

文章编号 1004-924X(2021)09-2189-21

## 相位敏感光时域反射仪的信号处理方法综述

田曼伶<sup>1,2</sup>, 刘东辉<sup>1,2</sup>, 曹晓敏<sup>1,2</sup>, 余凯禄<sup>1,2\*</sup>

(1. 北京交通大学 信息科学研究所, 北京 100044;

2. 现代信息科学与网络技术北京市重点实验室, 北京 100044)

**摘要:** 相位敏感光时域反射仪由于具备监测范围广、灵敏度高等优点被广泛应用于周界安防等领域。近年来, 研究者对其光学系统的改进使得传感距离更长、空间分辨率更高。但需要处理的数据量大大增加, 且环境噪声以及扰动类型多样给分布式扰动传感系统的实际应用带来了挑战。本文总结了提高该系统信号信噪比、扰动识别率等性能指标的信号处理方法, 包括降噪算法、特征提取算法、机器学习以及深度学习算法, 尽可能地对比各种算法的优劣, 并展望了未来该领域信号处理方法可能的发展方向。

**关键词:** 相位敏感光时域反射仪; 信号处理; 降噪算法; 特征提取; 机器学习

**中图分类号:** TP394.1; TH691.9 **文献标识码:** A **doi:** 10.37188/OPE.20212909.2189

## Signal processing methods of phase sensitive optical time domain reflectometer: a review

TIAN Man-ling<sup>1,2\*</sup>, LIU Dong-hui<sup>1,2</sup>, CAO Xiao-min<sup>1,2</sup>, YU Kuang-lu<sup>1,2\*</sup>

(1. *Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China;*

2. *Beijing Key Laboratory of Modern Information Science and Network Technology, Beijing 100044, China)*

\* *Corresponding author, E-mail: klyu@bjtu.edu.cn*

**Abstract:** Phase-sensitive optical time domain reflectometry is widely used in perimeter security and other fields because of its advantages of wide monitoring range and high sensitivity. In recent years, researchers have improved optical systems to increase sensing distance and spatial resolution, thus greatly increasing the amount of data that needs to be processed. In addition, strong environmental noise and diverse types of vibrations bring challenges to the practical application of distributed vibration sensing systems. This study summarizes the signal processing methods used to improve the signal-to-noise ratio and vibration recognition rate of the system, including noise reduction algorithms, feature extraction algorithms, machine learning, and deep learning algorithms; compares the advantages and disadvantages of different algorithms; and finally outlines the possible direction of signal processing methods in this field in the future.

**Key words:** phase sensitive optical time domain reflectometer; signal processing; noise reduction algorithm; feature extraction; machine learning

收稿日期: 2021-03-30; 修订日期: 2021-04-22.

基金项目: 中央高校基本科研业务费专项资金资助 (No. 2020YJS043, No. 2020JBM024); 国家自然科学基金 (No. 61805008)

## 1 引 言

近年来,基于相位敏感的光时域反射仪(Phase Sensitive-Optical Time Domain Reflectometry,  $\Phi$ -OTDR)或称分布式声学传感(Distributed Acoustic Sensing, DAS)具有抗电磁干扰、可分布式测量、监测范围广、灵敏度高、定位精度高以及结构紧凑等诸多优势,被广泛应用于周界安防、油气管道检测、土木结构健康监测等领域<sup>[1-5]</sup>。早期对  $\Phi$ -OTDR 的研究主要集中在光学系统(即信息采集系统),提高了  $\Phi$ -OTDR 系统的关键性能参数如传感距离、频率响应范围、空间分辨率、测量灵敏度等。有的方案通过增益信号提升性能,在  $\Phi$ -OTDR 系统中加上 EDFA 来提高瑞利散射信号强度<sup>[1,6]</sup>,也可利用双向一阶拉曼放大<sup>[7-9]</sup>、双向二阶拉曼放大<sup>[10]</sup>延长距离并提高空间分辨率;有的方案采用先进检测方案改善性能,如外差检测<sup>[9,11-13]</sup>、相干检测<sup>[14]</sup>等方案,也可通过与 Mach-Zehnder 干涉仪、Michelson 干涉仪组合的方案<sup>[15-18]</sup>来提高频率响应范围。

近十年来,随着  $\Phi$ -OTDR 系统性能的进一步提升(更高频率响应范围、更长传感距离、更高空间分辨率),单位时间内需要处理的数据量增加若干数量级,传统信号处理方案因处理速度慢而无法满足不同需求。此外,在实际应用中,传感光纤所处的环境复杂多变,动态环境变化如天气变化等因素使得采集到的信号信噪比(Signal Noise Ratio, SNR)低,且作用在光纤上的扰动信号种类繁多,识别扰动的难度增加。考虑到  $\Phi$ -OTDR 系统实际应用中,事件误报和漏报成本极大,而传统信号处理方案因为扰动识别准确率低、可正确识别的类别少,已无法满足  $\Phi$ -OTDR 系统的应用需求。因此,科研人员开始探索从信号处理方向提升系统性能,将图像处理、声音识别等领域的信号处理方法迁移至  $\Phi$ -OTDR 系统中,在提升信号的质量的同时,也避免增加成本和复杂性,实现对多种振动事件的实时识别。根据  $\Phi$ -OTDR 系统信号处理目的的不同,可分为降噪算法和扰动分类算法。分类算法可以根据是否需要预先提取时域特征<sup>[19]</sup>、频域特征<sup>[20-22]</sup>等分为两种:一种是传统机器学习算法,如:支持向量机(Support Vector Machine, SVM)<sup>[19,23-28]</sup>、相关

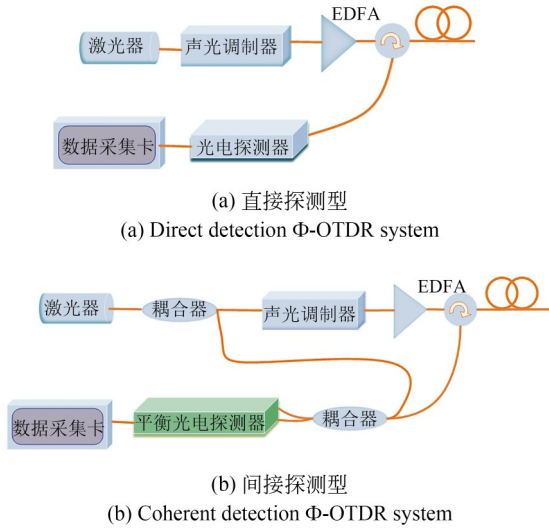
向量机(Relevant Vector Machine, RVM)<sup>[29-30]</sup>等;另一种是端到端的深度学习算法,如:卷积神经网络(Convolutional Neural Network, CNN)<sup>[31-35]</sup>。在实际应用中,降噪算法和分类算法通常结合使用,对系统采集到的信号做降噪处理后,再对其进行分类,如 Wu<sup>[35]</sup>利用谱减法抑制信号的宽带噪声,将处理后的信号输入到多尺度一维卷积神经网络(Multi-Scale 1-D CNN, MS 1-D CNN)识别扰动类型。

本综述查阅了 2000 年到 2020 年之间的相关资料(关键词: $\Phi$ -OTDR、DAS、信号处理、机器学习),主要分为六节:第一节和第六节分别是引言和总结;第二节介绍了最典型的  $\Phi$ -OTDR 系统及其事件检测的原理;第三节是降噪算法,介绍了用于  $\Phi$ -OTDR 信号降噪的算法原理及方案;第四节是分类算法和特征提取,介绍了用于分类扰动的算法,以及在部分算法中需要用到的信号的特征类型及特征提取方法;第五节是讨论,从系统传感距离、扰动类型数量等方面总结了应用于  $\Phi$ -OTDR 系统的分类算法,并从准确率、误报率(Nuisance Alarm Rate, NAR)等方面比较了不同算法的性能。本文还对  $\Phi$ -OTDR 系统信号处理方法的未来发展方向进行了展望,仅作抛砖引玉之用。

## 2 系统结构与原理

相位敏感光时域反射仪是一种基于后向瑞利散射原理的分布式光学传感系统,经典的  $\Phi$ -OTDR 系统如图 1 所示。光源发出的连续光经声光调制器(Acoustic-Optic Modulator, AOM)调制成探测光脉冲,再经掺铒光纤放大器(Erbium-Doped Optical Fiber Amplifier, EDFA)放大后进入到传感光纤中,产生的后向瑞利散射光经放大滤波后被光电探测器(Photodetector, PD)接收,随后由数据采集卡(Data Acquisition Card, DAQ)采集存储。当外界扰动或温度变化作用于传感光纤时,光纤的长度、折射率等参数变化,后向瑞利散射光的光程、相位随之变化,通过检测并分析反射光的光强或相位变化,即可检测识别扰动。 $\Phi$ -OTDR 可以分为直接探测型(图 1a)和相干探测型(图 1b)。

强度型  $\Phi$ -OTDR 系统工作原理介绍如

图1  $\Phi$ -OTDR 系统Fig. 1  $\Phi$ -OTDR system

下<sup>[36]</sup>,信号光场可表述为:

$$E_s(t) = A_s \cos(2\pi f_s t + \varphi_s), \quad (1)$$

其中: $A_s$ 是信号光光场的振幅, $\varphi_s$ 是信号光光场的相位,直接探测到的光场强度为:

$$I(t) = E_s^2(t) = \frac{A_s^2}{2} + \frac{A_s^2}{2} \cos(4\pi f_s t + 2\varphi_s). \quad (2)$$

当扰动作用于传感光纤上时,引起相位 $\varphi_s$ 的变化,体现为探测信号幅值的变化。

区别于直接探测型系统,在相干探测型系统中,平衡光电探测器还接收了频率为 $f_i$ 的本振光,其光场为:

$$E_i(t) = A_i \cos(2\pi f_i t + \varphi_i), \quad (3)$$

其中: $A_i$ 是本振光光场的振幅, $\varphi_i$ 是本振光光场的相位,两束光在光电探测器上混频后光强为:

$$I(t) = A_s^2 \cos^2(2\pi f_s t + \varphi_s) + A_i^2 \cos^2(2\pi f_i t + \varphi_i) + 2A_s A_i \cos(2\pi f_s t + \varphi_s) \cos(2\pi f_i t + \varphi_i), \quad (4)$$

滤除直流项后,最终探测器的输出电流为:

$$I(t) = \alpha A_s A_i \cos[2\pi(f_s - f_i)t + \varphi_s - \varphi_i], \quad (5)$$

其中: $\alpha = e\eta/h\nu$ 是光电变换的比例常数, $e$ 为电子电量, $\eta$ 为量子效率, $h\nu$ 为光子能量。由式(5)知,相干探测型系统,可以获取信号的相位与强度信息。通过分析两种系统采集到的信号的强度或相位变化可以获得扰动位置、扰动类型。

### 3 降 噪

原始信号的噪声主要是系统固有噪声以及环境噪声。低 SNR 是造成  $\Phi$ -OTDR 系统高 NAR 的关键因素之一。在传统处理手段中,在对扰动信号进行分类之前,通常需先降噪。根据降噪原理不同,本文将降噪算法分为谱减法<sup>[31,35,37-38]</sup>、小波分解(Wavelet Decomposition, WD)<sup>[2,39]</sup>、经验模态分解(Empirical Mode Decomposition, EMD)<sup>[40]</sup>、压缩感知降噪(Compressed Sensing, CS)<sup>[41]</sup>、图像降噪<sup>[29,42-44]</sup>、自适应时域匹配滤波(Adaptive Temporal Matched Filter, ATMF)<sup>[45]</sup>以及神经网络降噪算法<sup>[46]</sup>。

