计算机研究与发展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202330627
61(7): 1670−1680, 2024

基于半监督学习的未知异常检测方法

程渝栋　周　昉
（华东师范大学数据科学与工程学院　上海　200062）
（chengyd@stu.ecnu.edu.cn）

Semi-Supervised Learning-Based Method for Unknown Anomaly Detection

Cheng Yudong and Zhou Fang
（*School of Data Science and Engineering*, *East China Normal University*, *Shanghai* 200062）

Abstract　Anomaly detection aims to identify data that deviates from expected behavior patterns. Despite the potential of semi-supervised anomaly detection methods in enhancing detection accuracy by utilizing a limited amount of labeled data as prior knowledge, the labeled anomalies (i.e., seen anomalies) acquired are unlikely to cover all types of anomalies. In real-world scenarios, novel types of anomalies (i.e., unseen anomalies) often emerge, which may exhibit distinct characteristics from the known anomalies, thereby rendering them challenging to detect using existing semi-supervised anomaly detection methods. To address this issue, we propose a semi-supervised unknown anomaly detection (SSUAD) method, aimed at simultaneously identifying both known and unseen anomalies. This method utilizes a closed-set classifier for the classification of known anomalies and normal instances, and an unknown anomaly detector for the detection of unseen anomalies. Moreover, considering the extreme imbalance between anomalies and normal instances in the anomaly detection scenario, we design an effective data augmentation strategy to increase the number of anomaly samples. Experiments are conducted on UNSW-NB15 and KDDCUP99 datasets, as well as a real-world dataset SQB. The results reveal that, compared with existing anomaly detection methods, SSUAD exhibits significant improvement in the anomaly detection performance metrics AUC-ROC and AUC-PR, thereby verifying the effectiveness and reasonableness of the proposed method.

Key words　anomaly detection; out-of-distribution detection; semi-supervised learning; unseen anomalies; tabular data

摘　要　异常检测旨在识别偏离预期行为模式的数据. 虽然半监督异常检测方法可以充分利用有限的标签数据作为先验知识来提高检测准确性,但是收集到的标记异常（即已知异常）很难覆盖所有类型的异常并且在现实场景中往往存在着一些新型的异常（即未知异常),这些异常可能与已知异常表现出不同的特性,因此难以被现有的半监督异常检测方法识别. 针对该问题,提出了一种基于半监督学习的未知异常检测（semi-supervised unknown anomaly detection, SSUAD）方法,旨在同时识别已知异常和未知异常. 该方法利用闭集分类器对已知异常和正常分类,利用未知异常检测器检测未知异常. 此外,还考虑了异常场景中异常和正常极端不平衡的情况,设计了有效的数据增强方法来扩充异常样本的数量. 在 UNSW-NB15 和 KDDCUP99 数据集以及一个真实数据集 SQB 上进行了实验,实验结果表明,相较于现有的异常检测方法,SSUAD 在异常检测性能指标 AUC-ROC（area under receiver operating characteristic curve）和 AUC-PR（area under precision-recall curve）上都有明显的提升. 证明了 SSUAD 的有效性和合理性.

收稿日期: 2023-07-31; **修回日期**: 2023-12-07
基金项目: 上海市科技创新行动计划项目（23511100700）
This work was supported by the Municipal Science and Technology Innovation Action Plan of Shanghai (23511100700).
通信作者: 周昉（fzhou@dase.ecnu.edu.cn）

关键词 异常检测；分布外检测；半监督学习；未知异常；表格数据

中图法分类号 TP391

异常检测旨在识别偏离正常模式的异常行为. 在大数据的背景下，尤其是网络和互联网技术迅速发展的驱动下，异常检测已经成为一种关键的数据分析技术. 在诸如信用卡欺诈检测[1]、网络入侵检测[2]、罕见疾病检测[3]等多个领域，异常检测都发挥着至关重要的作用. 尽管在过去数十年中，已经提出了众多的异常检测方法并取得了显著的成果，但其中的大部分方法是基于完全未标记数据的无监督方法[4]. 由于这些方法缺乏真实异常标记数据的先验知识，可能会将很多噪声数据误判为异常，从而导致检测的准确率降低.

近年来，半监督方法[5]旨在利用少量标记的异常样本来训练模型的同时，以减少对大规模手动标注的依赖. 这类方法的研究动机主要源于实际应用，通常能以较小的成本获取一小部分标记的异常样本. 例如，在欺诈检测场景中，金融机构可能已掌握一小部分已确认为欺诈的交易信息. 这些已标记的异常样本提供了异常行为的重要指示，能够显著提高模型的检测精度[6]. 然而，异常行为本质上是未知的，所以这些已标记的异常样本通常是不完整的、不准确的. 目前的半监督方法大都侧重于识别在训练期间已经出现的异常（即已知异常），并不能推广到那些在训练期间未见过的新型异常（即未知异常）. 而在现实世界中，通常存在新型的异常，主要原因有 2 方面：1）有限的已标记的异常很难覆盖所有类型的异常；2）随着时间的推移和非法欺诈手段的不断创新，导致新型的异常不断涌出. 例如，欺诈者可能会滥用新的技术或设备进行非法获利. 此外，因为训练异常数据的数量比较少，所以这类方法检测已知异常的能力也常常受到限制. 因此，有必要提出一种量身定制的方法，可以有效地检测已知和未知的异常，以消除潜在危险.

然而，要实现这个目标是一件棘手的事情，主要存在 2 个方面的问题.

1）如何识别已知异常并检测混合在未标记数据集中的未知异常. 传统的半监督异常检测学习任务在处理未标记数据时，通常假定这些数据是正常样本. 然而在实际场景中，考虑到异常属于未知事件，并且具有新颖性、多样性以及不确定性，这种假设有可能会忽略那些潜藏在未标记数据中的未知异常. 而这些未知异常如果没有能被及时发现，将对模型的异常检测能力产生影响. 针对这个问题，Huang 等人[7]提出了鲁棒性的半监督学习方法，所有被检测到的新样本被简单地视为异常.

2）如何克服异常检测场景中由于正常和异常类不均衡而导致的异常样本学习不充分的问题. 在实际的异常检测任务中，由于正常样本通常占主导地位而异常样本相对稀疏，这种不平衡性可能导致模型在学习的过程中过度拟合正常样本的特征，从而无法捕捉和学习异常样本的特征. 为了解决这一挑战，可以采用数据增强技术以增强数据的多样性，从而有效地缓解样本不平衡问题. 例如，在图像异常检测领域，Li 等人[8]通过剪贴操作来创造空间上的局部不规则性. 在视频领域，Meng 等人[9]通过将视频帧进行随机裁剪和水平翻转来生成新的异常样本. 然而，由于表格数据缺乏丰富的上下文信息，因此很难将传统的数据增强方法应用于表格域中，甚至可能引入额外的噪声.

为了应对上述挑战，本文提出了基于半监督学习的未知异常检测（semi-supervised unknown anomaly detection, SSUAD）模型. 该模型首先引入一种表格自监督学习策略，设置恢复原始值和恢复用于损坏输入的掩码 2 个预设任务，来挖掘数据的内在结构和模式，以实现高效的表示学习. 其次，结合了基于树的特征选择方法与掩码机制，提出了一种新颖且可靠的数据增强方法，以增加异常数据的多样性. 最后，利用训练好的闭集分类器区分已知异常和正常，未知异常检测器来检测未知异常.

