MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **93** (2025) 599–629
ISSN: 0340–6253
doi: 10.46793/match.93-3.599X

Drug Repositioning Based on Dual Multi-Graph Regularization Nonnegative Matrix Factorization and Multi-kernel Neural Network

Hanxiao Xu[a], Da Xu[b], Yusen Zhang[a,*]

[a] *School of Mathematics and Statistics, Shandong University, Weihai, Shandong, 264200, China.*

[b] *School of Mathematical Sciences, Liaocheng University, Liaocheng, Shandong, 252000, China.*

hanxiaoxusdu@163.com, daxulcu@126.com, zhangys@sdu.edu.cn

(Received July 23, 2024)

Abstract

Drug repositioning is a valuable and efficient strategy to discover new applications for traditional medications. In contrast to experimental methods, developing accurate and effective computational methods is crucial. The identification of potential drug-disease associations is a vital aspect of drug repositioning. In the paper, we proposed a new computational model called DDNMFNN to identify potential drug-disease associations, combining nonnegative matrix factorization and neural networks. The sparsity of validated drug-disease associations leads to subpar model generalization performance. To address this issue, a novel dual multi-graph regularization nonnegative matrix factorization algorithm with adaptive weights is proposed to reconstruct the association matrix. An efficient optimization algorithm is designed and convergence proof is provided. Furthermore, a multi-kernel neural network is utilized to predict potential associations based on the multiple similarity matrices and the reconstructed association matrix. This network effectively combines the nonparametric flexibility of the multi-kernel

*Corresponding author.

method with the structural characteristics of deep learning. The experimental results of 10-fold cross-validation demonstrate the proposed model achieved the best performance by comparing it with state-of-the-art models on three datasets. Case studies of three diseases and prediction results of five real-world network datasets further indicate that the proposed model as a precise prediction tool that can facilitate drug repositioning efforts effectively.

1 Introduction

Bringing a drug successfully through all stages of drug development into clinical practice costs more than 1.5–2.5 billion dollars [1]. The clinical trial remain expensive, with a high failure rate. While national investment in new drug development has significantly increased, only a few drugs have been approved for marketing [2]. To expedite research and development while minimizing costs, drug repositioning technology focuses on identifying existing drugs that could potentially treat a specific disease, a field that is gaining momentum [3]. It utilizes publicly available databases and advanced computational techniques to predict large-scale associations between drugs and diseases. Drug repositioning significantly cuts down on the time and expenses typically associated with developing new drugs through biomedical means, while also offering valuable insights for guiding biological experiments.

The advancement of machine learning algorithms has led to widespread application and successful outcomes in identifying potential drug-disease associations [4]. These computational methods can be categorized into traditional machine learning-based methods, matrix factorization-based methods, and deep learning-based methods. The rapid progress in traditional machine learning has opened up new possibilities for predicting association between drugs and diseases. The PREDICT model proposed by Gottlieb et al. [5] input similarity data of drugs and diseases into a logistic regression classifier to obtain potential association. Oh et al. [6] used an integrative genetic network and integrated three classifiers (Decision Tree, Random Forest, and Multi-layer Perceptron) for association prediction. Matrix-based methods primarily rely on matrix completion and matrix

factorization techniques to calculate missing values in association matrix. Zhang et al. [7] constructed a similarity constraint matrix factorization model. Luo et al. [8] developed a drug repositioning recommendation system (DRRS) that used a singular value thresholding algorithm to identify potential associations. DisDrugPred model proposed by Xuan et al. [9] which used diverse prior knowledge and nonnegative matrix factorization (NMF) for prediction. Zhang et al. [10] applied Bayesian inductive matrix completion technique for drug repositioning. Although these approaches have shown some success in association prediction, they struggle to capture intricate nonlinear structures within the network and often yield subpar results in learning deep feature representations of data. Deep learning leverages multi-layer neurons with complex structures and nonlinear transformations to create high-level abstract models of data. This approach has been highly successful in diverse research areas including visual data processing, natural language processing, social network analysis, and so on [11, 12]. In recent years, there has been successful implementation of deep learning in drug repositioning [13–15]. Yu et al. [16] embedded an attention mechanism into different convolutional layers of graph convolutional networks. Fu et al. designed a multi-view graph convolutional network model based on a graph neural network (GNN), which integrated heterogeneous information through the proposed neighborhood information aggregation layer [17]. Gao et al. utilized an attention mechanism and a bilinear GNN to design a context-aware neighborhood aggregation for extracting local and global features of drugs and diseases [18]. DRWBNCF, a deep learning model proposed by Meng et al., can effectively encode local neighbors and interaction information [19]. While artificial neural networks have shown advanced performance, they are limited in their ability to leverage sparse association data for superior performance.

Graph regularization is an attractive strategy to preserve the inherent geometry and discriminant structure of the data space [20]. NMF excels at uncovering hidden features or structures, with non-negative constraints enhancing problem interpretability. Cai et al. [21] introduced graph regularization constraints to NMF to address the limitation that matrix factorization neglects the manifold structure of the dataset. Ai et al. [22]

incorporated multi-graph regularization into low-rank matrix factorization to capture valuable information in manifold space. Kernel methods are effective for capturing nonlinear patterns in data, with their success heavily reliant on kernel selection [23]. In contrast to traditional fixed kernel methods, multi-kernel learning exhibits flexibility in automatic kernel learning, which is beneficial for learning tasks involving diverse data sources [24, 25]. Deep kernel methods combine the nonparametric flexibility of kernel methods with the structural characteristics of deep learning, proving to be an efficient solution [26–28].

In this paper, we proposed a novel computational model for drug repositioning called DDNMFNN. Specifically, we proposed a dual multi-graph regularization nonnegative matrix factorization (DMGNMF) to relieve the sparsity of association dataset. Dual multi-graph regularization terms are incorporated to preserve geometric structure of drug and disease data during reconstruction of association matrix, while Tikhonov (L_2) regularization terms for low-rank matrices are introduced to prevent overfitting. This algorithm adaptively merges the multi-view neighborhood information of each node into the NMF framework in a conditionally optimal manner. Finally, the prediction results are obtained by a multi-kernel neural network (MKNN). Adequate experiments were implemented on three datasets to test the predictive performance of DDNMFNN. Case studies were performed to assess the ability of DDNMFNN to predict new drugs related to diseases, and comparative experiments were carried out on five real world datasets to verify the practical effect of model. All experimental results consistently demonstrate the effectiveness of our model in identifying potential drug-disease relationships. The main contributions can be outlined as follows:

1. We proposed a new algorithm called DMGNMF based on NMF and dual multi-graph regularization to alleviate the sparsity of association data. It can extract multi-view spatial information of both drug and disease, and adaptively assigns weights to each view.

2. We designed an efficient optimization algorithm to solve DMGNMF, and provided convergence proof.

3. We proposed a new computational model called DDNMFNN for drug repositioning, achieving the fusion of matrix factorization and neural networks. The effectiveness and robustness of the model have been fully verified through extensive experiments.

2 Materials and methods

In this section, we provide a detailed introduction for the proposed model. It consists of three parts, as shown in Figure 1. The first part is to calculate the similarity matrices of drugs and diseases separately and extract multi-view biological information (Figure 1A). In the second part, the similarity matrix and original association matrix are input into the DMGNMF algorithm to obtain the reconstructed drug-disease association matrix (Figure 1B). Finally, the prediction results are obtained through MKNN (Figure 1C).

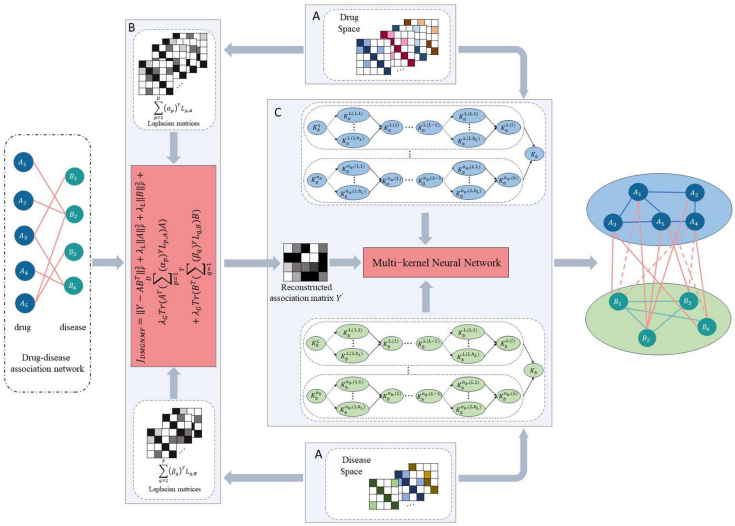

Figure 1. The overview of DDNMFNN model.

