

Beijing Forest Studio
北京理工大学信息系统及安全对抗实验中心



超图对比学习

硕士研究生 吴松凌

2024年07月28日

- 总结反思

- 算法和公式讲解不够细致，要明确输入输出关系、物理意义，并加以实例辅助
- 演讲技巧欠缺，可以调整语速，丰富语调，调动观众注意力
- 要加强互动，同时注重观众反馈，把控整体节奏

- 相关内容

- 2023.06.18 沈宇辉 《图匹配网络》
- 2023.05.07 周瑾洁 《基于图的知识追踪方法研究》
- 2022.09.25 李新帅 《异质图神经网络》
- 2022.02.27 张凌浩 《基于图结构处理的文本生成》

- 预期收获
- 题目内涵解析
- 研究背景与意义
- 研究历史与现状
- 知识基础
- 算法原理
 - TriCL
 - CCL
- 特点总结与工作展望
- 参考文献

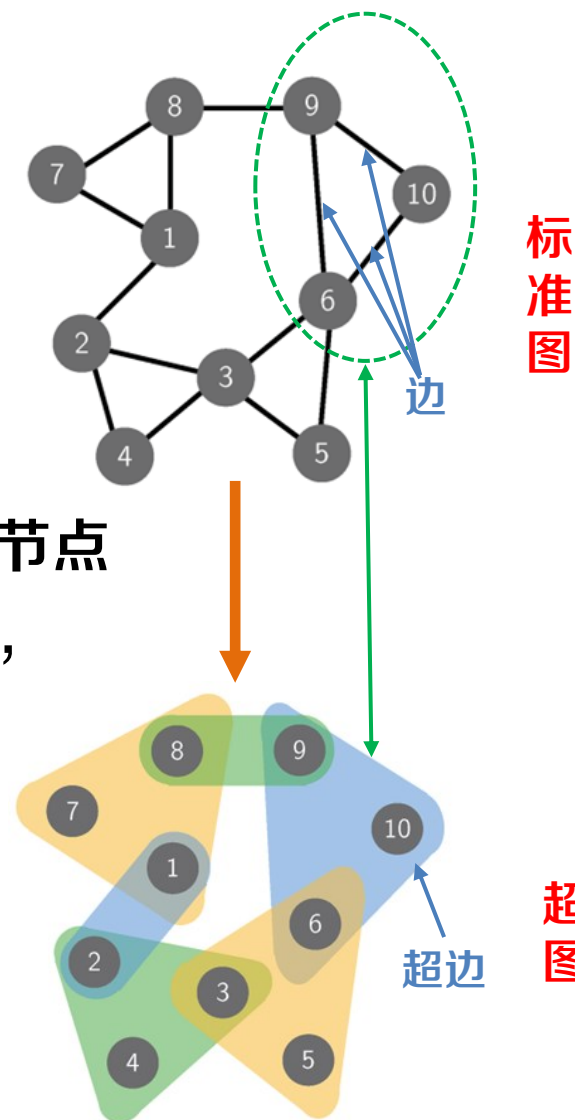
- 预期收获
 - 了解超图对比学习方法的基本概念
 - 熟悉超图对比学习方法的历史现状及应用场景
 - 理解超图增强、超图编码器、超图对比损失和对比学习的技术原理
 - 明确超图对比学习方法的发展趋势和未来前景

- 研究目标

- 面向**超图**的**表示学习**任务
- 利用**对比学习**来发现超图中的不变性和差异性，从而生成**区分度高、泛化能力强**的节点或超边表示

- 内涵解析

- 超图：标准图的扩展，由**超边**和**节点**组成，超边连接任意数量节点
- 对比学习：**自监督**学习方法，通过**正样本对**和**负样本对**的比较，学习出能够在不同下游任务中通用的高质量表示
 - 正样本对：来自**相同类别**的样本或是经过变换的**相同**样本
 - 负样本对：来自**不同类别**的样本或是经过变换的**不同**样本
- 超图对比学习：用对比学习方式**训练超图神经网络**，得到**节点或超边**的通用高质量表示



Collaborative contrastive learning for hypergraph node classification

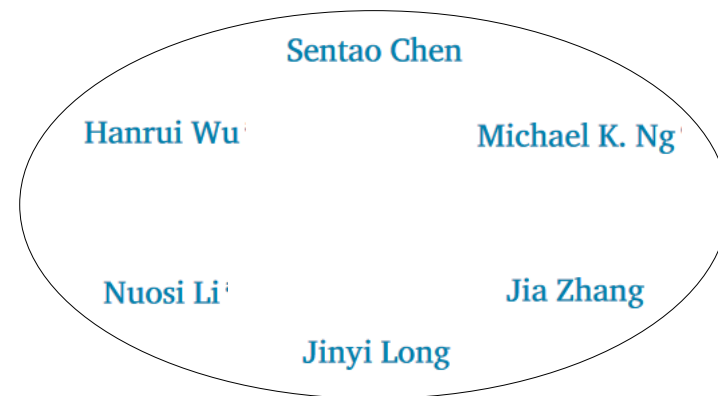
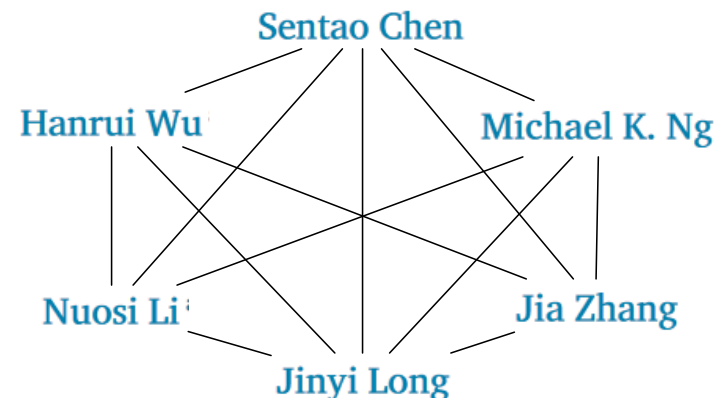
Hanrui Wu^{a,e}, Nuosi Li^a, Jia Zhang^{a,e}, Sentao Chen^b, Michael K. Ng^c, Jinyi Long^{a,d,e,*}

- 研究背景

- 数据复杂性：医疗、社交、推荐及文献分类等领域的数据存在**多对多**、**高阶**的关联关系
- **标注稀缺性**：监督或半监督学习需要大量标注数据，而这通常是时间和资源密集型的

- 研究意义

- 建模复杂关系：超边连接任意数量节点，能全面刻画节点间的**高阶关系**，提升表示学习性能
- 缓解**过平滑**问题：对比学习可以缓解超图神经网络中信息迭代传播导致的过平滑问题，即节点表示趋于相似
- 增强模型适应性：自监督方式训练的超图网络不易过拟合，可**泛化**到多种下游任务和不同数据中



Yu等人**开创性**地用自监督学习弥补多通道聚合引起的高阶连接信息损失，显著改善了**社交推荐**效果。其选择**子图采样**方式构建对比视图

2021

Wei等人首次提出使用**超图生成模型**来构建对比视图，并与超图神经网络共同训练，以提高低标签状态下超图神经网络的泛化能力

2022

Lee等人提出构建节点间、超边间以及节点与超边的**三向对比损失**，捕获节点嵌入中的节点级和超边级结构信息

2023

Qian等人构建了节点级和超边级的**双层对比损失**，并在损失计算中引入**自适应温度参数**以增强对比优化效果

2024

2021

Zhang等人提出了一种用于**社交推荐**的自监督超图学习框架，使用**双尺度节点丢弃**策略来创建自监督信号，从而对不同粒度的用户表示进行正则化以对抗稀疏性问题

2022

Xia等人提出了超图对比协同过滤框架，通过超图增强的**交叉视图**对比学习架构共同捕获局部和全局协作关系，以加强推荐系统的表示质量

2024

Wu等人提出用于超图节点分类的协同对比学习方法，根据超图导出节点间关系的标准图，在每个网络层中以协同交替的方式对比**标准图和超图**

超图对比学习

超图增强技术

对比损失形式

对比训练策略

• 超图

- $G = (V, E)$, 由节点和超边组成, 每个超边都是 V 的非空子集
- 关联矩阵 $H = \{0, 1\}^{|V| \times |E|}$, 表示节点 v 是否属于超边 e

