Multi-view Contrastive Learning Hypergraph Neural Network for Drug-Microbe-Disease Association Prediction

Luotao Liu[†], **Feng Huang**[†], **Xuan Liu**, **Zhankun Xiong**, **Menglu Li**, **Congzhi Song** and **Wen Zhang**[*]

College of Informatics, Huazhong Agricultural University, Wuhan 430070, China
{luotaoliu, fhuang233, lx666, xiongzk, mengluli, song_cz }@webmail.hzau.edu.cn,
zhangwen@mail.hzau.edu.cn

IJCAI-2023

Code: https://github.com/Liuluotao/MCHNN

Reported by Dongdong Hu

Introduction

There are plenty of computational methods for pair-wise association prediction, such as drug-microbe and microbe-disease associations, but few methods focus on the higher-order triple wise drug-microbe-disease (DMD)

However, the confrmed DMD associations are insuffcient due to the high cost of in vitro screening, which forms a sparse DMD hypergraph and thus brings in suboptimal generalization ability.
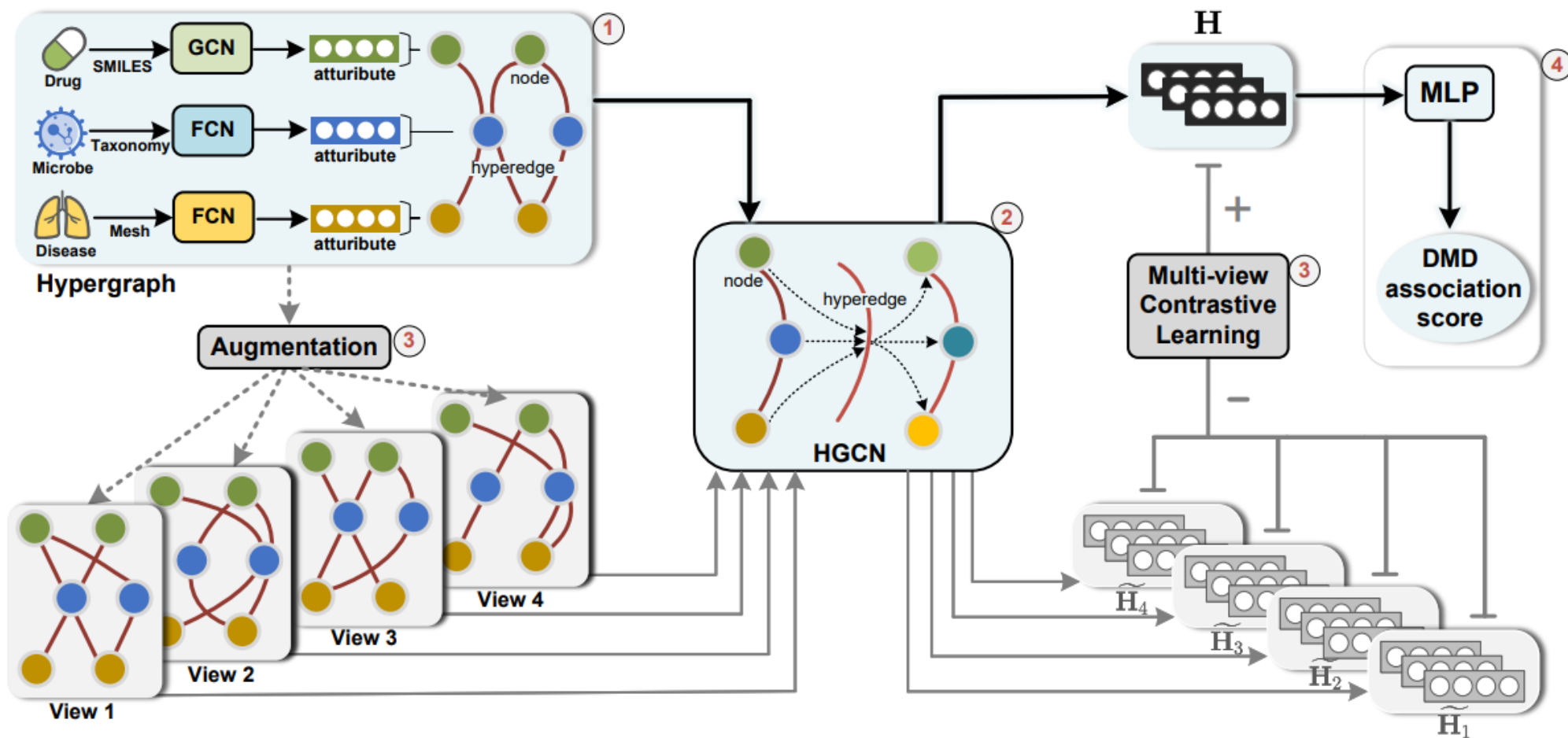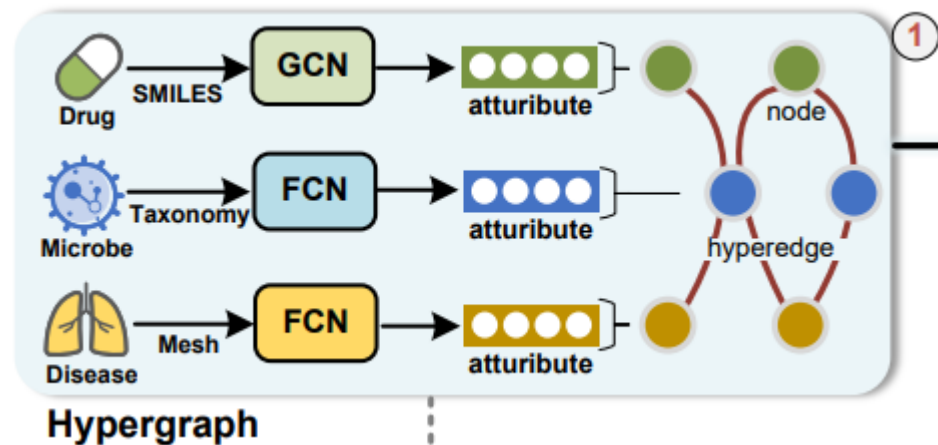
Method