2.1 Datasets and similarity measures

In this paper, we used three datasets as benchmark datasets, and the summary of three benchmark datasets is presented in Table 1. The first dataset is Fdataset [5], which is a gold standard dataset collected from the DrugBank [29] and OMIM [30] databases. The second dataset is the Cdataset collected by Luo et al. [31]. The third dataset is LRSSLdatase [32], which was collected from the DrugBank [29] and MeSH [33] databases. Adequate information about the biochemical properties of drugs and diseases helps to provide a high-quality and comprehensive representation for the model. We used seven similarity information of drugs and two similarity information of diseases, as shown in Table 2.

Table 1. Summary of three benchmark datasets.

Data sets	Drugs	Diseases	Dimensions	Associations	Sparsity
Fdataset	593	313	593×313	1933	1.04%
Cdataset	663	409	663×409	2353	0.87%
LRSSLdataset	763	681	763×681	3051	0.59%

Table 2. Summary of similarity construction methods of drug and disease.

Spaces	Similarities	Descriptions
Drug	KD_t	Clinical similarity between drug and drug [22]
	KD_{GIP1}	GIP similarity between drug and target [34]
	KD_{GIP2}	GIP similarity between drug-disease [34]
	KD_{LINGO}	LINGO similarity between drug and drug [35]
	KD_{edit}	Editing distance similarity between drug and drug
	KD_f	Chemical fingerprints similarity between drug and drug [31]
	KD_s	Chemical fingerprints similarity between drug and drug [36]
Disease	KS_{GIP}	GIP similarity between disease and drug [34]
	KS_{SEM}	Semantic similarity between disease and disease [37]

2.2 Dual multi-graph regularization nonnegative matrix factorization

2.2.1 The DMGNMF algorithm

Nonnegative Matrix factorization (NMF) is an effective data analysis tool, which was first proposed by Lee and Seung [38]. Due to its nonnegative

conditional limitations, it has a good explanation for the local characteristics of objects. We hope to maintain the intrinsic geometric structure of the sample data in high-dimensional space after performing low-dimensional projection. Based on manifold learning theory [20, 39] and spectral graph theory [40,41], we can discover that the Laplacian regularization constraint can effectively approximate the local geometric features of data. Cai et al. [21] proposed the GRNMF algorithm based on graph regularization, which can be formulated as:

$$\begin{aligned} J_{GRNMF} &= \left\|Y - AB^T\right\|_F^2 + \lambda Tr(B^T LB) \\ &s.t. A \geq 0, B \geq 0, \end{aligned} \tag{1}$$

where $Y \in R^{m\times n}$ is association matrix; $A \in R^{m\times k}$ and $B \in R^{n\times k}$ ($k \ll min(m,n)$) are nonnegative matrices; $\|\cdot\|_F^2$ means Frobenius norm of the matrix; $L = D - W$ is the Laplacian matrix, D is a diagonal matrix and $D_{ii} = \sum_j W_{ij}$, W represents similarity matrices; $Tr(\cdot)$ denotes the trace of the matrix.

The construction and selection of graphs are pivotal for the performance of the GRNMF algorithm. In this paper, we proposed the DMGNMF algorithm to reconstruct association matrix by merging dual multi-graph regularization terms into the NMF framework, which can adaptively assign appropriate weights for each graph regularization term. In addition, to adjust the smoothness of A and B and prevent overfitting, we applied L_2 regularization constraints on low-rank matrices. We can obtain the final DMGNMF objective function as follows:

$$\begin{aligned} J_{DMGNMF} = &\left\|Y - AB^T\right\|_F^2 + \lambda_L \left(\|A\|_F^2 + \|B\|_F^2\right) \\ &+ \lambda_G \left(Tr\left(A^T \left(\sum_{p=1}^{D} (\alpha_p)^\gamma L_{p,A} \right) A \right) \right. \\ &\left. + Tr\left(B^T \left(\sum_{q=1}^{T} (\beta_q)^\gamma L_{q,B} \right) B \right) \right) \end{aligned}$$

$$s.t. A \geq 0, B \geq 0; \sum_{p=1}^{D} \alpha_p = 1, \alpha_p \in [0,1]; \sum_{q=1}^{T} \beta_q = 1, \beta_q \in [0,1], \tag{2}$$

where λ_L and λ_G are the regularization coefficients; α_p and β_q are the weights of the graph Laplacian matrix $L_{p,A} = D_{p,A} - W_{p,A}$ and $L_{q,B} = D_{q,B} - W_{q,B}$, respectively; $W_{p,A}$ and $W_{q,B}$ are defined as similarity matrices of drugs and diseases, respectively; $\gamma > 1$ is the exponential term and applied to regulate the impact of differences in the smoothness of the graphs.

2.2.2 Optimization

The optimization objective function is nonconvex, and it is challenging to find its global minimum because it comprises multiple variables. For a single variable, the objective function is a convex function and differentiable. It is feasible to apply an iterative multiplicative updating algorithm to find a locally optimal solution. In other words, the optimization problem is divided into four subproblems, and each variable is alternately solved while the others are fixed.

Update rules for A and B

The solving processes for A and B are similar, therefore they are presented simultaneously in this part. Firstly, according to property of the Frobenius norm and trace, the objective function is transformed into:

$$\begin{aligned} J_{DMGNMF} = & Tr(Y^T Y) - 2Tr(YBA^T) + Tr(AB^T BA^T) \\ & + \lambda_L (Tr(A^T A) + Tr(B^T B)) \\ & + \lambda_G \left(Tr\left(A^T \left(\sum_{p=1}^{D} (\alpha_p)^\gamma L_{p,A} \right) A \right) \right. \\ & \left. + Tr\left(B^T \left(\sum_{q=1}^{T} (\beta_q)^\gamma L_{q,B} \right) B \right) \right). \end{aligned} \tag{3}$$

Let $\Phi = (\phi_{ik})$ and $\Psi = (\psi_{jk})$ as the Lagrange multiplier matrices of constraint $A \geq 0$ and $B \geq 0$, respectively, we can get Lagrange function

$\mathcal{L}_{DMGNMF}$:

$$\begin{aligned}\mathcal{L}_{DMGNMF} &= Tr(Y^TY) - 2Tr(YBA^T) + Tr(AB^TBA^T) \\ &+ \lambda_L(Tr(A^TA) + Tr(B^TB)) \\ &+ \lambda_G \left(Tr\left(A^T \left(\sum_{p=1}^{D} (\alpha_p)^\gamma L_{p,A} \right) A \right) \right. \\ &\left. + Tr\left(B^T \left(\sum_{q=1}^{T} (\beta_q)^\gamma L_{q,B} \right) B \right) \right) \\ &+ Tr(\Phi A^T) + Tr(\Psi B^T). \end{aligned} \tag{4}$$

The partial derivatives of $\mathcal{L}_{DMGNMF}$ with respect to A and B are as follows:

$$\frac{\partial \mathcal{L}_{DMGNMF}}{\partial A} = -2YB + 2AB^TB + 2\lambda_L A + 2\lambda_G \left(\sum_{p=1}^{D} (\alpha_p)^\gamma L_{p,A} \right) A + \Phi, \tag{5a}$$

$$\frac{\partial \mathcal{L}_{DMGNMF}}{\partial B} = -2Y^TA + 2BA^TA + 2\lambda_L B + 2\lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma L_{q,B} \right) B + \Psi. \tag{5b}$$

Finally, using the Karush Kuhn Tucker (KKT) complementarity condition [42], we can obtain the multiplication update rules for A and B:

$$a_{ik} \leftarrow a_{ik} \frac{\left(YB + \lambda_G \left(\sum_{p=1}^{D} (\alpha_p)^\gamma W_{p,A} \right) A \right)_{ik}}{\left(AB^TB + \lambda_L A + \lambda_G \left(\sum_{p=1}^{D} (\alpha_p)^\gamma D_{p,A} \right) A \right)_{ik}}, \tag{6a}$$

$$b_{ik} \leftarrow b_{ik} \frac{\left(Y^TA + \lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma W_{q,B} \right) B \right)_{ik}}{\left(BA^TA + \lambda_L B + \lambda_G \left(\sum_{q=1}^{T} (\beta_p)^\gamma D_{q,B} \right) B \right)_{ik}}. \tag{6b}$$

