

基于联邦学习与改进 CBAM-ResNet18 的脑肿瘤分类^①



吴波^{1,2}, 史东辉^{1,2}, 吕东来³, 胡涛^{1,2}

¹(安徽建筑大学 电子与信息工程学院, 合肥 230601)

²(质谱关键技术研发与临床应用安徽省联合共建学科重点实验室, 合肥 230601)

³(中国人民解放军联勤保障部队第 901 医院 肿瘤科, 合肥 230031)

通信作者: 史东辉, E-mail: shidonghui@ahjzu.edu.cn

摘要: 针对联邦学习框架下, 基于卷积注意力模块的多客户端脑肿瘤分类方法对于 MRI 图像中肿瘤区域细节提取能力不足、通道注意力与空间注意力相互干扰的问题, 以及针对多点医疗肿瘤数据分类准确性低的问题, 提出了一种融合联邦学习框架和改进的 CBAM-ResNet18 网络的脑肿瘤分类方法. 利用联邦学习特性联合多点脑肿瘤数据, 采用 Leaky ReLU 激活函数代替 ReLU 激活函数以减轻神经元死亡, 将卷积注意力模块中的通道注意力模块由先降维再升维改成先升维再降维, 充分提高网络对图像细节的提取能力, 将卷积注意力模块中的通道注意力模块与空间注意力模块由级联结构改为并联结构, 使得网络的特征提取能力不会受到二者先后顺序的影响. 通过在 Kaggle 公开的脑肿瘤 MRI 数据集上的进行实验, 该方法的准确率、精准度、召回率与 F1 值分别为 97.78%、97.68%、97.61% 与 97.63%, 比基准模型分别高 6.54%、4.78%、6.80%、7.00%. 实验结果证明, 该方法不仅能够打破数据孤岛, 实现多点数据融合, 而且比多数现有主流模型的性能更好.

关键词: 脑肿瘤分类; 联邦学习; 卷积注意力模块; 残差网络; 数据孤岛

引用格式: 吴波, 史东辉, 吕东来, 胡涛. 基于联邦学习与改进 CBAM-ResNet18 的脑肿瘤分类. 计算机系统应用, 2024, 33(4): 39-49. <http://www.c-s-a.org.cn/1003-3254/9469.html>

Brain Tumor Classification Based on Federated Learning and Improved CBAM-ResNet18

WU Bo^{1,2}, SHI Dong-Hui^{1,2}, LYU Dong-Lai³, HU Tao^{1,2}

¹(School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei 230601, China)

²(Key Laboratory of Mass Spectrometry Key Technology R&D and Clinical Application in Anhui Province, Hefei 230601, China)

³(Oncology Department, the 901st Hospital of the Joint Logistics Support Force of PLA, Hefei 230031, China)

Abstract: The multi-client brain tumor classification method based on the convolutional block attention module has inadequate extraction of tumor region details from MRI images, and channel attention and spatial attention interfere with each other under the federated learning framework. In addition, the accuracy in classifying medical tumor data from multiple points is low. To address these problems, this study proposes a brain tumor classification method that amalgamates the federated learning framework with an enhanced CBAM-ResNet18 network. The method leverages the federated learning characteristic to collaboratively work with brain tumor data from multiple sources. It replaces the ReLU activation function with Leaky ReLU to mitigate issues of neuron death. The channel attention module within the convolutional block attention module is modified from a dimension reduction followed by a dimension increment approach to a dimension increment followed by a dimension reduction approach. This change significantly enhances the

① 基金项目: 国家自然科学基金青年项目 (620011004); 安徽省科研编制计划 (2022AH050224); 安徽省重点研究与开发计划 (202004a07020050); 质谱关键技术研发与临床应用安徽省联合共建学科重点实验室开放课题 (2023ZPLH07)

收稿时间: 2023-10-26; 修改时间: 2023-11-27; 采用时间: 2023-12-07; csa 在线出版时间: 2024-03-04

CNKI 网络首发时间: 2024-03-08

network's ability to extract image details. Furthermore, the architecture of the channel attention module and spatial attention module in the convolutional block attention module has been shifted from a cascade structure to a parallel structure, ensuring that the network's feature extraction capability remains unaffected by the order of processing. A publicly available brain tumor MRI dataset from Kaggle is used in the study. The results demonstrate that FL-CBAM-DIPC-ResNet has a remarkable performance. It achieves impressive accuracy, precision, recall, and *F1* score of 97.78%, 97.68%, 97.61%, and 97.63%, respectively. These values of accuracy, precision, recall, and *F1* score are 6.54%, 4.78%, 6.80%, and 7.00% higher than those of the baseline model. These experimental findings validate that the proposed method not only overcomes data islands and enables data fusion from multiple sources but also outperforms the majority of existing mainstream models in terms of performance.

Key words: brain tumor classification; federated learning; convolutional block attention module (CBAM); residual network (ResNet); data island

脑肿瘤是位于脑部或其附近出现的任何肿瘤,其异质性强,占全球癌症病例的大约3%,男性比女性更常见^[1]。早期发现脑肿瘤患者所患肿瘤的类型,及时制定方案会提高治疗的生存率。评估患者的脑肿瘤影像是一项重要的工作,传统人工筛选肿瘤影像需要评估人具有一定的专业知识,而且工作人员工作量大,容易引起疲劳,进而影响评估结果。因此,寻找一种高效而准确的脑肿瘤自动诊断方法是当前医学影像领域的迫切需求之一。

近年来,计算机辅助诊断 (computer aided diagnosis, CAD) 对生物医学领域有重要影响,特别是将深度学习和医学影像结合能帮助医生提高诊断结果的准确率^[2]。各种深度学习模型在脑肿瘤分类中取得了一系列令人瞩目的结果。然而,现有研究往往在处理脑 MRI 图像时存在一定的局限性,对肿瘤区域特征关注不足,易损失细节特征,影响了诊断的准确性。在此背景下,已有学者提出了多种改进模型的方法。

Deepak 等^[3]将深度迁移学习用于三分类脑瘤识别,达到了 98.00% 的准确率,但存在小数据过度拟合问题。Ghassemi 等^[4]先将深度神经网络 (DNN) 作为生成对抗网络 (GAN) 在脑肿瘤数据集进行预训练,再替换全连接层,并将整个 DNN 用来区分脑肿瘤类别,达到了 95.60% 的准确率,但由于 GAN 的限制网络输入大小只能为 64×64。Brindha 等^[5]比较了人工神经网络 (ANN) 和卷积神经网络 (CNN) 在脑肿瘤分类中的性能, CNN 准确率为 89.00%, 高于 ANN 的 65.21%, 表明在脑肿瘤分类上, CNN 比 ANN 有更好的效果。Sultan 等^[6]提出一种基于 CNN 的深度学习模型,并在两个公

