

基于残差混合注意力机制的脑部 CT 图像 分类卷积神经网络模型

乔思波¹, 庞善臣¹, 王 敏², 翟 雪¹, 于世行³, 丁 桐⁴

(1. 中国石油大学(华东)计算机科学与技术学院, 山东青岛 266580; 2. 中国石油大学(华东)控制科学与工程学院, 山东青岛 266580;
3. 天津工业大学机械工程学院, 天津 300387; 4. 山东大学软件学院, 山东济南 250101)

摘 要: 针对阿尔兹海默症、病变(如脑肿瘤)和健康老化的3类脑部CT图像分类问题,本文提出了一种改进的ResNet-10卷积神经网络模型.该模型在网络的残差映射结构中加入残差混合注意力模块,解决了原模型提取的特征分辨性弱的问题,精确捕捉了脑部组织在CT图像中的位置和内容信息;此外,本文设计了全局平均池化层,简化了模型的复杂度,并在其后引入Dropout机制,缓解了过拟合.在训练阶段,该模型建立了标签平滑交叉熵损失函数,使模型在样本数量有限的情况下仍有较强的泛化能力.系列实验证明了改进后的ResNet-10网络模型在分类脑部CT图像时达到97.47%的分类精度.

关键词: 残差混合注意力模块; 标签平滑; 脑部CT; 卷积神经网络

中图分类号: TP399 **文献标识码:** A **文章编号:** 0372-2112(2021)05-0984-08

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.12263/DZXB.20200881

A Convolutional Neural Network for Brain CT Image Classification Based on Residual Hybrid Attention Mechanism

QIAO Si-bo¹, PANG Shan-chen¹, WANG Min², ZHAI Xue¹, YU Shi-hang³, DING Tong⁴

(1. College of Computer Science and Technology, China University of Petroleum, Qingdao, Shandong 266580, China;

2. College of Control Science and Engineering, China University of Petroleum, Qingdao, Shandong 266580, China;

3. College of Mechanical Engineering, Tiangong University, Tianjin 300387, China;

4. College of Software, Shandong University, Jinan, Shandong 250101, China)

Abstract: To classify three types of brain CT (computerized tomography) images in Alzheimer's disease, lesion (e.g., brain tumor) and healthy aging, an improved ResNet-10 convolutional neural network is proposed in this paper. A residual hybrid attention module is embedded in the residual identity mapping to capture the location and content information of brain tissue in brain CT images, solving the original model to extract weak distinguish features problems. In addition, to simplify the improved model and alleviate the overfitting, several techniques such as global average pooling and Dropout are used in the model. Moreover, to have strong generalization ability in the case of limited sample quantity, tag smoothing cross-entropy loss function is adopted to train the model. Experimental results show that the improved ResNet-10 achieves 97.47% accuracy in classifying brain CT images.

Key words: residual hybrid attention module; label smoothing; brain CT; convolutional neural network

1 引言

近些年,医学影像技术得到了快速发展,已成为研究各项疾病不可或缺的重要工具.电子计算机断层扫描(Computerized Tomography, CT)因其无创、廉价且方

便等特点被广泛应用于各项医学影像中,帮助医生早期评估诊断癌症^[1]、肿瘤^[2]以及骨头坏死^[3]等症状.同时,CT也是诊断脑部疾病的首选影像工具,能早期发现脑部潜在的病症,如阿尔兹海默症(Alzheimer's Disease, AD),使患者及时得到治疗.近年来,AD和病变

(脑肿瘤)发病率呈显著上升趋势,临床中确诊其病症是一个漫长且复杂的过程.为了帮助医生早期发现患者大脑出现异常症状,减少人工诊断的差异性,Gao 等人^[4]针对本文使用的脑部 CT 图像数据首次开展了对 AD、病变(如脑肿瘤)和健康老化的 3 类脑部 CT 图像分类的研究.因脑部 CT 影像含有大量噪声,与肝胆等部位的影像相比分辨率较低,不易提取相关特征,导致分类准确率较低.因此,提取脑部 CT 图像中分辨性强的特征对早期筛查脑部疾病至关重要.

深度学习作为机器学习的一个重要分支,由于其强大的特征学习能力被广泛应用于医学图像分析的各个领域,比如乳腺癌、肿瘤筛查^[5,6]、肺结节分类^[7]以及先心病的诊断^[8]等.当前,深度学习在脑部 CT、MRI (Magnetic Resonance Imaging)医学图像方面的应用也逐渐流行.Pereira 等人^[9]提出了自动分割脑部肿瘤的卷积神经网络模型,该模型采用小卷积核来缓解因增加模型深度而产生的过拟合问题.Derikvand 等人^[10]通过组合不同的卷积神经网络提取脑部肿瘤的局部和全局特征,提升肿瘤的分割效果.Kumar 等人^[11]利用集成方法对医学图像进行分类.通过集成不同的卷积神经网络模型学习脑部 CT 等医疗图像不同层次的特征,得到各种图像分辨性较强的特征信息.类似地,Farooq 等人^[12]设计了 4 个不同的卷积神经网络模型对 AD、轻度认知功能障碍、重度认知功能障碍以及健康的脑部 MRI 图像进行分类,并通过管道的方式集成分类结果,达到了较好的分类效果.可见,深度学习在脑部医疗图像方面的研究越来越成功.

