韦新怡. 基于多模态脑影像融合的阿尔茨海默病早期诊断方法[J]. 智能计算机与应用,2026,16(8):160-166. DOI:10.20169/j.issn.2095-2163.24121901

基于多模态脑影像融合的阿尔茨海默病早期诊断方法

韦新怡

(南京邮电大学 计算机学院软件学院网络安全学院, 南京 210023)

摘 要: 阿尔茨海默病(Alzheimer's Disease,AD)是一种严重影响全球老年人口健康的退行性疾病。早期发现可以防止患者进一步受损,这对治疗阿尔茨海默病至关重要。在过去的几十年里,已经证明神经影像学可以成为 AD 的关键诊断工具,不同神经影像学特征融合可以提高诊断性能。因此,本文提出了一种基于多模态数据的深度学习框架,用于阿尔茨海默病的早期诊断与分类。该框架基于结构磁共振成像(Magnetic Resonance Imaging,MRI)和正电子发射断层扫描(Positron Emission Tomography,PET)图像实现 AD 四分类任务。模型利用预训练的 VGG19 模型结合冻结策略,从预处理后的 MRI 和 PET 数据中提取关键特征。同时,特征融合模块引入 BLOCK 模型,通过超对角张量分解策略表示多模态特征的高阶交互。在 ADNI 数据集上的实验结果表明,本文框架在 AD 的早期诊断中具有显著优势,准确率达到 92.5%,精确率为 93.36%,召回率为 91.82%, *F*1 值为 92.17%。

关键词: 多模态; 深度学习; VGG; 阿尔茨海默病; 双线性模型

中图分类号: R749.1+6; TP181 **文献标志码**: A **文章编号**: 2095-2163(2026)08-0160-07

An early diagnosis method for Alzheimer's disease based on multimodal brain imaging fusion

WEI Xinyi

(School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Alzheimer's disease (AD) is a degenerative disease that severely affects the health of the elderly population worldwide. Early detection can prevent further damage to the patient, which is crucial for the treatment of Alzheimer's disease. Over the past decades, it has been demonstrated that neuroimaging can be a key diagnostic tool for AD and that fusion of different neuroimaging features can improve diagnostic performance. Therefore, a deep learning framework based on multi-modal data is proposed for early diagnosis and classification of Alzheimer's disease. This framework employs structural magnetic resonance imaging (MRI) and positron emission tomography (PET) data to perform four-class classification of AD. The model utilizes a pre-trained VGG19 model combined with a freezing strategy to extract key features from preprocessed MRI and PET data. Meanwhile, the feature fusion module introduces the BLOCK model to represent the higher-order interactions of multimodal features through the super-diagonal tensor decomposition strategy. The experimental results on the ADNI dataset show that the proposed framework has significant advantages in the early diagnosis of AD, with an accuracy of 92.5%, a precision of 93.36%, a recall of 91.82%, and an *F*1 value of 92.17%.

Key words: multimodal; deep learning; VGG; Alzheimer's disease; bilinear model

0 引 言

阿尔茨海默病(AD)是一种常见的神经退行性疾病,主要表现为记忆丧失、认知功能下降和行为改变[1]。随着全球人口老龄化的加剧,阿尔茨海默病的患病率逐年上升,给个人健康、家庭和社会带来沉重负担。早期诊断和干预被认为是延缓疾病进程的关键,但目前临床上对该病的诊断仍主要依赖于临床症状和认知功能评估,缺乏客观的生物标志物和有效的诊断手段[2]。

磁共振成像(MRI)和正电子发射断层扫描(PET)作为神经影像学中的重要工具,已成为阿尔茨海默病诊断的重要辅助手段。MRI 能够提供大脑结构的信息,如脑萎缩的程度,而 PET 能够反映脑

作者简介: 韦新怡(2000—),女,硕士研究生。主要研究方向:图像处理,计算机视觉。Email:1222045802@njupt.edu.cn。

收稿日期: 2024-12-19

内代谢活性，并可检测淀粉样斑块沉积等生物标志物[3]。结合这两种影像数据，可以更全面地评估阿尔茨海默病的病理特征，进而提高诊断的准确性和早期检测的灵敏度[4]。

近年来，深度学习在医学影像分析领域取得了显著进展[5]。尤其是卷积神经网络（Convolutional Neural Network，CNN）已被广泛应用于脑部影像数据的自动化分析，如脑组织分割、病变检测和疾病分类等任务[6]。

