

一种基于改进 Swin Transformer 模型的阿尔茨海默病 MRI 图像分类方法研究*

赵姗姗 史浩霖 王颖帅 万艳丽

(中国医学科学院/北京协和医学院医学信息研究所 北京 100020)

[摘要] 目的/意义 开发更有效的磁共振成像分类模型,提升阿尔茨海默病早期诊断准确率,提高患者生活质量。方法/过程 基于 Swin Transformer 模型,结合多分辨率特征融合(multi-resolution feature fusion, MRFF)模块,提出一种改进的网络结构 Swin Transformer + MRFF。采用数据增强和扩充策略,解决样本不足和类别不平衡问题。通过与其他深度学习模型的对比实验,评估改进模型在阿尔茨海默病识别中的优势。结果/结论 改进后的 Swin Transformer + MRFF 模型在 OASIS-1 数据集上表现显著优于基础模型,显著提高了轻度痴呆和无痴呆类别的识别能力。

[关键词] Swin Transformer 模型;多分辨率特征融合;磁共振成像医学图像分类;阿尔茨海默病

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2025.05.014

Study on Alzheimer's Disease MRI Image Classification Method Based on an Improved Swin Transformer Model

ZHAO Shanshan, SHI Haolin, WANG Yingshuai, WAN Yanli

Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China

[Abstract] **Purpose/Significance** To develop a more effective magnetic resonance imaging (MRI) classification model, so as to improve the early diagnosis accuracy of Alzheimer's disease (AD) and enhance the quality of life of patients. **Method/Process** Based on the Swin Transformer model, integrated with a multi-resolution feature fusion (MRFF) module, an improved network structure, Swin Transformer + MRFF, is proposed. Data augmentation and expansion strategies are adopted, and the problems of insufficient samples and category imbalance are solved. Through comparative experiments with other deep learning models, the advantages of the improved model in the recognition of AD are evaluated. **Result/Conclusion** The improved Swin Transformer + MRFF model significantly outperforms the base model on the OASIS-1 dataset, notably enhancing the recognition ability for mild dementia and non-dementia categories.

[Keywords] Swin Transformer model; multi-resolution feature fusion (MRFF); MRI medical image classification; Alzheimer's disease

1 引言

阿尔茨海默病 (Alzheimer's disease, AD) 是最常见的老年期痴呆类型,占老年痴呆总数的 60% ~

80%^[1],严重影响患者生活质量并加重经济负担。轻度认知功能障碍 (mild cognitive impairment, MCI) 是 AD 的早期阶段,也是疾病干预的关键期^[2]。磁共振成像能够有效检测病变,是 AD 诊断的重要技术手段。医学图像分类通过提取图像特征来预测标签类

[修回日期] 2025-05-09

[作者简介] 赵姗姗,硕士研究生,发表论文 2 篇;通信作者:万艳丽。

[基金项目] 国家自然科学基金资助项目 (项目编号:61971446);中国医学科学院创新工程项目 (项目编号:2022-I2M-1-019);中国医学科学院非营利性中央研究院基金 (项目编号:2024-ZHCH630-01)。

别，是计算机辅助诊断的重要内容。深度学习方法中卷积神经网络 (convolutional neural network, CNN) 广泛应用于医学图像分类，但面临准确性不足，泛化能力有限，建模长距离依赖关系方面存在局限性^[3]等问题。相反，Transformer^[4]具有优越的全局特征建模能力，在目标检测^[5]、语义分割^[6]、目标跟踪^[7]、图像生成^[8]和图像增强^[9]等任务中表现出色，显示出其在医学影像分析中的巨大潜力。

本研究基于 Swin Transformer 模型，结合多分辨率特征融合 (multi-resolution feature fusion, MRFF) 模块，提出一种改进的网络结构 Swin Transformer + MRFF，以提高阿尔茨海默病磁共振成像 (magnetic resonance imaging, MRI) 识别分类的准确性。探讨网络结构改进与数据增强对 Swin Transformer 模型性能的影响，并通过与其他深度学习模型的对比实验，评估改进模型在 AD 识别中的优势。

2 Swin Transformer 模型与方法

2.1 Swin Transformer 模型

2.1.1 概述 Transformer 是一种自然语言处理经典模型，于 2017 年被提出^[4]。2020 年 Dosovitskiy A 等^[10]提出的 Vision Transformer 模型首次将 Transformer 模型应用于图像分类任务。Swin Transformer 模型^[11]是 2021 年提出的计算机视觉领域 Transformer 模型，结合卷积局部化思想，并基于自注意力机制构建网络，可捕捉图像中的长距离依赖关系。与 Vision Transformer 模型不同，Swin Transformer 模型采用移动窗口自注意力机制，仅与相邻的局部 token 进行相关性计算，有效减少参数数量，显著降低模型的计算复杂度，并提高了模型效率。此外，Swin Transformer 模型引入层间连接机制，促进信息传递和特征提取，进一步提升了模型的整体性能。

