

改进残差网络的猴痘病毒皮肤病变分类^①



胡莹晖, 杜滨媛, 胡 成, 刘兴云

(湖北师范大学 物理与电子科学学院, 黄石 435002)

通信作者: 刘兴云, E-mail: liuxingyun@hbnu.edu.cn

摘 要: 目前猴痘病毒在全球范围内传播, 这种病毒在临床上与其他皮肤疾病难以区分, 特别是天花病毒和水痘病毒. 在确定性聚合酶链式反应技术和其他生物检测技术还没有完全成熟的情况下, 通过计算机辅助诊断技术检测猴痘病毒皮肤病变是一种可行的方法, 因此提出了一种基于残差网络的猴痘病毒皮肤病变分类算法. 该算法以残差网络为基本框架, 结合深度可分离卷积和轻量化注意力, 在降低模型计算量与复杂度的同时, 也提高了模型的性能. 实验结果表明, 该算法对猴痘病毒皮肤病变表现出较好的分类性能, 对猴痘皮肤病变的分类准确率、召回率和精度分别为 97.3%, 96.8% 和 97.2%, 且均优于实验中所对比的常见分类模型和其他研究方法.

关键词: 猴痘病毒; 计算机辅助诊断技术; 残差网络; 深度可分离卷积; 轻量化注意力

引用格式: 胡莹晖, 杜滨媛, 胡成, 刘兴云. 改进残差网络的猴痘病毒皮肤病变分类. 计算机系统应用, 2023, 32(6): 197-203. <http://www.c-s-a.org.cn/1003-3254/9151.html>

Classification of Monkeypox Virus Skin Lesions Based on Improved ResNet

HU Ying-Hui, DU Bin-Yuan, HU Cheng, LIU Xing-Yun

(School of Physics and Electronic Science, Hubei Normal University, Huangshi 435002, China)

Abstract: Monkeypox virus is currently circulating globally and is clinically indistinguishable from other skin diseases, particularly the smallpox virus and chickenpox virus. In the case that deterministic polymerase chain reaction technology and other biological detection technologies are not fully mature, it is a feasible method to detect skin lesions caused by the monkeypox virus by computer-aided diagnostic technology, so a classification algorithm for skin lesions caused by the monkeypox virus based on the residual network is proposed. Based on the residual network, the algorithm combines deep separable convolution and lightweight attention, which reduces the computational amount and complexity of the model and improves the classification performance of the model. The experimental results show that the algorithm shows excellent classification performance for skin lesions caused by the monkeypox virus, and the classification accuracy, recall, and precision of skin lesions caused by the monkeypox virus are 97.3%, 96.8%, and 97.2%, respectively, which are better than those of the common classification models and other research methods used in the experiment.

Key words: monkeypox virus; computer-aided diagnostic (CAD) techniques; residual network (ResNet); depth separable convolution; lightweight attention

猴痘病毒是在 1958 年发现并分离的, 当时在食蟹猴群体中爆发了两次类似于天花病毒的疾病. 在疾病爆发期之前没有发现临床症状, 患病的特征以斑丘疹为主, 由于它与其他已知的痘类病毒十分相似, 因此将

这种病毒命名为猴痘病毒^[1]. 在病毒出现早期常以动物传播为主, 很少在人类身上发现病毒, 因此人类对猴痘病毒的关注较少. 然而, 2022 年 5 月在欧洲发现了一系列人类感染猴痘的病例, 使猴痘病毒波及的范围越来越

① 收稿时间: 2022-12-09; 修改时间: 2023-01-17; 采用时间: 2023-02-03; csa 在线出版时间: 2023-04-25
CNKI 网络首发时间: 2023-04-26

越广泛,猴痘病毒已不容忽视。

猴痘病毒与天花和水痘病毒十分相似,但在临床症状中也有不少的区别。一般猴痘病毒病程分为3个阶段:潜伏期、发热期、特异性皮疹期^[2]。在发热期,约90%的病例都出现有颈部、耳后、腋窝等处淋巴结肿大,患者淋巴结肿大是区别于其他天花病毒最显著的体征^[3]。而这些症状都仅限于典型临床表现的患者,当临床表现不典型时,病毒诊断的可靠性和真实性则会大大降低^[4]。Jezek等人^[5]研究了977位皮疹患者,通过血清学和水疱性液体电子显微镜观察和培养后,在7.3%的非典型水痘患者和6.1%的不明原因皮疹患者中发现了猴痘。在诊断为水痘的730例病例中,有3.3%患有猴痘。这表明将猴痘病毒与其拥有相似皮肤病变体征的其他病毒区分是一件具有挑战的任务。

