

基于软约束潜在正则化对抗的高压并联电抗器异常声音检测

August 30, 2024

Abstract

本文提出了一种软约束潜在正则化对抗异常检测方法 (Soft-LRAAD)，旨在解决 LRAAD 方法中超参数 M 对生成器生成频谱图的 KL 散度上限的硬约束带来的问题。LRAAD 方法使用硬约束来控制 KL 散度上限，这样不仅使模型难以将正常数据和异常数据在潜在空间中进行有效分离，而且会导致模型训练不稳定，进而影响模型性能。为了解决这一问题，Soft-LRAAD 方法引入了软约束损失替代硬约束损失，即使用一个平滑的函数来逼近 KL 散度的上限约束损失。通过这种方式，Soft-LRAAD 方法提高了模型在潜在空间中区分正常数据和异常数据的能力，同时增强了模型训练的稳定性，从而提升了整体性能。本文在变电站实测声音数据集上进行了实验验证，结果显示 Soft-LRAAD 方法在异常检测任务上的 AUC 指标明显优于其他基线模型，验证了该方法的有效性和实用性，为异常声音检测在高压电抗器态监测领域提供了新的解决方案。

关键字：声学异常检测, 半监督学习, 对抗学习, 潜在变量正则化对抗

1 引言

高压并联电抗器在特高压变电站中起着至关重要的作用，其主要功能是补偿无功功率，稳定电网电压，防止电力系统中过电压的发生^[1-3]。由于特高压变电站运行环境复杂，高压并联电抗器在长期运行过程中可能会出现各种故障，例如电弧放电、绕组松动、机械振动等^[4-5]。这些故障往往伴随着异常的声音信号，因此对高压并联电抗器进行异常声音检测，对于保障电力系统的安全和可靠性具有重要意义。

目前，针对高压并联电抗器的异常检测方法主要包括油色谱检测、超声局放检测和高频局放检测。这些方法虽然在一定程度上能够帮助识别设备内部的潜在故障，但也存在着一些局限性：

- 油色谱检测：**油色谱检测是一种成熟的故障诊断方法，主要通过分析油中的气体组分来判断设备内部是否存在放电或过热问题^[6]。尽管油色谱分析能够提供关于放电或过热故障的早期预警信息，但其检测周期较长，实时性差，且难以反映电抗器的瞬时状态变化。此外，油色谱检测对早期和微小故障的敏感度较低，可能导致故障在发展到较严重阶段时才被检测到。
- 超声和高频局放检测：**超声检测通过捕捉电抗器内部局部放电产生的超声波信号来检测异常^[7]，高频检测利用高频电磁波信号检测电抗器内部的局部放电现象^[8]。这些方法在检测高压设备中的局部放电方面具有较高的灵敏度，尤其是在设备内部绝缘材料老化或发生局部放电时可以提供较为准确的诊断结果。然而，这些方法的检测精度和灵敏度容易受到环境噪声和外部干扰的影响，尤其是在复杂的电磁环境中，检测结果可能存在不稳定性。此外，超声和高频局放检测在检测机械故障和设备的振动问题时，效果较为有限，难以全面反映高压并联电抗器的运行状态。

相比之下，基于声学信号的检测方法能够同时监测设备运行中的机械和电气故障，且不受油质变化和局放信号的限制，具有更强的实时性和鲁棒性^[9-10]。本文提出的 Soft-LRAAD 方法通过软约束潜在正则化对抗异常检测技术，结合深度学习模型在声学信号分析中的优势，能够有效提高高压并联电抗器故障检测的精度和可靠性。

随着人工智能技术的快速发展，基于数据驱动异常检测方法逐渐成为研究热点。近年来，深度学习技术在图像识别、语音处理等领域取得了显著成果，这为高压并联电抗器的异常检测提供了新的思路。特别是对抗学习（Adversarial Learning）和潜在变量模型（Latent Variable Models）的结合，能够有效地提取数据的深层次特征，实现对异常模式的高精度识别。然而，现有的对抗学习方法，如潜在正则化对抗异常检测（LRAAD），在应用于高压并联电抗器的异常检测时，仍存在问题。LRAAD 方法通过硬约束来控制潜在空间的 KL 散度上限，这种硬约束虽然可以帮助模型在潜在空间中区分正常数据和异常数据，但也容易导致模型在训练过程中出现不稳定性，影响最终的检测效果。此外，硬约束的引入还可能使得生成的潜在空间不够灵活，难以捕捉异常数据的多样性和复杂性。

为了解决这些问题，本文提出了一种基于软约束潜在正则化对抗学习的高压并联电抗器异常检测方法（Soft-LRAAD）。该方法通过引入软约束损失函数，替代 LRAAD 方法中的硬约束损失，从而在潜在空间中更为灵活地处理正常数据和异常数据的分布问题。这种改进不仅提高了模型的训练稳定性，也增强了模型在不同工况下的异常检测性能。此外，本文在特高压变电站的实测声音数据集上，对所提出的方法进行了实验验证，结果表明，Soft-LRAAD 方法在异常检测任务中表现出优异的性能，显著优于传统方法。

本文的主要贡献包括以下几点：

1. 提出了异常检测方法 Soft-LRAAD，通过软边界的方式来控制生成器生成频谱图的 KL 散度上界，极大程度提高正常数据和异常数据在潜在空间的分离效果，从而提高了异常检测的准确性和鲁棒性。
2. 设计了基于 KL 散度损失和重构损失的异常分数，相比与仅使用重构损失作为异常分数的方法，基于 KL 散度和重构损失的异常分数能够更好地区分正常数据和异常数据，提高了异常检测的性能。
3. 使用特高压变电站的实测声音数据对模型进行了验证，实验结果表明，Soft-LRAAD 方法在不同负载条件下均表现出较高的异常检测准确率和鲁棒性。

本文的结构安排如下：第二章介绍了相关研究工作和背景知识；第三章介绍了 LRAAD 模型的背景知识，第四章详细描述了 Soft-LRAAD 模型的设计和实现，包括编码器和生成器的架构及其训练方法；第五章介绍了实验设置和数据集划分，并对实验结果进行分析和讨论；第六章总结了本文的研究工作，并提出了未来的研究方向。

