

摘要

随着生活方式和饮食习惯的变迁，糖尿病的流行趋势出现新动向，对个人精神健康和社会经济构成显著负担。由于尚未有根治糖尿病的有效医疗手段，频繁检测血糖成为必要措施。传统指尖取血法虽成本低廉，但患者体验差且易感染。随着信息技术的进步，智慧医疗平台得以快速发展，出现了基于物联网和人工智能等技术的可穿戴便携式设备，使得基于智能手机PPG信号的无创血糖检测成为可能。然而，现有的无创血糖检测技术仍存在成本高、响应慢和便携性差等问题。因此，开发一种简单快捷的无创血糖检测方法显得尤为重要。

本课题旨在通过PYTHON开发环境，运用深度学习算法，提出一种基于智能手机PPG信号的无创血糖检测技术，该技术采用LSTM模型进行训练，基于训练后的模型为用户进行血糖检测。本文首先阐述了相关的研究背景、意义及当国内外研究现状，然后介绍有关PPG信号的理论和关键技术，以及如何识别并提取PPG信号，计算血糖特征参数R。接着，详细分析了系统的功能需求和运行环境，判断是否能够搭建LSTM神经网络模型。最后，根据LSTM神经网络架构搭建测试模型，对该模型进行了细致的设计和测试，得到了训练集和测试集的结果。

关键词： 糖尿病，无创血糖检测，PPG信号，血糖检测，LSTM

ABSTRACT

With changes in lifestyle and eating habits, new trends in the prevalence of diabetes have emerged, posing a significant burden on individual mental health and socioeconomics. Since there is no effective medical treatment to cure diabetes, frequent blood glucose testing has become a necessary measure. Although the traditional fingertip blood collection method is low-cost, it has poor patient experience and is prone to infection. With the advancement of technology, such as wearable devices, the Internet of Things and artificial intelligence, smart medical platforms have developed rapidly, providing the possibility of non-invasive blood glucose detection based on smartphone PPG signals. However, current non-invasive blood glucose detection technology still faces challenges such as high cost, slow response, and poor portability. Therefore, it is particularly important to develop a simple and fast non-invasive blood glucose detection method.

This topic aims to use the PYTHON development environment and deep learning algorithms to propose a non-invasive blood glucose detection technology based on smartphone PPG signals. This technology uses the LSTM model for training, and uses the trained model to detect blood glucose for users. This article introduces the PPG signals theory, and how to identify and extract PPG signals. Then, the functional requirements and operating environment of the system were analyzed in detail to determine whether the LSTM neural network model could be built. Finally, the model was carefully designed and tested, and the results of the training set and test set were obtained.

KEY WORDS: diabetes, non-invasive blood glucose monitoring, PPG signal, blood glucose monitoring, LSTM

目 录

第一章 绪论.....	1
1.1 课题背景和研究意义	1
1.2 国内外研究现状	3
1.2.1 无创血糖检测的发展历程	3
1.2.2 无创血糖检测研究现状	4
1.3 论文主要研究内容及组织结构	5
第二章 无创血糖检测的理论基础.....	7
2.1 PPG 信号的特点及波形特征	7
2.1.1 PPG 信号的特点	7
2.1.2 PPG 信号的波形特征	8
2.2 PPG 信号检测	9
2.2.1 检测方式	9
2.2.2 波长的选择	10
2.3 朗伯-比尔定理的推导	13
2.4 本章小结	14
第三章 PPG 信号预处理及特征提取.....	15
3.1 数据采集和预处理	15
3.1.1 数据采集	15
3.1.2 信号提取	16
3.1.3 滤除 PPG 信号高频噪声干扰	17
3.1.4 消除 PPG 信号基线漂移	19
3.2 PPG 信号特征值提取	21
3.2.1 PPG 信号综合特征参数 R 的计算方法	21
3.2.2 初步建立血糖检测模型	23
3.3 特征参数提取与标准化	24
3.4 本章小结	25
第四章 建立无创血糖检测模型.....	26
4.1 相关技术简介	26
4.1.1 PYTHON 简介	26
4.1.2 NumPy 简介	26
4.1.3 Pandas 简介	26

西北工业大学本科毕业设计（论文）

4.1.4	Keras 简介	27
4.1.5	Sklearn 简介	27
4.1.6	LSTM 简介	28
4.2	系统分析	28
4.2.1	系统功能需求	28
4.2.2	技术可行性	28
4.2.3	操作可行性	29
4.3	LSTM 模型	29
4.3.1	LSTM 模型构建流程	29
4.3.2	LSTM 模型参数设置	30
4.4	系统模型设计	31
4.4.1	实验步骤	31
4.4.2	LSTM 模型设计	32
4.5	本章小结	32
第五章 程序运行结果及分析		33
5.1	建模所调用的库	33
5.2	程序设计	33
5.2.1	输入原始信号并进行预处理	33
5.2.2	计算特征系数	34
5.2.3	特征矩阵热力图可视化	36
5.2.4	数据归一化处理	36
5.2.5	构建 LSTM 模型	37
5.2.6	训练损失图	38
5.2.7	划分数据	38
5.3	训练结果及分析	39
5.4	本章小结	40
参考文献		41
致 谢		43
毕业设计小结		44

第1章第一章 绪论

1.1 课题背景和研究意义

在当今社会，由于人们普遍存在的不良饮食习惯和不良生活方式，许多疾病的患病率急剧攀升^[1]。糖尿病作为一种长期存在的健康挑战，其根源在于胰腺未能分泌足够的胰岛素，或者人体对于已产生的胰岛素无法高效利用。当糖尿病得不到有效控制时，高血糖状态便成为常态，随着时间的推移，这种持续的高血糖状态会逐步侵蚀人体的多个系统，特别是神经系统和血管系统，造成严重的健康损害。

据世界知名医学期刊柳叶刀 (*The Lancet*) 的数据^[2]如图1-1所示，到2021年，全球成人糖尿病患者数量将达到5.4亿，占成年人口的10%。

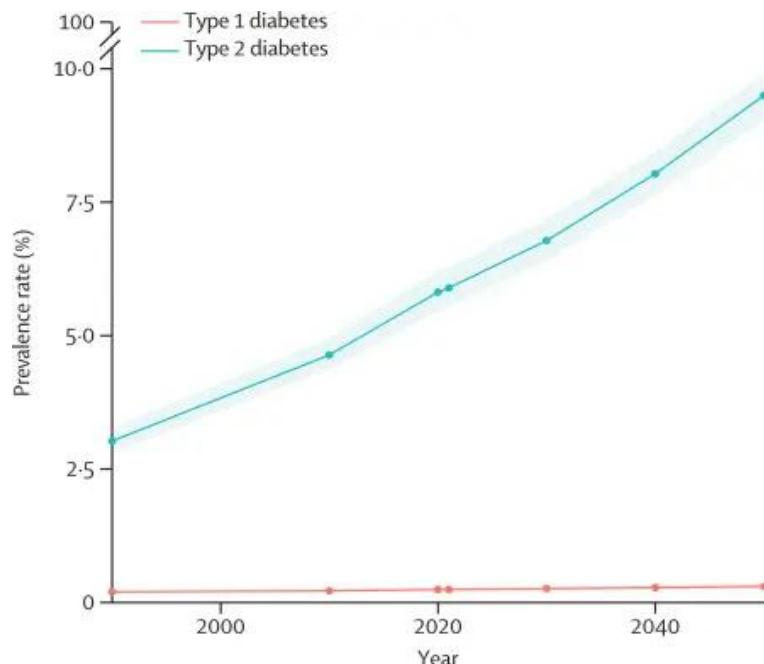


图 1-1 柳叶刀发布 1900 至 2050 年（预测）的糖尿病患病率

根据国际糖尿病联合会 (IDF) 近期发布的统计数据^[3]，2019年全世界范围内，身患糖尿病的人数大约达到了4.63亿。同时，该联合会还做出了预测，预计到2030年，全球将约有5.78亿成年人会身陷糖尿病的困境。在众多国家中，中国的糖尿病患者数量位居榜首。令人震惊的是，2019年中国有83万人因糖尿病及其引发的各种并发症而离世。在这些不幸离世的病患中，有一个现象特别值得关注：高达33.4%的患者去世时年龄还未到60岁。这一数据无疑揭示了糖尿病死亡病例

西北工业大学本科毕业设计（论文）

正逐渐趋向年轻化^[4]，这是一个需要社会各界高度警觉和积极应对的严峻问题。

与此同时，世界人口预计将增加20%，糖尿病患者人数将增加50%。超过70%的患者年龄在50岁以上。其中，糖尿病患病人数排名前三位的国家分别是中国、印度和巴基斯坦，中国约有1.5亿患糖尿病，约占人口总数的10%，预计到2040年将超过1.65亿。

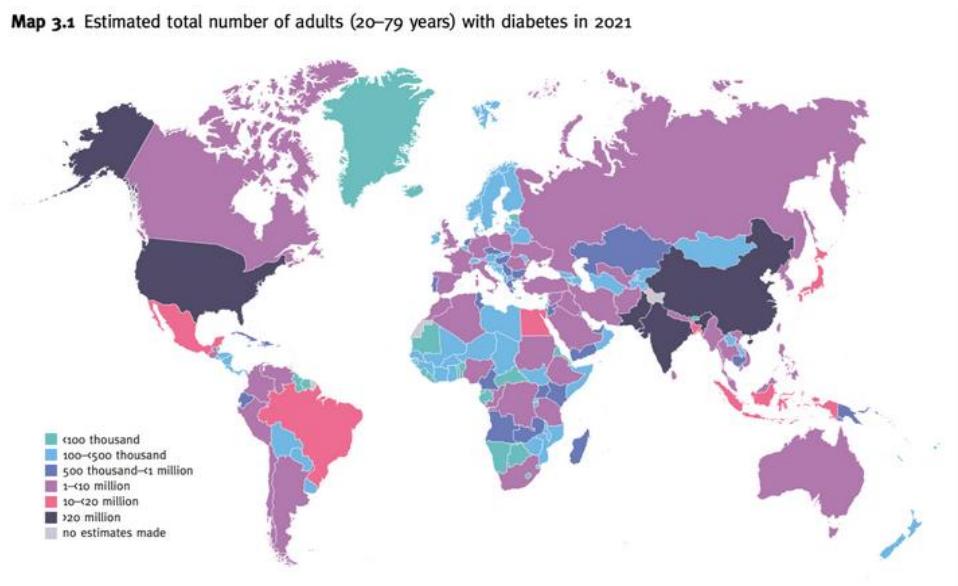


图 1-2 截至 2021 年世界范围内糖尿病成年患者地图

糖尿病是一种目前尚无法根治的慢性疾病，其根源在于身体调节血糖水平的能力出现障碍，即胰岛素功能受损。胰岛素失衡可能由多种因素导致，包括自然发生原因、不健康的饮食习惯以及缺乏足够的体力活动等。根据这些因素，糖尿病可被分类为1型和2型^[2]两种。

作为一种难以彻底治疗的慢性疾病，糖尿病的病情控制显得尤为重要。实现这一目标的关键是通过定期测量血糖水平来监测疾病的进展。由于糖尿病早期症状相对轻或毫无症状，因此血糖测试是诊断疾病的唯一有效手段，持续关注和检测血糖水平是糖尿病患者疾病控制和预防并发症的重要组成部分。

目前，血糖检测的主要手段是有创血糖检测，它通过采集人体血液样本直接测量血液中的葡萄糖浓度，因其准确度高（通常在95%以上）而广受信赖。然而，这种方法也有一些明显的局限性。为了监测血糖波动，患者每天至少需要进行7次抽血和化验，频繁的有创血液检测一方面给患者带来身体不适和感染风险，另一方面无法实现持续实时血糖监测，且耗时长、成本高。因此，寻找一种更加便捷、无痛、可持续的血糖检测方法成为当前该领域的研究重点。

本课题基于PPG技术提出了一种创新的无创血糖检测方法，该方法充分利用

西北工业大学本科毕业设计（论文）

智能手机捕捉PPG信号，以实现患者自我监控身体状况的移动医疗目标，不仅具有一定的技术创新性，在未来的医疗服务中也有重要的应用价值。

1.2 国内外研究现状

1.2.1 无创血糖检测的发展历程

无创血糖检测的研究始于20世纪80年代，可以分为3个阶段：1987年至1999年是第一阶段，Futrex、Biocontrol等大公司推出了无创血糖测量的原型^[5]。尽管血糖检测设备在技术上取得了很大进步，但这些设备的精度仍难以满足临床应用的要求。因此，众多科研团队在血糖检测领域持续深耕，不断探索新的技术与方法，以进一步提升血糖检测的准确性和可靠性，为患者提供更加精准的健康管理方案。

1999年至2005年是第二阶段，随着科研工作的深入与进展，该项目的巨大潜力和复杂性逐渐显现，引起了众多发达国家的重视，纷纷为该项研究投入巨额资金支持^[6]。在日本，为了攻克无创血糖检测技术，经济部联合松下电器、东京理工大学等单位共同组成了一个项目组，并投入了超过六亿元的资金^[7]。与此同时，1992年，美国的衣阿华大学研制成功了一套连续血糖检测设备，并开展了大量相关研究工作^[8]，但进展不大。

