Future Generation Computer Systems 161 (2024) 26–34

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

CTF-DDI: Constrained tensor factorization for drug–drug interactions prediction

Guosheng Han [a,c], Lingzhi Peng [a,c], Aocheng Ding [b], Yan Zhang [a,c], Xuan Lin [b,c,*]

[a] School of Mathematics and Computational Science, Xiangtan University, XiangTan, 411105, Hunan, China
[b] School of Computer Science, Xiangtan University, XiangTan, 411105, Hunan, China
[c] Key Laboratory of Intelligent Computing and Information Processing of Ministry of Education and Hunan Key Laboratory for Computation and Simulation in Science and Engineering, Xiangtan University, XiangTan, 411105, Hunan, China

ARTICLE INFO

Keywords:
Tensor factorization
Similarity fusion
Drug–drug interactions prediction

ABSTRACT

Computational approaches for predicting drug–drug interactions (DDI) can significantly facilitate combination therapy and drug discovery. Existing similarity-based methods often overlook simple yet valuable structural information or ignore multiple relationships from biological entities (e.g., target proteins and enzymes). Meanwhile, matrix factorization-based methods can alleviate the inherent sparsity issues in DDI data. However, this line of work usually only considers the original association information of DDI pairs. To address these issues, we proposed a novel tensor factorization strategy with effective constraint terms (CTF-DDI) for potential DDI prediction. Specifically, we first obtained drug features by constructing specific similarity matrices based on drug structure and drug-related biological associations. Then, a novel constrained tensor factorization (CTF) module was designed to further reconstruct drug similarity by introducing Hessian and $L_{2,1}$ regularization as constraints. Finally, we trained a deep neural network to extract nonlinear features for DDI prediction. Experimental results on two benchmark datasets demonstrated that the proposed CTF-DDI model outperforms classical tensor factorization and deep learning models. Furthermore, ablation and case studies validated the performance of CTF-DDI in DDI prediction. The source code of CTF-DDI is available at https://github.com/angelfacedac/CTF_DDI.

1. Introduction

Drug–drug interactions (DDI) refer to the alterations in pharmacological effects caused by the simultaneous administration of two or more drugs [1]. The combination therapy of multiple drugs provides promising prospects for disease treatment, but it also increases the risk of adverse effects from drug combinations [2]. For example, *methotrexate* and *nonsteroidal anti-inflammatory* drugs are commonly used to treat rheumatoid arthritis. However, these combinations are prone to delayed elimination of *methotrexate*, leading to bone marrow suppression and gastrointestinal toxicity [3]. The potential DDI prediction presents one of the most significant challenges in drug discovery, requiring considerable time and resources for in-vitro and clinical experiments.

Many computational methods have been developed for alleviating this problem in terms of feature extraction and representation. Conventionally, popular computational methods based on structural similarity are used to construct the feature matrix of drugs. This is followed by feature representation through effective encoding methods, such as non-negative matrix factorization (NMF), which can effectively obtain non-negative downscaling matrices for high-dimensional data and is commonly used in image analysis and processing [4,5]. Recent years have witnessed the success of machine learning and deep learning-based approaches proposed to address different issues related to DDI prediction [6–8]. For these computational methods, representing drug features simply yet effectively is one of the key steps towards obtaining satisfactory performance. Existing methods for representing drug features can be classified into two main categories.

(1) Similarity-based methods: This line of methods relies on the assumption that drugs with similar structure tend to interact with the same types of drugs [9]. DDI prediction can be performed based on drug structural similarity features such as chemical or ligand structure and side effects. For example, Ferdousi et al. [10] used a similarity measure based on 12 binary vectors and calculated drug pair similarity using the Rus-Rao method, validating that higher structural similarity among drugs correlates with a higher likelihood of DDI. However, these methods are solely based on drug structure similarity, often ignoring

* Corresponding author.
E-mail address: jack_lin@xtu.edu.cn (X. Lin).

https://doi.org/10.1016/j.future.2024.06.060
Received 17 May 2024; Received in revised form 20 June 2024; Accepted 28 June 2024
Available online 3 July 2024

information in drug-related biological relationships and their interactions. To enhance the accuracy of DDI prediction, researchers have integrated similarities in terms of drug-related biological relationships. For instance, Deng et al. [11] proposed DDIMDL, a multi-modal deep learning framework designed to incorporate four distinct drug features — chemical substructures, targets, enzymes, and pathways — sourced from various datasets. Subsequently, Lin et al. [12] also integrated multiple sources of drug-related information to predict potential DDIs. Unlike most methods that utilize only first-order similarity, Zhang et al. [13] proposed an integrated label propagation framework that takes into account higher-order interaction similarities for predicting potential DDI. In addition, neural network-based models have been developed to extract the similarity of drug nodes in an interaction network. For example, Deepwalk [14] and Node2vec [15] obtained node embedding of drugs to predict DDI by using random walking, while LINE [16], SDNE [17], and GAE [18] considered different encoding methods to capture the nonlinear features for DDI prediction. Although similarity-based approaches have demonstrated notable prediction performance, only a few methods take into consideration both aspects of drug structure similarity and drug-related biological relationships. Additionally, fully leveraging drug similarity remains a challenging task.

(2) Matrix factorization-based methods: In recent years, matrix factorization has shown satisfactory results in various bioinformatics tasks [19–21]. DDI prediction using matrix factorization methods typically involves decomposing the DDI matrix into several matrices. From these matrices, potential features are extracted, and the matrix is reconstructed to identify new DDI. For example, Vilar et al. [22] calculated the interaction probability matrix to predict new DDI by multiplying the DDI matrix and the IPF matrix, while Zhang et al. [23] introduced multi-relationship matrix factorization of drug properties into the matrix factorization to predict potential DDI. Traditional matrix factorization has many variants, such as singular value decomposition (SVD) and graph factorization (GF) [24]. Subsequently, GraRep [25] considered the higher-order proximity of networks and devised k-step transfer probability matrices for factorization. HOPE [26] used some well-known network similarity metrics. However, matrix factorization struggles with handling higher-dimensional data. In contrast, tensor factorization serves as a higher-order generalization of matrix factorization. It decomposes a tensor into the product of several smaller tensors to obtain its approximation. Therefore, tensor factorization not only inherits the characteristics of matrix factorization in handling data but also addresses the challenges of filling in sparse datasets and reducing the complexity of high-dimensional data. The canonical polyadic (CP) factorization model represents a standard tensor factorization method that solely considers the original drug association information for DDI prediction. Subsequently, Narita et al. [27] proposed the TFAI model, which introduced graph Laplacian regularization. Similarly, Huang et al. [28] proposed the TDRC model by incorporating similarity information as a constraint. With the remarkable successes of deep learning techniques across various fields like image classification [29,30] and natural language processing [31], deep neural networks (DNN) have emerged as powerful tools applicable to biomedical prediction, offering robust capabilities for learning and characterizing the nonlinear features of drugs. For example, to achieve overall high prediction accuracy and good generalization, Zhang et al. [32] proposed using weighted average and classifier ensemble models to integrate chemical, biological, phenotypic, and network data for predicting DDI. Moreover, considering the different forms of features in deep learning techniques, Ryu et al. [33] introduced the DeepDDI model using structural similarity features. Perozzi et al. [14] proposed the DeepWalk model, which uses drug pairs as nodes in an isomorphic graph for DDI prediction. Chen et al. [34] integrated features from drug molecular structures and knowledge graphs into the MUFFIN model. These matrix factorization-based methods demonstrate good results in DDI prediction and partially alleviate the limitations of similarity-based methods. However, they still do not fully utilize drug similarity information and may suffer from overfitting.

In summary, existing similarity-based methods often neglect the simultaneous construction of similarity matrices from both drug structure and drug-related biological relationships, failing to fully harness comprehensive drug similarity information. Moreover, matrix factorization-based approaches are prone to model overfitting due to their inadequate utilization of drug similarity information and other constraints. In this paper, we propose a novel constrained tensor factorization method, named CTF-DDI, for addressing drug–drug interactions prediction problem. Our method leverages diverse similarity calculation methods to extract information from both drug structure and drug-related biological relationships. These similarities are integrated through weighted averaging to construct comprehensive drug similarity matrices. Then, we use the CP factorization model, incorporating drug similarity information alongside Hessian regularization and $L_{2,1}$ regularization constraints to enhance model robustness. The resulting feature matrices are fed into a DNN for DDI prediction. This method for the first time learns the structural similarity and drug-related biological relationships in an end-to-end manner. Experimental results on two benchmark datasets demonstrate that CTF-DDI outperforms existing similarity-based and tensor factorization-based baseline methods, showcasing superior performance and better generalization ability.