2 相关工作

过去的研究已经涵盖了多种声学异常检测方法和相关技术。其中，一些方法依赖于传统的信号处理技术，而另一些方法则利用了深度学习和人工智能等新兴技术^[11-12]。基于传统信号处理的方法通常包括特征提取、模式识别和分类等步骤。这些方法通常包括使用滤波器、时频分析、频谱分析等技术来提取声音信号的特征，并使用分类器（如支持向量机、k 最近邻等）来识别异常。尽管传统方法在某些情况下取得了一定的成果，但其局限性在于依赖于人工特征设计和手动调整参数，且难以处理复杂的非线性关系和大规模数据^[13-15]。

近年来,随着深度学习技术的发展,基于潜在变量模型的方法在电力设备声学异常检测领域取得了显著的进展^[16-21]。Duman 等人^[18]提出了一种基于卷积自编码器 (CAE) 的工业过程声学异常检测方法。该方法使用对数刻度的 Mel 频谱特征作为模型的输入特征。与传统的异常检测方法 One-Class Support Vector Machine (OCSVM)^[22]相比, CAE 方法在工业过程中表现出更好的异常检测效果。基于变分自编码器 (VAE) 的异常检测方法^[19-20]通过引入潜在变量和 KL 散度来学习数据的潜在表示,从而提高模型的泛化能力。例如, An 等人^[19]提出了一种基于变分自编码器的异常检测方法,该方法通过训练一个 VAE 模型来学习正常数据的潜在表示,然后使用 KL 散度来衡量正常数据和异常数据之间的差异。

基于生成对抗网络的异常检测方法在图像异常检测领域取得了显著的进展。文献^[23]提出了一种基于生成对抗网络 (GANs) 的异常检测方法,通过对正常样本进行建模来识别异常。f-AnoGAN^[24]在 AnoGAN^[23]方法的改进, AnoGAN 使用迭代方法寻找图像在潜在空间的表示, f-AnoGAN 通过学习从图像到潜在空间的映射,显著提高了速度。文献^[25]提出了一种名为 GANomaly 的新型编码器-解码器-编码器架构模型,其同时使用重构损失、潜在表示损失和对抗损失来训练模型和进行异常分数计算,实验表明,其性能优于当代最先进的基于 GAN 和传统自编码器的异常检测方法,具有泛化到任何异常检测任务的能力。Jiang 等人^[26]提出了一种利用生成对抗网络 (GAN) 进行机器音频异常检测的无监督模型 AEGAN-AD。研究表明,通过引入鉴别器来提供特征级别的指导, AEGAN-AD 模型可以更深入地理解表示,而不仅仅关注表面噪音,解决了 CAE 可能会重构出异常信号的问题和 WaveNet 训练和预测时间长的问題。实验结果显示,该 AEGAN-AD 模型在 DCASE 2022 挑战中取得了优异的表现,达到了最先进的性能水平。

LRAAD 模型结合了对抗学习和潜在变量模型的优势,能够有效地提取数据的深层次特征,实现对异常模式的高精度识别。然而, LRAAD 方法中硬约束的引入容易导致模型训练不稳定,影响最终的检测效果。为了解决这一问题,本文提出了一种基于软约束潜在正则化对抗学习的高压并联电抗器异常检测方法 (Soft-LRAAD)。Soft-LRAAD 方法通过软约束损失替代硬约束损失,提高了模型在潜在空间中区分正常数据和异常数据的能力,增强了模型训练的稳定性,从而提升了整体性能。

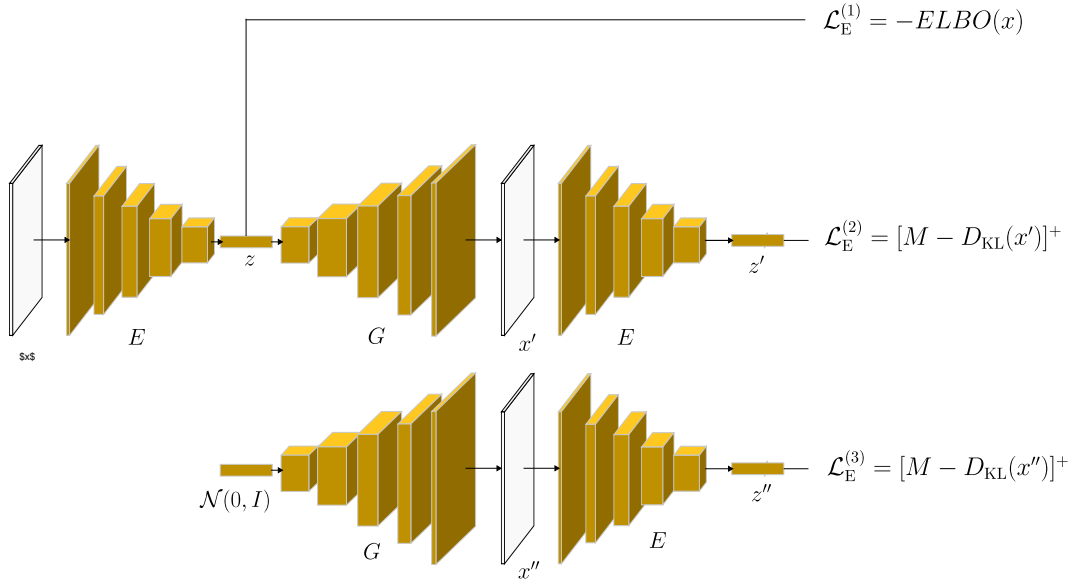
3 背景

证据下界: 证据下界 (Evidence Lower Bound, ELBO) 是变分自编码器 (Variational Autoencoder, VAE) 中的一个重要概念,用于衡量潜在变量的分布与真实分布之间的差异。ELBO 由两部分组成,一部分是重构误差,用于衡量生成器生成的数据与原始数据之间的差异;另一部分是 KL 散度,用于衡量潜在变量的分布与先验分布之间的差异。通过最大化 ELBO,可以使得生成器生成的数据更加接近真实数据,同时潜在变量的分布更加接近先验分布。证据下界的公式如下:

$$ELBO(x) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - D_{KL}(q_\phi(z|x)||p(z)) \quad (1)$$

