计算机研究与发展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202440177
61(8): 1968−1981, 2024

基于异常感知的变分图自编码器的图级异常检测算法

林 馥[1] 李明康[1] 罗学雄[2] 张书豪[1] 张 越[3] 王梓桐[3]
[1](武汉大学计算机学院 武汉 430072)
[2](麦考瑞大学计算机学院 澳大利亚 悉尼 2122)
[3](武汉大学网络安全学院 武汉 430072)
(linfu@whu.edu.cn)

Anomaly-Aware Variational Graph Autoencoder Based Graph-Level Anomaly Detection Algorithm

Lin Fu[1], Li Mingkang[1], Luo Xuexiong[2], Zhang Shuhao[1], Zhang Yue[3], and Wang Zitong[3]
[1](*School of Computer Science*, *Wuhan University*, *Wuhan* 430072)
[2](*School of Computing*, *Macquarie University*, *Sydney*, *Australia* 2122)
[3](*School of Cyber Science and Engineering*, *Wuhan University*, *Wuhan* 430072)

Abstract Graph anomaly detection plays a crucial role in identifying abnormal patterns within complex data structures, and finds applications in various domains such as malicious molecule identification, financial fraud detection, and social network analysis. While existing research in graph anomaly detection predominantly focuses on node-level anomaly detection, there is a scarcity of methods specifically designed for graph-level anomaly detection. Moreover, these methods often fail to adequately explore anomalous graph data, are sensitive to anomaly labels, struggle to capture features of anomalous samples effectively, exhibit poor model generalization, and suffer from performance reversal issues. There is a need for improvement in anomaly detection capabilities. we propose an anomaly-aware variational graph autoencoder based graph-level anomaly detection algorithm (VGAE-D), which utilizes an anomaly-aware variational graph autoencoder to simultaneously extract features from normal and anomalous graph data. This algorithm distinguishes the encoding information of normal and anomalous graphs in the encoding space, further exploring the graph encoding information to compute anomaly scores. Experimental evaluations on eight publicly available datasets from diverse domains demonstrate the effectiveness of the proposed graph-level anomaly detection algorithm. The results indicate that VGAE-D can efficiently identify anomalous graphs in different datasets, outperforming mainstream graph-level anomaly detection methods. Additionally, the algorithm exhibits a capability for learning from few anomalous samples, mitigating the issue of performance inversion to a large extent.

Key words graph-level anomaly detection; graph neural networks; variational graph autoencoder; graph representation learning; few-shot learning

摘 要 图异常检测在识别复杂数据结构的异常模式中具有重要作用,被广泛地应用于有害分子识别、金融欺诈检测、社交网络分析等领域.但目前的图异常检测研究大多数聚焦在节点级别的异常检测,针对图级别的异常检测方法仍然较少,且这些方法并不能对异常图数据进行充分挖掘,且对异常标签比较敏感,无法有效地捕捉异常样本的特征,存在模型泛化能力差、性能翻转问题,异常检测能力有待提升.提出了

收稿日期: 2024-03-15; **修回日期**: 2024-05-15
通信作者: 林馥(linfu@whu.edu.cn)

一种基于异常感知的变分图自编码器的图级异常检测算法（anomaly-aware variational graph autoencoder based graph-level anomaly detection algorithm, VGAE-D），利用具有异常感知能力的变分图自编码器提取正常图和异常图数据的特征，并差异化正常图和异常图在编码空间中的编码信息分布，对图编码信息进一步挖掘来计算图的异常得分. 在不同领域的 8 个公开数据集上进行实验，实验结果表明，提出的图级别异常检测方法能有效地对不同数据集中的异常图进行识别，异常检测性能高于目前主流的图级别异常方法，且具有少异常样本学习能力，较大程度上克服了性能翻转问题.

关键词 图级别异常检测；图神经网络；变分图自编码器；图表示学习；少样本学习

中图法分类号 TP183

图作为现实世界中一种广泛存在的数据结构，近些年来越来越受到学者和研究团队的关注. 从微观世界的分子结构到宏观世界的社交网络，都可以被抽象为图结构. 以化学分子为例，整个化学分子可以看作一张图[1]，分子中的原子可以视作节点，原子之间的化学键可以视为连接节点的边. 类似地，如果将一个社交网络视为图，则网络中的用户可以被视为节点[2]，用户之间的社交联系可以被当作边. 通过将真实世界中的物质结构和网络结构建模成图，不仅可以为学术研究提供新的视角，也能为解决现实世界中的实际问题提供新的思路.

除此之外，图的应用还延伸到了城市规划、交通流量优化等领域. 通过将城市的道路、建筑、人流等元素建模成图，研究者们能够更好地理解城市运行的规律，为城市设计和规划提供科学依据. 这种图模型在交通管理中也发挥着关键作用，能帮助优化道路网络、提高交通效率、减少拥堵问题.

由于图在现实世界的应用如此广泛，针对图的异常检测也变得尤为重要. 图异常检测的目标是找出与大多数图不同或不符合正常模式的图，这些异常的图可能代表潜在的问题、异常情况、欺诈行为、错误或其他重要事件[3]. 可以应用于有害分子鉴定、网络欺诈甄别等领域[4-5]. 例如，如果将不同种类的分子视为不同的图，将大多数无毒无害的分子视作正常图，而将少数的有害分子视作异常图，则有害分子鉴定问题可以转化为图异常检测问题. 在训练好图异常检测模型之后，遇到新种类的分子时就可以直接通过图异常检测模型来对该种分子是否有害进行初步的判断.

但是目前针对图的异常检测方法[6-8]大部分集中在节点级别的异常检测，即检测一张图中存在的异常节点，而对图级别异常检测方法的研究仍然比较少. 不同于节点级别异常检测，图级别异常检测旨在发现图集合中结构或者节点属性存在异常的图. 节点级别异常检测中的节点嵌入表示容易通过节点自身的属性及其邻接节点的属性聚合得到，而图级别异常检测的难点则在于如何有效地对整个图的结构信息和图中所有节点的属性信息进行挖掘来得到有效的图嵌入表示，并合理地利用图嵌入表示来评估图的异常程度.

随着深度学习技术的不断发展，许多基于图神经网络（graph neural networks, GNNs）的图级别异常方法被提出，这些方法利用 GNNs 来提取图在潜在空间的嵌入表示并取得了不错的效果，OCGIN[9] 利用 one-class classification 的思想，训练模型使得正常图在潜在空间的特征表示能聚集在样本中心点周围，而异常图则会远离样本中心点. OCGTL[10] 对 OCGIN 进行了改进，相较于 OCGIN 直接使用平均图嵌入表示作为聚类中心的方式，OCGTL 训练了多个 GNN 模型，聚类中心由其中一个 GNN 学习得到，并作为参照图嵌入表示，正常图的图嵌入表示会聚集在参照图嵌入表示周围. GlocalKD[11] 利用知识蒸馏的思想分别训练老师模型与学生模型，并通过比较图在 2 个模型上的重构误差来衡量图的异常程度.

OCGTL, OCGIN, GLocalKD 等都只针对正常图进行了数据挖掘而忽略了图数据集中存在的异常图，没有利用异常图数据中可能蕴含的异常信息. 此外，GLocalKD 和 GOOD-D[12] 都利用图自编码器（graph autoencoder, GAE）[13] 来对图进行编码，简单地将图经过 GAE 重构前后的属性重构误差和结构重构误差作为判断图异常程度的依据，图编码信息仅仅被用来进行图的生成任务，而未用于提取图的正常或异常信息. 这些做法都有可能对图异常检测的效果造成影响. OCGIN 首次对图级别异常检测中的性能翻转问题做了定义，性能翻转是指在真实世界的图数据集中指定不同类别的图作为异常图来训练图级异常检测模型时，模型的性能会出现较大差异. 值得注意的是，目前存在的大多数图级别异常检测方法都出

现了不同程度的性能翻转问题.

为了解决上述问题,本文提出了一种基于异常感知的变分图自编码器的图级别异常检测算法(anomaly-aware variational graph autoencoder based graph-level anomaly detection algorithm, VGAE-D),该算法能有效区分正常图和异常图的分布信息.不同于普通的GAE,本文采用具有异常感知能力的变分图自编码器(variational graph autoencoder, VGAE)[14]对图的嵌入表示进行提取,使用概率分布来建模图的潜在表示,在训练过程中增大异常图的编码信息方差,以便更好地捕捉异常图的多样性和不确定性,同时减小正常图的编码信息方差,以压缩正常图的编码空间.不同于简单依赖图重构误差来判定图正常与否的方法,本文使用异常评估器直接对图编码信息进行了进一步的提取,通过端到端的形式得出图的异常得分.本文的主要贡献有3个方面:

1)提出了一种端到端的图级别异常检测方法,将VGAE应用于图级别异常检测当中,通过改进VGAE并将VGAE的编码信息作为异常评估模块的输入,计算图的异常得分.

2)同时对正常图数据和异常图数据进行挖掘,可视化实验结果表明,本文提出的方法能较为有效地区分正常图和异常图在编码空间中的分布模式,进而有助于对图的编码信息进行更深层次的挖掘.

3)在公开的真实数据集上做了对比实验,实验结果表明,VGAE-D能够有效地对图级别异常进行检测,该方法的异常检测能力相较于当前主流的图异常检测方法有很大提升,较大程度上克服了性能翻转问题.

