

◇ 研究报告 ◇

应用原型网络的小样本次声信号分类识别方法*

赵子杰^{1,2,3} 程巍^{1,2} 姬培锋^{1,2} 滕鹏晓^{1,2} 吕君^{1,2} 杨军^{1,2,3†}

(1 中国科学院声学研究所 北京 100190)

(2 中国科学院噪声与振动重点实验室 北京 100190)

(3 中国科学院大学 北京 100049)

摘要: 地震、闪电、火箭发射、爆炸等活动都会伴随着次声信号的产生。为提升次声事件的监测能力,需要对小样本的次声信号进行正确分类识别。针对小样本集的次声事件的有效识别问题,结合长短期记忆模型提出了一种应用原型网络的次声信号分类方法。使用该方法分别对公开的次声信号数据集和实地采集的地震、爆炸、闪电、火箭再入产生的 4 类次声信号进行分类实验。实验结果表明,该方法相对于传统方法,简化了特征提取的过程,有效解决了小样本集次声信号的特征分析问题,取得较好的分类结果和泛化效果。

关键词: 次声; 小样本; 原型网络; 长短期记忆模型

中图法分类号: TN911.7

文献标识码: A

文章编号: 1000-310X(2024)06-1193-10

DOI: 10.11684/j.issn.1000-310X.2024.06.002

A method for classification of few-shot infrasound signals applying prototype network

ZHAO Zijie^{1,2,3} CHENG Wei^{1,2} JI Peifeng^{1,2} TENG Pengxiao^{1,2}

LYU Jun^{1,2} YANG Jun^{1,2,3}

(1 *Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China*)

(2 *Key Laboratory of Noise and Vibration Research, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China*)

(3 *University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: Events such as earthquakes, lightning, rocket launches, and explosions are accompanied by infrasound signals. In order to improve the monitoring capability of infrasound events, it is necessary to correctly classify small samples of infrasound signals. For the problem of effective identification of infrasound events with small samples and variable duration, a classification method of infrasound signals applying prototype network is proposed in combination with a long and short-term memory model. The method is used to conduct classification experiments on publicly available infrasound signal datasets and four types of infrasound signals generated by earthquakes, explosions, lightning, and rocket re-entry collected in the field. The experimental results show that the method simplifies the process of feature extraction and effectively solves the problem of feature analysis of variable duration infrasound signals compared with the traditional method, and achieves better classification results and generalization effects.

Keywords: Infrasound; Few-shot; Prototypical network; Long short-term memory

2023-02-28 收稿; 2023-04-07 定稿

*国家自然科学基金项目 (11874389)

作者简介: 赵子杰 (1997-), 男, 河北衡水人, 硕士研究生, 研究方向: 信号与信息处理。

†通信作者 E-mail: jyang@mail.ioa.ac.cn

0 引言

次声是指频率低于 20 Hz 的声波,具有能量衰减小、穿透能力强、传播距离远的特点^[1]。地震、泥石流、闪电、火山爆发等自然活动,化学爆炸、火箭发射等人类活动的发生过程中,都会伴随着低频次声信号的产生。利用次声对上述事件进行监测,对灾害预防和国防安全具有重要意义。次声信号分类作为次声事件监测的重要组成部分,是次声领域研究的重点。区别于其他领域的信号识别,低频次声信号易受环境风等背景噪声的影响,分类难度较大,同时次声事件样本稀缺,难以应用基于大规模样本的分类模型。因此,次声信号分类识别研究的难点在于小样本的前提下,提高次声事件的识别准确率。

在次声信号小样本分类识别研究的开始,研究人员主要应用信号分析处理方法,对次声信号的时域和频域进行特征分析,提取区分度明显的次声信号特征进行分类。1998年,Huang等^[2-3]提出了经验模态分解算法,引入希尔伯特谱分析,成为后来次声信号特征分析的主要方法,推动了次声信号识别的发展。2007年,Chilo等^[4]使用离散小波变换、连续小波变换的时间尺度谱结合倒谱系数、峭度等对车辆、爆炸、关门声、陨石产生的次声信号进行了特征提取,并用BP神经网络进行分类,比较了3种特征提取技术的优劣。Ham等^[5]利用倒谱系数作为特征,结合神经网络的方法对火山、火箭发射、大风等次声事件信号进行了分类实验。Liu等^[6]、Li等^[7]针对泥石流等地质自然灾害产生的次声信号,应用小波变换、希尔伯特黄变换,结合BP神经网络和支持向量机等识别算法,在特定的数据集上具有较高的识别准确率。胡志华等^[8]在经验模态分解的基础上采用集成经验模态分解的方法,将分形盒维数作为特征,对不同的类型的泥石流次声信号进行了研究。近年来随着深度学习方法在模式识别领域的优异表现,也有学者应用深度学习的方法对次声信号分类进行了研究,吴涓晖^[9-10]针对次声数据少的问题,采用生成对抗网络生成一批伪次声信号,扩充了数据集,同时将次声信号时频图送入卷积神经网络,取得了很好的识别效果。谭笑枫等^[11]则是采用一维卷积神经网络的方法对次声信号的Welch功率谱进行了分析和研究。戴翊靖等^[12]采用Fisher检测和非负矩阵分解的方法对火箭发射的次声信号