因此,本文采用卷积神经网络构建分类模型对 AD、病变(如脑肿瘤)以及健康老化的 3 类脑部 CT 影像数据进行分类.针对本文所使用的脑部 CT 图像数据集,Gao 等人^[4]首次采用深度学习的方法设计了一个简单的卷积神经网络模型,该模型对三种类别的分类准确率达到 86.32%.为了提高在该数据集的分类精度,Zahedinasab 等人^[13]提出了一个 7 层卷积神经网络模型,通过结合不同的激活函数,使得模型分类精度达到了 88.67%.显然,在该数据集上的分类精度远没有达到令人满意的效果.

实际上,卷积神经网络是深度学习的典型代表之一.通过深度卷积神经网络学习医疗图像深层次的特征,在表达图像时具有更好的识别性和鲁棒性.然而,现有的大部分深度学习方法均从整张 CT 图像中学习特征,这不仅包括有关脑部病变的特征,还包括背景区域等混杂特征.背景区域包含的其他组织的特征极可能混入提取的脑部特征,使得提取到的特征分辨性不强,模型分类不准确.

视觉注意力机制在图像分类^[14-16]和目标检测^[17]

等视觉任务中展示出了优越性.该机制符合人类的视觉系统,不会试图一次处理整个视觉场景信息,而是在需要的时候聚焦整个视觉空间中突出的部分.视觉注意力模块通过引导模型向具有特定信息的图像区域学习,帮助模型提取图像中最像大脑部位的动态特征.Wang 等人^[14]提出了一个残差注意力模块,该模块采用下采样和上采样的方式提取注意力特征映射.Hu 等人^[15]重点关注特征映射每个通道之间的联系,提出了一个通道注意力模块.将该模块简单地堆叠成卷积神经网络模型大幅提升分类效果.在卷积网络提取特征时,不仅要关注特征映射每个通道之间的相关性,其空间关系也很重要.因此,Woo 等人^[16]融合了通道信息和空间信息,提出了一个混合注意力模块(Convolutional Block Attention Module, CBAM),并在 ImageNet 数据集进行分类测试,验证了通道信息和空间信息结合的有效性.但该注意力模块主要采用池化操作处理特征映射,容易较多地损失特征映射的关键信息.

针对 AD、病变(如脑肿瘤)以及健康老化的 3 类脑部 CT 图像分类问题,本文的主要贡献如下:

(1)提出了一个残差混合注意力模块(Residual Hybrid Attention Module, RHAM).与 Woo 等人^[16]的方法不同,残差混合注意力模块主要采用卷积操作提取特征映射的相关信息,并加入残差跳连结构,在捕捉脑部组织的位置和内容信息时降低关键特征损失程度.该注意力模块具有良好的移植性,将其融入到网络模型中,可使模型提取更具分辨性的脑部特征,提高模型分类的准确率.

(2)改进了 ResNet-10^[18]残差网络模型.首先,使用全局平均池化层代替平均池化层,不再限制输入图像的尺寸.在全局平均池化层之后加入 Dropout^[19],进一步缓解了过拟合,得到了 MResNet-10;其次,在 MResNet-10 的残差映射结构中加入 RHAM 模块,并将改进后的模型命名为 RHAM-MResNet-10,该模型在分类脑部 CT 图像时大幅提升了分类精度.

2 基于残差混合注意力机制的脑部 CT 图像分类模型

针对 AD、病变以及健康老化的 3 类脑部 CT 图像分类问题,本文提出了一个基于残差混合注意力机制的卷积神经网络模型 RHAM-MResNet-10.本节将对该模型的设计进行细致阐述.

2.1 RHAM-MResNet-10

当前,深度卷积神经网络已经在提取图像特征上表现出了优越性.然而,随着卷积神经网络模型不断加深,极易出现梯度消失或爆炸,导致模型性能下降.残差网络模型因其特殊的跳连结构能够有效缓解梯度消失或爆炸,被广泛应用于各类任务.深度卷积神经网络拥

有较多的参数,在数据样本充足的情况下(如 ImageNet)能够达到极好的分类效果.但本文所使用的脑部 CT 图像样本数量较少,每个类分别约 1100 张.直接使用较深的网络模型极易产生过拟合,导致模型性能下降.因

此,本文提出了 RHAM-MResNet-10 网络模型,如图 1 所示,将残差混合注意力模块 RHAM 嵌入 MResNet-10 残差网络模型中,充分捕获脑部 CT 图像信息,提高模型分类的准确率.

