

基于可变形卷积AlexNet与软注意力机制的皮肤 病变识别算法

曹 迅, 冯艳玲, 马昭鹏, 胡铭铭

云南大学信息学院, 云南 昆明

收稿日期: 2024年4月26日; 录用日期: 2024年5月24日; 发布日期: 2024年5月31日

摘 要

为解决皮肤病识别领域中数据集类别不平衡、模型复杂度高以及准确率低的问题, 提出了一种基于可变形卷积AlexNet与软注意力机制的皮肤病变识别算法。首先, 提出改进的可变形卷积AlexNet网络模型, 提高模型辨析力的同时, 降低了模型的参数量, 加快模型的训练和测试效率。然后, 在改进的模型中集成了软注意力机制, 使模型聚焦于皮肤病的关键特征区域, 优化模型的特征提取和识别能力。最后, 提出了一种联合损失函数, 对焦点损失函数与交叉熵损失函数进行加权, 聚焦于困难样本和易出错样本, 解决因数据集类别不平衡而导致的网络朝着错误方向收敛的问题。在公开数据集进行实验, 主观和客观的实验结果表明, 提出算法在七种不同类别的皮肤病识别准确率高于对比算法, 具有较强的鲁棒性和泛化能力。

关键词

皮肤病识别, 可变形卷积AlexNet, 软注意力机制, 联合损失函数

Skin Disease Identification Algorithm Based on Deformable Convolutional AlexNet and Soft Attention Mechanism

Xun Cao, Yanling Feng, Zhaopeng Ma, Mingming Hu

School of Information Science & Engineering, Yunnan University, Kunming Yunnan

Received: Apr. 26th, 2024; accepted: May 24th, 2024; published: May 31st, 2024

Abstract

To address the issues of dataset class imbalance, high model complexity, and low accuracy in the

文章引用: 曹迅, 冯艳玲, 马昭鹏, 胡铭铭. 基于可变形卷积 AlexNet 与软注意力机制的皮肤病变识别算法[J]. 计算机科学与应用, 2024, 14(5): 229-238. DOI: 10.12677/csa.2024.145131

field of skin disease identification, this paper proposes a skin lesion recognition algorithm based on deformable convolutional AlexNet and soft attention mechanisms. First, an improved deformable convolutional AlexNet network model is introduced, which enhances the model's discriminative power while reducing the number of model parameters, thereby speeding up the model's training and testing efficiency. Subsequently, a soft attention mechanism is integrated into the improved model, focusing the model on key feature areas of skin diseases to optimize its feature extraction and recognition capabilities. Finally, a joint loss function is proposed, which applies weights to the focal loss function and cross-entropy loss function, concentrating on difficult and error-prone samples to solve the issue of network convergence in the wrong direction due to dataset class imbalance. Experiments on public datasets, with both subjective and objective results, demonstrate that the proposed algorithm achieves higher accuracy in identifying seven different categories of skin diseases compared to benchmark algorithms, exhibiting strong robustness and generalizability.

Keywords

Skin Disease Identification, Deformable Convolutional AlexNet, Soft Attention Mechanism, Joint Loss Function

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

皮肤病是一种常见的疾病，且致癌率极高，目前已经成为了全球性的公共卫生问题。近几十年来，其发病率和死亡率持续攀升，引起了广泛的关注。在所有皮肤病中，尽管黑色素瘤所占比例仅为 5%，但其晚期致死率却高达 85% [1]。值得注意的是，早期发现和诊断可以显著提高大多数皮肤癌类型的治疗效果和患者的生存几率。因此，准确的早期诊断在皮肤癌的治疗策略中占据核心地位，是当前临床和科研领域亟待解决的关键问题[2]。

传统的皮肤病变诊断方法，如 ABCD 规则、7 点检查表、Menzies 方法以及 CASH 特征等，长久以来辅助皮肤科医生进行临床诊断。然而，这些方法在实际应用中面临着诸多挑战。首先，诊断结果的准确性在很大程度上依赖于皮肤科医生的临床经验，而医生间的诊断水平存在显著差异。此外，手工提取病变特征不仅费时费力，而且难以捕捉到复杂皮肤病变的全部结构特征，严重限制了在处理复杂病例中的效能。

