

基于改进MobileViT模型的脑肿瘤图像分类研究

向思宇, 冯慧芳*

西北师范大学数学与统计学院, 甘肃 兰州

收稿日期: 2024年7月30日; 录用日期: 2024年8月27日; 发布日期: 2024年9月4日

摘要

针对目前基于深度学习的脑肿瘤分类算法参数多、计算复杂的问题, 提出了一种基于改进MobileViT的轻量级脑肿瘤图像分类模型。首先, 在轻量化模型MobileViT中加入卷积块注意力模块(CBAM)以有效地增强局部和全局特征。其次, 采用迁移学习方法加快网络模型在脑肿瘤图像上的学习速度, 并在训练过程中使用余弦退火算法来优化所提出的轻量化模型, 使得模型更好地收敛。最后, 在真实脑肿瘤数据集上对本文模型的有效性进行评估, 并与现有的最新基线模型(ResNet、DenseNet121、ShuffleNet、EfficientNet、MobileNet和MobileViT)进行比较。实验结果表明, 相比于基线模型, 本文所提出的模型不仅显著提高了脑肿瘤图像分类的准确性, 而且计算复杂度较低, 符合在边缘计算中部署深度学习模型的需求。

关键词

脑肿瘤, 图像分类, MobileViT, 轻量级, CBAM, 迁移学习

Study on Brain Tumor Image Classification Based on Improved MobileViT Model

Siyu Xiang, Huifang Feng*

College of Mathematics and Statistics, Northwest Normal University, Lanzhou Gansu

Received: Jul. 30th, 2024; accepted: Aug. 27th, 2024; published: Sep. 4th, 2024

Abstract

To address the challenges presented by current deep learning-based brain tumor classification algorithms, which involve numerous parameters and complex computations, we propose a lightweight

*通讯作者。

brain tumor image classification model based on an enhanced version of MobileViT. Firstly, a convolutional block attention module (CBAM) is added to the lightweight model MobileViT to effectively enhance the local and global feature maps. Secondly, a transfer learning approach is used to accelerate the learning speed of the network model on brain tumor images. Additionally, we employ the cosine annealing algorithm to optimize the training process of our proposed lightweight model, facilitating better convergence. Finally, we evaluate the effectiveness of our proposed model on a real brain tumor dataset, comparing it with several state-of-the-art baselines including ResNet, DenseNet121, ShuffleNet, EfficientNet, MobileNet, and MobileViT. The experimental results show that compared to the baseline model, the proposed model in this paper not only significantly improves the accuracy of brain tumor image classification, but also has a lower computational complexity, which meets the requirements of deploying deep learning models in edge computing.

Keywords

Brain Tumor, Image Classification, MobileViT, Lightweight, CBAM, Transfer Learning

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

脑肿瘤是在大脑、脑干、脊髓等中枢神经系统中形成的一类异常组织生长,是当今发病率和死亡率前十的恶性肿瘤之一[1]。脑肿瘤主要类型有:脑膜瘤、垂体瘤和脑胶质瘤。脑肿瘤严重威胁着人类的健康,随着肿瘤的发展,患者可能会出现颅内压升高,从而导致脑损伤,严重者甚至会致命。因此及时发现并准确判断脑肿瘤类型在治疗计划和患者护理中起着重要的作用。

头颅磁共振成像(Magnetic Resonance Imaging, MRI)是一种医学影像学检查技术, MRI 可以在不使用 X 射线的情况下,获取头颅的三维图像,在脑部疾病的诊断中具有重要的作用。脑肿瘤的治疗和预后通常依赖于病理类型和肿瘤分级等关键指标的确定,对脑肿瘤 MRI 图像进行自动分类,可以帮助医生更快速、准确地诊断脑肿瘤,从而提高治疗效果和患者生存率。

