

分类号:
学号: 20222108039

密级: 公开
单位代码: 10759

石河子大学

硕士学位论文



基于多模态医学数据的冠心病诊断研究 与系统实现

学位申请人	代云飞
指导教师	戴建国 教授
申请学位类别	专业硕士
专业名称	电子信息
研究领域	计算机技术
所在学院	信息科学与技术学院

中国·新疆·石河子
2025年6月

分类号:
学号: 20222108039

密级: 公开
单位代码: 10759

石河子大学

硕士学位论文



基于多模态医学数据的冠心病诊断研究 与系统实现

学位申请人	代云飞
指导教师	戴建国 教授
申请学位类别	专业硕士
专业名称	电子信息
研究领域	计算机技术
所在学院	信息科学与技术学院

中国·新疆·石河子
2025年6月

**Research on Coronary Artery Disease Diagnosis Based on
Multimodal Medical Data and System Implementation**

A Dissertation Submitted to

Shihezi University

In Partial Fulfillment of the Requirements

for the Degree of

Master of Engineering

By

Dai Yun-Fei

(Electronic Information)

Dissertation Supervisor: Prof. Dai Jian-Guo

June, 2025

石河子大学学位论文独创性声明及使用授权声明

学位论文独创性声明

本人所提交的学位论文是在我导师的指导下进行的研究工作及取得的研究成果。据我所知，除文中已经注明引用的内容外，本论文不包含其他个人已经发表或撰写过的研究成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确的说明并表示谢意。

研究生签名：

代云飞

时间： 2025 年 05 月 21 日

使用授权声明

本人完全了解石河子大学有关保留、使用学位论文的规定，学校有权保留学位论文并向国家主管部门或指定机构送交论文的电子版和纸质版。有权将学位论文在学校图书馆保存并允许被查阅。有权自行或许可他人将学位论文编入有关数据库提供检索服务。有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

研究生签名：

代云飞

时间： 2025 年 05 月 21 日

导师签名：

戴建国

时间： 2025 年 05 月 21 日

摘要

冠心病是中国主要的致死疾病之一，其发病机制涉及动脉内壁斑块沉积导致心肌缺血，若未及时干预，可能引发心源性猝死。因此，冠心病的早期诊断能够协助制定及时的治疗方案，从而显著提升患者的生存率。心音信号反映了心脏瓣膜运动和血流动力学变化，心电图记录了心脏电活动，基线资料涵盖患者的年龄、性别和既往病史等信息，共同为冠心病的诊断提供重要依据。因此，本研究探讨心音信号、12 导联动态心电图和基线资料的深度融合方法，构建基于多模态医学数据的冠心病诊断模型（Hierarchical Multimodal Synergy Network, HMSN），以无创、准确地诊断冠心病，并为临床医生提供高效的辅助决策支持。本文具体工作如下：

（1）提出了一种直接基于一维心音信号进行冠心病诊断的方法，以构建的多尺度注意力残差网络（Multi-Scale Attention Residual Network, MSARN）为基础，将 126 个医学多域特征（Medical Multi-Domain Feature, MMDF）和 206 个多域特征（Multi-Domain Feature, MDF）分别融入 MSARN 模型，构建了 MSARN-MMDF 和 MSARN-MDF 模型。实验结果表明，MSARN 模型避免了心音信号转为二维图像的信息损失，直接在一维信号上进行特征提取，准确率达到 83.45%。基于此，融合医学多域及多域特征的 MSARN-MMDF 和 MSARN-MDF 模型相较于基础 MSARN 模型，准确率分别提升了 2.15% 和 4.41%，这一结果验证了融合一维心音深度特征与多域特征在冠心病诊断中的有效性。

