each corresponds to a specific type of interaction. In such a matrix, the value of an element indicates how much a pair of corresponding substructures contributes to the interaction type. As we found, it is usual that a few elements are significantly larger than others in the matrix. Thus, substructure pairs having large values are important to the interaction type. In other words, we consider that they determine or cause the interaction type of interest.

The 'temperature increased' interaction was selected as a case study. Its most important substructure pair (with the probability $P^{s}_{228,478,1} = 0.69$) is $\langle s_{228}, s_{478}\rangle$, where s_{228} is '≥1 saturated or aromatic carbon-only ring size 8' and s_{478} is 'S-C=N-[#1]' in terms of PubChem fingerprints. After searching the database, we found two drugs, 'Docetaxel' and 'Paclitaxel' containing substructure s_{228}, as well as four drugs containing substructure s_{478}, which are 'Lansoprazole', 'Omeprazole', 'Pantoprazole' and 'Rabeprazole'. By counting all the eight drug pairs between s_{228}-contained drugs and s_{478}-contained drugs and their interaction statements in the dataset, we observed that 7 drug pairs leads to the interaction that patients' body temperature increases, except the drug pair ⟨*Paclitaxel*, *Pantoprazole*⟩ (Table 3g). This finding may conclude that the risk of 'temperature increased' could be triggered if two drugs containing (s_{228}, s_{478}) are taken together. Similar conclusions were also reported in other publications [32]. Thus, the detection of important substructure pairs is helpful to uncover why a DDI type of interest occurs.

Case study 2: important substructure pair subsets regarding a drug pair

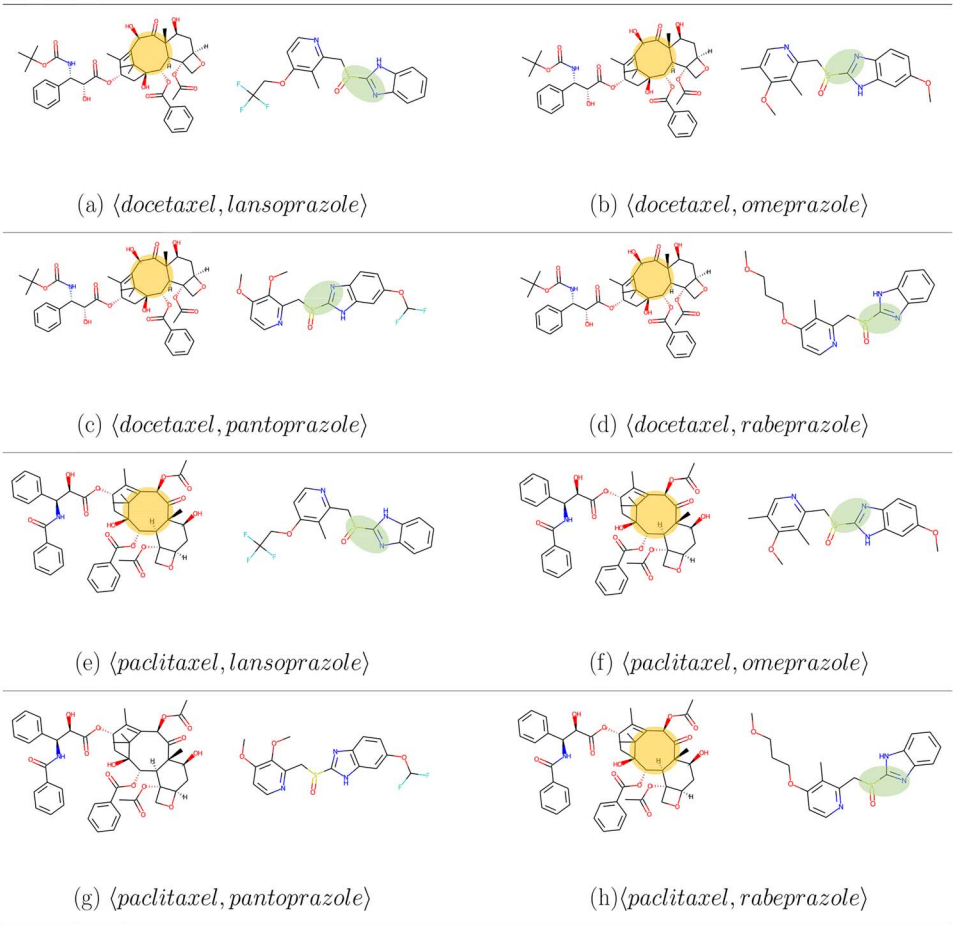
As we found, many drug pairs trigger more than one type of DDIs, referred to as multi-label DDIs. STNN-DDI can uncover how substructure pairs of two drugs contribute to multi-label DDIs. To illustrate this advantage, we carried out a visualization analysis by following [33] on the occurrence of interactions in a drug pair 'Carnitine' and 'Budesonide'. According to the interaction statements in