本文的主要贡献包括 3 个方面：

1）提出了一种基于半监督学习的未知异常检测（SSUAD）模型，该模型利用有限的标记异常样本和正常样本，训练闭集分类器和未知异常检测器来实现对已知异常和未知异常的有效检测；

2）提出了一种新的数据增强策略，通过融入特征重要性调整的掩码向量来扩充异常样本的数量，改善类别不平衡问题；

3）在公共和真实的数据集上评估了 SSUAD 方法，实验结果证明了 SSUAD 的有效性和合理性.

1 相关工作

1.1 异常检测

异常检测技术作为热点研究领域，近几十年来

研究人员提出了很多优秀的工作. 由于收集大量标记数据的难度和成本很大，有监督异常检测方法并不符合实际[4]. 目前，主流的异常检测方法可以分为2大类：无监督异常检测方法和半监督异常检测方法.

无监督异常检测方法是一种不依赖于标签信息，而是通过挖掘数据内在的分布特性和统计规律，来识别不符合整体数据分布特性的异常数据的方法. 虽然无监督异常检测方法避免了手动标注数据的昂贵成本，但是由于过度依赖异常分布的假设和缺乏异常的先验知识，这些方法在现实数据集上往往存在较高的误报率[10].

半监督异常检测方法，侧重于利用标记正常样本来学习正常类别的模式. 例如，OCSVM[11]通过建立超平面来分割样本，最大化正负样本之间的分离. 鉴于在许多实际应用中有少量的异常数据可用，最近的半监督异常检测方法则致力于利用这些少量的标记异常数据来学习模型，已证明这些有限的标记异常数据可以显著提高异常检测的准确性. 例如，Pang等人[12]提出的DevNet模型，借助神经偏差网络实现异常分数的端到端学习，用少量标记的异常样本和高斯先验，使异常样本的分数显著偏离正常样本分布的两端. 此外，该团队[13]还进一步提出了PreNet模型，通过基于配对关系预测的序数回归网络来学习异常分数. Zhang等人[14]遵循2阶段方法，先聚类标记异常样本，然后对无标记样本进行过滤，选取潜在异常和可靠正常. Ruff等人[15]提出的DeepSAD则利用神经网络将正常样本映射到超球体的中心区域，同时利用标记异常信息，将异常样本映射到超球体的边缘，从而在空间上清晰地区分正常和异常样本. Zhou等人[16]提出的FEAWAD通过自编码器对输入数据进行编码，并结合隐藏表示、重构残差向量和重构误差3个因子，为异常检测提供更有意义的数据表示. Li等人[17]在其工作Dual-MGAN中，通过将2个子GAN结合起来，充分利用已识别的异常的潜在信息，来检测离散的异常点和部分已识别的群体异常. Zong等人[18]提出一种全新的异常检测模型PIA-WAL，利用少量有标记的异常样本指导对抗学习. 然而，由于有限的标记异常很难覆盖所有类型异常，这些方法在检测未知异常的能力上仍然受限.

1.2 数据增强

数据增强作为一种有效扩充训练样本多样性的策略，能够在无需直接获取额外数据的情况下提高模型的泛化性能. 尽管在视觉和自然语言处理领域，诸如图像旋转以及句子重组等数据增强手段已被证明具有显著成效，但是在处理缺乏空间信息且高度异构的表格数据时，这些传统的数据增强方法往往难以被直接应用. 此外，还有一些不局限于数据类型的数据增强方法. 例如，SMOTE[19]通过在少数类样本之间进行插值来生成新的样本，以此来扩充少数类样本的数量. Mixup[20]则通过混合不同样本生成包含多种类别信息的新样本，增强了数据的多样性. 然而这些方法[19−20]都可能会生成分布外(out-of-distribution, OOD)的样本，从而引入额外的噪声.

1.3 分布外检测

OOD检测的目标在于识别与训练数据不同分布的样本. 在大多数情况下，OOD研究的核心是如何确定与训练数据集有明显分布差异的数据. 考虑到未知异常的新颖性、多样性以及不确定性，我们认为这些未知异常在特征空间上与已知异常和正常数据都存在显著差异. 这为我们提供了一个理论基础，即通过分布差异来识别未知异常. 在此背景下，本文不仅借鉴了传统OOD检测的思想，而且进一步推进了这一领域的研究. 我们不单纯地将未知异常视为OOD样本，而是结合传统的OOD检测方法和对未知异常独特性的深入分析，提出了一个新的方法来识别它们.

早期的OOD研究大多集中在识别并移除未标记数据中的OOD样本上. 例如，Chen等人[21]提出不确定性感知自蒸馏方法框架UASD，重点是从未标记数据中剔除OOD样本. Yu等人[22]采用多任务课程学习框架来同时检测未见过的类别并分类已见过的类别. 近年来，OOD研究开始考虑如何有效利用识别出的OOD样本，以减少对已知分类器的扰动，从而提高对未知类别的识别率. 例如，Huang等人[23]的TOOR模型虽然通过迁移学习将OOD样本分类为可循环利用和不可循环利用的OOD样本. 尽管上述方法[21−23]能够检测出与训练数据分布不同的样本，但未考虑到类别分布不均匀的问题.

2 SSUAD模型

2.1 问题定义

给定一个数据集 D 由标记部分 $D_{\mathrm{L}}=\{(\boldsymbol{X}^i,y^i)\}_{i=1}^{m}$ 和未标记部分 $D_{\mathrm{U}}=\{(\boldsymbol{X}^i)\}_{i=1}^{n}$ 构成. 其中 m 和 n 分别是标记数据和未标记数据的个数. 通常情况下，$n\gg m$. $\boldsymbol{X}^i=(x_1^i,x_2^i,\cdots,x_d^i)\in\mathbb{R}^d$ 是第 i 个样本，d 是特征维度. 记 $y\in C_{\mathrm{L}}=\{0,1,\cdots,C^{\mathrm{in}}\}$ 为标记数据中的标签类别. 特别地，$y=0$ 表示正常样本，而 $y\in\{1,2,\cdots,C^{\mathrm{in}}\}$ 则代表不同

类型的已知异常. 记C_{U}为未标记数据中的类别总数. 考虑到现实场景中, 未标记数据集中包含一些在标记数据中不存在的未知异常类, 即$C_{\mathrm{L}} \subset C_{\mathrm{U}}$, 则未标记数据中的未知异常类可以表示为$C^{\mathrm{out}} = C_{\mathrm{U}} \backslash C_{\mathrm{L}}$. 本文的目标是利用标记数据集$D_{\mathrm{L}}$和未标记数据集$D_{\mathrm{U}}$训练闭集分类器和未知异常检测器, 实现在正确区分正常和已知异常的同时, 也能够识别未知异常.

2.2 模型描述

图 1 展示了 SSUAD 的框架图. 该模型分为 2 部分: 表征学习和异常检测.