1 相关工作

1.1 图神经网络

图神经网络是一种流行的深度学习方法,可以用于处理以图结构形式建模表示的不规则数据[15],一些原本被广泛应用于计算机视觉和自然语言处理领域的神经网络模型如循环神经网络和前馈神经网络被迁移到图数据挖掘领域[16-17]用于对图表征进行学习,许多GNN模型被提出,如借鉴了卷积神经网络(convolutional network, CNN)的图卷积神经网络(graph convolutional network, GCN)[18-20]、图同构神经网络(graph isomorphism network, GIN)[21]、图自编码器[22-23]及图注意力网络(graph attention network, GAT)[24-25]等.

GCN借鉴了计算机视觉领域所提出的CNN概念,将卷积运算的操作对象由图像扩展到了图结构,在提取节点本身属性特征的同时还能聚合邻居节点的特征信息.GIN的设计灵感来自于图同构的概念,其核心思想是设计一个可排列不变(permutation-invariant)的网络结构,通过在节点更新的过程中采用池化和聚合的方法使得模型能够在不考虑节点具体排列的情况下学到图表示.

GAE由经典的自编码器(autoencoder, AE)改进迁移而来,能够学习图在低维空间的表示来进行特征提取,并通过对习得的图特征进行解码重构还原出图.VGAE在普通GAE的基础上增加了对图嵌入表示在潜在空间分布的学习能力,在节点链接预测等任务上相较于GAE表现更加优秀[26].GAT使用自注意力机制来学习每个节点与其邻居节点之间的关系,能够在图上进行灵活、自适应的信息聚合.这些网络模型为图特征表示提取任务提供了新的方法.

1.2 图级别异常检测

图级别异常检测是指从大量图结构数据构成的图数据集中识别出异常的图,目前这方面的研究工作仍然不多.大部分针对图的异常检测方法都集中在节点级别的异常检测,即检测单个图中的异常节点,这些主要关注图局部信息的节点级别异常检测方法不能很好地适用于需要关注全局信息的图级别异常检测,因此需要提出更多针对图级别异常检测的方法.早期传统的图级别异常方法[27-29]基于聚类的思想,对图的结构信息或节点信息进行聚类来区分正常图和异常图.另外一些算法如基于度中心性的异常检测、基于介数中心性的异常检测和基于紧密中心性的异常检测,聚焦于检测图的局部异常,通过挖掘图的结构信息来探测图中异常的节点和边,进而判断图的异常程度.后来有学者提出了一些两阶段方法[9, 30],首先经过图内核函数如WL核(weisfeiler-lehman graph kernel)[31]、传播核(propagation kernel, PK)[32]和Graph2Vec[33]等得到图的嵌入表示,之后应用传统的异常检测算法如孤立森林(isolation forest, IF)[34]算法、局部离群因子(local outlier factor, LOF)[35]算法和单类支持向量机(one-class support vector machine, OCSVM)[36]算法等进行异常检测.

上述基于浅层学习的图级别异常检测算法的缺点在于提取图嵌入表示的能力有限,仅依靠图的结构信息或节点属性信息进行聚类可能难以充分捕捉图的全局特征,并且无法定量评估图的异常程度.因此有学者提出可以利用GNNs,通过深度学习的方式对图嵌入表示进行学习提取[37-40].例如,GLAM[41]通

过 GIN 获取图的节点级别嵌入表示之后，利用 MMD 池化（matrix map decomposition pooling, MMD-pooling）进行异常检测；GOOD-D[12] 从节点属性和图结构 2 个视角进行对比学习和异常检测；GLADC[42] 借鉴对比学习的思想，对图编码信息进行增强，通过学习更全面的图嵌入表示增加模型对正常图和异常图的区分能力；GLADST[43] 基于知识蒸馏的思想，通过训练不同的学生模型对图的正常和异常程度进行感知.

尽管这些方法相较基于浅层学习的异常检测方法效果更好且更具普适性，但这些方法仅从节点属性和图结构的角度对图嵌入表示进行提取，并未关注图的分布信息，因此得到的图嵌入表示的语义信息仍不够丰富. 同时，主流的基于 GNNs 的图级别异常检测方法大都利用 GAE 来学习正常图在潜在空间的嵌入表示，在训练时只利用图数据集中的正常图，将正常图的重构损失作为其损失函数的重要组成部分，并未深入挖掘数据集中的异常图. 虽然这些方法中的 GAE 能够很好地重构正常图，但是它们往往忽略了异常图中蕴含的信息，因此都存在性能翻转的问题. 一些有监督的图异常检测方法如 OCGIN[9] 和 OCGTL[10]，虽然一定程度上克服了性能翻转的问题，但异常检测能力仍然有待提升.

2 问题定义

本节将给出图以及异常图的定义，并介绍图异常检测任务的目标.

2.1 图的定义

图G可以表示为$G=(\mathcal{V}_G,\mathcal{E}_G)$，其中$\mathcal{V}_G$是节点的集合，$\mathcal{E}_G$是边的集合. 假设$G$中有 n 个节点，使用邻接矩阵$\boldsymbol{A}\in\mathbb{R}^{n\times n}$来表示$G$的结构信息，当$(v_i,v_j)\in\mathcal{E}_G$ 时$A_{ij}=0$，否则，$A_{ij}=1$. 如果G是属性图，那么G中的每个节点$v_i\in\mathcal{E}_G$都对应着一个属性向量$\boldsymbol{x}_i\in\mathbb{R}^d$，每个属性图$G$对应一个属性矩阵$\boldsymbol{X}\in\mathbb{R}^{n\times d}$. 对于非属性图，将图中每个节点的度作为该节点的属性.

2.2 异常图的定义

异常图是指与数据集中其他大部分图明显不同的图，异常可能表现为拓扑结构异常，即图中节点之间存在不正常的连接关系. 属性信息异常，即图中某些节点的属性特征与数据集中大部分节点的特征差异较大；图全局异常，即图的整体特征与数据集中其他图的整体特征相比存在不正常的差异.

2.3 图级别异常检测任务

给定一个图的集合$\mathcal{G}=\{G_1,G_2,\cdots,G_m\}$，图级别异常检测任务的目标是识别上述不同异常模式的异常图，本文提出的图级别异常检测模型通过训练得到一个参数为θ的异常得分函数$f(G;\theta)$，使得对于G_i和G_j，当G_i较G_j更加符合$\mathcal{G}$中大部分图的模式时，有$f(G_i;\theta)<f(G_j;\theta)$，即可以通过函数$f(G;\theta)$来衡量$G$的异常程度. 具体而言，对于图$G$来说，如果$G$是正常图，应有$f(G;\theta)$趋近于 0；反之，如果$G$是异常图，则应有$f(G;\theta)$趋近于 1. 值得注意的是，对图中拓扑结构异常和属性信息异常的探测与节点级异常检测并不一样，前者是在不同的图之间对边或者节点进行异常检测，而后者是指在同一个图中识别出异常的边或节点.

3 VGAE-D 的实现

3.1 框架整体概述

为了能充分利用异常图数据中可能蕴含的信息，本文提出了一种端到端的图异常分数评估模型，利用具有异常感知能力的 VGAE 对图进行编码和重构，控制正常图和异常图编码信息的发散程度，并深层次挖掘图的编码信息来进行异常评估. 最终的实验结果表明本文提出的图异常检测方法可以高效准确地对异常图进行检测，并且较大程度上克服了性能翻转问题.

如图 1 所示，本文所提出的 VGAE-D 模型由 2 关键模块组成，分别是异常感知的 VGAE 模块和异常评估模块. 这 2 个模块协同工作，以实现对图数据中异常情况的准确检测和评估.

第 1 个模块是异常感知的 VGAE 模块，其主要任务是通过 GCN 编码从输入的图数据中提取图的结构和属性特征. 在这一过程中，借助高斯采样机制，模型能够在学习过程中生成更为平滑的编码信息. 同时对正常图和异常图的编码信息方差进行差异化处理，使得正常图的编码信息更加集中，而异常图的编码信息则更加分散. 改进后的 VGAE 具有感知正常图和异常图不同分布模式的能力. 第 2 个模块是异常评估模块，其任务是将第 1 个模块获取的图结构信息、图属性信息以及图编码分布信息映射为图的异常得分.