Update rules for $\boldsymbol{\alpha_p}$ and $\boldsymbol{\beta_q}$

The solving processes for α_p and β_q are similar, therefore they are presented simultaneously in this part. Fixed A and B, J_{MGNMF} can be

simplified as follows:

$$\begin{aligned} J_{DMGNMF}(\alpha_p, \beta_q) = \lambda_G \Bigg(& Tr\left(A^T \left(\sum_{p=1}^{D} (\alpha_p)^\gamma L_{p,A} \right) A \right) \\ & + Tr\left(B^T \left(\sum_{q=1}^{T} (\beta_q)^\gamma L_{q,B} \right) B \right) \Bigg) \\ & s.t. \sum_{p=1}^{D} \alpha_p = 1, \alpha_p \in [0,1]; \sum_{q=1}^{T} \beta_q = 1, \beta_q \in [0,1]. \end{aligned} \tag{7}$$

For α_p and β_q, we apply the Lagrange multiplier method separately to obtain the following solution formulas:

$$\alpha_p = \frac{\left(\frac{1}{Tr\left(A^T L_{p,A} A\right)} \right)^{\frac{1}{\gamma-1}}}{\sum_{p=1}^{D} \left(\frac{1}{Tr\left(A^T L_{p,A} A\right)} \right)^{\frac{1}{\gamma-1}}}, \tag{8a}$$

$$\beta_q = \frac{\left(\frac{1}{Tr\left(B^T L_{q,B} B\right)} \right)^{\frac{1}{\gamma-1}}}{\sum_{q=1}^{T} \left(\frac{1}{Tr\left(B^T L_{q,B} B\right)} \right)^{\frac{1}{\gamma-1}}}. \tag{8b}$$

2.2.3 Convergence analysis

We analyzed the convergence of DMGNMF utilizing the auxiliary function method. According to the update rules of low-rank matrices A and B, the following theorem is derived:

Theorem 1. *The objective function J_{DMGNMF} is nonincreasing under the updating rules in Eqs.(6a) and (6b).*

Obviously, function J_{DMGNMF} is larger than zero. We only need to prove that function J_{DMGNMF} is nonincreasing under update rules in Eqs.(6a) and (6b). The regularization terms of A and B in function J_{DMGNMF} are independent, so we can refer to the proof program in the original NMF [38]. First, we provide an auxiliary function and Lemma 1 as follows:

Definition 1. $G(h, h^t)$ is an auxiliary function for $F(h)$ if the conditions

$$G(h, h^t) \geq F(h),\ G(h, h) = F(h)$$

are satisfied.

Lemma 1. *If G is an auxiliary function of F, then F is nonincreasing under the update*

$$h^{t+1} = \arg\min_h G(h, h^t). \tag{9}$$

Proof. $F(h^{t+1}) \leq G(h^{t+1}, h^t) \leq G(h^t, h^t) = F(h^t)$. ■

Next, we will prove Theorem 1 by defining an appropriate auxiliary function $G(h, h^t)$ for J_{DMGNMF}. The objective function J_{DMGNMF} is rewritten into elemental form, as shown below:

$$\begin{aligned} J_{DMGNMF} = & \sum_{i=1}^{m}\sum_{j=i}^{n}(y_{ij} - \sum_{k=1}^{K} a_{ik}b_{jk})^2 + \lambda_L \sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij})^2 + \lambda_L \sum_{i=1}^{m}\sum_{j=1}^{n}(b_{ij})^2 \\ & + \lambda_G \sum_{k=1}^{K}\sum_{i=1}^{m}\sum_{l=1}^{m} a_{ik}\left(\sum_{p=1}^{D}(\alpha_p)^\gamma L_{p,A}\right)_{il} a_{lk} \\ & + \lambda_G \sum_{k=1}^{K}\sum_{j=1}^{n}\sum_{l=1}^{n} b_{jk}\left(\sum_{q=1}^{T}(\beta_q)^\gamma L_{q,B}\right)_{jl} b_{lk}. \end{aligned}$$

We suppose that b_{ij} is any element in B, and $F_{b_{ij}}$ is the part of J_{DMGNMF} that only involves the element b_{ij}. It is easy to obtain the first and second partial derivatives for J_{DMGNMF} with respect to B as follows:

$$\begin{aligned} F'_{b_{ij}} &= \left(\frac{\partial J_{DMGNMF}}{\partial B}\right)_{ij} \\ &= \left(-2Y^TA + 2BA^TA + 2\lambda_L B + 2\lambda_G\left(\sum_{q=1}^{T}(\beta_q)^\gamma L_{q,B}\right)B\right)_{ij}, \end{aligned} \tag{10}$$

$$F''_{b_{ij}} = 2(A^TA)_{jj} + 2\lambda_L + 2\lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma L_{q,B} \right)_{ii}, \tag{11}$$

Lemma 2. *The following function*

$$\begin{aligned} G(b, b_{ij}^t) =& F_{b_{ij}}(b_{ij}^t) + F'_{b_{ij}}(b_{ij}^t)(b - b_{ij}^t) \\ &+ \frac{\left(BA^TA + \lambda_L B + \lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma D_{q,B} \right) B \right)_{ij}}{b_{ij}^t} (b - b_{ij}^t)^2 \end{aligned} \tag{12}$$

is an auxiliary function for $F_{b_{ij}}(b)$.

Proof. Obviously, $G(b, b) = F(b)$. Then, we only need to prove that $G(b, b_{ij}^t) \geq F_{b_{ij}}(b)$. By using Eqs.(10) and (11), we can obtain the second-order Taylor expansion of $F_{b_{ij}}(b)$ at b_{ij}^t:

$$\begin{aligned} F_{b_{ij}}(b) =& F_{b_{ij}}(b_{ij}^t) + F'_{b_{ij}}(b_{ij}^t)(b - b_{ij}^t) \\ &+ \left((A^TA)_{jj} + \lambda_l + \lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma L_{q,B} \right)_{ii} \right) (b - b_{ij}^t)^2. \end{aligned} \tag{13}$$

Comparing Eqs.(12)and(13), we find that $G(b, b_{ij}^t) \geq F_{b_{ij}}(b)$ is equivalent to

$$\frac{\left(BA^TA + \lambda_L B + \lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma D_{q,B} \right) B \right)_{ij}}{b_{ij}^t} \geq (A^TA)_{jj} + \lambda_l + \lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma L_{q,B} \right)_{ii}.$$

Due to $\lambda_L > 0$ and $\lambda_G > 0$, $(\lambda_L B)_{ij} = \lambda_L b_{ij}^t$,

$$\begin{aligned} (BA^TA)_{ij} &= \sum_{k=1}^{K} b_{ij} (A^TA)_{kj} \geq b_{ij}^t (A^TA)_{jj}, \\ \left(\lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma D_{q,B} \right) B \right)_{ij} &= \lambda_G \sum_{l=1}^{n} \left(\sum_{q=1}^{T} (\beta_q)^\gamma D_{q,B} \right)_{il} b_{lj} \\ \geq \lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma D_{q,B} \right)_{ii} b_{ii}^t &\geq \lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma L_{q,B} \right)_{ii} b_{ii}^t. \end{aligned}$$

Thus, $G(b, b_{ij}^t) \geq F_{b_{ij}}(b)$. ■

Lemma 3. *The following function*

$$G(a, a_{ij}^t) = F_{a_{ij}}(a_{ij}^t) + F'_{a_{ij}}(a_{ij}^t)(a - a_{ij}^t) + \frac{\left(AB^TB + \lambda_L A + \lambda_G \left(\sum_{p=1}^{D} (\alpha_p)^\gamma D_{p,A}\right) A\right)_{ij}}{a_{ij}^t}(a - a_{ij}^t)^2, \tag{14}$$

is an auxiliary function for $F_{a_{ij}}(a)$. $F_{a_{ij}}$ is the part of J_{DMGNMF} that only involves the element a_{ij}.

The proof process is similar to Lemma 2, but due to space limitations, we will not provide a specific proof process here. We can now prove Theorem 1.