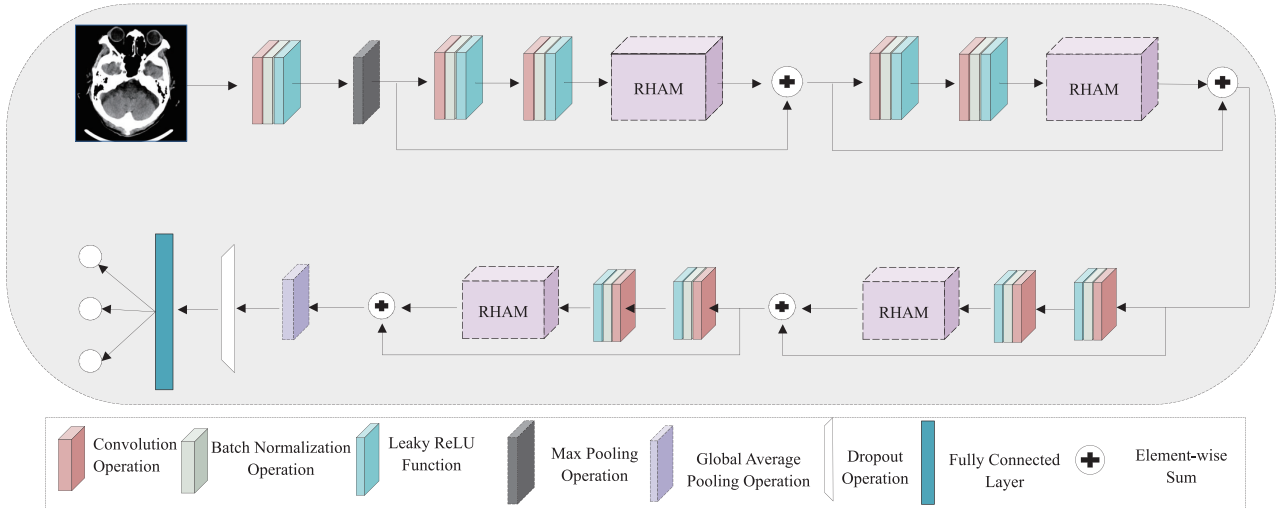


图1 RHAM-MResNet-10网络架构

从图 1 可以看出, RHAM-MResNet-10 网络模型的输入是一张脑部 CT 图像,输出是该 CT 图像属于 AD、病变(如脑肿瘤)以及健康老化 3 类的类别概率. RHAM-MResNet-10 首先采用 1 次卷积对输入的图像进行处理,过滤器的尺寸为 7×7 ,设置步长为 2,将输出的特征映射降维至原来的 $1/2$.在每次卷积后,均采用 Batch Normalization^[20]以进一步缓解过拟合;Leaky ReLU 激活函数充分利用梯度信息,保证模型不断收敛.之后,采用过滤器大小为 3×3 、步长为 2 的最大池化操作减少特征映射的参数,并将其降维至原始图像的 $1/4$.然后,采用了 4 个残差注意力卷积组提取特征映射的高级特征,每个残差注意力卷积组包含 2 次卷积和 1 次残差混合注意力操作.每组卷积操作的过滤器大小均为 3×3 ,过滤器的数量分别为 64、128、256 和 512.最后,利用全局平均池化层提取最后一次卷积输出特征映射的空间信息,降低了模型的参数数量.如同 Li 等人^[21],在其后加入了 Dropout 防止过拟合.经过全连接层,使用 Softmax 激活函数输出分类的结果概率.

2.2 残差混合注意力模块

在卷积神经网络中加入注意力机制能够让网络模型更加关注感兴趣区域,使得模型聚焦所要提取的重要特征,忽略不相关特征.本文提出的残差混合注意力模块包括残差通道注意力模块和残差空间注意力模块,同时用于关注特征映射的空间域和通道域信息,如图 2 所示.

从图 2 可以看出,残差混合注意力模块是把残差通道注意力模块与残差空间注意力模块按顺序堆叠.给定一个输入特征映射,残差混合注意力模块不仅能

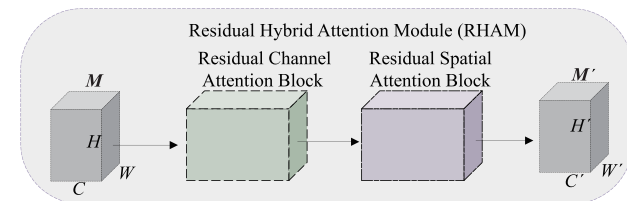


图2 残差混合注意力模块

够学习特征映射中特定目标有意义的信息,还能探索特定目标在特征映射中的位置,忽略其他不重要的目标特征.实际上,残差混合注意力机制并没有增加模型的深度,而是扩展了整个模型架构的宽度,在提升模型性能方面较大的帮助.接下来,分别对残差通道注意力模块和残差空间注意力模块进行详细地介绍.

残差通道注意力模块 (Residual Channel Attention Block, RCA) 为了研究特征映射通道之间的关系,本文提出了一个残差通道注意力模块,如图 3 所示.

残差通道注意力模块主要关注特征矩阵中有价值的信息,提取特定目标的分辨性较强的特征.在普通的卷积操作中,受到感受野范围的限制,卷积核每次只能关注特征映射的某一局部区域,不能捕获该局部区域以外的上下文信息.为了更加有效地探索特征映射整个空间的语义,在残差通道注意力模块中,本文首先采用全局平均池化操作,将输入的特征映射 $M \in \mathbb{R}^{H \times W \times C}$ 的整个空间信息融合为一个通道标识符 $I_c \in \mathbb{R}^{1 \times 1 \times C}$.从实现的角度,一个通道标识符 $I_c \in \mathbb{R}^{1 \times 1 \times C}$ 是特征映射从 2 维空间缩减到 1 维空间得到的.为了能够提取通道标识符的有效信息,本文使用了 2 个卷积操作捕获通道

