计 算 机 研 究 与 发 展
Journal of Computer Research and Development
https://crad.ict.ac.cn
63(10): 2666−2681, 2026

融合动态原型与自监督迁移学习的跨域图异常检测

胡耀平[1] 彭淑娟[1,2] 柳 欣[1,2] 崔 振[3] 朱 磊[4]

[1](华侨大学计算机科学与技术学院 福建厦门 361021)

[2](福建省大数据智能与安全重点实验室(华侨大学) 福建厦门 361021)

[3](北京师范大学人工智能学院 北京 100875)

[4](同济大学计算机科学与技术学院 上海 201804)

(hyaoping@stu.hqu.edu.cn)

Cross-Domain Graph Anomaly Detection via Fusion of Dynamic Prototypes and Self-Supervised Transfer Learning

Hu Yaoping[1], Peng Shujuan[1,2], Liu Xin[1,2], Cui Zhen[3], and Zhu Lei[4]

[1](*School of Computer Science and Technology*, *Huaqiao University*, *Xiamen*, *Fujian* 361021)

[2](*Fujian Key Laboratory of Big Data Intelligence and Security* (*Huaqiao University*), *Xiamen*, *Fujian* 361021)

[3](*School of Artificial Intelligence*, *Beijing Normal University*, *Beijing* 100875)

[4](*School of Computer Science and Technology*, *Tongji University*, *Shanghai* 201804)

Abstract Cross-domain graph anomaly detection aims to identify anomalies in a target-domain graph with limited labels, by leveraging both normal and anomalous samples from a source-domain graph. However, significant discrepancies in graph structures and feature distributions between different domains hinder effective knowledge transfer and limit cross-domain generalization. Moreover, the scarcity of labeled data in the target domain often induces ambiguous decision boundaries and unstable predictions during training. To address these challenges, we propose a cross-domain graph anomaly detection method that integrates dynamic prototypes with self-supervised transfer learning. Specifically, semantic prototypes are first constructed from the normal and anomaly samples in the source domain, and a bidirectional guidance strategy is then employed to align these prototypes with target-domain nodes in a shared semantic space, thereby alleviating cross-domain distribution shifts. To further enhance the stability and semantic discriminability, an exponential moving average mechanism is introduced for dynamic prototype updating. Subsequently, entropy minimization is applied to target-domain node embeddings to improve feature discriminability and reduce decision boundary ambiguity. Finally, pseudo anomaly scores can be well obtained by combining source-domain model predictions with unsupervised clustering results on the target domain data. Accordingly, high-confidence samples are selected as week supervision signals to iteratively refine the model, thereby improving both detection accuracy and robustness. Experimental results demonstrate that our method consistently outperforms state-of-the-art approaches in various cross-domain graph anomaly detection tasks, while validating its effectiveness and stability in cross-domain transfer scenarios.

收稿日期: 2025-11-28; **修回日期**: 2026-05-26

基金项目: 厦门市自然科学基金项目(3502Z202473043); 国家自然科学基金项目(62576143, 62476103); 福建省自然科学基金项目(2024J01096); 厦门市科技计划项目(3502Z20251019); 华侨大学中央高校基本科研业务费专项资金资助项目(2026HQYJ01)

This work was supported by the Natural Science Foundation of Xiamen City (3502Z202473043), the National Natural Science Foundation of China (62576143, 62476103), the Natural Science Foundation of Fujian Province (2024J01096), the Xiamen Science and Technology Project (3502Z20251019), and the Fundamental Research Funds for the Central Universities of Huaqiao University (2026HQYJ01).

通信作者: 彭淑娟(pshujuan@hqu.edu.cn)

Key words cross-domain graph anomaly detection; graph neural networks; unsupervised domain adaptation; prototype learning; information entropy

摘 要 跨域图异常检测旨在利用源域图中的正常与异常样本信息，在标注稀缺的目标域图上实现有效的异常识别。然而，源域与目标域之间通常存在显著的图结构差异与特征分布偏移，导致源域知识难以有效迁移至目标域，限制了跨域泛化能力。此外，目标域缺乏标注信息，易引发训练过程中的决策边界模糊和预测不稳定问题。为解决上述问题，提出了一种融合动态原型与自监督迁移学习的跨域图异常检测方法。首先，基于源域中的正常与异常样本构建语义原型，并结合双向引导策略，实现源域原型与目标域节点在共享语义空间中的知识对齐，从而缓解跨域图之间的分布偏移。为进一步增强原型的稳定性与语义判别性，引入指数移动平均机制实现原型的动态更新；随后，通过最小化目标域节点嵌入的熵值，增强特征表示的区分度，以缓解决策边界模糊问题；最后，综合利用源域模型与无监督方法在目标域数据上生成异常评分的伪标签，并筛选高置信度样本作为弱监督信号引导模型迭代优化，从而提升识别的准确性与鲁棒性。实验结果表明，该方法在多个跨域图异常检测任务中均优于现有主流方法，验证了其在跨域迁移场景下的有效性与稳定性。

关键词 跨域图异常检测；图神经网络；无监督领域自适应；原型学习；信息熵

中图法分类号 TP391

DOI: 10.7544/issn1000-1239.202550857 **CSTR:** 32373.14.issn1000-1239.202550857

图结构是一种以节点和边为基本组成要素的数据模型，能够有效抽象并刻画实体之间的复杂关联关系，因此在金融、电子商务和社交媒体等多个领域中被广泛应用于关联建模与智能分析等任务。例如，在金融领域，图结构可用于构建交易网络，从而辅助欺诈行为识别与风险传导路径的追踪；在电子商务场景中，通过建模用户与商品之间的多维交互关系，能够支持异常交易检测与用户画像构建；在社交媒体分析中，图结构可用于表征用户间的社交关系，进而服务于内容治理、社交行为分析及推荐系统的优化。

随着图数据分析技术的持续发展，以图结构为核心的信息提取、表征学习与异常检测等相关任务，已成为学术界与工业界共同关注的研究重点。在该领域中，图异常检测旨在识别图中显著偏离常规模式的节点或子结构，因其在安全风控、内容治理与系统状态监控等多个关键场景中具备重要应用价值，已成为当前研究的一项关键任务。然而，图数据通常具有复杂的拓扑关联与多样化的语义层次，如何在非欧几里得数据结构中，协同融合局部与全局特征以有效建模正常行为模式，并实现高精度的异常辨识，仍是当前图异常检测研究面临的关键挑战。

为应对上述挑战，研究者已提出多种图异常检测方法。早期方法主要基于浅层图模型，依赖人工构建的拓扑特征指标（如中心性、子图密度、节点孤立度等）进行结构特征提取，并结合子图模式识别、迭代评分机制或图传播策略实现异常判定。近年来，图神经网络（graph neural networks, GNNs）凭借其强大的节点与结构表征能力，逐渐成为图结构分析的主流技术路径。该类方法通过多层邻域信息聚合机制，能够有效捕捉图中的高阶依赖关系，从而显著提升了图异常检测的综合性能与跨场景泛化能力。

随着研究的深入与应用场景的不断拓展，现有基于图神经网络的图异常检测方法虽然在单一分布数据任务中表现良好，但其泛化能力仍受限于训练数据与测试数据同分布的假设。在实际应用环境中，目标域的图拓扑结构与节点属性分布往往与源域存在显著差异，导致模型在面对分布外图数据时性能不稳定甚至显著下降。为克服传统图异常检测方法局限于单一领域建模、难以实现跨域知识迁移的瓶颈，跨域图异常检测逐渐成为该领域的一个重要研究方向。该方向的核心思想是：将具备完整异常与正常标签的图数据作为源域，通过迁移其图结构信息与语义特征，辅助标签稀缺或完全缺失的目标域实现有效的异常识别。相较于传统单域方法，跨域策略的优势在于能够在无监督或弱监督条件下，突破领域壁垒，实现源域中丰富判别信息的有效复用。该方法不仅有助于缓解因异常样本稀缺而导致的模型性能限制，也为进一步提升检测精度和降低误报率提供了新的技术路径与研究思路。

近年来，跨域图异常检测领域涌现出一系列创新方法。例如，COMMANDER[1]通过引入图注意力编码器与域判别器实现跨域特征适配，并融合异常

分类器以识别共享异常,并同时借助属性解码器捕捉非共享异常,从而协同完成属性图的跨域异常检测任务;ACT[2]则采用边缘分布差异最小化策略,利用预训练模型生成伪标签以构建辅助学习目标,旨在增强模型的跨域判别能力。GTGD[3]通过图迁移与图解耦机制同时建模域共享异常与目标域特有异常,并结合节点级与结构级差异提升跨域检测能力;GADT3[4]则将测试时训练引入跨域图异常检测,通过基于同质性的自监督信号在推理阶段动态适配目标域分布,从而提升模型对异构图域的适应能力。这些方法在提升跨域图异常检测准确率方面取得了一定成效,也为后续研究提供了有益的技术借鉴与探索方向。尽管这些方法在跨域图异常检测中取得了初步进展,仍面临2个关键挑战:1)源域知识迁移性能受限问题。当源域与目标域之间存在显著图结构差异或特征分布偏移时,源域中的判别性特征难以有效保留与迁移,从而限制了模型的跨域泛化能力。2)目标域判别能力不足问题。由于目标域中标签稀缺甚至完全缺失,模型在判别过程中常面临决策边界模糊与预测结果不稳定的问题,进一步制约了检测效果的提升。

针对上述挑战,本文提出一种融合动态原型与自监督迁移学习的跨域图异常检测方法。首先,在源域数据上进行有监督预训练,以充分学习正常与异常样本的判别性特征表示。在此基础上,构建源域类别原型,并设计双向引导策略,实现源域原型与目标域节点在共享语义空间中的结构化对齐,从而缓解因图结构差异与分布偏移所带来的知识迁移障碍。为增强原型表示的稳定性与语义一致性,本文进一步引入指数移动平均策略,实现原型向量的动态更新。此外,通过最小化目标域节点嵌入的熵值,优化其特征表示的判别性,以缓解因目标域监督信息缺失所导致的决策边界模糊问题。最后,融合源域模型预测与无监督异常评分生成伪标签,筛选高置信度样本作为监督信号,驱动模型在目标域上进行迭代优化,从而提升异常识别的准确率与鲁棒性。本文的主要贡献概括如下:

1)提出一种面向跨域图异常检测的动态跨域原型对齐机制,基于源域类别均值构建监督语义原型,并通过指数移动平均机制进行动态更新,结合双向引导实现源-目标图在语义空间中的对齐,缓解因图结构差异与分布偏移带来的知识迁移障碍。

2)设计了一种基于嵌入熵最小化的跨域特征分布优化模块,通过最小化目标域节点嵌入的熵值,引导目标图节点在低熵区域分布,提升特征判别性,改善监督稀缺下的决策边界模糊问题。

3)构建基于源域判别知识与无监督评分融合的自监督迁移学习模块,通过伪标签生成与置信度筛选生成弱监督信号,用于引导模型迭代优化,在弱监督条件下显著提升了方法的识别准确性与模型鲁棒性。

4)在多个主流跨域图异常检测数据集上开展了系统的实验与对比分析,验证了本文方法在检测性能与模型鲁棒性方面均优于现有主流方法,展现出其在实际跨域场景下的有效性与适用性。

