

# 广州大学博士学位论文

## 临床特征驱动的智能辅助影像组学研究

Research of intelligent-assisted radiomics  
driven by clinical features

(终稿)

温金玉

导师：方美娥 教授

广州大学

2024年5月



分 类 号: \_\_\_\_\_ 密 级: \_\_\_\_\_

学校代码: 11078 保密日期: \_\_\_\_\_

学 号: 1112006005 保密期限: \_\_\_\_\_

## 广州大学博士学位论文

# 临床特征驱动的智能辅助影像组学研究

温金玉

学 科 专 业 (领 域) : 网络空间安全

研 究 方 向: 人工智能理论与应用

论 文 答 辩 日 期: 2024 年 5 月 22 日

指 导 老 师 (签 名) : \_\_\_\_\_

答 辩 委 员 会 主 席 (签 名): \_\_\_\_\_

答 辩 委 员 会 委 员 (签 名): \_\_\_\_\_



Classified Index: \_\_\_\_\_ U.D.C.: \_\_\_\_\_

School Number: 11078 Secrecy Date: \_\_\_\_\_

Student Number: 1112006005 Secrecy Period: \_\_\_\_\_

## Doctoral Dissertation of Guangzhou University

# Research of intelligent-assisted radiomics driven by clinical features

Jinyu Wen

<b>Supervisor:</b>	Prof. Meie Fang
<b>Academic Degree Applied for:</b>	Doctor of Science
<b>Specialty:</b>	Network Space Security
<b>Affiliation:</b>	School of Network Space Security
<b>Date of Defence:</b>	May 22, 2024
<b>Degree-Conferring-Institution:</b>	Guangzhou University



## 摘要

医学影像数据是疾病诊断的关键信息源，传统的影像诊断手段因人为因素易产生误诊和漏诊的问题。因此，人工智能辅助诊断的需求日益迫切。对于脑部疾病，特别是阿尔茨海默病（AD）和帕金森病（PD）等神经退行性疾病，这些疾病的诊断主要依赖于 CT、MRI、PET 等影像检查，这些检查能提供脑部结构和功能的详细信息。由于脑疾患类型多样且诊断流程复杂，因此整合多模态影像数据成为提升诊断精确度与全面性的关键环节。一方面，现行的多模态影像融合策略在一定程度上忽视了医学信息与影像特征间的内在关联性，倾向于简单套用自然图像融合技术，这种局限性在脑疾病诊断中限制了其实际效能。因此，本研究旨在从耦合医学信息与影像特征出发探究多模态融合技术新方法，增强对脑疾病诊断的准确性和医学可解释性。

另一方面，尽管现阶段基于医学影像数据的智能辅助诊断技术已取得了显著进步，但在很大程度上仍受制于对单一模态影像数据的过分依赖，以及忽略了其他潜在的诊断信息，这导致某些潜在的诊断线索可能被忽略。与此相比，基于医学影像组学的智能辅助诊断方法则展现出填补这一空白的潜力，它能够为临床医生提供更为精准的诊断依据，并有助于实现疾病的早期筛查，从而显著改善患者的预后。医学影像组学作为一门新兴的交叉研究领域，正在经历从传统的手工设计特征到特征学习的转型。深度学习方法的兴起赋予了医学影像组学全新的研究动能，使得模型能够直接从医学影像中发掘有价值的信息特征。本研究旨在通过人工智能和影像组学的综合应用，结合机器学习和深度学习技术，对脑肿瘤、AD 和 PD 的多模态医学数据进行智能辅助诊断。本研究立足临床实际需求，紧紧围绕着基于医学影像组学的智能辅助诊断算法研究，主要开展了以下工作：

(1) 针对现有的大多医学影像融合方法没有考虑医学影像的医学语义信息的问题，本文提出了一种基于医学语义信息引导的医学影像融合方法（MsgFusion）。首先在多模态脑影像的关键医学语义信息和影像特征之间建立关系，接着引导使用两个分支的特征提取和影像融合框架的设计。其中，在融合过程中使用了一种基于分层分类的策略，用于重建融合影像，以保持和增强反映解剖结构和功能代谢的突出的医学语义信息。相对现有的一些方法，本文提出的方法在 6/7 个客观评估指标上和 30 位临床医生主观评估上均占据优势。

(2) 因磁共振成像与正电子发射断层成像的融合，可以将生物解剖信息和生理代谢信息结合起来，对临床诊断和病变定位具有重要意义。因此，本文提出一种多维特征的自适应线性融合方法 (MdAFuse)。在特征提取阶段，构建三维特征独立的提取模块，可以有效地利用结构信息。在融合阶段，建立多维特征的仿射映射函数，以保持特征之间恒定的几何关系。此外，本文提出的方法中还含有关键特征可视化增强部分，旨在观察脑病变的动态生长，这可以促进脑肿瘤的早期诊断和治疗。实验结果表明，本文的方法得到了临床医生的高度评价和统计数据的支持。其中，本文提出的方法在 SSIM 和 VIF 指标上分别提升了 5.61% 和 13.76%。

(3) 针对绝大多数基于神经网络的分类方法都只能提取局部特征，感受野有限的问题，本文提出一种有效扩大感受野的小波卷积单元。另外，细粒度分类对于认知障碍的准确诊断和正确治疗具有重要意义。本文进一步设计小波卷积单元网络 (WCU-Net)。并成功应用于 AD 的细粒度多分类任务，在基于 AD 的脑 DTI 数据上首次实现了细粒度三分类。在细粒度三分类的实验中，均能获得超过 95% 的精确度。另外，在细粒度四 (AD、NC、EMCI、LMCI) 分类中，获得了 93.79% 的精确度，比现有的 SOTA 方法提高了 1.19%。

(4) 基于 PD 分类预测的研究，本文综合分析了国内外 PD 分类预测的智能辅助诊断技术，并梳理了利用机器学习模型进行辅助 PD 早期检测的研究工作，以指导早期干预和防止病情进展。首次尝试将帕金森病 (PD) 的诊断与疾病进展区分开来，不仅仅局限于简单的分类预测。另外，探讨了帕金森病预测领域的最新研究趋势，并统计了 PubMed 数据库中相关主题的文献数量。结合实际临床需求，概述了帕金森病预测的未来发展方向，这为智能辅助诊断提供了有价值的见解。

(5) 随着对评估和跟踪帕金森病进展的动态监测变得越来越重要，本文提出了一种基于跨模态数据融合的 PD 进展预测新方法 (CMFP)。该方法的独特之处在于采用纵向数据研究方法，结合机器学习技术，建立了早期 PD 进展预测模型。CMFP 方法对 PD 进展的预测达到 77.91% 的 AUC。与仅使用临床数据、DTI 数据和 DAT 数据的预测相比，AUC 分别提升了 24.48%、30.78% 和 32.7%。研究结果显示，跨模态数据融合显著提升了单模态预测的准确度，并且也说明了 DTI 数据与临床结合更有助于提高预测 PD 进展的性能，可辅助临床诊断。

本研究充分利用影像组学（多模态和异构数据），以机器学习或深度学习为基础，充分发挥人工智能在医疗领域的潜力，以辅助医生的临床决策。通过为医疗团队提供更加精准详实的数据分析及丰富的临床洞见，旨在为患者提供更个性化和有效的治疗方案，从而提升患者的生存质量、延长生存期。

**关键词：**多模态；医学影像分析；影像组学；深度学习；智能辅助诊断



## Abstract

Medical imaging data serves as a critical source of information for disease diagnosis, and traditional imaging diagnostic methods are prone to misdiagnosis and missed diagnosis due to human factors. As a result, the demand for artificial intelligence-assisted diagnosis is increasingly pressing. For brain diseases, particularly neurodegenerative disorders such as Alzheimer’s Disease (AD) and Parkinson’s Disease (PD), the diagnosis of these conditions heavily relies on imaging examinations like CT, MRI and PET, which provide detailed structural and functional information about the brain. Given the diversity of brain disorders and the complexity of their diagnostic processes, integrating multi-modal imaging data becomes a pivotal step towards enhancing both diagnostic precision and comprehensiveness. On one hand, current strategies for fusing multi-modal imaging often overlook the intrinsic relationship between medical information and imaging characteristics, opting instead for straightforward applications of natural image fusion techniques. This limitation constrains their practical effectiveness in diagnosing brain diseases. Therefore, this research aims to investigate novel multi-modal fusion methodologies by exploring the coupling between medical information and image characteristics, with the goal of enhancing diagnostic accuracy and interpretability for brain diseases.

On the other hand, despite significant advancements made in the field of intelligent auxiliary diagnostic technology based on medical imaging data at present, these systems remain largely constrained by an over-reliance on single-modal imaging data and, consequently, may overlook other potential diagnostic information, leading to the possibility of disregarding certain subtle diagnostic clues. By contrast, intelligent diagnostic approaches rooted in medical radiomics show particular promise in bridging this gap. These methodologies offer clinicians more accurate diagnostic foundations and play a crucial role in enabling early disease screening, thereby significantly improving patients’ prognoses. Medical radiomics, as an emerging interdisciplinary field, is undergoing a transformation from the conventional manual design of features to feature learning. The rise of

deep learning methodologies has infused medical radiomics with fresh research momentum, enabling models to directly extract valuable informative features from medical images. This research aims at integrating artificial intelligence (AI) with medical radiomics, coupling machine learning and deep learning techniques, to conduct intelligent auxiliary diagnosis on multi-modal medical data pertaining to brain tumors, AD, and PD. Grounded in actual clinical needs, this research focuses centrally on the investigation of intelligent auxiliary diagnostic algorithms based on medical radiomics, and it primarily encompasses the following tasks:

(1) In response to the prevalent issue where most existing medical image fusion methods do not adequately consider the medical semantic information within medical images, this thesis proposes a medical image fusion method guided by medical semantic information, termed MsgFusion. Initially, it establishes a correlation between essential medical semantic information and corresponding image features across multi-modal brain images. Subsequently, it guides the design of a dual-branch framework for feature extraction and image fusion. Within this framework, a hierarchical classification strategy is employed during the fusion process to reconstruct the fused image, ensuring that the prominent medical semantic information reflecting anatomical structures and functional metabolism is maintained and enhanced. Compared to several existing methods, the proposed MsgFusion method demonstrates superiority across six out of seven objective evaluation metrics and also garners favorable evaluations from thirty clinical experts in subjective assessments.

(2) The integration of Magnetic Resonance Imaging with Positron Emission Tomography enables the combination of biological anatomical information and physiological metabolic information, which holds considerable significance for clinical diagnosis and lesion localization. Consequently, this thesis presents a novel adaptive linear fusion method for multidimensional features, referred to as MdAFuse. During the feature extraction phase, a three-dimensional independent module is constructed to effectively exploit structural information. In the fusion stage, an affine mapping function is established for the multidimensional features to preserve constant geometric relationships among them. Moreover, the proposed methodology incorporates a crucial part dedicated to visual en-

hancement of key features, aimed at observing the dynamic growth of brain lesions, which can facilitate early diagnosis and treatment of brain tumors. Experimental results demonstrate that the method introduced in this thesis has been highly commended by clinical practitioners and substantiated by statistical data. Specifically, the proposed method achieved improvements of 5.61% and 13.76% respectively in the Structural Similarity Index (SSIM) and Visual Information Fidelity (VIF) metrics.

(3) Addressing the prevailing issue that the majority of neural network-based classification methods are limited in their ability to extract global features and possess finite receptive fields, this thesis introduces a novel Wavelet Convolution Unit (WCU) designed to effectively expand the receptive field. Furthermore, fine-grained classification plays a pivotal role in accurate diagnosis and appropriate treatment of cognitive impairments. this thesis further devises a Wavelet Convolutional Unit Network (WCU-Net), successfully applying it to fine-grained multi-classification tasks in the context of Alzheimer's Disease (AD). For the first time, it achieves fine-grained triple classification on AD-related brain Diffusion Tensor Imaging (DTI) data. In experiments involving fine-grained triple classification, the model consistently attains precision rates exceeding 95%. Moreover, in the case of fine-grained quadruple (AD, NC, EMCI, LMCI) classification, the model achieves a precision rate of 93.79%, representing an improvement of 1.19% over the current state-of-the-art (SOTA) methods.

(4) Based on studies concerning PD classification prediction, this thesis conducts a comprehensive analysis of both domestic and international intelligent auxiliary diagnostic technologies for PD classification and progression forecasting. It further reviews research endeavors that leverage machine learning models for assisting in the early detection of PD, aiming to guide timely interventions and prevent disease progression. For the first time, an attempt is made to distinguish between PD diagnosis and its disease progression, transcending mere binary or categorical predictions. Additionally, this thesis explores the latest research trends in the realm of PD prediction and statistically documents the number of related publications in the PubMed database. Grounded in practical clinical demands, this work delineates prospective avenues for PD prognosis, imparting

invaluable guidance for the advancement of intelligent diagnostic aids.

(5) With the increasing importance of dynamic monitoring for assessing and tracking the progression of PD, this thesis presents a novel method for predicting PD progression based on cross-modal data fusion, coined CMFP (Cross-Modal Fusion for Parkinson's Progression Prediction). The uniqueness of the CMFP approach lies in its longitudinal data research methodology, combining machine learning techniques to construct an early PD progression prediction model. The method achieves an AUC (Area Under Curve) of 77.91% in predicting PD progression. Comparatively, when using CMFP, there is a respective increase of 24.48%, 30.78%, and 32.7% in the AUC compared to predictions relying solely on clinical data, DTI data, and DAT (Dopamine Transporter) data alone. This research findings indicate that cross-modal data fusion significantly enhances the accuracy of predictions made using single modalities and further demonstrates that the integration of DTI data with clinical data particularly contributes to improving the performance in predicting PD progression. This supports its potential use as an auxiliary tool in clinical diagnosis.

This research fully harnesses the power of radiomics (multi-modal and heterogeneous data), building upon Machine Learning or Deep Learning foundations, to maximize the potential of Artificial Intelligence in the healthcare domain, thus supporting clinical decision-making for physicians. By furnishing medical teams with more precise and comprehensive data analyses, as well as rich clinical insights, the aim is to provide patients with more personalized and efficacious treatment plans, ultimately enhancing patient quality of life and prolonging survival periods.

**Keywords:** Multi-modal; Medical image analysis; Radiomics; Deep learning; Intelligent assisted diagnosis

# 目录

<b>摘要</b> .....	I
<b>Abstract</b> .....	V
<b>第 1 章 绪论</b> .....	1
1.1 研究背景及意义 .....	1
1.2 国内外研究现状 .....	3
1.3 本文的主要工作 .....	16
<b>第 2 章 理论与技术基础</b> .....	20
2.1 医学影像的成像技术原理 .....	20
2.2 智能影像组学技术的理论基础 .....	26
2.3 智能影像组学技术的性能评估 .....	35
2.4 本章小结 .....	43
<b>第 3 章 多模态脑影像的智能辅助融合</b> .....	45
3.1 基于脑影像的临床语义引导融合 .....	45
3.2 基于脑影像的多维特征自适应融合 .....	67
3.3 本章小结 .....	87
<b>第 4 章 阿尔茨海默症的智能辅助分类</b> .....	89
4.1 引言 .....	89
4.2 相关工作 .....	91
4.3 细粒度多分类的辅助诊断网络 .....	94
4.4 细粒度多分类的实验 .....	102
4.5 本章小结 .....	107
<b>第 5 章 帕金森疾病的智能辅助预测</b> .....	108
5.1 引言 .....	108
5.2 基于机器学习的智能辅助预测技术 .....	109
5.3 临床信息主导的智能辅助进展预测 .....	113
5.4 本章小结 .....	127
<b>第 6 章 总结与展望</b> .....	129

参考文献 .....	131
简称列表 .....	165
攻读博士学位期间发表的文章目录 .....	169
攻读博士学位期间参与的项目目录 .....	171
致谢 .....	172

# 第 1 章 绪论

自从影像组学这一概念被首次提出以来 [1]，基于医学影像组学的智能辅助诊断技术便持续成为计算机应用研究领域的研究热点 [1, 2]。在特定情况下，它能够为医生提供有效的决策支持，提高诊断准确性。随着数字化医学影像的广泛应用，这一技术在实际临床诊断中的需求和可控性大大增强，为深度学习在医学影像领域的落地提供了重要的突破口。这标志着智能辅助诊断技术将进一步与先进的计算机科学、机器学习相结合，为医学诊断领域带来更为准确、高效的解决方案，推动医学科技取得新的突破。

## 1.1 研究背景及意义

自 1985 年德国物理学家威廉·康拉德·伦琴 (Wilhelm Conrad Röntgen) 发现 X 射线以来 [3]，医学影像已经在医学诊断领域扮演着越来越重要的角色。在此之前，医生只能依赖解剖学来观察人体内部组织或通过确诊来了解病患身体内部的状况。然而，这种方式严重依赖医生个人经验，缺乏一致的直观和客观标准，同时存在一定的潜在风险。自 20 世纪 50 年代以来，电子科学技术的迅猛发展彻底改变了人类观察世界的方式。人们如今具备了制造精密传感器来表征各种物理量的能力。医学影像技术的崛起使得医生能够非侵入性地获得关于患者内部结构和病变的详细信息。随着影像技术的不断创新和进步，高清晰度影像设备（如 X 射线 (X Radiography, X-Ray) [4]、计算机断层扫描 (Computed Tomography, CT) [5]、磁共振成像 (Magnetic Resonance Imaging, MRI) [6]、显微影像 (Microscopic Imaging, Microscopy) [7]、单光子发射计算机断层扫描 (Single Photon Emission Computed Tomography, SPECT) [8]、超声成像 (Ultrasound, US) [9] 及正电子发射计算机断层扫描 (Positron Emission Tomography, PET) [10] 等）的不断进步为医学诊断开辟了崭新的篇章。

然而，尽管这些现代医学影像设备为诊断提供了强大的支持，单一模态的影像仍存在一定的局限性，难以满足准确诊断的需求。以脑肿瘤为例，CT、MRI 及 PET/SPECT 扫描成像差异明显，CT 可以观察到脑肿瘤的硬组织结构，MRI 则提供了更详细的软组织信息，而 PET/SPECT 则呈现了代谢活动。同样，对于阿尔茨海默病 (Alzheimer's disease, AD) 患者，MRI 可以观察到结构性改

变，而 PET/SPECT 则突显了代谢变化。对于帕金森疾病 (Parkinson's disease, PD)，DAT、DTI 与 PET 成像也呈现不同的信息，DAT 观察到神经元的损伤，DTI 显示了白质纤维损害，而 PET/SPECT 揭示了脑部代谢活动。因此，为了得出更为全面的结论，必须结合多种影像数据或者融入临床数据（脑脊液检查、神经性评估、生物标记物检测等）进行综合分析。

另外，放射科医生在面对庞大的病例图片时，需要处理复杂的影像信息，并进行主观判断。这一判断历程容易受到医生的经验、知识储备以及疲劳程度的影响，因而存在着漏诊和误诊的潜在风险。特别是在脑部疾病的诊断过程中，如阿尔茨海默病的轻度认知障碍容易被误诊为阿尔茨海默病，帕金森疾病的早期阶段也容易被误诊为其他神经系统疾病。这些微小而关键的变化往往在诊断过程中被忽略，进一步增加了临床决策的复杂性。随着计算机技术的不断提升和医学影像设备的普及，智能辅助诊断技术的发展为解决这些问题提供了崭新的可能性。

智能辅助诊断技术利用先进的计算方法和机器学习算法，结合影像学、医学影像处理技术以及其他手段，这项技术通过分析大量数据，辅助医生在诊断过程中做出更为精确的判断，从而提高患者治疗的效果 [11–14]。该领域的开端可追溯到 1950 年代，那时美国学者 Ledley 首次将数学模型应用于临床医学领域，运用了布尔代数和 Bayes 定理建立计算机诊断数学模型，首次成功应用于一组肺癌病例的诊断，为智能辅助诊断技术奠定了基础。在上世纪 70-80 年代，智能辅助诊断系统在医学领域取得了进一步的发展，广泛用于具有明显特征的病种，如先天性心脏病、胃癌、乳腺癌等。1983 年柏林举办的国际医学影像智能辅助诊断会议标志着医学影像智能辅助诊断学科的正式诞生。90 年代数字化 X 摄影设备的普及使得数字化影像能够直接进入智能辅助诊断系统，正式进入了临床辅助医生实现智能诊断的阶段。

在 21 世纪的来临之际，智能辅助诊断的研究进入了全新的时代，其内容变得更加广泛与深入。随着全球医疗水平的提升，一些疾病的诊断与治疗方式得到了显著改善。与此同时，人口老龄化的趋势使得老年病逐渐成为一个备受关注的焦点，例如 AD 和 PD 等。这些疾病不仅对患者的生活和家庭构成了巨大的困扰，还会对社会经济体系产生巨大的经济压力。因此，尽早确诊对于进行有效干预、延缓病程至关重要。在这一背景下，寻找可靠而有效的疾病诊断方式显得尤为重要。本文将深入探讨了脑肿瘤、AD 以及 PD 在影像组学的辅

助下进行诊断的手段，并结合人工智能方法对疾病特征进行融合、分类与预测，提出了多种创新的辅助诊断方法。

## 1.2 国内外研究现状

随着科技的飞速发展，基于医学影像组学的智能辅助诊断技术正引领着医学领域的革新浪潮。在这一领域的研究，深入挖掘了传统方法与深度学习方法两大方向，呈现出多层次、多维度的探索。传统医学影像组学诊断依赖于人工设计特征和构建分类器，虽然为医生提供了一定的支持，但其在面对复杂疾病或多样化影像时表现出的局限性逐渐凸显。然而，近年来深度学习的兴起，特别是卷积神经网络的引入，赋予了模型自主学习特征的能力，显著提升了诊断性能。智能辅助诊断方面，深度学习模型的广泛应用为研究者提供了前所未有的研究机会和成果。

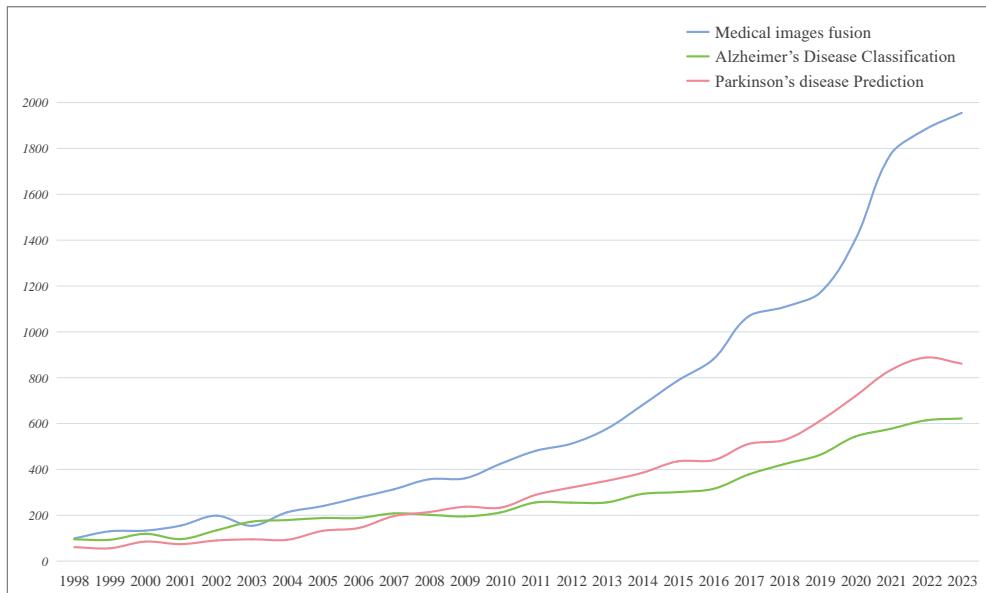


图 1.1: 按年搜索 PubMed 上三个关键词的发文量统计

基于智能辅助诊断的研究已经拓展至医学影像分析的多个领域，包括医学影像配准 [15, 16]、医学影像分割 [17–19]、医学影像融合 [20–22]、医学影像分类 [23–25]、疾病检测 [26, 27] 以及疾病预测 [28–30] 等多个层面。将深度学习方法融入智能辅助诊断系统，不仅将为医生提供更为准确、快速的诊断工具，也将为患者制定更加个性化和精确的治疗计划，推动医学影像技术不断走向新的高

峰。本文将深入探讨多模态的医学影像融合、AD 的影像分类及 PD 的疾病预测。图 1.1 统计了在 PubMed 上按年搜索 “Medical images fusion”、“Alzheimer’s Disease Classification” 以及 “Parkinson’s disease Prediction” 这几个关键词的文章数量，从中可以发现，在过去的 25 年里，基于智能辅助诊断的研究越来越受到关注。其中，多模态医学影像融合的研究正在以惊人的速度增长，其潜在的应用价值和社会影响不容小觑。

### 1.2.1 多模态医学影像融合

在临床影像诊断中，医学影像融合技术能够将来自不同模态的医学影像（如 CT、MRI、PET 等）有效地融合，提供更全面、立体的信息，有助于医生更准确地诊断病变或异常情况。此外，医学影像融合还可以用于手术规划和导航，在手术过程中为医疗专业人员提供即时且高清的影像资料，旨在提升外科手术的精确性及保障患者的安全。在当前的研究中，医学影像融合算法主要分为基于传统方法和基于深度学习方法两大类别。

#### 1.2.1.1 基于传统理论的融合方法

传统的医学影像融合方法涵盖了多种算法，其中包括基于稀疏表示 [31–33]、基于形态学 [34–36]、基于模糊理论 [37–39]、基于多尺度分解 [40–43] 及基于机器学习的融合算法 [44–48]。这些方法各自在不同场景下都表现出一定的优势，基于多尺度分解的融合方法由于其能够在不同尺度层面上进行信息处理，因而引起了广泛的研究兴趣。多尺度分解技术将影像转化为具有分层结构的多个尺度表示。这种分层结构的设计能在评估时兼顾影像的整体特性和具体细节，进而实现对影像深层意义的更全面理解。通过将影像分解成不同分辨率的子影像，多尺度分解允许独立处理不同层次的影像信息，有效地促进了对影像中关键特征的提取。这样的方法不仅有助于全局与局部信息的融合，还为影像处理提供了一种更为灵活、全面的手段，以适应不同尺度下的影像结构。基于多尺度分解的融合方法具备明显优势，能有效保留影像细节、融合全局与局部信息，同时增强特征表示，为影像处理任务提供强大支持。

拉普拉斯金字塔是典型的多尺度方法，特征金字塔的核心理念在于通过构建多尺度的影像表示，有助于系统更全面地理解影像中的对象特征，从而提升物体检测和识别的性能。Wang 等人 [49] 将拉普拉斯金字塔变换的原理作为影像融合策略，该方法主要包括三个步骤。首先，分别对每个源影像进行拉普拉

斯金字塔分解，然后采用不同的融合规则对新的每个拉普拉斯金字塔级别进行融合。对于最顶层，采用最大区域信息规则；对于其余级别，采用最大区域能量规则。最后，通过反向拉普拉斯金字塔变换得到融合影像。使用两组影像验证了提出的融合方法，并与其他融合方法进行比较。通过分析实验结果的分析显示，这种方法表现出了较好的性能，融合影像的质量优于其他方法的结果。在医学影像融合领域，基于多尺度分解的方法通过将医学影像分解为不同尺度的子影像，分别进行处理并最终融合，以确保在整个处理过程中更好地保留影像的结构和细节信息。这种策略使得医生在进行临床诊断时能够获得更全面的信息，从而提高了对患者病情的准确评估。多尺度分解的融合方法在医学影像处理中的引入，强调了在融合过程中对影像多层次特征的关注，这种思想的实际应用在 [50–52] 等文章中大放异彩。

传统的融合方法中，医学影像融合通常依赖于人工设计的特征和规则，通过影像处理技术实现对不同模态医学影像的融合。这种方法在早期为医学影像的综合分析提供了基础，但受限于人工设计的特征和规则，其应用范围和效果逐渐受到限制。

### 1.2.1.2 基于深度学习的融合方法

近年来，随着深度学习技术的迅速发展，基于深度学习方法的医学影像融合技术 [53–58] 取得了显著的突破。深度学习模型，特别是卷积神经网络 (CNN) [59, 60, 56, 61, 62] 以及生成对抗网络 (GAN) [63–66] 等，能够自动学习影像中的复杂特征和模式，从而实现对医学影像的更精细、准确的融合。这使得医学影像融合技术在影像诊断、手术导航等领域发挥着越来越重要的作用。

Lahoud 等人 [67] 利用预训练神经网络，通过提取深度特征图进行影像融合，实现将来自多模态源影像的信息整合到单一输出中。与传统方法不同，该方法采用了新颖的策略，通过比较特征图生成融合权重，推动多模态影像的融合过程。这一方法不仅适用于两幅影像的融合，还可以扩展到任意数量的输入来源。实验证明，该技术在视觉质量、客观评估和运行效率方面均有较好的效果。Kaur 等人 [55] 提出了一种采用多目标优化策略结合深度学习的医学成像处理方法，实现了不同模态的医学影像之间的有效整合。该方法首先使用非子采样轮廓波变换 (NSCT) 将影像分解为子带。随后，采用 Inception 的极端版本 (Xception) 对源影像进行特征提取。使用多目标差分演化算法进行特征优化。之后，通过决定系数和一种基于能量损耗的融合策略来计算融合权重。最

后，借助逆非子采样卷积变换构建融合后的影像。根据实验数据，这种方法在多模态影像融合技术中显示出明显的优势。

Zhou 等人 [63] 总结了生成对抗网络 (GAN) 在深度生成模型领域的研究热点，特别是在医学影像融合方面的广泛应用。首先，从基本模型和训练过程两个方面阐述了 GAN 的基本原理；其次，将变种 GAN 模型总结为三个方向（概率分布距离、整体网络架构、神经网络结构），从基于 f-散度的方法、基于 IPM 的方法、单生成器和双鉴别器 GAN、多生成器和单鉴别器 GAN、多生成器和多鉴别器 GAN、条件约束 GAN、卷积神经网络结构 GAN 和自动编码器神经网络结构 GAN 等八个维度总结了近年来的典型模型；第三，从三个方面探讨了 GAN 模型在医学影像融合领域的优势和应用；第四，介绍了 GAN 面临的主要挑战以及在医学影像融合领域所面临的挑战。Ma 等人 [65] 介绍了一种带有双重鉴别器的条件式生成对抗网络 (DDcGAN) 框架，专门用于整合具有不同清晰度的红外与可见光图像。此模型利用生成器和两个鉴别器之间的对抗性互动机制，实现了在融合影像中同时保留红外影像的热辐射和可见影像的纹理细节。为了融合不同分辨率的源影像，DDcGAN 采用降采样策略，以确保融合影像在保持信息清晰度的同时避免热辐射或纹理细节的丢失。该方法不仅在红外和可见光影像融合上表现出色，还在融合多模态医学影像方面取得了卓越的结果。

基于深度学习的医学影像融合不仅能够提高影像的清晰度和对比度，还能有效地在处理过程中维持影像的重要细节和内容，为医生提供更全面的诊断依据。此外，这些技术还有助于将不同模态影像之间的信息关联起来，为综合性医学分析提供更准确的数据支持。

### 1.2.2 阿尔茨海默症的分类

AD 作为老龄化人群中常见的慢性神经系统退行性疾病，其主要特征表现为认知障碍和记忆缺失。由于其发病机制缓慢且目前难以完全治愈，早期发现和区分中度认知功能障碍 (Mild Cognitive Impairment, MCI) 和 AD 对于实施有效的干预和治疗至关重要。值得注意的是，MCI 还分为两个亚型，即早期轻度认知功能障碍 (Early MCI, EMCI) 和晚期轻度认知功能障碍 (Late MCI, LMCI)，它们分别表示在认知功能下降中的不同阶段，EMCI 通常早于 LMCI 发生。这样的细致划分对于更准确的病程分析和制定干预方案至关重要。然而，AD 的影像诊断面临一系列挑战，因为实现精确分类依赖于能够明显区分不同

类别的特征。诊断主要依赖于对患者脑部扫描影像的分析，涉及解决一个目标分类问题。由于脑部影像的微小差异和结构变化幅度相对较小，特征不够明显，如图1.2，因此正确分类和判断阿尔茨海默症的病程变化具有较大难度。当前研究中，针对 AD 的影像分类算法主要分为两大类别：粗粒度分类（AD、NC 和 MCI 这这种类别的 2 分类或者是 3 分类）和细粒度分类（早期分类，主要是将 MCI 细分成了 EMCI 和 LMCI，再与 AD 或 NC 进行组合分类）。粗粒度分类主要关注整体疾病状态，而细粒度分类更注重对不同亚型的精准辨别。AD 分类算法的发展旨在提高对 MCI 和 AD 的早期检测，以实现及时的干预和治疗，更好地预防或延缓疾病的进展。

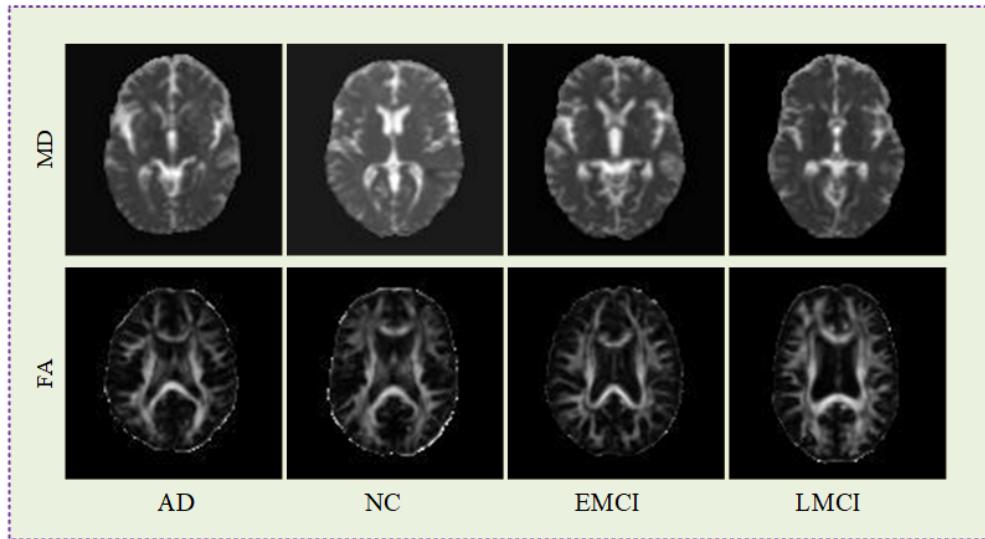


图 1.2: 四类 (AD、NC、EMCI、LMCI) 受试者的两种 DTI (弥散张量成像的分数各向异性 (FA) 和平均弥散率 (MD)) 脑部影像

### 1.2.2.1 基于粗粒度影像的分类应用

本文探讨的问题是如何在 AD、MCI 以及正常对照组 (Normal Control, NC) 之间进行有效的分类。一些学者侧重于从脑部医学影像中抽取特定的 AD 病理特征进行分类 [68–71]，例如海马、杏仁核及前额叶等关键脑区。Garali 等人 [69] 提出了一种新颖的方法，用于评估大脑区域在区分 AD 和健康脑影像方面的有效性。首先，将脑扫描数据分割并对齐至 116 个特定的解剖结构区域。然后计算这些区域的前四个时刻和直方图的熵。接着，使用接收器操作特性曲线对各个区域分离 PET 脑影像的能力进行排名。选取了 21 个区域作为输入，分别输入

到支持向量机和随机森林分类器中，评估结果基于 142 个脑 PET 影像进行。分类结果较使用最初的 116 个区域或输入整个脑体素时更好。Feng 等人 [70] 提出了一种基于 ROI 的小波子带能量 (ROICSE) 特征方法，以在频域中表示 sMRI 影像进行 AD 分类。具体而言，sMRI 扫描经过预处理步骤后，会使用预先定义的脑部掩膜进行分割，从而得到 90 个特定的兴趣区域 (ROI)。与传统方法在空间维度上直接提取这些 ROI 特征不同，这里采用的是小波变换技术，对每一个 ROI 进行分析，从而获得其能量分布的不同子带。然后，对于一个 ROI，构造一个子带能量 (SE) 特征向量来捕捉其能量分布和轮廓信息。随后，将 90 个 ROI 的 SE 特征向量串联起来形成 sMRI 影像的 ROICSE 特征。最后，使用支持向量机 (SVM) 将 880 名受试者的数据进行了分类，这些数据来源于 ADNI 和 OASIS 数据库。实验结果表明，AD vs. NC、MCI vs. NC 和 AD vs.MCI 的准确率分别是 93.57%、83.13% 和 82.73%，ROICSE 方法优于其他六种最先进的方法。Gray 等研究者 [71] 将每位参与者的脑部 MRI 扫描细致地分割为 83 个不同的解剖部分。随后，他们使用这些解剖区域作为参考框架，从同个人的 FDG-PET 扫描中提取了各 ROI 的平均放射性标记强度数据。最终，通过分析这些强度值作为特征，利用 SVM 算法对样本进行了分类处理。在 AD 患者和 NC 对照组之间取得了出色的区分度（准确率 82%），在 MCI 患者和 NC 对照组之间取得了良好的区分度（准确率 70%）。

尽管基于特定区域的兴趣点的方法来提取关键特征并对这些特征进行深入分析的策略在某些方面表现出效力，但也存在一些明显的限制。首先，特征提取过程可能会受到人为因素的影响，导致错误，同时分析手段也可能存在局限性。由于 AD 标志物不明确，因此在划分和选择 ROI 时可能发生疏漏，极大地影响 AD 诊断的效果。再者，采用基于 ROI 的技术在实现时，模型训练通常需要依赖大规模的实验数据来保证预测结果的精确度，这无疑会增加显著的时间和人力资源消耗。而在阿尔茨海默病分类研究的现状下，采用 ROI 方法的研究者们普遍采取了多样化的影像预处理手段，包括但不限于影像与参考模板的精确对位、应用刚性或非刚性变换来适配不同的参照框架，以及对影像进行详细的区域分割处理。面对这些挑战，深度学习方法在 AD 影像分析中展现出卓越的潜力 [72–76]。深度学习通过学习数据中的特征，自动发现并表征影像中的复杂模式，避免了手动提取特征的繁琐过程。这使得深度学习在处理 AD 影像分类任务时更具优势。其对大规模数据的高效处理能力，以及在学习表征方面的卓越性能，使得深度学习方法成为应对上述问题的有力工具。

Suk 等人 [72] 提出了基于深度学习的堆叠自编码器的特征表示方法。为了创建一个可靠的 AD/MCI 分类模型，作者综合考虑了深层信息和基础特征，以此提高模型的诊断性能。实验证明，所提出的方法对于 AD、MCI 的诊断分别具有 95.9%、85.0% 的准确率。通过对训练模型的可视化观察，一种利用深度学习技术来获取神经影像数据中高级共享潜在特征的新方法被提出 [74]。具体而言，该框架以受限的玻尔兹曼机（DBM）为单元，能够从三维图像片段中识别并提取层次化的潜在特征。接着，应用多模态 DBM 对 MRI 和 PET 图像的对应区块进行了联合特征提取。为了评估所提出的方法的效能，进行了一系列详尽的实验测试，并与最先进的方法进行了比较。在 AD vs. NC、MCI vs. NC 分类问题中，分别获得了 95.35%、85.67% 的最大准确率，优于竞争方法。So 等人 [73] 提出了一种结合纹理分析与深度学习的 AD 粗粒度分类方法。首先对 MRI 影像应用了三维灰度共生矩阵（GLCM）方法进行纹理分析。然后，使用 Fisher 系数选择适用于分类的特征。最后，利用多层感知器（MLP）模型将其分为 AD vs. MCI、AD vs. NC 和 MCI vs. NC 三类。该模型在分类准确性上表现出色，分别为 AD vs. MCI (72.5%)、AD vs. NC (85%) 和 MCI vs. NC (75%)，并在混淆矩阵、支持向量机（SVM）和 K 最近邻（KNN）分类器的评估中展现出优越性。所提出的方法相较于 SVM 和 KNN 分类器至少提高了 6-19% 的准确率。自动化的 AD 分类系统可以作为辅助工具，帮助医务人员诊断阿尔茨海默病的阶段，从而提供适当的医疗治疗。Khvostikov 等人 [76] 提出了一种基于 3D Inception 的卷积神经网络（CNN）的设计，用于 AD 的诊断。该网络注重内部资源的利用，通过在海马 ROI 上融合 sMRI 和 DTI 模态。与传统的基于 AlexNet 的网络相比，表现显著更好。AD vs. MCI vs. NC、AD vs. NC、AD vs. MCI 与 MCI vs. NC 这些分类的准确率分别是 68.9%、93.3%、86.7% 和 73.3%。

尽管在脑部医学影像中提取特定的 AD 病理特征的方法在精确定位疾病迹象方面取得了一定的成果，然而，与那些依赖深度学习的诊断技术相比，这些传统方法在精确度上稍显不足。这可能是因为这些传统方法主要关注单一或少数感兴趣区域（ROI），而对于复杂的神经影像数据，这种精细的特定特征提取方式难以涵盖疾病的的整体情况。特别是在处理 AD 的粗粒度分类问题时，通常涉及二分类（AD vs. NC、AD vs. MCI 和 MCI vs. NC）或三分类（AD vs. MCI vs. NC）的任务。这些任务主要旨在初步判别患者的整体状态，而未能提供关于疾病进展更详细信息的细致分析。相比之下，基于深度学习的方法通过

学习大量数据中的特征，具有更强大的表示学习能力，可以更全面地捕捉疾病的复杂模式和变化。这种方法的优势在于其对大规模和多模态数据的高效处理，以及对于特征学习的自动化处理。这种自动化处理避免了手动提取特征时可能引入的主观性和繁琐性。因此，基于深度学习的方法展现出更高的潜力，有望推动 AD 诊断领域向更为准确和全面的方向发展。这不仅提供了更有力的支持，还有助于更全面地理解和干预 AD 等神经系统疾病的发展过程。

### 1.2.2.2 基于细粒度影像的分类应用

尽管 AD 的粗粒度分类方法在筛查、早期检测和初步诊断方面具有一定的实用性，对于医务人员而言，在多阶段 AD 的诊断中进一步了解疾病的不同阶段、进展速度以及相关病理学变化，对个体化的治疗方案和干预手段更具有指导意义。目前，关于 AD 的智能辅助诊断的研究在多个 AD 进行性阶段方面的适用性相对较少。近年来，越来越多的学者对 AD 进行了细粒度的研究，例如 [77–85] 等。

Korolev 等人 [77] 采用残差网络和普通 3D 卷积神经网络结构，应用于 3D 结构性 MRI 脑扫描的 AD 与 MCI 以及 NC 的分类任务中，涉及到 AD 与 NC、EMCI、LMCI 之间的多类别分类。该研究展现了在细粒度分类上的一定成果，不仅涉及 AD vs. NC，还包括了 AD vs. EMCI、AD vs. LMCI、LMCI vs. NC 等任务，其结果表明了模型在处理这些复杂分类问题上的鲁棒性。实验测试的准确率分别是：AD vs. NC: 80%，AD vs. EMCI: 63%，AD vs. LMCI: 59%，LMCI vs. NC: 61%，LMCI vs. EMCI: 52%，EMCI vs. NC: 56%。Ruiz 等人 [78] 则提出了一种集成方法，采用 3D 密集连接卷积网络（3D DenseNets）模型进行四分类，即 AD、EMCI、LMCI 和 NC。这种方法通过密集连接改善模型内部数据的流动，同时采用概率融合方法合并每个独立分类器模型的概率输出。在测试中，该方法相对于处理 3D MRI 的其他最新方法表现更好，获得了 66.67% 的准确率，为细粒度分类提供了可行的解决方案。

Song 等人 [79] 使用图卷积神经网络（GCN）进行网络分类，将受试者分为 NC、EMCI、LMCI 和 AD 四个类别，实现了 89% 的准确率。Parmar 等人 [80] 通过使用经过修改的 3D 卷积神经网络设计了一种简化而准确的方法，从静息状态 fMRI 数据中提取特征并对 AD 进行四分类。该研究中的 CNN 设计保留了更多的空间和时间信息，获得了 93% 的准确率，为对 AD 进行多阶段诊断提供了一种潜在的有效途径。Helaly 等人 [83] 则采用两种基于深度学习的方法对

医学影像进行 AD 四个阶段 (AD vs. EMCI vs. LMCI vs. NC) 的分类。这两种方法分别基于简单的 CNN 架构和迁移学习原理，使用不同的模型进行医学影像分类。具体来说，第一种方法使用简单的 CNN 架构，该架构基于 2D 和 3D 卷积处理数据集为 2D 和 3D 的脑部扫描。第二种方法应用迁移学习原理，利用预训练的模型 (VGG19 模型) 进行医学影像分类。这两种方法分别取得了 93.61% 和 95.17% 的准确率，为多阶段 AD 分类提供了更多选择。

少数研究还对 AD 的进行性阶段 (包括 NC、主观记忆问题 (SMC)、EMCI、MCI、LMCI 和 AD) 进行了多类别分类 [85, 86]，进一步拓展了对 AD 多阶段的认知。Shamrat 等人 [85] 提出的 AlzheimerNet 分类器具有更细致的多阶段诊断能力。通过精细调整的卷积神经网络 (CNN) 架构，该模型可以识别阿尔茨海默病的五个阶段，包括 SMC、MCI、EMCI、LMCI 和 AD 以及 NC 类别。在六分类任务上，通过比较五个预训练模型 (VGG16、MobileNetV2、AlexNet、ResNet50 和 InceptionV3) 和 AlzheimerNet 模型在各种性能指标上的表现，该模型表现出 98.68% 的准确率，显示出了在多阶段 AD 诊断中的出色性能。

在 AD 的细粒度分类方法中，主要集中在二分类任务，例如 AD 与 EMCI、AD 与 LMCI、LMCI 与 NC、LMCI 与 EMCI 等。这些二分类任务在当前研究中占据主导地位，主要受到深度学习方法的广泛运用。就二分类而言，研究结果表明，这些方法在 AD 的早期诊断、疾病阶段划分以及患者与正常对照的区分等方面取得了显著的成果。例如，Korolev 等人 [77] 在 AD vs. NC、AD vs. EMCI、AD vs. LMCI 等任务中实现了可观的分类准确性，为早期 AD 的精准辨别提供了可能性。相比之下，三分类任务 (如 AD vs. EMCI vs. LMCI、NC vs. EMCI vs. LMCI) 相对较少，研究涉及更多的类别，需要更高的分类难度和挑战。在这方面，尚未形成广泛的研究共识。

然而，一些初步的研究显示，在处理多类别分类问题时，模型的性能仍然具有潜在的优势。通过更全面的分类，有望提供更丰富的信息，为医务人员提供更全面的诊断参考。六分类任务，涉及到 NC、SMC、EMCI、MCI、LMCI 和 AD 的分类，是相对较少见的研究方向。在这一领域，Shamrat 等人 [85] 提出的 AlzheimerNet 分类器在实现六分类任务上表现出色，具有 98.68% 的准确率，显示了对多阶段 AD 进行细致分类的强大潜力。这种细粒度的多分类方法有望为医疗决策提供更详细的信息，推动个体化治疗和干预的发展。绝大多数采用深度学习方法的 AD 细粒度分类方法，其准确率存在一定的差异，取决于模型

的架构设计、数据集质量以及训练过程中的参数调整等因素。在这方面，不同的研究方法可能呈现出高低不一的性能。然而，总体趋势显示出深度学习方法在解决 AD 细粒度分类问题上的潜在优势，通过学习丰富的影像特征，更全面地捕捉疾病的多样性，为提高 AD 诊断的准确性和全面性提供了有效途径。

### 1.2.3 帕金森疾病的预测

帕金森症 (PD) 是一种长期性的神经退行性疾病，它直接作用于那些在大脑中制造多巴胺的关键神经元。这种疾病在神经系统疾病中的发病率排在第二位，仅次于 AD[87, 88]。许多患有 PD 的患者在被诊断时已处于病情的中到晚期阶段，这使得他们失去了接受最佳治疗的理想时机。[89, 90]。因此，急需开发出更快捷、更精确、更具效能的检测技术，以便实现对 PD 的早期识别 [91]。可作为 PD 预测的数据主要包括医学影像数据和临床数据。医学影像数据可以呈现出脑部结构和活动的直观影像，有助于观察神经元损伤和多巴胺生成的变化。临床数据则提供了更全面的患者信息，包括症状描述、家族病史以及其他相关的临床特征，这些信息对于全面评估患者的健康状况至关重要。将医学影像数据和临床数据相结合，可从多个角度来分析和理解帕金森病的发病机制，从而构建出更精确、更可靠的早期诊断模型。这样的模型不仅能提高诊断的准确性，还能为患者提供及时的治疗建议，提升生活品质并减缓病症进展。

#### 1.2.3.1 基于医学影像的预测方法

在医学影像学领域，不断涌现新技术以助力帕金森病 (PD) 的诊断。这些技术因其独特的优势而备受瞩目，预计将大幅提升诊断的准确性。医学影像学提供了多种影像类型，如磁共振成像 (MRI) 揭示了大脑结构的细微差异；功能性磁共振成像 (fMRI) 捕捉了脑部活动模式；结构磁共振成像 (sMRI) 评估了大脑组织的形态改变；单光子发射计算机断层扫描 (SPECT) 和正电子发射断层成像 (PET) 则用于检测大脑代谢异常；弥散张量成像 (DTI) 描绘了神经纤维束的走向和完整性。这些技术共同为 PD 的早期发现与治疗提供了强有力的工具。

基于体素的形态测量学 (voxel-based morphometry, VBM) 技术利用 MRI 扫描快速且精确地量化脑部形态变化，如小脑的体积减少，并能在宏观层面迅速定位帕金森病 (PD) 的脑结构异常 [92, 93]。这项技术在成本效益和诊断准确性方面表现出色，已被临床研究所证实。Amoroso 等学者 [94] 通过分析大脑

肌肉萎缩与疾病严重度的关系，创建了一套全面的脑结构网络模型，其疾病分类准确度超过了 94%，优于标准 VBM 分析。另外，Rojas 等人开发了一种基于 SPECT 影像的经验模式分解新技术 [95]，结合 PCA 提取关键特征并运用 SVM 进行分类，实现了高达 95% 的准确率。Alexandra 等人 [96] 通过对两组独立的受试者的静息态功能磁共振影像进行分析，并运用随机逻辑回归挑选特征输入 SVM，在测试集中达到了 80% 的预测准确率。

Francisco 等研究者 [97] 开发了一种新技术，通过分析 FP-CIT SPECT 影像来预测 PD，并通过运用支持向量机（SVM）、K 最近邻（KNN）以及逻辑回归这三种不同的机器学习模型，实现了 97.9% 的高分类精度。此外，Martinez 等人 [98] 采用 FP-CIT SPECT 的医学影像数据结合 CNN 架构，包括 LeNet3D 与 Alexnet3D，对 PD 进行了分类，其中 Alexnet3D 的分类准确度为 94.1%，曲线下面积为 98.4%。Kim 等人 [99] 应用了 Inception V3 网络进行 PD 预测，获得了 87% 的曲线下面积。Castillo 等人 [100] 提出了一个集成学习的分类模型，结合了 SPECT 影像数据和多种生物学标记，构建了一个加权分类模型，该模型的分类准确度为 96%。这些研究成果显示，利用医学影像数据的 PD 预测技术正在持续改进，这些技术不仅提高了预测的准确性，还引入了多种先进技术和深度学习框架，为 PD 的早期诊断提供了更稳健和精确的工具。

近期研究通过 DTI 技术深入分析了患者的症状，进一步揭示了白质网络中神经元损伤的情况，并将其与正常对照组进行了比较。结果显示，PD 和 MCI 患者的额叶白质网络功能损伤较为集中，且认知网络的全面功能失调和多个神经元认知网络在频率域内所表现出的功能损伤之间存在直接联系 [101]。多模态 MRI 数据融合了磁共振成像的特点与早期疾病检测的优势，为 PD 的早期发现提供了有价值的影像学生物标志物 [91]。尽管 sMRI 和 DTI 单模态技术能够较好的反映血脑网络的宏观结构变化，但它们在功能形态层面上仍有不足。Rs-fMRI 主要用于评估缺血性痴呆患者的脑灰质异常，但目前尚未能够全面且系统地评估其对脑白质异常所导致的神经元活动失常以及相关的脑功能变化和特征 [91]。此外，针对大脑黑质皮层区域，现有的磁共振扫描成像技术也存在一定的局限性。尽管对该区域的异常表现具有较高的灵敏度，但在描述这些特征时仍然面临着复杂性、模糊性以及单一性的问题。

为了克服这些技术挑战，研究者们开始采用多模态检测和多重 MRI 异常检测技术，融合多种特征进行综合分析。该方法通过多角度、多层次地分析 PD 的

神经病理状况，实现了对疾病的全面诊断及治疗效果的综合评价。它为中国 PD 神经病理学的临床诊断和评估提供了一个较为完善、成熟、合理且可靠的新途径 [102–104]。Pyatigorskaya 等人 [102] 首个遵循国际准则对 NMS-MRI 信号进行详尽剖析的工作，旨在揭示黑质中的质子体积与信号强度的特定特征，并在 DTI 技术支持下，分析了局部各向异性的特性及其信号组合模式。成功将 PD 的预测准确率提升至 93%。Bowman 等人 [104] 则采用融合的策略，并分析和讨论了 rs-fMRI、sMRI 以及 DTI 中的组织结构特性。经综合考虑多种影像学指标如灰质密度、各向异性分数和功能连接等，以区分 PD 患者与正常对照组。通过这种多模态分析方法，他们达到了 98.9% 的分类准确率，这一结果显著优于仅依赖单一影像模态的传统方法。这些发现不仅增强了对 PD 神经生物学特征的理解，而且为临幊上 PD 的早期识别和治疗策略的制定提供了重要的影像学参考。

在医学影像领域，针对 PD 的早期诊断技术正在不断进步，涵盖了包括 sMRI、扩散张量成像 (DTI) 和静息态功能磁共振成像 (rs-fMRI) 在内的多种成像手段。这些技术能够揭示出具有高度特异性的影像学生物标志物，从而增强了对 PD 早期阶段的识别能力。尽管如此，每种技术也面临着自身的挑战；例如，sMRI 和 DTI 在捕捉功能性形态变化方面存在不足，而 Rs-fMRI 在评估脑白质皮层异常的神经元变化时则可能受到限制。为了克服这些限制，研究者们开始探索多模态检测和多重 MRI 异常检测技术的应用。这种方法通过综合不同 MRI 技术的优势，从多个维度和视角对 PD 进行全面的临床诊断和疗效评估。现有的研究 [102–104] 表明，这种多模态方法能够取得较高的预测准确性，为我国在 PD 神经病理学临幊诊断评估指标方面提供了新的思路和方法。

### 1.2.3.2 基于临幊数据的预测方法

医学领域的临幊资料丰富且多元，为深入研究 PD 提供了宝贵资源。临幊资料广泛涉及病患的生理和生化指标，其中蕴藏着揭示 PD 特征的关键信息，对于预测 PD 的发展趋势具有重要价值。基于这些临幊数据的 PD 预测方法，对于指导医学实践和治疗方案的制定具有重要意义。

在认知障碍领域，PD 患者常伴有帕金森病痴呆 (Parkinson's disease dementia, PDD)，及时识别其进展对患者的治疗和生活质量具有重大影响。为此，Kim 等人 [105] 采用了三种常用的认知评估测评——蒙特利尔认知评估 (MoCA)、马蒂斯痴呆分级量表 (DRS-2) 和简易的精神状态测评表 (MMSE)，并

使用回归模型评估了它们在预测 PD 认知进展方面的效能。研究发现，在条件一致情况下，MoCA 在诊断 PD 轻度认知障碍 (PD-MCI) 和 PDD 方面的准确性最高，分别达到了 AUC=0.79 和 AUC=0.89，敏感性分别为 76.4% 和 81.0%。这一成果首次证明了 MoCA 可以作为预测 PDD 发生的指标。尽管 MoCA 在检测和预测 PD-MCI 和 PDD 进展方面获得了较好的应用 [106–110]，但 MoCA 的低特异性限制了它作为诊断工具的实用性 [108, 109]，这反映了 PD 进展的复杂性以及当前缺乏可靠的客观生物标志物 [111]。因此，开发更先进的 PD 进展预测模型或更精确的诊断策略，对于选择合适的临床试验参与者显得尤为重要。

