

分类号： TP183

单位代码： 10431

密 级：

学 号： 10431200544

齐鲁工业大学

# 硕士学位论文

## (专业学位)

基于二维化十二导联 ECG 信号分类  
与冗余性研究

作 者 姓 名 \_\_\_\_\_ 李家豪 \_\_\_\_\_

领 域 \_\_\_\_\_ 电子信息 \_\_\_\_\_

所 在 学 部 \_\_\_\_\_ 电子电气与控制学部 \_\_\_\_\_

指导教师姓名 \_\_\_\_\_ 庞少鹏 \_\_\_\_\_

专业技术职务 \_\_\_\_\_ 副教授 \_\_\_\_\_

2023 年 05 月 31 日

分类号： TP183

单位代码： 10431

密 级：

学 号： 10431200544

齊魯工業大學

# 硕士学位论文

## (专业学位)

基于二维化十二导联 ECG 信号分类  
与冗余性研究

作 者 姓 名 \_\_\_\_\_ 李家豪 \_\_\_\_\_

领 域 \_\_\_\_\_ 电子信息 \_\_\_\_\_

所 在 学 部 \_\_\_\_\_ 电子电气与控制学部 \_\_\_\_\_

指导教师姓名 \_\_\_\_\_ 庞少鹏 \_\_\_\_\_

专业技术职务 \_\_\_\_\_ 副教授 \_\_\_\_\_

2023 年 05 月 31 日

**A Thesis Submitted for the Application of  
the Master's Degree of Engineering**

**Research on the Classification and  
Redundancy of 12-lead ECG Signals Based  
on Two-dimensional**

**Candidate: Li Jiahao**

**Specialty: Electronic Information**

**Supervisor: Professor Pang Shaopeng**

**Qilu University of Technology, Jinan, China**

**May, 2023**

## 学位论文独创性声明

本人声明，所呈交的学位论文系在导师指导下本人独立完成的研究成果。文中引用他人的成果，均已做出明确标注或得到许可。论文内容未包含法律意义上已属于他人的任何形式的研究成果，也不包含本人已用于其他学位申请的论文或成果，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谢意。

论文作者签名：\_\_\_\_\_

日期：\_\_\_\_\_年\_\_\_\_月\_\_\_\_日

## 学位论文知识产权权属声明

本人在导师指导下所完成的论文及相关的职务作品，知识产权归属齐鲁工业大学。齐鲁工业大学享有以任何方式发表、复制、公开阅览、借阅以及申请专利等权利，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，本人离校后发表或使用学位论文或与该论文直接相关的学术论文或成果时，署名单位仍然为齐鲁工业大学。

论文作者签名：\_\_\_\_\_

日期：\_\_\_\_\_年\_\_\_\_月\_\_\_\_日

导师签名：\_\_\_\_\_

日期：\_\_\_\_\_年\_\_\_\_月\_\_\_\_日

# 目 录

摘    要 .....	I
ABSTRACT .....	III
第 1 章 绪论 .....	1
1.1 研究背景 .....	1
1.2 研究目的和意义 .....	2
1.3 国内外研究现状 .....	2
1.4 本文的主要内容与整体结构 .....	5
第 2 章 心律失常、深度学习相关理论知识 .....	7
2.1 心律失常概述 .....	7
2.1.1 心律失常产生的原因 .....	7
2.1.2 常见心律失常类型 .....	8
2.1.3 心律失常的表现 .....	8
2.2 十二导联 ECG 概述 .....	9
2.2.1 十二导联 ECG 的起源和发展 .....	9
2.2.2 临床十二导联的作用 .....	11
2.3 常见心电图数据库 .....	11
2.3.1 部分导联数据库 .....	11
2.3.2 十二导联数据库 .....	12
2.4 深度学习 .....	13
2.4.1 深度学习简介 .....	13
2.4.2 深度学习原理 .....	14
2.4.3 常见神经网络 .....	17
2.5 本章小结 .....	20
第 3 章 基于 DSE-ResNet 的十二导联心律失常自动识别 .....	23
3.1 引言 .....	23

3.2 研究方法 .....	24
3.2.1 问题定义 .....	24
3.2.2 二维化十二导联 ECG .....	24
3.2.3 DSE-ResNet 模型构建 .....	27
3.2.4 正交试验 .....	30
3.2.5 集成模型 .....	31
3.3 实验细节 .....	31
3.3.1 实验设备环境 .....	31
3.3.2 数据预处理 .....	31
3.3.3 超参数组合选择 .....	33
3.4 结果 .....	34
3.4.1 评价标准 .....	34
3.4.2 在小批量测试集的表现 .....	34
3.4.3 在 CPSC2018 隐藏测试集的表现 .....	35
3.5 本章小结及讨论 .....	38
<b>第 4 章 基于二维化十二导联 ECG 信号冗余性分析 .....</b>	<b>41</b>
4.1 引言 .....	41
4.2 研究方法 .....	42
4.2.1 十二导联 ECG 计算方法 .....	42
4.2.2 融合数据集 .....	43
4.2.3 导联组合方式 .....	44
4.2.4 改进二维心电图 .....	45
4.2.5 模型构建 .....	47
4.2.6 训练和测试 .....	48
4.3 结果 .....	50
4.3.1 评价标准 .....	50
4.3.2 结果分析 .....	50
4.4 补充信息 .....	52
4.4.1 选择导联 aVL 的过程 .....	52
4.4.2 详细结果 .....	53
4.5 本章小结与讨论 .....	55

第 5 章 基于 Flask 的心律失常自动分类平台 .....	57
5.1 引言 .....	57
5.2 采用的框架和技术模块 .....	57
5.3 功能模块设计 .....	58
5.3.1 自动分类模块 .....	59
5.3.2 数据持久化 .....	61
5.3.3 FTP 工作站 .....	62
5.4 本章小节与讨论 .....	62
第 6 章 总结与展望 .....	65
6.1 全文工作总结 .....	65
6.2 未来研究展望 .....	66
参考文献 .....	69
致    谢 .....	77
在学期间主要科研成果 .....	79

## 摘 要

心律失常是心血管疾病中最常见的一组疾病。临床上心律失常的诊断由专业医生对病人的十二导联心电图（Electrocardiogram, ECG）进行分析推断得出。近些年数字医疗的发展使得心电数据呈现爆炸式增长。同时由于心律失常类型多，病理情况复杂，使得误诊时有发生。因此实现心律失常的自动化、智能化识别成为热点研究方向。

基于机器学习的心律失常自动识别算法严重依赖人工特征的设计，这使得识别过程及结果具有较高的主观性，且无法捕捉到心电图深层次的特征。基于深度学习的心律失常自动识别算法通常专注于单导联心电图的特征提取，继而融合不同导联训练后的特征进行再训练。这种方法在训练初期忽视了不同导联之间的相关联系，导致所提出的模型在部分心律失常类型分类性能低下。在算法不断精进的这段时间里，市面上出现的多种便携式设备也增添了实时记录心电图的功能，这为预防心律失常和提供自动识别创造了条件。受制于电极数量便携式设备无法同时获取完整十二导联 ECG 信号，这为利用十二导联 ECG 进行自动识别的算法造成了障碍。基于以上背景，本文主要进行了如下四项内容研究：

（1）本文提出了一种二维化十二导联 ECG 的方法并进行改进。十二导联 ECG 信号被转化为二维化数据作为深度学习模型的输入，该二维化数据既具备单导联信号在时间上的连续性又具有不同导联信号在空间上的相邻性。

（2）本文提出了一个通用的能够处理二维化数据的深度学习模型 DSE-ResNet。该模型能够在训练初期同时关注导联内部与导联之间的相关联系，实现二维化十二导联 ECG 在时间维度以及空间维度的特征提取。实验过程中引入了正交试验挑选超参数，并利用集成模型提升模型分类性能。

（3）本文对深度学习过程中导联信息的冗余性问题进行分析研究，旨在验证能否以损失小部分识别性能为代价换取兼容多种便携性设备进行高性能心律失常自动识别的可能性。换言之是在验证所有十二导联信息在智能识别过程是否需要全部使用。

（4）本文基于 Flask 框架构建了线上心律失常自动分类平台。用户可自主上传十二导联 ECG 采样文件，平台通过解析文件、加载数据和模型，最终给出识别结果并反馈至用户。

与近几年利用同一数据库的研究进行对比，结果表明本文建立的基于二维化心电图的深度学习模型在所有心律失常分类中取得的平均 $F_1 = 0.817$ ，在部分心律失常类型（例如心房颤动和传导阻滞）的自动识别中取得最高 $F_1$ 分数。针对导



联信号冗余性的研究表明，双极肢体导联和单极加压肢体导联的捆绑在深度学习过程中存在信号冗余。

综上，本文基于二维化十二导联 ECG 信号，使用 DSE-ResNet 模型增强了心律失常分类的精确性，可作为心律失常诊断方向的辅助检测算法。同时本文对导联信息的冗余性进行了研究，验证了深度学习过程中部分导联信息缺失的可能性，为便携式设备减少测量复杂性和增强识别性能提供理论基础。

**关键词：**十二导联；二维化；深度神经网络；心律失常；导联冗余性

## ABSTRACT

Cardiac arrhythmias are the most common group of cardiovascular diseases. The clinical diagnosis of arrhythmia is inferred from the analysis of a patient's 12-lead electrocardiogram (ECG) by a medical professional. Recent advances in digital health care have led to an explosion of ECG data. At the same time, the variety of cardiac arrhythmias and the complexity of the pathology make misdiagnosis a frequent occurrence. Therefore, the realization of automated and intelligent identification of cardiac arrhythmias has become a hot research direction.

The automatic recognition algorithm of cardiac arrhythmia based on machine learning relies heavily on the design of artificial features, which makes the recognition process and results highly subjective, and cannot capture the deep-level features of the ECG. The automatic cardiac arrhythmia recognition algorithm based on deep learning often focuses on the feature extraction of single-lead ECG, and then fuses the trained features of different leads for retraining. This method ignores the correlation between different leads in the early stage of training. This leads to the low performance of the proposed model in classification of some cardiac arrhythmia types. During this period of continuous improvement of the algorithm, various portable devices appearing on the market have also added the function of recording ECG in real time, which creates conditions for preventing cardiac arrhythmia and providing automatic identification. Restricted by the number of electrodes, portable devices cannot simultaneously acquire a complete 12-lead ECG signal, which creates obstacles for the automatic identification algorithm using 12-lead ECG. Based on the above background, this thesis mainly carried out the following four research contents:

(1) This thesis proposes a two-dimensional 12-lead ECG method and improves it. 12-lead ECG signals are transformed into a two-dimensional plane as the input of the deep learning model. The two-dimensional plane has both the temporal continuity of single-lead signals and the spatial adjacency of different lead signals.

(2) This thesis proposes a general deep learning model DSE-ResNet that can handle two-dimensional data. The model can pay attention to the correlation between the leads and the leads at the early stage of training, and realize the feature extraction of the two-dimensional 12-lead ECG in the time dimension and the space dimension. During the experiment, an orthogonal experiment was introduced to select

hyper-parameters, and ensemble learning was used to improve the classification performance of the model.

(3) This thesis analyzes and studies the redundancy of lead information in the process of deep learning, aiming to verify whether it is possible to exchange a variety of portable devices for high-performance automatic cardiac arrhythmia recognition at the cost of a small loss of recognition performance. In other words, it is to verify whether all 12-lead information needs to be fully used in the intelligent identification process.

(4) This thesis builds an online cardiac arrhythmia automatic classification platform based on the Flask framework. Users can independently upload 12-lead ECG sampling files, and the platform will analyze the files, load data and models, and finally give the recognition results and give feedback to users.

Compared with studies using the same database in recent years, the results show that the two-dimensional ECG-based deep learning model established in this thesis has achieved average  $F_1 = 0.817$  in all cardiac arrhythmia classifications, and in some cardiac arrhythmia types (such as atrial fibrillation and Highest score in automatic identification of conduction block). Research on lead signal redundancy has shown that bundling of bipolar and unipolar pressurized limb leads can be redundant during deep learning.

In summary, based on the two-dimensional 12-lead ECG signal, this thesis uses the DSE-ResNet model to enhance the accuracy of cardiac arrhythmia classification, which can be used as an auxiliary detection algorithm in the direction of cardiac arrhythmia diagnosis. At the same time, this thesis studies the redundancy of lead information, verifies the possibility of missing some lead information in the process of deep learning, and provides a theoretical basis for portable devices to reduce measurement complexity and enhance recognition performance.

**Keywords:** 12-lead, Two-dimensional, DNN, Cardiac Arrhythmia, Lead Redundancy

# 第 1 章 绪论

## 1.1 研究背景

心血管疾病是一种常见、多发和死亡率高的慢性病，具有病情隐蔽、危险性高和突发性强等特点。心律失常是心血管疾病中最常见的一组疾病，容易引起多种并发症，对人类的健康造成了极大的威胁。心电图（ECG）是医学日常实践中的基本工具，因其简单、无创、可靠等特点，多作为心律失常临床诊断的依据。全世界每年采集获得数百万的心电图记录<sup>[1]</sup>，且这个数据量逐年上升。专业医生增长的速度远不能匹配心电图的产出数量，导致病症诊断结果无法及时提供。在实际生活中对心律失常的诊断主要依赖医生对心电图的分析，呈指数剧增的心电数据加重了医生的工作负担。近几十年随着计算机算力的不断提升，越来越多的领域引入了计算机辅助识别诊断，计算机辅助识别在临床心电图工作流程中也变得越来越重要。然而，现有的商用 ECG 诊断算法仍然显示出较高的错误率<sup>[2]</sup>。随着 ECG 信号的数字化、计算机硬件计算力的提升和多种算法的提出，为大规模处理原始心电信号提供了条件，也为重新创建高性能的 ECG 智能识别方法展示了可能性。

由于心电图的复杂性，对心电图进行精确分析存在很大难度。目前专业医生对心电图的分析主要依靠常年累积的行医经验，识别异常心电图在波形结构中存在的问题。为客观、准确、快速地进行心电图自动分析，借助高性能计算设备和人工智能识别算法对心电图进行自动化分析，提升心电图分析的效率和准确率是智慧医疗在心电图领域发展的应有之义。同时随着便携式设备例如带有实时心电图记录功能的智能手表的不断出现，监测主体心电信号并利用智能识别算法反馈结果可为预防心血管疾病和及时就诊提供条件。

现如今国家大力推行人工智能和智慧医疗。“十四五”期间将持续推进“智慧医院”建设，通过信息化手段提升医疗服务的效率、质量与安全。《“健康山东 2030”规划纲要》提出加大研发健康医疗相关的人工智能技术和可穿戴设备，促进健康医疗智能装备产业升级。本研究旨在将人工智能技术应用于基于十二导联 ECG 的心律失常的智能识别研究中，同时分析十二导联 ECG 信号在深度学习过程中的冗余性。这有助于为便携式设备应用部分导联进行智能识别算法提供理论基础。因此本研究具有较高的临床意义和市场空间，具有较为广阔的应用前景，属于国家和山东省的重点发展领域。

## 1.2 研究目的和意义

根据 2021 中国心血管健康与疾病报告<sup>[3]</sup>, 中国患有高血压、血脂异常和糖尿病的人群已多达数亿, 依赖多种药物终生防治心血管病的人群越来越多。而心电图作为最常规的观测患者身体状态的依据, 一方面极度依赖于专业医生的结果诊断, 另一方面急剧增长的心电图数量导致仅有的专业医生无法及时给出诊断结果, 患者错过最佳治疗时间的情况时有发生。因此为高危人群提供心电信号实时记录, 并在此基础上实现心律失常的自动识别, 为医生提供辅助诊断变得尤为重要。

本文针对心电图的时序特性以及空间特性, 建立了一种二维化十二导联 ECG 的方法。其中二维代表有界平面, 有界平面中每个点代表某一导联的一个采样电压值。以二维卷积神经网络为基础构建深度学习模型实现导联内部特征和导联之间特征的提取。同时置入患者的性别和年龄作为辅助数据, 增强样本的总体特征。除此之外, 带有心电图记录功能的辅助设备通常只记录标准十二导联中的部分导联信号。而十二导联信息囊括了来自心脏不同方位的心电信息, 相较于部分导联, 利用完整十二导联实现心律失常的自动识别更加合理。然而全面的数据意味着庞大的记录量和计算量, 为了搞清楚在利用深度学习实现心律失常的自动识别过程中是否全部导联信号都发挥作用, 本论文对深度学习过程中不同导联信号组合的冗余性进行研究。综上本文的研究目的和意义如下:

(1) 研究成果提供了一种新的处理不同采样长度的十二导联 ECG 的方法。使得处理后的十二导联数据保持相同的长度, 同时让十二导联 ECG 既有时序上的连续性, 也具有空间上的相邻性。

(2) 本论文创建了一种二维数据通用的深度模型学习框架, 能够对十二导联 ECG 进行导联间特征和导联内部特征的提取。并实现较高评价体系的心律失常的自动识别。利用这种模型构建了一个基于 Flask 的心律失常自动分类平台, 用户可自主上传 ECG 数据, 实现心律失常在线识别。

(3) 本论文为自动识别心律失常方向上导联之间冗余性问题的研究提供新的视角。探究了不同深度学习模型和不同导联组合对分类性能的影响。为兼容多种便携式设备提供心律失常的自动识别创造条件。

## 1.3 国内外研究现状

过去几十年中对于心律失常自动分类的研究中, 基于机器学习的算法占据大多数, 这类算法通常分为构建特征工程和设计分类算法两部分。具体来说, 研究人员首先利用数学方法从原始 ECG 中提取大量具有医学意义的特征, 如小波特征<sup>[4]</sup>、P-QRS-T 复合特征<sup>[5,6]</sup>、心率变异性统计特征<sup>[7]</sup>、RR 间期相关统计特征<sup>[8]</sup>、高

阶统计特征<sup>[8,9]</sup>和形态学特征<sup>[10,11]</sup>。主成分分析<sup>[12,13]</sup>和独立成分分析<sup>[14,15]</sup>等降维方法实现心电图特征从高维空间向低维空间的映射。经过特征工程得到的信息利用分类算法进行分类,常见的分类算法包括支持向量机<sup>[16,17]</sup>、自组织映射<sup>[18]</sup>、聚类<sup>[19]</sup>等。尽管机器学习在心律失常分类方面取得了不错的表现,但仍有一些问题需要解决。例如,基于主观因素提取的特征导致一些潜在的、深层次的重要信息被忽略、ECG 的信息不能够被完整的使用等,这些情况会影响最终的分类性能。

近年来,深度神经网络(Deep Neural Network, DNN)以其强大的特征提取能力,极大地提升了对于语音识别、图像分类、疾病诊断等任务的技术水平。与机器学习方法不同,深度神经网络可以从原始输入数据中自动识别模式并学习有用的特征,而无需依赖大量的手动规则和特征工程,这使得它们特别适合应用于 ECG 数据。一些研究受此启发,将深度神经网络用于基于单导联或多导联心电图的心律失常自动分类中。

在这些研究中由于选择不同的数据库进行实验,导致实验样本的记录长度、导联数量不尽相同。但大体分为以下两类:基于单导联样本(每条记录有多个心拍,标签按照心拍标注)的分类算法和基于多导联样本(每条记录给出一个或多个标签)的分类算法。其相关研究包括 Ullah 等人<sup>[20]</sup>提出了一种基于单导联的二维卷积神经网络(Convolutional Neural Network, CNN)模型,利用短时傅立叶变换将一维心电时间序列信号转换为二维频谱图,每个频谱图对应一个心拍以及一个标签,经过训练的模型在 MIT-BIH 心律失常数据库进行评估取得了 99.11% 的平均准确率。Jun 等人<sup>[21]</sup>提出了一种使用二维深度卷积神经网络(CNN)的 ECG 心律失常分类方法,该方法将每个心跳间期的波形图转换成二维灰度图像作为 CNN 分类器的输入数据,并在 MIT-BIH 心律失常数据库进行测试得到了 99.05% 的平均准确度。Hannun 等人<sup>[22]</sup>开发了一种基于 ResNet 的一维 DNN 模型,实现了单导联心电数据端到端的学习,并完成正常节律以及 11 类心律失常类型的分类。训练完成的 DNN 模型在专用数据集上的敏感度超过了心脏病专家分类的平均敏感度,并且证明了端到端的深度学习方法可以从单导联心电图分类出多种不同类型的心律失常,其诊断性能与心脏病学家相似。

与单导联 ECG 相比,多导联 ECG 包含心脏不同方位的电位信号,一些研究表明充足的数据量更有利于心律失常的自动分类。例如 Wang 等人<sup>[23]</sup>提出了一种基于多尺度特征提取和十二导联 ECG 跨尺度信息互补的方法来捕捉心电信号中的异常状态,并在“The China Physiological Signal Challenge 2018 (CPSC2018)”提供的心律失常数据库中进行了测试,得到了 82.8% 的  $F_1$  分数。Chen 等人<sup>[24]</sup>提出了一种结合 CNN、循环神经网络(Recurrent Neural Network, RNN)和注意力机制(Attention)的用于心律失常分类的模型。实验过程中对十二导联 ECG 逐导联进行训练,然后融合逐导联训练的结果进行再训练。该模型应用于 CPSC2018 心

律失常数据库并取得了排名第一的测试结果。同时他们基于单导联 ECG 进行实验发现使用单导联数据产生的性能次于完整十二导联 ECG，其中利用导联 aVR 和 V1 的分类表现接近于十二导联。Zhang 等人<sup>[25]</sup>逐导联训练并融合特征再训练得到一维 DNN 模型，该模型在测试数据库平均得分  $F_1=0.813$ ，这个结果显示出优于四种基于专家特征的机器学习方法的性能。此外，在单导联 ECG 上训练的深度模型产生的性能低于同时使用所有十二导联，表现最好的是导联 I、aVR 和 V5。Chiou 等人<sup>[26]</sup>应用连续小波变换（CWT）将一维 ECG 信号转换为二维光谱，用于二维 CNN 分类，并评估了单个 ECG 导联对收缩性心力衰竭（HF）分类结果的贡献。结果表明将十二导联 ECG 信号中的导联 V5 和 V6 进行组合，可获得质量最高的评分。此外，zhao 等人<sup>[27]</sup>的文章表明，添加性别和年龄等临床数据作为辅助特征可以提高分类性能。王英龙等人<sup>[28]</sup>提出一种将多导联心电图数据应用于 CNN 和长短期记忆网络（Long Short-Term Memory, LSTM）的组合模型实现心律失常自动分类。一方面多导联心电图较单导联心电图蕴含更多心电信息，另一方面通过融合 CNN 及 LSTM，使得模型在学习时间序列数据和空间结构数据方面有着独特优势。利用多导联心电图数据训练 CNN-LSTM 提高了 DNN 的学习效率和心电图识别的精度。Hong 等人<sup>[29]</sup>提出了一种用于 ECG 分类的集成分类器，首先提取心电图的统计特征、信号处理特征和医学特征，然后构建 DNN 来自动提取深层特征，同时提出了一种新的算法来寻找中心波（一名患者心电图波形中最具代表性的波形）特征。最后将专家特征、深度特征和中心波特征结合起来，训练多个梯度决策树分类器，并将这些分类器集成在一起输出预测。结果表明集成分类器增强了模型的分类性能。

以上的部分研究利用单导联 ECG 进行实验，使用 MIT-BIH 心律失常数据库进行测试，这个数据库中记录了 48 名受试者的记录，每条记录按照心拍给出标注（48 条记录约合 11 万心拍），所以按照心拍切分样本后训练集中存在多个样本对应一位患者的情况，这导致训练后的模型过度拟合该数据库，使测试结果偏颇。基于十二导联作为数据源的研究，往往在 DNN 训练初期忽略导联间的关系，例如 DNN 使用一维卷积对各导联独立进行特征学习，再将不同导联得到的特征融合，这种模式虽然能够在训练初期提取各导联的内部特征，但是无法捕获导联之间的相关联系，从而割裂了十二导联数据之间的相关性。同时，这些模型对于部分心律失常类型识别的准确度仍然有进一步提高的可能性。在此基础上，随着各种能够记录不同导联 ECG 的便携式设备的出现，这些设备受制于电极数量少，无法记录完整十二导联信息。因此探索深度学习过程中导联信息的冗余性，利用部分导联信号实现心律失常高性能的自动识别成为新的研究方向。

## 1.4 本文的主要内容与整体结构

本论文利用深度学习相关技术对二维化十二导联ECG信号进行自动分类和对深度学习过程中导联信息的冗余性进行分析与研究。第一章依次介绍了课题的研究背景、研究目的和意义、国内外研究现状以及本文的主要内容与整体结构。

第二章对心律失常发生的原因、十二导联心电图、常见心电图数据库、心律失常和神经网络相关基础知识进行介绍，组建它们之间的联系。

第三章构建了用于处理二维化数据的通用神经网络模型 DSE-ResNet 实现心律失常的自动识别。首先对二维化十二导联 ECG 进行详细解释，包括二维化十二导联 ECG 的原理、数据的预处理等。然后阐述了二维化神经网络模型的构建过程、超参数的选择方式以及集成模型的原理。最终对分类结果做出解释和说明，并基于这个结果进行讨论。

第四章主要对深度学习过程中的十二导联 ECG 的信号冗余性进行实验分析。DNN 主体结构沿用第三章提出的 DSE-ResNet，并对这个结构进行延伸和扩展，得到五种不同结构的 DNN 模型。对第三章中的二维化方式进行改进产生更加复杂的数据维度结构从而保证样本在训练集的唯一性。通过擦除不同导联组合的信号值，探究在不同深度学习模型中不同导联组合对自动识别心律失常的影响，从而得出在深度学习过程存在冗余导联信号的结论，并对这种现象做出解释和讨论。

第五章建立了一个基于Flask的心律失常自动分类平台，本章节将数据处理算法与DNN分类算法打包为接口同前端页面实现数据交互。

第六章为本文的总结和展望。



## 第2章 心律失常、深度学习相关知识

### 2.1 心律失常概述

心律失常（Cardiac Arrhythmia）是指心脏电传导系统异常而引起的心跳过慢、过快或心跳不规则的总称。随着我国老龄化时代的到来和代谢危险因素持续流行，患有心律失常病症的患者群体与日俱增，且呈现年轻化的趋势<sup>[30]</sup>。它严重危害着人类的健康，了解心律失常产生的原因对预防和治疗心律失常有着重要的意义。

#### 2.1.1 心律失常产生的原因

心律失常根据获得类型可分为先天存在和后天获得。先天存在通常由家族遗传或基因突变导致；后天获得类型包括生理性因素例如运动、情绪、精神和睡眠问题等以及病理性因素例如心脏病、冠心病、心肌病、电解质紊乱、神经与体液调节功能失调等导致。