第三阶段是2006年至今，许多机构根据新的测量原理和方法开发了许多血糖测量设备^[9]。美国的维拉莱特公司在2011年成功研制ScoutDS无创血糖仪，并获得欧洲CE认证，该设备专为2型糖尿病患者设计微软公司与谷歌公司合作开发了一款智能隐形眼镜，通过检测眼泪中的葡萄糖浓度来估算人体的血糖水平。尽管这一研究成果颇具创意，但其技术水平尚临床应用中血糖检测的要求。

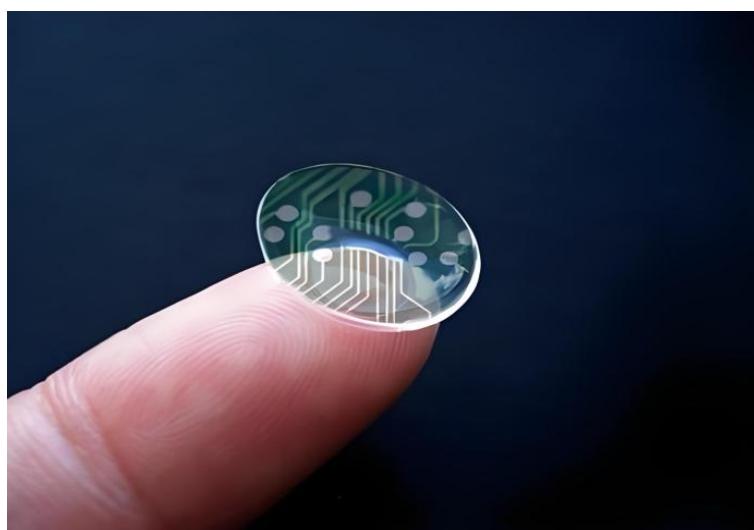


图 1-3 谷歌隐形眼镜式的血糖检测仪

1.2.2 无创血糖检测研究现状

目前，糖尿病患者尚不能得到根治，不得不依赖于药物治疗、规律运动和饮食调整等多重策略来管控病情。

传统的有创血糖检测需要频繁地进行手指部位采集血样，这种方法一方面给患者带来身体上的不适和心理压力，另一方面还会因操作不当而导致血液样本变异，引发测量误差甚至感染风险^[10]。



图 1-4 微创血糖仪

自本世纪初以来，微创血糖检测技术作为有创血糖检测技术的升级，逐渐进入了科研人员的视野。通过在身体某些部位（如手臂）安装传感器，实时监测血液中的葡萄糖浓度及其变化趋势^[11]，能够实时获取病人血液中的葡萄糖水平，使得血糖监测变得更加高效便捷，为糖尿病患者提供了更好的管理和治疗选择。然而，该技术对糖尿病前期患者的效用并不明显，且高昂的设备成本限制了其在日常持续血糖监测中的广泛应用。

当前，无创血糖检测技术是科研的重点。数十年间，此技术已获多项进展，主要分为光学式与非光学式检测。光学式检测利用特定波长（如红外光等）光线与人体组织的相互作用，观察入射光属性的变化。非光学式血糖检测技术聚焦于血糖波动的影响，涵盖超声波、电阻抗等多种方法，该技术正朝着更精准、便捷的方向发展。

红外光谱法因利用红外光与人体组织相互作用推算血糖浓度，已成为无创血糖检测的主流技术。科研机构和企业正致力于提高该方法的检测精度并减少外界干扰。尽管商业化产品如Veralight的ScoutDS已获得欧洲CE认证，且俄罗斯也研

发出高精度无创血糖检测仪，但红外光谱法的应用仍面临挑战，期待未来技术的更广泛应用和精确性的进一步提升。



图 1-5 Gluco Track 产品示意图

以色列的公司研制了无创血糖检测产品Integrity Applications Gluco Track, 如图1-5所示，雅培公司的血糖仪在血糖检测领域引入了能量代谢守恒法^[12]，该款血糖仪融合超声波和电磁波等技术，采用夹在耳垂的传感器，实现对血糖水平的非侵入式检测，为糖尿病患者带来了便捷的血糖管理体验。

我国在无创血糖检测技术方面也取得了很大的进展。李晓霞采用动态光谱频域提取法，有效降低了个体差异对检测的影响。唐飞团队则通过代谢守恒法，揭示了葡萄糖分解中生理参数与血糖浓度的关系。第三军医大学尝试基于唾液的检测，为无创血糖检测提供新思路。米永巍等人研究PPG信号与血糖关系，力求临床验证。尽管存在传感器误差等挑战，但科研人员的努力将持续推动技术优化^[13]。

综上所述，无创血糖检测技术仍有许多难题待解，市场上也缺乏公认的无创血糖检测仪。该技术易受干扰，人体复杂性、生理多变性和个体差异性都会增大检测误差。虽有研究者尝试用神经网络建模，但模型及精度需改进。因此，该技术还需时间发展。

1.3 论文主要研究内容及组织结构

本文旨在探讨通过光电容积描记图（PPG）信号实现连续无创血糖监测的方法。研究内容主要包括以下方面：

(1) 首先分析了PPG信号形成的生理基础，接着研究了PPG信号的波形特征，而且系统研究了基于PPG信号的无创血糖检测的基本原理、过程。同时，PPG信号的采集方法，包括确定采集部位、检测方式以及入射光波长的选择也是文章研究的重点。

(2) 为了提高采集到的信号的质量，文章中采用不同的滤波技术去除信号中的高频噪声干扰，对PPG信号进行预处理，例如中值滤波法和三次样条插值法。随后，从预处理后信号中提取与血糖值相关的特征参数，以备后面建模适用。

西北工业大学本科毕业设计（论文）

(3) 通过前文的理论基础，着手建立基于LSTM神经网络的无创血糖检测模型，根据提取出的特征参数建立于血糖值之间的数学关系，并且验证其之间的线性或非线性关系。

论文的主要结构如下：

第一章：绪论

首先概述了糖尿病作为慢性疾病的全球影响，阐述了如今对于糖尿病的预防方法就是检测并监控血糖。随后，介绍了有创血糖检测方法技术、微创血糖检测技术以及无创血糖检测技术，将三者进行对比，并结合如今检测血糖的研究方向，发现有无创血糖检测技术最适合当下医疗发展以及社会发展。紧接着，本文讲述了无创血糖检测的发展历史以及国内外研究现状。最后则阐明了文章的研究方向和基础架构。

第二章：无创血糖检测的理论基础

概述了无创血糖检测的相关理论基础。首先，介绍了PPG信号产生的生理因素以及波形特性；接着，通过引入Lambert Beer定理，通过数学公式的推导，探讨了PPG信号特征参数R与血糖值C之间的相关性。最后，文章深入研究了可能影响无创血糖检测的生理参数，以及选取恰当波长的入射光作为外界输入。

第三章：PPG信号预处理及特征值提取

主要介绍了PPG信号的预处理和特征值提取过程。首先，介绍了PPG原始信号的来源，以及如何采集信号的流程。其中，保留红色通道和绿色通道采集到的原始信号作为输入进行处理，通过中值滤波和三次样条插值两种信号处理方法，有效解决了PPG信号中的高频噪声干扰和基线漂移的问题，使得原始信号可以进一步优化作为模型输入。最后，计算了其中的特征参数。基于这些参数，初步建立了基于LSTM的血糖检测模型。

第四章：建立无创血糖检测模型

介绍了基于LSTM神经网络算法的无创血糖检测模型的构建。经过文献阅读，本文决定采用的是Python与Keras的结合来进行编程，以此为基础设计一款血糖监测系统。不仅如此，本文选用Adam作为优化策略来训练LSTM模型。本文模型主要围绕全连接层和LSTM层进行设计，并最终将这些特征投射到输出空间，以此研究不同的隐含层结构对模型预测准确度的影响。

第五章：程序运行结果及分析

通过编写代码创建和训练LSTM模型，实现了对血糖值的监测，并计算最终的识别率。在模型训练过程中，对识别效果进行了对比和优化。最后，对系统的各项功能进行了测试，展示了训练集和测试集的样例图。

第2章第二章 无创血糖检测的理论基础

2.1 PPG 信号的特点及波形特征

人体中的血液循环系统至关重要，它包括三个主要部分：血液、血管和心脏，称为心血管系统，对维持人体正常生命活动极为关键。心脏作为人体的中央泵站，通过收缩、舒张，将血液经动脉推送到各个器官和细胞中，为它们提供氧气和营养物质，再由静脉回流至心脏，完成一次血液循环^[14]。

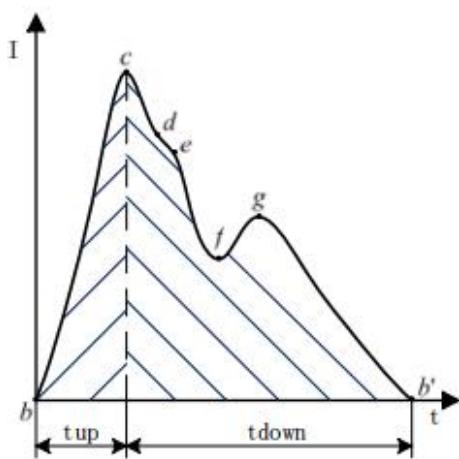


图 2-1 光电容积脉搏波理论图

血液循环在过程上中可分解为两大部分：“上升”和“下降”阶段，如上图所示，“上升”阶段是心脏的收缩过程，而“下降”阶段是心脏的舒张过程。经历这两大过程后，为周边器官提供了必要的血液供应，这一完整的泵血过程所产生的动脉内压力变化，称为“脉搏波”，即PPG信号。

2.1.1 PPG 信号的特点

(1) PPG信号的强度较弱，而且频率相对其他人体信号来说较低。具体的说，PPG信号的振幅很小，其频率分布范围在40Hz以下。

(2) PPG信号极易受到干扰。在收集或检测PPG信号的过程中，极易受到来自外界的干扰，例如自外界电磁场、人工或自然光源等。同时，人体的大幅度动作，甚至是呼吸等生理活动也可能对信号产生严重干扰。

(3) PPG信号具有不确定性。这种不确定性体现在不同个体之间，以及同一个人体在不同条件下采集到的PPG信号差异上。例如，当人体的情绪、呼吸等状态

发生改变时。

2.1.2 PPG 信号的波形特征

PPG 信号波形含有大量的人体生理信息，其结构通常包含两个主要组成部分：直流 (DC) 分量与交流 (AC) 分量。这两个分量共同塑造了PPG信号的独特且复杂的特性。直流分量反映了信号的平均水平，而交流分量则包含了与心脏跳动相关的周期性波动信息。

PPG 波形的 DC 分量对应于从组织中检测到的透射或反射光信号，并取决于组织的结构以及动静脉血和静脉血的平均血容量^[15]。

在人体生理信号分析中，直流成分通常表现为与呼吸频率相匹配的缓慢波动，而交流成分则更为活跃和动态。交流成分直接反映了心脏跳动周期内血容量的变化，这是心动周期不可或缺的要素。交流分量的基本频率与心率紧密相关，它叠加在DC成分之上，共同描绘出人体生理信号的完整图像。由于交流分量包含了身体部位中血容量随心脏跳动而呈现的周期性变化，因此血液流动信息表现在其中^[16]，如图 2-2 所示。

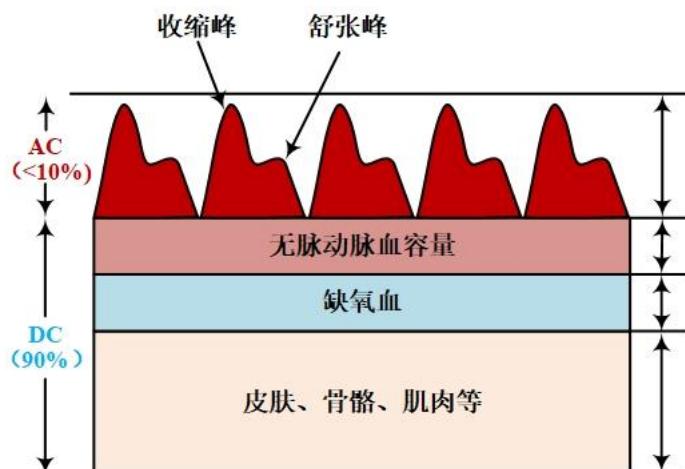


图 2-2 PPG 信号的波形成分组成

一个完整的单周期PPG信号如图2-3所示，由上升阶段up和下降阶段down组成，并具备B、C、D三个显著的特征点。具体来说，B点代表主波峰，即信号的最大值；C点为重博波谷，位于主波峰之后，表示信号的局部最小值；D点则为重博波峰，紧接着重博波谷之后，它标志着信号在下降支中的另一个峰值，这些波峰波谷在PPG信号的不同周期一直处于变化阶段，这同时也反映了人体不同的生理病理差异^[17]，以及PPG信号极易收到干扰的特点。

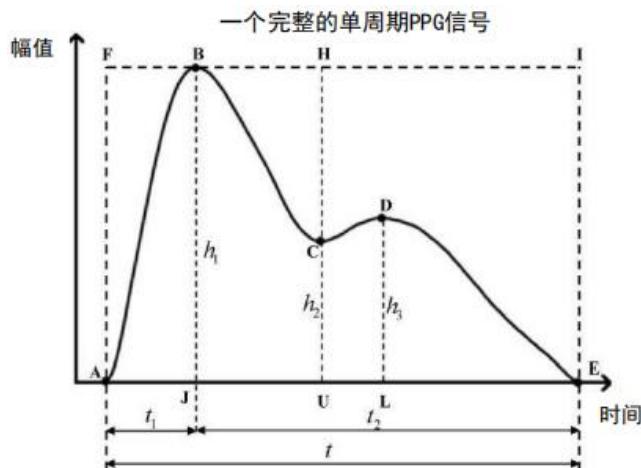


图 2-3 一个完整的单周期 PPG 信号

PPG波形所展现的独特特征，特别是其下降支的形态，与心脏外围阻力之间存在着密切的对应关系^[18]。因为其独特的性质，PPG信号检测装置已经成为研究人体生理状态的关键工具。通过对PPG信号的细致分析，不仅能够实现人体心血管疾病的无创检测，还能为疾病的早期预防、早期发现以及早期治疗提供强有力的支持，从而显著提高医疗诊断和治疗的效率与准确性。

