

分类号：TP391

密 级：公开

学校代码：10363

学 号：2220340114



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目 基于深度学习的阿尔兹海默症图像
智能识别系统设计

论 文 作 者	<u>陶新杰</u>
校 内 导 师	<u>魏利胜（教授）</u>
校 外 导 师	<u>刘小琿</u>
专业学位类别	<u>控制科学与工程</u>
研究方向（领域）	<u>模式识别与智能系统</u>

论文提交日期：2025 年 5 月 10 日

分类号: TP391

单位代码: 10363

密 级: 公开

学 号: 2220340114

题 目 基于深度学习的阿尔兹海默症图像智能识别系统设计

题 目 Design of Alzheimer's Disease Image Intelligent Recognition
System Based on Deep Learning

学 生 姓 名: 陶新杰

校 内 导 师: 魏利胜 (教授)

校 外 导 师: 刘小琿

学 科 (专 业): 控制科学与工程

研 究 方 向: 模式识别与智能系统

论文答辩日期: 2025 年 5 月 10 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：

日期： 年 月 日

安徽工程大学学位论文授权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，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，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 ___ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：

日期： 年 月 日

指导教师签名：

日期： 年 月 日

基于深度学习的阿尔兹海默症图像智能识别系统设计

摘 要

在人口老龄化态势日益严峻的背景下，阿尔兹海默症 (Alzheimer's Disease, AD) 已经成为影响老年人认知功能的主要神经性疾病之一，AD 会造成老年患者的认知功能损伤和行为能力减退。目前，对于阿尔兹海默症尚未有有效的治疗手段，但在发病的早期阶段进行医学干预可改善患者的认知功能并延缓患者的神经退行性病变的速度，所以寻找一种快速、准确地早期诊断方法具有重要意义。随着医学成像技术的发展，神经影像学检查成为诊断 AD 的常用方法之一，如核磁共振成像 (Magnetic Resonance Imaging, MRI) 和正电子发射断层成像 (Positron Emission Computed Tomography, PET)。而传统的 AD 诊断方法依赖于医生的经验和医技水平，很多经验不足的医生很难通过肉眼诊断患者的医学影像，从而造成误诊或者漏诊。随着深度学习技术的发展，AD 医学影像的智能诊断方法已逐渐从传统的机器学习过渡到深度学习。本文基于深度学习技术，结合基于 2D 影像的 AD 识别算法、基于 3D 影像的 AD 识别算法和多模态数据融合的 AD 识别算法设计一个能够智能诊断 AD 的辅助系统。具体研究内容如下：

1) 针对传统的 CNN 架构对 AD 图像分类效果差和 Transformer 架构模型参数太大等问题，提出改进的 Mobilevit 模型方法用于 AD 影像的自动分类。该模型通过设计多尺度分支模块来扩大感受野，以更好地捕获全局和上下文特征信息。其次，在模型中引入通道注意力机制，用来学习不同通道的特征信息，并集成空间注意力机制，进一步提取空间中不同位置的关键信息，以提高模型特征表示的质

量和判别性能，并通过实验验证该改进模型的有效性。

2) 针对当前基于深度学习的 AD 二维影像识别准确率不高、具有数据泄露等问题，提出一个改进的 3D-ShufflenetV2 网络模型。利用 3D 卷积提取三维图像的空间连续信息，从而提取更多的脑部病变信息。为增强模型的分类性能，在主干网络中引入 Squeeze-and-Excitation 注意力机制增加模型对重要特征的关注能力。此外，为增加模型对全局特征提取的能力，在网络结构中增加了全局特征融合分支。实验结果表明改进后的网络模型相比改进前分类性能得到了较大的提升，能够对 AD、CN、MCI 实现更加准确的分类。

3) 针对现有 AD 诊断模型准确率较低、基于单一模态影像提取病变特征能力不足等问题，提出一种基于深度学习的多模态影像融合方法用于 AD 分类。为更加高效地提取三维医学影像中的病变特征，提出一种新型的残差网络结构，然后利用改进后的残差网络分别对 MRI 图像和 PET 图像进行特征提取，并将两种图像特征进行融合后再进行后续分类。最后，为进一步提升网络模型在三维医学影像中捕捉空间位置与通道间关系的能力，引入坐标注意力机制。实验证明改进后的融合网络在 AD/CN、AD/MCI、CN/MCI 分类任务中分类性能相比基础的残差网络均有不同程度的提升，且分类准确率均显著高于基于单一模态影像数据方法，证明了改进后的融合网络的有效性。

4) 结合基于 2D 影像的 AD 识别方法、基于 3D 影像的 AD 识别方法和基于多模态影像融合的 AD 识别方法模型，设计一个智能辅助识别 AD 图像的系统模型，为专业人员提供一个快捷高效的诊断系统，具有一定的临床使用价值。

综上所述，本文使用计算机辅助技术开发了一个基于深度学习

技术的 AD 自动诊断系统，能够帮助专业医生更加高效地诊断出阿尔兹海默症，且该系统支持诊断多种模态的影像数据，具有较高的兼容性和临床使用价值。此外，本文也为基于深度学习技术的 AD 识别领域提供了新的诊断思路和实验基础。

关键词：阿尔兹海默症；深度学习；图像分类；智能诊断系统

Design of Alzheimer's disease image intelligent recognition system based on deep learning

ABSTRACT

Under the background of increasingly severe population aging situation, Alzheimer's Disease (AD) has become one of the primary neurodegenerative disorders affecting cognitive function in the elderly. AD primarily causes cognitive impairment and behavioral decline in older patients. Currently, there are no effective treatments for Alzheimer's disease, but early-stage medical intervention can improve cognitive function and slow the progression of neurodegeneration. Therefore, developing a rapid and accurate early diagnostic method is of significant importance. Advances in medical imaging technologies have established neuroimaging examinations, such as Magnetic Resonance Imaging (MRI) and Positron Emission Computed Tomography (PET), as common diagnostic tools for AD. Traditional AD diagnosis relies heavily on physicians' expertise and technical proficiency, with less-experienced clinicians often struggling to accurately interpret medical images, leading to misdiagnosis or oversight. As deep learning technologies evolve, intelligent recognition of AD medical imaging has shifted from traditional machine learning to deep learning-based methods. This study leverages deep learning technologies to design an auxiliary system capable of intelligently diagnosing AD, integrating 2D imaging-based AD recognition algorithms, 3D imaging-based AD recognition algorithms, and multi-modal data fusion-based AD recognition algorithms. The specific research contents are as follows:

1) Aiming at the problems that traditional CNN architecture has poor effect on AD image classification and Transformer architecture model parameters are too large, an improved Mobilevit model method is proposed for automatic classification of AD images. The model enlarges the receptive field by designing multi-scale branching modules to better capture global and contextual feature information. Secondly, the channel attention mechanism is introduced into the model to learn the feature information of different channels, and the spatial attention mechanism is integrated to further extract the key information of different positions in space, so as to improve the quality of the model feature representation and discriminant performance, and the effectiveness of the improved model is verified by experiments.

2) Aiming at the problems of low accuracy and data leakage of current AD 2D image recognition based on deep learning, an improved 3D-ShufflenetV2 network model is proposed. 3D convolution is used to extract spatial continuous information of three-dimensional images, so as to extract more brain lesion information. In order to enhance the classification performance of the model, Squeeze-and-Excitation attention mechanism is introduced into the backbone network to increase the attention ability of the model to important features. In addition, in order to increase the ability of the model to extract global features, a global feature fusion branch is added to the network structure. The experimental results show that the classification performance of the improved network model is greatly improved compared with that before the improvement, and it can achieve more accurate classification of AD, CN and MCI.

3) Aiming at the problems of low accuracy of existing AD diagnostic

models and insufficient ability to extract lesion features based on single modal images, a multi-modal image fusion method based on deep learning is proposed for AD classification. In order to extract lesion features from 3D medical images more efficiently, a new residual network structure is proposed, and then the improved residual network is used to extract features from MRI images and PET images respectively, and then the two image features are fused before subsequent classification. Finally, in order to further improve the ability of the network model to capture the relationship between spatial position and channel in three-dimensional medical images, the coordinate attention mechanism is introduced. Experiments show that the classification performance of the improved fusion network in AD/CN, AD/MCI and CN/MCI classification tasks is improved to varying degrees compared with the basic residual network, and the classification accuracy is significantly higher than that based on single modal image data methods, which proves the effectiveness of the improved fusion network.

4) Combining the AD recognition method based on 2D images, the AD recognition method based on 3D images and the AD recognition method model based on multi-modal image fusion, a system model of intelligent auxiliary recognition of AD images is designed, which provides a fast and efficient diagnosis system for professionals, and has certain clinical value.

In summary, this paper used computer-aided technology to develop an AD automatic diagnosis system based on deep learning technology, which can help professional doctors diagnose Alzheimer's disease more efficiently, and the system supports the diagnosis of multiple modes of

image data, with high compatibility and clinical use value. In addition, this paper also provides new diagnostic ideas and experimental basis for the field of AD recognition based on deep learning technology.

Tao Xinjie (Control Science and Engineering)

Supervised by Wei Lisheng

Key words: Alzheimer's disease; Deep learning; Image classification; Intelligent diagnostic system

目 录

摘 要.....	I
ABSTRACT.....	IV
第 1 章 绪论.....	1
1.1 研究背景和意义.....	1
1.2 国内外研究现状.....	2
1.2.1 基于 2D 影像的阿尔兹海默症识别研究现状.....	3
1.2.2 基于 3D 影像的阿尔兹海默症识别研究现状.....	4
1.2.3 基于多模态数据融合的阿尔兹海默症识别研究现状.....	5
1.3 主要研究内容.....	6
1.4 本章小结.....	8
第 2 章 基于 2D 影像的阿尔兹海默症识别.....	9
2.1 基于改进的 Mobilevit 网络 AD 识别算法原理.....	9
2.1.1 改进的 Mobilevit 主干网络.....	10
2.1.2 CBAM 注意力机制.....	11
2.1.3 改进的 Mobilevit block.....	12
2.2 实验分析.....	14
2.2.1 实验数据集与预处理.....	15
2.2.2 评价指标.....	17
2.2.3 消融实验.....	17
2.2.4 对比实验.....	19
2.3 本章小结.....	20
第 3 章 基于 3D 影像的阿尔兹海默症识别.....	21
3.1 基于改进的 ShufflenetV2 网络 AD 识别算法原理.....	21
3.1.1 改进的 ShufflenetV2 主干网络.....	21
3.1.2 基本单元和下采样单元.....	22
3.1.3 通道混洗模块.....	24
3.1.4 Squeeze-and-Excitation 注意力机制.....	25
3.2 实验分析.....	26
3.2.1 实验数据集与预处理.....	26
3.2.2 消融实验.....	26
3.2.3 对比实验.....	29
3.3 本章小结.....	32
第 4 章 基于多模态数据融合的阿尔兹海默症识别.....	34
4.1 基于多模态图像融合分类网络的 AD 识别算法原理.....	34
4.1.1 多模态图像融合分类网络.....	34
4.1.2 改进的残差网络.....	36
4.1.3 scSE 注意力机制.....	38
4.1.4 坐标注意力机制.....	40
4.2 实验分析.....	41
4.2.1 实验数据集与预处理.....	42

4.2.2 消融实验.....	43
4.2.3 对比实验.....	45
4.3 本章小结.....	51
第 5 章 基于深度学习的阿尔兹海默症智能辅助诊断系统.....	52
5.1 系统需求分析.....	52
5.2 系统功能设计.....	53
5.3 界面设计及功能实现.....	54
5.3.1 开发环境.....	54
5.3.2 界面实现.....	54
5.4 本章小结.....	59
第 6 章 总结与展望.....	60
6.1 总结.....	60
6.2 展望.....	61
参考文献.....	62
攻读学位期间发表的学术论文.....	67
攻读学位期间获得的奖励.....	67
致 谢.....	68

插图清单

图 1-1 主要研究框架图	6
图 2-1 MRI 影像的三个视角切片图	9
图 2-2 改进的 Mobilevit 网络结构图	10
图 2-3 CBAM 的结构图	11
图 2-4 CAM 结构图	11
图 2-5 SAM 结构图	12
图 2-6 改进的 Mobilevit block 结构图	13
图 2-7 RFB 结构图	14
图 2-8 图像预处理步骤示意图	16
图 2-9 原模型分类混淆矩阵图	18
图 2-10 改进的 Mobilevit 网络分类混淆矩阵图	18
图 3-1 改进的 ShufflenetV2 网络结构图	22
图 3-2 ShufflenetV2 网络基本单元结构图	23
图 3-3 ShufflenetV2 网络下采样单元结构图	24
图 3-4 通道混洗示意图	25
图 3-5 SE 注意力机制网络结构图	25
图 3-6 原模型分类混淆矩阵图	28
图 3-7 原模型 + SE 分类混淆矩阵图	28
图 3-8 原模型 + Feature fusion 分类混淆矩阵图	28
图 3-9 原模型 + SE + Feature fusion 分类混淆矩阵图	29
图 4-1 多模态图像融合分类网络结构图	35
图 4-2 改进的残差网络结构图	36
图 4-3 scSE 注意力机制结构图	38
图 4-4 CA 注意力机制结构图	40
图 4-5 原模型分类实验混淆矩阵图	43
图 4-6 原模型 + CA 分类实验混淆矩阵图	44
图 4-7 原模型 + CA + scSE 分类实验混淆矩阵图	44
图 4-8 原模型 + CA + scSE + Channel Shuffle 分类实验混淆矩阵图	44
图 5-1 系统的总体框架	53
图 5-2 系统登录界面	54
图 5-3 系统功能主界面	56
图 5-4 2D 图像诊断界面	56
图 5-5 3D 图像诊断界面	57
图 5-6 多模态数据诊断界面	58
图 5-7 患者信息界面	58
图 5-8 患者信息查询界面	59

插 表 清 单

表 2-1 消融实验结果	17
表 2-2 Mobilevit 与其他模型分类性能与参数量对比结果	19
表 2-3 不同 AD 分类方法的对比结果	20
表 3-1 数据集详细信息	26
表 3-2 消融实验结果	27
表 3-3 在 AD/MCI 数据集上与不同模型分类性能对比结果	30
表 3-4 在 CN/MCI 数据集上与不同模型分类性能对比结果	30
表 3-5 在 AD/CN/MCI 数据集上与不同模型分类性能对比结果	31
表 3-6 与经典深度学习模型对比结果	32
表 4-1 数据集详细信息	42
表 4-2 消融实验结果	43
表 4-3 不同网络模型基于不同模态数据的分类准确率(%)对比结果	45
表 4-4 基于不同模态数据分类的 AD 敏感性和特异性对比结果	46
表 4-5 性能对比结果	50
表 4-6 分类准确率对比结果	50

第 1 章 绪论

1.1 研究背景和意义

阿尔兹海默症(Alzheimer's Disease, AD)是一种常见的神经退行性疾病, 主要影响老年人的记忆、思维和行为功能, 严重影响患者的生活。随着人口老龄化形势的不断加剧, 阿尔兹海默症的发病率逐年增加, 预计到 2050 年, 患者数量将增加 2 到 3 倍, 该疾病已经成为全球公共卫生领域面临的重大挑战之一^[1]。AD 的发展是一个漫长的过程, 其发病历程主要经历三个阶段, 分别为正常认知状态 CN(Cognitively Normal)、轻度认知障碍 MCI(Mild Cognitive Impairment)和阿尔兹海默症 AD^[2]。而早期 AD 的临床表现并不明显, 主要表现于记忆力减退、难以集中注意力等, 因此大多数 AD 患者确诊时通常是在疾病的中晚期, 而且脑部的病变会随着年龄的增长越发严重^[3]。但如果能够诊断出早期的 AD, 患者可以及时地寻求专业医生的帮助, 明确疾病的发展趋势, 并给出对应的治疗方案, 从而在早期 AD 病程中帮助患者提高生活质量和延缓疾病的发展趋势^[4]。因此早期的 AD 诊断, 不仅能够为患者提供更早的干预机会, 还有助于缓解疾病的进展, 改善患者的生活质量。

近年来, 随着医学成像技术的发展, 神经影像学检查成为诊断 AD 的常用方法之一, 如核磁共振成像(Magnetic Resonance Imaging, MRI)^[5]和正电子发射断层成像(Positron Emission Computed Tomography, PET)^[6]。两种成像方法不同, 但都能够帮助医生更直观地分析患者的脑部病变区域。MRI 的优势在于其高空间分辨率, 可以提供脑结构的详细三维信息。此外, MRI 成像过程不涉及放射性物质, 因此是一种无创且安全的检查手段。MRI 能够显示出脑组织的结构变化, 例如早期 AD 的 MRI 影像特征表现为内侧颞叶的灰质萎缩^[7]。随着病情的发展, 脑部的顶叶、额叶、枕叶皮层等区域都会发生结构性变化^[8]。与 MRI 不同, PET 是一种通过注射放射性示踪剂并检测其在体内分布情况来获取图像的成像技术。PET 可以在分子水平上反映脑部的代谢活动, 因此, 能够提供有关大脑功能状态的信息^[9]。AD 患者的脑部代谢异常通常表现为葡萄糖代谢的降低, 尤其是在与记忆和认知功能相关的脑区, 如海马体和额叶区域。因此, PET 在 AD 的早期诊断中具有较高的敏感性, 尤其是在患者尚未表现出明显的认知症状时,

PET 能够通过监测脑部的代谢变化，提前发现潜在的病理变化^[10]。

传统的 AD 诊断方法通常依赖于医务人员通过分析医学影像或结合认知测试结果来进行判断。然而，影像分析和认知测试的结果容易受到医务人员的经验、知识水平和个人判断的影响，在 AD 的早期阶段，病变变化微妙，容易被忽视或误诊。由于 AD 的病理变化较为复杂，需要专业人员对脑部结构、代谢变化等进行细致的分析。然而，高水平的专业人才在全球范围内并不普遍，尤其是在一些偏远地区或发展中国家，医疗资源较为匮乏。这使得在这些地区的 AD 早期诊断面临较大的挑战，导致患者难以获得及时的诊断和治疗。因此，开发一种能够自动诊断 AD 的辅助系统，不仅能够提高早期 AD 诊断的准确性和一致性，还能够减少对专业人员的依赖，优化医疗资源的利用，对临床应用和 AD 自动诊断的发展具有深刻意义。

1.2 国内外研究现状

近年来，随着人工智能技术的发展，深度学习技术凭借其在图像识别和分类方面的卓越性能，为医学影像分析领域带来了革命性的变革。深度学习是一种以人工神经网络为基础的机器学习方法，通过构建多层次的神经网络模型，可以从大规模的数据中学习数据特征，并实现高效的分类和识别^[11]。在阿尔兹海默症的诊断中，医学影像数据通常具有高分辨率和大量的细节信息，但同时也面临着低识别率和信息复杂性高的问题。AD 的初期病理特征可能表现得非常细微，且患者的症状常常不明显，传统的影像分析方法依赖专业人员的经验，往往无法在早期阶段准确识别出微小的脑部变化。因此，AD 的早期诊断一直是医学影像领域中的一大挑战。深度学习技术的引入，为 AD 的自动诊断提供了巨大的契机。深度学习通过自动提取和学习影像中的高维特征，可以有效识别出与 AD 影像中相关的结构性变化和代谢异常，甚至在 AD 的早期阶段，提前预测患者大脑中的病理变化。与传统方法相比，深度学习技术能够在短时间内处理大量的影像数据，减少医务人员的工作负担，并大幅提高诊断的准确性和效率。目前，国内外已有大量学者投入到 AD 自动诊断的研究中，并取得了显著进展。研究主要集中在利用深度学习技术对 MRI、PET 等医学影像数据进行分析，通过构建卷积神经网络等深度学习模型，来提取影像特征，进行分类和预测。与此同时，多模态深度学习方法也得到了广泛应用，结合 MRI 和 PET 数据

进行更全面的疾病评估，从而提高 AD 识别的准确率。通过这些技术，研究人员能够有效提高 AD 的诊断精度，尤其是在早期诊断中的应用前景广阔。下面将对目前该领域的研究成果进行总结，并分析目前所面临的挑战。

1.2.1 基于 2D 影像的阿尔兹海默症识别研究现状

二维影像因其具有直观、便捷的特点，在临床诊断中经常被使用。在 AD 诊断中，二维影像通常指一些脑部的切片影像，如脑部核磁共振影像的矢状面、横断面、冠状面二维切片图像。深度学习模型通过提取这些切片图像的特征，然后基于提取的特征进行 AD 的分类识别。目前国内外学者在基于 2D 影像的 AD 识别领域进行了大量的研究，并取得了丰富的成果。例如 Billones 等人^[12]使用 VGGNet 模型对脑部的 MRI 冠状面切片进行分类识别，在 AD/MCI/CN 三分类任务中，识别的准确率达到 91.85%，在 AD/CN 分类中准确率达到 98.33%。但是该方法没有考虑到切片之间的特征关系，而是在选取切片时统一选择中间部分的切片图像，所以图像中存在着大量的冗余信息。为此，Jabason 等人^[13]利用图像熵的方法选取信息量最大的切片图像进行训练，实验表明该方法能够提高 AD 识别的准确率。同样也是使用图像熵的方法选取信息量最大的切片图像，Jain 等人^[14]使用迁移学习的方法预训练 VGG-16 网络，大幅度提升 AD 识别的准确率和模型的训练速度。但是在上述研究中，模型识别的准确率太过于依赖经过筛选的切片图像，所以对图像的细小特征信息提取能力较弱。为此，石磊等人^[15]使用特征增强金字塔网络(Feature Enhanced Pyramid Network, FEPN)提取切片图像的多尺度特征信息，增强了图像细小特征的提取能力。但是该方法的模型参数量较大，网络结构复杂。为简化模型的网络结构，朱剑波等人^[16]在 EfficientNet 网络中使用深层卷积和残差模块来减少模型的参数量，并在主干网络中添加自适应注意力机制，以增加模型对切片图像特征的提取能力。王宇春等人^[17]使用 Faster RCNN 分类框架，使用密集连接的特征金字塔网络提取海马体区域的 MRI 二维切片图像特征，实验表明这种方法能够显著提升 AD 识别的精度。Masud 等人^[18]提出了一种应用加权损失函数的轻量级卷积神经网络模型来解决当前医学影像数据集存在类别分布不均匀的问题，并通过实验验证该方法在数据分布不平衡的情况下的分类有效性。

