

融合改进 CBAM 机制和 ResNet 网络的肺炎 CT 图像分类研究

罗声平¹

LUO Shengping

摘要

为快速准确地对正常肺部、普通肺炎、新冠病毒肺炎 CT 图像进行识别分类,提出了一种融合改进的卷积块注意力模块(convolutional block attention module, CBAM)机制和 ResNet18 网络的新模型。对 CBAM 中的多层感知机(multilayer perceptron, MLP)进行升维改进,放大肺部 CT 图像关键特征;以 ResNet18 作为基础模型,将改进的 CBAM 机制融入 ResNet 模块中,以加强对关键细节特征的提取,并将 AlphaDropout 和 SeLU 激活函数融入网络中,防止其网络的过拟合化,加速模型收敛效果。通过混淆矩阵计算得出模型的准确率、精确率、召回率、F1 分数分别达到了 99.33%、99.34%、99.33% 和 0.984 5,相比改进前的 ResNet18 模型分别提高了 4.23%、4.88%、4.20%、0.042,且均高于 GoogLeNet、ResNet50 和 Xception 对照模型。研究结果表明,改进的 CBAM-ResNet18 模型对肺部 CT 图像具有良好的识别结果。

关键词

肺炎 CT 图像;残差网络;卷积块注意力模块;多层感知机;激活函数

doi: 10.3969/j.issn.1672-9528.2024.04.010

0 引言

采用逆转录聚合酶链反应(RT-PCR)方法区分普通肺炎和新冠病毒肺炎,耗时较多,假阴性率较高^[1],基于卷积神经网络的 CT 图像识别方法则可有效地开展肺部医学评估^[2]。针对肺结节检测中的大量假阳性问题,冯雨等人^[3]提出了一种基于三维卷积神经网络的肺结节识别方法。马圆等人^[4]采用深度信念网络方法对良恶性肺结节的 CT 图像开展分类研究。Polsinelli M 等人^[5]提出轻量级肺部 CT 图像识别网络模型 SqueezeNet,其准确率、敏感性、特异性和 F1 分数分别达到了 83%、85%、81% 和 0.833。

正常肺部、普通肺炎、新冠病毒肺炎三类 CT 图像特征差异性较小,对医学知识要求较高,高精度标注较为困难^[6-7]。针对高精度标注数据集缺乏,现有网络模型无法高效提取肺部 CT 图像特征,导致模型学习效率低、收敛不稳定的问题,本研究提出了改进的 CBAM-ResNet18 模型。该模型将 CBAM 机制中通道注意力的多层感知机 MLP 改进为先升维再降维 MLP,并将改进的 CBAM 机制融入 ResNet18 模型中,结合 AlphaDropout 正则化和 SeLU 激活函数,更好地避免过拟合现象,同时提高收敛速度和模型性能,实现了对正常肺部、普通肺炎、新冠病毒肺炎 CT 图像的精准识别。

1. 南昌师范学院物理与电子信息学院 江西南昌 330032

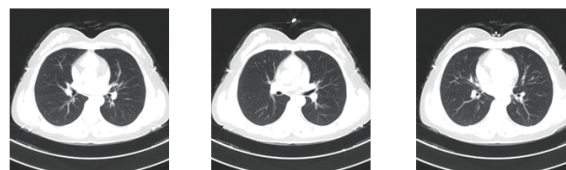
[基金项目]江西省教育厅自然科学研究项目(GJJ2202030);

南昌师范学院自然科学研究项目(22XJZR02)

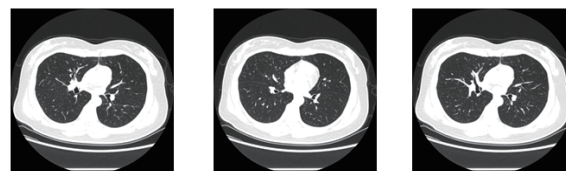
1 数据集及数据预处理

1.1 数据集

本研究依托国家生物信息中心的病毒信息库数据集,构建了 3 类 CT 图像数据,分别是正常肺部 CT 图像(Normal)、普通肺炎 CT 图像(CP)、新冠病毒肺炎(COVID-19) CT 图像(NCP),数据集示例如图 1 所示。



