

PAPER

ConvNTC: Convolutional neural tensor completion for predicting the disease-related miRNA pairs and cell-related drug pairs

Pei Liu,^{1,2} Xiao Liang,¹ Yue Li^{2,*} and Jiawei Luo^{1,*}

¹College of Computer Science and Electronic Engineering, Hunan University, Changsha, 410083, Hunan, China and ²School of Computer Science, McGill University, Montreal, H3A 0C6, Quebec, Canada

*Corresponding author. yueli@cs.mcgill.ca and luojiawei@hnu.edu.cn

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Systematic investigation of high-order molecular interactions can deepen our understanding of the mechanisms underlying biological systems. However, effectively capturing both multilinear and nonlinear relationships to accurately identify the complex triplet relationships remains a challenge. In this paper, we present a novel **Convolutinal Neural Tensor Completion** (ConvNTC) to model triplet network connections. ConvNTC consists of a multilinear module and nonlinear module. The former is a tensor decomposition approach that integrates multiple constraints to learn the tensor factor embeddings. The latter contains three components: an embedding generator to produce position-specific index embeddings for each tensor entry in addition to the factor embeddings, a convolutional encoder to perform nonlinear feature mapping while preserving the tensor's rank-one property, and a Kolmogorov-Arnold Network (KAN) based predictor to effectively capture high-dimensional relationships aligned with the intrinsic structure of real-world data. We demonstrate ConvNTC on two triplet prediction tasks of the disease-related miRNA pairs and cell-related drug pairs, respectively. Comprehensive experiments against ten state-of-the-art methods demonstrate the superiority of ConvNTC in terms of triplet imputation. ConvNTC reveals promising prognostic values of the miRNA-miRNA interactions on breast cancer and detects synergistic drug combinations in cancer cell lines. The source code is available at <https://github.com/Liangyushi/ConvNTC>.

Key words: synergistic miRNA pairs, synergistic drug combinations, convolutional neural network, tensor completion, Kolmogorov-Arnold Network (KAN).

Introduction

Investigations of synergistic states between biomolecules aim to reveal the interactions and regulatory mechanisms among various molecules within biological systems, which contribute to gaining a deeper insight into cellular functions, signal transduction and metabolic pathways. Among them, the exploration of synergy between miRNAs related to diseases primarily investigates the role of miRNA pairs in disease progression and development, concentrating on indentifying biomarkers at the gene regulation level (Lai et al., 2019; Xu et al., 2011). In contrast, the discovery of drugs combinations in diverse cellular/disease contexts examines the role of drug pairs in enhancing treatment efficacy, focusing on clinical applications and therapeutic outcomes (Cokol et al., 2011). Translating the identification of disease-related miRNA pairs and disease/cell line-associated drug pairs into the prediction task of miRNA-miRNA-disease and drug-drug-cell/disease triple relationships is an effective strategy.

Consequently, by reformulating the prediction of miRNA-miRNA-disease and drug-drug-cell/disease triple relationships

as a third-order tensor completion problem, numerous tensor-based computational models have been developed. For instance, Chen and Li (2018) presented a tensor completion method, called DrugCom, for capturing disease-related drug combinations using representation learning to integrate multiple additional information of drugs and diseases. Liu et al. (2020) integrated multi-view information on miRNAs and diseases to identify potential miRNA-miRNA-disease associations through a tensor completion framework known as miRCom. Although these methods are effective in identifying cooperative miRNA/drug pairs, they are limited to capturing only the multilinear structure of data, while real-world data often exhibit more complex properties, such as nonlinearity rather than just multilinearity.

Subsequently, deep learning algorithms were introduced to effectively capture nonlinear relationships. For example, Luo et al. (2021) proposed a graph attention mechanism-based neural tensor factorization framework, named GraphTF, to capture disease-related miRNA pairs. Although GraphTF can capture nonlinear factor embeddings, it fails to fully represent the nonlinear relationships within a tensor due to its reliance

on the Kronecker product for reconstructing the entire tensor. Sun et al. (2020) proposed a deep tensor factorization model, named DTF, which combines a classical tensor decomposition method CP_WOPT with multilayer perceptrons (MLPs) to predict the synergy state of drug pairs. Similarly, Han et al. (2024) proposed a novel constrained tensor factorization (CTF) integrated with a deep neural network, termed CTF-DDI, to identify interactions between drugs across different types. However, both DTF and CTF-DDI tend to overlook the low-rank structure of the tensor and may overfit to sparse training data, as they simply feed the learned low-rank factors from tensor decomposition into a MLP.

Recently, KAN (Liu et al., 2024) integrates Kolmogorov-Arnold mathematical theory with neural network architectures to improve the capacity and interpretability of neural networks. Unlike traditional MLP, KAN replaces linear weight matrices with learnable B-spline functions, aiming to efficiently learn and approximate complex high-dimensional functions by decomposing them into a series of simpler functions. KAN and its various variants have been widely applied across diverse scientific fields (Drokin, 2024; Li et al., 2024). However, to our best knowledge, KAN has not been effectively utilized in tensor decomposition.

In this study, we propose a hybrid deep tensor completion model, termed ConvNTC, which collaboratively captures both multilinear and nonlinear relationships to identify disease-related miRNA pairs and cell line-related drug pairs. Given the intricate and high-order dependencies inherent within a three-order tensor, two separate modules are designed for learning multilinear and nonlinear relationships, respectively. In the multilinear module, we incorporate similarity constraints, Hessian regularization and $L_{2,1}$ regularization into a traditional tensor decomposition framework to restrict the factors learning. In contrast, the nonlinear module is designed to capture latent high-order nonlinear interactions and enhance predictive performance, which combines convolutional neural networks (CNNs) with FastKAN (a variants of KAN). This module first employs an embedding generator to obtain initial embeddings for various entities, which are subsequently processed by a convolutional encoder and a FastKAN predictor, yielding predicting scores. Extensive comparative experiments against ten benchmark methods demonstrate that ConvNTC exhibits notable superiority, robustness and scalability. Additionally, case studies further illustrate the effectiveness of ConvNTC in identifying novel disease-related miRNA pairs and cell-related drug combinations.

