|  |  |  |  |
| --- | --- | --- | --- |
| 分类号： |  | 单位代码： | 10431 |
| 密 级： |  | 学 号： | 1043200544 |



**硕 士 学 位 论 文**

基于二维化多导联ECG信号分类方法

与导联冗余性研究

|  |  |
| --- | --- |
| 作者姓名 | 李家豪 |
| 专业 | 电子信息 |
| 所在学院 | 信息与自动化学院 |
| 指导教师姓名 | 庞少鹏 |
| 专业技术职务 | 副教授 |

|  |
| --- |
| 2023 年 4 月 9 日 |

**A Thesis Submitted for the Application of**

**the Master’s Degree of Engineering**

**Research on two-dimensional multi-lead ECG signal classification method and lead redundancy study**

|  |  |
| --- | --- |
| **Candidate:** | **Li Jiahao** |
| **Specialty:** | **Electronic and Information Engineering** |
| **Supervisor:** | **Professor Pang Shaopeng** |

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

**June, 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](#_Toc29646)

[ABSTRACT I](#_Toc17594)

[第1章 绪论 1](#_Toc30123)

[1.1 研究背景 1](#_Toc19692)

[1.2 研究目的和意义 2](#_Toc18777)

[1.3 国内外研究现状 3](#_Toc19507)

[1.4 本文的主要内容与整体结构 5](#_Toc7988)

[第2章 心律失常、深度学习相关的理论知识 7](#_Toc1971)

[2.1 心律失常概述 7](#_Toc3359)

[2.1.1 心律失常产生的原因 7](#_Toc21968)

[2.1.2 常见心律失常类型 8](#_Toc28628)

[2.1.3 心律失常的表现 8](#_Toc2227)

[2.2 十二导联心电图概述 9](#_Toc32707)

[2.2.1 十二导联心电图的起源和发展（【以下参考黄宛《临床心电图学》】） 9](#_Toc12854)

[2.2.2 临床十二导联的作用 11](#_Toc32232)

[2.3 常见心电图数据库 11](#_Toc13281)

[2.3.1 部分导联数据库 11](#_Toc16441)

[2.3.2 十二导联数据库 11](#_Toc25629)

[2.4 深度学习 13](#_Toc25242)

[2.3.1 深度学习的来源和发展 13](#_Toc6612)

[2.3.2 深度学习的原理 14](#_Toc18565)

[2.5本章小结 17](#_Toc14852)

[第3章 基金-股票网络映射模型 19](#_Toc15550)

[3.1 引言 19](#_Toc32542)

[3.2 建立基金-股票网络映射模型 19](#_Toc10182)

[3.3 结果 19](#_Toc12323)

[3.3.1 基金的投资水平 19](#_Toc23290)

[3.3.2 基金的投资分布 19](#_Toc28124)

[3.3.3 基金的投资倾向 19](#_Toc7540)

[3.3.4 投资策略 19](#_Toc11938)

[3.4 附录 19](#_Toc20778)

[3.4.1 基金-股票网络映射模型 19](#_Toc1420)

[3.4.2 改写的上证50(180)指数 19](#_Toc31206)

[3.5 本章小结 19](#_Toc30728)

[第4章 基于基金-股票网络映射模型的社团检测算法 21](#_Toc32002)

[4.1 引言 21](#_Toc32031)

[4.2 基于加权股票网络的社团检测 21](#_Toc26647)

[4.2.1 加权股票网络 21](#_Toc23542)

[4.2.2 基于加权股票网络的社团检测](#_Toc15861) **[错误！未定义书签。](#_Toc15861)**

[4.3 结果 21](#_Toc2895)

[4.3.1 社团分布 21](#_Toc1193)

[4.3.2 股票网络关键社团的演化 21](#_Toc533)

[4.4 本章小结 21](#_Toc25686)

[第5章 总结与展望 22](#_Toc10547)

[5.1 全文工作总结 22](#_Toc26154)

[5.2 未来研究展望 22](#_Toc24909)

[参考文献 23](#_Toc7102)

[致 谢 29](#_Toc25361)

[在学期间主要科研成果 31](#_Toc31383)

[一、发表学术论文 31](#_Toc14136)

# 

# 摘 要

心血管疾病严重危害着我国人民的身体健康，而心律失常又是心血管疾病中最常见的一组疾病，临床上的诊断通常通过医生对病人的十二导联心电图进行分析完成。然而心电图作为临床中最常见的检测手段之一以及拥有心脏相关疾病人群的增多，心电数据呈现爆炸式的增长，这无疑增加了现有医生的诊断负载。因此实现心律失常的自动化、智能化识别是智慧医疗发展应有之义。传统的智能识别严重依赖于人工特征的设计，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。近些年来，随着计算机性能的大幅提升，机器学习、深度学习等方法逐渐被应用在各个领域，并且表现出良好的性能。同时便携式设备也逐渐增添了监测心律的功能，为预防和提供预测诊断创造了条件。基于以上背景，本文主要研究了如下内容：

1. 针对现有研究往往从一维时序信号出发或者从二维心拍图出发进行研究而忽略心电信号原本在这两个维度都具有信息的问题。本文提出两种十二导联数据二维化的方法。将十二导联心电时序数据转化为二维平面，使得预处理后的心电数据既具有时间上的连续性又具有空间上的相邻性。
2. 针对现有研究往往忽略导联间关系的问题。本文提出基于ResNet和可分离式的SE模块的神经网络模型，称DSE-ResNet（ResNet with detached squeeze-and-excitation modules）。利用此模型可实现对二维化十二导联心电图导联内部与导联间的特征学习与提取。
3. 针对现有便携式设备只能计算少数几个导联信息，而无法利用全部十二导联的信息进行智能识别导致检测精度低、预测不准确的问题。本文对导联间的冗余性问题进行分析研究，旨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到。据此可减少便携式设备计算量，从而达到花费较少的计算成本得到相对可靠计算结果的目的。

