

## 基于 3D 残差网络的阿尔茨海默病辅助诊断模型研究\*

张 烁<sup>1△</sup> 张 荣<sup>2,3</sup> 郭颖婕<sup>4</sup> 余红梅<sup>3</sup>

**【摘要】目的** 针对阿尔茨海默病(AD)、轻微认知障碍(MCI)、正常人(CN)的多分类问题,融合 MRI 和 PET 图像特征,探究出一种用于 AD 早期辅助诊断的多分类深度学习模型。**方法** 利用两个独立的 3D-ResNeXt 网络模型对多模态医学图像进行训练。然后通过相关性分析来判断两个网络输出结果的一致性。最后通过计算 PET 与 MRI 图像的诊断结果之间的皮尔逊系数得到多模态神经影像辅助诊断结果。**结果** 模型准确率、模型 Weighted-Precision、模型 Weighted-Recall、皮尔逊系数在融合所有策略 3D-ResNeXt 网络模型的结果为 87.78%、88.32%、87.87%、86.40%;对于图像切片数为 80 的结果为 88.21%、84.7%、88.23%、85.7%、86.40%;融合多模态 MRI 和 PET 图像的结果为 87.78%、88.32%、87.87%、86.40%。**结论** 基于 3D-ResNeXt 网络模型的诊断效果在许多指标上都优于其他辅助诊断模型。结果表明,所提出的多模态辅助诊断模型能够达到良好的诊断率。本研究为 AD 早期辅助诊断模型的选择提供了一定的方法学参考。

**【关键词】** 阿尔茨海默病 磁共振影像 正电子发射断层影像 残差网络 多模态

**【中图分类号】** R195.1 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.06.016

近日, Lancet Public Health 发表研究,到 2050 年全球患痴呆的人数预计将增加到 1.53 亿人,其中 70%~80% 为阿尔茨海默病(Alzheimer's disease, AD)导致的痴呆<sup>[1]</sup>。AD 是一种退行性神经系统疾病,表现为认知和记忆不断丧失,严重影响病人的日常生活<sup>[2]</sup>。轻微认知障碍(mild cognitive impairment, MCI)为正常(cognitively normal, CN)和 AD 之间的一种状态, MCI 常被误诊为正常衰老,文献表明,44%的 MCI 可能在几年之内转化为 AD<sup>[3]</sup>。有效的药物治疗和心理治疗可以减缓 MCI 的恶化,提高患者的生活质量。因此, MCI 的准确诊断对于 AD 的早期诊断和治疗显得非常重要。

神经影像学诊断,可以看成是图像识别问题。近些年来,已经有一些研究使用机器学习和深度学习研究方法研究 MRI 和 PET 图像识别问题。如 Oh 等<sup>[4]</sup>使用了一种无监督的深度学习方法,利用卷积自动编码器完成了 AD 和 CN 的分类任务,准确率达为 86.6%,并使用深度迁移学习方法完成轻度认知功能障碍和稳定型轻度认知功能障碍的分类任务,准确率达到 73.95%。Xia 等<sup>[5]</sup>基于 3D 卷积长短期记忆网络使用 MRI 图像完成 AD 和 CN 的分类任务,该模型首先使用卷积神经网络(convolutional neural network, CNN)进行图像的特征提取,然后再使用相干长短期记忆网络进一步提取深层特性信息,分类准确率为 94.2%。

由于单模态神经影像图像只包含与脑异常的部分

信息,对于 CN、MCI 和 AD 的多分类效果可能会不尽如人意。多模态神经影像分类方法可以提供针对不同模态影像的互补特征,从理论上讲可以提高分类的准确性<sup>[6]</sup>。MRI 可以反映 AD 患者脑结构的变化, PET 可以得到大脑功能性特征,增强发现病变的能力<sup>[7]</sup>。MRI 图像和 PET 图像的融合是一种有效的多模态神经影像融合方法,可为医学临床诊断提供更可靠的数据<sup>[8]</sup>。Suk 等<sup>[9]</sup>使用堆叠式自动编码器分别从 MRI 和 PET 图像提取高层特征,然后将多模态融合的特征使用多核支持向量机(support vector machine, SVM)进行分类任务。Ju 等<sup>[10]</sup>提出了一种基于静息状态脑网络和深度学习的方法用于 AD 早期诊断。该方法将 MRI 图像特性和临床相关信息特征进行融合,用来区分 CN、MCI 和 AD,与传统方法相比,预测准确率提高了 31.21%。

