利用重要性加权损失缓解长尾异常分数分布

Jungi Lee研发部
*ELROILAB Inc.*首尔，大韩民国
ganbbang12@elroilab.com

Jungkwon Kim研发部 *ELROILAB Inc.*首尔，大韩民国
jkkim@elroilab.com

Chi Zhang研发部
*ELROILAB Inc.*首尔，大韩民国
czhang@elroilab.com

Sangmin Kim研发部
ELROILAB Inc. 首尔，大韩民国
smkim@elroilab.com

Kwangsun Yoo 研发部
ELROILAB Inc.
首尔，大韩民国
yks@elroilab.com

Seok-Joo Byun*研发部 *ELROILAB Inc.*首尔，大韩民国
sjbyun@elroilab.com

arXiv:2601.02440v1 [stat.ML] 5 Jan 2026

摘要—**异常检测对于工业应用至关重要，可用于识别罕见和未见过的模式，从而确保系统可靠性。传统模型基于单一正常数据类别进行训练，难以应对真实世界分布：正常数据呈现出多样化模式，导致类别不平衡和长尾异常分数分布（LTD）。这种不平衡会使模型训练产生偏差，并降低检测性能，尤其是对少数类实例的检测性能。为解决这一问题，我们提出了一种专为异常检测设计的新型重要性加权损失。与此前用于分类中长尾分布的方法相比，我们的方法不需要预先了解正常数据类别。相反，我们引入了一种加权损失函数，其中融入了重要性采样，以使异常分数分布与目标高斯分布相匹配，确保正常数据得到平衡表示。在三个基准图像数据集和三个真实世界高光谱成像数据集上进行的大量实验表明，我们的方法能够有效缓解长尾分布导致的偏差。我们的方法将异常检测性能提升了 0.043，凸显了其在真实世界应用中的有效性。**

***Index* 术语——长尾分布，异常检测 tion**

一、引言

异常检测是一门致力于将未见过的数据识别为异常样本的领域。通常，异常检测模型仅使用单一类别的正常数据进行训练，因为异常数据往往稀缺或无法获取。然而，在现实世界的工业场景中，正常数据通常由多种不同的对象或模式构成，而非单一同质类别。这种固有的多样性会导致正常数据中出现类别不平衡，进而形成右偏的异常分数分布，通常称为长尾分布（LTD）。例如，在高光谱异常检测中，数据集 [1]通常包含背景数据和多种不同的正常对象。图 1 使用 t 分布随机邻域嵌入（t-SNE） [2], 展示了这一现象，同时呈现数据及异常分数分布。可视化结果显示，数据形成了多个簇，而非单个紧凑簇，从而导致异常分数分布右偏。一个关键问题是尾部类别（少数类）与头部类别（多数类）的方差存在重叠，这会使模型在训练过程中偏向多数类。因此，模型检测尾部类别中异常样本的能力会显著下降。这一挑战凸显了采用先进方法的必要性，以有效应对类别不平衡，并增强异常检测模型在复杂工业环境中的鲁棒性。

在分类任务中，人们已经提出了许多方法来应对长尾分布带来的挑战。这些方法大致可分为三类 [3]: 类别再平衡、信息增强和模型增强。类别再平衡旨在通过调整训练数据的分布来缓解类别不平衡。常见技术包括根据类别频率进行欠采样 [4] 或过采样 [5], , 以及采用与各类别频率成反比缩放的加权损失函数 [6]。人们提出了混合增强 [7],等数据增强技术，以减轻长尾分布的影响 [8]。这些方法通过生成合成样本来改善少数类的表示。模型增强方法则包括改进模型架构或训练策略。例如，对比学习通过利用数据点之间的关系来强化特征表示 [9]。然而，这些方法主要针对多类分类而设计，且该任务具备多个类别标签的先验知识。在异常检测中，这一假设往往不切实际，因为该任务通常只涉及单一类别的正常数据，缺乏明确的多类信息。因此，将这些技术直接应用于异常检测会面临挑战，需要谨慎调整，以适应该领域独特的约束和特点。

在异常检测中，正常数据中存在多种类别

*通讯作者

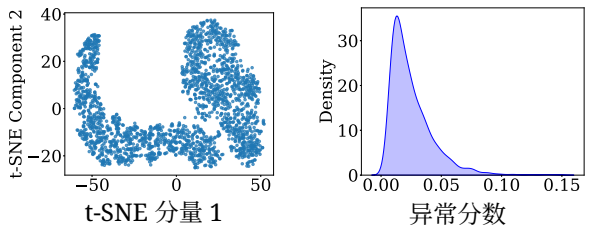

图 1. 高光谱异常检测数据集 [1]分析。我们分析高光谱异常检测数据集，并使用 t-SNE [2] （左）对数据进行可视化。异常分数分布（右）呈现长尾模式，凸显了固有的类别不平衡及其对异常检测的影响。

通常会导致长尾分布（LTD）。与已有成熟解决方案的分类问题不同，由于缺乏明确的多类别信息，在异常检测中解决这一问题要困难得多。为缓解这一问题，一些方法采用生成模型来扩充尾部类别实例，旨在平衡异常分数分布 [10]。此外， [10] 提出了一种利用 k-means 聚类从不平衡数据集中采样的方法。然而，这些方法存在若干局限，例如缺乏明确的尾部实例识别定义、k-means 聚类的限制，以及难以扩展到大型数据集。这些挑战凸显了对更具适应性和可扩展性解决方案的需求，以有效应对异常检测中的长尾分布。