综上，本文基于二维化十二导联心电数据，使用DSE-ResNet模型增强了心律失常分类的精确性，可作为医疗辅助诊断手段。同时本文对导联的冗余性进行研究，验证了深度学习过程中导联信息或可缺失的可能性，为便携式设备减少计算提供研究基础，具备较高的现实意义。

**关键词**： 十二导联；ResNet；心律失常；冗余性

# **ABSTRACT**

Complex Cardiovascular disease seriously endangers the health of our people, and arrhythmia is the most common group of diseases in cardiovascular disease. The clinical diagnosis is usually completed by the doctor's analysis of the patient's 12-lead electrocardiogram. However, the electrocardiogram is one of the most common detection methods in clinical practice, and the number of people with heart-related diseases is increasing, and the ECG data has shown explosive growth, which undoubtedly increases the diagnostic load of existing doctors. Therefore, the realization of automatic and intelligent identification of arrhythmia is the proper meaning of the development of smart medical care. Traditional intelligent recognition relies heavily on the design of artificial features, which makes the recognition process and results highly subjective, and cannot capture the deep-seated features of the ECG. In recent years, with the substantial improvement of computer performance, methods such as machine learning and deep learning have been gradually applied in various fields and have shown good performance. At the same time, portable devices have gradually added the function of monitoring heart rhythm, creating conditions for prevention and providing predictive diagnosis. Based on the above background, this paper mainly studies the following contents:

(1) In view of the fact that existing studies often start from one-dimensional time series signals or two-dimensional cardiograms, and ignore the problem that ECG signals originally have information in these two dimensions. In this paper, two methods for two-dimensionalization of twelve-lead data are proposed. The twelve-lead ECG time-series data is converted into a two-dimensional plane, so that the preprocessed ECG data has both temporal continuity and spatial adjacency.

(2) Aiming at the problem that existing research often ignores the relationship between leads. This paper proposes a neural network model based on ResNet and separable SE modules, called DSE-ResNet (ResNet with detached squeeze-and-excitation modules). This model can be used to realize the feature learning and extraction of the two-dimensional twelve-lead ECG inside and between the leads.

(3) For existing portable devices, only a few lead information can be calculated, and the information of all twelve leads cannot be used for intelligent identification, resulting in low detection accuracy and inaccurate prediction. This paper analyzes and researches the redundancy problem between leads, and aims to verify whether all twelve-lead information needs to be fully used in the intelligent recognition process. Accordingly, the calculation amount of the portable device can be reduced, so as to achieve the purpose of obtaining relatively reliable calculation results with less calculation cost.

In summary, based on the two-dimensional 12-lead ECG data, this paper uses the DSE-ResNet model to enhance the accuracy of arrhythmia classification, which can be used as an auxiliary medical diagnosis method. At the same time, this paper studies the redundancy of leads, verifies the possibility that lead information may be missing in the process of deep learning, and provides a research basis for portable devices to reduce calculations, which has high practical significance.

**Keywords**: 12-lead; DSE-ResNet; arrhythmia; 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和重要的一种疾病，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（Electrocardiogram，ECG）是临床医学日常实践中的基本工具，因其简单、无创、可靠等特点，多用于心律失常的临床诊断。在临床上，心律失常的诊断主要依赖医生对于心电图的分析，呈指数剧增的心电数据加重了医生的工作负担。全世界每年获得超过3亿次心电图[1]，且逐年上升，专业医生增长的速率远不能匹配心电图的产出数量，导致对病症的诊断无法及时提供。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，在许多临床设置中作为医生诊断的重要辅助手段。然而，现有的商用ECG诊断算法仍然显示出较高的错误率[2]。随着算力的不断提升、ECG信号的广泛数字化和多种算法的发展相结合，为原始心电信号进行大规模的处理提供了条件，也为重新创建ECG智能识别方法展示了可能性。

由于心电图的复杂性，对心电图的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计的行医经验，识别异常心电图在形态学中存在的问题。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、人工智能和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，检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。“十四五规划”中提及，加快优质医疗资源扩容和区域均衡布局，发展高端医疗设备，推广远程医疗等措施。“十四五”期间将推进“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》提出提高数字医疗设备、物联网等设备的生产制造水平,支持研发健康医疗相关的人工智能技术、医用机器人、健康和康复辅助器械、可穿戴设备以及相关微型传感器件,促进健康医疗智能装备产业升级。规划了人工智能在医疗影像辅助诊断系统中的应用研究。本研究是将人工智能应用于基于十二导联心电图的心律失常的智能识别研究中，同时分析十二导联在深度学习过程中的冗余性，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条件。因此本研究具有较高的临床意义和市场空间，具有广阔的应用前景，是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[3]，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说明，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结果，使得患者错过最佳治疗时间常有发生。因此为高危人群提供心电信号实时记录，实现心律失常的自动检测，为医生提供辅助诊断变得尤为重要。

在本文的研究中，我们针对心电图的时序特性以及空间特性，建立了一种新的二维化十二导联心电数据的方法。其中二维代表平面，平面中每个点代表某一导联的一个采样电压值。基于通用神经网络实现导联间特征和导联内部特征的提取，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，我们发现部分带有心电图记录功能的辅助设备通常只检测标准十二导联中的部分导联。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，利用十二导联实现心律失常的检测是合理的。然而全面的数据意味着庞大的检测量和计算量，为了搞清楚在利用深度学习实现心律失常的自动检测中，是否全部导联都发挥积极作用，我们对导联的冗余性进行研究。综上本文的研究目的和意义如下：

1. 研究成果提供了一个新的处理不同采样长度的十二导联数据的方法。使得处理后的十二导联数据保持相同的长度，同时让这些数据既具有时序上的连续性，也具有空间上的相邻性。
2. 本研究创建了一个通用的深度模型学习框架，能够对任意长度的十二导联数据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。
3. 本论文为研究导联之间冗余性的问题提供新的视角。为便携式设备提供实时检测心律失常结果的可能性。