#### 3.1 频谱相减法

频谱相减法<sup>[47]</sup>是处理宽带噪声的最常用方法。其原理是通过从含噪信号的幅度频谱中减去估计的噪声幅度频谱来获得纯信号的频谱。假设振动信号和噪声之间的相关性为零,则两者彼此独立,并且可以在频域中相加为 $Y_w(\omega)$ 。 $S_w(\omega)$ 、 $N_w(\omega)$ 是信号和噪声对应的短时傅立叶变换谱,可表示为:

$$|\hat{S}_w(\omega)|^2 = |\hat{Y}_w(\omega)|^2 - E[|N_w(\omega)|^2], \quad (6)$$

$$\hat{S}(m) = ITFT[|\hat{S}_w(\omega)|e^{i\varphi(\omega)}], \quad (7)$$

其中: $E[|N_w(\omega)|^2]$ 是背景噪声功率谱,可以从无振动的无声帧的短时功率谱中计算出来。谱减降噪算法优点在于约束少、计算量小。Xu<sup>[37]</sup>将谱减法用于 SVM 分类扰动类型之前,实验结果表明信号信噪比增益在 10 dB 以上;在之后其利用二维卷积神经网络(Two-Dimensional CNN, 2D CNN)分类扰动的研究<sup>[31]</sup>中,也使用了谱减法加强振动信号。Wu<sup>[35]</sup>使用了谱减法抑制宽带加性噪声,作为信号进入 MS 1-D CNN 前的预处理方法之一。Chen<sup>[38]</sup>利用谱减法处理信号,在不削弱信号强度的前提下,有效去除了宽带背景噪声。

#### 3.2 小波降噪

谱减法利用傅里叶变换将一维信号转换成频谱,通过含噪信号频谱减去噪声频谱提高 SNR。傅里叶变换在稳态信号的去噪中发挥出巨大的作用,但不能刻画非稳态信号的局部信息,这是因为傅里叶变换分解时使用的基没有局部化能力。而小波变换可以详细刻画信号时频局部特性,在去除噪声的同时较好地保留细节信

息,因此研究者利用小波变换对  $\Phi$ -OTDR 信号进行降噪<sup>[2,5,9,20,34]</sup>。

小波变换将信号分解为母小波函数的一系列放缩和平移后的基本函数:

$$X_{\text{CWT}}(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \psi^* \left( \frac{t-b}{a} \right) x(t) dt \quad a, b \in \mathbb{R} \& a \neq 0, \quad (8)$$

$a, b$  分别为膨胀因子和平移因子。连续小波变换 (Continuous Wavelet Transform, CWT) 包含大量冗余情况,因此,通常使用离散小波变换 (Discrete Wavelet Transform, DWT) 分解信号,以 2 的幂作为膨胀因子对 CWT 进行采样而获得:

$$X_{\text{DWT}}(j, k) = \sum_{j,k} x_{j,k}(t) 2^{-\frac{j}{2}} \psi \left( \frac{t-2^j k}{2^j} \right). \quad (9)$$

信号经 DWT 的过程可以看作是信号通过一组正交滤波器组,滤波后得到分布在各个尺度上的系数。其中有用信息通常分布在低频系数,噪声则一般分布在高频系数中,随着分解层数的增加,低频系数中保留的噪声越少。小波分解过程如图 2 实线所示。小波分解后,保留最大层数的低频系数(近似系数),对各个尺度上的高频系数(小波系数)进行阈值处理以去除其中的噪声,再对处理后的系数进行小波逆变换,即可得到去噪后的信号。

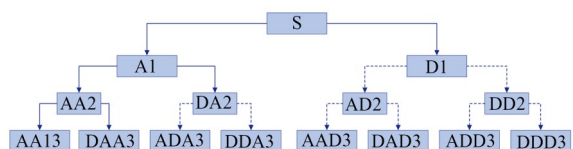


图 2 小波分解(实线)与小波包分解(实线加虚线)示意图

Fig. 2 Schematic diagram of wavelet decomposition (solid line) and wavelet packet decomposition (solid line plus dashed line)

利用 WD 对  $\Phi$ -OTDR 系统信号进行降噪通常可分为三个步骤:

Step 1: 获取振动信号原始数据,根据振动信号的数据特征,选择小波基函数,确定小波分解层数,对原始信号进行小波分解得到小波系数;

Step 2: 选择合适的阈值函数和阈值处理小波系数;

Step 3: 根据阈值处理后的小波系数重构

信号。

传统的硬、软阈值函数如图 3 示,  $\omega_{i,j}$  与  $\hat{\omega}_{i,j}$  分别为原小波系数和阈值量化后的估计小波系数。硬阈值函数将小于阈值  $\lambda$  的小波系数全部置零,将大于阈值  $\lambda$  的系数完全保留。因此在基本滤除噪声后,也会丢失细节信号。软阈值函数重新量化了大于阈值的小波系数,函数在定义域内连续无间断点,能较好地保留细节部分,但也保留了部分噪声。除了这两种阈值函数之外,还有改进的阈值函数如自适应阈值函数等。研究者们可以根据信噪比增益情况选择最佳的阈值函数与合适的阈值。

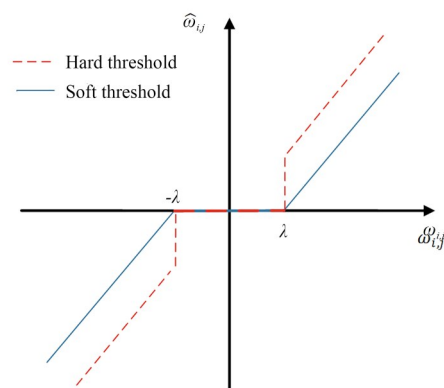


图 3 小波降噪中的阈值函数

Fig. 3 Threshold function in wavelet denoising

小波包分解 (Wavelet Packet Decomposition, WPD) 降噪算法的原理与 WD 类似。小波分解过程是从上到下的不平衡树,只对低频部分进行迭代分解,而小波包分解过程则类似从上到下的平衡树,对高频部分也进行分解,如图 2 实线加虚线示。与 WD 相比, WPD 能更好地分离高频细节。

2015 年, Wu<sup>[20]</sup> 使用多尺度小波分解方法分析了  $\Phi$ -OTDR 系统检测到的纵向时域信号,并通过有选择地重组相应的尺度分量,有效分离了系统自身变化缓慢的不利影响,成功将信号的信噪比提高了 35 dB。Qin<sup>[5]</sup> 使用一种新的小波收缩方法实现在 1 公里传感长度上用 5 ns 的光脉冲检测到 20 Hz 和 8 k-Hz 事件的分布式振动测量,实现了 0.5 米的空间分辨率。Chen<sup>[34]</sup> 对信号进行 WPD 降噪后,再将信号输入到一维卷积神经网络 (One-Dimensional Convolutional Neural Network, 1D CNN) 中进行扰动分类,平均分类准确



率达到95.6%以上。

Peng<sup>[9]</sup>认为离散小波变换收缩算法对使用EDFA放大的短距离Φ-OTDR系统具有很好的去噪效果,但因拉曼放大会引入其他的强度噪声,该方法不适用于采用拉曼放大的长距离Φ-OTDR中产生的低频干扰信号。因此,其在分布式拉曼放大与外差检测相结合的光纤长度为131.5 km的超长Φ-OTDR中使用了一种特殊的WD方法,利用小波的多分辨率分析法,在多尺度小波分解后,对低频系数和小波系数均设置阈值,小于设定阈值的系数为0,保持其余系数的值。利用两种系数调整方法分别去除高频噪声和系统缓慢变化带来的低频噪声。

### 3.3 经验模态分解

除了小波分解降噪算法外,也可以使用经验模态分解对信号进行降噪<sup>[40]</sup>。EMD将信号分解为固有模态分量(Intrinsic Mode Functions, IMF)和残差分量(Residual Component, RES),如式(10),其分解过程如图4示,IMF代表原始信号中的每个频率分量,并且从高频到低频排列,噪声信号被叠加在IMF信号上。对信号进行EMD分解,以获得IMFs和RES,计算 $IMF_i$ 和 $RES_i$ 之间的皮尔逊相关系数(Pearson Correlation Coefficient, PCC),PCC值在0.4~1.0之间的模态函数与残留分量信号叠加。Chen<sup>[40]</sup>该方法将SNR从7.32 dB提升到13.68 dB。

$$X(n) = \sum_{i=1}^L IMF_i(n) + res_i(n). \quad (10)$$

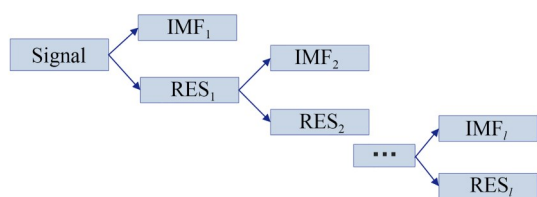


图4 EMD分解过程(IMF<sub>L</sub>:第L层分解后信号;res<sub>L</sub>:L层分解后的残差分量)

Fig. 4 EMD decomposition (IMF<sub>L</sub>: the signal after the decomposition of the L layer; res<sub>L</sub>: the residual component after the decomposition of the Lth layer)

### 3.4 压缩感知降噪

压缩感知理论基于信号的“感知—压缩—重建”理论,其核心在于如何使用尽量少的数据,尽可能地恢复信息。利用Φ-OTDR系统采集到的

传感信息“非满秩”而噪声均匀分布这一特性,Qu<sup>[41]</sup>采用压缩感知的降噪算法处理Φ-OTDR信号,在沿3 km传感光纤上对100 Hz振动事件的压缩比为18.9,信噪比改善为34.98 dB。算法原理及步骤如下示。

在特定的转换域中,信号是稀疏的,而无扰动的原始信号的频谱图没有明显的峰,这说明噪声是非稀疏的<sup>[41]</sup>。因此,在压缩与重构信号之后,大部分噪声得以去除。在对传感信号进行稀疏表示时,可以利用正交匹配追踪(Orthogonal Matching Pursuit, OMP)<sup>[48]</sup>等算法恢复信号。信号压缩过程如式(11)所示。其中X代表原始信号,Y代表压缩后的信号,S是X在稀疏矩阵下的表示,K代表S中非零元素的个数,即信号是K稀疏的,φ是观察矩阵,ψ代表稀疏矩阵,离散傅里叶变换(Discrete Fourier Transform, DFT)、离散余弦变换(Discrete Cosine Transform, DCT)以及DWT都是常见的稀疏矩阵。信号重构过程如图5所示。

$$Y = \phi X = \phi \psi S = AS. \quad (11)$$

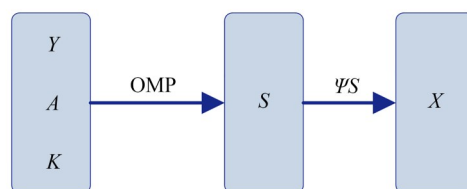


图5 信号重构流程

Fig. 5 Signal reconstruction process

压缩感知Φ-OTDR信号降噪过程<sup>[41]</sup>如下:

(1)计算阈值T并将DFT应用于原始瑞利反向散射曲线以获得原始S向量;

(2)计算K和 $K^* = K + 1$ 的值,K值是S向量中大于T的元素数;

(3)采用OMP算法和收缩阈值, $\hat{S}_1 = \text{OMP}(Y, A, K)$ , $\hat{S}_2 = \text{OMP}(Y, A, K^* - K)$ ;

(4)结合两种情况下的S向量的估计, $\hat{S} = \hat{S}_1 + \hat{S}_2 \eta$ , $\eta \sim 0.05$ ,并基于DFT获得降噪信号;

(5)处理传感光纤中另一个位置的信号时,重复步骤(1)至(4)。

### 3.5 图像降噪

科研人员将图像领域的算法用于Φ-OTDR

系统中,首先将传感信号转为二维图像,再使用图像降噪方案去噪。图像处理领域中的边缘检测算法<sup>[49]</sup>、分割算法<sup>[29]</sup>以及双边滤波算法<sup>[50]</sup>均被用于  $\Phi$ -OTDR 的降噪。

