

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2025)10-3230-12

论文引用格式: Hu X Y, Shi Y and Liu Y. 2025. Classification of Alzheimer's disease using multimodal brain imaging via DenseNet and Mamba. Journal of Image and Graphics, 30(10):3230-3241(胡馨悦, 石雨, 刘羽. 2025. 结合 DenseNet 和 Mamba 的多模态脑影像阿尔茨海默症分类. 中国图象图形学报, 30(10):3230-3241)[DOI:10.11834/jig.250028]

结合 DenseNet 和 Mamba 的多模态脑影像 阿尔茨海默症分类

胡馨悦¹, 石雨¹, 刘羽^{1,2*}

1. 合肥工业大学生物医学工程系, 合肥 230009; 2. 合肥工业大学测量理论与精密仪器安徽省重点实验室, 合肥 230009

摘要: 目的 阿尔茨海默症(Alzheimer's disease, AD)作为一种常见的老年性痴呆疾病,近年来已成为全球公共卫生面临的重大挑战,设计一种有效且精确的阿尔茨海默症早期诊断模型具有重要的临床意义和迫切需求。目前,阿尔茨海默症的临床诊断通常依赖于正电子发射断层扫描(positron emission tomography, PET)和核磁共振成像(magnetic resonance imaging, MRI)两种医学影像数据。然而,由于这两种模态间存在信息差异大、未精确配准等问题,现有的基于人工智能(artificial intelligence, AI)的诊断模型大多仅使用单一的 MRI 数据。这在一定程度上限制了多模态影像信息的充分利用和分类性能的进一步提升,制约了其临床实用性。针对上述问题,提出一种结合 DenseNet 和 Mamba 的多模态医学脑影像阿尔茨海默症早期诊断模型 DenseMamba。**方法** 该方法以经过标准预处理流程后的 PET 和 MRI 数据为输入,预处理流程包括:颅骨剥离、配准、偏置场校正和归一化。MRI 和 PET 级联后首先经过卷积层和激活层进行初步特征提取,提取到的特征再依次经过若干个交替的 Denseblock 和 TransMamba 模块分别进行局部和全局的特征提取。Denseblock 内的密集连接结构,增强了局部特征的提取和传播,从而能够捕捉影像中的细节信息;而 TransMamba 模块则基于状态空间模型,高效地建模全局依赖关系。交替的 Denseblock 和 TransMamba 使得模型能够更全面地理解多模态数据信息,充分挖掘多模态数据在临床诊断任务上的潜力。最后,将提取到的特征送入分类器得到疾病预测结果。**结果** 为验证方法的有效性,在公开的 ADNI(Alzheimer's disease neuro-imaging initiative)数据集上进行了评估。最终模型的准确率(accuracy)、精确度(precision)、召回率(recall)和 F1 值分别为 92.42%、92.5%、92.42% 和 92.21%。DenseMamba 在阿尔茨海默症分类任务中较对比算法表现优异,比次优算法准确率提升 0.42%。**结论** 实验结果表明,DenseMamba 能够充分发挥 PET 和 MRI 数据的潜力,显著提升分类性能,为阿尔茨海默症的早期诊断提供更精准的支持。

关键词: 阿尔茨海默症(AD);多模态医学图像;状态空间模型;Mamba;密集连接神经网络

Classification of Alzheimer's disease using multimodal brain imaging via DenseNet and Mamba

Hu Xinyue¹, Shi Yu¹, Liu Yu^{1,2*}

1. Department of Biomedical Engineering, Hefei University of Technology, Hefei 230009, China;

2. Anhui Province Key Laboratory of Measuring Theory and Precision Instrument, Hefei University of Technology, Hefei 230009, China

Abstract: Objective Alzheimer's disease (AD) is a common form of dementia in the elderly and has become a major chal-

收稿日期:2025-01-23;修回日期:2025-02-25;预印本日期:2025-03-04

* 通信作者:刘羽 yuliu@hfut.edu.cn

基金项目:国家自然科学基金项目(62176081, U23A20294)

Supported by: National Natural Science Foundation of China(62176081, U23A20294)

lenge in global public health in recent years. With the acceleration of global aging, the incidence of AD continues to rise, which places tremendous pressure on medical treatment and caregiving. The main characteristic of AD is the gradual decline in cognitive functions, particularly memory, language, and decision-making abilities, which are often accompanied by irreversible neuronal loss and brain atrophy. As the disease progresses, it becomes increasingly difficult to differentiate AD from other types of dementia, which makes early diagnosis more challenging. Therefore, early diagnosis is crucial for timely intervention and personalized treatment plans, as it can effectively delay disease progression. Currently, clinical diagnosis of AD typically relies on neuropsychological assessments, biomarker detection, and neuroimaging techniques. Among these approaches, positron emission tomography (PET) and magnetic resonance imaging (MRI) are the most commonly used imaging methods to evaluate the brain structure and function of AD patients. PET provides information on brain metabolism and amyloid plaques, while MRI offers high-resolution structural images that can reveal brain atrophy and other neuroanatomical changes related to the disease. However, the two imaging methods significantly differ in various aspects, such as resolution, contrast, and spatial alignment issues. As a result, existing artificial intelligence (AI)-based diagnostic models often rely on single-modality data, particularly MRI images. This reliance limits the full utilization of multimodal information and restricts further improvements in classification performance. To address this issue, this study proposes an early diagnosis model for AD based on multimodal imaging data (PET and MRI)—DenseMamba. This model combines the advantages of both imaging modalities, which overcomes the limitations of traditional single-modality methods. The model adopts a state-space model-based framework (Mamba) to effectively extract and fuse heterogeneous modal features, which achieves more robust classification performance. **Method** The DenseMamba model uses multimodal PET and MRI imaging data as input. To effectively integrate the two types of imaging data, the model first performs initial feature extraction through a convolutional layer and an activation layer. The extracted features are then passed through several alternating Denseblock and TransMamba modules for further local and global feature extraction. The Denseblock module, with its dense connection design, enhances feature learning capability by facilitating feature reuse across layers, which effectively mitigates the vanishing gradient problem in deep neural networks. This dense connectivity helps capture complex patterns and dependencies within the data. By contrast, the TransMamba module focuses on capturing global feature representations, modeling long-range dependencies within the data, and effectively integrating cross-space and -time related information. After multiple rounds of local and global feature extraction, the final features are processed by a classification head (composed of fully connected layers and a softmax output layer) to perform the multi-class classification task. The final output is the predicted category for each input image, which indicates whether the patient is classified as healthy control, mild cognitive impairment, or AD. **Result** To validate the effectiveness of the DenseMamba model, extensive experiments were conducted using the publicly available Alzheimer's disease neuroimaging initiative (ADNI) dataset. This dataset includes multimodal imaging data (PET and MRI), and the imaging data have been preprocessed to ensure alignment between PET and MRI images. Experimental results show that the DenseMamba model significantly outperforms traditional single-modality methods; specifically, it achieves an accuracy of 92.42%, precision of 92.5%, recall of 92.42%, and an F1 score of 92.21%. **Conclusion** DenseMamba provides an efficient and accurate solution for the early diagnosis of AD. By combining the complementary advantages of PET and MRI, DenseMamba overcomes some of the limitations of traditional single-modality methods, such as insufficient feature representation and data misalignment issues. Experimental results demonstrate that DenseMamba performs excellently on the ADNI dataset, which significantly improves classification performance and far surpasses existing single-modality imaging classification methods. This approach exhibits strong clinical application potential, which provides precise support for the early diagnosis of AD and facilitates earlier interventions and personalized treatments. In the future, further improvements to the DenseMamba model may include introducing advanced fusion techniques and expanding the dataset to include a broader patient population and additional imaging modalities. These enhancements will further improve the generalization ability and clinical applicability of the model, which offer more effective tools for the early diagnosis and treatment of AD and other neurodegenerative diseases.

