

基于分数阶图神经扩散的跨频域对齐对比学习方法

王友清¹ 徐世龙¹ 赵天祥¹ 王宇晨¹ 辛梦媛¹ 张琦² 苏烨³ 郭继鹏¹

摘要 图对比学习 (GCL) 作为一种强大的自监督表示学习范式, 能够通过有效利用无标签数据来增强半监督学习中的表示判别性和泛化能力. 然而, 现有的 GCL 方法在学习判别性嵌入表示以及图数据增强过程中实现对比多样性与语义一致性之间的平衡方面存在困难, 这导致在构建增强视图时关键信息的丢失. 为解决这些挑战, 提出一种新颖的跨频域对齐对比学习 (CfACL) 框架, 利用分数阶图神经扩散 (FGND) 进行图节点表示学习. FGND 利用切比雪夫多项式分数阶微分方程实现图信号中多阶邻域信息的远程扩散, 缓解过平滑问题并提高图嵌入表示的判别能力. 随后, 通过高频和低频滤波器分别构建两种不同的 FGND 形式, 形成自然的增强对比视图, 避免了随机增强引起的内在结构坍塌和语义漂移. CfACL 方法将高频滤波分量转换到低频域, 并在镜像的虚拟谱空间中进行对比学习, 从而能够在全局一致的语义空间中吸收有益的高频细节, 为下游任务生成全面的表示. 在同配性和异配性基准图数据集上的大量节点分类实验结果验证了所提方法的有效性.

关键词 图对比学习; 分数阶图扩散; 高频滤波; 低频滤波; 跨频域对齐

引用格式 王友清, 徐世龙, 赵天祥, 王宇晨, 辛梦媛, 张琦, 苏烨, 郭继鹏. 基于分数阶图神经扩散的跨频域对齐对比学习方法. 自动化学报, 2026, 52(6): 1189–1200

DOI 10.16383/j.aas.c250604 **CSTR** 32138.14.j.aas.c250604

Fractional-order Graph Neural Diffusion for Cross-frequency Alignment Contrastive Learning

WANG You-Qing¹ XU Shi-Long¹ ZHAO Tian-Xiang¹ WANG Yu-Chen¹
XIN Meng-Yuan¹ ZHANG Qi² SU Ye³ GUO Ji-Peng¹

Abstract Graph contrastive learning (GCL), a powerful self-supervised representation learning paradigm, could enhance representation discriminability and generalization in semi-supervised learning by effectively leveraging unlabeled data. However, existing GCL methods struggle to learn discriminative embedding and achieve better balance between contrastive diversity and semantic invariance during graph data augmentation, inevitably leading to the critical information loss when constructing augmented views. To address these challenges, this paper proposes a novel cross-frequency alignment contrastive learning (CfACL) framework with fractional-order graph neural diffusion (FGND) for graph node representation learning. The FGND leverages Chebyshev polynomial fractional differential equations to achieve the long-range diffusion of multi-order neighboring information, alleviating over-smoothing and improving the discriminability of graph representation. Then, two distinct FGNDs are characterized by high-frequency and low-frequency filters to form natural augmented contrastive views, avoiding the intrinsic structure collapse and semantic shift caused by random augmentation. The CfACL transforms high-frequency filtered components into the low-frequency domain and achieves the contrastive learning in mirrored virtual spectral space, which is capable of absorbing beneficial high-frequency details in a globally consistent semantic space and results in comprehensive representation for downstream tasks. Extensive node classification experiments demonstrate the effectiveness of the proposed method across homophilic and heterophilic benchmark graph datasets.

Keywords graph contrastive learning; fractional-order graph diffusion; high-frequency filtering; low-frequency filtering; cross-frequency alignment

Citation Wang You-Qing, Xu Shi-Long, Zhao Tian-Xiang, Wang Yu-Chen, Xin Meng-Yuan, Zhang Qi, Su Ye, Guo Ji-Peng. Fractional-order graph neural diffusion for cross-frequency alignment contrastive learning. *Acta Automatica Sinica*, 2026, 52(6): 1189–1200

收稿日期 2025-11-05 录用日期 2026-01-20

Manuscript received November 5, 2025; accepted January 20, 2026

国家重点研发项目 (2024YFB3311405), 国家自然科学基金 (6225303, 62403043) 资助

Supported by National Key Research and Development Program of China (2024YFB3311405) and National Natural Science Foundation of China (6225303, 62403043)

本文责任编辑 罗彪

Recommended by Associate Editor LUO Biao

1. 北京化工大学信息科学与技术学院 北京 100029 2. 山东农业大学信息科学与工程学院 泰安 271018 3. 中国科学院深圳先进技术研究院 深圳 518055

1. College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029 2. School of Information Science and Engineering, Shandong Agricultural University, Taian 271018 3. Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055

图结构数据是现实生活与工业活动中广泛存在的一类数据形式,如知识图谱^[1]、社交网络^[2]、多智能体系统^[3]、生物信息^[4]等,这些都可以用图结构来建模和表示.图结构数据自身的复杂性使得对其进行分析和处理是困难的^[5-7].近年来,图神经网络(graph neural networks, GNNs)因其对复杂拓扑图强大的结构表示学习能力而受到广泛关注,在节点分类^[8]、聚类^[9]、链路预测^[10]等任务中取得优异性能.图节点分类是一个经典的半监督任务,使用 GNNs 学习节点嵌入表示并利用交叉熵损失简单地将学习到的节点表示与标签类别进行匹配,实现无标签节点的类别预测^[11].然而,由于过平滑问题和标注样本的有限性,现有方法难以实现对无标签节点信息的有效传播,限制了分类模型的泛化能力.交叉熵损失仅对分类决策边界起作用,缺乏对特征表示本身的结构化约束,即现有方法忽视了无标签数据在提高结构判别性方面的潜力.近年来,对比学习通过以自监督的方式分析增强视图之间的相似性和差异性来提高表示判别性,使得模型能够辨别数据中丰富的底层结构模式^[12].对比学习进一步被扩展到图结构数据,即图对比学习(graph contrastive learning, GCL)^[13-14]被广泛关注.

数据增强和对比目标构建是对比学习的两个关键因素.根据增强方式,现有的 GCL 方法可分为两类:基于增强的方法和无增强方法.基于增强的方法^[15-18]通过对结构或属性特征添加扰动来构建对比学习的增强视图,这类方法对增强操作较敏感且容易产生语义漂移;无增强方法^[19-20]通过不同的特征提取器直接从原始图构建对比视图,避免了复杂且敏感的增强算子,简单有效.本文主要关注无增强的 GCL 方法,有两个必要因素需要考虑:1) 如何通过无增强的方式生成独特且高质量的对比视图? 2) 如何有效建模对比视图之间的语义关系以实现有益的自监督学习?

图扩散被广泛用于实现图上的全局消息传递,例如,基于常微分方程(ordinary differential equations, ODE)的图扩散^[21-23].与图卷积的离散消息传递方式不同,其本质是具有连续时间演化机制的热扩散,相应的节点特征更新与图结构的自然演化模式更加匹配^[19].然而,基于整数阶 ODE 的图扩散严格遵循马尔可夫性和局部邻接传播特性,难以捕获长程依赖并实现复杂的非局部状态传播.为解决这些问题,基于 ODE 的图扩散被推广到分数阶微分方程(fractional-order differential equations, FDE)^[24],引入分数阶微分算子 $\frac{d^\alpha}{dt^\alpha}$,其中 $\alpha \in (0, 1]$.Caputo 分数阶导数常被用于 FDE,其是一种具有

历史记忆特性的微分算子,能够有效建模长程依赖关系,因此,基于 FDE 的扩散可以融合来自状态空间的全部历史信息^[25].特别地,适当调整 $\alpha \in (0, 1]$ 可以控制历史信息和当前信息之间的权重差异.与 GNNs 类似,利用归一化拉普拉斯矩阵进行谱域信号滤波,将其作为图上的动态算子来更新节点特征.因此,基于 FDE 的图扩散通过探索时间和空间信息来学习图嵌入表示,可以作为无增强 GCL 方法的强大特征提取器.

大多数 GNNs 以及图扩散方法本质上使用低通滤波器来强调低频信息、探索节点的平滑表示,粗略地将高频信息视为噪声信息^[14, 26].实际上,高频信息并非完全是噪声,其可以记忆节点特定信息以提高节点表示的判别性.高低频域信息两者之间具有天然对立关系,可以作为对比学习的一种对比视图.从滤波器特性的角度来看,它们可以被粗略地视为负对比样本对,这违反了不同视图表示之间的互信息最大化原则,导致了在潜在特征空间中的语义差异.因此,对于有效的图对比学习,保持低频和高频表示之间的互补性和语义一致性是非常重要的.