1 相关工作

1.1 图异常检测

图异常检测(graph anomaly detection, GAD)旨在识别图数据中偏离正常模式的节点、边或子结构。与传统向量空间数据不同,图数据具有非欧几里得结构,使得传统方法难以直接建模,带来了独特建模挑战。然而,因其在金融反欺诈、社交媒体虚假账号识别、网络安全威胁监控和虚假新闻溯源等关键任务中具有广泛的应用价值,图异常检测已逐渐成为学术界与工业界广泛关注的重要研究方向[5]。

在现实场景中,图数据中的异常样本通常数量稀少且标注成本高昂,早期研究多采用无监督学习策略进行建模。传统方法通常依赖于人工设计的图结构特征(如中心性、孤立度、密度、稀疏性等),通过统计分析识别结构或属性上显著偏离的实体。典型方法包括基于密集子图识别的策略[6-7]、迭代评分机制[8-11]以及图上传播模型[12-13]等。然而,这些方法受限于人工特征工程的主观性与较差的泛化能力,难以有效捕捉图中的高阶结构信息和深层语义关联,从而影响其在复杂图结构场景中的表现力。

近年来,图神经网络的快速发展推动了图异常检测向深度学习方法演进[14]。通过聚合多跳邻域信息,图神经网络能更有效地学习具有语义表征能力的图嵌入表示,显著提升了异常模式识别的准确性。目前主流的深度检测方法大致可分为基于图自编码器的方法和基于图对比学习的方法2类。图自编码器类方法(如DOMINANT[15]和GUIDE[16])通常采用图卷积网络(GCN)[17]构建编码器-解码器结构,以节点属性和图结构的重建误差作为异常评分依据;相比之下,基于图对比的学习,如CoLA[18]和GRADATE[19],通过构建正负对样本(如节点-子图对),驱动模型学习更具判别性的节点表示,有效增强正负样本间的

语义区分度，从而优化异常识别的可解释性与模型泛化能力。

随着图数据形式的日益复杂，图异常检测的研究范畴正逐步从传统的静态同质图扩展到更具挑战性的场景，包括动态图[20]、异构图[21]以及跨域图[1-2]等。这些新兴场景不仅对模型的表征能力和泛化性能提出了更高要求，也催生了跨域图异常检测这一重要研究方向。

1.2 跨域图异常检测

跨域图异常检测（cross-domain graph anomaly detection, CDGAD）是图异常检测的重要延伸方向，其核心思想是在源域图具有充足标注数据的前提下，将源域图上的判别知识迁移到目标域图，辅助识别目标域图中标签未知的异常节点或子结构。

尽管当前图异常检测方法在单一领域无监督设定下表现良好，但其效果往往依赖于训练数据与测试数据在拓扑结构和特征分布上的相似性。在实际跨域场景中，源域与目标域之间常常存在显著的分布偏移和结构差异，这对模型的迁移与泛化能力提出新的挑战。因此，如何有效提取源域中具有语义一致性和判别能力的特征表示，并在目标域中有效复用，成为跨域图异常检测研究的关键难点。

在此背景下，Ding 等人[1]提出的 COMMANDER 框架，以图注意力编码器（GAT）[22]为核心编码器，构建共享图表示学习结构，对源域和目标域的属性图进行联合编码。通过与域判别器的对抗学习，缩小跨域差异，学习到具有域不变特性的节点表示；同时利用源域标签训练异常分类器来捕捉跨域共享异常，并引入属性解码器，通过节点属性重构误差检测目标域特有异常。Wang 等人[2]提出的 ACT 方法融合了对比学习与异常感知对齐机制，通过无监督对比学习优化目标图中正常节点的表示，使其在表示空间更具可分性。进一步引入跨域对齐策略，通过拉近目标域与源域中正常样本的距离，并拉远与异常样本的距离，提升模型在不同领域下的语义一致性表达能力。此外，该方法还采用了自标记偏差学习机制，进一步增强模型对目标域中潜在异常节点的识别性能。Huang 等人[3]提出的 GTGD，通过图迁移与图解耦机制同时建模域共享异常与目标域特有异常，并结合节点级与结构级差异提升跨域检测能力；Delaram 等人[4]提出的 GADT3，将测试时训练引入跨域图异常检测，利用基于同质性的自监督信号在推理阶段自适应目标域分布，从而提升模型对异构图域的适应能力。

除上述直接面向跨域图异常检测的方法外，近年来也有部分工作从原型学习或提示学习角度探索图异常检测的跨图泛化能力，例如 AnomalyGFM[23]和 UNPrompt[24]。前者主要在残差空间中学习任务无关的原型表示，后者则借助统一邻域提示增强模型在零样本或少样本场景下对未见图的泛化能力。总体来看，这类方法主要关注跨图泛化或低样本场景，而本文所研究的问题设定为源域有标注、目标域无标注的跨域图异常检测，二者在任务假设上存在差异。

尽管现有方法在跨域图异常检测方面已取得一定进展，但实际迁移过程中仍面临 2 个关键挑战：一方面，当源域与目标域间存在结构差异或特征分布偏移时，模型难以高效提取并复用源域中具有判别性的特征表示，导致知识迁移效率受限；另一方面，在目标域缺乏标注数据的情况下，模型难以构建有效的判别边界，从而影响对异常节点或边的识别精度。上述问题反映出当前方法在目标图特征适配性与判别能力方面仍有不足，为构建更具鲁棒性与自适应能力的跨域建模框架提供了研究动机。

1.3 无监督领域自适应

无监督领域自适应（unsupervised domain adaptation, UDA）旨在利用已标注的源域数据提升模型在目标域任务中的泛化性能，其核心思想是通过缓解源域与目标域在数据分布上的差异，从而增强模型在目标域上的适应能力。

目前主流的 UDA 方法大致可分为 2 类：一类是基于分布对齐的策略，通过引入最大均值差异[25-26]、Wasserstein 距离或 Gromov-Wasserstein 距离等统计量[27-28]，直接对齐源域与目标域的特征分布；另一类是基于对抗学习的策略，通过引入域分类器并通过对抗训练机制增强特征表示的跨域一致性。以 DANN[29]为代表，该方法通过最大化域分类器的预测熵，反向激励特征编码器生成对域不敏感的表示，从而实现源域和目标域之间的特征对齐。

近年来，研究者开始将上述 2 类策略拓展至图神经网络领域。例如，UDA-GCN[30]引入双图编码器，并结合对抗训练与注意力融合机制，实现节点表示的跨域对齐；AdaGCN[31]融合半监督与对抗训练策略以提升模型对目标图的适应能力。此外，GCLN[32]，CoCo[33]等方法引入图对比学习，通过构建正负样本对实现“语义吸引”与“语义排斥”，强化表示区分性。

图结构信息作为一个重要建模维度，也在跨域迁移中发挥了重要作用。例如，GraphAE[34]在传统特

征分布对齐的基础上引入基于结构信息的对齐策略，以缓解图结构演化带来的领域偏移；JHGDA[35]则基于图的层次结构，通过融合不同层级的信息差异，实现对结构偏移更全面的建模与处理。

尽管上述 UDA 方法在图结构数据中展现出一定的跨域泛化能力，但在应用于图异常检测任务时仍面临显著挑战。图异常检测任务具有数据高度不平衡、异常样本稀少以及图结构异质性强等问题，使得源域中学得的判别性模式在跨域迁移过程中容易被正常样本主导，导致异常信号的表达能力被弱化或淹没。

值得注意的是，在无监督条件下对齐特征分布，除了可以直接对齐特征本身，还可以利用分类器输出的置信度作为对齐的中间表征。例如，卓君宝等人[36]在跨域图像检索任务中提出的最小熵迁移对抗散列方法，其核心思想是在目标域上最小化特征分类响应的熵，促使目标域样本与源域样本在类别响应空间上分布对齐，从而学习到更具域不变性的特征。这一策略为在弱监督场景下优化目标域特征分布提供了有效的启发。

因此，在无监督跨域设定下，设计一种能够在缓解特征与结构差异的同时，保留并强化异常模式判别能力的对齐机制，已成为当前图异常检测中跨域迁移研究亟待突破的核心问题。

2 问题陈述

本文聚焦于跨域图异常检测任务，旨在解决在源域与目标域之间存在显著结构与特征分布差异时，如何有效利用源域图的监督信号，在目标域中无标签图及无标签的设定下，实现异常节点的高效识别。设$\mathcal{G}=(V,E,\boldsymbol{X})$表示一个属性图，其中$V=\{v_1,v_2,\cdots,v_n\}$为所有节点的集合，$|V|=n$，$E\subseteq V\times V$表示边集合，$\boldsymbol{X}\in\mathbb{R}^{n\times d}$表示节点特征矩阵，其第$i$行$\boldsymbol{x}_i$对应节点$v_i$的$d$维特征向量。

在跨域的设定下，设存在一个带有节点标签的源域图$\mathcal{G}_{\mathrm{s}}=(V_{\mathrm{s}},E_{\mathrm{s}},\boldsymbol{X}_{\mathrm{s}},\boldsymbol{Y}_{\mathrm{s}})$，其节点标签向量$\boldsymbol{Y}_{\mathrm{s}}\in\{0,1\}^{n_{\mathrm{s}}}$提供节点级别标注，其中标签值 0 表示节点正常，1 表示节点为异常。同时，设存在一个无标签的目标域图$\mathcal{G}_{\mathrm{t}}=(V_{\mathrm{t}},E_{\mathrm{t}},\boldsymbol{X}_{\mathrm{t}})$，用于目标域上的异常识别任务。

由于源域与目标域在图结构与特征分布上往往存在显著差异，使得二者在语义表示空间中可能呈现不一致性，增加了知识迁移与对齐的难度。为此，本文分别构建 2 个结构相同的的图神经网络模型ψ_{s}与ψ_{t}，分别用于学习源域图$\mathcal{G}_{\mathrm{s}}$和目标域图$\mathcal{G}_{\mathrm{t}}$的低维嵌入表示。表示学习过程为：

$$\begin{aligned}\boldsymbol{H}_{\mathrm{s}}&=\psi_{\mathrm{s}}(\mathcal{G}_{\mathrm{s}})=[\psi_{\mathrm{s}}(\boldsymbol{x}_1^{\mathrm{T}}),\psi_{\mathrm{s}}(\boldsymbol{x}_2^{\mathrm{T}}),\cdots,\psi_{\mathrm{s}}(\boldsymbol{x}_{n_{\mathrm{s}}}^{\mathrm{T}})]\,,\\ \boldsymbol{H}_{\mathrm{t}}&=\psi_{\mathrm{t}}(\mathcal{G}_{\mathrm{t}})=[\psi_{\mathrm{t}}(\boldsymbol{x}_1^{\mathrm{T}}),\psi_{\mathrm{t}}(\boldsymbol{x}_2^{\mathrm{T}}),\cdots,\psi_{\mathrm{t}}(\boldsymbol{x}_{n_{\mathrm{t}}}^{\mathrm{T}})]\,,\end{aligned}\tag{1}$$

其中$\boldsymbol{H}_{\mathrm{s}}\in\mathbb{R}^{n_{\mathrm{s}}\times d}$和$\boldsymbol{H}_{\mathrm{t}}\in\mathbb{R}^{n_{\mathrm{t}}\times d}$分别表示源域和目标域节点的嵌入表示矩阵，$n_{\mathrm{s}}$和$n_{\mathrm{t}}$分别为源域和目标域图的节点数量，$d$表示嵌入向量的维度。