Key words: Alzheimer's disease (AD); multimodal medical images; state-space model; Mamba; dense connectivity neural network

0 引言

阿尔茨海默症(Alzheimer's disease, AD)作为最常见的神经退行性疾病之一,其特征性的记忆力减退、认知功能障碍和行为改变对患者的生活质量构成了严重影响。随着全球人口老龄化的加速,AD的发病率不断攀升,对社会和家庭造成了沉重的负担。当前的AD诊断由于患者症状表现微弱、检测过程漫长且过度依赖医生经验,容易导致误诊,尤其是在轻度认知障碍阶段。轻度认知障碍(mild cognitive impairment, MCI)是介于正常人和阿尔茨海默症患者间的过渡状态,轻度认知障碍患者记忆认知功能出现问题但无明显痴呆表现,日常生活活动未受明显影响。尽早对MCI患者进行干预可以稳定患者病情,延缓病程。因此,早期诊断和干预对于延缓病程、改善患者状况显得尤为重要(王斌等,2023)。

在临床实践中,阿尔茨海默症的诊断主要依赖于核磁共振成像(magnetic resonance imaging, MRI)和正电子发射断层扫描(positron emission tomography, PET)两种影像学检查方法。MRI通过提供高分辨率的脑部结构影像,能够有效检测到大脑萎缩的区域,尤其是在海马和额叶等与认知功能相关的重要区域,帮助医生判断阿尔茨海默症的早期病变。PET则通过检测脑内代谢活动,揭示阿尔茨海默症早期的功能性变化,尤其是葡萄糖代谢的减少。PET扫描还能够通过标记 β -淀粉样蛋白沉积,进一步辅助确认阿尔茨海默症的病理特征。因此,MRI主要用于提供结构性信息,而PET则帮助揭示脑功能变化和病理特征,两者结合能够提供更为全面的诊断依据。

在医学影像分析领域,基于深度学习的医学影像智能计算是目前智慧医疗领域的研究热点,其中深度学习的方法已经应用于医学影像处理、分析的全流程。施俊等人(2020)介绍了深度学习在X射线、超声、计算机断层扫描(computed tomography, CT)和MRI四大影像技术中的应用现状,涵盖了多种分析处理任务。深度学习技术的发展为医疗诊断带来了新的曙光,尤其为包括阿尔茨海默症在内的神经退行性疾病的诊断开辟了新的道路。曾安等人(2019)提出一种基于卷积神经网络(convolutional neural network, CNN)和集成学习的多切片集成分类网络用

于该疾病的早期分类诊断。邓爽等人(2022)提出基于MRI切片的弱监督细粒度AD分类模型。Aderghal等人(2017)对MRI影像从横截面、纵截面和剖面3个维度进行切片后,通过3个通道的卷积神经网络模型进行特征抽取和分类。然而,大脑是一个高度复杂的器官,其功能和结构在不同脑区之间存在复杂的联系。2D切片无法充分反映这种复杂性,且切片过程可能会丢失大脑的三维空间信息,特别是对于大脑这种复杂结构的器官来说,空间信息的丢失可能会影响对疾病特征的全面理解,从而限制模型分类性能的进一步提升。

随着计算性能的不不断提升和深度学习技术的不断发展,基于人工智能(artificial intelligence, AI)的阿尔茨海默症早期诊断研究逐渐从传统的2D切片医学图像向3D医学图像数据转变。与2D切片相比,3D医学图像包含更加丰富的疾病信息,有助于提取更全面的空间特征,从而进一步提升诊断性能。Kruthika等人(2019)提出一个结合3D胶囊网络、3D-CNN和预训练的3D自动编码器技术的CBIR(content-based image retrieval)系统用于AD的检测。Xia等人(2020)结合了3D-CNN和3D卷积长短期记忆网络用于从sMRI中提取特征并进行AD检测。Billones等人(2016)在VGGNet(Visual Geometry Group net)的基础上改进网络,在每个池化层后加入dropout层以减少过拟合。Liu等人(2019)提出一种多任务多通道卷积神经网络用于阿尔茨海默症分类。Pan等人(2021)提出一个3D-CNN-SE(3D-convolutional neural network-squeeze-and-excitation)模块来提取3D医学图像的空间信息。Lian等人(2022)提出一种新颖的注意力引导的框架,用于从全脑sMRI中提取多级判别性特征,以进行阿尔茨海默症的诊断。Jin等人(2020)提出3DAN(3D attention network)模型,该模型结合残差神经网络(residual network, ResNet)和注意力机制,自动捕获脑图像中最具有区分性的位置,实现了对阿尔茨海默症更好的分类性能。洪欣等人(2023)提出的CTISS(ConvLSTM based on temporal images series slice)模型,采用两阶段脑影像分析框架,并在此基础上构建了一种分层时序卷积网络。该模型通过自适应学习率进行训练,并利用时序卷积层提取局部时序特征,随后利用双向长短时记忆网络捕捉长程时序依赖,最终通过注意力机制聚焦关键特征,从而

实现在分层大脑影像切片中提取有效的时序信息。

尽管基于 AI 的阿尔茨海默症分类方法不断涌现,其所获得的分类性能也不断提升,然而,现有的大多数研究仅仅基于单一模态的影像数据,即 MRI 或 PET。由于两者之间存在的模态差异和未配准等问题,多模态影像数据之间存在的互补信息未得到充分挖掘,基于多模态数据的方法研究尚处于起步阶段。Liu 等人(2024)提出一种名为 HAMMF(hierarchical attention-based multi-task and multi-modal fusion model)的新型深度学习模型,用于阿尔茨海默症的计算机辅助诊断。该模型通过上下文分层注意力模块(contextual hierarchical attention network, CHAN)有效融合 MRI 和 PET 多模态数据,并利用多任务学习框架同时进行分类和回归任务,从而提升诊断的准确率。Tang 等人(2024)提出一种基于 3DCNN 和改进 Transformer 的阿尔茨海默症多模态诊断模型。该模型利用 3DCNN 提取 sMRI 和 PET 的特征,并采用一种改进的 Transformer 网络逐层学习特征之间的全局相关信息。此外,通过卷积操作替换 Transformer 中的位置编码,以保留特征的空间信息并减少模型参数。Chen 等人(2023)提出一种基于注意力机制的多模态融合网络,该方法利用 3DCNN 分别提取 MRI 和 PET 的高维特征。同时,将人口统计学信息和临床数据进行向量化处理,并在此基础上融合通道注意力机制与空间注意力机制,构建一种自适应的特征融合策略,以有效整合影像数据和临床数据之间的互补信息。

