

中国科学技术大学

专业硕士学位论文

(专业学位)



基于声纹识别的输电铁塔螺栓松动检测 方法研究

作者姓名： 何创创

专业领域： 计算机技术

校内导师： 赵云波 教授 吕文君 副教授

实践导师： 李泽瑞 副研究员

完成时间： 二〇二六年五月八日

University of Science and Technology of China
A dissertation for master's degree

(Professional degree)



**Research on Bolt Loosening
Detection Methods for Transmission
Towers Based on Voiceprint
Recognition**

Author: Chuangchuang He

Speciality: Computer Technology

Supervisors: Prof. Yun-Bo Zhao, Prof. Wenjun Lv

Practice Supervisor: Assoc. Res. Fellow Zerui Li

Finished time: May 8, 2026

中国科学技术大学学位论文原创性声明

本人声明所呈交的学位论文，是本人在导师指导下进行研究工作所取得的成果。除已特别加以标注和致谢的地方外，论文中不包含任何他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的贡献均已在论文中作了明确的说明。

作者签名：_____

签字日期：_____

中国科学技术大学学位论文授权使用声明

作为申请学位的条件之一，学位论文著作权拥有者授权中国科学技术大学拥有学位论文的部分使用权，即：学校有权按有关规定向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅，可以将学位论文编入《中国学位论文全文数据库》等有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。本人提交的电子文档的内容和纸质论文的内容相一致。

控阅的学位论文在解密后也遵守此规定。

☒ 公开 ☐ 控阅（____年）

作者签名：_____

导师签名：_____

签字日期：_____

签字日期：_____

摘 要

螺栓作为输电铁塔角钢构件连接的核心紧固元件，其预紧状态直接关系到杆塔结构安全与线路稳定运行。受长期野外环境温度循环、风载振动及腐蚀等多因素耦合影响，螺栓连接易发生预紧力衰减，进而演化为松动甚至脱落，严重时可能引发杆塔失稳事故。由于传统接触式或视觉巡检方法在复杂空间中存在操作受限和盲区问题，近年来，结合深度学习与声学特征的结构健康监测方法逐渐成为研究热点。然而，在电力巡检实践中，野外极端环境噪声对声学特征稳定性的干扰，以及高质量标注样本匮乏导致的模型退化，严重制约了传统数据驱动方法在复杂工况下的识别精度与泛化能力。基于此，探索一种具备强抗噪能力与少样本跨场景学习特性的声纹检测方法，对于电力场景的高效巡检具有重要意义。

针对上述挑战，本文提出一种基于可调频声激励与声纹特征的输电铁塔螺栓松动检测与判别方法。通过构建多尺度能量注意力机制与半监督迁移学习框架，实现了极低标注比例下跨塔型场景的稳健识别与位置区分。本文的主要研究内容如下：

(1) 针对现场研究数据匮乏问题，构建了覆盖三处真实输电铁塔场景的螺栓声纹数据库，并确立了标准化巡检采样方案。通过采集多工况、不同松动形态下的声纹样本，评估了传感器布设方式对数据稳定性的影响；进一步提出基于稳态时间窗与幅值统计的共振频率识别方法，为声激励频率的优化选取提供了依据；同时，通过时频分析显示，紧固状态下声纹能量高度集中，而松动状态在特定频段呈现显著扩散特征。该数据库的建立为后续模型训练与性能评测提供了可靠的标准化支撑。

(2) 针对野外强噪声与工况变化导致特征易被淹没的问题，提出端到端多尺度能量注意力残差网络 (MSEAM-Net)。该网络以梅尔频谱图为输入，通过频带能量引导的多尺度注意力机制对关键频带进行自适应增强，并结合随机噪声扰动与一致性正则化提升模型泛化能力，从而实现复杂噪声环境下螺栓状态的鲁棒表征与判别。实验结果表明，该方法在多座铁塔复杂工况数据集上取得更优且更稳定的识别性能；在 500 kV 酒杯塔数据上，模型准确率与 F_1 分数相较于基线分别提升 11.09% 和 12.5%，验证了所提方法能够抑制真实环境中噪声与工况变化对声纹特征的干扰，有效提高了对螺栓状态的判别能力。

(3) 针对标签获取成本高、样本不足导致模型泛化能力受限的问题，构建了一种基于结构复杂度自适应四元组损失的半监督迁移学习框架。该方法通过联合源域标注数据与目标域少量标注及大量未标注数据进行协同优化，在两座 500 kV 输电铁塔的跨塔实验中，所提出的方法在不同标注比例下均表现出稳定

优势，尤其在 5%–10% 的极低标注条件下，准确率提升分别达到 5.57% 和 4.48%，验证了所提方法在少样本跨塔型巡检中的有效性与工程应用潜力，显著降低了现场标注成本。

关键词：输电铁塔；螺栓松动检测；声纹识别；可调频声激励；多尺度注意力；半监督迁移学习；度量学习

ABSTRACT

As the core fastening elements connecting angle-steel members in transmission towers, the preload condition of bolts is directly related to the structural safety of towers and the stable operation of transmission lines. Under the coupled effects of long-term outdoor temperature cycling, wind-induced vibration, corrosion, and other environmental factors, bolted joints are prone to preload attenuation, which may further develop into loosening or even detachment and, in severe cases, lead to tower instability accidents. Owing to the operational limitations and blind spots of traditional contact-based or visual inspection methods in complex spatial environments, structural health monitoring methods combining deep learning and acoustic features have gradually become a research hotspot in recent years. However, in practical power inspection scenarios, the stability of acoustic features is severely affected by extreme outdoor environmental noise, while the scarcity of high-quality labeled samples results in model degradation, both of which significantly restrict the recognition accuracy and generalization capability of conventional data-driven methods under complex operating conditions. Therefore, exploring an acoustic-signature-based detection method with strong noise robustness and few-shot cross-scene learning capability is of great significance for efficient inspection in power system applications.

To address the above challenges, this thesis proposes a transmission tower bolt looseness detection and identification method based on tunable-frequency acoustic excitation and acoustic signature features. By constructing a multi-scale energy attention mechanism and a semi-supervised transfer learning framework, robust recognition and location discrimination across different tower types are achieved under extremely low labeling ratios. The main contributions of this thesis are as follows:

(1) To address the scarcity of field data, a bolt acoustic signature database covering three real transmission tower sites was established, together with a standardized inspection and sampling protocol. By collecting acoustic samples under multiple operating conditions and different loosening states, the influence of sensor deployment strategies on data stability was evaluated. Furthermore, a resonance-frequency identification method based on steady-state time windows and amplitude statistics was proposed, which provides a basis for the optimal selection of acoustic excitation frequencies. In addition, time–frequency analysis showed that the acoustic energy under the tightened state was highly concentrated, whereas the loosened state exhibited significant energy

dispersion in specific frequency bands. The established database provides reliable and standardized support for subsequent model training and performance evaluation.

(2) To address the problem that discriminative features are easily submerged by strong outdoor noise and varying operating conditions, an end-to-end Multi-Scale Energy Attention Residual Network (MSEAM-Net) was proposed. Taking Mel spectrograms as input, the network employs a frequency-band-energy-guided multi-scale attention mechanism to adaptively enhance critical frequency bands, and combines random noise perturbation with consistency regularization to improve model generalization, thereby enabling robust representation and discrimination of bolt states in complex noisy environments. Experimental results show that the proposed method achieves superior and more stable recognition performance on datasets collected from multiple transmission towers under complex operating conditions. On the 500 kV goblet-shaped tower dataset, the proposed model improves accuracy and F_1 -score by 11.09% and 12.5%, respectively, compared with the baseline, demonstrating its effectiveness in suppressing the interference of real environmental noise and operating-condition variations on acoustic signature features and in enhancing bolt-state discrimination.

(3) To overcome the limitations in model generalization caused by the high cost of annotation and insufficient labeled samples, a semi-supervised transfer learning framework based on a structure-complexity-adaptive quadruplet loss was developed. By jointly optimizing labeled source-domain data with a small amount of labeled target-domain data and a large amount of unlabeled target-domain data, the proposed method achieved consistent advantages under different labeling ratios in cross-tower experiments involving two 500 kV transmission towers. In particular, under extremely low labeling conditions of 5%–10%, the accuracy improvements reached 5.57% and 4.48%, respectively. These results verify the effectiveness and engineering applicability of the proposed method for few-shot cross-tower inspection and significantly reduce the cost of on-site annotation.

Key Words: Transmission tower; Bolt looseness detection; Acoustic signature recognition; Tunable-frequency acoustic excitation; Multi-scale attention; Semi-supervised transfer learning; Metric learning

目 录

第 1 章 绪论	1
1.1 研究背景与意义	1
1.2 国内外研究现状	3
1.2.1 螺栓松动检测方法研究现状	3
1.2.2 多尺度注意力机制研究现状	6
1.2.3 半监督学习、迁移学习与度量学习研究现状	7
1.2.4 研究现状小结	8
1.3 研究内容与组织安排	9
1.3.1 研究内容	9
1.3.2 研究创新点	10
1.3.3 论文组织结构安排	10
第 2 章 相关基础知识	12
2.1 铁塔螺栓松动声纹识别原理	12
2.1.1 声纹信号的物理特性与诊断可观测性	12
2.1.2 从时域信号到可学习特征的表示方法	13
2.1.3 基于声纹信号的螺栓松动诊断流程	14
2.2 注意力机制	14
2.3 半监督学习	17
2.4 迁移学习、域自适应与半监督域自适应	18
2.5 度量学习	20
第 3 章 多场景螺栓声纹数据库构建与分析	22
3.1 引言	22
3.2 数据采集设备与场景描述	23
3.2.1 数据采集设备	23
3.2.2 数据采集场景	23
3.3 声激励装置与传感器布局分析	25
3.3.1 布设方案与实验设计	25
3.3.2 布局差异的特征建模与距离评价	26
3.3.3 布局实验结果分析	30
3.4 铁塔共振频率分析	32
3.4.1 基于幅值特征的共振频率确定方法	32

3.4.2 共振频率的实验设置与结果分析	33
3.5 基于声纹信号的铁塔螺栓松动与紧固特征分析	37
3.5.1 基于孤立森林的异常点检测	38
3.5.2 螺栓松动声纹特征对比分析	41
3.6 本章小结	45
第 4 章 基于多尺度能量注意力残差网络的螺栓松动识别方法	46
4.1 引言	46
4.2 现场采集信号干扰类型、特征及影响分析	47
4.3 多尺度能量注意力残差网络的方法设计	48
4.3.1 MSEAM-Net 的整体框架	48
4.3.2 基于噪声增强的一致性正则化训练	53
4.4 实验设计与结果分析	54
4.4.1 实验数据集与预处理	54
4.4.2 实验环境设置与评价指标	56
4.4.3 实验结果分析	57
4.5 本章小结	69
第 5 章 基于半监督迁移学习的少样本螺栓松动检测方法	70
5.1 引言	70
5.2 基于半监督迁移学习的螺栓松动检测方法设计	71
5.2.1 问题定义	71
5.2.2 结构复杂度自适应四元组损失设计	71
5.2.3 半监督迁移学习方法设计	74
5.3 实验设计与结果分析	77
5.3.1 实验数据集与预处理	77
5.3.2 实验环境设置与评价指标	78
5.3.3 实验结果分析	79
5.4 本章小结	87
第 6 章 总结与展望	88
6.1 工作总结	88
6.2 研究展望	90
参考文献	91
致谢	97
在读期间发表的学术论文与取得的研究成果	98

插图清单

图 1.1	输电铁塔螺栓松动导致的典型结构失效示意图	1
图 1.2	现有人工接触式输电铁塔螺栓紧固状态巡检	2
图 1.3	研究内容框架图	10
图 2.1	基于声纹信号的铁塔螺栓故障诊断流程	14
图 3.1	现场数据采集设备组成示意图	23
图 3.2	声激励装置与骨传导传感器安装示意图	26
图 3.3	声激励装置与传感器现场布局编号	27
图 3.4	不同工况下紧固与松动状态的 t-SNE 可视化	30
图 3.5	不同状态下的激发频率—幅值关系图	35
图 3.6	紫蓬 220kV 送出工程部分铁塔激发频率与幅值关系图	37
图 3.7	基于 t-SNE 与孤立森林的无标签音频异常检测流程	39
图 3.8	t-SNE 可视化与异常点检测结果	40
图 3.9	10 座铁塔音频信号语谱图对比	40
图 3.10	铁塔 (B2、B8、B20、C8、A20、B19) 音频信号频谱图对比	41
图 3.11	不同状态下的特征对比: 第一行为频谱图 (FFT), 第二行为语谱图 (Spectrogram)。	42
图 3.12	中部松动螺栓数量为 2、4、6 颗时的特征对比。上排为频谱图 (FFT), 下排为语谱图 (Spectrogram), 从左到右依次对应松动螺栓数量为 2 颗、4 颗、6 颗。	43
图 3.13	中部松动与顶部松动状态的特征对比。第一行为频谱图 (FFT), 第二行为语谱图 (Spectrogram)。	44
图 3.14	螺栓不同状态激励幅值频谱图	44
图 3.15	音频信号的时域与频域可视化分析	45
图 4.1	MSEAM 结构图	51
图 4.2	MSEAM-Net 网络结构图 (其中 Conv 表示卷积层, BN 表示批归一化, Dropout2d 表示二维随机失活层, GAP 表示全局平均池化。)	52
图 4.3	500 kV 干字塔和酒杯塔松动位置采集点示意图	56
图 4.4	不同注意力机制在 500 kV 酒杯塔数据集的分类结果	59
图 4.5	原始 log-Mel 谱图与残差谱图	63
图 4.6	原始谱图与残差谱图输入下的性能对比 (分类准确率/F1 分数)	63

图 4.7	多螺栓松动位置采集示意图	65
图 4.8	500 kV 干字塔各工况下螺栓松动可视化	65
图 4.9	500 kV 干字塔不同工况分类混淆矩阵	66
图 4.10	不同扭矩值下的螺栓松动可视化	67
图 4.11	不同扭矩值分类的混淆矩阵	68
图 5.1	伪标签引导的四元组挖掘与结构复杂度自适应损失	72
图 5.2	半监督迁移学习框架 SSTF	74
图 5.3	两座铁塔的数据采集位置和传感器/声学激励装置。	78
图 5.4	在目标域不同标注比例下, 各对比方法在目标域测试集上的平均 分类准确率	80
图 5.5	目标域 10% 标注条件下八种方法的混淆矩阵对比: (a) S+T; (b) SSDA; (c) ENT; (d) MME; (e) FADA; (f) CDAC; (g) APE; (h) 本文方法。	82
图 5.6	模型 1 (不使用 MSEAM 和 SQL) 和模型 4 (使用 MSEAM 和 SQL) 在目标域 10% 标注率下进行 5 次重复实验的分类结果箱线图。 . .	83
图 5.7	目标域 10% 标注率条件下测试集嵌入特征的 t-SNE 可视化对比。 .	84
图 5.8	目标域 10% 标注率条件下的混淆矩阵对比: 左为 Model2 (不含 SQL), 右为 Model4 (引入 SQL)。	84
图 5.9	500 kV 酒杯塔目标域样本的 Mel 频谱图对比: (a) 紧固状态; (b) 4 号位置能量衰减状态 (松动); (c) 4 号位置松动状态。	85
图 5.10	目标域标注比例为 10% 时自适应间隔参数 α/β 的均值随训练轮次 变化曲线 ($\alpha_0 = 0.2$, $\beta_0 = 0.4$, $\gamma = 1$)。	86
图 5.11	目标域标注比例为 10% 时, 不同 γ 取值下的 Top-1 准确率 (Top-1 Accuracy)、召回率 (Recall)、F1 分数 (F1 Score) 以及类别 4 的 召回率 (Recall_4) 与 F1 分数 (F1 Score_4)。	86

表 格 清 单

表 3.1	紫蓬送出工程部分塔型参数 ·····	24
表 3.2	张洼实训基地实验铁塔相关信息 ·····	24
表 3.3	宏源工业区实验铁塔相关信息 ·····	25
表 3.4	不同布设方式之间的特征差异（余弦距离） ·····	31
表 3.5	不同布设方式的类内离散度 $S_{\text{intra}}(g)$ 对比 ·····	31
表 4.1	现场采集信号主要干扰类型及其影响 ·····	48
表 4.2	MSEAM-Net 网络结构与参数配置（输入： $1 \times 64 \times 64$ ） ·····	52
表 4.3	500 kV 干字塔数据采集信息 ·····	55
表 4.4	500 kV 酒杯塔数据采集信息 ·····	55
表 4.5	220 kV 酒杯塔数据采集信息 ·····	56
表 4.6	不同方法在三个数据集上的性能对比（分类准确率：mean±std, F1 分数：mean）。 ·····	58
表 4.7	不同模型复杂度与推理效率对比 ·····	61
表 4.8	对 MSEAM 和一致性正则化 (CR) 的消融研究 ·····	62
表 5.1	不同塔型与螺栓状态的声学数据集样本分布 ·····	79
表 5.2	目标域不同标注比例下各方法的平均分类准确率 (%) ·····	81
表 5.3	10% 标注率下 MSEAM 与 SQL 的消融实验结果（单位：%） ·····	83

第1章 绪 论

1.1 研究背景与意义

随着我国经济的持续增长与能源结构的转型升级，电力系统的安全稳定运行对于保障国家能源安全及经济社会发展具有重要意义。截至 2024 年底，全国电网 220 千伏及以上输电线路回路长度已达到 96 万千米^[1]，这一规模的扩张不仅表明了电网基础设施建设的成效，也意味着作为输电线路主要支撑结构的输电铁塔数量持续攀升。根据《架空输电线路铁塔技术发展报告（2021 年）》统计，输电铁塔在我国铁塔制造行业中的占比约为 85%–90%。作为电力系统的重要组成部分，输电铁塔的结构安全性直接决定了电网的运行可靠性，而螺栓松动是引发铁塔结构失稳的主要诱因之一。如图 1.1 所示，尽管电网运维部门持续开展定期巡检与紧固作业，但由于输电铁塔长期暴露于野外复杂环境，螺栓连接在温度交变、风荷载、导线振动、雨雾腐蚀及人为扰动等多种复杂环境因素的长期共同作用下，极易发生连接界面摩擦系数衰减、微小间隙累积及局部塑性变形。这些微观损伤的积累将不可避免地导致螺栓预紧力衰退，进而演化为螺栓松动甚至脱落。该问题在“三跨”区段（即跨越铁路、一级及以上公路和重要输电通道）尤为突出。

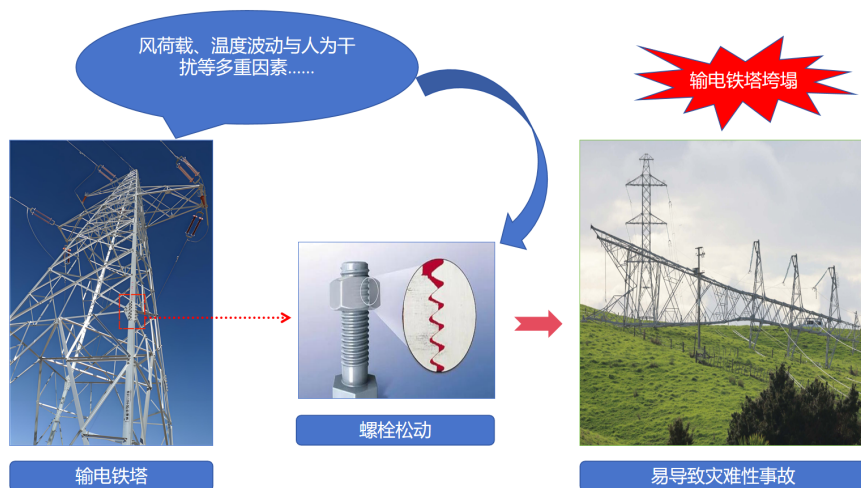


图 1.1 输电铁塔螺栓松动导致的典型结构失效示意图

螺栓松动对输电系统的危害具有显著的隐蔽性与连锁反应特征。首先，节点连接刚度的降低将直接削弱输电铁塔的整体结构稳定性，导致承载能力下降，在极端气象条件下极易引发塔体倾覆、垮塌等灾难性事故。其次，连接状态的劣化会改变结构的动力学特性，加剧风致振动响应与疲劳损伤累积，长期作用下可能诱发共振效应，加速结构失效进程。此外，铁塔结构的异常位移或剧烈振动可能导致绝缘子串受力不均及电气间隙改变，进而增加漏电、闪络及电弧击穿

的风险，直接威胁电力设备与运维人员的安全。因此，螺栓松动虽属微小物理缺陷，但其演化后果严重，亟需研究高效、可靠的检测技术以支撑输电铁塔的状态评估与风险预警。

然而，目前输电铁塔的螺栓状态检测仍严重依赖人工运维。据统计，常规铁塔螺栓数量可达 3000 个以上，其中塔头及横担等关键受力部位的螺栓尤为密集，约占总量的 45% 以上。如图 1.2 所示，现有的“人工登塔 + 抽检”模式不仅覆盖率低，难以客观反映整塔的紧固状态，且存在高空作业风险大、效率低下的弊端，已无法满足海量存量杆塔的全面排查需求。从工程应用与经济效益角度来看，研发新型自动化检测技术具有显著价值。一方面，通过早期发现预紧力衰退或潜在松动隐患，可大幅降低结构失稳与线路停运风险，减少事故损失与抢修成本；另一方面，基于信号感知与智能诊断的非接触式检测流程，可在不改变电网运行方式的前提下实现全覆盖评估，显著降低人力投入与作业风险，推动运维模式从传统人工巡检，向更高效、更精准的状态化运维转变。



图 1.2 现有人工接触式输电铁塔螺栓紧固状态巡检

在当前工程需求的推动下，基于声学振动特征的非接触式检测技术逐渐成为研究的热点。螺栓松动，尤其是主材连接松动会改变结构的局部刚度及振动传递路径，从而引起声学响应在频带能量分布、谐波结构及宽带扩散特性上的显著变化，这为基于声纹信息的松动状态识别与定位提供了物理依据。尽管目前已有一些基于声学原理的监测系统处于研发阶段，但在实际工程应用中仍面临诸多挑战。特别是在塔头、横担等螺栓密集区域，不同位置的松动所引发的声学特征高度相似，且容易受到环境噪声的干扰，这使得现有方法在识别准确率和定位精度上存在瓶颈。近年来，机器学习等数据驱动方法在故障诊断领域取得了显著成果，凭借对大量数据规律的学习，展现出了较高的诊断精度和强大的鲁棒性。然而，由于输电铁塔的特殊功能，它们通常建造在偏远的野外环境中，获取研究数据需要专业运维人员的配合，且存在一定的危险性，导致获得准确标签数据的

难度极大。因此，针对复杂野外工况和样本获取困难的现实问题，亟需在数据表征、模型鲁棒性以及少样本学习策略等方面进行深入研究。

基于以上背景，本文旨在探索高精度、高效率、强鲁棒性的智能检测方法，为输电线路的智能化运维与全寿命周期管理提供理论依据与技术支持。

1.2 国内外研究现状

1.2.1 螺栓松动检测方法研究现状

螺栓连接作为机械装备、土木结构及电力基础设施中最为通用的连接形式，是保障结构完整性与力学传递性能的关键节点。螺栓松动是此类连接结构中极为常见且隐蔽性强的失效模式，其演化过程通常呈现出“预紧力衰退—接触界面微动—连接刚度退化”的渐进特征。随着损伤的累积，结构系统往往伴随能量耗散增加与模态参数漂移，若未能及时识别，极易诱发构件脱落甚至导致整体结构失稳倒塌。针对螺栓松动的监测与评估问题，国内外学者已开展大量研究。现有技术路线总体可归纳为直接测量法与间接无损检测（Non-Destructive Testing, NDT）法两大类。其中，直接测量法侧重于物理量的直接获取，而 NDT 则通过分析结构响应特征反演连接状态，涵盖了压电阻抗、振动分析、超声导波、机器视觉及声学特征识别等主流方向。上述方法在检测灵敏度、实时性、定位精度及工程部署成本等方面各具优劣。然而，具体到输电铁塔运维场景，受限于螺栓基数庞大、分布空间密集、野外环境噪声干扰强以及典型故障样本稀缺等多重约束，现有的通用检测手段难以同时满足高可靠性、低成本与精确定位的工程需求，相关技术仍面临严峻挑战。

（1）直接预紧力/扭矩检测

扭矩扳手、角度法与拉伸测量等传统预紧力检测方法能够在施工验收或检修抽检中提供较直观的判据。扭矩扳手结构简单、操作直观，长期以来是工程一线的常用工具。然而，在实际工况下，螺栓扭矩与预紧力的对应关系易受螺纹摩擦、接触面润滑状态、表面粗糙度及材料屈服程度等多种因素影响，二者难以建立稳定可靠的量化关系。同时，输电铁塔单塔螺栓数量往往达到数千个，塔头及横担等关键区域高空作业风险大，若完全依赖人工逐一检测将带来巨大的人力成本与安全风险。因此，该类方法更适用于离线抽检和工程验收，难以支撑大规模、常态化的状态监测。

（2）压电阻抗与主动传感方法

压电阻抗（Electromechanical Impedance, EMI）方法依托压电陶瓷与结构之间的机电耦合特性，通过监测电导纳或电阻抗随结构机械阻抗的变化，间接反映螺栓连接状态。1994年，Liang 等人首次将压电片与结构简化为单自由度弹簧-质

量-阻尼 (Spring-Mass-Damper, SMD) 模型, 建立了电导纳与机械阻抗之间的函数关系, 为后续螺栓连接监测研究奠定了坚实基础^[2]。此后, Sun 等人^[3]将压电阻抗方法引入桁架结构健康监测领域, 进而推动了大量针对螺栓连接状态识别的相关研究。近年来, EMI 方法进一步向量化、智能化与全阶段识别发展。例如, Nguyen 等将 EMI 与深度学习结合, 实现了螺栓松动状态的自动评估^[4]; Lu 和 Shen 提出了非线性机电阻抗谱方法, 实现了螺栓连接从初始松动到严重松动的全阶段监测^[5]。在此基础上, 主动应力波、尾波干涉等衍生方法, 对螺栓松动“萌芽阶段”具备较高的检测灵敏度, 展现出精细化监测的应用潜力^[6-10]。总体而言, EMI 及主动传感类方法对早期微小损伤的识别灵敏度较高, 但应用于输电铁塔等户外工程场景时仍存在明显局限: 一是传感器粘贴质量与耦合一致性、温度漂移及长期可靠性等因素, 会影响检测特征的稳定性^[11]; 二是大规模布设传感器的成本较高, 且后续维护负担较重; 三是在复杂节点与多路径传播的野外环境中, 检测阈值与判别标准易受工况干扰而产生波动。因此, 该类方法更适用于少量关键节点的长期跟踪监测, 难以满足全塔海量螺栓的普适性监测需求。相关综述也指出, 早期灵敏感知与大规模工程可实施性之间仍存在明显矛盾^[12]。

(3) 超声导波与非线性声学方法

超声导波 (如 Lamb 波) 与非线性声学方法, 通过分析声波传播过程中的散射、反射及能量耗散特征变化, 实现对螺栓连接紧固状态的评估。部分研究引入能量成像、连续小波变换或非线性调制指数等量化指标, 进一步实现了螺栓松动程度的定量判断^[13-14]。其中, 非接触式激光超声与波能量耗散等方案^[15], 有效降低了对传感器与结构耦合状态的依赖, 提升了检测的便捷性。总体而言, 该类方法具有物理可解释性强、定量监测潜力突出的优势, 但对传感器布置、激励与接收同步性及声波传播路径稳定性的要求较高。应用于输电铁塔复杂桁架结构时, 声波传播存在多路径、多模态干扰及边界反射明显等问题, 导致信号解析难度大, 检测结果的可重复性易受现场工况影响; 同时, 野外环境中的噪声、电磁干扰及施工约束等因素, 进一步降低了其工程实用性。因此, 该类方法在实验室及可控结构场景中表现优异, 但在大尺度输电铁塔结构上的规模化部署, 仍需开展进一步的工程化改进工作。

(4) 振动-声学调制方法

基于振动-声学调制 (Vibration-Acoustic Modulation, VAM) 方法的检测路线, 并非直接以敲击声的整体模式差异作为判别依据, 而是从螺栓连接界面的接触非线性机理出发, 认为松动会导致连接刚度下降、摩擦滑移增强及能量耗散增加, 从而使低频振动激励与高频探测信号在接触界面处发生非线性耦合, 并在响应中形成边带、互调分量及调制畸变等可观测特征。围绕这一机理, 相关研究已提出基于频域离散度、边带特征、变分模态分解 (Variational Mode Decomposition,

VMD) 时频特征以及结合传统分类器或回归模型的技术策略, 实现对螺栓不同松动状态的有效识别^[16-20]。总体而言, VAM 方法更强调“松动界面非线性”这一物理本质, 具有机理清晰、对早期微弱松动敏感、可解释性较强以及具备在线监测潜力等优势。然而, VAM 路线对激励输入与传播条件的稳定性要求相对较高。已有研究表明, 界面弹塑性波行为及接触状态变化会引入额外不确定性, 从而影响 VAM 调制特征的稳定性与可重复性^[20]。对于输电铁塔这类大型空间桁架结构, 户外风载、导线舞动及复杂边界条件会进一步放大这种不确定性, 导致泵浦激励与探测响应波动, 削弱调制特征的一致性, 同时, 铁塔构件连接复杂、传播路径多变且衰减明显, 不同位置螺栓松动在观测端可能表现出相似的调制频带特征, 易造成相邻位置识别混淆和定位困难。尽管最新研究已通过扫频探测波、分数阶傅里叶变换滤波及机理特征与学习模型融合等方式提升了 VAM 特征提取的灵敏度与鲁棒性^[21-23], 但在面向真实铁塔场景的规模化应用中, 仍需进一步结合稳健激励设计、更高效的抗扰特征提取方法及多位置联合定位策略, 才能充分释放该路线在复杂工程环境中的应用潜力。

(5) 视觉检测与图像智能识别方法

随着深度学习在目标检测领域的突破, 视觉检测凭借非接触、可视化的优势, 成为螺栓状态监测的重要技术方向, 典型技术思路包括基于螺栓与螺母的相对旋转角度识别、标记编码(如 PAC 码)识别、RCNN 目标检测及几何参数估计等^[24-29]。该路线的突出优势是易与机器人、无人机巡检平台结合, 可实现大范围、高效化的批量检测, 但在输电铁塔实际场景中, 仍面临诸多现实局限: 螺栓多分布于塔头、横担等密集区域, 易被构件遮挡, 导致检测视角受限; 户外光照变化、雨雾、灰尘等环境因素, 会严重影响图像清晰度; 此外, 螺栓早期预紧力衰退阶段, 往往不会表现出明显的外观变化, 导致视觉检测对早期松动的识别敏感性不足, 难以满足工程上“早发现、早预警”的需求。因此, 视觉检测更适用于可见性强、无明显遮挡的螺栓部位, 难以实现全场景、全周期的松动监测。

(6) 声纹特征识别方法

声学检测方法以声信号为载体, 通过敲击或外部激励获取结构响应声, 并将其转换为功率谱、梅尔频率倒谱系数(Mel-Frequency Cepstral Coefficients, MFCC)、Mel 谱图等时频表征, 再结合机器学习或深度学习模型完成螺栓状态识别。与 VAM 路线侧重界面非线性调制机理不同, 声纹特征识别方法更侧重于将响应声本身视为“模式识别对象”, 通过构造有效的时频特征或深度表征来实现松动状态判别。该类方法具有非接触、部署成本低、易于快速巡检、适配户外场景等优势, 近年来在螺栓松动检测中受到广泛关注。早期研究多采用功率谱、MFCC 等手工提取特征, 结合传统机器学习分类器实现状态判别^[30-31]; 也有研究将 Mel 频谱应用于塔体巡检场景的螺栓松动识别, 并结合多通道采集与对抗训练增强

抗噪性能^[32]。此外,近年的研究还通过引入去噪、分段、平滑过滤、累计能量熵、相位特征以及注意力序列建模等策略,进一步提升了声学特征在复杂工况下的稳定性与识别能力。已有研究表明,相较于传统 MFCC 特征,这类改进方法在应对敲击力不一致、环境噪声干扰和未知异常状态时通常表现出更好的稳定性与适应性^[33-36]。不过,这一路线本质上仍属于“特征表征-模式分类”的数据驱动框架,对人工特征工程或训练数据分布具有较强依赖。当输电铁塔场景发生变化,如塔型不同、噪声环境改变、敲击力度波动、声波传播路径差异或节点边界条件变化时,同一松动状态下的声纹分布往往会显著离散,不同位置之间的类间边界也会变得模糊,导致模型泛化能力下降。因此,面对输电铁塔“跨塔型、多噪声、弱差异”的复杂工程需求,单纯依赖手工声学特征和浅层分类器已难以满足实际应用要求,仍需结合更稳健的时频表征、深度特征学习与跨场景泛化策略,进一步提升声纹识别方法在真实巡检场景中的可靠性与工程适用性。

1.2.2 多尺度注意力机制研究现状

近年来,深度学习凭借端到端的自动表征学习能力,在工业故障诊断领域快速发展^[37-38]。相较于传统依赖手工特征的机器学习方法,深度网络能够自动挖掘信号深层判别信息,更好地适应复杂噪声环境与工况波动^[39-40]。针对机械振动与声学信号普遍存在的弱差异、强噪声、多尺度特性,现有研究已在轴承等典型部件诊断中证实,采用“多尺度感知+注意力加权”的网络结构,可通过多分支、多核尺寸并行卷积捕捉多尺度特征,并利用注意力机制聚焦关键频带与时频片段,抑制冗余与噪声干扰,从而提升特征表征有效性^[41-45]。与之类似,输电铁塔螺栓松动声纹同样具有差异微弱、背景噪声复杂、特征尺度分布宽泛等特点,故障相关的有效信息极易被环境噪声、传播衰减与结构耦合效应所掩盖。然而,直接套用面向常规旋转部件的多尺度注意力结构,难以匹配铁塔声纹信号的传播特性与细粒度定位需求,仍需结合任务场景开展针对性设计。

多尺度建模的基本思想是通过不同感受野并行或逐级提取特征,使网络同时捕捉细粒度的局部模式与跨频带的整体结构信息。相关研究普遍采用多分支卷积、金字塔结构或不同核尺寸的并行卷积来增强尺度鲁棒性,从而更稳定地刻画复杂背景下的故障表征^[45-47]。在此基础上,注意力机制进一步通过“加权选择”提升特征利用效率:通道注意力可强调对特定频带纹理更敏感的响应通道,而空间注意力则在“时间-频率平面”上突出关键能量块与结构纹理,使网络在多尺度特征融合时更具针对性。为获得更强的表达能力,一些工作还将多种注意力模块进行组合或级联,以丰富表征并提升噪声环境下的可分性^[48]。

尽管多尺度与注意力机制在轴承、电机等部件级故障诊断中已取得良好效果^[49-50],但面向输电铁塔螺栓松动的细粒度定位任务仍存在明显不足。一方面,

邻近螺栓位置引起的声学差异更加微弱且易被传播路径与结构耦合效应所稀释,通用注意力模块未必能稳定提取“位置敏感”的关键频带与局部纹理;另一方面,现有网络设计往往缺少与结构传播机理相结合的先验约束,导致模型可解释性不足、跨塔型泛化不稳定。因此,面向铁塔场景需要构建更贴合声学能量分布与跨尺度响应特征的注意力结构,并与后续的结构约束学习策略协同设计,以提升细粒度可分性与工程可信度。

1.2.3 半监督学习、迁移学习与度量学习研究现状

在输电铁塔螺栓检测的工程现场,高质量标注数据的获取不仅成本高昂,且难以实现规模化采集,与之相对的是,无标签数据的获取则更为便捷、充足。这一“样本稀缺”难题在智能故障诊断领域普遍存在,半监督学习(Semi-Supervised Learning, SSL)与迁移学习(Transfer Learning, TL)、域自适应(Domain Adaptation, DA)已逐渐成为应对该问题的核心研究方向^[51-54],这类方法虽最初多应用于不同的故障诊断场景,但为输电铁塔螺栓检测的样本稀缺问题提供了重要技术借鉴。具体而言,半监督学习通过“少量有标签数据+大量无标签数据”的联合训练模式,借助一致性约束、自训练等核心机制,引导模型学习更平滑、更具泛化性的决策边界;域自适应作为迁移学习的重要实现方式,通过特征对齐与决策边界调整,将源域已学习到的知识有效迁移至目标域。在此基础上,半监督域自适应(Semi-Supervised Domain Adaptation, SSDA)将两者优势相结合,可同时利用源域充足的标注信息与目标域少量标注、大量未标注数据,为跨结构、跨工况的输电铁塔螺栓检测工程实际场景提供了更具可行性的技术路径。

针对螺栓松动检测中“数据稀缺”这一核心难题,国内外学者已从数据增强、小样本学习等角度开展了相关探索与尝试。例如,有研究将音频增强技术与卷积神经网络相结合,在小样本条件下有效提升了螺栓松动识别的性能与抗干扰能力^[55];另有研究引入 APGDF 等抗噪特征,结合原型网络设计,缓解了噪声干扰与样本不足双重因素带来的模型性能退化问题^[56]。但此类方法存在明显局限,其研究场景多集中于单一结构或相对可控的实验室环境,对输电铁塔“跨塔型、跨工况”所引发的特征分布偏移考虑不够充分,导致模型在迁移至新塔型或新布设条件时,易出现性能显著退化的问题,难以适配复杂的工程实际需求。

为进一步减少模型对真实标注数据的依赖,部分研究尝试采用有限元仿真与实际监测数据相结合的方式,通过“仿真到真实”的迁移学习实现知识复用,以此弥补真实标注数据的不足^[57]。但由于仿真场景与野外实际工况在材料参数、边界条件、声波传播路径及噪声分布等方面存在显著差异,当目标域标注数据极少时,单纯依靠简单的迁移对齐策略,往往难以实现模型的稳定泛化,甚至可能出现负迁移现象,导致检测性能下降。与此同时,尽管 SSDA 在视觉领域已有较

为成熟的范式，但将其针对性应用于输电铁塔螺栓松动检测的相关研究仍相对匮乏。相较于通用视觉任务，输电铁塔螺栓松动检测不仅存在跨域分布偏移问题，还面临相邻螺栓位置细粒度混淆、声波传播路径复杂、野外环境噪声强烈等独特的结构性挑战，单纯依靠传统的分布对齐或伪标签自训练策略，难以保证螺栓不同状态类别的清晰分离，无法满足细粒度定位的工程需求。

值得注意的是，度量学习为解决细粒度故障诊断问题提供了极具潜力的补充途径。其中，基于三元组损失 (triplet loss)、四元组损失 (quadruplet loss) 的度量学习方法，可在特征嵌入空间中同时强化类内紧致性与类间分离性^[58-59]，该思路已在风电齿轮等工业部件的故障诊断任务中得到验证，证实了其有效性^[60-61]；另有研究表明，通过度量学习结构与可视化解释方法相结合，可实现对螺栓微小角度松动状态的精准识别^[62]。然而，在输电铁塔螺栓检测“跨塔型域偏移、标注样本极少、相邻位置细粒度分类”的三重约束下，如何将半监督域自适应与度量学习进行有效融合，在实现跨域特征对齐的同时，不破坏类别结构，且能进一步增强目标域的细粒度可分性，仍是目前亟待深入研究的关键问题。

1.2.4 研究现状小结

综上所述，国内外已形成多技术路线并存的螺栓松动检测体系，但各方法应用于输电铁塔场景时均存在明显局限：直接预紧力与扭矩检测虽适用于工程验收与抽检，却无法全塔海量螺栓高效、低风险的巡检需求；压电阻抗、超声导波等主动传感类方法对早期微松动灵敏度高、物理可解释性强，但依赖传感器布设与精细激励配置，受环境漂移、耦合一致性及施工条件影响较大，难以在复杂桁架铁塔上大范围稳定部署；振动-声学调制方法机理明确、具备在线监测潜力，但野外不确定激励与多路径传播会降低特征稳定性，在塔头、横担等密集节点易出现相邻位置定位混淆；视觉检测虽具备非接触、可视化优势且适配机器人与无人机巡检，却受遮挡、光照、雨雾尘等环境因素制约，对早期预紧力衰退的敏感性不足。相比之下，声纹检测路线在设备轻量性、布设灵活性与户外适配性上更具工程吸引力，但传统手工特征结合机器学习的模式对特征工程依赖度高，跨塔型、跨工况的泛化能力薄弱，难以适配输电铁塔“跨结构、多噪声、弱差异”的复杂场景。

