

基于 CNN-LSTM 神经网络的声纹识别系统设计

牟俊杰, 姚刚, 孙涛

(海军航空大学 岸防兵学院, 山东 烟台 264001)

摘要: 为实现对心血管疾病的预警, 及早发现以心率、心肺音恶性变化为代表的危险前兆, 设计基于 CNN-LSTM 神经网络的声纹识别系统。利用物联网技术融合心率传感芯片、单片机、电子听诊器等设备, 对心率进行实时监测、辅助预警; 根据梅尔道普频率系数对心肺音信号进行特征提取, 构建基于 CNN-LSTM 算法的心肺音智能识别模型, 对部分心肺音进行智能检测诊断, 实验结果显示损失值为 0.082, 准确率达 0.908。开拓了人工智能技术在心血管疾病预警方面的应用空间, 前瞻性强、结构框架完整, 可有效避免医疗资源浪费, 前置对心血管疾病的应对措施, 市场应用前景广阔, 对于推动智慧医疗有重大作用。

关键词: 深度卷积神经网络; 长短期记忆网络; 特征提取; 梅尔道普频率系数; 心血管疾病; 声纹识别

中图分类号: TP399

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.200960

中文引用格式: 牟俊杰, 姚刚, 孙涛. 基于 CNN-LSTM 神经网络的声纹识别系统设计 [J]. 电子技术应用, 2021, 47(3): 75-78.

英文引用格式: Mu Junjie, Yao Gang, Sun Tao. Design of voiceprint recognition system based on CNN-LSTM neural network [J]. Application of Electronic Technique, 2021, 47(3): 75-78.

Design of voiceprint recognition system based on CNN-LSTM neural network

Mu Junjie, Yao Gang, Sun Tao

(Coastal Defense College, Naval Aviation University, Yantai 264001, China)

Abstract: For warning of cardiovascular disease, in order to early detect the change of heart and lung voice representing the signs of danger, the voiceprint recognition system based on CNN-LSTM is designed. Using the Internet of Things technology coalescing the heart rate sensor chip, single-chip computer, electronic stethoscope, such as equipments, it can monitor the heart rate in real-time, early warn. And the cardiopulmonary sound recognition model based on the CNN-LSTM algorithm is trained, results show that the loss value is 0.082, accuracy rate of 0.908. The system is forward-looking and has a complete structural framework, which can effectively avoid the waste of medical resources, preposite the countermeasures for cardiovascular diseases. It has a broad application prospect in the market, and plays a significant role in promoting smart medical treatment.

Key words: CNN; LSTM; features extraction; MFCC; cardiovascular disease; voiceprint recognition

0 引言

随着人工智能的突破性进展和“互联网+”技术的普及, 智慧医疗成为医疗技术发展的新引擎, 诞生了一系列智能医疗服务产品, 如智能药盒、智能手环等^[1]。但人工智能过高的成本导致寻找合适的切入方式显得尤为关键^[2]。

在人口老龄化日益严重的当下, 心血管疾病不断威胁老年人健康, 引发社会广泛关注。由于医疗知识欠缺、行动不便等原因, 部分老年人就医不及时, 错过了抢救的黄金时间, 留下永远的遗憾。开发心血管疾病方面的智能预警系统, 满足庞大的老年人群体需求迫在眉睫^[3]。在医疗实践中, 对心血管疾病的诊断常常以心率、心肺音数据为重要支撑, 国内外以 DSP^[4]、长短时记忆(Long Short Time Memory, LSTM)^[5]、卷积神经网络^[6](Convolutional Neural Network, CNN) 等方法算法为手段对心血管疾病

的信号诊断进行了相当多的分析, 但基本均停留在理论层面, 距离软硬件结合的实际应用尚有差距。各种医疗设备的聚焦点主要是信号的准确采集、分离^[7-8], 基于医疗伦理等原因, 对智能诊断设备的研制尚处于知识储备期, 有巨大的空白亟需填补。本文设计了基于 CNN-LSTM 的心血管疾病预警系统, 利用物联网技术采集心率和心肺音等健康指标数据, 对老人的健康状况进行实时监测、预警, 采用基于 CNN-LSTM 模型的智能算法对心肺音信号进行智能分析预警。系统着重考虑了适用性、稳定性和成本, 具有较高的实用价值和完整的结构框架, 是利用智慧医疗从应用层面解决心血管疾病问题的一次重要探索。