(a) 正常肺部 CT 图像



(b) 普通肺炎 CT 图像



(c) 新冠病毒肺炎(COVID-19) CT 图像

图 1 肺炎 CT 图像数据集示例

1.2 数据预处理

1.2.1 数据增强

为提高模型的泛化能力，同时避免过拟合现象的发生，可以采用数据增强方法^[8]。本研究采用增加随机噪声、镜像、旋转(90°、180°及270°)和亮度变化四类数据增强方式，分别得到正常肺部CT图像(Normal)3500张、普通肺炎CT图像(CP)3500张、新冠病毒肺炎(COVID-19)CT图像(NCP)3500张，共计10500张。将构建的实验数据集按照8:2的比例随机分为训练集和测试集，分布情况如表1所示。

表1 数据集分布情况(CT切片)

	Normal	CP	NCP	Total
训练集	2800	2800	2800	8400
测试集	700	700	700	2100

1.2.2 CT图像标注

本研究将每类肺部CT图片归为同一类型样本，只进行类别标注，无需像素级标注，标注过程相对来说难度较低。

2 改进的CBAM-ResNet18识别模型

2.1 ResNet网络模型

何恺明等人^[9]提出了基于跳跃连接和恒等映射的残差网络(residual network, ResNet)，核心思想是使用残差块(residual block)来搭建网络，很大程度上避免了梯度消失和梯度爆炸的发生。由于医疗诊断需要快速准确地识别肺部CT图像，对实时性要求较高，模型不宜过于复杂，故本研究将ResNet18作为基础网络模型。

2.2 CBAM注意力机制及其改进

CBAM^[10]引入了两种注意力机制：通道注意力机制CAM^[11]和空间注意力机制SAM^[12]。CAM对特征图分别进行平均池化和最大池化，降维至 $1 \times 1 \times C$ ，经过多层感知机MLP后维度降为 $1 \times 1 \times C/r$ 。随后，将输出的两个特征图进行相加和计算，再采用Sigmoid函数生成通道注意力权重 M_c ，其公式为：

$$M_c(F) = \sigma(MLP(MaxPool(F)) + MLP(AvgPool(F))) \quad (1)$$

式中： F 为输入特征图； σ 为Sigmoid激活函数； $AvgPool$ 与 $MaxPool$ 分别为平均池化与最大池化。

SAM则强化了空间关键特征信息的提取，弥补了通道注意力机制的不足，空间注意力权重 M_s 计算公式为：

$$M_s(F) = \sigma(Conv_{3 \times 3}(MaxPool(F); AvgPool(F))) \quad (2)$$

式中： $Conv_{3 \times 3}$ 为 3×3 卷积操作。

传统MLP将特征图降维至 $1 \times 1 \times C/r$ 后再升维至 $1 \times 1 \times C$ ，降维在一定程度上会造成关键信息的丢失，导致 M_c 对细节特征提取不足。准确快速地识别肺部CT图像对细节特征提取要求较高，本研究将CAM中的MLP进行改进，

如图2所示。改进后的MLP采取先升维再降维的处理思路，先分别将平均池化和全局池化后的 $1 \times 1 \times C$ 特征图升维至 $1 \times 1 \times rC$ ，再降维至 $1 \times 1 \times C$ ，提升CBAM机制对肺部CT图像的细节特征提取能力。

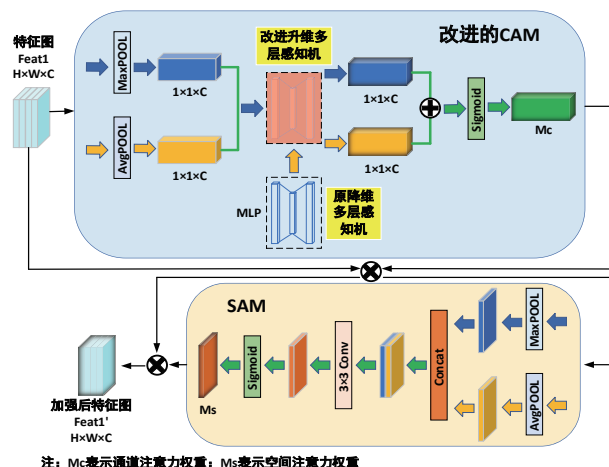


图2 改进的CBAM注意力机制

2.3 AlphaDropout正则化和SeLU激活函数

采用Dropout正则化让一部分神经元失活，可以提高模型的泛化能力，但采用Dropout后，每次迭代只有一部分参数更新，可能导致梯度下降变慢，减缓模型的收敛速度。Klambauer等人^[13]提出了AlphaDropout，它是一种保持输入均值和方差不变的Dropout，能保持数据的自规范性。