从以上文献研究分析可知，基于 2D 影像的 AD 分类方法具有模型结构简单、参数量较少和提取特征简单等优点，且国内外学者提出了多种分类模型，取得

了较多的研究成果。但目前基于 2D 影像的 AD 分类方法仍存在依赖手工筛选图像特征、分类准确率低等问题，所以基于 2D 影像的分类模型仍需改进。

1.2.2 基于 3D 影像的阿尔兹海默症识别研究现状

在 2D 影像分类方法中，分类模型对切片图像的特征提取依赖于切片的选择。但如何选择有价值的切片图像，目前尚未有可靠的方法。此外，由于每个切片都是二维平面的，数据的局限性可能导致一些重要信息的丢失。特别是在处理不同切片之间的信息传递时，2D 影像分类方法通常依赖于切片之间的推断，这容易导致局部区域的诊断错误。相比之下，3D 图像通过对大脑进行全方位的扫描，减少了信息丢失的可能性，提供了更加连续、完整的结构信息。所以有研究者提出使用 3D 影像进行 AD 的识别。深度学习技术在医学影像中的应用尤其依赖于影像数据的丰富性和完整性，3D 图像能够为深度学习模型提供更多的空间信息，这不仅增加了模型学习的输入维度，而且使得模型在进行分类时能够识别到更多的复杂模式。例如 Karasawa 等人^[19]使用 3D CNN 提取三维核磁共振影像的特征信息并进行分类，实验表明该方法在测试集上 AD 分类的准确率达到 87%。Goceri^[20]提出了一种新的 3D CNN 拓扑结构用于三维核磁共振影像的分类，该方法通过使用 Sobolev 梯度优化器和优化激活函数的方式来提高模型的性能。以上两种方法在 AD 早期阶段的识别准确率都很低，为此，Ai^[21]使用改进的 Densenet169 和 Resnet50 架构来诊断和分类 AD，这两种架构在正常、极轻度认知障碍、轻度认知障碍和中度认知障碍分类中准确率分别达到了 83.8% 和 81.9%，实验表明该方法能够用于早期 AD 的实时分析和分类。孙俊楠等人^[22]设计了一个 3D-SAMDensenet 算法用来对三维的核磁共振影像进行分类，该方法与传统的分类方法相比，在 MCI/CN 分类任务中更有优势。在以上的三维影像分类研究中，都是将三维的全脑影像作为深度模型的输入来进行分类的，这样会增加模型的计算量。为此，部分研究者提出了以脑部分割组织的局部差异来进行 AD 的分类识别。Sun 等人^[23]使用 3D CNN 从三维脑部核磁共振影像中分割双侧海马体并用于 AD 的分类，实验证明该方法能够有效诊断出早期的 AD 患者。曾安等人^[24]设计了一种基于脑部感兴趣区域的 3D CNN 架构用来辅助诊断 AD，研究发现该方法能够有效提升模型分类效果和泛化能力。

基于以上文献研究，目前基于 3D 影像的 AD 分类方法取得了显著成效，但仍然存在一些不足和挑战。首先，尽管深度学习模型能处理更为丰富的空间信

息,但由于全脑 3D 影像数据的高维特性,模型的计算量显著增加,训练过程需要大量的计算资源和时间,尤其是在临床应用中,计算效率仍然是一个瓶颈。其次,虽然目前出现了多种改进的网络架构,但在 AD 的早期阶段,对 AD 的不同病程分类准确率仍然偏低,缺少临床使用价值,且在不同患者之间的表现差异较大。因此,基于 3D 图像的 AD 识别方法仍需进一步改进。

1.2.3 基于多模态数据融合的阿尔兹海默症识别研究现状

以上两种识别方法都是基于单模态的影像信息识别 AD,而使用单模态的数据对 AD 进行诊断时,会因数据信息单一而不能够全面的分析病理信息,进而对 AD 识别的准确率偏低。如果能结合多模态数据特征进行 AD 识别,将不同模态的特征信息进行互补能够进一步提高 AD 识别的精度^[25]。Dai 等人^[26]使用两个三维神经网络分支分别学习 MRI 影像和 PET 影像的特征,然后将两者的特征进行融合,最后再基于融合特征进行 AD 分类,实验结果表明相比单模态影像识别,多模态融合识别 AD 的准确率大幅提升。Lu 等人^[27]结合 MRI 影像和 PET 影像分割出的灰质图像融合特征识别早期的 AD,研究发现使用融合特征识别方法性能更优。Song 等人^[28]将配准后的 MRI 影像与 PET 影像的灰质组织部分融合,然后使用 3D Simple CNN 对融合后的影像进行识别,该方法在 MCI/CN、MCI/AD 和 AD/CN 分类任务中准确率分别达到了 88.5%、84.8%和 94.1%。从以上研究可知,基于多模态数据融合识别方法的分类精度是优于基于单模态数据识别方法的。但是以上几种方法都是基于影像数据的融合,为了进一步提高 AD 分类的准确率,Wang 等人^[29]将患者的多模态数据(包括 3D 脑影像数据、人口统计学数据、遗传数据和血液的生物标记物数据)融合到一个深度学习框架中,并实现了早期 AD 的诊断。林雪峰等人^[30]设计了一种多模态分类网络,通过两个输入网络分别输入 MRI 影像和量表信息等一些临床数据,然后在池化层之后将多模态的数据特征进行融合,最后在全连接层进行分类识别,实验结果表明这种识别方法能够提高分类的准确性。Dwivedi 等人^[31]使用恶魔算法和离散小波变换对 MRI 图像和 PET 图像特征进行融合,然后使用 Resnet50 对图像进行特征提取,最后再进行分类。实验证明,这种改进方法能够更好地融合特征,并提高早期 AD 的分类准确率。Odusami 等人^[32]提出了一种早期特征融合方法,将 PET 图像和 MRI 图像特征串联,并使用改进的分类模型来解决两种图像数据的异质性问题,实验证明了这种分类方法的有效性。Zhang 等人^[33]为解决 PET 数

据缺失的问题，通过金字塔注意力机制和标准判别器生成 PET 图像，并将 MRI 图像与 PET 图像进行融合分类，实验结果显示这种方法在 AD/CN/MCI 分类任务中实现了 89.9% 的分类准确率。Rallabandi 等人^[34]是将基于图谱的分割 MRI 影像组织与 PET 影像融合，然后使用逆傅里叶变换和离散小波变换的方法重建融合后的影像，最后再基于融合的影像进行识别，该方法在多个病程分类任务中取得了优秀的识别性能。

基于以上文献研究可知，基于多模态数据融合的 AD 分类方法取得了显著成果，但在现有研究中，仍有许多未被充分解决的问题。例如现有的大多数多模态方法仍然依赖于手工设计的特征提取策略，且目前基于多模态数据融合的 AD 分类方法准确率普遍偏低，且存在模型复杂、参数量大、泛化能力不强等问题，因此需要进一步轻量化融合网络模型，提高模型的分类准确率。

1.3 主要研究内容

本课题旨在以 AD 不同模态的医学影像为研究对象，从二维图像、三维图像和多模态图像融合三个研究角度出发，探索基于不同模态图像的 AD 分类方法。最后，结合上述的三种分类方法，设计一个智能辅助识别 AD 图像的集成系统，实现对 AD 快捷高效的识别。课题的研究框图如图 1-1 所示，主要研究内容如下：

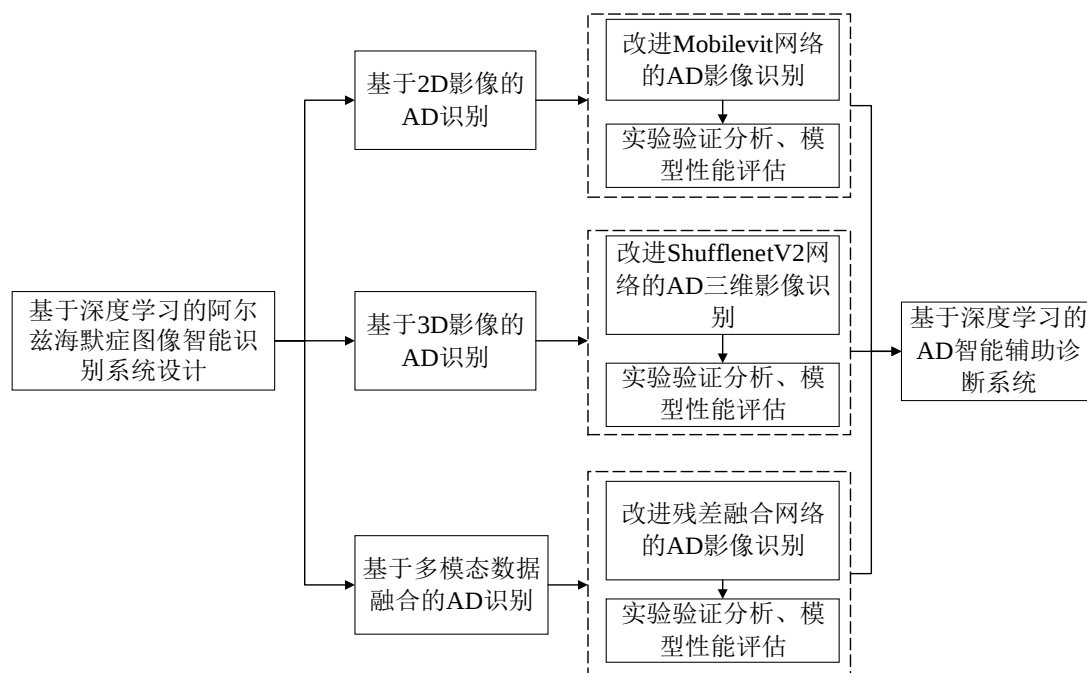


图 1-1 主要研究框架图

第 1 章：详细阐述了本课题的研究背景及意义，分析了当前基于 2D 影像的 AD 识别、基于 3D 影像的 AD 识别和基于多模态融合的 AD 识别的国内外研究现状，并总结当前研究工作的特点和不足，以及可能解决这些问题的方法，最后总结了本文主要的研究内容。

第 2 章：针对传统的 CNN 架构对高维图像分类效果差和 Transformer 架构模型参数太大等问题，提出了改进的 Mobilevit 模型用于 AD 核磁共振影像的识别。首先，通过设计多尺度分支模块来扩大感受野，以利于模型更好地捕获图像全局和上下文信息。其次，融合了通道注意力机制和空间注意力机制，用来学习不同通道的特征信息和提取空间中不同位置的关键信息，以提高模型特征表示的质量和判别性能，并通过实验验证分析该方法的有效性。

第 3 章：针对当前基于深度学习的 AD 二维影像识别准确率不高、具有数据泄露、识别效率低等问题，提出一个改进的 3D-ShufflenetV2 网络。首先，该网络利用 3D 卷积提取三维图像的空间连续信息，从而使网络能够提取更多的脑部病变信息。其次，在主干网络中引入 SE 注意力机制，通过基于全局信息重新校准通道级特征响应以增加模型对重要特征的关注能力，从而增强模型的性能。最后，为了增加模型对全局特征提取的能力，在网络结构中增加了全局特征融合分支。实验结果表明，改进网络能够实现对 AD、MCI 和 CN 进行准确的识别，改进后网络模型的性能比原网络有较大的提升，具有一定的应用价值。

第 4 章：针对现有 AD 诊断模型准确率较低、基于单一模态影像提取病变特征能力不足等问题，提出一种基于深度学习的多模态影像融合的 AD 分类方法。首先，提出一种新型的残差网络结构用来提取三维图像的病变特征。其次，利用改进后的残差网络作为特征提取器分别提取 MRI 图像和 PET 图像特征，然后将两种图像特征融合后进行分类。最后，为了使融合网络能更好地捕捉三维医学影像的空间位置与通道之间的关系，引入坐标注意力机制。实验结果证明，改进后的融合网络在多个分类任务中分类性能较基础的残差网络均有不同程度提升，且融合网络分类的准确率较单模态网络分类的准确率有较大提升，证明了改进后融合网络的有效性。

第 5 章：结合基于 2D 影像的 AD 识别方法、基于 3D 影像的 AD 识别方法和基于多模态影像融合的 AD 识别方法模型，设计一个智能辅助识别 AD 的系统

原型，该集成系统能够提供高精度的诊断结果和友好的用户控制界面，从而提高医生诊断的效率和准确率。

第 6 章：对本文工作进行概括和总结，并阐述本文研究工作的局限性，以及对未来工作和研究方向进行展望。

1.4 本章小结

本章首先阐述了研究 AD 辅助诊断系统的背景和实际意义，以及指出当前临床诊断 AD 存在依赖专业知识经验、消耗大量医疗资源等问题，这些问题亟需借助计算机辅助技术来解决。然后，从基于 2D 医学影像、3D 医学影像和多模态影像融合识别 AD 三个角度介绍该领域的国内外研究现状，并总结分析当前研究所存在的问题。最后总结本课题研究内容的总体框架，并阐述每章研究的主要内容，为后文详细论述做铺垫。

第 2 章 基于 2D 影像的阿尔兹海默症识别

目前，AD 辅助诊断主要以使用深度学习技术对 MRI 影像的切片数据进行分类识别为主，这种分类方法主要于卷积神经网络对 MRI 影像的切片数据进行分析 and 诊断。二维切片图像在处理过程中具有明显的优势，由于其结构简单，且二维图像的输入数据量较小，能够有效降低训练模型时的计算复杂度，从而减少模型的训练时间和计算资源需求。通常一个 MRI 的切片图像会从三个视角进行裁取，图 2-1 展示的是同一个 MRI 影像的三个不同视角的切片图像。传统 CNN 模型在处理高维度图像时，仍然面临着特征提取困难等问题。与此同时，近年来出现的 Transformer 架构^[35]虽然在图像处理领域展现出强大的性能，但其庞大的模型参数和计算量却成为其实际应用的瓶颈。因此，目前基于 2D 影像的 AD 识别仍存在一定的局限性。

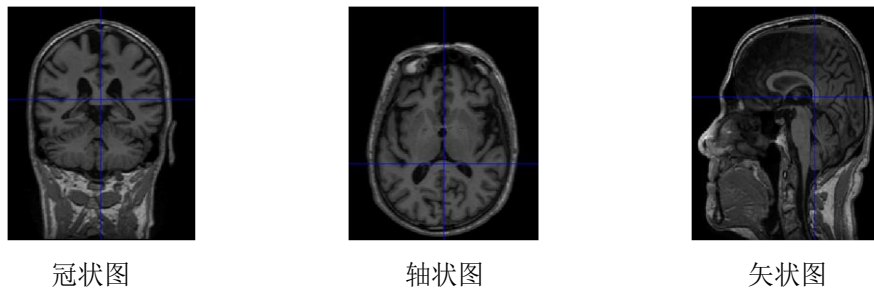


图 2-1 MRI 影像的三个视角切片图

2.1 基于改进的 Mobilevit 网络 AD 识别算法原理

深度学习技术在 AD 辅助诊断领域已经取得了较好的效果，但是传统的网络存在参数量较大、训练难度较大、准确率较低的缺点。所以针对以上问题，提出一种基于改进的 Mobilevit^[36]轻量级网络来提取 MRI 影像的切片特征。首先在 Mobilevit 主干网络中添加 CBAM(Convolutional Block Attention Module)^[37]注意力机制，以提高通道之间的关系，并在提高特征的表征能力和判别性的同时，减少计算成本。其次，在 Mobilevit block 中的局部特征提取部分引入局部感受野模块 RFB(Receptive Field Block)^[38]。通过引入多个不同尺度的感受野，可以增强局部特征的表达能力。此外，将原始输入与全局特征融合分支部分删除，改为局部特征与全局特征融合，以使网络更加全面地理解图像的语义信息，提高模型对全局图像的感知能力。

2.1.1 改进的 Mobilevit 主干网络

传统的 CNN 在视觉任务能够以较少的参数学习特征，但是这些特征都是局部的。而 Vit(Vision transformer)^[39]正好克服了 CNN 不能够提取全局特征的缺点，但是其参数量太大，算力要求高，这样就不适合部署到移动设备上。所以 Mobilevit 结合了 CNN 和 Vit 的优点，构建出一种既轻量又能够保持高精度识别的网络模型。改进后的 Mobilevit 主干网络结构如图 2-2 所示，为了增加模型对图像通道特征和空间特征的关注程度，在第一个卷积层之后添加 CBAM 模块。此外，为了进一步增强模型提取 MRI 切片图像特征的能力，将原始的 Mobilevit block 进行优化，引入 RFB 模块并改进特征融合分支模块。

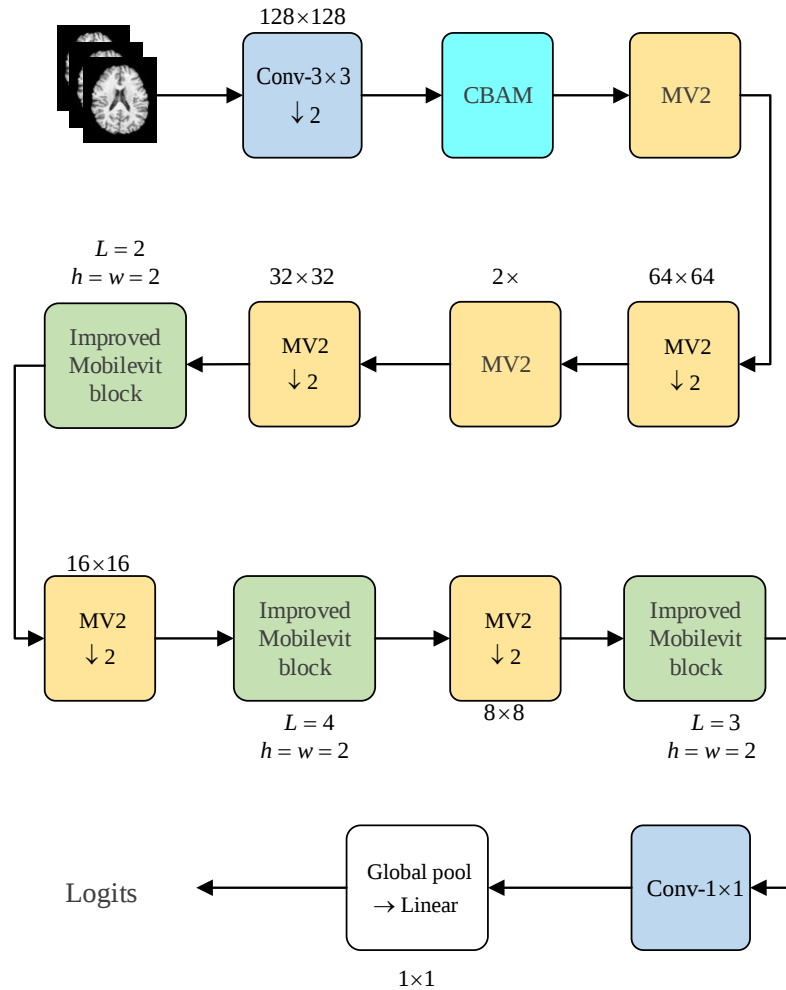


图 2-2 改进的 Mobilevit 网络结构图

图中的 MV2 是 MobileNetV2 网络中的倒残差结构的缩写， $\downarrow 2$ 代表执行步长为 2 的下采样， $2\times$ 表示执行两次该操作。在 Improved Mobilevit block 中， $L=2$ 表示 Transformer 模块执行的次数， $h=w=2$ 代表将每个 batch 设置成 2×2 的大小。

2.1.2 CBAM 注意力机制

为了提高模型的特征表示能力和辨别力，在主干网络的局部建模之后引入了 CBAM 模块。CBAM 的结构如图 2-3 所示，CBAM 主要由通道注意力模块(Channel Attention Module, CAM)^[40]和空间注意力模块(Spatial Attention Module, SAM)^[41]构成。CBAM 通过对输入特征的通道和空间信息进行不同的池化操作，输出更重要的通道特征和空间特征，以此来提高模型的特征表达能力。此外，CBAM 使用了两条分支结构，分别对空间特征和通道特征进行相应的处理，不仅能够进一步节省算力和减少模型的参数，还能够使其更加方便地嵌入到其他网络模型中。

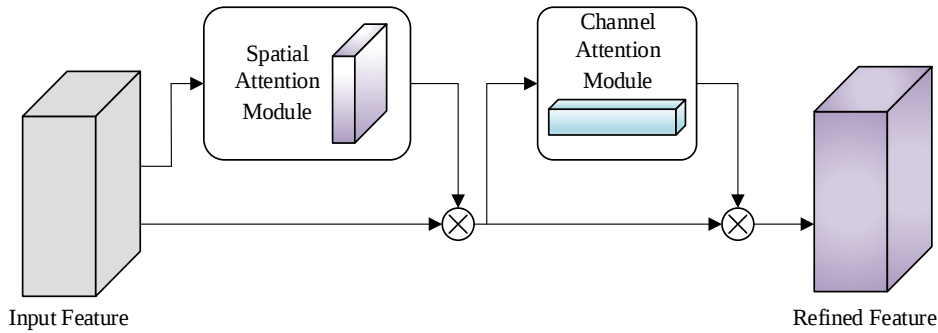


图 2-3 CBAM 的结构图

CAM 通过对通道特征进行全局平均池化和全连接层的处理，来学习通道之间的关系，然后得到每个通道的权重，最后对通道特征进行加权平均，得到具有重要通道特征的输出。其网络结构如图 2-4 所示，CAM 会将输入的特征 $F(H \times W \times C)$ 分别在特征图的高和宽两个维度上进行全局最大池化和全局平均池化，并得到两个 $1 \times 1 \times C$ 的张量，然后再将得到的两个张量分别输入一个两层的神经网络，这两层的神经网络是共享的。然后再将共享神经网络层输出的两个特征图进行加和并使用 Sigmoid 函数对特征进行激活，生成最终的通道注意力特征。最后将输出的特征与输入的特征进行相乘，完成对原始特征的加权，得到增强后的特征图。通道注意力特征计算过程如公式(2-1)所示。

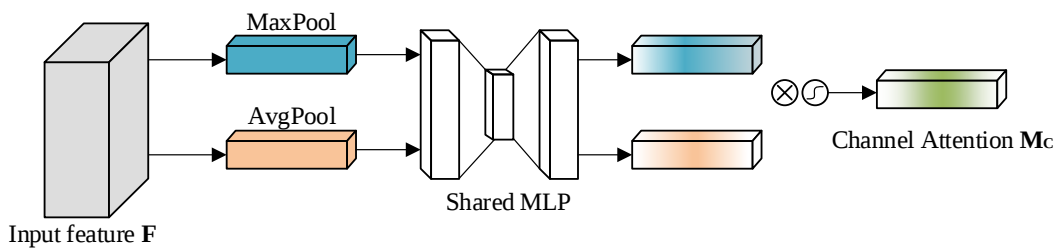


图 2-4 CAM 结构图

$$\begin{aligned} M_c(F) &= \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \\ &= \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))) \end{aligned} \quad (2-1)$$

其中， σ 表示激活函数， $W_0 \in \mathbb{R}^{C/r \times C}$ ， $W_1 \in \mathbb{R}^{C \times C/r}$ ， AvgPool 表示平均池化操作， MaxPool 表示最大池化操作， MLP 表示一个两层的神经网络， F_{avg}^c 和 F_{max}^c 分别表示平均池化特征和最大池化特征。

SAM 与 CAM 的核心区别在于 SAM 主要关注空间特征之间的关系，通过学习每个位置的权重，突出重要的空间特征并抑制无关区域。具体而言，SAM 首先对输入特征图在通道维度上分别进行全局最大池化和全局平均池化操作，从而生成两张空间特征图。然后将上一步得到的两个空间特征进行拼接得到组合特征。接着，将组合特征通过一个卷积层进行降维处理，并通过激活函数激活生成最终的空间注意力特征。最后，将生成的空间注意力特征与输入特征进行相乘，得到增强后的空间特征。SAM 的网络结构如图 2-5 所示，空间注意力特征计算过程如公式(2-2)所示。

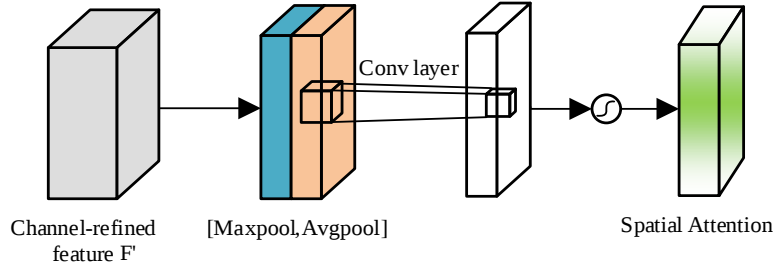


图 2-5 SAM 结构图

$$\begin{aligned} M_s(F) &= \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \\ &= \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])) \end{aligned} \quad (2-2)$$