标识符中每个通道之间的非线性关系. 为了控制两个卷积操作的复杂度, 将第一个卷积层设置为降维层, 正如 Hu^[15] 和 Woo^[16] 等人, 降维比率设置为 16 来降低通道标识符的通道维度. 第二个卷积层是用来把维度升到输入特征映射的维度; 然后, 利用 sigmoid 激活函数对两次卷积处理之后的通道标识符执行简单的门控控制, 得到一个逻辑化的通道信息结构体 U'_c . 之后, 逻辑化的空间结构体与输入的特征映射进行一次加权和操作得到通道注意力特征矩阵 $U_c \in \mathbb{R}^{H \times W \times C}$:

$$U'_c = \sigma(W_2 L(W_1 I)) \quad (1)$$

$$U_c = U'_c \times M \quad (2)$$

其中, $L(\cdot)$ 表示 Leaky ReLU 激活函数, $W_1 \in \mathbb{R}^{\frac{c}{r} \times c}$ 和 $W_2 \in \mathbb{R}^{c \times \frac{c}{r}}$ 分别表示第一次卷积和第二次卷积的权重, $\sigma(\cdot)$ 表示 sigmoid 激活函数.

最后, 为了减少输入特征映射信息的损失, 本文将输入的特征映射与得到的通道注意力特征映射通过残差结构进行信息融合, 得到最终的通道注意力特征映射 $\tilde{U}_c \in \mathbb{R}^{H \times W \times C}$:

$$\tilde{U}_c = U_c + M \quad (3)$$

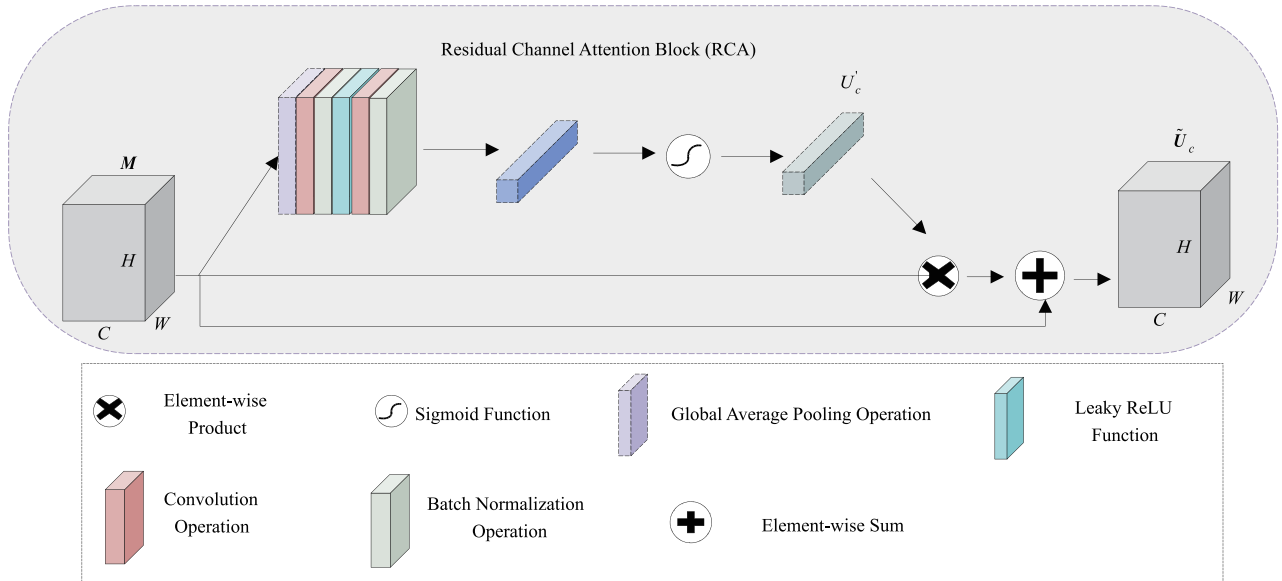


图3 残差通道注意力模块

残差空间注意力模块 (Residual Spatial Attention Block, RSA) 为了探索特征映射局部区域之间的空间关系, 本文提出了一个残差空间注意力模块, 如图 4 所示.

空间注意力模块的功能与通道注意力模块不同, 它主要关注的是目标物体在输入特征映射的位置信息. Woo 等人^[16] 在通道维度上采用了平均池化和最大池化操作得到了两种不同的特征映射. 本文认为直接采用池化的操作容易造成输入特征映射信息的大量丢失. 从图 4 可以看出, 首先采用了一层卷积核尺寸为 7×7 的卷积神经网络对输入特征映射进行初步处理. 通过融合输入特征映射通道维度的信息, 得到了一个空间标识符 $I_s \in \mathbb{R}^{H \times W \times 1}$. 然后, 利用 sigmoid 激活函数对空间标识符进行简单的门控控制, 得到一个逻辑化的空间信息结构体 U'_s . 之后, 逻辑化的空间信息结构体与输入的特征映射进行一次加权和操作得到空间注意力特征矩阵 $U_s \in \mathbb{R}^{H \times W \times C}$:

$$U'_s(M) = \sigma(\text{Conv}^{7 \times 7}(M)) = \sigma(I_s) \quad (4)$$

$$U_s = U'_s \times M \quad (5)$$

其中, $\sigma(\cdot)$ 代表 sigmoid 激活函数, $\text{Conv}^{7 \times 7}(\cdot)$ 表示卷

积核尺寸大小为 7×7 的卷积操作. 最后, 与残差通道注意力模块类似, 本文将输入的特征映射与得到的空间注意力特征映射通过残差结构进行信息融合, 得到最终的空间注意力特征映射 $\tilde{U}_s \in \mathbb{R}^{H \times W \times C}$:

$$\tilde{U}_s = U_s + M \quad (6)$$

2.3 损失函数

为了使模型在脑部 CT 数据有限的情况下仍有较好的泛化能力, 在训练模型时, 先对分类标签平滑处理, 再使用交叉熵计算模型的损失, 使模型不仅关注正确分类标签的损失值, 还能考虑错误分类标签的损失. 标签平滑后的交叉熵损失函数如下^[22]:

$$y'_k = (1 - v) \times y_k + v \times u(k) \quad (7)$$

$$\begin{aligned} H(y', y) &= - \sum_{k=1}^K y'_k \log(\hat{y}_k) \\ &= (1 - v)H(y, \hat{y}) + vH(u, \hat{y}) \end{aligned} \quad (8)$$

其中, k 表示第 k 个类别; y_k 表示第 k 个标签的真实值; v 表示平滑因子, 本文取值为 0.1; $u(k)$ 表示均匀分布, 本文取值为 $1/3$; y'_k 表示第 k 个平滑后的标签值; \hat{y} 表示预测类别概率.

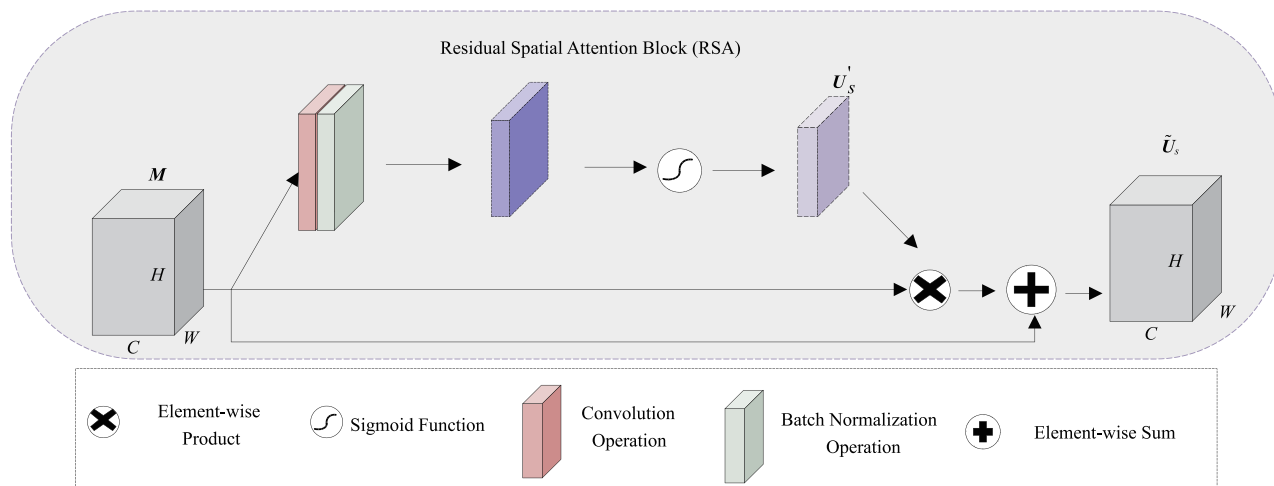


图4 残差空间注意力模块

3 实验

3.1 数据集与评估方法

本文使用的3D脑部CT数据集来自文献[4],该数据集共包含285个3D CT图像.其中,AD类的有57个,病变类的有113个,健康老化类的有115个,每个3D CT图像包含20张不同的2D大脑CT切片,本文使用2D大脑CT切片训练模型,部分样本图片如图5所示.

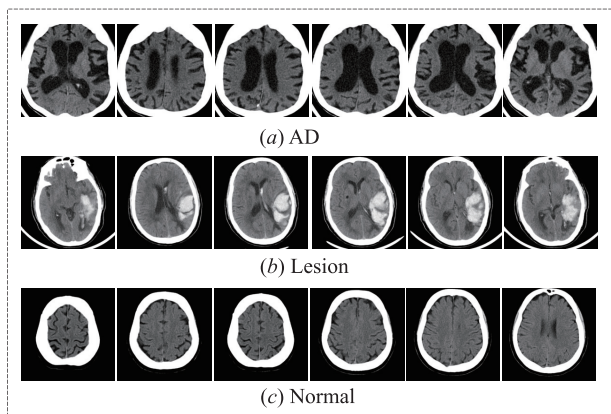


图5 部分脑部CT图像

对于AD和健康老化类,所有的2D大脑CT切片均用来训练模型,对于病变类的数据,只有包含病变信息(如脑肿瘤)的2D切片才被用来训练模型.

为了验证模型的性能,本文仍然采用了与文献[4,13]相同的评估方法:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

其中,TP指将正类预测为正类的数量;TN指将负类预测为负类的数量;FP指将负类预测为正类的数量;FN指将正类预测为负类的数量.

3.2 模型参数设置

在脑部CT图像分类的实验中,训练模型所需超参数设置如下:采用小批量样本的方式训练模型,小批量样本大小设置为32;采用Adam^[23]算法优化损失函数,初始的学习率设置为0.0001,同时加入了L2正则化对权重参数进行惩罚约束,惩罚因子设置为0.0001;训练过程中,采用指数衰减策略调整学习率,并将gamma设置为0.98;其他参数详细设置见表1.除此之外,采用随机水平翻转、随机垂直翻转以及调整亮度、对比度和饱和度的方法对训练数据进行增强,提高模型的泛化能力.本文所有的模型均是使用Pytorch深度学习框架实现,并在NVIDIA 2080Ti GPU训练得到的.