$\Phi$ -OTDR 系统采集到的信号是一维的瑞利后向散射曲线,因此,在使用图像降噪算法前,必须将一维信号转换成二维图像。研究者们通常累积一定时间段内的曲线构成二维矩阵,即瀑布图,并将各点的值归一化到 0~255 间得到二维的灰度图像,再使用图像处理算法检测出目标也是对信号进行降噪处理。

边缘检测法假设图像中强度函数不连续或者强度梯度非常陡峭的地方是图像的边缘,因此可以通过求出图像每个像素点的导数判断图像边缘,即导数显著大于周围像素点导数的点,即为目标边缘所在。边缘检测法常用 Sobel 算子<sup>[51]</sup>、Prewitt 算子<sup>[52]</sup>在图像上卷积。利用边缘检测算法处理  $\Phi$ -OTDR 信号,在定位扰动的位置同时能提升扰动位置处的信噪比。Zhu<sup>[42]</sup>使用 Y 方向(时间方向)的 Sobel 算子与图像卷积,通过对比  $3 \times 3, 4 \times 4, 5 \times 5$  不同尺寸的 Sobel 算子卷积后的实验结果发现:当卷积核尺寸增大时,信噪比提升,当 Sobel 算子与振动信号图像卷积求得的空间梯度已经达到最大值时,增大尺寸,信噪比下降;当采集卡的采样率和脉冲宽度匹配时,使用小尺寸的 Sobel 算子,振动检测的空间分辨率更小。实验结果证明,与传统差分法相比,基于 Sobel 算子的边缘检测法将扰动位置的信噪比提高至 8.4 dB。Wang<sup>[43]</sup>则利用 Prewitt 算子对信号图像进行两个方向的卷积,得到信号的时空分布信息,信噪比由 21.6 dB 提高至 39.07 dB。

图像分割根据灰度、色彩、纹理、几何形状等特征将目标从背景中分离出来,也可用于信号降噪。Sun<sup>[29]</sup>在分析振动信号二维时空图时发现,事件与背景的幅值存在较大的差异,因此其使用图像阈值分割算法去除噪声,阈值计算使用最大类间方差法(Otsu)<sup>[53]</sup>,分割后使用中值滤波去除图像上小的噪声点,最后利用图像膨胀策略来填补图像中的孔。

除此之外,图像滤波算法也是有效的降噪方法,通过重新分配图像像素值来平滑图像,进而

达到降噪的目的。2017 年 He<sup>[44]</sup>将二维图像自适应处理方法——双边滤波算法用于提高  $\Phi$ -OTDR 中扰动位置处的信噪比。该算法通过归一化加权平均来降噪,其权重主要取决于领域像素与中心像素之间的几何距离和像素灰度值差异,权重计算如式(12),为几何权重  $W_G$  和辐射权重  $W_R$  的乘积。 $W_G$  由中心像素与邻域像素之间的几何距离决定,与几何标准差  $\sigma_G$  有关, $W_R$  为中心像素与邻域像素之间的灰度值差异,计算如式(13)。在实际应用中,很难同时调整几何标准差  $\sigma_G$  和灰度标准差  $\sigma_R$  使 SNR 最优,因此将几何权重  $W_G$  调整为洛伦兹函数,见式(14), $m, n$  为邻域像素到目标像素的水平和垂直距离。随着  $\sigma_R$  的增加,SNR 增大,当  $\sigma_R$  达到最优时,SNR 达到峰值。最优  $\sigma_R$  与噪声标准差  $\sigma_n$  成线性正相关<sup>[54]</sup>,通过实验,He 得出算法迭代两次,滤波窗尺寸为  $\sim 41$  时,降噪效果最好,获得了  $\geq 14$  dB 的 SNR 提升。

$$W[i, j; m, n] = W_G[i, j; m, n] \cdot W_R[i, j; m, n], \quad (12)$$

$$W_R[i, j; m, n] = \exp \left\{ - \frac{(X[i, j] - X[i + m, j + n])^2}{2\sigma_R^2} \right\}, \quad (13)$$

$$W_G[i, j; m, n] = \frac{1}{1 + m^2 + n^2}. \quad (14)$$

### 3.6 自适应时域匹配滤波

自适应时域匹配滤波是能够有效提高扰动定位处 SNR 的降噪方法,通过估计扰动信号附近的噪声水平,设计最优的滤波器,以达到最佳 SNR。Ölçer<sup>[45]</sup>将 ATMF 应用于  $\Phi$ -OTDR 领域。采集  $\Phi$ -OTDR 系统数据  $X[M, N]$  ( $M$  代表瑞利后向散射曲线的数量,即采集时间/脉冲周期,  $N$  代表光纤方向上的传感点数),则滤波原理如下示<sup>[55]</sup>:

$y(k)$  是  $k$  位置的滤波后信号,  $W_{opt, k}$  是  $k$  位置处的  $Q \times 1$  的最优滤波器权重矩阵,  $Q$  是待确定的最优权重矩阵维度,即  $W_{opt, k}$  能处理的信号的时间长度。其计算方法见式(16),其中,  $R_k$  是根据由振动信号附近统计估计的噪声协方差矩阵,  $\alpha$  用于归一化输出的噪声功率,见式(17),  $S_k(\bar{\omega})$  代表施加在  $k$  位置处的振动信号,见式(18),  $f$  是

振动频率,  $f_s$  是探测脉冲的重复频率。

$$y(k) = \sum_{i=1}^Q w_{k,i}^* \cdot x_{k,i} = W_k^H X_k; Q \in (0, M], k \in (1, N), \quad (15)$$

$$W_{opt,k} = \alpha R_k^{-1} S_k(\bar{\omega}), \quad (16)$$

$$\alpha = \frac{1}{\sqrt{S_k^H(\bar{\omega}) R_k^{-1} S_k(\bar{\omega})}}, \quad (17)$$

$$S_k(\bar{\omega}) = [1; e^{j\bar{\omega}}; e^{j2\bar{\omega}}; \dots; e^{j2(Q-1)\bar{\omega}}]^T, \quad (18)$$

$$\bar{\omega} = \frac{2\pi f}{f_s}. \quad (19)$$

由式(19)知,ATMF必须先检测到扰动,才能计算出最佳的滤波器权重。与传统的时域滤波算法相比,ATMF通过估计扰动附近的噪声协方差矩阵,能够使扰动位置处的SNR最大,但滤波仅在原始信号的时间方向上进行,滤波后信号的时间维度会变小,有效信息丢失,且只能被用于低频振动信号的去噪。

### 3.7 神经网络降噪

除了传统的降噪算法外,也可以使用神经网络对信号进行降噪处理。Liehr使用CNN对 $\Phi$ -OTD频谱图进行降噪<sup>[46]</sup>,其使用监督学习的方法训练CNN,标签数据 $X$ 是在长达7 km的光纤长度的低噪声波长扫描相干光时域反射仪(Wavelength-Scanning Coherent Optical Time Domain Reflectometry, WS-COTDR)测量得到的距离-频率图像,噪声数据 $N$ 是未连接光纤的相同实验装置测量得到的低通滤波后的光电探测器噪声数据,带噪数据 $Y$ 即CNN的输入数据由 $X$ 与 $N$ 叠加得到。CNN降噪器能够以 $1.2 \text{ GB} \cdot \text{s}^{-1}$ 的速率实时降噪,使得系统在不使用反向散射增强光纤、分布式拉曼或分布式布里渊放大的情况下,在100 km的距离上进行 $\text{nm} \cdot \text{m}^{-1}$ 分辨率的定量应变测量。

综本节所述,在降噪算法中,谱减法和小波降噪算法适用于基于机器学习的分类算法之前,处理时间短且能提高整体信号信噪比;处理数据量较大时,可以选择压缩感知降噪算法,能够同时实现数据的压缩和降噪,也可使用神经网络降噪,处理速度快;图像降噪算法中的边缘检测以及双边滤波算法不仅能提高信号SNR还能提取振动事件,定位扰动,但会造成信号的失真;图像处理降噪算法和ATMF带来的信噪比增益相对

较小,且只能提高扰动位置处的SNR,无法提升整体信号质量。

## 4 分类算法

由第二节知,不同的扰动作用于光纤上时,系统采集到的光强或者相位会随之变化。因此,阈值判断被用于识别扰动。而在实际应用场景中,仅以信号强度/相位值作为判别扰动的唯一参数是不够的:一方面,环境噪声的干扰使得基于阈值的扰动识别方法极易受到影响,系统误报、漏报情况多;另一方面,动态阈值的计算导致系统的实时性差且无法识别光纤上的多点扰动。近年来,由于事件识别的传统方法的局限性,研究者们将复杂有效的分类算法应用于 $\Phi$ -OTDR信号处理端,包括传统机器学习算法以及深度学习算法。

传统机器学习算法,需要提取特定特征,再将其输入至分类器中识别扰动类型;而深度学习算法,将(降噪后的)原始信号或者转换得到的频谱图直接输入到神经网络中,模型同时承担了特征提取和分类的功能,实现端到端的信号处理。

### 4.1 传统机器学习算法

机器学习方法由于其强大的非线性学习能力被广泛应用于信号识别领域。研究者们开始使用机器学习领域中成熟的模型处理 $\Phi$ -OTDR系统信号,研究成果表明,机器学习方法在识别扰动事件上表现好<sup>[29,56-57]</sup>,识别时间短,识别率高,NAR也有所下降。

#### 4.1.1 特征提取

若采用SVM、随机森林(Random Forest, RF)、人工神经网络(Artificial Neural Network, ANN)等模型进行分类,原始信号在降噪后不能直接进入模型中,需要利用小波分解、短时傅里叶变换(Short-Time Fourier Transform, STFT)等方法提取特征,再将特征向量输入到模型中进行训练和测试。常用于分类的特征有时域特征、频域特征、时频域特征等,本文将文献特征提取方法总结如表1,后文将详细介绍。

##### 4.1.1.1 时域特征

表中时域特征通常由差分法或归一化处理后的光强或相位的时域信号计算得到,差分法和归一化是为了消除噪声对提取特征的影响,常见



表 1 特征提取方法

Tab. 1 Literature summary of feature extraction methods

特征提取类别	特征提取方法
[19] 时域	阈值判断、均值
	[58] 归一化
[22] 频域	STFT
	[24] 归一化
[37]	STE、水平交叉率(LCR)、Welch法
	[59] 快速傅里叶变换(FFT)
[57]	归一化、差分法、WPD、奇异值分解
	[26] 直方图、WPD、数理统计
[25] 时域+频域	WPD、ZCR、STE、经验模态分解
	[28] Welch法、周期图法
[38]	梅尔倒谱系(MFCC)、短时过零率、ST、ZCR、短时能量(STE)
	[20] 多尺度小波分解
[56]	ST-FFT
	[21] 小波分解、小波包分解
[60] 时频域	频谱图图像特征(SIF)
	[30] 小波能量谱分析
[23]	STE、LCR、功率谱分析、WD
	[61] 经验模态分解
[29] 其他	图像处理

的时域特征及其计算方法见表 2。

只将时域特征用于扰动分类的实例较少，

表 2 常见时域特征表

Tab. 2 Common time domain characteristics table

特征	定义
峰值(PK)	$X_{PK} = \max(X)$
最小值(Min)	$X_{Min} = \min(X)$
峰值差(PK-PK)	$X_{PK-PK} = \frac{[\max(X) - \min(X)]}{2}$
能量(E)	$X_E = 10 \lg \left[ \sum_{i=1}^n (x_i - \bar{X})^2 \right]$
均值(Mean)	$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$
平均整流值(Arv)	$X_{Arv} = \frac{1}{n} \sum_{i=1}^n  x_i $
均方根(RMS)	$X_{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$
方差(var)	$X_{var} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$
标准差(SD)	$X_{SD} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$