开的脑肿瘤数据集上验证进行实验,分别得到 96.13% 和 98.70% 的准确率,但没有在合并两个数据集的情况下实验。夏景明等^[7]将多尺寸卷积核、双通道池化层和多深度融合残差模块组合成 MDM-ResNet 网络,在脑肿瘤数据集上达到了 93.51% 的准确率,但多深度融合残差模块造成残差模块变宽,有较大的计算负担。由此可见,已有不少研究者进行了将深度学习与脑肿瘤医学影像结合的研究。但脑 MRI 图像肿瘤区域小,这些研究对肿瘤区域特征的关注不足,容易损失细节特征。

Özbay 等^[8]提出一种结合可解释性与特征融合的精确哈希方法,能够关注肿瘤区域,搜索精度较其他模型提升了 0.88%–5%。张晓倩等^[9]将高效通道注意力 (efficient channel attention, ECA) 和瓶颈注意力 (bottleneck attention module, BAM) 加入 SE-Net 网络,对脑肿瘤识别的准确率达到 99.47%。Xu 等^[10]将卷积注意力模块 (convolutional block attention module, CBAM) 加入 U-Net 网络,在脑肿瘤分类中的准确率达到 94.23%, 但原始 CBAM 注意力不足。Ljaz 等^[11]在癌组织病理学数据集中训练了 VGG 模型用作特征提取器,并在卷积块的第 1 个卷积层增加了 GAP 层和 CBAM, 实现了特殊的 VGG 模型,达到了 98.96% 的准确率,但其是二分类模型,无法保证在脑肿瘤多分类任务中的性能。安晨等^[12]将注意力机制、关系网络和元学习相结合,在无线胶囊内镜图像数据集上的准确率达到 90.28%, 但其仅针对小样本图像分类,模型泛化能力不足。由此可见,已有部分研究者将注意力机制用在医学图像分类任务中。尽管这些方法在一定程度上改进了诊断性能,但大

部分都没有改进注意力机制,仍存在一些共性问题,如注意力模块的不足和对小样本数据的局限性,并且少有处理多点脑肿瘤数据的研究。

Adnan 等^[13]针对组织病理学分类设计了联邦学习框架,准确率与集中训练相当,表明联邦学习在医疗方面的可靠性,但该框架仅针对二分类任务。Che 等^[14]提出一个通用多视图联邦学习框架来分析双向情感研究中的3种类型的元数据,最高达到88.92%的准确率,但其是针对文本数据,无法分析图像。许亚倩等^[15]基于联邦学习对多源数据训练疾病得分预测模型,虽然模型很稳定,但与其所对比的方法相比并不突出。段聪颖等^[16]提出一种基于联邦学习的皮肤病图像分类算法,准确度达到75.9%,但其没有关注图像细节。曹润芝等^[17]将联邦学习用于肺炎图像分类,得到95.2%的准确率,但该方法没有针对肺炎图像细节进行特征提取。由此可见,已有不少研究者将联邦学习用于医疗数据,但在脑肿瘤分类中应用联邦学习的研究较少。

由上可知,现实中仍然存在以下问题:(1)医疗数据存在数据孤岛现象,不同医院数据分布存在差异,单一医院训练的模型泛化能力不足;(2)传统深度学习大多只能进行集中式训练;(3)结合卷积注意力的残差网络特征提取能力不足,卷积注意力先降维再升维的设置会导致细节特征的损失,且空间注意力的计算会受到通道注意力的影响。

针对以上问题,本文提出了一种结合联邦学习框架(federated learning, FL)、改进残差网络(residual network, ResNet)和改进卷积注意力模块(convolutional block attention module-dimension increment-parallel connections, CBAM-DIPC)的方法(federated learning-convolutional block attention module-dimension increment-parallel connections-residual network, FL-CBAM-DIPC-ResNet)。用Leaky ReLU激活函数替换残差网络中的ReLU激活函数,减少神经元死亡的问题;将卷积注意力模块中通道注意力由原先的先降维再升维改成先升维再降维,减少图像特征细节的损失;将通道注意力模块(channel attention module, CAM)^[18]和空间注意力模块(spatial attention module, SAM)^[19]由级联结构替换为并联结构,使两个注意力模块直接学习原始输入特征图,而无须关注通道注意力和空间注意力的先后顺序;最后融合联邦学习框架处理多点医疗数据,解决数据孤岛现象。

1 相关理论

1.1 残差网络

自LeNet^[20]网络的提出给深度学习带来了新突破后,之后又涌现出许多优秀网络,如AlexNet^[21]、VGG^[22]、GoogLeNet^[23]等网络。但随着网络层数的增加开始出现梯度消失的情况。He 等^[24]提出深度残差网络(residual network, ResNet)网络,引入残差模块解决了梯度消失问题。残差网络能通过增加网络深度来提高模型的精度,其内部残差模块使用跳跃连接实现恒等映射,解决了过去深度神经网络由于加深网络层数而产生梯度消失的问题。

1.2 卷积注意力模块

Woo 等^[25]提出卷积注意力模块(convolutional block attention module, CBAM),CBAM由通道注意力模块和空间注意力模块处理输入特征图。CBAM首先计算输入特征图 F 的通道注意力,得到通道注意力加权的特征图 $F1$ 。然后在 $F1$ 的基础上进行空间注意力计算,得到空间注意力加权的特征图 $F2$ 。这是一个“级联”结构,公式见式(1)、式(2):

$$F1 = Mc(F) \otimes F \quad (1)$$

$$F2 = Ms(F1) \otimes F1 \quad (2)$$

其中, F 表示原始输入特征图, \otimes 表示特征图加权的乘法操作,式(1)中 $Mc(F)$ 表示特征图 F 经过通道注意力计算后的输出,式(2)中 $Ms(F1)$ 表示特征图 $F1$ 经过空间注意力计算后的输出。

1.3 联邦学习

联邦学习^[26]是一种特殊的分布式计算框架,由服务器与多个客户端组成,能让客户端在各自本地处理数据,并融合这些数据的特征进行建模分析。图1是联邦学习训练过程。首先进行任务初始化,参与者被选择并向其分配一个训练任务,然后参与者下载全局模型并在本地训练模型。在全局模型聚合阶段,参与者上传模型参数到服务器,全局模型被更新,以便下一轮训练使用。全局模型聚合公式如式(3):

$$\omega_{t+1} = \sum_{k \in V} \frac{n_k}{n} \omega_{t+1}^k + \sum_{k \notin V} \frac{n_k}{n} \omega_t \quad (3)$$

其中, k 是参与客户端的个数, V 是选中的客户端集合, n 是参与聚合客户端的样本总数, n_k 是每个客户端训练数据的数量, ω_t 是当前全局模型, ω_{t+1} 是更新的全局模型。