1 总体方案

该心血管疾病预警系统由心率检测分系统和体音

检测分系统两大功能模块构成。系统工作时,先通过穿戴式心率检测设备采集使用者的心率数据,通过蓝牙将数据传输给手机微信小程序。若判定心率数据表现异常,手机将提醒使用者使用电子听诊器,将心肺音数据传到PC端进行智能分析诊断,继而将分析诊断结果反馈给手机微信小程序。考虑到使用者的身体状况相对虚弱,若系统判断使用者需要就医,则手机将自动拨打急救电话、发送病人GPS信息给急救中心,并将使用者相关数据展示给医生,方便医生提前了解使用者病情,筹划后续诊疗工作。系统工作流程如图1所示。

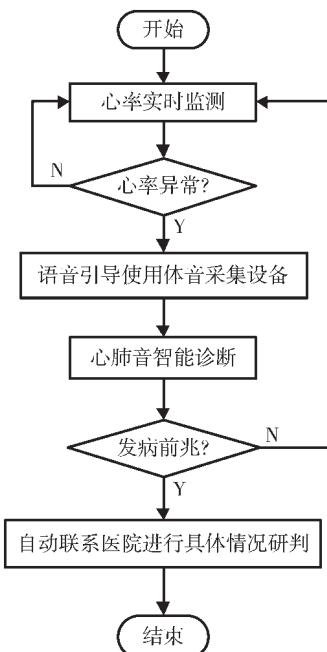


图1 工作流程图

2 硬件设计

检测系统由穿戴式心率检测设备、手机、PC端和电子听诊器等组成,如图2所示。

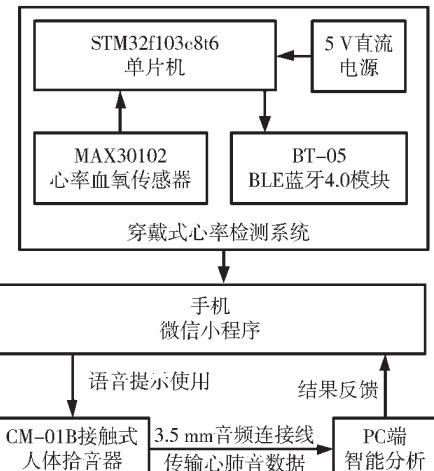


图2 系统组成框图

2.1 穿戴式心率检测设备

穿戴式心率检测设备包括心率传感芯片、单片机、蓝牙模块和电源模块等,将设备佩戴在手腕处,即可实时检测心率^[9]。其中MAX30102心率传感芯片集成了光电检测器、光器件以及带环境光抑制的低噪声电子电路,具有IIC通信接口,可以将采集到的数值传输给STM32单片机进行心率计算,还可通过软件关断模块,使待机电流接近为零,实现长时待机功能。

2.2 电子听诊器

CM-01B电子听诊器通过性能优越、传输数据稳定的PVDF压电薄膜同前置放大电路相连,将外部杂音干扰降至最低。当听诊器接触到振动,则反馈回一个高敏感型信号,以此通过电子脉冲的方式提供微小声音采样。将听诊器置于使用者心胸下两寸处,静置10 s,待PC端软件启动后,分别对心脏及肺部肺泡气管等处诊听,记录不同位置心肺音。电子听诊器通过3.5 mm音频数据线与电脑连接,由PC端利用深度学习技术完成对心肺音数据的分析判断。

3 软件设计

预警系统的软件部分包括心率检测模块和体音检测模块,二者前后衔接,共同实现心率传感器的驱动、IIC通信、蓝牙数据传输、心率异常预警、医疗引导、体音采集和智能分析诊断等功能。

3.1 心率检测模块

为避免繁杂的APP客户端,系统使用微信小程序实现多种功能的整合融合。心率检测模块可以全时段检测使用者的心率,并长时保留心率数据。当模块检测到心率超过设定的门限值,则判定使用者身体出现心血管疾病前兆,手机将立刻发出语音、文字预警警报,同时,为了避免因虚警就医而浪费医疗资源,模块会提醒使用者进一步检测心肺音。程序所实现的心率检测与预警如图3所示。



图3 心率检测界面(正常/异常)

IIC 总线是由时钟 SCL 和数据线 SDA 构成的串行总线, 在传送数据过程中共发送三种不同类型信号, 即开始信号、结束信号和应答信号, 总线中每个电路有唯一的地址, 只有选择正确的地址, 不同电路之间才能实现互联互通, 避免杂乱混乱, 其接口直接连接在组件之上, 因此大大减少了传统电路板的空间和芯片管脚的数量。

3.2 体音检测模块

体音检测是通过深度卷积神经网络-长短期记忆网络(CNN-LSTM)算法实现对心肺音类型的深度学习^[10-11], 利用训练出的识别模型对收集的数据进行智能诊断^[12-13], 以达到对爆破音、干啰音等心血管疾病高危前兆先期识别、居家听诊的目的。