2017年,曲洪权<sup>[19]</sup>提取时域信号的占空比和过均值两类特征用于识别扰动,其定义占空比为每帧信号过阈值的点数;过均值则是将帧内信号与均值作差,对作差后的振动信号进行邻近乘积处理,若乘积结果为负,则视为过均值一次。Wang<sup>[58]</sup>选取峰值、最小值、形状特征等 14 个时域特征用于训练扰动事件分类的随机森林模型,并且在生成模型的过程中,通过如式(20)的 OOB (Out of bag: 即测试集)误差来确定每个结点用于分离事件的最优特征,最终确定每个结点处使用的特征数为 5。

$$OOB \text{ error} = \frac{\text{分类错误样本数}}{\text{总样本数}}. \quad (20)$$

#### 4.1.1.2 频域特征

此外还可提取频域特征识别扰动进行分类,这是因为,不同类型的事件在接收信号中具有不同的频率成分,因此它们在分解分量中具有不同的分布,根据不同频率分量的能量及频域方差等频域特征,可以分类扰动事件。如:Tejedor<sup>[22]</sup>提取信号在不同频带对应的能量特征用于分类。

#### 4.1.2.3 时域+频域特征

本文调研发现,更多研究者倾向于将时域特征和频域特征共同组合成输入特征向量。2017年,张俊楠<sup>[24]</sup>选取时、频域归一化幅度平均值,方差和均方差,归一化幅度能量谱和时、频域的波形因子作为特征参量,在识别踩压、敲击、浇水、无扰动的多分类任务中,分别选取最佳的特征作为模型中二分类器的分类依据。同年,Xu<sup>[37]</sup>选择了短时能量、短时水平交叉率(Short-Time LCR, ST-LCR)、振动持续时间、功率谱能量比等作为分类特征,其中 ST-LCR 是提取信号的时频信息的基本方法,反映了在一定时期内在一定水平附近信号变化的频率,与传感信号的对称性和短时平均频率呈正相关。张伟<sup>[26]</sup>利用直方图法提取信号的短时主冲击与次冲击强度分辨特征并计算长时信号的平均冲击强度、冲击个数;利用小波包方法提取短时信号和长时信号的频率分辨特征。2019年 Wang<sup>[59]</sup>提取信号的时域、频域特征,将信号最大值、信号最小值、频域最大幅度、频率、信号中位数、信号能量、中值信号、主峰与次峰比组成特征向量,输入到模型中分类挖掘和



无扰动事件。Jia<sup>[27]</sup>分别在差分后的信号和归一化后的信号上提取峰值、最小值等15个时域上的特征,共30个时域特征,使用两层的WPD分解信号得到频域上低频部分的低频分量的能量(LLP Energy)、高频部分的低频分量能量(LHP Energy)等6个频域特征。在36个特征中选取类间距离大、类内距离小的峰值、最小值、均值、平均整流值(Average Rectified Value)、均方根、峰值间距、最小值间距、LLP能量以及小波信息量子共9个特征作为分类依据。Jia<sup>[57]</sup>还采用Fisher准则选取具有良好识别性能的时域、频域特征,使用了Fisher准则后,平均扰动识别准确率提高了0.83%,NAR下降了0.88%,识别时间缩短了0.0314 s。付群健<sup>[25]</sup>利用WPD提取小波能量特征,利用改进的经验模态分析(MEEMD)分解信号提取特征,此外,还提取了时域特征短时能量(STE)和短时过零率(ZCR)。2020年,Chen<sup>[38]</sup>在对信号进行降噪后,提取梅尔倒谱系数(MFCC)特征作为长短时记忆(Long Short Term Memory, LSTM)分类器的输入,实时计算传感光纤上每个位置的每个光信号帧的STE和ZCR,将两者与LSTM的隐藏状态连接作为特征向量,每个样本特征提取时间不超过0.1 s。姚媛媛<sup>[28]</sup>则从时域、频域提取信号特征,其提取的特征包括短时能量、短时过零率、短时过电平率、功率谱、频带方差、波形因子、脉冲因子、峰值因子、裕度因子等。

#### 4.1.1.4 时频域特征

除了分别提取时域、频域特征并组合成特征向量之外,时频特征也是常见的用于扰动事件分类的特征。不同于前两种特征,时频域特征中同时包含了时域信息和频域信息。小波分解法是最常见的 $\Phi$ -OTDR信号时频分析法。Wu<sup>[20]</sup>在2015年使用多尺度小波分解法提取分解信号,通过选择性地重新组合相应的频率分量可以分离和提取它不同的扰动事件。在 $\Phi$ -OTDR传感信号的小波分析中,其发现人类入侵事件处于较高频率成分中,中等频率区域代表环境干扰,系统参数缓慢变化导致的背景波动主要位于代表较低频率分量区域,因此Wu以各尺度的系数能量作为输入BP神经网络的特征向量,并获得了

89%的扰动识别准确率。2015年,Tejedor<sup>[56]</sup>在管道完整性威胁监视系统中设计了特征提取模块,利用短时快速傅里叶变换(ST-FFT)提取每帧的频谱内容并计算了频带的能量作为分类特征。2017年,Qian<sup>[21]</sup>对比了WD和WPD两种频域特征提取方法,将小波能量(Wavelet Energy, WE)特征和小波包能量(Wavelet Packet Energy, WPE)特征输入到ANN中进行扰动分类,实验结果表明WPE在分类中表现更好。这是由于WD在不同层级上以不均匀的比例分解信号频带,导致能量始终集中在较高频率的较宽频带中,不同类型事件的WE分布趋势十分相似。

2019年Wang<sup>[30]</sup>利用小波能量谱分析法提取特征,对原始信号进行分解并计算重构系数的能量。Zhang<sup>[60]</sup>提取信号的时频特征SIF作为时间检测的特征向量。徐铖晋<sup>[23]</sup>对信号进行时域、频域以及时频域分析,提取信号的持续时间、短时双过电平率、傅里叶能量谱以及小波能量谱组成复合特征向量,输入到分类器中。2020年,Wang<sup>[61]</sup>使用EMD处理归一化信号以及一阶差分信号,相比于小波能量谱分析,在相同分类模型中使用EMD方法提取的特征,分类准确率更高,由73.04%提高至93.01%。

#### 4.1.1.5 其他特征

在 $\Phi$ -OTDR系统中,从时域信号中分别提取时域、频域特征是很耗时的,也有研究者不将信号转换至频域,直接提取信号的时空域特征。2015年,Sun<sup>[29]</sup>使用了图像处理方法来提取时空域信号的特征,其通过聚类瑞利散射曲线来获取时空域信号,将2秒长的时空域信号矩阵转化为灰度图像,接着使用图像阈值分割方法<sup>[62-63]</sup>将事件区域信号与背景分开,在观察图像发现不同事件的区域形状存在差异后,利用边界跟踪方法找到边界,计算从质心到边界每个点的距离,以最大距离与最小距离的绝对值作为形状特征,还计算了7个区域描述符作为特征,如区域像素数、凸包像素数等。在提取特征之后,其利用协方差矩阵<sup>[64]</sup>选出最优的4个特征作为分类器的输入。

综合分析可知,时域特征、频域特征、时频域特征及图像上的形状特征都能在一定程度上反映出信号的扰动。在不同场景下,由于系统硬

件、环境、扰动类型等方面存在差异,可作为分类依据的最优特征不同。本文在总结文献时发现:时域特征通常选择经归一化处理后的信号的短时能量、短时过零率;而频域特征通常是信号各频带的能量;对于由原始信号处理得到的图像,扰动的形状是分类的主要依据;多个不同域特征更有利于扰动分类;提取信号的多个特征后,可以根据 Fisher 准则<sup>[57]</sup>、Relief 准则等选择最优特征数量及类型。在对  $\Phi$ -OTDR 信号进行特征提取后,所获得特征将被输入到模型中用于分类扰动类型,通常使用传统机器学习算法设计分类模型。

#### 4.1.2 分类模型

##### 4.1.2.1 人工神经网络

ANN 具有自学习、自适应以及强大的非线性函数逼近能力,可以从复杂的数据中提取出适用于分类的有效信息,不需要人为设计分类准则。ANN 结构如图 6 示,分为输入层、隐藏层及输出层,根据数据的复杂程度可以调整隐藏层的数量,隐藏层越多,参量越多,可表达的模型越复杂。2011 年, Jiang<sup>[65]</sup> 利用 BP 神经网络对机场入侵进行预警。2015 年, Wu<sup>[20]</sup> 使用三层 BP 神经网络来决定扰动事件:以各尺度的系数能量作为输入特征向量,输出为事件类型进行分类, NAR 低于 2%。2016 年 Qian<sup>[21]</sup> 将振动信号的小波包能量特征向量输入到 4 层 BP 神经网络中,在对背景噪声,人工挖掘事件和车辆通过干扰三种事件的识别中,平均识别率达到 94.4%。

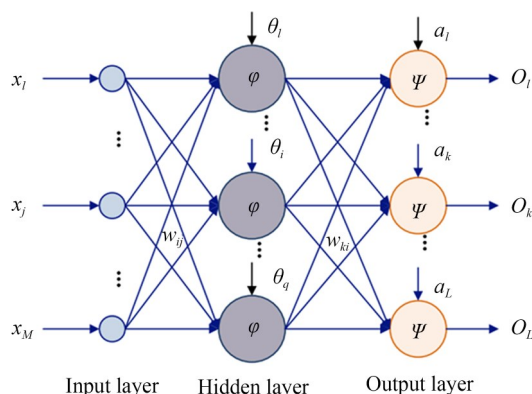


图 6 ANN 网络结构(单隐藏层)

Fig. 6 ANN network structure with single hidden layers

##### 4.1.2.2 相关向量机

相关向量机(RVM)是基于贝叶斯学习框架的学习机器,具有稀疏性高、概率输出大、核函数不需要满足 Mercer 条件的优点<sup>[66]</sup>。2015 年, Sun<sup>[29]</sup> 提取了振动信号的 10 个形态学特征,使用基于散布矩阵的特征选择方法,选取其中评估值最大的区域圆度等 4 个特征,将特征向量输入到 RVM 中识别行走、挖掘、车辆驶过三种扰动事件。由于 RVM 技术是针对二分类问题设计的,因此采用一对一的多类别方法来识别这三个事件即每个分类器识别两类事件,共三个 RVM 分类器:行走-挖掘二分类器、行走-车辆驶过二分类器以及挖掘-车辆驶过二分类器。特征向量输入到三个分类器中,如果两个分类器的结果相同,则可确定其类别。实验结果显示,该方法在低于 1s 的识别时间内,识别率可达 97.8%。2019 年 Wang<sup>[30]</sup> 将提取到的频域特征输入到 RVM 中,实现行走、敲击、慢跑三种振动事件,分类准确率达到 88.6%。

##### 4.1.2.3 支持向量机

支持向量机(SVM)是定义在特征空间上的间隔最大的二分类器,通常按监督学习方式求解最大边距超平面。在面对低维线性样本不可分时,利用非线性映射方法,使其转换到高维空间,通过寻找最优超平面,从而在高维空间对样本进行分类处理,其结构如图 7 示,  $x_n$  代表输入的第  $n$  维特征,  $K$  为核函数,将特征映射到高维空间中。因此, SVM 比较适用于非线性数据的划分,不易出现过拟合的问题。2017 年,徐铖晋<sup>[23]</sup> 从时域、频域及时频域上提取出复合特征并输入到 SVM 分类器中进行分类,实现了传感光纤上振动事件的精确的实时检测,光纤振动信号的识别正确率