Materials and methods

ConvNTC contains three parts: dataset module, multilinear relationship learning module and nonlinear relationship learning module. Taking 'disease-related miRNA pairs' as an example, the general outline of ConvNTC is shown in Fig. 1.

Dataset

Dataset1 for predicting the disease-related miRNA pairs, was derived from our previous work (Liu et al., 2020), which contains 14,679 miRNA-miRNA-disease triplet relationships among 351 miRNAs and 325 diseases. Additionally, it includes multi-type miRNA and disease similarity matrices. By applying an averaging operation, we obtained the final miRNA and disease similarity matrix. **Dataset2** for predicting cell-related drug pairs, was collected from the previous study (Preuer et al., 2018), which contains 46,124 drug-drug-cell triples with

continuous synergy scores among 38 drugs and 39 cell lines. It also contains the preprocessed features of drugs and cell lines. Based on these features, we employed cosine similarity method to obtain the similarity matrices of drugs and cells. More details of these two dataset are provided in Supplementary material 3.

Next, we take the prediction task of disease-related miRNA pairs as an example to describe details of other modules.

Problem formulation

Given n_m miRNAs, n_d diseases, and their triple associations (m_i, m_j, d_k) , a third-order incomplete tensor $\mathcal{X} \in \mathbb{R}^{n_m \times n_m \times n_d}$ with the first two symmetric modes can be constructed, where $\mathcal{X}_{i,j,k} = 1$ represents there exists cooperative interaction between miRNA pair (m_i, m_j) and disease n_d , and entry $\mathcal{X}_{i,j,k} = 0$ represents there is no interaction or missing value, where $i, j \in \{1, 2, \dots, n_m\}, k \in \{1, 2, \dots, n_d\}$ are the indices of miRNA and disease respectively. Based on this third-order tensor $\mathcal{X}$ with partially observed interactions between diseases and miRNA pairs, our purpose is to predict whether there exists synergy status of potential miRNA pairs related to a specific disease d_i or not, by filling in the missing values of incomplete third-order tensor $\mathcal{X}$.

Multilinear relationship learning

Multi-Constraint Tensor Decomposition (MCTD)

Given a third-order miRNA-miRNA-disease incomplete tensor $\mathcal{X} \in \mathbb{R}^{n_m \times n_m \times n_d}$ with the first two symmetric modes, and two similarity matrices $S_m \in \mathbb{R}^{n_m \times n_m}$ and $S_d \in \mathbb{R}^{n_d \times n_d}$ of miRNAs and diseases, we designed a multi-constraint tensor decomposition model, which incorporates similarity constraint, Hessian regularization and $L_{2,1}$ regularization into the traditional INDSCAL framework (Carroll and Chang, 1970) to enhance its overall efficacy and provide more effective model constraints. Specifically, similarity constraint can bring in additional biological information on miRNA and disease feature factors. Hessian regularization is more closely aligned with the inherent structure of data compared to conventional Laplace regularization, which can mitigate overfitting after dimensionality reduction. The $L_{2,1}$ norm enables to govern the row-wise sparsity, which can prevent overfitting and contribute to a more stable optimization process. The objective function of MCTD is formally defined as follows:

$$\begin{aligned} L_{MCTD} = & \left\| \mathcal{X} - \llbracket M, M, D \rrbracket \right\|_F^2 + \frac{\mu}{2} \left\| S_m - M A_m M^T \right\|_F^2 \\ & + \frac{\alpha}{2} \text{tr} \left(M^T U M \right) + \frac{\eta}{2} \left\| S_d - D A_d D^T \right\|_F^2 + \frac{\beta}{2} \text{tr} \left(D^T V D \right) \\ & + \frac{\lambda}{2} \left(\|M\|_{2,1}^2 + \|D\|_{2,1}^2 + \|A_m\|_F^2 + \|A_d\|_F^2 \right) \end{aligned} \quad (1)$$

where r is the rank of tensor $\mathcal{X}$, $\|\cdot\|_F$ is the Frobenius norm and $\llbracket \cdot \rrbracket$ is the outer product operation used to reconstruct tensor; $M \in \mathbb{R}^{n_m \times r}$, $D \in \mathbb{R}^{n_d \times r}$ are factor matrices for miRNAs and diseases respectively, while $A_m, A_d \in \mathbb{R}^{r \times r}$ are their projection matrices; μ and η are parameters to qualify the contribution of S_m and S_d respectively; $U \in \mathbb{R}^{n_m \times n_m}$, $V \in \mathbb{R}^{n_d \times n_d}$ are the Laplace matrices calculated from S_m and S_d respectively, $\text{tr}(\cdot)$ denotes the Hessian regularization term and α, β are parameters to control it; λ is a parameter to control $L_{2,1}$ regularization.