Table 2. The experimental results of different models. (Best is highlighted in bold)

Task	Model	AUC	AUPR	Acc	Pre
C1	DeepDDI	0.7293	0.632	0.3227	0.5102
	DDIMDL	0.806	0.7519	0.3998	0.5382
	DNN	0.9235	0.7595	0.4163	0.538
	Rescal	0.9017	0.9113	0.8828	0.829
	MDNN	0.9481	0.7896	0.5329	0.6224
	GoGNN	0.7774	0.7103	0.7076	0.6331
	STNN-DDI(M)	0.8844	0.7516	0.8586	0.8082
	STNN-DDI(P)	**0.9553**	**0.9212**	**0.8968**	**0.8517**
C2	DeepDDI	0.7546	0.1073	0.7147	0.0463
	DDIMDL	0.7622	0.1568	0.7672	0.0212
	DNN	0.7885	0.0243	0.7600	0.0499
	MDNN	0.7964	0.2376	**0.7833**	0.1493
	STNN-DDI(M)	**0.7971**	**0.7988**	0.7254	**0.6741**
	STNN-DDI(P)	0.7553	0.7239	0.6318	0.6024
C3	DeepDDI	0.7346	0.0193	0.6873	0.0173
	DDIMDL	0.7478	0.2404	0.7072	0.0203
	DNN	0.7373	0.028	0.7574	0.02
	MDNN	**0.7668**	0.0297	**0.7635**	0.0268
	STNN-DDI(M)	0.6945	**0.7216**	0.6300	**0.5980**
	STNN-DDI(P)	0.5632	0.4681	0.4638	0.4979

M represents using the MACCS fingerprint. P represents using the PubChem fingerprint.

Table 3. Visualization of the important substructure pairs regarding the 'temperature increased' interaction[a]



(a) $\langle docetaxel, lansoprazole \rangle$ (b) $\langle docetaxel, omeprazole \rangle$

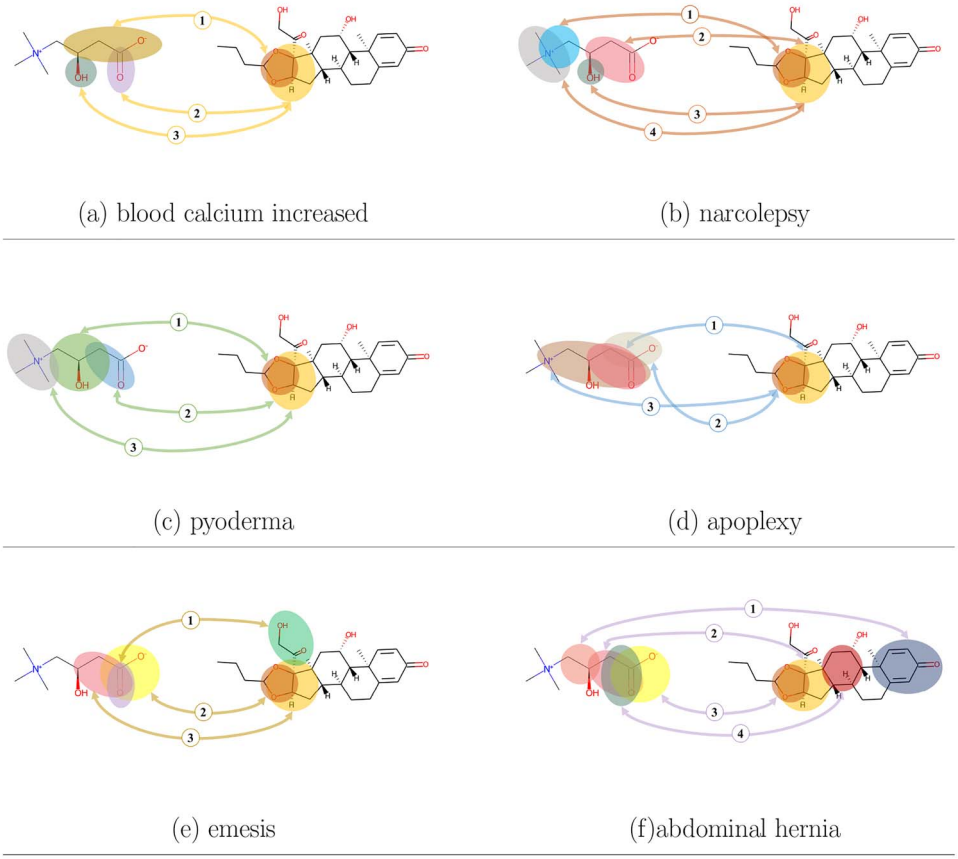
(c) $\langle docetaxel, pantoprazole \rangle$ (d) $\langle docetaxel, rabeprazole \rangle$