为克服这些挑战，我们提出了一种新颖的重要性加权损失（IWL），无需预先了解正常数据类别的数量或性质即可运行。我们的方法基于这样一个假设：异常分数分布应服从高斯分布。为此，我们利用重要性采样权重，这些权重根据当前异常分数分布与目标高斯分布之间的差异推导得出。这种自适应加权确保模型能够有效平衡尾部类别样本与头部类别样本，从而减轻长尾分布带来的偏差。通过使异常分数分布与目标高斯分布对齐，我们的方法改善了正常数据的表征，并提升了异常检测性能，即使在高度不平衡的数据集上也是如此。

我们作出以下贡献：

- 我们设计了一种专为长尾分布量身定制的异常检测损失函数，无需预先了解正常数据类别的数量或特征。
- 我们提出了一种重要性加权损失，它基于异常分数分布利用重要性采样，确保该分布符合高斯形式，从而提升异常检测性能。
- 我们通过在三个图像数据集和真实世界的高光谱图像数据集上进行大量实验，验证了所提出损失函数的必要性和有效性，证明了其在异常检测中处理长尾分布方面的优越性。

II. 相关工作

为解决长尾分布问题，具有代表性的方法包括欠采样和过采样。一种方法是随机选择头部类别样本，使其数量与尾部类别样本相同；另一种方法是复制尾部类别样本，使其数量与头部类别样本相同 [4], [11]。然而，这些方法可能导致头部类别欠拟合，或因过度复制尾部类别样本而产生过拟合。代价敏感学习通过为各类别分配不同的损失值来重新平衡：对头部类别施加惩罚，或为尾部类别赋予更高权重，以提升泛化性能 [12]–[14]。在迁移学习中，头尾知识迁移用于将信息从头部类别迁移到尾部类别 [15], [16]。近年来，许多研究采用了对比学习 [9] 和两阶段训练 [17] ，在这类方法中，特征提取器和下游任务模型会重新训练。然而，这些方法要么需要预先了解类别信息，要么仅适用于表格数据，因此无法用于通常仅包含单个类别的异常检测领域。

解决异常检测中类别不平衡问题的一种方法是在训练期间为长尾数据分配权重。然而，这种方法会导致构成长尾分布的少数类过拟合，从而降低性能 [18]。使用 k-均值聚类的 [10] 提出了一种针对不平衡问题的采样策略，但该方法依赖于 k 的数量，并且当正常类中的子类较少时会降低性能。使用生成模型进行数据增强可以缓解长尾数据稀缺的问题。在长尾分布中，生成模型反而会降低少数类的性能。文献 [19]通过采用文本衍生特征向量解决了这些问题，但这需要额外的模型。以往的方法需要大量研究或使用新模型，因此难以将其整合到传统的异常检测模型中。

III. 预备知识

考虑一个数据集 $X = \{x_i\}_{i=1}^n \subset \mathbb{R}^{n\times m}$，其中 n是数据的数量， m 是数据的维度。在训练自编码器（AE） $f(\cdot)$ 时，我们最小化误差$\|e_i\|_2^2$，其中 $e_i = x_i - f(x_i) \in \mathbb{R}^m$。

输出分布 最小化均方误差损失（均方误差） $\|e_i\|_2^2$ 等价于最大似然优化问题，该问题可表述如下：

$$\arg\min_{\theta} \sum_{j=1}^{m} ||x_{ij} - f(x_{ij})||_2^2 = \arg\max_{\theta} \prod_{j=1}^{m} P(f(x_{ij})|x_{ij}, \theta). \tag{1}$$

当模型训练得足够充分时，上述最大似然会近似输出分布，使得

$$\max_{\theta} \prod_{j=1}^{m} P(f(x_{ij})|x_{ij},\theta) = P(Y|X) \approx P(Y). \quad (2)$$

在 [18], 中，作者观察到，式（1）中的优化问题可以从Bregman理论 [20], 的角度进行解释，从而得到一个优化后的经验分布，该分布实际上可通过高斯核表示。具体而言，通过应用核密度估计， $P(Y)$ 的优化分布 $\hat{P}(Y)$ 可表示为

$$\hat{P}(Y) = \frac{1}{n}\sum_{j=1}^{m} \mathcal{N}(f(x_{ij}), \sigma^2), \quad (3)$$

这为底层分布提供了平滑估计。

异常分数的分布 利用上述关系，我们可以推导出异常分数的分布。事实上， x 与其重构 $f(x)$之间的残差误差服从高斯分布，即:

$$e_i = x_i - f(x_i) = \left(e_{i1}, e_{i2}, \cdots, e_{im}\right) \sim \mathcal{N}(0, \sigma^2). \quad (4)$$