LRAAD: 潜在正则化对抗异常检测 (Latent Regularized Adversarial Anomaly Detection, LRAAD) 是一种基于对抗学习的异常检测方法,通过编码器和生成器来学习正常数据的潜在表示,并通过对抗学习的方式来区分正常数据和异常数据。LRAAD 模型的核心思想是通过最小化正常数据的潜在表示与标准正态分布之间的 KL 散度,来学习将正常数据映射到潜在空间中的正态分布,同时将异常数据映射到潜在空间中的非正态分布。通过这种方式, LRAAD 模型可以更好地区分正常数据和异常数据,从而实现异常检测的目的。

Stage1: Train Encoder, Fix Generator



Stage2: Train Generator, Fix Encoder

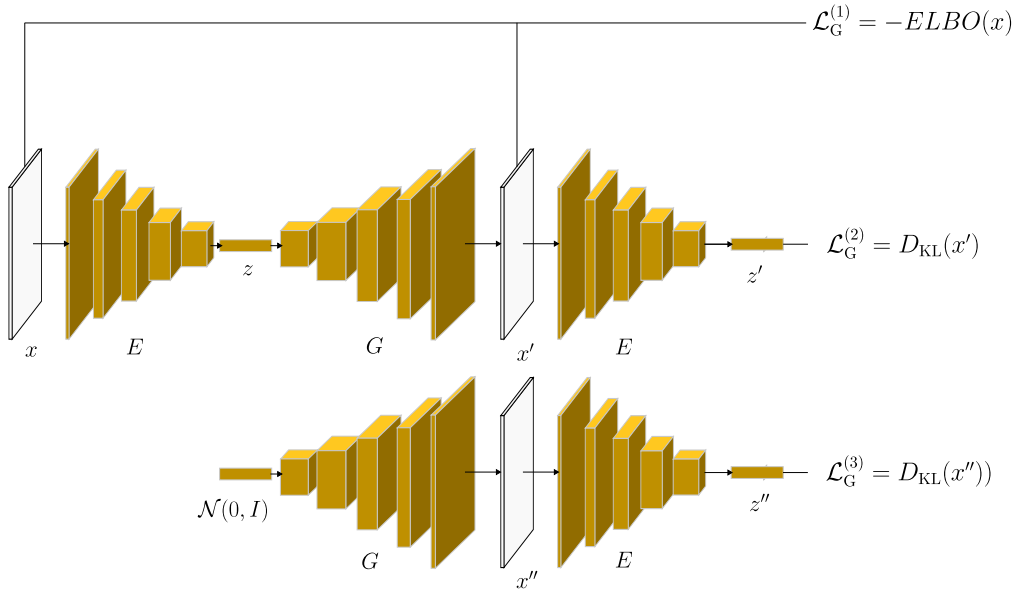


Figure 1: The Training Pipeline of LRAAD

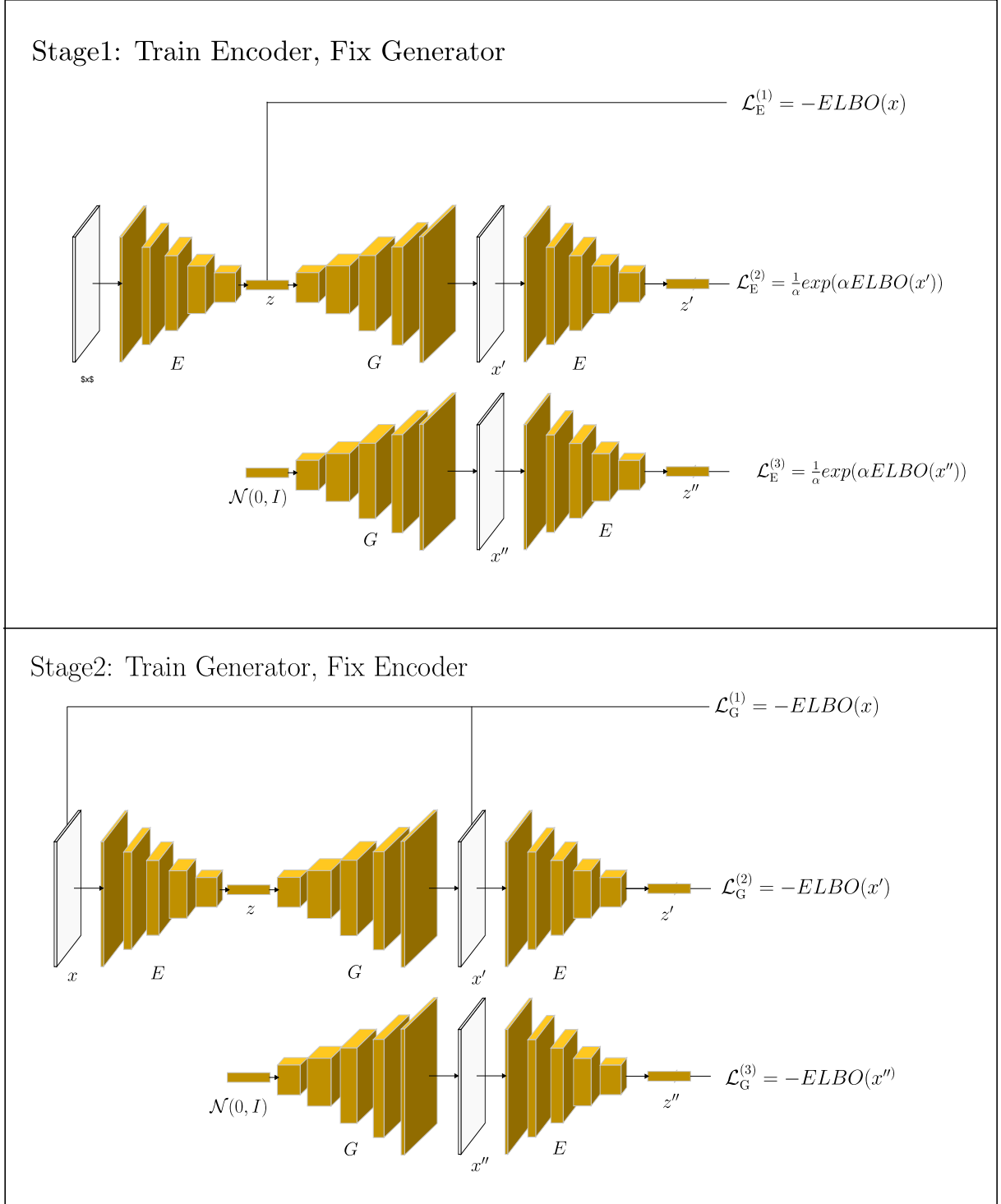


Figure 2: The Training Pipeline of Soft-LRAAD

记 $D_{\text{KL}}(x) = D_{\text{KL}}(q(\cdot | x) \| p(\cdot))$ 。其中， $q(\cdot | x)$ 表示编码器编码频谱图 x 的潜在表示的分布， $p(\cdot)$ 表示标准正态分布，记为 $\mathcal{N}(0, I)$ ， D_{KL} 表示 KL 散度。