• 超图神经网络——以HGNCN为例

- 核心思想: 节点特征先聚合到超边, 然后超边特征传播回节点

D_v : 节点度矩阵, 用于归一化

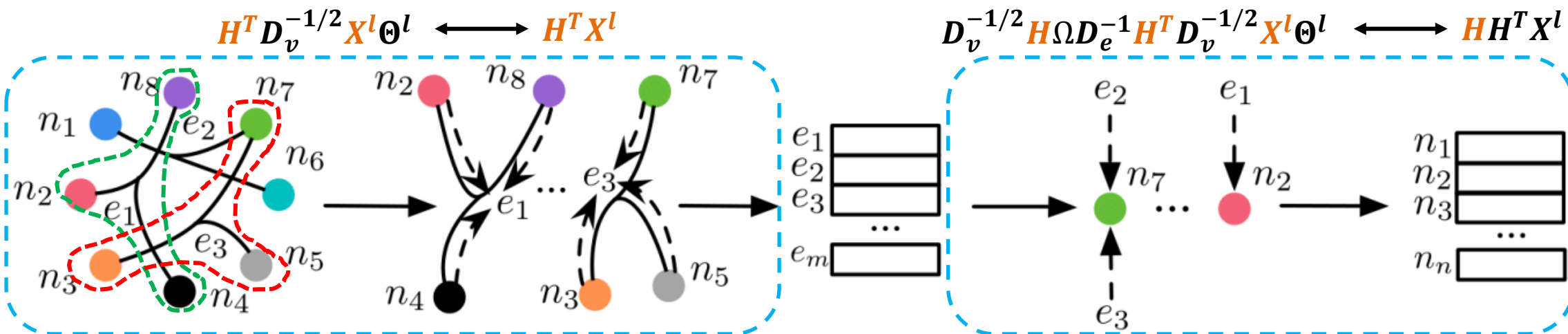
D_e : 超边度矩阵, 用于归一化

Ω : 超边权重矩阵

X^l : 节点第 l 层的特征

Θ^l : 全连接层, 对 X^l 进行特征转换

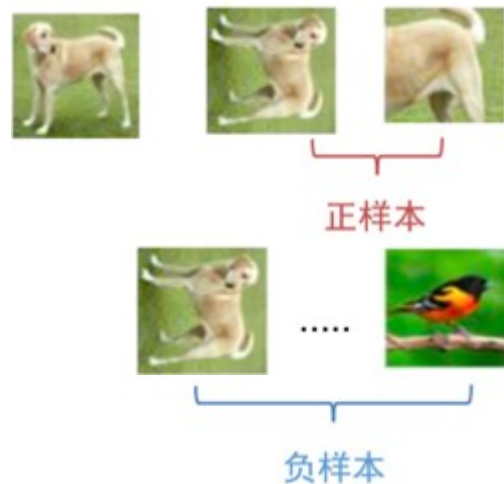
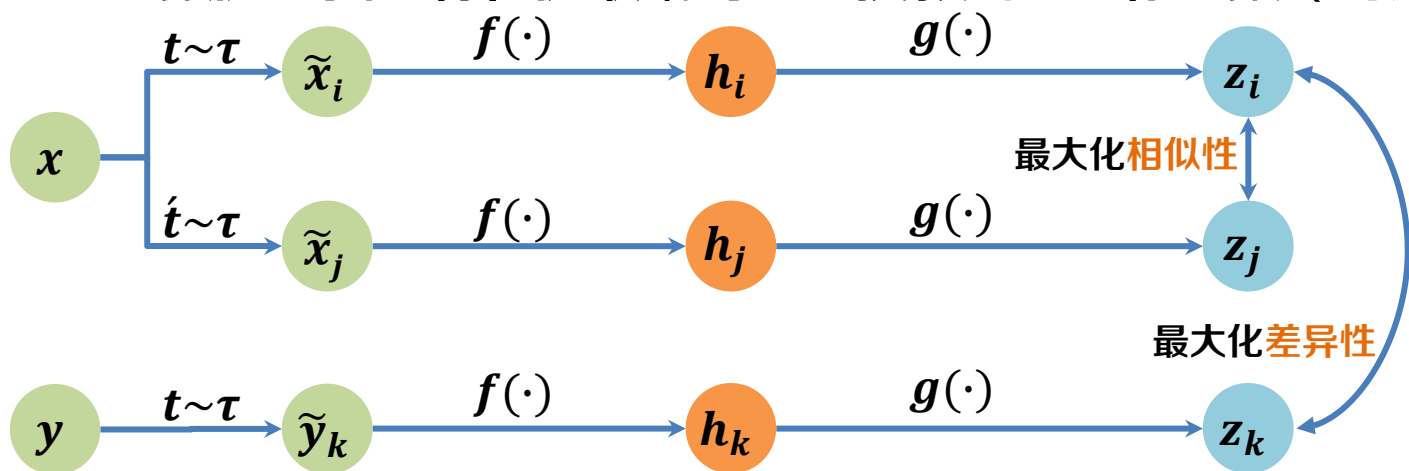
$$X^{l+1} = \sigma(D_v^{-1/2} H \Omega D_e^{-1} H^T D_v^{-1/2} X^l \Theta^l)$$



• 算法原理

- 原始样本经数据增强生成多个增强视图，设计对比损失训练模型，拉近正样本对、拉远负样本对

- 数据增强→特征提取编码器→投影层→目标函数（对比损失）



- 正样本对：同一原始样本的2个增强视图之间构成正样本对
- 负样本对：不同原始样本的2个增强视图之间构成负样本对
- 下游应用：用训练好的编码器处理原始样本，得到样本表示 $f(x)$ 、 $f(y)$

- 算法原理

- 超图增强→超图神经网络→投影层→超图对比损失

- 超图对比损失

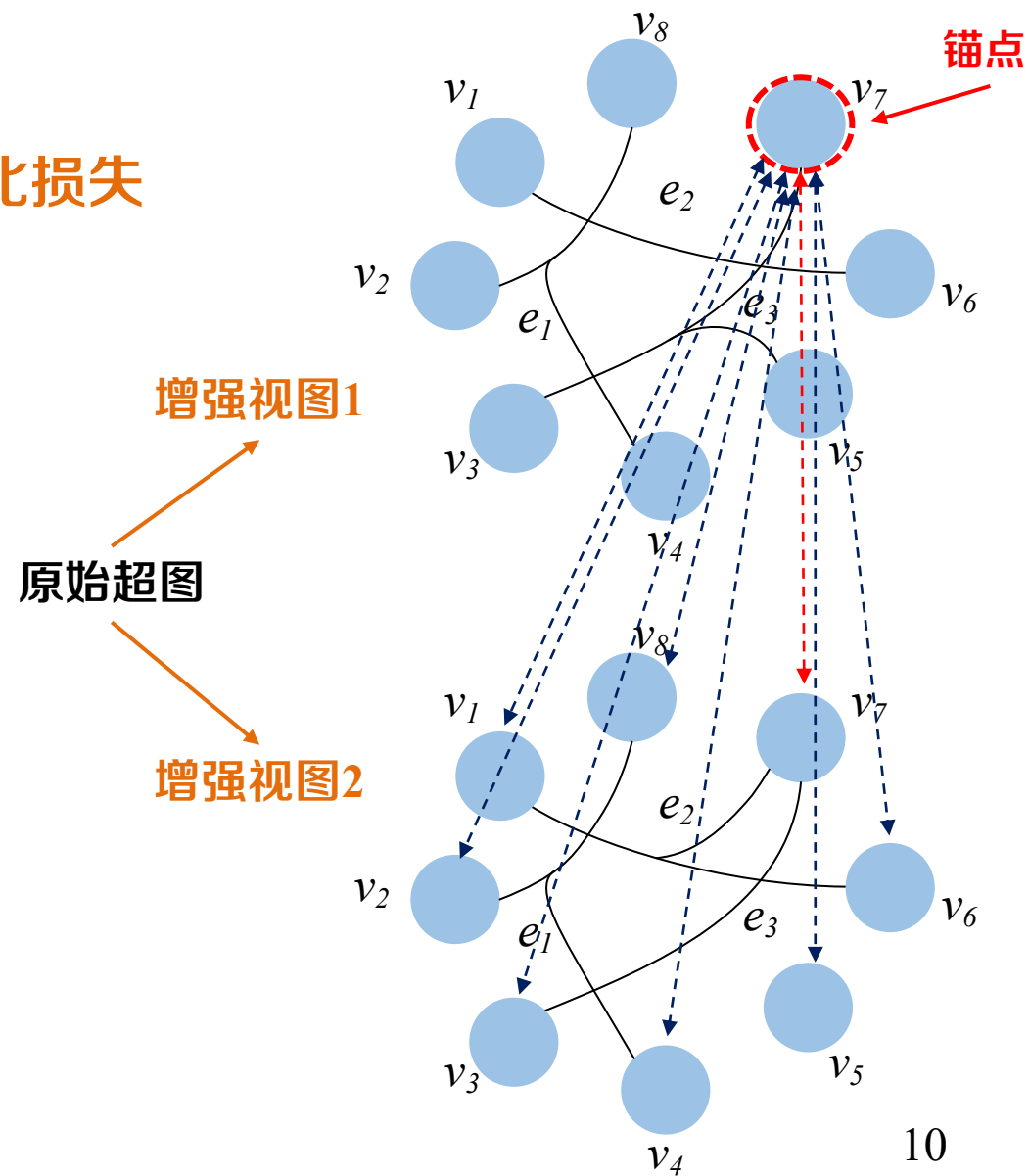
- 锚点：用于计算对比损失的基础点

- 锚点：增强视图1中节点 v_i 的表示
 - 正样本：增强视图2中节点 v_i 的表示
 - 负样本：增强视图2中其他节点的表示
 - 对应损失

$$l(z_{1,i}, z_{2,i}) = -\log \frac{\exp(\text{sim}(z_{1,i}, z_{2,i})/\tau)}{\sum_{k=1}^n \exp(\text{sim}(z_{1,i}, z_{2,k})/\tau)}$$