本研究提出一种跨频域对齐对比学习(cross-frequency alignment contrastive learning, CfACL)框架,其中改进的分数阶图神经扩散(fractional-order graph neural diffusion, FGND)作为强大的图嵌入表示编码器,以低频与高频表示作为对比目标实现自监督学习,提高表示判别性.为学习更合适的滤波方式并为无增强 GCL 获得判别性表示,使用切比雪夫多项式 FGND 来探索丰富的邻域信息,捕捉长程依赖和复杂传播模式.此外,切比雪夫多项式 FGND 易于推广,通过设置不同形式的可学习滑动余弦参数可以构建原始图信号的低频和高频滤波表示,实现对比表示增强.进一步,为避免直接最大化对比高低频视图间差异导致的语义漂移,CfACL 将高频信息映射到低频空间,并在语义一致的低频隐空间内将其与原始低频表示对齐,同时实现表示多样性和语义一致性.主要贡献如下:

- 提出一种简单而有效的自监督学习框架 CfACL, 直接利用低频和高频滤波器作为无增强的对比视图,并在语义对齐的虚拟空间中进行对比学习;
- 提出切比雪夫多项式 FGND 作为图信号消息传递方式,通过差异性参数化构造对应低频和高频滤波的两个 FGND 变体作为 CfACL 框架的图扩散编码器,获取对比视图;
- 在实际同配性和异配性图数据集上的大量实验证明了所提方法 CfACL 在分类性能和判别性表示学习方面的优越性.

1 预备知识

1.1 符号定义

给定一个包含 N 个节点 $\mathcal{V} = \{v_1, \dots, v_N\}$ 的无向图 $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X})$. 其中 \mathcal{E} 是边的集合, 当对应节点 v_i 和 v_j 连接时有 $e_{ij} \in \mathcal{E}$; $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbf{R}^{N \times d}$ 表示节点属性特征矩阵, d 是 \mathbf{x}_i 的特征维度. 节点对连接关系由拓扑邻接矩阵 $\mathbf{A} = [a_{ij}] \in \{0, 1\}^{N \times N}$ 表示, 如果 $e_{ij} \in \mathcal{E}$, 则 $a_{ij} = 1$, 否则 $a_{ij} = 0$. 归一化拉普拉斯矩阵定义为 $\tilde{\mathbf{L}} = \mathbf{I} - \tilde{\mathbf{A}}$, 其中 $\tilde{\mathbf{A}} = \mathbf{D}^{-1/2} \mathbf{A} \mathbf{D}^{-1/2}$ 是归一化邻接矩阵, $\mathbf{D} = \text{diag}\{d_1, \dots, d_N\}$ 是度矩阵, 满足 $d_i = \sum_{j=1}^N a_{ij}$.

图节点表示学习对于下游任务是重要的, 旨在通过训练一个编码器 f_θ 来获取原始图结构 \mathcal{G} 的判别嵌入表示 $\mathbf{Z} = f_\theta(\mathbf{A}, \mathbf{X}) \in \mathbf{R}^{N \times d'}$, 其中 θ 是可训练参数. 近年来, 由于对比学习与下游任务相互独立, 其为判别性表征学习提供了强大的自监督途径. 具有代表性的无增强 GCL 方法^[27-28] 利用不同的编码器 f_{θ_1} 和 f_{θ_2} 来构建对比视图, 并通过合适的对比策略在它们之间实现自监督, 从而探索互信息互补性和语义一致性. 本研究主要提出简单有效的编码器和对比机制, 实现深层次的自监督图对比学习. 在同配性和异配性数据的图节点分类任务上验证了所提方法 CfACL 的有效性.

1.2 图扩散模型

经典的 GNNs 方法利用邻域信息聚合方式来实现图上的消息传递, 学习判别性嵌入表示^[29]. 例如, 对于经典的图卷积网络 (graph convolutional networks, GCN), 第 $l+1$ 层的消息传递方式为 $\mathbf{Z}^{(l+1)} = \tilde{\mathbf{A}} \mathbf{Z}^{(l)}$, 其中 $\mathbf{Z}^{(0)} = \mathbf{X}$. 为了实现连续的消息传递且缓解深度 GNNs 面临的过平滑问题, Chamberlain 等^[22] 提出用于信息传播的连续热力学图扩散过程, 其中代表性的整数阶 ODE 图扩散模型可以表述为:

$$\frac{d}{dt} \mathbf{Z}(t) = \mathcal{F}(\mathbf{W}(\mathbf{Z}(t)), \mathbf{Z}(t)) \quad (1)$$

其中, t 是连续时间参数, 对应于 GNNs 中的网络层索引; 初始状态为 $\mathbf{Z}(0) = \mathbf{X}$; \mathcal{F} 是定义在图上的动力学算子. 在经典的图扩散方法中^[22], 动力学算子 \mathcal{F} 有两种形式, 包括线性形式 $\mathcal{F} := -\tilde{\mathbf{L}} \mathbf{Z}(t)$ 和非线性形式 $\mathcal{F} := (\mathbf{W}(\mathbf{Z}(t)) - \mathbf{I}) \mathbf{Z}(t)$, 这里 $\mathbf{W}(\mathbf{Z}(t))$ 表示一个关于 $\mathbf{Z}(t)$ 的可学习且时变的邻接权重矩阵. 相比于 GNNs 中的离散化卷积邻域信息聚合方式, 图扩散将节点特征传播建模为连续扩散过程, 避免了 GNNs 中的刚性逐层传播结构, 缓解过平滑并获

得更加平滑的表示. 此外, 调整扩散时间可以连续控制消息传递感受野, 提高可解释性.

1.3 图谱滤波

GNNs 中的图卷积算子在空间域执行邻域信息聚合, 其本质与谱域的图滤波一致. 图谱滤波利用拉普拉斯矩阵特征分解将图信号从空间域映射到谱域, 根据设计的滤波器函数调整不同频率分量的权重, 进一步映射回空间域以实现信息滤波. 图谱滤波器定义为 $\mathbf{U}g(\mathbf{\Lambda})\mathbf{U}^T \mathbf{X}$, 其中 \mathbf{U} 是归一化拉普拉斯矩阵 $\tilde{\mathbf{L}}$ 的特征向量矩阵, $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}$ 是特征值对角矩阵, $g(\cdot)$ 是滤波函数. GCN 本质是一种低通滤波器, 其滤波器函数 $g(\lambda) = 1 - \lambda/2$ 关于 λ 单调递减, 抑制高频信息. 为了避免矩阵特征分解和简化优化过程, 多项式图谱滤波被提出^[30], 即

$$\mathbf{U}g(\mathbf{\Lambda})\mathbf{U}^T \mathbf{X} \approx \sum_{k=0}^K w_k \tilde{\mathbf{L}}^k \mathbf{X} \quad (2)$$

其中, w_k 是动态可学习的参数.

式 (2) 中, 多项式滤波避免了拉普拉斯矩阵分解. 实际上, $\sum_{k=0}^K w_k \tilde{\mathbf{L}}^k \mathbf{X} = \mathbf{U}(\sum_{k=0}^K w_k \mathbf{\Lambda}^k) \mathbf{U}^T \mathbf{X}$, 即采用 K 阶多项式 $\sum_{k=0}^K w_k \lambda^k$ 近似任意形式的滤波函数 $g(\cdot)$. 然而, 普通多项式存在龙格现象, 导致滤波器稳定性差. 此外, 多项式图谱滤波器涉及矩阵幂运算, 导致高计算复杂度. 为此, 具有递推形式的切比雪夫多项式被提出^[31], 其具有较强的计算有效性和数值稳定性.

2 所提方法

在本节中, 提出跨频域对齐对比学习方法. 第 2.1 节推导了经典分数阶图扩散过程并分析其存在的问题; 第 2.2 节提出切比雪夫多项式分数阶图扩散并利用不同的可训练参数将其推广到低频域与高频域, 构造对比视图; 第 2.3 节介绍了跨频域对齐对比学习损失函数.

2.1 经典分数阶图扩散过程

在经典 GCN 中, 节点信息传播是通过离散的层间传递完成的, 第 $l+1$ 层特征更新方式如下:

$$\mathbf{Z}^{(l+1)} = \sigma(\tilde{\mathbf{A}} \mathbf{Z}^{(l)} \mathbf{W}) \quad (3)$$

其中, σ 为激活函数, 常见如 Sigmoid, ReLU 等. 这种更新方式符合马尔可夫过程, 即每一层状态仅依赖其前一层状态, 忽略了历史状态信息, 无法为信息传播中的记忆与滞后效应提供物理解释. 为了便

于理解, 忽略激活函数和权重矩阵, 式 (3) 变为:

$$\mathbf{Z}^{(l+1)} = \tilde{\mathbf{A}}\mathbf{Z}^{(l)} \quad (4)$$

将层数 l 视为连续时间变量 t , 当 Δt 趋近于 0 时, t 时刻导数可以用离散差分 $\mathbf{Z}^{(l+1)} - \mathbf{Z}^{(l)}$ 近似, 即

$$\frac{d\mathbf{Z}(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{Z}(t + \Delta t) - \mathbf{Z}(t)}{\Delta t} \approx \mathbf{Z}^{(l+1)} - \mathbf{Z}^{(l)} \quad (5)$$