在训练心肺音识别模型的过程中, 采用迁移学习技术, 引入一个公开的心肺音数据库。该心肺音数据库由葡萄牙和希腊的两个研究小组创建, 包括来自 126 名患者的 920 个不同长度的注释录音。录音共计 5.5 小时, 包含 6 898 个呼吸周期, 1 864 个噼啪声, 886 个喘息声, 506 个干啰音, 还有干净的呼吸音以及真实生活条件的嘈杂录音, 选取其中 800 个录音为训练样本, 120 个录音为测试样本。

梅尔道普频率系数(Mel-Frequency Cepstral Coefficients, MFCC)被广泛应用在声音识别等信号处理中, 能够很好地反映音频、音高变化^[14-15]。在构建识别模型时, 将按照 MFCC 对心肺音里包含的语音物理信息(频谱包络和细节)进行编码运算得到的特征向量记为 MFCC 特征。在处理数据集时, 从对应类型的每个文件中截取 2 s 的呼吸时长进行标注, 提取 MFCC 特征。特征提取由卷积层完成, 每个卷积核对应提取某一部分特征。在对声音信号里的 13 维 MFCC 特征进行提取的同时, 还抽取了 13 维 MFCC 特征的一阶方差和二阶方差, 一共形成了 39 维向量, 经过融合层作为 LSTM 模型的输入矩阵。

$$S = f(WZ + b)$$

式中: S 表示经过卷积操作抽取后的特征矩阵; f 为非线性激活函数 LReLU; 操作权重矩阵 W 和偏置向量 b 为该网络学习的参数。

卷积后将 MFCC 特征传递给池化层, 进行进一步聚合, 简化特征表达, 降低特征图的维度和参数的数据量。池化层大小为 4×4, 池化方法选择最大池化, 即选择 4×4 范围内最大的数值作为新的池化层特征。

LSTM 网络具有良好的捕捉时序特征信号的能力, 在分析所采集的体音数据时, 可有效结合以往长期采集的数据, 使分析结果更具科学性、准确性。

将提取 MFCC 特征形成的 39 维特征向量, 处理成(1 561, 39)的矩阵作为输入, 由于 LSTM 为循环神经网络, 为防止网络训练过程出现过拟合, 在循环的神经单元内部及循环之间分别加入 Dropout 和 Recurrent_dropout, 分别对同一 LSTM 单元中神经元之间的连接和不同循

环 LSTM 单元之间的连接进行一定比例的随机断开, 系统中是断开 20% 的内部及外部循环单元, 继而通过展平与全连接实现神经网络的整体结构。

为了使检测系统便于使用, 运用 C# 语言开发智能医疗电脑客户端, 并设计 C# 接口调用人工智能算法的 Python 文件, 集成了心肺音智能诊断、保存并发送检测数据等功能, 以提供智慧医疗、远程医疗服务, 体音检测客户端界面如图 4 所示。



图 4 体音检测界面

4 实验验证

4.1 穿戴式心率检测设备的通信测试

系统用 STM32 驱动心率传感器和蓝牙模块实现心率数据的采集与通信, 通过 USB 转 TTL 连接电脑, 用电脑串口助手测试心率传感器的数据采集情况, 测试结果如图 5 所示, 通信正常, 说明心率检测分系统能够正常采集数据。

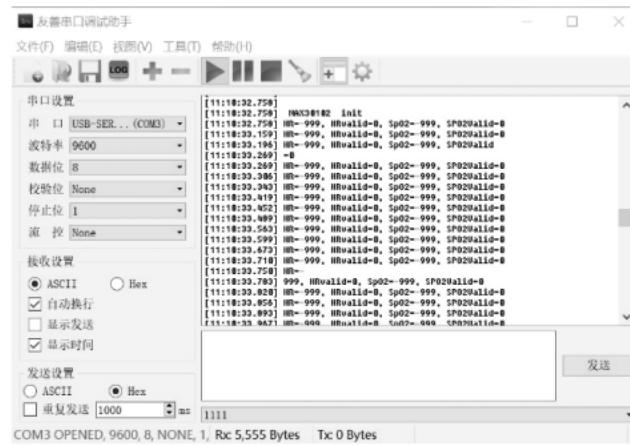


图 5 心率数据采集测试情况

与手机进行通信测试, 通过手机蓝牙串口 app 测试通信情况, 如图 6 所示, 通信正常, 说明能够将采集到的人体心率数据通过蓝牙发送到手机。

4.2 心肺音识别模型测试

在 800 个心肺音样本训练集上对 CNN-LSTM 网络



图 6 蓝牙串口通信情况

训练,训练采用 Adam 优化算法,学习率 lr 取 0.0001,训练迭代次数 steps=18 000,训练次数在 17 500 左右 loss 达到收敛,约为 0.082,训练曲线如图 7 所示。

