

## 基于原始肺音的端到端深度学习肺炎辅助诊断

肖欣,高玉清,张建敏

江汉大学人工智能学院,湖北 武汉 430056

**【摘要】**提出一种基于双路径不同卷积核的DCL-Net的端到端肺炎辅助诊断方法。该方法无需进行特征工程,将原始肺音信号直接输入模型,利用卷积核分别为 $1 \times 3$ 和 $1 \times 5$ 的双路径卷积网络,每个路径包含3个残差块,以便模型自动学习肺音信号不同尺度的特征,同时避免模型退化问题。为验证端到端方法的性能,将其与信号分析领域常用的梅尔倒谱图、短时傅里叶变换和小波变换这3种特征提取方法进行比较。结果显示,四分类任务(正常、普通、病重、病危)诊断准确率为61.4%,相比3种特征工程方法分别提高1.6%、5.0%和3.7%;二分类任务(正常、异常)诊断准确率为89.7%,相比3种特征工程方法分别提高11.0%、5.1%和11.2%。实验结果表明该方法可为肺炎病情评估提供更有效的诊断工具。

**【关键词】**原始肺音;肺炎;智能辅助诊断;端到端学习;DCL-Net

**【中图分类号】**R318;R563.1

**【文献标志码】**A

**【文章编号】**1005-202X(2025)02-0274-07

## End-to-end deep learning for auxiliary diagnosis of pneumonia using original lung sounds

XIAO Xin, GAO Yuqing, ZHANG Jianmin

School of Artificial Intelligence, Jiangnan University, Wuhan 430056, China

**Abstract:** An end-to-end auxiliary diagnosis method for pneumonia based on DCL-Net with dual-path of different convolutional kernels is proposed, in which no feature engineering is required, and the original lung sound signal is directly input into the model. The dual-path convolutional network with kernel sizes of  $1 \times 3$  and  $1 \times 5$ , with each path containing 3 residual blocks, allows the model to automatically learn features of lung sounds at different scales while avoiding model degradation. The performance of the end-to-end method is validated through the comparisons with 3 commonly used feature extraction methods in signal analysis, namely Mel-spectrogram, short-time Fourier transform, and wavelet transform. The results show that the proposed method has a diagnostic accuracy of 61.4% for the 4-class classification task (normal, moderate, severe, critical), which is 1.6%, 5.0%, and 3.7% higher than the other 3 feature extraction methods, and the diagnostic accuracy is 89.7% for the binary classification task (normal or abnormal), which is 11.0%, 5.1%, and 11.2% higher than the other 3 feature engineering methods, demonstrating that it can serve as an effective diagnostic tool for pneumonia.

**Keywords:** original lung sound; pneumonia; intelligent auxiliary diagnosis; end-to-end learning; DCL-Net

### 前言

肺炎是一种通常由细菌、病毒或真菌引起的肺部炎症性疾病,其传播途径主要有空气传播、飞沫传播和接触传播<sup>[1]</sup>。肺炎的常见症状有发热、咳嗽、呼吸困难、胸痛、头痛等,严重时甚至可能导致死亡。

全球每年约有300万人因肺炎死亡<sup>[2-3]</sup>,《中国卫生健康统计年鉴》显示2021年我国呼吸系统疾病死亡率位居前列,仅次于恶性肿瘤、心脏病、脑血管病,其中绝大多数是儿童和老人<sup>[4]</sup>,可见肺炎对这两个年龄段群体的健康影响尤为严重。

现代科技已经发展出几种常见的诊断肺炎的方法,如X线胸片、计算机断层扫描、肺功能检查。相比之下,肺部听诊提供一种无创、低成本、便捷式的工作方式<sup>[5]</sup>。将肺部听诊与计算机结合对肺音信号进行自动化分析是诊断和监测肺部疾病的有效方法,医生可以根据肺部听诊数据和患者的临床信息及时采取有效的治疗措施,从而提升肺炎诊断的准确率,有助于病情的早期治疗及控制疾病的发展。

