

冰裂声信号的多分量特征融合与 半监督学习检测方法*

张宇翔^{1,2} 卞忻然^{1,2,3,4} 李 理^{1,2†} 刘晓颖⁵ 金春燕⁵ 高家辉^{1,2,3}

(1 极地海洋声学与技术应用教育部重点实验室(哈尔滨工程大学) 哈尔滨 150001)

(2 哈尔滨工程大学 水声工程学院 哈尔滨 150001)

(3 哈尔滨工程大学 水声技术全国重点实验室 哈尔滨 150001)

(4 哈尔滨工程大学 三亚南海创新发展基地 三亚 572000)

(5 中国船舶集团有限公司 系统工程研究院 北京 100094)

2024 年 12 月 31 日收到

2025 年 3 月 24 日定稿

摘要 冰裂声信号作为极地冰盖动力学演化的“声学特征指纹”具有重要监测价值,而传统检测方法存在低信噪比下灵敏度不足以及对标注数据依赖性过高等问题,为此提出了一种基于多分量特征融合的半监督学习冰裂声信号检测框架,旨在提升低标注数据条件下冰裂声信号的自动化检测能力。首先,提出并对比混合波形图和堆叠时频图两种三维声场特征融合方案,优化三维弹性声场信息的有效利用。随后构建教师-学生模型的双分支迭代优化架构,引入半监督学习伪标签策略,利用未标注数据增强模型泛化能力以显著降低人工标注数据依赖。通过开展高纬度冰区试验,以 YOLOX 为基准模型验证所提方法的有效性:相同人工标注数据量下,检测性能较基准模型提升 26.2%;在标注数据减少 50% 的条件下,所提方法检测性能接近全标注模型(达成率 94.5%)。ROC 曲线分析进一步证实,堆叠时频图特征融合方案在冰裂声信号检测任务中具有明显优势。

关键词 冰裂信号, 半监督学习, 伪标签, 目标检测, 特征融合

PACS: 43.35, 43.60

DOI: 10.12395/0371-0025.2024425

CSTR: 32049.14.11-2065.2024425

Multi-component feature fusion and semi-supervised learning method for ice cracking signal detection

ZHANG Yuxiang^{1,2} BIAN Xinran^{1,2,3,4} LI Li^{1,2†} LIU Xiaoying⁵ JIN Chunyan⁵ GAO Jiahui^{1,2,3}

(1 Key Laboratory for Polar Acoustics and Application of Ministry of Education (Harbin Engineering University) Harbin 150001)

(2 College of Underwater Acoustic Engineering, Harbin Engineering University Harbin 150001)

(3 National Key Laboratory of Underwater Acoustic Technology, Harbin Engineering University Harbin 150001)

(4 Sanya Nanhai Innovation and Development Base of Harbin Engineering University Sanya 572000)

(5 Systems Engineering Research Institute, CSSC Beijing 100094)

Received Dec. 31, 2024

Revised Mar. 24, 2025

Abstract Ice cracking signals, which serve as acoustic fingerprints of the dynamic evolution of polar ice sheets, possess significant monitoring values, while traditional detection methods face challenges such as insufficient sensitivity in low signal-to-noise ratio environments and excessive dependence on annotated data. This paper proposes a semi-supervised learning framework based on multi-component feature fusion for ice cracking signal detection, aimed at enhancing automated detection capabilities under limited annotation conditions. First, two three-dimensional acoustic field feature fusion schemes, i.e. mixed waveform diagrams and stacked time-frequency representations, are proposed and compared to optimize the utilization of three-dimensional elastic acoustic field information. Then, a dual-branch iterative optimization architecture with teacher-student models is constructed, where a semi-

* 国家自然科学基金项目(62125104, 52171334, 62201166)和教育部春晖计划项目(HZKY20220309)资助

† 通讯作者: 李理, Li.li@hrbeu.edu.cn

supervised learning pseudo-labeling strategy is introduced to leverage unannotated data, thereby enhancing the model generalization capability while significantly reducing dependency on manual annotations. Through experiments conducted in high-latitude ice regions using YOLOX as a benchmark model, the effectiveness of the proposed technique is validated. With the same amount of manually annotated data, the detection performance is improved by 26.2% compared to the benchmark; when the annotated data is reduced by 50%, the proposed method maintains a performance comparable to fully-supervised models (achievement rate of 94.5%). The ROC curve analysis further confirms that the stacked time-frequency feature fusion scheme exhibits distinct advantages in ice cracking signal detection tasks.

Keywords Ice crack signal, Semi-supervised learning, Pseudo label, Target detection, Feature fusion

引言

极地冰盖作为地球气候系统的关键组分,对全球海平面波动、大气环流和生态系统平衡具有直接影响^[1]。冰裂声学信号是极地冰盖对环境扰动响应的声学特征指纹。外部环境变化(如温度骤降、潮汐应力耦合及风场驱动导致的冰体位移)会导致冰内应力局部累积,一旦超越临界阈值,冰体结构通过不可逆断裂过程将累积能量以声能量的形式释放,形成冰裂事件^[2]。通过对冰裂事件产生的声信号进行观测,获取声源激发机制信息(塑性变形、脆性断裂或动力碰撞等)和作为传播介质的冰层瞬态响应特征信息(声参数及频散特性等),可揭示冰盖在多时间尺度上的演化过程^[3-5]。因此,开展冰裂声信号的系统性研究,是后续冰层稳定性评估、震源定位、海冰动力学预测的先验条件,并可进一步服务极地航行和极地环境监测等需求^[6]。

冰裂声学信号与地震信号的传播介质接近,通过在冰面布设高精度检波器或地震仪阵列^[7]是被动观测冰裂事件的主要方式。近三十年来,随着极地观测网络规模的扩大和监测密度的提升,冰声学记录数据采集量呈指数级增长态势,单次观测实验产生的数据量往往达到数十至上百小时量级。面对海量数据中蕴含的丰富信息,人工筛选在效率和精度上均难以满足研究需求^[8]。基于上述背景,开发高效可靠的自动化冰裂事件检测算法成为当前极地冰声学亟待解决的问题。