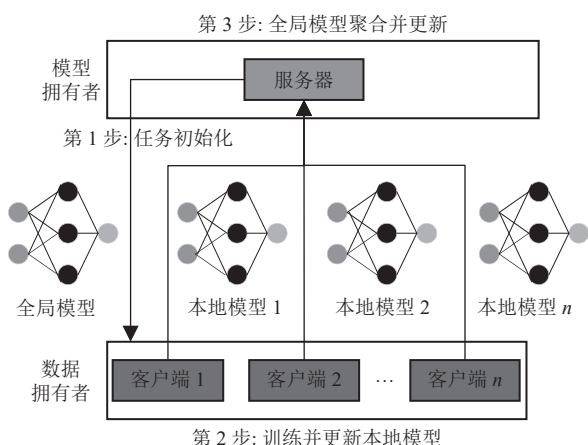


图1 联邦学习训练过程

2 FL-CBAM-DIPC-ResNet 方法

本文以建立多点脑肿瘤数据模型、提升脑肿瘤分类的准确性为目标。首先基于联邦学习框架，将残差网络中的 ReLU 激活函数替换成 Leaky ReLU 激活函数，然后把卷积注意力模块中通道注意力先降维再升维改成先升维再降维，并把通道注意力与空间注意力用并联方式连接，然后将改进后的卷积注意力模块引入基于联邦学习的改进残差网络中，最终实现脑肿瘤分类的 FL-CBAM-DIPC-ResNet 模型。

2.1 改进 ResNet 网络

本研究改进了 ResNet-18 网络，旨在提高模型对脑肿瘤图像特征的提取能力。标准的 ResNet-18 网络使用 ReLU 激活函数，然而当输入 x 为负的时候，ReLU 激活函数会出现神经元死亡的问题。为了解决这一问题，本研究选择采用 Leaky ReLU 激活函数，Leaky ReLU 是 ReLU 的优秀变体，拥有 ReLU 所有优点的同时能够保留输入的负值信息^[27]。Leaky ReLU 在 $x < 0$ 时得到一个负数梯度值，解决了 ReLU 神经死亡问题，如图 2 所示。ReLU 激活函数与 Leaky ReLU 激活函数公式分别如式 (4)、式 (5) 所示：

$$\text{ReLU}(x) = \begin{cases} \max(0, x), & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (4)$$

$$\text{Leaky ReLU}(x) = \begin{cases} x, & x \geq 0 \\ \alpha x, & x < 0 \end{cases} \quad (5)$$

其中， x 是输入， α 是泄露系数，且 α 小于 1。

采用 Leaky ReLU 激活函数的改进，有助于缓解梯度消失问题，使得 ResNet 网络更能够适应深层次

的学习任务，进而提高对脑肿瘤图像特征的建模能力。这一改进能确保模型在处理脑肿瘤图像时更加鲁棒和有效。

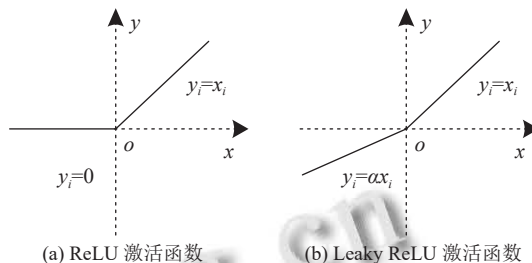


图2 激活函数对比图

2.2 改进卷积注意力模块

脑 MRI 是医学影像领域中一种重要的图像模态，脑组织具有复杂结构和多样性。在脑 MRI 图像中，由于其高分辨率和生物医学特性，涉及丰富的细节特征。原始卷积注意力模块先降维再升维会造成重要信息丢失和噪声引入，而且先计算通道注意力再计算空间注意力的方式会导致空间注意力的计算受到干扰，使模型不稳定。因此原始卷积注意力模块在处理脑 MRI 图像面临挑战，存在细节提取能力不足，以及模型稳定性易受到通道注意力与空间注意力计算先后顺序的影响。本节将针对这些问题对卷积注意力模块进行通道升维和并联结构的改进，以提高模型对脑肿瘤图像细节特征的提取能力以及提高模型的稳定性。

CBAM 由 CAM 和 SAM 组成。CAM 对输入特征图 F 进行平均池化和最大池化操作，并使其压缩成 $1 \times 1 \times C$ 格式。压缩后的特征图经过含有降维隐藏层的两次 MLP 网络，使压缩后的 $1 \times 1 \times C$ 特征图降维至 $1 \times 1 \times C/r$ 。由于脑肿瘤图像肿瘤区域尺寸小且数据集较少，因此该降维操作会导致 CBAM 不能很好地注意到图像中的细节。本研究将改进 CAM 中的通道降维操作，如图 3 所示。原 CAM 先将图像从 $1 \times 1 \times C$ 降维到 $1 \times 1 \times C/r$ ，再升维到 $1 \times 1 \times C$ 。改进的 CBAM 先升维再降维，即先将图像从 $1 \times 1 \times C$ 升维到 $1 \times 1 \times rC$ ，再降维到 $1 \times 1 \times C$ 。对于细节特征丰富的脑肿瘤图像，先升维有助于增加特征的细粒度表示，再降维可以将关注点重新集中到最重要的通道上。这样可以更好地保留原始特征，减少信息损失，提升 CBAM 提取脑肿瘤图像细节的能力。

原本的 CBAM 先进行通道注意力计算，再进行空

间注意力计算. 从某种程度上, 空间注意力的计算会受到通道注意力计算的影响, 对模型产生干扰, 导致模型不稳定. 本研究将原来 CBAM 中 CAM 和 SAM 由原来的“级联”方式改进为“并联”方式, 因此, CAM 和 SAM 将直接学习原始特征图, 一个注意力的计算不会受到另一个注意力计算影响. 改进的 CBAM 如图 4 所示,

并联公式如式 (6):

$$F1 = ((Mc(F) \otimes F) \oplus (Ms(F) \otimes F)) \otimes F \quad (6)$$

其中, F 表示原始输入特征图, $Mc(F)$ 表示特征图 F 经过通道注意力计算后的输出, $Ms(F)$ 表示特征图 F 经过空间注意力计算后的输出, \otimes 表示特征图加权的乘法操作, \oplus 表示特征图加权的加法操作.