（2）在心音信号分析的基础上，引入 12 导联动态心电图与基线资料，以探索多模态数据在冠心病诊断中的协同作用和诊断价值。针对基线资料，使用随机森林算法筛选出 40 个关键特征，并基于自注意力编码器框架构建冠心病诊断模型，最终获得 81.01% 的准确率。针对 12 导联动态心电图，转为一维心电信号后，使用融合时间-空间联合特征的自适应神经网络的冠心病诊断模型分别提取局部空间结构特征与动态时间变化特征，达到 80.00% 的准确率。在单模态研究的基础上，构建了基于多模态医学数据的冠心病诊断模型 HMSN，结合迁移学习策略，进一步挖掘心音信号、12 导联动态心电图及基线资料数据之间的潜在关联性与互补性。实验结果表明，相较于基于单一模态及两模态数据的冠心病诊断模型，HMSN 三模态模型效果最好，准确率达到 90.00%。

（3）设计并开发了一套基于 Python 和 PyQt 的交互式冠心病辅助诊断系统。该系统支持对心音信号、12 导联动态心电图和基线资料的数据编辑、可视化展示及多模态数据的预处理，能够基于单模态和多模态数据进行冠心病诊断，为临床医生提供准确的辅助决策支持工具，显著提升了诊断效率和准确性。

关键词：深度学习；心音信号；12 导联动态心电图；基线资料；冠心病诊断

Abstract

Coronary artery disease (CAD) is one of the leading causes of death in China. Its pathogenesis involves plaque deposition on the inner walls of arteries, leading to myocardial ischemia, which, if not treated in time, may result in sudden cardiac death. Therefore, early diagnosis of CAD can assist in formulating timely treatment plans, significantly improving patient survival rates. Heart sound signals reflect heart valve motion and hemodynamic changes, electrocardiograms (ECGs) record the electrical activity of the heart, and baseline data include patient information such as age, gender, and medical history, all of which provide critical diagnostic evidence for CAD. Therefore, this study explores a deep fusion approach integrating heart sound signals, 12-lead ambulatory ECGs, and baseline data to construct a CAD diagnostic model based on multimodal medical data, the hierarchical multimodal synergy network (HMSN), to enable non-invasive and accurate CAD diagnosis while providing efficient auxiliary decision support for clinicians. The specific contributions of this study are as follows:

(1) A method for directly diagnosing CAD using one-dimensional heart sound signals is proposed, based on the constructed multi-scale attention residual network (MSARN). The MSARN model avoids information loss caused by converting heart sound signals into two-dimensional images and directly extracts features from one-dimensional signals, achieving an accuracy of 83.45%. Building on this, the integration of 126 medical multi-domain features (MMDF) and 206 multi-domain features (MDF) into the MSARN model led to the development of the MSARN-MMDF and MSARN-MDF models. The experimental results demonstrate that, compared to the base MSARN model, the accuracies of the MSARN-MMDF and MSARN-MDF models improved by 2.15% and 4.41%, respectively, validating the effectiveness of incorporating deep one-dimensional heart sound features with multi-domain features for CAD diagnosis.

(2) Based on heart sound signal analysis, 12-lead ambulatory ECGs and baseline data were introduced to explore the synergy and diagnostic value of multimodal data in CAD diagnosis. For baseline data, 40 key features were selected using the random forest algorithm, and a CAD diagnostic model was constructed based on a self-attention encoder framework, ultimately achieving an accuracy of 81.01%. For 12-lead ambulatory ECGs, after converting them into one-dimensional ECG signals, a CAD diagnostic model incorporating a spatio-temporal adaptive neural network was employed to extract both local spatial structure features and dynamic temporal variation features, achieving an accuracy of 80.00%.

Building upon unimodal research, a multimodal medical data-based CAD diagnostic model, HMSN, was developed. By integrating a transfer learning strategy, the model further explored the potential correlations and complementarities among heart sound signals, 12-lead ambulatory ECGs, and baseline data. Experimental results indicate that, compared with CAD diagnostic models based on single-modality and dual-modality data, the HMSN three-modality model achieved the best performance, reaching an accuracy of 90.00%.

(3) An interactive CAD-assisted diagnosis system was designed and developed using Python and PyQt. This system supports data editing, visualization, and multimodal data preprocessing for heart sound signals, 12-lead ambulatory ECGs, and baseline data. It enables CAD diagnosis based on both unimodal and multimodal data, providing clinicians with an accurate auxiliary decision-support tool that significantly improves diagnostic efficiency and accuracy.