(e) $\langle paclitaxel, lansoprazole \rangle$ (f) $\langle paclitaxel, omeprazole \rangle$

(g) $\langle paclitaxel, pantoprazole \rangle$ (h) $\langle paclitaxel, rabeprazole \rangle$

[a] Only the 8 pairs of drugs (bottom in each panel) contain the substructure pair (s_{228}, s_{478}), and all the drug pairs can cause patient temperature increased except ⟨*Paclitaxel*, *Pantoprazole*⟩ (highlighted by semitransparent colorful ellipses).

Table 4. Multi-label DDIs between Carnitine and Budesonide[a]



[a] Six types of interactions (bottom in each panel) between 'Carnitine' (left side in each panel) and 'Budesonide' (right side in each panel) are highlighted by different colors. Important substructures are highlighted by semitransparent colorful ellipses. Curves with double arrows denote important structure pairs, where ①, ②, ③ and ④ are their indices of important structure pairs in each of the interaction types.

the dataset, these two drugs trigger 6 types of interactions, including 'blood calcium increased', 'narcolepsy', 'pyoderma', 'apoplexy', 'emesis' and 'abdominal hernia'.

After enumerating all their substructure pairs in each type of their interactions, we picked up important substructure pairs, where a set of substructures coming from 'Carnitine' and another set of substructures coming from 'Budesonide' (Table 4). All of them are called 'multiple to multiple' SSIs.

In detail, for the interaction type of 'blood calcium increased', there are 3 pairs between three important substructures 'C(~C)(~C)(~H)(~O)', 'C(-C)(=O)', 'C-C-C-O-[#1]' from 'Carnitine' and two 'OC1C(O)CCC1', '(C(=O)CO)[C@@]' from 'Budesonide', these substructure pairs form three SSIs marked by ①, ② and ③, respectively (Table 4a), which are the mainly causes of the interaction type of 'blood calcium increased' between drug 'Carnitine' and drug 'Budesonide'. Similarly, there are 4 important substructure pairs associated with 'narcolepsy', 3 associated with 'pyoderma', 3 associated with 'apoplexy', 3 associated with 'emesis' and 4 associated with 'abdominal hernia' (Table 4b–f). Especially, 4 interactions of 'abdominal hernia' involve 4 substructures coming from 'Carnitine' and another 4 substructures coming from 'Budesonide'. Moreover, the two substructure 'OC1C(O)CCC1', '(C(=O)CO)[C@@]' in 'Budesonide' attend all types of its interactions.

This illustration demonstrates that STNN-DDI can uncover interaction type-specific substructure pairs in a given DDI. The exclusive advantage of STNN-DDI, compared with other models, enables the explanation of why two drugs trigger multi-type even multi-label interactions. Its interpretability can provide valuable guidance in the development of new drugs and poly-drug therapy.

Conclusions

In this paper, we have designed a novel end-to-end learning model, STNN-DDI, for multi-type DDI prediction, based on the assumption that the interaction between two drugs is determined by the occurrence of their substructures and their DDI types are determined by the linkages between different substructure sets. STNN-DDI leverages a 3-D substructure-aware tensor and a fully connected neural network to construct a

SSI space, where drugs are embedded to discriminate what types of interactions they trigger and how likely they trigger a specific type of interaction. Due to the SSI space, STNN-DDI is not only a model of transductive DDI prediction but also a model of inductive DDI prediction. Moreover, it can handle the case that two drugs trigger one type of interaction as well as multiple types of interactions simultaneously because the DDI types are determined by the association/interaction between different substructure sets. More importantly, since predefined substructures own specific chemical meanings, it can explicitly reveal an important substructure pair across drugs regarding a DDI type of interest and uncover major interaction type-specific substructure pairs in a given DDI. To summarize, STNN-DDI contributes not only an effective predicting model of multiple-type DDIs but also an explorer to the DDI mechanism.