Optimization strategy of MCTD

We employed the ADMM method as optimization strategy of MCTD. To facilitate the solution of M and D , letting $C = M, Z = M, H = M$ and $F = D, P = D$, the augmented

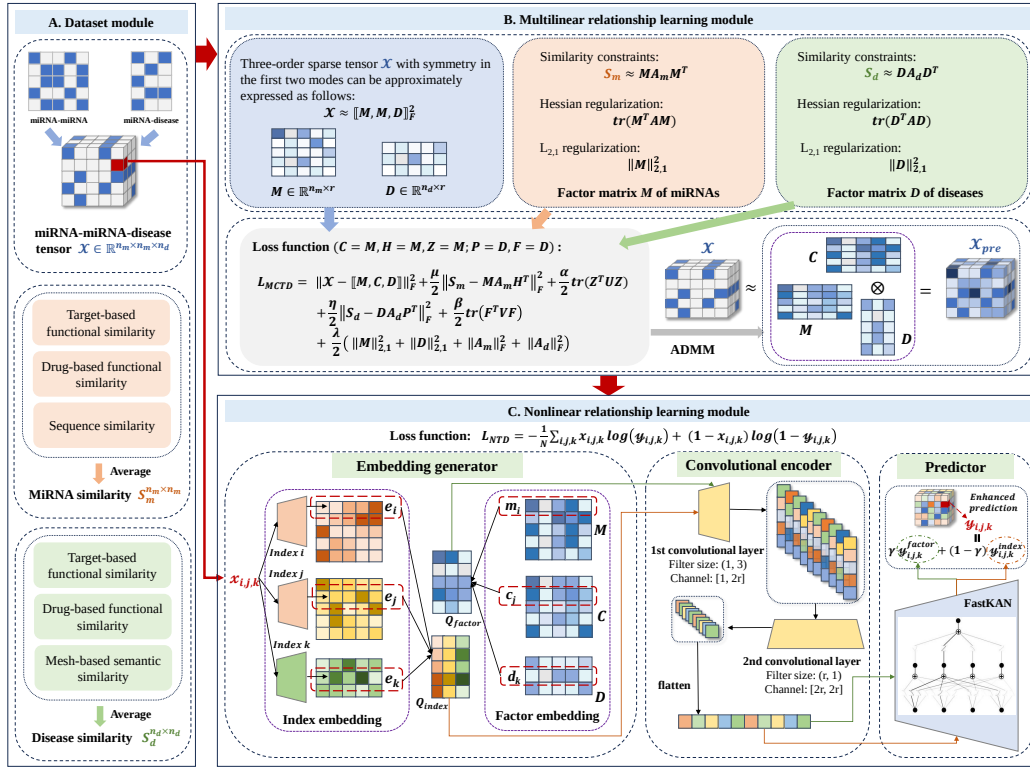


Fig. 1. The general outline of ConvNTC by taking the predicting task of disease-related miRNA pairs as an example. It consists of three modules: **A. Dataset module.** We constructed a third-order tensor $\mathcal{X}$ from miRNA-disease and miRNA-miRNA associations to represent miRNA-miRNA-disease triplets, and obtained miRNA and disease similarity matrices (S_m and S_d) using average operation to integrate multi-type similarities. **B. Multilinear relationship learning module.** We learn the multilinear factor embeddings (M , C and D) with r components across three dimensions of the tensor $\mathcal{X}$ by introducing three constraints into a traditional tensor decomposition. **C. Nonlinear relationship learning module.** An embedding generator first generates the index and factor embeddings of three dimensions for each given entry $\mathcal{X}_{i,j,k}$. Then, a convolutional encoder consisting of two-layer convolutional network is designed to capture the nonlinear features. Finally, a predictor with one-layer FastKAN is adopted to further learn the high-dimensional relationships and obtain the enhanced prediction $\mathcal{Y}_{i,j,k}$.

Lagrangian function of MCTD is constructed as follows:

$$\begin{aligned}
 L_{MCTD} = & \left\| \mathcal{X} - \llbracket M, C, D \rrbracket \right\|_F^2 + \frac{\rho_1}{2} \left\| M - C + \frac{Y_1}{\rho_1} \right\|_F^2 \\
 & + \frac{\mu}{2} \left\| S_m - M A_m H^T \right\|_F^2 + \frac{\theta_1}{2} \left\| M - H + \frac{R_1}{\theta_1} \right\|_F^2 \\
 & + \frac{\alpha}{2} \text{tr}(Z^T U Z) + \frac{\rho_2}{2} \left\| M - Z + \frac{Y_2}{\rho_2} \right\|_F^2 \\
 & + \frac{\eta}{2} \left\| S_d - D A_d P^T \right\|_F^2 + \frac{\theta_2}{2} \left\| D - P + \frac{R_2}{\theta_2} \right\|_F^2 \\
 & + \frac{\beta}{2} \text{tr}(F^T V F) + \frac{\rho_3}{2} \left\| D - F + \frac{Y_3}{\rho_3} \right\|_F^2 \\
 & + \frac{\lambda}{2} \left(\|M\|_{2,1}^2 + \|D\|_{2,1}^2 + \|A_m\|_F^2 + \|A_d\|_F^2 \right) \quad (2)
 \end{aligned}$$

where $Y_1, Y_2, R_1 \in \mathbb{R}^{n_m \times r}$, $Y_3, R_2 \in \mathbb{R}^{n_d \times r}$ are the Lagrange multipliers; $\rho_1, \rho_2, \rho_3, \theta_1, \theta_2$ are the penalty parameters of them. Subsequently, each variable is updated by fixing others until convergence. The updating equations of all variables are provided in Supplementary material 1, and the pseudocode of MCTD is shown in Supplementary algorithm S1. To this end, we can obtain the learned factor matrices M, C, D until the above optimization process converges.

Nonlinear relationship learning

The core of nonlinear relationship learning module is a Neural Tensor Decomposition (NTD) method, which consists of three

components: embedding generator, convolutional encoder and FastKAN predictor. For a third-order incomplete tensor $\mathcal{X} \in \mathbb{R}^{n_m \times n_m \times n_d}$ with the first two symmetric modes, the input of NTD is an index vector $t = [i, j, k]$ and the output is an enhanced prediction score $\mathcal{Y}_{i,j,k}$.