$$\mathcal{L}_{\text{E}} = \|x - x'\|^2 + D_{\text{KL}}(x) + [M - D_{\text{KL}}(x')]^+ + [M - D_{\text{KL}}(x'')]^+ \quad (2)$$

$$\mathcal{L}_{\text{G}} = \|x - x'\|^2 + D_{\text{KL}}(x') + D_{\text{KL}}(x'') \quad (3)$$

等价地,

$$\mathcal{L}_E = -ELBO(x) + [M - D_{KL}(x')]^+ + [M - D_{KL}(x'')]^+ \quad (4)$$

$$\mathcal{L}_G = -ELBO(x) + D_{KL}(x') + D_{KL}(x'') \quad (5)$$

其中, $[M - D_{KL}]^+$ 表示生成图像的 KL 散度的约束项, 其定义为 $\max(0, M - D_{KL})$, 其中 M 是一个超参数, 用于控制 KL 散度的上限, 避免生成图像的 KL 散度无穷大。

4 Soft-LRAAD

4.1 Soft-LRAAD 方法概述

Soft-LRAAD (Soft-Constrained Latent Regularization Adversarial Anomaly Detection) 是一种改进的异常检测方法, 旨在克服 LRAAD 方法中因硬约束导致的潜在空间区分能力不足和模型训练不稳定的问题。通过引入软约束损失函数, Soft-LRAAD 不仅能够更有效地区分正常数据和异常数据, 还提高了模型的训练稳定性。

LRAAD 方法采用硬约束来控制生成图像的 KL 散度上限, 硬约束的引入常常导致模型在潜在空间中难以区分正常数据和异常数据。此外, 硬约束还可能导致训练过程中的梯度不稳定, 从而影响模型的整体性能。为了解决这些问题, Soft-LRAAD 方法提出使用软约束损失函数来替代硬约束损失。

在 Soft-LRAAD 方法中, 我们使用一个平滑的函数 $\frac{1}{\alpha} \exp(-\alpha D_{KL}(x))$ 来逼近 KL 散度的上限约束损失 $[M - D_{KL}(x)]^+$, 它们的函数图像如图 3 和图 4 所示。具体来说, 我们使用软约束损失函数 $\frac{1}{\alpha} \exp(-\alpha D_{KL}(x')) + \frac{1}{\alpha} \exp(-\alpha D_{KL}(x''))$ 来替代 LRAAD 中的硬约束损失 $[M - D_{KL}(x')]^+ + [M - D_{KL}(x'')]^+$ 。硬约束的引入会在潜在空间中设置严格的边界, 使得生成图像的 KL 散度限制在边界 M 之下, 限制了模型的表现。而软约束通过使用平滑函数来逼近 KL 散度的上限, 从而缓解了这种限制。使得模型能够在保持生成图像质量的同时, 更好地在潜在空间中分离正常数据和异常数据。

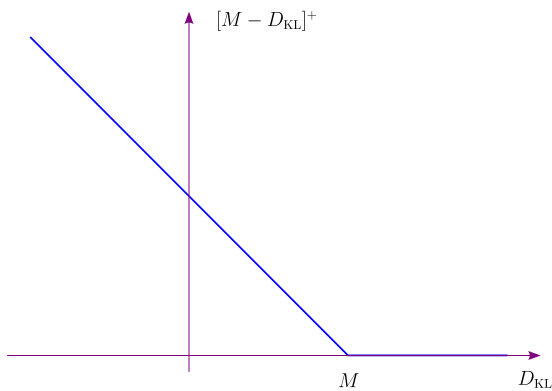


Figure 3: The hard constraints function $[M - D_{KL}]^+$ of KL divergence in LRAAD

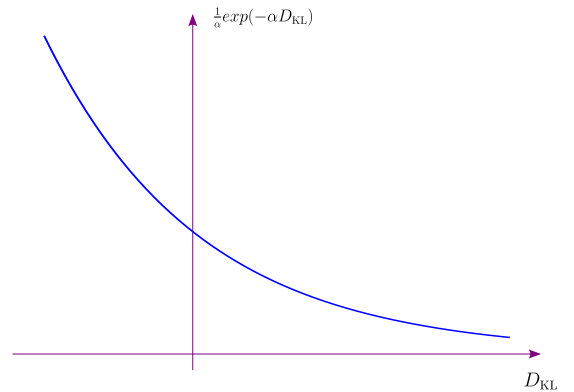


Figure 4: The soft constraints function $\frac{1}{\alpha} \exp(-\alpha D_{KL})$ of KL divergence in Soft-LRAAD

4.2 软约束损失

Soft-LRAAD 方法中的软约束损失是对 LRAAD 方法的重要改进之一。LRAAD 方法使用硬约束来控制生成图像的 KL 散度上限，这导致了模型在潜在空间中难以有效区分正常数据和异常数据，同时也可能造成训练不稳定的问题。为了解决这些挑战，Soft-LRAAD 引入了软约束损失，通过使用一个平滑的函数来逼近 KL 散度的上限约束损失，从而缓解了硬约束带来的问题。

为了克服硬约束的缺点，Soft-LRAAD 方法引入了软约束损失。软约束损失函数的形式为 $\frac{1}{\alpha} \exp(-\alpha D_{\text{KL}}(x))$ ，其中 α 是一个可调节的超参数。相比于硬约束，软约束的函数图像呈现出平滑递减的趋势，这样的设计使得损失在 KL 散度接近阈值时依旧缓慢降低，而不是硬约束那样保持不变，这样可以使得生成图像的 KL 散度可以变大，但又不至于无穷大。通过这种方式，Soft-LRAAD 方法能够更好地保持生成图像的质量，同时提高模型的异常检测性能。