目前,对于脑肿瘤 MRI 图像的分类方法大致分为两类,一种是基于深度学习的分类方法,另一种则是基于轻量化模型分类方法。Mengash 等[2]开发了一种基于 CNN 的脑肿瘤分类技术,并使用 k 倍交叉验证法和 Hold-out 测试法去评估模型的性能和泛化能力。El Kader 等[3]提出了一种基于混合深度卷积神经网络和深度分水岭自动编码的脑肿瘤分类模型(CNN-DWA)并在数据集上的分类准确率达 98%。Vankdothu 等[4]将卷积神经网络与长短期记忆网络相结合,提出了一种 CNN-LSTM 网络模型用于 MRI 图像中的脑肿瘤检测,其分类效果优于标准的 CNN 模型。Simo 等[5]基于深度学习和全卷积神经网络设计了一种脑肿瘤检测模型,并探究了四种优化器对模型的优化效果,实验证明所提出的基于 Nesterov 动量优化算法的序列模型在脑肿瘤检测和分类中实现了高精度。Sandhiya 等[6]利用 Inception V3 和 DenseNet201 深度学习架构用于提取基本特征,并使用粒子群优化核极限学习机(Particle swarm optimized kernel Extreme Learning Machine, PSO-KELM)模型对四种脑部 MRI 图像进行分类预测。

虽然基于 CNN 的网络模型在脑肿瘤图像数据集上表现出了出色的分类准确性,但传统的 CNN 架构往往参数量较大,并不利于实际应用。因此,越来越多的研究人员致力于轻量化脑肿瘤分类模型的研究。Lu 等[7]将 MobileNetV2 的特征学习与经典分类器的参数优化相结合,构建了一种具有随机权值的前馈网络并用于脑肿瘤分类。Vaiyapuri 等[8]采用了 EfficientNet、DenseNet 和 MobileNet 的集成模型用于脑肿瘤

图像的特征提取。Luo 等[9]针对 CNN 模型中参数量大和计算复杂度高的问题, 提出了一种多视角提取和密集注意力的轻量级脑肿瘤分割网络。

本文以 MobileViT_xxs [10]网络模型为主干网络, 提出基于改进 MobileViT 模型的脑肿瘤图像分类模型, 在 MobileViT 模型中增加 CBAM 模块和残差连接机制有效地增强局部和全局特征, 并使用迁移学习的方法提升模型的训练效率。本文所提出的模型不仅显著提高了脑肿瘤图像分类的准确性, 而且计算复杂度较低, 符合在边缘计算中部署深度学习模型的需求。

2. 基于改进 MobileViT 的脑肿瘤图像分类模型

2.1. 模型概述

为了通过整合全局信息来提升模型的性能, 我们提出了一种新的基于改进 MobileViT 模型的脑肿瘤图像分类方法, 该方法集成了 MobileViT、CBAM 和迁移学习技术等多种技术。具体的改进包括: (1) 在 MobileViT 模型中添加了 CBAM 模块和残差连接机制; (2) 将原始 MV2 模块中的 ReLU6 激活函数替换为 SiLU 激活函数; (3) 在训练过程中使用余弦退火算法更新学习率并使用迁移学习的方法加速模型收敛。图 1 详细描述了本文提出模型的结构。

