

中国科学技术大学

工程硕士论文工作中期报告考核表

学号	SA21219079	姓名	方成峰	导师	陈香
工程领域	电子科学与技术		研究方向	生物医学信号处理	
论文题目	基于 PPG 技术的手势交互系统实现				
论文类型	① 工程设计（√）		② 研究论文（）		

论文中期小结

一、研究目标

完成如图 1 所示的基于 PPG 技术的手势交互系统的整体开发。针对系统所需的人机交互任务，设计一套 PPG 信号的前端采集设备并通过低功耗蓝牙与手机端软件组成一个完整的系统；针对现有穿戴式设备所遇到的电极偏移问题，提出一套针对性的算法流程：利用梯度提取及时域增强等预处理操作；同时针对交互系统的需求，研发一款由蓝牙扫描模块、数据传输模块、信号处理模块、分类算法模块、人机交互模块组成的手机端应用软件。

PPG传感器

信号采集模块

数字采样单元

信号采集前端

主控单元

无线传输模块
(低功耗蓝牙)

数据缓存

数据预处理

分类算法手势识别

受控APP

手机端

图 1 基于 PPG 技术的手势交互系统

二、研究内容

1. 采集前端设计与实现

2. 针对电极偏移问题的分类算法研究

3. 安卓平台的交互应用程序设计

三、论文工作进展

1. PPG 采集前端的设计及其实现

PPG 前端的器件选型指标及解决方案如表 1 所示，传感器部分选用的是支持三种光源的 PPG 传感器 MAX30101 及支持加速度与角速度采集的惯性传感器 JY901，这两种传感器都支持 IIC 协议，可以方便地和主控单元 STM32F103 进行集成。而数据传输模块则是采用低功耗蓝牙芯片 JDY-31，其传输距离及吞吐率都可以很好地满足本采集前端的需求，且其可以通过蓝牙透传的功能便捷地将数据传输到手机端。

	传感器阵列		传输方式	控制核心	尺寸重量
指标要求	通道数：≥6		无线传输	充足IO口	体积： ≤6.5*4.5 *2.0cm 质量： ≤80g
	PPG传感器	惯性传感器		体积小、易于佩戴	
	可选光源： 绿/红/红外	支持多轴加速度和角速度			
	支持IIC协议	支持IIC协议		高速数据传输接口	
解决方案	MAX30101 红光+红外+绿光 支持IIC协议 支持100Hz以上采样率 5.6mm × 3.3mm × 1.55mm	JY901 3轴加速度+3轴角速度 支持IIC协议 支持100Hz以上采样率 15.24mm × 15.24mm × 2mm	JDY-31 蓝牙芯片 传输频率2.4GHz 传输距离30m 吞吐率16K bytes/s	STM32 F103芯片	-

表 1 PPG 前端设计选型

PPG 采集前端的实体图如图 2 所示，主体由 4 个 MAX30101 模块和 JY901 模块组成，每个 MAX30101 模块搭载有 3 种 PPG 光源（红光、红外、绿光）的 PPG 信号，而 JY901 则是搭载 3 通道的加速度及 3 通道的角速度信号。而数据的采集及传输控制则由主控单元 STM32F103 及蓝牙模块完成。采样率目前设置为 100HZ，也即每秒由 MAX30105 及 JY901 采样 100 次，通过 STM32F103 进行整合，最后由蓝牙模块传输到手机端进行处理。

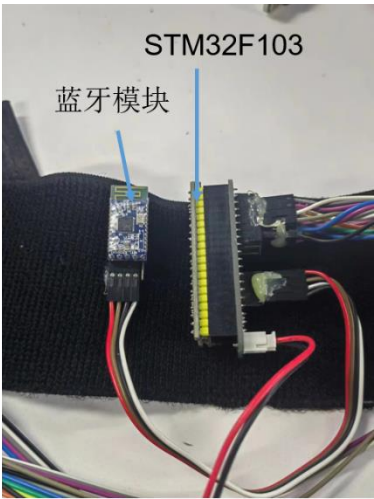


图 2 PPG 采集前端设计

而其 PPG 采集模块电路设计如图 3 所示其由采集芯片 MAX30101 和给其供电的电源模块组成，而整体的采集设备则由四个 PPG 采集模块和一个惯性传感单元 JY901 环绕而成，这些不同芯片的输出由 IIC 协议进行采集，采样率目前设置为 100HZ，也即每秒由 MAX30105 及 JY901 采样 100 次，通过 STM32F103 依照 IIC 协议进行采集，最后由蓝牙模块传输到手机端进行处理。而为了扩展 STM32F103 的 GPIO 口以满足本采集前端的需求，本研究还设计了一个转接板，其电路结构如图 4 所示，其主要依照图 3 的采集前端主体电路，扩展了一些如 GPIOA、GPIOB 等 GPIO 口。