与的 LRAAD 方法不同,Soft-LRAAD 采用了软约束的方式来处理 KL 散度项,使得模型在训练时对 KL 散度的惩罚更加灵活。具体来说,Soft-LRAAD 的编码器损失函数 \mathcal{L}_E 和生成器损失函数 \mathcal{L}_D 分别定义如下:

$$\mathcal{L}_E = -ELBO(x) + \frac{1}{\alpha} \exp(-\alpha D_{\text{KL}}(x')) + \frac{1}{\alpha} \exp(-\alpha D_{\text{KL}}(x'')) \quad (6)$$

$$\mathcal{L}_G = -ELBO(x) + D_{\text{KL}}(x') + D_{\text{KL}}(x'') \quad (7)$$

其中, α 一个可调节的超参数,用于控制软约束项的影响程度。

对于 KL 散度项,加一个重构损失项,这样可以更好地保持生成图像的质量,同时提高模型的异常检测性能。通过这种方式,Soft-LRAAD 方法能够更好地区分正常数据和异常数据,提高异常检测的准确性和鲁棒性。其公式如下所示。

$$\mathcal{L}_E = -ELBO(x) + \frac{1}{\alpha} \exp(\alpha ELBO(x')) + \frac{1}{\alpha} \exp(\alpha ELBO(x'')) \quad (8)$$

$$\mathcal{L}_G = -ELBO(x) - ELBO(x') - ELBO(x'') \quad (9)$$

软约束对模型训练的影响主要体现在几个方面。首先,软约束使得潜在变量在潜在空间中分布更加灵活,增强了对正常数据和异常数据的区分能力。其次,软约束的引入有助于模型在训练过程中更加稳定,避免了硬约束可能导致的梯度不连续和训练不稳定问题。最后,通过调节 α 参数,可以灵活控制软约束对 KL 散度的惩罚力度,进一步优化模型性能。

4.3 模型训练

Soft-LRAAD 的结构与 LRAAD 类似,包括编码器 E 、生成器 G 两个部分。编码器 E 负责将输入的梅尔频谱图 x 映射到潜在空间,生成器 G 负责将潜在变量映射回梅尔频谱域,实现频谱图的重建。Soft-LRAAD 的整体架构如图 2所示。

Soft-LRAAD 的训练过程主要包括编码器 E 和生成器 G 的训练。具体的训练算法如 Algorithm 1 所示。Soft-LRAAD 的训练与 LRAAD 类似,编码器与生成器以对抗的方式进行训练,通过最小化编码器和生成器的损失函数来优化模型参数。通过交替训练编码器和生成器,Soft-LRAAD 模型的编码器能更

好地区分正常数据和异常数据，生成器能更好地重建输入的频谱图，从而提高异常检测的准确性和鲁棒性。

Algorithm 1 Training Soft-LRAAD model

- 1: Initialize network parameters
 - 2: **For** number of epochs **do**
 - 3: Random mini-batch X from dataset
 - 4: Compute latent representation of X : $Z = E(X)$
 - 5: Sample Z_p from prior distribution $p(z) = \mathcal{N}(0, I)$
 - 6: Compute reconstructed spectrogram of Z and Z_p : $X' = G(Z)$, $X'' = G(Z_p)$
 - 7: Compute latent representation of X' and X'' : $Z' = E(X')$, $Z'' = E(X'')$
 - 8: Compute Encoder loss \mathcal{L}_E according to Eq. 8
 - 9: Update Encoder parameters by minimizing \mathcal{L}_E
 - 10: Compute Generator loss \mathcal{L}_G according to Eq. 9
 - 11: Update Generator parameters by minimizing \mathcal{L}_G
 - 12: **End For**
-

4.4 异常检测

我们采用重构误差和 KL 散度之和作为异常分数。重构误差衡量生成频谱图与原始输入频谱图之间的差异，而 KL 散度则衡量潜在表示与标准正态分布之间的差异。对于正常数据，其重构误差较小，潜在表示接近标准正态分布；而对于异常数据，其重构误差较大，潜在表示远离标准正态分布。通过结合这两个指标，模型能够更准确地识别异常数据。异常分数定义的数学表达式如下：

$$\mathcal{A}(x) = \|x - G(E(x))\|^2 + D_{\text{KL}}(E(x) \parallel \mathcal{N}(0, I)) \quad (10)$$

其中， $\mathcal{A}(x)$ 表示输入频谱图 x 的异常分数， $E(x)$ 表示输入频谱图 x 的潜在表示， $G(E(x))$ 表示潜在表示 $E(x)$ 的重构频谱图， $D_{\text{KL}}(E(x) \parallel \mathcal{N}(0, I))$ 表示潜在表示 $E(x)$ 与标准正态分布之间的 KL 散度。通过计算异常分数，我们可以对输入频谱图 x 进行异常检测，判断其是否为异常数据。

5 Experiments

本节我们将介绍实验的设置和结果。我们首先介绍实验的数据集和预处理方法，然后详细描述实验的设置和评估指标，最后给出实验结果和分析。

5.1 Data Preprocessing

在本研究中，为了确保高压并联电抗器声音信号的异常检测能够获得准确和可靠的结果，对实验数据进行了系统而细致的预处理。数据预处理的质量直接影响到模型的训练效果和检测性能，因此，我们采用了一系列步骤来优化声音信号数据的处理过程。以下是数据预处理的具体步骤：

Table 1: Soft-LRAAD Model Hyperparameter Settings

Hyperparameter	Setting
Latent Space Dimension	32
Learning Rate (Encoder)	0.0002
Learning Rate (Generator)	0.0002
Loss Weight (Reconstruction)	1
Loss Weight (Adversarial)	1
Loss Weight (KL Divergence)	1
Alpha (Soft Constraint Coefficient)	128×128
Training Epochs	100
Batch Size	32
Dataset Split	80% Training, 30% Testing

实验数据来自中国某特高压变电站的高压并联电抗器实测声音数据。这些声音信号是通过安装在变电站内的高灵敏度传感器采集的，录音文件以 WAV 格式存储，每个文件的采样频率为 44.1kHz，持续时间为 10 秒。由于原始音频文件较长且包含多种工况信息，为了提高数据处理的精度和模型的训练效率，我们将每个录音文件切分为多个长度为 1.36 秒的音频片段。这种切分方式确保每个片段能够独立反映电抗器的运行状态，并且便于后续的特征提取和模型训练。