Key words: Deep learning; Heart sound signals; 12-lead ambulatory electrocardiogram; Baseline data; Coronary artery disease diagnosis

目录

摘要	I
Abstract	II
目录	IV
第 1 章 绪论	1
1.1 研究背景及意义	1
1.1.1 研究背景	1
1.1.2 研究意义	2
1.2 国内外研究现状	3
1.2.1 基于心音的心血管疾病诊断研究现状	3
1.2.2 基于心电的心血管疾病诊断研究现状	4
1.2.3 基于多模态数据的心血管疾病诊断研究现状	5
1.2.4 研究现状评述	6
1.3 研究内容与技术路线	6
1.3.1 研究目标	6
1.3.2 研究内容	6
1.3.3 技术路线	7
1.4 论文组织结构	8
第 2 章 相关技术理论	10
2.1 预处理相关技术	10
2.1.1 小波降噪技术	10
2.1.2 巴特沃斯带通滤波器技术	11
2.2 卷积神经网络	12
2.3 长短时记忆网络	13
2.4 SHAP 理论	14
2.5 迁移学习技术	15
2.6 模型评价指标	16
2.7 本章小结	16
第 3 章 基于心音的冠心病诊断研究	17
3.1 引言	17

3.2 数据获取与预处理	17
3.2.1 数据集概况	17
3.2.2 研究人群	18
3.2.3 数据增强与心音信号预处理	19
3.3 心动周期切割	20
3.4 MMDF 和 MDF 提取	21
3.5 模型构建	26
3.5.1 MSARN 模型	26
3.5.2 总体模型结构	29
3.5.3 实验平台	30
3.6 实验结果与讨论分析	30
3.6.1 基准测试	30
3.6.2 消融实验	32
3.6.3 基于 SHAP 算法的模型可解释性分析	35
3.6.4 对比分析	36
3.7 本章小结	38
第 4 章 基于多模态医学数据的冠心病诊断研究	39
4.1 引言	39
4.2 数据获取	39
4.2.1 基线资料获取	39
4.2.2 心电信号获取	40
4.3 数据预处理	41
4.3.1 基线资料预处理	41
4.3.2 心音、心电信号预处理	42
4.4 基于基线资料的冠心病诊断方法	42
4.4.1 基线资料特征筛选	42
4.4.2 模型构建	44
4.4.3 实验结果与讨论分析	45
4.5 基于心电信号的冠心病诊断方法	45
4.5.1 STANN 模型	45
4.5.2 实验结果与讨论分析	47
4.6 基于多模态数据的冠心病诊断方法	49
4.6.1 模型构建	49
4.6.2 迁移学习策略	50

4.6.3 实验结果与讨论分析	50
4.7 基于 SHAP 算法的模型可解释性分析	52
4.8 本章小结	54
第 5 章 冠心病辅助诊断系统实现	56
5.1 系统需求分析	56
5.1.1 功能需求分析	56
5.1.2 非功能需求分析	57
5.2 系统设计	57
5.2.1 系统整体架构	57
5.2.2 系统详细设计	58
5.3 系统实现	60
5.3.1 系统环境	60
5.3.2 系统展示	60
5.4 本章小结	63
第 6 章 总结与展望	64
6.1 总结	64
6.2 展望	65
参考文献	66
致谢	72
作者简介	73

第 1 章 绪论

1.1 研究背景及意义

1.1.1 研究背景

随着科技和社会经济的快速发展，世界人民生活水平在得到显著提升的同时，生活方式也发生了巨大变革。然而，这种变革间接导致了不健康生活方式的盛行，包括饮食结构失衡、运动量不足以及吸烟等，增加了患心血管疾病的风险。根据世界卫生组织的统计，心血管疾病已成为全球主要致死性疾病之一^[1]。在 2019 年，全球因心血管疾病死亡的人数达到 1770 万，占总死亡人数的 32%。与此同时，根据《中国心血管健康与疾病报告 2023》流行病学调查^[2]，随着人口老龄化进程的加速和不健康生活方式的普及，心血管疾病的患病率在我国逐年上升，死亡率也呈现上升趋势，已成为我国目前患病率最高的疾病之一。

