

北京交通大学

硕士学位论文

基于深度学习的心电图类型识别和个体识别

ECG Classification and Identification
Based on Deep Learning

作者：陈 滨

导师：郭宇春

北京交通大学

2019年5月

学位论文版权使用授权书

本学位论文作者完全了解北京交通大学有关保留、使用学位论文的规定。特授权北京交通大学可以将学位论文的全部或部分内容编入有关数据库进行检索，提供阅览服务，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交论文的复印件和磁盘。学校可以为存在馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。

(保密的学位论文在解密后适用本授权说明)

学位论文作者签名：陈斌

签字日期：2019年5月27日

导师签名：邵宇

签字日期：2019年5月27日

学校代码：10004

密级：公开

北京交通大学

硕士学位论文

基于深度学习心电图类型识别和个体识别

ECG Classification and Identification
Based on Deep Learning

作者姓名：陈 滨

学 号：16120044

导师姓名：郭宇春

职 称：教 授

学位类别：工 学

学位级别：硕 士

学科专业：通信与信息系统

研究方向：信息网络

北京交通大学

2019年5月

致谢

本论文的研究工作是在我的导师郭宇春教授的悉心指导下完成的。郭宇春教授科学的工作方法、开阔的视野以及渊博的知识给了我极大的帮助。郭宇春教授严谨的治学态度、精益求精的学术风范、恪尽职守的工作作风以及豁达的人生理念，深深的感染和激励着我不断进取，对我以后的工作和学习有着很大的影响。在此衷心感谢三年来郭宇春教授对我的悉心指导和关怀。

感谢实验室里的所有老师。衷心感谢陈一帅老师、赵永祥老师、李纯喜老师、郑宏云老师和张立军老师在我研究生学习阶段对我的无私帮助和关怀，我所有的科研成果都凝结着各位老师的辛勤汗水。在此向各位老师表示诚挚的谢意。

另外，在实验室工作和撰写论文期间，谭晓颖师姐、刘翔师兄、李乃鹏师兄、贾凯凯师兄、李梦月师兄、田丽红师姐、李俊峰师兄等师兄师姐们帮助我解决了很多疑惑，李勇宏、唐伟康、张大富、胡玮、周伟、张莹、孙欢、王亚珊等同学也给予了我很多帮助，在此向他们一并表达我的感谢之意。

最后，特别感谢一直给予我无尽的爱和支持的父母和女友，正是来自他们的默默奉献，才使得我顺利的完成学业，成为社会的有用之才。

摘要

传统上，医生采用人工的方式判读心电图以评估患者病情。近年来，人工智能技术促进了心电图数据的智能信息挖掘，实现了心电图类型自动识别及异常发现，也实现了以心电图为生物特征的个体识别系统。

典型的基于人工智能的心电图分析系统包括三个主要处理过程，包括心电图样本生成、特征提取和模型分类。现有研究在这三个处理过程上均存在明显不足。样本生成阶段使用固定序列长度切分心拍且切分的起始位置固定，没有考虑心拍长度的动态性和导致切分得到的样本数量非常有限，严重影响了系统的分类性能。特征提取阶段提出的不同特征提取方法，大都基于单个心拍提取特征，忽略了心拍内部的局部特征和心拍之间的内在关联。最后，分类阶段所使用简单的神经网络，不能准确获取不同特征向量的权重，导致模型分类效果受限。综上，现有基于深度学习的心电研究还存在明显的性能提升空间。

为此，本文完成了基于深度学习的心电图类型识别（工作 1）和心电图个体识别（工作 2）的两项研究。本文思想是提出准确的心拍切分方法和有效的特征提取方法，改善神经网络分类器，提高现有心电系统的性能，对基于心电图的系统实用化具有重要价值和理论意义。论文的主要贡献如下。

(1) 提出了一种自适应动态心拍切分方法，准确切分心拍，提升了工作 1 的识别准确率。该方法利用动态的 RR 间隔长度表示单个心拍长度，并依据此切分心拍，体现心拍长度的动态性。此外，工作 2 改进了一种样本提取方法，解决了现有研究样本数量不足的问题，同时避免了现有研究中训练数据和测试数据重叠导致的评估效果虚高的问题。

(2) 提出了基于多尺度的特征提取方法，从多维度挖掘到心电图的显著特征，明显改善了工作 1 的模型分类效果。方法的主要特点是，从局部心拍、单个心拍和多个心拍三种尺度上，分别利用深度学习自动提取特征和心电图医学知识手动提取特征，从而实现了对样本更为丰富和准确的表达。

(3) 将计算机视觉中的通道注意力方法引入到神经网络模型中，自动给不同特征向量分配权重，进一步提升了模型识别效果。最终实验表明，本文在工作 1 和 2 中的识别准确率分别高达 96.94%和 95.86%，比现有研究分别提高了 3.5%和 3.0%。

关键词：类型识别；个体识别；心拍切分；特征提取；通道注意力模块

ABSTRACT

Traditionally, doctors used manual methods to interpret ECG to assess patients' condition. In recent years, artificial intelligence technology has promoted information mining of ECG data, achieved ECG automatic classification and abnormal findings, and realized ECG individual identification system as well.

A typical AI-based ECG analysis system consists of three main processes, ECG sample generation, feature extraction, and classification. Existing research has obvious shortcomings in these three processes. Above all, sample generation, the fixed sequence length is used to segment the beat and the starting position of the segmentation is fixed, ignoring the dynamics of the beat length and resulting in the insufficiency of samples, which seriously affects the classification performance. Secondly, the different methods proposed to extract features are mostly considered from single beat scale, ignoring the local features of a beat and the intrinsic association between beats. Finally, the simple neural network used for classification cannot obtain accurate weights of different feature vectors, resulting in that the model classification accuracy is limited. In summary, with the utilization of deep learning, there is still a significant space for improvement.

To this end, this thesis concentrates on the study of ECG based on deep learning, and completes two tasks, i.e. ECG classification (task 1) and ECG identification (task 2). The core idea of this thesis is to improve the performance of the ECG system by proposing accurate beat segmentation and effective feature extraction methods, improving neural network classifier, which manifests significant value both in theory and application aspects. The main contributions of the thesis are as follows.

(1) An adaptive dynamic beat segmentation method is proposed to accurately segment the beat and improve the classification accuracy corresponding to task 1. The method uses dynamic RR interval length for representation of single beat length, through which effective segmentation and dynamic representation are realized simultaneously. In addition, task 2 improves a sample extraction method, which solves the problem of samples deficiency, and avoids the unreliability of high evaluation accuracy caused by the overlap of training data and test data in existing research.

(2) A multi-scale feature extraction method is proposed. Significant features of ECG captured from multiple dimensions obviously improve the classification accuracy

for task 1. The main idea of the method is that features scaling on the local beat, the single heart beat and the multiple beats are extracted automatically and manually, respectively. For the former, it is achieved with the use of deep learning, while for the latter, ECG medical knowledge is involved. Therefore, the expression of the sample achieves variety and accuracy at the same time.

(3) The channel attention method in computer vision is introduced into the neural network we used, and weights are automatically assigned to different feature vectors, which further enhances the model accuracy. The final experiments show that the accuracy of this thesis in task 1 and 2 reaches 96.94% and 95.86%, which is 3.5% and 3.0% higher than related work, respectively.

KEYWORDS : ECG classification; ECG biometric; beat segmentation; feature extraction; channel attention module

目 录

摘要	III
ABSTRACT.....	IV
1 引言	1
1.1 研究背景与意义	1
1.2 国内外研究现状	2
1.2.1 心电图类型识别研究现状	3
1.2.2 心电个体识别研究现状	4
1.3 本论文的主要研究内容及贡献	6
1.4 本论文的组织结构	7
2 论文相关知识介绍	8
2.1 心电图简介	8
2.1.1 心拍简介	8
2.1.2 常见异常心拍简介	12
2.2 生物识别系统简介	13
2.2.1 常见生物识别系统	13
2.2.2 心电生物识别系统的优缺点	14
2.3 深度学习简介	15
2.3.1 深度学习概念简介	15
2.3.2 常用神经网络简介	16
2.3.3 深度学习框架简介	19
2.4 数据集简介	20
2.5 本章小结	21
3 心电图类型识别系统	23
3.1 现有工作局限及改进思路	23
3.2 心拍切分	25
3.2.1 自适应的单拍切分	25
3.2.2 重叠采样的子拍切分	26
3.2.3 多拍的切分	28
3.3 多尺度特征提取	28
3.4 分类器的改进	29
3.5 实验验证	30
3.5.1 实验设置	30

3.5.2 实验网络设计	31
3.5.3 不同切分方法实验	33
3.5.4 多尺度特征实验	34
3.5.5 各类型心拍的实验表现	36
3.6 本章小结	37
4 心电图个体识别系统	39
4.1 现有工作局限及改进思路	39
4.2 样本生成	41
4.3 小波变换	42
4.4 通道注意力模块的改进	46
4.5 实验研究	47
4.5.1 实验网络设计	47
4.5.2 参数选定实验	48
4.5.3 相关工作对比	51
4.5.4 算法稳定性分析	53
4.6 本章小结	56
5 结论	57
5.1 本文工作总结	57
5.1.1 心电图类型识别研究总结	57
5.1.2 心电图个体识别研究总结	58
5.2 未来工作展望	59
参考文献	60
作者简历及攻读硕士/博士学位期间取得的研究成果	64
独创性声明	65
学位论文数据集	66

缩略词表

英文缩写	英文全称	中文全称
AI	Artificial Intelligence	人工智能
ECG	Electrocardiogram	心电图
CVD	Cardiovascular Disease	心血管疾病
NCD	Non-Communicable Disease	非传染性疾病
AAMI	The Association for the Advancement of Medical Instrumentation	美国医疗仪器促进协会
CNN	Convolutional Neural Network	卷积神经网络
RNN	Recurrent Neural Network	循环神经网络
MLP	Multi-Layer Perceptron	多层感知机
SVM	Support Vector Machine	支持向量机
AP	Average Pooling	平均池化
MP	Max Pooling	最大池化
ReLU	Rectified Linear Unit	修正线性单元
BS	Beat Segmentation	固定长度的心拍切分
ABS	Adaptive Beat Segmentation	自适应的心拍切分
OWS	Overlapping Wave Segmentation	重叠采样的子波切分
NWS	Non-overlapping Wave Segmentation	非重叠采样的子波切分
CAM	Channel Attention Module	通道注意力模块
MBS	Multiple Beat Segmentation	多拍切分
MF	Multi-scale Features	多尺度特征
HF	Hand-crafted Features	手动特征
STFT	Short Time Fourier Transform	短时傅里叶变换
FFT	Fast Fourier Transform	快速傅里叶变换
ELU	Exponential Linear Units	指数线性单元

1 引言

1.1 研究背景与意义

随着城镇化的发展，人们的生活更加便捷，同时，生活节奏也在逐渐加快，伴随而来的生活压力也不断增加，致使患心血管疾病的人群也大大增加。根据 2018 年世界卫生组织报告统计，2016 年，全球由于心血管疾病（简称 CVD）造成了约 1790 万人死亡，这个比例占到非传染性疾病（简称 NCD）造成死亡的 44%，在所有的非传染性疾病中，排名第一。在国内，根据中国心血管疾病报告 2017 的描述，2016 年，全国大约有 2.9 亿人患有心血管疾病，并且农村居民患有心血管疾病的比例超过了城镇居民，二者之间的差距还在进一步扩大。同时，居民死亡人数中，超过 40% 是由心血管疾病造成的，居首位，远高于癌症和其他疾病。同时，另一方面，根据国际统计年鉴 2017，2014 年，我国每千人口医生数为 1.9 人，每千人口病床数为 3.8 张，这意味着，我国平均每 1000 人，共享 1.9 名医生，2.2 张病床。而国际高收入国家的每千人口医生数为 2.9，每千人口病床数为 4.2 张，这两个数值都远远高于我国的水平。这反映了我国医疗水平与世界高收入国家还有较大差距，考虑到我国资源分配的不平衡，在农村，也就是心血管疾病高发的地区，这两项指标的数值会大打折扣。

一方面是患有心血管疾病的人越来越多，另外一方面，是医疗资源的短缺，尤其是在心血管疾病高发的农村，这使得对心脏活动的动态实时监测的需求在不断增加，研究出一种可穿戴式或便携式的心电图（简称 ECG）监测设备，以及心电图信号类型自动识别系统显得十分必要和有意义，这样就可以及时向患者通报心电图异常，及时就医，保障人们的生命财产安全。同时，一个自动心电图类型识别系统，还可以为医生提供帮助，为医生的诊断做出辅助判断，提高医生看病效率，缓解医疗资源紧张的问题。

近些年来，随着科技的不断发展，传统的身份识别手段，例如 ID 卡，签名等身份识别技术已经无法满足人们的日常需求，同时，不法分子利用高科技复制 ID 卡和签名的新闻屡见不鲜，导致人们的个人财产安全不断受到威胁，寻求新的身份识别手段迫在眉睫。而生物识别（又可称个体识别）技术由于其高唯一性，难复制性，高可识别率的特征被逐渐用作身份识别工具。在各行各业中，各种生物识别系统都得到了广泛的发展与应用，例如指纹识别技术，在各种公司打卡制度和电子设备的用户认证中都受到了广泛的青睐；人脸识别，在电子设备中被大量应用为密码，对使用电子设备的个体进行验证，以保障用户电子产品的安全与

隐私；虹膜识别，在一些银行和个人保险柜的场景中，也得到了较多应用。但是目前使用的生物识别系统，多为一些静态生物特征，像人脸、指纹、虹膜都是固定的特征，在生理上很少发生变化，这样的生物特征优缺点都很明显。优点是由于其物理特征固定，方便提取有效的稳定的特征，对识别准确率的提升有很大的帮助。同时，由于其物理特征固定，也为不法分子提供了可乘之机，可以利用高科技复制这些信息，只需要几张照片，就可以利用 3D 打印技术打印出一个三维立体的人脸，指纹的提取更为简单，只要是用户徒手碰过的东西，几乎都会留下指纹。

生物识别被视为是目前最安全的身份识别技术之一，是当前研究和应用的热点，但是传统的生物识别系统，都是对静态生物特征进行识别，使得生物识别技术的安全性受到了极大地威胁。而一些动态生物特征，例如心电图，这种随环境、心情和个体不同而产生变化的动态生物特征，是未来利用生物特征进行身份识别的趋势所在。动态生物特征由于本身是变化的，使得不法分子想复制这些生物特征是极其困难的，并且，就算不法分子得以复制，在不了解生物识别系统提取特征的方式的情况下，依旧很难骗过生物识别系统，几乎无法复制。所以，研究利用心电图进行个体识别的技术非常有意义，并且具有广泛的应用前景。

1.2 国内外研究现状

随着医疗技术和水平的不断提升，有很多医疗数据不断积累，越来越多的研究者也开始从事医疗数据的研究了。以心电图为例，对心电图的研究有着较长的时间，总的来看，心电图的研究工作主要可以分为两类：一类是对心电图类型进行识别。所谓对心电图类型进行识别，就是区分出不同类型的心电图，最主要的识别方式分为两种，一种是心电图正异常的二类识别，即只需要辨认出心电图是正常的还是不正常的（不正常又称为异常）。虽然该任务难度相对较小，但是这种研究在实际应用中也有很重要的意义，即当发现患者有心电图异常时，就及时提醒患者就医，起到了医疗监护的作用，并且降低模型错误判断带来的风险；另外一种是对心电图正异常的多类型识别，即分为正常和多种异常的类别，把异常类细化成了多种类型，根据异常心电图的严重程度，给患者不同的建议。这种研究的实际意义更大，因为心电图异常的类型很多，有些异常可能是由于熬夜等短暂不规律生活造成的心律异常，通过一两天自我调节即可恢复正常，无需就医，但是有些异常可能是一些重大疾病的前兆，需要立即就医。心电图多类型识别就能根据不同的情况给出不同的建议。第二类是利用心电图进行个体识别。心电图信号因为是一种动态特征，很难复制，高安全性，非常适合个体识别。个体识别的工

作也可以分为两种，一种是授权系统，即通过模型识别一个用户在不在指定的用户集合内，属于二分类问题。另外一种识别系统，即识别用户是指定用户集合内的哪一个具体用户。这两者都有对应的实际应用场景。授权系统多用于一些权限系统，比如对于某些高机密文件，只有到达某种行政级别的人才有权查看，否则就拒绝查看请求。识别系统可用于打卡签到系统，通过心电图识别出来签到的是某个具体的人。接下来对这两类工作进行详细的介绍。

1.2.1 心电图类型识别研究现状

心电图类型识别的工作大致可以分为两个时期，第一个是在深度学习被广泛使用之前，传统的机器学习分类器（例如随机森林，支持向量机）一般是广泛的选择。第二个时期是随着深度学习的发展，越来越多的研究选择神经网络作为分类器。两个时期的工作，其心电图分类系统都可以分为三个步骤：心拍切分、特征提取、分类^[1]。接下来，本文将分别从这三个方面阐述研究现状。

心拍切分。现有的心拍切分方法的基本原理是，以 R 波尖峰为参照点，在参照点两边分别取一个固定的长度，以此来表示一个心拍的长度。文献[2]首先对心电信号做下采样，然后取 R 波尖峰前 61 个点和 R 波尖峰后 38 个点，总计 100 个点，用来表示一个心拍。文献[3]同样先对心电信号时间序列做下采样，然后在 R 波尖峰两边各取 64 个点。除此之外，文献[4]没有对原始心电信号时间序列做下采样，而是直接以 R 波尖峰为中心，在其前后各取 128 个点来表达一个心拍。文献[5]在 R 波尖峰前取 100 个点，R 波尖峰后取 200 个点。其他很多文献也是用诸如此类的方法来切分心拍，即以 R 波尖峰为一个参照点，取一定长度用以表示一个心拍，本文称之为固定长度的心拍切分方法。这种方法忽略了心拍长度随时间、场景或个体而变化，使得切分得到的心拍残缺或者冗余，影响特征提取和分类的效果。

特征提取。在特征表达上，现有工作的特征大致可以分为两类，一类是手动特征，另一类是深度特征。手动特征主要可以通过以下方法获取，包括傅里叶变换^[6]，小波变换^[7]，埃尔米特函数变换(Hermitian function)^[8]，时域幅值与时长^[9]，以及形态学知识^[10]。深度特征主要通过神经网络获取，例如卷积神经网络^[4]，递归神经网络^[11]，但是目前大多数特征提取都是针对单个心拍级别的，这样的特征在尺度上过于单一，信息损失过大。

分类。分类器的选择是多种多样的，但总的来说，可以分为两大类，第一类是传统的机器学习分类器，例如随机森林^[12]，支持向量机^[13]，K 近邻分类器^[14]。另一类是神经网络分类器，常用的神经网络分类器主要包括：人工神经网络^[15]，

卷积神经网络^[16,17]，递归神经网络^[11]，长短期记忆网络^[18]，以及基于块的神经网络^[19]。但是，现有工作仅仅是对神经网络一个简单的应用，忽略了基本的神经网络对特征重要度的学习能力不足。除此之外，分类器所针对的对象也有所不同，以文献[3,17]为例的研究，针对每一个个体，训练了一个分类器，但是以文献[4]为例的研究，对所有个体训练了一个分类器。虽然对每个个体训练一个分类器，对这个个体而言，会得到一个更高的分类准确率，进而使得总体的分类准确率也会更高。但是，这意味着，如果我们有一个新的个体，首先要收集这个个体大量的心电数据，然后训练，接着才可以预测。但是这样的问题对于文献[4]所训练的分类器而言，显然是不存在的。因此，对所有个体训练一个分类器的实际意义更大，所以，本文的研究会对所有的个体训练一个分类器。

如上所述，现有工作的心电类型识别系统多是对心拍做固定的切分，忽略心拍长度随时间、场景或个人而改变。同时，提取特征的尺度过于单一，特征没有结合医学知识和医生的经验。除此之外，分类器也没有经过精心设计，只是选用了一些常用的分类器。本文打算动态的切分心拍，并且提取多尺度的特征，最后利用改进后的卷积神经网络进行分类。

1.2.2 心电个体识别研究现状

类似于类型识别工作，心电个体识别的工作也可以根据分类器的不同大致可以分为两类，第一类是利用手动特征和传统机器学习分类器进行分类。第二类是利用深度学习进行特征提取和分类。整个心电个体识别系统可以分为三个部分，分别是样本生成、特征提取和分类，但是在样本生成之前，数据获取也是一个值得讨论的问题。数据获取指的是获取用来训练模型的数据，多为具有一段长度，含有多个心拍的心电记录。接下来，将从四个方面阐述心电个体识别研究现状。

数据获取。现有工作数据获取主要通过两个途径，第一个途径是自己收集数据构建自己的心电个体识别数据库，如文献[20]收集了 11 个身体健康的志愿者的数据，并且每个志愿者有多条记录，文献[21]构建了一个含有 13 个健康个体的数据库，文献[22]构建了含有 10 个健康个体的数据库。第二个途径是利用公开数据集，比如 MIT-BIH 心律异常数据库，PTB 心电数据库，ECG-ID 心电数据库。文献[23]利用 MIT-BIH 心律异常数据库上的所有个体（47 人）验证了他们的算法，文献[24]选用了 MIT-BIH 心律异常数据库上含有导联 I 的 33 个个体，文献[25]在 MIT-BIH 上选用了 10 个健康个体，而文献[26]选用了 10 个心律异常的个体；文献[27]在 PTB 心电数据库上选用了两个子数据库，分别是 52 个健康的个体和 48 个不健康的个体，文献[28]则在 PTB 上选取了 60 个个体，文献[29]选用了 ECG-ID

数据库上的 90 个个体。可以看出，自建数据库大都选用健康的个体，并且数据库总人数较少。而公开数据集具有更多的个体，并且有些公开数据集还含有心律异常的个体。

样本生成。现有绝大多数的样本生成方法都可称为基准的样本生成方法，即把一个心拍、类似一个心拍的一个心动周期（如一个 RR 间隔）、或部分心动周期（如一个 QRS 波群）作为一个样本。如文献[29-32]都是把一个心拍切分成一个样本，文献[26,33,34]都把一个 RR 间隔切分成一个样本，文献[35]则把一个 QRS 波群切分成一个样本。在众多工作中，只有文献[36]提出了一种非基准的方法抽取样本，即在一个心电记录上，随机抽样一个指定长度的子心电记录作为一个样本。基准方法受心电记录长短的限制，无法获取大量的样本，若采用神经网络作为分类器，很容易由于样本不足，造成模型无法得到充分的训练。

特征提取。根据分类器的不同，特征提取方法也分为两种，一种是手动提取的特征，文献[37]根据检测到 PQRSST 波峰提取了 17 个手动特征，文献[38]利用经验模式分解获取了大量手动特征，还有其他常用方法如小波变换^[39,40]、傅里叶变换^[41]、形态学^[42,43]、主成分分析^[24,30]、以及数据编码^[44]。另外一种则是利用深度学习提取特征，如文献[31,35,36]都是利用神经网络提取特征，省去了大量的特征工程，并且能够得到高维度和高抽象的特征。虽然特征提取的方法很多，对于不同方法获得的样本，不同特征提取方法得到的特征的分度也有所区别，需要依据样本选择合适的特征提取方法。

分类。分类器可分为两类，一类是传统机器学习分类器，如 K 近邻分类器^[33,34]、支持向量机^[45]、线性分类器^[43]、随机森林^[46]等。另一类是深度学习分类器，如深度神经网络^[31]、反向传播神经网络^[47]、卷积神经网络^[48]等。而基本神经网络分类器的通病就是对特征重要度提取能力的不足。

现有工作的数据来源混乱，尤其是自建数据库，不仅含有个体都很少，数据质量也有限，缺乏实际意义，而大多数选用公开数据库的研究，也选用了健康的个体，算法在心律异常个体上的性能尚未验证。同时，提取特征的基础大多是基于一个心拍或者基于一个 RR 间隔的，即一个样本是一个心拍或者一个 RR 间隔，由于一条记录的心拍数量十分有限，这导致了样本数也较少，而过少的样本会使得模型欠拟合。除此之外，虽然深度学习在心电识别领域应用广泛，但是现有工作只是简单利用深度学习，并未设计出一个针对心电个体识别问题的更好的神经网络。本文选择在公开的含有心律异常个体的数据集上验证算法，同时在样本生成上，采用了一种随机抽样的方法，这样就可以生成大量样本，使得神经网络可以得到充分的训练。同时，引入通道注意力模块，搭建一个更适合心电个体识别问题的神经网络。