3.2 异常感知的 VGAE 模块

给定一个图数据集$\mathcal{G}$作为输入，编码器会把图映射到低维潜在空间，而解码器则尝试通过图在潜在空间的表示重构出原始图，训练 VGAE 的过程可以表示为最小化损失函数的过程：

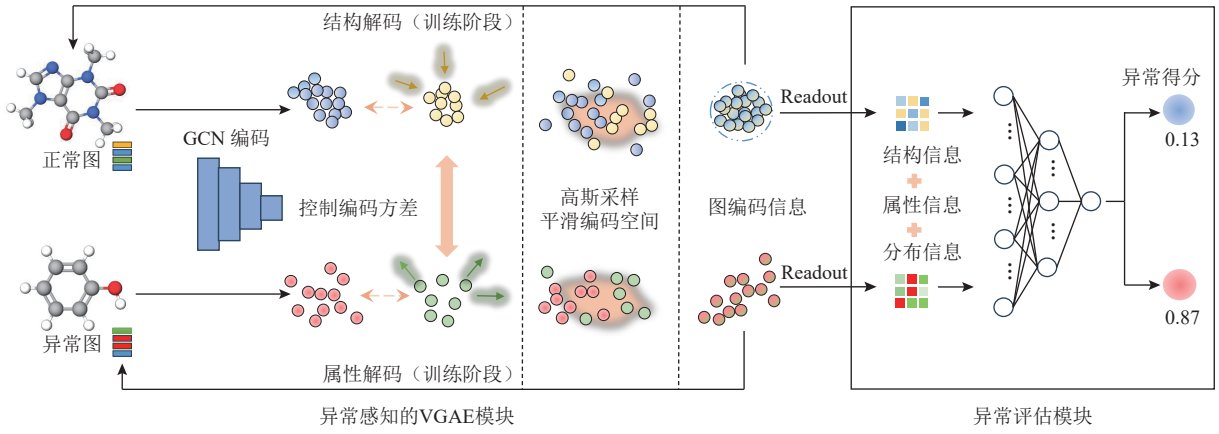


Fig. 1 Framework diagram of VGAE-D

图 1 VGAE-D 框架图

$$\frac{1}{|\mathcal{G}|}\sum_{G\in\mathcal{G}} dist(G, De(En(G))), \tag{1}$$

其中，$dist(\cdot,\cdot)$是预定义的距离函数，这里选用L_2范数距离函数，$En(\cdot)$代表编码器而$De(\cdot)$代表解码器.

本文利用 GCN 来提取低维潜在空间中的节点嵌入表示，GCN 可以将节点及其邻居节点的属性信息和结构信息进行聚合. GCN 的卷积过程可以表示为：

$$H^{l+1} = f(H^l, \boldsymbol{A}|\boldsymbol{W}^l), \tag{2}$$

其中H^l是第 l 层卷积的输入，H^{l+1}是输出的节点嵌入. $\boldsymbol{W}^l$是需要训练的权重矩阵，采用属性图G的属性矩阵$\boldsymbol{X}$为H^0作为第 1 层卷积的输入. 如果图G没有节点属性，那么将节点的度数作为节点的属性，单层 GCN 卷积函数$f(H^l, \boldsymbol{A}|\boldsymbol{W}^l)$的定义为：

$$f(H^l, \boldsymbol{A}|\boldsymbol{W}^l) = \delta(\bar{\boldsymbol{D}}^{-\frac{1}{2}}\bar{\boldsymbol{A}}\bar{\boldsymbol{D}}^{-\frac{1}{2}}H^l\boldsymbol{W}^l), \tag{3}$$

其中$\bar{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{I}$，$\bar{\boldsymbol{D}}$是$G$的度矩阵，为对角矩阵，定义为$\bar{D}_{ii} = \sum_j A_{ij}$，$\delta(\cdot)$是非线性激活函数.

3.2.1 图编码器

使用 GCN 来分别提取节点嵌入的均值和方差并在高斯分布上进行采样以获取最终在潜在空间的节点嵌入表示. 图编码器对图的编码流程可以形式化为：

$$\boldsymbol{H}^l = GCN(\boldsymbol{X}, \boldsymbol{A}), \tag{4}$$

$$\boldsymbol{\mu} = f_{\mathrm{ReLU}}(\boldsymbol{H}^l, \boldsymbol{A}|\widehat{\boldsymbol{W}}), \tag{5}$$

$$\log(\boldsymbol{\sigma}) = f_{\mathrm{ReLU}}(\boldsymbol{H}^l, \boldsymbol{A}|\breve{\boldsymbol{W}}), \tag{6}$$

$$Z = \boldsymbol{\mu} + \beta\cdot\boldsymbol{\sigma}, \text{ 且}\beta \sim \mathcal{N}(0,1), \tag{7}$$

这里，$\boldsymbol{X}$是图G的属性矩阵，$\boldsymbol{A}$是图G的邻接矩阵，$\boldsymbol{H}^l \in \mathbb{R}^{n\times h_1}$是经过 GCN 网络 l 次卷积之后的节点嵌入表示，h_1是卷积之后节点的特征维度，$\boldsymbol{\mu} \in \mathbb{R}^{n\times h_2}$和$\boldsymbol{\sigma} \in \mathbb{R}^{n\times h_2}$是从$\boldsymbol{H}^l$中学习到的图嵌入表示的均值和方差，$\widehat{\boldsymbol{W}}$和$\breve{\boldsymbol{W}}$是需要训练的权重矩阵，$\beta$是从标准高斯分布随机采样得到的值，$\boldsymbol{Z} \in \mathbb{R}^{n\times h_2}$是编码器最终采样得到的编码信息.

3.2.2 结构重构解码器

为了重构图的结构信息，结构重构解码器利用图编码器得到的节点嵌入表示$\boldsymbol{Z}$对图的邻接矩阵进行还原得到重构的邻接矩阵$\widehat{\boldsymbol{A}}$，计算过程为：

$$\widehat{\boldsymbol{A}} = Sigmoid(\boldsymbol{Z}\cdot\boldsymbol{Z}^{\mathrm{T}}). \tag{8}$$

值得注意的是，这里使用内积运算的结果作为 2 个节点之间有连接关系的概率，即：

$$p\left(\widehat{A}_{ij} = 1|\boldsymbol{z}_i\cdot\boldsymbol{z}_j\right) = Sigmoid(\boldsymbol{z}_i\cdot\boldsymbol{z}_j^{\mathrm{T}}). \tag{9}$$

3.2.3 属性重构解码器

属性重构解码器的任务是从潜在空间中的节点嵌入表示逆向还原出图的节点属性特征，以节点的嵌入表示$\boldsymbol{Z}$作为输入，经过 GCN 计算得到重构的属性矩阵$\widehat{\boldsymbol{X}}$，计算过程为：

$$\widehat{X} = f_{\mathrm{ReLU}}(\boldsymbol{Z}, \boldsymbol{A}|\boldsymbol{W}), \tag{10}$$

这里$\boldsymbol{W}$是需要学习的权重矩阵.

3.3 异常评估模块

在获得了图在潜在空间的编码信息$\boldsymbol{Z}$之后，通过异常评估模块来评估图中异常的可能性. $\boldsymbol{Z}$包含了图的结构信息和节点属性信息，这里首先通过$Readout(\cdot)$函数从$\boldsymbol{Z}$中得到图的图级表示$\boldsymbol{H}^G$，再由$\boldsymbol{H}^G$得到图的异常得分S. 具体而言，异常得分S的计算过程为：

$$\boldsymbol{H}^G = Readout(\boldsymbol{Z}), \tag{11}$$

$$S = Sigmoid(\boldsymbol{H}^G\cdot\boldsymbol{W}^G), \tag{12}$$

这里的$Readout(\cdot)$函数选用$cat(\cdot)$拼接函数，即将图中各节点的潜在表示$\boldsymbol{z}_i$拼接为一个整体的向量作为整个图G的图级表示$\boldsymbol{H}^G$，之后再由$\boldsymbol{H}^G$得到其异常得分.

3.4 损失函数

VGAE-D 的损失函数由 3 部分构成：变分图自编

码器损失、方差平衡编码损失以及异常得分的评估损失.

对于 VGAE 来说，其损失包括属性重构损失、结构重构损失和 KL 散度损失（Kullback-Leibler divergence loss）.

属性重构损失和结构重构损失用于训练编码器，使其能够更准确地从图中提取出节点的属性特征和图的结构特征. KL 散度损失用于约束编码信息 $\boldsymbol{Z}$ 的分布，使其符合预设的先验分布，以确保潜在空间的平滑性，这里设定先验分布为高斯分布.

用欧式距离计算重构前后节点属性的差异作为属性重构损失 L_{attr}，计算过程为：

$$L_{\text{attr}}=\sqrt{\sum_{i=1}^{n}(\boldsymbol{x}_i-\hat{\boldsymbol{x}}_i)^2}, \tag{13}$$

其中 $\boldsymbol{x}_i$ 和 $\hat{\boldsymbol{x}}_i$ 分别代表原始的节点特征向量和重构后的节点特征向量.

