

DOI:10.19651/j.cnki.emt.2519435

# 基于半非负矩阵分解神经网络的异常检测方法<sup>\*</sup>

韩威<sup>1,2</sup> 吴黎明<sup>2</sup>

(1. 广州铁路职业技术学院电气工程学院 广州 511300; 2. 广东工业大学机电工程学院 广州 510006)

**摘要:** 异常检测旨在识别数据中的异常,在工业检测等领域中具有显著的应用价值。目前主流异常检测方法采用的是自编码器等无监督模型,这类模型采用全连接层或卷积层进行编码、解码的数据处理过程,会导致模型面临缺乏解释性、语义误差等问题。为此,本文提出将半非负矩阵分解模型和神经网络训练方式相结合,设计了一个用于异常检测的半非负矩阵分解神经网络。由于半非负矩阵分解模型具有“局部叠加构成整体”的特性,该网络能更好的保留语义信息,且具有可解释性。此外,该网络的特征矩阵作为权值随着网络训练而更新,这能有效解决传统半非负矩阵分解模型存在的局部最优解问题。在三个数据集上,测试了该网络的异常检测性能,实验结果表明其在应对连续数据时,比主流自编码器和变分自编码器方法的检测指标高3%,在离散数据上也取得了不弱于主流方法的效果;与基于传统半非负矩阵分解模型的异常检测方法相比,该网络在所有检测指标上均有提升,升幅最高达12%。该网络是利用传统矩阵分解模型构建神经网络的有益探索,能有效解决异常检测问题。

**关键词:** 异常检测;无监督学习;半非负矩阵分解;神经网络

中图分类号: TP274; TN98 文献标识码: A 国家标准学科分类代码: 510.4030

## Anomaly detection based on semi-nonnegative matrix factorization neural network

Han Wei<sup>1,2</sup> Wu Liming<sup>2</sup>

(1. School of Electrical Engineering, Guangzhou Railway Polytechnic, Guangzhou 511300, China;

2. School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China)

**Abstract:** Anomaly detection aims to identify abnormal patterns in data, and it has significant application value in various fields such as industrial inspection. The current mainstream anomaly detection methods employ unsupervised models such as auto-encoder. These models use fully connected layers or convolutional layers for the data processing during encoding and decoding, which can lead to problems such as lack of interpretability and semantic errors. This paper proposes a combination method of the semi-nonnegative matrix factorization model and network training to design a semi-nonnegative matrix factorization neural network for anomaly detection. Due to the characteristic of the semi-nonnegative matrix factorization model that “local superposition constitutes the whole”, this network can better preserve semantic information and is also interpretable. Additionally, the feature matrix of this network is updated as weights during the training of the network, which effectively solves the problem of local optimal solutions existed in the traditional semi-nonnegative matrix factorization model. The anomaly detection performance of this network was tested on three datasets. The experiment results show that it outperforms mainstream auto-encoder and variational auto-encoder methods by more than 3 percentage points in continuous data, and achieves comparable results in discrete data. Compared with the detection method based on the traditional semi-nonnegative matrix factorization model, this network has significantly improved in all detection metrics, with the highest improvement reaching 12%. This network is a beneficial exploration that utilizes traditional matrix factorization model to construct neural network, and it can effectively solve the anomaly detection problem.

**Keywords:** anomaly detection;unsupervised learning;semi-nonnegative matrix factorization;neural network

## 0 引言

很多实际场景存在正负样本不均衡问题,例如判断产品质量是否合格、设备状态是否正常、计算机系统是否被入侵等。此问题的一般表现为:能采集到的正常样本比异常样本数量显著多,异常类别多且异常样本不易大量采集。

收稿日期:2025-07-24

\* 基金项目:广东省自然科学基金面上项目(2022A1515011559)、广州市教育局高校科研项目(202234875)、广州市高等教育教学质量与教学改革工程项目(2024CJRHD001)资助

目前异常检测是解决正负样本不均衡问题的常用方法,它旨在识别数据中的异常<sup>[1]</sup>。异常检测在入侵检测<sup>[2]</sup>、产品质量检测<sup>[3]</sup>、系统状态监测<sup>[4]</sup>等领域获得了很强的研究兴趣,表现出良好的应用前景。