Figure 1: Workflow of MCHNN: ① DMD Hypergraph construction, ② Hypergraph representation Learning, ③ Multi-View Contrastive Learning, ④ Model training.

Method

Hypergraph

drug set $\mathcal{D}$, microbe set $\mathcal{M}$ disease set $\mathcal{N}$,

their Cartesian product $\mathcal{S} = \mathcal{D} \times \mathcal{M} \times \mathcal{N}$ is a set of all possible DMD triplets.

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}), \quad \mathcal{V} = \mathcal{D} \cup \mathcal{M} \cup \mathcal{N}$$

known DMD associations are represented as hyperedges $\mathcal{E}$

$$X \in \mathbb{R}^{|\mathcal{V}| \times F} \qquad Y \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{E}|} \quad \mathcal{E} \subset \mathcal{S}$$

DMD Hypergraph Construction

$$Y_{ve} = \begin{cases} 1, & \text{if } v \in e \\ 0, & \text{if } v \notin e \end{cases} \qquad (1)$$

$$X \qquad X_{\mathcal{D}} \quad X_{\mathcal{M}} \quad X_{\mathcal{N}}$$

GIN

$$Z^{(k)} = \text{MLP}^{(k)} \left((A + (1 + \epsilon)I) Z^{(k-1)} \right) \qquad (2)$$