冠心病（Coronary Artery Disease, CAD）是心血管疾病中最常见的一种类型^[3]，发病机制通常始于动脉内壁的脂质沉积，形成动脉粥样斑块。根据病情的稳定性，CAD 可分为稳定性冠心病（Stable Coronary Artery Disease, SCAD）和不稳定性冠心病（Unstable Coronary Artery Disease, UCAD）。SCAD 通常表现为冠状动脉的轻度或中度狭窄，血管通畅性较好，症状较为可预测，患者在静息状态下通常无明显不适。SCAD 的症状多在运动或情绪激动时出现，且疼痛较短暂，休息或药物治疗后可缓解。然而，SCAD 可能随着时间推移逐渐恶化，冠状动脉的狭窄程度可能加重，导致心肌缺血更加严重，进而发展为缺血性心脏病或急性心肌梗死。一旦 SCAD 演变为 UCAD 或急性心肌梗死，症状可能突发且剧烈，常常不易通过常规治疗缓解，迅速威胁患者生命。因此，定期进行 CAD 检测至关重要，能够及时发现早期病变。通过早期诊断，可以制定合理的治疗方案，降低疾病恶化的风险，从而提高患者的生存率。

根据最新统计^[2]，中国心血管疾病的患者总数已达到 3.3 亿，其中 CAD 患者约为 1139 万。目前，冠状动脉造影被广泛认可为诊断 CAD 的金标准。该技术涉及在患者冠状动脉内注入对比剂，随后运用 X 射线成像技术来精确评估血管的狭窄程度。通常情况下，当冠状动脉血管狭窄程度达到 50% 以上时，就可被诊断为 CAD。然而，这种方法属于有创性操作^[4]，不仅需要较高水平的医疗设施和专业技术支持，还会对患者

带来一定的侵入性风险，同时设备成本和就医费用较为昂贵。这些因素使得冠状动脉造影的广泛应用受到限制，尤其是在医疗资源匮乏的基层地区。综上所述，目前的 CAD 诊断方法在准确性、便捷性以及普适性方面存在不足。因此，迫切需要一种能够快速、低成本、无创并且易于推广的检测方法，来满足早期筛查和诊断 CAD 的需求。

1.1.2 研究意义

有研究表明^[5]，CAD 引起的冠脉狭窄能够导致心音和心电图的异常变化，进而带来临床表现上的差异。通过分析这些变化，可以为 CAD 的早期诊断提供重要依据。心音信号是心脏在血液循环过程中产生的声音，主要由心脏瓣膜开闭引起的振动以及血液流动的紊乱产生，能够反映心脏功能状态。在正常情况下，心音^[6]主要由第一心音和第二心音两个基本成分组成。然而，对于 CAD 患者，由于冠状动脉狭窄导致心肌供血不足，常会引发心音异常，如杂音的产生或心音时序的紊乱。心电图^[7]是通过电极记录心脏在跳动过程中产生的电活动的波形图，广泛用于心血管疾病的诊断。对于 CAD 患者，心电图可能出现如 ST 段抬高或压低等特征性变化，是判断 CAD 的重要依据。基线资料^[8]是患者在入院时采集的基础信息，涵盖了患者的个人特征、既往病史和生活习惯等多方面内容。这些资料包括但不限于性别、年龄、身高、体重、血压、血脂水平、过往病史（如高血压、糖尿病等）、家族病史等。基线资料为临床医生提供患者整体健康状况的重要背景信息，是诊断和治疗方案制定的重要依据。

尽管心音信号、心电图和基线资料在 CAD 诊断中具有重要价值，但传统诊断方式存在显著局限。由于心音信号的复杂性及其特征的微弱变化，仅依靠人工听诊易受医生主观经验影响，可能导致漏诊或误诊。心电图虽然是心脏疾病检测的常见手段，但其波形分析需要医生具备丰富的专业经验，仍存在误诊和漏诊的风险。基线资料为 CAD 的风险评估提供了重要依据，但传统的人工处理和分析方法效率低且容易出错。近年来，计算机辅助诊断技术经历了迅猛的发展，尤其在深度学习技术在医疗领域得到广泛应用的背景下，其在疾病检测方面的高效性和准确性已经得到了充分的实证支持。利用计算机自动化处理技术，可以显著提高数据分析的效率，而深度学习算法则通过多层非线性模型从复杂的医疗数据中提取关键特征，为疾病的精确诊断提供强有力的支撑。