Keys Points

- We provides a unified form STNN-DDI for the DDI prediction in both transductive and inductive scenarios.
- STNN-DDI provides an integrated solution for the case that two drugs trigger single interaction and the case that they trigger multiple-type interactions.
- STNN-DDI has a good interpretability and it enables to both reveal an important substructure pair across drugs regarding a DDI type of interest and uncover major interaction type-specific substructure pairs in a given DDI.

Acknowledgements

The authors thank the anonymous reviewers for their constructive comments that helped us to improve our manuscript and make it clearer. We also thank the Center for High Performance Computation, Northwestern Polytechnical University for providing computation resource.

Code and data availability

Source code is freely available at https://github.com/zsy-9/STNN-DDI.

Funding

National Nature Science Foundation of China (Grant No. 61872297); Shaanxi Provincial Key Research & Development Program, China (Grand No. 2020KW-063).

References

1. Horikawa N, Sugimoto S. Comorbid disease and concomitant drug information in clinical researches. *Rinsho yakuri/Japanese J Clin Pharmacol Ther* 2019;**50**:9–16.
2. Fuseau E, Sheiner LB. Simultaneous modeling of pharmacokinetics and pharmacodynamics with a nonparametric pharmacodynamic model. *Clin Pharmacol Ther* 1984;**35**(6):733.
3. Paterno MD, Maviglia SM, Gorman PN, *et al.* Tiering drug-drug interaction alerts by severity increases compliance rates. *J Am Med Inform Assoc: JAMIA* 2009;**16**(1):40–6.
4. Giacomini KM, Krauss RM, Roden DM, *et al.* When good drugs go bad. *Nature* 2007;**446**:975–7.
5. Whitebread S, Hamon J, Bojanic D, *et al.* Keynote review: in vitro safety pharmacology profiling: an essential tool for successful drug development. *Drug Discov Today* 2005;**10**(21):1421–33.
6. Tatonetti NP, Ye P, Daneshjou R, *et al.* Data-driven prediction of drug effects and interactions. *Sci Transl Med* 2012;**4**:125ra31–1.
7. Gottlieb A, Stein GY, Oron Y, *et al.* Indi: a computational framework for inferring drug interactions and their associated recommendations. *Mol Syst Biol* 2012;**8**:592–2.
8. Rayhan F, Ahmed S, Mousavian Z, *et al.* Frnet-dti: deep convolutional neural network for drug-target interaction prediction. *Heliyon* 2020;**6**:e03444. 10.1016/j.heliyon.2020.e03444.
9. Huang K, Xiao C, Hoang T, *et al. Caster: predicting drug interactions with chemical substructure representation*, Vol. **34**, 2020a, 702–9.
10. Zhang P, Wang F, Jianying H, *et al.* Label propagation prediction of drug-drug interactions based on clinical side effects. *Sci Rep* 2015;**5**:12339.
11. Zhang W, Chen Y, Liu F, *et al.* Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data. *BMC Bioinformatics* 2017;**18**(1):18.
12. Xu Chu, Yang Lin, Yasha Wang, Leye Wang, Jiangtao Wang, and Jingyue Gao. Mlrda: a multi-task semi-supervised learning framework for drug-drug interaction prediction. In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019*, Macao, China: Morgan Kaufmann, San Francisco, 2019, 4518–4524.
13. Ryu JY, Kim HU, Sang YL. Deep learning improves prediction of drug-drug and drug-food interactions. *Proc Natl Acad Sci U S A* 2018;**115**(18):E4304.
14. Deng Y, Xinran X, Qiu Y, *et al.* A multimodal deep learning framework for predicting drug-drug interaction events. *Bioinformatics* 2020;**36**(15):4316–22. 10.1093/bioinformatics/btaa501.
15. Ferdousi R, Safdari R, Omidi Y. Computational prediction of drug-drug interactions based on drugs functional similarities. *J Biomed Inform* 2017;**70**:54–64.
16. Hanchen Wang, Defu Lian, Ying Zhang, *et al.* Gognn: graph of graphs neural network for predicting structured entity interactions. *Yokohama, Japan, July 11–17, 2020*. Morgan Kaufmann, San Francisco. 2020 pages 1317–1323.
17. Huang K, Xiao C, Glass L, *et al.* Skipgnn: predicting molecular interactions with skip-graph networks. *Sci Rep* 2020b;**10**:1–16.
18. Takeda T, Hao M, Cheng T, *et al.* Predicting drug-drug interactions through drug structural similarities and interaction networks incorporating pharmacokinetics and pharmacodynamics knowledge. *J Chem* 2017;**9**(1):1–9.
19. Wang Y, Chen R, Ghosh J, *et al.* (eds). Rubik: knowledge guided tensor factorization and completion for health data analytics. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Sydney, NSW, Australia: ACM, New York, NY, 2015, 1265–74.
20. Acar E, Kolda TG, Dunlavy DM. All-at-once optimization for coupled matrix and tensor factorizations. In MLG'11: *Proceedings of Mining and Learning with Graphs*. San Diego, USA: ACM, New York, NY, 2011.
21. Narita A, Hayashi K, Tomioka R, *et al.* Tensor factorization using auxiliary information. *Data Min Knowl Discov* 2012;**25**:298–324.