- 整体对比损失

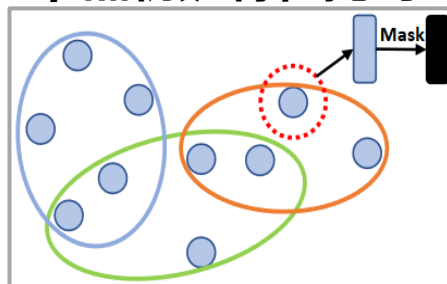
$$L = \frac{1}{2|V|} \sum_{i=1}^{|V|} \{l(z_{1,i}, z_{2,i}) + l(z_{2,i}, z_{1,i})\}$$



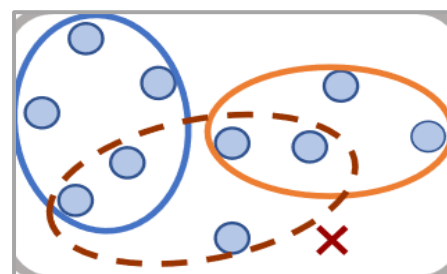
• 手动规则的超图增强

- 节点初始特征掩码
- 超边删除
- 节点删除
- 超边成员掩码
- 随机游走生成子图
- 标准图转化

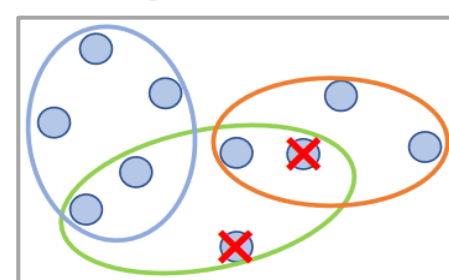
节点初始特征掩码



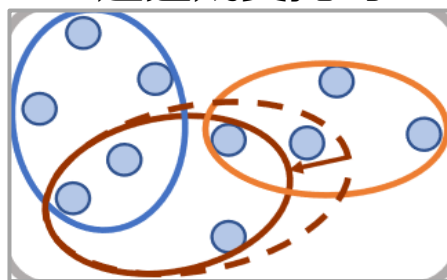
超边删除



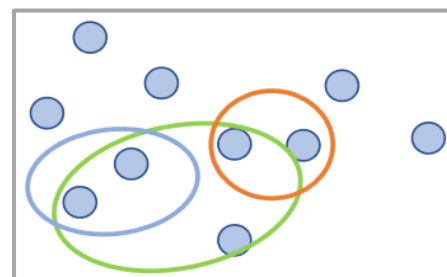
节点删除



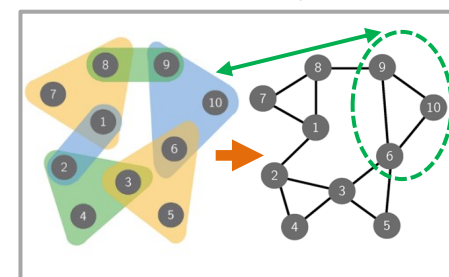
超边成员掩码



随机游走生成子图

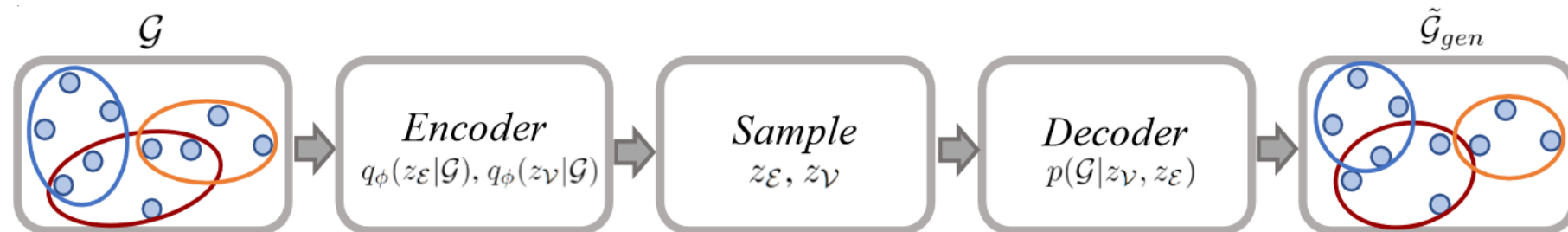


标准图转化



• 超图生成模型

- 类似变分自编码器 (VAE)





I'm Me, We're Us, and I'm Us: Tri-directional Contrastive Learning on Hypergraphs

T	目标	编码高阶结构信息，生成 区分度高、泛化性强 的节点和超边表示
I	输入	节点及其特征（ $1k-2w$ ）、关联矩阵*1
P	处理	1.增强原始超图，生成2个增强视图 2.编码2个增强视图，分别生成节点和超边表示 3.投影节点表示到损失计算空间 4.最小化 三向对比损失 ，以优化编码器
O	输出	节点表示

P	问题	现有方法仅使用节点级对比损失，只能捕捉到节点之间的简单关系
C	条件	无监督 信息
D	难点	如何选取负样本和构建对比损失，以充分利用超图的高阶信息
L	水平	AAAI Conference on Artificial Intelligence（2023 CCF-A）

• 超图增强

– 结合**节点初始特征掩码**与**超边成员掩码**，生成超图的**增强视图***2

- 按预设比例 p_f ，随机掩去所有节点的特征 X ，同视图中的节点特征被掩去**位置相同**
- 按预设比例 p_m ，随机掩去关联矩阵 H 中的非零值

• 超图编码器

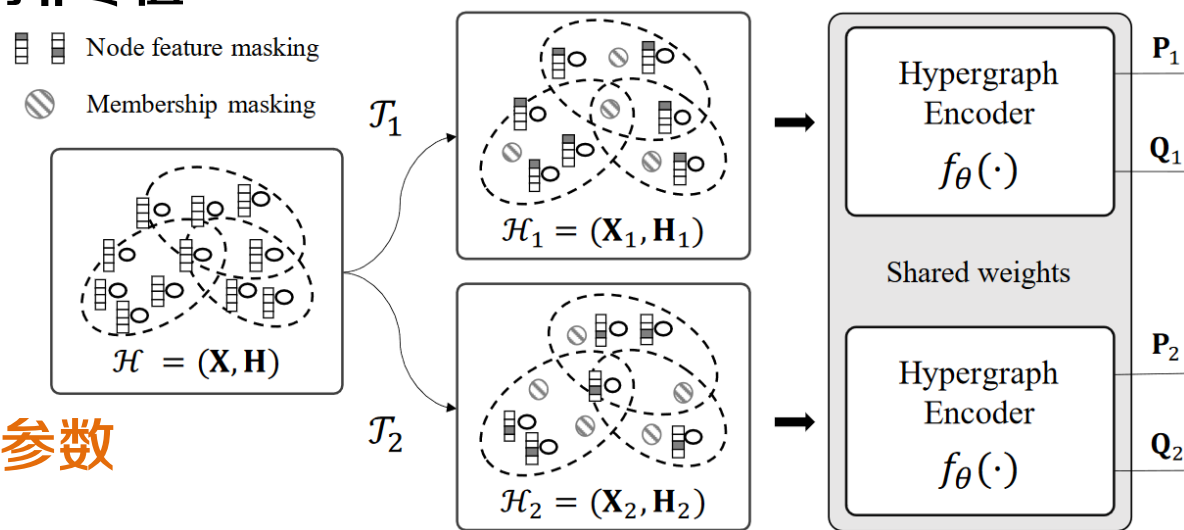
– 操作：分别为2个增强视图生成节点表示和超边表示（ P_1, P_2 和 Q_1, Q_2 ）