将式 (4) 代入式 (5) 得到:

$$\frac{d\mathbf{Z}(t)}{dt} = (\tilde{\mathbf{A}} - \mathbf{I})\mathbf{Z}(t) = -\tilde{\mathbf{L}}\mathbf{Z}(t) \quad (6)$$

正如式 (6) 所示, 在整数阶 ODE 图扩散过程中, 未来状态 $\mathbf{Z}(t + \Delta t)$ 仅依赖于当前状态 $\mathbf{Z}(t)$. 整数阶 ODE 图扩散将图信号传播描述为一个连续时间马尔可夫过程, 但它阻碍了模型捕捉图中长程依赖关系和历史演化趋势. 为克服这个限制, 将整数阶 ODE 推广到 FDE, 在传播扩散过程中引入记忆效应, 即当前时刻 t 的状态更新可以利用从初始状态到 t 时刻的完整历史状态信息. 特别地, Caputo 分数阶导数方程定义为:

$$D_t^\alpha \mathbf{Z}(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\tau)^{-\alpha} \mathbf{Z}'(\tau) d\tau \quad (7)$$

其中, $\Gamma(\cdot)$ 是 Gamma 函数, $\alpha \in (0, 1)$ 是衰减参数. 与整数阶 ODE 相比, 式 (7) 不具有马尔可夫性质, 积分项可以捕获从 0 到 t 整个时间区间内的历史信息. 此外, $(t-\tau)^{-\alpha}$ 为历史导数 $\mathbf{Z}'(\tau)$ 分配合适的权重, 其关于 τ 单调递增. 因此, Caputo 分数阶导数方程具有记忆衰减效应, 即近期的历史信息比长期信息具有更大的权重. 衰减率由阶数 α 控制, 较小的 α 导致较慢衰减, 赋予模型更持久的记忆; 较大的 α 导致快速衰减, 退化为无记忆的马尔可夫过程. 进一步, FGND 过程被形式化为:

$$D_t^\alpha \mathbf{Z}(t) = \mathcal{F}(\mathbf{Z}(t)) = -\tilde{\mathbf{L}}\mathbf{Z}(t) \quad (8)$$

学习到的节点嵌入 $\mathbf{Z}(t)$ 同时包含了时间维度和空间域信息, 是更具有判别性的图嵌入表示.

2.2 切比雪夫多项式分数阶图扩散

观察式 (8) 右侧滤波算子, 其本质是固定的一阶线性滤波器, 只能单调地抑制高频分量, 是一个低通滤波器. 进一步, 将分数阶微分方程 (8) 的解在谱域中表示如下:

$$\mathbf{Z}(t) = \mathbf{U} \text{diag}\{E_\alpha(-\lambda_1 t^\alpha), \dots, E_\alpha(-\lambda_N t^\alpha)\} \mathbf{U}^T \mathbf{Z}(0) \quad (9)$$

其中, $E_\alpha(z) = \sum_{k=0}^{\infty} z^k / (\Gamma(\alpha k + 1))$ 是 Mittag-Leff-

ler 单参数函数, 而对应于 FGND 的 $E_\alpha(-\lambda t^\alpha)$ 则展现出独特的代数衰减特性. 当 $t \rightarrow \infty$ 时, 有:

$$E_\alpha(-\lambda t^\alpha) \sim \frac{1}{\Gamma(1-\alpha)\lambda t^\alpha} + O\left(\frac{1}{\lambda^2 t^{2\alpha}}\right) \quad (10)$$

其中, $O(\cdot)$ 为高阶无穷小. 由式 (10) 可知, 分数阶微分方程图扩散对高频分量的衰减速率约为 $1/(\lambda t^\alpha)$. 对于经典的整数阶常微分方程图扩散过程 (6), 其解在谱域可表示为:

$$\mathbf{Z}(t) = \mathbf{U} \text{diag}\{e^{-\lambda_1 t}, e^{-\lambda_2 t}, \dots, e^{-\lambda_N t}\} \mathbf{U}^T \mathbf{Z}(0) \quad (11)$$

显然, 整数阶微分方程图扩散对高频分量的衰减率是 $e^{-\lambda t}$, 呈指数衰减形式, 远快于代数衰减, 高频分量会被迅速抑制. 因此, FGND 过程弱化了对图信号中高频分量的抑制, 这与其只能进行低通滤波存在一定矛盾. 为提高 FGND 过程对图信号中高低频信息调整的灵活性, 将滤波函数扩展为可学习的多项式形式, 即

$$g(\lambda) \approx \sum_{k=0}^K w_k \phi_k(\lambda) \quad (12)$$

其中, w_k 是动态可学习的参数, $\phi_k(\lambda)$ 是一组多项式基函数. 为了便于扩展滤波算子获取高频和低频节点表示构造对比学习视图, 使用参数化的切比雪夫多项式作为 FGND 过程滤波算子, 即

$$D_t^\alpha \mathbf{Z}(t) = \left(\sum_{k=0}^K w_k T_k(\hat{\mathbf{L}}) \right) \mathbf{Z}(t) \quad (13)$$

其中, 参数 K 表示切比雪夫多项式的最高阶数; $\hat{\mathbf{L}} = 2\tilde{\mathbf{L}}/\lambda_{max} - \mathbf{I}$ 将特征值归一化到范围 $[-1, 1]$, 确保适应切比雪夫多项式的定义域; $T_k(\cdot)$ 是一个在区间 $[-1, 1]$ 内满足正交性的切比雪夫递推多项式, 以递归形式定义如下:

$$\begin{cases} T_0(x) = 1 \\ T_1(x) = x \\ T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x) \end{cases} \quad (14)$$

式 (13) 中, w_k 是一个可学习参数, 定义为:

$$w_k = \frac{2}{K+1} \sum_{j=0}^K \gamma_j T_k(x_j) \quad (15)$$

其中, $x_j = \cos(\pi(j+1/2)/(K+1))$, $j = 0, \dots, K$ 是一个切比雪夫采样节点. 切比雪夫插值表明 γ_j 与相应切比雪夫节点 x_j 处的滤波器值存在直接对应关系. 因此, 通过控制可学习的滑动余弦参数 γ_j 来构造高频与低频滤波器^[32], 其参数 γ_j^h 和 γ_j^l 分别定义为:

$$\begin{cases} \gamma_j^h = \sigma(\beta_a^h) + \\ \quad \frac{1}{2}\sigma(\beta_b^h) \left(1 + \cos\left(1 + \frac{\frac{1}{2}\tanh(\delta_h) + j}{K}\pi\right)\right) \\ \gamma_j^l = \sigma(\beta_a^l) - \\ \quad \frac{1}{2}\sigma(\beta_b^l) \left(1 + \cos\left(1 + \frac{\frac{1}{2}\tanh(\delta_l) + j}{K}\pi\right)\right) \end{cases} \quad (16)$$

其中, $\sigma = \text{ReLU}(\cdot)$ 是激活函数, 以保证参数 γ_j 的非负性; $\beta_a^l, \beta_a^h, \beta_b^l, \beta_b^h, \delta_h$ 和 δ_l 均为可学习参数, 其初始值设置如下: $\beta_a^l = \delta_h = \delta_l = 0, \beta_a^h = \beta_b^l = \beta_b^h = 2$. 在式 (16) 中, 增益参数 β^l 和 β^h 控制滤波器的整体幅度, 相位偏移参数 δ_l 和 δ_h 控制通带中心位置, 实现对高频或低频的针对性增强. 滑动余弦策略允许通带在频谱上平滑移动, 从而灵活地关注不同频段的信息. 观察式 (13), 切比雪夫多项式滤波器对由 FGND 过程演化得到的节点嵌入 $\mathbf{Z}(t)$ 进行空间信息聚合操作. FGND 过程捕获的时间记忆为后续的空间频率分解提供了全面的、具有历史感知的特征基础, 这种时空关系挖掘使模型能够同时利用长程依赖性和多尺度结构模式的信息, 从而产生更具表达力的节点表示. 此外, 通过控制参数 δ_h 和 δ_l 简便地实现对图信号高低频域信息的挖掘, 获取有效的对比表示.