2. Methods

In this section, we first introduce the flowchart of the proposed model called CTF-DDI (Section 2.1). Then, we further describe the two main modules of CTF-DDI model, namely similarity fusion module (Section 2.2) and constrained tensor factorization (CTF) module (Section 2.3). Finally, we present the details of DDI task prediction, called DNN module (Section 2.4).

2.1. Flowchart of CTF-DDI

Fig. 1 illustrates the flowchart of CTF-DDI, which can be divided into main three parts. Specifically, we first computed and integrated distinct drug similarities sourced from DS1 and DS3 datasets, yielding two similarity matrices pertaining to drug structure and drug-related biological relationships. These matrices, combined with Hessian and $L_{2,1}$ regularization, which serve as constraints in the traditional CP factorization model. Next, we decomposed the tensor $\mathcal{X}$ of known DDI types to derive the feature matrices $\boldsymbol{C}$, $\boldsymbol{P}$, $\boldsymbol{F}$ utilizing the CTF approach, with the column vector $(\boldsymbol{c}_i\ \boldsymbol{p}_j\ \boldsymbol{f}_k)'$ serving as input to the deep neural networks (DNN) model. Finally, we obtained the reconstructed tensor $\mathcal{X}'$, representing the drug–drug interactions types and their respective probabilities.

2.2. Similarity fusion module

In this paper, $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ represent similarity matrices obtained by calculating different drug similarities in terms of both drug structure and drug-related biological relationships, where $\boldsymbol{S} \in R^{m\times m}$ and m denotes the number of drugs. For the DS3 dataset, seven similarity matrices are calculated. Drug structure similarity is based on chemical similarity, ligand-based similarity, side-effects similarity, and similarity based on the anatomical, therapeutic, and chemical (ATC) classification system. Drug-related biological similarity is constructed based on gene ontology (GO) semantic similarity, sequence similarity, and distance on the protein-protein interactions (PPI) network. The specific computational process for obtaining drug similarity matrices is outlined below.

(1) Chemical similarity: The standard Simplified Molecular Input Line Entry System (SMILES) representations of drug molecules are obtained from the DrugBank. Subsequently, hash fingerprints are generated using the Chemical Development Kit (CDK) with default parameters. The similarity between drugs is then determined by calculating

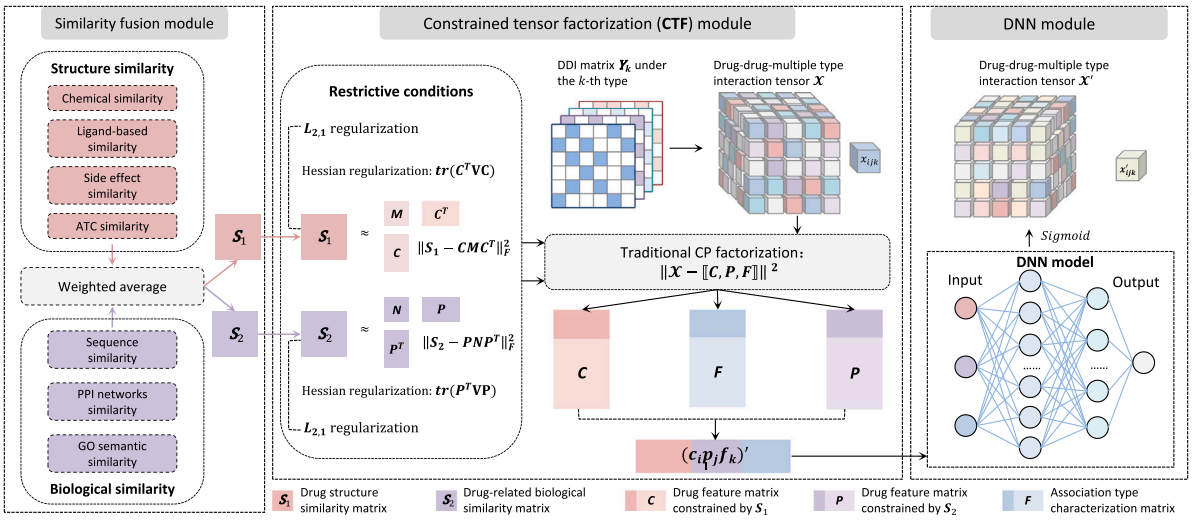


Fig. 1. Flowchart of the CTF-DDI model.

a two-dimensional Tanimoto score based on the hash fingerprints, equivalent to the Jaccard score of the fingerprints . In essence, each fingerprint is treated as a specific set of elements, and the similarity between two drugs is expressed as the magnitude of their intersection.

(2) Ligand-based similarity: The Similarity Ensemble Approach (SEA) correlates protein receptors by evaluating the chemical 2D similarity of ligand sets that regulate their function. Utilizing the standard SMILES representations of drugs, the SEA search tool compares them against a compendium of ligand sets and computes the E-values of these sets. Drugs are queried against two ligand databases provided in the tool (MDL Drug Data Report and WOMBAT), and drug fingerprints are calculated using two different methods (Scitegic ECFP4 and Daylight) to generate a list of four similar ligand sets. After harmonizing the four lists and screening the drug-ligand pairs using E-values greater than 10^{-5}, a compilation of related protein receptor families for each drug is obtained. Finally, the Jaccard score between receptor families is computed to determine the similarity between drugs.

(3) Side-effect similarity: The side effect of drugs is sourced from SIDER database, employing text mining techniques to extract associations between drugs and their side effects from package inserts. To enhance the comprehensiveness of the side effect profile, predictions were made for drugs absent in SIDER based on their chemical properties. Subsequently, the similarity between drugs was established by computing Jaccard scores between the known side effects, or the first 13 predicted side effects in the case of drugs lacking known data. However, side effects related to renal failure or renal insufficiency, renal tubular acidosis, and papillary necrosis were deliberately omitted from consideration to mitigate bias towards side effects directly linked to drug interactions.

(4) ATC-related similarity: ATC-related data was obtained from DrugBank database, and similarities among ATC classifications were determined using a semantic similarity algorithm . The semantic similarity algorithm establishes a relationship between the probability $p(x)$ and all nodes (i.e., ATC levels) x within the ATC hierarchy. This is accomplished by tallying the number of levels below x and subsequently computing the maximum ATC level c shared by all their common ancestors, expressed as $-logp(c)$.

(5) Sequence similarity: The similarity between drugs is computed using the Smith-Waterman sequence comparison score between drug-target pairs. Following the previous normalization method, the Smith-Waterman score is derived by computing the geometric mean of the scores obtained from comparing each sequence with itself.

(6) PPI network distance similarity: Distances between each drug-target pair are computed based on the shortest paths in the PPI network. And these calculated shortest path distances are transformed into corresponding similarity values using the formula outlined below.

$$S(P, P') = Ae^{-D(p,p')} \tag{1}$$

where $S(P, P')$ represents the computed similarity value between two proteins, and $D(p, p')$ denotes the shortest path between these proteins in the PPI network. The parameter A and self-similarity are set to 0.9 and 1, respectively.

(7) GO semantic similarity: We calculated GO semantic similarities between drug-target pairs using the csbl.goR package, selecting the option to use all three gene ontologies.

For the DS1 dataset, chemical, side effect, and offside effect similarities were selected to compute the structural similarity matrix $\boldsymbol{S}_1$. And sequence, transporter, enzyme, pathway, and indication similarities were selected to compute the drug-related biological similarity matrix $\boldsymbol{S}_2$.

2.3. Constrained tensor factorization (CTF) module

We first introduce one of the most commonly used factorization methods, CP factorization, to reconstruct the drug–drug–type tensor. Next, we describe our CTF method that incorporates drug similarity and other constraints into the CP model. Subsequently, we develop an efficient optimization method for solving the CTF objective function.