综上所述，本研究以冠心病患者的心音信号、12 导联动态心电图和基线资料为研究对象，系统地开展了基于深度学习、信号处理及多模态特征融合的冠心病诊断研究，并开发了一套冠心病辅助诊断系统。该系统能够实现对 CAD 的无创、快速、低成本诊断，显著提高了诊断的准确性和一致性。通过早期诊断与病变识别，该系统有助于预防冠心病向更严重的心血管疾病发展，降低疾病恶化的风险，从而提高患者的生存

率。此外，系统为临床医生提供了高效的辅助决策支持，优化了诊断流程，显著减轻了医生的工作负担。该系统的良好适应性和普及性使其在医疗资源匮乏地区具有显著应用优势，能够降低诊断成本，惠及更多患者，推动无创检测技术的发展。

1.2 国内外研究现状

1.2.1 基于心音的心血管疾病诊断研究现状

心音是心脏及心血管系统运动时所发出的声音，其中包含着心血管系统各个部分相互之间的生理和病理信息。通过对心音定量采集与分析，提取关键参数，将患者的监测序列与标记数据库进行比较，不仅可以自动获得更直观的诊断^[6]结果，而且专家可以利用临床知识进一步推断出潜在的心血管疾病。基于心音的心血管疾病诊断研究始于 20 世纪 90 年代，Wang 等人^[9]使用神经网络和时间分析进行心音异常检测，此后的研究进一步借助心音数据提取手工设计特征，作为不同机器学习分类器的输入。例如，Aziz^[10]等人采用经验模态分解^[11]对心音信号进行去噪处理^{[12][13][14]}，提取时域和频域特征，并使用支持向量机等机器学习方法进行分类；Yadav 等人^[15]从时域和频域提取过零率、能量熵等手工设计特征，选择最显著特征输入到支持向量机、朴素贝叶斯、k 近邻等机器学习分类器；Sawant 等人^[16]使用 SMOTE 填充^[17]进行数据平衡^{[18][19][20]}，从时频域中提取特征，放入 LightGBM 分类器；Abduh 等人^[21]根据梅尔倒谱系数（Mel-Frequency Cepstral Coefficients, MFCCs）提出了改进的分数 MFCCs 特征，采用 k 近邻法、集成分类器等方法进行分类。上述研究表明，使用手工设计特征的检测方法依赖于不同域特征的组合，从而尽可能地揭示与疾病相关的特征。然而，手工设计特征可能无法充分捕获心音的所有关键信息，从而限制检测模型的性能，这促使研究者寻求更自动化、高效的特征提取技术。

随着技术的快速发展，深度学习^[22]在医学领域得到了广泛应用，其通过多层非线性建模，自动从复杂医疗数据中提取关键特征，并逐层抽象形成高级语义特征，从而实现自动诊断。在基于深度学习的心血管疾病诊断研究^[23]中，主要采用了卷积神经网络（Convolutional Neural Network, CNN）。CNN 通过局部感受野、权重共享等机制，能够有效提取数据中的局部特征，因此研究人员已将其广泛应用于基于心音的心血管疾病诊断领域。通常，经过预处理后的心音数据是一维时序数据，因此，研究人员常将其转换为特征图^[24]作为模型的输入。例如，Wang 等人^[25]将一维心音序列转为梅尔图、对数梅尔图、短时傅立叶变换图；Dhar 等人^[13]将一维心音序列转为交叉小波图；Zhou 等人^[26]将一维心音信号转换为二维 MFCC 特征图进行分类，分别取得 94.1%的