实现异常检测功能的传统机器学习方法主要采用聚类<sup>[5]</sup>、单类支持向量机<sup>[6]</sup>、孤立森林<sup>[7]</sup>等离群点检测模型,这些模型通常分为两个步骤:首先从数据中提取低维特征,然后将低维特征输入检测模型得到异常判别结果。然而,基于数据降维的低维特征提取过程一般是根据人工经验设计,有些数据的高维性和动态性使得设计合适特征变得极为困难且耗时<sup>[8]</sup>。

近年来,随着深度学习方法在自动学习原始数据的特征方面具有良好的表现,其在异常检测任务中的应用也越来越广泛,相关大多数研究集中在基于编码器-解码器框架的无监督学习模型<sup>[9]</sup>上。这些模型的结构一般为:先使用编码器学习数据的低维特征,然后使用解码器重构数据,其训练及应用过程为:首先使用正常数据训练模型,然后将待检数据输入模型,并得到数据重构误差,最后根据误差值是否超过一个阈值做出异常判别。目前相关主流方法有自编码器(auto-encoder, AE)<sup>[10-11]</sup>、变分自编码器(variational auto-encoder, VAE)<sup>[12-13]</sup>等。在 AE 模型中,编码器一般是多个全连接层或者卷积层,用于将原始数据逐渐降维至低维特征,解码器与编码器相反,将低维特征逐渐升维以重构原始数据,最终根据重构误差来发现异常。VAE 模型引入概率建模,通过编码器输出低维特征的分布参数(均值和方差),然后从该分布重构数据,当输入异常样本时,其重构误差或概率似然会偏离正常范围,从而被标记为异常。AE 模型的损失函数本质上是逐个数据元素的误差,会导致语义重构问题<sup>[14]</sup>,即数据重构得很好,但语义已丢失。VAE 模型将重构精度、变分正则项作为优化目标,会导致牺牲重构精度以换取平滑的低维特征空间,也面临语义丢失问题。此外,AE 和 VAE 作为深度学习模型,也缺乏可解释性。

矩阵分解可对原始数据进行基于部分的表示,是一种有效的特征提取和数据重构方法。常用矩阵分解模型有非负矩阵分解(nonnegative matrix factorization, NMF)<sup>[15]</sup>和半非负矩阵分解(semi-nonnegative matrix factorization, Semi-NMF)<sup>[16]</sup>。NMF 是将非负的数据分解为特征矩阵和系数矩阵,并限定两个分解因子也是非负的。Semi-NMF 在具有 NMF 的优点的同时,还放宽了 NMF 的非负性约束,允许数据矩阵和特征矩阵具有混合符号。近年来,有一些研究将矩阵分解用于解决异常检测任务,其中大多是将其用于数据降维,即提取原始数据的特征,然后将低维特征输入聚类<sup>[5]</sup>、单类支持向量机<sup>[6]</sup>等模型进行异常识别。此外,也有个别研究利用 NMF 的重构误差来进行异常判别<sup>[17]</sup>。然而传统 NMF 模型和 Semi-NMF 模型的优化目标函数都是非凸的,会导致矩阵分解因子容易陷入局部最优解<sup>[16]</sup>,从而影响异常检测效果。

针对上述情况,本文将神经网络训练方式引入 Semi-NMF 模型,设计了一个适用于异常检测任务的半非负矩阵分解神经网络(semi-nonnegative matrix factorization neural network, SNMFNN)。SNMFNN 将数据矩阵的重构误差作为损失函数,将 Semi-NMF 的特征矩阵作为网络权值。之所以选择 Semi-NMF 构建此网络,是因为其允许数据矩阵和特征矩阵具有混合符号,这对原始数据具有很好的适用性;此外在网络训练过程中,权值需可正可负,其特征矩阵也符合这一条件。本文详细介绍了 SNMFNN 的结构及应用方法,通过与传统 Semi-NMF 的数据重构效果对比,说明了其合理性,并在三个数据集上评估和探讨了它的异常检测性能。