Embedding generator

Given a tensor entry $\mathcal{X}_{i,j,k}$ at index (i, j, k) , the embedding generator takes an index vector $t = [i, j, k]$ as input, and outputs the corresponding index and factor embeddings Q_{index}, Q_{factor} of these indices. The index embedding of each indice is produced by a trainable operator $emb(\cdot)$ as follows:

$$\begin{aligned}
 e_i &= emb_{miRNA}(t[0]) \in \mathbb{R}^{1 \times r}, t[0] = i \\
 e_j &= emb_{miRNA}(t[1]) \in \mathbb{R}^{1 \times r}, t[1] = j \\
 e_k &= emb_{disease}(t[2]) \in \mathbb{R}^{1 \times r}, t[2] = k \\
 Q_{index} &= [[e_i][e_j][e_k]] \in \mathbb{R}^{3 \times r} \\
 Q_{index} &= reshape(Q_{index}) \in \mathbb{R}^{1 \times r \times 3} \quad (3)
 \end{aligned}$$

where $emb(\cdot)$ is defined as `nn.Embedding` in torch, which can map discrete inputs (i.e., miRNA and disease) to a continuous vector. Its factor embedding $Q_{factor} \in \mathbb{R}^{1 \times r \times 3}$ is straightforwardly generated from the learned factors M, C, D .

Convolutional encoder

The convolutional encoder is composed of two 2D convolution layers: $\text{conv}2d_1(\cdot)$ with filter size $(1, n)$ and $\text{conv}2d_2(\cdot)$ with a filter size $(r, 1)$, where $n = 3$ and r are the dimension and rank of association tensor, respectively. Let n_C be the number of channels in the convolutional layers and δ_1, δ_2 are convolution filters of the two convolution layers. It should be emphasized that factor embedding Q_{factor} shares the convolutional encoder with index embedding Q_{index} . Taking Q_{index} as an example, its convolutional coding is:

$$\begin{aligned} Q_{index}^{conv1} &= \sigma_c(\text{conv}2d_1(Q_{index}; \delta_1)) \in \mathbb{R}^{n_C \times r \times 1} \\ Q_{index}^{conv2} &= \sigma_c(\text{conv}2d_2(Q_{index}; \delta_2)) \in \mathbb{R}^{n_C \times 1 \times 1} \\ Q_{index}^{conv2} &= \text{flatten}(Q_{index}^{conv2}) \in \mathbb{R}^{1 \times n_C} \end{aligned} \quad (4)$$

where $\sigma_c = \text{ReLU}(\cdot)$ is nonlinear activation function. Similarly, the convolutional coding of Q_{factor} is $Q_{factor}^{conv2} \in \mathbb{R}^{1 \times n_C}$.

FastKAN predictor

The FastKAN predictor comprises a one-layer of FastKAN (Li, 2024) that utilizes radial basis functions (RBFs) with Gaussian kernels to approximate third-order B-spline bases, thereby significantly accelerating model computations. More explanation are in Supplementary Material 2. Taking index embedding Q_{index}^{conv2} as the input, the output $\mathcal{Y}_{i,j,k}^{index}$ is:

$$\begin{aligned} Q_{index}^{laynorm} &= \text{layernorm}(Q_{index}^{conv2}) \in \mathbb{R}^{1 \times n_C} \\ Q_{index}^{rbf} &= \text{rbf}(Q_{index}^{laynorm}) \in \mathbb{R}^{n_C \times n_g} \\ Q_{index}^{rbf} &= \text{reshape}(Q_{index}^{rbf}) \in \mathbb{R}^{1 \times (n_C \times n_g)} \\ \mathcal{Y}_{i,j,k}^{index} &= w_b b(Q_{index}^{conv2}) + w_s Q_{index}^{rbf} \end{aligned} \quad (5)$$

where $\text{rbf}(\cdot)$ denotes RBFs; $\text{layernorm}(\cdot)$ is a layer normalization that ensures the input remains within a RBF domain; $b(\cdot) = \text{SiLU}(\cdot)$ is a base activation function similar to residual connection; w_b, w_s are trainable factors to better control the overall magnitude of $b(\cdot)$ and $\text{rbf}(\cdot)$ functions. Similarly, the output $\mathcal{Y}_{i,j,k}^{factor}$ of factor embedding Q_{factor}^{conv2} is calculated in the same way, and the final enhanced predicting score is represented as:

$$\mathcal{Y}_{i,j,k} = \gamma \mathcal{Y}_{i,j,k}^{factor} + (1 - \gamma) \mathcal{Y}_{i,j,k}^{index} \quad (6)$$

where γ is a hyperparameter to measure the weight of index and factor embeddings. To this end, the binary cross-entropy is used to train NTD in a mini-batch way as follows:

$$L_{NTD} = -\frac{1}{N_T} \sum_t^T \mathcal{X}_t \log(\mathcal{Y}_t) + (1 - \mathcal{X}_t) \log(1 - \mathcal{Y}_t) \quad (7)$$

where N_T is the number of training entries in tensor $\mathcal{X}$, and $T \in \mathbb{R}^{N_T \times 3}$ is a training index matrix, and t is an index vector of an entry in $\mathcal{X}$.

Results

To systematically evaluate the performance of ConvNTC, numerous experiments were conducted for ConvNTC and ten baseline methods (including five linear models, i.e., CANDECOMP/PARAFAC (CP) (Kolda and Bader, 2009), TFAI (Narita et al., 2012), DrugCom (Chen and Li, 2018), TDRC (Huang et al., 2021), CTF (Han et al., 2024), and five nonlinear models, i.e., DeepSynergy (Preuer et al., 2018), Costco (Liu et al., 2019), DTF (Sun et al., 2020),

GraphTF (Luo et al., 2021), CTF-DDI (Han et al., 2024)). Details about experimental setup (including dataset division, baseline methods, evaluation metrics and parameter setting) are provided in Supplementary Material 3. Among them, the parameters analysis was conducted to gain the optimal parameter combination of ConvNTC (Details and results are provided in Supplementary Material 4).