在智能诊断技术层面，深度学习的端到端特征学习能力为解决上述难题提供了支撑，其中多尺度建模与注意力机制可在弱差异、强噪声背景下突出关键特征，提升表征质量；半监督学习与迁移学习/域自适应则有效应对了“少量标注+大量无标签”以及“源域与目标域分布偏移”的工程现实。但现有技术仍存在明显研究空白：多尺度注意力结构多面向通用部件级故障诊断，未针对输电铁塔螺栓松动“位置敏感、邻近类别差异极弱”的细粒度定位需求进行定制化设计；

跨塔型域偏移与极少标注条件下，单纯依靠分布对齐或伪标签自训练难以保证类结构清晰分离，易出现相邻位置识别混淆问题。

基于上述研究空白与工程实际需求，本文以声纹信息为检测载体，围绕输电铁塔螺栓松动检测“跨塔域偏移显著、环境噪声复杂、标注样本稀缺、需细粒度定位”的核心矛盾展开研究。在表征层面，设计贴合声纹能量分布与跨尺度响应特征的多尺度注意力结构，强化关键频带信息提取并抑制背景干扰；在学习策略层面，融合半监督迁移对齐与结构约束度量学习机制，充分利用无标签数据的同时强化类内紧致性与类间分离性，重点提升相邻螺栓的判别能力与模型跨塔型泛化稳定性，后续章节将通过实验严格验证所提方法的有效性与优越性。

1.3 研究内容与组织安排

1.3.1 研究内容

为实现对长期处于复杂野外环境中的输电铁塔结构健康状态的精准感知，本文围绕输电铁塔螺栓紧固状态检测的关键技术与工程应用中的优化问题，提出一种基于可调频声激励与声纹传感器的螺栓松动故障声纹信号感知方法。研究首先面向不同场景的输电铁塔开展实测数据采集，并基于真实数据对检测方案中的关键参数进行实证分析；随后从不同尺度提取螺栓松动诱发的声波响应特征；最后结合深度学习故障诊断模型，实现螺栓松动状态的智能识别与分类。如图 1.3所示，本文的具体研究内容如下：

1、基于声纹传感器阵列的螺栓松动信号采集与分析。研究声纹传感器阵列的布设方案，优化声波激励源与传感器在三维空间的布局形式；在此基础上研究可变频声波激励源在不同铁塔结构上的最佳共振频率确定方法，探索不同结构铁塔的共振频率范围；进一步获取螺栓松动位置的异常声纹信号特征，明确不同工况环境下声纹信号的差异规律，并设计相应的声纹特征提取与标注策略，采集不同型号铁塔、不同螺栓松动程度及不同工况条件下的声纹数据，建立螺栓松动声纹数据库，为后续模型训练与评估提供数据支撑。

2、基于可调频声激励的螺栓松动声纹特征提取算法研究。针对可调声激励条件下螺栓松动信号的非平稳特性，研究螺栓松动声纹信号的特征提取方法，实现对声波响应信号的多尺度分解与表征；在不同尺度上提取具有诊断意义的能量与频带特征，从而获得对不同松动状态更敏感、对环境扰动更鲁棒的声纹表示。

3、少样本情景下的螺栓松动定位与跨塔泛化方法研究。面向工程应用中标注样本获取成本高、目标铁塔标签稀缺的问题，研究少样本条件下的螺栓松动定位技术。基于半监督学习与迁移学习理论，在仅有少量标注样本的情境下充分利

用大量无标签数据与源域先验知识，实现对铁塔关键部位松动螺栓的识别与定位，降低训练成本并提升模型跨塔泛化能力与鲁棒性。

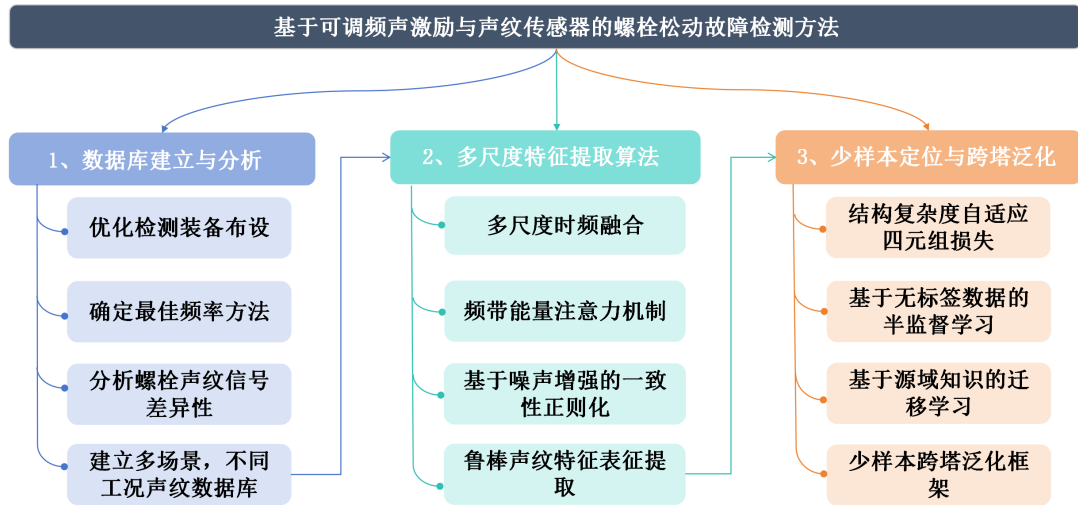


图 1.3 研究内容框架图

1.3.2 研究创新点

本文的主要创新点概括如下：

(1) 提出一种面向铁塔螺栓松动检测的可调频声激励—声纹传感协同感知方案。通过可调频率的声波激励驱动源，以及具备宽频响应与高动态范围的声纹传感器，提高了松动信号的可检测性与判别准确性。在此基础上设计了基于声纹信号的螺栓松动特征提取算法，能够在多尺度上有效识别并提取松动信号中的关键信息，实现对螺栓紧固状态的更准确分析与评估。

(2) 针对现有螺栓松动检测方法在工程应用中存在的局限，以及声纹识别方法普遍依赖大规模标签数据库的问题，提出一种极少标注样本条件下的半监督迁移学习诊断框架。该框架能够在少量标注数据支撑下充分挖掘无标签数据价值，并借助源域知识提升目标铁塔的识别与定位性能；与传统分类方法相比，在极少样本情境下表现出更高精度与更强鲁棒性，有效缓解了大规模标签数据难以获取的问题，为输电铁塔螺栓故障诊断提供了新的研究思路。

1.3.3 论文组织结构安排

论文的结构共分为六章，具体安排如下：

第1章：绪论。介绍输电铁塔螺栓松动检测的研究背景与工程意义，阐述螺栓松动机理及其危害；系统回顾国内外螺栓松动检测方法的研究现状，总结现有研究的不足与挑战，进而明确本文的研究内容、创新点与整体框架。

第2章：相关基础知识。介绍本文涉及的理论基础与关键技术方法，包括基

于声纹信号的螺栓松动检测原理、注意力机制、半监督学习与迁移学习等，为后续章节方法设计提供理论支撑。

第3章：多场景螺栓声纹数据库构建与分析。介绍本文涉及的三座铁塔数据采集场景与实验系统，分析声激励装置与传感器布设方案，并提出一种基于幅值响应的铁塔共振频率确定方法；利用实测数据分析不同工况下紧固与松动状态的差异规律，为后续算法设计与模型训练提供数据依据与机理支撑。

第4章：基于多尺度能量注意力残差网络的螺栓松动识别方法。结合第3章实证分析结果，针对性设计多尺度能量注意力机制，并引入基于噪声的一致性正则化训练策略；在三座铁塔数据集上开展实验验证所提方法的有效性，并对识别结果进行深入分析以提升模型可解释性。

第5章：基于半监督迁移学习的少样本螺栓松动检测方法。为降低标注与训练成本，基于第4章的特征提取网络，构建了少样本条件下的螺栓松动检测与定位框架；在所采集的两座输电铁塔数据上进行实验验证，并结合实验结果进一步分析模型的作用机制与泛化特性。

第6章：总结与展望。对全文研究内容与主要结论进行总结，分析研究中存在的不足，并展望未来可能的改进方向与应用前景。

第2章 相关基础知识

本章围绕输电铁塔螺栓松动声纹检测与少样本跨塔型识别的关键理论基础展开。首先介绍声纹信号在螺栓松动场景下的物理特性及诊断机理，并给出常用时频特征表示方法；随后阐述注意力机制在深度特征提取中的作用；最后介绍半监督学习与迁移学习的核心思想与典型方法，为本文提出的半监督迁移学习框架奠定基础。

2.1 铁塔螺栓松动声纹识别原理

2.1.1 声纹信号的物理特性与诊断可观测性

螺栓连接是输电铁塔角钢节点的关键受力部件，其工作状态直接影响节点的受力性能和铁塔整体结构安全。螺栓发生松动后，连接界面的接触状态发生改变，会直接造成接触刚度、摩擦耗能特性与等效阻尼参数等因素的劣化，进而诱发结构动力学响应特征的显著变化。针对螺栓松动的检测需求，传统静态力学测量方法难以适配输电铁塔带电运维、大范围巡检的工程场景，而声纹检测方法凭借非接触、远距离、对结构损伤敏感的优势，成为当前该领域的研究热点。该方法通过外部激励诱发结构产生可传播的弹性波与声振耦合响应，经传感器采集响应信号后，通过提取并分析信号的时频域特征，实现对螺栓连接状态变化的量化表征与识别。以主动声激励检测机理为例，当声振波穿越螺栓连接界面时，松动引发的界面微间隙、接触非线性与摩擦滑移行为，会导致波的能量散射增强、特定频带内的能量重分布，同时激发谐波、调制边带等非线性响应成分，上述由松动引发的信号特征变化，会在频谱与时频图谱中形成具有显著区分度的声纹差异，可作为螺栓松动识别的核心依据。此类由界面接触状态变化引发的能量散射与非线性响应规律，已在基于扫频激励的螺栓松动检测相关研究中得到广泛验证^[63-64]，也为本文提出的“可调频主动声激励 - 声纹传感”螺栓松动检测方案提供了核心理论支撑与物理基础。

从信号表现角度看，螺栓声纹通常具有以下两类可利用特性：

- **频域与谱分布特性：**声纹信号由多频成分组成，包含基频、谐波以及与结构模态相关的峰值响应。螺栓松动会引起某些模态频带的能量增强或衰减，并可能产生额外的宽带噪声与非线性侧带，表现为频谱峰位、峰宽与能量分布的变化。
- **时域与瞬态特性：**在冲击、脉冲激励或扫频激励下，连接界面会出现短时瞬态响应，例如摩擦冲击、局部滑移，其在时域上表现为波形畸变、包络

波动与脉冲簇；在时频域上则表现为局部能量“团簇”或能量轨迹改变。

因此，若能在多工况下稳定采集声纹信号，并通过合适的时频表示与特征学习模型提取差异模式，就可实现螺栓松动状态识别与定位。

2.1.2 从时域信号到可学习特征的表示方法

螺栓松动引发的声纹信号差异具有显著非平稳特性，若直接对时域信号进行特征学习，易受环境噪声与工况扰动的干扰，因此在工程实践中，通常先将时域声纹信号转换为时频表示形式，再输入深度神经网络进行特征学习与状态识别，以提升诊断的准确性与鲁棒性。

1. 短时傅里叶变换与谱图

短时傅里叶变换（Short-Time Fourier Transform, STFT）通过滑动窗函数 $w[\cdot]$ 对时域信号进行分帧处理，实现时域与频域信息的联合表征，其局部频谱计算公式如下：

$$X(m, \omega) = \sum_{n=-\infty}^{+\infty} x[n] w[n-m] e^{-j\omega n}, \quad (2.1)$$

其中， $x[n]$ 为离散时间信号， n 为采样点索引； $w[n-m]$ 为滑动窗函数， m 表示时帧位置； ω 为角频率； j 为虚数单位。 $X(m, \omega)$ 表示信号在时帧 m 处的局部频谱，其幅值和功率谱分别反映对应时刻下各频率分量的强度及能量分布。窗函数的引入使 STFT 能够在局部时间范围内分析频谱变化，从而实现对非平稳声纹信号的时频表征。该表示方法能够清晰反映声纹信号的模态峰值与频带能量分布特征，是声纹识别领域应用最广泛的基础时频表示之一，适用于初步捕捉螺栓松动引发的声纹能量变化。

2. Mel 频谱与 MFCC

人耳对不同频率声音的感知具有非线性特性，Mel 滤波器组可模拟人耳的频率感知特性，在低频段采用更精细的频率分辨率、高频段采用更粗糙的频率分辨率对频谱能量进行聚合，进而构造 Mel 频谱（Mel-spectrogram）。Mel 频谱不仅具备工程降维、增强特征稳健性的优势，还能与二维卷积神经网络的输入格式良好适配，便于深度网络进行特征挖掘。在工程实践中，可通过 librosa 等开源音频处理工具库快速实现 Mel 频谱与梅尔频率倒谱系数特征的计算^[65]。

MFCC 在 Mel 频谱能量的基础上，进一步经过对数压缩与离散余弦变换（Discrete Cosine Transform, DCT）处理，得到更紧凑的倒谱系数表示。该特征能够有效提取声纹信号的关键判别信息，在复杂噪声背景下仍能保持较强的故障判别能力，是故障声纹识别领域应用成熟的经典手工特征。结合本文研究需求，Mel 频谱（或其变体）更适合作为深度神经网络的输入，其保留的丰富时频纹理

信息，可支撑网络在多尺度上学习螺栓松动引发的声纹能量重分布与纹理差异特征。

2.1.3 基于声纹信号的螺栓松动诊断流程

综合上述声纹信号表示方法的分析，基于声纹信号的输电铁塔螺栓松动诊断流程如图2.1所示，其核心环节依次为：激励设计—信号采集—预处理—时频表示—特征学习—状态判别。

其中，激励环节可采用冲击锤敲击或可调频声发装置激励，通过可控激励方式提升螺栓松动引发的结构声纹响应的可观测性；信号采集环节采用宽频响应、高动态范围的声纹传感器，采集结构振动传播至传感点的声纹响应信号，形成完整且可靠的声纹数据库。随后对采集到的原始声纹信号进行分帧、滤波、归一化等预处理操作，去除噪声干扰并统一信号尺度，再将其转换为谱图或 Mel 频谱等时频表示形式后输入深度模型；最后通过监督或半监督学习算法完成螺栓不同松动状态的分类与故障预警，实现螺栓松动的精准诊断。

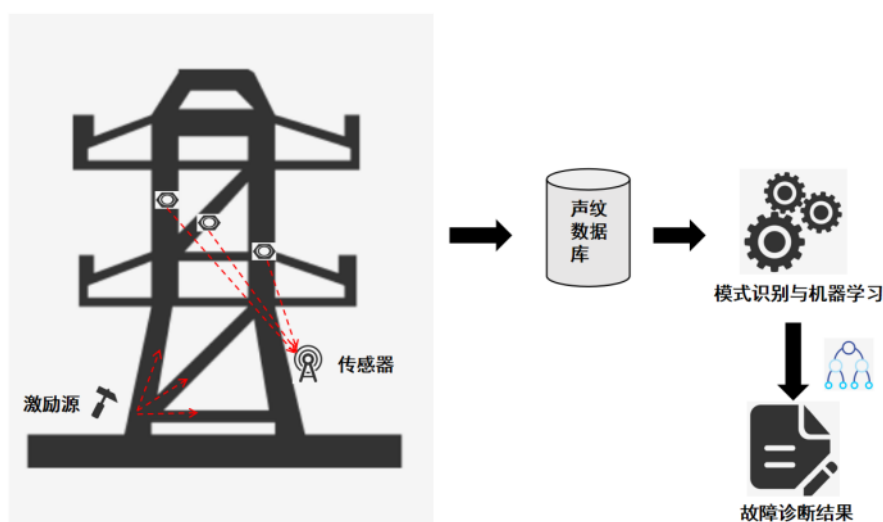


图 2.1 基于声纹信号的铁塔螺栓故障诊断流程

2.2 注意力机制

注意力机制（Attention）的核心目标是在特征表示过程中，对故障判别关键信息赋予更高权重，同时抑制冗余信息与噪声干扰，进而提升模型的特征判别能力。对于基于声纹谱图（STFT 谱图、Mel 频谱等）的输电铁塔螺栓松动识别任务而言，故障相关有效信息通常集中在三个维度：（1）与铁塔结构模态、局部共振特性相关的关键频带；（2）由螺栓松动引发的摩擦滑移、碰撞等现象产生的瞬态时间片；（3）跨塔型、跨工况变化下依然稳定存在的能量聚集区域与时频纹理结构。因此，通过注意力机制对声纹时频图进行自适应重加权，能够在“复杂

传播路径 + 强噪声背景 + 跨塔型域偏移”的工程条件下，增强故障特征的可分性，提升模型的泛化鲁棒性。从数学计算形式上看，假设某层卷积网络输出特征为 $\mathbf{F} \in \mathbb{R}^{C \times H \times W}$ （其中 C 为通道数， H 、 W 分别为时频图的高度与宽度），注意力机制通常通过学习一个权重映射，与原特征进行逐元素或逐通道相乘，实现特征重标定，其通用表达式如下：

$$\mathbf{F}' = \mathcal{A}(\mathbf{F}) \odot \mathbf{F}, \quad (2.2)$$

其中 $\mathcal{A}(\cdot)$ 表示注意力计算模块， $\odot$ 表示逐元素乘法运算。根据权重作用维度的不同，注意力机制可分为通道注意力、空间注意力以及多尺度注意力等典型类型，下文将分别对其核心原理及在螺栓松动声纹识别中的应用价值进行阐述。

(1) 通道注意力机制

通道注意力（Channel Attention）的核心是聚焦“哪些特征通道更具故障判别价值”。在卷积神经网络中，不同特征通道对应不同卷积核学到的模式响应，可理解为对声纹信号中不同频带纹理、局部能量团簇、谱线结构的“探测单元”。在输电铁塔螺栓松动声纹识别任务中，部分特征通道会对螺栓松动引发的时频能量重分配、特定纹理变化更敏感，而另一部分通道则可能主要响应于野外环境噪声或非故障相关区域。因此，通道注意力的核心目标是自适应增强与故障相关的通道响应，抑制无关通道或噪声主导通道的干扰。

Squeeze-and-Excitation (SE) 网络是通道注意力机制的经典实现结构^[66]，其核心流程可概括为“压缩 (Squeeze) — 激励 (Excitation) — 重标定 (Reweight)”三步。首先，对特征图的空间维度进行全局平均池化，提取每个通道的全局响应特征，得到通道描述向量：

$$s_c = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W F_c(i, j), \quad c = 1, \dots, C, \quad (2.3)$$

其中 $\mathbf{s} \in \mathbb{R}^C$ 为通道描述向量，用于表征每个特征通道的全局响应强度。随后，通过两层全连接网络构建瓶颈结构，挖掘通道间的依赖关系，生成通道权重：

$$\mathbf{w} = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{s})), \quad (2.4)$$

其中 $\delta(\cdot)$ 表示 ReLU 激活函数， $\sigma(\cdot)$ 表示 Sigmoid 激活函数， $\mathbf{w} \in (0, 1)^C$ 为归一化后的通道权重向量。最后，将通道权重与原特征通道进行逐通道相乘，完成特征重标定：

$$F'_c = w_c \cdot F_c. \quad (2.5)$$

SE 模块的突出优势在于结构轻量、计算开销小，易于嵌入残差网络等深度模型，且能有效提升特征选择性。对于螺栓松动声纹识别任务而言，SE 模块可使网络在训练过程中自动学习并强化对故障相关频带纹理敏感的通道，从而提高故障判别特征在高噪声背景下的可见性，为后续细粒度识别奠定基础。

(2) 空间注意力机制

空间注意力 (Spatial Attention) 与通道注意力互补，其核心是聚焦时频图的哪些空间位置包含关键故障信息。对于 STFT 谱图、Mel 频谱等而言，空间位置对应“时间—频率”二维平面上的局部区域，螺栓松动导致的声纹差异，往往表现为特定频带在某一时间段内出现能量增强、纹理变粗、边带信号出现等局部特征，空间注意力能够帮助模型精准定位这些关键区域，提升对局部弱故障特征的捕捉能力。

卷积块注意力模块 (Convolutional Block Attention Module, CBAM) 在 SE 模块的基础上，进一步引入空间注意力，采用“通道注意力”和“空间注意力”的串联结构，实现对特征的双重重标定^[67]，其计算流程如下：

$$\mathbf{F}_1 = \mathcal{M}_c(\mathbf{F}) \odot \mathbf{F}, \quad \mathbf{F}_2 = \mathcal{M}_s(\mathbf{F}_1) \odot \mathbf{F}_1, \quad (2.6)$$

其中 $\mathcal{M}_c(\cdot)$ 表示通道注意力模块 (如 SE 模块)， $\mathcal{M}_s(\cdot)$ 表示空间注意力模块。空间注意力的实现通常先对特征图的通道维度进行平均池化与最大池化操作，得到两张 $H \times W$ 的空间特征图，将其拼接后通过卷积运算生成空间权重图：

$$\mathcal{M}_s(\mathbf{F}) = \sigma(f^{7 \times 7}([\text{AvgPool}(\mathbf{F}); \text{MaxPool}(\mathbf{F})])), \quad (2.7)$$

其中 $f^{7 \times 7}$ 表示 7×7 卷积运算，用于捕捉空间全局依赖关系。该结构能够引导网络在“时间—频率”二维平面上，精准突出与螺栓松动相关的关键能量块与时频纹理结构，更贴合声纹信号故障差异的时频定位特性，进一步提升特征的判别能力。

需要注意的是，输电铁塔结构复杂、声纹传播路径长，不同塔型、不同工况会导致声纹频带响应存在显著差异。单一尺度卷积结构难以同时兼顾局部纹理细节 (如瞬态摩擦条纹、边带纹理) 与全局能量分布特征，因此在实际模型设计中，通常引入多尺度分支或特征金字塔结构，并结合注意力机制对多尺度特征进行融合加权。本文提出的多尺度能量注意力模块正是基于上述工程需求设计：在多尺度维度上，同时刻画声纹信号的局部纹理变化、关键频带能量聚集特征以及跨尺度一致的稳定结构响应；通过“多尺度特征抽取 + 能量加权 + 注意力融合”的协同设计，在噪声扰动与跨塔型域偏移条件下，获得更稳定、更具可迁移性的特征表征基础，为后续螺栓松动精准识别与定位提供支撑。

2.3 半监督学习

在输电铁塔螺栓松动检测的工程场景中，数据采集与标注面临显著挑战：螺栓松动样本采集成本高、标注过程繁琐且耗时，尤其是不同塔型、不同螺栓位置的标注样本极为稀缺，难以满足深度模型大规模监督训练的需求。半监督学习作为一种高效的学习范式，基于“少量有标签数据 + 大量无标签数据”的联合训练模式，能够充分挖掘无标签数据中的有效信息，有效降低标注成本，同时提升模型的泛化能力，为解决上述工程痛点提供了重要技术途径。半监督学习的常用理论假设主要包括三类：聚类假设（同类样本在特征空间中呈现聚簇分布）、低密度分割假设（模型决策边界应避开样本密度较高的区域）与流形假设（高维数据本质上分布在低维流形上），这些假设为半监督学习的算法设计提供了理论支撑。

(1) 伪标签思想与自训练

伪标签（Pseudo-Labeling）是自训练（Self-Training）策略中一类经典且应用广泛的方法^[68]。其核心思想是：利用当前训练阶段的模型对无标签样本进行类别预测，将置信度较高的预测结果作为伪标签，作为监督信号纳入模型训练过程，从而逐步扩大有效训练集规模，实现模型性能的迭代提升。设模型表达式为：

$$f_{\theta,\phi}(x) = h_{\phi}(g_{\theta}(x)), \quad \mathbf{p}(x) = \text{softmax}(f_{\theta,\phi}(x)), \quad (2.8)$$

其中， $f_{\theta,\phi}(x)$ 表示模型整体输出， $g_{\theta}(x)$ 为特征提取网络（参数为 θ ）， $h_{\phi}(\cdot)$ 为分类头（参数为 ϕ ）， $\mathbf{p}(x)$ 为经 softmax 函数归一化后的类别概率分布。对于无标签样本 x ，其伪标签的生成过程具体如下：

$$\hat{y} = \arg \max_k p_k(x), \quad \pi = \max_k p_k(x), \quad (2.9)$$

其中， $\hat{y}$ 为预测类别（即伪标签）， $p_k(x)$ 为样本 x 属于第 k 类的概率， π 为预测置信度， τ 为预设置信度阈值。当预测置信度 $\pi \geq \tau$ 时，认为该预测结果可靠，将样本与对应伪标签 $(x, \hat{y})$ 纳入伪标签集合，参与下一轮模型训练。伪标签方法的关键技术难点在于置信度阈值的选择与错误伪标签的累积控制：阈值设置过低会引入大量噪声伪标签，导致模型训练偏移；阈值设置过高则会使可利用的无标签样本数量不足，无法充分发挥半监督学习的优势。针对这一问题，本文采用“双阈值/动态保留”策略，其核心目的在于平衡伪标签的可靠性与样本覆盖度，并通过多轮迭代训练逐步提升伪标签质量，确保模型训练的稳定性与有效性。从半监督学习的思想归类而言，伪标签与自训练属于自训练思想范畴，其设计核心正是依托前文提及的聚类假设，通过模型迭代生成可靠伪标签，从而实现无标签数据的有效挖掘与利用。

(2) 一致性正则化

一致性正则化是半监督学习的另一类核心方法，其思想归类属于一致性正则化范畴，且依托前文所述的流形假设，核心假设为：对同一样本施加不同的扰动或数据增强操作后，样本的语义类别保持不变，模型的预测输出也应保持一致。该方法通过约束模型对扰动样本的输出一致性，引导模型学习更鲁棒、更具泛化能力的特征表征。设 $a_w(\cdot)$ 表示弱数据增强操作（如轻微噪声添加、小幅平移）， $a_s(\cdot)$ 表示强数据增强操作（如随机裁剪、幅度缩放、频率掩蔽），则一致性正则化的损失函数可表示为：

$$\mathcal{L}_{cons} = \mathbb{E}_{x \in \mathcal{D}_t^u} \text{KL}(\mathbf{p}(a_w(x)) \parallel \mathbf{p}(a_s(x))), \quad (2.10)$$

其中， $\text{KL}(\cdot \parallel \cdot)$ 为 KL 散度，用于衡量两个概率分布的差异， $\mathcal{D}_t^u$ 表示目标域无标签数据集， $\mathbb{E}$ 表示数学期望。

在一致性正则化的具体实现中，FixMatch 与 MixMatch 是两类经典且有效的算法。其中，FixMatch 算法将弱增强样本的预测结果作为伪标签，要求模型对强增强样本的输出与该伪标签保持一致^[69]，其无标签损失函数定义如下：

$$\hat{y} = \arg \max \mathbf{p}(a_w(x)), \mathcal{L}_u = \mathbf{1}(\max \mathbf{p}(a_w(x)) \geq \tau) \cdot \text{CE}(\mathbf{p}(a_s(x)), \hat{y}), \quad x \in \mathcal{D}_t^u. \quad (2.11)$$

式中， $\mathbf{1}(\cdot)$ 为指示函数，当括号内条件满足时取值为 1，否则为 0， $\text{CE}(\cdot, \cdot)$ 为交叉熵损失函数， τ 为置信度阈值。MixMatch 算法则综合了标签猜测、数据增强与 MixUp 混合训练策略，通过对样本与标签的混合增强，使模型决策边界更平滑，进一步提升模型的泛化能力^[70]。

在输电铁塔螺栓松动识别任务中，一致性正则化的核心意义在于：野外场景中存在大量非语义扰动，如不同强度的风噪、背景噪声，以及不同的激励幅度、声纹传播衰减条件，通过一致性约束，可引导模型学习对这类非语义扰动不敏感的稳定特征表征，从而显著提升模型在真实野外复杂场景中的识别鲁棒性，适配跨塔型、多工况的工程需求。

2.4 迁移学习、域自适应与半监督域自适应

仅依赖目标域半监督学习，往往难以在极少标签条件下学到稳健的特征表示。迁移学习通过引入源域已有的知识，提升目标域模型的泛化能力；而当源域与目标域存在明显分布差异时，域自适应技术则通过学习域不变特征表示，有效缓解域偏移问题。对于本文输电铁塔螺栓松动检测任务而言，源域可对应“标注数据较充分的塔型或工况”，目标域则对应“新塔型或新布设条件下，标注样本

极少但无标签数据充足的场景”。此处的域偏移主要来源于结构模态差异、声纹传播路径差异、环境噪声背景差异等多个方面。

(1) 对抗式域对齐

深度域自适应网络 (Domain-Adversarial Neural Network, DANN) 是对抗式域自适应方法的典型代表^[71]。该方法在编码器 $g_\theta(\cdot)$ 之后引入域判别器 $q_\psi(\cdot)$ ，用于区分样本来自源域还是目标域，同时通过梯度反转层 (Gradient Reversal Layer) 约束编码器 g_θ ，使其生成域不可区分的特征表示。设域标签 $r \in \{0, 1\}$ (其中源域标签为 0，目标域标签为 1)，其域判别损失函数可表示为：

$$\mathcal{L}_{dom}(\theta, \psi) = \mathbb{E}_{x \sim D_s} \text{CE}(q_\psi(g_\theta(x)), 0) + \mathbb{E}_{x \sim D_t} \text{CE}(q_\psi(g_\theta(x)), 1). \quad (2.12)$$

模型的整体优化目标为：

$$\min_{\theta, \phi} \max_{\psi} \mathcal{L}_{cls}(\theta, \phi) - \lambda \mathcal{L}_{dom}(\theta, \psi), \quad (2.13)$$

其中 $\mathcal{L}_{cls}$ 为监督分类损失， λ 为平衡分类损失与域判别损失的权重系数。其直观意义在于：分类器需学会精准区分样本类别，而编码器则被迫“隐藏”样本的域归属信息，以此减少域分布差异对模型性能造成的不利影响。

当目标域标注样本极少时，单纯的“整体分布对齐”策略易引发类别混叠问题。Saeid Motiian 等提出了基于“样本配对 + 四类判别”的少样本域自适应思路^[72]：该思路将样本两两配对，划分为同域同类、跨域同类、同域异类、跨域异类四类，通过训练域-类判别器识别样本配对类型，显式引入类语义约束，使得域对齐过程不仅追求域间特征分布一致，更能保证跨域同类样本聚合、异类样本分离，有效避免类别混叠。

(2) 半监督域自适应

半监督域自适应 (SSDA) 同时利用源域充足的标注数据、目标域少量标注数据及大量无标签数据，是更贴近输电铁塔螺栓检测工程现实的学习设置。SSDA 的核心挑战在于，既要实现源域与目标域的分布对齐，又要避免在对齐过程中破坏类别结构，尤其是针对输电铁塔螺栓相邻位置这种细粒度类别。

因此，典型的 SSDA 方法多采用“域对齐 + 自训练 + 结构约束/边界约束”的组合策略：在利用源域标注数据进行监督训练的同时，通过目标域少量标注样本校准模型决策边界，并借助伪标签与一致性约束机制，逐步挖掘目标域无标签数据中的有效信息，最终获得在目标域具有稳定可分性的特征表示空间。本文后续提出的半监督迁移学习框架，正是基于这一思路，将对抗式域对齐与伪标签自训练方法相结合，旨在缓解跨塔型域偏移问题，提升模型在低标注条件下的自适应能力。

2.5 度量学习

度量学习 (Metric Learning) 的核心目标是学习一个合适的特征嵌入空间, 使同类样本在该空间内分布更紧致, 异类样本分布更分散, 进而提升细粒度识别精度与模型鲁棒分类能力。结合本文输电铁塔螺栓松动检测任务, 相邻螺栓位置对应的类别在声学特征上差异微弱, 单纯依靠交叉熵损失学习的分割超平面, 易在邻近类别处产生混淆; 而度量学习通过直接约束样本间的距离关系, 更适合刻画“类内一致、类间分离”的结构目标, 恰好适配这一细粒度识别需求。本文统一采用

$$\mathbf{z} = g_{\theta}(x) \quad (2.14)$$

表示样本 x 的特征嵌入表示, 其中 $g_{\theta}(\cdot)$ 为特征提取网络 (参数为 θ), 并以 $\text{dist}(\cdot, \cdot)$ 表示嵌入空间中的距离度量。

(1) 三元组损失

三元组损失 (Triplet Loss) 通过构造三元组样本 (x^a, x^p, x^n) 施加约束, 其中 x^a 为锚点样本 (anchor), x^p 为与锚点同类的正样本 (positive sample), x^n 为与锚点异类的负样本 (negative sample)。其损失函数形式如下^[58]:

$$\mathcal{L}_{tri} = [\text{dist}(\mathbf{z}^a, \mathbf{z}^p) - \text{dist}(\mathbf{z}^a, \mathbf{z}^n) + \alpha]_+, \quad (2.15)$$

其中 $[\cdot]_+ = \max(\cdot, 0)$, α 为预设间隔。该损失的核心作用是推动锚点与正样本的距离 $\text{dist}(a, p)$, 比锚点与负样本的距离 $\text{dist}(a, n)$ 小至少 α , 从而实现类内样本收缩、类间样本分离的目标。

(2) 四元组损失

四元组损失 (Quadruplet Loss) 在三元组损失的基础上, 增加了一个额外负样本, 目的是同时约束“跨类间距”与“不同负类之间的分离度”, 进而在复杂细粒度识别任务中, 帮助模型学习到更稳定的类别结构^[59]。其一种常见的损失函数形式为

$$\mathcal{L}_{quad} = [\text{dist}(\mathbf{z}^a, \mathbf{z}^p) - \text{dist}(\mathbf{z}^a, \mathbf{z}^n) + \alpha]_+ + [\text{dist}(\mathbf{z}^a, \mathbf{z}^p) - \text{dist}(\mathbf{z}^n, \mathbf{z}^{n'}) + \beta]_+, \quad (2.16)$$

其中 $\mathbf{z}^{n'}$ 为另一类负样本的特征嵌入, α 与 β 为不同层次的间隔参数。与三元组损失相比, 四元组损失通过增加额外负样本约束, 更有利于在多类别、邻近类别难分的场景下, 形成边界清晰的类别簇结构, 适配本文螺栓相邻位置细粒度识别的需求。

度量学习的效果好坏, 很大程度上依赖于间隔的设置与样本构造策略的合理性。在输电铁塔螺栓声纹识别任务中, 部分类别 (例如声纹传播距离较远、能量衰减显著的螺栓位置) 的类内离散度较大, 判别难度更高。若对所有类别采

用固定间隔，易产生两类问题：要么对“简单类”约束过强，导致模型训练不稳定；要么对“困难类”约束不足，造成类别边界模糊，影响识别精度。针对这一问题，本文在后续章节提出结构复杂度自适应四元组损失，其核心思路是根据类别结构复杂度，动态调整间隔参数 α 与 β ，对判别困难的类别赋予更强的分离约束，对易判别的简单类保持更稳健的收敛行为，最终提升相邻螺栓位置等细粒度类别的可分性，以及模型的跨域适应稳定性。

第3章 多场景螺栓声纹数据库构建与分析

本章介绍了多场景现场声纹数据采集、数据预处理与分析工作。针对输电铁塔服役场景分散、工况差异大的特点，数据采集选取三类典型的输电铁塔，覆盖了不同地理环境、电压等级与运行工况。针对采集的原始声纹数据，本章完成了异常值剔除，筛选有效样本等预处理工作。在此基础上，从三个核心维度开展数据分析，首先是分析声激励装置参数、传感器布设方式对声纹信号采集质量的影响规律，然后提取并验证不同场景下输电铁塔的共振频率特征，最后挖掘螺栓不同工况下的声纹信号差异化特征。基于上述分析结果，本章优化形成了标准化的现场数据采集方案，同时明确了螺栓松动声纹检测的核心敏感特征，为后续章节的检测模型构建、训练与泛化性验证提供了坚实的数据支撑与理论依据。

3.1 引言

在基于数据驱动的螺栓松动检测研究中，如何高效且准确地获取具有代表性的声纹信号数据，是构建可靠故障诊断模型的前提条件。然而，由于螺栓松动状态通常具有工况复杂、特征微弱以及现场样本稀缺等特点，传统监督学习方法在实际应用中面临“样本不足”与“泛化能力弱”的双重挑战。因此，建立一个覆盖多工况、具备高区分度的螺栓声纹样本库，不仅是保障模型训练与评估质量的基础，更是少样本学习方法实现工程落地的关键所在。值得注意的是，声纹数据的采集过程受多种因素影响，例如如何选择合适的声学激励频率、声激励装置与传感器的布设方式是否影响信号质量、以及在不同工况下紧固与松动状态的声纹特征差异如何体现等问题，均直接影响到最终数据的判别能力与有效性。因此，基于实地采集样本的数据分析，不仅有助于构建更加贴近实际的模型输入，还可为工程应用提供技术指导，同时也有助于从机理层面深入理解螺栓松动现象，为后续模型设计提供可靠的先验知识。

针对上述问题，本章围绕三个不同现场采集的数据样本，引入多种声纹特征提取与分析方法，包括 MFCC、语谱图、频谱图等时频域特征表达方式，并结合 t-SNE 降维可视化、余弦距离计算与特征分布聚类等方法进行综合分析评估。

本章结构安排如下：第3.2节介绍了三个典型数据采集场景的基本信息；第3.3节分析声激励装置与传感器布局对采样结果的影响；第3.4节提出一种基于幅值变化特征的铁塔共振频率识别方法；第3.5节针对不同工况下的螺栓紧固与松动状态进行声纹特征对比分析；第3.6节对本章工作进行总结与归纳。

3.2 数据采集设备与场景描述

3.2.1 数据采集设备



图 3.1 现场数据采集设备组成示意图

如图 3.1 所示，输电铁塔声纹数据采集系统主要由声激励装置、骨传导传感器以及数据采集终端三部分组成。其中，声激励装置的工作频率范围为 20 Hz–20 kHz，可通过磁吸方式固定于输电铁塔主材表面，用于向塔体注入主动声学激励信号。骨传导传感器采用贴片式安装方式，能够高灵敏度感知沿塔体结构传播的声学响应信号，并在一定程度上抑制空气传播环境噪声及非目标噪声的干扰。声激励装置发出的激励信号经塔体传播后，由骨传导传感器进行接收。各装置通过有线方式连接至数据采集终端，终端用于显示声激励装置和传感器的工作状态，并完成参数调节、数据存储及导出等功能。

3.2.2 数据采集场景

(1) 安徽省合肥市紫蓬 220 kV 送出工程

为获取接近实际运行工况下的输电铁塔螺栓声纹数据，本文联合国家电网相关单位，对合肥紫蓬 220 kV 送出工程中的典型塔型开展了现场数据采集工作。紫蓬 220 kV 送出工程为一项高压输电线路工程，已于 2023 年 3 月正式投运。该线路工程包含多类螺栓连接结构组件，包括塔架主材、绝缘子挂点以及导线固定装置等，为研究高压电力设施中螺栓连接状态的声学响应特性提供了理想的现场环境基础。本研究选取 220-HB21S-Z2 塔型作为主要测试对象，实际完成了对共计 16 座输电铁塔螺栓紧固状态的现场样本采集工作，其结构参数详见表 3.1。鉴于紫蓬 220 kV 工程处于运行状态，为保障电网系统的稳定性及采集人员的作业安全，采集过程中未对螺栓状态进行干预操作，故所获取数据均为无标签数据。为提高数据利用率，最大程度挖掘其潜在信息，本文进一步在部分铁塔上实施了“一塔多测”策略，结合同型号多塔对比采样的方法，试图从野外真实场景下不同工况条件中提取螺栓声纹的共性特征。

表 3.1 紫蓬送出工程部分塔型参数

杆塔编号	塔型	KML 编号	备注	呼高/m	全高/m
A2	220-HB21S-Z2	ZA2	—	36	52.4
A6	220-HB21S-Z2	ZA6	—	30	46.4
A8	220-HB21S-Z2	ZA8	—	33	49.4
A11	220-HB21S-Z2	ZA11	—	30	46.4
A20	220-HB21S-Z2	ZA20	—	36	52.4
B2	220-HB21S-Z2	ZB2	—	36	52.4
B6	220-HB21S-Z2	ZB6	—	39	55.4
B8	220-HB21S-Z2	ZB8	—	33	49.4
B11	220-HB21S-Z2	ZB11	—	33	49.4
B16	220-HB21S-Z2	ZB16	—	33	49.4
B19	220-HB21S-Z2	ZB19	—	39	55.4
B20	220-HB21S-Z2	ZB20	—	36	52.4
C2	220-HB21S-Z2	ZC2	—	39	55.4
C6	220-HB21S-Z2	ZC6	—	33	49.4
C7	220-HB21S-Z2	ZC7	—	33	49.4
C8	220-HB21S-Z2	ZC7+1 [*]	—	36	52.4