(1) CP factorization: For one dataset, given the drug–drug–type tensor $\mathcal{X}$, the CP factorization model can be represented as the following optimization problem.

$$\min_{\boldsymbol{C}', \boldsymbol{P}', \boldsymbol{F}} \left\| \mathcal{X} - [[\boldsymbol{C}', \boldsymbol{P}', \boldsymbol{F}]] \right\|^2 \tag{2}$$

where $\|\cdot\|$ is the norm of a tensor. $\boldsymbol{C}', \boldsymbol{P}' \in R^{m \times r}$ are two drug feature matrices, where $\boldsymbol{C}'$ (or $\boldsymbol{P}'$) is not constrained by $\boldsymbol{S}_1$ (or $\boldsymbol{S}_2$). $\boldsymbol{F} \in R^{t \times r}$ represents feature matrix of association type. $[[\boldsymbol{C}', \boldsymbol{P}', \boldsymbol{F}]]$ is the reconstructed tensor and r is the rank.

(2) Objective function of CTF: The traditional CP model solely relies on original association information, failing to fully exploit drug similarity data. Therefore, we proposed the CTF, which integrates drug similarity into the CP to maximize its utilization and incorporates Hessian and $L_{2,1}$ regularization terms to constrain the model. On one hand, compared to traditional Laplace regularization, Hessian regularization better aligns with the data characteristic, enhancing performance in preventing overfitting after data dimensionality reduction. On the other hand, unlike the sparsity requirement of the L_1 norm, the $L_{2,1}$ norm also regulates the sparsity of rows and not only combats overfitting but

also stabilizes the optimization process. The objective function of CTF is defined as follows.

$$\begin{aligned}\mathcal{J}(\boldsymbol{C},\boldsymbol{P},\boldsymbol{F}) = &\min_{\boldsymbol{C},\boldsymbol{P},\boldsymbol{F}} \|\mathcal{X} - [[\boldsymbol{C},\boldsymbol{P},\boldsymbol{F}]]\|^2 + \frac{\mu}{2}\left\|\boldsymbol{S}_1 - \boldsymbol{C}\boldsymbol{M}\boldsymbol{C}^T\right\|_F^2 \\ &+ \frac{\eta}{2}\left\|\boldsymbol{S}_2 - \boldsymbol{P}\boldsymbol{N}\boldsymbol{P}^T\right\|_F^2 + \frac{\alpha}{2}\operatorname{tr}(\boldsymbol{C}^T\boldsymbol{U}\boldsymbol{C}) + \frac{\beta}{2}\operatorname{tr}(\boldsymbol{P}^T\boldsymbol{V}\boldsymbol{P}) \\ &+ \frac{\lambda}{2}\left(\|\boldsymbol{C}\|_{2,1}^2 + \|\boldsymbol{P}\|_{2,1}^2 + \|\boldsymbol{F}\|_{2,1}^2 + \|\boldsymbol{M}\|_F^2 + \|\boldsymbol{N}\|_F^2\right)\end{aligned} \tag{3}$$

where $\|\cdot\|_F$ is the Frobenius norm of a matrix. $\boldsymbol{C}$ and $\boldsymbol{P}$ are drug feature matrices constrained by $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$. $\boldsymbol{M}$ and $\boldsymbol{N}$ represent the projection matrices, and a real-valued function $f(x,y) = x\boldsymbol{M}y^T$, with x and y representing the row vectors of $\boldsymbol{C}$ (or $\boldsymbol{P}$), is utilized to approximate $\boldsymbol{S}_1$ (or $\boldsymbol{S}_2$). Besides, μ and η control the contribution of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$, respectively. $tr(\cdot)$ corresponds to the Hessian regularization term. α and β represent the parameters of the Hessian regularization term. λ is the regularization coefficient.

(3) Optimization method: We employed an alternately updating rule as optimization method for the objective optimization. To facilitate the solution of $\boldsymbol{C},\boldsymbol{P},\boldsymbol{F}$, we first expressed the objective function in terms of the augmented Lagrangian.

$$\begin{aligned}\mathcal{L} = &\min_{\boldsymbol{C},\boldsymbol{P},\boldsymbol{F}} \|\mathcal{X} - [[\boldsymbol{C},\boldsymbol{P},\boldsymbol{F}]]\|^2 + \frac{\mu}{2}\left\|\boldsymbol{S}_1 - \boldsymbol{C}\boldsymbol{M}\boldsymbol{C}^T\right\|_F^2 + \frac{\eta}{2}\left\|\boldsymbol{S}_2 - \boldsymbol{P}\boldsymbol{N}\boldsymbol{P}^T\right\|_F^2 \\ &+ \frac{\alpha}{2}\operatorname{tr}(\boldsymbol{Z}_1^T\boldsymbol{U}\boldsymbol{Z}_1) + \frac{\beta}{2}\operatorname{tr}(\boldsymbol{Z}_2^T\boldsymbol{V}\boldsymbol{Z}_2) + \langle\boldsymbol{Y}_1, \boldsymbol{Z}_1 - \boldsymbol{C}\rangle + \langle\boldsymbol{Y}_2, \boldsymbol{Z}_2 - \boldsymbol{P}\rangle \\ &+ \frac{\theta}{2}(\|\boldsymbol{Z}_1 - \boldsymbol{C}\|_F^2 + \|\boldsymbol{Z}_2 - \boldsymbol{P}\|_F^2) \\ &+ \frac{\lambda}{2}\left(\|\boldsymbol{C}\|_{2,1}^2 + \|\boldsymbol{P}\|_{2,1}^2 + \|\boldsymbol{F}\|_{2,1}^2 + \|\boldsymbol{M}\|_F^2 + \|\boldsymbol{N}\|_F^2\right)\end{aligned} \tag{4}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product of the two. $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ are auxiliary variables. $\boldsymbol{U}$ and $\boldsymbol{V}$ are the Laplace matrices computed from the matrices $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$, respectively. $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are the Lagrange multiplier. θ is the penalty parameter. $\boldsymbol{I}$ is an identity matrix of size (rxr).

Update the feature matrix $\boldsymbol{F}$. When the other variables are fixed, the problem reduces to solving the following function.

$$\min_{\boldsymbol{F}} \left\|\mathcal{X}_{(3)} - \boldsymbol{F}(\boldsymbol{C}\otimes\boldsymbol{P})^T\right\|_F^2 + \frac{\lambda}{2}\|\boldsymbol{F}\|_{2,1}^2 \tag{5}$$

where $\mathcal{X}_{(3)}$ is the mode-3 matricization of tensor $\mathcal{X}$, and $\otimes$ denotes the Khatri–Rao product. The closed-form solution for $\boldsymbol{F}$ is as follows.

$$\boldsymbol{F} = \mathcal{X}_{(3)}(\boldsymbol{C}\otimes\boldsymbol{P})((\boldsymbol{C}\otimes\boldsymbol{P})^T(\boldsymbol{C}\otimes\boldsymbol{P}) + \frac{\lambda}{2}\boldsymbol{I})^{-1} \tag{6}$$

Update the feature matrix $\boldsymbol{C}$. Similarly, the matrix $\boldsymbol{C}$, constrained by $\boldsymbol{S}_1$, can be obtained through solving the function.

$$\begin{aligned}&\min_{\boldsymbol{C}} \left\|\mathcal{X}_{(1)} - (\boldsymbol{P}\otimes\boldsymbol{F})^T\right\|^2 + \langle\boldsymbol{Y}_1, \boldsymbol{Z}_1 - \boldsymbol{C}\rangle + \frac{\theta}{2}\|\boldsymbol{Z}_1 - \boldsymbol{C}\|_F^2 \\ &+ \frac{\mu}{2}\left\|\boldsymbol{S}_1 - \boldsymbol{C}\boldsymbol{M}\boldsymbol{C}^T\right\|_F^2 + \frac{\lambda}{2}\|\boldsymbol{C}\|_{2,1}^2\end{aligned} \tag{7}$$

where $\mathcal{X}_{(1)}$ is the mode-1 matricization of tensor $\mathcal{X}$. Inspired by [35], the problem can be further transformed as follows.