– **不限制**超图神经网络的选择

– 参数限制：2个视图的超图神经网络**共用参数**

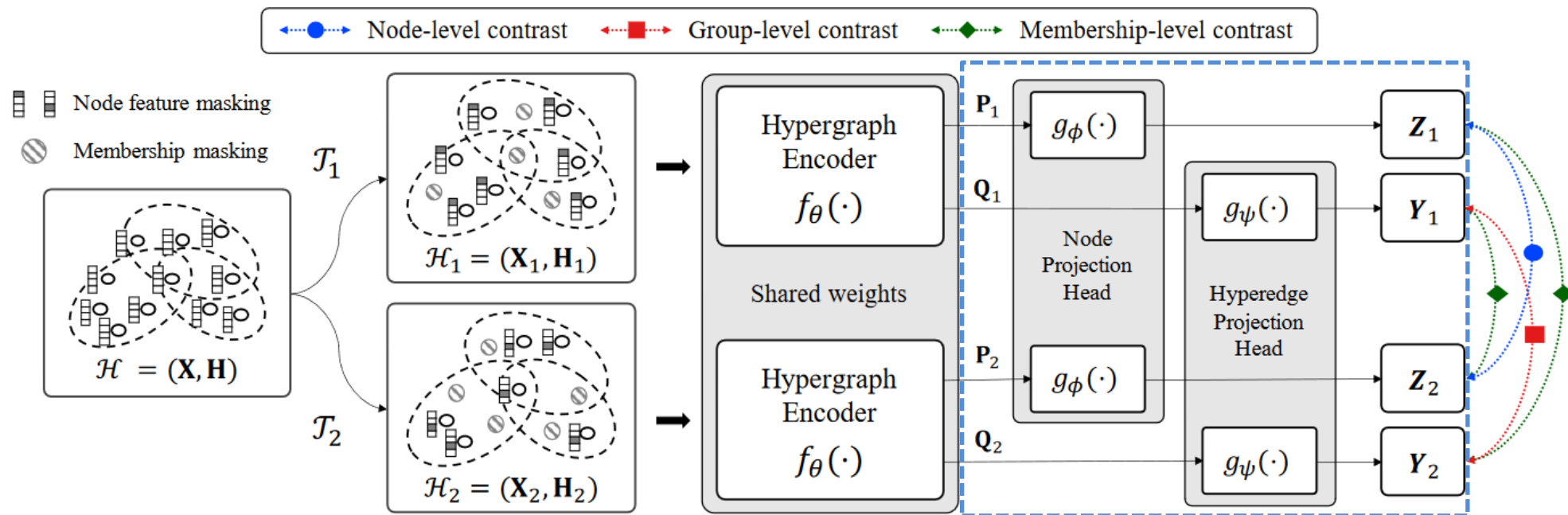
• 超边表示： $Q^{l+1} = \sigma(D_e^{-1} H^T P^l \Theta_e^l)$

• 节点表示： $P^{l+1} = \sigma(D_v^{-1} H \Omega Q^{l+1} \Theta_v^l)$



- 投影层

- 操作：将 P_1 , P_2 和 Q_1 , Q_2 分别投影到用于对比损失计算的 Z_1 , Z_2 和 Y_1 , Y_2
- 结构：节点投影层和超边投影层均由2层MLP和ELU激活函数组成
- 参数限制：2个视图的节点和超边投影层分别共用参数
- 作用：分离表示学习和对比目标；减少信息瓶颈；提升模型的鲁棒性



- 核心思想

- 最大化同一节点之间、同一超边之间以及每个超边及其成员节点之间的一致性
- 加权聚合3种对比损失作为优化目标: $L = L_n + \omega_g L_g + \omega_m L_m$

- 节点级对比

- 以第1个视图的节点*i* 的表示 $z_{1,i}$ 为锚点, 第2个视图中来自同一节点*i*的潜在表示 $z_{2,i}$ 作为正样本, 来自不同节点*k* ($k \neq i$)的潜在表示 $z_{2,k}$ 作为负样本
- 第1个视图的节点*i*为锚点的节点级对比损失为:

$$l_n(z_{1,i}, z_{2,i}) = -\log \frac{\exp(\text{sim}(z_{1,i}, z_{2,i})/\tau_n)}{\sum_{k=1}^{|V|} \exp(\text{sim}(z_{1,i}, z_{2,k})/\tau_n)}$$

- 考虑对称性, 整体的节点级对比损失为:

$$L_n = \frac{1}{2|V|} \sum_{i=1}^{|V|} \{l_n(z_{1,i}, z_{2,i}) + l_n(z_{2,i}, z_{1,i})\}$$

- 超边级对比——与节点级对比的思路相似

- 第1个视图的超边 j 为**锚点**的对比损失和整体的超边级对比损失分别为：

$$l_g(y_{1,j}, y_{2,j}) = -\log \frac{\exp(\text{sim}(y_{1,j}, y_{2,j})/\tau_g)}{\sum_{k=1}^{|E|} \exp(\text{sim}(y_{1,j}, y_{2,k})/\tau_g)}, \quad L_g = \frac{1}{2|E|} \sum_{j=1}^{|E|} \{l_g(y_{1,j}, y_{2,j}) + l_g(y_{2,j}, y_{1,j})\}$$

- 成员级对比

- 核心思想：在原始超图中，节点 i 是超边 j 的成员，则在2个增强视图中，二者的表示被鉴别为**具有成员关系的概率**应该越高

- 节点 i 和超边 j 分别为**锚点**，**不包含节点 i 的超边和不属于超边 j 的节点**都作为**负样本**，对比损失为：

$$l_m(z_{1,i}, y_{2,j}) = -\log \frac{\exp(D(z_{1,i}, y_{2,j})/\tau_m)}{\exp(D(z_{1,i}, y_{2,j})/\tau_m) + \sum_{k: i \notin k} \exp(D(z_{1,i}, y_{2,k})/\tau_m)} - \log \frac{\exp(D(z_{1,i}, y_{2,j})/\tau_m)}{\exp(D(z_{1,i}, y_{2,j})/\tau_m) + \sum_{k: k \notin j} \exp(D(z_{1,i}, y_{2,k})/\tau_m)}$$

- 整体的成员级损失： $L_m = \frac{1}{2|K|} \sum_{h_{i,j}=1} \{l_m(z_{1,i}, y_{2,j}) + l_m(z_{2,i}, y_{1,j})\}$

$D(,)$ ：将2个向量嵌入成一个1维向量，用作**鉴别函数**

- 数据集

- 共被引数据集: Cora-C、Citeseer 和 Pubmed
- 共同作者数据集: Cora-A 和 DBLP
- 计算机视觉和图形数据集: NTU2012 和 ModelNet40

	Cora-C	Citeseer	Pubmed	Cora-A	DBLP	NTU2012	ModelNet40
# Nodes	1,434	1,458	3,840	2,388	41,302	2,012	12,311
# Hyperedges	1,579	1,079	7,963	1,072	22,363	2,012	12,311
# Memberships	4,786	3,453	34,629	4,585	99,561	10,060	61,555
Avg. hyperedge size	3.03	3.20	4.35	4.28	4.45	5	5
Avg. node degree	3.34	2.37	9.02	1.92	2.41	5	5
Max. hyperedge size	5	26	171	43	202	5	5
Max. node degree	145	88	99	23	18	19	30
# Features	1,433	3,703	500	1,433	1,425	100	100
# Classes	7	6	3	7	6	67	40

- 对比方法

- 有监督方法*10，无监督方法*4

- 评价指标

- 将所有方法输出的表示用于节点分类和聚类任务
- 节点分类指标ACC；节点聚类指标NMI（归一化互信息）和F1

节点分类任务对比实验

- 实验设计：无监督方法先获取节点表示再进行监督分类训练
- 结论1：3种对比损失都可以提高性能
- 结论2：标准图对比因为**结构信息丢失**，在**平均超边尺寸**较大的数据集上效果较差