2.2 PPG 信号检测

2.2.1 检测方式

PPG装置一般由LED灯和探测反射或透射光的光电探测器组成^[19]。当血液循环发生波动时，血液对于光吸收量也会发生波动，从而产生关键的PPG信号。通过对这些信号进行精确的处理和细致的分析，能够深入洞察人体的血流动力学特征，并揭示血液成分中新颖而重要的信息。

当外界仪器的入射光照在人体部位时，光会在表皮，毛孔，肌肉，血管等部位发生反射，吸收，折射，散射等现象^[20]，如图2-4所示：

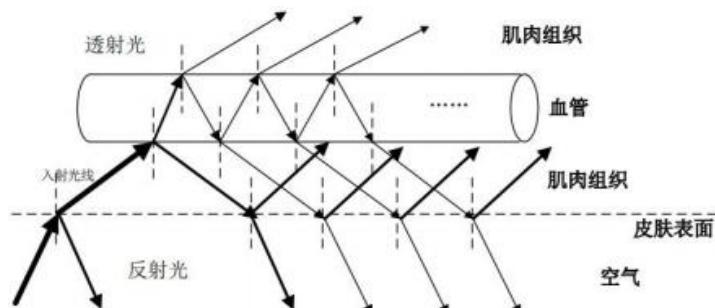


图 2-4 入射光进入血液示意图

由于探测光的反射和透射占据了接收光线的大部分，PPG 信号的传感衍生出两种不同的测量方式^[21]：一个是反射模式，另一个是透射模式，如图2-5所示。

光电探测器的光源和接收器位于同一侧的方法是反射模式；而光电探测器的光源和接收器位于身体部位的相向一侧的则是透射模式。透射式 PPG 传感器通常放置在耳唇部分，而反射式 PPG 传感器可以放置在血管靠近皮肤表面的指尖、手腕部附近^[22]，如图2-5所示，模拟人手指接受外界入射光源检测PPG信号。

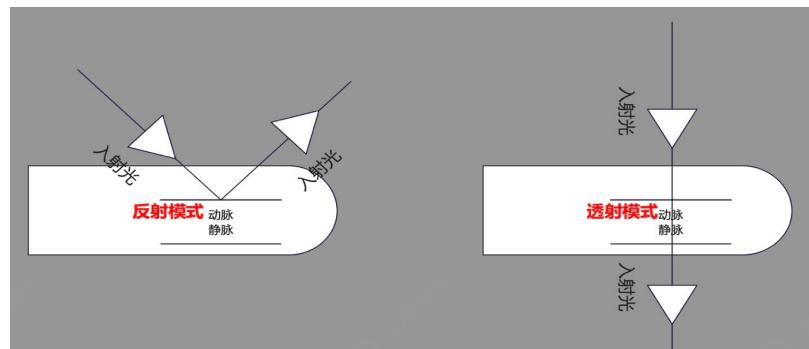


图 2-5 PPG 信号的两种不同测量方式

透射式检测经常用于检测较为薄弱的部位，因为这些区域的透光性能优越，使得血液容积的微小变化能够直接、清晰地被捕捉到。而在面对手臂、大腿等组织相对丰厚的部位时，反射式检测方式则更具优势。尽管这些区域的透光性相对较弱，但反射式检测能够精准地捕捉到从深层组织反射回来的光信号，从而有效地反映出血液容积的波动。

在反射式PPG信号检测中，LED光源发出的光信号需历经多层组织，直至抵达光电探测器（PD）^[23]。这些组织层次对光信号进行不同程度的散射和吸收，对信号的强度和质量产生直接影响。根据Lambert-Beer定律，长波长光因其更强的穿透力，能深入人体深层组织。然而，在实际应用中，LED光源发出的光功率在穿越这些组织层时，会由于散射和吸收作用而发生衰减，这种衰减可以通过吸收系数的变化来精确衡量。

本文聚焦于反射模式，通过精心策划的实验设计和数据处理流程，更精确地捕捉和反映血液容积的动态变化及血液动力学的核心特征，为医学研究和临床实践提供有力支持。

2.2.2 波长的选择

在检测PPG信号时，要获得预期的检测精度，必须选择合适的检测波长^[24]，首要考虑的就是在接收外界入射光后的能被葡萄糖分子高效吸收的光波长。

作为对照，需要选定一个葡萄糖的非敏感波长，即在该波长下，葡萄糖对光

西北工业大学本科毕业设计（论文）

的吸收度较低。为了确保测量的精确性，波长应对除葡萄糖以外的其他物质展现出极高的排斥度，也就是不容易被其他物质所吸收。而作为评估健康状况的关键指标，人体血糖浓度，也直接反映了葡萄糖在血液中的浓度。如图2-6所示，葡萄糖的分子结构由C₆H₁₂O₆构成是进行波长选择时需要考虑的重要化学基础。

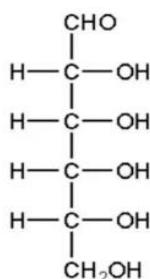


图 2-6 葡萄糖分子式

如图2-6所示，葡萄糖分子的构成主要是甲基和羟基，对于自然光有十分特殊的频率吸收的特性^[24]，选择自然光作为检测光，以便能够准确地捕捉这些基团对光的响应。

具体来说，葡萄糖在1500-1800nm的波段内展现出一阶倍频吸收，而在1000-1350nm的波段则呈现二阶倍频吸收。然而，值得注意的是，血液中水的含量占据很大比重，且水在这些波段内也有特定的吸收峰，因此，在选择检测波长时，需要尽量避开水的吸收峰，水的吸收峰为1550-2000nm波段^[25]。

在进行反射式检测时，为确保测量精度，需要皮肤组织对检测光的吸收度保持在较低水平。如图2-7所示，皮肤组织在550至1350nm的光谱范围内显示出较低的吸收特性，这使得这一区间成为潜在的检测波长选择范围。然而，值得警惕的是，脂肪在特定波长范围内，尤其是接近1000nm和2100nm时，展现出较高的吸收率。这一发现对于选择最佳检测波长具有重要的指导意义，确保避开这些高吸收区域，以获得更准确的测量结果。

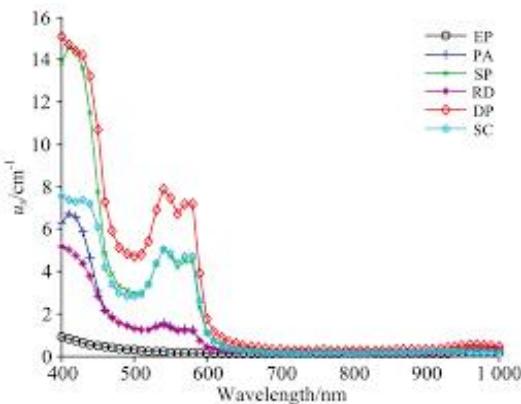


图 2-7 吸收检测波长示意图

除了上述要考虑的物质，还需要考虑血液中其他物质对光的吸收，如蛋白质，甘油三酯，尿素等^[26]，如表2-1所示：

表 2-1 血液中各种物质的光吸收范围

中文名称	英文名称	一般范围	吸收范围
球蛋白	GLO	20~30 g/L	500 ~ 600nm
白蛋白	ALB	40~55 g/L	450nm
总蛋白	TP	60~80 g/L	1000 ~ 1500nm
胆固醇	CHO	4~5 mmol/L	0 ~ 500nm
总肝红素	TBIL	3~17 mmol/L	0 ~ 200nm
甘油三酯	TG	0.25~1.8 mmol/L	0 ~ 900nm
肌酐	CRE	0.05~0.13mmol /L	0 ~ 100nm
尿素	BUN	2.8~7.5 mmol /L	0 ~ 80nm
尿酸	UA	0.18~0.3 mmol /L	0 ~ 30nm

深入分析表2-1的数据，可以发现血液中绝大多数物质的吸收峰值都集中在低于1000nm的波长范围内。基于这一关键发现，将检测波长的选择范围限定在经过综合考虑，选定了波长为660nm的光作为本研究的检测波长。

而智能手机LED灯光的波长也恰好在这个范围之内^{[27][28]}，故此，选取智能手机作为PPG信号采集设备是可行的。

本研究旨在深入探究并验证特定参数间的关联性及智能手机在PPG信号采集方面的潜力。智能手机上的闪光灯与摄像头成为了实现PPG信号采集的关键硬件组件。通过本研究中研发的信号提取算法，这些通用设备能够高效且准确地提取出PPG信号。

因此，本研究的创新之处在于将智能手机应用于PPG信号的收集，并进一步实现了血糖检测的功能。这一方法不仅具有高度的便携性，让用户能够随时随地监测自己的健康状况，而且成本效益显著，为日常健康管理提供了经济有效的解决方案。

2.3 朗伯-比尔定理的推导^[29]

朗伯-比尔定理建立了液体浓度与吸收光频之间的线性关系：溶液浓度越高，光吸收越强。此外，该定理还表明光照强度不会改变溶液对光的吸收比例，为不同光照下的血糖测量提供了理论支撑。

根据Lambert-Beer定理的概念，一束光强为 I_0 ，波长为 λ 的入射光，当它照射到人体表面时，会发生透射，而透射后的光强 I 为：

$$I = I_0(e^{-\varepsilon_0 c_0 L} e^{-\varepsilon_1 c_1 L} e^{-\varepsilon_2 c_2 L}) \quad (2-1)$$

L 表示光路长度； ε_0 和 c_0 分别表示其他物质对光的吸收率以及所吸收物质的浓度； ε_1 、 c_1 分别表示人体血液中葡萄糖的吸光系数和浓度； ε_2 、 c_2 分别表示人体血液中参考物质的吸光系数和浓度。

值得注意的是，人体血管壁具有弹性，而这种弹性会使得血管会随着心脏收缩与舒张发生形变，最终会使得光路长度增加，记为 ΔL ，由此可知光强的变化量是：

$$I + \Delta I = I_0 e^{-\varepsilon_0 c_0 L} * e^{-(\varepsilon_1 c_1 + \varepsilon_2 c_2)(L + \Delta L)} \quad (2-2)$$

将以上两个式子相除，可得：

$$\frac{I + \Delta I}{I} = e^{-(\varepsilon_1 c_1 + \varepsilon_2 c_2) \Delta L} \quad (2-3)$$

当直流分量占据主导地位时，即 $(\Delta I / I) \ll 1$ ，对式 (2-3) 两边同时取对数：

$$\ln(1 + \frac{\Delta I}{I}) = -(\varepsilon_1 c_1 + \varepsilon_2 c_2) \Delta L \quad (2-4)$$

本文采用了两路不同波长的光源来采集PPG信号，波长分别为 λ_1 和 λ_2 ， I^1 和 I^2 表示两束入射光的光强， $\varepsilon^{\lambda 1}$ 和 $\varepsilon^{\lambda 2}$ 表示两束入射光的吸光度，因此，公式 (2-4) 可写为：

$$\frac{\Delta I^{\lambda_1}}{I^{\lambda_1}} = -(\varepsilon_1^{\lambda_1} c_1 + \varepsilon_2^{\lambda_1} c_2) \Delta L \quad (2-5)$$

$$\frac{\Delta I^{\lambda_2}}{I^{\lambda_2}} = -(\varepsilon_1^{\lambda_2} c_1 + \varepsilon_2^{\lambda_2} c_2) \Delta L \quad (2-6)$$

用式 (2-5) 减去式 (2-6)，可得：

$$\frac{\Delta I^{\lambda_1}}{I_1} - \frac{\Delta I^{\lambda_2}}{I^{\lambda_2}} = c_1 \Delta L (\varepsilon_1^{\lambda_2} - \varepsilon_1^{\lambda_1}) + c_2 \Delta L (\varepsilon_2^{\lambda_2} - \varepsilon_2^{\lambda_1}) \quad (2-7)$$

为排除葡萄糖以外物质对光强的吸收，应当选取在波长 λ_1 和 λ_2 下吸光系数相同的参考物质，即 $\xi_2^{\lambda_2} - \xi_2^{\lambda_1} = 0$ ，则变换式 (2-7) 可得血糖浓度：

$$C_1 = \frac{1}{(\varepsilon_1^{\lambda_2} - \varepsilon_1^{\lambda_1}) \Delta L} \left(\frac{\Delta I^{\lambda_1}}{I^{\lambda_1}} - \frac{\Delta I^{\lambda_2}}{I^{\lambda_2}} \right) \quad (2-8)$$

将式中 $\frac{\Delta I^{\lambda_1}}{I^{\lambda_1}} - \frac{\Delta I^{\lambda_2}}{I^{\lambda_2}}$ 记为综合特征参数 R ，则血糖浓度可表示为：

$$C_1 = \frac{1}{(\varepsilon_1^{\lambda_2} - \varepsilon_1^{\lambda_1}) \Delta L} R \quad (2-9)$$