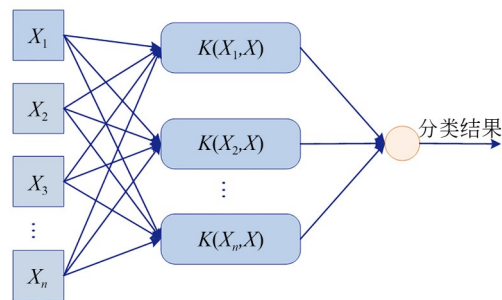


图 7 SVM 结构

Fig. 7 SVM structure

达到93%。2017年,张俊楠<sup>[24]</sup>将SVM算法用于 $\Phi$ -OTDR系统扰动信号特征分类中,并对二叉树结构的分类器进行优化,通过提取信号的平均值、均方差、波形因子以及能量谱等特征,对三种扰动踩压、浇水和敲击进行识别,分别设计四个分类器对事件进行识别,突破了单一特征参量目标识别的局限,提高了扰动源的准确性,降低了运算的时间复杂度,最终正确的扰动识别率在96%以上,漏报率和NAR在4%以下。2017年,曲洪权<sup>[19]</sup>提出一种改进的SVM算法,是基于样本风险代价和先验概率调节超平面阈值的方法,通过占空比和过均值的特征提取方法对数据进行处理,与BP神经网络相比平均分类准确率提高11.51%,与标准的SVM相比平均错误代价降低8.28。2019年,Jia<sup>[27]</sup>提出了一种基于近类支持向量机(NC-SVM)的事件识别方法,该方法利用k近邻(kNN)算法将现有的二元SVM分类器扩展到多类问题。对于25.05 km长的 $\Phi$ -OTDR实验系统,该方案可以有效地识别出五种干扰事件,包括浇水、攀爬、敲击、挤压和虚假干扰事件。5种事件的平均识别率超过94%,识别时间为0.55 s,NAR为5.62%。与一对一的多类支持向量机分类器相比,该方法具有识别率高、识别时间短、NAR低等优点。

#### 4.1.2.4 高斯混合模型

高斯混合模型(Gaussian Mixture Model, GMM)是多个高斯分布函数的线性组合,理论上GMM可以拟合出任意类型的分布,通常被用于解决聚类任务。因此,GMM被用来分类分布式光纤传感系统的振动事件。2017年,Tejedor<sup>[56]</sup>在关于管道完整性威胁监视系统的在线和盲场测试中,选择GMM作为其中的分类模块识别机器活动(不同的挖掘机、钻机等机械实行的移动、打击等活动)以及识别对管道完整性产生威胁/没有威胁的活动。将20 s信号划分成415帧数据进入到GMM中得到每帧的分类结果,再根据投票决策得到信号最终分类结果。为了避免虚假预警以及持续预警,还引入了时空分析程序,即在40 m周围传播不超过80 s的活动的信号被认为是虚警,且被归类为威胁的信号,与先前的威胁相隔小于80 s且小于40 m时,将被归类为同一威胁,系统将不再预警。2017年,Tejedor<sup>[22]</sup>在此

基础上增加了上下文特征模块,该模块将归一化后的频带能量特征向量输入到多层感知器(Multilayer Perceptron, MLP)。研究人员使用三个输入层维度不同的MLP对短、中和长时间上下文的信息进行建模,对应的时间长度分别是5 s, 12.5 s和20 s,将提取到的短、中、长时特征与频带能量特征向量串联后输入到各自对应的GMM中。实验结果表明:提取不同时间长度的上下文特征能有效提高整体分类性能,如:短时特征将不同机器活动的分类准确率由39.5%提高至54.0%;单一时间长度的上下文特征在某些类的识别表现反而较差,如:在识别小型挖掘机的击打活动时,短时特征的识别准确率由13.8%降至11.6%。基于此,Tejedor进一步将不同时间长度的上下文特征的分类结果利用求和法、乘法、最大值法组合起来,实验结果表明决策组合将准确率提高了15.72%。这种提取信号不同尺度特征共同用于分类的方法在后文中提到的深度学习算法MS 1-D CNN<sup>[35]</sup>也出现过。

#### 4.1.2.5 随机森林

随机森林算法(RF)是一个包含多个决策树的分类器,每棵树都是自上而下生成的二元树,最终分类结果由每棵树的输出共同决定。2018年,Wang<sup>[59]</sup>利用RF算法进行挖掘事件和噪声(无扰动)分类,提取信号的时域特征作为分类依据,并利用网格方法<sup>[67]</sup>优化了RF算法。2019年,Wang<sup>[58]</sup>提取信号的14个包含时域特征在内的特征作为RF算法的输入,RF中所有树都是自上而下生成的二元树,从根结点开始,每个结点随机从14个特征中选择 $m_{\text{try}}$ 个特征用于分离,其中最优的特征将通过阈值判断将样本分到不同的结点中,当样本属于同一类时,结点将不再分离。树数 $n$ 与每个节点处的 $m_{\text{try}}$ 是RF分类器的两个重要参数,这两个参数可以通过分析OOB误差式(20)来确定。

#### 4.1.2.6 梯度提升决策树

梯度提升决策树<sup>[68]</sup>(Gradient Boosting Decision Tree, GBDT)跟RF一样都是包含多个决策树的分类器,两者的不同之处在于GBDT是多个决策树串行,所有分类结果相加得到其预测值,每一棵新树的建立是为了使之前模型的残差梯度下降。2020年,Wang<sup>[61]</sup>将EMD分解 $\Phi$ -OT-



DR 信号得到的能量特征输入到极端梯度提升 (Extreme Gradient Boosting, XGBoost) 算法中, 对喷水、敲门、攀爬、按压、假干扰等五种干扰事件的识别正确率分别为 97.96%, 95.90%, 91.1%, 94.84% 和 99.69%, NAR 为 4.1%, 识别时间为 0.093 s。

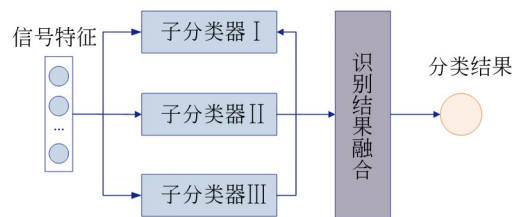
#### 4.1.2.7 极限学习机

极限学习机 (Extreme Learning Machine, ELM) 即单隐藏层前馈网络, 通过最小二乘法进行优化, 没有迭代过程, 仅需要确定输出权重, 是一种训练时间短的灵活的非线性分类器。输入层信号和大量隐藏的神经元通过随机初始化的连接权重连接在一起, 输出神经元是线性的, 输出权重的优化值可以在一次迭代中简单地实现。2019 年, Zhang<sup>[60]</sup> 提取信号的时频特征 SIF, 将其输入到 ELM 和 SVM 两种分类器中, 并选取了线性核函数和 RBF (Radial Basis Function Kernel) 核函数作为前两者的核函数, 构成 4 种分类器, 并对五种事件 (气锤、电钻、金属锤、木锯、打桩机) 进行分类并比较其性能, 实验结果表明 RBF 核函数分类表现优于线性核函数, ELM 分类器表现总体优于 SVM。带有 RBF 核函数的 ELM 能达到 93.9% 的识别准确率, 比带 RBF 核函数的 SVM 分类准确率提高了 1.3%; 平均训练、测试时间分别为 0.87 s 和 1.07 s, 远小于 SVM。Jia<sup>[57]</sup> 提取了振动信号的 40 个时域与频域特征, 利用 Fisher 评分准则选取其中适用于分类的 25 个显著特征, 分别使用筛选后的特征和所有特征训练 ELM 分类器, 结果表明利用 Fisher 准则进行特征选择可以有效地提高  $\Phi$ -OTDR 实验系统中五种干扰事件 (浇水, 爬升, 爆震, 挤压, 假扰动) 的识别率, 超过 95%, 并减少识别时间, 最终识别时间低于 1 s。

#### 4.1.2.8 多分类器组合

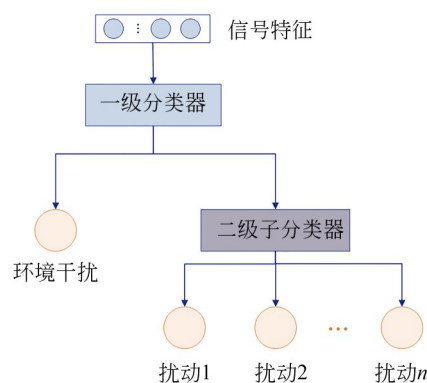
除了使用上述算法识别分类不同的扰动事件外, 也有研究者使用多个分类器组合的算法。不同的分类器在输入同样的信号特征后, 可能会输出不同的分类结果, 且在不同类扰动事件上的识别性能也不一样, 但多个分类器同时分类错误的情况是少见的。因此, 与单个分类器相比, 多分类器组合的方法不仅能提高扰动的识别准确率, 还能有效降低 NAR。多分类器的组合方法如图 8 示, 可分为两种: 一种是融合子分类器的输

出得到最终分类结果, 各子分类器间相互独立; 另一种是选取不同的分类器作为一级分类器与二级分类器, 两者承担的分类任务不同, 前者负责区分环境干扰与扰动事件, 后者进一步识别不同的扰动类型。



(a) 子分类器融合方案示意图

(a) Schematic diagram of sub-classifiers fusion



(b) 分类器分级组合示意图

(b) Schematic diagram of hierarchical classifiers combination

图 8 多分类器组合

Fig. 8 Combination of multiple classifiers

2017 年, 张伟<sup>[26]</sup>将短时信号与长时信号相结合, 从信号的时间相关性和空间相关性出发, 在统计学的基础上, 提取信号在时域以及变换域上的多个短时分辨特征和长时分辨特征, 设计和融合了 SVM 分类器、BP 神经网络、贝叶斯网络的三种子分类器算法的识别结果, 最终 NAR 降至 0%, 事件平均识别率达到 97.41%。2020 年, 姚媛媛<sup>[28]</sup>提出基于分类融合理论的加权分类方法用于分布式光纤系统的模式识别, 提升分类识别精度。即在搭建  $\Phi$ -OTDR 系统后, 根据数据特性和信号特征用 SVM 算法、BP 神经网络、Ada-boost 三种不同的分类器进行分类。根据不同类型扰动信号的特性, 按照加权投票的方式对分类器进行融合, 进而将不同的扰动类型进行准确分



类。对比单个分类器识别结果,分类融合算法的平均分类准确率在94%以上。2019年,付群健<sup>[25]</sup>针对分布式光纤传感系统的灵敏度高、NAR高的特点采用一种二级分类器设计。通过SVM进行一级分类,区分环境干扰及人为干扰,再使用集成学习算法作为第二级分类器,采用随机森林和GBDT两种分类器,有效提高了识别准确率,达到97%以上。

综上所述,传统机器学习算法在扰动识别率、系统NAR上的表现均很优秀,识别准确率高至97.41%,NAR低至0%,但其不具备提取特征的能力。不同文献中最终用于分类的特征、所使用的实验系统、实验数据集均不完全一样,即适用于 $\Phi$ -OTDR系统扰动分类的最优特征没有统一的标准。此外,为了达到更好的分类效果以及实现多扰动分类,该类方案通常需要组合多个不同或相同的分类器。

## 4.2 深度学习算法

### 4.2.1 卷积神经网络

卷积神经网络(CNN)是一种常见的深度学习网络架构,受生物自然视觉认知机制启发而来。CNN可以看作一种特殊的多层感知机或前馈神经网络,具有局部连接、权值共享的特点,其中大量的神经元按照一定的方式组织起来,对视野中的交叠区域产生反应。经典的网络模型包括:LeNet-5<sup>[69]</sup>, AlexNet<sup>[70]</sup>, GoogLeNet<sup>[71]</sup>, VGG-Net<sup>[72]</sup>等。该网络被广泛用于图像分类、目标检测与跟踪、图像复原等领域。起初,研究人员将 $\Phi$ -OTDR振动信号转换为二维图像,再输入到CNN中进行识别<sup>[31-33]</sup>。常用的信号处理方法是STFT<sup>[31]</sup>以及MFCC特征提取方法,得到的频谱图以及时间-MFCC图<sup>[73]</sup>将作为CNN的输入向量,对五种类型的入侵实现了高于90%的分类准确率。除了利用特征提取方法将 $\Phi$ -OTDR传感信号转换为图像外,还可以将信号的瀑布图直接转换为灰度图像<sup>[33,74]</sup>。Sun<sup>[75]</sup>将不同光脉冲得到的瑞利后向散射曲线按时序排列得到三维光脉冲扫描图像,将处理三维图像得到的信号包络面输入到CNN中,5种事件的平均识别准确率>98%,算法识别时间为0.6 s。为了增强模型对信号的分类能力,可以用SVM代替CNN中的Softmax层<sup>[31]</sup>,分类准确率从88%提高到93.3%;