平均排名

	Method	Cora-C	Citeseer	Pubmed	Cora-A	DBLP	NTU2012	ModelNet40	A.R.↓
标准图网络	MLP	60.32 ± 1.5	62.06 ± 2.3	76.27 ± 1.1	64.05 ± 1.4	81.18 ± 0.2	65.17 ± 2.3	93.75 ± 0.6	14.4
	GCN*	77.11 ± 1.8	66.07 ± 2.4	82.63 ± 0.6	73.66 ± 1.3	87.58 ± 0.2	71.17 ± 2.4	91.67 ± 0.2	11.1
	GAT*	77.75 ± 2.1	67.62 ± 2.5	81.96 ± 0.7	74.52 ± 1.3	88.59 ± 0.1	70.94 ± 2.6	91.43 ± 0.3	10.1
超图网络	HGNN	77.50 ± 1.8	66.16 ± 2.3	83.52 ± 0.7	74.38 ± 1.2	88.32 ± 0.3	72.03 ± 2.4	92.23 ± 0.2	9
	HyperConv	76.19 ± 2.1	64.12 ± 2.6	83.42 ± 0.6	73.52 ± 1.0	88.83 ± 0.2	72.62 ± 2.6	91.84 ± 0.1	10.1
	HNHN	76.21 ± 1.7	67.28 ± 2.2	80.97 ± 0.9	74.88 ± 1.6	86.71 ± 1.2	71.45 ± 3.2	92.96 ± 0.2	10.3
	HyperGCN	64.11 ± 7.4	59.92 ± 9.6	78.40 ± 9.2	60.65 ± 9.2	76.59 ± 7.6	46.05 ± 3.9	69.23 ± 2.8	16.3
	HyperSAGE	64.98 ± 5.3	52.43 ± 9.4	79.49 ± 8.7	64.59 ± 4.3	79.63 ± 8.6	OOT	OOT	16
	UniGCN	77.91 ± 1.9	66.40 ± 1.9	84.08 ± 0.7	77.30 ± 1.4	90.31 ± 0.2	73.27 ± 2.7	94.62 ± 0.2	5.9
	AllSet	76.21 ± 1.7	67.83 ± 1.8	82.85 ± 0.9	76.94 ± 1.3	90.07 ± 0.3	75.09 ± 2.5	96.85 ± 0.2	6.3
标准图对比学习	Node2vec*	70.99 ± 1.4	53.85 ± 1.9	78.75 ± 0.9	58.50 ± 2.1	72.09 ± 0.3	67.72 ± 2.1	84.94 ± 0.4	16
	DGI*	78.17 ± 1.4	68.81 ± 1.8	80.83 ± 0.6	76.94 ± 1.1	88.00 ± 0.2	72.01 ± 2.5	92.18 ± 0.2	8.3
	GRACE*	79.11 ± 1.7	68.65 ± 1.7	80.08 ± 0.7	76.59 ± 1.0	OOM	70.51 ± 2.4	90.68 ± 0.3	11
超图对比学习	S ² -HHGR	78.08 ± 1.7	68.21 ± 1.8	82.13 ± 0.6	78.15 ± 1.1	88.69 ± 0.2	73.95 ± 2.4	93.26 ± 0.2	6.6
Unsupervised	Random-Init	63.62 ± 3.1	60.44 ± 2.5	67.49 ± 2.2	66.27 ± 2.2	76.57 ± 0.6	74.39 ± 2.6	96.29 ± 0.3	12.9
	TriCL-N	80.23 ± 1.2	70.28 ± 1.5	83.44 ± 0.6	81.94 ± 1.1	90.88 ± 0.1	75.20 ± 2.6	97.01 ± 0.2	3.3
	TriCL-NG	81.45 ± 1.2	71.38 ± 1.2	83.68 ± 0.7	82.00 ± 1.0	90.94 ± 0.1	75.25 ± 2.5	97.02 ± 0.1	2
	TriCL	81.57 ± 1.1	72.02 ± 1.2	84.26 ± 0.6	82.15 ± 0.9	91.12 ± 0.1	75.23 ± 2.4	97.08 ± 0.1	1.1

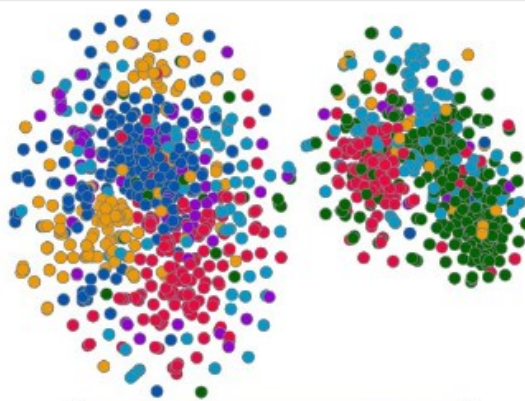
节点分类任务消融实验

- 结论1: 节点级对比损失对节点表示区分度的影响最大, 成员级对比影响最小
- 结论2: 单独使用成员级对比损失会导致模型崩溃

$\mathcal{L}_n$	$\mathcal{L}_g$	$\mathcal{L}_m$	Cora-C	Citeseer	Pubmed	Cora-A	DBLP	NTU2012	ModelNet40	A.R.↓
✓	-	-	80.23 ± 1.2	70.28 ± 1.5*	83.44 ± 0.6	81.94 ± 1.1	90.88 ± 0.1	75.20 ± 2.6	97.01 ± 0.2	4
-	✓	-	79.69 ± 1.6	71.02 ± 1.3*	80.20 ± 1.3	78.98 ± 1.4	88.60 ± 0.2	74.41 ± 2.6	96.66 ± 0.2	5.6
-	-	✓	76.76 ± 1.8	63.98 ± 2.0	79.86 ± 0.9	76.77 ± 1.1	63.95 ± 7.2	73.11 ± 2.8	96.57 ± 0.2	7
✓	✓	-	81.45 ± 1.2	71.38 ± 1.4	83.68 ± 0.7	82.00 ± 1.0	90.94 ± 0.1	75.25 ± 2.5	97.02 ± 0.1	2.3
✓	-	✓	80.49 ± 1.3	70.46 ± 1.5	83.98 ± 0.7	81.62 ± 1.0	90.75 ± 0.1	75.12 ± 2.5	97.03 ± 0.1	3.6
-	✓	✓	80.80 ± 1.1	71.73 ± 1.4	82.81 ± 0.7	80.24 ± 1.0	90.17 ± 0.1	73.76 ± 2.5	96.74 ± 0.1	4.4
✓	✓	✓	81.57 ± 1.1	72.02 ± 1.4	84.26 ± 0.6	82.15 ± 0.9	91.12 ± 0.1	75.23 ± 2.4	97.08 ± 0.1	1.1

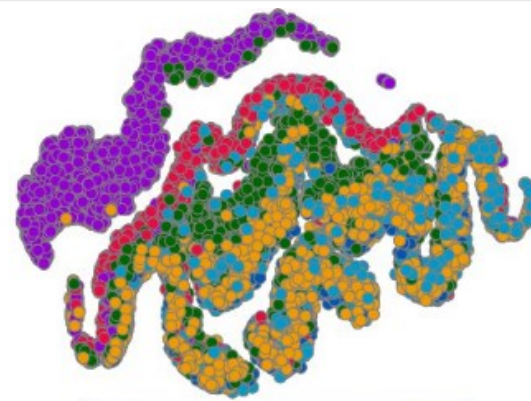
模型崩溃分析

- 缺少节点和超边对比, 使得表示没有区分度



Silhouette score: -0.031

(a) Citeseer



Silhouette score: 0.002

(b) DBLP

节点聚类任务对比实验

- 实验设计：K-means聚类
- 聚类性能平均排名第一，证明TriCL对不同下游任务的泛化效果

Method	Cora-C		Citeseer		Pubmed		Cora-A		DBLP		NTU2012		ModelNet40		A.R.↓
	NMI↑	F1↑	NMI↑	F1↑	NMI↑	F1↑	NMI↑	F1↑	NMI↑	F1↑	NMI↑	F1↑	NMI↑	F1↑	
features	20.0	28.8	21.5	36.1	19.5	53.4	17.2	29.2	37.0	47.3	81.7	69.0	90.6	86.5	4.3
Node2vec*	39.1	44.5	24.5	38.5	23.1	40.1	16.0	34.1	32.4	37.8	78.3	57.7	72.9	53.1	5.0
DGI*	54.8	60.1	40.1	51.7	30.4	53.0	45.2	52.5	58.0	57.7	79.6	61.7	85.0	73.7	2.9
GRACE*	44.4	45.6	33.3	45.7	16.7	41.9	37.9	43.3	16.7	41.9	74.6	47.5	79.4	59.9	4.7
S ² -HHGR	51.0	56.8	41.1	53.1	27.7	53.2	45.4	52.3	60.3	62.7	82.7	71.2	91.0	90.6	2.5
TriCL	54.5	60.6	44.1	57.4	30.0	51.7	49.8	56.7	63.1	63.0	83.2	71.5	95.7	94.7	1.7