交叉熵损失函数是一种用于测量 2 个概率分布之间差异的常用损失函数，定义如式（14）所示，其中 Y 是模型对样本的输出的集合，X 是样本真实标签的集合. 模型还原出的图邻接矩阵中的元素是 2 节点之间有连接关系的概率，因此图结构的重构损失 L_{struct} 利用交叉熵损失函数定义如式（15）所示.

$$F(X,Y)=-(Y\cdot\log(X)+(1-Y)\cdot\log(1-X)), \tag{14}$$

$$L_{\text{struct}}=F(\hat{\boldsymbol{A}},\boldsymbol{A}). \tag{15}$$

预设的图编码信息的先验分布为高斯分布，因此 KL 散度损失 L_{KL} 的计算方法如式（16）所示，其中 $\boldsymbol{\mu}$ 和 $\boldsymbol{\sigma}$ 是图编码信息的均值和方差：

$$L_{\text{KL}}=-\frac{1}{2}(1+2\log(\boldsymbol{\sigma})-\boldsymbol{\mu}^2-\boldsymbol{\sigma}^2). \tag{16}$$

整个 VGAE 模块的损失 L_{VGAE} 是图结构重构损失. 图属性重构损失与散度损失之和，可以表示为：

$$L_{\text{VGAE}}=L_{\text{attr}}+L_{\text{struct}}+L_{\text{KL}}. \tag{17}$$

方差平衡编码损失 L_{std} 是对正常图和异常图编码信息之间方差的调整，使得正常图的编码信息更加集中，而异常图的编码信息更加分散. 方差平衡编码损失的定义如式（18）所示，其中 ϕ 是图样本方差的集合，T 是对图样本异常得分的标注，当图 G_i 是正常图时有 T_i=0 否则 T_i=1：

$$L_{\text{std}}=F(Sigmoid(\phi),T). \tag{18}$$

这样可以在训练过程中增大异常图的编码方差和减少正常图的编码方差，进而差异化正常图和异常图的编码信息. 同样使用交叉熵损失函数来训练异常评估模块，使得其对正常图的评分趋向于 0 而对异常图的评分能够趋近于 1. 异常得分 L_{score} 的评估损失计算为：

$$L_{\text{score}}=F(S,T). \tag{19}$$

VGAE-D 模型总的损失函数 $Loss$ 是上述 3 部分损失函数之和，可以表示为：

$$Loss=\alpha L_{\text{VGAE}}+L_{\text{score}}+L_{\text{std}}. \tag{20}$$

L_{VGAE} 由图的结构重构损失、属性重构损失及 KL 散度损失构成，其大小主要受数据集中图的平均节点个数和稠密程度的影响，图的平均节点个数和稠密程度越大，初始训练时的 L_{VGAE} 可能就会越大. L_{score} 和 L_{std} 由叉熵损失函数计算得出，为了控制 L_{VGAE} 的大小与它们在相同数量级，利于模型收敛，这里引入超参数 α 以缩放 L_{VGAE} 的取值范围.

4 实　　验

4.1 数据集

为了更加准确地对 VGAE-D 进行评估，本文选用了真实世界中不同领域的 8 个数据集[①]进行实验.

分子数据集 AIDS, BZR, COX2, DHFR, MUTAG 和 NCI1 是在药物发现和分子活性预测领域常用的数据集. 蛋白质数据集 ENZYMES 和 PROTEINS 是蛋白质结构分类和预测任务常用的数据集. 各数据集的信息如表 1 所示.

Table 1　Information of Datasets

表 1　数据集信息

数据集	图数	平均节点数	平均边数
AIDS	2 000	15.69	16.20
BZR	405	35.75	38.36
COX2	467	41.22	43.45
DHFR	756	42.43	44.54
ENZYMES	600	32.63	62.14
NCI1	4 110	29.87	32.30
PROTEINS	1 113	39.06	72.82
MUTAG	188	17.93	19.79

1）AIDS 数据集中的图来源于 AIDS 抗病毒筛选数据库，该数据库包含 2 个类别，活性和非活性，分别代表能对抗艾滋病毒的分子和不能对抗艾滋病毒的分子. 这些分子被转化为图，其中原子表示为节点，

① 本文实验所用数据集均来自于：https://chrsmrrs.github.io/datasets/

而共价键表示为节点之间的边. 节点属性信息中含有相应原子的信息. 整个数据集包含有 1 600 个非活性样本和 400 个活性样本.

2)BZR 数据集汇集了一组包含 405 种生物分子的受体配体, 包括活性和非活性 2 种配体, 数据集中的每个图代表 1 种配体, 这些数据主要来源于 Haefely[44] 和 Huang 等人[45] 的工作. 他们通过测量配体对地西泮分子与受体结合的抑制效果来衡量配体的体外结合亲和力并以 pIC_{50} 计量, 计量范围从 0.34 nmol~70 μmol. 该数据集综合考虑 2 类配体数量的平衡, 并选定 pIC_{50} = 7.0作为配体是否具有活性的阈值.

3)COX2 数据集汇集了 467 种环氧合酶-2(COX-2) 抑制剂, 该数据集将每种抑制剂视作图, 将抑制剂分成活性和非活性 2 类. 这些抑制剂在体外对人类重组酶活性的抑制效果同样以 pIC_{50} 计量, pIC_{50} 值范围从 1 nmol~100 μmol 以上. 以 pIC_{50} = 6.5为阈值将抑制剂分类为活性(365 种)和非活性(102 种).

4)DHFR 数据集汇集了 756 种二氢叶酸还原酶(DHFR)抑制剂. 数据集中的数据可以用于研究和评估 DHFR 抑制剂的活性和效力, 这在药物研发和药理学研究中可能具有重要意义. 数据集中的每个图都代表了一种抑制剂, 有活性(461 种)和非活性(295 种)2 类.

5)ENZYMES 数据集中, 每个蛋白酶被描述为一个图, 其中图的节点代表蛋白质中的氨基酸残基(或称为二级结构元素), 而边表示这些残基之间的接近关系. 图的标签共有 7 类, 表示 7 种不同的蛋白酶.

6)NCI1 数据集是能够抑制肺癌细胞的化合物的合集, 数据集中的每个图都代表一种化合物, 根据化合物对肺癌细胞的抑制效果将其分为活性(2 057 种)和非活性(2 053 种)2 类.

7)PROTEINS 数据集中, 每个蛋白质结构被表示为一个图. 在这个图中, 节点代表蛋白质中的氨基酸残基, 边表示这些残基之间的相互作用关系, 例如氢键和相对位置等. 每个蛋白质结构都有一个标签, 代表了其属于的蛋白质家族, 数据集中的图共有 2 类, 分属 2 种不同的蛋白质家族.

8)MUTAG 数据集中, 每种化合物被表示为一个图, 整个数据集由 188 种化合物组成, 根据化合物对细菌的诱变作用被分为 2 类.

4.2 评估策略

在图异常检测问题中, 评估模型性能的主要标准是 ROC-AUC(receiver operating characteristic-area under the curve, 以下简称 AUC). AUC 是一种广泛使用的指标, 用于评估模型在不同异常分类阈值下的性能. 在图级别异常检测任务中, 通常正常图样本数量远多于异常图数量, AUC 不受类别不平衡的影响, 更适合在这种情况下评估模型性能.

AUC 是以真阳率(true positive rate, TPR) 为纵轴和假阳率(false positive rate, FPR)为横轴绘制的 ROC 曲线下的面积. 当 AUC 接近于 1 时, 表示模型在不同阈值下能够高度区分正常样本和异常样本, 保持较高的性能. 相反, 当 AUC 接近于 0.5 或更低时, 表示模型的性能与随机猜测相当.

本文选择 AUC 作为主要的评估标准, 以评估提出的图级别异常检测方法在各数据集和不同实验设置下的性能.

4.3 基线方法

为了验证所提方法的有效性, 本文在 8 个数据集上利用 AUC 作为评价指标, 将提出的 VGAE-D 模型与其他模型进行对比. 为保证实验结果的普遍性与稳定性, 本文设定每个模型在实验过程中都使用相同的随机数种子. 并且, 每个模型都存在 2 种类型的异常检测, 即认定真实数据集中标签是 0 为异常的检测和认定标签是 1 为异常的检测. 接下来介绍参与实验的其他模型:

1)两阶段模型. Graph2Vec-IF, Graph2Vec-LOF 和 Graph2Vec-OCSVM 都属于此类. 第 1 个阶段使用无监督学习算法 Graph2Vec[33] 捕捉图的属性特征和结构特征, Graph2Vec 通过将子图视为“单词”, 将图视为“文章”, 利用 Word2Vec 生成图级嵌入向量. 第 2 阶段利用传统异常检测器在嵌入向量空间中检测出异常值, 实验选取 IF, LOF 和 OCSVM 作为异常检测方法.

IF[34] 是一种基于树的集成方法, 用于识别异常值. 它通过构建随机树并测量样本的路径长度来确定样本的孤立程度. 较短的路径长度表明样本更容易被孤立, 被认为是异常的概率更高.

LOF[35] 是一种基于局部邻域密度的异常检测方法. 它测量每个样本相对于其邻域的密度偏差, 密度较低的样本被认为是异常值. LOF 考虑了样本相对于整体数据集的局部行为, 能够捕捉数据分布中的局部不规则性.

OCSVM[36] 是支持向量机的变种, 专门设计用于单类别问题, 即仅有正例没有负例的情况. 它通过学习正例的边界来识别异常值, 将那些远离正例区域的样本视为异常.

2)端到端模型. 本文选取了比较具有代表性的利用深度学习来进行端到端图级别异常检测的 4 种

模型 OCGIN, OCGTL, GLocalKD 以及 GOOD-D 来作为对比方法.

OCGIN[9] 利用 GIN 学习得到节点的嵌入表示后，通过加和聚合的方式从节点嵌入表示中得到图级别嵌入表示，并借鉴 DeepSVDD[46] 来设计其损失函数，最终将在潜在空间中图嵌入表示与图嵌入表示中心点的距离作为图的异常得分.

OCGTL[10] 对 OCGIN 做了改进，同样基于 one-classification 的思想，但是选取的聚类参照中心不再是所有训练样本的中心点，而是也由模型学习得到，同时不同于只用 1 个 GIN 得到图嵌入表示的做法，OCGTL 使用 1 个 GNN 集合得到多个不同的图嵌入表示，使得正常图经过这些网络得到的图嵌入表示能够聚集在参照中心周围，并使用这些图嵌入表示到参照中心点的距离来衡量图异常的程度.