其中， $\xi_1^{\lambda_1}$ 和 $\xi_1^{\lambda_2}$ 分别表示葡萄糖对波长为 λ_1 和 λ_2 的入射光的吸收度，为常数。 ΔL 可以视为一个相对恒定的常量。则式 (2-9) 可写为：

$$C_1 = a * R \quad (2-10)$$

其中， C_1 为本文计算的血糖初值。

为提高血糖检测的准确性，需有效排除检测过程中的内外干扰因素，以确保 PPG 信号的准确提纯，并识别出影响 PPG 波形的人体参数。这对优化检测方法和提升精度至关重要。

2.4 本章小结

首先，本章详细阐述了 PPG 信号产生的生理原理，接着阐述了光学传感器如何通过对人体组织透射或反射光的检测来采集 PPG 信号。不仅如此，本章还结合人体生理参数，对 PPG 信号的波形特征进行了深入分析，详细研究了 PPG 信号采集的技术细节。最后，引入了 Lambert Beer 定理并推导，探讨了 PPG 信号特征参数 R 与血糖值 C_1 之间的定量关系。

第3章第三章 PPG信号预处理及特征提取

3.1 数据采集和预处理

为了满足实验需求，本文所使用的数据来源于Azure开放数据集服务。以下是对数据采集和预处理步骤的客观描述，旨在确保数据集的可信度和可用性。

3.1.1 数据采集

鉴于PPG信号易受温度和环境光影响，数据采集过程在特定环境下进行：环境保持安静，温度控制在 $24 \pm 2^{\circ}\text{C}$ ，且无强光源干扰，从而确保数据的准确性和可靠性。在本次研究中，80名受试者的生理信息被征集作为数据源^[30]。其中，30名受试者的血糖水平属于正常范围，占比37.5%；而50名受试者被诊断为高血糖，占比62.5%。表3-1是受试者的信息，表3-2是血糖范围。

表 3-1 受试者信息

项目	最大值	最小值	平均值	标准偏差值
年龄	73	42	47.62	5.80
体重	76.2	48.3	63.86	10.67
身高	1.83	1.43	1.612	0.15
BMI	26.37	16.67	22.46	5.34
血糖	9.64	3.87	7.75	3.17

表 3-2 血糖范围

组别	血糖范围(mmol/L)	等级
M1	4.0 ~ 6.2	正常
M2	6.3 ~ 7.7	临界
M3	7.8 ~ 10	警告

在数据采集阶段，研究人员使用iPhone X Plus手机相机，设定在120帧率的模式下，对每位受试者的左手食指进行了60秒的视频拍摄。在拍摄过程中，要求受试者的指尖同时覆盖闪光灯和相机镜头，以确保拍摄到足够清晰的手指血容量变化。这些视频记录了手指血容量变化引起的明暗切换现象，为后续的数据分析

西北工业大学本科毕业设计（论文）

提供了重要依据。完成视频拍摄后，根据每位受试者测得的血糖值，他们被唯一编号并分配到不同的组别：G1组编号为1-30，G2组编号为31-58，G3组编号为59-80。接着，这些视频被传输到计算机上，使用MATLAB软件进行详细的数据预处理，以便后续进行深入的生理信息分析^[30]。

3.1.2 信号提取

随着手指血容量的变化，相同区域的像素强度在相邻帧之间会发生变化。为了从这些变化中提取出连续的PPG（光电容积描记图）信号，采取以下步骤：

- (1) 视频帧分析：首先，从正确录制的视频中提取出每一帧的图像数据。
 - (2) 颜色通道分离：将每一帧图像中的红、绿、蓝（RGB）三色通道分离出来。由于不同颜色通道对血液变化的敏感度可能不同，这一步骤有助于更精确地分析血容量变化。
 - (3) 阈值设定：为了从像素强度的变化中提取出PPG信号，需要为每一帧设定一个阈值。这个阈值的选择是基于实验经验和波形精度的考量。当阈值设置得较低时，输出信号可能会变得过于平滑，失去一些重要的波动信息；而当阈值设置得较高时，信号中可能会出现不必要的峰值，影响信号质量。
 - (4) 信号计算：在设定好阈值后，通过计算每一帧中超过阈值的像素强度变化，可以得到连续的PPG信号。这个信号反映了手指血容量的动态变化。
 - (5) 阈值优化：在本文中，研究人员通过多次实验和比较，最终将阈值设定为像素强度范围的1.01倍。这个阈值在保持信号波形精度的同时，也避免了不必要的噪声和峰值。PPG信号提取的具体操作如图3-1、3-2、3-3所示：



图 3-1 使用手机的摄像功能录制视频，提取视频帧图片

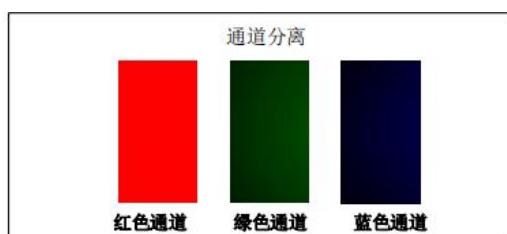


图 3-2 颜色通道分离

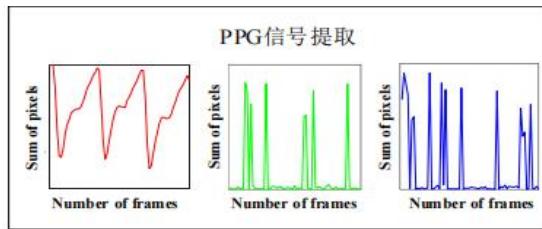


图 3-3 分别从三个通道中提取PPG信号

从图中可以看出，在PPG信号提取的三色通道中，只有红色通道获得的 PPG 信号是最清晰的。作为以后建立并训练血糖模型的原始信号输入，选取红色和绿色通道提取出的PPG信号，以此作为比对，测试建立的模型识别是否准确。

3.1.3 滤除 PPG 信号高频噪声干扰

PPG信号属于0.1Hz-40Hz的低频信号，而电磁干扰等属于50Hz以上的噪声信号，因此需要滤除信号中的高频噪声^[31]。

对于高频噪声的处理，通常根据不同的场合选取不同的方法，这些方法各有其优缺点和适用场景。

低通滤波器是一种简单且直接的方法，它能快速地将高频噪声滤除。然而，当面对较为复杂的噪声干扰时，低通滤波器可能无法完全排除噪声，尤其是在噪声与信号频谱重叠时。小波变换法则是一种更为先进的方法，它相较于传统方法在处理非平稳信号和时变信号时具有显著优势。小波变换可以通过分解信号到不同的频率和时间尺度上，然后对每个尺度上的信号进行去噪处理，从而实现更精细的噪声抑制。然而，小波变换需要准确的建模，且计算时间较长，这限制了其在一些实时性要求较高的应用中的使用。中值滤波则是一种基于统计排序的非线性信号处理技术。它通过取信号中某一点邻域内所有点值的中值来替代该点的值，从而消除孤立的噪声点。中值滤波具有计算简单、实现方便等优点，因此在实时图像处理等领域得到了广泛应用。但中值滤波会出现丢失波峰和波谷的问题，从而导致信号趋于平滑^[28]。

中值滤波是一种常用的非线性滤波技术，然而，传统的中值滤波方法可能会误将正常信号当作噪声而丢失，导致信号的波峰和波谷等重要信息被平滑掉。因此，本文对原有的中值滤波技术进行改进，使得其可以尽可能地保留信号中的重要信息。

该算法的基本原理是：在用邻域窗口中的中值替换信号值之前，先对信号点进行判断，以确定其是否真正需要被替换。通过比较信号点与邻域窗口内其他点的差异，或者利用信号的其他特性（如频率、幅度等）来判断该点是否为噪声点。如果判断该点为噪声点，则使用中值替换；否则，保留其原始值。

在PPG信号分析中，对于长度为 n 的信号，采用长度为 $2k+1$ (k 为整数) 的

西北工业大学本科毕业设计（论文）

滑动窗口。窗口两端补充 k 个端点值以保持信号完整性。窗口内数据排序后得中位数 $X(Z)$ 。随后计算绝对中位差（MAD），量化数据离散程度。MAD计算公式如下：

$$MAD = med(|X(i) - X(Z)|) \quad (3-1)$$

为了数据去噪，定义 $b*MAD$ ($2 < b < 3$) 为阈值，其中MAD是数据与中位数 $X(Z)$ 之差的绝对值的中位数，如图3-4所示：

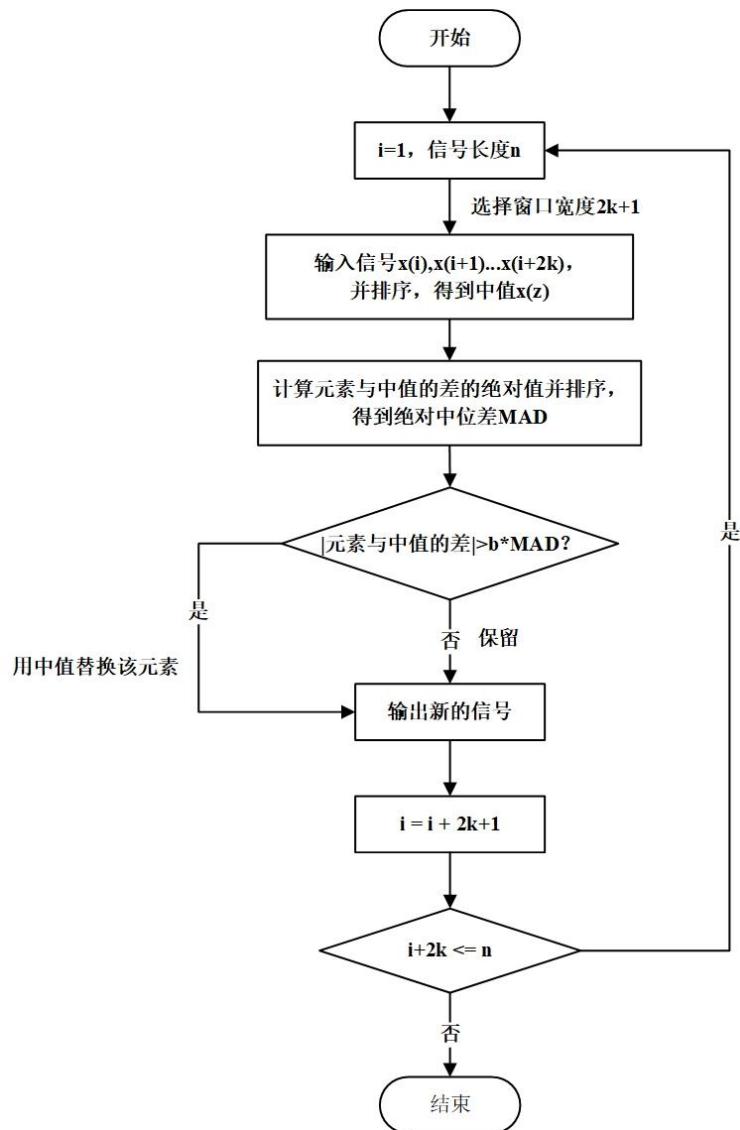


图 3-4 改进中值滤波流程图

改进后的中值滤波技术对一段原始信号进行处理，效果如图3-5所示：

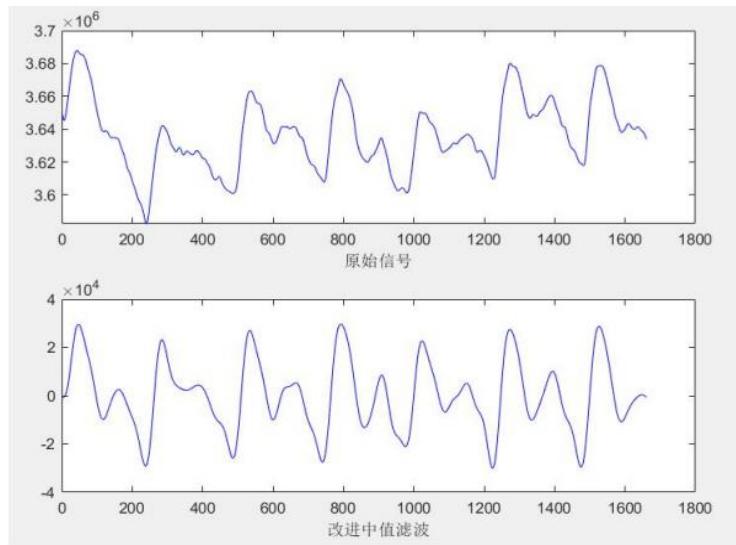


图 3-5 改进后的效果图

3.1.4 消除 PPG 信号基线漂移

PPG信号在采集的过程中，容易受到来自肢体运动，呼吸，心率甚至情绪的影响，从而导致信号出现整体漂移^[29]。

基线漂移是一种低频噪声，其频率主要集中在0.1Hz附近，这一频率范围与PPG信号的频率有所重叠，因此传统的数字滤波方法在处理时效果有限。文章选定三次样条插值法作为去除基线漂移的方法，而该方法的核心思路在于通过拟合PPG信号的基线来消除基线漂移。