- 超图对比学习加入引文网络的全局信息，增强了聚类全局结构的一致性，NMI更高；
- 但是加入引文信息可能导致某些文献被归入相关性更强但类别不同的簇中，从而使得具体类别的匹配度下降，F1反而降低



Collaborative contrastive learning for hypergraph node classification

T	目标	利用对比学习进行针对 节点分类 任务的超图表示学习
I	输入	节点及其特征、关联矩阵*1
P	处理	1.根据关联矩阵计算 标准图 的邻接矩阵 2. 交替 使用超图和标准图增强节点表示，进行 协作对比学习 3.预测节点标签 4.有监督的预测损失和对比损失 联合优化
O	输出	节点预测标签

P	问题	现有方法的 超图增强技术 对结构信息利用不足，可能丢失有价值信息
C	条件	针对 有监督 情况下的节点分类任务
D	难点	如何构建节点的增强表示
L	水平	Pattern Recognition（2024 SCI 一区）

• 标准图转化

- 计算表示节点之间关系的邻接矩阵 $A = HH^T$

$$\begin{array}{|c|c|c|c|} \hline 1 & 0 & 1 & 0 \\ \hline 0 & 1 & 1 & 1 \\ \hline 0 & 1 & 0 & 1 \\ \hline \end{array} \times \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 1 \\ \hline 1 & 1 & 0 \\ \hline 0 & 1 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 2 & 1 & 0 \\ \hline 1 & 3 & 2 \\ \hline 0 & 2 & 2 \\ \hline \end{array}$$

H H^T A

节点和哪些超边有关 H^T 节点和哪些节点有关
超边和哪些节点有关

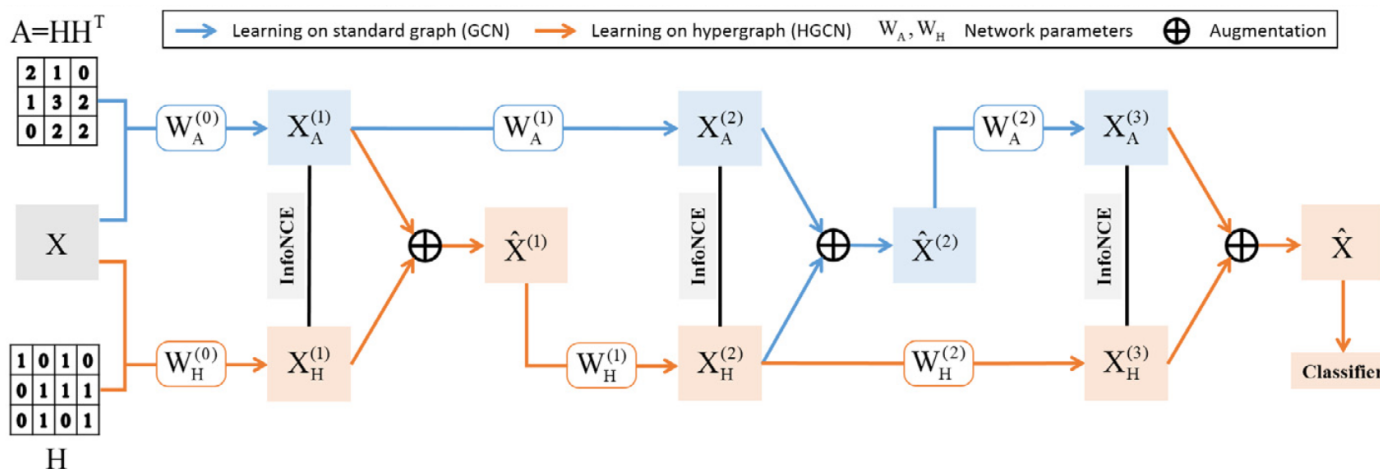
• 协作对比学习

- 超图编码器HGCN: $X_H^{(l+1)} = \sigma(D_v^{-1/2} H \Omega D_e^{-1} H^T D_v^{-1/2} X_H^{(l)} W_H^{(l)})$
- 标准图编码器GCN: $X_A^{(l+1)} = \sigma(D^{-1/2} A D^{-1/2} X_A^{(l)} W_A^{(l)})$
- 在3层网络中，将超图与标准图分别用作锚点视图和增强视图，交替进行对比与增强，将第3层网络的增强表示用于节点分类任务

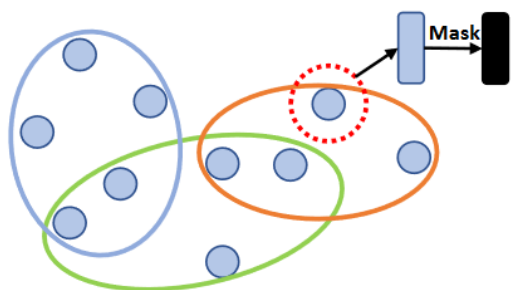
• 模型优化

- 节点预测交叉熵损失 L_{ce}
- 3层对比损失 $L_{cl}^{(i)}$
- 模型优化目标

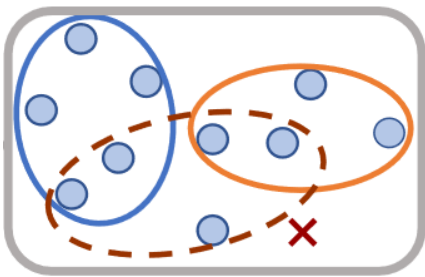
$$L = L_{ce} + \lambda \sum_{i=1}^3 L_{cl}^{(i)}$$



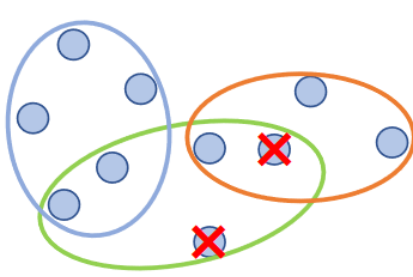
- 核心思想
 - 将高阶的超边关系转换为标准图中的边关系，使用标准图作为**增强视图**与超图进行对比学习
- 优势分析——相比更改特征与结构的增强方法
 - 保持全局结构一致性：保留超图中的**所有关系信息**，避免更改操作引入的**噪声**
 - 有效视图生成：标准图生成的视图能更**清晰直接地体现节点间关系**