$$\begin{aligned}&\min_{\boldsymbol{C}}\left\|\mathcal{X}_{(1)} - \boldsymbol{C}(\boldsymbol{P}\otimes\boldsymbol{F})^T\right\|^2 + \frac{\theta}{2}\left\|\boldsymbol{Z}_1 - \boldsymbol{C} + \frac{\boldsymbol{Y}_1}{\theta}\right\|_F^2 \\ &+ \frac{\mu}{2}\left\|\boldsymbol{S}_1 - \boldsymbol{C}\boldsymbol{M}\boldsymbol{J}_1^T\right\|_F^2 + tr(\boldsymbol{R}_1^T(\boldsymbol{C} - \boldsymbol{J}_1)) \\ &+ \frac{\rho_1}{2}\|\boldsymbol{C} - \boldsymbol{J}_1\|_F^2 + \frac{\lambda}{2}\|\boldsymbol{C}\|_{2,1}^2\end{aligned} \tag{8}$$

where $\rho_1 > 0$ is referred to as the penalty parameter, and $\boldsymbol{J}_1$ is an auxiliary variable. The solution is then obtained by taking the partial derivative.

$$\boldsymbol{J}_1 = (\mu\boldsymbol{S}_1\boldsymbol{C}\boldsymbol{M} + \rho_1\boldsymbol{C} + \boldsymbol{R}_1)(\mu(\boldsymbol{C}\boldsymbol{M})^T(\boldsymbol{C}\boldsymbol{M}) + \rho_1\boldsymbol{I})^{-1} \tag{9}$$

$$\boldsymbol{R}_1 = \boldsymbol{R}_1 + \rho_1(\boldsymbol{C} - \boldsymbol{J}_1) \tag{10}$$

The feature matrix $\boldsymbol{C}$ is obtained as follows.

$$\begin{aligned}&\boldsymbol{C} = (\mathcal{X}_{(1)}(\boldsymbol{P}\otimes\boldsymbol{F}) + \mu\boldsymbol{S}_1\boldsymbol{J}_1\boldsymbol{M}^T + \rho_1\boldsymbol{J}_1 - \boldsymbol{R}_1 + \theta\boldsymbol{Z}_1 + \boldsymbol{Y}_1) \\ &((\boldsymbol{P}\otimes\boldsymbol{F})^T(\boldsymbol{P}\otimes\boldsymbol{F}) + \mu\boldsymbol{M}\boldsymbol{J}_1^T\boldsymbol{J}_1\boldsymbol{M}^T + \rho_1\boldsymbol{I} + \lambda\boldsymbol{I} + \lambda\boldsymbol{D}\boldsymbol{I} + \theta\boldsymbol{I})^{-1}\end{aligned} \tag{11}$$

where $\boldsymbol{R}_1$ is the Lagrange multiplier. $\boldsymbol{D}$ is a diagonal matrix with the ith diagonal element $\boldsymbol{D}_{ii} = 1/2\left\|(\boldsymbol{C})^i\right\|_2$, i.e., half of the L_2 norm of the ith column vector in matrix $\boldsymbol{C}$.

The update rules of $\boldsymbol{Z}_1$ and $\boldsymbol{Y}_1$ are as follows.

$$\begin{cases}\boldsymbol{Z}_1(t+1) = (\theta\boldsymbol{I} + \alpha\boldsymbol{U})^{-1}(\theta\boldsymbol{C} - \boldsymbol{Y}_1(t)) \\ \boldsymbol{Y}_1(t+1) = \boldsymbol{Y}_1(t) + \theta(\boldsymbol{Z}_1(t+1) - \boldsymbol{C})\end{cases} \tag{12}$$

Update the feature matrix $\boldsymbol{P}$. The process shares a similar optimization structure to Eq. (7). The update process is as follows.

$$\boldsymbol{J}_2 = (\eta\boldsymbol{S}_2\boldsymbol{P}\boldsymbol{N} + \rho_2\boldsymbol{P} + \boldsymbol{R}_2)(\eta(\boldsymbol{P}\boldsymbol{N})^T(\boldsymbol{P}\boldsymbol{N}) + \rho_2\boldsymbol{I})^{-1} \tag{13}$$

$$\boldsymbol{R}_2 = \boldsymbol{R}_2 + \rho_2(\boldsymbol{P} - \boldsymbol{J}_2) \tag{14}$$

$$\begin{aligned}&\boldsymbol{P} = (\mathcal{X}_{(2)}(\boldsymbol{C}\otimes\boldsymbol{F}) + \eta\boldsymbol{S}_2\boldsymbol{J}_2\boldsymbol{N}^T + \rho_2\boldsymbol{J}_2 - \boldsymbol{R}_2 + \theta\boldsymbol{Z}_2 + \boldsymbol{Y}_2) \\ &((\boldsymbol{C}\otimes\boldsymbol{F})^T(\boldsymbol{C}\otimes\boldsymbol{F}) + \eta\boldsymbol{N}\boldsymbol{J}_2^T\boldsymbol{J}_2\boldsymbol{N}^T + \rho_2\boldsymbol{I} + \lambda\boldsymbol{I} + \lambda\boldsymbol{Q}\boldsymbol{I} + \theta\boldsymbol{I})^{-1}\end{aligned} \tag{15}$$

where $\boldsymbol{R}_2$ is the Lagrange multiplier, and ρ_2 is the penalty parameter. $\boldsymbol{Q}$ is a diagonal matrix with the ith diagonal element $\boldsymbol{Q}_{ii} = 1/2\left\|(\boldsymbol{P})^i\right\|_2$, i.e., half of the L_2 norm of the ith column vector in matrix $\boldsymbol{P}$.

The update rules of $\boldsymbol{Z}_2$ and $\boldsymbol{Y}_2$ are as follows.

$$\begin{cases}\boldsymbol{Z}_2(t+1) = (\theta\boldsymbol{I} + \beta\boldsymbol{V})^{-1}(\theta\boldsymbol{P} - \boldsymbol{Y}_2(t)) \\ \boldsymbol{Y}_2(t+1) = \boldsymbol{Y}_2(t) + \theta(\boldsymbol{Z}_2(t+1) - \boldsymbol{P})\end{cases} \tag{16}$$

Update the projection matrices $\boldsymbol{M}$ and $\boldsymbol{N}$. The updates for $\boldsymbol{M}$ and $\boldsymbol{N}$ employ an efficient solver developed using the conjugate gradient method (CG). Before the iteration, set $\boldsymbol{M}^{(0)} = 0$, $\boldsymbol{R}^{(0)} = \mu\boldsymbol{C}^T\boldsymbol{S}_1\boldsymbol{C} - \mu\boldsymbol{C}^{(T)}\boldsymbol{C}\boldsymbol{M}^{(0)}\boldsymbol{C}^{(T)}\boldsymbol{C} - \lambda\boldsymbol{M}^{(0)}$. Repeat the iteration until $\left\|\boldsymbol{R}^{(k)}\right\|_F^2$ is sufficiently small. The update process is as follows.

$$\min_{\boldsymbol{M}}(\frac{\mu}{2}\left\|\boldsymbol{S}_1 - \boldsymbol{C}\boldsymbol{M}\boldsymbol{C}^T\right\|_F^2) + \frac{\lambda}{2}\|\boldsymbol{M}\|_F^2 \tag{17}$$

$$\min_{\boldsymbol{N}}(\frac{\eta}{2}\left\|\boldsymbol{S}_2 - \boldsymbol{P}\boldsymbol{N}\boldsymbol{P}^T\right\|_F^2) + \frac{\lambda}{2}\|\boldsymbol{N}\|_F^2 \tag{18}$$

$\boldsymbol{M}$ and $\boldsymbol{N}$ are updated through the equations.

$$\boldsymbol{M} = CG(\boldsymbol{S}_1, \boldsymbol{C}, \boldsymbol{C}, \mu, \lambda), \boldsymbol{N} = CG(\boldsymbol{S}_2, \boldsymbol{P}, \boldsymbol{P}, \eta, \lambda) \tag{19}$$

Algorithm 1 outlines the key steps of the CTF method. The input consists of the tensor $\mathcal{X}$ representing the known DDI types, along with the drug structure similarity matrix $\boldsymbol{S}_1$, the drug-related biological similarity matrix $\boldsymbol{S}_2$, the rank r, the weighting parameters μ, η, the Hessian regularization parameters α, β, the penalty parameter θ, the regularization coefficient λ, and the maximum iterations t_{max}. The output consists of drug feature matrix $\boldsymbol{C}$ constrained by $\boldsymbol{S}_1$, drug feature matrix $\boldsymbol{P}$ constrained by $\boldsymbol{S}_2$, and association type feature matrix $\boldsymbol{F}$, respectively. Specifically, initialization of the algorithm involves stochastic-ally initializing the matrices, setting up the initial parameters, and defining the iteration index (Line 1). The iterative process of the algorithm is shown in Line 2 onward. Through the optimization derivation of the objective function, the projection matrices (Line 3), followed by the feature matrices are updated (Lines 4–8), until convergence is achieved and the loop terminates, as indicated by the convergence condition.