1)表征学习. 为了充分地利用标记数据和未标记数据来预训练潜在空间, SSUAD 通过设置恢复原始值和恢复用于损坏输入的掩码这 2 个预设任务, 从损坏的输入数据中学习到有效的特征表示, 提高数据表征学习的鲁棒性.

2)异常检测.SSUAD 利用闭集分类器和未知异常检测器对已知异常和未知异常进行检测, 其中闭集分类器主要区分已知异常和正常, 未知异常检测器通过对比未知类别与已知类别在特征空间中的差异来识别未知异常.

2.3 表征学习

为了使用标记数据和未标记数据的预训练潜在空间, 本文采用自监督学习, 通过引入特征向量估计和掩码向量估计 2 个预设任务, 使得特征向量估计器能够从被损坏的样本中恢复输入样本, 同时通过掩码向量估计器来预测应用于样本的掩码向量.

1)掩码向量估计. 掩码向量估计器使用编码器产生的低维特征表示作为输入, 并输出一个预测向量, 该向量预测输入样本中被噪声替代的特征的位置, 即掩码向量. 具体来说, 首先通过掩码生成器生成一个二元掩码$\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_d)^{\mathrm{T}} \in \{0,1\}^d$, 其中$\mu_j$从概率为$p_{\mathrm{m}}$的伯努利分布中随机采样. 然后构造损坏的向量$\tilde{\boldsymbol{X}}$, 这通过将$\mu_j = 1$的$\boldsymbol{X}$的每个维度$j$替换为从训练集中随机采样的实例的维度$j$实现, 具体表现为$\tilde{\boldsymbol{X}} = (\boldsymbol{1} - \boldsymbol{\mu}) \odot \boldsymbol{X} + \boldsymbol{\mu} \odot \bar{\boldsymbol{X}}$, 其中⊙表示元素间的乘积, $\bar{\boldsymbol{X}}$是随机采样的样本, $\bar{\boldsymbol{X}}$的每个$\bar{x}_j$都是从第j个特征的经验边际分布中随机采样的. 掩码向量估计器的目标是使得估计的掩码向量$\hat{\boldsymbol{\mu}}$尽可能地接近实际的掩码向量$\boldsymbol{\mu}$. 本文采用二元交叉熵损失作为掩码向量估计的损失函数, 损失函数如式(1)所示:

$$L_{\mathrm{mask}} = -\frac{1}{d}\sum_{j=1}^{d}[\mu_j \ln(\hat{\mu}_j) + (1-\mu_j)\ln(1-\hat{\mu}_j)]. \tag{1}$$

2)特征向量估计. 特征向量估计器利用编码器的输出来预测输入样本的特征向量. 其目标是使损坏的样本$\tilde{\boldsymbol{X}}$尽可能地接近未被损坏的原始输入$\boldsymbol{X}$. 本文采用均方误差作为特征向量估计的损失函数, 该函数衡量的是预测的特征向量$\hat{\boldsymbol{X}}$和原始特征向量$\boldsymbol{X}$之间的差距. 其损失函数如式(2)所示:

$$L_{\mathrm{recons}} = -\frac{1}{d}\sum_{j=1}^{d}(x_j - \hat{x}_j)^2. \tag{2}$$

最终, 编码器、掩码向量估计器以及特征向量估计器将通过优化损失函数进行训练:

$$L_{\mathrm{vime}} = L_{\mathrm{recons}} + \lambda_{\mathrm{mask}} L_{\mathrm{mask}}, \tag{3}$$

其中, λ_{mask}为平衡因子.

2.4 异常检测

2.4.1 异常样本数据增强

为解决异常检测任务中正负样本极度不平衡的问题, 本文引入了一种基于特征重要性的数据增强策略来扩充数据的多样性. 考虑到表格数据的特征可能具有显著的相关性. 例如在异常商户检测的实际应用场景中, 日交易额度、日余额等关键特征不仅

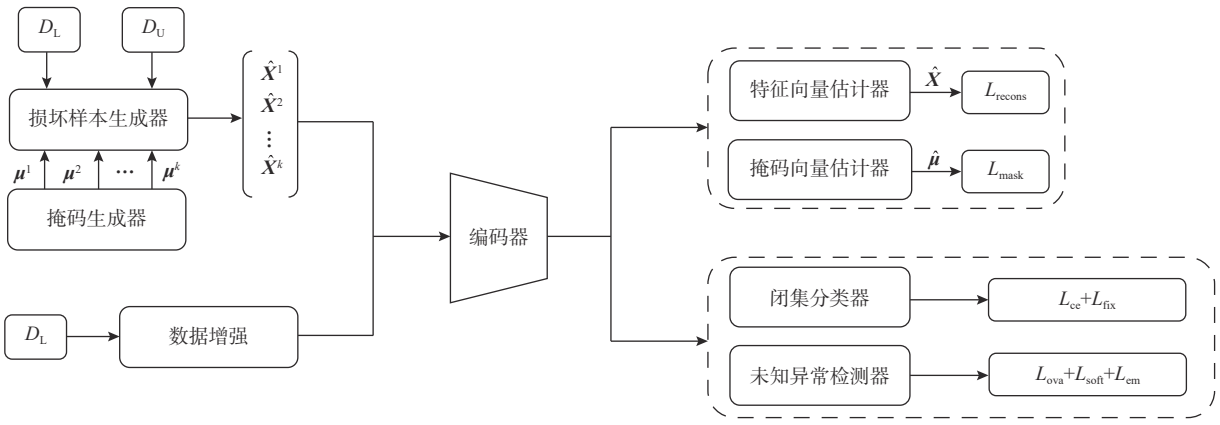

Fig. 1 The frame diagram of SSUAD

图 1 SSUAD 框架图

存在相关性，而且对于商户的异常状态判断起到决定性的作用. 因此，本文引入一种基于特征重要性的掩码生成器，该生成器产生多个掩码向量，且针对重要特征的掩码操作进行了限制.

本文采用基于树的模型，如决策树或随机森林来进行特征选择. 这是因为树模型在分裂时会优先选择最具区分性的特征，并通过评估特征在分裂点的纯度提升来为每个特征赋予一个重要性评分. 这种特征重要性评分不仅反映了每个特征对模型预测性能的贡献度，还考虑了特征间的交互效应. 为选择合适的树模型，通过比较不同模型在验证集上的性能来确定最优的树模型. 在生成掩码向量的过程中，利用树模型内置的特征重要性为每个特征赋予重要性评分，根据得到的特征重要性评分来调整每个特征位置的处理方式. 具体而言，先将所有特征的重要性评分从高到低进行排序，然后根据预设的比例阈值来选择需要保留的特征，对剩余的特征进行掩码操作. 当生成了考虑特征重要性的掩码向量后，进一步将这些向量与已标记的异常样本结合. 对于被掩码的特征维度，添加不同程度的高斯噪声，以生成增强后的异常样本. 这些增强后的异常样本不仅更接近原始数据的真实分布，而且强调了关键特征的作用.