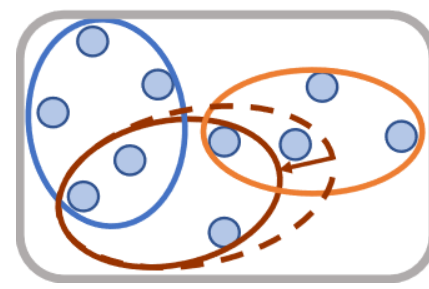
节点初始特征掩码



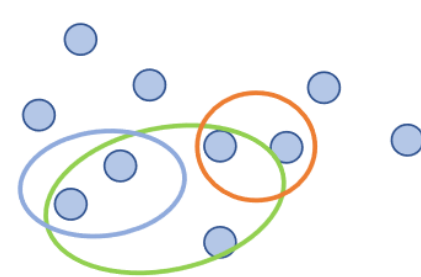
超边删除



节点删除

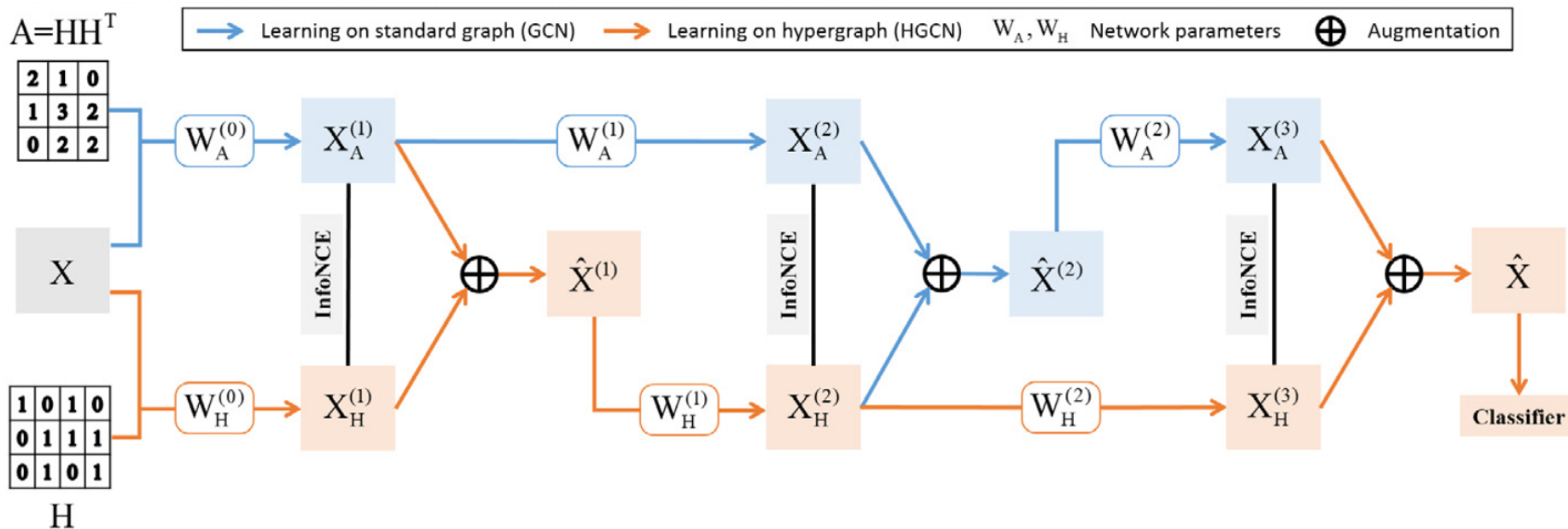


超边成员掩码



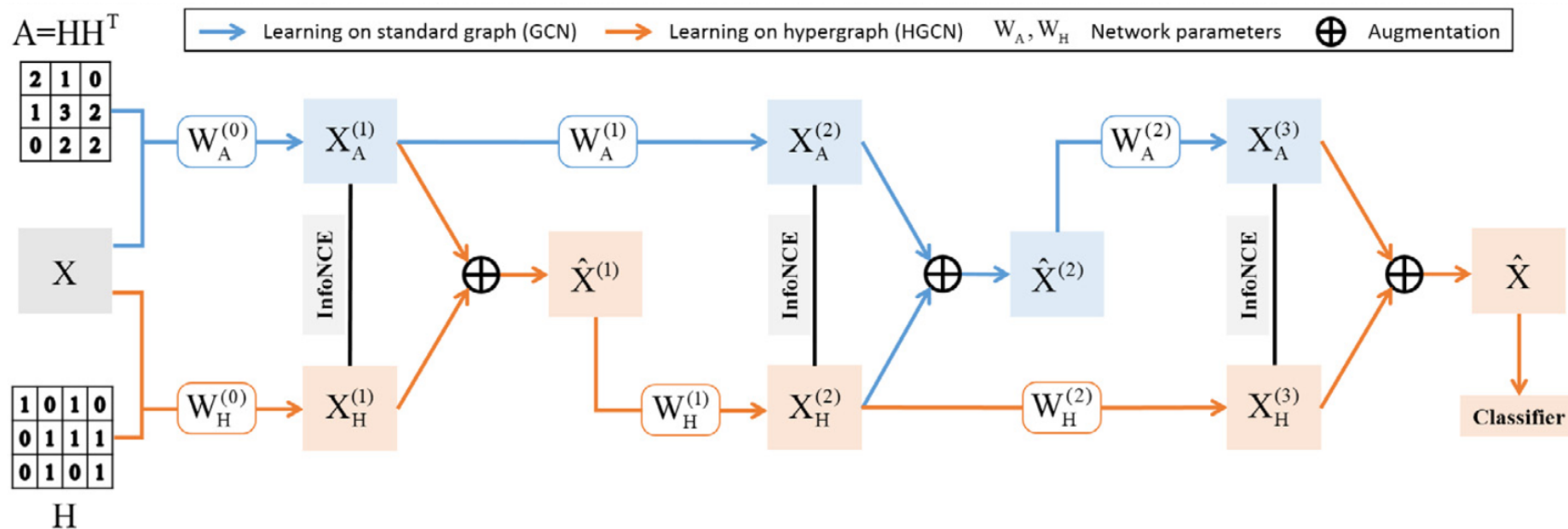
随机游走生成子图

- 协作增强
 - 在每一层网络中，将GCN和HGCN的视图 $X_A^{(i)}$ 和 $X_H^{(i)}$ 进行拼接以获得增强的节点表示 $\hat{X}^{(i)}$ ，替换二者之一，作为下一层GCN和HGCN的输入
- 对比损失
 - 在每层网络中，HGCN和GCN视图中的节点交替作为锚点，计算节点级损失函数



• 核心思想

- 将超图与标准图分别用作锚点视图和增强视图，交替进行**对比与增强**
- 通过对比，优化HGCN和GCN参数，提取和强化**节点间关系与高阶的结构信息**，得到高质量的表示
- 通过增强，使得每个视图能从另一视图中**获得更丰富的信息**，提高最终**分类性能**



- 数据集

- 共被引数据集: Cora-C、Citeseer 和 Pubmed
- 共同作者数据集: Cora-A 和 DBLP

	Cora-C	Citeseer	Pubmed	Cora-A	DBLP
# Nodes	1,434	1,458	3,840	2,388	41,302
# Hyperedges	1,579	1,079	7,963	1,072	22,363
# Memberships	4,786	3,453	34,629	4,585	99,561
Avg. hyperedge size	3.03	3.20	4.35	4.28	4.45
Avg. node degree	3.34	2.37	9.02	1.92	2.41
Max. hyperedge size	5	26	171	43	202
Max. node degree	145	88	99	23	18
# Features	1,433	3,703	500	1,433	1,425
# Classes	7	6	3	7	6

- 对比方法

- 非图网络方法*2, 标准图网络*1, 超图网络*8, 超图对比学习方法*1

- 评价指标

- 节点分类指标ACC

非图网络: NN、SVM

标准图网络: GCN

超图对比学习: TriCL

超图网络: FastHyperGCN、HyperGCN、HGNN、HNHN、UniGCN、UniGAT、HCNH、HCoN

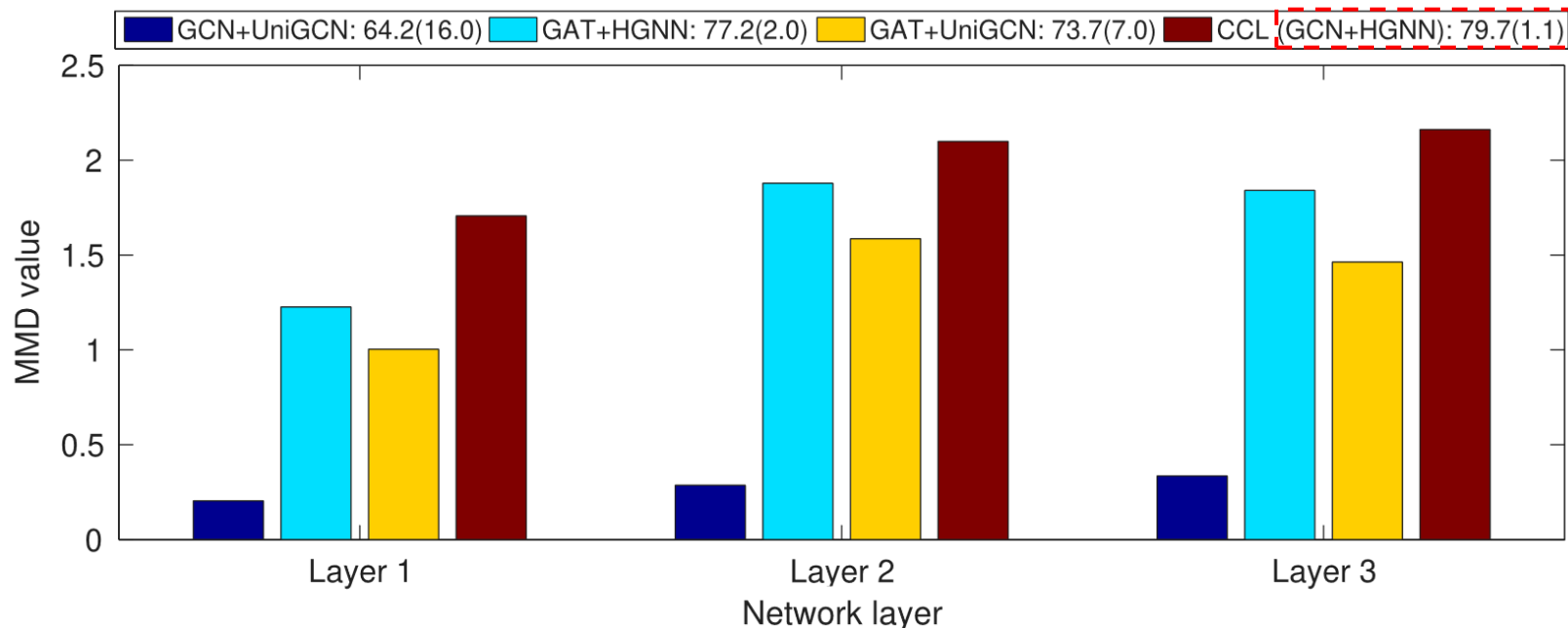
实验分析

- NN、SVM效果最差，说明图结构信息不能被省略
- 利用对比学习使标准图和超图的卷积**协作**有助于增强表示实例的能力
- 在每个网络层中进行对比会产生**信息更丰富**的表示
- CCL在Cora-C数据集上的效果可能受限于GCN的性能