数据的异常分数 s_i 定义为 e_i 的平方欧几里得范数 x_i ，即

$$s_i = ||x_i - f(x_i)||_2^2 = \sum_{j=1}^{m} e_{ij}^2. \quad (5)$$

因此，我们关注 e_{ij} 的分布，以推断s_i的分布。由于每个 $e_{ij} \sim \mathcal{N}(0,\ \sigma^2)$对于 $1 \le j \le m$而言都是独立的，平方项 e_{ij}^2 服从缩放卡方分布$\sigma^2\chi_1^2$。结合每个 e_{ij} 均相互独立这一假设，异常分数满足

$$\frac{s_i}{\sigma^2} = \frac{1}{\sigma^2}\sum_{j=1}^{m} e_{ij}^2 \sim \chi_m^2, \quad (6)$$

其中， χ_m^2 是自由度为 m的卡方分布。因此，异常分数 s_i 的分布可表示为 $\sigma^2 \cdot \chi_m^2$。当随机变量 $Z \sim \chi_m^2$时，众所周知， $\mathbb{E}(Z) = m$ 且$\mathrm{Var}(Z) = 2m$。因此，我们得到

$$\mathbb{E}[s_i] = \sigma^2 m \quad \text{and} \quad \mathrm{Var}(s_i) = 2\sigma^4 m. \quad (7)$$

根据中心极限定理（CLT），异常分数可表示为求和形式

$$s_i = \sum_{j=1}^{m} e_{ij}^2 \quad (8)$$

当特征数量 m足够大时，会收敛到正态分布。因此，采用均方误差损失函数时，异常分数的分布服从高斯分布。

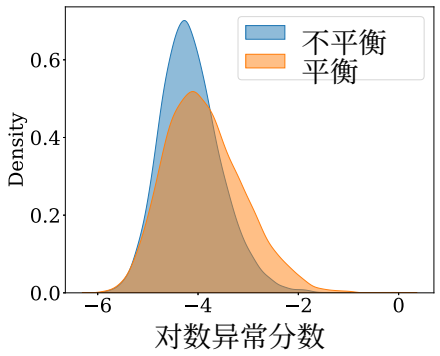

图2。MNIST数据集中对数异常分数的分布。我们分别将类别“6”和“7”指定为多数类和少数类。为构建平衡数据集，为两个类别分配相同数量的样本。相反，在不平衡数据集中，少数类的样本数仅为多数类的十分之一。类别不平衡的引入使偏度从0.504增加到0.524，表明类别不平衡会在异常分数分布中固有地诱导出长尾分布。

IV. 长尾分布损失

我们证明了异常分数分布预期会服从高斯分布。然而，分类问题中的类别不平衡会导致长尾分布，因此，正常类内部的数据不平衡也会使异常检测呈现长尾分布。图2展示了在MNIST数据集上训练得到的异常分数分布，其中类别6和类别7被指定为正常数据。“平衡分布”表示两个类别的数据占比相同，而“不平衡分布”则是将类别6定义为多数类、类别7定义为少数类后得到的分布。由于这种不平衡，对数异常分数分布的偏度从0.504增加到0.524，这表明少数类的存在会导致长尾分布。使用传统的均方误差损失，难以有效地从长尾数据中学习。此外，在训练过程中为尾部类别实例赋予较高权重可能会导致过拟合 [18]。过拟合的发生是因为模型过度关注少数类，而忽略了多数类。这表明，适当调整权重可以平衡过拟合与欠拟合，从而缓解长尾分布带来的挑战。由于过拟合问题源于少数类占主导地位的扭曲分布，我们提出了一种基于正态分布重要性采样的加权损失函数。该方法动态调整权重，确保学习过程更加平衡，同时有效应对长尾分布在异常检测中的局限性。

为实现我们的目标，必须通过适当加权来减弱分布中的极端变化。在这一过程中，根据初步结果，在保留尾部数据原有位置的同时，异常分数分布应呈现轻尾行为。为此，我们必须谨慎调整加权

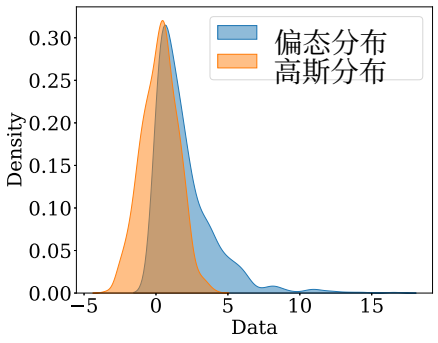

图 3. 对偏态分布应用 Box-Cox 变换以使其正态化，从而得到类似高斯分布的分布。

方案，确保异常分数分布符合轻尾正态分布。这一目标面临两个关键挑战：（1）如何调整权重，使异常分数分布呈现正态形式？（2）如何在保持异常分数固有排序的同时，将其转换为正态分布，确保尾部数据仍处于尾部？为应对这些挑战，我们结合使用重要性采样与 Box-Cox 变换。

当直接估计概率分布 $p(x)$ 面临困难时，可以采用重要性采样；而另一种分布 $q(x)$ 则更容易获得。重要性采样的基本公式如下：

$$\begin{aligned}\mathbb{E}_{p(x)}[f(x)] &= \int f(x)p(x)dx \\ &= \int f(x)p(x)\frac{q(x)}{q(x)}dx \\ &= \mathbb{E}_{q(x)}[f(x)\frac{p(x)}{q(x)}].\end{aligned} \tag{9}$$