2.3 跨频域对齐对比学习

高频和低频滤波表示提供了原始图信号不同维度的特性, 在表示层面存在明显的多样性, 但在语义层面存在结构一致性. 经典的对比表示学习方法通常采用正样本对构造自监督信号, 本文提取图信号的高低频滤波表示作为自然的对比信号, 其在语义上彼此互斥, 无法作为正样本对. 若对其直接进行高低频负样本对比, 模型会过多关注高低频滤波的频率差异特性, 而非语义判别信息, 进一步会导致表示空间中同一节点的不同表示产生语义冲突, 造成语义坍塌, 这与多视图学习遵循最大化互信息准则相违背^[33]. 为此, 设计了跨频域对齐对比学习机制, 将高低频表示映射到虚拟的“同频”谱空间进行对比学习, 有助于实现表示多样性和语义一致性的和谐. 具体地, 将高频表示 \mathbf{Z}_h 映射到低频空间, 获取原始高频分量的低频表示 \mathbf{Z}_h^l :

$$\mathbf{Z}_h^l = f(\mathbf{Z}_h) \quad (17)$$

其中, $f(\cdot)$ 是特征变换函数, 本文采用经典的低频拉普拉斯谱滤波算子实现对高频分量到低频空间的映射. 因此, \mathbf{Z}_h^l 能够保留高频表示中的判别性信息, 且可以与低频表示在虚拟空间中进行语义对齐, 实

现对比学习. 在同频空间中, \mathbf{Z}_h^l 和 \mathbf{Z}_l 具有一致的语义, 将它们构建为对比正样本对, 采用 InfoNCE 对比损失来最大化它们之间的语义一致性, 即

$$\mathcal{L}_{\text{info}} = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp(\frac{1}{\tau_t} \text{sim}(\mathbf{Z}_h^l, \mathbf{Z}_l, i))}{\sum_{j=1}^N \exp(\frac{1}{\tau_t} \text{sim}(\mathbf{Z}_h^l, \mathbf{Z}_l, j))} \quad (18)$$

其中, $\text{sim}(\cdot, \cdot)$ 表示余弦相似性度量函数, τ_t 是温度系数控制对比度. 高频分量 \mathbf{Z}_h 通过函数 f 进行跨频域映射得到 \mathbf{Z}_h^l , 实际上保留了与低频分量 \mathbf{Z}_l 共享的语义信息. 因此, 跨频域对齐对比学习有效地将频率特定的变化与语义内容解耦, 平衡地表示互补与语义一致问题. 此外, 通过将高频表示视为低频表示的“困难扰动”, 而非直接对立面, 可在频域上实现语义一致性, 避免特征坍塌与语义漂移. 综上, 跨频域对齐对比学习避免了直接将高低频表示对立带来的信息割裂, 实现了跨频域信息自蒸馏, 增强了图表征的鲁棒性与判别性.

跨频域对齐对比学习使得高低频嵌入在表示层面具有互补性, 在语义层面具有结构一致性. 因此, 通过下述方式对高低频表示融合, 为下游任务获取更加全面的表示, 即

$$\mathbf{Z}_f = \alpha_1 \mathbf{Z}_l + (1 - \alpha_1) \mathbf{Z}_h \quad (19)$$

其中, α_1 是自注意可学习权重系数, 用于权衡 \mathbf{Z}_l 和 \mathbf{Z}_h 两者之间的重要性. 融合高频与低频能够在全局平滑性与局部差异性之间取得平衡, 既保留图结构的一致语义, 又强化边界与细节表达, 从而提升图模型的判别性、鲁棒性与泛化能力. 对于分类任务, 将更加全面的表示 \mathbf{Z}_f 输入到一个分类器中, 获取概率形式的伪标签:

$$\hat{\mathbf{Y}} = \text{softmax}(\mathbf{Z}_f \mathbf{W} + \mathbf{b}) \quad (20)$$

其中, \mathbf{b} 为偏置向量.

采用交叉熵损失函数作为训练目标, 即

$$\mathcal{L}_c = -\frac{1}{|\mathcal{V}_{tr}|} \sum_{v_i \in \mathcal{V}_{tr}} \sum_{c=1}^C y_{i,c} \log \hat{y}_{i,c} \quad (21)$$

其中, \mathcal{V}_{tr} 是训练集, $y_{i,c}$ 和 $\hat{y}_{i,c}$ 分别表示样本 v_i 属于类别 c 的 one-hot 真实标签指示和预测概率.

3 实验结果及分析

3.1 实验设置

数据集

为了全面评估所提出方法的泛化能力与鲁棒性, 选取了具有不同图结构特性的真实数据集进行

验证, 包括三个同配性引文网络数据集: Cora、Citeseer 和 PubMed, 六个异配性数据集: Wisconsin、Texas、Cornell、Chameleon、Squirrel 和 Actor. 对于所有数据集, 遵循多项式滤波图对比学习 (polynomial filters GCL, PolyGCL)^[34] 中采用的数据划分. 对于同配性数据集 Cora、Citeseer 和 PubMed, 采用标准划分方法; 对于异配性数据集, 随机选取 48% 样本为训练集, 32% 样本为验证集, 剩余的 20% 样本为测试集. 此外, 为了衡量一个图的同配性, 边同配性比被用来度量同类节点相连边数与总边数的比值, 定义如下:

$$\mathcal{H}_{\text{edge}} = \frac{|\{e_{ij} \in \mathcal{E}: y_i = y_j\}|}{|\mathcal{E}|} \quad (22)$$

其中, y_i 是节点 v_i 的真实标签. 更多数据集详细统计信息如表 1 中所示.

表 1 数据集统计信息
Table 1 The statistics of the datasets

数据集	节点数量	边数量	特征数量	类别数量	边同配性比
Cora	2708	5429	1433	7	0.81
Citeseer	3327	4732	3703	6	0.73
PubMed	19717	88651	500	3	0.80
Wisconsin	251	466	1703	5	0.17
Texas	183	309	1793	5	0.06
Cornell	183	295	1703	5	0.12
Chameleon	2277	36101	2325	5	0.23
Squirrel	5201	217073	2089	5	0.20
Actor	7600	33391	932	5	0.21

基线方法

为了全面评估提出的 CfACL 方法的有效性, 选择下述具有代表性的基于 GCL 的图节点分类方法作为对比, 包括: 深度图信息最大化 (deep graph infomax, DGI)^[35]、基于自适应增强的图对比学习 (graph contrastive learning with adaptive aug-

mentation, GCA)^[36]、基于典型相关性分析的图自监督学习 (canonical correlation analysis inspired self-supervised learning on graphs, CCA-SSG)^[37]、自举式图表示学习 (bootstrapped graph representation learning, BGRL)^[38]、单通图对比学习 (single-pass GCL, SP-GCL)^[39]、不对称图对比学习 (asymmetric contrastive learning for graphs, GraphACL)^[19] 和 PolyGCL^[34].

3.2 实验结果分析

图节点分类任务的实验结果如表 2 所示, 最优结果以粗体标出, 次优结果以下划线标出. 所提出的方法 CfACL 在所有数据集上均获得最优的节点分类结果. 相比于次优的对比方法, 本方法所取得的性能提升是明显的, 平均提高了 5.56%, 充分验证了本方法在判别性图节点表示学习及其下游节点分类任务上的有效性. 为了从实验结果角度验证本方法设计的有效性, 分析并总结如下:

1) 本文提出的 CfACL 在同配性和异配性图数据集上均取得了最优的节点分类性能, 验证了其在复杂数据上的鲁棒性. 具体而言, 在异配性 Wisconsin、Cornell、Texas、Chameleon、Squirrel 和 Actor 数据集上, 所提出的方法相较于次优方法分别提高了 4.26%、12.37%、10.98%、3.17%、8.49% 和 2.39%. 在异配图中, 异类节点之间存在错误的连接关系, 因此节点本身的特异性信息相比于邻居信息更加可靠, 需要发挥节点本身的潜力以排除异配连接节点的干扰. 本文提出的 CfACL 方法引入高频滤波器关注异配节点自身的特征信息, 并且采用切比雪夫多项式分数阶图扩散可以扩大信息聚合的感受野, 从全局提取判别信息, 因此在异配性数据集上效果较好. 此外, 该方法在同配性数据集 Cora、Citeseer 和 PubMed 上仍然保持了有竞争力的性能.

2) 相比基于数据增强的图对比学习方法 GCA、CCA-SSG 和 BGRL, 本文所提方法取得了显著的

表 2 在同配性和异配性数据集上的节点分类实验结果 (%)
Table 2 The node classification experiment results on homophilic and heterophilic datasets (%)

方法	Cora	Citeseer	PubMed	Wisconsin	Cornell	Texas	Chameleon	Squirrel	Actor
DGI	82.30 ± 0.60	71.80 ± 0.70	76.80 ± 0.60	55.21 ± 1.02	45.33 ± 6.11	58.53 ± 2.98	60.27 ± 0.70	26.44 ± 1.12	28.30 ± 0.76
GCA	82.93 ± 0.42	72.19 ± 0.31	80.79 ± 0.45	59.55 ± 0.81	52.31 ± 1.09	52.92 ± 0.46	63.66 ± 0.32	48.09 ± 0.21	28.47 ± 0.29
CCA-SSG	84.00 ± 0.40	73.10 ± 0.30	81.00 ± 0.40	58.46 ± 0.96	52.17 ± 1.04	59.89 ± 0.78	62.41 ± 0.22	46.76 ± 0.36	27.82 ± 0.60
BGRL	82.70 ± 0.60	71.10 ± 0.80	79.60 ± 0.50	51.23 ± 1.17	50.33 ± 2.29	52.77 ± 1.98	64.86 ± 0.63	36.22 ± 1.97	28.80 ± 0.54
SP-GCL	83.16 ± 0.13	71.96 ± 0.42	79.16 ± 0.84	60.12 ± 0.39	52.29 ± 1.21	59.81 ± 1.33	65.28 ± 0.53	52.10 ± 0.67	28.94 ± 0.69
GraphACL	<u>84.20 ± 0.31</u>	<u>73.63 ± 0.22</u>	<u>82.02 ± 0.15</u>	69.22 ± 0.40	<u>59.33 ± 1.48</u>	71.08 ± 0.34	<u>69.12 ± 0.24</u>	<u>54.05 ± 0.13</u>	30.03 ± 0.13
PolyGCL	81.97 ± 0.19	71.97 ± 0.29	77.48 ± 0.39	<u>76.08 ± 3.33</u>	43.78 ± 3.51	<u>72.16 ± 3.51</u>	46.84 ± 1.53	34.25 ± 0.66	<u>34.37 ± 0.69</u>
CfACL	85.17 ± 1.51	76.67 ± 1.38	86.37 ± 1.26	80.34 ± 0.47	71.70 ± 7.43	83.14 ± 5.34	72.29 ± 1.50	62.54 ± 1.14	36.76 ± 1.34