进行了分析。孟子轩^[13]在此基础上采用基于非负矩阵分解的特征提取方法,自动挖掘目标信号的隐含结构,后通过支持向量机和卷积神经网络进行分类实验,对4类小样本次声信号数据进行分类实验,取得了83.13%的平均识别准确率。近年来,随着数据科学研究的深入,小样本成为人工智能领域绕不开的问题,学者们提出了多种方法尝试解决该问题,基于度量的元学习方法就是解决小样本问题的重要突破。基于度量的元学习方法通过学习相似性度量来“学习”学习策略。由于其具有鲁棒性强、学习速度快、可解释性强等优点,在许多现实世界的应用中具有很高的实用价值。

现有的小样本次声信号分类方法大多需要对次声信号进行时域、频域或时频域特征分析。时域上采集到的次声信号常包含大量的背景噪声,频域上多数次声事件的频率集中在5 Hz以下,目标次声信号与噪声信号在低频窄带中混叠严重,因此对次声信号的特征分析都需要进行大量的工作。采用简单的卷积神经网络对次声信号进行编码,虽然可以深度挖掘次声信号的特征,避免了特征分析的过程,但是深度学习的方法常需要大量的样本,小样本的次声信号并不能满足该条件。

针对以上问题,本文提出了一种应用原型度量的小样本次声信号分类方法,将原型网络^[14]引入次声信号分类模型中,将深度学习中的长短期记忆(Long short-term memory, LSTM)^[15]模型与原型网络相结合,作为原型网络的编码单元,解决了可变时长次声信号编码问题,避免了复杂的特征分析过程。通过编码计算次声信号的嵌入表示,进一步计算每类次声信号的原型表示,即各类次声信号的类中心。采用欧几里得距离度和余弦相似度度量的方式,计算次声信号编码后的嵌入表示与各类次声信号的原型表示(类原型)的度量距离,进一步地对次声信号样本进行预测。在小样本次声信号的情况下,有效提高了次声信号的识别准确率。

1 分类模型构建

分类模型的构建主要包括基于“演绎法”的规则(专家)系统与基于“归纳法”的机器(统计)学习系统两类。基于“演绎法”的规则(专家)系统,如支持向量机等,需要进行复杂的特征分析,提取区分度明显的次声信号特征,使次声信号分类研究变得复杂。因此本文考虑采用基于“归纳法”的机器(统计)

学习系统构建分类模型。根据实际应用情况，针对可变时长的次声信号构建四分类模型，模型应当解决可变时长次声信号统一处理，对次声信号进行特征编码和次声信号最终分类判别的问题。

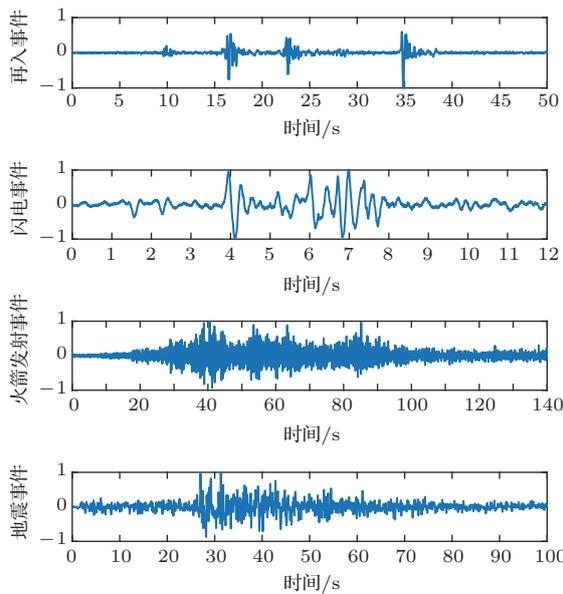
1.1 次声信号概述

首先，激发次声事件的声源不同，导致次声信号的持续时间，即长度也不相同。因此在次声事件分类模型中应当考虑不同持续时间的次声信号，充分利用每段次声信号包含的全部信息，避免截取次声信号片段导致信息不全，从而影响识别准确率。不同类别的次声信号波形如图1(a)所示，时长如表1所示。

此外，由于次声源、传输环境和传输距离的不同，接收到的次声信号的波形也不同，反映它们的频谱结构不同。同一类次声(化爆)的波形如图1(b)所示。多数次声事件的频率在5 Hz以下。

表1 次声信号持续时间
Table 1 Duration of infrasound signals
(单位: s)

类型	事例1	事例2	事例3	事例4
地震	7	1.5	3	4
闪电	2	3	2.5	4
火箭再入	4	2	3	6
爆炸	3.5	10	6	2



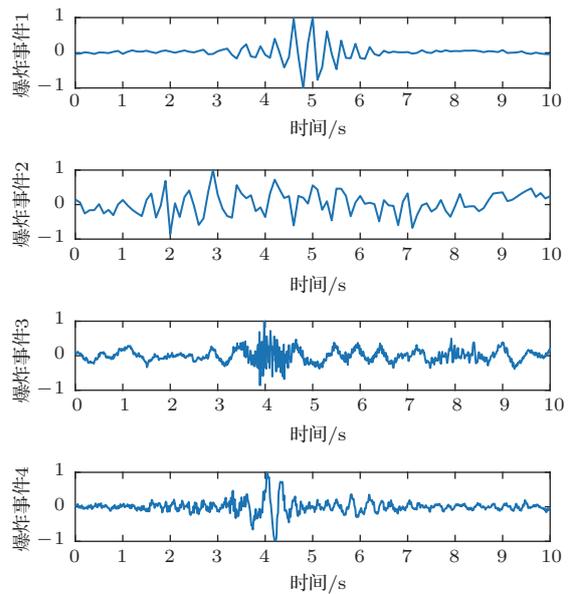
(a) 异类次声事件持续时间

根据图1和表1,次声信号持续时间较长且持续时间不尽相同。采用传统方法的可变时长次声信号分类模型,主要有以下几个问题:(1)对可变时长的次声信号,通常采用对数据进行截取的方法,得到相同长度的信号片段,进而进行后续操作,增加了数据预处理阶段的工作量。(2)次声信号通常是衰减的,因此长时次信号的前后时间顺序关联的信息也是次声信号的一个重要特征,对信号进行截取,则会舍去这一关键特征,造成该类型样本识别率不高。