近年来，深度学习成为皮肤病检测的一个理想的解决方案。Medhat 等[3]通过对传统 AlexNet 网络结构进行简化，并采用迭代幅度修剪技术，显著降低了模型的参数量和计算资源消耗，从而加速了模型的运行速度；Hoang 等[4]提出了一种结合熵权重(EW)和第一累积矩(FCM)的图像阈值技术，旨在提高分割精度，并通过优化 ShuffleNet 网络结构，实现了模型的轻量化，缩短了训练和测试时间；Raghavendra 等[5]设计了一种新型卷积神经网络，通过整合全局平均池化层(GAP)和预处理技术，既轻量化了模型，又提升了其泛化能力。然而，尽管这些方法提高了模型效率和速度，它们在捕获输入图像的关键特征区域方面仍显示出不足，影响了识别的准确性，并且未能有效应对皮肤病种类间的高相似度、类内差异性大以及数据集类别不平衡的问题。Yan 等[6]在 VGG-16 网络的基础上进行了简化和改进，通过引入预处理层(CBRM 层)并整合残差学习技术，增强了模型对皮肤病变图像中复杂细节和变化的捕获能力，从而提

升了皮肤病变分类的准确性；Wei 等[7]开发了一个融合空间注意力与通道注意力的网络架构，并引入辅助学习机制，专注于捕捉皮肤病变区域的关键局部模式，以辅助分类过程，显著提高了模型的特征表示能力和分类性能；Zhang 等[8]提出了基于注意力机制与残差学习的创新卷积神经网络模型，旨在提高分类的准确性；Carcagnì 等[9]对 DenseNet 中的最后一个 Dense Block 进行了多级细节调整和模型融合，有效提升了对不同皮肤病变分类的准确性；Nguyen 等[10]通过在 DenseNet201、ResNet50、VGG16 等主干模型中引入软注意力机制，增强了模型对重要特征的关注，提高了特征提取和分类性能。虽然这些方法在提升皮肤病变分类性能方面取得了显著进展，但同时也增加了模型结构的复杂度，为模型训练和优化带来了新的挑战。

综上所述，当前在皮肤病识别领域研究仍面临以下主要挑战：1) 网络模型复杂，参数冗余，耗费计算资源；2) 难以捕获到疾病关键特征，影响识别的准确性；3) 数据集类别不平衡的问题未得到有效解决，影响模型的训练和泛化能力。为了解决以上问题，本文提出了一种基于可变形卷积 AlexNet 与软注意力机制的皮肤病变识别算法，其主要贡献是：1) 提出改进的可变形卷积 AlexNet 网络模型，提高模型辨识力的同时，降低了模型的参数量，加快模型的训练和测试效率；2) 在改进的模型中集成了软注意力机制，使模型聚焦于皮肤病的关键特征区域，提高在复杂环境下提取皮肤病的关键特征的能力，优化模型的特征提取和识别能力；3) 提出加权的焦点损失函数与交叉熵损失函数，避免由于数据的类别不平衡导致模型朝着错误方向收敛，增强模型的泛化能力和稳定性。