心脏电活动由心脏电信号传导系统进行控制。该传导系统是由位于心肌内能够产生和传导激动的多种心肌细胞构成，主要组成部分包括窦房结、房间束、结间束、房室结、房室束、右束支、左束支和浦肯野（Purkinje）纤维等<sup>[31]</sup>。正常情况下心电信号（激动）产生于窦房结，即窦房结是心脏的正常起搏点。产生的电信号首先经过房间束和心房肌，然后传导至房室结，再经由左右束支传导到心室束，心室束再将电活动传导至Purkinje纤维，最终电活动被Purkinje传导至每一个心室肌细胞，引发心脏收缩。心肌细胞通过连续不断地接受电信号刺激产生有规律的收缩舒张，使得心脏能够为全身提供源源不断的血液供给，保障人体血液循环。然而心脏的各个传导系统组织结构都存在病变可能性，导致产生不同类型的心律失常。

从心脏结构及心脏电信号传导过程分析，心律失常产生的原因主要来自于以下几个方面：

（1）窦房结激动异常：窦房结发生病变或结构性变化，导致电信号产生异常或者不产生电信号，引发心搏频率和节律异常。

（2）激动产生于窦房结之外：窦房结是正常心脏起搏点，电信号产生于窦房结之外，可能导致电信号传导缓慢，引起心律失常。

（3）传导系统结构阻滞：电信号受心脏电信号传导系统通道结构影响，导致电信号传导过快、过慢或者阻滞，引起心律失常。常见阻滞包括窦房阻滞、房室传导阻滞以及室内阻滞。

### 2.1.2 常见心律失常类型

由于心电传导系统的复杂性，心律失常的根源可能来自于不同部位。因此心律失常类型较多，以下展示按照发生部位分类的心律失常类型：

（1）窦房结发生的心律失常：窦房结发生的心律失常常以“窦性”命名。例如窦性心动过速、窦性心动缓慢、窦性心律不齐、窦性停搏等。

（2）心房发生的心律失常：心房发生的心律失常常以“房性”命名。例如心房颤动、心房扑动、阵发性房性心动过速等。

（3）房室交界区发生的心律失常：房室交界区心律失常是由房室结及周围组织引起的心律失常。常见例如房室交界区期前收缩、非阵发性交界区心动过速等。

（4）心室发生的心律失常：心室发生的心律失常常以“室性”命名。例如心室颤动、心室扑动、室性心动过速等。

### 2.1.3 心律失常的表现

不同心律失常类型在不同患者身上的表现可能不同，相同心律失常类型测量出的 ECG 波形结构也可能不相同。这就造成心律失常识别难度较高，误诊率也较高。从生理症状看，心律失常患者常伴有视力模糊、头晕、昏厥、腹胀、腹泻、尿频尿急、胸闷气短等表现。而在 ECG 中则表现为波形结构、波段持续时间和各种峰值等标量值异于常人。

一条正常心电图信号由多个心跳节拍信号组成，每个心跳节拍信号包含 P 波、QRS 复合波和 T 波。这意味着一个心跳节拍信号存在多个波形峰值（P 波、Q 波、R 波、S 波、T 波和 U 波）、多组间隔（PR 间隔、RR 间隔、QRS 间隔、ST 间隔和 QT 间隔）和两个段（PR 段和 ST 段）。ECG 波形反映了心脏兴奋过程中心脏不同方位的电位变化<sup>[32]</sup>。不同峰值、间隔和段都有其正常振幅或持续时间值。这些峰值、间隔和段称为 ECG 特征。任意 ECG 特征出现异常，是特定类型心律失常的心电图表现。图 2.1 展示了一个正常心跳节拍信号波形结构。

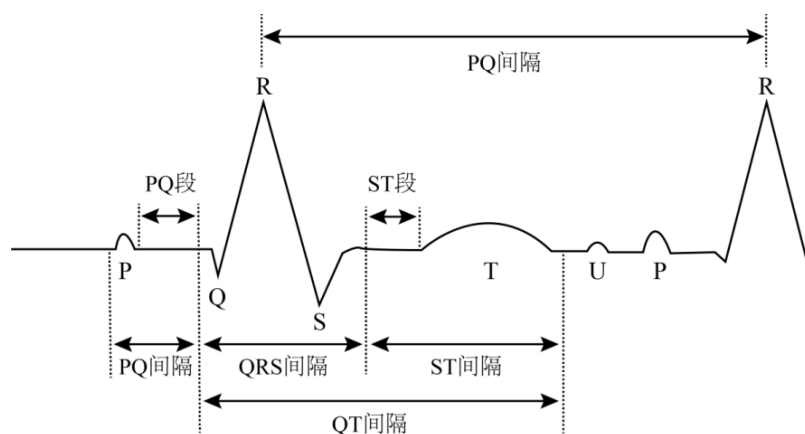


图 2.1 正常心跳节拍信号波形结构

## 2.2 十二导联 ECG 概述

十二导联 ECG 具有无创、便捷等特点，是临床中最广泛使用的诊断心律失常的手段。为全方位反映心脏电活动情况，我国临床常用标准十二导联组成体表心电图，所用的导联体系为 Einthoven-Wilson 体系<sup>[33]</sup>。

### 2.2.1 十二导联 ECG 的起源和发展

二十世纪初，体表心电图由三个导联组成，称标准导联 I、II 和 III，也称双极肢体导联 I、II 和 III。它们的电极连接方式如下，标准导联 I：将测量设备的阳极连接至人体左腕，阴极连接至人体右腕，能够记录人体左臂与右臂间电压差。标准导联 II：阳极连接人体左踝，阴极连接人体右腕，能够记录了人体左腿与右臂间电压差。标准导联 III：阳极连接人体左踝，阴极连接人体右腕，能够记录人体左腿与左臂电极间电压差。其电极摆放位置如图 2.2 所示。标准导联 I、II、III 记录了心脏额状面不同方位的电位变化。

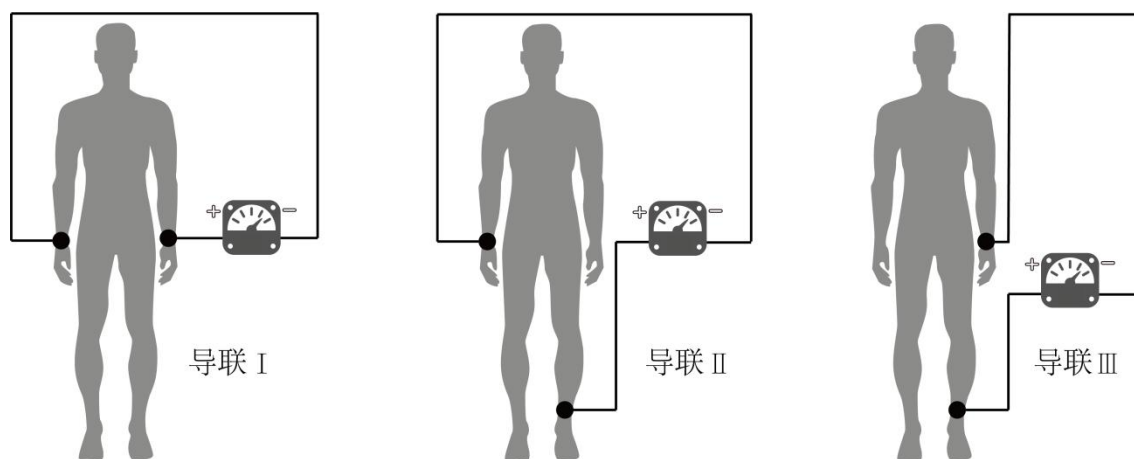


图 2.2 双极肢体导联电极放置方式图

二十世纪三十年代末至四十年代初，位于美国密歇根州的以 Wilson 为首的科研团队不满足于现有的只能观测心脏额状面电位变化的体表心电图系统。实验过程中，他们首先把测量双极肢体导联 I、II、III 的左、右手及左踝的电极连在一起，发现综合电位几乎等于 0，于是他们便把这个综合电极定义为“中心电端”，用它连接测量设备的阴极端，另外用一个“记录电极”放在身体的不同部位。“记录电极”记录的 ECG 便被定义为“单极肢体导联心电图 VR、VL 和 VF”。为了测量心脏水平方向上的电位变化，Wilson 团队还提出了胸前导联系统 V1、V2、V3、V4、V5 和 V6。胸前六导联电极摆放位置如图 2.3 所示，它们记录了心脏横断面不同方位的电位变化。

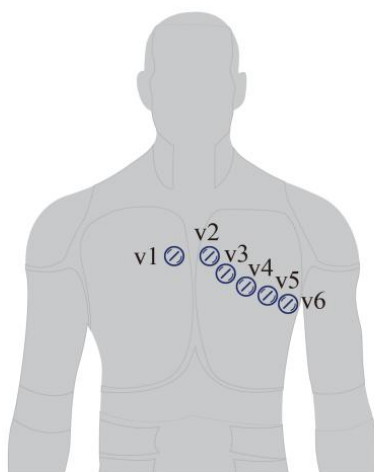


图 2.3 胸前六导联电极放置方式图

由于单极肢体导联心电图 VR、VL 和 VF 所展示的心电图波形不明显，Goldberger<sup>[34]</sup>提出了单极加压肢体导联 aVR、aVL 和 aVF（a 代表 augmented，V 代表单极导联，R、L、F 分别代表右臂、左臂和左腿）。具体原理是在测量右腕的单极导联时，便把中心电端中的右腕电极拔除。实际上阳极连接人体右腕，阴极连接人体左腕、左踝，这样得到的心电图与 VR 相同，但整体波形增大了 0.5 倍，故称为单极加压肢体导联（Augmented unipolar limb lead）。同样的记录 VL 时，也只把阳极连左腕，阴极连接右腕、左踝；记录 VF 时，把阳极连在左踝，阴极连接左、右腕。这样便得到了单极加压肢体导联 aVR、aVL 和 aVF。导联 aVR、aVL 和 aVF 的电极位置图如图 2.4 所示。这三个导联也记录了心脏额状面不同方位的电位变化。

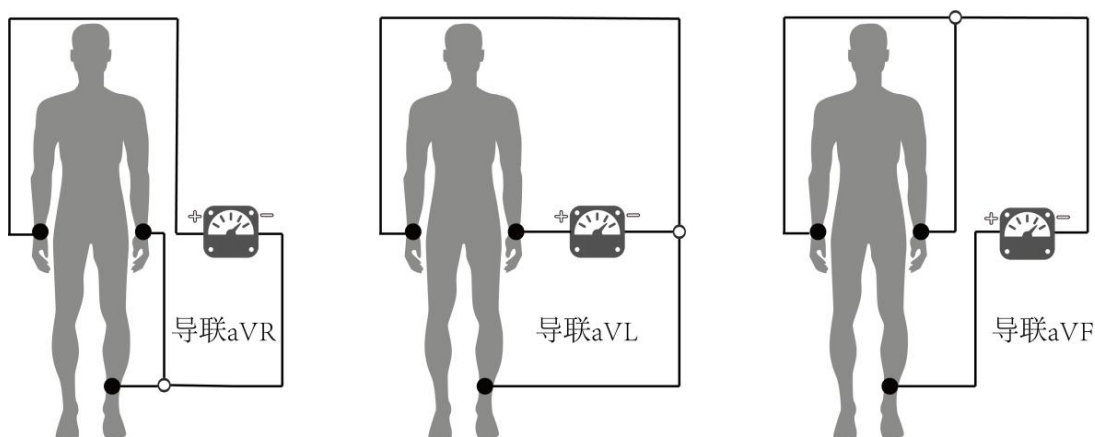


图 2.4 单极加压肢体导联电极放置方式图

至此，标准十二导联系统正式产生，它共由三部分导联信号构成。分别为双极肢体导联 I、II 和 III；单极加压肢体导联 aVR、aVL 和 aVF；胸前导联 V1、V2、

V3、V4、V5 和 V6。它能够测量来自于心脏额状面和横断面的多方位的 ECG 信号，是检查相关心脏疾病的最基本、最常见的检查手段，也是临床中最广泛应用的心电图检测系统。

### 2.2.2 临床十二导联的作用

相较于美国临床使用较多的三导联系统，十二导联系统不仅能够精确记录心律失常的发生，例如房颤、室性心动过速等，而且能够提高心肌缺血检测效率，能够帮助医生定位心肌缺血或者梗死的部位，为心脏相关疾病的检查提供有力的保障。同时我国国产动态心电图（Holter）仪器占有较高市场份额，更加速了十二导联 ECG 系统在临床和日常生活中的应用。

## 2.3 常见心电图数据库

心电图数据受信息私密性的保护，获取医院私有数据库难度较大。但为进行研究，部分数据库被开源供给大众进行参考开发。常见的开源数据库分为两种，给定部分导联信号的小型数据库和给定十二导联 ECG 信号的大型数据库。以下对这些数据库以及本研究使用到的数据库展开介绍。

### 2.3.1 部分导联数据库

部分导联数据库指的是提供的样本只包含十二导联中某几个导联的心电信号数据的数据库。目前最常见的部分导联权威数据库包括麻省理工的 MIT-BIH 数据库，欧洲的 ST-T 心电数据库，这两个数据库都只记录了两个导联的心电信息。详细信息如下：

（1）MIT-BIH 心律失常数据库：MIT-BIH 心律失常数据库是全球首个用以评估心律失常诊断仪器性能的测试数据库。该数据库包含了 48 组采样率为 30Hz、长度大于 30 分钟、导联数量为二导联的动态心电图信号样本。每组样本给定的导联信息为改良后的导联 II、导联 V1、V2、V4 和 V5 中的两个。且由两位专业医生按照心拍对 48 组样本进行了详细标注，总共约 110000 个标注。自 1980 年以来，它已被用于全球约 500 个地点的心脏动力学基础研究<sup>[35]</sup>。目前相关基础研究和心脏相关医疗设备的开发很大程度上受到了 MIT-BIH 数据库的影响。随着时间流逝，由于这些数据库存在记录导联数量过少，患者样本个数稀缺等问题，越来越多的研究开始选择其他数据库进行实验。

（2）ST-T 心电数据库：ST-T 心电数据库主要用于评价和分析 ST 段和 T 波检测的算法性能<sup>[36]</sup>。该数据库由 79 名心肌缺血疾病患者的 90 组带注释的动态心电图样本组成。每条样本记录持续 2 小时，包含两个导联信号。由于 ST-T 数据库专用于心肌缺血方向的研究，故而在心律失常方向应用该数据库的研究较少。

### 2.3.2 十二导联数据库

随着计算机算力的不断提升和多个十二导联 ECG 数据库的开源,越来越多的研究着手于蕴含心电信息更加丰富的十二导联数据库进行数据挖掘,尝试借助优良的算法和卓越的计算机性能实现心律失常的自动识别。近几年被广泛使用的十二导联心电数据库包括 2018 年中国生理信号挑战赛的心律失常数据库 (CPSC2018)、圣彼得堡 INCART 心律失常数据库、Physikalisch Technische Bundesanstalt 心律失常数据库 (PTB) 和乔治亚州心律失常数据库 (G12EC) 等,这些数据库均可从 PhysioNet<sup>[37,38]</sup>获取。以下详细介绍这些数据库:

(1) CPSC2018 数据库: CPSC2018 心律失常数据库的数据采集自中国的 11 家医院,信号采样率为 500Hz<sup>[39]</sup>。共计 9831 组十二导联心电图样本,样本持续时间从 6 秒到 144 秒不等,其中 6877 组(男性: 3699, 女性: 3178)样本被公开用于训练,2954 组样本未被开源用于评价测试。每组样本包含十二导联 ECG 信号、患者基本信息(年龄和性别)以及标签,其中标签对应 9 个类别: 正常节律、心房颤动 (AF)、一度房室传导阻滞 (I-AVB)、左束支传导阻滞 (LBBB)、右束支传导阻滞 (RBBB)、房性早搏 (PAC)、室性早搏 (PVC)、ST 段压低 (STD) 和 ST 段抬高 (STE)。不同样本给定的标签数量不同,绝大多数样本只有一个标签,少数样本存在两个或者三个标签。

(2) CPSC-Extra 数据库: 该数据库是 CPSC2018 数据库未曾公布和使用的数据。它在 2021 年的 PhysioNet 举办的名为 “Will Two Do? Varying Dimensions in Electrocardiography: The PhysioNet/Computing in Cardiology Challenge 2021 (PCC2021)” 的比赛中被公开使用。

(3) G12EC 心律失常数据库: G12EC 数据库样本收集自美国东南部。该数据集包含 10344 (男性: 5551, 女性: 4793) 组十二导联 ECG 样本,每条样本的记录长度为 10 秒,信号采样率为 500Hz。

(4) 圣彼得堡 INCART 十二导联心律失常数据库: 圣彼得堡 INCART 十二导联心律失常数据库包含来自 32 个动态心电图记录仪中提取的 74 组带标注的样本,每条样本的记录长度为 30 分钟,信号采样率为 257Hz;

(5) PTB 心律失常数据库: PTB 心律失常数据库是一个大型公开可用的心电图数据集<sup>[40]</sup>,收集自柏林的本杰明富兰克林医学大学的心脏内科。它包含了两部分的数据,第一部分 PTB 数据库包含 516 组十二导联心电图样本 (男性: 377, 女性: 139), 信号采样率为 1000Hz; 第二部分 PTB-XL 数据库包含 21837 组样本 (男性: 11379, 女性: 10458), 信号采样率为 500Hz。

这五个数据库包含约有 27 类心律失常类型。与 MIT-BIH 数据库不同的是这些十二导联 ECG 数据库的标签按照样本给出,而不是基于心拍给出。即每组十二

导联心电图样本对应一种或者多种心律失常类型。这些数据集为训练和评估心律失常自动分类算法提供资源。

## 2.4 深度学习

深度学习是近几年来人工智能领域出现频次较高的名词。伴随着算力的不断提升、科技化社会的到来和各种各样的开源医疗数据的出现都促进了深度学习在智慧医疗领域的发展和应用。它的优势在于能够自动学习提取表格等结构化数据以及视频、音频、图片等非结构化数据的不同层次的特征，使得分类或者预测模型在很大程度上减小了专家系统所带来的主观影响。

### 2.4.1 深度学习简介

近些年来技术不断突破、算法层层优化，越来越多的行业开始应用人工智能，例如经济金融领域的量化交易模型，短视频平台和购物平台的智能喜好推荐，大语言模型 ChatGPT 以及智慧交通、智慧医疗等。人工智能在技术上的实现核心是针对不同业务场景实现不同的算法，而这类算法的统称就叫做机器学习（Machine Learning）。

神经网络可以看作是机器学习的一类算法，它设计的初衷是为了模拟大脑神经元的处理方式，使原始流程式的计算执行方式能够转化为像人脑的逻辑迭代进行推演。神经网络的发展已经从最初的感知器模型到浅层神经网络再到现如今使用较为广的深度学习神经网络。为了更好地拟合复杂问题，神经网络逐渐从线性化模型过渡到非线性模型，算法的复杂度也在不断提升。图 2.5 展示了一个 N 输入 M 输出的三层神经网络的结构图。它由输入层、输出层和隐藏层三部分组成，每一层包含数量不完全相同的神经元。

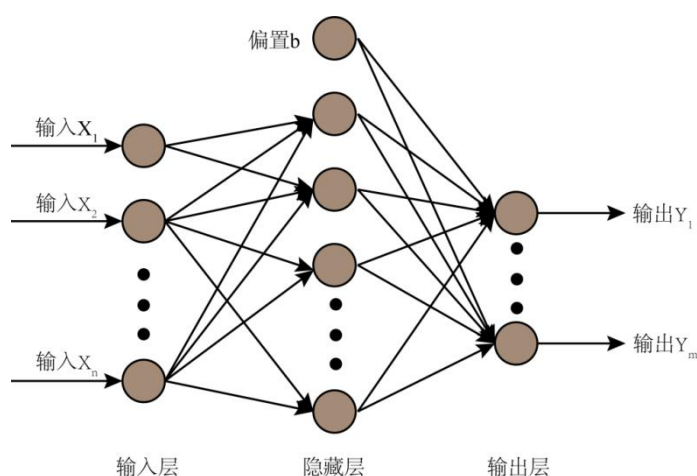


图 2.5 三层神经网络结构图

深度神经网络是指通过堆叠多层功能不相同的网络层，实现数据不同层次特

征的提取。分层组合带来的是指数级表达空间，使得模型可以更容易地学习和表示各种特征。理论上越深的神经网络能够提取的特征越抽象，越能贴近人脑所能提取到的信息的轮廓。然而神经网络的深度直接决定了参数的数量，参数越多需求计算机的算力越高。因此对深度的设计也应该具有适度性。相较于浅层神经网络，深层网络所包含的隐藏层数量更多，而隐藏层神经元数量相对较少，浅层的网络想要达到深层网络同样的结果需要指数级增长神经元数量才能达到。

### 2.4.2 深度学习原理

针对复杂状况下的数据，例如视频流、音频流等，深度学习与机器学习相比最大的优势在于原始数据可直接输入模型进行特征学习，而不再需要专家系统的支持，这在一定程度上避免主观选择特征给模型造成学习误区。现如今深度学习最广泛的实现方式是深度神经网络。

深度神经网络中最著名的结构就是卷积神经网络，卷积神经网络参考了人脑对于视觉听觉类任务的处理原则，分析局部再规划整理全局。由于卷积神经网络具有“局部连接”和“权值共享”的特点，减少了参数量，解决了全连接网络参数太多，训练时间复杂度高的问题。深度学习的整体算法流程是输入数据逐一经过功能不同的隐藏层正向传播将结果送至输出层，输出层通过损失函数计算正向传播的结果与真实值的误差，该误差经过反向传播反馈到各层，计算各层参数在损失函数上的梯度，各层参数根据梯度的反方向更新，通过多次轮回得到最小化损失函数值，以此建立最优模型。以下具体介绍深度学习的相关原理：

（1）深度神经网络的结构：深度神经网络的全局结构类似于图 2.5 所展示的三层神经网络的结构，不同的是 DNN 隐藏层数量较多，且存在卷积网络层、池化层等不同功能的层用以提取不同层次的特征。深度神经网络能够从大量数据中学习到相关特征用于预测或者分类任务，其关键在于算法的设计和优化，包括激活函数的设计、损失函数的设计、梯度下降算法的设计以及参数优化和参数调整等。

（2）激活函数：现实生产生活中需要深度神经网络解决的问题常常是来自于不同行业的，这些数据存在较大的差异，而且数据之间大多不在线性关系，这使得线性模型结构不适合这些任务。为了给模型添加非线性的处理单元，激活函数的概念被提出。神经网络运行过程中上一层的输入经过加权求和后作用于另一个非线性函数并将结果传递到下一层，这里的非线性函数就指的是激活函数。为了满足反向传播算法的条件，激活函数需要具备以下性质：连续且可导（允许存在部分点位不可导）；激活函数的形式以及它的导数形式应尽可能的简单，能够提高计算效率；激活函数的导函数的值域应在合理的区间，防止梯度消失和梯度爆炸。常见的激活函数包括：



Sigmoid 激活函数，常使用 $\sigma$ 表示。它能够将输入的值映射在区间(0,1)上，可等价理解为将神经元的输入进行了归一化，输出结果往往接近于区间上下限，因此常用作二分类任务。在实际使用过程中，由于 Sigmoid 在靠近 0 和 1 处的曲线较为平滑，在这些点附近的梯度趋近于 0，导致在反向传播过程中输出结果靠近 0 或 1 的神经元的梯度也趋近于 0，使得权重参数无法更新，产生梯度消失的问题<sup>[41]</sup>。Sigmoid 激活函数公式如下所示，相比其他激活函数计算成本较高。常见激活函数的函数图像如图 2.6 所示。

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.1)$$

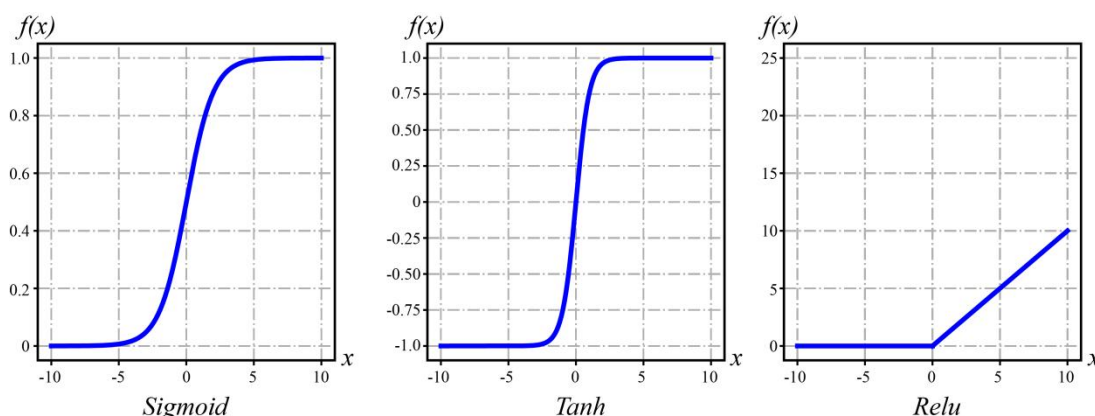


图 2.6 常见激活函数图

Tanh 激活函数，它将输入的值映射在区间[-1,1]。引进了负输入被映射为负数，接近 0 的输入被映射在接近 0 的位置的特性。但由于函数曲线类似于 Sigmoid<sup>[42]</sup>，仍旧存在反向传播过程中梯度消失的问题。Tanh 通常应用在隐藏层中。Tanh 激活函数公式如下：

$$f(x) = \frac{2}{1 + e^{-2x}} \quad (2.2)$$

Relu (Rectified Linear Unit) 激活函数在深度神经网络中使用最为广泛，是简单的分段线性函数<sup>[43]</sup>。实际传递参数过程中，当神经元给定的输入为正数时，输出为 1，使梯度消失的问题得到缓解，且计算成本较低。Relu 激活函数的公式如下：

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2.3)$$

(3) 损失函数：损失函数是深度神经网络在学习过程中更新参数的指标，也可理解为损失函数是表示当前状态下神经网络性能“恶劣程度”的指标。不同任务也对应着不同的损失函数。回归任务中，经常使用的是均方差 Mean Squared Error (MSE) 损失函数，其定义如下：

$$Loss_{MSE}(W, b) = \frac{1}{N} \sum_x \|y(x) - \hat{y}(x)\|^2 \quad (2.4)$$

其中  $W$  代表权重参数,  $b$  代表偏置参数,  $N$  代表批处理输入训练数据的个数,  $x$  表示输入数据,  $y(x)$  表示真实值即标签,  $\hat{y}(x)$  代表预测值。