**【收稿日期】**2024-10-26

**【基金项目】**武汉市科技计划项目(2020020601012320);老年医学与现代康养学科(群)项目([鄂教研函[2021]5号])

**【作者简介】**肖欣,硕士研究生,研究方向:智慧医疗和人工智能应用, E-mail: bellaxiao129@163.com

**【通信作者】**张建敏,博士,教授,研究方向:智慧医疗和人工智能应用, E-mail: zhangjm@jhun.edu.cn

传统的基于肺音听诊的肺炎辅助诊断主要是基于信号分析和人工特征工程的方法,通过对肺部声音的采集和分析,识别并提取肺音的特征信息,然后进行人工挑选显著特征的方法进行后续的肺音分类识别。石陆魁等<sup>[6]</sup>对正常音、爆裂音和哮鸣音进行分类,将肺音信号进行小波分解得到的小波系数转换成小波能量特征向量,然后利用线性判别分析将获取的小波能量进行降维,并用支持向量机结合低维特征进行分类,得到更高的识别精度。Altan等<sup>[7]</sup>使用Hilbert-Huang变换提取的频率调制统计特征分析多通道肺音,并将深度学习算法用于分类阶段,区分慢性阻塞性肺病(Chronic Obstructive Pulmonary Disease, COPD)患者与健康受试者。Demir等<sup>[8]</sup>采用短时傅里叶变换将ICBHI2017肺音信号转换为频谱图图像,利用两种基于深度学习的方法对肺音进行四分类。Kim等<sup>[9]</sup>将肺音转换为梅尔频谱图后利用迁移学习和卷积神经网络(Convolutional Neural Networks, CNN)对临床记录的呼吸音进行四分类,结合时间序列、呼吸音和CNN分类的预训练图像特征提取,实现较高的异常声音检测准确率。余辉等<sup>[10]</sup>设计一种改进的19层CNN,对COPD患者的临床肺音数据进行预处理后,利用二阶谱分析技术提取肺音特征后输入网络模型,在健康受试者和COPD患者的二分类中得到很高的准确率。张乙鹏等<sup>[11]</sup>提出一种CNN-LSTM的混合深度学习模型,使用小波变换对ICBHI2017肺音数据集进行特征提取,将获得的空

间维度特征和时间维度特征融合,通过模型完成肺音分类。TaghiBeyglou等<sup>[12]</sup>提出一种结合梅尔尺度频谱图和声音录音原始时间序列的新型频谱时间深度神经网络TRespNet,在SPRSound数据库中完成七类呼吸事件分类任务。

基于端到端学习的肺炎辅助诊断是一种利用深度学习技术直接从原始肺音数据中学习特征并进行诊断的方法,这种方法绕过了传统的特征工程步骤,

直接从原始肺音数据中提取相关信息。目前,基于端到端学习的肺炎辅助诊断研究都集中在医学影像或者原始肺音特征的二维图像上,将其输入网络模型进行分类识别。Ardila等<sup>[13]</sup>提出一种端到端的方法,使用深度CNN学习特征,而不是使用手工设计的特征,并构建一个三维CNN模型,对整个CT体积进行端到端分析。Liu等<sup>[14]</sup>提出一种基于端到端神经网络的肺音分类系统,采用时延神经网络,将从肺音样本中提取的梅尔频率倒谱系数、倒梅尔频率倒谱系数、对数梅尔滤波器能量这3种声学特征作为端到端系统的输入,并在网络的输出层上使用Softmax对正常和异常肺音进行分类。Cao等<sup>[15]</sup>提出一种端到端特征金字塔深度多实例学习模型(E2EFP-MIL),通过端到端学习自动提取广泛的形态特征,并识别具有区分性的组织形态学模式,在肺癌全玻片图像上训练,并得到较高的测试精度。然而,基于原始肺音的肺炎辅助诊断研究仍是空白。

本研究提出一种基于端到端学习的DCL-Net模型,将不做任何特征工程的原始肺音音频经过数据处理后直接输入到DCL-Net进行训练,让网络模型自动学习肺音的特征,并完成肺音的分类识别任务。实验证明DCL-Net模型能满足肺炎患者病情辅助诊断的需求。

## 1 方法

### 1.1 主要流程

本研究提出的端到端肺音辅助诊断方法主要流程如图1所示,主要包括数据处理和分类识别过程。首先将收集到的肺音数据集进行数据处理,以5 s为一段切分肺音数据集为肺音数据段;其次利用音频信号数据增强的方法,将原始肺音数据集进行扩充,使其数量足够大以满足训练需求;然后将经过处理后的肺音数据段输入DCL-Net模型,微调超参数后进行模型训练验证和测试,完成病情评估。