目前,绝大多数分类模型依然基于传统的卷积神经网络架构,虽然 CNN 在局部特征提取方面表现出色,但这种架构的局限性在于其主要聚焦于图像的局部特征的提取。这虽然有助于捕捉细节,但却忽略了全局信息,使得模型在处理复杂模式和长距离依赖关系时存在一定的瓶颈。因此,单纯依赖卷积操作的网络往往难以充分挖掘图像中的全局信息,进而影响了分类性能的进一步提升。

针对上述问题,本研究提出一种基于 DenseNet(densely connected convolutional network)和 Mamba 的多模态医学数据阿尔茨海默症分类方法。该方法利用卷积层进行浅层特征提取,随后通过卷积神经网络与状态空间模型的联合建模,分别提取影像数据的局部细节特征和全局依赖关系。最终,基于提取的特征,通过分类器实现对疾病的有效预测。

本文的贡献主要体现在以下两个方面:1)在进行阿尔茨海默症分类时,对 MRI 和 PET 进行了标准的预处理:颅骨剥离、配准和偏置场校正。颅骨剥离有助于去除脑部无关区域,减少噪声并聚焦于大脑区域;配准可以确保不同模态的数据在空间上的对齐,增强跨模态数据整合效果,从而提升分类性能;偏置场校正则能去除图像中的低频亮度不一致问题,改善图像的质量。预处理流程通过去除噪声、消除不相关信息和扫描误差,确保不同模态图像的对齐,帮助模型提取更加准确的特征,从而最终提升分类任务的整体表现。2)在网络结构设计层面结合了 DenseNet(Huang 等,2017)和 Mamba 架构,以进一步增强多模态特征传递和全局上下文信息流动能力。针对医学数据稀缺这一问题,DenseNet 通过特征共享有效缓解了过拟合问题。针对医学图像中病灶特征细微的特点,DenseNet 的密集连接结构能够有效融合浅层和深层的特征,从而显著增强网络对细微特征的捕捉与表示能力,进而提高其在复杂医学图像任务中的学习和泛化能力,使其在医学图像分析领域更具优势。Mamba 能够有效捕捉全局和局部特征之间的长距离依赖关系,而鉴于医学图像数据的全局信息和局部信息常常具有互补性,Mamba 的加入增强了 AD 分类模型在阿尔茨海默症分类任务中的准确性和可解释性。此外,Mamba 的选择性机制与传统卷积使用固定感受野的方式不同,其能够根据输入内容自适应地关注关键特征,聚焦 MRI 中反映结构性变化的病灶区域以及 PET 中反映代谢活动的异常区域,降低图像噪声和无关信息的干扰,从而更精准地捕捉细微病理特征,提升诊断灵敏度。多模态医学数据的联合使用和 Mamba 架构的加入不仅能够提升模型的疾病分类性能,也为神经科学的智能诊断提供了新的思路。

1 数据准备

1.1 阿尔茨海默症影像数据集

本文实验使用的影像数据来自阿尔茨海默症神经影像计划 ADNI(Alzheimer's disease neuroimaging initiative)公开数据库,为了确保数据的质量和一致性,选取同一受试者的采集间隔不超过一个月的 MRI 和 PET 影像数据,共选用 631 组样本对。总体而言,收集的影像数据根据疾病状态分为 3 类:AD、

CN(cognitively normal)和MCI,共使用了219名CN、185名MCI和227名AD受试者的多模态影像数据。所使用影像数据的基本信息如表1所示。

表1 ADNI数据集中受试者的基本信息
Table 1 Basic information about subjects in the ADNI dataset

类别	样本数	年龄/岁
CN	219	72.6±7.0
MCI	185	74.3±6.0
AD	227	74.8±8.1

1.2 数据预处理

所有的MRI和PET数据都按照标准流程进行处理,具体预处理流程如图1所示,包括:1)使用FSL(FMRIB's software library)工具中的BET(brain extraction tool)软件对脑部进行颅骨及无关组织剥离,减少无关影像特征干扰,从而提高后续处理的准确性;2)使用ANTs(advanced normalization tools)软件将MRI和PET配准到MNI152_T1_1mm模板上(Wang等,2019),以消除不同个体间的全局线性差异,包括全局平移、比例和旋转差异,使得图像在空间上对齐,便于后续的多模态分析;3)将配准后的图像数据使用ANTs软件的N4BiasField Correction函数进行偏置场校正,去除由磁场不均匀性引起的低频亮度不一致,增强图像的均匀性和可比性;4)将MRI图像和PET进行Max-Min归一化,消除多模态特征之间的像素值数量级差异。最终,经过预处理得到的图像数据尺寸为182×218×182。

图2展示了MRI和PET图像在不同预处理阶段的效果。经过预处理后,脑部无关区域已被有效去

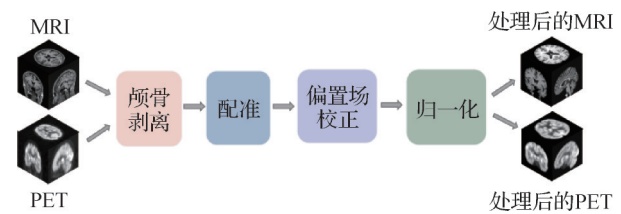


图1 数据预处理流程

Fig. 1 Data preprocessing process

除,图像仅保留了大脑区域的相关信息,背景噪声显著减少。这一过程确保了大脑组织有效提取,为后续分析提供了更加清晰的图像。此外,两个模态(MRI和PET)的数据在空间上也得到了准确对齐,确保了同一大脑区域在不同模态图像中的一致性。这种空间对齐有效消除了由于扫描角度或位置差异导致的误差,使得多模态数据能够更好地进行比较和整合,从而提高分析的准确性和可靠性。

2 算法描述

所提出的DenseMamba网络架构如图3所示。首先,MRI和PET影像按照标准流程处理后进行级联作为模型的输入,之后经过一个卷积层和激活层进行初步特征提取,提取到的特征再依次经过若干个交替的Denseblock和TransMamba模块分别进行局部和全局的特征提取,最后将特征送到分类器进行预测得到分类结果。

