

基于改进 MobileNetV3-Small 的皮肤肿瘤分类^①



石兴, 方睿, 罗鸣, 刘天锴

(成都信息工程大学 计算机学院, 成都 610103)
通信作者: 方睿, E-mail: fangrui@cuit.edu.cn

摘要: 许多皮肤癌疾病具有明显的初期患病征兆. 目前皮肤癌诊断主要依靠具有专业知识的医务工作者进行诊断, 其存在着耗时长、复用性低等问题. 针对上述问题, 该研究提出一种基于改进 MobileNetV3-Small 的轻量级皮肤肿瘤识别模型. 首先提出了一种基于坐标注意力机制模块 (coordinate attention, CA) 的 CaCo 注意力模块. 其次针对皮肤肿瘤数据集样本分布不均衡, 提出了一种联合多损失函数来增强模型对少样本病例的学习能力. 实验结果表明, 改进的 MobileNetV3-CaCo 模型精确率、平衡准确率和模型参数量分别为 93.39%、86.35% 和 2.29M, 达到了理想的识别效果.

关键词: 皮肤癌; 轻量级模型; 注意力模块; 联合损失函数

引用格式: 石兴, 方睿, 罗鸣, 刘天锴. 基于改进 MobileNetV3-Small 的皮肤肿瘤分类. 计算机系统应用, 2023, 32(12): 120-128. <http://www.c-s-a.org.cn/1003-3254/9350.html>

Skin Tumor Classification Based on Improved MobileNetV3-Small

SHI Xing, FANG Rui, LUO Ming, LIU Tian-Kai

(College of Computer Sciences, Chengdu University of Information Technology, Chengdu 610103, China)

Abstract: Many skin cancer diseases have obvious early symptoms. Currently, the diagnosis of skin cancer mainly relies on medical workers with professional knowledge, bringing the problems such as long time consumption and low reusability. In response to these problems, a lightweight skin disease recognition model based on improved MobileNetV3-Small is proposed in this study. Firstly, a CaCo attention module based on coordinate attention (CA) mechanism is proposed. Secondly, for the uneven distribution of the samples of skin-cancer datasets, a combination of multiple loss functions is proposed to enhance the learning ability of the model for cases with few samples. The improved MobileNetV3-CaCo model has an accuracy, balance accuracy, and model parameter quantity of 93.39%, 86.35%, and 2.29M, respectively, thus ideal recognition results are achieved.

Key words: skin cancer; lightweight model; attention module; joint loss function

皮肤肿瘤是全世界常见的健康问题, 也是造成全球疾病负担的主要原因^[1]. 2013 年, Karimkhani 等在调查全球疾病负担时, 皮肤病造成的负担占全球疾病负担的 1.79%, 且成为非致命性疾病负担和残疾的第 4 大原因^[2].

常见的皮肤癌筛查手段为皮肤镜检查与病理组织

检查. 皮肤镜又称表皮透光显微镜, 其能对活体表皮等肉眼难以观察的细微结构提供清晰的图像. 相比于肉眼诊断的准确率 (约 70%), 经过皮肤镜诊断的皮肤病变, 其病理符合率可高达 95%. 病理组织检查是诊断皮肤恶性肿瘤的金标准^[3]. 通过有创性取样, 得到病变部位的组织, 病理医生通过在显微镜下观察进行病理诊

① 基金项目: 国家重点研发计划 (2020YFA0608000)

收稿时间: 2023-06-08; 修改时间: 2023-07-12, 2023-07-27; 采用时间: 2023-08-11; csa 在线出版时间: 2023-11-17

CNKI 网络首发时间: 2023-11-20

断. 以上两种诊断方法都需要在医院进行专业的诊断, 存在着不可复用性、经济成本高、耗时长等问题. 因此寻找一种快速诊断、准确率适中、复用性高的诊疗手段非常具有现实意义.

随着人工智能的快速发展, 深度学习被广泛地应用于皮肤癌分类领域. 基于深度学习的医学影像处理技术目前已被证明在某些领域达到甚至超过了人类专家水平, 因此这项技术也逐渐走向临床, 开始辅助医生的日常工作^[4]. Han等^[5]使用 ResNet152 模型对 12 种不同的皮肤疾病的临床图片进行了分类. 他们的研究表明, 该算法在某些情况下与专业皮肤科医生的诊断性能相媲美, 甚至更好, 为自动化皮肤疾病诊断系统的开发提供了基础. 王诗琪等^[6]使用 5 094 幅色素痣和脂溢性角化病 (SK) 的皮肤镜图像对 CNN 网络 ResNet-50 通过迁移学习进行训练, 建立 CNN 二分类模型. 同时在测试时对两种疾病各 30 张图片进行自动分类, 通过对该模型预测结果与 95 位有经验的皮肤科医师进行比较. 结果显示, 该 CNN 自动分类模型在这两种疾病的分类性能上与有经验的皮肤科医师水平相当. 斯坦福大学工程学院和医学院团队在 Nature 上发布了在皮肤癌诊断领域的最新突破^[7], 利用在 ImageNet 上预训练后的 Google Inception-v3 模型对 129 450 张皮肤表面图片进行训练. 最后通过模型预测后的识别精度和 21 位皮肤科医生的识别准确度分别达到 69.4% 和 66% 左右. 但由于其选用的基准模型参数量较大, 在移动端部署存在困难. Han等^[8]使用传统卷积神经网络对 174 种疾病合计 20 余万张图片进行训练, 该模型可准确预测恶性肿瘤, 提出主要治疗方案. 通过在 134 种疾病中进行多类别分类, 提高医疗专业人员的诊断能力. Fujisawa等^[9]开发了一个基于 GoogLeNet 皮肤病变分类器, 该分类器旨在诊断皮肤肿瘤. 该论文证明了深度学习和计算机视觉技术在医学领域, 特别是皮肤科诊断方面的潜力. 尽管他们使用了一个相对较小的数据集且该模型, 但他们的模型仍能实现高准确率的皮肤肿瘤诊断. 其模型主要缺点是在不同疾病识别性能上差异较大. 王璐^[10]提出了一种多特征融合的方式, 利用典型相关分析法对单一特征进行融合后训练 SVM 分类器, 同时通过选择不同的核函数优化分类器. 在自建的含有 90 种常见皮肤病共计 4 702 张图片的皮肤病临床图像数据集上 Top-1 准确率达到 65.7%. 由于该数据集为自建数据集, 在实验结果对比评判上存在不确定性. 孔祥勇^[11]、Li等^[12]提出了一种基于乘法的数据