从现有研究来看,当前冰裂声学事件检测方法主要包括基于短期平均/长期平均算法(STA/LTA)^[9],通过观察特定频段的能量分布进行冰裂事件识别的频谱检查法,通过测量冰层中传播波速度和衰减特性推断冰裂事件的弥散测量法,以及利用互相关函数计算信号之间相似的模板匹配检测法^[10,11]。然而,这些方法在面对真实观测记录时存在不同程度的应用局限。STA/LTA算法作为一种基于阈值的触发检

测算法^[12],在处理远距离传播或地震级冰裂导致的低信噪比事件时灵敏度不足^[13],且其计算复杂度随数据规模呈超线性增长^[14]。频谱检查法被证实在检测冰裂事件中比STA/LTA方法更有效,但对低信噪比事件的敏感度较低,且通常需要额外的手动检查或信号过滤^[1]。弥散测量法通过识别表面波的弥散性推断冰裂事件,但依赖于冰层速度梯度假设且实施复杂度过高^[10]。现有模板匹配算法可应用于微小信号检测^[15,16],然而在该场景下对先验模板具有高度依赖性。冰声信号作为一种天然信号,其形成机理和传播过程具有随机性,这种随机性会对时域波形产生复杂调制,制约了模板匹配方法的推广应用^[17]。

近年来,基于深度神经网络学习的信号检测及分类算法已被广泛应用。例如在地震学的信号检测中,使用卷积神经网络搜索滤波器最优参数,以区分地震和地震噪声^[18]。区别于上述传统检测算法,神经网络模型从大型训练数据集中自动学习并提取事件信号的共同特征,针对具体任务可以拥有适应性更强的检测能力,但在冰裂信号检测领域应用深度学习网络仍面临挑战。虽然冰裂信号与地震信号在物理特性上相似,但深度学习在地震领域实现了高精度检测结果的前提是地震学及地球物理领域在世界范围内累积多年获得的大量标注数据,以及相对冰裂事件,地震事件在时、空间范围内的相对稀疏性。而海冰声学与冰裂声学信号研究尚属新兴学科,相比于地震研究领域,缺少系统性的机器学习应用经验且全球范围内尚未建立公开标注数据集。因此,深度学习在冰裂事件检测领域的应用仍然面临许多现实问题。

针对上述问题,本研究提出了一种基于多分量特征融合的半监督学习冰裂声信号检测方法。首先,利用冰裂声信号多模态、多分量信号特点分别从时域和时频域构建其对应声学信号的可视化呈现方式,将不同分量的特征进行有效整合,实现与后期特征提取环节的紧密融合,从而提升神经网络对于冰裂事件多维度信息的综合感知能力。其次,得益于半

监督学习方式的框架设计,模型在训练过程中可以降低 50% 的数据标注和人工依赖。此外,引入伪标签策略,使模型在自主学习过程中学会判断特征信息并生成伪标签,从而增强模型对复杂冰裂信号的适应性并保持高精度检测能力。

1 典型冰裂信号特征

海冰声传播特性受弹性波导影响显著^[5],前期理论研究与试验观测均指出,海冰中能够实现有效长程传播的低频段弹性波类型有 3 种:板间纵波(S_0')、水平剪切波(SH)和泄漏模式流固耦合界面波(QS)^[19-21]。上述 3 种类型弹性波的传播速度、偏振特性和频散特性均存在较大差别: S_0' 波场传播速度最快(约 3800 m/s)且频散特性微弱,在水平平面内呈近线性偏振;SH 波场传播速度稍慢(1800 m/s)且无频散,呈现水平线性偏振且主偏振方向与 S_0' 波场在水平面内相互垂直;QS 模式弹性波场传播速度最慢且在低频段频散现象严重(<1500 m/s),整体在垂直平面内呈现椭圆偏振。

前期研究指出,温度快速下降导致的热应力(Thermal Traction)是冰裂事件的主要形成机理,是局部弹性势能快速释放形成的瞬态声能量激发。其声信号特点与地震、敲击等脉冲震源激发信号高度类似,主要呈现宽频、瞬时性。结合前述海冰声传播特

点与激发条件,冰裂声能量在经由冰中传播过程后逐步汇聚至 SH、 S_0' 和 QS 三个模态,其传播速度的差异进一步导致其在时间域上的逐渐分化^[22]。海冰声信号主要通过三分量振动传感器采集,它将局部振动信息投影到 3 个正交的检波器分量之上实现冰声波场的三维记录。

图 1(a) 为本研究对应实测单次典型冰裂信号的三分量波场记录。对比北极实测冰声信号^[23]及相关理论与数值仿真分析可知,图中 3 个波包对应上述 3 种模态的声场。 S_0' 和 SH 模态信号呈现短脉冲特点,能量在时域分布紧凑。同时由于其能量主要分布在水平平面,因此 Z 分量(垂直分量)记录的信号中上述模态几乎不可见。QS 模态导波显著的频散特性导致其脉冲波包的时域展宽,高频波分量传播速度更快因此更先到达,但是其整体传播速度显著低于其他两个模态导波。QS 模态呈现椭圆二维偏振,其能量主要分布在垂直分量上,但在对应的水平分量上也有所体现。将上述三分量信号经短时傅里叶变换(STFT)以直观展示声能量的时频动态特性,如图 1(b) 所示,其整体上与前述理论分析一致,但是不同模态声能量在三分量时频空间上的分布差异更为突出,SH 和 S_0' 模态主要分布在水平分量上,而 QS 模态主要分布在垂直分量上。