表1 RHAM-MResNet-10 参数设置

Layer name	Output size	Parameters
Conv1	100 × 100	7 × 7, 64, stride 2
Max Pool	50 × 50	3 × 3, stride 2
Residual Unit 1	50 × 50	$\begin{pmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{pmatrix}$
Attention Module	50 × 50	RHAM
Residual Unit 2	25 × 25	$\begin{pmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{pmatrix}$
Attention Module	25 × 25	RHAM
Residual Unit 3	13 × 13	$\begin{pmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{pmatrix}$
Attention Module	13 × 13	RHAM
Residual Unit 4	7 × 7	$\begin{pmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{pmatrix}$
Attention Module	7 × 7	RHAM
Global Average Pool		1 × 1
Dropout		0.5
FC, Softmax		3

3.3 实验结果

目前,主流的残差卷积神经网络模型有 ResNet-10、ResNet-18、ResNet-34 以及 ResNet-50,模型的深度依次增加.为了选择匹配本文脑部 CT 数据复杂度的模型,分别采用上述残差网络模型对脑部 CT 图像数据集进行训练,得到的分类结果如表 2 所示.

表 2 残差网络模型分类结果对比

Model	Accuracy (%)
ResNet-50	86.32 ± 1.3
ResNet-34	87.25 ± 1.2
ResNet-18	89.21 ± 0.6
ResNet-10	90.31 ± 0.4

表 2 表明,ResNet-10 在脑部 CT 数据集上的分类效果最好,说明该模型与数据集的复杂度基本匹配.因此,选择 ResNet-10 进行改进得到 MResNet-10,作为本文的骨干网络模型.

分别在上述残差网络模型中加入抑制过拟合的方法,得到 MResNet-10、MResNet-18、MResNet-34 以及 MResNet-50 等骨干网络模型.将这些骨干网络模型分别对脑部 CT 图像进行分类,分类结果如表 3 所示.

表 3 骨干网络模型分类结果对比

Model	Accuracy (%)
MResNet-50	88.64 ± 1.1
MResNet-34	89.86 ± 1.2
MResNet-18	90.02 ± 1.0
MResNet-10	92.41 ± 0.9

表 3 表明,深度卷积神经网络虽然在大规模数据集上能够表现出好的分类效果.但面对小规模数据集时易产生过拟合现象,导致分类性能变差.因此,设计一个与数据集复杂度相匹配的网络模型尤为重要.

将本文提出的 RCA、RSA、RHAM 以及 Woo 等人^[16]提出的 CBAM 分别加入到上述骨干网络模型,在脑部 CT 数据集上进行分类,结果见表 4. 训练模型时,验证集的损失变化曲线以及准确率变化曲线见图 6 和图 7.

通过对比表 4 中加入注意力模块后的骨干模型分类准确度,可以得到以下 2 点结论:

(1) 与单一通道域注意力模块或空间域注意力模块相比,混合域注意力模块更加有效地提升模型分类效果.

(2) 与 Woo 等人^[16]提出的 CBAM 混合域注意力模块相比,本文提出的 RHAM 混合域注意力模块采用卷积操作代替了池化操作,引入了残差跳连结构,在提取注意力特征映射时有效避免了关键信息的丢失,保证了模型的稳定性.与 CBAM 相比,RHAM 在帮助骨干模型提升脑部 CT 分类效果时更加有效.