2.1.2 模型结构 Swin Transformer 模型标准网络结构，见图 1。先通过图像分割模块 Patch Partition 将输入的 RGB 图像分割成多个不重叠的 patch，每个 patch 的特征为其原始像素的 RGB 值串联表示。然后，patch 特征通过线性嵌入层 (linear embedding) 映射到目标特征维度，进入 Swin Transformer

Block 进行处理，完成阶段 1 的操作。在阶段 2 至阶段 4 中，图像数据通过 Patch Merging 层进行下采样，合并相邻的 patch 以减少 token 数量，同时将图像分辨率减半并将特征通道数增加一倍。重复该过程，逐步捕获多尺度的特征信息。

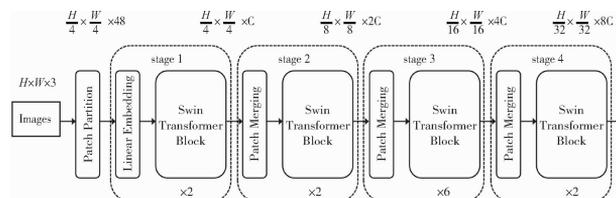


图 1 Swin Transformer 模型标准网络结构

Swin Transformer Block 是该网络的核心模块，包括两个步骤，见图 2。

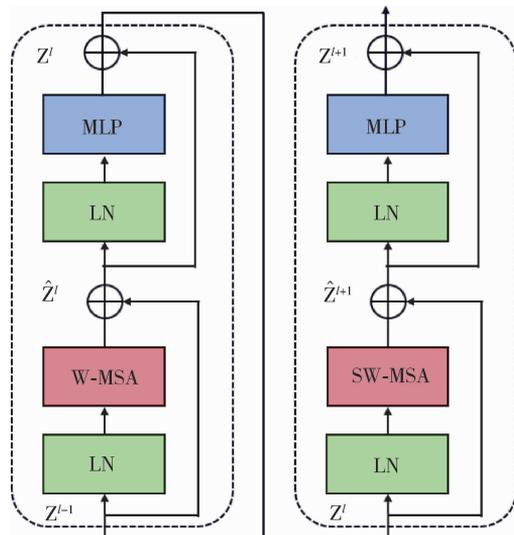


图 2 Swin Transformer Block 结构

步骤 1 使用基于窗口的多头自注意力机制 (window-based multi-head self-attention, W-MSA)，通过设置固定窗口大小限制计算范围，只计算窗口内部的注意力关系，减少计算量并保留局部依赖性。步骤 2 引入平移窗口的多头自注意力机制 (shifted window-based multi-head self-attention, SW-MSA)，使相邻窗口间能够交换信息、捕捉全局特征，增强了模型对空间信息的感知能力。通过局部自注意力和窗口偏移的结合，Swin Transformer 模型有效减少了计算量，同时保留了对局部和全局特征的强大表达能力。层归一化 (layer normalization, LN) 被用来规

范每一层的输出，稳定训练过程并加速收敛，有助于减轻梯度消失和梯度爆炸问题，并提高模型的稳定性。多层感知器（multilayer perceptron, MLP）结构是一个由两层全连接层组成的模块，用于进一步增强模型的非线性表达能力。

2.2 模型搭建

采用 Swin - Transformer 模型的一个轻量级版本 Swin - Transformer - Tiny 作为基础模型，在 Patch Partition 结构前加入 MRFF 模块，通过提取和融合不同分辨率的特征，增强模型对多尺度信息的感知能力^[12]。MRFF 模块包含 5 个基本结构。一是多尺度缩放因子。使用不同的缩放因子（如 1.0、0.5、0.25）对输入的不同分辨率特征图进行缩放，生成相应分辨率的特征图。低分辨率特征图有助于捕捉全局信息，高分辨率特征图有助于捕捉细节^[13]。假设原始图像尺寸为 $W \times H$ ，缩放因子为 s ，则缩放后的图像尺寸为：