1.3 本论文的主要研究内容及贡献

本文主要研究基于深度学习的心电图类型识别和个体识别问题。从 1.2 节研究现状可以看出，两种识别研究具有很高的相关性，主要体现如下：(1) 在样本生成上，都需要从一整条心电记录中通过某种方法获取训练样本，不同的是，类型识别要求准确的切分心拍，而个体识别问题只需要一段心电记录即可；(2) 在特征提取上，很多特征提取方法都在两类识别工作上有应用，比如小波变换、形态学等，二者的区别在于提取什么样的特征，类型识别主要是提取不同心拍之间的差别，重点在于心拍的局部细节，而个体识别主要是提取不同人的心拍表现差别，重点在于心拍的整体信息，不重视是否有疾病等；(3) 在分类器的选择上，二者的选择是相似的，且最近的工作都倾向于选择神经网络。

针对 1.2 节提出的现有工作的局限性，结合上述的相似性，本文进行了大量研究和改善，具体分为类型识别和个体识别两个方面进行阐述。

关于心电图类型识别问题，具体的研究包括以下几点。

(1) 心电图中心拍的动态切分。根据心拍的长度会随场景、时间或者个体变化而变化的特点，提出对心拍进行动态的切分，同时，动态的切分子拍和多拍。除此之外，为了满足特征提取和分类器固定长度输入的要求，把动态的心拍重整成一个固定长度的心拍，也是研究的一个重点。

(2) 更为丰富的特征表达。心拍是由多个部分组成的，一个心拍类型也需要通过观察邻近的心拍才能确定，研究从多个尺度表达心拍对心拍分类尤为重要。除此之外，深度学习得到的抽象特征和手动提取的特征都很重要，抽象特征能更好的表达心拍，手动特征是医学知识的体现，二者互为弥补，研究二者结合的特征集合能有助于分类。

(3) 学习能力更强的深度学习分类器。为了弥补神经网络对特征重要度学习能力的不足，引入通道注意力模块，并针对本研究的问题对通道注意力模块进行改善，提高模型分类准确率。最终类型识别准确率达 96.94%，比现有相关工作约高 3%。

关于心电图个体识别的问题，具体的研究包括以下几点。

(1) 选择高质量的数据集。自建数据库质量较差，个体数较少，且都是健康的个体，算法对心律异常敏感。因此选择公开的且含有较多个体和大量心律异常的数据库，有助于研究泛化能力更强的识别算法。

(2) 非基准的样本生成方法。为了得到大量的样本，改善一种随机抽样方法，以固定长度而非心拍为基准生成样本。可以通过多次抽取，获得多个样本，满足神经网络分类器训练时的样本需求。

(3) 区分度高的特征表达。由于上述的样本生成方法会使得生成的样本规律性差，需要通过域转换来获得更好的数据表达，选择合适的域变换十分重要，通过对比常用的域变换方法，选择最适合样本的特征提取方法。

(4) 能力更强的深度学习分类器。和心电图分类问题相同，普通的神经网络对特征重要度的学习能力非常有限，限制了识别的准确率，通过引入和改善通道注意力模块，提升分类器的性能。最终模型个体识别准确率高达 95.86%，同样比现有相关工作高了约 3%。

1.4 本论文的组织结构

本论文的整体组织结构如下：

第二章详细介绍本研究所用到的相关技术和背景知识。医学知识部分包括心动周期的过程、心电图的组成部分和心电生物识别系统，深度学习部分包括深度学习理论、常用的网络以及常用的深度学习框架，同时，还将对本文用到的数据集进行简介。

第三章详细阐述提出的新的心电图类型识别系统。首先，将分析现有工作局限性以及可以改进的方向，接着详细解释本文提出的改进算法，然后利用实验验证提出来的算法的性能，并从多个角度与现有相关工作进行对比，考查的算法的性能。

第四章详细介绍提出的心电个体识别系统。首先，分析现有工作的不足和以及可改进方向，第二步是阐述本文提出的改进算法，第三步通过实验证明提出的算法的性能优越性，并与现有工作进行对比。

第五章是对全文的总结，阐述本文两方面工作的主要贡献，并对未来的研究工作进行展望，介绍可以继续研究的方向。

2 论文相关知识介绍

本章主要介绍论文中会使用到的一些理论知识、技术背景以及数据集。包括心电图相关医学知识简介，生物识别系统简介、深度学习简介以及数据集简介，最后对本章小结。

2.1 心电图简介

2.1.1 心拍简介

所谓心动周期，顾名思义就是心脏跳动的周期，指从当前心跳的起始到下一个心跳的起始，人体的心血管系统所经历的一个机械的周期过程，心脏的结构示意图如图 2-1 所示。

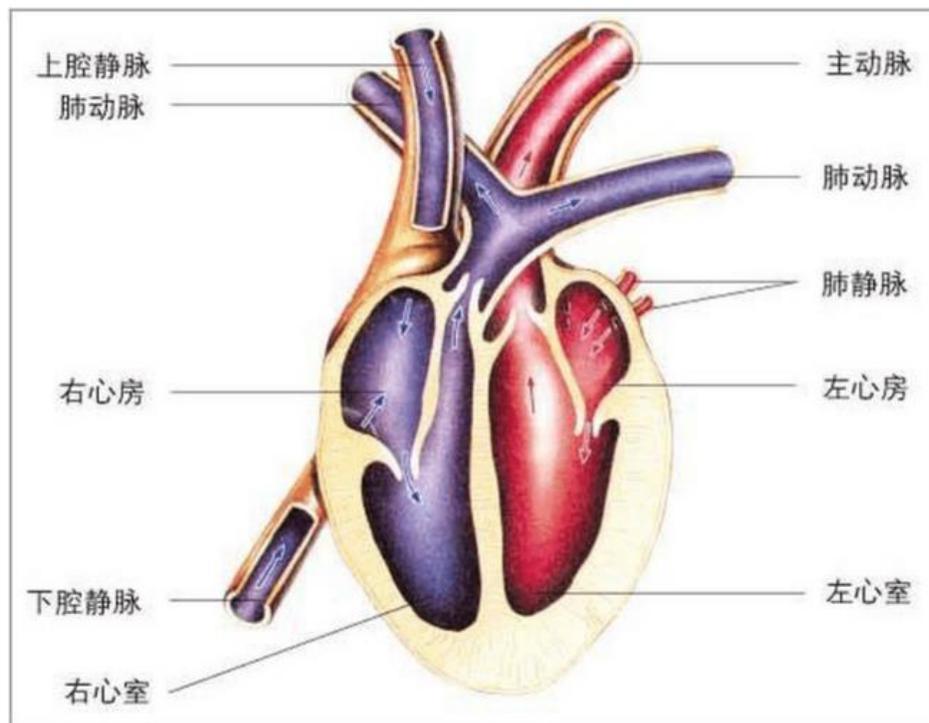


图 2-1 心脏结构示意图^[49]
Figure 2-1 Heart Structure Diagram^[49]

心动周期的主要过程如下：一开始，心脏舒张时会导致内压降低，接着，腔静脉血液就会回流入心。心脏收缩时就会导致内压升高，然后将血液泵到动脉。心脏的这个收缩和舒张一次所形成的行为就构成了一个所谓的心动周期。对于一个正常心动周期中，最先开始的是两个心房收缩，值得注意的是，其中右心房的

收缩略先于左心房。然后心房开始舒张后，两个心室就收缩，与之前不同的是，左心室的收缩会略先于右心室。在心室舒张的后期心房又开始收缩。

心电图（electrocardiogram，简称 ECG）是使用称为心电图机的设备来记录一段时间内，身体表面上随时间变化的电压。心电图一般记录的是由多个心动周期组成的一段身体表面电压变化的曲线。通过该方法获得的信号一般是含有大量噪声的，并且收集的心电图信号的噪声可以被分类为以下类型的干扰的噪声^[50]。

(1) 肌电干扰，除了 ECG 信号之外，许多其他生物信号存在于人体表面，多种生物电信号混合在一起，由于其一般是快速且无规律的变化，所以可以把肌电干扰理解为高斯白噪声。

(2) 基线漂移，即在 ECG 测量期间测量的基线随时间缓慢变化。这相当于在整个心电图测量中，给曲线增加或减少一个强度随着时间变化的直流分量。基线漂移是所有干扰中最严重的，基线漂移的频率通常小于 1Hz，能量主要集中在 0.1Hz。

(3) 工频干扰，是指由于我们使用到的心电图机接的是市电系统的电，加之整个电磁环境都是在市电电力系统中，使得我们的整个工作系统都在市电系统的干扰之中。我国市电频率在 50Hz，所以工频干扰也集中在 50Hz。

在图 2-2 中，画出了 MIT-BIH 心律失常数据库上某个病人的一部分原始数据，图 2-3 是滤波之后的数据，可以明显的看到，肌电干扰、工频干扰和基线漂移都已经被滤掉了，曲线也变得更加平滑。

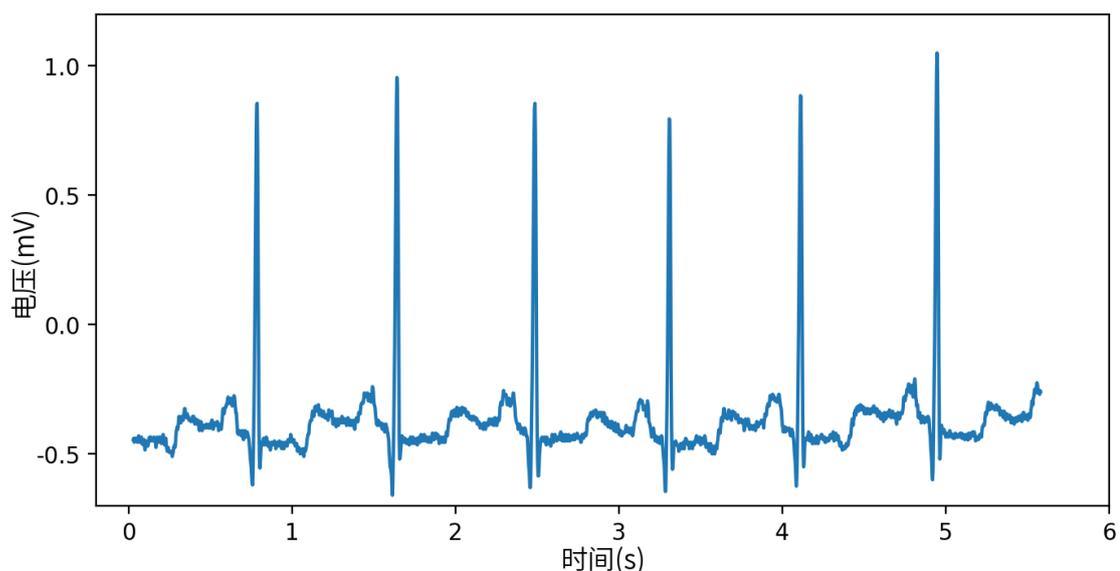


图 2-2 原始心电图记录
Figure 2-2 Original ECG Record

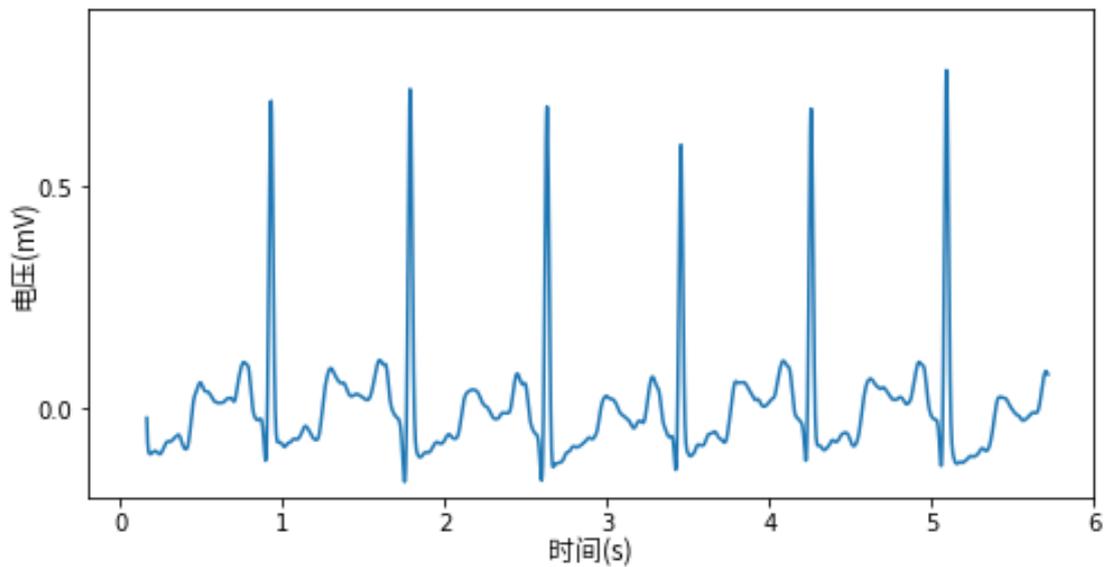


图 2-3 滤波之后的心电图记录
Figure 2-3 ECG Record After Filtering

本文把一个心动周期所产生的生物心电信号称为一个心拍。一个完整的心拍如图 2-4 所示，其主要由一个 P 波，一个 QRS 波群和一个 T 波组成。下面，将简要介绍各个波、波段。

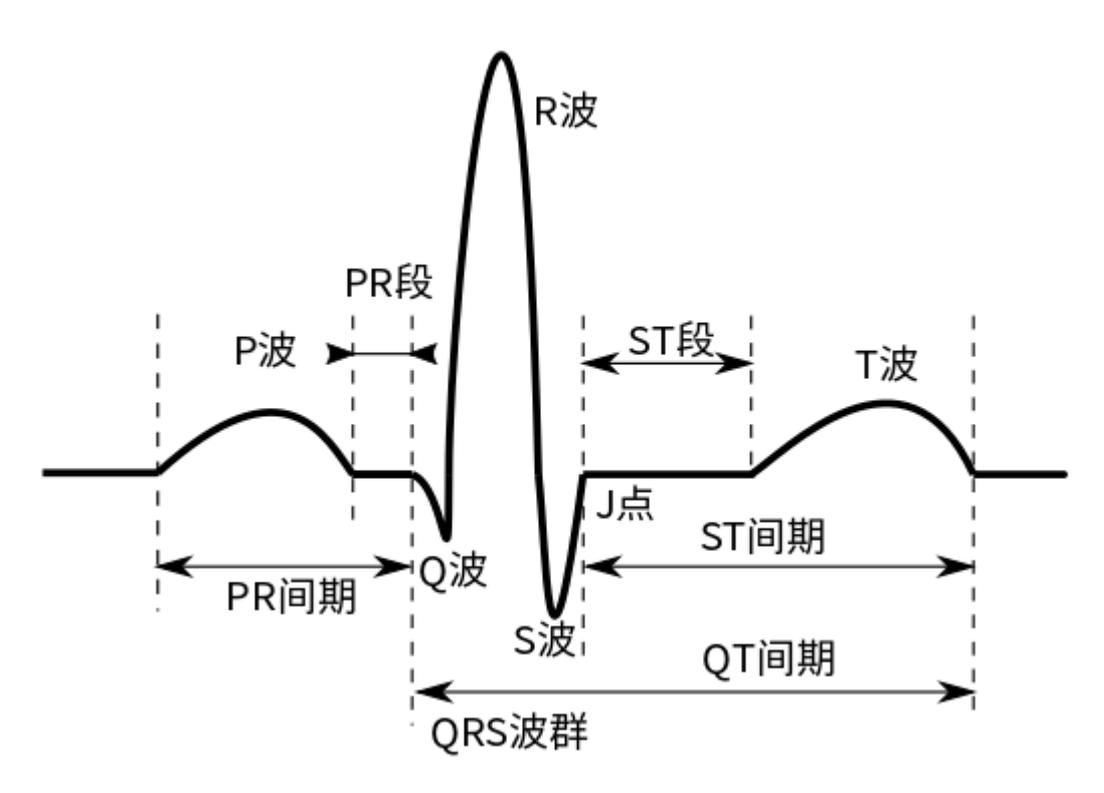


图 2-4 单个心拍示意图
Figure 2-4 Single Beat Diagram

(1) P 波: P 波的起始也就是一个心拍的起始。窦房结是正常的的心脏电激动开始的地方。然后,从右心房通过房室束再传导到左心房,这就形成了心电图上的 P 波。对于正常的 P 波,其时长和幅度都有限制,一般其时长不超过 0.12 秒,其高度不超过 0.25 毫伏。正常的 P 波表现为较为光滑的单峰波形,但是当心房扩大,两房之间的传导出现异常时,P 波的波形也会出现异常,主要可表现为具有高尖峰或者是具有两个峰的 P 波。

(2) PR 间期:指的是从 P 波的起点到 QRS 波起点之间的时长,也就是从心房除极开始到心室除极开始之间的时长。PR 间期激动沿着前中后结间束传导到房室结。由于房室结的传导速度相对缓慢,这也就形成了图 2-4 中的 PR 段。一个正常的 PR 间期的时长大约在 0.12~0.2 秒,其时长和年龄有一定关系,一般表现为年龄越大的个体,相应也会有越大的 PR 间期时长。如果心房传导到心室的这一段出现了阻滞,一般表现为 PR 间期的延长或者 P 波之后的心室波的消失。

(3) QRS 波群:QRS 波群是整个心拍过程中幅度最大,特征最明显的一个波群。激动向下经希氏束、左右束支同步激动左右心室形成 QRS 波群。QRS 波群的波形代表着心室的除极过程,其主要由三个波组成,一个向下的振幅较小的 Q 波,一个向上的幅度最大的 R 波,以及一个向下的振幅较小的 S 波。一个正常的 QRS 波群的长度不超过 0.11 秒。但是当出现心脏左束支或者右束支传导阻滞,或心室扩大或肥厚等情况,往往就会出现 QRS 波群的增宽,变形以及时长变长的情况。

(4) J 点:在 QRS 波结束与 ST 段开始的交点,该处被称为 J 点,其代表心室肌细胞全部除极完毕。

(5) ST 段:该段指的是 QRS 波群的结束到 T 波开始之前的那一段。对应心室肌细胞全部除极完毕到复极尚未开始的这一段时间。处在此时的各部位心室肌都处于除极的状态,细胞之间还没有电位差。一般情况下,这一段的波形是处于电位线上的,其时长一般在 0.05~0.15 秒的范围内。当出现一些异常情况,比如缺血或者坏死的情况出现在了某些部位的心肌,那么在心室除极完毕时,ST 段仍存在电位差,对应到心电图上,就是 ST 段会发现偏移,不在等电位线上。

(6) T 波:T 波代表的是心室的复极,其幅度相较于 QRS 波群较小,但是相较于 P 波较大,其时长较长。值得注意的是,在同一导联上,T 波的方向应该与 QRS 波群的主方向一致。在心电图上,T 波的改变容易受到多种因素的影响,例如低平或者倒置的 T 波可能是由于心肌缺血导致的,高尖的 T 波可能是高血钾、急性心肌梗死导致的。医学专家分析 T 波主要观察其形态与振幅。

(7) U 波:U 波不是一个一定出现的波段,如果有,其一般跟随在 T 波之后。目前对于 U 波的产生机制以及其作用还没有一个完全明确的解释,目前比较认可的一种说法是 U 波极可能与心室的复极有关,同时,U 波与某些疾病的诊断也有

一定的关系。

(8) QT 间期：该间期代表了从心室除极到复极的时间。在心电图上表示着从 QRS 波群的开始到 T 波的结束，一个正常的 QT 间期时长大约为 0.44 秒，但是由于心率的变化，会导致 QT 间期也有变化，QT 间期一般与心率成反相关。为了消除 QT 间期受心率的影响，因此引入了矫正的 QT 间期 (QTC)，其中的一种计算方法如公式(2-1)所示，其中 RR 表示 RR 间期或者 RR 间隔。QT 间期的延长，往往与恶性心律失常有关。

$$QTC = \frac{QT}{\sqrt{RR}} \quad (2-1)$$

(9) RR 间隔：RR 间隔指的是从当前 R 波尖峰到下一个心拍 R 波尖峰的时长，一般也认为这就是一个心动周期的长度，也认为这就是一个心拍的长度。因为 R 波的幅度较高，容易用导数法检测，现在医学上计算心率的方法大都是基于计算 RR 间隔的距离去计算的。

总而言之，通过以上的介绍和分析，可以了解到一个心动周期的组成以及各个时期的特点，这对进行心电类别识别以及心电个体识别工作都有着十分重要的意义，也为改善现有各种心电相关的智能系统提供了理论基础，尤其是医学知识和常识的介绍，为本研究数据提供了一个专业的角度，这对引入一些医生们赖以信任的特征提供了非常大的帮助。在后文中，也将利用这些医学知识来弥补分类器专业知识不足的问题。

2.1.2 常见异常心拍简介

心拍异常，在医学上又称为心律异常。心律异常的原因有很多种，一般认为由于窦房结激动异常或者激动产生于窦房结以外，激动的传导缓慢、阻滞或者经由一些异常通道传导，就是心脏活动的起源或者传导障碍导致心脏搏动的频率或者节律异常。心律异常是心血管疾病中重要的组成部分，它既可以单独发病，又可以与其他心血管疾病共同发作。是非常棘手和难缠的疾病，严重时可危及人们的生命财产，是当今非传染病异常死亡的一个重大原因。心拍的异常可分为很多种类，大大小小的可分为几百种，但是在实际之中，分为这么多种是没有意义的，况且无论是对医生还是患者，这也是无法记住的。因此，有必要简化心拍分类的标准。

在国外，美国有一个协会叫做美国医疗仪器促进协会（英文全称 The Association for the Advancement of Medical Instrumentation，简称 AAMI），成员包

括临床和生物医学工程师、技师，医师，护士和医院管理人员，教育工作者和研究人员，制造商，分销商，政府代表及其他医护专业人员。他们关于心拍的分类提出了一种五分类的标准，主要包括正常心拍（N），室上性异位心拍（S），室性异位心拍（V），融合心拍（F），不可分类心拍（Q）。这个标准在学术界的应用非常广泛，尤其是在心拍正异常多分类上。

在国内临床医学上，把常见的心电图分为 11 类，称之为 11 种常见心电图，主要是：正常心电图，窦性心动过缓，窦性心动过速，房室传导阻滞，房早与室早，心房扑动，房颤，室速，室扑，室颤，心肌梗死。这些疾病有的可以直接根据心电图一次性诊断出来，有的可能需要结合以前的心电图或者病人疾病史。有的可能不会危及生命，有的可能导致病人的死亡。辨别这些疾病需要大量的医学知识，也需要医生长年累月的经验积累。

2.2 生物识别系统简介

2.2.1 常见生物识别系统

所谓生物识别系统指的是利用生物识别技术来判别不同的个人的鉴别系统，其核心内容就是生物识别技术。生物识别技术是指通过计算结合其他物理学（比如声学，光学），再通过生物传感器以及生物统计学原理等各种生物学技术，利用人体等个体具有的固有的唯一的生理特征（例如虹膜、人脸、声音、指纹等）和行为特征进行个体身份识别与鉴定。生物识别系统是一种软硬件结合，各种高科技手段结合的个人身份识别系统，具有广泛的应用前景。

比较成熟且较为广泛使用的生物识别系统有：指纹识别、掌纹识别、签名识别、声音识别、虹膜识别和人脸识别。

指纹识别。指纹识别是目前最为成熟的生物识别系统，其优点在于其高可靠性、方便与便于被接受，设备价格低廉，小巧方便，这也是指纹识别占据大部分市场的主要原因。缺点是人们容易在接触各种物品的时候留下指纹，这也造成了指纹数据容易被盗取的特点。

掌纹识别。掌纹类似于指纹，每个人的手的形状都是不同的，掌纹也都是不同的。但是相比于其他生物识别应用，掌纹识别很难大规模应用，因为不同个体之间的掌纹差距不是很大，并且由于人的变化和和设备原因，很难获取到一个高精度的特征，这阻碍了掌纹识别的大规模推广。

签名识别。优点在于简单实用，易于辨识。缺点是人写的字会根据场景的不同发生变化，同时，字体也容易遭到别人模仿，安全性较低。

声音识别。声音识别也是一种较为传统的识别方式，但是一直没有得到广泛的使用。原因主要是一方面其安全性较低，便于模仿和收集，另一方面是其稳定性较低，普通人可能由于感冒而导致声音发生变化，致使系统无法辨认该个体。

虹膜识别。虹膜识别的优点是具有很高的精度。但缺点是对于某些特定的人（比如某些类别的盲人没有虹膜），无法使用，加上采集设备的极其昂贵的价格，阻碍了其广泛的应用。

人脸识别。人脸识别的优点是设备成本低廉，在特定的前提下识别率较高。但是人脸识别也存在着一些缺点，是否佩戴眼镜，往往会被判别成两个人，不成熟的人脸识别技术也无法在灯光昏暗时使用，而红外摄像头的成本又较高。除此之外，人脸日常暴露在外，其安全性较差，已多次发生过利用 3D 打印技术复制人脸，解锁个人账户的案例了。