## 1 半非负矩阵分解神经网络

### 1.1 半非负矩阵分解

本文设计的 SNMFNN 选用 Semi-NMF 进行数据降维以获得低维特征。Semi-NMF 的基本原理为<sup>[16]</sup>:假设数据  $\mathbf{X}$  是  $n \times p$  的矩阵,即是  $n$  个样本数据的集合,每一个样本都有  $p$  维特征;在满足一定约束条件下,将数据矩阵  $\mathbf{X}$  分解为两个因子,即特征矩阵  $\mathbf{F}$  ( $\mathbf{F} \in \mathbb{R}^{k \times p}$ ) 和系数矩阵  $\mathbf{G}$  ( $\mathbf{G} \in \mathbb{R}^{n \times k}$ ),并使得数据矩阵的重构误差最小。Semi-NMF 允许数据矩阵  $\mathbf{X}$  和特征矩阵  $\mathbf{F}$  的元素有混合符号,同时限制系数矩阵  $\mathbf{G}$  仅由非负元素组成,其近似于以下因式分解:

$$\mathbf{X} \approx \mathbf{GF} \quad (1)$$

优化 Semi-NMF 的两个分解因子所采用的目标函数为:

$$\min \Gamma(\mathbf{G}, \mathbf{F}) = \|\mathbf{X} - \mathbf{GF}\|_{\text{F}}^2, \text{s. t. } \mathbf{G} \geqslant 0 \quad (2)$$

基于乘法规则,按式(3)和式(4)对  $\mathbf{F}$  和  $\mathbf{G}$  进行迭代优化。采用的迭代规则为:两个因子交替迭代更新,即在更新其中一个因子的同时,固定另一个因子。

$$\mathbf{F} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{X} \quad (3)$$

$$\mathbf{G} = \mathbf{G} \sqrt{\frac{(\mathbf{X}\mathbf{F}^T)^{\text{pos}} + \mathbf{G}(\mathbf{F}\mathbf{F}^T)^{\text{neg}}}{(\mathbf{X}\mathbf{F}^T)^{\text{neg}} + \mathbf{G}(\mathbf{F}\mathbf{F}^T)^{\text{pos}}}} \quad (4)$$

式中:符号  $\mathbf{A}^{\text{pos}}$  表示矩阵  $\mathbf{A}$  的负元素被 0 替换,符号  $\mathbf{A}^{\text{neg}}$  表示矩阵  $\mathbf{A}$  的正元素被 0 替换,可用式(5)所示的定义进行表述。

$$\forall i, j, \quad \mathbf{A}^{\text{pos}} = \frac{|\mathbf{A}_{i,j}| + \mathbf{A}_{i,j}}{2}, \quad \mathbf{A}^{\text{neg}} = \frac{|\mathbf{A}_{i,j}| - \mathbf{A}_{i,j}}{2} \quad (5)$$

### 1.2 网络结构、损失函数及异常检测应用

SNMFNN 是将基于 Semi-NMF 的低维特征提取过程和数据重构过程进行直接连结,其结构包括输入层、Semi-NMF 层、重构层,如图 1 所示。

输入层用于接受样本数据矩阵  $\mathbf{X}$ ,接着由 Semi-NMF 层在固定特征矩阵  $\mathbf{F}$  的情况下对数据矩阵进行半非负矩阵分解,通过对式(4)进行迭代运算以获得系数矩阵  $\mathbf{G}$ ,

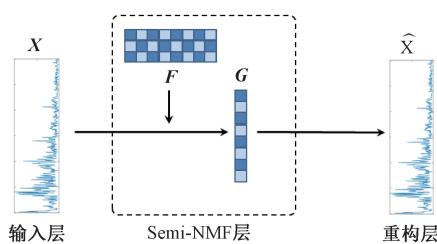


图 1 SNMFNN 的结构

Fig. 1 Structure of the SNMFNN

Semi-NMF 层的运算过程如式(6)所示。

$$\mathbf{G} = f^{(m)}(\mathbf{G}) \quad (6)$$

其中,  $f(\mathbf{G})$  为式(4),  $m$  为迭代次数。然后在重构层,根据特征矩阵和系数矩阵按式(7)对数据矩阵进行重构,得到重构后的数据矩阵  $\hat{\mathbf{X}}$ 。

$$\hat{\mathbf{X}} = \mathbf{G}\mathbf{F} \quad (7)$$

Semi-NMF 的分解残差即是训练 SNMFNN 的损失函数,如式(8)所示。

$$\Omega = \frac{1}{n} \sum_{i,j} (\mathbf{X} - \hat{\mathbf{X}})_{i,j}^2 \quad (8)$$

其中,  $n$  是每次训练的样本数,  $1 \leq i \leq n, 1 \leq j \leq p$ 。

在 SNMFNN 的训练过程中,仅有特征矩阵  $\mathbf{F}$  是需要训练的网络权值,  $\mathbf{F}$  中元素的初始值在区间  $(-1, 1)$  范围,系数矩阵  $\mathbf{G}$  中元素的初始值在区间  $(0, 1)$  范围。