Gunawardena 等[7]在 ADNI 数据集上比较支持向量机（SVM）和卷积神经网络（CNN）模型，结果表明 CNN 在阿尔茨海默病医学影像数据处理上效果更优。Solano-Rojas 等[8]基于磁共振成像数据使用三维 DenseNet 模型检测阿尔茨海默病，并构建了深度神经网络分类器。Awang 等[9]基于预训练的 DenseNet-201 模型，利用 MRI 切片数据诊断阿尔茨海默病的不同阶段，并通过数据增强方法提升了 DenseNet-201 的准确率。Yagis 等[10]提出了一种端到端的阿尔茨海默病检测方法，通过卷积自编码器结合监督预测和无监督表征学习，在 OASIS 脑部 MRI 数据集上的实验表明，该方法在准确性上优于传统分类器。Venugopalan 等[11]结合 MRI、遗传数据和临床数据，采用堆叠去噪自编码器提取临床和遗传数据特征，使用 3D 卷积神经网络处理影像数据。随后将三种特征进行特征拼接，将 AD 患者分为 AD、MCI 和 NC 三类。

然而，尽管这些方法在某些任务中取得了令人鼓舞的成果，但仍面临许多挑战，尤其是在多模态数据融合与小样本学习等方面。由于医学数据样本通常较为有限，易导致模型过拟合和泛化能力不足等问题。此外，现有方法通常忽略了不同模态之间潜在的交互信息，而这种信息的有效融合对于提高诊断性能至关重要。

本文提出了一种基于多模态数据融合的深度学习框架，旨在提升阿尔茨海默病的早期诊断能力。具体而言，本文利用 MRI 和 PET 数据，通过特征提取和融合模块结合深度学习模型，提供了一种新的多模态诊断思路。数据预处理模块采用 SPM12 工具对 MRI 和 PET 数据进行标准化处理；特征提取模块基于预训练 VGG19 模型并结合冻结策略进行训练，充分利用预训练模型学到的通用知识；特征融合模块引入 BLOCK 模型对 MRI 与 PET 数据进行有效融合，进一步提高了诊断的准确性和鲁棒性。

1　材料和方法

1.1　数据集

本研究使用阿尔茨海默病神经影像学倡议（Alzheimer’s Disease Neuroimaging Initiative，ADNI）数据库[12]中的 280 例样本作为实验数据。每例样本均包含 MRI 和 PET 影像数据。数据集按照病理状态分为 4 个类别：正常对照组（NC）、早期轻度认知障碍（EMCI）、晚期轻度认知障碍（LMCI）和阿尔茨海默病（AD）。本次实验的数据集规格见表 1。简易精神状态检查（MMSE）是一种常用于评估个体认知功能的量表。根据 MMSE 评分的数值大小，可以有效反映患者的认知水平，通常得分较高者认知功能正常，而得分较低则提示可能存在认知障碍，得分显著低于正常值则可能与阿尔茨海默病等神经退行性疾病相关。

表 1　数据集统计表

Table 1　Dataset statistics

类别	数量	MMSE	男性数量	女性数量
AD	49	23.10±1.91	27	22
NC	65	28.98±1.15	43	22
EMCI	117	28.32±1.74	59	58
LMCI	53	27.18±1.65	33	20

1.2　方法

本研究框架包含 4 个模块。（1）数据预处理模块：使用 SPM12 工具对 MRI 和 PET 数据进行处理。（2）特征提取模块：基于预训练 VGG19 模型，并结合冻结策略进行训练。（3）特征融合模块：采用 BLOCK 模型对 MRI 与 PET 数据进行融合。（4）分类模块：通过正则化和全连接层完成分类分数输出。框架图如图 1 所示。

1.2.1　数据预处理

在 MRI 和 PET 数据的采集过程中，由于个体差异和环境因素的影响，图像强度、运动伪影等问题会对数据的质量产生影响。此外，直接处理原始的三维数据不仅计算复杂，且耗时较长。因此，为了提高数据处理的效率和质量，本文对数据进行了预处理，以去除原始数据中的不相关信息，并获得统一大小和强度的二维切片数据。