其中， σ 表示激活函数， $f^{7 \times 7}$ 表示卷积核为 7×7 的卷积操作， F_{avg}^s 表示平均池化特征， F_{max}^s 表示最大池化特征。

2.1.3 改进的 Mobilevit block

改进的 Mobilevit block 结构如图 2-6 所示，Mobilevit block 有三个主要模块，分别为局部特征表达模块、全局特征表达模块和融合特征模块。改进的 Mobilevit block 在局部特征表达部分添加了 RFB 模块，并删除了原来的原始输入数据与全局特征融合分支，改为局部特征与全局特征融合。输入特征首先经过一个 3×3 的卷积层提取局部特征，然后再将特征输入到 RFB 模块进行下一步

的特征提取，以增强局部特征提取能力。为了减少模型的参数量，通过重组特征结构的方式进行全局的特征建模，在通道方向将局部特征与全局特征拼接，以提高模型的鲁棒性。最后将输出特征经过一个 3×3 的卷积层进行特征融合输出。为了让改进的 Mobilevit block 提取到更多的特征信息，将局部特征与全局特征进行融合。在局部特征提取部分提取完特征之后，将提取完的特征向量沿通道方向与全局特征提取部分输出的特征向量拼接，这样相比将原始数据与全局特征拼接，能够提取更多的图像特征。

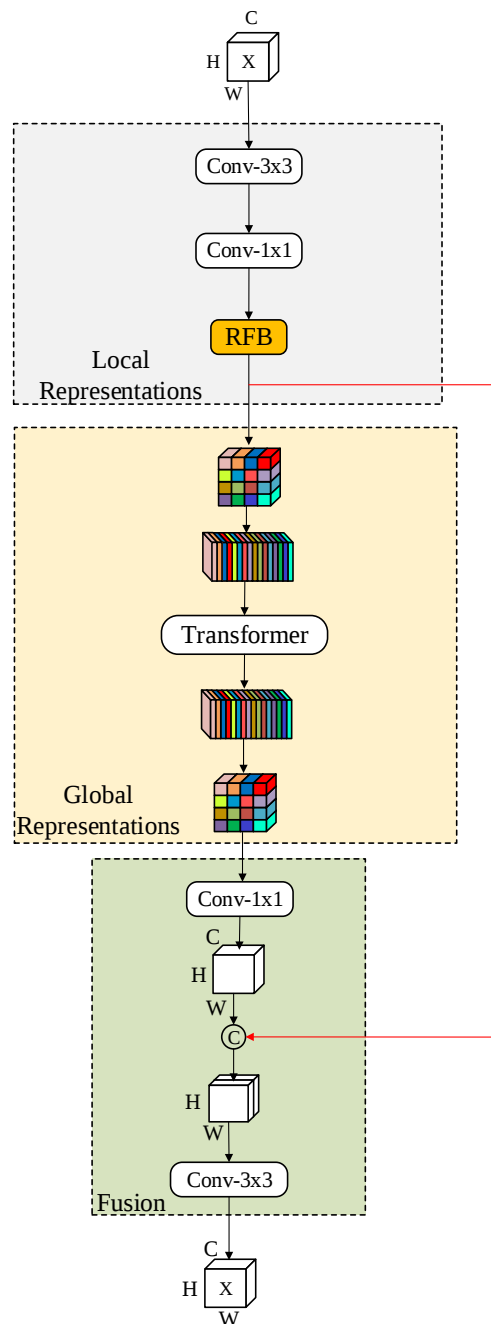


图 2-6 改进的 Mobilevit block 结构图

Mobilevit block 只使用了一个 3×3 的卷积层提取局部特征，这样在多尺度特征提取方面就失去了一定的感知能力。RFB 结构如图 2-7 所示，RFB 模块具有四个分支结构，其中三个分支网络由不同尺寸卷积核的卷积层组成，然后在每个分支的卷积层之后添加不同扩张率的空洞卷积来提取特征，最后将三个分支提取的特征进行融合。这样能够进一步扩大感受野，使局部特征表达部分能够更好地学习到图像上下文信息。

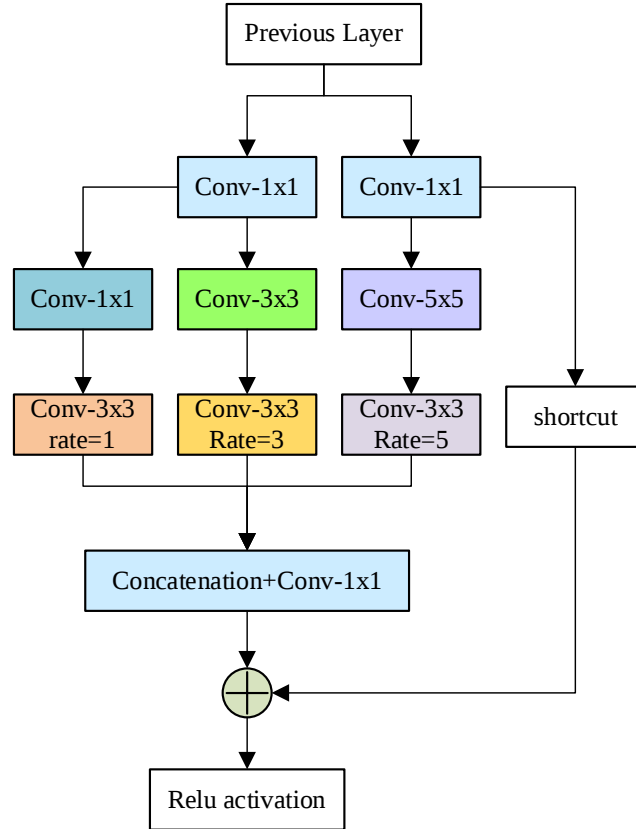


图 2-7 RFB 结构图

2.2 实验分析

本节对改进的 Mobilevit 网络模型进行实验验证分析，首先，介绍实验使用的公开数据集以及数据集数据组成的详细信息，并介绍了数据集的预处理步骤。其次，介绍了本章实验所使用评估模型性能的评价指标，最后对改进的 Mobilevit 网络模型进行消融实验和对比实验，消融实验通过逐步加入不同改进模块，定量分析各模块对模型分类性能的具体影响。对比实验通过在同一数据集上使用其他经典网络模型与改进的 Mobilevit 网络模型进行性能对比，全面展示了改进后 Mobilevit 网络在 AD 识别任务中的优势。

本章实验所使用的实验环境为 Pytorch 深度学习框架，采用的硬件配置为 Windows11 操作系统，Intel i5 CPU，NVIDIA GeForce GTX3060 显卡，12G 显存，编程语言为 Python 3.8。数据集划分比例为训练集：验证集：测试集=7：2：1。使用 AdamW 优化器，初始学习速率设置为 0.0002，weight_decay 设置为 0.02，学习率下降方式采用余弦退火衰减策略，训练迭代次数为 200 次，批量大小设置为 32。

2.2.1 实验数据集与预处理

本章所使用的实验数据均来自公开的数据集阿尔兹海默症神经影像计划(The Alzheimer's Disease Neuroimaging Initiative, ADNI)^[42]。ADNI 由美国国家生物医学成像与生物工程研究院(NIBIB)联合国家老龄化研究所(NIA)于 2004 年共同发起，针对 AD 的早期检测和跟踪构建了 ADNI 数据库。该计划通过系统构建多模态生物医学数据库，为 AD 病理机制研究与早期诊断技术开发提供关键数据支撑。历经 ADNI-1、ADNI-GO、ADNI-2 及 ADNI-3 四个研究阶段，目前已整合形成涵盖影像数据、临床认知测试量表、生物标志物和遗传学数据等多维度数据集。本章从中随机抽取 588 受试者的 1.5T 类型 T1 加权的 MRI 数据，所有的受试者都经过了基本临床评估，并经过专业医生诊断，确保实验数据准确。每个 MRI 影像取十张切片作为实验图像数据，其中包括 AD 图像 1860 张，MCI 图像 1960 张和 CN 图像 2060 张，其分辨率统一为 308×369 像素，图片格式为 PNG。

从 ADNI 数据库获取的原始数据都是全脑的影像，原始的影像数据包含了颅骨、头皮和其他头部组织，这些冗余信息会影响网络模型对脑部病变特征的提取。为了提高影像的信噪比，提升后续处理和分析的准确性，本章对所有原始 MRI 影像数据进行了预处理操作，预处理步骤示意图如图 2-8 所示，预处理步骤如下：

1) AC-PC（前交叉-后交叉）校正

AC-PC 校正，全称为 Anterior Commissure-Posterior Commissure 校正，是一种在脑部影像处理中常用的标准化方法。这种方法基于脑内两个解剖标志点：前交叉(Anterior Commissure, AC)和后交叉(Posterior Commissure, PC)。通过将这两个点在图像中对齐到一个标准位置，可以实现不同个体脑影像的标准化对齐。其目的是减少不同受试者之间头部位置的差异，这种校正确保所有影像数据在

进行分析时都是在相同的空间参考系内，从而提高了数据分析的一致性和准确性。本章实验使用 Matlab 软件中的 SPM12 工具箱完成所有数据的 AC-PC 校正。

2) 模板配准

模板配准是将每个受试者的脑影像数据与一个预定义的标准脑模板进行对齐的过程。这一步骤可以进一步减少个体间的解剖结构差异，有助于在群体水平上进行比较和分析。通过模板配准，可以更容易地识别出 AD 患者脑部的典型变化特征，如海马体体积缩小等，从而提高分类模型的准确性。本章实验使用 FSL 软件对所有影像数据配准到 MNI152 标准空间，这步操作可以减少图像间的差异，使模型更容易学习到疾病的共性特征。

3) 去除颅骨

为了增加图像数据的信噪比，去除非脑组织对分类模型提取特征的干扰，将所有影像数据都进行去颅骨处理。尤其是对于冗余信息较多的 MRI 数据来说，去颅骨操作对提高模型的 AD 识别准确率具有重要意义。本章实验使用 FSL 软件中的 BET 方法将所有实验数据做去除颅骨操作。

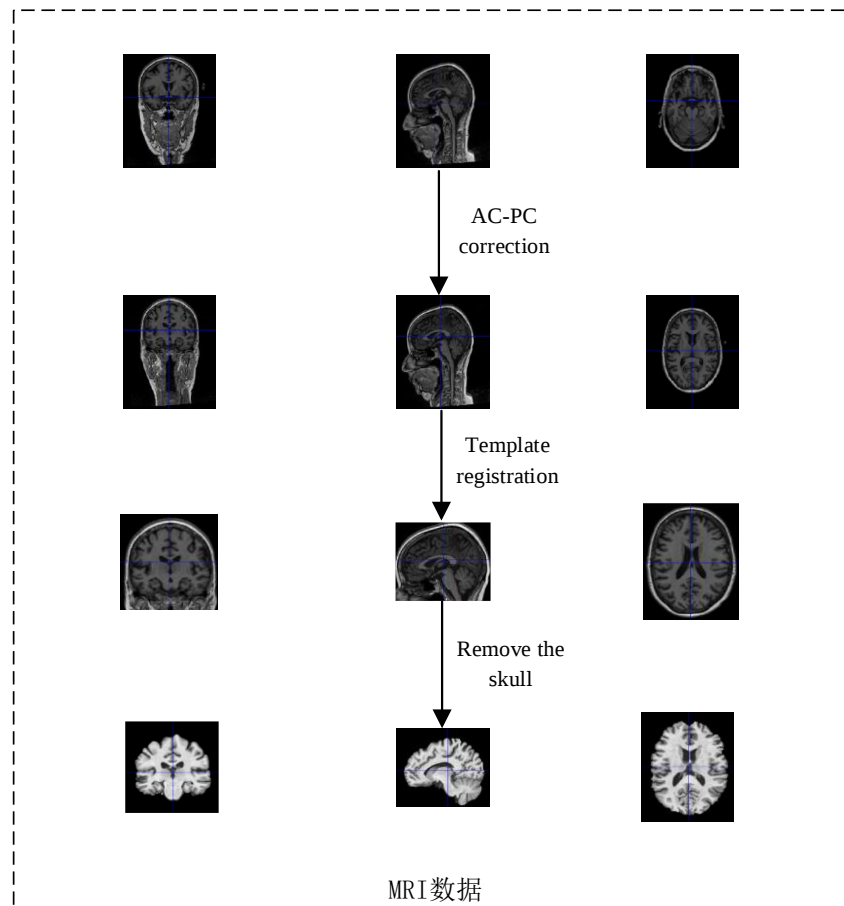


图 2-8 图像预处理步骤示意图

2.2.2 评价指标

本章实验使用图像分类任务中常用的准确率(Accuracy)、敏感性(Sensitivity)和特异性(Specificity)三个指标来评估模型，准确率指的是模型分类正确的样本数量在总体样本中的占比。敏感性指模型分类为正类样本中，真实正类样本的比例。特异性指模型分类识别为负类样本的数量占总负类样本数量的比值。上述指标的计算公式如下：

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (2-3)$$

$$Sen = \frac{TP}{TP + FN} \quad (2-4)$$

$$Spe = \frac{TN}{TN + FP} \quad (2-5)$$

其中， TP 、 TN 、 FP 、 FN 分别表示真正类、真负类、假正类和假负类。

2.2.3 消融实验

为了更好地展示改进后的 Mobilevit 网络在分类性能上的提升，以 Mobilevit 模型为基准网络，通过逐步叠加的方式，将 CBAM 模块和 RFB 模块分别集成到 Mobilevit 模型中，并在测试集上进行消融实验来评估各模块的作用和改进效果。实验结果通过 Accuracy 这一经典评价指标进行对比分析，消融实验的对比结果如表 2-1 所示。

表 2-1 消融实验结果

Baseline	CBAM	RFB	Accuracy(%)	准确率提升(%)
✓			95.24	
✓	✓		96.26	1.02↑
✓	✓	✓	96.43	0.17↑

由表 2-1 可知，在基准网络的基础上引入 CBAM 模块后，模型的分类准确率提升了 1.02 个百分点。这一结果表明，将通道注意力机制与空间注意力机制结合到主干网络的局部建模模块中，可以显著增强模型对有价值的通道和空间信息的关注，从而进一步提高模型的特征表达能力。在此基础上，通过在 Mobilevit Block 中添加 RFB 模块，分类准确率提升了 0.17 个百分点。这一提升说明，多尺度分支模块的引入有效扩展了模型的感受野，使其能够捕获更多的特征信息，从而增强了模型对不同类别 MRI 图像的特征提取能力。

此外，为了更直观地展示改进后的 Mobilevit 网络在分类性能上的提升，本章实验绘制了模型改进前后分类结果的混淆矩阵，如图 2-9 和图 2-10 所示。混淆矩阵清晰地展示了分类模型在不同类别上的分类准确率及错误分类分布，有助于深入分析改进后的网络模型在不同分类任务中的表现与潜在优化方向。

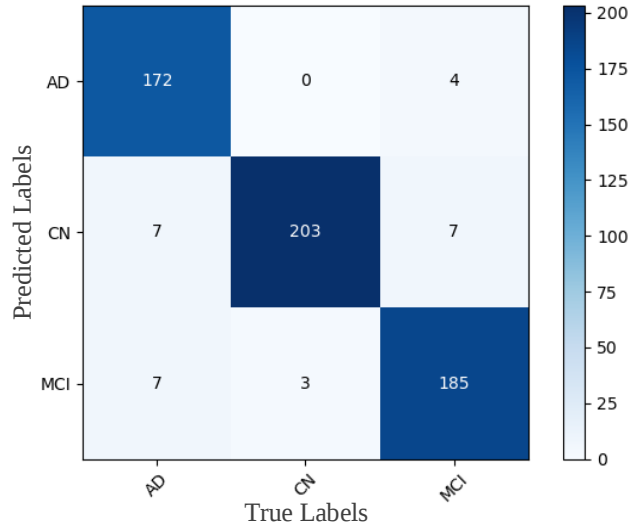


图 2-9 原模型分类混淆矩阵图

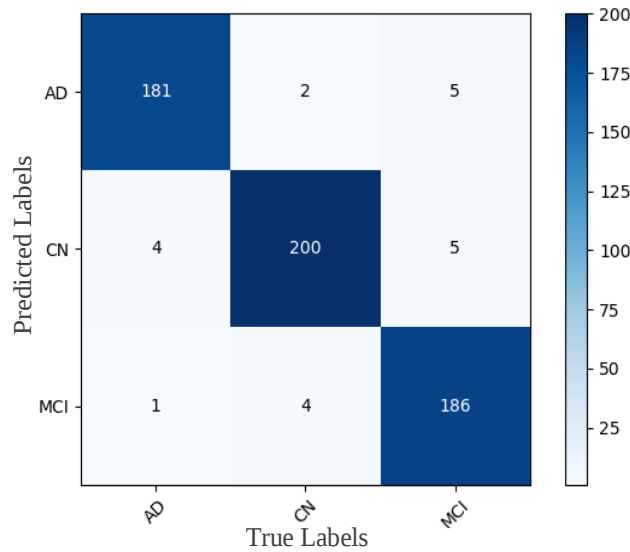


图 2-10 改进的 Mobilevit 网络分类混淆矩阵图

由上图可以看出，改进前后的网络模型对每个类别的图像均未实现完全正确分类。然而，改进后的 Mobilevit 网络模型在 AD 和 MCI 类别的分类任务中，识别样本的正确率相比原始模型有显著提升。在 CN 类别的分类中，尽管识别的错误样本数略多于原始模型，但整体识别正确的样本数仍然高于原始模型。

这一结果表明，改进后的分类模型能够有效提升模型的分类性能，特别是在复杂类别（如 AD/MCI）的识别任务中表现出了更高的可靠性与准确性，进一步验证了本章方法的可行性与优势。

2.2.4 对比实验

为了证明改进模型的优秀分类性能，将改进后的模型与原模型以及 MobilenetV2^[43]、MobilenetV3^[44]、Restnet50^[45]和 Efficientnet^[46]进行了对比实验。所有的模型训练都采用相同的训练参数，不加载预训练权重。各模型在该数据集上的分类性能指标如表 2-2 所示。

表 2-2 Mobilevit 与其他模型分类性能与参数量对比结果

模型	Accuracy/%	Sensitivity/%	Specificity/%	Params/MB
Mobilevit	95.24	95.13	97.57	7.63
MobilenetV2	89.29	89.17	94.60	8.50
MobilenetV3	94.89	94.97	97.47	6.14
Restnet50	94.22	94.17	97.10	89.70
Efficientnet	96.43	96.43	98.20	20.02
本章方法	96.43	96.43	98.20	8.32

根据表 2-2 的实验数据对比分析可知，在 AD/MCI/CN 三分类任务中，改进后的 Mobilevit 模型相较于基线模型展现出更加优秀的分类性能，其分类准确率、敏感性和特异性指标分别提高了 1.19、1.30 和 0.63 个百分点。这一结果表明，改进后的模型在保持模型诊断敏感性的同时，有效提升了分类模型的整体准确率与特异性，这对于医学图像分类中降低假阳性率和假阴性率具有重要临床意义。而改进后的模型由于增加了部分特征表达层，模型参数相较于基线模型略有提升，但改进后的模型在整体的分类性能指标上表现的更加优秀。与 MobilenetV2 和 MobilenetV3 两个轻量级的网络模型相比，在参数量相差不大的情况下，模型的分类性能均得到了较大提升。与经典的分类网络 Restnet50 相比较，改进后的 Mobilevit 网络准确率提升了 2.21 个百分点，而模型的参数量只为 Restnet50 网络参数量的 1/10。在分类性能与 Efficientnet 相当的情况下，改进后的 Mobilevit 网络模型的参数量下降了 140%。综上所述，改进后的 Mobilevit 模型能够在保持参数量小的同时，拥有较高的识别准确率。

为了进一步验证改进后的 Mobilevit 模型在 AD 分类领域的优越性能，将其

与同样使用 ADNI 数据集的其他文献方法进行比较，比较结果如表 2-3 所示。为了与其他文献保持一致，本实验选择 Accuracy 作为评价标准。

表 2-3 不同 AD 分类方法的对比结果

方法来源	分类目标	分类准确率
文献[47]	AD/CN	84%
文献[48]	AD/CN	83%
文献[49]	AD/MCI/CN	95.74%
文献[50]	AD/CN	88.4%
文献[51]	AD/CN	88.02%
本章方法	AD/MCI/CN	96.43%

从表 2-3 可以得出结论，目前在基于深度学习的 2D 切片图像分类领域中，模型的分类准确率达到 80% 以上。与其他文献的分类方法相比，本章方法的识别准确率更高，诊断 AD 更加准确，其分类准确率达到 96.43%，这验证了本章方法的优越性和普遍适用性，并表明本章方法在 AD 分类领域具有良好的应用前景。

2.3 本章小结

本章提出了改进的 Mobilevit 模型用于 AD 的核磁共振影像识别，通过对其主干网络和 Mobilevit block 进行改进，并在 ADNI 数据集上进行了实验对比。实验结果表明，在小幅增加模型参数量的情况下，模型的分类性能得到了大幅提升，在 AD/MCI/CN 分类任务中，识别的平均准确率达到 96.43%。相比较于其他轻量级网络，在参数量相当的情况下，模型的分类性能能够显著提高。与其他分类网络模型相比，在极小参数量的情况下，分类性能还能够显著提高。综上所述，本章方法可以更准确、更快速的识别出 AD 的类别，满足在节省算力的情况下，部署到医疗设备上。改进模型能够为智能医疗设备研究提供参考，辅助医生对 AD 进行快速诊断。在未来的工作中，将寻求更轻量、准确率更高的网络模型，促进轻量级的网络在医学图像处理领域的发展。

第3章 基于3D影像的阿尔兹海默症识别

3D 影像通过对脑部进行全方位扫描，能够提供更加丰富、连续的结构信息，为深度学习模型提供了更全面的特征输入。因此，使用基于 3D 影像数据进行 AD 诊断能够增加模型提取脑部病变的特征信息，从而增加识别的准确率。本章将在第二章的基础上，使用改进的 ShufflenetV2 网络模型^[52]对三维的 MRI 影像进行分类，同时结合注意力机制和特征融合机制来提高模型的识别能力，进一步实现更加准确的 AD 诊断。

3.1 基于改进的 ShufflenetV2 网络 AD 识别算法原理

基于 2D 影像的 AD 识别方法具有数据量小、模型结构简单等优点，但是其在 AD 早期识别准确率仍然不高，且由于数据集划分易存在数据泄露导致分类准确率虚高等问题。为了解决这些问题，本章提出一种基于 3D 影像的 AD 识别方法。使用基于 ShufflenetV2 网络的 AD 识别框架，利用 3D 卷积充分提取三维影像的特征信息，进一步提高模型识别的准确率。为了提高模型对重要特征通道的信息提取能力，引入 SE 注意力机制。此外，为了增加模型对三维影像全局特征的感知能力，在原模型的基础上引入一条全局特征融合分支网络，以保证模型提取更多有用的特征信息，从而进一步提高模型识别 AD 的准确率。

3.1.1 改进的 ShufflenetV2 主干网络

ShufflenetV2 网络的初衷是为了进一步提高在移动端和边缘计算设备上的深度学习模型的效率和性能。ShufflenetV2 网络主要由基本单元和下采样单元组成，输入的特征信息通过通道混洗机制，更有效地进行特征混合。在卷积层之间，它采用重新排列输入特征图通道的方式，以确保跨通道的信息能够充分交换，提高模型的表达能力。另外，ShufflenetV2 网络为了让信息在网络中有效传递，同时避免复杂度和参数数量的大幅增加，在主干网络中使用了一种轻量级的残差连接方式。但是由于原网络结构只适用于处理二维的平面图像，这对处理具有三维特征信息的医学核磁共振影像是友好的。所以，本章方法在原网络结构的基础上，使用 3D 卷积代替了原网络中的 2D 卷积，以提取三维影像病变相关的特征信息。另外，为增强模型对重要的特征通道的关注能力，在模型的第一层 3D 卷积和最后一层 3D 卷积之后引入了 SE 注意力机制，用来获取病变信