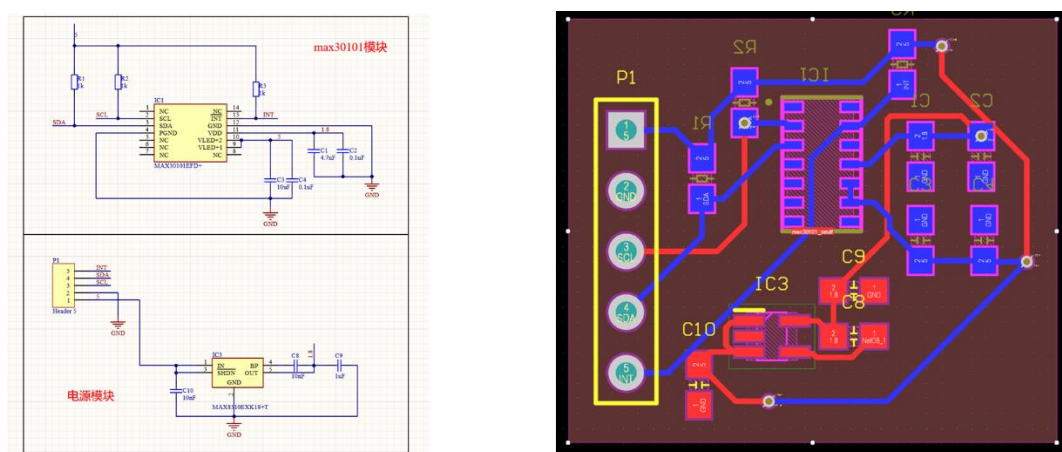


图 3 PPG 采集模块电路

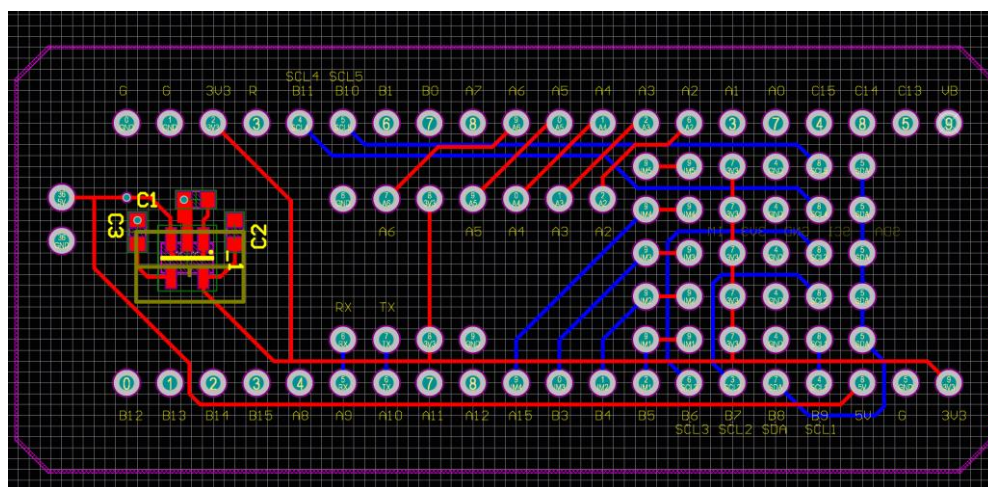


图 4 转接板电路

2. 针对电极偏移问题的分类算法流程及其性能验证

2.1 针对电极偏移问题的分类算法流程

电极偏移问题是指 PPG 信号这种由穿戴式设备所采集到的数据，数据分布会随着穿戴位置而产生变化，从而导致训练数据和测试数据必须是同一位置的数据，否则识别率会大幅下降。而为了克服 PPG 手势识别中的这种电极偏移问题，本研究提出了如图 5 所示的一套分类算法流程，其主要由预处理与位置分类两个步骤组成。

其流程具体如下所示：

(1) 采集多个传感器穿戴位置处的 PPG 手势数据

使用 $N(N>1)$ 通道PPG信号采集设备采集用户执行手势动作时产生的PPG信号，构建包含多个不同传感器穿戴位置的PPG手势数据集。

具体地，针对特定任务所需的包含 K 种动作的手势集，让用户在不同传感器穿戴位置下依次执行各手势动作。不同穿戴位置的实现可让用户进行一次实验后先取下传感器，让其休息几分钟后再次佩戴传感器进行实验，如此反复，以保证应尽量多地覆盖用户可能的传感器穿戴位置，获得不同穿戴位置下的多样性PPG手势数据。

对于采集到的PPG手势数据，进行活动段划分以提取出手势活动段的数据，再将不同的活动段归一化到相同的尺寸。

(2) 基于时间窗采样的样本增强

基于时间窗采样的样本增强步骤可分为以下几步：1) 确定窗长和移动步长。窗长定义每个时间窗包含的连续数据点的数量，移动步长为窗口之间的间隔；2) 将原始活动段数据切割成多个重叠的时间窗。通过移动窗口的起始位置，逐步滑动窗口来生成一系列的子样本；3) 采用与原始活动段相同的标签对增强后的子样本打标签；

(3) 时间维度和空间维度的梯度提取

假定经过时间窗采样增强后的子样本为origin，其尺寸为 (W,N) ， W 为时间窗采样中所设置的窗长大小， N 则为PPG传感器通道数。假设经过时间维度的梯度提取操作后的数据为timeGradient，而经过空间维度的梯度提取操作后的数据为spaceGradient。则有：

$$\begin{cases} timeGradient_{i,j} = data_{i+1,j} - data_{i-1,j} \\ spaceGradient_{i,j} = timeGradient_{i,j+1} - timeGradient_{i,j-1} \end{cases} \quad (1)$$

其中 $1 \leq i \leq W$, $1 \leq j \leq N$ 。

(4) 手势分类器训练

首先, 构建出一个基于特定算法的分类器, 可以是传统的机器学习模型, 如支持向量机 (Support Vector Machine, SVM)、线性判别器 (Linear Discriminant Analysis, LDA)、随机森林等, 也可以是神经网络模型如卷积神经网络 (Convolutional Neural Network, CNN)、长短记忆网络 (Long Short-Term Memory, LSTM) 等。

随后, 把前述经过处理的PPG手势数据集按一定比例划分为训练集和验证集。其中训练集用来训练分类器、验证集用来调整模型参数。经过训练后, 可得到本发明提出的可克服传感器位置偏移的PPG手势分类器。