随着人们对数据安全和隐私保护的需求日益增长，个人识别系统的重要性与日俱增，这其中，由于生物识别系统具有很强的唯一性与不可复制性，有较高的安全性，在实际生活中得到了广泛的应用，较为成熟的生物识别系统有指纹识别系统、人脸识别系统、虹膜识别系统等，这些已经商用的系统不仅有着很高的识别率，检测也很便捷。但是，由于技术的发展，一些不法分子利用高科技来复制指纹，人脸和虹膜等生物特征，使得这些生物识别系统的安全性不断受到挑战，因此急需发展一种高安全性和不可复制性的生物识别系统，心电生物识别系统就是这样的一种可发展的技术。

2.2.2 心电生物识别系统的优缺点

与传统的生物识别系统（例如指纹识别，人脸识别等）相比，心电识别系统主要有以下几点优势：(1) 高可获取性，任何活着的人都会有自己的心跳，而指纹、虹膜，人声对于某些特定的人群，可能并不具备该特征，但心电数据没有这些问题，获取更为方便；(2) 高安全性，心电数据私密性较高，类比于指纹和人脸，心电很少“暴露”在外，从源头上短了被盗用的可能性；(3) 低可复制性，相比于其他生物识别所具有的稳定性的特点，即其特征不随时间变化，心电特征随时间和环境有较大变化，造成了其特征具有“不稳定”的特点，降低了其可复制性；(4) 低检测费用，相比于虹膜和指纹识别，心电的检测费用较低，不需要大量的高科技仪器来收集数据，为其广泛应用奠定了经济基础。以上的这些优点，显示了心电生物识别存在和发展的意义，为生物识别家族添加了新的可选方向，也是目前研究的热点所在。

但是，与其他生物识别系统相比，心电识别系统也有以下的几点缺陷：(1) 项

对较低的准确率，由于心电特征的不稳定性，使得在判别时会带来不小的难度，从目前的工作来看，其识别准确率相比于指纹系统较低；(2) 难以大规模应用，相比于其他的生物识别系统，心电识别系统的总数目目前仍有较大的限制，这成为其实用化道路上的一个阻碍；(3) 公开的数据库较少，由于心电数据的私密性，目前公开的数据库，尤其是包含个体数较多的公开数据集较少，使得研究可用的数据选择较少。以上这些问题是阻止心电识别系统实用的障碍，解决这些问题，对心电识别系统的实用化有很大帮助。

表 2-1 中，本文总结了常用的生物识别系统在多个维度上的比较，从表中可以看出，心电生物识别系统在安全性和数据可获取性上比较高，同时成本较低，但是准确性仍需提升。

表 2-1 常见生物识别系统的比较
Table 2-1 Comparison of Common Biometric Systems

生物信号	心电	人脸	虹膜	指纹	声音
准确性	中	高	高	高	低
安全性	高	低	中	低	低
可获取性	高	高	中	中	低
可复制性	低	高	高	高	高
检测花费	低	低	大	中	低
应用规模	中	大	大	大	低
公开数据库	少	多	中	多	多

2.3 深度学习简介

2.3.1 深度学习概念简介

深度学习这个概念自从 2006 年被正式提出以后，在过去的 13 年间，有了巨大的发展和改进，给我们的生活带来了巨大的影响。关于深度学习的概念没有明确的定义。一般认为，深度学习的起源来自于多层感知机。最简单的感知机就是一个只有两层的神经网络，即一个输入层和一个输出层。在图 2-5 中，展示了带有隐层的多层感知机，图中的网络结构具有两个隐层，一般带有至少一个隐层的多层感知机就可以称之为深度神经网络，也就是深度学习的一个实例。一般的，随着网络层数增多，所提取到的特征也更为抽象。在图 2-5 中，每一层网络都有一定数量的神经元组成，这些神经元的基本结构都是一致的，其计算公式可用公式(2-2)

表示。

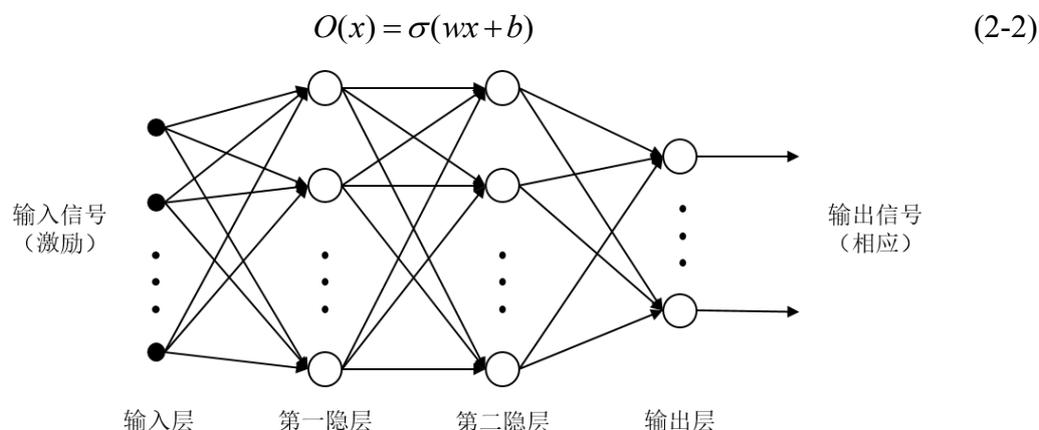


图 2-5 多层感知机^[51]
Figure 2-5 Multilayer Perceptron(MLP)^[51]

在公式(2-2)中, x 表示一个神经元的输入, w, b 分别是需要学习的权重和常量, 一般也可把 b 看成是 w 的一维。其中 σ 是激活函数, 是一个非线性的函数, 常用的非线性函数有 sigmoid 函数, tanh 函数, relu 函数等。激活函数, 也就是非线性函数, 是神经网络中非常重要的一个组成部分, 如果没有激活函数, 神经网络只能是一个简单地线性网络, 而非线性函数让神经网络的表达能力大大的增强了, 能拟合更多种的分类空间的划分。

深度学习发展较为长久, 一般认为到目前为止, 深度学习的发展有三个浪潮, 分别是 20 世纪 40 年代到 60 年代深度学习的雏形出现在控制论中, 20 世纪 80 年代到 90 年代深度学习以联结主义 (connectionism) 为代表, 而从 2006 开始, 随着 GPU 的发展, 大大提高了运算能力, 也开始以深度学习的名义走上了“复兴之路”^[52]。

2.3.2 常用神经网络简介

神经网络的类型多种多样, 但是经常使用的主要分为卷积神经网络 (CNN), 递归神经网络 (RNN) 和全连接网络 (也可称为多层感知机, MLP)。

卷积神经网络, 指的是一种专门用来处理具有网格结构特征的数据的网络。常见的具有网格特征数据有一维的时间序列, 二维的图像以及三维的视频数据。卷积神经网络在诸多应用领域都有着非常优异的表现, 尤其是对这种网格结构的数据。

卷积神经网络里的一种基本运算就是卷积运算, 也就是信号处理里面的卷积运算, 卷积是对两个实变函数的一种数学运算。对于连续函数, 公式定义如公式(2-3)

所示。

$$s(t) = \int x(a)w(t-a)da \quad (2-3)$$

其中 $x(a), w(a)$ 表示是两个实变的连续函数。通常卷积运算有一个通用的符号表示，即“*”，所以公式(2-3)也可以表达成公式(2-4)的形式。

$$s(t) = (x * w)(t) \quad (2-4)$$

如果两个函数是离散函数，则上面的积分符号会变成求和符号，对于两个一维离散函数，其卷积计算如公式(2-5)所示。

$$s[k] = \sum_{a=-\infty}^{\infty} x[a]w[k-a] \quad (2-5)$$

如果被卷积的两个函数是二维的，其计算如公式(2-6)所示。

$$s[i, j] = (X * W)[i, j] = \sum_m \sum_n X[m, n]W[i-m, j-n] \quad (2-6)$$

因为卷积满足交换律，所以公式(2-6)也可以写成公式(2-7)的形式。

$$s[i, j] = (X * W)[i, j] = \sum_m \sum_n X[i-m, j-n]W[m, n] \quad (2-7)$$

虽然理论上说，卷积神经网络最重要的一种运算是卷积运算，但是在实际应用中，很多神经网络的库函数实现的并不是卷积运算，而是互相关运算，因为这两者的差别只是卷积运算对卷积核进行了翻转，但是互相关运算没有，本质上是没区别的，并且互相换运算减少了一个翻转操作，所以互相关运算应用的更为普遍。公式(2-8)就是一个互相关运算。

$$s[i, j] = (X * W)[i, j] = \sum_m \sum_n X[m, n]W[i+m, j+n] \quad (2-8)$$

从公式(2-8)中可以看出，虽然很多神经网络的库实现的是互相关函数，但是仍然叫做卷积，本文也把这种互相关运算简称为卷积。

在图 2-6 中，本文演示了一个卷积核和输入函数实现卷积的过程。该卷积是一个二维卷积的过程，并且设置了只当卷积核完全与图像重叠的时候才进行卷积（也就是本文所说的“有效”卷积），还把步长设置为了 1（即卷积一次，卷积核移动一个方格）。从图中可以看到，离散卷积，也可以理解为一种的矩阵的乘法，对应位置元素相乘相加，得到新矩阵的一个元素。然后卷积核向右平移一个单位，继续重复上述操作。当卷积核到了最右的位置，下一步就是把卷积核挪到下一行的最左边。最后得到的输出特征计算如图 2-6 最右边所示。

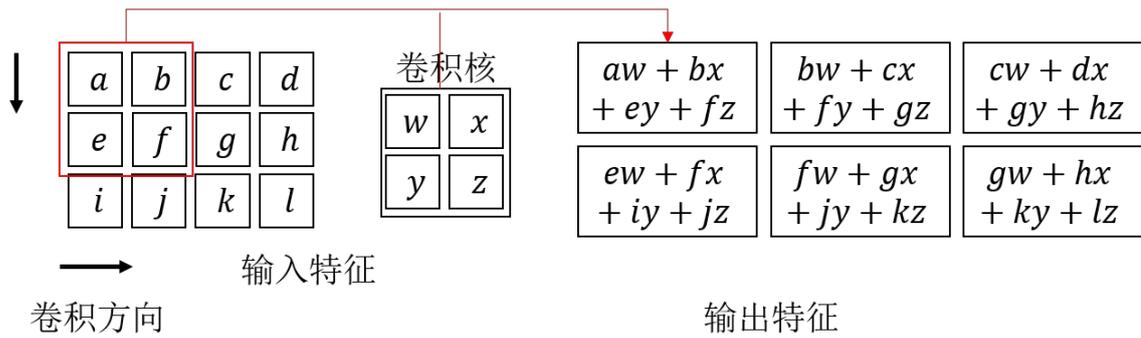


图 2-6 卷积运算过程^[52]
Figure 2-6 Convolution Operation^[52]

循环神经网络指的是专门用来处理序列数据的神经网络，主要是处理类似于 x_1, x_2, \dots, x_t 的数据，像文本数据就是一种序列数据。循环神经网络的特点就是网络结构中具有自循环的结构。递归神经网络是循环神经网络的一种推广和发展，其将链状结构改为了树状结构。如图 2-7 所示，左边是一种循环神经网络的机构，右边是一种递归神经网络的结构。再循环神经网络中，黑色的方块可以认为是一种自循环的计算，一般也是一种门结构，用来控制循环部分的输出。在递归神经网络中， V, U, W 都是权重矩阵，最后的 o 代表输出部分。

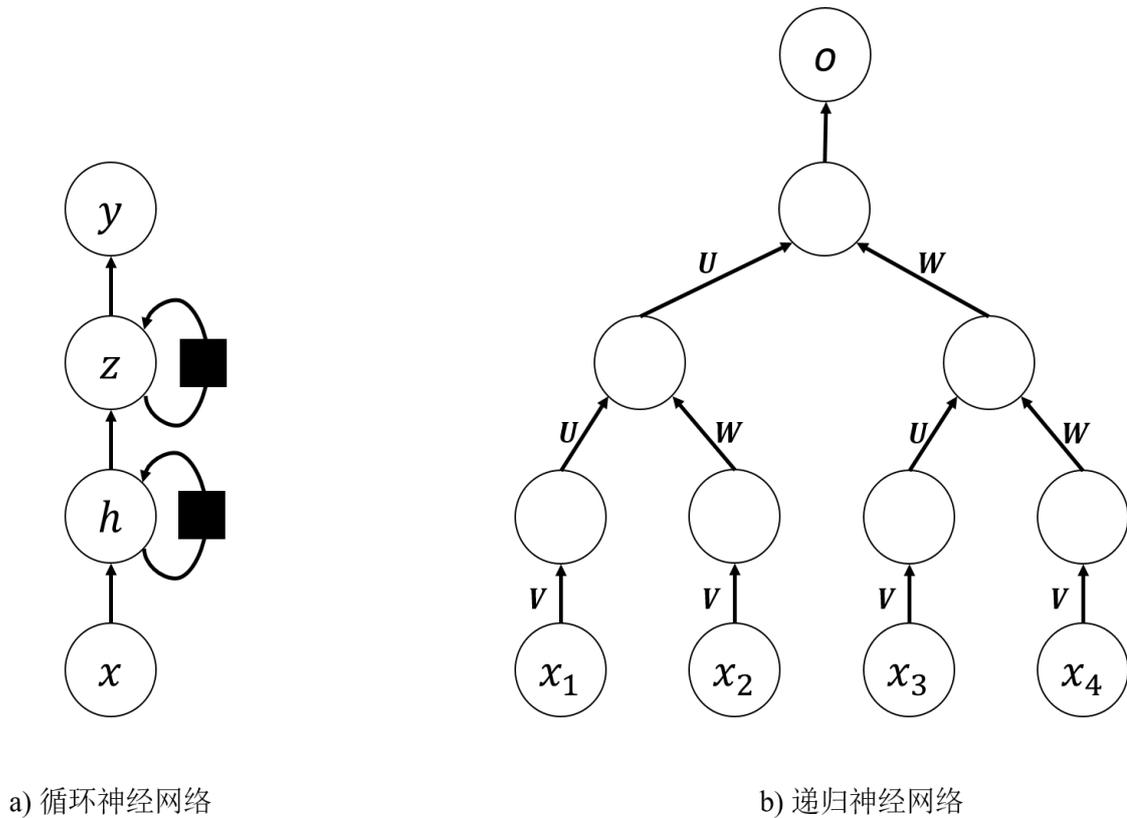


图 2-7 递归神经网络^[52]
Figure 2-7 Recurrent neural network^[52]

在序列模型中，比较经典的应用就是门控 RNN（又称 gated RNN）包括基于长短期记忆（long short-term memory）和基于门控循环单元（gated recurrent unit）的网络。以长短期记忆网络的神经元为例，在图 2-8 中本文画出了长短期记忆的内部结构。长短期记忆的特点是具有输入门、遗忘门和输出门，这使得它能记住较长时间上的信息，其自循环的权重是根据上下文而变化的，而不是固定的，这是其很重要的一个特点。长短期记忆网络在各种语言模型中获得了巨大的成功，是最为典型的一种序列模型。由于序列模型有自循环的模块，这导致了其计算量比起卷积神经网络要大得多，在实际应用中，会发现序列模型训练所花费的时间往往要比同等大小的卷积神经网络大得多。

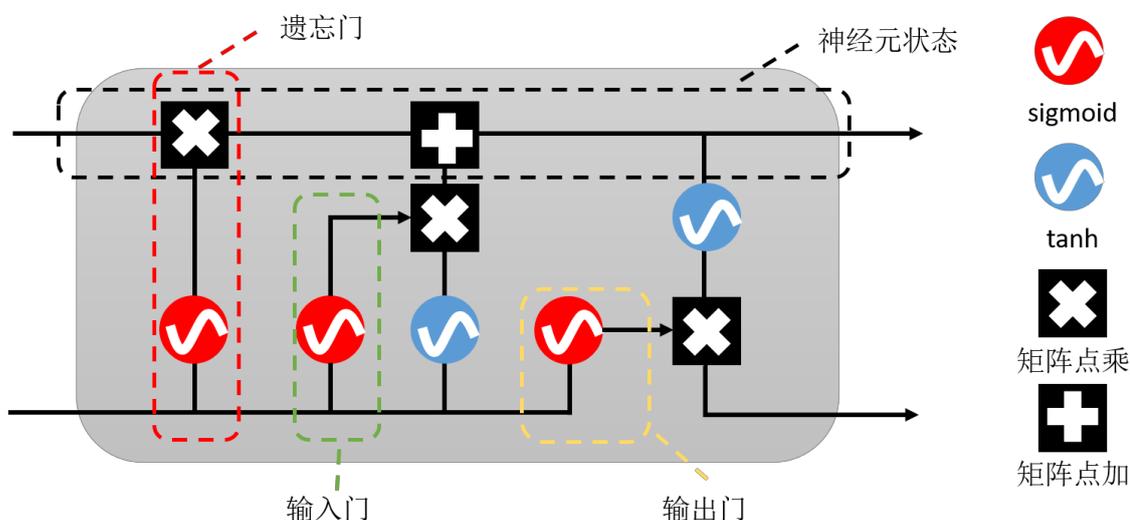


图 2-8 长短期记忆的结构

Figure 2-8 Structure of Long Short-Term Memory

除了卷积神经网络和序列模型，还有一种发展更早更基础的神经网络，一般称为全连接网络，也就是多层感知机。在深度学习中，卷积神经网络和序列模型中一般都会包含全连接网络。全连接网络一般被用来做分类器或者回归网络，而卷积模块和序列模块都是为了提取到合适的特征，以让全连接网络拥有足够好的分类或者回归的特征。

2.3.3 深度学习框架简介

近些年，随着深度学习和 GPU 的快速发展，深度学习框架也得到了大量的发展，目前市场上的深度学习框架多种多样，以 python 语言的编程库为例，比较主流的有 TensorFlow、PyTorch、Keras、Theano、Caffe 等等。接下来将简要介绍各个深度学习框架。

TensorFlow。该框架最初是谷歌的 Google Brain Team 研发出来的，最初的 TensorFlow 更像是一个数据计算流图。在实现计算前，必须定义好一个完整的计算图（也称为静态计算图），之后才可以进行计算。不过，后来版本的 TensorFlow 也可以在只有局部图的情况下进行计算。其最大的优点是社区比较完善。但是 TensorFlow 需要编写大量的代码，耗费太多的时间。

PyTorch。这是 FaceBook 开发的，与 python 具有很强相融性的张量计算和动态神经网络的框架。PyTorch 一开始以其代码量少，具有动态计算图（该特点也直接导致了 TensorFlow 开发了动态计算图）的特点而著名，在市场上快速占领了一席之地。但是由于其开发时间较晚，社区相对于 TensorFlow 较为单调。

Keras。其定位是简化用户编程的过程，提供一些常用的功能，其后端一般为 TensorFlow 或者 Theano。在 Keras 中，很多常用功能都已经被编写成一个类了，用的时候直接调用就行，这使得开发者的代码量大大减少。但是也有一定的局限性，因为当开发者想实现一些 Keras 不具备的功能时，就需要去了解其底层代码。鉴于其定位与功能与本文研究较为符合，本文采用 Keras 实现深度学习算法。

Theano。这是比较老牌的深度学习框架，可以高效的处理用户定义、优化以及计算有关多维数组的数学表达式。但是，Theano 早期不支持多块 GPU 的并行运算，这限制了其发展，另外一方面，由于和 TensorFlow 功能的高重合度，使得大部分市场被 TensorFlow 侵占。目前，Theano 已经停止更新。

Caffe。也是最老牌的框架之一，起初仅仅关注计算机视觉，所以应用范围有限，但是在有限的范围内，其具有非常好的通用性，这也是其被广泛应用的原因。类似于 Keras，Caffe 在灵活性上也有局限性，一旦想做一些改变，还需要自己大量的用 C++ 或者 CUDA 编程。除此之外，安装 Caffe 需要大量的依赖包，这对新手而言，是极其不友善的事情。

以上，本文介绍了一些常用的深度学习框架，并且简述了各种框架的优缺点，不同的框架各有特点，根据不同的任务可以选择不同的深度学习框架，本研究的重点是提出了一些新的心拍类型识别算法和心电个体识别系统的算法，深度学习只是本研究实现算法的方式，所以本文选择了 Keras 作为深度学习的框架，利用其代码量少，能快速搭建的特点。

2.4 数据集简介

为了验证提出的心电类型识别和心电个体识别算法，本文采用 MIT-BIH 心律失常数据库的心电数据进行验证，该数据库被广泛认为是最常用的标准心电数据库之一，采用该数据进行验证有以下几点优势：(1) 该数据库含有大量异常心拍，

不仅可以进行心电图类型识别算法验证，也可验证心电个体识别算法是否对心律异常敏感；(2) 该数据一共含有 47 个人，对于心电个体识别系统是非常合适的个体数；(3) 大量分类和识别问题都是在该数据库进行的，方便与其他算法比较性能。

MIT-BIH 数据库是由贝斯以色列医院心律失常实验室从 1975 到 1979 年收集获得的，由 48 条时长略长于半小时的心电记录组成，其中 60%来自住院患者。这 48 条记录从 47 个个体获取，其中男性 25 个，年龄从 32 岁到 89 岁，女性 22 个，年龄从 22 岁到 89 岁，其中记录号 201 和 202 来自同一个人。所有的记录都经过一个通带为 0.1-100Hz 的带通滤波器，并且采样频率为 360Hz。该数据库的标签有两个专家独立工作打标，之后进行一一比对，如果发现不一致，将进行协商并解决，并且之后经过多次修正。数据库中的心拍类别超过 40 类，但是很多类别的心拍只有很少的数量。在使用数据之前，需要对原始数据进行滤波。主要需要滤除的是肌电干扰，工频干扰和基线漂移，图 2-3 展示了其中某条记录部分数据的滤波结果。在两类工作中，对该数据库的使用略有不同，接下来将分开介绍。

对于心电类型识别系统，为了和其他工作保持一致，本文采用 AAMI 所提倡的标签类别，并把 MIT-BIH 数据库主要的 16 类转换为 AAMI 对应的 5 类，如表 2-2 所示。这五类标签分别是：正常心拍 (N)，室上性异位心拍 (S)，室性异位心拍 (V)，融合心拍 (F)，不可分类心拍 (Q)。根据 AAMI 的心拍分类标准，48 条记录中，有 4 条记录（记录号为 102,104,107,217）对于实验是没有意义的，故舍弃这 4 条记录。剩下的 44 条记录，采用标准肢体导联 II 的心电数据做实验。

对于心电个体识别系统，由于每条记录只有两个导联，并且各不相同，本文选择了含有记录数最多的导联来实验，即导联 II。由于记录号 102 和记录号 104 没有导联 II 的数据，为了保持数据一致，本文抛弃了这两条记录，剩下的 45 条记录就是本文的实验数据。因为 MIT-BIH 心率异常数据库含有大量的异常心拍，如果本文算法能够在该数据库上取得较好的结果，就说明本文算法对心律异常不敏感或者敏感度较低，具有较好的泛化能力。

2.5 本章小结

本章介绍了一些论文中所涉及的医学背景知识、生物识别系统的背景知识、深度学习背景知识和数据集。主要分为以下几点：(1) 心电图的形成和特点介绍；(2) 常见生物识别系统的优缺点的介绍；(3) 深度学习相关知识的介绍；(4) 对本文将要用到的数据库进行了简单地介绍。这些原理与技术是进行研究必备的背景知识，是本文工作理论与技术支持，只有在了解了这些之后，才能更高效的进行后续的研究工作，这些具备十分重要的实际意义。

表 2-2 MIT-BIH 数据库标签转 AAMI 标签
Table 2-2 MIT-BIH Database Tag to AAMI Tag

AAMI 标签	MIT-BIH 标签				
正常心拍(N)	正常心拍(N)	左束支传导阻滞(L)	右束支传导阻滞(R)	心房逃逸(e)	交界逃脱心拍(j)
室上性异位心拍(S)	心房早搏(A)	畸形房性早搏(a)	交界早搏(J)	室上性早搏(S)	
室性异位心拍(V)	室性早搏(V)	心室逃逸心拍(E)			
融合心拍(F)	心室和正常搏动的融合(F)				
不可分类心拍(Q)	起搏心拍(/)	起搏与正常搏动融合(f)	不可分类心拍(Q)		

3 心电图类型识别系统

本章将详细介绍提出的心电图类型识别算法。首先，本文通过分析现有工作的不足和改进思路，接着提出改进算法，并对算法进行详细介绍，然后利用多组实验验证算法的有效性，最后对本章进行小结。