息的浅层特征。

此外，考虑到三维图像数据拥有丰富的边缘和纹理信息，这些信息对区分不同种类的疾病是至关重要的，而原网络结构中缺少对初期层全局特征信息的提取。在深度网络中，随着层级的增加，初期层的细节特征可能会逐渐丢失，而增加了全局特征融合分支模块后，网络模型初期层包含的高级语义信息可以与后期层的抽象特征信息结合起来，从而丰富了特征表示，有助于网络更好地进行分类任务。因此本章提出一种全局特征融合分支模块，将原始的三维影像数据通过一层 3D 卷积进行全局特征建模，将提取的全局特征与局部特征进行融合，以实现不同尺度的特征信息之间进行交互，改进后的 ShufflenetV2 网络结构如图 3-1 所示。

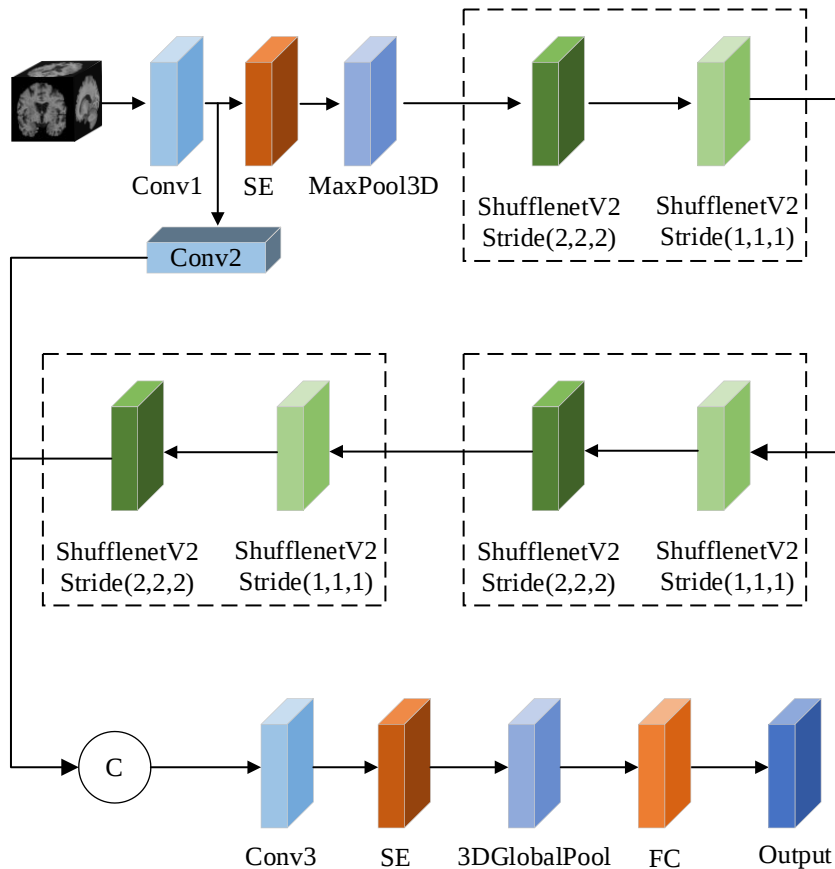


图 3-1 改进的 ShufflenetV2 网络结构图

3.1.2 基本单元和下采样单元

hufflenetV2 网络结构最主要的两个组成部分是基本单元和下采样单元，其中其本单元结构如图 3-2 所示，输入特征被通道分割模块平均分成原来通道数一

半的两个部分，其中一部分经过右侧的第一层卷积核为 $1\times 1\times 1$ 的三维卷积用来调整通道数。紧接着将特征输入到一个卷积核大小为 $3\times 3\times 3$ 的三维深度可分离卷积，在保证特征提取能力的基础上进一步减少模型参数和计算复杂度，然后将输出特征经过一个 3D 的逐点卷积混合不同通道的特征，从而增加通道间的特征交互。最后与另一半的辅助分支进行特征融合，形成一个更深的特征图，这样操作的目的是使模型能够维持足够的特征信息，为下一步的通道混洗提供更丰富的特征表示。

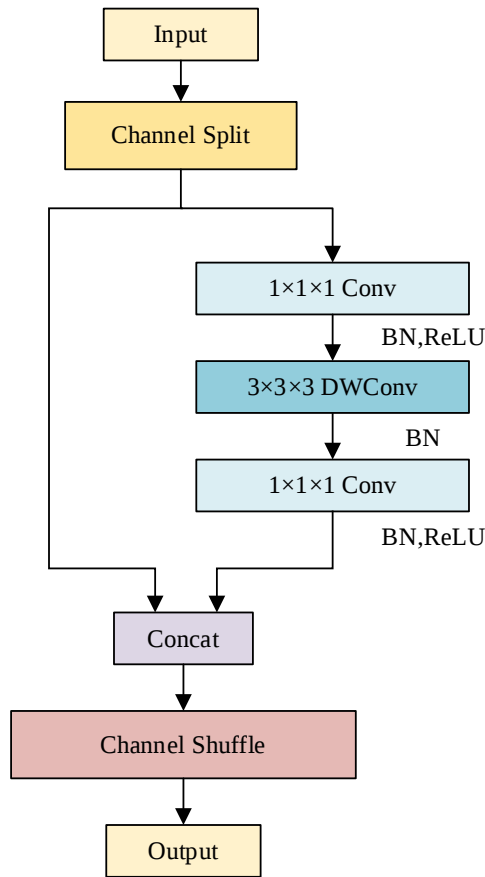


图 3-2 ShuffleNetV2 网络基本单元结构图

为了保证网络能够捕获更高层次的图像特征，ShuffleNetV2 网络中还使用了下采样单元结构，下采样单元结构如图 3-3 所示。下采样单元中在两个分支上均使用了步长为 2，卷积核为 $3\times 3\times 3$ 的深度可分离卷积，然后都经过一个卷积核为 $1\times 1\times 1$ 的卷积层调整通道数，最后再将两个分支的特征进行拼接操作。这样操作的目的是在减少参数数量和计算复杂度的同时，还能够保持网络对空间特征的有效提取。

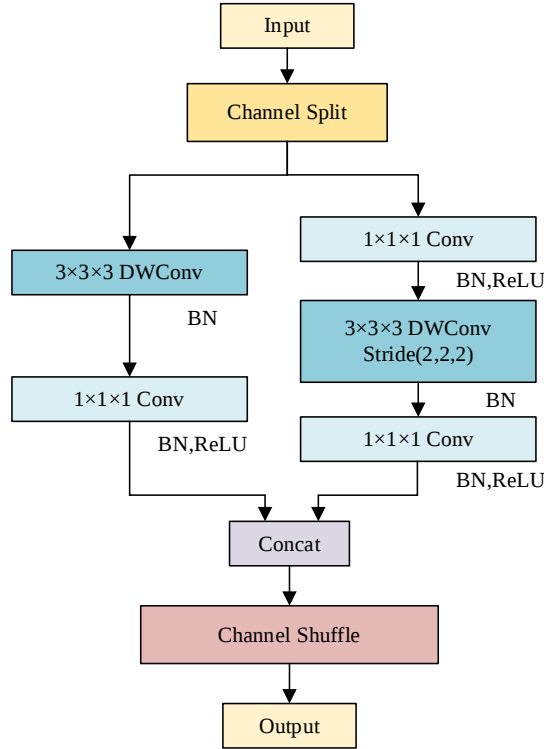


图 3-3 ShuffleNetV2 网络下采样单元结构图

3.1.3 通道混洗模块

在 ShuffleNetV2 网络中，通道混洗模块是一个关键技术，其主要是为了解决分组卷积导致的各组通道间信息无法交换的问题，该模块的主要作用是提高特征表示的多样性和网络模型的泛化能力。通道混洗操作通过打破网络中通道间的固定连接模式，促进不同通道之间的信息交互，其网络结构如图 3-4 所示，通道混洗操作主要有三个步骤：

- 1) 重塑：首先，将输出通道按组重塑为一个新的维度，即将形状为 $[N, C, H, W, D]$ 的特征图重塑为 $[N, G, C/G, H, W, D]$ 。
- 2) 转置：然后，对这个六维的张量进行转置操作，即交换第二维和第三维，得到 $[N, C/G, G, H, W, D]$ 的张量。
- 3) 平坦化：最后，将转置后的张量平坦化回 $[N, C, H, W, D]$ ，完成通道的混洗。

其中， N 表示批次大小， G 表示分组数量， H 表示特征图的高度， W 表示特征图的宽度， D 表示特征图的深度。

通过上述过程，原本属于同一组的通道被重新分布到不同的组中，从而实现了跨组的信息交流。通过通道混洗后，原始通道信息会被转移到新的位置，

这个新位置使得后续的分组卷积能够接触到来自所有原始组的信息。这种通道混洗和分组卷积的结合使用，在 ShufflenetV2 网络提取特征信息中非常关键。它既保持了模型的轻量级特性，又通过促进不同通道间的信息交流，提高了模型的性能和泛化能力。

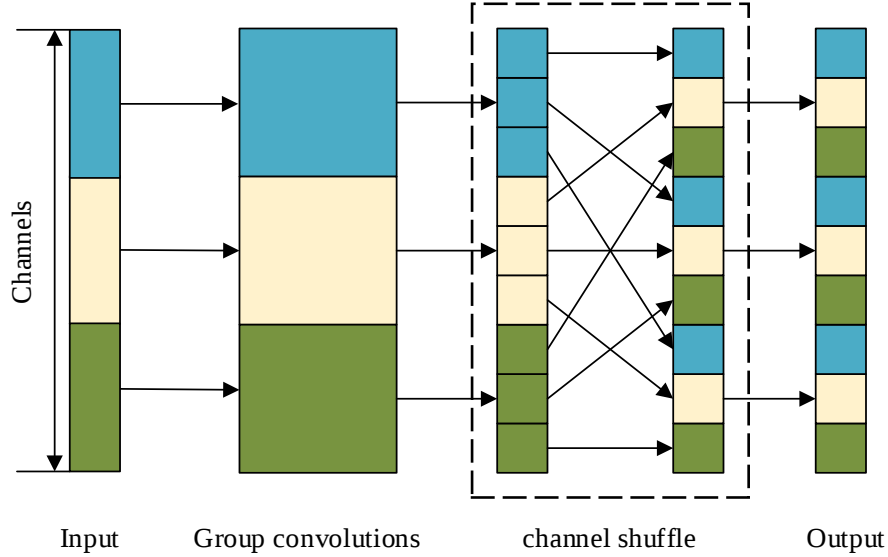


图 3-4 通道混洗示意图

3.1.4 Squeeze-and-Excitation 注意力机制

Squeeze-and-Excitation 注意力机制^[53]是一种增强网络特征表达能力的方法，它通过显式地建模通道间的依赖关系来提高网络的性能。其网络结构如图 3-5 所示，SE 注意力机制包含 Squeeze 和 Excitation 两个主要步骤。

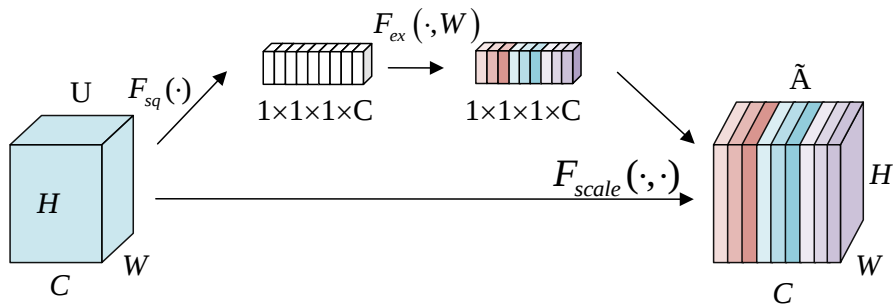


图 3-5 SE 注意力机制网络结构图

图中的 $F_{sq}(\cdot)$ 为 Squeeze 操作，输入的特征图 U 首先会被全局平均池化操作，生成一个 $1 \times 1 \times 1 \times C$ 的特征向量； $F_{ex}(\cdot, W)$ 为 Excitation 操作，通过两个全连接层对上一步挤压操作得到的特征向量进行处理，得到不同通道的权重值。 $F_{scale}(\cdot, \cdot)$ 表示将上一步生成的权重向量对特征图 U 进行权重赋值，最终输出具

有特征注意力的特征图 $\tilde{A}$ 。SE 注意力机制的核心思想是通过全连接网络根据损失自适应地调整通道的特征权重，而不是直接根据特征通道的数值分配来判断特征通道权重的大小，这样操作的好处是使网络能够重点关注更有用的特征。

3.2 实验分析

为了验证改进的 ShufflenetV2 网络模型在 3D 影像数据集上优异的分类性能，本章实验在 Pytorch 深度学习框架下进行，采用硬件配置为 Windows11 操作系统，Intel i5 CPU，NVIDIA GeForce GTX3060 显卡，12G 显存，编程语言为 Python 3.8 版本。为了使模型快速收敛，使用 AdamW 优化器对模型进行优化，学习率下降方式采用余弦退火衰减策略，损失函数使用交叉熵损失函数，训练迭代次数为 30 次，批量大小设置为 8。

3.2.1 实验数据集与预处理

本章所使用的实验数据均来自公开的数据集 ADNI，从 ADNI 数据集中随机抽取 816 位受试者脑部的 1.5T 类型 T1 加权的 MRI 数据，其中包括 187 个 AD 影像数据、400 个 MCI 影像数据和 229 个 CN 影像数据，数据格式为 NIFTI，所有实验数据按照第二章的预处理方法进行了 AC-PC 校正、模板配准和去颅骨操作。所有的受试者都经过了基本临床评估，并经过专业医生诊断，确保实验数据准确，数据集的详细信息如表 3-1 所示。实验过程中将数据集按照训练集：验证集：测试集比例为 7：2：1 进行随机划分。此外，为了防止数据泄露，数据集按照患者个体进行划分，输入数据的大小为 $80 \times 80 \times 80$ 。

表 3-1 数据集详细信息

类别	数量	年龄	性别	
			男	女
AD	187	75.4(7.5)	98	89
MCI	400	75.9(5.1)	257	143
CN	229	74.8(7.4)	110	119
总计	816	75.4(6.9)	465	351

3.2.2 消融实验

本章实验的分类任务分为两种，分别是二分类(AD/MCI, CN/MCI)和三分类(AD/CN/MCI)。为验证本章方法的有效性，首先在 AD/MCI 二分类任务上进

行了消融对比实验，以 ShufflenetV2 网络模型为基准模型，逐步在主干网络中添加 SE 模块和全局特征融合分支网络模块，通过层层叠加的方式，探讨各模块对模型分类性能的提升效果。选取 Accuracy 作为评价指标，实验结果如表 3-2 所示。

表 3-2 消融实验结果

Baseline	SE	Feature fusion	Accuracy(%)	准确率提升(%)
✓			91.38	
✓	✓		93.11	1.73↑
✓		✓	94.83	3.45↑
✓	✓	✓	96.55	5.17↑

在表 3-2 中，Baseline 代表了未经任何改进的 ShufflenetV2 基础网络。根据表中的数据，引入 SE 注意力机制后，改进的 ShufflenetV2 网络在 AD/MCI 的分类任务中，准确率提升了 1.73 个百分点。这一改进得益于 SE 模块的通道加权机制，在网络前向传播过程中，通过全局平均池化获取特征通道的全局信息，再通过两个全连接层构建通道间的非线性关系，最终生成通道注意力权重。这种自适应特征校准机制使模型能够强化对 AD 影像中早期病理特征的敏感度，同时抑制非关键区域的特征响应。实验结果表明，注意力机制特别增强了网络对浅层特征的捕获能力，这对识别早期 AD 的细微病理改变至关重要。

在引入全局特征融合模块之后，改进后的 ShufflenetV2 网络的分类准确率相较于原模型提高了 3.45 个百分点。这一提升充分证明了在深度网络中，初期层所提取的细节特征对于 AD/MCI 分类任务具有重要价值。当同时整合 SE 注意力机制与全局特征融合模块时，改进后的 ShufflenetV2 网络展现出显著的协同效应，最终分类准确率达到 96.55%，较基础模型提升 5.17 个百分点。这种性能提升源于两个模块的互补作用。SE 模块通过通道维度的动态特征重标定优化了特征质量，而全局融合模块则从空间维度实现了多尺度特征的有机整合。这表明，经过综合改进后的模型在分类性能上实现了显著提升，具备更高的实际应用潜力。

为直观地展示改进的 ShufflenetV2 网络模型在分类性能上的提升，本章实验绘制了消融实验的混淆矩阵，如图 3-6 至 3-9 所示。通过混淆矩阵可清晰地观察到模型在引入不同改进模块后的分类效果，进一步验证改进措施的有效性。

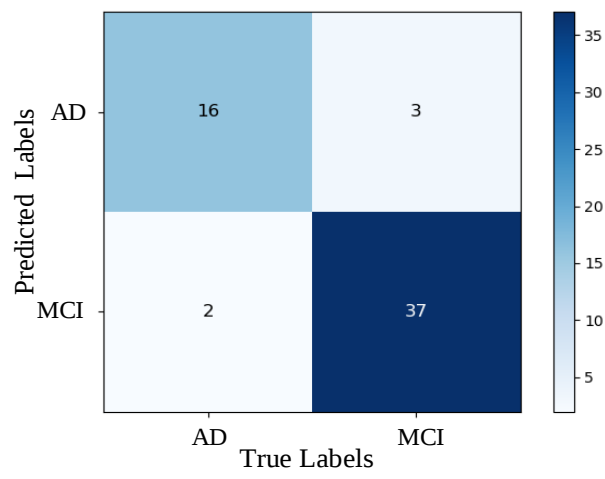


图 3-6 原模型分类混淆矩阵图

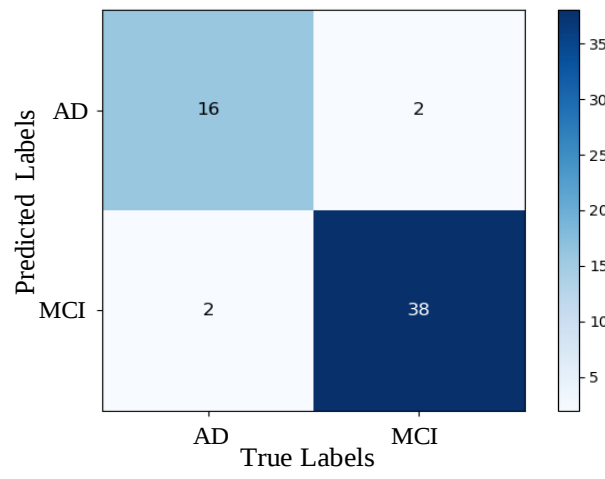


图 3-7 原模型 + SE 分类混淆矩阵图

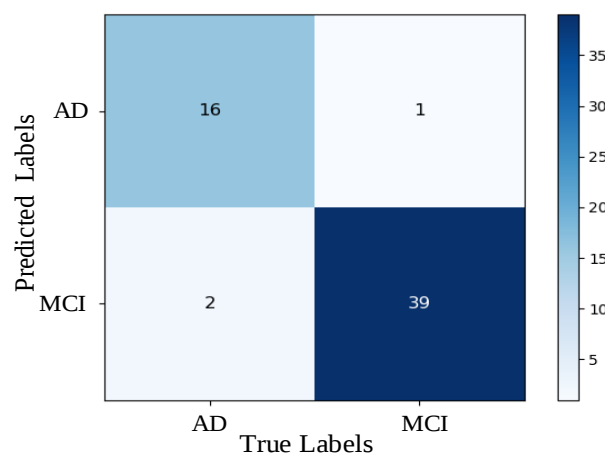


图 3-8 原模型 + Feature fusion 分类混淆矩阵图

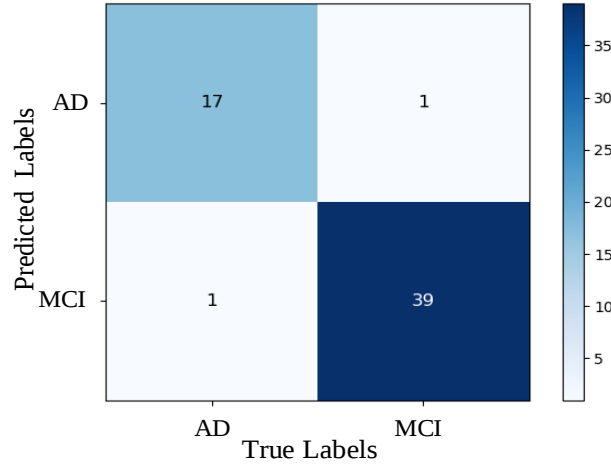


图 3-9 原模型 + SE + Feature fusion 分类混淆矩阵图

从图 3-7 中的混淆矩阵可以观察到，在基准模型的基础上添加 SE 注意力机制后，模型在 MCI 样本的识别能力方面得到了显著提升。具体而言，Baseline + SE 模型对 MCI 类别的识别正确数增加，同时错误分类为 AD 的样本数减少，表明 SE 模块通过引入通道间的权重分配机制，使得模型能够更加关注重要的通道特征信息，从而提升了对 MCI 类别的特征提取能力。由图 3-8 可知在原模型的基础上引入特征融合模块后，模型进一步提升了对 AD 和 MCI 样本的识别性能。从图中可以观察到引入了特征融合模块后的分类模型对 MCI 类别的识别正确数增加，对 AD 类别的误分类样本数进一步减少，这一实验结果体现出特征融合模块通过全局和局部特征的有效结合，提高了模型对复杂特征的感知能力，优化了特征表达。由图 3-9 的混淆矩阵分析可得，在结合 SE 模块与特征融合模块的改进模型中，整体分类性能实现了最优。这表明通过整合注意力机制和特征融合策略，改进后的模型在分类任务中具有更强的鲁棒性和更优的分类能力。综合以上实验结果可以得出结论，逐步引入的改进模块能够有效提升模型在 AD 和 MCI 类别上的分类性能，验证了本章方法在特征提取和分类任务中的可行性和优越性。

3.2.3 对比实验

为了进一步验证本章方法的优异性能，在 AD/MCI 数据集、CN/MCI 数据集和 AD/CN/MCI 数据集上，设置相同的实验条件与传统的分类网络模型 Resnet50、Densenet169^[54]和轻量化网络模型 Mobilenet^[55]、MobilenetV2 进行对比，其中在 AD/MCI 数据集上的对比结果如表 3-3 所示。

表 3-3 在 AD/MCI 数据集上与不同模型分类性能对比结果

Models	Acc/%	Sen/%	Spe/%
ShufflenetV2	91.4	88.9	92.5
Resnet50	94.8	88.9	97.5
Densenet169	94.8	88.9	97.5
Mobilenet	81.3	83.3	80.0
MobilenetV2	91.4	83.3	95.0
本章方法	96.6	94.4	97.5