(5) 位置分类器训练

同手势分类器的训练步骤, 首先构建出一个基于特定算法的分类器, 可以是机器学习模型或者神经网络模型。以位置作为标签, 把前述经过处理的PPG手势数据集按一定比例划分为训练集和验证集。其中训练集用来训练分类器、验证集用来调整模型参数。经过训练后, 可得到本研究所需的PPG位置分类器。

(6) 手势样本预测

对于一个需要预测的手势动作样本, 得到预测结果的流程如下:

- a、通过时间窗采样增强、时间维度梯度提取操作和空间维度梯度提取操作, 生成多个子样本;
- b、每个子样本通过位置分类器, 选取分类概率较大的x个位置的数据用于训练手势分类器, 得到可克服传感器位置偏移的PPG手势分类器;
- c、对每个子样本, 采用训练好的可克服传感器位置偏移的PPG手势分类器得到对应的预测结果;
- d、对所有子样本的预测结果采用投票等操作, 得到最终的测试结果

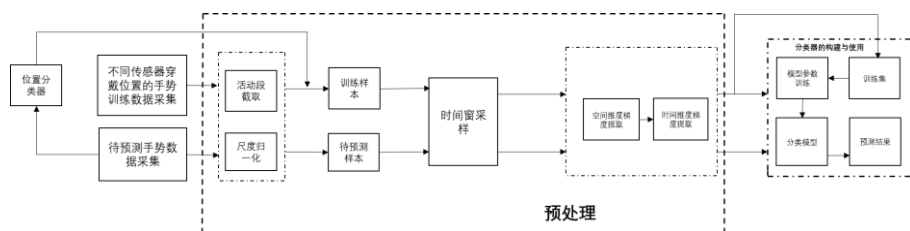


图 5 针对电极偏移问题的分类算法

2.2 手势集构建

为了验证本研究所提出的分类算法针对电极偏移问题的有效性，我们设计并完成了 10 名受试者的 10 批次 8 手势数据采集实验。具体步骤如下：

(1) 设计手势集

依照如下原则：

1. 产生明显的动脉血流量的变化；
2. 覆盖腕、指关节的多种运动；
3. 符合日常使用习惯，简单易于操作；
4. 具备实际应用场景中的交互意义。

我们设计了如图 6 所示的 8 手势集。



图 6 八手势集

(2) 数据采集

实验共招募了 10 名健康的右利手参与者进行数据采集实验，其中男性参与者和女性参与者各 5 名，年龄均在 22~26 岁之间，所有参与者没有任何的神经和肌肉骨骼疾病或病史。为满足对传感器偏移问题研究的需要，需获得含有充分传感器偏移信息的手势 PPG 数据。为此，对每位受试者在不同时间点开展了 10 批次数据采集。在每批次数据采集实验中，受试者在

实验员的指示下依次执行 8 种手势动作，每种手势动作连续重复执行 50 次，对于手势动作的执行时间没有严格的要求，受试者按照自己感觉舒适的速度执行即可，不同手势动作间提供 1~2 分钟的休息时间。每一批次数据采集实验完成后，将采集设备从受试者的手腕取下。为避免连续过长时间执行手势动作而导致肌肉疲劳，影响所采集到信号的质量，在每一批次采集结束后，受试者需休息大约 10~15 分钟，然后再次重新佩戴采集设备，开始下一批次的数据采集实验。

在实际生活中佩戴诸如智能手表等设备时难免会因表带的松紧程度或者剧烈运动等因素造成表盘位置发生小范围内的移动，而这种小范围移动的方向可以是上、下、左、右以及其它各种角度。因此，本研究对每位受试者采集的 10 批次数据充分包含了各个方向的偏移，分别为：参考位置、上偏、下偏、参考位置、左偏、右偏、左上偏、左下偏、右上偏、右下偏。首先，所谓的参考位置是按照以下方式进行确定的：根据大多数人群的佩戴习惯，表盘在掌背一面，以便抬腕时能直接看到表盘的信息，将此面称为手腕正面，对应掌心的一面则称为手腕背面；在佩戴采集设备时，手腕正面和背面分别分布有两个 PPG 传感器芯片，手腕同一面的两个 PPG 传感器芯片一个靠近尺骨、另一个靠近桡骨，四个 PPG 传感器芯片距离手腕上凸起的腕骨约一指宽度的距离，将在此位置佩戴采集设备定义为参考位置。然后，当伸出手掌，手腕正面朝向自身时，直视所看到的上、下、左、右方位即为上述的偏移方向，而左上、左下、右上、右下则表示两个方位的 45°夹角方向，所有方向的偏移距离参考位置约为 1 厘米。参考位置和各偏移位置的示意图如图 7 所示。



图 7 参考位置和各偏移位置

2.3 预处理效果验证

为验证预处理对解决电极偏移问题的作用，本研究采用受试者 1 的上偏移位置和下偏移位

置的数据做 t-SNE 可视化，预处理前后的可视化结果分别如图 8 及图 9 所示，以手势 4 为例，可以看到经过预处理，上偏移位置与下偏移位置的手势 4 数据被聚集到了一起。