MRI 数据预处理流程：

（1）AC-PC 配准：通过对图像进行标准化处理，使不同受试者的解剖结构对齐。

（2）图像配准：通过将图像配准到大脑标准模板 MNI152 空间中，提高群体研究的准确性和可比性。

(3)颅骨去除:使用自动化算法去除脑外的颅骨部分,减少不相关的组织信息干扰,保留脑部区域,提高后续分析的有效性。

(4)平滑处理:进行图像平滑处理,减少噪声和伪影,确保图像的细节更加清晰,从而提高特征提取和分析的可靠性。

(5)强度归一化:采用强度归一化方法对图像数据进行处理,将图像的灰度值范围统一,消除不同扫描条件下的强度差异,确保数据在后续处理中的一致性和可比性。此外,SPM12 可以将脑组织分割为灰质、白质和脑脊液。在本研究中只关注灰质切片来检测 AD。

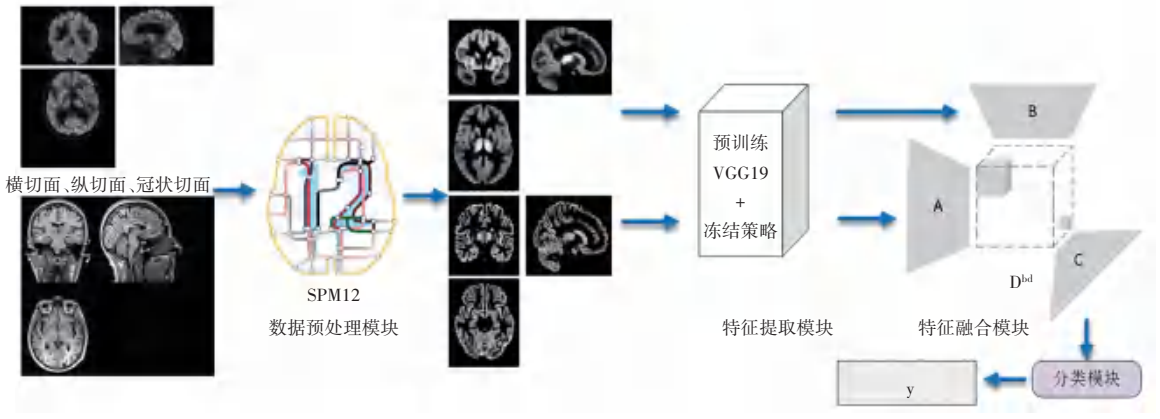


图 1 系统框架图

Fig. 1 System architecture diagram

PET 数据预处理流程:

(1)数据校正:时间校正用于调整扫描过程中因放射性核素衰变导致的信号强度变化,运动校正通过配准算法矫正患者扫描过程中的位置偏移,降低运动伪影对图像质量的影响。

(2)图像配准。

(3)强度归一化。

在数据预处理过程中,所有的二维切片经过处理后统一调整为 224×224 的尺寸。通过这一预处理步骤,保证了输入数据的一致性,从而提高了模型的训练效率和稳定性。考虑到切片中包含的信息量不同,选择富有信息的切片对于模型的学习至关重要。本文对每个样本从横切面、纵切面和冠状切面中各选取 20 张高质量的图像,模型的输入特征包含每个切面的信息。这种三通道的输入特征不仅可以更全面地反映大脑的空间结构信息,还能够为模型提供更多的多维度数据,有助于提高模型对复杂脑部特征的识别能力。

1.2.2 结合迁移学习与冻结策略的 VGG19 模型

迁移学习是一种将预训练模型应用于新任务,以加速模型训练并提升性能的方法。在医学图像处理领域,迁移学习尤为重要,因为构建一个高性能的深度学习模型通常需要大量标注数据,而这类数据在实际应用中往往难以获取[13]。通过迁移学习,可以利用在大规模公开数据集(如 ImageNet)上预训练得到的模型,在规模较小或类别不平衡的医学数据集上进行微调,从而显著缩短训练时间并增强模型的泛化能力。

在迁移学习的应用中,冻结策略常被用于进一步提升训练效率和模型性能。该策略通过固定模型部分层的参数,使其在训练过程中不参与更新。通常,这些冻结的层是预训练模型的低层网络,其学习到的特征较为通用,并能在不同任务中共享。通过冻结这些层,既能避免通用特征被过度调整,让模型专注于训练任务特定的高层特征,又能减少计算负担。不仅提升训练效率,还能缓解小样本学习中的过拟合问题,增强模型的泛化能力。