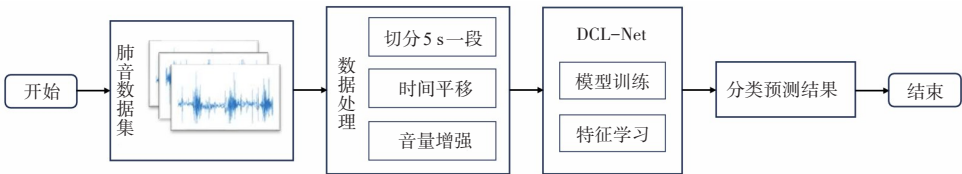


图1 端到端肺炎辅助诊断方法主要流程

Figure 1 Main workflow of the end-to-end method for auxiliary diagnosis of pneumonia

### 1.2 基于DCL-Net肺炎辅助诊断模型

#### 1.2.1 残差学习 肺音信号是长的一维非平稳复杂信

号<sup>[16]</sup>,通常包含多个频率和振幅的成分,这些成分可能源于肺部不同结构的振动(如气道、支气管和肺泡

等)以及与呼吸相关的生理变化。在使用深度网络模型进行肺炎诊断时,需要更深层次的网络模型,使其能提取到较为复杂的特征,模型的性能也会更好。然而,先前的研究表明,随着网络模型层数的增多,准确率反而开始下降,但是其原因并不是过拟合,而是深度网络的退化问题<sup>[17-18]</sup>。

残差学习是深度神经网络中常用的技术,旨在解决深度网络训练过程中的梯度消失和梯度爆炸等问题<sup>[19]</sup>。为了更形式化地描述残差学习,引出残差块的概念,残差块通常由两个连续的卷积层组成,每个卷积层后有批量归一化层和激活函数。残差块结构如图2所示。

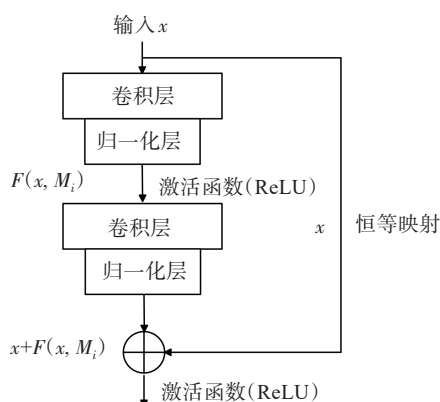


图2 一个残差块结构

Figure 2 A residual block structure

残差学习相比传统特征学习更容易,残差块的输出 $y$ 通过恒等映射得到:

$$y = x + F(x, W_i) \quad (1)$$

其中, $x, y$ 表示输入和输出向量。 $F(x, W_i)$ 为学习到的残差映射,以图2为例,残差块的数学表达式可以表示:

$$F = W_2 \cdot \text{ReLU}(W_1 \cdot x + b_1) + b_2 \quad (2)$$

其中, $W_1$ 和 $W_2$ 是两个卷积核的权重, $b_1$ 和 $b_2$ 分别是第一个和第二个卷积层的偏置项,ReLU是修正线性单元激活函数,因能显著提高训练速度而被广泛应用<sup>[20]</sup>。

残差块的设计使得网络可以轻松地学习恒等映射,即将输入直接传递到输出,无需进行任何变换,即使添加了更多的层,网络也能保持良好的性能。具体来说,如果残差值不为零,增加网络的层数可以提高网络性能;如果残差值为零,则当前层只是恒等映射,网络性能既不会提高也不会降低,这样就可以避免退化问题,从而可以构建更深层次的网络。

**1.2.2 一维CNN** CNN作为一种典型的前馈神经网络,

能从原始输入数据中自动学习和提取特征。虽然CNN最初是为图像分类任务设计的,但CNN对一维数据的适应,如时间序列或序列数据,已经显示出显著的有效性<sup>[21-23]</sup>。一维CNN是基于传统CNN的基本原理构建的,其主要架构包括卷积层、池化层和全连接层。其中,卷积操作是一维CNN的关键步骤之一。卷积层使用一维卷积核实现卷积运算:

$$x_j^i = f\left(\sum_{i=1}^{M^{l-1}} x_i^{l-1} * k_{ij}^l + b_j^i\right) \quad (3)$$