2.4.2 识别未知异常

本文借鉴 OOD 检测的思想，将标记数据集中的样本类别视为已知类别，包括正常样本和已知异常样本，同时将未标记数据集中的未知异常视作未知类别. 本文利用未知异常检测器来检测未标记数据集中的未知类. 针对每一个已知类都会构建子分类器，输出距离表示样本与对应类别的距离，以此判断样本是否属于其对应的类别. 每个子分类器都会输出一个 2 维向量，该向量代表输入样本是该类别的内部样本还是外部样本. 对于标记数据集 D_{L} 中的样本，最小化损失函数为

$$L_{\mathrm{ova}}=-\frac{1}{m}\sum_{i=1}^{m}-\ln(p^{y^i}(t=0|\boldsymbol{X}^i))-\min_{i\neq y^i}\ln(p^i(t=1|\boldsymbol{X}^i)),\tag{4}$$

其中 $p^{y^i}(t=0|\boldsymbol{X}^i)$ 和 $p^i(t=1|\boldsymbol{X}^i)$ 分别为样本 $\boldsymbol{X}$ 被视为类别 y^i 的内部样本和外部样本的概率，这 2 个概率之和为 1.对于未标记样本，如果该样本在闭集分类预测的最高概率类别对应的子分类器中被判断为外部样本的概率大于 0.5，则将该未标记样本分类为未知类.

考虑到未知类不仅可能与已知类在特征空间的分布不同，还可能具有不同的内在模式或规律，这种差异性会降低未知异常检测器分类的准确性. 为了解决这个问题，对每一个子分类器使用熵最小化策略. 这种熵最小化方法不仅可以提升模型对已知类预测的置信度，还可以保持未知类实例的未知性质，因为未知类并没有被错误地归类到任何一个已知类别中. 对于未标记数据集中的样本，熵最小化损失函数如式(5)所示.

$$L_{\mathrm{em}}=-\frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{C^{\mathrm{in}}}p^j(t=0|\boldsymbol{X}^i)\ln(p^j(t=0|\boldsymbol{X}^i))+p^j(t=1|\boldsymbol{X}^i)\ln(p^j(t=1|\boldsymbol{X}^i)).\tag{5}$$

此外，由于未知类缺乏标签信息，这使得传统半监督学习中所采用的硬一致性损失方法表现不理想. 为了克服这个难点，本文采取优化子分类器对数据增强的平滑性来改善训练信号的传播质量. 具体而言，对同一未标记样本的 2 个增强视图，最小化其在子分类器下的预测结果之间的距离，以达到平滑决策边界的效果. 所采用的软一致性损失函数为

$$L_{\mathrm{soft}}=-\frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{C^{\mathrm{in}}}\sum_{t\in(0,1)}|p^j(t|T_1(\boldsymbol{X}^i))-p^j(t|T_2(\boldsymbol{X}^i))|^2,\tag{6}$$

其中 $T_1(\boldsymbol{X}^i)$ 和 $T_2(\boldsymbol{X}^i)$ 是对应未标记样本 $\boldsymbol{X}^i$ 的 2 个不同的视图，通过向未标记样本 $\boldsymbol{X}^i$ 添加随机高斯噪音来实现的.

2.4.3 识别已知异常和正常

闭集分类器用于对已知类别进行分类识别. 闭集分类器的好坏会影响到未知异常检测器识别未知类的效果. 为了进一步提高闭集分类器的性能，对于标记数据集中的样本 $(\boldsymbol{X}^i,y^i)$，采用标准交叉熵损失函数来优化，损失函数如式(7)所示.

$$L_{\mathrm{ce}}=-\frac{1}{m}\sum_{i=1}^{m}\sum_{j=0}^{C^{\mathrm{in}}}y^j\ln(p^{y^j}).\tag{7}$$

对于未标记数据集中的样本，借鉴半监督学习方法 FixMatch.先通过闭集分类器对未标记样本进行预测，然后利用子分类器进行判断. 当且仅当样本被准确归类为已知类别，并且其预测结果的置信度超过某一特定阈值时，才会将预测结果视为伪标签. 最后计算伪标签的损失函数如式(8)所示.

$$L_{\mathrm{fix}}=-\frac{1}{n}\sum_{i=1}^{n}(\max(q\geqslant\tau)H(\hat{q},F(y|\beta(\boldsymbol{X}^i)))),\tag{8}$$

其中 $q=F(y|\alpha(\boldsymbol{X}^i))$ 表示闭集分类器预测的类别概率，$\hat{q}=\arg\max_j(q)$ 表示闭集分类器预测的最大概率所对应的类别，$\alpha(\cdot)$ 表示数据增强中的弱增强，$\beta(\cdot)$ 表示数据增强中的强增强，τ 表示对应的阈值，本文 τ=0.95.

为了进一步提高模型对异常数据的识别能力，本文采用以下训练策略. 首先，引入特征向量估计和掩码向量估计 2 个预设的任务对表格数据进行预训练，再利用 L_{mask} 和 L_{recons} 这 2 个损失函数进行训练. 然后通过融入特征重要性的掩码向量对标记异常样本进行数据增强，将得到的增强后的数据与原始标记异常数据集合并，采用标准交叉熵损失函数 L_{ce} 和一对多损失函数 L_{ova} 分别训练闭集分类器和未知异常检测器. 对于未标记数据，引入 FixMatch 损失函数 L_{fix} 来进一步优化闭集分类器，同时，通过熵最小化损失函数 L_{em} 和软一致性损失函数 L_{soft} 对未知异常检测器进行精细化调整.

值得注意的是，本文模型采取了阶段性的训练策略. 考虑到在训练的初始阶段，模型尚未达到理想的稳定性和预测准确性. 若过早地对未标记数据中选择的已知类别样本打上伪标签可能会引入噪音，从而对模型造成干扰. 因此，在训练的前 $Epoch_{\text{fix}}$ 个周期内，模型集中在标记数据的学习和优化方面，在模型经历前 $Epoch_{\text{fix}}$ 个周期的训练后，其泛化性能得到提高. 此时，再对未标记数据进行预测并生成相应的伪标签. 这种策略有助于提高模型在未标记数据上的预测能力，使模型更好地检测异常.

3 实　　验

3.1 环境和数据集

本文实验的操作系统平台为 CentOS 7.5 版本. 硬件环境为 2 个 Intel Xeon Gold 6240R 处理器，每个处理器为 24 核 2.40 GHz 主频，以及 4 块 NVIDIA Tesla V100 显卡，每块显卡的显存容量为 32 GB.所有的机器学习方法基于 Scikit-Learn 来实现，所有的深度学习方法基于 Pytorch 来实现.

本实验使用了 2 个公共数据集 UNSW-NB15[24] 和 KDDCUP99[25]. 除此之外，为了将本文模型应用于商户交易异常检测. 本文还在真实数据集 SQB 上进行了相关实验. 下面介绍这 3 个数据集：

1）UNSW-NB15 数据集是最近发布的一组网络入侵数据集，是由澳大利亚网络安全中心的网络范围实验室所提供的. 该数据集共有 107 687 个数据实例，每个实例特征有 196 维，其除了正常网络流量外，还包含了 Generic, Backdoor, Analysis 等 7 种不同的网络攻击类型，将这些网络攻击行为视为异常情况，在整个数据集中占比 21.5%.