$$X_{\mathcal{D}} \in^{|\mathcal{D}| \times F}$$
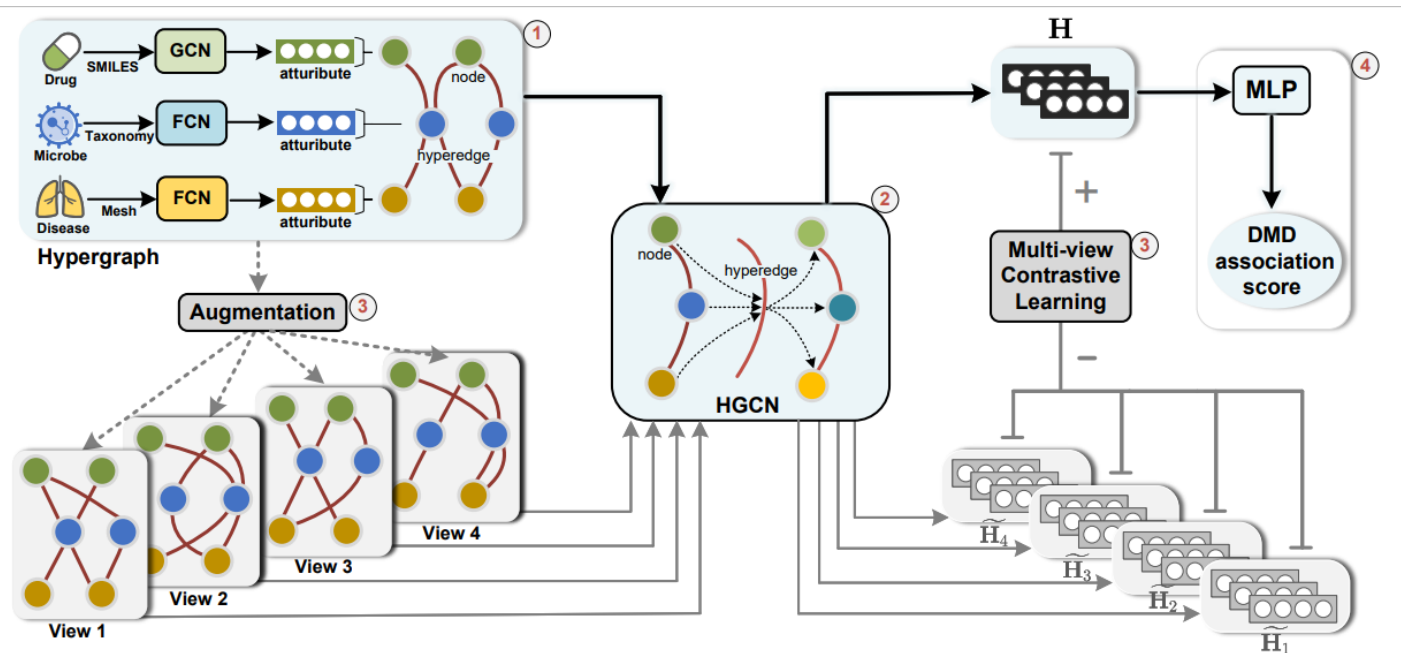
similarity matrices

$$S_{\mathcal{M}} \in \{1, 0\}^{|\mathcal{M}| \times |\mathcal{M}|} \qquad S_{\mathcal{N}} \in \mathbb{R}^{|\mathcal{N}| \times |\mathcal{N}|}$$

$$X_{\mathcal{M}} \in \mathbb{R}^{|\mathcal{M}| \times F} \qquad X_{\mathcal{N}} \in \mathbb{R}^{|\mathcal{N}| \times F}$$

fully-connected networks

Method

Figure 1: Workflow of MCHNN: ① DMD Hypergraph construction, ② Hypergraph representation Learning, ③ Multi-View Contrastive Learning, ④ Model training.

$$H \in \mathbb{R}^{|\mathcal{V}| \times F} \rightarrow s \in \mathbb{R}^F \qquad \text{global mean pooling layer}$$

$$\mathcal{L}_c = -\frac{1}{5|\mathcal{V}|} \left(\sum_{v \in \mathcal{V}} \log \Psi(h_v, s) + \sum_{z=1}^{4} \sum_{v \in \mathcal{V}} \log \left(1 - \Psi(\tilde{h}_v^z, s) \right) \right) \qquad (4)$$

HGCN

$$H^{(l)} = \sigma \left(D^{-1} Y W B^{-1} Y^\top H^{(l-1)} \Theta^{(l-1)} \right) \qquad (3)$$

where $\sigma(\cdot)$ indicate a nonlinear activation function (ReLU), Θ is a learnable weight matrix, $H^{(l)}$ is the node embeddings at the l-th layer and $H^{(0)}$ is initialized with X; D and B are diagonal matrices respectively corresponding to the sums of rows and columns in Y called degrees of nodes and hyperedges;