$$W' = S \times W, H' = S \times H \quad (1)$$

二是卷积处理。包含多个卷积块，每个卷积块含一个卷积层、一个批归一化层和一个 ReLU 激活函数。每个卷积块用于对不同分辨率下的特征图分别进行卷积操作，以提取每个尺度上的特征^[14]。其标准卷积操作的数学公式表示如下。其中 W_{ij} 是卷积核的权重， $x_{i,j}$ 是输入特征图的位置值， b 是偏置。

$$f(x) = \text{Conv2d}(x) \sum_i^j W_{ij} \times x_{i,j} + b \quad (2)$$

三是特征对齐。通过自适应池化将不同分辨率的特征图调整为与原始输入图像相同的高宽尺寸，以便有效融合。假设 f 是经过卷积处理的特征图，池化后的特征图 f' 可如下表示。其中， n 是池化区域的大小， f_i 是池化区域内的像素值。

$$f' = \frac{1}{n} \sum_{i=1}^n f_i \quad (3)$$

四是特征拼接。将不同尺度的特征图沿通道维度拼接，将各尺度提取到的特征信息合并到一个特征图中，假设有两个特征图 f_1 和 f_2 ，拼接后的结果 out 的形状如下。其中 C_1 和 C_2 是两个特征图的通道数。

$$f_1 \in R^{C_1 \times H \times W}, f_2 \in R^{C_2 \times H \times W},$$

$$\text{out} = \text{concat}(f_1, f_2), \text{out} \in R^{(C_1+C_2) \times H \times W} \quad (4)$$

五是通道数压缩。通过 1×1 卷积将拼接后的特

征图通道数压缩回输入特征图的通道数，确保最终输出的特征图与输入的通道数一致，表示方式如下。其中， W_{ij} 是 1×1 卷积核的权重， $x_{i,j}$ 是输入的特征图。这样可以避免特征图维度膨胀，提高计算效率，并确保后续模块可以接收正确维度的输入^[15]。

$$f_{\text{out}}(x) = \sum_i^j W_{ij} \times x_{i,j} + b \quad (5)$$

MRFF 模块在多特征融合方法中具有显著优势，其模块结构，见图 3。

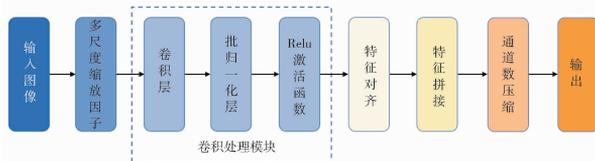


图 3 MRFF 模块结构

首先，通过不同的缩放因子处理图像，捕获多尺度信息，既能提取全局特征，又能关注局部细节。与传统方法如特征金字塔网络相比，MRFF 模块通过在多个尺度上处理图像，而不是仅依赖于特定尺度，能更全面地捕捉不同层次的特征信息。其次，MRFF 使用自适应池化将每个分辨率的特征调整为统一尺寸，确保在处理不同分辨率时不会丢失关键信息，增强了模型的灵活性和适应性。相较于一些固定尺寸的多尺度方法，如金字塔池化，MRFF 提供了更强的灵活性和动态调整能力。最后，特征拼接操作能够有效融合不同尺度的特征，并通过 1×1 卷积将通道数降回原始维度，避免了计算复杂度过度增加、多卷积层或额外计算步骤带来的计算负担。卷积层与激活层的组合进一步提升了每个尺度的特征提取能力，使多尺度特征更加精细，具有更强的非线性表达能力。本研究将 MRFF 模块置于 Swin Transformer 网络的前端，融合多分辨率特征信息，确保从输入图像开始便能获取丰富的多尺度特征，为后续 Swin Transformer 模型提供更全面的输入特征，见图 4。

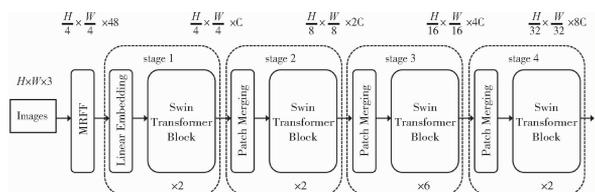


图 4 添加 MRFF 后的 Swin Transformer 模型网络结构

3 实验结果与分析

3.1 实验环境和参数设置

本研究基于 Pytorch 架构, 使用 Python 3.8 和 PyCharm 进行开发, 运行在 Linux 系统下, 服务器配置包括 32 核 CPU、512GB 内存、6TB 存储和 4 块 32GB 显存的 NVIDIA Tesla V100 GPU。网络输入大小为 224×224 , 训练时初始学习率为 $8e-3$, 优化器采用 SGD, momentum 为 0.95, 学习率使用 cos 衰减, 权重衰减为 $1e-3$, 损失函数为交叉熵, Batch size 为 32, 迭代次数为 350。