计算机辅助诊断(CAD)技术是通过机器学习方法来分析成像或非成像患者数据,评估患者病情和帮助临床医生进行决策的一项新兴技术^[6]。随着深度学习的出现,同时也成功推动了计算机辅助诊断技术的研究和开发。利用深度学习进行CAD在医学领域十分常见,应用包括疾病和正常模式的分类、恶性和良性病变的分类以及预测未来癌症的高风险和低风险模式的分类^[7]。Allugunti^[8]使用著名的黑色素瘤数据集构建并评估了用于CAD诊断皮肤癌症的卷积神经网络模型,其CNN分类器的模块化和分层结构显著减少了计算工作量同时优于其他先进的机器学习技术。Zuluaga-Gomez等人^[9]提出了一种基于卷积神经网络的计算机辅助诊断系统,作为热图像诊断乳腺癌的替代诊断方法。实验结果表明,该系统可以获得较低的假阳性和假阴性分类率。Saygili^[10]提出了一种利用胸部CT和X线图像对新冠肺炎患者进行分类的模型,该模型可以在不到一分钟的时间内以较高的成功率检测到新冠肺炎。创建了一个决策支持系统将有助于放射科医生的决策,有助于病毒的早期诊断,并能将新冠肺炎病毒引起的肺炎与其他疾病引起的肺炎区分开来。计算机辅助诊断技术或许可以成为临床医生决策过程中有效的支持信息。

深度卷积神经网络是最常用的深度学习网络,它为图像分类带来了一系列突破。网络深度的增加可以丰富提取图像的不同特征,但同时也会带来一些问题。随着网络深度的增加,模型性能逐渐饱和和退化,导致梯度爆炸和消失等问题,He等人^[11]提出的残差网络解决

了该问题。假设 x 表示层的输入, $H(x)$ 为几个堆叠层需要学习隐藏的底层映射函数,而残差函数 $F(x)$ 则近似为 $H(x)-x$,原函数则为 $F(x)+x$ 。网络中的残差结构可以从增加的网络深度中获得增益使网络性能大大增加,即使网络在很深的情况下也能获得很好的性能。残差网络由残差块组成,通过短连接实现对网络层恒等映射。短连接是跳过一个层或者更多个网络层的连接,允许网络将图像特征从早期卷积层传递到后续卷积层。以残差网络为例,假设 x 和 y 为卷积层的输入和输出向量,则有 $y=F(x)+x$,其中 $F(x)$ 表示要学习的残差映射。短连接可以解决网络堆叠导致退化的问题,并且成本较低。既没有引入额外的计算参数也没有增加计算的复杂度。

目前常见的残差网络有ResNet18、ResNet34、ResNet50、ResNet101等,名字后的数字代表着各个网络的层数。其中ResNet50被广泛引用于各个领域的图像识别分类研究^[12-15]。ResNet50网络由Stem卷积层,残差块卷积层,平均全局池化层和全连接层组成,残差块卷积层由普通残差块和下采样残差块组成。本文利用所设计的残差网络模型来提取医学图像特征和特定疾病结果之间的关联,当临床医生通过输入新病例的图像,模型预测新病例的诊断结果,从而起到计算机辅助诊断作用。

综上所述,本文提出了一种基于残差网络结合深度可分离卷积和轻量化注意力的分类算法,来辅助诊断猴痘病毒皮肤病变。工作的主要贡献如下:(1)将ResNet50中 3×3 步长为1的标准卷积替换为深度可分离卷积,通过在结构中加入瓶颈层能够显著的降低计算成本,既能减少输入张量又不会影响模型整体的性能。(2)在残差块中加入无参数简单注意力模块(SimAM),它是无参数的,既能获得更好的性能,又能使模型更加轻量化。(3)本文将Swish激活函数替换成ReLU激活函数,在网络越深的条件下,Swish激活函数在训练期间能够防止梯度饱和,且具有更好的泛化能力。(4)在网络的残差块之间加入短连接,既解决了因残差块堆积而所带来的梯度爆炸和消失问题,又不会额外地增加模型参数和计算复杂性。整个网络可以通过带有反向传播的随机梯度下降进行端到端训练。