敏感性、88.7%的特异性；Chen 等人^[27]将心音信号转换为梅尔图和对数梅尔图，取得71.4%的敏感性、95.1%的特异性；Xiang 等人^[24]将心音信号转换成多种二维特征图形式，包括包络图、时域波形图、对数梅尔图和对数功率谱图，在 InceptionV3 模型中的准确率分别达到 84.9%、86.6%、93.4%和 93.8%。在转换过程中，研究者们深入探讨了不同域特征图对诊断任务性能的影响。例如，Wang 等人^[25]基于公开数据集 PhysioNet，系统评估了多种时频特征图对 CNN 诊断模型性能的影响。结果显示，将短时傅里叶变换、对数梅尔变换、希尔伯特-黄变换、小波变换、梅尔变换和斯托克韦尔变换特征图用于所设计模型的输入时，斯托克韦尔变换特征图在综合性能评估中表现最优，其综合得分达到 65.2，显著优于其他时频特征图。Chen 等人^[28]进一步研究了梅尔图和对数梅尔图在基于心音的心血管疾病诊断中的性能差异。结果显示，对数梅尔图、梅尔图的平均准确率为 91.74%、87.42%，这表明了对数梅尔图在提高诊断模型性能方面的有效性。同时 Yin 等人^[29]在小规模数据集下也表明了梅尔图在冠心病无创诊断中的可用性。此外，研究人员也在其他领域进行了更多研究，如 Er^[30]创新地将局部二值模式和局部三值模式应用于一维心音信号的特征提取中，通过组合两种特征，获得了信息量更丰富的特征表示。与此同时，深度学习与信号处理技术的结合也成为研究热点。Fan 等人^[31]提出了一种名为“可学习的提升小波变换块”的方法，该方法将可训练的 CNN 嵌入到提升小波变换框架中，既保留了小波变换的多分辨率时频分析能力，又充分发挥了 CNN 的非线性特征学习优势，为信号处理与深度学习的融合提供了新的思路。上述研究表明，高质量的时频图有利于疾病检测。然而，不同二维时频图在捕获原始心音信号特征的能力上存在显著差异。此外，二维时频图像的转换过程通常依赖于经验判断，可能导致关键信息的丢失，从而影响模型性能。

1.2.2 基于心电图的心血管疾病诊断研究现状

在早期基于心电图（Electrocardiograms, ECG）信号的心血管疾病诊断研究中，研究者通常采用信号分解与特征工程的策略，将 ECG 信号转化为时域、频域或时频域特征，再结合传统机器学习算法进行分类。例如，Deng 等人^[32]通过提取 ECG 信号中的 ST 段特征，采用统计方法实现了 84.6%的准确率；Kaveh 等人^[33]从 ECG 信号中提取了时域、频域，利用支持向量机等分类器实现了 88%的准确度；Lee 等人^[34]提取时域、频域和非线性特征，使用支持向量机、决策树获得了 90%的准确率。尽管这些方法在一定程度上推动了 ECG 自动化诊断的发展，但其局限性亦日益显著：ECG 波形存在显著的个体差异和瞬时变化特性，传统特征工程难以捕捉复杂的时空关联性，且在噪声干扰下鲁棒性不足^[35]。

近年来，深度学习技术通过端到端的特征学习和模式识别能力，显著提升了基于

心电的心血管疾病诊断的准确性与泛化性。Rajpurkar 等人^[36]构建了 34 层二维 CNN 模型,在心律失常的自建数据集中实现了 77.6%的分类准确率;Lai 等人^[37]采用一维 CNN 结构,通过原始 12 导联心电图和 F 波频谱,在房颤检测任务中达到 93.1%的准确率。然而,单一 CNN 模型对长程时序依赖性的建模能力有限,研究者开始探索结合循环神经网络的混合架构。例如,Yao 等人^[38]提出了一种基于时空特征融合的深度神经网络,通过注意力机制自适应评估不同长度的 ECG 信号的重要性,在包含房颤、传导阻滞和 ST 段异常的 9 类心律失常检测中达到 81.2%的准确率,显著优于传统特征工程方法。随着算法创新,深度学习在复杂心血管疾病的早期预警与精准诊断中展现出更广阔的应用前景。例如,针对肿瘤治疗相关心功能不全的早期筛查,Jacobs 等人^[39]开发了基于 AI-ECG 的预测模型,通过分析乳腺癌患者化疗前后的 ECG 信号,成功预测射血分数低于 50%和 35%的异常状态,AUC 分别达到 0.93 和 0.94。上述研究表明,深度学习与 ECG 的结合,可作为无创、低成本的心脏异常筛查工具,辅助临床决策并优化随访策略。