二分类任务中, 常使用的损失函数是二分类交叉熵损失函数。该模型通常使用 Sigmoid 激活函数, 它能将模型的输出映射在  $(0,1)$  区间内, 即给定任意输入  $x$ , 模型预测为正类的概率  $\hat{y} \in (0,1)$ 。因此模型得到正类负类的概率分别为  $\hat{y}$  和  $1 - \hat{y}$ 。因此  $N$  输入的情况下的条件概率公式为:

$$p(y_i|x_i) = \prod_{i=1}^N \hat{y}_i^{y_i} (1 - \hat{y}_i)^{1-y_i} \quad (2.5)$$

二分类交叉熵损失函数在此基础上取最小化负对数即:

$$Loss(W, b) = -\sum_{i=1}^N (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) \quad (2.6)$$

式中  $i$  代表输入类别序号,  $N$  代表批处理输入样本个数,  $y_i$  代表真实值,  $\hat{y}_i$  代表预测概率,  $W$  和  $b$  表示参数。

多分类任务中使用的损失函数是多分类交叉熵损失函数, 与二分类相比区别在于真实值  $y_i$  不是具体的数, 而变成了一组 One-hot 向量<sup>[44]</sup>, 向量中真实分类下标所对应的值为 1, 其余为 0。使用的激活函数是 Softmax, 它将样本所有可能的分类预测概率都映射在  $(0,1)$  上, 且这组概率数据的和为 1。其表达形式如下:

$$Loss(W, b) = -\sum_{i=1}^N \sum_{j=1}^M y_i^j \log(\hat{y}_i^j) \quad (2.7)$$

式中  $N$  代表批处理样本个数,  $M$  代表类别个数,  $y_i$  代表 One-hot 向量,  $\hat{y}_i$  代表预测概率向量。

(4) 神经网络的梯度下降算法: 梯度下降算法是神经网络模型的核心, 神经网络模型依赖它完成各层权重参数的自动更新和调整, 参数经过多次调整后模型的分类和预测能力越趋近于样本总体真实情况。在数学领域梯度方向代表了函数增长速度最快的方向, 因此函数减小速度最快的方向即为梯度的反方向。梯度下降算法原理是计算函数在某一点沿梯度方向的最大变化率, 根据这一方向更新参数。常见的梯度下降算法包括 SGD、带有动量的 SGD、AdaGrad、Adam 等。以随机梯度下降法 (SGD) 为例, 假设损失函数为  $Loss$ , 则随机梯度下降的权重参数和偏置更新表达式如下:

$$\begin{cases} W \leftarrow W - \eta \frac{\partial Loss}{\partial W} \\ b \leftarrow b - \eta \frac{\partial Loss}{\partial b} \end{cases} \quad (2.8)$$

式中  $W$  代表权重参数,  $b$  代表偏置,  $\eta$  表示学习率。

### 2.4.3 常见神经网络

(1)卷积神经网络 CNN: 卷积神经网络 CNN 通常由卷积操作(Convolution)、池化操作(Pooling)和全连接网络(Fully connected network)组成。CNN 具有局部连接和权值共享的特点, 这些特性得力于卷积操作的存在。卷积操作是 CNN 的核心部分, 利用卷积核连接输入特征图和输出特征图。它的计算逻辑依赖于卷积核与等行等列的输入特征图局部信息对应位置相乘再相加完成, 不断重复移动卷积核在特征图的水平和竖直方向进行操作得到输出特征图, 这个过程类似于图像处理中的“滤波器运算”。单通道卷积运算计算方式如下图所示。

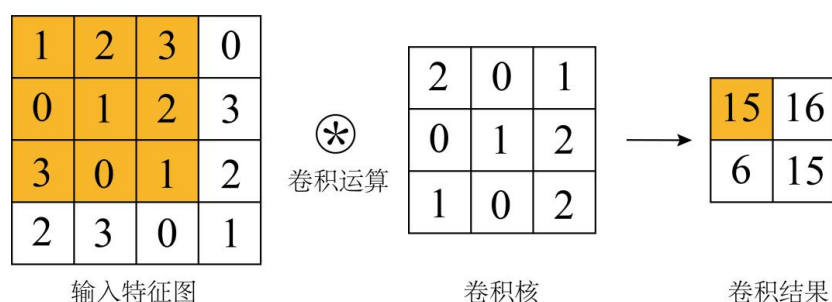


图 2.7 单通道卷积计算方式图

在这个过程中局部信息与卷积核的对应位置相乘再相加的操作称局部连接; 在移动的过程中卷积核不发生改变, 同时作用于整张特征图, 称权值共享。相较于传统的全连接神经网络, 这二者的结合大大降低了参数量, 也解决了全连接神经网络忽视输入数据形状的问题。多通道的卷积操作的原理图如下所示。

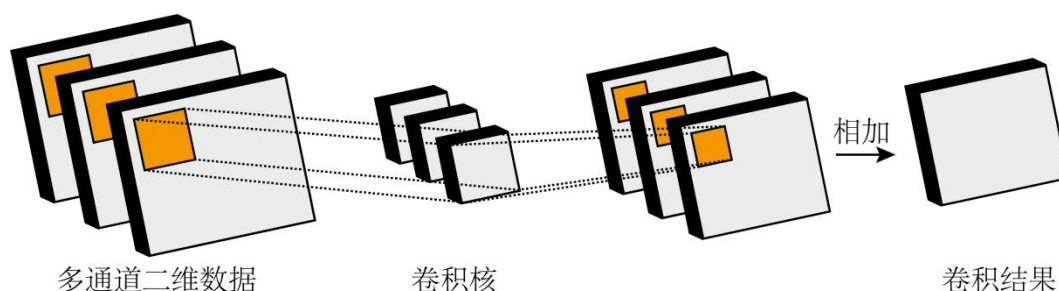


图 2.8 多通道卷积操作原理图

池化操作通常用于特征图的降维。类似于卷积操作中移动卷积核的方式, 对特征图的局部信息进行特征再提取, 通过这种方式减小特征图的长宽, 达到更深层次特征提取的目的。常见的池化操作分为最大值池化和平均池化, 最大值池化更关注局部信息反映强烈的地方, 下图展示了最大值池化的原理示意图, 其中池化大小为 2, 步长为 2。

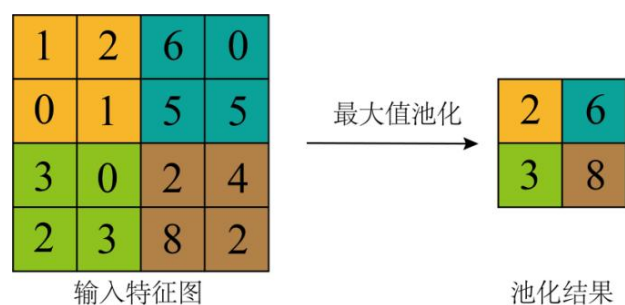


图 2.9 最大值池化原理示意图

下图展示了平均池化的原理示意图，其中池化大小为 2，步长为 2。平均池化更聚焦于局部信息的整体平滑度，所以平均池化输出局部信息的平均值。

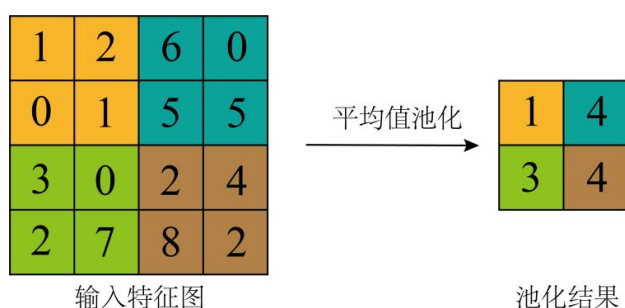


图 2.10 平均值池化原理示意图

全连接网络一般位于 CNN 结构的最后一部分，用于输出最终的结果，整体架构类似于图 2.5 中的神经元排列方式。全连接层由多个单一的神经元组成，每个神经元都连接着上一层的所有节点。神经元的个数直接决定了输出维度的大小。在实际过程中为了提高模型的非线性能力通常需要和激活函数搭配使用。

(3) 循环神经网络 RNN: 循环神经网络 RNN 是一类能够高效处理序列数据的 DNN, 它能够挖掘到数据中的时序信息以及语义信息。不同于全连接神经网络, RNN 引入循环结构, 能够针对性的处理输入之间前后关联的问题, RNN 的展开结构如下图所示。

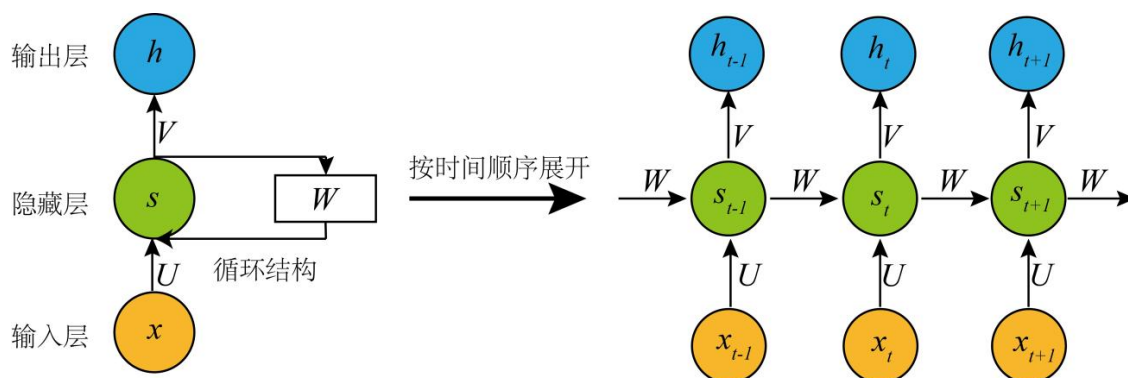


图 2.11 RNN 结构示意图

RNN 能够记住前向时刻的信息，是因为任一时刻隐藏层的输出  $h_t$  不仅由该时刻的输入  $x_t$  决定，还由上一时刻的隐藏层的状态  $s_{t-1}$  决定，这导致 RNN 对于受前文影响较大的数据十分敏感。传播过程的具体计算公式如下所示。

$$h_t = g(V \cdot S_t) \quad (2.9)$$

$$S_t = f(U \cdot x_t + W \cdot S_{t-1}) \quad (2.10)$$

式中  $h_t$  代表  $t$  时刻的输出， $S_t$  代表  $t$  时刻隐藏层的值， $W$ 、 $U$  和  $V$  表示权重参数，权重参数的值在一轮传播过程中不发生改变。

实践过程中发现针对长序列的数据应用 RNN 得到的预测结果和分类表现并不好。这主要是由于 RNN 无法解决长期依赖的问题。即当前系统状态依赖于很早之前出现的信息，RNN 的结构导致在经过多轮次传播后的梯度容易发生消失或爆炸的情况，从而丢失了长期信息。RNN 的变体长短期记忆网络 (Long Short-Term Memory, LSTM) 和双向 LSTM 的出现缓解了这个问题，下图展示了 LSTM 的结构组成。

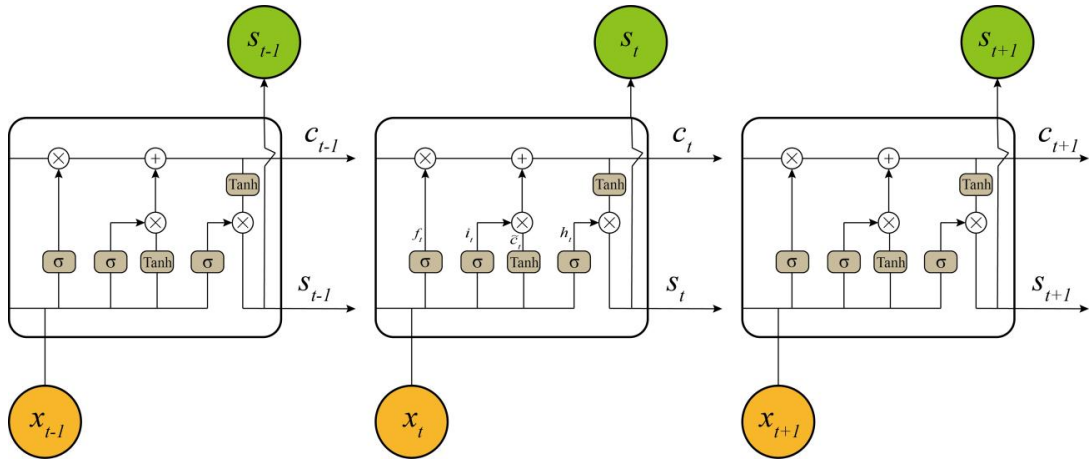


图 2.12 LSTM 的结构示意图

图 2.11 中展示了  $t$  时刻附近的 LSTM 模块，每个 LSTM 模块的结构是相同的。相较于传统 RNN 的隐层状态  $S_t$ ，LSTM 模块增加了细胞状态  $C_t$ ，该状态表示神经元在  $t$  时刻过后的记忆，即  $C_t$  收集了  $t+1$  时刻前神经网络对于所有输入信息的概括，这使得 LSTM 能够解决长期依赖的问题。 $t$  时刻的所有输入由上一时刻的细胞状态  $C_{t-1}$ 、上一时刻的隐层状态  $S_{t-1}$  和  $t$  时刻输入向量  $x_t$  三部分组成。任一时刻的隐层状态  $S_t$  都有上一时刻细胞状态  $C_{t-1}$  参与计算，并通过“门”逻辑结构进行细胞状态修改。不同“门”逻辑结构的运算如下列公式所示：

$$f_t = \sigma(W_f \cdot [S_{t-1}, x_t] + b_f) \quad (2.11)$$

$$i_t = \sigma(W_i \cdot [S_{t-1}, x_t] + b_i) \quad (2.12)$$

$$\tilde{C}_t = \text{Tanh}(W_c \cdot [S_{t-1}, x_t] + b_c) \quad (2.13)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (2.14)$$

$$h_t = \sigma(W_h \cdot [S_{t-1}, x_t] + b_h) \quad (2.15)$$

$$S_t = h_t * \text{Tanh}(C_t) \quad (2.16)$$

式中  $W$  和  $b$  表示权重参数和偏置。式(2.11)表示遗忘门的计算过程,式(2.12)-式(2.14)表示更新门的计算过程,式(2.15)和式(2.16)表示输出门的计算过程。

双向 LSTM 是 RNN 基于 LSTM 的一种变体,它实现了记忆信息的双向传播,使得神经元能够同时感知序列前方和序列后方两组信息。它的组成结构本质上是由两条方向相反的 LSTM 链组成。下图展示了这种结构。

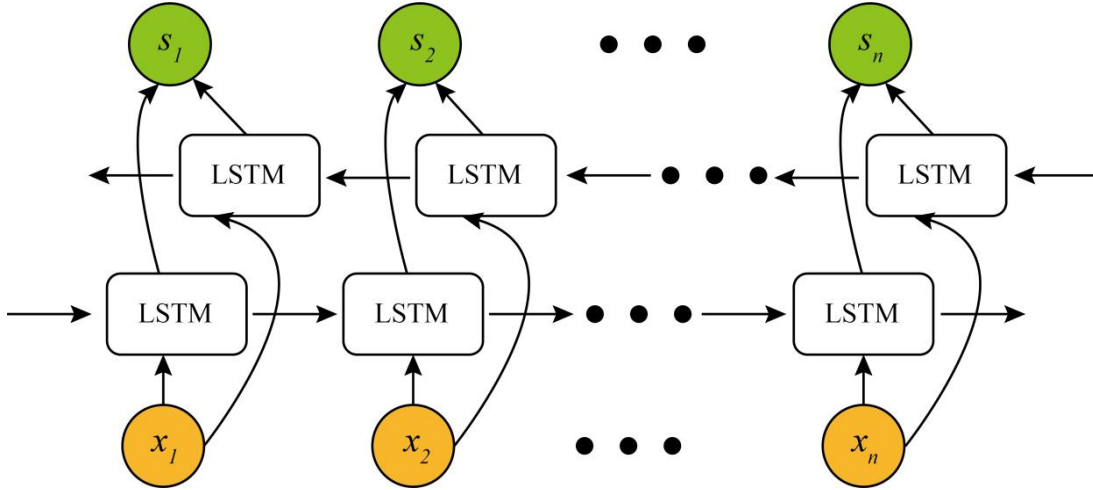


图 2.13 BiLSTM 的结构示意图

## 2.5 本章小结

本章首先对心律失常相关基础知识进行阐述,详细介绍了心律失常产生的原因,常见的心律失常类型以及心律失常患者在日常生活的症状以及正常心跳的心电图波形结构展示。

其次介绍了十二导联 ECG 的起源和发展、十二导联的组成部分、每一部分的作用以及在测量不同导联时电极的安置方式,同时阐述了临床上十二导联的作用。之后列举了到目前为止常见的开源部分导联 ECG 数据库以及十二导联 ECG 数据库。最后介绍了深度学习这项技术的来源、发展以及对当前科技社会的影响。并

详细说明了深度学习的工作原理和实现方式。通过这一章的介绍，可以初步理解应用深度学习的发展和心电图领域的关系，为后续研究的详细阐述提供理论依据。



## 第3章 基于 DSE-ResNet 的十二导联心律失常自动识别

### 3.1 引言

十二导联 ECG 由于其简单、无创和可靠等优点,常用于心律失常的临床诊断。全世界每年采集获得数百万的心电图记录<sup>[1]</sup>。巨大的诊断工作量导致专业医生诊断心律失常效率低下,误诊现象增多。因此 ECG 的数字化与自动分类算法的结合越来越受到关注。近些年来,机器学习相关算法的出现缓解了这种情况,它针对这类任务通常包含了专家特征构建和分类器分类两个模块,尽管机器学习在心律失常分类方面有着广泛的研究应用,但仍然存在一些问题需要解决。例如基于主观因素的特征工程忽略了 ECG 中潜在的重要特征,这可能会影响最终的分类性能。

与机器学习方法不同,DNN 可以从原始输入数据中自动学习特征,而无需大量特征工程。许多研究已经将 DNN 用于基于单导联或多导联 ECG 的心律失常自动分类中。

例如,Ullah<sup>[20]</sup>等人将单导联心电图转换为二维频谱图像,并使用二维 CNN 学习频谱图的特征来实现心律失常的自动分类。他们的模型在 MIT-BIH 数据集实现了 99.11% 平均分类准确率。Hannun<sup>[22]</sup>等人开发了一种可根据单导联心电图对十二种心律失常类别进行分类的 DNN。实验结果显示他们的 DNN 模型的平均  $F_1$  得分(0.837)超过了心脏病专家的平均得分(0.780)。这表明端到端的深度学习方法可以基于单导联识别各种心律失常。同时他们提到有限的采样时长和数据库只包含单导联 ECG 等因素限制了得出更多有效结论。与单导联相比,多导联心电图包含了更多有价值的信息<sup>[45]</sup>,更有利于心律失常的自动分类。Zhang<sup>[25]</sup>等人提出了一种可解释的 DNN,用于十二导联 ECG 的自动诊断。他们的实验表明,在单导联 ECG 上训练的 DNN 的性能低于同时使用十二导联产生的性能。Zhao<sup>[26]</sup>等人将患者的年龄和性别作为辅助信息输入到 DNN 中,DNN 模型在 The PhysioNet/Computing in Cardiology Challenge 2020<sup>[46]</sup>中获得了第二名的测试结果。

这些研究促进了深度学习在心律失常自动分类中的应用。然而一些基于单导联心律失常自动分类的研究表明,利用单导联 ECG 导致 DNN 分类性能降低。这促使本研究选择十二导联而不是单导联作为实验数据。部分基于十二导联 DNN 的训练过程分为两步,首先对导联进行逐条训练,然后融合每个导联的训练特征,最后得到分类结果。这使得在训练开始时没有注意导联之间的关系。基于这些问题,本章提出了二维化十二导联 ECG 和 DSE-ResNet 模型。本章工作的主要贡献可以概括如下:



(1) 提出了一种多通道时序数据的二维化转换方法。将原始的十二导联 ECG 拼接成类似灰度图一样的二维化数据，每一列代表单导联的时间序列，每个“像素”代表导联某一时刻的电压值。

(2) 提出了一种二维 DNN 模型 DSE-ResNet 用于处理多通道时间序列 ECG 信号。基于二维卷积神经网络的特性，DSE-ResNet 可以在训练阶段学习导联内部和导联之间的特征。

(3) 提出了一种针对十二导联 ECG 的切片规则来扩展训练集。

(4) 采用正交试验选择超参数。在评估模型阶段，使用基于投票策略的集成模型增强分类性能。

最终模型在 CPSC2018 隐藏测试集的实验结果表明，DSE-ResNet 模型对正常节律和 8 类心律失常进行分类的平均  $F_1 = 0.817$ 。同时与 CPSC2018 竞赛中最优的几种模型相比，本文提出的模型在 2 个亚异常类型中取得了最优的  $F_1$  分数。这表明基于二维化十二导联 ECG 和 DSE-ResNet 在检测特定心律失常方面具有优势，有潜力作为辅助工具帮助医生进行心律失常分析。

## 3.2 研究方法

### 3.2.1 问题定义

本章节旨在基于二维化十二导联 ECG 实现正常心律和 8 类心律失常的自动分类。所提出模型的输入  $x$  包括二维化 ECG 信号和患者的年龄、性别信息，输出是  $y$ ，对应于正常节律和 8 类心律失常的预测值。输入信息和真实标签  $y_i$  构成训练集：

$$X = \{(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)\} \quad (3.1)$$

式中代表训练集样本个数。模型的训练目标是在有限的训练时期内最小化 Softmax 交叉熵损失函数，其中 Softmax 交叉熵损失函数如公式 (3.2) 所示：

$$Loss(X) = -\frac{1}{n} \sum_{i=1}^n \log \frac{e^{p(x_i, y_i)}}{\sum_j e^{p(x_i, y_j)}} \quad (3.2)$$

式中  $p(x_i, y_i)$  和  $p(x_i, y_j)$  表示模型将输入  $x_i$  预测为真实标签  $y_i$  和不同标签  $y_j$  的概率。由于分类类别为 9 类，故  $j$  的取值为  $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ 。

### 3.2.2 二维化十二导联 ECG

#### (1) 数据来源

本章节研究内容使用的数据库是来自 CPSC2018 的十二导联 ECG 数据集，采样频率为 500Hz。CPSC2018 数据集共有 9831 组样本，其中 6877 组样本被开源用于训练，2954 组样本被保密用于测试。每组样本包含了原始十二导联心电图信号、患者年龄、性别和信号对应的标签。其中标签分别对应了 9 种类别：正常节

律、心房颤动（AF）、一度房室传导阻滞（I-AVB）、左束支传导阻滞（LBBB）、右束支传导阻滞（RBBB）、房性早搏（PAC）、室性早搏（PVC）、ST 段压低（STD）和 ST 段抬高（STE）。这些心律失常类型还可分为四类亚异常类型，例如 LBBB 和 RBBB 可以统称为传导阻滞类型（Block）；PAC 和 PVC 可以统称为过早收缩（PC）；STD 和 STE 可划分为 ST 段改变（ST）。通过对比模型在亚异常类型和具体类型的分类性能，可以整体评价该模型对不同心律失常类型的识别敏感度。表 3.1 中显示了 CPSC2018 数据源的详细信息。

表 3.1 CPSC2018 数据库详细信息

No.	心律失常类型	样本数量	采样时间长度（单位 s）			训练集	小样本测试集	隐藏测试集
			平均	最小	最大			
1	正常	918	15.43	10	64	895	69	394
2	AF	1098	15.04	9	74	1112	79	466
3	I-AVB	704	14.27	10	54	695	45	295
4	LBBB	207	14.94	9	65	203	15	97
5	RBBB	1695	14.62	10	118	1691	124	756
6	PAC	574	19.43	9	74	546	47	250
7	PVC	653	20.92	6	144	826	44	276
8	STD	826	15.50	8	138	825	58	340
9	STE	202	17.15	10	60	216	19	80
	总计	6877	15.95	6	144	7117	500	2954

## （2）二维化处理原理

临床医学实践中，心脏病专家通常需要观察多个导联 ECG 作为判断心律失常的依据。如 PAC 患者的心电图异常主要表现在导联 V1、II 和 aVF，而 LBBB 患者典型的心电图异常主要表现在导联 I、V1、V2、V5、V6 和 aVR<sup>[47]</sup>。

不同心律失常的检测需要十二导联心电图的综合信息，这意味着导联内部和导联间特征在心律失常的分类中起着重要作用。为了利用二维卷积神经网络局部连接的特性同时提取十二导联 ECG 的导联内部和导联间特征，本研究对十二导联 ECG 进行二维化处理。具体来说，对于一组样本的任一导联  $a \in \mathbb{R}^{L \times 1}$ ，组合所有导联拼接在一起形成二维心电图  $A \in \mathbb{R}^{L \times 12}$ ，其中表示原始十二导联 ECG 的长度（采样点个数）。尽管 CPSC2018 数据库提供的不同样本信号的采样长度不同，但是同一样本不同导联的长度是相同的。图 3.1（a）展示了原始十二导联 ECG 信号，按照从左至右的排列顺序依次为导联 I、II、III、aVR、aVL、aVF、V1、V2、V3、V4、V5 和 V6。如图 3.1（b）所示，原始十二导联 ECG 被拼接成灰度图一样的有界二维化数据，每列代表一个导联的时间序列，每列中的“像素”表示该