^{*} C8 对应 KML 编号为 “ZC7+1”。

(2) 张洼实训基地

由于紫蓬 220 kV 送出工程为在运营状态的工程项目，其采集的数据样本均不包含螺栓松动状态的标签信息，所能提供的有效信息相对有限。为了更系统地研究输电铁塔螺栓松动的响应特性及其内在机理，本文选取张洼实训基地内一座 220 kV 酒杯型输电铁塔作为实验对象，并开展了一系列实证性研究。该塔型的结构参数详见表3.2。相较于紫蓬工程现场，该实训基地具备相对封闭、可控的环境条件。在确保操作安全的前提下，借助现场运维人员的协同配合，能够获取包含明确标签信息的多组实验样本，为后续建模和分析提供了更加丰富的数据支撑。

表 3.2 张洼实训基地实验铁塔相关信息

参数名称	酒杯塔
塔形	直线塔
电压等级/kV	220
呼高/m	21
全高/m	24.5
正面根开/m	7.92

(3) 国网宏源工业区

为了获取更加多样化的螺栓声纹信号数据，并进一步开展螺栓松动定位相关研究，本文选取国家电网宏源工业区内的两种典型输电铁塔塔型（干字塔与酒杯塔）作为实验对象，开展系统性的数据采集工作。两种塔型的结构参数详见

表3.3。在此基础上，通过对多个典型螺栓位置实施不同程度的松动操作，采集相应声纹信号，对原有数据集进行补充和扩展，以提升模型训练数据的代表性与多样性。此外，通过本次实地实验采集的数据，进一步对所提出的松动定位算法进行可行性验证，并尝试对螺栓的紧固状态进行量化分析，从而为后续的实际应用奠定数据基础。

表 3.3 宏源工业区实验铁塔相关信息

参数名称	干字塔	酒杯塔
塔形	耐张塔	直线塔
电压等级/kV	500	500
呼高/m	18	21
全高/m	30	25
正面根开/m	11.2	8.9

本文螺栓声纹数据库主要来源于上述三类典型实测场景，覆盖了不同电压等级、塔型结构及螺栓松动形态，可为多场景实验验证提供基础。但受现场采集条件和安全作业要求限制，当前数据集尚未系统覆盖季节温变、沿海高湿、山区强风及极端天气等复杂环境因素。因此，本文结论主要适用于现有数据覆盖范围。

3.3 声激励装置与传感器布局分析

输电铁塔螺栓主要分布于塔腿法兰连接处、主材与斜材连接节点、横担连接部位以及塔身节段拼接区域，整体呈现出沿塔高分层分布、在关键连接节点局部集中的特点。由于螺栓连接点通常对应杆件交汇或局部受力集中的区域，其松动状态会对结构传播特性及局部响应特征产生明显的影响。因此，从理论上讲，声激励装置与传感器的布局应结合目标螺栓所在节点的三维空间位置进行综合考虑。需要说明的是，本文试验主要面向铁塔基座典型连接区域开展，且声激励装置与骨传导传感器均安装于基座同一高度层附近。为保证实验条件一致性并突出基座平面内的相对位置对采集结果的影响，本将布局问题简化为二维平面内的布设分析，即在控制安装高度基本一致的前提下，重点比较不同平面相对位置下的布局差异。因而，本文关于布局优选的结论主要适用于铁塔基座典型节点区域的一发一收采集场景，对于跨高度层、跨节点类型的三维布局优化问题，仍有待后续进一步研究。

3.3.1 布设方案与实验设计

本文采用“一发一收”配置，即一个变频声激励装置输出指定频率的声波激励，一个固定骨传导传感器同步采集结构声纹响应。在铁塔基座平面内的相对位

置上，重点考察两类典型布局：

- **对角布设**：声激励装置与传感器分别位于铁塔基座对角位置，如图3.2a所示；
- **邻角布设**：声激励装置与传感器位于相邻基柱位置，如图3.2b所示。

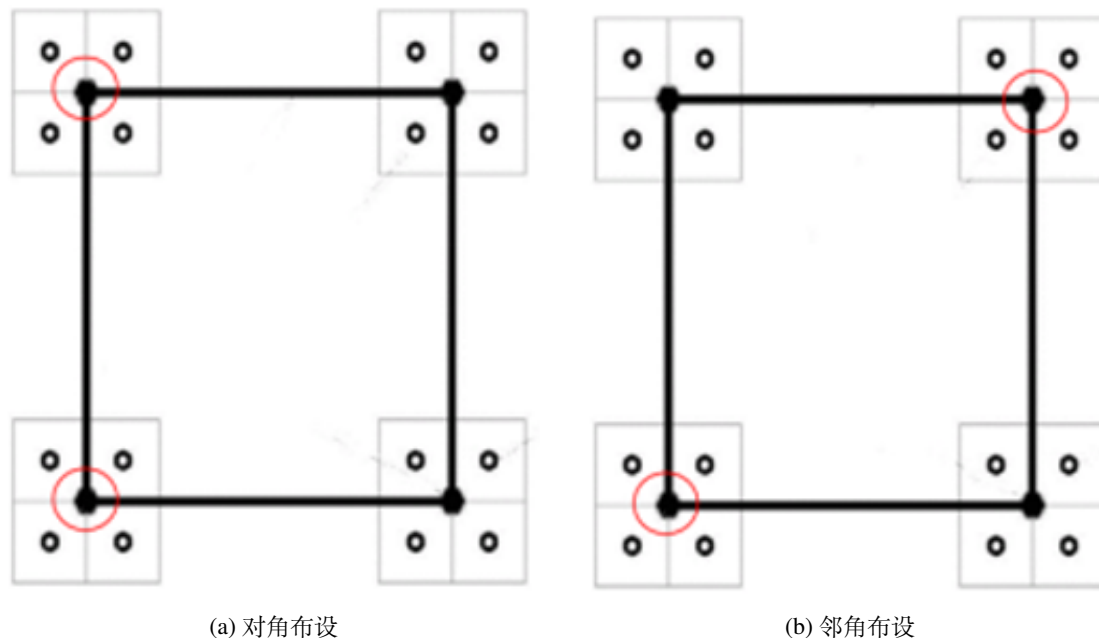


图 3.2 声激励装置与骨传导传感器安装示意图

为进一步刻画邻角布设在真实铁塔几何条件下的差异性，如图 3.3 所示，在张洼实训基地试验中对该基座四个基柱进行编号：将声激励装置所在基柱标记为 1 号位，逆时针依次标记为 2、3、4 号位。则传感器位于 3 号位对应对角布设（记为 13），位于 2 号或 4 号位对应邻角布设（记为 12、14）。本文选取张洼实训基地内的 220kV 酒杯塔作为实验对象，在相同频率扫描策略下，针对不同工况分别采集不同布局数据，用于验证布局差异的影响。

虽然本文布局比较主要在基座周向平面内展开，但其对应的对象仍是铁塔基座典型螺栓连接节点。换言之，本文所分析的二维布局实质上是针对特定高度层和典型节点区域的一种局部简化描述，其目的在于在控制变量条件下考察声激励装置与传感器相对位置对采集结果的影响，而非对全塔范围内所有螺栓节点的三维最优布局进行全面求解。

3.3.2 布局差异的特征建模与距离评价

(1) t-SNE 可视化分析

t-SNE (t-distributed Stochastic Neighbor Embedding) 是一种非线性降维方法，其通过在高维与低维空间分别构造样本相似度的概率分布，并最小化两者之间的 KL 散度，使低维表示尽可能保持高维空间的局部邻域结构。该方法能够较好



图 3.3 声激励装置与传感器现场布局编号

揭示样本的聚类与交叠关系，适用于对不同布设方式下声纹样本分布进行直观对比。在本研究中，t-SNE 主要用于对不同布设方式（对角 13、邻角 12、邻角 14）以及不同工况（紧固/松动）下采集到的声纹样本进行定性对比。为避免幅值尺度差异对嵌入结果造成主导影响，本文首先对每段稳态声纹信号构建统一维度的频域表征向量，并采用标准化对特征进行归一化处理；随后将标准化后的高维特征输入 t-SNE，得到二维嵌入结果并进行可视化。

考虑到 t-SNE 不保证全局距离可解释性，且不同参数设置可能导致可视化形态发生变化，本文不直接依据 t-SNE 图对“布局差异大小”进行定量结论，而是将其用于提出与验证假设，例如“某布局类内更集中或者更分散”或“不同布局之间是否存在明显分界”。为了进一步量化声激励装置与传感器不同布局方式对采集声纹信号的影响，本文从通用谱表征差异与定频激励响应差异两条路径构建特征并定义布局间距离指标，通过基于 MFCC 的平均谱形态相似度度量和面向定频激励的同步检波特征度量两个角度入手，同时引入类内离散度评价采集稳定性。

(2) MFCC 平均谱表征向量与布局间余弦距离

MFCC 是语音与声学信号处理中常用的谱特征之一，其基于短时频谱的倒谱特征，能够紧凑表征信号的谱包络与能量分布形态。本文利用 MFCC 对不同布设方式采集到的声纹信号进行特征化表示，并通过布局内均值与布局间余弦距离实现差异量化。MFCC 主要用于构建布局层面的“平均谱表征向量”：对每段稳态声纹信号提取 MFCC 序列，并在时间帧维度求均值得到该段信号的 MFCC 统计向量；进一步对同一布局下所有样本的统计向量求均值，得到该布局的平均

MFCC 中心向量。最后采用余弦距离对不同布局的中心向量进行比较，以量化不同布设方式对声纹平均谱形态的影响程度。设在布局条件 g 下采集到第 j 段稳态声纹信号为 $\tilde{x}_{g,j}(n)$ ，采样率为 f_s 。对信号进行分帧与加窗处理：第 t 帧记为

$$x_{g,j}^{(t)}(n) = \tilde{x}_{g,j}(n + tH)w(n), \quad n = 0, 1, \dots, L-1, \quad (3.1)$$

其中 L 为帧长、 H 为帧移、 $w(n)$ 为窗函数（即 Hanning 窗）。对每帧作离散傅里叶变换并计算功率谱：

$$X_{g,j}^{(t)}(k) = \sum_{n=0}^{L-1} x_{g,j}^{(t)}(n)e^{-i2\pi kn/L}, \quad P_{g,j}^{(t)}(k) = \frac{1}{L} \left| X_{g,j}^{(t)}(k) \right|^2. \quad (3.2)$$

采用 B 个 Mel 滤波器组 $\{H_b(k)\}_{b=1}^B$ 对功率谱加权得到 Mel 频带能量：

$$E_{g,j}^{(t)}(b) = \sum_k P_{g,j}^{(t)}(k) H_b(k), \quad b = 1, \dots, B. \quad (3.3)$$

取对数后通过离散余弦变换得到第 m 维 MFCC 系数（取前 D 维）：

$$c_{g,j}^{(t)}(m) = \sum_{b=1}^B \log \left(E_{g,j}^{(t)}(b) + \epsilon \right) \cos \left[\frac{\pi m}{B} \left(b - \frac{1}{2} \right) \right], \quad m = 1, \dots, D, \quad (3.4)$$

其中 ϵ 为防止 $\log(0)$ 的微小常数。将每帧 MFCC 组成向量 $\mathbf{c}_{g,j}^{(t)} \in \mathbb{R}^D$ ，对时间帧求均值得到该段音频的 MFCC 统计表征向量：

$$\mathbf{m}_{g,j} = \frac{1}{T_{g,j}} \sum_{t=1}^{T_{g,j}} \mathbf{c}_{g,j}^{(t)} \in \mathbb{R}^D, \quad (3.5)$$

其中 $T_{g,j}$ 为该段信号的帧数。若布局 g 下共有 M_g 段有效样本，则定义布局 g 的平均 MFCC 中心向量为

$$\bar{\mathbf{m}}_g = \frac{1}{M_g} \sum_{j=1}^{M_g} \mathbf{m}_{g,j}. \quad (3.6)$$

对两种布局 g_1, g_2 ，用余弦距离衡量其平均谱形态差异：

$$d_{\cos}^{\text{MFCC}}(g_1, g_2) = 1 - \frac{\bar{\mathbf{m}}_{g_1}^T \bar{\mathbf{m}}_{g_2}}{\|\bar{\mathbf{m}}_{g_1}\| \|\bar{\mathbf{m}}_{g_2}\|}. \quad (3.7)$$

当 $d_{\cos}^{\text{MFCC}} \rightarrow 0$ 时，表示两种布局在 MFCC 均值表征层面的谱形态高度相似。

(3) 面向定频激励的同步检波特征与布局稳定性评价

除通用谱表征外，本文试验采用已知频率的声发激励（ $f_0 = 450 \text{ Hz}$ ），稳态响应可视作同频窄带分量叠加噪声。为更稳健地估计激励频率处的响应强度，

本文引入同步检波思想，将稳态信号投影到 $\cos(2\pi f_0 n/f_s)$ 与 $\sin(2\pi f_0 n/f_s)$ 基上。设稳态信号长度为 N ，定义

$$I = \frac{2}{\sum_{n=0}^{N-1} w^2(n)} \sum_{n=0}^{N-1} (\tilde{x}_{g,j}(n)w(n)) \cos(2\pi f_0 n/f_s), \quad (3.8)$$

$$Q = \frac{2}{\sum_{n=0}^{N-1} w^2(n)} \sum_{n=0}^{N-1} (\tilde{x}_{g,j}(n)w(n)) \sin(2\pi f_0 n/f_s), \quad (3.9)$$

则激励频率 f_0 处的幅值与相位估计为

$$A_1 = \sqrt{I^2 + Q^2}, \quad \phi_1 = \arctan 2(Q, I). \quad (3.10)$$

同理可得到 $2f_0, 3f_0$ 处的谐波幅值 A_2, A_3 。为刻画可能由螺栓松动引入的非线性增强，定义二、三次谐波比及总谐波失真为

$$r_2 = \frac{A_2}{A_1 + \epsilon}, \quad r_3 = \frac{A_3}{A_1 + \epsilon}, \quad \text{THD} = \frac{\sqrt{A_2^2 + A_3^2}}{A_1 + \epsilon}. \quad (3.11)$$

此外，若连接松动引入调制或碰撞效应，激励频率附近能量可能由主峰向两侧扩散。设 $P(f)$ 为功率谱密度估计，定义以 f_0 为中心、半宽为 Δ 的邻域能量以及主峰极窄带（半宽 δ ）能量：

$$E_{\text{band}} = \int_{f_0-\Delta}^{f_0+\Delta} P(f) df, \quad E_{\text{main}} = \int_{f_0-\delta}^{f_0+\delta} P(f) df, \quad (3.12)$$

则旁带能量比定义为

$$\rho = \frac{E_{\text{band}} - E_{\text{main}}}{E_{\text{main}} + \epsilon}. \quad (3.13)$$

综上，将每段信号的定频特征组成向量 $\mathbf{z}_{g,j} \in \mathbb{R}^p$ ，即

$$\mathbf{z}_{g,j} = [\log(A_1 + \epsilon), \phi_1, \log(r_2 + \epsilon), \log(r_3 + \epsilon), \log(\text{THD} + \epsilon), \log(\rho + \epsilon)]^\top. \quad (3.14)$$

与 MFCC 类似，定义布局 g 的定频特征中心向量

$$\bar{\mathbf{z}}_g = \frac{1}{M_g} \sum_{j=1}^{M_g} \mathbf{z}_{g,j}. \quad (3.15)$$

并以余弦距离刻画两布局在定频响应特征层面的差异：

$$d_{\cos}^{(z)}(g_1, g_2) = 1 - \frac{\bar{\mathbf{z}}_{g_1}^\top \bar{\mathbf{z}}_{g_2}}{\|\bar{\mathbf{z}}_{g_1}\| \|\bar{\mathbf{z}}_{g_2}\|}. \quad (3.16)$$

为进一步衡量同一布局下数据的可重复性，即稳定性，引入类内离散度指标，定义为样本到其布局中心的平均余弦距离：

$$S_{\text{intra}}(g) = \frac{1}{M_g} \sum_{j=1}^{M_g} \left(1 - \frac{\mathbf{z}_{g,j}^\top \bar{\mathbf{z}}_g}{\|\mathbf{z}_{g,j}\| \|\bar{\mathbf{z}}_g\|} \right). \quad (3.17)$$

$S_{\text{intra}}(g)$ 越小表示该布局下特征分布越集中、采集结果越稳定。

(4) 指标解释与使用方式

本文将 t-SNE 可视化作为定性分析，用于观察不同布局样本在特征空间中的聚集和交叠情况；同时采用 $d_{\cos}^{\text{MFCC}}$ 与 $d_{\cos}^{(z)}$ 从不同角度量化布局间差异。其中， $d_{\cos}^{\text{MFCC}}$ 主要反映不同布局对声纹平均谱形态的影响，而 $d_{\cos}^{(z)}$ 与 $S_{\text{intra}}(g)$ 更侧重定频激励下响应幅值、谐波非线性与能量扩散等机制性差异。若出现“可视化分布差异较明显但 $d_{\cos}^{\text{MFCC}}$ 较小”的情况，通常意味着不同布局更可能改变了类内离散度而非平均谱中心位置，此时 $S_{\text{intra}}(g)$ 可用于补充说明。

3.3.3 布局实验结果分析

t-SNE 嵌入结果更侧重保持样本的局部邻域关系，其坐标轴本身不具有明确物理意义，且全局距离不保证一致可解释性。因此，本文主要依据二维图像中散点的聚集程度与交叠、分离趋势对不同布设方式的差异进行定性判断。若同一布设方式样本点在嵌入空间中更为聚集，通常表明该布局下数据类内一致性更强、可重复性更好；若散点分布更为分散，则意味着该布局下样本波动更大，可能更易受到传播路径、结构耦合条件或环境扰动的影响。另一方面，若不同布局的散点分布呈现明显分离，可视为布局差异导致特征分布发生变化的直观证据；若多类散点分布高度交叠，则说明布局对样本总体分布的影响相对有限。

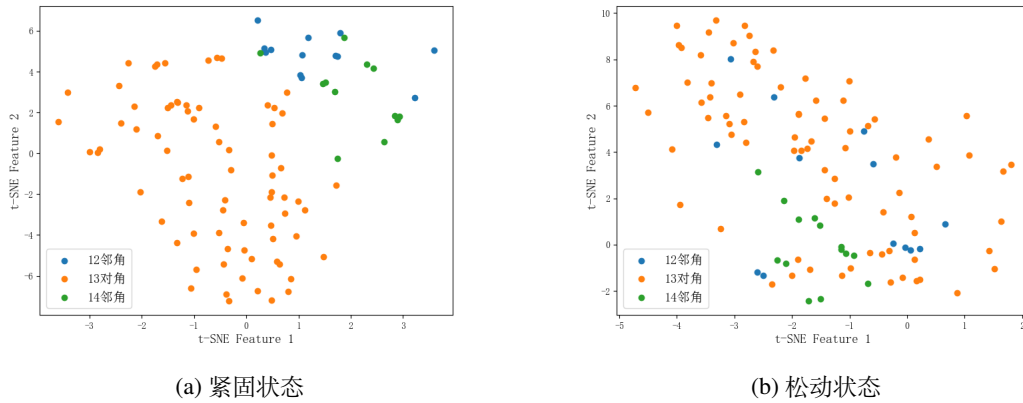


图 3.4 不同工况下紧固与松动状态的 t-SNE 可视化

结合图3.4a与图3.4b可以观察到：在松动工况下，不同布设方式的样本点分布总体交叠程度更高，表明松动状态可能引入更强的非线性与能量扩散，使得样本对传播路径与微小扰动更敏感，从而导致类内波动增大；在紧固工况下，样布点的分离趋势相对更为清晰，邻角 12 与邻角 14 样本更多集中于嵌入空间的相近区域，而对角 13 样本整体偏离，说明在结构连接更稳定时，布局差异更可能以整体分布偏移的形式体现出来。

表3.4与表3.5给出了不同布设方式在两类特征空间中的布局间差异（余弦距离）以及布局内稳定性（类内离散度）对比结果。总体来看，无论紧固或松动工

表 3.4 不同布设方式之间的特征差异（余弦距离）

布设对比	$d_{\cos}^{\text{MFCC}}$ (紧固)	$d_{\cos}^{\text{MFCC}}$ (松动)	$d_{\cos}^{(z)}$ (紧固)	$d_{\cos}^{(z)}$ (松动)
邻角 12 vs 对角 13	0.007494	0.005294	0.000705	0.000184
邻角 12 vs 邻角 14	0.004266	0.013184	0.001490	0.001825
邻角 14 vs 对角 13	0.007193	0.009815	0.003087	0.001358

表 3.5 不同布设方式的类内离散度 $S_{\text{intra}}(g)$ 对比

布设方式	紧固状态	松动状态
邻角 12	0.003191	0.002428
邻角 14	0.004948	0.005048
对角 13	0.003457	0.003928

况，基于 MFCC 的中心距离 $d_{\cos}^{\text{MFCC}}$ 以及基于定频响应特征的中心距离 $d_{\cos}^{(z)}$ 均处于较小量级，表明在本文采集条件与特征表征方式下，声激励装置与传感器的周向布局变化对声纹特征中心位置的影响总体有限，不同布局在平均谱形态与定频机制性响应层面的差异并不显著。

进一步地，从定频响应特征距离 $d_{\cos}^{(z)}$ 可见，邻角 12 与对角 13 在紧固与松动两种工况下均呈现较高一致性，说明在这两种布设方式下，激励频率附近的幅值、谐波与能量扩散等响应特征具有较强的一致性与可迁移性。相比之下，邻角 14 与其他两类布设方式之间的 $d_{\cos}^{(z)}$ 略大，提示其传播路径或局部结构耦合条件可能存在一定差异，但整体仍处于较小范围。与“中心差异较小”不同，表3.5显示不同布局的类内离散度 $S_{\text{intra}}(g)$ 存在更明显差异，说明布局对数据的影响更主要体现为类内一致性的变化，而非特征均值的显著偏移。具体而言，邻角 12 在两种工况下均具有最小的 S_{intra} ，表现出最佳的采集稳定性；对角 13 次之；邻角 14 在两种工况下 S_{intra} 相对较大，表明其采集数据波动更强、可重复性相对较弱。

综上所述，对角布设与邻角 12 布设在特征中心位置上均表现出较高一致性，但对角布设在覆盖范围与传播表征能力方面更具优势。相较于邻角布设，对角布设的激励与接收路径跨越基座对角方向，能够覆盖更广的典型连接区域，从而获得更具代表性的结构传播响应特征。虽然邻角 12 在类内离散度指标上略优于对角布设，但对角布设的稳定性仍处于可接受范围内，能够满足后续共振频率扫描与松动识别试验对数据一致性和可重复性的要求。同时需要指出的是，不同邻角位置之间的采集结果并不完全一致，其中邻角 14 在两种工况下均表现出较高的类内离散度，说明邻角布设对局部传播路径和结构耦合条件更为敏感，采集结果稳定性相对不足，不利于后续统一建模。综合考虑采集稳定性、响应覆盖性和传播表征能力，本文最终选取**对角布设**作为后续试验的标准布局方案。该结论主要适用于铁塔基座典型节点区域的一发一收采集场景；对于不同高度层、不同节点类型条件下的三维布局优化问题，仍需在后续研究中进一步展开。

3.4 铁塔共振频率分析

输电铁塔作为典型的高耸空间结构，其在外部激励作用下表现出明显的动力学特性。结构在振动过程中通常存在若干固有频率，当外部激励频率与结构某一固有频率接近时，系统将产生共振现象，此时振动响应幅值显著放大。上述与结构振动幅值显著增强相对应的激励频率，通常称为结构的共振频率，其大小受结构质量分布、刚度特性以及连接状态等因素的共同影响。在基于声纹信号的输电铁塔螺栓松动检测研究中，声纹数据的质量在很大程度上取决于数据采集阶段的激励方式与参数设置。其中，声发激励频率的选择对铁塔结构振动响应强度及声纹特征的显著性具有直接影响。若激励频率与结构动力学特性不匹配，系统响应较弱，采集到的声纹信号中与结构状态相关的有效信息将被噪声掩盖，从而不利于后续松动状态的识别。而当外部激励频率接近铁塔结构的共振频率时，系统振动响应幅值将显著增强。此时，结构连接状态（如螺栓松动）引起的刚度变化，更容易通过振动及声辐射特性反映到声纹信号中。因此，在开展声纹数据采集之前，有必要对铁塔结构在不同工作状态下的共振特性进行分析，从而为声发激励频率的合理选择提供依据。为此，本节中提出一种基于时域幅值特征的共振频率识别方法，并以实验数据对其有效性进行验证，确定适用于后续声纹信号采集与螺栓松动检测的声发激励频率范围。

3.4.1 基于幅值特征的共振频率确定方法

设在声发激励频率为 f_i 条件下采集到的声纹信号表示为离散时间序列：

$$x_{f_i}(n), \quad n = 0, 1, \dots, N - 1 \quad (3.18)$$

其中， N 为信号总采样点数。

声纹信号 $x_{f_i}(n)$ 可视为铁塔结构在外部激励作用下的综合响应，其幅值大小与结构动力学特性以及激励频率密切相关。为消除声发启动及系统过渡过程的影响，仅选取系统进入稳定振动状态后的信号用于分析。

在声发激励作用初期，铁塔结构的响应信号通常包含明显的瞬态成分，该部分信号主要反映系统从静止状态向稳态振动状态的过渡过程，不利于共振特性的准确表征。因此，本文在分析过程中引入稳态时间窗 $[t_s, t_e]$ ，并将该时间窗内的声纹信号视为铁塔结构在给定激励频率下的稳态响应。设采样频率为 f_s ，则对应的采样点区间为：

$$n_s = \lfloor t_s f_s \rfloor, \quad n_e = \lfloor t_e f_s \rfloor \quad (3.19)$$

稳态声纹信号定义为原始声纹信号在稳态时间窗内的截取结果，即：

$$\tilde{x}_{f_i}(n) = x_{f_i}(n) w(n) \quad (3.20)$$

其中，窗函数 $w(n)$ 定义为：

$$w(n) = \begin{cases} 1, & n_s \leq n \leq n_e, \\ 0, & \text{其他}. \end{cases} \quad (3.21)$$

在稳态振动条件下，铁塔结构对外部激励的响应强度可通过声纹信号幅值进行表征。本文将稳态声纹信号的平均绝对幅值作为系统在激励频率 f_i 下的等效响应幅值，其定义为：

$$A(f_i) = \frac{1}{L} \sum_{n=1}^L |\tilde{x}_{f_i}(n)| \quad (3.22)$$

其中， $L = n_e - n_s + 1$ 为稳态时间窗内的采样点数。

上述幅值特征可视为铁塔结构在激励频率 f_i 条件下频率响应函数幅值的实验估计值。通过对不同激励频率下幅值特征的对比分析，可以刻画铁塔结构的频率响应特性。

为减小环境噪声和单次实验随机因素对分析结果的影响，在相同激励频率条件下进行多次重复实验。设在激励频率 f_i 下共进行 M 次实验，第 k 次实验得到的幅值特征记为 $A_k(f_i)$ ，则该频率下的统计平均幅值定义为：

$$\bar{A}(f_i) = \frac{1}{M} \sum_{k=1}^M A_k(f_i) \quad (3.23)$$

统计平均处理能够提高幅值特征的稳定性和可靠性，使其更准确地反映铁塔结构在给定激励频率下的整体响应水平。上述建模基础上，将铁塔结构在不同声发激励频率条件下的统计平均绝对幅值进行对比分析。当激励频率接近结构共振频率时，系统响应幅值达到最大值。因此，本文将幅值特征达到极大值时所对应的激励频率，定义为铁塔结构在当前工作状态下的共振频率：

$$f_r = \arg \max_{f_i \in F} \bar{A}(f_i) \quad (3.24)$$

所确定的共振频率用于指导后续声纹数据采集阶段声发激励频率的选取。

3.4.2 共振频率的实验设置与结果分析

为进一步对不同工况下输电铁塔的共振频率开展系统性研究，本文在张洼实训基地开展声发激励试验，采集铁塔在不同工作状态下的声纹响应信号。试验对象为基地内 220 kV 酒杯型输电铁塔。通过人为改变关键连接部位螺栓的紧固状态以模拟典型工况，并在各工况下进行声发激励频率扫描，获取对应的声纹响应数据，为后续共振频率分析与声发频率选择提供数据支撑。

在声纹响应信号采集过程中，本文统一设置声纹信号的采样率为 16000 Hz，以满足对铁塔结构振动声信号的时域分析需求。每次声发激励频率条件下的录

制时长均为 10 s，其中包含声激励装置启动阶段、系统过渡阶段以及稳定振动阶段的完整响应过程。为保证所提取声纹信号能够真实反映铁塔结构在给定激励频率下的稳定响应特性，本文在数据处理阶段引入稳态时间窗，对每段声纹信号中 2–8 s 区间内的信号进行截取，并作为稳态声纹信号用于后续幅值特征计算与共振频率分析。在相同工况与相同声发激励频率条件下，进行多次重复采集以增强统计可靠性。设在激励频率 f_i 下共采集 M 段有效声纹样本，其中 M 的取值根据该频率下可获得的有效样本数量确定。对每一段稳态声纹信号分别计算幅值特征，并采用统计平均方式得到该频率下的平均响应幅值，用于绘制频率—幅值关系曲线。考虑螺栓连接状态变化对结构动力学特性的影响，本文设置三种典型工况：

- **紧固状态**：整座铁塔连接构件相关螺栓均处于正常紧固状态，作为基准工况；
- **中部松动状态**：铁塔中部位置连接构件螺栓设置为松动状态，以模拟局部连接刚度降低；
- **顶部松动状态**：铁塔顶部位置连接构件螺栓设置为松动状态，以模拟不同空间位置松动对结构响应的影响。

为保证激励与响应信号的有效耦合，声激励装置与声纹采集传感器采用对角布设方式。即声激励装置与声学传感器分别布设在铁塔结构的对角位置，使激励信号能够充分作用于结构整体，同时采集到具有代表性的响应信号。在不同工况下，声激励装置与传感器的安装位置保持不变，以确保不同状态下采集数据之间的可比性。上述三种工况均在同一铁塔、相同布设条件下实施，保证不同工况数据的可比性。为模拟输电铁塔在实际运行过程中可能出现的螺栓松动工况，本文通过人为调节对角连接构件处螺栓的紧固程度来构造不同工作状态。其中，螺栓松动状态定义为螺栓扭矩基本为零，即螺栓处于无有效预紧力状态，但仍保持螺栓与构件之间的基本连接关系。该设置能够显著改变局部连接刚度，从而对铁塔整体动力学特性产生影响。

在每种工况下，通过调节声激励装置的输出频率，对铁塔进行离散频率激励并采集响应信号。声发激励频率范围设置为 200–1000 Hz。考虑低频段共振特性变化可能更为敏感，同时兼顾试验效率，本文采用分辨率分段设置的频率扫描方案：

- 在 200–700 Hz 频段内，以 50 Hz 为步长递增（即 200, 250, …, 700 Hz）；
- 在 700 Hz 以上频段内，以 100 Hz 为步长递增（即 800, 900, 1000 Hz）。

考虑户外试验环境对声纹信号可能产生影响，本文在数据采集过程中同步记录气象条件。各工况的采集环境如下：

- 对角紧固状态采集：晴，温度 16–19℃，微风；

- 对角中部松动状态采集：温度 21℃，微风；
- 对角顶部松动状态采集：温度 22℃，微风。

总体而言，各工况采集时均处于微风条件且温度变化范围较小，可认为环境扰动在可控范围内。为进一步降低环境噪声影响，采集过程中尽量避开强噪声源，如车辆通过、人员高频走动等。通过上述方式构建对角紧固、对角中部松动以及对角顶部松动三种典型工况，为分析螺栓松动对铁塔共振特性及声纹响应的影响提供实验基础。

基于前述共振频率确定方法，对铁塔结构在不同工作状态下的声纹响应信号进行分析，计算各声发激励频率对应的统计平均幅值特征，并绘制频率—幅值关系曲线。

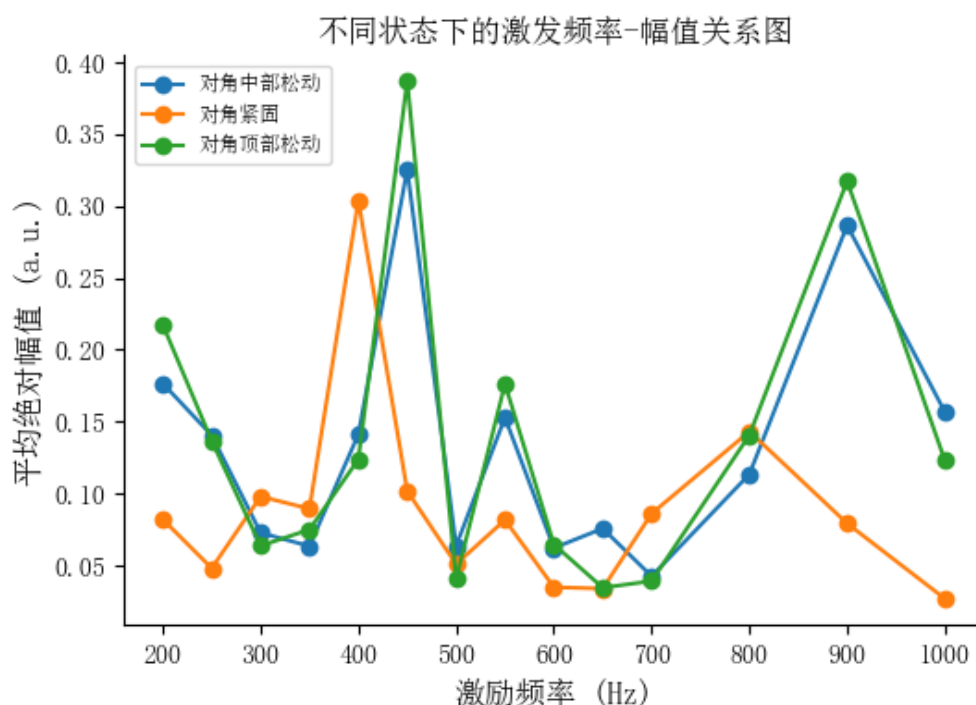


图 3.5 不同状态下的激发频率—幅值关系图

如图 3.5 所示，实验结果表明，在各工作状态下，铁塔结构的声纹响应幅值随声发激励频率变化呈现出明显的峰值特征。综合分析对角紧固状态、对角中部松动状态以及对角顶部松动状态下的频率—幅值关系曲线可以发现，不同工况下铁塔结构的声纹响应特性既具有一定共性，又表现出与螺栓连接状态相关的差异性。在对角中部松动状态和对角顶部松动状态下，系统在约 450 Hz 附近均出现明显的幅值峰值，且二者的整体变化趋势保持一致，这表明，450 Hz 对应铁塔结构的低阶主导共振模态，其振型以结构整体振动为主，对局部松动位置不敏感。而在对角紧固状态下，该频率处的响应幅值相对较低，其对应的峰值激励频率为 400 Hz。

其次,螺栓紧固和松动两种情况下峰值频率的幅值不同,如果排除设备的可靠性和测量误差的影响,可能原因是由于螺栓的松动改变了铁塔的结构,从而改变了它的共振频率与相应的谐振响应。这是因为塔的振动相当于一个复杂的弹簧-质量系统,它的共振频率是由其质量分布和刚度决定的,螺栓的松动可能导致铁塔的刚度(即抵抗弯曲或形变的能力)下降。刚度的下降会降低系统的共振频率,增加振幅。螺栓的松动可能会改变铁塔的质量分布。质量分布的改变也会改变系统的共振频率。这就需要结合结构动力学进一步研究,包括模态分析,例如有有限元分析,进一步详细的实地测试和分析,以确定螺栓松动如何具体影响铁塔的共振频率和振动响应。

有趣的是,从不同工况下的频率—幅值曲线可以观察到,对角紧固状态在约 400 Hz 与 800 Hz 附近呈现较高响应,而对角松动状态在约 450 Hz 与 900 Hz 附近呈现较高响应。该现象表明铁塔系统在上述频段存在较强的频率选择性响应。考虑到螺栓松动将引起连接刚度降低并改变结构动力学特性,峰值频率在不同工况下发生一定偏移是合理的,尤其在低频段可理解为结构主导模态共振频率的变化。同时,由于 400/800 及 450/900 具有近似倍频关系,上述高频段响应亦可能叠加了系统非线性或激励—声学—测量链路的频响特性。为避免将链路特性误判为结构模态,本文将该频段响应视为“结构局部模态与测量链路共同作用”的结果,并在后续声发频率选取时优先考虑低频段稳定峰值附近的频率,以保证激励效果的稳定性与可复现性。

为进一步探究输电铁塔的共振频率分布范围,本文选取紫蓬 220kV 送出工程中的部分典型铁塔,在声激励装置与传感器呈对角布设状态下,采集其声纹响应信号并利用前述共振频率识别方法处理并进行可视化分析,结果如图 3.6 所示。分析结果表明,A2、C2、B2、B20 四座铁塔的最大幅值所对应的共振频率分别为 500 Hz、300 Hz、450 Hz 和 500 Hz。尽管这些铁塔结构型号相同,但其共振频率仍存在一定差异,说明即使在同一线路工程中,由于铁塔自身状态差异及所处环境条件的不同,其动力学响应也存在一定波动性。造成该差异的原因可能包括铁塔构件材料的局部老化、螺栓连接状态的差异、塔基与地基耦合程度不同,以及外部环境因素(如风荷载、湿度、温度等)的变化。这也再次印证了输电铁塔作为一种大型复杂机械结构,其共振特性受到多种耦合因素共同影响。

通过对实训基地封闭环境与紫蓬 220 kV 工程现场的大量样本数据进行整合与比对分析,初步可将当前 220 kV 系列输电铁塔的共振频率区间归纳为 300 Hz 至 500 Hz。该频率区间可作为后续现场检测中的参考频段,若对其他型号铁塔进行数据采集和声纹检测时,亦可基于本文提出的共振频率提取方法实现快速初筛与频段划定,从而为状态评估与故障定位提供数据支持。

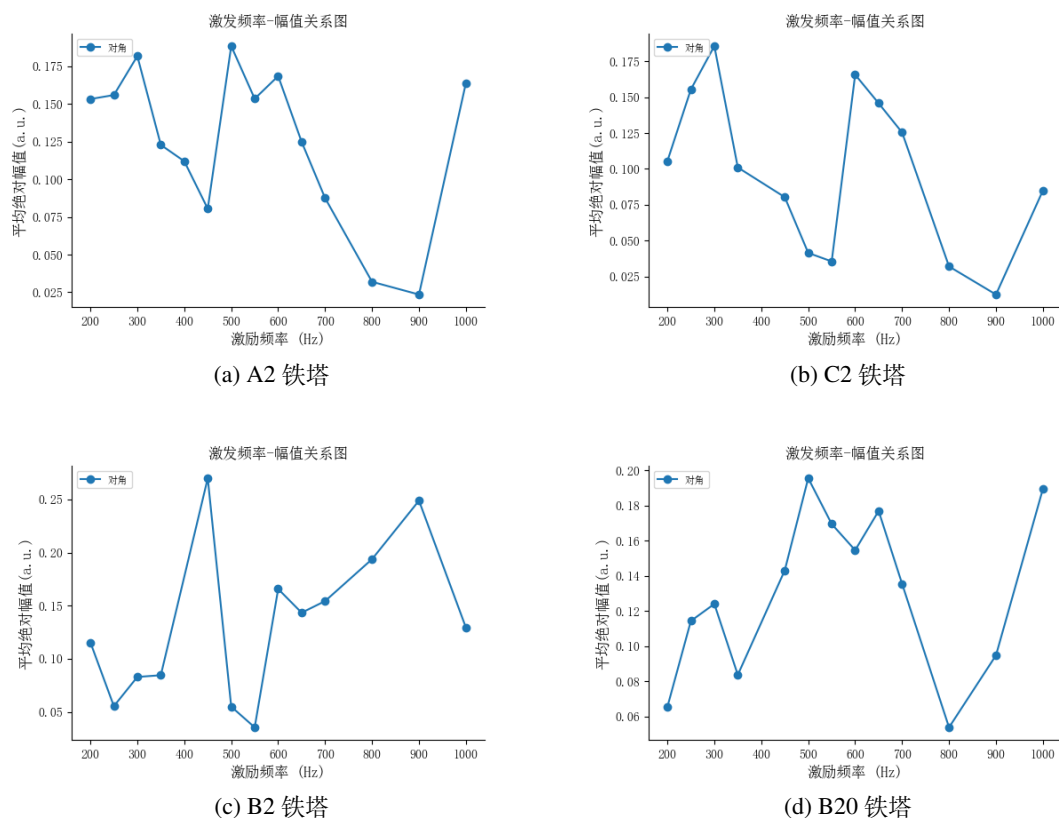


图 3.6 紫蓬 220kV 送出工程部分铁塔激发频率与幅值关系图

3.5 基于声纹信号的铁塔螺栓松动与紧固特征分析

在螺栓松动检测的研究中，声纹数据分析作为一种重要的技术手段，通过捕捉材料内部微观结构变化引起的弹性波，能够提供关于螺栓状态的关键信息。声发射信号在受到外力作用时，因螺栓连接状态的变化而产生不同的频率成分和能量分布。因此，分析这些信号的频谱和时间-频率特性成为螺栓松动诊断的核心任务。语谱图和频谱图是分析声纹信号的两种重要工具，它们能够从不同的角度揭示信号的频率和能量变化，从而帮助识别螺栓在不同状态下的行为特征。