在异常检测应用中,先用正常数据对 SNMFNN 进行训

练,让 SNMFNN 记住正常数据的特征,以得到 SNMFNN 模型;然后将待检数据输入该模型,得到数据重构误差,误差值若高于阈值,则待检数据被判为异常,反之则为正常。

### 1.3 网络合理性分析

传统 Semi-NMF 是采用乘法迭代的方式更新分解因子(记为传统 SNMF),而本文是通过网络训练的方式更新 Semi-NMF 的分解因子(记为网络 SNMF)。将传统 SNMF 与本文提出的网络 SNMF 进行对比,以说明 SNMFNN 的合理性。让传统 SNMF 和网络 SNMF 分别对相同的图像进行分解和重构,通过原始图像和重构图像的相似度来对比二者的分解效果。本文选用 MSE(mean squared error)、PSNR (peak signal-to-noise ratio)、SSIM (structural similarity index) 等图像质量评价的典型指标<sup>[18]</sup>,用来衡量图像相似度。MSE 是两幅图像的均方误差,此值越小,说明图像越相似;PSNR 即峰值信噪比,基于 MSE 指标,衡量信号最大功率与噪声功率的比值,单位是 dB,此值越高,说明图像越相似,一般达到 25 dB 以上即认为图像接近;SSIM 从亮度、对比度、图像结构三方面综合评估图像相似性,此值越接近 1,表示图像越相似。

本文选择动物图像(像素值均为非负数据)和 Mel 频率倒谱图(像素值有正有负,即为半非负数据)进行相似性评价实验。设置相同的矩阵分解初始值,分别利用传统 SNMF 和网络 SNMF 对动物图像(大小为  $1024 \times 1024$  像素)进行分解和重构,原始图像和重构图像如图 2 所示,图像相似性评价参数的计算结果如表 1 所示。



(a) 原始图像

(a) The original image

(b) 传统SNMF的重构图像

(b) The image reconstructed by the traditional SNMF

(c) 网络SNMF的重构图像

(c) The image reconstructed by the SNMF network

图 2 图像重构效果

Fig. 2 Reconstruction effect of the image

表 1 图像重构后的相似性参数值

Table 1 Values of the image similarity parameters

方法	MSE	PSNR	SSIM
传统 SNMF	$4.31 \times 10^{-4}$	33.65 db	0.829
网络 SNMF	$5.41 \times 10^{-4}$	32.66 db	0.816

Slider 数据集的声学特征,每张图的大小是  $309 \times 320$ ,从中随机选择 100 幅图。按照相同的矩阵分解初始值,将这些图输入传统 SNMF 和网络 SNMF 分别进行分解和重构,原始图像和重构图像的相似性评价参数 MSE、PSNR、SSIM 的计算结果如图 3 所示;三个参数的均值及方差如表 2 所示,其中符号“ $\pm$ ”的前者为均值、后者为标准差。

从相似性评价实验结果可以看出,本文设计的网络

用于实验的 Mel 频率倒谱图来源于 2.1 节介绍的

表 2 100 幅倒谱图重构后的相似性参数值

Table 2 Similarity parameters of 100 reconstructed spectrograms

方法	MSE	PSNR	SSIM
传统 SNMF	$8.77 \times 10^{-4} \pm 3.70 \times 10^{-4}$	$30.87 \pm 1.54$ db	$0.9993 \pm 5.30 \times 10^{-4}$
网络 SNMF	$0.001 \pm 2.92 \times 10^{-4}$	$30.17 \pm 1.38$ db	$0.9989 \pm 5.40 \times 10^{-4}$