从表 3-3 可以看出，在 AD/MCI 分类任务中，改进后的模型在分类准确率上提高了 5.2%，敏感性提高了 5.5%，特异性提高了 5%。实验结果表明，模型的整体分类性能得到了大幅提升。此外，与经典的分类网络 ResNet50 和 DenseNet169 相比，本章方法在分类准确率和敏感性方面均表现得更为优异，而在特异性方面则与这两个网络相当。这说明本方法在保持较高特异性的同时，进一步提升了分类准确率和敏感性，从而在综合性能上更具优势。而相较于两个轻量级的分类网络，本章方法由于采用了更深层次的特征提取网络，因此在分类准确率、敏感性和特异性这三个指标上均显著优于轻量级网络。这一结果充分证明了改进后的模型在特征提取和分类性能上的优越性，能够更有效地处理复杂的医学图像分类任务。

为了验证改进后的模型在其他 AD 病程的分类任务中的有效性，保持相同的实验条件在 CN/MCI 数据集上进行对比实验，实验结果如表 3-4 所示。

表 3-4 在 CN/MCI 数据集上与不同模型分类性能对比结果

Models	Acc/%	Sen/%	Spe/%
ShufflenetV2	92.1	82.6	97.5
Resnet50	93.7	87.0	97.5
Densenet169	95.2	95.7	95.0
Mobilenet	92.1	87.0	95.0
MobilenetV2	88.9	78.3	95.0
本章方法	95.2	95.7	97.5

由表 3-4 的数据可以看出，本章方法在 CN/MCI 分类任务中相较于其他分类网络仍表现出了优异的分类性能。与改进前的分类模型相比，特异性保持不变，但分类的准确率和敏感性分别提高了 3.1%和 13.1%。由此可见，改进后的模型

在 CN/MCI 数据集上保持了低误诊率的同时，显著提高了诊断的准确率和降低了漏诊率。而相比 Densenet169 网络，由于该网络模型使用了更深层次的网络，所以在分类准确率和敏感性两个分类指标上，改进后的模型没有明显提升，但是由于本章方法引入了特征融合模块和注意力机制，在特异性方面显著优于 Densenet169 网络。相比 Resnet50 网络，本章方法在特异性指标上与 Resnet50 网络相当，但是准确率和敏感性分别提高了 1.5%和 8.7%。由此可见本章方法在 CN/MCI 分类任务中，整体的分类性能优于表中的其他分类网络，证明了改进后的模型在 CN/MCI 数据集上分类的有效性。

以上两个实验在二分类任务的 AD/MCI 和 CN/MCI 数据集上证明了改进后模型分类性能的优越性，为了进一步证明本章方法在多分类任务中优异的分类性能，同样采取相同的实验条件，在 AD/CN/MCI 数据集进行对比实验，实验结果如表 3-5 所示。

表 3-5 在 AD/CN/MCI 数据集上与不同模型分类性能对比结果

Models	Acc/%	Sen/%	Spe/%
ShufflenetV2	70.4	64.8	87.2
Resnet50	75.3	66.0	87.6
Densenet169	71.6	61.9	86.3
Mobilenet	63.0	53.3	80.2
MobilenetV2	71.6	63.7	85.7
本章方法	75.3	77.5	87.4

由表 3-5 中的实验结果可知，在三分类任务中，本章方法相比其他分类模型表现出了优秀的分类性能。具体而言，本章方法与改进前的模型进行相比，在三个分类指标上均有提升，尤其是分类准确率和敏感性显著提升，这表明改进模块的引入使得模型对三种图像数据的识别精度大幅提升，且漏诊率大幅减小。相比 Densenet169、Mobilenet 和 MobilenetV2 三个分类网络，本章方法在三个分类指标上均达到最高，说明本章方法在整体性能上的领先。而相较于经典的 Resnet50 网络，本章方法在特异性略低于 Resnet50 网络，这得益于 Resnet50 网络的残差结构和较深的网络层次，能够获取更详细的浅层特征信息，因此 Resnet50 网络的误诊率更低。但是在敏感性指标上，本章方法显著优于 Resnet50 网络，这是因为本章方法融入了通道混洗机制和注意力机制，使得模

型更加关注有用的特征信息，从而进一步降低漏诊率。综上所述，改进后的模型在分类性能上得到了很大的提升，在当前主流的分类模型中分类性能也处于领先水平，以上实验结果证明了本章改进网络模型在三分类任务中的有效性和实用性。

为了进一步证明本章方法在 AD 分类领域中的优秀性能，将本章方法的实验分类性能结果与其他文章方法进行比较，选取通用的分类准确率作为评价指标，影像数据均使用 ADNI 公开数据集，对比结果如表 3-6 所示。

表 3-6 与经典深度学习模型对比结果

方法来源	Acc/%		
	AD/MCI	CN/MCI	AD/CN/MCI
文献[56]	85.3	91.8	—
文献[57]	70.5	85.9	62.5
文献[58]	78.0	81.0	—
文献[59]	76.6	79.4	—
本章方法	96.6	95.2	75.3

表 3-6 展示了本章方法与其他文章方法对比的结果，从表中可以看出在基于深度学习的 AD 分类工作中研究者已经取得了一定的成果。本章提出的方法在 AD/MCI 分类任务中，分类的准确率比文献[56]的方法高了 11.3 个百分点，而在 CN/MCI 分类任务中，分类的准确率比文献[56]高了 3.4 个百分点，在三分类任务中，本文的方法比文献[57]高了 12.8 个百分点。在三个分类任务中，本章所提出的改进方法的分类准确率均高于其他文献中的分类方法。综上所述，本章所提出的改进方法，在 AD 分类任务中具有较为先进的性能，具有一定的实际应用价值。

3.3 本章小结

针对基于二维影像的 AD 识别准确率不高、具有数据泄露、效率低等问题，本章提出了一个改进的 3D-ShufflenetV2 网络模型。首先，该网络利用 3D 卷积提取三维图像的空间连续信息，这样能够提取更多的脑部病变信息。其次，在主干网络中引入 SE 注意力机制，通过基于全局信息重新校准通道级特征响应以增加模型对重要特征的关注能力，从而增强模型的分类性能。最后，为了增加模型对全局特征提取的能力，在网络结构中增加全局特征融合分支模块，充分

利用图像的全局特征信息，从而进一步增加模型的分类能力。并通过消融实验和对比实验证明所改进模块的有效性和 AD 分类性能的先进性。实验结果证明，本章提出的改进网络能够实现对 AD、MCI 和 CN 的准确识别，改进后网络模型的分类性能比原网络有较大的提升，具有一定的实际应用价值。

第4章 基于多模态数据融合的阿尔兹海默症识别

早期 AD 的临床表现不明显,如果能够诊断出早期的 AD,患者可以及时的寻求专业医生的帮助,明确疾病的发展趋势,并给出对应的治疗方案,从而在早期 AD 病程中帮助患者提高生活质量和延缓疾病的发展趋势。基于单模态的影像分类方法已经取得了一定的成果,但单模态影像数据由于其病理特征的单一性,在 AD 早期阶段的诊断中准确率往往较低,这是由于单一模态的 MRI 或 PET 影像难以捕捉到全面的病理变化,导致其诊断能力受到限制。因此如果能够结合多模态影像数据特征进行 AD 诊断,通过融合不同模态的特征实现信息互补,则可以显著提升 AD 的诊断精度^[60]。本章将探讨基于多模态影像数据融合的 AD 识别方法,使用两个特征提取分支网络分别提取 MRI 影像和 PET 影像的特征,并将特征融合后再进行分类,以提高 AD 诊断的准确率。

4.1 基于多模态图像融合分类网络的 AD 识别算法原理

针对当前早期 AD 诊断准确率不高、基于单模态影像提取病变特征信息能力弱等问题,提出一种基于多模态图像融合分类网络的 AD 识别算法。首先,提出了一种新型的残差网络分支结构用来提取三维图像的病变特征。其次,利用改进后的残差网络作为特征提取器分别提取 MRI 影像和 PET 影像的特征,然后将两种影像特征在特征层融合后再进行分类。最后,为了使融合网络能更好地捕捉三维医学影像的空间位置与通道之间的关系,在主干融合网络中引入了坐标注意力机制。实验证明,改进后的融合网络在 AD/MCI/CN 分类任务中达到了较高的准确率,且改进后的网络模型与改进前的基础网络和基于单模态影像的 AD 分类方法相比,AD 各病程的分类准确率显著提升。

4.1.1 多模态图像融合分类网络

多模态图像融合分类网络由特征提取模块、特征融合模块和特征分类模块三部分组成,其总体框架如图 4-1 所示。特征提取模块由两个改进的特征提取器组成,MRI 图像和 PET 图像首先会分别经过一个卷积核为 7 的 3D 卷积层进行初步的局部特征提取,紧接着输出的特征图会经过一个坐标注意力机制(Coordinate Attention, CA)^[61],用于加强三维影像的空间位置与通道信息之间的关系,这对于提取具有复杂空间结构的三维医学影像特征信息尤为重要。然后

特征图会进入改进的残差网络进行进一步的特征提取，在特征提取器模块中，为了减少模型的过拟合风险，将模型的神经元从 32 维递增至 256 维。

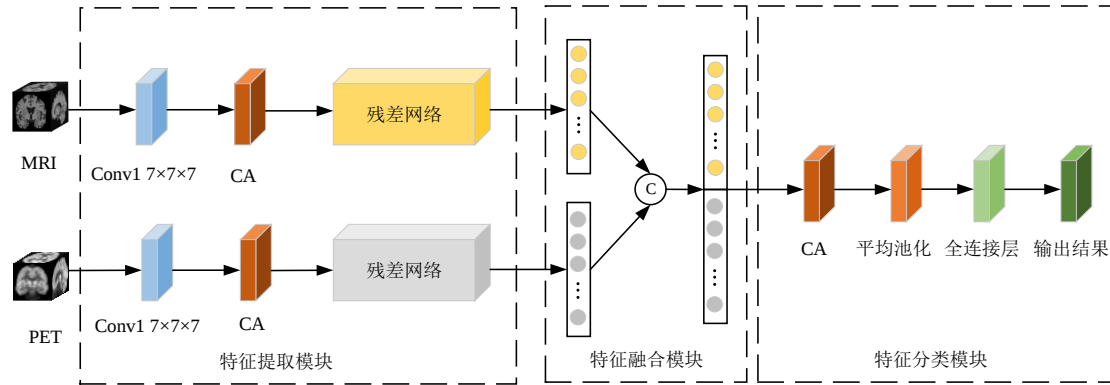


图 4-1 多模态图像融合分类网络结构图

在特征融合模块中，MRI 影像和 PET 影像分别含有不同的脑部特征信息。MRI 影像能够提供详细的脑部结构信息，而 PET 影像则能够提供脑部代谢活动的信息，这两种特征信息在识别 AD 方面具有十分重要的作用，通过特征拼接和维度扩展，可以降低单模态特征中的冗余信息，使得模型更加高效地利用输入数据^[62]。在高维特征空间的三维医学影像数据中，扩展特征维度可以增加特征空间的自由度，使得模型能够更好地捕捉数据中的复杂关系和模式。相比其他融合策略，如加权求和方法可能在特征较为冗余或某一模态特征较为复杂的情况下，导致信息丢失或不充分利用^[63]。而特征拼接通过直接扩展维度，使得每种模态的信息都能被充分保留，增强了模型对数据变化的适应性。与基于注意力机制的融合策略相比，虽然基于注意力机制的融合策略能够自动地为不同模态特征赋予不同的权重，从而突出重要模态，但这种方法依赖于注意力模型的设计和训练，可能在某些情况下无法充分捕捉到所有模态的细节信息^[64]。特征拼接通过直接连接不同模态的特征，不依赖于权重计算，而是通过硬连接直接扩展特征维度，使得模型对不同模态的变化更具适应性。这种方法能够减少在复杂数据集上的过拟合风险，尤其是在医学图像分析中，噪声和变异可能较大，避免了因注意力模型训练不当可能带来的性能损失。因此在特征融合模块中将两个特征提取器分支提取的 MRI 影像特征与 PET 影像特征进行拼接，神经元从 256 维扩展到 512 维。在特征融合模块输出一个拼接的多模态特征之后，使用坐标注意力机制对融合后的特征进行处理以进一步增强特征表示，使模型

能够捕捉更关键的特征，并抑制无关或冗余信息，从而提高特征表示的质量，最后再经过一个平均池化层和全连接层实现分类。

4.1.2 改进的残差网络

分支网络中的残差网络是在 Residual Network (Resnet)的基础上改进而来的，其网络结构如图 4-2 所示。改进的残差网络主要分为四层，每一层都由两个残差结构组成，用于提取输入图像的特征和进行特征学习。Layer1 是网络的初始层，用于初步的提取输入数据的初级特征。Layer2 在 Layer1 的基础上通过残差块中的卷积层和残差连接进一步提取特征。此外，该层包含步幅为 2 的卷积操作，用于下采样，从而减少特征图的空间维度，同时增加特征的抽象程度。Layer3 在 Layer4 重复 Layer2 的操作，通过更多的残差块，来学习更深层次的特征，这些深层次的特征对于识别复杂的医学图像数据至关重要。由于三维的医学影像数据具有丰富的空间信息，所以在第一层中的两个残差结构不对输入特征图做下采样处理，以保证保留影像更多的边缘、纹理等结构特征。而在后面三层结构中，每一层的第一个残差结构都进行了下采样处理，这样做的目的是在深层网络结构中逐渐减小特征图的大小，从而在深层次的网络中逐步聚合和压缩特征信息，以便提升模型的判断能力。

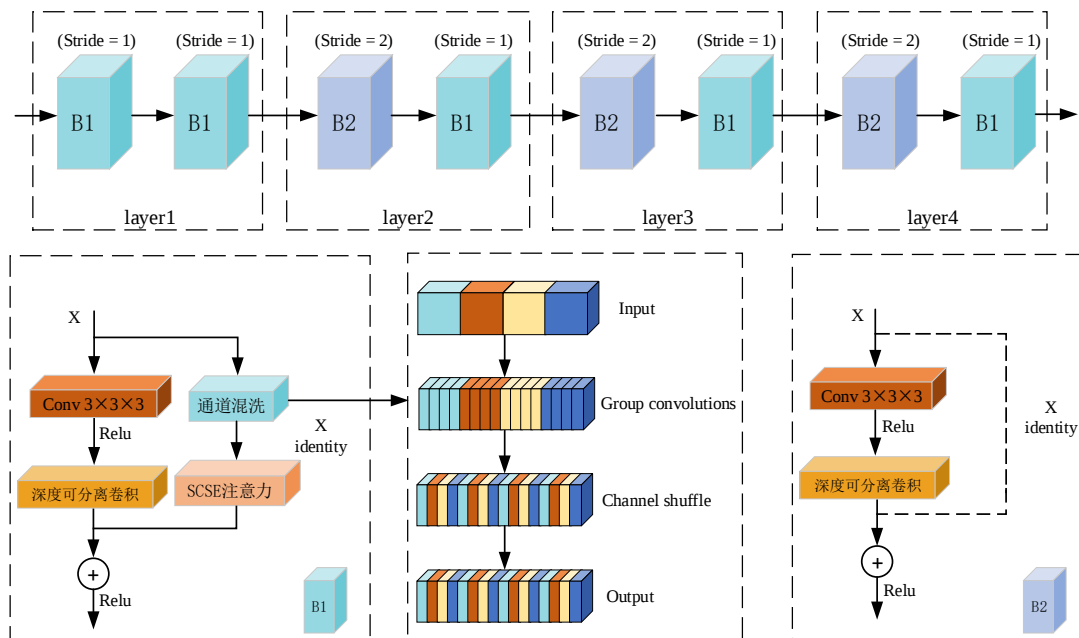


图 4-2 改进的残差网络结构图

在改进的残差网络结构中，第二层的普通卷积被替换为深度可分离卷积。

这一改进显著减少了模型的参数量，从而提高了计算效率。对于 3D 图像数据，由于输出通道数较大，深度可分离卷积在降低计算复杂度方面的优势尤为明显。此外，深度可分离卷积不仅有效减少了计算量，还能够精简特征，降低冗余度，进而使得提取到的特征更加简洁且具有更强的表达能力。深度可分离卷积通过深度卷积和逐点卷积，分步提取数据的空间特征和通道特征，可以更加灵活地捕捉不同尺度的信息。这种特性对于处理三维图像尤为重要，因为三维图像包含更多的空间层次和细节信息。

下面对深度可分离卷积在降低计算复杂度和减少模型参数方面做具体分析，假设输入图像的尺寸为 $H \times W \times D$ ，其中 H 、 W 、 D 分别表示图像的高度、宽度和深度，并且假设该层的输入通道为 C_{in} ，输出通道为 C_{out} ，卷积核大小为 $K_h \times K_w \times K_d$ ，在不考虑偏置的情况下，则标准的 3D 卷积计算复杂度为：

$$Flops = H \times W \times D \times K_h \times K_w \times K_d \times C_{in} \times C_{out} \quad (4-1)$$

其参数数量为：

$$Parameter = K_h \times K_w \times K_d \times C_{in} \times C_{out} \quad (4-2)$$

而深度可分离卷积的计算复杂度为：

$$Flops = H \times W \times D \times (K_h \times K_w \times K_d \times C_{in} + C_{in} \times C_{out}) \quad (4-3)$$

参数量可表示为：

$$Parameter = K_h \times K_w \times K_d \times C_{in} + C_{in} \times C_{out} \quad (4-4)$$

本章实验使用的图像数据大小为 $128 \times 128 \times 128$ ，假设输入通道为 64，输出通道为 128，卷积核大小为 $3 \times 3 \times 3$ ，则使用普通 3D 卷积的计算复杂度约为 $4.4 \times 10^9 Flops$ ，参数量约为 2.2×10^5 。而使用深度可分离 3D 卷积的计算复杂度约为 $2.1 \times 10^9 Flops$ ，参数量约为 0.9×10^5 。由此可见，相同大小的输入数据深度可分离 3D 卷积的计算复杂度和参数量约为普通 3D 卷积的一半，所以将普通 3D 卷积替换成深度可分离 3D 卷积能够进一步减少模型参数和计算复杂度。

当步长为 1 时，特征图将进行 B1 模块的处理，首先输入特征先经过一个 3D 卷积和深度可分离卷积的处理，然后原特征图经过通道混洗后，通过打乱特征图的通道排序再重组来获得不同通道之间的信息交互。重组后的特征再经过空

间和通道挤压激励(spatial and channel Squeeze & Excitation, scSE)^[65]注意力机制进行两个维度的特征加权，最后再与主干结构输出的特征进行相加。当步长不为 1 时，B2 模块会将特征图进行下采样，然后再与通过 3D 卷积和深度可分离卷积的特征图进行相加，输出通道变成原来的两倍，从而增加网络捕捉特征信息的能力。

4.1.3 scSE 注意力机制

scSE 注意力机制模块主要由两个子模块组成，分别是空间挤压激励(spatial Squeeze & Excitation block, sSE)和通道挤压激励(channel Squeeze & Excitation block, cSE)。scSE 将经过通道混洗后的特征图以并行的方式进行空间和通道双重特征加权，其网络结构如图 4-3 所示。scSE 注意力机制的核心思想是通过两个子模块对输入特征图的每个空间位置和通道进行加权，突出重要的位置特征和通道特征，抑制不重要的特征来增加模型的特征表达能力。使用 scSE 注意力机制结合三维特征信息的通道混洗技术，使改进后的残差模块能够学习到更有用的三维影像特征。

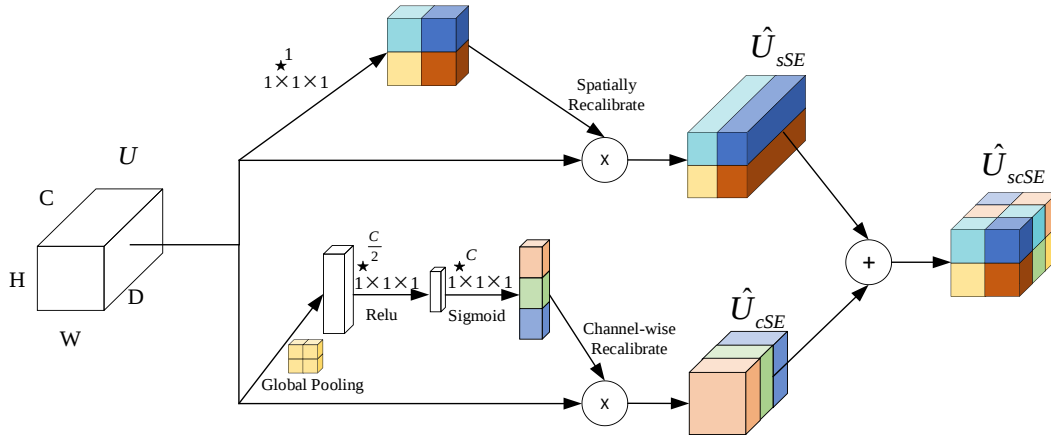


图 4-3 scSE 注意力机制结构图

图 4-3 中的上半部分是 sSE 部分，首先对特征图 U (假设特征图的尺寸为 $H \times W \times D$) 进行最大池化操作并进行特征维度压缩，得到每个位置的全局特征，提取出每个位置的全局空间信息 z ，然后再使用 Sigmoid 激活函数对这些全局空间特征进行非线性变换，生成每个位置的空间注意力权重 S ，最后再与原始的特征相乘进行空间重新校准，其特征向量的计算公式如下：

$$U = [u_{1,1,1}, u_{1,2,1}, \dots, u_{i,j,k}, u_{H,W,D}] \quad (4-5)$$

$$z_{H,W,D} = \text{conv}(U) \quad (4-6)$$

$$S = \sigma(z_{H,W,D}) \quad (4-7)$$

$$\hat{U}_{sSE} = S \cdot U \quad (4-8)$$

其中， $u_{i,j,k}$ 表示特征图 U 在坐标 (i, j, k) 处的通道特征， conv 表示卷积核为 $1 \times 1 \times 1 \times C$ 的卷积操作， σ 表示 Sigmoid 激活函数， S 表示空间注意力权重， $\hat{U}_{sSE}$ 表示输出的空间注意力特征。

图 4-3 中的下半部分是 cSE 部分，cSE 首先会通过全局平均池化操作对特征图的通道维度进行压缩，得到每个通道的全局特征 z ，这一步骤提取了每个通道的全局语义信息，其特征向量计算如公式(4-9)所示。

$$z_q = \frac{1}{H \times W \times D} \sum_{i=1}^H \sum_{j=1}^W \sum_{k=1}^D u_q(i, j, k) \quad (4-9)$$

其中， z_q 表示特征图第 q 个通道被全局平均池化操作后的值， i, j, k 表示的是该通道的空间位置坐标， u_q 表示的是该通道特征图的值。

在上一步的全局语义信息提取之后，将提取的特征经过两次全连接层处理。经过第一层的全连接层之后，使用 Relu 激活函数进行神经元激活，通道数变为原来的一半，以帮助模型学习更复杂的特征关系。然后再经过第二个全连接层，使用 Sigmoid 激活函数进行激活，重新将特征映射为原始通道数，得到每个通道的重要性权重。最后再与原始特征相乘进行校准，其计算公式如下所示：

$$\hat{U}_{cSE} = U \cdot \sigma(W_2 \delta(W_1 Z)) \quad (4-10)$$

其中， $\sigma(\bullet)$ 表示 Sigmoid 激活函数， $\delta(\bullet)$ 表示 Relu 激活函数， W_1 表示第一个全连接层的权重矩阵， W_2 表示第二个全连接层的权重矩阵， $\hat{U}_{cSE}$ 表示输出的通道注意力特征。