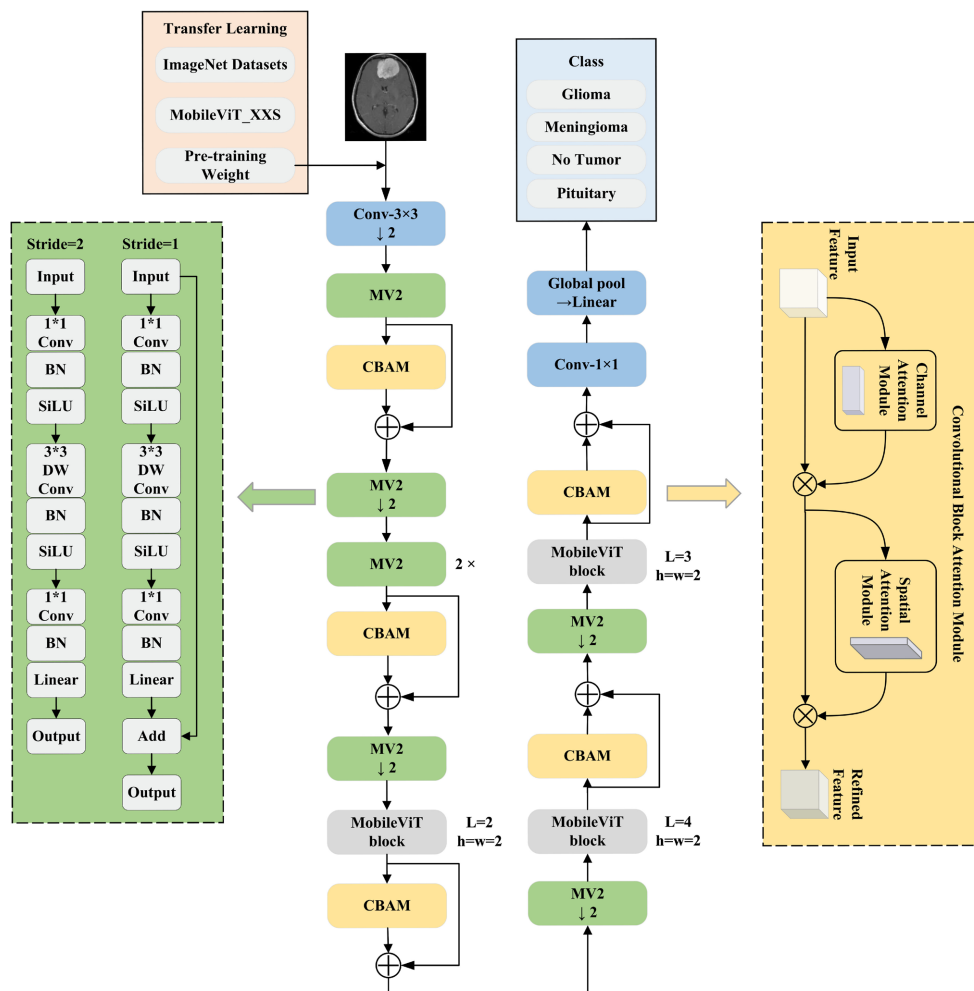


Figure 1. The structure of proposed model
图 1. 模型结构图

2.2. MobileViT Block

MobileViT block 是 MobileViT 模型的核心部分, 其结构如图 2 所示。MobileViT block 的主要计算过程分为以下四个步骤:

1) 将输入矩阵 $X \in R^{H \times W \times C}$ 先通过一个 $n \times n$ 的卷积层以捕捉特征图中的局部空间信息, 再通过一个 1×1 的卷积层将特征图投影到高维特种空间中, 得到具有局部空间信息的 $X_L \in R^{H \times W \times d}$, 其中 H, W, C 表示输入特征矩阵的高、宽和通道数, d 表示经过 1×1 的卷积层调整后的通道数。

2) 将 X_L 转换为一系列不重叠的 patches 表示为 $X_U \in R^{N \times P \times d}$, 其中 $P = wh$, $N = HW/P$, w, h 是 patches 的高和宽。

3) $X_U \in R^{N \times P \times d}$ 通过 Transformer 模块[11]得到具有全局信息的序列 $X_G \in R^{N \times P \times d}$, 其具体计算公式可以表示为:

$$X_G(p) = \text{Transformer}(X_U(p)), 1 \leq p \leq P \tag{1}$$

4) 将 X_G 还原为 $H \times W \times d$ 大小的序列 $X_F \in R^{H \times W \times d}$ 。随后通过 1×1 卷积将 X_F 调整为 $H \times W \times C$ 的大小并与最初的输入特张图 X 进行拼接。最后通过一个 3×3 的卷积层进行特征融合。