Proof. Replacing $G(h, h_{ij}^t)$ in Eq.(9) by Eq.(12) results in the update rule:

$$b_{ij}^{t+1} = \arg\min_b G(b, b_{ij}^t) = b_{ij}^t \frac{\left(Y^TA + \lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma W_{q,B}\right) B\right)_{ij}}{\left(BA^TA + \lambda_L B + \lambda_G \left(\sum_{q=1}^{T} (\beta_q)^\gamma D_{q,B}\right) B\right)_{ij}}. \tag{15}$$

According to Lemma 2, Eq.(12) is an auxiliary function for $F_{b_{ij}}(b)$. $F_{b_{ij}}(b)$ is nonincreasing under the update rule Eq.(15) according to Lemma 1.

Similarly, replacing $G(h, h_{ij}^t)$ in Eq.(9) by Eq.(14) results in the update rule:

$$a_{ij}^{t+1} = \arg\min_a G(a, a_{ij}^t) = a_{ij}^t \frac{\left(YB + \lambda_G \left(\sum_{p=1}^{D} (\alpha_p)^\gamma W_{p,A}\right) A\right)_{ij}}{\left(AB^TB + \lambda_L A + \lambda_G \left(\sum_{P=1}^{D} (\alpha_p)^\gamma D_{p,A}\right) A\right)_{ij}}. \tag{16}$$

Thus, $F_{a_{ij}}(a)$ is nonincreasing under the update rule Eq.(16) according to Lemma 1 and Lemma 3. To sum up, the proof of Theorem 1 is completed. ■

2.3 Multi-kernel neural network

We used a multi-kernel neural network (MKNN) proposed by Ai et al. [43] for association prediction. It is different from the traditional neural network in that MKNN takes the matrix as the input of the network and its activation function is a kernel function (such as RBF kernel, linear kernel, or polynomial kernel). In detail, its workflow is to input each similarity matrix into different activation functions, and then conduct weighted summation to obtain the output of the network of this layer. Therefore,

the output of the l-layer of MKNN can be calculated using the following formula:

$$\mathcal{K}_a^{i,(l)} = \sum_{k=1}^{h^l} \alpha^{i,(l,k)} g_k\left(\mathcal{K}_a^{i,(l-1)}\right) = \sum_{k=1}^{h^l} \alpha^{i,(l,k)} K_a^{i,(l,k)}, \tag{17a}$$

$$\mathcal{K}_b^{i,(l)} = \sum_{k=1}^{h^l} \beta^{i,(l,k)} g_k\left(\mathcal{K}_b^{i,(l-1)}\right) = \sum_{k=1}^{h^l} \beta^{i,(l,k)} K_b^{i,(l,k)}, \tag{17b}$$

where l represents l-th layer of network, i represents i-th kernel network, $\mathcal{K}_{a/b}^{i,(l)}$ is defined as the output for drugs/diseases, respectively; $K_{a/b}^{i,(l,k)}$ is defined as the k-th kernel matrix for drug/disease; $g_k(\cdot)$ is the k-th kernel function, $\alpha^{i,(l,k)}$ and $\beta^{i,(l,k)}$ are the weight parameter. We can obtain the drug feature representation K_a after forward propagation. K_a is calculated by weighted fusion of the outputs of each layer in the network. In the same way, the low-dimensional feature representation of drugs K_b can also be obtained:

$$K_a = \sum_{k=1}^{n_a} \alpha^{i,(L+1)} \mathcal{K}_a^{i,(L)}, \tag{18a}$$

$$K_b = \sum_{k=1}^{n_b} \beta^{i,(L+1)} \mathcal{K}_b^{i,(L)}, \tag{18b}$$

where $\alpha^{i,(L+1)}$ and $\beta^{i,(L+1)}$ are the weight coefficients.

Another difference from the traditional neural network is the loss function of MKNN, as shown below:

$$Loss = \left\| Y' - K_a \theta K_b^T \right\|_F^2 + \lambda \left\| \theta \right\|_F^2 \tag{19}$$

where Y' means the reconstruction association matrix by DMGNMF algorithm, and T represents the transposition of matrix K_b.The detailed procedure of the DDNMFNN are summarized in Algorithm 1.

3 Results and discussion

3.1 Implementation details and evaluation metrics

To evaluate the predictive performance of DDNMFNN, the 10-fold cross-validation (10-CV) was performed 10 times on three benchmark datasets.

Algorithm 1 : Algorithm of DDNMFNN.

Iput: Association matrix $Y_{train} \in R^{m\times n}$, drug similarity matrices $\{KD_t, KD_{ATC}, KD_{GIP}, KD_{LINGO}, KD_{edit}, KD_{f1}, KD_{f2}\} \in R^{7\times m\times m}$, disease similarity matrices $\{KS_{DIP}, KS_{SEM}\} \in R^{2\times n\times n}$, regularization parameters λ_L,λ_G,λ; the number iterations of DMGNMF algorithm N; the number iterations of MKNN M.

Output: Predicted association matrix Y^*.

Initialize matrices $A \in R^{m\times k} \geq 0$,$B \in R^{n\times k} \geq 0$, parameters α_p, β_q;
Construct the Laplacian matrices by the similarity matrices of drugs and diseases, respectively;
for $i = 1 \rightarrow N$ **do**
    Use Eq.(6a) to update low rank matrix A;
    Use Eq.(6b) to update low rank matrix B;
    Use Eq.(8a) to update parameter α_p;
    Use Eq.(8b) to update parameter β_q;
end for
Obtain the reconstruction matrix $Y' = A \times B^T$;
Initialize weight coefficients α^0 and β^0 in the MKNN;
Input drug similarity matrices, disease similarity matrices and Y' into MKNN;
for $i = 1 \rightarrow M$ **do**
    Use Eqs.(17a) and (18a) to calculate the output of the drug kernel network K_a;
    Use Eqs.(17b) and (18b) to calculate the output of the disease kernel network K_b;
    Solve θ in Eq.(19);
    Update weight coefficients α and β in the MKNN;
end for
return K_a, K_b and θ, and obtain $Y^* = K_a\theta K_b^T$.

At the same time, we used the area under the precise recall curve (AUPR) as evaluation metrics and the area under the receiver operating characteristic curve (AUC) as evaluation metrics. All experiments were conducted on Intel(R) Core(TM) i7-9750H CPU @ 2.60GHz by MATLAB R2018b.

3.2 Parameter sensitivity

DDNMFNN model consists of balance parameters λ_L, λ_G and λ. To explore the impact on the performance of the proposed model under different values of the parameters, we calculated the AUPR value under different parameters and set λ_L, λ_G and λ with the value range of $\{0.01, 0.1, 1, 10, 100\}$,

as shown in Figure 2. The specific operation is to fix the value of one variable and change the value of the other two variables. For example, Figure 2(a) displays the effect of parameters λ_G and λ on Fdataset when parameter λ_L is fixed. When the value of λ is adjusted, the value of AUPR shows significant changes and similar patterns of change occur on the three datasets. The AUPR value increases as the value of λ changes from 0.01 to 0.1, and markedly decreases when λ varies from 0.1 to 100. Final, we recommend setting $\lambda_L \in [0.1, 1]$, $\lambda_G \in [0.01, 1]$ and $\lambda \in [0.1, 1]$.

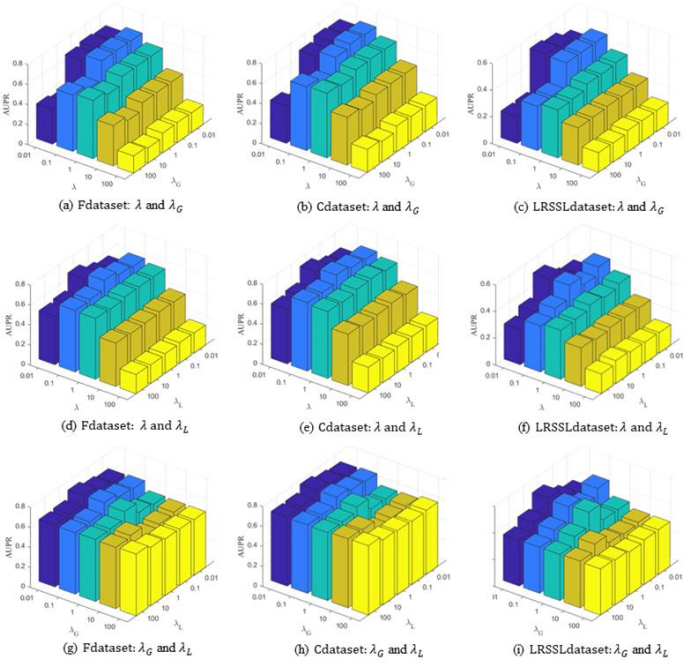

Figure 2. Parameter sensitivity on three datasets.