1 研究方法

1.1 注意力模块

SimAM注意力是一个轻量化的注意力模块^[16]。与

现有集中在通道域和空间域串行或并行组合的大多数注意力模块不同, SimAM 注意力模块参照人脑的协同机制, 提出了一种统一权值的三维注意力, 以便每个神经元分配到唯一的权重. 为了更好地实现注意力, 需要估计每个神经元的权重. 信息丰富的神经元通常与其他神经元表现出不同的放电模式, 较活跃的神经元还有可能抑制周围神经元的活动. 因此在视觉处理上, 应该给予这些神经元更高的权重. 而找到这些神经元的方法就是测量目标神经元和其他神经元之间的线性可分性, 由此定义了以下能量函数^[16]:

$$e_t(w_t, b_t, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{x}_i)^2 \quad (1)$$

其中, $\hat{t} = w_t t + b_t$, $\hat{x}_i = w_t x_i + b_t$. t 和 x_i 是输入特征 X 的单个通道中的目标神经元和其他神经元, M 为该通道上神经元的个数. 当 $y_t = \hat{t}$ 且 $y_0 = \hat{x}_i$ 时能量函数有最小值, 对 y_t 和 y_0 采用二进制标记 (即 1 和 -1), 并在式 (1) 中添加正则化符, 得到最终能量函数:

$$e_t(w_t, b_t, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_t x_i + b_t))^2 + (1 - (w_t t + b_t))^2 + \lambda w_t^2 \quad (2)$$

它的快速解析解为:

$$w_t = \frac{-2(t - \mu_t)}{(t - \mu_t)^2 + 2S_t^2 + 2\lambda} \quad (3)$$

$$b_t = -\frac{(t + \mu_t) \times w_t}{2} \quad (4)$$

其中, μ_t , S_t^2 分别为该通道中除了目标神经元以外所有神经元的平均值和方差. 假设单个通道中的所有像素遵循相同的分布, 可以计算出所有神经元的平均值和方差, 并重新使用该通道上的所有神经元. 它可以显著的降低计算成本, 避免重复计算每个位置的平均值和方差. 最小能量公式:

$$e_{t_{\min}} = \frac{4(\lambda + \hat{S}^2)}{2\hat{S}^2 + 2\lambda + (t - \hat{\mu})^2} \quad (5)$$

由式 (5) 可知, 当 $e_{t_{\min}}$ 的值越小, 则目标神经元与其他神经元的区别越大, 对视觉处理越重要. 因此, 每个神经元的重要性可以通过 $\frac{1}{e_{t_{\min}}}$ 来表示. 整体优化公式为:

$$\tilde{X} = \text{Sigmoid}\left(\frac{1}{E}\right) \cdot X \quad (6)$$

其中, E 是跨通道和空间维度的所有 $e_{t_{\min}}$ 的所有组.

SimAM 模块无需添加任何额外的参数, 与 SE^[17]、CBAM^[18] 等注意力模块相比, 这是一个很大的优势且可以有效增强各种网络的表达能力. 对于构建轻量化的网络使用 SimAM 注意力模块是一个很好的选择.

1.2 激活函数

本文模型选用的激活函数为 Swish 函数, 在更深层的网络模型中, 它比目前应用的最广泛的修正线性单元 (ReLU) 更有效. 且在图像分类和机器翻译等领域上, 性能始终与 ReLU 激活函数相匹配或优于 ReLU 激活函数^[19]. Swish 激活函数定义为:

$$f(x) = x \cdot \sigma(\beta x) \quad (7)$$

其中, $\sigma(x)$ 是 Sigmoid 函数, β 是常数或可训练参数, 在本文中 $\beta=1$. Swish 函数可以看作是一个平滑且非单调函数, 它有利于防止训练期间梯度逐渐接近为 0 导致饱和, 同时它平滑的特点让其具有更好的泛化能力和学习优化.

1.3 深度可分离卷积

在设计卷积神经网络模型中, 深度可分离卷积应用的十分广泛. Howard 等人^[20] 在 MobileNet 网络使用深度可分离卷积旨在减少模型的计算量以构造一个轻量化的网络, 适用于移动和嵌入式视觉应用. Chollet 等人^[21] 在 Xception 网络将 Inception 模块替换成深度可分离卷积旨在相似的网络参数量中, 网络能获得更好的性能. 深度可分离卷积包括两个部分, 一部分是输入的每个信道上独立执行的空间卷积, 用于滤波; 另一部分是输出上信道投影至新的信道空间的逐点卷积, 用于合并^[21]. 这种结构能有效地减少模型的计算量和规模. 以本文为例, 如图 1 所示, Stem 层结构中的卷积核为 3, 步长为 1, 滤波器数为 32 的深度可分离卷积, 与相同条件下的标准卷积相比, 它们的计算成本如下:

$$\text{cost} = K \times K \times C_{\text{in}} \times H_{\text{out}} \times W_{\text{out}} \times C_{\text{out}} \quad (8)$$

$$\text{cost}_C = 3 \times 3 \times 32 \times 110^2 \times 32 = 288 \times 110^2 \times 32 \quad (9)$$

$$\text{cost}_D = (3 \times 3 + 32) \times 110^2 \times 32 = 41 \times 110^2 \times 32 \quad (10)$$

其中, K 为滤波器尺寸, C_{in} 为输入通道数, C_{out} 为输出通道数, H_{out} 和 W_{out} 分别为输出特征图的长和宽. cost_C 为标准卷积计算成本, cost_D 为深度可分离卷积计算成本. 在相同条件下, 深度可分离卷积的计算成本降低了约 85.8%. 本文使用部分深度可分离卷积替代标准卷积, 其目的是在不影响模型整体性能的前提下, 在一定程度上将模型的复杂度降低.

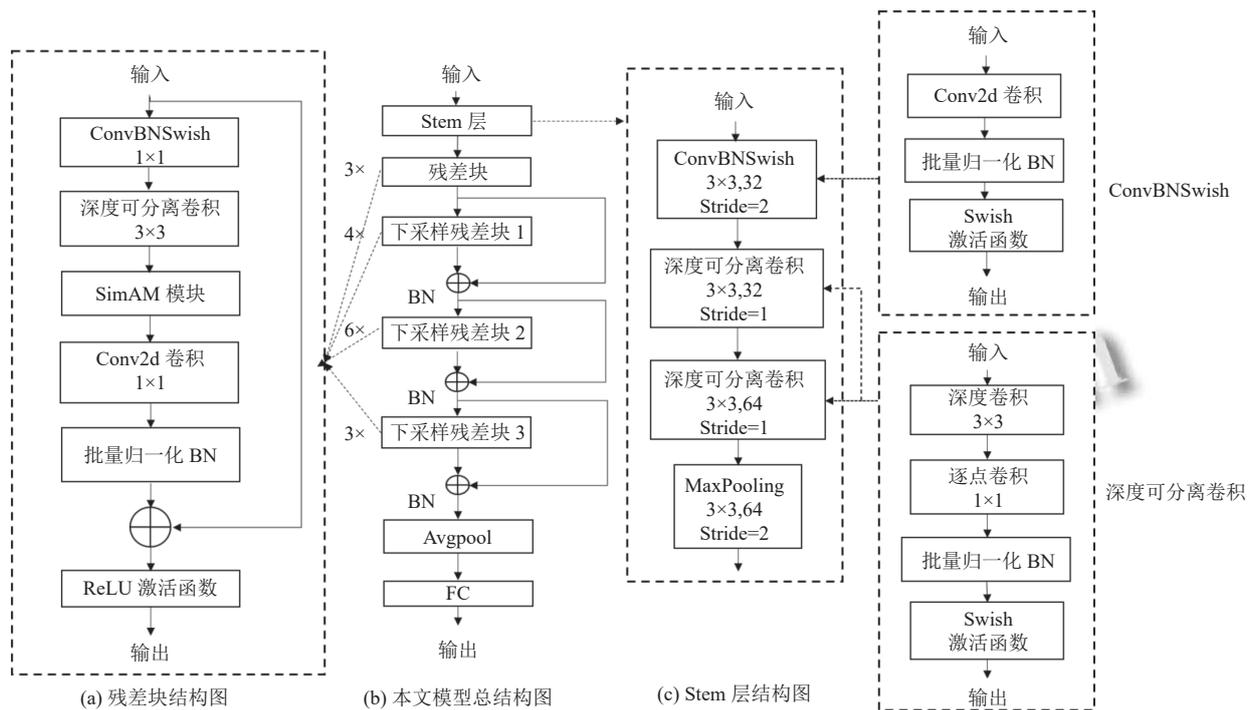


图1 模型总框图

1.4 网络模型设计

本文方法基于 ResNet50 网络模型, 其中每个残差块所拥有的残差数目分别为 3, 4, 6, 3^[11]. 将预处理过尺寸为 224×224×3 的图像作为改进网络模型的输入. 改进的网络模型如图 1 所示. ResNet50 中 Stem 层中使用一个 7×7 卷积核步长为 2 的卷积层和一个 3×3 卷积核步长为 2 的最大池化层, Stem 层的目的是为了降低输入图像的像素, 减少模型计算成本. 本文则将 Stem 层中的 7×7 卷积层替换成 3 个 3×3 的卷积层, 减少了模型的参数量, 且将其中步长为 1 的 2 个卷积层改为深度可分离卷积, 能够显著的降低计算成本, 既能减少输入张量又不会影响模型整体的性能. 在残差块中, 加入 SimAM 注意力模块, 在无任何参数的基础上, 提取有效的特征信息, 让模型更加轻量化. 模型中有 3 个下采样的残差块, 在残差块之间加入短连接可以防止梯度爆炸和加快网络模型的收敛. 在所有的单个卷积中都加入批量归一化 BN^[22] 和 Swish 激活函数, 防止过拟合和加快网络的训练和收敛.

2 实验分析

2.1 实验环境及参数

本文实验使用的计算机为 AMD Ryzen 9 5950X

16 核处理器, 显卡 GTX3080Ti, 显卡内存为 12 GB. 在 Windows 10 操作系统, Python 3.7 和 PyTorch 框架开发环境下运行. 所有实验的网络都使用随机梯度下降 (SDG) 优化器进行训练, 训练周期为 60, 使用批次大小为 32. 损失函数为交叉熵函数, 动量设置为 0.9, 权重衰减为 5E-4. 初始学习率设置为 0.001, 使用 PyTorch 框架中 ReduceLROnPlateau 类动态调整学习率, 其中参数 patience 设置为 5, factor 设置为 0.2.