3.1 现有工作局限及改进思路

现有心电信号类型识别系统主要由心拍切分，特征提取和分类三部分组成，如图 3-1 所示。心拍切分是将 ECG 记录切割成一系列单个心拍的集合。单个心拍主要包括 P 波，QRS 波和 T 波，这些波被称为子拍。心拍或子拍的长度随时间，个体和环境条件而变化。准确的心拍切分是特征提取和分类的关键基础。特征作为心拍的抽象表达，主要分为两类，即手工制作的特征和深度特征。手工制作的特征是在子拍，心拍或 RR 间隔的长度上定义的统计数据。这类特征通常具有明确的医学意义。深度特征是具有高维度和潜在含义的深度神经网络（例如 CNN）的输出。可以利用深度神经网络或传统机器学习分类器（例如，SVM）对心拍进行分类。特征提取和分类都需要一致形状的输入数据，即心拍或子波需要具有固定长度。现有工作在心电类型识别系统的三个组成部分都有一定的局限性，主要体现在以下几个方面。



图 3-1 心电图分类系统流程图
Figure 3-1 ECG Classification System Flow

心拍切分。如 1.2.1 小节所述，现有的心拍切分方法的基本原理是，以 R 波尖峰为参照点，在参照点两边分别取一个固定的长度，以此来表示一个心拍的长度，如文献[2-5]，这种方法本文称之为固定长度的心拍切分方法。这种方法在子拍切分中同样也有使用，文献[17]首先对原式数据进行下采样，然后在 R 波尖峰前后分别去 26 和 37 个点。基于获取到的心拍，他们又提出了子拍的切分。前 20 个点代表 P 波，紧接着的 20 个点代表 QRS 波群，最后 24 个点代表 T 波。很显然，这种子拍切分的方法也是固定长度的切分方法。这种方法的特点是每个切分得到的心拍或者子拍都具有相同的长度，这样做的目的是为了满足不同分类器和特征提取工具的要求，即输入数据具有相同的长度。但是，这种固定长度的心拍切分方法有一个严重的缺陷，忽略了心拍长度的动态性。由于时间、个体及环境条件的变化都

会引起心拍或子拍长度的变化，但是，固定长度切分方法却忽略了心拍或子拍长度的动态性。实际上，因为心拍或子拍的长度是变化的，现有的固定长度切分方法可能导致心拍或子拍是不完整的或冗余的，而这种不完整或冗余会影响特征提取的准确性，进而影响分类器性能。

特征提取。在特征表达上，手动特征和深度特征都被分别证明了其有效性，但是很少有研究将这两种特征大规模的结合使用。鉴于手动特征结合了大量的医学知识，而深度特征提取了大量抽象的信息，二者结合可以互补彼此没有的信息。除此之外，现有研究还有一个局限性，无论是手工特征还是深度特征，现有工作大多数都是基于单个心拍提取到的，然后对于一个医学专家而言，子拍和多拍也是心电图分类的重要特征。文献[17]引入了子拍，即 P 波，QRS 波群和 T 波，并证明了子拍对心拍分类效果的有提升作用。由于他们切分子拍采用的是固定长度切分方法，所以切分得到的子拍并不准确，其效果也有限。除此之外，多拍还没有被引入。本文定义多拍为多个心拍之间的联系，在后文中，将会详细介绍多拍特征的使用。

分类。近些年，比起传统的机器学习分类器，神经网络得到了更广泛的使用，因为神经网络可以获得一个更好的分类效果。值得注意的一点是，现有工作所使用的神经网络都是基础的神经网络，而文献[53]证明了分类器中的通道信息对很多图片任务，例如图片分割，都有显著效果，同时，也点出了现有基础的神经网络对通道信息提取的能力非常有限。文献[54]在空间上延伸了这种思想，并且证明了其有效性。对基本的神经网络而言，每个通道特征具有相同的重要性，显然，这是不合理的。所以可以引入通道注意力模块，给每个通道赋予不同的权重，进而改善分类器的学习能力。

总的来说，现有心拍类型识别算法的缺点与局限性有以下几点，(1) 在心拍切分上，用的是固定长度心拍切分方法，忽略了心拍的长度是动态变化的，从而导致了心拍切分的不完整或者冗余，进而影响后面特征提取和分类器的性能；(2) 在特征提取上，用到的特征比较单一，都是心拍尺度上的特征，子拍的特征很少用到，多拍的特征更是没有提及，同时，手动特征和深度特征结合做的不是很好；(3) 在分类器的选择上，虽然深度学习被广泛使用，但是现有分类器的对通道信息不敏感的特点，并没有得到改善。

针对以上问题，本文提出了相应的改进算法：(1) 对于心拍切分，提出一种根据 RR 波峰间隔来自动调整心拍长度的方法，对于子拍切分，根据子拍的特点，提出重叠采样的采样方法，为了提取多拍特征，本文也提出动态多拍切分方法，以上得到的心拍都是动态变化的，为了满足特征提取和分类器固定长度的要求，采取填充和截断的方法，但是对于心拍、子拍和多拍，填充方法的具体使用却各有

不同；(2) 在特征提取阶段，提出大规模结合手动特征和深度特征，其中手动特征需要结合医学知识和实验观察，除此之外，本文从多个尺度上提取手动特征和深度特征；(3) 为了改善分类器对于通道信息的学习能力，引入并改善了通道注意力模块。

3.2 心拍切分

3.2.1 自适应的单拍切分

为了切分一个心拍，首先需要找到一个心拍的起点，和该心拍的长度。如图 3-2 所示，T 波和 P 波之间的波形十分的平坦，所以当前心拍和上一个心拍的边界是模糊的，难以确定。但是利用文献[55]的方法，R 波的波峰是很容易并且快速准确的检测到。RR 间隔的长度，即当前 R 波尖峰与后一个 R 波尖峰之间的长度，被广泛用来去估计一个心拍的长度。所以，本文定义，第 i 个心拍的长度就是第 i 个 RR 间隔的长度。

$$l_s^i = R_{i+1} - R_i \quad (3-1)$$

其中 R_i 表示第 i 个心拍的 R 峰， l_s^i 表示第 i 个心拍的真实长度。从公式(3-1)中可知，心拍的真实长度是动态变化的。

通过公式(3-1)，确定了心拍的真实长度，但是心拍的起点依旧难以确定。如图 3-2 所示，以 R 峰作为一个参照点，心拍的起点可以通过把 RR 间隔的长度向左偏移一个 l_f 的长度。众所周知，P 波，QRS 波群和 T 波在一个心拍中各自占据一个比较固定的比例。所以，本文设定参数 n_f ，然后即可算出 l_f ，如下，

$$l_f^i = n_f \times l_s^i \quad (3-2)$$

其中 n_f 代表 P 波在一个心拍中所占比例，并且对所有的心拍都是同一个值。该参数的选择会影响心拍切分的准确性，将会在后文中进行详细的讨论。

通过以上两个步骤，获得的心拍的真实长度是动态变化的，但是由于特征提取和分类都需要心拍是固定长度的，所以需要把心拍的长度从非固定的重整成固定的。一种直观的解决方法是缩放，但是由于缩放操作会导致不同的心拍会有不同程度的缩放，进而导致不同的心拍会有不同程度的变形和损伤，这对后期特征提取和分类是极其不利的。所以，本文用填充和截断的方法。如图 3-3 中 b) 子图所示，本文设置了一个长度参数 L_s ，且称为有效长度。当 l_s^i 比 L_s 大，因为 T 波尾部对分类而言是无意义的，所以直接在心拍的尾部进行截断。否则，本研究采用填充

策略。首先，把每个心拍的 R 峰固定在一个心拍序列的指定位置，即对所有的心拍，R 峰具有相同的位置下标。然后在心拍的头部和尾部补零。同时， L_s 的选择必须十分谨慎和合理，因为如果 L_s 太小了，绝大多数的心拍都会是不完整的。相反的，如果 L_s 太大了，虽然绝大多数的心拍都是完整的，但是又会出现新的问题，即大部分心拍序列中尾部都有太多的零，而这些无效数据对于分类器而言就是一种噪声。在后文中，将会根据实验数据集心拍真实长度 L_r 的分布来选择合适的有效长度 L_s 。

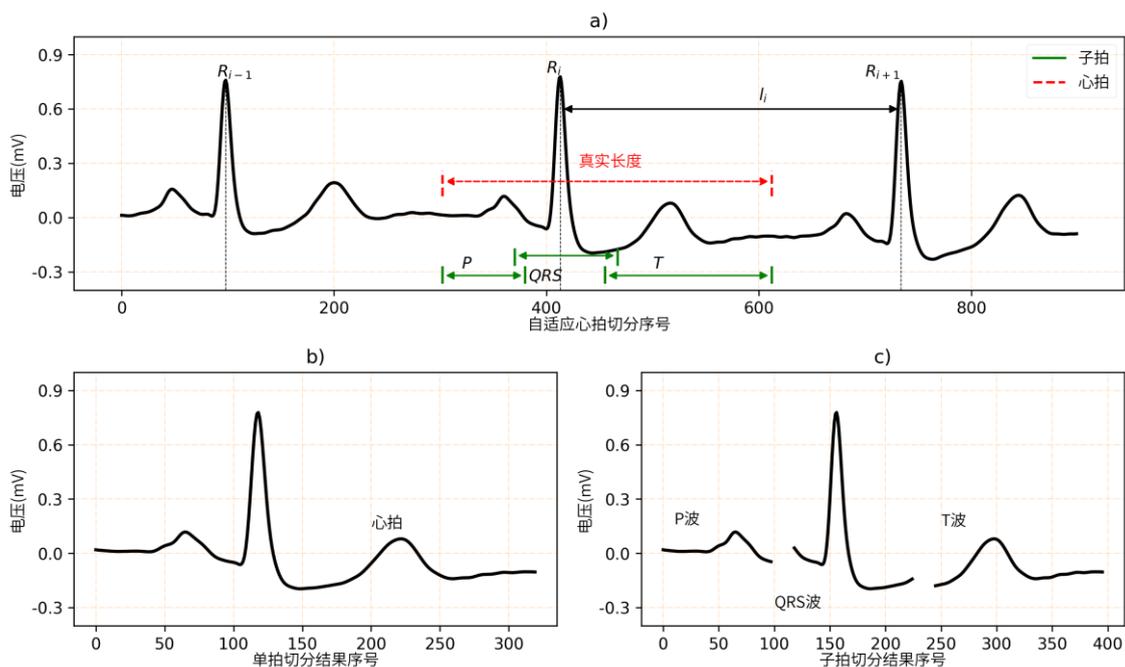


图 3-2 连续心拍及其切分

Figure 3-2 Continuous Beats and Segmentation

使用提出的自适应的心拍切分算法，获得的心拍会具有一个固定的长度 L_s ，即有效长度，并且心拍的真实长度 L_r 是动态变化的。在图 3-3 中，本文举例展示了不同切分算法的结果。其中图 3-3 的 a) 子图是传统固定长度切分方法的结果，显然上一个心拍的 T 波被错误的切分到了本心拍中；图 3-3 的 b) 子图是本文提出的自适应切分方法的结果，显然，心拍切分的比较准确，并且为了获取固定长度，在该心拍的前后分别补了一定数量的零。

3.2.2 重叠采样的子拍切分

为了精准地获取 P 波，QRS 波群和 T 波，每个波的起点和长度也是必要的。然而，如图 3-2 所示，与单个心拍识别所遇到的问题一样，各个波的边界依旧不够清晰，以至于不能用识别算法自动识别，但是每个波在心拍中所占比例相对比

较固定，这样就可以获取每个波的长度了。为了解决每个波的起点不好确定的问题，本文采用重采样来解决该问题。可以通过公式(3-3)来获取每个波的长度。

$$\begin{cases} l_p^i = (n_p + n_o) \times l_s^i \\ l_r^i = (n_r + 2 \times n_o) \times l_s^i \\ l_t^i = (n_t + n_o) \times l_s^i \\ n_p + n_r + n_t = 1 \end{cases} \quad (3-3)$$

其中 l_p^i, l_r^i, l_t^i 分别代表P波, QRS波群, T波的真实长度。 n_p, n_r, n_t 是三个分别代表P波, QRS波群, T波在一个心拍中所占比例参数。 n_o 是一个重叠参数, 代表重叠采样区间占单个心拍长度的比例。以上四个比例参数的取值对所有的心拍都是相同的。获取到的子拍长度同样也是动态变化的, 同样需要重整到固定长度。如图3-3的d)子图所示, 与单个心拍的重整操作不同, 补零操作是直接补到各个子拍的尾部。值得注意的是, 当单个心拍的长度是大于 L_s 的, 本文的子拍提取是基于截断后的心拍, 所以对于子拍而言, 没有截断操作。之所以这样操作的原因是因为发现这样获取到的子拍更为完整。图3-3的c)子图展示了利用文献[4]的方法获得的子拍, 他们通过直接把单个心拍分成三个互不重叠的固定长度的部分来获得子拍, 并且他们的心拍也是用固定长度的切分方法获得的。这种方法忽略了单拍和子拍长度的动态性, 可能导致子拍不完整或者冗余, 甚至错误地把前一个心拍的T波认为是当前心拍的P波, 这对特征提取和分类器都是一个严重的干扰。显然, 本文的方法获取到的子拍更加准确。

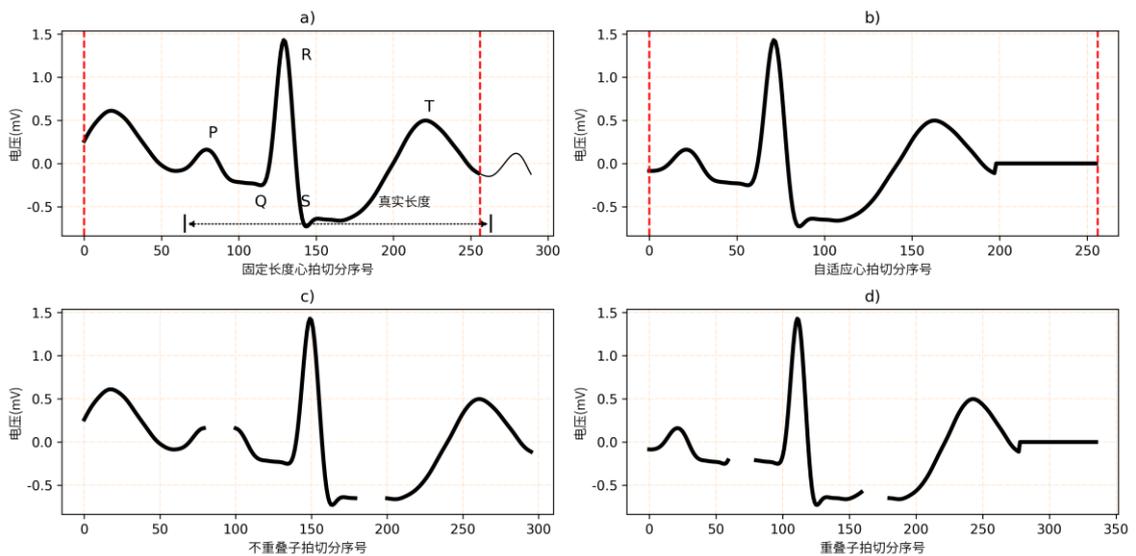


图 3-3 不同切分方法的结果
Figure 3-3 Results of Different Segmentation Methods

3.2.3 多拍的切分

为了切分多拍，多拍的起点和长度同样是必要的。但是首先需要定义多拍。对于医学专家，多拍是指提取多个心拍之间的联系，实际上就是找心拍之间的异同点。本文定义多拍为当前心拍的三倍真实长度的心拍。值得注意的是，这儿的定义和三个连续心拍的概念是不同的。因为主要目的是对当前心拍进行分类，所以把多拍定义为当前心拍的三倍更为合理。对于心拍 i ，本文在当前心拍的之前和之后各取一个真实长度 l_s^i 。于是多拍的真实长度定义如下，

$$l_r^i = 3 \times l_s^i \quad (3-4)$$

其中 l_r^i 表示多拍 i 的真实长度，并且是基于 l_s^i 而动态变化的。为了满足特征提取和分类的要求，即输入数据要求是相同的长度，同样需要采取重整操作。与之前的方法略有不同，补零和截断都会在多拍的尾部进行。

3.3 多尺度特征提取

基于心拍，子拍和多拍的切分，多尺度的手动特征可以快速的提取。手动特征主要分为两大类。第一类是为了反映子拍和心拍的平均电压强度和电压强度的变化程度，本文计算了五种参数，即均值，方差，绝对值均值，平方均值和绝对值的平方根均值。因此，本文一共获得了 20 个手动特征。第二类是和 RR 间隔相关的特征。首先，单个心拍的 RR 间隔被当作一个特征。为了反映心拍之间的统计特征，五种相邻 RR 间隔的比例也被计算当作手动特征，具体如下，

$$\left\{ \begin{array}{l} r_1^i = \frac{RR_{i-1}}{RR_i} - 1 \\ r_2^i = \frac{RR_i}{RR_{i+1}} - 1 \\ r_3^i = \frac{RR_{i-1}}{RR_{i+1}} - 1 \\ r_4^i = \frac{|RR_{i-1} - RR_{i+1}|}{RR_{i-1} + RR_{i+1}} \\ r_5^i = \frac{RR_i}{RR_{i-1} + RR_{i+1}} \end{array} \right. \quad (3-5)$$

其中 RR_i 表示第 i 个心拍的 RR 间隔， $r_1^i, r_2^i, r_3^i, r_4^i, r_5^i$ 都是和 RR 间隔相关的数值。综上，本文一共设计了 26 个手动特征。这些手动特征会弥补分类器缺乏医学知识的

缺点。

在 3.2 节中，本文提出了子拍、单拍和多拍的切分方法，利用此方法，可以准确的得到不同尺度上的心拍，再利用一维卷积神经网络，可以快速获得多尺度的深度特征。

3.4 分类器的改进

为了解决 CNN 在学习通道信息时能力有限的问题，本文引入了通道注意力模块，该模块可以帮助 CNN 更有效的提取通道信息，并且在多种图像任务上被证明是非常有效的，比如图像分割任务。如图 3-4 所示，输入的特征图是尺寸是 $C \times L$ ，其中 C 表示通道数， L 表示每个通道的特征长度。为了学习各个通道的重要性，本文采取了与文献[54]相似的方法。对于每个通道，全局最大池化和全局平均池化都被采用了。然后，一个共享的有一个隐藏层的全连接模块紧接其后。隐藏层的神经元数是 C/r ，其中 r 是一个超参数，被称作压缩率，输出层的神经元数是 C 。接着，把两个张量相加，并将得到的新张量与原始特征图相乘，最终得到了加权之后的新特征图。用公式表达本文的通道注意力模块如下，

$$\begin{aligned} F_n &= F \times [\sigma(MLP(AP(F))) + \sigma(MLP(MP(F)))] \\ &= F \times [\sigma(W_1 \sigma(W_0(F_a))) + \sigma(W_1 \sigma(W_0(F_m)))] \end{aligned} \quad (3-2)$$

其中 F 代表原始特征图， F_n 代表新特征图， AP 和 MP 分别代表全局平均池化和全局最大池化， F_a 代表通过全局平均池化操作获得的张量， F_m 代表通过全局最大池化操作获得的张量， MLP 代表全连接模块。 σ 代表 ReLU 激活函数， W_0 和 W_1 是全连接模块的权重矩阵，全局最大池化和全局平均池化获得张量共享这两个矩阵。

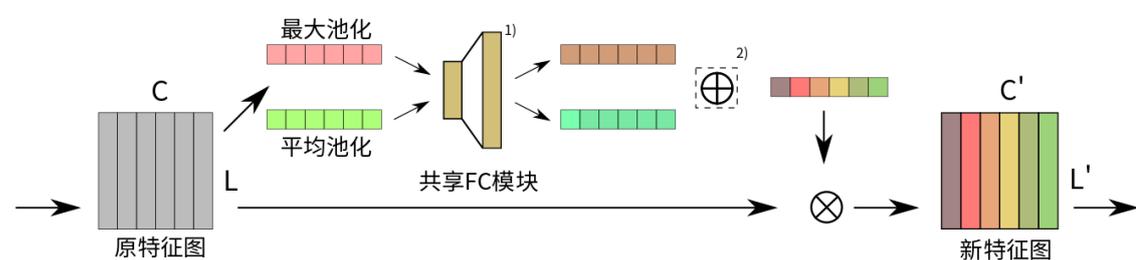


图 3-4 改进后的通道注意力模块
Figure 3-4 Improved Channel Attention Module

在图 3-4 中，本文的通道注意力模块与文献[54]的通道注意力模块有几处明显的不同。第一，在文献[54]的通道注意力模块中，共享全连接层的输出层的激活函数是 sigmoid 函数，这里改用了 ReLU 函数，因为通过实验发现，修改之后的模块会对最后准确率有 1% 的提升。第二，在文献[54]的通道注意力模块中，两个由共

享全连接层产生的张量相加之后还要经过一个激活函数，本研究把这个激活函数丢弃了，相加之后就是本文的权重张量，因为这样会使最后的准确率更高。

3.5 实验验证

本节将对提出的算法进行验证，首先对实验的一些参数设置进行说明，其次介绍使用到的一维卷积神经网络，然后进行不同心拍切分方法实验和多尺度特征实验，最后有相关工作进行对比分析。

3.5.1 实验设置

数据集的划分。根据 AAMI 的推荐标准，每个个体的训练数据总时长不超过 5 分钟。同时，由于本文选择的对所有个体训练一个分类器，为了和其它文献有可比性，本文参照文献[4]的数据集划分方法，做了同样的数据集划分。对 MIT-BIH 数据库的前 20 条记录（记录号 100-124），从类别 N、类别 S 和类别 V 各取 75 个心拍，类别 F（13 个）和类别 Q（7 个）的心拍全部取出，总计 245 个心拍。这 245 个心拍和后 24 条记录（记录号 200-232）的前五分钟的心拍被用来做训练集。剩余的所有心拍被用来做测试集，同时，从测试集里面随机抽取 5000 个心拍做验证集。值得注意的是，整个数据集大约有 11 万个心拍，其中训练集略多于 1 万个心拍，不到全数据集的 10%。同时，为了降低数据冗余，对滤波后的数据做了三倍下采样，即采样到 120Hz。此外，后面所有的实验都是基于这种数据集划分的方法。

众所周知的是，神经网络具有一定的波动性，其最后的结果也有一定的随机性，为了降低这种随机性，本研究把所有的实验重复十次，然后对实验结果取十次的均值。同时，为了量化这种波动性，本文定义了一个可以测量模型稳定性的参数，并称之为模型的稳定性 s ，其定义如下，

$$s = -\log_{10} \left[\sum_{i=1}^n (a_i - \bar{a}) \right] \quad (3-7)$$

其中 n 是实验重复的总次数，本文中是 10。 a_i 是第 i 次实验最终的测试准确率。 \bar{a} 是 n 次实验的平均测试准确率。显然， s 的值越大，模型越稳定。

在前几节，介绍了本文提出的各种算法，同时还有各种参数的值需要确定。在本节，将详细解释各种参数值的选取。参数的选择主要是基于两种原则：第一种是对于某些参数，其本身不具备具体的医学含义，则本文对该参数值的选取，

是基于实验结果选取的；第二种是对于某些参数，具备具体的医学含义，则根据其医学含义选取该参数值。

关于 L_s 的选取。 L_s 指一个心拍的有效长度，对所有的心拍都是同一个值，该值过大过小都是不合适的。所以基于 MIT-BIH 数据库上心拍的真实长度的分布来取值，全数据集上心拍的真实长度分布如图 3-5 所示。首先，计算出全数据集上心拍的真实长度，然后以 50 为心拍真实长度区间，统计每个区间内的心拍数量，值得注意的是，某些区间的样本数少于 100，就没有画出该区间。从图中可以看出，区间 300-350 上，拥有最多的心拍，同时，该区间左边包含了大部分的心拍，所以本文把 L_s 的值设为 300。实际上，因为本文对原始数据做了 3 倍下采样，所以 L_s 设为了 100。

关于 n_f, n_p, n_r, n_e, n_o 的选取。这几个参数都和子拍的占比有关，根据其医学含义选取合适的值比较合理。以 n_f 为例，首先，根据医学知识给 n_f 设定一个范围，然后再以一定的间隔取一些值。然后画出 44 条记录的前 5 分钟心拍在不同参数值下的切分结果，根据切分结果的表现选择最合适的值。具体的，依次从以下序列中取值，即[0.20,0.25,0.30,0.35,0.40,0.45]，然后画出不同参数值下的各心拍，并判断所有心拍的完整性。最后发现，当 n_f 取值 0.35 时，切分的心拍最为完整。同理，本文得到 n_p, n_r, n_e, n_o 的最佳取值分别为 0.25,0.25,0.5 和 0.03。除此之外，经过多次试验，最终发现当通道注意力模块中的压缩率 r 为 4 时，网络能获得最好的实验结果。