## 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，这类算法通常分为特征工程和结果分类两部分。具体来说，研究人员首先利用数学方法从原始心电数据中手动提取了大量具有医学意义的心电图特征，如小波特征[4]、P-QRS-T复合特征[5,6]、心率变异性统计特征[7]，RR 间期相关统计特征[8,11]、高阶统计特征[8]和形态学特征[10,11]。主成分分析[12,13]和独立成分分析[14,15]等数学方法实现从高维空间向低维空间提取心电图特征。经过特征工程得到的特征可输入分类算法中进行分类，常见的分类算法有支持向量机[16,17]、自组织映射[18]、聚类[19]等机器学习算法用于分析人工特征，并给出预测结果。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，基于主观因素的特征工程会导致一些潜在的重要特征被忽略、心电信息不能够被完整的使用，这些情况可能会影响最终的分类性能。

近年来，深度神经网络以其强大的特征提取能力，极大地提升了语音识别、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解释ECG数据。一些研究受到启发，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。

在这些研究中由于选择不同的数据库进行实验，导致样本长度、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，标签以心拍给出）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[20]提出了一种基于单导联的二维卷积神经网络(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。将ECG信号分为八类即正常搏动、室性早搏、起搏搏动、右束支传导阻滞搏动、左束支传导阻滞搏动、心房早搏、心室扑动波搏动、心室逸搏搏动。得到的模型在MIT-BIH心律失常数据库进行评估并且取得99.11%的平均准确率。Jun等人[21]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转换成二维灰度波形图像作为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[22]开发了一种深度神经网络模型，实现了单导联心电数据端到端的学习，并完成12个心律失常类别的分类。得到的DNN模型在专用数据集上的敏感性超过了心脏病专家分类的平均敏感性，并且证明了端到端的深度学习方法可以从单导联心电图中分类出范围广泛的不同心律失常，其诊断性能与心脏病学家相似。

与单导联相比，多导联心电图包含更多有价值的信息，更有利于心律失常的自动分类。Wang等人[23]提出了一种基于多尺度特征提取和12导联心电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CPSC心律失常数据库中进行了测试，得到了82.8%的分数。Chen等人[24]提出了一种结合了卷积神经网络 (CNN)、循环神经网络和注意力机制的用于心律失常分类神经网络模型，应用于CPSC心律失常数据库并取得了排名第一的测试结果，同时基于单导联进行实验，发现使用单导联数据产生的性能仅比使用完整12导联数据稍差，其中导联aVR和V1导联最为突出。Ribeiro等人[25]展示了一个DNN模型，该模型在超过200万次标记检查的数据集中训练，并发现该模型在识别6种类型的心律失常方面优于心脏病医生。Yu-An Chiou等人[28]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，并评估单个ECG导联对收缩性心力衰竭（HF）分类结果的贡献，设计综合评分方法以改善结果。结果表明将12导联心电图信号中的V5导联和V6导联进行组合，可获得质量最高的结果。此外，zhao等人[26]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[27]提出一种基于卷积神经网络和长短期记忆网络的心电图分类方法，通过多导联心电图数据和卷积神经网络和长短期记忆网络（CNN-LSTM）组合模型实现心律失常自动分类的方法。一方面多导联心电图较单导联心电图蕴含更多信息，另一方面CNN-LSTM组合模型结合CNN和LSTM的优点，在学习空间数据结构和时间序列结构方面有着独特优势，利用多导联心电图数据训练CNN-LSTM能够提高网络的学习效率和心电图识别的精度。Hong等人[29]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。

基于深度学习的部分研究基于单导联ECG心电数据，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10万心拍），所以在切分样本后存在多个样本对应一位患者的情况，这可能导致模型心律失常分类模型过拟合，使测试结果偏颇。基于十二导联作为数据源的研究，往往忽略导联间的关系，例如深度神经网络模型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关系，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。

## 1.4 本文的主要内容与整体结构

本学位论文主要基于深度学习对二维化的十二导联心电数据进行自动分类和对深度学习过程中导联的冗余性进行分析与研究。第一章我们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。建立他们之间的联系。

第三章我们构建了通用的神经网络模型实现心律失常的自动识别。首先对二维化十二导联数据进行详细解释，包括二维化十二导联数据的原理、数据的预处理等。然后阐述了的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。

第四章我们对十二导联的冗余性进行实验分析。主体结构任然沿用第三章提出的神经网络模型，并对这个结构进行延伸和扩展，最终扩充为为五种不同结构的模型。对第三章中构建的二维化方式进行变革产生更加复杂的数据结构保证样本的唯一性。通过对冗余导联填充零的方式，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的问题，并对这种现象做出解释。

第五章为本文的总结和展望。

# 

# 第2章 心律失常、深度学习相关的理论知识

## 2.1 心律失常概述

心律失常（cardiac arrhythmia）是心血管疾病中最常见的一组疾病。它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。随着我国高龄化时代的到来和代谢危险因素持续流行[30]，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。它严重的危害着人类的健康，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

心律失常的主要病因根据获得类型可分为先天遗传和后天获得两种类型。遗传类型主要受家族基因和基因突变影响明显；后天获得类型主要包括生理性因素例如运动、情绪、精神和睡眠问题等和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传导，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[31]。正常情况下心电信号产生于窦房结，窦房结是正常心脏起搏点，产生的电信号首先经过房间束、心房肌传导到房室结，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的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都有几率产生病变，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，不同部位均可能导致心律失常，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. 房室交界区域发生的心律失常：
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心电图波形结构也有可能是不相同的。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。而在心电图信号中则表现为波形结构、波段持续时间、峰值等异于常人。

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）。心电波形反映了心脏兴奋过程中心脏不同方位的电位变化[32]。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，都有可能引发心律失常。图2.1展示了一个正常心跳节拍信号波形结构。