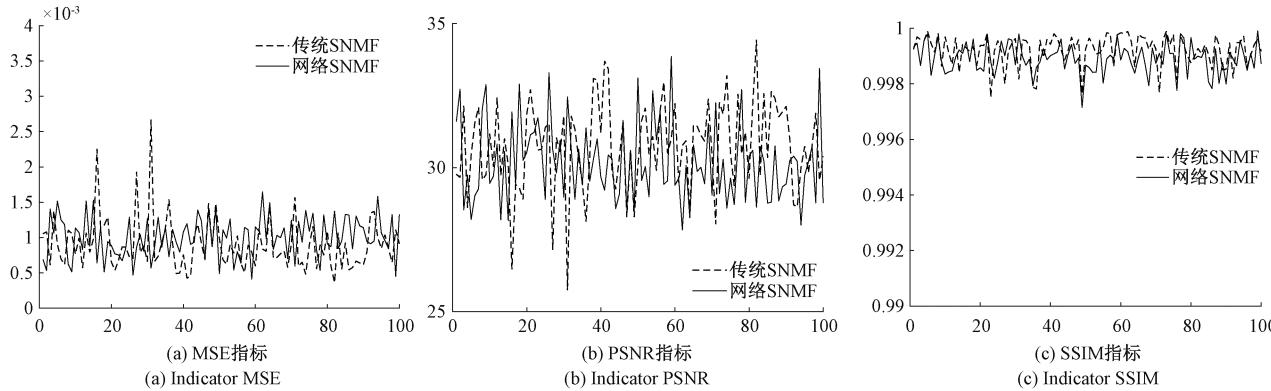


图 3 Mel 频率倒谱图相似性评价参数的计算结果

Fig. 3 Similarity parameters of Mel spectrograms

SNMF 和传统 SNMF 在数据分解和重构方面的效果非常接近,说明网络 SNMF 和传统 SNMF 一样,在数据分解和重构过程中没有引起数据信息失真,进而说明利用网络训练的方式对 Semi-NMF 模型的因子进行更新,与传统 Semi-NMF 模型采用乘法迭代更新分解因子的典型方式趋于一致,进而说明了本文设计的 SNMFNN 具有合理性。

## 2 实验结果

### 2.1 数据集及数据处理

本文在 3 个数据集上对 SNMFNN 的异常检测性能进行评估,数据集及相应数据特征提取过程简要介绍如下。

KDDCup 数据集<sup>[19]</sup>。此数据集是一个计算机网络入侵检测的经典数据集,由美国加州大学欧文分校和加州大学伯克利分校提供,并国际知识发现和数据挖掘竞赛(KDD Cup)中使用。该数据集来源于对美国空军网络流量的模拟,包括正常网络流量和多种类型的入侵流量,用于识别网络入侵行为。每个数据由网络连接的各种特征组成,如连接的持续时间、服务类型、源和目标地址、传输层协议等共计 121 维特征,数据集中的类别标签指示了每个连接是正常连接还是某种类型的入侵行为(包括拒绝服务攻击、欺骗攻击、恶意代码攻击等多种攻击类型)。本文选用 KDDCup 的子集 kddcup.data\_10\_percent。

Slider 数据集<sup>[20]</sup>。该数据集于 2019 年发布,目的是为工业设备异常声音检测研究提供真实工厂环境下的声音数据。由 8 个麦克风组成的均匀圆形阵列采集工业导轨的工作声音,采样率 16 kHz,采样精度 16 bit,本文选择数据集中 Slider02 的声音进行实验,选用第一个麦克风的录音。共有 1 068 个正常声音样本、267 个异常声音样本,每

个声音样本的时长为 10 s,异常情况为轨道损坏、皮带松动、缺少润滑剂等。实验中,训练集选择 801 个正常样本,其余正常样本和异常样本用作测试集。按照此数据集提供方<sup>[20]</sup>的数据处理方式进行声学特征提取,每个样本的特征是一个大小为  $309 \times 320$  的 Mel 频率倒谱图。

Motor 数据集。此数据集是本文作者对国内某空调电机制造企业提供的一批电机样品的运行声音进行采集而得。电机样品包含 46 个正常样品、20 个故障样品,故障情况为转子摩擦故障、电磁故障、轴承故障。对每个样品分别采集了 30 个声音样本,每个样本的时长为 1 s,采样率 48 kHz,采样精度 32 bit。对每个样本提取声学特征的过程为:采用窗长为 25 ms、步长为 10 ms 的滑动窗将时长为 1 秒的声音信号分割为 99 个数据块;按 Mel 刻度将每个数据块划分为 24 个频段,计算每个频段的 Mel 频率倒谱系数,选择前 12 个 MFCCs (mel-frequency cepstrum coefficients, MFCCs)作为特征;将每个数据块的特征按时间顺序连接在一起,可得到一个大小为  $99 \times 12$  的 MFCCs 特征图,即为一个样本的声学特征。随机选择每个正常样品的 15 个样本作为训练集,正常样品的其余样本和异常样品的全部样本作为测试集,因此训练集有 690 个样本、测试集有 1 290 个样本(正常样本 690 个、异常样本 600 个)。