需要说明的是,以上内容描述的信号特点与对

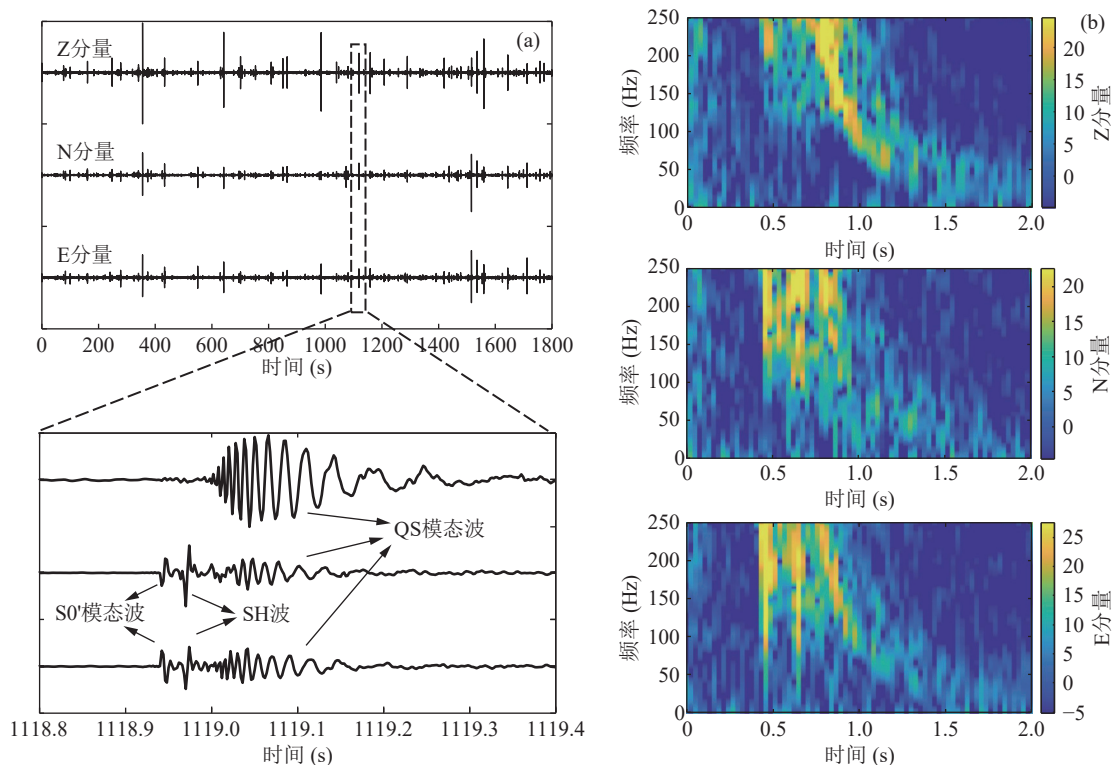


图1 冰裂信号三分量特征图 (a) 单次典型冰裂信号三分量波场记录; (b) 单次典型冰裂信号三分量时频图

应理论研究结果相一致,但并不能完整代表所有冰源噪声。图 1 描述的波场特点需要冰裂所释放的声能量在冰中经过充分传播才可以形成,不能代表冰体在挤压、摩擦等作用下形成的其他类别冰源噪声。本研究聚焦于上述信号,因为它同时体现了声源与传播介质特性,未来研究中可用于表征与监测冰层特性。对于冰裂激发但未经充分传播的冰裂信号,由于其信号时频特性缺乏对传播介质的描述能力,因此暂时不纳入本研究讨论范围。

2 冰裂声信号检测框架

本文提出的冰裂声信号检测方法由两阶段的设计框架构成。图 2(a) 为第一阶段的融合特征输入模块,在对多分量冰裂声信号进行预处理后,整合时域、时频域以保留冰裂声信号多模态、多维度的特征丰富性,为后续检测任务提供多样化的特征表示。图 2(b) 是第二阶段框架的核心处理单元,即半监督学习伪标签策略模块,通过教师模型和学生模型间的参数共享和知识传递构成闭环学习系统,扩充训练样本空间以提升最终输出的目标检测模型的泛化应用能力。该系统中的教师模型和学生模型均采用基于深度残差网络的多尺度特征目标检测模型,如图 2(c) 所示。两个模型初始状态相同,经过训练后,将完成参数优化的教师模型作为目标检测模型输出。

2.1 半监督学习伪标签策略模块

如图 2(b) 所示,本研究采用 Efficient Teacher 框

架应用伪标签策略,以解决大量冰声被动观测数据与有限冰裂数据标注能力之间的矛盾,该策略基于双分支迭代优化架构设计^[24-26],构成本检测模型设计概念的核心。在 Efficient Teacher 框架中,教师模型和学生模型共享相同的网络架构但承担不同的任务角色,教师模型与学生模型之间建立知识传递过程可分为以下几个关键步骤:

首先,利用标注数据对学生模型进行监督训练,通过最小化交叉熵损失函数 $LOSS_{sup}$ 以优化其参数。而后在学生模型与教师模型间建立知识传递,为确保其稳定性,采用指数滑动平均 (Exponential Moving Average, EMA) 机制更新教师模型的参数。具体而言,在时间步 t ,教师模型的参数 $\theta_{teacher}$ 通过如下动态更新方程进行优化:

$$\theta_{teacher}^t = \alpha \cdot \theta_{teacher}^{t-1} + (1 - \alpha) \cdot \theta_{student}^t, \quad (1)$$