最后将 sSE 和 cSE 输出的特征进行拼接，得到最终的特征，输出的特征在空间和通道两个维度对特征进行相加，从而增强模型对重要特征的关注，提高模型的性能。

$$\hat{U} = \hat{U}_{sSE} + \hat{U}_{cSE} \quad (4-11)$$

其中， $\hat{U}$ 表示拼接后的特征。

4.1.4 坐标注意力机制

在基于脑部医学影像识别 AD 的深度学习网络模型中，三维医学影像数据包含丰富的空间结构特征信息，对提升识别的准确率起着至关重要的作用。因此，本章方法在特征提取网络中引入了 CA 注意力机制，以进一步增强模型的特征表达能力。CA 注意力机制是一种通过显式引入空间坐标信息来优化特征提取的注意力机制，其独特之处在于能够同时捕获通道信息和空间位置信息，从而更精确地定位关键区域并突出具有诊断意义的特征。CA 注意力机制在捕获全局依赖和空间关系的同时，具有更低的计算复杂度，更适合三维医学影像的特征建模，其网络结构如图 4-4 所示。

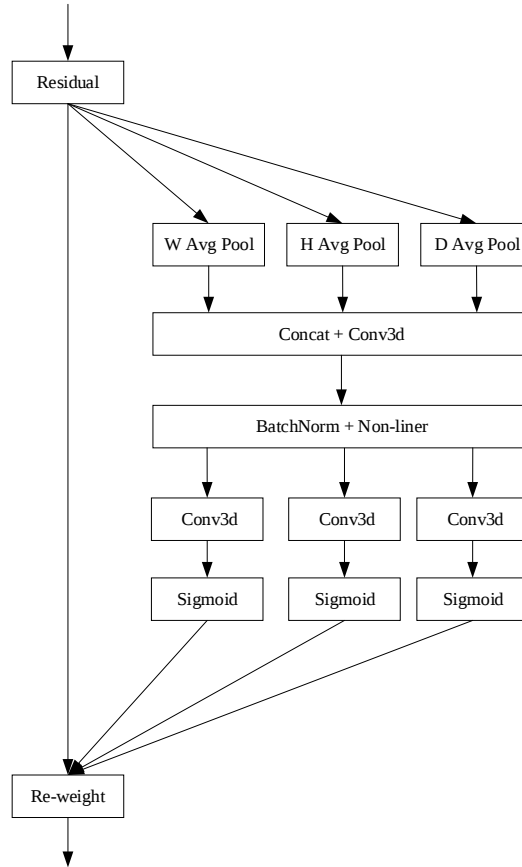


图 4-4 CA 注意力机制结构图

假设输入特征图 X 的尺寸是 $H \times W \times D \times C$ ，CA 首先对输入的特征图沿深度、高度和宽度三个方向进行全局平均池化，这样有利于模型捕捉到特征图中的长程依赖关系，从而提高模型对 AD 病变信息的提取能力。其具体公式如下：

$$Z_c^d = \frac{1}{H \times W} \sum_{0 \leq h \leq H} \sum_{0 \leq w \leq W} X_c(d, h, w) \quad (4-12)$$

$$Z_c^h = \frac{1}{D \times W} \sum_{0 \leq d \leq D} \sum_{0 \leq w \leq W} X_c(d, h, w) \quad (4-13)$$

$$Z_c^w = \frac{1}{D \times H} \sum_{0 \leq d \leq D} \sum_{0 \leq h \leq H} X_c(d, h, w) \quad (4-14)$$

其中， Z_c^d ， Z_c^h ， Z_c^w 分别表示第 c 个通道沿深度、高度和宽度三个方向进行全局平均池化后的特征。

然后将三个方向上生成的特征沿着不同方向进行拼接，最后再通过一个卷积核为 $1 \times 1 \times 1$ 的 3D 卷积和激活函数进行通道降维，具体公式如下：

$$y = \delta(\text{Conv}[Z_c^d, Z_c^h, Z_c^w]) \quad (4-15)$$

其中， Conv 表示卷积核为 $1 \times 1 \times 1$ 的 3D 卷积， $[\cdot]$ 表示特征拼接操作， δ 表示 Relu 激活函数， y 表示输出特征。

在上述特征降维操作之后，再将特征进行批量归一化操作，然后将归一化后的特征沿着不同的方向进行通道升维和分离处理，并使用 Sigmoid 激活函数对特征进行加权生成三个维度的独立特征，最后与输入特征逐元素相乘得到最终的特征。其具体公式如下：

$$g^d = \sigma(\text{Conv}(y^d)) \quad (4-16)$$

$$g^h = \sigma(\text{Conv}(y^h)) \quad (4-17)$$

$$g^w = \sigma(\text{Conv}(y^w)) \quad (4-18)$$

$$G = X \times y^d \times y^h \times y^w \quad (4-19)$$

其中， g^d ， g^h 和 g^w 表示在三个维度上的独立特征， σ 表示 Sigmoid 激活函数， X 表示原始输入特征， G 表示输出特征。

4.2 实验分析

本章共进行了四组分类实验，分别是二分类(AD/MCI, CN/MCI, AD/CN)和

三分类(AD/CN/MCI)，并在 AD/CN/MCI 分类任务中进行了消融实验来验证改进模型的有效性。为了证明本章方法在 AD 识别领域的先进性能，在 AD/CN/MCI 数据集上与经典的分类模型进行了对比实验。

实验环境为 Pytorch 深度学习框架，采用的硬件配置为 Intel i5 CPU，NVIDIA GeForce GTX3060 显卡，12G 显存，编程语言为 Python 3.8。为使模型快速收敛，使用 AdamW 优化器对模型进行优化，学习率下降方式采用指数衰减策略，衰减因子设置为 0.9，训练迭代次数为 50 次。本章实验所使用的评价指标为 Accuracy、Sensitivity、Specificity 和曲线下面积(Area Under the Curve)四个指标。前三个评价指标已经在 2.2.2 小节详细介绍，曲线下面积衡量了模型在所有可能的分类阈值下区分正例和负例的能力，其计算公式如下：

$$AUC = \frac{\sum_{ins_i \in positiveclass} rank_{ins_i} - \frac{M \times (M+1)}{2}}{M \times N} \quad (4-20)$$

其中， M, N 分别代表正样本和负样本的个数， $rank_{ins_i}$ 表示第 i 个样本的序号。

4.2.1 实验数据集与预处理

本章实验所使用的图像数据均来自于 ADNI 数据库，从中选取 67 岁~76 岁的 561 位受试者脑部的 MRI 影像和 PET 影像，其中包括 244 个 AD 患者、108 个 CN 正常对照组和 209 个 MCI 患者，数据格式为 NIFTI。所有实验数据按照第二章的预处理方法进行了 AC-PC 校正、模板配准和去颅骨操作，且将所有图像输入数据大小裁剪成 $80 \times 80 \times 80$ 。所有的受试者都经过了基本临床评估，并经过专业医生诊断，确保实验数据准确，数据集的详细信息如表 4-1 所示。在构建实验数据集时，将数据按照患者个体划分成训练集、验证集和测试集，保证同一患者的 MRI 数据和 PET 数据只在三种数据集中出现一次，防止数据泄露，且训练集：验证集：测试集比例为 7：2：1。

表 4-1 数据集详细信息

类别	数量	年龄	性别	
			男	女
AD	244	75.1(7.8)	143	101
MCI	209	75.4(4.9)	138	71
CN	108	74.9(7.4)	65	43
总计	561	75.1(6.7)	346	215

4.2.2 消融实验

为了证明改进后的残差网络分支结构的有效性，使用多模态融合网络在 AD/CN/MCI 数据集进行了消融实验，逐步添加 CA 注意力机制、scSE 注意力机制和通道混洗机制，以分类准确率作为评价指标，实验结果如表 4-2 所示。

表 4-2 消融实验结果

Baseline	CA	scSE	Channel Shuffle	Accuracy(%)	准确提升(%)
✓				83.93	
✓	✓			85.71	1.78↑
✓	✓	✓		87.50	1.79↑
✓	✓	✓	✓	91.07	3.57↑

由表 4-2 的实验结果分析可知，在基础残差网络中引入 CA 注意力机制后，模型的分类准确率提升了 1.78 个百分点，表明 CA 注意力机制有助于提取更多三维影像数据的特征信息，从而提高分类准确率。在残差分支网络中加入 scSE 注意力机制后，相比仅增加 CA 注意力机制，准确率提升了 1.79 个百分点，实验结果说明 scSE 注意力机制使模型能够同时关注影像数据的空间和通道特征信息，进一步提升分类性能。此外，在引入通道混洗技术后，模型的分类准确率比增加双注意力机制的模型提高了 3.57 个百分点，这归功于通道间信息交互能力的增强。综上所述，改进后的残差网络模型能更准确地识别 AD，有效性和实用性得到了充分证明。为直观展示改进模型在分类任务上的性能提升，本章实验绘制了消融实验的混淆矩阵，如图 4-5 至图 4-8 所示。

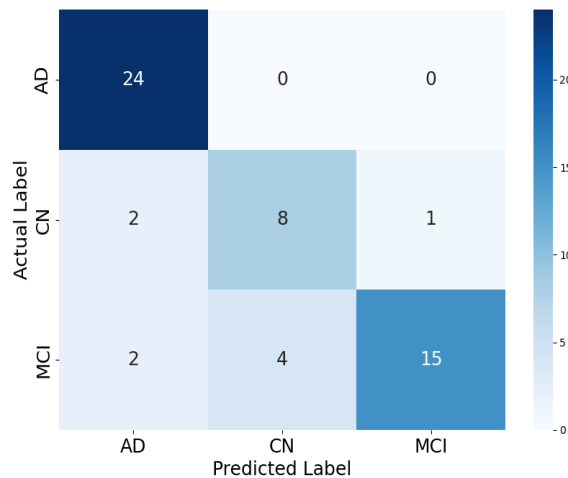


图 4-5 原模型分类实验混淆矩阵图

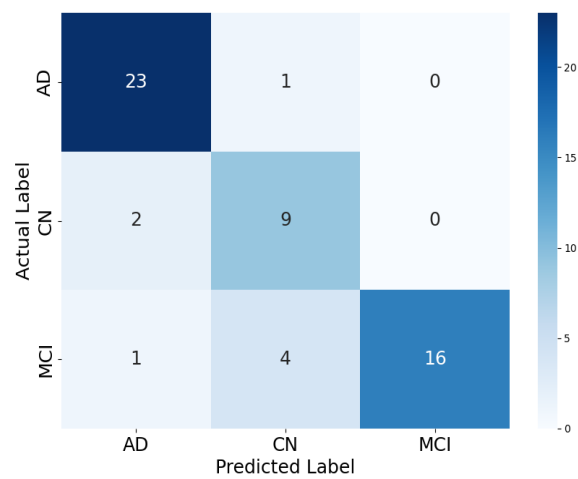


图 4-6 原模型 + CA 分类实验混淆矩阵图

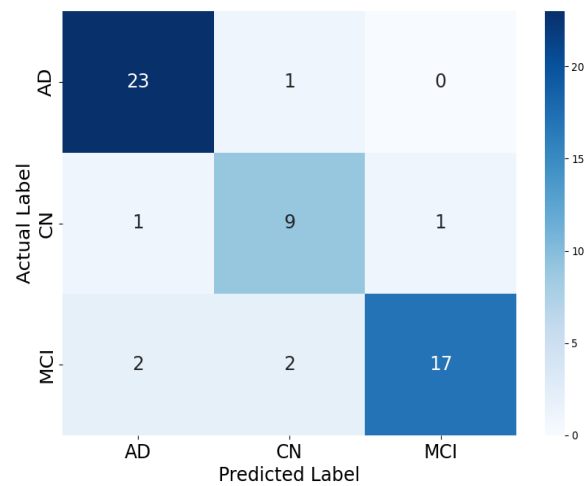


图 4-7 原模型 + CA + scSE 分类实验混淆矩阵图

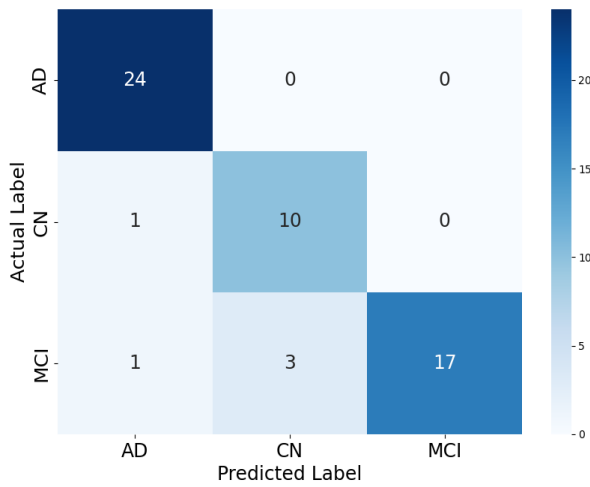


图 4-8 原模型 + CA + scSE + Channel Shuffle 分类实验混淆矩阵图

从上图中可知,通过引入不同改进模块后,模型的分类性能得到了显著提升。当加入CA注意力机制后,模型对CN和MCI类别的识别准确率明显提高,表明CA注意力机制模块能够有效利用空间坐标信息,增强模型对特征的表达能力。在进一步引入scSE注意力机制模块后,模型对MCI类别的识别能力进一步增强,不仅正确识别的样本数量增加,误检的数量也显著减少,这表明scSE模块能够在通道和空间特征之间实现更加精准的特征选择和增强。此外,当在网络中增加通道混洗技术后,模型的整体分类性能得到了进一步提升。通道混洗技术通过优化特征的流通和交互,提高了特征表示的效率,最终使分类正确的样本数量显著增加,漏检和误检的数量进一步减少。在基础模型中,CN类别被误分类为MCI的样本较多,这与临床诊断中CN和MCI类别存在部分特征重叠或差异不显著的现象一致。通过结合注意力机制和通道混洗技术,有效增强了模型对特征的提取与区分能力,从而降低了CN的分类错误率,并提升了MCI类别的分类准确率。

4.2.3 对比实验

为了验证本章改进方法在AD分类中具有较高的分类准确率,分别在基于单模态影像数据和多模态影像数据集上进行二分类(AD/MCI, CN/MCI)实验和三分类(AD/CN/MCI)实验,并与基础的Resnet18网络作为特征提取器的分类网络进行对比实验,具体的实验结果如表4-3所示。

表 4-3 不同网络模型基于不同模态数据的分类准确率(%)对比结果

数据模态	主干网络	AD/CN	AD/MCI	CN/MCI	AD/CN/MCI
MRI	Resnet18	80.00	75.56	75.00	64.29
	本章方法	85.71	80.00	84.38	69.64
PET	Resnet18	91.43	88.89	84.38	76.79
	本章方法	94.29	91.11	87.50	78.57
MRI+PET	Resnet18	94.29	93.33	87.50	83.93
	本章方法	97.14	95.55	90.63	91.07

由表4-3可以得出结论,本章改进的主干网络在基于单模态影像数据的分类任务中准确率相较于基础的Resnet18网络均有明显提升。本章方法在基于MRI单模态影像数据分类的准确率相较于Resnet18在AD/CN分类中,提升了5.71%,在AD/MCI分类任务中提升了4.44%,在CN/MCI和AD/CN/MCI分类任务中分

别提升了 9.38%和 5.35%。由上述分析可知，改进后的残差网络能够提取更多的 MRI 图像特征，从而提高模型对 AD 识别的准确率。

在基于单模态的 PET 数据分类中，本章方法仍表现出了优于原网络模型的性能，在四种分类任务中，准确率均高于原网络。证明本章方法不仅能提取具有丰富结构信息的 MRI 影像特征，也能有效提取具有生物代谢信息的 PET 影像特征。而且在四种分类任务中，基于 PET 影像分类的准确率是高于基于 MRI 影像的，这也符合临床诊断的特点。

在 AD/CN、AD/MCI、CN/MCI 和 AD/CN/MCI 四个分类任务中，使用改进后的残差网络作为特征提取器进行多模态数据融合分类时，准确率均高于基于 Resnet18 网络，准确率分别提高了 2.85%、2.22%、3.13%和 7.14%。综合四组对比实验结果，改进后的残差网络具备更强的特征提取能力，有效提升了模型的分​​类准确率。

在四个分类任务中，基于多模态融合分类方法的分类准确率均优于基于单模态数据分类方法。在三个二分类任务中，多模态融合分类方法的准确率相较于基于 MRI 影像分类方法分别提高了 17.14%、19.99%和 15.63%，在三分类任务中提高了 26.78%。与基于 PET 影像分类方法相比，多模态融合分类方法在四个任务中分别提高了 5.71%、6.66%、6.25%和 14.28%。这表明，多模态数据融合分类方法显著提升了分类准确率，尤其在三分类任务中优势明显。实验结果证明，本章方法能有效融合两种模态数据的特征信息，充分利用其互补性，在 AD 多种病程分类中表现出色，相较于传统单模态数据分类方法具有显著优势。

为了进一步证明本章方法在 AD 识别方面的有效性，对基于单模态 MRI 数据分类方法和基于单模态 PET 数据分类方法在 AD/CN、AD/MCI 和 AD/CN/MCI 分类中 AD 的敏感性和特异性与基于多模态融合方法进行比较，所使用的特征提取器均为改进后的残差网络，具体的实验结果如表 4-4 所示。

表 4-4 基于不同模态数据分类的 AD 敏感性和特异性对比结果

数据模态	AD/CN		AD/MCI		AD/CN/MCI	
	Sen%	Spe%	Sen%	Spe%	Sen%	Spe%
MRI	91.67	72.73	75.00	85.72	79.17	81.25
PET	95.83	90.91	95.83	85.71	87.50	93.75
MRI+PET	100.00	90.91	95.83	94.24	100.00	93.75

由表 4-4 可以看出, 基于多模态数据融合分类方法在三个分类任务中, 对 AD 识别的敏感性和特异性均表现出了优异性能。在 AD/CN 分类中, 基于多模态数据融合分类方法较基于单模态 MRI 数据分类方法 AD 敏感性提高了 8.33%, 特异性提高了 18.18%。相比基于单模态 PET 数据分类方法 AD 敏感性提高了 4.17%。在 AD/MCI 分类中, 多模态数据融合分类方法相较于两种单模态数据分类方法, AD 特异性分别提高了 8.52%和 8.53%。AD 敏感性相较于基于单模态 MRI 数据分类方法提高了 20.83%。而在三分类任务中, 多模态数据融合分类方法对 AD 分类的敏感性和特异性仍表现出了优异的性能, 相比较于两种单模态数据分类方法, AD 敏感性分别提高了 20.83%和 12.5%, AD 特异性相比基于单模态 MRI 数据分类方法提高了 12.5%。从以上分析可知, 多模态数据融合分类方法对 AD 的敏感性和特异性优于基于单模态数据分类方法, 这意味着本章提出的方法能够最大限度地避免漏诊和误诊, 这对于实际临床应用和患者管理来说都是至关重要的。