### 2.2 评价指标

选择精确率(Precision)、召回率(Recall)、F1 分数(F1-Score)作为评价指标,计算公式如式(9)~(11)所示。精确率衡量被模型预测为异常的样本中真实异常样本的比例,高精确率可减少因将正常样本误报为异常而导致的冗余维护成本。召回率表示模型对所有真实异常样本的正确识别能力,高召回率意味着漏检率低,对预防突发性异常

至关重要。F1 分数是精确率和召回率的调和平均数,能够反映模型在精确率和召回率之间的平衡程度。此外,为了更全面地测试检测效果,本文还选用 auc 和 p\_auc 指标。

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$F1\text{-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

其中,N 为异常,P 为正常,用大写字母 T 和 F 分别表示样本的预测标签与其实际标签相符或相反,TP 表示被正确预测为异常的样本数,FP 表示被错误预测为异常的样本数(实际为正常),FN 表示被错误预测为正常(实际为异常)的样本数。

### 2.3 对比方法

本文提出的基于 SNMFNN 的异常检测方法,是将数据重构误差值作为异常判别的依据,因此当前流行的同类异常判别方法 AE、VAE 是用于对比的主要基线。此外,SNMFNN 是将传统 Semi-NMF 模型和神经网络训练方式相结合,其中 Semi-NMF 的系数矩阵仍然采用传统乘法迭代方式进行更新,特征矩阵则随着网络训练而更新,这有利于解决传统 Semi-NMF 模型易陷入分解因子是局部最优解的问题。为了凸显此改进的积极作用,本文还设计了基于传统 Semi-NMF 模型的异常检测方法(记为 SNMFA 方法),即利用传统 Semi-NMF 对数据按照式(3)、(4)、(7)进行分解和重构,然后根据重构误差进行待检数据的异常判别。

在实验中,SNMFNN 方法及其对比方法 AE、VAE、SNMFA 在各数据集上的网络结构及超参数设置如下。在 KDDCup 数据集的实验中,所有方法的输入层均有 121 个节点,AE 方法和 VAE 方法的网络结构中的编码过程为:第一个全连接层 40 个节点、激活函数为 tanh,第二个全连接层 5 个节点、线性激活函数,解码过程与编码过程相逆;SNMFNN 方法特征矩阵的维度为 8,式(6)中的迭代次数 m 为 11。对于 Slider 数据集和 Motor 数据集,所有方法的输入层相应分别有 320 个节点、1 188 个节点,AE 方法和 VAE 方法的编码器和解码器结构相同,编码器结构具体为:第一个全连接层 150 个节点、激活函数为 tanh,第二个全连接层 50 个节点、激活函数为 tanh,第三个全连接层 20 个节点、激活函数为 tanh,第四个全连接层 10 个节点、线性激活函数,解码器与编码器的结构相反;SNMFNN 方法中,特征矩阵的维度分别设为 20 和 30,式(6)中的迭代次数 m 为 11。网络训练相关的 minibatch 是 128、learning rate 是 0.001、epoch 是 100。此外,为了更好的对比 SNMFNN 方法和 SNMFA 方法,二者在所有实验中的矩阵维度和初始值均相同。

### 2.4 实验环境与实验结果

本文方法及其对比方法涉及的特征提取、模型训练及

验证代码均在 TensorFlow2.10 和 Python3.11 框架下进行编写,在 Windows 11 系统上运行,各种方法在三个实验数据集上的测试结果如表 3~5 所示。

表 3 KDDCUP 数据集的实验结果

Table 3 Experimental results of the KDDCUP dataset

方法	Precision	Recall	F1-Score	auc	p_auc
AE	0.955	0.965	0.960	0.977	0.967
VAE	0.960	0.970	0.965	0.980	0.972
SNMFA	0.931	0.941	0.936	0.962	0.947
SNMFNN	0.961	0.971	0.966	0.981	0.972

表 4 Slider 数据集的实验结果

Table 4 Experimental results of the Slider dataset

方法	Precision	Recall	F1-Score	auc	p_auc
AE	0.642	0.668	0.655	0.648	0.619
VAE	0.636	0.662	0.649	0.642	0.613
SNMFA	0.656	0.682	0.669	0.662	0.632
SNMFNN	0.681	0.708	0.694	0.688	0.657