3.2 实验数据

3.2.1 实验数据集 使用 OASIS 数据库中的 OASIS-1 数据集。该数据集由 416 名 18—96 岁受试者的横截面数据组成^[16]。OASIS-1 数据集中, MRI 图像被划分为以下 4 类: 非痴呆 (NonDemented), 非常轻度痴呆 (VeryMildDemented), 轻度痴呆 (MildDemented), 中度痴呆 (ModerateDemented)。将非常轻度痴呆分类数据并入轻度痴呆类别。按照 8:1:1 的比例分配训练集、验证集和测试集, 见表 1。其中训练集中中度痴呆类别图像扩充为原本的 50 倍。3 种 AD 的 MRI 图像, 见图 5。

表 1 数据集分配详情 (张)

数据集	非痴呆	轻度痴呆	中度痴呆
训练集	2 560	2 509	52 × 50
验证集	320	314	6
测试集	320	314	6

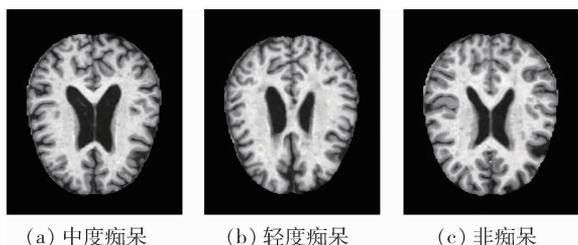


图 5 3 种阿尔茨海默病的 MRI 图像

3.2.2 数据预处理 在预处理阶段对数据进行增强^[17]和数据量扩增^[18]。数据增强是深度学习图像分类任务中的必要且关键组成部分^[19]。采用 Cu-

buk E D 等^[20]提出的 AutoAugment 自动搜索改进数据增强策略, 通过随机选择变换子策略及其应用概率和强度, 提升数据集的多样性, 每个操作有独立的应用概率 (p_1, p_2) 和变换强度 ($magnitude_1, magnitude_2$), 随机性使每次训练中的图像变换具有高度多样性。具体增强操作的详细情况, 见表 2。

表 2 图像增强操作参数与说明

增强操作	增强范围	具体描述
剪切	$[-0.3, 0.3]$	沿 x 轴或 y 轴对图像进行剪切
平移	$[0, 150/331]$	沿 x 轴或 y 轴平移图像, 最大平移距离约为图像尺寸的 15%
旋转	$[-30, 30]$	改变图像的角度, 最大旋转角度为 30 度
颜色增强	$[-1, 1]$	调整图像的颜色饱和度, 增强颜色效果, 值越大, 颜色越鲜艳
对比度增强	$[-1, 1]$	增强图像的对比度, 值越大, 对比度越强
锐度增强	$[-1, 1]$	提高图像的锐度, 增强细节, 值越大, 图像越锐利
亮度增强	$[-1, 1]$	调整图像的亮度, 值越大, 图像越亮
太阳化	$[0, 255]$	反转图像的颜色, 增加负片效果, 值越大, 太阳化效果越强
自动对比度		自动调整图像的对比度, 去除最亮和最暗部分的极端值
直方图均衡化		对图像进行直方图均衡化, 改善图像的亮度分布
图像反转		将图像的颜色反转, 生成负片效果

此外, 对数据量进行扩增, 对于 OASIS-1 数据集中占比较小的中度痴呆类别 MRI 图像进行过采样增强^[21-23], 将数据量扩增为原本的 50 倍, 以此来平衡数据, 使训练集中非痴呆: 轻度痴呆: 中度痴呆 $\approx 1:1:1$ 。部分数据预处理结果, 见图 6。

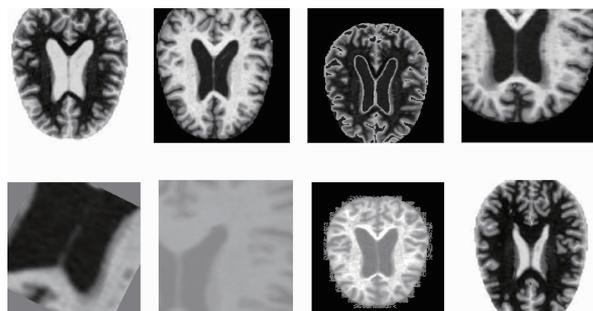


图 6 数据增强效果 (部分)