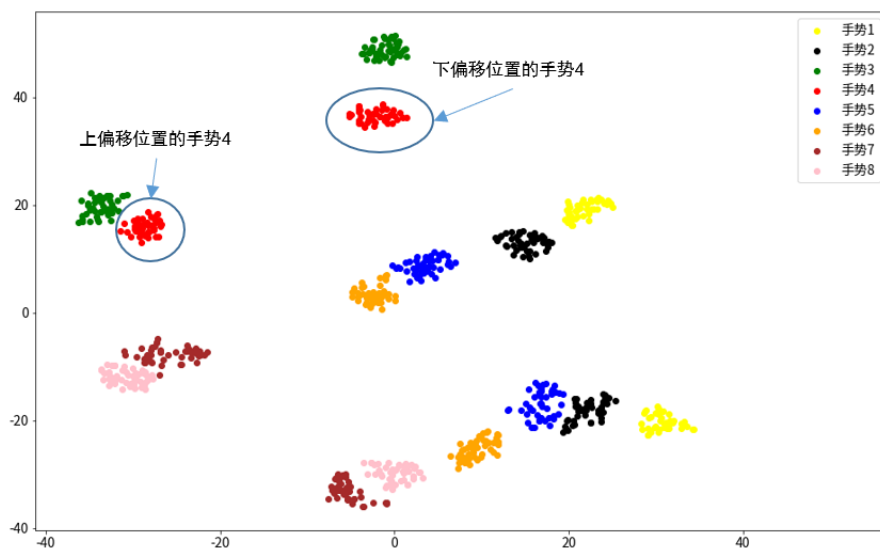


图 8 预处理前的 t-SNE 可视化结果

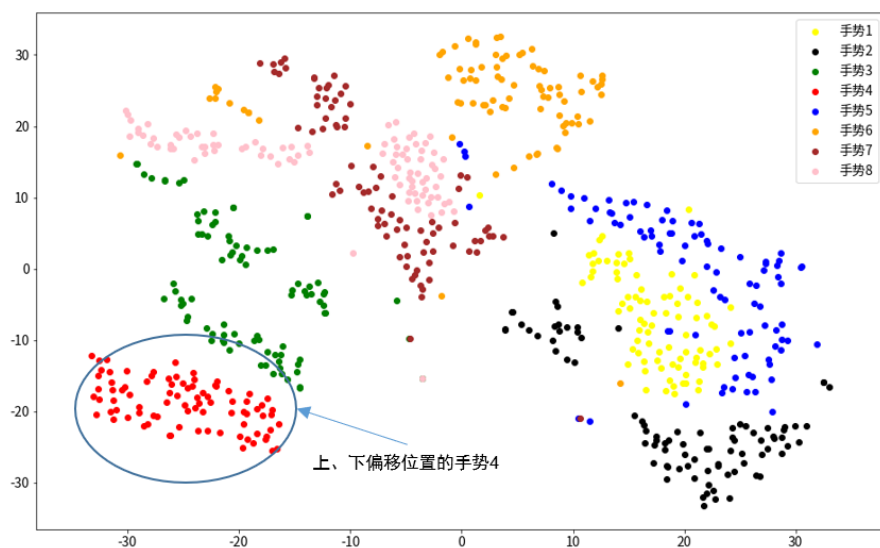


图 9 预处理后的 t-SNE 可视化结果

2.4 分类器结构

考虑到采集的多通道 PPG 信号同时具有时序信号和空间信号的特点，本研究提出了如图 10 所示的基于 CNN 和 LSTM 的神经网络结构作为位置分类器和手势分类器。其由两层卷积层、两层 LSTM 层及一层全连接层顺序连接而成，其中卷积层又由一维卷积、批归一化、relu

激活、池化及 dropout 层顺序连接而成。

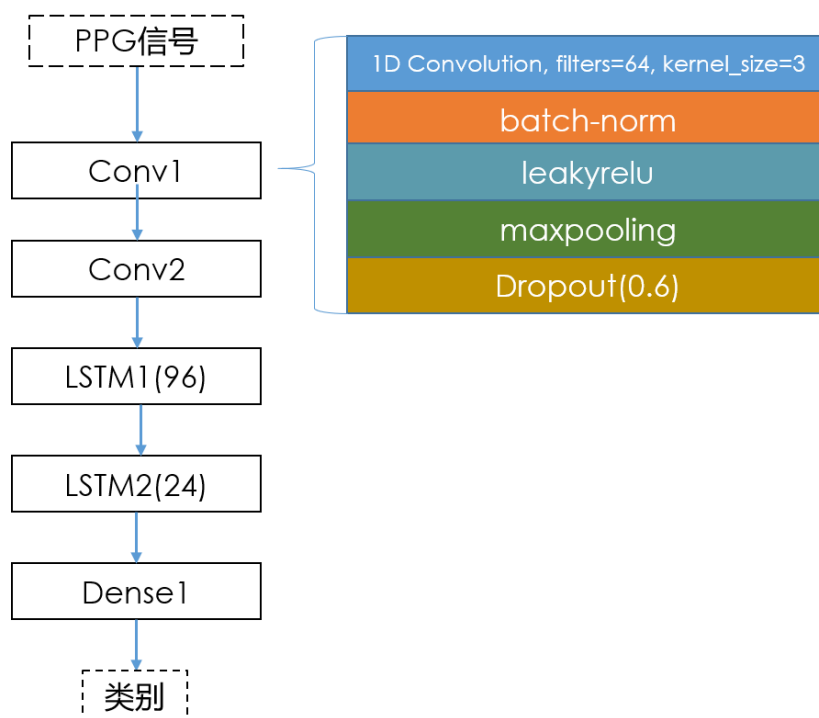


图 10 基于 CNN 和 LSTM 的神经网络结构

2.5 位置作为标签的数据可分性

按照如图 5 所示的分类算法流程，为了得到测试数据对应的分类位置，首先需要以位置为标签训练一个位置分类器，这就要求位置标签数据具有良好的可分性。因此本研究采用了受试者 1 的 10 位置数据进行 t-SNE(t-Distributed Stochastic Neighbor Embedding)可视化处理。其结果如图 11 所示，可以看到其具有良好的可分性，而按照图 10 所示的神经网络模型，以位置作为标签，训练测试比=3: 7 对全部 10 位受试者的数据其进行分类，结果如图 12 所示，其测试识别率为 $99.96\% \pm 0.05\%$ ，也说明了其可分性完全满足我们的需求。