基于上述嵌入表示，本文的目标是：在目标域无监督的设定下，充分利用源域与目标域图的数据，构建一个具有强泛化能力和良好异常感知性能的异常评分函数$\eta(\cdot)$。其形式如下：

$$\hat{y}_i=\eta(\psi_{\mathrm{t}}(\boldsymbol{x}_i))\in[0,1]\,,\tag{2}$$

其中$\hat{y}_i$数值越高，表示节点v_i越可能为异常节点。

3 本文方法

本文提出了一种融合动态原型与自监督迁移学习的跨域图异常检测方法。方法的整体框架如图 1 所示，主要包括 3 个关键阶段：源域异常建模、语义空间对齐以及自监督迁移学习。

第 1 阶段，基于源域的有监督数据，构建正常样本与异常样本的嵌入原型，并设计具有判别能力的异常评分机制，从而为目标域的跨域迁移提供了可靠的先验知识与判别基础。第 2 阶段，引入动态原型对齐策略，以缓解源域与目标域在图结构与特征分布上的差异，实现嵌入空间的有效对齐。同时，结合嵌入熵最小化机制，约束目标域节点的表示分布，提升判别稳定性，增强模型在未标注目标域中异常模式的识别与适应能力。第 3 阶段，在对齐后的统一嵌入空间基础上，结合源域的判别先验和无监督异常评分策略来生成伪标签，并筛选出高置信度伪标签作为弱监督信号，用于对目标域的图编码器与异常打分器进行联合优化与微调，从而进一步提升模型在无标签目标域上的稳定性与检测性能。

通过上述 3 个阶段的协同优化，本文方法能够有效应对图结构差异和特征空间偏移所带来的挑战，显著提升了跨域环境下的模型判别能力及鲁棒性。图 2 直观展示了训练过程中目标域图节点嵌入的分布演化，进一步印证了本文方法在异常特征分离与检测性能上的优势。

3.1 源域异常建模阶段

在源域异常建模过程中，本文构建了一个基于图神经网络ψ_{s}的结构化学习模块，用于对图中节点

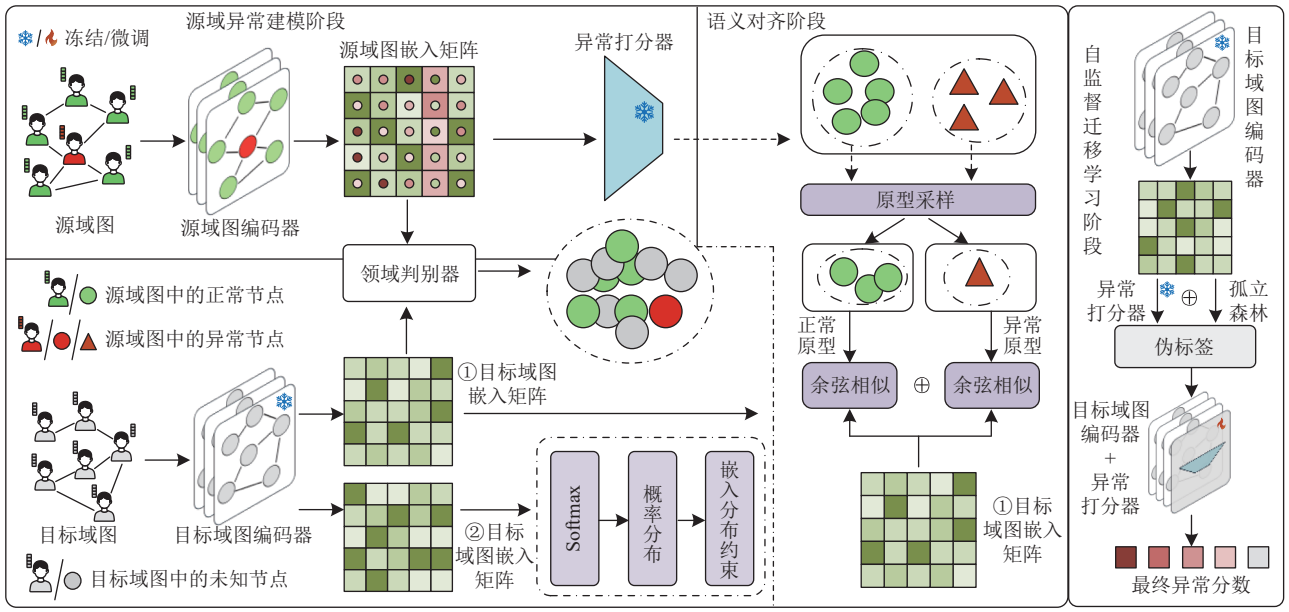
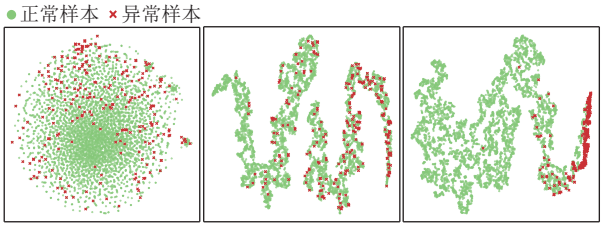


Fig. 1 Framework of the cross-domain graph anomaly detection model fusing dynamic prototypes and self-supervised transfer learning

图 1 融合动态原型与自监督迁移学习的跨域图异常检测模型框架



(a) 原始分布 (b) 语义对齐后的分布 (c) 自监督训练后的分布

Fig. 2 Feature distributions of each training stage

图 2 各训练阶段的特征分布

及其邻域的多语义特征进行建模，并生成具备良好判别能力的嵌入表示$\boldsymbol{H}_s$。该模块采用基于归纳式学习的 GraphSAGE[37] 框架，通过多层邻域聚合与特征变换的方式提取节点的上下文依赖关系。具体而言，在第l层传播过程中，节点v_i的表示$\boldsymbol{h}_i^{(l)}$通过其自身的上一层嵌入$\boldsymbol{h}_i^{(l-1)}$与其 1 阶邻域$\mathcal{N}(v_i)$内节点的聚合信息共同更新。其传播机制为：

$$\begin{aligned}\boldsymbol{a}_i^{(l)} &= AGG\left(\left\{\boldsymbol{h}_i^{(l-1)} | \forall j \in \mathcal{N}(v_i)\right\}\right), \\ \boldsymbol{z}_i^{(l)} &= CONCAT\left(\boldsymbol{h}_i^{(l-1)}, \boldsymbol{a}_i^{(l)}\right), \\ \boldsymbol{h}_i^{(l)} &= \sigma\left(\boldsymbol{W}^{(l)} \cdot \boldsymbol{z}_i^{(l)}\right),\end{aligned} \tag{3}$$

其中$AGG(\cdot)$表示邻域特征聚合函数，$CONCAT(\cdot)$表示向量拼接操作，σ为激活函数，$\boldsymbol{W}^{(l)}$为第l层的可学习权重矩阵。

为缓解深层网络中可能出现的表示退化问题，本文进一步引入了残差连接机制，使节点在多层聚合过程中能保留原始信息结构，进而增强模型的鲁棒性与训练稳定性。

在图表示学习的基础上，本文构建了异常打分器η^s，用于建模源域中正常节点与异常节点在嵌入空间中的分布特征。异常打分器η^s是一个将d维节点嵌入向量映射至 1 维异常评分的线性函数，定义为：

$$\hat{y}_i = \eta^s(\boldsymbol{h}_i) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{h}_i + b, \tag{4}$$

其中$\boldsymbol{w} \in \mathbb{R}^d$和$b \in \mathbb{R}$分别表示可学习的权重向量与偏置项，$\hat{y}_i$表示节点$v_i$的异常得分。

为了优化打分器在源域中的判别能力，本文采用一种基于偏差的损失函数[38]，其核心设计在于促使正常样本的异常得分分布尽可能集中，从而使得异常样本的得分显著偏离该分布，实现更好的判别效果。具体形式为：

$$\mathcal{L}_{dev} = (1-y)|dev(\hat{y}_i)| + y \cdot \max(0, a - dev(\hat{y}_i)) \text{。} \tag{5}$$

如果v_i是一个异常节点则$y=1$，否则$y=0$；$dev(\cdot)$是一个基于偏差的 Z-Score 函数：

$$dev(\boldsymbol{h}_v) = \frac{\hat{y}_i - \mu_s}{\sigma_s}, \tag{6}$$

其中μ_s，σ_s，a分别是异常分数的平均值、标准差和偏离平均值的距离。

图神经网络ψ_s与异常打分器η^s通过联合训练不断优化，最终构建出一个判别性强的嵌入空间，使得源域中的正常节点与异常节点在该空间中可以被一个线性超平面有效区分。

在源域上训练完成后，本文保留异常打分器η^s的权重$\boldsymbol{w}$和b，并将其迁移至目标域，用于对目标域节点嵌入进行异常评分与伪标签生成。该设计为目标域

在缺乏真实标注信息的情况下,提供了一种可靠的异常识别先验信息,从而增强了模型在无监督迁移过程中的判别能力与泛化性能。

3.2 语义空间对齐阶段

针对跨域图异常检测任务中面临的源域知识迁移不充分以及目标域异常判别边界模糊的问题,本文设计了动态原型对齐机制和嵌入熵最小化的分布优化策略2个关键模块,旨在提升模型的跨域迁移能力和判别鲁棒性。

3.2.1 动态原型对齐

为充分挖掘源域图中节点的类别语义信息,并有效引导目标域节点的嵌入分布与源域对齐,本文提出动态原型对齐(dynamic prototype alignment, DPA)机制。不同于已有方法依赖无监督原型学习或聚类重估原型中心的建模方式,本文直接利用源域已标注正常样本与异常样本的嵌入均值来构建类别原型,从而赋予原型明确的监督语义。同时,在原型更新过程中,本文采用指数移动平均(exponential moving average, EMA)进行平滑维护,以增强原型表示的稳定性并提升跨域对齐效果。具体而言,本文基于源域中正常样本与异常样本,分别构建正常类别原型 $\boldsymbol{P}_{\text{normal}} \in \mathbb{R}^d$ 与异常类别原型 $\boldsymbol{P}_{\text{anomaly}} \in \mathbb{R}^d$,并结合双向引导与动态更新机制,实现源域原型与目标域节点在共享语义空间中的结构化对齐。此设计有效缓解了跨域图数据之间由于图结构差异与特征分布偏移所导致的知识迁移障碍,显著增强了跨域场景下的异常检测鲁棒性与判别稳定性。

在训练初始阶段,本文基于源域已标注的正常样本集 $\mathcal{N}_{\text{s}}$ 和异常样本集 $\mathcal{A}_{\text{s}}$,分别计算其嵌入表示的均值作为初始类别原型:

$$\boldsymbol{P}_{\text{normal}}^{(0)} = \frac{1}{|\mathcal{N}_{\text{s}}|}\sum_{v_i \in \mathcal{N}_{\text{s}}} \boldsymbol{h}_i,\quad \boldsymbol{P}_{\text{anomaly}}^{(0)} = \frac{1}{|\mathcal{A}_{\text{s}}|}\sum_{v_i \in \mathcal{A}_{\text{s}}} \boldsymbol{h}_i \,, \tag{7}$$