2）KDDCUP99 是异常检测通用的一个经典数据集，由 1998 年美国国防高级研究计划局资助的 MIT Lincoln 实验室创立. 本文使用其中的一个数据子集进行实验，该子集包括 77 255 个样本，每个样本包括 34 维特征，其中包含 DOS, R2L, U2R 共 3 种网络攻击行为，并将这些网络攻击行为视为异常情况.

3）SQB 是从聚合支付平台收集的商户交易流水数据的真实数据集，共有 319 133 条交易记录，每条记录包括 183 维特征，商户常见的异常类型包含欺诈、赌博、被动销赃，新型的异常类型包含套现、线上交易、刷单等，所有的异常占整个数据集的 0.6%.

实验所用到的数据集统计信息如表 1 所示.

Table 1　Dataset Statistics

表 1　数据集统计信息

数据集	维度	样本数量	异常比例/%
UNSW-NB15	196	107 687	21.5
KDDCUP99	34	77 255	5
SQB	183	319 133	0.6

为了复制场景需求，首先对整体数据集随机打乱后，按照 8∶2 划分为训练集和测试集. 以 UNSW-NB15 数据集为例，本文设置默认的已知异常类型的个数为 3，然后从 7 种类型的异常中随机抽取 3 个异常类，并分别从选定的异常类中随机抽取 100 个异常样本作为少量有标记的已知异常. 对于训练集中未标记数据集，分别从 7 个异常类中随机抽取异常样本与正常训练样本混合，从而生成含有已知异常和未知异常的未标记数据集. 同时考虑到异常的罕见性，控制未标记数据集中异常样本占整个数据的 5%.

3.2 评价指标

本文采用 2 种在异常检测领域中被广泛采用的评估指标，即接收者操作特性曲线下的面积（AUC-ROC）和精确度-召回率曲线下的面积（AUC-PR）来评估本文方法的性能. 在这 2 个指标中，数值越大代表模型的性能越好. 其中 AUC-ROC 是对真阳性率和假阳性率曲线的概括，但是通常会对模型的性能展示出过于乐观的视角. 而 AUC-PR 则更加贴近异常检测的实际场景，因为它专门对异常类别的精确率和召回率进行概括.

3.3 对比模型

为了证明本文模型的有效性和优越性，SSUAD 将与 4 个被应用于异常检测的模型进行对比实验，实验包含经典的无监督异常检测方法孤立森林（isolation forest, iForest）以及最先进的半监督异常检测方法：ADOA[14], DevNet[12], FEAWAD[16], PreNet[13]. 下

面对这些模型进行简单的介绍.

1) iForest[26]. 是一个基于集成模型的无监督异常检测方法，通过构建孤立树并计算孤立样本所需的步骤数量来检测异常情况. 该方法假设分布稀疏且距离高密度群体较远的点为离群点.

2) ADOA[14]. 遵循 2 阶段步骤：先对标记的异常样本进行聚类，同时对未标记样本进行过滤选取潜在异常和可靠正常实例；然后训练一个加权分类器以进一步检测.

3) DevNet[12]. 是一种基于深度网络的半监督的方法，结合少量标记的异常样本和高斯先验，训练端到端异常分数器.

4) FEAWAD[16]. 通过自编码器对输入数据进行编码，并结合隐藏表示、重构残差向量和重构误差 3 个因子，为异常检测提供更有意义的数据表示.

5) PreNet[13]. 通过基于配对关系预测的序数回归网络来学习异常分数，学习 2 个随机采样的训练实例的关系来推广未知异常.

3.4 实验参数

本文采用随机梯度下降 (stochastic gradient descent, SGD) 方法优化所有深度学习模型，对于 2 个公开数据集 UNSW-NB15 和 KDDCUP99，学习率设置为 0.001，学习器迭代次数设置为 30.对于真实数据集 SQB，学习率设置为 0.000 1，学习器迭代次数设置为 50，惩罚因子 λ_{mask} 设置为 1；对于数据集 UNSW-NB15 和 SQB，设置已知异常类别数为 3，扰动特征概率 p_{m} 设置为 0.09；对于数据集 KDDCUP99，扰动特征概率 p_{m} 设置为 0.3.为了公平地比较不同的模型，本文对数据进行标准化操作，并采用网格搜索策略对模型的超参数进行调优. 每个模型都进行 5 次随机运行实验，并报告了测试性能的平均值. 不同数据集下的基本运行参数设定如表 2 所示.

Table 2 Parameters Setting of Different Datasets

表 2 不同数据集的参数设置

数据集	C^{in}	$Epoch_{\text{fix}}$	p_{m}
UNSW-NB15	3	10	0.09
KDDCUP99	2	10	0.3
SQB	3	15	0.09

SSUAD 在不同数据集下的实现架构见表 3. 其中 E, E_{m}, E_{r}, C, D 分别表示编码器、掩码向量估计器、特征向量估计器、闭集分类器、未知异常检测器，$FC(W_{\text{in}},W_{\text{out}},\sigma)$ 表示具有 W_{in} 个输入神经元、W_{out} 个输出神经元以及激活函数为 σ 的全连接多层神经网络.

3.5 实验结果

3.5.1 模型性能

表 4 展示了 SSUAD 模型和其他对比模型在数据集上的 AUC-ROC 和 AUC-PR 的结果. 从表 4 中可以看出 SSUAD 模型相较于其他方法在这 2 项指标上都获得了明显的提升. 具体地，对比无监督异常检测方法 (iForest)，SSUAD 在 UNSW-NB15，KDDCUP99，SQB 这 3 个数据集上的 AUC-ROC 指标分别提高了 25.8 个百分点、1.9 个百分点、3.7 个百分点. 在 AUC-PR

Table 3 Framework Setup for Different Datasets

表 3 不同数据集下的框架设置

模型	UNSW-NB15	KDDCUP99	SQB
E	*FC*(196,64,ReLu)	*FC*(34,17,ReLu)	*FC*(183,64,ReLu)
	FC(64,64,ReLu)	*FC*(17,17,ReLu)	*FC*(64,64,ReLu)
E_{m}	*FC*(64,196,Tanh)	*FC*(17,34,Tanh)	*FC*(64,183,Tanh)
E_{r}	*FC*(64,196,Tanh)	*FC*(17,34,Tanh)	*FC*(64,183,Tanh)
C	*FC*(64,64,ReLu)	*FC*(17,17,ReLu)	*FC*(64,64,ReLu)
	BatchNorm1d(64)	*BatchNorm1d*(17)	*BatchNorm1d*(64)
	FC(64,4,Softmax)	*FC*(17,3,Softmax)	*FC*(64,4,Softmax)
D	*FC*(64,64,ReLu)	*FC*(17,17,ReLu)	*FC*(64,64,ReLu)
	BatchNorm1d(64)	*BatchNorm1d*(17)	*BatchNorm1d*(64)
	FC(64,8,Softmax)	*FC*(17,6,Softmax)	*FC*(64,8,Softmax)

Table 4 Comparative Results of Anomaly Detection

表 4 异常检测对比结果

数据集	模型	AUC-ROC/%	AUC-PR/%
UNSW-NB15	iForest	68.9	32.2
	ADOA	73.5	28.8
	DevNet	90.6	80.1
	FEAWAD	89.2	75.0
	PreNet	84.6	74.9
	SSUAD(本文)	**94.7**	**86.3**
KDDCUP99	iForest	96.6	71.8
	ADOA	87.8	19.7
	DevNet	94.4	75.8
	FEAWAD	92.4	73.7
	PreNet	95.1	79.2
	SSUAD(本文)	**98.5**	**87.7**
SQB	iForest	92.1	7.6
	ADOA	92.7	11.3
	DevNet	91.9	35.7
	FEAWAD	91.3	25.8
	PreNet	91.6	39.2
	SSUAD(本文)	**95.8**	**42.1**

注：黑体数值表示最优结果.