View 1 (drug-mode perturbation): $\qquad (d', m, n)$

View 2 (microbe-mode perturbation) $\qquad (d, m', n)$

View 3 (disease-mode perturbation) $\qquad (d, m, n')$

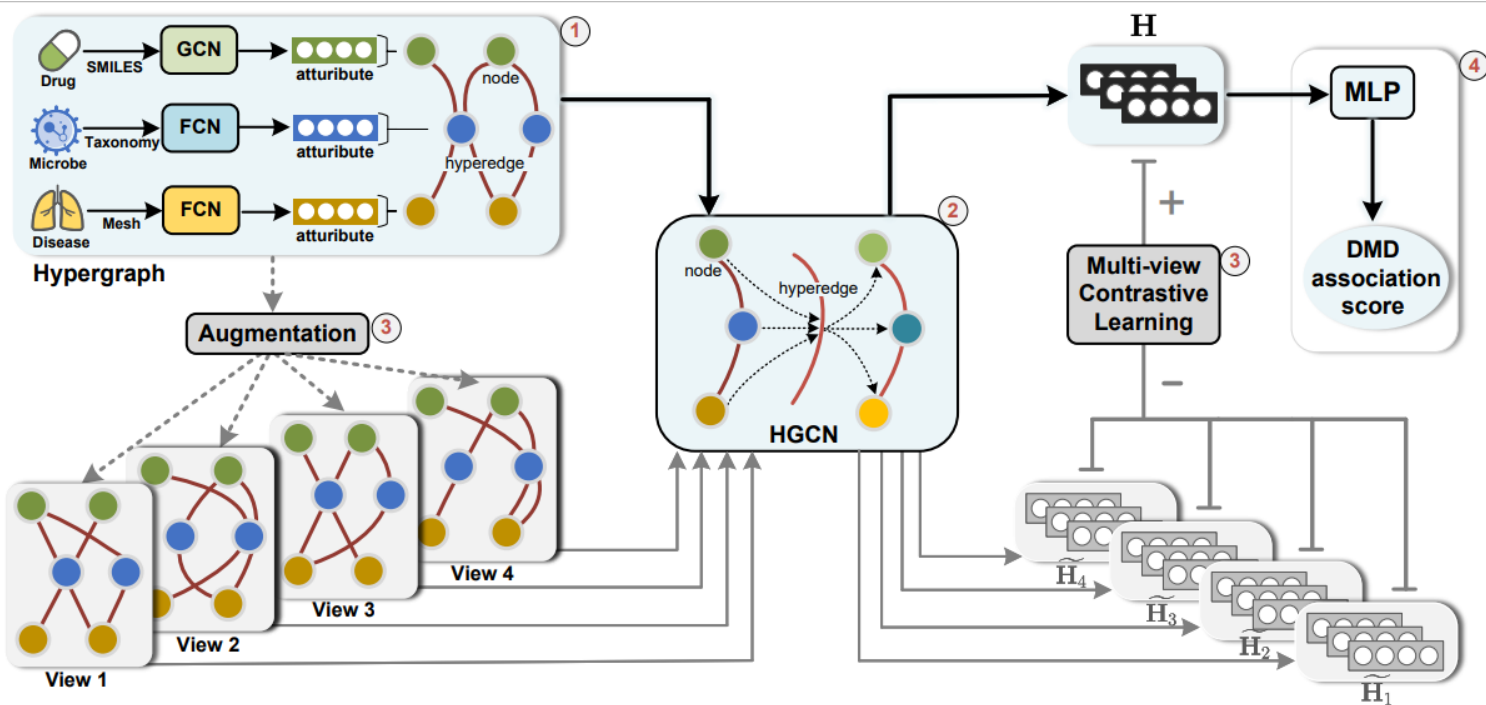**View 4** (random perturbation) $\qquad (d', m', n')$

$$\{(X, \tilde{Y}_z)\}_{z=1}^{Z} \qquad \{\tilde{H}_z\}_{z=1}^{Z}$$

Method

Figure 1: Workflow of MCHNN: ① DMD Hypergraph construction, ② Hypergraph representation Learning, ③ Multi-View Contrastive Learning, ④ Model training.