其中 $\boldsymbol{h}_i$ 为节点 v_i 的嵌入向量。

考虑到目标域中样本分布可能随训练进程发生演化,本文采用EMA策略对类别原型进行动态更新,在保障更新稳定性的同时提升模型对目标域分布演化的适应能力:

$$\begin{aligned}
\boldsymbol{P}_{\text{normal}}^{(l+1)} &= \boldsymbol{P}_{\text{normal}}^{(l)} + \alpha \cdot \left(\frac{1}{\min(a, |\mathcal{N}^{(l)}|)} \sum_{v_i \in \mathcal{N}^{(l)}} \boldsymbol{h}_i \right), \\
\boldsymbol{P}_{\text{anomaly}}^{(l+1)} &= \boldsymbol{P}_{\text{anomaly}}^{(l)} + \alpha \cdot \left(\frac{1}{\min(b, |\mathcal{A}^{(l)}|)} \sum_{v_i \in \mathcal{A}^{(l)}} \boldsymbol{h}_i \right),
\end{aligned} \tag{8}$$

其中 α 是控制原型更新步长与稳定性的衰减权重因子,a 和 b 为控制聚合数量的超参数,$\mathcal{N}^{(l)}$ 和 $\mathcal{A}^{(l)}$ 表示第 l 次迭代中源域中正常样本与异常样本的嵌入集合。min(·)是取最小值,确保原型更新的稳定。

基于所构建的原型,本文设计一种面向目标域节点的双向引导对齐损失函数,从“靠近正常原型”与“远离异常原型”2个方向共同优化目标域节点在嵌入空间中的分布。

具体而言,正常类别对齐损失旨在引导目标域样本向正常原型聚集。该损失通过对目标节点与正常原型相似度低于预设阈值的情况施加惩罚,增强其在语义空间中与正常类别的匹配程度:

$$\mathcal{L}_{\text{normal}} = E\left[\max\left(0, M - \frac{\psi_{\text{t}}(\boldsymbol{x}_{\text{t}})^{\text{T}} \cdot \boldsymbol{P}_{\text{normal}}}{\|\psi_{\text{t}}(\boldsymbol{x}_{\text{t}})\|_2 \cdot \|\boldsymbol{P}_{\text{normal}}\|_2}\right)\right], \tag{9}$$

其中 $\psi(\cdot)$ 表示目标域图编码器,$\boldsymbol{x}_{\text{t}}$ 表示目标域图的节点属性,M 为相似度边距超参数。

另一方面,为防止目标域节点的嵌入表示过度接近源域异常类原型,本文引入异常类对齐损失,通过抑制目标节点与异常原型之间的语义相关性,有效限制其向异常语义区域的偏移:

$$\mathcal{L}_{\text{anomaly}} = E\left[\max\left(0, M + \frac{\psi_{\text{t}}(\boldsymbol{x}_{\text{t}})^{\text{T}} \cdot \boldsymbol{P}_{\text{anomaly}}}{\|\psi_{\text{t}}(\boldsymbol{x}_{\text{t}})\|_2 \cdot \|\boldsymbol{P}_{\text{anomaly}}\|_2}\right)\right]。 \tag{10}$$

为进一步平衡正常与异常类别对齐损失之间的相互作用,并增强模型对复杂跨域分布的适应性,本文将上述损失项进行加权融合,构建最终的原型对齐损失函数:

$$\mathcal{L}_{\text{proto}} = \mathcal{L}_{\text{normal}} + \delta \cdot \mathcal{L}_{\text{anomaly}} \,, \tag{11}$$

其中 δ 为动态调节系数,可根据训练阶段或任务需求灵活调整正常与异常类别损失的相对权重,从而实现对跨域特征对齐过程的自适应优化。

通过该损失函数的优化,目标域节点的嵌入表示逐渐向源域正常类原型靠近,并远离异常类原型,从而在共享嵌入空间中形成具有可判别性的语义对齐结构,增强跨域对齐效果与模型判别鲁棒性。

3.2.2 嵌入熵最小化的分布优化

在跨域图异常检测中,由于目标域缺乏真实标签,模型在进行无监督判别时通常面临预测置信度不足的问题,导致在嵌入空间中造成类别边界模糊,进而影响判别稳定性。为缓解此问题,本文引入嵌入熵最小化(embedding entropy minimization, EEM)机制,引导目标域节点的嵌入表示生成更加集中且语义判别性强的分布,提升模型识别的稳定性与可靠性。

具体而言,对于目标域节点的嵌入表示 $\boldsymbol{h}_i$,本文在特征维度上进行归一化处理,构建其概率分布 $\boldsymbol{p}_i = (p_{i1}, p_{i2}, \cdots, p_{id})$,数学表达为

$$p_{ij}=\frac{\exp(h_{ij})}{\sum_{k=1}^{d}\exp(h_{ik})}。\tag{12}$$

随后，基于该分布定义信息熵损失函数：

$$\mathcal{L}_{\mathrm{EEM}}=-\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{d}p_{ij}\ln(p_{ij}+\varepsilon),\tag{13}$$

其中n表示目标域节点的数量，ε是为保障数值稳定性引入的小量正则项，p_{ij}表示目标域中第i个样本在第j个维度上的归一化值。

通过最小化$\mathcal{L}_{\mathrm{EEM}}$，模型被引导生成更加集中且语义判别性强的嵌入表示，在共享语义空间中形成更加明确的类别分界，从而增强对目标域中异常样本的判别能力与预测稳定性。

3.3 自监督迁移学习阶段

在完成领域对齐之后，为进一步提升模型在目标域上的表征能力与判别性能，本文设计了自监督迁移学习机制，通过生成伪标签引导模型迭代优化。该机制主要包含 3 个步骤：

1）目标域特征编码。基于已完成跨域语义对齐的目标域编码器ψ_{t}，对目标图$\mathcal{G}_{\mathrm{t}}$的节点进行特征编码，得到其在对齐后的特征空间中的低维嵌入表示$\boldsymbol{H}_{\mathrm{t}}$。得益于前期的原型对齐与嵌入熵最小化训练，编码器$\psi_{\mathrm{t}}$已具备初步的跨域语义感知与判别性特征提取能力，使得目标域节点嵌入$\boldsymbol{H}_{\mathrm{t}}$蕴含更易于异常识别的特征信息。

2）伪标签生成与高置信候选筛选。鉴于目标域缺乏真实标签，本文提出一种融合迁移知识与本地结构的伪标签生成策略。该策略结合源域预训练的异常打分器η^{s}和通用的无监督检测方法$f(\cdot)$（本文采用孤立森林[39]），为目标域节点生成综合伪异常分数。具体计算方式为

$$\hat{y}_i=\beta\cdot\frac{\eta^{\mathrm{s}}(\boldsymbol{H}_{\mathrm{t}}^{(i)})-\mu(\eta^{\mathrm{s}}(\boldsymbol{H}_{\mathrm{t}}))}{\sigma(\eta^{\mathrm{s}}(\boldsymbol{H}_{\mathrm{t}}))}+\frac{f(\boldsymbol{H}_{\mathrm{t}}^{(i)})-\mu(f(\boldsymbol{H}_{\mathrm{t}}))}{\sigma(f(\boldsymbol{H}_{\mathrm{t}}))},\tag{14}$$

其中$\eta(\boldsymbol{H}_{\mathrm{t}}^{(i)})$和$f(\boldsymbol{H}_{\mathrm{t}}^{(i)})$分别表示异常打分器和无监督方法对目标域图节点$v_i$的异常评分。$\mu(\cdot)$和$\sigma(\cdot)$分别表示求平均值和标准差的函数。加权调节系数$\beta$用于权衡来自迁移知识的语义判断与本地结构的拓扑判断对最终决策的贡献。

为从伪异常得分$\{\hat{y}_i\}_{i=1}^{|V|}$中筛选出高置信度的异常候选，受 ACT[2] 和 Pang 等人[40] 工作的启发，本文采用基于 Cantelli 不等式的阈值筛选机制。该机制定义伪异常样本集$\mathcal{A}_{\mathrm{pseudo}}$为

$$\mathcal{A}_{\mathrm{pseudo}}=\left\{v|\hat{y}(v)>\mu_{\hat{y}}+\gamma\cdot\sigma_{\hat{y}},\forall v\in\mathcal{G}_{\mathrm{t}}\right\},\tag{15}$$

其中$\mu_{\hat{y}}$和$\sigma_{\hat{y}}$分别表示$\hat{y}$伪异常分数的均值和标准差，$\gamma>0$为控制置信度的超参数。该策略为模型联合优化提供高质量的弱监督信号，显著提升目标域在无标注条件下的异常识别与建模能力。

3）联合微调与决策边界优化。基于上述步骤生成的高置信度伪异常样本集$\mathcal{A}_{\mathrm{pseudo}}$及其补集（即伪正常样本集$\mathcal{N}_{\mathrm{pseudo}}=\mathcal{V}_{\mathrm{t}}\backslash\mathcal{A}_{\mathrm{pseudo}}$），本文对目标域编码器$\psi_{\mathrm{t}}$和异常打分器$\eta^{\mathrm{t}}$进行联合微调。其中，异常打分器$\eta^{\mathrm{t}}$的初始参数继承自源域异常打分器$\eta^{\mathrm{s}}$。在优化过程中，目标图$\mathcal{G}_{\mathrm{t}}=(V_{\mathrm{t}},E_{\mathrm{t}},\boldsymbol{X}_{\mathrm{t}})$通过编码器$\psi_{\mathrm{t}}$编码为低维嵌入表示$\boldsymbol{H}_{\mathrm{t}}'$，再由异常打分器$\eta^{\mathrm{t}}$得到异常评分$\boldsymbol{S}_{\mathrm{t}}=\eta^{\mathrm{t}}(\boldsymbol{H}_{\mathrm{t}}')$。通过最小化式（5），模型在学习目标域数据分布的同时，得以在嵌入空间中重塑更具判别性的决策边界，进而提升异常识别的准确性。该联合优化过程显著增强了模型在目标域上的检测精度与鲁棒性。

3.4 复杂度分析

为说明本文方法的计算代价与可拓展性，下面从时间复杂度和空间复杂度 2 个方面进行分析。

本文的主要计算开销来自图编码器ψ_{s}和ψ_{t}的向前传播。设输入图节点数为N，边数为m，特征维度为F，图编码器共有L层，第l层输出维度为d_l。在邻居采样机制下，一次 GraphSAGE 编码的时间复杂度为

$$O\left(\sum_{l=1}^{L}E_l d_{l-1}+N_l d_{l-1}d_l\right),\tag{16}$$

其中N_l和E_l分别表示第l层采样子图中的节点数和边数。由于本文采用固定规模的局部邻居采样，整体计算开销主要由采样子图规模决定，而不会随整图规模产生额外的 2 次增长。