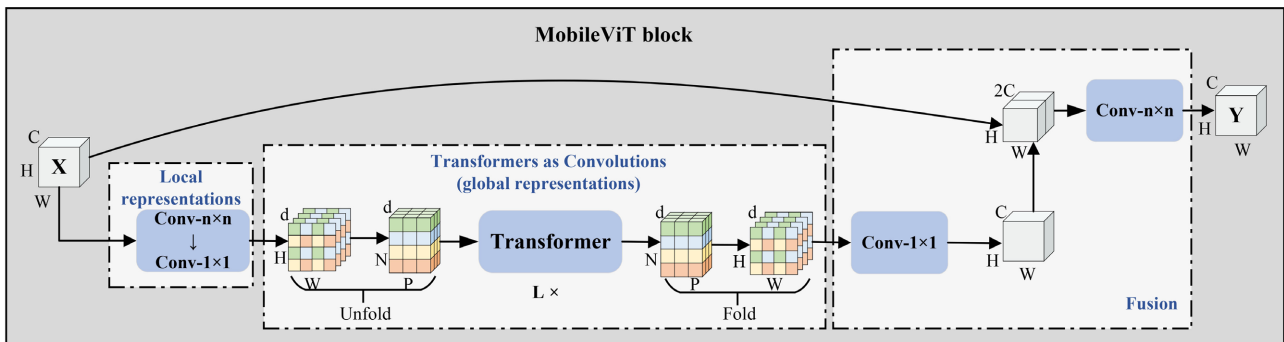


Figure 2. The structure of MobileViT block

图 2. MobileViT block 结构图

2.3. 改进的 MV2 模块

MV2 模块是 MobileNetV2 [12]中的倒残差模块。在 MobileNetV2 中, 传统的卷积层被深度可分离卷积所取代, 它由两个子模块组成: 深度卷积(Depthwise Convolution)和逐点卷积(Pointwise Convolution)。这种技术有效地分解了传统的卷积运算, 在保持表达能力的同时提高了计算效率。本文中的 MV2 模块首先使用 1×1 的卷积进行维度扩张, 然后通过 3×3 的卷积进行特征提取, 最后再经过一个 1×1 的卷积来对特征图进行降维操作。

原 MV2 模块中的激活函数为 ReLU6 激活函数, 是 ReLU 激活函数的变种之一。虽然 ReLU 函数解决了正输入的梯度消失问题, 但它在负输入时遇到了导数恒为 0 的难题, 导致负区间的梯度消失问题, 从而阻碍了许多神经元的更新。因此, 本文使用 SiLU 激活函数代替 ReLU6 激活函数, SiLU 激活函数是 Sigmoid 和 ReLU 的改进版本, 具有无上界有下界、在接近零时更平滑的特点, 在深度网络中的效果优于 ReLU。SiLU 激活函数的计算如公式(2)所示:

$$\text{SiLU}(x) = \frac{x}{1 + e^{-x}}. \tag{2}$$

2.4. 卷积注意力模块

卷积注意力模块(Convolutional Block Attention Module, CBAM) [13]是一个综合了通道注意力机制和空间注意力机制的复合模型, 它由通道注意力模块和空间注意力模块两个子模块组成, 其结构如图 3 所示。

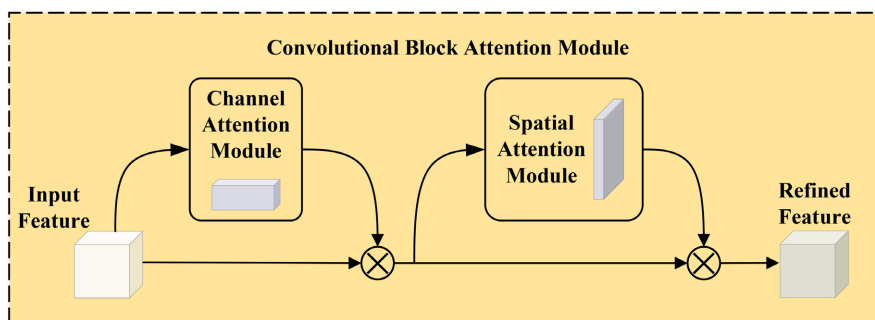


Figure 3. The structure of CBAM
图 3. CBAM 模块结构图

通道注意力模块旨在帮助网络识别各通道的重要性, 具体计算过程如等式(3)、(4)所示。首先, 输入特征分别通过全局平均池化(AvgPool)和全局最大池化(MaxPool)操作提取全局信息。然后, 分别通过多层感知机(MLP)对提取到的全局信息进行处理, 随后将两部分信息相加再通过一个 Sigmoid 激活函数得到每个通道的权重向量 $M_c(F)$ 。最后, 将最初的输入特征 F 与通道权重向量 $M_c(F)$ 相乘, 得到通道注意力特征图 $F1$ 。空间注意力模块则是帮助网络识别不同空间维度的重要性, 同理可根据公式(5)、(6)得到空间注意力特征图 $F2$ 。