导联在不同采样点的电压值。

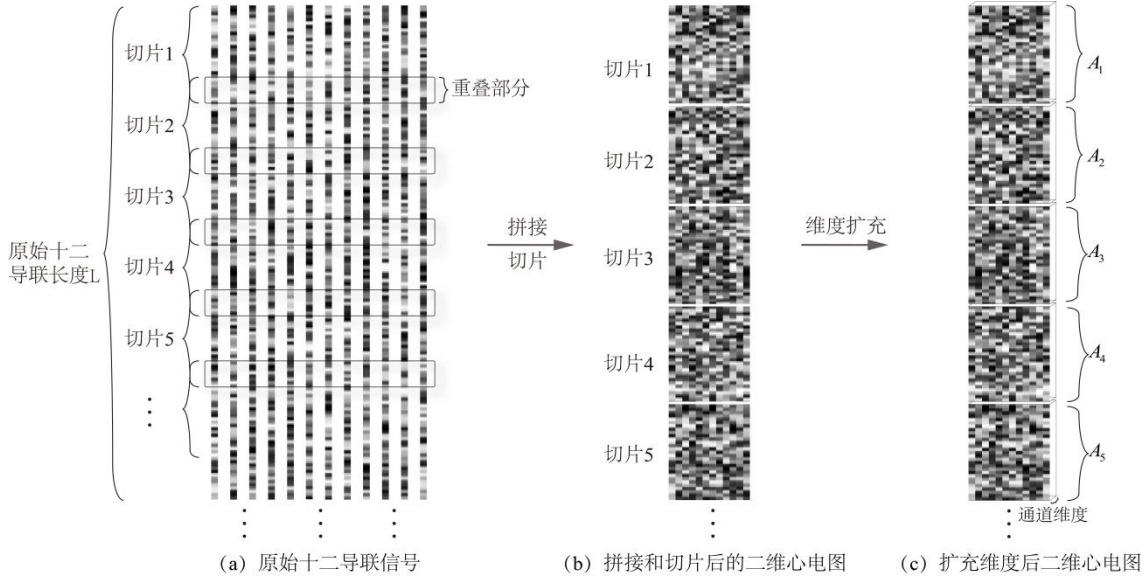


图 3.1 二维化原理示意图

注：（a）原始十二导联 ECG；（b）切片和拼接十二导联信号；（c）切片后的二维心电图

### （3）切片规则

从表 3.1 可以看出，正常心律和 8 类心律失常的样本数量差距较大，不同样本之间原始十二导联 ECG 的长度也不相同。为了充分利用 CPSC2018 的数据，统一训练集十二导联 ECG 的长度，并对二维化后的心电图进行切片。图 3.1（a）和图 3.1（b）展示了切片的原理示意图。

由于 CPSC2018 的测试集不对外开放，本研究从 6877 组开放获取数据中分离出前 500 组数据作为离线小批量测试集。应用少量测试集的主要作用是比较最优子模型和集成模型的性能。其余 6377 组样本中的十二导联 ECG 被切片并用于训练。不同样本的原始十二导联 ECG 长度对应不同切片方法，切片的具体步骤如下：

原则 1：如果在原始十二导联 ECG 长度  $A \in \mathbb{R}^{L \times 12}$  中  $L < 8192$ ， $A$  在第一维的长度补 0 至  $L=8192$ 。

原则 2：如果在  $A \in \mathbb{R}^{L \times 12}$  中， $L$  满足  $8192 \leq L < 1.5 \times 8192$ ，将  $A$  尾部多余的数据截掉至  $L=8192$ 。

原则 3：如果在  $A \in \mathbb{R}^{L \times 12}$  中， $L$  满足  $L \geq 1.5 \times 8192$ ，则将  $A$  切成  $n$  片。切片长度为 8192（16.384s），切片之间的重叠长度为 4096。经过切片二维心电图  $A$  变为  $n$  个二维心电图样本。其中切片数量如式（3.3）所示，式中  $[x]$  表示小于  $x$  的最大整数。

$$n = \lfloor (2 \times L) / 8192 \rfloor \quad (3.3)$$

需要注意的是，切片长度决定了输入到 DNN 中的十二导联 ECG 的长度。

DSE-ResNet 中有多个 0.5 倍的下采样过程。为了便于下采样后的维度统计，选择 2 的指数幂的长度作为切片长度。在 500Hz 的采样率下，8192 的切片长度表示大约 16.384s 的十二导联 ECG 信号长度。本研究统计了 CPSC2018 中原始样本的长度分布，样本的平均长度为 15.95s，因此选择最接近的 8192（16.384s）作为切片长度。12.7% 的样本长度超过平均长度的 1.5 倍，这部分样本具有更多的 ECG 信息。为充分利用数据以及防止过拟合通过切片来扩充训练集。切片后训练集中的不同心律失常样本个数如表 3.1 所示。

#### （4）维度扩充

切片后的二维心电图  $A \in \mathbb{R}^{8192 \times 12}$  被增加了一个维度，使得十二导联信号的维度满足二维卷积（Conv2D）层对输入数据维度的要求。新添加的维度被称为通道维度，将二维心电图  $A \in \mathbb{R}^{8192 \times 12}$  扩充为  $A \in \mathbb{R}^{8192 \times 12 \times 1}$ ，其中导联长度为 8192，导联个数为 12，通道维度为 1。在维度扩充的过程中，实际电压值没有发生改变，而是在逻辑上增加了一个维度。训练过程中每个卷积层输出特征图的通道数与卷积核的数量同步变化。图 3.1（c）展示了对二维心电图扩充维度的过程。

#### 3.2.3 DSE-ResNet 模型构建

异常心电信号主要表现为波形形状和周期性节律的变化<sup>[48]</sup>。一些异常心电信号是周期性的，另一些异常心电信号只在少数心跳周期中产生。因此，DSE-ResNet 包含用于提取导联内部和导联间特征的残差结构 ResNet 和用于提取二维心电图全局特征 DSE 部分。此外，患者年龄和性别被引入作为训练的辅助特征。

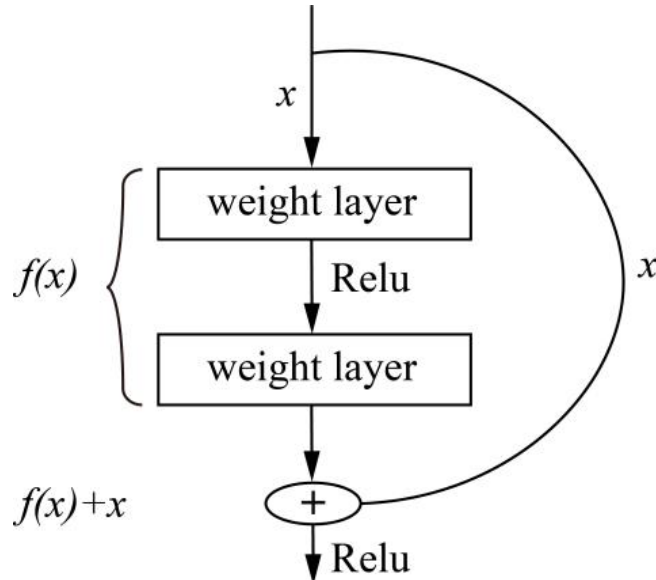


图 3.2 残差块结构

ResNet 是如今深度学习领域中最常见的模型，它引入快捷连接（shortcut）实现特征的恒等映射，缓解了反向传播过程中出现梯度消失和梯度爆炸的现象。关

于 Resnet 的残差块结构如图 3.2 所示。弯曲的路径即为快捷连接，中间的 weight layer 指代具有参数的不同神经网络层，激活函数使用 Relu,  $f(x)$  表示经过正向传播后提取的特征。

残差块 (Residual Block) 通常被引入 CNN 结构中，以改善通过网络的梯度流并支持更深层网络的训练。本研究提出模型中的 ResNet 由 1 个残差块 Residual Block-1 和 9 个 Residual Block-2 组成。每个残差块都有 2 个二维卷积 (Conv2D) 层，用于提取二维心电图局部特征包括导联内部和导联间特征。整个 ResNet 部分有二十个 Conv2D 层，其中卷积核的大小为 (32,1)。第一个 Conv2D 卷积层有十二个卷积核，最后四个 Conv2D 层分别有 192 个卷积核，中间每四个 Conv2D 层卷积核数量翻倍。除此之外，每个 Residual Block 还包括了 Relu 激活函数层，主要用于提升模型的非线性能力。批量归一化层 (Batch Normalization, BN) 可以将输入特征映射到数据范围较小的区间内，减小数据的发散程度，缓解了 S 型激活函数出现的梯度消失的问题。同时加入了 BN 的神经网络在训练时学习率的衰减速度比较大，算法收敛更快，因而模型的训练速度更快。批量归一化的具体计算方式如下：

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.4)$$

$$\rho = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \quad (3.5)$$

$$\tilde{x}_i \leftarrow \frac{x_i - \mu}{\sqrt{\rho^2 + \varepsilon}} \quad (3.6)$$

$$y_i \leftarrow \gamma \tilde{x}_i + \beta \quad (3.7)$$

式中  $n$  表示批处理样本个数， $x_i$  表示输入样本， $\mu$  表示样本均值， $\rho$  表示方差， $\tilde{x}_i$  表示归一化后的样本， $y_i$  表示 BN 层的输出， $\gamma$  和  $\beta$  表示可训练的参数， $\varepsilon$  是一个无限趋近于 0 的数，保证分母不为 0。

Residual Block 中的丢弃层 (Dropout) 主要用于解决过拟合的问题。其工作原理是在每个小批量的正向传播过程中屏蔽指定比例的神经元参与训练，利用这种方式能够减少隐藏层神经元之间的相互作用，使权值的更新不再依赖于存在固定关系的隐藏层节点的共同作用，迫使网络去学习更加鲁棒的特征。

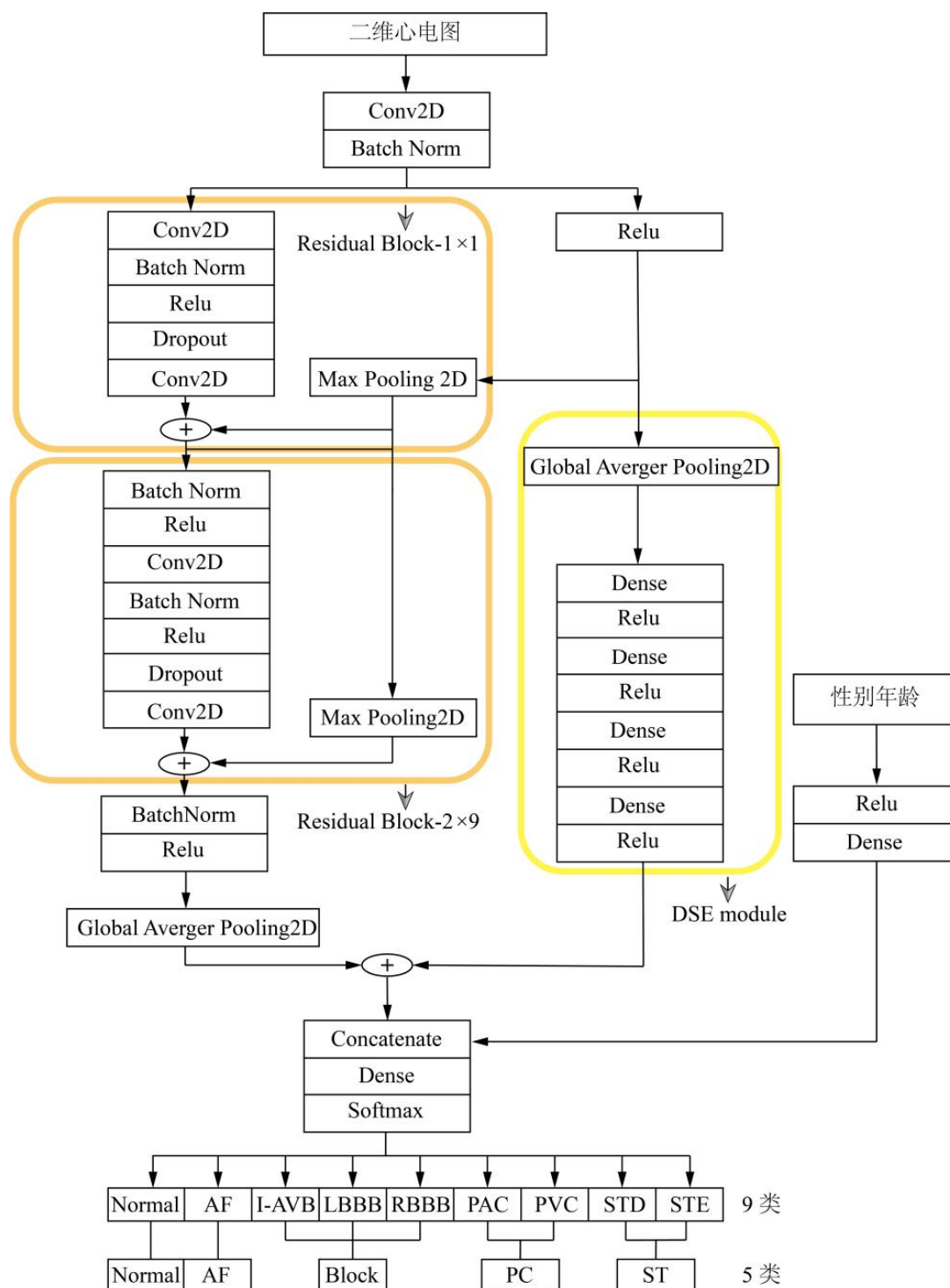


图 3.3 DSE-ResNet 整体结构

Squeeze-and-excitation (SE) module 是一种通道注意力模型。它可以对不同通道维度给出不同的重要性评分<sup>[49,50]</sup>。通常与 ResNet 结构穿插连接。其结构如图 3.4 所示。

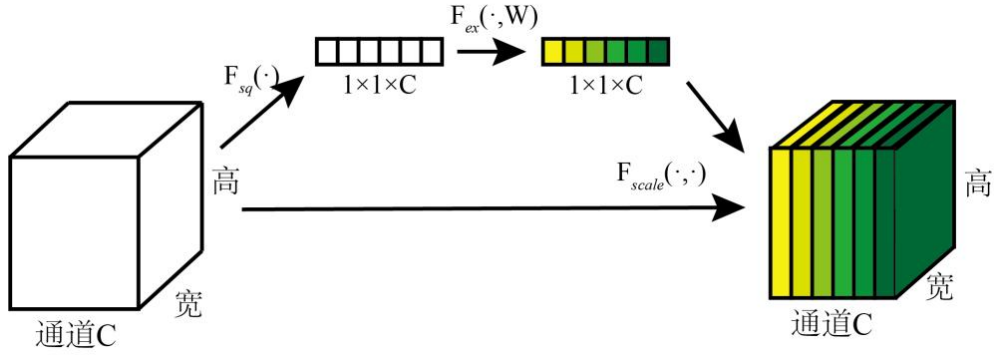
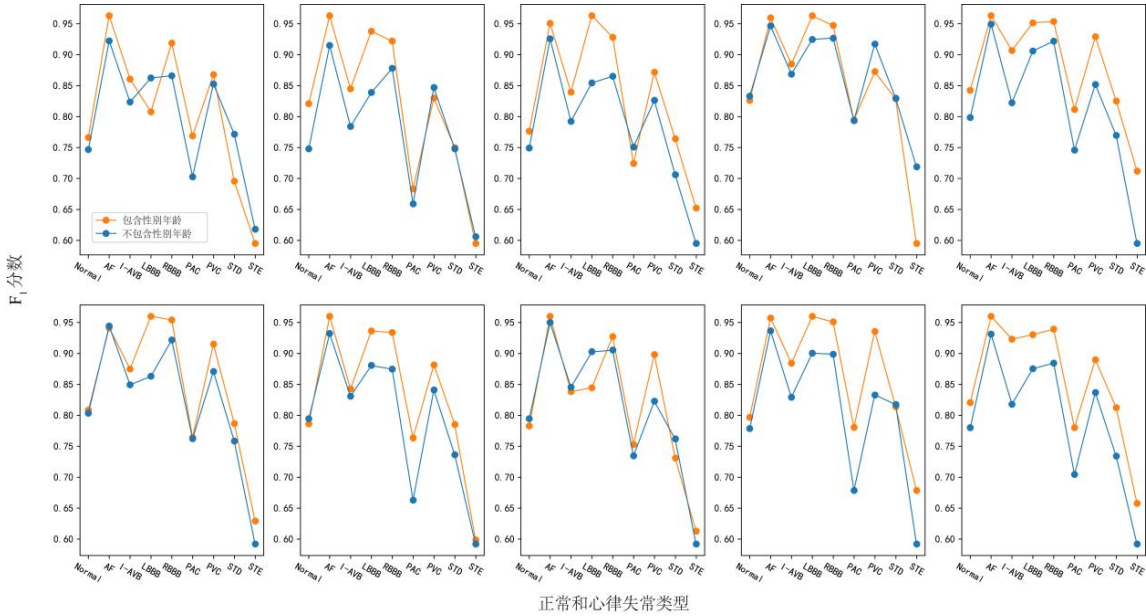


图 3.4 SE module 的结构示意图

不同于常见的和 ResNet 结构穿插连接的方式，本章提出模型中的 Detached SE (DSE) 模块独立于任何残差块。仅使用它提取二维心电图不同通道维度的全局特征。然后 DSE 模块中的四个全连接层将提取的全局特征映射到一个新的特征空间。尽管加入 DSE 模块会增加整个模型的计算复杂度，但二维心电图不同通道维度的整体特征是值得关注的。图 3.3 展示了 DSE-ResNet 的整体结构。不同年龄和性别的患者在同一种心律失常上可能表现出不同的波形状态。图 3.5 显示多组模型在包含年龄和性别的情况下获得的心律失常的  $F_1$  分数优于不包含年龄和性别的情况。因此，将年龄和性别作为辅助特征引入训练中，有助于 DSE-ResNet 捕捉患者基本信息对心律失常的影响。

图 3.5 具有与不具有年龄和性别的十个模型的  $F_1$  分数对比

### 3.2.4 正交试验

适当的超参数可以提升模型学习的性能和效果。赵哲耘等人<sup>[51]</sup>的研究表明基于正交试验优化卷积神经网络模型的超参数具备可行性。正交试验设计



（Orthogonal experimental design, OED）在本研究中用来选择超参数组合。OED 是一种研究多因素、多水平问题的设计方法。其中因素表示影响实验结果的各个变量，在本章节中表示超参数；水平表示每个因素的不同取值或处理条件，在本章节中表示不同超参数的取值。OED 基于正交性从整个测试集合中选取一些分布均匀的代表点进行测试。选择代表点的过程是通过构建正交表来实现的。正交表是一种特殊设计的表格，具有均匀分布和正交性的特点。通过正交表的选择，可以确保各个因素之间的相互独立性，从而消除因素间的相互干扰。应用正交试验的优点是进行试验次数少，效率高。通过汲取部分基于深度学习的研究中关于超参数设置的相关经验<sup>[22]</sup>，本章节选取了一些超参数进行调整并给出各超参数的一组估计值。这些估计值构成了正交表的整个测试点。成对独立组合测试（PICT）<sup>[52,53]</sup>工具被选定用来构建正交表，以获得多组超参数的代表性组合。与常见的优化超参数的随机选择和网格搜索方法不同，PICT 是一种用于软件测试领域的选择组合参数技术，用于减少系统测试用例输入的数量。神经网络中大量超参数的选择是 PICT 首选的应用场景。

### 3.2.5 集成模型

集成模型通过构造和组合多个学习器来完成学习任务<sup>[29]</sup>。与单个模型的性能相比，集成模型往往能取得更好的分类性能和泛化能力<sup>[54]</sup>。集成模型被用来减少模型的整体误差。集成模型包含多个学习器，每个学习器都是基于超参数的代表性组合训练的最优 DSE-ResNet 模型。

个体学习器在本文中被称为单一最优模型。集成模型使用投票策略来集成所有单一最优模型。每个单一最优模型都会对同一个测试样本给出一个预测值，基于少数服从多数的多模型投票策略，集成模型将得票最多的预测值作为最终输出值。集成模型的使用可以有效提高模型的性能和容错能力。

## 3.3 实验细节

### 3.3.1 实验设备环境

本章节所提出的模型使用 Keras 框架构建训练。所有实验均在配备 Quadro P2200 显卡和显存为 5G 的服务器上运行。软件环境采用 3.6 版本的 Python 编译器，使用 Pycharm 构建项目。

### 3.3.2 数据预处理

#### （1）去噪

原始信号中存在肌肉噪声、电流噪声和基线漂移，这些噪声可能影响模型的性能。为了去除这些噪声对十二导联信号的影响，使用 Butterworth<sup>[55]</sup>带通滤



波器滤除了频率为 0.5Hz 至 49Hz 之外的噪音信号，这个区间源自 CPSC2018<sup>[63]</sup> 比赛中各团队滤除噪音的大致区间。图 3.6 显示了用 Welch<sup>[56]</sup> 方法计算异常样本导联 I 利用 Butterworth 带通滤波器滤波前后的功率谱密度曲线，根据曲线能够观察到高频噪声被衰减。功率谱密度曲线的可视化应用了不同窗口和不同窗口长度的 Welch 方法。

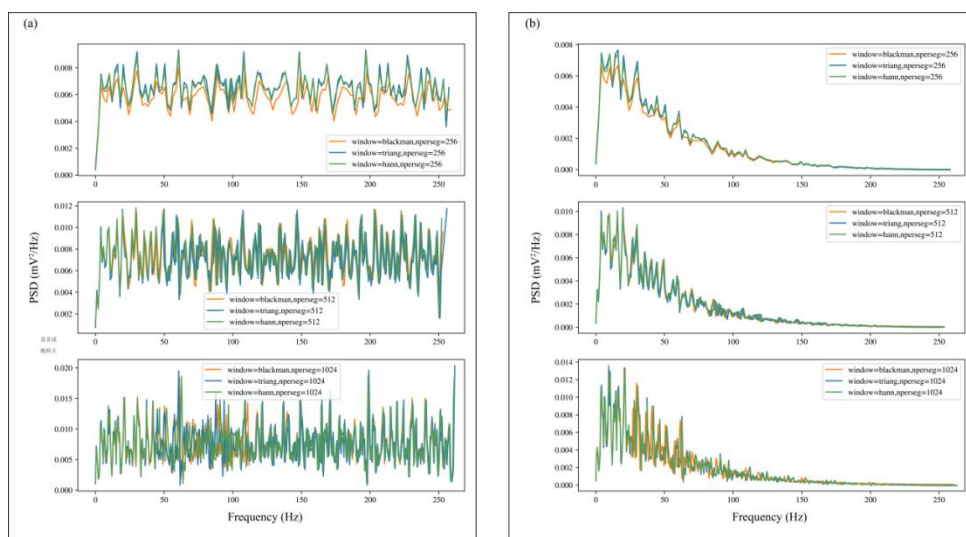


图 3.6 不同窗口和不同窗口长度应用 Welch 得到的功率谱密度曲线

注：（a）异常样本导联 I 信号滤波前的功率谱密度曲线；（b）异常样本导联 I 信号滤波后的功率谱密度曲线。每个子图使用相同的窗口长度和不同的窗口。window 表示窗函数，包括布莱克曼窗、汉宁窗和三角窗。nperseg 表示窗口长度包括 256、512 和 1024。

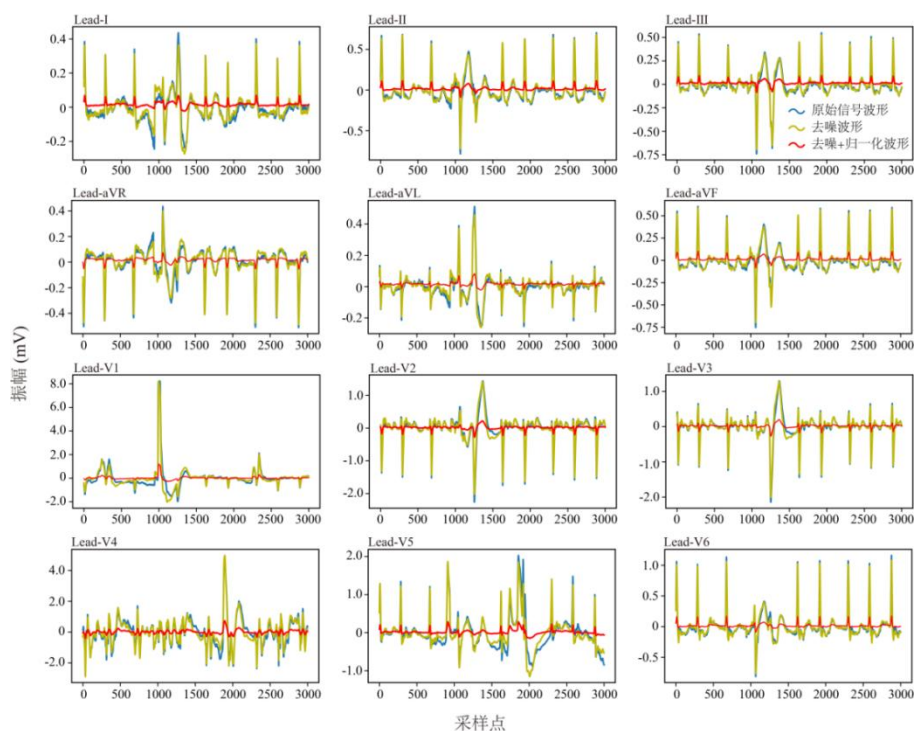


图 3.7 十二导联预处理对比波形图

## （2）最大最小值归一化

时间序列数据在特定情况下取值范围很广，因此需要将其缩放到固定值区间以加快学习过<sup>[57]</sup>。原始十二导联 ECG 中电压值的幅值为 $[-20.9\text{mV}, 20.7\text{mV}]$ ，不同导联间的幅值差异较大。从图 3.7 可以看出，原始十二导联 ECG 的最大和最小振幅呈对称区间分布。因此本研究使用最大最小值归一化<sup>[27,58]</sup>将二维心电图的电压值幅值缩放到对称区间 $[-3\text{mV}, 3\text{mV}]$ ，最大最小归一化如公式（3.8）所示：

$$A_{ij} \leftarrow R_{min} + \frac{(R_{max} - R_{min})(A_{ij} - A_{min})}{(A_{max} - A_{min})} \quad (3.8)$$

式中 $R_{max}=3\text{mV}$ ， $R_{min}=-3\text{mV}$ ，二者表示归一化区间的边界值， $A_{ij}$ 为二维心电图中第  $i$  行第  $j$  列的电压值， $A_{max}$ 和 $A_{min}$ 分别表示二维心电图中最大和最小电压值。图 3.7 显示了随机某个样本经过去噪和归一化的十二导联波形图。

### 3.3.3 超参数组合选择

OED 被用来确定超参数值的组合。首先批处理大小（batch-size）控制在显卡能承受的最大极限。其次，选取学习率（learning rate）、丢弃神经元比例（dropout）和梯度下降算法中的动量（momentum）三个超参数进行正交试验。根据先前研究的经验<sup>[22]</sup>，learning rate 的值集是 $\{0.05, 0.1, 0.15\}$ ；dropout 的值集是 $\{0.3, 0.5, 0.8\}$ ；momentum 的值集是 $\{0.5, 0.7, 0.9\}$ 。使用 PICT 构建正交表来组合和匹配预设值。表 3.2 给出了通过 PICT 配置的超参数预设值的组合。对每组超参数组合的模型进行五折交叉验证，然后选择验证集中平均损失最低的一折模型作为单一最优模型。