SeLU激活函数在输入为负数时具有指数增长，而在输入为正数时近似于线性，如公式(3)所示。SeLU引入了自标准化机制，使神经网络的输出在训练过程中保持均值和方差接近于1，能在一定程度上避免神经网络中的梯度爆炸和梯度消失问题。

$$SeLU(x) = \lambda \begin{cases} x & x > 0 \\ \alpha e^x - \alpha & x \leq 0 \end{cases} \quad (3)$$

式中： λ 约为1.0507， α 约为1.67326。

本研究提出的网络模型结合AlphaDropout正则化和SeLU激活函数，可以减少过拟合现象的发生，同时提升收敛速度和识别性能。

2.4 改进后的CBAM机制与ResNet网络的融合

综上所述，本研究将改进的CBAM机制融入ResNet18网络中，更好地提取肺部CT图像关键特征，过滤无关信息。改进的肺部CT图像识别模型CBAM-ResNet18如图3所示。

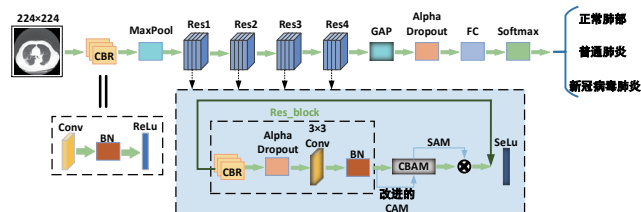


图3 改进的肺部CT图像识别模型CBAM-ResNet18

图3中,输入为224×224大小的CT图像,进入卷积层和池化层后,得到56×56大小的特征图,随后经过4个ResNet模块继续提取关键信息^[14]。进入ResNet模块后,在改进的CAM中,特征图先升维,再降维,获得通道注意力权重 M_c 。4个ResNet模块采用数量分别为64、128、256和512的3×3卷积核,特征图大小分别为56×56、28×28、14×14和7×7。

本研究结合AlphaDropout正则化和SeLU激活函数,输入的分布保持不变,减少过拟合的发生。最后,采用Softmax函数将原始的线性得分转换为概率分布,实现肺部CT图像的精准分类,其公式为:

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_i^N e^{x_i}}, i \in \{i, \dots, N\} \quad (4)$$

式中: $\sum_i^N e^{x_i}$ 为所有输入信号 x_i 的指数函数和。

3 实验结果与分析

3.1 训练环境

本研究基于Windows10操作系统和PyTorch搭建了CBAM-ResNet18训练平台,Python版本为3.8.13,PyTorch版本为1.12.1。模型训练采用的处理器参数为Intel(R)Core(TM) i7-9750H CPU@ 2.60 GHz 2.59 GHz, GPU为NVIDIA GeForce GTX 1650,显存为16 GB,内存为24 GB。

在训练阶段,输入CT图像大小为224×224; batch size设置为16, epochs设置为80轮,采用随机梯度下降算法SGD进行模型优化,采用交叉熵损失函数计算损失值。为加速模型收敛,初始学习率设置为 10^{-3} ,当学习率衰减超过3次后,权重衰减为0。

3.2 CBAM-ResNet18 模型特征可视化

将CBAM-ResNet18模型指定层的特征图进行可视化,可以直观理解每层网络提取出的特征信息。特征图可视化需要对特征图进行通道分离,将tensor展开后以图片的形式保存。为直观展示CBAM-ResNet18模型的整体性及网络层特点,将第5网络层输出的64通道特征图[56,56,64]和第9网络层输出的128通道特征图[28,28,128]进行可视化,如图5所示。