$$\hat{p} = \mathrm{MLP}(\boldsymbol{h}_d \parallel \boldsymbol{h}_m \parallel \boldsymbol{h}_n) \qquad (5)$$

$$\mathcal{L}_p = -\frac{1}{|\mathcal{T}|} \sum_{i \in \mathcal{T}} (p_i \log \hat{p}_i + (1 - p_i) \log(1 - \hat{p}_i)) \qquad (6)$$

$$\mathcal{L} = \alpha \mathcal{L}_p + (1 - \alpha)\mathcal{L}_c \qquad (7)$$
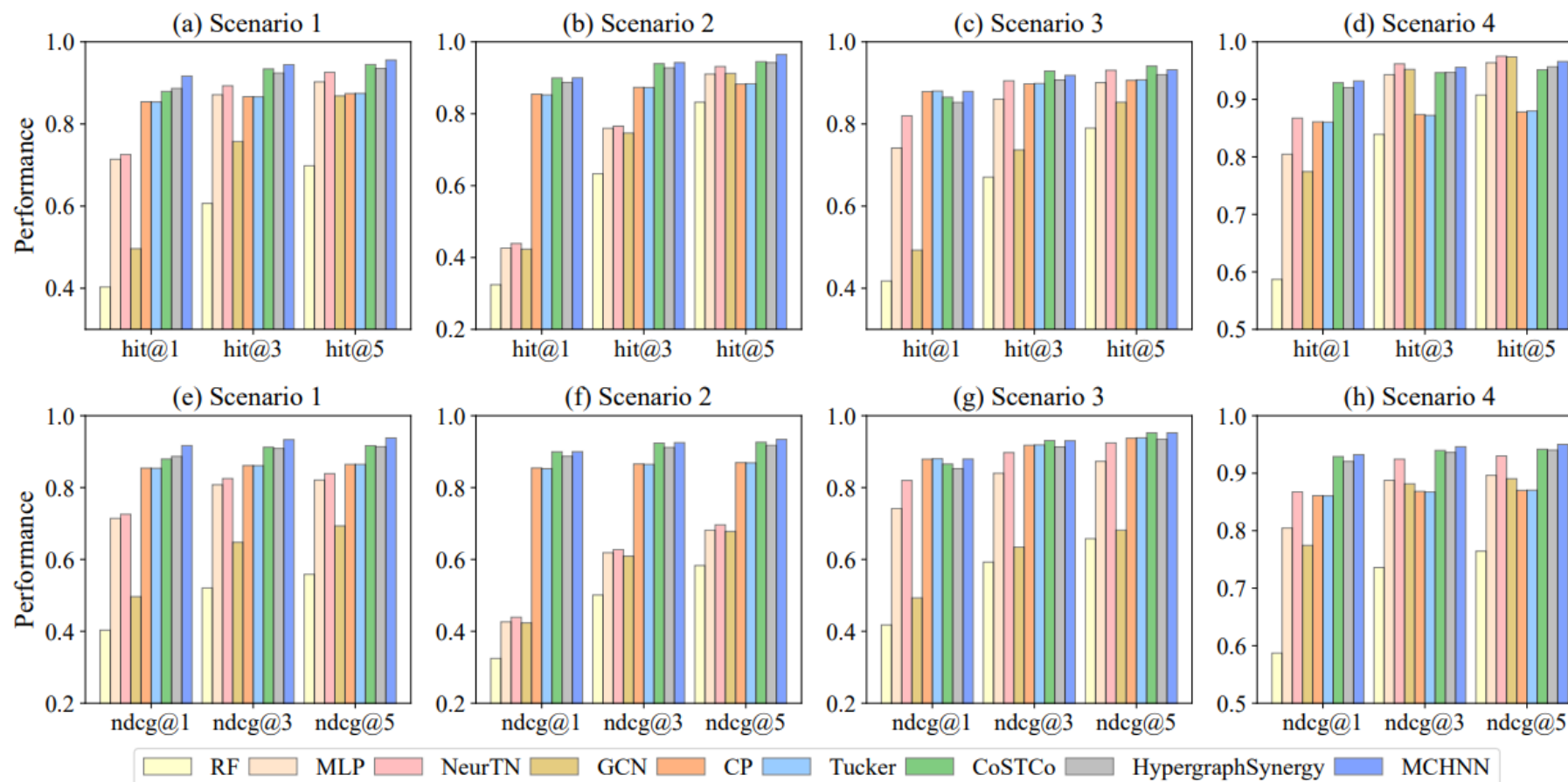
Experiments