3.3 评价标准

采用经典分类评价指标对该模型进行评价，包括准确率 (accuracy)、精度 (precision)、召回率 (recall)、F1 指数^[24]，计算方式如下。其中真阳性 (true positive, TP) 代表实际为阳性且被模型正确分类的样本数，真阴性 (true negative, TN) 代表实际为阴性且被模型正确分类的样本数，假阳性 (false positive, FP) 代表实际为阴性但被模型错误分类的样本数，假阴性 (false negative, FN) 代表实际为阳性但被模型错误分类的样本数；阳性指患者，阴性指正常人。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (6)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

3.4 实验结果分析

3.4.1 不同模型分类结果对比 比较多种模型在相同数据集和平台下的性能表现，见表 3。在 AD 的 MRI 图像分类任务中，Swin Transformer + MRFF 模型表现最佳，准确率为 87.26%，明显优于 VGG 16 bn (81.47%)、mobilenetv2 (79.20%) 和

vit_b_16 (73.89%)。Swin Transformer + MRFF 模型的召回率为 91.43%，显示出更好的灵敏性；精度为 91.45%，远超 vit_b_16 (75.39%) 和 mobilenetv2 (76.35%)，表明其对负类样本的判别能力更强；F1 指数为 91.42%，表现最优。与未加入 MRFF 的 Swin Transformer 模型相比，加入 MRFF 后准确率从 84.83% 提升至 87.26%，F1 指数从 87.26% 提升至 91.42%，显著提升了模型性能。

表 3 不同模型分类结果对比 (%)

模型	准确率	召回率	精度	F1
VGG16bn	81.47	79.29	80.05	79.23
mobilenetv2	79.20	77.87	76.35	77.10
vit_b_16	73.89	68.84	75.39	71.48
Swin Transformer	84.83	89.82	85.27	87.26
Swin Transformer + MRFF	87.26	91.43	91.45	91.42

绘制基线 Swin Transformer 模型和 Swin Transformer + MRFF 模型的损失曲线，见图 7。加入 MRFF 后，训练损失和验证损失的下降更为平稳，且最终验证集的损失值更低。特别是在训练初期，加入 MRFF 的模型显著减少了损失的波动，验证了 MRFF 机制在提升模型稳定性和泛化性能方面的有效性。整体上，Swin Transformer + MRFF 相比原模型在收敛速度和最终性能方面均有提升。

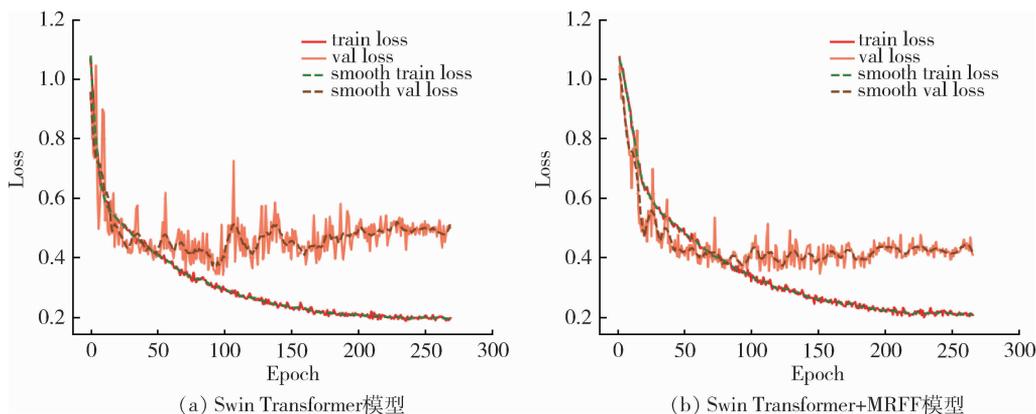


图 7 两种模型的损失曲线

Swin Transformer 模型和 Swin Transformer + MRFF 模型在 AD 的 MRI 图像分类任务中的分类结果，

见表 4。相比基线模型，Swin Transformer + MRFF 模型的精确率、召回率和 F1 指数均显著提升。Swin

Transformer + MRFF 模型具有更强的泛化能力，能更准确地捕捉不同类别的痴呆症患者特征。

表 4 两种模型各类别 MRI 图像分类性能对比 (%)

模型	评价指标	非痴呆	轻度痴呆	中度痴呆
Swin Transformer	精确率	88.02	82.23	86.31
	召回率	81.12	88.97	100
	F1	84.14	85.26	92.28
Swin Transformer + MRFF	精确率	88.98	86.45	100
	召回率	84.89	88.88	100
	F1	86.78	86.85	100