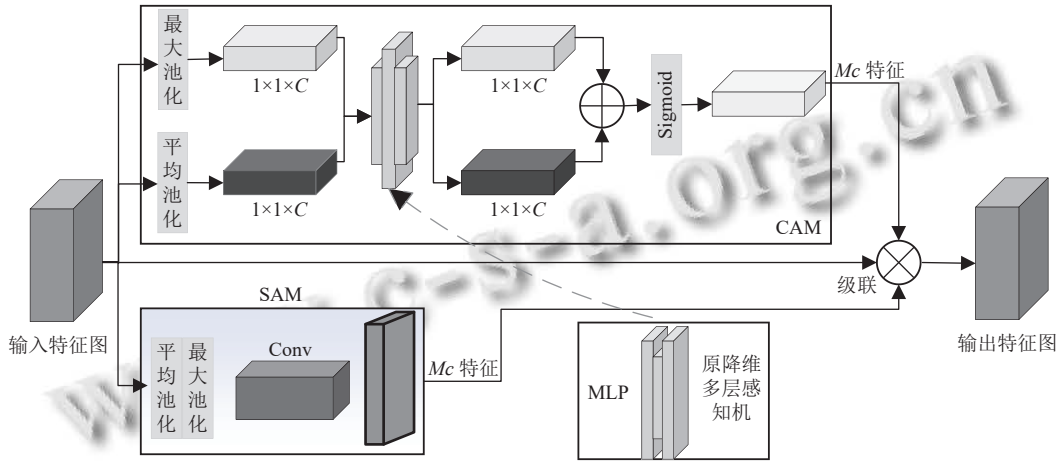


图3 CBAM 升维感知机

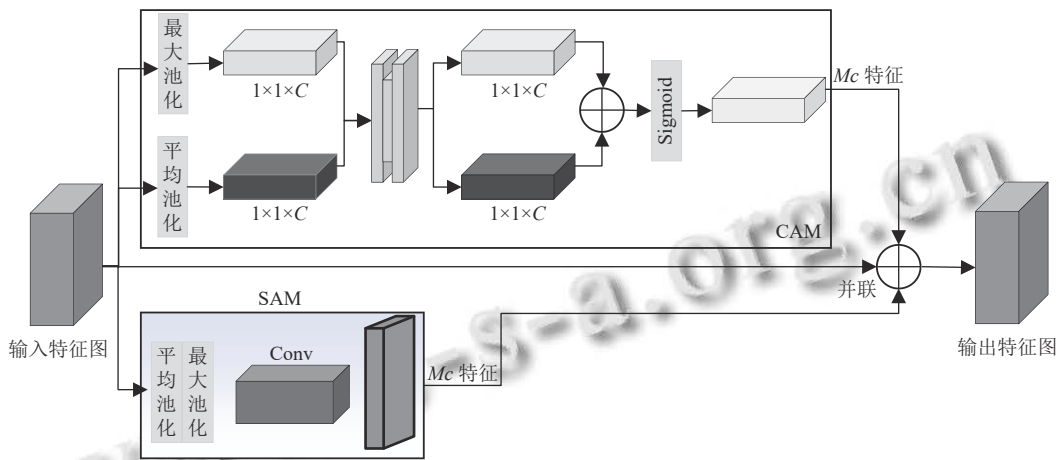


图4 CBAM 并联替换级联

同时将 CBAM 由先降维再升维改进为先升维再降维, 并将 CAM 与 SAM 由“级联”改成“并联”, 其整体结构如图 5 所示. 通道注意力计算公式如式 (7)、空间注意力公式如式 (8).

$$Mc(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \\ = \sigma(W_1(W_2(F_{\text{avg}}^c) + W_1(W_0(F_{\text{max}}^c)))) \quad (7)$$

$$Ms(F) = \sigma(f^{7 \times 7}[\text{AvgPool}(F); \text{MaxPool}(F)]) \\ = \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])) \quad (8)$$

其中, σ 表示 Sigmoid 激活函数, W_0 、 W_1 表示升维 MLP 权重, F_{avg}^c 和 F_{max}^c 表示平均池化特征, F_{max}^c 和 F_{max}^c 表示最大池化特征, $f^{7 \times 7}$ 表示 7×7 的带过滤器的卷积操作.

2.3 FL-CBAM-DIPC-ResNet 框架

不同医院数据分布存在差异, 用单一医院数据训练出来的模型泛化能力较弱, 整合多个医院的数据有利于提升模型的泛化能力. 联邦学习是一种服务器-客户端架构, 可以整合多个医院的数据, 得到鲁棒性更强的模型. 在训练过程中, 每个医院的客户端使用本地数

据进行模型的训练, 而无需将敏感数据传输到中心服务器. 通过联邦学习, 模型可以从多个医院的数据中学

到更全面和具有代表性的特征, 从而提高对不同医院数据的适应能力.

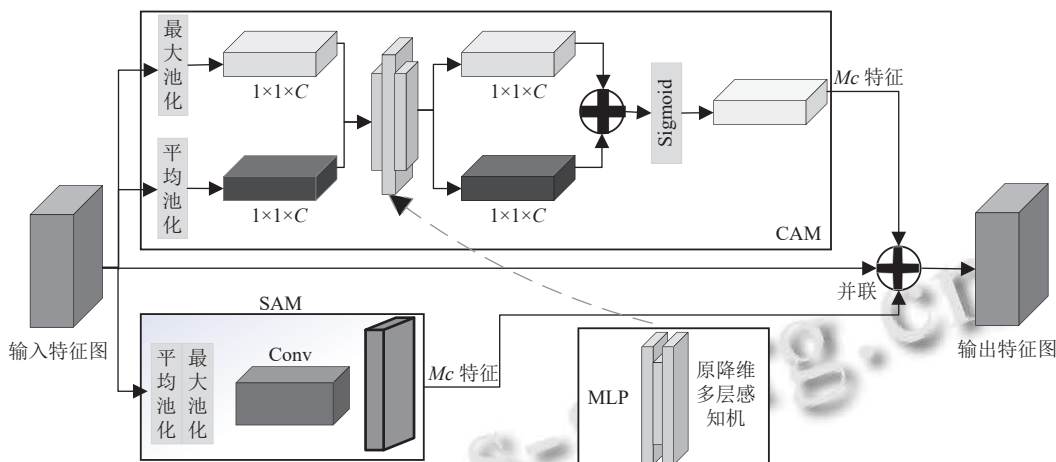


图5 改进的CBAM

本研究将联邦学习框架与改进的CBAM-ResNet18相结合, 使用一个中心服务器协调多个客户端进行分布式协同训练, 使得模型能够在考虑不同医院数据特点的同时, 通过改进的网络结构更好地捕获脑肿瘤图像的细节特征.

以脑肿瘤MRI图像四分类为例, FL-CBAM-DIPC-ResNet算法的流程图如图6所示. 首先对原始数据进行数据预处理, 将数据按照联邦学习要求划分成多个客户端, 然后FL-CBAM-DIPC-ResNet算法对训练集进行训练. 在全局聚合时, 每个客户端的数据仅在本地训练, 客户端仅上传训练好的模型参数. 通过该算法, 最终实现脑肿瘤MRI图像的分类. 这种综合方法有望提高脑肿瘤分类模型的性能, 并使其更适用于多医院的实际应用场景.

另外本研究的联邦学习算法如算法1所示.

算法1. FL-CBAM-DIPC-ResNet 算法

输入: 从每轮聚合客户端比例 C (0-1); 客户端 $client_j$ ($1 \leq j \leq J$); 学习率 e , 本地迭代次数 l_s , 全局拟合轮数 $t \in T$.
输出: 全局 FL-CBAM-DIPC-ResNet 模型权重.