针对上述问题,提出一种针对于可变时长次声信号分类的模型结构,既可以无需考虑次声信号的持续时间,一定程度上保留时间序列的前后关联的信息,也无需进行复杂的特征分析和特征提取过程。第一个需求需要借助循环神经网络的优势,第二个需求则通过原型网络对次声信号进行特征编码与度量计算。基于以上考虑,本文通过循环神经网络编码得到次声信号特征,应用原型网络的方法,构建次声事件分类器。具体方法和模型结构将在1.2节进行说明。

1.2 基于LSTM和原型网络的次声事件分类模型

基于LSTM和原型网络的次声事件分类模型流程图如图2所示。



(b) 同类次声信号持续时间

图1 异类、同类次声事件持续时间比较

Fig. 1 Duration of infrasound signals

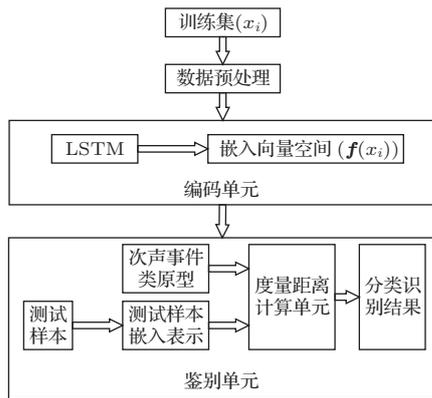


图2 基于LSTM和原型网络的次声事件分类模型流程图

Fig. 2 Flow chart of a prototype network-based infrasound event classification model

模型的结构主要包含编码单元与度量计算单元。输入的次声信号数据 x_i 通过 LSTM 编码后得到样本在嵌入空间的向量表示 $f(x_i)$ 。原型网络是一种基于度量的元学习方法^[16]，是一种简单高效的小样本学习方法，该网络的目标是学习到一个向量空间来实现分类任务。基于原型网络的次声事件分类模型的基本思想是对每一个需要分类的次声事件来创建一个次声事件类中心(原型表示)。对于一个需要进行分类的查询样本，采用度量计算各分类类别的原型表示与查询样本点的距离来进行确定。因此，将上述次声信号样本的嵌入向量 $f(x_i)$ 的均值作为各自的次声事件类原型。表达式如式(1)所示：

$$c_k = \frac{1}{m} \sum_{i=1}^m f(x_i), \quad (1)$$

其中， c_k 表示第 k 类的原型， m 表示第 k 类样本训练集的个数。

通过度量计算单元，计算各个次声信号样本 x_i 的嵌入向量 $f(x_i)$ 到每个次声事件类原型 c_k 的距离 $d(f(x_i), c_k)$ ，利用 softmax 生成次声信号样本 x_i 为各类别的概率。表达式如(2)所示：

$$p(y = k|x_i) = \frac{\exp(-d(f(x_i), c_k))}{\sum_{k'} \exp(-d(f(x_i), c_{k'}))}. \quad (2)$$

进一步的，得到损失函数，如式(3)所示：

$$J = -\lg(p(y = k|f(x_i))). \quad (3)$$

分类模型中的编码单元采用 LSTM 实现。由于次声信号具有长时的特征，在分类任务中需要兼顾

次声信号的上下文特征。同时不同类别、不同强度的次声源的次声信号持续时间并不相同，如地震产生的次声信号通常会持续 20 s 以上，而闪电产生的相关次声信号通常只会持续几秒钟的时间。为解决可变时长次声信号的编码问题，充分利用长时次声信号携带的信息，本方法采用 LSTM 模型作为基于原型网络的次声事件分类模型的编码单元，对输入次声信号进行特征编码。LSTM 网络是一种时间循环神经网络，但与一般的循环神经网络不同，LSTM 网络解决了一般的循环神经网络存在的长期依赖问题，形式上 LSTM 网络具有一种重复神经网络模块的链式结构。LSTM 模型结构框图如图 3 所示。

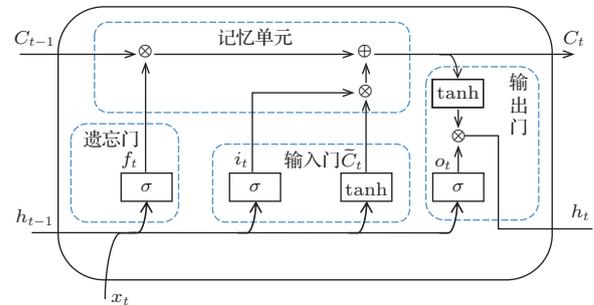


图3 LSTM网络结构

Fig. 3 Long and short-term memory network

标准 LSTM 单元由遗忘门、输入门、输出门和记忆单元组成。遗忘门的作用是决定从上一个细胞状态中丢弃什么信息。如式(4)所示：

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (4)$$

其中， h_{t-1} 为上一时刻的隐层输出， x_t 表示 t 时刻输入， W_f 为待训练参数， b_f 为偏置项， $\sigma(\cdot)$ 表示 sigmoid 激活函数。