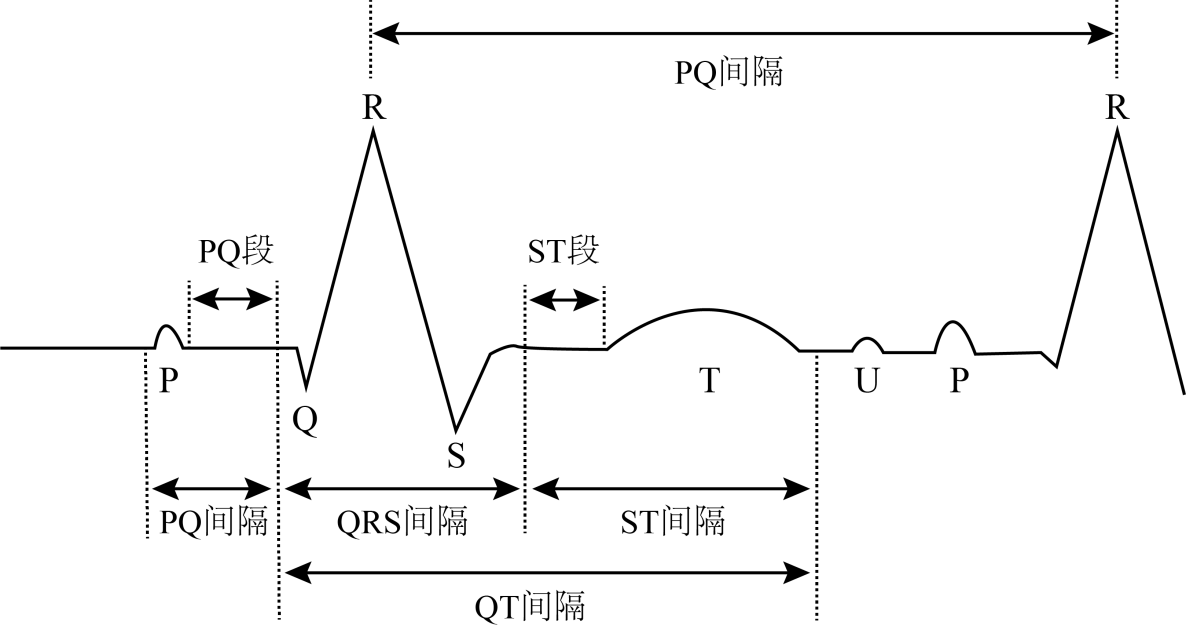


图2.1 正常心跳节拍信号波形结构

## 2.2 十二导联心电图概述

心电图具有无创、便捷等特点，是临床中应用最广泛的检查之一。而十二导联心电图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，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[33]。

### 2.2.1 十二导联心电图的起源和发展

二十世纪初，体表心电图由三个导联组成，称标准导联Ⅰ、Ⅱ、Ⅲ，也称双极肢体导联Ⅰ、Ⅱ、Ⅲ。它们的连接方式如下，标准导联Ⅰ：连接方式是将心电信号测量机阳极连接至人体左腕，阴极连接在人体右腕，达到记录人体左臂与右臂电极间的电压差的目的。标准导联Ⅱ：阳极连接人体左踝，阴极连接人体右腕，记录了人体左腿与右臂电极间的电压差。标准导联Ⅲ：阳极连接人体左踝，阴极

连接人体右腕，记录了人体左腿与左臂电极间的电压差。其具体测量位置图如图2.2所示。标准导联Ⅰ、Ⅱ、Ⅲ记录了心脏纵切面不同方位的电位变化。

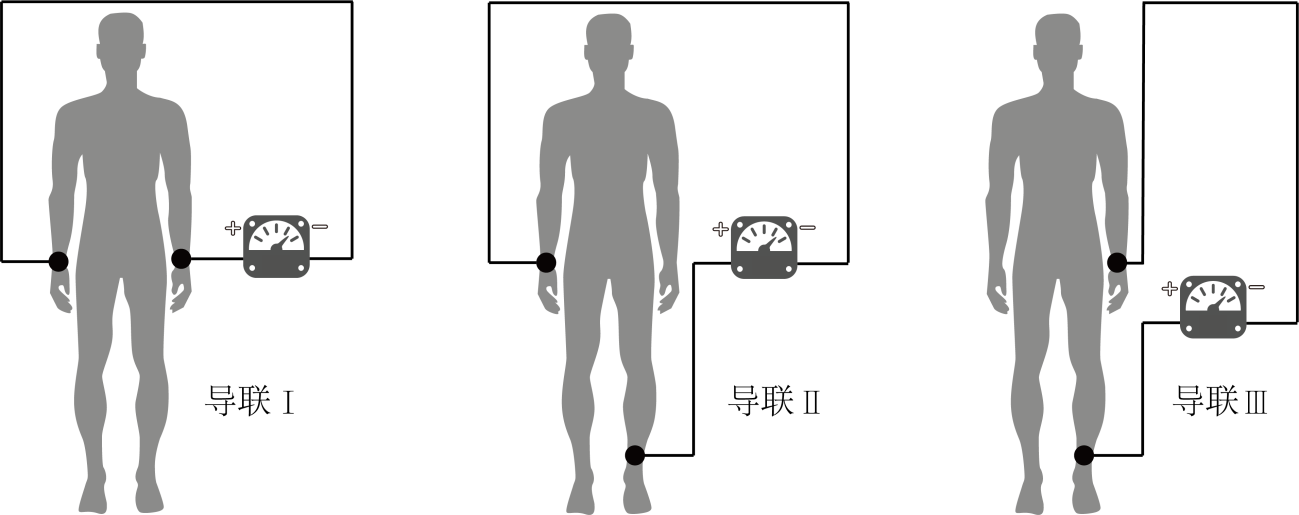


图2.2 标准导联Ⅰ、Ⅱ、Ⅲ的测量方式

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纵向切面电位变化的体表心电图系统。实验过程中，他们首先把标准导联Ⅰ、Ⅱ、Ⅲ的左、右手及左踝的电极板连在一起，发现综合电位几乎等于零，于是它们便把这个综合电极定义为“中心电端”，用它连于心电信号测量机的阴极端，另外用一个记录电极放在身体的不同部位。由于阴极电位近于零，因而记录电极录出的心电图，便被定义为“单极肢体导联心电图VR、VL、VF”，以便与先前的三个标准导联心电图相区别。为了测量心脏水平方向上的电位变化，Wilson团队提出了胸前导联系统V1、V2、V3、V4、V5、V6，胸前六导联电极位置图如图2.3所示，它们记录了心脏水平面不同方位的电位变化。