设想如下：

(1)首先，将PPG信号按照其周期进行分段，利用每个周期内的波谷作为分段点，对每个周期内的信号进行三次样条插值，以拟合出该周期内的基线。

(2)最后，将原始PPG信号减去拟合出的基线信号，从而得到去除了基线漂移的纯净信号。这种方法能够精确地拟合并消除基线漂移，保留PPG信号中的有用信息。

具体步骤如下：

在PPG信号的处理中，首先按照其周期性特点将信号划分为 n 个区间，从而得到 $n+1$ 个端点： $[x_0, x_1, \dots, x_{n-1}, x_n]$ 。其中， x_0 是PPG信号的起始点，而 x_n 是信号的终点。这些端点将信号划分为 n 个区间： $[(x_0, x_1), (x_1, x_2), \dots, (x_{n-2}, x_{n-1}), (x_{n-1}, x_n)]$ ，每个区间内包含 m 个点。为了在每个小区间内构造一个平滑的曲线，采用三次样条插值法，意味着在每个小区间 $[x_{i-1}, x_i]$ 内，信号将由一个三次方程来近似。

在区间 $[x_{i-1}, x_i]$ 内，三次样条插值函数 $S(x_i)$ 可以表示为：

$$S(x_i) = a_i + b_i x_i + c_i x_i^2 + d_i x_i^3 \quad (3-2)$$

其中 a_i 、 b_i 、 c_i 、 d_i 为需要求得的三次样条插值函数系数，一共有 n 个区间。

因为三次样条插值函数 $S(x_i)$ 的端点均为区间的端点，所以：

$$S_1(x_i) = y_1 \quad (3-3)$$

$$S_m(x_i) = y_m \quad (3-4)$$

其中 $S_1(x_i)$ 和 $S_m(x_i)$ 表示三次样条插值函数 $S_1(x_i)$ 的左右两个端点， y_1 和 y_m 表示区间 $[x_{i-1}, x_i]$ 两端对应的信号值。由此共可以得到 $2n$ 个方程。

在利用三次样条插值法对PPG信号进行分区处理时，每个区间内的三次样条插值函数 $S(x_i)$ 需要确保其一阶导数和二阶导数在区间交界处保持连续。这意味着，对于任意相邻的两个区间 $[x_{i-1}, x_i]$ 和 $[x_i, x_{i+1}]$ ，第 i 个区间末端（即 x_i 点）处的一阶导数和二阶导数，应分别等于第 $i+1$ 个区间开端（即 x_{i+1} 点）处的一阶导数和二阶导数。这样的连续性要求确保了整个插值函数在信号上的平滑过渡，从而更准确地拟合原始信号，即：

$$S'_m(x_i) = S'_1(x_{i+1}) \quad (3-5)$$

$$S''_m(x_i) = S''_1(x_{i+1}) \quad (3-6)$$

其中， $S'_m(x_i)$ 和 $S''_m(x_i)$ 分别代表三次样条插值函数 $S(x_i)$ 在末端 m 的一阶和二阶导数， $S'_1(x_{i+1})$ 和 $S''_1(x_{i+1})$ 分别代表三次样条插值函数 $S(x_{i+1})$ 在起始端的一阶和二阶导数。根据这些公式，可以构建出包含 $2n-2$ 个方程的方程组。利用自然边界条件，设定端点的二阶导数为 0，来进一步确定方程组的解。这一条件的具体公式如下：

$$S''_1(x_1) = S''_m(x_n) = 0 \quad (3-7)$$

本文共构建了 $4n$ 个方程，通过解这些方程，得到了 $4n$ 个系数，这些系数用于构建完整的三次样条插值函数。该函数即为PPG信号的基线。为了消除基线漂移的影响，将原始的PPG信号减去这条基线信号，最终得到的是去除基线漂移后的PPG信号，效果如图3-6所示。

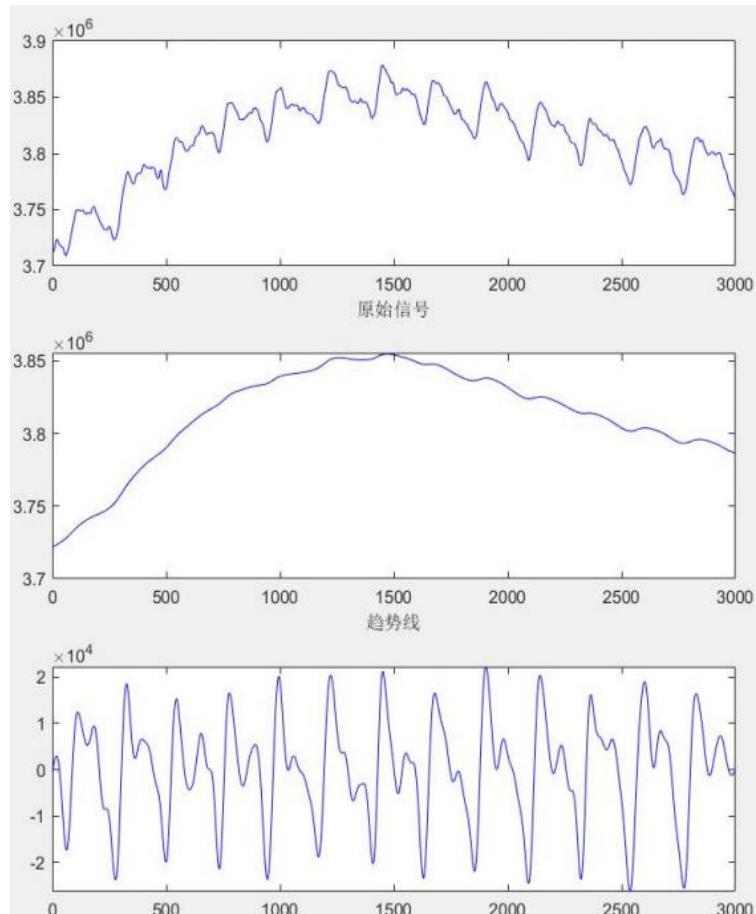


图 3-6 经过高频噪声处理的原始信号使用三次样条插值后的效果图

3.2 PPG 信号特征值提取

3.2.1 PPG 信号综合特征参数 R 的计算方法

由第二章的Lambert-Beer定理推导可知，人体血糖浓度 C_I 和血糖特征参数 R 之间的关系如下：

$$C_I = \frac{1}{(\varepsilon_1^{\lambda_1} - \varepsilon_1^{\lambda_2})\Delta L} R \quad (3-8)$$

C_I 代表血糖的初始数值，而 R 则代表红光AC分量与DC分量之比与红外光交流分量与直流分量之比的差值。

显然，血糖初始浓度 C_I 与特征参数 R 之间存在着线性关联。关于 R 值的计算，

其公式如下:

$$R = I_{AC}^{\lambda_1} / I_{AD}^{\lambda_1} - I_{AC}^{\lambda_2} / I_{AD}^{\lambda_2} \quad (3-9)$$

在上述公式中, $I_{AC}^{\lambda_1}$ 和 $I_{AC}^{\lambda_2}$ 代表红光与红外光PPG信号中由人体动脉血液吸收的部分, 该部分通过PPG信号的波峰值 Q_{max} 与波谷值 Q_{min} 的差值来表示。而 $I_{AD}^{\lambda_1}$ 和 $I_{AD}^{\lambda_2}$ 则代表被人体肌肉、静脉血液等吸收的PPG信号部分, 该部分表现为稳定的直流分量, 可通过PPG信号的平均值 Q_m 来反映。

为了计算 R 值, 需参考图3-7, 并从PPG信号中提取出主波峰值 Q_{max} 、波谷值 Q_{min} , 同时计算出信号的平均值 Q_m 。

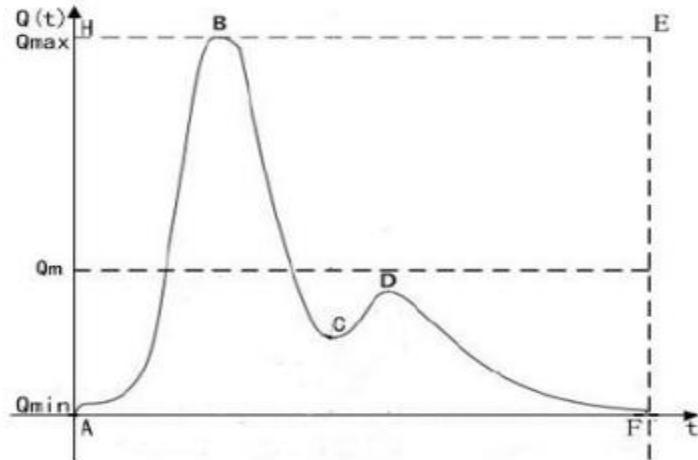


图 3-7 PPG 信号示意图

信号平均值 Q_m 计算公式如下:

$$Q_m = \frac{1}{F - A} \int_A^F Q(t) dt \quad (3-10)$$

由于PPG信号的上升和下降沿均呈现出陡峭的特性, 因此采用差分法能够便捷地定位信号的转折点。基于此, 本文运用差分阈值法成功提取了PPG信号中的主波峰值 Q_{max} 与波谷值 Q_{min} , 具体检测效果如图3-8所示:

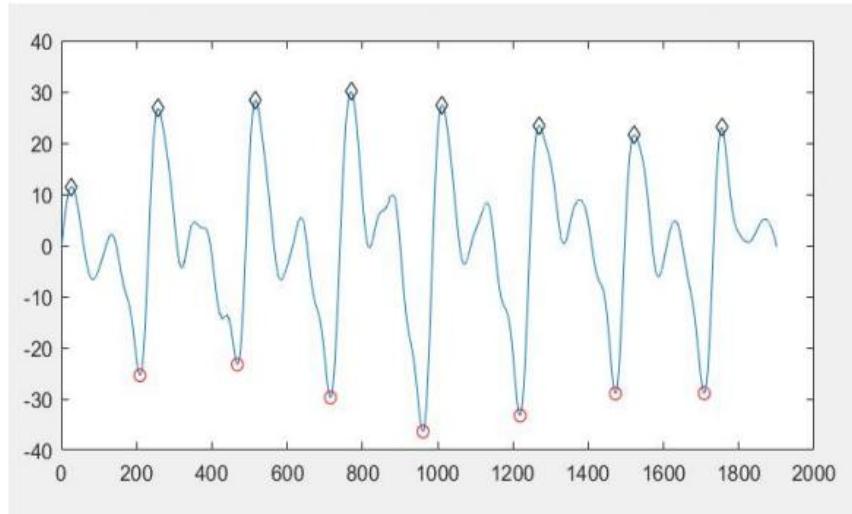


图 3-8 PPG 信号的主波峰和波谷检测示意图

提取PPG信号的主波峰和波谷值后，综合特征参数R可表示为如下：

$$R = \frac{Q_{max}^{\lambda_1} - Q_{min}^{\lambda_1}}{Q_m^{\lambda_1} - Q_{min}^{\lambda_1}} - \frac{Q_{max}^{\lambda_2} - Q_{min}^{\lambda_2}}{Q_m^{\lambda_2} - Q_{min}^{\lambda_2}} \quad (3-11)$$

其中， $Q_{max}^{\lambda_1}$ ， $Q_{min}^{\lambda_1}$ 和 $Q_m^{\lambda_1}$ 表示红光PPG信号的主波峰，波谷和信号平均值。

$Q_{max}^{\lambda_2}$ ， $Q_{min}^{\lambda_2}$ 和 $Q_m^{\lambda_2}$ 表示红外光PPG信号的主波峰，波谷和信号平均值。

3.2.2 初步建立血糖检测模型

上文的推导已经可以明确血糖值和特征参数之间的关系，为了更精确地描述这种关系，本文采用最小二乘法对综合特征参数R与血糖值进行拟合，得到了关于综合特征参数R与血糖值关系的拟合公式，具体如下：

$$C_1 = a_1 * R + a_2 \quad (3-12)$$

$$C_1 = b_1 * R^2 + b_2 * R + b_3 \quad (3-13)$$

$$C_1 = d_1 * R^3 + d_2 * R^2 + d_3 * R + d_4 \quad (3-14)$$

其中， C_1 代表血糖的初始值，而 a_1 至 a_2 、 b_1 至 b_3 以及 d_1 至 d_4 则是通过多组综合特征参数R与对应的血糖值进行标定后得到的拟合系数。为了确定最佳的初步血糖模型，需要计算平均绝对误差来进行评估，MAE的具体计算公式如下：

$$MAE = \frac{1}{n} \sum_{i=1}^n |C - C_i| \quad (3-15)$$

其中， n 表示样本数， C 表示血糖标准值。

分别通过单人和多人血糖检测实验，验证上述拟合公式的 MAE ，实验对比结果如表3-3所示：

表 3-3 初步拟合结果

MAE	一次拟合(mmol/L)	二次拟合(mmol/L)	三次拟合(mmol/L)
单人	0.72	0.86	0.93
多人	1.58	1.88	2.30

根据上表数据显示，一次拟合的误差为最小，因此本文选择采用一次拟合作为拟合公式，也表明了特征参数 R 与血糖值 C 之间的数学关系是一次线性关系，验证了第二章的对于朗伯比尔定律的延伸推导是正确的。

3.3 特征参数提取与标准化

综合考量上述多个特征参数，包括综合特征参数R、血糖初值 C^1 、心率E、年龄G以及体质指数M，可以发现这些参数各自具有不同的量级，若直接对这些参数进行比较或加权处理，很可能会导致某些参数的权值过大或过小，从而影响模型的训练，进一步影响结果的准确性和可靠性。