表 3.2 经过 PICT 选择出的超参数组合

试验组序号 No.	learning rate	dropout	momentum
1	0.1	0.8	0.5
2	0.15	0.3	0.9
3	0.05	0.5	0.5
4	0.15	0.8	0.7
5	0.05	0.3	0.7
6	0.1	0.5	0.9
7	0.15	0.3	0.5
8	0.05	0.8	0.9
9	0.1	0.3	0.7
10	0.15	0.5	0.7

### 3.4 结果

#### 3.4.1 评价标准

算法的分类性能可以通过准确率（Accuracy）、精确率（Precision）、特异度（Specificity）、灵敏度（Sensitivity）和 $F_1$ 分数来评价<sup>[59,60]</sup>。对于多分类任务，平均 $F_1$ 分数<sup>[61]</sup>是衡量分类性能的重要指标。 $F_{1i}$ 表示第 $i$ 类心律失常的 $F_1$ 分数，它是精确率 $F_p$ 和召回率 $F_R$ 的调和平均值，其中 $F_p$ 描述了预测的正样本中有多少是真正的正样本， $F_R$ 描述了有多少真正的正样本被挑选出来。 $F_{1i}$ 的定义如下：

$$F_{1i} = \frac{2(F_p \times F_R)}{F_p + F_R} \quad (3.9)$$

式中 $F_p = TP/(TP + FP)$ ， $F_R = TP/(TP + FN)$ 。 $TP$ 表示正样本被分类为正的个数， $FP$ 表示负样本被分类为正的个数， $FN$ 表示正样本被分类为负的个数。所有类型的平均 $F_1$ 得分是评价模型整体性能的综合评价指标，定义为：

$$F_1 = \frac{1}{9} \sum_{i=1}^9 F_{1i} \quad (3.10)$$

根据 CPSC2018 评分细则还应计算亚异常类型的 $F_1$ 分数，即 AF、Block、PC 和 ST。此外，准确率、敏感度和特异度也被用作评价 DSE-ResNet 性能指标，它们被定义为：

$$Acc_i = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.11)$$

$$Se_i = \frac{TP}{TP + FN} \quad (3.12)$$

$$Sp_i = \frac{TN}{TN + FP} \quad (3.13)$$

式中 TN 是分类为负样本而实际也是负样本的数量。应该注意的是，召回率 $F_R$ 和敏感度在数值上是相同的。

#### 3.4.2 在小批量测试集的表现

本研究比较了单一最优模型和集成模型（Ensemble Model, EM）基于小批量测试集（500 个样本）的 $F_1$ 分数，其中每个单一最优模型是基于超参数的代表性组合训练的最优模型（超参数组合见表 3.2）。集成模型是基于投票策略来整合所有单一最优模型。

表 3.3 集成模型与单一最优模型在小样本测试集上的F<sub>1</sub>分数

No.	平均 $F_1$	正常及 8 类心律失常								亚异常类型			
		正常	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE	Block	PC	ST
1	0.783	0.739	0.962	0.846	0.786	0.912	0.742	0.854	0.659	0.545	0.885	0.793	0.629
2	0.816	<b>0.821</b>	0.963	0.845	0.938	0.922	0.683	0.83	0.75	0.595	0.903	0.761	0.709
3	0.81	0.745	0.955	0.821	<b>0.97</b>	0.928	0.682	0.86	0.73	0.595	0.905	0.773	0.695
4	0.776	0.738	0.933	0.824	0.938	0.915	0.692	0.806	0.742	0.4	0.896	0.747	0.688
5	0.835	0.787	0.954	0.876	0.938	<b>0.941</b>	0.744	<b>0.907</b>	0.763	0.606	<b>0.926</b>	0.826	0.728
6	0.824	0.783	0.919	0.851	0.938	0.932	0.738	0.892	0.761	0.6	0.912	0.814	0.727
7	0.817	0.763	<b>0.969</b>	0.83	0.941	0.938	0.736	0.876	0.762	0.541	0.911	0.807	0.704
8	0.78	0.743	0.938	0.804	0.811	0.902	0.71	0.87	0.686	0.556	0.867	0.789	0.652
9	0.832	0.76	0.938	0.857	0.941	0.931	0.742	0.914	<b>0.779</b>	0.629	0.914	0.824	0.743
10	0.828	0.787	0.942	<b>0.901</b>	0.909	0.919	0.742	0.864	0.778	0.606	0.914	0.802	<b>0.738</b>
EM	<b>0.843</b>	0.787	0.949	0.87	<b>0.97</b>	0.935	<b>0.764</b>	0.897	0.748	<b>0.667</b>	0.922	<b>0.83</b>	0.729

### 3.4.3 在 CPSC2018 隐藏测试集的表现

图 3.8 展示了单一最优模型（learning rate=0.15, dropout=0.5, momentum=0.7）在训练集和验证集上的损失和准确率的变化曲线。验证集主要用来观察训练过程中模型在非训练集的损失和准确率曲线变化情况，通过观察曲线平稳性来控制模型训练的次数，防止出现过拟合现象。图 3.8 中模型损失和准确率曲线从在训练周期（epoch）为 30 次时开始趋于稳定。实验过程中尝试将 epoch 增加到 70 次发现过拟合现象，因此采用早停（early stopping）方法，将训练次数减少到 50 个训练周期。

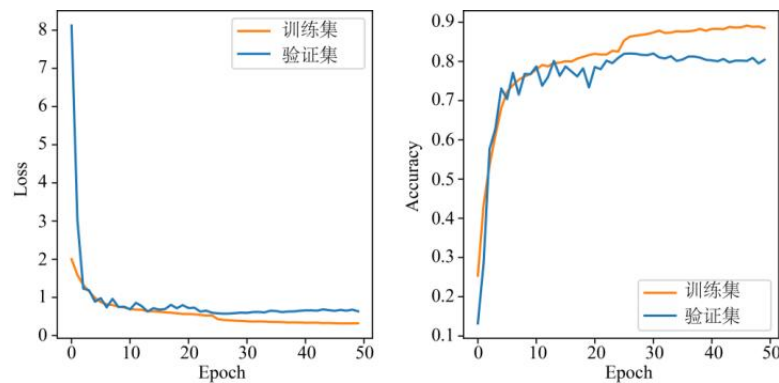


图 3.8 损失和准确率变化曲线

十组不同超参数组合的最优单一模型分别训练完成后，将这些模型组合形成集成模型并提交给 CPSC2018 官方工作人员，得到了基于隐藏测试集（2954 组样本）的测试结果。图 3.9 显示了集成模型在测试集得到的混淆矩阵。对于亚异常型 ST（STD 和 STE 的统称），根据混淆矩阵的结果显示 53 个标有 STD 标签的样本和 27 个标有 STE 标签的样本被预测为正常（Normal），19 个标有 Normal 标签的样本被预测为 STD。引起 ST 段改变的疾病在病理学角度不局限于心律失常还包括心肌梗死、心肌缺血、心包疾病以及药物作用等，这种复杂性导致了能够确定由心律失常而产生 ST 段改变的样本在临床中较少。基于混淆矩阵对 ST 类型的错误判断能够表明 DSE-ResNet 模型对亚异常类型 ST 的识别不敏感，这可能是由于 ST 的训练样本数量稀少导致 DSE-ResNet 能够提取到的特征有限。此外，专业医生对 ST 的诊断意见不一<sup>[62]</sup>，导致样本标注错误也是原因之一。对于亚异常类型 AF 和 Block，所提出的模型分别获得了 0.944 和 0.913 的  $F_1$  分数。

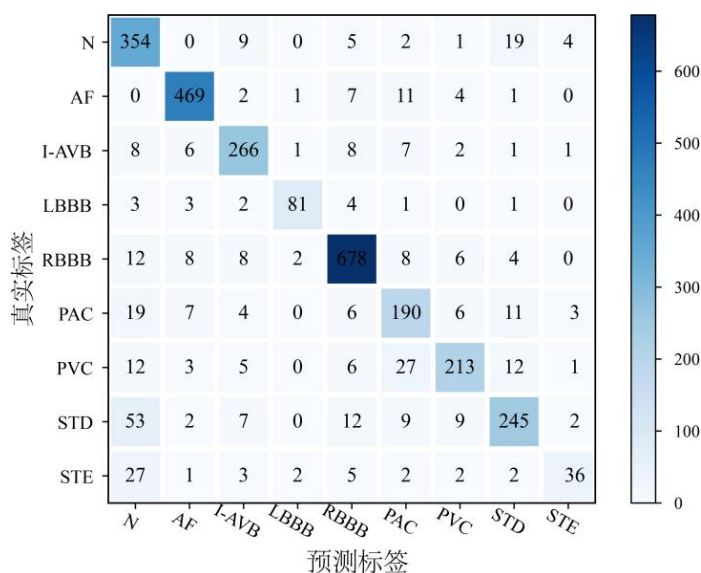


图 3.9 混淆矩阵

根据混淆矩阵计算得到了 DSE-ResNet 在隐藏测试集上的具体分类性能。表 3.4 显示了在 CPSC2018 隐藏数据库中模型对于不同心律失常的准确度、精确率、灵敏度和特异度分数。正常心律和 8 类心律失常的平均准确度和平均特异度分别为 0.965 和 0.979，且均在 LBBB 上达到最大值。值得注意的是 LBBB 在训练集中是所有分类中样本数量最少的类型（仅有 203 组），而 DSE-ResNet 对 LBBB 识别的误诊率非常低。通过寻找医学方向的依据表明，这可能是由于 LBBB 的识别在临床判断中具有多种关键特征，例如 QRS 波的时限异常（男性  $\geq 140\text{ms}$ ，女性  $\geq 130\text{ms}$ ）、QRS 波形态异常（导联 V1 的 QRS 波呈 QS 形）以及在导联 I、aVL、V1、V2、V5 和 V6 中至少有两个或两个以上导联存在 QRS 波的切迹或顿挫<sup>[66]</sup>。

多重特征的综合判定表现在二维心电图图中使得 DSE-ResNet 能够学习足够多的特性用以识别 LBBB 类型。

表 3.4 DSE-ResNet 在 CPSC2018 隐藏测试集的相关性能指标

名称	平均	正常及 8 类心律失常								亚异常类型			
		正常	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE	Block	PC	ST
$Acc_i$	0.965	0.936	0.978	0.972	<b>0.992</b>	0.962	0.954	0.963	0.946	0.979	0.946	0.943	0.930
$F_P$	0.845	0.725	<b>0.94</b>	0.869	0.931	0.927	0.739	0.877	0.828	0.766	0.912	0.806	0.819
$Se_i$	0.803	0.898	<b>0.947</b>	0.887	0.853	0.934	0.772	0.763	0.723	0.45	0.914	0.768	0.671
$Sp_i$	0.979	0.942	0.986	0.983	<b>0.998</b>	0.972	0.972	0.987	0.978	0.996	0.952	0.970	0.975

表 3.5 展示了 DSE-ResNet 和 CPSC2018 竞赛排行榜中排名前五的队伍模型的平均 $F_1$ 分数（针对正常节律和 8 类异常类型的平均）以及四种亚异常类型的 $F_1$ 分数。该对比基于相同的隐藏测试集，测试结果表明 DSE-ResNet 平均 $F_1=0.817$ ，仅落后最先进模型 0.02。值得注意的是，DSE-ResNet 在两个亚异常类型中取得了最好的测试结果，分别为 $F_{AF}=0.944$  和 $F_{Block}=0.913$ 。同时，基于隐藏测试集的测试结果说明，该模型从二维心电图图中学习内部和导联间特征，对 AF 和 Block 识别能力更敏感。

表 3.5 DSE-ResNet 与 CPSC2018 排名前五的模型对比结果

模型	平均 $F_1$	$F_{AF}$	$F_{Block}$	$F_{PC}$	$F_{ST}$
DSE-ResNet	0.817	<b>0.944</b>	<b>0.913</b>	0.786	0.738
Chen 等人 <sup>[24,63]</sup>	<b>0.837</b>	0.933	0.899	<b>0.847</b>	<b>0.779</b>
Cai 等人 <sup>[63]</sup>	0.830	0.931	0.912	0.817	0.761
He 等人 <sup>[63]</sup>	0.806	0.914	0.879	0.801	0.742
Yu 等人 <sup>[63]</sup>	0.802	0.918	0.890	0.789	0.718
Yan 等人 <sup>[63]</sup>	0.791	0.924	0.882	0.779	0.709

表 3.6 比较了 DSE-ResNet 和近些年基于 CPSC2018 隐藏测试集的研究模型的性能。表中的结果表明，本文提出的模型在 AF、I-AVB、LBBB 和 PAC 中的 $F_1$ 得分分别为 0.944、0.878、0.89 和 0.755，在平均 $F_1$ 的得分也是最高的。与其他

方法相比, 本文使用的同时学习导联内部和导联间特征有助于识别多种类型的心律失常。综上所述 DSE-ResNet 在检测心律失常某些具体类型方面具有一定的优势。

表 3.6 同其他相关研究进行对比

相关研究	$F_1$ 分数									
	正常	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE	平均 $F_1$
DSE-ResNet										
CNN+通道注意力+集成模型	0.803	<b>0.944</b>	<b>0.878</b>	<b>0.89</b>	0.931	<b>0.755</b>	0.816	0.72	0.567	<b>0.817</b>
Yao 等人 <sup>[61]</sup>	0.753	0.9	0.809	0.874	0.922	0.638	0.832	0.762	0.462	0.772
CNN+LSTM										
Yao 等人 <sup>[61]</sup>	0.789	0.92	0.85	0.872	<b>0.933</b>	0.736	<b>0.861</b>	0.789	0.556	0.812
CNN+注意力机制										
Liu 等人 <sup>[64]</sup>	<b>0.82</b>	0.91	0.87	0.87	0.91	0.63	0.82	<b>0.81</b>	<b>0.6</b>	0.81
CNN+专家特征										
Liu 等人 <sup>[64]</sup>	0.8	0.89	0.87	0.77	0.9	0.65	0.79	0.8	0.56	0.78
CNN										
Wang 等人 <sup>[65]</sup>	0.79	0.93	0.85	0.86	0.93	0.75	0.85	0.8	0.56	0.813
CNN+注意力机制										

### 3.5 本章小结及讨论

本章节提出了一种基于二维化十二导联 ECG 的心律失常自动分类模型 DSE-ResNet, 能够实现正常节律和 8 类心律失常的自动分类。二维处理方法将原始十二导联 ECG 合并到一个二维空间中, 借助 CNN 局部连接的特性使得 DSE-ResNet 可以同时提取十二导联 ECG 的内部和导联间特征。采取正交试验代替网格搜索来选择超参数降低参数组合的数量, 经过筛选的超参数组合被训练为最优单一模型作为集成模型的组成部分。基于小样本测试集的实验表明, 基于投票策略的集成模型的性能远优于单一模型。与 CPSC2018 中前 5 名模型的结果对比表明, DSE-ResNet 模型取得的平均 $F_1=0.817$ , 并且在两个亚异常类型上取得了最优的测试结果。

这表明基于深度学习模型对二维化十二导联 ECG 进行特征学习在部分心律失常类型的自动识别占据优势。尤其是 AF 和 Block, 它们的自动分类可能更取决于导联之间的关系。这也意味着使用 DSE-ResNet 处理多导联 ECG 信号以捕获内

部导联和导联间特征对于心律失常的自动识别是有效的。这种结果不仅为基于十二导联 ECG 心律失常的自动识别提供了新的视角，而且提出了一些研究方向，例如基于二维心电图探索如何进一步提高预测的准确率、分析十二导联 ECG 之间的相关关系、探索深度学习过程导联信号的冗余性等问题。



## 第4章 基于二维化十二导联 ECG 信号冗余性分析

### 4.1 引言

第三章的研究中提及 PAC 患者的 ECG 异常主要表现在导联 V1、II 和 aVF，而 LBBB 患者典型的心电图异常主要表现在导联 I、V1、V2、V5、V6 和 aVR<sup>[47]</sup>。这可能表示十二导联 ECG 中不同导联对于不同或者相同心律失常的识别贡献是不同的，即有的导联信号对特定类型心律失常的诊断可能是冗余的。

随着各种带有心电监测功能的便携式设备的出现例如 Apple Watch 手表、心电带、动态心电监护仪<sup>[67]</sup>和可穿戴式的肢体六导联心电测量设备<sup>[68]</sup>，使得用户实时监测和记录心电状态成为可能。然而这种设备在心律异常预警上仍存在局限性，相较于医院测量十二导联 ECG 使用的十电极系统，便携式设备受制于本身具有方便携带的性质，不能够安装过多的电极导致只能测量部分导联信号。这启发本研究通过擦除十二导联 ECG 的不同导联组合来研究导联信号冗余，以保证在便携式设备缺失一部分导联信息的情况下，仍能够给出较高准确度的心律失常诊断。

近些年的部分研究涉及到使用少数几个导联完成心律失常的自动识别，并进行了对比。例如 Chen 等人<sup>[24]</sup>开发了一种 CNN 模型来识别心律失常。仿真结果表明，仅使用单导联心电图进行自动心律失常识别的性能低于使用完整的十二导联心电图。其中以导联 aVR 和 V1 最为突出。Chiou 等人<sup>[26]</sup>通过使用连续小波变换将每个导联转换为二维频谱，并设计了一个二维 CNN 模型来识别心力衰竭。结果表明，导联 V6 在单导联模式下实现了最高的准确率、灵敏度、特异度和  $F_1$  分数，而导联 V5 和 V6 的组合在多导联模式下表现最佳。Cho 等人<sup>[69]</sup>开发了一种变分自编码器，发现基于肢体六导联利用变分自编码器诊断心肌梗死是具备可行性的。Chen 等人<sup>[45]</sup>提出了多分支卷积和残差网络来识别心律失常。通过比较单导联和八导联融合的识别性能，他们发现八导联融合可以提高识别性能。此外，PhysioNet 还发起了一项名为“Will Two Do? Varying Dimensions in Electrocardiography: The PhysioNet/Computing in Cardiology Challenge 2021 (PCC2021)”的基于导联信号冗余探索的竞赛<sup>[37,38]</sup>，该竞赛要求参与者构建一种算法，该算法可以分别使用十二导联、肢体六导联（I、II、III、aVR、aV 和 aVF）、四导联（I、II、III 和 V2）、三导联（I、II 和 V2）和二导联（I 和 II）诊断心律失常。

本次竞赛的一些研究具有指导意义。Philip 等人<sup>[70]</sup>基于散射循环网络（Scattering-Recurrent Networks）利用组合导联实现心律失常识别。他们发现模型

的性能随着导联数量的减少而略有下降，这表明导联之间存在很强的相关性。Matteo 等人<sup>[71]</sup>使用自动化机器学习 (AutoML) 来对不同导联组合的 ECG 信号进行分类，多个模型实例的累积结果表明使用肢体六导联时性能最佳。Niels 等人<sup>[72]</sup>使用卷积循环神经网络来识别不同导联组合的心电信息异常，发现使用二导联可以获得更好的分类性能。比赛结果表明，不同导联组合在隐藏测试集中的表现差异较大。

先前的研究已经训练得到了 DSE-ResNet，在本章中对该模型进行了延伸和扩充，得到多种不同组成结构的深度神经网络模型。这方便探究在不同 DNN 模型中，缺失不同导联组合的信息在心律失常自动识别上的分类趋势是否一致。为了保持模型的统一性，本研究控制传递到 DNN 的十二导联 ECG 信号在维度上是统一的，并不是像比赛 PCC2021 中完全去除导联，而是通过用 0 填充来代替要擦除的导联信息。

相较于第三章的算法，本章节在初期对数据的处理上做了如下改进：

(1) 抛弃去噪算法。使 DNN 基于原始信号完成端到端的分类流程，这更加符合现实情景。

(2) 降低采样率，缩短样本长度，缩短训练时间。需要训练的模型为 30 个，需要大量的训练时间，降低采样率是节省训练时间的一种常见做法。

(3) 收集来自世界各地的多个数据库形成融合数据库，增强模型泛化性。

(4) 改进二维心电图的处理方式，保证样本在训练数据集的唯一性。相较于第三章扩充通道维度实现逻辑上的二维化，本章节提出的改进后的二维心电图在通道维度包含同一原始十二导联 ECG 不同时间段的信号片段。

## 4.2 研究方法

### 4.2.1 十二导联 ECG 计算方法

十二导联心电图包括六个肢体导联和六个胸前导联，分别记录心脏在额状面和横断面的电活动。六个肢体导联包括三个双极肢体导联 I、II、III 和三个单极加压肢体导联 aVR、aVL 和 aVF。六个胸前导联包括导联 V1、V2、V3、V4、V5 和 V6。十二导联 ECG 可以通过十电极系统测量得到。图 4.1 (a) 显示了十个电极在体表的摆放位置。标准十二导联 ECG 的数值计算和变形公式如图 4.1 (b) 所示。图 4.1 (c) 展示了不同导联测量心脏电位的方向。

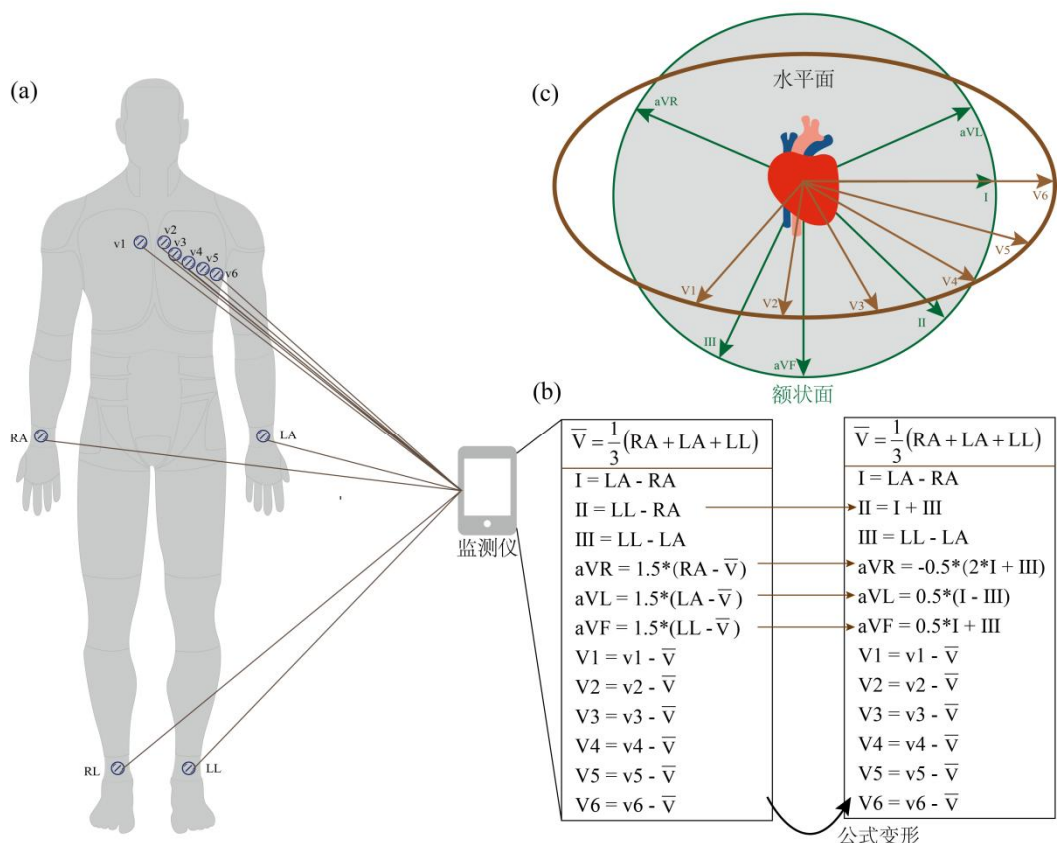


图 4.1 十电极系统示意图

注：（a）十电极摆放位置示意图；（b）十二导联计算公式及公式变形；（c）十二导联测量心脏电位方向示意图

## 4.2.2 融合数据集

本章节所使用的十二导联 ECG 样本来自 PCC2021 提供的不同国家地区的四个开源大型数据库。所有样本信号的采样率为 500Hz。四个数据库分别是 CPSC2018、CPSC-Extra、PTB 和 G12EC。数据库的详细信息见表 4.1。

表 4.1 融合数据库详细信息

数据库	总计	女性样本数量	男性样本数量	信号长度	采样率
CPSC2018	6877	3699	3178	6s-144s	500Hz
CPSC-Extra	3453	1843	1610	8s-98s	500Hz
PTB	21837	11379	10458	10s	500Hz
G12EC	10344	5551	4793	10s	500Hz

这些数据库中包含的心律失常类型并不完全相同。参照 CPSC2018 标记的 9 种类别，从其他三个数据库中提取了包含相同 9 种类别的样本。年龄异常（即没有给定或给定年龄异常）和原始信号长度小于 8s 的样本被剔除。正常节律和 8

类心律失常的有效样本见表 4.2。融合四个数据库经过过滤后共获得了 19883 组样本。每组样本都包含原始十二导联 ECG 信号，以及相应的年龄、性别和标签。

表 4.2 融合数据库后不同心律失常类型数量

标签	类型	CPSC2018	CPSC-Extra	PTB	G12EC	总计
1	正常	898	—	1732	7871	10501
2	AF	1219	40	234	525	2018
3	IABV	721	96	338	148	1303
4	LBBB	199	9	199	365	772
5	RBBB	1671	—	388	—	2059
6	PAC	542	102	265	132	1041
7	PVC	626	—	27	469	1122
8	STD	784	29	3	—	816
9	STE	185	35	10	21	251
	总计	6845	311	3196	9531	19883

#### 4.2.3 导联组合方式

根据图 4.1 (b) 中十二导联的计算公式以及变形公式，可以发现导联 II、aVR、aVL 和 aVF 的值可以从导联 I 和 III 推导出，相反导联 I 和 III 也可被推导出。从特征选择的角度考虑，这些导联之间的电压标量值存在线性逻辑关系，即特征相关度高。因此这些导联在心律失常的自动识别中可能存在冗余。为了探究这个问题，在本节中建立了六种不同导联擦除组合以探索在不同 DNN 模型下丢失不同导联组合信息对自动识别的影响。

DNN 的参数（如过滤器的数量、卷积核大小）针对不同的输入维度需要进行调整，从而导致输入数据的维度和模型参数都会影响实验结果。在这种情况下，比较缺少不同导联对自动心律失常识别的影响是不合理的。不同于 PCC2021 直接删除导联，本研究使用填充 0 的方式代替需要擦除的导联信息。这是为了保证输入 DNN 数据的维度一致，而选择 0 是由于任一导联在采样时间内的平均值约为 0。