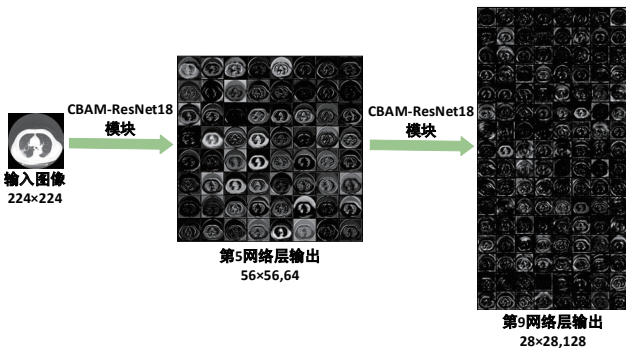


图4 改进的CBAM-ResNet18模型特征可视化

3.3 CBAM-ResNet18 模型性能评估结果

本研究为三分类任务,采用改进的CBAM-ResNet18模型对3种肺部CT图像测试集进行分类,得到混淆矩阵如图5所示。由图5可知,本研究所采用的改进CBAM-ResNet18模型针对2100个肺部CT图像样本,正确识别数为2086,误识别数为14,其中700个新冠病毒肺炎CT图像全部正确识别,总体误识别数较少,能够实现肺炎CT图像的精准确识别。

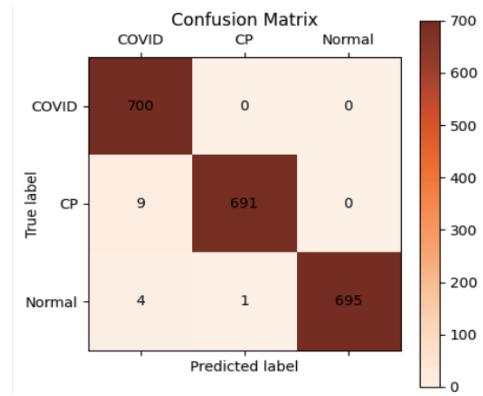


图5 肺部CT图像测试集的混淆矩阵

评价模型性能的指标通常有准确率(Accuracy)、精确率(Precision)、召回率(Recall)、 F_1 分数(F_1 Score),其公式为:

$$Accuracy = \frac{T_p + T_N}{T_p + T_N + F_p + F_N} \quad (5)$$

$$Precision = \frac{T_p}{T_p + F_p} \quad (6)$$

$$Recall = \frac{T_p}{T_p + F_N} \quad (7)$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (8)$$

式中: T_p 和 T_N 表示肺部CT图像样本被正确识别的数量, F_p 和 F_N 表示被错误识别的数量。为了更好地评估模型性能,本研究采用了对比实验,在数据预处理方法、训练环境、参数设置完全一致的前提下,与其它4类经典网络模型(包括GoogLeNet、ResNet18、ResNet50和Xception)进行指标对比,得到3类肺部CT图像识别性能表现,如表2所示。由表2可知,改进的CBAM-ResNet18模型在准确率、精确率、召回率和 F_1 分数上均有较好的提升。

表2 不同模型的肺炎CT图像识别性能表现

网络模型	准确率 (Accuracy)	精确率 (Precision)	召回率 (Recall)	F_1 分数 (F_1 Score)
GoogLeNet	93.12%	92.48%	93.26%	0.926 8
ResNet18	95.10%	94.46%	95.13%	0.951 5
ResNet50	96.70%	95.82%	96.55%	0.962 4
Xception	96.78%	96.08%	96.65%	0.963 2
CBAM-ResNet18	99.33%	99.34%	99.33%	0.993 3

为了进一步评估本研究所述模型的性能,绘制出3类肺部CT图像样本的ROC曲线,如图7所示,效果良好,进一步计算AUC面积,正常肺部、普通肺炎、新冠病毒肺炎CT图像测试集对应的值分别为0.999 95、0.999 94、1.000 00,均非常接近或等于1,证明该模型性能较为优越。