随着海量数据集的涌现，特别是那些涵盖纵向研究和多维度信息的数据集 [112–114]，为 PD 进展预测模型的构建迎来了新的发展契机。例如，Diba 等研究者 [115] 通过机器学习技术分析 PD 的血清样本的研究，揭示了外周炎症标志物与 PD 症状间的联系。研究结果表明，患者初始阶段的外周炎症水平对于预测其病情的未来演变具有重要意义。Chan 等人 [116] 提出的预测模型，证实了诸如左旋多巴之类的药物在控制 PD 发展和改善患者症状方面的显著疗效。也有学者利用生物力学测量，包括步态和姿态的平衡能力，构建了预测 PD 患者两年之内进展的网络模型 [117]，该模型采用 CNN 技术预测 MDS-UPDRS 第三部分评分的两年变化率，为针对快速进展型 PD 的药物治疗提供了新思路。另外，Latourelle 等人 [112] 的研究指出，临床基线的运动评分数据是预测 PD 未来发展趋势的关键因素之一。

在 PD 的初期阶段，及时检测非运动症状对于控制病情发展极为重要。Nilshi 及其同事 [118] 开发了一种新的远程监控 PD 进展的技术，该技术融合了深度学习与聚类分析两种方法。具体实施时，研究团队采用了深度信念网络(DBN) 结合支持向量回归 (SVR) 的策略来预测 PD 的 UPDRS 评分，并构建了多个时间尺度上的 DBN 预测模型。为了提升预测的性能，研究中还引入了自组织映射 (SOM) 的聚类技术，并通过 SVR 对 DBN 模型的预测结果进行整合优化。研究成果表明，这种结合聚类与 DBN，再通过 SVR 进行结果融合的方法，在预测 Total-UPDRS 和 Motor-UPDRS 方面均优于单一的 SOM、DBN 或 SVR 技术。具体而言，在预测 Total-UPDRS 时，测试集上的均方根误差 (RMSE) 与决定系数 (R<sup>2</sup>) 分别达到了 50.8% 与 93.3%；在预测 Motor-UPDRS 得分方面，模型的均方根误差 (RMSE) 为 52.1%，决定系数 (R<sup>2</sup>) 则达到 91.4%。这一技术开辟了远程监控 PD 患者病情进展的新途径。

在医学研究领域，对 PD 的多样化临床数据进行深入分析，是揭示疾病预测因素的关键。当前研究重点之一是对比不同的认知评估测试，其中蒙特利尔认知评估 (MoCA) 在检测 PD 认知障碍方面显示出较强的能力，尽管如此，其低特异性问题仍待克服。结合大型数据集及 ML 技术，PD 的进展预测模型得到了显著提升，特别是在生物标志物和基因与疾病关联的探索上。总体来看，PD 的研究正步入充满希望的新阶段，为更有效的疾病干预和治疗策略提供了坚实的科学支撑。

### 1.3 本文的主要工作

本项研究聚焦于智能辅助影像组学分析技术在医学辅助诊断领域中的应用，特别关注脑肿瘤、AD 和 PD。关键技术手段涵盖影像融合、影像分类与疾病预测。本文的主要工作包括：

- 在多模态医学影像融合方面，本文分析了 MRI、CT、PET、SPECT 及其他医学影像，归纳总结了其各自隐藏的医学语义信息，并与影像特性对应，设计了一个医学语义引导的双分支网络。另外，本文从多个维度如低频、高频和多尺度出发，分析不同的空间、通道特征，提出三种特征提取模块。并结合自适应线性融合策略，提出一种多模态脑影像融合的网络模型。这两种方法都得到了临床医师较高的评价。
- 在 AD 诊断方面的研究上发现了 AD 不止有三类 (AD、NC 和 MCI)，其中 MCI 还可以分期，细分为 EMCI 和 LMCI。EMCI 阶段是可逆的，及时发现和干预可以避免 AD 的发展，而在 LMCI 阶段及时诊断和治疗可以延缓 AD 的发展或治愈 AD。另外，由于绝大多数基于深度学习方法的应用都难于获得全局的特征，而频域的处理可以做到。因此，本文提出了一种小波卷积单元，并将其应用于 AD 的细粒度多分类。
- 在 PD 诊断领域，本文归纳总结了现阶段国内外学者在 PD 预测上的研究工作，将其定义为静态和动态预测两种类型。针对当前的研究现状，本文讨论和分析了 PD 智能诊断未来的研究趋势。另外，本文在 PD 基线的跨模态（除了考虑影像层面的数据外，还结合了非影像类数据）数据中，筛选了代表 PD 进展的关键特征。其次将跨模态的数据融合，应用于 PD 的进展预测，并提出了一种跨模态融合预测进展的方法，有利于临床预测性能的提升。

本研究课题内的各工作环节间存在着紧密的相互联系与支撑。首先，通过引入首个关于多模态医学影像融合的任务，确立了医学语义的核心概念，这一理念继而被融入第二个融合方法中，显著增强了影像的可解释性，为病灶的精确定位与动态追踪提供了有力支持。在实施这两项多模态融合技术的同时，现有神经网络模型在捕获全局特征方面的局限性被意识到，进而促成了第三个工作——即通过优化网络结构以有效拓展感受野，并成功应用于AD的细粒度分类，实现了诊断精度的新突破。进一步地，本文巧妙地将多模态影像融合的思路迁移至跨模态数据融合的领域，特别是在帕金森病（PD）病情进展预测的研究中，这一思路的应用极大提高了预测的准确性。综上所述，这些工作的依次开展不仅在技术层面上形成了层层递进的创新链，还拓宽了技术应用范围，共同促进了智能辅助影像组学技术的发展与临床价值提升。

全文共分为六个章节，整体的篇章结构如图1.3所示。

第一章绪论。本章深入介绍了基于医学影像组学的智能辅助技术的研究背景与意义。强调了基于医学影像组学的智能辅助诊断技术的关键任务，包括影像融合、影像分类与疾病预测等方面。具体而言，本章详细分析了国内外在多模态医学影像融合方面的研究，包括基于传统方法和深度学习方法的不同取向。此外，本章对AD影像分类进行了深入讨论，涵盖了粗粒度分类（AD、MCI、NC的二分类或三分类）和细粒度分类（AD、EMCI、LMCI、MSC、MCI、NC的二分类、四分类和六分类）。最后，本章聚焦于PD预测，探讨了基于影像数据和基于临床数据两个层面的研究现状。整体而言，该章节不仅对研究内容进行了明晰概述，还为后续章节的深入探讨奠定了坚实基础。

第二章基础理论。本章深入研究了医学影像的多模态特性，对不同类型的医学影像成像技术原理进行了详细介绍，包括CT、MRI、PET等。本章详细分析了智能辅助诊断技术的理论基础，涵盖了传统机器学习方法（如SVM）和深度学习方法（如CNN）在医学影像分类中的应用。强调了这两者的融合是提高医学诊断准确性和效率的前沿趋势。另外，本章着重考虑了多模态数据的复杂性，讨论了信息融合的策略，包括像素级、特征级和决策级的融合方法，为实际应用提供了有益的启示。最后，本章还介绍了多种质量评估指标，包括主观与客观评价方法。对于有关医学影像基础理论和信息融合策略的全面认识，为后续章节中的具体技术方法提供了理论指导。

第三章聚焦于多模态影像融合。设计了两种新型的结合两种医学成像技术

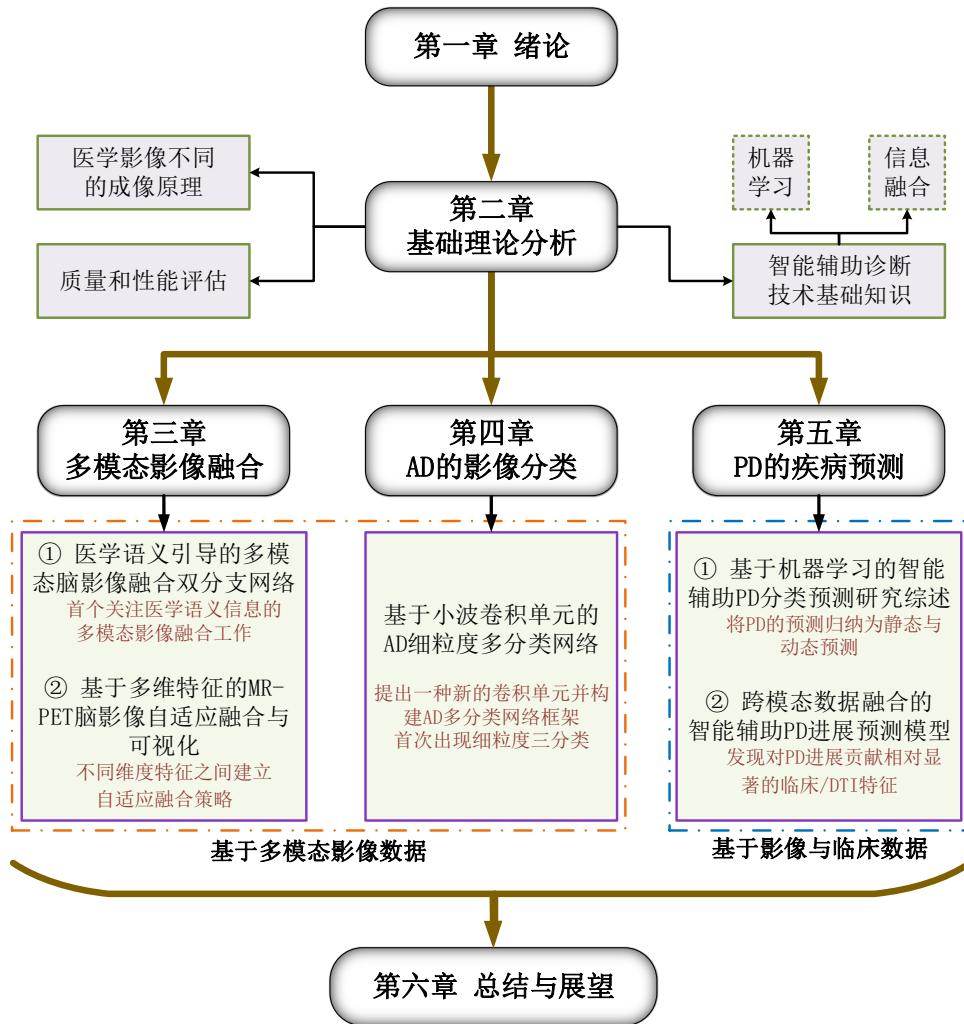


图 1.3: 本研究的内容布局与章节安排

的融合方法，以实现更全面的信息整合。一种是名为 MsgFusion 的基于医学语义信息 (MS-Info) 引导的融合方法。通过分析 MR/CT/PET/SPECT 的关键 MS-Info，该方法获取相应的影像特征，采用 SF 分支和 GV 分支的双分支网络，成功提高了 CNN 的泛化能力。在医学脑影像的处理和分析中，包括 MR-CT 影像融合、MR-SPECT 影像融合和 MR-PET 影像融合等方面，该方法相较于现有方法表现出显著优势，得到了临床医生的高度评价和统计数据的支持。另一种也是基于深度学习的融合框架。它结合了多维特征、空间特征和通道特征，并引入了三种不同的特征提取模块（粗特征模块（coarse feature module, CFM）、精细特征模块（coarse feature module, FFM）和多尺度特征模块（multi-scale

feature module, MFM)) 和自适应线性融合机制。该方法还提出了一种关键特征增强策略，可以增强同一病例不同时期融合影像的可视化效果，有助于肿瘤定位、分割和疾病跟踪等临床应用。

第四章关注 AD 影像分类。首先，提出了一种创新的卷积单元（wavelet convolution unit, WCU），它将小波变换与传统卷积函数结合，以获得更广泛的非局部感受野，显著提升卷积神经网络性能。在 WCU 的引导下，本章设计了影像分类的网络框架（WCU-Net）成功应用于 AD 的细粒度和多分类任务。基于 AD 的脑 DTI 数据（NC、MCI、EMCI、LMCI、AD）实现了所有粗粒度和细粒度的 12 种组合分类，包括首次引入的细粒度三分类。该框架在所有 12 种分类中取得了高准确率，尤其在 8 种细粒度分类中表现卓越。细粒度四分类的准确率仍有提升空间，并计划未来将 WCU-Net 应用于不同疾病的分类和分期。

第五章探讨 PD 疾病预测。首先，本章详细回顾了基于机器学习的 PD 分类预测的研究。全面介绍了静态和动态两种研究方法，而且突出了机器学习在增强 PD 智能辅助诊断和健康评估中的巨大应用前景。本章深入分析了在 PD 预测中所采用的数据类型，并探讨了 PD 预后领域未来的发展潜力和可能的演变路径。其次，本章提出了一种基于跨模态数据融合的智能辅助 PD 进展预测的新方法（CMFP）。它通过集成基线 DTI 和临床特征数据，该方法的独特之处在于采用纵向数据研究方法，结合机器学习技术，建立了早期 PD 基线 DTI 及临床特征的疾病进展模型。研究结果显示，跨模态数据融合显著提升了单模态预测的准确度，并且 DTI 数据对临床预测性能的提升具有显著作用。这一发现为未来的 PD 研究指明了方向，强调了跨模态数据融合在提高预测性能方面的潜在优势。

第六章回顾总结了全文研究内容，阐述了在当前研究工作中所存在的不足，并展望了未来可能的研究方向和工作设想，为基于医学影像组学的智能诊断技术发展提供了实践经验和研究思路。

## 第 2 章 理论与技术基础

因不同的医学影像来源于不同的成像原理，本章深入分析了七种医学影像成像技术，包括 CT、MRI、X 射线等，以满足医学领域辅助诊断的需求。其次，探讨了智能辅助诊断技术理论基础，如传统机器学习方法（SVM）和深度学习方法（CNN）在医学影像分类中的应用，它们的结合是提高医学诊断准确性和效率的前沿趋势。并研究了多种层面的影像融合策略，包括像素级、特征级和决策级的方法，为实际应用提供了启示。最后，引入了较全面的评估方法，包括主观和客观评价，用于评估多模态影像融合与分类算法的性能。

### 2.1 医学影像的成像技术原理

在医学领域，确实如“工欲善其事，必先利其器”所言，一台优秀的智能辅助诊断机器对于提高医疗效率至关重要。然而，很多医院面临的现实是缺乏多模态医学影像扫描设备，限制了对患者的全面诊断。尽管在影像科室，大多数医院也只配备了单一类型的扫描设备。但如果能够使用一种简单的算法实现多模态影像的融合，将会是一项令人振奋的技术进步。不同多模态医学影像具有不同的表现形式，分别如图2.1所示。

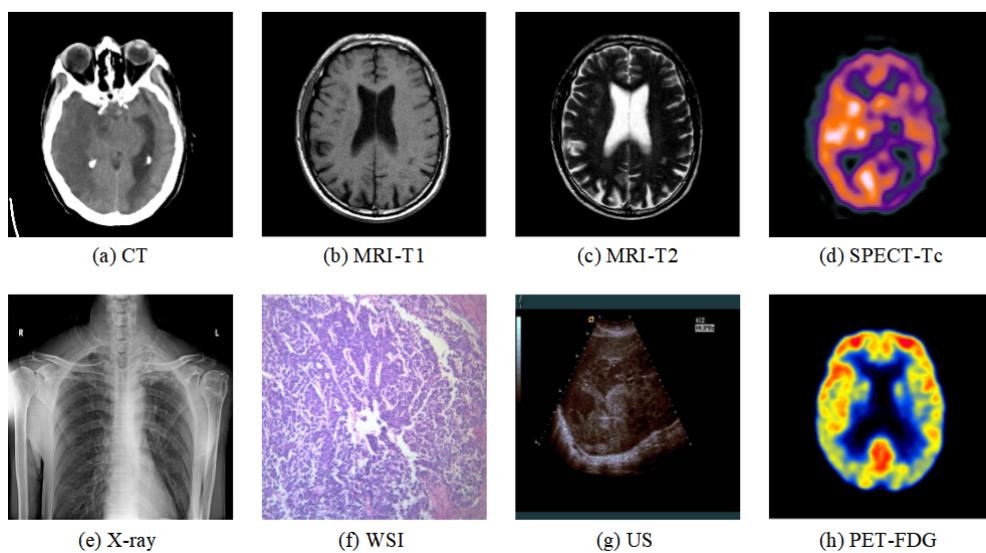


图 2.1：七种不同类型的医学影像

### 2.1.1 基于解剖结构的成像技术

**计算机断层扫描 (CT)** [5] 是一项以 X 射线成像为基础的医学影像技术，通过对人体特定部位进行 X 射线扫描，利用探测器接收透过人体的 X 射线并转换为电信号，最终生成数字化的横断面影像。其工作原理包括 X 射线透过人体产生的信息采集、数字影像处理和影像重建。CT 影像的构建涉及到对扫描区域的体素进行 X 射线吸收系数或衰减系数的测量，这些数据通过计算机处理后形成数字矩阵，再转换为灰度影像。影像的空间分辨力取决于像素大小和数目，CT 影像的像素通常为  $1.0 \times 1.0 \text{ mm}$  或  $0.5 \times 0.5 \text{ mm}$ ，数目可达  $256 \times 256$  或  $512 \times 512$ ，因此，CT 能够提供高度细致的解剖影像，具有较高的空间分辨力。

CT 影像的特色在于其灰度表示方式，黑色代表低吸收区（低密度，如肺部含气体），白色代表高吸收区（高密度，如骨骼），而相对于传统 X 射线影像，CT 的密度分辨力更高，能够清晰地显示组织密度差异较小的软组织结构。这使得 CT 在展示由软组织构成的器官（如脑、脊髓、肺等）以及病变的影像上具有突出的优势。CT 在中枢神经系统疾病的诊断上表现出色，广泛应用于颅内肿瘤、脑梗塞、脑出血、外伤性脑损伤等疾病的检测。尽管 CT 在医学诊断领域取得了显著进展，但其使用也存在一些问题，如辐射暴露和对软组织对比度较低。随着技术的不断改进，CT 仍然是一种重要的影像学工具，对于临床诊断和治疗的支持具有不可替代的作用。

**核磁共振成像 (MRI)** [6] 是一种利用磁场和无线电波探测人体组织的详细横截面影像的技术。在进行磁共振成像 (MRI) 检查时，患者被置于一个强度高且分布均匀的静态磁场内，利用特制的无线电频率脉冲激活体内氢原子的自旋核，从而引发共振反应并捕获相应的信号。这些信号运用计算机的处理后生成高分辨率的影像，展示了人体内部的详细结构。

MRI 的两种主要成像方式：MRI-T1 (T1 反映的是纵向弛豫时间) 和 MRI-T2 (T2 反映的是横向弛豫时间)，通过对比度和信号强度的不同分别对应于不同的组织特性。MRI-T1 成像对脂肪组织具有较高的信号强度，呈现亮白特征，有利于显示解剖结构和脂肪组织的分布，适用于解剖结构和骨髓等的观察。相反，MRI-T2 成像对水分组织有较高的信号强度，呈现亮白特征，更适用于显示水肿、炎症和软组织病变。这种高度区分不同组织的成像方式使 MRI 在临幊上得到广泛应用，涵盖了中枢神经系统、心血管系统、关节、软组织、盆腔脏器等多个部位的病变检查。MRI 在医学影像学中的主要优势之一是对软组织的高

分辨率成像。在头颅、纵隔、腹腔等实质性脏器的检查中，MRI 表现出卓越的成像性能。由于人体组织成分的差异，MRI 能够产生不同的信号，成为一种精准而全面的影像诊断工具。

MRI 技术的不断发展和创新进一步丰富了其应用领域，涵盖多种变体，如静息态磁共振成像 (rsMRI)、sMRI 和 fMRI。每一种变体都在特定领域中展现出独特的优势。sMRI 通过施加外部磁场和无线电波，激发人体组织中的氢原子核，然后测量它们在磁场中的共振信号。通过复杂的数学算法，特别是傅里叶变换，这些信号被转换成详细的影像，揭示了身体内部的精细解剖结构，为肿瘤、损伤和先天性异常等病变提供详细的解剖学信息，在神经外科和器官移植等手术前的评估中得到广泛应用。fMRI 通过测量血氧水平变化，观察大脑在执行任务时的活动，提供时空分布的神经活动信息。在神经科学的研究中，fMRI 起到关键作用，用于研究认知功能、感觉知觉和运动执行，同时在神经精神病学和神经内科等领域用于评估神经系统疾病的功能改变。rsMRI 是一项测量大脑在安静状态下血流和神经活动的成像技术。通过揭示大脑静息状态下的功能连接性，rsMRI 为研究大脑的静息状态、神经网络连接性以及精神疾病、脑功能障碍等提供有益信息。这些 MRI 的变体在医学和研究领域中各自发挥着独特的作用，医生和研究人员可根据不同的研究和诊断需求选择适当的 MRI 变体，共同推动医学影像领域的不断发展，为深入理解人体生理和病理提供了丰富的信息资源，有望为未来的医学研究和诊疗提供更精准、全面的支持。

**超声成像 (US)** [9] 是一种利用超声波对人体进行扫描的技术，它通过捕捉和分析反射回来的声波信号来构建体内器官的影像。医学中的超声仪器涵盖了多种类型，其中 B 型超声是目前应用最为广泛的诊断方法。医学超声波成像是通过超声探头发出的超声波携带不同组织的回波，以构建人体内部结构的影像。医学超声波成像是通过超声探头发出的超声波携带不同组织的回波，以构建人体内部结构的影像。B 模式超声影像 (B 超) 是其中最常见的一种，其优势在于无电磁辐射、便于携带且提供实时成像。

超声诊断依赖于细致的深入观测与评估，识别不同的特征，并将这些信息整合以推断可能的病理原因。超声检查广泛应用于内脏器官的定位、测量及形态分析，同时能够帮助诊断病灶的具体范围及其性质，并提供有关腺体等组织的精确解剖影像，已在眼科、心血管以及妇产科等多个医疗领域得到普遍采用。医生在超声检查中综合评价器官的形态、尺寸、质地、病变的声学特性、对后

壁声波的影响、内部结构、位置关系及功能表现。对于不同类型的病理性影像特点，例如囊性与实质性病变、均质性与非均质性病变、钙化性与含气性病变、炎性与纤维化病变以及良性与恶性病变，医生能够通过综合观察进行诊断和鉴别诊断，为制定治疗方案提供重要参考。

### 2.1.2 基于功能性的成像技术

**单光子发射计算机断层成像术 (SPECT)** [8] 是一种重要的医学成像技术，通过利用放射性同位素的  $\gamma$  射线，提供了对生物体内部功能和生物学过程的独特洞察。该技术基于放射性同位素的放射性衰变原理，通过测量  $\gamma$  光子的产生和探测，以获取有关生物体内部结构和功能的信息。

SPECT 在分辨率方面存在一些限制，如相对较低的空间分辨率和较差的时间分辨率，但其在医学影像学中的广泛应用使其成为疾病诊断和治疗方案制定的不可或缺的工具。SPECT 的优势在于其提供的是功能性信息，而非仅仅是解剖结构。这对于疾病的早期诊断和治疗效果评估至关重要。如：在神经科学领域，局部脑血流断层显像是 SPECT 的一项关键应用。它对缺血性脑血管意外、癫痫致痫灶的定位等提供了高度准确的诊断。这项技术在判断脑肿瘤的血运、鉴别术后或放疗后的复发和瘢痕等方面也具有独特优势。AD 的早期诊断是 SPECT 在神经影像学上的又一亮点。通过对 AD 患者的局部脑血流进行研究，SPECT 为该疾病的早期诊断提供了一致的共识。相较于传统的 CT，SPECT 在痴呆程度和认知状况相近的两类痴呆的鉴别上更为可靠。

然而，SPECT 也存在一些挑战和局限性。其中之一是辐射暴露，因为该技术需要使用放射性同位素。尽管通常在安全范围内，但这仍然是需要考虑的问题。此外，SPECT 的设备和维护成本相对较高，这可能限制其在某些医疗机构的推广和应用。同时，SPECT 依赖放射性同位素的摄入，这可能涉及到一些限制和患者安全的问题。

**正电子发射断层成像术 (PET)** [10] 作为一种核医学成像手段，它依赖于正电子在湮灭过程中释放的辐射光子。在这个过程中，回旋加速器将带电粒子加速，并轰击特定的靶核，进而生成正电子放射性核素。通过精心设计的探测器阵列和高效的符合电路，PET 能够检测到这些湮灭光子对，并据此绘制出体内正电子核素的截面分布图。这种影像能够清晰地展示病变的位置、形状、大小，以及相关的代谢功能，为医生提供宝贵的诊断信息。

在肿瘤学中，PET 成像被广泛用于癌症的早期诊断、分期、治疗效果评估

和复发监测。<sup>18</sup>F-FDG (2-氟-18-氟-2-脱氧-D-葡萄糖) 是一种常用的显像剂，通过显示异常的葡萄糖代谢水平，PET 可以精准定位和描绘肿瘤。在实际应用中，患者接受 <sup>18</sup>F-FDG 药物注射后，PET 扫描仪通过探测正电子湮灭事件获取体内放射性核素分布情况。影像分析利用标准化摄取值 (SUV) 等定量测量，对病变的代谢活性进行评估，为肿瘤生物学特性、治疗反应和复发监测提供关键信息。

在神经学领域，<sup>18</sup>F-FDG PET 在 AD 早期诊断中取得显著进展。通过评估脑部葡萄糖代谢情况，PET 揭示患者大脑功能异常，为 AD 的诊断和病程监测提供关键信息。PET 的高分辨率和功能性信息使其在医学中的应用不断拓展，为多领域的疾病管理提供了更全面、精准的影像学支持。

### 2.1.3 基于其他类型的成像技术

**X 射线** (X-ray) [4] 是一种电磁辐射，在医学影像学中，X 射线被广泛用于探测和成像，广泛用于识别骨骼问题、胸部及腹部疾病，包括肺部与肠道堵塞等状况。X 射线成像主要依赖于光电效应，通过测量透射或散射的 X 射线来获取关于人体内部结构的信息。

X 射线成像的优点之一是其广泛的应用范围和相对较低的成本。X 射线设备相对便宜且易于使用，使得它在许多医疗机构中得到了广泛的普及。此外，X 射线成像速度快，可以提供即时的影像结果，对急诊情况和危重患者的救治具有重要意义。然而，X 射线成像也存在一些缺点和风险。首先，X 射线是一种电离辐射，长期频繁暴露可能对人体产生辐射损伤的风险，特别是对于儿童和孕妇。其次，X 射线成像对柔软组织的分辨率较低，不适合检测早期病变或微小的异常。

**显微镜** [7] 作为一种基于凸透镜成像的仪器，在医学中发挥着广泛的作用。通过物镜和目镜的双重放大过程，显微镜能够将微小物体放大到人眼可分辨的尺寸，为医学研究提供了重要的支持。在医学领域，显微镜的应用涉及组织学、细胞学和病理学等多个方面。通过观察组织切片 (whole slide image, WSI)，显微镜可以提供高分辨率的影像，帮助医生深入了解微观结构，从而支持准确的疾病诊断和治疗方案的制定。

特别是在癌症研究和诊断领域，显微成像技术发挥着至关重要的作用。如：显微镜下的成像可用于对癌症病理学进行详细分析，包括观察组织的完整性、细胞浸润深度以及细胞变化的程度。这些信息对于判断肿瘤的恶性程度至关重要。

要。尽管显微镜具有高分辨率、实时观察和非侵入性等优点，但也存在一些挑战，如有限的穿透深度、专业训练要求和较高的设备成本。近年来，随着技术的不断创新，显微成像在医学领域的应用前景将继续扩展。新兴技术如活体显微镜技术使得在活体动物或人体上进行实时显微观察成为可能，为研究疾病的生长过程、细胞动态和治疗效果提供了更为深入的理解。

表 2.1: 不同的医学影像成像机制、分辨率及优缺点对比

<i>Imaging</i>	成像机制	分辨率	优势	局限性
X-ray	X 射线	[0.1mm, 1mm]	成像速度快、价格低、广泛应用	X 射线有害、软组织成像效果差
CT	X 射线	[0.5mm, 2mm]	成像速度快、价格低、骨组织敏感	X 射线有害、软组织成像效果差、有伪影
MRI	磁共振	[0.5mm, 1.7mm]	软组织成像效果好、无骨伪影	硬组织成像效果差、时间较长
US	超声波	[1.5mm, 2mm]	成本低、无损耗、使用方便	分辨率较低、无法穿透骨骼、空气、存在折射效应
SPECT	$\gamma$ 射线	[3.5mm, 10mm]	对比度高、提供功能性信息、价格较低	分辨率较低、速度慢、放射性同位素的摄入
PET	核	[4mm, 6mm]	高分辨率、提供功能性信息	放射性同位素的摄入、成本较高
WSI	凸透镜	[200nm, 0.2 $\mu$ m]	高分辨率、实时观察、非侵入性	有限穿透深度、需要专业训练、成本较高

不同的成像机制可以获得不同的医学影像，并根据不同的特性应用于不同的疾病诊断中，各自的特性比较如表2.1所示，表中的分辨率通常在给出的参考范围内，具体值取决于设备的类型和扫描参数等使用的条件。X-ray 主要用于颅骨成像，对于骨折和骨质疾病的检测具有优势。CT 提供高分辨率的骨髓和软组织影像，对脑肿瘤的大小和位置具有关键意义。MRI 适用于探测脑部软组织病变，如脑肿瘤、AD 等。US 在产前检查和婴儿脑部检查方面有优势。SPECT 在 AD 早期诊断方面有帮助，而 PET 在脑肿瘤代谢活性和 AD 脑代谢方面表现出

色。显微成像对神经系统疾病的基础研究和理解细胞水平变化至关重要。在实际应用中，这些成像技术的综合运用可以更全面地观察和诊断疾病。

在医学影像领域，倘若有一台能够扫描多模态的医学影像数据的机器会极大地提高诊断效率和准确性。然而，由于成本和设备更新的限制，在国内医疗环境中，许多医院只拥有特定类型的扫描仪，其产生的影像局限于该仪器的特性。这就迫使医生必须根据需要使用不同的设备，增加了工作的复杂性。若有一种简便且高效的算法，能够将来自不同设备的影像进行融合，将为医学影像的多模态融合提供一种经济高效的解决方案。因此，开发出能够实现多模态融合的算法，不仅能够在医学影像的全面性诊断上大显身手，同时也能为医院节省设备购置成本和提高工作效率。这种简单而实用的技术创新，无疑将为医学领域带来福音。

## 2.2 智能影像组学技术的理论基础

### 2.2.1 医学影像组学关键技术介绍

医学影像组学是一个交叉学科领域，它结合了医学影像学、生物信息学、计算机科学和数据科学等多个学科的知识和技术。在过去，医学影像组学的分析往往依赖于传统的手工设计特征，即研究者根据以往的经验和知识来选择和定义特征。在智能辅助诊断方面，最初主要依赖于基于机器学习的方法，这些方法通常需要人工设计特征，即研究者需要根据专业知识来选择或构造用于模型训练的特征。然而，近年来深度学习技术的兴起和发展，使得医学影像组学能够通过神经网络来自动学习和提取特征，这一变革减少了对专家知识的依赖，显著提高了特征提取的效率和准确性。

#### 2.2.1.1 基于传统的机器学习方法

基于传统的 ML 方法，如支持向量机 (SVM)、决策树、随机森林等，以其稳定性和可解释性在医学影像分类任务中取得了显著的成果。本章将多种常见的 ML 方法进行比较，见表2.2，并详细分析经典的 SVM 分类器。

在二分类模型中，支持向量机 (SVM) 是一种广泛应用的方法。其核心分类思想是寻找一个能够将样本分割并具有最大间隔的超平面。这里的“间隔最大”指的是确保超平面与其最近的样本点的距离最大。通过最大化间隔，SVM

表 2.2: 常见的基于传统的机器学习方法对比

ML 方法	优势	局限性
支持向量机	SVM 模型只与支持向量有关、高维数据中只需少量的样本、节省了内存，鲁棒性强	非线性 SVM 需要进行核函数映射、计算开销较大
贝叶斯	实现简单、收敛速度快、聚类效果较优、无参数训练模型	$k$ 需要人为设定、只能得到局部最优解、抗干扰差、非凸数据集难以收敛
决策树	对数据要求度低、速度快、准确性高、适合高维数据	在训练数据上比较耗时、在面对复杂的高维数据时容易过拟合
随机森林	实现简单、准确性较高、适合高维数据	整组数据的决策树上容易过拟合、训练时间较长
逻辑回归	计算速度快、适合二分类、内存占用小、模型可解释性好	不能解决非线性问题、很难处理数据不平衡的问题、准确率不高
K-最近邻	实现简单、可直接在测试数据上查找最近的邻居来进行分类	依赖于距离度量、计算成本高、不适用于高维稀疏数据

能够提高分类预测的可信度。如图2.2，中间的实线即为所求的最优超平面，虚线边界上的点被定义为支持向量。最优超平面到两侧虚线边界的距离是相同的，这个距离称为几何间隔，Margin 即为两倍的几何间隔，其计算为  $2/\|\mathbf{w}\|$ 。y 是分类面，其计算为  $\mathbf{w}x+b$ 。线性可分 SVM 可归为凸二次规划问题，针对该类问题可以考虑用 Lagrange 乘子法进行求解，目标函数为公式 (2.1)。

$$\min \frac{1}{2} \|W\|^2 \quad (2.1)$$

$$s.t. \quad y_i (W^T x_i + b) \geq \pm 1$$

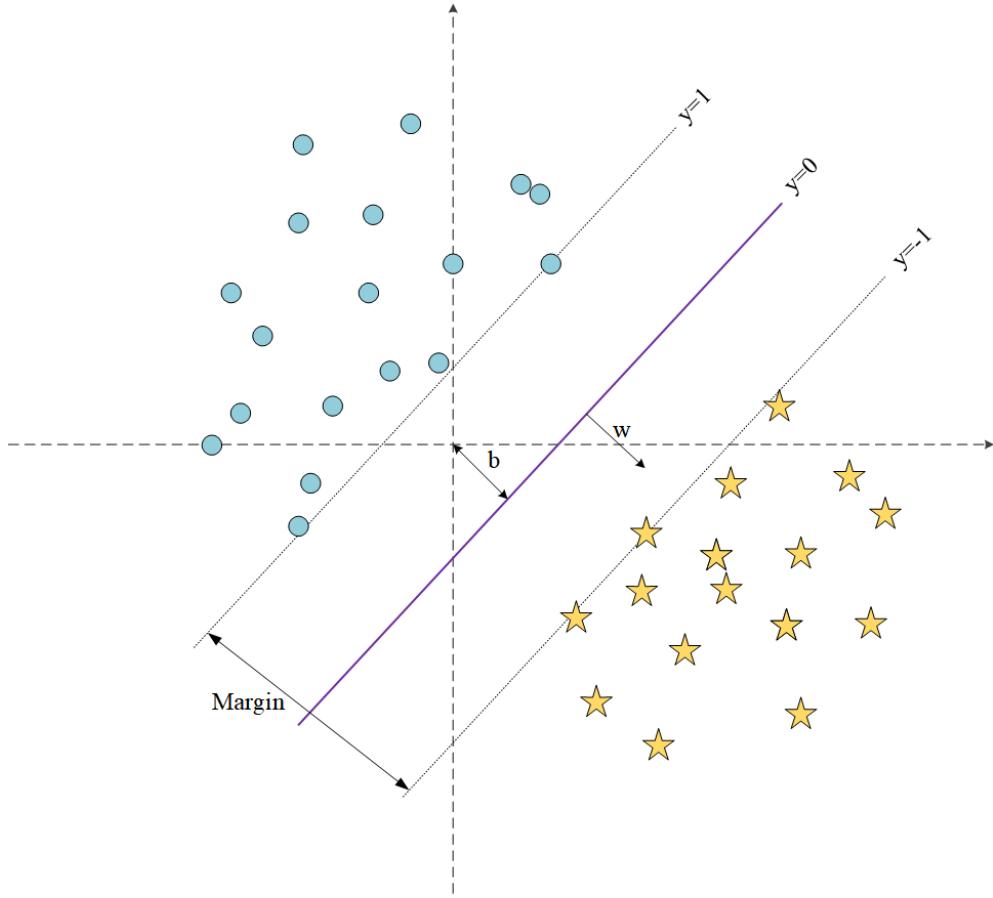


图 2.2: 线性可分 SVM

线性可分 SVM 在面对非线性可分的样本集时显然存在局限性。为了解决这一问题，SVM 采用了核函数  $\phi(x)$  的方法，典型的核函数包括 RBF（径向基函数），poly（多项式核函数），linear（线性核函数），sigmoid 等。SVM 使用这些核方法把数据映射到高维，以便更好地分离。具体而言，非线性可分 SVM 的模型可以表示为式 (2.2):

$$f(x) = W^T \phi(x) + b \quad (2.2)$$

其中， $\phi(x)$  表示样本在高维空间中的映射， $W$  是对应的权重向量， $b$  是偏置项。通过这样的映射，原本在低维空间中无法线性分隔的样本，在高维空间中找到了一个线性超平面，从而完成了分类任务。这种非线性可分 SVM 的处理方式为算法赋予了更大的灵活性，使其能够更好地适应实际场景中存在复杂

结构和非线性关系的数据。

SVM 作为一个多才多艺的分类器，不仅在 ML 领域取得了成功，而且在医学影像处理中表现出色，为医学领域的疾病诊断和预测提供了强大的支持。其能够同时应对线性和非线性的分类任务，为医学影像的自动化分析和医生的决策提供了可靠的辅助。

### 2.2.1.2 基于前沿的深度学习方法

基于 DL 方法的发展更是推动了医学影像处理的前沿。卷积神经网络 (CNN) 和循环神经网络 (RNN) 等深度学习架构通过学习影像的高级特征表达，使得模型能够自动从数据中学习到更抽象、更复杂的表示，从而在医学影像分类中取得了卓越的性能。DL 方法的主要优势在于其端到端的学习能力，对于复杂和大规模数据集的处理能力更为突出。

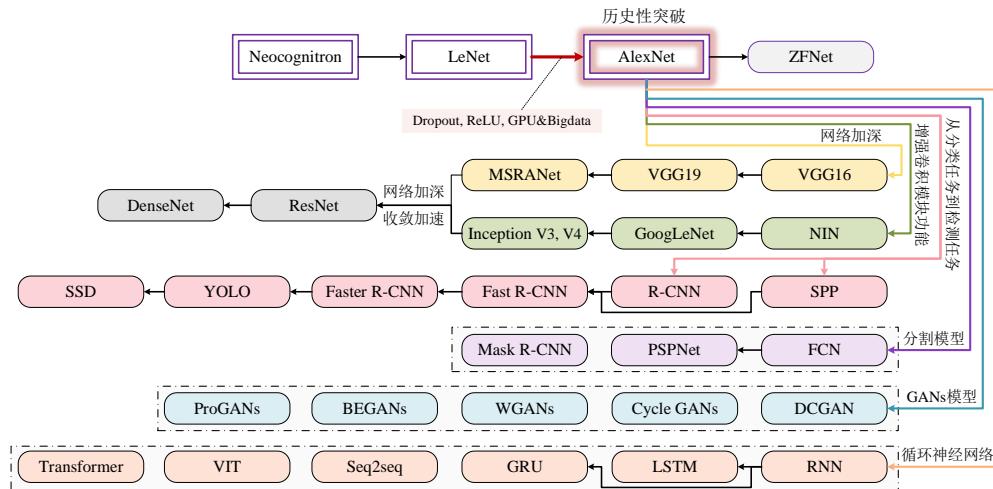


图 2.3: 深度神经网络发展脉络

CNN 代表了 DL 领域的重要里程碑，其模拟了生物神经元结构的并行计算方式，能够在分类任务中自动学习和提取数据中的抽象特征，从而无需依赖手工设计的特征。CNN 相对于传统机器学习方法在影像融合、分类、预测等任务中具有显著优势。其核心技术包括卷积、激活、批归一化、池化和全连接。在影像分类领域，CNN 通过卷积操作实现对输入影像的特征提取，利用局部感知和权值共享机制捕捉影像的局部特征，使得模型能够更好地适应数据的复杂结构。激活函数引入非线性，如 Sigmoid、Tanh、Softmax 和 ReLU，赋予模型对复杂关系的学习能力。批归一化在神经网络的中间层中规范输入，加速训练过

程，提高模型的稳定性。池化操作通过降采样减小特征图的尺寸，有助于保留重要信息并降低计算复杂度。全连接层整合局部特征，并将其映射到样本标签空间，完成最终分类。

CNN 在实际应用中展现出卓越的性能，尤其在计算机视觉领域，为处理庞大而复杂的数据集提供了高效而强大的解决方案。图2.3为 CNN 的大致发展脉络。其中，LeNet[119] 作为 CNN 的经典之作，为 DL 领域的发展奠定了坚实基础，成为影像分类和识别领域的开创性工作。AlexNet[120] 在某种程度上继承并发展了 LeNet 的思想，并且引入了一些新的技术，如在 CNN 中广泛使用了 ReLU 激活函数，使用了 Dropout 技术防止过拟合，并采用了 GPU 加速训练过程。AlexNet 的模型结构如图2.4所示，它在 2012 年的 ImageNet 大规模视觉挑战赛中取得了冠军，其不仅在性能上大幅超越了当时的其他方法，而且其成功也证明了 DL 在处理大规模视觉识别任务中的巨大潜力，极大地促进了 DL 在各个领域的应用和发展。

AlexNet 等 CNN 网络模型在医学领域的应用为医学影像分析、疾病诊断和个性化医疗等方面带来了巨大的进步，推动了医学影像组学分析技术的发展，为医学诊断和治疗提供了全新的思路和工具。一些经典的网络模型在医学领域取得了显著的成果，如 FCN[121] 作为一种专门用于影像分割的模型，特别适用于医学影像的病灶分割等任务。U-Net[122] 是一种用于生物医学影像分割的经典网络模型，结构简单，同时在分割任务中表现出色。因此，医学领域广泛采纳并应用了这项技术。总的来说，CNN 网络模型在医学领域的应用为医疗诊断、疾病预测和治疗方案的制定带来了更多的可能性和机会，有助于提高医疗保健的水平和效率。

## 2.2.2 影像组学技术的信息融合策略

信息融合涵盖了不同医学影像之间的融合（多模态融合），例如 CT 与 SPECT 的融合、CT 与 MRI 的融合、MRI 与 SPECT 的融合甚至是三种影像的融合。同时也包括影像和其他异构数据之间的融合（跨模态融合），如临床指标与医学影像之间的融合。作为一种关键的影像处理策略，尤其在多模态医学影像领域，影像信息融合通过综合多种影像信息并采取相应的融合策略，可以显著提高影像的质量和信息量。常见的融合方法主要分为三种：像素级、特征级以及决策级融合。

这三种融合方法在多模态影像融合、AD 分类和 PD 预测的研究中备受关

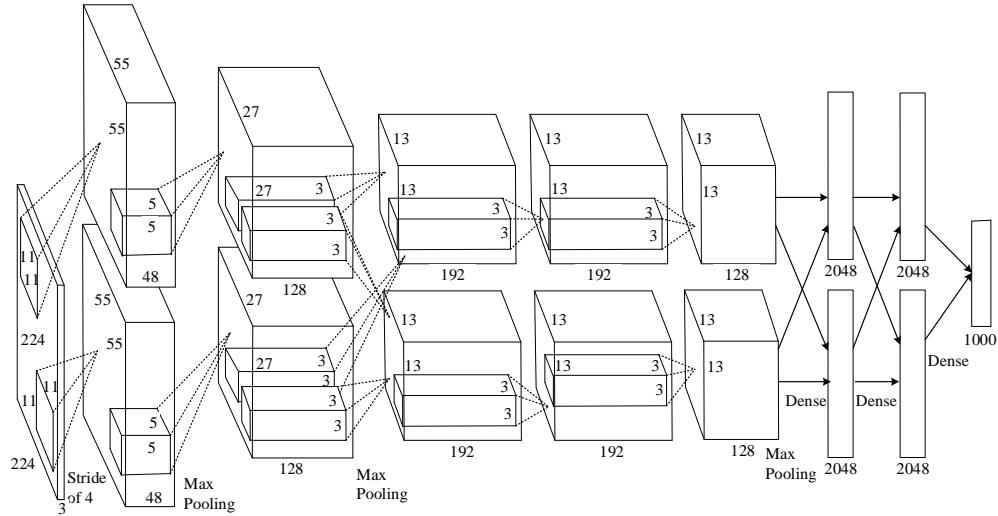


图 2.4: AlexNet[120] 网络结构

注。在多模态影像融合方面，这些方法为整合不同成像模态提供了高效手段，有助于充分利用各种影像信息，提高影像的全面性和信息量。在 AD 分类和 PD 疾病预测领域，融合方法的应用为更准确、可靠的疾病诊断和预测提供了新的思路和工具。通过综合不同模态的信息，可以更全面地了解患者的病理状态，为个性化治疗和康复提供更为精准的支持。这些研究不仅推动了医学影像处理技术的进步，也在提升脑部疾病的辅助诊断水平方面取得了显著的研究成果。

### 2.2.2.1 基于像素级的技术方案

像素级融合技术涉及在影像处理的基础层面上，对原始影像数据的像素点进行直接操作和整合，形成新的像素值，进而生成融合影像。这一方法在影像信息融合中起到重要作用，特别是在像 PET 和 MRI 影像的融合中。像素级融合技术能够有效地结合不同影像源的信息，从而增强影像细节，这有助于提升疾病诊断的准确性。将 PET-MRI 影像对采取像素级融合可以充分发挥 PET 在发现受试者脑部葡萄糖代谢变化方面的优势，尤其在 AD 患者中，PET 可揭示脑部葡萄糖摄取降低、导致脑萎缩的情况。虽然 PET 影像本身的细微解剖结构分辨率不甚理想，但通过与 MRI 技术的联合应用，能够显著提升影像的整体清晰度和细节丰富度。二者的融合可以克服各自的局限性，为痴呆患者的病程分析提供更全面的信息。通过综合 PET 和 MRI 的数据，可以更准确地评估 AD 患者的病变程度，为医生提供更为精确的诊断和治疗决策支持。因此，像素级

影像融合在神经学领域中的应用有望在脑部疾病的诊断和治疗方面发挥关键作用。

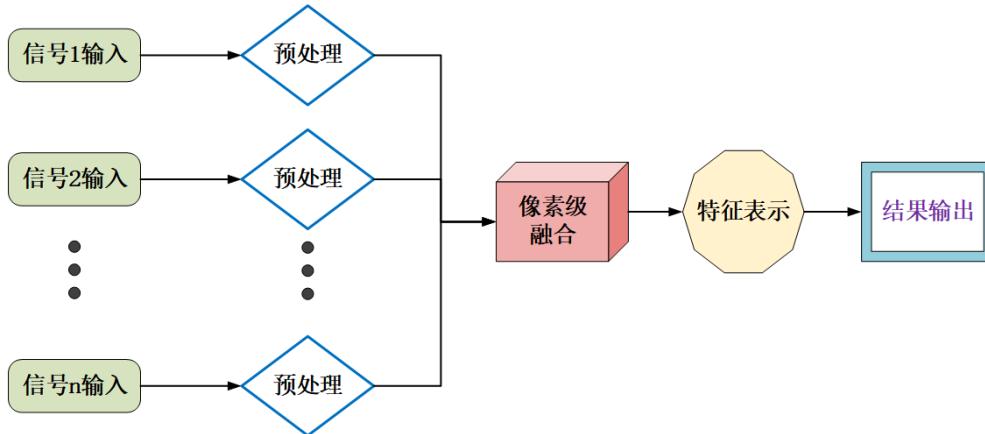


图 2.5: 像素级信息融合机制

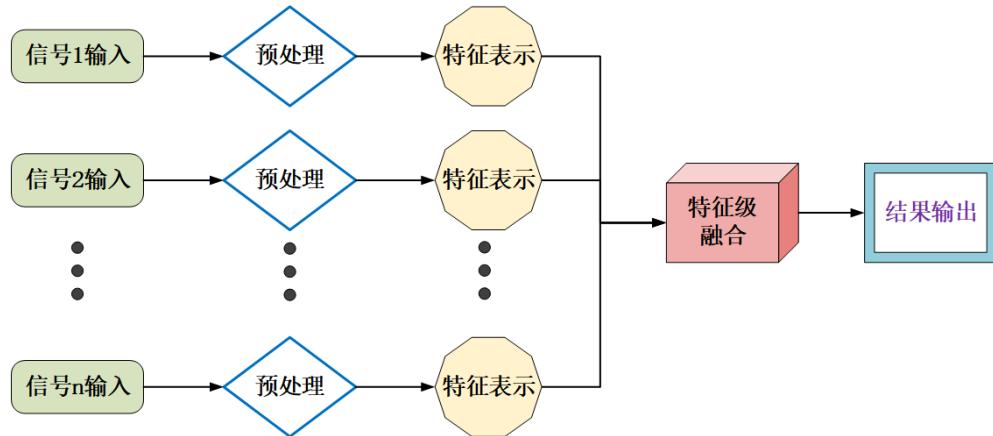
如图2.5所示，像素级影像融合通过各类不同和类型的输入数据执行预处理操作，并采用特定的计算策略计算新的像素点数值，最终生成融合影像。这种融合策略有效地整合了来自不同来源的信息，引入更加细致的信息，并强化了影像的相关特征，具备高度的精度。像素层面的影像融合的独特之处在于能够最大限度地保留源影像中的内容，使得融合后的影像内容和细节都得到了增强。该方法有助于突显潜在目标，判断和识别潜在目标的像素点，为影像的理解和分析处理提供有力支持。

尽管像素级影像信息融合具有显著的优势，但也面临一些局限性。首先，由于像素级融合涉及大量数据的计算处理，可能导致较长的处理时间，实现实时处理与显示是一项挑战；其次，由于数据量庞大，在通信传输过程中，此类信息更易受到噪声影响。此外，如果参与融合的源影像没有经过精准配准处理，则融合后的影像可能出现模糊不清、目标定位不准确以及细节信息失真的问题。

### 2.2.2.2 基于特征级的技术方案

特征级影像融合通过先对不同模态原始数据进行预处理，接着运用各类模型提取各自特有的信息特征，再依据特定融合规则，将这些各异模态数据的特征有效整合起来，如图2.6所示。通过深度学习等技术，特征层面的影像融合技术可以自动学习并提取不同模态的有重要信息的特征，从而更好地揭示目标的

多方面信息。这对于医学影像处理领域尤为关键，例如在神经影像学中，结合 CT 和 MRI 的特征级融合有助于更全面地理解脑部结构和病变。



从源影像中提取特征通常指的是识别和量化影像中的关键信息，这既包括整个影像的全局特征，也包括特定区域的局部特征（如感兴趣区域，即 ROI）。例如，在医学成像领域，MRI 和 CT 影像的特征提取旨在识别出有助于诊断的关键信息，并将这些信息进行分析和处理，以便用于后续的影像融合。通过这种方法，可以从 MRI 和 CT 影像中提取出有用的特征，然后将这些特征结合起来，形成一个更为全面的特征表示。这种特征级的影像融合方法能够有效地减少数据的冗余，提高信息的密度，从而在分类和其他影像处理任务中达到更高的准确度。相比于直接在像素级别进行融合，特征融合能够显著降低所需的计算资源和存储空间，因为它只关注影像中的重要信息，而不是每一个单独的像素。同时，特征级融合的计算速度较快，虽然对影像匹配精确度的要求相对较低。这种方法为在医学影像处理领域提供了更高效的影像信息融合策略。相比于其他方法，特征级影像融合在内存和时间上的需求较低，因此，其具有更迅速的计算效能，并能够有效提升影像处理的实时响应能力。但也可能损失部分细节性特征。

### 2.2.2.3 基于决策级的技术方案

决策级影像融合技术被视为一种高层次的认知整合策略，它代表了影像融合技术中最复杂与最抽象的手段。此方法先将每幅影像依据各自的特征信息进行独立识别与分类处理，然后结合每个源影像的决策结果，按照一定的规则和

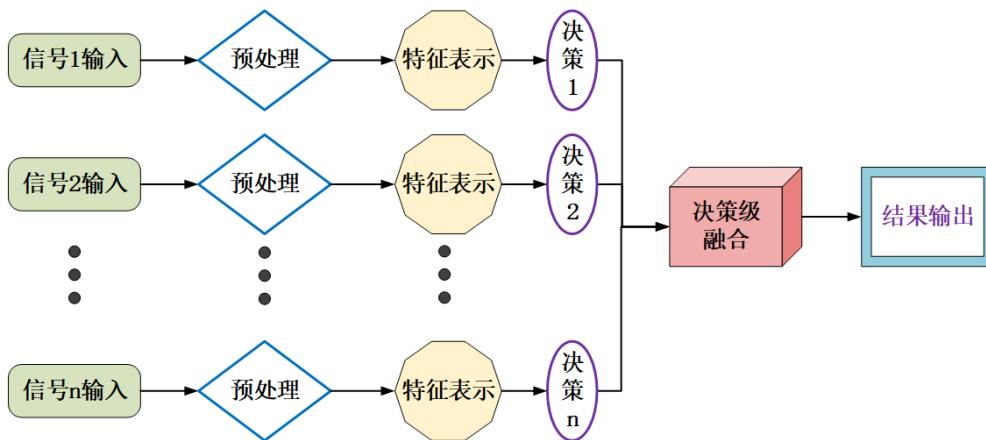


图 2.7: 决策级信息融合机制

置信度权重，如采用贝叶斯推理或是集合决策机制等，对这些决策结果进行综合集成，从而得出最优的整体判断。以 SPECT 和 CT 影像为例，在这一层级上，首先分别对两种不同模态的数据完成独立的诊断决策，随后遵循特定准则，将这两种影像的决策输出相结合，目的是为了实现整体最佳的诊断效果，进而准确区分 AD 病患的不同严重程度，该策略的实现如图2.7所示。

这种决策级影像融合方法在医学影像领域具有重要意义。通过将多个模态的独立决策结果合并，可以得到更全面、准确的诊断信息，提高对疾病状态的认知。例如，在 AD 的诊断中，结合 SPECT 和 CT 影像的决策结果，可以更全面地评估患者的脑部结构和功能，有助于提高对疾病的早期诊断和分级。这种影像融合方法的优势在于能够综合不同模态影像的信息，克服单一模态的局限性，为更准确的医学影像诊断提供支持。

决策层面的影像融合技术位于影像融合方法的顶层，它以认知理论为核心，针对具体问题需求实施深层次、有目的性的信息整合。该融合技术具有数据处理要求低、抗扰动稳定性强等优势。然而，与影像的像素级或特征级的策略相比，决策级影像融合存在明显的缺点。其关键挑战在于对前序步骤输出结果的高度依赖性，且对于影像集成的要求更为复杂。在影像集成的决策层，不仅要对每幅影像独立作出判断，而且还必须针对合并前的多项复杂预处理操作作出决策，这样的处理使得决策级融合的过程变得更加精细和复杂。

## 2.3 智能影像组学技术的性能评估

在智能辅助诊断应用研究中，对方法性能的评估是至关重要的一环。评估手段可以划分为基于主观判断的评估方法和依赖客观数据的评估方法两类。主观评估通常采用主观视觉评价方法，特别适用于多模态影像融合方法的评估。这种评估方法依赖于人眼对影像质量和信息丰富度的感知，通过人工观察和比较来评估影像融合结果。主观评估的优势在于能够捕捉到人的感知差异，但缺点在于受主观因素和个体差异的影响，结果可能不够客观。

客观评估则采用一系列已有的评估指标来量化方法的性能。在多模态影像融合中，常用的客观指标包括信息熵（Entropy, EN）[123]、互信息（Mutual Information, MI）[124]、结构相似性（Structural Similarity Index, SSIM）[125]、标准差（Standard Deviation, SD）[126]、空间频率（Spatial Frequency, SF）[127]、视觉信息保真度（Visual Information Fidelity, VIF）[128]等。这些指标能够从不同的角度和维度评估影像融合的效果，如影像的清晰度、对比度、保真度等。在涉及分类或预测性能进行衡量时，通常采用的客观标准包括准确率（Accuracy, ACC）、召回率、F1分数、ROC（Area Under the ROC Curve, AUC）、均方误差（Mean Squared Error, MSE）等。这些指标直接衡量了模型在任务完成过程中的性能，为研究人员提供了客观、量化的评价标准。

通过主观和客观两种评估方法的结合，可以更全面地评价医学影像融合方法的性能，确保其在实际应用中具有良好的可靠性和准确性。这种评估方式有助于选择最适合特定医学应用场景的影像融合方法，旨在为临床医学领域研发更为精确且值得信赖的辅助诊断手段。

### 2.3.1 客观评估手段

#### 2.3.1.1 影像质量的分析与评估

为了精确评价融合后影像的质量，通常采用一系列相关的影像指标来进行客观评价。这些指标涵盖了影像的不同方面，基于信息理论、影像特征、结构、和基于人类视觉感知[128]。以下是一些常用的客观评价指标，它们有助于全面了解影像融合结果的性能：