2.1 网络结构

CNN网络只能关注图像中的局部特征,而Mamba具有长距离依赖特征的捕捉能力,能够有效地提升网络对全局信息的理解,这意味着CNN和

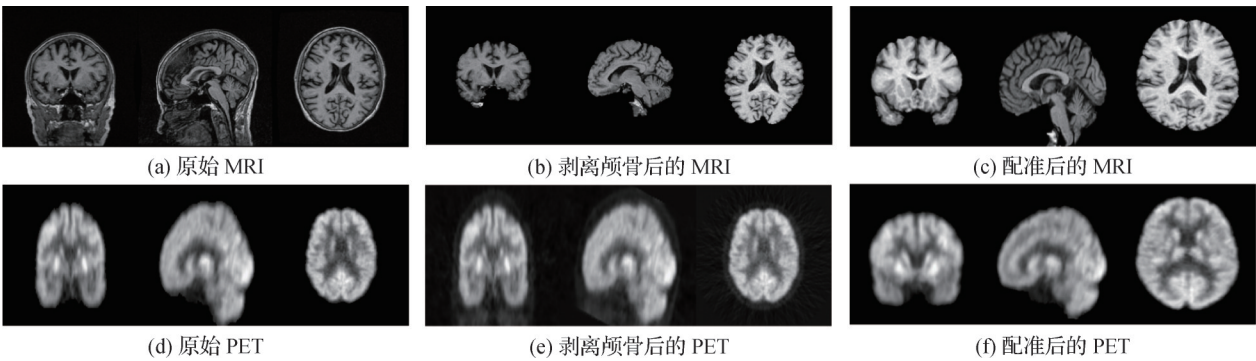


图2 预处理流程效果展示

Fig. 2 Preprocessing results display ((a) original MRI;(b) MRI after skull dissection;(c) MRI after registration;(d) original PET;(e) PET after skull dissection;(f) PET after registration)

Mamba 在特征学习任务上具有潜在的互补性。DenseMamba 整合了 DenseNet 和 Mamba 在特征学习方面的优势,且本文方法设计中使用的 3D 网络模型架构能够更好地保留多模态医学影像中的空间信息,能够更好地识别 AD 患者多模态影像中的解剖标志和病理特征。

图 3 展示了 DenseMamba 的具体结构, MRI 和 PET 数据经过级联后首先经过一个卷积核为 $7 \times 7 \times 7$ 的卷积层,较大的卷积核能够在更大范围内提取特征,增加感受野,从而捕捉到图像中更丰富的上下

文信息与大尺度的结构和模式。提取到的初步特征接着经过若干个交替的 Denseblock 和 TransMamba 模块进一步进行特征细化。具体而言, DenseMamba 包含 4 个 Denseblock 和 3 个 TransMamba 模块。Denseblock 通过其密集连接的设计,能够有效地增强特征传递,而 TransMamba 则结合了卷积神经网络提取局部特征和状态空间模型可以提取全局特征的优势,进一步协同处理,以获得更具判别性的特征表示。最后输出特征经过全连接层将卷积得到的特征映射输出,并通过 softmax 层输出 3 个类别的预测概率。

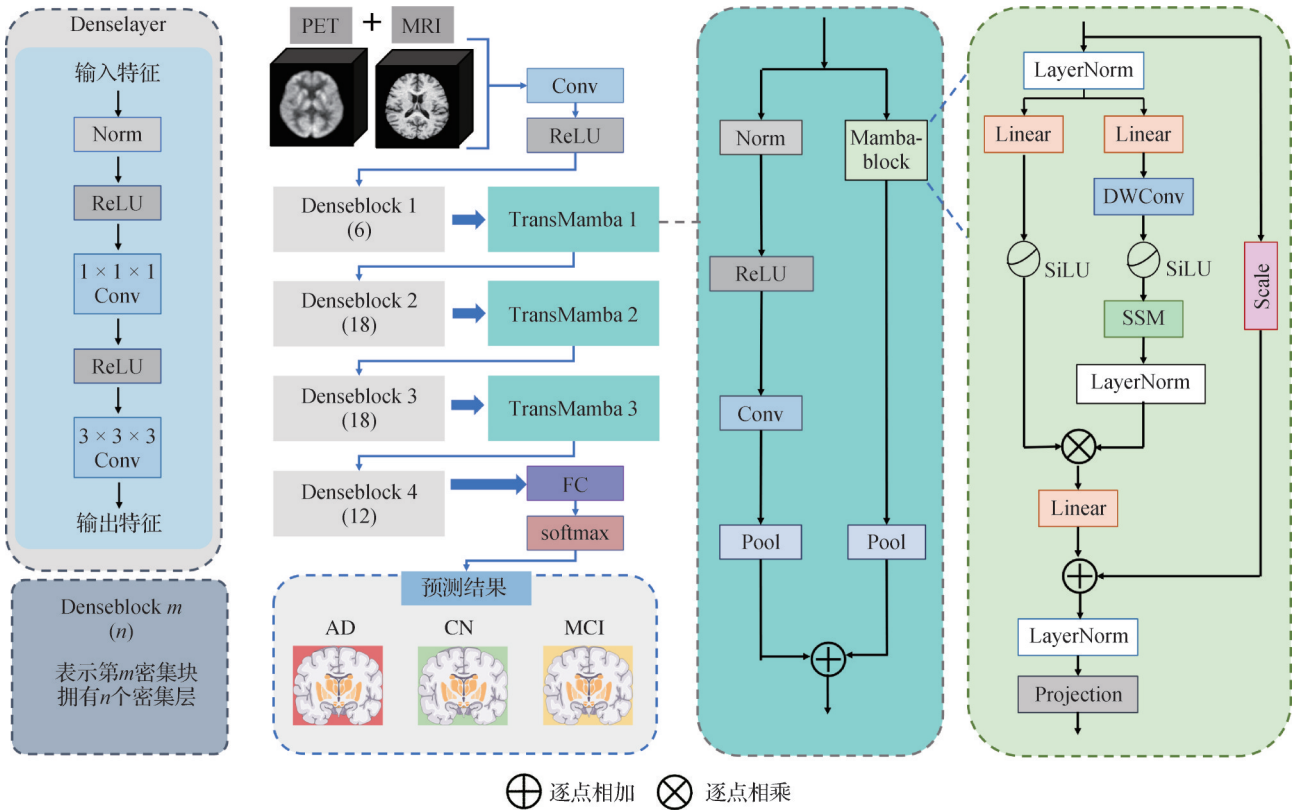


图 3 DenseMamba 架构示意图

Fig. 3 DenseMamba architecture diagram

2. 1. 1 Denseblock

每个 Denseblock 内部包含若干个 Denselayer, 具体而言, 网络中 4 个 Denseblock 分别包含 6、18、18 和 12 个 Denselayer。每个 Denselayer 由归一化层 (Norm)、激活函数 (ReLU) 和卷积层 (Conv) 组成。关键在于, Denselayer 之间采用密集连接方式: 每个 Denselayer 的输入不仅包含前一个 Denselayer 的输出, 还包含前面所有层的特征映射, 这种密集连接方式使得特征在网络中得到充分复用, 显著提升了特征表示的丰富性。

2. 1. 2 TransMamba

TransMamba 包含并行的两个分支, 如图 3 中浅绿色框图所示。第 1 个分支包含 1 个归一化层、1 个 ReLU 激活函数、1 个卷积层、1 个池化层 (Pool); 第 2 个分支包含 1 个 Mambablock 和 1 个池化层。

经过 Denseblock 块提取到的特征 M_{in} 作为 TransMamba 的输入并行经过两个分支后融合。在第 1 个分支, M_{in} 首先通过归一化层对特征进行归一化处理, 以加速训练过程, 接着应用 ReLU 激活函数, 增强模型对复杂特征的捕捉能力, 然后通过卷积层进