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))), \quad (3)$$

$$F1 = F \otimes M_c(F), \quad (4)$$

$$M_s(F) = \sigma(\text{Conv}([\text{AvgPool}(F); \text{MaxPool}(F)])), \quad (5)$$

$$F2 = F1 \otimes M_s(F), \quad (6)$$

其中, $F, F1, F2$ 分别表示输入特征图, 通道注意力特征图和空间注意力特征图。AvgPool(\cdot) 表示全局平均池化, MaxPool(\cdot) 表示全局最大池化。MLP(\cdot) 表示多层感知机, $\sigma(\cdot)$ 是 Sigmoid 激活函数, \otimes 表示逐点乘法, Conv(\cdot) 则表示卷积操作。

2.5. 迁移学习

迁移学习是一种机器学习方法, 旨在将从一个任务或领域学到的知识和经验应用到另一个相关的任务或领域中, 以改善学习效果。迁移学习已广泛应用于自然语言处理、计算机视觉等领域。本文提出的分类模型是为计算机视觉中的图像分类任务而设计的。在本文中, 我们使用 ImageNet 数据集对 MobileViT 进行了预训练, 该数据集包含 1000 个类别和 126 万张自然图像。尽管自然图像与脑肿瘤图像可能不同, 但它们仍然具有相关性。通过迁移学习, 模型能够从 ImageNet 数据集中学习到角、边、颜色、纹理等特征。这些学习到的特征有助于提高模型在脑肿瘤图像分类中的有效性。

2.6. 余弦退火算法

学习速率是影响深度学习模型性能的超参数之一。学习速率过大往往会导致模型的振荡, 而学习速

率过小则会导致非常缓慢的训练过程。余弦退火[14]是一种优化算法, 在训练深度神经网络时, 作为调整学习速率的策略。它基于余弦函数的周期变化, 通过在训练过程中逐渐降低学习率, 帮助模型更好地收敛。采用余弦退火算法, 学习速率在训练过程中会出现周期性的变化。该策略有助于模型在训练早期更好地探索损失函数的空间, 避免在训练后期落入局部最优点。

3. 实验

3.1. 数据集介绍

本文所使用的数据集来自于 Kaggle 数据集, 是 2020 年 7 月由印度杜尔加布尔国家理工学院的 Sartaj Bhuvaji 等人公开的脑肿瘤 MRI 四分类数据集[15]。该数据集一共包含四种类别的脑部 MRI 图像, 其中胶质瘤(glioma tumor) 926 张、脑膜瘤(meningioma tumor) 937 张、无肿瘤(no tumor) 500 张以及垂体瘤(pituitary tumor) 901 张。数据集中的部分图像如图 4 所示。

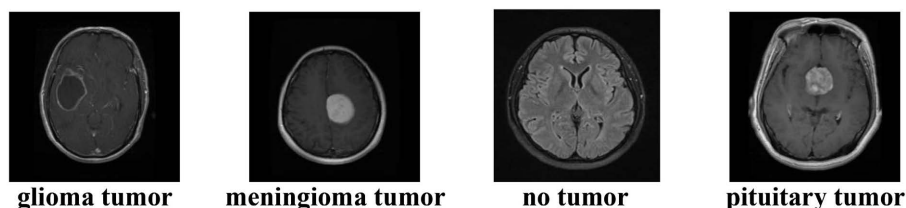


Figure 4. Classes of the dataset
图 4. 数据集类别

3.2. 数据预处理

为便于模型训练, 本文将图像重新分配标签: 胶质瘤标记为 0, 脑膜瘤标记为 1, 无肿瘤标记为 2, 垂体瘤标记为 3。并且将图像随机裁剪到 224×224 像素的分辨率大小, 然后以 0.5 的概率随机水平翻转图像, 以增强模型的鲁棒性。随后将图像转换为张量类型, 并对图像进行标准化处理。标准化后的图像 X_n 可表示为:

$$x_n = \frac{x_i - \mu}{\sigma} \quad (7)$$

其中, X_i 表示图像中每个通道的像素值, μ 表示图像每个通道中像素值的均值, σ 表示图像每个通道中像素值的标准差。将处理好的数据集的 70% 作为训练集, 30% 作为测试集, 其数据分布如表 1 所示。