其中, θ 表示 t 时刻的网络参数, α 是平滑系数,通常取值在 0.99 至 0.999 之间。

随后进入模型迭代优化阶段,应用双流结构将未标注数据同时输入两个模型。其中,学生模型产生预测结果用于计算无监督损失;教师模型生成伪标签,应用置信度阈值筛选机制保证伪标签质量,最终得到无监督损失函数 $LOSS_{unsup}$ 。基于此前步骤所得监督训练与无监督训练损失函数,计算得到总损失函数:

$$LOSS = LOSS_{sup} + \lambda \cdot LOSS_{unsup}, \quad (2)$$

其中, λ 为自适应权重因子,用于动态平衡监督学习

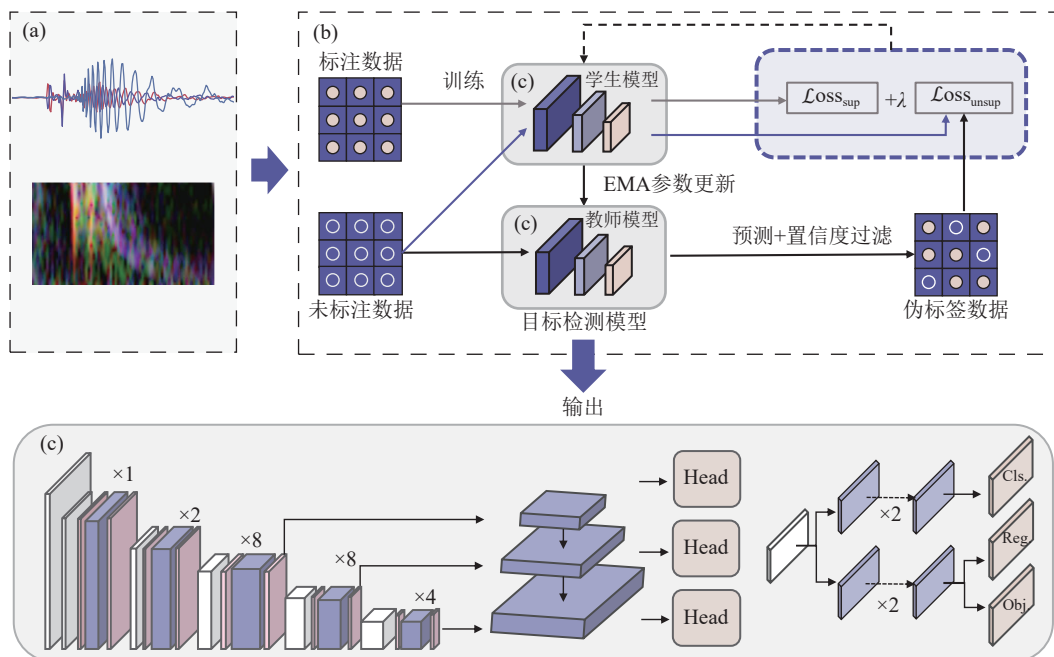


图 2 基于多分量特征融合的半监督学习检测框架 (a) 融合特征输入模块; (b) 半监督学习伪标签策略模块; (c) 目标检测模块

与无监督学习之间的超参数。

在参数更新阶段, 该框架通过随机梯度下降算法优化学生模型参数, 并使用 EMA 机制更新教师模型。这种双向迭代的优化机制使得整个训练过程形成闭环。即教师模型通过生成高质量伪标签来指导学生模型学习, 而学生模型的参数经 EMA 平滑后又反过来更新教师模型。通过这种良性互补的训练方式, 模型能够有效利用未标注数据中蕴含的特征信息, 在显著降低标注依赖的同时保持检测性能。

2.2 目标检测模型

研究使用的目标检测模型 (基于深度残差网络的多尺度特征目标检测模型) 通过级联卷积层组实现特征层次化表达, 再利用特征金字塔网络 (FPN) 融合多尺度特征, 有效地整合不同感受野内的上下文信息, 促进检测头网络通过密集预测策略实现目标的精确定位。如图 2(c) 所示, 该架构包括主干特征提取网络、多尺度特征融合网络和解耦检测头三部分。其中, 主干网络采用深度残差结构设计, 通过残差连接有效缓解了深层网络的梯度消失问题。网络包含五个阶段的下采样操作, 在每个阶段结合 1×1 与 3×3 卷积核的复合应用及残差连接进行特征提取, 保证提取有效性的同时缓解了梯度消失问题。而后引用局部注意力机制, 进一步增强模型对关键特征的捕捉能力^[27], 以提高冰裂事件检测精度。

特征融合层面, 模型引入了一种自适应特征融合机制特征金字塔网络。模块通过自顶向下的路径增强横向连接机制, 实现了多尺度特征的有效融合。高层特征通过上采样逐步传递到低层特征图, 同时通过 1×1 卷积进行通道对齐, 确保不同层级特征的语义一致性。这种设计充分利用了不同层级特征的互补性, 增强了模型对不同尺度目标的检测能力。

检测头的设计采用了解耦头结构 (Decoupled Head)^[28], 其中分类分支 (Cls.) 引入了改进的多尺度分类头, 通过动态卷积增强特征提取能力; 回归分支 (Reg.) 采用了自适应 IOU 感知机制, 提升了边界框回归的精度; 置信度分支 (Obj.) 用于评估目标框中是否存在目标物体。每个分支独立优化, 避免了任务之

间的相互干扰, 提高了模型的泛化性能。同时, 模型采用了无锚点 (anchor-free) 的检测范式, 直接预测目标的中心点位置及边界框偏移量, 有效降低了计算复杂度。为了进一步提升检测精度, 检测头中引入了自注意力机制, 增强了模型对目标上下文信息的理解能力^[29]。

2.3 融合特征输入模块

融合特征输入模块的功能是实现多分量冰裂声信号特征的协同学习。对冰裂声信号特有的多模态、多分量特征在时域、时频域进行融合, 让模型充分利用多维度信息, 提高检测的全面性和准确性。经特征融合预处理的数据以混合波形和堆叠时频图两种特征表示方式作为后续半监督学习模块的输入。