2.2 数据集与预处理

本文研究所采用的数据集来自猴痘皮肤病变数据集 (MSLD)^[23], 其中包括猴痘、水痘和麻疹的皮肤病变图像, 这些图像主要是从网站、新闻门户网站和公众可访问的病例报告中收集的. 由于原样本量较少, 因此使用数据增强的方法增加样本量, 包括图片的旋转、平移、缩放和剪切; 改变图像的饱和度和对比度; 增加亮度抖动和噪声, 部分数据增强样本如图 2 所示. 数据增强后的样本其数量约原始样本数量的 14 倍, 即“猴痘”类有 1428 张图片和“其他”类有 1764 张图片, 数据集分布情况见表 1 所示.

将数据增强后的图像按照 7:2:1 的比例将数据集分为训练集、测试集和验证集. 网络模型的输入图像大小为 224×224×3, 所有输入图像均进行了归一化处理.

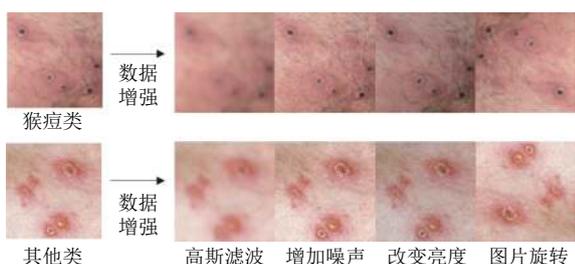


图2 部分数据增强图片

表1 数据集分布情况

| 类别标签 | 原始图像 | 数据增强图像 |
|------|------|--------|
| 猴痘类 | 102 | 1428 |
| 其他类 | 126 | 1764 |
| 总计 | 228 | 3192 |

2.3 评价指标

本文模型将训练后得到的准确率、精度、召回率和 $F1-score$ 作为评价模型训练的性能指标. 准确率表示预测结果的精准度.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (11)$$

精度表示在预测结果中, 真样本被正确预测成正样本的概率.

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

召回率表示在全部真样本中, 真样本被正确预测成正样本的概率.

$$Recall = \frac{TP}{TP + FN} \quad (13)$$

$F1-score$ 为精度和召回率的评估指标, 在一般情况下它们的关系是负相关的. 理想状态下, 精度和召回率越高越好, 即 $F1-score$ 的值越高越好.

$$F1-score = 2 \times \frac{Recall \times Accuracy}{Recall + Accuracy} \quad (14)$$

其中, P 为真样本, 即为猴痘病毒皮肤病变样本, N 为假样本, 即为其他皮肤病变样本. TP 表示真阳性样本, 即猴痘病毒皮肤病变被准确识别; TN 表示真阴性样本, 即其他皮肤病变被准确识别; FN 表示假阴性样本, 即猴痘病毒皮肤病变识别成其他皮肤病变; FP 表示假阳性样本, 即其他皮肤病变识别成猴痘病毒皮肤病变.