Table 1. Distribution of categories in the dataset

表 1. 数据集中类别分布

数据集	胶质瘤	脑膜瘤	无肿瘤	垂体瘤	总计
训练集	649	656	350	631	2286
测试集	277	281	150	270	978
总计	926	937	500	901	3264

3.3. 实验环境及参数设置

实验采用 Windows 10 平台进行训练, 并使用 Pytorch 深度学习框架进行网络构建。该实验基于 GPU 型号 NVIDIA GeForce GTX 1080, 15 GB 内存。GPU 数量为 1, CPU 数量为 8 核。实验设置迭代次数为

20, BatchSize 为 30。使用余弦退火算法在训练过程中对学习率进行调整, 其中学习率在每个周期开始的最大值设置为 0.001, 学习率在每个周期结束的最小值设置为 0.00001。对比模型的学习率固定为 0.0001。所有模型均采用了迁移学习的方法提升模型训练性能。

3.4. 评价指标

为评估模型的性能, 本文采用准确率(Accuracy)、精确率(Precision)、召回率(Recall)和 F1 值作为评估指标, 具体的计算规则如公式(8)、(9)、(10)、(11)所示。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (8)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (9)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (10)$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

其中 TP(True Positive)表示原本为正类样本且预测为正类样本的数量, TN(True Negative)表示原本为负类样本且预测为负类样本的数量, FP(False Positive)表示原本为负类样本但预测为正类样本的数量, FN(False Negative)表示原本为正类样本但预测为负类样本的数量。

3.5. 基线网络

为验证本文方法的有效性, 分别使用 ResNet34 [16], ResNet50 [16], DenseNet121 [17], ShuffleNetV2_x2_0 [18], EfficientNet [19], MobileNetV3_large [20], MobileNetV3_small [20], MobileNetV2 [11], MobileViT [10]和改进的 MobileViT 模型对脑肿瘤图像进行分类, 并对各模型的性能结果进行比较。在这些基线中, ResNet 是深度卷积神经网络, DenseNet、ShuffleNet、EfficientNet 和 MobileNet 是轻量级深度学习模型。

4. 实验结果与讨论

4.1. 对比实验

本文采用了 10 种不同的深度学习模型对脑肿瘤图像数据集进行分类。表 2 为多种模型的实验结果, 从表 2 中可以得出以下结论:

- 1) 本文提出模型在脑肿瘤图像数据集上的分类准确率为 97.0%, 精确率为 97.1%, 召回率为 97.0%, F1 值为 97.0%。经过比较, 本文所提出模型的所有性能指标都是最佳的, 甚至分类准确率比 ResNet50 高出 0.4%。
- 2) 在基线模型中, ResNet 系列网络的分类性能明显优于其他轻量级深度学习模型(DenseNet, ShuffleNet, EfficientNet, MobileNet 和 MobileViT), 这归因于 ResNet 作为深度残差网络, 在图像分类、目标检测和语义分割等许多计算机视觉任务中都表现出了良好的性能。
- 3) 在基线模型中, MobileViT 是轻量级模型中表现最佳的模型, 其分类准确率为 90.6%, 精确率为 90.8%, 召回率为 90.6%, F1 值为 90.6%。然而与我们提出的模型相比, MobileViT 的四个指标都低了近 7%, 这表明我们提出的模型优于 MobileViT。