- (1) 基于信息理论的评估指标包含 EN[123]、MI[124]、FMI[129] 等。

信息熵（Entropy, EN）作为一种度量手段，可用来量化影像所包含的信息内容多少，计算公式为 (2.3)，其中  $P(x_i)$  是影像的灰度级别的概率。熵越高，

影像信息越丰富 [123]。

$$EN = - \sum_{i=1}^n P(x_i) \log P(x_i) \quad (2.3)$$

互信息 (Mutual Information, MI) 一种用来评估两个图像所含信息量的相互依存程度的度量，数值越大表示它们共享的信息量越丰富 [124]。公式 (2.4) 定义了 MI 的计算方法，它涉及联合概率密度函数  $p(m, n)$  及其对应的边缘概率密度函数  $p(m)$  与  $p(n)$ 。

$$MI(M, N) = \sum_{m \in M} \sum_{n \in N} p(m, n) \log \left( \frac{p(m, n)}{p(m)p(n)} \right) \quad (2.4)$$

特征互信息 (Feature Mutual Information, FMI) 用来衡量特征间的互信息，值越高表示效果越好 [129]。其定义为 (2.5)，其中  $F$  表示融合影像， $A$  和  $B$  表示源影像， $H_F$ 、 $H_A$  和  $H_B$  分别是影像  $F$ 、 $A$  和  $B$  基于直方图的熵， $I_{FA}$  和  $I_{FB}$  分别是融合影像  $F$  与源影像之间的特征信息量。

$$FMI_F^{AB} = \frac{I_{FA}}{H_F + H_A} + \frac{I_{FB}}{H_F + H_B} \quad (2.5)$$

(2) 基于影像特征的评估指标主要包括 SF[127]、SD[126]、 $Q_E$ [130]、AG[131]、SCD[132] 等。

标准差 (Standard Deviation, SD) 衡量的是影像灰度级别的离散程度，较高的标准差代表影像对比度更强 [126]。计算公式为 (2.6)。其中， $\mu$  是融合影像  $F$  的均值。

$$SD = \sqrt{\frac{1}{H \times W} \sum_{x=1}^H \sum_{y=1}^W (F(x, y) - \mu)^2} \quad (2.6)$$

边缘相关融合质量 (Edge Correlation Fusion Quality,  $Q_E$ ) 用于衡量边缘信息的融合质量，值越高越好 [130]。计算公式为 (2.7)。其中， $X_i$  和  $Y_j$  是两影像的像素值， $\bar{X}$  和  $\bar{Y}$  是平均像素值。

$$Q_E = \frac{\sum_{i=1}^N \sum_{j=1}^N (X_i - \bar{X})(Y_j - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2 \sum_{j=1}^N (Y_j - \bar{Y})^2}} \quad (2.7)$$

空间频率 (Spatial Frequency, SF) 衡量影像的空间变化程度, 值越高越好 [127]。计算公式为 (2.8)。其中,  $F$  表示融合影像, 水平与垂直方向上的空间频率 (RF 与 CF) 分别定义为 (2.9) 和 (2.10):

$$SF = \sqrt{RF^2 + CF^2} \quad (2.8)$$

$$RF = \sqrt{\frac{1}{H \times W} \sum_{x=1}^H \sum_{y=2}^W [F(x, y) - F(x, y-1)]^2} \quad (2.9)$$

$$CF = \sqrt{\frac{1}{H \times W} \sum_{x=2}^H \sum_{y=1}^W [F(x, y) - F(x-1, y)]^2} \quad (2.10)$$

平均梯度 (Average Gradient, AG) 可以用于衡量融合影像的清晰程度, 可以认为平均梯度越大, 影像清晰度越好, 融合质量越好 [131]。计算公式为 (2.11)。其中,  $I$  表示融合影像,  $M$  与  $N$  分别表示影像的高和宽  $I(i, j)$  表示在  $(i, j)$  位置的像素值。

$$AG = \frac{1}{(M-1)(N-1)} \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} \sqrt{\frac{(I(i+1, j) - I(i, j))^2 + (I(i, j+1) - I(i, j))^2}{2}} \quad (2.11)$$

(3) 基于结构相似性的评估指标主要包括 SSIM[125]、 $Q_S$ [128]、MSE[133]。

结构相似性 (Structural Similarity, SSIM) 用于综合考虑亮度、对比度和结构信息评价影像质量 [125]。值越接近 1 越好, 计算公式为 (2.12)。其中,  $\mu_X$  和  $\mu_Y$  是两影像的均值,  $\sigma_X$  和  $\sigma_Y$  是标准差,  $\sigma_{XY}$  是协方差,  $C_1$  和  $C_2$  是常数。

$$SSIM(X, Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \quad (2.12)$$

$Q_S$ [134] 是基于 SSIM 的另一种结构策略, 计算如公式 (2.13) 所示。其中,  $w$  表示为窗口尺寸, 局部权重  $\lambda_w$  定义如式 (2.14)。 $s(A_w)$ 、 $s(B_w)$  分别指代两影像  $A$  与  $B$  在窗口为  $w$  的方差。

$$Q_S = \begin{cases} \lambda_v SSIM(A_v, F_w) + (1 - \lambda_w) SSIM(B_w, F_w), & SSIM(A_w, B_w|w) \geq 0.75 \\ \max(SSIM(A_v, F_w), SSIM(B_w, F_w)), & SSIM(A_w, B_w|w) < 0.75 \end{cases} \quad (2.13)$$

$$\lambda_w = \frac{s(A_w)}{s(A_w) + s(B_w)} \quad (2.14)$$

均方误差 (Mean Square Error, MSE) 是一种基于像素差异的影像质量客观评估标准, 它衡量的是融合影像与理想参考影像间的偏差程度 [133]。当 MSE 值越低时, 表明融合影像的质量更接近或优于参考影像。计算公式为 (2.15),  $n$  表示样本数量,  $y_i$  表示实际值,  $\hat{y}_i$  表示预测值。

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (2.15)$$

(4) 基于人类视觉感知的评估指标, 主要包括 VIF[128]、以及基于源影像与生成影像的评估指标 (CC[135]、 $Q^{ab/f}$ [130] 等)。

基于视觉信息保真度 (Visual Information Fidelity, VIF) 用于衡量影像的视觉信息保真度, 值越高越好 [128]。计算公式为 (2.16)。 $k$  和  $b$  分别表示子带和块的索引。 $g_{k,b}$  是第  $k$  子带处的第  $b$  块中的标量增益,  $s_{k,b}^2$  是从像素的局部方差计算的最大似然估计。 $\sigma_{k,b}^r$  表示第  $b$  个块中的源影像  $I_r$  在第  $k$  个子带处的标准偏差。 $\sigma_N^2$  表示噪声  $N$  模型的协方差。

$$VlF(I_r, I_d) = \frac{\sum_k \sum_b \log_2 \left( 1 + \frac{s_{k,b}^2 (\sigma_{k,b}^r)^2}{((\sigma_{k,b}^d)^2 - s_{k,b}^2 \cdot (\sigma_{k,b}^r)^2 + \sigma_N^2)} \right)}{\sum_k \sum_b \log_2 \left( 1 + \frac{(\sigma_{k,b}^r)^2}{\sigma_N^2} \right)} \quad (2.16)$$

$$g_{k,b} = \sigma_{k,b}^{r,d} / (\sigma_{k,b}^r)^2 \quad (2.17)$$

相关系数 (Correlation Coefficient, CC) 评价融合影像, 就是用的皮尔逊

相关系数 [135]，计算公式如 (2.18)。

$$CC = \frac{(r_{AF} + r_{BF})}{2} \quad (2.18)$$

$$r_{XF} = \frac{\sum_{i=1}^M \sum_{j=1}^N (X(i,j) - \bar{X})(F(i,j) - \mu)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N (X(i,j) - \bar{X})^2 \left( \sum_{i=1}^M \sum_{j=1}^N (F(i,j) - \mu)^2 \right)}} \quad (2.19)$$

$Q^{ab/f}$  是一种创新的无需参考影像即可对融合影像进行客观质量评估的方法 [130]，计算  $Q^{ab/f}$  的算法通过局部性指标来估算融合影像中所保留的来自源影像显著信息的程度，该值越高，则表明融合影像的质量越优。计算公式为 (2.20)。其中， $(i, j)$  表示坐标信息， $w_{i,j}^I$  是  $Q_{i,j}^{IF}$  ( $I = A, B$ ) 的权重， $Q_{i,j}^{IF} = Q_{g,i,j}^{IF} \cdot Q_{\alpha,i,j}^{IF}$ 。 $Q_{k,j}^{IF}$  ( $k = g, \alpha$ ) 评估源影像 ( $I$ ) 与融合影像 ( $F$ ) 在边缘尺寸及方向上的一致性。

$$Q^{AB/F} = \frac{\sum_{i=1}^M \sum_{j=1}^N (Q_{i,j}^{AF} w_{i,j}^A + Q_{i,j}^{BF} w_{i,j}^B)}{\sum_{i=1}^M \sum_{j=1}^N (w_{i,j}^A + w_{i,j}^B)} \quad (2.20)$$

在影像融合的性能评价中，主要关注两个核心要素：融合影像的质量和算法的处理效率。评估融合影像质量时，需同时兼顾主观的视觉体验和客观的量化指标。融合算法的客观评估指标往往只能从特定角度对融合结果进行评价，缺乏全面性。尤其是在医学影像领域，通用指标的匹配度受到限制。此外，融合的医学影像主要用于辅助临床诊断，因此其主观评价结果同样具有重要性。总体而言，医学影像融合算法的综合评估需要综合考虑主观和客观两个层面，以确保融合结果在视觉感受和实际医学应用中的双重可信度。这种综合评估方法有助于更全面、准确地衡量影像融合算法的性能，为进一步的改进和优化提供有力的指导。

### 2.3.1.2 模型性能的检验与评定

在机器学习中有很重要的一点就是评级指标，这是判断算法性能很重要的、很有必要的一个评判标准，如 MSE、RMSE、MAE、MAPE 等 [133]。

均方根误差 (Root Mean Square Error, RMSE) 的计算方法是先平方、再

平均、然后开方，计算如 (2.21)。均方根误差 (RMSE) 是一个衡量观测数据与真实值之间偏差大小的度量。类似地，根据公式 (2.15) 所示，预测值与实际值的接近程度决定了模型的准确度；预测值与实际值差异越小，模型准确度越高；差异越大，模型准确度越低。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (2.21)$$

平均绝对误差 (Mean Absolute Error, MAE)，计算公式是 (2.22)。其本质为各测量值绝对偏差的平均程度，即对所有测量值与真实值间绝对差值求平均。相较于其他误差度量方式，MAE 避免了正负误差相互抵消的情况，能够更直接、精确地反映预测结果与实际值之间的差距大小。和 MSE、RMSE 类似，当预测值和真实值的差距越小，则模型越好；相反则越差。

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (2.22)$$

平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE)，计算公式是 (2.23)。该指标常被用于评估预测准确性，特别是在时间序列预测等场景中，因其能够以百分比形式表示预测值与真实值之间的平均偏离程度，从而有效量化模型的预测精确性。MAPE 为 0 表示完美模型，MAPE 大于 1 则表示劣质模型。

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (2.23)$$

在医学领域执行分类和预测等任务中，分类准确率 (Accuracy, ACC) 是一个常见的衡量标准；而在医学疾病检测领域中，除了依赖准确率外，还特别重视敏感性 (Sensitivity, SEN) 和特异性 (Specificity, SPE) 这两个指标作为评价疾病检出效果的重要依据。ACC、SEN 和 SPE 三种评价指标 [136] 的计算方式分别定义为 (2.24)、(2.25) 和 (2.26)。

$$ACC = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (2.24)$$

$$SEN = \frac{TP}{(TP+FN)} \quad (2.25)$$

$$SPE = \frac{TN}{(TN+FP)} \quad (2.26)$$

在上述三个公式中，真正例 (True Positive, TP) 是指被模型正确识别为患病的个体数量，真负例 (True Negative, TN) 是指被模型准确判断为健康状态的非患者个体数量，假正例 (False Positive, FP) 表示正常人被错误预测为患者的数据，假负例 (False Negative, FN) 表示患者被错误地预测为正常的数量。ACC 是指在所有样本中正确预测的比例；SEN 是指在所有病人样本中被正确识别为病人的比例；SPE 是指在所有正常样本中被正确识别为正常的比例。这些指标共同用于评估分类模型的性能。ACC 反映了整体的分类性能，而 SEN 和 SPE 则更专注于模型对患者和正常人的区分程度。

在医学决策中，这些指标的合理权衡对于确保模型对患者和正常人的辨别能力至关重要，以避免误诊或漏诊的风险。因此，在模型评估中，综合考虑这些指标可以更全面地评估分类模型的性能。另外，在多模医学影像分类中也为了评估分类器的性能，通常会采用接收者操作特征 (ROC) 曲线。该曲线的横轴表示假阳性率 (FPR)，也就是误诊率，它衡量的是错误判断的阴性案例占总阴性案例的比例，其数学表达式如 (2.27) 所示。而曲线的纵轴代表真阳性率 (TPR)，它反映了正确识别的阳性案例占总阳性案例的比例，其定义如 (2.28) 所述。

$$FPR = \frac{FP}{FP+TN} \quad (2.27)$$

$$TPR = \frac{TP}{TP+FN} \quad (2.28)$$

AUC 指的是 ROC 曲线下的面积，可以定量分析分类器的好坏，AUC 值越大，则分类器的性能越好，公式定义为 (2.29)。其中， $M$  表示正例总数， $N$  表示负例总数， $P$  表示正例， $Rank_i$  是样本  $i$  的排名，得分最高的排名为  $N$ ，以此递减，累加正例样本的排名可最终计算出 AUC 的值。

$$AUC = \frac{\sum_{i \in P} Rank_i - \frac{M(1+M)}{2}}{M \times N} \quad (2.29)$$

不同的评估指标从多个维度综合评价算法性能，有助于全面理解算法在特定任务中的表现。这样的多指标评估不仅能够提供更全面的性能视角，还能够更有效地指导算法的优化和改进。在医学领域，这些评估指标的综合运用尤为重要，因为它们能够为医生提供多方面信息，有助于更准确地进行辅助诊断。

### 2.3.2 主观评估手段

主观质量评价是一种依赖于人类感知系统来评估影像融合效果优劣的测量手段。1974 年，国际电信联盟无线电通信部门（ITU-R）发布了技术规范 BT.500.1，该建议书为建立统一的影像质量主观评价实验设备配置、测试环境标准以及相应的评测方法提供了指导原则。随着技术的不断发展，这一技术建议已经发展到第 13 版 BT.500.13[137]。主观评价测试依赖众多参与者对显示技术进行评估，同时对测试条件如屏幕、观看距离和光照环境有严格规定，导致这类测试成本较高，不仅包括人员费用，还涵盖了硬件和软件支出，这限制了其在广泛的视觉通讯系统中的应用。

表 2.3: 影像融合效果的主观视觉评分参考 [138, 139]

质量尺度	妨碍尺度	分值
非常差	无法清楚观察影像	1 分
差	影响观看	2 分
一般	影像质量变化明显，影响观看程度比 较轻微	3 分
好	影像质量存在变化但不影响观看	4 分
非常好	几乎看不出影像质量变换	5 分

鉴于影像的视觉品质最终需依赖人类感知来判定，因此在进行定性分析时，主观质量评价方法具有不可忽视的重要性。然而，主观评价存在主观性强、个体差异大等问题，为减少人眼主观判断可能带来的不确定性，对目标影像的定

量分析变得至关重要，即还需客观质量评价。当前，主观的质量评价技术常被应用来构建影像及其相应的主观质量评分数据库，然而，在医学影像融合领域，针对这一领域的主观评价数据库仍处于缺失状态。与客观评价指标不同，影像的主观评价缺乏量化标准，主要依据影像的视觉效果以及对源影像显著信息的保留程度进行分析与评价。

主观评价的方法是通过相关领域的专家对融合影像进行评估，评估可分为五个等级，表2.3显示了国际上规定的五个等级的质量尺度和妨碍尺度评价[138, 139]。主观评价方法通过观察者对融合结果进行观察，符合人体的视觉感知。然而，由于观察者对融合结果的观察重点存在差异，对同一融合算法的评估也可能有所不同。因此，为了更全面有效地评估融合影像，目前的影像融合评估中，需要结合客观的质量评价方法进行综合评价。

## 2.4 本章小结

在这一章节中，本章以满足医学领域辅助诊断的迫切需求为切入点，深入探讨了七种医学影像技术，包括CT、MRI、X射线、SPECT、PET、显微成像和超声成像。通过对这些成像技术的机制、各自的优缺点以及在不同疾病方面的应用进行仔细分析，以便更好地应对医学影像领域的复杂性和多样性。在考虑医学影像处理的多任务需求的同时，本章分析了基于传统机器学习方法(SVM)和基于新兴的深度学习方法(CNN)。医学领域对影像技术的需求不断增长，这两种类型的分类模型都在不同的场景中展现了强大的潜力。传统方法在数据较为有限或需要解释性强的情况下具备优势，而深度学习方法则在处理大规模数据和复杂特征提取方面取得了显著的成就。因此，这两者的结合和并行应用成为医学影像分类领域的一个前沿趋势，为更准确、高效的医学诊断提供了新的可能性。

另外，本章进一步深入研究了不同层面的融合策略，包括基于像素级、特征级和决策级的信息融合方法。这些策略的介绍有利于更好地理解多模态医学影像融合与分类任务的实际应用场景，同时为未来算法设计提供了启示和指导。针对多模态影像融合与分类算法的评估，本章引入了丰富的质量和性能评估方法。这包括主观与客观的评价手段，特别强调对客观评估指标的详尽分析，以更准确地理解算法在不同场景下的性能表现。通过综合考虑多个评估指标，本章有助于深入了解算法的优劣，并为在实际医学应用中提高其有效性提供更全

面的评估。

## 第3章 多模态脑影像的智能辅助融合

不同疾病依赖的影像检查类别不同，不同的影像在揭示不同脑部病变的特异性表现上具有不可替代的价值。本文的工作聚焦于脑部疾病这一多元且复杂的诊断领域，鉴于其多样化的病理类型与诊断挑战，单一影像技术往往不足以实现精确诊断。因此，核心探讨的是如何有效地整合运用多种影像模态，旨在最大程度地提升脑疾病诊断的精确度与完整性。在这一章中，提出了两种针对多模态医学影像融合算法。一种是医学语义引导的脑影像融合双分支网络（Msg-Fusion）[140]，另外一种是多维特征自适应线性融合网络（MdAFuse）[141]，分别于3.1和3.2小节详细介绍。

### 3.1 基于脑影像的临床语义引导融合

#### 3.1.1 引言

自20世纪90年代以来，影像融合技术在医学领域得到了发展和应用。然而，自然图像和医学影像之间存在许多差异，例如信噪比、分辨率、相关区域大小和影像场景[142–145]。直接将传统的自然影像网络应用于医学影像往往效率低下，因为这会导致性能下降。因此，本节认为有必要深入分析医学影像的特点，设计一个专用的网络模型。医学影像融合可以方便医生对病例进行观察、评估和对病变进行分析，从而做出更准确的诊断。一般来说，CT/MRI/SPECT/PET常被医生用来观察脑部疾病。不同的医学影像模态具有独特的医学语义信息（medical semantic information, MS-Info）。不同模态的融合结果应该保持并增强相关MS-Info在每个模态中对诊断有意义。因此，本节提出了一个医学语义引导的双分支网络。在特征提取阶段，根据不同模态的MS-Info引导影像特征，采用有利于提取相应影像特征的网络分支策略，保证融合结果中各模态的MS-Info得到保持和增强。影像融合技术将两幅源影像中的这些信息融合并增强到一幅影像中，使医生能够更方便、更准确地观察、理解和诊断疾病。因此，脑CT/MRI/SPECT/PET影像的融合技术对帮助医生诊断脑部疾病具有重要意义。

近年来，许多医学影像融合方法被提出。这些现有的方法主要包括两类，即，传统的人工融合方法[146–151]和基于深度学习的融合方法[152–160]。传统的

方法是由人工认知驱动的。传统的特征提取方法主要依靠人工提取，需要专业的领域知识和复杂的参数调整过程。此外，每种方法都针对特定的应用场景。基于深度学习的方法是数据驱动的。它们可以通过学习大量的样本来获得深度抽象的特征。数据集的表达更加高效和准确，提取的抽象特征展现出较高的稳健性和广泛的适用性。目前，DL 方法已成功应用于影像处理的许多领域。然而，基于 DL 的医学影像分析仍处于早期发展阶段。相关的研究正成为热点，具有广泛的应用前景。本文提出了一种新的 CT/MRI/SPECT/PET 脑影像融合方法，目的是保留和增强原始影像中重要的 MS-Info 信息，以辅助医生诊断脑部疾病。本节工作作出了以下贡献：

- 这是第一个关注多模态医学影像 MS-Info 的影像融合工作，并将其映射到相应的影像特征上。设计了一种医学语义引导的双分支网络，以有效地学习与不同模态（MRI/PET/SPECT/CT）的 MS-Info 对应的深度特征。
- 在 SF-Branch 中，本节提出了一种空间域和频域相结合的特征提取方案，在保留原始影像信息的基础上，更方便地从 MRI 中提取关键 MS-Info 的相应特征，这在医学影像融合的智能辅助设计中尚属首次。
- 在 GV-Branch 中，本节提出了一种结合灰度空间和改进的 HSV 颜色空间中的亮度分量信息的特征提取方案，提取对应于 PET/SPECT 影像的关键 MS-Info 的基础上，保留原始影像信息的深层次特征。
- 本节提出了一个新的临床评估，其是来自医生的评价。他们评估融合影像的质量取决于有多少 MS-Info 从源影像中被保留和增强。同时，本研究采用了七种常见的客观评价指标。实验结果表明，在所有评价指标结果中，所提出的影像融合方法均优于 9 种有代表性的融合方法。

### 3.1.2 相关工作

现有的融合技术涵盖了传统的和以深度学习为基础的各类方法。在这一部分中，本节主要介绍了医学影像融合的相关工作，影像处理中的傅里叶变换和 HSV 颜色空间变换的理论。

#### 3.1.2.1 医学影像融合

医学影像具有自身的特殊性，医学影像融合在影像引导的医学诊断、治疗等计算机视觉任务中发挥着重要作用。其中，多模态影像融合模块也被设计用

于疾病辅助诊断，如 [161]。Algarni 提出了一种基于多模态影像融合的诊断系统，适用于融合 MRI 和 CT 影像 [162]，但仅适用于融合 MRI 和 CT 影像。为了开发一种即使受损影像也能准确保留细节信息的融合方法，Li 等人将低秩稀疏矩阵字典学习方法应用于医学影像融合 [163]，可以同时达到去噪和增强的效果。Panigrahy 等人提出了一种利用双通道 PCNN 的医学影像融合方法 (WPADCPCNN) [164]，该方法可以获得良好的融合效果，但仅适用于 MRI-SPECT 融合。Parvathy 提出了一种基于优化阈值和深度学习的融合模型 [165]，可以为专家提供解剖和生理数据，以促进诊断过程。针对彩色空间域的伪彩色影像，提出了一种基于 Otsu 自适应阈值 (atsIF) 的双尺度影像融合方法 [149]。

Hermessi 等人提出了一种基于卷积神经网络相似性学习的多模态 MRI 和 CT 影像融合方法 [166]。Das 等人提出了一种新的基于低秩纹理先验分解和超像素分割的影像融合 [167]，该方法结合了三种方案，即，采用灰度优化、优化的低秩纹理先验知识和像素相关的高斯混合模型，有效地提高了融合影像的视觉保真度。Kumar 等人提出了一种新的 CNN 方法，专门用于 MRI 和 PET 影像融合 [155]，并在训练过程中使用结构相似性指数 (SSIM) 作为损失函数。Liang 等人提出了多层级联融合网络 (MCFNet) [168]。该网络补充了融合医学影像在两次下采样过程中丢失的空间信息。Yu 等人提出了网络 (IFCNN) [157]，Li 等人提出了通用网络 (NestFuse) [158]。这两种方法均采用卷积神经网络进行特征提取，然后采用特征值最大化策略进行多模态医学影像融合。它们可以获得很好的融合效果，但由于卷积网络不能捕捉长距离的相关性，无法避免全局纹理信息的丢失。

目前，大多数医学影像融合方法都有局限性。首先，MRI/CT/PET/SPECT 是常用的脑成像技术，然而，在大多数医学影像融合方法中仅考虑两种模态。其次，大多数方法主要遵循自然图像融合的思想。事实上，医学影像与自然图像之间存在差异，不同模态的医学影像之间也存在很大差异。没有充分考虑不同模式的关键 MS-Info。第三，现有方法的融合策略采用加权求和策略和最大值策略。权重系数的选取需要经过多次试验，由结果来确定。求和策略会导致融合影像模糊，而最大值策略会丢失一些关键特征。这种简单、单一的融合策略会导致源影像中关键 MS-Info 的丢失。第四，医学影像融合的最终目的是方便临床医生阅读，融合影像的质量会影响医生对疾病的诊断。因此，临床医生的质量评估应当是金标准。现有的医学影像融合方法很少考虑临床医生的评价。最主要的一点是没有充分考虑源影像的 MS-Info。因此，本文构建了一种新的融

合网络，重点是脑医学影像的 MS-Info 和临床评价的融合目标。

### 3.1.2.2 频域变换在影像处理中的应用

在图像处理的过程中，要将图像转换到频域进行处理，常用的是傅立叶变换。在图像处理领域，可以利用傅立叶变换来获得空间图像的频率分布，然后在频域进行各种图像处理，可以有目的地实现许多功能。近年来，傅立叶变换在图像编码 [169]、图像检测 [170]、图像压缩 [171]、图像分析 [172, 173]、图像配准 [174] 和图像重建 [175] 中有着广泛的应用。傅里叶变换在图像融合中具有潜在的优势。Naidu 等人 [176] 提出了一种基于快速傅立叶变换 (MFFT) 的多分辨率图像像素级融合算法。然而，这种方法是基于多分辨率自然图像的融合。据广泛阅读文献发现，目前还没有利用傅立叶变换进行医学影像融合的案例。此外，频域分析方法在基于深度学习的方法中也具有不可否认的潜力。例如，Kai 等人 [177] 提出了一种基于频率的学习方法，证明了频域学习方法在分类、检测和分割任务中的普遍性和优越性。在频域对影像进行处理，可有效地保存影像信息，提高精度。因此，本节也采用频域处理对医学影像进行处理，并联合深度学习方法，以达到更好的融合效果。

在数字图像处理技术中，傅立叶变换占据着至关重要的地位，它可以将图像从空间域变换到频率域。从本质上讲，变换后的图像与原始图像相同，并且可以更直观地分析图像的幅度和相位分量，即图像的全局和局部信息。这些信息对于 MRI 的全局纹理和局部几何形状都是非常重要的信息，代表了软组织边缘和内部结构的语义特征。利用傅立叶变换对 MRI 进行处理，对融合后的病灶定位起着关键作用。

### 3.1.2.3 彩色空间变换在影像处理中的应用

RGB、YUV 和 HSV 颜色空间通常用于影像处理。RGB 颜色空间是最常用的影像颜色表示空间。YUV 易于压缩，便于传输和处理。HSV 颜色空间最适合人类的视觉感知，因为在该空间中的颜色变化很容易被人类区分。H (Hue) 表示色调，指影像的颜色偏好，S (Saturation) 表示颜色的强度或纯度。V (Value) 表示颜色的亮度。RGB 和 HSV 转换通常用于影像处理以增强影像 [178]。(R, G, B) 分别是颜色的红色、绿色和蓝色坐标，其值是 0 和 1 之间的真实的数。如果 M 等于 R、G 和 B 中的最大值，并且 m 等于这些值中的最小值，则 RGB 和 HSV 之间的转换过程如等式 (3.1) 所示：

$$\begin{aligned}
 H &= \begin{cases} 0^\circ & \text{if } M = m \\ 60^\circ \times \frac{G-B}{M-m} + 0^\circ & \text{if } M = R \text{ and } G \geq B \\ 60^\circ \times \frac{G-B}{M-m} + 360^\circ & \text{if } M = R \text{ and } G < B \\ 60^\circ \times \frac{B-R}{M-m} + 120^\circ & \text{if } M = G \\ 60^\circ \times \frac{R-G}{M-m} + 240^\circ & \text{if } M = B \end{cases} . \\
 S &= \begin{cases} 0 & \text{if } M = 0 \\ \frac{M-m}{M} = 1 - \frac{m}{M} & \text{otherwise} \end{cases} \\
 V &= M
 \end{aligned} \tag{3.1}$$

HSV 颜色空间相较于 RGB 颜色空间，在模拟人类视觉感知上更为贴近 [179]，HSV 颜色空间中多个通道的使用可以单独处理，并且彼此独立。所以，在 HSV 色彩模式下，能够显著简化影像分析与处理的复杂度。它也用于影像融合 [180]。HSV 颜色空间在影像处理中得到了广泛的应用，通常使用不同的通道来解决不同的问题。与 RGB 空间相比，HSV 空间可以非常直观地表达颜色、色调和明暗度，便于颜色之间的对比和情感的交流。因此，本文还将 HSV 颜色空间的优势应用到医学影像融合中。为了更好地提取有用的 MS-Info，本节采用了一个自定义的  $V$  分量，这将在下一节中详细介绍。

Modality	Key MS-Info	Image Feature	Extraction Strategy	Branch
MR	Clear shape of soft tissue	Boundary zone	High frequency band in Frequency Domain	SF-branch
	Clear internal structure of soft tissue	Internal texture detail	Low frequency band in Frequency Domain	
	NIL	More source image information	Spatial Domain	
CT	High-resolution global anatomical structure of hard tissue	Global contour lines; Local shape location of tiny area; More source image information	Multi-scale; Concatenate; Gray color space	GV-branch
	High-resolution local anatomical structure of hard tissue			
PET	Obvious display of early small lesion	Brightness	Brightness component of HSV color space	
	High-distinguished functional-metabolic abnormality tissue			

图 3.1: 多模态医学影像的关键医学语义信息引导两个特征提取分支的设计

### 3.1.3 医学语义引导的双分支网络

不同模态的医学影像中包含着不同的医学语义信息 (medical semantic information, MS-Info), 这对临床疾病诊断具有重要意义。因此, 应该保持和加强各自在其多模态医学影像融合结果的 MS-Info。为了实现这一目标, 本节首先根据临床医学和成像理论分析每种模态的 MS-Info。然后, 映射关键 MS-Info 到影像特征, 并为不同的功能设计有效的提取策略。MRI/CT/PET/SPECT 的关键医学语义信息如何指导 MsgFusion 的两个分支设计的细节在表3.1中给出, 其引导构建所提出的医学语义引导的双分支融合网络 (Medical Semantic Guided Fusion, MsgFusion) 的两个分支。

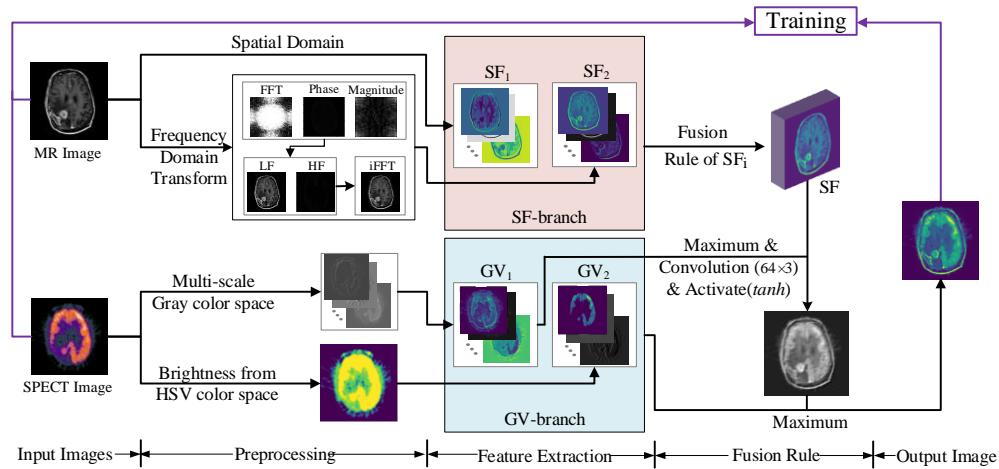


图 3.2: MsgFusion 方法的框架: 医学语义引导的多模态脑影像融合双分支网络

MsgFusion 的整体框架如图3.2所示, 其中预处理和特征提取是关键步骤。首先, 在深度 MS-Info 提取阶段, 网络结合了两个特征提取分支, 空间域和频率域结合的分支 (the branch combines spatial domain and frequency domain, SF-Branch) 和灰度与自定义 V 分量结合的分支 (the branch combines gray color space and self-defined brightness information, GV-Branch), 如图3.2所示。SF-Branch 采用傅立叶变换, GV-Branch 采用 HSV 颜色空间, 不仅充分利用了影像的空间和频率关系, 而且从影像的重要信息中提取了丰富的语义特征。第二, 是本节的分层融合策略。这两个特征提取分支的结合不仅提高了算法的性能, 而且有效地获得了多模态医学脑影像的重要的 MS-Info。

### 3.1.3.1 空间域和频率域结合的方案

SF-Branch 被设计用于从 MRI 中提取深度特征。如表3.1所示, MRI 的 MS-Info, 即, 软组织的清晰形状和内部结构在频域中更容易区分为高频带信息和低频带信息。为了有效地提取 MS-Info 对应的深层特征和更多的源影像信息, 本节采用了频域和空间域相结合的策略。图3.3示出了 SF-Branch 的过程。

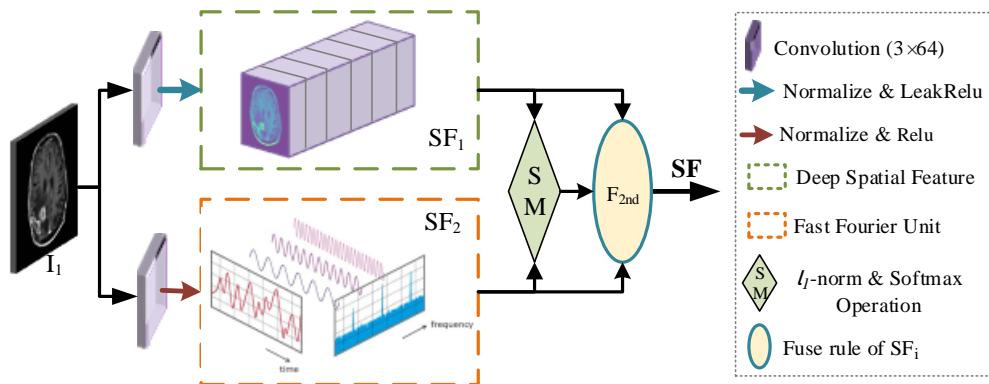


图 3.3: SF-Branch 提取特征的过程

特征提取的 SF-Branch 包括两部分, 一部分是从神经网络的卷积特性中获取源 MRI 的深层特征; 另一部分是从频域中获取 MS-Info 对应的特征。在第一部分中, 通道被放大到 64, 卷积核的大小设置为  $7 \times 7$ , 步幅设置为 1, 填充量为 3。批量归一化并使用 LeakyReLU (alpha=0.2, inplace=true) 激活卷积特征。该部分的输出记录为  $SF_1$ , 如图3.3所示。另一部分先采用频域处理, 其输出记为  $SF_2$ 。频域处理的示例显示在由图3.2中用频域变换标记的箭头所指向的正方形中。具体地, 一幅  $M \times N$  影像的二维离散傅立叶变换和逆变换表示为等式 (3.2) 和式 (3.3)。其中,  $x$  和  $y$  为空间域中的影像变量,  $f(x, y)$  表示点  $(x, y)$  处的灰度值,  $u$  和  $v$  为频率变量, 当  $u$  和  $v$  为 0 时, 为原点处的傅立叶变换, 相当于一幅影像的平均灰度值。

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M+vy/N)} \quad (3.2)$$

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M+vy/N)} \quad (3.3)$$

假设  $Re$  和  $Im$  分别用于表示  $F$  的实部的和虚部，则它们的计算是根据等式 (3.4) 和等式 (3.5)：

$$\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \left(2\pi \left(\frac{wx}{N} + \frac{vy}{N}\right)\right) \quad (3.4)$$

$$\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \sin \left(2\pi \left(\frac{wx}{N} + \frac{vy}{N}\right)\right) \quad (3.5)$$

影像经过傅立叶变换后，每个像素都是一个包含实部的和虚部的复数。通过计算每个像素的复数的幅值和相位，可以得到影像的幅值和相位。傅立叶频谱、相位角和振幅被定义如下：

$$P(u, v) = |F(u, v)|^2 = Re(u, v)^2 + Im(u, v)^2, \quad (3.6)$$

$$\phi(u, v) = \arctan \left[ \frac{Im(u, v)}{Re(u, v)} \right], \quad (3.7)$$

$$|F(u, v)| = [Re(u, v)^2 + Im(u, v)^2]^{\frac{1}{2}}. \quad (3.8)$$

这是首次使用傅立叶变换进行医学影像融合。傅里叶变换的作用是把影像的空间域灰度级分布转换成影像的频率域表示，逆傅里叶变换是将影像的频率表示变換回灰度分布级函数。通过傅立叶变换可以得到磁共振影像的幅值和相位。影像的幅度包含影像的全局信息，而相位包含影像的局部信息。利用傅里叶变换的性质，该方法能够取得较好的融合效果，并通过实际实验验证了该方案的切实有效性。

当完成上述步骤时，接下来本节需要在空间域和频率域之间融合特征。本文提出的加权图特征是融合多尺度深度特征来获得空间特征的详细结构。权重映射通过  $l_1$  范数和 softmax 运算通过目标函数完成，如下式 (3.9) 所示：

$$\xi_k(x, y) = \frac{\|\psi_k(x, y)\|_1}{\sum_{i=1}^2 \|\psi_i(x, y)\|_1}, \quad (3.9)$$

其中， $\|\cdot\|_1$  表示  $l_1$ -norm， $k \in 1, 2$ 。 $(x, y)$ ，并且在深度特征中，每个位置对应一个维度为  $dim$  的向量表示。 $\psi$  表示具有  $dim$  维度的向量。最终的融合特征图  $\varphi$ ，是两个增强的特征图的叠加，其由式 (3.10) 表示：

$$\varphi = \sum_{i=1}^2 \xi_{k(x,y)}^{(i)} \times \psi_i. \quad (3.10)$$

### 3.1.3.2 灰度与 V 颜色分量结合的方案

GV-Branch 设计用于从 CT/PET/SPECT 提取深层特征。一方面，GV-Branch 采用灰度空间的多尺度级联策略，从源影像中提取全局轮廓和局部形状特征，补偿不同尺度下的信息损失。另一方面，根据表3.1中的分析，PET/SPECT（高分辨功能代谢异常组织）的关键 MS-Info 主要反映在影像的亮度水平上。为了获得亮度信息，本节使用 HSV 颜色空间变换来计算 V 分量，并将其改进为新的亮度值以突出功能影像（PET/SPECT）的 MS-Info。将 CT 的 MS-Info（硬组织的高分辨率全局和局部解剖结构）和 PET/SPECT 的部分 MS-Info（早期小病变的明显显示）引导到多尺度策略，在灰度颜色空间中级联。为了从不同的尺度和层次捕获信息，以较少的信息损失，采用了多尺度和跳跃连接策略。CT 不包含功能代谢信息。当 CT 输入到 GV-Branch 时，仅需要灰度颜色空间中的深度特征。总的来说，为了有效地提取对应于 MS-Info 和更多源影像信息的深度特征，本节采用了 HSV 颜色空间和灰度颜色空间相结合的策略。

1) 通过多尺度策略获得深度卷积特征：由于不同层次和分辨率的卷积运算可以提取不同重要程度的信息，因此提出了一种针对脑影像的功能组织特征提取方法。该方法的功能组织特征分支网络如图3.4所示。网络块主要包括卷积层、池化层、上采样和跳跃连接操作。图3.4给出了获取 MS-Info 的具体步骤，不同的色块表示不同数量的输入和输出通道，其对应于不同厚度。

多次使用拼接可以扩展高维特征，更准确地发现更重要的语义特征。当多条分支合并时，采用最大值法突出高频纹理信息。由于各种脑部疾病是不均匀的，使用这种特征提取策略可以使病变程度和正常组织之间的分类更加明显。然后，可以更容易地确定 PET 影像中的局部区域是否异常，并在 CT 影像中观察到清晰的轮廓和全局结构信息。并且，使用长短跳来增强图3.4中带有紫色虚线的特征的转移，可以充分融合不同层次的视觉特征，减少特征转移过程中的特征损失。这个想法受到 DenseNet[181] 的启发，DenseNet 主要由许多密集的网络块实现。为了解决消失梯度问题，网络只级联卷积操作的前后层的权重系数，而不级联池化操作。在本节的网络中，将前一层和前一层的特征输入到下一层，连接模式表示为式 (3.11)：

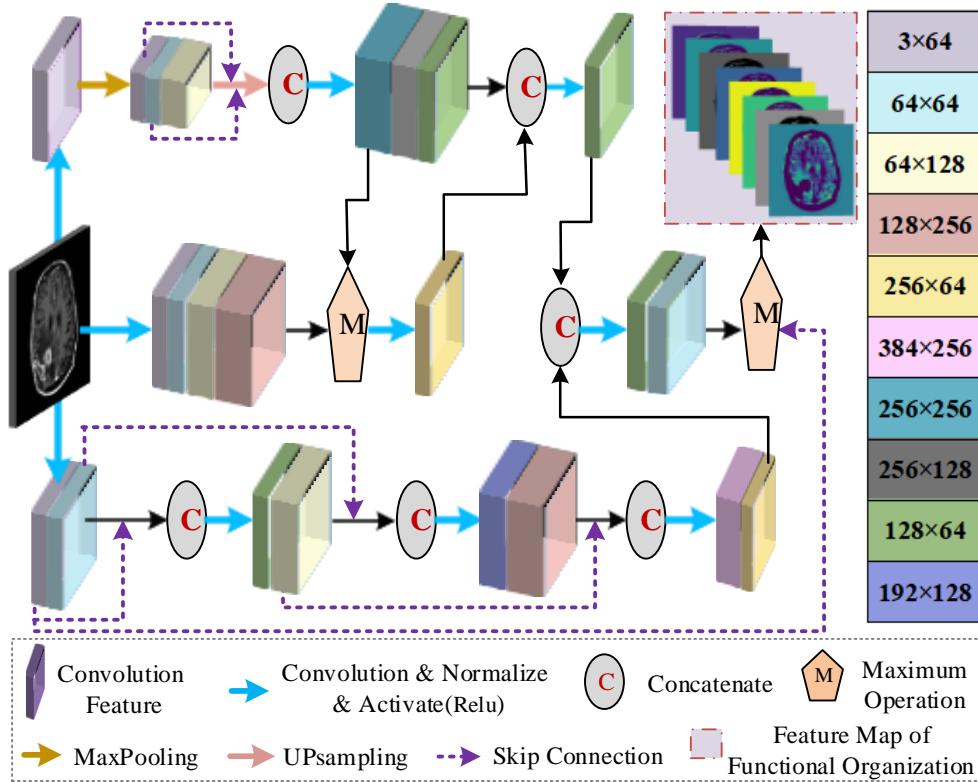


图 3.4: GV-Branch 提取特征的过程一

$$f_d = H_d([f_0, f_1, \dots, f_{d-1}]), \quad (3.11)$$

其中,  $f$  表示输出,  $d$  表示网络中的层数,  $H_d(*)$  表示非线性传递函数的组合, 包括级联、BN、激活和卷积操作。多层特征 (不一定是相邻层) 的并行输入由组合函数操作以获得新特征。新特征的通道数由不同的卷积参数决定。详细步骤如图3.5所示。此操作有助于捕获跨区域的长期和多层次依赖关系, 并减少特征转移过程中的损失。因此, 可以保留更完整的脑影像结构。

2) 通过颜色空间变换获取亮度信息: HSV 颜色模型, 基于对色彩的直觉理解构建而成, 由色调 (Hue)、饱和度 (Saturation) 和亮度 (Value) 三个参数组成, 并以其几何表示而著称于六角锥体模型。HSV 颜色空间可以单独处理, 并且彼此独立。因此, 通过采用 HSV 颜色模型进行影像分析与处理时能够显著降低复杂度, 并且相比 RGB 色彩空间, HSV 更贴合人类视觉系统的自然感知方式。在此特征提取部分, 本节计算功能影像的 HSV 空间的 V 分量, 以提取其

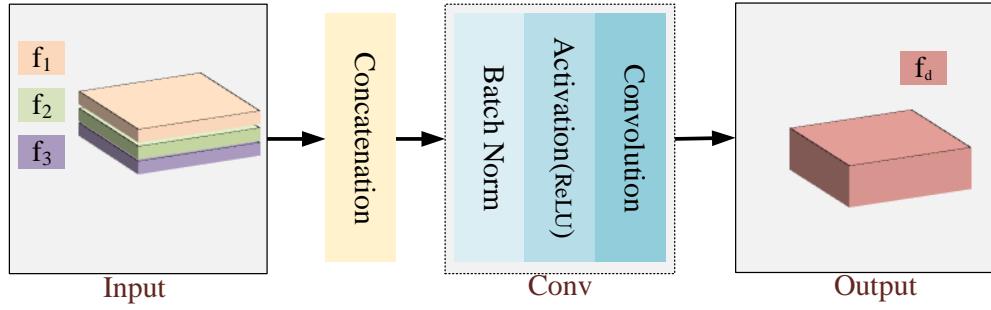


图 3.5: 特征跳跃策略

亮度特征用于后续处理,如图3.6所示。在该图中,  $I_p$  表示原始 RGBA 影像,  $I_A$  表示 RGB 影像,  $V$  表示将  $I_p$  变换到 HSV 空间之后的亮度分量。为了进一步增强亮度明显区域的局部信息,本节定义了一个新的亮度值  $V'$ , 如算法1所示。

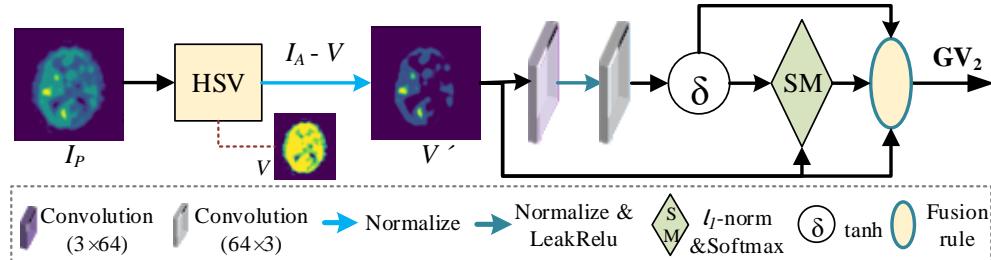


图 3.6: GV-Branch 提取特征的过程二

如图3.6所示,  $I_p$  是 PET 影像的一部分。图中可以发现三个不同的色块(黄色, 浅绿色和深蓝色), 它们表明能量高或新陈代谢旺盛的区域。病变最有可能存在于这些区域。这些颜色块可以从  $I_p$  的前景中辨别出来, 主要是因为它们的亮度信息。然而, 从亮度影像  $V$  中可以看出, 这些有意义的区域还不够清晰, 无法准确地定位病变。在算法 1 中, 本节通过  $I_A - V$  计算差分影像, 以进一步突出这些色块。然后, 病变可以清楚地定位在  $V'$  影像的右侧。基于新的亮度影像, 通过后续过程提取 HSV 颜色空间中的深度特征  $GV_2$ , 如图3.6所示。

### 3.1.3.3 双分支的分层融合与重构

为了充分利用从两个分支中提取的深层 MS-Info 特征, 本节采用了基于分类和分层的融合策略, 如图3.2的融合规则阶段所示。在第一层融合解剖结构特征 (SF 和  $GV_1$ ), 然后在第二层融合功能代谢特征 ( $GV_2$ ), 有利于保留和增强

---

**算法 1** New Luminance Computation

---

**输入:** RGBA image  $I_p : (R \ G \ B \ A)$   
**输出:** New luminance value  $V'$

```

1: newLum ( $I_p$ ):
2:    $I_A \leftarrow \text{Compress} (I_p);$ 
3:    $I_{HSV} : (H \ S \ V) \leftarrow \text{Conver} (I_p);$ 
4:    $V' \leftarrow I_A - V;$ 
5:   return  $V';$ 
6: Compress ( $I_p$ ):
7:    $I_A \leftarrow I_p : (R) + I_p : (G) + I_p : (B);$ 
8:   return  $I_A;$ 
9: Conver ( $I_p$ ):
10:   $I_A \leftarrow \text{Compress} (I_p);$ 
11:   $I_{HSV} \xleftarrow{\text{Eq. (3.1)}} I_A;$ 
12:  return  $I_{HSV} : (H \ S \ V);$ 

```

---

待融合医学影像的关键 MS-Info。本节的融合策略不是一个简单的加权和。为了保持和增强多模态影像的重要纹理特征，使用最大化操作。同时，使用卷积和激活函数来获得局部细节结构特征。如图3.2所示，采用  $\tanh$  作为激活函数。

### 3.1.3.4 损失函数构建及参数设置

由于医学数据的保密性，大量的多模态医学数据不易获取。通常，许多基于 ImageNet 数据训练的网络模型，如 ResNet[182] 和 VGG[183]，被选择来生成预训练参数。为了在解决梯度破碎问题和收敛速度方面获得更好的性能，本节选择 ResNet101 来迁移 ImageNet 上的最后一个卷积特征，也就是说，使用在 ImageNet 数据上预训练的 ResNet101 的最后一层作为本节的第一个卷积层来提取有效的影像特征。由于预训练模型是针对分类任务设计的，因此在本节的网络结构中，第一层的参数会根据本节的任务进行调整，以达到更好的融合效果。为了获得源影像的更精确的重建，本节联合结构和像素信息来构造损失函数如下式 (3.12):

$$L = \omega L_S + (1 - \omega) L_P \quad (3.12)$$

总损失函数由结构相似度  $L_S$  和像素损失  $L_P$  组成， $\omega$  为调整系数，其值由训练和测试效果根据不同取值而定。 $I_f$  表示融合的结果， $I_i, i = 1, 2$  表示输入的源影像。SSIM 表示两个影像的结构相似性 [44]。对于这两种类型的影像，

计算结构相似度并将其组合。假设两个影像的结构相似度为  $F_{SSIM_i}$ ,  $F_{SSIM_i} = 1 - SSIM(I_f, I_i), i = 1, 2$ , 总结构相似度函数计算如下:

$$L_S = \sum_{i=1}^2 \alpha_i \left( 1 - \frac{(2\mu_{I_f}\mu_{I_i} + c_1)(2\sigma_{I_f}\sigma_{I_i} + c_2)}{(\mu_{I_f}^2 + \mu_{I_i}^2 + c_1)(\sigma_{I_f}^2 + \sigma_{I_i}^2 + c_2)} \right) \quad (3.13)$$

其中,  $\mu_{I_f}$  是  $I_f$  的平均值,  $\mu_{I_1}$  和  $\mu_{I_2}$  分别是  $I_1$  和  $I_2$  的平均值。 $\mu_{I_f}^2$  是  $I_f$  的方差,  $\mu_{I_1}^2$  和  $\mu_{I_2}^2$  分别是  $I_1$  和  $I_2$  的方差。 $\sigma_{I_f I_1}$  和  $\sigma_{I_f I_2}$  是原始影像和融合影像之间的协方差,  $c_i$  是常数。像素损失是  $l_2$  范数。假设两幅影像的像素损失为  $F_{M_i}$ ,  $F_{M_i} = \|I_f - I_i\|_2$ , 则总像素损失函数计算如下:

$$L_P = \beta \sum_{j=1}^n (I_{f_j} - I_{1_j})^2 + (1 - \beta) \sum_{j=1}^n (I_{f_j} - I_{2_j})^2 \quad (3.14)$$

在式 (3.13) 和式 (3.14) 中,  $\alpha_i, i = 1, 2$ ,  $\beta$  是设置在区间  $[0, 1]$  中的权重参数。在本节中, 这三个参数都设置为 0.5。在方程式 (3.12) 中,  $\omega \in [0, 1]$  的取值由训练效果和测试效果决定。参数  $\omega$  决定了不同模态原始影像与融合结果影像之间相关性的程度。在其他参数不变的情况下, 通过观察不同  $\omega$  的融合影像的效果来确定它的取值, 如图3.7所示。

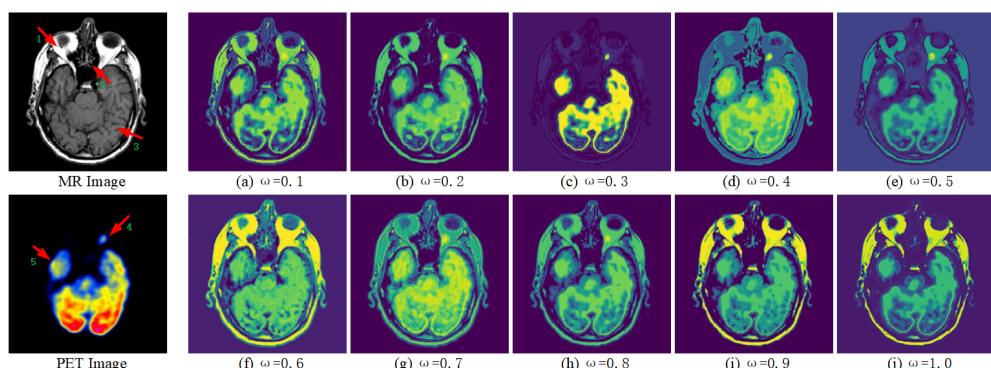


图 3.7: 不同的  $\omega$  值应用于 MRI-PET 影像融合的结果

本节通过实验来解释  $\omega$  的影响, 并为  $\omega$  设置了一组值 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0。在图3.7中, 第一列显示了 MRI 和 PET 影像, 其他影像是采用不同  $\omega$  的融合结果。从图中可以直观地观察到, 当  $\omega$  小于 0.5 时, 融合结果中的 MRI 信息损失更大, 尤其是当 0.3 最严重。当  $\omega$  大于 0.5 时, 融

表 3.1: 针对不同  $\omega$  取值下的 MRI-PET 影像融合的质量评估

$\omega$	EN	SD	VIF
0.1	3.3390	46.1341	0.2065
0.2	3.1554	44.4309	0.2087
0.3	3.2173	47.9444	0.0395
0.4	3.7952	<b>51.3496</b>	0.2017
0.5	4.0602	26.4541	0.0443
0.6	<b>4.5183</b>	46.2509	0.2243
0.7	<b>4.3093</b>	<b>51.8038</b>	<b>0.2801</b>
0.8	3.1933	45.3808	0.2378
0.9	3.1667	47.9978	<b>0.2871</b>
1.0	3.9747	35.0954	0.1738

合结果中 PET 的亮度信息或多或少会丢失或失真。只有当  $\omega$  为 0.7 时，PET 的亮度信息才能得到最好的保留，MRI 的边缘轮廓信息也很清晰。表3.1显示了图3.7中不同  $\omega$  下 MRI-PET 融合的三个指标。EN 表示信息熵，SD 表示标准差，VIF 是对人类视觉感知的评估。这三个值越大，融合影像质量越好。融合影像质量评估指标的计算结果与上述观察结果一致。当  $\omega$  为 0.7 时，三个指标值达到最佳平衡和优化状态，其中 SD 最好，EN 和 VIF 接近最佳值。

综上所述，当  $\omega$  取 0.7 时，得到的结果最好。因此，本节在实验中选择该值来测试融合效果。在训练过程中，初始学习率为 0.001，250 epoch，批量大小为 2，采用批量归一化，优化函数为 Adam。此外，本节实施了一种自适应损失调整方案，即每隔 50 个训练周期将学习率调整为初始值的 10%，因为在训练过程中使用较小的学习率来调整权重以获得更好的权重。本节的方法已经在 PyTorch 版本 1.5.0 及其相应的 CUDA 版本 10.1 下的 Python 3.7.9 中实现。

### 3.1.4 多模态脑影像融合的实验

为了进一步验证所提出融合技术的效能，实验在不同的大脑影像对（MRI-SPECT/MRI-CT/MRI-PET）上进行。从 ADNI(<http://adni.loni.usc.edu/data-samples/access-data/>) 中获得了总共 555 个 MRI-PET 影像对作为训练集，其大小为  $256 \times 256$ 。在测试 MRI-CT、MRI-SPECT 和 MRI-PET 影像对中使用了全脑图谱 (<http://www.med.harvard.edu/aanlib/>) 上的 30 对影像，包括脑血管疾病（中风）和肿瘤疾病（脑肿瘤）。所有实验都在 GeForce RTX 3090 Ti 12 Intel Core Interl (R) Xeon (R) CPU E5-2678 v3 2.50 GHz 64 GB RAM 设备