3.4.2 数据增强效果 数据增强和扩充显著提升了模型性能，见表 5。Swin Transformer 模型的准确率从 70.52% 提升至 84.83%，有效避免过拟合，并增强了泛化能力；召回率从 47.39% 提升至 89.82%，大幅提高了对少数类样本的识别能力。数据增强还提高了 mobilenetv2 的 F1 指数。VGG16bn 和 Swin Transformer 的 F1 指数也分别从 50.23% 和 47.00% 提升至 79.23% 和 87.26%，证明了该模型在复杂数据下的鲁棒性。

表 5 数据增强效果指标

模型	数据增强	准确率 (%)	召回率 (%)	精度 (%)	F1 (%)
VGG16bn	×	75.53	50.94	52.61	50.23
VGG16bn	√	81.47	79.29	80.05	79.23
mobilenetv2	×	80.92	79.01	74.30	76.26
mobilenetv2	√	82.72	80.21	88.45	83.70
vit_b_16	×	70.60	47.54	47.18	47.28
vit_b_16	√	73.89	68.84	75.39	71.48
Swin Transformer	×	70.52	47.39	47.42	47.00
Swin Transformer	√	84.83	89.82	85.27	87.26
Swin Transformer + MRFF	×	73.45	49.35	49.18	49.26
Swin Transformer + MRFF	√	87.26	91.43	91.45	91.42

4 讨论

本研究通过改进 Swin Transformer 模型并结合 MRFF 模块，显著提升了 AD 的 MRI 图像分类准确性和鲁棒性。对比实验结果显示，改进后的 Swin Transformer + MRFF 模型在 OASIS - 1 数据集上取得了 87.26% 的准确率和 91.43% 的召回率，F1 指数从

87.26% 提高至 91.42%，表明 MRFF 模块有效增强了模型对多尺度特征的捕捉能力，特别是轻度痴呆和非痴呆类别的识别效果得到了显著改善。此外，数据增强和扩充策略有效解决了样本不足和类别不平衡问题，改善了过拟合现象^[25]。证明基于 Swin Transformer 的深度学习方法在 AD 的 MRI 图像分类中具有优势，为早期诊断提供了强有力的支持。

Swin Transformer 模型展现了强大性能，未来混合模型架构^[26-27]、自监督学习^[28]和少样本学习^[29-30]、融合多模态信息^[31-32]等方面的研究将推动该模型在更广泛的实际应用中发挥更大的作用。一是探索将 Swin Transformer 模型与其他轻量级卷积模型或轻量级 Transformer 架构结合，形成更高效的混合结构。二是探索如何减少对大规模标签数据的依赖，利用自监督学习和少样本学习是两个值得探讨的方向。三是探索 Swin Transformer 模型在跨模态任务中的应用和改进。

在数据处理方面，本文使用的数据增强方法容易受到对抗性攻击，未带来新的图像视觉特征。未来可考虑在图像中添加随机白噪声或随机擦除图像的一部分^[33]、生成对抗网络^[34]等。本研究集中在单一数据集上进行测试，未来研究将扩展至多个医学图像数据集，进一步验证本研究模型在不同数据源下的表现，从而提升其泛化能力和鲁棒性。

5 结语

本研究提出了一种改进的 Swin Transformer 模型，结合多分辨率特征融合模块，显著提升了阿尔茨海默病 MRI 图像分类准确性和鲁棒性。对比实验结果表明，改进后的模型在 OASIS - 1 数据集上表现优异，准确率、召回率和 F1 指数均显著提高。该方法已显示出较强的分类性能，为阿尔茨海默病的早期诊断提供了有力支持，具有广阔的应用前景。