除图编码外，动态原型对齐、嵌入熵最小化以及线性异常打分器均在低维嵌入空间中进行，其额外计算开销较小，整体时间复杂度仍由图神经网络编码过程主导。空间复杂度主要包括节点特征、采样子图、中间层表示及模型参数存储，可表示为

$$O\left(NF+\sum_{l=1}^{L}(N_l d_l+E_l)+|\theta|\right),\tag{17}$$

其中$|\theta|$表示模型参数规模。总体而言，本文方法的时间与空间开销均主要随采样子图规模近似线性增长，因此具有较好的可扩展性。

4 实验分析

为验证本文提出的融合动态原型与自监督迁移

学习的跨域图异常检测方法的有效性，本文在多个真实世界的图数据集上开展了系统而全面的实验研究。本节首先介绍所用数据集的构建方式与统计特性，随后说明对比基线方法、评估指标及实现细节，并对实验结果进行全面分析。为进一步验证模型中关键模块的贡献，本文还开展了消融实验与参数敏感性分析。各数据集的详细统计数据如表 1 所示。

Table 1 Details for Datasets

表 1 数据集详细信息

数据集	节点个数	属性	异常个数
RES[13]	5 012	8 000	250
HTL[13]	4 322	8 000	250
NYC[13]	21 040	10 000	1 000
AMZ[41]	18 601	10 000	726

4.1 实验数据集

本文选用 4 个来源于现实平台的真实图数据集：HTL(yelphotel)[13]，RES(yelpres)[13]，NYC(yelpnyc)[13]，AMZ(amazon)[41]。这些数据集源自不同平台的用户评论与行为数据，覆盖酒店、餐饮及电子商务等多样化场景，为模型提供了极具挑战性的跨域测试环境。

所有数据集均采用统一的构建范式：数据集图中节点代表用户，若 2 个用户评论过同一项(如酒店、餐厅、商品)，则在二者之间建立边，形成“用户—项—用户”的结构。这一设计有效建模用户之间的行为共现模式，反映潜在的协同异常或欺诈行为，数据集中的异常节点均由平台反欺诈机制或人工审核标注。4 个数据集的简要说明如下：

1)HTL。来源于芝加哥地区的酒店评论数据，标注了由平台识别的欺诈用户。

2)RES。来源于芝加哥地区的餐厅评论数据，其图构建与异常标注方式与 HTL 一致。

3)NYC。来源于纽约地区的餐厅评论数据，其结构与 RES 相同。由于地域差异，NYC 与 RES 之间的跨域任务常用于评估模型在不同城市区域下的迁移性能。

4)AMZ。来源于电子商务平台的商品评论数据，其领域异质性更强、结构更复杂。异常用户被人工审核为评论至少 2 个被标记为欺诈的商品。

本文共设计 4 组跨域实验组合，包括 HTL 与 RES、NYC 与 AMZ、RES 与 NYC 以及 HTL 与 NYC。在每一组合中，2 个数据集互为源域和目标域，构建双向跨域评估机制。这一设定有助于全面评估所提出方法在不同领域、不同结构之间进行知识迁移与异常检测的能力。

4.2 比较方法和评价指标

为系统评估本文所提方法在跨域图异常检测任务中的性能，本文共选择了 2 类共 13 种具有代表性的图异常检测方法参与对比实验：一类是基于无监督学习的图异常检测方法，另一类为面向跨域设置的图异常检测方法。2 类方法从不同角度验证了本文方法的判别能力与跨域知识迁移效果。

在无监督图异常检测方法方面，本文将图表示学习模型 DGI[42] 分别与传统异常检测算法 LOF[43] 和 IF[39] 结合，构建了 2 种基于表示学习的检测方法：DGI+LOF 和 DGI+IF。同时，本文还选取了具有代表性的无监督图异常检测方法 DOMINANT[15]，COLA[18]，GAAN[44]，ANOMALOUS[45]，ADONE[46] 进行对比。由于这些方法主要适用于单域图结构，实验中我们将其直接应用于目标域，以评估其在未引入跨域适配机制情况下的跨域泛化能力。

在跨域图异常检测方法方面，本文选用了 ADDA[47]，COMMANDER(s)[1]，COMMANDER(u)，ACT[2]，GTGD[3]，GADT3[4] 等典型模型作为对比基准。其中，COMMANDER(s) 为原始方法，其通过学习源域与目标域共享的特征表示，实现跨域知识迁移；而 COMMANDER(u) 则是其改进版本，采用分别建模源域和目标域特征表示的方式，尤其适用于 2 个域之间的特征空间存在较大差异的场景。

在评估过程中，本文采用受试者工作特征曲线下面积(receiver operating characteristic area under the curve，AUC-ROC)和准确率-召回率曲线下面积(precision-recall area under the curve，AUC-PR)作为主要评价指标，二者在模型性能评估中具有良好的互补性。AUC-ROC 反映模型在不同分类阈值下的整体区分能力，适用于类别分布较为均衡的场景。然而，在异常样本极其稀少的情况下，其表现可能无法准确反映模型对异常样本的识别能力，甚至可能高估模型的实际性能。相比之下，AUC-PR 更侧重于评估模型对少数类(即异常类)的识别能力，尤其是在类别严重不平衡的数据集中，其评估结果更具实际参考价值。因此，结合 AUC-ROC 与 AUC-PR 进行综合评估，可以更全面地反映模型在跨域图异常检测任务中的性能表现。总体而言，AUC-ROC 与 AUC-PR 的得分越高，表明模型在异常检测任务中的表现越优越。所有对比实验结果详见表 2 和表 3。

4.3 实验设置

所有对比模型均基于 PyTorch 深度学习框架进

Table 2 AUC-ROC in Comparative Experiments

表 2 对比实验中的 AUC-ROC

类型	方法	RES→HTL	NYC→HTL	HTL→RES	NYC→RES	RES→NYC	HTL→NYC	AMZ→NYC	NYC→AMZ	平均值
无监督方法	DGI+LOF	0.778±0.009	0.778±0.009	0.850±0.026	0.850±0.026	0.612±0.015	0.612±0.015	0.612±0.015	0.728±0.015	0.742±0.016
	DGI+IF	0.667±0.009	0.667±0.009	0.843±0.017	0.843±0.017	0.844±0.007	0.844±0.007	0.844±0.007	0.537±0.015	0.723±0.009
	ANOM	0.186±0.002	0.186±0.002	0.417±0.009	0.417±0.009	0.536±0.001	0.536±0.001	0.536±0.001	0.496±0.003	0.409±0.004
	DOM	0.694±0.000	0.694±0.000	0.767±0.000	0.767±0.000	0.692±0.000	0.692±0.000	0.692±0.000	0.867±0.000	0.755±0.000
	ADONE	0.738±0.035	0.738±0.035	0.477±0.024	0.477±0.024	0.632±0.036	0.632±0.036	0.632±0.036	0.847±0.052	0.671±0.037
	GAAN	0.644±0.010	0.644±0.010	0.668±0.041	0.668±0.041	0.406±0.006	0.406±0.006	0.406±0.006	0.861±0.015	0.645±0.018
	COLA	0.485±0.034	0.485±0.034	0.555±0.063	0.555±0.063	0.811±0.006	0.811±0.006	0.811±0.006	0.496±0.003	0.587±0.027
跨域方法	ADDA	0.624±0.064	0.589±0.109	0.787±0.144	0.726±0.345	0.750±0.076	0.750±0.062	0.697±0.126	0.640±0.051	0.684±0.118
	CMDR(s)	0.690±0.009	N/A	0.774±0.007	N/A	N/A	N/A	0.699±0.006	0.859±0.007	0.756±0.007
	CMDR(u)	0.699±0.009	0.707±0.009	0.763±0.024	0.780±0.017	0.694±0.006	0.693±0.002	0.695±0.001	0.848±0.009	0.751±0.012
	ACT	0.792±0.018	0.804±0.006	0.892±0.015	0.948±0.014	0.831±0.005	0.830±0.005	0.830±0.002	0.925±0.004	0.868±0.009
	GTGD	0.848±0.022	0.862±0.017	0.959±0.008	**0.992±0.007**	0.859±0.011	0.861±0.009	**0.861±0.013**	0.917±0.006	0.892±0.012
	GADT3	0.736±0.003	N/A	0.898±0.002	N/A	N/A	N/A	0.464±0.004	0.473±0.002	0.643±0.003
	本文	**0.880±0.013**	**0.891±0.017**	**0.974±0.008**	0.984±0.004	**0.864±0.004**	**0.866±0.002**	0.853±0.003	**0.959±0.013**	**0.909±0.003**

注：黑体数值表示最优值。“±”前数值表示平均值；“±”后数值表示标准差。ANOM 表示 ANOMALOUS，DOM 表示 DOMINANT，CMDR(s) 表示 COMMANDER(s)，CMDR(u) 表示 COMMANDER(u)。“N/A”并非实验失败，而是表示该方法在对应实验中不适用。

Table 3 AUC-PR in Comparative Experiments

表 3 对比实验中的 AUC-PR

类型	方法	RES→HTL	NYC→HTL	HTL→RES	NYC→RES	RES→NYC	HTL→NYC	AMZ→NYC	NYC→AMZ	平均值
无监督方法	DGI+LOF	0.247±0.016	0.247±0.016	0.269±0.036	0.269±0.036	0.133±0.006	0.133±0.006	0.133±0.006	0.097±0.008	0.186±0.017
	DGI+IF	0.197±0.013	0.197±0.013	0.296±0.031	0.296±0.031	0.366±0.014	0.366±0.014	0.366±0.014	0.042±0.003	0.226±0.009
	ANOM	0.053±0.000	0.053±0.000	0.040±0.001	0.040±0.001	0.091±0.001	0.091±0.001	0.091±0.001	0.037±0.000	0.055±0.001
	DOM	0.216±0.000	0.216±0.000	0.264±0.000	0.264±0.000	0.145±0.000	0.145±0.000	0.145±0.000	0.252±0.000	0.219±0.000
	ADONE	0.244±0.029	0.244±0.029	0.183±0.031	0.183±0.031	0.155±0.029	0.155±0.029	0.155±0.029	0.259±0.076	0.210±0.041
	GAAN	0.152±0.006	0.152±0.006	0.089±0.017	0.089±0.017	0.039±0.001	0.039±0.001	0.039±0.001	0.203±0.035	0.121±0.015
	COLA	0.082±0.009	0.082±0.009	0.109±0.011	0.109±0.011	0.128±0.003	0.128±0.003	0.128±0.003	0.037±0.000	0.089±0.006
跨域方法	ADDA	0.227±0.028	0.171±0.062	0.260±0.126	0.254±0.140	0.254±0.140	0.181±0.021	0.239±0.110	0.051±0.002	0.177±0.057
	CMDR(s)	0.210±0.007	N/A	0.268±0.006	N/A	N/A	N/A	0.145±0.001	0.242±0.019	0.216±0.008
	CMDR(u)	0.216±0.008	0.207±0.009	0.253±0.025	0.267±0.015	0.144±0.002	0.145±0.002	0.145±0.001	0.220±0.024	0.209±0.014
	ACT	0.287±0.006	0.284±0.010	0.330±0.018	0.477±0.065	0.249±0.012	0.241±0.009	0.243±0.003	0.497±0.020	0.358±0.002
	GTGD	0.368±0.032	0.379±0.020	0.563±0.024	0.776±0.031	0.388±0.017	0.380±0.012	**0.371±0.011**	0.649±0.016	0.484±0.020
	GADT3	0.168±0.002	N/A	0.223±0.001	N/A	N/A	N/A	0.048±0.005	0.035±0.001	0.119±0.003
	本文	**0.412±0.026**	**0.401±0.015**	**0.701±0.084**	**0.795±0.044**	**0.408±0.022**	**0.420±0.005**	0.338±0.012	**0.707±0.017**	**0.526±0.008**