Jin 等<sup>[11]</sup>提出了一种改进的 3D 残差网络用来辅助诊断 AD,并在模型中加入了注意力机制提高分类性能,并标注了与 AD 诊断有联系的脑部区域。针对以上研究,虽然在 AD 分类中取得了较好性能,但也存在以下问题:首先,未充分利用 MRI 和 PET 图像的各自互补的特征进行多模态特征的融合;其次,目前大多数研究都是将三维图像切片转化为二维图像后进行网络学习,这样会导致重要的结构信息丢失,影响分类的准确率。针对以上问题,本文提出了一种基于 3D 残差网络,并融合 MRI 和 PET 图像特征的 AD、MCI、CN 的多分类诊断模型。

## 对象与方法

## 1. 数据来源

本文使用的神经影像图像来源于公共数据库阿尔茨海默病神经影像学研究数据库(Alzheimer's disease neuroimaging initiative, ADNI)<sup>[12]</sup>(<https://adni.loni>。

\* 基金项目:国家自然科学基金(82273742);山西省基础研究计划(202203021222261)

1.山西青年职业学院计算机与信息工程系(030032)

2.山西医科大学计算机教学部

3.山西医科大学公共卫生学院

4.山西大学大数据科学与产业研究院

△通信作者:张烁, E-mail: 14034374@qq.com

usc.edu/)上的数据集。本文实验数据分为 AD、MCI、CN 三类,共 820 名受试者的图像信息,其中包含 MCI 300 名,AD 220 名,CN 300 名,每名受试者均有 MRI 和 PET 图像,MRI 图像使用 Axial 方向图像,受试者统计信息如表 1 所示。

表 1 数据集信息统计表

类别	图像数	年龄/岁	性别(男/女)
AD	220	72.99±1.2	98/122
CN	300	72.57±1.4	198/102
MCI	300	72.72±1.4	181/119

### 2. 图像预处理

理论上来说,由于 3D 残差网络是端到端的学习训练过程,可以将原始图像直接送入模型进行训练,从而“自动地”从低到高提取图像特征。但由于神经影像图像样本数量少,图像特征丰富,对于深层网络训练困难,并且很容易发生过拟合的现象,所以,对于合理的图像预处理可以提高网络模型的性能。

对于 MRI 图像,图像预处理分为四个步骤:①将原始图像的格式转换为 NIFTI 格式;②通过放射变化和仿射变换将所有个体的 T1 结构图像在空间上互相配准,所有图像的分辨率和相同的解剖结构与模板对应,这样可以去除个体大脑结构上的差异;③对脑组织进行分割,首先去除头骨等非脑组织,之后根据不同脑组织所属的体素值,运用迭代算法进行参数估计,最

终将大脑组织分为白质、灰质、脑脊液三个部分。④将分割后的图像使用 dartsel 工具将每个个体的灰度密度图标准化到模板,以消除形状的差异。

对于 PET 图像,与 MRI 图像预处理过程类似,对于每个个体的 PET 图像,首先将原始图像转换成 NIFTI 格式;完成预处理后,将每位个体的 PET 图像和对应的 MRI 图像进行图像配准和仿射对齐,防止因为仪器的不同产生差异;然后通过线性和非线性 3D 变换归一化到 MNI 空间,使得图像在空间上大小和分辨率一致;最后,使用高斯核对归一化的 PET 图像进行平滑处理,平滑处理可以模糊图像边缘,提高信噪比。图 1 为神经影像预处理流程图。

由于 ADNI 数据库中,供训练的数据集数量有限,为了有足够的神经影像图像数据进行训练,对预处理后的 MRI 和 PET 图像进行数据扩增。对于 MRI 和 PET 图像,使用 python 中的 albumentations 包来实现图像的数据扩增。①图像翻转,对图像进行 90 度翻转,防止图像出现黑边;②图像裁剪,使用临近插值法对图像进行随机裁剪;③图像变形缩放,对图像进行非刚性转换,包括图像的弹性变换、网格畸变、光学畸变;④图像非空间性变换,包括图像对比度均衡;⑤改变图像的亮度与对比度。将预处理后的每个 MRI 和 PET 图像都进行以上 5 种数据扩增操作,相当于将原始数据集数量扩大 5 倍。