作者贡献：赵姗姗负责实验实施、论文撰写；史浩霖负责文献调研；王颖帅负责技术指导；万艳丽负责提供指导。

利益声明：所有作者均声明不存在利益冲突。

参考文献

- 1 FILIOU R, BIER N, SLEGGERS A, et al. Connected speech assessment in the early detection of Alzheimer's disease and mild cognitive impairment: a scoping review [J]. *Aphasiology*, 2020, 34 (6): 723 - 755.
- 2 ROSENBERG P B, LYKETSOS C. Mild cognitive impairment; searching for the prodrome of Alzheimer's disease [J]. *World psychiatry*, 2008, 7 (2): 72.
- 3 CHEN L, LI S, BAI Q, et al. Review of image classification algorithms based on convolutional neural networks [J]. *Remote sensing*, 2021, 13 (22): 4712.
- 4 VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [EB/OL]. [2024 - 05 - 11]. <https://arxiv.org/abs/1706.03762>.
- 5 CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [EB/OL]. [2024 - 05 - 11]. <https://arxiv.org/abs/2005.12872>.
- 6 WANG H, ZHU Y, ADAM H, et al. MaX - DeepLab: end-to-end panoptic segmentation with mask transformers [C]. Online; The IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
- 7 MA F, SHOU M Z, ZHU L, et al. Unified transformer tracker for object tracking [C]. New Orleans: The IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.
- 8 JIANG Y, CHANG S, WANG Z. Transgan: two pure transformers can make one strong gan, and that can scale up [EB/OL]. [2025 - 01 - 11]. https://proceedings.neurips.cc/paper_files/paper/2021/hash/7c220a2091c26a7f5e9f1cfb099511e3-Abstract.html.
- 9 CHEN H, WANG Y, GUO T, et al. Pre-trained image processing transformer [C]. Online; The IEEE/CVF conference on computer vision and pattern recognition, 2021.
- 10 DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words; transformers for image recognition at scale [EB/OL]. [2024 - 05 - 11]. <https://arxiv.org/pdf/2010.11929/1000>.
- 11 LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [C]. Online; The IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
- 12 DUARTE D, NEX F, KERLE N, et al. Multi-resolution feature fusion for image classification of building damages with convolutional neural networks [J]. *Remote sensing*, 2018, 10 (10): 1636.
- 13 RAMIREZ J M, ARGUELLO H. Multiresolution compressive feature fusion for spectral image classification [J]. *IEEE transactions on geoscience and remote sensing*, 2019, 57 (12): 9900 - 9911.
- 14 CHEN J, WANG W, XING F, et al. Multi - feature fusion attention network for single image super - resolution [J]. *IET image processing*, 2023, 17 (5): 1389 - 1402.
- 15 HUO X, SUN G, TIAN S, et al. HiFuse: hierarchical multi-scale feature fusion network for medical image classification [EB/OL]. [2024 - 05 - 11]. <https://www.sciencedirect.com/science/article/pii/S1746809423009679>.
- 16 MARCUS D S, WANG T H, PARKER J, et al. Open access series of imaging studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults [J]. *Journal of cognitive neuroscience*, 2007, 19 (9): 1498 - 1507.
- 17 SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning [J]. *Journal of big data*, 2019, 6 (1): 1 - 48.
- 18 ZHANG X, WANG Y, ZHANG N, et al. Spectral - spatial fractal residual convolutional neural network with data balance augmentation for hyperspectral classification [J]. *IEEE transactions on geoscience and remote sensing*, 2021, 59 (12): 10473 - 10487.
- 19 ZAMIR S W, ARORA A, KHAN S, et al. Learning enriched features for real image restoration and enhancement; computer Vision - ECCV 2020 [C]. Glasgow: European Conference on Computer Vision, 2020.
- 20 CUBUK E D, ZOPH B, MANE D, et al. Autoaugment: learning augmentation strategies from data [C]. Long Beach: The IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
- 21 AHAMED B S, ARYA M S, NANCY A O V. Diabetes mellitus disease prediction using machine learning classifiers with oversampling and feature augmentation [J]. *Advances in human - computer interaction*, 2022 (1): 9220560.
- 22 ZUBAIR M, WOO S, LIM S, et al. Deep representation learning with sample generation and augmented attention module for imbalanced ECG classification [J]. *IEEE journal of biomedical and health informatics*, 2023, 28 (5): 2461 - 2472.
- 23 JIANG X, GE Z. Data augmentation classifier for imbalanced fault classification [J]. *IEEE transactions on automation science and engineering*, 2020, 18 (3): 1206 - 1217.

