

基于深度神经网络的新型肠鸣音检测分析方法

卫子然¹, 张鑫¹, 李东¹, 操家庆², 桂坤², 蔡清萍^{1*}

¹上海长征医院肠胃外科, 上海

²宁波江丰生物技术有限公司, 浙江 宁波

Email: lidong3012012022@sina.com, caiqingpingwcwk@163.com

收稿日期: 2020年8月5日; 录用日期: 2020年8月26日; 发布日期: 2020年9月2日

摘要

肠鸣音是人体重要的生理信号, 对肠鸣音的检测和分析具有重要的临床价值。传统的听诊器检测法主观性强, 且无法做到连续、动态的监测, 导致数据的时效性及精确性差。另一种方法是借助声音传感器采集和数字化肠鸣音, 然后利用计算机对肠鸣音进行处理和特征分析, 以实现肠鸣音客观、定量检测。但由于肠鸣音具有环境噪声干扰大、周期性差、随机性强等特点, 人工提取普适稳定的特征极为困难, 导致这种识别方法检测准确率较低。项目组提出一种有效的肠鸣音检测方法, 即先提取肠鸣音的MFCC特征, 然后采用深度神经网络提取更加稳定抽象的特征, 最后采用softmax识别出肠鸣音出现的位置和肠鸣音具体的类别。实验表明, 这种方法能够准确检测出肠鸣音出现的时刻, 同时也有较高准确率识别肠鸣音类别, 值得在临床中推广应用。

关键词

肠鸣音, 深度神经网络, 特征提取

A New Method for Detection and Analysis of Bowel Sounds Based on Deep Neural Networks

Ziran Wei¹, Xin Zhang¹, Dong Li¹, Jiaqing Cao², Kun Gui², Qingping Cai^{1*}

¹Department of Gastrointestinal Surgery, Shanghai Changzheng Hospital, Shanghai

²Konfoong Biotech International Co., Ltd., Ningbo Zhejiang

Email: lidong3012012022@sina.com, caiqingpingwcwk@163.com

Received: Aug. 5th, 2020; accepted: Aug. 26th, 2020; published: Sep. 2nd, 2020

*通讯作者。

文章引用: 卫子然, 张鑫, 李东, 操家庆, 桂坤, 蔡清萍. 基于深度神经网络的新型肠鸣音检测分析方法[J]. 仪器与设备, 2020, 8(3): 93-99. DOI: 10.12677/iae.2020.83012

Abstract

Bowel sounds is an important physiological signal of the human body, and it has important clinical value for the detection and analysis of bowel sounds. The traditional stethoscope detection method is highly subjective and cannot be continuously and dynamically monitored, resulting in poor timeliness and accuracy of data. Another method is to use a sound collector to collect and digitize bowel sounds, and to use a computer to reduce noise, amplify and manually extract features, and identify them through a classifier. However, due to the characteristics of large environmental noise interference, poor periodicity, and strong randomness, artificial extraction of universal features is extremely difficult, resulting in lower classifier recognition accuracy. In this paper, an effective method is proposed for the detection of bowel sounds, namely, the MFCC feature extraction of bowel sounds, and then the deep neural networks is used to detect the category of bowel sounds. The confirmative study of large samples based on the bowel sounds database shows that this method has high recognition accuracy and can avoid the high complexity of human extraction features, so it is worth popularizing in clinical application.

Keywords

Bowel Sounds, Deep Neural Networks, Feature Extraction

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

肠鸣音是肠蠕动时肠管内气体和液体随之流动而产生的一种断断续续咕噜声或气过水声，它反映了人体肠道的运动状态。作为一个重要的临床体征，肠鸣音的监测在临床上，尤其是胃肠外科、急诊外科具有重要的意义[1]。传统方法采用听诊器诊断，但这种方法由于不同临床医生的经验水平等不同主观因素，导致其准确率低，临床价值大大降低[2]。为了客观，量化的诊断肠鸣音，一些研究者提出先通过传感器采集并数字化肠鸣音，然后通过借助数字信号处理的技术来提取分析不同肠鸣音的不同特征来诊断肠鸣音方法[3]。但由于肠鸣音的特异性和敏感性差，外部噪声强，如何提取稳定、普适的特征是数字化肠鸣音诊断的难点。