输入门的作用是决定 t 时刻有多少新的信息需要记忆，包括两部分 i_t 与 \tilde{C}_t ，如式(5)、式(6)所示：

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (5)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C). \quad (6)$$

记忆单元的输出 C_t 为遗忘门与输入门的结合，如式(7)所示：

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t. \quad (7)$$

LSTM 的输出为输出门 o_t 与记忆单元 C_t 的结合，如式(8)、式(9)所示：

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (8)$$

$$h_t = o_t * \tanh(C_t). \quad (9)$$

次声信号以帧级为单位输入至编码单元 LSTM 中，可建立针对于次声信号的长期的上下文关系，即充分利用当前次声信号帧与先前次声信号帧的特征关系。同时可以采用 LSTM 作为原型网络的编码单元，可以灵活地处理不同长度的次声信号序列，提高了次声信号编码的适应性和泛化能力。

1.3 次声信号分类器训练流程

常见的多分类器训练方法为设置训练集和测试集，通过训练集对应用原型度量方法的次声信号分类器进行优化，并在测试集中验证次声信号分类器的效果。程序流程如图 4 所示。

训练阶段：

(1) 预处理后的训练集数据有着不同的帧数，将这些数据输入至 LSTM 中，数据的每帧特征都会得到对应的隐层输出；

(2) 每个样本数据的最后一帧的隐层输出作为该数据的嵌入表示进行后续操作；

(3) 在有标签的情况下，对同一标签类别的样本求均值，得到各类次声事件的原型表示；

(4) 对每个样本的嵌入表示与各类次声事件的原型表示作度量计算；

(5) 根据度量计算的结果，对各个训练样本进行预测，得到预测标签；

(6) 将预测标签与真实标签做交叉熵损失，根据交叉熵损失进行优化；

(7) 优化至收敛后，最终得到各类次声事件的原型表示。

测试阶段：

(1) 将预处理后的待测样本数据输入至在训练阶段训练好的 LSTM 中，得到各帧特征对应的隐层输出；

(2) 待测样本数据的最后一帧的隐层输出作为该数据的嵌入表示；

(3) 将待测样本数据的嵌入表示与训练阶段得到的各类次声事件的最终原型表示作度量计算；

(4) 根据度量计算的结果，对待测样本进行预测，得到该次声信号分类器对该待测样本的预测标签。

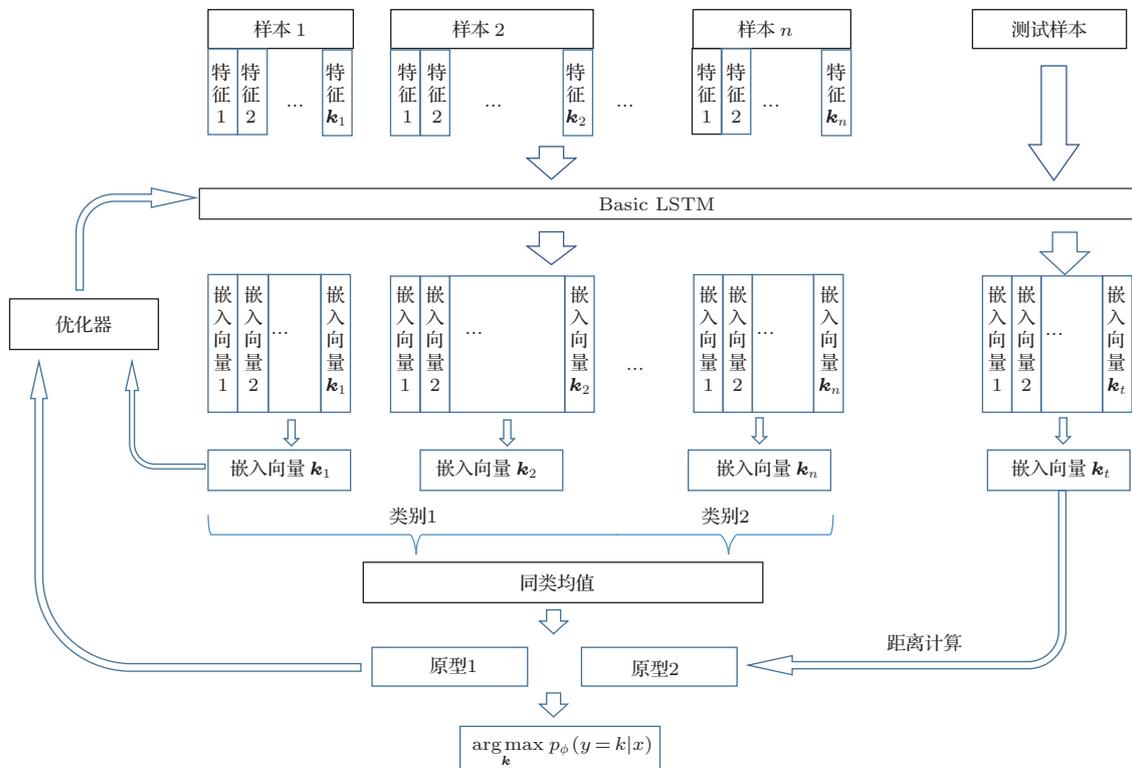


图 4 原型度量次声分类程序流程图

Fig. 4 Flow chart of infrasound classification by prototype metric

2 分类实验与结果分析

在实验开始前需对次声数据进行预处理,包括高通滤波滤除低频噪声,分帧进行短时傅里叶变换。本文使用的数据集中训练集与测试集的比例为7:3,在训练时采用四折交叉验证的方式进行模型选择。识别分类实验包括两部分,分别是公开次声事件数据集的分类识别实验和实地采集的次声事件数据集的分类识别实验。在应用原型度量进行次声信号分类时,除需要确定原型网络编码单元LSTM的相关参数外,还需要考虑度量计算单元的度量函数的选择。本文比较了以欧几里得距离和余弦相似度为度量计算单元的次声信号分类效果以及训练阶段损失函数的收敛速度,以确定度量计算单元的度量函数的最终选择。