请注意，我们的目标是使该分布服从正态分布。因此，我们将概率 $p(x)$ 定义为正态分布，将 $q(x)$ 定义为异常分数的分布。通过将重要性采样用作加权机制，我们可以有效地重塑异常分数分布，使其近似于高斯分布。然而，这种方法带来了下一个挑战：如何从正态分布中提取 $p(x)$，同时保留异常分数的顺序。我们采用 Box-Cox 变换，它可以有效地将偏态分布转换为高斯分布，如图 3 所示。Box-Cox 变换的公式如下：

$$f_b(s_i) = \begin{cases}\frac{s_i^\lambda - 1}{\lambda}, & \text{if } \lambda \neq 0, \\ \ln(s_i), & \text{if } \lambda = 0,\end{cases} \quad \text{for} \quad s_i > 0, \tag{10}$$

其中，s_i 是异常分数，λ 是决定变量所取指数的参数。通过确定 λ 的最优值，可以获得最佳的正态分布曲线近似。应用此变换后，我们从异常分数中得到一个高斯分布，同时保留分数的相对排序。利用根据原始异常分数分布及其变换后的高斯分布计算得到的重要性采样权重，我们使模型能够学习不具有长尾特征的分布，从而提高其稳健性和异常检测性能。

算法 1 概述了所提出的损失函数。首先，计算异常分数（第 5 行）。随后，为确定权重向量$\mathbf{w}$的上界 T，使用以下公式计算异常分数 $f_s(\mathbf{s})$ 的偏度（第 6 行）：

$$f_s(\mathbf{s}) = \frac{1}{N}\sum_{i=1}^{N}(\frac{s_i - \mu}{\sigma})^3. \tag{11}$$

其中，μ 和 σ 分别表示异常分数的均值和标准差。为确保变换后的分布保留与原始分布相似的特征，我们首先对异常分数进行缩放，使其满足 Box-Cox 变换的条件（第 7 行）。接下来，我们应用 Box-Cox 变换（第 8 行）对分布进行归一化。对于概率估计，我们采用高斯估计而非核密度估计（KDE），因为与真实分布相比，KDE 往往会高估概率（第 9–10 行）。然而，传统的高斯估计容易受到离群值的影响。为缓解这一问题，我们引入了一种改进的 z 分数离群值检测机制，其公式如下。

$$\begin{aligned}f_p(\mathbf{s}) &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\big(-\frac{(\mathbf{s}-\mu)^2}{2\sigma^2}\big), \\ \mu &= \mathbb{E}[\hat{s}_i], \\ \sigma^2 &= \mathrm{Var}(\hat{s}_i), \\ \hat{\mathbf{s}} &= \{s_i | s_i \leq t\},\end{aligned} \tag{12}$$

其中，μ 和 σ 分别表示集合 $\hat{\mathbf{s}}$ 中分数的均值和标准差。集合$\hat{\mathbf{s}}$ 中的元素表示移除离群值后的异常分数，而 t 是用于识别离群值的阈值。使用改进的 Z 分数离群值检测器，阈值 t由下式确定

$$\begin{aligned}t &= 3.5 * \mathrm{MAD}/0.6745 + \hat{s}, \\ \mathrm{MAD} &= \mathrm{median}_{i\in\{1,\cdots,N\}}(|s_i - \bar{s}|), \\ \bar{s} &= \mathrm{median}_{i\in\{1,\cdots,N\}}(s_i).\end{aligned} \tag{13}$$

这里，MAD 是中位数绝对偏差，$\bar{s}$是$\mathbf{s}$中元素的中位数。随后，我们根据原始分布和变换后的分布所推导出的概率，计算重要性采样权重（第 11 行）。

为确保稳定收敛并缓解权重过高导致的梯度爆炸问题，我们对权重设置上界。该上界根据分布的偏度动态确定，并进一步受硬性上限约束，T_0（第 12–13 行）。最后，使用这些权重，我们通过加权损失函数训练模型 f（第 14–15 行）。

表 I 使用 $\beta = 200$对三个长尾数据集进行评估。最高性能值以粗体突出显示。

模型	数据集	Loss	多数类										
			0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Avg.**
AE	MNIST	MSE	0.976	0.797	0.710	0.657	0.651	0.682	0.712	0.647	**0.687**	0.625	0.714
		IWL	**0.979**	**0.805**	**0.715**	**0.662**	**0.658**	**0.689**	**0.737**	**0.647**	0.682	**0.654**	**0.723**
	F-MNIST	MSE	0.719	0.798	0.752	0.783	**0.575**	**0.628**	0.559	0.766	0.659	0.721	0.696
		IWL	**0.730**	**0.811**	**0.759**	**0.789**	0.558	0.626	**0.572**	**0.779**	**0.664**	**0.764**	**0.705**
	CIFAR10	MSE	**0.470**	**0.460**	0.610	**0.622**	0.639	0.553	0.516	**0.506**	**0.520**	**0.465**	0.536
		IWL	0.466	**0.460**	**0.612**	**0.622**	**0.644**	**0.560**	**0.531**	0.501	0.516	0.458	**0.537**
深度支持向量数据描述	MNIST	MSE	0.849	0.790	0.730	0.655	0.624	0.653	0.717	0.722	0.796	0.712	0.725
		IWL	**0.880**	**0.832**	**0.772**	**0.666**	**0.705**	**0.723**	**0.774**	**0.761**	**0.805**	**0.763**	**0.768**
	F-MNIST	MSE	**0.807**	0.824	0.755	**0.833**	0.651	**0.669**	**0.595**	0.858	0.682	0.797	0.747
		IWL	0.790	**0.835**	**0.784**	0.827	**0.675**	0.662	0.593	**0.868**	**0.689**	**0.822**	**0.754**
	CIFAR10	MSE	**0.508**	**0.515**	0.575	0.616	**0.604**	0.602	0.557	0.514	0.564	0.506	0.556
		IIWL	0.501	0.513	**0.577**	**0.636**	0.597	**0.613**	**0.568**	**0.519**	**0.565**	**0.535**	**0.562**