指标上分别提高了 54.1 个百分点、15.9 个百分点、34.5 个百分点. 这表明无监督异常检测方法缺乏对标记数据的利用, 由于缺乏有效的监督信息指导模型的训练, 对噪声比较敏感, 导致其识别能力下降. 此外, 无监督异常检测方法过度依赖于异常分布的假设, 这也会导致模型无法准确地识别异常. 与基于半监督的异常检测方法(ADOA, DevNet, FEAWAD, PreNet)相比, SSUAD 在 3 个数据集上的 AUC-ROC 指标上分别提高了 4.1~21.2 个百分点、3.4~10.7 个百分点、3.1~4.5 个百分点. 在 AUC-PR 指标上分别提高了 6.2~57.5 个百分点、8.5~68 个百分点、2.9~30.8 个百分点. 其中 ADOA, DevNet, FEAWAD 只能识别已知异常, 并没有考虑到对未知异常的检测, PreNet 虽然通过配对关系推广未知异常, 但是在学习拼接实例上并不能很好地捕捉到正常和异常的特性. 表 4 的实验结果证明了 SSUAD 不仅能有效地利用标记数据实现对不同异常模式的精确识别, 还能够处理不同异常比例的数据集, 展示出了良好的稳健性和适应性.

表 5 展示了本文提出的数据增强方法和传统的数据增强方法在数据集 UNSW-NB15 上的 AUC-ROC 和 AUC-PR 的结果. 从表 5 中可以看出本文提出的增强方法相较于其他方法在这 2 项指标上都获得了明显的提升. 具体地, SMOTE 通过在少数类样本之间生成新的合成样本, 但是生成的合成样本并不总是代表真实的数据模式; Borderline-SMOTE[27] 是在 SMOTE 基础上改进的过采样方法, 该方法仅使用边界上的少数样本来合成新样本. Kmeans-SMOTE[28] 先对样本进行聚类操作, 然后根据簇密度的大小分别对不同簇的样本进行合成. Mixup 通过随机地结合来自 2 个原始样本的特征以生成新的训练样本. 生成的数据可能超出实际数据的分布范围, 且可能引入额外的噪声.

Table 5 Comparative Results of Data Augmentation Methods

表 5 数据增强方法对比结果

数据增强方法	AUC-ROC/%	AUC-PR/%
SMOTE	90.2	81.5
Borderline-SMOTE	92.6	82.0
Kmeans-SMOTE	93.8	84.6
Mixup	93.1	82.8
SSUAD（本文）	**94.7**	**86.3**

注：黑体数值表示最优结果.

表 5 中的数据增强方法主要聚焦于生成新的样本点, 但这些方法不直接对特定特征的重要性进行考虑. 在高维数据中, 不是所有特征对分类或异常检测等任务都同等重要. 而本文的数据增强方法利用树结构来评估特征的重要性, 并基于这些评分进行数据增强, 确保了生成的样本更具代表性, 更贴近真实数据分布. 此外, SMOTE 及其变体方法可能会产生距离原始数据分布较远的合成样本, 这可能导致模型的过度拟合或引入噪声. 而基于树的特征选择方法可以有效减少生成样本的噪声, 并减少过度拟合的可能性.

3.5.2 消融实验

为了验证模型中各部分的有效性, 本节使用不同的模型框架在 UNSW-NB15 数据集上进行消融实验, 实验参数设置同 3.4 节. 不同的模型框架为:

1)SSUAD(without vime). 未使用自监督学习框架, 没有使用预设任务优化编码器.

2)SSUAD(without augmentation). 没有使用数据增强, 只使用原始的标记异常训练模型.

3)SSUAD(without detector). 没有使用未知异常检测器, 包括删除其对应的损失函数. 模型会将未标记数据集中的样本全部当作已知类别进行分类, 其余同本文模型.

4)SSUAD(without soft loss). 没有使用软一致性损失函数, 其余同本文模型.

5)SSUAD(without fixmatch loss). 没有使用损失函数 FixMatch, 其余同本文模型.

6) SSUAD 模型. 本文最终模型.

实验结果如表 6 所示.

Table 6 Our Ablation Study

表 6 本文消融实验

模型	AUC-ROC/%	AUC-PR/%
SSUAD（without vime）	92.1	81.5
SSUAD（without augmentation）	92.7	82.3
SSUAD（without detector）	91.8	79.8
SSUAD（without soft loss）	93.2	82.7
SSUAD（without fixmatch loss）	94.1	83.6
SSUAD（本文）	**94.7**	**86.4**

注：黑体数值表示最优结果.

SSUAD(without detector)的各项指标均不如其他模型, 未知异常检测器的目的是检测那些不属于已知类别的样本. 没有未知异常检测器, 模型会默认地将所有输入样本分类给已知的类别, 即使这些输入可能完全不符合任何已知模式, 这会导致模型对未知类别的误报. SSUAD(without vime)的各项指标比含

有自监督模块的模型低，没有使用自监督学习框架，可能导致模型不能充分利用大量的未标记数据来探索数据中的潜在结构和关系，从而无法捕获到一些重要、但在标记数据中不明显的模式. 这也验证了自监督学习能够帮助模型学习更好的特征表示. 当删除数据增强这一部分时，SSUAD(without augmentation)的性能会下降，数据增强不仅缓解类别不平衡的问题，还能够为模型提供更加丰富和多样化的数据视角. 没有数据增强的模型可能没有很好的泛化能力. 此外，SSUAD(without soft loss)由于缺乏软一致性损失函数，导致模型的决策边界可能对数据变化过于敏感. SSUAD(without fixmatch loss)则无法充分利用未标记数据所提供的潜在信息进行训练. 然而，二者的所有指标均不如最终 SSUAD，这进一步证明了软一致性损失函数和 FixMatch 损失函数能够提升模型检测的综合性能.

表 6 的实验结果表明，SSUAD 中的各模块对模型的性能提升都有正面的作用.

3.5.3 敏感性分析

本节实验对损坏样本生成器中样本所要扰动的特征数p_m取不同值进行实验，以验证参数对模型的影响. 实验将p_m的取值区间设定为 [0,1]，在 UNSW-NB15 数据集上进行实验，每组实验运行 5 次，取结果的平均值. 结果如图 2 所示.