一步提取局部特征,最后通过池化层降低特征维度,减少计算量并抑制过拟合。 M_{in} 在第2个分支先经过 Mambablock 的处理,而后通过池化层。两个分支的处理结果合并得到最终融合的输出特征 M_{out} 。具体为

$$M_{out} = Pool(Conv(ReLU(Norm(M_{in})))) + Pool(Mambablock(M_{in})) \quad (1)$$

式中, M_{in} 表示 TransMamba 模块的输入。

状态空间模型(state space model, SSM)是一种用于描述动态系统状态随时间演变的数学模型。SSM通过一组矩阵和状态变量来描述系统如何随时间推进,通常被表述为线性常微分方程,具体为

$$h'(t) = Ah(t) + Bx(t) \quad (2)$$

$$y(t) = Ch(t) + Dx(t) \quad (3)$$

式中, A 是状态转移矩阵,描述了系统状态随时间的变化, B 是输入矩阵,描述了外部输入对系统状态的影响, C 是输出矩阵,描述了系统状态如何影响输出, D 是传递矩阵,描述了输入直接对输出的影响。

Mamba 是一种用于时间序列建模的模型,通过利用选择性状态空间模型,Mamba 能够基于内容的重要性动态地筛选与处理信息,实现对输入部分的选择性记忆或忽略。这一机制使得 Mamba 在处理速度和扩展性方面尤其针对长序列数据表现出显著的优势。

Mambablock 的具体结构如图 3 右侧部分所示,通过融合残差连接和调整因子,在不增加模型参数和计算复杂度的前提下,显著提升了 SSM 在长距离空间建模方面的性能。

对于输入特征 W_{in} ,Mambablock 首先进行归一化处理,然后将其送入两个独立处理分支并行处理。在第 1 个分支,首先利用线性层(Linear)对特征通道进行扩展,随后进行深度可分离卷积(depthwise separable convolution, DWConv)和 SiLU (sigmoid-weighted linear unit)激活,紧接着是 SSM 和归一层。具体为

$$W_1 = LayerNorm(SSM(SiLU(DWConv(Linear(W_{in})))))) \quad (4)$$

式中, W_{in} 是 Mambablock 的输入特征。DWConv 通过将标准卷积操作分解为深度卷积和逐点卷积来减少计算量和模型参数。SSM 通过状态变量和观测变量来描述系统的动态行为,其将输入特征映射到一个

连续的状态空间中,进而通过状态转移矩阵和观测矩阵来捕捉特征之间的动态关系。在第 2 个分支中,特征通道同样通过线性层扩展后直接应用 SiLU 激活,具体为

$$W_2 = SiLU(Linear(W_{in})) \quad (5)$$

式中, W_{in} 是 Mambablock 的输入特征。然后将这两个分支的输出特征 W_1 和 W_2 通过 Hadamard 乘积融合得到 W'_{out} ,Hadamard 乘积是一种逐元素相乘的操作,即对输入特征图中每个对应位置的元素进行乘法运算。它不同于常规的矩阵乘法或卷积运算,与传统的特征拼接或加法方法相比,Hadamard 乘积具有计算简单、高效且参数量较少的优点,避免了额外的复杂度和计算负担。更为重要的是, W_2 可以被视为一个动态的注意力掩码,它能够根据分支之间的相互作用,自动对 W_1 中的各个特征分量进行缩放,从而实现对特征的选择。通过这种方式,Hadamard 乘积不仅保留了各分支特征的独特信息,还有效促进了特征间的互补性,有效地结合两个分支的信息,增强模型对重要特征的捕捉能力(Wu 等,2025)。

经过线性层处理后的 W'_{out} 与经由调整因子 $s \in \mathbf{R}$ 缩放后的 W_{in} 进行残差连接,连接结果随后通过归一层进行规范化,最终通过投影层(Projection)转换得到更深层的特征表示,以便在后续网络层中继续处理(Wu 等,2025)。具体为

$$W_{out} = projection(LayerNorm(W'_{out} + s \times W_{in})) \quad (6)$$

式中, $projection$ 表示投影层,投影层本质上是一种线性变换,其目的是将输入向量从一个维度空间映射到另一个维度空间。更具体地,一个投影层由一个权重矩阵 W 和一个偏置向量 b 组成,具体为

$$y = Wx + b \quad (7)$$

式中, x 是输入向量, y 是输出向量。权重矩阵 W 决定了输入向量的各个维度如何组合以形成输出向量,而偏置向量 b 则提供了一个可学习的偏移量。通过调整权重矩阵和偏置向量,投影层可以学习到输入数据中最相关的特征表示。

投影层旨在实现特征的有效提取和转换。具体而言,在经过归一化处理和残差连接之后,投影层负责将特征向量映射到更适于后续网络层操作的维度空间,从而优化整体信息的传递效率。此外,通过学习权重矩阵的参数,投影层能够提取输入中的深层特征,将原始特征转化为更具代表性的表示形式。

2.2 损失函数

交叉熵损失函数(cross-entropy loss)是多分类任务中广泛应用的一种损失度量,其通过衡量模型预测的类别分布与真实标签之间的差异来评估分类性能。交叉熵损失函数具体为

$$L = \frac{1}{N} \sum_i L_i = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (8)$$

式中, M 表示类别的数量, y_{ic} 表示符号函数, 如果样本 i 的真实类别等于 c 取 1, 否则取 0。 p_{ic} 表示观测样本 i 属于类别 c 的预测概率。

3 实验结果与分析

3.1 实验设置

实验硬件设备包括: Intel Core i9-10900X CPU 和 NVIDIA RTX A6000 GPU; 软件配置方面, 操作系统为 64 位 Ubuntu20.04.6 LTS, 深度学习框架选用 PyTorch, 编程语言为 Python 3.10。

实验中, 将预处理过的 ADNI 数据集按 7:2:1 的比例分为训练集、验证集和测试集。训练过程中迭代次数设置为 200, 批次数设置为 6, 学习率一开始设置为 0.0001, 分别在迭代次数为 100 和 180 时按照衰减因子 0.1 进行衰减, 采用 Adam 优化器, 采用准确度、精确率、召回率和 F1 分数等指标对分类性能进行评价。

3.2 实验结果

DenseMamba 模型旨在精准区分阿尔茨海默症的 3 种状态: 正常认知、轻度认知障碍 (MCI) 和阿尔茨海默症 (AD)。为评估其性能, 实验采用分类准确率 (accuracy)、精确率 (precision)、召回率 (recall) 和 F1 值 (F1 score) 4 项指标, 将 DenseMamba 与多种先进的 AD 诊断方法进行比较。这些对比方法包括: Cui 等人 (2019) 的增强型 Inception (V3) 网络、Liu 等人 (2019) 的 3DCNN、Puente-Castro 等人 (2020) 的改进 ResNet、Venugopalan 等人 (2021) 的多模态融合分类方法、Saratxaga 等人 (2021) 的 Brain3D、Kong 等人 (2022) 结合 3D 稀疏自编码器和 3D 卷积网络的方法、Mulyadi 等人 (2023) 的 XADLiME、Chen 等人 (2023) 的基于注意力机制的融合算法、Li 等人 (2024) 提出的基于改进 Med-3D 网络和注意力机制的算法以及 Hu 等人 (2025) 提出的 3D-SEConvNeXt。表 2 展示了不同方法与 DenseMamba 的实验结果,