Ablation study

ConvNTC comprises five key components (i.e., MCTD, index embedding, factor embedding, convolutional encoder and FastKAN predictor), affecting on its performance. To assess the impact of these parts, we systematically designed five variants: ConvNTC-nfact, ConvNTC-nind, ConvNTC-nconv, ConvNTC-mlp and ConvNTC-nntd. Summarization is provided in Supplementary Table S3. The results in Supplementary Table S4 depicted that ConvNTC achieved the best performance on both datasets using all five parts. Specifically, we found that the entire nonlinear relationship learning module (i.e., NTD) is able to significantly enhance the performance of ConvNTC. Moreover, convolutional encoder and FastKAN predictor are indispensable parts of NTD, playing a crucial role in capturing the complex nonlinear characteristics. Together, the results clearly demonstrated the advantages of collaborative learning of multilinear and nonlinear modules of ConvNTC. More detailed analysis is in Supplementary Material 5.

Comparison with SOTA methods

Performance analysis

Table 1 demonstrated that ConvNTC achieved the best results across all metrics on both datasets. Notably, nonlinear methods generally outperformed linear models, highlighting their effectiveness in capturing complex tensor structures. However, most nonlinear approaches struggled to perform consistently well across both datasets, with the exception of DTF. This indicated that incorporating multilinear factor embeddings aided in modeling nonlinear relationships, thereby improving prediction accuracy. ConvNTC's superior performance over DTF further corroborated the efficacy of its design. The convolutional encoder preserves the tensor's rank-one structure during nonlinear mapping, and the FastKAN predictor effectively captures complex high-order relationships more closely aligned with tensor's intrinsic property, resulting in enhanced predictive capability. The p-values of all paired t-test were less than 0.05, further demonstrating that ConvNTC was significantly better than baselines in statistical testing.

Robustness analysis

To systematically assess the robustness of ConvNTC, we explored the influence of the negative sample size and training ratio on Dataset1. We first randomly sampled the negative samples with a size of $[1n, 2n, 4n, 6n, 8n, 10n]$, where n is the number of positive samples. Supplementary Figure S2 depicted the AUPRC values of ConvNTC and baselines, which indicated that all methods decreased as the increasing of negative sample sizes, and ConvNTC performed best. Additionally, the decrease in ConvNTC was significantly lower than others, indicating that ConvNTC exhibited strong robustness against imbalanced data. To assess the model performance under different training ratio, we set the training ratio as 10%, 30%, 50%, 70%, 90% and 100% of the original known data. Supplementary Figure S3 showed that the AUROC values of most models increase with higher training ratios and ConvNTC consistently outperformed

Table 1. Comparison results of ConvNTC and baselines on Dataset1 and Dataset2. Bolded items indicate the optimal performance.

Methods		AUPRC	AUROC	F1	Accuracy	Recall	Specificity	Precision	Paired t-test
Dataset1									
Linear	TFAI	0.9283	0.8901	0.8507	0.8518	0.8439	0.8597	0.859	0.00002
	CTF	0.9435	0.9163	0.8699	0.8737	0.8445	0.9029	0.8972	0.00014
	CP	0.9553	0.9409	0.8717	0.8734	0.86	0.8868	0.885	0.00027
	DrugCom	0.9682	0.9804	0.9437	0.9436	0.9457	0.9415	0.9418	0.00010
	TDRC	0.9732	0.9601	0.9238	0.9259	0.898	0.9539	0.9513	0.00133
	MCTD (ours)	0.978	0.9706	0.9234	0.9263	0.8905	0.9621	0.9594	0.00523
Nonlinear	Costco	0.9227	0.8906	0.8784	0.8509	0.9412	0.7606	0.8427	0.00121
	GraphTF	0.9496	0.956	0.9069	0.9034	0.9283	0.8785	0.8874	0.00022
	DeepSynergy	0.9788	0.9781	0.9274	0.9272	0.929	0.9255	0.9259	0.00071
	CTFDDI	0.982	0.9813	0.9305	0.9302	0.934	0.9264	0.9271	0.00102
	DTF	0.9832	0.9836	0.9447	0.9445	0.9489	0.9401	0.9407	0.00068
	ConvNTC (ours)	0.9988	0.9991	0.9926	0.9926	0.9946	0.9905	0.9906	
Dataset2									
Linear	CP	0.2150	0.6430	0.2838	0.8559	0.3165	0.9094	0.2625	0.00401
	TDRC	0.2696	0.6971	0.3407	0.8642	0.3887	0.9114	0.3066	0.00428
	DrugCom	0.3877	0.8004	0.4289	0.8710	0.5030	0.9075	0.3837	0.00470
	TFAI	0.4075	0.8232	0.4656	0.8935	0.5140	0.9312	0.4282	0.00629
	CTF	0.6360	0.9179	0.5986	0.9258	0.6127	0.9569	0.5887	0.00836
	MCTD (ours)	0.7563	0.9574	0.7155	0.9476	0.7298	0.9692	0.7033	0.01071
Nonlinear	GraphTF	0.2421	0.6267	0.2960	0.8464	0.3459	0.8977	0.2633	0.00323
	DeepSynergy	0.4578	0.8576	0.4657	0.9011	0.4766	0.9432	0.4591	0.00734
	CTFDDI	0.6900	0.9301	0.6321	0.9342	0.6261	0.9648	0.6404	0.00885
	Costco	0.8712	0.9338	0.8480	0.9737	0.8119	0.9898	0.8911	0.00885
	DTF	0.9176	0.9838	0.8805	0.9789	0.8632	0.9903	0.8989	0.01513
	ConvNTC (ours)	0.9269	0.9853	0.8901	0.9806	0.8691	0.9917	0.9126	

all other baselines, which underscored that ConvNTC exhibited resilience to variations in training ratio.