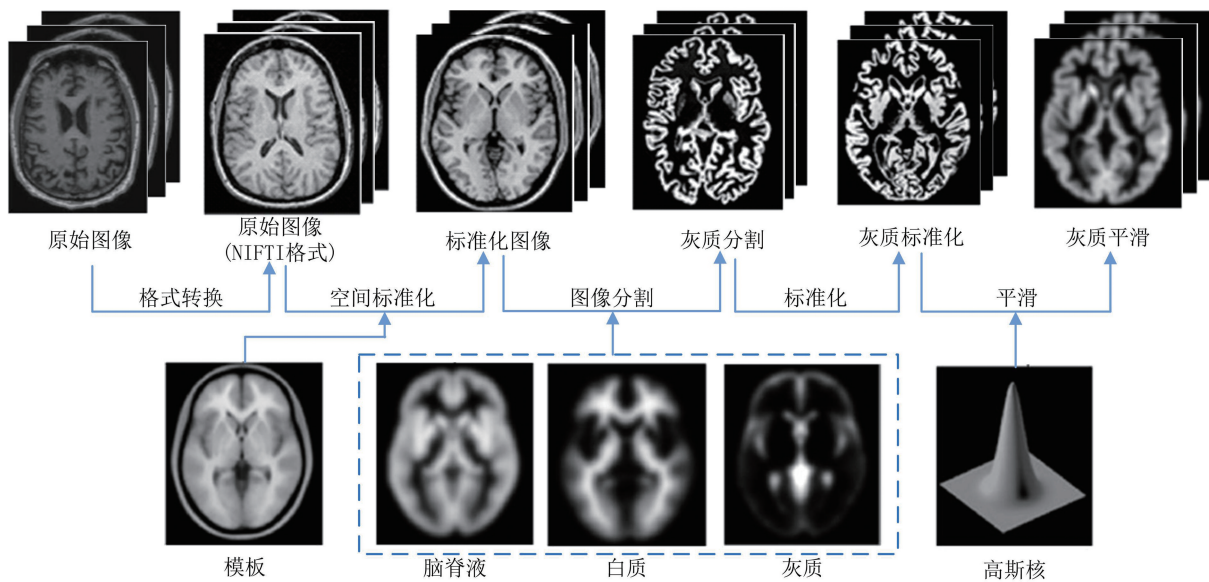


图 1 神经影像预处理流程图

### 3. 改进的 3D 残差网络模型

本文提出了一种用于 AD 辅助诊断的 3D 残差网络模型(3D-ResNeXt),通过一系列的卷积和池化过程,分别从 MRI 和 PET 图像提取特征,然后通过相关分析来判断两个 3D 残差网络输出的一致性。对于输出结果,计算 MRI 和 PET 诊断结果之间的相关性,作

为多模态影像诊断的结果的权重,多模态 3D-ResNeXt 网络辅助诊断模型如图 2 所示。

#### (1) Conv1 层

Conv1 模块通常是由多个卷积层和池化层构成的,该模块主要为了降低图像的分辨率,减少计算。在 ResNet18、ResNet34、ResNet50、ResNet101 等经典残差

网络中,Conv1 是由一个卷积核为  $7 \times 7$ ,步长为 2 的卷积层和一个池化核为  $3 \times 3$ ,步长为 2 的最大池化层构成的。由于 MRI 和 PET 图像为 3 维图像,预处理后的 MRI 和 PET 图像大小统一变为  $224 \times 224 \times n \times 3$  ( $n$  为切片大小)。

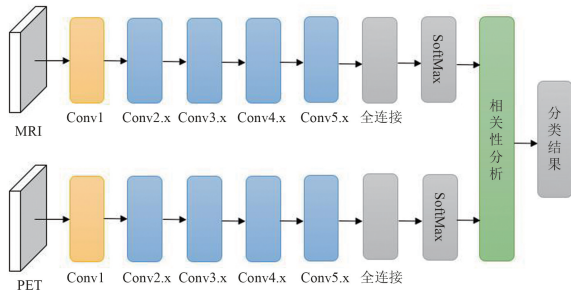


图 2 多模态 3D-ResNeXt 网络辅助诊断模型

为了降低训练参数的数量,将 Conv1 设计为由 3 个卷积层和 1 个最大池化层串联的结构。第 1 个为步长为 2 的  $3 \times 3 \times 3$  卷积层,后串联 2 个步长为 1 的  $3 \times 3 \times 3$  的卷积层,最后串联 1 个步长为 2 的  $3 \times 3 \times 3$  的最大池化层。由感受野计算公式可知,将 1 个步长为 2 的  $7 \times 7 \times 3$  的卷积层修改为 3 个  $3 \times 3 \times 3$  卷积层,感受野保持不变,训练参数减少。