2.4. DNN module

To leverage the advantages of deep learning in feature extraction, we combined with the CTF method and a DNN, resulting in a hybrid model named the tensor factorization strategy with effective constrained terms for potential drug–drug interactions prediction (CTF-DDI). In this study, the DNN contains D fully connected layers, and the propagation process uses a forward propagation algorithm, where the output of the previous layer is used as the input of the next layer.

Algorithm 1: The algorithm of CTF

Input : $\mathcal{X}$, $\boldsymbol{S}_1$, $\boldsymbol{S}_2$, γ, μ, η, α, β, λ, θ and t_{max}
Output: feature matrices $\boldsymbol{C}$, $\boldsymbol{P}$, and $\boldsymbol{F}$
Initialization: $t \leftarrow 0$, $\boldsymbol{C}$, $\boldsymbol{P}$, $\boldsymbol{F}$, $\boldsymbol{Y}_1(t), \boldsymbol{Y}_2(t) = 0$, $\boldsymbol{Z}_1(t), \boldsymbol{Z}_2(t) = 0$, $\boldsymbol{R}_1, \boldsymbol{R}_2 = 0$, $\rho_1, \rho_2 = 1$;
while $t \leq t_{max}$ **do**
  Update $\boldsymbol{M}$, $\boldsymbol{N} \leftarrow$ Eq. (19)
  Update $\boldsymbol{F} \leftarrow$ Eq. (6)
  Update $\boldsymbol{J}_1$, $\boldsymbol{R}_1$, $\boldsymbol{C} \leftarrow$ Eqs. (9), (10), (11)
  Get $\boldsymbol{Z}_1(t+1)$, $\boldsymbol{Y}_1(t+1) \leftarrow$ Eq.
  Update $\boldsymbol{J}_2$, $\boldsymbol{R}_2$, $\boldsymbol{P} \leftarrow$ Eqs. (13), (14), (15)
  Get $\boldsymbol{Z}_2(t+1)$, $\boldsymbol{Y}_2(t+1) \leftarrow$ Eq. (16)
  if *the convergence condition is met* **then**
    break
  $t \leftarrow t + 1$
return $\boldsymbol{C}$, $\boldsymbol{P}$, $\boldsymbol{F}$

Table 1
Description of the experimental datasets.

Datasets	Drugs	Pairs	Interactions	Label value
DS1	548	300 304	97 168	[0, 1]
DS3: CYP	807	651 249	10 078	[0, 1]
DS3: NCYP	807	651 249	40 904	[0, 1]

We used binary cross-entropy (BCE) as loss function and the model is continuously optimized to minimize the loss function to obtain the predicted DDI probabilities.

Specifically, the feature matrices $\boldsymbol{C}$, $\boldsymbol{P}$, and $\boldsymbol{F}$, corresponding to each element x_{ijk} of the original tensor $\mathcal{X}$ via CTF, are concatenated into a column vector $(\boldsymbol{c}_i \ \boldsymbol{p}_j \ \boldsymbol{f}_k)'$ by selecting the ith row of $\boldsymbol{C}$, the jth row of $\boldsymbol{P}$, and the kth row of $\boldsymbol{F}$, respectively. Subsequently, a forward propagation algorithm traverses through each fully connected layer, employing the ReLU activation function $h(\cdot)$, i.e., $h(x) = relu(x) = max(x, 0)$. Following the learning process of the multi-layer fully-connected neural network, the output x^n is obtained. The $Sigmoid$ activation function is then applied to obtain the output as the predicted interaction x'_{ijk} between drug i and drug j under each association type. Consequently, for each element in the tensor $\mathcal{X}$, the corresponding row concatenation of its feature matrix serves as a separate input layer, which is learned by the DNN to obtain the prediction results of DDI under various association types. As a result, a new tensor $\mathcal{X}'$ is obtained, from which potential interactions and their association types can be inferred.

3. Results

3.1. Materials

(1) Datasets: To evaluate the performance of our proposed CTF-DDI model, we adopted two widely-used benchmark datasets. One is the DS1 [36], which contains 548 drugs and 97168 DDI. We used DS1 as a dataset for binary-class DDI prediction, and in the tensor $\mathcal{X}$, the value is 1 if there is an interaction between drugs, and 0 otherwise. The other is the DS3 [37] for multi-class DDI prediction, which involves two types of DDI, the Cytochrome P450 involved (CYP) and without involving Cytochrome P450 (NCYP). Table 1 summarizes the details of these two datasets.

(2) Evaluation metrics: To assess the predictive performance of the CTF and the CTF-DDI model, we conducted a five-fold cross-validation analysis on the datasets. And we employed several evaluation metrics, including Precision, Accuracy, Recall, F1-score (F1), the area under the ROC curve (AUC), and the area under the PR curve (AUPR).

Table 2
Comparison results of models on the DS1 dataset.

Model	AUC	AUPR	F1	Accuracy	Recall	Precision
SVD	0.9244	0.9287	0.8453	0.8436	0.8546	0.8362
HOPE	0.9379	0.9435	0.8572	0.8566	0.8608	0.8538
GraRep	0.9336	0.9391	0.8538	0.8531	0.8577	0.8500
Deepwalk	0.9322	0.9375	0.8513	0.8511	0.8524	0.8502
Node2vec	0.9297	0.9354	0.8478	0.8479	0.8473	0.8483
LINE	0.9173	0.9234	0.8307	0.8304	0.8322	0.8292
SDNE	0.9173	0.9204	0.8338	0.8321	0.8424	0.8254
GAE	0.8941	0.8673	0.8227	0.8137	0.8647	0.7846
CTF-DDI	**0.9553**	**0.9579**	**0.8843**	**0.8830**	**0.8944**	**0.8745**

(3) Baselines: We conducted comparative experiments in terms of both matrix factorization-based and drug similarity-based DDI prediction models.

Matrix factorization-based models: (i) **SVD** is one of the traditional methods in matrix factorization. Given a drug–drug similarity matrix, a reconstructed approximation matrix can be obtained [24]. (ii) **HOPE** adopts some well-known network similarity measures such as Katz Index and Common Neighbors to preserve network structures [26]. (iii) **GraRep** considers the high-order proximity of network and designs k-step transition probability matrices for factorization [25].

Similarity-based models: (i) **DeepWalk** takes drug pairs as node inputs to the isomorphic graph and feeds them into the classifier used for DDI prediction [14]. (ii) **Node2vec** adopts a flexible biased random walk procedure that smoothly combines breadth-first sampling and depth-first sampling to generate node sequences [15]. (iii) **LINE** directly models node embedding vectors by approximating the first-order proximity and second-order proximity of nodes, which can be seen as a single-layer MLP model [16]. (iv) **SDNE** adopts a deep auto-encoder to preserve the second-order proximity by reconstructing the neighborhood structure of each node [17]. (v) **GAE** utilizes a GCN encoder and an inner product decoder to learn node embedding [18].

(4) Parameter setting: In the CTF method, six parameters are involved: γ, μ, η, α, β, θ, and λ. Here, γ is the rank of the reconstruction tensor. Due to the involvement of numerous crucial parameters in tensor factorization models, different combinations of parameter settings can, to some extent, affect predictive performance. Therefore, λ is set to 0.5, γ ranges from 1 to 100, and μ, η, α, β, and θ vary within the range of 0 to 1. The method is iterated for 200 times, with the automatic stopping criterion being the convergence of predicted variances. Ultimately, the optimal performance is obtained when $\gamma = 51$, $\mu = 0.5$, $\eta = 0.2$, $\alpha = 0.5$, $\beta = 0.5$, and $\theta = 0.5$. Similarly, setting the number of hidden layers in the DNN of the CTF-DDI model to 3, and setting the batch size and epoch to 1000 and 300, respectively, yields the best model performance.