Shi<sup>[74]</sup>也使用了CNN+SVM的分类策略,8种事件的分类准确率达到94.71%,与只使用CNN分类相比,提高了2.04%;也可以改良分类效果好结构复杂的GoogLeNet,在保留分类精度的前提下,移除重复模块与并行路径,缩短模型训练时间<sup>[33]</sup>,对5种事件的分类准确率为96.67%,模型在新应用场景下的训练时间仅为7 min。为了进一步节省信号处理时间,提高系统的实时性,Chen<sup>[34]</sup>将降噪后的一维信号输入到1D-CNN中,网络结构如图9示,石油管道安全监测的现场实测数据的实验结果显示,与2D-CNN相比,1D-CNN在识别指标与速度方面表现更好,平均分类准确率从89.12%提高至95.55%,同时节省了92.8%的处理时间。2020年,Ruan<sup>[76]</sup>搭建了一个端到端的自适应滤波卷积神经网络(Adaptive Filter-CNN, AF-CNN),原始信号先划分为不同帧,堆叠后进行一维卷积处理,每个卷积核相当于一个过滤器,通过神经网络的迭代更新来调整过滤器参数。该一维卷积层相当于自适应滤波层,得到不同频带的滤波信号,将其输入到CNN中进行特征提取和分类,对4种扰动事件的平均识别准确率达到96.7%。

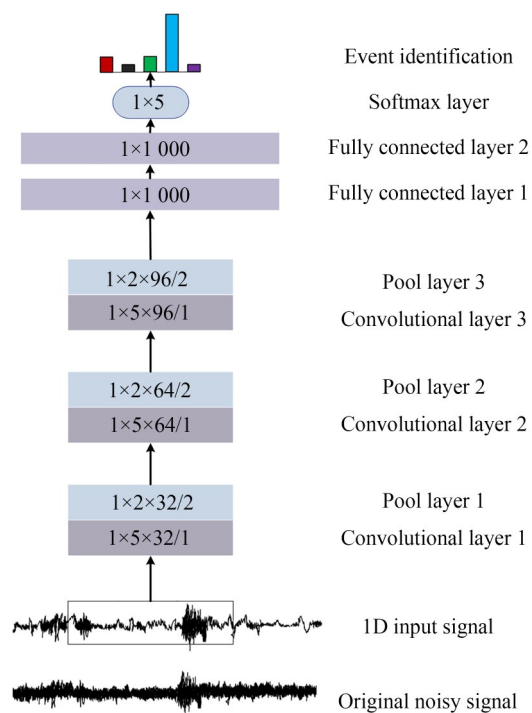


图9 基于CNN的 $\Phi$ -OTDR数据处理方案

Fig. 9 CNN-based  $\Phi$ -OTDR data processing scheme

在 CNN 处理振动信号中,选取固定长度的窗函数会出现时间分辨率和频率分辨率无法较好均衡的问题。从学习局部感受野中特征信息的角度考虑可知:卷积核尺度和跨度较小,信号的时间分辨率较好,对高频特征敏感,但无法较好地学习信号中存在的低频特征;尺度较大的卷积核及对应的大跨度,其效果则相反。为了兼顾高频与低频特征,2019 年, Guan<sup>[35]</sup>将 MS 1-D CNN 用于振动事件多分类中,分类准确率为 96.59%,与 2D-CNN 相比,仅相差 0.19%,但处理速度更快,网络单次迭代时间节省 50%,为 2.41 s。Lyu<sup>[77]</sup>也使用 MS CNN 模型,信号先输入到卷积层和软注意层中,学习传感信号中不同特征的权重分布,该分布再进入到三个并行的不同尺度的卷积块中提取特征,提取的不同尺度特征融合后进入到全连接层,根据原型学习<sup>[78]</sup>(Prototype Learning)的方法进行分类,原型学习可以识别不属于训练集中的类的其他未知事件,具有良好的可接受性和拒绝性,实验结果表明,该方案对 6 个已知类别的目标数据集的平均识别准确率可以达到 84.67%,对 4 个未知类别的自然干扰数据集的拒绝率(将未知样本识别为异常的比率)约为 83.75%,识别响应时间仅为 0.017s 左右。

#### 4.2.2 长短时记忆神经网络

长短时记忆神经网络(LSTM)<sup>[79]</sup>是研究人员对循环神经网络做出的改进,通过引入门控机制来解决长期依赖问题。被广泛应用于语音识别、机器翻译以及文本生成等。不同于 CNN, LSTM 的输入多是从振动信号中提取到的特征,其结构如图 10 示。 $x_t$ 为  $t$  时刻的空域特征和隐藏层状态。

2019 年, Chen<sup>[38]</sup>提取了每帧扰动信号的梅尔频率倒谱系数作为基于注意力机制的长短时记忆网络(Attention Based LSTM, A-LSTM)的输入,每帧信号的隐藏状态与短时能量比及 ZCR 共同作为 softmax 层的输入,用于分类五种扰动事件。

LSTM 还可以与 CNN 结合使用。2019 年, Bai<sup>[80]</sup>考虑到了扰动信号在时间上的连续性,将

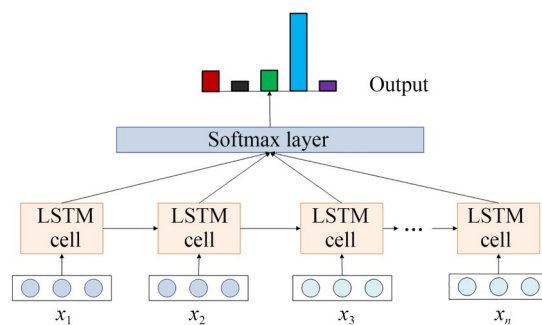


图 10 基于 LSTM 的  $\Phi$ -OTDR 数据处理方案

Fig. 10 LSTM-based  $\Phi$ -OTDR data processing scheme

LSTM 与 1-D CNN 结合,采用了深度学习中的 CLDNN 模型,原始信号直接作为模型输入,省去了预处理的步骤,平均识别正确率达到了 97%。2020 年, Li<sup>[81]</sup>提出了一种基于 CNN 和 LSTM 的  $\Phi$ -OTDR 信号识别框架,即 ConvLSTM, CNN 提取原始数据空域的特征, LSTM 则用来分析特征的时域关系。将该框架部署在高铁站,在强背景噪声的情况下进行了实地实验, ConvLSTM 模型的入侵威胁检测率为 85.6%, NAR 为 8.0%。对于入侵事件分类,平均识别率为 69.3%,平均 NAR 为 13.2%,均检测响应时间缩短到 8.25s。Li<sup>[82]</sup>还设计了 3D CNN 和 ConvLSTM 组合的扰动识别模型,将  $\Phi$ -OTDR 系统采集的信号沿时间和距离两个方向以相同的长度切割得到的时空窗数据作为模型输入。与待识别时空窗处于同一位置但不同时间段的时空窗输入到 ConvLSTM 中提取时域特征,与待识别时空窗处于同一时间段但不同位置的时空窗一起输入到 3D CNN 中提取空域特征,两种特征融合后输入到全连接层与 Softmax 层中得到最终分类结果,在实际场景下进行的实验中,对三种入侵事件的检测率  $>90\%$ , NAR 约为 10%。Wang<sup>[83]</sup>提出了端到端的多支长短时记忆卷积神经网络(Multi-branch LSTM CNN, MLSTM-CNN),使用不同移动差分距离和恒等映射处理原始信号得到的数据输入到 MLSTM 中获得不同时间尺度的特征,在 CNN 中将所有分支 LSTM 提取到的特征连接起来并分类,对 5 种事件的平均识别率达到 95.7%, NAR 为 4.3%。Wu<sup>[84]</sup>构建了 1DCNNs-BiLSTM 网络结构,一个传感点上的信号经过预

处理并输入到单独的1D-CNN网络中,以提取局部和全局结构特征,该网络对应于LSTM的一个单元;然后将提取的1D-CNN特征向量并行馈入双向LSTM网络中,以继续挖掘不同节点处信号之间的空间关联。最后,将提取的时空特征输入到全连接层以识别事件类型,分类准确率达到97%。

综上所述,不同扰动类型信号的频谱图存在差异意味着从图像层面可以识别扰动类型,且原始信号的波形与声音信号有一定的相似之处,研究者大多是将图像处理、语音识别领域的成熟的神经网络应用在 $\Phi$ -OTDR信号处理端。通过调研文献发现,从使用短时傅里叶变换得到频谱图及提取MFCC作为模型输入到仅对原始信号进行降噪处理甚至不做处理直接输入到分类模型中,研究者在使用深度学习算法时越来越关注信号的“端到端”、“自动处理”,即信号在进入分类模型前尽可能减少信号预处理,减少对分类的人为干扰,进而减少工作量,并获得更出色的准确率和NAR指标。此外,研究者在选择分类模型时,更关注原始信号在时域和空域上的特征,例如:利用CNN提取时空矩阵的特征、利用LSTM提取信号的时序性以及利用1D-CNN提取信号的空域双向相关性等。

## 5 讨论

在介绍 $\Phi$ -OTDR系统结构及原理后,本综述在第三节中介绍了用于 $\Phi$ -OTDR信号处理的各种降噪算法及其原理,在第四节中介绍了用于信号分类的各类特征的提取方法与其组合使用的传统机器学习算法,本综述还介绍了近年来流行的基于深度学习的扰动识别方案。表3整理了第四节中提到的 $\Phi$ -OTDR信号的分类算法,比较了传感距离、分类准确率、事件种类、NAR、处理时间等性能参数。尽管由于不同分类算法各自应用的场景不同(传感距离不同、扰动类型不一致等)、实验系统不同、数据集不同,其效果无法进行更深入、准确量化的比较,但分析各自的实验结果可知,无论是传统机器学习算法还是

深度学习算法,其在 $\Phi$ -OTDR系统上的表现均很优秀。在小样本数据、扰动事件种类少的情况下,传统机器学习算法的表现好,平均扰动识别准确率 $>90\%$ 。其中,多个弱分类器集成的强分类器在系统NAR上表现较好,如<sup>[26]</sup>在油气管道监测中使用三个分类器,当三个分类器输出结果均为人为挖掘偷油事件时才预警,严格的预警条件使NAR降低到了0%。在不考虑前期特征选择耗时的情况下,传统机器学习算法由于复杂度低,处理时间较短,最低达到 $0.093\text{ s}$ <sup>[61]</sup>。在多分类情况下,神经网络的平均准确率要优于大部分的传统机器学习算法,扰动识别的平均准确率可高达 $98\%$ <sup>[75]</sup>,而传统机器学习算法GMM用于信号8分类的平均准确率极低,仅为 $69.7\%$ <sup>[22]</sup>。此外,深度学习算法由于不需要设计特征提取模块,其应用于不同场景下的 $\Phi$ -OTDR系统的迁移性较传统机器学习算法好,在实时监测多种扰动事件的 $\Phi$ -OTDR系统上更有优势。

从作者研究中来看,部分小样本、特征不明显的事件识别率还可以进一步提升;系统误报率、系统实时性等指标还有一定的优化空间。现有分类算法大多暂未考虑 $\Phi$ -OTDR信号的时空双关性,研究者可以充分利用这一特性设计合理的分类算法,进一步改善信号处理算法的性能。此外,近年来流行的深度学习算法分类扰动需要较长时间,训练大量数据且需要对网络结构与参数做出适当的调整以获得最优的分类效果,考虑到目前不同系统采集到的数据格式不一致,各自使用的神经网络输入的维度也不一样,现有算法的迁移性和普适性仍不够好。如果能够在花费少量时间的前提下充分利用神经网络迁移学习的能力,使神经网络学习到不同场景下 $\Phi$ -OTDR信号的共有特征,深度学习算法很可能在 $\Phi$ -OTDR信息提取中有更出色的表现。最后,从各科研团队的数据集来看,数据集还较小,数据格式并未统一,代码尚未见公开,如果能规范本领域数据集格式,创建统一的、公开的数据集,并公开代码,相信对 $\Phi$ -OTDR数据处理研究发展、产业化应用定有所裨益。