为了更好地捕捉声音信号中的特征信息，我们将切分后的音频片段转换为梅尔频谱图 (Mel-spectrogram)。梅尔频谱图是一种在横轴上采用 Mel 刻度、纵轴上表示功率谱的时频表示形式，它能够更符合人类听觉特性地反映声音信号的频谱特征。相比于传统的频谱图，梅尔频谱图能够更有效地聚焦于声音信号的关键频段，增强低频部分的信息，使得信号特征更加显著。

在生成梅尔频谱图后，为了进一步处理和分析信号中的细微变化，我们对梅尔频谱图进行了对数变换。对数变换能够压缩频谱中的幅度动态范围，减弱高频部分的影响，同时突出低频信号的细节信息。这一变换使得频谱数据更加平滑稳定，有助于模型更好地学习声音信号中的异常特征。

5.2 Experimental Settings

在训练 Soft-LRAAD 模型时，我们需要设置多个超参数来调节模型的性能。例如，编码器和生成器的学习率、损失函数的权重系数（如重构损失、对抗损失、KL 散度损失的权重）、潜在空间的维度大小、软约束系数、训练的迭代次数等。这些超参数的选择对模型的训练和性能具有重要影响，需要通过实验和验证来确定最佳的超参数组合。在实验中，我们采用的超参数设置如 Table 1所示：

5.3 Evaluation Metrics

在本研究中，我们使用了多种评估指标来量化和比较不同异常检测模型的性能，特别是针对高压并联电抗器声音信号的检测效果。这些评估指标包括 ROC 曲线 (Receiver Operating Characteristic Curve)、AUC (Area Under the ROC Curve) 值和 p-AUC (Partial Area Under the ROC Curve) 值。以下是对这些评估指标的详细描述。

ROC 曲线: ROC 曲线是一种常用的评估二分类模型性能的工具。它通过绘制真阳性率 (True Positive Rate, TPR) 与假阳性率 (False Positive Rate, FPR) 之间的关系, 直观展示模型在不同阈值下的分类能力。对于异常检测任务, ROC 曲线能够反映出模型在识别异常数据 (阳性) 与正常数据 (阴性) 之间的平衡情况。一般来说, ROC 曲线越接近左上角 (即 TPR 高且 FPR 低), 说明模型的检测性能越好。在本实验中, 我们通过对不同模型生成的异常分数设定多个阈值, 计算各个阈值下的 TPR 和 FPR, 并绘制 ROC 曲线, 从而对比不同模型的异常检测能力。

AUC (Area Under the Curve) 值: AUC 值是 ROC 曲线下方的面积, 是衡量模型分类效果的一个综合指标。AUC 的取值范围在 0.5 到 1 之间, 其中 0.5 表示模型的性能与随机分类等同, 1 表示模型能够完美区分正常数据与异常数据。因此, AUC 值越接近 1, 说明模型在异常检测中的分类能力越强。AUC 指标在本实验中被广泛用于比较不同异常检测模型的整体性能。通过计算不同模型的 AUC 值, 我们能够定量分析它们在高压并联电抗器声音信号检测任务中的表现。

p-AUC (Partial Area Under the Curve) 值: p-AUC 值是对 AUC 的进一步细化。它关注的是 ROC 曲线中 FPR 小于某个特定阈值时的曲线下面积。通常在实际应用中, 我们对异常检测模型的要求是, 假阳性率应尽可能低。因此, p-AUC 值能够更好地反映模型在低 FPR 区域的性能, 即在严格条件下模型的异常检测能力。在本实验中, 我们分别计算了 FPR 小于 0.3、0.2 和 0.1 时的 p-AUC 值。这些 p-AUC 值能够帮助我们评估模型在不同假阳性率要求下的性能表现, 从而更加准确地反映模型在实际应用中的检测能力。

5.4 Experimental Results

本节将详细介绍基于 Soft-LRAAD 模型在高压并联电抗器声音信号异常检测任务中的实验结果, 并与其他传统模型和深度学习模型进行对比分析。我们主要关注模型的检测精度、鲁棒性以及在不同负载条件下的表现。为了量化各模型的性能, 本文使用了 ROC 曲线、AUC 值和 p-AUC 值等评估指标。实验结果显示, Soft-LRAAD 模型在多个实验数据集上均表现出优异的检测性能。

Figure 5展示了 Soft-LRAAD 模型与其他模型在异常检测任务上的 ROC 曲线。实验中, Soft-LRAAD 模型的 ROC 曲线明显优于传统方法和其他深度学习模型, 曲线整体上更接近左上角, 表明其在不同阈值下均能有效区分正常数据和异常数据。在六个不同的数据集上, Soft-LRAAD 模型的 ROC 曲线始终处于其他模型之上, 表现出更高的真阳性率 (TPR) 和更低的假阳性率 (FPR)。这表明, Soft-LRAAD 模型在检测电抗器运行中产生的异常声音信号时, 能够更准确地识别出异常, 并有效降低误报率。

Table 2、3、4和 5分别展示了不同模型在不同数据集上的 AUC 值和 p-AUC 值。从表中可以看出, Soft-LRAAD 模型在大多数数据集上均取得了最高的 AUC 值, 表明其在异常检测任务中具有更强的分类能力。此外, Soft-LRAAD 模型在 FPR 小于 0.3、0.2 和 0.1 时的 p-AUC 值也明显优于其他模型, 表明其在低 FPR 区域的性能更加出色。这些结果表明, Soft-LRAAD 模型在高压并联电抗器声音信号异常检测任务中具有更高的检测精度和鲁棒性, 能够更好地识别正常数据和异常数据。