在模型选择上,VGG19[14] 作为基础架构,原因在于其在深度学习中应用广泛和性能优异。与 Inception-Net、Xception-Net、ResNet 等其他网络架构相比,VGG19 拥有 1.38 亿参数,具备强大的特征提取能力,适配复杂的医学图像任务。可对 MRI 和 PET 切片数据进行关键特征提取并分类。VGG-19 包含 19 层网络深度,涵盖卷积层、池化层和全连接层。其设计理念是通过使用小的卷积核(3×3)和多层堆叠的卷积层来加深网络的层次结构,从而增加模型的非线性表达能力。

VGG19 通常采用最大池化策略。在 2×2 的池化中,池化操作可表示为:

$$I'_{i,j} = \max(I^*_{2i-1,2j-1}, I^*_{2i-1,2j}, I^*_{2i,2j-1}, I^*_{2i,2j}) \quad (1)$$

其中,i 和 j 为特征图中位置;I' 是池化后的输出特征图;I^* 表示卷积后的输出特征。

ReLU 激活函数是 VGG19 中常用的非线性激活函数，数学表达式为：

$$\mathrm{ReLU} = \begin{cases} 0, & \text{if } v < 0 \\ x, & \text{if } x \geqslant 0 \end{cases} \tag{2}$$

本研究中使用的 VGG-19 模型采用了 ImageNet 数据库的预训练权重，并基于迁移学习进行训练。在训练过程中，通过迁移学习方法，冻结了模型中的 block1、block3、block5 层的权重，以加速训练过程并提高模型在目标任务上的性能。此外，在进行单模态 AD 预测时，还需要冻结部分全连接层的权重并添加新的全连接层和 Softmax 层以进行 4 分类，网络结构如图 2 所示。

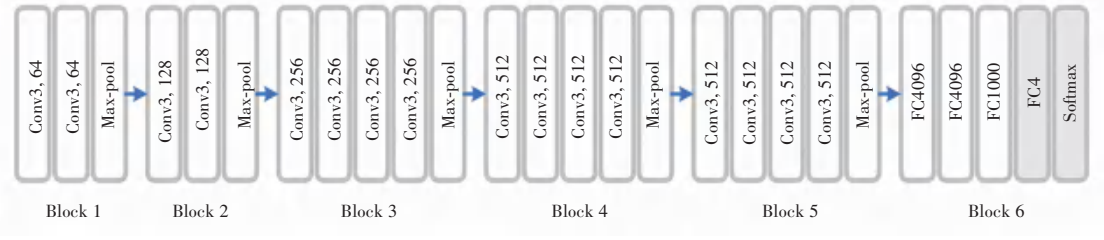


图 2　VGG 框架图

Fig. 2　VGG architecture diagram

1.2.3　双线性超对角融合模型 BLOCK

BLOCK 模型最早用于视觉问答和视觉关系检测领域[15]。该模型结合双线性映射与超对角矩阵结构强化两种模态之间的交互作用。

双线性映射可融合视觉特征与语言特征，使得模型能够捕捉两种模态之间更为细致和丰富的交互信息。具体而言，双线性模型将两个向量通过张量积 $\boldsymbol{T}$ 将两个输入向量投影到 K 维空间中去，计算公式为：

$$y = \boldsymbol{x}_1 \boldsymbol{T} \boldsymbol{x}_2^{\mathrm{T}},\ \boldsymbol{T} \in R^{\dim(x_1)\times\dim(x_2)\times K} \tag{3}$$

其中，$\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 为输入向量；$\boldsymbol{T}$ 为投影矩阵；y 为输出结果。

然而，直接使用双线性映射可能会导致模型参数数量急剧增加，特别是在处理大规模数据时。因此，BLOCK 模型引入了块项分解策略来表示双线性映射。如图 3 所示，块项分解通过将张量积分解为多个较小的矩阵或向量，从而有效减少了计算复杂度和模型参数数量。这一策略不仅降低了模型的训练和推理成本，还能提升其在大规模数据集上的泛化能力和计算效率。具体而言，$\boldsymbol{T}$ 在秩 (L, M, N) 项中的分解定义为：

$$\boldsymbol{T} := \sum_{r=1}^{R} \boldsymbol{D}_r \times \boldsymbol{A}_r \times \boldsymbol{B}_r \times \boldsymbol{C}_r \tag{4}$$