为了进一步说明多模态融合分类方法在分类性能上优于单模态数据分类方法, 使用 ROC 曲线对模型进行评估, 具体的评估结果如图 4-9 至 4-12 所示。

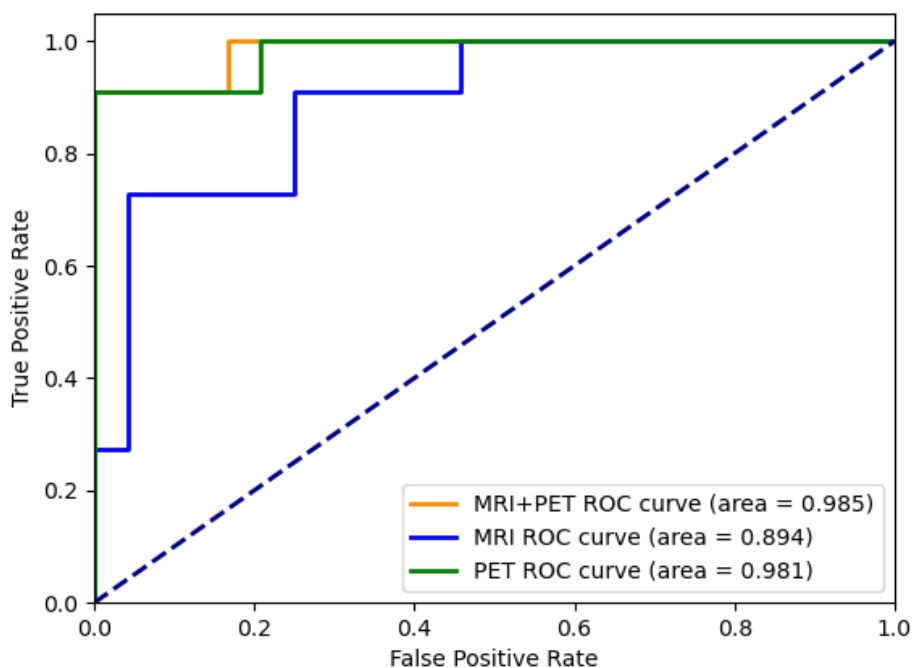


图 4-9 AD/CN 分类实验的 ROC 曲线

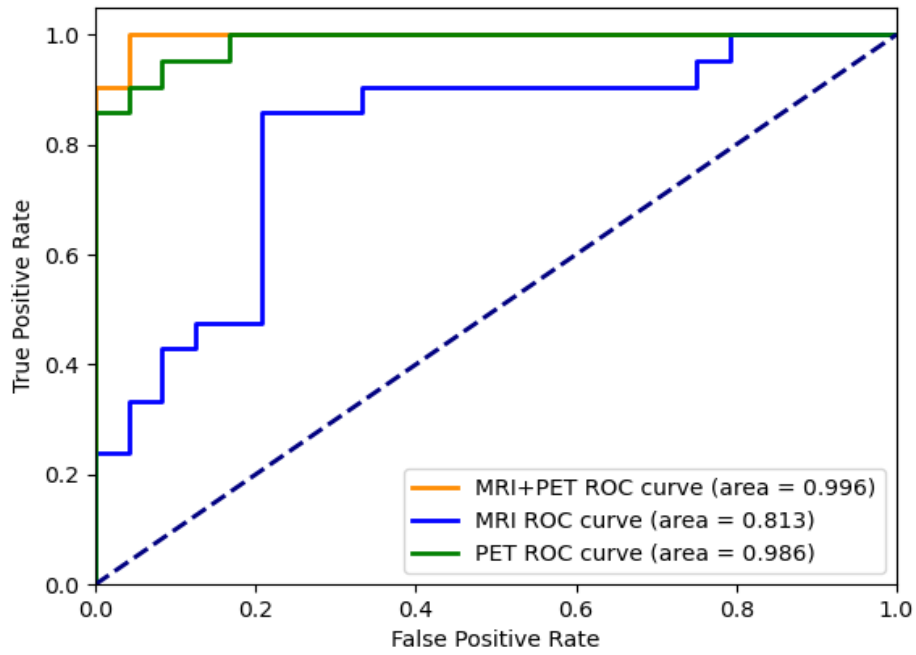


图 4-10 AD/MCI 分类实验的 ROC 曲线

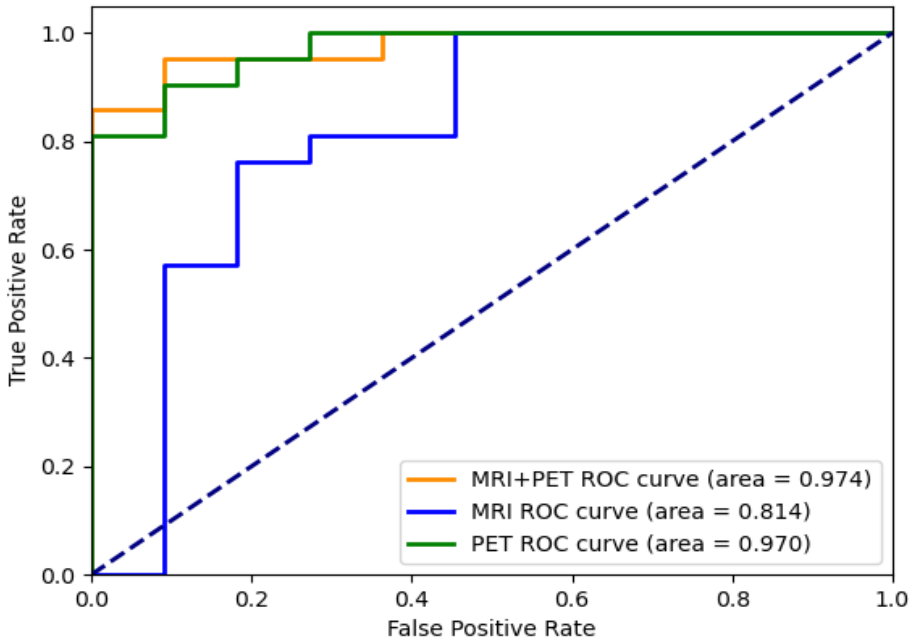


图 4-11 CN/MCI 分类实验的 ROC 曲线

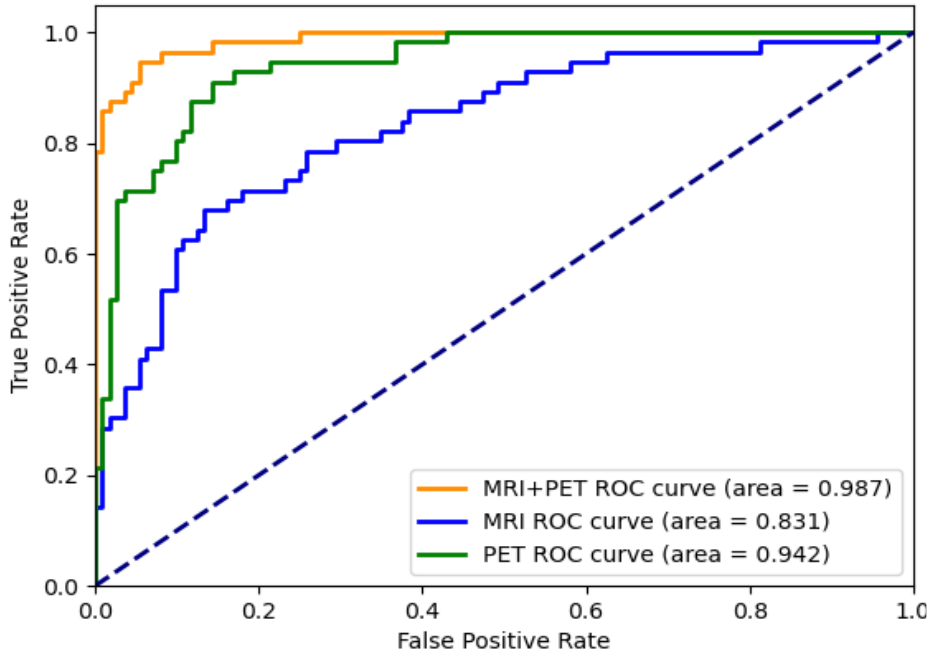


图 4-12 AD/CN/MCI 分类实验的 ROC 曲线

由图 4-9 至 4-12 的 ROC 曲线可知，在四个分类任务中，基于多模态数据融合分类的 AUC 值分别达到了 0.985、0.996、0.974 和 0.987。其中 AD/MCI 分类任务的 AUC 值最高，说明本章改进后的分类模型在区分 AD/MCI 分类任务中表现最好。而在 CN/MCI 分类任务中 AUC 值最小，这与上述混淆矩阵实验结果吻合。但是本章方法通过在两个分支网络中引入坐标注意力机制，增加模型对重要特征信息的关注能力，从而进一步增加了网络模型在 CN/MCI 分类任务中的准确率。此外，基于 PET 数据的分类方法的 AUC 值均高于基于 MRI 数据的分类方法，这是由于 PET 通过检测脑区葡萄糖代谢率或淀粉样蛋白沉积，能够捕捉 AD 早期的功能性改变。而 MRI 主要反映脑体积萎缩，但这些结构性变化通常在疾病中晚期才显著。在四个实验中，基于多模态数据融合分类方法的 AUC 值更接近于 1，明显高于其他两种基于单模态数据分类方法。这表明将两种模态数据进行融合分类能够提升模型的整体分类能力，且相比较于基于单模态数据分类具有更好的泛化能力和更高的鲁棒性，充分证明了本章方法的有效性。

为了检验本章改进网络在 AD 分类任务中的优越性，使用 Densenet、Resnext^[66]和 Shufflenet^[67]三个模型作为融合分类网络的分支网络在 AD/CN/MCI 数据集上进行对比实验，实验结果如表 4-5 所示。

表 4-5 性能对比结果

分类模型	Acc(%)	Sen(%)	Spe(%)
Resnet18	83.93	100.00	87.50
Resnext50	71.43	91.67	90.63
Densenet121	82.14	79.17	93.75
Shufflenet	87.50	100.00	84.38
本章方法	97.14	100.00	93.75

根据表 4-5 的实验结果可以看出，相较于两个经典的残差网络，本章方法在分类准确率上显著提升。具体来说，与 Resnet18 网络相比，本章方法的分类准确率提高了 13.21 个百分点，而与 Resnext50 相比，准确率提高了 25.71 个百分点。在敏感性指标上，本章方法与 Resnet18 网络相当，但在特异性上提高了 6.25 个百分点。与 Densenet121 网络相比，本章方法在分类准确率和敏感性两个指标上分别提高了 15 个百分点和 20.83 个百分点。此外，与同样采用通道混洗技术的 Shufflenet 网络相比，本章方法在分类准确率和特异性上分别提高了 9.64 个百分点和 9.37 个百分点。实验结果表明，本章方法与经典的分类网络相比在特征提取和分类性能上具有显著优势，能够更有效地整合多模态数据的特征信息，充分利用不同模态数据的互补性。因此，在 AD 的多种病程分类任务中，本章方法表现更加出色，相较于传统的单模态数据分类方法具有明显的优势，进一步验证了其在实际应用中的潜力和有效性。

为了证明本章方法在 AD 分类领域具有优异的分类性能，将本章方法与其他基于多模态数据融合的 AD 分类方法进行比较，比较的结果如表 4-6 所示。

表 4-6 分类准确率对比结果

方法来源	数据模态	分类准确率(%)			
		AD/CN	AD/MCI	CN/MCI	AD/CN/MCI
文献[28]	MRI+PET	94.11	84.83	88.48	74.54
文献[33]	MRI+PET	93.40	93.10	92.50	89.90
文献[34]	MRI+PET	95.90	94.10	95.50	-
文献[68]	MRI+PET	93.21	85.63	86.52	87.67
文献[69]	MRI+PET+CSF	88.02	84.14	-	-
文献[70]	MRI+PET+CSF+Genetic	91.80	-	79.50	60.20
本章方法	MRI+PET	97.14	95.55	90.63	91.07

由表 4-6 可以看出, 本章方法在 AD/CN、AD/MCI 和 AD/CN/MCI 分类任务中均取得了最高的分类准确率, 充分显示了所提出方法在 AD 分类任务中的优势。具体而言, 在 AD/CN 分类任务中, 本章方法的准确率为 97.14%, 相较于文献[34]的方法提升了 1.24 个百分点, 在 AD/MCI 分类任务中, 本章方法的准确率为 95.55%, 与文献[34]的方法相比提高了 1.45 个百分点。在 AD/CN/MCI 分类任务中, 本章方法的准确率达到 91.07%, 较文献[33]的方法提高了 1.17 个百分点。以上实验结果表明, 本章方法在二分类和三分类任务中的分类性能均优于当前主流方法, 验证了本章模型的有效性和优越性。

然而, 在 CN/MCI 分类任务中, 本章方法的分类准确率为 90.63%, 略低于文献[34]的 95.50%。这一结果可能与 CN 和 MCI 数据在特征分布上的高度重叠有关, 导致模型在区分这两个类别时的难度较大。相比之下, 文献[34]的方法可能更适合于 CN/MCI 分类任务中对重叠特征的提取。此外, 与一些融合多模态数据的方法(如文献[69]和文献[70])相比, 本章方法仅融合了 MRI 和 PET 两种影像数据, 而未引入如 CSF(脑脊液)或遗传信息等其他数据模态。这在一定程度上限制了模型对多模态特征的全面学习, 可能是 CN/MCI 分类任务中表现欠佳的原因之一。尽管如此, 本章方法在整体分类性能上的表现仍然显著优于其他方法, 特别是在 AD/CN 和 AD/MCI 任务中表现出明显的优势。这表明, 本章模型在提取关键特征信息、增强分类能力方面具有更好的效果。同时, 本章方法的网络结构相较于文献[34]等方法具有更高的灵活性, 为模型的进一步优化和简化提供了可能性。

4.3 本章小结

本章提出了一种基于多模态影像数据特征融合的 AD 识别方法, 使用改进的残差网络作为融合分类模型的特征提取分支网络, 分别提取 MRI 数据特征和 PET 数据特征。本章详细介绍了多模态图像融合分类网络的架构, 并阐述了网络改进模块的原理, 最后通过实验验证分析改进后的融合分类网络在各种 AD 分类任务中的性能表现。实验结果证明, 基于多模态融合分类方法的 AD 识别准确率显著优于基于单模态数据的分类方法, 改进后残差网络在多种数据集上分类的性能也均优于原始的残差网络结构。本章所提出的多模态融合分类方法在 AD 分类任务上取得了优异结果, 且具有一定的实际应用价值。

第5章 基于深度学习的阿尔兹海默症智能辅助诊断系统

随着计算机技术的不断发展,智能辅助诊断系统逐渐成为医务工作者常用的医疗诊断工具^[70]。传统的 AD 诊断容易受到医生的个人经验、专业知识和主观判断的影响,尤其是早期的 AD 病变特征不明显,极易造成误诊和漏诊。因此,为了提高 AD 诊断的准确率,去除医生主观判断对诊断的影响。本章在第二章的 2D 识别算法、第三章的 3D 识别算法和第四章的多模态识别算法的基础上设计一个能够根据患者不同模态的影像自动诊断 AD 的辅助系统。该系统能够为医生提供一个良好的交互页面和快捷的诊断流程。

5.1 系统需求分析

阿尔兹海默症是一种常见的神经退行性疾病,患者的主要症状为记忆力减退、认知功能障碍和行动能力减弱等,这些症状随着病情的发展逐渐加重,严重影响患者的生活质量。随着医学影像技术的发展,MRI 影像和 PET 影像成为诊断 AD 的重要工具,这些影像技术能够显示大脑结构和功能的变化,为 AD 的早期诊断提供了可能。而传统的 AD 影像诊断方法主要依赖于医生的专业经验,但是早期的 AD 患者的影像特征不明显,一些经验不足的医生难以做出正确的判断,这种依赖于个人经验的诊断方式容易造成患者的漏诊和误诊,给患者带来不可挽回的损失。此外,传统的 AD 影像诊断方法还依赖于医生手工提取特征,这一过程不仅耗时耗力,而且对医生的专业技能要求极高。在医疗资源紧张背景下,这种诊断方式给医生造成了巨大的工作负担,限制了医疗服务的效率和可及性。因此,开发出一种能够自动诊断 AD 影像的辅助系统能够极大地提高 AD 诊断的效率、节省医疗资源、减轻医生的工作负担,并有助于提高诊断的准确性和一致性。本章旨在设计一款基于不同模态影像的 AD 自动诊断集成系统,该系统将利用深度学习技术,结合 MRI 和 PET 影像数据,自动分析和识别 AD 的特征。

该系统由三种自动诊断模型组成,每种模型都针对不同的影像模态进行了优化,以适应不同的诊断需求。系统能够根据使用者提供的不同模态的影像数据自动进行 AD 诊断,提供快速、准确的诊断结果。此外,系统设计了一个用户友好的交互界面,使得医生和技术人员能够轻松地上传影像数据、启动诊断

流程，并查看诊断结果。系统还设计了一个患者信息管理系统，用于存储和管理患者的医疗记录和诊断结果，确保信息的安全性和隐私性。通过这些功能，系统不仅提高了医生的工作效率，还为患者提供了更加精准和个性化的医疗服务。随着技术的不断发展和完善，这款 AD 自动诊断集成系统有望成为未来 AD 诊断和治疗的重要工具。

5.2 系统功能设计

为便捷医务人员管理系统和保障患者隐私，AD 辅助诊断系统采用了 ID 登录界面，输入正确的账户和密码才能使用该系统。系统的整体框架如图 5-1 所示，进入主界面之后，用户可以根据不同模态的影像选择诊断模型，诊断结束之后可以将诊断结果储存到患者信息数据库中，在主页面也可查看已经诊断过的患者信息。

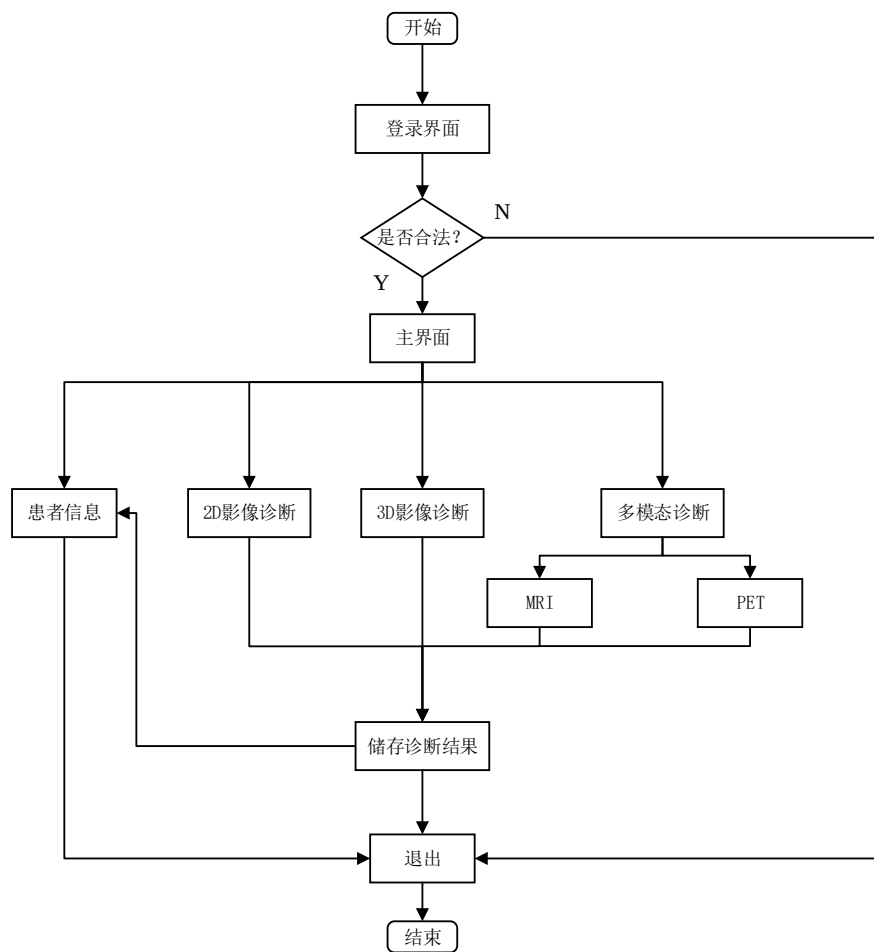


图 5-1 系统的总体框架

5.3 界面设计及功能实现

5.3.1 开发环境

本章系统设计与实现所使用的操作系统是 64 位 Windows11，处理器为 intel i5 12400F，运行内存为 16GB，显卡为 GTX 3060，显存为 12GB。开发平台为 Pycharm，使用的开发语言为 Python 3.9，界面设计使用的工具为 PyQt5 库和 Qt Designer，选用 MySQL 作为数据库系统，方便统计和管理患者的诊断信息。

5.3.2 界面实现

为了保障系统安全可靠，在进入功能主界面之前需要用户进行账号和密码验证，当用户名和密码正确时，点击“登录”按钮会进入功能主界面，否则会弹出“密码错误！”提示，点击“退出”按钮则会退出系统，登录页面如图 5-2 所示。



图 5-2 系统登录界面

系统登录界面的部分代码：

```

////////////////////////////////////
class LoginWindow(QMainWindow):
    def __init__(self):
        super().__init__()
        self.ui = loginUI.Ui_LoginWindow()
        self.ui.setupUi(self)

```

```

self.shadow.setOffset(5, 5)
self.shadow.setBlurRadius(10)
self.shadow.setColor(QtCore.Qt.gray)
self.ui.frame.setGraphicsEffect(self.shadow)
self.ui.pushButton_Login.clicked.connect(self.login_in)
self.ui.pushButton_Exit.clicked.connect(LoginWindow.close)
self.show()
def login_in(self):
    account = self.ui.lineEdit_L_account.text()
    password = self.ui.lineEdit_2_L_password.text()
    if account == "Administrator" and password == "123456":
        self.win = main_homeUI()
        self.close()
    else:
        print("password wrong!")
def retranslateUi(self, LoginWindow):
    _translate = QtCore.QCoreApplication.translate
    LoginWindow.setWindowTitle(_translate("LoginWindow", "MainWindow"))
    self.label.setText(_translate("LoginWindow", "阿尔兹海默症辅助诊断系统"))
    self.lineEdit_L_account.setPlaceholderText(_translate("LoginWindow", "账号：
"))
    self.lineEdit_2_L_password.setPlaceholderText(_translate("LoginWindow", "密
码： "))
    self.pushButton_L_sure.setText(_translate("LoginWindow", "登录"))
    self.pushButton_Exit.setText(_translate("LoginWindow", "退出"))

```

////////////////////////////////////

当系统验证用户名和密码正确之后，系统会自动跳转到功能主界面，功能主界面主要包括“2D 图像诊断”、“3D 图像诊断”、“多模态诊断”和“患者信息”四个按钮，当点击不同的按钮之后，会进入相应的功能页面，功能主界面如图 5-3 所示。



图 5-3 系统功能主界面

当用户选择使用 2D 图像诊断功能时，用户可以选择相应的按钮上传患者的影像数据，在界面右侧会展示上传的图像。然后点击“开始诊断”，系统会在内部调用第二章的基于 2D 影像的 AD 识别算法进行诊断，诊断结束之后在界面的提示框会展示出诊断的疾病类型和预测的概率。最后点击“储存结果”，系统会将此次诊断的结果储存到数据库中。2D 图像诊断的界面如图 5-4 所示。



图 5-4 2D 图像诊断界面

3D 图像诊断界面与 2D 图像诊断界面相似，用户需要将待诊断的影像文件

通过上传按钮上传到诊断系统中，且只支持上传.NII 后缀格式的医学影像文件。文件上传成功之后，诊断系统会在界面上展示用户所上传的 3D 影像三个方向的中间切片图像。当用户点击“开始诊断”按钮之后，诊断系统会通过第三章的改进 ShufflenetV2 网络对此次上传的医学影像进行 3D 诊断，诊断结果会在预览图像的下方显示，如果诊断失败或者图像解析失败，会在界面下方的提示框中展示错误信息。最后可通过储存结果信息功能将诊断结果储存到数据库中，3D 诊断界面如图 5-5 所示。



图 5-5 3D 图像诊断界面

多模态数据诊断界面整体与 3D 图像诊断界面相似，但由于多模态数据诊断需要上传多种模态的数据，所以在上传图像功能上有所区别。当点击多模态诊断界面的上传图像按钮时，诊断系统支持一次上传两个图像文件，且只支持.NII 后缀的医学影像文件。文件上传成功之后，诊断系统会在功能主界面上展示 MRI 图像和 PET 图像三个方向上的中间切片图像，如果文件上传有误，诊断系统会在功能界面上展示错误信息。当用户点击“开始诊断”按钮之后，诊断系统会将用户上传的两种模态的影像数据加载到本文第四章的多模态分类网络中预测诊断结果，并将诊断结果输出到功能界面上。用户点击“储存结果”按钮，诊断系统会将此次预测的结果保存到数据库中，且会在功能界面上提示储存成功的信息，多模态诊断界面如图 5-6 所示。

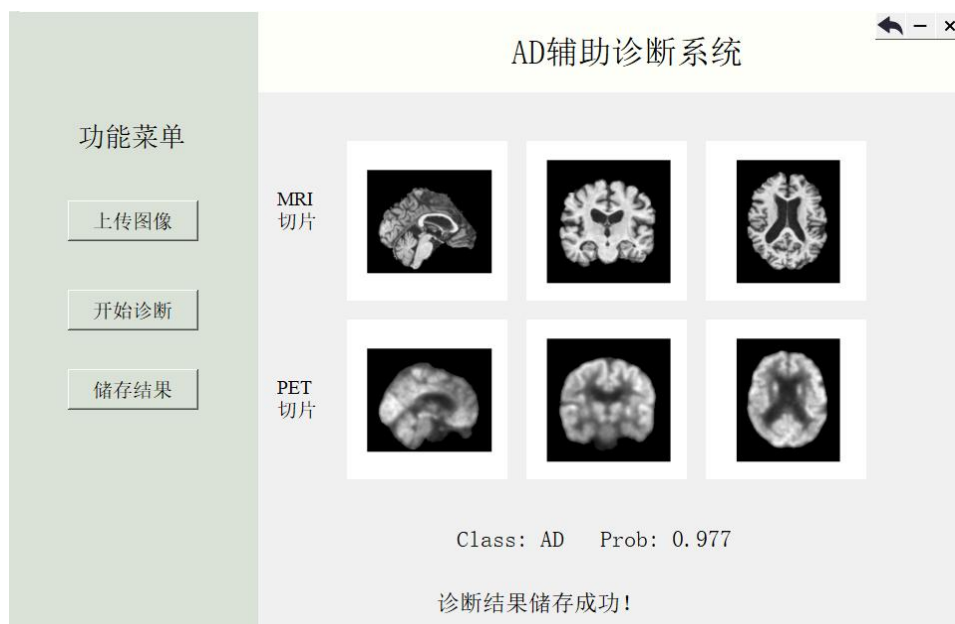


图 5-6 多模态数据诊断界面

为了方便医务人员对患者的诊断信息进行科学管理，该系统还集成了数据库管理系统。每次诊断的信息都会保存在本地数据库中，在系统功能主页面中点击“患者信息”按钮即可查看已经诊断过的患者信息，包括患者的姓名、年龄、诊断的疾病类型和概率等信息，患者信息界面如图 5-7 所示。