Table 2. Classification results for each model
表 2. 各模型的分类结果

模型	准确率	精确率	召回率	F_1	参数量	FLOPs
ResNet50	0.966	0.967	0.966	0.966	23,516,228	4.09826048G
ResNet34	0.963	0.964	0.963	0.963	21,286,724	3.66699008G
DenseNet121	0.892	0.891	0.892	0.891	6,957,956	2.848985856G
ShuffleNetV2_x2_0	0.793	0.801	0.793	0.795	5,353,192	0.584959696G
EfficientNet	0.771	0.777	0.771	0.770	4,012,672	0.393804448G
MobileNetV3_large	0.877	0.877	0.877	0.876	4,207,156	0.220317112G
MobileNetV3_small	0.843	0.844	0.843	0.841	1,521,956	0.057191512G
MobileNetV2	0.773	0.776	0.773	0.774	2,228,996	0.306178784G
MobileViT	0.906	0.908	0.906	0.906	952,308	0.304916367G
本文提出模型	0.970	0.971	0.970	0.970	955,542	0.306991777G

计算复杂度也是深度学习模型的一个重要度量指标,特别是对于部署在边缘设备上的深度学习模型,在本文中我们使用参数量和 FLOPs (每秒浮点运算次数)来评估各模型的计算复杂度。从表 2 中可以明显看出,尽管 ResNet50 实现了 96.6%的高分类准确率,但它的参数量为 23,516,228 大约是参数最多的轻量级网络 DenseNet121 的 3.38 倍,是参数最少的轻量级网络 MobileViT 的 26.69 倍。从表 2 中还可以看到,本文所提出模型的参数量为 955542,仅次于 MobileViT 模型。本文模型的 FLOPs 比 MobileNet 和 MobileViT 等轻量级模型略大,但仅比 MobileViT 高 0.65%。虽然我们提出的模型复杂度略高于 MobileViT,但它仍然属于轻量级网络,满足实时或部署在边缘设备上的要求,也具有相对较高的分类精度。

4.2. 消融实验

为验证本文提出的模型的有效性,我们进行了消融实验,以验证改进模块对模型分类性能的影响。实验结果如表 3 所示,对 MV2 模块进行改进后模型分类准确率提高了 5.2%,其他性能指标也得到了显著的改进。再相继加入余弦退火算法和 CBAM 模块,这两种改进方法都在不同程度上提升了模型分类效果。

Table 3. Ablation experiment
表 3. 消融实验

模型	准确率	精确率	召回率	F_1
MobileViT	0.906	0.908	0.906	0.906
MobileViT + Improved MV2	0.958	0.958	0.958	0.958
MobileViT + Improved MV2 + Cosine Annealing	0.966	0.966	0.966	0.966
MobileViT + Improved MV2 + Cosine Annealing + CBAM	0.970	0.971	0.970	0.970

5. 结论

为了使脑肿瘤图像分类模型能够应用于移动设备和嵌入式设备,我们提出了一种基于改进的 MobileViT 的轻量级图像分类模型。首先,利用 CBAM 模块来提高模型捕获局部信息和融合全局信息的能力。其次,将 MV2 模块中的激活函数 ReLU6 替换为更稳健的 SiLU 激活函数。第三,采用余弦退火算

法来更新模型的学习率, 以避免模型陷入局部最优点。最后, 采用了迁移学习方法, 将在 ImageNet 数据集上预训练好的权重应用于模型中, 以提升模型学习速率。大量的实验表明, 我们提出的模型在脑肿瘤图像分类方面具有良好的竞争性能。我们开发的低成本智能诊断工具, 不仅帮助医学专家和放射技师提供脑疾病的早期诊断, 而且适合部署在边缘设备。

在未来的研究中, 我们将考虑在更多更大的数据集上进行实验研究。目前我们只研究了三种脑肿瘤, 计划通过增加更多的脑肿瘤类别来验证我们的模型的有效性。此外, 本文提出的模型是一个轻量级的深度学习模型, 虽然这个模型的分类精度高, 计算复杂性不是最小的, 所以在不降低分类精度的情况下进一步减少计算时间是我们另一个潜在的研究方向。