其中，$\boldsymbol{T}$ 被分解为 R 个较小的矩阵；$\boldsymbol{D}_r$ 是一个大小为 $L \times M \times N$ 的块 - 超对角张量；$\boldsymbol{A}_r$、$\boldsymbol{B}_r$ 和 $\boldsymbol{C}_r$ 分别是大小为 $I \times L$、$J \times M$ 和 $K \times N$ 的矩阵。更简洁的表达方式为：

$$\boldsymbol{T} = \boldsymbol{D}^{bd} \times \boldsymbol{A} \times \boldsymbol{B} \times \boldsymbol{C} \tag{5}$$

然后通过矩阵分解将输入特征 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 分别通过映射矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 进行投影，得到 $\boldsymbol{x}_1' = \boldsymbol{A}\boldsymbol{x}_1$，$\boldsymbol{x}_2' = \boldsymbol{B}\boldsymbol{x}_2$。然后将这两个投影与块 - 超对角张量融合，得到 $\boldsymbol{z}_r$。

$$z_r = \boldsymbol{D}_r \times \boldsymbol{x}_{1r}' \times \boldsymbol{x}_{2r}' \tag{6}$$

最后将 R 个 $\boldsymbol{z}_r$ 合并得到 $\boldsymbol{z}$，最后的预测向量是 $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{z}$。通过这种块项分解方法，BLOCK 模型有效减少了模型的参数数量和计算复杂度，同时保留了双线性映射中关键的交互信息。

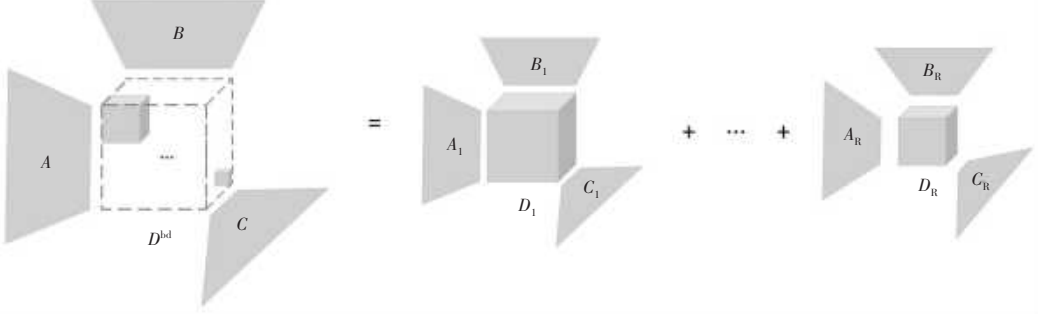


图 3　块项分解技术

Fig. 3　Block decomposition technique

2 实验与分析

2.1 评价指标

为评估本文方法的有效性,采用准确率(Accuracy)、精确率(Precision)、召回率(Recall)和 $F1$ 值作为评价指标。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{7}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

$$F1 = 2 \times \frac{\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{10}$$

其中,TP 为正确预测的正类样本;FP 为负类样本错误预测为正类样本;FN 为正类样本错误预测为负类样本;TN 为正确预测为负类样本。

为进一步评估模型在不同类别上的表现,本文引入了宏平均指标,即对每个类别的精度、召回率和 $F1$ 值分别计算,取其平均值。这种方法能够综合反映模型在各类别上的整体表现,尤其适用于类别不平衡场景,可避免模型评价偏向大类的问题。

2.2 实验环境与参数设置

本研究模型基于 PyTorch 2.0.1 实现,所有实验均在 Windows 11 操作系统上运行,硬件环境包括 NVIDIA GeForce RTX 3060 Ti GPU、Intel(R) Core(TM) i5-13400F 处理器以及 16 GB 内存。模型优化使用了 Adam 优化器,单模态单次训练的数据批量大小为 32,多模态单次训练的批次大小为 16,初始学习率设置为 1e-4,实验训练次数为 100。为动态调整学习率,采用了基于 Lambda 函数的学习率调度器(LambdaLR),该调度器根据训练轮数以 0.95 的衰减系数逐步减少学习率,从而帮助模型在训练过程中更好地收敛。