2. 方法

2.1. 识别整体模型

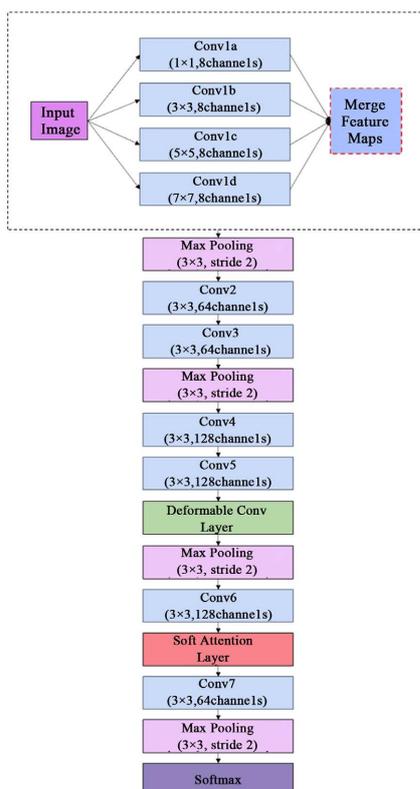


Figure 1. Network model structure

图 1. 网络模型结构

如图 1 所示, 提出的基于可变形卷积 AlexNet 和软注意力机制的网络模型架构包括 10 个卷积层、5 个最大池化层、3 个全连接层, 以及集成的 1 个可变形卷积层和 1 个软注意力层。为增强模型的表达能力, 所有卷积层均采用 ReLU 函数进行非线性激活。

在第一个卷积块通过设定 4 种不同尺寸(1×1 、 3×3 、 5×5 、 7×7)的卷积核, 且每种尺寸的通道大小均设置为 8, 来并行提取图像的特征。随后将这些特征合并为单一张量, 进行进一步的特征提取, 有效减少了模型的参数。继续的六个卷积层使用 3×3 卷积核, 通道大小依次为 64、64、128、128、128、64, 通过逐步精化的特征表达来深化模型的学习。所有最大池化层均配置为 3×3 的大小, 步长设置为 2, 减少特征维度并降低计算负担。

模型深层部分, 在第 8 和第 9 个卷积层后分别嵌入可变形卷积层和软注意力层, 使模型能够自适应皮肤病变图像的不规则区域, 并专注于关键特征区域。最后, 模型通过 softmax 层输出分类结果及其对应的概率分布。

2.2. 可变形卷积

在皮肤病变区域内, 特征往往呈现出多样化的尺度、不同的形态以及较大的形变。标准的卷积神经网络采用固定的几何架构进行运算, 其特征提取过程亦遵循固定的几何结构, 缺少对形态变化适应的机制, 这一限制降低了模型对形变较大病变区域的识别能力。

为解决此问题, 本文集成了可变形卷积[11]以优化图像特征提取。该技术使卷积神经网络能够适配不规则物体形状, 可变形卷积根据目标的具体形态动态调整, 更准确地匹配目标的自然特征, 避免了无关信息的捕获。

通过允许卷积核在特征区域附近自由采样, 而不受传统固定采样策略的约束, 并根据得到的特征图的实际情况动态调整其形状, 可变形卷积显著提升了特征图的解析能力。这种方法有效地解决了标准卷积在泛化能力及对形态变化适应性方面的缺陷, 图 2 展示了标准卷积与可变形卷积采样点的对比:

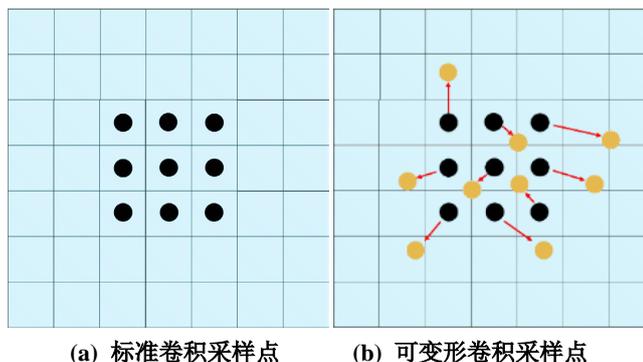


Figure 2. Comparison of standard convolution and deformable convolution sampling
图 2. 标准卷积与可变形卷积采样对比

可变形卷积的运算公式如下:

$$y(p) = \sum_{k=1}^K w_k \cdot x(p + p_k + \Delta p_k) \cdot \Delta m_k \quad (1)$$

式中, $x(p)$ 与 $y(p)$ 分别代表了输入与输出特征图在位置 p 的特征值。 K 旨在定量表示单元采样点的总数。对于特定的采样位置 k , w_k 表示相应的采样权重, 而 p_k 指代预设的偏移量。 Δp_k 与 Δm_k 分别代表在位置 k 上通过学习得到的偏移量和调制系数。