将声信号转化为图像是机器学习算法在相关领域应用中的常见做法, 其核心在于通过可视化手段突出和展示声信号的特征信息。在基于伪标签的半监督学习目标检测框架中, 特征信息的视觉呈现质量直接决定了模型的检测性能。在生物声学、语音分析及地震学等传统机器学习算法应用场景, 该过程通常通过声信号时域波形和时频域能量分布两种方法实现^[30,31]。然而, 与上述场景有所区别, 冰裂事件的声学特征蕴含于其三维弹性声场中, 任意单分量时域信号均难以完整呈现其相对复杂的物理特性。因此, 需要针对典型冰裂声信号特点将三分量信号特征融合, 适配上述网络结构的同时支持实现模态特征协同表征, 最终提升模型在复杂冰裂信号识别中的表现。

研究延续当前主流声信号可视化路径, 设计了两种改进的图像化特征表示方法, 以提升模型的特征提取与识别能力, 图 3(a) 是基于时域波形的可视化方案。以相同时间窗口内三分量冰裂声信号最大绝对振幅为参照进行归一化处理, 消除不同信号片段之间的幅值差异。最后, 采用统一时间轴和不同的颜色, 将归一化后的信号绘制于相同图窗中完成混合。混合后的三通道特征图保留了冰裂信号三个分量上的时域波形信息和信息独立性, 优化多维特征的可视化表达的同时避免了信息流失。

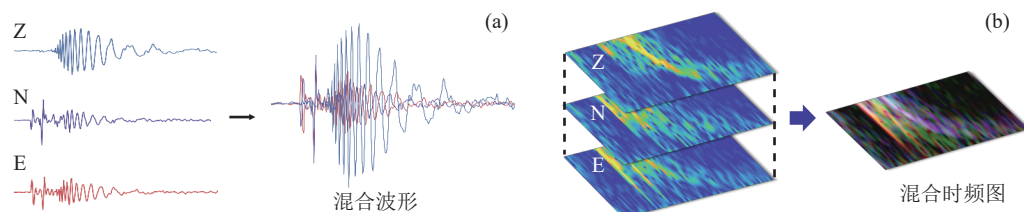


图3 冰裂声信号三分量两种可视化特征表示方案 (a) 混合波形图可视化; (b) 堆叠时频图可视化

如图 3(b) 所示, 第二种方案是基于短时傅里叶分析的时频域可视化方案。利用 STFT 将声信号的能量分布以色图的形式映射在时间-频率空间是传统声音可视化方法之一, 但上述过程中的颜色引用并非必须, 单一信号时频域能量分布可用单色图(黑白)完整展示, 只是针对人眼视觉系统对色彩的灵敏度优势而选择彩色图予以呈现。因此, 研究选择充分利用 RGB 显色系统特性, 将典型冰裂声信号记录分别进行 STFT 处理后整合成为彩色时频谱图, 利用颜色呈现时频空间内声能量在三个正交分量上的分布特征。

具体做法是, 将三分量时域信号片段分别进行 STFT 处理得到其时频谱对数表达, 而后基于累积统计分布情况以 0.1% 为阈值剔除极端幅值。上述处理的必要性在于, 为了适配 RGB 显色系统, 上述时频谱数值需要归一化至 $[0, 255]$ 的整数空间, 通过剔除偶发性极值有效规避归一化流程带来的不良影响。本研究中选择将 N 分量信号对应 R 色谱, E 分量对应为 G 色谱, 垂直分量对应为 B 色谱。图 3(b) 为冰裂声信号的三分量堆叠时频图可视化, 横轴为时间, 纵轴为频率, 图像平均亮度体现了三分量记录中的平均声能量大小, 而颜色则表示了三个正交分量对应方向上的声能量分布特点。

3 实验

3.1 数据获取

2024 年初在吉林省吉林市松花湖区杉松背附近(图 4)湖面冰层上开展了数据采集试验。实验区

域冰面平整, 冰层厚度实测为 50 cm。实验期间大气温度日变化在 $-15\text{ }^{\circ}\text{C}$ ~ $25\text{ }^{\circ}\text{C}$ 左右, 有少量降雪。冰声采集设备为型号 A1 的三分量检波器, 其有效工作频带为 0.1~240 Hz, 采样频率设定为 500 Hz, 与冰面采用少量清水冻结耦合。本次数据采集试验中观测到大量典型冰裂事件, 对应冰声记录质量良好。本次实验设备所在位置(红色五角星标注)西侧 200 m 位置存在冰层挤压、破裂后形成的隆起, 南北贯穿湖面, 是上述冰裂事件的主要潜在成因。

3.2 数据检测效率对比验证实验

实验选取 2024 年 2 月 1 日凌晨 2 点至 3 点间的 1 小时冰声连续观测记录作为训练数据。采用基于滑动窗口的时序分割策略对连续信号分帧, 设置时间窗口长度为 5 s、重叠率为 50% 以降低漏检概率。而后, 参照 2.3 节所述流程对所得三分量信号片段可视化预处理。利用人工初筛和基于三分量信号偏振分析的二次核实完成人工标注过程, 将上述信号中包含典型冰裂信号的时间片段构建为正样本数据集, 其余时间片段对应观测划分为负样本数据集。最终, 正样本数为 297, 负样本数为 1502。

为验证伪标签目标检测框架降低训练资源开销方面的有效性, 将常规目标检测算法 YOLOX 作为基准检测器。通过评估两种方法在典型冰裂声信号检测应用场景下的表现, 明确本文所提框架能否在维持与基准模型相当的检测性能的同时减少人工标注量需求。