研究过程中使用 CPSC2018 赛后开源的七个 DNN 模型基于同一测试集进行实验，发现在逐个擦除单个导联信号时，擦除导联 aVL 对心律失常识别的影响最小（见 4.4.1）。这促使本研究选择导联 aVL 进行冗余性探究。

此外，还研究了双极肢体导联、单极加压肢体导联和胸前导联在信号被擦除后对心律失常识别的影响。擦除方式根据被擦除的导联数量依次排序为：

- (1) case-1: 没有导联信号被擦除，即完整的十二导联 ECG 信号。
- (2) case-2: 导联 aVL 的所有采样点的信号值被填充为 0。

(3) **case-3**: 双极肢体导联被擦除。即导联I、II和III的所有采样点的信号值被填充为0。

(4) **case-4**: 单极加压肢体导联被擦除。即导联 aVR、aVL 和 aVF 的所有采样点的信号值被填充为0。

(5) **case-5**: 导联II、aVR、aVL 和 aVF 的所有采样点的信号值被填充为0。

(6) **case-6**: 胸前导联被擦除。即导联 V1、V2、V3、V4、V5 和 V6 的所有采样点的信号值被填充为0。图 4.2 展示了不同擦除导联组合方式的原理图。

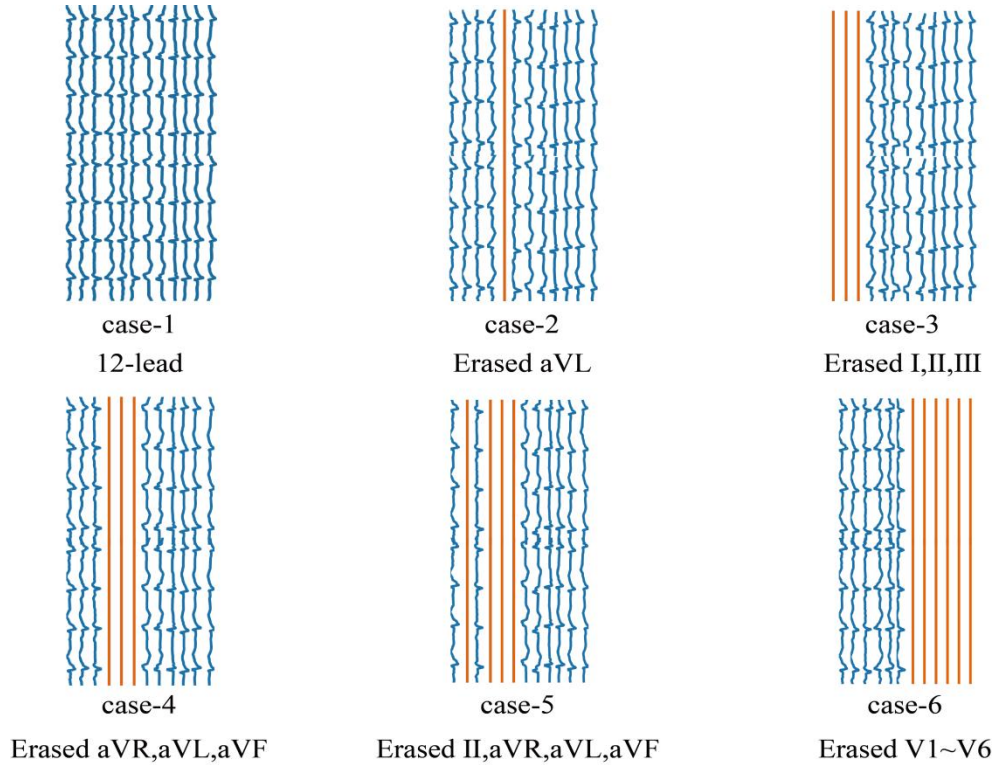


图 4.2 不同擦除导联组合方式原理图

注：导联的排列方式从左至右依次为导联I、II、III、aVR、aVL、aVF、V1、V2、V3、V4、V5 和 V6

#### 4.2.4 改进二维心电图

融合数据库样本的原始十二导联 ECG 信号长度不同。为了保证输入 DNN 的十二导联 ECG 具有相同的维度，需要进行数据预处理。常用方法包括对短时长的样本填充 0 到指定长度，以及截断过长的样本以确保所有样本具有相同的长度。第三章的研究采取了这种方式，但是十二导联 ECG 不同时间段的波形结构不完全相同，填充可能会破坏原始十二导联 ECG 的结构，而裁剪十二导联 ECG 可能会抹除用于确定心律失常类型的关键波形。因此，通过填充或截断来保持样本长度一致是不合理的。

第三章中提及的切片操作，原始样本的十二导联 ECG 经过切片后会变成训练集中的多组样本，使得训练数据中存在同一患者的多个二维心电图样本。过多的

重复同一患者的样本可能导致模型过拟合。且经过维度扩充后二维心电图的维度变为  $A_i \in \mathbb{R}^{8192 \times 12 \times 1}$ ，扩充通道维度为 1 只是为了适应二维 CNN 的训练要求，在本章节中提出了一种改进的二维心电图，称为“Block”。能够保证训练数据中患者样本的唯一性，同时在通道维度添加导联信息实现类似于图片具有 RGB 三通道数据的构造。

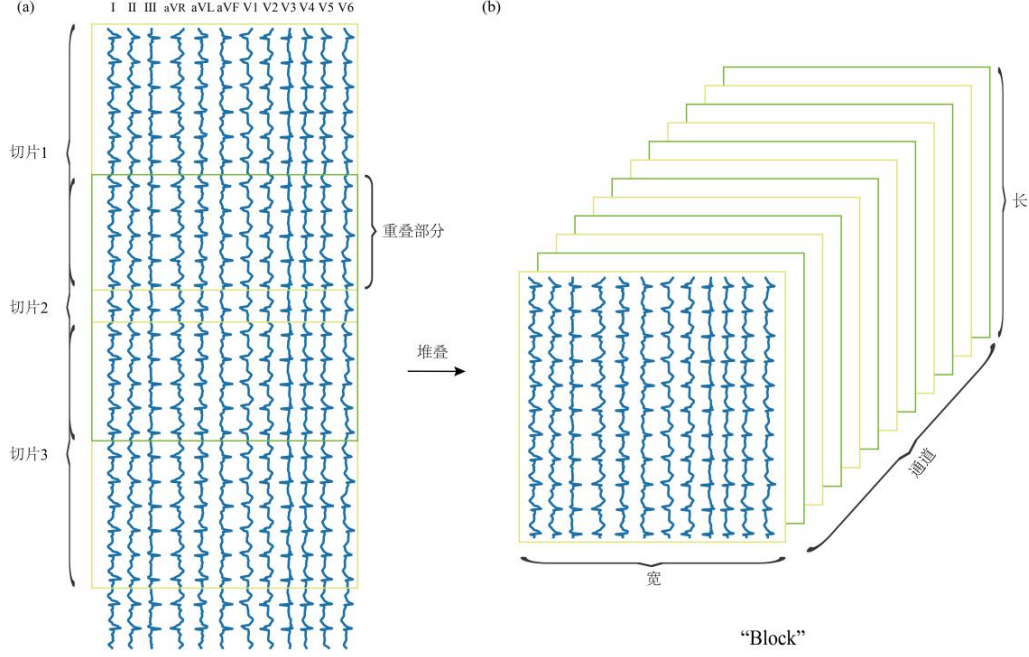


图 4.3 改进的二维化原理示意图

注：（a）切片原则示意图；（b）“Block”示意图

如图 4.3 所示，按照时序顺序十二导联 ECG 被切分并堆叠成一个“Block”。每个 Block 有三个维度：分别表示切片的长度、导联的个数和切片的个数。每个切片包含同一样本不同时间段的十二导联 ECG，经过切片堆叠可以最大程度地保留样本蕴含的信息。每组样本的原始十二导联 ECG 被切分为十二片，切片的长度为 2048（4.096s）。重叠部分的长度由原始信号长度  $L$  决定，当原始信号长度  $L \in (2048, 12 \times 2048]$  时，重叠部分长度为：

$$Overlap = \lfloor (12 \times 2048 - L) / 11 \rfloor \quad (4.1)$$

当原始样本长度  $L > 12 \times 2048$ ，重叠部分长度为 0，截取该样本前  $12 \times 2048$  的长度作为有效信号部分，其余舍弃。经过数据处理共得到 19883 组二维心电图  $A \in \mathbb{R}^{2048 \times 12 \times 12}$ 。不同擦除方式的二维心电图结构示意图如图 4.4 所示。



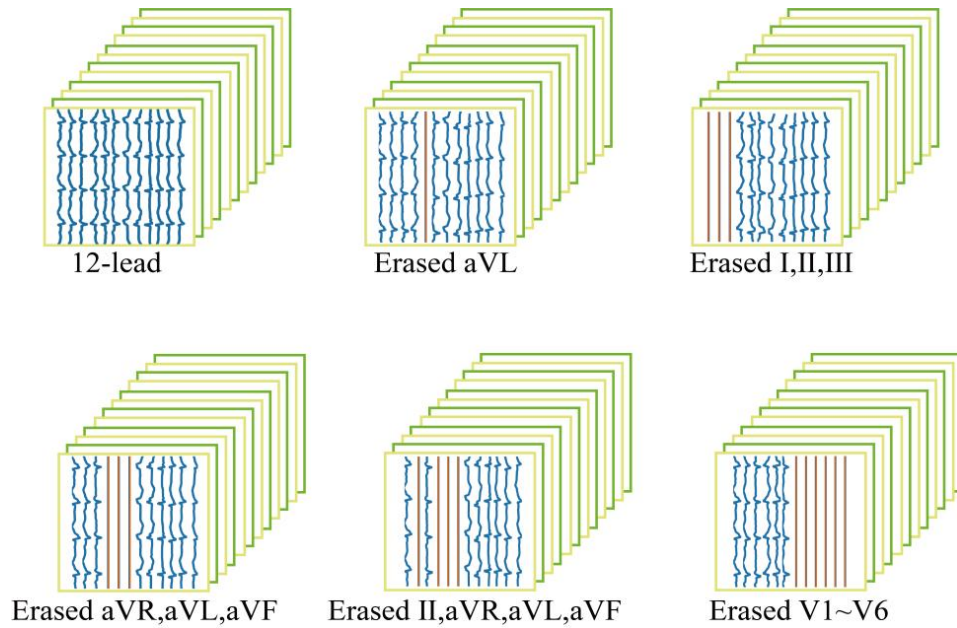


图 4.4 不同擦除方式的二维心电图结构示意图

#### 4.2.5 模型构建

随着算力的提高，能够实现更深层次神经网络的学习。在一定范围内神经网络的深度越深，提取的特征越抽象，特征表达能力越强。在各种神经网络结构中，CNN 由于其具有局部连接和权重共享的特性，可以有效地提取局部特征。RNN 常用于处理一维序列数据<sup>[73]</sup>，其变体包括长短期记忆网络 LSTM<sup>[74]</sup>、双向长短期记忆网络 BiLSTM<sup>[75]</sup>等解决了 RNN 长期依赖的问题。此外，ResNet<sup>[76]</sup>因为克服了深度学习过程中的梯度消失和爆炸问题而成为深度学习领域最常用的框架。

本章节提出的模型以第三章提出的 DSE-ResNet 模型为基线，延伸和扩展了五种不同结构的 DNN 模型来实现自动心律失常识别。构建的五个 DNN 模型的结构如图 4.5 所示，具体如下：

（1）model-1：该模型为 DSE-ResNet，其中 ResNet 用于提取导联内部和导联间的特征，DSE 用于提取十二导联 ECG 不同时间片段的全局特征，年龄和性别作为辅助特征提高分类性能。

（2）model-2：该模型取自 DSE-ResNet 中的部分结构，包括 ResNet 结构和患者的年龄和性别信息。

（3）model-3：该模型由 ResNet 和 LSTM 组成。LSTM 是 RNN 的一种变体，通过引入门机制（输入门、遗忘门、输出门）和细胞状态，解决了 RNN 无法处理长期依赖的问题，其结构如图 2.11 所示。LSTM 通过各种门函数保留重要特征，以确保它们在长期传播过程中不丢失。由于其独特的特性，LSTM 适用于对时间序列数据（例如十二导联 ECG）进行建模。

(4) model-4: 该模型由 ResNet 和 BiLSTM 组成。BiLSTM 是基于 LSTM 的变体, 由前向 LSTM 和后向 LSTM 组成, 其结构如图 2.12 所示。其优点是可以同时捕获时序数据的前向信息和后向信息。

(5) model-5: 该模型为 SE-ResNet。SE 在通道维度上实现了注意力机制<sup>[50]</sup>。其结构如图 3.4 所示。在本文中, 通道维度所存储的数据表示十二导联不同时间段的采样信号。SE 可以通过增加重要切片的权重和弱化次要切片的权重来提高模型的注意力。

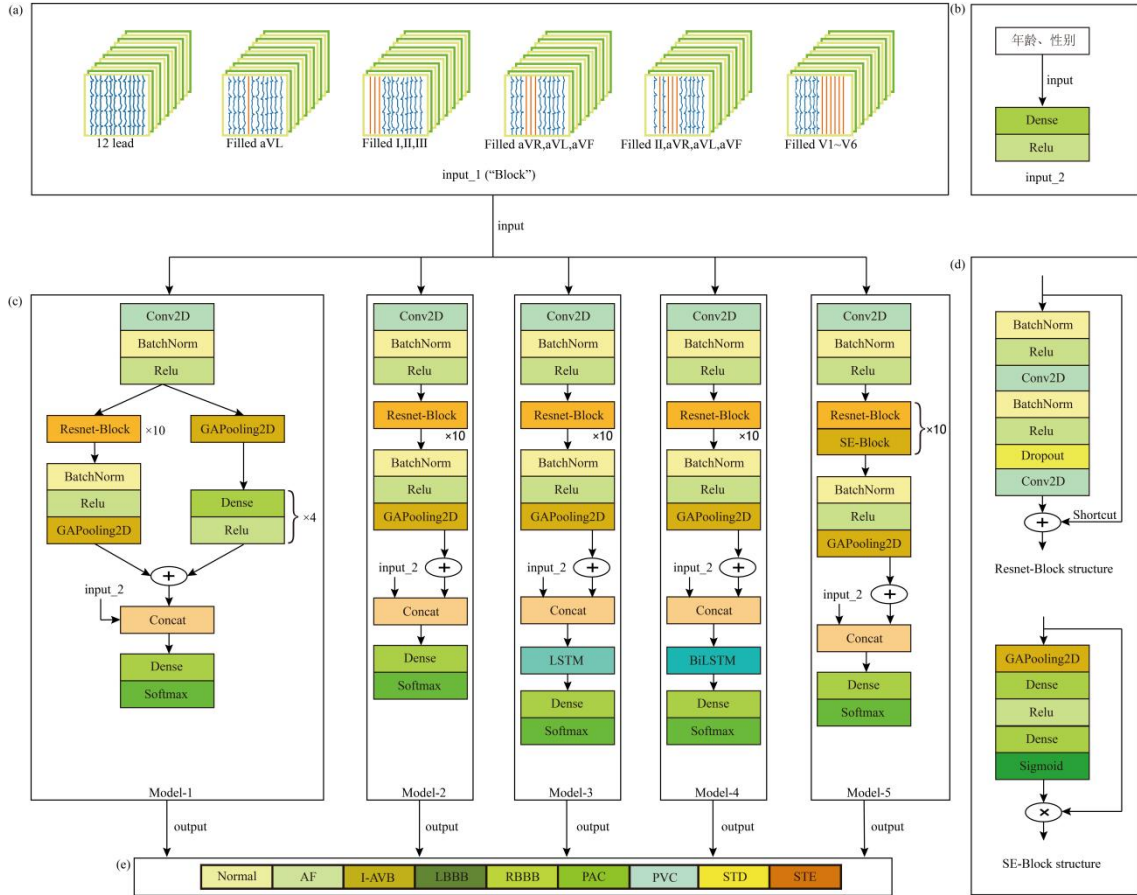


图 4.5 DNN 模型结构

注: (a) 包含 5 种不同的擦除方法的二维心电图 A; (b) DNN 的第二部分输入源, 包含患者的年龄和性别; (c) 不同结构的 DNN 模型结构; (d) ResNet-Block 和 SE-Block 的组成结构; (e) 心律失常分类类别

#### 4.2.6 训练和测试

融合数据库样本被划分为三组: 训练集、验证集和测试集。所有样本随机打乱并按照标签类别分配 77% 为训练集, 8% 为验证集, 15% 为测试集。验证集用于监测模型训练过程中的过拟合现象, 验证集和测试集中的样本不参与训练过程。

DNN 模型的输入数据来自两部分如图 4.5 (a) 和图 4.5 (b) 所示, 分别代表二维心电图 “Block” 和辅助特征。辅助特征包括性别和年龄, 其中男性和女性进



行二值转换分别表示为 1 和 0，年龄为整数。样本的标签使用 0 到 8 代替，其中 0 代表正常心律，1 到 8 代表 8 类不同类型的心律失常。

本章节基于具有六种（包括完整十二导联）不同擦除组合的十二导联 ECG 训练每个 DNN 模型。训练后得到 30（即  $5 \times 6 = 30$ ）个模型，并通过测试集对其进行性能评价。训练过程中的相关超参数包括带动量的随机梯度下降（SGD），其中动量为 0.2，每个模型训练 50 个 epoch，dropout 为 0.4，批处理大小为 32，学习率在前 30 个 epoch 中为 0.1，在后 20 个 epoch 中为 0.01。LSTM 和 BiLSTM 的单元数为 64，损失函数使用 Softmax 交叉熵损失函数。五个 DNN 模型的卷积层参数如表 4.3 所示。

表 4.3 DNN 模型的卷积层参数

Conv2D 层	卷积核大小	步长	滤波器数量	输出尺寸
Conv2D-1	12×12	1×1	12	2048×12×12
Conv2D-2	32×1	1×1	12	2048×12×12
Conv2D-3	32×1	1×1	12	2048×12×12
Conv2D-4	32×1	2×1	12	1024×12×12
Conv2D-5	32×1	1×1	12	1024×12×12
Conv2D-6	32×1	1×1	24	1024×12×24
Conv2D-7	32×1	1×1	24	1024×12×24
Conv2D-8	32×1	2×1	24	512×12×24
Conv2D-9	32×1	2×1	24	256×12×24
Conv2D-10	32×1	1×1	48	256×12×48
Conv2D-11	32×1	1×1	48	256×12×48
Conv2D-12	32×1	2×1	48	128×12×48
Conv2D-13	32×1	1×1	48	128×12×48
Conv2D-14	32×1	1×1	96	128×12×96
Conv2D-15	32×1	1×1	96	128×12×96
Conv2D-16	32×1	2×1	96	64×12×96
Conv2D-17	32×1	1×1	96	64×12×96
Conv2D-18	32×1	1×1	192	64×12×192
Conv2D-19	32×1	1×1	192	64×12×192
Conv2D-20	32×1	2×1	192	32×12×192
Conv2D-21	32×1	1×1	192	32×12×192

### 4.3 结果

#### 4.3.1 评价标准

由于存在多个模型，每个模型有多个分类结果，所以采用精确率 *Precision*、召回率 *Recall*、宏平均  $F_1$  分数、接受者操作特征曲线（Receiver Operating Characteristic Curve, ROC）和 ROC 曲线下面积（Area Under Curve, AUC）对模型的性能进行综合评价<sup>[61]</sup>。第  $i$  类心律失常精确率的计算方法如式（4.2）：

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \quad (4.2)$$

对于同一 DNN 模型不同导联擦除方式在所有分类类别上的平均 *Precision* 计算方法如下：

$$Precision = \frac{1}{9} \sum_{i=1}^9 Precision_i \quad (4.3)$$

同理，第  $i$  类心律失常召回率以及不同导联擦除方式分类平均 *Recall* 如下：

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \quad (4.4)$$

$$Recall = \frac{1}{9} \sum_{i=1}^9 Recall_i \quad (4.5)$$

基于 *Precision* 和 *Recall* 可以计算一个 DNN 模型的宏平均  $F_1$  分数，如下所示：

$$F_1 = \frac{1}{9} \sum_{i=1}^9 \frac{2(Precision_i \times Recall_i)}{Precision_i + Recall_i} \quad (4.6)$$

ROC 分析和 AUC 指标在医疗决策系统中已经得到广泛应用<sup>[77]</sup>。对于样本数量不平衡的多分类问题，宏平均（macro-average）ROC 通过绘制多类敏感度和特异度之间的平均定量关系来衡量模型的性能。AUC 是 ROC 曲线下的面积。当 ROC 曲线越接近左上角，AUC 越接近 1 时，DNN 模型分类性能越好。

#### 4.3.2 结果分析

表 4.4 显示了不同擦除组合在不同 DNN 模型的 *Precision*、*Recall* 和宏平均  $F_1$  分数的详细信息。根据表 4.4 发现每个 DNN 模型的 *Recall* 和宏平均  $F_1$  的最佳结果通常出现在 case-1（没有导联信息被擦除）中。这表明当完整十二导联 ECG 用于 DNN 模型的训练时，通常会提高模型的识别性能。值得注意的是，部分 DNN 模型的最佳性能出现在其它擦除组合中，例如 case-2（擦除导联 aVL）在 model-1（model-3）取得最优的  $Precision=0.797$  和宏平均  $F_1=0.758$ （ $Precision=0.795$  和宏平均  $F_1=0.756$ ），case-3（擦除导联 I、II 和 III）和 case-5（擦除导联 II、aVR、aVL 和 aVF）在 model-2 中达到最高的  $Recall=0.759$ ，case-3 在 model-4 中取得最优

$Precision=0.788$ 。

表 4.4 不同擦除组合在不同 DNN 模型的相关评价分数

<i>Precision</i> 表	case-1	case-2	case-3	case-4	case-5	case-6
model-1 (DSE-ResNet)	0.77	<b>0.797</b>	0.794	0.78	0.761	0.735
model-2 (ResNet)	<b>0.816</b>	0.781	0.772	0.747	0.767	0.78
model-3 (ResNet+LSTM)	0.755	<b>0.795</b>	0.771	0.677	0.771	0.708
model-4 (ResNet+BiLSTM)	0.774	0.731	<b>0.788</b>	0.769	0.774	0.741
model-5 (SE-ResNet)	<b>0.802</b>	0.692	0.801	0.758	0.798	0.739
平均值	0.783	0.759	<b>0.785</b>	0.746	0.774	0.741
<i>Recall</i> 表	case-1	case-2	case-3	case-4	case-5	case-6
model-1 (DSE-ResNet)	<b>0.746</b>	0.732	0.71	0.7	0.717	0.719
model-2 (ResNet)	0.737	0.73	<b>0.759</b>	0.707	<b>0.759</b>	0.674
model-3 (ResNet+LSTM)	<b>0.737</b>	0.73	0.728	0.636	0.7	0.712
model-4 (ResNet+BiLSTM)	<b>0.753</b>	0.694	0.707	0.675	0.701	0.734
model-5 (SE-ResNet)	<b>0.76</b>	0.689	0.725	0.715	0.744	0.743
平均值	<b>0.747</b>	0.715	0.726	0.687	0.724	0.716
宏平均 $F_1$ 表	case-1	case-2	case-3	case-4	case-5	case-6
model-1 (DSE-ResNet)	0.752	<b>0.758</b>	0.744	0.727	0.732	0.717
model-2 (ResNet)	<b>0.768</b>	0.75	0.763	0.719	0.761	0.707
model-3 (ResNet+LSTM)	0.743	<b>0.756</b>	0.739	0.654	0.727	0.707
model-4 (ResNet+BiLSTM)	<b>0.761</b>	0.71	0.739	0.714	0.731	0.734
model-5 (SE-ResNet)	<b>0.779</b>	0.683	0.74	0.734	0.766	0.738
平均值	<b>0.761</b>	0.731	0.745	0.71	0.743	0.721

这表明缺少某些导联信息可能会提高 DNN 模型的 *Precision*、*Recall* 或宏平均 $F_1$ 分数。不同模型对正常和 8 类类型心律失常的 $F_1$ 分数如表 4.5 所示。显然擦除不同导联组合训练后的不同 DNN 模型对于具有足够样本的类型都显示出优异的性能。然而对于样本数量过少的 STE 类型所有模型都表现出较差的性能。出乎意料的是 PAC 类型在所有擦除组合中也都表现出较差的性能，结合第三章的实验对于 PAC 的分类性能表现良好推断可能是改进后的二维心电图使得深度学习模型对这种类型的特征学习造成阻碍。

不同擦除组合在所有模型的平均 *Precision*、*Recall* 和 $F_1$ 分数如图 4.6 (a) 所

示。统计结果表明,在 case-1 中得到最优平均  $F_1=0.761$  和  $Recall=0.747$ 。与预先想象的结果相反,发现在 case-3 中获得了最佳平均值  $Precision=0.785$ 。同时 case-3 的模型平均  $Recall$  和  $F_1$  仅次于 case-1 (没有导联信息被擦除)。case-5 相较于 case-1 模型平均  $F_1$  性能降低了 0.018, 与 case-3 相比仅降低 0.002。

图 4.6 (b) - (f) 给出了每个 DNN 模型对于不同擦除组合的 ROC 曲线。对比发现 case-3 在 model-3 中取得最高 AUC, 在多数模型中达到了次高的 AUC。case-5 的性能略低于 case-3。