2.1 实验次声数据来源

本文实验采用的数据有两个,分别为开源的典型次声事件信号库(Library of typical infrasonic signals, LOTIS)^[17]和实地布阵采集的次声信号。

LOTIS数据集包括极光引发的大气重力波(Atmosphere gravity wave, AGW)115条、山体相关次声(Mountain associated wave, MAW)110条、微气压相关次声(Microbaroms, MB)67条、火山爆发产生的次声(Volcanic eruption, VE)65条,4类次声信号数据。

实采数据集是由次声监测台阵采集到的次声时域信号,包括爆炸次声信号(Explosion, EP)17条、地震次声信号(Earthquake, EQ)22条、闪电信号(Lighting, LT)115条、火箭再入信号(Rocket re-entry, RE)10条。

2.2 LOTIS数据集分类结果及分析

LOTIS作为公开数据集,数据集中同类次声事件样本的长度是一致的,这里对LOTIS数据集进行实验,主要目的是利用拥有更多数据的数据集,确定最终的模型结构,随后在实采数据集中验证本方法的泛化性能。LOTIS数据集训练集、测试集划分如表2所示。

编码单元LSTM不同的隐层层数以及隐层维度,在欧式距离度量单元和余弦相似度度量单元下的平均识别准确率如表3和表4所示(平均识别准确率为在相同条件下进行100次分类实验后得到的平均准确率)。

表2 LOTIS训练集、测试集划分
Table 2 The training set and test set division of LOTIS

	大气重力波 (AGW)	山体相关次声 (MAW)	微气压相关次声 (MB)	火山爆发次声 (VE)
训练集	80	77	47	45
测试集	35	33	20	20
总计	115	110	67	65

表3 欧式距离度量
Table 3 Euclidean distance

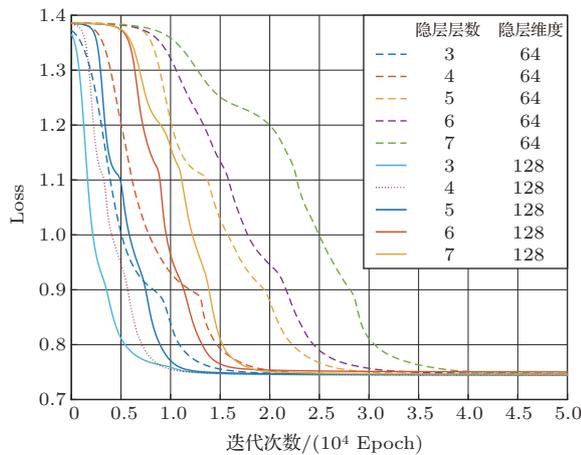
隐层层数	隐层维度 (单位: %)			
	32	64	128	256
1	67.54	74.12	75.36	77.35
2	69.16	73.29	86.11	88.52
3	79.58	92.62	98.32	98.67
4	89.44	95.27	97.89	93.47
5	92.32	96.49	91.65	88.90
6	91.89	95.58	87.53	82.43
7	92.17	96.05	88.26	84.88

表4 余弦相似度度量
Table 4 Cosine similarity

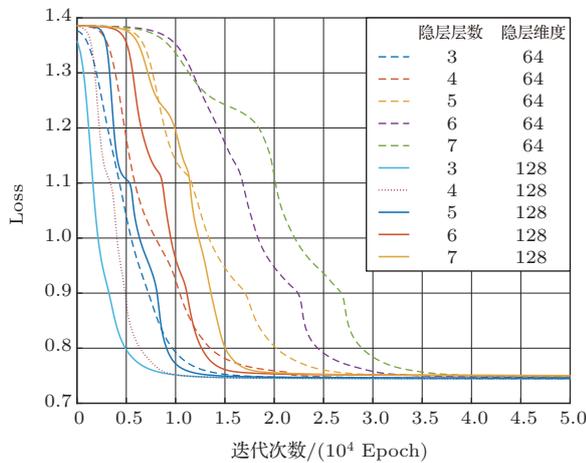
隐层层数	隐层维度 (单位: %)			
	32	64	128	256
1	68.71	74.55	74.68	76.53
2	68.46	74.49	84.08	87.62
3	79.17	90.15	97.96	97.54
4	88.38	96.45	98.15	90.77
5	91.41	96.37	93.39	89.52
6	91.92	95.67	85.47	83.63
7	90.11	96.18	87.62	81.28

LOTIS数据集次声分类实验主要对编码单元LSTM的隐层层数以及隐层维度进行分析。根据表3和表4的实验结果,当编码单元LSTM的隐层层数为3层及以上时,应用原型度量方法的次声信号分类器可以取得较好结果。编码单元LSTM的隐层维度为32维时,分类效果远低于编码单元LSTM采用更高维度时的次声信号分类器,后续分析中不予考虑。当编码单元LSTM的隐层维度为256维时,虽然取得较好的分类效果,但消耗了大量的计算资