因此，在进行参数运算之前，必须对数据进行标准化处理，以确保所有参数在相同的尺度上进行比较和运算。常见的标准化方法有两种，其中一种是z-score标准化，另一种是min-max标准化。这两种方法都能有效地将数据转换到统一的尺度上，便于后续的数据分析和模型建立。

z-score标准化，或称标准差标准化，是数据处理中的一种常用手段。该方法能够将数据转化为标准正态分布，转化后的数据标准差为1，均值为0。在进行z-score标准化之前，需要先行计算出数据的平均值与标准差。

计算方法如下：

$$Y = \frac{X - \mu}{\sigma} \quad (3-16)$$

其中， Y 表示标准化后的数据， X 表示原数据， μ 表示原数据的平均值， σ 表示原数据的方差。

min – max 标准化，利用数据的最大值与最小值将数据缩放到区间[0,1]上。

标准化前需要求得数据的最大值和最小值。

$$Y = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3-17)$$

其中， Y 表示标准化后的数据， X 表示原数据， X_{min} 表示原数据中的最小值， X_{max} 表示原数据中的最大值。

上述两种标准化方法均不要求数据全为正数，且操作简便。但经过实际数据分析，可以看出 $min-max$ 标准化处理会导致结果偏差。与此相比，z-score标准化能将原始数据转换成无单位的标准化数据，提高数据的可靠性和准确性。因此，在全面考虑后，本文使用z-score标准化处理后的数据作为后续无创血糖检测模型的输入。

3.4 本章小结

本章首先说明了PPG信号的数据集来源，以及介绍了数据集获取的步骤，说明了数据集的可靠性和可信度。紧接着，对相应的PPG原始信号几种预处理方法进行了介绍，根据其优缺点选取了中值滤波和三次样条插值法，并对样本进行试验并验证其正确性。接下来，计算得出了综合特征参数 R ，并以此为基础建立了综合特征参数 R 与血糖之间的初步数学模型，并通过一次拟合验证了其线性关系的正确。最后，对相关的特征参数进行了标准化处理。

第4章第四章 建立无创血糖检测模型

在得到上述的特征参数后，需要结合对应的血糖标准值进行分析，寻找特征参数与血糖值之间的关系，从而建立血糖检测模型。

4.1 相关技术简介

4.1.1 PYTHON 简介

Python是一种高级解释型编程语言，以“优美、简洁、易懂”著称。其设计理念强调简洁易懂，应用领域非常广泛，包括网络爬虫、嵌入式程序开发、大数据应用和云计算等领域。Python语法简洁明了，代码量相较其他语言更少，且拥有丰富的库和工具，能迅速推进项目开发。此外，Python还具备灵活性、可移植性、跨平台性和易维护性等诸多优点。Python的庞大社区为开发者提供了丰富的解决方案，对初学者尤为友好。同时，Python也是谷歌、Dropbox、Uber和Instagram等大公司首选的开发语言之一。

4.1.2 NumPy 简介

NumPy是Python的一个重要扩展库，它主要支持对大量多维数组和矩阵的数据进行操作，是Python数据科学领域中的核心基础库。由于众多数据科学库都建立在NumPy的基础之上，因此掌握NumPy有助于深入理解Python的数据科学生态系统。

NumPy的优势包括：

(1) 强大的数组操作功能：NumPy的显著优点在于其出色的数组操作能力，能处理存储相同类型多维数据的数组，并执行多样的数学运算、数值分析及线性代数等任务。

(2) 高性能表现：得益于大量使用C语言编写的底层代码，NumPy在数值计算和数据操作方面表现出色，运行速度快。

(3) 丰富的算法库：NumPy提供了众多的数值计算函数、统计函数以及金融和统计类方法，能满足数据科学家的多样化需求。

(4) 跨平台兼容性：NumPy可在不同操作系统和平台上顺畅运行，并能与其他Python库无缝集成。

(5) 开源特性：作为开源项目，NumPy拥有庞大的社区支持和开源开发实践，为用户提供了广阔的资源与帮助。

4.1.3 Pandas 简介

Pandas是一个开源的数据处理和分析库，以NumPy为基础，并由量化金融研

西北工业大学本科毕业设计（论文）

究员Wes McKinney创立。它提供了数据框架（DataFrame）和一套处理工具，使得数据清洗、分组、聚合及分析等任务得以迅速完成。

Pandas的主要功能包括：

- (1) 数据清洗：Pandas可以轻松应对缺失、重复或异常数据，提供数据转换、整理及空值填充等功能，有效防止数据错误和损失。
- (2) 数据分组：利用Pandas，用户可以根据一个或多个条件灵活地将数据分组，便于进行多维度分析、聚合运算等。
- (3) 数据聚合：通过groupby操作，Pandas能迅速对数据进行聚合，支持使用mean()、count()、sum()等统计函数，也允许自定义聚合函数。
- (4) 时间序列处理：Pandas内置强大的时间序列工具，能高效解析和转换日期数据，并按日期进行时间序列分析、汇总等。
- (5) 数据可视化：Pandas结合matplotlib等可视化库，可便捷地创建图表，助力用户更直观地理解数据。
- (6) Pandas的核心数据结构包括Series和DataFrame。Series是一维标记数组，而DataFrame则是二维表格式结构，每列可具有不同的数据类型。Pandas支持多种基本操作，如选择、排序、过滤及数据清洗等。

4.1.4 Keras 简介

Keras是一个以Python编写的高级神经网络API，它提供了一个简洁且统一的接口，用于构建和训练包括CNN、RNN和常见DNN在内的深度学习模型。尽管许多人将Keras视为一个深度学习框架，但它实际上能够在TensorFlow、Theano和CNTK等多个后端上运行。

利用Keras API，用户可以方便地构建不同类型的神经网络模型。Keras使得编写代码更为简洁，从而显著提升了模型开发和迭代的效率。此外，Keras支持自动求导功能，用户仅需定义模型结构和损失函数，而反向传播算法和优化器的生成则由Keras自动完成。同时，Keras还提供了丰富的常用层和模型，以及详尽的文档和教程，为用户提供了全面的学习资源和工具支持。

总体而言，Keras是一个易于学习和使用的高级神经网络API，其简洁一致的接口为深度学习模型的构建、训练和调试提供了极大的便利。由于能够在多个后端运行，且代码简洁易懂，Keras已成为机器学习和研究领域中开发者的实用高效工具。

4.1.5 Sklearn 简介

Sklearn是一个开源的Python机器学习库。它涵盖了数据挖掘和数据分析的多种工具和算法，如分类、回归、聚类、降维等，并以NumPy、SciPy和matplotlib等库为基础，既简单又高效。

Sklearn的主要功能有数据预处理、特征操作、模型选取与评估。在数据预处

西北工业大学本科毕业设计（论文）

理方面，它能帮助清洗数据、处理缺失值、归一化特征等。对于特征操作，Sklearn可以进行特征选择和转换，以便提取和构建有价值的特征。在模型选取与评估上，它实现了多种经典的机器学习算法，如决策树、支持向量机等，并提供了交叉验证和网格搜索等方法来选定和评估最佳模型。

此外，Sklearn配备了详尽的文档和示例，使用户能迅速掌握各种机器学习算法和工具。同时，它与Pandas、TensorFlow等其他库和工具能够无缝集成，为用户在数据处理和模型开发上提供了更大的灵活性。

4.1.6 LSTM 简介

LSTM是一种特殊的递归神经网络，可以利用时间序列对输入进行分析其单元结构包括输入门、遗忘门、输出门及一个能长期存储信息的细胞状态。这些门分别负责新信息的输入、细胞状态的更新及信息的输出。这种机制使LSTM在处理序列数据时表现出色，广泛应用于语言建模、时间序列分析、机器翻译等领域。其出色的记忆与长期依赖建模能力，使LSTM成为深度学习中的关键工具。

4.2 系统分析

4.2.1 系统功能需求

系统主要是通过系统的开发可以实现通过创建和训练LSTM模型，用训练好的模型进行血糖检测。系统用例需求如图 4-1 所示：

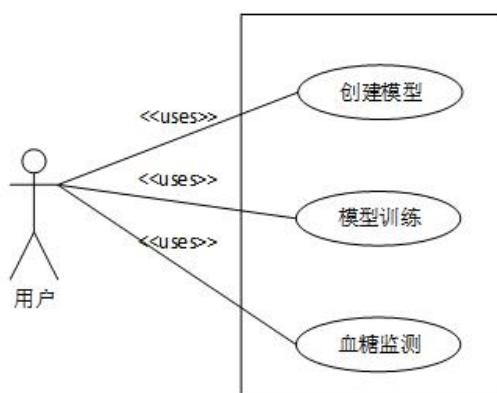


图 4-1 系统用例图

4.2.2 技术可行性

(1) 硬件可行性分析

本系统对硬件无特殊要求，普通配置即可满足需求，确保系统高效稳定运行。开发平台选用酷睿i5或以上配置的PC，其性能和速度均达标。因此，从硬件角度看，系统开发完全可行。

(2) 软件可行性分析

本血糖检测系统基于PYTHON技术，结合机器学习算法进行开发。通过构建和训练LSTM模型来实现血糖检测功能。PYTHON的稳定性、可扩展性、开发效率以及便捷的管理特性，极大地简化了系统开发流程。综上所述，该系统的软件开发同样具有可行性。

4.2.3 操作可行性

血糖检测系统的用户界面设计简洁直观，便于用户轻松上手。即便是日常使用电脑的基础用户，也能无障碍地进行操作访问。本系统采用PYTHON语言进行开发，借助其强大的开发环境，使系统更加完备。该系统不仅操作简单，管理方便，还具备良好的交互性。综上所述，本系统的开发是切实可行的。。

4.3 LSTM 模型

4.3.1 LSTM 模型构建流程

构建LSTM模型进行训练和预测的具体步骤如图4-2所示。

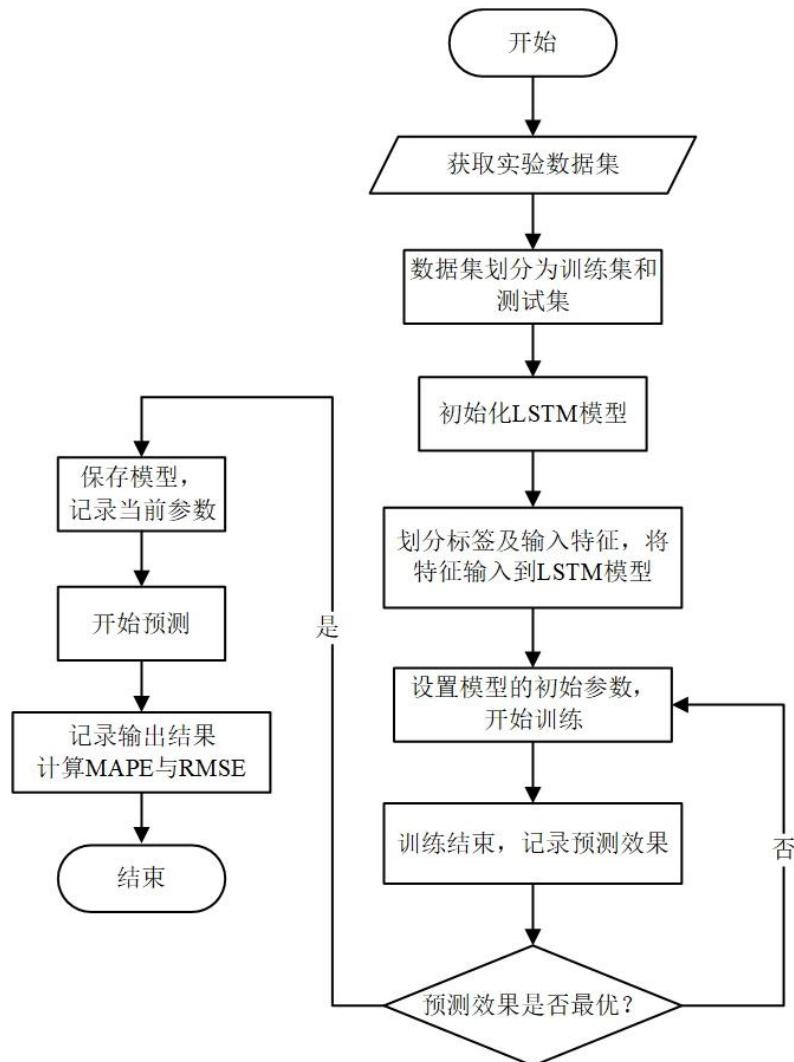


图 4-2 LSTM 模型流程图

详细操作步骤：

- (1) 采集实验数据，并将其合理地分割成训练数据集和测试数据集；
- (2) 构建长短期记忆网络（LSTM）模型，并将提取的特征数据输入到该模型中；
- (3) 配置模型训练的迭代次数、学习率等关键参数，并启动训练过程；
- (4) 训练完成后，对模型的预测效果进行评估，并根据评估结果对模型参数进行相应调整；
- (5) 反复执行步骤(3)和(4)，直至模型的预测效果达到最优。一旦达到最优效果，即保存当前的最优模型；
- (6) 将特征数据输入到已保存的最优模型中，记录并分析模型的预测结果。此外，还需利用平均绝对百分比误差（MAPE）和均方根误差（RMSE）对预测结果进行评估。

4.3.2 LSTM 模型参数设置

为提升LSTM模型在处理PPG信号时的准确率，关键在于选取恰当的激活函数。常用的激活函数包括Sigmoid函数、Tanh函数以及ReLU函数^[23]。

Sigmoid函数是一个在生物学中常见的S型函数，也称为S型生长曲线。在深度学习中，由于其单增以及反函数单增等性质，Sigmoid函数常被用作神经网络的激活函数，将变量映射到[0,1]之间，其表达式如下：