的 GNU/Linux x86 64 系统上进行。

对于每对影像的实验，本节使用所提出的 MsgFusion 和其他九种代表性方法 (LatLrr[148], IFCNN[157], NestFuse[158], atsIF[149], FusionDN[159], FusionGAN[156], FunFuseAn[155], WPADCPNN[164] 和 OLTPSPS[167]) 得到融合结果。并采用 EN[123]、SD[126]、MI[124]、rSFe[184]、SM[185]、VIF[128] 等 6 种常用的评价指标和一种相对新颖的指标  $R_Q^{(F/(AB))}$  (简称 RQ) [186] 来评价这 10 种方法的融合效果。RQ 通过计算分数阶微分来反映融合影像中保留了多少边缘信息，它取代了 [187] 中使用的噪声敏感的 Sobel 算子，文章中采用三个 Sigmoid 函数 ( $\tanh$ 、 $\arctan$ 、 $\text{Logistic}$ ) 来得到三个度量。考虑到它们共同的函数单调性，本节只选择了一个基于  $\text{Logistic}$  函数的度量应用于本节的实验中。由于不同指标的值相差很大，本节进行了适当的线性变换，图3.9和图3.11的左上角标出了这些变换，以方便在同一图中进行比较。

### 3.1.4.1 MRI-SPECT 影像对的融合

图3.8中的第一列显示源 MRI 和 SPECT 影像。MRI 清晰显示脑脊液和其他软组织的纹理细节。SPECT 上的不同颜色和亮度可以反映代谢信息。图3.8 (a) - (i) 示出了所考虑的十种融合方法的结果。图3.8 (j) 展示出了通过 MsgFusion 获得的融合结果。聚焦于由七个箭头指示的区域，可以看出，与由其他七种方法获得的融合影像中的对应区域相比，显示出更清晰的结构和纹理细节。结合临床信息，本节分析了 MRI 和 SPECT 的 MS-Info 是否很好地保留在图3.8 (j) 中蓝色矩形框内箭头指示的 7 个区域的融合结果中。

箭头 1 指向小脑结节区域，MRI 源影像上可见一个轮廓清晰的白色小三角形，右上角略有模糊。然而，在 SPECT 源影像中，没有明显的特征。在 LatLrr 的融合结果中，影像显得模糊，与周围组织无法区分。在 IFCNN、NestFuse 和 FusionDN 的融合结果中，它们都比 SPECT 保留了更多的 MRI 信息。右上角的增强效果不明显。FusionGan 和 FunFuseAn 结果中的小三角形扭曲。在 OLTPSPS 的融合结果中，亮度信息明显丢失。

箭头 2 为左小脑半球区域，主要来自 SPECT 影像，表现为边缘模糊的亮点，在 MRI 上看起来像一个小黑洞。在 LatLrr、IFCNN、NestFuse、atsIF、WPADCPNN 和 OLTPSPS 的融合结果中，不容易找到那个亮点。在 FusionGAN 和 FunFuseAn 的融合结果中，亮点增强，但略有模糊。在 FusionDN 的融合结果中，这一亮点很明显，但与周边机构融合，具体位置无法确定。MsgFusion 的融

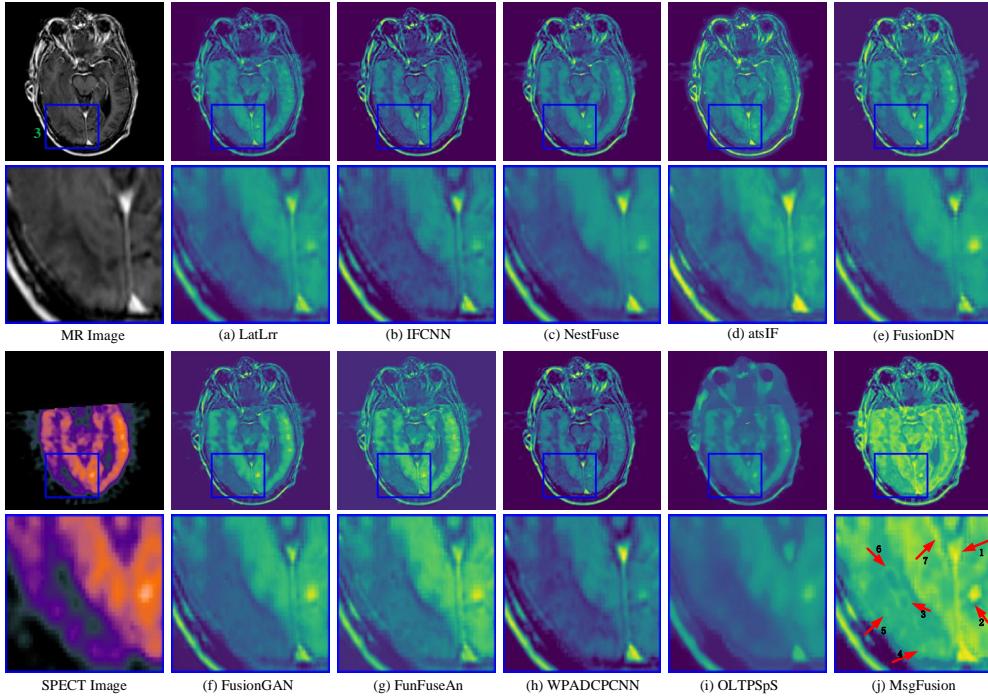


图 3.8: 用于脑 MRI-SPECT 影像融合的不同方法 (LatLrr、IFCNN、NestFuse、atsIF、FusionDN、FusionGAN、FunFuseAn、WPADCPCNN、OLTPSpS 和本节提出的方法) 性能比较

合结果中，有明显的亮度和轮廓，位置准确。

箭头 3 和 6 代表右小脑半球区域。在 MRI 中，它是具有灰度值，形状像一片树叶。在 SPECT 影像中，有三个相邻的灰色空环。从 SPECT 影像中可以发现，在箭头 6 指示的位置有一个清晰的灰色空环。在 NestFuse 和 atsIF 获得不均匀分布区域的融合影像中可以识别它，但不是很清晰，箭头 3 处的环消失不见。在 IFCNN、FusionDN、FusionGAN、FunFuseAn 和 WPADCPCNN 的融合结果中，它们在第 6 个 ROI 的环更清晰，而在第 3ROI 则不清晰。在本节提出的方法中，几个环可以很好地显示出来，并且很容易找到它们的位置。

箭头 4 和 5 分别是枕骨和右小脑半球的边缘区域。MRI 中的这个区域看起来像花边，但不容易检测到。它在 SPECT 中表现为一个不明显的亮度小点。相比之下，MsgFusion 保留了 MRI 的边缘，并且可以很容易地在 SPECT 中找到几个点。区域 7 是小脑结节附近的右鞍半球边缘。在由浅白色边缘信息表示的 MRI 中，在 SPECT 影像中似乎存在一条直线。对于所有方法的不同结果中，只有 MsgFusion 可以清楚地找到其边缘。

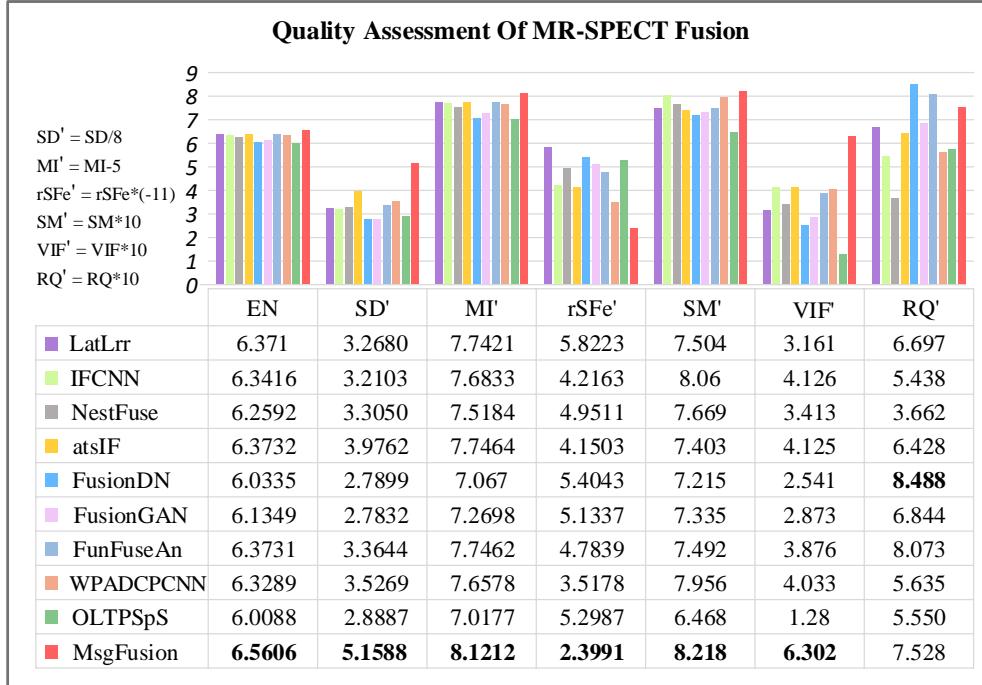


图 3.9: 计算不同方法对 MRI-SPECT 融合的质量评估指标值

图3.9显示了图3.8中融合结果的评估指标，不同的颜色柱条对应不同融合方法。EN、SD、MI、SM、VIF、RQ 分别表示融合的熵、标准差、互信息、结构相似度、视觉信息保真度和边缘信息，这些值越大，融合效果越好。rSFe 反映了改善的空间频率，其绝对值越小，融合效果越好。除 RQ 外，其他 6 个评价指标均表明 MsgFusion 方法达到最佳融合效果。因此，本节的融合结果具有更合适的亮度、更清晰的轮廓和更精细的纹理。此外，本节的结果保持并增强了重要的医学信息。当大脑出现异常时，可以通过观察 SPECT 的颜色和亮度信息，并结合 MRI 的纹理进行有效的分析，来确定可能的疾病类型。

### 3.1.4.2 MRI-CT 影像对的融合

图3.10中的第一列显示了原始 CT 和 MRI。很容易在 CT 影像上找到硬轮廓，如骨骼，在 MRI 上找到软组织结构。右边是不同融合方法的结果。图3.10 (j) 展示出了 MsgFusion 的融合结果和获得的局部放大影像，主要关注由箭头指向的六个 ROI。

箭头 1 指向的区域代表侧脑室的下角，在 CT 影像上为深灰色，在 MRI 上为明亮而清晰。MRI 中 ROI 1 的面积较大，具有重要的 MS-Info 信息，常被用

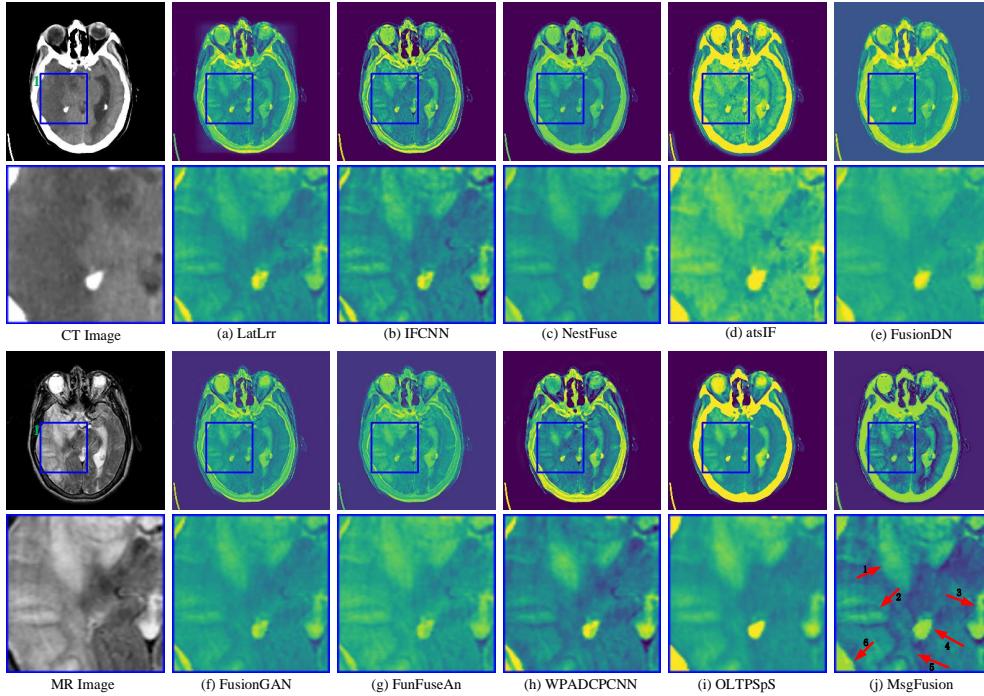


图 3.10: 用于脑 MRI-CT 影像融合的不同方法 (LatLrr、IFCNN、NestFuse、atsIF、FusionDN、FusionGAN、FunFuseAn、WPADCPCNN、OLTPSpS 和本小节提出的 MsgFusion) 性能比较

来判断是否存在病变。在 LatLrr、IFCNN、FusionDN、FunFuseAn 和 OLTPSpS 的融合结果中，轮廓不明显且不容易找到。在 atsIF 结果中，位置的亮度得到改善，但边缘出现了平滑现象。在 FusionGAN 和 WPADCPCNN 的融合结果中，位置的边缘分类明显，但亮度不够。在本文的融合结果中，可以很容易地找到区域及其边界。箭头 2 指向的区域呈岛状，主要反映在 MRI 上。这一区域并不像正常脑部影像中显示的那样完整，这表明组织失去了水分或萎缩。在图3.10中可以发现，只有 MsgFusion 融合结果中的区域明显，并显示出有别于其他组织的清晰轮廓。箭头 3 指向的区域在 CT 影像上为第四脑室，在 MRI 上也有脑桥和基底动脉的部分信息相连。atsIF 方法的亮度具有良好的亮暗影像对比度，然而，该区域的结构完整性被破坏。ROI 3 中的其他算法的结果清晰但不够明亮，而 WPADCPCNN 和 MsgFusion 更好。

箭头 4 指向的区域为侧脑室后角，主要表现在 CT 影像上。它在 MRI 上不是特别明显，但是，通过与周围组织仔细区分可以发现。对于 LatLrr、IFCNN、Nesthood 和 FusionGAN，可以找到 ROI 4 的轮廓，但其边界附近存在伪影，亮

度信息不足。对于 FusionDN，这个位置很明显，但周围组织丢失。FunFuseAn 融合结果表明，融合区域变得模糊。OLTPSpS 的结果有足够的亮度，但面积略小。在本节提出的方法中，ROI 4 不仅具有清晰的轮廓和明显的边界，而且具有足够的亮度。在本节提出的方法中，侧脑室后角的重要医学特征存在并得到增强。

箭头 5 指向的区域为中央沟，主要表现在 MRI 上。该位置的密度和亮度信息可以帮助医生确定是否存在异常。对于 MsgFusion 区域，它具有清晰的轮廓和亮度信息，很容易找到。箭头 6 指向的区域为顶骨，主要在 CT 影像上反映。当这个区域有边缘缺损或不完整时，就可以判断是否有脑损伤。通过以上分析，可以表明 MsgFusion 具有相对更好的融合效果。

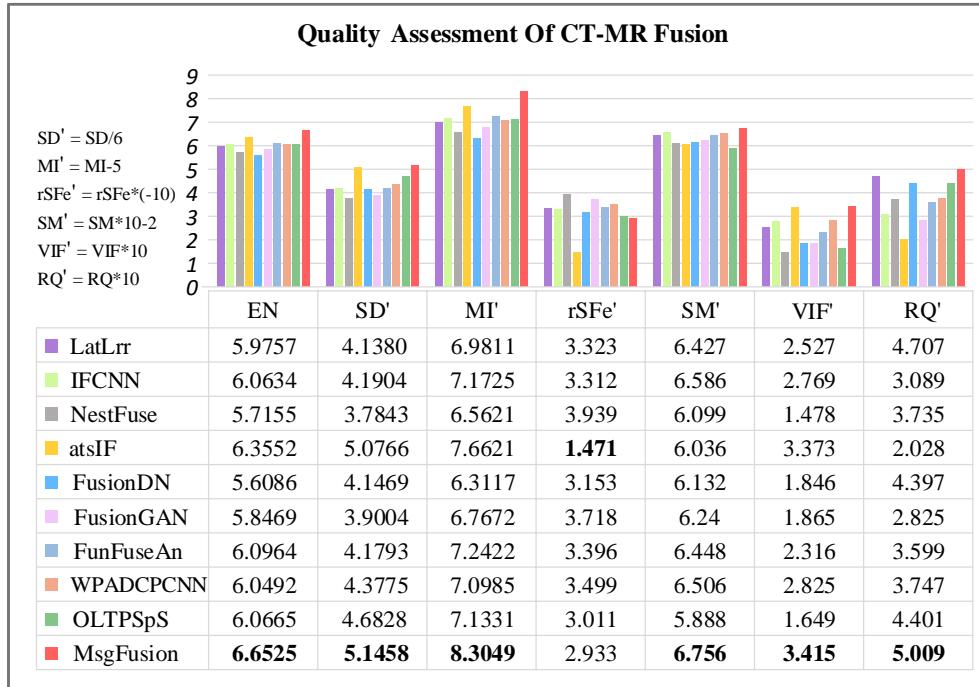


图 3.11: 计算不同方法对 MRI-CT 融合的质量评估指标值

此外，图3.11展示了对应于图3.10的质量评估指标值。结果表明，本节的 MsgFusion 方法在 6 个指标值上占优势，在 rSFe 上是次优的。这意味着原始 CT 和 MRI 的信息不仅保持良好，而且在本节的方法的融合结果中也得到了增强。当通过观察 MRI 密度和面积可以确定脑出血时，CT 影像可以发现脑出血的大致范围。

### 3.1.4.3 影像融合的问卷调查

对于医学影像融合，最终的目标是为医生提供易于观察的融合影像，以便对病变进行定性、定量和定位分析。因此，这种主观评价可以证实 MsgFusion 方法在临床意义上的优势。为了验证算法的临床有效性，本节对融合效果进行了在线问卷调查，并分发给了不同医院神经内科和医学影像科的 30 名医生。这些医生的临床经验：10 年以上 15 人，5-10 年 3 人，3-5 年 6 人，3 年以下 6 人。在这份问卷中，本节根据 15 组融合实验设计了 15 个问题。对于每个问题，将 6 种有代表性的方法（即 LatLrr、NestFuse、atsIF、FusionGAN、FunFuseAn 和 MsgFusion）的融合影像作为选项。不同方法得到的融合影像的顺序是随机排列的。每个受访者被允许选择一个或两个具有最佳融合结果的选项作为每个问题的答案。医生不需要确定融合后的影像保留了多少原始影像信息。他们只需要根据他们的临床经验来判断哪些融合结果更有利于他们的观察和临床诊断。最终，收到了 29 名参与者的有效答案。

表 3.2: 统计问卷调查表中临床医生选择各种方法的融合影像次数

	<b>LatLrr</b>	<b>NestFuse</b>	<b>atsIF</b>	<b>FusionGAN</b>	<b>FunFuseAn</b>	<b>MsgFusion</b>
$Q_1$	6	16	9	7	0	7
$Q_2$	12	10	7	8	1	5
$Q_3$	9	7	7	7	1	12
$Q_4$	5	9	5	0	6	17
$Q_5$	3	8	10	6	4	12
$Q_6$	3	10	2	4	4	21
$Q_7$	7	8	9	2	5	13
$Q_8$	4	6	8	10	4	12
$Q_9$	8	7	12	4	9	4
$Q_{10}$	4	1	5	15	7	11
$Q_{11}$	7	6	9	3	4	14
$Q_{12}$	4	5	15	2	6	10
$Q_{13}$	10	5	3	11	6	9
$Q_{14}$	6	6	10	7	1	13
$Q_{15}$	10	2	4	2	12	11
$\sum$	98	106	115	88	70	172

统计结果如表3.2所示，其中记录了从每种方法中为融合影像选择的次数。在用  $Q_i(i=1, \dots, 15)$  标记的每一列中，以下数字对应于医生为六种方法的融合影像选择的次数。在 15 组实验中，MsgFusion 产生的融合影像在 8 组中被最频繁地选为最佳融合影像，在 4 组中被第二频繁地选为最佳融合影像。从临床医

生的角度来看，MsgFusion 的融合效果远远超过其他任何考虑的方法。标记有  $\Sigma$  的最后一列列出了其融合影像被选择的每种融合方法的总次数。计算结果表明，本文方法比次优方法选择的次数分别高 26.5% 和 10.17%。

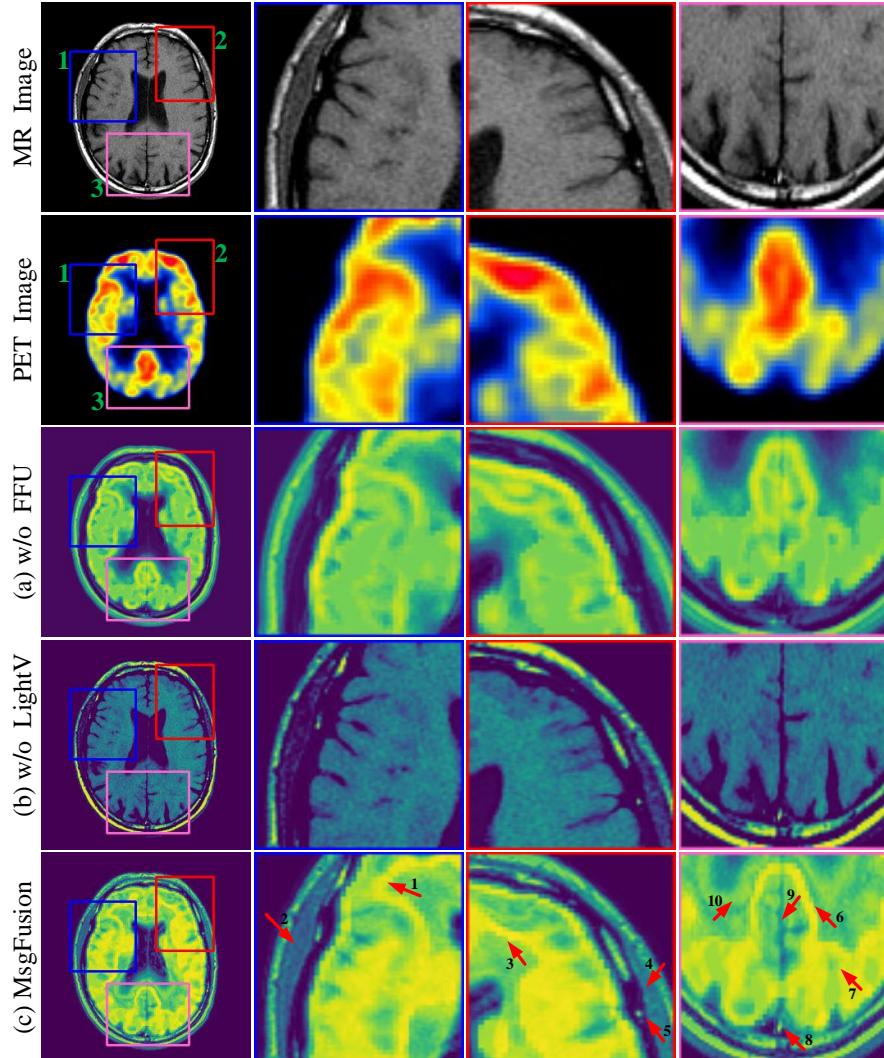


图 3.12：本节所提出的影像融合方法的消融实验

#### 3.1.4.4 影像融合的消融实验

为了说明在 SF-Branch 中结合频域和在 GV-Branch 中结合 HSV 颜色空间的必要性和有效性，本节对 MRI-PET 融合进行了消融实验。实验结果如图3.12所示。前两行显示 MRI 和 PET 的源影像及其各自的三个局部放大的

ROI。下面的三行分别显示的是：(a) 缺少频域处理，(b) 缺少来自 HSV 颜色空间的改进亮度分量，(c) 本节提出的 MsgFusion 方法。在最后一行局部放大影像中，标记了十个箭头，指向具有明显医学特征信息的区域。箭头 1 和 3 指向额叶，箭头 2 指向的是额骨，箭头 4 指向的是额骨内面，箭头 5 指向的是额骨和额叶之间的空间，顶叶区域由箭头 6、7 和 10 指示，箭头 8 指向的是上级矢状窦，并且箭头 9 指向大脑纵隔。当融合过程中不涉及 FFT 时，融合结果可以更完整地保留 PET 中的结构特征，但会丢失更多的 MRI 信息。当只使用 FFT 而不考虑 HSV 颜色空间的改进  $V'$  分量时，融合结果能较好地保留 MRI 的完整特征，但不能反映 PET 的功能信息。如图3.12的第五行所示，当同时考虑两者时，融合结果明显变得更好。

表 3.3: 图3.12中对应的三种融合策略的质量评估

<b>Whole image</b>	EN	SD	MI	rSFe	SM	VIF
w/o FFU	4.5027	56.8141	9.0054	-0.5018	0.2704	0.3219
w/o LightV	3.5768	36.7745	7.1537	-0.2977	0.3523	0.1262
MsgFusion	4.476	67.5634	8.952	-0.4103	0.3205	0.4076
<b>Average ROIs</b>	EN	SD	MI	rSFe	SM	VIF
w/o FFU	6.9201	52.0514	13.8401	-0.4204	0.5594	0.3227
w/o LightV	5.9916	34.0278	11.9832	-0.3802	0.6118	0.1738
MsgFusion	6.9598	54.9381	13.9195	-0.3467	0.5830	0.3652

表3.3示出了图3.12中的 6 个评价指标的相应计算结果。表中的红色值是最佳值，蓝色值是次佳值。计算了整幅融合影像的评价指标，结果表明，本文所提出的方法效果较好（最优方法有 2 个指标，次优方法有 4 个指标）。此外，本节还计算了图 10 所示的三个 ROI 区域的评价指标的平均值，如表3.3的接下来的四行所示。从三个 ROI 区域的评价指标的平均值也可以发现，本文的方法优于其他两种方法（缺少傅立叶应用和缺少  $V'$  分量计算）。

### 3.1.5 小结

在这一节中，介绍了一种 MS-Info 引导下的脑部疾病影像深度特征融合方法 MsgFusion。本节分析了 MRI/CT/PET/SPECT 的关键 MS-Info，以获取其对应的影像特征，从而找到最有效的提取策略。为此，设计了一个包括 SF-Branch 和 GV-Branch 的双分支网络。SF-Branch 结合了空间域和频域信息，GV-Branch 结合了 HSV 颜色空间的多尺度灰度影像和亮度。双网络机制成功

地提高了 CNN 的泛化能力，充分体现了频域信息和颜色空间信息的重要性，保证了融合结果的有效性。对医学脑影像进行了处理和分析，包括 MRI-CT 影像融合、MRI-SPECT 影像融合和 MRI-PET 影像融合。实验表明，与现有方法相比，该方法具有明显的优势。本节还请求临床医生通过问卷调查的方式对融合结果进行评估。统计数据也证明了所提出的 MsgFusion 达到了最好的融合效果。在未来将考虑扩展框架，将 CT、MRI、PET、SPECT、DTI 和两种或两种以上其他成像手段整合在一起，并将其应用于临床诊断。

## 3.2 基于脑影像的多维特征自适应融合

### 3.2.1 引言

医学成像在各种临床应用中发挥着重要作用。其中，MRI 和 PET 为多种疾病提供了影像学依据，广泛应用于各种疾病的临床诊断，如脑良恶性肿瘤、抑郁症、早发性 AD、脑缺血等。MRI 和 PET 脑影像的融合已被证明具有临床意义 [188, 189]。各种脑部疾病的诊断通常依赖于 MRI-PET 成像。脑病变的动态观察和脑肿瘤的定性分析也非常重要 [161]。MRI 技术是通过捕捉人体内的电磁信号以实现成像的一种手段；该技术能清晰显示软组织，空间分辨率高，有利于病变范围的确定。PET 扫描能够揭示人体内在的详尽生理活动和代谢状态信息，并且尽管空间分辨率低，但能显示人体组织中病变的异常变化。因此，MRI 和 PET 的融合可以联合解剖结构信息和生理代谢信息，帮助医务人员更好地诊断脑部异常。MRI 与 PET 结合的广泛需求也促使了一体化 MRI-PET 成像设备的发展（混合 PET/MRI），但仍存在几个问题：PET 和 MRI 系统相互干扰，购买和维护成本高，采集后的融合方法对设备性能至关重要，而目前应用于混合 PET/MRI 的融合方法多为传统方法，存在诸多局限性。因此，通过从不同设备开发更好的 MRI 和 PET 融合算法以及提高混合 PET/MRI 的性能来开发更有效的 MRI-PET 融合方法是重要的。由于不同模态的医学影像各自具有独特的特征属性，MRI-PET 影像的融合具有挑战性：1. 尽可能保留和增强 MRI 和 PET 的关键信息。2. 选择适当的客观评价指标具有挑战性。3. 在融合任务中，缺少真实对照组，因此损失层的正确选择变得至关重要。

目前，用于 MRI-PET 医学影像融合的传统方法有简单加权 [148]、多分辨率金字塔 [190]、小波变换 [191]、颜色空间 [149]、主成分分析 [192]、人类视觉系统 [193] 等。近年来，深度学习已成为影像融合的代表性方法和研究热点。各

种基于深度学习的影像融合方法相继被提出 [153, 194, 36]。利用深度学习进行影像融合是一种有效的解决方案，为了追求更好的感知效果，影像融合模型的设计主要是基于一些复杂的设计规则的定义来增强变换和融合策略。本小节提出了一种基于无监督深度学习的 MRI-PET 医学影像融合新思路。首先建立三维特征提取模块，提取源影像的浅层、深层和多尺度特征，然后建立多维特征的仿射映射函数，融合多维特征。本节工作可以总结为以下三个主要贡献：

- 影像的多维分析分为三个模块，分别从源影像中提取浅层特征、深层特征和多尺度特征。这三个模块分别针对源影像中不同信息的重要性，尝试提取不同层次的特征，减少特征传递过程中的特征损失，提高后续特征融合的准确性。
- 建立仿射映射变换函数以保持不同维度特征之间的几何关系，并通过学习自适应地生成相关系数。从而在融合后的影像中充分保留 MRI 影像的空间纹理信息和 PET 影像的功能代谢信息等多维特征。
- 为了进一步增强融合影像的可视化效果，本文提出了一种基于能量的颜色增强算法，主要是增强融合影像中来自原始 PET 影像的能量信息。当用它来显示同一病例不同时间序列的异常区域时，病变的演变过程（例如，脑肿瘤）可以更好地追踪。

### 3.2.2 相关工作

目前，可应用于 MRI-PET 医学影像的融合方法包括一些传统的融合方法和基于深度学习 (DL-based) 的融合方法，每种类型都有自己的优缺点。在本节中，将讨论一些代表性的方法。

1) 传统融合方法：在传统的影像融合方法中，简单加权平均法、小波变换应用和颜色空间交换是三种具有代表性的像素级融合方法，常应用于医学影像融合。Li 等人 [148] 提出了一种基于潜在低秩表示 (LatLrr) 的简单有效的影像融合方法，旨在最大程度地保持原始影像中的有价值信息。该方法采用一种简单的加权平均融合策略。利用低秩聚类的思想，将影像分为低秩部分和重要部分，并将低秩部分和重要部分的加权和作为融合结果。LatLrr 能突出影像的全局结构信息，但局部结构保持能力和细节提取能力较差。Hla 等人 [195] 将低秩分解方法用于含噪影像融合，主要通过最小秩正则化得到融合结果。

小波变换也是一种经典的方法。小波不仅具有正交性、双正交性和紧性，而

且具有多分辨特性。Zhan 等人 [196] 提出了一种基于相位一致性融合 (PCF) 的影像融合方法; 他们使用相位一致性来提取影像中的局部和剧烈变化。该方法在空间上使用 Gabor 小波滤波器来改善相位一致性。小波变换的应用 [197, 191] 也有利于理解影像, 特别是医学影像。小波变换具有多分辨率的特点, 可以从粗到细观察信号, 但小波变换方法需要一个合适的母小波和一个可行的分解层次。为了便于医生理解影像, 特别是观察 PET 影像中描绘的生理代谢变化, 一些研究人员已经开发了融合策略来保持伪彩色。Du 等人 [149] 使用双尺度策略和色域变换来保持伪彩色信息, 从而整合灰度影像 (如 MRI) 和伪彩色影像 (如 PET 和 SPECT), 并专注于保留伪彩色信息。然而, 颜色空间转换也会造成一些信息损失。综上所述, 一些传统的融合方法可以达到高质量的融合效果, 但大多数融合方法依赖于人工对特定影像类型的特征提取规则, 包括参数设置。随着影像类型和数量的增加, 特征提取变得越来越复杂。同时, 传统方法的泛化能力很弱。

2) 基于 DL 的融合方法: 在多模态影像融合中, 将传统的技术手段与深度学习策略相融合, 可以提高模型性能。Zhong 等人 [194] 提出了一种基于 CNN 的联合影像融合和超分辨率方法。Rajalingam 等人 [198] 提出了一种深度引导混合多模医学影像融合 (HMMIF) 方法, 该方法已应用于神经囊虫病 (一种退行性疾病) 的神经病理学。经典的传统方法和深度学习方法的结合可以有效地提高网络的性能。然而, 一些传统的方法需要先验知识, 并结合传统的方法可能会增加时间和空间复杂度。最近, 一些端到端的深度神经网络已经应用于医学影像融合 [168, 199]。Rajalingam 等人 [200] 提出的新的医学影像融合方法使用了联合卷积神经网络来生成加权图, 以融合来自多模态医学影像的像素运动信息。Zhao 等人 [201] 提出了一种基于表示学习的通用融合框架, 该框架研究的是基于边缘细节和对比度的特定领域的无参考感知度量损失, 以优化学习过程, 使融合影像呈现出更具体的表现。

为了加强 U-Net 在捕捉全局特征上的能力, Xiao 等研究者 [201] 设计了两个单元 (特征提取单元 (GFPE) 和注意力联合单元 (GACU) ), 旨在高效地捕获并利用影像中的全局语义和边缘特性。无监督深度神经网络模型, 如 DeepVTF[202] 和 VIFNet[203] 也被提出。DeepVTF 方法在输入常规真彩色和融合输出之间建立视觉相似性度量, 以获得自然直观的影像。VIFNet 是定义了一个强大的混合损失函数, 它是由一个修改的结构相似性度量和总变分组成。Nishant 等人 [155] 提出了一种名为 FunFuseAn 的深度神经网络模型, 该模型

使用 SSIM 作为损失函数来融合 MRI 和 PET 影像。Xu 等人 [204] 介绍了一种基于梯度和连通域多焦点影像的深度融合模型。为了克服深度网络应用于多聚焦影像融合时梯度丢失的障碍，设计了一种掩模网络直接生成二值掩模。采用一致性验证策略，通过调整初始二值模板生成融合结果。Yu 等人 [157] 提出了融合模型 IFCNN，它使用了迁移学习技术和最大张量策略。Xu 等人 [159] 介绍了一种无监督和统一的密集连接网络，称为 FusionDN，用于不同类型的影像融合任务。该方法基于源影像的保留程度（即，源影像的质量和信息量）。

特别是近两年来，几种新的方法，如 [205–207] 被提出。Zhao 等人 [205] 提出了一种用于多模态影像融合的双分支变换器-CNN 网络架构 (CDDFuse)。为了提取特定的模态特征和模态共享特征，他们采用了 Restormer、Lite Transformer 和可逆神经网络模块。Xu 等人 [206] 构建了一种全新的端到端统一无监督影像融合架构 (U2Fusion)。通过特征提取和信息测量，U2Fusion 自动估计源影像中对应关系的重要性，并提供自适应信息保留。Liang 等人 [207] 提出了一种强大的影像分解模型 (DeFusion)，它通过自监督表示学习来执行融合任务，而无需任何配对数据或复杂的损失函数。深度学习方法的另一个重要分支是基于生成对抗网络 (GAN) 的方法 [208, 209]。受条件生成对抗网络 (CGAN) 的启发，Ma 等人 [156] 将 GAN 引入影像融合和无监督 GAN 网络影像融合框架 FusionGAN。

与传统方法相比，深度学习方法具有自主学习、灵活的实时处理和良好的泛化性能的特点。在基于深度学习的方法中已经发现了相当大的潜力。但目前，专门为医学影像融合设计的深度神经网络仍处于初期。可应用于混合 PET/MRI 的相应方法更是罕见。在这小节中，专注于从 MRI 和 PET 影像中提取和保留关键特征。为了实现这一目标，设计了多维特征模型和自适应线性融合策略。

### 3.2.3 多维特征自适应线性融合网络

如图3.13所示，本文的工作包括两个方面。首先，本节提出了一个基于多维特征自适应线性融合的 MRI 和 PET 脑影像融合网络 (adaptive linear fusion method for multi-dimensional features, MdAFuse)。其次，本节提出了一种基于能量的可视化方法，以进一步增强融合影像。在融合网络中，本节采用了多维特征提取方法和自适应线性融合策略，这两方面将分别在3.2.3.1小节和3.2.3.2小节中详细介绍。3.2.3.3小节将解释如何设置该网络的损失函数。随后，将在小节3.2.3.4中描述所提出的可视化方法。

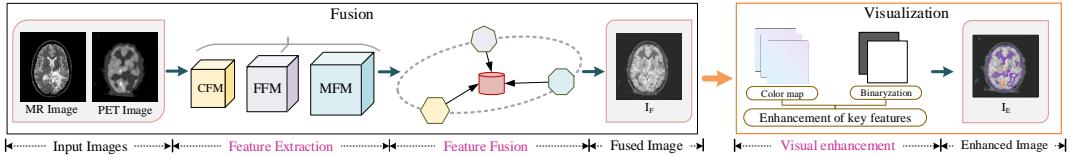


图 3.13: 本小节提出的方法流程图（包括两个阶段：影像融合（图3.14）和可视化增强（图3.15））

### 3.2.3.1 多维度特征的提取策略

为了从源影像中提取尽可能多的特征，本节采用多维度策略，充分利用空间和通道特征，建立了浅层特征模块 (coarse feature module, CFM)，深层特征模块 (coarse feature module, FFM) 以及多尺度特征模块 (multi-scale feature module, MFM)，分别提取浅层特征信息、深层特征信息以及多尺度特征信息。图3.14中的蓝色虚线矩形表示特定的特征提取过程。其中，CFM 为紫红色实线边缘黄底区域，FFM 为蓝色实线边缘紫底区域，MFM 为紫色边缘蓝底区域。在图3.14中，不同的小方块代表三个步骤的组合 (卷积运算，批量归一化和 ReLU 激活)，不同颜色的方块代表不同大小的卷积核，黄色是 1，蓝色是 3，绿色是 5，紫色是 7，粉红色是 3。不同大小的正方形表示不同的特征图大小和输出通道数。图3.14右下角的灰色虚线框中有不同通道号的定义，即不同宽度的长方形代表不同通道的输出数量，并且颜色越深，通道的数量越大。对于输入影像，特征提取模块将获得三个维度的特征，即  $CF_1$ 、 $FF_1$  和  $MF_1$ 。这些特征提取模块是根据人类视觉感知的感知特性和影像的物理特性设计的。

在图3.14中，CFM 由紫红色实线包围的黄色背景区域表示。CFM 代表浅层特征模块，它负责通过卷积运算从影像中提取浅层特征。该模块利用较大的卷积核和较小的特征通道来捕捉影像的主要形状结构和边缘轮廓信息，代表影像背景和物体轮廓等全局低频信息。从图3.14所示的特征图中，可以观察到，CFM 模块提取的 CF 特征图表现出更清晰的轮廓。类似地，在图3.14中，FFM 由蓝色实线包围的紫色背景区域表示。FFM 指的是深层特征模块，它专注于深层特征提取。该模块采用较小的卷积核和较大的特征通道来捕获纹理细节，例如边缘，精细纹理细节和噪声，表示影像内具有显著变化的局部高频信息。用于深层特征提取的通道数远大于用于浅层特征提取的通道数。从图3.14所示的特征图中，可以观察到由 FFM 模块提取的 FF 特征图显示了更明显的纹理细节。

在图3.14中，MFM 是紫色包围的蓝色背景的区域。MFM 主要是从金字塔

和分类的思想中获取影像的空间结构,从而获得多尺度影像之间的互补信息,利用多分辨率和多尺度影像之间的互补信息实现精细融合。在 MFM 中,还使用了深度分离卷积网络,便于挖掘更丰富有用的信息,为后续的影像理解和应用分析提供了良好的基础。另外,本节在构建网络时使用跳跃连接方法,与参考文献 [210, 211] 一样。在本节的特征提取模块的设计中,更强调局部细粒度的信息,而不仅是全局特征。因此,使用短跳跃连接有利于局部信息的传输和保存,使网络能够更好地学习和利用这些局部化特征。因此,在特征提取过程中采用了短跳级联的方法,以获得更详细的信息,实现特征的逐步融合。

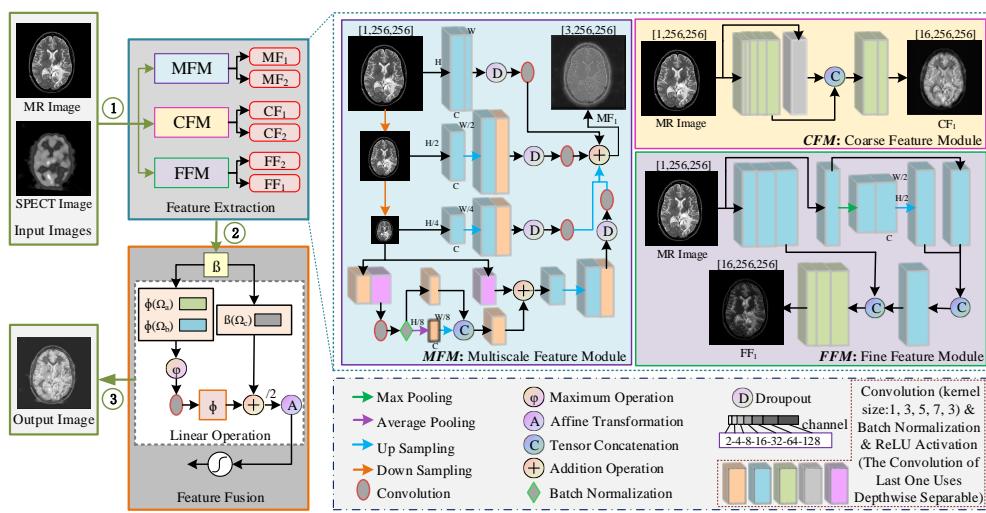


图 3.14: 多模态影像融合网络的框架及具体的实现步骤

### 3.2.3.2 多维度特征的融合策略

MdAFuse 从三个维度提取特征。浅层特征主要体现在边缘轮廓信息上,而深层特征则能保留纹理细节。为了从一幅影像中同时获得高频信息和低频信息,将每幅影像的粗特征和细特征相加,以获得具有更多信息的影像特征。为了突出两幅影像中的显著特征,特别是 PET 影像上的突出信息(通常对应于异常区域),本节使用最大值来获得更突出的部分。多尺度特征可以提供更多的结构信息,也可以补偿在跨尺度上丢失的局部细节。因此,浅层和深层特征必须是多尺度的。对于每个输入,将获得三个特征(CF、FF 和 MF)。那么,当输入两个模态影像时,共提取六个特征。并提出了一种自适应融合策略,使融合结果

同时包含6个特征。本节设计的目标函数如(3.15):

$$F = \delta(\Phi((v(\varphi(\Phi(\Omega_a), \Phi(\Omega_b))) + \beta(\Omega_c))/2)), \quad (3.15)$$

其中,  $F$  是融合的结果。 $\delta$  是激活函数, 本文采用  $\tanh$ 。 $\Omega_a = \{CF_1\} \cup \{FF_1\}$ ,  $\Omega_b = \{CF_2\} \cup \{FF_2\}$ ,  $\Omega_c = (MF_1 + MF_2)/2$ . 当计算  $\Omega_a$  的  $\Phi$  时,  $\Gamma_1$  和  $\Gamma_2$  分别表示  $CF_1$  和  $FF_1$  特征。类似地, 如果计算  $\Omega_b$  的  $\Phi$ ,  $\Gamma_1$  和  $\Gamma_2$  分别表示  $CF_2$  和  $FF_2$  特征。 $\Phi$  是通过仿射变换获得的两个特征之和, 如等式(3.16)。在这个等式中,  $U$  表示特征图的大小。 $A = Diag(B)$ ,  $\beta(\Gamma) = A\Gamma(i) + e$ ,  $B$  和  $e$  是可学习参数。设置大小的规则与后归一化的规则相同[212]。初始化  $B$  的值为1,  $e$  为0。 $\beta$  是仿射变换函数, 用于表示特征之间的相关性。在几何学领域, 仿射变换被描述为一种结合了非奇异线性映射与平移的变换过程, 它连接了两个向量空间并保持了它们的仿射结构, 可用于保持二维影像之间的位置关系。因此, 仿射变换被用作融合特征, 以充分利用特征之间的相关性并保持它们。

$$\Phi(\Gamma) = \beta\left(\sum_{i=1}^U \Gamma_1(i)\right) + \beta\left(\sum_{i=1}^U \Gamma_2(i)\right) = A \sum_{j=1}^2 \sum_{i=1}^U \Gamma_j(i) + 2e, \quad (3.16)$$

$v$  表示卷积运算, 以保持输入和输出维度一致。 $\varphi$  是取最大值操作为突出显示重要信息。

### 3.2.3.3 损失函数的构建

为了更准确地重建源影像, 在训练阶段考虑了特征提取阶段和融合过程, 并获得了最小损失函数如式(3.17):

$$L = \alpha L_S + \beta L_M \quad (3.17)$$

同3.1小节中的损失函数(3.12)设计相同, 总损失函数由结构相似度  $L_S$  和像素损失  $L_M$  组成, 具体计算同式(3.13)和(3.14):

在等式(3.17)中,  $\alpha$  和  $\beta$  是设置在区间  $[0, 1]$  中的权重参数。通过对比分析, 我将  $\alpha$  设为 0.3,  $\beta$  设为 0.7, 结果最佳。在训练过程中, 初始学习率为 0.0001, 60 epoch, 批量大小为 2, 采用批量归一化, 优化函数为 Adam。

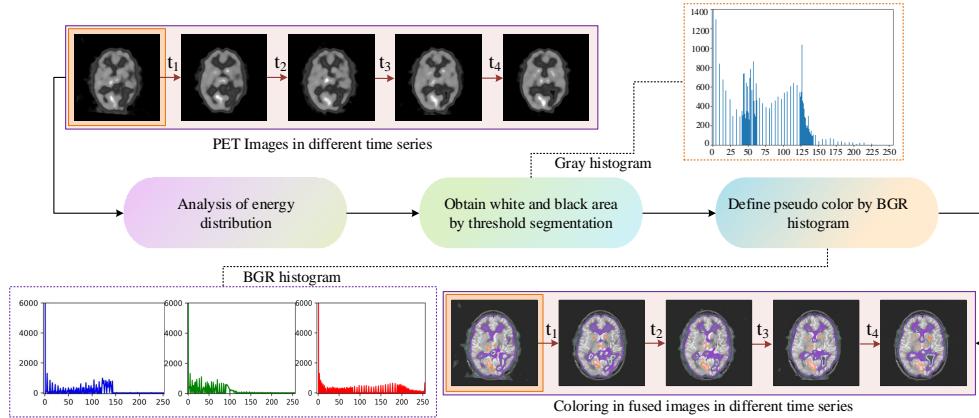


图 3.15: 增强融合影像中的关键特征策略

### 3.2.3.4 关键特征的增强方案

PET 或 SPECT 影像的空间分辨率有限，边界灰度区分不清，限制了异常和正常区域的区分。为了提高可视性，伪彩色处理通常应用于临床环境，帮助医生识别边界细节并提高疾病诊断的准确性。

在通过 MRI-PET/SPECT 融合生成灰度融合影像时，来自 MRI 的空间纹理是清晰的，而低分辨率灰度 PET/SPECT 影像难以揭示葡萄糖代谢的显著异常。现有的伪彩色方法涉及基于灰度的离散颜色映射和从灰度到颜色属性的连续转换。然而，这些技术中没有一种侧重于突出 PET/SPECT 影像中的这些显著异常。为了强调这些异常，基于 PET/SPECT 能量的灰度直方图，本节提出了一种颜色增强算法。该算法增强了异常信息的视觉效果，详细描述见算法2和图3.15。

其中， $I_g$  表示灰度 PET/SPECT， $I_f$  为融合的影像， $I_e$  为增强的影像， $\eta_w$  和  $\eta_b$  分别为阈值二值化过程中的白色和黑色区域。 $\rho_w$  和  $\rho_b$  是白色和黑色区域的两种自定义颜色。 $I_c[a_{ij}](i, j)$  表示  $I_c$  中对应于  $I_a$  的区域。关键特征增强过程如图3.15所示。在这部分工作中，主要目的是获得可以代表葡萄糖含量或生理代谢的 PET/SPECT 影像和增强融合结果中的重要信息。

首先，分析灰度 PET/SPECT 影像中的能量分布，包括不同像素的大小分布、不同位置像素的分布、有用信息的集中区域以及代表相似信息的区域数量。在 PET/SPECT 中主要存在两种形式的能量分布。一种是葡萄糖含量高或代谢旺盛的组织，显示在白色区域。另一种是葡萄糖含量低或代谢弱的组织，用黑

---

**算法 2** Energy-based color enhancement

---

**输入:** Images  $I_g, I_f$

**输出:** Image  $I_e$

```

1:  $x \leftarrow$  the weight of  $I_g$ 
2:  $y \leftarrow$  the height of  $I_g$ 
3:  $\rho_w \leftarrow$  self-defined color for image of white area
4:  $\rho_b \leftarrow$  self-defined color for image of black area
5:  $\eta_w, \eta_b \leftarrow \gamma(I_g)$ 
6:  $\sigma \leftarrow \text{COLORUP}(I_g, I_f, \eta_w, \rho_w)$ 
7:  $I_e \leftarrow \text{COLORUP}(I_g, \sigma, \eta_b, \rho_b)$ 
8: function COLORUP( $I_g, I_f, I_a, I_c$ )
9:    $I \leftarrow I_f.\text{copy}()$ 
10:  if  $I_c == \rho_w$  then
11:     $\mu \leftarrow 5\sigma - 20, \sigma = 0, 2, 6, 8$ 
12:     $\xi \leftarrow 5\mu - 20, 0 \leq \mu \leq 4$ 
13:  else if  $I_c == \rho_b$  then
14:     $\mu \leftarrow 5\sigma - 20, 0 \leq \sigma \leq 11 \& \sigma \neq 4$ 
15:     $\xi \leftarrow 5\mu + 10, 0 \leq \mu \leq 11$ 
16:  end if
17:  for  $i=1$  to  $x$  do
18:    for  $j=1$  to  $y$  do
19:      if  $I_a(i, j) \neq 0$  then
20:         $a_{ij} \leftarrow I_a(i, j)$ 
21:         $\psi \leftarrow I_f(i, j) - I_g(i, j)$ 
22:        for  $k=1$  to ( $\text{len}(\mu)$ ) do
23:           $I' \leftarrow I[a_{ij}]$ 
24:           $I'_c \leftarrow I_c[a_{ij}]$ 
25:          if  $\psi \leq \mu[k]$  then
26:             $I'(i, j) \leftarrow I'_c(i, j) + \xi[k]$ 
27:          else
28:             $I'(i, j) \leftarrow I'_c(i, j) + \xi[-1]$ 
29:          end if
30:        end for
31:      end if
32:    end for
33:  end for
34: return  $I'$ 
35: end function

```

---

色区域表示。其次，需要在灰色 PET/SPECT 中获得黑色和白色区域。使用阈值法来提取这两种类型的影像。阈值化的关键在于阈值的设置， $\tau_w$  是提取白色区域块的阈值，其由等式 (3.19) 获得， $\zeta$  是指直方图中峰值的灰度值， $\kappa$  表示指峰的总数。在公式 (3.18) 中的函数表示直方图的计算， $\phi$  表示其转换到一维的操作。参数  $\chi$  表示峰与其周围谷之间的最小高度差，在图3.15的示例中设置为 100。 $\vartheta$  的作用是寻找直方图中的峰值。对于黑色区域，设置了两个阈值。为了消除一些噪声并避免假溢出，添加另一个阈值  $\tau_{b1}$ ，其被设置为 10。 $\tau_{b2}$  是通过 Otsu 方法获得的全局阈值。算法 1 中的  $\gamma$  表示白色特征和黑色特征的两个二值影像，这两个二值影像是由公式 (3.20) 和 (3.21) 分别的计算获得。在式 (3.20) 和 (3.21)， $I_g(i, j)$  表示影像  $I_g$  中第  $i$  行第  $j$  列的位置处的像素值。

$$\zeta = \vartheta(\chi(\phi(I_g)), p), \quad (3.18)$$

$$\tau_w = \zeta^l, \quad l = \kappa - 1, \quad (3.19)$$

$$I_w(i, j) = \begin{cases} 0, & I_g(i, j) \leq \tau_w \\ 255, & I_g(i, j) > \tau_w \end{cases}, \quad (3.20)$$

$$I_b(i, j) = \begin{cases} 0, & I_g(i, j) < \tau_{b1} \parallel I_g(i, j) > \tau_{b2} \\ 255, & I_g(i, j) > \tau_{b1} \& I_g(i, j) \leq \tau_{b2} \end{cases}, \quad (3.21)$$

最后，根据 BGR 直方图定义伪彩色，具体过程如算法 2 所示，在图3.15中包含三幅影像的直方图（根据伪彩色 PET/SPECT 影像计算的相应 BGR 的三种颜色的直方图）。为了便于观察，纵轴的值仅小于 6000。从 BGR 直方图中，可以确定蓝色、绿色和红色像素值的大致范围。因此，白色区域呈现黄色显示，黑色区域呈现深蓝色显示。 $\rho_w$  和  $\rho_b$  分别用于表示黄色和深蓝色区域。亮度根据像素的大小而变化，并且对于没有辅助诊断的区域不显示颜色。然后，将融合影像与原始 PET/SPECT 灰度影像之间的像素值进行比较，并做出相应的改变。具体过程如算法2的 *ColorUp* 所示，以及得到的影像的着色效果如图3.15所示。

### 3.2.4 多模态影像融合的实验

本节使用 Python 3.7.9 和 Pytorch 版本 1.7.1+cu110 及其相应的 CUDA 版本 11.1。GNU/Linux x86, GeForce RTX 3090 Ti GPU 20GB RAM-64 系统。在训练过程中，使用的数据来自 ADNI (<https://adni.loni.usc.edu/data-samples/access-data/>)，本节下载了 555 个 MRI 和 PET 影像对。样本的年龄从 55 岁到 90 岁不等，样本的性别有男有女。所有影像均作为轴向切片进行分析，体素尺寸为  $1.0 \times 1.0 \times 1.0 \text{ mm}^3$ 。通过非相交数据集测试对本节的训练模型进行交叉验证。从 ADNI 和全脑图谱 (<http://www.med.harvard.edu/AANLIB/home.html>) 中获得了 80 个独立不重复受试者的 137 个影像对进行测试。其中，74 对 MRI-SPECT, 42 对 MRI-PET, 21 对 MRI-CT。在图3.16和图3.18的融合实验中，测试数据是具有胶质瘤的 MRI 和 SPECT 脑影像，这是脑神经系统中的典型应用。MRI 与 PET/SPECT 的融合可以整合生物解剖信息和生理代谢信息，帮助医生对病变进行定位和诊断。实验中还分析和比较了现有的几种基于传统方法和深度学习方法的融合算法。

为了定量衡量不同融合方法的融合效果，本节采用了 9 种质量评价指标，信息熵 (EN) [213]，互信息 (MI) [214]，标准差 (SD) [215]，视觉信息保真度 (VIF) [216]，结构相似性 (SSIM) [125]，差异相关性之和 (SCD) [132]，相关系数 (CC) [135]，均方误差 (MSE) [133] 和空间频率误差比 (rSFe) [184]。对于 MSE，较小的值表示更好的性能。对于 rSFe，其较小绝对值对应于较好的融合效果。而其他七个评价指标则相反。rSFe 是一个相对不常见的评价指标，用于衡量影像的整体活动水平，它由像素点中的四个空间频率（行、列、主对角线、次对角线）和四个方向的一阶梯度值（水平、垂直、主对角线、次对角线）整合而成，其数值绝对值越小，则表明融合效果越佳。为了便于比较和分析，本节对部分指标的计算结果进行了归一化处理。如表3.4所示，本节对每个度量执行线性变换，其中  $k_i$  和  $b_i$  ( $i = 1, 2, 3$ ) 分别表示图3.17、图3.19和图3.22中的变换系数。