2.3. 软注意力机制

当处理皮肤病变图像时，通常只有图像中的一小部分像素与病变识别直接相关。皮肤病变的特征，包括对比度、颜色、形态、尺寸，以及各种伪影，显著提升了皮肤病分类任务的复杂度。受 Shaikh 等[12]的研究启发，本文将软注意力机制集成到皮肤病识别当中，以提高皮肤病识别的准确性和效率。

软注意力机制通过对输入图像中的每个像素赋予不同的权重，使模型能够主动聚焦于与诊断密切相关的特征，同时排除对诊断价值较低的信息。这种机制的引入，不仅优化了特征提取过程，而且增强了模型对于皮肤病变区域的敏感度和识别能力。

特征张量 $t \in \mathbb{R}^{h \times w \times d}$ 作为输入传递至三维卷积层，该层的权重维度定义为 $W_k = \mathbb{R}^{h \times w \times d \times K}$ 。随后使用 softmax 函数对该卷积输出进行归一化处理，生成 $K = 128$ 个分离注意力图。

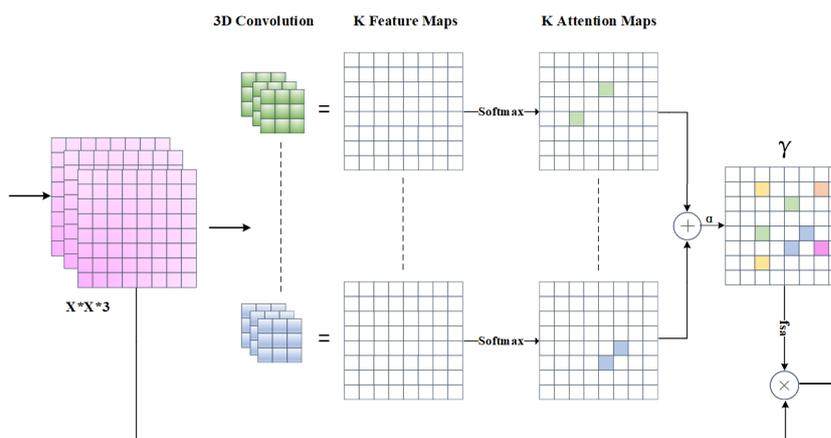


Figure 3. Soft attention unit
图 3. 软注意单元

软注意单元如图 3 所示，这些注意力特征图被聚集以产生一个统一的注意力图 α 作为加权函数，用以精确地调整特征张量 t 中的显著特征值。然后加权特征通过乘以一个可学习的标量 γ 对其进行缩放以产生细微调整后的特征 f_{sa} 。该过程细微调整的特征 f_{sa} 其原始特征采用残差连接方式结合，促进了信息的融合与保留。

在模型训练阶段，标量 γ 初始化为 0.01，使网络能够渐进地适应并主动学习调节所需的注意力强度，以优化网络对特征的专注度。

软注意的输出如下：

$$f_{sa} = \gamma t \left(\sum_{k=1}^K \text{soft max} (W_k * t) \right) \quad (2)$$

2.4. 联合损失函数

数据集的类别不平衡问题一直是皮肤病识别领域面临的一个主要挑战，为解决这一难题，本文提出了一种加权了交叉熵损失函数和焦点损失函数的联合损失函数。交叉熵损失函数使模型偏向多数类，负责优化分类任务性能。焦点损失函数则通过降低易分类样本的损失贡献，增加类别不平衡和难分类样本的损失贡献，提升模型的性能和鲁棒性。

交叉熵损失函数是一种广泛应用于分类问题的损失函数，其主要目的是最小化模型输出与真实标签之间的差异。随着预测值与真实标签之间的差异扩大，其“惩罚”程度按非线性规律递增，导致损失值

显著提高。反之，当模型预测的结果与真实标签更为接近时，交叉熵损失的数值则相应降低。

交叉熵损失函数表达式如下：

$$L_{CE}(p, y) = -\frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} y_{i,j} \log p_{i,j} \quad (3)$$

其中， N 表示总样本数， M 表示分类标签数， $y_{i,j}$ 表示样本 i 是否属于类别 j 的真实标签(1 或 0)， $p_{i,j}$ 表示模型预测样本 i 属于类别 j 的概率。