融合方法, 通过在元数据的一维卷积序列提取系数来提升分类精度. 尽管通过结合元数据训练可以提高模型的识别准确率, 但这种识别模式适用性不是很强, 需要综合多元信息进行预测. 殷文君等^[13]提出一种融合皮肤肿瘤临床患者元数据的密集卷积网络分类模型, 但其使用的网络模型参数较大, 分类精度不高. Hekler等^[14]使用深度神经网络对组织病理学图像进行分类, 特别是对黑色素瘤的组织病理学图像进行分类. 尽管该网络对皮肤肿瘤图像的分类精度不高, 但也证明深度学习对黑色素瘤组织病理学诊断存在潜在帮助. 随着深度学习技术的进一步发展, 训练速度更快、参数量更少的轻量级模型不断出现. 其中 MobileNet 系列模型^[15-17]能够快速准确地识别类间差异, 具有速度快、灵活度高的特点, 其可以嵌入到移动设备. 目前已被广泛应用于病虫害识别^[18]、果实识别^[19]等领域, 但在皮肤镜图像识别的应用较少.

基于以上神经网络所存在的模型识别精度不高、模型参数量大、不易复现等不足之处, 本研究通过实验对比后采用轻量级网络模型 MobileNetV3-Small 作为基准模型来使模型达到轻量化的需求. 首先采用迁移学习方法, 从 PyTorch 官方代码仓库中获得预训练权重进行模型参数初始化, 使模型训练收敛速度更快, 缩短训练时间; 其次将 MobileNet 模型中的原 SE 注意力模块更换为改进的 CaCo 注意力模块, 增强模型对皮肤肿瘤图像的位置信息感知能力. 然后针对皮肤肿瘤数据集分布不均衡引入联合损失函数来增强模型对少样本病例的学习能力. 最后通过实验验证构建的 MobileNetV3-CaCo 皮肤肿瘤识别模型的效果.

1 模型构造方法

1.1 MobileNetV3-Small 模型

2019 年 Google 团队提出了继 MobileNetV2^[16]之后的新一代轻量级神经网络 MobileNetV3^[17]. 该网络模型在 ImageNet^[20]分类中实现了比 MobileNetV2 低 20% 的时延, 同时提高了 3.2% 精度. 对比 MobileNetV2 模型, MobileNetV3 加入了 SE 注意力模块, SE 模块主要作用是通过学习的方式获取每个 channel 的重要程度, 然后依照这个重要程度来对各个通道上的特征进行加权, 从而突出重要特征, 抑制不重要的特征. 简单来说就是训练一组权重, 对各个 channel 的特征图加权. 与此同时还更新了激活函数, 引入了新的非线性 h-swish 激活函数. MobileNetV3 的 block 模块如图 1 所示.

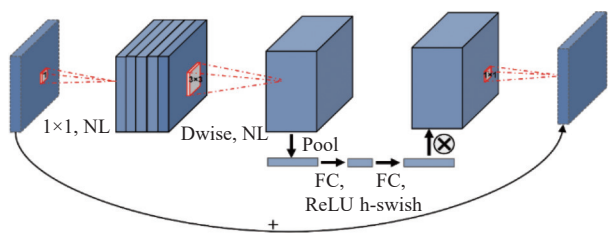


图1 MobileNetV3 block 结构图

除了网络模型上的更新外,还使用了NAS (neural architecture search) 搜索参数,作者通过使用NAS 搜索之后得到的网络,对网络结构的推理时间进行逐层分析,针对某些耗时的层结构做了进一步优化. MobileNet 系列主要通过深度卷积 (depthwise convolution) 和逐点卷积 (pointwise convolution) 替代传统卷积操作提高模型效率,使模型更容易用在移动端.

1.2 迁移学习

迁移学习是一种机器学习方法,它利用已有知识来解决不同但相关领域的问题.这种方法放宽了传统机器学习的两个基本假设:首先,传统机器学习要求训

练样本和测试样本在统计分布上是独立且相同的;其次,传统方法通常需要大量的训练本来构建良好的模型^[21].当前,迁移学习方法主要基于卷积神经网络 (convolutional neural network, CNN).这意味着我们可以利用在源领域中训练得到的权重信息.然后通过结合目标领域的分类任务,重新构建模型的全连接层,并使用目标领域的数据对模型进行微调,以适应新的分类要求.这种方法不仅可以减少训练时间,还能够实现更好的分类性能^[22].

2 MobileNetV3 的改进

2.1 改进 CA

计算机视觉中的注意力机制是让模型更加关注于图像中的重点信息,忽略干扰信息,其已经被广泛应用于多个领域^[23-25]. MobileNetV3-Small 在瓶颈结构中加入了SE 注意力模块,但SE 模块主要作用是增强对通道的注意力.而在特征图中位置信息也是尤为重要的,坐标注意力模块^[26]的基本结构如图2所示.