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (4-1)$$

Tanh函数，又叫做双曲正切函数，是双曲函数的一种。双曲正切函数在数学语言上一般写作Tanh。它解决了Sigmoid函数的不以0为中心输出问题，然而，梯度消失的问题和幂运算的问题仍然存在：

$$f(x) = \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4-2)$$

ReLU函数，全称为：Rectified Linear Unit，是一种人工神经网络中常用的激活函数，通常意义上，其指代数学中的斜坡函数，即：

$$f(x) = \max(0, x) \quad (4-3)$$

相较于Sigmoid和Tanh函数，ReLU函数更能有效解决梯度消失问题，并且能显著提升训练速度，因此被选为本次模型的激活函数。

在本课题中，选用Adam作为优化策略来训练LSTM模型。相较于Adagrad和RMSprop，Adam在运算速度、资源占用、处理大型数据集及参数调整，以及应对高噪声和梯度稀疏性挑战上，均显示出其卓越性能，非常适合大规模数据的训

练任务。

在深度神经网络的搭建与训练过程中，不仅要慎重选取适合的激活函数和优化策略，还需对网络深度、训练轮次、隐藏节点数量等核心配置进行精确调整。这些要素对训练成效具有重大影响。典型的网络框架包含输入层、隐含层和输出层，其中隐含层负责对数据进行线性分割。虽然增加隐含层数可以提升模型的复杂度和理论上的精度，但层数过多也可能引发模型参数冗余。

本文模型主要围绕全连接层和LSTM层进行设计。全连接层在网络中扮演着对提取的特征进行非线性转换的角色，深入挖掘特征之间的内在关系，并最终将这些特征投射到输出空间。在预设的1000轮训练和每批次100个数据样本的条件下，使用L代表LSTM层，F代表全连接层，以此研究不同的隐含层结构对模型预测准确度的影响。

4.4 系统模型设计

4.4.1 实验步骤

在本实验中，采用的是Python与Keras的结合来进行编程。Keras，作为一款以Python为基础编写的高级神经网络API，为深度学习模型的构建与训练提供了简洁统一的接口。该技术框架运用了先进的图计算方式，即通过有向图的节点与边来明确描述数学计算过程。本实验在Windows操作系统环境下进行，实验环境的详细配置参见表4-1：

表 4-1 实验环境

CPU	英特尔酷睿11代i5-1135G7
GPU	GTX2060 Ti@2GB
内存	16GB
显卡	Arc显卡
操作系统	Windows 11
深度学习框架	Keras
编程语言	Python

详细步骤如下：

- (1) 实验目的：进行参照性实验；
- (2) 实验设置：设定了试验设定值，作为后续调整的基准；
- (3) 技术挑战：
 - a) 权重缩减通过增加损失函数的规整化项实现；
 - b) 非确定性训练数据集可能导致规整化项的大幅随机变化；
 - c) 规整化项的大幅变化会影响梯度调整指向的稳定性，增加训练误差。

表 4-2 基本实验参数

参数	取值
初始学习速率 (n_0)	0.1
学习速率衰减率 (d)	0.1
衰减间隔 (NUM_EPOCHS_PER_DECAY)	350
动量系数 (u)	0.9
Dropout 比例 (1-p)	0.5
权值衰减项权重 (k)	[0,0.001]
最大迭代步数 (max steps)	1000

4.4.2 LSTM 模型设计

LSTM在序列建模方面具有显著优势，其能够捕获长期上下文信息，并拥有神经网络拟合非线性的特性，这些是条件随机场(CRF)模型所无法比拟的。LSTM的结构简图如图4-3所示：

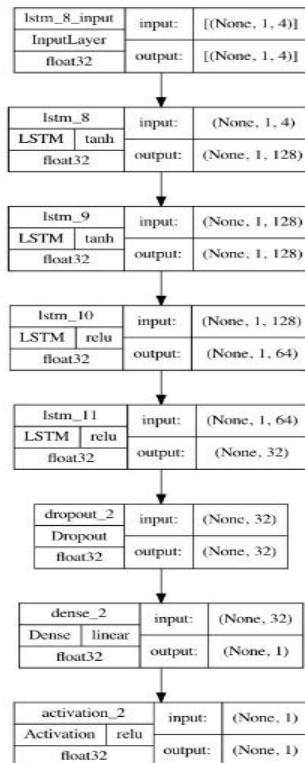


图 4-3 LSTM 结构图

4.5 本章小结

本章首先阐述了建立模型所需要的各项技术，包括Python、NumPy、Pandas、Keras、Sklear和LSTM等，接着对开发模型系统的功能需求和可行性进行了分析，然后构建LSTM模型和各种参数配置，并介绍了建立模型所需硬件软件的配置。

第 5 章第五章 程序运行结果及分析

5.1 建模所调用的库

- (1) pandas: 用于数据处理和分析, 这里主要用于将 NumPy 数组转换为 DataFrame。
- (2) numpy: 用于高效的数值计算, 处理数组和矩阵。
- (3) seaborn: 用于数据可视化, 特别是热力图。
- (4) sklearn.preprocessing: 用于数据的归一化处理, 将数据缩放到[0, 1]的范围内。
- (5) keras.layers: 用于构建和定义神经网络模型。
- (6) heartpy: 用于处理和分析心电图数据的库。

5.2 程序设计

5.2.1 输入原始信号并进行预处理

输入绿色通道和红色通道提取出的PPG信号, 对信号进行预处理, 并利用heartpy 库对PPG信号进行加载、可视化、处理和特征提取。

(1) 其中, 使用matplotlib库来绘制信号的波形图。这有助于直观地查看信号的原始形态, 从而对其质量进行初步评估, 并识别任何异常或噪声。对于每一组数据, 使用matplotlib库, 通过plt创建一个大小为12*4的图形窗口, 并绘制出加载的信号数据, 部分代码复现如下:

```
data, timer = hp.load_exampledatal(0)      #加载第一组示例数据
plt.figure(figsize=(12, 4))      #窗口大小为12*4
plt.plot(data)
```

(2) 数据的采样率被设定为固定的120.0Hz, 部分代码复现如下:

```
wd, m = hp.process(data, sample_rate=120.0)
plt.figure(figsize=(12, 4))      #创建一个新的图形窗口
hp.plotter(wd, m)      #绘制处理后的信号及其分析结果
for measure in m.keys():
    print("%s: %f" % (measure, m[measure]))      #遍历度量字典
```

(3) 中值滤波

```
for h in range(height):
    row = []
```

```
for w in range(width):
    value=im.getpixel((h,w))      #用getpixel读取这一点像素值
    row.append(value)            #把它加到这一行的list中去
data.append(row)          #把记录好的每一行加到data的子list中去
data=signal.medfilt(data,kernel_size=3)
data=np.int32(data) #转换为int类型
```

(4) 三次样条插值

```
ipo3=spi.splrep(x,y,k=3)      #生成模型参数
iy3=spi.splev(ix3,ipo3)        #生成插值点
plt.plot(x,y)
plt.plot(ix3,iy3)
plt.legend(['origin','interp'],loc='upper left') plt.show()
```

运行结果如图5-1、5-2所示：

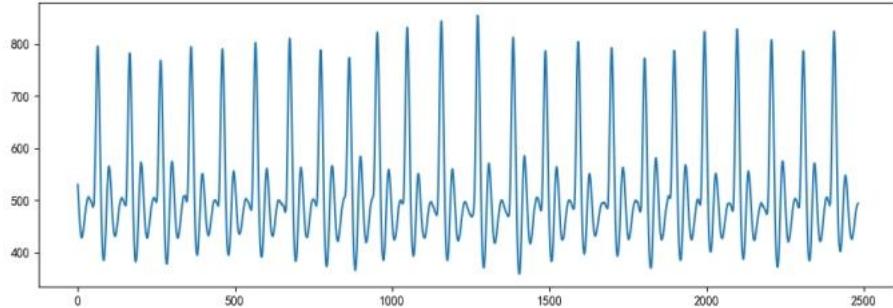


图 5-1 信号 1

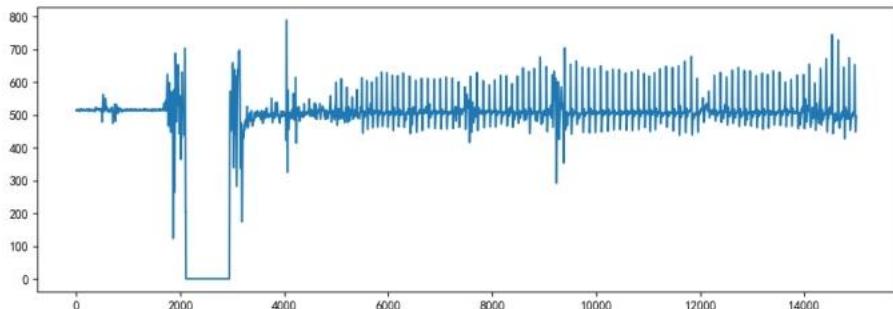


图 5-2 信号 2

5.2.2 计算特征系数

确定两个信号数据的峰值，并以此为根据，通过差分阈值法计算特征系数。

部分代码如下：

- (1) 确定峰值，使用scipy.signal.find_peaks (峰值检测)

```
peaks, _ = find_peaks(x, distance=150)
plt.plot(x)
plt.plot(peaks, x[peaks], "x")
```

(2) 计算特征系数:

```
feature=pd.DataFrame(x_data)
#使用 pandas 的corr()方法来计算特征之间的相关系数。
corr = feature.corr()      #计算各列之间的相关系数。
```

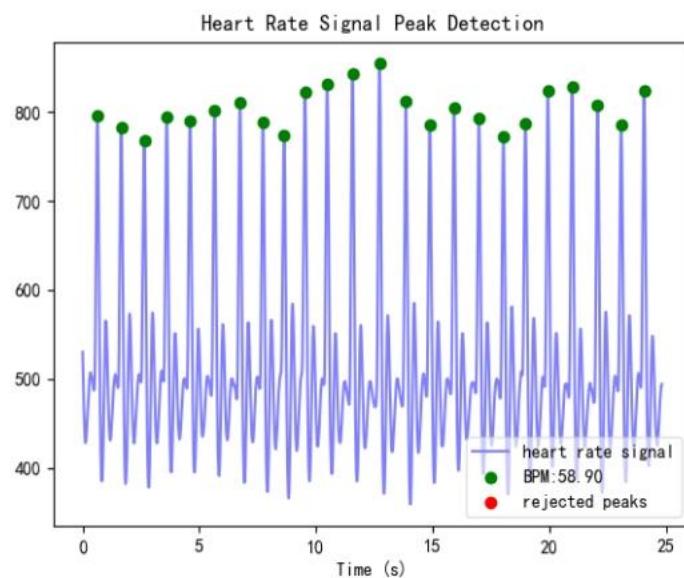


图 5-3 信号峰值确定 (1)

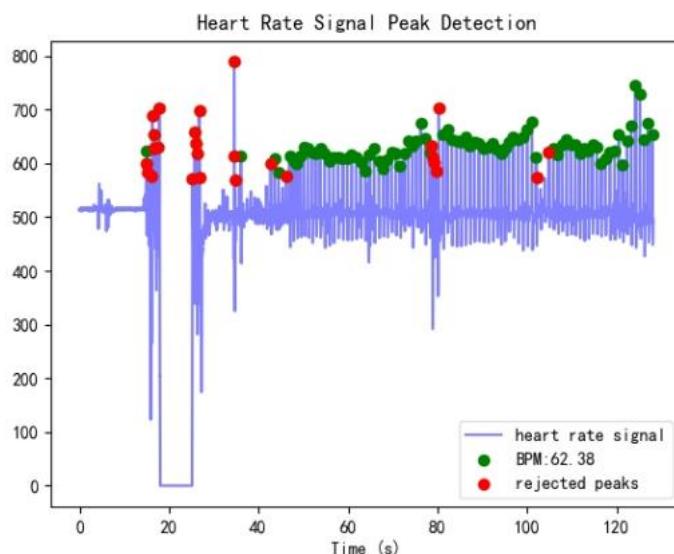


图 5-4 信号峰值确定 (2)

5.2.3 特征矩阵热力图可视化

热力图利用颜色分布直观展示人群分布、密度及变化趋势等复杂数据。它通过为每个离散点建立缓冲区，并使用渐变的灰度填充，灰度值可叠加，交叉区域叠加后颜色更亮，表示“热度”更高。最后，根据灰度值从色带中映射颜色并重新着色生成热力图，通常从冷到暖色表示点密度状态。此过程需启用地图的Alpha通道。

通过上述计算，展示相关系数的热力图。热力图中的颜色深浅表示了特征之间相关性的强弱，越接近1说明越相关，如图5-5所示。

部分代码如下：

```
plt.figure(figsize=(10,6))
ax = sns.heatmap(corr, xticklabels=corr.columns, yticklabels=corr.columns,
linewidths=0.2, cmap="RdYlGn", annot=True)
plt.title("变量间相关系数")
```



图 5-5 不同特征之间的关系图

5.2.4 数据归一化处理

为了消除数据之间量纲带来的影响，程序需要对数据统一进行归一化处理，使得不同指标之间具有可比性。

例如，分析受试者的身高和体重对血糖的影响，如果使用米 (m)和千克 (kg)