为系统评估本文所提模型的检测效率, 首先随机抽取正样本数据集的 80% 作为对照组(YOLOX

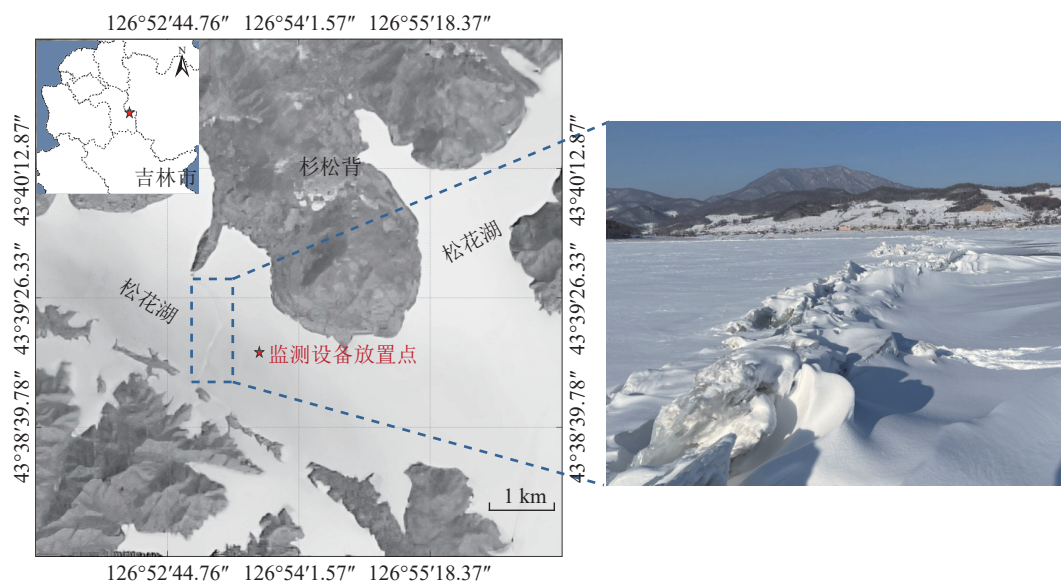


图 4 数据采集实验卫星图^[32]

模型训练) 的训练集。对于实验组 (伪标签目标检测框架), 基于 2.1 节所述流程, 随机抽取正样本数据集的 40% 作为标注数据; 从下个小时 (2024 年 2 月 1 日凌晨 3 点至 4 点) 的数据里随机抽取相同数量样本作为“未标注数据”混合构成其模型训练的训练集。这种设计确保了实验组使用的人工标注训练样本量仅为对照组的 50%。对模型训练过程中的收敛情况进行追踪, 发现所有实验的训练损失和测试损失均快速下降, 随后在 200~300 轮迭代后逐渐平稳收敛且波动幅度合理的趋势。本文应用了早停策略, 以避免模型过拟合。

在测试过程中, 为确保模型评估的可靠性和公平性, 从正样本和负样本中抽取相同数量的样本混合形成测试集, 消除样本类别不平衡对模型评估的影响。通过计算测试集识别结果的精确率 (Precision)、召回率 (Recall) 和 F1-Score 完成模型检测效率的评估。其中, 精确率表示正确识别样本数占所有识别结果的比率; 召回率表示正确识别样本占所有正样本的比率, F1-Score 作为精确率和召回率的调和平均值, 反映模型在准确性与召回性能上的平衡性。三个评估指标的数学表达式如下:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (4)$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (5)$$

其中, TP (True Positive) 表示正确识别的冰裂信号样本数, FP (False Positive) 表示负样本被误判为正样本的数量, FN (False Negative) 表示未能被检测出的实际冰裂信号样本数, TN (True Negative) 表示负样本被正确判定为负样本的数量。

为进一步评估模型在实际应用中的检测表现, 引入假阳率 (False Positive Rate, FPR) 作为补充指标, 该指标用于反映检测过程中, 将负样本误判为正样本的数量占总负样本数量的比率:

$$\text{FPR} = \frac{FP}{FP + TN}. \quad (6)$$

表 1 实验结果表明, 应用伪标签策略的半监督

目标检测框架在仅使用一半人工标注数据的情况下, 混合波形图的精确率下降 9.29%, 召回率提升 6.67%, F1-score 仅下降 1.64%; 堆叠时频图的精确率下降 0.24%, 召回率下降 10%, F1-score 下降 4.41%。以 F1-score 作为综合性能指标考量可以发现, 不论选择哪种特征表示方法, 应用伪标签策略后的实验结果相对基准模型仅出现轻微下降。在整体检测性能相当的条件下, 大幅度降低了人工标注带来的训练开销。对比两种特征表示方法可以看到, 堆叠时频图拥有更高的 F1-score 和更低的假阳率, 初步体现了该方法的应用优越性。

在检测任务中, 可以通过改变判定阈值来调整检测敏感度, 在此过程中召回率和假阳率通常始终呈现正相关关系, 因此无法实现对假阳率和漏检率的同步一致。针对冰裂声信号检测场景, 综合考虑研究价值和获取成本后, 在模型训练和参数优化过程中选择相对低的判定阈值以保证较高召回率, 实现尽可能多的信号捕获。在此过程中, 假阳率仍保持在与基准模型相近的水平 (30%)。需要说明的是, 上述参数设置导致的过敏感特性属于策略选择, 而非模型固有缺陷, 这种选择对于全面检测和研究冰裂现象具有积极意义。

3.3 检测精度验证对比实验

相对 3.2 节实验中的非对等实验条件, 本节实验验证相同标注资源限制条件下伪标签策略带来的检测性能提升。因此, 整体实验参数保持一致的条件下, 将基准模型的训练数据量降低至与伪标签优化模型相同水平, 即正样本数据集的 40%, 实验结果如表 2 所示。

对比表 2 和表 1 的第一行数据可知, 利用 40% 样本条件下 YOLOX 模型的分类性能受到了显著劣化, 说明该条件下 YOLOX 模型尚处于数据稀疏状态。而相较于仅使用标注数据的基准模型, 半监督框架通过有效利用伪标签机制, 在检测精度、召回率等指标上均取得显著改善。具体而言, 在混合波形图特征表示下, 该框架的 F1-score 相对基准模型提升了 13.8%。特别值得注意的是, 当采用堆叠时频图特征表示时, 性能差异进一步扩大至 20.8%。因此在下