2. 相关工作

传统先采用自适应滤波器[4]和小波变换滤波器[5]对肠鸣音进行降噪，然后采用小波分析，归一化香浓能量分布，功率谱密度估计等多种方法对肠鸣音进行特征提取[6]。这些特征能够很好区分出机械性肠梗阻引起的亢进类型肠鸣音。然而现有的滤波器无法完全过滤肠鸣音中所有的噪声，这些噪声对减弱和消失等肠鸣音等的特征提取造成了较大的干扰，导致这类肠鸣音提取到的特征无法用简单线性分类器识别。

近年，深度神经网络已被证明语音识别邻域有效性，主要原因在于它是非线性的分类器，即利用多层非线性的感知器，不依赖的专门人工特征提取就能到达较高的识别率。此外经过特定的训练的神经网络具有很高的抗噪性。鉴于此，项目组提出一种深度神经网络来检测肠鸣音方法，即先提取肠鸣音的 MFCC 特征以降低计算量，然后采用一种基于 wavenet [7]深度神经网络提取更加稳定抽象的特征，并

添加注意力层，对神经网络每个输出节点特征进行加权。最后采用 softmax 识别每个输出节点对应时刻是否出现肠鸣音，同时融合每个输出节点输出的特征，采用 softmax 识别该肠鸣音具体类别。

3. 方法

本节具体介绍我们提出肠鸣音识别方法 3.1 和训练方式 3.2，随后将在第 4 节通过实验介绍该方法每个模块对识别效果产生的影响。

肠鸣音识别方法

MFCC 特征提取：一般肠鸣音传感器的采样率(fs)为 8000 HZ，采集时间为 2~4 分钟，那么采集到肠鸣音的数据长度高达 96 万。如果直接用于识别，计算量大难以达到实时要求。幸运的是 Davis 等人基于听觉模型提出将语音转化为 MFCC 特征的方法[8]，不仅能提高语音识别准确率，而且转化后的 MFCC 特征的数据长度与原始语音数据相比大幅缩短，降低后续语音处理的计算量。为了验证 MFCC 特征也能够用于肠鸣音的识别，本文随机提取不同类型(亢进，活跃，正常，减弱，消失)肠鸣音的 MFCC 特征，并可视化其特征图(如图 1)，结合病理医生标注的肠鸣音位置信息，可看出大部分肠鸣音出现时刻巧好在图 1 中高亮的橘黄色区域。本文根据经验选取长度为移动步长(step)为 32 ms 的滑窗来提取肠鸣音 MFCC 特征(其他移动步长的滑窗应该也能取得相似识别效果)。在该参数下，每秒的音频数据长度(length)从原来的 $\text{length} = \text{fs}$ 降低至 $1000/\text{step} = 32$ 左右长度的数据。