Figure 2: 5-CV performance of MCHNN and baselines in four scenarios in terms of hit@n and ndcg@n.
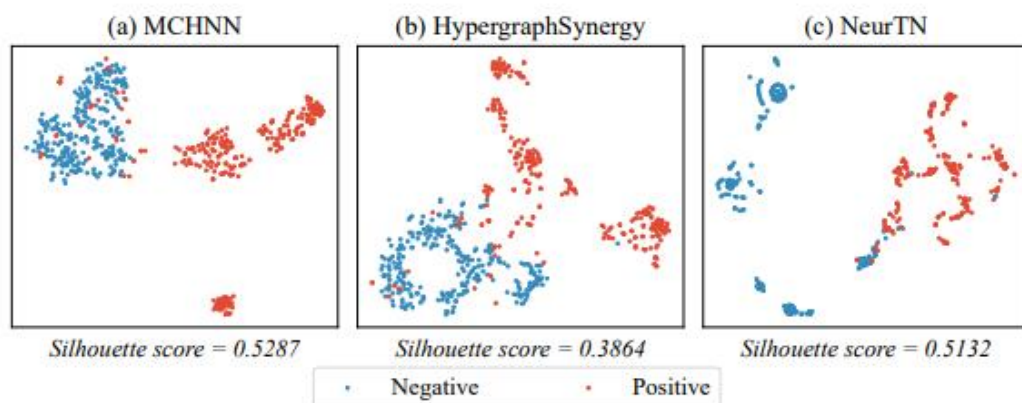
Experiments

Figure 3: The t-SNE visualization of three models on scenario 4.
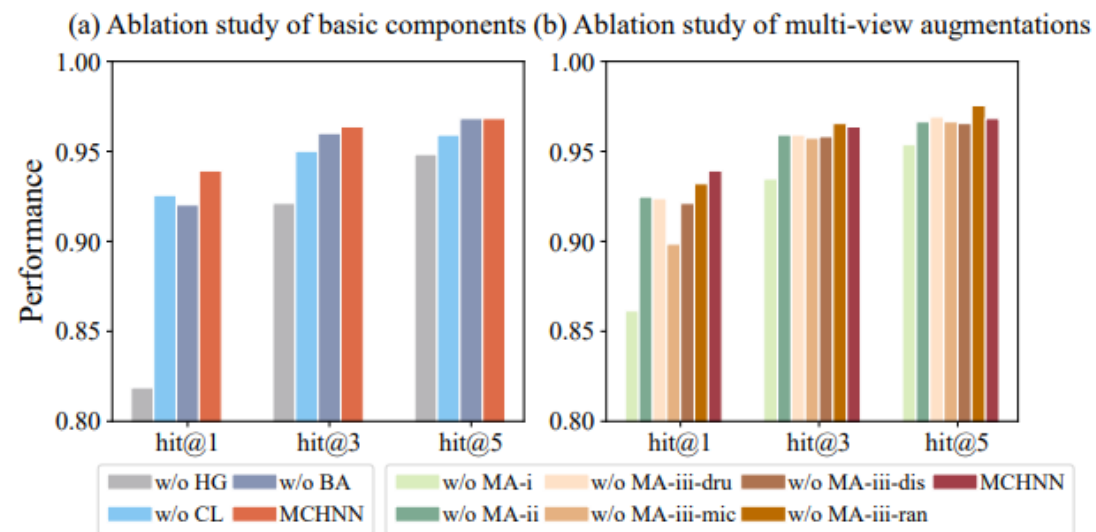

Figure 4: Average values of MCHNN and its variants on four scenarios in ablation study.

Experiments

Methods	MCHNN			w/o CL		
	hits@1	hits@3	hits@5	hits@1	hits@3	hits@5
$[0, 50]$	0.6736	**0.8299**	**0.8785**	**0.6840**	0.7882	0.8160
$[51, 100]$	**0.9229**	**0.9840**	**0.9894**	0.9016	0.9335	0.9548
$[101, +\infty]$	0.9909	0.9955	**1.0000**	**0.9955**	**1.0000**	1.0000

Table 1: Average values of MCHNN and w/o CL on four scenarios in terms of average degrees of nodes in triplets.

Thanks