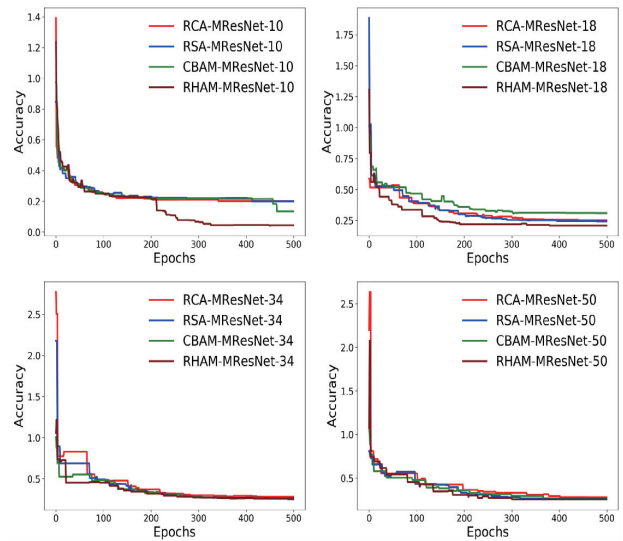


图 6 损失变化曲线

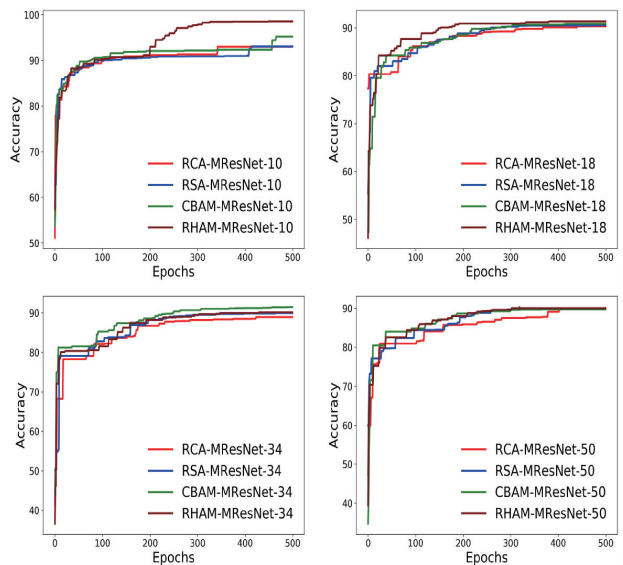


图 7 准确率变化曲线

将本文提出的 RHAM-MResNet-10 网络模型应用于脑部 CT 图像分类任务,并与 Gao 等人^[4]和 Zahedinasab 等人^[13]提出的方法进行比较,比较结果见表 5.

通过对比表 5 中所有模型分类准确度,可以得到以下 2 点结论:

(1) 与 Gao 等人^[4]的分类结果相比,说明了在普通的卷积神经网络模型中加入注意力机制,能够有效提升模型的性能.同时也验证了本文提出的残差混合注意力模块能够帮助模型聚焦脑部相关组织的位置和内容信息,提取更具有分辨性的特征,提升模型分类精度.

(2) 与 Zahedinasab 等人^[13]的分类结果相比,说明了加入注意力机制同时结合抑制过拟合方法的模型,在提升模型分类效果时比仅修改激活函数更有效.

表 4 不同的注意力模块分类结果对比

Model	Accuracy (%)
RCA-MResNet-50	89.32 ± 0.6
RSA-MResNet-50	89.60 ± 0.7
CBAM-MResNet-50	89.71 ± 0.2
RHAM-MResNet-50	89.89 ± 1.2
RCA-MResNet-34	89.65 ± 0.4
RSA-MResNet-34	89.95 ± 0.1
CBAM-MResNet-34	90.48 ± 0.3
RHAM-MResNet-34	90.18 ± 0.4
RCA-MResNet-18	90.21 ± 0.2
RSA-MResNet-18	90.42 ± 0.1
CBAM-MResNet-18	90.76 ± 0.4
RHAM-MResNet-18	90.91 ± 0.2
RCA-MResNet-10	92.98 ± 0.2
RSA-MResNet-10	93.02 ± 0.8
CBAM-MResNet-10	95.20 ± 1.0
RHAM-MResNet-10	97.47 ± 1.1

表 5 不同的模型分类结果对比

Module	Accuracy (%)
2D SIFT ^[4]	85.61 ± 1.1
2D KAZE ^[4]	86.31 ± 1.0
2D CNN ^[4]	86.32 ± 0.0
Mixed LReLU/ELU CNN ^[13]	87.12 ± 8.1
GATED LReLU/ELU CNN ^[13]	87.44 ± 9.0
MixedPReLU/PELU CNN ^[13]	87.85 ± 8.3
GATED PReLU/PELU ^[13]	88.59 ± 9.3
RHAM-MResNet-10	97.47 ± 1.1

4 总结与展望

为了帮助医生早期诊断患者大脑出现异常症状,减少人工诊断的差异性,本文提出了一个 RHAM-MResNet-10 残差网络模型对 AD、病变(如脑肿瘤)和健康老化的 3 类脑部 CT 图像进行分类.在该模型的残差映射结构中加入残差混合注意力模块,不仅能够关注大脑在图像中的位置信息,还能重点捕捉脑部相关组织的特征信息,极大降低噪声的影响,提升模型的分类准确率.在样本数量有限的情况下,加入全局平均池化以及 Dropout 等技巧,并结合标签平滑交叉熵损失函数,缓解了过拟合现象,提升了模型的泛化能力.系列实验证明了提出的 RHAM-MResNet-10 残差网络模型在脑部 CT 图像分类问题的有效性.未来的工作中,将继续探索脑部 CT 图像包含的病理信息,采用 3D 卷积神经网络同时处理每个病人的 20 张切片,不仅提取切片的静态信息,还要捕获切片之间的渐变信息,能够更好地预测疾病的发展趋势.

参考文献

[1] 王旭,段辉宏,裴生东.基于 CT 影像组学的非小细胞肺癌预后分析方法[J].电子学报,2020,48(4):637-642.
Wang X,Duan H H,Nie S D. Prognostic analysis method for

non-small cell lung cancer based on CT radiomics[J]. Acta Electronica Sinica,2020,48(4):637-642. (in Chinese)

[2] 桂春燕.应用 CT 和磁共振成像诊断盆腔肿瘤的临床价值[J].影像研究与医学应用,2020,4(15):40-41.
Gui C Y. The clinical value of CT and magnetic resonance imaging in the diagnosis of pelvic tumors[J]. Journal of Imaging Research and Medical Applications,2020,4(15):40-41. (in Chinese)

[3] 朱立军. CT、MRI 进行股骨头坏死放射诊断的效果观察[J].影像研究与医学应用,2020,4(15):136-137.
Zhu L J. The effect of CT and MRI in radiographic diagnosis of femoral head necrosis[J]. Journal of Imaging Research and Medical Applications,2020,4(15):136-137. (in Chinese)

[4] Gao W X,Hui R,Tian Z. Classification of CT brain images based on deep learning networks[J]. Computer Methods and Programs in Biomedicine,2017,138:49-56.