2.3 实验结果

在模型训练过程中,随机选取数据集的 70%作为训练集,其余部分按验证集和测试集各占 15%的比例进行划分,用于模型性能的验证与评估。为全面评估本文提出的多模态方法的有效性,并与单模态方法进行对比,表 2 总结了不同模态下模型的性能指标。本研究分别基于 PET、MRI 以及多模态 PET+MRI 数据开展实验,计算了准确率、宏精确率、宏召回率和宏 $F1$ 值。结果显示,多模态方法(PET+MRI)在各项指标上表现最佳,准确率为 0.925 0,宏平均精度为 0.933 6,宏平均召回率为 0.918 2,宏平均 $F1$ 值达到了 0.921 7。其次是 PET 数据,模型的表现也较为优秀。相比之下,MRI 单模态数据的性能稍逊于 PET。以上结果表明,多模态方法通过有效融合不同模态的数据优势,显著提升了阿尔茨海默病早期诊断的准确性和鲁棒性。

表 2 不同模态及其融合模型的性能比较

Table 2 Performance comparison of different modalities and multi-modal fusion models

模态	准确率	宏精确率	宏召回率	宏 $F1$ 值
PET	0.904 8	0.904 6	0.918 0	0.905 0
MRI	0.857 1	0.843 0	0.860 4	0.850 4
PET+MRI	**0.925 0**	**0.933 6**	**0.918 2**	**0.921 7**

通过混淆矩阵(见图 4 和图 5),可以直观地观察到模型在不同类别之间的区分能力。具体而言,混淆矩阵提供了模型在各类别上的真阳性、假阳性、真阴性和假阴性的详细信息。此外,通过矩阵可视化,能够直观地呈现模型在各个类别之间的预测分布情况。这种可视化方式有助于深入分析模型在不同类别上的分类表现,揭示其在区分各类样本时的优劣势。如图 4(a)所示,模型在区分 NC 和 EMCI 时存在一定的混淆。

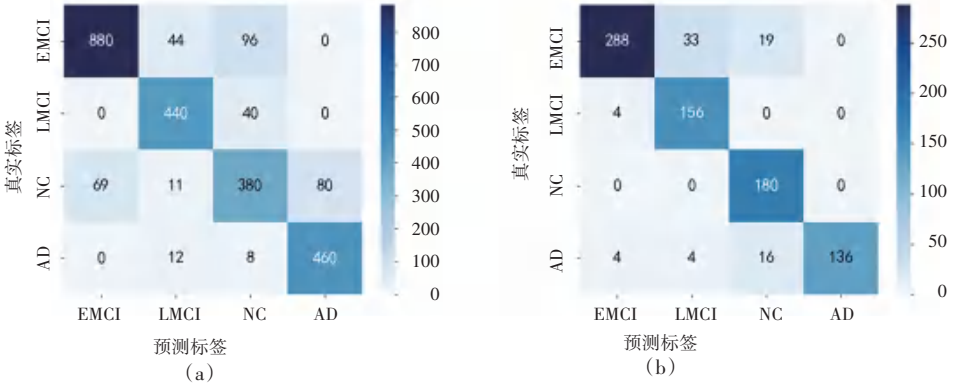


图 4 基于单模态数据的混淆矩阵

Fig. 4 Confusion matrix based on single-modal data

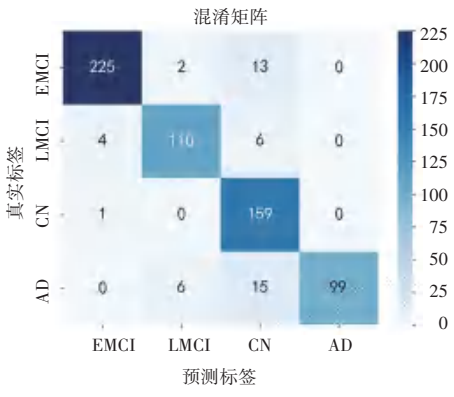


图 5　基于多模态数据的混淆矩阵

Fig. 5　Confusion matrix based on multi-modal data

为验证本文方法的有效性，将本文方法与在 ADNI 数据集上进行的 5 种现有方法进行对比。Ahmed 等[16]从 MRI 中提取局部影像生物标志物，采用多核学习方法以识别 AD 患者。Glozman 等[17]主要分析海马体和脑室中组织退化引起的细微形状变化来对 AD 进行分类。Chen 等[18]提出一种基于注意力机制的多模态融合网络，首先通过 CNN 提取 MRI、PET 和临床数据的特征，然后通过通道注意力和空间注意力机制进行数据融合。Kong 等[19]通过将 PET 图像与其相对应的 MRI 图像配准，生成融合影像，然后利用 3D CNN 网络进行 AD 分类分析。Song 等[20]首先对 PET 和 MRI 进行预处理，然后提出一种简单的三维卷积神经网络，用于快速评估影像融合方法的效果。对比结果见表 3，本文方法在性能上优于单模态方法，且在多模态 AD 分类任务中也优于现有的相关方法。