#### 3.2.4.1 与传统方法的实验对比分析

在本小节测试了三种传统的融合方法，并将它们与本节提出的方法进行了比较。如图3.16所示，第一列和第二列分别显示来自具有脑异常的一些病例的 MRI 和 SPECT 源影像。以下各列显示了通过三种传统方法和建议方法获得的结果，每种方法各一列。在图3.16中，可以观察到来自不同融合方法的融合影

表 3.4: 每个度量指标值上的线性变换, 其中  $k_i$  和  $b_i$  ( $i = 1, 2, 3$ ) 分别表示图3.17、图3.19和图3.22中的变换系数

	EN	MI	SD	VIF	SCD	SSIM	CC	MSE	rSFe
$k_1$	8	6	1/2.1	48	29	16	42	18	-37
$b_1$	-11	0	0	0	2	0	5	5	8
$k_2$	5.6	5	0.4	31	18	16	26	240	-32
$b_2$	0	0	0	0	0	0	0	3	6
$k_3$	1	1	1/13	5	2	10	100	15	-8
$b_3$	-1	0	0	0	0	-14	90	0	0

像之间的差异。PCF 是一种利用小波变换原理实现影像融合的方法, 融合后的影像信息丢失严重。LatLrr 是一种低秩表示方法, 其融合策略是简单的加权平均。这种方法的结果可以保持良好的纹理细节, 但较少的 SPECT 影像的信息。atsIF (自适应双尺度影像融合) [149] 是一种通过颜色空间转换的融合策略, 可以保留更多的 SPECT 信息, 但 MRI 细节丢失。与其他方法相比, 本节提出的方法可以更好地保持两个源影像的信息。在第一行中, 每个影像的三个局部区域使用三个彩色正方形包围并被放大, 并分别显示在第二, 第三和第四行中, 以进行更清晰的比较。

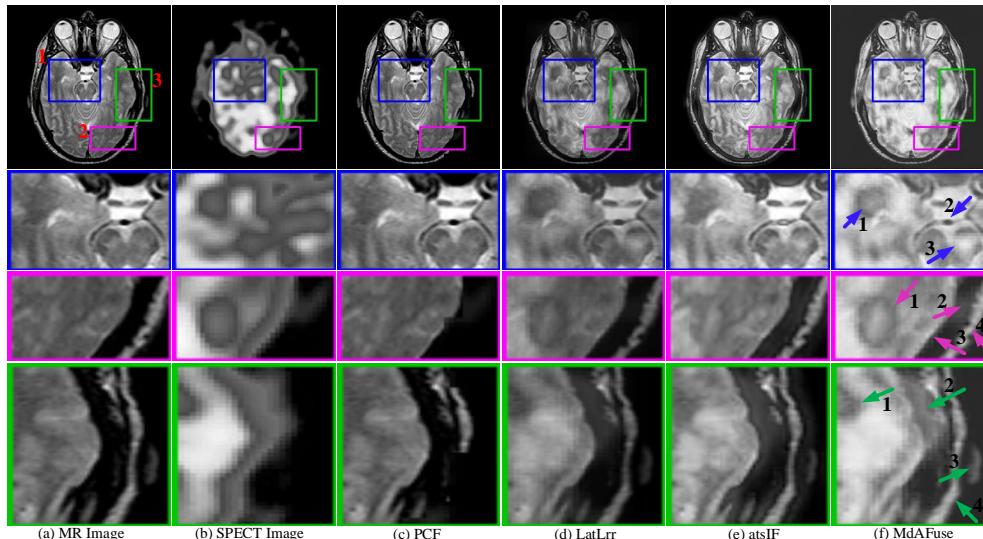


图 3.16: 比较传统的融合方法和本节提出的 MdAFuse 方法的 MRI-SPECT 融合结果

图3.16中的第二行列出了与第一行中蓝色正方形内的部分相对应的放大影

像。在最后一列，本节用三个箭头标记了颞叶、基底动脉和脑桥。颞叶在 MRI 上呈灰色，不规则。SPECT 显示较浅的灰色圆圈，具有明显的色差。在 PCF 结果中，该区域不明显，MRI 上的特征完全保留，但 SPECT 上没有。在 LatLrr 融合的结果中，该区域可以保留两个源影像的信息，但也丢失了部分 SPECT 信息，例如颞叶周围的像素值，这是没有区别的。atsIF 方法的效果在亮度方面有了很大的提高，可以在 SPECT 中保留更多的信息，但合并后的源影像会出现信息丢失。与其他方法相比，本节提出的方法可以很好地保留颞叶区域的特征。对于基底动脉区，也注意到明显的差异。在 MRI 中，该区域是一个小的黑色椭圆，而在 SPECT 中，它是一个均匀的灰色区域，没有形状特征。PCF 完全保留了 MRI 上的信息，LatLrr 和所提出的方法显示该位置区域为深灰色，形状没有改变。atsIF 的融合结果变成了一个白色区域，在小椭圆的外面有一个黑色的圆圈很明显。箭头 3 指向脑桥区，这是 MRI 上灰度分布不均匀的区域，SPECT 影像上亮度明显且类似三角形的小区域。同样，在 PCF 的融合结果中找不到 SPECT 的特征信息，在 LatLrr 和 atsIF 结果中该位置区域的亮度信息也不明显。相对而言，本小节提出的方法的结果能清晰地反映 MRI 和 SPECT 的特征。

图3.16中的第三行列出了与第一行中紫色方块内的区域相对应的放大影像。同样，在最后一列中，标记了指向枕叶、枕骨、侧窦和枕叶轮廓的四个箭头。从四个箭头的显示也证明了该方法的融合效果更好，尤其是在第一个箭头区域。图3.16中的第四行列出了与第一行中蓝色方块内的部分对应的放大影像。最后一列中的四个箭头指向颞叶、颞肌、耳廓和外侧枕骨。从四个箭头的显示也证明了本小节方法的融合效果更好。

传统的影像融合方法主要属于像素级的融合操作，这种方法追求的是最大限度地保留信息。为了充分验证所设计的影像融合模型的实际效能，使用 9 个度量 (EN, MI, SD, VIF, SSIM, SCD, CC, MSE, rSFe) 来评估融合结果的质量，如图3.17所示。在图3.17中，可以观察到本节的 MdAFuse 方法的大多数度量都具有优势。具体地，SCD 值比次优值高 9%。

### 3.2.4.2 与新型方法的实验对比分析

在这一部分中，对几种基于 DL 的影像融合方法进行了实验，并对它们的融合结果进行了比较。如图3.18所示，第一列和第二列示出了要融合的 MRI 和 SPECT 源影像，并且随后的 6 列展示了通过 6 种融合方法（包括所提出的

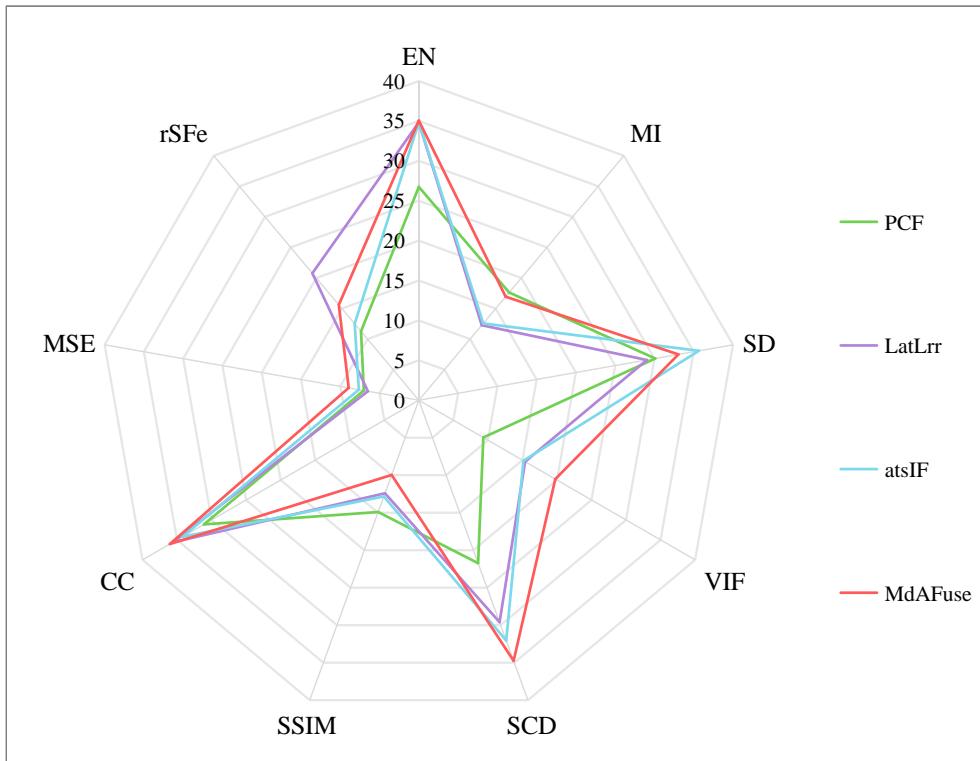


图 3.17: 图3.16中与基于传统医学影像融合方法的客观评估比较分析

方法: MdAFuse) 获得的融合结果。从这些结果中, 可以观察到不同的融合方法之间的差异。总体而言, 可以观察到 U2Fusion 和 CDDFused 生成的融合影像似乎没有考虑 SPECT 的信息。虽然 FusionDN 和 DeFusion 方法保留了源影像中的 SPECT 信息, 但仍然可以观察到一些信息丢失。FunFuseAn 方法的影像融合效果相对较好, 与本文提出的 MdAFuse 方法在视觉上具有相似性。因此, 放大了三个小区域, 以进一步比较和分析。在第一行中, 使用三种颜色标记三个不同的区域并将其放大, 如第二至第四行所示。

首先, 图3.18中蓝色方块标记的 ROI 主要反映侧脑室前角、丘脑和颞叶的信息。在 MRI 中, 侧脑室的前角呈白色且对称, 而在 SPECT 影像中, 左侧脑室是灰色的。左侧丘脑区域在 MRI 中为深灰色, 而晶状体区域在 SPECT 影像中为明亮。右侧丘脑区域呈白色, 在 MRI 中看起来像丛林, 而在 SPECT 影像中, 它呈黑色, 不规则。颞叶的面积在蓝色方块的左右两侧测量, 主要反映在 MRI 上, 是连续的。在图3.18的最后一列中, 四个区域用彩色方块标记。区域 1 为侧脑室左前角。U2Fusion 和 DeFusion 中这个位置的结构相对完整, 但影

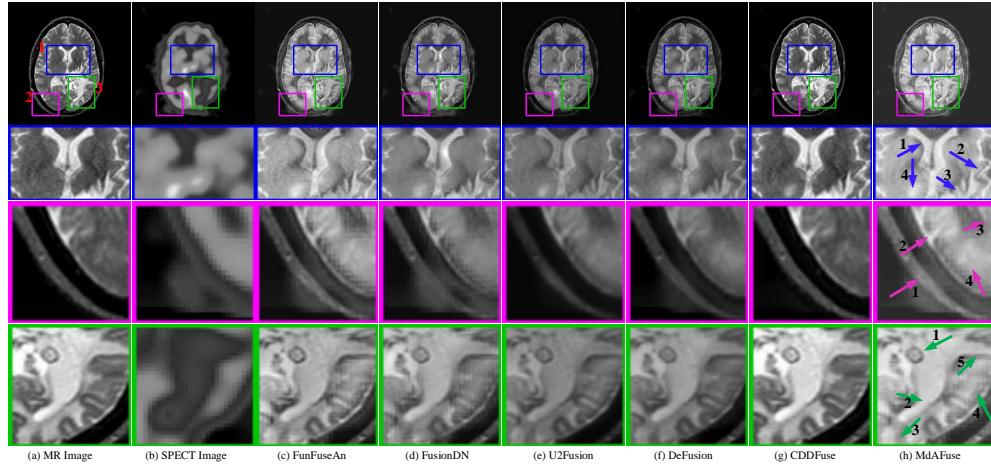


图 3.18: 本节提出的 MdAFuse 方法与现有的基于 DL 的方法性能比较

像看起来模糊。CDDFuse 突出显示 MRI 信息。FunFuseAn 类似于 FusionDN，两者都有色差。侧脑室左前角 FusionDN 较完整，而侧脑室右前角像素值分布不均匀。相比之下，MdAFuse 可以更完整地保留 MRI 特征。区域 2 是颞叶。在 FunFuseAn 和 FusionDN 的结果中，结构相对完整和连续，但颜色相对较暗。U2Fusion 和 DeFusion 的结果可以很好地保留 MRI 特征，但存在伪影。区域 3 是丘脑右侧的位置，在 MRI 上应该反映更多的纹理细节。在 U2Fusion 和 DeFusion 的结果中，位置不明显。FunFuseAn、FusionDN 和 CDDFuse 很相似，组织结构不完整。相比之下，MdAFuse 可以完全保留右侧丘脑。区域 4 为左侧丘脑，应显示一个亮点的扩散区。FusionDN 和 DeFusion 的结果显示无明显状态。U2Fusion 的结果可以识别出一个亮度块，但该区域的灰度值比较均匀，没有发现差异变化；其他方法可以明显地发现亮度块，可以检测出更多层次的亮度特征。

其次，在图3.18中放大的第二个紫色方块中，最后一列标记的四个区域是脑膜、枕叶、侧脑室枕角和大脑皮层（灰质）。在 MRI 中，ROI 1 显示为粗灰色柱，而 SPECT 影像显示不均匀的波浪状灰色块。FusionDN 和 DeFusion 的结果都显示出一些丢失，并且没有保留 MRI 的完整结构信息。U2Fusion 和 CDDFuse 主要保留 MRI 信息，忽略 SPECT 特征。相比之下，FunFuseAn 和 MdAFuse 考虑到了两幅原始影像中更重要的信息，并更完整地保留有用的特征。ROI 2 是枕叶，在 U2Fusion 和 DeFusion 结果中不容易找到其。在 FusionDN 的结果中，未观察到位置区域与周围区域之间的显著差异。ROI 3 是侧脑室的左枕角。

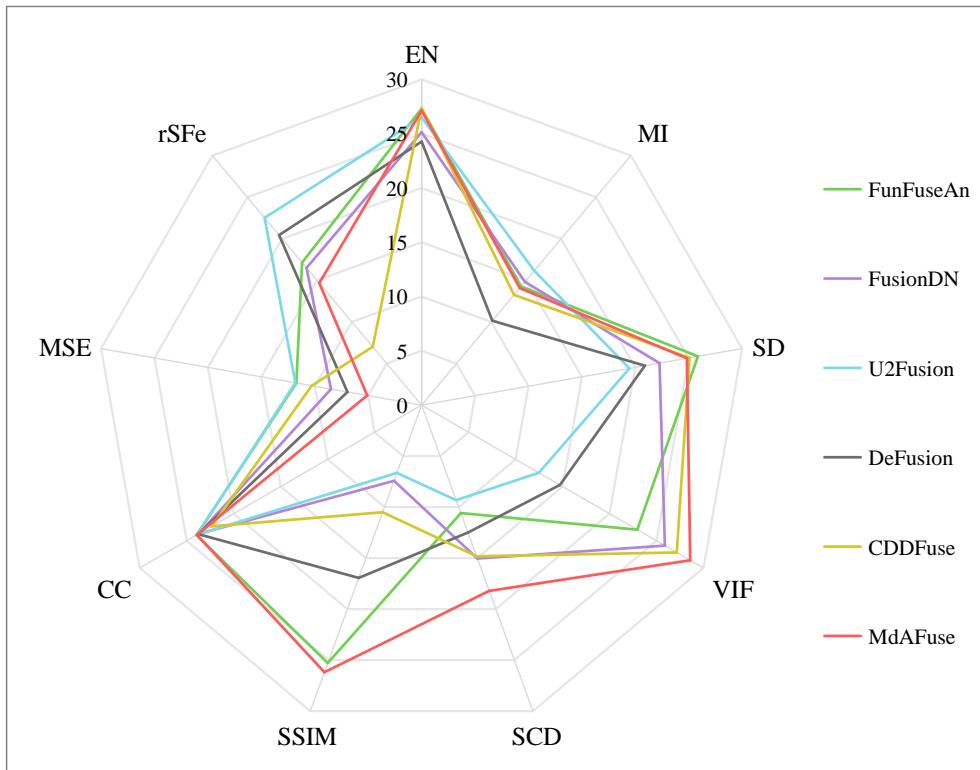


图 3.19: 图3.18中与基于 DL 医学影像融合方法的客观评估比较分析

可以在 MRI 上找到枕角的特征信息，这是 SPECT 影像上灰度值均匀的区域。FunFuseAn 可以发现细微的变化，但本节的方法的结果更明显。箭头 4 指的是大脑皮层的一部分（灰质）。在这一区域的 MRI 中，可以发现一个灰度值分布不均匀的散在区域，并且可以观察到精细的纹理特征。在 SPECT 影像中，该区域是具有均匀灰度值的浅灰色。从测试结果来看，只有 FunFuseAn 和 MdAFuse 的灰度值层次结构比较突出，能够比较完整地保留 MRI 的特征信息，而其他方法的灰度值层次结果都难以识别。

最后，对于图3.18中第三个蓝色正方形区域，信息量丰富，包括脉络丛、侧脑室枕角、枕叶、岛叶和小脑。五个箭头指向本节的方法的结果，这是明显不同于其他方法。区域 1 是脉络丛。在 MRI 中，整个位置是两个灰色同心圆。外圆的灰度值比内圆的灰度值深，差异明显。圆心的灰度值仅次于外圆，也比较突出。在 SPECT 影像中，灰度值更均匀。U2Fusion 和 DeFusion 的结果显示，内圆的半径相对较大，圆心不明显，内圆区域的像素值没有差异。FunFuseAn, FusionDN, CDDFuse 和本节的 MdAFuse 方法可以更好地保留 MRI 信息，并

且可以清楚地找到内圈和外圈以及圆心之间的明显差异。类似地，其他箭头指向在本节的 MdAFuse 方法中获得更好的效果。

基于深度学习的融合方法主要可以从它们在特征级的操作来区分，同样地使用 9 个质量评估度量 (EN、MI、SD、VIF、SSIM、SCD、CC、MSE、rSFe) 来评估融合影像结果的质量，如图3.19所示。从图3.19中，可以观察到，与其他方法相比，本节的方法在 6 个指标上具有优势，特别是在 SSIM, SCD, VIF 和 MSE 中，与次优值的差异非常明显。对于 rSFe 度量，CDDFuse 方法的值最小，但这个相对较小的值似乎有点异常。这可能是由于其较少保留 SPECT 影像中的信息，如图3.18 (g) 所示。相比之下，本节的方法在这个指标上表现得更好。虽然在 SD 和 MI 指标上的表现并不突出，但与最大值和第二大值的差距并不明显。

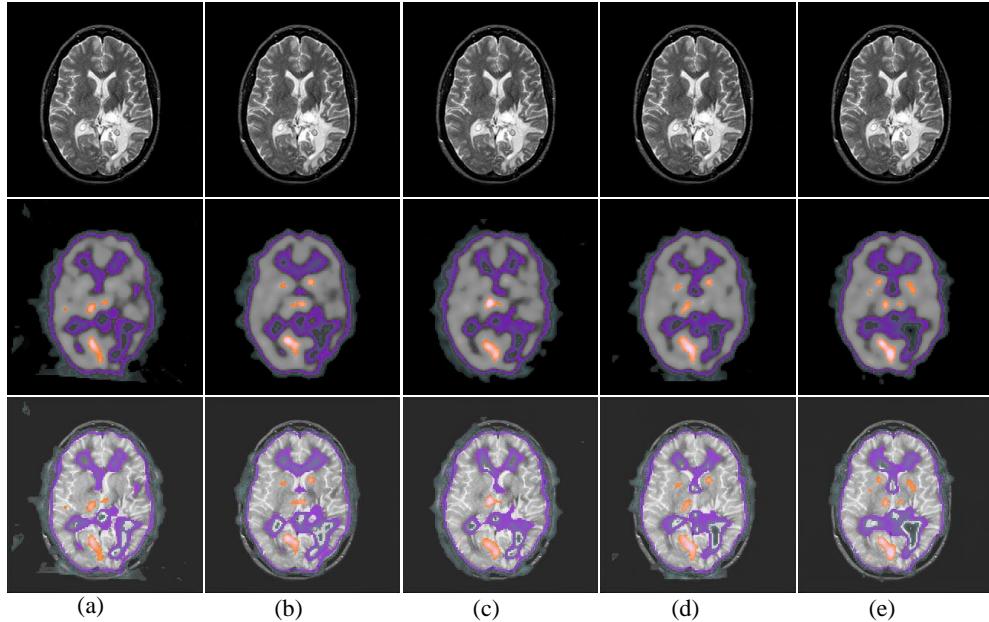


图 3.20: 增强某一受试者融合影像中的关键特征

采用 74 对 MRI-SPECT、42 对 MRI-PET 和 21 对 MRI-CT 医学影像融合结果与几种 SOTA 方法的定量分析结果见表3.5。表中的值表示每种方法在各种度量中的平均性能。在表3.5中，以粗体显示的值表示最佳值，而那些带下划线的值表示次优值。总体而言，本节的 MdAFuse 方法优于其他方法，在大多数指标中显示出上级结果，特别是在 EN、VIF、SSIM 和 MSE 上。CDDFuse 是目前最先进的方法，当应用于 MRI-SPECT 融合时，本节提出的方法在 VIF 和

SSIM 中分别获得了 11.14% 和 18.38% 的提升。此外，本节提出的方法将 MSE 降低了 2.63%。在 MRI-PET 融合的情况下，本节的方法使得 VIF 和 SSIM 分别提高了 13.76% 和 5.61%，而 MSE 降低了 4.50%。

表 3.5: 多模态医学影像融合结果与 SOTA 方法的定量比较与分析

<i>MRI-SPECT</i>	EN	MI	SD	VIF	SCD	SSIM	CC	MSE	rSFe
FunFuseAn	4.2348	2.6155	55.3604	0.8947	1.0452	1.5794	0.8917	0.0535	-0.3340
FusionDN	3.9001	<b>2.6912</b>	56.5726	0.5216	0.9051	0.4167	0.8756	0.0388	-0.2623
U2Fusion	4.0581	2.4404	42.2308	0.3188	1.7674	0.3016	0.8832	0.0344	-0.5399
DeFusion	3.7698	1.7543	49.9900	0.5393	0.7180	1.4480	0.8654	0.0225	-0.3775
CDDFuse	3.9151	2.4925	<b>58.3698</b>	0.9666	1.3458	1.4717	0.8381	0.0375	<b>-0.0449</b>
<b>MdAFuse</b>	<b>4.4135</b>	2.5343	58.0256	<b>1.0780</b>	<b>1.1483</b>	<b>1.6555</b>	0.8838	<b>0.0112</b>	-0.2508
<i>MRI-PET</i>	EN	MI	SD	VIF	SCD	SSIM	CC	MSE	rSFe
FunFuseAn	4.5540	2.5407	57.5225	0.7334	1.1915	1.4668	<b>0.8122</b>	0.1123	-0.3633
FusionDN	4.2210	<b>2.6952</b>	64.5598	0.4297	1.4188	0.4453	0.8024	0.0756	-0.2153
U2Fusion	4.4952	2.0185	50.1283	0.3383	1.5538	0.2741	0.7760	0.0453	-0.5511
DeFusion	4.1463	1.6765	63.5379	0.5199	1.4278	1.4335	0.7909	3.3017	-0.2963
CDDFuse	4.2248	2.0258	<b>70.7307</b>	0.7056	<b>1.6863</b>	1.4905	0.7958	0.0700	<b>-0.0316</b>
<b>MdAFuse</b>	<b>4.7241</b>	2.3880	60.5385	<b>0.8432</b>	1.2910	<b>1.5466</b>	0.8094	<b>0.0250</b>	-0.3086
<i>MRI-CT</i>	EN	MI	SD	VIF	SCD	SSIM	CC	MSE	rSFe
FunFuseAn	5.0245	3.1844	58.5545	0.6295	0.9037	1.5134	0.8263	0.0940	-0.4218
FusionDN	4.8268	<b>3.2174</b>	69.0312	0.3832	0.9328	0.4696	0.7698	0.1007	-0.4246
U2Fusion	4.9148	2.1846	55.1028	0.3055	<b>1.6707</b>	0.4371	<b>0.8347</b>	0.0460	-0.5763
DeFusion	4.6002	2.1969	66.1716	0.4649	1.1152	1.2976	0.8153	3.1185	-0.4319
CDDFuse	4.8309	2.2431	<b>79.4820</b>	0.6100	1.4100	1.3094	0.8285	0.0753	<b>-0.0255</b>
<b>MdAFuse</b>	<b>5.1195</b>	3.0399	56.9286	<b>0.6790</b>	0.9250	<b>1.5666</b>	0.8343	<b>0.0241</b>	-0.4661

### 3.2.4.3 脑影像的关键特征增强实验

图3.20中的前两行显示了在五个周期内从具有脑异常的同一病例中采样的 MRI-SPECT 源影像。SPECT 影像第二行的白色和黑色变化可用于确定病变的位置和病变的严重程度，有助于肿瘤的定位、分割和辅助诊断。第三行是前两行对应的 MRI-SPECT 影像对的融合结果，代表本节提出的方法的着色效果。从左到右的每列表示不同时间的成像结果。从图3.20中可以发现 SPECT 中能量信息的变化和容易判断异常的关键位置，主要体现在黄色轮廓和紫色区域。从左到右，融合结果上的黄色区域数量增加，左下顶叶的黄色区域增加，白色区域的亮度也提高。如果这种亮度代表较高的葡萄糖能量，则很可能在该位置发生糖尿病和高血糖。对于有黑点的区域，黑色块的数量随时间减少，但面积和浓度增加。例如，侧脑室枕角在图3.20 (a) 中只有一个点，在图3.20 (b) 中形成两个小的蓝黑色区域。图3.20 (c) 中的黑蓝色区域的面积变大并向下移动，而图3.20 (d) 中的两个小区域合并，最后在图3.20 (e) 中面积扩大。表明该区存在异常。如果彩色区域代表细胞代谢能力，那么葡萄糖和蛋白质的严重缺乏可能会引起侧脑室枕角相应的疾病。因此，对融合结果进行着色可以增强融合

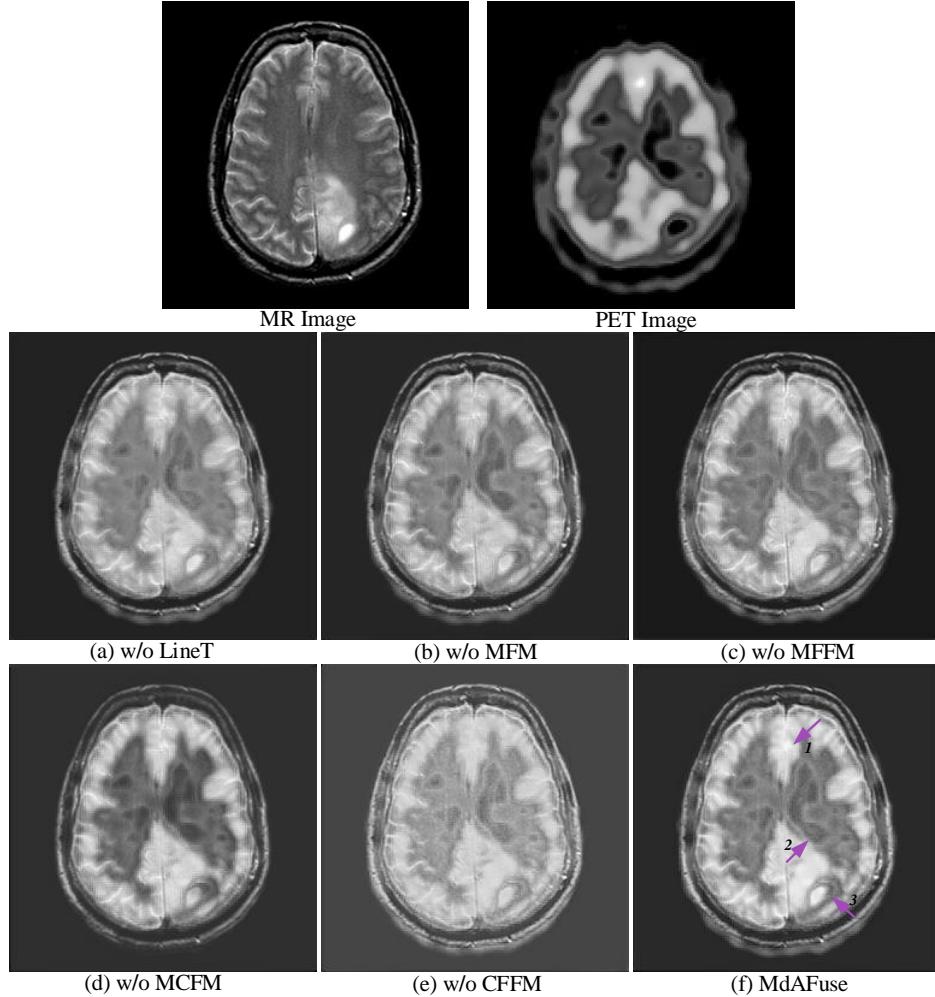


图 3.21: 本节所提出的影像融合方法的消融实验

后的效果，不仅可以揭示异常组织结构位置，还可以揭示能量变化，有助于准确诊断疾病演变。

#### 3.2.4.4 影像融合的消融实验

为了充分说明浅层特征提取、深层特征提取、多尺度特征提取和线性变换融合策略的必要性和有效性，本节进行了五组消融实验，结果如图3.21所示。图3.21中的第一行表示一对多模态（MRI-PET）影像，第二至第三行展示出缺失线性变换操作的融合结果（w/o LinearT）、缺失多尺度特征提取模块（w/o MFM）、缺失多尺度和深层特征提取模块（w/o MFFM）、缺失多尺度和浅层特征提取模块（w/o MCFM）、缺失浅层和深层特征提取模块（w/o CFFM）以及

所提方法的融合策略。在图3.21 (f) 中, 三个箭头包含不同结果中的不同信息。在区域 1 中, 应保留 PET 上的亮度信息。在图3.21 (b) 和图3.21 (c) 中, 该位置的亮度不明显。在图3.21 (a) 和图3.21 (d) 中, 位置模糊, 这可能是由低亮度强度引起的。区域 2 是 PET 影像中的黑色区域, 并且在 MRI 中灰度不均匀。在图3.21 (b) 和图3.21 (d) 中, 它是具有均匀灰度级的区域。在图3.21 (a) 和图3.21 (c) 中, 灰度级以特定梯度布置, 但是边缘出现平滑现象。在图3.21 (f) 中, 效果更好, 并且保留了更多的结构信息。MRI 上箭头 3 所指的区域是一个比扁豆大的小区域, 亮度强而均匀, 与周围的色差有层次关系。在这六种情况下, 本节的 MdAFuse 方法都能反映出该位置的增强效果, 取得了较好的效果。

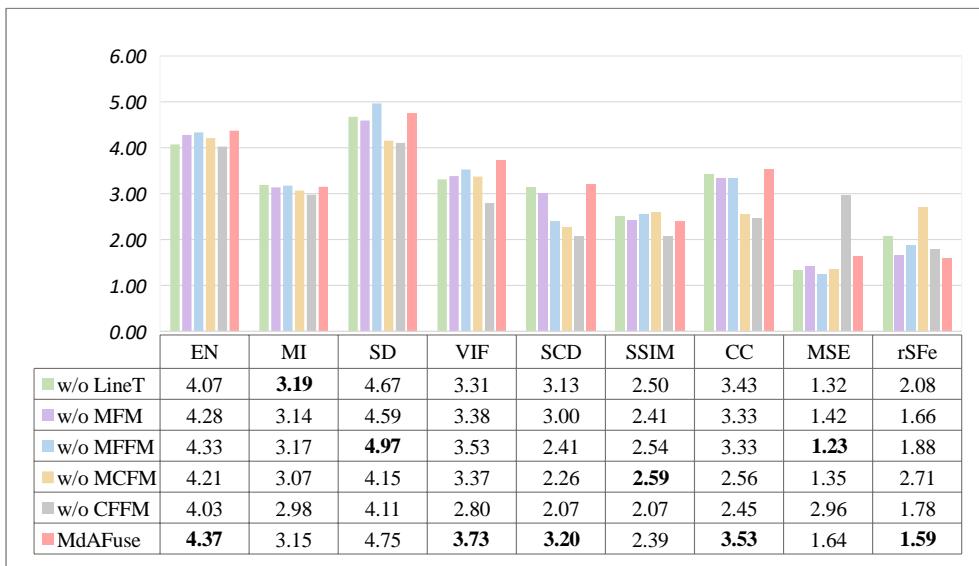


图 3.22: 消融实验结果 (图3.21) 的 9 种质量评价指标显示

图3.22显示了一些用于质量评估的指标值。在图3.22中, 下半部分是原始的六个度量值, 上半部分是条形图显示。为了更好地观察和比较同一图表中不同的度量值, 对每个度量值进行了线性变换, 变换系数如表3.4所示。在图3.22中, 可以观察到本节的方法在 5 个指标中具有最优值, 在 1 个指标中具有次优值。因此, 相比较而言, 本节提出的方法表现出了最佳的性能。因此, 也说明了本节提出的 MdAFuse 方法对于其他融合方法在 9 种客观评估指标上的结果是最好的, 这也证明了通过增加多维特征并使用自适应融合策略构建的网络可以获得更好的融合结果。

### 3.2.5 小结

在这小节中，本节提出了一种新的基于 DL 的融合框架，结合空间特征和通道特征的多维特征。同时，采用深度分离卷积网络，挖掘更丰富的有用信息，可为源影像提供显著特征和丰富细节特征，供后续影像理解和应用分析。采用三种不同的特征提取模块和一种基于各维特征相关性的自适应线性融合机制来保留 MRI 的空间纹理信息和 PET/SPECT 影像的生理代谢信息。此外，本节还提出了一种关键特征增强方法，该方法可以增强同一病例不同时期融合影像的可视化效果，有助于肿瘤定位、分割和疾病跟踪等临床应用。不同的疾病采用不同的常用影像学检查进行评价，多模态医学影像融合也是多种多样的。后续工作将持续研究不同类型的医学融合方法，并将其应用于人工智能医学诊断。

虽然本节的方法在 MRI-PET/SPECT 影像融合中表现出了一定的优势，但仍然存在一些局限性。本节的方法只关注 MRI 和 PET/SPECT 影像的关键特征，而不考虑某种疾病的特异性。此外，对于大脑影像的融合，没有增加对潜在疾病区域的神经科学分析。将联合神经科学与人工智能方法更深入地结合将是今后研究的一个有意义的方向。

## 3.3 本章小结

在3.1节中，本节介绍了一种基于 MS-Info 引导的脑部疾病影像深度特征融合方法，被称为 MsgFusion。该方法通过分析 MRI/CT/PET/SPECT 的关键 MS-Info，获取相应的影像特征，并设计了一个包括 SF-Branch 和 GV-Branch 的双分支网络。SF-Branch 结合了空间域和频域信息，GV-Branch 结合了 HSV 颜色空间的多尺度灰度影像和亮度。这一双网络机制成功提高了 CNN 的泛化能力，充分考虑了频域信息和颜色空间信息的重要性，确保了融合结果的有效性。实验证明，在医学脑影像的处理和分析中，包括 MRI-CT 影像融合、MRI-SPECT 影像融合和 MRI-PET 影像融合等方面，该方法相较于现有方法具有显著优势。通过临床医生的问卷调查评估和统计数据的支持，MsgFusion 的融合效果被证实为最佳。

在3.2小节中，本节提出了一种新的基于 DL 的融合框架，被称为 MdAFuse，采用了空间特征和通道特征的多维度结合。引入深度分离卷积网络，旨在挖掘更丰富的有用信息，为源影像提供显著特征和丰富细节特征，以支持后续影像理解和应用分析。提出了三种不同的特征提取模块和一种基于各维特征相关性

的自适应线性融合机制，以同时保留 MRI 的空间纹理信息和 PET 影像的生理代谢信息。此外，本节提出了一种关键特征增强方法，可提升同一病例不同时期融合影像的可视化效果，有助于肿瘤定位、分割和疾病跟踪等临床应用。考虑到不同疾病使用不同常用影像学检查进行评价，未来将持续研究和拓展多模态医学影像融合方法，以适应人工智能医学诊断的不断发展。尽管本节的方法在 MRI-PET/SPECT 影像融合中表现出一定的优势，但仍存在一些局限性。本节的方法侧重于提取 MRI 和 PET/SPECT 影像的关键特征，而未考虑某种疾病的特异性。此外，在大脑影像的融合中，本节未加入对潜在疾病区域的神经科学分析。

未来，计划扩展框架，将两种以上的医学影像进行融合，并应用于更广泛的临床诊断场景。另外，将更深入地探讨如何将神经科学与人工智能方法更加紧密地结合，以实现更全面和准确的医学影像融合。

## 第4章 阿尔茨海默症的智能辅助分类

当前神经网络驱动的图像分类算法普遍存在仅能捕获局部空间特征的问题，这在一定程度上制约了特征抽取的有效性，进而可能影响到分类性能的整体准确性。特别是在AD早期阶段的诊断分类任务中，面对从EMCI、MCI到LMCI这一连续进展阶段的精细化分类任务，亟需发展更加稳健且智能的分类策略，以实现对疾病进程的精准评估与预测。在这一章中，将介绍一种以影像为导向的新小波卷积单元神经网络。一方面，为了获得非局部的感受野以及避免信息丢失，本章定义了一个新的卷积运算单元。使用该小波卷积单元，进一步设计了一个基于它的细粒度AD多分类网络[217]，在细粒度多分类任务上实现了更高的分类精度。

### 4.1 引言

AD是在老年人群中非常常见的一种疾病。目前，还没有有效的治疗方法可以治愈AD或改变其进展。MCI是介于AD和NC之间的中间阶段。临床研究表明，MCI可分为EMCI和LMCI。EMCI阶段是可逆的，及时发现和干预可以避免AD的发展，而在LMCI阶段及时诊断和治疗可以延缓AD的发展或治愈AD，因此痴呆的早期发现和诊断将成为主要目标。AD的早期准确诊断是一个有意义和挑战性的任务。

卷积神经网络在非医疗影像分类领域取得了显著成效，因此众多研究开始探索将其技术应用于医学影像分类以及智能化辅助疾病的诊断。尤其是在过去的十年间，学者们提出了一系列基于深度学习技术针对AD分类的方法，如表4.1所示。但同时，这些现有方法也暴露出一些局限性问题。

- 当前大多数运用深度学习技术对AD进行分类的方法，通常采用的是为非医学影像设计的标准卷积操作，其局限性在于局部感受野受限。尽管空洞卷积能有效拓宽局部感知范围，但同时也会导致大量信息的丢失问题。与自然图像不同，整个医学影像中的细节特征都比较重要。另外，细粒度AD分类的实现和准确性面临着一定的挑战和瓶颈。为了更有效地提取医学影像的深度特征，需要一种新的非局部感受野卷积运算。

- 现有的大多数 AD 分类方法只能实现粗粒度的分类。只有少数方法，例如 [218, 219]，研究了某些类型的细粒度分类，而细粒度的分类具有更重要的临床意义。此外，AD 分类方法很多，但它们所能实现的分类组合参差不齐。需要一种能够实现 12 种全分类的方法，为该领域的研究提供参考。
- 大多数现有的工作采用 MRI 数据进行 AD 分类。根据临床医学理论，DTI 反映了脑内纤维束间组织结构的连续性，与 AD 密切相关。因此，本章选择 DTI 进行细粒度的 AD 分类。

关于 DTI 数据的选择，本章给予以下详细的解释。神经影像学检查的方式（包括 DTI 扫描、正电子发射断层扫描（PET）扫描、MRI 扫描等）通常用于 AD 的诊断。DTI[220] 也称为基于水分子运动的特殊 MRI 模态。DTI 成像的原理是水分子在梯度场中的扩散改变磁矩。因此，这项创新技术能够清晰展示大脑白质区域中神经束的布局，以揭示人体中枢神经系统的纤维构造细节。脑中纤维束之间组织结构的连续性可以从水中扩散纤维束的成像中推断 [221]。在 AD 患者的研究中，DTI 影像显示为灰质、白质等的 FA（各向异性分数）异常降低和 MD（平均扩散率）值异常增加。因此，与其他影像数据相比，DTI 数据更有利于实现早期的细粒度分类。

为了突破现有 AD 分类方法的上述局限性，本章将小波理论应用于卷积运算中，定义了一种新的卷积单元，并将单尺度和多尺度小波分解的特征融合在一起，以获得非局部的感受野，避免信息丢失。采用新的小波卷积单元（wavelet convolution unit, WCU），本章基于 DTI 数据实现了 12 种所有组合分类，特别是对于细粒度分类，提高了分类精度。总之，本章提出了一个小波卷积单元来构建一个有效的 WCU 网络，用于细粒度和多 AD 分类。使用 WCU 的深度学习方法和 DTI 数据主要证明本章提出的方法实现了更高精度的细粒度 AD 分类。本节工作的三个主要创新点：

- 本章提出了一种新的卷积单元（WCU），将小波变换与传统的卷积函数相结合，以获得非局部的感受野，避免信息丢失。嵌入 WCU 的网络大大提高了卷积神经网络的性能。WCU 是首个将小波分析嵌入到初级的卷积神经网络中的智能辅助阿尔茨海默症分类的工作。
- 本章提出了一种基于 WCU 和脑影像 DTI 数据的 AD 分类网络框架（WCU-Net）。首次实现了所有粗粒度和细粒度的组合分类。大多数现有的 AD 分类方法限于 AD 与 MCI 与 NC 之间的粗粒度分类，少数方法实

现了一种或几种细粒度分类。本章所提出的方法实现了所有八种细粒度分类，其中，首次出现细粒度三分类。该方法为以后的 AD 分类研究提供了参考标准。

- 本章的工作实现的所有 12 种分类获得了高准确率，特别是对 8 种细粒度分类的准确率从 92.5%，92.6%，93.5%，90.9%，81%，无，无，92.6% 提高到 97.30%，95.78%，95.00%，94.00%，97.89%，95.71%，95.07%，93.79%。首先，提出了细粒度的三分类方法，准确率达到 95% 以上。细粒度二分类 EMCI 与 LMCI 相比，分类准确率提高了 16.89%。细粒度四分类的准确率达到 93.79%，比现有方法提高了 1.19%。此外，通过使用来自合适的预训练经典网络模型的初始权重，可以进一步提高准确性。

4.2 节简要介绍了包括小波分析在内的频域分析方法及其在图像处理中的应用，并介绍了基于神经网络的 AD 辅助诊断和分类的研究现状。4.3 节详细介绍了新的小波卷积单元、基于 WCU 的神经网络框架和训练方法。4.4 节描述了基于 WCU-Net 的 AD 细粒度多分类方法在脑 DTI 影像数据上的实现，并进行了一系列的实验和结果分析。最后，4.5 节对本小节的工作进行了总结，并提出了未来的研究方向。

## 4.2 相关工作

### 4.2.1 基于传统智能辅助分类技术

在 AD 的分类任务中，传统的方法大多集中于 AD、NC 和 MCI 的粗分类 [222]。Xiao 等人 [223] 提出了一种新的分类框架，以准确识别 AD 患者。该方法分析了多特征组合相关技术，并利用协方差法对 SVM-RFE 算法进行了改进，结果表明多特征组合方法优于单特征方法。白质（White matter, WM）损伤是 AD 病理级联反应的重要组成部分，在 AD 的诊断中常被忽视 [224]。研究白质最有力的工具是 DTI，它是一种无创的体内成像技术。通过色散特性识别白色物质纤维趋势和损伤程度可以揭示 AD 的结构完整性并描述白质的恶化 [225]。Ben Ahmed 等人 [226] 提出从 DTI 和 sMRI 中提取局部影像衍生的生物标志物，以构建多模态 AD 特征。很少有学者采用常规方法进行细粒度的二分类 [227]。Ashburner 和 Friston [228] 提出了一种基于体素形态测量的传统特征表示的方法。Zhang 等人 [229] 提出了一种基于地标的形态特征方法。

#### 4.2.2 基于新型智能辅助分类技术

患者往往在被确诊患有 AD 时已经发展到疾病的中期或晚期。鉴于此，在利用人工智能进行 AD 的早期诊断时，迫切需要一个既具备高灵敏度又具有高效率的诊断方案 [230–232]。近年来，随着 DL 在各个领域取得的优异成绩，DL 逐渐在医学领域得到广泛应用。CNN 是最常见的深度学习方法，由于其在影像分析和分类领域的成功而受到广泛关注 [233, 234]。尽管如此，深度学习技术在诊断 AD 时面临巨大挑战，这主要归因于医学影像数据缺乏适当的预处理、存在采集误差以及知识库的不足。Aderghal 及其同事 [235] 设计了一种专门针对结构性磁共振成像 (sMRI) 扫描特性的数据增强方法，以优化连续切片样本的训练与分类过程。另一方面，Lei[236] 探索了一种融合 MRI 与 PET 数据的判别式特征学习技术。多模态成像融合手段在增强对 AD 的诊断准确度方面具有明显优势，但其涉及的多步骤影像处理流程常常导致预处理阶段所需时间增加。为克服这一问题，Fang 等人的研究 [237] 中创新性地构建了一个新框架，该框架整合了三种最先进的深度卷积神经网络模型，并应用于多模态影像数据，旨在提高 AD 分类任务的精确度。

在 AD 分类任务中，主要使用 MRI 数据，如 AD 和 NC 的二分类 [238, 239]，以及 AD, MCI 和 NC 的三种成对组合 [240–242, 81, 243]。文献 [244] 描述了一种用于检测和识别 AD 的 MRI 和 PET 影像的自动化和稳健的方法。Bi 等人 [245] 提出了一种新的深度学习技术应用于 AD 的预测。该方法在 AD 与 NC 分类中的准确度比较，MCI 分型为 91.25%。DTI 已间歇性用于识别 AD 的分类任务 [246, 247]，实现了二分类 (AD 与 NC)，准确度分别为 0.885, 0.8235。Ebadi 等人 [248] 研究了关于通过使用集成分类模块执行分类任务来应用大脑连接模式辅助诊断 AD 和 MCI 的潜力。表4.1列出了这些方法和相应的分类精度，表中的每种粗粒度分类的最高准确率和次高准确率分别用红色和蓝色标记 (D-1: AD vs. NC, D-2: AD vs. MCI, D-3: MCI vs. NC, T-1: AD vs. NC vs. MCI)。

随着越来越多的学者开始关注 AD、NC 和 MCI 的粗粒度二分类问题，其方法也在不断改进，并达到了令人满意的准确率。也有学者使用 DL 方法对 AD 进行细粒度分类 [251]，将 MCI 进一步细化为 EMCI 和 LMCI。Liu 等人 [232] 提出了一种通过深度多任务多通道学习框架进行 AD 诊断的联合分类和回归框架。Basaia 等人 [252] 提出了一种基于 CNN 的 MRI 数据 MCI 细粒度分类方

表 4.1: AD 的粗粒度分类的相关工作及其相应的准确率

Methods	Modality	Accuracy (" - " : not classified)			
		D-1	D-2	D-3	T-1
Ben Ahmed et al. [240]	MRI	0.838	0.695	0.621	-
Payan and Montana [222]	MRI	0.954	0.868	0.921	0.895
Aderghal et al. [235]	MRI	0.914	0.695	0.656	-
Yang et al. [241]	MRI	0.91	0.877	0.855	-
Xiao et al. [223]	MRI	0.8571	0.7944	0.8611	0.75
Madusanka et al. [242]	MRI	0.8661	0.8205	0.7896	-
Bi et al. [245]	MRI	0.8915	0.9701	0.926	0.9125
Duc et al. [249]	MRI	-	-	-	0.8715
Xing et al. [238]	MRI	0.92	-	-	-
Pan et al. [81]	MRI	0.89	0.83	0.68	-
Ebrahimi and Luo [239]	MRI	0.9375	-	-	-
Liu et al. [243]	MRI	0.9896	0.9537	0.926	-
Lei et al. [236]	MRI+PET	0.969	-	0.866	-
Vu et al. [244]	MRI+PET	0.988	0.93	0.95	0.9113
Ben Ahmed et al. [226]	MRI+DTI	0.902	0.766	0.794	-
Ebadi et al. [248]	DTI	0.8	0.833	0.7	-
Qu et al. [246]	DTI	0.8235	-	-	-
Lella et al. [247]	DTI	0.885	-	-	-
Bigham et al. [250]	DTI	0.958	0.833	0.833	-
<b>Ours</b>	DTI	<b>0.992</b>	<b>0.9668</b>	<b>0.9616</b>	<b>0.9362</b>

法，实现了 6 种组合的二分类器，分别为 EMAD、LMCI、EMCI 和 NC。Fang 等人 [218] 提出了一种通过二次迁移学习的细粒度分类方法，在 5 种细粒度分类任务上实现了精度的提高。De 和 Chowdhury[219] 提出了 DTI 数据中 AD 与 NC 与 EMCI 与 LMCI 的四分类方法。VoxCNN [77] 网络分别用于 FA, MD 和 EPI 数据的训练，并对 FA 和 MD 的平均数据使用随机森林进行分类。对四种网络模型的输出结果进行融合，提出了一种分层平均融合决策来实现分类任务，分类准确率达到 92.6%，优于现有方法。表4.2列出了这些方法和相应的分类精度，其中，D-1: AD VS. NC, D-4: AD VS. EMCI, D-5: AD VS. LMCI, D-6: NC VS. EMCI, D-7: NC VS. LMCI, D-8: EMCI VS. LMCI, T-2: AD VS. EMCI VS. LMCI, T-3: NC VS. EMCI VS. LMCI, QC: AD VS. NC VS. EMCI VS. LMCI。表中红色值表示最高准确度，蓝色值表示第二高准确度。显然，本章提出的方法对所有列出的分类都达到了最高准确度。

### 4.2.3 基于频域理论分析的技术

频域分析是从频率的角度进行分析，可以发现与时域不同的特性。其中，傅立叶变换和小波变换是两种典型的频域分析方法。小波由小波基函数组成，可以描述信号时间（空间）和频率（尺度）域的局部性质，而傅里叶变换只具有频

表 4.2: AD 的细粒度分类的相关工作及其相应的准确率

Methods	Modality	D-1	D-4	D-5	D-6	D-7	D-8	T-2	T-3	QC
Ashburner and Friston [228]	MRI	0.884	-	-	-	-	-	-	-	0.404
Zhang et al. [229]	MRI	0.870	-	-	-	-	-	-	-	0.431
Liu et al. [232]	MRI	0.937	-	-	-	-	-	-	-	0.518
Basaia et al. [252]	MRI	0.992	0.754	0.859	0.871	0.761	0.751	-	-	-
Wen et al. [251]	MRI	0.91	-	-	-	-	0.81	-	-	-
Prasad et al. [227]	DTI	0.78	-	-	-	-	0.63	-	-	-
De and Chowdhury [219]	DTI	-	-	-	-	-	-	-	-	0.926
Fang et al. [218]	DTI	0.946	0.925	0.926	0.935	0.909	0.808	-	-	-
<b>Ours</b>	DTI	<b>0.9920</b>	<b>0.9730</b>	<b>0.9578</b>	<b>0.9500</b>	<b>0.9400</b>	<b>0.9789</b>	<b>0.9571</b>	<b>0.9507</b>	<b>0.9379</b>

率分析的性质。此外，相较于快速傅立叶变换，小波变换在速度上具有快一个数量级的优势。假设信号长度为  $L$ ，则傅立叶变换和小波变换的计算复杂度分别为  $L \log_2 L$  和  $L$ 。小波分析在图像分析中有着广泛的应用，如图像特征提取、图像压缩、图像水印等。

近年来，有少许论文将小波变换引入 CNN[253–255]。Liu 等人 [253] 应用小波分解来获得高频和低频信息，低频信息被连续分解 4 次。并且高频信息与下一层中的低频信息连续地结合以被更新。最后，对最后的高频和低频信息进行了融合。所有这些操作都出现在卷积之前。Fujieda 等人 [254] 使用小波分析来代替最大池化操作，以优化 U-Net 架构，这是扩张滤波和子采样的推广。但在现有的工作中，小波变换只是作为一个固定系数的滤波器，参与或替代网络结构中的某个过程，例如下采样。他们并没有将小波变换和卷积运算作为一个新的单元，而是将它们结合在一起。

### 4.3 细粒度多分类的辅助诊断网络

在细粒度和多分类任务中，本章提出了一种新的 AD 分类方法。接下来，将介绍本章自定义的小波卷积单元 (Wavelet Convolution Unit, WCU)、细粒度多分类网络 (WCU-Net)，以及 WCU-Net 对 AD 的细粒度多分类问题，分别对

应4.3.1节、4.3.2节与4.3.3节。

### 4.3.1 小波卷积单元的设计思路

小波分析在空间域和频率域都具有良好的局部化特性，并且小波变换具有多分辨率的特点，有利于提取各个分辨率的不同特征。通过二维小波函数分解，将一定尺度的低频部分分解为四个部分：高阶尺度的近似信息和三个方向的详细信息（即，水平、垂直和对角线），如公式（4.1）：

$$\begin{aligned} f_{wav}(x, y) = & \frac{1}{\sqrt{MN}} \sum_m \sum_n W_\varphi(0, m, n) \varphi_{0,m,n}(x, y) + \\ & \frac{1}{\sqrt{MN}} \sum_{j=0}^{\infty} \left( \sum_m \sum_n W_\psi^H(j, m, n) \psi_{j,m,n}^H(x, y) \right. \\ & + \sum_m \sum_n W_\psi^V(j, m, n) \psi_{j,m,n}^V(x, y) + \\ & \left. \sum_m \sum_n W_\psi^D(j, m, n) \psi_{j,m,n}^D(x, y) \right), \end{aligned} \quad (4.1)$$

其中  $M$  和  $N$  表示特定影像的大小。 $\varphi_{0,m,n}(x, y)$  表示尺度函数，计算如式（4.2）所示。 $\psi_{j,m,n}^{Dir}(x, y)$  表示小波基函数，计算如式（4.3）所示。 $j$  表示阶数，它决定了范围和缩小范围。 $Dir$  显示方向，可以是水平、垂直或对角线。 $m$  和  $n$  表示运动的位置。在4.1中， $W_\varphi(0, m, n)$  表示近似系数，其计算如式（4.4）所示。 $W_\psi(j, m, n)$  表示细节系数，其计算如式（4.5）所示。

$$\varphi_{0,m,n}(x, y) = 2^{\frac{j}{2}} \varphi(2^j x - m, 2^j y - n), \quad (4.2)$$

$$\psi_{j,m,n}^{Dir}(x, y) = 2^{\frac{j}{2}} \psi^{Dir}(2^j x - m, 2^j y - n), \quad Dir = \{V, H, D\}, \quad (4.3)$$

$$W_\varphi(0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \varphi_{0,m,n}(x, y), \quad (4.4)$$

$$W_\psi(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \psi_{j,m,n}^{Dir}(x, y). \quad (4.5)$$

小波变换可以将影像分解成不同大小、位置和方向的分量。在小波变换的基础上，本章对它进行了改进，如式（4.6）所示。小波分解后，在近似和细节

系数上使用了  $\xi(\cdot)$  函数。 $\xi(\cdot)$  表示将卷积、BN 和激活三个操作进行组合，如式 (4.7)，它可以使小波分解的系数比获得的整体信息更局部、更完整。在式 (4.7) 中， $Fm$  表示特征图。

$$\begin{aligned}\xi_{wav}(x, y) &= \xi(Approx) + \xi(Detail) \\ &= \xi\left(\frac{1}{\sqrt{MN}} \sum_m \sum_n W_\varphi(0, m, n) \varphi_{0,m,n}(x, y)\right) + \\ &\quad \xi\left(\frac{1}{\sqrt{MN}} \sum_{j=0}^{\infty} \left( \sum_m \sum_n W_\psi^H(j, m, n) \psi_{j,m,n}^H(x, y) + \right.\right. \\ &\quad \left.\sum_m \sum_n W_\psi^V(j, m, n) \psi_{j,m,n}^V(x, y) + \right. \\ &\quad \left.\left. \sum_m \sum_n W_\psi^D(j, m, n) \psi_{j,m,n}^D(x, y) \right)\right),\end{aligned}\quad (4.6)$$

$$\xi(Fm) = \text{ReLU}(BN(Conv(Fm))). \quad (4.7)$$

众所周知，初始卷积具有有限的局部感受野，这往往会导致在深层抽象特征中丢弃一些非局部结构信息。通常采用空洞卷积来扩大局部感受野，信息损失较大。而对于医学影像，非局部结构信息和局部细节信息往往具有重要的临床意义。因此，本章定义了上述新的卷积，将小波变换集成到卷积运算中，以同时获得非局部感受野并避免信息丢失。