表 1 效率对比评估指标结果

模型	特征表示	精确率 (%)	召回率 (%)	F1-score (%)	假阳率 (%)
YOLOX	混合波形图	75.00	70.00	72.41	23.33
	堆叠时频图	72.97	90.00	80.60	33.33
伪标签优化	混合波形图	65.71	76.67	70.77	40.00
	堆叠时频图	72.73	80.00	76.19	30.00

表 2 精度对比评估指标结果

模型	特征表示	精确率 (%)	召回率 (%)	F1-score (%)	假阳率 (%)
YOLOX	混合波形图	62.07	60.00	61.02	36.67
	堆叠时频图	69.57	53.33	60.38	23.33
伪标签优化	混合波形图	65.71	76.67	70.77	40.00
	堆叠时频图	72.73	80.00	76.19	30.00

一小节中将探究不同特征表示方法对模型性能的影响。

为进一步验证模型在实际应用场景中的泛化能力,从实验采集数据中额外选取一小时原始波形数据进行测试。分别应用三种方法(传统的 STA/LTA 算法、基准模型 YOLOX 和伪标签优化模型)对该数据进行冰裂声信号检测。在 STA/LTA 算法的参数选择上,参考 Rööslä 等的研究^[33],依据冰裂信号的持续时间特征,将短时窗设置为 1 s,长时窗设置为 10 s,触发阈值设为 3 以平衡检测灵敏度和虚警率。由于 STA/LTA 算法仅适用于单分量信号检测,考虑到 QS 模态在垂直方向的显著性,最终选择垂直分量数据进行测试,这一选择也与 Rööslä 等^[33]的研究思路一致。经专家人工复核后,最终 STA/LTA、YOLOX 和伪标签优化方法的精确率分别为 33.91%、60.45% 和 65.14%。传统 STA/LTA 方法在处理复杂的冰裂信号时表现不佳,本研究提出的伪标签优化模型表现最好。

综合表 2 的实验结果可以看出,无论是在相同标注资源限制条件下还是在处理全新波形数据时,伪标签优化模型都有着良好表现,证明了其在实际应用场景中的泛化能力和适用性。

3.4 特征融合对比实验

在目标检测任务中,数据集的质量对模型性能具有决定性影响。从表 1 和表 2 的四种指标数据可以看出,不论是基准模型还是伪标签策略模型,堆叠时频图的结果大部分都优于混合波形图。为了全面评估混合波形图和堆叠时频图两种特征表示方法在冰裂声信号检测任务中的性能优劣,引入 ROC (Receiver Operating Characteristic) 曲线作为评估工具。

ROC 曲线通过描绘真阳率 (True Positive Rate) 与假阳率 (False Positive Rate) 的权衡关系,全面展示检测器在不同决策阈值下的性能表现。降低分类阈值时,更多的真实冰裂信号会被正确检测出来,同时更多的非冰裂信号会被误判为冰裂信号。ROC 曲线下面积 (Area Under Curve, AUC) 是 ROC 曲线性能的定量度量, AUC 越大则相同假阳率下的真阳率更高。因此,曲线越接近左上角,表明检测器的检测性

能更好。如图 5 所示,堆叠时频图表示方法的 AUC 值达到 0.738,相较于混合波形图的 0.662 提升了 11.5%。这一结果表明,时频域特征相比时域特征能更好地捕捉冰裂声信号的特征模式。

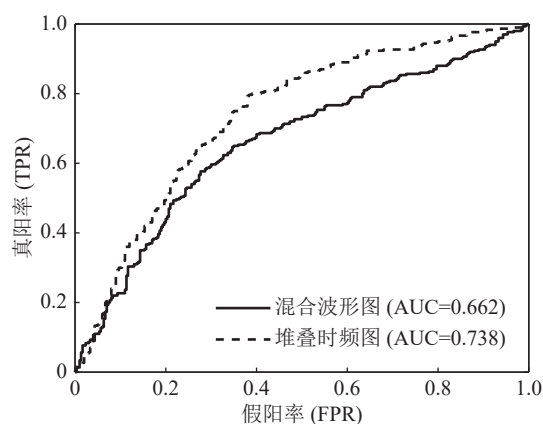


图 5 ROC 曲线评估

4 结论

本文应用基于多分量特征融合的半监督学习方法,结合高纬度冰区实测数据,实现了低标注数据条件下冰裂声信号的高效检测。对所提方法在混合波形图和堆叠时频图两种特征融合方案下的检测性能进行了系统评估,得到以下结论:(1) 半监督学习伪标签策略显著降低了对标注数据的依赖性,使用 50% 的人工标注量即可实现接近基准 YOLOX(全标注模型) 94.5% 的检测性能;(2) 在相同人工标注量条件下,检测性能较基准模型 YOLOX 提升 26.2%;(3) 堆叠时频图特征融合方案在本次冰裂声信号检测任务中表现更加优异,相比混合波形图方案具有更好的综合检测能力。

相比内陆湖冰,极地海冰在材料特性和微观结构等方面均有所差异,可能导致其冰裂声源激发特征存在不同。但研究表明两者在低频段传播特性以及导波模态结构上具有较高的一致性。同时前述结果充分说明,本文所提方法通过对于伪标签策略的合理应用,有效降低了模型检测性能和训练集标签样本数量的强依赖性,从而进一步提升其在极地场景下的适用性。未来将深入研究多维度信息感知对