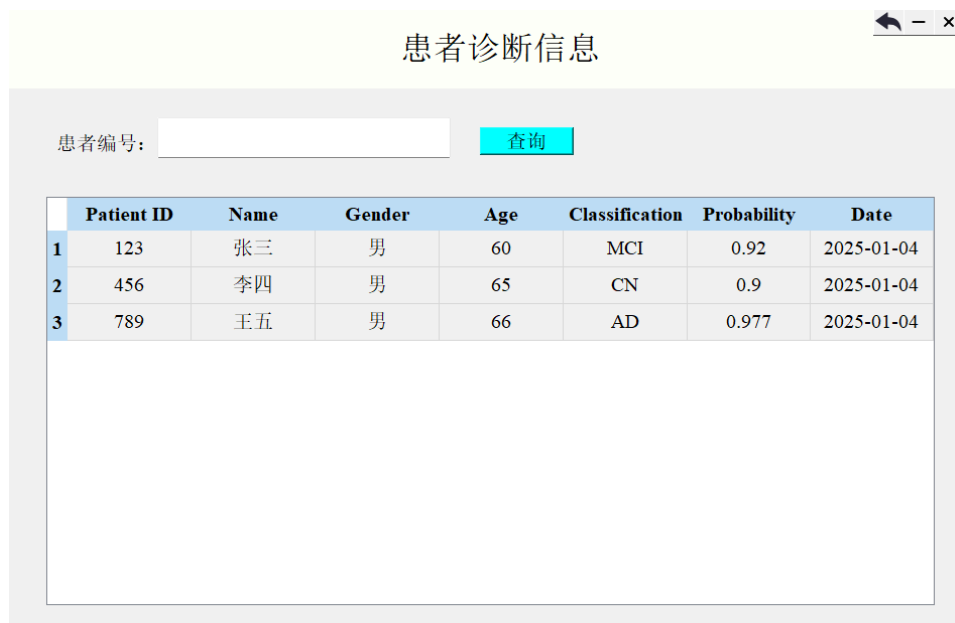


图 5-7 患者信息界面

在患者信息界面中，还支持查询功能，当医务人员想在数据库中查找已经诊断过的患者信息，即可在查询框中输入患者的 ID 编号。为保证数据的唯一性，

系统设置在诊断患者影像数据时会生成唯一的 ID 编号，查询患者信息结果的界面如图 5-8 所示。



查询结果							
	Patient ID	Name	Gender	Age	Classification	Probability	Date
1	123	张三	男	60	MCI	0.92	2025-01-04

图 5-8 患者信息查询界面

5.4 本章小结

本章在基于深度学习的 AD 自动识别算法的基础上设计了一个能够辅助诊断 AD 的集成系统。首先，在当前医疗资源匮乏的情况下，分析了 AD 辅助诊断系统的需求。然后介绍了整体的系统功能设计，简述了系统的流程框架。最后介绍了系统开发的环境信息并详细介绍了每个功能页面的实现过程。该系统集成了 2D 图像识别算法、3D 图像识别算法和多模态数据融合识别算法，能够为专业医生提供一个简单便捷的诊断界面，且具有较高的识别准确率，在临床诊断中具有一定的实际价值。

第 6 章 总结与展望

6.1 总结

阿尔茨海默症是一种常见的慢性神经退行性疾病，其在早期发病特征并不明显，因此在早期的诊断中极易误诊和漏诊。临床的 AD 诊断方法依赖于医生的医技水平和个人经验，对于一些经验不足的医生来说，诊断 AD 的效率低下。此外，在一些医疗资源匮乏的地方，早期 AD 诊断的准确率更低。因此，使用深度学习技术进行早期的 AD 诊断来解决这些问题具有十分重要的意义。本文使用深度学习技术对不同模态的 AD 影像进行识别，并结合三种不同的识别算法开发了一个能够辅助诊断 AD 的集成系统。首先，针对基于 2D 影像识别模型参数量较大、难训练、准确率较低的问题，提出一种结合注意力机制和特征融合模块的改进 Mobilevit 网络，从而在提高模型分类准确率的同时，使模型更加轻量化。其次，针对 2D 图像含有的特征信息少、易存在信息泄露等问题，提出了一种改进的 ShufflenetV2 网络对 3D MRI 影像进行识别。最后，针对基于单模态识别算法特征单一，识别准确率偏低的问题，提出了一种基于多模态数据的残差融合分类模型。本文的具体研究内容如下：

1) 提出了一种改进的 Mobilevit 模型方法用于 MRI 影像的识别。首先，通过设计多尺度分支模块来扩大感受野，以利于模型更好地捕获全局和上下文信息。其次，融合了通道注意力机制和空间注意力机制，用来学习不同通道的特征信息和提取空间中不同位置的关键信息，以提高模型特征表示的质量和判别性能。并通过实验验证该方法能够在保持高分类准确率的同时，进一步轻量化模型。

2) 探讨了一种改进的 3D-ShufflenetV2 网络用于 3D MRI 影像的识别。该网络利用 3D 卷积提取三维影像的空间连续信息，使其能够提取更多的脑部病变信息。其次，在主干网络中引入 SE 注意力机制，通过基于全局信息重新校准通道级特征响应以增加模型对重要特征的关注能力，从而增强模型的分类性能。最后，为了增加模型对全局特征提取的能力，在网络结构中增加了全局特征融合分支。实验结果表明，改进网络能够实现对 AD、MCI 和 CN 进行准确的识别，改进后网络模型的性能比原网络有较大的提升，具有一定的应用价值。

3) 研究了一种基于深度学习的多模态影像融合的 AD 识别方法。首先，提出一种新型的残差网络结构用来提取三维图像的病变特征。其次，利用改进后

的残差网络作为特征提取器分别提取 MRI 和 PET 的图像特征，然后将两种图像特征融合后进行分类。最后，为了使融合网络能更好地捕捉三维医学影像的空间位置与通道之间的关系，引入了坐标注意力机制。实验证明，改进后的融合网络在多个分类任务中分类性能较基础的残差网络均有不同程度的提升，且融合网络分类的准确率较单模态分类网络的分类准确率有较大提升，证明了改进后的融合网络的有效性。

最后，结合基于 2D 影像的 AD 识别算法、基于 3D 影像的 AD 识别算法和基于多模态数据融合的 AD 识别算法设计了一个能够辅助诊断 AD 的集成系统，该系统能够为医务人员提供便捷高效的诊断界面，且能够科学地管理患者信息。此外，该系统支持多种模态的影像数据诊断，使诊断更加高效，从而解决医生工作负担重和地域医疗资源分配不平衡的问题。

6.2 展望

本文设计了一种能够辅助诊断 AD 的集成系统，该系统结合了三种基于不同模态影像的 AD 识别算法模型。该智能诊断系统能够为医生提供一个快捷高效的交互页面，且具有较高的 AD 诊断准确率，具有一定的实际使用价值。但是仍有一些改进之处，具体如下：

1) 本文的基于 2D 影像的 AD 识别算法在多分类任务中准确率仍有较大的提升空间，且模型结构仍有进一步简化的空间。未来可针对提升多分类任务准确率和简化模型结构的方向进一步研究。

2) AD 辅助诊断系统对 AD 病程分类只有三种，但仍可以继续对 AD 的病程进行细分，如将分类扩展为不同亚型 AD 的预测，以便诊断出更早期的 AD。未来可针对细分 AD 病程进行分类研究，使辅助诊断系统更加完善，进一步提升系统的临床实用性。

3) 在多模态数据融合分类算法中融合的数据模态只有两类，虽然该算法的识别准确率较高，但未来可以融合更丰富的模态数据来进行 AD 识别，以解决当前的问题。此外，还可以针对不同的患者群体（如年轻患者与老年患者、不同的遗传背景）调整模型，提高其个性化诊断的能力。

参考文献

- [1] Fan K, Xu H, Gao L, et al. Efficient and privacy preserving access control scheme for fog-enabled IoT [J]. Future Generation Computer Systems, 2019, 99: 134-142.
- [2] Anderson N D. State of the Science on Mild Cognitive Impairment [J]. The Journals of Gerontology: Series B, 2020, 75(7): 1359-1360.
- [3] Boyle P A, Yu L, Leurgans S E, et al. Attributable risk of Alzheimer's dementia attributed to age-related neuropathologies[J]. Annals of neurology, 2019, 85(1): 114-124.
- [4] Bron E E, Smits M, Van Der Flier W M, et al. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge[J]. NeuroImage, 2015, 111: 562-579.
- [5] Chandra A , Dervenoulas G , Politis M. Magnetic resonance imaging in Alzheimer's disease and mild cognitive impairment[J]. JNeurol, 2019, 266(6) :1293 - 1302.
- [6] Jack C R, Barrio J R, Kepe V. Cerebral amyloid PET imaging in Alzheimer's disease[J]. Acta neuropathologica, 2013, 126: 643-657.
- [7] Pini L, Pievani M, Bocchetta M, et al. Brain atrophy in Alzheimer's disease and aging[J]. Ageing research reviews, 2016, 30: 25-48.
- [8] McDonald C R, McEvoy L K, Gharapetian L, et al. Regional rates of neocortical atrophy from normal aging to early Alzheimer disease[J]. Neurology, 2009, 73(6): 457-465.
- [9] Ding Y, Zhao K, Che T, et al. Quantitative radiomic features as new biomarkers for Alzheimer's disease: An amyloid PET study[J]. Cerebral Cortex, 2021, 31(8): 3950-3961.
- [10] Riederer I, Bohn K P, Preibisch C, et al. Alzheimer disease and mild cognitive impairment: integrated pulsed arterial spin-labeling MRI and 18F-FDG PET[J]. Radiology, 2018, 288(1): 198-206.
- [11] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [12] Billones, C. D, Demetria, O. J. L. D, Hostallero, D. E. D, et al. DemNet: a convolutional neural network for the detection of Alzheimer's disease and mild cognitive impairment [C]//Proceedings of the 2016 IEEE Region 10 Conference (TENCON). IEEE, Singapore, 22-25 November 2016: 3724-3727.
- [13] Jabason E, Ahmad, M. O, Swamy, M. N. S. Hybrid feature fusion using rnn and pre-trained cnn for classification of Alzheimer's disease (Poster)[C]//Proceedings of the 2019 22th International Conference on Information Fusion (FUSION). IEEE, Ottawa, ON, Canada, 2-5 July 2019: 1-4.
- [14] Jain R, Jain N, Aggarwal A, et al. Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images[J]. Cognitive Systems Research, 2019, 57: 147-159.
- [15] 石磊,彭少康,张亚萌,等. 基于特征增强金字塔网络的阿尔茨海默症早期诊断研究[J].数据采集与处理,2022, 37(4): 727-735.
- [16] 朱剑波,葛明锋,董文飞. 基于改进 EfficientNet 的阿尔兹海默症影像分类[J].计算机与现代化,2023, (06): 56-61.
- [17] 王宇春,郭浩,王波,等.基于 Faster RCNN 算法的阿尔兹海默症分类研究[J].工业控制计算机,2023,36(01):102-104.

- [18] Masud M, Almars A M, Rokaya M B, et al. A novel light-weight convolutional neural network model to predict Alzheimer's disease applying weighted loss function[J]. *Journal of Disability Research*, 2024, 3(4): 20240042.
- [19] Karasawa H, Liu C L, Ohwada H. Deep 3D convolutional neural network architectures for Alzheimer's disease diagnosis[C]//*Proceedings of the 10th Asian Conference on Intelligent Information and Database Systems (ACIIDS 2018)*. Springer, Dong Hoi City, Vietnam, 19-21 March 2018: 287-296.
- [20] Goceri E. Diagnosis of Alzheimer's disease with Sobolev gradient-based optimization and 3D convolutional neural network[J]. *International journal for numerical methods in biomedical engineering*, 2019, 35(7): e3225.
- [21] Al Shehri W. Alzheimer's disease diagnosis and classification using deep learning techniques[J]. *PeerJ Computer Science*, 2022, 8: e1177.
- [22] 孙俊楠,李岳阳,罗海驰.基于结合 SimAM 模块的三维 Densenet 的阿尔茨海默症分类算法研究[J].*计算机与数字工程*,2024,52(08):2279-2283.
- [23] Sun J, Yan S, Song C, et al. Dual-functional neural network for bilateral hippocampi segmentation and diagnosis of Alzheimer's disease[J]. *International journal of computer assisted radiology and surgery*, 2020, 15: 445-455.
- [24] 曾安,邹超,潘丹. 基于 3D 卷积神经网络-感兴趣区域的阿尔茨海默症辅助诊断模型[J].*生物医学工程研究*,2020, 39(2): 133-138.
- [25] Bailey D L, Pichler B J, Gückel B, et al. Combined PET/MRI: multi-modality multi-parametric imaging is here[J]. *Molecular Imaging & Biology*, 2015, 17(5): 1-14.
- [26] Dai Y, Bai W, Tang Z, et al. Computer-Aided Diagnosis of Alzheimer's Disease via Deep Learning Models and Radiomics Method[J]. *Applied Sciences*, 2021, 11(17): 8104-8124.
- [27] Lu D, Popuri K, Ding G W, et al. Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images[J]. *Scientific reports*, 2018, 8(1): 5697-5710.
- [28] Song J, Zheng J, Li P, et al. An effective multimodal image fusion method using MRI and PET for Alzheimer's disease diagnosis[J]. *Frontiers in digital health*, 2021, 3: 637386.
- [29] Wang Q, Xu R. AANet: attentive all-level fusion deep neural network approach for multi-modality early Alzheimer's disease diagnosis[C]//*Proceedings of the AMIA Annual Symposium Proceedings (AMIA)*. American Medical Informatics Association, Chicago, IL, USA, 5-9 November 2022: 1125-1134.
- [30] 林雪峰,李炜. 基于深度学习的阿尔兹海默症多模态辅助诊断研究[J].*工业控制计算机*,2020,33(03):58-60+109.
- [31] Dwivedi S, Goel T, Tanveer M, et al. Multimodal fusion-based deep learning network for effective diagnosis of Alzheimer's disease[J]. *IEEE MultiMedia*, 2022, 29(2): 45-55.
- [32] Odusami M, Maskeliūnas R, Damaševičius R, et al. Explainable deep-learning-based diagnosis of Alzheimer's disease using multimodal input fusion of PET and MRI Images[J]. *Journal of Medical and Biological Engineering*, 2023, 43(3): 291-302.
- [33] Zhang M, Sun L, Kong Z, et al. Pyramid-attentive GAN for multimodal brain image complementation in Alzheimer's disease classification[J]. *Biomedical Signal Processing and Control*, 2024, 89: 105652.

- [34] Rallabandi V P S, Seetharaman K. Deep learning-based classification of healthy aging controls, mild cognitive impairment and Alzheimer's disease using fusion of MRI-PET imaging[J]. Biomedical Signal Processing and Control, 2023, 80: 104312.
- [35] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Neural Information Processing Systems Foundation, Long Beach, CA, USA, 4-9 December 2017: 6000-6010.
- [36] Mehta S, Rastegari M. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer [C]//Proceedings of the International Conference on Learning Representations (ICLR). Virtual Event, 25-29 April 2022:1-26.
- [37] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional Block Attention Module [C]//Proceedings of the European Conference on Computer Vision (ECCV). Springer, Munich, Germany, 8-14 September 2018: 3-19.
- [38] Liu S, Huang D. Receptive field block net for accurate and fast object detection [C]//Proceedings of the European Conference on Computer Vision (ECCV). Springer, Munich, Germany, 8-14 September 2018: 385-400.
- [39] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at scale[C]//Proceedings of the International Conference on Learning Representations (ICLR). Virtual Event, 3-7 May 2021: 1-22.
- [40] Zhou B, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Las Vegas, NV, USA, 27-30 June 2016: 2921-2929.
- [41] Zhu X, Cheng D, Zhang Z, et al. An empirical study of spatial attention mechanisms in deep networks [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, Seoul, Korea (South), 27 October - 2 November 2019: 6688-6697.
- [42] Jack Jr C R, Bernstein M A, Fox N C, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods[J]. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine, 2008, 27(4): 685-691.
- [43] Sandler M, Howard A, Zhu M, et al. MobileNetV2: inverted residuals and linear bottlenecks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Salt Lake City, UT, USA, 18-22 June 2018: 4510-4520.
- [44] Howard A, Sandler M, Chu G, et al. Searching for MobileNetV3 [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, Seoul, Korea (South), 27 October - 2 November 2019: 1314-1324.
- [45] 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 (CVPR). IEEE, Las Vegas, NV, USA, 27-30 June 2016: 770-778.
- [46] Tan M, Le Q. EfficientNet: rethinking model scaling for convolutional neural networks [C]//Proceedings of the International Conference on Machine Learning (ICML). PMLR, Long Beach, CA, USA, 9-15 June 2019: 6105-6114.
- [47] Pan D, Zeng A, Jia L, et al. Early detection of Alzheimer's disease using magnetic resonance imaging: a novel approach combining convolutional neural networks and ensemble learning[J]. Frontiers in Neuroscience, 2020, 14: 1-12.

- [48] Yiğit A, Işık Z. Applying deep learning models to structural MRI for stage prediction of Alzheimer's disease [J]. Turkish Journal of Electrical Engineering & Computer Sciences, 2020, 28(1): 196-210.
- [49] Zhao Z, Chuah J H, Chow C O, et al. Machine learning approaches in comparative studies for Alzheimer's diagnosis using 2D MRI slices[J]. Turkish Journal of Electrical Engineering and Computer Sciences, 2024, 32(1): 93-107.
- [50] 胡喜园, 邓洪敏, 徐泽林, et al. 基于改进残差网络的阿尔兹海默症分类 [J]. 计算机应用, 2022, 42(S1): 71-79.
- [51] 林伟铭, 高钦泉, 杜民. 卷积神经网络诊断阿尔兹海默症的方法 [J]. 计算机应用, 2017, 37 (12): 3504-3508.
- [52] Ma N, Zhang X, Zheng H T, et al. ShuffleNet V2: practical guidelines for efficient cnn architecture design [C]//Proceedings of the European Conference on Computer Vision (ECCV). Springer, Munich, Germany, 8-14 September 2018: 116-131.
- [53] Hu J, Shen L, Sun, G. Squeeze-and-Excitation networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Salt Lake City, UT, USA, 18-22 June 2018: 7132-7141.
- [54] Huang G, Liu Z, Van Der Maaten, L., et al. Densely connected convolutional networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Honolulu, HI, USA, 21-26 July 2017: 4700-4708.
- [55] Howard A G, Zhu M, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Honolulu, HI, USA, 21-26 July 2017: 1704-1713.
- [56] Altaf T, Anwar S M, Gul N, et al. Multi-class Alzheimer's disease classification using image and clinical features[J]. Biomedical Signal Processing and Control, 2018, 43: 64-74.
- [57] 胡众义, 张夏彬. 利用 3D-RepVGG 进行阿尔兹海默症诊断[J]. 计算机应用, 2023, 43 (S1): 26-32.
- [58] De Mendonça L J C, Ferrari R J, Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease classification based on graph kernel SVMs constructed with 3D texture features extracted from MR images[J]. Expert Systems with Applications, 2023, 211: 118633.
- [59] Ahmed O B, Benois-Pineau J, Allard M, et al. Recognition of Alzheimer's disease and Mild Cognitive Impairment with multimodal image-derived biomarkers and Multiple Kernel Learning[J]. Neurocomputing, 2017, 220: 98-110.
- [60] Falahati F, Westman E, Simmons A. Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging[J]. Journal of Alzheimer's disease, 2014, 41(3): 685-708.
- [61] Hou Q, Zhou D, Feng J. Coordinate Attention for Efficient Mobile Network Design [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Virtual Event, 19-25 June 2021: 13713-13722.
- [62] Zhang W, Yang G, Huang H, et al. ME-Net: multi-encoder net framework for brain tumor segmentation[J]. International Journal of Imaging Systems and Technology, 2021, 31(4): 1834-1848.
- [63] 董小伟, 韩悦, 张正, 等. 基于多尺度加权特征融合网络的地铁行人目标检测算法[J]. 电子与信息学报, 2021, 43(7): 2113-2120.

- [64] 张宸嘉,朱磊,俞璐.卷积神经网络中的注意力机制综述[J].计算机工程与应用,2021,57(20):64-72.
- [65] Roy A G, Navab N, Wachinger C. Concurrent spatial and channel 'Squeeze & Excitation' in fully convolutional networks [C]//Medical Image Computing and Computer Assisted Intervention (MICCAI) 2018. Springer, Granada, Spain, 16-20 September 2018: 421-429.
- [66] Xie S, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Honolulu, HI, USA, 21-26 July 2017: 1492-1500.
- [67] Zhang X, Zhou X, Lin M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Salt Lake City, UT, USA, 18-22 June 2018: 6848-6856.
- [68] Zhu Q, Yuan N, Huang J, et al. Multi-modal AD classification via self-paced latent correlation analysis[J]. Neurocomputing, 2019, 355: 143-154.
- [69] Tong T, Gray K, Gao Q, et al. Multi-modal classification of Alzheimer's disease using nonlinear graph fusion[J]. Pattern recognition, 2017, 63: 171-181.
- [70] Mintz Y, Brodie R. Introduction to artificial intelligence in medicine[J]. Minimally Invasive Therapy & Allied Technologies, 2019, 28(2): 73-81.

攻读学位期间发表的学术论文

- [1] **陶新杰**, 魏利胜. 基于深度学习的阿尔兹海默症影像分类研究综述[J].常州工学院学报.(三类, 已录用)
- [2] **Xinjie Tao**, Lisheng Wei, Shaoyu Tang, et al. Image Recognition of Alzheimer's disease Based on Improved MobileVit Network[C]//2024 43rd Chinese Control Conference (CCC). IEEE, 2024: 8002-8007.(EI:20244117155666)
- [3] **Xinjie Tao**, Lisheng Wei, Xiaodong Ma, et al. Design on Teaching Case of Visual Magnetic Levitation Ball with Robust Control[C]//2024 4th International Conference on Computer, Control and Robotics (ICCCR). IEEE, 2024: 201-206. (EI:20243116776554)
- [4] **Xinjie Tao**, Lisheng Wei, Qin Wang. Alzheimer's Disease Imaging Recognition Based on Improved 3D-ShufflenetV2 Network[C]//International Conference on Life System Modeling and Simulation(LSMS). Singapore: Springer Nature Singapore, 2024: 276-289. (EI: 20250917965963)
- [5] **Xinjie Tao**, Lisheng Wei, Shengbo Zhu. Alzheimers Disease Recognition Based on Multimodal Image Fusion [J]. International Journal of Biomedical Engineering and Technology. (EI, 已录用)

攻读学位期间获得的奖励

- [1] 获得 2022-2023 学年安徽工程大学研究生**三等奖学金**
- [2] 获得 2023-2024 学年安徽工程大学研究生**二等奖学金**
- [3] 获得 2024-2025 学年安徽工程大学研究生**一等奖学金**
- [4] 获得 2023 年中国传感器创新创业大赛(华东赛区)**二等奖**
- [5] 获得 2023 年安徽工程大学第三届研究生数学建模竞赛**三等奖**

致 谢

行文至此，意味着我的研究生生涯即将结束。首先我要感谢我的研究生导师魏利胜教授，在他的细心指导下我顺利完成了毕业论文，在我遇到方向性难题的时候能给我耐心的指导，让我在科研过程中少走很多弯路。魏老师学识渊博、治学严谨，能够成为魏老师的学生，我倍感荣幸。此外，还要感谢我的校外导师上海欧朔智能包装科技有限公司刘小琿老师，感谢刘老师对我论文算法方面的耐心指导。

父母之爱子则为之计深远。借此机会我要特别感谢我的父母和家人，感谢父母二十多年对我无微不至的照顾与疼爱，在学业上给予我充分的理解与支持。身无饥寒，父母无愧于我；人无长进，我以何待父母。至此大鹏展翅，一举千里，我必将以人之长进报效国家，回报父母。

另外，我还要感谢我的女朋友赵濠江女士。“斯人若彩虹，遇上方知有。”一路走来，实属不易。非常感谢赵女士七年的异地陪伴，也感谢她给予了我无限的支持和关爱。更感谢她对我的包容和理解，“两情若是久长时，又岂在朝朝暮暮。初见乍惊欢，久处仍怦然，爱之所及，向阳而生，往后余生，唯愿汝与吾白头执手赏夕霞。”

最后我还要感谢课题组的各位同门，在人生中最重要的一段时间里，遇到了这个温暖的集体，我们互相学习和帮助，让我拥有一段美好的研究生时光，衷心祝愿所有同门前程似锦！万事顺遂！

陶新杰

安徽工程大学电气工程学院

二〇二五年五月

尚德敏学 唯实惟新