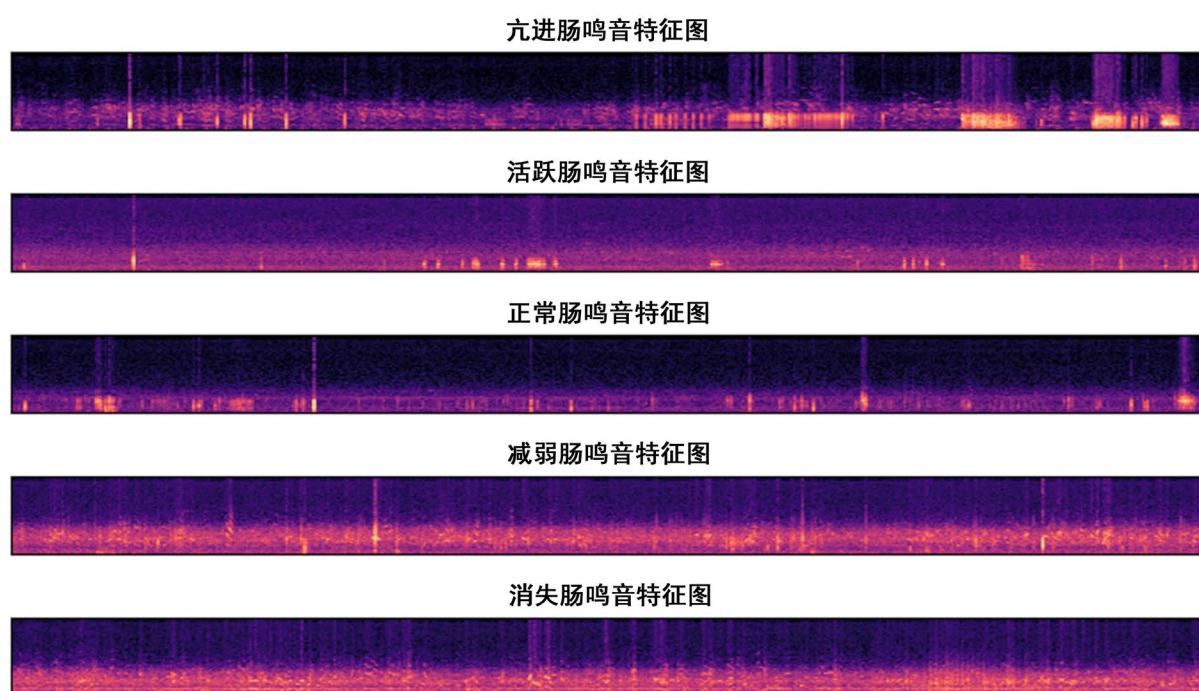


Figure 1. MFCC feature map of different types of bowel sounds

图 1. 不同类型肠鸣音的 MFCC 特征图

基于 wavenet 的肠鸣音特征提取：尽管 MFCC 特征能过直观的区别响亮肠鸣音和非肠鸣音特征，但和其他传统的特征提取法一样无法直观区分噪声和轻微的肠鸣音，我们用深度神经网络进一步提取更加稳定抽象的特征。为了确保神经网络能够提取到有效的肠鸣音特征，则需要神经网络最后一层的每个输出节点能够获取足够的上下文信息，即足够大的感受野(Receptive Field)。我们采用 wavenet 作为肠鸣

音的特征提取层，因为 wavenet 每个基础层采用空洞卷积，以指数增长的方式扩大感受视野。如图 2 所示，假设 wavenet 某一基础层含有 n 个隐藏层，每个隐藏层的卷积大小为 k ，那么该基础层的每个输出节点的感受视野为 $(2^{n+1} - 2)(k - 1) + k$ 。