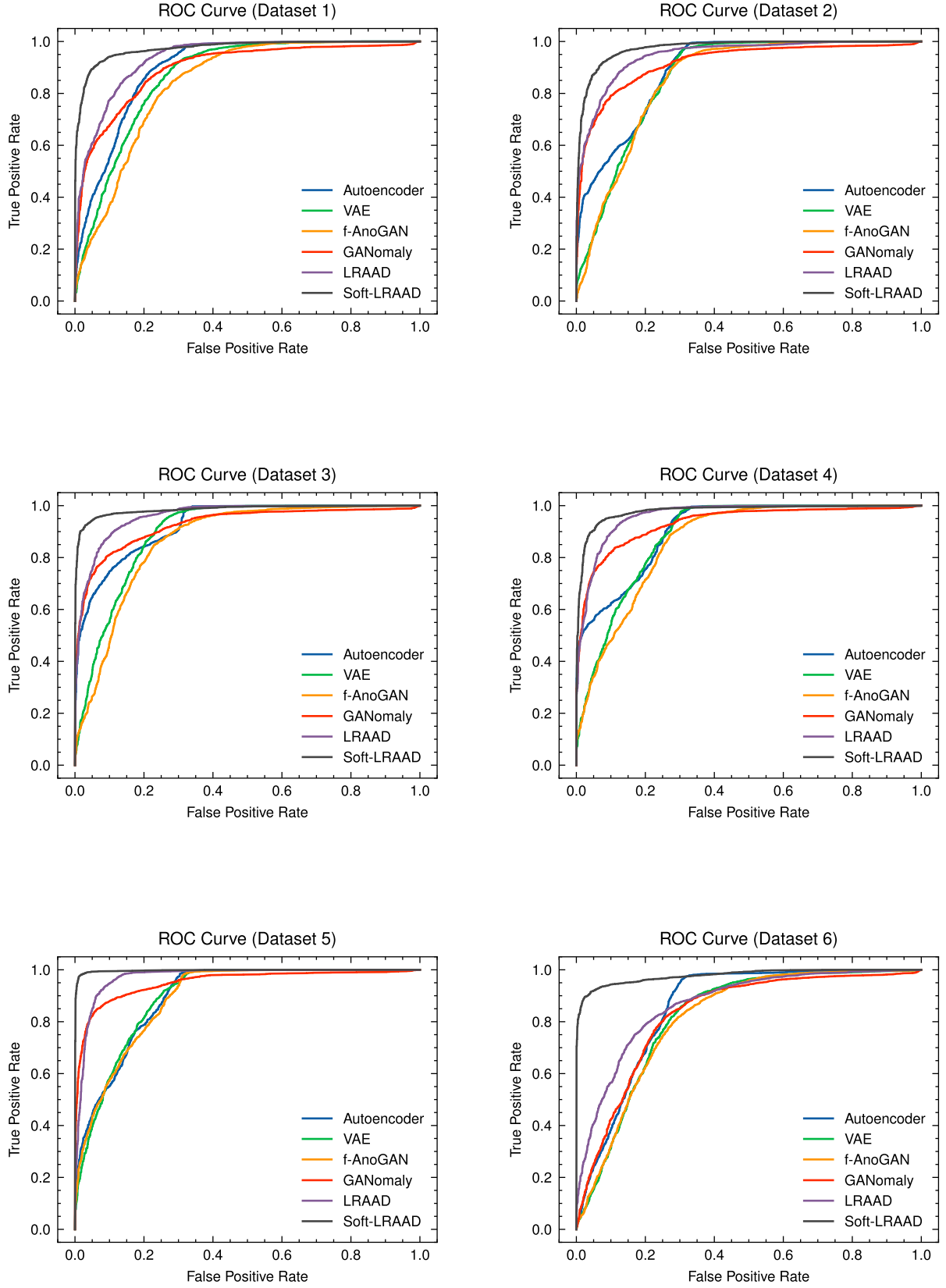


Figure 5: ROC Curves of Different Models on Different Datasets

Table 2: AUC Comparison of Different Models on Different Datasets

Dataset	CAE	VAE	f-AnoGAN	GANomaly	LRAAD	Soft-LRAAD(Our)
Dataset1	0.8975	0.8662	0.8400	0.8966	0.9361	0.9742
Dataset2	0.8911	0.8652	0.8560	0.9195	0.9475	0.9755
Dataset3	0.9266	0.8958	0.8662	0.9266	0.9607	0.9851
Dataset4	0.9087	0.8868	0.8654	0.9333	0.9626	0.9788
Dataset5	0.8962	0.8965	0.8910	0.9511	0.9721	0.9975
Dataset6	0.8523	0.8178	0.8126	0.8270	0.8674	0.9782

Table 3: p-AUC Comparison of Different Models on Different Datasets (FPR<0.3)

Dataset	CAE	VAE	f-AnoGAN	GANomaly	LRAAD	Soft-LRAAD(Our)
Dataset1	0.6664	0.5824	0.5161	0.7232	0.7930	0.9162
Dataset2	0.6394	0.5559	0.5487	0.7843	0.8430	0.9152
Dataset3	0.7605	0.6576	0.5935	0.8076	0.8639	0.9538
Dataset4	0.6972	0.6252	0.5790	0.8164	0.8342	0.9118
Dataset5	0.6529	0.6569	0.6423	0.8705	0.8866	0.8794
Dataset6	0.5262	0.4656	0.4589	0.5243	0.6335	0.9377

Table 4: p-AUC Comparison of Different Models on Different Datasets (FPR<0.2)

Dataset	CAE	VAE	f-AnoGAN	GANomaly	LRAAD	Soft-LRAAD(Our)
Dataset1	0.5431	0.4503	0.3809	0.6417	0.7078	0.8706
Dataset2	0.5355	0.4192	0.4037	0.7324	0.7865	0.8798
Dataset3	0.7041	0.5225	0.4598	0.7496	0.8176	0.9399
Dataset4	0.6142	0.5014	0.4532	0.7609	0.8173	0.9108
Dataset5	0.5465	0.5390	0.5371	0.8332	0.8115	0.8306
Dataset6	0.3801	0.3190	0.3227	0.3935	0.5377	0.9097

Table 5: p-AUC Comparison of Different Models on Different Datasets (FPR<0.1)

Dataset	CAE	VAE	f-AnoGAN	GANomaly	LRAAD	Soft-LRAAD(Our)
Dataset1	0.3640	0.2736	0.2261	0.5268	0.5629	0.8279
Dataset2	0.4405	0.2472	0.2318	0.6353	0.6658	0.8068
Dataset3	0.6024	0.3372	0.2651	0.6810	0.7020	0.9167
Dataset4	0.5487	0.3345	0.3136	0.6635	0.6993	0.8059
Dataset5	0.4090	0.3700	0.3964	0.7626	0.7400	0.6544
Dataset6	0.2168	0.1566	0.1624	0.2330	0.3791	0.8936