GlocalKD[11] 利用知识蒸馏的思想，分别使用预测和随机目标模型同时对图和节点的嵌入表示进行学习，让正常图通过预测模型得到的图级和节点级别的嵌入表示能贴近随机目标模型，并利用 2 个模型学习得到的图嵌入表来衡量图的异常程度.

GOOD-D[12] 假设对于一个图数据集，其中的正常图的嵌入表示应当落在某个群落之内，而异常图的嵌入表示则相对孤立，并采用对抗学习的思想，从特征视图和结构视图对图级嵌入表示做提取并最终得到整体的图级嵌入表示，对于正常图来说，其最终的图嵌入表示应该聚集在某个图群落之内，而异常图的嵌入表示则距离群落较远.

4.4 实验参数设置

本文在实验中设定 VGAE-D 模型的学习率为 0.01，训练次数为 100 轮次，批大小为 64，网络中 VGAE 的卷积层数设置为 2，且第 1 层卷积输出的维度 h_1 设置为 128，第 2 层卷积输出的维度 h_2 设置为 64，超参数 α 设置为 0.05 并设定随机数种子为 1，使用 K-Flod 交叉验证进行实验，交叉验证次数为 5. 数据集中 80% 的数据作为训练集，其余数据作为测试集，最终的实验结果取 5 次交叉验证中在测试集上计算得到的 AUC 值的平均值.

4.5 对比实验结果

如表 2 和表 3 所示，VGAE-D 异常检测模型在异常标签为 0 和为 1 时 AUC 指标表现几乎都优于基线方法，除了异常标签为 0 时在 PROTEINS 数据集上的表现略低于 OCGIN，异常标签为 1 时在 NCI1 数据集上的表现低于 OCGIN，在 MUTAG 数据集上的表现低于 OCGTL，其他情况都是 VGAE-D 表现最佳. 端到端模型在各数据集上的 AUC 曲线如图 2 和图 3 所示，可以直观地看出在大部分数据集上 VGAE-D 的 AUC 曲线面积都更大，即异常检测性能更好. 这是因为 VGAE 提取得到的图编码信息除包含有图的结构信息和属性信息之外还包含有分布信息，更加全面，后续对图编码信息中的异常进行挖掘时会更加准确.

OCGIN 和 OCGTL 在进行图异常检测时，都基于一个先验假设，即正常图的图嵌入表示是相似的且占多数，异常图的图嵌入表示是各有差异的. 异常标签为 1 时，NCI1 数据集中的数据较为符合这种假设，因此两者的异常检测性能较好，但当异常标签为 0 时，前述先验假设不再适用，因此两者的异常检测性能均出现了大幅度下降，类似的情况也可以在 PROTEINS 和 MUTAG 数据集上观测到. VGAE-D 没有使用这个先验假设知识，虽然异常标签为 0 时在 PROTEINS 数据集上的异常检测性能略低于 OCGIN，异常标签为 1 时在 NCI1 数据集上的异常检测性能低于 OCGIN，

Table 2 Comparison of AUC of Each Model when the Anomaly Label is 0

表 2 异常标签为 0 时各模型的 AUC 对比 %

数据集	模型							
	Graph2Vec-IF	Graph2Vec-LOF	Graph2Vec-OCSVM	OCGIN	GLocalKD	GOOD-D	OCGTL	VGAE-D（本文）
AIDS	91.47	91.89	91.26	94.70	90.66	93.22	95.85	**99.82**
BZR	56.25	55.01	59.57	81.94	61.80	76.14	78.81	**85.51**
COX2	56.22	55.70	50.17	60.81	54.48	56.19	60.00	**82.61**
DHFR	56.30	57.53	55.92	63.46	65.18	60.63	52.82	**83.24**
ENZYMES	57.72	57.72	56.29	52.80	54.88	64.13	49.00	**74.70**
NCI1	34.51	60.67	34.67	53.83	68.32	61.56	61.78	**69.35**
PROTEINS	55.64	63.57	57.15	**74.22**	68.93	63.35	65.4	70.86
MUTAG	58.53	58.61	55.99	65.38	88.46	77.20	82.05	**90.38**

注：黑体数值表示最优值.

Table 3 Comparison of AUC of Each Model when the Anomaly Label is 1

表 3 异常标签为 1 时各模型的 AUC 对比 %

数据集	模型							
	Graph2Vec-IF	Graph2Vec-LOF	Graph2Vec-OCSVM	OCGIN	GLocalKD	GOOD-D	OCGTL	VGAE-D（本文）
AIDS	3.22	42.51	2.98	98.06	95.76	13.50	97.57	**99.73**
BZR	34.34	53.33	27.69	73.95	66.50	33.40	62.15	**85.59**
COX2	39.00	56.95	40.84	51.08	61.90	35.56	50.27	**79.76**
DHFR	46.84	56.32	35.76	60.96	55.64	31.17	60.38	**81.92**
ENZYMES	33.61	34.90	38.20	64.32	48.62	50.29	62.60	**68.17**
NCI1	63.39	58.55	63.52	**72.72**	31.77	32.87	71.77	69.68
PROTEINS	44.08	40.28	44.44	56.31	54.26	52.41	64.11	**73.72**
MUTAG	56.67	57.44	51.61	92.78	87.79	16.81	**94.87**	91.45

注：黑体数值表示最优值.

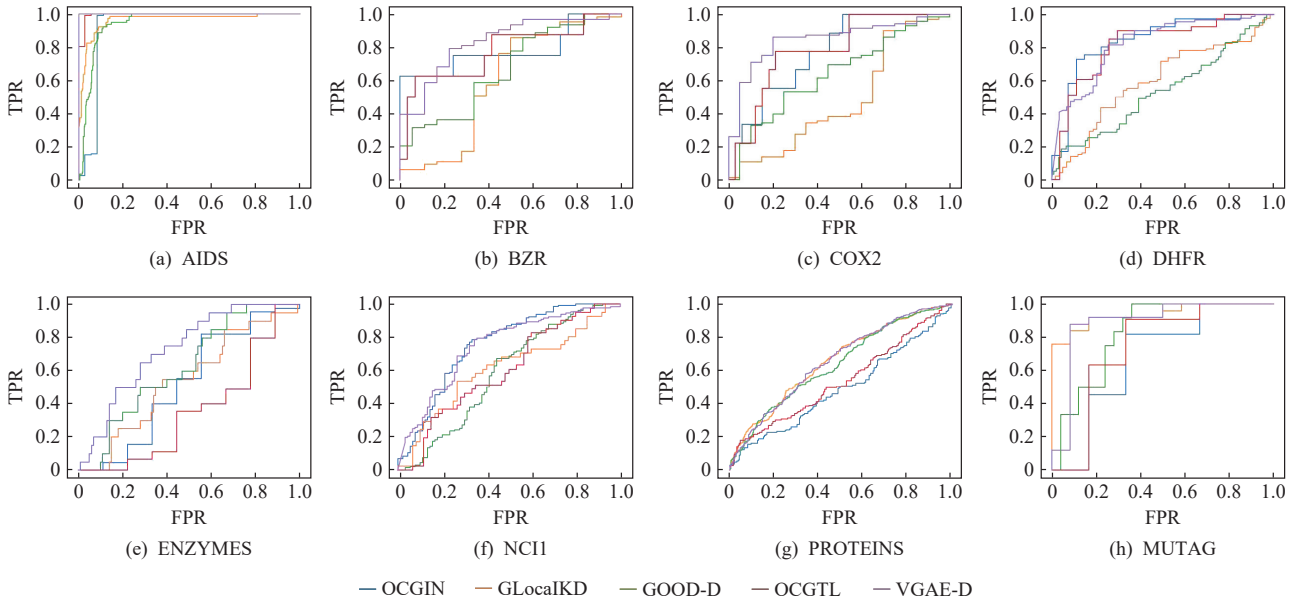


Fig. 2 The AUC curve of the end-to-end model on each dataset when the anomaly label is 0

图 2 异常标签为 0 时端到端模型在各数据集上的 AUC 曲线

在 MUTAG 数据集上的表现低于 OCGTL, 但仍有较强的异常检测能力, 且不易受异常标签变化的影响. 当异常标签发生变化时, VGAE-D 的异常检测性能明显优于 OCGIN 和 OCGTL.

同时可以发现, 对比实验中的两阶段方法及 GOOD-D 模型在大部分数据集上都存在严重的性能翻转问题, 在异常标签为 0 时的异常检测性能较异常标签为 1 时更好, 这是因为在大部分数据集中标签为 0 的图样本数量要远大于标签为 1 的图样本数量, 而这些方法在样本数量较少时并不能有效地提取到图的特征信息. 同样地, OCGIN, GLocalKD 和 OCGTL 在部分数据集上也存在这种情况, 三者在不同数据集上由异常标签变化导致的最大性能差异分别为 27.40%, 36.55% 和 16.66%, 而 VGAE-D 仅为 6.53%. VGAE-D 在异常标签为 0 和 1 时都表现出良好的异常检测能力, 这说明其对异常标签并不敏感, 在异常图样本较少的情况下也能准确地捕捉图编码信息中的异常, 相较于对比方法有更好的泛化能力.