	Method	Co-citation			Co-authorship	
		Citeseer	Cora	Pubmed	Cora	DBLP
非图网络	NN	27.0 $\pm$ 9.0	24.5 $\pm$ 5.7	41.2 $\pm$ 2.7	28.9 $\pm$ 8.1	20.6 $\pm$ 6.2
	SVM	21.0 $\pm$ 1.3	32.6 $\pm$ 5.2	41.3 $\pm$ 4.1	52.1 $\pm$ 1.1	27.2 $\pm$ 0.1
标准图网络	GCN [27]	63.0 $\pm$ 1.9	72.1 $\pm$ 1.9	69.6 $\pm$ 6.3	40.0 $\pm$ 3.3	83.2 $\pm$ 2.6
	FastHyperGCN [42]	56.1 $\pm$ 11.2	50.4 $\pm$ 11.7	54.4 $\pm$ 10.0	45.2 $\pm$ 12.9	70.5 $\pm$ 14.3
超图网络	HyperGCN [42]	54.7 $\pm$ 9.8	52.2 $\pm$ 11.4	60.0 $\pm$ 10.7	55.0 $\pm$ 0.9	71.3 $\pm$ 1.2
	HGNN [12]	61.1 $\pm$ 2.2	76.9 $\pm$ 1.8	63.3 $\pm$ 2.2	58.2 $\pm$ 0.3	77.6 $\pm$ 0.4
	HNHN [25]	64.8 $\pm$ 1.6	71.3 $\pm$ 1.9	75.9 $\pm$ 1.5	63.9 $\pm$ 0.8	85.1 $\pm$ 0.2
	UniGCN [43]	70.9 $\pm$ 1.0	78.3 $\pm$ 1.7	78.8 $\pm$ 1.7	61.4 $\pm$ 0.9	88.3 $\pm$ 0.2
	UniGAT [43]	70.7 $\pm$ 1.0	78.5 $\pm$ 1.8	78.7 $\pm$ 1.7	60.6 $\pm$ 0.9	87.8 $\pm$ 0.3
	HCNH [26]	71.4 $\pm$ 1.2	78.5 $\pm$ 1.7	77.1 $\pm$ 3.6	65.9 $\pm$ 0.5	–
	HCoN [8]	71.2 $\pm$ 2.7	79.0 $\pm$ 1.5	80.4 $\pm$ 1.1	66.5 $\pm$ 0.5	88.0 $\pm$ 0.1
	TriCL [36]	71.1 $\pm$ 0.8	81.0 $\pm$ 1.0	80.3 $\pm$ 1.6	64.2 $\pm$ 0.5	88.5 $\pm$ 0.1
超图对比学习	CCL (ours)	71.9 $\pm$ 0.8	79.7 $\pm$ 1.1	81.9 $\pm$ 0.9	64.6 $\pm$ 0.5	88.6 $\pm$ 0.1

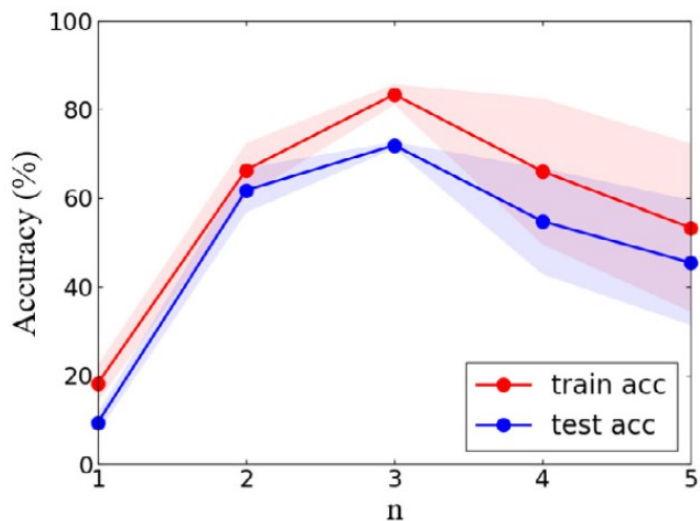
- 编码器变体实验结果

- 实验设计：使用不同的图神经网络组合，分别计算3层中的增强表示的分类性能以及标准图和超图表示的MMD值
- 指标MMD：2个视图表示的最大平均差异，代表2个视图表示的多样性
- 结论：高多样性会带来更好的分类性能，这说明越能提供同一节点的不同视角，越能学习有效的表示

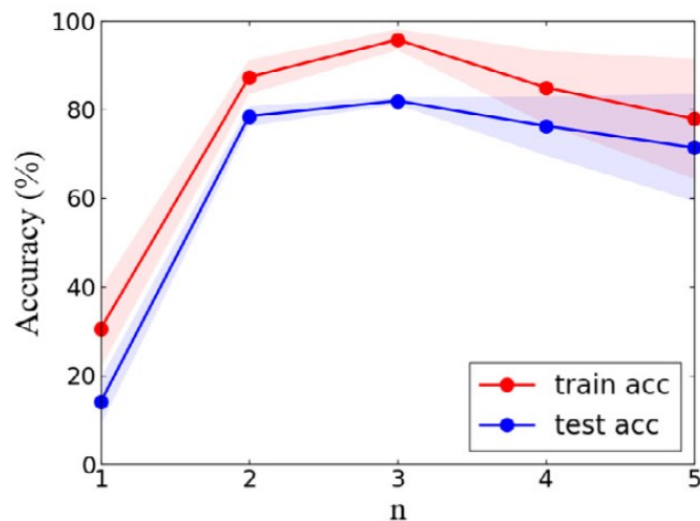


- 模型层数实验结果

- 实验设计：将协作对比层数分别设为1、2、3、4、5，每种设置进行100次实验
- 结论1：3层模型实现了最佳性能，明显优于单层模型，表明在每一层中执行对比的有效性
- 结论2：对于深度超过3层的模型，性能下降，推测随着模型深度的增加，训练会变得困难，并且参数数量的增加可能会导致过拟合



(a) Results w.r.t. n on Citeseer



(b) Results w.r.t. n on Pubmed



特点总结与未来展望

• 特点总结

算法	TriCL	CCL
优势	1.从3个层次构建对比损失，针对性地学习超图的 高阶关系 2.以自监督方式学习超图表示，可以很好地 泛化 到多种下游任务	1.标准图转化 不会引入噪声 ，且直接体现节点间关系，更具对比意义 2.协作对比方式通过多层对比与增强，能更充分地提取和聚合有用信息
劣势	1.涉及的 超参数过多 ，需要大量实验寻找并验证最优设置 2.多项对比损失导致 训练时间相对过长	1.多层对比损失导致 训练时间相对过长 2.性能受限于基础方法，通用性不足 3.协作对比方式 可解释性不足

• 未来展望

- 课程负采样：模仿人类课程中的学习顺序，以一种**由易到难**的方式提供负样本以进行对比学习，进一步提高模型的**泛化能力和收敛速度**
- 优化对比损失：基于同类样本的中心点进行对比，以高效率处理**大规模数据集**

- [1] Lee D, Shin K. I'm me, we're us, and i'm us: Tri-directional contrastive learning on hypergraphs[C]. Proceedings of the AAAI Conference on Artificial Intelligence. 2023, 37(7): 8456-8464.**
- [2] Wu H, Li N, Zhang J, et al. Collaborative contrastive learning for hypergraph node classification[J]. Pattern Recognition, 2024, 146: 109995.**
- [3] Wei T, You Y, Chen T, et al. Augmentations in hypergraph contrastive learning: Fabricated and generative[J]. Advances in neural information processing systems, 2022, 35: 1909-1922.**

知人者智，自知者明。胜人者有力，自胜者强。知足者富。强行者有志。不失其所者久。死而不亡者，寿。

谢谢！