3.3 Convergence study of DMGNMF

According to the discussion in Section 2.2, the DMGNMF algorithm is solved through an alternating iterative scheme, and its convergence is theoretically guaranteed. To determine the number of iterations N, we plotted convergence curves on three benchmark data, as shown in Figure 3. It can

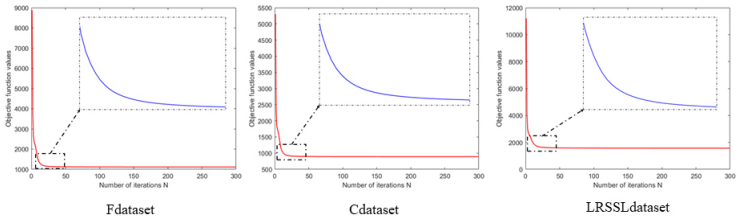

Figure 3. Convergence curves of the DMGNMF algorithm on three datasets.

be seen that N is less than 50 at convergence. The proposed model can be suitable for larger datasets.

3.4 Experimental setup of MKNN

As is well known, the number of layers in a neural network can affect its expressive power. Therefore, we conducted experiments to determine the structure of multi-kernel neural networks. Firstly, we tested the performance of different basic kernels applied to drug-disease datasets. The base kernel functions in our study are linear kernel (K_{lin}), quadratic polynomial kernel (K_{poly2}), cubic polynomial kernel (K_{poly3}), and radial basis function kernel (K_{RBF}). Table 3 displays the AUPR and AUC values for each kernel. We can observe that the difference between the maximum and minimum values is lower than 0.02. So, according to the order of AUPR from high to low, we generate four different node types, namely N_1(the model contains K_{poly2}), N_2(the model contains K_{poly2} and K_{poly3}), N_3(the model contains K_{poly2}, K_{poly3} and K_{lin}), N_4 (the model contains K_{poly2}, K_{poly3}, K_{lin} and K_{RBF}). We investigated the performance of neural networks under different node types and different layers on benchmark datasets, as shown in Table 4. The number of layers for MKNN is set to 2, and the node type is set to N_4 in this paper.

3.5 Ablation studies

DDNMFNN consists of the following two key components to enhance the drug-disease association identification: (1) DMGNMF algorithm adap-

Table 3. Performance of four base kernels.

Kernels	AUPR	AUC
K_{lin}	0.6827	0.9466
K_{poly2}	0.6927	0.9546
K_{poly3}	0.6926	0.9525
K_{RBF}	0.6542	0.9364

Table 4. Summary of similarity construction methods of drug and disease.

Layers	Node Types	Fdatase		Cdatase		LRSSLdataset	
		AUPR	AUC	AUPR	AUC	AUPR	AUC
1	N_1	0.6927	0.9546	0.7364	0.9629	0.4822	0.9120
	N_2	0.6937	0.9538	0.7377	0.9620	0.4744	0.9074
	N_3	0.6943	0.9527	0.7400	0.9615	0.4897	0.9061
	N_4	0.6960	0.9531	0.7388	0.9623	0.4843	0.9072
2	N_1	0.6747	0.9548	0.7192	0.9643	0.4309	0.9168
	N_2	0.6899	0.9546	0.7326	0.9634	0.4521	0.9121
	N_3	0.6952	0.9540	0.7385	0.9626	0.4760	0.9096
	N_4	0.6919	0.9532	0.7350	0.9622	0.4604	0.9052

tively integrates various geometric information of diseases and drugs to alleviate the sparsity problem of the original association matrix; (2) MKNN can map the features of diseases and drugs into the kernel space in a multi-layer multi-kernel learning framework and output prediction results. In this section, we designed an ablation study to investigate the contribution of DMGNMF and MKNN. We input three benchmark datasets into DMGNMF, MKNN, DDNMFNN and the experimental results under 10 times 10-CV test were displayed in Figure 4. We can observe that DDNMFNN achieved the best predictive performance on three datasets. This indicates that the combination of DMGNMF and MKNN is an effective strategy.

To further research the effectiveness of DMGNMF algorithm, we compared it with four representative NMF-based algorithms: NMF [38], GRNMF [21], CWNMF [44], and RNMF [45]. CWNMF is a weighted NMF algorithm with collaborative regularization terms and regularization terms. RNMF is an algorithm that integrates the joint regularized NMF and the symmetric NMF. We combined them with MKNN to obtain

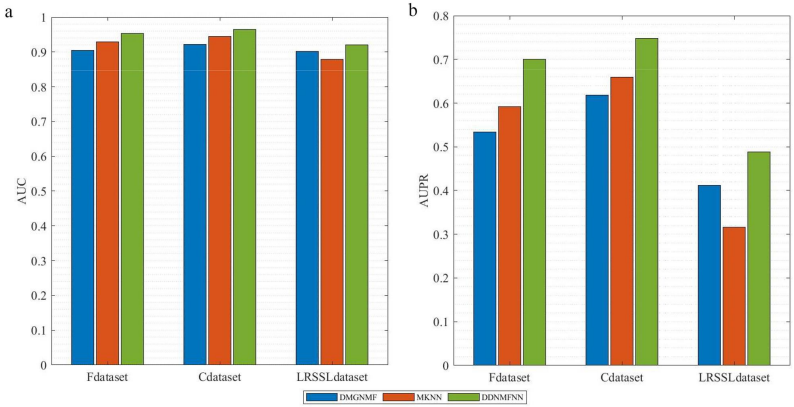

Figure 4. The results of ablation experiment on three benchmark datasets. (a) The AUC values of 10-CV test. (b) The AUPR values of 10-CV test.

four models (NMF-MKNN, GRNMF-MKNN, CWNMF-MKNN, RNMF-MKNN). To ensure fairness, the regularization coefficients of all matrix factorization algorithms are randomly set to 1. The experimental results of all models under 10 times 10-CV test were shown in Figure 5. It can be observed that a meaningful phenomenon is that the predictive performance of models with regularization constraints has significantly improved. This completely reveals that regularization constraints are of considerable significance. In addition, our proposed model achieved optimal predictive performance and significantly improved the AUPR value on all three datasets. This sufficiently indicates the effectiveness of DMGNMF algorithm.

3.6 Comparison with other prediction models

In this section, we compared the state-of-the-art models with DDNMFNN. To ensure the objectivity of the comparative experiment, we compared the predictive performance of all models under a 10-CV framework. All models and experimental results are summarized in Table 5, where SCMFDD [7], DRRS [8], DisDrugPred [9], DRIMC [10], LAGCN [16], MVGCN [17], SMGCL [18], DRWBNCF [19], and MGRMF [22] have already introduced

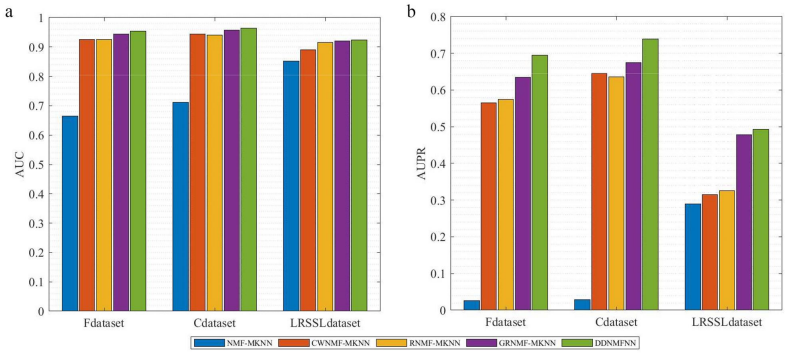

Figure 5. The experimental results of five NMF-based models combined with MKNN on three benchmark datasets. (a) The AUC values of 10-CV test. (b) The AUPR values of 10-CV test.

in the Introduction. To comprehensively compare the performance of DDNMFNN, we also compared five models for drug-target interaction prediction based on matrix factorization, which are KBMF [46], GRGMF [47], NRLMF [48], MGRNNM [49], and RRSSVD [50], and a model for predicting miRNA-disease association applying graph convolutional network and matrix completion, namely NIMCGCN [51]. In addition, two representative models, MBiRW [31] and BNNR [52], were employed for comparison. MBiRW is a random walk algorithm applied to similarity networks. BNNR utilized the bounded kernel norm to complete association matrix.