表 3　与现有相关方法比较

Table 3　Comparison with existing methods

模型	模态	准确率/%	类别
多核学习[16]	MRI	90.20	AD, NC
统计学习[17]	MRI	83.50	AD, MCI, NC
注意力机制融合网络[18]	MRI+PET+临床数据	81.10	AD, MCI, NC
3D CNN[19]	MRI+PET	87.67	AD, MCI, NC
3D Simple CNN[20]	MRI+PET	74.54	AD, NC, MCI
本文方法	MRI+PET	**92.50**	AD, NC, LMCI, EMCI

3　可视化研究

计算机辅助诊断在促进临床应用中具有显著潜力。AD 的诊断通常依赖于分析大脑的结构变化，如识别异常萎缩区域等。本文提出的方法作为一种有效的辅助诊断工具，可自动定位 MRI 图像中的潜在病理区域，为医生提供参考，以识别病变位置。为更好地解释模型的决策逻辑，本文对模型学习的特征进行了研究，并直观展示了其关注的关键区域。通过采用 Grad-CAM[21]技术生成热力图，可以揭示卷积层特征激活所对应的重点区域。这些热力图不仅揭示了模型关注的图像部分，还有助于理解模型进行分类的依据，如图 6 所示，不同阶段的热力图关注的区域各不相同。对于 AD 患者，模型在冠状切面上更关注大脑上部区域；而对于 NC，模型则聚焦于大脑中央区域。此外，如图 6(a)、(b)所示，不同轻度认知障碍阶段的关注区域亦存在显著差异。

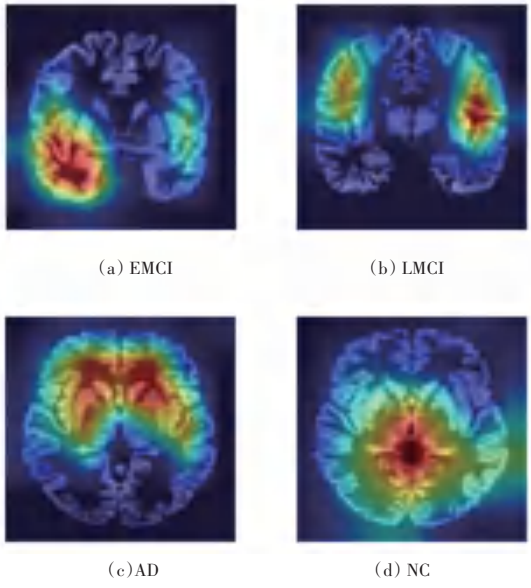


图 6　使用 Grad-CAM 的 AD 可视化结果

Fig. 6　AD visualization results using Grad-CAM

4　结束语

本文提出了一种基于迁移学习方法与双线性模型相结合的阿尔茨海默病诊断模型。该模型通过结合预训练模型与冻结策略，成功缓解了小样本数据在深度学习应用中的挑战，显著提高了模型在有限数据条件下的泛化能力。在多模态数据的处理上，本文引入了双线性模型用于高效融合 MRI 与 PET 影像数据，能够最大程度地提取两种模态之间的深层特征交互关系。为进一步提升模型的表示能力，本文采用了超对角张量分解策略，对多模态数据的高阶交互进行建模，从而在保持数据复杂性表达的同时有效降低了计算复杂度和模型参数规模。此外，为了确保影像数据的质量和一致性，本文使用了 SPM12 工具对原始影像数据进行了标准化预处理。该步骤为后续模型训练提供了高质量的输入数据，有助于提高模型的稳定性和可靠性。综上所述，本

文提出的基于迁移学习与双线性模型的 AD 诊断框架,不仅有效解决了小样本数据和多模态数据融合问题,还在提高诊断精度和降低计算负担方面展现了显著的优势。未来工作可考虑进一步优化特征选择和模型细节,以推动该方法在临床实践中的应用。

参考文献

[1] Alzheimer′s Association. 2024 Alzheimer′s disease facts and figures[J]. Alzheimers Dement, 2024, 20(5): 3708-3821.

[2] Jin Boru, Cheng Xiaoqin, Fei Guoqiang, et al. Identification of diagnostic biomarkers in Alzheimer′s disease by integrated bioinformatic analysis and machine learning strategies [J]. Frontiers in Aging Neuroscience, 2023, 15: 1169620.