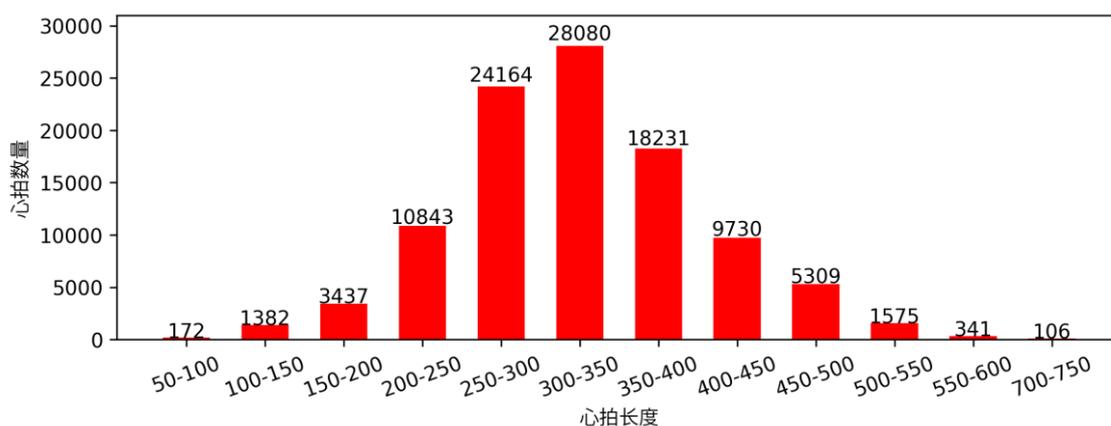


图 3-5 MIT-BIH 数据库的心拍真实长度分布

Figure 3-5 The Distribution of Beat Real Length on MIT-BIH Database

3.5.2 实验网络设计

除了选取合适的参数值，还需要合理设计网络结构，合理选择网络参数。为

了对比有通道注意力模块的网络和没有通道注意力模块的网络，本文设置了两个相似度非常高的网络。如图 3-6 所示，他们唯一的区别就是在最后一层卷积层，是否采用带有通道注意力模块的卷积层。

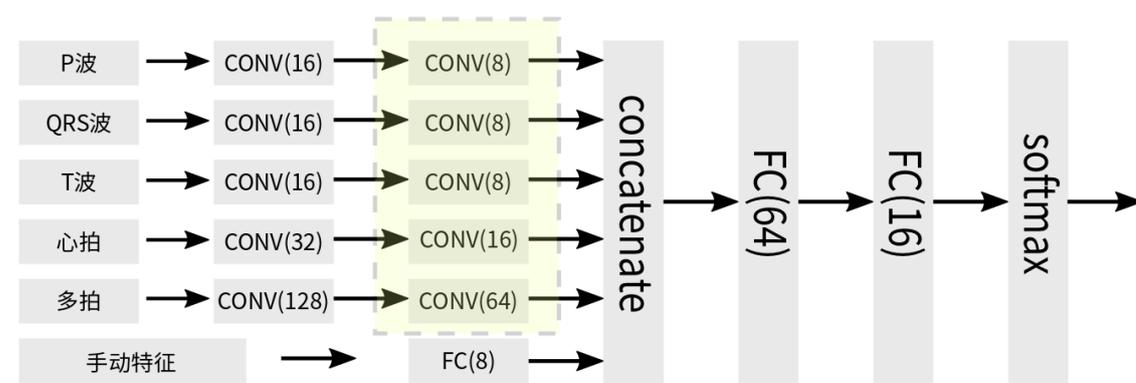


图 3-6 卷积神经网络结构
Figure 3-6 The Structure of CNN

由于心拍的表达是多尺度的，所以需要设计一个多输入的一维卷积神经网络，其结构如图 3-6 所示，这是没有加入通道注意力模块的网络图，有通道注意力模块的网络图就在原图的第二层卷积上修改即可。在图 3-6 中，一个 CONV 模块表示一个卷积模块，其组成结构如图 3-7 所示，一个卷积层，一个随机失活层，然后是一个最大池化层，后面接着一个批归一化层。值得注意的是，在第一层的 CONV 模块，本文删除了随机失活层，因为随机失活层的本质是为了削弱模型可能训练过拟合的作用，而其效果更倾向于在网络较深层起到更明显的作用。同时，为了考虑训练模型的时间成本和最终模型效果，本文选择了只有两层卷积层的卷积神经网络，因为通过实验发现，验证效果在用一层或者两层卷积时，会有较大的差距，并且训练时间差距不是很大，但是当卷积层增加到三层及以上时，验证效果提升的非常缓慢，已经接近饱和，但是训练时间会有十分明显的增加。综合考虑，本文选择了两层卷积层。除此之外，文献[53]证明了通道注意力模块只在较深层使用时，会得到一个较好的效果，即只在第二层添加通道注意力模块。事实上，也通过实验证明，在所有层加上通道注意力模块的效果都不及只在第二层加上通道注意力模块的效果好。因为在更深层获取的特征更为抽象，也更为高级，而更高级的特征对不同类别的心拍有更好的区分度。所以，只在第二层加上了通道注意力模块。在全连接模块层数的选择上，采用了和选择卷积层层数时相同的原则，并且最终确定了用两个隐藏层加上一个输出层，就能取得很好的结果，并且训练时间也是可以接受的。图 3-6 中的一个 FC 模块的组成情况如图 3-7 中所示，一个全连接层，一个随机失活层和一个批归一化层。



图 3-7 模块组成
Figure 3-7 Module Configuration

每一层的神经元数都在图 3-6 的括号内写出, 这些参数都是经过大量的实验和调参确认的。同时, 所有的随机失活层的参数都被设为了 0.2, 最大池化层的参数都设为了 2, 参数初始化选择了 Xavier 初始化方法, 迭代次数为 30 次。为了防止模型过拟合, 在输出层添加了 L2 正则化和早停法 (early stopping), 其中 L2 正则化的权重是 0.01。

3.5.3 不同切分方法实验

为了证明提出的心拍和子拍切分方法有更好的性能, 本节会做一系列的实验。为了后文表述方便, 本文把自适应的心拍切分方法称为 ABS, 现有的固定长度心拍切分方法称为 BS, 基于 ABS 的重叠采样子拍切分方法称为 OWS, 基于 BS 的非重叠采样的子拍切分方法称为 NWS。同时, 称通道注意力模块为 CAM。

在图 3-8 中, 展示了不同切分方法的实验结果。比较 ABS 算法和 BS 算法的结果可知, 无论是在准确率上还是模型稳定性上, ABS 都比 BS 更胜一筹, 说明了相比于 BS 方法, ABS 不仅能够准确的切分心拍, 并且对最后的结果有着明显的提升, 证明了 ABS 的有效性。同理, 对比 NWS 和 OWS 可知, OWS 在两项指标上同样优于 NWS, 证明了本文提出的 OWS 的有效性。有趣的是, 通过对比 ABS 和 NWS, 会发现 ABS 的两项指标均比 NWS 效果要好, 这说明了心拍切分的重要性, 当心拍切分不精准时, 及时用了多尺度的特征, 其效果也不一定好, 也验证了不准确的心拍切分会影响特征提取并且混淆分类器。

同时, 本文对 BS, ABS 和 OWS 都做了不同网络的对照实验, 从三组对照实验中可以看出, 加入 CAM 对各种方法的准确率都有明显的提升, 尤其是在特征表达尺度单一的时候, 这种优势更为明显, 并且 CAM 的优势会被充分的特征表达有所弥补。对比有无 CAM 的 BS 方法的结果, 可知即使心拍切分的不好, 通道注意

力模块也能筛选出更为稳定有效的特征。但是对比 BS 和 ABS 两种方法在不同网络中的准确率增量，前者是 0.5%，后者是 1%，ABS 是 BS 的两倍，证明了 CAM 在精准的心拍切分上能取得更好的结果，再次证明了精准切分的重要性。

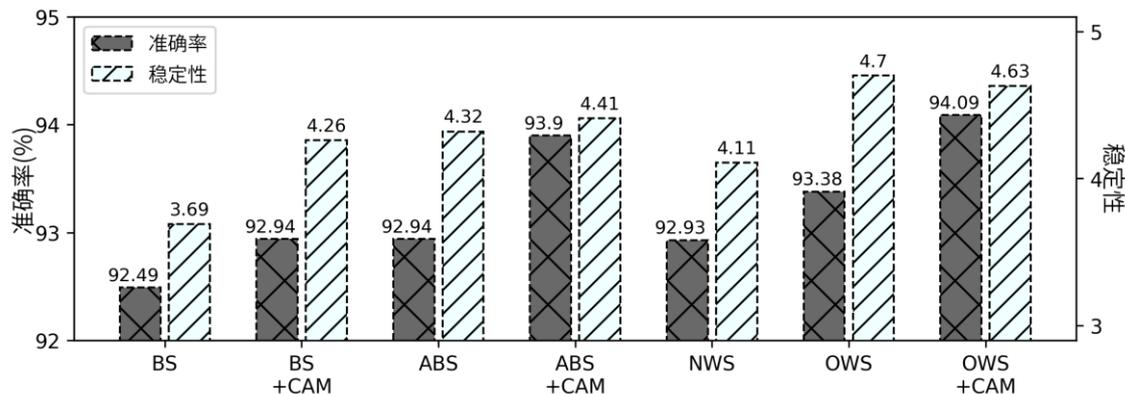


图 3-8 不同切分方法的实验结果

Figure 3-8 Experimental Results of Different Segmentation Methods

除此之外，CAM 对模型稳定性参数也有着明显的提升效果，仅仅在 OWS 下，CAM 的表现略逊，但是差距也很小（实际上，两者的方差都是 10^{-5} 的级别，非常的小）。这是因为首先即使十次试验会减弱神经网络的随机性，但是无法消除，所以这种结果也具有一定的随机性，其次，由于本文在网络中加入了早停法，加上为了考虑时间成本，训练次数很少，而 CAM 的网络参数要比不加通道注意力模块的网络要多，所以有可能在某些次试验中，CAM 的网络未得到充分的训练，导致其模型稳定性出现了一定的欠拟合。

3.5.4 多尺度特征实验

为了证明多尺度特征（即子拍、心拍和多拍特征）的有效性，本文也进行了大量的实验。多尺度特征分为两部分，一部分是由一维卷积神经网络提取的多尺度特征，另一部分是手动提取的多尺度特征。同上，本文把多拍切分方法称之为 MBS，把只有一维卷积神经网络提取的多尺度特征称为 MF（也就是 ABS，OWS 和 MBS 方法的结合），把手动提取的多尺度特征称为 HF。

如图 3-9 所示，本文依次用只有心拍特征（ABS），心拍和子拍特征（ABS+OWS），心拍、子拍和多拍特征（ABS+OWS+MBS，即 MF）以及 MF 叠加 HF 的特征分别进行多次实验。并且针对 MF 叠加 HF，还利用加入 CAM 的网络作为对照实验，以继续验证 CAM 的有效性。从图中可以观察到，准确率随着更多特征的加入而提升。同时，在相同的情况下，即用 MF 叠加 HF，CAM 的网络

依旧比普通的神经网络有更好的学习能力，不过对准确率的提升也越来越小。这主要是两点原因造成的，一是 CAM 网络参数更多，网络也更为复杂，但是为了保持两种网络的其他超参数的一致性，本文设置的训练次数非常的有限，导致 CAM 网络没有得到充分的训练；二是随着特征的加入，心拍表达也变得更加丰富和高效，这种充分的数据表达会弥补网络学习能力有限的缺点，可以理解为充分的数据表达和通道注意力模块具有相似的作用，均弥补网络学习能力的不足。从图中可以看出，随着多拍特征的加入，使得模型准确率有了极大的提升，证明了前后心拍的特征对当前心拍的类型判断有着非常大的影响，而这在之前的工作都被忽略了。并且，多尺度手动特征和多尺度深度特征结合，对准确率也有着十分明显的提升，说明医学知识的加入也会弥补网络的不足，这在之前的工作也是被忽略的。

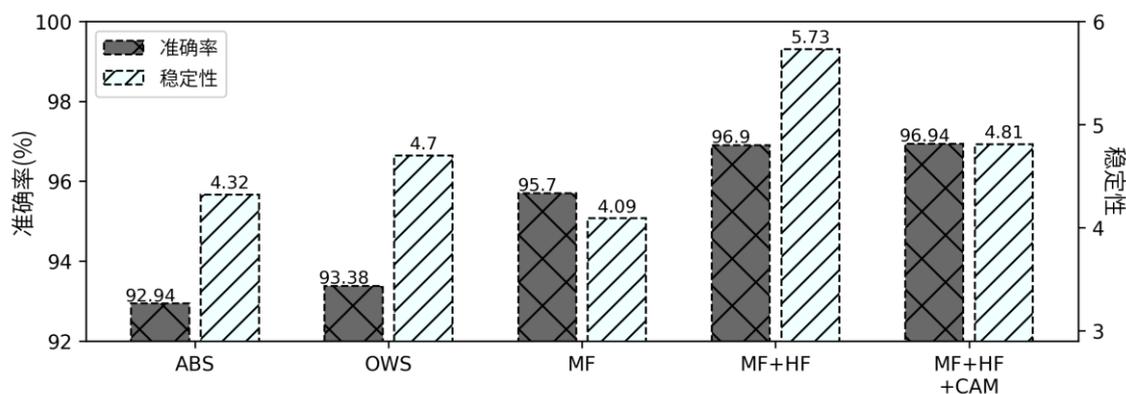


图 3-9 多尺度特征实验结果
Figure 3-9 Experimental Results of Multi-Scale Features

同时，随着特征的加入，模型稳定性也成上升趋势。因为随着特征的加入，一方面每个特征的权重都会对应的减小，使得那些原本表现不稳定的特征对模型的影响变得更小，另一方面更多特征的加入使得心拍表达更为丰富，不同类别心拍之间的差异性也增大了。但是，CAM 模型稳定性略有下降，其原因与 3.5.3 小节相同。

除此之外，为了验证本文算法的优越性，在表 3-1 中，把本文的结果和现有的较为出色的工作的结果进行了对比（其中括号里的是 CAM 的表现），由于之前的工作都是只进行了一次实验，并未对实验进行多次重复，所以无法计算其模型稳定性，只能对准确率进行比较。在表 3-1 中可以看出，本文的工作明显优于其他同类工作。和现有工作中最好的结果比较，本文把心拍的错误分类率降到了仅为原有错误率的 50%左右，鉴于错误分类率提升空间很小，只有不到 7%，本文的改进是巨大的。

表 3-1 相关工作对比（小括号里的是通道注意力模块的效果）
Table 3-1 Comparison of Related Work (Results of CAM Showing in Parentheses)

方法	年份	准确率
文献[56]	2014	91.7%
文献[4]	2016	92.7%
文献[57]	2017	93.47%
文献[58]	2018	92.06%
本文	—	96.90% (96.94%)

3.5.5 各类型心拍的实验表现

以上的实验都是从整体角度分析了提出的算法对整体效果的提升，本节将分析每个类型的心拍的具体提升情况。AAMI 建议的五类分类标准，除去训练集和验证集，在本文的测试集中有 90000 多个样本，其中 N 类约 80000 个样本，S 类约 3000 个样本，V 类约 6000 个样本，F 类不到 800 个样本，Q 类仅 8 个样本，是一个分布非常不均匀的样本集。考虑到 Q 类的样本量太小，且本身就是不可分类的类型，没有实际意义，所以，本文不给予分析。

因为涉及到每一类心拍，就需要有混淆矩阵，但是文献[58]的混淆矩阵只显示了后 24 条记录的分类结果，也就是产生训练数据的记录，明显会得到一个更高的结果，并且这样的混淆矩阵也不具备太多的意义，同时文献[57]在他们的训练数据和测试数据中加入了大量的伪造数据，使得原本分布非常不平均的样本集变得均衡，使得每一类的表现都非常相似，但是由于里面有大量非真实样本，其实际意义有待考证。鉴于本文的工作和文献[4]的工作使用的数据集划分方法相同，所以和文献[4]进行比较是非常合适的。另外，考虑到 CAM 的网络可能为得到充分的训练，所以使用 MF 叠加 HF 的结果进行比较。

在表 3-2 中，计算了每个类别的精确率，召回率和 F1 值，其中小括号内的是文献[4]的结果。由于本文的实验重复了十次，所以先计算每次实验中，每个类型的精确率，召回率和 F1 值，然后对十次结果取均值，得到了最后的结果。从表中可以看出，本文在各个类别的各个指标上都比文献[4]的值有着明显的提升，尤其是对于异常的三类，虽然其样本数较少，但是各类参数值都有着明显的提升，这样的提升显然是非常有意义的，因为对于心电自动分类系统，把异常心拍判断出来更为重要。其中对于 S 类，文献[4]的各个指标都不超过 0.5，本文的算法把精确率和召回率和 F1 指数都提升了接近 0.25，这显然是个十分明显的提升。同理，对 V，F 类也有类似的提升，证明了本文的算法确实在各个方面的表现都优于现有的

算法。在看 N 类，虽然现有工作在各个指标上原本就有着很高的值，但是本文的算法在各个参数上仍有明显的提升，鉴于各个参数值原本就很高，这样的提升意义也十分重大，F1 值甚至达到了 0.99。

表 3-2 各类别心拍精确率、召回率、F1 值（文献[4]的结果在小括号里）
Table 3-2 Precision, Recall, F1 of each class(Results of Reference[4] Showing in Parentheses)

类别	N	S	V	F
精确率	98.03% (97.24%)	72.33% (44.66%)	92.08% (64.26%)	74.59% (64.50%)
召回率	98.98% (96.54%)	57.14% (35.08%)	89.47% (79.20%)	64.06% (61.02%)
F1 值	0.99 (0.97)	0.64 (0.39)	0.91 (0.71)	0.69 (0.63)

3.6 本章小结

本章，本文的工作主要体现在以下几个方面。

首先，本文对现有心电自动类型识别系统分析，发现了现有心电类型识别系统的局限与不足，主要体现为以下几点：(1) 心拍切分的不准确，把心拍切成固定长度，忽略了心拍的长度随着时间、场景和个体的不同而改变，这是现有心电自动分类系统最大的不足，因为不准确的心拍切分，会影响后续的特征提取和分类；(2) 特征单一，一方面是忽略了在心拍分类时，多拍和子拍的特征也十分重要，另一方面是深度特征和手动特征具有互补特性，应将两者结合起来用于分类；(3) 分类器学习能力的不足没有得到解决，传统的卷积神经网络对通道信息的学习能力十分有限，不能区分不同通道的信息有不同的重要度。

其次，为了解决现有的心电自动类型识别系统的局限和不足，本文提出了以下的算法来改进类别识别系统：(1) 把固定长度的心拍切分改为自适应的心拍切分，根据心拍长度来切分心拍，通过截断和填充来实现把动态长度的心拍转换为固定长度的心拍，并且基于不定长度的心拍获取子拍和多拍，为提取多尺度特征做准备；(2) 从单一的特征扩展为多尺度特征，主要包括两方面，一方面是用卷积神经网络提取多尺度的深度特征，另一方面结合医学知识，提取多尺度的手动特征，这二者的结合可以相互弥补；(3) 引入通道注意力模块，改善卷积神经网络对通道信息提取能力不足的问题。

最后，本文在 MIT-BIH 心律异常数据库上，进行了多项实验，一一验证了本

文提出的算法的有效性，并且和先有工作做了大量对比，从各个维度来看，本文的工作都是优于现有工作的。

4 心电图个体识别系统

本章进行心电图个体识别研究。具体而言，首先通过分析现有工作的局限与不足，接着提出本文的改进算法，并对本文的算法进行详细的介绍，然后利用实验验证，最后进行本章小结。

4.1 现有工作局限及改进思路

目前，心电个体识别系统是学术界的一个研究热点，其系统可分为四个模块：数据获取，样本生成，特征提取，分类器。其中数据获取指的是获取用来训练模型的数据，多为具有一段长度，含有多个心拍的心电记录。其中样本生成决定了后面需要何种类型的特征提取工具和分类器的选择。该系统流程如图 4-1 所示。现有的心电个体识别系统识别率虽然已经达到了一定的高度，但是在各个模块仍有一定的局限性。接下来，将分点详细阐述。



图 4-1 心电生物识别系统流程图
Figure 4-1 ECG Biometric System Flow

数据获取。如 1.2.2 小节介绍，现有工作的数据获取主要分为两种途径，第一种是自己收集数据，构建自己的数据库，第二种是公开数据集。自己构建的数据库主要有以下几个特点：(1) 总人数较少，受制于自建数据库的成本过高，这种自己构建的数据库总人数较少，一般维持在 20 个个体以内，而一般研究要 40-50 个个体才为合适；(2) 数据质量参差不齐，由于每个人收集心电数据的设备不同，导致收集到的数据质量参差不齐，导致算法泛化能力差；(3) 自我构建的数据库几乎都为健康的志愿者，导致算法极有可能对心率异常较为敏感。公开数据集在包含个体数量和数据质量上都得到了保证，其算法的有效性也能够的到验证。为了让提出的算法具有一定的普适性，本文研究将在 MIT-BIH 心律失常数据库进行验证。

样本生成。现有的样本生成方法，主要是基准的方法，这种方法的前提是，需要检测出一个心拍或者 R 波尖峰，如第 3 章所述，检测一个完整的心拍是极其困难的事，如果心拍是异常的，则心拍识别率会下降的更厉害，如文献[27]样本生成采用的是基准的方法，其识别准确率在 PTB 数据集的 52 个健康个体上，高达 96.15%，但是在 PTB 数据集的 48 个心律异常个体上，降到了 89.48%，在 MIT-BIH 心律异常数据库的 47 个个体上，更是降到了 83%，这样的下降是不可接受的，尤

其是在两个心律异常的数据库上，总人数减少了，还有超过 10% 的下降，这就是基准方法的一个缺点。除此之外，由于深度学习需要大量的样本进行模型训练，而基准的方法由于心电记录长度的有限，导致不能产生足够多的样本，使得模型容易欠拟合。文献[36]提出的非基准的方法，即用固定的长度在心电记录上抽样，避免了以上的两个问题，即模型对心律异常是低敏感度的，同时可以产生大量的样本，使神经网络得到充分的训练。但是文献[36]仍然有不足之处：(1) 抽样时，训练样本的抽样区间和测试样本的抽样区间重合，导致测试样本的某些数据极有可能被模型训练时“看过”，导致模型识别准确率存在偏高的可能性，应当保证在时间上，所有的训练区间在测试区间之前；(2) 取样数量和样本时长作者并未进行严谨的实验证明，尤其是训练样本，作者随机取了 500 个，并不合理，因为过于稠密或者过于稀疏的抽样，都会导致真真实分布被改变，需要实验确定合理值。

特征提取。在分类之前，需要对生成的样本进行特征提取，现有特征提取的方法有很多种，但是并不是所有的特征提取方法都有相同的效果，比如傅里叶变换和小波变换，都可做特征提取，但是由于傅里叶变换是固定窗长，其时域分辨率较差，而小波变换窗长可变，在时频域都有较好的分辨率，这导致了大多数情况下小波变换的识别率都要比傅里叶变换的识别率高。根据样本生成方法，选择不同类型样本区分度高的特征提取方式很重要。

分类。近些年，随着深度学习的发展，越来越多的研究开始选择神经网络作为分类器，但是和第 3 章分类器的问题相似，基础的神经网络通道信息提取能力不足，限制了识别准确率，这是急需解决的一个问题。

总的来说，现有的心电生物识别主要有以下几点问题。

(1) 在数据源上，使用的数据源杂乱无章，有用自己建的，也有用公开数据的。而自己建的数据库往往导致算法不可复现，并且算法泛化能力差，其可信度受到质疑。

(2) 在样本生成上，有使用基准的方法，也有使用非基准的方法。基准的方法虽然能使后续的操作变得简单，但是识别一个心拍的难度较大，导致了算法对异常心拍敏感，并且这会使得训练样本非常有限，致使深度学习可能无用武之地。对比之下，非基准的方法，对数据的要求较低，可获得大量的样本，且省去了检测心拍的步骤，具有很强的实用性，但是现在仍然需要改进。

(3) 在特征提取上，虽然多种方法被证明有效，但是各种方法的优劣性和样本生成方式有关，需要结合样本生成方式提取特征。

(4) 在选择分类器上，比起传统的机器学习分类器，深度学习分类器的分类能力更强，但是神经网络的通道信息提取能力有限，需要改进。

针对以上发现的问题，本文提出了以下算法进行改善。

(1) 在公开的数据集上验证算法，本文采用 MIT-BIH 心率异常数据库，该数据库一方面拥有 47 个个体，总人数较多，适合进行算法验证。另一方面还有大量异常心拍，还可验证算法是否对心律异常敏感。