表 5 Motor 数据集的实验结果

Table 5 Experimental results of the Motor dataset

方法	Precision	Recall	F1-Score	auc	p_auc
AE	0.821	0.812	0.816	0.829	0.809
VAE	0.820	0.810	0.815	0.827	0.807
SNMFA	0.727	0.718	0.723	0.742	0.728
SNMFNN	0.853	0.843	0.848	0.858	0.849

### 2.5 实验结果讨论

从表 3~5 中的实验结果可以看出,与当前流行的异常检测方法 AE、VAE 相比,SNMFNN 方法在 KDDCUP 数据集上的优势不明显,而在 Slider 数据集和 Motor 数据集上具有明显更优的效果。这是因为:KDDCUP 数据集是网络流量数据,具有离散特性,即使相邻数值也容易存在突变,而 Slider 数据集和 Motor 数据集中的数据是声音数据,具有连续性,且从声音数据中提取的 MFCCs 图(即声音特征),其相邻像素值一般也是连续的;Semi-NMF 模型假设数据是特征矩阵和系数矩阵的线性组合,连续数据容易满足“局部叠加构成整体”的特性,与 AE 方法、VAE 方法相比,SNMFNN 方法会更好的保留语义特征,从而有利于 SNMFNN 方法在声音数据上取得更出色的表现;而离散数据可能不是很符合此假设,且 SNMFNN 中没有 tanh、sigmoid 等具有平滑效果的激活函数,这可能也限制了其用于离散数据时的表现。

在 3 个数据集上,SNMFNN 方法都表现的比 SNMFA 方法明显更为出色。这是因为传统 Semi-NMF 的优化目标函数(如式(2))是非凸的,导致矩阵分解因子容易陷入

局部最优解,从而可能影响 SNMFA 方法的异常检测效果,而 SNMFNN 方法是采用神经网络训练的方式优化矩阵分解因子,虽然网络训练也面临着陷入局部最优解的问题,但是相比于矩阵分解,此问题会明显较弱。

### 3 结 论

本文将网络训练方式引入经典 Semi-NMF 模型,将 Semi-NMF 的分解因子(特征矩阵)由传统上基于乘法迭代更新变为随着网络训练而更新,从而设计了一个用于解决异常检测任务的新网络 SNMFNN,给出了网络结构及其优化目标函数。SNMFNN 和 Semi-NMF 模型对数据的分解和重构效果基本一致,这体现了 SNMFNN 的设计合理性。将 SNMFNN 在三个数据集上进行了异常检测性能测试,实验结果表明,在针对连续性数据时,与当前流行的误差重构法 AE、VAE 相比,它能较明显的提升异常检测效果,这是因为 SNMFNN 能更好的保留语义特征;SNMFNN 能有效缓解 Semi-NMF 模型面临的分解因子局部最优解问题,因此其在所有测试结果上都比 SNMFA 方法更好。此外,由于 Semi-NMF 模型具有“局部叠加构成整体”的可解释性,让 SNMFNN 具有了深度学习神经网络所没有的可解释性。异常检测在很多领域具有显著的应用价值,本文是利用传统矩阵分解模型构建神经网络的有益尝试,未来将进一步将矩阵分解研究成果拓展至深度学习领域,提升异常检测效果。