3.2. Performance comparison

To evaluate the predictive performance of the CTF-DDI model proposed in this study, a five-fold cross-validation was conducted on both the DS3 dataset and the DS1 dataset. Classic models are selected for comparison based on the characteristics of the two datasets. Therefore, we compared CTF-DDI with other models, including SVD, HOPE, GraRep, DeepWalk, Node2vec, LINE, SDNE, and GAE, for the DDI prediction task. On the DS3 dataset, interactions are categorized into CYP and NCYP types. To assess the predictive performance of the CTF-DDI model under these two interaction types, the experimental results were classified accordingly. On the DS1 dataset, this study comprehensively evaluated the overall predictive performance of the CTF-DDI model.

As shown in Table 2, the CTF-DDI model achieved the highest average AUPR of 95.8% and average AUC of 95.5% on the DS1 dataset. It can be clearly observed that CTF-DDI outperformed other models in all metrics. Specifically, CTF-DDI outperformed other models by 1.82%

Table 3
Comparison results of models under CYP and NCYP on the DS3 dataset.

	Model	AUC	AUPR	F1	Accuracy	Recall	Precision
CYP	SVD	0.8714	0.8687	0.7899	0.7895	0.7915	0.7883
	HOPE	0.9113	0.8974	0.8391	0.8391	0.8391	0.8391
	GraRep	0.9122	0.8984	0.8335	0.8332	0.8352	0.8318
	Deepwalk	0.9000	0.8928	0.8236	0.8232	0.8252	0.8220
	Node2vec	0.8949	0.8848	0.8161	0.8163	0.8153	0.8169
	LINE	0.8636	0.8603	0.7894	0.7870	0.7894	0.7806
	SDNE	0.8949	0.8750	0.8380	0.8366	0.8451	0.8311
	GAE	0.7977	0.7220	0.7671	0.7219	0.9156	0.6600
	CTF-DDI	**0.9807**	**0.9802**	**0.9288**	**0.9281**	**0.9368**	**0.9212**
NCYP	SVD	0.8698	0.8496	0.7896	0.7892	0.7910	0.7883
	HOPE	0.8886	0.8674	0.8166	0.8138	0.8289	0.8047
	GraRep	0.8850	0.8659	0.8114	0.8090	0.8215	0.8015
	Deepwalk	0.8791	0.8626	0.8010	0.8010	0.8044	0.7976
	Node2vec	0.8622	0.8414	0.7862	0.7881	0.7790	0.7935
	LINE	0.8499	0.8307	0.7705	0.7647	0.7900	0.7519
	SDNE	0.8651	0.8360	0.8029	0.7919	0.8474	0.7628
	GAE	0.7830	0.7307	0.7521	0.7092	0.8824	0.6553
	CTF-DDI	**0.9733**	**0.9754**	**0.9227**	**0.9229**	**0.9205**	**0.9251**

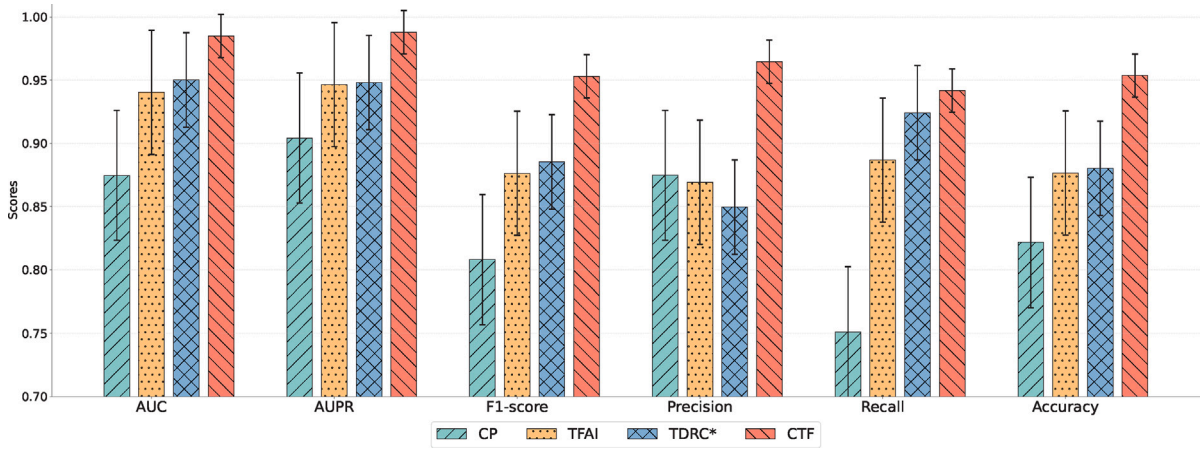


Fig. 2. Comparison results of other models on DS3 dataset.

to 6.40% on the AUC, 1.50% to 9.46% on the AUPR, and 3.06% to 6.97% on the F1-score.

Table 3 depicts the comparison results of these models on the DS3 dataset, where the best results are highlighted in bold. The results indicate that the CTF-DDI model demonstrates superior performance in AUC, AUPR, F1-score, Recall, Precision, and Accuracy. For the CYP interaction, the AUC and the AUPR reached 0.9807 and 0.9802, respectively. Similarly, for the NCYP interaction, the model exhibited slightly better performance compared to other models, with the AUC and the AUPR reaching 0.9733 and 0.9754, respectively. Furthermore, CTF-DDI achieved relative improvements of at least 9.66% and 11.5% in F1-score for CYP and NCYP interactions, respectively.

These superior prediction performances indicate that our proposed CTF-DDI model is better suited for capturing potential features of drugs. The CTF method fully utilizes drug information and provides a more efficient representation of features. This presents an efficient and reliable model choice for bioinformatics and drug research in related fields, offering robust support for research and applications within biomedical domain.

3.3. Ablation study

We performed ablation experiments with the CTF method using the conventional tensor factorization method: (i) **CP** is one of the common forms of standard tensor decomposition without adding any other constraints. (ii) **TFAI** introduces graph regularization based on the CP model but does not make use of similarity information [27]. (iii) **TDRC** is a proposed tensor factorization method based for miRNA-disease association prediction, which only uses similarity information as a constraint based on the CP model [28]. In this paper, we also used it as a baseline to validate the CTF method, denoted as TDRC*.

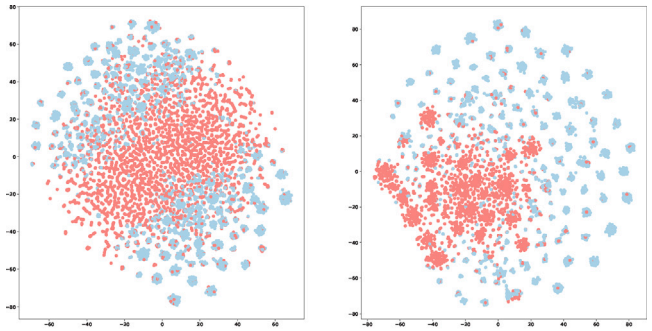
Since parameter settings have a certain impact on the predictive performance of the models, for the sake of fairness, the parameters of various tensor factorization models are set to their optimal values for comparison. As shown in Fig. 2, the CTF model significantly outperforms the CP, TFAI, and TDRC* models across all evaluation metrics, with scores of 0.9844, 0.9879, 0.9530, 0.9645, 0.9418, and 0.9535, respectively. Compared to the other models, CTF shows at least 6.97%, 9.31%, and 7.49% improvement in F1-score, Precision, and Accuracy, respectively. The error bars represent standard deviation, and the CTF model's shorter error bars indicate more stable performance. The final results show that optimal DDI prediction performance is achieved by effectively combining various drug features and corresponding constraints (i.e., drug similarity, Hessian regularization, and $L_{2,1}$ regularization), emphasizing the importance of each component and supporting the model's effectiveness in handling DDI prediction tasks.