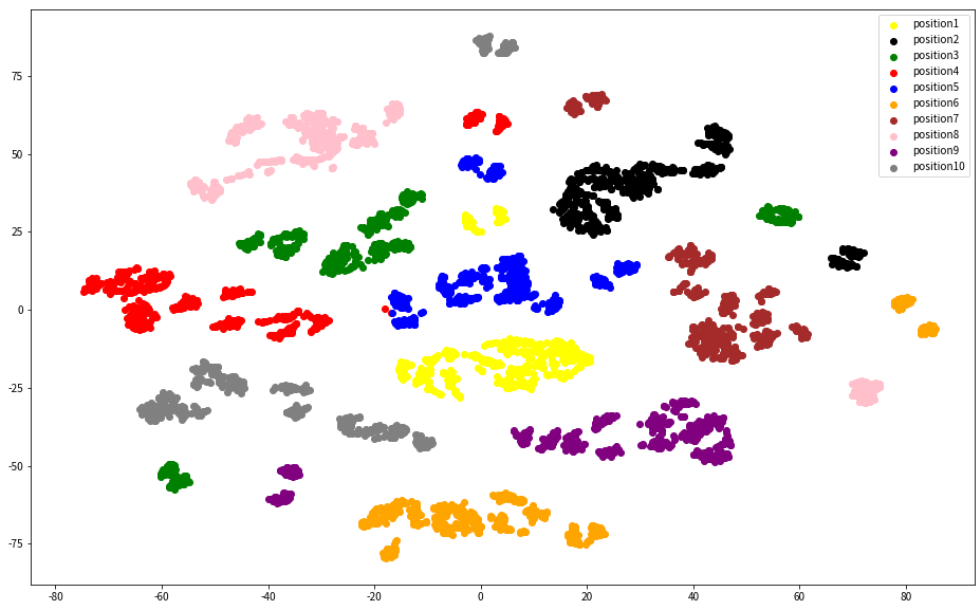


图 11 位置标签数据的 t-SNE 可视化结果

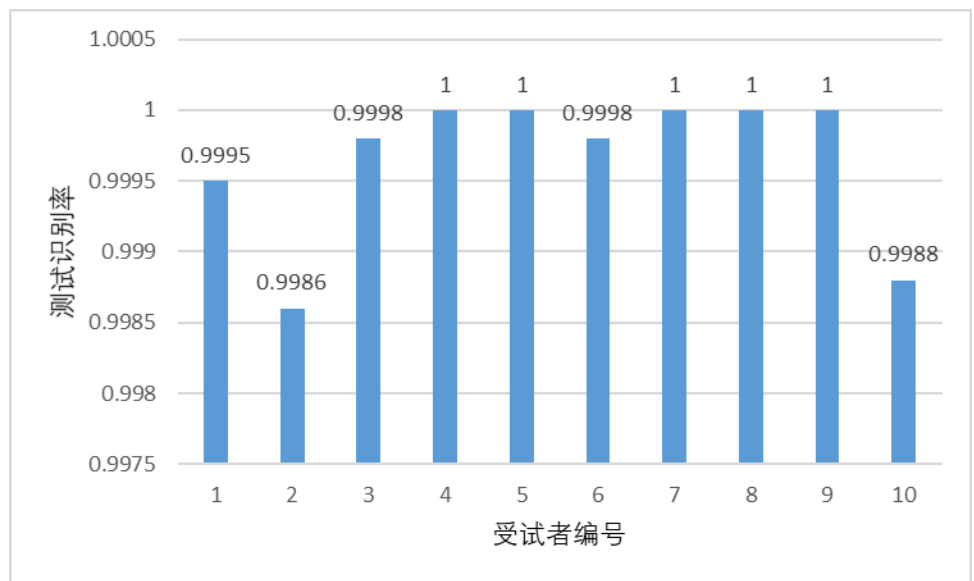


图 12 十受试者位置分类结果

2.6 确定用于训练的位置数量

为了确定图 5 流程中所应采取的训练位置的最佳数量, 本研究开展了探索识别率随训练位置数量变化关系的相关实验: 训练位置数量 x 从 1 到 10 变化, 每个受试者的 10 个偏移位置, 依次取 1 个做测试位置, 对每条测试数据, 让其通过位置分类器, 得到分类概率前 x 个位置做训练位置, 训练位置的数据训练得到相对应的手势分类器, 测试数据再通过手势分类器得到测试结果, 最后取 10 个位置的测试识别率的均值作为受试者的平均识别率, 取全部受试者均