其中, $k_{ij}^l$ 表示从第 $(i-1)$ 层的第 $i$ 个神经元到第 $l$ 层的第 $i$ 个神经元的卷积核,卷积核以步幅 $s$ 在序列上滑动,每次移动一个步长,计算出一个新的特征值 $x_j^i$ 。 $x_j^i$ 和 $x_i^{l-1}$ 分别是第 $l$ 层的输出序列和输入序列。 $j$ 表示卷积核的个数, $M^{l-1}$ 表示 $(l-1)$ 层的神经元个数。 $*$ 表示卷积操作, $f(\cdot)$ 表示激活函数, $b_j^i$ 表示第 $i$ 个神经元在 $(l-1)$ 层的偏置。

池化层通常在卷积层之后进行,对卷积层的输出特征序列进行下采样,降低其维度,减少网络参数,提高模型的计算效率,同时保留最显著的信息。常用的池化操作包括最大池化和平均池化。最大池化是指在滑动窗口中的每个子区域中选择最大值:

$$y_{i,j}^{\text{maxpool}} = \text{MaxPooling}(\max y_{i,j}) \quad (4)$$

其中, $y_{i,j}$ 表示经过卷积层后的输出,MaxPooling( $\cdot$ )表示池化操作, $y_{i,j}^{\text{maxpool}}$ 表示该区域内的最大值。平均池化将滑动窗口中的每个子区域的平均值作为输出值:

$$y_{i,j}^{\text{avepool}} = \text{AveragePooling}\left(\frac{1}{w \times h} \sum_{i=1}^w \sum_{j=1}^h y_{ij}\right) \quad (5)$$

其中, $y_{i,j}$ 表示经过卷积层后的输出,AveragePooling( $\cdot$ )表示池化操作, $y_{i,j}^{\text{avepool}}$ 表示这些区域内的平均值。池化层没有可学习的参数,并且具有计算简单的特点,可以有效减轻网络内部的过拟合风险。

在一维CNN中,全连接层是网络的最后一层,通常用于将卷积层和池化层提取的特征通过一系列加权连接和偏置将其映射为最终的输出。全连接的作用是将上一层的所有神经元连接到下一层的每个神经元,使得网络能根据学习到的特征进行预测和分类,最终产生期望的输出。

**1.2.3 DCL-Net模型** 基于肺音信号的肺炎辅助诊断是一个一维时间序列信号的识别分类问题,为适应肺音信号的输入及更有效地捕捉肺音信号不同感受野的特征。本研究提出一种双路径不同卷积核的DCL-Net,使用该模型捕获原始肺音信号中不同尺度的特征,并使用残差学习有效减轻梯度消失及退化问题。两个路径分别使用 $1*3$ 和 $1*5$ 的卷积核,每个