语谱图（Spectrogram）通过短时傅里叶变换对信号进行时间-频率分析，能够展示信号在时间和频率上的分布，尤其适用于分析非平稳信号，如螺栓在松动过程中由于接触非线性和摩擦变化引起的频率成分变化。语谱图通过颜色的深浅表示信号的强度变化，能够直观地反映螺栓状态的时变特性，尤其是松动过程中信号的频率和能量分布的变化。相比之下，频谱图是通过快速傅里叶变换（Fast Fourier Transform, FFT）将时域信号转换为频域表示，能够为松动与紧固状态下的螺栓提供静态的频率分布视图。频谱图有助于揭示在不同工况下，螺栓的频率响应差异，尤其是在松动状态下，频率成分的改变通常伴随较强的非线性效应，能够作为松动状态诊断的有力依据。

为了深入分析不同工况下的声纹数据差异，本文采集了紫蓬 220kV 送出工

程、张洼实训基地和宏源工业区三个场景的不同数据。数据预处理过程包括去除直流偏移、归一化处理、滤波以及应用窗函数，保证数据的平稳性与频率成分的清晰度。然后，使用快速傅里叶变换和短时傅里叶变换分别得到频谱图和语谱图。这些图形能够有效展示螺栓在紧固与松动状态下的特征差异。由于紫蓬220kV送出工程场景中所采集的数据均为无标签数据，因此采用孤立森林模型对其中的异常点进行初步筛选。

3.5.1 基于孤立森林的异常点检测

孤立森林（Isolation Forest）是一种常用的无监督学习方法，专门用于异常点检测。其基本原理是通过随机选择特征并在该特征上随机选取切分点，将数据逐步划分为若干子集。异常点由于其在数据分布中的稀有性，通常在划分过程中较快地被“孤立”出来。因此，孤立森林能够有效地识别数据中的异常点。该方法的计算效率较高，适用于大规模数据集的异常检测，且无需标签信息。

（1）孤立森林的数学原理

孤立森林算法通过构建多棵随机树来逐步将数据进行划分。每棵树的构建过程如下：1. 随机选择一个特征，并在该特征上随机选择一个切分点，将数据分成两个子集。2. 对每个子集重复上述过程，直到每个子集仅包含一个数据点，或者达到预设的树深度。对于每个数据点，孤立森林算法计算其路径长度，即从根节点到该点的路径长度。异常点由于与其他点的差异较大，其路径通常较短，而正常点的路径较长。孤立森林的异常评分基于路径长度进行计算，路径越短，异常评分越高，表明该数据点越可能是异常点。

孤立森林的异常评分计算公式为：

$$s(x) = 2^{-\frac{E_h(x)}{c(n)}} \quad (3.25)$$

其中， $E_h(x)$ 表示样本 x 在树 h 中的平均路径长度， $c(n)$ 是一个基于样本数量 n 的常数，通常定义为：

$$c(n) = 2 \cdot \left(\frac{n-1}{2} \right) - 1 \quad (3.26)$$

最终，孤立森林的异常评分是多棵决策树对该点评分的平均值。

（2）铁塔音频数据的异常点检测

在本研究中，结合紫蓬送出工程中的12座铁塔的无标签数据，我们应用孤立森林方法对t-SNE降维后的音频数据进行异常点检测。具体来说，音频数据首先经过预处理，提取频域特征，形成1000维的特征向量。每个特征向量表示该信号在不同频段上的能量分布。然后通过孤立森林模型，我们能够识别出与其他数据分布显著不同的异常点，进而推测出该铁塔螺栓的松动或紧固状态。

如图3.7所示，异常点的检测过程包括以下步骤：具体步骤如下：

- **数据预处理**：将时长 15 秒的音频信号划分为每段 1.5 秒的时间窗口，并计算每个频段的能量，得到一个 10×1000 维的特征向量。
- **特征标准化**：对所有特征进行标准化处理，以确保不同频段之间的可比性。
- **t-SNE 降维**：对所有铁塔的音频数据进行 t-SNE 降维，得到二维特征向量，用于可视化分析。
- **异常点检测**：使用孤立森林算法对降维后的数据进行异常点检测，识别出与其他样本明显不同的音频数据点。

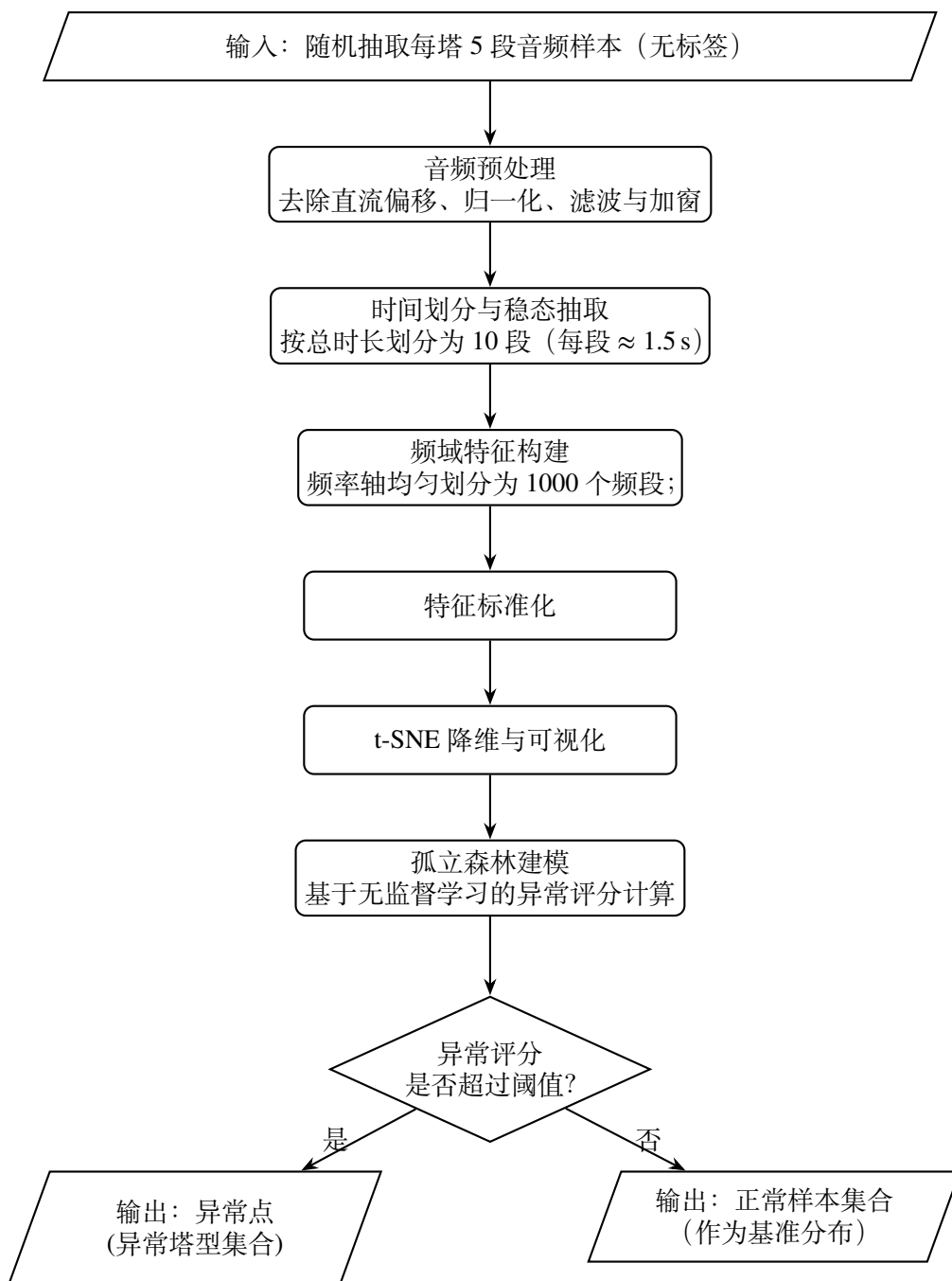


图 3.7 基于 t-SNE 与孤立森林的无标签音频异常检测流程

基于该方法，我们识别出了在降维后特征空间中与其他铁塔数据分布显著不

同的铁塔样本，图3.8a显示了这 12 座铁塔在 t-SNE 降维后的数据分布，图3.8b展示了这些异常点的分布情况。这些异常样本的分布特点表明它们可能处于不同的螺栓状态下，特别是 B2、B8、B20 和 C8 四座铁塔，需要优先关注这些铁塔的松动或紧固状态。

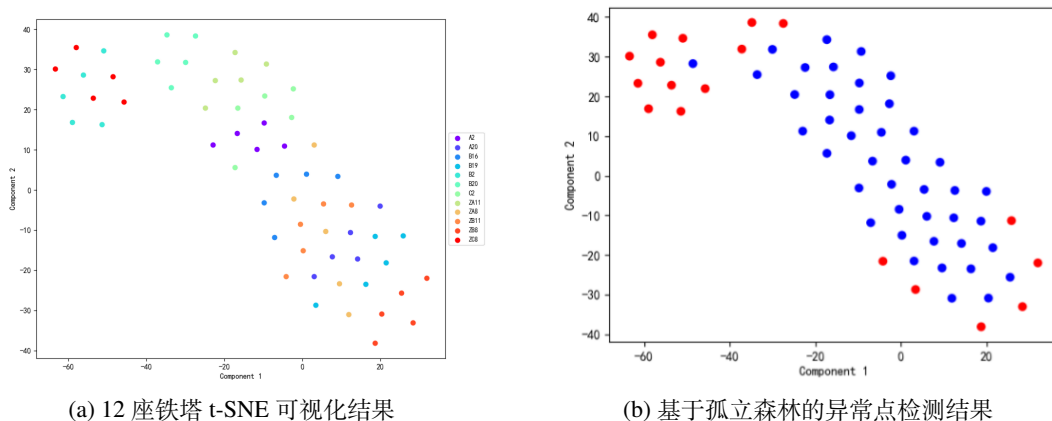


图 3.8 t-SNE 可视化与异常点检测结果

为进一步刻画上述四座铁塔与其余塔型在声学响应上的差异性，本文在 450 Hz 激励条件下，采用短时傅里叶变换对四座铁塔及其他铁塔的音频数据进行时频分析，并以语谱图形式进行可视化展示。如图 3.9所示，对比结果表明，塔型 B2、B20 与 C8 的语谱图在时频能量分布上与其他塔型存在显著差异：其背景宽带成分相对较弱、整体能量分布更为集中，呈现出更“干净”的时频结构。相比之下，B8 塔型的语谱图与其余铁塔语谱图在能量成分分布上较为相似，表现为整体能量水平偏高且宽带成分更为明显，因此在该工况下被视作与多数塔型不同的特殊样本。基于上述时频特征，可初步将 B2、B20 与 C8 归为一类，而将 B8 与其余塔型归为另一类。在此基础上，进一步对归一化后的信号进行快速

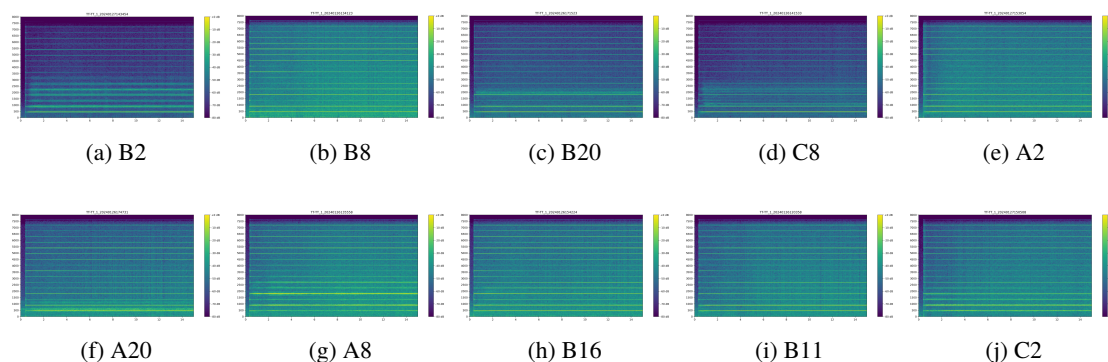


图 3.9 10 座铁塔音频信号语谱图对比

傅里叶变换频谱分析。如图3.10所示，B2、B20 与 C8 三座铁塔在低频与高频区间的谱线分量相对较少，频谱结构更为简洁，表明其频率成分较为单一。相比之下，其余铁塔的部分样本在 0–1000 Hz 范围内存在不同程度的低频成分，同时伴

随较强的高频分量，频率成分更为复杂。并且在 450 Hz 激励下，B2、B20 与 C8 对应主频幅值分别约为 30000、25000 与 40000，而在其他激励频率下的幅值均低于 20000，B8 塔型主频幅值整体低于 10000。

结合相关文献中关于螺栓松紧状态对结构振动和声学响应影响的结论^[73]，本文推测 B2、B20 与 C8 塔型的多数螺栓处于相对紧固状态；而其余塔型可能存在不同程度的螺栓松动，其中 B8 塔型的松动特征更为突出。上述推断仍需在后续工作中结合现场检测或力矩测量等手段进一步验证。

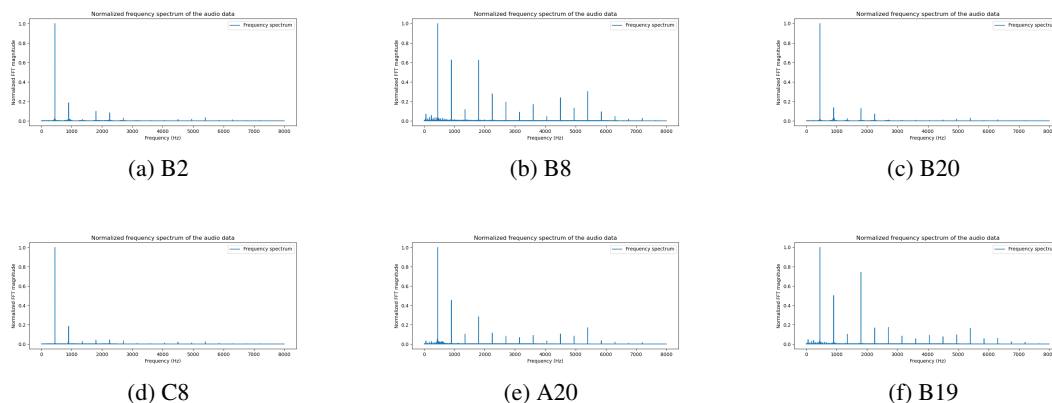


图 3.10 铁塔（B2、B8、B20、C8、A20、B19）音频信号频谱图对比

3.5.2 螺栓松动声纹特征对比分析

为进一步验证螺栓紧固与松动状态下声纹特征的差异性，并对前文基于紫蓬送出工程无标签数据分析得到的结论进行现场佐证，本文在另外两个场景中开展了螺栓松动状态的有标签实验研究。与工程现场采集的自然运行数据不同，本实验基于明确标注松动状态的声纹数据集，通过人为控制螺栓松动程度和位置，构造多种可复现工况，并采集对应声纹信号，系统分析不同工况下的声学响应特征。该部分试验的目的在于观察不同松动状态下声纹信号在频域与时频域的变化规律，并且提炼具有工程可解释性的差异特征，为后续松动识别模型与激励策略提供依据。

1. 试验工况设计与标签体系

考虑铁塔结构受力传递路径及连接形式的差异，本文将试验铁塔按空间位置划分为底部、中部与顶部三个关键区域，并分别设计松动工况以模拟实际运行中可能出现的典型状态：

- **底部区域：**从完全紧固到不同部件松动，涵盖铁塔杆件单侧松动、两侧松动，以及防盗螺栓、接地螺栓等典型螺栓类型的松动情形。
- **中部区域：**重点研究松动螺栓数量对响应的影响，设置中部 2/4/6 颗螺栓松动等工况。

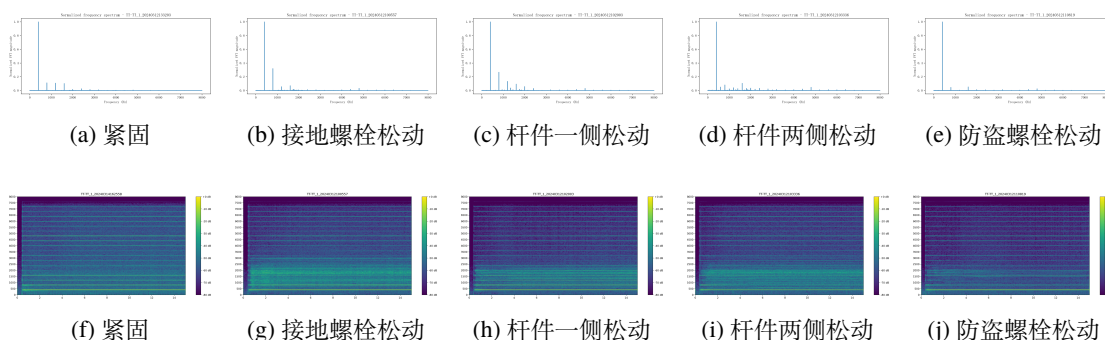


图 3.11 不同状态下的特征对比：第一行为频谱图（FFT），第二行为语谱图（Spectrogram）。

- **不同空间位置对比：**在底部、中部、顶部等不同位置分别设置松动，比较位置差异对声纹响应的影响。

上述工况均为带标签数据（紧固、松动类型/松动数量/松动位置），因此能够在“状态已知”的前提下对声纹差异进行针对性分析，与紫蓬工程无标签数据形成互补。

2. 频谱图与语谱图对比分析

为了系统探究铁塔螺栓松动的特征差异，本文采用快速傅里叶变换与短时傅里叶变换对张洼实训基地所采集的三类典型工况声纹数据进行频域与时频域可视化分析。图3.11给出了不同零部件在紧固与松动状态下的频谱图与语谱图示例。通过对多组样本的频谱图进行综合对比可以发现：在紧固状态下，频谱能量主要集中于声发激励频率及其相关分量，除主峰及少量谐波外谱图整体较为“干净”，与紫蓬 220 kV 送出工程中 B2、B20 与 C8 三座铁塔紧固样本的频谱形态具有一致性；而在松动状态下（包括接地螺栓松动、杆件松动等），在 0–2000 Hz 频段内可观察到更为明显的谱线波动与能量扩散现象，即除激励频率及谐波成分外出现更多非主导频率分量，表明松动引入了额外的振动模式或接触非线性效应，使能量由主峰向更宽频带扩散。

与频谱图提供“静态频率分布”不同，语谱图能够反映信号能量在时间与频率上的联合分布，从而更适用于观察松动引起的非平稳成分与能量弥散趋势。在语谱图层面，整体能量随时间变化较为稳定，说明试验采用的连续稳定声发激励具有较好的稳态特性；然而，相较于紧固状态，松动状态下在低频段，尤其是 0–2000 Hz 范围内出现更强烈的能量响应，表现为更明显的离散噪点与能量弥散区域。这一现象可以理解为螺栓松动导致预紧力下降，使连接界面产生微滑移、摩擦或间隙碰撞，从而引入宽带能量成分并增强低频响应。

值得注意的是，本文在对不同部件松动样本进行对比时发现，防盗螺栓松动所对应的频谱与语谱变化相对不如接地螺栓或杆件松动明显。输电铁塔防盗螺栓作为电力系统的重要安全防护装置，通常采用双母防盗结构：内侧为普通紧固螺帽，主要承担连接强度与预紧力提供；外侧为专用防盗螺帽，具有特殊结构设

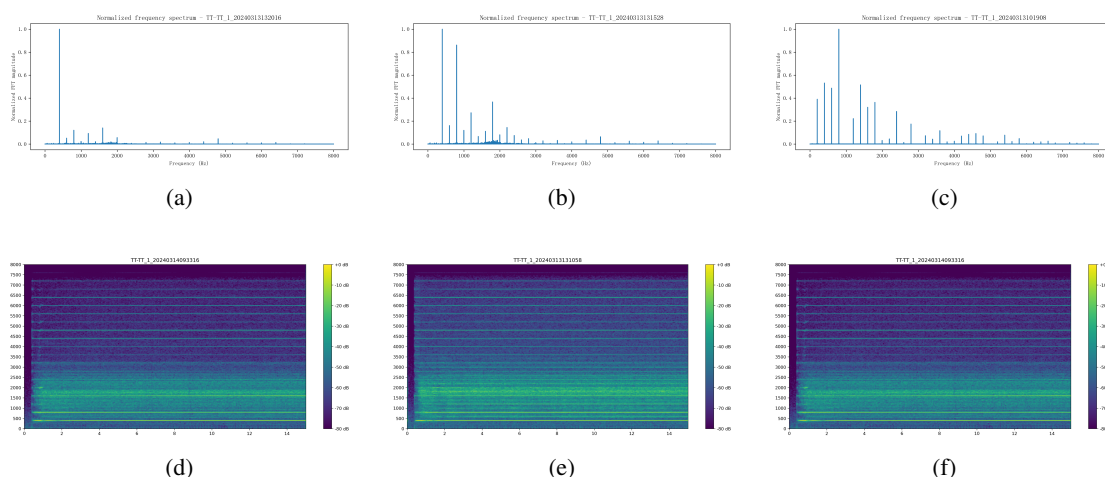


图 3.12 中部松动螺栓数量为 2、4、6 颗时的特征对比。上排为频谱图 (FFT)，下排为语谱图 (Spectrogram)，从左到右依次对应松动螺栓数量为 2 颗、4 颗、6 颗。

计，用于防盗与防拆。由于该结构的受力路径与接触关系更为复杂，且外侧防盗螺帽的松动不一定直接对应连接预紧力的显著下降，因此其松动引起的结构等效刚度变化可能相对有限，从而导致声纹信号差异弱于其他关键受力部位的松动。这一观察说明在后续基于声纹的松动检测研究中，需要针对防盗螺栓这类特殊结构设置更具针对性的特征表征或激励策略，以避免漏检或误判，但本文研究侧重于威胁到铁塔结构稳定性的关键部位螺栓松动，而防盗螺栓松动往往不会对铁塔的结构稳定造成明显的影响，故不做过多的研究拓展。

进一步地，对于中部不同数量螺栓松动工况，图3.12表明其频谱与语谱变化规律与上述结论一致：随着松动螺栓数量从 2 颗增加到 4 颗、6 颗，频谱图中非主导谱线分量呈增加趋势，谱线形态更为杂乱，语谱图中低频段与宽带区域的能量响应亦更为显著。该现象说明以松动数量表征的松动程度与系统非线性程度或者能量扩散水平具有一定相关性，松动数量增多会降低连接整体刚度并增强接触非线性效应，从而使响应能量更易分散到更宽频带。

对于不同空间位置的螺栓松动对比结果，如图 3.13 所示，可观察到随着松动位置与接收传感器之间距离增大，频谱与语谱中松动引起的差异特征逐渐减弱。这一规律现象可能的物理解释为：声纹信号在结构传播与接收过程中会受到传播距离、路径复杂性以及连接节点不确定性等因素影响，导致能量衰减与随机扰动增强；当松动位置距离传感器更远时，局部松动引起的差异成分在传播过程中更易被耗散或被背景噪声掩盖，从而在可视化结果中呈现“特征淡化”的趋势。该结论也提示在工程应用中，传感器布设位置与监测覆盖范围会直接影响对特定部位松动的可检测性，后续研究需结合布设优化与特征增强方法提高远端松动的检出能力。

图3.14给出了底部不同状态下的频谱图激励幅值变化。通过图片可观察到：

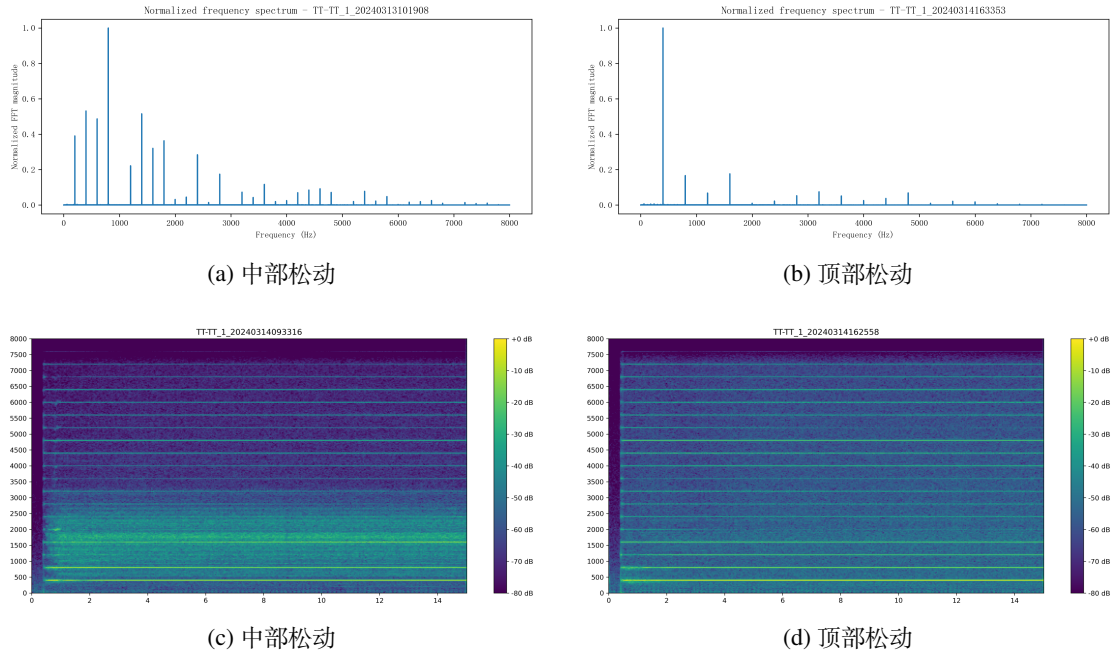


图 3.13 中部松动与顶部松动状态的特征对比。第一行为频谱图 (FFT)，第二行为语谱图 (Spectrogram)。

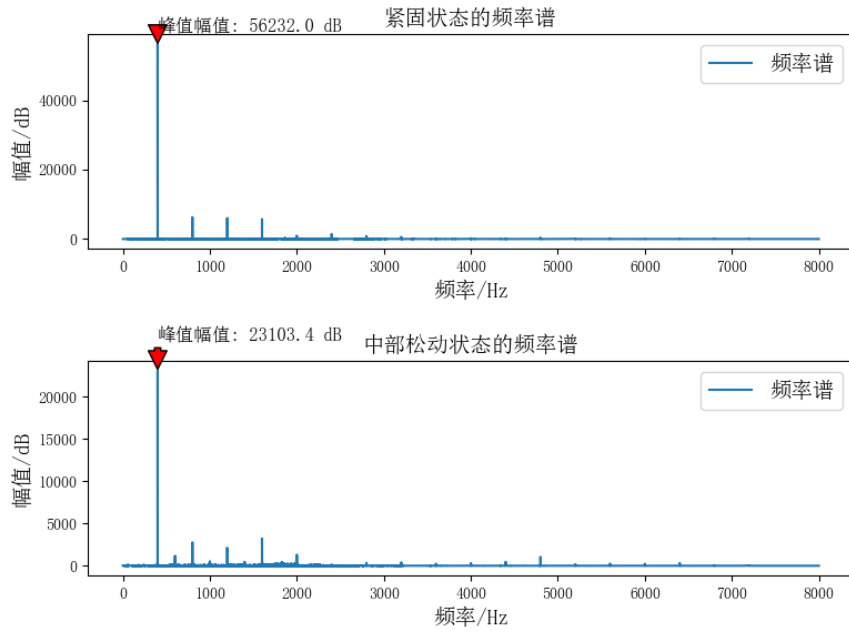


图 3.14 螺栓不同状态激励幅值频谱图

在紧固状态下，激励频率对应的主峰幅值整体更高，部分样本峰值可达到 5×10^4 量级；而松动状态下，激励主峰幅值显著下降，部分样本低于 2×10^4 量级。由于试验采用固定频率输出连续稳定声波，声发输入能量可视为相对恒定。出现松动时，连接界面可能产生微滑移、碰撞或摩擦等接触非线性效应，从而引入额外的“杂讯”成分。该现象在频谱上表现为，除激励频率及其谐波附近外，其他频率处出现更多非零能量成分，导致能量从主峰向更宽频带扩散；相应地，激励主

峰幅值会因能量分流而降低。

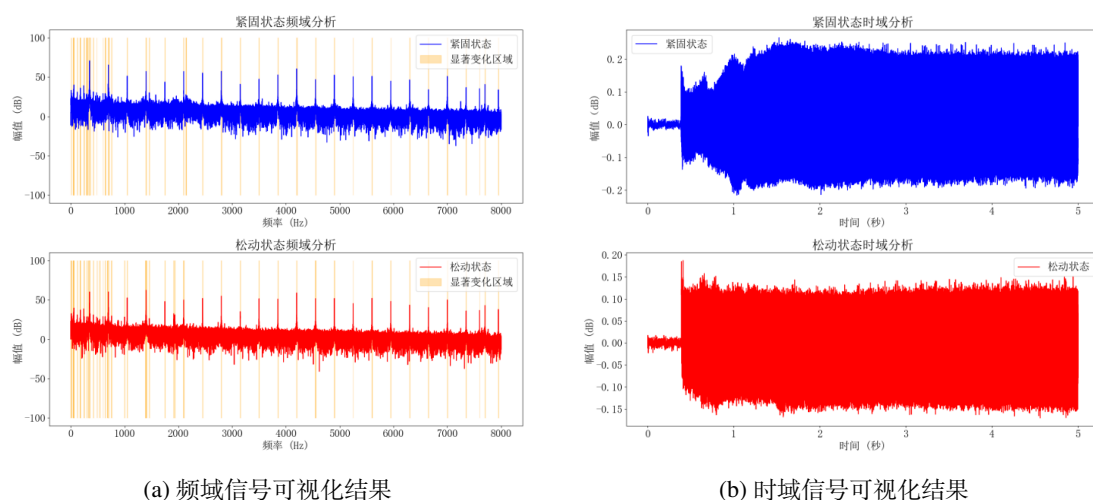


图 3.15 音频信号的时域与频域可视化分析

除张洼基地可控工况外,本文进一步对宏源工业区采集数据进行频域与时域可视化分析。图3.15a的频域可视化结果也呈现一种趋势,在低频段(约0–2000 Hz)范围内,松动状态呈现更多频率成分与更显著的变化区域;从图3.15b对应的时域波形中可以看出,紧固状态的频谱能量整体更集中、幅值相对更高,松动状态的幅值整体较低,表明在连接不稳定条件下,系统有效响应能量可能因摩擦或碰撞等机制而产生更强的耗散。

3.6 本章小结

本章围绕输电铁塔螺栓松动声纹检测技术,完成了数据采集方案的设计与多维度特征分析,针对现场复杂环境与结构响应耦合问题,提出了有效的解决方案,为后续算法建模与工程应用提供了实验基础与参数依据。具体而言,构建了覆盖紫蓬 220 kV 送出工程、张洼实训基地及宏源工业园区的声纹数据采集体系,确保了无标签真实数据与有标签数据的互补支持;通过对比分析声激励装置与传感器布设方式,确定了对角布设方案作为标准,保障了实验数据的稳定性;提出并验证了基于稳态时间窗与幅值统计特征的共振频率识别方法,为声激励频率的优化选取提供了依据;最后,通过频谱与语谱图分析,验证了螺栓松动状态下声纹特征的差异,并揭示了螺栓类型、松动位置与传感器布设对识别效果的影响机制。总的来说,本章为后续的螺栓松动状态识别与模型构建奠定了坚实的数据支持和关键参数依据。

第4章 基于多尺度能量注意力残差网络的螺栓松动识别方法

针对实际场景中的输电铁塔螺栓松动检测问题,本章提出一种面向声纹识别任务的鲁棒特征学习方法:在网络结构层面,设计多尺度能量注意力模块(Multi-Scale Energy Attention Module, MSEAM),通过多尺度池化获取不同时间-频率分辨率下的响应信息,并基于频带能量统计生成自适应权重,实现对关键频带特征的增强与对无关成分的抑制,从而提高模型对松动引起的能量扩散与频带选择性变化的表征能力;在训练策略层面,引入基于噪声增强的一致性正则化(Consistency Regularization),对输入声谱图施加随机信噪比噪声与 SpecAugment 遮挡扰动,并约束同一样本在“干净/扰动”条件下的特征表示保持一致,以减少模型对噪声与遮挡的敏感性,提升模型泛化能力。基于上述设计,本文构建了端到端的声纹识别网络(MSEAM-Net),并在三类数据集上开展系统实验验证。

4.1 引言

在上一章中,本文围绕多场景声纹数据采集、声激励装置与传感器布设方式、铁塔共振频段识别以及紧固/松动状态的时频特征差异进行了实证分析,验证了声纹信号在螺栓状态表征方面的可行性。然而,在面向工程现场的螺栓松动识别任务中,仍存在若干关键挑战,例如现场采集信号往往受到风噪、环境噪声、设备链路频响与传播路径差异等因素影响,导致同一状态样本在特征空间中出现较大波动;另一方面,螺栓松动属于典型的界面接触非线性问题,松动会引发声信号的能量扩散、谐波增强与宽带噪声纹理变化,而上述差异化特征多集中在特定频带或多尺度时频结构中,若检测模型缺乏针对性的特征增强机制,这类微弱特征极易被背景扰动掩盖,最终降低识别结果的稳定性与工程鲁棒性。

针对上述问题,本章构建了一种端到端的声纹识别模型。通过结合3.5节中对紧固与松动状态时频差异的分析结果,模型在整体结构上引入了注意力机制,构建以频带能量扰动为导向的特征选择路径,从而刻画螺栓松动状态下关键频带能量的波动特征。同时,利用一致性正则化的训练策略来进一步提升模型对环境噪声干扰的适应能力。

本章的主要研究内容与组织结构如下:第4.2节对现场采集信号中主要干扰因素的类型、特征及影响进行分析;第4.3节介绍 MSEAM-Net 的整体框架与残差骨干网络,并阐述基于噪声增强的一致性正则化训练目标与优化方法;第4.4节介绍了所用数据集、预处理流程与评价指标,并与 VGG16、ResNet18 等典型网

络进行对比实验；进一步开展鲁棒性分析、模型参数量与计算复杂度分析以及消融实验；最后在更贴近工程应用的任务上进行扩展验证，从而为后续工程部署与应用提供依据；第4.5节对全章内容进行总结。

4.2 现场采集信号干扰类型、特征及影响分析

在工程现场条件下，输电铁塔螺栓松动声纹信号除包含目标结构状态信息外，还会受到多种外部干扰和系统性变化因素的共同影响。前者主要表现为加性噪声，后者则更多体现为测量链路及传播过程引起的响应失真或样本分布偏移。若不对上述因素进行分析并在模型设计中加以考虑，则同一螺栓状态下的样本可能在特征空间中产生较大波动，从而降低模型的稳定性与跨场景泛化能力。

如表4.1所示，从来源上看，现场干扰主要可分为两类：一类为加性干扰，包括风噪、环境噪声以及采集链路背景噪声等；另一类为系统性差异，包括设备链路频响差异、结构传播路径差异以及安装耦合条件变化等。其中，风噪通常来源于自然风场及塔体周围气流扰动，具有明显的随机性和非平稳性；环境噪声则主要来自车辆、人为活动、机械运行及鸟鸣等外界声源，常表现为宽带或突发性纹理干扰。与此不同，设备链路频响差异和传播路径差异并不直接表现为附加噪声，而是会改变目标响应在不同频带上的幅值分布、时延结构和能量扩散形态，从而使得不同场景、不同点位下采集到的同类样本呈现出不一致性。

为便于后续方法设计，记理想结构响应为 $s(t)$ ，设备链路冲激响应为 $h_c(t)$ ，结构传播路径响应为 $h_p(t)$ ，加性噪声为 $n(t)$ ，则现场采集信号可简化表示为

$$x(t) = h_c(t) * h_p(t) * s(t) + n(t), \quad (4.1)$$

其中 $n(t)$ 综合表征风噪、环境噪声及背景电路噪声等加性干扰，而 $h_c(t)$ 与 $h_p(t)$ 则分别表征设备链路频响差异和结构传播路径差异对目标响应的调制作用。

进一步地，考虑到不同类型加性干扰在时频域中的表现并不完全一致，本文将 $n(t)$ 细化表示为

$$n(t) = n_w(t) + n_e(t) + n_b(t), \quad (4.2)$$

其中， $n_w(t)$ 表示风噪声， $n_e(t)$ 表示外部环境噪声， $n_b(t)$ 表示采集链路背景噪声。风噪声通常具有低频能量占优、随时间变化明显和非平稳性较强等特点，容易在低频区域掩盖螺栓敲击响应中的弱能量成分；环境噪声则可能呈现宽带分布或短时突发特征，会在时频图中形成局部异常纹理；背景电路噪声一般幅值较低但持续存在，会降低整体信噪比。由于螺栓松动状态的判别信息主要体现在敲击响应的瞬态能量分布、频带衰减差异以及局部时频纹理变化中，上述加性干扰可能导致有效特征被淹没、局部频带响应失真以及同类样本内部差异增大。因此，仅

将加性噪声视为统一的随机扰动是不充分的，需要在训练阶段结合其时频特征进行更有针对性的鲁棒性约束。

基于上述分析，本文在模型设计中采用结构增强与训练约束的联合策略来提高对现场干扰的适应能力。针对设备链路频响差异和传播路径差异可能引起的频带能量重分布，采用多尺度能量注意力机制，对不同频带的判别性响应进行自适应加权，以突出与螺栓松动相关的关键能量模式，而对风噪、环境噪声和局部频带缺失等不确定扰动，在训练阶段引入噪声增强、遮挡增强以及一致性正则化约束，使模型在“干净样本—扰动样本”之间学习到更加稳定的表征。由此，前述干扰分析不仅为后续自适应增强策略提供了理论依据，也为 MSEAM-Net 在复杂现场条件下的稳健识别能力提供了方法支撑。

表 4.1 现场采集信号主要干扰类型及其影响

干扰类型	主要来源	主要特征	主要影响
风噪	自然风场、局部气流扰动	非平稳、低频占优、能量随时间波动	掩盖低频及弱瞬态响应，降低信噪比并引起特征波动
环境噪声	车辆、人为活动、机械运行等	宽带、突发、时变、局部纹理明显	引入非目标时频成分，干扰局部纹理判别并增大类内波动
设备链路频响差异 传播路径差异	激励器、传感器、线缆及采集链路差异 测点位置、高度层及节点连接关系差异	频带增益不一致、样本间存在偏置 时延、衰减及能量扩散不同	改变谱形分布，降低跨场景一致性 增大同类样本离散度，影响模型判别边界

4.3 多尺度能量注意力残差网络的方法设计

本节针对输电铁塔声纹识别在实际工况下面临的噪声扰动等问题，从网络结构与训练策略设计两方面提出基于多尺度能量注意力与一致性正则化的螺栓松动识别模型（MSEAM-Net）。其核心思想主要包括：一方面在特征提取阶段引入多尺度能量驱动的频带注意力机制，增强与螺栓松动相关的关键频带响应；另一方面通过噪声增强与遮挡扰动构造一致性约束，促使模型对输入扰动具有稳定表征，从而提升鲁棒性与泛化能力。

4.3.1 MSEAM-Net 的整体框架

在前述多场景数据分析中可观察到：紧固状态下的声谱能量分布通常较为平滑，主要集中于激励频率及其相关分量；而松动状态往往伴随低频窄带高能条纹与特定频段宽带抬升等现象，为便于后续注意力机制的建模与表述，本文将松动状态下常见的能量异常形态概括为“窄带峰值”和“宽带抬升”两类。所谓窄带峰值，是指声谱能量在少量频点或极窄频带内显著增强，表现为频谱图中的尖

峰或语谱图中的细条纹，其通常与外部激励频率、谐波分量或结构共振模态响应相关；而宽带抬升是指在较宽频段内能量整体水平上移，表现为语谱图背景亮度提升或频谱底噪抬高，反映松动引入的摩擦、微滑移与间隙碰撞等接触非线性效应导致的能量扩散。因此，若仅依赖单尺度卷积提取特征，模型容易受到共振偏移与背景扰动影响，难以稳定捕获“窄带峰值 + 宽带抬升”这一复合能量模式。为此，本文引入多尺度池化构建不同感受野的谱图表征，并进一步以频带能量统计作为驱动信号，构建频带-通道联合注意力，以实现关键频带的自适应加权。