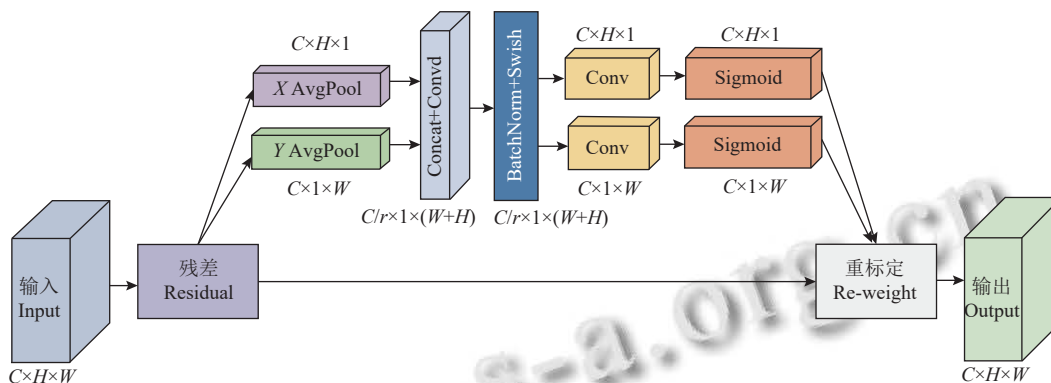


图2 坐标注意力模块基本结构

图2中X/Y AvgPool为X/Y方向平均池化;Concat代表拼接;BatchNorm代表批量归一化;Swish和Sigmoid代表非线性激活函数;C为通道数;H为特征图的高度;W为特征图的宽度;r为缩减系数.对于给定的输入X的通道数为C,高度为H,宽度为W.沿着输入特征图的水平方向(H,1)和垂直方向(1,W)分别进行池化操作,从而获得输入特征图的x,y轴相关位置信息.水平方向上得到的一维特征如式(1)所示:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \quad (1)$$

垂直方向上得到的一维特征如式(2)所示:

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w) \quad (2)$$

其中,X为特征图在第c通道具体位置的特征信息;h为特征图的具体高度,i为特征图的宽度,取值范围是[0,W];w为特征图的具体宽度,j为特征图的高度,取值范围为[0,H].

接下来,将2个方向上的特征图进行空间拼接,并使用1x1卷积变换将维度变为原来的C/r,在经过批量归一化和Swish激活操作得到包含2个方向信息的中间特征图,如式(3)所示:

$$f = \delta(F_1([Z^h, Z^w])) \quad (3)$$

其中, f 为空间信息在 2 个方向编码得到的中间特征图, δ 为激活函数 Swish, F_1 为 1×1 的卷积变换函数. 首先沿 2 个方向的空间维度将 f 分解为 2 个单独的张量 f^h 和 f^w . 然后利用 2 个 1×1 卷积变换函数将 f^h 和 f^w 转换成与输入 X 具有相同通道数的张量, 再经过激活函数 σ 后得到在 h 和 w 上的注意力权重. 最后将拓展后的注意力权重与 X 做乘法加权操作, 得到坐标注意力模块的输出, 如式 (4) 所示:

$$y_c = x_c(i, j) \cdot (\sigma[F_h(f^h)]) \cdot (\sigma[F_w(f^w)]) \quad (4)$$

其中, y_c 为第 c 通道的输出, σ 为激活函数 Sigmoid, F_h 和 F_w 为在高度和宽度上的卷积变换函数.

坐标注意力模块能够通过像素坐标系上的有效定位, 使模型能集中对感兴趣区域的注意力, 获取兼顾皮肤镜图像中通道和位置的信息, 减少对干扰信息的关注, 从而提升模型的特征表达能力. 但坐标注意力模块仅关注空间范围内的信息, 缺乏对通道的注意力. 两种不同的注意力各自的作用是不同的. 序列化地使用两种注意力机制, 要比并行化使用效果要好^[27]. 同时先使用通道注意力要优于先使用空间注意力. 因此对 CA 进行改进, 在其前面添加通道注意力机制, 由此得到改进的 CA 模块—CaCo, 让模型可以兼顾关注通道和空间两个维度的信息, 增强模型的识别能力. 改进 CaCo 注意力模块的结构如图 3 所示.

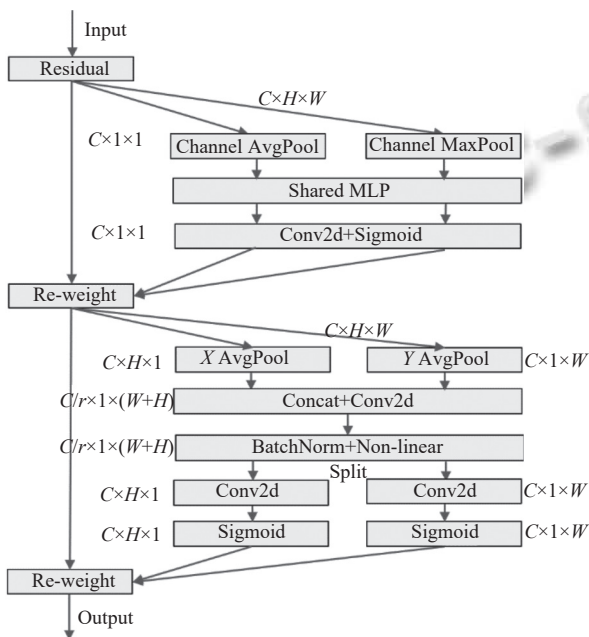


图 3 改进 CaCo 注意力模块结构

2.2 改进损失函数

Focal-loss 损失函数是 He 等^[28]在目标检测任务提出的, 其主要目的是增加那些少样本的损失权重, 降低多样本损失权重, focal-loss 函数是基于交叉熵损失函数改进而来的. Focal-loss 损失函数的定义如式 (5) 所示:

$$L_F = f(x) = \begin{cases} -\alpha(1-P)^\gamma \log P, & y = 1 \\ -\alpha P^\gamma \log(1-P), & y \neq 1 \end{cases} \quad (5)$$