22. Hui Y, Mao K-T, Shi J-Y, *et al*. Predicting and understanding comprehensive drug-drug interactions via semi-nonnegative matrix factorization. *BMC Syst Biol* 2018;**12**(1):101–10.
23. Rohani N, Eslahchi C, Katanforoush A. Iscmf: Integrated similarity-constrained matrix factorization for drug-drug interaction prediction. *Netw Model Anal Health Inform Bioinform* 2020;**9**:1–8.
24. Sobhanam H, Mariappan A. Addressing cold start problem in recommender systems using association rules and clustering technique. In: *2013 International Conference on Computer Communication and Informatics*. Craiova, Romania: IEEE Piscataway, NJ, 2013, 1–5.
25. Zhang W, Jing K, Huang F, *et al*. Sflln: a sparse feature learning ensemble method with linear neighborhood regularization for predicting drug-drug interactions. *Inform Sci* 2019;**497**: 189–201.
26. Harrold MW, Zavod RM. Basic concepts in medicinal chemistry. *Drug Dev Ind Pharm* 2013;**40**:988–8.
27. Arachchige PCM, Chen Y-PP. Medfused: a framework to discover the relationships between drug chemical functional group impacts and side effects. *Comput Biol Med* 2021;**133**:104361.
28. Fu T, Xiao C, Sun J. Core: automatic molecule optimization using copy & refine strategy. *32nd AAAI Conf Artif Intell AAAI* 2020;**34**(1):638–45.
29. Kolda TG, Bader BW. Tensor decompositions and applications. *SIAM Rev* 2009;**51**:455–500.
30. Hitchcock FL. The expression of a tensor or a polyadic as a sum of products. *J Math Phys* **6**:164–89.
31. Sidiropoulos N, De Lathauwer L, Xiao F, *et al*. Tensor decomposition for signal processing and machine learning. *IEEE Trans Signal Process* 2017;**65**:3551–82.
32. Rabanser S, Shchur O, Günnemann S. Introduction to tensor decompositions and their applications in machine learning. *CoRR*, abs/1711.10781, 2017 . http://arxiv.org/abs/1711.10781.
33. Zhu J, Liu Y, Wen C. Mtma: multi-task multi-attribute learning for the prediction of adverse drug-drug interaction. *Knowl Based Syst* 2020;**199**:105978.
34. Harpaz R, DuMouchel W, LePendu P, *et al*. Performance of pharmacovigilance signal-detection algorithms for the fda adverse event reporting system. *Clin Pharmacol Ther* 2013;**93**(6): 539–46.
35. Health Products Government of Canada, Health Canada. Therapeutic Effectiveness Food Branch, Marketed Health Products Directorate, and Policy Bureau. Medeffect canada - advisories, warnings and recalls - health canada. *Reporting Adverse Reactions*.
36. Lee G, Park C, Ahn J. Novel deep learning model for more accurate prediction of drug-drug interaction effects. *BMC bioinformatics* 2019;**20**(1):1–8.
37. Lyu T, Gao J, Tian L, *et al*. Mdnn: a multimodal deep neural network for predicting drug-drug interaction events. *IJCAI 2021*. Virtual Event / Montreal, Canada: Morgan Kaufmann, San Francisco, 2021, 3536–3542.
38. Nickel M, Tresp V, Kriegel H-P. A three-way model for collective learning on multi-relational data. *ICML 2011*, Bellevue, Washington, USA: Omnipress, New York, NY, 2011, 809–816.