路径包含3个具有残差连接的卷积块,这两种不同的卷积核尺寸使得网络能学习不同尺度的特征,以更全面地理解和表征肺音信号的信息。DCL-Net模型结构如图3所示。

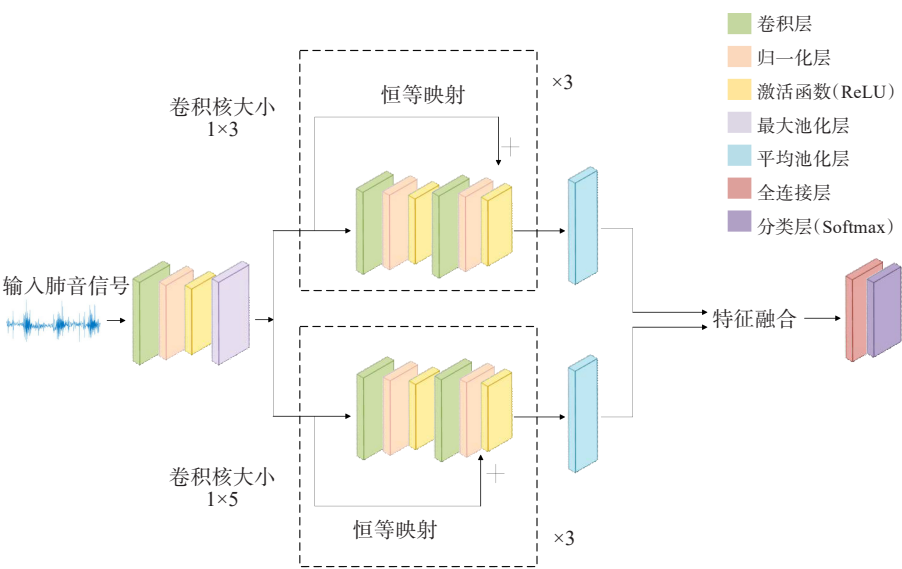


图3 DCL-Net模型结构  
Figure 3 DCL-Net model structure

首先,肺音信号经过一个一维卷积层,这个卷积层采用大小为7的卷积核,并将输出通道设为64,随后引入一个批量归一化层对卷积层的输出进行归一化,再经过ReLU激活层增加网络的非线性;为了进行下采样,接着使用池化核为3的最大池化层。接下来是并行的双路径的残差模块,每个残差模块由两个卷积层组成,且卷积层后面跟着批量归一化层和ReLU激活层。这些卷积层分别采用1\*3和1\*5的卷积核,使得网络能学习不同尺度的特征。在残差模块内部,残差连接有助于信息流在模块内更顺畅地传递,从而避免梯度消失问题。经过双路径的残差模块后,进入全局池化层阶段,由于在卷积层后特征数量增加,导致数据维数的膨胀,因此选择对每个特征使用池化操作,以减少参数量并增强模型的鲁棒性。考虑到信号的连续性,本研究使用平均池化层对特征进行全局平均池化。将双路径输出的张量序列进行连接,然后输入全连接层,用于对连接后的张量序列进行分类,得出类别的预测概率。本研究还在每个池化层后引入dropout层,以防止网络过拟合,提高模型的泛化能力。

2 实验设计

2.1 实验平台

本研究所用硬件设施如下:CPU型号为Intel®Core™ i7-11390H;显卡型号为NVIDIA

GeForce RTX 3090;系统内存为16.0 GB;操作系统为Windows 11;深度学习环境为Pytorch2.0.0、Python3.8、Cuda11.8。

2.2 实验数据

本研究采用2020年3月31日至4月5日在华中科技大学同济医院接受治疗的COVID-19确诊病例,共172例,患者平均年龄65岁,建立两个肺音数据库,分别是:(1)根据患者的病情程度的四分类;(2)根据肺音是否正常的二分类。各病例都通过下一代测序、实时荧光定量PCR方法<sup>[24]</sup>或根据临床诊断标准确认。研究符合1975年赫尔辛基宣言<sup>[25]</sup>的法令,并得到华中科技大学同济医院伦理委员会的批准,机构审查委员会批准编号为TJ-C20201202,并已获得患者或患者近亲的知情同意。

根据诊断指南,利用听诊器在每位患者的10个部位上记录20~30 s自然状态下呼吸情况,对于不能坐起或翻身的气管插管患者,只记录6个部位。采样频率为4 000 Hz,采集数据为wav格式。

2.3 数据处理

首先,将原始肺音以5 s为一段切分,如图4所示,保证每段肺音数据中至少包含一个完整的呼吸周期;然后,利用时间平移和音量增强的数据增强方法对肺音数据进行扩充,确保数据量足够多以满足训练需求。

将原始数据集和生成的新数据集合并在一起,

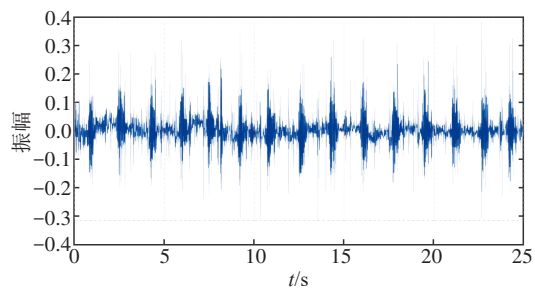


图4 以5 s为一段分割肺音  
Figure 4 Segmenting lung sounds into 5-second intervals

形成增强后的数据集,如表1所示。使用增强后的数据集训练模型可以增强其对时间变化和噪声的鲁棒性。

表1 数据处理前后肺音数据段数量统计  
Table 1 Statistics of lung sound segment counts before and after data processing

肺音数据库	类型	数据处理前	数据处理后
四分类	正常	131	2 343
	普通	61	2 088
	病重	134	2 346
	病危	94	2 192
二分类	正常	373	8 928
	异常	361	8 352

2.4 评价指标

准确率(Accuracy)表示模型预测正确的样本数占总样本数的百分比:

Accuracy = (TP + TN) / (TP + TN + FP + FN) (6)

若以病情分为正常和异常二分类为例,TP代表模型诊断为正常,真实类型为正常;FP代表模型诊断为正常,真实类型为异常;FN代表模型诊断为异常,真实类型为正常;TN代表模型诊断为异常,真实类型为异常。

精确率(Precision)是指模型在预测为正例的样本中,实际为正例的比例,用来衡量模型诊断精度:

Precision = TP / (TP + FP) (7)

灵敏度(Sensitivity)衡量真实值中诊断为正常占样本中实际正常的比例:

Sensitivity = Recall = TP / (TP + FN) (8)

F1分数是精确率和召回率的调和平均数,综合考量模型的准确性和召回能力:

F1 = (2 \* Precision \* Recall) / (Precision + Recall) (9)

3 结果与分析

3.1 端到端方法结果分析

对采集到的肺音信号进行数据处理,并将数据扩充到一定数量以满足模型训练需求。将处理后的肺音数据直接输入DCL-Net模型。在训练过程中,初始学习率设置为0.000 01,批量尺寸设置为32,模型训练500次,Dropout参数为0.2,以ReLU为激活函数,使用SGD优化器,选择交叉熵作为损失函数。

为了从不同感受野学习肺音信号的特征,采用DCL-Net模型进行肺炎患者病情诊断。测试结果显示,四分类任务的诊断准确率为61.4%,二分类任务的诊断准确率为89.7%,模型测试的评价指标数值及混淆矩阵如表2、图5所示。针对肺音四分类任务,模型在区分“正常”类和“病危”类的性能较好,在区分“普通”类和“病重”类的性能较差,由于“普通”类样本较少,模型只能从有限的样本中学习特征。医生在听诊时存在病情由“病重”转“普通”和“病危”转“病重”的现象,从而将“普通”类和“病重”类错误分类。医生对患者的听诊结果标注有听出异常和未听出异常两种情况,这也进一步说明肺音听诊的主观性。针对肺音二分类任务,模型在“正常”类别和“异常”类别上都取得较高的精确率、灵敏度和F1分数,这表明模型对于区分正常情况和异常情况具有较好的能力,整体性能良好。

表2 端到端方法诊断结果  
Table 2 Diagnostic results of the end-to-end method

数据库	类型	准确率	精确率	灵敏度	F1分数
四分类	正常	0.614	0.660	0.878	0.753
	普通		0.615	0.255	0.360
	病重		0.474	0.500	0.485
	病危		0.693	0.803	0.744
二分类	正常	0.897	0.902	0.896	0.899
	异常		0.892	0.898	0.895

3.2 对比实验

为验证本研究提出的基于端到端方法对肺炎患者病情诊断的有效性,因此将端到端学习方法与时频域分析中常用的3种特征提取方法进行比较,分别是梅尔倒谱图<sup>[26]</sup>、短时傅里叶变换<sup>[27]</sup>、小波变换<sup>[27]</sup>。