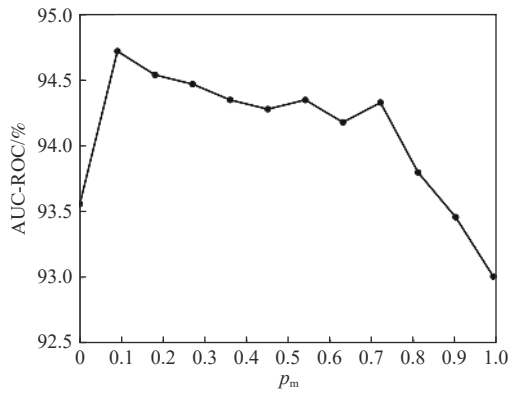

Fig. 2 Sensitivity analysis of hyperparameter p_m

图 2 超参数p_m敏感性分析

从图 2 可以看出，当p_m较大时，意味着有更多的特征被掩码或损坏，这可能使得编码器、掩码向量估计器和特征向量估计器的组合面临更大的挑战，因为它们必须基于部分损失的信息来解决预设任务. 在这种情况下，模型可能会在学习复杂特征表示时遇到困难，从而导致预设任务的解决效果不理想. 另一方面，如果p_m较小，那么被掩码或损坏的特征较少，这使得这 3 个网络可以较为轻松地解决预设任务. 然而，这可能会带来另一个问题，即模型得到的表征可能不够丰富或不具有足够的信息量. 原因在于，模型可以借助大量完整的特征来解决预设任务，不需要过多学习从部分或损坏的信息中提取有用的特征，这可能会导致最后得到的表征相对简单，对后续的任务并无太大的帮助.

综上所述，对于损坏特征的比例p_m的选择，需要在 2 个方面之间寻找平衡. 一方面，应当有足够多的特征被掩码或损坏，以便推动模型学习更复杂、更具有信息量的特征表征. 另一方面，又不能有过多的特征被掩码或损坏，以免使得预设任务变得过于困难，影响模型的学习效率和效果. 因此，选择合适的p_m值是至关重要的，应该根据具体的任务和数据集的特点来灵活确定. 在 UNSW-NB15 实验中，本文选择了$p_m \in [0,1]$的范围，并通过交叉验证的方式确定了最优的p_m值.

此外，本文也对式(3)中的平衡因子λ_{mask}进行相关实验，λ_{mask}主要用于调整预测特征值和预测掩码位置这 2 个前置任务之间的平衡，实验将λ_{mask}的取值区间设定为 [0.1,10]，在 UNSW-NB15 数据集上进行实验. 结果如图 3 所示.

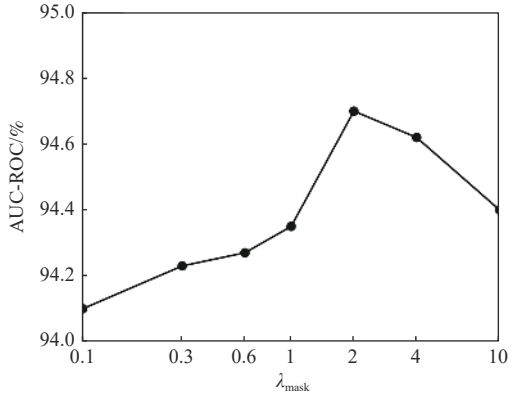

Fig. 3 Sensitivity analysis of hyperparameter λ_{mask}

图 3 超参数λ_{mask}敏感性分析

从图 3 可以看出，如果λ_{mask}过小或过大，即过度强调一个任务就会导致另一个任务的性能下降，从而影响整个模型的性能平衡和准确性. 在λ_{mask}=2 时，模型的整体性能最好，如果继续增加λ_{mask}则会导致模型过度优化掩码向量而忽略了重构误差的准确性.

3.5.4 鲁棒性分析

在本文研究的场景中，标记异常数量的变动对模型鲁棒性的测试尤为关键. 当标记异常样本数量

增加时，需要观察模型的性能是否能保持稳定. 为了验证模型在不同数量已知异常样本的训练集下的鲁棒性，本节进行了以下实验. 通过对 UNSW-NB15 数据集的标记异常样本进行变化，每类标记异常样本的个数从 20 增至 60，再增至 100，比较在拥有不同数量的标记异常样本情况下与其他模型在异常检测能力上的差异.

图 4 展示了所有异常检测模型在不同数量标记异常样本下 AUC-PR 的结果. 可以明显看到除了 iForest 和 ADOA，其他异常检测模型性能都随着标记异常样本数量的增加而上升，这表明更多的标记异常样本可以提高模型的异常检测能力. 值得注意的是，在不同数量的已知异常的情况下，SSUAD 模型在 AUC-PR 性能上的表现始终高于其他异常检测模型. 这是因为 SSUAD 通过预测掩码位置和对应的特征，成功捕捉了正常样本和异常样本之间的差异性，并且有效的数据增强策略大大提高了标记数据的利用效率，为模型的训练提供了更多有价值的信息，从而在整体中增强了模型的异常检测能力.

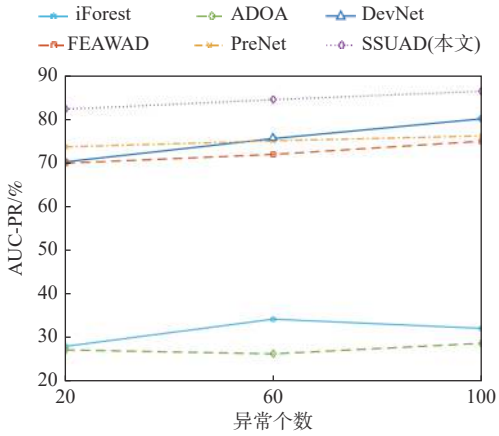

Fig. 4 Robustness experiment of anomaly count

图 4 异常个数鲁棒性实验

在未标记的训练数据集中，异常污染的存在是一个普遍现象. 为了深入探讨 SSUAD 在应对不同污染率的鲁棒性，本文将 SSUAD 与半监督基线模型进行了对比实验，在不同污染率下的 AUC-PR 结果如图 5 所示. 从图 5 可以看出，尽管存在不同比例的异常污染，但是 SSUAD 始终优于其他基线模型. 随着污染率的继续上升，SSUAD 表现出良好的稳定性；而基线模型仅关注已知异常，使得它们无法在未标记训练数据中识别未知异常，导致污染率增加时性能明显下降.

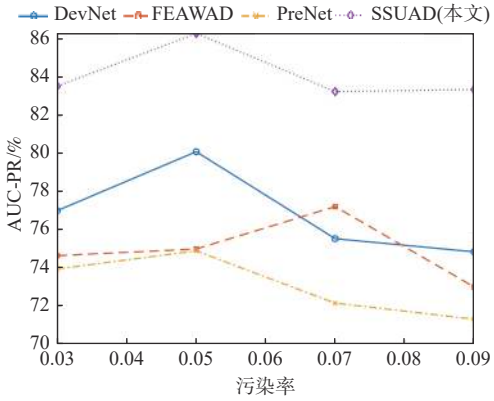

Fig. 5 Robustness experiment of contamination rate

图 5 污染率鲁棒性实验

4 结 论

本文针对现有的半监督异常检测方法在识别未知异常方面的不足以及表格数据的特性，提出了一种基于半监督学习的未知异常检测方法(SSUAD). 该方法首先利用自监督学习，从损坏的输入数据中抽取出有效的特征表示. 接着，通过数据增强来解决在异常检测方法中正常数据与异常数据间的极度不平衡问题. 最后，训练了闭集分类器和未知异常检测器用于区分已知异常与未知异常. 经过系列对比实验和鲁棒性实验，验证了 SSUAD 在有效性和实用性方面的优势.