(2) 由于基准的样本生成方法存在大量的缺点，本文采用非基准的样本生成方法，并对现有的非基准样本生成方法进行改善和更深入的探究，研究各种参数对识别准确率的影响。

(3) 由于样本是随机抽样的，不具备基准方法产生的样本的高规律性，这为特征提取带来了一定的困难，可通过域变换来解决这一问题。通过对比不同的域变换方法，本文最终采取小波变换进行样本域变换和特征提取。

(4) 神经网络由于对通道信息提取能力不足，需要解决这一问题，本文引入在第 3 章已被证明非常有效的通道注意力模块，并针对本问题继续做具体改善。

接下来，将详细介绍本文的改进算法。

4.2 样本生成

本文所提的非基准的方法即使在指定的区间长度 s 内，给定一个样本长度 l 和样本数 n ，则在区间长度 s 内，抽取 n 个长度为 l 的样本即可。这里需要注意的是，由于训练样本和测试样本都需要抽取，为了不出现类似于文献[53]的错误，即训练区间和测试区间重叠，导致测试样本可能出现在训练集中，使得模型“看过”测试样本，并进行了训练，造成了识别准确率偏高的问题，本文让训练区间和测试区间不交叉，并且让训练区间在时间轴上早于测试区间（因为对任何一个个体，肯定是先训练模型，得到他的特征和权重，才可以识别），这样就能避免出现这种问题。

如图 4-2 所示，在一整条记录上，首先根据自己的需求，定义出一个长度为 s_1 的长度区间作为训练区间，然后在训练区间上，取 n_1 个长度为 l_1 的样本，作为训练样本。同时，在该条记录上，在定义一个长度区间为 s_2 的长度区间作为测试区间，然后再测试区间上，取 n_2 个长度为 l_1 的样本，作为测试样本。需要注意的是，这里的 s_1 和 s_2 不一定等长， n_1 和 n_2 也不一定相等，因为测试样本和训练样本都没必要相等。但是最好保证式 4-1 成立。

$$\frac{n_1}{S_1} \approx \frac{n_2}{S_2} \quad (4-1)$$

这是为了保证训练区间和测试区间有相似的抽样密度，即二者分布大致相似，这是机器学习的前提。但是通过实验发现，即使这个比例关系不成立，对最终的结果影响也不大。除此之外，需要把测试区间放在训练区间之后，这是符合正常

逻辑的，因为一般心电生物识别是识别模型训练之后的时间的，而非之前，也没有实际意义，这也是文献[53]的问题。而训练样本长度和测试样本长度需要一致，这样模型才可以预测。

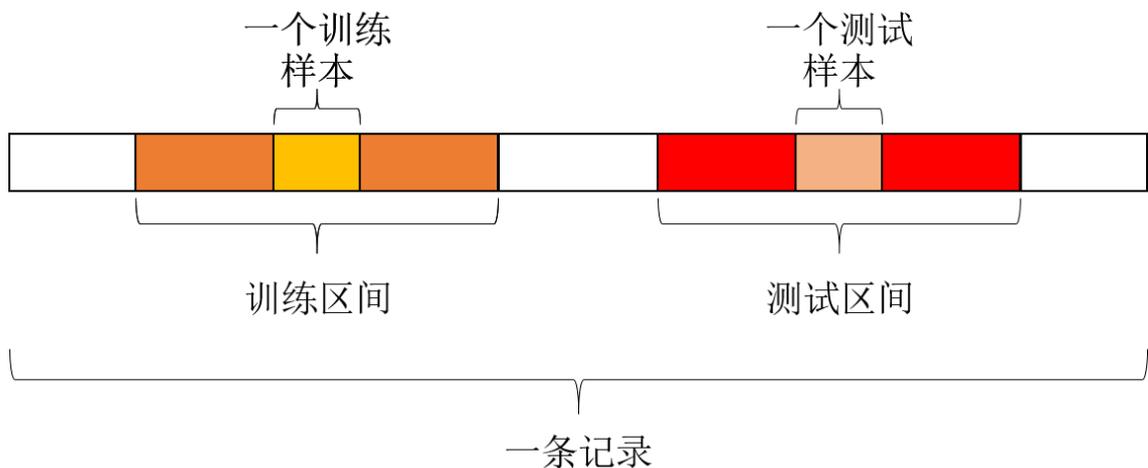


图 4-2 样本生成原理

Figure 4-2 Sample Generation Principle

一般情况下，参数区间长度 s 的选择，与一条记录的总长度有关。而样本长度 l 的选择，则要求至少包含一个完整心拍，因为如果一个样本不能包含至少一个心拍，就会造成信息极大地损失，不同个体之间的区别也就变得更小，使得模型更加难以识别不同个体，一般的，可取不小于 2 秒的值。样本数量 n 的选择，也是根据模型需要而决定的，一般可取不小于 100 的值。具体的最优值需要进行多次试验获得，将在后面进行详细的实验，并做更多的解释。

4.3 小波变换

由于本文是随机抽取样本，不像基准的方法能够知道每个波段大概在什么位置，所以用时域的信号直接卷积意义不大，会使得模型学习非常困难，所以本文用域变换来解决这个问题。常用的域变换有小波变换和傅里叶变换，下面将对这二者原理进行介绍。

小波变换是一种信号变换分析方法，其继承和发展了短时傅里叶变换局部化的思想，同时又克服了窗口大小不随频率变化等缺点，能够提供一个可变的时频窗口，这使得它是进行信号时频分析和处理的理想工具。它的主要特点是通过变换能够充分突出问题某些方面的特征，能对时间频率的局部化分析，通过伸缩平移计算对信号逐步进行多尺度细化，最终达到高频处时间细分，低频处频率细分，能自动适应时频信号分析的要求，从而可聚焦到信号的任意细节，解决了傅里叶

变换的所无法处理的情况。

傅里叶变换的是一种全局性的变化，不具备局部分析能力，同时，也不能分析非平稳信号。这抑制了傅里叶变换的适用范围，使得在很多场合，根本无法使用傅里叶变换。于是人们对傅里叶变换做了各种改进，如短时傅里叶变换（STFT）。但是由于短时傅里叶变换的滑动窗函数一旦选定就固定不变了，这导致了其视频分辨率固定不变，不具备自适应能力，但是小波变换可以很好地解决这个问题。

首先，傅里叶变换的公式如下，

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) \times e^{-i\omega t} dt \quad (4-2)$$

对比小波变换的公式如下：

$$WT(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \times \psi\left(\frac{t-\tau}{a}\right) dt \quad (4-3)$$

可以看见，小波变换函数就是修改的傅里叶变换函数，把原有的基函数换成了新的基函数，用一个时间轴上长度有限的且幅度衰减的小波基函数来替换了原本时间轴上长度无限且幅度恒定的三角函数基函数。这样做的好处就是不但可以获得频率，还可以定位到时间。

从公式(4-3)中可以看出，相比于傅里叶变换只有频率 ω 一个参数，小波变换有两个参数，分别是尺度 a 和平移量 τ 。尺度参数 a 是用来控制小波函数的伸缩，而平移量参数 τ 则是用来控制小波函数平移的多少。此处的频率则与尺度相对应（反比），时间则与平移量相对应。此外，常用的小波基函数有 haar 小波，daubechies 小波，symlet 小波，coiflet 小波，biorthogonal 小波，gaussian 小波，mexicanhat 小波等。

如图 4-3 和图 4-4 所示，利用随机抽样的方法，在 MIT-BIH 心律异常数据库上，任意选择了某两个个体，并在相同的时间区间上各抽取了一个样本，其中样本时长为 2 秒，并分别绘制了通过快速傅里叶变换（图 4-3）和离散小波变换（图 4-4）得到的特征图，其中红色和蓝色的曲线分别表示两个样本。

从图 4-3 中可以看出，两条曲线重合度非常大，除了轮廓稍有区别之外，几乎没有任何差距，但是随着个体数的增加，不同个体经过快速傅里叶变换之后的轮廓差距也会越来越小，使得难以辨认个体和样本的对应关系。因此，这是个不稳定的特征，快速傅里叶变换对本文的样本而言，不是一种合适的特征提取方法。除此之外，特征并不是一条平滑的曲线，相邻点之间的幅值变化大，使得曲线看起来杂乱无章，这对卷积神经网络而言，提取特征也是一个较难的问题。

再看图 4-4，从 5 级离散小波变换的结果来看，无论是在低频还是高频，两个样本都有着肉眼可见的区别，尤其是在高频部分，这种区别非常明显，很容易就可以看出样本和个体的关系。当随着样本的增多，虽然这种差距可能会减小，但是可以通过增加离散小波变换的级数来扩大不同个体之间的特征距离，使得模型能够准确的辨认出不同的个体。

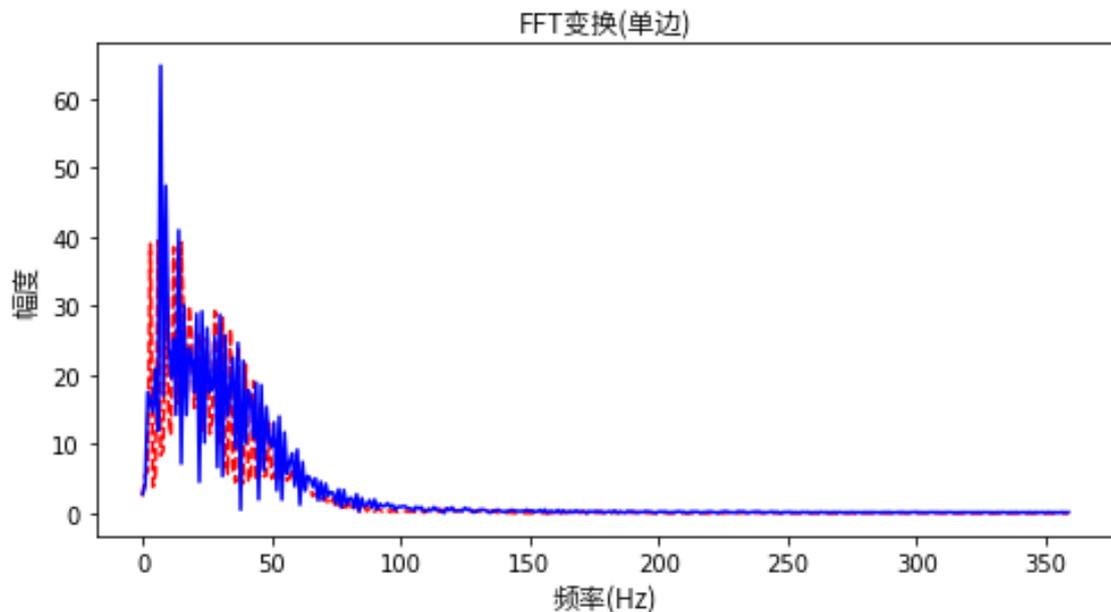


图 4-3 不同个体的 FFT 结果
Figure 4-3 Results of FFT from Different Person

通过以上的实例分析可知，针对本文利用随机抽取的样本而言，相较于时间分辨率较差的快速傅里叶变换，在时频域都具有较好分辨率的离散小波变换是更好的选择，通过离散小波变换，可以让属于不同个体的样本，在经过特征提取之后，具有更大区分度的特征，所以本文采用离散小波变换进行特征提取。

还有一点值得注意的是，离散小波变换后得到数据与快速傅里叶变换之后所得到的数据有所不同。快速傅里叶变换得到的是一组数据。且是对称的。而是离散小波变换之后得到的是多组数据，且长度有所不同。对于 n 级离散小波变换而言，得到的是 $n + 1$ 组数据。小波分解树如图 4-5 所示，每一级的高频和低频部分数据等长，即 CA_i 与 CD_i 等长，上一级的数据长度是下一级的两倍，即 CA_i 的长度是 $CA_{(i+1)}$ 的长度的 2 倍，这就是前文说的数据长度不等。在实际应用中， n 是一个很重要的参数，其取值代表着得到的特征的精度。取小了，导致精度不够，信息损失厉害，无法为不同样本提供足够的区分度；取大了，导致数据量太多，增大了计算负担。本文会通过实验选择最合适数值，后面的实验部分将会做详细介绍。

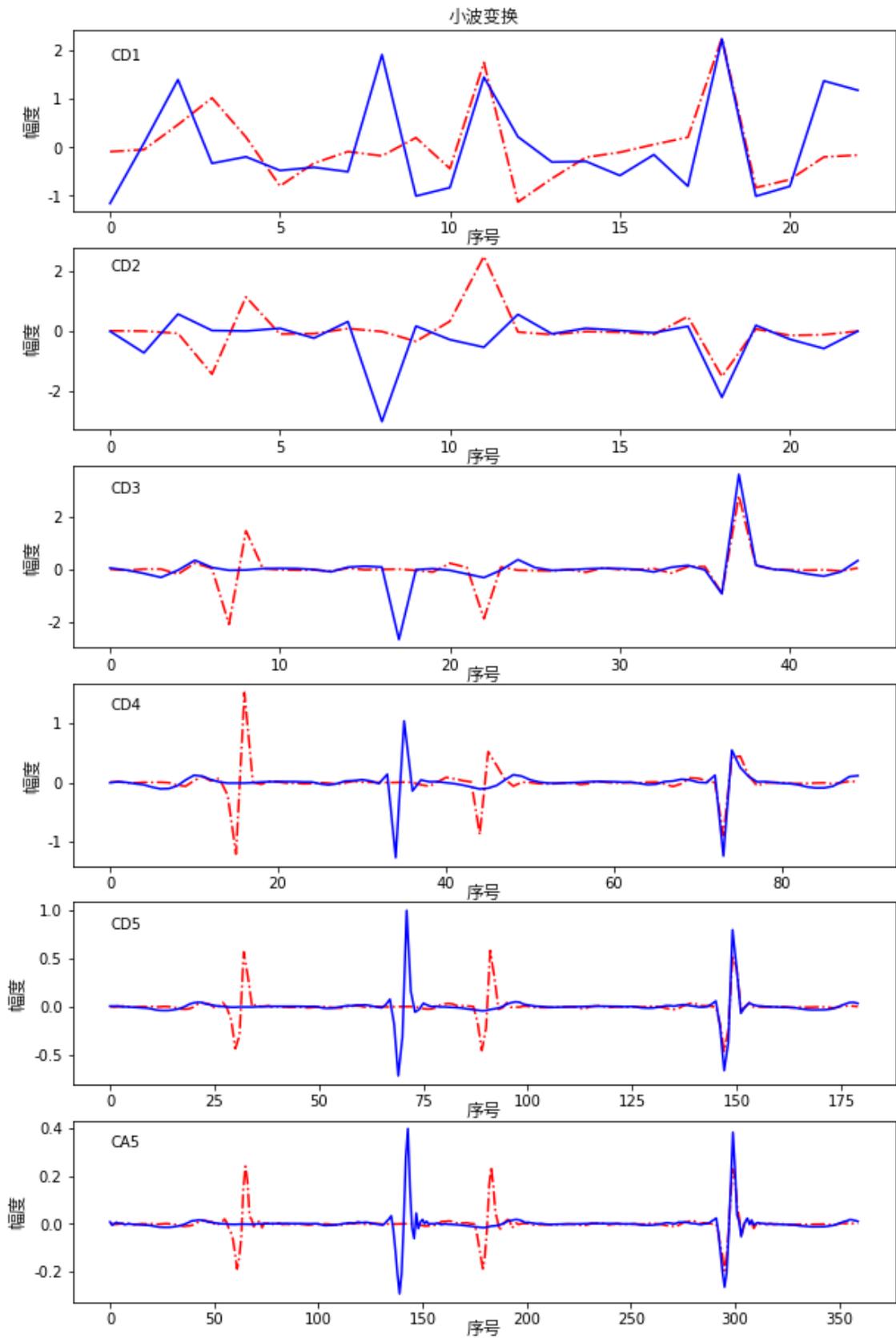


图 4-4 不同个体的 DWT 结果
Figure 4-4 Results of DWT from Different Person

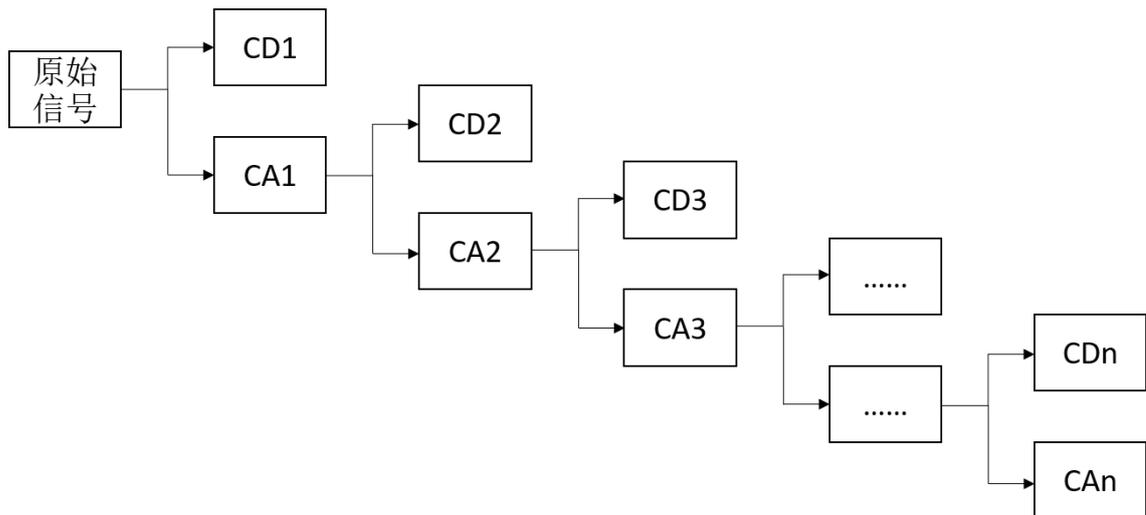


图 4-5 小波分解树
Figure 4-5 Wavelets Decomposition Tree

4.4 通道注意力模块的改进

在 3.2.5 小节中，本文引入了通道注意力模块，其结构如图 4-6 所示。由在 3.2.5 中已经介绍过了，本处不在对原文赘述。与原有的通道注意力相比，本文做了两处改动，分别把所有的激活函数都改为了 ReLU 函数，并删除了图中 2) 的张量叠加操作。但是在本章实验发现，模型在训练过程中，经常会出现 loss 周期性暴增的问题，这使得模型不能趋于稳定，难以确定是否已经达到了最优。经过研究发现，这是由于部分神经元失活导致的。为了解决神经元失活的问题，本文把 ReLU 激活函数换成了 ELU 函数，两个激活函数的区别如图 4-7 所示。

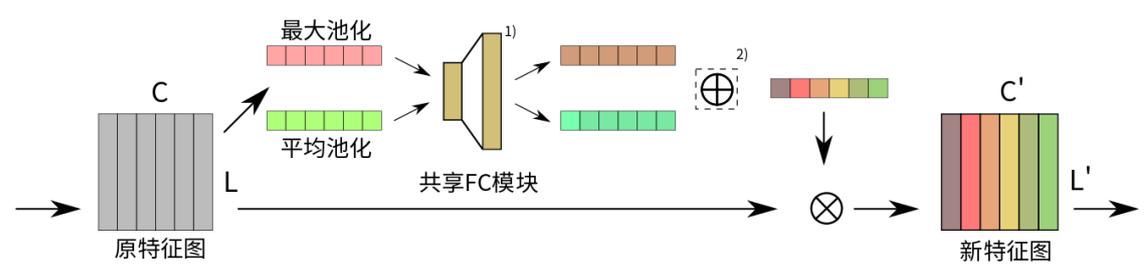


图 4-6 通道注意力模块
Figure 4-6 Channel Attention Module

在图 4-7 中可以看到，二者的主要区别是在负半轴，ReLU 函数的负半轴激活值都为 0，其梯度也为 0，这也是神经元失活的原因，即当神经元的梯度为 0 时，神经元不进行梯度更新了，从而产生了失活的现象。而 ELU 函数在负半轴还是有一个很小的梯度的，这使得所有的神经元都处在“工作”的状态，不需要担心大量的神经元失活，所以也就没有 loss 值暴涨的现象了。ELU 激活函数的映射关系如公

式(4-4)所示。

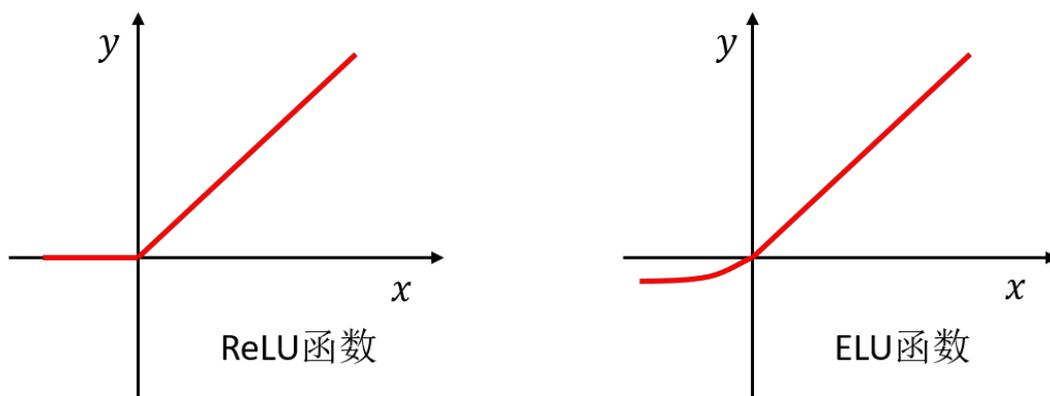


图 4-7 ReLU 和 ELU 激活函数
Figure 4-7 ReLU and ELU Activation Function

$$f(x) = \begin{cases} x, & x \geq 0 \\ \alpha(e^x - 1), & x < 0 \end{cases} \quad (4-4)$$

其中 α 是一个尺度参数，本文采用了默认值，也就是 1。除此之外，后面的实验中详细介绍在卷积神经网络的卷积模块的浅层、中层和深层分别使用通道注意力模块的异同点，并对不同的现象做出解释。

4.5 实验研究

本节将进行具体的实验研究。首先，简单介绍数据集，然后描述所设计实验网络结构，接着进行大量参数选取实验，最后，将结果与其他相关工作对比，进行分析。

4.5.1 实验网络设计

本节将介绍实验部分用到的网络结构。图 4-8 展示了本文的实验网络结构，首先，根据大量的实验，发现 3 层卷积模块（CONV）、2 层全连接模块（DENSE）和一个输出层，已经有足够的学习能力，虽然 4 层卷积在最后的分类准确率上有一定提升，但是提升效果不明显，同时带来的代价是训练时间成本和内存消耗的大幅度增加，综合考虑最后选择了 3 层卷积神经网络。后文在没有特别提及的情况下，都是使用的这个网络结构，当出现不一致的时候，将会给出详细的改动地方。

在图 4-8 中，除了基本的卷积神经网络，还加入了通道注意力模块（即 CAM），

但是并不是所有层的卷积都会使用 CAM（故图中用虚线链接），一般都会在最后一层卷积加上 CAM，当进行通道注意力模块位置实验时，本研究会尝试第一层或者第二层加入 CAM（同时，取消第三层的 CAM）。从图 4-8 可以看出，该网络结构是一个多输入结构，输入的支路数和离散小波变换的级数有关。卷积模块的组成与第三章实验网络所用的卷积模块组成相同，全连接模块与第三章用全连接模块也相同。

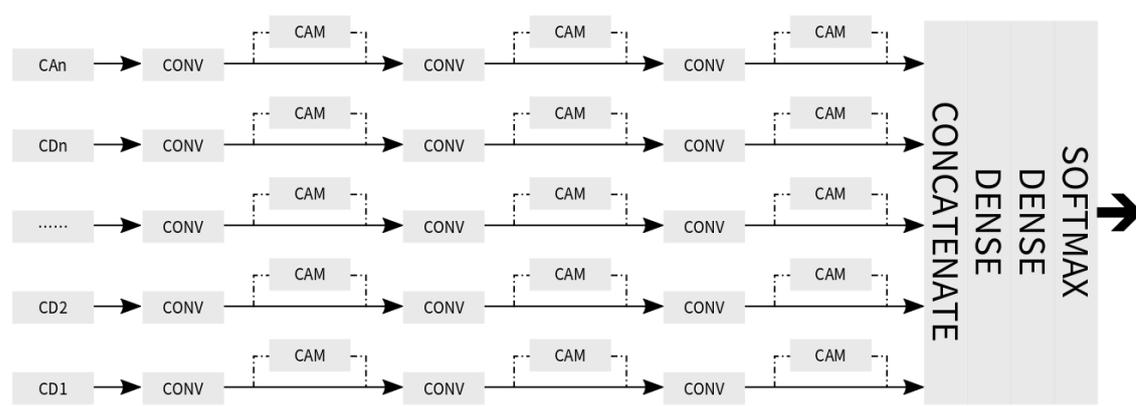


图 4-8 实验网络结构

Figure 4-8 Structure of Experimental Network

在所有的实验中，本文把训练的次数(epochs)设为了 100，学习率设成了 0.01，并且在输出层增加了一个权重为 0.002 的 L2 正则化。为了降低神经网络的随机性对实验结果的影响，每个实验都重复了十次，最后的结果取十次的均值。实验的评价标准是识别率，也就是深度学习中的准确率（Accuracy）指标。