Generalizability analysis

To assess the generalizability of ConvNTC, we designed a new experiment to investigate whether ConvNTC can predict synergy status for new unseen data. Specifically, we conducted five-fold cross-validation in five times on the entities (i.e., removing all triplets involving the entity from the training data) rather than triple associations. As show in Supplementary Figure S4 and Supplementary Table S5, ConvNTC performed best on Dataset1, while MCTD performed best on Dataset2, followed by ConvNTC. This indicated that our proposed method can effectively predict miRNA/drug pairs of new diseases/cells to a certain extent. Additionally, We hypothesize that the inferior performance of ConvNTC compared to MCTD in Dataset 2 may stem from the 5-fold cross-validation conducted on entity-level data, resulting in significant data loss and feature noise. More analysis is in Supplementary Material 6. Overall, our method has good generalization in predicting new disease/cell-associated miRNA/drug pairs.

Case studies

Predicting novel miRNA-miRNA-disease associations

To demonstrate the capability of ConvNTC to predict novel miRNA-miRNA-disease associations, we carried out case studies on Breast Neoplasms. Given a triple association (m_i, m_j, d_k) , we validated it by verifying its pairwise associations i.e., (m_i, m_j) , (m_i, d_k) , (j, d_k) . Specifically, we confirmed the miRNA-disease associations by querying the existing public databases, i.e., dbDEMC 2.0 (Yang et al., 2017),

miRCancer (Xie et al., 2013) and PhenomiR 2.0 (Ruepp et al., 2010) and HMDD v4.0 (Cui et al., 2024), and verified the miRNA-miRNA pairs using a miRNA set enrichment analysis tool TAM 2.0 (Li et al., 2018). More details on TAM 2.0 are provided in Supplementary Material 7.1.1. Table 2 showed that all predicted top-10 miRNA pairs were confirmed, and their association network in Supplementary Figure S5 indicated that most miRNAs act on diseases in the form of pairs (or modules).

To further evaluate the biological significance of triple predictions, we perform survival analysis of breast cancer from The Cancer Genome Atlas (TCGA) on the top-10 miRNA pairs. More details on the execution of survival analysis are in Supplementary Material 7.1.3. Table 2 showed that three miRNA pairs (hsa-mir-125a, hsa-mir-99b), (hsa-mir-99a, hsa-mir-99b) and (hsa-mir-181a-2, hsa-mir-181d) have a significant impact on the patient's survival probability. Their survival curves and Pearson correlation of miRNA expression were shown in Supplementary Figure S6. For the pair (hsa-mir-125a, hsa-mir-99b), the combined analysis revealed more statistically significant segregation between high-risk and low-risk groups compared to the individual analyses of hsa-mir-125a and hsa-mir-99b. Similar results were obtained for the other two miRNA pairs. These observations indicated that these miRNA pairs have a synergistic effect on breast cancer and exhibit promising prognostic value. Interestingly, the expressions of pairs (hsa-mir-125a, hsa-mir-99b) and (hsa-mir-181a-2, hsa-mir-181d) were positively correlated, whereas the pair (hsa-mir-99a, hsa-mir-99b) was negatively correlated, implying synergistic and antagonistic role of these miRNA pairs, respectively.

Additionally, we performed the enrichment analysis on the common target genes of these three miRNA pairs. More

Table 2. The detailed evidences of the top-10 predicted miRNA pairs for **Breast neoplasms**. Noting that TF, TS and CS are the abbreviations of Transcription Factor, Tissue Specificity and Cell Specificity respectively. The bolded items indicate the max values.

MiRNA1	MiRNA2	MiRNA1-MiRNA2 Evidences							MiRNA1-Disease Evidences	MiRNA2-Disease Evidences	Triplets Evidences
		Enrichment Analysis using TAM 2.0 (-log10(p-value))									
		Function	Disease	TF	Cluster	Family	TS	CS			
hsa-mir-125a	hsa-mir-99b	2.39362	3.30437	3.57840	5.70997	4.73993	3.15366	2.95468	dbDEMC;miRCancer; PhenomiR;HMDDV4	dbDEMC;PhenomiR; HMDDV4	3.25964
hsa-mir-10a	hsa-mir-99b	0.00000	2.82136	3.57840	0.00000	4.73993	0.00000	0.00000	dbDEMC;miRCancer; PhenomiR;HMDDV4	dbDEMC;PhenomiR; HMDDV4	0.52288
hsa-mir-19b-1	hsa-mir-19b-2	2.64839	3.49975	3.73060	0.00000	5.70997	3.86646	3.11691	dbDEMC;PhenomiR	dbDEMC;PhenomiR	0.12494
hsa-mir-99a	hsa-mir-99b	3.82391	3.67684	0.00000	0.00000	4.73993	0.00000	0.00000	dbDEMC;miRCancer; PhenomiR;HMDDV4	dbDEMC;PhenomiR; HMDDV4	1.50864
hsa-mir-19a	hsa-mir-19b-2	2.64839	3.17420	3.73060	0.00000	5.70997	3.86646	3.11691	dbDEMC;miRCancer; PhenomiR;HMDDV4	dbDEMC;PhenomiR	0.43180
hsa-mir-181c	hsa-mir-181d	3.32808	2.93417	0.00000	5.18709	5.01100	0.00000	3.27327	dbDEMC;PhenomiR; HMDDV4	dbDEMC;PhenomiR; HMDDV4	1.22915
hsa-mir-154	hsa-mir-494	0.00000	2.55261	0.00000	3.25259	3.95468	0.00000	0.00000	dbDEMC;miRCancer; PhenomiR;HMDDV4	dbDEMC;miRCancer; PhenomiR;HMDDV4	0.21467
hsa-mir-181a-2	hsa-mir-181d	3.45204	2.92905	0.00000	0.00000	5.01100	0.00000	0.00000	dbDEMC;PhenomiR	dbDEMC;PhenomiR; HMDDV4	2.16115
hsa-mir-125b-2	hsa-mir-99b	2.39362	3.20235	0.00000	0.00000	4.73993	0.00000	0.00000	dbDEMC;PhenomiR; HMDDV4	dbDEMC;PhenomiR; HMDDV4	1.23657
hsa-mir-100	hsa-mir-99b	3.10876	3.89752	0.00000	0.00000	4.73993	0.00000	0.00000	dbDEMC;miRCancer; PhenomiR;HMDDV4	dbDEMC;PhenomiR; HMDDV4	0.50864