算法 1 用于异常检测的长尾分布损失

输入： 样本 X，模型 f，超参数，α 和 T_0
$\epsilon = 0.0001$
对于每个训练轮次**执行**
  对于每个小批量**x** ⊂**X 执行**
    $\mathbf{s} = |\mathbf{x} - f(\mathbf{x})||_2^2$ // 计算异常分数
    $T = \max(|f_s(\mathbf{s})|, \epsilon)$ // 计算偏度 (11)
    $\mathbf{s} = \mathbf{s} -$ 最小值$(\mathbf{s}) + \epsilon$
    $\mathbf{b} = f_b(\mathbf{s})$ // Box-Cox 变换 (10)
    $\mathbf{p}_s = f_p(\mathbf{s})$ // 估计**s**(12)的概率
    $\mathbf{p}_b = f_p(\mathbf{b})$ // 估计**b**(12)的概率
    $\mathbf{w} = \mathbf{p}_b/\mathbf{p}_s$ // 重要性采样权重
    $T = \min(\alpha T, T_0)$ //计算上界
    $\mathbf{w} =$ 裁剪$(\mathbf{w}\,0, T)$
    $L = \frac{1}{N}\sum_{i=1}^{N} w_i * ||x_i - f(x_i)||_2^2$
    使用 L 更新模型参数
  结束循环
结束循环

V. 实验

A. 数据集

我们在三个广泛使用的基准数据集上评估 IWL：MNIST [21], FashionMNIST（F-MNIST） [22],和 CIFAR-10[23]。在传统的异常检测设置中，每个类别都被视为正常类，而所有其他类别则被视为异常类 [24], [25]。为模拟长尾分布，我们将第 i 类指定为多数类，并将相邻的 mod$(i+1, 10)$ 类指定为少数类。少数类中的样本数量，N_{minority}，为

定义为 $\frac{N}{\beta}$，其中 N 表示多数类中的样本数量，β 由 $\frac{N}{N_{\text{minority}}}$给出。对于图像数据集，我们采用受试者工作特征曲线下面积（AUROC）作为评估指标。实验结果报告了全部十个

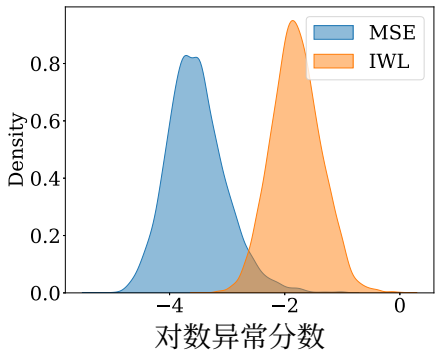

图4. 使用深度支持向量数据描述（DSVDD）可视化 MNISTtraining 数据集上经对数缩放的分数分布。

类别在三个不同随机种子下的平均 AUROC，其中依次将每个类别视为多数类。

为评估我们的方法在工业场景中的实际适用性，我们进一步在三个用于食品异常检测的高光谱成像（HSI）数据集 [1]上进行评估，其中正常对象为杏仁、蒜薹和开心果。该数据集具有独特的挑战：正常训练数据主要由背景和目标正常对象组成，因而存在显著的光谱变异性。因此，该数据集呈现长尾分布。数据集由 224 波段高光谱图像组成，分辨率为 512 ×400，每个类别各有一张训练图像和一张测试图像。我们仅使用 10% 的训练数据进行训练。测试数据包含罕见异常样本，仅占测试集总量的 8.9%。鉴于类别不平衡严重，我们使用精确率-召回率曲线下面积（AUPR）而非受试者工作特征曲线下面积（AUROC），因为在高度不平衡的场景中，AUPR 能更可靠地评估异常检测性能。