4.5.2 参数选定实验

在 4.3 节中，阐述了小波变换的原理，同时也注意到小波变换级数是实验的一个重要参数，需要通过实验来选择。除此之外，还有很多实验参数需要选定，比如训练集每个类的样本数，每个样本的时长等等，本节将通过实验结果来选定。选定参数的流程主要如下：(1) 根据实际情况，对某一参数，先选定几个待选数值；(2) 对待选数值进行实验；(3) 根据实验结果选定最合适的值。下面，具体讲述部分参数的选取过程。

样本数量的选择。在样本生成阶段，每个类生成多少样本较为合适。显然，在一定的范围内，样本数越多，模型就能得到更好的训练，但是样本数越多，训练所需时间和存储空间也就越大，并且随着数据量的增加，效果提升也越来越不明显，并且由于抽样区间固定，样本数超出一定范围，可能会带来过拟合现象。所以，合适的样本数量显得十分重要。在选定样本数量之前，首先先预选取几个值，

受到计算资源的限制，预选定的样本数量值分别为 100,200,300,400。同时，选定了单个样本的时长为 4 秒，训练区间和测试区间都设定为 40 秒，同时训练样本数量和测试样本数量相同，通道注意力模块都只加在了第三层卷积上。图 4-9 展示了不同样本数量的训练结果。在图 4-9 中可以看到，样本数量从 100 增加到 200 和从 200 增加到 300，最后测试识别准确率增幅都很明显，但是从 300 增加到 400 时，测试准确率增幅已经非常不明显了，仅为从 200 增加到 300 时测试准确率增幅的 6%左右，这说明样本数到 300 时，这个参数已经达到了一个饱和程度，测试准确率提升的主要影响因素已经不是样本数量了，而是其他的因素了。另一方面，每当样本增加 100 个时，发现在全训练集上训练一次的训练时间也会对应增加 10s（具体的数值和机器、batch 大小有关），占用的内存也会相应的增加。综合测试准确率、训练时间、内存消耗的情况，本文选择 300 为样本数量，这样既能取得一个较好的测试准确率，又能有较低的训练时间和较小的内存消耗。

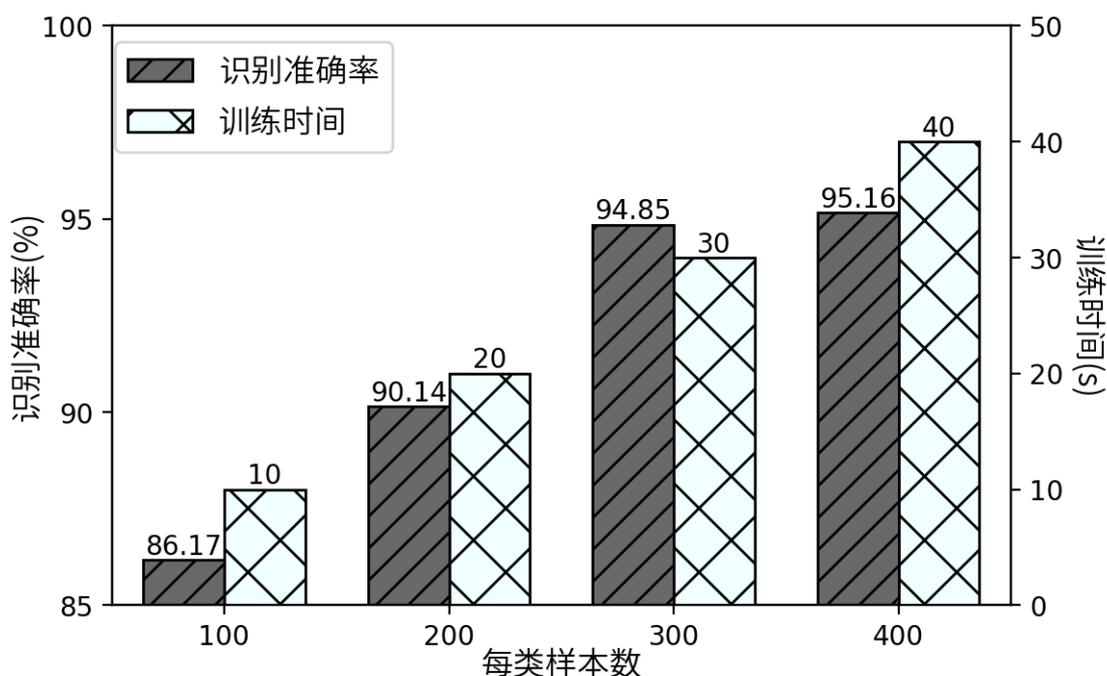


图 4-9 样本数量实验结果

Figure 4-9 Experimental Results of Sample Number

通道注意力模块的使用。为了探究通道注意力模块的作用以及用在何处能取得最好的效果，本文做了一个探究实验。通过不使用通道注意力模块和把通道注意力模块放在第一层卷积，第二层卷积和第三层卷积，来探究通道注意力模块在浅层、中层和深层带来的影响。本实验中，把样本时长设为 4 秒，训练区间和测试区间设为 120 秒，训练和测试数据每个类别都是 100 个，实验结果如表 4-1 所示。在表 4-1 中，可以看到，首先，没有通道注意力模块的识别准确率远远低于有通道

注意力模块的识别准确率，这证明了通道注意力模块的作用十分明显。其次，通道注意力模块用在深层时，效果最好。而通道注意力模块所在的层越浅，实验效果也越差。原因是，底层的卷积主要是提取低维度的特征，目的主要是提取有效的特征，但是这些特征不具有区分度，每个个体都具有这样的特征，例如，第一层卷积可能都是为了提取到 R 波尖峰的位置。随着卷积层的叠加，提取到的特征也更为抽象，区分度也更好，此时通道注意力模块发挥的作用也越大。

表 4-1 通道注意力模块实验结果
Table 4-1 Experimental Results of Channel Attention Module

模块位置	无	浅层 (第一层)	中层 (第二层)	深层 (第三层)
识别准确率	77.22%	85.40%	87.24%	90.38%

样本时长的选择。除了上述的样本数量，单个样本的时长也是一个很重要的参数。样本时长过短，不能包含一个完整的心拍，提取不到完整的特征，对识别准确率有很大影响，样本时长过长，测试者在接受测试时，等待时间也过长，可接受性较差，同时，过长的样本包含过多的心拍，导致噪声过多，提取到的特征不够稳定。所以，对样本时长的预选取值为 2 秒，4 秒，6 秒，8 秒。同时，训练样本取值为 100，测试样本也为 100 个，训练区间和测试区间都为 40 秒，通道注意力模块都放在第三层，最后的结果如表 4-2 所示。从表 4-2 中可知，并不是样本时长越长越好。时长越长，噪声也越多，经过小波变换之后，噪声的影响也更大。同时，2 秒时长和 4 秒时长之间的实验结果差距远小于 4 秒和 6 秒的实验结果差距，证明类似于样本数量变量，当样本时长小于 4 秒时，样本时长对实验结果的影响不是很大，但是一旦超过 4 秒，样本时长对实验结果的影响就很大。显然，对于测试者而言，2 秒的等待时间明显是优于 4 秒的，同时 2 秒时长的样本也有更好的表现。综上，本文把样本时长参数定为 2s。

表 4-2 样本时长实验结果
Table 4-1 Experimental Results of Sample Length

样本时长/s	2	4	6	8
识别准确率	88.98%	86.17%	82.81%	81.51%

其他参数用类似的方法进行选取，通过大量的实验，本文把训练区间和测试区间定为 40 秒，小波变换的级数选择了 5 级。至此，大部分需要实验来选定的参数都已经选定了。

4.5.3 相关工作对比

通过上一小节的大量实验，确定了大部分参数的最优值，在 MIT-BIH 数据集上，本文把单个样本时长的值定为 2 秒，训练区间和测试区间的长度都定为 40 秒，训练样本和测试样本每个类的数量都定为 300 个，并且又在测试区间上每个类抽取了 150 个样本作为验证集，通道注意力模块放在了最后一层卷积，小波变换级数定为 5 级。最终，对最优参数进行了十次实验，并对十次结果进行了平均，最终在 45 个个体的系统中，识别准确率为 95.86%。

图 4-10 展示了某一次实验的混淆矩阵，从图中可以看出，绝大多数的样本都被正确识别了，只有一两个个体的极少数样本被错误分类，证明了模型有较高的识别性能。表 4-3 展示了本文和其他相关工作的对比。从表中可以看出，本文的工作有比较明显的优势，需要注意的是，本文之所以比其他工作少 2 个个体，是因为这两个个体数据(记录号 102, 104)不具备心电图导联II，和其他数据不是一个导联，缺乏训练的实际意义，所以本文舍弃了这两条数据。但是这对本文提出的系统的影响并不大，下一小节将介绍算法的稳定性，在一定范围内，系统总人数对算法性能的影响较小。同时可知，提出的算法对心律异常是不敏感的，因为在 MIT-BIH 心率异常数据库上，本文仍能达到接近 96%的识别准确率，就算在心律正常数据库上本文的算法能达到 100%的性能，中间也只有 4%的差距，这样的波动是完全可以接受的，从而证明了本文的算法对心律异常是低敏感度的。

表 4-3 相关工作对比
Table 4-3 Comparison of Related Work

方法	年份	识别总人数	识别准确率
文献[23]	2012	47	74.17%
文献[34]	2015	47	93.1%
文献[27]	2017	47	83%
文献[36]	2017	47	91.1%
本文	—	45	95.86%

以上是和一些在 MIT-BIH 心律失常数据库上总人数规模相似的工作进行比较，表 4-4 是和一些工作相似度较低的文献进行的比较。从表中可以看出，现有大部分研究都是基准的方法，这也使得大部分系统在构建时，特意选取了健康的个体，造成了很多算法对异常心律十分敏感。并且，很少有工作对分类器进行过改进。同时，自建数据库的个体数都普遍偏少。表中，文献[24]识别准确率要好于本文的工作，但是他们的识别总人数只有 33 人，同时是健康的个体，我们用本文的

方法做了 33 个健康个体的实验，识别准确率为 97.59%，虽然比该工作略差，考虑到，本文的算法在含有心律异常个体时，识别准确率仍较高（30 个人为 96.77%），对心律异常仍不敏感，而该工作对于心律异常的敏感度尚不知道，因为他们的算法是基于健康个体设计的，所以本文的算法仍有优势。

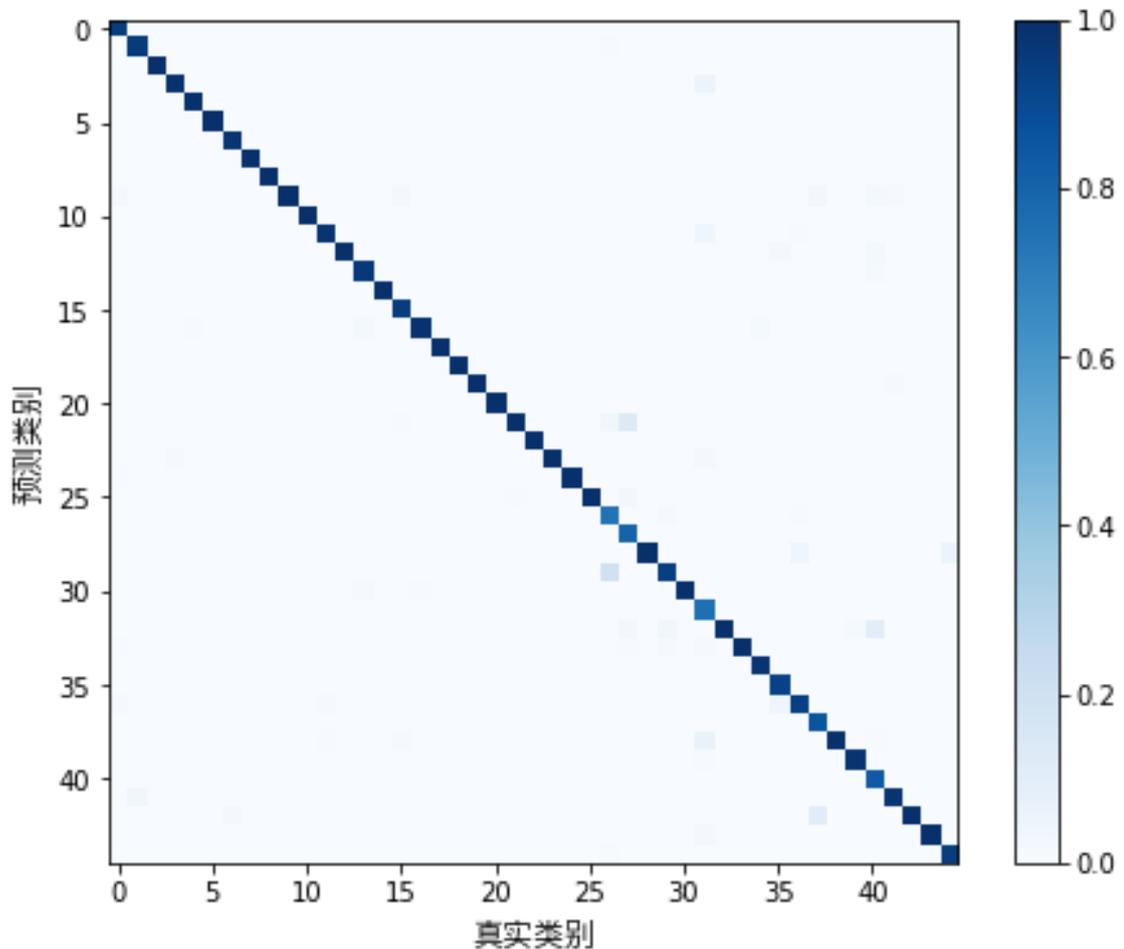


图 4-10 心电生物识别混淆矩阵（45 人）
Figure 4-10 Confusion Matrix of ECG Biometrics(45 Subjects)

表 4-4 其他工作对比
Table 4-4 Comparison of Other Work

方法	数据来源	数据规模	含有心律异常个体	样本生成方法	分类器改进	识别准确率
文献[24]	MIT	33	否	基准方法	无	99.64%
文献[21]	自建	13	否	基准方法	无	90%
文献[31]	自建	18	否	基准方法	无	88.97%
文献[32]	PTB	38	未知	基准方法	无	73.60%
本文	MIT	45	是	非基准方法	有	95.86%

4.5.4 算法稳定性分析

个体识别系统除了需要实现较高的识别准确率，还需要有较高的稳定性，即识别系统总人数的多少对系统的识别准确率不会有太大的影响。为了验证提出的算法的稳定性，本文设计了不同人数的个体识别系统，即 10, 20, 30, 40 个个体的系统，并对结果进行了分析。在图 4-11 中，展示了不同人数的个体识别系统的识别准确率。从图中可以看出，虽然随着总人数的增加，系统的识别准确率虽然有一定的下降，但是下降的范围很小，不到 4%，这是正常的，因为随着类别的增加，类间距离逐渐变小，某些异常的样本更有可能被错误识别。因为最后的识别准确率没有比较大的下降，所以提出的算法有比较高的稳定性，在一定范围内，对系统总人数不敏感。

图 4-12 至图 4-15 展示了 10, 20, 30, 40 个个体时的混淆矩阵。从图中可以看出，在人数较少时，几乎所有的样本都被正确识别了，只有当人数较多时，才有极个别的个体有少量的样本识别失败，这与系统人数增加，类间距离减少和某些个体的一些心律异常都有一定关系。但是从图中可以看出，只有很少的样本被错误识别了，这可以仍可以说明本文提出的算法在一定范围内对系统总人数和心律异常是低敏感的。

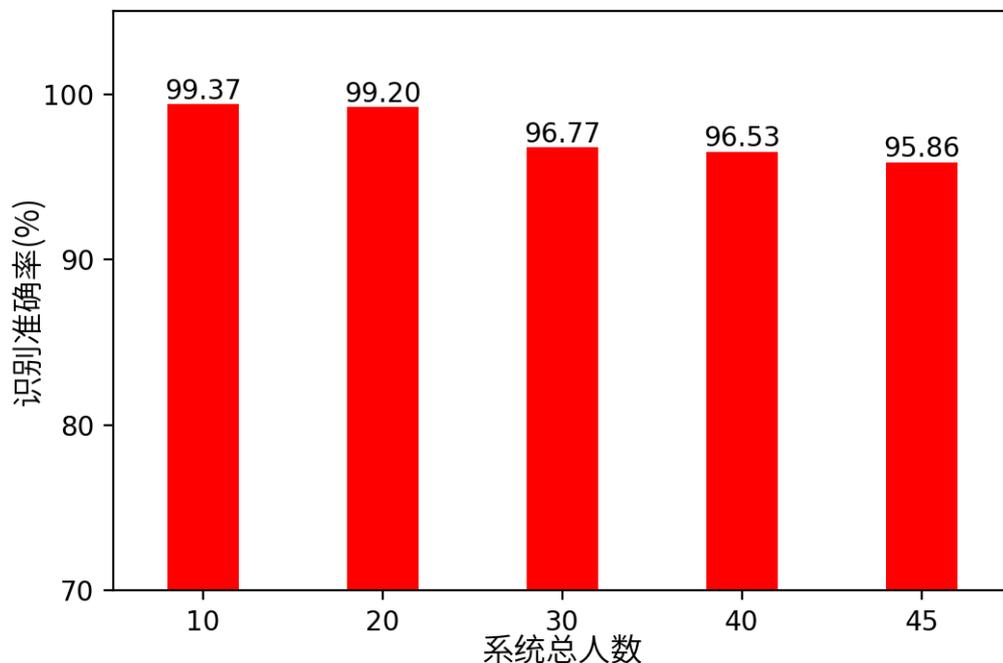


图 4-11 总人数实验结果

Figure 4-11 Experimental Results of Subject Number

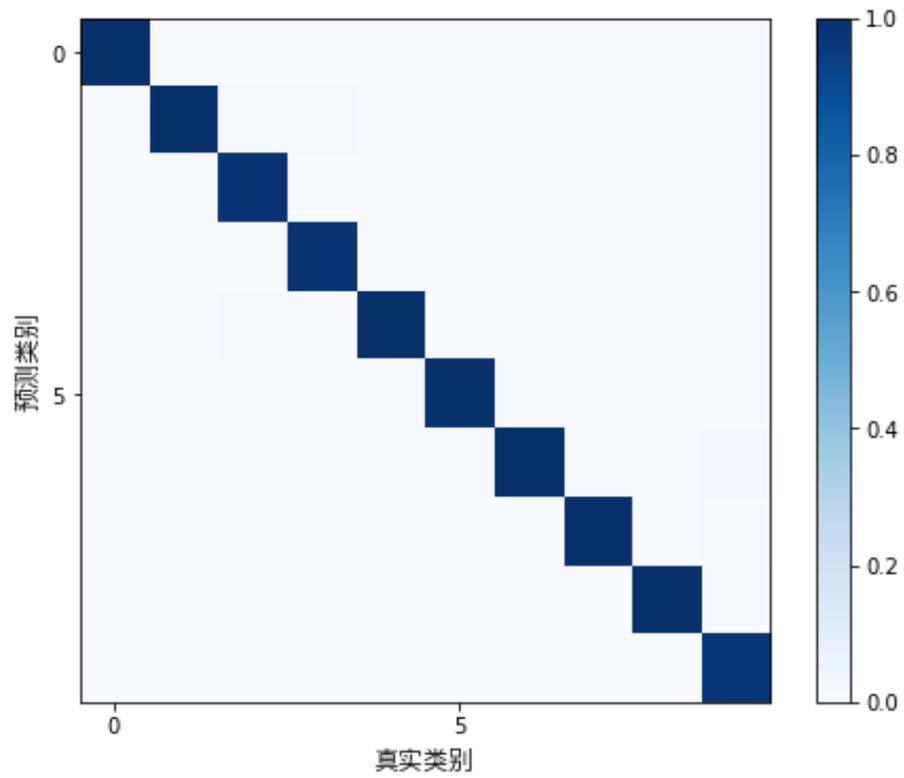


图 4-12 心电生物识别混淆矩阵（10 人）
Figure 4-12 Confusion Matrix of ECG Biometrics(10 Subjects)

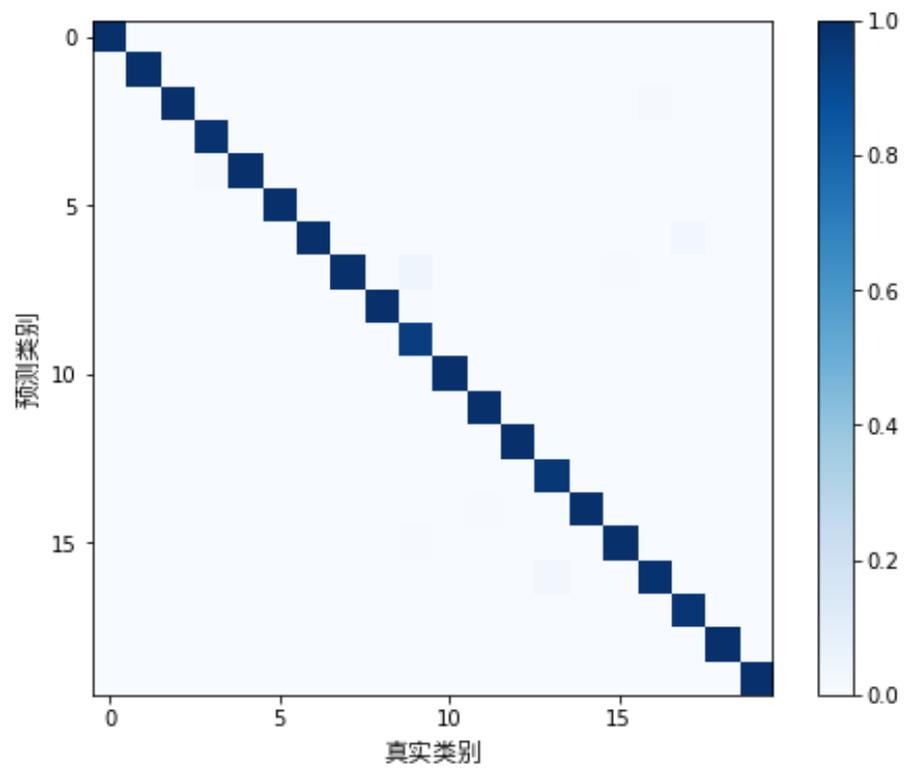


图 4-13 心电生物识别混淆矩阵（20 人）
Figure 4-13 Confusion Matrix of ECG Biometrics(20 Subjects)

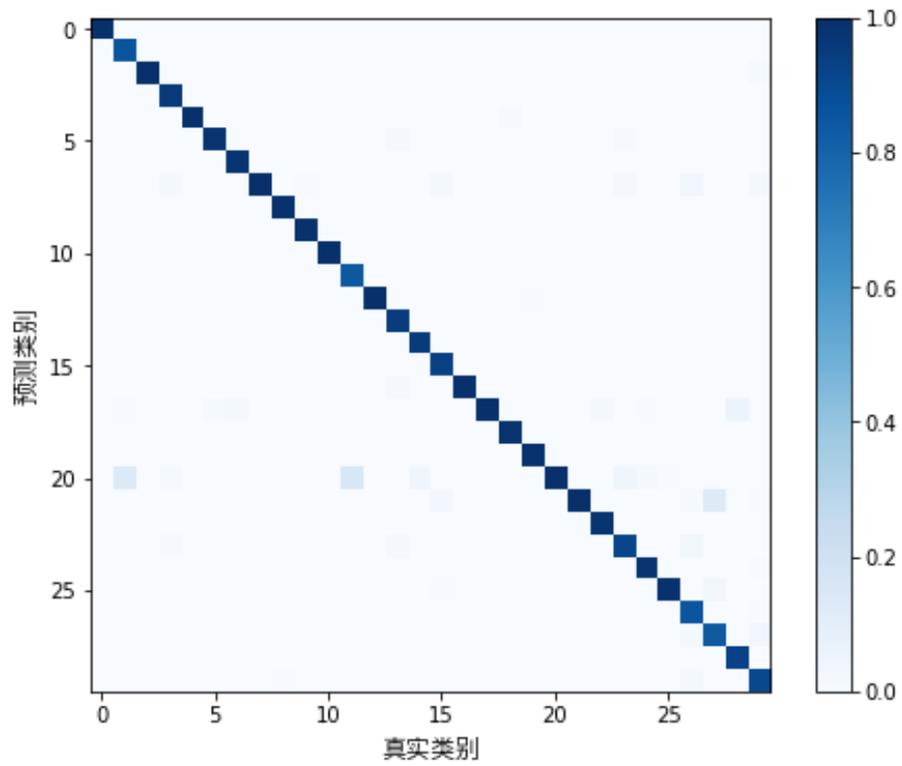


图 4-14 心电生物识别混淆矩阵（30 人）
Figure 4-14 Confusion Matrix of ECG Biometrics(30 Subjects)