感受野计算公式如下:

$$f_i = (f_{i-1} - 1) \times stride + size \quad (1)$$

其中,  $f_i$  为第  $i$  层感受野,  $stride$  为步长,  $size$  为卷积核大小。

Conv1 模块如图 3 所示。

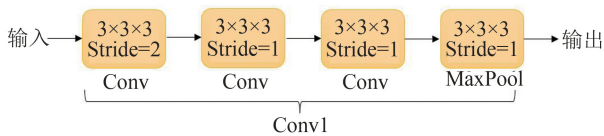


图 3 Conv1 模块

(2) 提出一种 3D 卷积核拆分结构

图 2 中 Conv2.x、Conv3.x、Conv4.x、Conv5.x 为残差网络的瓶颈层。每个瓶颈层是由多个卷积核组成的。对于 3D 卷积核与 2D 卷积核相比,对 1 个维度的信息,会导致网络参数的数量增加,会使模型训练的时间变长。在 Inception-v3 模型<sup>[13]</sup>中,将一个  $3 \times 3$  的卷积核拆分成  $1 \times 3$  和  $3 \times 1$  的 2 种卷积核可以加快模型参数的训练速度。本文提出了一种 3D 卷积核拆分结构(3D convolution kernel split, 3DCK)。对于尺寸为  $d \times k \times k$  的 3D 卷积核( $d$  是卷积核的切片深度, $k$  是卷积核的空间大小)可拆分为  $1 \times k \times k$  卷积核和  $k \times 1 \times 1$  卷积核两种形式,其中  $1 \times k \times k$  卷积核对 MRI 或 PET 图像每个切片图像进行二维特征提取,对于  $k \times 1 \times 1$  对 MRI 或 PET 图像进行深度卷积,提取相邻切片之间的特征信息。3DCK 将参数量更少的切片深度一维信息和二维图像特征充分融合,二维图像的卷积结果作为一维

深度卷积的输入,可以保留丰富的特征信息,从而减少模型尺寸和提高网络训练的速度。图 4 为 3D 卷积核拆分示意图。

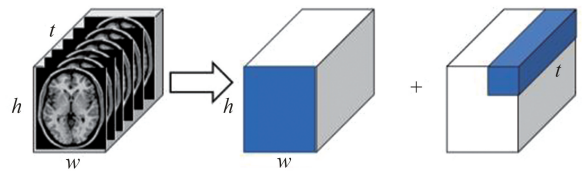


图 4 3D 卷积核拆分示意图

(3) 3D-ResNeXt 网络模型结构设计

残差网络主要解决了网络退化的问题,可以将网络的深度扩展到上千层,并且网络仍具有良好的性能<sup>[14]</sup>。残差网络是由多个残差块堆叠而成的,每一个残差块表示为:

$$x_{i+1} = x_i + f(x_i, w) \quad (2)$$

其中,  $x_i$  和  $x_{i+1}$  分别表示第  $i$  个残差块的输入和输出,  $f$  表示残差函数,  $w$  表示第  $i$  个残差块的网络参数。残差网络保证了数据流从浅层向深层的有效传播,从而提高网络的性能。

本文提出了一种新的残差结构(如图 5 所示),该结构从 2D 向 3D 进行扩展,融合了 3DCK,加快了网络训练速度,提高了网络性能,该结构表示为:

$$x_{i+1} = x_i + f_L(g([x_0, \dots, x_n]), [w_0, \dots, w_n]) + f_R(g([x_0, \dots, x_n]), [w_0, \dots, w_n]) \quad (3)$$

其中,  $x_i$  和  $x_{i+1}$  分别表示第  $i$  个残差块的输入和输出,  $[x_0, \dots, x_n]$  表示第  $i$  个残差块中从第 0 个卷积层到第  $n$  个卷积层特征映射的连接,  $[w_0, \dots, w_n]$  表示第  $i$  个残差块的所有网络参数,  $f_L$  表示网络左边的残差函数,  $f_R$  表示网络右边的残差函数。

如图 5 所示,输入数据分别经过两个分支进行卷积运算,对左边部分进行 2 次  $1 \times 3 \times 3$  卷积操作,用来提取 3D 图像每个切片的特征,或者通过卷积核为  $1 \times 1 \times 1$  捷径连接;对右半部分进行 2 次  $3 \times 1 \times 1$  卷积操作,用来提取 3D 图像切片之间的深度特征。