4.6 少异常样本学习实验

为了探究 VGAE-D 在少量异常训练样本下的学习能力, 本文分别将训练数据中的异常图样本数量减少到原来的 10%, 20%, 50%, 即分别将训练用异常图样本比率设置为 r=0.1, 0.2, 0.5, 并保持各端到端模型训练用异常样本数量为原来的 100% 不变进行对比实验. 实验结果如表 4 和表 5 所示, 可以发现 VGAE-D 在训练用异常图样本数量减少到 50% 时, 其异常检测性能就已经与其他方法持平或者比其他方法更好. 这就表明 VGAE-D 的收敛速度较快, 在训

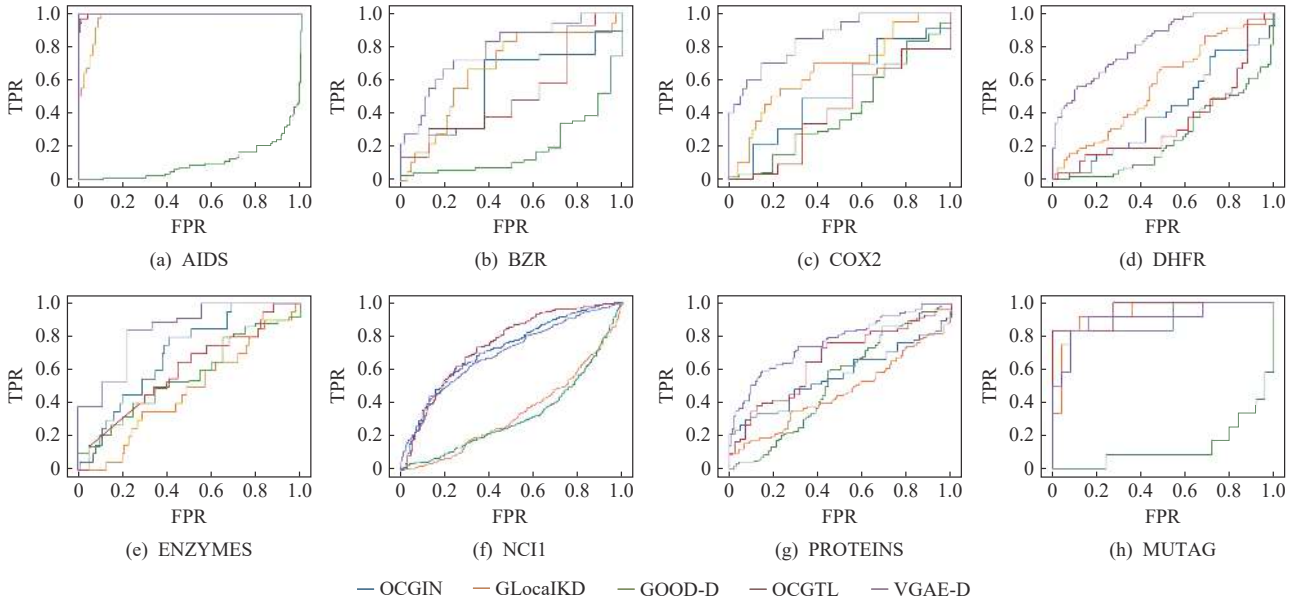


Fig. 3 The AUC curve of the end-to-end model on each dataset when the anomaly label is 1

图 3 异常标签为 1 时端到端模型在各数据集上的 AUC 曲线

Table 4 Comparison of AUC at Different Training Anomalous Samples Ratios when Anomaly Label is 0

表 4 异常标签为 0 时不同训练用异常样本比率下的 AUC 对比 %

数据集	模型						
	OCGIN	GLocalKD	GOOD-D	OCGTL	VGAE-D $r=0.1$	VGAE-D $r=0.2$	VGAE-D $r=0.5$
AIDS	94.70	90.66	93.22	95.85	99.64	99.68	**99.83**
BZR	**81.94**	61.80	76.14	78.81	58.36	61.06	75.98
COX2	60.81	54.48	56.19	60.00	63.73	61.59	**65.37**
DHFR	63.46	65.18	60.63	52.82	58.52	67.94	**71.86**
ENZYMES	52.80	54.88	**64.13**	49.00	54.78	51.30	57.80
NCI1	53.83	68.32	61.56	61.78	64.55	67.76	**68.80**
PROTEINS	**74.22**	68.93	63.35	65.4	65.93	66.02	67.02
MUTAG	65.38	88.46	77.20	82.05	90.84	90.32	**91.43**

注：黑体数值表示最优数值.

Table 5 Comparison of AUC at Different Training Anomalous Samples Ratios when Anomaly Label is 1

表 5 异常标签为 1 时不同训练用异常样本比率下的 AUC 对比 %

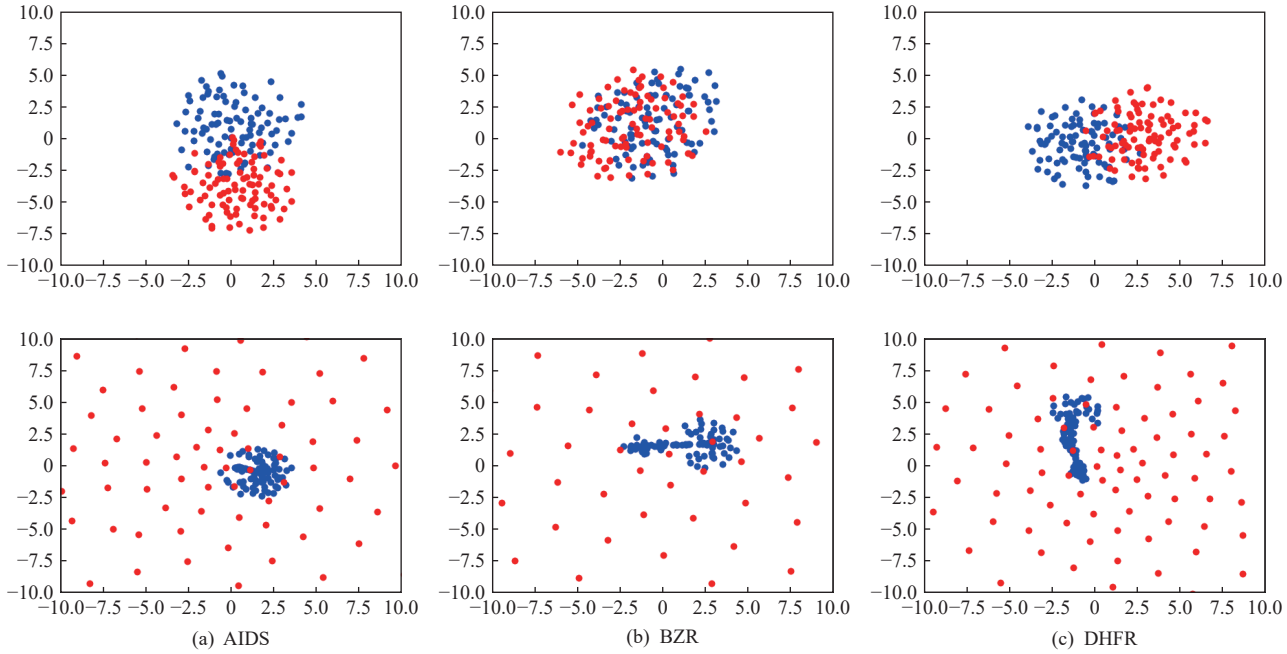
数据集	模型						
	OCGIN	GLocalKD	GOOD-D	OCGTL	VGAE-D $r=0.1$	VGAE-D $r=0.2$	VGAE-D $r=0.5$
AIDS	98.06	95.76	13.50	97.57	99.44	99.58	**99.74**
BZR	73.95	66.50	33.40	62.15	61.56	67.26	**76.22**
COX2	51.08	61.90	35.56	50.27	64.18	62.28	**65.88**
DHFR	60.96	55.64	31.17	60.38	57.36	67.69	**73.97**
ENZYMES	**64.32**	48.62	50.29	62.60	53.84	51.50	57.90
NCI1	**72.72**	31.77	32.87	71.77	64.20	66.03	68.81
PROTEINS	56.31	54.26	52.41	64.11	70.14	65.28	**76.24**
MUTAG	92.78	87.79	16.81	**94.87**	91.01	91.32	90.83

注：黑体数值表示最优数值.

练时模型能从少量的异常图样本数据上学习到图异常的潜在原因和表现形式,在较少异常数据的数据集上也能够达到不错的异常检测效果,因此能在较大程度上克服性能翻转问题.

4.7 可视化实验

VGAE-D 图级别异常检测方法的关键在于差异化正常图和异常图在编码空间中的分布模式,为了检验模型对正常图和异常图编码信息的区分效果,本文在 AIDS, BZR 和 DHFR 三个数据集上各随机选取了一组样本进行可视化实验,每组样本包含 1 个正常图样本和 1 个异常图样本. 实验所用的可视化降维方法为 t-SNE[47],实验结果如图 4 所示. 其中,蓝色节点代表正常图的图编码信息,红色节点代表异常图的图编码信息.