2.4 实验结果与分析

2.4.1 与不同激活函数对比

为了证明 Swish 激活函数在改进模型中的有效性,

实验将常见的几种激活函数进行对比, 其中激活函数包括 LReLU^[24]、ReLU^[25]、Mish^[26]. 在改进模型中分别使用这几种激活函数, 对比模型的性能, 实验结果如表 2 所示.

表2 不同激活函数在改进模型中对比

| 模型 | Recall | Precision | F1-score | Accuracy |
|-------|--------------|--------------|--------------|--------------|
| LReLU | 0.919 | 0.974 | 0.946 | 0.953 |
| ReLU | 0.937 | 0.967 | 0.952 | 0.958 |
| Mish | 0.954 | 0.982 | 0.968 | 0.972 |
| 本文方法 | 0.968 | 0.972 | 0.970 | 0.973 |

本文方法使用的 Swish 激活函数在召回率、准确率、 $F1-score$ 的评价指标中优于其他激活函数, 而精度略低于 Mish 激活函数. 观察 Mish 激活函数的其他 3 个指标可以发现: Mish 激活函数总体稳定性不如 Swish 激活函数, 在召回率和精度之间差距较大. 因此将 Swish 激活函数用于本文改进的算法更为合理.

2.4.2 与其他研究方法对比

实验对比了其他关于猴痘病毒皮肤病变分类的研究方法, 结果如表 3 所示.

表3 不同研究方法在数据集上的对比

| 模型 | Recall | Precision | F1-score | Accuracy |
|---------------------------|-------------|-------------|-------------|--------------|
| Ali等人 ^[23] | 0.83 | 0.87 | 0.84 | 0.830 |
| Sitaula等人 ^[27] | 0.85 | 0.85 | 0.85 | 0.871 |
| Sahin等人 ^[28] | 0.90 | 0.90 | 0.90 | 0.911 |
| 本文方法 | 0.97 | 0.97 | 0.97 | 0.973 |

从结果来看, 本文方法分类准确率达到 97.3%, $F1-score$ 达到了 97%, 均显著高于其他研究方法. 该对比表明本文提出的算法对猴痘病毒皮肤病变是十分有效的. 该模型能够在猴痘病毒皮肤病变图像上表现出较好的分类性能, 能为临床医生提供一个良好的客观的辅助诊断意见.

2.4.3 与不用网络模型对比

为了使实验对比数据更具说服力, 将常用的网络模型使用本文的实验参数进行训练, 在数据集上得到的结果如表 4 所示.

表4 不同网络模型在数据集上的对比

| 模型 | Recall | Precision | F1-score | Accuracy |
|-------------|--------------|--------------|--------------|--------------|
| DenseNet121 | 0.747 | 0.926 | 0.827 | 0.860 |
| InceptionV3 | 0.825 | 0.894 | 0.858 | 0.878 |
| ResNet34 | 0.884 | 0.930 | 0.906 | 0.918 |
| ResNet18 | 0.905 | 0.912 | 0.908 | 0.918 |
| ResNet50 | 0.930 | 0.933 | 0.931 | 0.939 |
| MobileNetV3 | 0.926 | 0.940 | 0.933 | 0.940 |
| 本文方法 | 0.968 | 0.973 | 0.970 | 0.973 |

ResNet18、ResNet34 和 ResNet50 三个残差网络的评价指标都要高于 DenseNet121 和 InceptionV3, 其中 ResNet50 在残差网络中性能最好, 只有 MobileNetV3 的评价指标略高于 ResNet50. MobileNetV3 中采用了深度可分离卷积和 SE 注意力模块. 由表 5 可知, 当基准模型加入注意力模块和深度可分离卷积后评价指标准确率、精度和 $F1-score$ 均高于 MobileNetV3. 综上, ResNet50 网络结构更适合本文的研究, 基于 ResNet50 改进的网络算法能使网络模型的性能进一步提升, 并达到较好的效果.