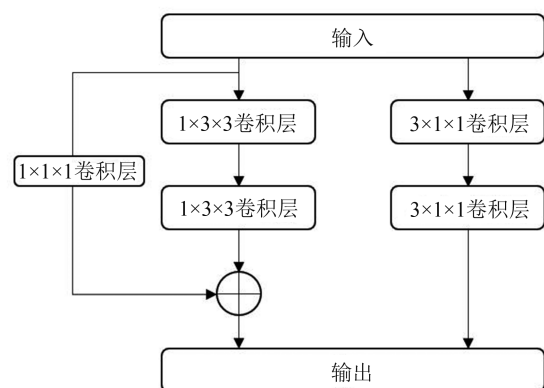


图 5 残差基本结构

本文以 ResNeXt 作为基础网络进行改进:首先,将 Conv1 层改造成 3 个卷积核为 3×3×3 的串联卷积层;其次,将原 ResNeXt 网络基础模块中 2 个 3×3×3 卷积核改造成对图像进行两个分支的卷积运算,对左边部分进行 2 次 1×3×3 卷积操作,对右半部分进行 2 次 3×1×1 卷积操作,进一步降低网络参数数量,提高训练速度。表 2 为 3D-ResNeXt 网络结构参数。

表 2 3D-ResNeXt 网络结构参数

网络结构	网络构成	通道数	重复次数
Conv1	Conv, 3×3×3, stride(2,2,2)	64	1
	Conv, 3×3×3, stride(1,2,2)	64	1
	Conv, 3×3×3, stride(1,2,2)	64	1
	MaxPool, 3×3×3, stride(2,2,2)	64	1
Conv2.x	左 1×3×3	256	3
	右 3×1×1		
Conv3.x	左 1×3×3	512	4
	右 3×1×1		
Conv4.x	左 1×3×3	1024	6
	右 3×1×1		
Conv5.x	左 1×3×3	2048	3
	右 3×1×1		

(4) 相关性分析

相关分析是一种研究两个或两个以上随机变量之间相关性的统计分析方法。相关系数是研究变量之间线性相关程度的量度。两个变量之间的皮尔逊系数定义为两个变量协方差和标准差之间的商。MRI 和 PET 图像分别经过 3D-ResNeXt 网络的输出的神经影像诊断结果分为三类:0 代表 CN,1 代表 MCI,2 代表 AD。

卷积神经网络的输出可以被理解为神经成像诊断的结果。经过深度学习训练后,卷积神经网络将图像(MRI 或 PET)分为三类:0 代表认知 CN,1 代表 MCI,2 代表 AD。如果通过 MRI 和 PET 图像的诊断结果一致,则神经影像学诊断结果可靠。如果 MRI 和 PET 图像的诊断结果不一致,则说明 3D-ResNeXt 网络模型的诊断结果不可靠。本文基于皮尔逊相关系数计算 MRI 和 PET 图像预测结果之间的相关性。相关性分析的目的在于确定两种神经影像学诊断的结果是否一致,其计算公式如下:

$$\gamma = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4)$$

其中,  $\bar{x}$  和  $\bar{y}$  表示概率的平均值,  $n$  表示  $x$  和  $y$  的类别数,  $\gamma$  的值在 -1 和 1 之间。通过  $\gamma$  来判断两个 3D-ResNeXt 网络预测结果的一致性。如果  $\gamma$  的值接近 1,则表示 3D-ResNeXt 网络对 MRI 和 PET 图像预测结果是一致的,多模态深度学习模型的诊断结果是可靠的,如果  $\gamma$  的值接近 -1,则多模态深度学习模型

的诊断结果是不可靠的。

4. 实验环境与参数配置

本实验软件环境包括:Python 3.6, TensorFlow 2.8 开源框架和 sklearn 包,操作系统:Windows 11;硬件环境为: Intel i7 2.8GHz 6 核处理器、16GB 内存、GPU 为 NVIDIA GeForce DTX 2060 显存 6GB。图像预处理使用 Matlab R2019a 中的 CAT 12 和 SPM12 工具、dartel 工具箱。使用 TensorFlow2.8 中的 tf.keras 搭建 3D-ResNeXt 网络模型。输入图像为 224×224× $n$  的三维图像(其中  $n$  为切片数量,取值为 8、32、80),超参数 epochs 为 50、训练的批次 batch\_size 为 20、初始学习率为 0.0001,并使用 cosine decay 的衰减方式进行学习率调整、dropout 丢掉率为 0.5、参数优化算法为 Adam、采用 5 折交叉验证。

5. 评价指标

本文使用 4 个评价指标对多分类训练结果进行评估,包括:准确率 (accuracy, ACC)、加权精确度 (weighted-precision, WP)、加权召回率 (weighted-recall, WR)、皮尔逊系数 ( $\gamma$ )。