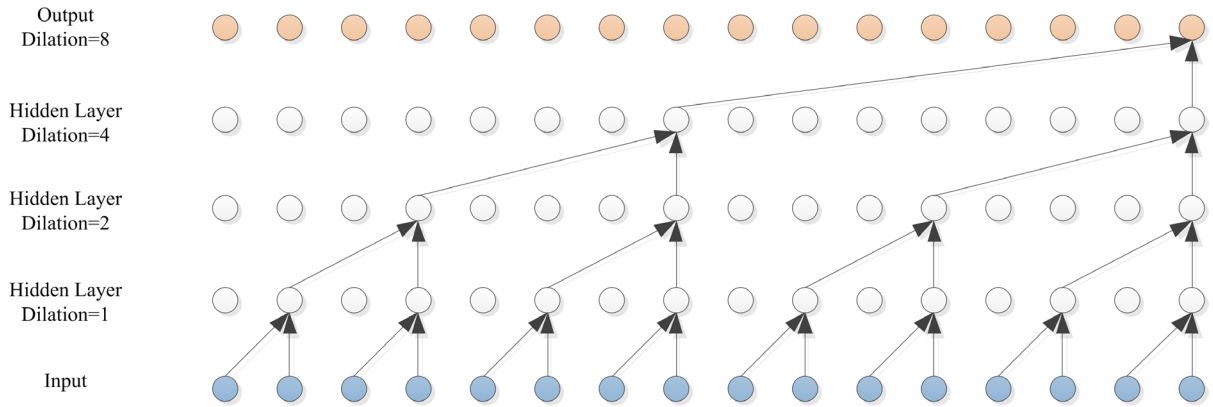


Figure 2. Wavenet basic network structure
图 2. Wavenet 基础网络结构

为了确定肠鸣音的最大感受视野，即确定 wavenet 中 n 和 k 的取值范围，本文从鸣音数据库中随机抽取了不同类型(正常, 活跃, 亢奋, 减弱, 消失)的肠鸣音各 100 条，然后统计了这些肠鸣音中每一小段的肠鸣音持续时间。如图 3 所示，大部分的肠鸣音持续时间在 4 s 以下，因此肠鸣音的网络特征输出感受视野(ReceptiveField)应在 $4 * \text{length} = 128$ 左右。根据经验，空洞卷积核的大小 k 取值范围为 3, 5, 7，则对应的隐藏层数 n 取值范围为 6, 5, 4。我们发现卷积核大小 $k = 7$ 时，层数 $n = 4$ ，识别效果最好。

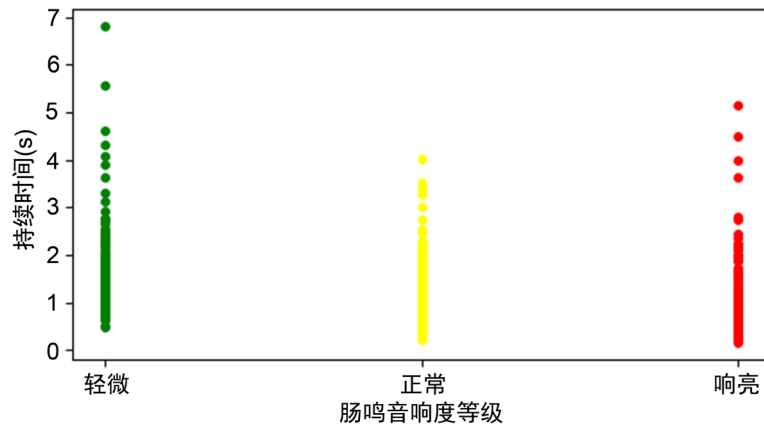


Figure 3. Distribution of bowel sounds with different loudness
图 3. 不同响度肠鸣音持续时间分布图

注意力层：我们在采集肠鸣音发现，消失，减弱等大部分情况下肠鸣音的持续时间不到 10 s，占采集音频的总时长不足 10%。于是我添加注意层，以降低降低背景声音的比重。即在 wavenet 特征输出层添加一个全连接层，预测每个输出节点输出特征的贡献度，并用 softmax 将这些贡献度进行归一化，然后将归一化的贡献度乘以原输出特征，得到注意力加权后的特征。

识别结果输出：我们对加权后的特征，采用 softmax 预测每个输出节点是否出现肠鸣音(肠鸣音定位)，然后将所有输出节点累加后用 softmax 输出肠鸣音音频的类别。

数据采集：神经网络能不能具有较长抗噪性，取决于训练数据集中是否含有多样的噪声数据。为了确保神经网络在训练时能够学习到噪声和肠鸣音的区别，我们额外采集了 200 个不同非肠鸣音用于训练，这些数据包括心跳声，远处的说话声和汽笛声等。我们发现，通过添加额外的数据训练有效降低肠鸣音的错误率。