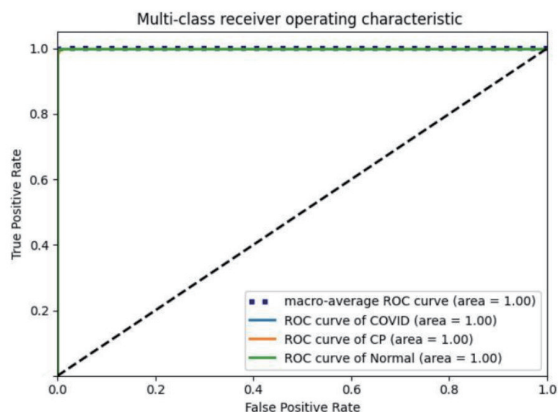


图6 肺部CT图像测试集的ROCROCROC曲

4 结语

针对高精度标注数据集缺乏且现有网络模型无法高效提取肺部CT图像细节信息,模型性能指标偏低的问题,本研究提出了一种改进的肺部CT图像识别模型CBAM-ResNet18,整体识别性能表现得到提升,主要结论如下。

(1) 将CBAM机制中的MLP进行改进,从传统的先降维后升维改为先升维后降维,放大了肺部CT图像的关键特征,提升了特征提取效率;以ResNet18为基础模型,将改进的CBAM机制融入ResNet残差块中,结合AlphaDropout正则化和SeLU激活函数,提升了模型的整体性能。

(2) 通过对混淆矩阵计算分析,改进的CBAM-ResNet18模型对3类肺部CT图像识别的准确率、精确率、召回率、F1分数分别达到了99.33%、99.34%、99.33%和0.984 5,性能评估结果较好。

(3) 改进的CBAM-ResNet18模型能够更好地提取肺部CT图像的有效信息,实现学习效率和识别准确率的提升,但本模型的大小与参数量仍有一定的完善空间。

参考文献:

- [1] FANG Y, ZHANG H, XIE J, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR[J]. Radiology, 2020, 296(2): E115-E117.
- [2] 郭保苏, 庄集超, 吴凤和, 等. 基于CT图像卷积神经网络处理的新冠肺炎检测[J]. 计量学报, 2021, 42(4): 537-544.
- [3] 冯雨, 易本顺, 吴晨玥, 等. 基于三维卷积神经网络的肺结节识别研究[J]. 光学学报, 2019, 39(6): 256-261.

- [4] 马圆, 王凤, 韩勇, 等. 基于深度信念网络检测PET/CT图像肺结节良恶性[J]. 中国医学影像技术, 2020, 36(1): 77-80.
- [5] POLSINELLI M, CINQUE L, PLACIDI G. A light CNN for detecting COVID-19 from CT scans of the chest[J]. Pattern recognition letters, 2020, 140: 95-100.
- [6] 任冬伟, 王旗龙, 魏云超, 等. 视觉弱监督学习研究进展[J]. 中国图象图形学报, 2022, 27(6): 1768-1798.
- [7] MEI S A, YANG H A, YIN Z P. An unsupervised-learning based approach for automated defect inspection on textured surfaces[J]. IEEE transactions on instrumentation and measurement, 2018, 67(6): 1266-1277.
- [8] 谢未央, 陈彦博, 王季勇, 等. 基于卷积神经网络的CT图像肺结节检测[J]. 计算机工程与设计, 2019, 40(12): 3575-3581.
- [9] MUHAMMAD S, ZHAOQUAN G. Deep residual learning for image recognition: a survey [J]. Applied sciences, 2022, 12(18): 8972-8972.
- [10] YANG X, ZHANG Q, WANG S, et al. Detection of solar panel defects based on separable convolution and convolutional block attention module[J]. Energy sources, part A: recovery, utilization, and environmental effects, 2023, 45(3): 7136-7149.
- [11] FU J, ZHENG H, MEI T. Look closer to see better: recurrent attention convolutional neural network for finegrained image recognition[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 4476-4484.
- [12] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 7132-7141.
- [13] YAO L, STEPHEN G, THALAIYASINGAM A. Bidirectionally self-normalizing neural networks[J]. Neural networks: the official journal of the international neural network society, 2023, 167: 283-291.
- [14] 张文景, 蒋泽中, 秦立峰. 基于弱监督下改进的CBAM-ResNet18模型识别苹果多种叶部病害[J]. 智慧农业(中英文), 2023, 5(1): 111-121.

【作者简介】

罗声平(1990—),男,江西吉安人,硕士,讲师,高级工程师,研究方向:信号与信息处理、模式识别。

(收稿日期:2024-01-24)