精确度 (precision) 表示预测为正确的样本中真正正确的比例,公式如式(5)所示。

$$P = \frac{TP}{TP+FP} \quad (5)$$

召回率 (recall) 表示预测为正确样本中真正正例占有所有正例的比例,公式如式(6)所示。

$$R = \frac{TP}{TP+FN} \quad (6)$$

评价多分类模型 WP 的方法,公式如式(7)所示。

$$WP = \frac{1}{n} \sum_{i=1}^n P_i W_i \quad (7)$$

评价多分类模型 WR 的方法,公式如式(8)所示。

$$WR = \frac{1}{n} \sum_{i=1}^n R_i W_i \quad (8)$$

评价多分类模型 ACC 的方法,公式如式(9)所示。

$$ACC = \frac{\sum_{i=1}^n TP_i}{\sum_{i=1}^n TP_i + \sum_{i=1}^n FP_i} \quad (9)$$

其中,TP 表示真阳性 (true positive)、FP 表示假阳性 (false positive)、TN 表示真阴性 (true negative)、FN 表示假阴性 (false negative)、 $W_i$  表示第  $i$  类测试集总数与测试集总数的比值、 $n$  为类别个数、 $P_i$  为第  $i$  个类别计算的精确度、 $R_i$  为第  $i$  个类别计算的召回率。

结果与讨论

1. 实验结果与分析

(1) 3D-ResNeXt 网络的性能验证

为了验证本文 3D-ResNeXt 网络模型的有效性,采用多组对比的消融实验。实验策略如下:①仅将传统 ResNeXt 网络修改为三维结构,不融合任何策略,记为 3D-ResNeXt<sup>(1)</sup>;②仅将传统 ResNeXt 网络中 Conv1 的 7×7×3 的卷积层修改为 3 个 3×3×3 串联的卷积层,记为 3D-ResNeXt<sup>(2)</sup>;③仅将传统 ResNeXt 网络瓶颈层中的卷积核使用 3DCK 拆分结构,记为 3D-ResNeXt<sup>(3)</sup>;④融合以上所有策略,构成本文使用的方法。输入图像大小为 224×224×32(有 32 个切片组成的三维图像),使用以上方法结果如表 3 所示。

表 3 不同改进网络方法的结果

网络模型	ACC (%)	WP (%)	WR (%)	γ (%)
3D-ResNeXt <sup>(1)</sup>	77.03	77.78	77.04	78.3
3D-ResNeXt <sup>(2)</sup>	80.37	80.76	80.37	77.2
3D-ResNeXt <sup>(3)</sup>	85.56	86.26	85.56	84.7
本文方法	87.78	88.32	87.87	86.4

通过表 3 可以看出对传统 ResNeXt 网络不同改进方法的结果。只使用单一策略都可以提升网络模型的性能,如果使用全部策略构建 3D-ResNeXt 网络,ACC、WP、WR、γ 的结果都好于 3D-ResNeXt<sup>(1)</sup>、3D-ResNeXt<sup>(2)</sup>、3D-ResNeXt<sup>(3)</sup>。实验策略②相对于实验策略①网络提升幅度最小,ACC 只提高了 3.34%;实验策略③相对于实验策略①网络参数量得到了降低,ACC 提高了 8.53%;本文方法使用了全部策略,综合性能也是最优的。各方法在测试集的混淆矩阵如图 6 所示。

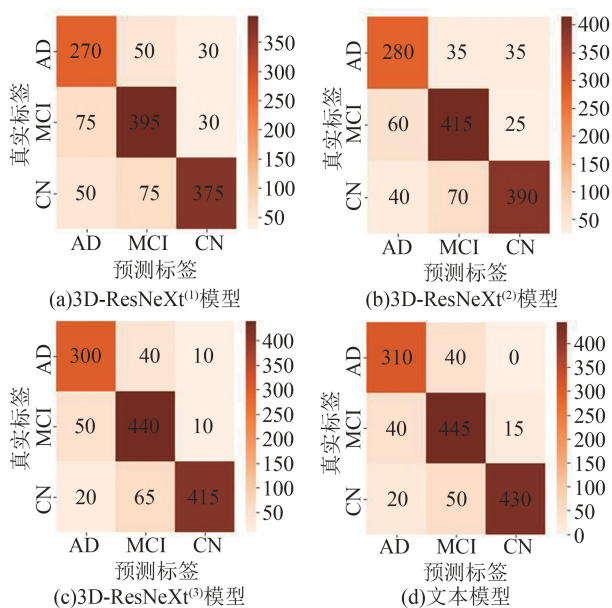


图 6 不同方法在测试集的混淆矩阵

(2) 不同图像切片数的 3D-ResNeXt 网络性能对比

在 ADNI 数据库中,不同受试者的 MRI 和 PET 图像的切片数有可能不相同,为了在训练中保证每一个

输入的图像大小相同,分别将 224×224×8、224×224×32、224×224×80 大小的图像输入融合所有策略的 3D-ResNeXt 网络进行训练,对比不同切片数图像对网络性能的影响。实验结果如表 4 所示。