DenseMamba 在三分类任务中表现出显著的优越性。

具体而言, DenseMamba 取得了 92.42% 的分类准确率, 比次优算法高出 0.42%。同时, DenseMamba 的精确率、召回率和 F1 分数分别达到 92.50%、92.42% 和 92.21%。准确率、召回率和 F1 分数均优于其他比较方法。精确率和召回率的高值表明, DenseMamba 在识别正确类别和避免误分类方面的表现优异, 而 F1 分数的高值则综合反映了其在精确率和召回率之间的良好平衡。

密集连接有效地增强了特征在网络中的流动性和复用性, 而 Mambablock 则能同时捕捉多尺度特征, 进而揭示 MRI 和 PET 多模态图像中的局部与全局特征之间的内在关联, 最终实现诊断精度的显著提升。

表 2 ADNI 数据集多分类 (CN vs. MCI vs. AD) 的性能比较

Table 2 Performance comparison of multiple classification on ADNI dataset (CN vs. MCI vs. AD)

方法	准确率	精确率	召回率	F1 值
Cui 等人 (2019)	0.857 0	-	-	-
Liu 等人 (2019)	0.701 0	-	-	-
Puente-Castro 等人 (2020)	0.786 4	0.688 7	0.582 8	0.603 0
Venugopalan 等人 (2021)	0.850 0	0.840 0	0.790 0	0.810 0
Saratxaga 等人 (2021)	0.880 0	-	-	-
Mulyadi 等人 (2023)	0.683 5	-	-	-
Kong 等人 (2022)	0.876 7	-	-	-
Chen 等人 (2023)	0.841 0	-	-	-
Li 等人 (2024)	0.920 0	0.930 0	0.890 0	-
Hu 等人 (2025)	0.897 1	-	-	-
本文	0.924 2	0.925 0	0.924 2	0.922 1

注: 加粗字体表示各列最优结果。“-”表示无相关数据。

3.3 消融实验及分析

为测试多模态医学图像数据联合输入、DenseNet 网络和 Mambablock 模块设计的有效性, 设计了一系列消融实验。

3.3.1 多模态医学数据的有效性验证

为了验证多模态医学图像输入的有效性, 进行了如下实验: 将分别输入单一模态的医学图像 (即 MRI 图像或 PET 图像) 的结果和同时输入两种模态图像的结果进行比较。图 4 展示了不同的输入在

DenseMamba上的性能结果。结果显示,相比于仅使用单模态数据而言,当同时输入PET和MRI两种模态时,其各项指标均取得了明显的提升。具体而言,仅输入MRI图像时的分类准确率为84.85%,仅输入PET图像时的分类准确率为80.30%,而同时输入MRI和PET图像能够提供更全面的解剖信息和功能信息,分类准确率提升至92.4%。

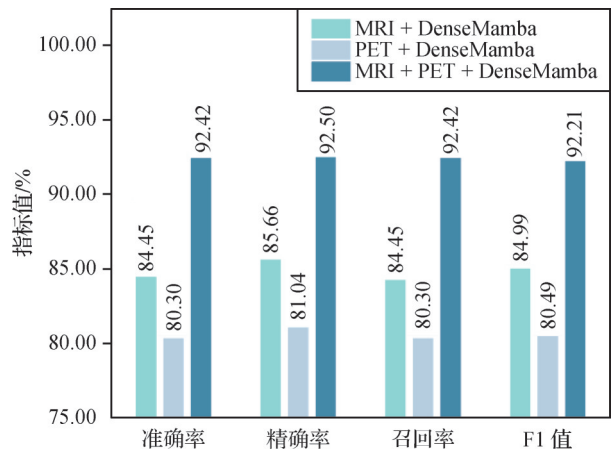


图4 不同输入的实验结果
Fig. 4 Different input experimental results

3.3.2 DenseNet 的有效性验证

为了评估不同深度学习模型在阿尔茨海默症分类任务中的性能差异,本研究对比了ResNet和DenseNet两种模型的表现。实验中,两种模型均使用MRI和PET数据的联合输入进行训练和测试。图5展示了不同模型实现的阿尔茨海默症分类结果。结果表明,相较于ResNet,DenseNet模型在各项关键指标上均展现出显著优势,实现了更优越的分类性能。具体而言,DenseNet模型在准确率、精确率、召回率以及F1 Score上分别达到89.39%、90.21%、89.39%和89.53%,而ResNet模型对应的指标分别为84.62%、87.37%、83.30%和83.65%。DenseNet在阿尔茨海默症识别的准确性和可靠性方面均优于ResNet模型。因此,DenseNet模型更适合处理阿尔茨海默症分类任务中涉及的多模态医学影像数据。此外,未来将进一步探讨DenseNet模型所重点关注的特征区域,以期提升模型的诊断能力。

3.3.3 Mambablock 的有效性验证

为进一步探究Mambablock对网络性能的影响,将MRI和PET两个模态的数据输入到包含Mamba-

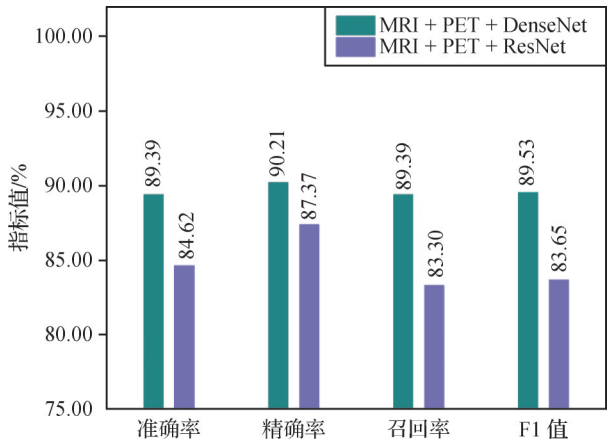


图5 ResNet和DenseNet结果对比
Fig. 5 Comparison results of ResNet and DenseNet

block的网络与不包含Mambablock的网络。图6展示了Mambablock消融实验结果。结果表明,不包含Mambablock的网络在准确率上明显低于包含Mambablock的网络。具体而言,不包含Mambablock的网络准确率为89.39%,而包含Mambablock的网络准确率提高至92.42%。这一结果表明,Mambablock与传统的卷积结构的整合能够显著提升网络的分类性能。