details on the operation of enrichment analysis are provided in Supplementary Material 7.1.4. Figure 2 depicted the enriched terms and pathways that are known to be implicated in the breast cancer progression. Processes like lysosome and vacuole organization support cancer cell survival, while chromosome segregation and mitotic regulation maintain genomic stability. The ‘Breast cancer pathway’ involves hormonal signaling, cell growth and survival mechanisms specific to breast cancer. Other pathways like PI3K/Akt/mTOR promotes tumor growth, and ATM signaling aids DNA repair. Together, these enriched terms and pathways highlight specific metabolic and signaling processes in breast neoplasm.

Predicting novel synergistic drug pairs

We employed ConvNTEC to predict novel synergistic drug pairs after training on Dataset2. We showcased the 855 potential drug pairs with 42 drugs for malignant melanoma cancer cell line A375 from paper (Wang et al., 2022). The supports of top-10 predicted drug combinations in Table 3 showed that 9 out of the 10 predicted combinations have been corroborated by prior studies or clinical trials. These combinations focus on inhibiting growth signals, DNA repair mechanisms and cellular metabolism to enhance treatment efficacy for A375 cells. More analysis are provided in Supplementary Material 7.2. Therefore, our drug-drug-cell predictions hold therapeutic potential and warrant further experimental validation.

Table 3. The supports of the top-10 predicted novel synergistic combinations on A375 cancer cell line.

Drug 1	Drug 2	Evidences	
		Publications (PMID)	
Erlotinib	Talazoparib	36601757; 26124641	
Erlotinib	Lenvatinib	36066408	
Erlotinib	Enasidenib	NA	
Erlotinib	Everolimus	22378048; 22968184	
Erlotinib	Copanlisib	34734007	
Erlotinib	Ruxolitinib	25065853	
Abemaciclib	Erlotinib	33194700	
Erlotinib	Crizotinib	28961830; 31960422; 31078602; 32540560	
Erlotinib	Pazopanib	29645086; 22688250; 23135778	
Erlotinib	Regorafenib	25907508	

Conclusion

By building on the premise of reformulating the task of identifying disease-associated miRNA pairs and cell-related drug pairs as a third-order tensor completion problem, we presented a two-stage hybrid tensor completion framework ConvNTEC. Different from prior work, ConvNTEC modeled the tensor’s underlying multilinear structure by developing a multi-constraint tensor decomposition method with three constraints. Furthermore, ConvNTEC integrated CNN and FastKAN into a unified architecture, which can comprehensively capture the intrinsic high-order complexity of the tensor and the potential nonlinear interactions between objects. As the key contribution, we proposed a convolutional encoder that enabled nonlinear feature mapping while maintaining the tensor’s rank-one property, alongside a FastKAN predictor that effectively captured high-dimensional relationships aligned with the intrinsic structure of the real-world data. This provides a way to enhance model representation capabilities and facilitate feature interaction learning in this domain. Numerous experiments demonstrated the superiority and effectiveness of ConvNTEC in predicting triple relationships, and case studies exhibited the significant potential of ConvNTEC in advance precision medicine and drug development.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (#62372165 and #62032007) and the Natural Sciences and Engineering Research Council (NSERC) Discovery Grant (RGPIN-2019-0621).

References

- J. D. Carroll and J.-J. Chang. Analysis of individual differences in multidimensional scaling via an n-way generalization of “eckart-young” decomposition. *Psychometrika*, 35(3):283–319, 1970.
- H. Chen and J. Li. Drugcom: Synergistic discovery of drug combinations using tensor decomposition. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 899–904. IEEE, 2018.
- M. Cokol, H. N. Chua, M. Tasan, B. Mutlu, Z. B. Weinstein, Y. Suzuki, M. E. Nergiz, M. Costanzo, A. Baryshnikova,