From Table 5, we can see that DDNMFNN ranks first in AUPR values on three datasets, and has significantly enhanced prediction performance compared to other models. Specifically, the AUPR value of DDNMFNN achieved 0.7 on the Fdataset, which is 14.1% higher than the highest value of other models. On the Cdataset, the highest AUPR value obtained by MGRMF on the Fdataset is 0.661, while our model is 0.749 and improves by 8.8%. Similar results occurred on the LRSSLdataset. The AUPR value of our model is 0.488, which is 4% higher than MGRMF. However, our model did not achieve the highest AUC value. The AUC of DDNMFNN on three datasets is 0.954, 0.965, and 0.920, which are 0.2%, 0.3% and 3.4% lower than DRIMC. Meanwhile, although DRIMC achieved the high-

est AUC, its AUPR did not have an advantage compared to other models. For imbalanced datasets, AUPR can objectively reflect the performance of the model. These experimental results demonstrate the proposed model is effective and reliable.

Table 5. Comparison of performance between DDNMFNN and other models.

Models	AUC			AUPR		
	Fdataset	Cataset	LRSSL dataset	Fdataset	Cdataset	LRSSL dataset
DisDrugPred	0.890	0.908	0.921	0.070	0.067	0.069
SCMFDD	0.712	0.711	0.761	0.004	0.004	0.004
MBiRW	0.911	0.932	0.920	0.129	0.199	0.067
DRRS	0.929	0.948	0.899	0.140	0.216	0.051
KBMF	0.862	0.860	0.759	0.164	0.219	0.060
NRLMF	0.935	0.947	0.913	0.226	0.289	0.163
DRIMC	0.956	0.968	0.954	0.290	0.377	0.161
LAGCN	0.883	0.920	0.935	0.130	0.191	0.114
DRWBNCF	0.926	0.941	0.936	0.491	0.566	0.349
MGRNNM	0.897	0.909	0.860	0.540	0.625	0.440
RRSSVD	0.928	0.946	0.930	0.484	0.555	0.357
MGRMF	0.917	0.939	0.849	0.559	0.661	0.448
BNNR	0.930	0.934	0.927	0.437	0.470	0.315
GRGMF	0.805	——	0.816	0.550	——	0.440
NIMCGCN	0.828	0.851	0.829	0.339	0.433	0.267
MVGCN	0.853	0.862	0.849	0.558	0.630	0.443
SMGCL	0.935	0.947	0.926	0.549	0.626	0.390
SMGCL-NS	0.928	0.937	0.914	0.524	0.582	0.437
DDNMFNN	0.954	0.965	0.920	0.700	0.749	0.488

3.7 Performance on five real world datasets

To further evaluate the predictive performance of DDNMFNN, we applied it to five real-world network datasets: (i) G-protein coupled receptors (GPC): the biological network of drugs binding GPC; (ii) Enzymes: the biological network of drugs binding enzyme proteins; (iii) Ion channels: the network of drugs binding ion channel proteins; (iv) Drug–target: the chemical network of drug–target association; (v) Southern Women (SW): the social relations network of events and women. Because the above datasets without prior similarity information are only interaction matrices, we apply linear kernels, GIP similarity, and polynomial kernels to construct the similarity matrix as model inputs.

To objectively evaluate the predictive performance of DDNMFNN, we compared it with SRNMF [53] and MGRMF [22]. Based on different sim-

ilarity measures, five different SRNMF-baesd models were obtained for comparison. The same data was input into all models to ensure fairness in the experiment. Table 6 displays the AUC values of different models on five real-world networks. The proposed model attained the best predictive results on five datasets. This strongly indicates that DDNMFNN has excellent predictive performance in real-world association networks.

Table 6. The AUC values of DDNMFNN and other compared model on five real-world datasets.

Models	GPC	Enzymes	Ion channel	Drug-target	SW
DDNMFNN	0.91	0.97	0.98	0.96	0.83
MGRMF	0.88	0.92	0.95	0.96	0.82
SRNMF-CN	0.84	0.88	0.94	0.93	0.83
SRNMF-AA	0.83	0.88	0.94	0.93	0.82
SRNMF-CAA	0.83	0.87	0.94	0.92	0.82
SRNMF-CJC	0.83	0.87	0.95	0.93	0.80
SRNMF-JC	0.83	0.88	0.93	0.92	0.85

3.8 Case studies

To verify the performance of the proposed model to infer disease-related drugs without any known associated drugs, we conducted case studies about breast cancer, lung cancer, and Alzheimer's disease on the Fdataset. Specifically, for a specific disease, we removed all known drugs associated with it from the dataset. Then, the DDNMFNN was trained using the remaining known associations and candidate drugs were tested to find drugs related to the disease. The purpose of the above operation was to ensure independence between the train set and the validation set. Finally, we ranked the drugs in descending order of predicted values and validated the top 10 drugs using the CTD database [54].

Alzheimer's disease (AD) is a common neurodegenerative disease that seriously threatens human health. At present, there are at least 50 million dementia patients worldwide and about 60-70% of them have AD [55, 56]. Its etiology is still unclear. Table 7 shows the prediction results of AD, and only Modafinil and Tetrabenazine among the top 10 drugs did not be validated in the CTD database. Haloperidol possesses a good therapeu-

tic effect in a dose of 2-3 milligrams per day [57]. Modafinil is a stimulant that partly works by inhibiting the reuptake of the neurotransmitter norepinephrine. Related studies have shown that reusing established noradrenergic drugs is most likely to provide effective treatment for general cognition and apathy in AD [58].

Table 7. Prediction results of the top 10 associated drugs with Alzheimer's disease.

Index	DrugBank ID	Drug name	Confirmed
1	DB01219	Dantrolene	Y
2	DB00502	Haloperidol	Y
3	DB00163	Vitamin E	Y
4	DB00745	Modafinil	N
5	DB00989	Rivastigmine	Y
6	DB00393	Nimodipine	Y
7	DB00324	Fluorometholone	Y
8	DB00313	Valproic acid	Y
9	DB00181	Baclofen	Y
10	DB04844	Tetrabenazine	N

Breast cancer is one of the most common malignant tumors and the leading cause of cancer death in women worldwide. According to statistics, there were 2.1 million new breast cancer patients and 627000 deaths worldwide in 2018 [59]. Table 8 lists the 10 drugs highly related to breast cancer predicted by the model. Comparing the predicted results to the CTD database, the top 10 drugs predicted were found to be associated with breast cancer. Al-Tweigeri et al. [60] treated 59 patients with locally advanced cancer with doxorubicin followed by docetaxel/cisplatin. The experimental results indicate that combination therapy is feasible, safe, and effective. The literature published by Lu et al. [61] suggested that a combination of Cisplatin, capecitabine, and docetaxel can be safely administered without prophylactic G-CSF, and may be an effective neoadjuvant in patients diagnosed with locally advanced breast cancer.

According to the 2023 US cancer statistics published in the journal CA-A Cancer Journal for Clinicians, approximately 350 people die from lung cancer every day, almost 2.5 times the number of deaths from the second deadliest cancer [62]. The top 10 drugs predicted by the model have been identified in the CTD database, as shown in Table 9. Cisplatin-etoposide

Table 8. Prediction results of the top 10 associated drugs with breast cancer.

Index	DrugBank ID	Drug name	Confirmed
1	DB00997	Doxorubicin	Y
2	DB00515	Cisplatin	Y
3	DB01196	Estramustine	Y
4	DB00286	Conjugated estrogens	Y
5	DB00783	Estradiol	Y
6	DB01101	Capecitabine	Y
7	DB01248	Docetaxel	Y
8	DB00007	Leuprolide	Y
9	DB00014	Goserelin	Y
10	DB00499	Flutamide	Y

combined chemotherapy was developed in the early 1980s for the treatment of non-small cell lung cancer [63]. Until now, the first-line treatment for small cell lung cancer has remained etoposide combined with cisplatin chemotherapy [64]. Three case studies demonstrate that DDNMFNN can be a promising tool to exploit potential drug-disease associations.