表 4 网络模型在不同切片数下的性能 (%)

图像切片	ACC	WP	WR	γ
8	84.23	81.32	84.22	83.3
32	87.78	88.32	87.87	86.4
80	88.21	84.70	88.23	85.7

对于不同图像的切片数,模型训练准确率变化,如图 7 所示。

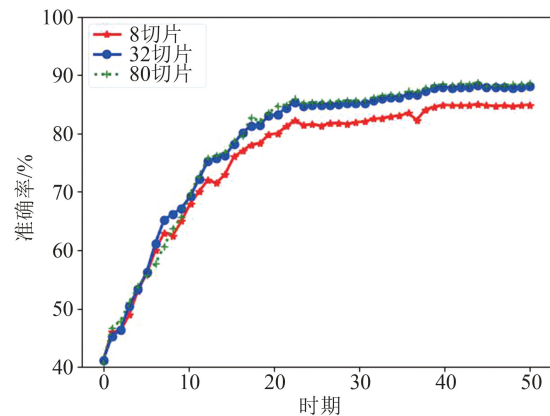


图 7 不同切片数模型准确率变化

通过表 4 可以看出,随着输入图像切片数的增大,模型在 ACC、WR 方面都有所提升,但提升幅度有限,但在 WP 和 γ 方面性能有所下降,使用 224×224×80 大小的图像进行训练时间反而是使用 224×224×32 大小的图像训练的 2.3 倍。通过图 6 可以看出,在模型训练的前期,不同图像切片数输入的数据在准确率上升较快,在经历了 25 个迭代周期后,准确率趋于稳定,输入切片数为 8 的图像模型准确率低于输入切片数为 32 和 80 的图像模型大约 3.55%,对于输入切片数为 32 和 80 的图像模型的准确率变化没有太大区别,图像中准确率曲线基本重合。综上所述,输入切片为 32 的图像训练模型不论在训练时间和模型准确率方面性能最优。

(3) 多模态和单模态性能比较

本文分别选取输入图像大小为固定的 224×224×32 的单模态 MRI 图像、单模态 PET 图像、多模态 MRI 和 PET 图像在 3D-ResNeXt 网络进行分类实验,对比多模态和单模态图像训练的 ACC、WP、WR。实验结果如表 5 所示。

表 5 不同模态性能对比 (%)

模态	图像切片	ACC	WP	WR
单模态 MRI 图像	32	83.11	82.32	83.21
单模态 PET 图像	32	85.51	84.31	85.52
多模态 MRI+PET	32	87.78	88.32	87.87

通过表 5 可以看出,多模态 MRI 和 PET 图像在本

文网络进行分类实验的性能优于任意单模态图像。使用单模态 PET 图像进行分类实验的性能优于使用单模态 MRI 图像。图 8 绘制了单模态 MRI 图像、单模态 PET 图像、多模态 MRI 和 PET 图像分类的受试者工作特征(receiver operating characteristic, ROC)曲线。ROC 曲线下的面积反映了诊断的准确性。从 ROC 曲线可以看出,在使用 3D-ResNeXt 网络对图像进行分类的前提下,多模态神经影像诊断优于单模态神经影像诊断。

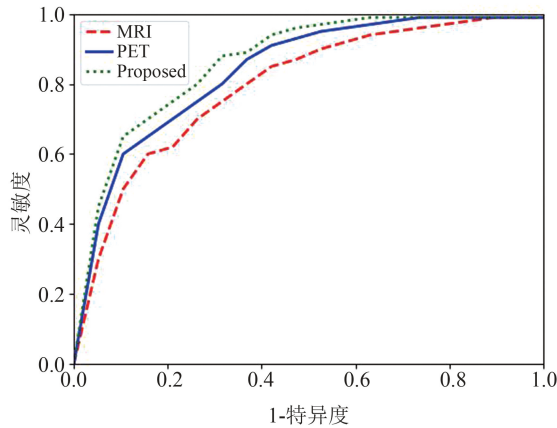


图 8 单模态 MRI 图像、单模态 PET 图像、多模态 MRI 和 PET 图像分类的 ROC 曲线

(4)不同模型性能比较

本文选取输入图像大小为固定的 224×224×32 的 MRI 和 PET 图像在不同网络模型进行训练,对比不同模型分类的 ACC、WP、WR、 $\gamma$  系数。实验结果如表 6 所示。