表 3 分类算法性能比较  
Tab. 3 Classification algorithm performance comparison

文献	分类算法	传感距离/km	分类准确率	事件种类	误报率	识别时间/s
[65]	ANN		95%	3		
[20]	ANN		89.19%	3	1.75%	
[21]	ANN	65	94.4%	3	6.7%	
[29]	RVM	20	97.8%	3		<1
[30]	RVM	10	88.6%	3		
[23]	SVM	20	93.8%	5		
[24]	二叉树结构的 SVM	25.05	>94%	4	<4%	
[19]	改进 SVM		95.72%	3		
[27]	NC-SVM	25.05	94.3%	5	5.62%	0.55
[56]	GMM	45	80%	2		
[22]	GMM	45	69.7%	8	31.2%	
[59]	RF		98.67%	2		
[58]	RF		96.58%	4		
[61]	XGBoost	25.05	95.9%	5	4.1%	0.093
[60]	ELM	10	93.9%	5		1.070 6
[57]	F-ELM	25.05	95.3%	5	4.67%	<0.1
[26]	SVM+BPNN+贝叶斯网络	65	98%	2	0%	
[25]	SVM(Ⅰ级分类器)	10	99.41%	2	0.69%	
	RF/GBDT(Ⅱ级分类器)		98.02%/97.89%	4		
[28]	SVM+BPNN+AdaBoost	25.05	94.1%	5		
[31]	2D CNN+SVM	40	93.3%	4		0.6
[32]	2D CNN		98.02%			
[33]	2D CNN	1	96.67%	5		
	(改进的 Googlenet)					
[74]	2D CNN+SVM	1	94.17%	8		
[75]	2D CNN	2	>98%	5		
[34]	1D CNN	34	95.55%	5		
[35]	MS 1-D CNN	40	96.59%	3		
[77]	MS CNN		84.67%	6		0.017
[76]	AF-CNN		96.7%	4		
[38]	ALSTM	50	94.3%	5		0.91
[80]	CLDNN	33	97.2%	3	8%	
[81]	ConvLSTM	40	85.6%	3		8.25
[82]	3D CNN+ConvLSTM	40	>90%	3	~10%	
[83]	MLSTM-CNN	25.05	95.7%	5	4.3%	1.20
[84]	1D-CNNs+Bi LSTM	40	97%	5		

6 结 论

针对目前  $\Phi$ -OTDR 测量距离长、空间分辨率高、测量频率高等带来的海量数据,信号处理技术将是未来的重要发展方向。本文综述了近

年来用于  $\Phi$ -OTDR 系统的信号处理方法,信号处理算法无需增加硬件设备,即可改善  $\Phi$ -OTDR 性能指标,包括:提高信号质量的降噪算法、分类扰动的各类机器学习算法以及相应的特征提取算法。本文的调研发现,降噪算法大多与分类算



法配合使用,旨在提高扰动识别的准确率;传统分类算法计算量小,在小样本、种类少的扰动事件识别上表现好,但往往需要人为设计特征提取方案;深度学习算法更适用于多扰动分类,准确

率更高,鲁棒性更强,本文还对比了各分类算法的传感距离、准确率等参数,进而分析了不同分类算法在不同场景下的优劣,展望了未来 $\Phi$ -OTDR系统信号处理的发展方向。

#### 参考文献:

- [1] JUAREZ J C, TAYLOR H F. Field test of a distributed fiber-optic intrusion sensor system for long perimeters [J]. *Applied Optics*, 2007, 46 (11): 1968-1971.
- [2] QIN Z G, CHEN L, BAO X Y. Continuous wavelet transform for non-stationary vibration detection with phase-OTDR [C]. *Proc SPIE 8421, OFS2012 22nd International Conference on Optical Fiber Sensors*, 2012, 8421: 8421A0.
- [3] JUAREZ J C, MAIER E W, CHOI K N, *et al.* Distributed fiber-optic intrusion sensor system [J]. *Journal of Lightwave Technology*, 2005, 23 (6): 2081-2087.
- [4] JUAREZ J C, TAYLOR H F. Polarization discrimination in a phase-sensitive optical time-domain reflectometer intrusion-sensor system [J]. *Optics Letters*, 2005, 30(24): 3284-3286.
- [5] QIN Z G, CHEN L, BAO X Y. Wavelet denoising method for improving detection performance of distributed vibration sensor [J]. *IEEE Photonics Technology Letters*, 2012, 24(7): 542-544.
- [6] CHOI K N, TAYLOR H F. Spectrally stable Er-fiber laser for application in phase-sensitive optical time-domain reflectometry [J]. *IEEE Photonics Technology Letters*, 2003, 15(3): 386-388.
- [7] RAO Y J, LUO J, RAN Z L, *et al.* [C]. 2009 20th International Conference on Optical Fibre Sensors, Edinburgh UK, 2009: p750310.
- [8] MARTINS H F, MARTÍN-LÓPEZ S, CORREDERA P, *et al.* Phase-sensitive optical time domain reflectometer assisted by first-order Raman amplification for distributed vibration sensing over >100 km [J]. *Journal of Lightwave Technology*, 2014, 32 (8): 1510-1518.
- [9] PENG F, WU H, JIA X H, *et al.* Ultra-long high-sensitivity  $\Phi$ -OTDR for high spatial resolution intrusion detection of pipelines [J]. *Optics Express*, 2014, 22(11): 13804-13810.
- [10] MARTINS H F, MARTIN-LOPEZ S, FILOGRANO M L, *et al.* Comparison of the use of first and second-order Raman amplification to assist a phase-sensitive optical time domain reflectometer in distributed vibration sensing over 125 km [C]. *Proc SPIE 9157*, 2014, 9157: 91576K.
- [11] LU Y L, ZHU T, CHEN L, *et al.* Distributed vibration sensor based on coherent detection of phase-OTDR [J]. *Journal of Lightwave Technology*, 2010, 28(22): 3243-3249.
- [12] WANG Z N, LI J, FAN M Q, *et al.* Phase-sensitive optical time-domain reflectometry with Brillouin amplification [J]. *Optics Letters*, 2014, 39 (15): 4313-4316.
- [13] WANG Z N, ZENG J J, LI J, *et al.* Ultra-long phase-sensitive OTDR with hybrid distributed amplification [J]. *Optics Letters*, 2014, 39 (20): 5866-5869.
- [14] PAN Z Q, LIANG K Z, YE Q, *et al.* Phase-sensitive OTDR system based on digital coherent detection [C]. 2011 Asia Communications and Photonics Conference and Exhibition (ACP). November 13-16, 2011, Shanghai, China. IEEE, 2011: 1-6.
- [15] HE H J, SHAO L Y, LUO B, *et al.* Multiple vibrations measurement using phase-sensitive OTDR merged with Mach-Zehnder interferometer based on frequency division multiplexing [J]. *Optics Express*, 2016, 24(5): 4842-4855.
- [16] HE Q, ZHU T, XIAO X H, *et al.* All fiber distributed vibration sensing using modulated time-difference pulses [J]. *IEEE Photonics Technology Letters*, 2013, 25(20): 1955-1957.
- [17] SHAN Y Y, DONG J Y, ZENG J, *et al.* A broadband distributed vibration sensing system assisted by a distributed feedback interferometer [J]. *IEEE Photonics Journal*, 2018, 10(1): 1-10.
- [18] ZHU T, HE Q, XIAO X H, *et al.* Modulated pulses based distributed vibration sensing with high frequency response and spatial resolution [J]. *Optics Express*, 2013, 21(3): 2953-2963.
- [19] 曲洪权, 夏雨, 毕福昆. 一种基于改进型SVM算

- 法的光纤入侵信号识别研究[J]. 北方工业大学学报, 2017, 29(2): 33-38.
- QU H Q, XIA Y, BI F K. An improved SVM method to recognize harmful intrusion signal for optical fiber pre-warning system [J]. *Journal of North China University of Technology*, 2017, 29(2): 33-38. (in Chinese)
- [20] WU H J, XIAO S K, LI X Y, *et al.* Separation and determination of the disturbing signals in phase-sensitive optical time domain reflectometry ( $\Phi$ -OTDR) [J]. *Journal of Lightwave Technology*, 2015, 33(15): 3156-3162.
- [21] QIAN Y, WU H J, ZHANG W, *et al.* Feature extraction with WD and WPD in distributed optical-fiber vibration sensing system for oil pipeline safety monitoring[C]. *Asia Pacific Optical Sensors Conference*. Shanghai. Washington, D. C.: OSA, 2016: W4A. 37.
- [22] TEJEDOR J, MACIAS-GUARASA J, MARTINS H, *et al.* A novel fiber optic based surveillance system for prevention of pipeline integrity threats[J]. *Sensors*, 2017, 17(2): 355.
- [23] 徐铖晋. 分布式光纤传感系统的信号处理技术研究[D]. 杭州: 浙江大学, 2017.
- XU CH J. *Research on Signal Processing Technology of Distributed Optical Fiber Sensing System* [D]. Hangzhou: Zhejiang University, 2017. (in Chinese)
- [24] 张俊楠, 姜淑琴, 梁生. 基于SVM算法的 $\varphi$ -OTDR分布式光纤扰动传感系统模式识别研究[J]. 红外与激光工程, 2017, 46(4): 0422003.
- ZHANG J N, LOU SH Q, LIANG SH. Study of pattern recognition based on SVM algorithm for  $\varphi$ -OTDR distributed optical fiber disturbance sensing system [J]. *Infrared and Laser Engineering*, 2017, 46(4): 0422003. (in Chinese)
- [25] 付群健. 分布式光纤振动传感系统模式识别方法研究[D]. 长春: 吉林大学, 2019.
- FU Q J. *Research on Pattern Recognition Method of Distributed Optical Fiber Vibration Sensing System* [D]. Changchun: Jilin University, 2019. (in Chinese)
- [26] 张伟. 基于分布式光纤振动传感器的管道监测信号处理方法[D]. 成都: 电子科技大学, 2017.
- ZHANG W. *Methods of Pipeline Monitoring Signal Processing Based on Distributed Optical Fiber Vibration Sensor* [D]. Chengdu: University of Electronic Science and Technology of China, 2017. (in Chinese)
- [27] JIA H Z, LIANG S, LOU S Q, *et al.* A k-nearest neighbor algorithm-based near category support vector machine method for event identification of  $\Phi$ -OTDR [J]. *IEEE Sensors Journal*, 2019, 19(10): 3683-3689.
- [28] 姚媛媛. 分布式光纤传感系统的振动信号识别研究[D]. 北京: 北京交通大学, 2020.
- YAO Y Y. *Research on Vibration Signal Recognition of Distributed Optical Fiber Sensing System* [D]. Beijing: Beijing Jiaotong University, 2020. (in Chinese)
- [29] SUN Q, FENG H, YAN X, *et al.* Recognition of a phase-sensitivity OTDR sensing system based on morphologic feature extraction [J]. *Sensors (Basel, Switzerland)*, 2015, 15(7): 15179-15197.
- [30] WANG Y, WANG P F, DING K, *et al.* Pattern recognition using relevant vector machine in optical fiber vibration sensing system [J]. *IEEE Access*, 2019, 7: 5886-5895.
- [31] XU C J, GUAN J J, BAO M, *et al.* Pattern recognition based on time-frequency analysis and convolutional neural networks for vibrational events in  $\varphi$ -OTDR [J]. *Optical Engineering*, 2018, 57(1): 1.
- [32] JIANG F, LI H L, ZHANG Z H, *et al.* An event recognition method for fiber distributed acoustic sensing systems based on the combination of MFCC and CNN [C]. *Proc SPIE* 10618, 2018, 1061: 1061804.
- [33] SHI Y, WANG Y Y, ZHAO L, *et al.* An event recognition method for  $\Phi$ -OTDR sensing system based on deep learning [J]. *Sensors*, 2019, 19(15): 3421.
- [34] CHEN J P, WU H J, LIU X R, *et al.* A real-time distributed deep learning approach for intelligent event recognition in long distance pipeline monitoring with DOFS [C]. 2018 *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*. October 18-20, 2018, Zhengzhou, China. *IEEE*, 2018: 290-296.
- [35] 吴俊, 管鲁阳, 鲍明, 等. 基于多尺度一维卷积神经网络的光纤振动事件识别[J]. 光电工程, 2019, 46(5): 79-86.
- WU J, GUAN L Y, BAO M, *et al.* Vibration events recognition of optical fiber based on multi-