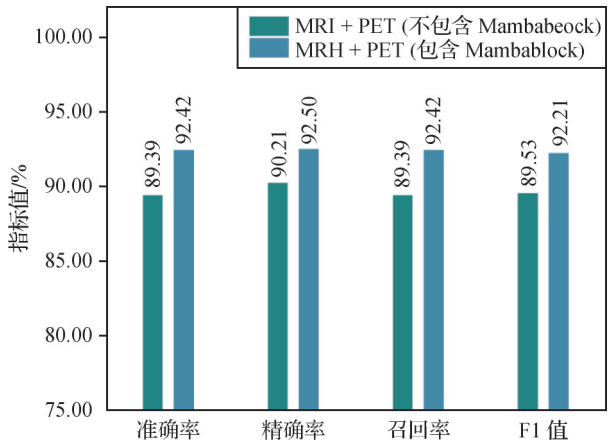


图6 Mambablock消融实验结果
Fig. 6 Mambablock ablation experiment results

本研究对比了不同输入数据和模型架构在阿尔茨海默症分类任务中的表现。图7展示了4组实验的混淆矩阵。结果表明,单独使用MRI作为输入时,DenseMamba在识别AD方面具有较强的能力,但容易将CN样本误分类为MCI;单独使用PET作为输入时,DenseMamba在区分AD与MCI时略显不足。当同时输入MRI与PET时,模型性能显著提升。此外,不包含Mambablock的多模态模型在分类性能上也

优于单模态模型,这一结果进一步验证了多模态数据联合使用的有效性。同时,包含 Mambablock 的多

模态模型分类性能优于不包含 Mambablock 的模型,进一步验证了 Mambablock 的有效性。

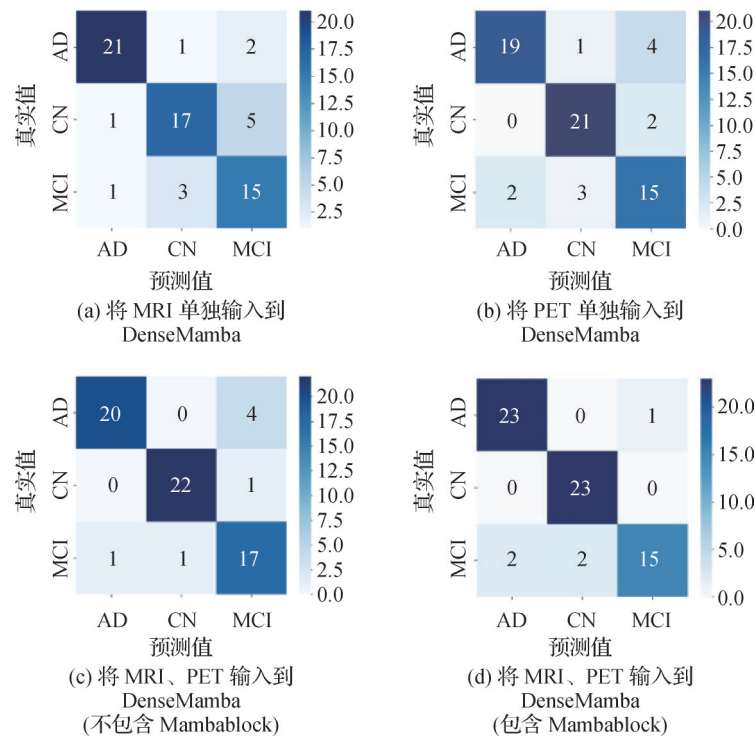


图7 消融实验的混淆矩阵

Fig. 7 Confusion matrices of ablation experiment ((a) the MRI input to DenseMamba; (b) the PET input to DenseMamba; (c) the MRI-PET input to DenseMamba without Mambablock; (d) the MRI-PET input to DenseMamba with Mambablock)

4 结 论

针对阿尔茨海默症(AD)患者数量逐年增加,以及医生在临床医学图像分析中面临的时间压力日益加大的现状,本文提出一种基于3D卷积技术、密集连接神经网络(DenseNet)和Mamba的综合诊断模型,旨在为医生提供高效的辅助诊疗工具,以提高诊疗效率并减轻医生的工作负担。

该模型充分利用MRI和PET数据,实现对AD、CN和MCI 3类患者的分类任务。3D卷积技术能够有效提取空间信息并捕捉图像中的深层特征;密集连接网络强化了特征的传递和复用;Mamba善于捕捉长距离依赖特征方面,进一步增强了分类的准确性。

通过实验数据的分析与对比,结果表明所提出的模型在分类任务中表现优异,取得了较好的分类效果。这表明该方法能够提升分类准确率,同时有

效降低了人工识别的时间成本,具有较强的临床应用潜力。

在未来的研究中,计划进一步优化该模型,探索更多的改进方向。例如,结合更丰富的医学影像数据、引入自监督学习等,以及提升模型的计算效率,使其更适应实际临床应用中的需求。通过持续优化与实践验证,以期能够为阿尔茨海默症的早期诊断和精准医疗提供更加可靠和高效的辅助工具。

参考文献(References)

- Aderghal K, Benois-Pineau J and Afdel K. 2017. Classification of sMRI for Alzheimer's disease diagnosis with CNN: single Siamese networks with 2D+? Approach and fusion on ADNI//Proceedings of 2017 ACM on International Conference on Multimedia Retrieval. Bucharest, Romania: Association for Computing Machinery: 494-498 [DOI: 10.1145/3078971.3079010]
- Billones C D, Demetria O J L D, Hostallero D E D and Naval P C.