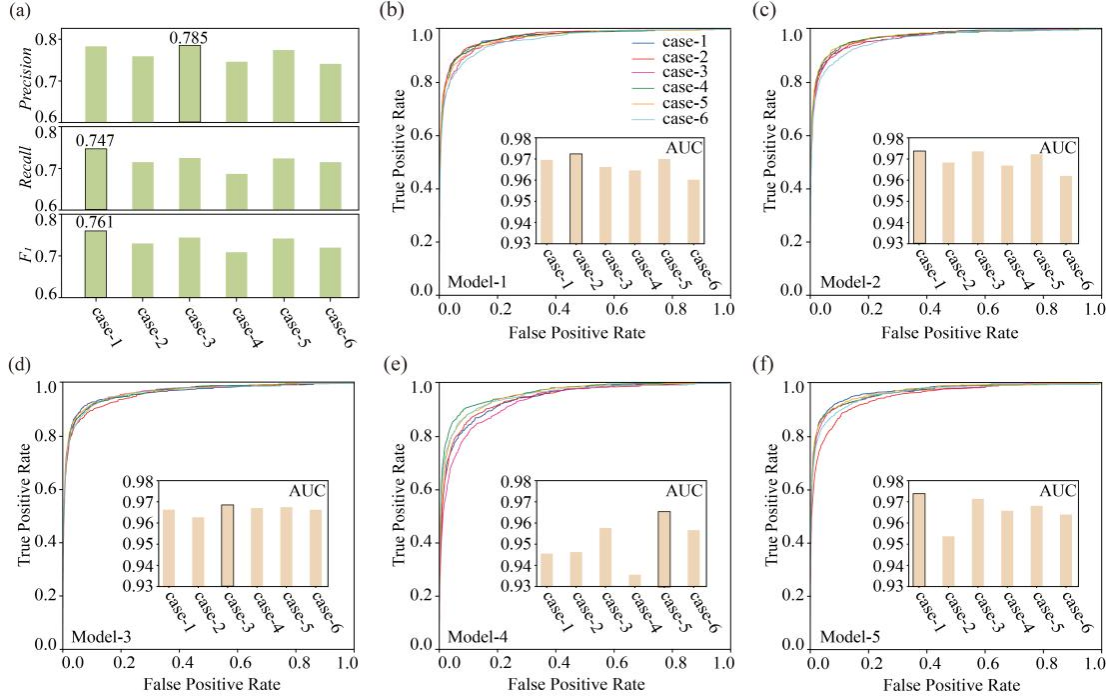


图 4.6 不同 DNN 模型应用不同导联擦除组合的分类性能

注: (a) 模型平均 Precision、Recall 和  $F_1$  分数图; (b) - (f) 不同擦除组合在不同 DNN 模型的 ROC 曲线和 AUC

综合表明,导联 I、II、III 与导联 aVR、aVL、aVF 存在的逻辑计算关系使得在深度学习过程中存在导联信息冗余。这种冗余仅略微影响 DNN 模型的性能。和最初构想的相同,存在逻辑换算关系的导联相关性较强,导致在深度学习过程中相关性较强的导联信号可能变为冗余特征。这为测量不同 ECG 信号的便携式设备应用部分导联信息诊断心律失常类型成为可能。

## 4.4 补充信息

### 4.4.1 选择导联 aVL 的过程

第四章主要实验开始前,根据现有的实验环境,对 CPSC2018 挑战赛的前二十五名的模型调试成功七个 DNN 模型<sup>[63]</sup>,包括 CPSC0223、CPSC0183、CPSC0235、

CPSC0228、CPSC0204、CPSC0212 和 CPSC0166。将 CPSC2018 数据库的样本复制成十三份，一份是原始数据，另外十二份逐导联擦除信息（用 0 代替）。由于 CPSC2018 挑战赛只公开了这些队伍的测试代码，不包含训练代码。因此使用了这些团队在完整的十二导联数据上训练的模型来测试逐导联擦除的另外十二份数据。基于十三份数据测试的七个 DNN 模型的平均 $F_1$ 如图 4.7 所示。仿真结果表明当 aVL 导联被擦除时，平均 $F_1$ 分数略低于使用完整十二导联信号，这可能表示导联 aVL 在自动心律失常识别中可能是冗余的。而当导联 V1 被擦除时，对分类结果的影响最严重，这间接表明导联 V1 对于心律失常识别的重要性。

除此之外，图 4.7 还展示了擦除单个导联在这些开源模型上的平均 $F_1$ 分数的数据分布，可以发现排名较高的模型例如 CPSC0223 几乎在所有的擦除方式中表现良好的性能，而排名较低的模型例如 CPSC016 全部表现较差的性能。这也从另一个角度反映出模型的优劣直接决定了分类性能的好坏。

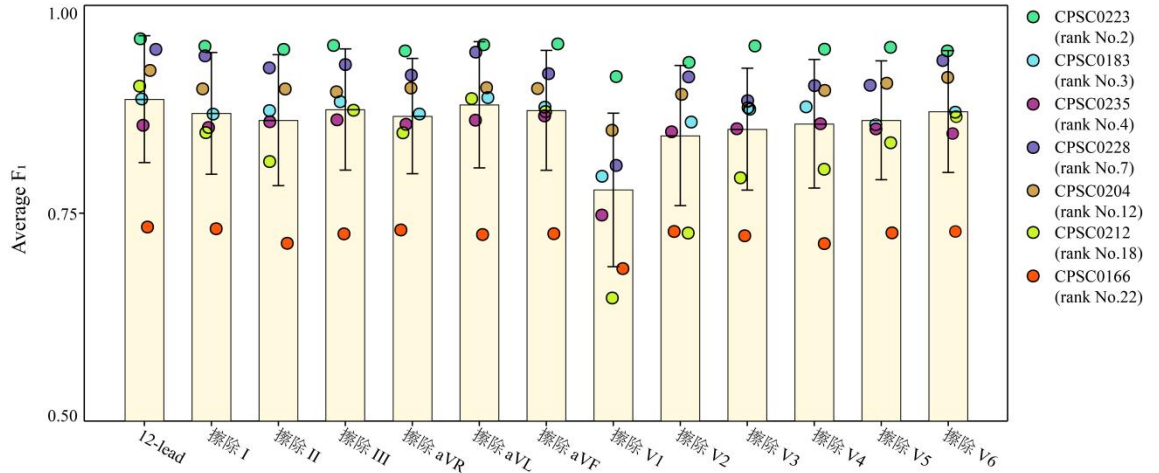


图 4.7 基于十三份测试数据在七个 DNN 模型的平均 $F_1$

#### 4.4.2 详细结果

受空间的限制以及阅读观感，关于第四章实验的全部结果放置在此。不同模型对不同擦除组合所表现出的结果呈现一致化的趋势，对于正常类型的识别所有模型都表现出 0.9 以上的 $F_1$ 分数，而对于 PAC 和 STE 均表现出较低的评分，意外的是 model-3 (ResNet+LSTM) 在 case-4 (擦除导联 aVR、aVL 和 aVF) 作为输入的情况下对于 STE 类型的评分为 0，这个结果说明在这种情况下模型无法学习到关于 STE 合理的特征，也从另一个角度反映出不同擦除方式的二维心电图对特定类型的心律失常的自动识别存在误区，这种原因仍需进一步研究。-

表 4.5 不同擦除组合在不同 DNN 模型的不同心律失常的 $F_1$ 分数

模型	填充方式	$F_1$ 分数								
		正常	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE
model-1	case-1	<b>0.939</b>	0.874	0.752	0.9	0.792	0.57	0.75	0.743	0.452
	case-2	<b>0.934</b>	0.87	0.794	0.897	0.814	0.58	0.772	0.705	0.453
	case-3	<b>0.926</b>	0.863	0.774	0.911	0.763	0.539	0.755	0.688	0.473
	case-4	<b>0.926</b>	0.871	0.709	0.878	0.784	0.502	0.758	0.664	0.449
	case-5	<b>0.927</b>	0.865	0.764	0.924	0.797	0.527	0.733	0.651	0.4
	case-6	<b>0.924</b>	0.855	0.694	0.829	0.753	0.6	0.765	0.688	0.343
model-2	case-1	<b>0.937</b>	0.873	0.834	0.915	0.803	0.526	0.765	0.74	0.519
	case-2	<b>0.93</b>	0.89	0.806	0.902	0.754	0.573	0.758	0.707	0.433
	case-3	<b>0.937</b>	0.883	0.779	0.914	0.794	0.612	0.754	0.741	0.452
	case-4	<b>0.931</b>	0.863	0.757	0.871	0.745	0.547	0.734	0.662	0.364
	case-5	<b>0.936</b>	0.887	0.814	0.916	0.788	0.545	0.753	0.736	0.474
	case-6	<b>0.914</b>	0.857	0.772	0.898	0.717	0.589	0.78	0.642	0.195
model-3	case-1	<b>0.936</b>	0.849	0.746	0.884	0.811	0.537	0.789	0.735	0.4
	case-2	<b>0.939</b>	0.872	0.78	0.91	0.806	0.564	0.789	0.695	0.453
	case-3	<b>0.942</b>	0.862	0.753	0.896	0.801	0.508	0.791	0.667	0.431
	case-4	<b>0.926</b>	0.845	0.741	0.884	0.779	0.4	0.707	0.604	0
	case-5	0.921	0.864	0.733	<b>0.928</b>	0.778	0.492	0.709	0.744	0.37
	case-6	<b>0.927</b>	0.869	0.754	0.872	0.735	0.585	0.777	0.662	0.184
model-4	case-1	<b>0.939</b>	0.868	0.782	0.923	0.806	0.536	0.783	0.718	0.494
	case-2	<b>0.931</b>	0.851	0.722	0.883	0.782	0.484	0.742	0.718	0.273
	case-3	<b>0.931</b>	0.874	0.773	0.909	0.784	0.576	0.801	0.721	0.286
	case-4	<b>0.919</b>	0.872	0.745	0.904	0.782	0.517	0.727	0.673	0.291
	case-5	<b>0.924</b>	0.85	0.765	0.879	0.795	0.521	0.722	0.661	0.462
	case-6	<b>0.926</b>	0.873	0.78	0.861	0.736	0.633	0.757	0.693	0.349
model-5	case-1	<b>0.939</b>	0.889	0.798	0.917	0.831	0.639	0.776	0.688	0.537
	case-2	<b>0.926</b>	0.831	0.634	0.882	0.797	0.474	0.719	0.553	0.333
	case-3	<b>0.942</b>	0.875	0.766	0.907	0.845	0.588	0.759	0.733	0.244
	case-4	<b>0.931</b>	0.875	0.799	0.892	0.785	0.569	0.734	0.673	0.348
	case-5	<b>0.936</b>	0.893	0.744	0.893	0.801	0.627	0.767	0.695	0.536
	case-6	<b>0.928</b>	0.838	0.79	0.873	0.743	0.567	0.771	0.702	0.426
model-6	case-1	<b>0.939</b>	0.874	0.752	0.9	0.792	0.57	0.75	0.743	0.452
	case-2	<b>0.934</b>	0.87	0.794	0.897	0.814	0.58	0.772	0.705	0.453
	case-3	<b>0.926</b>	0.863	0.774	0.911	0.763	0.539	0.755	0.688	0.473
	case-4	<b>0.926</b>	0.871	0.709	0.878	0.784	0.502	0.758	0.664	0.449
	case-5	<b>0.927</b>	0.865	0.764	0.924	0.797	0.527	0.733	0.651	0.4
	case-6	<b>0.924</b>	0.855	0.694	0.829	0.753	0.6	0.765	0.688	0.343

## 4.5 本章小结与讨论

针对十二导联 ECG 之间部分导联存在计算逻辑关系以及现有便携式心电监测设备存在电极较少无法得到完整十二导联 ECG 的问题,本章节通过设计擦除不同组合的十二导联 ECG 研究深度学习模型对自动识别心律失常的导联信号冗余性问题。

实验前期首先收集并融合了来自中国、德国和美国的四个开放获取的大型十二导联数据库,在某种程度上融合后的数据库代表了来自世界范围内的心律失常样本的集合。在此基础上经过适度训练好的模型具有良好的泛化性能。

其次根据十二导联不同的组成部分构建了六种不同擦除导联信息的组合方式,包括不进行擦除的原始十二导联 ECG、擦除双极肢体导联、单极加压肢体导联、胸前六导联、擦除导联 aVL 和擦除计算逻辑冗余的四个导联。通过改进第三章提出的二维心电图,使得在二维心电图的通道维度包含样本不同时间段的多片十二导联 ECG 信号,既保证了训练样本与患者对应的唯一性,也保证了十二导联 ECG 信号在时间和空间维度两方面的信息。

最后延伸和扩展了第三章中提出的 DSE-ResNet 模型,通过添加 LSTM、BiLSTM 以及利用 ResNet、SE 块得到五种提取信息不同的神经网络结构。

实验结果表明五个 DNN 模型得到的 *Precision*、*Recall* 和  $F_1$  分数除去正常类型,其余心律失常类型基本分布在区间[0.7,0.8]内。相较于第三章 DSE-ResNet 的分类性能,这个结果低于最初的预期。一个主要原因是目前使用的融合数据集的地理覆盖范围广,人群特性差异大,样本数据库的标注标准不一,导致同一心律失常类型的样本存在较大差异,诊断 PAC 类型结果较差也能印证这一点。符合预期的是基于五个 DNN 模型的仿真结果表明,从十二导联 ECG 中擦除部分导联确实会降低 DNN 模型的识别性能。与使用完整的十二导联 ECG 作为输入相比,当使用本章节设计的多种不同擦除组合的导联作为输入时,这些组合的最优  $F_1$  分数仅降低了 0.016。这个减少量是非常有限的,尤其是当只有导联 I、II、III 和导联 II、aVL、aVR 和 aVF 被擦除时。这印证了图 4.1 (b) 中这些导联之间存在逻辑换算关系而导致自动识别过程中导联信息冗余的问题。

在最初的假设中擦除导联 II、aVL、aVR 和 aVF 与其他擦除组合相比影响应该最小。但实际结果表明,导联 I、II、III 被擦除是除使用完整十二导联数据外的最佳选择。经过分析造成这种情况的原因可能来自三个方面。首先模型可能在很大程度上决定了分类性能,本章节只是根据二维心电图的特性进行二维卷积神经网络的设计,并未能穷尽网络模型,因此可能存在更加适合类似于二维心电图这种结构的模型;其次对于某些特定的心律失常类型,可能需要特定导联的参与,如 ST 型。模型根据十二导联进行特征学习,给出的判断也总是依赖于整体十二

导联,如果能够针对不同心律失常类型给定不同导联权重进行学习可能会有较好的结果;最后基于这种擦除导联的方式可能会让模型学习到不合理的特征,尽管训练集和测试集都保持了一样的擦除方式。例如填充 0 在整体上会让模型认为某些导联旁边的导联特定值就为 0,但是为了保证输入维度的统一性和模型参数结构的统一性,最终依旧选择了这种方式。

Alday 等人的研究<sup>[78]</sup>表明,使用深度学习进行心律失常分类,分类误差不限于年龄、种族、性别,还可能存在模型结构、优化函数等因素。在本研究中,选择的模型最大限度地保持了主体结构,并且只改变了一些层的结构。显然结果表明,这些略有不同的模型在相同导联组合的情况下得到的分类性能差别不大。

因此以牺牲较小的识别性能为代价,换取便携式设备基于擦除导联组合方式进行心律失常诊断是合理的。一方面不需要便携式设备添加过多的电极,能够记录的导联与填充 0 的其他导联组成二维心电图,输入到通用的、统一的 DNN 模型中进行识别,既不要求硬件设备有过高的算力,也不必花费大量的时间训练不同组合的模型。

综上所述,在基本保持自动心律失常识别性能的前提下,擦除适当的导联信息以降低测量复杂度是可行的。未来的工作可以扩展到研究所有可能的导联组合对自动心律失常识别的影响、研究识别特定心律失常所需的导联组合,并提高便携式设备的识别性能等。



## 第5章 基于 Flask 的心律失常自动分类平台

### 5.1 引言

实际生活中，用户很少直接接触算法，通常通过操控硬件或软件界面发出指令完成需求。算法附着于软件或者硬件模块，展示良好的接口给用户是作为技术人员应具备的素养。本章节旨在提升心律失常分类算法的可用性和易操作性。并将第三章中训练得到的 DSE-ResNet 集成模型以及十个单一最优模型进行完整接口封装，并基于 Flask 框架构建心律失常在线分类平台。

### 5.2 采用的框架和技术模块

基于 Flask 的心律失常自动分类平台采用三层浏览器/服务器（B/S）架构。相较于传统的客户端/服务端（C/S）架构，B/S 架构能够屏蔽客户端硬件设备差异，尤其是涉及深度学习这种强依赖硬件设备型号的计算。不同客户端可通过浏览器访问服务器实现任务交互。而系统功能的核心部分集中在服务器进行开发，保证了系统的安全性。图 5.1 展示了三层 B/S 架构模型图。

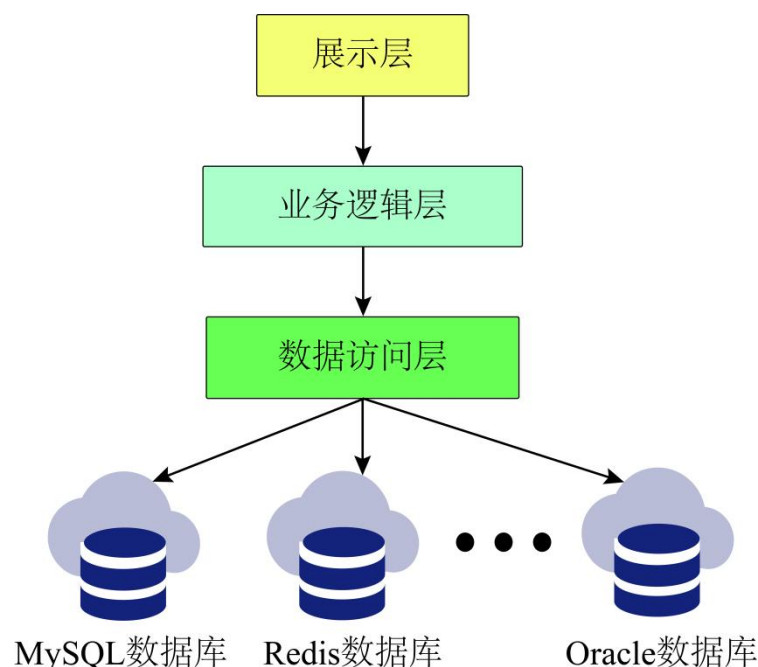


图 5.1 三层 B/S 架构模型图

三层 B/S 架构包含展示层、业务逻辑层和数据访问层三部分。展示层主要承

担了设计和显示页面视图、接收参数信息、返回与渲染数据模型等任务。业务逻辑层用于处理具体业务，即实现功能的详细设计。数据访问层基于不同的数据库实现数据的存储（持久化）和提取。本系统前端采用 Bootstrap 封装超文本标记语言（HTML）页面，后端使用 Python 实现业务逻辑处理，前后端通过 Flask 框架实现交互，使用 Mysql 数据库实现数据持久化。

Bootstrap 是一个用于快速开发 Web 应用程序和网站的前端 UI 框架。它提供了 HTML 和层叠样式表（CSS）的常用结构模型，使得开发者通过简单修改和调用就能构建个性化的页面。

Python 主要服务于业务逻辑的实现以及前后端页面的交互。在本系统中业务逻辑的实现包括文件的读取、数据提取、数据预处理、DSE-ResNet 模型加载、数据读入模型并分类。前后端页面交互包括用户点击操作、用户提交表单和回显分类结果到 HTML 页面。

前后端页面交互依赖于 Flask 框架实现。Flask 是 Python 编写的一个轻量级 Web 开发框架，主要包括 Werkzeug 模块和 Jinja2 模板引擎等开源库。Werkzeug 主要用于网页跳转请求和 Web 服务器网关接口系统。Jinja2 模板引擎具备强大的自动 HTML 转义系统和模板继承机制，能够有效防范跨站脚本攻击和提升开发效率。

MySQL 是当前最流行的关系型数据库管理系统之一。它具备体积小、处理数据量巨大、支持多线程、保证较高的数据安全性和数据一致性等优势。常应用于中小型网站数据库。

### 5.3 功能模块设计

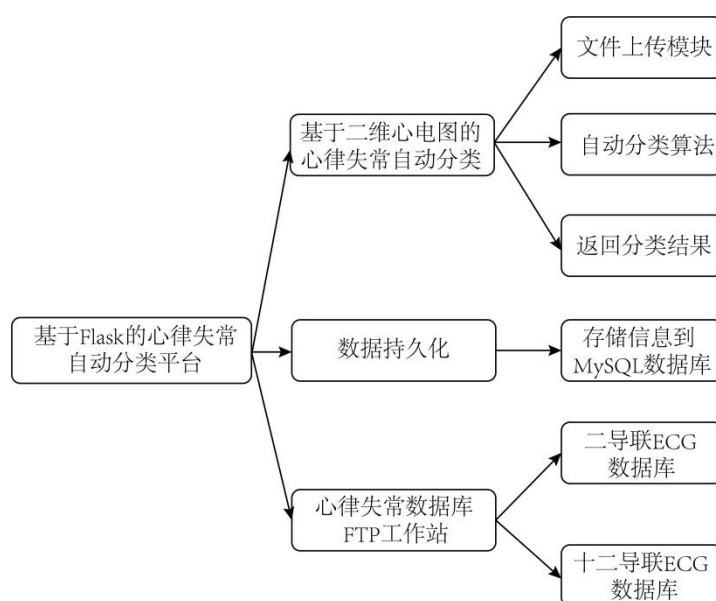


图 5.2 基于 Flask 的心律失常自动分类平台功能模块图

基于Flask的心律失常自动分类平台提供心律失常自动分类、数据持久化和开源数据库介绍与下载三个功能模块。图 5.2 展示了这些功能模块。

### 5.3.1 自动分类模块

自动分类模块包括文件上传模块、自动分类算法和返回分类结果三部分。文件上传模块是分类平台的核心功能，用户可上传由十电极系统产生的十二导联 ECG 信号文件到服务器，服务器完成数据处理和算法运行，然后返回分类结果到 HTML 页面，其具体处理逻辑如图 5.3 所示。

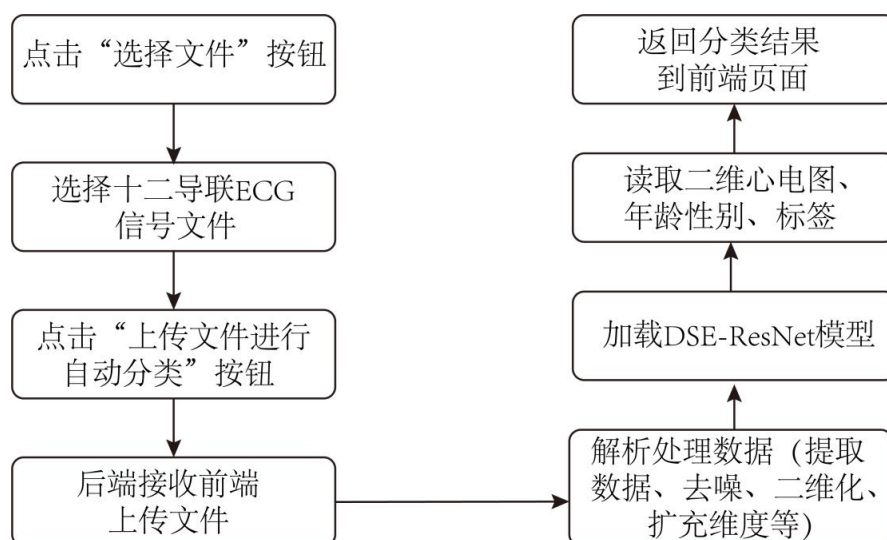


图 5.3 文件上传模块处理逻辑示意图

以CPSC2018 数据库提供的样本文件为基准，文件上传模块要求用户上传的文件名称后缀以“.mat”结尾。Mat文件中存储变量名为“ECG”的 1×1 结构体，结构体包含性别字段（值为“Male”或者“Female”）、年龄字段以及十二导联 ECG 信号值。图 5.4 和图 5.5 分别展示了Mat文件存储结构体内容及十二导联 ECG 存储格式。



图 5.4 Mat 文件存储结构体内容示意图

变量 - ECG.data									
ECG.data									
	1	2	3	4	5	6	7	8	9
1	-0.0166	-0.0177	-0.0082	-0.0069	-0.0036	-0.0242	-0.0508	-0.0496	-0.0316
2	-0.0407	-0.0470	-0.0375	-0.0411	-0.0417	-0.0608	-0.0733	-0.0675	-0.0601
3	-0.0241	-0.0292	-0.0294	-0.0342	-0.0382	-0.0366	-0.0226	-0.0179	-0.0285
4	0.0282	0.0323	0.0231	0.0242	0.0221	0.0429	0.0623	0.0583	0.0456
5	0.0033	0.0057	0.0100	0.0130	0.0174	0.0055	-0.0138	-0.0151	-0.0017
6	-0.0319	-0.0374	-0.0337	-0.0378	-0.0402	-0.0488	-0.0482	-0.0428	-0.0441
7	0.0369	0.0411	0.0313	0.0403	0.0520	0.0623	0.0604	0.0514	0.0453
8	0.0372	0.0562	0.0462	0.0518	0.0574	0.0557	0.0449	0.0380	0.0380
9	0.0560	0.0819	0.0679	0.0692	0.0707	0.0686	0.0591	0.0477	0.0422
10	0.0406	0.0565	0.0426	0.0469	0.0519	0.0461	0.0348	0.0256	0.0151
11	-0.0131	-0.0217	-0.0281	-0.0304	-0.0316	-0.0373	-0.0413	-0.0454	-0.0431
12	0.0330	0.0434	0.0320	0.0345	0.0322	0.0294	0.0275	0.0254	0.0213
13									

图 5.5 十二导联 ECG 存储格式示意图

图 5.6 展示了平台用户操作界面，用户依次点击“选择文件”和“上传文件进行自动分类”，后端接收前端提交的 Mat 文件并进行数据处理等操作，等待运行结束，返回文件上传成功状态以及分类结果。



图 5.6 平台用户操作界面

为提供给用户不同DNN模型的心律失常分类结果，后端同时返回DSE-ResNet集成模型的输出及十个最优单一模型的输出。图 5.7 展示了一次用户和平台之间的正常交互过程，用户选择了CPSC2018 数据中名为“A0127.mat”的样本文件，点击上传文件进行自动分类，后端算法运行结束后，返回结果格式如下：

“文件上传成功！”。

“集成模型的预测的结果为:右束支传导阻滞LBBB”。

“10 个最优单一模型的预测的结果分别为:['房性早搏PAC', '室性早搏PVC', '右束支传导阻滞LBBB', '正常Normal', '右束支传导阻滞LBBB', '右束支传导阻滞LBBB', 'ST段抬高', '右束支传导阻滞LBBB', 'ST段抬高', '右束支传导阻滞LBBB']”。



图 5.7 用户正常操作示意图

### 5.3.2 数据持久化

基于Flask的心律失常自动识别平台使用MySQL数据库存储数据。用户上传的文件存储在服务器目录下，每上传一次文件数据库会增加一条信息。该信息包括唯一标识该条记录的id、上传文件的名称、上传文件的日期和自动识别的结果。表 5.1 展示了数据库表的逻辑结构设计。

表 5.1 数据库存储信息表

字段名称	字段类型	数据类型及长度	说明
id	主键 Primary Key	int(11)	自增,不为空
file_name	字符 Char	varchar(255)	文件名称
date	时间 DateTime	varchar(255)	上传文件日期
detect_result	字符 Char	varchar(255)	自动识别结果