焦点损失函数是为了解决分类中的类别不平衡问题而提出的，特别是在一个类的样本远多于其他类的情况。通过减少易分类样本的权重来提高模型对难分类样本的关注度。焦点损失函数的定义为：

$$L_{FL}(p, y) = -\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \alpha (1 - p_{i,j})^{\varrho} y_{i,j} \log p_{i,j} \quad (4)$$

其中， ϱ 是调节参数，用以控制易分类样本对损失的贡献程度，从而强化对难分类样本的关注，本文取值为 3.0。

对上述损失函数进行加权后，提出的联合损失函数表达式为：

$$L_{combined} = (1 - \beta) L_{CE}(p, y) + \beta L_{FL}(p, y) \quad (5)$$

其中， β 表示平衡两种损失函数贡献的权重系数，本文取值为 0.0002。

3. 实验结果与分析

提出算法与对比算法参数设置保持一致性：批处理尺寸设定为 32，训练周期设定为 20；为训练网络模型，采用 Adam 优化器并将学习率设为 0.0002；输入图像尺寸均裁剪为 $224 \times 224 \times 3$ ，并做归一化处理。

实验环境的硬件配置使用 Intel(R) Xeon(R) CPU E5-2678 v3 @ 2.50GHz 作为处理器，并配备 NVIDIA GeForce RTX 2080Ti 显存为 11.0 GB 的 GPU。软件环境方面，系统基于 64 位 Windows 10 操作系统，采用 Python 3.7 作为编程语言，且选用 TensorFlow 2.0 框架进行深度学习模型的开发与测试。

3.1. 数据集

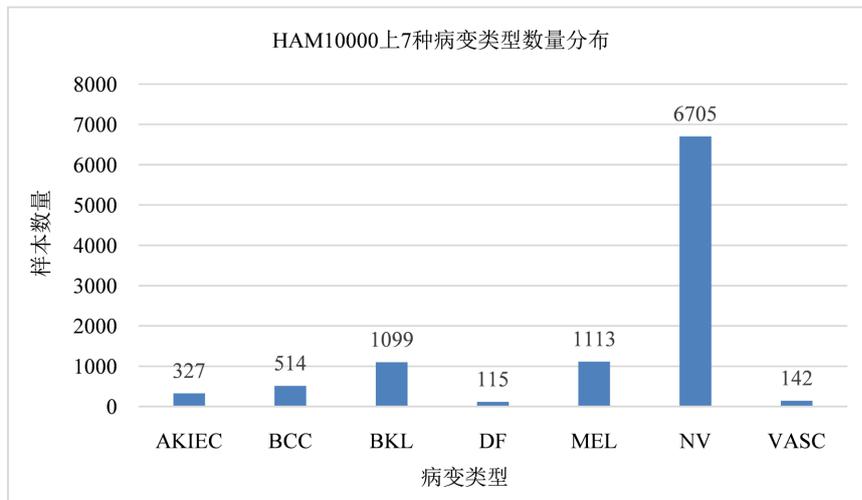


Figure 4. Number distribution of 7 types of skin lesions in HAM10000

图 4. HAM10000 上 7 种病变类型数量分布

HAM10000 [13]数据集,是由国际皮肤成像组织在2018年发布的公开数据集,包含10015张具有 450×600 像素分辨率的皮肤显微图像。此数据集包含了光化性角化病(AKIEC)、基底细胞癌(BCC)、良性角化病(BKL)、皮肤纤维瘤(DF)、黑色素瘤(MEL)、黑素细胞痣(NV)以及血管性皮肤病变(VASC)常见的7种病变类型。图4展示了数据集中7种病变类型的数量分布。

为了更好地评价模型训练和测试的有效性,将可访问带注释的数据集当中85%作为训练集,15%作为测试集。

3.2. 实验结果

本文将提出算法与 AlexNet [3]、VGG16 [6]、ResNet50 [10]、ARL14 和 ARL50 [8]进行了对比分析,其中所有对比算法均采用交叉熵损失函数进行优化。为了进一步增强模型预测结果的可解释性,本研究采用梯度加权类激活映射(Grad-CAM) [14]技术,通过分析目标类别对于卷积层输出的梯度信息,生成热力图(heatmap),以突出了网络在进行特定类别预测时所依赖的关键区域。