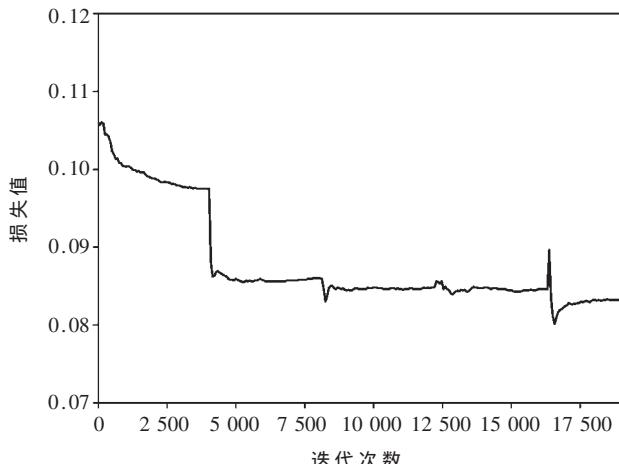


图 7 训练过程 loss 的变化趋势

在 120 个录音测试集上对最终训练的网络模型进行测试, 正确识别的录音数量为 109, 识别的平均准确率达到了 0.908, 满足实际应用需求。

5 结论

该声纹识别系统，利用物联网技术实现了心率、心肺音检测设备与通信设备、数据处理设备的融合，可实时监测心率以先期预警，并支持基于心肺音分析的虚警验证。利用深度学习技术结合医学大数据，对心肺音数据进行 MFCC 特征提取，基于 CNN-LSTM 算法训练出心肺音识别模型，可准确判断部分心血管疾病前兆，实用性强、准确性高，在居家独立进行心血管疾病先期诊断

处理方面应用前景广阔。下一步工作是对穿戴设备进行小型化设计,加大医疗数据库的体量,进一步提高训练模型的识别精度,使系统更加便捷高效。

参考文献

- [1] 袁紫藤,陶金婷,谈莹,等.国内外医疗人工智能应用现状及相关政策[J].医学信息学杂志,2019,40(5):2-9.
 - [2] 史加荣,马媛媛.深度学习的研究进展与发展[J].计算机工程与应用,2018,54(10):1-10.
 - [3] 向运华,王晓慧.人工智能时代老年健康管理研究[J].新疆师范大学学报(哲学社会科学版),2019,40(4):98-107.
 - [4] 陈茜茹,钟清华,张涵,等.基于 DSP 的心肺音分离系统研究[J].现代电子技术,2019,42(6):92-95.
 - [5] 雷志彬,陈俊霖.基于全连接 LSTM 的心肺音分离方法[J].自动化与信息工程,2018,39(6):25-30.
 - [6] POEMOMO A, KANG D K. Biased dropout and crossmap dropout: learning towards effective dropout regularization in convolutional neural network[J]. Neural Networks, 2018, 104: 60-67.
 - [7] 李昊宇,潘欣裕,郁凯,等.使用压电传感器的心肺音采集系统研究[J].单片机与嵌入式系统应用,2019(3):62-66.
 - [8] NERSISSON R, NOEL M M. Heart sound lung sound separation algorithms:a review[J].Journal of Medical Engineer and Technology, 2017, 41(1): 13-21.
 - [9] 黄衍标,陈华珍,黄键鹏,等.基于腕部可穿戴设备的跌倒监护系统设计与实现[J].计算机测量与控制,2019,27(1): 102-106.
 - [10] 董莺艳.基于深度学习的声纹识别方法研究[D].重庆:重庆理工大学,2019, 11-18.
 - [11] 谭朝文,王威廉,宗容,等.卷积神经网络应用于先心病心音信号分类研究[J].计算机工程与应用,2019,55(12): 174-180.
 - [12] 罗元,吴承军,张毅,等.Mel 频率下基于 LPC 的语音信号深度特征提取算法[J].重庆邮电大学学报(自然科学版),2016,2: 174-179.
 - [13] 杜永萍,赵晓铮,裴兵兵.基于 CNN-LSTM 模型的短文本情感分类[J].北京工业大学学报,2019,45(7): 662-670.
 - [14] 朱春利,李昕.基于多特征融合与动态阈值的语音端点检测方法[J].计算机工程,2019,45(2): 250-257.
 - [15] 冯沛,白静,薛珮芸,等.基于多特征组合的普通话塞音识别[J].现代电子技术,2019,42(8): 159-163.

(收稿日期:2020-09-29)

作者简介：

牟俊杰(1990-),男,硕士,助教,主要研究方向:嵌入式系统应用。

姚刚(1985-),男,博士,讲师,主要研究方向:嵌入式系统应用。

孙涛(1980-),男,博士,副教授,主要研究方向:武器系统控制。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所