值作为总体识别率。其结果如图 13 所示，可以看到其在训练位置数量取 7 时效果最佳。

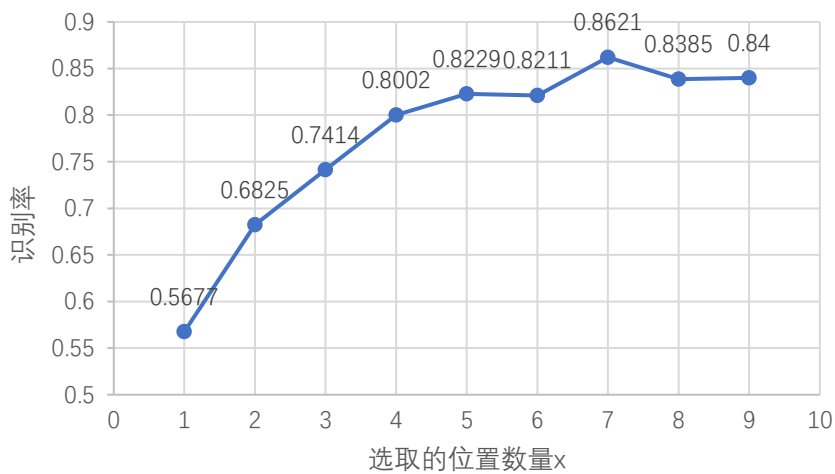


图 13 识别率所训练位数数量 x 关系

2.7 分类算法流程有效性验证

每个受试者的 10 个偏移位置，依次取 1 个做测试位置，对每个测试数据，选取分类概率前 7 的 7 个位置做训练位置，由训练位置的数据训练得到手势分类器，测试数据再通过手势分类器得到测试识别率，最后取 10 个位置的测试识别率的均值作为受试者的平均识别率。得到结果如图 14 所示，相比于采用全位置+未预处理方案的对照组，本研究采用的位置分类+预处理的方案的手势识别率做到了约 35%的提升：从均值 $51\% \pm 8.2\%$ 提升到了均值 $86\% \pm 3.8\%$ 。

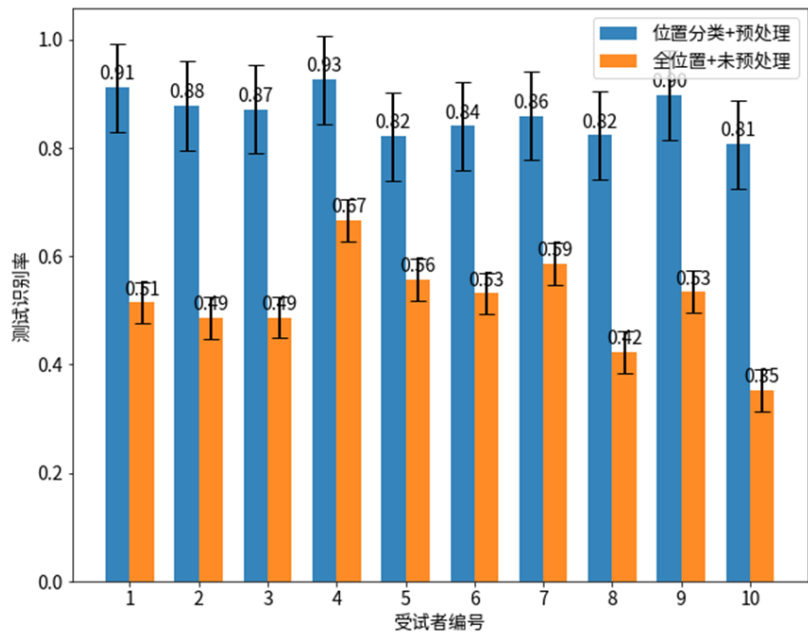


图 14 分类算法流程及对照组识别率

3. 安卓端交互应用程序设计及其进展

3.1 交互应用程序整体设计

本研究设计的基于 PPG 技术的人机交互系统的移动端应用程序如图 15 所示，主要由数据传输模块、信号处理模块、分类算法模块及人机交互模块组成。初始时，由数据传输模块也即低功耗蓝牙模块与采集前端建立通信，负责完成数据由前端到手机端的传输；随后信号处理模块负责将从前端接受的 PPG 信号进行图 5 中的一系列预处理操作：活动段截取、尺度归一化、时间窗采样、梯度提取；经过这一系列的预处理之后，再由分类算法模块利用训练好的分类模型去识别，以得到最后的预测结果，同时分类算法模块还负责模型的训练、存储以及读取；人机交互界面利用得到的预测结果与用户进行交互操作，例如在游戏中控制人物移动或执行手机应用的操作（如打开淘宝/QQ 等）。

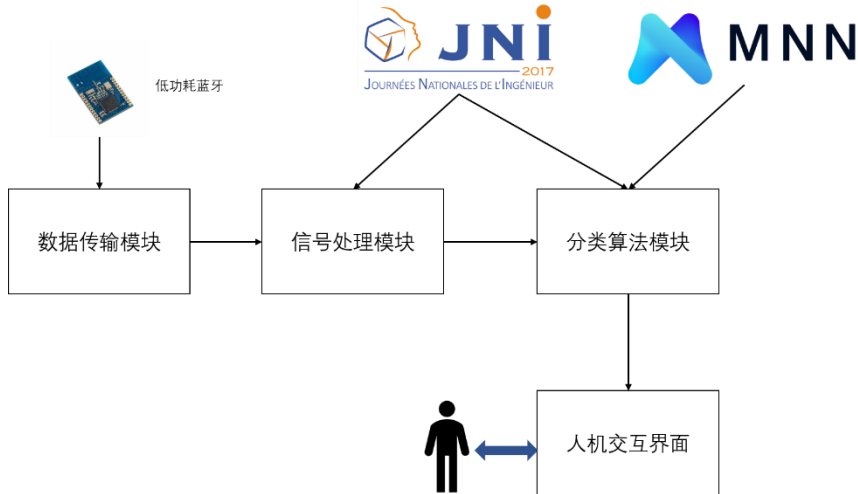


图 15 PPG 人机交互系统终端方案

而具体的软件接口设计则如图 16 所示，底层的低功耗蓝牙接口负责和前端进行通信，从前端接受 PPG 信号，同时给上层的预处理及分类算法模块提供信号读取接口；而信号处理模块则提供一系列预处理算法：活动段截取、尺度归一化、时间窗采样、梯度提取等，其向下使用低功耗蓝牙接口与前端进行通信，向上提供各种预处理算法实现，且其会实时监测是否出现了手势活动段，若出现活动段，则将捕捉到的手势活动段进行一系列预处理及分类算法流程；分类算法模块则是提供了一个分类器接口，其实现可以是传统机器学习算法：SVM 或者神经网络算法：CNN、LSTM 等；而人机交互界面则包括如画图、游戏、蓝牙选择等一系列界面。为了系统整体的实时性，对于耗时较大的信号处理及分类算法模块，本研究采用的是利用 C++ 进行原生开发，然后用 JNI 技术嵌入安卓端的技术方案；而人机交互界面的部分则依旧采用 Java 进行原生开发。而对于分类算法接口的神经网络实现，本研究引入了阿里开源的神经网络