性能优势, 并且其余无增强的图对比学习方法 SP-GCL、GraphACL 和 PolyGCL 在大多数数据集上也取得了较优的性能. GCA、CCA-SSG 和 BGRL 采用随机增强、动态增强等策略构造对比视图, 其性能严重依赖增强方式及其质量. 不同于此, 无增强的图对比学习方法通过不同的特征编码器自适应学习增强视角, 能够获得表达能力更强的对比视图. 尤其对于提出的 CfACL 和 PolyGCL, 利用图信号中自然存在的高低频信息作为对比视角, 充分利用数据内在的信息, 避免随机增强等引入对比偏差, 提高了对比学习的判别性.

3) 相比于无数据增强图对比学习方法 SP-GCL、GraphACL 和 PolyGCL, 本文提出的 CfACL 仍然取得了最优的节点分类性能. 尤其相比于重要的基准 PolyGCL, 本方法在所有数据集上平均提高了 12.90%. 实际上, PolyGCL 和本文提出的 CfACL 均使用图信号内在的高低频信息构造图对比视角, CfACL 性能的优越性主要归结于判别性的节点表示学习方法和对比学习机制. 对于节点表示学习, 尽管 PolyGCL 采用了灵活的多项式滤波器作为特征提取器, 但是其依旧遵循 GNN 的层级结构, 仅能捕获局部邻域信息且易面临过平滑问题, 限制了表示学习的判别性. 本文提出的切比雪夫多项式分数阶图扩散本质上是一个具有记忆特性的动态系统, 在节点表示更新过程中, 可以有效挖掘历史信息中的长程依赖关联并结合多项式滤波器捕获复杂的多尺度空间结构信息, 从时空双维度获取判别性的图节点表示. 此外, 新颖的跨频域对齐对比损失的设计, 在虚拟“同频”空间中促进语义一致性并保持原始高低频表示本质的互补性信息, 避免了两者直接对比导致的语义漂移, 增强了图表示学习的判别性.

3.3 消融实验结果

3.3.1 高频和低频表示的影响

为了验证高频和低频表示对于获取全面且具判别性的节点表示的重要性, 设计如下消融实验:

1) “CfACL w/o 高频”. 采用跨频域对齐对比学习获取高低频表示, 但仅使用低频表示作为分类器输入, 验证高频表示的重要性.

2) “CfACL w/o 低频”. 采用跨频域对齐对比学习获取高低频表示, 但仅使用高频表示作为分类器输入, 验证低频表示的重要性.

在 PubMed、Cornell 和 Actor 数据集上的消融实验结果如图 1 所示, 完整的 CfACL 方法相较于消融版本“CfACL w/o 高频”和“CfACL w/o 低频”均取得了更高的分类性能, 验证了基于跨频域对齐

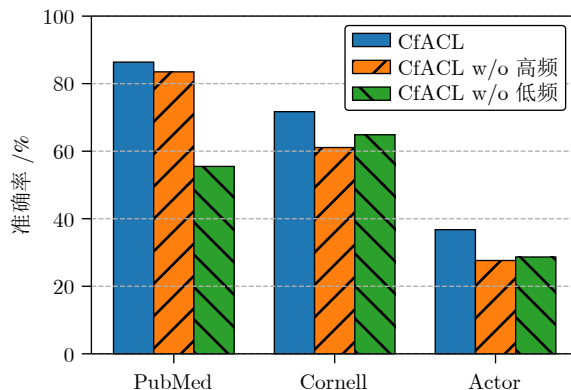


图 1 消融实验结果

Fig.1 The results of ablation experiment

对比学习获得的高低频信息对于判别性的图节点表示是重要的, 在保持语义一致的前提下提供了互补的表示基础. 进一步, 在同配性 PubMed 数据集上, “CfACL w/o 高频”的性能优于“CfACL w/o 低频”, 说明低频滤波通过捕捉图结构的平滑特征可以为同配图获取更具判别性的节点表示. 在异配性 Cornell 和 Actor 数据集上, “CfACL w/o 低频”的性能优于“CfACL w/o 高频”, 说明富含节点特异性信息的高频表示对于异配图更加重要. 尽管高低频信息对于同配和异配图具有不同的重要性, 但是本文提出的跨频域对齐对比学习方法可以平衡其低层次表示互补与高层次语义一致性, 充分发挥两者的积极作用.

3.3.2 跨频域对齐对比损失函数有效性评估

上述实验结果已经验证了跨频域对齐对比学习策略在获取判别性图节点表示和提高节点分类性能方面的有效性, 本节重点评估设计的跨频域对齐对比损失机制在防止表示崩塌和语义漂移方面的有效性. 为此, 在训练过程中选取了几种常用的对比损失函数: 欧几里得 (Euclidean) 损失、Barlow Twins 损失^[40]、VICReg (variance-invariance-covariance regularization) 损失^[41] 以及本文设计的跨频域对齐 (cross-frequency alignment, CfA) 损失. 具体地, 对高低频滤波表示采用这些损失函数进行直接对比, 自监督设置如下所示:

1) Euclidean 损失. 将 \mathbf{Z}_h 和 \mathbf{Z}_l 视为原始节点的两个表示视图, 采用欧几里得距离度量相似度, 即

$$\mathcal{L}_{Euclidean} = \frac{1}{N} \sum_{i=1}^N \|\mathbf{Z}_{h,i} - \mathbf{Z}_{l,i}\|_2^2 \quad (23)$$

2) Barlow Twins 损失. 采用余弦距离度量相似度, 同一样本的跨频域表示相似度趋近于 1, 不同样本的趋近于 0, 定义如下:

$$\mathcal{L}_{Barlow\ Twins} = \sum_{i=1}^N (1 - C_{ii})^2 + \beta \sum_{i=1}^N \sum_{j=1, j \neq i}^N C_{ij}^2 \quad (24)$$

其中, $\mathbf{C} = \mathbf{Z}_h^T \mathbf{Z}_l$ 是归一化特征表示 \mathbf{Z}_h 和 \mathbf{Z}_l 的余弦相似度矩阵, β 是权衡参数.

3) VICReg 损失. 旨在确保每个表示方差稳定并最大程度地减小跨频域表示冗余, 实现表示对齐. 本实验中, VICReg 损失 \mathcal{L}_{VICReg} 由方差正则化 \mathcal{L}_{var} , 不变性损失 \mathcal{L}_{inv} 和协方差正则化 \mathcal{L}_{cov} 三部分组成:

$$\mathcal{L}_{VICReg} = \eta_1 \mathcal{L}_{var} + \eta_2 \mathcal{L}_{inv} + \eta_3 \mathcal{L}_{cov} \quad (25)$$

其中, η_1 , η_2 和 η_3 是平衡系数. 方差正则化 \mathcal{L}_{var} 确保节点表示的每个特征维度都具有足够方差, 定义如下:

$$\mathcal{L}_{var} = \frac{1}{d} \sum_{j=1}^d \max\left(0, \varepsilon - \sqrt{\text{Var}(\mathbf{Z}_h[:, j])}\right) + \max\left(0, \varepsilon - \sqrt{\text{Var}(\mathbf{Z}_l[:, j])}\right) \quad (26)$$

其中, $\text{Var}(\mathbf{Z}[:, j]) = (1/N) \sum_{i=1}^N (\mathbf{Z}_{i,j} - \mu_j)^2$, $\mu_j = (1/N) \sum_{i=1}^N \mathbf{Z}_{i,j}$; ε 是无穷小量, 确保方差非零. 不变性损失 \mathcal{L}_{inv} 定义为 \mathbf{Z}_h 和 \mathbf{Z}_l 中相应样本的均方欧氏距离, 与第 3.3.2 节 1) 中 Euclidean 损失相同:

$$\mathcal{L}_{inv} = \frac{1}{N} \sum_{i=1}^N \|\mathbf{Z}_{h,i} - \mathbf{Z}_{l,i}\|_2^2 \quad (27)$$

协方差正则化 \mathcal{L}_{cov} 定义如下:

$$\mathcal{L}_{cov} = \frac{1}{d} \sum_{i \neq j} (\text{Cov}(\mathbf{Z}_h)_{j,k}^2 + \text{Cov}(\mathbf{Z}_l)_{j,k}^2) \quad (28)$$