Table 9. Prediction results of the top 10 associated drugs with lung cancer.

Index	DrugBank ID	Drug name	Confirmed
1	DB00563	Methotrexate	Y
2	DB00515	Cisplatin	Y
3	DB00773	Etoposide	Y
4	DB00398	Sorafenib	Y
5	DB01030	Topotecan	Y
6	DB00997	Doxorubicin	Y
7	DB01005	Hydroxyurea	Y
8	DB00444	Teniposide	Y
9	DB00441	Gemcitabine	Y
10	DB01101	Capecitabine	Y

4 Conclusions

In this paper, we introduced a novel computational model, namely DDNMFNN, which can effectively predict drug-disease associations. To fully extract prior knowledge related to association information, we generated multiple similarity matrices for drugs and diseases to feed into DMGNMF

and MKNN. Different from traditional NMF algorithms, our key innovation lies in incorporating dual multi-graph regularization. It not only compensates for the shortcomings of NMF in discovering the inherent geometric and discriminative structures of the data by modeling multi-view similarity information of drugs and diseases, but also enables adaptive learning of the weights of each graph by iterations. We applied L_2 regularization term to the low-rank in the proposed model to improve the effectiveness of the node representations. In addition, we selected MKNN for association prediction, combining the strengths of deep learning and NMF to capture nonlinear structures in sparse association data.

To evaluate the predictive ability of our model, we conducted a considerable number of experiments. Compared with existing computational models, DDNMFNN demonstrated superior predictive performance on three benchmark datasets under 10-CV. Furthermore, the model was applied on five real world datasets to verify its robustness. The case studies of three diseases further indicated that the proposed model can effectively predict unknown disease-related drugs. In brief, the excellent performance of DDNMFNN in the above experiments exhibited that it is a promising tool for drug repositioning.

***Acknowledgment*:** This work was supported by the National Natural Science Foundation of China under Grant (No. 61877064, U1806202 and 61533011) and project ZR2024QF189 supported by Shandong Provincial Natural Science Foundation.

References

[1] V. Subbiah, The next generation of evidence-based medicine, *Nat. Med.* **29** (2023) 49–58.

[2] S. Pushpakom, F. Iorio, P. A. Eyers, K. J. Escott, S. Hopper, A. Wells, A. Doig, T. Guilliams, J. Latimer, C. McNamee, A. Norris, P. Sanseau, D. Cavalla, M. Pirmohamed, Drug repurposing: Progress, challenges and recommendations, *Nat. Rev. Drug Disc.* **18** (2018) 41–58.

[3] S. S. Sadeghi, M. R. Keyvanpour, An analytical review of computational drug repurposing, *IEEE/ACM Trans. Comput. Biol. Bioinf.* **18** (2021) 472–488.

[4] X. Pan, X. Lin, D. Cao, X. Zeng, P. S. Yu, L. He, R. Nussinov, F. Cheng, Deep learning for drug repurposing: Methods, databases, and applications, *WIRES* **12** (2022) #e1597.

[5] A. Gottlieb, G. Y. Stein, E. Ruppin, R. Sharan, PREDICT: A method for inferring novel drug indications with application to personalized medicine, *Mol. Syst. Biol.* **7** (2011) 1–9.

[6] M. Oh, J. Ahn, Y. Yoon, A network-based classification model for deriving novel drug-disease associations and assessing their molecular actions, *PLoS One* **9** (2014) 1–12.

[7] 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 Bioinf.* **19** (2018) 1–12.

[8] H. Luo, M. Li, S. Wang, Q. Liu, Y. Li, J. Wang, Computational drug repositioning using low-rank matrix approximation and randomized algorithms, *Bioinf.* **34** (2018) 1904–1912.

[9] P. Xuan, Y. Cao, T. Zhang, X. Wang, S. Pan, T. Shen, Drug repositioning through integration of prior knowledge and projections of drugs and diseases, *Bioinf.* **35** (2019) 4108–4119.

[10] W. Zhang, H. Xu, X. Li, Q. Gao, L. Wang, DRIMC: An improved drug repositioning approach using Bayesian inductive matrix completion, *Bioinf.* **36** (2020) 2839–2847.

[11] Q. Zhang, L. T. Yang, Z. Chen, P. Li, A survey on deep learning for big data, *Inf. Fus.* **42** (2018) 146–157.

[12] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M. L. Shyu, S. C. Chen, S. S. Iyengar, A survey on deep learning: Algorithms, techniques, and applications, *ACM Comput. Surv.* **51** (2018) 1-36.

[13] X. Zeng, S. Zhu, X. Liu, Y. Zhou, R. Nussinov, F. Cheng, DeepDR: A network-based deep learning approach to in silico drug repositioning, *Bioinf.* **35** (2019) 5191–5198.

[14] J. Peng, Y. Wang, J. Guan, J. Li, R. Han, J. Hao, Z. Wei, X. Shang, An end-to-end heterogeneous graph representation learning-based framework for drug-target interaction prediction, *Brief. Bioinf.* **25** (2021) 1–9.

[15] Y. Chen, T. Ma, X. Yang, J. Wang, B. Song, X. Zeng, MUFFIN: Multi-scale feature fusion for drug-drug interaction prediction, *Bioinf.* **37** (2021) 2651–265.

[16] Z. Yu, F. Huang, X. Zhao, W. Xiao, W. Zhang, Predicting drug-disease associations through layer attention graph convolutional network, *Brief. Bioinf.* **22** (2021) 1–11.

[17] H. Fu, F. Huang, X. Liu, Y. Qiu, W. Zhang, MVGCN: data integration through multi-view graph convolutional network for predicting links in biomedical bipartite networks, *Bioinf.* **38** (2022) 426–434.

[18] Z. Gao, H. Ma, X. Zhang, Y. Wang, Z. Wu, Similarity measures-based graph co-contrastive learning for drug – disease association prediction, *Bioinf.* **39** (2023) 1–8.

[19] Y. Meng, C. Lu, M. Jin, J. Xu, X. Zeng, J. Yang, A weighted bilinear neural collaborative filtering approach for drug repositioning, *Brief. Bioinf.* **23** (2022) 1–13.

[20] B. Geng, D. Tao, C. Xu, L. Yang, X. S. Hua, Ensemble manifold regularization, *IEEE Trans. Pattern Anal. Mach. Intell.* **34** (2012) 1227–1233.

[21] D. Cai, X. He, J. Han, T. S. Huang, Graph regularized nonnegative matrix factorization for data representation, *IEEE Trans. Pattern Anal. Mach. Intell.* **33** (2011) 1548–1560.

[22] C. Ai, H. Yang, Y. Ding, J. Tang, F. Guo, Low rank matrix factorization algorithm based on multi-graph regularization for detecting drug-disease association, *IEEE/ACM Trans. Comput. Biol. Bioinf.* **20** (2021) 3033–3043.

[23] J. Li, Y. Liu, H. Lin, Y. Yue, W. Wang, Efficient kernel selection via spectral analysis, in: C. Sierra (Ed.), *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, Melbourne, 2017, pp. 2124–2130.

[24] M. Gönen, E. Alpaydin, Multiple kernel learning algorithms, *J. Mach. Learn. Res.* **12** (2011) 2211–2268.

[25] S. S. Bucak, R. Jin, A. K. Jain, Multiple kernel learning for visual object recognition: A review, *IEEE Trans. Pattern Anal. Mach. Intell.* **36** (2014) 1354–1369.

[26] M. R. Mohammadnia–Qaraei, R. Monsefi, K. Ghiasi–Shirazi, Convolutional kernel networks based on a convex combination of cosine kernels, *Pattern Recognit. Lett.* **116** (2018) 127–134.

[27] A. L. Afzal, S. Asharaf, Deep kernel learning in core vector machines, *Pattern Anal. Appl.* **21** (2018) 721–729.

[28] A. L. Afzal, S. Asharaf, Deep multiple multilayer kernel learning in core vector machines, *Expert Syst. Appl.* **96** (2018) 149–156.

[29] D. S. Wishart, C. Knox, A. C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B. Gautam, M. Hassanali, DrugBank: A knowledgebase for drugs, drug actions and drug targets, *Nucleic Acids Res.* **36** (2008) D901–D906.