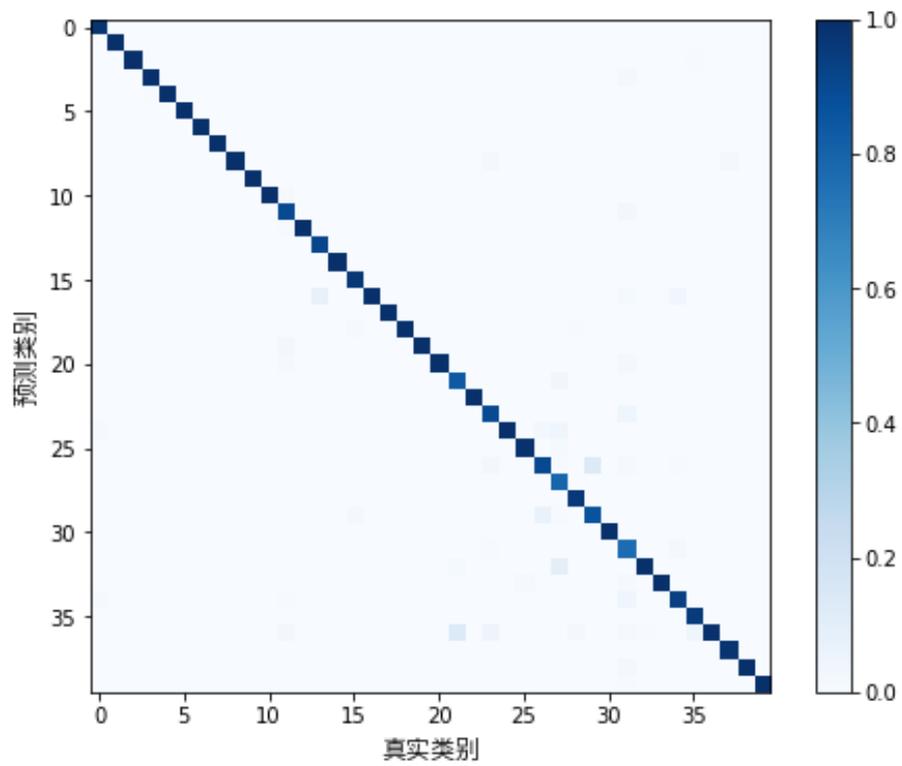


图 4-15 心电生物识别混淆矩阵（40 人）
Figure 4-15 Confusion Matrix of ECG Biometrics(40 Subjects)

4.6 本章小结

本章，本文工作主要体现在以下几个方面。

首先，本文对现有心电个体识别系统进行了总结，分析了现有心电个体识别系统的局限与不足，主要体现为以下几点。(1) 目前各研究所使用数据库质量层次不齐，实验结果难以复现。(2) 样本生成多以一个心拍为一个样本，由于心电图记录十分有限，导致样本也非常有限，而神经网络需要大量的样本训练，才能得到一个较好的模型，这一突出矛盾并未解决。除此之外，这种以心拍为样本的算法对心律异常都比较敏感，实用性大打折扣。(3) 目前数据表达的方式有很多种，但是很多的表达方式信息损失较大，选择一种合适的表达方法很重要。(4) 虽然很多研究使用了神经网络，但是都是用一些基本的神经网络，并未针对本问题进行结构改善。(5) 虽然很多研究都取得了较好的成果，但是系统稳定性较差，一旦加入心律异常个体或者个体数过多，很多研究成果对心律异常和系统总人数非常敏感。

其次，为了解决现有的心电个体识别系统的局限和不足，本文提出了以下的算法改进心电个体识别系统。(1) 在数据来源上，选择公开的数据集。本文选用了 MIT-BIH 心律异常数据集，该数据集具有较高质量，并且易于获取。除此之外，该数据库拥有较多个体和大量的心律异常个体，方便检测算法对心律异常和系统总人数的敏感度。(2) 在样本生成方法上，本文用随机抽样的方法替代了单个心拍即单个样本的方法。这样理论上可以抽取无限多的样本，根据实际需要，抽取一定数目，解决了样本不足的问题。(3) 在数据表达上，选用了经过验证十分有效的小波变换。相较于傅里叶变换，小波变换用的是长度变换的窗口，在时域和频域上，都有较好的分辨率。(4) 为了解决卷积神经网络的通道信息提取能力有限的问题，引入了通道注意力模块，并对该模块进行了改进。(5) 通过改变系统总人数和加入大量心律异常的个体，可以验证算法对系统总人数和异常心律的敏感度。

最后，本文在 MIT-BIH 心律异常数据库上，进行了大量实验，验证了本文提出的算法的有效性与低敏感度，并且和先有工作做了大量对比，从各个维度来看，本文工作的成果都是优于现有工作成果的。

5 结论

5.1 本文工作总结

本节主要是对本文工作成果进行细致的总结归纳，关键性的研究成果主要分为两个部分：基于深度学习的心电图类型识别研究，以及基于深度学习的心电个体识别研究。下面，将对这两部分进行重点阐述。

5.1.1 心电图类型识别研究总结

本节主要总结本文在心电图类型识别方面的工作成果，具体的贡献如下：

(1) 通过大量阅读文献和学习基本医学知识，结合生活常识，发现了现有研究的几个突出问题：心拍切分都采取的固定长度切分，忽略了每个人的心拍长度不一致，同时，同一个人不同场景下心拍长度也会发生变化；在特征提取时，现有研究多是基于一个心拍去提取特征的，但是医生判断一个心拍是否正常，往往需要结合前后心拍特征，还有当前心拍各个子拍（PQRST波）的特征；在分类器的选择上，现有工作多是简单的选择一个卷积神经网络或者递归神经网络，没有考虑到该网络是否和当前研究问题符合，是否有改进的空间和价值。针对上述问题，本文对各部分工作都做了大量改进。

(2) 在心拍切分上，本文用自适应的动态切分来代替固定长度切分。首先用动态RR间隔来表示一个心拍长度，通过平移操作来获取一个动态的真实的心拍。同时，为了满足特征提取和分类器要求输入是一个长度和维度固定的数据，本文提出了一个填充和截断的方法，以使心拍长度从非固定的转换到固定的。同时，为了能够提取多尺度的特征，本文提取了子拍，即P波，QRS波群和T波，以及多拍。

(3) 在特征提取上，本文提出了多尺度特征的提取，主要包括手动多尺度特征的提取和深度多尺度特征的提取。手动多尺度特征的提取主要是结合医生判断心电图类别的信息，在时间和幅度上，对子拍和单拍提取多个特征。深度多尺度特征主要是通过深度学习（一维卷积神经网络）对子拍、单拍和多拍提取深度特征。这两部分特征可以在分类时起到相互弥补的作用，既结合了医学特征的稳定性，又结合了深度特征的高维度和抽象性。

(4) 在分类器上，本文选用了一维卷积神经网络，但是考虑到神经网络对通道

之间的信息学习能力较差，而心电图具有丰富的通道信息这一矛盾，引入了通道注意力模块，并针对本研究的具体问题，对通道注意力模块进行了改进。最后，在 MIT-BIH 心律失常数据库上，按照 AAMI 建议的五分类和数据集划分标准，验证了本文提出的算法，并取得了 96.94% 的分类准确率，同时，各类别的精确率、召回率、F1 值都有提升，尤其是在样本较少的类别上，这三类指标的值有大幅度的提升。

以上研究成果，对于进一步提升心电图类型自动识别系统的性能有很大的帮助，同时，也为我们继续研究心电图类型自动识别系统提供了方向。除此之外，本文的算法有着较好的性能，对于实时数据也可以精确预测，对实时心电图类型自动识别系统也有较大的实际意义。

5.1.2 心电图个体识别研究总结

本节主要总结本文在心电图个体识别方面研究的工作成果，具体的贡献如下：

(1) 通过大量阅读文献和对个体识别系统的深入了解，本文对现有心电个体识别系统的主要问题做了总结：数据获取渠道杂乱，数据质量参差不齐，导致很多算法根本无法验证和复现其效果；样本生成方法多为用一个心拍作为一个样本（即基准的方法），但是心电数据十分有限，导致样本数量不足，分类器欠拟合，并且检测心拍本身就是一个极其困难的过程；特征提取的方法虽然有很多，但是很多方法提取到的特征都不稳定，尤其是当系统中有个体患有心脏疾病时，这样的特征提取方法容易导致模型不稳定；分类器的选择过于简单，没有选择合适的分类器，针对具体的问题，也没有对分类器进行改善。针对上述问题，本文提出了新的心电个体识别算法，对现有工作做了大量的改进。

(2) 针对数据来源混乱的问题，我们决定在验证算法时，使用公开的且被大量使用的数据集，比如 MIT-BIH 心律失常数据库。使用该数据库还有一个优点是，该数据库含有大量的患有心脏疾病的病人，通过最后的实验结果，可以验证算法是否对心律异常的个体敏感，如果不敏感，这说明算法具有较好的稳定性。

(3) 在样本生成阶段，本文用随机抽样的方法代替了基准的方法。即在一个指定的时间区间内，用固定的时长去抽样，通过多次抽样，就可以获取大量的样本。而且，这种获取样本的方法省去了检测心拍的过程，减少了大量的计算和可能带来的误差。同时，对于神经网络分类器而言，可以得到充分的训练。

(4) 抽样方法获取的样本在时域上规律比较差，为了得到较好的数据表达。本文引入了小波变换，相比于傅里叶变换，小波变换在时域和频域都有较好的分辨率，可以得到更多更有效的特征。

(5) 分类器的选择是一维卷积神经网络。为了改善神经网络较差的通道信息提取能力，本文同样引入了通道注意力模块，并且，相比于心电图类型识别，针对心电个体识别，本文对通道注意力模块做了不同的改变。最后，在 MIT-BIH 数据集上进行了验证，通过大量的实验，包括参数选择实验和人数实验，本文的算法在 MIT-BIH 心律失常数据库上，当系统总人数是 45 个时，识别准确率为 95.86%。同时，本文还通过改变系统总人数，验证了提出的算法的稳定性。

以上的研究成果，对于心电图个体识别系统的研究有很大的帮助。同时，本文提出了心电图个体识别算法应该对心律失常和系统总人数不敏感或者敏感性较低，也为之后的心电图个体系统识别提供了新的研究方向。

5.2 未来工作展望

随着深度学习技术的不断发展以及医学水平的不断提升，人们对心脏健康会有越来越高的要求。另一方面，随着人们压力不断增大以及环境的不断变化，心血管疾病可能也会有越来越多的变种。同时，心电图个体识别由于其数据安全性较高，难以复制，也会得到越来越多的关注的。于是，基于深度学习的心电数据研究还将继续是一个热点问题。因此，作者未来的工作重点会继续放在心电图类型识别和个体识别上：对于心电类型识别，设计更多的异常类型，细化心律异常的类型，使得系统功能更加强大；对于心电个体识别，设计容纳更多个体，并且对心律失常敏感度更低的识别系统，加大其商用化的可能性。

参考文献

- [1] Jun T J, Nguyen H M, Kang D, et al. ECG arrhythmia classification using a 2-D convolutional neural network[J]. arXiv preprint arXiv:1804.06812, 2018.
- [2] Cimen E, Ozturk G. Arrhythmia classification via k-means based polyhedral conic functions algorithm[C]//2016 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2016: 798-802.
- [3] Kiranyaz S, Ince T, Gabbouj M. Real-time patient-specific ECG classification by 1-D convolutional neural networks[J]. IEEE Transactions on Biomedical Engineering, 2016, 63(3): 664-675.
- [4] Zubair M, Kim J, Yoon C. An automated ECG beat classification system using convolutional neural networks[C]//2016 6th international conference on IT convergence and security (ICITCS). IEEE, 2016: 1-5.
- [5] Ye C, Kumar B V K V, Coimbra M T. Combining general multi-class and specific two-class classifiers for improved customized ECG heartbeat classification[C]//Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). IEEE, 2012: 2428-2431.
- [6] Kropf M, Hayn D, Schreier G. ECG classification based on time and frequency domain features using random forests[C]//2017 Computing in Cardiology (CinC). IEEE, 2017: 1-4.
- [7] Cruz C I M, Marasigan J P, Perez A P G, et al. A comparative study between DWT-ANFIS and DWT-SVM in ECG classification[C]//2016 IEEE Region 10 Conference (TENCON). IEEE, 2016: 2052-2055.
- [8] Ebrahimzadeh A, Ahmadi M, Safarnejad M. Classification of ECG signals using Hermite functions and MLP neural networks[J]. Journal of AI and Data Mining, 2016, 4(1): 55-65.
- [9] Naseer N, Nazeer H. Classification of normal and abnormal ECG signals based on their PQRST intervals[C]//2017 International Conference on Mechanical, System and Control Engineering (ICMSC). IEEE, 2017: 388-391.
- [10] Ye C, Kumar B V K V, Coimbra M T. Heartbeat classification using morphological and dynamic features of ECG signals[J]. IEEE Transactions on Biomedical Engineering, 2012, 59(10): 2930-2941.
- [11] Limam M, Precioso F. Atrial fibrillation detection and ECG classification based on convolutional recurrent neural network[C]//2017 Computing in Cardiology (CinC). IEEE, 2017: 1-4.
- [12] Li T, Zhou M. ECG classification using wavelet packet entropy and random forests[J]. Entropy, 2016, 18(8): 285.
- [13] Smíšek R, Hejč J, Ronzhina M, et al. SVM based ECG classification using rhythm and morphology features, cluster analysis and multilevel noise estimation[C]//2017 Computing in Cardiology (CinC). IEEE, 2017: 1-4.
- [14] Safie S, Yusof M I, Kadir K, et al. Multiple pulse K-nearest neighbors authentication for malay ECG based class attendance system[C]//2014 4th International Conference on Engineering Technology and Technopreneuship (ICE2T). IEEE, 2014: 156-160.

- [15] Vishwa A, Lal M K, Dixit S, et al. Classification of arrhythmic ECG data using machine learning techniques[J]. IJIMAI, 2011, 1(4): 67-70.
- [16] Kiranyaz S, Ince T, Hamila R, et al. Convolutional neural networks for patient-specific eeg classification[C]//2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2015: 2608-2611.
- [17] Xiang Y, Luo J, Zhu T, et al. ECG-based heartbeat classification using two-level convolutional neural network and RR interval difference[J]. IEICE TRANSACTIONS on Information and Systems, 2018, 101(4): 1189-1198.
- [18] Warrick P, Homsy M N. Cardiac arrhythmia detection from ECG combining convolutional and long short-term memory networks[C]//2017 Computing in Cardiology (CinC). IEEE, 2017: 1-4.
- [19] Jiang W, Kong S G. Block-based neural networks for personalized ECG signal classification[J]. IEEE Transactions on Neural Networks, 2007, 18(6): 1750-1761.
- [20] Porée F, Gallix A, Carrault G. Biometric identification of individuals based on the ECG. Which conditions?[C]//2011 Computing in Cardiology. IEEE, 2011: 761-764.
- [21] Zheng G, Zhang H Y, Ji S Z. ECG based identification under different physical status[C]//2018 International Conference on Machine Learning and Cybernetics (ICMLC). IEEE, 2018, 1: 197-203.
- [22] Lee J N, Kwak K C. Personal identification using a robust eigen ECG network based on time-frequency representations of ECG signals[J]. IEEE Access, 2019.
- [23] Chen H, Tseng K K, Zeng F, et al. A new ECG identification with neural network[C]//2012 12th International Conference on Hybrid Intelligent Systems (HIS). IEEE, 2012: 427-430.
- [24] Wu J J, Zhang Y. ECG identification based on neural networks[C]//2014 11th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP). IEEE, 2014: 92-96.
- [25] Naraghi M E, Almasi A, Shamsollahi M B. Human identification using ECG feature extracted from innovation signal of extended kalman filter[C]//2012 5th International Conference on BioMedical Engineering and Informatics. IEEE, 2012: 545-549.
- [26] Naraghi M E, Shamsollahi M B. ECG based human identification using wavelet distance measurement[C]//2011 4th International Conference on Biomedical Engineering and Informatics (BMEI). IEEE, 2011, 2: 717-720.
- [27] He C, Li W, Chik D. waveform compensation of ECG data using segment fitting functions for individual identification[C]//2017 13th International Conference on Computational Intelligence and Security (CIS). IEEE, 2017: 475-479.
- [28] Bashar M K, Ohta Y, Yoshida H. ECG-based biometric authentication using multiscale descriptors: ECG-based biometric authentication[C]//2015 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS). IEEE, 2015: 1-4.
- [29] Khoma V, Pelc M, Khoma Y. Artificial neural network capability for human being identification based on ECG[C]//2018 23rd International Conference on Methods & Models in Automation & Robotics (MMAR). IEEE, 2018: 479-482.
- [30] Zhang Z, Wei D. A new ECG identification method using bayes' theorem[C]//TENCON 2006-2006 IEEE Region 10 Conference. IEEE, 2006: 1-4.
- [31] Wiclaw L, Khoma Y, Fałat P, et al. Biometric identification from raw ECG signal using deep

- learning techniques[C]//2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). IEEE, 2017, 1: 129-133.
- [32] Tantawi M, Revett K, Tolba M F, et al. A novel feature set for deployment in ECG based biometrics[C]//2012 Seventh International Conference on Computer Engineering & Systems (ICCES). IEEE, 2012: 186-191.
- [33] Ferdinando H, Seppänen T, Alasaarela E. Bivariate empirical mode decomposition for ECG-based biometric identification with emotional data[C]//2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017: 450-453.
- [34] Dar M N, Akram M U, Usman A, et al. ECG biometric identification for general population using multiresolution analysis of DWT based features[C]//2015 Second International Conference on Information Security and Cyber Forensics (InfoSec). IEEE, 2015: 5-10.
- [35] Xu J, Li T, Chen Y, et al. Personal identification by convolutional neural network with ECG signal[C]//2018 International Conference on Information and Communication Technology Convergence (ICTC). IEEE, 2018: 559-563.
- [36] Zhang Q, Zhou D, Zeng X. HeartID: A multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications[J]. IEEE Access, 2017, 5: 11805-11816.
- [37] 付乾坤. 基于心电信号的生物识别技术研究[D]. 哈尔滨工业大学, 2015.
- [38] 杨雷. 基于心电信号的身份识别算法研究[D]. 杭州电子科技大学, 2012.
- [39] Wu Y T, Wu H T, Chang W W. Compressed domain ECG biometric identification using JPEG2000[C]//2015 12th International Joint Conference on e-Business and Telecommunications (ICETE). IEEE, 2015, 5: 5-13.
- [40] Pathoumvanh S, Airphaiboon S, Prapochanung B, et al. ECG analysis for person identification[C]//The 6th 2013 Biomedical Engineering International Conference. IEEE, 2013: 1-4.
- [41] Saechia S, Koseeyaporn J, Wardkein P. Human identification system based ECG signal[C]//TENCON 2005-2005 IEEE Region 10 Conference. IEEE, 2005: 1-4.
- [42] Rezgui D, Lachiri Z. Integrating EMD attributes for person identification from electrocardiographic signals[C]//2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). IEEE, 2016: 478-482.
- [43] Fattah S A, Shahnaz C, Jameel A S M M, et al. ECG signal based human identification method using features in temporal and wavelet domains[C]//TENCON 2012 IEEE Region 10 Conference. IEEE, 2012: 1-4.
- [44] Chen H, Zeng F, Tseng K K, et al. ECG human identification with statistical support vector machines[C]//2012 International Conference on Computing, Measurement, Control and Sensor Network. IEEE, 2012: 237-240.
- [45] Zhao Z, Yang L. ECG identification based on matching pursuit[C]//2011 4th International conference on biomedical engineering and informatics (BMEI). IEEE, 2011, 2: 721-724.
- [46] Liu J, Yin L, He C, et al. A multiscale autoregressive model-based electrocardiogram identification method[J]. IEEE Access, 2018, 6: 18251-18263.

- [47] Loong J L C, Swee S K, Bear R, et al. Effects of diseased ECG on the robustness of ECG biometric systems[C]//2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES). IEEE, 2010: 307-310.
- [48] Byeon Y H, Pan S B, Kwak K C. Intelligent deep models based on scalograms of electrocardiogram signals for biometrics[J]. Sensors, 2019, 19(4): 935.
- [49] 佚名. 心脏结构图. <https://www.wendangwang.com/doc/369d58eaf493da10a1612e27>. 2013.
- [50] 樊承柱. 基于深度神经网络的心电图自动分类方法研究与实现[D]. 山东大学, 2016.
- [51] 王晓海, 吴志刚. 多层感知机 (1). <http://book.51cto.com/art/201302/380103.htm>. 2013.
- [52] Ian Goodfellow, Yoshua Bengio, Aaron Courville [著]. 赵申剑, 黎彧君, 符天凡等[译]. 深度学习[M]. 北京:人民邮电出版社,2017.
- [53] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
- [54] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
- [55] Pan J, Tompkins W J. A real-time QRS detection algorithm[J]. IEEE Trans. Biomed. Eng, 1985, 32(3): 230-236.
- [56] Tang X, Shu L. Classification of electrocardiogram signals with RS and quantum neural networks[J]. International Journal of Multimedia and Ubiquitous Engineering, 2014, 9(2): 363-372.
- [57] Acharya U R, Oh S L, Hagiwara Y, et al. A deep convolutional neural network model to classify heartbeats[J]. Computers in biology and medicine, 2017, 89: 389-396.
- [58] Zhai X, Tin C. Automated ECG classification using dual heartbeat coupling based on convolutional neural network[J]. IEEE Access, 2018, 6: 27465-27472.

作者简历及攻读硕士/博士学位期间取得的研究成果

一、作者简历

陈滨，男，1994年3月生。2012年9月至2016年6月，就读于北京交通大学电子信息工程学院通信工程专业，取得工学学士学位。2016年9月至2019年6月，就读于北京交通大学电子信息工程学院通信与信息系统专业，取得工学硕士学位。

二、发表论文

- [1] Bin Chen, Yuchun Guo, Yishuai Chen, et al. "ECG Classification Based on Unfixed-Length Segmentation of Heartbeat" [C]//IEEE International Conference on Consumer Electronics-taiwan. 2019.

三、参与科研项目

- [1] 社会化信息网络的需求感知与需求整形，国家自然科学基金面上项目，No.61572071

独创性声明

本人声明所提交的学位论文是本人在导师指导下进行的研究工作和取得的研究成果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京交通大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名： 陈斌 签字日期： 2013年 5月 27日

学位论文数据集

表 1.1: 数据集页

关键词*	密级*	中图分类号	UDC	论文资助
类型识别, 个体识别, 心拍切分, 特征提取, 通道注意力模块	公开			国家自然科学基金 No.61572071
学位授予单位名称*		学位授予单位代码*	学位类别*	学位级别*
北京交通大学		10004	工学	硕士
论文题名*		并列题名		论文语种*
基于深度学习的心电图类型识别和个体识别				中文
作者姓名*	陈滨		学号*	16120044
培养单位名称*		培养单位代码*	培养单位地址	邮编
北京交通大学		10004	北京市海淀区西直门外上园村 3 号	100044
学科专业*		研究方向*	学制*	学位授予年*
通信与信息系统		信息网络	3 年	2019
论文提交日期*	2019.06.03			
导师姓名*	郭宇春		职称*	教授
评阅人	答辩委员会主席*		答辩委员会成员	
	李纯喜		赵永祥, 郑宏云, 张立军, 陈一帅	
电子版论文提交格式 文本 () 图像 () 视频 () 音频 () 多媒体 () 其他 () 推荐格式: application/msword; application/pdf				
电子版论文出版 (发布) 者		电子版论文出版 (发布) 地		权限声明
论文总页数*	63			
共 33 项, 其中带*为必填数据, 为 21 项。				