B. 设置

为评估 IWL 的普适性，我们使用两种具有代表性的异常检测

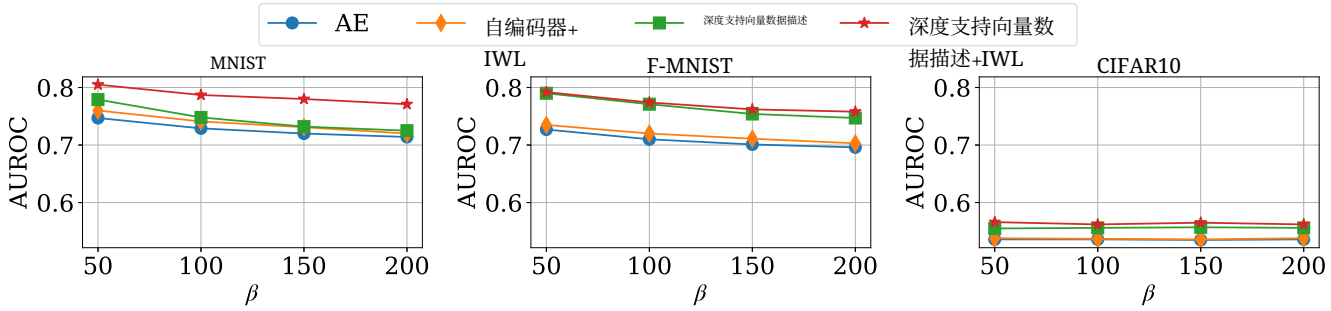

图 5. 三种图像数据集上不同 β value 的评估结果。

模型：自编码器（AE） [26], [27] 和深度支持向量数据描述（DSVDD） [25]。AE 是一种基于重构的模型，通过重现输入进行训练，并将重构误差作为异常检测标准。相比之下，DSVDD 旨在将正常数据映射到潜在空间中围绕中心点的紧凑区域。

对于图像数据集，自编码器的架构遵循 [24] 提出的设计。编码器由三个卷积模块组成，每个模块包含一个卷积层、批归一化 [28], 和带泄漏的ReLU 激活函数 [29],，滤波器数量分别为 16、32 和 64。卷积核大小和步幅大小分别设为 3 和 2。类似地，对于深度支持向量数据描述，自编码器架构遵循 [25] 中概述的框架。编码器由两个卷积层构成，分别包含 8×5×5 和 4×5×5 个滤波器，后接一个含 32 个单元的全连接层。卷积层之后依次应用批归一化、带泄漏的ReLU 激活和 (2×2) 最大池化。按照 [25] 的建议，为避免平凡解，移除了深度支持向量数据描述各层的偏置。解码器采用与编码器对称的设计，以反卷积操作替代卷积，并以上采样替代最大池化。最后一个反卷积层不包含批归一化等额外操作。对于高光谱图像数据集，自编码器和深度支持向量数据描述的编码器架构均由五个全连接层组成，除最后一层外，每层后均接批归一化和带泄漏的ReLU 激活。与图像数据集采用的方法类似，为避免平凡解，深度支持向量数据描述中移除了偏置。此外，解码器采用与编码器对称的设计。

在模型训练中，我们对所有模型均训练了 100 个训练轮次。此外，DSVDD 的预训练在图像数据集上进行 150 个训练轮次，在高光谱图像数据集上进行 20 个训练轮次。所有模型均使用 Adam优化器 [30]进行优化，学习率为 0.0001，权重衰减为 10^{-6}。上界 T 的超参数 α 和 T_0 分别设置为 4 和 20。

C. 可视化分布变换

图4展示了使用深度支持向量数据描述在MNIST训练数据集上得到的对数缩放异常分数分布。采用均方误差的分布偏度为0.623，而采用IWL的分布偏度为0.176。IWL将偏度降低了0.447，证明其能够有效地使异常分数分布趋近于高斯形式。

D. 长尾图像数据集上的评估

我们使用常规 AE 和 DSVDD 模型评估 IWL，并以均方误差（MSE）损失作为基线。表 I 展示了在三个数据集上，分别使用 MSE 和 IWL 训练的 AE 与 DSVDD 的性能比较，其中使用了 $\beta = 200$。值得注意的是，IWL 始终优于 MSE，使平均受试者工作特征曲线下面积提升了 0.43。这一效果在 MNIST数据集上的 DSVDD 中尤为明显，因为该数据分布表现出较强的偏度。随着数据集逐渐变得更加复杂——从 MNIST 到 CIFAR-10——IWL 所带来的性能增益逐渐减小，异常分数分布也越来越接近高斯分布。图 5 展示了平均受试者工作特征曲线下面积随少数类样本数量变化的情况。尽管 IWL 始终优于 MSE，但随着 β 减小，两种方法之间的性能差距逐渐缩小。这些结果表明，IWL 能够有效减轻长尾分布的不利影响，从而提升其在不同数据集上的异常检测性能。

E. 在真实世界数据集上的评估

我们在一个真实世界的工业数据集上进行评估，该数据集包含三个高光谱成像（HSI）子集。该数据集专为识别食品中的各种异物而设计。与标准基准数据集不同，该数据集中的正常数据既包括背景元素，也涵盖原材料的多种光谱变化，因此天然呈现长尾分布（LTD）。这一特性使其成为评估我们所提算法有效性的理想数据集。该高光谱成像数据集包含三个子集，分别以杏仁、蒜薹和开心果为主要对象。表 II 列出了 AE 和 DSVDD 在该数据集上的精确率-召回率曲线下面积（AUPR）结果。值得注意的是，与各自的基线相比，采用 IWL 的 AE 提升了 0.12，而采用 IWL 的 DSVDD 则提升了 0.033。这些

表 II 使用五个不同随机种子在真实世界高光谱异常检测数据集上的评估。最高性能值以粗体突出显示。

模型	Loss	杏仁	蒜薹	开心果	平均值
AE	MSE	0.716	0.885	**0.793**	0.797
	IWL	**0.728**	**0.897**	0.780	**0.801**
深度支持向量数据描述	MSE	0.705	0.816	0.769	0.763
	IWL	**0.738**	**0.846**	**0.799**	**0.794**