其中, γ 为调节系数; α 为平衡因子. $\gamma > 0$ 时, 可以减小易于分类样本的损失, 使模型注重于更加分类困难和复杂的样本. 此外, α 作为平衡因子可以平衡数据集中因正负样本分布不均衡问题.

Label-distribution aware margin (LDAM)^[29]损失函数是另一种缓解类不平衡问题的损失函数. 一般来说, 同一类别样本的特征在特征空间上的距离是比较接近的, 不同类别样本的特征在特征空间上距离是比较远的. 于是分类问题本质上是在找一条决策边界, 使这些特征点能被正确的划分到对应的类别. LDAM 核心思想就是对少数类引入比多数类更强的正则化, 以减少它们的泛化误差, 从而使其保持模型学习多数类并强调少数类的能力. LDAM 损失只使少数类获得相对较大的决策边距, 更加偏向于少数类地学习, 以达到最优地分类效果. LDAM 损失函数的定义如式 (6) 所示:

$$L_{LDAM}((x, y); f) = -\log \frac{e^{z_y - \Delta_j}}{e^{z_y - \Delta_j} + \sum_{j \neq y} e^{z_j}} \quad (6)$$

其中, $\Delta_j = \frac{c}{n_j^{1/4}}$, $j \in \{1, \dots, k\}$, 相较于交叉熵损失函数, LDAM 损失函数可以增大少数类的分类间距, 并且该距离由样本数量调节.

Focal-loss 使模型更加注重于难以分类的样本, 而 LDAM 损失函数则是通过数据集的类数量来判断权重. 由此可以通过联合 focal-loss 和 LDAM 损失函数来同时监督 MobileNetV3-CaCo 网络的训练, 提高模型对皮肤肿瘤图像的整体识别精度. 联合损失函数如式 (7) 所示:

$$L = \alpha L_F + (1 - \alpha) L_{LDAM} \quad (7)$$

其中, α 是用来调节损失的权重因子, 通过实验证明, 使用联合损失函数有效提高了皮肤肿瘤分类的准确性, 并且当 $\alpha=0.5$ 时模型识别效果最好.

2.3 MobileNetV3-CaCo 整体结构

为了准确定位皮肤镜图像中的感兴趣区域, 引入改进后兼顾通道和坐标信息的注意力机制. 具体的来说是将 MobileNetV3-Small 原通道注意力挤压-激励 (squeeze-and-excitation, SE) 模块替换为改进后的 CaCo 注意力模块, 增强其对位置信息的关注度, 提升模型的识别效果. 改进后的 MobileNetV3-CaCo 模型整体结构如图 4 所示.

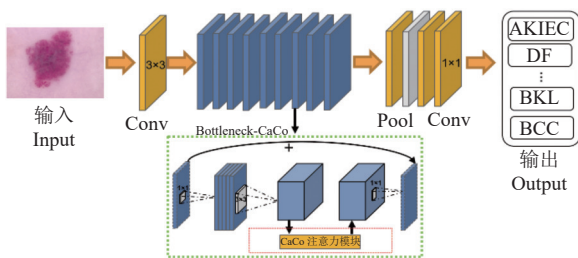


图 4 MobileNetV3-CaCo 模型整体结构

3 实验结果与分析

3.1 数据集介绍

本文训练与测试所采用的数据集是 HAM 数据集^[30], 该数据集由 10015 张分辨率为 600×450 像素的皮肤镜图像组成, 包含了 7 个不同的类别, 分别是: 黑色素细胞痣 (melanocytic nevi, NV)、基底细胞癌 (basal cell carcinoma, BCC)、黑色素瘤 (melanoma, MEL)、良性角化病 (benign keratosis, BKL)、皮肤纤维瘤 (dermatofibroma, DF)、血管性皮肤损伤 (vascular skin lesion, VASC) 和光化性角化病 (actinic keratosis, AKIEC). 其中基底细胞癌 (BCC) 和黑色素瘤 (MEL) 是恶性的, 其余皮肤肿瘤都是良性的, 所有黑色素瘤的诊断都是通过活检的组织病理学评估来证实的, 而黑色素细胞痣的诊断则是通过组织病理学检查 (24%)、专家共识 (54%) 或其他诊断方法来证实的. 数据集类别示例展示如图 5 所示. 数据统计情况如表 1 所示.