为实现复杂野外场景下的螺栓松动识别，本文首先将原始声纹波形转换为统一规格的时频表征，并在此基础上构建“编码器 + 分类器”的两阶段识别框架。

(1) 声学前端与输入表征

设第 b 条声纹样本的原始波形为 $\mathbf{x}_b(t)$ 。本文采用 Mel 滤波器组将其映射为 N 带 Mel 谱图 $\mathbf{S}_b \in \mathbb{R}^{1 \times N \times T}$ ，其中 N 为 Mel 频带数， T 为时间帧数。为压缩动态范围并减弱幅值尺度差异的影响，对 Mel 谱取对数并进行样本级标准化：

$$\tilde{\mathbf{S}}_b = \frac{\log(\mathbf{S}_b) - \mu_b}{\sigma_b + 10^{-9}}, \quad 0 \leq b < B, \quad (4.3)$$

其中 μ_b 与 σ_b 分别为样本 $\log(\mathbf{S}_b)$ 的均值与标准差。

考虑卷积核感受野与输入尺度一致性，本文将时间维截断/补齐到固定长度 N ，从而得到统一尺寸的输入张量：

$$\mathbf{X} \in \mathbb{R}^{B \times 1 \times H \times W}, \quad H = W = N, \quad (4.4)$$

其中 B 为批量大小， H 与 W 分别对应频带维与时间帧维。该设置使得不同采集样本在输入尺度上可比，并便于后续卷积网络稳定训练。

(2) 多尺度谱表征构建

设浅层残差阶段输出特征为

$$\mathbf{F}_1 \in \mathbb{R}^{B \times C \times H \times W}, \quad (4.5)$$

其中 B 为批量大小， C 为通道数， H 与 W 分别表示频带维与时间帧维。给定尺度集合 $\mathcal{S} = \{1, 2, 4\}$ ，对每一尺度 $s \in \mathcal{S}$ 执行自适应平均池化并上采样回原尺寸：

$$\mathbf{Z}^{(s)} = \text{Up}(\text{AvgPool}_s(\mathbf{F}_1)), \quad \mathbf{Z}^{(s)} \in \mathbb{R}^{B \times C \times H \times W}, \quad (4.6)$$

其中 $\text{AvgPool}_s(\cdot)$ 表示将特征池化至 $(H/s, W/s)$ ， $\text{Up}(\cdot)$ 表示双线性插值上采样至 (H, W) 。

将各尺度特征沿通道维拼接得到多尺度融合表示：

$$\mathbf{X}_{\text{MS}} = [\mathbf{Z}^{(1)} \parallel \mathbf{Z}^{(2)} \parallel \mathbf{Z}^{(4)}] \in \mathbb{R}^{B \times C' \times H \times W}, \quad C' = mC, \quad (4.7)$$

该操作同时保留了局部纹理信息与全局上下文信息，为后续能量驱动的注意力建模提供更丰富的输入。

(3) 频带能量统计与能量映射

为量化松动诱发的频带能量变化，本文在时间维上对 $\mathbf{X}_{\text{MS}}$ 进行能量统计，采用 RMS 形式定义频带能量图 $\mathbf{E}$ ：

$$E_{b,c,h} = \sqrt{\frac{1}{W} \sum_{t=1}^W (\mathbf{X}_{\text{MS}})_{b,c,h,t}^2 + \varepsilon}, \quad \varepsilon = 10^{-6}, \quad (4.8)$$

其中 $E_{b,c,h}$ 表示第 b 个样本在通道 c 、频带 h 处的能量强度。由此得到

$$\mathbf{E} \in \mathbb{R}^{B \times C' \times H}, \quad (4.9)$$

其在保持频带维 H 的同时压缩时间维 W ，能够更直接反映不同频带的能量差异。

(4) 频带-通道联合注意力生成

在声纹诊断任务中，不同卷积通道往往对应不同模态响应、谐波结构或纹理模式。若对通道进行全局压缩（如传统 SE 结构）容易导致关键频带信息被均匀化或被局部强响应掩盖。为此，如图 4.1 所示，本文引入频带-通道联合注意力，为每个频带 h 学习一组独立的通道权重，实现对“窄带峰值”和“宽带抬升”两类现象的兼容建模。

具体地，对能量图 $\mathbf{E}$ 首先采用 1×1 的一维卷积沿通道维进行压缩，得到每个频带的低维嵌入：

$$\mathbf{U} = \mathbf{W}_1 \mathbf{E} \in \mathbb{R}^{B \times r \times H}, \quad \mathbf{W}_1 \in \mathbb{R}^{r \times C'}, \quad (4.10)$$

其中 r 为压缩维度。随后经 ReLU 激活并通过第二层一维卷积恢复至 C' 维，得到注意力权重：

$$\mathbf{A} = \sigma(\mathbf{W}_2 \varphi(\mathbf{U})) \in \mathbb{R}^{B \times C' \times H}, \quad \mathbf{W}_2 \in \mathbb{R}^{C' \times r}, \quad (4.11)$$

其中 $\varphi(\cdot)$ 与 $\sigma(\cdot)$ 分别表示 ReLU 与 Sigmoid 函数。由于该过程对每个频带 h 保持独立建模，模型能够学习到频带相关的注意力模式，例如在低频段强调窄带能量峰值，在其他频段捕获更粗粒度的宽带抬升。

最后，将 $\mathbf{A}$ 在时间维上广播扩展为 $\tilde{\mathbf{A}} \in \mathbb{R}^{B \times C' \times H \times W}$ ，并对多尺度特征进行逐点加权：

$$\mathbf{F}'_1 = \mathbf{X}_{\text{MS}} \odot \tilde{\mathbf{A}}, \quad (4.12)$$

其中 $\odot$ 表示 Hadamard 乘积， $\mathbf{F}'_1$ 即 MSEAM 模块输出特征。该机制能够在保持多尺度上下文的同时，突出与松动状态相关的关键频带能量模式，抑制与诊断无关的背景扰动，从而提高跨铁塔、跨场景条件下的特征稳健性与判别能力。

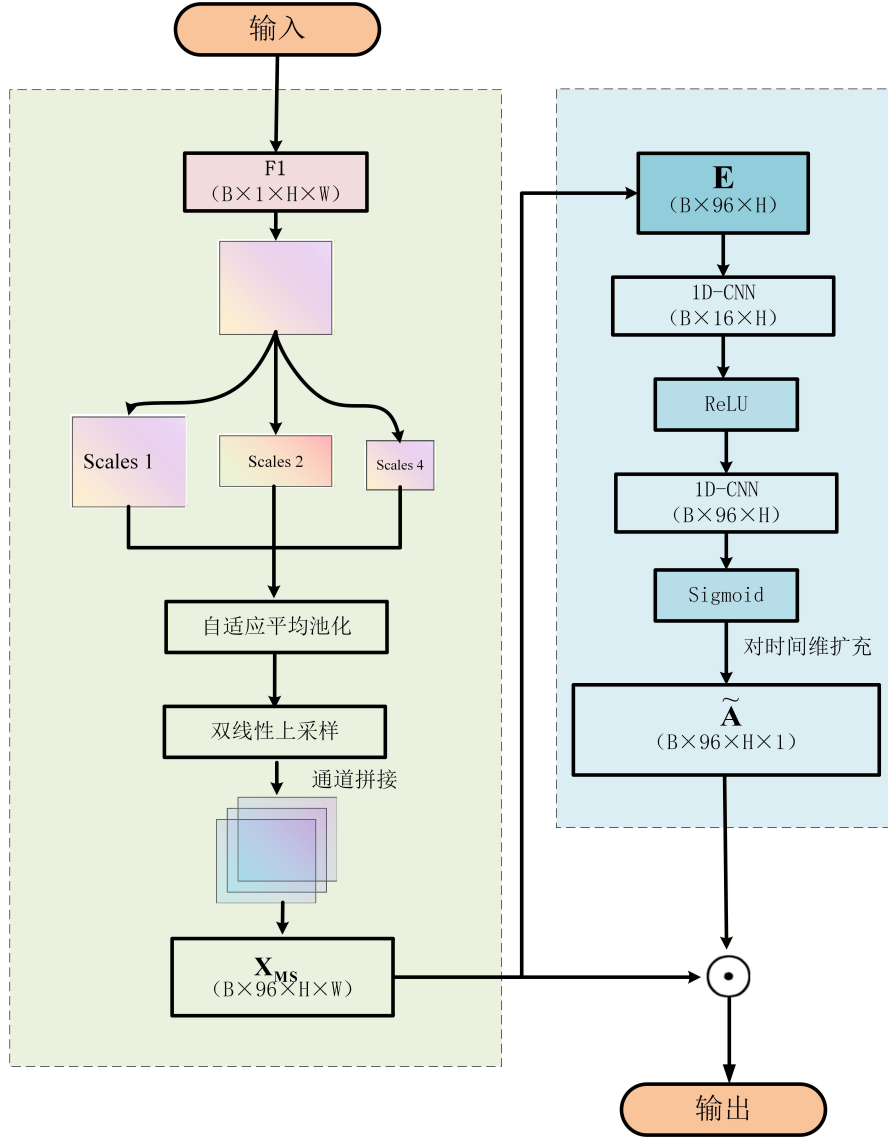


图 4.1 MSEAM 结构图

(5) 编码器 + 分类器两阶段框架

如图4.2, 本文采用“编码器 $f_{\theta}(\cdot)$ + 分类器 $g_{\phi}(\cdot)$ ”的两阶段结构：编码器负责从声谱图中提取判别性表征，输出嵌入向量 $\mathbf{z} \in \mathbb{R}^{B \times D}$ ；分类器将嵌入映射为类别概率 $\hat{\mathbf{y}} \in \mathbb{R}^{B \times K}$ ，其中 K 为类别数。整体前向过程表示为

$$\mathbf{z} = f_{\theta}(\mathbf{X}), \quad \hat{\mathbf{y}} = g_{\phi}(\mathbf{z}). \quad (4.13)$$

(6) 编码器结构

在编码器设计上, 本文以残差网络 (ResNet) 作为主干以获得稳定的深层表征能力, 并在浅层阶段引入多尺度能量注意力模块 (MSEAM)。该模块对频带维生成自适应权重, 用于突出松动状态引起的关键频带能量变化, 从而提升跨塔型场景下的特征一致性与可分性。

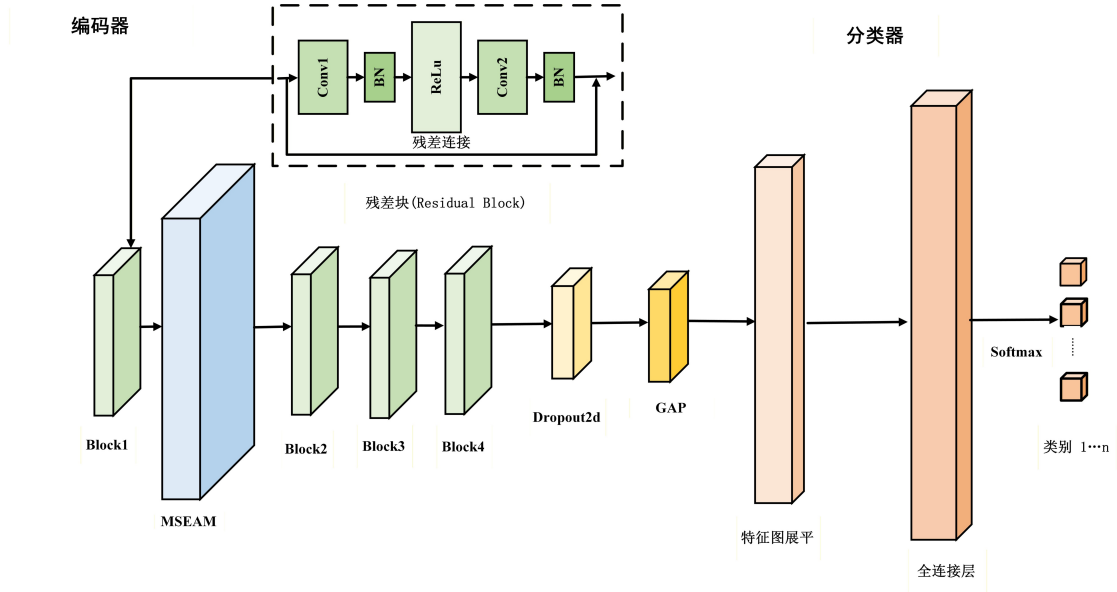


图 4.2 MSEAM-Net 网络结构图（其中 Conv 表示卷积层，BN 表示批归一化，Dropout2d 表示二维随机失活层，GAP 表示全局平均池化。）

具体而言，编码器由四个残差阶段构成，其计算流程为

$$\mathbf{F}_1 = \mathcal{R}_1(\mathbf{X}), \quad \mathbf{F}'_1 = \text{MSEAM}(\mathbf{F}_1), \quad \mathbf{F}_2 = \mathcal{R}_2(\mathbf{F}'_1), \quad \mathbf{F}_3 = \mathcal{R}_3(\mathbf{F}_2), \quad \mathbf{F}_4 = \mathcal{R}_4(\mathbf{F}_3), \quad (4.14)$$

其中 $\mathcal{R}_i(\cdot)$ 表示第 i 个残差阶段。最终通过全局平均池化获得固定长度嵌入：

$$\mathbf{z} = \text{GAP}(\mathbf{F}_4) \in \mathbb{R}^{B \times D}. \quad (4.15)$$

为降低小样本条件下的过拟合风险，本文在池化前引入二维 Dropout 进行空间随机失活，从而提升模型的泛化能力与训练稳定性。

如表 4.2 所示，在本文网络中，输入声谱图首先经 3×3 卷积与浅层残差块提取低阶时频纹理得到 $\mathbf{F}_1$ ，随后引入 MSEAM 生成加权特征 $\mathbf{F}'_1$ 并送入后续残差阶段进行逐级下采样与高层语义提取。最终通过自适应全局平均池化获得定长嵌入向量用于分类。由于 MSEAM 显式建模了频带能量与多尺度上下文，其在后续层中可为松动识别提供更加稳定的判别性线索，并降低因共振频段偏移导致的特征失配风险。

表 4.2 MSEAM-Net 网络结构与参数配置（输入： $1 \times 64 \times 64$ ）

阶段	层/模块	输出尺寸 ($C \times H \times W$)	卷积参数 (核/步幅/填充)	通道变化 (in→out)
S0	残差块 (Residual Block1)	$32 \times 64 \times 64$	$3 \times 3/1/1$	$1 \rightarrow 32$
S1	多尺度能量注意力模块 (MSEAM)	$96 \times 64 \times 64$	—	$32 \rightarrow 96^a$
S2	残差块 (Residual Block2)	$64 \times 32 \times 32$	$3 \times 3/2/1$	$96 \rightarrow 64$
S3	残差块 (Residual Block3)	$128 \times 16 \times 16$	$3 \times 3/2/1$	$64 \rightarrow 128$
S4	残差块 (Residual Block4)	$256 \times 8 \times 8$	$3 \times 3/2/1$	$128 \rightarrow 256$
S5	二维 Dropout (Dropout2d) ($p = 0.5$)	$256 \times 8 \times 8$	—	—
S6	全局平均池化 (GAP)	$256 \times 1 \times 1$	—	—
S7	特征图展平 (Flatten)	256 (向量)	—	—
H1	全连接层 (Fully Connected Layer)	$K = \text{类别数}$	—	$256 \rightarrow K$

^a 由于 MSEAM 包含多尺度特征拼接，通道数由 32 扩展为 $32 \times 3 = 96$ （对应尺度集合 $\{1, 2, 4\}$ ）。

4.3.2 基于噪声增强的一致性正则化训练

在本文的整体框架中，噪声增强与一致性正则化训练被集成在模型的输入数据增强与编码器特征约束阶段。结合第4.2节中对现场采集信号干扰类型及影响的分析，本文在训练阶段通过构造扰动样本来模拟风噪、环境噪声及局部频带缺失等非目标干扰，并通过一致性正则化约束模型学习对上述扰动不敏感的稳定表示。同时，结合上一小节提出的 MSEAM，利用多尺度频带加权突出与螺栓松动相关的关键能量模式，从而使模型既能够适应复杂场景下的噪声扰动，又能够保持对目标状态特征的判别敏感性。

1. 噪声增强与遮挡扰动

本文在训练阶段对输入声谱图施加组合扰动 $\mathcal{T}(\cdot)$ ，以模拟风噪、环境噪声、局部频带缺失以及场景变化引起的非目标扰动，并降低模型对这些因素的敏感性。具体而言，采用两类增强方式：（1）按随机信噪比（SNR）添加高斯噪声，用于模拟现场采集中的随机背景扰动；（2）采用 SpecAugment 在时间维与频率维上进行随机遮挡，用于模拟局部频带缺失、突发噪声覆盖以及部分响应失真的情况。对样本 $\mathbf{X}$ 构造扰动版本

$$\tilde{\mathbf{X}} = \mathcal{T}(\mathbf{X}). \quad (4.16)$$

其中，SNR 在 $[\text{SNR}_{\min}, \text{SNR}_{\max}]$ 范围内随机采样，遮挡宽度与遮挡次数按实验设置确定。

2. 一致性正则化目标函数

在监督分类任务下，传统训练仅最小化交叉熵损失 $\mathcal{L}_{ce}$ ，本文在此基础上增加一致性正则项：要求同一样本的“干净版本”和“扰动版本”在特征空间中保持接近。设编码器输出为 $\mathbf{z} = f_{\theta}(\mathbf{X})$ 与 $\tilde{\mathbf{z}} = f_{\theta}(\tilde{\mathbf{X}})$ ，为消除特征幅值尺度的影响，先进行 ℓ_2 归一化：

$$\hat{\mathbf{z}} = \frac{\mathbf{z}}{\|\mathbf{z}\|_2}, \quad \hat{\tilde{\mathbf{z}}} = \frac{\tilde{\mathbf{z}}}{\|\tilde{\mathbf{z}}\|_2}. \quad (4.17)$$

一致性损失采用均方误差：

$$\mathcal{L}_{con} = \|\hat{\mathbf{z}} - \hat{\tilde{\mathbf{z}}}\|_2^2. \quad (4.18)$$

同时，为增强扰动样本的判别学习能力，本文对 $\mathbf{X}$ 与 $\tilde{\mathbf{X}}$ 均施加监督分类损失：

$$\mathcal{L}_{ce}(\mathbf{X}, y) = - \sum_{k=1}^K \mathbb{I}(y = k) \log \hat{y}_k, \quad \mathcal{L}_{ce}(\tilde{\mathbf{X}}, y) = - \sum_{k=1}^K \mathbb{I}(y = k) \log \tilde{\hat{y}}_k. \quad (4.19)$$

其中, $\hat{y}_k$ 表示干净样本 $\mathbf{X}$ 经模型预测的第 k 类概率输出, $\tilde{\hat{y}}_k$ 表示扰动样本 $\tilde{\mathbf{X}}$ 经模型预测的第 k 类概率输出。 $\mathbb{I}(\cdot)$ 为指示函数, 满足

$$\mathbb{I}(y = k) = \begin{cases} 1, & y = k \\ 0, & \text{其他} \end{cases}$$

最终的训练目标为

$$\mathcal{L} = \mathcal{L}_{ce}(\mathbf{X}, y) + \mathcal{L}_{ce}(\tilde{\mathbf{X}}, y) + \lambda \mathcal{L}_{con}, \quad (4.20)$$

其中 λ 为一致性约束权重。当 $\lambda = 0$ 时, 模型退化为常规监督训练; 当 $\lambda > 0$ 时, 模型在保证分类判别性的同时, 进一步学习对噪声与遮挡扰动不敏感的稳定表示。

3. 训练流程描述

综上, 如算法 4.1 所示, 本文的训练流程可概括为: 对每个批次输入 $\mathbf{X}$ 生成扰动版本 $\tilde{\mathbf{X}}$, 分别前向得到 $\mathbf{z}, \tilde{\mathbf{z}}$ 与预测 $\hat{\mathbf{y}}, \tilde{\hat{\mathbf{y}}}$, 计算式(4.20)并反向更新参数。该策略结合强数据增强和表征一致性约束, 能够有效缓解实际工程条件下的分布偏移与噪声干扰, 提高模型在多场景数据集上的泛化性能。

算法 4.1 基于噪声增强的一致性正则化训练流程

Input: 训练数据 $D = \{(\mathbf{X}, y)\}$; 编码器 f_θ ; 分类器 g_ϕ ; 增强算子 $\mathcal{T}(\cdot)$; 一致性权重 λ

Output: 训练后的参数 (θ, ϕ)

```

1 for epoch = 1, ..., E do
2     foreach  $(\mathbf{X}, y) \sim D$  do
3          $\tilde{\mathbf{X}} \leftarrow \mathcal{T}(\mathbf{X});$                                      // 噪声 + 遮挡增强
4          $\mathbf{z} \leftarrow f_\theta(\mathbf{X}), \tilde{\mathbf{z}} \leftarrow f_\theta(\tilde{\mathbf{X}});$ 
5          $\hat{\mathbf{y}} \leftarrow g_\phi(\mathbf{z}), \tilde{\hat{\mathbf{y}}} \leftarrow g_\phi(\tilde{\mathbf{z}});$ 
6          $\mathcal{L} \leftarrow \mathcal{L}_{ce}(\hat{\mathbf{y}}, y) + \mathcal{L}_{ce}(\tilde{\hat{\mathbf{y}}}, y) + \lambda \|\text{norm}(\mathbf{z}) - \text{norm}(\tilde{\mathbf{z}})\|_2^2;$ 
7         更新  $(\theta, \phi)$  以最小化  $\mathcal{L}$ ;
8     end
9 end
    
```

通过这种一致性正则化训练与噪声增强相结合的方式, 本文进一步提升了网络的泛化能力和鲁棒性, 尤其在存在噪声和遮挡的情况下, 能够有效减少模型对非本质扰动的依赖, 确保对实际故障特征的准确识别。

4.4 实验设计与结果分析

4.4.1 实验数据集与预处理

本文选取的实验数据来源于第三章所述采集场景中的张洼实训基地与宏源工业园区, 共包含 **3 座铁塔** 的有标签声纹数据集, 三个数据集的规模、类别设置与标注信息汇总如表 4.3, 4.4, 4.5 所示。

表 4.3 500 kV 干字塔数据采集信息

工况及位置信息	样本数量/个	离地垂直距离/m
C 主材 1 号连接处松动 2 颗	600	8.11
C 主材 2 号连接处松动 2 颗	560	16.48
C 主材 3 号连接处松动 2 颗	430	18.00
C 主材 4 号连接处松动 2 颗	425	20.24
C 主材 5 号连接处松动 2 颗	542	28.08
C 主材 6 号连接处松动 2 颗	590	30.00
D 主材 1 号连接处松动 2 颗	1477	8.11
D 主材 2 号连接处松动 2 颗	875	16.48
D 主材 3 号连接处松动 2 颗	850	18.00
D 主材 4 号连接处松动 2 颗	1010	20.24
中部斜材松动 2 颗	605	10.19
所有螺栓紧固	650	— [*]

^{*} “—” 表示该工况不对应固定的离地垂直距离（或不适用）。

为增强结构响应并提高声纹特征的可辨识性，基于第 3.4.1 小节提出的共振频率确定方法，本文分别选取 450 Hz、400 Hz、400 Hz 作为 220 kV 酒杯塔、500 kV 干字塔与 500 kV 酒杯塔的数据采集激励频率。为保证不同样本之间的可比性，三座铁塔均采用“一发一收”的采集方式，并统一采用 AC 对角布设模式安装声激励装置与骨传导传感器。具体而言，以声激励装置安装所在主材为起始点记为 A 号主材，沿塔体基座逆时针方向依次标记其余三根主材为 B、C、D 号主材。对角布设条件下，声激励装置与传感器分别固定于 A 号与 C 号主材位置，以实现塔体整体振动的有效激励与代表性响应采集。

表 4.4 500 kV 酒杯塔数据采集信息

工况及位置信息	样本数量/个	离地垂直距离/m
C 主材 1 号连接处松动 2 颗	745	15.48
C 主材 2 号连接处松动 2 颗	700	17.65
D 主材 1 号连接处松动 2 颗	910	15.48
D 主材 2 号连接处松动 2 颗	905	17.65
所有螺栓紧固	985	— [*]

^{*} “—” 表示该工况不对应固定的离地垂直距离（或不适用）。

在任务设置方面，张洼实训基地 220 kV 酒杯塔数据集重点用于评估模型对不同区域（如底部/中部/顶部）及不同部件（如主材连接点、横担连接处、接地/防盗螺栓等）松动状态的识别能力；而宏源工业园区 500 kV 干字塔与 500 kV 酒杯塔数据集则侧重于验证模型对具体空间位置螺栓松动的定位能力。考虑到现场传感器可安装位置及结构对称性，宏源数据采集阶段重点覆盖 C 号与 D 号主材区域及其相邻连接部位，以兼顾定位分辨率与采集效率。

针对宏源工业区两座 500 kV 铁塔，在采集点位标注方面，如图 4.3 所示，本文结合铁塔运维人员经验，重点关注主材连接节点与横担连接处等螺栓松动高风

表 4.5 220 kV 酒杯塔数据采集信息

工况及位置信息	样本数量/个
底部接地螺栓松动 2 颗	147
底部防盗螺栓松动 2 颗	210
杆件松动	616
中部螺栓松动 2 颗	154
顶部螺栓松动 2 颗	910
所有螺栓紧固	1050

险区域。对每个关注位置采用“自下而上”的编号策略进行记录，并同步标注其在结构中的相对高度信息，以便在后续实验中开展位置级别的分类评估。每段原始音频的采集时长为 10 s。为提升样本数量并增强模型训练的有效数据规模，本文将每段音频进一步裁剪为若干段长度为 2 s 的短时片段作为训练样本。螺栓松动状态的判定依据为扭矩阈值，所有松动状态下的螺栓扭矩小于 30 N·m。

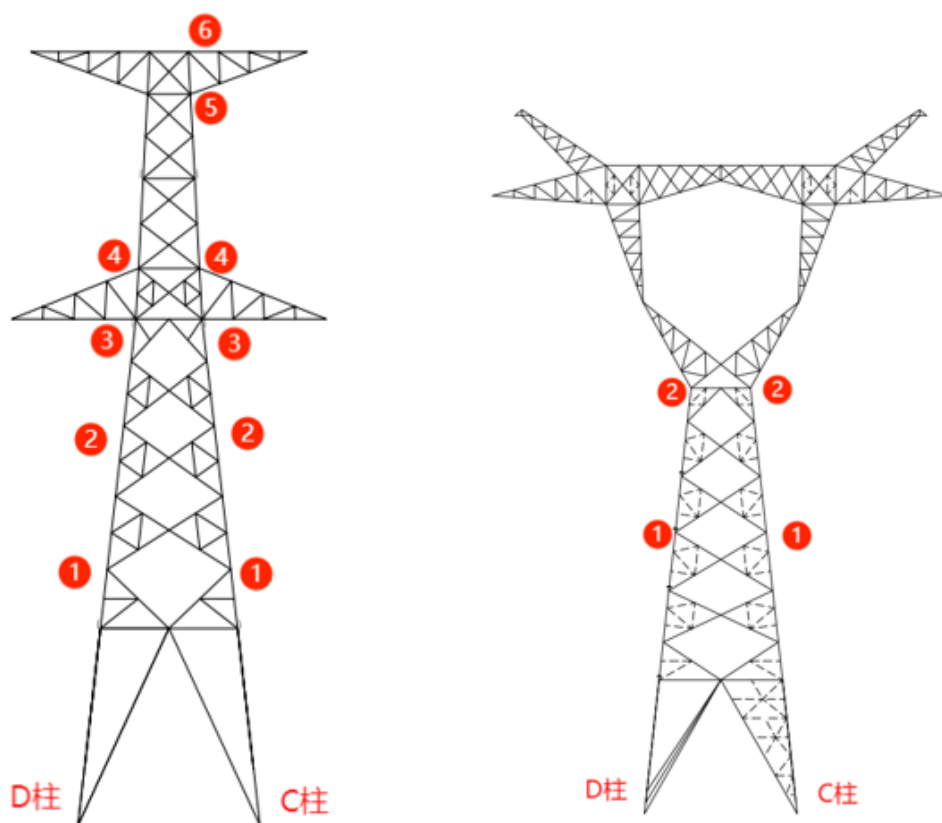


图 4.3 500 kV 干字塔和酒杯塔松动位置采集点示意图

4.4.2 实验环境设置与评价指标

(1) 实验环境是实现细节

本文所有实验均在配备 NVIDIA GeForce RTX 2080 Ti (12 GB 显存) 的计算平台上完成。软件环境方面，实验采用 Python 3.10 开发，实现基于 PyTorch 深度

学习框架；GPU 加速环境为 CUDA 12.2。为保证结果的可复现性，训练过程中固定数据划分策略，并在不同随机种子条件下进行重复实验。

(2) 数据集划分与重复实验策略

为评估模型在已知场景下的识别性能与泛化能力，本文将每个数据集按照 70%/15%/15% 划分为训练集、验证集与测试集。其中，训练集用于模型参数学习，验证集用于超参数选择与早停策略（early stopping）控制，测试集仅用于最终性能报告。为降低单次划分或随机初始化带来的偶然性影响，本文对每个模型在 5 个不同随机种子下独立训练与测试，并将 5 次结果取平均值作为最终实验性能。

(3) 评价指标

本文采用分类准确率（Accuracy）与 F1 分数（F1-score）作为主要评价指标，以同时反映模型整体判别能力与对类别不均衡情形下的综合表现。

设测试集共有 N 个样本，模型预测正确的样本数为 N_c ，则分类准确率定义为

$$\text{Accuracy} = \frac{N_c}{N}. \quad (4.21)$$

对于第 k 类，记真正例（True Positive）、假正例（False Positive）、假反例（False Negative）分别为 TP_k, FP_k, FN_k ，则该类的精确率（Precision）与召回率（Recall）定义为

$$\text{Precision}_k = \frac{TP_k}{TP_k + FP_k}, \quad \text{Recall}_k = \frac{TP_k}{TP_k + FN_k}. \quad (4.22)$$

F1 分数为精确率与召回率的调和平均：

$$\text{F1}_k = \frac{2 \text{Precision}_k \text{Recall}_k}{\text{Precision}_k + \text{Recall}_k}. \quad (4.23)$$

在多分类任务中，本文报告宏平均 F1（Macro-F1），即对各类别 F1 分数取算术平均：

$$\text{Macro-F1} = \frac{1}{K} \sum_{k=1}^K \text{F1}_k, \quad (4.24)$$

其中 K 为类别数。相较于仅使用 Accuracy，Macro-F1 能够在类别样本量不均衡时更客观地反映模型对各类别的综合识别能力。

4.4.3 实验结果分析

为验证本文所提出的模型在输电铁塔螺栓松动识别任务中的有效性与工程适用性，本节围绕该任务的实际特点设计对比实验与分析框架。输电铁塔长期运行于野外环境，受风噪、环境噪声及测量链路波动等因素影响显著；同时，螺

栓松动的带标签样本获取成本高、标注难度大，导致数据规模相对有限。在此条件下，过于复杂的模型容易产生过拟合并降低跨场景泛化能力。因此，本文在保证识别性能的同时，强调模型结构的轻量化与鲁棒性。在特征提取器层面，本文选取三种参数规模较为适中、训练稳定的图像识别骨干网络作为对比，包括 VGG16、ResNet18 与 MobileNetV2，以评估通用 CNN 结构在声谱图输入下的基线性能。在方法层面，进一步引入三种主流声纹识别框架作为对照实验：CRNN、X-vector TDNN 与 ECAPA-TDNN，用以检验本文方法相对于典型声纹建模范式的优势与差异。此外，为突出 MSEAM 模块在诊断中的关键作用，本文构建了多种注意力机制的结构级别对比实验，涵盖空间注意力、通道-空间联合注意力及其多尺度扩展形式，并与本文提出的多尺度能量驱动频带注意力进行对比，以分析不同注意力范式对声谱图判别特征的影响。与此同时，本文进一步统计并比较各模型的参数量与复杂度，并通过消融实验验证一致性正则化与 MSEAM 模块的有效性与贡献。为增强模型可解释性，本文基于频带响应机理对特征进行可视化与对比分析，通过比较异常（松动）样本与正常（紧固）样本在关键频带上的能量差异，验证频带特征在识别过程中的关键作用。最后，为使模型更贴近工程应用场景，本文在更接近实际运维需求的设置下开展扩展实验：一方面评估模型对单一松动与复合松动混合工况的识别能力，另一方面结合不同扭矩水平样本，探索并确定扭矩松动判据阈值，为后续现场快速判别提供参考依据。

1. 模型性能分析

表 4.6 给出了本文提出的 MSEAM-Net 与多种经典网络在三个数据集上的分类性能对比结果，其中分类准确率以五次独立实验的均值 $\pm$ 标准差表示，F1 分数为五次实验的均值。总体而言，MSEAM-Net 在三组数据集上均取得了较高的准确率与 F1 分数，且在复杂场景数据集上表现出更好的稳定性与鲁棒性。

表 4.6 不同方法在三个数据集上的性能对比（分类准确率：mean $\pm$ std，F1 分数：mean）。

方法	500 kV 干字塔		500 kV 酒杯塔		220 kV 酒杯塔	
	分类准确率 (%)	F1 分数 (%)	分类准确率 (%)	F1 分数 (%)	分类准确率 (%)	F1 分数 (%)
MSEAM-Net	96.56 $\pm$ 0.29	96.80	94.36 $\pm$ 0.54	95.30	99.92 $\pm$ 0.11	99.61
ResNet18	94.36 $\pm$ 1.34	95.20	89.92 $\pm$ 1.22	88.40	99.88 $\pm$ 0.18	99.60
VGG16	96.30 $\pm$ 0.65	96.40	92.88 $\pm$ 1.94	92.90	99.88 $\pm$ 0.18	99.60
CRNN	91.37 $\pm$ 1.04	91.40	83.27 $\pm$ 3.70	82.80	99.53 $\pm$ 0.60	99.40
MobileNetV2	94.15 $\pm$ 0.28	94.20	92.40 $\pm$ 1.36	92.00	99.66 $\pm$ 0.24	99.48
Xvector_TDNN	95.17 $\pm$ 1.28	95.20	93.05 $\pm$ 1.54	92.80	99.91 $\pm$ 0.19	99.80
ECAPA_TDNN	95.33 $\pm$ 0.69	95.20	92.46 $\pm$ 1.43	92.40	99.96 $\pm$ 0.09	99.90

在 500 kV 干字塔数据集上，MSEAM-Net 取得 96.56 $\pm$ 0.29% 的分类准确率与 96.80% 的 F1 分数，较 ResNet18 的 94.36 $\pm$ 1.34% 在准确率上提升约 2.20 个百分点，同时标准差显著更小，表明本文方法不仅精度更高，而且对随机初始化与训练扰动更不敏感。与 VGG16、MobileNetV2 等图像识别骨干相比，MSEAM-Net 通过多尺度能量驱动的频带注意力弱化了对纹理式谱图细节的依赖，更聚焦于

螺栓松动引起的关键频带能量异常，从而在噪声与链路差异存在时仍能保持稳定收益。与 X-vector TDNN、ECAPA-TDNN 等主流声纹识别模型相比，本文方法无需强时序聚合即可有效刻画稳态激励下的能量分布特征，在该数据集上性能处于最优梯队。在 500 kV 酒杯塔数据集上，MSEAM-Net 取得 $94.36 \pm 0.54\%$ 的分类准确率与 95.30% 的 F1 分数，较 ResNet18 的 $89.92 \pm 1.22\%$ 提升约 4.44 个百分点，优势更为显著；相较 CRNN ($83.27 \pm 3.70\%$) 提升幅度更大。该结果说明在数据分布更复杂、频带漂移更明显的条件下，传统仅依赖局部卷积纹理或序列建模的方法更易受噪声影响，而本文将频带注意力与一致性正则化训练结合，可学习到更稳定的特征表示，从而获得更可靠的泛化性能。在 220 kV 酒杯塔数据集上，各方法的分类准确率与 F1 分数整体接近饱和，不同模型之间差异较小。MSEAM-Net 达到 $99.92 \pm 0.11\%$ 的分类准确率与 99.61% 的 F1 分数，说明在任务难度较低、样本可分性较强时，本文方法仍能保持接近最优的识别能力；同时也表明此时结构改进带来的边际收益会被压缩，进一步提升更依赖于更细粒度标签、更强干扰条件或更严格的跨塔/跨场景划分。综上，MSEAM-Net 在不同数据集上均表现出较强泛化性，尤其在工况复杂的 500 kV 系列数据集中优势更为明显，验证了多尺度能量注意力和一致性正则化训练在提升鲁棒性方面的有效性。后续小节将从模型复杂度与消融实验角度进一步分析各模块对性能提升的贡献，并讨论不同频带特征的重要性与诊断可解释性。

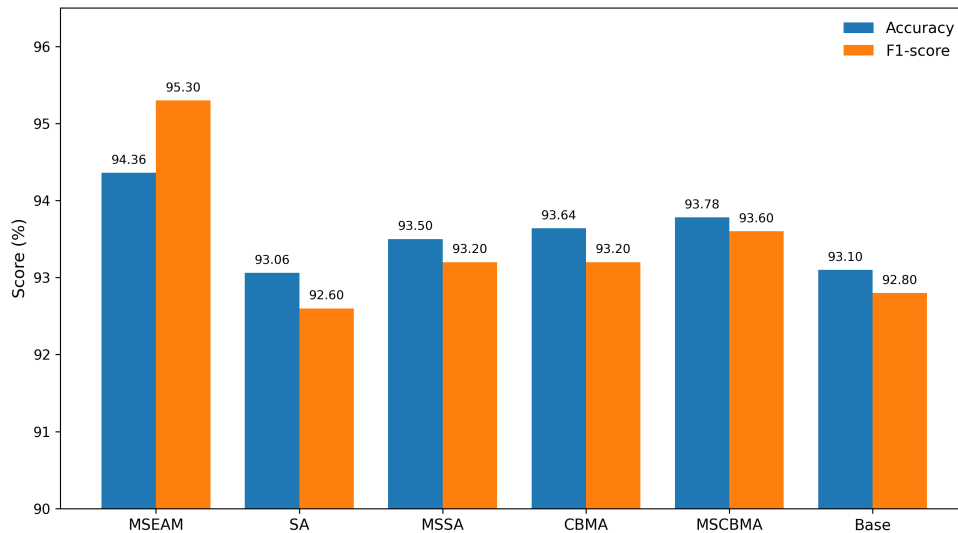


图 4.4 不同注意力机制在 500 kV 酒杯塔数据集的分类结果

为进一步验证所提出多尺度能量注意力机制的有效性，本文在 500 kV 酒杯塔数据集上，将 MSEAM-Net 与多种典型注意力结构进行结构级别对比，包括：仅引入空间注意力（SA）、多尺度空间注意力（MSSA）、通道-空间联合注意力（CBAM）以及多尺度融合的通道-空间注意力（MSCBAM），并以不含注意力的 Base 模型作为基线。各方法的分类准确率与 F1 分数结果如图 4.4 所示。

从结果上看,所有注意力机制均优于 Base,说明在声谱图建模中引入显著性加权能够抑制冗余背景并提升类别可分性。Base 模型在该数据集上分类准确率为 93.10%, F1 分数为 92.80%;引入 SA 后分类准确率提升至 93.06%,但 F1 分数仍为 92.60%,增益有限。这表明仅依赖空间显著性可能更偏向局部纹理或块状结构,对松动相关的频带能量模式捕捉不足,且对类别间不平衡或难类的改善不明显。相比之下, MSSA、CBAM 与 MSCBAM 均带来更稳定的提升: MSSA 达到 93.50%/93.20%, CBAM 达到 93.64%/93.20%, MSCBAM 达到 93.78%/93.60% (分类准确率/F1 分数)。这说明加入通道维建模以及多尺度上下文融合能够更好地增强声谱图的判别信息。然而,这些机制本质仍是将声谱图作为二维图像处理,其注意力权重主要由卷积特征的统计量驱动,更擅长突出纹理与形状模式,而对本文任务中更关键的窄带峰值与宽带抬升的能量异常缺少显式建模。

与上述机制相比, MSEAM 在分类准确率与 F1 分数上均取得最优,相对 Base 分别提升 1.26 与 2.50 个百分点;相对次优的 MSCBAM 也分别提升 0.58 与 1.70 个百分点。MSEAM 并非仅做通道或者空间显著性加权,而是通过多尺度池化获得不同感受野下的频谱表征,并在频带维计算能量强度以形成能量驱动的频带-通道联合权重。该设计更贴合声发激励稳态响应的物理机理,能够在外届噪声干扰存在时,仍稳定强调与松动相关的关键频带能量变化,从而明显地提升对各类更均衡的识别能力。