- 1) 初始化 FL-CBAM-DIPC-ResNet 模型;
- 2) 每轮选择一部分客户端参与训练;
- 3) for t in T :
 - 3.1. $m \leftarrow \max(C \cdot J, 1)$;
 - 3.2. $V_t \leftarrow$ 选择的 m 个客户端的子集;
 - 3.3. $client_j \in V_t$, 训练轮数 $i \in l_s$, 本地训练;
 - 3.4. $\omega_{i,j,t} = \omega_{i-1,j,t} - e \nabla \ell_{j,t}(\omega_{i-1,j,t}; b)$, $i \in l_s$, 各客户端通过下载全局模型参数进行本地训练;
 - 3.5. $\omega_{i+1} \leftarrow \sum_{k \in V} \frac{n_k}{n} \omega_{i+1}^k$, 根据每个 $client_j$ 的参数 $\omega_{j,t}$ 选择 m 个客户端进行模型聚合;
- 4) 训练结束.

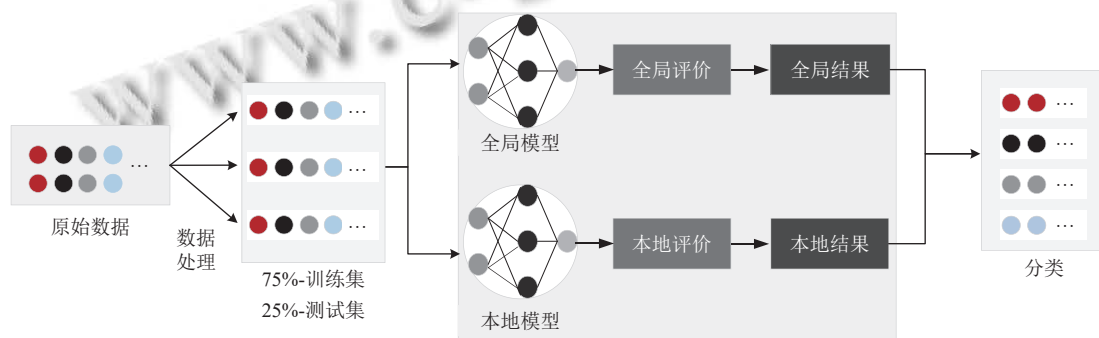


图6 FL-CBAM-DIPC-ResNet 算法流程图

3 实验分析

3.1 实验环境

本研究实验环境为 Ubuntu 16.04.6 LTS, Intel(R)

Core(TM) i7-6850K CPU @ 3.60 GHz 处理器, 8 GB NVIDIA GeForce GTX 1070Ti 显卡, 64 GB 内存, 1 TB 硬盘, 编程语言为 Python, 实验框架基于 PyTorch.

3.2 数据集与预处理

本研究使用的数据集来自 Kaggle 公开的脑肿瘤 MRI 数据集^[28]. 该数据集由 Figshare、SARTAJ 和 Br35H 这 3 个数据集组成, 分为 4 类: 神经胶质瘤、脑膜瘤、无肿瘤和垂体瘤. 数据集共有 7 023 张人脑 MRI 图像, 其中神经胶质瘤 1 621 例、脑膜瘤 1 645 例、无肿瘤 2 000 例和垂体瘤 1 757 例. 部分人脑 MRI 图像如图 7 所示. 参考文献[6,8,9]对脑肿瘤数据集的划分方式, 本实验将训练集与测试集比例划分为 3:1. 由于该数据集图像统一且分布均匀, 因此裁剪图像尺寸为 224×224, 并进行归一化操作. 同时将数据集分割成数量相同的 3 个客户端, 当图像数量无法均分时, 将多余的图像划分到第 3 个客户端, 分别是 client0、client1、client2, 如表 1 所示.

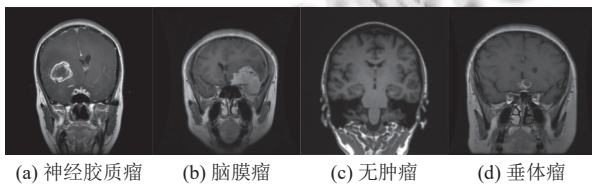


图 7 部分人脑 MRI 图像

表 1 数据集划分(张)

客户端	神经胶质瘤	脑膜瘤	无肿瘤	垂体瘤
client0	540	548	666	585
client1	540	548	666	585
client2	541	549	668	587

3.3 评估指标

为全面评价 FL-CBAM-DIPC-ResNet 的性能, 采用准确率 (Accuracy, ACC)、精确率 (Precision, P)、召回率 (Recall) 与 $F1$ 值 ($F1$) 作为评价标准, 公式分别见式 (9)–式 (12):

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

$$P = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = \frac{2P \times Recall}{P + Recall} \quad (12)$$

其中, TP 指分类为正且实际为正确的数量, TN 指分类有误且实际为反例的数量, FP 指分类为正但实际为反例的数量, FN 指分类有误但实际为正确的数量.

3.4 实验结果

本节使用 Kaggle 公开脑肿瘤 MRI 数据集^[28]进行实验. 其中, 客户端设置为 3 个, 批次大小为 32, 学习率为 0.01, 本地迭代次数为 5, 全局拟合轮数为 30. 实验基准模型为基于联邦学习框架下的 ResNet-18 网络, 命名为 FL-ResNet. 本研究完成以下实验内容: (1) 比较 CBAM 在 FL-ResNet 中不同位置的影响; (2) 比较 ReLU 与 Leaky ReLU 激活函数的性能; (3) 比较卷积注意力模块中 CAM 升至不同维度的性能; (4) 比较 CAM 与 SAM 级联与并联的性能; (5) 比较 FL-CBAM-DIPC-ResNet 模型与其他主流模型的性能.

3.4.1 CBAM 在 FL-ResNet 不同位置的性能

本节将 CBAM 分别放在 FL-ResNet 网络的不同位置, CBAM 在残差模块内部的网络命名为 CBAM-In, CBAM 在 ResNet-18 网络的第 1 层与最后一层的网络命名为 CBAM-ResNet. 实验结果如表 2, CBAM-ResNet 的准确率、精确率、召回率、 $F1$ 值比基准模型 FL-ResNet 分别高 4.26%、2.41%、4.42%、4.57%, 且比 CBAM-In 分别高 1.59%、1.01%、2.00%、1.81%. 结果表明 CBAM 能够增强网络性能, 且由于 ResNet-18 残差模块前后网络小, CBAM 对 ResNet-18 网络的残差模块前后影响更大.