注：黑体数值表示最优值。“±”前数值表示平均值；“±”后数值表示标准差。ANOM 表示 ANOMALOUS，DOM 表示 DOMINANT，CMDR(s) 表示 COMMANDER(s)，CMDR(u) 表示 COMMANDER(u)。“N/A”并非实验失败，而是表示该方法在对应实验中不适用。

行实现。本文所提方法在 ACT 框架基础上进行开发，并沿用其训练流程，以便在相同实验条件下进行公平对比。与 ACT 的原始领域对齐策略不同的是，本文采用 DPA，通过为源域中的正常样本与异常样本分别构建语义原型，并引入了一个额外的调控因子δ用于调节不同原型对目标域节点嵌入分布的影响。在自监督迁移学习阶段，首先将目标域图通过ψ_t编码获得其节点嵌入表示，利用η^s与经典的无监督异常检测方法——孤立森林[39]分别对目标域进行异常评分，并通过加权因子β将二者结果进行融合，生成综合异常评分，并以此筛选出高置信度的伪标签样本作为弱监督信号。随后，将ψ_t与η^t构建形成一个新的模型，并使用所构建的弱监督信息对其进行微调，以提升模型在目标域上的异常检测性能与鲁棒性。除

非特别说明,本文在所有实验中采用统一参数设置:在动态原型对齐模块中,取 $\alpha = 0.9$, $\delta = 0.3$, $M = 1.0$,其余参数a与b在所有实验中保持固定;在自监督迁移模块中,β和γ亦在各任务中保持一致。总体而言,本文仅对少数关键参数进行小范围调节,其余参数均在不同数据集与跨域任务上共享统一配置。

本文实验在 Ubuntu20.04 LTS 操作系统环境下运行,使用的硬件平台包括 NVIDIA RTX 3090 GPU 与 Intel® Xeon® Gold 6330 CPU, CUDA Toolkit 版本为 570.124.04。

4.4 实验结果分析

通过与 13 种现有图异常检测方法的对比实验,所提方法在跨域图异常检测任务中展现出良好的有效性与显著的性能优势。如表 2 和表 3 所示,本文方法在 AUC-ROC 和 AUC-PR 指标上分别取得 0.909 与 0.526 的平均得分,相较于当前最优基线 GTGD 分别提升了 1.9% 与 8.7%,充分表明本文方法在跨域设定下具有更强的整体判别能力和对异常样本的识别准确性。

进一步地,在 8 个跨域任务中,本文方法在 AUC-ROC 上有 6 个任务取得最优结果,在 AUC-PR 上有 7 个任务取得最优结果。以 AUC-ROC 为例,本文方法在 RES→HTL, NYC→HTL 与 NYC→AMZ 任务较最优基线提升约 3.8%, 3.4%, 3.7%。在 AUC-PR 指标上,HTL→RES, RES→HTL, HTL→NYC 这 3 项任务分别相较基线分别取得 24.5%, 12.0%, 10.5% 的增益,进一步表明本文方法在目标域与源域间存在明显图结构差异和语义偏移时,依旧可以保持稳定且高效的异常判别能力。

实验表明,传统单域无监督方法在跨域图异常检测任务中整体性能显著弱于跨域建模方法。以 DGI+IF 和 DGI+LOF 为例,二者本质上并非专门为图异常检测任务设计,因而其在目标域难以准确识别异常模式。而其余无监督方法是面向单域的图异常检测设计,并未引入有效的跨域迁移机制,无法利用源域中的语义与结构信息,导致其在多个设定下性能受限。上述结果说明,当前主流无监督图异常检测方法难以满足源域与目标域之间分布异质性带来的迁移建模需求,其结构设计在跨域背景下存在显著局限。

在与面向跨域建模的对比方法中,本文方法仍展现出较强竞争力。以经典跨域模型 ADDA[47] 为例,尽管其在图像任务中具备良好的领域对齐能力,但由于未设计图结构感知模块,因此在图域迁移任务中提升有限。相比之下,COMMANDER[1] 和 ACT[2] 虽针对跨域图异常检测任务进行了专门设计,但在源域语义信息与目标域嵌入分布的融合方面仍存在一定不足,性能仍有较大提升空间。需要说明的是表 2 和表 3 中的“N/A”并非实验失败,而是表示该方法在对应任务中不适用。这主要是因为部分基线要求源域与目标域具有相同的节点属性维度,而本文部分跨域任务存在属性维度不一致的情况。若额外加入输入投影或特征对齐模块,会改变原始模型设定并影响比较的公平性,因此本文将此类情形统一记为“N/A”。以 ACT 为例,其采用自监督偏移学习机制,在完成源域与目标域特征空间映射后,仅对目标域编码器进行了微调,并引入了独立的判别器模块,限制了源域中的异常特征在目标域中的表达与迁移能力。

综上所述,本文方法在多个跨域图异常检测任务中均展现出稳定的性能优势,尤其在源域与目标域存在较大结构差异与语义偏移的情况下,其异常识别表现仍然处于较高水平,表明本文方法具备较强的迁移适应能力。对比分析也进一步表明,提升跨域能力的关键在于如何有效迁移并融合源域中包含的判别性异常信息。该发现为未来研究提供了方向指引:应在建模过程中更充分地利用源域中的监督信息,并探索更具判别性的领域对齐与融合机制,以提升跨域检测的整体性能。

4.5 消融实验

为评估 DPA 与 EEM 模块在跨域语义对齐与判别能力增强方面的作用,本文构建了 2 类消融模型:分别移除 DPA 模块与 EEM 模块,并在此基础上开展 8 项跨域图异常检测任务中与完整模型的对比实验。如表 4 所示,在大多数任务中,完整模型的表现优于其消融版本(w/o DPA 和 w/o EEM),这验证了 DPA 与 EEM 模块在跨域语义对齐与判别能力增强中的重要作用。

当模型移除 DPA 模块后,整体性能在多个任务中显著下降。特别地,在 AMZ→NYC 与 NYC→AMZ 的跨域设定中,由于源域与目标域之间在节点属性分布和图拓扑结构上存在较大的领域差异,模型在缺失 DPA 模块的情况下无法生成具有判别性的伪标签,预测结果趋于正常类别集中,导致异常检测失效(对应结果标记为“N/A”)。这一现象表明,DPA 为跨域知识迁移提供了稳定且可解释的语义引导,是实现鲁棒迁移的基石。同时,我们注意到在 HTL→RES 任务中,移除 DPA 后模型性能略优于完整模型。

Table 4 Ablation Study on Relative Modules in Semantic Alignment Stage

表 4 语义对齐阶段相关模块的消融实验

指标	数据集	w/o DPA	w/o EEM	本文方法
AUC-ROC	RES→HTL	0.739	0.853	**0.880**
	NYC→HTL	0.806	0.832	**0.891**
	HTL→RES	**0.978**	0.955	0.973
	NYC→RES	0.982	0.958	**0.984**
	RES→NYC	0.854	0.861	**0.864**
	HTL→NYC	0.821	0.865	**0.865**
	AMZ→NYC	N/A	**0.855**	0.853
	NYC→AMZ	N/A	0.903	**0.959**
AUC-PR	RES→HTL	0.270	0.362	**0.412**
	NYC→HTL	0.269	0.349	**0.401**
	HTL→RES	**0.814**	0.562	0.701
	NYC→RES	0.743	0.558	**0.795**
	RES→NYC	0.371	0.406	**0.408**
	HTL→NYC	0.338	0.417	**0.420**
	AMZ→NYC	N/A	**0.343**	0.338
	NYC→AMZ	N/A	0.544	**0.707**

注：黑体数值表示最优值。

该结果表明，DPA 并非在所有场景下都能带来一致增益。其可能原因在于，RES 目标域本身已具有较强的结构判别信息，而来自源域的原型约束在该场景下可能对目标域特有异常模式产生干扰，从而引入轻微的负迁移。这表明，DPA 的有效性与源域和目标域之间的语义一致性密切相关，在分布偏移较大的场景下仍存在一定适用边界。

移除 EEM 模块后，虽然模型未完全失效，但在 HTL→RES，HTL→NYC 等任务中表现出明显预测波动，AUC-PR 指标下降显著。这表明，EEM 并非实现跨域对齐的必需条件，其核心功能是为了提升模型分类的确定性。通过最小化嵌入熵，将目标域节点表征从模糊的低置信度区域约束到更紧凑的高置信度区域。这种分布约束能显著提升预测的确定性，对于异常样本稀缺和类别不平衡的严苛场景至关重要，可有效减少误判并提升对少数类的识别精度。

在自监督迁移学习阶段，本文采用源域异常打分器η^s的参数对目标域异常打分器η^t进行初始化，旨在提升目标域模型在早期训练阶段的判别能力和收敛效率。为验证该初始化策略的有效性，本文将完整模型与目标域异常打分器随机初始化版本的模型在 4 组典型跨域任务中进行对比实验，记录二者 AUC-ROC 与 AUC-PR 指标随训练迭代次数的变化趋势，如图 3 所示。

实验表明，在大多数任务中，采用源域参数初始化的完整模型在训练初期展现出显著优于或与随机初始化模型持平的性能表现，该模型可以快速且稳定地收敛至较高性能水平；相比之下，消融版本模型不仅整体性能偏低，在训练过程中还出现了显著的性能波动，特别是在对类别不平衡更为敏感的 AUC-PR 指标上，二者性能差距尤为显著。结果证实，源域参数初始化策略能够在训练初期提供高质量语义起点，有助于避免模型陷入局部最优或训练偏差，是实现稳定跨域知识迁移的重要保障。

综上所述，DPA 与 EEM 模块在语义对齐、嵌入分布优化与判别增强方面发挥了重要作用，二者联合使用有助于提升模型在跨域图异常检测任务中的

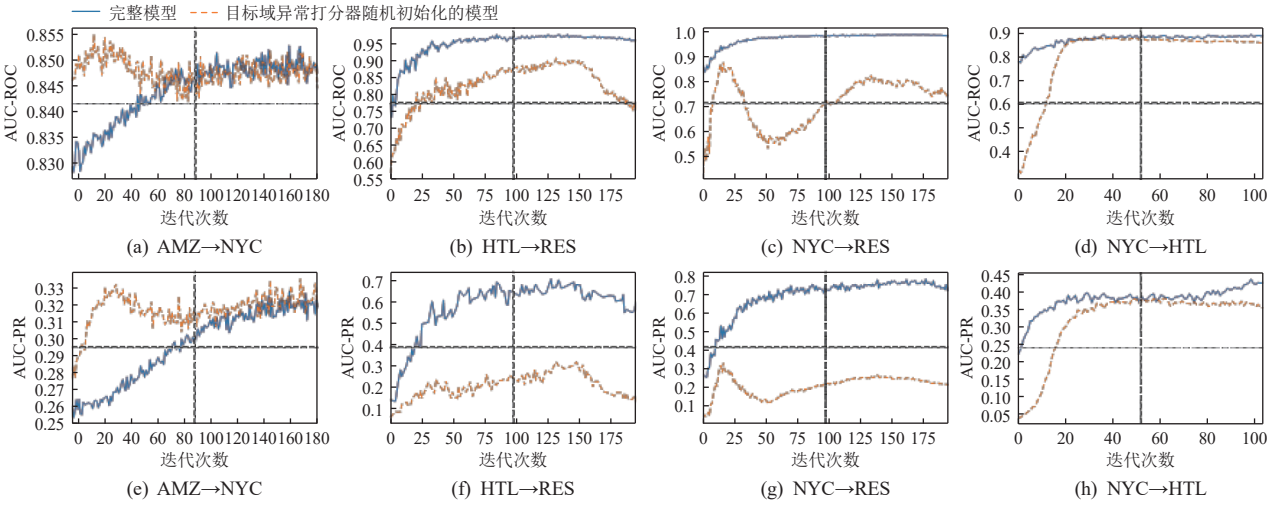


Fig. 3 Ablation study on initialization strategies for the target domain anomaly scorer