3.4. Visualization experiment

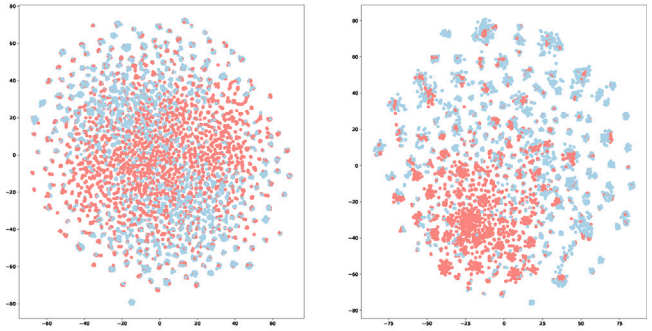
We iteratively trained the proposed CTF method and examined the results by projecting the features of the DS3 dataset into a two-dimensional space using t-distributed stochastic neighbor embedding

Table 4
Top 5 DDI and their evidence for interaction with drug *Amphetamine* and *Bromocriptine*.

	Name	Type	Evidence	Description
Amphetamine	*Almasilate*	72	DrugBank	The combination of *Amphetamine* and *Almasilate* may raise serum levels.
	Selexipag	60	DrugBank	*Amphetamine* may decrease the antihypertensive activities of *Selexipag*.
	Thonzylamine	76	DrugBank	*Amphetamine* may decrease the sedative and stimulatory activities of *Thonzylamine*.
	Desoxyn	49	–	unconfirmed.
	Oxprenolol	60	DrugBank	The therapeutic efficacy of *Oxprenolol* can be decreased when used in combination with *Amphetamine*.
Bromocriptine	*Ubidecarenone*	73	DrugBank	The risk or severity of hypoglycemia is increased when *Ubidecarenone* is combined with *Bromocriptine*.
	Lisuride	19	DrugBank	The risk or severity of adverse effects can be decreased when *Lisuride* is combined with *Bromocriptine*.
	Lorcaserin	49	DrugBank	The risk or severity of adverse effects can be increased when *Lorcaserin* is combined with *Bromocriptine*.
	Phenobarbital	4	DrugBank	The metabolism of *Bromocriptine* can be increased when combined with *Phenobarbital*.
	Trovafloxacin	9	DrugBank	The therapeutic efficacy of *Bromocriptine* can be increased when used in combination with *Trovafloxacin*.



(a) The CYP interaction



(b) The NCYP interaction

Fig. 3. t-SNE visualization of learned representations under different interaction.

(t-SNE). Specifically, we first randomly selected 10,000 drug–drug–type data samples $(c_i\ p_j\ f_k)$ from the original tensor $\mathcal{X}$, with 50% positive and 50% negative samples. Then the high-dimensional feature representations $(c_i\ p_j\ f_k)$ and $(c_i\ p_j\ f_k)'$, before and after CTF reconstruction, were projected into the two-dimensional space by t-SNE. Finally the results before and after training were compared for two interaction types, CYP and NCYP. The results are shown in Fig. 3, where red nodes represent negative samples and blue nodes represent positive samples. During the initial training process, the initial data samples are projected in the feature space, showing a lack of clear distinction between different sample categories. After iterative reconstruction of CTF, the similarity features of drugs can be effectively fused, and the Hessian and $L_{2,1}$ regularization constraint terms make the similar drugs closer in the potential space while the data sparsity is lower. From the results shown in Figs. 3(a) and 3(b), we can see that the reconstructed data samples, when projected under different interaction types, show a tendency to be classified and exhibit more obvious clustering patterns. In addition, we observe significant clustering of samples with the same labels in specific regions, which suggests that the method succeeds in categorizing similar samples within the learned feature space, thus facilitating better differentiation between similar samples. This not only emphasizes the effectiveness of CTF in predicting DDI, but also implies that CTF-DDI provides a meaningful representation of the data.

3.5. Case study

To further validate the predictive capability of the CTF-DDI model for predicting DDI, we conducted a case analysis on two drugs, *Amphetamine* (DB00182) and *Bromocriptine* (DB01200), selected from the DrugBank 5.0 dataset. During the experiment, each drug was tested individually by removing its interactions with other drugs. In other words, each drug was treated as a new drug with no interactions with any other drugs in the original tensor, ensuring the accuracy of the experiments. Subsequently, the predicted results were analyzed, and the top 20 drug pairs interacting with *Amphetamine* and *Bromocriptine* were selected. Evidence supporting these predicted drug pairs was then sought using PubMed, DrugBank, and Drug Interactions Checker provided by Drugs.com. Table 4 shows the top 5 drug pair results for *Amphetamine* and *Bromocriptine*, and the complete results are included in Table B.2 and Table B.3 in the Appendix. The research findings indicate that out of the 20 predicted DDI for *Amphetamine*, 18 have supporting evidence, while for *Bromocriptine*, 19 out of the 20 predicted interactions are confirmed. These above cases show that our proposed CTF-DDI model can effectively predict potential drug–drug pairs and thus contribute to drug combination therapy.

4. Conclusion

In this paper, we proposed a novel tensor factorization strategy with effective constraint terms (CTF-DDI) for predicting DDI and their

types. The proposed CTF-DDI captures multiple features of drug similarity information through tensor factorization with constraints, and it subsequently extracts nonlinear features to predict DDI by using a DNN. Based on CP factorization, the constrained tensor factorization method (CTF) leverages drug similarity information and uses Hessian and $L_{2,1}$ regularization terms as constraints. Compared with the traditional tensor factorization methods, our method performed better in prediction. In addition, we constructed the CTF-DDI model by training a DNN, which can predict DDI more effectively. Through comparative analysis with other DDI prediction models, the superior prediction performance of CTF-DDI model was finally verified. However, it is important to acknowledge the limitations of our approach. One such limitation lies in the methods employed for fusing drug similarity information, particularly concerning the availability of reliable negative samples. Additionally, the data also lacks reliable negative samples, and information on non-interaction pairs is not accurately obtained. In future, we aim to address these limitations by refining our selection and fusion methods for similarity information. In summary, the CTF-DDI model offers an effective solution for DDI prediction, enabling the inference of complex disease-drug combinations.

CRediT authorship contribution statement

Guosheng Han: Methodology, Conceptualization. **Lingzhi Peng:** Writing – original draft, Formal analysis. **Aocheng Ding:** Visualization, Software. **Yan Zhang:** Writing – original draft. **Xuan Lin:** Writing – review & editing, Supervision.

Declaration of competing interest

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

Data availability

The data that has been used is confidential.

Acknowledgments

We thank the editors and reviewers for their efforts in reviewing this manuscript. This work is supported in part by Natural Science Foundation of Hunan Province of China (No. 2021JJ30684, 2021JJ10020, No. 2022JJ40451), Key Foundation of Hunan Educational Committee, China (No. 19A497), National Natural Science Foundation of China (No. 62202413), Excellent Youth Funding of Hunan Provincial Education Department, China (No. 23B0129).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.future.2024.06.060.