表 2 CBAM 在残差模块不同位置 (%)

模型	ACC	P	Recall	$F1$
FL-ResNet	91.24	92.90	90.81	90.63
CBAM-In	93.91	94.30	93.23	93.39
CBAM-ResNet	95.50	95.31	95.23	95.20

3.4.2 ReLU 与 Leaky ReLU 性能比较

本节将加入了 CBAM 的 FL-ResNet 网络中的激活函数换成 Leaky ReLU 激活函数, 并命名为 CBAM-LR. CBAM-LR 与 CBAM-ResNet 实验对比结果如表 3. CBAM-LR 中准确率、精确率、召回率、 $F1$ 值比 CBAM-ResNet 分别高 0.35%、0.29%、0.33%、0.37%. 该结果表明 Leaky ReLU 激活函数能够减轻神经元死亡对模型性能造成的影响.

表 3 ReLU 与 Leaky ReLU 的对比 (%)

模型	ACC	P	Recall	$F1$
CBAM-LR	95.85	95.60	95.56	95.57
CBAM-ResNet	95.50	95.31	95.23	95.20

3.4.3 通道注意力升维的最佳维度

本节将探究卷积注意力中 CAM 升维的最佳维度,

第(1)组为第3.4.1节中的CBAM-ResNet模型, CBAM-ResNet是先降维到 $1 \times 1/16$,再升维到 $1 \times 1 \times C$.第(2)–(6)组分别是在第(1)组的基础上,将通道维度先升维到 $1 \times 1 \times 2C$ 、 $1 \times 1 \times 3C$ 、 $1 \times 1 \times 4C$ 、 $1 \times 1 \times 5C$ 、 $1 \times 1 \times 10C$,再降维到 $1 \times 1 \times C$.第(7)组是不采用CBAM的模型,实验结果如表4.该结果表明对通道注意力模块升维能够提升网络对特征图的提取能力,并且综合分析参数量、训练时长与准确率,在不明显提升参数量与训练时长的情况下,对通道注意力模块升维的最佳维度是3维.

表4 CBAM中CAM不同升维维度性能

组	升维策略	特征图大小	参数量	训练时长(轮/s)	ACC (%)
(1)	原CBAM	$1 \times 1 \times C/16$	123 380 540	122.5	95.50
(2)	$1 \times 1 \times 2$	$1 \times 1 \times 2C$	134 729 020	124.4	96.41
(3)	$1 \times 1 \times 3$	$1 \times 1 \times 3C$	140 586 300	124.9	96.76
(4)	$1 \times 1 \times 4$	$1 \times 1 \times 4C$	146 443 580	125.2	96.53
(5)	$1 \times 1 \times 5$	$1 \times 1 \times 5C$	152 300 860	125.7	95.33
(6)	$1 \times 1 \times 10$	$1 \times 1 \times 10C$	181 587 260	127.6	93.00
(7)	不采用CBAM	—	123 012 304	100.0	91.24

3.4.4 CAM与SAM的级联与并联性能比较

本节将比较卷积注意力模块中CAM与SAM级联与并联之间的性能差异.级联模型是第3.4.1节中的CBAM-ResNet模型,将改进后的并联模型命名为CBAM-ResNet-PC.实验结果如表5所示,CBAM-ResNet-PC中准确率、精确率、召回率、F1值比CBAM-ResNet分别高1.20%、1.15%、1.27%、1.27%.该结果表明CAM与SAM并联将直接学习原始特征图,后一个注意力的计算不会受到前一个注意力的影响,因此表现的性能更加好.

表5 CAM、SAM级联与并联对比 (%)

模型	ACC	P	Recall	F1
CBAM-ResNet-PC	96.70	96.46	96.50	96.47
CBAM-ResNet	95.50	95.31	95.23	95.20

3.4.5 消融实验

为了验证模块改进的有效性,本节进行了消融实验.消融实验结果如表6所示,其中第(1)组是基于联邦学习框架的ResNet-18基准模型FL-ResNet.第(2)组是在FL-ResNet模型残差模块的前后加入CBAM的模型CBAM-ResNet,该模型有利于提升FL-ResNet模型的注意力.第(3)组是将CBAM-ResNet模型中的激活函数替换为Leaky ReLU激活函数的模型CBAM-LR,该模型能够缓解网络中神经元死亡问题.第(4)组

是将CBAM-LR中通道注意力升3维,充分提取脑肿瘤图像的细节特征.第(5)组是对CBAM-LR中通道注意力与空间注意力进行并联操作,避免空间注意力的计算被通道注意力影响.第(6)组是将CBAM-LR中通道注意力升3维的同时与空间注意力并联的最终模型FL-CBAM-DIPC-ResNet.改进后的模型准确率达97.78%,在FL-ResNet模型中加入CBAM,更换Leaky ReLU激活函数,将通道注意力维数提升3维并与空间注意力并联后,较基准模型FL-ResNet提升了6.54%,充分验证了改进后模型的有效性.

表6 FL-CBAM-DIPC-ResNet消融实验

组	加入CBAM	使用Leaky ReLU	CAM通道升维	CAM与SAM并联	ACC (%)
(1)	×	×	×	×	91.24
(2)	√	×	×	×	95.50
(3)	√	√	×	×	95.85
(4)	√	√	√	×	97.32
(5)	√	√	×	√	97.10
(6)	√	√	√	√	97.78

3.4.6 与其他模型的对比

为了验证FL-CBAM-DIPC-ResNet模型的效果,将其与目前主流的经典模型进行比较.多数模型是基于ResNet和DenseNet改进的,本文将与ResNet-50、Res2NeXt-50^[29]、ResNeXt-50^[30]、ResNeSt-50^[31]、DenseNet-121^[32]、DPN-107^[33]、文献[9]进行比较.各项指标结果如表7所示.

表7 本文方法与经典模型对比 (%)

模型名	ACC	P	Recall	F1	数据分布	数据孤岛现象
ResNet-50	82.30	81.72	81.20	80.97	单点分布	存在
Res2NeXt-50	83.30	83.37	82.08	82.23	单点分布	存在
ResNeXt-50	88.02	88.03	87.24	87.49	单点分布	存在
ResNeSt-50	93.97	94.19	93.59	93.82	单点分布	存在
DenseNet-121	96.02	95.86	95.82	95.82	单点分布	存在
DPN-107	97.56	97.44	97.61	97.52	单点分布	存在
文献[9]	99.47	99.45	99.42	99.43	单点分布	存在
本文方法	97.78	97.68	97.61	97.63	多点分布	打破数据孤岛

由表7可以看出本文方法的各项指标高于大多数主流模型,仅比文献[9]的指标低.其中ResNet-50准确率最低,仅82.30%,因其网络结构不足以表示多尺度和多样化的特征,感受野范围有限.Res2NeXt-50准确率也仅83.30%,虽然Res2Net增加了网络层感受野的范围,但提升效果不明显.ResNeXt-50引入了变换集合大小来补充深度和宽度之外的维度,使得结果准确率提升到88.02%.尽管ResNeSt-50在不同分支上应用了