表 6 不同模型性能对比(%)

模型	切片数	ACC	WP	WR	$\gamma$
VGG19	32	63.81	60.91	63.88	59.3
ResNet50	32	80.32	75.32	80.41	79.2
Inception-v3	32	82.61	78.13	82.58	82.6
C3D	32	85.21	82.42	85.19	84.9
DenseNet121	32	84.71	80.17	87.68	85.2
本文模型	32	87.78	88.32	87.87	86.4

通过表 6 可以看出,由于二维的 VGG19 和 ResNet50 模型无法处理连续切片之间的特征信息,所以在 ACC、WP、WR、 $\gamma$  系数的性能最低。Inception-v3 网络主要使用了分解卷积的思想,将大卷积分解为多个小卷积,在训练参数量降低的同时也提高了网络的性能。C3D 是一种使用 3D 卷积运算从 3D 图像提取空间和时间特征的分类模型,使用此模型不仅可以学习到 3D 图像的空间信息,还可以学习到连续切片之间的特性信息,所以在模型性能方面对于 VGG19、ResNet50、Inception-v3 有了较大的提升。对于密集连接卷积网络 DenseNet121 和传统网络 VGG19 和 ResNet50 相比,减少了梯度消失和网络参数数量,网络深

度大大深于 VGG19 和 ResNet50,所以在模型性能方面优于 VGG19 和 ResNet50,但此网络同样无法处理连续切片之间的特征信息,模型性能弱于 C3D。本文设计的模型使用 3D 卷积核同时考虑到获取空间和时间的特征,还使用了大卷积分解为多个小卷积串联的方法,减少了模型的训练参数,所以本文模型综合性能相对于其他模型性能最优。

结 论

本文提出了一种用于 AD 早期辅助诊断的深度学习模型,本文设计了多模态 3D-ResNeXt 网络辅助诊断模型用于训练 MRI 和 PET 图像。利用这两个独立的 3D-ResNeXt 网络,通过卷积、池化和 softmax 等一系列计算,提取 MRI 图像特征和 PET 图像特征。通过相关分析来判断两个独立的 3D-ResNeXt 网络输出的一致性。如果两种网络模型的结果相似,则可以直观地看出对同一患者的诊断与不同的诊断方式相一致。在此基础上,提出了一种新的相关性计算方法。本文计算了 PET 图像的诊断结果与 MRI 图像的诊断结果之间的皮尔逊系数,通过分析得到多模态神经影像辅助诊断结果。对数据集图像通过图像预处理和数据增强扩展数据集。在 ADNI 开放数据库上进行的大量实验表明,本文方法的诊断效果在许多指标上都优于其他辅助诊断模型。实验结果表明,所提出的多模态辅助诊断模型能够达到良好的诊断率。下一步,将对分类任务细化,研究 pMCI 和 sMCI 的分类任务。

参 考 文 献

[ 1 ] Collaborators GDF. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the global burden of disease study 2019 [ J ]. Lancet Public Health, 2022, 7 ( 2 ): e105-e125.

[ 2 ] Thies W, Bleiler L. 2020 Alzheimer's disease facts and figures [ J ]. Alzheimer's Dementia, 2020, 16( 3 ):391-460.

[ 3 ] Alzheimer's Association. 2018 Alzheimer's disease facts and figures [ J ]. Alzheimers Dement, 2018, 14( 3 ): 367-429.

[ 4 ] Oh K, Chung YC, Kim KW, et al. Classification and visualization of Alzheimer's disease using volumetric convolutional neural network and transfer learning [ J ]. Scientific Reports, 2019, 9( 1 ): 1-17.

[ 5 ] Xia Z, Yue G, Xu Y, et al. A novel end-to-end hybrid network for Alzheimer's disease detection using 3D CNN and 3D CLSTM [ C ]. 2020 IEEE 17th International Symposium on Biomedical Imaging, 2020: 1-4.

[ 6 ] Bailey DL, Pichler BJ, Gckel B, et al. Combined PET/MRI: multi-modality multi-parametric imaging [ J ]. Mol Imaging Biol, 2015, 17 ( 5 ): 1-14.

[ 7 ] Riederer I, Bohn KP, Preibisch C, et al. Alzheimer disease and mild cognitive impairment: integrated pulsed arterial spin-labeling MRI and 18F-FDG PET [ J ]. Radiology, 2018, 288( 1 ): 198-206.

( 下转第 892 页 )