表 5 消融实验对比

| 模型 | Recall | Precision | F1-score | Accuracy |
|-------|--------------|--------------|--------------|--------------|
| base | 0.891 | 0.951 | 0.920 | 0.931 |
| base2 | 0.898 | 0.948 | 0.923 | 0.932 |
| base3 | 0.930 | 0.957 | 0.943 | 0.950 |
| 本文方法 | 0.968 | 0.972 | 0.970 | 0.973 |

2.4.4 消融实验

为了验证各个策略和模块的有效性, 我们进行了消融实验. 本文所使用的策略和方法包括深度可分离卷积、无参数简单注意力模块 (SimAM)、残差网络的残差块中加入短连接, 将这些策略和方法逐一加入到模型中. 实验将传统 ResNet50 中 Stem 层中的 7×7 卷积层替换成 3 个 3×3 卷积层, 并将传统 ResNet50 中 ReLU 激活函数替换成本文使用的 Swish 激活函数作为基准模型 (base). 在基准模型后加入深度可分离卷积, 模型记作 base2; 在 base2 模型中加入 SimAM 模块, 模型记作 base3; 在 base3 模型的部分残差块中加入短连接, 模型即为本文方法, 结果如表 5 所示.

从表 5 可知, 在逐一加入不同模块和策略之后, 分类的各项评价指标总体逐步提升. 基准模型单独加入深度可分离卷积后, 评价指标中的精度有所下降, 其他分类评价指标提升幅度较小, 分析原因是本文基于残差网络的规模较小, 深度可分离卷积在减少了计算量的情况下, 影响了模型的一部分性能. 但在模型中加入了短连接和注意力模块之后, 模型的性能大大提升, 其准确率和 $F1-score$ 分别比基准模型提升了 4.32% 和 5.15%. 分析其原因是猴痘病毒皮肤病变的位置较为分散, 而加入了无参数注意力模块可以帮助网络将注意力主要集中在病变区域, 使之更有效的提取猴痘病毒皮肤病变特征. 而在残差块之间加入短连接, 可以有效地解决网络中残差块堆积而造成网络退化的问题. 以

上分析说明本文提出的策略和模块是有效的.

2.4.5 模型复杂度分析

本文方法是基于 ResNet50 网络改进的, 对比本文方法和 ResNet50 模型的复杂度, 结果见表 6.

表 6 模型复杂度对比

| 模型 | Accuracy | F1-score | 参数量 (M) | MACs (G) |
|----------|--------------|--------------|--------------|-------------|
| ResNet50 | 0.939 | 0.931 | 25.56 | 4.13 |
| 本文方法 | 0.973 | 0.970 | 16.25 | 3.51 |

本文模型复杂度的评判指标为模型参数量和累积操作数 (MACs). 在卷积层中的计算公式如下:

$$\text{参数量} = K \times K \times C_{in} \times C_{out} \quad (15)$$

$$MACs = (C_{in}K^2 + 1)C_{out}H_{out}W_{out} \quad (16)$$

其中, K 为滤波器尺寸, C_{in} 为输入通道数, C_{out} 为输出通道数, H_{out} 和 W_{out} 分别为输出特征图的长和宽. 对比 ResNet50, 本文方法减少了 36.4% 的参数量和 15.0% 的 MACs, 其准确率和 $F1-score$ 分别比 ResNet50 提高了 5.17% 和 4.02%. 说明本文方法在不降低性能指标的同时, 又可以降低模型的复杂度.

3 结论与展望

本文提出了一种基于残差网络与轻量化注意力结合的猴痘病毒皮肤病变分类算法. 在 MSLD 数据集上对本文网络模型与现有常见的网络模型和研究方法对比, 实验结果表明本文模型具有更优的性能. 在未来的工作中, 将使用本文模型进一步解决其他皮肤病医学图像的分类任务, 从简单的二分类任务扩展到更为复杂的多分类任务.

参考文献

- Gessain A, Nakoune E, Yazdanpanah Y. Monkeypox. The New England Journal of Medicine, 2022, 387(19): 1783–1793. [doi: 10.1056/NEJMra2208860]
- 王会芳, 张利萍, 张纪元, 等. 非典型猴痘: 一场新的疫情? 解放军医学杂志, 2022, 47(6): 533–537. [doi: 10.11855/j.issn.0577-7402.2022.06.0533]
- Di Giulio DB, Eckburg PB. Human monkeypox: An emerging zoonosis. The Lancet Infectious Diseases, 2004, 4(1): 15–25. [doi: 10.1016/S1473-3099(03)00856-9]
- 姚开虎, 杜倩倩, 胡亚红. 人感染猴痘的诊断和治疗. 中国当代儿科杂志, 2022, 24(9): 960–966.
- Jezek Z, Szczeniowski M, Paluku KM, et al. Human