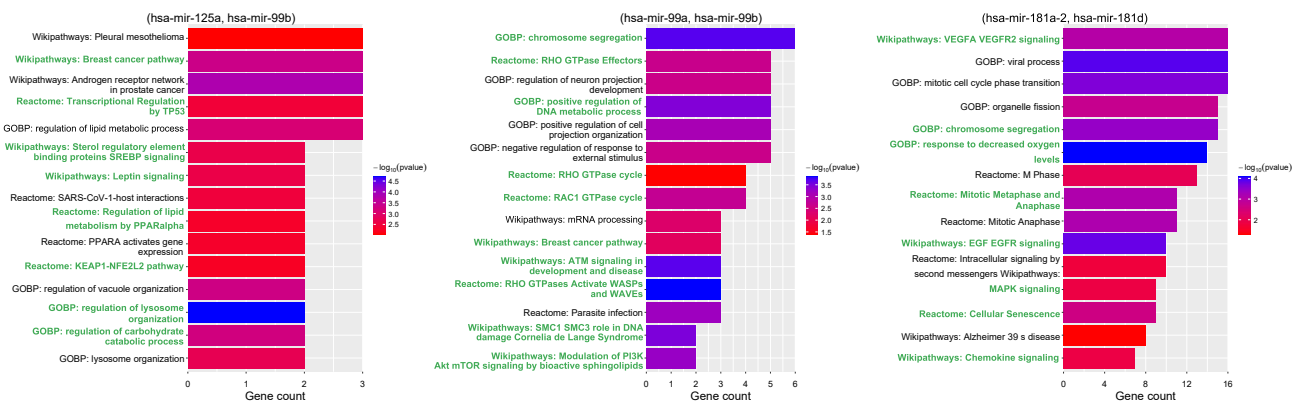


Fig. 2. Enrichment analysis of miRNA pairs (hsa-mir-125a, hsa-mir-99b), (hsa-mir-99a, hsa-mir-99b) and (hsa-mir-181a-2, hsa-mir-181d) related to Breast Neoplasms on 'GO_BP', 'Wikipathways' and 'Reactome' using R package 'clusterProfiler' (Yu et al., 2012) and 'ReactomePA' (Yu and He, 2016). Noting that green fonts indicate key processes and pathways related to breast cancer.

- G. Giaever, et al. Systematic exploration of synergistic drug pairs. *Molecular systems biology*, 7(1):544, 2011.
- C. Cui, B. Zhong, R. Fan, and Q. Cui. Hmdd v4. 0: a database for experimentally supported human microrna-disease associations. *Nucleic Acids Research*, 52(D1): D1327–D1332, 2024.
- I. Drokin. Kolmogorov-arnold convolutions: Design principles and empirical studies. *arXiv preprint arXiv:2407.01092*, 2024.
- G. Han, L. Peng, A. Ding, Y. Zhang, and X. Lin. Ctf-ddi: Constrained tensor factorization for drug–drug interactions prediction. *Future Generation Computer Systems*, 161:26–34, 2024.
- F. Huang, X. Yue, Z. Xiong, Z. Yu, S. Liu, and W. Zhang. Tensor decomposition with relational constraints for predicting multiple types of microrna-disease associations. *Briefings in bioinformatics*, 22(3):bbaa140, 2021.
- T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM review*, 51(3):455–500, 2009.
- X. Lai, M. Eberhardt, U. Schmitz, and J. Vera. Systems biology-based investigation of cooperating micrnas as monotherapy or adjuvant therapy in cancer. *Nucleic acids research*, 47(15):7753–7766, 2019.
- C. Li, X. Liu, W. Li, C. Wang, H. Liu, and Y. Yuan. U-kan makes strong backbone for medical image segmentation and generation. *arXiv preprint arXiv:2406.02918*, 2024.
- J. Li, X. Han, Y. Wan, S. Zhang, Y. Zhao, R. Fan, Q. Cui, and Y. Zhou. Tam 2.0: tool for microrna set analysis. *Nucleic acids research*, 46(W1):W180–W185, 2018.
- Z. Li. Kolmogorov-arnold networks are radial basis function networks. *arXiv preprint arXiv:2405.06721*, 2024.
- H. Liu, Y. Li, M. Tsang, and Y. Liu. Costco: A neural tensor completion model for sparse tensors. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 324–334, 2019.
- P. Liu, J. Luo, and X. Chen. mircom: tensor completion integrating multi-view information to deduce the potential disease-related mirna-mirna pairs. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 19(3):1747–1759, 2020.
- Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark. Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756*, 2024.
- J. Luo, Z. Lai, C. Shen, P. Liu, and H. Shi. Graph attention mechanism-based deep tensor factorization for predicting disease-associated mirna-mirna pairs. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 189–196. IEEE, 2021.
- A. Narita, K. Hayashi, R. Tomioka, and H. Kashima. Tensor factorization using auxiliary information. *Data Mining and Knowledge Discovery*, 25:298–324, 2012.
- K. Preuer, R. P. Lewis, S. Hochreiter, A. Bender, K. C. Bulusu, and G. Klambauer. Deepsynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics*, 34(9): 1538–1546, 2018.
- A. Ruepp, A. Kowarsch, D. Schmidl, F. Buggenthin, B. Brauner, I. Dunger, G. Fobo, G. Frishman, C. Montrone, and F. J. Theis. Phenomir: a knowledgebase for microrna expression in diseases and biological processes. *Genome biology*, 11:1–11, 2010.
- Z. Sun, S. Huang, P. Jiang, and P. Hu. Dtf: deep tensor factorization for predicting anticancer drug synergy. *Bioinformatics*, 36(16):4483–4489, 2020.
- J. Wang, X. Liu, S. Shen, L. Deng, and H. Liu. Deepdds: deep graph neural network with attention mechanism to predict synergistic drug combinations. *Briefings in Bioinformatics*, 23(1):bbab390, 2022.
- B. Xie, Q. Ding, H. Han, and D. Wu. mircancer: a microrna–cancer association database constructed by text mining on literature. *Bioinformatics*, 29(5):638–644, 2013.
- J. Xu, C.-X. Li, Y.-S. Li, J.-Y. Lv, Y. Ma, T.-T. Shao, L.-D. Xu, Y.-Y. Wang, L. Du, Y.-P. Zhang, et al. Mirna–mirna synergistic network: construction via co-regulating functional modules and disease mirna topological features. *Nucleic acids research*, 39(3):825–836, 2011.
- Z. Yang, L. Wu, A. Wang, W. Tang, Y. Zhao, H. Zhao, and A. E. Teschendorff. dbdenc 2.0: updated database of differentially expressed mirnas in human cancers. *Nucleic acids research*, 45(D1):D812–D818, 2017.
- G. Yu and Q.-Y. He. Reactomepa: an r/bioconductor package for reactome pathway analysis and visualization. *Molecular BioSystems*, 12(2):477–479, 2016.
- G. Yu, L.-G. Wang, Y. Han, and Q.-Y. He. clusterprofiler: an r package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology*, 16(5):284–287, 2012.