- (9): 88 - 90, 94.
- 5 罗玲霞, 李婧, 吴芳, 等. 基于 MBI - GS 量表调查急诊科医务人员职业倦怠现状及影响因素 [J]. 中南大学学报 (医学版), 2024, 49 (6): 981 - 988.
 - 6 王静, 宋慧敏, 彭志勇, 等. 中国大陆中南地区 ICU 医护职业倦怠现状及危险因素 [J]. 中华重症医学电子杂志 (网络版), 2020, 6 (4): 418 - 423.
 - 7 戴涛明, 刘军军, 吴丽华. 江苏省儿科医务人员职业倦怠现状及影响因素分析 [J]. 江苏卫生事业管理, 2023, 34 (7): 984 - 988.
 - 8 楼小花, 林梅芳, 潘雨萍. 三甲妇产科医院医护人员职业倦怠状况调查及影响因素分析 [J]. 现代实用医学, 2018, 30 (10): 1392 - 1394.
 - 9 孙彦彦, 王瑾, 刘晓曼, 等. 中国医务人员职业倦怠水平及影响因素 [J]. 环境与职业医学, 2023, 40 (4): 405 - 412.
 - 10 吴志伟, 陈雨, 李东升, 等. 医务人员职业倦怠和组织支持感典型相关分析 [J]. 中国职业医学, 2019, 46 (3): 322 - 325, 330.
 - 11 陈海梅, 张敏, 沙悦, 等. 医生职业倦怠现状及危险因素研究: 基于系统性综述 [J]. 职业卫生与应急救援, 2024, 42 (3): 385 - 391.
 - 12 毛会, 桑振修, 梁锦铭, 等. MBI - HSS 量表评估基本公共卫生服务人员职业倦怠的信度和效度 [J]. 预防医学情报杂志, 2021, 37 (12): 1711 - 1715, 1721.
 - 13 杨正敏, 黄雪娟, 梁雪萌, 等. 医学生情绪感知与一般心理健康的影响因素调查 [J]. 昆明医科大学学报, 2024, 45 (9): 136 - 143.
 - 14 SELIGMAN M E. Flourish: a visionary new understanding of happiness and well - being [M]. New York: Simon and Schuster, 2011.
 - 15 聂玉真, 王蓓蓓, 田书亚. 基于 PERMA 模型的积极心理干预对冠心病经皮冠状动脉介入治疗患者的影响 [J]. 河南医学研究, 2024, 33 (11): 2089 - 2094.
 - 16 王芹, 吴秋盈, 王琳, 等. 基于 PERMA 模式的护理对乳腺癌患者干预效果的 Meta 分析 [J]. 护理管理杂志, 2024, 24 (8): 705 - 709.
 - 17 刘俐君, 安俊红, 张英, 等. 基于 PERMA 模式的积极心理干预改善糖尿病足病人身心症状的效果 [J]. 护理研究, 2024, 38 (4): 714 - 718.

(上接第 85 页)

- 24 GRANDINI M, BAGLI E, VISANI G. Metrics for multi - class classification: an overview [EB/OL]. [2024 - 05 - 11]. <https://arxiv.org/abs/2008.05756>.
- 25 WANG J, PEREZ L. The effectiveness of data augmentation in image classification using deep learning [EB/OL]. [2024 - 05 - 11]. <https://arxiv.org/abs/1712.04621>.
- 26 FU B, ZHANG M, HE J, et al. StoHisNet: a hybrid multi - classification model with CNN and transformer for gastric pathology images [EB/OL]. [2024 - 05 - 11]. <https://www.sciencedirect.com/science/article/pii/S0169260722003066>.
- 27 LU T, WAN L, QI S, et al. Land cover classification of UAV remote sensing based on transformer - CNN hybrid architecture [J]. Sensors, 2023, 23 (11): 5288.
- 28 XIE Z, LIN Y, YAO Z, et al. Self - supervised learning with swin transformers [EB/OL]. [2024 - 05 - 11]. <https://arxiv.org/abs/2105.04553>.
- 29 HILLER M, MA R, HARANDI M, et al. Rethinking generalization in few - shot classification [EB/OL]. [2024 - 05 - 11]. https://proceedings.neurips.cc/paper_files/paper/2022/file/1734365bbf243480dbc491a327497cfl_Paper_Conference.pdf.
- 30 WU J, TIAN X, ZHONG G. Supervised contrastive representation embedding based on transformer for few - shot classification [J]. Journal of physics: conference series, 2022, 2278 (1): 1 - 9.
- 31 IQBAL A, SHARIF M. BTS - ST: swin transformer network for segmentation and classification of multimodality breast cancer images [EB/OL]. [2024 - 05 - 11]. <https://www.sciencedirect.com/science/article/pii/S0950705123001430>.
- 32 CAI G, ZHU Y, WU Y, et al. A multimodal transformer to fuse images and metadata for skin disease classification [J]. The visual computer, 2023, 39 (7): 2781 - 2793.
- 33 XIE P, SHI S, YANG S, et al. Improving the transferability of adversarial examples with a noise data enhancement framework and random erasing [EB/OL]. [2024 - 05 - 11]. <https://doi.org/10.3389/fnbot.2021.784053>.
- 34 CRESWELL A, WHITE T, DUMOULIN V, et al. Generative adversarial networks: an overview [J]. IEEE signal processing magazine, 2018, 35 (1): 53 - 65.