其中 $\text{Cov}(\mathbf{Z})_{j,k}^2 = (1/N) \sum_{i=1}^N (\mathbf{Z}_{i,j} - \mu_j)(\mathbf{Z}_{i,k} - \mu_k)$ 表示节点嵌入 \mathbf{Z} 的协方差矩阵的非对角元素. 在代表性 Cornell 数据集上的实验结果如图 2 所示. 在

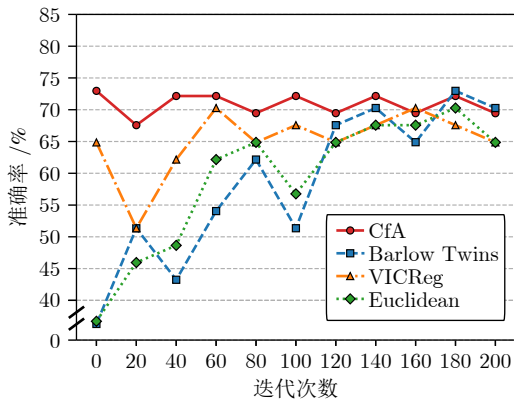


图 2 CfACL 使用不同对比学习损失在数据集 Cornell 上的实验结果

Fig.2 The experimental results of CfACL with various contrastive learning losses on the Cornell dataset

迭代训练过程中, 提出的 CfA 损失确保准确率相对稳定, 而基于其余对比损失的准确率出现了不同程度的波动, 表明训练过程中出现了特征坍塌和语义漂移, 影响了训练过程的稳定性.

3.4 过平滑分析

经典 GNNs 方法的离散邻居信息聚合机制易面临过平滑的问题, 即随着网络层数的增加, 所有节点的表示会趋于相似, 从而失去判别性. 因此, 经典的 GNNs 通常设置为 2~3 层. 不同于经典的 GNNs, 本研究设计的切比雪夫多项式分数阶图扩散是模拟连续的动力学过程实现信息的传递, 其时间 t 类似于 GNNs 中的层数. 为此, 本节测试提出的 CfACL 方法在 $t \in \{2, 3, \dots, 10\}$ 设置下的性能表现, 实验结果如图 3 所示. 可以观察到, 随着 t 的增大, 分类准确率有一定的波动, 但是没有出现断崖式的下降. 尤其在 PubMed 数据集上, 分类性能提升且趋于相对稳定. 进一步, 为了观察在更深层数下的过平滑现象, 测试 CfACL 方法在 $t \in \{4, 8, 16, 32, 64\}$ 设置下的分类性能. 正如图 4 所示, 当 $t \geq 8$ 时, 准确率出现一定程度的下降, 但当 t 一直增大到 64 时, 分类准确率未出现断崖式下降, 仍能够实现相对满意的分类性能.

上述实验结果表明本方法能够在一定程度上缓解过平滑问题, 这主要归结于切比雪夫多项式分数阶图扩散机制所具有的历史记忆特性与可控的滤波方式, 能够有效抑制节点信息传播过程中的特征平均化趋势. 具体地, 从时间维度来看, 传统的离散式层间传播过程中, 特征更新仅依赖每层的邻居节点, 节点表示会过快同化. 本文提出的分数阶图神经扩散 FGND 引入分数阶微分方程, 通过积分项引入从初始状态到当前时刻的全部历史信息, 节点更新不仅依赖于每层的邻居节点, 还会考虑历史信息的影

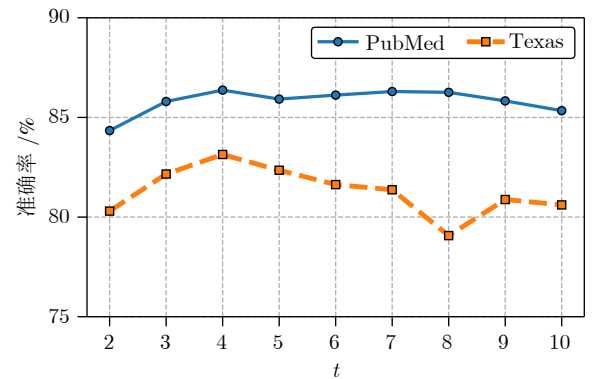


图 3 CfACL 关于不同时间 t 的实验结果

Fig.3 The experimental results of CfACL with various time t

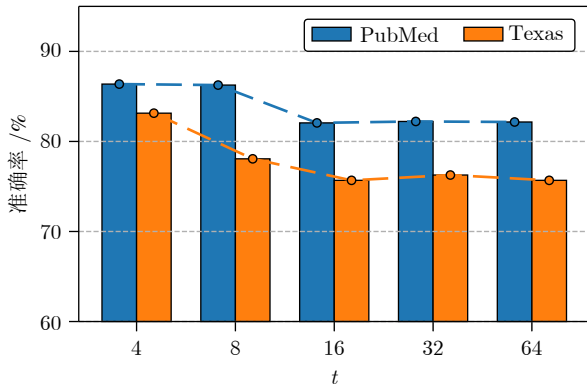


图4 CfACL在时间 $t \in \{4, 8, 16, 32, 64\}$ 下的分类性能
Fig.4 The classification performance of CfACL when time $t \in \{4, 8, 16, 32, 64\}$

响, 因此信息并不会被过快地同化, 能够缓解过平滑问题. 从空间维度来看, 本方法将滤波算子扩展为可学习的切比雪夫多项式, 相较于传统的低通滤波器, 切比雪夫多项式 FGND 在进行低频筛选时也能够增强有益的高频成分, 保留节点特异性信息, 进而捕捉不同尺度的结构信息, 节点信息更加丰富. 综上, 本文提出的方法可以使节点特征在传播中仍然能够保持多样性和判别性, 缓解过平滑问题.

3.5 表示可视化

为了直观验证 CfACL 在判别性节点学习方面的优越性, 采用 t-SNE 可视化技术在二维空间内展

示原始属性特征 \mathbf{X} 和 PolyGCL、GraphACL 及 CfACL 所学图节点表示的分布情况, 在 Cora 和 Cornell 数据集上的可视化实验结果如图 5 所示. 可以观察到, 相比于原始属性特征 \mathbf{X} , 上述方法学习的图节点表示的判别性明显提高. 在上述图表示学习方法中, CfACL 学习的图节点表示在低维空间中呈现出更加紧凑的类内聚集和更加清晰的类间分离. 特别是在结构复杂的 Cornell 数据集上, CfACL 更能将不同类别的节点划分到不同的区域, 而其他方法的结果则存在较多的重叠与混淆. 该可视化实验结果证明了 CfACL 方法能够学习高质量、高判别性的图表示, 为提高下游任务性能提供了表示基础.

3.6 模型训练计算时间对比

为了从实验角度分析本文所提方法的计算复杂度, 本节展示并对比所提方法的训练成本. 选取主流的图对比学习方法 PolyGCL^[34] 和 GraphACL^[19] 作为对比基准, 在部分同配性和异配性数据集上测试实际训练时间. 特别地, 在相同实验环境下, 严格遵循各方法的最优参数配置, 测量上述方法实际运行时间十次并汇报时间开销的均值. 实验结果如表 3 所示, 从表中可以看出, 所提方法 CfACL 在不同规模与结构的数据集上所需的训练时间是较短的, 具有相对明显的计算成本优势. 这一结果表明, 提出的方法不仅在表示学习能力上具有显著竞争力, 而且在计算效率方面也具有一定的优势.