综上,空间注意力对本任务的提升有限,多尺度与通道-空间联合机制可带来一定增益,但仍受图像式注意力的模式限制,而本文提出的 MSEAM 通过多尺度和能量驱动的频带注意力更直接刻画松动诱发的谱能量异常,因此在结构级别对比中表现最佳。

2. 模型复杂度分析与消融实验

(1) 模型复杂度分析

为验证所提 MSEAM-Net 在提升故障识别性能的同时,不会引入过高的工程部署开销,本文从模型规模、理论计算复杂度及实际推理效率三个维度,对模型的轻量化特性与落地适用性展开全面分析。选取总参数量 (Params)、乘加运算量 (MACs)、单样本推理时延 (Latency) 以及推理吞吐量 (Throughput) 作为模型复杂度评价指标。其中,总参数量用以表征模型的存储占用开销,乘加运算量用于刻画模型的理论计算复杂度,单样本推理时延与推理吞吐量则用于衡量统一软硬件环境下,模型的实际推理速度与工程部署效率。

卷积层的参数量和计算量可近似写为

$$\text{Params}_{\text{conv}} = k^2 C_{\text{in}} C_{\text{out}}, \quad \text{MACs}_{\text{conv}} \approx HW \cdot k^2 C_{\text{in}} C_{\text{out}}, \quad (4.25)$$

其中, C_{in} 和 C_{out} 分别表示输入通道数和输出通道数, $k \times k$ 表示卷积核大小,

$H \times W$ 表示输出特征图的空间尺寸。

MSEAM 模块带来的新增开销主要来自注意力分支。多尺度池化和上采样不会引入额外的可学习参数。设输入通道数为 C ，尺度数为 m ，拼接后的通道数为 $C' = mC$ ，压缩维度为 r ，则注意力分支新增参数量可近似写为

$$\text{Params}_{\text{attn}} \approx C'r + rC' = 2C'r = 2mCr. \quad (4.26)$$

这个式子说明，注意力分支的参数量与通道数 C 近似成线性关系，当 r 明显小于 C' 时，模块新增参数量不会过大。本文中的一致性正则化只作用于训练阶段，它不会增加推理阶段的参数量和计算量。

表 4.7 不同模型复杂度与推理效率对比

模型	总参数量 / M	MACs / G	推理时间 / ms	推理吞吐量 / (次/s)
MSEAM-Net	1.27	0.2602	5.95	168.0
ResNet18	11.30	0.1420	4.10	243.7
MobileNetV2	2.55	0.0247	7.30	136.9
CRNN	0.36	0.0392	1.39	720.1
Base	1.22	0.2160	1.78	560.8
VGG16	15.25	1.2497	3.22	310.4
Xvector_TDNN	1.01	0.0434	1.80	555.4
ECAPA_TDNN	1.06	0.0518	10.43	95.9

注：各模型分类器参数量相同，均为 1797。因此，表中的总参数量差异主要来自编码器部分。

表 4.7 给出了 MSEAM-Net 与 ResNet18、VGG16、MobileNetV2、CRNN、Base、Xvector_TDNN 以及 ECAPA_TDNN 的复杂度与推理效率对比结果。可以看出，MSEAM-Net 的总参数量为 1.27 M，显著低于 ResNet18 的 11.30 M 与 VGG16 的 15.25 M，也低于 MobileNetV2 的 2.55 M，说明所提出模型在模型规模上具有较好的紧凑性。与此同时，MSEAM-Net 的理论计算量为 0.2602 G MACs，虽高于 CRNN、Xvector_TDNN、ECAPA_TDNN 和 MobileNetV2 等轻量化结构，但远低于 VGG16 的 1.2497 G，表明本文方法是在可接受的计算开销下提升特征表达能力，而非单纯依赖大规模参数堆叠。

从实际部署效率看，MSEAM-Net 的单样本推理时间为 5.95 ms，吞吐量为 168.0 次/s，推理速度快于 MobileNetV2 的 7.30 ms 和 ECAPA_TDNN 的 10.43 ms，满足实时推理需求。相比基础网络 Base，MSEAM-Net 的参数量仅由 1.22 M 增加至 1.27 M，增幅约为 3.87%，说明所引入的多尺度增强与注意力机制并未显著扩大模型规模，其 MACs 由 0.2160 G 增加至 0.2602 G，增幅约为 20.51%，反映出更强特征建模能力是以适度增加理论计算量为代价实现的。综合来看，MSEAM-Net 在保持较小模型规模的同时，引入了对关键频带能量分布的显式建模，取得了较好的复杂度与效率折中。结合前文识别性能结果，说明本文模型能够在较低参数规模和可接受部署开销下获得更优的识别表现与更强的跨场景鲁棒性。

综合来看，MSEAM-Net 并未追求极致的轻量化水平，也未依托大规模参数

堆叠的方式提升模型性能。实验结果表明,该模型始终保持 1M 级的参数规模,并在此基础上引入了多尺度特征提取与频带-通道联合注意力建模机制。该设计在将模型规模维持在较小区间的同时,也带来了一定的理论计算量上升与推理时延增加。由此可见,本文所提 MSEAM-Net 的核心设计理念并非追求极致轻量化,而是在较小的参数量级下,融入更精细、更具针对性的特征建模能力,这与前文的模型结构设计思路完全契合。

(2) 消融实验设置

为进一步量化 MSEAM 模块与一致性正则化 (CR) 对性能提升的贡献,本文构建如下对比模型:

- **Model1:** 仅使用残差主干网络进行特征提取与分类;
- **Model2:** 移除 MSEAM 模块,其余结构与训练策略保持一致;
- **Model3:** 保留 MSEAM 模块,但移除一致性正则项 (令 $\lambda = 0$),退化为常规监督训练;
- **Model4:** 完整模型 (MSEAM + CR)。

(3) 消融实验结果与分析

表 4.8给出了不同消融版本在测试集上的平均分类准确率与宏平均 F1 (Macro-F1)。可以看出相比 Model1-Model3, MSEAM-Net 在准确率与 F1 分数上均取得最优结果。同时, Model2 和 Model3 相较于完整模型均存在性能下降,说明多尺度能量驱动的频带注意力能够有效增强与螺栓松动相关的关键频带信息并抑制背景扰动,同时噪声增强的一致性约束有助于学习对噪声与遮挡扰动更稳定的特征表征,从而提升模型鲁棒性与泛化能力。总体而言, MSEAM 与 CR 两者在模型中起到互补作用,前者提升判别性,后者提升鲁棒性,二者结合带来更稳定的性能增益。

表 4.8 对 MSEAM 和一致性正则化 (CR) 的消融研究

Model	MSEAM	CR	Accuracy (%)	Macro-F1 (%)
Model1 (w/ MSEAM and CR)	✗	✗	92.38	93.20
Model2 (w/ MSEAM)	✓	✗	93.10	92.80
Model3 (w/ CR)	✗	✓	93.18	92.60
Model4 (MSEAM-Net, ours)	✓	✓	94.36	95.20

(4) 频带特征重要性分析

为验证 MSEAM 中多尺度能量加权对螺栓松动识别的贡献,并进一步明确模型依赖的关键判别信息,本文从输入表征层面对频带能量模式进行干预实验。如图 4.5所示,本文在原始 log-Mel 谱图 S 上引入滑动窗平滑以估计稳态背景成分 $\text{Smooth}(S)$,并构造残差谱图 (residual spectrogram):

$$S_{\text{res}} = S - \text{Smooth}(S), \quad (4.27)$$

该操作等效于抑制稳态高能谱峰并突出局部波动成分。随后保持网络结构与训练策略不变，仅将输入由原始谱图 $\mathbf{S}$ 替换为残差谱图 $\mathbf{S}_{\text{res}}$ 重新训练模型，以评估稳态高能频带对跨塔判别的影响。

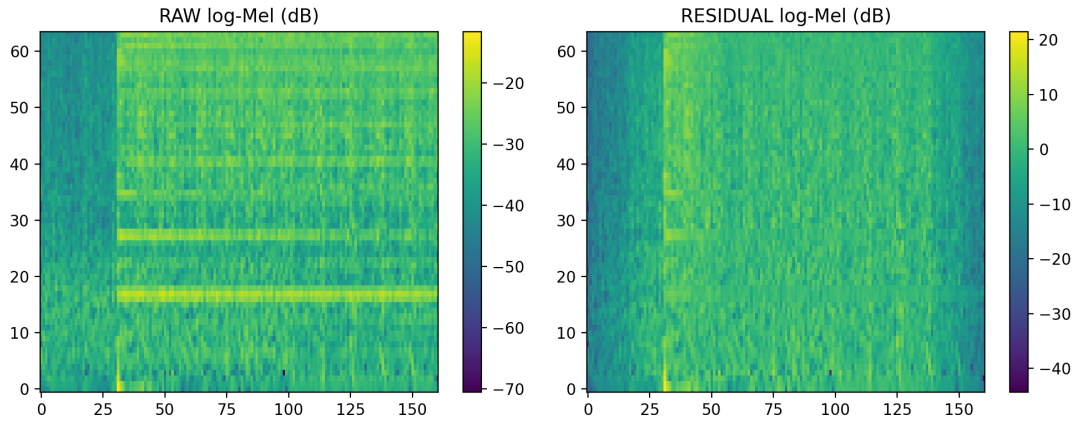


图 4.5 原始 log-Mel 谱图与残差谱图

实验结果表明，抑制稳态谱峰会显著导致模型性能下降。以 MSEAM-Net 为例，在 500 kV 干字塔与 500 kV 酒杯塔数据集上，使用原始谱图 (Raw) 输入时分类准确率分别为 96.56% 与 94.36%；当输入替换为残差谱图 (Residual) 后，性能下降至 500 kV 干字塔为 90.50%，500 kV 酒杯塔为 81.70%，并且相应的 F1 分数也分别下降至 91.00% 和 81.00%。上述结果从实证角度说明，稳态高能频带及其对应的能量分布是螺栓松动跨铁塔结构泛化识别的重要依据；当该部分信息被削弱后，模型难以形成稳定的判别表征，从而产生显著性能退化。

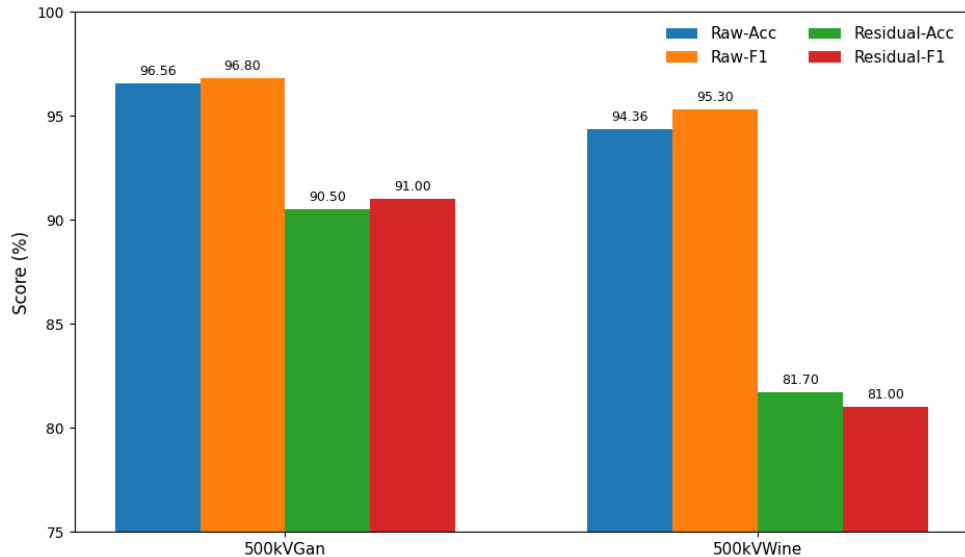


图 4.6 原始谱图与残差谱图输入下的性能对比 (分类准确率/F1 分数)

从机理层面看，螺栓松动会改变连接刚度与接触条件，进而引起结构动力学响应发生变化。Wang 等^[74]指出，螺栓松动将改变连接处动力学并导致主导模态响应发生频移，这类共振频移与能量重分配通常会表现为特定声学频带的能量

增强、谱峰位置变化或能量弥散。因此，在声发激励的稳态段内，原始谱图中呈现的高能、稳定频带往往对应结构主导响应，是松动诊断中最具信息量的频域线索。残差谱图抑制该部分稳态能量后，等效于削弱了与结构模态紧密相关的“物理可解释成分”，从而导致跨塔泛化能力明显下降。

进一步地，该现象也解释了传统注意力模块在本任务中提升有限的原因。SA、CBAM 等方法通常依赖全局平均池化获取通道描述，容易压平频带内的能量差异，降低对窄带谱峰与频移现象的敏感性。相比之下，MSEAM 通过多尺度自适应池化构建不同感受野下的谱表征，并基于能量强度形成频带-通道注意力矩阵，能够同时实现窄带峰值放大与宽带能量抬升，对不同塔型与不同共振带宽下的能量模式变化更具鲁棒性。换言之，MSEAM 的改进并非仅来自图像式显著性加权”，而是来自对稳态能量模式这一关键物理特征的定向增强，从而提升识别性能并降低难例混淆。

综上，残差谱图输入导致的显著性能下降验证了稳态高能频带是螺栓松动识别的核心判别信息。MSEAM 对这些关键频带的针对性增强在性能提升与跨结构泛化中起到了关键作用，同时也提升了模型的可解释性。

3. 复杂工况与工程任务验证

为了进一步探究模型在真实复杂工况下的诊断能力，研究完善了宏源工业区干字型铁塔所采集的铁塔松动数据集，补充了多个不同位置的螺栓同时松动的相关数据。如图 4.7 所示，在进行螺栓多位置松动采集时，通过采集 C 柱 1 号与 2 号，4 号，5 号，以及 C 柱 5 号和 6 号的相关数据探究同一主材上不同位置松动时分布差异。同时，通过采集 C 柱 1 号和 D 柱 1 号以及 C 柱 2 号与 D 柱 1 号松动数据探究相邻主材同一高度和不同高度之间的差异。

利用 t-SNE 对完善后的数据集进行可视化，如图 4.8 所示，从 t-SNE 可视化图中可以看出不同工况下螺栓松动的数据在二维空间中的分布。图中显示了多个不同标签的点，这些点表示不同的螺栓松开工况。每个标签的点在 t-SNE 图上形成了明显的集群。这表明不同的螺栓松开工况在特征空间中有着明显的区别。同一个标签的点大多聚集在一起，表明同一工况下的数据具有相似的特征，标签为“斜材松动”的点在图中分布较为分散，可能表明斜材松动的特征多样性较大，不同数据样本之间的特征相似度较低。标签为“螺栓紧固”的点也形成了独立的集群，说明螺栓紧固状态下的特征与松动状态有明显区别。同一主材上不同松动位置呈现明显的集群，标签为“C 柱 5 号”和“C 柱 6 号”与其他“C 柱”系列的点（如“C 柱 2 号”和“C 柱 3 号”等）在 t-SNE 图中有一定的距离，这表明即使是在同一主材下，不同高度位置也会导致特征的差异。标签为“C15 松动”和“C14 松动”等复合松开工况的点在图中也形成了独立的集群，且与单一松开工况的点有明显的区别，表明复合松开工况的特征较为复杂，可能涉及多种

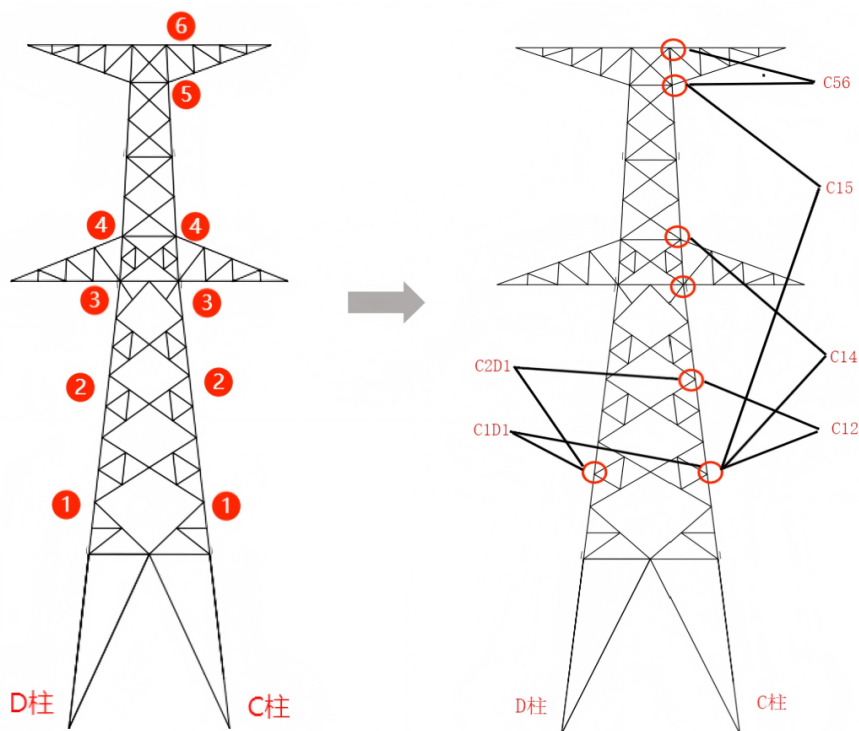


图 4.7 多螺栓松动位置采集示意图

松动模式的叠加效应，这与单一松动工况有显著差异。标签为“C1D1 松动”的点与“C2D1 松动”的点分布在较近的位置，表明若松动点位于两个相邻主材上的工况之间存在某些相似性。

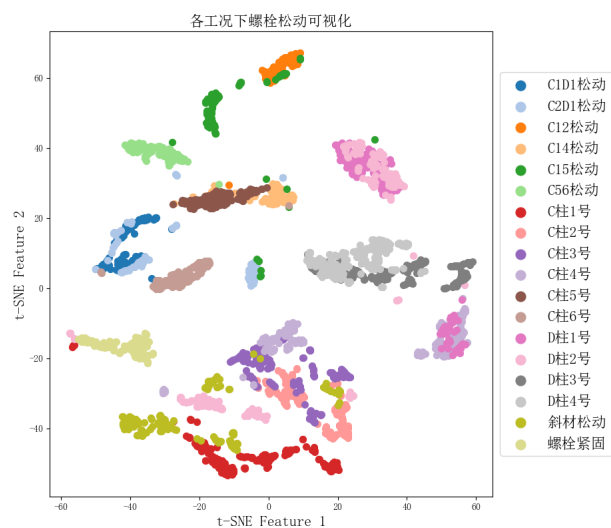


图 4.8 500 kV 干字塔各工况下螺栓松动可视化

将该干字型铁塔的 18 种工况数据输入到网络中进行诊断，其分类结果的混淆矩阵如图 4.9 所示，除了标签 6 (D 柱 1 号) 和 7 (D 柱 2 号)，8 (D 柱 3 号) 和 9 (D 柱 4 号) 产生混淆外，对于该场景中的其他位置的螺栓松动工况，模型均能保持较高的诊断精度，同时混淆矩阵的结果也进一步说明对于同一主材的相

邻位置螺栓松动进行诊断是一个挑战。

基于上述多位置螺栓松动实验结果可以看出，模型不仅能够区分紧固状态与松动状态，还能够在预设检测点位范围内对不同主材、不同高度位置及部分复合松动工况进行类别识别。从采集点位的空间分布来看，500 kV 干字塔中部分相邻检测点的离地高度间隔约为 1.5–2.2 m，模型对其中多数位置能够保持较好的分类效果，说明声纹特征中包含与松动位置相关的空间差异信息，能够在当前实验条件下将疑似松动风险缩小至约 2 m 量级的局部连接区段。需要指出的是，该结果仍属于预设点位条件下的区段级位置识别，而非任意螺栓坐标意义上的精确定位。尤其是混淆矩阵中 D 柱部分相邻位置仍存在一定混淆，表明当螺栓位于同一主材、相邻高度或传播路径相近区域时，其声纹响应差异较弱，模型的螺栓级定位能力仍存在提升空间。因此本文方法现阶段更适合作为输电铁塔螺栓松动巡检中的辅助初筛手段，用于判断预设检测区域是否存在松动风险，并为运维人员缩小人工复核范围提供参考。对于实际运维决策，仍需结合人工复检、扭矩扳手检测、历史巡检记录以及结构重要性等级等信息进行综合判断。

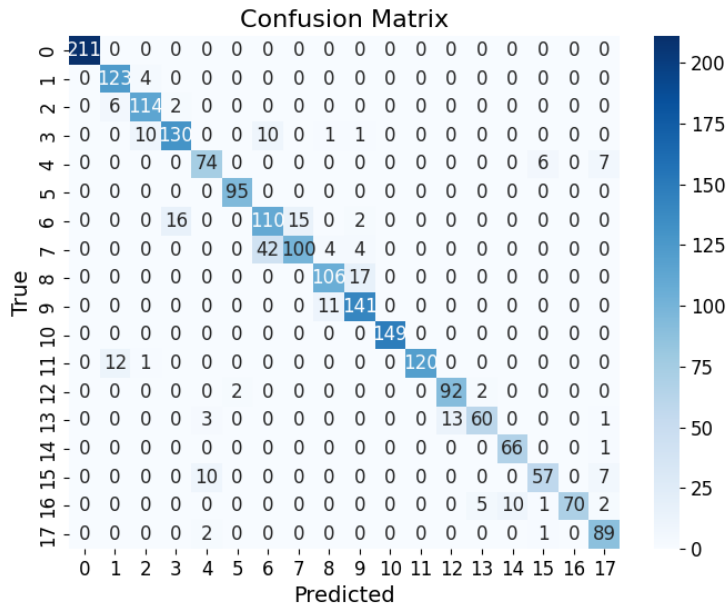


图 4.9 500 kV 干字塔不同工况分类混淆矩阵

4. 螺栓松动程度量化实验

上述 18 种工况标签主要用于验证模型对不同螺栓松动位置、相邻主材松动位置以及复合松动工况的识别能力，其标签含义侧重于“松动位置”和“松动组合”的区分，并未进一步量化不同松动程度之间的差异。为了回应工程检测中“螺栓松动到何种程度能够被算法有效识别”的问题，本文在完成多工况位置诊断实验的基础上，进一步设计了不同扭矩值下的补充实验，用于分析螺栓松动程度变化对声纹特征分布及模型分类结果的影响。该实验并非用于重新定义前述

18 类工况标签，而是作为对“螺栓松动”状态的量化说明和算法检测能力边界的补充验证。

在实验过程中，选取铁塔中特定位置的 M24 型号螺栓作为测试对象。首先由专业运维人员将该螺栓调整至松动状态，随后利用数显扭矩扳手对其进行不同扭矩值下的逐级紧固，并分别采集对应状态下的声纹数据。根据 GB 50233—2014《110 kV 750 kV 架空输电线路施工及验收规范》，M24 螺栓的标准紧固扭矩值为 250 N·m。基于该标准，本文设置 25 N·m、50 N·m、75 N·m、125 N·m、150 N·m 和 250 N·m 共 6 组扭矩状态，其中 250 N·m 作为正常紧固状态，其余扭矩值用于模拟不同程度的松动状态。通过该组实验，可以进一步分析螺栓由明显松动逐渐过渡到标准紧固过程中，声纹特征是否呈现可分辨的变化规律。

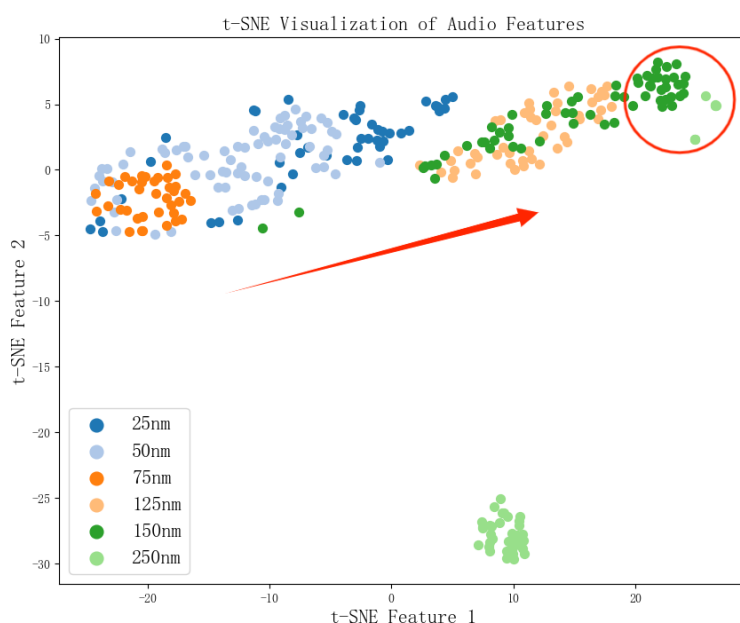


图 4.10 不同扭矩值下的螺栓松动可视化

如图4.10所示，通过对不同扭矩值下采集数据的特征进行 t-SNE 可视化可以看出，不同扭矩状态下的数据点在二维空间中呈现出一定的聚类差异。250 N·m 数据点形成了相对独立的聚类，与低扭矩状态下的数据分布存在明显区分，说明标准紧固状态下的声纹特征与松动状态具有较大差异。低扭矩状态，如 25 N·m、50 N·m 和 75 N·m，对应的数据点主要集中在相近区域，而较高扭矩状态，如 150 N·m 和 250 N·m，对应的数据点逐渐向另一侧区域过渡，表明随着扭矩值增大，螺栓连接状态趋于紧固，其声纹特征也发生连续变化。同时，部分相邻扭矩值之间存在重叠区域，例如 25 N·m、50 N·m 和 75 N·m 之间，以及 125 N·m 和 150 N·m 之间的数据分布存在一定交叠，说明当扭矩差异较小时，不同松动程度下的声纹特征可能具有相似性，进而增加模型分类难度。

进一步将不同扭矩值下的数据输入本文模型进行分类诊断，得到的混淆矩阵如图4.11所示。其中，类别1至类别6分别代表 25 N·m、50 N·m、75 N·m、125 N·m、150 N·m 和 250 N·m。由混淆矩阵可以看出，模型能够较好地地区分低扭矩松动状态与标准紧固状态，但在相邻扭矩等级之间仍存在一定混淆。例如，25 N·m 与 50 N·m、50 N·m 与 75 N·m、125 N·m 与 150 N·m 之间均出现了不同程度的误判。这说明当扭矩值间隔较小时，螺栓声纹特征的差异相对有限，模型对细粒度松动程度的区分能力会受到影响。相比之下，当扭矩差异增大时，不同状态之间的特征差异更加明显，分类结果也更加稳定。

结合 t-SNE 可视化结果和混淆矩阵可以发现，当扭矩值达到 150 N·m 时，部分样本在特征空间中已逐渐接近 250 N·m 紧固状态样本，说明该扭矩附近可能处于松动状态与紧固状态之间的过渡区域。由于本实验所采集的扭矩状态是离散设置的，且尚未覆盖 150 N·m 至 250 N·m 之间更细间隔的连续扭矩数据，因此本文不将该结果作为普适性的工程判据，而仅基于当前实验条件，将 150 (± 5) N·m 作为本文数据集中 M24 螺栓松动状态与紧固状态的参考分类阈值。在本文实验条件下，当 M24 螺栓扭矩低于或接近标准紧固扭矩 250 N·m 的 60% 时，模型能够较为有效地识别其与正常紧固状态之间的差异。该结果为前述螺栓松动标签提供了进一步的量化解释，也为后续开展更细粒度的松动程度评估提供了实验依据。

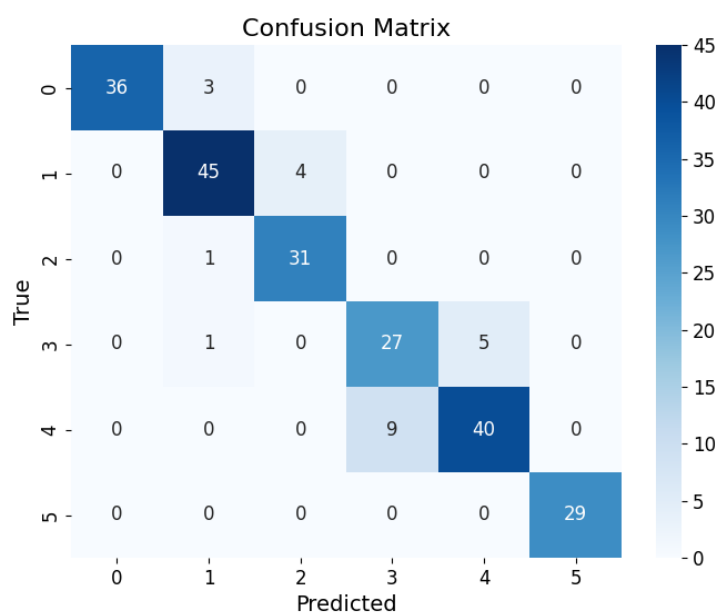


图 4.11 不同扭矩值分类的混淆矩阵

4.5 本章小结

本章首先分析了输电铁塔现场采集信号中风噪、环境噪声、设备链路频响差异与传播路径差异等干扰因素的来源、特征与影响，并在此基础上，针对松动状态下“窄带峰值 + 宽带抬升”的多尺度谱能量异常难以稳定表征的问题，提出端到端的多尺度能量注意力残差网络 MSEAM-Net。在结构层面引入多尺度能量注意力模块，通过多尺度池化融合不同感受野信息，并以频带 RMS 能量统计驱动频带—通道联合注意力，实现对关键频带的自适应增强与背景成分抑制。在训练层面结合随机 SNR 噪声与 SpecAugment 遮挡构造扰动样本，引入一致性正则化约束同一样本在干净和扰动条件下的表征一致，从而提升模型鲁棒性与跨场景泛化能力。基于三座铁塔多工况数据集的实验表明，MSEAM-Net 在 500 kV 系列复杂数据集上相较 ResNet18、VGG16 及典型声纹框架取得更优且更稳定的分类准确率与 Macro-F1，并在结构对比与消融实验中验证了 MSEAM 与一致性正则化的互补贡献，进一步的频带干预与工程扩展实验说明稳态高能频带及其能量分布是螺栓松动识别的重要物理线索，模型能够支持多位置复合松动与不同扭矩等级判别，并给出在垂直方向约 2m 范围的诊断精度和 M24 规格螺栓约 $150(\pm 5) \text{ N} \cdot \text{m}$ 的有效松动区分阈值，为后续工程部署提供了方法与实证依据。值得注意的是，本文实验数据仍主要来源于有限数量的塔型和采集地区，模型泛化验证重点集中在已有数据范围内的跨塔型、跨场景和多工况识别。对于季节温变导致的结构材料参数和连接界面状态变化、沿海高湿或盐雾环境导致的腐蚀影响、山区强风和导线舞动引起的持续振动扰动等复杂因素，当前实验尚未形成系统化、长期化的数据覆盖。

第5章 基于半监督迁移学习的少样本螺栓松动检测方法

针对输电铁塔螺栓松动检测任务，在实际工程应用和部署中仍然存在一些挑战。首先，不同塔型在结构模态、共振频带与传播路径上存在差异，使得在某一塔型上训练的模型难以直接泛化到另一塔型；其次，在真实的野外场景下获得大规模、精确标签的成本高、周期长，且存在一定的安全隐患，而无标签数据的获取则相对廉价便捷。因此，本章面向跨塔型少样本螺栓松动检测任务，围绕域对齐、伪标签自训练与结构复杂度自适应度量学习三条主线展开，通过引入域对抗对齐策略减小源域和目标域的分布差异，并针对标签稀缺问题提出双阈值伪标签策略以稳定引入未标注数据；最后提出结构复杂度自适应四元组损失，通过动态间隔与结构化采样增强近邻类别的可分性，从而提升跨域与少样本条件下的判别能力。

5.1 引言

输电铁塔螺栓松动会导致连接刚度下降与结构动力学特性改变，是影响线路安全的重要隐患。相较于传统接触式检测，基于声纹信号的松动诊断具有非接触、可在线监测与成本较低等优势。第4章通过三个数据集实际验证了模型的有效性，并且在标签充足的情况下训练出的模型具有较高的精度。然而，针对输电铁塔这一特定对象，获取标签数据存在诸多困难，例如输电铁塔所处的环境条件复杂多变，包括温度变化、风速、湿度等自然因素，这些因素都可能影响螺栓的松动状态，使得在动态环境下获取准确标签变得更加困难，同时工作人员需要在高空作业，对不同松动状态下的螺栓进行标记，这不仅增加了操作的难度，也提高了安全风险。因此，亟需一种更加鲁棒和减少训练成本的方法。

针对上述问题，本文在前一章提出的多尺度能量注意力编码器基础上，进一步构建半监督迁移学习（semi-supervised transfer learning）框架，具体而言，利用 MSEAM-Net 结构，提取稳定的时频能量表征，并结合所设计的半监督迁移学习训练策略及其损失，通过源域丰富标注与目标域少量标注和大量未标注数据，实现跨塔型条件下的少样本螺栓松动检测与定位。

本章的主要研究内容与组织结构如下：第5.2节面向跨塔型少样本螺栓松动检测任务，首先给出问题定义，在此基础上构建半监督迁移学习框架，并重点介绍结构复杂度自适应四元组损失及其联合优化策略；第5.3节介绍实验数据集、实现细节与对比设置，并从分类性能、混淆矩阵、消融实验、类间分离能力以及参数敏感性等方面对所提方法进行验证与分析；第5.4节对全章内容进行总结。

5.2 基于半监督迁移学习的螺栓松动检测方法设计

5.2.1 问题定义

本文考虑跨塔型少样本迁移场景：源域为具有充足标注样本的塔型数据集 $\mathcal{D}_s^{lab} = \{(x_s^i, y_s^i)\}_{i=1}^{n_s}$ ，目标域为仅含少量标注与大量未标注样本的塔型数据集 $\mathcal{D}_t^{lab} = \{(x_t^j, y_t^j)\}_{j=1}^{n_t}$ 、 $\mathcal{D}_t^{unlab} = \{x_t^k\}_{k=1}^{n_u}$ ，且满足 $|\mathcal{D}_t^{lab}| \ll |\mathcal{D}_s^{lab}|$ 。两域共享标签空间，即 $y \in \{0, 1, \dots, K-1\}$ 。训练数据整体可表示为

$$\mathcal{D} = \mathcal{D}_s^{lab} \cup (\mathcal{D}_t^{lab} \cup \mathcal{D}_t^{unlab}). \quad (5.1)$$

模型采用“特征提取器 + 分类器”的结构，记特征提取器为 $g(\cdot)$ ，分类器为 $h(\cdot)$ ，整体预测函数为 $f(x) = h(g(x))$ 。其中 $g(\cdot)$ 采用第4章所述的 MSEAM-Net 结构，本章重点在于半监督迁移学习训练策略及其损失设计。

5.2.2 结构复杂度自适应四元组损失设计

1. 动机与基本思想

在跨塔型少样本螺栓松动识别任务中，目标域类别往往呈现类内离散度不均衡与近邻类别高度相似的特征，例如相邻位置（同一主材相邻高度/相邻连接点）的声学响应受传播路径与结构模态影响，频带能量分布差异细微，导致类别边界模糊。传统三元组/四元组度量学习通常采用固定间隔，但在少样本且类内结构差异显著的情况下，固定间隔会出现两难，即间隔过小将导致复杂类别分离不足，间隔过大则可能对简单类别施加过强约束、降低泛化能力。为此，如图5.1所示，本文提出结构复杂度自适应四元组损失（Structure Complexity-Adaptive Quadruplet Loss, SQL），通过结构化四元组采样与复杂度驱动的双间隔自适应实现对不同类别差异度的动态匹配，从而提升少样本条件下的嵌入空间判别性。

2. 伪标签引导的结构化四元组采样（PL-SQ）

当目标域仅有少量标注样本时，未标注数据蕴含大量可用于优化类间边界的信息。本文设计 Pseudo-Label Structured Quadruplet (PL-SQ) 策略，将高置信伪标签样本引入四元组构造，以同时覆盖类内多样性与近邻类难样本。

(1) 双阈值伪标签生成

设特征提取器-分类器为 $f = h \circ g$ 。对每个未标注目标域样本 x_t^i ($i = 1, \dots, N_u$)，计算 softmax 概率：

$$p_i = \text{softmax}(f(x_t^i)), \quad c_i = \arg \max_k p_{ik}, \quad \pi_i = \max_k p_{ik},$$

其中 p_{ik} 表示样本 x_t^i 属于第 k 类的预测概率， c_i 为伪标签， π_i 为置信度。为在伪标签可靠性与样本覆盖度之间取得平衡，采用双阈值筛选：

$$\Omega = \{i \mid \pi_i \geq \tau_s\} \cup \{\text{top}-(\tau_d N_u) \text{ by } \pi_i\},$$

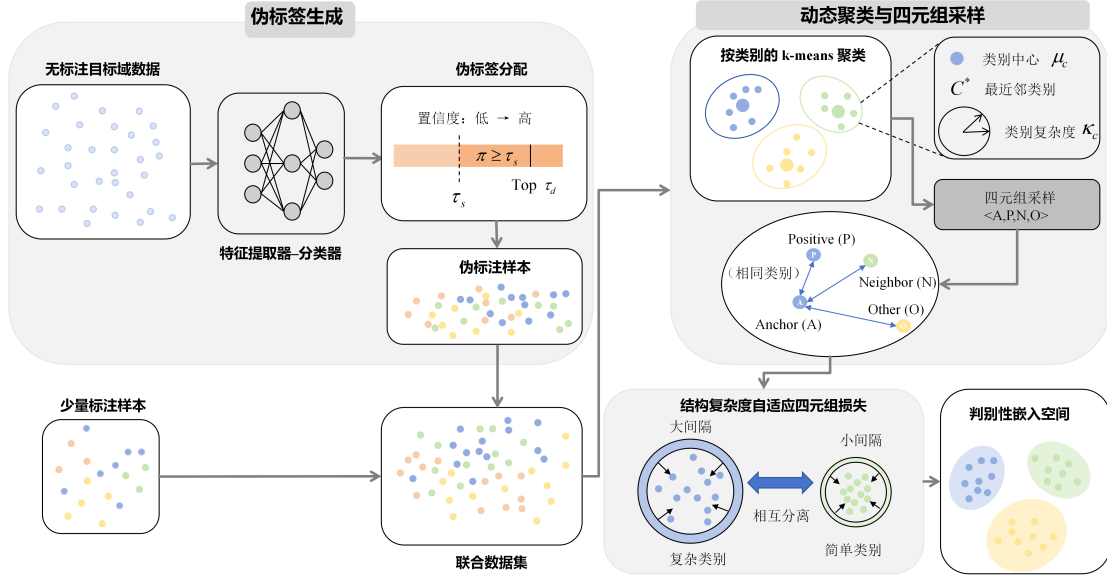


图 5.1 伪标签引导的四元组挖掘与结构复杂度自适应损失

其中 τ_s 为静态置信阈值, τ_d 为动态保留比例。得到伪标签集合

$$\hat{D}_t = \{(x_t^i, c_i)\}_{i \in \Omega}, \quad \Pi = \{\pi_i\}_{i \in \Omega}.$$

该机制在训练早期可避免因模型未充分训练所导致的伪标签枯竭, 在训练后期可抑制模型过度自信导致的误差累积。

(2) 动态聚类与四元组构造

将目标域标注样本与伪标注样本的嵌入表示记为 $Z = [Z_{\text{lab}}, Z_{\text{pseudo}}]$ 。对每个类别 c 的样本集合 Z_c 执行类内 k -means 聚类, 聚类数设置为

$$k_c = \min \{ \lceil \log_2 |Z_c| \rceil, |Z_c| \},$$

从不同簇中抽取 (A, P) 以保证类内多样性。类中心定义为

$$\mu_c = \frac{1}{|Z_c|} \sum_{z \in Z_c} z,$$

并定义最近邻类别为

$$c^* = \arg \min_{q \neq c} \|\mu_c - \mu_q\|_2.$$

进一步, 采用类内平方半径刻画类别结构复杂度:

$$\kappa_c = \left(\frac{1}{|Z_c|} \sum_{z \in Z_c} \|z - \mu_c\|_2^2 \right)^{1/2}. \quad (5.2)$$

基于上述定义, 构造结构化四元组 $\langle A, P, N, O \rangle$:

- A, P : 来自不同簇的同类样本;