虽然 SSUAD 已经表现出较好的性能，但仍然存在一些问题和待改进之处. 闭集分类器的准确性和伪标签的阈值选择会影响异常检测的识别率. 因此未来的工作将集中于优化分类器的结构和训练策略以及研究自适应的伪标签阈值选择方法，以进一步增强其对未知异常的检测能力.

作者贡献声明：程渝栋负责完成实验并撰写论文；周昉提出指导意见并修改论文.

参 考 文 献

[1] Dal Pozzolo A, Boracchi G, Caelen O, et al. Credit card fraud detection: A realistic modeling and a novel learning strategy[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 29(8): 3784−3797

[2] Liao H J, Lin C H R, Lin Y C, et al. Intrusion detection system: A comprehensive review[J]. Journal of Network and Computer Applications, 2013, 36(1): 16−24

[3] Fernandes G, Rodrigues J J P C, Carvalho L F, et al. A comprehensive survey on network anomaly detection[J]. Telecommunication

Systems, 2019, 70(3): 447−489
[4] Pang Guansong, Shen Chunhua, Cao Longbing, et al. Deep learning for anomaly detection: A review[J]. ACM Computing Surveys, 2021, 54(2): 1−38
[5] Ding Kaize, Zhou Qinghai, Tong Hanghang, et al. Few-shot network anomaly detection via cross-network meta-learning[C]// Proc of the 30th Int Conf on World Wide Web. New York: ACM, 2021: 2448−2456
[6] 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 Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2018: 2041−2050
[7] Huang Junkai, Fang Chaowei, Chen Weikai, et al. Trash to treasure: Harvesting OOD data with cross-modal matching for open-set semi-supervised learning[C]// Proc of the 18th IEEE/CVF Int Conf on Computer Vision(ICCV). Piscataway, NJ: IEEE, 2021: 8310−8319
[8] Li C L , Sohn K, Yoon J, et al. Cutpaste: Self-supervised learning for anomaly detection and localization[C]//Proc of the IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 9664−9674
[9] Meng Debin, Peng Xiaojiang, Wang Kai, et al. Frame attention networks for facial expression recognition in videos[C]//Proc of the 28th IEEE Int Conf on Image Processing (ICIP). Piscataway, NJ: IEEE, 2019: 3866−3870
[10] Campos G O, Zimek A, Sander J, et al. On the evaluation of unsupervised outlier detection: Measures, datasets, and an empirical study[J]. Data Mining and Knowledge Discovery, 2016, 30(4): 891−927
[11] Li K L, Huang H K, Tian S F, et al. Improving one-class SVM for anomaly detection[C]//Proc of the 2003 Int Conf on Machine Learning and Cybernetics. Piscataway, NJ: IEEE, 2003: 3077–3081
[12] Pang Guansong, Shen Chunhua, Van Den Hengel A. Deep anomaly detection with deviation networks[C]//Proc of the 25th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2019: 353−362
[13] Pang Guansong, Shen Chunhua, Jin Huidong, et al. Deep weakly-supervised anomaly detection[C]//Proc of the 29th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2023: 1795−1807
[14] Zhang Yalin, Li Longfei, Zhou Jun, et al. Anomaly detection with partially observed anomalies[C]//Proc of the 27th Int Conf on World Wide Web. New York: ACM, 2018: 639−646
[15] Ruff L, Vandermeulen R A, Görnitz N, et al. Deep semi supervised anomaly detection[C/OL]//Proc of the 8th Int Conf on Learning Representations. 2020[2023-06-11]. https://openreview.net/pdf?id=HkgH0TEYwH
[16] Zhou Yingjie, Song Xucheng, Zhang Yanru, et al. Feature encoding with autoencoders for weakly supervised anomaly detection[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(6): 2454−2465
[17] Li Zhe, Sun Chunhua, Liu Chunli, et al. Dual-MGAN: An efficient approach for semi-supervised outlier detection with few identified anomalies[J]. ACM Transactions on Knowledge Discovery from Data, 2022, 16(6): 1−30
[18] Zong Weixian, Zhou Fang, Pavlovski M, et al. Peripheral instance augmentation for end-to-end anomaly detection using weighted adversarial learning[C]//Proc of the 27th Int Conf on Database Systems for Advanced Applications. Berlin: Springer, 2022: 506−522
[19] Chawla N V, Bowyer K W, Hall L O, et al. SMOTE: Synthetic minority over-sampling technique[J]. Journal of Artificial Intelligence Research, 2002, 16(1): 321−357
[20] Zhang Hongyi, Cisse M, Dauphin Y N, et al. Mixup: Beyond empirical risk minimization[C/OL]//Proc of the 6th Int Conf on learning Representations. 2018[2023-05-30]. https://openreview.net/pdf?id=r1Ddp1-Rb
[21] Chen Yanbei, Zhu Xiatian, Li Wei, et al. Semi-supervised learning under class distribution mismatch[C]//Proc of the 14th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2020: 3569−3576
[22] Yu Qing, Ikami D, Irie G, et al. Multi-task curriculum framework for open-set semi-supervised learning[C]//Proc of the 16th European Conf on Computer Vision (ECCV 2020). Berlin: Springer, 2020: 438−454
[23] Huang Zhuo, Yang Jian, Gong Chen. They are not completely useless: Towards recycling transferable unlabeled data for class-mismatched semi-supervised learning[J]. IEEE Transactions on Multimedia, 2022, 25: 1844−1857
[24] Moustafa N, Slay J. UNSW-NB15: A comprehensive data set for network intrusion detection systems[C/OL]//Proc of the Conf on Military Communications and Information Systems Conf. 2015[2023-07-09]. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7348942
[25] Tavallaee M, Bagheri E, Lu Wei, et al. A detailed analysis of the KDD CUP 99 data set[C/OL]//Proc of the IEEE Symp on Computational Intelligence for Security and Defense Applications. Piscataway, NJ: IEEE, 2009[2023-04-20]. https://ieeexplore.ieee.org/abstract/document/5356528
[26] Liu F T, Ting K M, Zhou Zhihua. Isolation-based anomaly detection[J]. ACM Transactions on Knowledge Discovery from Data, 2012, 6(1): 1−39
[27] Han Hui, Wang Wenyuan, Mao Binghuan. Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning[C]//Proc of the Int Conf on Intelligent Computing. Berlin: Springer, 2005: 878−887
[28] Douzas G, Bacao F, Last F. Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE[J]. Information Sciences, 2018, 465: 1−20

Cheng Yudong, born in 1998. Master. His main research interest includes anomaly detection.
程渝栋, 1998 年生. 硕士. 主要研究方向为异常检测.

Zhou Fang, born in 1983. PhD, associate professor. Her main research interests include data mining and machine learning.
周　昉, 1983 年生. 博士, 副教授. 主要研究方向为数据挖掘、机器学习.