### 参考文献

- [1] 李艳霞,柴毅,胡友强,等. 不平衡数据分类方法综述[J]. 控制与决策, 2019, 34(4): 673-688.  
LI Y X, CHAI Y, HU Y Q, et al. Review of imbalanced data classification methods [J]. Control and Decision, 2019, 34(4): 673-688.
- [2] WANG S, BALAREZO J F, KANDEEPAN S, et al. Machine learning in network anomaly detection: a survey[J]. IEEE Access, 2021, 9: 152379-152396.
- [3] 张玥,陈锡伟,陈梦丹,等. 基于对比学习生成对抗网络的无监督工业品表面异常检测[J]. 电子测量与仪器学报, 2023, 37(10): 193-201.  
ZHANG Y, CHEN X W, CHEN M D, et al. Unsupervised surface anomaly detection of industrial products based on contrastive learning generative adversarial network [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(10): 193-201.
- [4] ZHANG Y CH, DONG ZH Y, KONG W C, et al. A composite anomaly detection system for data-driven power plant condition monitoring [J]. IEEE Transactions on Industrial Informatics, 2019, 16(7): 4390-4402.
- [5] PU G, WANG L J, SHEN J, et al. A hybrid unsupervised clustering-based anomaly detection method[J]. Tsinghua Science and Technology, 2020, 26(2): 146-153.
- [6] JI Y H, LEE H C. Event-based anomaly detection using a one-class SVM for a hybrid electric vehicle[J]. IEEE Transactions on Vehicular Technology, 2022, 71(6): 6032-6043.
- [7] 魏新园,周京欢,钱牧云,等. 随机森林算法在超声缺陷识别中的应用研究[J]. 电子测量与仪器学报, 2024, 38(5): 47-55.  
WEI X Y, ZHOU J H, QIAN M Y, et al. Research on the application of random forest algorithm in ultrasonic defect recognition[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38 (5): 47-55.
- [8] HAN W, XIE SH L, YANG Z Y, et al. Heart sound classification using the SNMFNet classifier [J]. Physiological Measurement, 2019, 40(10): 105003.
- [9] DARBAN Z Z, WEBB G I, PAN SH R, et al. Deep learning for time series anomaly detection: A survey[J]. ACM Computing Surveys, 2024, 57(1): 1-42.
- [10] NAWAZ A, KHAN S S, AHMAD A. Ensemble of autoencoders for anomaly detection in biomedical data: A narrative review [J]. IEEE Access, 2024, 12: 17273-17289.
- [11] 刘韵婷,李绅科,郭辉,等. 基于神经批采样的轮胎 X 光图像异常检测研究[J]. 电子测量技术, 2023, 46(5): 157-163.  
LIU Y T, LI SH K, GUO H, et al. Research on tire X-ray image anomaly detection based on neural batch sampling [J]. Electronic Measurement Technology, 2023, 46(5): 157-163.
- [12] XIE T M, XU Q F, JIANG C X, et al. A robust anomaly detection model for pumps based on the spectral residual with self-attention variational autoencoder [J]. IEEE Transactions on Industrial Informatics, 2024, 20(6): 9059-9069.
- [13] 李桢煜,宋宇晨,彭喜元,等. 基于对比序列重构的卫星遥测数据异常检测方法[J]. 仪器仪表学报, 2024, 45(4): 17-26.  
LI ZH Y, SONG Y CH, PENG X Y, et al. Contrastive time-series reconstruction method for satellite anomaly detection [J]. Chinese Journal of Scientific Instrument, 2024, 45(4): 17-26.
- [14] CHEN M, SHI X B, ZHANG Y, et al. Deep feature learning for medical image analysis with convolutional autoencoder neural network[J]. IEEE Transactions on

- Big Data, 2017, 7(4): 750-758.
- [15] HE CH B, FEI X, CHENG Q W, et al. A survey of community detection in complex networks using nonnegative matrix factorization [J]. IEEE Transactions on Computational Social Systems, 2022, 9(2): 440-457.
- [16] SHU ZH Q, SUN Y W, TANG J L, et al. Adaptive graph regularized deep semi-nonnegative matrix factorization for data representation [J]. Neural Processing Letters, 2022, 54: 5721-5739.
- [17] 张焱, 蔡有鑫, 王平, 等. 单时频谱非负矩阵编码与解调的特征提取[J]. 仪器仪表学报, 2022, 43(12): 238-247.  
ZHANG Y, CAI Y X, WANG P, et al. Feature extraction using nonnegative matrix coding and demodulation of single spectrogram [J]. Chinese Journal of Scientific Instrument, 2022, 43 (12): 238-247.
- [18] SETIADI D R I M. PSNR vs SSIM: Imperceptibility quality assessment for image steganography[J]. Multimedia Tools and Applications, 2021, 80: 8423-8444.
- [19] SIDDIQUE K, AKHTAR Z, KHAN F A, et al. KDD Cup 99 data sets: a perspective on the role of data sets in network intrusion detection research[J]. Computer, 2019, 52(2): 41-51.
- [20] BOZTAS G, TUNCER T, AYDOGMUS O, et al. A DCSLBP based intelligent machine malfunction detection model using sound signals for industrial automation systems [J]. Computers and Electrical Engineering, 2024, 119: 109541.

### 作者简介

韩威,博士,高级工程师,主要研究方向为智能传感与在线检测。

吴黎明(通信作者),硕士,教授,主要研究方向为机器视觉检测。

E-mail: ghanwei@yeah.net