图 3 目标域异常打分器初始化策略的消融实验

整体表现。同时,基于源域的参数初始化策略进一步提升了模型的收敛效率与稳定性。上述工作从模型设计有效性和训练过程动态演化2个角度验证了本文方法的实践优越性。

4.6 参数敏感性分析

为进一步验证关键超参数设置的合理性,本文对δ和β分别进行了敏感性分析。

首先,针对式(11)中的超参数,本文在RES→HTL和HTL→NYC这2个跨域任务上,对$\delta\in\{0.0,0.1,0.3,0.5,0.7,0.9,1.0\}$进行了敏感性分析。如图4所示,参数$\delta$的变化会对模型性能产生一定影响。当$\delta$取较小或中等值时,模型通常能够取得较好的检测效果,说明适当引入异常原型项有助于提升跨域对齐质量与异常表征能力。而当δ逐渐增大时,模型性能在部分任务上出现明显下降,表明过强的异常原型约束可能破坏正常与异常语义之间的平衡,从而影响目标域表示学习的稳定性。此外,HTL→NYC任务对δ的变化相对更稳定,而RES→HTL任务的波动更为明显,这说明不同跨域场景下模型对该参数的敏感程度存在一定差异。综合来看,δ在合理区间内能够保持较稳定的性能表现,因此本文将其默认设置为0.3,以兼顾检测效果与参数鲁棒性。

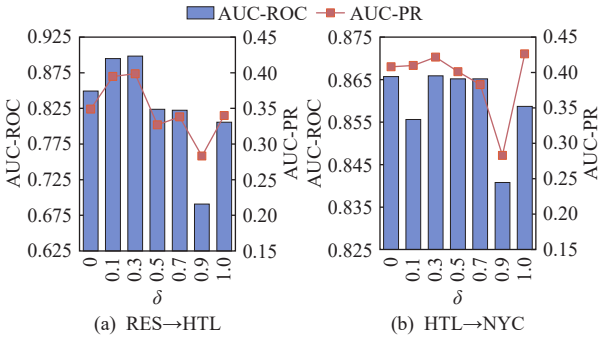


Fig. 4 Sensitivity analysis of parameter δ on representative cross-domain tasks

图4 参数δ在代表性跨域任务上的敏感性分析

其次,在自监督迁移学习阶段,为合理调节源域训练的异常打分器η和孤立森林模型所生成评分在综合伪异常得分中的权重,本文引入了超参数β来控制二者的贡献比例。为系统评估β对模型性能的影响,本研究在RES→HTL的实验配置下,对$\beta\in\{0.0,0.1,0.3,0.5,0.7,0.9\}$进行了敏感性分析实验。

如图5所示,实验结果清晰地展示了随着β值的增加,伪标签生成过程中异常评分分布的演化趋势。当β较小时,模型输出的异常评分呈现出与理想情况相符的右偏长尾分布:大部分样本得分较低,仅少数样本具有显著的高异常评分。这种分布形态有助于提升异常类别的判别能力,同时保留目标域数据的潜在结构特性。然而,随着β值增大,异常评分分布结构发生了显著变化,原有的峰尾特征逐渐弱化,最终演变为更对称且集中的类高斯分布。这种变化偏离了异常检测任务所期望的判别性分布,削弱了模型对异常的识别能力。

进一步地,如图6所示,不同β取值下模型最终

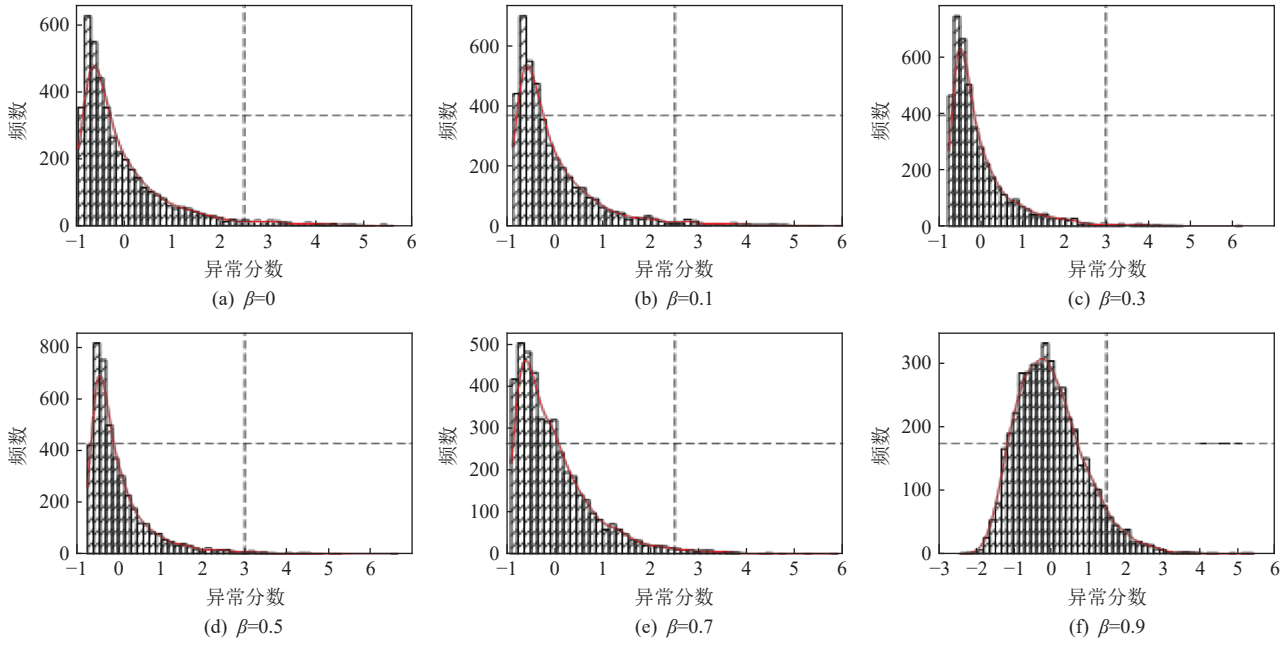


Fig. 5 The impact of different β values on pseudo-label generation

图5 不同β值对伪标签生成的影响

的 AUC-ROC 和 AUC-PR 呈现出先上升后下降的总体优势，即当β取较小或中等值时，模型性能较优，而当β继续增大时，检测性能出现明显退化。该结果与图 5 中的分布变化规律是一致的，说明过强的源域预训练异常打分器输出会破坏目标域异常评分原有的判别结构，从而影响最终的伪标签质量与迁移检测效果。

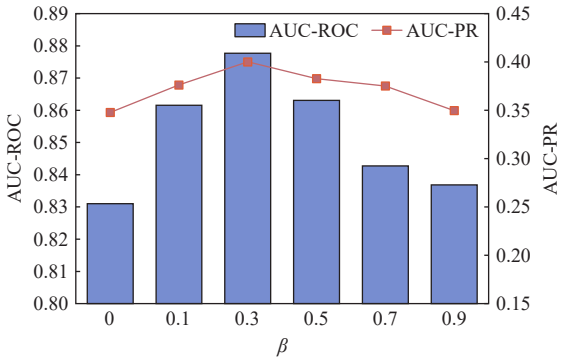


Fig. 6 Sensitivity analysis of model on RES→HTL task with different values of β

图 6 不同 β 取值下模型在 RES→HTL 任务上的敏感性分析

综合实验结果分析，模型在$\beta = 0.3$附近通常取得最优或次优性能。基于此，本文在不使用目标域标签的前提下，根据源域验证表现与训练稳定性选择β值作为最终配置。这一策略在保留目标域数据结构信息的基础上，适度融合源域模型的打分信息，实现了对潜在异常的更稳定识别。实验结果证实，最有效的迁移机制并非单纯依赖源域信息重新定义异常属性，而是在源域模型有效引导下，适度融合目标域自身的无监督异常检测机制，从而在保留目标域分布特征的同时，实现对异常模式的稳健建模。

5 结　论

针对跨域图异常检测中源域与目标域之间存在的图结构异质性与特征分布偏移挑战，本文提出了一种融合动态原型学习与自监督迁移机制的新型检测框架。该框架首先在源域中构建正常与异常的语义原型，通过双向引导策略实现跨域共享语义空间下的高效知识对齐，显著提升跨域建模中的语义一致性。为增强其稳定性与判别能力，引入指数移动平均机制动态更新原型；同时，通过最小化目标域节点嵌入熵，优化目标域的特征空间结构，有效缓解因监督信号缺失带来的决策边界模糊问题。在此基础上，框架设计了融合源域评分与无监督检测的伪标签生成机制，筛选高置信度样本作为弱监督信号以引导目标域模型的迭代优化，最终在弱监督条件下实现了模型识别准确性与鲁棒性的双重提升。

在 8 种典型的跨域图异常检测实验配置下，本文方法在 AUC-ROC 和 AUC-PR 指标上分别取得了 0.909 与 0.526 的平均性能，较当前最先进的基线方法分别提升了 1.9% 与 8.7%，充分证明了该方法的性能优势与迁移稳健性。综合评估结果表明，本文方法在复杂图结构和特征分布偏移场景下，具备更强的泛化能力和更高的判别一致性。

未来，本文将进一步从多源域协同迁移与未知目标域泛化 2 个方向拓展跨域图异常检测模型。一方面，探索多源域原型融合与自适应对齐机制，以提升模型对复杂异构图数据的适应能力；另一方面，将进一步结合零样本图异常检测与通用图异常检测的研究思路，增强模型在未见图上的直接检测能力，从而进一步提升方法的通用性与实用性。

作者贡献声明：胡耀平负责算法设计与实验；彭淑娟负责模型可行性分析；柳欣负责模型优化和算法分析；崔振负责实验的多样性分析；朱磊负责算法优化。

参　考　文　献

[1] Ding Kaize, Shu Kai, Shan Xuan, et al. Cross-domain graph anomaly detection[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(6): 2406−2415

[2] Wang Qizhou, Pang Guansong, Salehi M, et al. Cross-domain graph anomaly detection via anomaly-aware contrastive alignment[C]//Proc of the 37th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2023: 4676−4684

[3] Huang Changqin, Shi Xinxing, Gao Chengling, et al. Cross-domain graph anomaly detection via graph transfer and graph decouple[J]. CAAI Transactions on Intelligence Technology, 2025, 10(4): 1089−1103