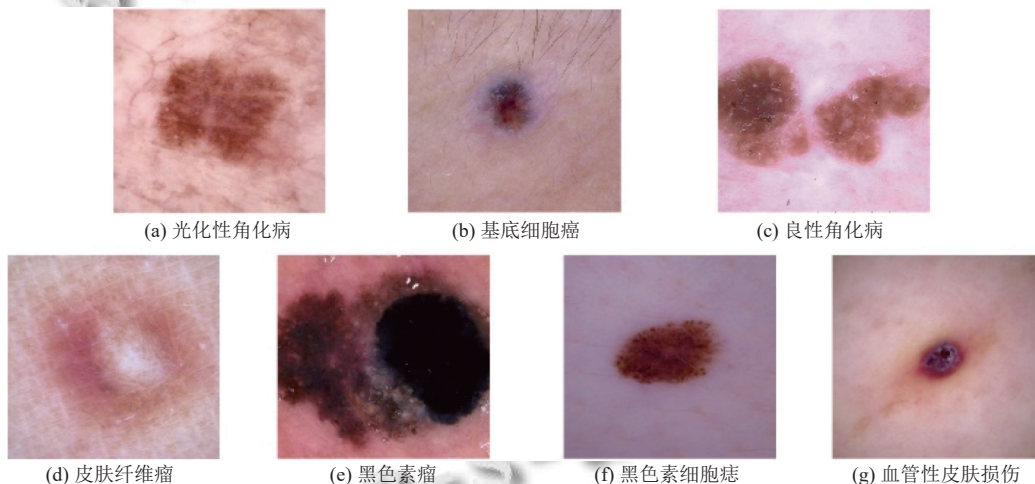


图 5 数据集类别示例

表 1 皮肤肿瘤样本分布情况

临床诊断	样本数
黑色素细胞痣 (NV)	6705
基底细胞癌 (BCC)	514
黑色素瘤 (MEL)	1113
良性角化病 (BKL)	1099
皮肤纤维瘤 (DF)	115
血管性皮肤损伤 (VASC)	142
光化性角化病 (AKIEC)	327

在进行数据预处理时, 采用随机水平翻转、随机垂直翻转、旋转 (90°, 180°)、随机颜色变换、随机缩放等方法对 HAM 数据集进行数据增强.

3.2 实验环境及参数设置

本文算法基于 PyTorch 框架在 Python 中实现. 训练时将数据集按照 8:1:1 的比例分为训练集、验证集和测试集, 训练时训练集与验证集采取交叉验证方式进行训练, 测试集的数据与训练时的数据集不重叠. 实验过程中, 统一设置 epoch 为 100, 优化器采用 AdamW 优化器, 学习率设置为 5E-4, weight decay 设置为 5E-2, batchsize 设置为 128. 训练时均使用 PyTorch 官方代码仓库中的预训练权重参数初始化模型权重.

实验使用的硬件配置为: 华硕 X570-PRO 主板、AMD Ryzen 9 3900x 12-core Processor 的 CPU、16 GB

DDR4 2 133 MHz 内存、NVIDIA Geforce GTX3080Ti 10 GB 显卡, 软件配置为: CUDA 11.7、Python 3.9、Cudnn 8.2.1、并采用 PyTorch 1.13.1 深度学习框架以及 Ubuntu 18.04 操作系统。

3.3 评价指标

为了全面评价 MobileNetV3-DA 模型的性能, 本实验选取正确率 (Accuracy)、特效度 (Specificity)、灵敏度 (Sensitivity) 和平衡准确性 (balance accuracy, BACC) 等指标综合评估模型的识别效果。计算公式如式 (8)–式 (11) 所示:

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

$$Specificity = \frac{TN}{TN + FP} \quad (9)$$

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

$$BACC = \frac{Sensitivity + Specificity}{2} \quad (11)$$

其中, TP 、 FP 、 FN 和 TN 分别为混淆矩阵中分类模型对不同皮肤肿瘤的分类情况统计。其中, TP (true positive) 代表真实值为正样本且识别也为正样本的数量, FP (false positive) 代表真实值为负样本但识别为正样本的数量, FN (false negative) 代表真实值为正样本但识别为负样本的数量, TN (true negative) 代表真实值为负样本且识别也为负样本的数量。进行皮肤肿瘤识别时, 把要识别样本的实际类别数视为正样本数, 把其他所有类别数之和视为负样本数。

3.4 MobileNetV3 与其他模型性能对比

在医学图像处理领域, 医学图像数据集的数据量通常比较小, 使用预训练策略可以有效提升模型的泛化能力。He 等^[31]从随机初始化权重的情况下训练不同的神经网络。然后用 COCO 数据集进行目标检测和示例分割任务, 结果显示预训练策略在目标任务上数据不足的时候可以带来大幅提升, 还能规避一些目标数据的优化问题, 同时还可以缩短研究周期。本实验同样在不同的初始化策略下对不同参数量的 MobileNetV3 以及 ResNet18 进行测试, 实验结果如表 2 所示。

对不使用预训练权重的 ResNet18 与 MobileNetV3 进行比较, 表 2 中的结果表明, 在训练策略相同的情况下, ResNet18 的性能表现优于 MobileNetV3。这是因为前者拥有更多的参数量, 在模型精度上表现得更好。同样, 分别对使用预训练策略的 ResNet18 与 Mobile-

NetV3 和未使用预训练策略的 ResNet18 与 MobileNetV3 进行比较。表 2 结果表明, 在其余训练策略相同的情况下, 使用预训练策略要比未使用预训练策略具有更好的泛化能力。

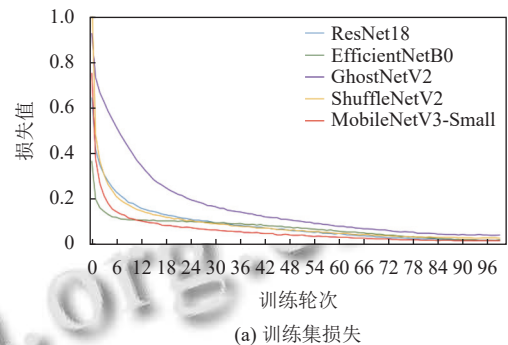
表 2 不同初始化策略的模型对比

模型	准确率 (%)	平衡准确性 (%)	模型参数量 (M)
ResNet18 (without pre-training)	89.88	80.67	11.18
ResNet18 (with pre-training)	91.48	84.23	11.18
MobileNetV3 (without pre-training)	87.17	73.71	2.18
MobileNetV3 (with pre-training)	90.88	83.44	2.18