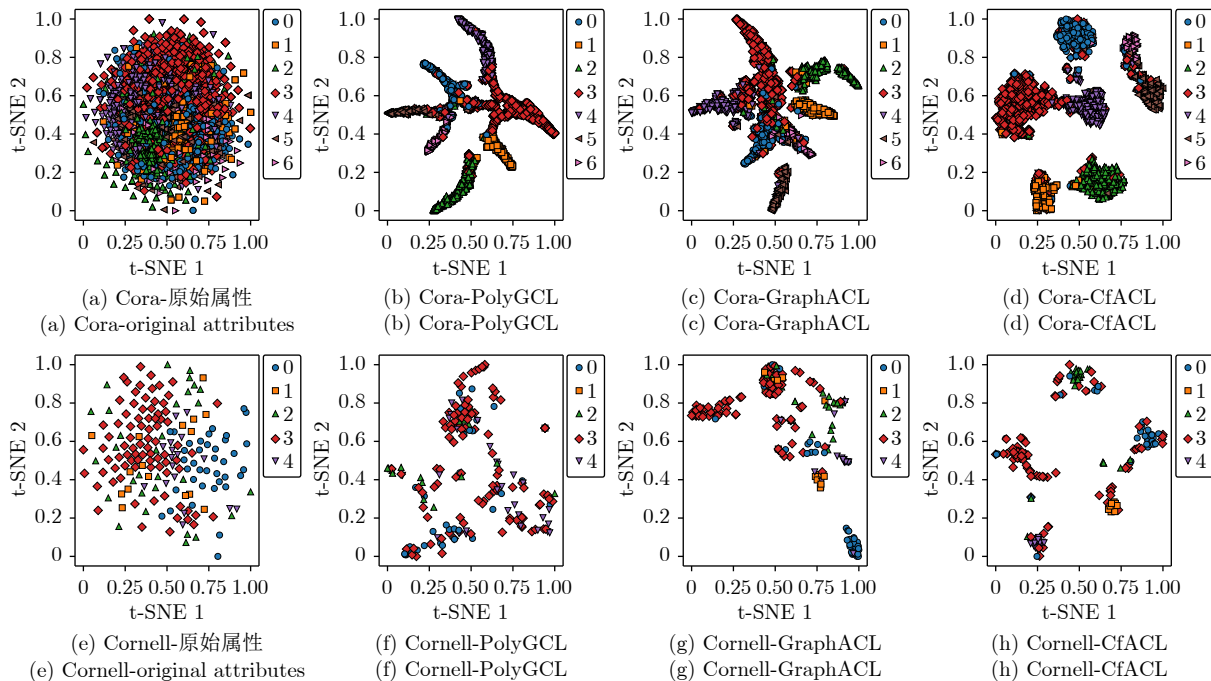


图5 在 Cora 和 Cornell 数据集上的 t-SNE 可视化结果

Fig.5 The t-SNE visualization results on the Cora and Cornell datasets

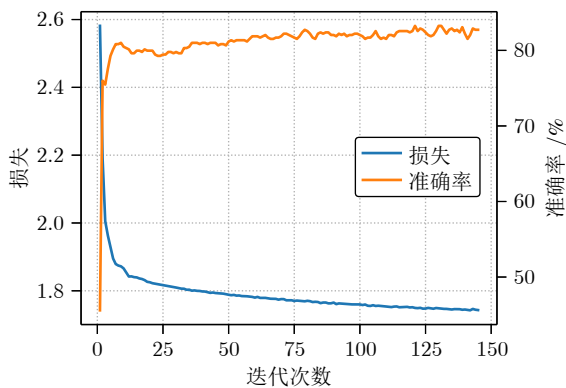
3.7 收敛性分析

为了验证所提方法在训练过程中的收敛性和稳定性, 选取同配性数据集 Cora 和异配性数据集 Chameleon 作为代表性数据集, 展示模型训练过程中目标损失和分类准确率的变化情况. 实验结果如图 6 所示, 可以观察到, 分类准确率可以很快达

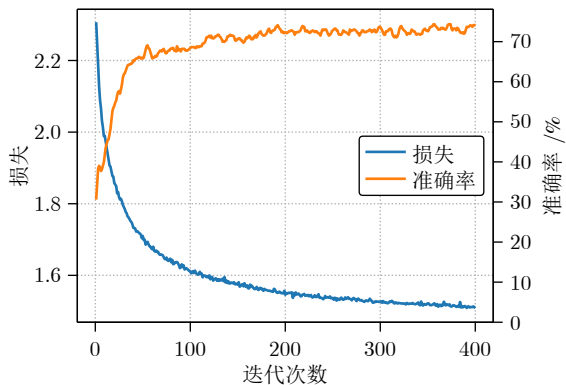
表 3 CfACL 与代表性对比方法 PolyGCL 和 GraphACL 在不同数据集上的训练成本对比 (s)

Table 3 The comparison of training costs for CfACL versus the representative comparison methods PolyGCL and GraphACL across different datasets (s)

方法	Cora	PubMed	Chameleon	Cornell	Texas	Actor
PolyGCL	23.42	4392.96	19.62	32.29	24.03	54.71
GraphACL	23.08	340.35	212.19	28.10	27.64	3888.55
CfACL	19.58	108.86	19.58	22.79	29.96	32.99



(a) Cora 数据集
(a) Cora dataset



(b) Chameleon 数据集
(b) Chameleon dataset

图 6 在 Cora 和 Chameleon 数据集上的训练损失和分类准确率随训练迭代步数的变化

Fig. 6 The variation of training loss and classification accuracy with training iteration epochs on Cora and Chameleon datasets

到较高水平并保持相对稳定. 此外, 随着迭代训练的进行, 损失也迅速下降并在最后趋于稳定. 上述实验结果表明了所提模型具有较强的收敛性和稳定性.

4 结束语

本文提出一种基于分数阶图神经扩散的跨频域对齐对比学习方法 CfACL, 通过集成分数阶微分方程的长程记忆特性和切比雪夫多项式谱滤波的多尺度结构信息建模能力, 构建了一个时空双维度的图表示学习方法. 进一步, 基于高频和低频的双路径切比雪夫多项式滤波过程构造图信号中自然的对比视角, 并通过跨频域对齐对比策略在保持语义一致性的同时有效融合多源对比信息. 丰富的实验结果表明, 提出的 CfACL 可以为下游节点分类任务获取全面且具有判别性的图节点表示, 提高分类准确率. 然而, 在开放环境中, 大规模的分数阶图扩散需要进一步被考虑, 以提高方法的可拓展性.

参考文献

- Liang K, Meng L Y, Liu M, Liu Y, Tu W X, Wang S W, et al. A survey of knowledge graph reasoning on graph types: Static, dynamic, and multi-modal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, **46**(12): 9456–9478
- Li M R, Zhang Y, Zhang W, Zhao S Y, Piao X L, Yin B C. CSAT: Contrastive sampling-aggregating transformer for community detection in attribute-missing networks. *IEEE Transactions on Computational Social Systems*, 2024, **11**(2): 2277–2290
- Luo Biao, Hu Tian-Meng, Zhou Yu-Hao, Huang Ting-Wen, Yang Chun-Hua, Gui Wei-Hua. Survey on multi-agent reinforcement learning for control and decision-making. *Acta Automatica Sinica*, 2025, **51**(3): 510–539
(罗彪, 胡天萌, 周育豪, 黄廷文, 阳春华, 桂卫华. 多智能体强化学习控制与决策研究综述. *自动化学报*, 2025, **51**(3): 510–539)
- Wu Z H, Lu J L, Yu J J, Zhou S, Pi Y Y, Wang H S. Divide and conquer: Coordinating multiplex mixture of graph learners to handle multi-omics analysis. In: Proceedings of the 34th International Joint Conference on Artificial Intelligence. Montreal, Canada: IJCAI, 2025. Article No. 736
- Li M R, Zhang P Y, Xing W B, Zheng Y J, Zaporjets K, Chen J Z, et al. A survey of large language models for data challenges in graphs. *Expert Systems With Applications*, 2026, **298**: Article No. 129643
- Wu Bo, Liang Xun, Zhang Shu-Sen, Xu Rui. Advances and applications in graph neural network. *Chinese Journal of Computers*, 2022, **45**(1): 35–68
(吴博, 梁循, 张树森, 徐睿. 图神经网络前沿进展与应用. *计算机学报*, 2022, **45**(1): 35–68)
- Wu Z H, Cai J Y, Zhang Y H, Lu J L, Chen Z L, Zhuang S M, et al. Where graph meets heterogeneity: Multi-view collaborative graph experts. In: Proceedings of 39th Annual Conference on Neural Information Processing Systems. San Diego, USA: Open-Review.net, 2025. 1–22
- Wu Z H, Pan S R, Chen F W, Long G D, Zhang C Q, Yu P S. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, **32**(1): 4–24
- Zhang Jian-Peng, Pei Yu-Long, Liu Cong, Li Shao-Mei, Chen Hong-Chang. A semi-supervised clustering algorithm based on