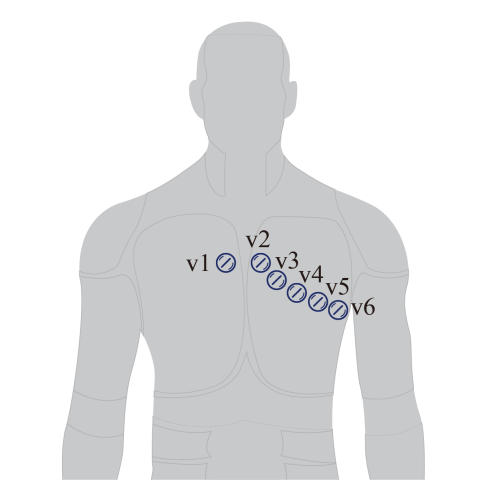


图2.3 胸前六导联电极位置图

介于单极肢体导联心电图VR、VL、VF所展示的心电图波形不明显，Goldberger[34]提出了加压单极肢体导联aVR、aVL、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，故称为加压单极肢体导联（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。这样便得到了加压单极肢体导联aVR、aVL及aVF。导联aVR、aVL、aVF的电极位置图如图2.4所示。这三个导联也记录了心脏水平面不同方位的电位变化。

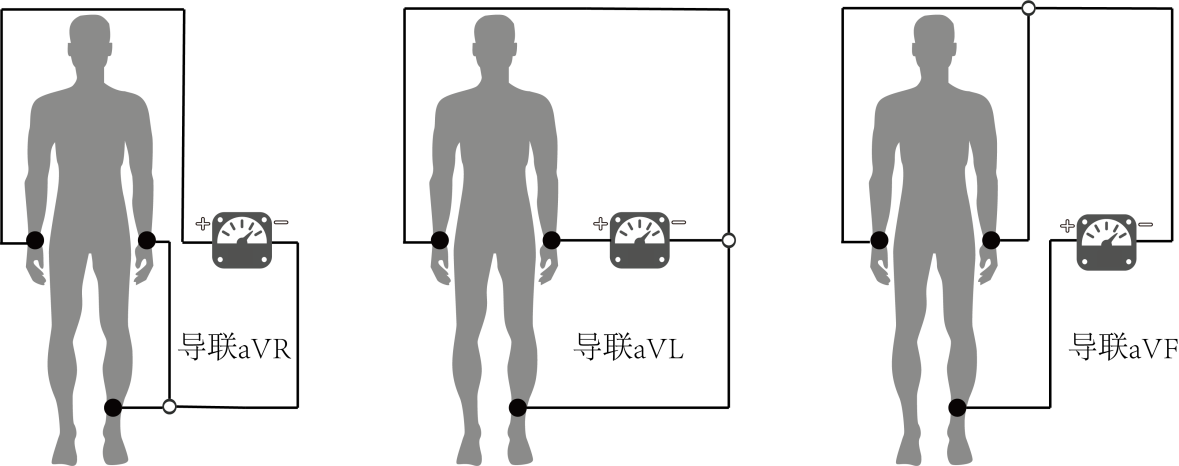


图2.4 加压单极肢体导联aVR、aVL及aVF的测量方式

至此，标准十二导联系统正式产生，它由十个电极组成如图2.4所示，共产生三部分导联信号，分别为双极肢体导联Ⅰ、Ⅱ、Ⅲ；单极加压肢体导联aVR、aVL、aVF；以及胸前导联V1、V2、V3、V4、V5、V6。它能够测量出来自于心脏横向切面和纵向切面的多种电信号，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的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ST段发生异常的检测效率，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（Hotler）仪器占有较高市场份额，更加速了十二导联心电图系统在临床中的应用。

## 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，给定部分导联数据的小型数据库和给定十二导联数据的大型数据库。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、V5中的两个。且由两位专业医生按照心拍对48组样本进行了详细注释，总共约110000个注释。自1980年以来，它已被用于该目的以及全球约500个地点的心脏动力学基础研究[35]。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，患者样本个数稀少等问题，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[36]。该数据库由79名患有心肌缺血疾病的患者的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。最近今年被广泛使用的十二导联心电数据库包括包括2018年中国生理信号挑战赛的CPSC2018心律失常数据库、圣彼得堡INCART心律失常数据库、Physikalisch Technische Bundesanstalt (PTB) 心律失常数据库、乔治亚州心律失常数据库等，这些数据库均可从PhysioNet Challeng[37,38]获取。以下详细介绍这些数据库。

1. CSPC2018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[39]。共计9831组12导联心电图样本，样本持续时间从6秒到144秒不等，其中6877个（女性：3178个和男性：3699个）样本被公开用于训练，2954个样本未被开源用于评价测试。每组样本包含12导联心电图信号、患者基本信息（年龄和性别）以及标注，其中标注对应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)、室性质早搏 (PVC)、ST段压低 (STD) 和 ST段抬高 (STE)。不同样本给定的标注数量不同，大多数样本只有一个标注，少数样本存在两个或者三个标注。
2. CPSC2018-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”的比赛中被公开使用。
3. 乔治亚州心律失常数据库：乔治亚州心律失常数据库是来自美国东南部的一组心电数据，该数据集包含10344（男性：5551，女性：4793）组十二导联心电图样本，每条记录长度为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 12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，每条样本记录长30分钟，信号采样率为257Hz；
5. PTB心律失常数据库：Physikalisch Technische Bundesanstalt (PTB) 心律失常数据库是一个大型公开可用的心电图数据集[40]。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。