本研究将从肺音中提取的这3种特征分别输入到DCL-Net网络模型,采取二维卷积。为保证实验结果更

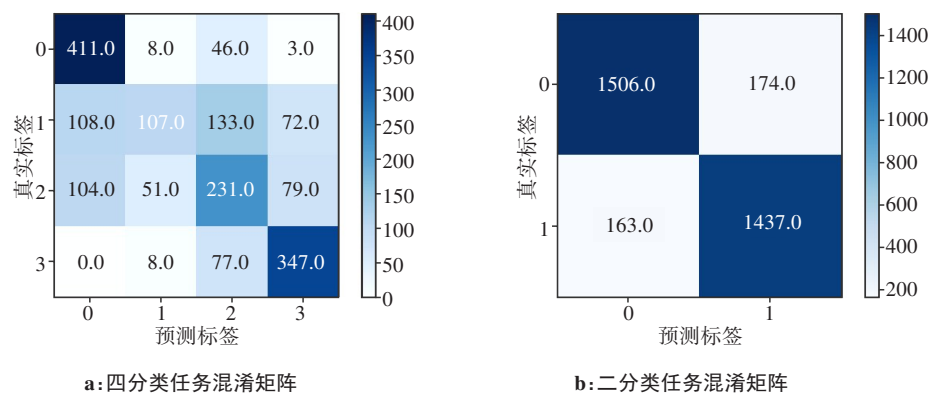


图5 四分类任务和二分类任务混淆矩阵

Figure 5 Confusion matrixes for 4-class and binary classification tasks

具说服力,在实验过程中,端到端学习方法及3种特征提取方法均使用相同的数据集,实验结果如图6所示。结果表明没有经过显式特征提取的端到端方法相较于传统的特征工程方法有较高的识别精度。本研究提出的端到端方法针对肺炎患者病情严重程度的四分类任务的最终诊断准确率为61.4%,相比3种特征工程方法分别提高1.6%、5.0%和3.7%;针对是否患有肺炎的二分类任务最终诊断准确率为89.7%,相比3种特征工程方法分别提高11.0%、5.1%和11.2%(图6)。考虑到这些都是针对真实临床领域中记录中的分析结果,端到端方法达到的诊断结果较为理想,有利于肺炎患者的初步筛查和后续病情诊断工作。由于参与者人数较少,测试结果不是决定性的,可以获取更多肺炎患者的肺音信号样本进行优化调整,增加网络模型的稳健性,更好地完成基于肺音的端到端肺炎智能辅助诊断工作。

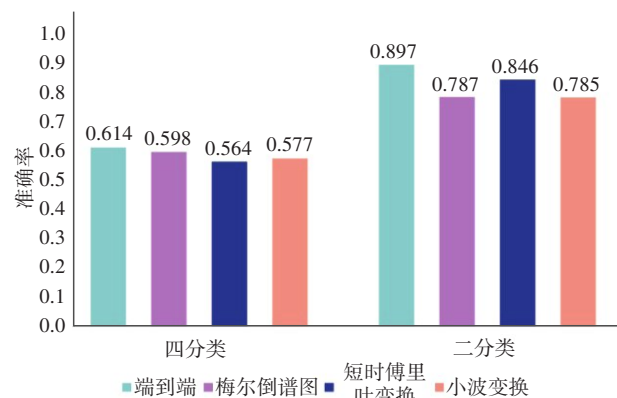


图6 端到端方法与3种特征工程方法准确率对比

Figure 6 Comparison of the end-to-end method with 3 feature engineering methods

4 结 语

本研究提出一个基于端到端方法的肺炎辅助诊

断模型,该方法不进行显式的特征提取,直接让模型从采集到的原始肺音信号中学习特征并进行分类。结果表明,相较于传统的人工提取特征的方法,这种端到端的诊断方法不仅提高了诊断准确性,而且在噪声环境下也具有稳定性。因此,这种端到端的辅助诊断模型有望成为未来肺炎诊断的有利工具,为医生提供更可靠、高效的诊断支持。未来的研究方向可能包括进一步优化模型性能,验证在更广泛人群中的适用性,并探索将该方法与临床实践相结合,以提高肺炎诊断的精准性和及时性。

【参考文献】

[1] Mathioudakis AG, Fally M, Hansel J, et al. Clinical trials of pneumonia management assess heterogeneous outcomes and measurement instruments[J]. J Clin Epidemiol, 2023, 164: 88-95.

[2] GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the global burden of disease study 2015[J]. Lancet, 2016, 388(10053): 1459-1544.

[3] Safiri S, Carson-Chahhoud K, Noori M, et al. Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990-2019: results from the global burden of disease study 2019[J]. BMJ, 2022, 378: e069679.

[4] GBD 2019 LRI Collaborators. Age-sex differences in the global burden of lower respiratory infections and risk factors, 1990-2019: results from the global burden of disease study 2019[J]. Lancet Infect Dis, 2022, 22(11): 1626-1647.

[5] Sgalla G, Larici AR, Sverzellati N, et al. Quantitative analysis of lung sounds for monitoring idiopathic pulmonary fibrosis: a prospective pilot study[J]. Eur Respir J, 2019, 53(3): 1802093.

[6] 石陆魁, 刘文浩, 李站茹. 基于LDA和小波分解的肺音特征提取方法[J]. 计算机工程与应用, 2017, 53(22): 116-120.

Shi LK, Liu WH, Li ZR. Feature extraction method of lung sound based on LDA and wavelet decomposition [J]. Computer Engineering and Applications, 2017, 53(22): 116-120.