- scale 1-D CNN[J]. *Opto-Electronic Engineering*, 2019, 46(5): 79-86. (in Chinese)
- [36] 于森. 基于双重同源外差相干检测的 $\Phi$ -OTDR系统的研究及应用[D]. 长春: 吉林大学, 2017.
- YU M. *Research and Application of Phase-sensitive Optical Time-domain Reflectometric System Based on Single-source Dual Heterodyne Detection Scheme*[D]. Changchun: Jilin University, 2017. (in Chinese)
- [37] XU C J, GUAN J J, BAO M, *et al.* Pattern recognition based on enhanced multifeature parameters for vibration events in  $\phi$ -OTDR distributed optical fiber sensing system[J]. *Microwave and Optical Technology Letters*, 2017, 59(12): 3134-3141.
- [38] CHEN X, XU C J. Disturbance pattern recognition based on an ALSTM in a long-distance  $\phi$ -OTDR sensing system[J]. *Microwave and Optical Technology Letters*, 2020, 62(1): 168-175.
- [39] QIN Z G, CHEN L, BAO X Y. Wavelet denoising method for improving detection performance of distributed vibration sensor[J]. *IEEE Photonics Technology Letters*, 2012, 24(7): 542-544.
- [40] CHEN W, MA X H, MA Q L, *et al.* Denoising method of the  $\Phi$ -OTDR system based on EMD-PCC[J]. *IEEE Sensors Journal*, 2021, 21(10): 12113-12118.
- [41] QU S, CHANG J, CONG Z H, *et al.* Data compression and SNR enhancement with compressive sensing method in phase-sensitive OTDR[J]. *Optics Communications*, 2019, 433: 97-103.
- [42] ZHU T, XIAO X H, HE Q, *et al.* Enhancement of SNR and spatial resolution in  $\phi$ -OTDR system by using two-dimensional edge detection method[J]. *Journal of Lightwave Technology*, 2013, 31(17): 2851-2856.
- [43] WANG Y, JIN B Q, WANG Y C, *et al.* Real-time distributed vibration monitoring system using  $\Phi$ -OTDR[J]. *IEEE Sensors Journal*, 2017, 17(5): 1333-1341.
- [44] HE H J, SHAO L Y, LI H C, *et al.* SNR enhancement in phase-sensitive OTDR with adaptive 2-D bilateral filtering algorithm[J]. *IEEE Photonics Journal*, 2017, 9(3): 1-10.
- [45] ÖLÇER İ, ÖNCÜ A. Adaptive temporal matched filtering for noise suppression in fiber optic distributed acoustic sensing[J]. *Sensors*, 2017, 17(6): 1288.
- [46] LIEHR S, BORCHARDT C, MÜNZENBERGER S. Long-distance fiber optic vibration sensing using convolutional neural networks as real-time denoisers[J]. *Optics Express*, 2020, 28(26): 39311-39325.
- [47] BOLL S. Suppression of acoustic noise in speech using spectral subtraction[J]. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1979, 27(2): 113-120.
- [48] TROPP J A, GILBERT A C. Signal recovery from random measurements via orthogonal matching pursuit[J]. *IEEE Transactions on Information Theory*, 2007, 53(12): 4655-4666.
- [49] ROBERTS, G L. *Machine Perception of Three-dimensional Solids*[M]. 1965.
- [50] ELAD M. On the origin of the bilateral filter and ways to improve it[J]. *IEEE Transactions on Image Processing*, 2002, 11(10): 1141-1151.
- [51] SOBEL I E. *Camera Models and Machine Perception*[D]. Stanford University, 1970.
- [52] PREWITT. Object enhancement and extraction[J]. *Picture Processing and Psychopictorics*, 1971.
- [53] ZHU Q D, JING L Q, BI R S. Exploration and improvement of Otsu threshold segmentation algorithm[C]. 2010 8th World Congress on Intelligent Control and Automation. July 7-9, 2010, Jinan, China. *IEEE*, 2010: 6183-6188.
- [54] ZHANG M, GUNTURK B K. Multiresolution bilateral filtering for image denoising[J]. *IEEE Transactions on Image Processing*, 2008, 17(12): 2324-2333.
- [55] SHAO L Y, LIU S Q, BANDYOPADHYAY S, *et al.* Data-driven distributed optical vibration sensors: a review[J]. *IEEE Sensors Journal*, 2019, 20(12): 6224-6239.
- [56] TEJEDOR J, MACIAS-GUARASA J, MARTINS H F, *et al.* Towards detection of pipeline integrity threats using a smart fiber optic surveillance system: PIT-STOP project blind field test results[C]. 25th International Conference on Optical Fiber Sensors. Jeju, Korea, Republic of. *SPIE*, 2017.
- [57] JIA H Z, LOU S Q, LIANG S, *et al.* Event identification by F-ELM model for  $\phi$ -OTDR fiber-optic distributed disturbance sensor[J]. *IEEE Sensors Journal*, 2020, 20(3): 1297-1305.

- [58] WANG X, LIU Y, LIANG S, *et al.* Event identification based on random forest classifier for  $\Phi$ -OTDR fiber-optic distributed disturbance sensor [J]. *Infrared Physics & Technology*, 2019, 97: 319-325.
- [59] WANG J, HU Y Z, SHAO Y C. The digging signal identification by the random forest algorithm in the phase-OTDR technology [J]. *IOP Conference Series: Materials Science and Engineering*, 2018, 394: 032005.
- [60] ZHANG M J, LI Y C, CHEN J, *et al.* Event detection method comparison for distributed acoustic sensors using  $\varphi$ -OTDR [J]. *Optical Fiber Technology*, 2019, 52: 101980.
- [61] WANG Z D, LOU S Q, LIANG S, *et al.* Multi-class disturbance events recognition based on EMD and XGBoost in  $\varphi$ -OTDR [J]. *IEEE Access*, 2020, 8: 63551-63558.
- [62] CREMERS D, ROUSSON M, DERICHE R. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape [J]. *International Journal of Computer Vision*, 2007, 72(2): 195-215.
- [63] SANKUR B. Survey over image thresholding techniques and quantitative performance evaluation [J]. *Journal of Electronic Imaging*, 2004, 13 (1) : 146-168.
- [64] BIANCO A, BOENTE G, PIRES A M, *et al.* Robust discrimination under a hierarchy on the scatter matrices [J]. *Journal of Multivariate Analysis*, 2008, 99(6): 1332-1357.
- [65] JIANG L H, LIU X M, YANG R Y. Application of the HHT method to the airport fiber fence warning [C]. 2011 *International Conference on Electronics, Communications and Control (ICECC)*. September 9-11, 2011, Ningbo, China. *IEEE*, 2011: 1337-1340.
- [66] TIPPING M E. Sparse Bayesian learning and relevance vector machine [J]. *Journal of Machine Learning Research*, 2001, 1: 211-244.
- [67] 赵发林, 张涛, 李康. 基于遗传算法的随机森林模型在特征基因筛选中的应用 [J]. *中国卫生统计*, 2016, 33(4): 559-562, 566.
- ZHAO F L, ZHANG T, LI K. An optimized random forest based on genetic algorithm and its application to feature selection for gene data [J]. *Chinese Journal of Health Statistics*, 2016, 33 (4) : 559-562, 566. (in Chinese)
- [68] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine [J]. *The Annals of Statistics*, 2001, 29(5): 1189-1232.
- [69] LECUN Y, BOTTOU L, BENGIO Y, *et al.* Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86 (11): 2278-2324.
- [70] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [71] SZEGEDY C, LIU W, JIA Y Q, *et al.* Going deeper with convolutions [C]. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 7-12, 2015, Boston, MA, USA. *IEEE*, 2015: 1-9.
- [72] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-scale Image Recognition [J]. *Computer Science*, 2014.
- [73] LI H L, ZHANG Z H, JIANG F, *et al.* An event recognition method for fiber distributed acoustic sensing systems based on the combination of MFCC and CNN [C]. 2017 *International Conference on Optical Instruments and Technology: Advanced Optical Sensors and Applications*. October 28-30, 2017. Beijing, China. *SPIE*, 2018: 1061804.
- [74] SHI Y, WANG Y Y, WANG L Y, *et al.* Multi-event classification for  $\Phi$ -OTDR distributed optical fiber sensing system using deep learning and support vector machine [J]. *Optik*, 2020, 221: 165373.
- [75] SUN Q, LI Q J, CHEN L, *et al.* Pattern recognition based on pulse scanning imaging and convolutional neural network for vibrational events in  $\Phi$ -OTDR [J]. *Optik*, 2020, 219: 165205.
- [76] RUAN S S, MO J Q, XU L, *et al.* Use AF-CNN for end-to-end fiber vibration signal recognition [J]. *IEEE Access*, 2021, 9: 6713-6720.
- [77] LYU C G, HUO Z Q, LIU Y G, *et al.* Robust intrusion events recognition methodology for distributed optical fiber sensing perimeter security system [J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-9.
- [78] GIRAUD-CARRIER C, VILALTA R, BRAZDIL P. Introduction to the special issue on meta-learning [J]. *Machine Learning*, 2004, 54



- (3): 187-193.
- [79] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [80] BAI Y, XING J C, XIE F, *et al.* Detection and identification of external intrusion signals from 33 km optical fiber sensing system based on deep learning [J]. *Optical Fiber Technology*, 2019, 53: 102060.
- [81] LI Z Q, ZHANG J W, WANG M N, *et al.* Fiber distributed acoustic sensing using convolutional long short-term memory network: a field test on high-speed railway intrusion detection [J]. *Optics Express*, 2020, 28(3): 2925-2938.
- [82] LI Z Q, ZHANG J W, WANG M N, *et al.* An anti-noise  $\phi$ -OTDR based distributed acoustic sensing system for high-speed railway intrusion detection[J]. *Laser Physics*, 2020, 30(8): 085103.
- [83] WANG Z D, LOU S Q, WANG X, *et al.* Multi-branch long short-time memory convolution neural network for event identification in fiber-optic distributed disturbance sensor based on  $\varphi$ -OTDR[J]. *Infrared Physics & Technology*, 2020, 109: 103414.
- [84] WU H J, YANG M R, YANG S Q, *et al.* A novel DAS signal recognition method based on spatio-temporal information extraction with 1DCNNs-BiLSTM network [J]. *IEEE Access*, 2020, 8: 119448-119457.

#### 作者简介:



田曼伶(1998—),女,湖南湘西人,硕士生,2019年于北京语言大学获得学士学位,现为北京交通大学计算机与信息技术学院硕士生,研究方向为分布式光纤传感,深度学习和信号处理。  
E-mail: 19120309@bjtu.edu.cn



刘东辉(1996—),男,河北保定人,硕士生,2019年于河北科技大学获得学士学位,现为北京交通大学计算机与信息技术学院硕士生,研究方向为分布式光纤传感的信号处理,机器学习。  
E-mail: 19125140@bjtu.edu.cn