注:蓝色节点为正常图的图编码信息,红色节点为异常图的图编码信息,第 1 行图片为样本组在训练前的编码信息分布,第 2 行图片为样本组在训练后的编码信息分布.

Fig. 4 Visualization of graph encoding information distribution

图 4 图编码信息分布可视化

从图 4 可以看出,训练后的模型能使正常图的编码信息更加聚集,异常图的编码信息更加发散,从而对正常图和异常图的编码信息的分布模式进行区分,这将便于异常评估模块准确区分正常图和异常图.

4.8 消融实验

在 VGAE-D 中, VGAE 是重要的构成部分,为了探究 VGAE-D 中不同模块对模型效果的影响以及模型架构设计的合理性,本文着重针对具有异常感知能力的 VGAE 进行了消融实验,并构造了多个变体模型进行对比:

1)ours-evl. 该变体模型通过消融 L_{std},将具有异常感知能力的 VGAE 退化为普通的 VGAE,不再控制编码空间中正常图和异常图编码信息的离散程度.

2)ours-gae. 该变体模型将原网络架构中使用的 VGAE 替换为普通的 GAE,不再通过采样的方式得到图编码信息,这会破坏编码信息的平滑性.

3)ours-struct. 该变体模型通过消融 L_{struct} 移除了对图结构重构过程的控制,即在重构图时重点关注图节点的属性信息而不关注图的结构信息.

4)ours-attr. 该变体模型通过消融 L_{attr} 移除了对图节点属性重构过程的控制.

5)ours-vgae. 该变体模型通过消融 L_{VGAE} 来探究 VGAE 对模型性能的影响.

消融实验的实验结果如图 5 和图 6 所示,当消融掉 L_{attr}, L_{struct} 和 L_{VGAE} 时,模型性能在大部分数据集上都出现了不同程度的下降. 这些损失的存在能增强模型对图结构和属性信息的学习能力,当消融掉这些损失时,图编码信息的丰富性受损,因此模型的异常检测能力出现了下降. 从实验结果中也可以看出,当消融掉 L_{std} 后,由于不再区分正常图和异常图的分布模式,不利于后续对图进行异常评估,因此模型的异常检测能力出现了下降. 同时,使用 VGAE 比用普通的 GAE 效果更好,这是因为 VGAE 中的高斯采样机制能使图编码空间更加平滑和连续. 总的来说,本

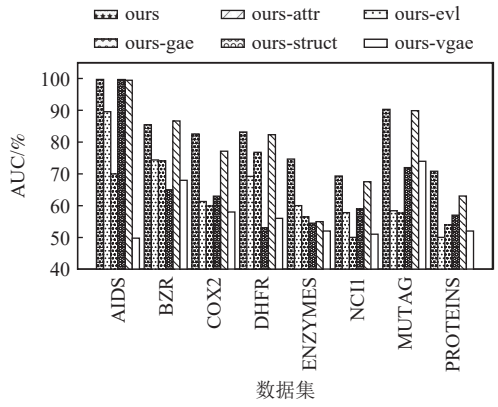


Fig. 5 Results of ablation experiments when the anomaly label is 0

图 5 异常标签为 0 时消融实验结果

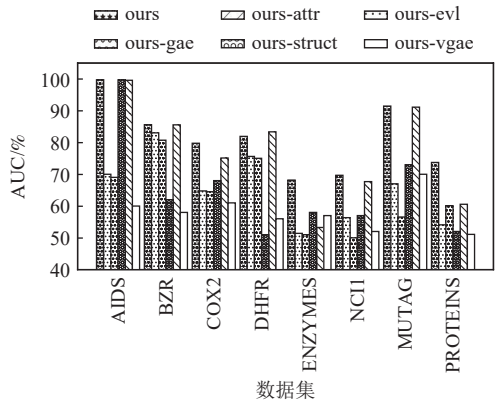


Fig. 6 Results of ablation experiments when the anomaly label is 1

图 6 异常标签为 1 时消融实验结果

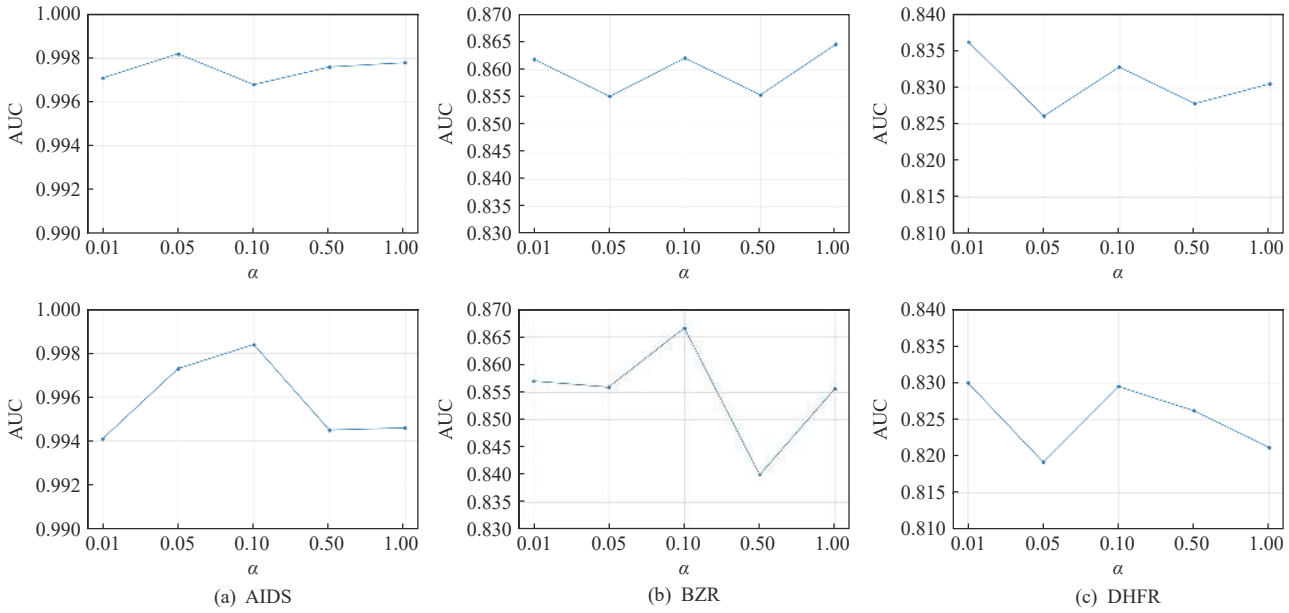
节的消融实验很好地验证了 VGAE-D 模型架构设计的合理性和有效性.

4.9 超参数实验

超参数α能对 VGAE 部分的损失进行缩放进而影响模型的收敛效果, 为了探究其对模型性能的影响, 本节同样选取 AIDS, BZR 和 DHFR 三个数据集展开参数实验. 图 7 展示了异常标签分别为 0 和 1 时, α在不同取值下的实验结果, 可以发现α的取值并不会对模型最终的异常检测性能产生显著影响, 综合来看, α取值在 0.01~0.1 之间较为合理. 同时从表 1 和图 7 可以观察到, 3 个数据集中, DHFR 数据集的平均节点数和边数最多, 当α=0.01 时, 模型的异常检测效果最好. 这是因为α的具体取值与数据集规模无关, 只与数据集中图的平均节点个数和边数有关. 当数据集中图的平均节点个数和边数越多时, 模型初始对图进行重构产生的重构损失和 KL 散度损失可能就会越大, 即L_{VGAE}可能就会越大, 此时减小α的取值可以控制L_{VGAE}与其他 2 部分损失在一个数量级, 利于模型收敛.

5 结 论

本文提出的 VGAE-D 图级别异常模型能同时挖掘正常图和异常图的特征信息, 在模型训练过程中引导正常图的编码表示聚集在相对集中的区域而让异常图的编码信息分散到更广泛的区域, 有效地差



图注：第 1 行图片为异常标签为 0 时的实验结果，第 2 行图片为异常标签为 1 时的实验结果.

Fig. 7 Experimental resluts of hyper parameter

图 7 超参数实验结果

异化了正常图和异常图的图编码信息,进而提高了模型的异常检测能力.实验结果表明,VGAE-D 的图异常检测能力优于目前主流的图级别检测方法.

同时得益于 VGAE-D 在少异常样本下的快速学习能力,VGAE-D 在异常标签发生变化时不会出现明显的性能翻转问题,因而能更广泛地应用于图级别异常检测领域.但是 VGAE-D 也存在一些局限性,当训练数据样本没有标签信息或者只有极少量的标签信息时,VGAE-D 将不能很好地适用,未来我们将继续探索如何在图异常样本稀缺以及异常标签缺失的情况下进行图级别异常检测任务.

作者贡献声明:林馥提出了算法思路和实验方案;李明康完成主要实验并撰写论文;罗学雄提供了指导意见并修改了论文;张书豪、张越和王梓桐完成了部分实验.

参 考 文 献

[1] Jiang Qiangrong, Ma Jiajia. A novel graph kernel on chemical compound classification[J]. Journal of Bioinformatics and Computational Biology, 2018, 16(6): 1850026

[2] Newman M E J, Watts D J, Strogatz S H. Random graph models of social networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2002, 99(suppl_1): 2566−2572.