这五个数据库大约包含有27类心律失常的标注，且标注基于样本给出，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自动分类心律失常算法提供资源。

## 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。算力的不断提升、科技化社会的到来、各种各样的数据源都促进了深度学习的优势在于能够自动学习和提取例如结构化数据以及例如视频、音频、图片等非结构数据的特征，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。

### 2.3.1 深度学习的来源和发展

提到深度学习，脑海中就会浮现与它密切相关的几个词例如人工智能、机器学习、神经网络等。现实生活中作为普通人很难直接感受到这些听起来晦涩难懂的名词有什么作用，但实际上这部分知识正在催促着整个社会迈向虚拟化和科技化。

最早出现的名词是人工智能，目前我们正在经历人工智能的第三次发展热潮。1956年在达特茅斯会议中正式将这类能够代替人力思考和劳动的技术命名为人工智能，经历一些年的研究发展，人工智能技术确实帮助企业解决了部分简单问题，但受制于能够解决的问题范围窄，一旦碰到需求复杂度高的项目便无法研究下去，这使得致力于发展人工智能的国家停止对人工智能的投资，人工智能的第一次发展高潮停滞了。随着专家系统的问世，带有学者主观判断逻辑的人工智能系统迅速开启了第二波浪潮，同时第四代计算机系统也在这个阶段出现，第四代计算的发展吸引了空前的目光，人工智能的研究市场被抢占，间接导致人工智能第二波发展高潮的没落。随着计算机时代的来临，计算机硬件软件发展趋于稳定，算力的不断提升使得研究人员重新审视人工智能，各行各业源源不断的产生数据，如何利用好这些数据，进行数据挖掘和大数据分析结合人工智能助力实现高效快捷的企业增长和创新型社会成为主流。近些年来技术的不断突破、算法的层层优化，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。

而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）。例如在浏览淘宝过程中，当我们点击某类产品进行浏览时，客户端会收集这些信息发送到服务器，服务器经过一系列的特征总结，将用户喜好反馈到客户端，客户端根据返回的信息展示更多相关的商品信息给用户，从而提升交易成功率。而这个算法就融合了基于机器学习的推荐算法。

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进行推演。神经网络的发展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为的深度神经网络，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个输入M个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。

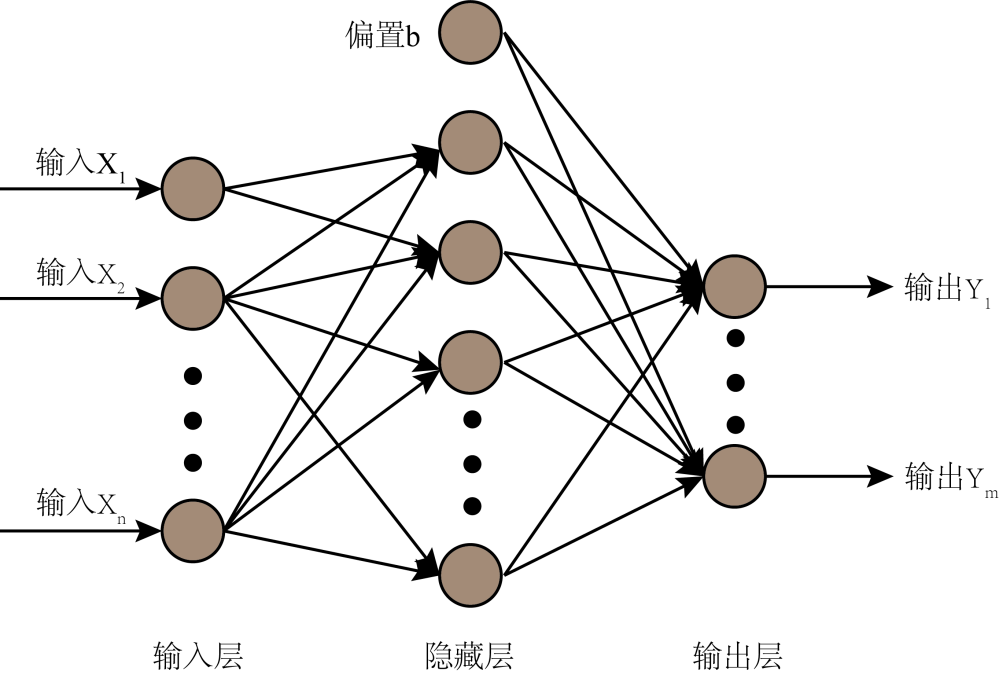


图2.5 三层神经网络结构图

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。在结构上具有更大容量，分层组合带来的是指数级表达空间，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。

### 2.3.2 深度学习的原理

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，分析局部在规划整理全局。由于卷积神经网络具有“局部感受野”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，不同的是神经网络隐藏层数量较多，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、参数调整等。

1. 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，而且根据历史数据是可以发现数据之间大多不存在线性关系，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域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尽可能的简单，提交计算效率；激活函数的导函数的值域在合理的区间，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,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，带式Sigmoid的梯度趋近于0，直接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[41]。Sigmoid激活函数公式如公式（2.1）所示，相比其他激活函数计算成本较高。

（2.1）

Tanh激活函数，它将输入的值映射在区间[-1,1]上，是对Sigmoid的增强，引进了负输入依旧被映射为负数，接近0的输入被映射在接近0的位置。但由于函数曲线类似于Sigmoid[42]，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：