源,分类模型效率不高,同时增加了模型复杂度,分类器出现了过拟合现象,后续分析中也不考虑256维以及更高维度的隐层维度。基于以上分析结果,当编码单元LSTM的隐层层数大于等于3层,隐层维度大于等于64维且小于256维时,各次声信号分类器的平均识别准确率相近,进一步的,对具有相近分类效果的分类型器,分析其在训练阶段的收敛速度,进而确定具体参数。下面对设置不同编码单元隐层层数和隐层维度以及不同度量方式的分类器的收敛情况进行分析,如图5所示。



(a) 采用欧式距离度量的各分类器损失曲线



(b) 采用余弦相似度度量的各分类器损失曲线

图5 各个分类器的损失曲线

Fig. 5 Loss curve of classifiers

根据图5中各分类器的损失曲线,各分类器均可以收敛。但不同的编码单元LSTM隐层层数与隐层维度将影响训练阶段的收敛速度。损失函数曲线中虚线代表隐层维度为64时的收敛情况,实线代表隐层维度为128时的收敛情况。如图5所示,隐层维度相同时,隐层层数为3时收敛速度最快,并且随着

隐层层数的增加,收敛速度逐渐变慢,这与隐层层数增加、模型变得更加复杂是相关的。隐层层数大于等于5且隐层维度为128维时,虽然模型在训练集中可以收敛,但结合表3和表4,此时模型在测试集中的识别准确率下降,出现过拟合现象。隐层层数相同时,隐层维度为64时收敛速度要慢于隐层维度为128时的收敛速度。由此说明,特征编码的维度选择128维时,可以得到比64维更具区分度的次声信号嵌入表示。综上所述,分析以欧式距离和余弦相似度为度量单元的分类器,编码单元LSTM的隐层层数为3层、隐层维度为128维时,分类器收敛速度最快,分类效果最好。因此应用原型度量方法将次声信号分类模型编码单元LSTM的参数确定为3层隐层层数,128维隐层维度。

如图6所示,对于度量计算单元,采用欧式距离度量和余弦相似度度量在识别准确率上几乎没有差别,在收敛速度上差距较小,采用余弦相似度度量的方式相对较快,因此后续模型的度量单元将采用余弦相似度的方式进行度量计算。编码单元LSTM采用3层隐层层数,128维隐层维度,度量计算单元采用余弦相似度度量的次声信号分类器的测试集混淆矩阵如图7所示。

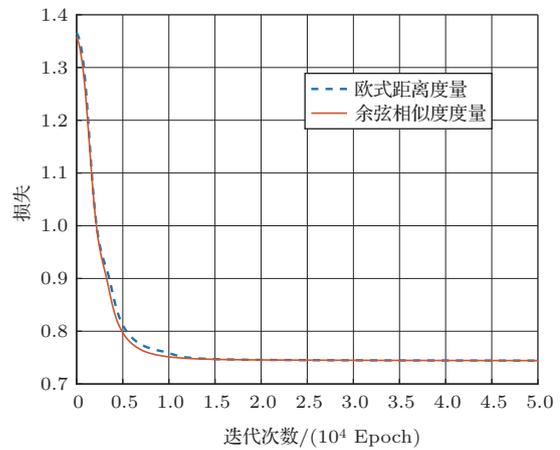


图6 采用欧式距离度量和余弦相似度度量的损失曲线

Fig. 6 Loss curve of Euclidean metric and Cosine similarity

根据图7,迭代次数为0、未采用LOTIS训练集对模型进行训练时,AGW、MAW、VE均与MB有相对明显区分。经过5000次迭代,AGW的识别准确率明显提升,而MAW和VE仍需进一步区分。迭代次数为7500次时,结合图6,此时在训练集中损失已接近收敛,但是除MB外,其他次声事件识别准确率仍有提升的空间。迭代至收敛后,即10000次迭

代,识别准确率可达99.07%,且未出现过拟合,模型效果显著。

结合表4、图5与图6,可以确定最优的应用原型度量方法的次声信号分类器,编码单元采用3层隐层层数,128维隐层维度,度量单元采用余弦相似

度的方式。该次声信号分类器具有较快的收敛速度,同时消耗较少的计算资源。对LOTIS数据集进行100次分类实验(1次分类实验指划分训练集和测试集后,经过训练集训练后的模型对测试集进行一次分类),该分类器可实现97.96%的平均识别准确率。