图4.1以一幅原始 MRI 为例的实现过程。根据新定义的卷积，首先通过小波变换生成四个系数分量 (A: 原始影像的近似表示，D: 原始影像的对角边缘特征，V: 垂直方向的奇异特征，H: 水平方向的奇异特征)。然后对每个分量进行卷积，可得到四幅特征图。由于小波分解的四个系数表示不同的影像特征，因此在图4.1的原始 MRI 中，将它们堆叠后的区域用黄色框标记。初始卷积的局部感受野也用红框标记以供比较。根据小波变换理论，进一步的频域逐点更新将对小波变换中涉及的所有输入特征产生全局特征。当进行逆小波变换以获得最终的特征图 (见图4.1 (b)) 时，扩展的局部感受野进一步扩散到非局部感受野。并且，可以看到特征图影像非常接近原始影像，没有太多的信息损失。初始卷积的最终特征图如图4.1 (a) 所示，与原始影像有很大的不同，有很多信息丢失。

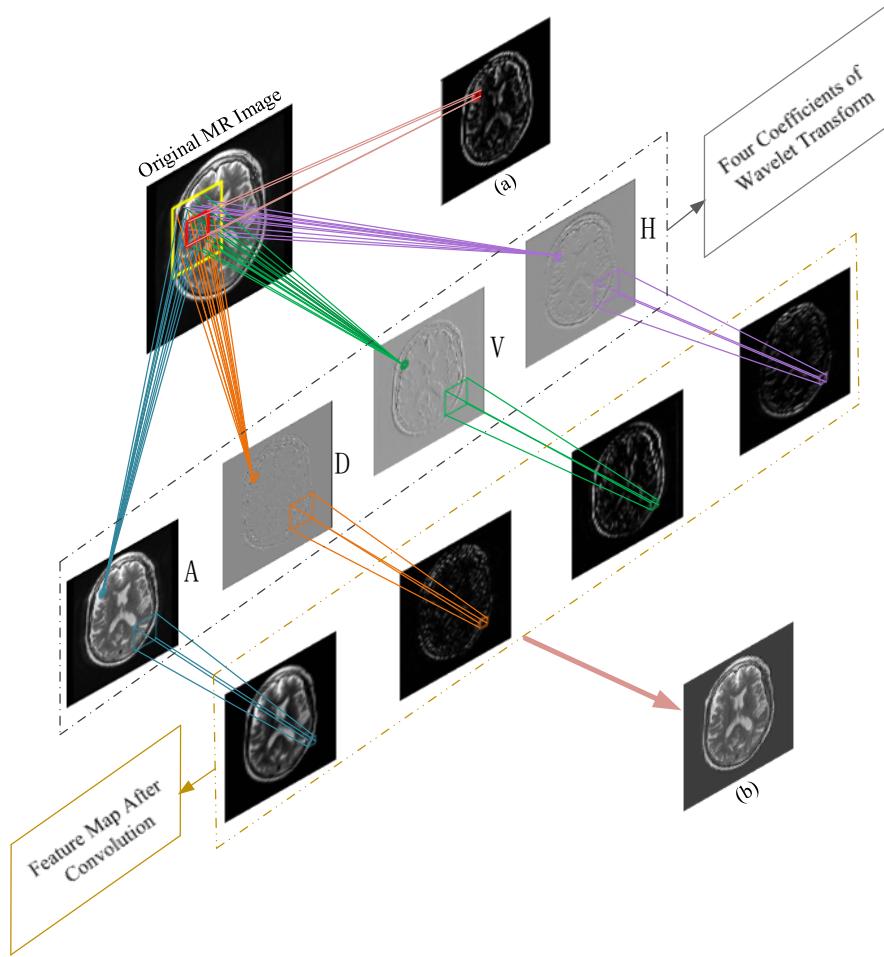


图 4.1: 卷积嵌入小波提取非局部感受野

基于自定义的小波分解函数，本章提出了一个小波卷积单元，如图4.2中黄色方块所示。它由四部分组成：编码器，WCUs，解码器，全连接。WCUs (wavelet convolution unit) 是一个自定义的跨尺度信息提取模块，它利用小波分析和嵌入卷积来提取跨尺度信息。为了获得非局部的感受野，本章对单尺度小波分解系数进行卷积运算。此外，多尺度小波分解可以获得更全面的多尺度特征信息。此外，为了解决奇异点定位不准确的问题，本章将这些跨尺度特征融合到小波卷积单元中。本章的 WCUs 如算法3所示。

首先对影像进行单尺度分解，得到低频近似系数和三个（水平、垂直和对角）高频细节系数，并对每个系数应用  $\xi$  函数，然后对系数进行逆变换。在逆变换之前，先用  $\xi$  函数改变小波变换域的局部系数，可以选择性地扩大局部细节的分

---

**算法 3** Wavelet Convolution Unit

---

```

1: WCU ( $X$ )
2:    $W_s \leftarrow \text{SinScaT} (X, \text{wave});$ 
3:    $W_m \leftarrow \text{MulScaT} (X, \text{wave});$ 
4:    $W \leftarrow \text{Concatenate} ([W_s, W_m], 1);$ 
5:    $X_r \leftarrow \xi_{\text{wav}} (W);$ 
6:   return  $X_r;$ 
7: Function SinScaT ( $X, \text{wave}$ )
8:    $C : \{c_i\}_{i=0}^3 \leftarrow \text{dwt} (X, \text{wave});$ 
9:    $\hat{C} : \{\hat{c}_i\}_{i=0}^3 \leftarrow \xi_{\text{wav}} (c_0, c_1, c_2, c_3);$ 
10:   $\Gamma_s \leftarrow \text{idwt} ((\hat{c}_0, \hat{c}_1, \hat{c}_2, \hat{c}_3), \text{wave});$ 
11:  return  $\Gamma_s;$ 
12: Function MulScaT ( $X, \text{wave}$ )
13:    $\hat{C} \leftarrow \text{wavedec} (X, \text{wave}, 3);$ 
14:    $\Gamma_{0,1} \leftarrow \text{idwt} (\hat{c}_0, \hat{c}_1, \text{wave});$ 
15:    $\Gamma_{3,2} \leftarrow \text{idwt} (\Gamma_{0,1}, \hat{c}_2, \text{wave});$ 
16:    $\Gamma_m \leftarrow \text{idwt} (\Gamma_{3,2}, \hat{c}_3, \text{wave});$ 
17:   return  $\Gamma_m;$ 

```

---

类，减少无用成分。同时将影像分解为三个尺度，根据不同尺度系数进行重构，得到变换后的特征图。最后，将两次重构的结果连接在一起，并再次应用  $\xi$  函数，以获得小波卷积单元的输出结果。在实验测试中，本章选择的小波基函数是 Haar 小波（一阶 Daubechies 小波），这是唯一直接适用于离散二维影像的不连续小波。Haar 小波具有正交性和对称性。此外，在多尺度分解部分，本章设置了 3 尺度分解。在单尺度小波分解的卷积中，设置  $size = 1, stride = 1, padding = 0$  在单尺度和多尺度组合的卷积中，设置  $kernelsize = 3, stride = 1, padding = 1$ 。

### 4.3.2 细粒度多分类的网络结构

CNN 作为深度学习中的一种典型网络结构，它由几层组成，包括但不限于卷积层、池化层、激活层和全连接层。本章方法的框架主要由四部分组成，包括编码，小波卷积单元，解码和全连接层，如图4.2所示。为了有效地处理变长/短序列，使用编码和解码的思想。在编解码阶段，主要采用初级卷积运算来不断调整网络层的深度。在这两个过程中，使用了卷积层、BN 和激活层的组合来提高网络模型的拟合能力。深度 CNN 的卷积层用于提取影像局部区域的尺度/位移不变特征。卷积层的主要好处是在同一特征图中共享权重，这减少了参数并

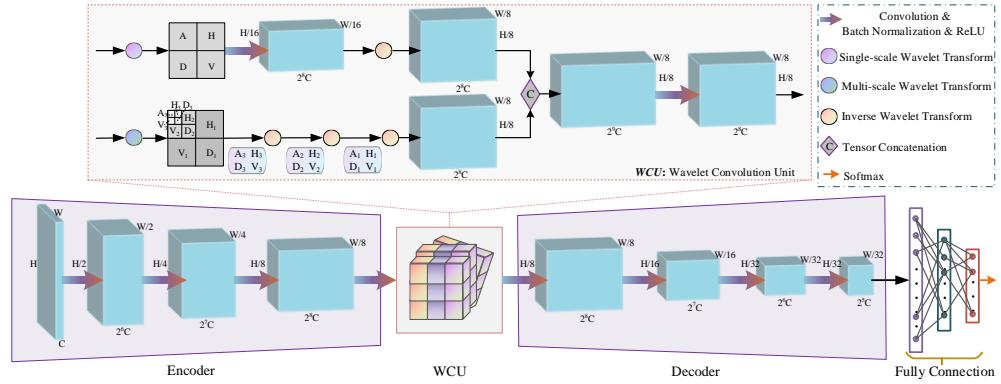


图 4.2: 本节提出的小波卷积网络的框架图

导致模型的简单化。卷积层的不同层中的输入和输出通道的数量如图4.2中蓝色框周围的值所示。在所有卷积层中， $kernelsize = 4, stride = 2, padding = 1$ 。卷积层后面是 BN 层，它本质上是一个归一化的网络层。

接着，将每个隐藏层神经元的输入逐步对接到非线性激活函数，最终归入该函数的值区间的极限饱和区域。将其设计为一般正态分布，其中，0 为平均值，1 为方差值。为了避免梯度消失，非线性函数被设计来处理输入特征。接下来是为模型引入激活层，以学习复杂的表示。为了提高收敛速度，本节中使用的激活函数（除了最终输出）是非线性激活函数 ReLU。CNN 的性能主要取决于层结构和滤波器集合，研究表明网络结构设计是提高 CNN 性能的有效途径 [234]。本节基于小波的局部化和多分辨率特性，设计了一种小波卷积单元。最后一部分是全连接层。在前面的部分中处理的特征图已经变成了一个矢量，它不再是空间定位的。通过增加全连接层，可以发现卷积层中局部特征之间的非线性关系。本章使用了 3 层全连接，并通过 softmax 输出预测的类别。为了提高网络模型的泛化能力，在全连接前后采用了 dropout 方法，并将系数设置为 0.5。在训练过程中，本章使用经典的交叉熵损失函数计算真实的标签和预测输出之间的损失。

### 4.3.3 临床特征驱动型辅助分类

#### 4.3.3.1 临床特征驱动的脑影像数据

众所周知，数据预处理一直是深度学习模型的重要步骤，尤其是在医学领域。DTI 是 MRI 的一种特殊表现形式，是一种能有效观察和追踪白色纤维束

的无创性方法。MRI 主要用于 AD 和其他症状的诊断。一般来说，DTI 能更好地反映脑部疾病的特征，有利于诊断。

在三维视图中，用对称矩阵的概念量化扩散各向异性的信号数据，定义为  $D$ ，如等式 (4.8)。其中， $D_{xx}$ 、 $D_{yy}$ 、 $D_{zz}$  是沿着空间直角坐标系的  $x$ 、 $y$ 、 $z$  轴三个相互正交方向上传播的弥散系数。DTI 表现为一个  $3 \times 3$  对称正定矩阵，它包含三个特征值  $(\lambda_1, \lambda_2, \lambda_3)$  和相关的特征向量  $V = (v_1, v_2, v_3)^T$ 。这 3 个特征向量分别指示了水分子扩散的三个主要方向，而与之对应的特征值则体现了水分子在这三个方向上的弥散强度。

$$D = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} = V^T \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} V. \quad (4.8)$$

在本章的方法中，使用 DTI 的 FA 和 MD 指标 ( $Ind_{FA}, Ind_{MD}$  分别由式 (4.9) 和 (4.10) 表达)，这两个指标可以使用 FSL 软件或脚本对 DTI 数据进行预处理获得。如图4.3中黄色背景的方块所示。MD 影像在第一行，第二行是 FA 影像。FA 的值与白质内部髓鞘的完整性、纤维束的密集程度以及纤维排列的平行性呈正相关关系。因此，FA 影像对脑内白质纤维结构的观察最为清晰，灰质边界明显。 $Ind_{FA}$  的计算公式如 (4.9) 所示。其中， $\bar{\lambda}$  是三个特征值的平均值。MD 反映了分子的整体扩散水平和阻力。MD 仅反映了扩散的整体程度，而不涉及扩散的具体方向。 $Ind_{MD}$  越大，组织中包含的游离水分子越多。 $Ind_{MD}$  的计算公式如 (4.10) 所示。 $Tr(D)$  是  $D$  的迹，它是一个矩阵不变量。

$$Ind_{FA} = \sqrt{\frac{\sum_{i=1}^3 3(\lambda_i - \bar{\lambda})^2}{\sum_{i=1}^3 2\lambda_i^2}}, \quad (4.9)$$

$$Ind_{MD} = Tr(D)/3 = (D_{xx} + D_{yy} + D_{zz}) / 3. \quad (4.10)$$

#### 4.3.3.2 细粒度与多分类的技术路线

本节建立了一个新的框架，通过 WCU-Net 实现 AD 的细粒度和多分类。对于二、三和四分类框架通用，在实验装置中，只有它们的输入和输出不同。如图4.3所示，该框架包括三个部分：数据输入，编码器、WCU、解码器、FC 和激活函数的网络架构，以及类别预测的输出。在第一部分中，对 DTI 数据进行

了预处理。FSL (<https://fsl.fmrib.ox.ac.UK/fsl/fslwiki/Fsl Installation/Linux>) 工具对原始 DTI 数据进行 b0 影像提取、剥脑、涡流校正、张量计算等预处理操作。从所获得的张量指数中选择 FA 和 MD 作为输入的数据。在第二部分中，依次输入 4 个类别 (AD, NC, EMCI, LMCI) 的 FA 和 MD，并分别通过编码器，WCU，解码器，FC 和激活函数 (Softmax)。在这一部分中，主要对 WCU 中的流程进行了扩展和展示。WCU 采用自定义的小波卷积单元，结合单尺度小波分解和多尺度小波分解的特点，获得完整的跨尺度特性。最后一部分是网络的预测输出，其将概率分布转换为标签数据。最终的预测类别取决于具有最高概率的类别。

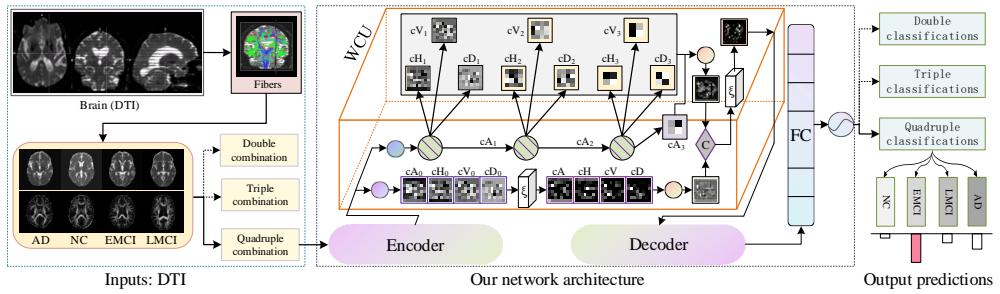


图 4.3: WCU-Net 应用于细粒度多分类任务中的具体步骤

CNN 是深度智能系统中最常用于检测 AD 的网络，一些研究倾向于设计 CNN 的结构 [231]。经典的网络结构，如 ResNet, DenseNet, SqueezeNet, Inception, AlexNet, VGGNet 和 GoogLeNet，已成功应用于影像分类任务 [256]，这有助于权重初始化过程。在本节中，利用这 7 个经典的网络模型初始化本章提出的网络结构权重。表4.3列出了 WCU-Net 与预训练 CNN 的分类准确度，其中，D-1: AD vs. NC, D-4: AD vs. EMCI, D-5: AD vs. LMCI, D-6: NC vs. EMCI, D-7: NC vs. LMCI, D-8: EMCI vs. LMCI, T-2: AD vs. EMCI vs. LMCI, T-3: NC vs. EMCI vs. LMCI, QC: AD vs. NC vs. EMCI vs. LMCI)，2-9 列的每一列表示不同的网络模型参与的应用 (Net\_1: ResNet101, Net\_2: DenseNet161, Net\_3: SqueezeNet1\_1, Net\_4: Inception\_v3, Net\_5: AlexNet, Net\_6: VGG13, Net\_7: GoogLeNet, Net\_8:w/o pretrained CNNs)。在权重初始化过程中是否采用预训练 CNNs 模型的分类准确率也列在最后一列以供比较。从表4.3可以看出，每个预训练模型都能在一定程度上提高 2-4 种组合分类的准确率，在第 2-8 列用红色或蓝色值标记。特别是对于细粒度的 NC vs. EMCI 和

NC vs. LMCI 二分类，预训练的效果明显。总之，九种组合分类的最高准确度值分别达到 0.9920, 0.9730, 0.9611, 0.9870, 0.9611, 0.9900, 0.9636 和 0.9379, 其中 AD 与 NC, AD 与 EMCI 和 AD 与 LMCI 与 EMCI 与 NC 的最佳分类准确度由 WCU-Net 得到，无预训练 CNNs 模型参与。而在本节中，所提出的方法则是使用未经预训练网络初始化参数的 WCU-Net，并获得的结果与现有的 AD 分类方法进行比较。

表 4.3: 权重初始化过程中是否采用预训练模型的精度比较

Class	Net_1	Net_2	Net_3	Net_4	Net_5	Net_6	Net_7	Net_8
D-1	0.9840	0.9890	0.9790	0.9710	0.9701	0.9850	0.9890	<b>0.9920</b>
D-4	0.9630	0.9550	0.9690	0.9670	0.9500	0.9480	0.9680	<b>0.9730</b>
D-5	0.9533	0.9556	0.9533	0.9511	0.9478	<b>0.9611</b>	0.9522	0.9578
D-6	<b>0.9750</b>	<b>0.9720</b>	<b>0.9650</b>	<b>0.9710</b>	<b>0.9690</b>	<b>0.9870</b>	<b>0.9700</b>	0.9500
D-7	<b>0.9611</b>	<b>0.9533</b>	<b>0.9456</b>	<b>0.9489</b>	<b>0.9567</b>	0.9089	<b>0.9590</b>	0.9400
D-8	<b>0.9800</b>	<b>0.9800</b>	0.9722	<b>0.9833</b>	0.9644	0.9778	<b>0.9900</b>	0.9789
T-2	0.9350	<b>0.9586</b>	0.9490	0.9521	<b>0.9600</b>	<b>0.9586</b>	<b>0.9636</b>	0.9571
T-3	0.9257	0.9243	0.9464	<b>0.9833</b>	0.9193	0.9407	0.9429	0.9507
QC	0.9353	0.9168	0.9284	0.9068	0.9358	0.9337	0.9053	<b>0.9379</b>

## 4.4 细粒度多分类的实验

### 4.4.1 实验验证的数据说明

本节中使用的数据来自 ADNI ([www.adni-info.org](http://www.adni-info.org))。由于每个类别的数据量不相等，并且数据的总数不足以训练深度学习模型，因此使用数据扩大来增加每个少数类别的样本数量。另外，医学影像可视化后是灰度影像，为了不对影像造成较大的改变，数据增强方法主要是通过加入随机噪声和亮度进行扩大。为了增加数据的大小，本章将使用数据增强来扩展所获得的数据。数据增强的方法与文献 [218] 一致，数据增强后的情况见表4.4。

DTI 是 MRI 的一种特殊形式，但其获取比 MRI 困难。DTI 可以捕捉脑内的白色物质束和灰质束，并通过测量活体组织内水分子的扩散来工作。由于扩散张量是一个对称的  $3 \times 3$  矩阵，它可以用它的特征值 ( $\lambda_i$ ) 和特征向量 ( $V_i$ ) 来描述，然后使用特征值和特征向量来处理标量指数。两个主要的扩散指数 (FA

表 4.4: 增强前后的数据集详情

Diagnosis	AD	EMCI	LMCI	NC
Age	75.23	73.95	73.45	74.52
Gender(M/F)	99 / 54	225 / 138	108 / 59	110 / 109
Subjects	153	363	167	219
MMSE	18-27	24-30	24-30	25-30
Train data	9273	9748	10154	10113
Valid data	2100	2088	2100	2200
Test data	500	400	500	500

和 MD) 是基于本征值和代表的扩散过程的幅度。年龄是 AD 的主要危险因素，因此选择年龄在 73-76 岁之间且认知功能评分差异显著的人群 (AD: 18-27, EMCI/LMCI: 24-30, NC: 25-30) 进行研究分析。所下载的数据集来自 902 个样本，分为四类：153 个 AD 患者，167 个 LMCI，363 个 EMCI 和 219 个 NC。选择 DTI 数据集作为模型的训练集、验证集和测试集，数量如表4.4所示。

#### 4.4.2 实验验证的参数说明

在网络训练过程中，影像的大小为  $128 \times 128$ 。在训练过程中，初始学习率为 0.001，40 个 epoch，批量大小为 32，使用批量归一化。优化函数是 SGD，动量的值为 0.85。此外，实施了一种动态调整的学习率策略，每隔 8 个训练周期将学习率降低至初始值的 10%。本章使用 Python 3.8.12 和 Pytorch 1.10.2 版本，CUDA 版本为 10.2，系统为 Tesla T4 16 G RAM 64 GNU/Linux x86。

#### 4.4.3 早期细粒度分类应用

临床研究表明，AD 的细粒度分型具有重要意义。关于 AD 的分类有很多研究，然而，大多数现有的方法局限于粗粒度的二分类，即，AD vs. NC、AD vs. MCI 和 MCI vs. NC。一些方法实现了粗粒度的三分类，即，AD 与 NC 与 MCI 分类。近年来，一些学者对 MCI 进行了进一步的细粒度分类。早期很少有学者利用 DTI 数据做细粒度的二分类 [227, 251]，其较高的准确率为 0.81。一些学者使用 MRI 数据对 AD 进行了细粒度的四分类。Ashburner 和 Friston[228]、Zhang 等人 [229] 和 Liu 等人 [232] 提出应用于 AD 细粒度四分类并使用 MRI 数据的方法，但准确性较低。也有一些方法使用 MRI 数据应用于三分类和四分类，可以获得更好的性能。例如，在 [257] 中，三分类准确率可以达到 68.8%，四分类

准确率可以达到 59.1%。然而，三分类是 AD vs NC vs (EMCI+LMEC)，这与传统的粗粒度三分类相似。在文献 [258] 中四分类的准确率可达 61.9%。Basaia 等 [252] 提出了基于 CNN 的 MRI 数据 MCI 细粒度分类方法，实现了 AD、LMCI、EMCI、NC 六种组合双分类。Fang 等人 [218] 提出了一种通过二次迁移学习的细粒度方法，提高了 5 种细粒度二分类的准确性。目前，De 和 Chowdhury[219] 实现了细粒度的四分类，精度为 0.926。

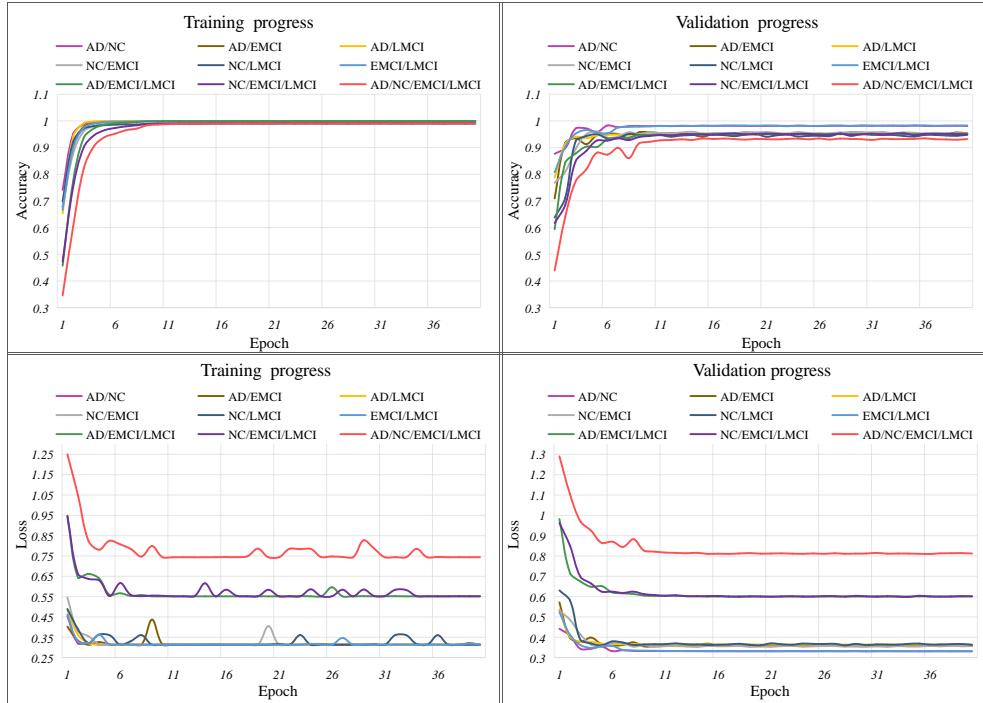


图 4.4: 准确度曲线和损失曲线

在本节中，为了便于尽可能准确地诊断 AD 分类，展开了 12 种组合分类，包括 4 种粗粒度分类和 8 种细粒度分类，如表4.1和4.2的所示。本章提出的方法是首次出现两种细粒度的三分类 (AD vs. LMCI vs. EMCI 和 NC vs. LMCI vs. EMCI)。表4.2列出了本章提出的方法对 AD 与 NC 分类以及 8 种粒度多分类的准确性。为便于比较，表4.2还列出了现有细粒度分类方法的准确性。可以观察到，对于实现的 6 种二分类，本章提出的方法的准确性比现有的方法 [252, 218] 要好得多。对于细粒度的四分类，本章提出的方法准确率为 0.9379，也比 [219] 提高了 1.19%，首次实现了细粒度的三分类，其中 AD vs. EMCI vs. LMCI、NC vs. EMCI vs. LMCI 的准确度分别达到 0.9571 和 0.9507。

图4.4中的第一行显示了在没有预训练的网络模型作为初始化权重参数的情况下，所提出的方法的预测精度曲线分布，并且在训练阶段和验证阶段之间的分布存在一定的偏差。在训练过程中，准确率曲线没有波动，并且随着训练时期的增加，在不同分类组合中的准确率更接近于1。在验证过程中，1-8轮出现了小范围的振荡，但总体上升趋势是正确的。在验证过程中，参数不断更新。随着轮数的增加，AD vs. NC 和 EMCI vs. LMCI 分类的准确率接近1，四分类的准确率稍低，但也在0.9-1之间。在其他数据组合分类的情况下，精度也稳定在第11轮。图4.4中的第二行显示了训练和验证阶段的损失曲线。二分类的损失较小，而四分类的损失相对较大。而训练和验证过程中的损失明显减少，并趋于稳定。因此，表明本节提出的网络模型具有良好的泛化性能。

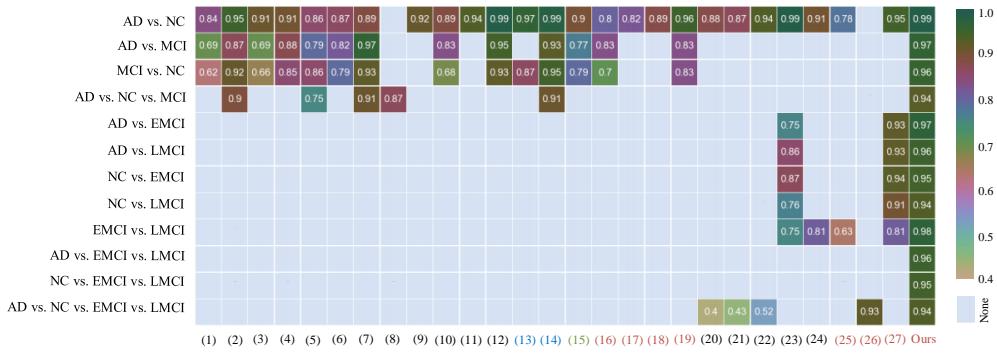


图 4.5: 各种粗粒度和细粒度 AD 分类方法精度比较的热力图

#### 4.4.4 多类别分类应用分析

图4.5显示了本文所涉及的所有粗粒度和细粒度AD分类的准确性比较的热力图。图中的小矩阵单元颜色越深，其相应方法的准确度越高。淡蓝色单元格表示没有相应的分类工作。大面积的浅蓝色单元表明，大多数现有的方法只关注于少数组合分类。但本章提出的方法实现了所有12种分类，对于所有细粒度分类都具有较高精度，为以后该领域的研究工作提供了完整的参考。图中横轴上颜色相同的数字表示这些方法使用的数据模态是相同的。红色、蓝色、绿色和黑色分别对应于DTI、MRI+PET、MRI+DTI和MRI。(1)-(27)列分别指下列参考文献中提出的不同方法：Ben Ahmed et al. [240], Payan and Montana [222], Aderghal et al. [235], Yang et al. [241], Xiao et al. [223], Madusanka et al. [242], Bi et al. [245], Duc et al. [249], Xing et al. [238], Pan et al. [81], Ebrahimi and Luo [239], Liu et al. [243], Lei et al. [236], Vu et al. [244], Ben Ahmed et al.

[226], Ebadi et al. [248], Qu et al. [246], Lella et al. [247], Bigham et al. [250], Ashburner and Friston [228], Zhang et al. [229], Liu et al. [232], Basaia et al. [252], Wen et al. [251], Prasad et al. [227], De and Chowdhury [219], Fang et al. [218]。从这张热图中，可以得出以下结论：

- 大多数方法只能实现粗粒度的分类。一些方法实现了一种或几种细粒度分类。本章的方法实现了所有 12 种的组合分类，而且只有本章的方法实现了细粒度的三分类。
- 大多数方法对于不同类别组合的精度不稳定，甚至存在许多精度小于 0.70。本章的方法实现了稳定的分类且精度高于 0.935 的有 12 种组合分类。

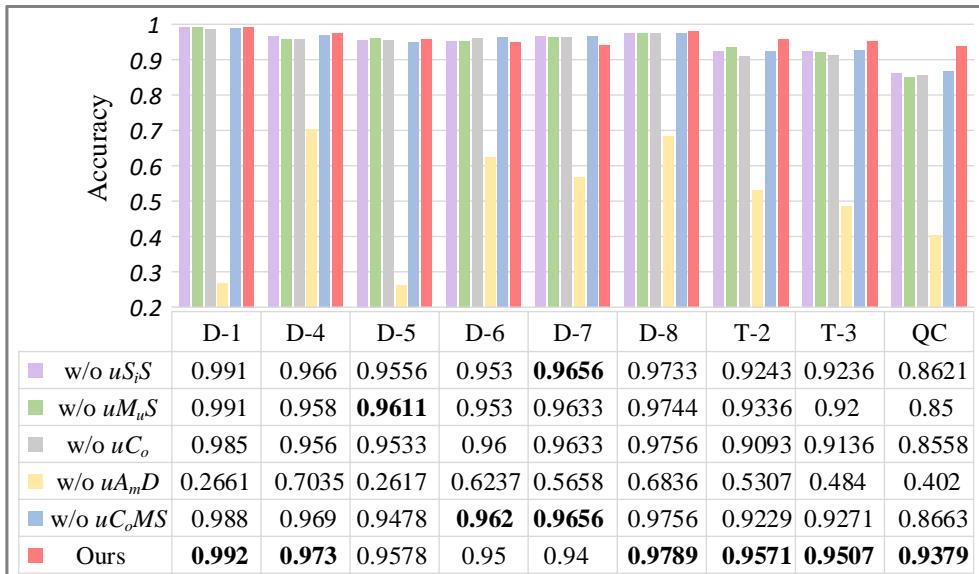


图 4.6: 本章提出的方法的消融实验

关于 AD 分类的研究工作似乎很多，但他们通常专注于不同的数据模态和类别的组合，各种方法之间的可比性不是很强。本章方法实现了所有 12 种粗粒度和细粒度分类，具有较高的准确率，为后续基于 DTI 数据的 AD 分类研究提供了一个合适的参考标准。不过，有一点需要特别说明。在本章中，主要集中在 DL 方法的准确性，但忽略了神经科学的部分，潜在的目标对的 ROI 可以关联到 AD 的不同阶段。实际上，这对于基于神经科学的疾病研究和基于深度学习方法的智能辅助诊断的可解释性非常重要。在未来的研究中，将神经科学与

人工智能方法更深入地结合将是一个有意义的研究方向。

#### 4.4.5 细粒度多分类的消融实验

如第4.3节所述，本章的方法的最佳性能主要取决于单尺度小波分解( $uS_iS$ )、多尺度小波变换( $uM_uS$ )、 $\xi$ 函数运算( $uC_o$ )和数据扩增 $uA_mD$ 。为了进一步证明这四个部分的有效性，本章设计了五组方法进行消融实验，即，缺少 $uS_iS$ 、缺少 $uM_uS$ 、缺少 $uC_o$ 、缺少 $uA_mD$ 和缺少 $uC_oMS$ (缺少 $uM_uS$ 和 $uC_o$ )。所有的实验结果都显示在图4.6中，与本章的方法进行比较。表格中的每行表示的是不同的方法，每列代表不同的组合分类，各组合同章节4.3.3.2定义的相同。从图4.6可以发现，数据扩增至至关重要，其他部分对于细粒度分类也很重要。特别是在细粒度的四分类中，各组件的结合使用，分类精度有了明显的提高。

### 4.5 本章小结

本章提出了一种新的网络 WCU-Net，它在小波卷积单元中结合了单尺度和多尺度变换。该策略不仅获得了非局部感受野，而且实现了跨尺度信息融合。然后将 WCU-Net 成功地应用于 AD 细粒度和多分类中，并达到 SOTA 精度。细粒度分类对于认知障碍的准确诊断和正确治疗具有重要意义。细粒度分类特别是细粒度的四分类准确率还有待进一步提高。后续将会尝试在 AD 的更多分期中应用 WCU-Net，以及基于医学影像的不同性质疾病的分类，并进行更多研究。

## 第 5 章 帕金森疾病的智能辅助预测

PD 的发生机制因素除复杂性外，还涵盖了遗传背景、环境暴露及生活习惯等多个层次的交织影响。其中， $\alpha$ -突触核蛋白聚合现象是其生物学标识的核心组成部分。现有多模态影像整合技术在探究医学信息与影像特征深层次相关性方面存在不充分的问题，这在一定程度上限制了其在脑部疾病诊断中的表现力。为了提升对 PD 等多种脑疾病的精确诊断能力，亟需引入跨模态数据集成分析的方法。机器学习（ML）技术的进步为智能辅助诊断与健康评估提供了智能的决策支持，有望应对特殊情况下的误诊或难以确诊的挑战。第5.2节综述了 ML 的智能辅助 PD 分类预测方法 [91]。ML 模型通过挖掘大量数据中的隐藏规律，为早期辅助诊断提供了新的机会。在第5.3节中，提出了一种跨模态数据融合的 ML 方法（CMFP），它是通过综合 DTI 和临床数据的一种 PD 早期进展预测的新方法，可显著提高单模态预测的精确度。

### 5.1 引言

PD 是一种常见的神经系统疾病，主要影响人类的神经系统和运动控制 [259]，表现为静止时的颤抖、动作缓慢、身体活动受限以及平衡失调 [260]。PD 随时间缓慢恶化，诊断依赖于患者的临床症状；肢体功能障碍和非运动症状不仅损害生活品质，还与较高死亡率相关，故早期预测和及时治疗至关重要 [261]。至今为止，尚未发现能够准确反映 PD 的疾病进程的临床数据或生物指标 [260]。而且，对临床效果的评估往往需要较长的时间，且其结果易受到患者当时健康状况的影响而发生变化 [260]。标准的 MRI 能够展现清晰的组织差异，但在诊断疑似 PD 患者方面，它的主要价值在于帮助排除其他脑部疾病，而不是直接用来确认是否为 PD[262]。PD 是一个异质性疾病，根据起病的年龄、临床表现及进展的快慢可以被分成不同的亚型 [263, 264]。PD 疾病的发展会随着时间的演变而发生不同的变化，一些患者病情会保持稳定，而一些则进展为帕金森病痴呆（PDD）。因此，有必要预测 PD 疾病的进展。

研究发现，白质微观结构的变化在皮层神经元显著退化前就已发生，且这种变化未必与灰质体积减少相关 [260]。通过使用扩散张量成像（DTI）技术，研究者能够监测到帕金森病患者的大脑白质损伤，并发现这些损伤的程度与患者

接下来一年中运动能力和认知功能的改变有着紧密的联系 [260, 265]。最新研究运用 DTI 技术揭示，在早期 PD 患者中，全脑体素分析指出胼胝体的分数各向异性 (FA) 和平均弥散度 (MD) 出现了明显的变化，特别是在膝部和压部区域更为显著 [266]。研究还发现，胼胝体 FA 值的降低与大脑广泛皮质及皮质下区域的 FA 和 MD 值降低存在相关性 [266]。进行多元回归分析后，结果表明在研究的起始点及两年后的跟进中，患者的运动刚性程度与胼胝体的 FA 值之间存在显著的负向关联，且这种关联在胼胝体前部区域最为突出。这些发现提示研究者们，在 PD 的早期阶段，胼胝体的微观结构变异可能是一个有用的生物标志物，用于监测运动刚性症状及其疾病进程的变化 [266]。

ML，特别是与数据挖掘技术相结合时，关注的是一种算法的发展，它可以从已知数据中学习规律形成模型，然后将模型应用到未知数据中进行预测 [278]。因此，ML 被广泛应用于 PD 及其进展的预测，以提高其预测 [279–281] 的性能。目前的统计模型预测 PD 的临床进展具有挑战性。过去来自横断面试验的单变量纵向或多变量分析在预测单个结果或单个时刻的能力有限 [282]。因此，跨模态及混合模型的设计将有利于 PD 进展的研究。

## 5.2 基于机器学习的智能辅助预测技术

本节将 PD 的预测归纳为静态与动态 (进展) 预测 [91]。静态预测方法通常依赖于某一特定时间点的临床基线数据来进行分类预测。这种方法只考虑了单一时刻的数据，而没有将随时间变化的数据纳入考量。多数相关研究采用这种方式来尝试预测疾病发展情况，如 [267, 98–100, 268, 269, 271–277]。另外，动态指代涉及到在一段时间内收集的多个连续的临床状态数据，用以进行分类预测。这种动态预测方式通过分析患者在不同时间点的数据变化，来更好地预测疾病的进展情况，例如 PD 的发展趋势。这种方法比静态预测更能反映疾病随时间的演变过程，因而对于制定长期的治疗计划和管理策略尤为重要。如文献 [112, 113, 118, 283–289]。

本节在 PubMed 检索平台通过搜索 PD 的静/动态预测并统计出了近 33 年 (1990—2022 年) 相关研究成果，如图 5.1 所示。图 5.1 显示，在过去 33 年间，PD 的预测研究主要集中在静态预测领域，且自 2010 年以来研究数量呈现持续上升趋势，至 2022 年已增长至 648 篇。尽管动态预测的相关成果较少，但自 2016 年起其增长速度显著加快。这些研究表明，借助智能辅助技术可以有效促进临

表 5.1: 基于机器学习技术在 PD 某一时刻智能辅助预测的相关工作

文献	方法	数据及来源	ACC(%)	AUC(%)
文献 [267]	Transfer Learning (AlexNet)	PPMI 中的 MRI	88.90	-
文献 [98]	Lenet 和 Alexnet	PPMI 中的 FP-CIT SPECT	94.10	98.40
文献 [99]	Transfer Learning (Inception V3)	PPMI 中的 SPECT	-	87.00
文献 [100]	SVM 和 KNN 的集 成分类	PPMI 中的 SPECT 和多种异质生物标志 物 (CSF, Plasma, RNA 和 Serum)	96.00	-
文献 [268]	RF、XGBT 和 CatBoost 的集成分 类	UCI 库中的声学特征	86.25	-
文献 [269]	KNN, MLP, OPF 和 SVM	文献 [270] 中的发音 与语言数据	94.55	87.00
文献 [271]	GBT, XGBT, Bagging 和 ExtraTree	UCI 库中的声学特征	82.35	-
文献 [272]	DT, LR 和 KNN 的集成分类	UCI 库中的声学特征	84.59	-
文献 [273]	文献 [273] 中提出 的一种新颖的深度 学习方法	PPMI 中的 REM, CSF、嗅觉丧失和 SPECT 数据	96.45	-
文献 [274]	NB, SVM, Boosting 和 RF	PPMI 中的睡眠行为 障碍、CSF、嗅觉丧 失和 DAT 数据	96.40	98.88
文献 [275]	DNN	UCI 库中的声学特征	81.67	-
文献 [276]	文献 [276] 中提出 的一种新颖的深度 学习方法 (CROWD)	UCI 库中的声学特征	96.00	-
文献 [277]	CNN	PPMI 中的 MRI, SPECT 和 CSF	93.33	-

表 5.2: 基于机器学习在 PD 智能辅助进展预测的相关工作

文献	方法	数据及来源	ACC(%)	AUC(%)
文献 [112]	BN	PPMI 中的临床、分子和基因数据	-	-
文献 [113]	RF	PPMI 中的受试者人口统计学、疾病特征、CSF 和 DAT 数据	-	-
文献 [118]	DBN, SVR 和 SOM	基于真实世界 PD 数据的 Total-UPDRS 和 Motor-UPDRS	-	-
文献 [283]	CERNNE	PPMI 和 ADNI 中的 SNP 和 fMRI	88.60	-
文献 [284]	SVM	PPMI 中的临床评分 (睡眠、嗅觉)、MRI 和 DTI	92.08	94.44
文献 [285]	CNN	英国脑库中的 MRI	96.80	99.50
文献 [286]	CNN	PPMI 中的 DAT, SPECT 和非成像临床指标 (UPDRS_III 评分)	83.00	90.00
文献 [287]	KNN, DT, RF, NB, SVM, K-Means 和 GMM	PhysioNet 网站上的步态周期 vGRFs	90.00	-
文献 [288]	KNN, DT, RF, SVM 和 GMM	PhysioNet 网站上的步态周期 CDTW	97.00	-
文献 [289]	多源集成学习与 CNN	mPower 数据集文献 [210] 中的敲击、行走、声音和记忆数据	82.00	-

床诊断和疾病治疗。动态预测在医疗健康领域的研究与实践正在逐步扩展，并日益受到重视。

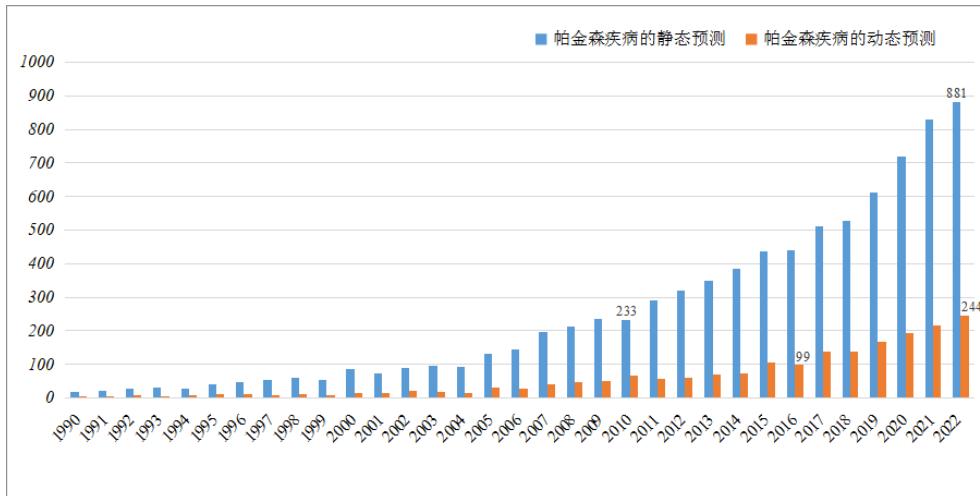


图 5.1: PD 的静态及动态智能辅助预测发文量统计

数据的丰富性提高、计算技术的进步以及神经网络设计的优化共同促进了深度学习技术的发展和普及。作为 ML 领域中的先进技术，深度学习通过构建更加复杂的神经网络来增强模型的表现力和准确度。在 PD 的预测领域，深度学习方法同样取得了令人瞩目的成效 [267, 274, 270, 290, 276, 277, 291, 292, 286]。表 5.1 和表 5.2 汇总了近年来运用机器学习技术对帕金森病 (PD) 进行静态和动态评估的多项研究成果，涵盖了所采用的不同 ML 技术、数据来源以及预测的准确率 (ACC 和 AUC)。其中，大多数研究倾向于使用传统 ML 技术，而采用深度学习技术的研究相对少见。数据主要来自 PPMI，并且大多数医学影像数据采用的是 SPECT。在这些研究中，仅使用单一数据类型进行预测的成功率通常较低，而使用步态周期数据进行预测的成功率较高。同时，将脑脊液 (CSF) 等临床评分与医学影像数据相结合，能够显著提高预测的性能。

当前，多模态分类预测模型在帕金森病 (PD) 的研究中正逐渐受到重视并展现出发展潜力。尽管如此，这一领域仍面临诸多挑战，特别是在收集和整合 PD 患者全面数据方面存在限制。因此，开发无需人工干预即可从多种模态中自动提取特征的早期 PD 分类预测模型显得尤为重要。这样的模型能够利用更为丰富的数据资源，从而构建更精确的分类预测网络。此外，表 5.2 的分析表明，目前只有少数研究尝试结合基因数据与医学影像数据来探索它们与 PD 之间的

关联。这种跨模态的数据融合方法有望进一步加深学者对 PD 遗传背景及其与疾病发展之间联系的理解。

## 5.3 临床信息主导的智能辅助进展预测

随着人口老龄化的增加，帕金森病患者也在迅速增加，因此，对评估和跟踪帕金森病进展的动态监测变得越来越重要。为了找到一种更为高效的帕金森病进展预测方法，本节通过分析单模态 (DTI、DAT 及临床) 特征在 PD 疾病 5 年后进展预测的临床价值，并提出一种跨模态数据融合 (DIT& 临床、DAT& 临床) 对帕金森病的进展进行预测的方法 (Cross-Modal Fusion prediction model, CMFP)。该方法的步骤主要是数据准备，建模，预测。数据涉及到临床，DTI 及 DAT 这 3 种模态，不同模态数据类型均使用 Lasso 方法进行特征筛选。各单模态采用 AdaBoost 进行分类，将结果应用于新的融合策略 CMF 中得到新的模型。最后使用新的模型进行预测。CMFP 在 PD 进展上的预测获得的 AUC 为 0.7791。与单独的临床数据、DTI 数据和 DAT 数据预测所获得的 AUC 分别提高了 0.2448、0.3078、0.327。测试临床与 DTI 结合预测相对于单独使用临床数据预测的结果具有统计学显著性， $p$  值为  $9.183e-4$ 。该方法还确定了与 PD 相关的关键的脑区和重要的临床指标。本节的研究结果表明，CMFP 是有效的，其有助于改善单模态数据预测性能低的不足，提高 PD 进展预测的准确性。并能确定 PD 的潜在生物标志物信息，为了解 PD 提供了新的途径，值得进一步研究。

### 5.3.1 智能辅助预测的数据与方案

#### 5.3.1.1 进展预测的数据说明

本项研究利用了帕金森病进展标志物倡议 (PPMI) 的丰富资源，该资源汇集了多个研究中心的 PD 患者纵向数据。所有参与 PPMI 的研究站点均已获得相应伦理审批，且所有参与者在加入前都签署了知情同意书。研究中，本节采用 HYS 量表会考虑运动功能的退化、生活质量的降低以及多巴胺水平的减少等多个因素，并据此将 PD 分为五个不同的阶段 (1 至 5 期)。研究入选标准包括：至少有 5 年的纵向随访记录且 HYS 评分完整的 PD 患者；基线数据中含有完整的 MRI 影像资料，包括 3D T1WI 和 DTI 扫描；使用 3T Siemens MRI 进行扫描。筛选条件包括临床信息不完整、成像质量不佳以及图像处理出现失误等情况。

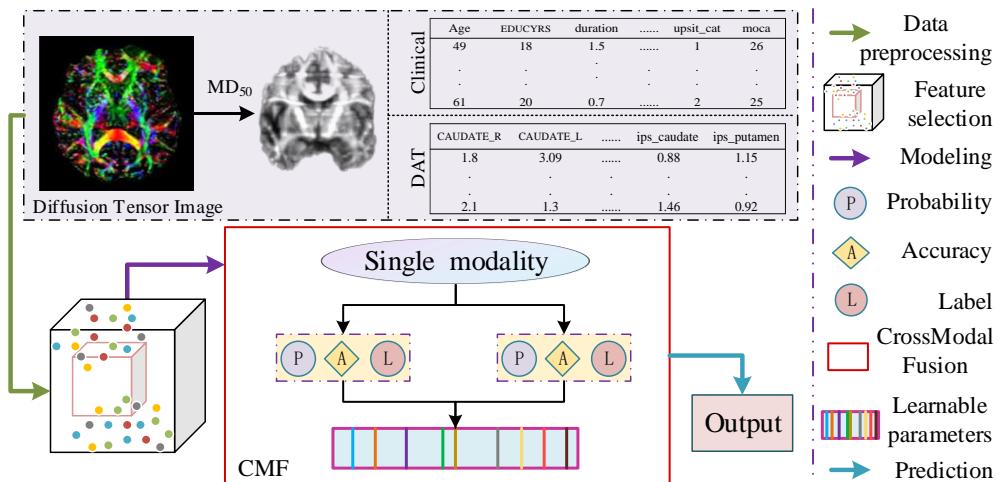


图 5.2: 本节提出的方法 (CMFP) 框架图

在本研究中，共有 123 名 PD 患者符合入选条件。通过分析 PPMI 数据库中这些患者 5 年的 HYS 评分动态，本节观察到 74 例患者的评分高于初始水平，46 例患者评分持平，而 3 例患者评分低于起始水平。据此，本节将评分提升的患者定义为进展组（共 72 例），评分未变或下降的患者则归为稳定组（共 49 例）。研究还收集了一系列临床信息，包括性别、年龄、教育背景以及 PD 运动评定量表等，并测量了脑脊液中的生物标志物：A $\beta$ 42、 $\alpha$ -syn、t-tau 和 p-tau 的蛋白浓度。123 位受试者接受了 3T Siemens MRI 的非对比增强 3D T1 加权成像和 DTI 扫描。DTI 影像的预处理使用了 PANDA 软件，包括去除非脑组织、纠正涡流效应和轻微头部移动、计算 DTI 矩阵以及应用连续跟踪 (FACT) 算法进行纤维追踪。若追踪过程中遇到曲率角度大于 45° 或体素分数各向异性 (FA) 值小于 0.2，则追踪终止。在构建 ML 模型时，本节选取了与 PD 进展相关的基线期临床指标、DAT 水平及 DTI 脑网络特征作为输入变量，并运用 Lasso 的方法进行特征选择，旨在提升模型的预测准确度和泛化能力。

### 5.3.1.2 进展预测的特征选择

在构建人工智能机器学习模型时，特征选择扮演着至关重要的角色，其重要性不容忽视。不同的特征选择会直接影响模型的预测准确性和在未见数据上的泛化能力。在本节的研究工作中，通过整合基线时的临床参数、多巴胺转运体 (DAT) 扫描数据以及扩散张量成像 (DTI) 所揭示的网络特性，本节创建了一个机器学习框架，用以预测五年后 PD 患者疾病的演变情况。为了提高模型的预测

效果，本节对数据进行了严格的预处理，剔除了无效、不规范和错误的数据。在原始的临床数据中，在本研究中，对 16 项临床指标进行了系统分析，这些指标包括三类基础信息、九种用于评估运动能力和认知水平的测试量表，以及四项反映脑脊液中蛋白质浓度的生化测量值。通过应用 Lasso 方法进行特征筛选，本节成功识别并保留了 4 个最为关键的临床变量 (UPDRS Part III Score、UPDRS Total Score、Age、UPSIT score)，这些变量与 PD 的疾病进展有着较高的相关性。对于 DTI 变量为脑区 MD (50) 矩阵，也使用 Lasso 方法进行数据选择，最终选择 7 个特征 (Splenum of Corpus Callosum、Fornix (Column and Body of Fornix)、Inferior Cerebellar Peduncle (Left)、Superior Cerebellar Peduncle (Left)、Fornix (Crescent)/Stria Terminalis (Right)、Tapetum Left、Tapetum Right)，分别代表终纹、胼胝体压部、穹窿、左侧小脑下脚、左侧小脑上脚、左内矢状层、右内矢状层这几个脑区。对于 DAT 数据，也使用 Lasso 方法进行数据选择，最终选择 4 个特征：Putamen Left、High Putamen、Low Striatum 和 High Striatum，它们分别代表了左侧腹核、高腹核、低纹状体和高纹状体。这样的特征选择策略不仅提高了模型的解释能力，还有助于增强模型预测效果及应对不同情况的广泛适用性。

### 5.3.1.3 跨模态数据融合的进展预测

本节通过整合基线时的临床参数、多巴胺转运体 (DAT) 扫描数据以及扩散张量成像 (DTI) 所揭示的网络特性来训练机器学习模型，用以预测五年后 PD 患者疾病的演变情况。在单模态建模过程中，本节采用 AdaBoost 算法进行的分类。AdaBoost (自适应增强算法) 是一种集成方法，它能够将多个表现一般的弱分类器协同工作以构建一个强大的综合分类器。其“自适应”体现在：每当一个弱分类器在训练中对某些样本判断错误时，这些样本的权重会被相应增加。随后，在下一轮迭代中，所有样本根据更新后的权重再次参与训练新的弱分类器。此过程不断循环，每轮迭代都会依据全体样本调整权值并确定每个新加入的弱分类器的贡献度，继续迭代直到满足预定错误率条件或达到最大循环次数。AdaBoost 能够根据弱分类器提供的信息动态调整错误率预期，并且以高效率运作，已有研究使用 AdaBoost 对 PD 进行早期预测 [293]。

为了进一步改进单变量基线数据对 5 年后进展预测的性能，本节采取跨模态数据组合的方式，并提出一种新的决策策略以提高预测性能。图 5.2 是本研究的总框架图，包括数据预处理、特征筛选、分类器建模、预测 4 部分构成，具体步

骤：

1. 设置参数 epoch=120, 使用 Adam 作为优化算法, 初始学习率为 0.01, 每 50 个 epoch 学习率下降 0.1, 损失的计算采取的是 MSE 的计算过程,  $\beta$  为 1 ,  $\alpha$  为 0.4;
2. 使用 Lasso 方法对多模态数据进行特征筛选, 将筛选出的特征作为训练集、验证集和测试集, 并按照 6: 3: 1 分配三个子集;
3. 将不同的单模态的训练集数据使用 AdaBoost 方法进行训练, 得到的模型使用验证集数据进行预测, 得到预测的概率、预测的标签和精确度, 并分别保存单模态训练好的模型;
4. 将不同模态预测的结果（预测的概率、预测的标签和精确度）输入多模态融合算法（CMF: CrossModalFusion）中, 得到一组新的预测结果（预测的概率、预测的标签和精确度）, 并保存融合模型;
5. 重复 3-5 步骤, 直到 epoch=120 结束循环;
6. 查找单模态训练过程中保存的精确度最高的模型  $M_s$  和两模态融合过程中保存的精确度最高的模型  $M_c$ 。
7. 将测试集数据输入模型  $M_s$ , 得到预测的结果（预测的概率、预测的标签和精确度）, 并将这组结果输入到模型  $M_c$  中进行测试, 得到两模态融合的预测概率和标签。

在决策之前本节获取到了两种单模态数据的预测结果, 包括预测标签, 预测概率及精确度, 分别记作  $\rho_a$ ,  $\rho_b$ ,  $\eta_a$ ,  $\eta_b$ ,  $\varphi_1$  and  $\varphi_2$ 。根据这些单模态预测结果, 使用本节提出的决策策略 CMF 可以得到决策融合的预测结果  $\Phi$ , 它由预测概率和预测标签组成, 即  $\Phi = [\rho, \eta]$ 。本节定义预测类别的差异  $\Omega = \text{abs}(\rho_a - \rho_b)$ , 即两个模型预测的类别之间的绝对差值。准确率的差异的符号  $\sigma = \text{sign}(\varphi_1 - \varphi_2)$ , 即两个模型准确率之差的符号。sign 函数返回-1、0 或 1, 表示差异为负、零或正。预测概率的差异  $v = \text{abs}(\eta_a - \eta_b)$ , 即两个模型预测的概率之间的绝对差值。 $\lambda$  是通过概率差异  $v$  计算的权重。 $\lambda = \iota / (1 + e^{-\alpha * v})$ 。其中,  $\iota$  为 1 ,  $\alpha$  为 0.4; 最终的预测概率和预测标签由公式 (5.1) 和 (5.2) 获得, 根据  $\Omega$ ,  $\sigma$  和  $v$  的取值, 选择不同的分支来组合预测结果。 $\zeta(A, B, C)$  表示条件函数, 如果  $A$  为真, 则结果为  $B$ , 否则为  $C$ 。

$$\rho = \zeta(\Omega = 0, \zeta(\sigma > 0, \eta_a, \eta_b), \zeta(\sigma > 0, \lambda + \eta_a, \lambda + \eta_b)) \quad (5.1)$$

$$\eta = \zeta(\Omega = 0, \zeta(\sigma > 0, \rho_a, \rho_b), \zeta(\rho < 0, -1, \rho > 0.5, 1, 0)) \quad (5.2)$$

### 5.3.2 智能辅助预测的实验分析

本节使用的是 Python 3.7.9 语言编程，在 GNU/Linux x86 64 system of GeForce RTX 3090 Ti 12 Intel Core Interl(R) Xeon(R) CPU E5-2678 v3 2.50GHz 64GB RAM 平台上运行。为了提高模型的性能，在输入模型进行训练与验证之前样本顺序都进行了随机的打乱。

表 5.3: 临床数据、DTI 数据和 DAT 数据的特征组合

$\delta_2$	Age	UP3s					
$\delta_3$	Age	UP3s	PTs				
$\delta_4$	Age	UP3s	PTs	Uts			
$\delta_5$	Age	UP3s	PTs	Uts	STs		
$\gamma_2$	SoCC	Fcb					
$\gamma_3$	SoCC	Fcb	ICPl				
$\gamma_4$	SoCC	Fcb	ICPl	SCPl			
$\gamma_5$	SoCC	Fcb	ICPl	SCPl	FcSTr		
$\gamma_6$	SoCC	Fcb	ICPl	SCPl	FcSTr	TaR	
$\gamma_7$	SoCC	Fcb	ICPl	SCPl	FcSTr	TaR	TaL
$\vartheta_2$	PuL	HPu					
$\vartheta_3$	PuL	HPu	LSt				
$\vartheta_4$	PuL	HPu	LSt	HSt			

#### 5.3.2.1 不同类型数据预测分析

使用单一或双模态数据来预测帕金森病 (PD) 的进展会得到不同的结果。表5.4展示了使用三种模态分别预测 PD 进展的平均 AUC 值，可以观察到每个 AUC 值相对较低。其中， $\Gamma$  指的是特征选择，而  $\Gamma_i$  指代特定的特征组合， $i$  表示组合的序号。不同数据集对应不同的特征组合，具体的特征组合显示在表5.3中。在表 5.3 中， $\delta_i$ ,  $\gamma_i$  和  $\vartheta_i$  分别代表临床数据、DTI 数据和 DAT 数据的特征组合，其中  $i$  表示特征组合的数量。 $\Gamma_{21}$ ,  $\Gamma_{23}$  和  $\Gamma_{50}$  代表每个模态的所有特征组合。UP3s: UPDRS Part III score, PTs: UPDRS Total Score, Uts: UPSIT score, STs: STAI scores. SCC: Splenium of Corpus Callosum, Fcb: Fornix (Column and Body of Fornix), ICl: Inferior Cerebellar Peduncle (Left), SCl: Superior

Cerebellar Peduncle (Left), FcSr: Fornix (Crescent)/Stria Terminalis (Right), TaR: Tapetum Right, TaL: Tapetum Left. PuL: Putamen Left, HPu: High Putamen, LSt: Low Striatum, HSt: High Striatum.