（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[43]。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：

（2.3）

常见激活函数的函数图像如图2.6所示。

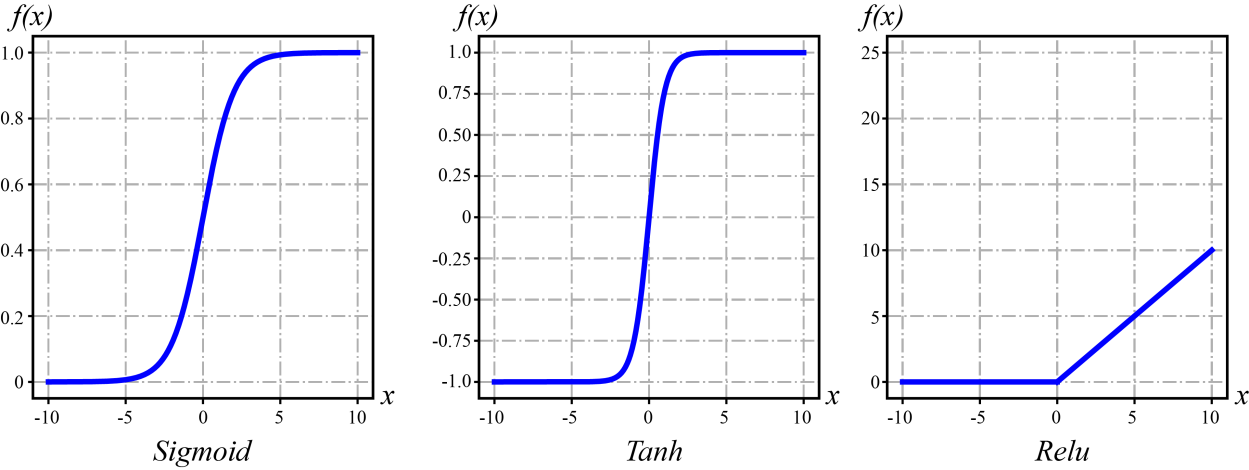


图2.6 常见激活函数图

1. 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：

（2.4）

其中*w*代表权重参数，*b*代表偏置参数，N代表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，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。因此模型得到正类负类的概率分别为和。

因此N个输入的情况下的条件概率公式为：

（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：

（2.6）

式中*i*代表输入第几个，N代表输入个数，代表真实值，代表预测概率。

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不是具体的数，而变成了一组One-hot向量[44]，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：

（2.7）

式中N代表样本个数，M代表类别个数，代表One-hot向量，代表预测概率向量。

1. 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*，则随机梯度下降的参数更新表达式如下：

（2.8）

式中*w*代表权重参数，*b*代表偏置，*η*表示学习率。

## 2.5本章小结

本章第一节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。

然后针对心律失常的检测引出十二导联心电图的概念，介绍了十二导联心电图的起源和发展、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二导联心电图数据库以及十二导联心电图数据库。最后我们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。通过这一章的介绍，我们可以初步理解应用深度学习在心电图领域的重要性，为后续研究的详细阐述提供理论依据。