- N : 来自最近邻类别 c^* 的样本;
- O : 根据类别复杂度自适应选择的负样本, 以控制优化难度。

若 κ_c 低于全局复杂度中位数 $\check{\kappa}$, 即类别较简单, 则从最近邻类 c^* 中选取距离 A 最远的样本作为半难负样本: $O = \arg \max_{z \in \mathcal{Z}_{c^*}} \|z - z_A\|_2$; 否则从更远类别集合

$$\mathcal{F}_c = \{q \mid \|\mu_q - \mu_c\|_2 > \|\mu_{c^*} - \mu_c\|_2\}$$

中随机抽取 O , 以避免过难负样本在少样本条件下引起训练震荡。该“局部聚类-全局近邻”采样机制使四元组同时具备类内多样性与边界难度。

3. 复杂度自适应双间隔与 SQL 损失

SQL 的核心在于根据类别结构复杂度动态调整类内间隔 α_c 与类间间隔 β_c 。设基础间隔为 α_0, β_0 , 缩放因子为 γ , 并设置上界 $\alpha_{\max}, \beta_{\max}$, 则

$$\alpha_c = \min(\alpha_0 + \gamma \kappa_c, \alpha_{\max}), \quad (5.3)$$

$$\beta_c = \min(\beta_0 + \gamma \max\{\kappa_c, \kappa_{c^*}\}, \beta_{\max}). \quad (5.4)$$

其中 α_c 随类别自身复杂度增大而增大, 以增强复杂类别的类内紧致约束; β_c 由相邻类别的最大复杂度决定, 以在边界类别均复杂时提供更大安全区。为保证数值稳定, α_c, β_c 在前向计算中更新, 但不参与梯度反传。

对四元组 (a, p, n, o) , 记归一化嵌入的平方欧氏距离为 $d_{xy} = \|\hat{z}_x - \hat{z}_y\|_2^2$, 则损失项定义为

$$\mathcal{L}_{ap} = \varphi(d_{ap} - d_{an} + \alpha_c), \quad \mathcal{L}_{an} = \varphi(d_{an} - d_{ao} + \beta_c),$$

其中 $\varphi(\cdot)$ 为 ReLU 函数。对 mini-batch 四元组集合 Q , 四元组平均损失为

$$\mathcal{L}_{quad} = \frac{1}{|Q|} \sum_{(a,p,n,o) \in Q} (\mathcal{L}_{ap} + \mathcal{L}_{an}).$$

为进一步提升类内紧致与类间中心分离, 引入两项正则:

$$\begin{aligned} \mathcal{L}_{compact} &= \frac{1}{K} \sum_{c=1}^K \frac{1}{|\mathcal{Z}_c|} \sum_{z \in \mathcal{Z}_c} \|z - \mu_c\|_2, \\ \mathcal{L}_{sep} &= \frac{1}{K(K-1)} \sum_{c \neq c'} \varphi(v - \|\mu_c - \mu_{c'}\|_2), \end{aligned} \quad (5.5)$$

其中 v 为类中心最小分离阈值。最终 SQL 定义为

$$\mathcal{L}_{SQL} = \mathcal{L}_{quad} + \rho (\mathcal{L}_{compact} + \mathcal{L}_{sep}), \quad (5.6)$$

ρ 为正则权重。由此, SQL 可在少样本条件下实现对复杂类别的更强分离, 同时对简单类别保持稳定收敛。

5.2.3 半监督迁移学习方法设计

本节在 Few-Shot Adversarial Domain Adaptation (FADA) 思想基础上, 构建半监督迁移学习框架 SSTF (Semi-Supervised Transfer Framework)。如图 5.2 所示, SSTF 将源域监督预训练、域-类对抗对齐与伪标签、结构度量细化有机结合, 形成一个可迭代优化的统一流程, 以应对跨塔型声纹信号的域偏移与目标域标注稀缺问题。

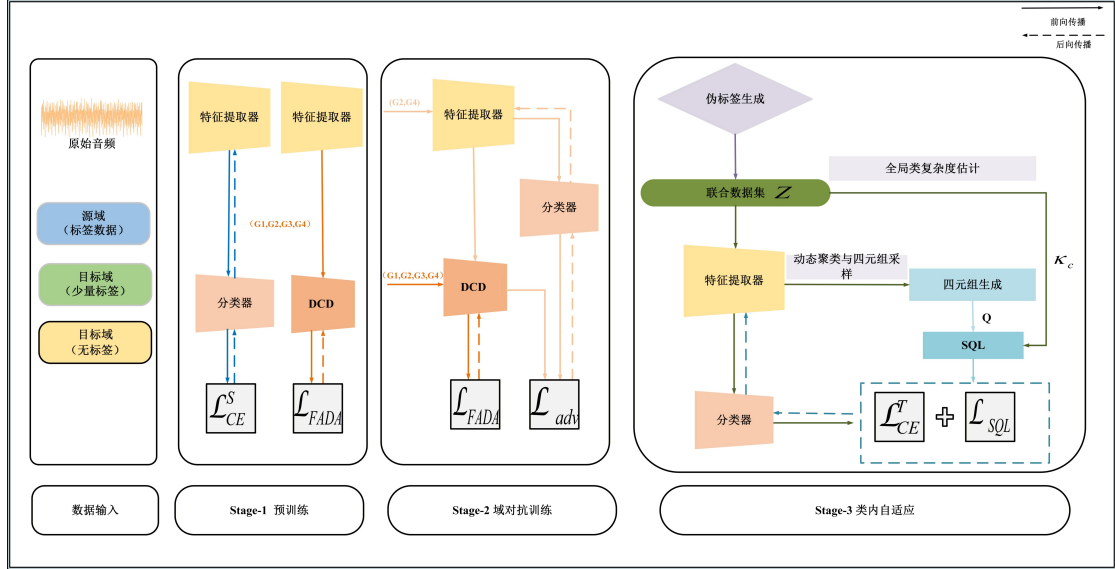


图 5.2 半监督迁移学习框架 SSTF

1. 阶段一：源域监督初始化与域判别器预热

为获得稳定的初始化并降低少样本条件下的训练不确定性, 本文在进入对抗对齐之前先进行两步预训练: (1) 源域监督预训练特征提取器与分类器; (2) 固定特征提取器后对域-类判别器 (Domain-Class Discriminator, DCD) 进行预热训练, 具体过程如算法 5.1 所示。

(1) 源域监督预训练

首先利用源域标注数据 D_s^{lab} 对特征提取器 g 与分类器 h 进行监督训练, 最小化源域交叉熵损失:

$$\mathcal{L}_{CE}^S = -\mathbb{E}_{(x_s, y_s) \sim D_s^{lab}} \sum_{k=1}^K \mathbb{I}(y_s = k) \log \hat{y}_{s,k}, \quad \hat{y}_s = \text{softmax}(h(g(x_s))), \quad (5.7)$$

其中 K 为螺栓状态类别数。该步骤使 $g(\cdot)$ 学到对螺栓松动状态具有判别性的基础表征, 为后续跨塔迁移提供基础。

(2) 域-类判别器 (DCD) 预训练

在完成监督预训练后, 固定特征提取器参数 θ_g (分类器 h 可冻结也可保持不更新, 本文默认冻结), 引入域-类判别器 D 对样本对进行四分类学习。按照

FADA 的配对规则，将样本对划分为四组：

$$\begin{aligned} G_1 &: (x_s, x_s), y_i = y_j; & G_2 &: (x_s, x_t), y_i = y_j; \\ G_3 &: (x_s, x_s), y_i \neq y_j; & G_4 &: (x_s, x_t), y_i \neq y_j, \end{aligned}$$

其中 $x_s \in \mathcal{D}_s^{lab}$, $x_t \in \mathcal{D}_t^{lab}$, y 为真实标签。对每一对样本 (x_i, x_j) ，将其特征拼接为

$$\mathbf{u} = [g(x_i) \parallel g(x_j)], \quad (5.8)$$

并训练 D 输出其所属的组别 G_k 。DCD 的预训练目标为：

$$\mathcal{L}_D = -\mathbb{E}_{(x_i, x_j, G_k)} \sum_{k=1}^4 \mathbf{1}_{\{(x_i, x_j) \in G_k\}} \log D_k(\mathbf{u}). \quad (5.9)$$

该步骤的作用是在进入对抗学习前，使判别器具备基本的域/类配对识别能力，从而为后续的交替对抗优化提供更稳定的梯度信号，避免训练初期出现判别器过弱导致的对齐失效。

算法 5.1 阶段一：源域监督预训练与判别器预热

Input: 源域有标签集 $\mathcal{D}_s^{lab}$ ，目标域少量有标签集 $\mathcal{D}_t^{lab}$ ，特征提取器 g ，分类器 h ，域-类判别器 D ，学习率 η

Output: 预训练后的参数 θ_g, θ_h ，判别器初始参数 θ_d

```

1 (1) 源域监督预训练 (更新  $g, h$ ) ;
2 初始化优化器 Adam( $\{\theta_g, \theta_h\}, \eta$ );
3 for epoch = 1 to  $n_{pre}$  do
4     foreach  $(x_s, y_s) \in \mathcal{D}_s^{lab}$  do
5          $z \leftarrow g(x_s)$ ;
6          $\hat{y} \leftarrow h(z)$ ;
7          $\mathcal{L}_{CE}^S \leftarrow \text{CE}(\hat{y}, y_s)$ ;
8         反向传播并更新  $\theta_g, \theta_h$ ;
9     end
10 end
11 (2) 判别器预热 (冻结  $g$ , 更新  $D$ ) ;
12 冻结  $\theta_g$  (必要时也可冻结  $\theta_h$ ) ;
13 初始化优化器 Adam( $\{\theta_d\}, \eta$ );
14 for epoch = 1 to  $n_{warm}$  do
15     从  $\mathcal{D}_s^{lab}, \mathcal{D}_t^{lab}$  采样样本对  $(x_i, x_j, G_k) \leftarrow \text{采样配对}(\cdot)$ ;
16      $\mathbf{u} \leftarrow [g(x_i) \parallel g(x_j)]$ ;
17      $\mathcal{L}_D \leftarrow \text{CE}(D(\mathbf{u}), G_k)$ ;
18     反向传播并更新  $\theta_d$ ;
19 end
20 return  $\theta_g, \theta_h, \theta_d$ ;
```

2. 阶段二：域-类判别器与对抗对齐的交替优化

仅依赖源域监督初始化难以直接适配目标塔型的结构差异与噪声差异，因此本文在阶段一预训练基础上，引入域-类判别器 D 实现跨域对齐。与传统仅判别“域标签”的对抗方法不同，DCD 采用 FADA 的配对四分类策略，在推动源/目

标域特征靠拢的同时显式约束类别语义一致性，从而减少“对齐但错类”的负迁移风险。

如算法 5.2 所示，阶段二采用“更新特征提取器-更新判别器”的交替迭代方式。其核心思想是：在固定判别器 D 时，更新 g 以及 h ，使跨域同类对 (G_2) 更难被判别器区分为“跨域”，从而促使同类样本在特征空间中对齐，同时保持异类对 (G_4) 仍可被区分，以维持类别判别性。

具体而言，首先固定 D ，采样跨域配对组 G_2 与 G_4 ，构造对抗损失：

$$\mathcal{L}_{adv}^{(g,h)} = \mathcal{L}_{cls} + \eta \mathcal{L}_D^{(G_2, G_4)}, \quad \eta > 0, \quad (5.10)$$

其中 $\mathcal{L}_{cls}$ 为监督分类损失，用于维持嵌入空间的类别判别结构， η 表示平衡超参数，用于调节分类损失与对抗损失在总目标中的相对权重；在实现上可由源域有标签数据与目标域少量有标签数据共同计算：

$$\mathcal{L}_{cls} = \mathcal{L}_{CE}^S + \xi \mathcal{L}_{CE}^T, \quad (5.11)$$

ξ 为目标域监督项权重，通常取 $[0, 1]$ 的较小值以适配少样本场景。 $\mathcal{L}_D^{(G_2, G_4)}$ 表示仅在跨域配对组上计算的 DCD 交叉熵：

$$\mathcal{L}_D^{(G_2, G_4)} = -\mathbb{E}_{(x_i, x_j, G_k) \in (G_2 \cup G_4)} \sum_{k \in \{2, 4\}} \mathbf{1}_{\{(x_i, x_j) \in G_k\}} \log D_k(\mathbf{u}). \quad (5.12)$$

通过最小化式 (5.10) 更新 θ_g, θ_h ，可推动跨域同类对的特征分布靠近，同时在 $\mathcal{L}_{cls}$ 约束下避免类别语义坍塌。

随后固定 g, h ，沿用阶段一的配对规则，将样本对划分为四组，使用四组配对样本 $G_1 \sim G_4$ 更新判别器 D ，其损失为：

$$\mathcal{L}_D = -\mathbb{E}_{(x_i, x_j, G_k)} \sum_{k=1}^4 \mathbf{1}_{\{(x_i, x_j) \in G_k\}} \log D_k(\mathbf{u}). \quad (5.13)$$

上述两步交替迭代，使得 g 不断学习“域不变且类可分”的表征，而 D 持续提升其对配对组别的辨识能力，从而形成稳定的对抗博弈过程，实现跨塔型特征对齐与语义一致性的协同优化。

3. 阶段三：目标域类内自适应

仅依靠对抗对齐往往只能实现粗粒度分布匹配，难以解决目标域内部的细粒度类间混淆。因此，如算法 5.3 所示，SSTF 在对齐基础上引入双阈值伪标签策略生成高置信伪标注集 $\hat{D}_t$ ，并将其与少量真实标注 D_t^{lab} 合并为统一训练集。目标域监督损失为

$$\mathcal{L}_{CE}^T = - \sum_{(x,y) \in (D_t^{lab} \cup \hat{D}_t)} \sum_{k=1}^K \mathbb{I}(y=k) \log \hat{y}_k. \quad (5.14)$$

算法 5.2 阶段二：域-类判别器与对抗域对齐

Input: 源域有标签集 D_s^{lab} , 目标域少量有标签集 D_t^{lab} , 预训练 g, h, D , 对抗权重 η_{adv} , 学习率 η

Output: 对齐后的参数 $\theta_g, \theta_h, \theta_d$

- 1 初始化优化器 $\text{Adam}(\{\theta_g, \theta_h\}, \eta)$ 与 $\text{Adam}(\{\theta_d\}, \eta)$;
- 2 **for** epoch = 1 **to** n_{adv} **do**
 - // (a) 更新 g, h : 固定 D , 主要使用跨域组 G_2, G_4
 - 3 冻结 θ_d , 解冻 θ_g, θ_h ;
 - 4 从 $G_2 \cup G_4$ 采样样本对 $(x_i, x_j, y_i, y_j, G_k)$;
 - 5 $z_i \leftarrow g(x_i)$, $\hat{y}_i \leftarrow h(z_i)$; $z_j \leftarrow g(x_j)$, $\hat{y}_j \leftarrow h(z_j)$;
 - 6 $\mathcal{L}_{cls} \leftarrow \text{CE}(\hat{y}_i, y_i) + \text{CE}(\hat{y}_j, y_j)$;
 - 7 $\mathcal{L}_D^{(2,4)} \leftarrow \text{CE}(D(\text{GRL}([z_i \| z_j])), G_k)$;
 - 8 反向传播并更新 θ_g, θ_h , 损失为 $\mathcal{L}_{cls} + \eta_{adv} \mathcal{L}_D^{(2,4)}$;
 - // (b) 更新 D : 固定 g, h , 在四组配对 $G_1 \sim G_4$ 上训练判别器
 - 9 冻结 θ_g, θ_h , 解冻 θ_d ;
 - 10 从 $G_1 \sim G_4$ 采样样本对 (x_i, x_j, G_k) ;
 - 11 $\mathbf{u} \leftarrow [g(x_i) \| g(x_j)]$;
 - 12 $\mathcal{L}_D \leftarrow \text{CE}(D(\mathbf{u}), G_k)$;
 - 13 反向传播并更新 θ_d ;
- 14 **end**
- 15 **return** $\theta_g, \theta_h, \theta_d$;

与此同时, 利用 $D_t^{lab} \cup \hat{D}_t$ 的嵌入构造 PL-SQ 四元组集合 Q , 并在线估计类别复杂度 κ_c 、自适应间隔 α_c, β_c , 引入 SQL 损失 $\mathcal{L}_{SQL}$ (式 (5.6)) 以增强近邻类别分离与类内紧致。

4. 联合优化目标

综合上述机制, SSTF 在目标域联合微调阶段采用如下总体目标函数:

$$\mathcal{L} = \mathcal{L}_{CE}^S + \mathcal{L}_{CE}^T + \mathcal{L}_D + \mathcal{L}_{adv} + \lambda(t) \mathcal{L}_{SQL}, \quad (5.15)$$

其中 $\lambda(t)$ 为随训练轮次逐步增大的权重, 用于在特征尚不稳定的早期抑制度量约束、在后期增强类间边界学习。由此, SSTF 能够在 $|D_t^{lab}| \ll |D_s^{lab}|$ 的条件下, 充分利用源域知识与目标域未标注数据, 实现跨塔型少样本螺栓松动检测的鲁棒迁移与精细定位。

5.3 实验设计与结果分析

5.3.1 实验数据集与预处理

如图 5.3 所示, 实验采用宏源工业区的两座 500 kV 输电铁塔数据集: 源域为 500 kV 干字塔型, 目标域为 500 kV 酒杯塔型。声激励装置-传感器采用对角布局进行数据采集, 每个塔型设置 5 类工况 (0 表示紧固, 其余表示不同位置松动, 且松动定义为扭矩小于 30 N·m)。原始音频裁剪为 2 s, 并重采样至 16 kHz, 提取 64 维 log-Mel 频谱并按式 (4.3) 进行归一化。实验数据量统计见表 5.1。

算法 5.3 阶段三：伪标签引导的类内自适应与 SQL 优化

Input: 目标域有标签集 D_t^{lab} , 目标域无标签集 D_t^{unlab} , 已对齐的特征提取器 g 与分类器 h , 阈值 (τ_s, τ_d) , 权重调度 $\lambda(t)$, 学习率 η

Output: 精炼后的参数 θ_g, θ_h

```

1 初始化优化器 Adam( $\{\theta_g, \theta_h\}, \eta$ );
2 for epoch = 1 to  $n_{ref}$  do
    // (1) 伪标签生成与筛选：双阈值策略
3   $\hat{D}_t \leftarrow$  高置信伪标签筛选( $D_t^{unlab}, f = h \circ g, \tau_s, \tau_d$ );
4   $D_t^{mix} \leftarrow D_t^{lab} \cup \hat{D}_t$ ;
    // (2) 结构统计：提取嵌入、类内聚类、构造四元组并计算复杂度与自适应间隔
5   $Z \leftarrow$  提取嵌入特征( $g, D_t^{mix}$ );
6   $Q, \{\kappa_c\}, \{\alpha_c, \beta_c\} \leftarrow$  PL-SQ + 间隔自适应( $Z$ );
    // (3) 联合优化：分类监督 + 结构复杂度自适应四元组损失
7  foreach mini-batch  $B \subset D_t^{mix}$  do
8       $\mathcal{L}_{CE}^T \leftarrow 0$ ;
9      foreach  $(x, y) \in B$  do
10          $z \leftarrow g(x); \hat{y} \leftarrow h(z)$ ;
11          $\mathcal{L}_{CE}^T \leftarrow \mathcal{L}_{CE}^T + CE(\hat{y}, y)$ ;
12     end
13      $\mathcal{L}_{SQL} \leftarrow SQL(Q; \{\alpha_c\}, \{\beta_c\})$ ;
14     反向传播并更新  $\theta_g, \theta_h$ , 损失为  $\mathcal{L}_{CE}^T + \lambda(t)\mathcal{L}_{SQL}$ ;
15 end
16 end
17 return  $\theta_g, \theta_h$ ;
```

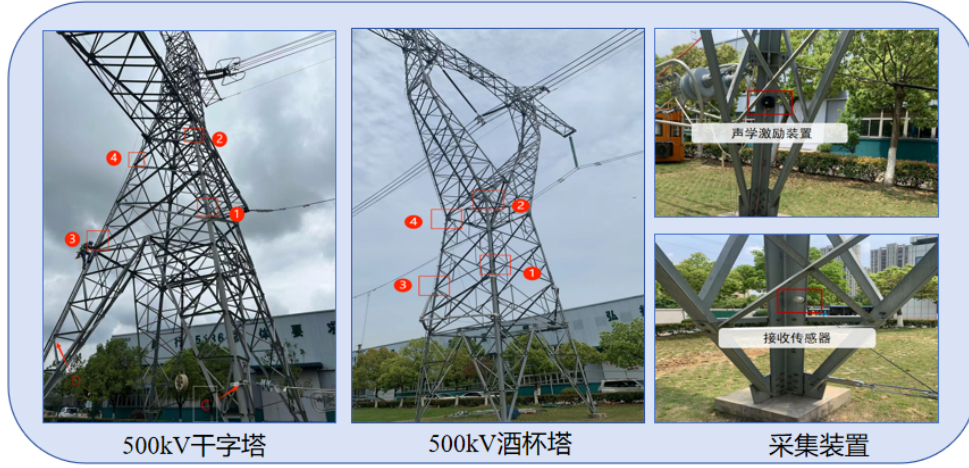


图 5.3 两座铁塔的数据采集位置和传感器/声学激励装置。

5.3.2 实验环境设置与评价指标

1. 训练实现细节

本文的实验环境与第 4.4.2 小节中的设置一致，所有实验在 NVIDIA GeForce RTX 2080 Ti GPU（显存 12GB）上完成。所提出的半监督迁移学习框架（SSTF）采用三阶段训练策略，整体流程见算法 5.1–5.3。

(1) **源域监督预训练。**首先在源域有标签数据集 D_s^{lab} 上对特征提取器 g 与

表 5.1 不同塔型与螺栓状态的声学数据集样本分布

塔型	类别 0 (紧固)	类别 1	类别 2	类别 3	类别 4	总计
500 kV 干字塔	1001	870	955	850	875	4551
500 kV 酒杯塔	985	745	700	910	905	4245

分类器 h 进行监督预训练, 以获得稳定的基础判别表征。该阶段训练轮数设为 10, 优化目标为源域交叉熵损失 $\mathcal{L}_{CE}^S$ (式 (5.7))。

(2) **域-类判别器预训练**。在冻结特征提取器参数 θ_g 的条件下, 对域-类判别器 DCD 进行预训练, 使其能够对配对样本的“域关系与类别关系”进行四分类判别 (对应组 $G_1 \sim G_4$)。该阶段训练轮数设为 100, 损失函数为 $\mathcal{L}_D$ (式 (5.9))。

(3) **目标域联合微调与类内自适应**。在完成初步对齐后, 使用目标域少量标注样本 D_t^{lab} 与大量无标签样本 D_t^{unlab} 进行联合微调与类内自适应, 训练轮数设为 1000。该阶段综合优化三类目标: 监督分类损失、对抗域对齐损失以及结构复杂度自适应四元组损失 $\mathcal{L}_{SQL}$ (式 (5.6))。伪标签采用双阈值筛选策略: 静态置信阈值为 $\tau_s = 0.9$, 动态保留比例为 $\tau_d = 0.5$ 。SQL 的基础间隔设为 $\alpha_0 = 0.2$ 、 $\beta_0 = 0.4$, 结构复杂度缩放系数为 $\gamma = 0.5$, 并采用 $\lambda(t)$ 对 $\mathcal{L}_{SQL}$ 权重进行渐进调度 (式 (5.15)), 以避免早期伪标签噪声对度量学习造成不稳定影响。

上述三个阶段均使用 Adam 优化器, 学习率为 1×10^{-3} , 并设置 $\beta_1 = 0.9$, $\beta_2 = 0.999$, 批大小为 64。数据集按 70%/15%/15% 划分为训练集、验证集和测试集。训练过程中以验证集损失为准进行早停 (patience=30), 从而抑制过拟合并提升模型选择的可靠性。

2. 对比实验与消融设置

为全面评估本文方法在少样本、跨塔迁移场景下的有效性与鲁棒性, 本文设置了半监督学习和半监督域自适应两组对比实验。

为分析在不同标注比例下 SQL 与整体框架的有效性, 本文在目标域设置 5%, 10%, 20%, 30%, 40%, 50% 六种标注比例, 其余样本作为无标签数据参与训练。对比方法分为两类: (i) 纯半监督学习方法: Pseudo-Label^[68]、FixMatch^[69]、MixMatch^[70], 仅利用目标域少量标注与大量无标签样本进行训练; (ii) 半监督域自适应方法: 选取 8 种代表性方法进行对比 (包括将 DANN^[75] 与伪标签自训练结合的 SSDANN 等)。为保证公平性, 所有方法统一采用相同的特征提取器与分类器结构 (见第 4.3.1 小节) 并在 5 个不同随机种子下重复实验, 并以 Top-1 准确率作为核心指标进行比较。

5.3.3 实验结果分析

图 5.4 给出了在目标域不同标注比例下, 各对比方法在目标域测试集上的平均分类准确率变化趋势, 详细数值结果见表 5.2。

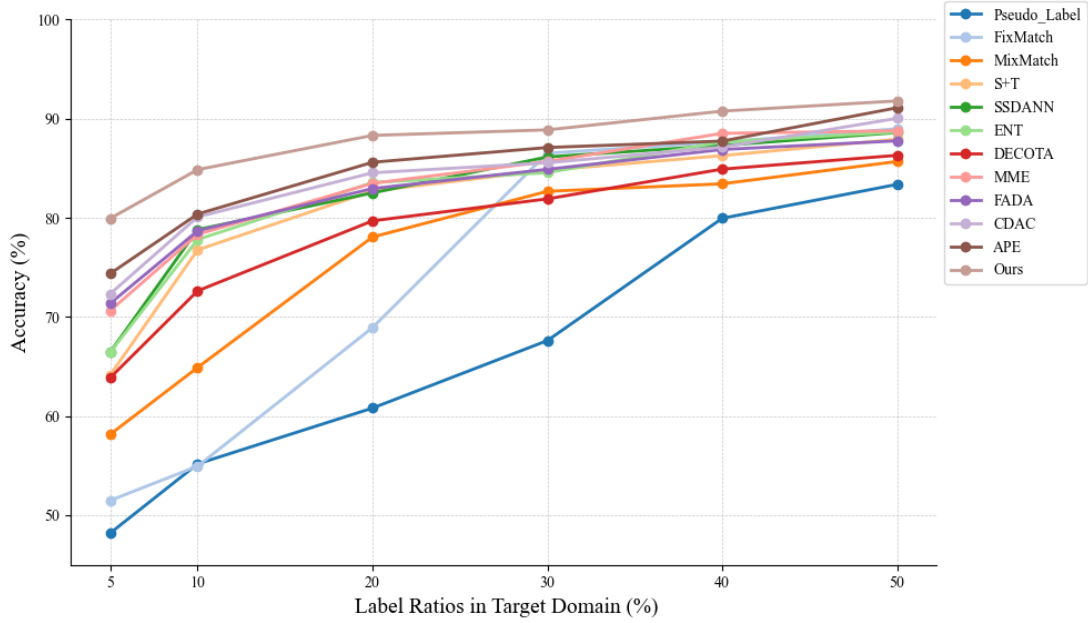


图 5.4 在目标域不同标注比例下，各对比方法在目标域测试集上的平均分类准确率

综合图表结果，可以得到以下结论：

1. **总体趋势与低标注优势显著。**随着目标域标注比例从 5% 提升至 50%，各方法的准确率均呈现稳定上升趋势，符合半监督学习“标注越多、性能越高”的一般规律。在最具挑战性的低标注场景下，本文方法优势尤为明显。当标注比例为 5% 和 10% 时，本文方法分别取得 **79.94%** 与 **84.86%** 的准确率，相比第二优方法 APE 分别提升 5.57% 与 4.48%。当标注比例达到 50% 时，本文方法仍保持领先，表明其在不同标注规模下均具有较好的稳定性与可扩展性。
2. **纯半监督方法在少样本条件下表现受限。**从表 5.2 可见，Pseudo-Label、Fix-Match、MixMatch 等仅依赖目标域“少量标注 + 大量无标签”的方法在 5% ~ 20% 标注阶段整体性能明显落后，说明在输电铁塔螺栓松动这种跨结构差异显著、标注稀缺、噪声复杂的任务中，仅靠目标域自训练容易受到伪标签噪声影响，难以学习到判别性强的特征表征。相比之下，引入源域先验知识与跨域对齐约束的半监督域自适应方法整体优于纯半监督基线，验证了“源域知识迁移”对缓解少标注瓶颈的必要性。
3. **与主流半监督域自适应方法相比，本文方法具有持续领先优势。**在全部标注比例下，本文方法均优于 8 种半监督域自适应对比方法。尤其在 5% 标注比例下，本文方法相较 FADA 提升 8.58%，体现出本文框架在极少标注场景下的有效性。进一步观察图 5.4 的性能曲线可以发现，本文方法的曲线更平滑、提升更稳定，说明其能够更有效地利用无标签数据并逐步修正伪标签分布，同时避免了少样本条件下常见的震荡与过拟合问题。这主要

得益于：一方面，域-类判别器引导的对抗对齐缓解了源/目标域分布差异；另一方面，SQL 通过结构复杂度自适应间隔强化了类间边界、提升了类内聚合，从而增强了目标域的判别性表示学习。

表 5.2 目标域不同标注比例下各方法的平均分类准确率 (%)

方法	5% 标注	10% 标注	20% 标注	30% 标注	40% 标注	50% 标注
A. 仅目标域半监督学习 (SSL)						
Pseudo-Label ^[68]	48.19	55.17	60.80	67.61	79.95	83.37
FixMatch ^[69]	51.48	54.94	68.92	86.52	87.65	88.98
MixMatch ^[70]	58.15	64.90	78.06	82.67	83.43	85.68
B. 半监督域自适应 (SSDA)						
S+T (源域 + 目标域联合训练)	64.14	76.77	82.70	84.80	86.27	87.93
DECOTA ^[76]	63.86	72.63	79.69	81.91	84.91	86.30
SSDANN	66.46	78.81	82.51	86.14	87.30	88.60
ENT ^[77]	66.43	77.81	83.55	84.55	87.65	88.65
MME ^[78]	70.63	78.33	83.48	85.68	88.52	88.78
FADA ^[72]	71.36	78.69	82.95	84.89	86.89	87.76
CDAC ^[79]	72.33	80.09	84.54	85.55	87.13	90.03
APE ^[80]	74.37	80.38	85.60	87.09	87.74	91.10
本文方法 (MSEAM + SQL)	79.94	84.86	88.31	88.87	90.76	91.59

总体而言，本文提出的半监督迁移学习框架在不同标注比例下均表现出稳定且领先的识别性能，尤其在极少标注条件下优势突出，验证了“对抗域对齐 + 伪标签引导类内自适应 + 结构复杂度自适应度量学习”联合优化策略的有效性与鲁棒性。

1. 混淆矩阵分析

为进一步探究不同方法在目标域低标注条件下的误判模式，本文在目标域标注比例为 10% 的设置下，对性能排名前八的方法绘制混淆矩阵进行对比分析，如图 5.5 所示。总体来看，多数方法在同一主材相邻松动位置的判别上存在较明显混淆。这一现象与输电铁塔结构声学响应的物理特性密切相关，即相邻连接点的结构模态与共振特征高度相似，导致其声谱特征在频带能量分布上呈现较强重叠，从而使分类边界更加模糊。

具体而言，位置“1”与“2”之间的互相误判是主要错误来源之一。除本文方法外，多数模型在该对相邻类别上的误判比例均超过 20%，其中 CDAC 的误判比例甚至超过 30%，表明在细粒度定位任务中，部分迁移学习方法对“邻近类”可分性刻画不足，空间分辨能力受限。类似地，位置“3”与“4”的总误判比例普遍超过 10%。相比之下，本文方法在相邻类区分上表现更优：位置“1”与“2”的误判比例降低至 17.7%，位置“3”与“4”的误判比例降低至 7.6%。同时，本文方法在五个类别上均能保持 80% 以上的单类识别准确率，说明在标注稀缺条件下，所提出的 MSEAM+SQL 框架能够学习到更具判别性且可迁移的结构声学表征，从而提升相邻类别的分离能力并增强跨域鲁棒性。

此外，图 5.5 中还可观察到位置“0”（完全紧固）与“4”之间存在一定程度

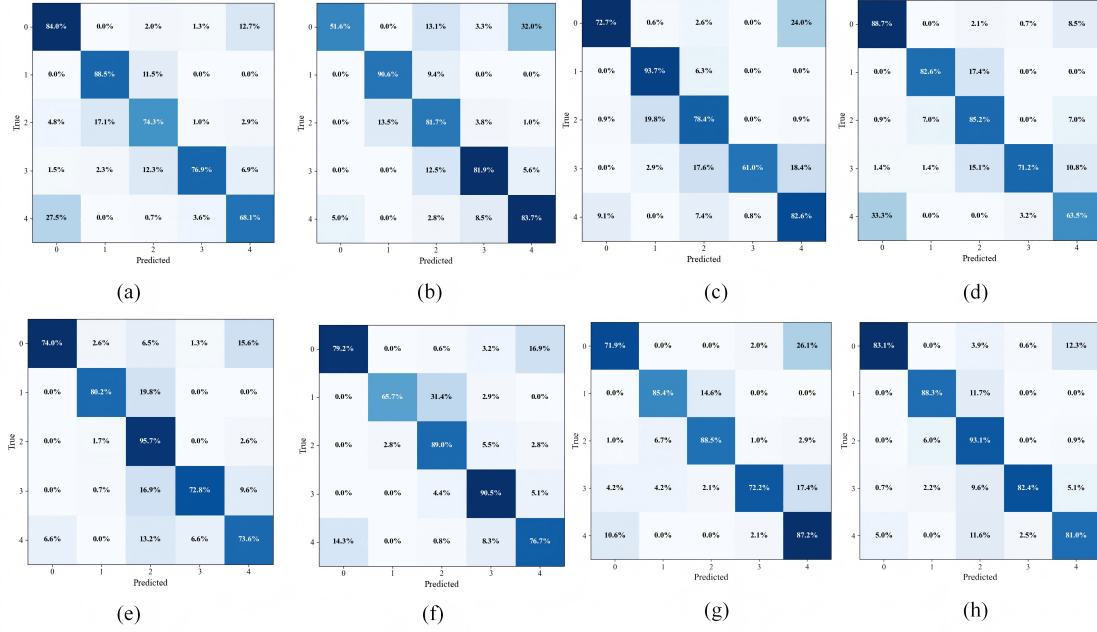


图 5.5 目标域 10% 标注条件下八种方法的混淆矩阵对比：(a) S+T；(b) SSDA；(c) ENT；(d) MME；(e) FADA；(f) CDAC；(g) APE；(h) 本文方法。

的混淆。由于二者在空间位置上并不相邻，该误判更可能源自实验布置所带来的激励-响应路径相似性，在“对角布置”的声激励装置与传感器配置下，部分工况可能产生相近的传播衰减与频带响应形态，从而造成声学特征的类型相似。关于该现象的更深入讨论将在下一节结合声谱对比进一步展开。

综上，混淆矩阵结果表明在目标域低标注场景下，本文方法能够显著降低相邻位置之间的误判，并保持更均衡的类别识别性能，从而验证了结合 MSEAM 与 SQL 的半监督迁移学习框架在少样本跨塔螺栓松动定位任务中的有效性与鲁棒性。

2. 注意力机制消融与 SQL 有效性验证

为验证所提出多尺度能量注意力模块 (MSEAM) 与结构复杂度自适应四元组损失 (SQL) 的有效性，本文在目标域标注比例为 10% 的少样本条件下开展消融实验。具体地，将 MSEAM 与 SQL 的“是否启用”作为两个二元因素，构造四种模型变体：Model1 (同时移除 MSEAM 与 SQL)、Model2 (仅引入 MSEAM)、Model3 (仅引入 SQL) 以及 Model4 (同时引入 MSEAM 与 SQL 的完整模型)。各模型均在相同设置下训练，并采用 Top-1 准确率、召回率与 F1 分数作为评价指标，结果汇总于表 5.3。

从表 5.3 可见，Model1 的 Top-1 准确率、召回率和 F1 分数分别为 80.41%、80.4% 与 79.60%，作为基线模型在少标注条件下存在明显性能瓶颈。仅引入 MSEAM 的 Model2 将 Top-1 准确率提升至 83.27%，F1 分数提升至 83.35%，说明多尺度能量注意力能够有效突出与松动相关的关键频带能量模式，增强特征表

征的可分性。仅引入 SQL 的 Model3 同样带来增益,表明结构复杂度自适应间隔能够改善类内紧致性与类间分离度,缓解相邻位置或相近工况造成的类间混淆。

表 5.3 10% 标注率下 MSEAM 与 SQL 的消融实验结果 (单位: %)

模型	MSEAM	SQL	Top-1 准确率	召回率	F1 分数
Model1	✗	✗	80.41	80.4	79.60
Model2	✓	✗	83.27	83.8	83.35
Model3	✗	✓	83.10	83.1	83.00
Model4	✓	✓	85.58	85.6	85.30

当同时引入 MSEAM 与 SQL 时, Model4 取得最优结果,相较于基线模型 (Model1) 在准确率上提升 5.17 个百分点,在 F1 上提升 5.70 个百分点,体现出两者在关键频带增强与判别边界拉开方面具有互补性。进一步地,图 5.6 的箱线图结果表明,完整模型在五次重复实验中具有更高的中位数与均值,同时方差更小;其中 Model1 出现明显离群点,反映出训练不稳定或对初始化较敏感,而 Model4 的结果分布更集中,说明本文方法在低标注条件下能提供更稳定、鲁棒的识别性能。

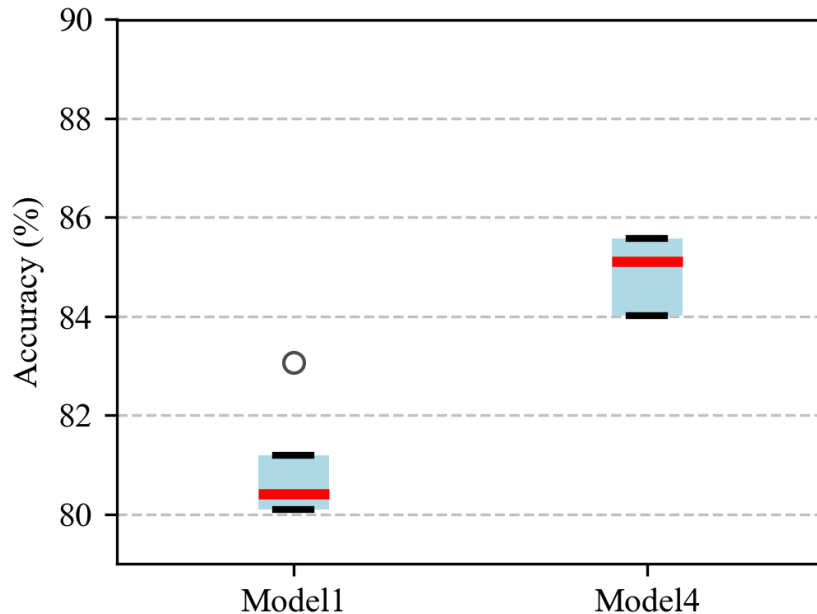


图 5.6 模型 1 (不使用 MSEAM 和 SQL) 和模型 4 (使用 MSEAM 和 SQL) 在目标域 10% 标注率下进行 5 次重复实验的分类结果箱线图。

3. SQL 在类间分离中的有效性分析

为进一步验证结构复杂度自适应四元组损失在少样本条件下提升类间可分性的作用,本文选取 10% 目标域标注率场景,对两种模型变体的测试集嵌入特征进行可视化对比: Model2 (不含 SQL) 与 Model4 (引入 SQL)。具体地,采用 t-SNE 将高维嵌入投影至二维平面,并结合混淆矩阵对类别混淆来源进行分析,结果如图 5.7 与图 5.8 所示。

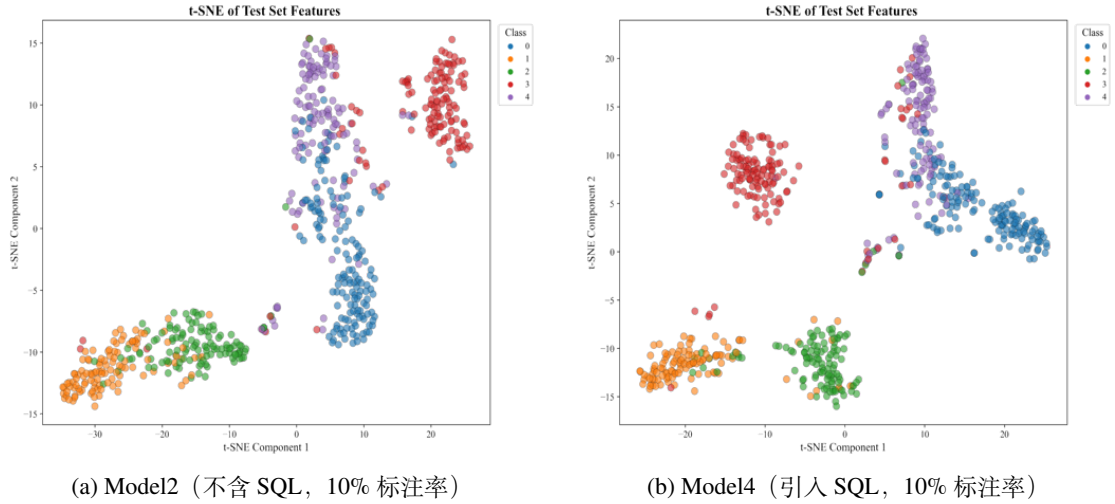


图 5.7 目标域 10% 标注率条件下测试集嵌入特征的 t-SNE 可视化对比。

从图 5.7 可以观察到，引入 SQL 后（Model4）同类样本在特征空间中呈现更为紧致的聚集形态，类间边界更清晰；相较之下，未引入 SQL 的 Model2 存在更明显的类间重叠现象，尤其在空间相邻位置对应的类别上更易发生混叠。这表明 SQL 通过结构化四元组采样与基于结构复杂度的自适应间隔机制，能够在保持类内多样性的同时，进一步拉大难分样本对的判别距离，从而改善嵌入空间的几何结构。