[30] A. Hamosh, A. F. Scott, J. S. Amberger, C. A. Bocchini, V. A. McKusick, Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders, *Nucleic Acids Res.* **33** (2005) D514–D517.

[31] H. Luo, J. Wang, M. Li, J. Luo, X. Peng, F. X. Wu, Y. Pan, Drug repositioning based on comprehensive similarity measures and bi-random walk algorithm, *Bioinf.* **32** (2016) 2664–2671.

[32] X. Liang, P. Zhang, L. Yan, Y. Fu, F. Peng, L. Qu, M. Shao, Y. Chen, Z. Chen, LRSSL: Predict and interpret drug-disease associations based on data integration using sparse subspace learning, *Bioinf.* **33** (2017) 1187–1196.

[33] I. Dhammi, S. Kumar, Medical subject headings (MeSH), *Indian J. Orthop.* **88** (2000) 265–266.

[34] T. van Laarhoven, S. B. Nabuurs, E. Marchiori, Gaussian interaction profile kernels for predicting drug-target interaction, *Bioinf.* **27** (2011) 3036–3043.

[35] D. Vidal, M. Thormann, M. Pons, LINGO, an efficient holographic text based method to calculate biophysical properties and intermolecular similarities, *J. Chem. Inf. Model.* **45** (2005) 386–393.

[36] C. Steinbeck, Y. Han, S. Kuhn, O. Horlacher, E. Luttmann, E. Willighagen, The Chemistry Development Kit (CDK): An open-source Java library for chemo- and bioinformatics, *J. Chem. Inf. Comput. Sci.* **43** (2003) 493–500.

[37] M. A. van Driel, J. Bruggeman, G. Vriend, H. G. Brunner, J. A. M. Leunissen, A text-mining analysis of the human phenome, *Eur. J. Hum. Genet.* **14** (2006) 535–542.

[38] L. Daniel D., S. H. Sebastian, Learning the parts of objects by non-negative matrix factorization, *Nature* **401** (1999) 788–791.

[39] D. Cai, X. He, X. Wu, J. Han, Non-negative matrix factorization on manifold, *IEEE Xplore* (2008) 63–72.

[40] D. A. Spielman, Spectral graph theory, in: U. Naumann, O. Schenk (Eds.), *Combinatorial Scientific Computing*, CRC Press, Boca Raton, 2012, pp. 1–29.

[41] D. Cai, X. He, Y. Hu, J. Han, T. Huang, Learning a spatially smooth subspace for face recognition, in: *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.*, 2007, pp. 1–7.

[42] F. Facchinei, C. Kanzow, S. Sagratella, Solving quasi-variational inequalities via their KKT conditions, *Math. Program.* **144** (2014) 369–412.

[43] C. Ai, H. Yang, Y. Ding, J. Tang, F. Guo, A multi-layer multi-kernel neural network for determining associations between non-coding RNAs and diseases, *Neurocomputing* **493** (2022) 91–105.

[44] D. Xu, H. Xu, Y. Zhang, R. Gao, Novel Collaborative weighted non-negative matrix factorization improves prediction of disease-associated human microbes, *Front. Microbiol.* **13** (2022) 1–12.

[45] D. Xu, H. Xu, Y. Zhang, W. Chen, R. Gao, LncRNA-protein interaction prediction based on regularized nonnegative matrix factorization and sequence information, *MATCH Commun. Math. Comput. Chem.* **85** (2021) 555–574.

[46] M. Gonen, S. Kaski, Kernelized Bayesian matrix factorization, *IEEE Trans. Pattern Anal. Mach. Intell.* **36** (2014) 2047–2060.

[47] Z. C. Zhang, X. F. Zhang, M. Wu, L. Ou-Yang, X. M. Zhao, X. L. Li, A graph regularized generalized matrix factorization model for predicting links in biomedical bipartite networks, *Bioinf.* **36** (2020) 3474–3481.

[48] Y. Liu, M. Wu, C. Miao, P. Zhao, X. L. Li, Neighborhood regularized logistic matrix factorization for drug-target interaction prediction, *PLoS Comput. Biol.* **12** (2016) 1–26.

[49] A. Mongia, A. Majumdar, Drug-target interaction prediction using multi graph regularized nuclear norm minimization, *PLoS One* **15** (2020) 1–19.

[50] A. G. Sorkhi, Z. Abbasi, M. I. Mobarakeh, J. Pirgazi, Drug–target interaction prediction using unifying of graph regularized nuclear norm with bilinear factorization, *BMC Bioinf.* **22** (2021) 1–23.

[51] J. Li, S. Zhang, T. Liu, C. Ning, Z. Zhang, W. Zhou, Neural inductive matrix completion with graph convolutional networks for miRNA-disease association prediction, *Bioinf.* **36** (2020) 2538–2546.

[52] M. Yang, H. Luo, Y. Li, J. Wang, Drug repositioning based on bounded nuclear norm regularization, *Bioinf.* **35** (2019) i455–i463.

[53] W. Wang, X. Chen, P. Jiao, D. Jin, Similarity-based regularized latent feature model for link prediction in bipartite networks, *Sci. Rep.* **7** (2017) 1–12.

[54] A. P. Davis, T. C. Wiegers, J. Wiegers, B. Wyatt, R. J. Johnson, D. Sciaky, F. Barkalow, M. Strong, A. Planchart, C. J. Mattingly, CTD tetramers: a new online tool that computationally links curated chemicals, genes, phenotypes, and diseases to inform molecular mechanisms for environmental health, *Toxicol. Sci.* **195** (2023) 155–168.

[55] S. S. Kang, L. Meng, X. Zhang, Z. Wu, A. Mancieri, B. Xie, X. Liu, D. Weinshenker, J. Peng, Z. Zhang, K. Ye, Tau modification by the norepinephrine metabolite DOPEGAL stimulates its pathology and propagation, *Nat. Struct. Mol. Biol.* **29** (2022) 292–305.

[56] Christina Patterson, World Alzheimer report 2018 - The state of the art of dementia research: New frontiers, *Alzheimer's Dis. Int.* (2018) 1–48.

[57] Devanand, K. Marder, K. S. Michaels, H. A. Sackeim, K. Bell, M. A. Sullivan, T. B. Cooper, G. H. Pelton, R. Mayeux., A Randomized, placebo-controlled dose-comparison trial of haloperidol for psychosis and disruptive behaviors in Alzheimer's disease, *Am. J. Psychiatr.* **155** (1998) 1512–1520.

[58] M. C. B. David, M. Del Giovane, K. Y. Liu, B. Gostick, J. B. Rowe, I. Oboh, R. Howard, P. A. Malhotra, Cognitive and neuropsychiatric effects of noradrenergic treatment in Alzheimer's disease: systematic review and meta-analysis, *J. Neurol. Neurosurg. Psychiatry* **93** (2022) 1080–1090.

[59] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA. Cancer J. Clin.* **68** (2018) 394–424.

[60] T. A. Al-Tweigeri, D. S. Ajarim, A. A. Alsayed, M. M. Rahal, M. O. Alshabanah, A. M. Tulbah, O. A. Al-Malik, D. M. Fatani, G. A. El-Husseiny, N. B. Elkum, A. A. Ezzat, Prospective phase II study of neoadjuvant doxorubicin followed by cisplatin/docetaxel in locally advanced breast cancer, *Med. Oncol.* **27** (2010) 571–577.

[61] Y. S. Lu, D. R. Chen, L. M. Tseng, D. C. Yeh, S. T. Chen, C. M. Hsieh, H. C. Wang, H. T. Yeh, S. H. Kuo, C. S. Huang, Phase II study of docetaxel, capecitabine, and cisplatin as neoadjuvant chemotherapy for locally advanced breast cancer, *Cancer Chemother. Pharmacol.* **67** (2011) 1257–1263.

[62] R. L. Siegel, K. D. Miller, N. S. Wagle, A. Jemal, Cancer statistics 2023, *CA. Cancer J. Clin.* **73** (2023) 17–48.

[63] A. Ardizzoni, G. Antonelli, F. Grossi, L. Tixi, M. Cafferata, R. Rosso, The combination of etoposide and cisplatin in non-small-cell lung cancer (NSCLC), *Ann. Oncol.* **10** (1999) S13–S17.

[64] J. Chen, X. Pan, F. Na, X. Chen, C. Chen, Epigenetic reprogramming in small cell lung cancer, *Cancer Biol. Med.* **19** (2022) 1111–1116.