框架 MNN，其在移动端的表现要远远优于市面上其他主流开源框架。

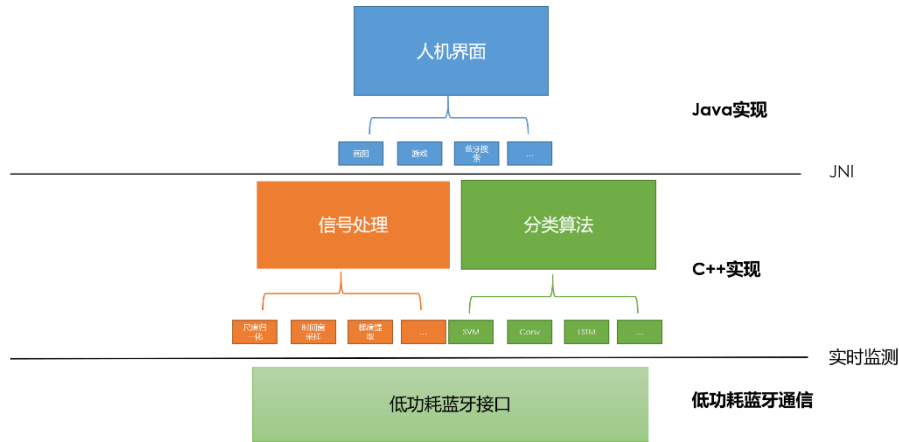


图 16 软件设计接口

3.2 实时监测算法实现

实时监测算法的主要依据的原理是手势段的起始点和终止点恰好对应于梯度能量值的波峰两端，如图 17 所示。其中梯度能量值 $energy$ 的定义如下式所示：

$$\begin{cases} energy_j = \sum_{i=j}^{j+w} gradientData_i^2 \\ gradientData_i = \sum_k data_{i+1,k} - \sum_k data_{i-1,k} \end{cases} \quad (2)$$

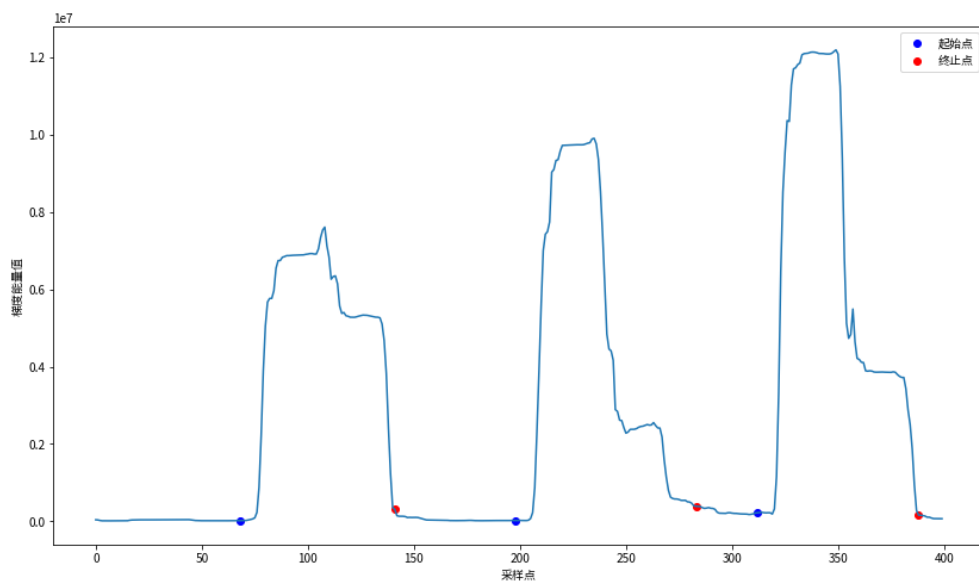


图 17 手势活动段起始与终止点

而其主要流程如图 18 所示：（1）首先按照公式(2)算出静息状态下的梯度能量均值作为 **threshold**；（2）按照公式(2)实时计算 **data** 对应的 **energy** 值，然后判断其是否大于 **threshold**；（3）若其大于 **threshold** 且还未找到起始点，则将该点作为起始点；（4）若其小于 **threshold** 且已经找到起始点，则将该点作为终止点。