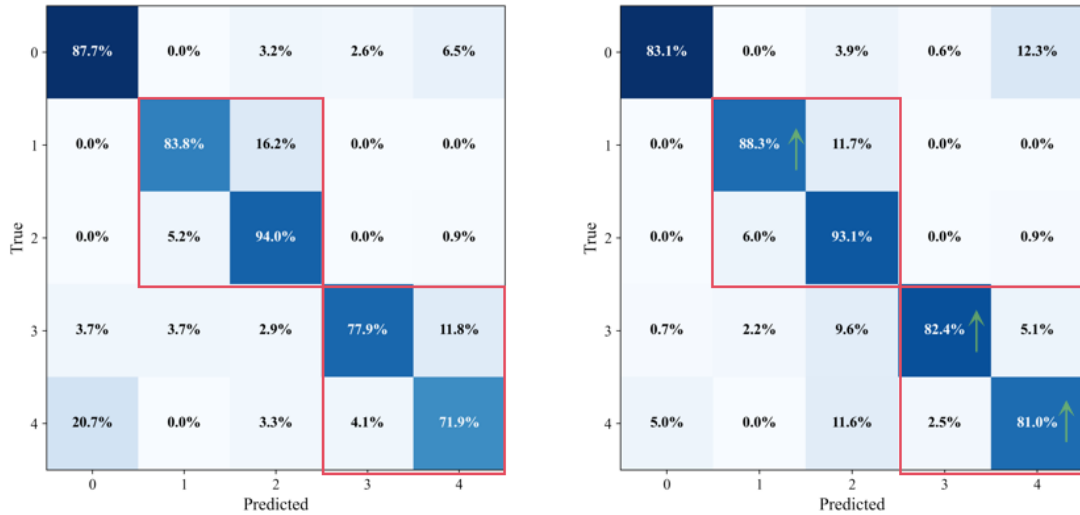


图 5.8 目标域 10% 标注率条件下的混淆矩阵对比：左为 Model2（不含 SQL），右为 Model4（引入 SQL）。

如图 5.8 所示，混淆矩阵结果进一步从定量角度支持上述结论。与 Model2 相比，Model4 在多个易混类别对上误判率显著降低。类别 1 与类别 2 的互相误判减少 3.7%，类别 3 与类别 4 的混淆减少 8.3%，类别 0 与类别 4 的混淆减少 9.9%；同时，Model4 在类别 4 上的识别精度提升最高达 9.1%。值得注意的是，Model2 将 20.7% 的松动样本误判为“紧固态”（类别 0），该类错误在输电铁塔状态监测

等安全关键应用中具有较高风险；而 SQL 的引入显著降低了此类误判，体现出其对“难分/高风险类别”的强化判别能力。

尽管同一主材上相邻位置（如类别 1 与类别 2、类别 3 与类别 4）在声学响应上具有一定相似性，导致存在一定程度的混淆属于合理现象，但类别 0 与类别 4 的频繁误判更值得关注：二者在空间上并非相邻位置，其混淆更可能来源于传播路径与传感器布置带来的信号衰减效应。结合图 5.9 的谱图对比可见，类别 4（D 主材约 18 m 高度处螺栓松动）在局部频带表现出更强的能量波动，但由于信号需跨越较长传播距离到达安装在 C 主材约 1 m 高度处的传感器，信号衰减会削弱其与紧固态（类别 0）之间的差异性，进而降低诊断可分性。这一现象揭示了当前单传感器对远距离位置的感知存在一定局限，也为后续引入多传感器协同布设、优化激励和采集配置提供了依据。

综上，SQL 在少标注与跨域条件下能够针对类内结构差异较大、边界更复杂的类别分配更强的判别约束，从而实现更细粒度的类间边界建模，缓解由结构共振相似性、传播衰减以及域偏移引起的类别混淆，其中在类别 1 与类别 4 等易错类别上的提升，进一步验证了 SQL 对复杂类别与高风险误判的抑制作用。

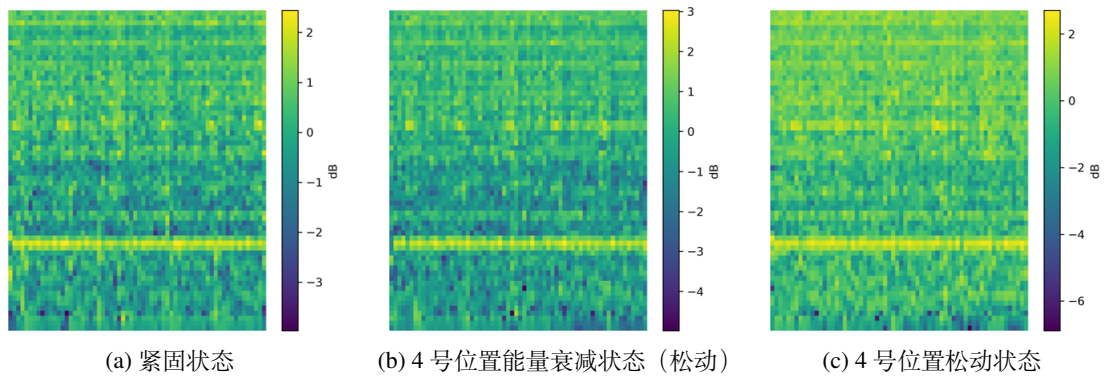


图 5.9 500 kV 酒杯塔目标域样本的 Mel 频谱图对比：(a) 紧固状态；(b) 4 号位置能量衰减状态（松动）；(c) 4 号位置松动状态。

4. 参数敏感性分析

为验证结构复杂度自适应四元组损失中自适应间隔参数的动态行为及关键超参数 γ 的敏感性，本文在目标域标注比例为 10% 的设置下进行参数分析。图 5.10 展示了在固定 $\alpha_0 = 0.2$ 、 $\beta_0 = 0.4$ 、 $\gamma = 1$ 时，训练过程中类内间隔 α 与类间间隔 β 的均值变化趋势。

从图 5.10 可见，训练初期由于特征空间尚未收敛、类别分布较为离散，各类别结构复杂度 κ_c 相对较高，使得自适应间隔在 γ 的缩放作用下整体偏大。随着训练推进，特征逐步形成更紧凑的类内聚集与更清晰的类间分离， α 与 β 的均值呈现缓慢下降趋势，反映出类内离散度降低、类间边界逐渐稳定。值得注意的是，在约第 72 个 epoch 附近出现明显拐点， α/β 均值发生快速下降并随后趋于平

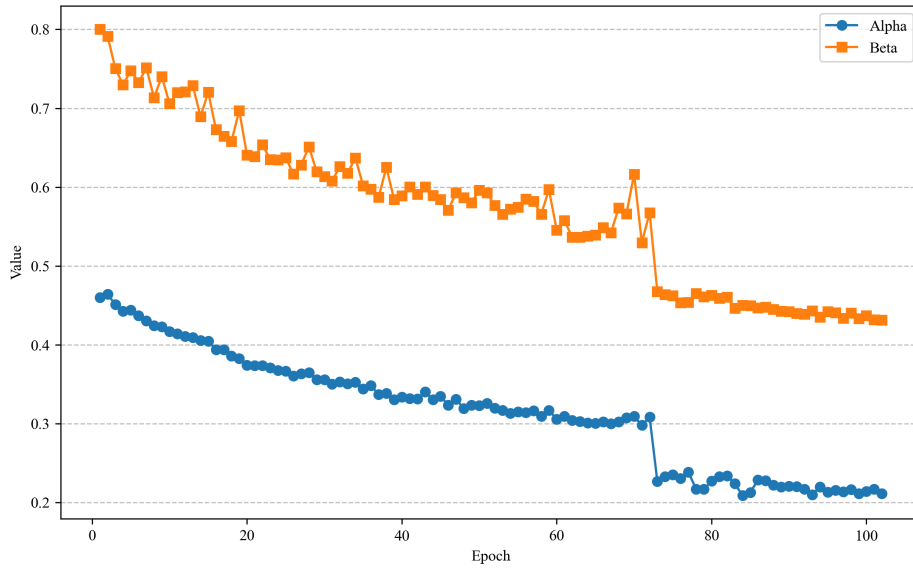


图 5.10 目标域标注比例为 10% 时自适应间隔参数 α/β 的均值随训练轮次变化曲线 ($\alpha_0 = 0.2$, $\beta_0 = 0.4$, $\gamma = 1$)。

稳。该现象通常与伪标签质量阶段性提升或聚类结构突然收敛有关：当高置信伪标签数量增加、类中心估计更加准确时，特征簇会出现“收缩—重排”，从而触发间隔参数的同步重配置。这一过程表明 SQL 具有“结构—间隔”自适应机制，即结构复杂度的变化会反馈到间隔更新中，进而继续推动更稳定的判别特征学习。

在训练后期， α 与 β 均稳定在接近基准间隔的范围内，其中 $\alpha \in [0.21, 0.24]$ 、 $\beta \in [0.43, 0.47]$ 。少量波动主要来源于伪标签的迭代更新与周期性聚类重估，但整体已进入稳定收敛阶段。这说明所设计的自适应策略满足“前期动态、后期收敛”的期望，早期提供更强的判别约束以快速拉开易混类别，后期逐渐回归稳态以避免过度约束导致的训练不稳定。

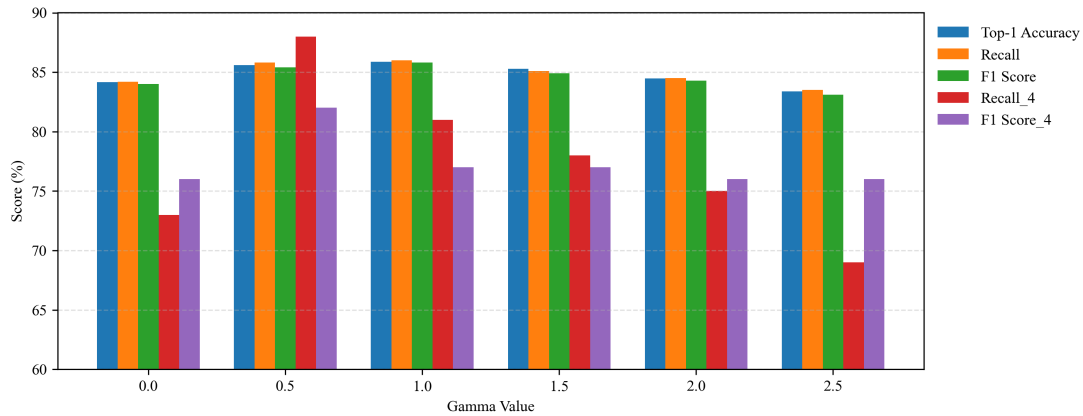


图 5.11 目标域标注比例为 10% 时，不同 γ 取值下的 Top-1 准确率 (Top-1 Accuracy)、召回率 (Recall)、F1 分数 (F1 Score) 以及类别 4 的召回率 (Recall_4) 与 F1 分数 (F1 Score_4)。

进一步地,图 5.11 给出了不同 γ 取值下的整体性能 (Top-1 准确率/召回率/F1 分数) 以及类别 4 的召回率与 F1 分数,用于评估 γ 对复杂度缩放强度的影响。结果表明, γ 过小时,自适应间隔对复杂度差异的响应不足,难以针对结构复杂或易混类别施加足够的判别压力,从而限制泛化性能;而 γ 过大时,会导致间隔膨胀过快并频繁触发上界裁剪,使优化目标变得过于激进,进而造成类内/类间距离分布失衡,甚至出现对困难类别 (如类别 0) 的过拟合,表现为整体精度下降,且类别 4 的召回率下降更为明显。综合考虑整体指标与关键类别表现, γ 的较优取值范围约为 [0.5, 1.0]。在该区间内,自适应间隔能够兼顾对复杂类别的增强分离和对简单类别的稳定收敛,从而获得更优且更稳定的跨域少样本识别效果。

5.4 本章小结

本章面向跨塔型、少样本与弱标注条件下的螺栓松动检测问题,提出了一种基于半监督迁移学习的统一框架。在前一章所提出 MSEAM-Net 的基础上,本章重点引入基于 FADA 的域对抗对齐,以减小源/目标塔型间分布差异;采用双阈值伪标签策略,在保证可靠性的同时维持未标注样本覆盖度;所提出的结构复杂度自适应四元组损失,通过动态间隔与结构化采样增强目标域细粒度类别分离。实验结果表明,该框架在 5%–10% 极低标注条件下优势尤为明显,相比其他方法提升 5.57% 与 4.48%。并在各标注比例下保持稳定领先。混淆矩阵进一步揭示了相邻位置与远距离衰减带来的分类难点,为多传感器布设与工程化部署提供了依据。

第6章 总结与展望

6.1 工作总结

作为输电铁塔的关键连接部件，螺栓的紧固程度是杆塔结构稳固、线路安全运维的重要保障。野外服役环境复杂多变，各类外界荷载与侵蚀作用相互耦合，极易导致铁塔螺栓出现预紧力损失、松动缺失等问题，甚至诱发杆塔倾覆、电网停运等安全事故，带来巨大的经济损耗与负面社会影响。为此，研发贴合现场巡检需求、兼具经济性与实用性的螺栓松动检测定位技术，成为保障输电线路安全运行亟待解决的工程问题。现阶段常用的螺栓紧固状态检测技术，如预紧力直接检测、压电阻抗检测、超声检测等，应用于输电铁塔野外场景时仍存在诸多局限，例如设备购置与布设成本偏高、现场检测操作繁琐、施工安全管控要求严苛，且在大规模杆塔线路巡检中难以兼顾检测效率与判别可靠性。近年来，声纹识别技术逐步从语音信号处理领域拓展至工业装备故障诊断场景，凭借非侵入式检测、部署便捷、抗干扰能力较强等优势，为输电铁塔螺栓松动的现场巡检采样与离线状态识别提供了全新的技术思路。但受制于输电铁塔野外工况复杂、环境噪声干扰强烈且不同塔型的结构模态差异显著，松动声纹特征的稳健提取、模型在跨场景与跨塔型条件下的泛化性能提升，以及少样本约束下的模型高效训练，成为该技术落地应用亟需攻克的关键难题。针对以上工程痛点与技术挑战，本文围绕真实巡检场景下的铁塔螺栓声纹信号采集、特征增强处理以及少样本迁移学习方法展开深入研究，主要内容总结如下：

1. **多场景声纹数据采集与实证分析，形成可复现的采集方案与物理依据。**针对数据驱动检测方法受到数据质量的影响较大，本文与国家电网相关单位合作，深入三处真实输电铁塔场景采集多工况螺栓声纹样本，并围绕声激励装置与传感器布设方式开展对角与邻角方案对比。通过 t-SNE 可视化以及余弦距离、类内离散度等指标综合评估发现：不同布设方式对特征中心位置影响有限，但会影响类内稳定性。综合激励覆盖性、数据可比性与工程可复现性，本文在后续试验中统一采用对角布设作为标准方案。为增强松动引起的差异响应并提升有效信息含量，本文提出基于稳态时间窗与幅值统计特征的共振频率识别方法，即通过扫描不同激励频率并计算稳态段平均绝对幅值，确定铁塔在不同工况下的主导共振响应频段。实验表明，螺栓连接状态变化会引起峰值响应幅度与峰值频率差异，并在不同铁塔样本中观察到一定波动性。综合现场多样本结果，初步将 220 kV 系列输电铁塔的主导共振频率区间归纳为 300–500 Hz，该方法为后续声发激励频率选取与采集方案优化提供依据。此外，本文基于频谱图与语谱图对螺栓紧固和

松动状态差异进行了验证与解释：紧固状态下能量更集中于激励频率及相关分量，谱图整体更“干净”；松动状态在低频段出现更明显的谱线波动与能量扩散，语谱图中离散噪点与能量弥散区域增强，并随松动程度提高而增强。针对防盗螺栓等特殊结构样本，其差异表现弱于接地螺栓与杆件松动，提示后续方法设计需考虑螺栓类型与受力或接触机理差异。进一步地，空间位置对比显示随着松动位置与传感器距离增大，差异特征减弱，反映结构传播损耗与路径不确定性对可检测性的影响，为传感器布设与覆盖范围优化提供了工程启示。

2. **提出端到端多尺度能量注意力残差网络，提升复杂噪声背景下的识别鲁棒性。**针对野外环境噪声强、工况变化大的特点，本文在第三章差异性分析的基础上，提出端到端的多尺度能量注意力残差网络 MSEAM-Net。该方法在结构层面引入多尺度能量注意力模块：通过多尺度池化融合不同感受野信息，并以频带 RMS 能量统计驱动频带-通道联合注意力，实现关键频带自适应增强与背景成分抑制；在训练层面结合随机 SNR 噪声与 SpecAugment 遮挡构造扰动样本，引入一致性正则化约束同一样本在干净和扰动条件下表征一致，从而提升鲁棒性与跨场景泛化能力。在三座铁塔多工况数据集上的系统实验表明，MSEAM-Net 在 500 kV 系列复杂数据集上相较典型图像与声纹识别框架取得更优且更稳定的分类准确率与 Macro-F1；其中在 500 kV 酒杯塔数据上，本文方法的准确率与 F1 分数较 ResNet18 与 CRNN 分别提升 4.44%、6.9% 和 11.09%、12.5%。消融实验进一步验证了 MSEAM 与一致性正则化的互补贡献。频带干预与工程扩展实验说明稳态高能频带及其能量分布是故障识别的重要物理线索；模型能够支持多位置复合松动与不同扭矩等级判别，且垂直距离的诊断精度达到约 2m，并给出 M24 螺栓约 150($\pm$ 5) N·m 的松动/紧固有效区分阈值，为巡检场景下的现场采样与离线判别提供了方法与实证依据。
3. **构建结构复杂度自适应四元组损失与半监督迁移学习框架，提升少样本跨塔型适应能力。**针对实际工程应用中“跨塔型域偏移显著、标签获取成本高、无标签数据相对充足”的典型矛盾，本文在上一章提出的能量引导频带注意力编码器基础上，进一步提出结构复杂度自适应四元组损失 (SQL) 用于增强目标域细粒度类间可分性，半监督迁移学习框架 (SSTF) 用于联合利用源域标注数据与目标域少量标注和大量未标注数据，通过域对齐、自训练与结构约束协同优化，缓解跨塔型分布差异并提高低标注条件下的泛化能力。在两座 500 kV 系列输电铁塔的跨塔实验中，SSTF 在各标注比例下均保持稳定领先；尤其在 5%–10% 极低标注条件下优势最为明显，相比次优方法分别提升 5.57% 与 4.48%。上述结果验证了 SQL 对类结构建模

与 SSTF 对跨域适应的协同增益，为少样本条件下的巡检应用与工程试用提供了可行方案。

6.2 研究展望

尽管本文围绕输电铁塔螺栓松动的巡检采样与检测开展了系统研究，并在多场景数据与少样本跨塔迁移设置下验证了方法有效性，但仍存在提升与拓展的空间。结合工程应用需求与当前研究局限，后续工作可从以下几个方面展开：

1. **数据与场景覆盖的进一步扩展。**本文数据来源于有限数量的塔型与工况，尽管已覆盖不同电压等级与多种松动形态，但在季节变化（温湿度、雨雾）、风速风向、背景噪声类型以及不同地形（山地、平原、沿海）等因素下仍可能出现新的域偏移。后续可通过多区域、多季节的持续采集，构建更大规模、更高多样性的声纹数据库，并建立与气象、载荷等外部因素关联的元数据标注机制，以支撑跨环境泛化能力评估与提升。
2. **松动位置与程度的精细化定位与量化评估。**本文以多类别状态识别为主，已具备一定的邻近类别分离能力，但在实际应用中，运维更关注“松动发生在何处、松动程度如何”。后续可引入多实例学习、弱监督定位或图结构建模等方法，将节点拓扑与传播衰减规律纳入模型，进一步实现松动位置的细粒度定位与松动程度的连续量化，从而提供更可靠的检修决策依据。
3. **多传感器与多模态信息融合的螺栓松动感知方法。**本文主要基于声纹信号开展螺栓松动检测研究，虽在复杂工况下取得了较为稳定的识别效果，但单一模态在野外复杂场景中仍可能存在信息受限的问题。后续研究可考虑融合振动、加速度、图像或红外等多源传感信息，构建多模态协同感知框架，通过特征级或决策级融合充分挖掘不同模态在时域、频域及空间域上的互补优势。
4. **面向离线应用的模型轻量化与流程工程化。**本文方法当前主要用于离线训练与离线推理的巡检判别场景。为便于现场端使用，后续可从模型压缩、知识蒸馏、算子优化等角度开展轻量化研究，形成“采样-预处理-推理-报告生成”的一体化工具链；同时建立更完善的阈值校准与风险分级机制，使输出结果更贴合运维工作流，提升工程可用性。

总体而言，本文验证了声纹感知与深度学习方法在输电铁塔螺栓松动巡检中的可行性与潜力。未来在数据体系扩展、布设与激励优化、定位量化、鲁棒迁移与工程化落地等方面进一步深入研究，有望持续提升复杂环境下螺栓松动检测的准确性与实用价值。

参 考 文 献

- [1] 中华人民共和国国家发展和改革委员会. 加强能源产供储销体系建设夯实能源安全供应保障基础[EB/OL]. (2025-12-01). https://www.ndrc.gov.cn/xwdt/ztzl/jxpl/202512/t20251204_1402219.html.
- [2] LIANG C, SUN F P, ROGERS C A. Coupled electro-mechanical analysis of adaptive material systems—determination of the actuator power consumption and system energy transfer[J]. *Journal of Intelligent Material Systems and Structures*, 1994, 5(1): 12-20.
- [3] SUN F P, CHAUDHRY Z, LIANG C, et al. Truss structure integrity identification using PZT sensor-actuator[J]. *Journal of Intelligent Material Systems and Structures*, 1995, 6(1): 134-139.
- [4] NGUYEN T T, TA Q B, HO D D, et al. A method for automated bolt-loosening monitoring and assessment using impedance technique and deep learning[J]. *Developments in the Built Environment*, 2023, 14: 100122.
- [5] LU R, SHEN Y. Entire loosening stage monitoring of bolted joints via nonlinear electro-mechanical impedance spectroscopy[J]. *Structural Health Monitoring*, 2026, 25(1): 76-92.
- [6] JIANG T, WU Q, WANG L, et al. Monitoring of bolt looseness-induced damage in steel truss arch structure using piezoceramic transducers[J]. *IEEE Sensors Journal*, 2018, 18(16): 6677-6685.
- [7] CHEN D, SHEN Z, FU R, et al. Coda wave interferometry-based very early stage bolt looseness monitoring using a single piezoceramic transducer[J]. *Smart Materials and Structures*, 2022, 31(3): 035030.
- [8] TAN B, WANG T, FANG Q, et al. Bolt looseness monitoring using dynamic mode decomposition with piezoelectric active sensing[J]. *Measurement*, 2024, 238: 115204.
- [9] CHEN X, HAN W, ZHANG Z. Loosening state monitoring and identification of multi-bolted flange joints based on nonlinear wave energy transmission[J]. *Mechanical Systems and Signal Processing*, 2025, 224: 112114.
- [10] WANG L, YI S, SUN X, et al. Quantitative evaluation of bolt pre-load using coda wave interferometry and nonlinear coda wave interferometry: a comparative study[J]. *Journal of Civil Structural Health Monitoring*, 2025, 15(2): 563-573.
- [11] SHELGAONKAR T, NARAYANAN A, BASAPPA U. A review on the impact of thermal effects on electro-mechanical impedance technique-based structural health monitoring[J]. *Journal of Nondestructive Evaluation*, 2025, 44: 76.
- [12] HUANG J, LIU J, GONG H, et al. A comprehensive review of loosening detection methods

- for threaded fasteners[J]. *Mechanical Systems and Signal Processing*, 2022, 168: 108652.
- [13] TOLA K D, LEE C, PARK J, et al. Bolt looseness detection based on ultrasonic wavefield energy analysis using an Nd:YAG pulsed laser scanning system[J]. *Structural Control and Health Monitoring*, 2020, 27(9): e2590.
- [14] LUO L, FAN J, MA J, et al. A nonlinear ultrasonic wave mixing method for looseness detection of bolted joints[J]. *Ultrasonics*, 2024, 142: 107402.
- [15] ZHOU Y, YUAN C, SUN X, et al. Monitoring the looseness of a bolt through laser ultrasonic [J]. *Smart Materials and Structures*, 2020, 29(11): 115022.
- [16] 万书亭, 贾东亮, 孙瑞滨, 等. 基于振动信号频域离散特征的螺栓完全松动检测方法研究[J]. *中国工程机械学报*, 2024, 22(02): 247-251.
- [17] 谢锋云, 符羽, 刘昆, 等. 基于多域特征的螺栓松动检测方法研究[J]. *电子测量与仪器学报*, 2021, 35(04): 109-117.
- [18] ZHOU Y, WANG S, ZHOU M, et al. Percussion-based bolt looseness identification using vibration-guided sound reconstruction[J]. *Structural Control and Health Monitoring*, 2022, 29(2): e2876.
- [19] QIN X, PENG C, ZHAO G, et al. Full life-cycle monitoring and earlier warning for bolt joint loosening using modified vibro-acoustic modulation[J]. *Mechanical Systems and Signal Processing*, 2022, 162: 108054.
- [20] LI J, WEN B, ZHANG Z, et al. Interface behaviors of elastic-plastic waves and its impact on uncertainties in vibro-acoustic modulation (VAM) for structural health monitoring (SHM) of bolt loosening[J]. *Mechanical Systems and Signal Processing*, 2024, 212: 111268.
- [21] FAN J, CHEN W, CHEN D, et al. Vibro-acoustic modulation-based bolt looseness detection method using frequency sweep probe waves[J]. *Nondestructive Testing and Evaluation*, 2024, 39(5): 665-683.
- [22] WANG X, YANG H, QI J, et al. Research on bolt loosening detection based on fractional fourier transform and vibro-acoustic modulation method[J]. *Processes*, 2026, 14(7): 1130.
- [23] LI J, HE Y, LI Q, et al. Artificial intelligence (AI)-based evaluation of bolt loosening using vibro-acoustic modulation (VAM) features from a combination of simulation and experiments [J]. *Applied Sciences*, 2022, 12(24): 12920.
- [24] 张天龙, 闫子权, 乔廷强, 等. 基于视觉测振和卷积神经网络的螺栓松动检测方法[J]. *强度与环境*, 2022, 49(02): 48-56.
- [25] PAN Y, MA Y, DONG Y, et al. A vision-based monitoring method for the looseness of high-strength bolt[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-14.
- [26] HUYNH T C, PARK J H, JUNG H J, et al. Quasi-autonomous bolt-loosening detection method using vision-based deep learning and image processing[J]. *Automation in Construction*, 2019,

- 105: 102844.
- [27] HUYNH T C. Vision-based autonomous bolt-looseness detection method for splice connections: Design, lab-scale evaluation, and field application[J]. *Automation in Construction*, 2021, 124: 103591.
- [28] LEI W, YUAN F, GUO J, et al. Vision-based real-time bolt loosening detection by identifying anti-loosening lines[J]. *Sensors*, 2024, 24(20): 6747.
- [29] WU T, SHANG K, DAI W, et al. High-resolution cross-scale transformer: A deep learning model for bolt loosening detection based on monocular vision measurement[J]. *Engineering Applications of Artificial Intelligence*, 2024, 133: 108574.
- [30] KONG Q, ZHU J, HO S C M, et al. Tapping and listening: A new approach to bolt looseness monitoring[J]. *Smart Materials and Structures*, 2018, 27(7): 07LT02.
- [31] ZHANG Y, ZHAO X, SUN X, et al. Bolt loosening detection based on audio classification[J]. *Advances in Structural Engineering*, 2019, 22(13): 2882-2891.
- [32] QIN Y, YUAN L. Intelligent bolt loosening detection in transmission towers using acoustic signature analysis and machine learning[J]. *Processes*, 2025, 13(4): 1111.
- [33] WANG S, ZHOU Y, KONG Q. A force-adaptive percussion method for bolt looseness assessment[J]. *Journal of Civil Structural Health Monitoring*, 2024, 14: 831-841.
- [34] CHENG L, ZHANG Z, LACIDOGNA G, et al. Sound sensing: Generative and discriminant model-based approaches to bolt loosening detection[J]. *Sensors*, 2024, 24(19): 6447.
- [35] CHEN X, ZHOU W, YANG J, et al. A percussion method with attention-based sequence modeling for bolt looseness detection and adaptive anomaly recognition[J]. *Measurement*, 2025, 256: 118394.
- [36] CHEN Y, CHEN H, WEI J, et al. Automatic detection of loose bolts in pipeline structures based on anti-noise mel cepstrum differential features and deep learning[J]. *Journal of Civil Structural Health Monitoring*, 2026, 16: 7.
- [37] BAO Z, LIU C, YANG H, et al. From theory to industry: A survey of deep learning-enabled bearing fault diagnosis in complex environments[J]. *Engineering Applications of Artificial Intelligence*, 2026, 163: 113068.
- [38] LI J, ZHENG C, WU Z, et al. Review of fault diagnosis for rotating machinery: Prior knowledge integration in data-driven methods benefits model interpretability and generalizability[J]. *Mechanical Systems and Signal Processing*, 2026, 242: 113623.
- [39] WANG F, SONG G. A novel percussion-based method for multi-bolt looseness detection using one-dimensional memory augmented convolutional long short-term memory networks [J]. *Mechanical Systems and Signal Processing*, 2021, 161: 107955.
- [40] YUAN C, WANG S, QI Y, et al. Automated structural bolt looseness detection using deep

- learning-based prediction model[J]. *Structural Control and Health Monitoring*, 2022, 29(3): e2899.
- [41] LI F, WANG L, WANG D, et al. An adaptive multiscale fully convolutional network for bearing fault diagnosis under noisy environments[J]. *Measurement*, 2023, 216: 112993.
- [42] HU B, LIU J, XU Y. A novel multi-scale convolutional neural network incorporating multiple attention mechanisms for bearing fault diagnosis[J]. *Measurement*, 2025, 242(Part B): 115927.
- [43] SHEN J, ZHAO D, LIU S, et al. Multiscale attention feature fusion network for rolling bearing fault diagnosis under variable speed conditions[J]. *Signal, Image and Video Processing*, 2024, 18(Suppl 1): 523-535.
- [44] JIN Z, HU X, WANG H, et al. Rolling bearing fault diagnosis model based on multi-scale depthwise separable convolutional neural network integrated with spatial attention mechanism [J]. *Sensors*, 2025, 25(13): 4064.
- [45] JIA L, CHOW T W S, WANG Y, et al. Multiscale residual attention convolutional neural network for bearing fault diagnosis[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-13.
- [46] TANG Y, REN Y P, WU J, et al. Deep learning-based bearing fault diagnosis using a trusted multi-scale quadratic attention-embedded convolutional neural network[J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73: 1-15.
- [47] LIU X, LU J, LI Z. Multiscale fusion attention convolutional neural network for fault diagnosis of aero-engine rolling bearing[J]. *IEEE Sensors Journal*, 2023, 23(17): 19918-19934.
- [48] HU B, LIU J, XU Y. A novel multi-scale convolutional neural network incorporating multiple attention mechanisms for bearing fault diagnosis[J]. *Measurement*, 2025, 242(Part B): 115927.
- [49] SHEN J, ZHAO D, LIU S, et al. Multiscale attention feature fusion network for rolling bearing fault diagnosis under variable speed conditions[J]. *Signal, Image and Video Processing*, 2024, 18(Suppl 1): 523-535.
- [50] MIAO Z, FENG W, LONG Z, et al. Motor fault diagnosis using attention-based multisensor feature fusion[J]. *Energies*, 2024, 17(16): 4053.
- [51] KONG Q, XIONG Q, XIONG H, et al. Semi-supervised networks integrated with autoencoder and pseudo-labels propagation for structural condition assessment[J]. *Measurement*, 2023, 214: 112779.
- [52] ZHAO B, CHENG C, ZHAO S, et al. Hybrid semi-supervised learning for rotating machinery fault diagnosis based on grouped pseudo labeling and consistency regularization[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-12.
- [53] YU X, WANG Y, LIANG Z, et al. An adaptive domain adaptation method for rolling bearings' fault diagnosis fusing deep convolution and self-attention networks[J]. *IEEE Transactions on*

- Instrumentation and Measurement, 2023, 72: 1-14.
- [54] YANG L, YANG Z, LI F, et al. Single-layer feedforward neural networks with dynamic width for domain adaptation[J]. Science China Information Sciences, 2025, 68(8): 1-15.
- [55] CHELIMILLA N, CHINTHAPENTA V, KORLA S. Addressing data scarcity using audio signal augmentation and deep learning for bolt looseness prediction[J]. Smart Materials and Structures, 2024, 33(8): 085012.
- [56] LIU P, WANG X, WANG Y, et al. Research on percussion-based bolt looseness monitoring under noise interference and insufficient samples[J]. Mechanical Systems and Signal Processing, 2024, 208: 111013.
- [57] LIU Z, HUANG X, ZHAO L, et al. Research on online monitoring technology for transmission tower bolt looseness[J]. Measurement, 2023, 223: 113703.
- [58] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: A unified embedding for face recognition and clustering[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015: 815-823.
- [59] CHEN W, CHEN X, ZHANG J, et al. Beyond triplet loss: A deep quadruplet network for person re-identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017: 403-412.
- [60] QU F, LIU J, LIU X, et al. A multi-fault detection method with improved triplet loss based on hard sample mining[J]. IEEE Transactions on Sustainable Energy, 2020, 12(1): 127-137.
- [61] GUI X, ZHANG J, TANG J, et al. A quadruplet deep metric learning model for imbalanced time-series fault diagnosis[J]. Knowledge-Based Systems, 2022, 238: 107932.
- [62] BOLTON T, BASS J, GABER T, et al. A novel triplet loss architecture with visual explanation for detecting the unwanted rotation of bolts in safety-critical environments[J]. Engineering Applications of Artificial Intelligence, 2025, 156: 111097.
- [63] 鲁光涛, 刘哲, 李友荣, 等. 扫频非线性超声调制的螺栓联接状态识别[J]. 机械设计与制造, 2026(02): 223-226+232.
- [64] WANG F, SONG G. Bolt early looseness monitoring using modified vibro-acoustic modulation by time-reversal[J]. Mechanical Systems and Signal Processing, 2019, 130: 349-360.
- [65] DAVIS S B, MERMELSTEIN P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1980, 28(4): 357-366.
- [66] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018: 7132-7141.
- [67] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.

- [68] LEE D H, et al. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks[C]//Proceedings of the Workshop on Challenges in Representation Learning, International Conference on Machine Learning (ICML): volume 3. 2013: 896.
- [69] SOHN K, BERTHELOT D, CARLINI N, et al. FixMatch: Simplifying semi-supervised learning with consistency and confidence[C]//Advances in Neural Information Processing Systems (NeurIPS): volume 33. 2020: 596-608.
- [70] BERTHELOT D, CARLINI N, GOODFELLOW I, et al. MixMatch: A holistic approach to semi-supervised learning[C]//Advances in Neural Information Processing Systems (NeurIPS): volume 32. 2019.
- [71] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks [J]. Journal of Machine Learning Research, 2016, 17(59): 1-35.
- [72] MOTIIAN S, JONES Q, IRANMANESH S, et al. Few-shot adversarial domain adaptation [C]//Advances in Neural Information Processing Systems (NeurIPS): volume 30. 2017.
- [73] 何生成. 铁塔螺栓松动检测技术研究[D]. 北京交通大学, 2018.
- [74] WANG F, HO S C M, SONG G. Modeling and analysis of an impact-acoustic method for bolt looseness identification[J]. Mechanical Systems and Signal Processing, 2019, 133: 106249.
- [75] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks [J]. Journal of Machine Learning Research, 2016, 17(59): 1-35.
- [76] YANG L, WANG Y, GAO M, et al. Deep co-training with task decomposition for semi-supervised domain adaptation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2021: 8906-8916.
- [77] GRANDVALET Y, BENGIO Y. Semi-supervised learning by entropy minimization[C]//Advances in Neural Information Processing Systems (NeurIPS): volume 17. 2004.
- [78] SAITO K, KIM D, SCLAROFF S, et al. Semi-supervised domain adaptation via mini-max entropy[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 2019: 8050-8058.
- [79] LI J, LI G, SHI Y, et al. Cross-domain adaptive clustering for semi-supervised domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021: 2505-2514.
- [80] KIM T, KIM C. Attract, perturb, and explore: Learning a feature alignment network for semi-supervised domain adaptation[C]//European Conference on Computer Vision (ECCV). 2020: 591-607.

致 谢

光阴荏苒，寒暑易节，转眼间在科大的三载求学时光已接近尾声。伫立在研究生阶段的终点回首望去，这不仅是一段知识积累的过程，更是一场关于意志、心态与认知的深刻蜕变。

首先，诚挚感谢我的导师赵云波教授。赵老师深厚的学术造诣、敏锐的科研直觉以及严谨治学的态度，为我树立了标杆。在论文选题、实验方案设计及稿件反复打磨的过程中，赵老师总是能一针见血地指出症结所在，他不仅教会了我如何“做学问”，更教会了我如何以谦逊执着的姿态面对未知的挑战。

由衷感谢吕文君老师和李泽瑞老师。特别感谢吕老师，作为我科研道路上的引路人，您在我最迷茫的“萌新”阶段给予了耐心的指导与包容。从最初对研究课题的一知半解，到如今能够独立梳理逻辑、解决科研难题，每一点进步都离不开您的倾心点拨与支持。李老师在实验设备调试与项目推进中的无私帮助，也令我受益匪浅。

感谢实验室同窗们的陪伴。科学研究从来不是孤岛，感谢大家在每周组会上的思想碰撞，以及在无数个科研攻关日夜里的相互勉励。那些在信智楼 10 层共同奋斗的瞬间，是我科研生涯中最温暖的底色。

深切感谢我的父母。感谢你们多年来无条件的支持与理解，你们是我避风的港湾，让我能够心无旁骛地在学术世界里探索。也要感谢我的朋友们，在我情绪波动、科研受阻时提供的倾听与陪伴，是你们的关心让我始终保持积极向上的心态。

最后，我想感谢那个曾经在实验室挑灯夜战、不言放弃的自己。感谢自己在无数次实验失败后的复盘，感谢自己在深夜科大校园里追寻真理的身影。

路漫漫其修远兮，愿这段科大时光永远化作我前行路上的星光，指引着我未来的方向。

在读期间发表的学术论文与取得的研究成果

已发表论文

1. Fei Xie, **Chuangchuang He**, Ji Chang, Chen Hu, Xiaofei Wang, Yu Kang, Wenjun Lv. Physics-Informed Semi-Supervised Transmission Towers Bolt Detection and Localization Using Acoustic Wave Generator [J]. IEEE Transactions on Instrumentation and Measurement, 2026, 75: 1-15.

已授权发明专利

1. 吕文君, **何创创**, 康宇. 一种基于声纹信号的设备故障检测方法 [P]. 中国发明专利, 授权公告号: CN118197297B, 授权公告日: 2024-11-22.
2. 赵云波, **何创创**, 康宇, 吕文君. 一种无监督域自适应的设备故障诊断方法 [P]. 中国发明专利, 授权公告号: CN106152260B, 授权公告日: 2025-08-12.

参加科研项目

1. 国网安徽省电力有限公司经济技术研究院科技项目, “基于声纹特征提取算法的输电铁塔螺栓紧固状态检测及定位技术研究”, 项目编号: B6120923001G, 已结项, 主要参与人。

待发表论文

1. **Chuangchuang He**, Fei Xie, Wenjun Lv, Ji Chang, Zerui Li, Yun-Bo Zhao. Semi-supervised cross-tower acoustic diagnosis for bolt-loosening localization in power transmission towers [J]. Journal of Civil Structural Health Monitoring, 2026. (Under Review)