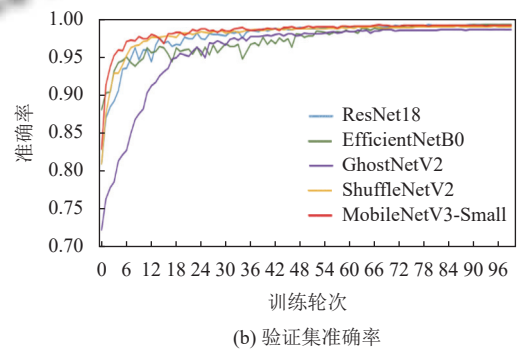
为了验证轻量级卷积神经网络 MobileNetV3 的有效性, 对不同深度学习模型进行比较。相关结果对见表 3 及图 6。

表 3 不同模型性能对比

模型	准确率 (%)	平衡准确性 (%)	模型参数量 (M)
ResNet18	91.48	84.23	11.18
ShuffleNetV2	91.38	84.25	5.36
GhostNetV2	88.88	78.63	6.16
EfficientNetB0	91.29	83.42	5.29
MobileNetV3	90.88	83.44	2.18



(a) 训练集损失



(b) 验证集准确率

图 6 模型实验结果对比

由表 3 可以看出, ResNet18 模型利用残差网络构建了较深的神经网络, 具有较好的特征提取能力和反向传播能力, 在实验中表现出较好的分类性能, 但是训练的时间长, 耗费的内存和计算资源较大; ShuffleNetV2

利用分组卷积的思想降低了参数量,但与此同时其收敛过程不稳定,导致模型的稳定性能受影响. MobileNetV3-Small 模型以较小的模型参数量取得较好的分类效果,训练结果接近 EfficientNetB0, 优于 GhostNetV2. 由图 6 可以看出,在经过 100 个 epoch 后,各个模型的损失值和准确率趋于稳定,这表明模型已经训练充分. 与其他模型相比 MobileNetV3-Small 在训练过程中表现出更平稳的收敛过程,损失值和准确率也较早的稳定在最优值附近,以较少的参数量实现了较好的分类识别效果.

3.5 MobileNetV3-CaCo 与其他模型性能对比

通过对改进后的 MobileNetV3-CaCo 结合不同的损失函数进行对比,相关结果对比见表 4.

表 4 MobileNetV3-CaCo 结合不同结合损失函数性能对比

模型	准确率 (%)	平衡准确性 (%)	模型参数量 (M)
MobileNetV3	90.88	83.44	2.18
MobileNetV3-CaCo	91.78	85.70	2.29
MobileNetV3-CaCo+focal	91.98	85.74	2.29
MobileNetV3-CaCo+LDAM	91.78	86.31	2.29
MobileNetV3-CaCo+改进损失函数	93.39	86.35	2.29

表 4 中前面 2 种未标注损失函数的表示统一使用交叉熵函数优化模型,后面 3 种表示改进后的 MobileNetV3-CaCo 模型分别采用 focal-loss 和 LDAM 以及改进后的混合损失函数提升模型的性能. 从表 4 中可知,通过使用 focal-loss 或 LDAM 损失函数后,模型的准确性以及平衡准确性都有所提升. 这是因为这两种损失函数分别注重于更加分类困难的样本和更加偏向于少数类地学习,在模型训练过程中都可以更好的调节因数据量不均衡而对不同样本数量的关注度差距. 通过结合这两种损失函数改进的混合损失函数在模型识别精度以及平衡准确性上均有所上升,能够有效识别少数量类的皮肤肿瘤图像.

通过对不同注意力模块组合进行比较,相关结果对比见表 5 及图 7. 其中 MobileNetV3-CBAM 表示 MobileNetV3-Small 的 SE 注意力模块替换为 CBAM 模块, MobileNetV3-SeCo 表示在原 SE 模块后引入坐标注意力模块, MobileNetV3-Eca 表示将 MobileNetV3-Small 的 SE 注意力模块替换为 Eca 模块, MobileNetV3-CaCo 表示本文构建模型. 从表 5 可知, MobileNetV3-CBAM、MobileNetV3-SeCo、MobileNetV3-Eca 和 MobileNetV3-CaCo 模型的识别准确率为 91.48%、89.68%、90.98% 和 91.78%, 相比其他 3 类模型,

MobileNetV3-CaCo 能够更加准确的识别皮肤镜图像,有效提升模型精度.

表 5 不同注意力模块组合加损失函数性能对比

模型	准确率 (%)	平衡准确性 (%)	模型参数量 (M)
MobileNetV3	90.88	83.44	2.18
MobileNetV3-CBAM	91.48	85.72	2.20
MobileNetV3-SeCo	89.68	78.70	2.64
MobileNetV3-Eca	90.98	82.64	2.08
MobileNetV3-CaCo	91.78	85.70	2.29
MobileNetV3-CaCo+改进损失函数	93.39	86.35	2.29

从图 7 可以看出,通过对不同注意力模块组合或者引入不同的注意力模块,对模型的识别效果影响参差不齐. 其中较另外两种模型比较,引入 CBAM 模块效果较好,尤其体现在对皮肤纤维瘤识别效果较好. 本文构建的 MobileNetV3-CaCo 模型的识别果优于其余模型的识别效果. 特别是在引入联合损失函数后,模型对每种皮肤镜图像识别效果更均衡,明显提升模型对少数量类的识别能力. 在少量增加模型参数量的情况下,模型的正确率和平衡准确性均得到提升.