表 5.4: 单模态数据预测 PD 进展的平均 AUC 值

	$\Gamma_2$	$\Gamma_3$	$\Gamma_4$	$\Gamma_5$	$\Gamma_6$	$\Gamma_7$	$\Gamma_{21}$	$\Gamma_{23}$	$\Gamma_{50}$
Clinical	0.4077	0.4371	0.4515	0.4478	-	-	-	0.4629	-
DTI	0.4208	0.4476	0.4489	0.4310	0.4770	0.4191	-	-	0.4381
DAT	0.5017	0.4767	0.4631	-	-	-	0.4506	-	-

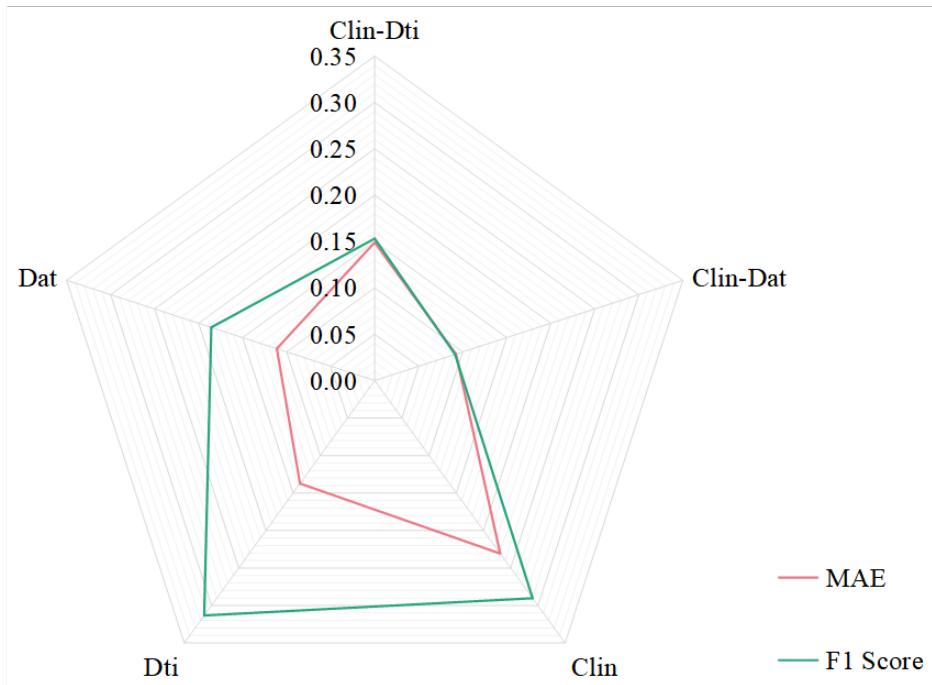


图 5.3: 不同策略进行五次预测的 MAE 和 F1 score 的方差

通过本节提出的跨模态融合预测方法 (CMFP)，对 DTI 和 DAT 的临床组合进行了测试，AUC 结果如表5.5所示。其中， $\delta_i$ ， $\gamma_i$  和  $\vartheta_i$  分别代表临床数据、DTI 数据和 DAT 数据的特征组合。它们的定义与表 5.3 中的相同。表5.5中的 A 表示临床数据与 DTI 的结合预测；B 表示临床数据与 DAT 的结合预测。不同数量的特征选择会导致 AUC 值的变化，当选择了 4 个临床特征和 7 个 DTI 特征时，结果相对较高，达到了 0.7791。因此，在随后的分析中，将以该组合作为实验对象。然而，临床和 DAT 的联合预测效果相对较差。当选择 2 个临

床特征和 DAT 特征进行预测时，AUC 最高为 0.6。使用本节提出的方法，对于预测 4 个临床特征和 7 个 DTI 特征的组合，方差为 0.1085。相反，在预测 2 个临床特征和 2 个 DAT 特征的组合八次后，方差为 0.1929。总体而言，临床和 DTI 的联合预测效果优于临床和 DAT 的联合预测效果。

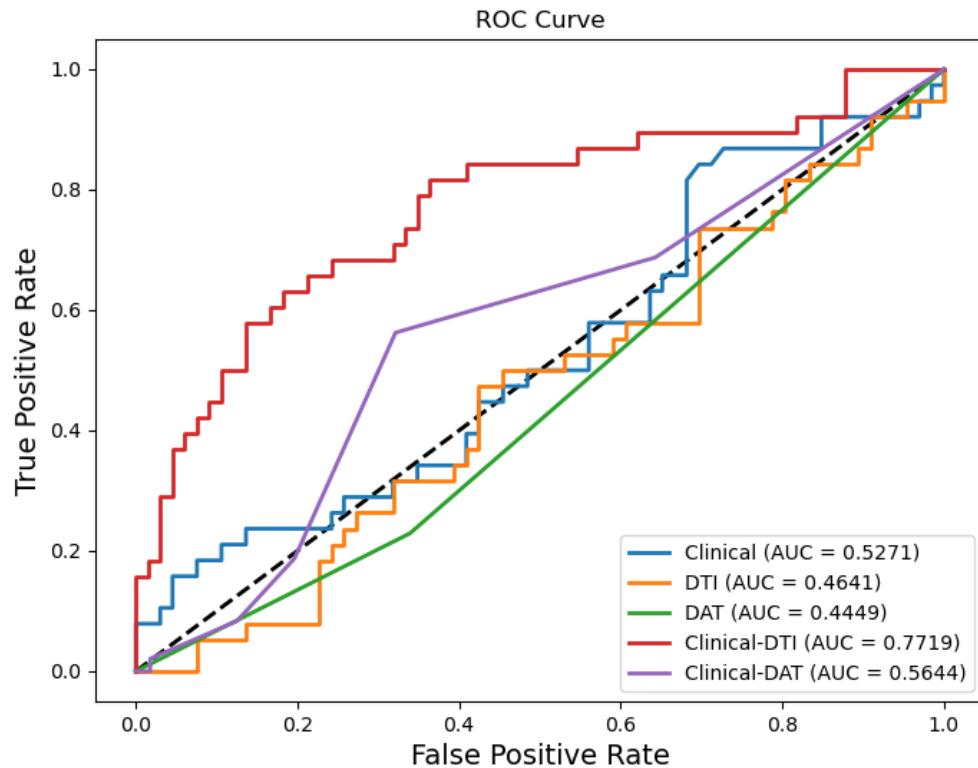


图 5.4：五种不同策略预测 PD 进展的 ROC 曲线对比分析

在表 5.5 中可以观察到，当将临床数据的 4 个特征与 DTI 数据的 7 个特征结合起来进行预测时，所得到的 AUC 值在所有特征组合中表现最佳。同样地，当将临床数据的 2 个特征与 DAT 数据的 2 个特征结合起来进行预测时，也得到了更优异的 AUC 值。为了更深入地分析单模态数据预测和双模态数据预测的模型性能，计算了 MAE 和 F1 分数。图 5.5(a) 和 (b) 分别展示了五种预测情况的 MAE 和 F1 分数评估结果。其中，Clin-Dti 代表临床数据与 DTI 结合的预测结果，Clin-Dat 代表临床数据与 DAT 结合的预测结果。MAE 用于衡量预测结果与实际观察值之间的平均偏差。较小的 MAE 值表示模型的预测结果更接近真实值。而 F1 分数是基于精确度和召回率的调和平均值，较高的值表

示模型的分类性能更好。

表 5.5: 临床数据分别与 DTI 和 DAT 融合预测的 AUC 值

<b>A</b>	$\delta_{23}$	$\delta_2$	$\delta_3$	$\delta_4$	$\delta_5$
$\gamma_{50}$	0.6409	0.4647	0.5539	0.4873	0.4859
$\gamma_2$	0.4955	0.5113	0.5506	0.5102	0.6459
$\gamma_3$	0.5315	0.5161	0.5053	0.5193	0.5858
$\gamma_4$	0.5458	0.4963	0.5435	0.4566	0.6038
$\gamma_5$	0.5288	0.4936	0.4623	0.6019	0.4905
$\gamma_6$	0.4872	0.6439	0.7096	0.5681	0.6135
$\gamma_7$	0.6677	0.5228	0.4717	<b>0.7791</b>	0.5172
<b>B</b>	$\delta_{23}$	$\delta_2$	$\delta_3$	$\delta_4$	$\delta_5$
$\vartheta_{21}$	0.5759	0.4670	0.4144	0.4606	0.4960
$\vartheta_2$	0.5592	<b>0.6000</b>	0.5683	0.5694	0.5842
$\vartheta_3$	0.5591	0.5990	0.4982	0.4657	0.5146
$\vartheta_4$	0.5729	0.5207	0.5518	0.5130	0.5882

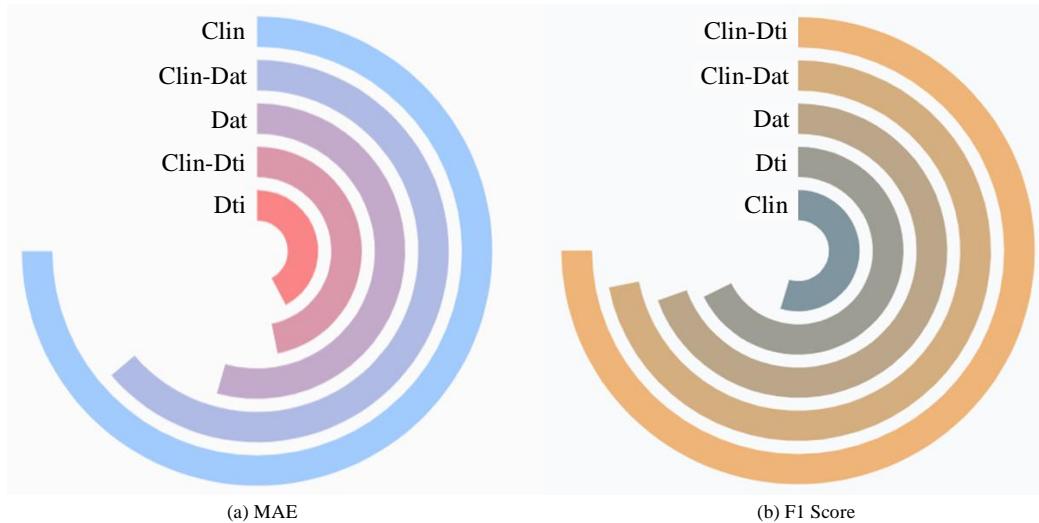


图 5.5: 采用不同策略进行五次预测 PD 进展的 MAE 和 F1 得分的平均值

在图 5.5(a) 的展示中, 可以发现在三种单模态数据预测方法中, 基于 DTI 数据的预测方法在平均绝对误差 (MAE) 方面取得了最优表现。而在两种双模态数据融合预测中, 结合临床数据与 DTI 数据的预测方法在 MAE 上展现出了更佳的性能。进一步地, 图 5.5(b) 揭示了五种预测方法中, 临床数据和 DTI 数据结合的预测方法在 F1 得分上位居首位。此外, 单模态预测的 F1 得分显著低

于双模态预测的 F1 得分。另外，对不同预测方法在 MAE 和 F1 得分上的方差进行计算，结果如图 5.3 所示。

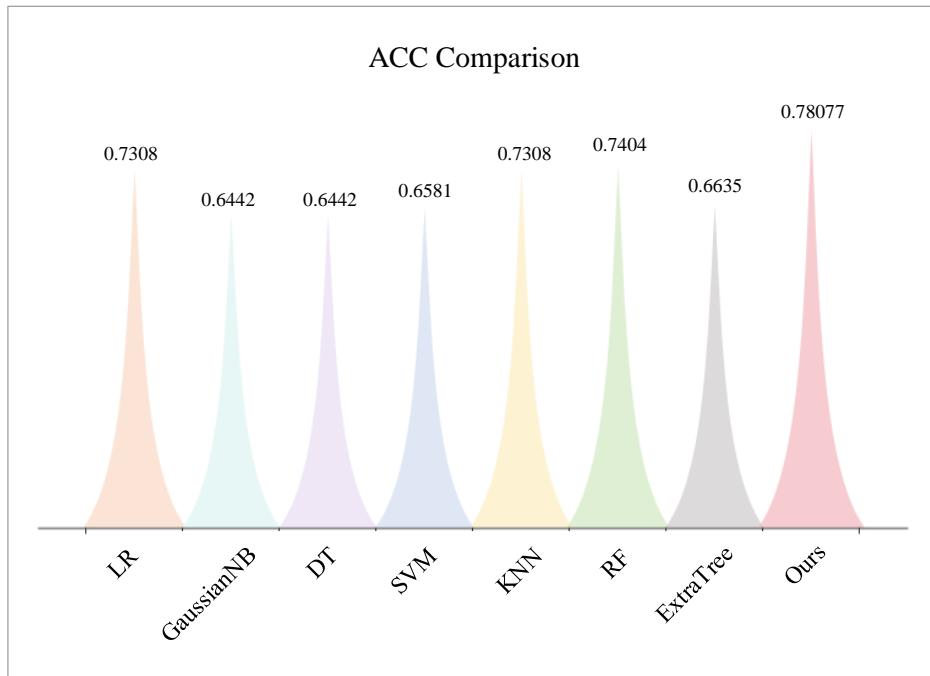


图 5.6: 嵌入不同 ML 技术的模型 ACC 对比分析

### 5.3.2.2 集成机器学习技术分析

在本研究中，提出了一个集成了 Adaboost 算法的新模型。为了验证该模型的有效性，对其进行了一系列的性能测试，并将结果与其他几种流行的 ML 技术（支持向量机（SVM）、逻辑回归（LR）、高斯朴素贝叶斯（GaussianNB）、K 最近邻（KNN）、决策树（DT）、随机森林（RF）、极端树（ExtraTree））进行了对比，如图 5.6 和图 5.7 所示。本节采用了准确率（ACC）、均方根误差（RMSE）和平均绝对误差（MAE）这三个核心指标来衡量模型的预测能力。ACC 是衡量分类模型性能的常用标准，它反映了模型正确预测的样本占总样本数的比例。当准确率接近 1 时，表明模型具有极高的预测准确性。RMSE 则提供了预测值与实际值之间差异的量化度量。由于它是误差的平方根，因此对于较大的误差值，RMSE 会更加敏感。一个较低的 RMSE 值意味着预测模型在数据拟合上更为精准。MAE 同样衡量预测值与实际值之间的差距，但与 RMSE 不同的是，MAE 对所有误差值赋予相同的权重，因此它对异常值的影响较小。一个较低的

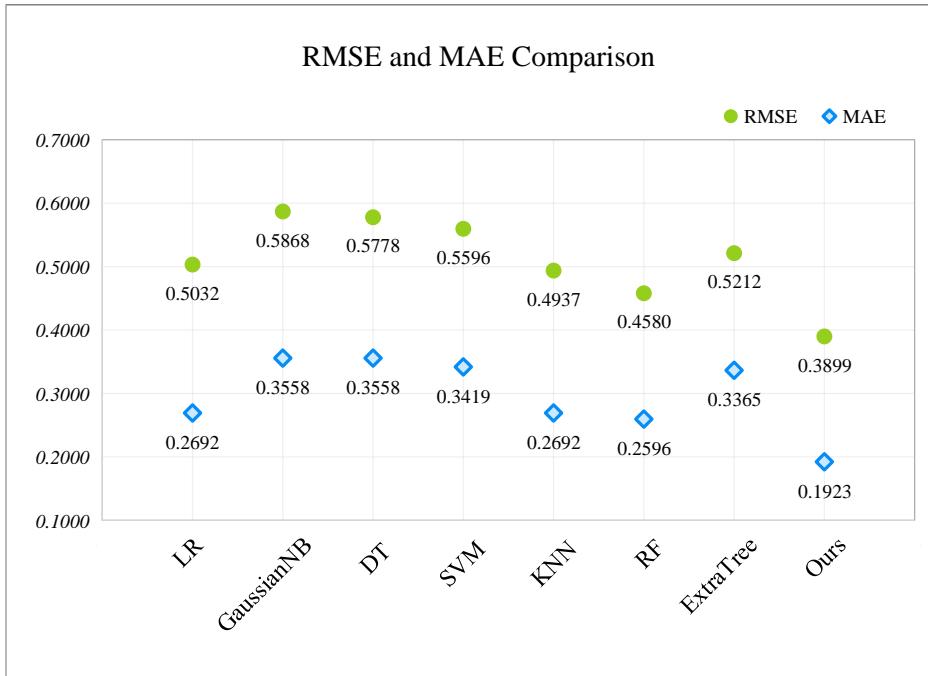


图 5.7: 嵌入不同 ML 技术的模型 RMSE 与 MAE 对比分析

MAE 值通常指示着模型的预测误差较小，即模型性能较好。

通过图 5.6 到图 5.9 的直观展示，本节的方法在所有三个性能指标上都表现出色。特别是在 ACC 方面，本节的模型比其他算法表现得更为突出，比排名第二的随机森林算法（RF）高出了 6.73 个百分点。在 RMSE 和 MAE 的数值对比上，本节的模型也显示出了相对于 RF 算法分别低 6.81% 和 6.73% 的显著优势。综上所述，本节的模型不仅在理论上设计精巧，而且在实际应用中也展现出了卓越的预测性能。这些结果充分证明了本节提出的方法在处理相关问题时的有效性和优越性。

### 5.3.3 进展预测模型的探索与发现

弥散张量成像 (DTI) 作为一项先进的磁共振成像技术，已在近年来得到迅速发展。它利用多个不同强度的扩散敏感梯度，能够精确地描绘出人体内水分子的扩散能力及其移动方向，从而深入揭示生物组织，尤其是脑组织的微观结构特点。这种成像技术不需要对活体进行任何侵入性操作，便能直观地观察和分析白质纤维束的完整性和方向性，这对于研究神经系统的疾病具有重要意义。一些学者基于弥散张量成像 (DTI) 的研究探索了白质的变化，并发现与对照组

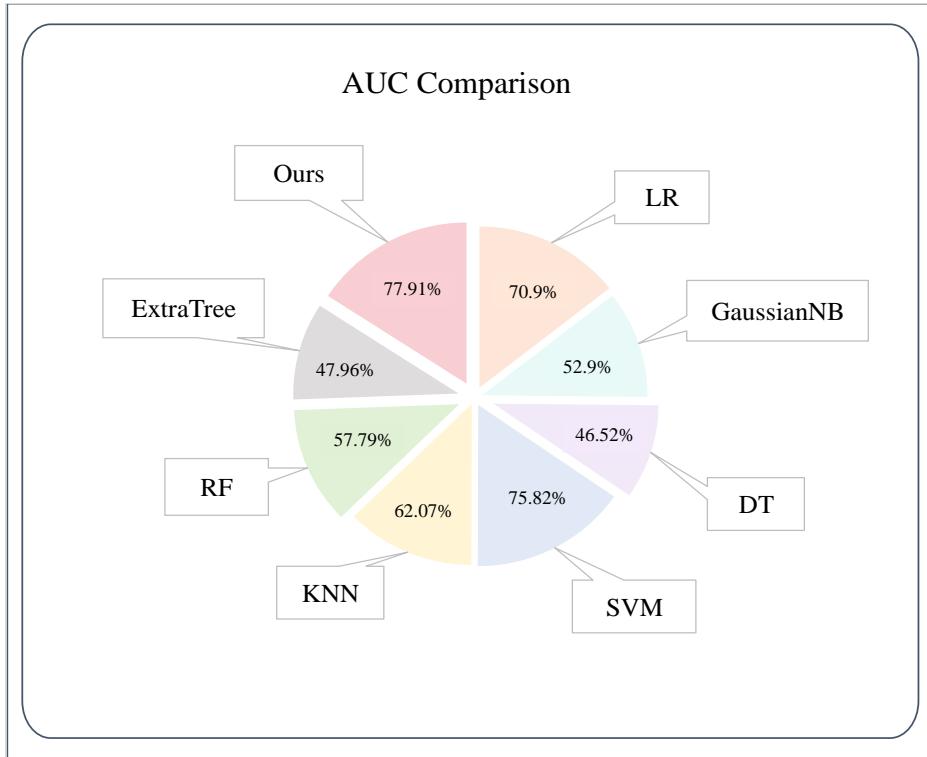


图 5.8: 嵌入不同 ML 技术的模型 AUC 对比分析

相比，PD-MCI 患者的额叶白质损伤更为显著，且与整体认知功能的损害以及多个认知领域的损害存在关联 [101]。对于 PD 的复杂多样的神经机制，通过整合多种模态 MRI 特征，从多个层面和不同角度进行综合评估和分析，为进一步提高 PD 早期诊断和干预提供了潜在的生物标志物 [102–104]。Pyatigorskaya 等人 [102] 研究者结合了 NMS-MRI 中的黑质体积和信号强度特征以及 DTI 中的部分各向异性特征，成功地将 PD 诊断的准确率提升至 93%。Lei 等人 [294, 284] 开发的基于 MRI 和 DTI 的预测模型能够智能识别 PD，并对临床评分进行准确预测，这在临床实践中具有重要的应用价值。Chougar 等人 [295] 则进一步探索了如何利用脑区（13 个大脑区域）体积和 DTI 参数作为输入，并使用监督 ML 算法准确预测 PD、PSP 和 MSA-P 的分类。Shin 等人 [296] 的研究关注了 MRI 皮层厚度在预测 PD 患者从轻度认知障碍向痴呆转变中的作用。他们的研究结果显示，MRI 皮层厚度能够有效预测个体的疾病进展，并且当与临床数据相结合时，其预测性能得到了进一步提升。这一发现强调了 MRI 技术在个体化医疗和精准治疗策略制定中的重要作用。

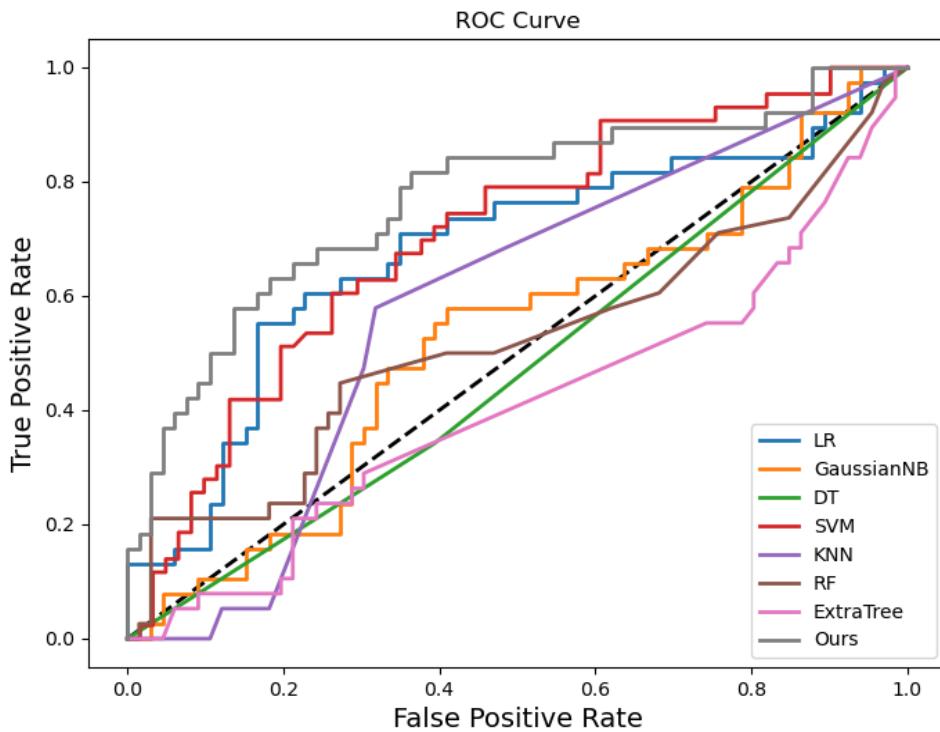


图 5.9: 嵌入不同 ML 技术的模型 ROC 对比分析

通过合并表 5.4 和表 5.5 中的数据，可以确定对 PD 预测贡献相对显著的临床指标，进而推断出与 PD 进展相关的潜在特征。这些关键的临床数据特征包括年龄、UPDRS 第三部分分数、UPDRS 总分以及 UPSIT 分数。

当前，可靠的网络表征技术的发展为在全脑连接层面理解神经疾病提供了新的途径。然而，目前关于使用白质数据来预测帕金森病（PD）进展的研究相对有限。Wee 等人 [297] 开发了一种基于网络的多变量分类算法，该算法利用白质纤维束数据，能够有效区分正常对照组和轻度认知障碍（MCI）患者。研究结果显示，这种分类框架能够为临床诊断以及与认知障碍相关的脑结构变化提供一种创新的替代和辅助手段。尽管如此，这项研究主要集中在阿尔茨海默病领域，对于 PD 的应用尚需进一步探索。

在本研究的预测模型中，除了纳入临床特征，还挑选了七个来自 DTI 变量的特征，这些特征涉及到脑白质的 MD。研究发现，这些脑白质的特征在预测 PD 的发展进程方面具有重要作用，并且当与传统的临床特征相结合时，能够

显著提升模型的预测准确性。具体来说，仅依赖 DTI 全球白质特征构建的预测模型获得了 0.44035 的 AUC 值（方差为 0.0185），而独立使用临床特征的模型 AUC 值稍高，为 0.4414（方差为 0.0187）。然而，当这两类特征被整合到同一个预测模型中时，模型的性能得到了显著的增强，AUC 值跃升至 0.7791（方差为 0.1085），这表明 DTI 白质特征与临床特征之间存在着互补作用，共同为预测 PD 提供了更加丰富和精确的信息。

多模态融合技术通过整合来自不同来源和类型的数据，能够提供更为丰富和立体的信息，从而在理解复杂的生物医学问题时发挥着至关重要的作用。在 PD 的研究中，多模态融合尤其展现了其独特优势。例如，结合临床症状评估、扩散张量成像 (Diffusion Tensor Imaging, DTI) 和多巴胺转运体 (Dopamine Transporter, DAT) 显像等多种数据源，不仅能够揭示 PD 患者大脑结构和功能的细微变化，还能帮助医生更准确地评估病情的严重程度和进展速度。图5.4所展示的五条 ROC 曲线直观地证明了多模态融合的有效性。在图5.4中，不同的颜色用于表示不同数据组合的预测结果。特别是在图 (b) 中，阴影区域展示了置信区间。分析结果指出，将临床数据与 DTI 或 DAT 影像数据结合使用的多模态方法，相较于单一数据源的预测方法，在预测帕金森病进展方面更为精确。尤其是将临床数据与 DTI 数据相结合的方法，展现出了最优的预测效果。单独使用某一种数据源的曲线（如临床数据、DTI 数据或 DAT 数据）通常只能提供有限的信息，而一旦将这些数据结合起来，ROC 曲线就会明显上升，表明模型的预测能力得到了显著提升。特别是当临床数据与 DTI 数据结合时，预测性能的提升尤为显著，这进一步验证了多模态融合在提高 PD 预测准确性方面的潜力。

表 5.6: 临床数据分别与 DTI 和 DAT 相结合进行预测

<i>Avg</i>	<b>AUC</b>	<b>ACC</b>	<b>SEN</b>	<b>SPE</b>
Clin-Dti	<b>0.7791</b>	<b>0.8077</b>	<b>0.9464</b>	<b>0.7470</b>
Clin-Dat	0.6000	0.7385	0.8679	0.6298
<i>Vac</i>	<b>AUC</b>	<b>ACC</b>	<b>SEN</b>	<b>SPE</b>
Clin-Dti	<b>0.1085</b>	0.1490	0.1417	0.1890
Clin-Dat	0.1929	<b>0.0923</b>	<b>0.1041</b>	<b>0.1516</b>

正如表5.4、表5.5以及图5.4所展示的，将临床数据与 DTI 数据结合起来进行预测，其效果优于结合临床数据与 DAT 数据。此外，本节还进行了 Mann-Whitney U 检验，该检验进一步证实了临床数据与 DTI 数据结合预测的显著

性，其结果在统计学上显著优于单独使用临床数据， $p$  值达到了  $9.183e-4$ 。这一结果强调了多模态数据融合在提高预测准确性方面的潜力。

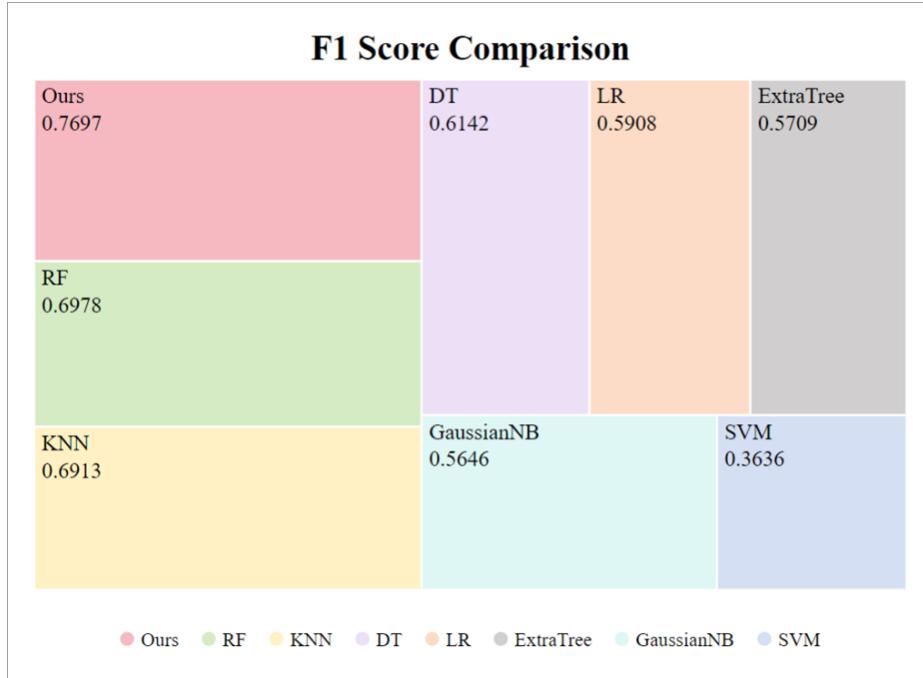


图 5.10: 嵌入不同 ML 技术的模型 F1 score 对比分析

为了深入比较临床数据与 DTI 及 DAT 数据结合的预测效果，采用了 ACC (准确率)、SEN (灵敏度) 和 SPE (特异性) 这三项关键指标进行了详尽的分析，并将结果呈现于表5.6中。该表格所展示的数据是在经过 8 次测试后，各项指标的平均值 (Avg) 和方差值 (Vac)。通过观察可以清晰地发现，在将临床数据与 DTI 数据相结合的情况下，无论是哪一项指标，其表现均显著优于其他组合，这一结果凸显了多模态融合在提高预测准确性方面的重要价值。

为了展现本节方法的卓越性，对不同机器学习算法替代情境下的 AUC 值进行了详细分析。AUC 作为衡量模型分类能力的指标，其值从 0 至 1 不等，数值越大表明模型的性能越优。图5.8和图5.9可见，即使在 AUC 值上 SVM 算法排名第二，本节的方法仍领先 SVM 2.09 个百分点。

本节还计算了各模型的 F1 分数，并将结果呈现于图5.10。从图中可以明显看出，本节的方法在 F1 分数上取得了最佳成绩，相比 RF 高出了 7.19%。虽然 SVM 在 AUC 指标上位居第二，但在 F1 分数方面却表现不佳。根据多维度的

评估结果，本节的方法展现出了卓越的性能，证实了其在实际应用中的可靠性和稳定性。

综上所述，本节的研究发现将 DTI 的白质纤维束数据与临床特征相结合，能够显著提高对 PD 进展的预测能力。通过采用本文提出的 CFMP 方法构建的网络模型，不仅能够有效评估 PD 患者的疾病严重度，还能实时监控疾病的进展情况。这一突破性成果对于改进 PD 的早期诊断和治疗方案具有深远的影响，有望为患者提供更加精准和个性化的医疗服务，从而改善他们的生活质量。

### 5.3.3.1 小结

PD 的发展是一个缓慢而持续的过程，其确诊通常依赖于临床表现和体征。随着疾病进入中期至晚期，治疗变得复杂且困难。因此，早期预测疾病进展对于实施及时有效的干预方案极为关键。但是，迄今为止还未发现精准的生物指标可以预判 PD 的病程发展，因此，构建一个能早期智能诊断该疾病的模型对于临床实践具有重大的意义。DTI 是一种能够量化大脑内水分子扩散性能的技术，它能够准确地衡量大脑中白质束的结构完好性，并且有助于识别早期病症的潜在迹象。本小节的研究采用基于纵向数据的研究方法，结合 ML 技术，整合早期 PD 患者的 DTI 数据和临床特征，构建了一个疾病进展预测模型。该模型有望为 PD 的早期诊断和治疗决策提供有力的影像组学支持。该小节主要局限性在于仅利用了脑白质的平均扩散率 (MD) 指标进行分析，未能纳入其他相关指标，如分数各向异性 (FA) 和纤维密度 (FN) 等。此外，研究未从其他数据库获取数据来验证模型的泛化能力。在未来的研究中，将探索多模态融合分析，包括脑灰质结构数据、基因信息等，以期对 PD 及其他神经退行性疾病进行更深入的理解和研究。

## 5.4 本章小结

本章节主要对 PD 的预测展开了研究与分析，包括 PD 的静态与动态预测。其中，在第5.2节中，回顾了基于 ML 的智能辅助帕金森疾病分类预测，强调 ML 技术的发展为帕金森病的智能诊断与健康测评提供了解决误诊或无法确诊的解决方案。通过对分类器训练的概述，强调了 ML 模型在挖掘大量数据中潜在规律方面的关键作用。在第5.3节中，提出了一种 CMFP 方法，基于跨模态数据融合的 ML，用于智能辅助 PD 进展预测。该方法结合了 DTI 和临床特征数据，

强调了 PD 疾病进展的漫长过程和对早期预测的需求。研究发现跨模态结合预测提高了单模态预测的准确性, DTI 数据对临床预测性能的提升具有显著作用。然而, 本节的研究未涵盖神经科学机制, 未来需要更深入地结合方法与神经科学, 以提高准确性。

## 第 6 章 总结与展望

人工智能的飞速进步为人类探索生命奥秘开辟了广阔天地，尤其在处理生老病死等根本问题时，人们对健康和长寿的追求从未停止。在医疗领域，人工智能的广泛应用正引领着一场深刻的变革：它不仅在临床诊断中协助医生和放射科专家提高工作效率、降低误诊概率，同时让患者得以获取更多的医学信息，增强了对自身病情的理解，显著提升了他们的健康意识。

影像组学的兴起为医疗影像分析带来了前所未有的革新，它为医生提供了全新的诊断工具。通过整合不同模态的影像数据与关键的临床信息，如基因数据和病理信息，影像组学能够为医生呈现更全面、更精确的疾病影像。深度学习在影像组学中的运用为整合这些数据提供了强大的支持。深度学习模型能够自动学习并挖掘多模态数据间的复杂联系，发现隐藏在数据背后的模式与规律。借助深度学习技术，影像组学能够更精确地定位和分析疾病特征，为医生提供更为可靠的辅助诊断信息。这种高度智能化的方法不仅提高了诊断的准确度，还加速了医学界对疾病机理和治疗方法的深入理解。

本研究专注于基于医学影像组学的辅助诊断算法以及其在临床实践中的应用。通过紧密结合临床需求，运用融合多模态数据及深度学习技术的手段，成功实现了对脑肿瘤、阿尔茨海默病（AD）和帕金森病（PD）的智能辅助诊断。在临床医生的指导和支持下，本文完成了以下研究工作：

1. 提出两种多模态融合的方法，成功在两模态的医学影像数据融合方面取得显著效果。这为更全面、多角度地理解患者病情提供了可行性，并且在提高诊断准确性方面具有重要意义；
2. 引入一种新的网络模型，该模型在对 AD 进行细粒度多分类任务中表现出较高的精度。这一模型的应用不仅提高了诊断的准确性，同时也为更细致的疾病分类奠定了坚实的基础；
3. 深入调研并综述了帕金森病分类预测的研究现状及趋势，并针对跨模态数据融合提出了一种创新的 PD 进展预测模型。这种模型有望为未来更加精准的帕金森病预测提供新的思路和方法。

当前研究尚处于初级阶段，未来有望探索以下几个研究方向：

- (1) 全面整合临床数据：未来的研究将更加强调对临床数据的全面整合，特

别是将基因数据与影像数据相结合用于辅助诊断。在影像组学的理念下，将医学数据看作一个复杂的网络，通过将来自不同来源的数据相互关联，协同构建患者的整体健康画像。这种多模态数据的整合，尤其是与基因数据和病理信息的结合，将为医生提供更全面的疾病理解，为个性化治疗方案的制定提供更准确的依据，从而提高治疗的针对性和有效性。

(2) 网络模型的优化与设计：将进一步优化网络模型，设计更适应医疗诊断需求的结构。考虑采用更先进的策略，例如卷积神经网络（CNN）与图卷积网络（GCN）的结合，以提高模型性能。通过深入研究和改进网络结构，期望能实现更准确、可靠且高效的医学影像辅助诊断模型。这将有助于提高诊断的准确性，使其更适应不同疾病特征的处理，并更好地满足临床医生的实际需求。

(3) 神经科学与深度学习的融合：未来的研究将致力于将神经科学的原理与深度学习应用相融合，以更深层次地探索潜在的疾病与感兴趣区域（ROI）之间的关系，从而进一步提高预测精度。通过深入研究大脑神经网络的结构和功能，可以更好地理解不同疾病对特定脑区域的影响。深度学习模型在此过程中将学习并捕捉微妙的神经影像特征，有望揭示潜在的疾病标志物。

(4) 强化医疗 AI 安全框架：鉴于医疗数据敏感性，未来的医学影像系统需内置强大的安全措施，如专用加密确保数据安全，区块链技术保护隐私与数据完整性，及动态监控防御网络攻击。通过融合深度学习与先进网络安全技术，构建智能安全的医疗数据生态，推动精准医疗同时，维护数字医疗环境的安全信任，保障医患利益。

这些研究方向将为医学影像组学辅助诊断领域提供新的动力和方向，有助于更好地迎接临床挑战，为患者提供更为个性化、精准的医疗服务。

## 参考文献

- [1] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. *European Journal of Cancer*, 2012, 48(4): 441-446.
- [2] Lo C M, Hung P H, Hsieh K L C. Computer-aided detection of hyperacute stroke based on relative radiomic patterns in computed tomography[J]. *Applied Sciences*, 2019, 9(8): 1668.
- [3] Bunaciu A A, Udriștioiu E G, Aboul-Enein H Y. X-ray diffraction: instrumentation and applications[J]. *Critical Reviews in Analytical Chemistry*, 2015, 45(4): 289-299.
- [4] Warren B E. X-ray Diffraction[M]. Courier Corporation, 1990: 1-215.
- [5] Boyd D P, Lipton M J. Cardiac computed tomography[J]. *Proceedings of the IEEE*, 1983, 71(3): 298-307.
- [6] Young I. Nuclear magnetic resonance imaging[J]. *Electronics and Power*, 1984, 30(3): 205-210.
- [7] Haider M, Uhlemann S, Schwan E, et al. Electron microscopy image enhanced[J]. *Nature*, 1998, 392(6678): 768-769.
- [8] Jaszczak R J, Coleman R E, Lim C B. SPECT: Single photon emission computed tomography[J]. *IEEE Transactions on Nuclear Science*, 1980, 27 (3): 1137-1153.
- [9] Newman P G, Rozycki G S. The history of ultrasound[J]. *Surgical Clinics of North America*, 1998, 78(2): 179-195.
- [10] Raichle M E. Positron emission tomography[J]. *Annual Review of Neuroscience*, 1983, 6(1): 249-267.

- [11] Xenakis T A, Gelalis I D, Koukoubis T D, et al. Neglected congenital dislocation of the hip: role of computed tomography and computer-aided design for total hip arthroplasty[J]. *The Journal of Arthroplasty*, 1996, 11(8): 893-898.
- [12] Weiss S M, Kulikowski C A, Amarel S, et al. A model-based method for computer-aided medical decision-making[J]. *Artificial Intelligence*, 1978, 11(1-2): 145-172.
- [13] Kanthraj G, Srinivas C R, Shenoi S D, et al. Comparison of computer-aided design and rule of nines methods in the evaluation of the extent of body involvement in cutaneous lesions[J]. *Archives of Dermatology*, 1997, 133(7): 922-923.
- [14] Economou G P, Lymberopoulos D, Karavatselou E, et al. A new concept toward computer-aided medical diagnosis-a prototype implementation addressing pulmonary diseases[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2001, 5(1): 55-65.
- [15] Balakrishnan G, Zhao A, Sabuncu M R, et al. VoxelMorph: a learning framework for deformable medical image registration[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(8): 1788-1800.
- [16] Ashburner J. A fast diffeomorphic image registration algorithm[J]. *Neuroimage*, 2007, 38(1): 95-113.
- [17] Sharma N, Aggarwal L M. Automated medical image segmentation techniques[J]. *Journal of Medical Physics*, 2010, 35(1): 3-14.
- [18] Zhou Z, Rahman Siddiquee M M, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. 2018: 3-11.

- [19] Chen J, Lu Y, Yu Q, et al. Transunet: Transformers make strong encoders for medical image segmentation[EB/OL]. 2021: 1-13. <https://doi.org/10.48550/arXiv.2102.04306>.
- [20] Anwar S M, Majid M, Qayyum A, et al. Medical image analysis using convolutional neural networks: a review[J]. *Journal of Medical Systems*, 2018, 42: 1-13.
- [21] Dogra A, Kumar S. Multi-modality medical image fusion based on guided filter and image statistics in multidirectional shearlet transform domain[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2023, 14(9): 12191-12205.
- [22] Tang L, Yuan J, Ma J. Image fusion in the loop of high-level vision tasks: A semantic-aware real-time infrared and visible image fusion network[J]. *Information Fusion*, 2022, 82: 28-42.
- [23] Dai Y, Gao Y, Liu F. Transmed: Transformers advance multi-modal medical image classification[J]. *Diagnostics*, 2021, 11(8): 1384:1-1384:15.
- [24] Frid-Adar M, Diamant I, Klang E, et al. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification[J]. *Neurocomputing*, 2018, 321: 321-331.
- [25] Wang X, Peng Y, Lu L, et al. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2017: 2097-2106.
- [26] Gaitonde D Y, Cook D L, Rivera I M. Chronic kidney disease: detection and evaluation[J]. *American Family Physician*, 2017, 96(12): 776-783.
- [27] Ferentinos K P. Deep learning models for plant disease detection and diagnosis[J]. *Computers and Electronics in Agriculture*, 2018, 145: 311-318.

- [28] Dahiwade D, Patle G, Meshram E. Designing disease prediction model using machine learning approach[C]//International Conference on Computing Methodologies and Communication. 2019: 1211-1215.
- [29] Shah D, Patel S, Bharti S K. Heart disease prediction using machine learning techniques[J]. *SN Computer Science*, 2020, 1: 1-6.
- [30] Krishnamoorthi R, Joshi S, Almarzouki H Z, et al. A novel diabetes health-care disease prediction framework using machine learning techniques[J]. *Journal of Healthcare Engineering*, 2022, 2022: 1-10.
- [31] Liu Y, Liu S, Wang Z. A general framework for image fusion based on multi-scale transform and sparse representation[J]. *Information Fusion*, 2015, 24: 147-164.
- [32] Maqsood S, Javed U. Multi-modal medical image fusion based on two-scale image decomposition and sparse representation[J]. *Biomedical Signal Processing and Control*, 2020, 57: 101810:1-101810:8.
- [33] Li Q, Wang W, Chen G, et al. Medical image fusion using segment graph filter and sparse representation[J]. *Computers in Biology and Medicine*, 2021, 131: 104239:1-104239:10.
- [34] Matsopoulos G, Marshall S, Brunt J. Multiresolution morphological fusion of MR and CT images of the human brain[J]. *IEE Proceedings-Vision, Image and Signal Processing*, 1994, 141(3): 137-142.
- [35] Mukhopadhyay S, Chanda B. Fusion of 2D grayscale images using multi-scale morphology[J]. *Pattern Recognition*, 2001, 34(10): 1939-1949.
- [36] Liu Y, Chen X, Ward R K, et al. Medical image fusion via convolutional sparsity based morphological component analysis[J]. *IEEE Signal Processing Letters*, 2019, 26(3): 485-489.
- [37] Ali H, Elmogy M, El-Daydamony E, et al. Multi-resolution MRI brain image segmentation based on morphological pyramid and fuzzy c-mean

- clustering[J]. *Arabian Journal for Science and Engineering*, 2015, 40: 3173-3185.
- [38] Hu M, Zhong Y, Xie S, et al. Fuzzy system based medical image processing for brain disease prediction[J]. *Frontiers in Neuroscience*, 2021, 15: 714318:1-714318:15.
- [39] Velmurugan S P, Sivakumar P, Rajasekaran M P. Multimodality image fusion using centre-based genetic algorithm and fuzzy logic[J]. *International Journal of Biomedical Engineering and Technology*, 2018, 28(4): 322-348.
- [40] Jin Z, Wang Y, Chen Z, et al. Medical image fusion in gradient domain with structure tensor[J]. *Journal of Medical Imaging and Health Informatics*, 2016, 6(5): 1314-1318.
- [41] Li S, Kang X, Fang L, et al. Pixel-level image fusion: A survey of the state of the art[J]. *Information Fusion*, 2017, 33: 100-112.
- [42] Singh R, Khare A. Fusion of multimodal medical images using Daubechies complex wavelet transform—A multiresolution approach[J]. *Information Fusion*, 2014, 19: 49-60.
- [43] Yin M, Liu X, Liu Y, et al. Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain [J]. *IEEE Transactions on Instrumentation and Measurement*, 2018, 68(1): 49-64.
- [44] Jasti V D P, Zamani A S, Arumugam K, et al. Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis[J]. *Security and Communication Networks*, 2022, 2022: 1-7.
- [45] Raja R, Kumar S, Rani S, et al. Artificial intelligence and machine learning in 2D/3D medical image processing[M]. CRC Press, 2020: 1-195.

- [46] Alseelawi N, Hazim H T, Salim ALRikabi H T. A Novel Method of Multimodal Medical Image Fusion Based on Hybrid Approach of NSCT and DTCWT[J]. *International Journal of Online and Biomedical Engineering*, 2022, 18(3).
- [47] Tang Z, Chen K, Pan M, et al. An augmentation strategy for medical image processing based on statistical shape model and 3D thin plate spline for deep learning[J]. *IEEE Access*, 2019, 7: 133111-133121.
- [48] Diwakar M, Tripathi A, Joshi K, et al. Latest trends on heart disease prediction using machine learning and image fusion[J]. *Materials Today: Proceedings*, 2021, 37: 3213-3218.
- [49] Wang W, Chang F. A Multi-focus Image Fusion Method Based on Laplacian Pyramid[J]. *Journal of Computers*, 2011, 6(12): 2559-2566.
- [50] Sahu A, Bhateja V, Krishn A, et al. Medical image fusion with Laplacian pyramids[C]//International Conference on Medical Imaging, M-health and Emerging Communication Systems. 2014: 448-453.
- [51] Mi S, Zhang L, Peng H, et al. Medical image fusion based on DTNP systems and Laplacian pyramid[J]. *Journal of Membrane Computing*, 2021, 3: 284-295.
- [52] Jiang Q, Jin X, Cui X, et al. A lightweight multimode medical image fusion method using similarity measure between intuitionistic fuzzy sets joint laplacian pyramid[J]. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2023, 7(3): 631-647.
- [53] Rajalingam B, Priya R. Multimodal medical image fusion based on deep learning neural network for clinical treatment analysis[J]. *International Journal of ChemTech Research*, 2018, 11(06): 160-176.
- [54] Liu Y, Chen X, Wang Z, et al. Deep learning for pixel-level image fusion: Recent advances and future prospects[J]. *Information Fusion*, 2018, 42: 158-173.

- [55] Kaur M, Singh D. Multi-modality medical image fusion technique using multi-objective differential evolution based deep neural networks[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2021, 12: 2483-2493.
- [56] Xia K j, Yin H s, Wang J q. A novel improved deep convolutional neural network model for medical image fusion[J]. *Cluster Computing*, 2019, 22: 1515-1527.
- [57] Zhou T, Cheng Q, Lu H, et al. Deep learning methods for medical image fusion: A review[J]. *Computers in Biology and Medicine*, 2023: 106959:1-106959:15.
- [58] Rajalingam B, Al-Turjman F, Santhoshkumar R, et al. Intelligent multimodal medical image fusion with deep guided filtering[J]. *Multimedia Systems*, 2022, 28(4): 1449-1463.
- [59] Wang K, Zheng M, Wei H, et al. Multi-modality medical image fusion using convolutional neural network and contrast pyramid[J]. *Sensors*, 2020, 20 (8): 2169.
- [60] El-Sayed F, El-Shafai W, Elsayed Taha T, et al. Efficient fusion of medical images based on CNN[J]. *Menoufia Journal of Electronic Engineering Research*, 2021, 30(2): 79-83.
- [61] Li X, Zhao J. A novel multi-modal medical image fusion algorithm[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2021, 12: 1995-2002.
- [62] Zhang Y, Wang M, Xia X, et al. Medical image fusion based on quasi-cross bilateral filtering[J]. *Biomedical Signal Processing and Control*, 2023, 80: 104259:1-104259:10.
- [63] Zhou T, Li Q, Lu H, et al. GAN review: Models and medical image fusion applications[J]. *Information Fusion*, 2023, 91: 134-148.

- [64] Huang J, Le Z, Ma Y, et al. MGMDcGAN: medical image fusion using multi-generator multi-discriminator conditional generative adversarial network[J]. *IEEE Access*, 2020, 8: 55145-55157.
- [65] Ma J, Xu H, Jiang J, et al. DDcGAN: A dual-discriminator conditional generative adversarial network for multi-resolution image fusion[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4980-4995.
- [66] Wang C, Yang G, Papanastasiou G, et al. DiCyc: GAN-based deformation invariant cross-domain information fusion for medical image synthesis[J]. *Information Fusion*, 2021, 67: 147-160.
- [67] Lahoud F, Süsstrunk S. Zero-learning fast medical image fusion[C]// International Conference on Information Fusion. 2019: 1-8.
- [68] Gray K R, Wolz R, Heckemann R A, et al. Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease[J]. *NeuroImage*, 2012, 60(1): 221-229.
- [69] Garali I, Adel M, Bourennane S, et al. Region-based brain selection and classification on pet images for Alzheimer's disease computer aided diagnosis[C]//IEEE international Conference on Image Processing. 2015: 1473-1477.
- [70] Feng J, Zhang S W, Chen L. Extracting ROI-based contourlet subband energy feature from the sMRI image for Alzheimer' s disease classification[J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021, 19(3): 1627-1639.
- [71] Gray K R, Wolz R, Keihaninejad S, et al. Regional analysis of FDG-PET for use in the classification of Alzheimer's disease[C]//2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. 2011: 1082-1085.

- [72] Suk H I, Shen D. Deep learning-based feature representation for AD/MCI classification[C]//Medical Image Computing and Computer-Assisted Intervention. 2013: 583-590.
- [73] So J H, Madusanka N, Choi H K, et al. Deep learning for Alzheimer's disease classification using texture features[J]. *Current Medical Imaging*, 2019, 15(7): 689-698.
- [74] Suk H I, Lee S W, Shen D, et al. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis[J]. *NeuroImage*, 2014, 101: 569-582.
- [75] Tufail A B, Ma Y K, Kaabar M K, et al. Classification of initial stages of Alzheimer's disease through pet neuroimaging modality and deep learning: quantifying the impact of image filtering approaches[J]. *Mathematics*, 2021, 9(23): 3101:1-3101:16.
- [76] Khvostikov A, Aderghal K, Krylov A, et al. 3D inception-based CNN with sMRI and MD-DTI data fusion for Alzheimer's disease diagnostics[EB/OL]. 2018: 1-13. <https://doi.org/10.48550/arXiv.1809.03972>.
- [77] Korolev S, Saifiullin A, Belyaev M, et al. Residual and plain convolutional neural networks for 3D brain MRI classification[C]//International Symposium on Biomedical Imaging. 2017: 835-838.
- [78] Ruiz J, Mahmud M, Modasshir M, et al. 3D DenseNet ensemble in 4-way classification of Alzheimer's disease[C]//Brain Informatics: International Conference. 2020: 85-96.
- [79] Song T A, Chowdhury S R, Yang F, et al. Graph convolutional neural networks for Alzheimer's disease classification[C]//IEEE International Symposium on Biomedical Imaging. 2019: 414-417.
- [80] Parmar H, Nutter B, Long R, et al. Spatiotemporal feature extraction and classification of Alzheimer's disease using deep learning 3D-CNN for fMRI data[J]. *Journal of Medical Imaging*, 2020, 7(5): 056001-056001.