冰裂声信号检测性能的影响机制, 进一步优化特征融合方案, 以提高检测系统在极地环境监测中的适应性和稳定性。

参 考 文 献

- O'Neel S, Marshall H P, McNamara D E, *et al.* Seismic detection and analysis of icequakes at Columbia Glacier, Alaska. *J. Geophys. Res.: Earth Surface*, 2007; **112**(F3): 1–11
- Podolskiy E A, Walter F. Cryoseismology. *Rev. Geophys.*, 2016; **54**(4): 708–758
- Veitch S A, Nettles M. Assessment of glacial-earthquake source parameters. *J. Glaciol.*, 2017; **63**(241): 867–876
- Scholander P F, Nutt D C. Bubble pressure in Greenland icebergs. *J. Glaciol.*, 1960; **3**(28): 671–678
- 殷敬伟, 马丁一, 张宇翔, 等. 极地海冰声波导建模综述. *物理学报*, 2022; **71**(8): 162–172
- Seroussi H, Nowicki S, Payne A J, *et al.* ISMIP6 Antarctica: A multi-model ensemble of the Antarctic ice sheet evolution over the 21st century. *Cryosphere*, 2020; **14**: 3033–3070
- Moreau L, Boué P, Serripieri A, *et al.* Sea ice thickness and elastic properties from the analysis of multimodal guided wave propagation measured with a passive seismic array. *J. Geophys. Res.: Oceans*, 2020; **125**(4): 1–18
- Latto R B, Turner R J, Reading A M, *et al.* Towards the systematic reconnaissance of seismic signals from glaciers and ice sheets, Part I: Event detection for cryoseismology. *Cryosphere*, 2024; **18**(4): 2061–2079
- Binder G, Chakraborty D. Detecting microseismic events in downhole distributed acoustic sensing data using convolutional neural networks. In: SEG Technical Program Expanded Abstracts, 2019: 4864–4868
- Hudson T S, Smith J, Brisbourne A M, *et al.* Automated detection of basal icequakes and discrimination from surface crevassing. *Ann. Glaciol.*, 2019; **60**(79): 167–181
- Allen R V. Automatic earthquake recognition and timing from single traces. *Bull. Seismol. Soc. Am.*, 1978; **68**(5): 1521–1532
- Van Trees H L. Detection, estimation, and modulation theory, Part I: detection, estimation, and linear modulation theory. John Wiley & Sons, 2004
- 唐伟, 刘俊民, 王晓明, 等. 次声信号的时频信号能量统计检测方法. *声学学报*, 2014; **39**(1): 85–92
- Stork A L, Baird A F, Horne S A, *et al.* Application of machine learning to microseismic event detection in distributed acoustic sensing data. *Geophysics*, 2020; **85**(5): 149–160
- Schaff D P, Waldhauser F. One magnitude unit reduction in detection threshold by cross correlation applied to Parkfield (California) and China seismicity. *Bull. Seismol. Soc. Am.*, 2010; **100**(6): 3224–3238
- Ross Z E, Trugman D T, Hauksson E, *et al.* Searching for hidden earthquakes in Southern California. *Science*, 2019; **364**(6442): 767–771
- Zhu W, Biondi E, Li J, *et al.* Seismic arrival-time picking on distributed acoustic sensing data using semi-supervised learning. *Nat. Commun.*, 2023; **14**(1): 8192
- Perol T, Gharbi M, Denolle M. Convolutional neural network for earthquake detection and location. *Sci. Adv.*, 2018; **4**(2): 1–8
- Ewing M, Crary A P. Propagation of elastic waves in ice. Part II. *Physics*, 1934; **5**(7): 181–184
- Ewing M, Crary A P, Thorne A M. Propagation of elastic waves in ice. Part I. *Physics*, 1934; **5**(6): 165–168
- Stein P J. Acoustic monopole in a floating ice plate. *J. Acoust. Soc. Am.*, 1986; **80**(4): 1263–1263
- Gao J, Zhang Y, Ma D, *et al.* In-situ characterization of wave velocity in ice cover with seismic observation on guided wave. *Cold Reg. Sci. Technol.*, 2024; **231**(6): 104392
- Stein P J, Euerle S E, Parinella J C. Inversion of pack ice elastic wave data to obtain ice physical properties. *J. Geophys. Res.: Oceans*, 1998; **103**(C10): 21783–21793
- Yang X, Song Z, King I, *et al.* A survey on deep semi-supervised learning. *IEEE Trans. Knowl. Data Eng.*, 2022; **35**(9): 8934–8954
- Deng J, Li W, Chen Y, *et al.* Unbiased mean teacher for cross-domain object detection. IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, Online, 2021: 4091–4101
- Zhou H, Ge Z, Liu S, *et al.* Dense teacher: Dense pseudo-labels for semi-supervised object detection. European Conference on Computer Vision, 2022: 35–50
- Niu Z, Zhong G, Yu H. A review on the attention mechanism of deep learning. *Neurocomputing*, 2021; **452**: 48–62
- Peng H, Tan X. Improved YOLOX's anchor-free SAR image ship target detection. *IEEE Access*, 2022; **10**: 70001–70015
- Tang R, Sun H, Liu D, *et al.* EYOLOX: An efficient one-stage object detection network based on YOLOX. *Appl. Sci.*, 2023; **13**(3): 1506
- Xie J, Zeng X, Liang C *et al.* Ice plate deformation and cracking revealed by an in situ-distributed acoustic sensing array. *Cryosphere*, 2024; **18**(2): 837–847
- Deng S, Srinivasan R. Semantic and acoustic analysis of speech by functional networks with distinct time scales. *Brain Res.*, 2010; **1346**: 132–144
- USGS, 松花湖卫星影像, 比例尺 1 km. “LandsatLook”. (2024-2-1). <https://landsatlook.usgs.gov>
- Röösli C, Walter F, Husen S, *et al.* Sustained seismic tremors and icequakes detected in the ablation zone of the Greenland ice sheet. *J. Glaciol.*, 2014; **60**(221): 563–575