References

[1] M. Huang, Z. Jiang, S. Guo, Phar-LSTM: A pharmacological representation-based LSTM network for drug-drug interaction extraction, PeerJ 11 (2023) e16606.
[2] S. Lin, Y. Wang, L. Zhang, Y. Chu, Y. Liu, Y. Fang, M. Jiang, Q. Wang, B. Zhao, Y. Xiong, et al., MDF-SA-DDI: Predicting drug-drug interaction events based on multi-source drug fusion, multi-source feature fusion and transformer self-attention mechanism, Brief. Bioinform. 23 (1) (2022) bbab421.
[3] J.K. Amaral, G. Lucena, R.T. Schoen, Chikungunya arthritis treatment with methotrexate and dexamethasone: A randomized, double-blind, placebo-controlled trial, Current Rheumatology Reviews. 20 (3) (2024) 337–346.
[4] C. Li, H. Che, M. Leung, C. Liu, Z. Yan, Robust multi-view non-negative matrix factorization with adaptive graph and diversity constraints, Inform. Sci. 634 (2023) 587–607.
[5] Z. Liu, X. Luo, M. Zhou, Symmetry and graph bi-regularized non-negative matrix factorization for precise community detection, IEEE Trans. Automat. Sci. Eng. 21 (2) (2024) 1406–1420.
[6] Z. Li, S. Zhu, B. Shao, X. Zeng, T. Wang, T.Y. Liu, DSN-DDI: An accurate and generalized framework for drug-drug interaction prediction by dual-view representation learning, Brief. Bioinform. 24 (1) (2023) bbac597.
[7] S. Liu, Y. Zhang, Y. Cui, Y. Qiu, Y. Deng, Z. Zhang, W. Zhang, Enhancing drug-drug interaction prediction using deep attention neural networks, IEEE/ACM Trans. Comput. Biol. Bioinform. 20 (2) (2022) 976–985.
[8] X. Shen, Z. Li, Y. Liu, B. Song, X. Zeng, PEB-DDI: A task-specific dual-view substructural learning framework for drug-drug interaction prediction, IEEE J. Biomed. Health Inform. 28 (1) (2024) 569–579.
[9] S. Vilar, R. Harpaz, E. Uriarte, L. Santana, R. Rabadan, C. Friedman, Drug-drug interaction through molecular structure similarity analysis, J. Am. Med. Inform. Assoc. 19 (6) (2012) 1066–1074.
[10] R. Ferdousi, R. Safdari, Y. Omidi, Computational prediction of drug-drug interactions based on drugs functional similarities, J. Biomed. Inform. 70 (2017) 54–64.
[11] Y. Deng, X. Xu, Y. Qiu, J. Xia, S. Liu, A multimodal deep learning framework for predicting drug-drug interaction events, Bioinformatics 36 (15) (2020) 4316–4322.
[12] Q. Jin, J. Xie, D. Huang, C. Zhao, H. He, MSFF-MA-DDI: Multi-source feature fusion with multiple attention blocks for predicting drug–drug interaction events, Comput. Biol. Chem. 108 (2024) 108001.
[13] P. Zhang, F. Wang, J. Hu, R. Sorrentino, Label propagation prediction of drug-drug interactions based on clinical side effects, Sci. Rep. 5 (1) (2015) 12339.
[14] B. Perozzi, R. Al-Rfou, S. Skiena, Deepwalk: Online learning of social representations, in: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014, pp. 701–710.
[15] A. Grover, J. Leskovec, Node2vec: Scalable feature learning for networks, in: Proceedings of the 22th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 855–864.
[16] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, Q. Mei, Line: Large-scale information network embedding, in: Proceedings of the 24th International Conference on World Wide Web, 2015, pp. 1067–1077.
[17] D. Wang, P. Cui, W. Zhu, Structural deep network embedding, in: Proceedings of the 22th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1225–1234.
[18] T.N. Kipf, M. Welling, Variational graph auto-encoders, 2016, ArXiv.org.
[19] N. Rohani, C. Eslahchi, A. Katanforoush, ISCMF: Integrated similarity-constrained matrix factorization for drug-drug interaction prediction, Netw. Model. Anal. Health Inform. Bioinform. 9 (2020) 1–8.
[20] S. Sadeghi, J. Lu, A. Ngom, A network-based drug repurposing method via non-negative matrix factorization, Bioinformatics 38 (5) (2022) 1369–1377.
[21] Y. Ma, Y. Zhao, Y. Ma, Kernel Bayesian nonlinear matrix factorization based on variational inference for human-virus protein-protein interaction prediction, Sci. Rep. 14 (1) (2024) 5693.
[22] S. Vilar, E. Uriarte, L. Santana, N.P. Tatonetti, C. Friedman, Detection of drug-drug interactions by modeling interaction profile fingerprints, PLoS One 8 (3) (2013) e58321.
[23] W. Zhang, Y. Chen, D. Li, X. Yue, Manifold regularized matrix factorization for drug-drug interaction prediction, J. Biomed. Inform. 88 (2018) 90–97.
[24] A. Ahmed, N. Shervashidze, S. Narayanamurthy, V. Josifovski, A.J. Smola, Distributed large-scale natural graph factorization, in: Proceedings of the 22nd International Conference on World Wide Web, 2013, pp. 37–48.
[25] S. Cao, W. Lu, Q. Xu, Grarep: Learning graph representations with global structural information, in: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, 2015, pp. 891–900.
[26] M. Ou, P. Cui, J. Pei, Z. Zhang, W. Zhu, Asymmetric transitivity preserving graph embedding, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1105–1114.
[27] A. Narita, K. Hayashi, R. Tomioka, H. Kashima, Tensor factorization using auxiliary information, Data Min. Knowl. Discov. 25 (2012) 298–324.
[28] F. Huang, X. Yue, Z. Xiong, Z. Yu, S. Liu, W. Zhang, Tensor decomposition with relational constraints for predicting multiple types of microRNA-disease associations, Brief. Bioinform. 22 (3) (2021) bbaa140.
[29] H. Jiang, Z. Diao, T. Shi, Y. Zhou, F. Wang, W. Hu, X. Zhu, S. Luo, G. Tong, Y.-D. Yao, A review of deep learning-based multiple-lesion recognition from medical images: Classification, detection and segmentation, Comput. Biol. Med. 157 (2023) 106726.
[30] U.A. Bhatti, M. Huang, H. Neira-Molina, S. Marjan, M. Baryalai, H. Tang, G. Wu, S.U. Bazai, MFFCG: Multi feature fusion for hyperspectral image classification using graph attention network, Expert Syst. Appl. 229 (2023) 120496.
[31] D. Khurana, A. Koli, K. Khatter, S. Singh, Natural language processing: State of the art, current trends and challenges, Multimedia Tools Appl. 82 (3) (2023) 3713–3744.
[32] W. Zhang, X. Yue, W. Lin, W. Wu, R. Liu, F. Huang, F. Liu, Predicting drug-disease associations by using similarity constrained matrix factorization, BMC Bioinform. 19 (2018) 1–12.
[33] J.Y. Ryu, H.U. Kim, S.Y. Lee, Deep learning improves prediction of drug-drug and drug-food interactions, Proc. Natl. Acad. Sci. 115 (18) (2018) E4304–E4311.

[34] Y. Chen, T. Ma, X. Yang, J. Wang, B. Song, X. Zeng, MUFFIN: Multi-scale feature fusion for drug-drug interaction prediction, Bioinformatics 37 (17) (2021) 2651–2658.
[35] Z. Kang, Y. Lu, Y. Su, C. Li, Z. Xu, Similarity learning via kernel preserving embedding, in: Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, vol. 33, (no. 01) 2019, pp. 4057–4064.
[36] W. Zhang, Y. Chen, F. Liu, F. Luo, G. Tian, X. Li, Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data, BMC Bioinform. 18 (2017) 1–12.
[37] N. Rohani, C. Eslahchi, Drug-drug interaction predicting by neural network using integrated similarity, Sci. Rep. 9 (1) (2019) 13645.

Guosheng Han is currently a professor at the School of Mathematics and Computational Sciences, Xiangtan University, Xiangtan, China. He received his Ph.D. degree in Applied Mathematics from Xiangtan University in 2013. He has published dozens of papers in international conferences and journals, and chaired the National Natural Science Foundation of China (Youth Program) and the Key Research and Development Program of Hunan Province. His research interests include machine learning, deep learning and bioinformatics.

Lingzhi Peng is currently a M.S. candidate at the School of Mathematics and Computational Sciences, Xiangtan University, Xiangtan, China. Her research interests include machine learning, deep learning and bioinformatics.

Aocheng Ding is currently an undergraduate student at the School of Computer Science, Xiangtan University, Xiangtan, China. His research interests include machine learning, graph neural networks and bioinformatics.

Yan Zhang received her M.S. degree from the School of Mathematics and Computational Sciences, Xiangtan University, Xiangtan, China. Her research interests include machine learning, deep learning and bioinformatics.

Xuan Lin is currently a lecturer at the School of Computer Science, Xiangtan University, Xiangtan, China. Before joining Xiangtan University, he received the Ph.D. degree in computer science from Hunan University, Changsha, China, in 2021. He was visiting scholar in University of Illinois at Chicago, from 2019 to 2020. His main research interests include machine learning, graph neural networks and bioinformatics. He has published several research papers in these fields including IJCAI, AAAI, BIBM, Briefings in Bioinformatics, IEEE TKDE, TCBB etc.