通道注意力机制,但其对于脑肿瘤图像细节提取能力不足,准确率只有93.97%。DenseNet-121将每一层与前馈方式下的每一层连接起来,虽然准确率达到96.02%,但是会非常消耗内存与时间。DPN-107结合了DenseNet和ResNeXt的优点,准确率达到97.56%。文献[9]在Se-Net中加入ECA和BAM,达到了99.47%的准确率。但是和现有其他方法一样,该方法针对单点脑肿瘤数据进行深度学习分析,而本文方法利用联邦学习技术建立处理多点数据融合的模型,并且兼具一定的隐私保护功能。

FL-CBAM-DIPC-ResNet的混淆矩阵对比如图8所示。可以看出,模型正确地预测了371个神经胶质瘤、402个脑膜瘤、520个无肿瘤和425个垂体瘤,错误地预测了19个神经胶质瘤、14个脑膜瘤、1个无肿瘤和4个垂体瘤。由于神经胶质瘤属于恶性肿瘤,而垂体瘤是垂体附近的良性肿瘤,因此神经胶质瘤被误分类成垂体瘤的概率小;神经胶质瘤生长快速,脑膜瘤生长缓慢,因此,刚开始生长的神经胶质瘤与脑膜瘤具有较高相似性,此时神经胶质瘤可能被误判成脑膜瘤;由于垂体瘤的病灶区域较小,模型将其误判成脑膜瘤;由于脑膜瘤生长缓慢,因此刚生长发育的脑膜瘤有可能会被误判成无肿瘤,而其他肿瘤即使被误判也能确定有肿瘤,从而最大程度地提醒患者。

真实标签	预测标签			
	神经胶质瘤	脑膜瘤	无肿瘤	垂体瘤
神经胶质瘤	371	17	0	2
脑膜瘤	3	402	5	6
无肿瘤	0	1	520	0
垂体瘤	1	4	0	425

图8 FL-CBAM-DIPC-ResNet混淆矩阵

另外对比FL-CBAM-DIPC-ResNet模型和第3.4.1节中未改进的FL-ResNet模型,分析二者对神经胶质瘤、脑肿瘤、无肿瘤、垂体瘤分类的准确率、精准率、召回率和F1值,各项指标数据如表8,其中第

(1)组为FL-CBAM-DIPC-ResNet模型,第(2)组为FL-ResNet模型。由表8可知,FL-CBAM-DIPC-ResNet对神经胶质瘤、脑肿瘤、无肿瘤、垂体瘤的分类准确性分别为95.13%、96.63%、99.81%和98.84%,而FL-ResNet的分类准确性分别为83.08%、94.23%、99.62%和85.58%。FL-CBAM-DIPC-ResNet对肿瘤的识别率比FL-ResNet分别高出12.05%、2.4%、0.19%、13.26%。相较于FL-ResNet,FL-CBAM-DIPC-ResNet在神经胶质瘤和垂体瘤分类中展现了更显著的识别率提升。

表8 FL-CBAM-DIPC-ResNet(1)与FL-ResNet(2)指标对比(%)

类别	组	指标对比(%)			
		ACC	P	Recall	F1
神经胶质瘤	(1)	95.13	98.95	95.26	97.05
	(2)	83.08	83.83	83.83	88.65
脑膜瘤	(1)	96.63	94.59	96.55	95.55
	(2)	94.23	94.32	94.32	86.24
无肿瘤	(1)	99.81	99.02	99.81	99.41
	(2)	99.62	99.62	99.62	97.45
垂体瘤	(1)	98.84	98.15	98.84	98.49
	(2)	85.58	85.46	85.46	90.20
均值	(1)	97.60	97.68	97.62	97.63
	(2)	90.63	90.81	90.81	90.63

FL-CBAM-DIPC-ResNet在神经胶质瘤、脑肿瘤、无肿瘤、垂体瘤分类下表现优异,其准确率、精准率、召回率与F1值的均值分别达到了97.60%、97.68%、97.62%和97.63%。相较之下,FL-ResNet在相同分类任务上的性能相对较低,其准确率、精准率、召回率和F1值的均值仅为90.63%、90.81%、90.81%和90.63%。FL-CBAM-DIPC-ResNet比FL-ResNet分别高出了6.97%、6.87%、6.81%和7.00%,凸显了其在图像分类任务中的卓越性能和对多类别疾病的高效识别能力。

上述实验结果充分证明,FL-CBAM-DIPC-ResNet相对于FL-ResNet在医学图像分类领域具有显著的优越性,为神经胶质瘤、脑肿瘤、无肿瘤、垂体瘤等多类别疾病的准确分类提供了有力的支持。尽管脑MRI图像因其具有挑战性的细微特征而备受关注,然而FL-CBAM-DIPC-ResNet在面对这些挑战时仍然表现出色,显著超越了FL-ResNet的性能水平。这进一步突显了FL-CBAM-DIPC-ResNet对输入图像细节特征的更加细致和全面的提取能力,为脑肿瘤的精准分类奠定了可靠的基础。这一研究为深度学习在医学影像领域的应用提供了实质性的贡献,拓展了对于脑部疾病自

动诊断的技术边界,为未来医学影像分析领域的研究和应用提供了有益的参考与借鉴。

4 结论与展望

本文提出了一种新的网络模型 FL-CBAM-DIPC-ResNet,建立处理多点脑肿瘤数据的模型,更大程度地保留脑肿瘤的特征,从而提高了脑肿瘤识别的准确率。

该方法的主要创新点是:(1) ReLU 激活函数容易造成神经元死亡,而用 Leaky ReLU 激活函数替换后减轻了神经元死亡造成的影响;(2) 脑肿瘤 MRI 图像细节特征小,原本的 CBAM 在提取特征时通道注意力先降维再升维会损失图像细节特征,而先升维再降维能够较大程度地提取到肿瘤的细节特征;(3) 原本的 CBAM 先计算通道注意力再计算空间注意力,后者难免会受到前者的影响,而将二者并联起来能够最大程度地减轻它们先后顺序的影响;(4) 以往关于脑肿瘤的研究很少关注医疗多点数据融合的问题,而使用联邦学习框架能够打破数据孤岛、建立处理医疗多点数据的模型。