4. 评估

4.1. 评价指标

我们采用了准确率(accuracy)评估神经网络的肠鸣音音频分类性能，并采用 F1-Score 评估肠鸣音定位性能。如图 4 所示，红色曲线为人工标注的阳性区间。绿色曲线为人工标注的阴性区间，红色背景为模型输出的阳性片段，绿色背景为模型输出的阴性片段。其中模型将人工标注的阴性区间预测为阳性总长定义为 FP ，模型将人工标注的阳性区间预测为阴性总长定义为 FN ，模型将人工标注的阳性区间预测为阳性总长定义为 TP ，那么实验的评估指标为：

$$F1 = \frac{2 * TP}{(TP + FP) + (TP + FN)}$$

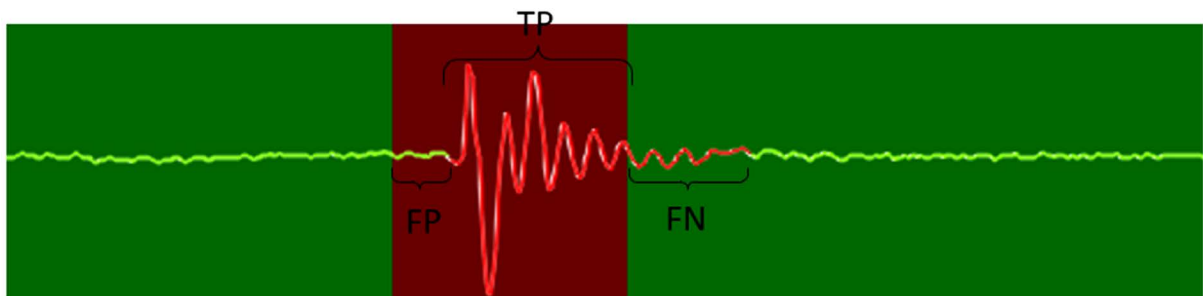


Figure 4. Pointing and prediction of bowel sounds

图 4. 肠鸣音标注和预测图

4.2. 实验

4.2.1. 数据

1) 采集方法

项目组通过骨传导采集器，能够将测试患者肠鸣音存储为音频播放文件。应用该采集器，对上海长征医院 3000 个不同的病人，进行肠鸣音数据采集，每段音频时长在 2~4 分钟，采样频率 f_s 为 8000 HZ。

2) 数据分布

随机选取其中 2000 个数据用于训练，另外 1000 个用于测试。样本集分布情况如表 1 所示。

Table 1. Sample set distribution

表 1. 样本集分布情况

分类	样本种类	数量	平均时长
训练集	正常病例	200	2.35 min
	肠梗阻术前	800	3.15 min
	肠梗阻术后	800	3.23 min
	噪音	200	4 min

Continued

测试集	正常病例	100	2.43 min
	肠梗阻术前	400	3.08 min
	肠梗阻术后	400	3.32 min
	噪音	100	4 min

4.2.2. 模型

wavenet 由多层基础层组成。我们尝试多种不同层数验证其效果。其结果如图 5 所示，当层数增加到 3 以后，测试集分类准确率和肠鸣音定位 F1 几乎不再增加。