[5] Cheng J Z,Ni D,Chou Y H,et al. Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in US images and pulmonary nodules in CT scans[J]. Scientific Reports. 2016,6:1-13.

[6] 王磊,徐涛,宋传东,等.基于深度学习的 miRNA 与疾病相关性预测算法[J].电子学报,2020,48(5):870-877.
Wang L,Xu T,Song C D,et al. Prediction algorithm of association between miRNAs and diseases based on deep learning[J]. Acta Electronica Sinica,2020,48(5):870-877. (in Chinese)

[7] Ciompi F,Hoop D B,Riel S J,et al. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box[J]. Medical Image Analysis,2015,26(1):195-202.

[8] Oktay O,Ferrante E,Kamnitsas K,et al. Anatomically constrained neural networks (ACNNs): Application to cardiac image enhancement and segmentation[J]. IEEE Transactions on Medical Imaging,2018,37(2):384-395.

[9] Leclerc S,Smistad E,Pedrosa J,et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography[J]. IEEE Transactions on Medical Imaging,2019,38(9):2198-2210.

[10] Pereira S,Pinto A,Alves V,et al. Brain tumor segmentation using convolutional neural networks in MRI images[J]. IEEE Transactions on Medical Imaging,2016,35(5):1240-1251.

[11] Kumar A,Kim J,Lyndon D,et al. An ensemble of fine-tuned convolutional neural networks for medical image classification[J]. IEEE Journal of Biomedical and Health Informatics,2017,21(1):31-40.

[12] Farooq A,Anwar S,Awais M,et al. A deep CNN based

- multi-class classification of Alzheimer's disease using MRI [A]. IEEE International Conference on Imaging Systems and Techniques [C]. New York, USA: IEEE, 2017. 1 – 6.
- [13] Zahedinasab R, Mohseni H. Using deep convolutional neural networks with adaptive activation functions for medical CT brain image classification [A]. International Iranian Conference on Biomedical Engineering [C]. New York, USA: IEEE, 2018. 1 – 6.
- [14] Wang F, Jiang M Q, Qian C, et al. Residual attention network for image classification [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE, 2017. 3156 – 3164.
- [15] Hu J, Shen L, Sun G. Squeeze-and-Excitation networks [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE, 2018. 7132 – 7141.
- [16] Woo S, Park J, Lee J, et al. CBAM: Convolutional block attention module [A]. Proceedings of the European Conference on Computer Vision [C]. Berlin, Germany: Springer, 2018. 3 – 19.
- [17] Cao Y H, Chen K, Loy C C, et al. Prime sample attention in object detection [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE, 2020. 11583 – 11591.
- [18] Kaul C, Manandhar S, Pears N. Focusnet: An attention-based fully convolutional network for medical image segmentation [A]. IEEE 16th International Symposium on Biomedical Imaging [C]. New York, USA: IEEE, 2019. 455 – 458.
- [19] 唐海桃, 薛嘉宾, 韩纪庆. 一种多尺度前向注意力模型的语音识别方法 [J]. 电子学报. 2020, 48(7): 1276 – 1283.
- Tang H T, Xue J B, Han J Q, et al. A method of multi-scale forward attention model for speech recognition [J]. Acta Electronica Sinica, 2020, 48(7): 1276 – 1283. (in Chinese)
- [20] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE, 2016. 770 – 778.
- [21] Xiang L, Shuo C, Xiao L H, et al. Understanding the disharmony between dropout and batch normalization by variance shift [A]. IEEE Conference on Computer Vision [C]. New York, USA: IEEE, 2019. 2677 – 2685.
- [22] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE, 2016. 2818 – 2826.
- [23] Kingma D, Ba J. Adam: A method for stochastic optimization [J]. arXiv Preprint, 2015, arXiv:1412.6980.

作者简介



乔思波 男, 1993 年 7 月生, 山东聊城人。现为中国石油大学(华东)计算机技术与资源信息工程专业博士研究生, 主要研究方向为医疗图像处理、人工智能算法及其应用。

E-mail: siboqiao@qq.com

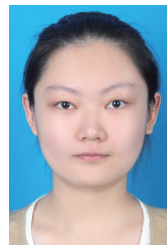


庞善臣(通信作者) 男, 1974 年 10 月生, 山东济宁人。现为中国石油大学(华东)计算机科学与技术学院三级教授、博士生导师, 山东省泰山产业领军人才。主要从事 Petri 网理论、可信计算、人工智能算法、医疗图像处理等方面的研究。

E-mail: pangsc@upc.edu.cn



王敏 女, 1994 年 12 月生, 江苏泰州人。现为中国石油大学(华东)控制科学与工程专业博士研究生。主要研究方向为图像处理。



翟雪 女, 1995 年 8 月生, 山东聊城人。现为中国石油大学(华东)计算机技术与资源信息工程专业博士研究生。主要研究方向为区块链、图像分析。



于世行 男, 1992 年 11 月生, 山东聊城人。现为天津工业大学机械工程专业博士研究生。主要研究方向为数据分析、图像处理和推荐系统。



丁桐 男, 1994 年 7 月生, 宁夏银川人。现为山东大学人工智能专业博士研究生。主要研究方向为人工智能算法及其应用、医疗图像处理。