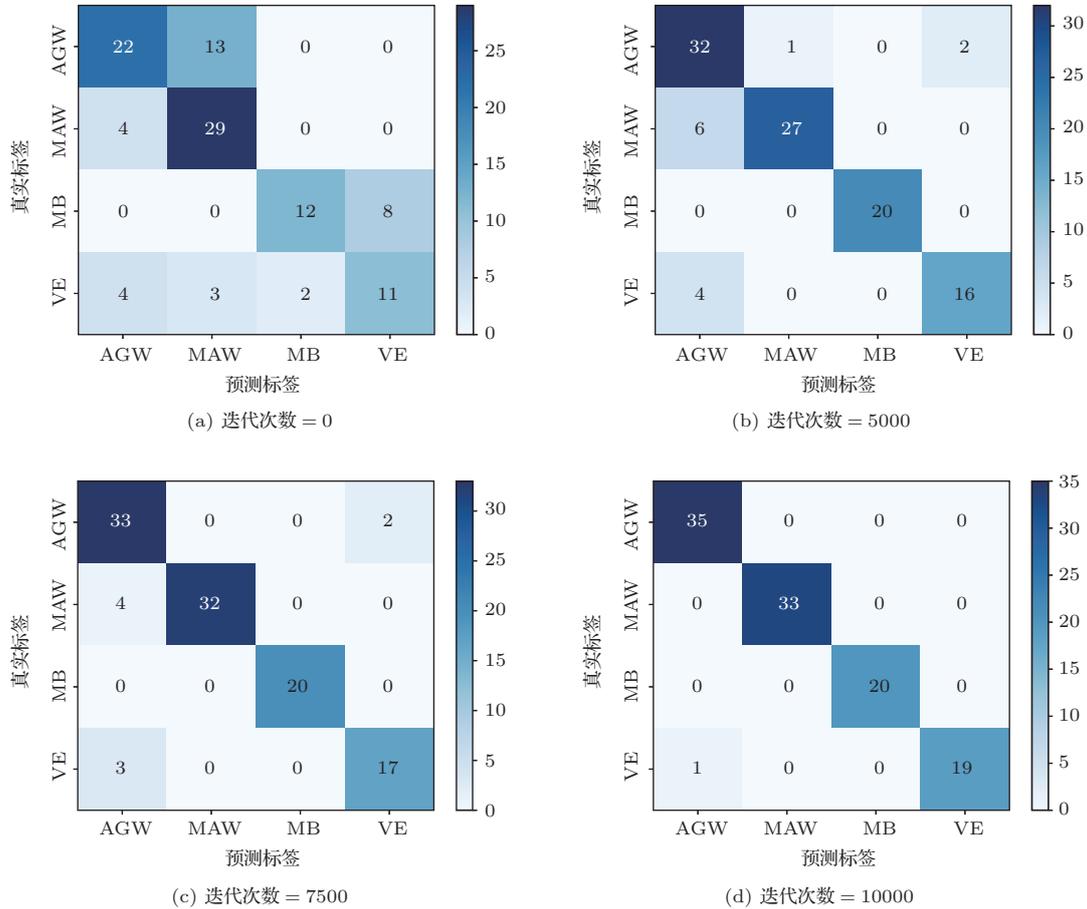


图7 不同迭代次数 LOTIS 测试集结果混淆矩阵

Fig. 7 Confusion matrix of LOTIS test dataset

2.3 实地采集次声数据分类结果及分析

本文使用的实采数据集是通过实地布阵自行采集的次声时域信号。与LOTIS数据集不同,实采数据集包含的4类次声信号数据,均为可变时长的次声信号。在实地采集的数据集中进行应用本文提出的次声信号分类方法,更加具有工程上的实际意义,可以充分验证该方法对于其他各类可变时长次声信号分类的泛化能力。实采数据集训练集、测试集划分如表5所示。

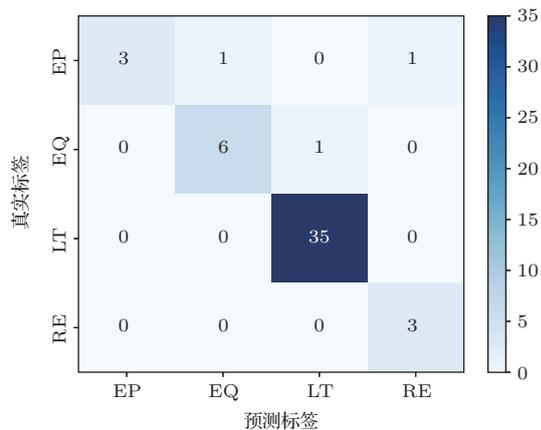
在2.1节,通过对LOTIS数据的分类实验,确定了应用原型度量方法的次声信号分类器的最优模型结构,对实采数据集将继续使用这种模型

结构,即分类器编码单元采用3层隐层、128维隐层维度的LSTM模型,度量单元采用余弦相似度的方式。

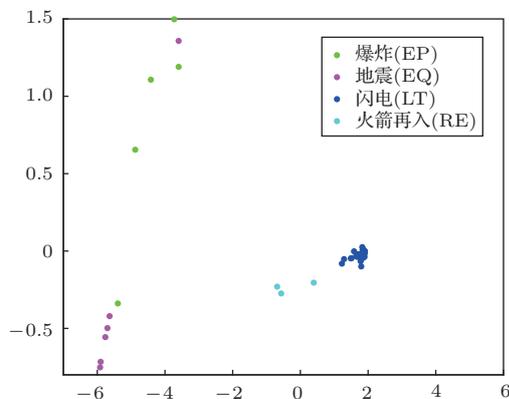
表5 实采训练集、测试集划分
Table 5 The training set and test set division of self-datasets

	爆炸次声信号 (EP)	地震次声信号 (EQ)	闪电次声信号 (LT)	火箭再入次声信号 (RE)
训练集	12	15	80	7
测试集	5	7	35	3
总计	17	22	115	10

图8为一次分类实验后的结果示意图,图8(a)为测试集预测结果与真实标签的混淆矩阵,图8(b)为测试集数据编码后的嵌入表示进行主成分分析降维后的二维散点分布。如图所示,EQ与其他3类次声信号的有着明显区别,分析地震次声信号,其具有持续时间长、能量高的特点,所以经过编码单元编码后的次声信号嵌入表示,与其他3类信号的嵌入表示有着较大的度量距离,因此可实现100%的识别准确率。分析图8(b),LT与EP编码后的嵌入表示有明显的区分度,未出现相互错判的情况。由于采集到的RE较少,用于计算RE原型的样本较少,因此RE的次声信号原型表示未能与LT和EP的类原型完全区分,会有将LT和EP错判为RE的情况,随着RE数据的增加,这种误判情况将有效减少。最后,对实采数据集进行100次分类实验,在测试集上的平均识别准确率可达95.36%。相较于孟子轩^[13]提出的基于非负矩阵分解的次声信号分类方法,在相同的数据集下,平均识别准确率有效提升了10%以上,充分证明本方法的优越性。