根据数据库的逻辑结构设计，通过以下SQL语句建立数据库和表，数据库名为“ecg\_detector”，表名为“information”。

```
create database ecg_detector;
create table information (
    id int(11) primary key not null auto_increment,
    file_name varchar(255),
    date varchar(255),
    detect_result varchar(255)
)engine=INNODB default charset=utf8;
```

图 5.8 展示了 information 表的最终结构以及存储信息的格式。样本“A0127.mat”经过平台自动识别后会在数据库表中插入一条新的记录。

```
mysql> select * from information;
```

id	file_name	date	detect_result
1	mydata.mat	2022-11-01 14:30:38.944061	集成模型的预测的结果为:ST段抬高
2	test.mat	2022-11-01 14:31:22.654815	集成模型的预测的结果为:右束支传导阻滞LBBB
3	12lead.mat	2022-11-01 14:35:46.941341	集成模型的预测的结果为:正常Normal
4	9.mat	2022-11-01 14:48:15.154263	集成模型的预测的结果为:正常Normal
5	12.mat	2022-11-01 14:55:22.412572	集成模型的预测的结果为:正常Normal
6	李四.mat	2022-11-01 15:30:38.226154	集成模型的预测的结果为:正常Normal
7	张三.mat	2023-02-01 22:28:58.115381	集成模型的预测的结果为:ST段抬高
8	A0072.mat	2023-02-15 22:40:58.991705	集成模型的预测的结果为:ST段抬高
9	A0127.mat	2023-02-15 22:43:49.925486	集成模型的预测的结果为:右束支传导阻滞LBBB

9 rows in set (0.01 sec)

图 5.8 information 表结构

### 5.3.3 FTP 工作站

本平台收集了不同心律失常数据库的下载地址，建立了一个类似于阿里源提供第三方数据源下载的服务。图 5.9 展示了 FTP 工作站页面。

<b>FTP工作站</b>	
MIT-BIH数据库.....	<a href="#">下载</a>
ST-T数据库.....	<a href="#">下载</a>
CPSC2018数据库.....	<a href="#">下载</a>
PTB数据库.....	<a href="#">下载</a>
乔治亚州心律失常 (G12EC) 数据库.....	<a href="#">下载</a>
INCART数据库.....	<a href="#">下载</a>

图 5.9 FTP 工作站页面

平台初始页面提供了四个常见心律失常数据库的介绍与下载链接。同时用户可点击“相关ECG数据库”页面跳转至FTP工作站，点击“下载”即可超链接至相关数据库下载地址。

## 5.4 本章小节与讨论

本章基于 Flask 构建了一个在线的心律失常自动分类平台，用户可通过自主上传十二导联 ECG 文件得到正常类型和 8 类心律失常类型的自动识别，同时对于研究人员可通过本平台获取不同心律失常数据库的数据。本平台的优势在于分类

算法的实现集中在服务器，从而屏蔽了客户端硬件设备的差异。其次，采用框架式结构的开发使得系统的可扩展性较强。

当然本系统也存在局限性，例如系统目前提供的功能较少、能够兼容的文件格式仅限于以“.mat”为后缀的文件、能够识别的心律失常种类较少。未来的工作包括平台功能扩展与算法持续优化。



## 第 6 章 总结与展望

### 6.1 全文工作总结

随着计算机算力的不断提升以及人工智能的不断发展，越来越多的领域倾向于使用人工智能助力行业发展。智慧医疗领域也开始进入人工智能时代，心电图作为临床最常见的检查，它的智能检测占据了智慧医疗发展中重要的一环。随着 ECG 的数字化，动态心电监护仪全天候二十四小时不断产生数据。而心律失常存在种类多、突发性高、形成原因复杂等问题使得专业医生对心律失常的诊断存在误诊率高等问题。实现高精确率心律失常的自动识别来辅助医生进行诊断迫在眉睫。同时，科技的发展也给人们带来便利，越来越多的人开始在工作之余注重自己的身体状况，便携式心电监测设备的出现使得实时监测主体心电状态并提供心律失常预警成为可能。这些设备记录的信息使得人们能够实时观测自己的心电状态。但是由于电极较少而无法记录完整十二导联 ECG，因此无法获得全方位心脏状态信息。许多研究表明 DNN 模型基于十二导联 ECG 进行心律失常的识别性能优于单导联，这激发了研究深度学习过程中十二导联信息冗余性问题的兴趣，旨在验证利用较少导联组合实现高分类性能的心律失常自动识别的可行性。

近些年针对十二导联 ECG 的多数深度学习模型在进行训练时分为两步，首先对导联进行逐条训练，然后融合每个导联的训练特征，最后得到分类结果。这使得在训练开始时没有注意导联之间的关系。基于以上这些问题，本研究进行了实验，并得到了一些结论。本文工作的主要贡献可以概括如下：

（1）提出了一种多通道时间序列信号的二维化转换方法。将原始的十二导联 ECG 拼接成灰度图一样的二维化数据，每一列代表单导联的时间序列，每个“像素”代表心电图的一个电压值。二维化后的十二导联 ECG 既具备时间维度的连续性，又具备空间维度的相邻性。之后的研究对这种二维化方法进行了扩展，使得二维化后的十二导联 ECG 具备类似于图片 RGB 三个通道的数据，每个通道代表同一样本不同采样时间段的十二导联 ECG 信号。

（2）提出了一种二维 DNN 模型 DSE-ResNet 用于处理多通道时间序列 ECG 信号。DSE-ResNet 可以在整个训练阶段学习导联内部和导联之间的特征。对比 CPSC2018 中前五名模型的结果以及近些年的研究表明，本研究提出的模型在平均  $F_1$  值上表现良好的性能，并且在两个亚异常类型上取得了最好的测试结果。在实验过程中采用正交试验代替常见的网格搜索和随机选择用于选择超参数，这种方式能够有效减少调参时间。在评估模型阶段，基于小样本测试集进行的实验表



明使用基于投票策略的集成模型获得的分类性能更好。

(3) 提出了深度学习过程中导联信号存在冗余的问题,并针对这个问题提出一套具体的研究方案。通过建立六种不同的擦除导联信息的组合方式,擦除导联信号是为了模拟便携式设备无法测量这些信息,从而观测不同导联组合方式在不同 DNN 模型上的分类性能能否达到类似于完整十二导联 ECG 作为输入时的性能。并验证不同 DNN 模型使用相同擦除方式所得到的心律失常分类趋势是否一致。结果表明深度学习过程中导联 I、II 和 III 被擦除是除使用完整十二导联 ECG 信号外的最佳选择,平均  $F_1$  分数相较于完整十二导联仅降低了 0.016。产生这种现象的主要原因是导联 I、II、III 和导联 aVL、aVR、aVF 存在逻辑计算关系,使得只要存在部分导联信号 DNN 模型就能学习到足够的特征。这为 DNN 模型兼容不同型号的便携式设备创造了理论可能。

(4) 建立了一个基于 Flask 的线上心律失常自动识别平台。该平台为用户提供自助服务接口,用户可上传自己测量的十二导联 ECG 数据进行心律失常类别的自动识别。

## 6.2 未来研究展望

当然,经过长时间的研究和实验也发现了研究过程中出现的一些无法解决或者来得及解决的问题,例如:

(1) 二维化过程中十二导联的排序方式按照双极肢体导联 I、II、III;单极加压肢体导联 aVR、aVL、aVF;胸前导联 V1、V2、V3、V4、V5、V6 的顺序进行排序,排序方式是否影响导联间特征的学习目前并不清楚。由于十二导联全排序产生的组合过多,这需要花费很长的时间进行进一步研究。

(2) 导联信号冗余性的探究过程中,设计的擦除组合是从十二导联 ECG 的分类以及逻辑计算冗余的角度出发选择的。其他情况的擦除组合能够何种性能目前尚不清晰。

(3) 实验过程中,对于不同心律失常的识别给定不同导联的权重是一致的。但实际上特定类型的心律失常在某些导联表现更明显,因此十二导联不同的权重来判断指定类型的心律失常是更合理的。

(4) 对于 ST 类型的诊断,仍需要大量的样本供给。在对 CPSC2018 数据库的研究中,几乎所有的模型对 ST 尤其是 STE 的识别都呈现出较低的评价分数

(5) 目前基于 Flask 的心律失常自动识别平台能够兼容用户提交的十二导联 ECG 文件格式有限,还有待进行进一步的接口拓展和功能开发。

结合上述本研究忽略的问题,未来的研究方向包括:

(1) 研究二维化十二导联 ECG 在深度学习过程中的排序对分类结果的影响。

（2）研究擦除导联的其他组合方式对心律失常的影响,尽可能匹配到现有便携式心电监测设备能够记录的导联,从而实现不同设备的心律失常的高性能自动识别。

（3）研究不同心律失常类型和能够诊断它的导联之间的对应关系,从而在自动识别算法中给予特定心律失常对应导联的权重,加大识别的准确率。

（4）可视化不同导联擦除方式在训练过程中被关注的部分,以此判定 CNN 关注的局部信息主要依赖的导联是什么。

（5）收集更多的十二导联数据库,将不同心律失常类型的数量保持在同一水平,屏蔽因缺乏样本而导致模型学习特征不足够的问题。

（6）持续优化基于 Flask 的心律失常自动分类平台的功能和算法性能。同时关注系统的实时性和数据库的安全性。

## 参考文献

- [1] Weimann K, Conrad T O F. Transfer learning for ECG classification[J]. Scientific reports, 2021, 11(1): 1-12.
- [2] Shah A P, Rubin S A. Errors in the computerized electrocardiogram interpretation of cardiac rhythm[J]. Journal of electrocardiology, 2007, 40(5): 385-390.
- [3] 马丽媛,王增武,樊静等.《中国心血管健康与疾病报告 2021》概要[J].中国介入心脏病学杂志,2022,30(07):481-496.
- [4] Wang J, Wang P, Wang S. Automated detection of atrial fibrillation in ECG signals based on wavelet packet transform and correlation function of random process[J]. Biomedical Signal Processing and Control, 2020, 55: 101662.
- [5] Tsipouras M G, Fotiadis D I, Sideris D. An arrhythmia classification system based on the RR-interval signal[J]. Artificial intelligence in medicine, 2005, 33(3): 237-250.
- [6] Haseena H H, Mathew A T, Paul J K. Fuzzy clustered probabilistic and multi layered feed forward neural networks for electrocardiogram arrhythmia classification[J]. Journal of Medical Systems, 2011, 35: 179-188.
- [7] Mondéjar-Guerra V, Novo J, Rouco J, et al. Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers[J]. Biomedical Signal Processing and Control, 2019, 47: 41-48.
- [8] Afkhami R G, Azarnia G, Tinati M A. Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals[J]. Pattern Recognition Letters, 2016, 70: 45-51.
- [9] Martis R J, Acharya U R, Prasad H, et al. Application of higher order statistics for atrial arrhythmia classification[J]. Biomedical signal processing and control, 2013, 8(6): 888-900.
- [10] de Oliveira L S C, Andreão R V, Sarcinelli-Filho M. Premature ventricular beat classification using a dynamic Bayesian network[C]//2011 annual international conference of the IEEE engineering in medicine and biology society. IEEE, 2011: 4984-4987.
- [11] De Chazal P, O'Dwyer M, Reilly R B. Automatic classification of heartbeats using ECG morphology and heartbeat interval features[J]. IEEE transactions on biomedical engineering, 2004, 51(7): 1196-1206.

- [12] Ince T, Kiranyaz S, Gabbouj M. A generic and robust system for automated patient-specific classification of ECG signals[J]. IEEE Transactions on Biomedical Engineering, 2009, 56(5): 1415-1426.
- [13] Wang J S, Chiang W C, Hsu Y L, et al. ECG arrhythmia classification using a probabilistic neural network with a feature reduction method[J]. Neurocomputing, 2013, 116: 38-45.
- [14] Martis R J, Acharya U R, Min L C. ECG beat classification using PCA, LDA, ICA and discrete wavelet transform[J]. Biomedical Signal Processing and Control, 2013, 8(5): 437-448.
- [15] Martis R J, Acharya U R, Prasad H, et al. Automated detection of atrial fibrillation using Bayesian paradigm[J]. Knowledge-Based Systems, 2013, 54: 269-275.
- [16] 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.
- [17] Osowski S, Hoai L T, Markiewicz T. Support vector machine-based expert system for reliable heartbeat recognition[J]. IEEE transactions on biomedical engineering, 2004, 51(4): 582-589.
- [18] Lagerholm M, Peterson C, Braccini G, et al. Clustering ECG complexes using Hermite functions and self-organizing maps[J]. IEEE Transactions on Biomedical Engineering, 2000, 47(7): 838-848.
- [19] Guo G, Wang H, Bell D, et al. KNN model-based approach in classification[C]//On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7, 2003. Proceedings. Springer Berlin Heidelberg, 2003: 986-996.
- [20] Ullah A, Anwar S M, Bilal M, et al. Classification of arrhythmia by using deep learning with 2-D ECG spectral image representation[J]. Remote Sensing, 2020, 12(10): 1685.
- [21] 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.
- [22] Hannun A Y, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network[J]. Nature medicine, 2019, 25(1): 65-69.

- [23] Wang R, Fan J, Li Y. Deep multi-scale fusion neural network for multi-class arrhythmia detection[J]. IEEE journal of biomedical and health informatics, 2020, 24(9): 2461-2472.
- [24] Chen T M, Huang C H, Shih E S C, et al. Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model[J]. Iscience, 2020, 23(3): 100886.
- [25] Zhang D, Yang S, Yuan X, et al. Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram[J]. Iscience, 2021, 24(4): 102373.
- [26] Chiou Y A, Syu J Y, Wu S Y, et al. Electrocardiogram lead selection for intelligent screening of patients with systolic heart failure[J]. Scientific reports, 2021, 11(1): 1-12.
- [27] Zhao Z, Fang H, Relton S D, et al. Adaptive lead weighted ResNet trained with different duration signals for classifying 12-lead ECGs[C]//2020 Computing in Cardiology. IEEE, 2020: 1-4.
- [28] 王英龙,成曦,舒明雷,朱清,周书旺. 基于卷积神经网络和长短期记忆网络的心电图分类方法[P]. 山东省: CN110179453A,2019-08-30.
- [29] Hong S, Wu M, Zhou Y, et al. ENCASE: An ENsemble CLASSifiEr for ECG classification using expert features and deep neural networks[C]//2017 Computing in cardiology (cinc). IEEE, 2017: 1-4.
- [30] 《中国心血管健康与疾病报告 2021》概述[J]. 中国心血管病研究,2022,20(07):577-596.
- [31] 白杰云,王宽全,张恒贵.基于心脏电生理模型的心律失常机制研究进展[J].生物化学与生物物理进展,2016,43(02):128-140.
- [32] 窦春江,徐盛开,寇炜.临床心电图的解析与应用方法[J].实用心电学杂志,2015,24(05):353-357.
- [33] 杨静,刘鸣.动态心电图导联系统的前世今生[J].实用心电学杂志,2017,26(04):263-266+270.
- [34] Goldberger E. A simple, indifferent, electrocardiographic electrode of zero potential and a technique of obtaining augmented, unipolar, extremity leads[J]. American Heart Journal, 1942, 23(4): 483-492.
- [35] Moody G B, Mark R G. The impact of the MIT-BIH arrhythmia database[J]. IEEE engineering in medicine and biology magazine, 2001, 20(3): 45-50.
- [36] Taddei A, Distant G, Emdin M, et al. The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography[J]. European heart journal, 1992, 13(9): 1164-1172.

- [37] Reyna M A, Sadr N, Alday E A P, et al. Will two do? Varying dimensions in electrocardiography: the PhysioNet/Computing in Cardiology Challenge 2021[C]//2021 Computing in Cardiology (CinC). IEEE, 2021, 48: 1-4.
- [38] Reyna MA, Sadr N, Perez Alday EA, et al. Issues in the automated classification of multilead ECGs using heterogeneous labels and populations. *Physiol. Meas*, 2022.
- [39] Liu F, Liu C, Zhao L, et al. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection[J]. *Journal of Medical Imaging and Health Informatics*, 2018, 8(7): 1368-1373.
- [40] Wagner P, Strodthoff N, Bousseljot R D, et al. PTB-XL, a large publicly available electrocardiography dataset[J]. *Scientific data*, 2020, 7(1): 154.
- [41] Han J, Moraga C. The influence of the sigmoid function parameters on the speed of backpropagation learning[C]//From Natural to Artificial Neural Computation: International Workshop on Artificial Neural Networks Malaga-Torremolinos, Spain, June 7–9, 1995 Proceedings 3. Springer Berlin Heidelberg, 1995: 195-201.
- [42] Kalman B L, Kwasny S C. Why tanh: choosing a sigmoidal function[C]//[Proceedings 1992] IJCNN International Joint Conference on Neural Networks. IEEE, 1992, 4: 578-581.
- [43] Ide H, Kurita T. Improvement of learning for CNN with ReLU activation by sparse regularization[C]//2017 international joint conference on neural networks (IJCNN). IEEE, 2017: 2684-2691.
- [44] Bagui S, Nandi D, Bagui S, et al. Machine learning and deep learning for phishing email classification using one-hot encoding[J]. *Journal of Computer Science*, 2021, 17(7): 610-623.
- [45] Chen B, Guo W, Li B, et al. A study of deep feature fusion based methods for classifying multi-lead ECG[J]. *arXiv preprint arXiv:1808.01721*, 2018.
- [46] Alday E A P, Gu A, Shah A J, et al. Classification of 12-lead ecgs: the physionet/computing in cardiology challenge 2020[J]. *Physiological measurement*, 2020, 41(12): 124003.
- [47] Surawicz, Borys, et al. "AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: part III: intraventricular conduction disturbances: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the

- Heart Rhythm Society: endorsed by the International Society for Computerized Electrocardiology." *Circulation* 119.10 (2009): e235-e240.
- [48] Zhang J, Liu A, Gao M, et al. ECG-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network[J]. *Artificial Intelligence in Medicine*, 2020, 106: 101856.
- [49] Zhu J, Zhang Y, Zhao Q. Atrial fibrillation detection using different duration ECG signals with SE-ResNet[C]//2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP). IEEE, 2019: 1-5.
- [50] Zhu Z, Wang H, Zhao T, et al. Classification of cardiac abnormalities from ECG signals using SE-ResNet[C]//2020 Computing in Cardiology. IEEE, 2020: 1-4.
- [51] 赵哲耘,刘玉敏,孙静静.基于混合正交试验的卷积神经网络参数优化[J].*统计与决策*,2021,37(08):46-50.
- [52] McCaffrey J D. Generation of pairwise test sets using a genetic algorithm[C]//2009 33rd annual IEEE international computer software and applications conference. IEEE, 2009, 1: 626-631.
- [53] 郑根让,林正春.基于 Pairwise 算法的组合测试应用研究[J].*电子技术与软件工程*,2022(15):49-52.
- [54] Vollmer M, Sodmann P, Nath N, et al. A Convolutional Neural Network for ECG Annotation as the Basis for the Classification of Cardiac Rhythms[J]. *Ulmer Informatik-Berichte*, 2019, 3(2): 4.
- [55] Daud S S, Sudirman R. Butterworth bandpass and stationary wavelet transform filter comparison for electroencephalography signal[C]//2015 6th international conference on intelligent systems, modelling and simulation. IEEE, 2015: 123-126.
- [56] Welch P. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms[J]. *IEEE Transactions on audio and electroacoustics*, 1967, 15(2): 70-73.
- [57] Bhanja S, Das A. Impact of data normalization on deep neural network for time series forecasting[J]. *arXiv preprint arXiv:1812.05519*, 2018.
- [58] Patro S, Sahu K K. Normalization: A preprocessing stage[J]. *arXiv preprint arXiv:1503.06462*, 2015.
- [59] Houssein E H, Ibrahim I E, Neggaz N, et al. An efficient ECG arrhythmia classification method based on Manta ray foraging optimization[J]. *Expert systems with applications*, 2021, 181: 115131.

- [60] Houssein E H, Hassaballah M, Ibrahim I E, et al. An automatic arrhythmia classification model based on improved marine predators algorithm and convolutions neural networks[J]. Expert Systems with Applications, 2022, 187: 115936.
- [61] Yao Q, Wang R, Fan X, et al. Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network[J]. Information Fusion, 2020, 53: 174-182.
- [62] McCabe J M, Armstrong E J, Ku I, et al. Physician accuracy in interpreting potential ST-segment elevation myocardial infarction electrocardiograms[J]. Journal of the American Heart Association, 2013, 2(5): e000268.
- [63] CPSC2018[EB/OL]. <http://2018.icbeb.org/Challenge.html>.
- [64] Liu Z, Meng X A, Cui J, et al. Automatic identification of abnormalities in 12-lead ECGs using expert features and convolutional neural networks[C]//2018 International Conference on Sensor Networks and Signal Processing (SNSP). IEEE, 2018: 163-167.
- [65] Wang R, Yao Q, Fan X, et al. Multi-class arrhythmia detection based on neural network with multi-stage features fusion[C]//2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). IEEE, 2019: 4082-4087.
- [66] Almer J, Zusterzeel R, Strauss D G, et al. Prevalence of manual Strauss LBBB criteria in patients diagnosed with the automated Glasgow LBBB criteria[J]. Journal of Electrocardiology, 2015, 48(4): 558-564.
- [67] 汪康宁. 基于 Android 智能终端的实时 Holter 系统的软件设计[D].东南大学,2016.
- [68] Walsh III J A, Topol E J, Steinhubl S R. Novel wireless devices for cardiac monitoring[J]. Circulation, 2014, 130(7): 573-581.
- [69] Cho Y, Kwon J, Kim K H, et al. Artificial intelligence algorithm for detecting myocardial infarction using six-lead electrocardiography[J]. Scientific reports, 2020, 10(1): 20495.
- [70] Warrick P A, Lostanlen V, Eickenberg M, et al. Arrhythmia classification of reduced-lead electrocardiograms by scattering-recurrent networks[C]//2021 Computing in Cardiology (CinC). IEEE, 2021, 48: 1-4.
- [71] Bodini M, Rivolta M W, Sassi R. Classification of ECG signals with different lead systems using AutoML[C]//2021 Computing in Cardiology (CinC). IEEE, 2021, 48: 1-4.



- [72] Osnabrugge N, Rustemeyer F, Kaparakis C, et al. Multi-label classification on 12, 6, 4, 3 and 2 lead electrocardiography signals using convolutional recurrent neural networks[C]//2021 Computing in Cardiology (CinC). IEEE, 2021, 48: 1-4.
- [73] Medsker L R, Jain L C. Recurrent neural networks[J]. Design and Applications, 2001, 5: 64-67.
- [74] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [75] Zhou P, Shi W, Tian J, et al. Attention-based bidirectional long short-term memory networks for relation classification[C]//Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short thesis). 2016: 207-212.
- [76] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [77] Zweig M H, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine[J]. Clinical chemistry, 1993, 39(4): 561-577.
- [78] Alday E A P, Rad A B, Reyna M A, et al. Age, sex and race bias in automated arrhythmia detectors[J]. Journal of Electrocardiology, 2022, 74: 5-9.

## 致 谢

三年求学生涯行进至此，这段行程即将画上句点，我心中充满不舍。闭上眼睛，过去的一幕幕浮现眼前。作为一个山西人，大学求学在河南，研究生求学在山东，兜兜转转七年行程跨遍中部地区。从家乡小县城出发，见识到了周围省会城市便利的交通、先进的教育、开放的理念和充足的医疗资源。让我不禁感慨家乡与这些地方的差距甚大。在工大这几年的生活也让我对未来的具体规划发生了不小的转变。

研究生三年生活中，很有幸能够和刘文彬、于洋、王泽均等几位舍友轻松快乐友好的度过三年同住生活，尽管大家生活作息存在异同，但互相包容的生活态度、共同的部分爱好使得我们能够克服这些问题。感谢刘昆老师在疫情留校期间对自己的关心和帮助。同时感谢学校和国家为我提供的助学金。感谢父母仍然每月给予我充足的生活费，使我的生活不算那么拮据。更要感谢生源地贷款让我不再为学费发愁。

学习上，很感谢工大给我们提供了良好的学习场地，充足的学习资源，安静的学习环境，使得我能够在研一的课程学习以及后续的专业学习中不断提升专注度与效率。

科研上，首先非常感谢能够来到庞少鹏老师的课题组进行学习。很荣幸我能成为庞老师带的第一批学生中的一个。同门还有周岳、田志两个好兄弟，尽管在科研上我们三个人分属不同的方向，但我们经常会讨论一些小的算法设计和程序实现，闲暇之余经常互诉衷肠，互相勉励不断进步。在这三年的学习中让我记忆最深刻的是，每周四雷打不动的事情就是庞老师对我们一对一的指导，这使我们能够在自己的研究方向上突飞猛进。庞老师对科研的严肃严谨、对学生的认真负责、对生活的积极乐观时时刻刻都在影响着我。我相信我未来也一定能够成为一个雷厉风行、恪尽职守、积极向上的人。

研三的最后一个小尾巴，也真心祝愿我的师弟陈真琦、王天宇、邢云龙、王兴辉等能够在未来几年时间里不断进步，尽早出自己的成果。

最后，再次感谢所有帮助过我的老师，同学，朋友们。我们江湖再见！

## 在学期间主要科研成果

### 一、发表学术论文

- [1] Jiahao Li, Shao-peng Pang, Fangzhou Xu, Peng Ji, Shuwang Zhou, Minglei Shu. Two-dimensional ECG-based cardiac arrhythmia classification using DSE-ResNet[J]. Scientific Reports (中科院三区, IF=4.996, 与第三章研究内容相关)
- [2] Chuanzhe Zhang, Jiahao Li, Shao-peng Pang, Fangzhou Xu, Shuwang Zhou. A 12-lead ECG correlation network model exploring the inter-lead relationships[J]. EPL (中科院三区, IF=1.958, 与第四章研究内容相关)
- [3] Shao-peng Pang, Jiahao Li, Fangzhou Xu, Shuwang Zhou. Redundancy analysis of 12-lead ECG in automatic arrhythmia recognition[J]. Physiological Measurement (审中, 中科院三区, IF=2.688, 与第四章研究内容相关)

### 二、其它科研成果

- [1] 发明专利: 庞少鹏 李家豪 舒明雷. 一种基于十二导联心电图数据二维化的多输入残差神经网络的 ECG 信号分类方法, 实审阶段 (与第三章研究内容相关)
- [2] 发明专利: 庞少鹏 李家豪 李广良. 一种基于 Opencv 的菲涅尔透镜中心定位方法, 实审阶段