[7] Altan G, Kutlu Y, Allahverdi N. Deep learning on computerized analysis of chronic obstructive pulmonary disease [J]. IEEE J Biomed Health Inform, 2020, 24(5): 1344-1350.

[8] Demir F, Sengur A, Bajaj V. Convolutional neural networks based efficient approach for classification of lung diseases[J]. Health Inf Sci Syst, 2020, 8(1): 4.

[9] Kim Y, Hyon YK, Jung SS, et al. Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning[J]. Sci Rep, 2021, 11(1): 17186.

- [10] 余辉, 赵婧, 仇兆禹, 等. 基于深度学习的慢性阻塞性肺病的诊断模型研究[J]. 中国生物医学工程学报, 2022, 41(5): 558-566.  
Yu H, Zhao J, Qiu ZY, et al. Diagnosis model of chronic obstructive pulmonary disease based on deep learning[J]. Chinese Journal of Biomedical Engineering, 2022, 41(5): 558-566.
- [11] 张乙鹏, 孙文慧, 陈扶明. 基于小波变换和CNN-LSTM的肺音分类算法[J]. 中国医学物理学杂志, 2024, 41(3): 356-364.  
Zhang YP, Sun WH, Chen FM. Lung sound classification algorithm based on wavelet transform and CNN-LSTM[J]. Chinese Journal of Medical Physics, 2024, 41(3): 356-364.
- [12] TaghiBeyglou B, Assadi A, Elwali A, et al. TRespNET: a dual-route exploratory CNN model for pediatric adventitious respiratory sound identification[J]. Biomed Signal Process Control, 2024, 93: 106170.
- [13] Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography[J]. Nat Med, 2019, 25(6): 954-961.
- [14] Liu LL, Li L, Li S, et al. An end-to-end system based on TDNN for lung sound classification [C]//2020 IEEE 14th International Conference on Anti-counterfeiting, Security, and Identification (ASID). Piscataway, NJ, USA: IEEE, 2020: 20-24.
- [15] Cao L, Wang J, Zhang YY, et al. E2EFP-MIL: end-to-end and high-generalizability weakly supervised deep convolutional network for lung cancer classification from whole slide image[J]. Med Image Anal, 2023, 88: 102837.
- [16] Kwon AM, Kang K. A temporal dependency feature in lower dimension for lung sound signal classification[J]. Sci Rep, 2022, 12(1): 7889.
- [17] Srivastava RK, Greff K, Schmidhuber J. Highway networks[EB/OL]. (2015-11-03)[2024-05-01]. <https://arxiv.org/abs/1505.00387>.
- [18] He KM, Sun J. Convolutional neural networks at constrained time cost[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2015: 5353-5360.
- [19] Sun XJ, Liu W, Wang K, et al. An integrated skip convolutional network with residual learning and feature extraction for point and interval prediction of solar radiation[J]. Appl Soft Comput, 2024, 159: 111621.
- [20] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks [J]. Commun ACM, 2017, 60(6): 84-90.
- [21] Abdoli S, Cardinal P, Lameiras Koerich A. End-to-end environmental sound classification using a 1D convolutional neural network[J]. Expert Syst Appl, 2019, 136: 252-263.
- [22] Kiranyaz S, Ince T, Hamila R, et al. Convolutional neural networks for patient-specific ECG classification [C]//2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Piscataway, NJ, USA: IEEE, 2015: 2608-2611.
- [23] Chen XZ, Jia JF, Yang J, et al. A vibration-based 1DCNN-BiLSTM model for structural state recognition of RC beams[J]. Mech Syst Signal Process, 2023, 203: 110715.
- [24] WHO. Clinical management of severe acute respiratory infection when novel coronavirus (nCoV) infection is suspected[EB/OL]. (2020-01-12) [2024-05-05]. <https://www.who.int/publications/i/item/10665-332299>.
- [25] World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects[J]. JAMA, 2013, 310(20): 2191-2194.
- [26] Dianat B, La Torraca P, Manfredi A, et al. Classification of pulmonary sounds through deep learning for the diagnosis of interstitial lung diseases secondary to connective tissue diseases [J]. Comput Biol Med, 2023, 160: 106928.
- [27] Motamedi-Fakhr S, Moshrefi-Torbati M, Hill M, et al. Signal processing techniques applied to human sleep EEG signals-a review[J]. Biomed Signal Process Control, 2014, 10: 21-33.

(编辑:谭斯允)