[3] 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综述[J]. 计算机研究与发展, 2021, 58(7): 1436−1455)

[4] Jiang Jiangguo, Chen Jiuming, Gu Tianbo, et al. Anomaly detection with graph convolutional networks for insider threat and fraud detection[C]//Proc of IEEE Military Communications Conf. Piscataway, NJ: IEEE, 2019: 109−114

[5] Pourhabibi T, Ong K L, Kam B, et al. Fraud detection: A systematic literature review of graph-based anomaly detection approaches[J]. Decision Support Systems, 2020, 133: 1−15

[6] Zhang Ge, Wu Jia, Yang Jian, et al. Fraudre: Fraud detection dual-resistant to graph inconsistency and imbalance[C]//Proc of IEEE Int Conf on Data Mining. Piscataway, NJ: IEEE, 2021: 867−876

[7] Zhao Tong, Jiang Tianwen, Shah N, et al. A synergistic approach for graph anomaly detection with pattern mining and feature learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(6): 2393−2405

[8] 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, 2021, 33(6): 2378−2392

[9] Zhao Lingxiao, Akoglu L. On using classification datasets to evaluate graph outlier detection: Peculiar bservations and new insights[J]. Big Data, 2023, 11(3): 151−180

[10] Qiu Chen, Kloft M, Mandt S, et al. Raising the bar in graph-level anomaly detection[C]//Proc of Int Joint Conf on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, 2022: 2196−2203

[11] Ma Rongrong, Pang Guangsong, Chen Lingpang, et al. Deep graph-level anomaly detection by glocal knowledge distillation[C]//Proc of the 15th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2022: 704−714

[12] Liu Yixin, Ding Kaize, Liu Huan, et al. GOOD-D: On unsupervised graph out-of-distribution detection[C]//Proc of the 16th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2023: 339−347

[13] Pan Shirui, Hu Ruiqi, Long Guodong, et al. Adversarially regularized graph autoencoder for graph embedding[C]//Proc of Int Joint Conf on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, 2018: 2609−2615

[14] Kipf T N, Welling M. Variational graph auto-encoders[J]. Stat, 2016, 1050: 21

[15] Reiser P, Neubert M, Eberhard A, et al. Graph neural networks for materials science and chemistry[J]. Communications Materials, 2022, 3(1): 1−18

[16] Scarselli F, Gori M, Tsoi A C, et al. The graph neural network model[J]. IEEE Transactions on Neural Networks, 2008, 20(1): 61−80

[17] Micheli A. Neural network for graphs: A contextual constructive approach[J]. IEEE Transactions on Neural Networks, 2009, 20(3): 498−511

[18] Chen Ming, Wei Zhewei, Huang Zengfeng, et al. Simple and deep graph convolutional networks[C]//Proc of Int Conf on Machine Learning. New York: ACM, 2020: 1725−1735

[19] Gao Hongyang, Wang Zhengyang, Ji Shuiwang. Large-scale learnable graph convolutional networks[C]//Proc of the 24th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2018: 1416−1424

[20] Wu Fe, Souza A, Zhang Tianyi, et al. Simplifying graph convolutional networks[C]//International Conference on Machine Learning. New York : ACM, 2019: 6861−6871.

[21] Xu K, Hu Weihua, Leskovec J, et al. How powerful are graph neural networks?[C]//Proc of Int Conf on Learning Representations. https://doi.org/10.48550/arXiv.1810.00826

[22] Fan Shaohua, Wang Xiao, Shi Chuan, et al. One2multi graph autoencoder for multi-view graph clustering[C]// Proc of The Web Conf. New York: ACM, 2020: 3070−3076

[23] Hou Zhenyu, Liu Xiao, Cen Yukuo, et al. Graphmae: Self-supervised masked graph autoencoders[C]//Proc of the 28th ACM SIGKDD Conf on Knowledge Discovery and Data Mining. New York: ACM, 2022: 594−604

[24] Velickovic P, Cucurull G, Casanova A, et al. Graph attention networks[J]. Stat, 2018, 1050: 4

[25] Wang Xiao, Ji Houye, Shi Chuan, et al. Heterogeneous graph attention network[C]//Proc of the World Wide Web Conf. New York: ACM, 2019: 2022−2032

[26] Ji Ziwei, Lee N, Frieske R, et al. Survey of hallucination in natural

language generation[J]. ACM Computing Surveys, 2023, 55(12): 1−38

[27] Chakrabarti D. Autopart: Parameter-free graph partitioning and outlier detection[C]//Proc of European Conf on Principles of Data Mining and Knowledge Discovery. Berlin: Springer, 2004: 112−124

[28] Kuang Da, Ding C, Park H. Symmetric nonnegative matrix factorization for graph clustering[C]//Proc of the 2012 SIAM Int Conf on Data Mining. Philadelphia, PA: SIAM, 2012: 106−117

[29] Nikulin V, Huang T H. Unsupervised dimensionality reduction via gradient-based matrix factorization with two adaptive learning rates[C]//Proc of ICML Workshop on Unsupervised and Transfer Learning. New York: ACM, 2012: 181−194

[30] Nguyen H T, Liang P J, Akoglu L. Anomaly detection in large labeled multi-graph databases[J]. arXiv preprint, arXiv: 2010.03600, 2020

[31] Shervashidze N, Schweitzer P, Van Leeuwen E J, et al. Weisfeiler-Lehman graph kernels[J]. Journal of Machine Learning Research, 2011, 12(9): 2539−2561

[32] Neumann M, Garnett R, Bauckhage C, et al. Propagation kernels: Efficient graph kernels from propagated information[J]. Machine Learning, 2016, 102: 209−245

[33] Scherer P, Lio P. Learning distributed representations of graphs with Geo2DR[J]. arXiv preprint, arXiv: 2003.05926, 2020

[34] Liu F T, Ting Kaiming, Zhou Zhihua. Isolation forest[C]// Proc of Int Confon Data Mining. Piscataway, NJ: IEEE, 2008: 413−422

[35] 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

[36] Schölkopf B, Williamson R C, Smola A, et al. Support vector method for novelty detection[J]. Advances in Neural Information Processing Systems, 1999, 12: 583−588

[37] Eswaran D, Faloutsos C, Guha S, et al. Spotlight: Detecting anomalies in streaming graphs[C]//Proc of the 24th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2018: 1378−1386

[38] Zheng Li, Li Zhenpeng, Li Jianli, et al. AddGraph: Anomaly detection in dynamic graph using attention-based temporal GCN[C]//Proc of Int Joint Conf on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, 2019: 3−7

[39] Manzoor E, Milajerdi S M, Akoglu L. Fast memory-efficient anomaly detection in streaming heterogeneous graphs[C]//Proc of the 22nd ACM SIGKDD International Conf on Knowledge Discovery and Data Mining. New York: ACM, 2016: 1035−1044

[40] 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

[41] Zhao Lingxiao, Sawlani S, Srinivasan A, et al. Graph anomaly detection with unsupervised GNNs[J]. arXiv preprint, arXiv: 2210.09535, 2022

[42] Luo Xuexiong, Wu Jia, Yang Jian, et al. Deep graph level anomaly detection with contrastive learning[J]. Scientific Reports, 2022, 12(1): 19867

[43] Lin Fu, Luo Xuexiong, Wu Jia, et al. Discriminative graph-level anomaly detection via dual-students-teacher model[C]//Proc of Int Conf on Advanced Data Mining and Applications. Berlin: Springer, 2023: 261−276

[44] Haefely W. Tranquilizers[J]. psychopharmacolosy, 1983, 2: 92-182.

[45] Sutherland J J, O'brien L A, Weaver D F. Spline-fitting with a genetic algorithm: A method for developing classification structure−activity relationships[J]. Journal of Chemical Information and Computer Sciences, 2003, 43(6): 1906−1915

[46] Zhang Zheng, Deng Xiaogang. Anomaly detection using improved DeepSVDD model with data structure preservation[J]. Pattern Recognition Letters, 2021, 148: 1−6

[47] Van der Maaten L, Hinton G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9(11): 2579−2605



Lin Fu, born in 1978. PhD, associate professor. His main research interests include graph neural network, data mining, and machine learning.

林　馥, 1978 年生. 博士, 副教授. 主要研究方向为图神经网络、数据挖掘、机器学习.



Li Mingkang, born in 2001. Master candidate. His main research interests include graph neural network and machine learning.

李明康, 2001 年生. 硕士研究生. 主要研究方向为图神经网络、机器学习.



Luo Xuexiong, born in 1995. PhD candidate. His main research interests include graph neural network, graph anomaly detection, and brain graph analysis.

罗学雄, 1995 年生. 博士研究生. 主要研究方向为图神经网络、图异常检测、脑图分析.



Zhang Shuhao, born in 2003. Undergraduate. His main research interests include knowledge graph and machine learning.

张书豪, 2003 年生. 本科生. 主要研究方向为知识图谱、机器学习.



Zhang Yue, born in 2000. Master candidate. His main research interests include graph model and machine learning.

张　越, 2000 年生. 硕士研究生. 主要研究方向为图模型、机器学习.



Wang Zitong, born in 2001. Master candidate. His main research interests include knowledge graph and intelligent computing.

王梓桐, 2001 年生. 硕士研究生. 主要研究方向为知识图谱、智能计算.