- factor graph model for dynamic graphs. *Acta Automatica Sinica*, 2020, **46**(4): 670–680
(张建朋, 裴雨龙, 刘聪, 李邵梅, 陈鸿昶. 基于因子图模型的动态图半监督聚类算法. *自动化学报*, 2020, **46**(4): 670–680)
- 10 Cai L, Li J D, Wang J, Ji S W. Line graph neural networks for link prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, **44**(9): 5103–5113
 - 11 Gong Pei-Liang, Ai Li-Hua. Two-order approximate spectral convolutional model for semi-supervised classification. *Acta Automatica Sinica*, 2021, **47**(5): 1067–1076
(公沛良, 艾丽华. 用于半监督分类的二阶近似谱图卷积模型. *自动化学报*, 2021, **47**(5): 1067–1076)
 - 12 Zhang Chong-Sheng, Chen Jie, Li Qi-Long, Deng Bin-Quan, Wang Jie, Chen Cheng-Gong. Deep contrastive learning: A survey. *Acta Automatica Sinica*, 2023, **49**(1): 15–39
(张重生, 陈杰, 李岐龙, 邓斌权, 王杰, 陈承功. 深度对比学习综述. *自动化学报*, 2023, **49**(1): 15–39)
 - 13 Ju W, Wang Y F, Qin Y F, Mao Z Y, Xiao Z P, Luo J Y, et al. Towards graph contrastive learning: A survey and beyond. arXiv preprint arXiv: 2405.11868, 2024.
 - 14 Zhao T X, Wang Y Q, Xu S L, Yang T C, Gao J B, Guo J P. Dual-level noise augmentation for graph clustering with triplet-wise contrastive learning. *Pattern Recognition*, 2026, **172**: Article No. 112463
 - 15 Chen J S, Liu H P, Hopcroft J, He K. Leveraging contrastive learning for enhanced node representations in tokenized graph transformers. In: Proceedings of the 38th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc., 2024. Article No. 2725
 - 16 Lee H K, Zhang Q C, Yang C, Xiong L. Node-level contrastive unlearning on graph neural networks. arXiv preprint arXiv: 2503.02959, 2025.
 - 17 Zhao Y H, Wang Y J, Wang Z K, Shan W, Huang M M, Wang X W. Graph contrastive learning with progressive augmentations. In: Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Toronto, Canada: ACM, 2025. 2079–2088
 - 18 Liang H D, Du X J, Zhu B L, Ma Z J, Chen K, Gao J B. Graph contrastive learning with implicit augmentations. *Neural Networks*, 2023, **163**: 156–164
 - 19 Xiao T, Zhu H S, Chen Z Y, Wang S H. Simple and asymmetric graph contrastive learning without augmentations. In: Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc., 2023. Article No. 710
 - 20 Zhao T X, Wang Y Q, Wang J L, Wang J P, Cui M L, Gao J B, et al. Hybrid-collaborative augmentation and contrastive sample adaptive-differential awareness for robust attributed graph clustering. arXiv preprint arXiv: 2510.02731, 2025.
 - 21 Wang D, Zhou W, Hu S. Information diffusion prediction with graph neural ordinary differential equation network. In: Proceedings of the 32nd ACM International Conference on Multimedia. Melbourne, Australia: ACM, 2024. 9699–9708
 - 22 Chamberlain B, Rowbottom J, Gorinova M I, Bronstein M, Webb S, Rossi E. GRAND: Graph neural diffusion. In: Proceedings of the 38th International Conference on Machine Learning. Virtual Event: PMLR, 2021. 1407–1418
 - 23 Thorpe M, Nguyen T M, Xia H D, Strohmer T, Bertozzi A L, Osher S J, et al. GRAND++: Graph neural diffusion with a source term. In: Proceedings of the 10th International Conference on Learning Representations. Virtual Event: OpenReview.net, 2022.
 - 24 Kang Q Y, Zhao K, Ding Q X, Ji F, Li X H, Liang W F, et al. Unleashing the potential of fractional calculus in graph neural networks with FROND. In: Proceedings of the 12th International Conference on Learning Representations. Vienna, Austria: OpenReview.net, 2024.
 - 25 Zhao K, Li X H, Kang Q Y, Ji F, Ding Q X, Zhao Y N, et al. Distributed-order fractional graph operating network. In: Proceedings of the 38th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc., 2024. Article No. 3286
 - 26 Wang J L, Guo J P, Sun Y F, Gao J B, Wang S F, Yang Y C, et al. DGNN: Decoupled graph neural networks with structural consistency between attribute and graph embedding representations. *IEEE Transactions on Big Data*, 2025, **11**(4): 1813–1827
 - 27 Lee N, Lee J, Park C. Augmentation-free self-supervised learning on graphs. In: Proceedings of the 36th AAAI Conference on Artificial Intelligence. Virtual Event: AAAI Press, 2022. 7372–7380
 - 28 Zhao Y N, Ji F, Zhao K, Li X H, Kang Q Y, Liang W F, et al. Simple graph contrastive learning via fractional-order neural diffusion networks. arXiv preprint arXiv: 2504.16748, 2025.
 - 29 Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks. In: Proceedings of the 5th International Conference on Learning Representations. Toulon, France: OpenReview.net, 2017.
 - 30 Yu W D, Hou Z C, Liu X R. Automated polynomial filter learning for graph neural networks. In: Proceedings of the International Conference on Big Data. Washington, USA: IEEE, 2024. 680–689
 - 31 He M G, Wei Z W, Wen J R. Convolutional neural networks on graphs with Chebyshev approximation, revisited. In: Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc., 2022. Article No. 527
 - 32 Zou Z Y, Jiang Y H, Shen L, Liu J, Liu X R. LOHA: Direct graph spectral contrastive learning between low-pass and high-pass views. In: Proceedings of the 39th AAAI Conference on Artificial Intelligence. Philadelphia, USA: AAAI Press, 2025. 13492–13500
 - 33 Guo J P, Sun Y F, Gao J B, Hu Y L, Yin B C. Logarithmic Schatten- p norm minimization for tensorial multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, **45**(3): 3396–3410
 - 34 Chen J Y, Lei R L, Wei Z W. PolyGCL: Graph contrastive learning via learnable spectral polynomial filters. In: Proceedings of the 12th International Conference on Learning Representations. Vienna, Austria: OpenReview.net, 2024.
 - 35 Veličković P, Fedus W, Hamilton W L, Liò P, Bengio Y, Hjelm R D. Deep graph infomax. In: Proceedings of the 7th International Conference on Learning Representations. New Orleans, USA: OpenReview.net, 2019.
 - 36 Zhu Y Q, Xu Y C, Yu F, Liu Q, Wu S, Wang L. Graph contrastive learning with adaptive augmentation. In: Proceedings of the Web Conference. Ljubljana, Slovenia: ACM, 2021. 2069–2080
 - 37 Zhang H R, Wu Q T, Yan J C, Wipf D, Yu P S. From canonical correlation analysis to self-supervised graph neural networks. In: Proceedings of the 35th International Conference on Neural Information Processing Systems. Virtual Event: Curran Associates Inc., 2021. Article No. 7
 - 38 Thakoor S, Tallec C, Azar M G, Azabou M, Dyer E L, Munos R, et al. Large-scale representation learning on graphs via bootstrapping. In: Proceedings of the 10th International Conference on Learning Representations. Virtual Event: OpenReview.net, 2022.
 - 39 Wang H N, Zhang J Y, Zhu Q, Huang W, Kawaguchi K J, Xiao X K. Single-pass contrastive learning can work for both homophilic and heterophilic graph. In: Proceedings of the 21st Annual Conference of the International Speech Communication Association. Shanghai, China: ISCA, 2020.
 - 40 Zbontar J, Jing L, Misra I, LeCun Y, Deny S. Barlow twins: Self-supervised learning via redundancy reduction. In: Proceedings of the 38th International Conference on Machine Learning. Virtual

Event: PMLR, 2021. 12310–12320

- 41 Bardes A, Ponce J, LeCun Y. VICReg: Variance-invariance-covariance regularization for self-supervised learning. In: Proceedings of the 10th International Conference on Learning Representations. Virtual Event: OpenReview.net, 2022.



王友清 北京化工大学信息科学与技术学院教授. 主要研究方向为工业大数据分析.

E-mail: wang_youqing@buct.edu.cn
(**WANG You-Qing** Professor at the College of Information Science and Technology, Beijing University

of Chemical Technology. His main research interest is industrial big data analysis.)



徐世龙 北京化工大学信息科学与技术学院硕士研究生. 主要研究方向为图神经网络和图表示学习.

E-mail: xushilong@buct.edu.cn
(**XU Shi-Long** Master student at the College of Information Science and Technology, Beijing University

of Chemical Technology. His research interests include graph neural network and graph representation learning.)



赵天祥 北京化工大学信息科学与技术学院硕士研究生. 主要研究方向为机器学习.

E-mail: tianxiangzhao@buct.edu.cn
(**ZHAO Tian-Xiang** Master student at the College of Information Science and Technology, Beijing

University of Chemical Technology. His main research interest is machine learning.)



王宇晨 北京化工大学信息科学与技术学院硕士研究生. 主要研究方向为图神经网络和图异常检测.

E-mail: 2025200789@buct.edu.cn
(**WANG Yu-Chen** Master student at the College of Information Science and Technology, Beijing University

of Chemical Technology. His research interests

include graph neural networks and graph anomaly detection.)



辛梦媛 北京化工大学信息科学与技术学院博士研究生. 主要研究方向为对抗样本学习.

E-mail: 2025400300@buct.edu.cn
(**XIN Meng-Yuan** Ph.D. candidate at the College of Information Science and Technology, Beijing

University of Chemical Technology. Her main research interest is adversarial example learning.)



张琦 山东农业大学信息科学与工程学院副教授. 主要研究方向为图机器学习, 知识图谱和大语言模型.

E-mail: QiZhang2025@sdau.edu.cn
(**ZHANG Qi** Associate professor at the School of Information Science and Engineering, Shandong Agricultural

University. His research interests include graph machine learning, knowledge graph and large language models.)



苏 烨 中国科学院深圳先进技术研究院博士研究生. 主要研究方向为机器学习与模式识别.

E-mail: ye.su@siat.ac.cn
(**SU Ye** Ph.D. candidate at Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences.

His research interests include machine learning and pattern recognition.)



郭继鹏 北京化工大学信息科学与技术学院副教授. 主要研究方向为机器学习、模式识别和工业过程数据分析. 本文通信作者.

E-mail: guojipeng@buct.edu.cn
(**GUO Ji-Peng** Associate professor at the College of Information Science and Technology, Beijing University of Chemical

Technology. His research interests include machine learning, pattern recognition and industrial process data analysis. Corresponding author of this paper.)