4 结论与展望

针对皮肤镜图像难分类的问题,本研究对 MobileNetV3-Small 模型进行了改进,构建了 MobileNetV3-CaCo 皮肤镜图像识别模型. 将 MobileNetV3-Small 中 SE 注意力模块更换为改进后的 CaCo 注意力模块,以此在保留模型对通道的注意力的同时增强模型对位置信息的感知能力,提升了模型的信息表达能力;同时采用改进联合损失函数来提升模型的识别精度. 结果表明:

1) MobileNetV3-CaCo 模型能够准确识别皮肤镜图像. 针对数据集改进的混合损失函数可以有效提升模型识别精度. MobileNetV3-CaCo 在测试集上的识别准确率为 92.29%, 对比 ResNet18、GhostNetV2、EfficientNetB0 和 ShuffleNetV2 这些常见轻量级分类模型分别高 1.91、4.71、2.51 和 2.01 个百分点.

2) MobileNetV3-CaCo 模型满足了轻量化的要求. MobileNetV3-CaCo 的模型参数量为 2.29M, 较好的平衡了模型的复杂度和识别效果.

模型测试结果表明,本研究构建的 MobileNetV3-CaCo 模型能够快速、准确地识别皮肤镜图像. 下一步将考虑将模型部署到移动终端,实现皮肤镜图像的快速智能识别.

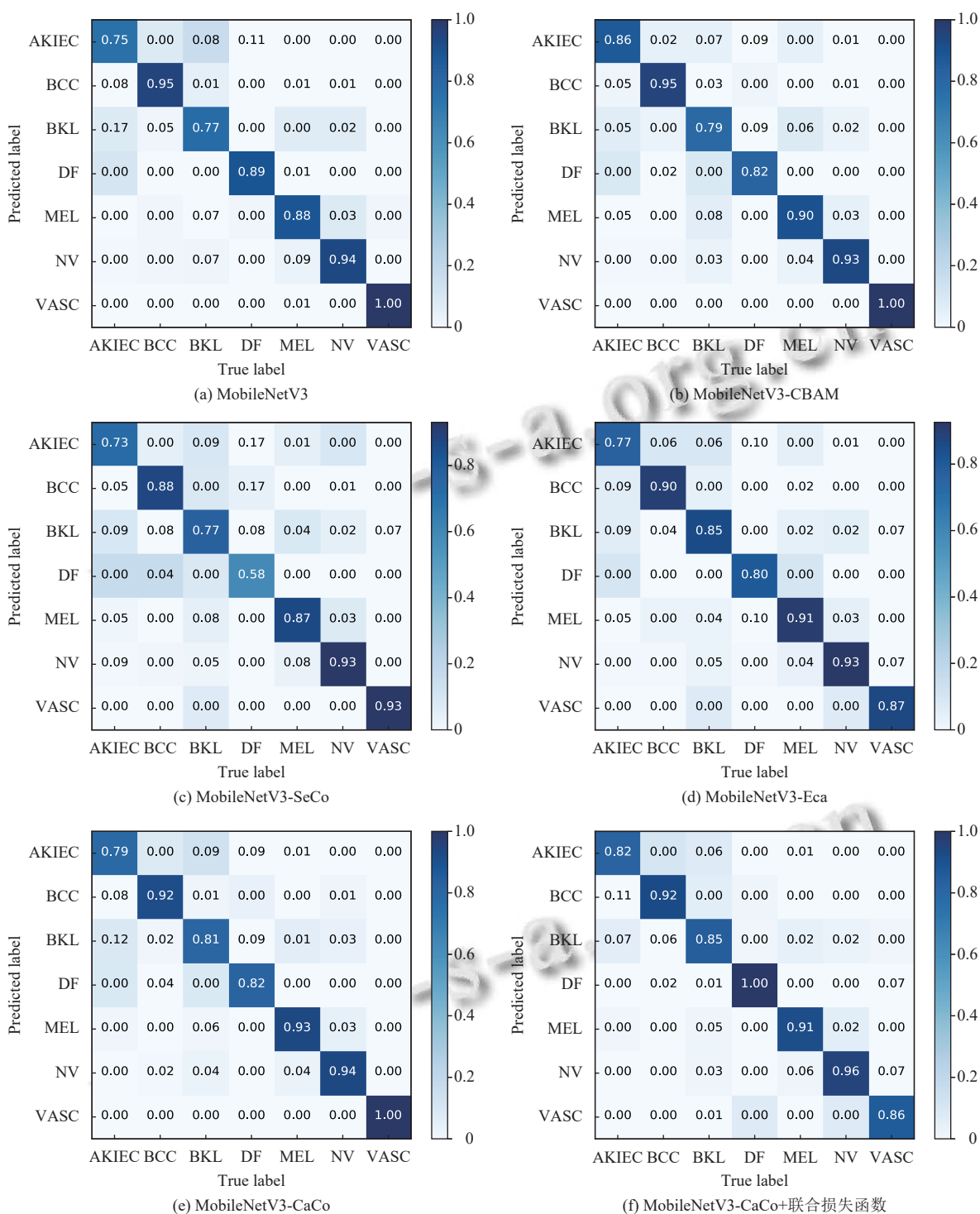


图7 引入不同注意力机制模型的混淆矩阵

参考文献

- 1 樊文龙, 王红心, 杨晓雨, 等. 1990–2019年亚洲 60–89岁老年人皮肤病的疾病负担. 中国皮肤性病学杂志, 2022, 36(11): 1257–1264.
- 2 Karimkhani C, Dellavalle RP, Coffeng LE, et al. Global skin disease morbidity and mortality: An update from the global