6 Conclusion

在本研究中，我们提出了一种基于软约束潜在正则化对抗学习的高压并联电抗器声音信号异常检测方法 (Soft-LRAAD)。该方法旨在解决现有异常检测技术在面对复杂工况下时的灵敏度和鲁棒性不足的问题。通过引入软约束损失函数，Soft-LRAAD 模型在潜在空间中实现了对正常数据和异常数据的更有效区分，并在模型训练的稳定性和检测精度方面取得了显著提升。

实验结果表明，Soft-LRAAD 方法在不同负载和工况条件下均表现出优异的异常检测性能。相比于传统的油色谱分析、超声波检测和高频局放检测等方法，Soft-LRAAD 不仅提高了检测的实时性和精度，还能够更好地应对复杂电磁环境中的干扰问题。此外，与其他深度学习模型相比，Soft-LRAAD 在 AUC 和 p-AUC 等指标上取得了更高的分数，显示出其在低假阳性率区域的检测优势。

References

- [1] DAS S, SIDHU T S, ZADEH M R D, et al. A novel hybrid differential algorithm for turn to turn fault detection in shunt reactors[J]. IEEE Transactions on Power Delivery, 2017, 32(6): 2537-2545.
- [2] MAGDALENO-ADAME S, ESCARELA-PEREZ R, OLIVARES-GALVAN J C, et al. Temperature reduction in the clamping bolt zone of shunt reactors: Design enhancements[J]. IEEE Transactions on Power Delivery, 2014, 29(6): 2648-2655.
- [3] TÜMAY M, DEMIRDELEN T, BAL S, et al. A review of magnetically controlled shunt reactor for power quality improvement with renewable energy applications[J]. Renewable and Sustainable Energy Reviews, 2017, 77: 215-228.
- [4] YAO C, ZHAO Z, LI C, et al. Noninvasive method for online detection of internal winding faults of 750 kv ehv shunt reactors[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2015, 22(5): 2833-2840.
- [5] VELÁSQUEZ R M A, LARA J V M. Root cause analysis for shunt reactor failure in 500 kv power system[J]. Engineering Failure Analysis, 2019, 104: 1157-1173.

- [6] ALI M S, OMAR A, JAAFAR A S A, et al. Conventional methods of dissolved gas analysis using oil-immersed power transformer for fault diagnosis: A review[J]. *Electric Power Systems Research*, 2023, 216: 109064.
- [7] WANG Y, LI X, GAO Y, et al. Partial discharge ultrasound detection using the sagnac interferometer system[J]. *Sensors*, 2018, 18(5): 1425.
- [8] ZHENG S, CUI Y, QI B. Feasibility of detecting partial discharge in smoothing reactor using uhf method[C]//*Proceedings of the 21st International Symposium on High Voltage Engineering: Volume 1*. Springer, 2020: 993-1002.
- [9] ILKHECHI H D, SAMIMI M H. Applications of the acoustic method in partial discharge measurement: A review[J]. *IEEE Transactions on Dielectrics and Electrical Insulation*, 2021, 28(1): 42-51.
- [10] GAO S, XING C, ZHANG Z, et al. Early warning of high-voltage reactor defects based on acoustic-electric correlation[J]. *Energies*, 2022, 15(19): 7196.
- [11] MNASRI Z, ROVETTA S, MASULLI F. Anomalous sound event detection: A survey of machine learning based methods and applications[J]. *Multimedia Tools and Applications*, 2022, 81(4): 5537-5586.
- [12] JOMBO G, ZHANG Y. Acoustic-based machine condition monitoring—methods and challenges[J]. *Eng*, 2023, 4(1): 47-79.
- [13] PANG G, SHEN C, CAO L, et al. Deep learning for anomaly detection: A review[J]. *ACM computing surveys (CSUR)*, 2021, 54(2): 1-38.
- [14] CHALAPATHY R, CHAWLA S. Deep learning for anomaly detection: A survey[A]. 2019.
- [15] SHI Z, ZHAO Q, SU L, et al. Equipment anomaly detection in power grids using deep learning [C]//*2021 International Conference on Intelligent Computing, Automation and Systems (ICICAS)*. IEEE, 2021: 271-275.
- [16] SUEFUSA K, NISHIDA T, PUROHIT H, et al. Anomalous sound detection based on interpolation deep neural network[C]//*ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020: 271-275.
- [17] HUANG C, YANG Z, WEN J, et al. Self-supervision-augmented deep autoencoder for unsupervised visual anomaly detection[J]. *IEEE Transactions on Cybernetics*, 2021, 52(12): 13834-13847.
- [18] DUMAN T B, BAYRAM B, İNCE G. Acoustic anomaly detection using convolutional autoencoders in industrial processes[C]//*14th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2019) Seville, Spain, May 13–15, 2019, Proceedings 14*. Springer, 2020: 432-442.

- [19] AN J, CHO S. Variational autoencoder based anomaly detection using reconstruction probability [J]. Special lecture on IE, 2015, 2(1): 1-18.
- [20] WU Y, CARSAULT T, NAKAMURA E, et al. Semi-supervised neural chord estimation based on a variational autoencoder with latent chord labels and features[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020, 28: 2956-2966.
- [21] BAYRAM B, DUMAN T B, INCE G. Real time detection of acoustic anomalies in industrial processes using sequential autoencoders[J]. Expert Systems, 2021, 38(1): e12564.
- [22] ERFANI S M, RAJASEGARAR S, KARUNASEKERA S, et al. High-dimensional and large-scale anomaly detection using a linear one-class svm with deep learning[J]. Pattern Recognition, 2016, 58: 121-134.
- [23] SCHLEGL T, SEEBÖCK P, WALDSTEIN S M, et al. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery[C]//International conference on information processing in medical imaging. Springer, 2017: 146-157.
- [24] SCHLEGL T, SEEBÖCK P, WALDSTEIN S M, et al. f-anogan: Fast unsupervised anomaly detection with generative adversarial networks[J]. Medical image analysis, 2019, 54: 30-44.
- [25] AKCAY S, ATAPOUR-ABARGHOUEI A, BRECKON T P. Ganomaly: Semi-supervised anomaly detection via adversarial training[C]//Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part III 14. Springer, 2019: 622-637.
- [26] JIANG A, ZHANG W Q, DENG Y, et al. Unsupervised anomaly detection and localization of machine audio: a gan-based approach[C]//ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023: 1-5.