表 I 三种真实世界数据集上采用最先进方法的评估结果（跨五个不同随机种子）。最高性能值以粗体突出显示。

模型	杏仁	蒜薹	开心果	平均值
AE	0.716	0.885	0.789	0.797
深度支持向量数据描述	0.705	0.816	0.769	0.763
RAPP	0.701	0.659	0.674	0.678
MemAE	0.327	0.131	0.275	0.244
DIF	0.073	0.044	0.043	0.053
SLAD	0.449	0.465	0.440	0.452
AE+IWL	0.728	**0.897**	0.780	**0.801**
DSVDD+IWL	**0.738**	0.846	**0.799**	0.794

结果验证了我们的方法在提升真实世界数据集上的异常检测性能方面的有效性，并凸显了其在各种工业领域中的广泛适用性。

表 III 展示了我们的方法与最先进的异常检测模型之间的比较分析。除了自编码器和深度支持向量数据描述，我们还评估了多种近期的异常检测方法，包括沿投影路径重构（RAPP） [31], 记忆增强自编码器（MemAE） [24], DIF [32], 和 SLAD [33]。实验结果表明，现有模型难以从长尾数据中有效学习，常常将正常样本误分类为异常样本，导致性能大幅下降。相比之下，IWL 成功缓解了长尾分布带来的挑战，显著提升了其基础模型的性能。这些发现表明，IWL 不仅通过解决长尾分布相关的局限性来改善异常检测，还能作为一种稳健且具有良好泛化能力的解决方案，适用于各种异常检测框架。

VI. 结论

在本研究中，我们提出了重要性加权损失（IWL）函数，以应对异常检测中长尾分布（LTD）这一现实挑战。我们的方法基于一个理论假设：异常分数应近似服从高斯分布。基于这一前提，我们采用 Box-Cox 变换使分布呈现类高斯特征，并利用重要性采样提升异常检测性能。在三个基准图像数据集和三个高光谱成像数据集上进行的大量实验表明，IWL 可将自编码器和深度支持向量数据描述的性能提升 0.043，使其相较于传统方法能够更好地处理长尾分布。尽管 IWL 成功缓解了 LTD 问题，但本研究目前并未明确着眼于优化训练过程或增强模型泛化能力，这些方向将留待未来探索。此外，我们计划整合数据增强技术和对比学习，进一步提升性能。除了这些改进之外，我们还打算评估 IWL 在不同应用中应对 LTD 相关挑战的有效性，例如医疗异常检测和金融欺诈检测。总之，IWL 有效缓解了 LTD 引发的欠拟合问题，显著提升了异常检测性能。尽管该方法只需修改损失函数，而无需改变底层模型架构，但它为现实世界中的工业应用提供了一种切实可行且高效的解决方案。

参考文献

[1] J. Lee, M. Kim, J. Yoon, K. Yoo, 和 S.-J. Byun, 《用于食品安全检测的高光谱成像异常检测》, 2024年。 [在线]。可访问：https://dx.doi.org/10 21227 963 1 34 。/ e- d G. E. Hinton 和 [2] S. Roweis, 《随机邻域嵌入》, 神经信息处理系统进展, 第15卷, 2002年。[3] Y. Zhang, B. Kang, B. Hooi, S. Yan, 和 J. Feng, 《深度长尾学习：综述》, *IEEE*模式分析与机器智能汇刊, 第45卷, 第9期, 第10 795–10 816页, 2023年。[4] X.-Y. Liu, J. Wu, 和 Z.-H. Zhou, 《用于类别不平衡学习的探索性欠采样》, *IEEE*系统、人类与控制论汇刊*B*辑（控制论）, 第39卷, 第2期, 第539–550页, 2008年。[5] N. V. Chawla, K. W. Bowyer, L. O. Hall, 和 W. P. Kegelmeyer, 《SMOTE：合成少数类过采样技术》, 人工智能研究杂志, 第16卷, 第321–357页, 2002年。[6] Z. Zhang 和 T. Pfister, 《在没有奖励数据的情况下学习快速样本重加权》, 载于*IEEE/CVF*国际计算机视觉会议论文集, 2021年, 第725–734页。[7] H. Zhang, M. Cisse, Y. N. Dauphin, 和 D. Lopez-Paz, 《mixup：超越经验风险最小化》, 载于国际学习表征会议, 2018年。 [在线]。可访问：https://openreview.net/forum?id=r1Ddp1-Rb [8] Z. Zhong, J. Cui, S. Liu, 和 J. Jia, 《改进长尾识别的校准》, 载于*IEEE/CVF*计算机视觉与模式识别会议论文集, 2021年, 第16 489–16 498页。[9] Y. Yang 和 Z. Xu, 《重新思考标签在改进类别不平衡学习中的价值》, 神经信息处理系统进展, 第33卷, 第19 290–19 301页, 2020年。[10] J. Kim, K. Jeong, H. Choi, 和 K. Seo, 《不平衡问题中的基于GAN的异常检测》, 载于计算机视觉——*ECCV 2020*研讨会：英国格拉斯哥, *2020*年*8*月*23*日至*28*日, 论文集, 第*VI*部分*16*。Springer, 2020年, 第128–145页。[11] A. Estabrooks, T. Jo, 和 N. Japkowicz, 《一种从不平衡数据集中学习的多重重采样方法》, 计算智能, 第20卷, 第1期, 第18–36页, 2004年。[12] C. Elkan, 《代价敏感学习的基础》, 载于国际人工智能联合会议, 第17卷, 第1期。Lawrence Erlbaum Associates Ltd, 2001年, 第973–978页。[13] Z.-H. Zhou 和 X.-Y. Liu, 《采用解决类别不平衡问题的方法训练代价敏感神经网络》, *IEEE*知识与数据工程汇刊, 第18卷, 第1期, 第63–77页, 2005年。[14] Y. Sun, M. S. Kamel, A. K. Wong, 和 Y. Wang, 《用于不平衡数据分类的代价敏感提升》, 模式识别, 第40卷, 第12期, 第3358–3378页, 2007年。[15] Y.-X. Wang, D. Ramanan, 和 M. Hebert, 《学习对尾部进行建模》, 神经信息处理系统进展, 第30卷, 2017年。[16] J. Wang, T. Lukasiewicz, X. Hu, J. Cai, 和 Z. Xu, 《RSG：一种用于不平衡数据集学习的简单而有效的模块》, 载于*IEEE/CVF*计算机视觉与模式识别会议论文集, 2021年, 第3784–3793页。