- burden of disease study 2013. JAMA Dermatology, 2017, 153(5): 406–412. [doi: 10.1001/jamadermatol.2016.5538]
- 3 中华医学会皮肤性病学分会皮肤肿瘤研究中心, 中国医师协会皮肤科医师分会皮肤肿瘤学组. 皮肤基底细胞癌诊疗专家共识 (2021). 中华皮肤科杂志, 2021, 54(9): 757–764.
- 4 姚侃敏, 潘自来, 宋琦, 等. 影像云在放射诊断质控工作中

- 的应用价值. 中国医学计算机成像杂志, 2018, 24(5): 401–405. [doi: 10.3969/j.issn.1006-5741.2018.05.009]
- 5 Han SS, Kim MS, Lim W, *et al.* Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. *Journal of Investigative Dermatology*, 2018, 138(7): 1529–1538. [doi: 10.1016/j.jid.2018.01.028]
- 6 王诗琪, 刘洁, 朱晨雨, 等. 皮肤科医师与深度卷积神经网络诊断色素痣和脂溢性角化病皮肤镜图像比较. *中华皮肤科杂志*, 2018, 51(7): 486–489. [doi: 10.3760/cma.j.issn.0412-4030.2018.07.002]
- 7 Esteva A, Kuprel B, Novoa RA, *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 2017, 542(7639): 115–118. [doi: 10.1038/nature21056]
- 8 Han SS, Park I, Chang SE, *et al.* Augmented intelligence dermatology: Deep neural networks empower medical professionals in diagnosing skin cancer and predicting treatment options for 134 skin disorders. *Journal of Investigative Dermatology*, 2020, 140(9): 1753–1761. [doi: 10.1016/j.jid.2020.01.019]
- 9 Fujisawa Y, Otomo Y, Ogata Y, *et al.* Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. *British Journal of Dermatology*, 2019, 180(2): 373–381. [doi: 10.1111/bjd.16924]
- 10 王路. 基于机器学习的皮肤病图像分类 [硕士学位论文]. 杭州: 浙江大学, 2018.
- 11 孔祥勇. 基于深度学习的黑色素瘤分类问题的研究 [硕士学位论文]. 杭州: 浙江工业大学, 2019.
- 12 Li WP, Zhuang JX, Wang RX, *et al.* Fusing metadata and dermoscopy images for skin disease diagnosis. *Proceedings of the 17th IEEE International Symposium on Biomedical Imaging*. Iowa City: IEEE, 2020. 1996–2000.
- 13 殷文君, 黄建华, 纪元法. 基于改进密集卷积网络的皮肤肿瘤分类方法. *计算机工程*, 2023, 49(7): 288–294.
- 14 Hekler A, Utikal JS, Enk AH, *et al.* Pathologist-level classification of histopathological melanoma images with deep neural networks. *European Journal of Cancer*, 2019, 115: 79–83. [doi: 10.1016/j.ejca.2019.04.021]
- 15 Howard AG, Zhu ML, Chen B, *et al.* MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv:1704.04861*, 2017.
- 16 Sandler M, Howard A, Zhu ML, *et al.* MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 4510–4520.
- 17 Howard A, Sandler M, Chu G, *et al.* Searching for MobileNetV3. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019. 1314–1324.
- 18 孙俊, 朱伟栋, 罗元秋, 等. 基于改进 MobileNet-V2 的田间农作物叶片病害识别. *农业工程学报*, 2021, 37(22): 161–169. [doi: 10.11975/j.issn.1002-6819.2021.22.018]
- 19 张伏, 陈自均, 鲍若飞, 等. 基于改进型 YOLOv4-LITE 轻量级神经网络的密集圣女果识别. *农业工程学报*, 2021, 37(16): 270–278. [doi: 10.11975/j.issn.1002-6819.2021.16.033]
- 20 Deng J, Dong W, Socher R, *et al.* ImageNet: A large-scale hierarchical image database. *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*. Miami: IEEE, 2009. 248–255.
- 21 陈立福, 武鸿, 崔先亮, 等. 基于迁移学习的卷积神经网络 SAR 图像目标识别. *中国空间科学技术*, 2018, 38(6): 45–51.
- 22 苏静雷, 王红军, 王政博, 等. 多通道卷积神经网络和迁移学习的燃气轮机转子故障诊断方法. *电子测量与仪器学报*, 2023, 37(3): 132–140.
- 23 曹城硕, 袁杰. 基于 YOLO-Mask 算法的口罩佩戴检测方法. *激光与光电子学进展*, 2021, 58(8): 0810019.
- 24 鲍海龙, 万敏, 刘忠祥, 等. 基于区域自我注意力的实时语义分割网络. *激光与光电子学进展*, 2021, 58(8): 0810018.
- 25 陈子涵, 吴浩博, 裴浩东, 等. 基于自注意力深度网络的图像超分辨率重建方法. *激光与光电子学进展*, 2021, 58(4): 0410013.
- 26 Hou QB, Zhou DQ, Feng JS. Coordinate attention for efficient mobile network design. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 13713–13722.
- 27 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.
- 28 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.
- 29 Cao KD, Wei C, Gaidon A, *et al.* Learning imbalanced datasets with label-distribution-aware margin loss. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*. Vancouver: Curran Associates Inc., 2019. 140.
- 30 Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions. *Scientific Data*, 2018, 5(1): 180161. [doi: 10.1038/sdata.2018.161]
- 31 He KM, Girshick R, Dollár P. Rethinking ImageNet pre-training. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019. 4918–4927.

(校对责编: 孙君艳)