- [81] 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: 501050:1-501050:19.
- [82] Basaia S, Agosta F, Wagner L, et al. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks[J]. *NeuroImage: Clinical*, 2019, 21: 101645:1-101645:8.
- [83] Helaly H A, Badawy M, Haikal A Y. Deep learning approach for early detection of Alzheimer's disease[J]. *Cognitive Computation*, 2021: 1-17.
- [84] Fu' adah Y N, Wijayanto I, Pratiwi N, et al. Automated classification of Alzheimer's disease based on MRI image processing using convolutional neural network (CNN) with AlexNet architecture[C]//Journal of Physics: Conference Series: Vol. 1844. 2021: 012020:1-012020:9.
- [85] Shamrat F J M, Akter S, Azam S, et al. AlzheimerNet: An effective deep learning based proposition for Alzheimer's disease stages classification from functional brain changes in magnetic resonance images[J]. *IEEE Access*, 2023, 11: 16376-16395.
- [86] Ramzan F, Khan M U G, Rehmat A, et al. A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks[J]. *Journal of Medical Systems*, 2020, 44: 1-16.
- [87] Prusiner S B. Neurodegenerative diseases and prions[J]. *New England Journal of Medicine*, 2001, 344(20): 1516-1526.
- [88] Leroy E, Anastasopoulos D, Konitsiotis S, et al. Deletions in the Parkin gene and genetic heterogeneity in a Greek family with early onset Parkinson's disease[J]. *Human Genetics*, 1998, 103: 424-427.

- [89] De Pablo-Fernández E, Lees A J, Holton J L, et al. Prognosis and neuropathologic correlation of clinical subtypes of Parkinson disease[J]. *Journal of the American Medical Association Neurology*, 2019, 76(4): 470-479.
- [90] Papagno C, Trojano L. Cognitive and behavioral disorders in Parkinson's disease: an update. I: cognitive impairments[J]. *Neurological Sciences*, 2018, 39(2): 215-223.
- [91] 温金玉, 方美娥. 基于机器学习的计算机辅助帕金森疾病分类预测研究综述[J]. 图学学报, 2023, 44: 1065-1079.
- [92] Li X, Xing Y, Martin-Bastida A, et al. Patterns of grey matter loss associated with motor subscores in early Parkinson's disease[J]. *NeuroImage: Clinical*, 2018, 17: 498-504.
- [93] de Schipper L J, van der Grond J, Marinus J, et al. Loss of integrity and atrophy in cingulate structural covariance networks in Parkinson's disease [J]. *NeuroImage: Clinical*, 2017, 15: 587-593.
- [94] Amoroso N, La Rocca M, Monaco A, et al. Complex networks reveal early MRI markers of Parkinson's disease[J]. *Medical Image Analysis*, 2018, 48: 12-24.
- [95] Rojas A, Górriz J, Ramírez J, et al. Application of empirical mode decomposition (EMD) on DaTSCAN SPECT images to explore Parkinson disease[J]. *Expert Systems with Applications*, 2013, 40(7): 2756-2766.
- [96] Abós A, Baggio H C, Segura B, et al. Discriminating cognitive status in Parkinson's disease through functional connectomics and machine learning [J]. *Scientific Reports*, 2017, 7(1): 1-13.
- [97] Oliveira F P, Faria D B, Costa D C, et al. Extraction, selection and comparison of features for an effective automated computer-aided diagnosis of Parkinson's disease based on  $[^{123}\text{I}]$  FP-CIT SPECT images[J]. *European Journal of Nuclear Medicine and Molecular Imaging*, 2018, 45: 1052-1062.

- [98] Martinez-Murcia F J, Górriz J M, Ramírez J, et al. Convolutional neural networks for neuroimaging in Parkinson's disease: Is preprocessing needed? [J]. *International Journal of Neural Systems*, 2018, 28(10): 1850035:1-1850035:19.
- [99] Kim D H, Wit H, Thurston M. Artificial intelligence in the diagnosis of Parkinson's disease from ioflupane-123 single-photon emission computed tomography dopamine transporter scans using transfer learning[J]. *Nuclear Medicine Communications*, 2018, 39(10): 887-893.
- [100] Castillo-Barnes D, Ramírez J, Segovia F, et al. Robust ensemble classification methodology for I123-Ioflupane SPECT images and multiple heterogeneous biomarkers in the diagnosis of Parkinson's disease[J]. *Frontiers in Neuroinformatics*, 2018, 12: 1-16.
- [101] Wang W, Mei M, Gao Y, et al. Changes of brain structural network connection in Parkinson's disease patients with mild cognitive dysfunction: a study based on diffusion tensor imaging[J]. *Journal of Neurology*, 2020, 267: 933-943.
- [102] Pyatigorskaya N, Magnin B, Mongin M, et al. Comparative study of MRI biomarkers in the substantia nigra to discriminate idiopathic Parkinson disease[J]. *American Journal of Neuroradiology*, 2018, 39(8): 1460-1467.
- [103] Gu Q, Zhang H, Xuan M, et al. Automatic classification on multi-modal MRI data for diagnosis of the postural instability and gait difficulty subtype of Parkinson's disease[J]. *Journal of Parkinson's Disease*, 2016, 6(3): 545-556.
- [104] Bowman F D, Drake D F, Huddleston D E. Multimodal imaging signatures of Parkinson's disease[J]. *Frontiers in Neuroscience*, 2016, 10: 131:1-131:11.
- [105] Kim H M, Nazor C, Zabetian C P, et al. Prediction of cognitive progression in Parkinson's disease using three cognitive screening measures[J]. *Clinical Parkinsonism and Related Disorders*, 2019, 1: 91-97.

- [106] Hu M T, Szewczyk-Królikowski K, Tomlinson P, et al. Predictors of cognitive impairment in an early stage Parkinson's disease cohort[J]. *Movement Disorders*, 2014, 29(3): 351-359.
- [107] Pigott K, Rick J, Xie S X, et al. Longitudinal study of normal cognition in Parkinson disease[J]. *Neurology*, 2015, 85(15): 1276-1282.
- [108] Marras C, Armstrong M J, Meaney C A, et al. Measuring mild cognitive impairment in patients with Parkinson's disease[J]. *Movement Disorders*, 2013, 28(5): 626-633.
- [109] Hoops S, Nazem S, Siderowf A, et al. Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease[J]. *Neurology*, 2009, 73(21): 1738-1745.
- [110] Dalrymple-Alford J, MacAskill M, Nakas C, et al. The MoCA: well-suited screen for cognitive impairment in Parkinson disease[J]. *Neurology*, 2010, 75(19): 1717-1725.
- [111] Miller D B, O' Callaghan J P. Biomarkers of Parkinson's disease: present and future[J]. *Metabolism*, 2015, 64(3): 40-46.
- [112] Latourelle J C, Beste M T, Hadzi T C, et al. Large-scale identification of clinical and genetic predictors of motor progression in patients with newly diagnosed Parkinson's disease: a longitudinal cohort study and validation [J]. *The Lancet Neurology*, 2017, 16(11): 908-916.
- [113] Simuni T, Long J D, Caspell-Garcia C, et al. Predictors of time to initiation of symptomatic therapy in early Parkinson's disease[J]. *Annals of Clinical and Translational Neurology*, 2016, 3(7): 482-494.
- [114] Holford N H, Chan P L, Nutt J G, et al. Disease progression and pharmacodynamics in Parkinson disease—evidence for functional protection with levodopa and other treatments[J]. *Journal of Pharmacokinetics and Pharmacodynamics*, 2006, 33: 281-311.

- [115] Ahmadi Rastegar D, Ho N, Halliday G M, et al. Parkinson' s progression prediction using machine learning and serum cytokines[J]. *npj Parkinson's Disease*, 2019, 5(1): 14:1-14:8.
- [116] Chan P L, Nutt J G, Holford N H. Levodopa slows progression of Parkinson' s disease. External validation by clinical trial simulation[J]. *Pharmaceutical Research*, 2007, 24: 791-802.
- [117] Raval V, Nguyen K P, Gerald A, et al. Prediction of individual progression rate in Parkinson' s disease using clinical measures and biomechanical measures of gait and postural stability[C]//IEEE International Conference on Acoustics, Speech and Signal Processing. 2020: 1319-1323.
- [118] Nilashi M, Ahmadi H, Sheikhtaheri A, et al. Remote tracking of Parkinson's disease progression using ensembles of deep belief network and self-organizing map[J]. *Expert Systems with Applications*, 2020, 159: 113562:1-113562:13.
- [119] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [120] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. *Advances in Neural Information Processing Systems*, 2012, 25.
- [121] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3431-3440.
- [122] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention. 2015: 234-241.
- [123] Roberts J W, Van Aardt J A, Ahmed F B. Assessment of image fusion procedures using entropy, image quality, and multispectral classification[J]. *Journal of Applied Remote Sensing*, 2008, 2(1): 023522:1-023522:28.

- [124] Qu G, Zhang D, Yan P. Information measure for performance of image fusion[J]. *Electronics Letters*, 2002, 38(7): 1-2.
- [125] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [126] Shi W, Zhu C, Tian Y, et al. Wavelet-based image fusion and quality assessment[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2005, 6(3-4): 241-251.
- [127] Eskicioglu A M, Fisher P S. Image quality measures and their performance [J]. *IEEE Transactions on Communications*, 1995, 43(12): 2959-2965.
- [128] Han Y, Cai Y, Cao Y, et al. A new image fusion performance metric based on visual information fidelity[J]. *Information Fusion*, 2013, 14(2): 127-135.
- [129] Haghigat M B A, Aghagolzadeh A, Seyedarabi H. A non-reference image fusion metric based on mutual information of image features[J]. *Computers and Electrical Engineering*, 2011, 37(5): 744-756.
- [130] Piella G, Heijmans H. A new quality metric for image fusion[C]// International Conference on Image Processing: Vol. 3. 2003: 111-173.
- [131] Wu J, Huang H, Qiu Y, et al. Remote sensing image fusion based on average gradient of wavelet transform[C]//IEEE International Conference Mechatronics and Automation: Vol. 4. 2005: 1817-1821.
- [132] Aslantas, V., Bendes, et al. A new image quality metric for image fusion: The sum of the correlations of differences[J]. *Aeu-international Journal of Electronics and Communications*, 2015, 69(12): 160-166.
- [133] Willmott C J, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance[J]. *Climate Research*, 2005, 30(1): 79-82.

- [134] Li S, Hong R, Wu X. A novel similarity based quality metric for image fusion[C]//International Conference on Audio, Language and Image Processing. 2008: 167-172.
- [135] Han S, Li H, Gu H, et al. The study on image fusion for high spatial resolution remote sensing images[J]. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXVII. Part B*, 2008, 7: 1159-1164.
- [136] Khedher L, Ramírez J, Górriz J M, et al. Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images[J]. *Neurocomputing*, 2015, 151: 139-150.
- [137] Radiocommunication I. Methodology for the subjective assessment of the quality of television pictures[J]. *Recommendation ITU-R BT. 500-11*, 2002: 1-48.
- [138] 韦玉春, 汤国安, 杨昕. 遥感数字图像处理教程[M]. 遥感数字图像处理教程, 2007: 1-298.
- [139] Tang L, Tian C, Li L, et al. Perceptual quality assessment for multimodal medical image fusion[J]. *Signal Processing: Image Communication*, 2020, 85: 115852:1-115852:8.
- [140] Wen J, Qin F, Du J, et al. MsgFusion: Medical semantic guided two-branch network for multimodal brain image fusion[J]. *IEEE Transactions on Multimedia*, 2023, 26: 944-957.
- [141] Wen J, Khan A, Chen A, et al. High-Quality Fusion and Visualization for MR-PET Brain Tumor Images via Multi-Dimensional Features [J]. *IEEE Transactions on Image Processing*, May 2024, published online, DOI: 10.1109.
- [142] Wei T, Aviles-Rivero A I, Wang S, et al. Beyond fine-tuning: Classifying high resolution mammograms using function-preserving transformations[J]. *Medical Image Analysis*, 2022, 82: 102618:1-102618:10.

- [143] Shen Y, Wu N, Phang J, et al. An interpretable classifier for high-resolution breast cancer screening images utilizing weakly supervised localization[J]. *Medical Image Analysis*, 2021, 68: 101908:1-101908:17.
- [144] Guo J, Zhou Z, Wang L. Single image highlight removal with a sparse and low-rank reflection model[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 268-283.
- [145] Xia W, Chen E C, Pautler S E, et al. A global optimization method for specular highlight removal from a single image[J]. *IEEE Access*, 2019, 7: 125976-125990.
- [146] Yu Z, Bai X, Tao W. Boundary Finding Based Multi-focus Image Fusion through Multi-scale Morphological Focus-measure[J]. *Information Fusion*, 2017, 35: 81-101.
- [147] Li H, Wu X J, Durrani T. Multi-focus Noisy Image Fusion using Low-Rank Representation[EB/OL]. 2018: 1-17. <https://doi.org/10.48550/arXiv.1804.09325>.
- [148] Li H, Wu X J. Infrared and visible image fusion using latent low-rank representation[EB/OL]. 2018: 1-6. <https://doi.org/10.48550/arXiv.1804.08992>.
- [149] Du J, Fang M, Yu Y, et al. An adaptive two-scale biomedical image fusion method with statistical comparisons[J]. *Computer Methods and Programs in Biomedicine*, 2020, 196: 105603:1-105603:13.
- [150] Diwakar M, Singh P, Shankar A. Multi-modal medical image fusion framework using co-occurrence filter and local extrema in NSST domain[J]. *Biomedical Signal Processing and Control*, 2021, 68: 102788:1-102788:10.
- [151] Wang B, Niu H, Zeng J, et al. Latent Representation Learning Model for Multi-band Images Fusion via Low-Rank and Sparse Embedding[J]. *IEEE Transactions on Multimedia*, 2021, 23: 3137-3152.

- [152] Hui L, Wu X J. DenseFuse: A Fusion Approach to Infrared and Visible Images[J]. *IEEE Transactions on Image Processing*, 2018, 28(5): 2614-2623.
- [153] Liu Y, Chen X, Wang Z, et al. Deep learning for pixel-level image fusion: Recent advances and future prospects[J]. *Information Fusion*, 2018, 42: 158-173.
- [154] Yan X, Gilani S Z, Qin H, et al. Unsupervised deep multi-focus image fusion[EB/OL]. 2018: 1-11. <https://doi.org/10.48550/arXiv.1806.07272>.
- [155] Kumar N, Hoffmann N, Oelschlägel M, et al. Structural similarity based anatomical and functional brain imaging fusion[M]//Multimodal Brain Image Analysis and Mathematical Foundations of Computational Anatomy: Vol. 11846. 2019: 121-129.
- [156] Ma J, Wei Y, Liang P, et al. FusionGAN: A generative adversarial network for infrared and visible image fusion[J]. *Information Fusion*, 2019, 48: 11-26.
- [157] Yu Z A, Yu L B, Peng S C, et al. IFCNN: A general image fusion framework based on convolutional neural network[J]. *Information Fusion*, 2020, 54: 99-118.
- [158] H. Li T D, X. Wu. NestFuse: An Infrared and Visible Image Fusion Architecture based on Nest Connection and Spatial/Channel Attention Models [J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69 (12): 9645-9656.
- [159] Xu H, Ma J, Le Z, et al. FusionDN: A Unified Densely Connected Network for Image Fusion[J]. *AAAI Conference on Artificial Intelligence*, 2020, 34 (7): 12484-12491.
- [160] Zhao F, Zhao W. Learning Specific and General Realm Feature Representations for Image Fusion[J]. *IEEE Transactions on Multimedia*, 2021, 23: 2745-2756.

- [161] Zhou T, Thung K H, Zhang Y, et al. Inter-modality dependence induced data recovery for MCI conversion prediction[C]//Medical Image Computing and Computer Assisted Intervention: Vol. 11767. 2019: 186-195.
- [162] Algarni A D. Automated medical diagnosis system based on multi-modality image fusion and deep learning[J]. *Wireless Personal Communications*, 2020, 111(2): 1033-1058.
- [163] Li H, He X, Tao D, et al. Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning[J]. *Pattern Recognition*, 2018, 79: 130-146.
- [164] Panigrahy C, Seal A, Mahato N K. MRI and SPECT image fusion using a weighted parameter adaptive dual channel PCNN[J]. *IEEE Signal Processing Letters*, 2020, 27: 690-694.
- [165] Parvathy V S, Pothiraj S, Sampson J. A novel approach in multimodality medical image fusion using optimal shearlet and deep learning[J]. *International Journal of Imaging Systems and Technology*, 2020, 30(4): 847-859.
- [166] Hermessi H, Mourali O, Zagrouba E. Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain[J]. *Neural Computing and Applications*, 2018, 30(7): 2029-2045.
- [167] Das M, Gupta D, Radeva P, et al. Optimized Multimodal Neurological Image Fusion based on Low-rank Texture Prior Decomposition and Superpixel Segmentation[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022.
- [168] Liang X, Hu P, Zhang L, et al. MCFNet: Multi-Layer Concatenation Fusion Network for Medical Images Fusion[J]. *IEEE Sensors Journal*, 2019, 19(16): 7107-7119.
- [169] Abiko K, Uruma K, Sugawara M, et al. Image Segmentation Based Graph-Cut Approach to Fast Color Image Coding via Graph Fourier Transform [C]//IEEE Visual Communications and Image Processing. 2019: 1-4.

- [170] Kuo C C, Chuang H C, Liao A H, et al. Fast Fourier transform combined with phase leading compensator for respiratory motion compensation system[J]. *Quantitative Imaging in Medicine and Surgery*, 2020, 10(5): 907–920.
- [171] Gnutti A, Guerrini F, Leonardi R, et al. Symmetry-Based Graph Fourier Transforms: Are They Optimal For Image Compression?[C]//IEEE International Conference on Image Processing. 2021: 1594-1598.
- [172] Fang L, Au O C, Chen Y, et al. Joint Demosaicing and Subpixel-Based Down-Sampling for Bayer Images: A Fast Frequency-Domain Analysis Approach[J]. *IEEE Transactions on Multimedia*, 2012, 14(4): 1359-1369.
- [173] Hu G, Cui B, Yu S. Joint Learning in the Spatio-Temporal and Frequency Domains for Skeleton-Based Action Recognition[J]. *IEEE Transactions on Multimedia*, 2020, 22(9): 2207-2220.
- [174] Atta-Fosu T, Guo W. Joint segmentation and nonlinear registration using fast fourier transform and total variation[M]//Research in Shape Analysis. 2018: 111-132.
- [175] Pezzotti E. Efficient Non-uniform Fast Fourier Transform (NuFFT) Implementation for MRI Processing on FPGA[D]. University of Illinois at Chicago, 2017.
- [176] Naidu V. Multi-resolution image fusion by FFT[C]//International Conference on Image Information Processing. 2011: 1-6.
- [177] Xu K, Qin M, Sun F, et al. Learning in the frequency domain[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2020: 1740-1749.
- [178] Peng L, HUANG Y, Kunlun Y. Multi-algorithm fusion of RGB and HSV color spaces for image enhancement[C]//Chinese Control Conference. 2018: 9584-9589.

- [179] Jin X, Chen G, Hou J, et al. Multimodal sensor medical image fusion based on nonsubsampled shearlet transform and S-PCNNs in HSV space [J]. *Signal Processing*, 2018, 153: 379-395.
- [180] Liu J, Wei M. Scene sparse recognition method via intra-class dictionary for visible and near-infrared HSV image fusion[J]. *Journal of Computer Applications*, 2018, 19: 3355-3359.
- [181] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2017: 4700-4708.
- [182] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[C]//AAAI Conference on Artificial Intelligence. 2017: 4278-4284.
- [183] Wang L, Guo S, Huang W, et al. Places205-VGGNet models for scene recognition[EB/OL]. 2015: 1-2. <https://doi.org/10.48550/arXiv.1508.01667>.
- [184] Zheng Y, Essock E A, Hansen B C, et al. A new metric based on extended spatial frequency and its application to DWT based fusion algorithms[J]. *Information Fusion*, 2007, 8(2): 177-192.
- [185] Piella G, Heijmans H. A new quality metric for image fusion[C]// International Conference on Image Processing: Vol. 3. 2003: 173-176.
- [186] Sengupta A, Seal A, Panigrahy C, et al. Edge information based image fusion metrics using fractional order differentiation and sigmoidal functions [J]. *IEEE Access*, 2020, 8: 88385-88398.
- [187] C. Xydeas V P. Objective Image Fusion Performance Measure[J]. *Electronics Letters*, 2000, 36(4): 308-309.
- [188] Nakamoto Y, Tamai K, Saga T, et al. Clinical value of image fusion from MR and PET in patients with head and neck cancer[J]. *Molecular Imaging and Biology*, 2009, 11: 46-53.

- [189] Heiss W D. The potential of PET/MR for brain imaging[J]. *European Journal of Nuclear Medicine and Molecular Imaging*, 2009, 36(1): 105-112.
- [190] Fedorova A N, Zeitlin M G. Quantum multiresolution: tower of scales[C]// International Conference on Mathematical Modeling in Physical Sciences: Vol. 490. 2014: 012216:1-012216:5.
- [191] Gambhir D, Manchanda M. Waveatom transform-based multimodal medical image fusion[J]. *Signal, Image and Video Processing*, 2019, 13(2): 321-329.
- [192] Benjamin J R, Jayasree T. Improved medical image fusion based on cascaded PCA and shift invariant wavelet transforms[J]. *International Journal of Computer Assisted Radiology and Surgery*, 2018, 13(2): 229-240.
- [193] Yang Y, Yang M, Huang S, et al. Multifocus image fusion based on extreme learning machine and human visual system[J]. *IEEE Access*, 2017, 5: 6989-7000.
- [194] Zhong J, Yang B, Li Y, et al. Image fusion and super-resolution with convolutional neural network[C]//Chinese Conference on Pattern Recognition. 2016: 78-88.
- [195] Hla B, Xh C, Zya B, et al. Noise-robust image fusion with low-rank sparse decomposition guided by external patch prior - ScienceDirect[J]. *Information Sciences*, 2020, 523: 14-37.
- [196] Zhan K, Li Q, Teng J, et al. Multifocus image fusion using phase congruency[J]. *Journal of Electronic Imaging*, 2015, 24(3): 033014:1-033014:13.
- [197] Wen J, Xuan S, Li Y, et al. Image-segmentation algorithm based on wavelet and data-driven neutrosophic fuzzy clustering[J]. *The Imaging Science Journal*, 2019, 67(2): 63-75.
- [198] Rajalingam B, Al-Turjman F, Santhoshkumar R, et al. Intelligent multimodal medical image fusion with deep guided filtering[J]. *Multimedia Systems*, 2022, 28(4): 1449-1463.

- [199] YIN X. Identification of subthreshold depression based on deep learning and multimodal medical image fusion[J]. *Chinese Journal of Medical Imaging Technology*, 2020, 36(8): 1158-1162.
- [200] Rajalingam B, Priya R. Multimodal medical image fusion based on deep learning neural network for clinical treatment analysis[J]. *International Journal of ChemTech Research*, 2018, 11(06): 160-176.
- [201] Xiao B, Xu B, Bi X, et al. Global-feature encoding U-Net (GEU-Net) for multi-focus image fusion[J]. *IEEE Transactions on Image Processing*, 2021, 30: 163-175.
- [202] Shopovska I, Jovanov L, Philips W. Deep Visible and Thermal Image Fusion for Enhanced Pedestrian Visibility[J]. *Sensors (Basel, Switzerland)*, 2019, 19(17): 3727:1-3727:21.
- [203] Hou R, Zhou D, Nie R, et al. VIF-Net: An unsupervised framework for infrared and visible image fusion[J]. *IEEE Transactions on Computational Imaging*, 2020, 6: 640-651.
- [204] Xu H, Fan F, Zhang H, et al. A deep model for multi-focus image fusion based on gradients and connected regions[J]. *IEEE Access*, 2020, 8: 26316-26327.
- [205] Zhao Z, Bai H, Zhang J, et al. CDDFuse: Correlation-Driven Dual-Branch Feature Decomposition for Multi-Modality Image Fusion[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2023: 5906-5916.
- [206] Xu H, Ma J, Jiang J, et al. U2Fusion: A Unified Unsupervised Image Fusion Network[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(1): 502-518.
- [207] Liang P, Jiang J, Liu X, et al. Fusion from Decomposition: A Self-Supervised Decomposition Approach for Image Fusion[C]//European Conference on Computer Vision: Vol. 13678. 2022: 719-735.

- [208] Xu H, Ma J, Zhang X P. MEF-GAN: Multi-exposure image fusion via generative adversarial networks[J]. *IEEE Transactions on Image Processing*, 2020, 29: 7203-7216.
- [209] Li J, Huo H, Li C, et al. AttentionFGAN: Infrared and visible image fusion using attention-based generative adversarial networks[J]. *IEEE Transactions on Multimedia*, 2021, 23: 1383-1396.
- [210] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.
- [211] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2017: 2261-2269.
- [212] Touvron H, Cord M, Sablayrolles A, et al. Going deeper with image transformers[C]//IEEE International Conference on Computer Vision. 2021: 1-30.
- [213] Aardt V, Jan. Assessment of image fusion procedures using entropy, image quality, and multispectral classification[J]. *Journal of Applied Remote Sensing*, 2008, 2(1): 1-28.
- [214] Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(8): 1226-1238.
- [215] Yun-Jiang, Rao. In-fibre Bragg grating sensors[J]. *Measurement Science and Technology*, 1997, 8(4): 355-375.
- [216] Han Y, Cai Y, Cao Y, et al. A new image fusion performance metric based on visual information fidelity[J]. *Information Fusion*, 2013, 14(2): 127-135.

- [217] Wen J, Li Y, Fang M, et al. Fine-Grained and Multiple Classification for Alzheimer's Disease With Wavelet Convolution Unit Network[J]. *IEEE Transactions on Biomedical Engineering*, 2023, 70(9): 2592-2603.
- [218] Fang M, Jin Z, Qin F, et al. Re-transfer learning and multi-modal learning assisted early diagnosis of Alzheimer's disease[J]. *Multimedia Tools and Applications*, 2022, 81(20): 29159-29175.
- [219] De A, Chowdhury A S. DTI based Alzheimer's disease classification with rank modulated fusion of CNNs and random forest[J]. *Expert Systems with Applications*, 2021, 169: 114338:1-114338:12.
- [220] Huynh K M, Chen G, Wu Y, et al. Multi-Site Harmonization of Diffusion MRI Data via Method of Moments[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(7): 1599-1609.
- [221] Frisoni G B, Fox N C, Jack C R, et al. The clinical use of structural MRI in Alzheimer disease[J]. *Nature Reviews Neurology*, 2010, 6(2): 67-77.
- [222] Payan A, Montana G. Predicting Alzheimer's disease: A neuroimaging study with 3D convolutional neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2015: 1-9.
- [223] Xiao Z, Ding Y, Lan T, et al. Brain MR Image Classification for Alzheimer's Disease Diagnosis Based on Multifeature Fusion[J]. *Computational and Mathematical Methods in Medicine*, 2017, 2017: 1952373:1-1952373:13.
- [224] Rathore S, Habes M, Iftikhar M A, et al. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages[J]. *NeuroImage*, 2017, 155: 530-548.
- [225] Horgusluoglu-Moloch E, Xiao G, Wang M, et al. Systems modeling of white matter microstructural abnormalities in Alzheimer's disease[J]. *NeuroImage: Clinical*, 2020, 26: 102203:1-102203:15.

- [226] Ben Ahmed O, 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.
- [227] Prasad G, Joshi S H, Nir T M, et al. Brain connectivity and novel network measures for Alzheimer's disease classification[J]. *Neurobiology of Aging*, 2015, 36: 1-11.
- [228] Ashburner J, Friston K J. Voxel-Based Morphometry—The Methods[J]. *NeuroImage*, 2000, 11(6): 805-821.
- [229] Zhang J, Gao Y, Gao Y, et al. Detecting Anatomical Landmarks for Fast Alzheimer's Disease Diagnosis[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(12): 2524-2533.
- [230] Shehata M, Khalifa F, Soliman A, et al. Computer-Aided Diagnostic System for Early Detection of Acute Renal Transplant Rejection Using Diffusion-Weighted MRI[J]. *IEEE Transactions on Biomedical Engineering*, 2019, 66(2): 539-552.
- [231] Bringas S, Salomón S, Duque R, et al. Alzheimer's Disease stage identification using deep learning models[J]. *Journal of Biomedical Informatics*, 2020, 109: 103514:1-103514:9.
- [232] Liu M, Zhang J, Adeli E, et al. Joint Classification and Regression via Deep Multi-Task Multi-Channel Learning for Alzheimer's Disease Diagnosis[J]. *IEEE Transactions on Biomedical Engineering*, 2019, 66(5): 1195-1206.
- [233] Golla A K, Bauer D F, Schmidt R, et al. Convolutional Neural Network Ensemble Segmentation With Ratio-Based Sampling for the Arteries and Veins in Abdominal CT Scans[J]. *IEEE Transactions on Biomedical Engineering*, 2021, 68(5): 1518-1526.

- [234] Liang Z, Cai J, Cao Z, et al. CameraNet: A Two-Stage Framework for Effective Camera ISP Learning[J]. *IEEE Transactions on Image Processing*, 2021, 30: 2248-2262.
- [235] Aderghal K, Boissenin M, Benois-Pineau J, et al. Classification of sMRI for AD Diagnosis with Convolutional Neuronal Networks: A Pilot 2-D+ $\epsilon$  Study on ADNI[C]//International Conference on Multimedia Modeling. 2017: 690-701.
- [236] Lei B, Chen S, Ni D, et al. Discriminative Learning for Alzheimer's Disease Diagnosis via Canonical Correlation Analysis and Multimodal Fusion[J]. *Frontiers in Aging Neuroscience*, 2016, 8: 77:1-77:17.
- [237] Fang X, Liu Z, Xu M. Ensemble of deep convolutional neural networks based multi-modality images for Alzheimer's disease diagnosis[J]. *IET Image Processing*, 2020, 14(2): 318-326.
- [238] Xing X, Liang G, Blanton H, et al. Dynamic Image for 3D MRI Image Alzheimer's Disease Classification[C]//European Conference on Computer Vision. 2020: 355-364.
- [239] Ebrahimi A, Luo S. Convolutional neural networks for Alzheimer's disease detection on MRI images[J]. *Journal of Medical Imaging*, 2021, 8(2): 024503:1-024503:18.
- [240] Ben Ahmed O, Mizotin M, Benois-Pineau J, et al. Alzheimer's disease diagnosis on structural MR images using circular harmonic functions descriptors on hippocampus and posterior cingulate cortex[J]. *Computerized Medical Imaging and Graphics*, 2015, 44: 13-25.
- [241] Yang J, Li S, Xu W. Active Learning for Visual Image Classification Method Based on Transfer Learning[J]. *IEEE Access*, 2018, 6: 187-198.
- [242] Madusanka N, Choi H K, So J H, et al. Alzheimer's Disease classification based on multi-feature fusion[J]. *Current Medical Imaging*, 2019, 15(2): 161-169.

- [243] Liu Z, Lu H, Pan X, et al. Diagnosis of Alzheimer's disease via an attention-based multi-scale convolutional neural network[J]. *Knowledge-Based Systems*, 2022, 238: 107942:1-107942:11.
- [244] Vu T D, Ho N H, Yang H J, et al. Non-white matter tissue extraction and deep convolutional neural network for Alzheimer's disease detection [J]. *Soft Computing*, 2018, 22(20): 6825-6833.
- [245] Bi X, Li S, Xiao B, et al. Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology[J]. *Neurocomputing*, 2020, 392: 296-304.
- [246] Qu Y, Wang P, Liu B, et al. AI4AD: Artificial intelligence analysis for Alzheimer's disease classification based on a multisite DTI database[J]. *Brain Disorders*, 2021, 1: 100005:1-100005:7.
- [247] Lella E, Pazienza A, Lofù D, et al. An Ensemble Learning Approach Based on Diffusion Tensor Imaging Measures for Alzheimer's Disease Classification[J]. *Electronics*, 2021, 10(3): 249:1-249:16.
- [248] Ebadi A, Dalboni da Rocha J L, Nagaraju D B, et al. Ensemble Classification of Alzheimer's Disease and Mild Cognitive Impairment Based on Complex Graph Measures from Diffusion Tensor Images[J]. *Frontiers in Neuroscience*, 2017, 11: 56:1-56:17.
- [249] Duc N T, Ryu S, Qureshi M N I, et al. 3D-Deep Learning Based Automatic Diagnosis of Alzheimer's Disease with Joint MMSE Prediction Using Resting-State fMRI[J]. *Neuroinformatics*, 2020, 18(1): 71-86.
- [250] Bigham B, Zamanpour S A, Zare H. Features of the superficial white matter as biomarkers for the detection of Alzheimer's disease and mild cognitive impairment: A diffusion tensor imaging study[J]. *Heliyon*, 2022, 8(1): e08725:1-e08725:9.

- [251] Wen J, Thibeau-Sutre E, Diaz-Melo M, et al. Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation[J]. *Medical Image Analysis*, 2020, 63: 101694:1-101694:19.
- [252] Basaia S, Agosta F, Wagner L, et al. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks[J]. *NeuroImage: Clinical*, 2019, 21: 101645:1-101645:8.
- [253] Liu P, Zhang H, Lian W, et al. Multi-Level Wavelet Convolutional Neural Networks[J]. *IEEE Access*, 2019, 7: 74973-74985.
- [254] Fujieda S, Takayama K, Hachisuka T. Wavelet Convolutional Neural Networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2018: 1-10.
- [255] Salyers J B, Dong Y, Gai Y. Continuous Wavelet Transform for Decoding Finger Movements From Single-Channel EEG[J]. *IEEE Transactions on Biomedical Engineering*, 2019, 66(6): 1588-1597.
- [256] Ananda A, Ngan K H, Karabağ C, et al. Classification and Visualisation of Normal and Abnormal Radiographs; A Comparison between Eleven Convolutional Neural Network Architectures[J]. *Sensors*, 2021, 21(16): 5381:1-5381:21.
- [257] Sørensen L, Nielsen M. Ensemble support vector machine classification of dementia using structural MRI and mini-mental state examination[J]. *Journal of Neuroscience Methods*, 2018, 302: 66-74.
- [258] Dimitriadis S I, Liparas D, Tsolaki M N. Random forest feature selection, fusion and ensemble strategy: Combining multiple morphological MRI measures to discriminate among healthy elderly, MCI, cMCI and Alzheimer's disease patients: From the Alzheimer's disease neuroimaging initiative (ADNI) database[J]. *Journal of Neuroscience Methods*, 2018, 302: 14-23.

- [259] Kumar T, Sharma P, Prakash N. Comparison of Machine learning models for Parkinson's Disease prediction[C]//IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference. 2020.
- [260] 陈阿梅. 基于 DTI 的早期 PD 伴快动眼睡眠行为障碍神经机制研究及 PD 疾病进展预测模型构建[D]. 暨南大学, 2022.
- [261] Agosta F, Weiler M, Filippi M. Propagation of pathology through brain networks in neurodegenerative diseases: from molecules to clinical phenotypes[J]. *CNS Neuroscience and Therapeutics*, 2015, 21(10): 754-767.
- [262] Post B, Speelman J D, De Haan R J. Clinical heterogeneity in newly diagnosed Parkinson' s disease[J]. *Journal of Neurology*, 2008, 255(5): 716-722.
- [263] Thenganatt M A, Jankovic J. Parkinson disease subtypes[J]. *JAMA neurology*, 2014, 71(4): 499-504.
- [264] Rajput A, Voll A, Rajput M, et al. Course in Parkinson disease subtypes: a 39-year clinicopathologic study[J]. *Neurology*, 2009, 73(3): 206-212.
- [265] Schwarz S T, et al. Diffusion tensor imaging of nigral degeneration in Parkinson's disease: a region-of-interest and voxel-based study at 3 T and systematic review with meta-analysis[J]. *NeuroImage: Clinical*, 2013, 3: 481-488.
- [266] Amandola M, Sinha A, Amandola M J, et al. Longitudinal corpus callosum microstructural decline in early-stage Parkinson' s disease in association with akinetic-rigid symptom severity[J]. *npj Parkinson's Disease*, 2022, 8 (1): 1-10.
- [267] Sivaranjini S, Sujatha C. Deep learning based diagnosis of Parkinson' s disease using convolutional neural network[J]. *Multimedia Tools and Applications*, 2020, 79(21): 15467-15479.

- [268] Al-Sarem M, Saeed F, Boulila W, et al. Feature selection and classification using CatBoost method for improving the performance of predicting Parkinson' s disease[M]//Advances on Smart and Soft Computing. 2021: 189-199.
- [269] Almeida J S, Rebouças Filho P P, Carneiro T, et al. Detecting Parkinson' s disease with sustained phonation and speech signals using machine learning techniques[J]. *Pattern Recognition Letters*, 2019, 125: 55-62.
- [270] Vaiciukynas E, Verikas A, Gelzinis A, et al. Detecting Parkinson' s disease from sustained phonation and speech signals[J]. *PloS One*, 2017, 12(10): e0185613:1-e0185613:16.
- [271] Nahar N, Ara F, Neloy M, et al. Feature Selection Based Machine Learning to Improve Prediction of Parkinson Disease[C]//International Conference on Brain Informatics. 2021: 496-508.
- [272] Patra A K, Ray R, Abdullah A A, et al. Prediction of Parkinson' s disease using Ensemble Machine Learning classification from acoustic analysis[C]//Journal of Physics: Conference Series: Vol. 1372. 2019: 012041:1-012041:8.
- [273] Wang W, Lee J, Harrou F, et al. Early detection of Parkinson' s disease using deep learning and machine learning[J]. *IEEE Access*, 2020, 8: 147635-147646.
- [274] Prashanth R, Roy S D, Mandal P K, et al. High-accuracy detection of early Parkinson's disease through multimodal features and machine learning[J]. *International Journal of Medical Informatics*, 2016, 90: 13-21.
- [275] Grover S, Bhartia S, Yadav A, et al. Predicting severity of Parkinson' s disease using deep learning[J]. *Procedia Computer Science*, 2018, 132: 1788-1794.
- [276] Masud M, Singh P, Gaba G S, et al. CROWD: crow search and deep learning based feature extractor for classification of Parkinson' s disease [J]. *ACM Transactions on Internet Technology*, 2021, 21(3): 1-18.

- [277] Pahuja G, Prasad B. Deep learning architectures for Parkinson's disease detection by using multi-modal features[J]. *Computers in Biology and Medicine*, 2022, 146: 105610:1-105610:10.
- [278] Kuang B, Tekin Y, Mouazen A M. Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content[J]. *Soil and Tillage Research*, 2015, 146: 243-252.
- [279] Hill-Burns E M, Debelius J W, Morton J T, et al. Parkinson's disease and Parkinson's disease medications have distinct signatures of the gut microbiome[J]. *Movement Disorders*, 2017, 32(5): 739-749.
- [280] Nilashi M, Ibrahim O, Ahmadi H, et al. A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques[J]. *Biocybernetics and Biomedical Engineering*, 2018, 38(1): 1-15.
- [281] Salmanpour M R, Shamsaei M, Saberi A, et al. Machine learning methods for optimal prediction of motor outcome in Parkinson's disease[J]. *Physica Medica*, 2020, 69: 233-240.
- [282] Wang M, Li Z, Lee E Y, et al. Predicting the multi-domain progression of Parkinson's disease: a Bayesian multivariate generalized linear mixed-effect model[J]. *BMC Medical Research Methodology*, 2017, 17(1): 147.
- [283] Bi X a, Hu X, Xie Y, et al. A novel CERNNE approach for predicting Parkinson's Disease-associated genes and brain regions based on multimodal imaging genetics data[J]. *Medical Image Analysis*, 2021, 67: 101830:1-101830:12.
- [284] Lei H, Huang Z, Zhang J, et al. Joint detection and clinical score prediction in Parkinson's disease via multi-modal sparse learning[J]. *Expert Systems with Applications*, 2017, 80: 284-296.

- [285] Kiryu S, Yasaka K, Akai H, et al. Deep learning to differentiate Parkinsonian disorders separately using single midsagittal MR imaging: a proof of concept study[J]. *European Radiology*, 2019, 29(12): 6891-6899.
- [286] Adams M P, Rahmim A, Tang J. Improved motor outcome prediction in Parkinson's disease applying deep learning to DaTscan SPECT images[J]. *Computers in Biology and Medicine*, 2021, 132: 104312:1-104312:8.
- [287] Khoury N, Attal F, Amirat Y, et al. Data-driven based approach to aid Parkinson's disease diagnosis[J]. *Sensors*, 2019, 19(2): 242:1-242:27.
- [288] Khoury N, Attal F, Amirat Y, et al. CDTW-based classification for Parkinson's Disease diagnosis[C]//European Symposium on Artificial Neural Networks. 2018: 621-626.
- [289] Prince J, Andreotti F, De Vos M. Multi-source ensemble learning for the remote prediction of Parkinson's disease in the presence of source-wise missing data[J]. *IEEE Transactions on Biomedical Engineering*, 2018, 66(5): 1402-1411.
- [290] Naranjo L, Perez C J, Campos-Roca Y, et al. Addressing voice recording replications for Parkinson's disease detection[J]. *Expert Systems with Applications*, 2016, 46: 286-292.
- [291] Sharanyaa S, Sambath M, Renjith P. Optimized Deep Learning for the Classification of Parkinson's Disease Based on Voice Features[J]. *Critical Reviews™ in Biomedical Engineering*, 2022, 50(5): 1-28.
- [292] Oh S L, Hagiwara Y, Raghavendra U, et al. A deep learning approach for Parkinson's disease diagnosis from EEG signals[J]. *Neural Computing and Applications*, 2020, 32: 10927-10933.
- [293] Anisha C, Arulanand N. Early Prediction of Parkinson's Disease (PD) Using Ensemble Classifiers[C]//International Conference on Innovative Trends in Information Technology. 2020: 1-6.

- [294] Lei H, Huang Z, Elazab A, et al. Longitudinal and multi-modal data learning via joint embedding and sparse regression for Parkinson's disease diagnosis[C]//International Workshop on Machine Learning in Medical Imaging. 2018: 310-318.
- [295] Chougar L, Faouzi J, Pyatigorskaya N, et al. Automated categorization of Parkinsonian syndromes using magnetic resonance imaging in a clinical setting[J]. *Movement Disorders*, 2021, 36(2): 460-470.
- [296] Shin N Y, Bang M, Yoo S W, et al. Cortical thickness from MRI to predict conversion from mild cognitive impairment to dementia in Parkinson disease: a machine learning-based model[J]. *Radiology*, 2021, 300(2): 390-399.
- [297] Wee C Y, Yap P T, Li W, et al. Enriched white matter connectivity networks for accurate identification of MCI patients[J]. *Neuroimage*, 2011, 54(3): 1812-1822.

## 简称列表

英文简称	中文含义及英文全称
AI	人工智能 (Artificial intelligence)
DL	深度学习 (Deep Learning)
AD	阿尔茨海默症 (Alzheimer's disease)
PD	帕金森疾病 (Parkinson's disease)
CT	计算机断层成像 (Computed Tomography)
MRI	磁共振成像 (magnetic Resonance imaging)
SPECT	单光子发射计算机断层显像 (single-photon emission computed tomograph)
PET	正电子发射型计算机断层显像 (Positron Emission Tomography)
fMRI	功能性磁共振成像 (functional magnetic resonance imaging)
sMRI	结构磁共振成像 (structural Magnetic Resonance Imaging)
X-Ray	X 射线 (X Radiography)
Microscopy	显微影像 (Microscopic Imaging)
US	超声成像 (Ultrasound)
DTI	弥散张量 (Diffusion Tensor Imaging)
DAT	多巴胺转运蛋白 (Dopamine transporter)
NC	正常认知 (Normal Cognition)
EMCI	早期轻度认知障碍 (Early Mild Cognitive Impairment)
LMCI	晚期轻度认知障碍 (Late Mild Cognitive Impairment)
SOTA	最前沿的技术水平 (state-of-the-art)
NSCT	非子采样轮廓波变换 (Non-Subsampled Contourlet Transform)
CNN	卷积神经网络 (Convolutional Neural Network)
GAN	生成对抗网络 (Generative Adversarial Network )

英文简称	中文含义及英文全称
DDcGAN	带有双重鉴别器的条件式生成对抗网络 (Dual-Discriminator Conditional Generative Adversarial Network)
ROI	感兴趣区域 (Region of Interest)
SVM	支持向量机 (support vector machine)
DBM	玻尔兹曼机 (Deep Boltzmann Machine)
GLCM	灰度共生矩阵 (Gray-Level Co-occurrence Matrix)
SCD	主观认知下降 (Subjective Cognitive Decline)
VBM	基于体素的形态测量学 (voxel-based morphometry)
PCA	主成分分析 (Principal Component Analysis)
rs-fMRI	静息态功能磁共振成像 (Resting-state functional magnetic resonance imaging)
PDD	PD 患者常伴有帕金森病痴呆 (Parkinson's disease dementia)
MoCA	蒙特利尔认知评估 (Montreal Cognitive Assessment)
DRS-2	马蒂斯痴呆分级量表 (Mattis Dementia Rating Scale, 2nd Edition)
MMSE	简易的精神状态测评表 (Mini-Mental State Examination)
SOM	自组织映射 (Self-Organizing Map)
RMSE	均方根误差 (Root Mean Square Error)
R2	决定系数 (Coefficient of Determination)
MAE	平均绝对误差 (Mean Absolute Error)
MSE	均方误差 (Mean Squared Error)
RMSE	均方根误差 (Root Mean Squared Error)
MAPE	平均绝对百分比误差 (Mean Absolute Percentage Error)
EN	信息熵 (Entropy)
MI	互信息 (Mutual Information)
FMI	特征互信息 (Feature Mutual Information)
$Q_E$	边缘相关融合质量 (Edge Correlation Fusion Quality)

## 简称列表

---

英文简称	中文含义及英文全称
AG	平均梯度 (Average Gradient)
SSIM	结构相似性 (Structural Similarity Index)
SD	标准差 (Standard Deviation)
SF	空间频率 (Spatial Frequency)
rSFe	空间频率误差比 (the ratio of spatial frequency error )
VIF	视觉信息保真度 (Visual Information Fidelity)
CC	相关系数 (Correlation Coefficient)
ACC	准确率 (Accuracy)
ROC	受试者工作特征曲线 (Receiver Operating Characteristic)
AUC	曲线下面积 (Area Under the Curve)
SEN	敏感性 (Sensitivity)
SPE	特异性 (Specificity)
TP	真正例 (True Positive)
TN	真负例 (True Negative)
FP	假正例 (False Positive)
FN	假负例 (False Negative)
MS-Info	医学语义信息 (medical semantic information)
WCU	小波卷积单元 (Wavelet Convolution Unit)
WM	白质 (White matter)
BN	批归一化 (Batch Normalization)
FA	分数各向异性 (Fractional Anisotropy)
MD	平均扩散系数 (Mean Diffusivity)
CSF	脑脊髓液 (Cerebro spinal fluid)
RNA	核糖核酸 (ribonucleic acid)
REM	眼皮快速颤动 (arapid eye movement)
UCI	University of CaliforniaIrvine
PPMI	The Parkinson' s Progression Markers Initiative
GMM	高斯混合模型 (Gaussian mixture models)
OPF	优化路径森林 (optimum-path forest)
DNN	深度神经网络 (deep neural network)

英文简称	中文含义及英文全称
NB	贝叶斯 (Naive Bayes)
GaussianNB	高斯朴素贝叶斯 (Gaussian Naive Bayes)
MLP	多层感知机 (multi-layer perceptron)
KNN	K 近邻算法 (K-nearest neighbour)
LR	逻辑回归 (logistic regression)
GBT	梯度提升 (gradient boosting tree)
DBN	深度信念网络 (Deep Belief Network)
SVR	支持向量回归 (support vector regression)
RF	随机森林 (random forest)
DT	决策树 (decision tree)
ExtraTree	极端随机树 (extremely randomized trees)
XGBT	极端梯度提升 (eXtreme gradient boosting tree)
CROWD	乌鸦搜索和深度学习组合分类模型 (crow search and deep learning)
CERNNE	聚类进化随机集成神经网络 (clustering evolutionary random neural network ensemble)
vGRFs	垂直地面反作用力 (vertical ground reaction forces)
CDTW	连续动态时间归整 (continuous dynamic time warping)

## 攻读博士学位期间发表的文章目录

1. **Jinyu Wen**, Asad Khan, Amei Chen, Weilong Peng, Meie Fang, Ping Li and C. L. Philip Chen. High-Quality Fusion and Visualization for MR-PET Brain Tumor Images via Multi-Dimensional Features[J]. *IEEE Transactions on Image Processing*, May 2024, published online, DOI: 10.1109/TIP.2024.3404660. (中科院一区, IF:10.6)
2. **Jinyu Wen**, Feiwei Qin, Jiao Du, Meie Fang, Ping Li, Xinhua Wei, and C. L. Philip Chen. MsgFusion: Medical Semantic Guided Two-Branch Network for Multi-Modal Brain Image Fusion[J]. *IEEE Transactions on Multimedia*, 26, 944-957, 2024. (中科院一区, IF:7.3)
3. **Jinyu Wen**, Yang Li, Meie Fang, Lei Zhu, David Dagan Feng, and Ping Li. Fine-Grained and Multiple Classification for Alzheimer's Disease with Wavelet Convolution Unit Network[J]. *IEEE Transactions on Biomedical Engineering*, 70(9), 2592-2603, 2023. (中科院二区, IF:4.6)
4. **Jinyu Wen**, Amei Chen, Meie Fang, Xinhua Wei. Computer aided prediction of Parkinson's disease progression by machine learning from cross-modal data[J]. *npj Parkinsons Disease*. (审稿中, 中科院一区, IF:8.7)
5. 温金玉, 方美娥. 基于机器学习的计算机辅助帕金森疾病分类预测研究综述 [J]. 图学学报, 44(6), 2023. (中文核心)
6. Danyang Yao, **Jinyu Wen**, Amei Chen, Meie Fang, Xinhua Wei, Zhigeng Pan. Trimodal Fusion Network Combined Global-Local Feature Extraction Strategy and Spatial-Frequency Fusion Strategy. *ML4CS* (3): 212-226, 2022. (共一, EI 会议收录)
7. 陈阿梅, 温金玉, 黄俊祥, 成小芳, 阮秀杭, 方美娥, 魏新华. 血管周围间隙扩散张量成像指数在评估帕金森病患者脑类淋巴系统功能中的价值 [J]. 放射学实践, 8: 954-959, 2022. (中文核心)

8. Jipeng Hu, **Jinyu Wen**, Meie Fang. A survey on adversarial attack and defense of deep learning models for medical image recognition[J]. *Metaverse* 2023; 4(1): 17, 2023.
9. Asad Khan, Sakander Hayat, Muhammad Ahmad, **Jinyu Wen**, Muhammad Umar Farooq, Meie Fang, Wenchao Jiang. (2022). Cross-modal retrieval based on deep regularized hashing constraints[J]. *International Journal of Intelligent Systems*, 37(9), 6508-6530. (中科院二区, IF:7)
10. 方美娥, 姚丹阳, 温金玉. 一种基于深度学习的多模态医学图像融合方法. 2022-08-05, ZL 202210936228.6

## 攻读博士学位期间参与的项目目录

1. 主持广州大学研究生“基础创新”项目《基于张量计算的多模态医学图像脑网络构建及 AD/MCI 细粒度分类诊断应用研究》（项目编号：2021GDJC-D16）
2. 参与国家自然科学基金项目《CAGD 与 CAD/CAE/AI 深度融合的类 B 样条等几何方法研究》（项目号：62072126）
3. 参与广州市重点研究计划项目《基于类 B 样条的等几何边界元分析及形状优化》（项目号：202102010439）

## 致 谢

白驹过隙，时光匆匆。行文至此，多年的学生时代即将谢幕，回首过去，仿佛一场梦。四年的博士生活在紧张而充实的科研中默默度过，回顾这段难忘的时光，心头留下了痛苦的泪水，也回荡着欢乐的笑声，然而更多的是对经历困难后那抹明媚阳光的深刻感受。

我清晰地记得，在开始写作的第一周里，每天都凝望着论文题目下的一片空白，愁眉不展，默许着心底涌动的渴望。随着日子一天天流逝，当我在反复修改论文框架时，心急如焚，突然想起年幼时母亲对我说的一句普通却蕴含哲理的话语。我出生于群山环抱的乡村，在高中之前时常与黄土为伍。小时候，我喜欢和小伙伴在家乡的田野间玩家家，用泥土捏出各种形状的食物，搭建简陋的灶台，甚至用泥巴堆起小墙，划定自己的领地。到了不再玩家家的年纪，每天放学后都帮助家人从事各种农活。其中最难忘的是在炎热的夏天下地处理烟杆。或许很多人对烟杆不了解，也不清楚它的意义。我的老家位于闽西北地区，大多数家庭都务农，上半年种烟叶，下半年转种水稻。没错，烟叶就是那个“吸烟有害健康”包装盒里的原材料，但它让我们的村实现了小康。种烟叶不仅能够维持生计，还能积攒一些存款。每年暑假烟叶都已基本采收结束，每到这个时候，就逃不过处理剩余烟杆，将田地改成水田以便种植水稻。回忆那时，望着一片又一片的烟田，心头充满泪水。面对未来的日子，气温只会越来越高，而烟杆处理任务还有很多。我是家中的长女，其他的弟弟妹妹们都能尽情玩耍，而我一想到这就不满，不愿再继续。直到有一天，母亲走过来，轻轻拍了拍我的肩膀。她的眼神里充满了理解和鼓励，她说：“孩子，累了就休息一下吧。每天在太阳升起之前努力工作，总能完成的。”这番话让我感到内疚。父母辛勤劳作，几乎每天都在田间辛勤劳作，年复一年地重复同样的农事，只为了让我们接受更好的教育，走出大山。回首儿时的生活，虽然没有芭比娃娃、琴棋书画等多种多样的玩具和兴趣爱好，但那个时候的生活中充满了勤劳和淳朴的气息。从回忆起母亲的话开始，我努力克制内心的不安，每天不少于8小时呆在电脑前看论文和写论文，累了就休息。尽管这期间难免感到孤独，但每天都过得充实。如今，我已经走出了大山，来到了城市求学。每当我想起儿时的生活和母亲的教诲，我都会感到一股无形的力量在支撑着我前进。我知道，无论未来遇到多少困难和挑战，只要我坚持不懈地努力下去，就一定能够实现自己的梦想。

我是一个性格介于内向和外向之间的人，在公共场合也会显得胆怯。踏入广州大学求学生涯，可谓是我的幸运之选，因为在这里，我得以认识了一群非凡的才华横溢之人。在即将毕业的时刻，我首先要感谢母校为我提供的培养，母校为我的科研学习提供了一片滋润的土壤，更使我明白了“博学笃行，与时俱进”不仅是一种态度，更是一种科研方法。我要向我的导师方美娥教授表达我最崇高的敬意和最真挚的谢意！在攻读博士学位的四年里，方老师为我提供了宽松与自由的学术研究环境，同时以一丝不苟、精益求精的学术风范和认真严谨的治学态度教会我如何做科研。正是导师的严格要求和在论文撰写过程中的耐心指导，我才得以在科研的海洋中乘风破浪，逐渐成长。因此，我真诚地向方美娥教授表达我的敬意和谢意，感谢她在这四年中对我的栽培和帮助，这将是我一生中最珍贵的记忆。

感谢这四年多来与我并肩作战、共同探索知识的同窗们：法代东、李聪、杨逸伦、胡济鹏、姚丹阳、毛家辉、余柳浩、陈庆丰、张泽初、罗屹、尹俊杰、杨志豪、钟承志、周昊、解梦达、龙芳、杨家谋、鲜俊兰、黄晚桃、杨伟、魏伟明、陈泓宇。感谢你们在工作和生活中给予我的帮助与关心，使得我的科研生活既充实又充满乐趣。希望在再见之际，我们都能保持那份青春的热情与活力。此外，衷心感谢友谊团队尤其是魏新华医师和陈阿梅医师，在医学数据分析和临床难题解答上给予的大力支持和专业指导。

感谢多年来家人在我的学业和生活中所给予的重要支持和帮助，特别是要感谢我的先生——李玉琪。感谢你十年之久的陪伴，无论是顺风顺水还是逆境挫折，你始终如一地陪伴在我的左右。感谢在我攻读博士期间对我无尽的包容和理解，让我得以全身心投入学业和科研。是你的支持、关爱和鼓励，让我在面对困难时更有信心去克服。感谢你毫不犹豫地成为我坚强的后盾，让我可以轻松愉快地渡过博士学习生活。

同时，我也要感谢自己在每次遇到挫折时，我都能勇敢地站起来，继续前进。感谢自己一直坚守理想信念，认真对待每一项工作。更要感谢自己做出了正确的选择，引领我走向了今天的生活。

最后，感谢在百忙之中参与评阅本文和学位答辩的各位专家和老师们！

温金玉

2024年5月于广州大学