[17] B. Kang, S. Xie, M. Rohrbach, Z. Yan, A. Gordo, J. Feng, 和 Y. Kalantidis, 《解耦长尾识别中的表征与分类器》, 载于国际学习表征会议, 2020年。 [在线]。可访问: https://openreview.net/forum?id=r1gRTCVFvB [18] D. Ding, M. Zhang, X. Pan, M. Yang, 和 X. He, 《时间序列预测中的极端事件建模》, 载于第25届ACM SIGKDD知识发现与数据挖掘国际会议论文集, 2019年, 第1114–1122页。[19] C.-H. Ho, K.-C. Peng, 和 N. Vasconcelos, 《借助可学习类别名称进行长尾异常检测》, 载于IEEE/CVF计算机视觉与模式识别会议（CVPR）论文集, 2024年6月, 第12 435–12 446页。[20] A. Banerjee, S. Merugu, I. S. Dhillon, J. Ghosh, 和 J. Lafferty, 《使用Bregman散度进行聚类》。机器学习研究杂志, 第6卷, 第10期, 2005年。[21] Y. LeCun, C. Cortes, 和 C. Burges, 《MNIST手写数字数据库》, ATT Labs [在线]。可访问: http://yann.lecun.com/exdb/mnist, 第。[22] H. Xiao, K. Rasul, 和 R. Vollgraf, 《Fashion-MNIST: 一个用于基准测试机器学习算法的新型图像数据集》, CoRR, 卷号abs/1708.07747, 2017年。 [在线]。可访问: http://arxiv.org/abs/1708 07747 。[23] A. Krizhevsky, 《从微小图像中学习多层特征》, 技术报告, 2009年。[24] D. Gong, L. Liu, V. Le, B. Saha, M. R. Mansour, S. Venkatesh, 和 A. v. d. Hengel, 《记忆正常性以检测异常: 用于无监督异常检测的记忆增强深度自编码器》, 载于IEEE/CVF国际计算机视觉会议论文集, 2019年, 第1705–1714页。[25] L. Ruff, R. Vandermeulen, N. Goernitz, L. Deecke, S. A. Siddiqui, A. Binder, E. M¨uller, 和 M. Kloft, 《深度单类分类》, 载于国际机器学习会议。PMLR, 2018年, 第4393–4402页。[26] G. E. Hinton, S. Osindero, 和 Y.-W. Teh, 《一种用于深度置信网络的快速学习算法》, 神经计算, 第18卷, 1554, 2006 7, 1527期, 第–页。[27] P. Bergmann, S. L¨owe, M. Fauser, D. Sattlegger, 和 C. Steger, 《通过将结构相似性应用于自编码器来改进无监督缺陷分割》, arXiv预印本arXiv:1807.02011, 2018年。[28] S. Ioffe 和 C. Szegedy, 《批归一化: 通过减少内部协变量偏移加速深度网络训练》, 载于国际机器学习会议。PMLR, 2015年, 第448–456页。[29] B. Xu, N. Wang, T. Chen, 和 M. Li, 《卷积网络中整流激活函数的实证评估》, arXiv预印本arXiv:1505 00853, 2015年。[30] D. P. Kingma 和 J. Ba, 《Adam: 一种随机优化方法》, arXiv预印本arXiv:1412.6980, 2014年。[31] K. H. Kim, S. Shim, Y. Lim, J. Jeon, J. Choi, B. Kim, 和 A. S. Yoon, 《RAPP: 沿投影路径重构的新颖性检测》, 载于国际学习表征会议, 2020年。[在线]。可访问: https://openreview.net/forum?id=HkgeGeBYDB [32] H. Xu, G. Pang, Y. Wang, 和 Y. Wang, 《用于异常检测的深度孤立森林》, IEEE知识与数据工程汇刊, 第35卷, 第12期, 第12 591–12 604页, 2023年。[33] H. Xu, Y. Wang, J. Wei, S. Jian, Y. Li, 和 N. Liu, 《令人瞩目的监督信号及其来源: 通过尺度学习进行深度异常检测》, 载于国际机器学习会议。PMLR, 2023年, 第38 655–38 673页。