作为单位，那么身高特征会在 $1.6 \sim 1.8m$ 的数值范围内，体重特征会在 $50 \sim 100kg$ 的范围内。分析出来的结果显然会倾向于数值差别比较大的体重特征。

想要得到更为准确的结果，就需要进行特征归一化 (Normalization) 处理，使各指标处于同一数值量级，以便进行分析，如图5-6、5-7所示。

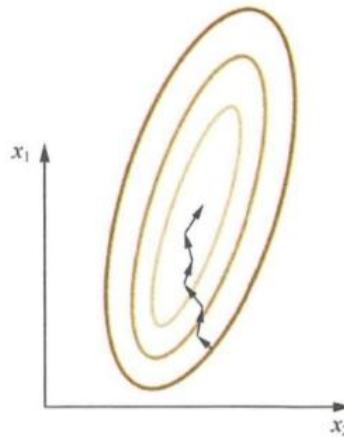


图 5-6 未归一化的数据的梯度下降过程

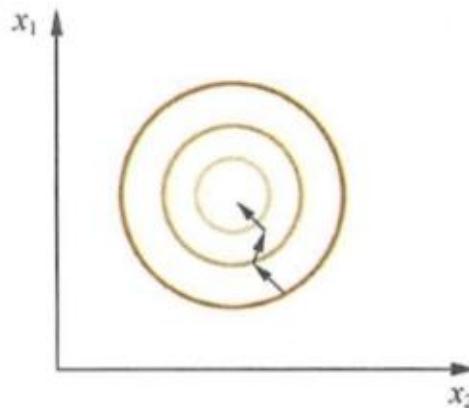


图 5-7 归一化的数据的梯度下降过程

部分代码如下：

```
def feature_normalize(X):
    mean = np.mean(X, axis=0)      #计算每个特征（即每列）的平均值。
    std = np.std(X, axis=0)        #计算每个特征的标准差。
    return (X - mean) / std       #对数据进行标准化处理。
```

5.2.5 构建 LSTM 模型

使用Keras构建LSTM模型，模型包含一个LSTM层 (32个神经元)，一个Dropout

层（防止过拟合），和一个全连接层（输出一个神经元）。模型使用均方误差作为损失函数，并使用Adam优化器进行优化。

部分代码复现如下：

```
model = Sequential()
model.add(LSTM(32, input_dim=1))
model.add(Dropout(0.1))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam')
```

5.2.6 训练损失图

训练损失图 (Training Loss Plot, TLP) 是一种在机器学习和深度学习过程中用来监控模型训练进度的可视化工具。通过观察损失值的变化，可以评估模型的拟合效果，调整超参数，以及确定合适的训练停止条件。

部分代码如下：

```
#绘制损失曲线
loss = history.history['loss']
plt.figure()
plt.plot(loss)
plt.legend(['loss'])
plt.xlabel('epoch')
```

结果如图5-8所示。

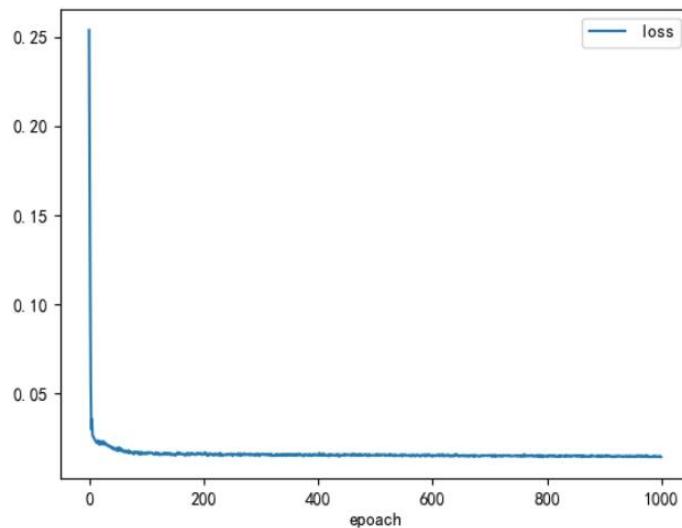


图 5-8 训练的 loss 图 (训练集的误差减小收敛)

5.2.7 划分数据

对数据集数据进行数据分割，划分训练集和测试集，比例为80%和20%。使用

训练数据对模型进行训练，总共训练1000个周期，每个批次包含100个样本。

部分代码如下：

```
history=model.fit(x_train,y_train,epochs=1000,batch_size=100,verbose=2)
```

5.3 训练结果及分析

结果如下所示：

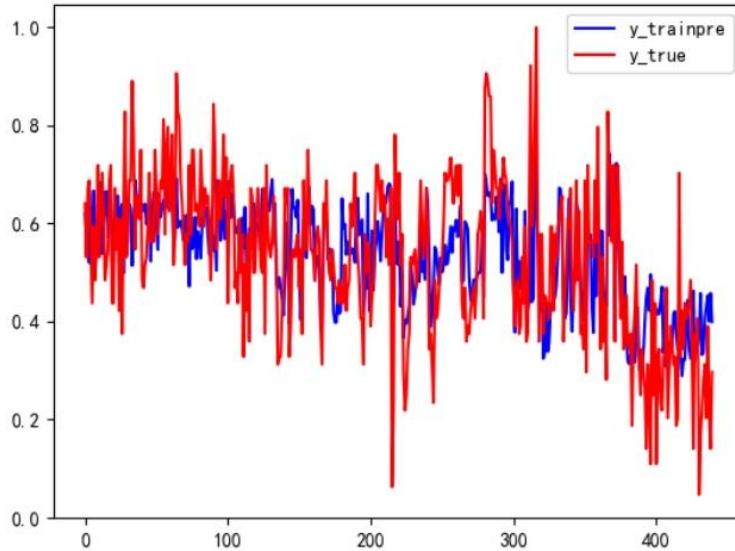


图 5-9 模型在训练集上的预测和实际对比图

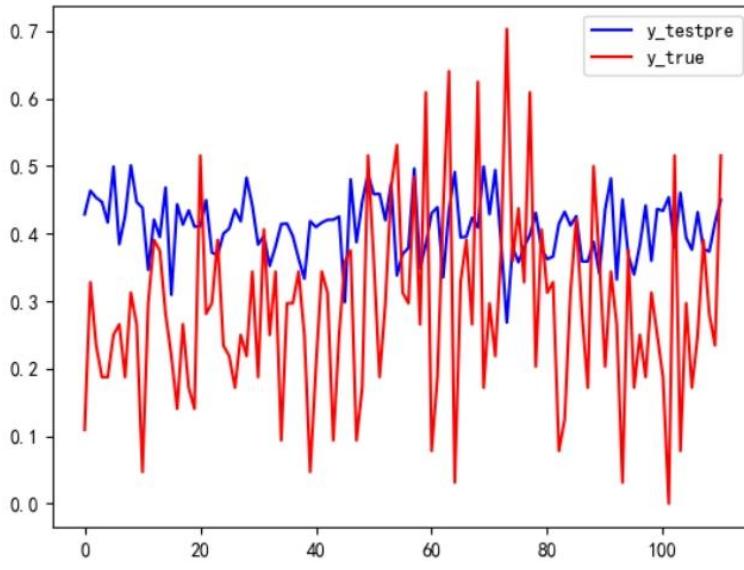


图 5-10 模型在测试集上的预测和实际对比图

最终经过计算，该模型的平均绝对百分比误差 (MAPE) 达到30%，模型对于血糖浓度以及趋势的预测识别率达到65%以上，图5-9是模型在训练集上的预测和实

际对比效果图，图5-10是模型在测试集上的预测和实际对比效果图。从图中可以看出，相比传统的RNN，LSTM模型的计算复杂度高，需要大量数据进行训练，本文训练出的模型精度基本达到预期效果，试验结果也证实基于光电容积脉搏波描记法进行无创血糖检测的可行性。

5.4 本章小结

本章详细介绍了建模所用的库及主要过程的源代码，经过多次迭代和训练，最终得到了模型在训练集和测试集上的对比效果图，结果基本符合预期，由于本文样本数量的原因且训练模型单一，在后期研究中仍需进一步优化。

参考文献

- [1] 梁晓峰. 我国慢性病防控工作及展望[J]. 中国慢性病预防与控制,2022,30(06):408-409.
- [2] Tönnies Thaddäus, Quantifying the underestimation of projected global diabetes prevalence by the International Diabetes Federation (IDF) Diabetes Atlas.[J]. BMJ open diabetes research & care,2021,9(1).
- [3] MD Jenks. Causes Of Diabetes[J]. Electronics & communication engineering journal,2008,3(10):28.
- [4] 顾楠,郭晓蕙.1型糖尿病规范化管理的难点与展望[J].中华医学杂志,2022,102(16):1148-1152.
- [5] 刘志栋,周时高,张柯培,张笑,张意,陈俊岚,许洁睿,王森. Analysis and Research Progress of Type 2 Diabetes Combined with Obesity in Traditional Chinese Medicine[J]. 中医学,2021,10(04).
- [6] 刘月姣. 《中国居民营养与慢性病状况报告(2020 年)》发布[J]. 农产品市场,2021(02):58-59.
- [7] 韦哲,张秉玺,石恒兵,赵刚,王能才.无创血糖检测技术的发展[J].中国医学装备,2020,17(12):196-199.
- [8] 刘志栋,周时高,张柯培,张笑,张意,陈俊岚,许洁睿,王森. Analysis and Research Progress of Type 2 Diabetes Combined with Obesity in Traditional Chinese Medicine[J]. 中医学,2021,10(04).
- [9] 田瀚. 基于近红外光谱法分析人体血液成分的技术研究[D]. 中国矿业大学,2020.DOI:10.27623/d.cnki.gzkyu.2020.001737.
- [10] PEDERSEN, AMANDA. GlucoTrack offers painless glucose-monitoring option. (cover story)[J]. Medical Device Daily, 2013.
- [11] 吕晓凤, 张婷琳, 肖锋, 等. 基于近红外反射光谱的无损血糖分析[J]. 光谱学与光谱分析, 2016, 36(07):2312-2317.
- [12] 杜玉宝. 便携式无创血糖检测仪的实现[D]. 重庆大学.
- [13] 陈剑虹, 何菲, 雷苏力, 刘泽晨. 一种基于频谱分析的无创血糖检测方法[P]. 陕西省: CN110338813A,2019-10-18.
- [14] Dantu V, Vempati J, Srivilliputhur S. Non-invasive blood glucose monitor based on spectroscopy using a smartphone. Annual International Conference of the IEEE Engineering in Medicine and Biology Society[C]. Chicago: IEEE, 2014.

西北工业大学本科毕业设计（论文）

-
- [15] Wrobel M S. Non-invasive blood glucose monitoring with Raman spectroscopy: prospects for device miniaturization[J]. IOP Conference Series: Materials Science and Engineering, 2016.
 - [16] 官智样. 即时检验血糖仪检测指尖血及静脉全血血糖的结果对比[J]. 中国医疗器械信息, 2019, 25(22):23-24.
 - [17] 杨星, 季忠, 杨力. 基于近红外光谱法的无创血糖检测原理与研究现状[J]. 生物医学工程学杂志, 2013, 30(01):204-207.
 - [18] Siddiqui S A, Zhang Y, Lloret J, et al. Painfree blood glucose monitoring using wearable sensors: Recent advancements and future prospects[J]. IEEE Reviews in Biomedical Engineering, 2018, 11:21-35, 2018.
 - [19] 马爽. 基于 PPG 的无创血糖检测系统模型与算法研究[D]. 中国科学院研究生院(沈阳计算技术研究所), 2016.
 - [20] 李刚, 周梅, 吴红杰. 无创人体血糖检测光学方法的研究现状与发展[J]. 光谱学与光谱分析, 2013, 30(10): 2744-2747.
 - [21] 付冠军. 近红外无创血糖检测技术研究[D]. 西北工业大学, 2005.
 - [22] 董健. 无创血糖检测及其实验研究[D]. 南昌: 江南科技师范大学, 2014.
 - [23] 李硕, 张松, 杨琳. 脉搏波信号采集质量实时分析研究及设备开发[D]. 北京工业大学, 2014.
 - [24] 范强. 心血管生理参数非接触式检测关键技术研究[D]. 武汉大学, 2017.
 - [25] 周树然. 基于弹性腔与弹性管模型的上肢血管系统建模分析[D]. 东北大学, 2016.
 - [26] 代娟. 近红外光谱无创血糖检测模型研究[D]. 重庆大学, 2019.
 - [27] 冯培华, 凌永权. 基于信号处理技术的无创血糖估计研究[D]. 广东工业大学, 2019.
 - [28] 刘丽梅, 孙玉荣, 李莉. 中值滤波技术发展研究[J]. 云南师范大学学报(自然科学版), 2004(01):23-27.
 - [29] 孙斌, 王成超, 陈建飞, 张永芳, 陈小惠. 受运动伪影干扰 ppg 序列的优质信号提取算法[J]. 仪器仪表学报, 2018, 39(09):171-178.
 - [30] 刘雪玉. 人体皮肤背景差异下的无创血糖测量校正方法研究[D]. 天津大学, 2019.
 - [31] 陶军, 靳亚非, 王礼春, 刘磊, 廖新学, 高修仁, 曾群英, 陈国伟, 马虹. 年龄对血管弹性和内皮细胞功能的影响[J]. 中华心血管病杂志, 2003(04):13-16.