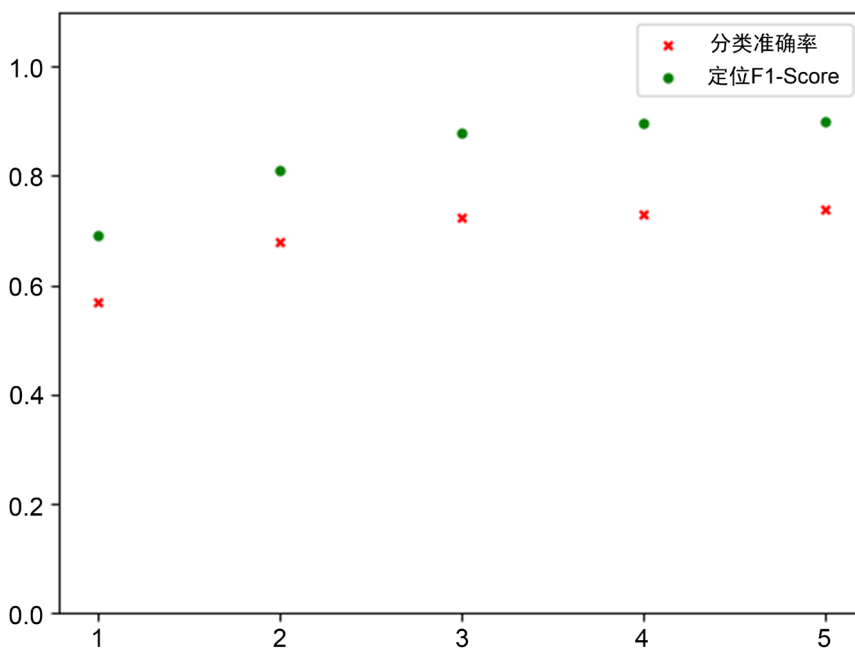


Figure 5. Accuracy under different layers

图 5. 不同层数下识别效果

我们用三层基础层的 wavenet 提取肠鸣音特征，尝试添加注意力层和自适应滤波来检测肠鸣音。识别效果如表 2 所示，注意力层对模型的分类的提升效果显著，其原因在与注意力层让神经网络更加专注于音频中肠鸣音部分的特征提取，此外从表中也看出自适应滤波对肠鸣音提升效果不大，这说明我们通过增加噪声训练的神经网络具有与自适应滤波相同的抗噪性能。

Table 2. Impact of different modules on the recognition of bowel sounds

表 2. 不同模块增加对肠鸣音识别效果的影响

wavenet	网络结果组合			肠鸣音识别效果	
	注意力层	自适应滤波	分类准确率	定位 F1-Score	
✓	°	°	0.56	0.91	
✓	✓	°	0.72	0.88	
✓	✓	✓	0.72	0.89	

5. 总结与展望

实验证明, 对肠鸣音 MFCC 特征提取后, 基于 wavenet 神经网络对肠鸣音识别, 能够达到较好的识别效果, 其中肠鸣音音频分类准确率达 72%, 音频肠鸣音定位的 F1-score 高达 88%, 能够极大的提高临床诊断效率和准确率。

基金项目

军队重点项目(BWS13C011); 上海申康“市级医院新兴前沿技术联合攻关项目”(SHDC12017126); 上海市卫健委面上项目(201940493)。

参考文献

- [1] Cannon, W. (1905) Auscultation of the Rhythmic Sounds Produced by the Stomach and Intestines. *American Journal of Physiology*, **14**, 339-353. <https://doi.org/10.1152/ajplegacy.1905.14.4.339>
- [2] 侯峰, 刘思达, 段降龙. 肠鸣音检测技术的应用进展[J]. 世界最新医学信息文摘, 2018, 18(A5): 128-129.
- [3] 胡勇, 刘文辉. 肠鸣音的采集与分析[J]. 中国生物医学工程学报, 1995, 14(3): 266-275.
- [4] Mansy, H.A. and Sandler, R.H. (1997) Bowel-Sound Signal Enhancement Using Adaptive Filtering. *IEEE Engineering in Medicine and Biology*, 105-116. <https://doi.org/10.1109/51.637124>
- [5] Hadjileontiadis, L.J. (2000) Enhancement of Bowel Soounds by Wavelet-Based Filtering. *IEEE Transactions on Bio-medical Engineering*, **47**, 155-159. <https://doi.org/10.1109/10.846681>
- [6] 张和华. 人体肠鸣音检测系统研制及肠鸣音信号处理方法研究[D]: [硕士学位论文]. 重庆: 第三军医大学, 2009.
- [7] Oord, A.V.D., Dieleman, S., Zen, H., *et al.* (2016) WaveNet: A Generative Model for Raw Audio.
- [8] Davis, S. and Mermelstein, P. (1980) Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **28**, 357-366. <https://doi.org/10.1109/TASSP.1980.1163420>