后续将会验证该模型在更多医疗数据集上的泛化能力,并且会针对各个客户端进行个性化算法设计。

参考文献

- 1 Bondy ML, Scheurer ME, Malmer B, *et al.* Brain tumor epidemiology: Consensus from the brain tumor epidemiology consortium. *Cancer*, 2008, 113(S7): 1953–1968. [doi: [10.1002/ncr.23741](https://doi.org/10.1002/ncr.23741)]
- 2 Zhao Z, Chuah JH, Lai KW, *et al.* Conventional machine learning and deep learning in Alzheimer's disease diagnosis using neuroimaging: A review. *Frontiers in Computational Neuroscience*, 2023, 17: 1038636. [doi: [10.3389/fncom.2023.1038636](https://doi.org/10.3389/fncom.2023.1038636)]
- 3 Deepak S, Ameer PM. Brain tumor classification using deep CNN features via transfer learning. *Computers in Biology and Medicine*, 2019, 111: 103345. [doi: [10.1016/j.compbiomed.2019.103345](https://doi.org/10.1016/j.compbiomed.2019.103345)]
- 4 Ghassemi N, Shoeibi A, Rouhani M. Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images. *Biomedical Signal Processing and Control*, 2020, 57: 101678. [doi: [10.1016/j.bspc.2019.101678](https://doi.org/10.1016/j.bspc.2019.101678)]
- 5 Brindha PG, Kavinraj M, Manivasakam P, *et al.* Brain tumor detection from MRI images using deep learning techniques. *IOP Conference Series: Materials Science and Engineering*, 2021, 1055(1): 012115. [doi: [10.1088/1757-899X/1055/1/012115](https://doi.org/10.1088/1757-899X/1055/1/012115)]
- 6 Sultan HH, Salem NM, Al-Atabany W. Multi-classification of brain tumor images using deep neural network. *IEEE Access*, 2019, 7: 69215–69225. [doi: [10.1109/ACCESS.2019.2919122](https://doi.org/10.1109/ACCESS.2019.2919122)]
- 7 夏景明, 邢露萍, 谈玲, 等. 基于 MDM-ResNet 的脑肿瘤分类方法. *南京信息工程大学学报(自然科学版)*, 2022, 14(2): 212–219.
- 8 Özbay E, Özbay FA. Interpretable features fusion with precision MRI images deep hashing for brain tumor detection. *Computer Methods and Programs in Biomedicine*, 2023, 231: 107387. [doi: [10.1016/j.cmpb.2023.107387](https://doi.org/10.1016/j.cmpb.2023.107387)]
- 9 张晓倩, 罗建, 杨梅, 等. 基于改进 SE-Net 网络与多注意力的脑肿瘤分类方法. *西华师范大学学报(自然科学版)*, 2024, 45(1): 93–101.
- 10 Xu QQ, Xu HC, Liu J, *et al.* MR image classification for brain tumor texture based on pseudo-label learning and optimized feature extraction. *Computational and Mathematical Methods in Medicine*, 2022, 2022: 7746991.
- 11 Ijaz A, Raza B, Kiran I, *et al.* Modality specific CBAM-VGGNet model for the classification of breast histopathology images via transfer learning. *IEEE Access*, 2023, 11: 15750–15762. [doi: [10.1109/ACCESS.2023.3245023](https://doi.org/10.1109/ACCESS.2023.3245023)]
- 12 安晨, 汪成亮, 廖超, 等. 基于注意力关系网络的无线胶囊内镜图像分类方法. *计算机工程*, 2021, 47(10): 252–259, 268.
- 13 Adnan M, Kalra S, Cresswell JC, *et al.* Federated learning and differential privacy for medical image analysis. *Scientific Reports*, 2022, 12(1): 1953. [doi: [10.1038/s41598-022-05539-7](https://doi.org/10.1038/s41598-022-05539-7)]
- 14 Chè SC, Kong ZM, Peng H, *et al.* Federated multi-view learning for private medical data integration and analysis. *ACM Transactions on Intelligent Systems and Technology*, 2022, 13(4): 1–23.
- 15 许亚倩, 崔文泉, 程浩洋. 基于联邦学习和重要性加权的疾病得分预测. *计算机系统应用*, 2022, 31(12): 375–382. [doi: [10.15888/j.cnki.csa.008871](https://doi.org/10.15888/j.cnki.csa.008871)]
- 16 段聪颖, 陈思光. 基于联邦深度学习的皮肤病智能诊断研究. *生物信息学*. <http://kns.cnki.net/kcms/detail/23.1513.Q.20230228.1441.002.html>. (在线出版)(2023-02-28).
- 17 曹润芝, 韩斌, 刘嘎琼. 基于联邦学习和改进 ResNet 的肺炎辅助诊断. *计算机系统应用*, 2022, 31(2): 227–233. [doi: [10.15888/j.cnki.csa.008353](https://doi.org/10.15888/j.cnki.csa.008353)]
- 18 Fu JL, Zheng HL, Mei T. Look closer to see better: Recurrent attention convolutional neural network for fine-grained image recognition. *Proceedings of the 2017 IEEE*

- Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 4476–4484.
- 19 Zhu MH, Jiao LC, Liu F, *et al.* Residual spectral-spatial attention network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(1): 449–462. [doi: [10.1109/TGRS.2020.2994057](https://doi.org/10.1109/TGRS.2020.2994057)]
- 20 Li LF, Wang YR. Improved LeNet-5 convolutional neural network traffic sign recognition. *International Core Journal of Engineering*, 2021, 7(4): 114–121.
- 21 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Proceedings of the 26th International Conference on Neural Information Processing Systems. Lake Tahoe: NIPS*, 2012. 1106–1114.
- 22 Tammina S. Transfer learning using VGG-16 with deep convolutional neural network for classifying images. *International Journal of Scientific and Research Publications*, 2019, 9(10): 143–150.
- 23 Yu ZG, Dong YY, Cheng JH, *et al.* Research on face recognition classification based on improved GoogleNet. *Security and Communication Networks*, 2022, 2022: 7192306.
- 24 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE*, 2016. 770–778.
- 25 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. *Proceedings of the 15th European Conference on Computer Vision. Munich: Springer*, 2018. 3–19.
- 26 Yang Q, Liu Y, Chen TJ, *et al.* Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 2019, 10(2): 12.
- 27 Nayef BH, Abdullah SNHS, Sulaiman R, *et al.* Optimized Leaky ReLU for handwritten Arabic character recognition using convolution neural networks. *Multimedia Tools and Applications*, 2022, 81(2): 2065–2094. [doi: [10.1007/s11042-021-11593-6](https://doi.org/10.1007/s11042-021-11593-6)]
- 28 Nickparvar M. Brain tumor MRI dataset. <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>. [2023-04-10].
- 29 Gao SH, Cheng MM, Zhao K, *et al.* Res2Net: A new multi-scale backbone architecture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(2): 652–662. [doi: [10.1109/TPAMI.2019.2938758](https://doi.org/10.1109/TPAMI.2019.2938758)]
- 30 Xie SN, Girshick R, Dollár P, *et al.* Aggregated residual transformations for deep neural networks. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE*, 2017. 5987–5995.
- 31 Zhang H, Wu CR, Zhang ZY, *et al.* ResNeSt: Split-attention networks. *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE*, 2022. 2735–2745.
- 32 Huang G, Liu Z, van der Maaten L, *et al.* Densely connected convolutional networks. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE*, 2017. 2261–2269.
- 33 Chen YP, Li JN, Xiao HX, *et al.* Dual path networks. *Proceedings of the 31st International Neural Information Processing Systems. Long Beach: Curran Associates Inc.*, 2017. 4470–4478.

(校对责编: 孙君艳)