(a) 实采数据集一次分类实验测试集结果



(b) 测试集降维后的散点分布

图8 实采数据集分类结果

Fig. 8 Classification result of self-dataset

对于小样本次声信号分类存在的样本分布不均衡的问题应当予以考虑。由于闪电事件发生较多,所以在实采数据集中,LT的样本数远多于其他事件,使得训练集中的数据分布不均衡。分析实采数据集分类实验,不均衡的LT数据,并未造成严重的过拟合问题,即EP、EQ、RE未被大量错判为LT。同时对LT数据集进行降采样(即对数据进行抽样实现各类样本数据均衡),保证数据均衡后进行分类实验,在测试集中的平均识别准确率为95.45%,无明显提升。因此现有的115条LT导致的数据不均衡问题,可以认为对次声信号分类结果无显著影响。

3 结论

本文针对小样本条件下可变时长的次声信号分类识别问题,结合LSTM模型提出了一种应用原型度量的次声信号分类方法。分别在公开的次声信号数据集和实地采集的地震、爆炸、闪电、火箭再入产生的4类次声信号进行分类实验,确定了以3层隐层、128维隐层维度的LSTM模型作为编码单元,以余弦相似度度量作为度量单元的次声信号分类器,并在两个数据集中分别取得了97.96%和95.36%的平均识别准确率。实验证明,应用原型度量的次声信号分类识别方法简化了传统次声信号特征分析的过程,解决了可变时长次声信号的特征提取问题,同时能够提高分类识别结果,取得了较好的泛化效果。

参 考 文 献

- [1] 童娜. 次声的特点及其应用[J]. 声学技术, 2003, 22(3): 199-202.
Tong Na. Features and applications of infrasound[J]. Technical Acoustics, 2003, 22(3): 199-202.
- [2] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings Mathematical Physical & Engineering Sciences, 1998, 454(1971): 903-995.
- [3] Huang E, Shen S. 希尔伯特-黄变换及其应用[M]. 北京: 国防工业出版社, 2017: 4-15.
- [4] Chilo J, Jabor A, Liszka L, et al. Filtering and extracting features from infrasound data[C]. IEEE-NPSS Conference on Real Time. IEEE, 2005.
- [5] Ham F M, Leeney T A, Canady H M, et al. An infrasonic event neural network classifier[C]. IJCNN'99. Inter-

- national Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339), 1999, 6: 3768–3773.
- [6] Liu X, Li M, Tang W, et al. A new classification method of infrasound events using Hilbert-Huang transform and support vector machine[J]. *Mathematical Problems in Engineering*, 2014: 456818.
- [7] Li M, Liu X, Liu X. Infrasound signal classification based on spectral entropy and support vector machine[J]. *Applied Acoustics*, 2016, 113: 116–120.
- [8] 胡至华, 袁路, 马东涛, 等. 基于EEMD分形和LS-SVM的次声信号识别泥石流类型[J]. *山地学报*, 2020, 38(4): 619–629.
- [9] 吴湏晖, 邹士亚, 庞新良, 等. 应用支持向量机和神经网络对大气次声信号识别的初步实验[J]. *应用声学*, 2020, 39(2): 207–215.
Wu Yunhui, Zou Shiya, Pang Xinliang, et al. Experimental study on atmospheric infrasound signal recognition using SVM and ANN[J]. *Journal of Applied Acoustics*, 2020, 39(2): 207–215.
- [10] 吴湏晖, 赵子天, 陈晓雷, 等. 大气低频声信号识别深度学习研究方法研究[J]. *电子科技大学学报*, 2020, 49(5): 758–765.
Wu Yunhui, Zhao Zitian, Chen Xiaolei, et al. Research on deep learning method of atmospheric low frequency acoustic signal recognition[J]. *Journal of University of Electronic Science and Technology of China*, 2020, 49(5): 758–765.
- [11] 谭笑枫, 李夕海, 刘继昊, 等. 基于一维卷积神经网络的化爆和地震次声分类[J]. *应用声学*, 2021, 40(3): 457–467.
Tan Xiaofeng, Li Xihai, Liu Jihao, et al. Classification of chemical explosion and earthquake infrasound based on 1-D convolutional neural network[J]. *Journal of Applied Acoustics*, 2021, 40(3): 457–467.
- [12] 戴翊靖, 滕鹏晓, 吕君, 等. 火箭发射的次声信号分析[J]. *应用声学*, 2021, 40(5): 676–683.
Dai Yijing, Teng Pengxiao, Lyu Jun, et al. Analysis of the infrasound signals from rocket launch[J]. *Journal of Applied Acoustics*, 2021, 40(5): 676–683.
- [13] 孟子轩. 基于非负矩阵分解的次声信号识别分类技术研究[D]. 北京: 中国科学院声学研究所, 2022.
- [14] Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning[J]. *Advances in Neural Information Processing Systems*, arXiv: 1703.05175, 2017.
- [15] Karim F, Majumdar S, Darabi H, et al. LSTM fully convolutional networks for time series classification[J]. *IEEE Access*, 2018, 6: 1662–1669.
- [16] Huisman M, van Rijn J N, Plaat A. A survey of deep meta-learning[J]. *Artificial Intelligence Review*, 2021, 54(6): 4483–4541.
- [17] Bryan K J, Smith K E, Solomon M L, et al. Deep wavelet scattering features for infrasonic threat identification[C]. *Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XIX*. SPIE, 2018, 10629: 125–135.