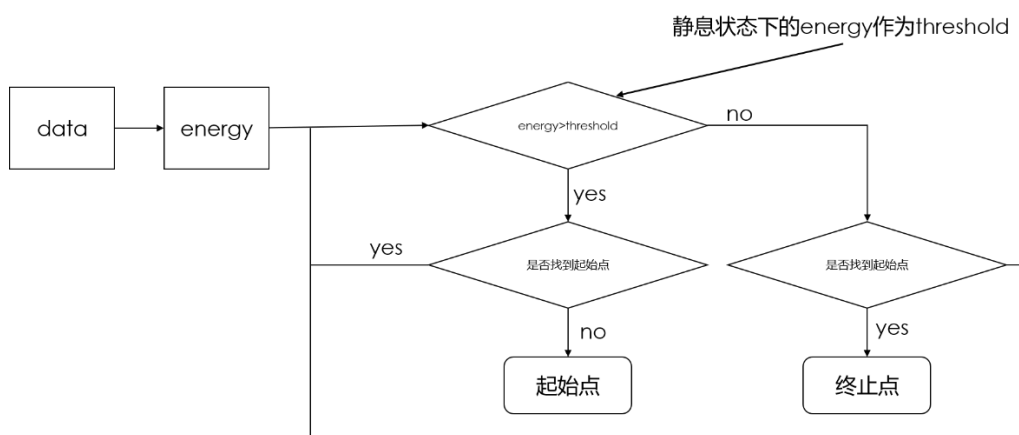


图 18 实时监测算法流程

3.3 基于神经网络的分类器移植

本研究采用的分类器（位置分类器及手势分类器）是如 2.2 节所示的基于 CNN 与 LSTM 的神经网络结构，而神经网络移植到手机端的整体流程如图 19 所示，其中 MNN(Mobile Neural Network)是阿里开源的一个全平台轻量级高性能深度学习引擎。其主要步骤如下：

- （1）依据 2.1 节所述流程在电脑端利用 Tensorflow 框架训练得到对应的神经网络模型；
- （2）在电脑端编译 MNN 源码获取相应的转换工具及共享库；
- （3）利用转换工具将步骤（1）中训练好的模型转换为对应的 MNN 模型；
- （3）利用步骤（2）中编译得到的共享库将对应的 MNN 模型部署到手机端。

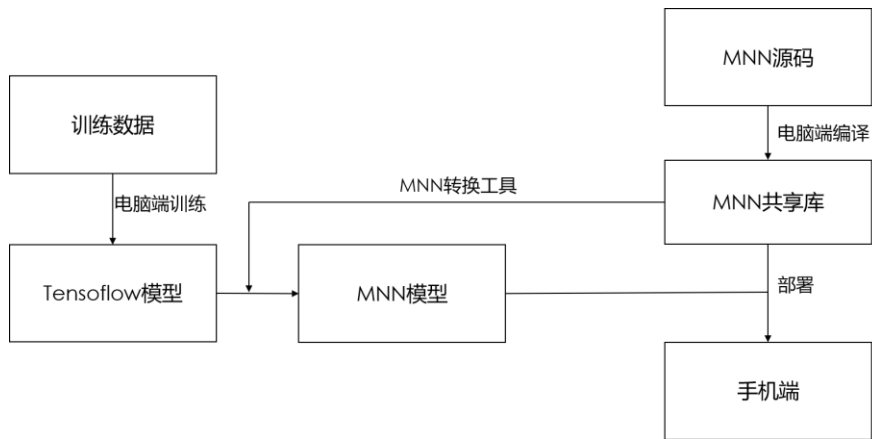


图 19 神经网络模型移植流程

1.3 系统有效性及实时性验证

为了验证软件系统的有效性及时时性，本研究开展了如下实验：让用户按照前文所述的步骤构建一个多批次的 8 手势训练集，从而训练得到分类器（位置分类器及手势分类器），随后用户佩戴前端采集设备，8 手势，每个手势重复 50 次左右，进行实时识别测试。

其测试识别率结果如图 20 所示，整体的识别率在 81%左右，已经和图 14 电脑端线下处理的结果相近；而实时性的结果如图 17 所示，其系统的整体耗时：从用户执行手势到手机端响应，在 200ms 以内，完全满足实时性的要求。

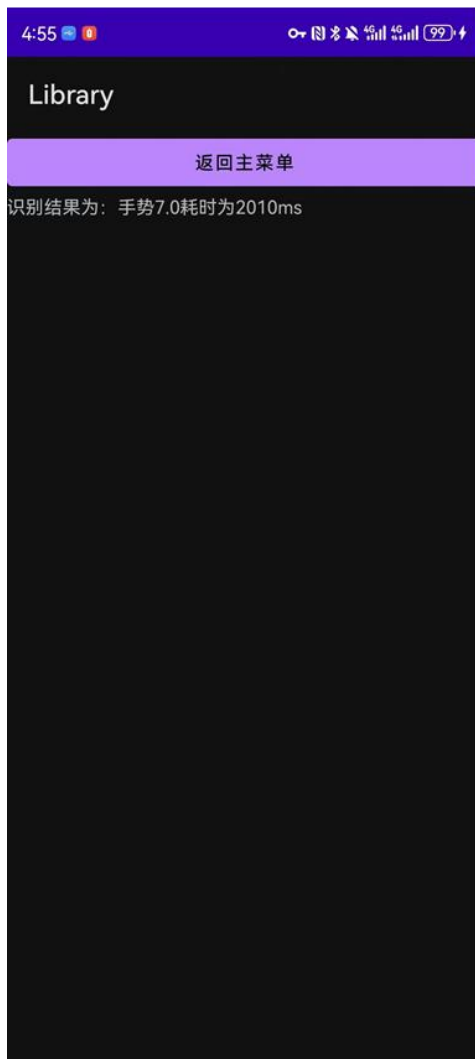


图 20 模型训练及验证

下一步工作计划:

- 1、进行移动端的人机交互界面开发;
- 2、进一步完善克服电极偏移问题的手势分类算法;
- 3、进行系统整体调试, 完善整体系统性能。

在学期间所参加科研项目、获奖情况、发表文章及发明、专利等

项目名称	时 间	成果鉴定或发表刊物、出版单位
专利：一种可克服传感器位置偏移的光电容积脉搏波手势识别方法	2023	专利公开号：CN116861325A
论文中期报告考核评语		
<div style="text-align: right;">导师签字：(学校) (企业)</div> <div style="text-align: right;">培养单位负责人签字： 年 月 日</div>		