1.2.3 基于多模态数据的心血管疾病诊断研究现状

近年来,多模态数据融合技术的快速发展为心血管疾病的早期诊断和精准检测提供了新的思路。传统基于单模态数据的分析方法虽然能够捕捉特定维度的病理特征,但受限于个体生理差异和疾病复杂性,其诊断灵敏度和特异性常难以满足临床需求。因此,整合多模态数据^{[40][41]}的互补信息已成为提升诊断效能的关键路径,其技术路线主要分为特征级融合与模型级融合两类。

特征融合是指在特征提取阶段,将来自不同模态的数据特征进行整合,综合各模态之间的关联性与互补性,从而构建更丰富的特征表示。这一方法通过直接操作特征向量,能够在模型输入阶段实现信息的高度共享,常见技术包括特征拼接、加权组合和基于注意力机制的融合。例如,Li 等人^[42]提出将 ECG 和心音信号进行特征融合,通过时频变换将两种信号转换为频谱图、斯托克韦尔变换图和梅尔变换图后输入 CNN,在自建数据集上获得了 96.51%的准确率;Lee 等人^[43]整合基线资料与 12 导联心电图,开发了三种人工智能模型,在以 ECG、高风险因素和基线资料为输入的堆叠模型 C 中获得了 75%的敏感性,显著优于传统方法;Morshed 等人^[44]将心音信号和 ECG 输入到 1 维 CNN 中以提取一维时间变化,在心脏瓣膜缺陷检测任务中获得了 95.06%的准确率,优于现有的一些方法。

模型融合则侧重于异构数据的协同决策优化,既不同模态数据由各自的独立模型进行训练,随后通过集成学习方法对最终输出结果进行决策^[45]。常用的决策融合策略包括加权投票、软投票以及 Stacking 等方法。相比特征融合,模型融合的优势在于可

以充分发挥各模态独立特性^[46]，同时通过综合多个模型的决策结果，提高整体的鲁棒性与准确性。例如，Li 等人^[47]使用心音数据和心电图数据作为 BiLSTM 模型的输入，得到一个分类结果，再利用小波变换将两种数据转换为时频图用于 GoogleNet 模型的输入，得到 2 个分类结果。最终，通过改进的 D-S 理论对以上 3 个结果做出决策，在 PhysioNet 数据集上的特异性为 90.8%。综上所述，通过多模态数据融合，能够更全面地揭示疾病的病理特征，进一步提升诊断的准确性和可靠性。

1.2.4 研究现状评述

已有研究表明，在心血管疾病诊断中，单模态模型（如心音信号、心电图或基线资料）由于特征域存在局限性，精度很难进一步提高。多模态数据通过融合多源医学信息，的确能够进一步提高诊断的准确性。当前多模态诊断研究多采用双模态数据，信息维度仍然有限，仍需融合更多模态的医学数据，以进一步提高检测性能，从而探索多模态数据在 CAD 诊断中的潜力。

1.3 研究内容与技术路线

1.3.1 研究目标

本研究旨在综合应用信息处理和深度学习技术，结合心音信号、12 导联动态心电图和基线资料，研究多模态医学数据融合方法，构建基于多模态医学数据的冠心病诊断模型，实现低成本早期冠心病诊断。研究将以新疆医科大学胸痛中心为主要数据来源和示范应用单位，开展系统应用与模型评估工作，为冠心病的无创诊断研究和临床应用提供新的方法与工具。

1.3.2 研究内容

本研究以冠心病患者的心音信号、12 导联动态心电图和基线资料为研究对象，构建无创、高精度的冠心病诊断模型，开发冠心病辅助诊断系统进行成果验证和示范应用。通过本研究为临床医生早期无创诊断冠心病提供参考。该研究主要分为以下 3 个方面：

(1) 基于心音的冠心病诊断研究：以冠心病患者的心音信号为核心，深入探讨其特征变化与冠心病之间的关联。首先，采用 Daubechies 8 小波函数对收集到的心音信