# 

# 第3章 基金-股票网络映射模型

## 3.1 引言

复杂

## 3.2 建立基金-股票网络映射模型

我

## 3.3 结果

### 3.3.1 基金的投资水平

为

### 3.3.2 基金的投资分布

为

### 3.3.3 基金的投资倾向

### 3.3.4 投资策略

季

## 3.4 附录

### 3.4.1 基金-股票网络映射模型

为

### 3.4.2 改写的上证50(180)指数

## 3.5 本章小结

为

# 

# 第4章 基于基金-股票网络映射模型的社团检测算法

## 4.1 引言

目

## 4.2 基于加权股票网络的社团检测

### 4.2.1 加权股票网络

## 4.3 结果

### 4.3.1 社团分布

### 4.3.2 股票网络关键社团的演化

## 4.4 本章小结

本

# 第5章 总结与展望

## 5.1 全文工作总结

本

## 5.2 未来研究展望

结

# 参考文献

1. Holst, H., Ohlsson, M., Peterson, C. and Edenbrandt, L., 1999. A confident decision support system for interpreting electrocardiograms. Clinical Physiology, 19(5), pp.410-418.
2. Shah, A.P. and Rubin, S.A., 2007. Errors in the computerized electrocardiogram interpretation of cardiac rhythm. Journal of electrocardiology, 40(5), pp.385-390.
3. 马丽媛,王增武,樊静,胡盛寿.《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. and Wang, S., 2020. Automated detection of atrial fibrillation in ECG signals based on wavelet packet transform and correlation function of random process. Biomedical Signal Processing and Control, 55, p.101662.
5. Tsipouras, M.G., Fotiadis, D.I. and Sideris, D., 2005. An arrhythmia classification system based on the RR-interval signal. Artificial intelligence in medicine, 33(3), pp.237-250.
6. Haseena, H.H., Mathew, A.T. and Paul, J.K., 2011. Fuzzy clustered probabilistic and multi layered feed forward neural networks for electrocardiogram arrhythmia classification. Journal of Medical Systems, 35(2), pp.179-188.
7. Mondéjar-Guerra, V., Novo, J., Rouco, J., Penedo, M.G. and Ortega, M., 2019. Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers. Biomedical Signal Processing and Control, 47, pp.41-48.
8. Afkhami, R.G., Azarnia, G. and Tinati, M.A., 2016. Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals. Pattern Recognition Letters, 70, pp.45-51.
9. Martis, R.J., Acharya, U.R., Prasad, H., Chua, C.K., Lim, C.M. and Suri, J.S., 2013. Application of higher order statistics for atrial arrhythmia classification. Biomedical signal processing and control, 8(6), pp.888-900.
10. de Oliveira, L.S., Andreão, R.V. and Sarcinelli-Filho, M., 2011, August. Premature ventricular beat classification using a dynamic Bayesian network. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 4984-4987). IEEE.
11. De Chazal, P., O'Dwyer, M. and Reilly, R.B., 2004. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE transactions on biomedical engineering, 51(7), pp.1196-1206.
12. Ince, T., Kiranyaz, S. and Gabbouj, M., 2009. A generic and robust system for automated patient-specific classification of ECG signals. IEEE Transactions on Biomedical Engineering, 56(5), pp.1415-1426.
13. Wang, J.S., Chiang, W.C., Hsu, Y.L. and Yang, Y.T.C., 2013. ECG arrhythmia classification using a probabilistic neural network with a feature reduction method. Neurocomputing, 116, pp.38-45.
14. Martis, R.J., Acharya, U.R. and Min, L.C., 2013. ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. Biomedical Signal Processing and Control, 8(5), pp.437-448.
15. Martis, R.J., Acharya, U.R., Prasad, H., Chua, C.K. and Lim, C.M., 2013. Automated detection of atrial fibrillation using Bayesian paradigm. Knowledge-Based Systems, 54, pp.269-275.
16. Ye, C., Kumar, B.V. and Coimbra, M.T., 2012. Heartbeat classification using morphological and dynamic features of ECG signals. IEEE Transactions on Biomedical Engineering, 59(10), pp.2930-2941.
17. Osowski, S., Hoai, L.T. and Markiewicz, T., 2004. Support vector machine-based expert system for reliable heartbeat recognition. IEEE transactions on biomedical engineering, 51(4), pp.582-589.
18. Lagerholm, M., Peterson, C., Braccini, G., Edenbrandt, L. and Sornmo, L., 2000. Clustering ECG complexes using Hermite functions and self-organizing maps. IEEE Transactions on Biomedical Engineering, 47(7), pp.838-848.
19. Guo, G., Wang, H., Bell, D., Bi, Y. and Greer, K., 2003, November. KNN model-based approach in classification. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems" (pp. 986-996). Springer, Berlin, Heidelberg.
20. Ullah, A., Anwar, S.M., Bilal, M. and Mehmood, R.M., 2020. Classification of arrhythmia by using deep learning with 2-D ECG spectral image representation. Remote Sensing, 12(10), p.1685.
21. Jun, T.J., Nguyen, H.M., Kang, D., Kim, D., Kim, D. and Kim, Y.H., 2018. ECG arrhythmia classification using a 2-D convolutional neural network. arXiv preprint arXiv:1804.06812.
22. Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P. and Ng, A.Y., 2019. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature medicine, 25(1), pp.65-69.
23. Wang, R., Fan, J. and Li, Y., 2020. Deep multi-scale fusion neural network for multi-class arrhythmia detection. IEEE journal of biomedical and health informatics, 24(9), pp.2461-2472.
24. Chen, T.M., Huang, C.H., Shih, E.S., Hu, Y.F. and Hwang, M.J., 2020. Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model. Iscience, 23(3), p.100886.
25. Zhang, D., Yang, S., Yuan, X. and Zhang, P., 2021. Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram. Iscience, 24(4), p.102373.
26. Zhao, Z., Fang, H., Relton, S.D., Yan, R., Liu, Y., Li, Z., Qin, J. and Wong, D.C., 2020, September. Adaptive lead weighted ResNet trained with different duration signals for classifying 12-lead ECGs. In 2020 Computing in Cardiology (pp. 1-4). IEEE.
27. 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：CN110179453B,2020-01-03.
28. Chiou, Y.A., Syu, J.Y., Wu, S.Y., Lin, L.Y., Yi, L.T., Lin, T.T. and Lin, S.F., 2021. Electrocardiogram lead selection for intelligent screening of patients with systolic heart failure. Scientific reports, 11(1), pp.1-12.
29. Hong, S., Wu, M., Zhou, Y., Wang, Q., Shang, J., Li, H. and Xie, J., 2017, September. ENCASE: An ENsemble ClASsifiEr for ECG classification using expert features and deep neural networks. In 2017 Computing in cardiology (CinC) (pp. 1-4). IEEE.
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.DOI:10.16476/j.pibb.2015.0302.
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.DOI:10.13308/j.issn.2095-9354.2015.05.012.
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.DOI:10.13308/j.issn.2095-9354.2017.04.006.
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, Distante 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 MA, Sadr N, Perez Alday EA, Gu A, Shah AJ, Robichaux C, Rad AB, Elola A, Seyedi S, Ansari S, Ghanbari H, Li Q, Sharma A, Clifford GD. Will Two Do? Varying Dimensions in Electrocardiography: The PhysioNet/Computing in Cardiology Challenge 2021. Computing in Cardiology 2021; 48: 1-4,
38. Reyna MA, Sadr N, Perez Alday EA, Gu Annie, Shah AJ, Robichaux C, Rad AB, Elola A, Seyedi S, Ansari S, Ghanbari H, Li Q, Sharma A, Clifford GD. Issues in the automated classification of multilead ECGs using heterogeneous labels and populations. Physiol. Meas. 2022.
39. F. F. Liu, C. Y. Liu\*, L. N. Zhao, X. Y. Zhang, X. L. Wu, X. Y. Xu, Y. L. Liu, C. Y. Ma, S. S. Wei, Z. Q. He, J. Q. Li and N. Y. Kwee. An open access database for evaluating the algorithms of ECG rhythm and morphology abnormal detection. 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.

# 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，写这段话的过程中眼睛甚至被打湿。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，一心想要为建设家乡做出自己的贡献。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。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，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。让我记忆犹新的是，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。我相信我未来一定能够成为一个恪尽职守，尽职尽责的人。

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. ChuanzheZhang,ShaopengPang,HaoYu,GuozhengHan，A fund-stock network projection model，Physica A，566（2021)，125630.（SCI收录，影响因子，2.243，已录用）（第一作者）

## 专利相关

## 参与项目

## 四、参与竞赛