[3] Nordberg A, Rinne J, Kadir A, et al. The use of PET in Alzheimer disease[J]. Nature Reviews Neurology, 2010, 6(2): 78-87. DOI: 10.1038/nrneurol.2009.217.

[4] Sharma S, Mandal P K. A comprehensive report on machine learning-based early detection of Alzheimer′s disease using multi-modal neuroimaging data [J]. ACM Computing Surveys, 2022, 55(2): 1-44.

[5] Yao Zhaomin, Wang Hongyu, Yan Wencheng, et al. Artificial intelligence-based diagnosis of Alzheimer′s disease with brain MRI images [J]. European Journal of Radiology, 2023, 165: 110934.

[6] Zhang Li, Wang Mingliang, Liu Mingxia, et al. A survey on deep learning for neuroimaging-based brain disorder analysis[J]. Frontiers in Neuroscience, 2020, 14: 779.

[7] Gunawardena K A N N P, Rajapaks E R N, Kodikara N D. Applying convolutional neural networks for pre-detection of Alzheimer′s disease from structural MRI data [C]//Proceedings of 2017 24th International Conference on Mechatronics and Machine Vision in Practice. Piscataway, NJ: IEEE, 2017: 1-7.

[8] Solano-Rojas B, Villalon-Fonseca R. A low-cost three-dimensional DenseNet neural network for Alzheimer′s disease early discovery [J]. Sensors, 2021, 21(4): 1302.

[9] Awang M K, Zia-ur-Rehman M K, Rashid J, et al. Classification of Alzheimer disease using DenseNet-201 based on deep transfer learning technique [J]. PLoS One, 2024, 19(9): e0304995.

[10] Yagis E, Seco de Herrera A G, Citi L. Convolutional autoencoder based deep learning approach for Alzheimer′s disease diagnosis using brain MRI [C]//Proceedings of 2021 IEEE 34th International Symposium on Computer-based Medical Systems. Piscataway, NJ: IEEE, 2021: 486-491.

[11] Venugopalan J, Tong L, HassanzaDeh H R, et al. Multimodal deep learning models for early detection of Alzheimer′s disease stage [J]. Scientific Reports, 2021, 11(1): 3254.

[12] Mueller S G, Weiner M W, Thal L J, et al. Ways toward an early diagnosis in Alzheimer′s disease: The Alzheimer′s Disease Neuroimaging Initiative[J]. Alzheimer′s & Dementia, 2005, 1(1): 55-66.

[13] 潘璐璐, 余勇夫, 秦国友. 迁移学习简介及其在医学研究领域中的应用[J]. 复旦学报(医学版),2024, 51(6): 1016-1020.

[14] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [C]//Proceedings of the International Conference on Learning Representations. 2015: 1-14.

[15] Ben-Younes H, Cadene R, Thome N, et al. Block: Bilinear superdiagonal fusion for visual question answering and visual relationship detection [C]//Proceedings of the AAAI Conference on Artificial Intelligence. AAAI ,2019: 8102-8109.

[16] Ahmed O B, Benois-Pineau J, Allard M, et al. Recognition of Alzheimer′s disease and mild cognitive impairment with multimodal image-derived biomarkers and multiple kernel learning [J]. Neurocomputing, 2017, 220: 98-110.

[17] Glozman T, Solomon J, Pestilli F, et al. Shape-attributes of brain structures as biomarkers for Alzheimer′s disease [J]. Journal of Alzheimer′s Disease, 2017, 56(1): 287-295.

[18] Chen Hui, Guo Huiru, Xing Longqiang, et al. Multimodal predictive classification of Alzheimer′s disease based on attention-combined fusion network: Integrated neuroimaging modalities and medical examination data [J]. IET Image Processing, 2023, 17(11): 3153-3164.

[19] Kong Zhaokai, Zhang Mengyi, ZHU Wenjun, et al. Multi-modal data Alzheimer's disease detection based on 3D convolution [J]. Biomedical Signal Processing and Control, 2022, 75: 103565.

[20] Song Juan, Zheng Jian, Li Ping, et al. An effective multimodal image fusion method using MRI and PET for Alzheimer′s disease diagnosis [J]. Frontiers in Digital Health, 2021, 3: 637386.

[21] Selvaraju R R, Cogswell M, Das A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization [C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2017: 618-626.