- monkeypox: Confusion with chickenpox. *Acta Tropica*, 1988, 45(4): 297–307.
- 6 Chan HP, Hadjiiski LM, Samala RK. Computer-aided diagnosis in the era of deep learning. *Medical Physics*, 2020, 47(5): e218–e227.
- 7 Chan HP, Samala RK, Hadjiiski LM, *et al.* Deep learning in medical image analysis. In: Lee G, Fujita H, eds. *Deep Learning in Medical Image Analysis*. Switzerland: Springer, 2020. 3–21.
- 8 Allugunti VR. A machine learning model for skin disease classification using convolution neural network. *International Journal of Computing, Programming and Database Management*, 2022, 3(1): 141–147.
- 9 Zuluaga-Gomez J, Al Masry Z, Benagoune K, *et al.* A CNN-based methodology for breast cancer diagnosis using thermal images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2021, 9(2): 131–145.
- 10 Saygılı A. A new approach for computer-aided detection of coronavirus (COVID-19) from CT and X-ray images using machine learning methods. *Applied Soft Computing*, 2021, 105: 107323. [doi: [10.1016/j.asoc.2021.107323](https://doi.org/10.1016/j.asoc.2021.107323)]
- 11 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.
- 12 胡喜园, 邓洪敏, 徐泽林, 等. 基于改进残差网络的阿尔兹海默症分类. *计算机应用*, 2022, 42(S1): 71–79.
- 13 黄英来, 艾昕. 改进残差网络在玉米叶片病害图像的分类研究. *计算机工程与应用*, 2021, 57(23): 178–184. [doi: [10.3778/j.issn.1002-8331.2105-0321](https://doi.org/10.3778/j.issn.1002-8331.2105-0321)]
- 14 谢圣桥, 宋健, 汤修映, 等. 基于迁移学习和残差网络的葡萄叶部病害识别. *农机化研究*, 2023, 45(8): 18–23, 28. [doi: [10.13427/j.cnki.njyi.2023.08.048](https://doi.org/10.13427/j.cnki.njyi.2023.08.048)]
- 15 李航, 余镇, 倪东, 等. 基于深度残差网络的皮肤镜图像黑色素瘤的识别. *中国生物医学工程学报*, 2018, 37(3): 274–282. [doi: [10.3969/j.issn.0258-8021.2018.03.003](https://doi.org/10.3969/j.issn.0258-8021.2018.03.003)]
- 16 Yang LX, Zhang RY, Li LD, *et al.* SimAM: A simple, parameter-free attention module for convolutional neural networks. *Proceedings of the 38th International Conference on Machine Learning*. PMLR, 2021. 11863–11874.
- 17 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 7132–7141.
- 18 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. *Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Munich: Springer, 2018. 3–19.
- 19 Ramachandran P, Zoph B, Le QV. Searching for activation functions. *Proceedings of the 6th International Conference on Learning Representations*. Vancouver: OpenReview.net, 2018.
- 20 Howard AG, Zhu ML, Chen B, *et al.* Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861, 2017.
- 21 Chollet F. Xception: Deep learning with depthwise separable convolutions. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 1800–1807.
- 22 Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on International Conference on Machine Learning*. Lille: JMLR.org, 2015. 448–456.
- 23 Ali SN, Ahmed M, Paul J, *et al.* Monkeypox skin lesion detection using deep learning models: A feasibility study. arXiv:2207.03342, 2022.
- 24 Maas AL, Hannun AY, Ng AY. Rectifier nonlinearities improve neural network acoustic models. *Proceedings of the 30th International Conference on Machine Learning*. Atlanta: JMLR, 2013. 3.
- 25 Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. *Proceedings of the 27th International Conference on International Conference on Machine Learning*. Haifa: Omnipress, 2010. 807–814.
- 26 Misra D. Mish: A self regularized non-monotonic neural activation function. arXiv:1908.08681v2, 2019.
- 27 Sitaula C, Shahi TB. Monkeypox virus detection using pre-trained deep learning-based approaches. *Journal of Medical Systems*, 2022, 46(11): 78. [doi: [10.1007/s10916-022-01868-2](https://doi.org/10.1007/s10916-022-01868-2)]
- 28 Sahin VH, Oztel I, Oztel GY. Human monkeypox classification from skin lesion images with deep pre-trained network using mobile application. *Journal of Medical Systems*, 2022, 46(11): 79. [doi: [10.1007/s10916-022-01863-7](https://doi.org/10.1007/s10916-022-01863-7)]

(校对责编: 孙君艳)