基金项目

国家自然科学基金资助项目(72461030)。

参考文献

- [1] Louis, D.N., Perry, A., Wesseling, P., *et al.* (2021) The 2021 WHO Classification of Tumors of the Central Nervous System: A Summary. *Neuro-Oncology*, **23**, 1231-1251. <https://doi.org/10.1093/neuonc/noab106>
- [2] Mengash, H.A. and Mahmoud, H.A.H. (2021) Brain Cancer Tumor Classification from Motion-Corrected MRI Images Using Convolutional Neural Network. *Computers, Materials & Continua*, **68**, 1551-1563. <https://doi.org/10.32604/cmc.2021.016907>
- [3] El Kader, I.A., Xu, G., Shuai, Z., *et al.* (2021) Brain Tumor Detection and Classification by Hybrid CNN-DWA Model Using MR Images. *Current Medical Imaging*, **17**, 1248-1255. <https://doi.org/10.2174/1573405617666210224113315>
- [4] Vankdothu, R., Hameed, M.A. and Fatima, H. (2022) A Brain Tumor Identification and Classification Using Deep Learning Based on CNN-LSTM Method. *Computers and Electrical Engineering*, **101**, Article 107960. <https://doi.org/10.1016/j.compeleceng.2022.107960>
- [5] Simo, A.M.D., Kouanou, A.T., Monthe, V., *et al.* (2024) Introducing a Deep Learning Method for Brain Tumor Classification Using MRI Data towards Better Performance. *Informatics in Medicine Unlocked*, **44**, Article 101423. <https://doi.org/10.1016/j.imu.2023.101423>
- [6] Sandhiya, B. and Raja, S.K.S. (2024) Deep Learning and Optimized Learning Machine for Brain Tumor Classification. *Biomedical Signal Processing and Control*, **89**, Article 105778. <https://doi.org/10.1016/j.bspc.2023.105778>
- [7] Lu, S.Y., Wang, S.H. and Zhang, Y.D. (2020) A Classification Method for Brain MRI via MobileNet and Feedforward Network with Random Weights. *Pattern Recognition Letters*, **140**, 252-260. <https://doi.org/10.1016/j.patrec.2020.10.017>
- [8] Vaiyapuri, T., Jaiganesh, M., Ahmad, S., *et al.* (2023) Ensemble Learning Driven Computer-Aided Diagnosis Model for Brain Tumor Classification on Magnetic Resonance Imaging. *IEEE Access*, **11**, 91398-91406. <https://doi.org/10.1109/ACCESS.2023.3306961>
- [9] Luo, H., Zhou, D., Cheng, Y., *et al.* (2024) MPEDA-Net: A Lightweight Brain Tumor Segmentation Network Using Multi-Perspective Extraction and Dense Attention. *Biomedical Signal Processing and Control*, **91**, Article 106054. <https://doi.org/10.1016/j.bspc.2024.106054>
- [10] Mehta, S. and Rastegari, M. (2021) Mobilevit: Light-Weight, General-Purpose, and Mobile-Friendly Vision Transformer.
- [11] Dosovitskiy, A., Beyer, L., Kolesnikov, A., *et al.* (2020) An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale.
- [12] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L. (2018) Mobilenetv2: Inverted Residuals and Linear Bottlenecks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 4510-4520. <https://doi.org/10.1109/cvpr.2018.00474>
- [13] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [14] Loshchilov, I. and Hutter, F. (2016) SGDR: Stochastic Gradient Descent with Warm Restarts.
- [15] Muezzinoglu, T., Baygin, N., Tuncer, I., Barua, P.D., Baygin, M., Dogan, S., *et al.* (2023) Patchresnet: Multiple Patch Division-Based Deep Feature Fusion Framework for Brain Tumor Classification Using MRI Images. *Journal of Digital Imaging*, **36**, 973-987. <https://doi.org/10.1007/s10278-023-00789-x>

- [16] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/cvpr.2016.90>
- [17] Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017) Densely Connected Convolutional Networks. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 2261-2269. <https://doi.org/10.1109/cvpr.2017.243>
- [18] Zhang, X., Zhou, X., Lin, M. and Sun, J. (2018) Shufflenet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 6848-6856. <https://doi.org/10.1109/cvpr.2018.00716>
- [19] Tan, M. and Le, Q.V. (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning, ICML 2019*, Long Beach, 9-15 June 2019, 6105-6114.
- [20] Howard, A., Sandler, M., Chen, B., Wang, W., Chen, L., Tan, M., *et al.* (2019) Searching for MobileNetV3. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 1314-1324. <https://doi.org/10.1109/iccv.2019.00140>