[4] Delaram P, Arlei S. Cross-domain graph anomaly detection via test-time training with homophily-guided self-supervision[J/OL]. Transactions on Machine Learning Research, 2025[2026-05-11]. https://openreview.net/forum?id=sB3LqdOlNb

[5] Chen Bofeng, Li Jingdong, Lu Xingjian, et al. Survey of deep learning based graph anomaly detection methods[J]. Journal of Computer Research and Development, 2021, 58(7): 1436−1455 (in Chinese)
(陈波冯, 李靖东, 卢兴见, 等. 基于深度学习的图异常检测技术综述[J]. 计算机研究与发展, 2021, 58(7): 1436−1455)

[6] Hooi B, Song H A, Beutel A, et al. Fraudar: Bounding graph fraud in

the face of camouflage[C]//Proc of the 22nd ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2016: 895−904

[7] Shin K, Hooi B, Faloutsos C. M-zoom: Fast dense-block detection in tensors with quality guarantees[C]//Proc of the Joint European Conf on Machine Learning and Knowledge Discovery in Databases. Berlin: Springer, 2016: 264−280

[8] Kumar S, Hooi B, Makhija D, et al. Rev2: Fraudulent user prediction in rating platforms[C]//Proc of the 17th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2018: 333−341

[9] Li Ronghua, Xu Yu J, Huang Xin, et al. Robust reputation-based ranking on bipartite rating networks[C]//Proc of the 2012 SIAM Int Conf on Data Mining. Philadelphia, PA: SIAM, 2012: 612−623

[10] Wang Guan, Xie Sihong, Liu Bing, et al. Identify online store review spammers via social review graph[J]. ACM Transactions on Intelligent Systems and Technology, 2012, 3(4): 1−21

[11] Zhang Feng, Zhai Jidong, Shen Xipeng, et al. TADOC: Text analytics directly on compression[J]. The International Journal on Very Large Data Bases, 2021, 30(2): 163−188

[12] Akoglu L, Chandy R, Faloutsos C. Opinion fraud detection in online reviews by network effects[C]//Proc of the 17th Int AAAI Conf on Web and Social Media. Palo Alto, CA: AAAI, 2013: 2−11

[13] Rayana S, Akoglu L. Collective opinion spam detection: Bridging review networks and metadata[C]//Proc of the 21st ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2015: 985−994

[14] Ma Shuai, Liu Jianwei, Zuo Xin. Survey on graph neural network[J]. Journal of Computer Research and Development, 2022, 59(1): 47−80 (in Chinese)
(马帅, 刘建伟, 左信. 图神经网络综述[J]. 计算机研究与发展, 2022, 59(1): 47−80)

[15] Ding Kaize, Li Jundong, Bhanushali R, et al. Deep anomaly detection on attributed networks[C]//Proc of the 2019 SIAM Int Conf on Data Mining. Philadelphia, PA: SIAM, 2019: 594−602

[16] Yuan Xu, Zhou Na, Yu Shuo, et al. Higher-order structure based anomaly detection on attributed networks[C]//Proc of the 2021 IEEE Int Conf on Big Data. Piscataway, NJ: IEEE, 2021: 2691−2700

[17] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[J]. arXiv preprint, arXiv: 1609.02907, 2016

[18] Liu Yixin, Li Zhao, Pan Shirui, et al. Anomaly detection on attributed networks via contrastive self-supervised learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(6): 2378−2392

[19] Duan Jingcan, Wang Siwei, Zhang Pei, et al. Graph anomaly detection via multi-scale contrastive learning networks with augmented view[C]//Proc of the 37th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2023: 7459−7467

[20] Yu Wenchao, Cheng Wei, Aggarwal C C, et al. Netwalk: A flexible deep embedding approach for anomaly detection in dynamic networks[C]//Proc of the 24th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2018: 2672−2681

[21] Yang C, Xiao Yuxin, Zhang Yu, et al. Heterogeneous network representation learning: Survey, benchmark, evaluation, and beyond[J]. arXiv preprint, arXiv: 2004.00216, 2020

[22] Velickovic P, Cucurull G, Casanova A, et al. Graph attention networks[J]. arXiv preprint, arXiv: 1710.10903, 2017

[23] Qiao Hezhe, Niu Chaoxi, Chen Ling, et al. AnomalyGFM: Graph foundation model for zero/few-shot anomaly detection[C]//Proc of the 31st ACM SIGKDD Conf on Knowledge Discovery and Data Mining V2. New York: ACM, 2025: 2326−2337

[24] Niu Chaoxi, Qiao Hezhe, Chen Ling, et al. Zero-shot generalist graph anomaly detection with unified neighborhood prompts[C]//Proc of the 34th Int Joint Conf on Artificial Intelligence. New York: ACM, 2025: 3226−3234

[25] Gretton A, Borgwardt K, Rasch M, et al. A kernel method for the two-sample-problem[C]//Proc of the 20th Int Conf on Neural Information Processing Systems. Cambridge, MA: MIT, 2006: 513−520

[26] Long Mingsheng, Zhu Han, Wang Jianmin, et al. Unsupervised domain adaptation with residual transfer networks[C]//Proc of the 30th Int Conf on Neural Information Processing Systems. New York: Curran Associates Inc, 2016: 136−144

[27] Lee C Y, Batra T, Baig M H, et al. Sliced wasserstein discrepancy for unsupervised domain adaptation[C]//Proc of the 2019 IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019: 10277−10287

[28] Shen Jian, Qu Yanru, Zhang Weinan, et al. Wasserstein distance guided representation learning for domain adaptation[C]//Proc of the 32nd AAAI Conf on Artificial Intelligence and 30th Innovative Applications of Artificial Intelligence Conf and 8th AAAI Symp on Educational Advances in Artificial Intelligence. Palo Alto, CA: AAAI, 2018: 4058−4065

[29] Ajakan H, Germain P, Larochelle H, et al. Domain-adversarial neural networks[J]. arXiv preprint, arXiv: 1412.4446, 2014

[30] Wu Man, Pan Shirui, Zhou Chuan, et al. Unsupervised domain adaptive graph convolutional networks[C]//Proc of the Web Conf. New York: ACM, 2020: 1457−1467

[31] Dai Quanyu, Wu Xiaoming, Xiao Jiaren, et al. Graph transfer learning via adversarial domain adaptation with graph convolution[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 35(5): 4908−4922

[32] Wu Man, Pan Shirui, Zhu Xingquan. Attraction and repulsion: Unsupervised domain adaptive graph contrastive learning network[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2022, 6(5): 1079−1091

[33] Yin Nan, Shen Li, Wang Mengzhu, et al. CoCo: A coupled contrastive framework for unsupervised domain adaptive graph classification[C]// Proc of the 40th Int Conf on Machine Learning. Brookline, MA: Microtome Publishing, 2023: 40040−40053

[34] Guo Gaoyang, Wang Chaokun, Yan Bencheng, et al. Learning adaptive node embeddings across graphs[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 35(6): 6028−6042

[35] Shi Boshen, Wang Yongqing, Guo Fangda, et al. Improving graph domain adaptation with network hierarchy[C]//Proc of the 32nd ACM Int Conf on Information and Knowledge Management. New York: ACM, 2023: 2249−2258

[36] Zhuo Junbao, Su Chi, Wang Shuhui, et al. Min-entropy transfer

adversarial hashing[J]. Journal of Computer Research and Development, 2020, 57(4): 888−896 (in Chinese)
(卓君宝, 苏驰, 王树徽, 等. 最小熵迁移对抗散列方法[J]. 计算机研究与发展, 2020, 57(4): 888−896)
[37] Hamilton W, Ying Zhitao, Leskovec J. Inductive representation learning on large graphs[C]//Proc of the 31st Int Conf on Neural Information Processing Systems. New York: Curran Associates Inc, 2017: 1025−1035
[38] Pang Guansong, Ding Choubo, Shen Chuncha, et al. Explainable deep few-shot anomaly detection with deviation networks[J]. arXiv preprint, arXiv: 2108.00462, 2021
[39] Liu F T, Ting Kaiming, Zhou Zhihua. Isolation forest[C]//Proc of the IEEE Int Conf on Data Mining. Piscataway, NJ: IEEE, 2008: 413−422
[40] Pang Guansong, Cao Longbing, Chen Ling, et al. Learning representations of ultrahigh-dimensional data for random distance-based outlier detection[C]//Proc of the 24th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2018: 2041−2050
[41] Kaghazgaran P, Caverlee J, Squicciarini A. Combating crowdsourced review manipulators: A neighborhood-based approach[C]//Proc of the 11th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2018: 306−314
[42] Veličković P, Fedus W, Hamilton W L, et al. Deep graph infomax[C/OL]//Proc of the Int Conf on Learning Representations. 2019[2026-05-11]. https://openreview.net/forum?id=rklz9iAcKQ
[43] Breunig M M, Kriegel H P, Ng R T, et al. LOF: Identifying density-based local outliers[C]//Proc of the 2000 ACM SIGMOD Int Conf on Management of Data. New York: ACM, 2000: 93−104
[44] Chen Zhenxing, Liu Bo, Wang Meiqing, et al. Generative adversarial attributed network anomaly detection[C]//Proc of the 29th ACM Int Conf on Information & Knowledge Management. New York: ACM, 2020: 1989−1992
[45] Peng Zhen, Luo Minna, Li Jundong, et al. ANOMALOUS: A joint modeling approach for anomaly detection on attributed networks[C]//Proc of the 27th Int Joint Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2018: 3513−3519
[46] Bandyopadhyay S, N L, Vivek S V, et al. Outlier resistant unsupervised deep architectures for attributed network embedding[C]//Proc of the 13th Int Conf on Web Search and Data Mining. New York: ACM, 2020: 25−33
[47] Wilson G, Cook D J. A survey of unsupervised deep domain adaptation[J]. ACM Transactions on Intelligent Systems and Technology, 2020, 11(5): 1−46



Hu Yaoping, born in 2001. Master candidate. His main research interest includes graph anomaly detection.
胡耀平, 2001 年生。硕士研究生。主要研究方向为图异常检测。



Peng Shujuan, born in 1982. PhD, professor, PhD supervisor. Member of CCF. Her main research interests include computer vision, pattern recognition, and machine learning.
彭淑娟, 1982 年生。博士, 教授, 博士生导师。CCF 会员。主要研究方向为计算机视觉、模式识别、机器学习。



Liu Xin, born in 1982. PhD, professor, PhD supervisor. Senior member of CCF. His main research interests include information retrieval, pattern recognition, and deep learning.
柳　欣, 1982 年生。博士, 教授, 博士生导师。CCF 高级会员。主要研究方向为信息检索、模式识别、深度学习。



Cui Zhen, born in 1981. PhD, professor, PhD supervisor. Senior member of CCF. His main research interests include image processing, pattern recognition, and machine learning.
崔　振, 1981 年生。博士, 教授, 博士生导师。CCF 高级会员。主要研究方向为图像处理、模式识别、机器学习。



Zhu Lei, born in 1987. PhD, professor, PhD supervisor. Senior member of CCF. His main research interest includes high-efficiency multi-modal large models and spatial intelligence.
朱　磊, 1987 年生。博士, 教授, 博士生导师。CCF 高级会员。主要研究方向为高效能多模态大模型与空间智能。