2016. DemNet: a convolutional neural network for the detection of Alzheimer's disease and mild cognitive impairment//Proceedings of 2016 IEEE Region 10 Conference (TENCON). Singapore, Singapore: IEEE: 3724-3727 [DOI: 10.1109/TENCON.2016.7848755]
- Chen H, Guo H R, Xing L Q, Chen D, Yuan T, Zhang Y P and Zhang X D. 2023. Multimodal predictive classification of Alzheimer's disease based on attention-combined fusion network: integrated neuro-imaging modalities and medical examination data. *IET Image Processing*, 17(11): 3153-3164 [DOI: 10.1049/ipr2.12841]
- Cui Z Y, Gao Z A, Leng J X, Zhang T L, Quan P and Zhao W. 2019. Alzheimer's disease diagnosis using enhanced inception network based on brain magnetic resonance image//Proceedings of 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). San Diego, USA: IEEE: 2324-2330 [DOI: 10.1109/BIBM47256.2019.8983046]
- Deng S, He X H, Qing L B, Chen H G and Teng Q Z. 2022. Weakly supervised fine-grained classification method of Alzheimer's disease based on improved visual geometry group network. *Journal of Computer Applications*, 42(1): 302-309 (邓爽, 何小海, 卿麟波, 陈洪刚, 滕奇志. 2022. 基于改进 VGG 网络的弱监督细粒度阿尔兹海默症分类方法. *计算机应用*, 42(1): 302-309) [DOI: 10.11772/j.issn.1001-9081.2021020258]
- Hong X, Huang K F and Yang C H. 2023. Temporal series features extraction based on Bi-ConvLSTM of Alzheimer's disease prediction CTISS model. *Journal of Image and Graphics*, 28(4): 1146-1156 (洪欣, 黄铠洋, 杨晨晖. 2023. 基于 Bi-ConvLSTM 时序特征提取的阿尔兹海默症预测 CTISS 模型. *中国图象图形学报*, 28(4): 1146-1156) [DOI: 10.11834/jig.211186]
- Hu Z Y, Wang Y H and Xiao L. 2025. Alzheimer's disease diagnosis by 3D-SEConvNeXt. *Journal of Big Data*, 12(1): #15 [DOI: 10.1186/s40537-025-01088-8]
- Huang G, Liu Z, Van Der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Jin D, Zhou B, Han Y, Ren J J, Han T, Liu B, Lu J, Song C Y, Wang P, Wang D W, Xu J, Yang Z Y, Yao H X, Yu C S, Zhao K, Wintermark M, Zuo N M, Zhang X Q, Zhou Y Y, Zhang X, Jiang T Z, Wang Q and Liu Y. 2020. Generalizable, reproducible, and neuroscientifically interpretable imaging biomarkers for Alzheimer's disease. *Advanced Science*, 7(14): #2000675 [DOI: 10.1002/adv.202000675]
- Kong Z K, Zhang M Y, Zhu W J, Yi Y, Wang T and Zhang B C. 2022. Multi-modal data Alzheimer's disease detection based on 3D convolution. *Biomedical Signal Processing and Control*, 75: #103565 [DOI: 10.1016/j.bspc.2022.103565]
- Kruthika K R, Rajeswari, Maheshappa H D and Alzheimer's Disease Neuroimaging Initiative. 2019. CBIR system using capsule networks and 3D CNN for Alzheimer's disease diagnosis. *Informatics in Medicine Unlocked*, 14: 59-68 [DOI: 10.1016/j.imu.2018.12.001]
- Li Y M, Tang J H, Ding W W, Luo J, Ahmad N and Kumar R. 2024. An assisted diagnosis of Alzheimer's disease incorporating attention mechanisms Med-3D transfer modeling. *Computers, Materials and Continua*, 78(1): 713-733 [DOI: 10.32604/cmc.2023.046872]
- Lian C F, Liu M X, Pan Y S and Shen D G. 2022. Attention-guided hybrid network for dementia diagnosis with structural MR images. *IEEE Transactions on Cybernetics*, 52(4): 1992-2003 [DOI: 10.1109/TCYB.2020.3005859]
- Liu M X, Zhang J, Adeli E and Shen D G. 2019. Joint classification and regression via deep multi-task multi-channel learning for Alzheimer's disease diagnosis. *IEEE Transactions on Biomedical Engineering*, 66(5): 1195-1206 [DOI: 10.1109/TBME.2018.2869989]
- Liu X, Li W M, Miao S, Liu F Y, Han K and Bezabih T T. 2024. HAMMF: hierarchical attention-based multi-task and multi-modal fusion model for computer-aided diagnosis of Alzheimer's disease. *Computers in Biology and Medicine*, 176: #108564 [DOI: 10.1016/j.combiomed.2024.108564]
- Mulyadi A W, Jung W, Oh K, Yoon J S, Lee K H and Suk H I. 2023. Estimating explainable Alzheimer's disease likelihood map via clinically-guided prototype learning. *NeuroImage*, 273: #120073 [DOI: 10.1016/j.neuroimage.2023.120073]
- Pan Q, Ding K and Chen D H. 2021. Multi-classification prediction of Alzheimer's disease based on fusing multi-modal features//Proceedings of 2021 IEEE International Conference on Data Mining (ICDM). Auckland, New Zealand: IEEE: 1270-1275 [DOI: 10.1109/ICDM51629.2021.00156]
- Puente-Castro A, Fernandez-Blanco E, Pazos A and Munteanu C R. 2020. Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques. *Computers in Biology and Medicine*, 120: #103764 [DOI: 10.1016/j.combiomed.2020.103764]
- Saratxaga C L, Moya I, Picón A, Acosta M, Moreno-Fernandez-de-Leceta A, Garrote E and Bereciartúa-Pérez A. 2021. MRI deep learning-based solution for Alzheimer's disease prediction. *Journal of Personalized Medicine*, 11(9): #902 [DOI: 10.3390/jpm11090902]
- Shi J, Wang L L, Wang S S, Chen Y X, Wang Q, Wei D M, Liang S J, Peng J L, Yi J J, Liu S F, Ni D, Wang M L, Zhang D Q and Shen D G. 2020. Applications of deep learning in medical imaging: a survey. *Journal of Image and Graphics*, 25(10): 1953-1981 (施俊, 汪琳琳, 王珊珊, 陈艳霞, 王乾, 魏冬铭, 梁淑君, 彭佳林, 易佳锦, 刘盛锋, 倪东, 王明亮, 张道强, 沈定刚. 2020. 深度学习在医学影像中的应用综述. *中国图象图形学报*, 25(10): 1953-1981) [DOI: 10.11834/jig.200255]
- Tang Y, Xiong X, Tong G, Yang Y and Zhang H. 2024. Multimodal

- diagnosis model of Alzheimer's disease based on improved transformer. *BioMedical Engineering OnLine*, 23; #8 [DOI: 10.1186/s12938-024-01204-4]
- Venugopalan J, Tong Li, Hassanzadeh H R and Wang M D. 2021. Multimodal deep learning models for early detection of Alzheimer's disease stage. *Scientific Reports*, 11 (1) : #3254 [DOI: 10.1038/s41598-020-74399-w]
- Wang B, Wu X H, Gu R, Qing L B and He X H. 2023. Alzheimer's disease classification network based on improved residual network. *Intelligent Computer and Applications*, 13(3): 69-76, 82 (王斌, 吴晓红, 辜蕊, 卿琳波, 何小海. 2023. 基于改进ResNet的阿尔兹海默症分类网络. *智能计算机与应用*, 13(3): 69-76, 82)
- Wang H F, Shen Y Y, Wang S Q, Xiao T F, Deng L M, Wang X Y and Zhao X Y. 2019. Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease. *Neurocomputing*, 333: 145-156 [DOI: 10.1016/j.neucom.2018.12.018]
- Wu R K, Liu Y H, Liang P C and Chang Q. 2025. UltraLight VM-UNet: parallel vision Mamba significantly reduces parameters for skin lesion segmentation. *Patterns*, #101298 [DOI: 10.1016/j.patter.2025.101298]
- Xia Z M, Yue G H, Xu Y W, Feng C Y, Yang M Y, Wang T F and Lei B Y. 2020. A novel end-to-end hybrid network for Alzheimer's disease detection using 3D CNN and 3D CLSTM//*Proceedings of the 17th IEEE International Symposium on Biomedical Imaging (ISBI)*. Iowa City, USA: IEEE: 1-4 [DOI: 10.1109/ISBI45749.2020.9098621]
- Zeng A, Jia L F, Pan D and Song X W. 2019. Early prognosis of Alzheimer's disease based on convolutional neural networks and ensemble learning. *Journal of Biomedical Engineering*, 36 (5) : 711-719 (曾安, 贾龙飞, 潘丹, Song X W. 2019. 基于卷积神经网络和集成学习的阿尔茨海默症早期诊断. *生物医学工程学杂志*, 36(5): 711-719) [DOI: 10.7507/1001-5515.201809040]

作者简介

胡馨悦,女,硕士研究生,主要研究方向为医学图像识别。

E-mail:xinyuehu@mail.hfut.edu.cn

刘羽,男,通信作者,教授,主要研究方向为多源图像融合和医学图像分析。E-mail:yuliu@hfut.edu.cn

石雨,男,博士研究生,主要研究方向为医学图像融合。

E-mail:yushi@mail.hfut.edu.cn