图5(a)展示了基底细胞癌(BCC)的原始皮肤病变图像,其中病变周边的毛发影响了识别过程。图5(b)正确识别了病变为BCC,却将焦点放在了图像右上方的毛发区,显示出较低的关注度。图5(c)热力图虽然覆盖了病变区域,但误将其分类为良性角化病(BKL)。图5(d)的热力图主要集中在右上方非病变区域,错误将病变识别为黑素细胞痣(NV)。图5(e)成功精准聚焦于病变区域,但误将其分类为良性角化病(BKL)。图5(f)的热力图主要集中在左上方非病变区域,错误将病变识别为黑素细胞痣(NV)。图5(g)是本文提出方法的预测结果,其热力图的中心精确覆盖了病变区域,且94.36%的识别准确率也展示了模型有较高的识别性能。

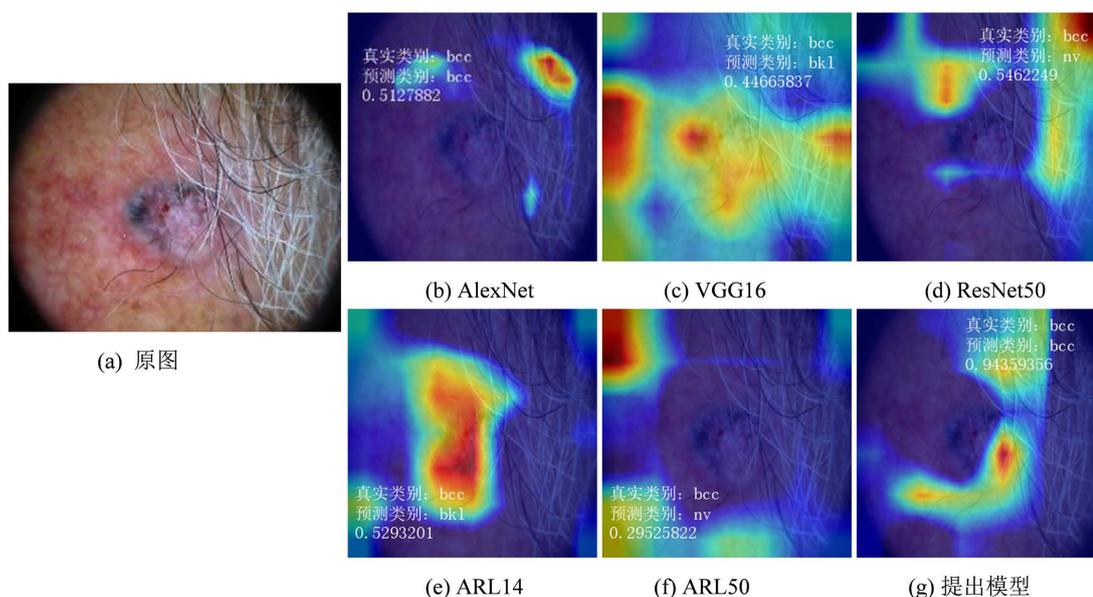


Figure 5. Experiment 1: Grad-CAM heat map analysis

图5. 实验1: Grad-CAM 热力图分析

图6(a)展示了良性角化病(BKL)的原始病变图像,病变区域分布相对较为分散。图6(b)~图6(g),各算法均成功将病变类型识别为BKL,展现了这些算法在病变识别方面的有效性。本文提出算法的预测结果如图6(g)所示,预测准确率达到97.04%,明显高于其他算法,包括图6(b)的83.21%、图6(c)的78.70%、图6(d)的73.85%、图6(e)的73.85%以及图6(f)的86.28%。这一结果不仅凸显了本文提出算法的优越性,也体现了在处理具有较大分散区域的BKL病变时,具有更高的识别精度和鲁棒性。

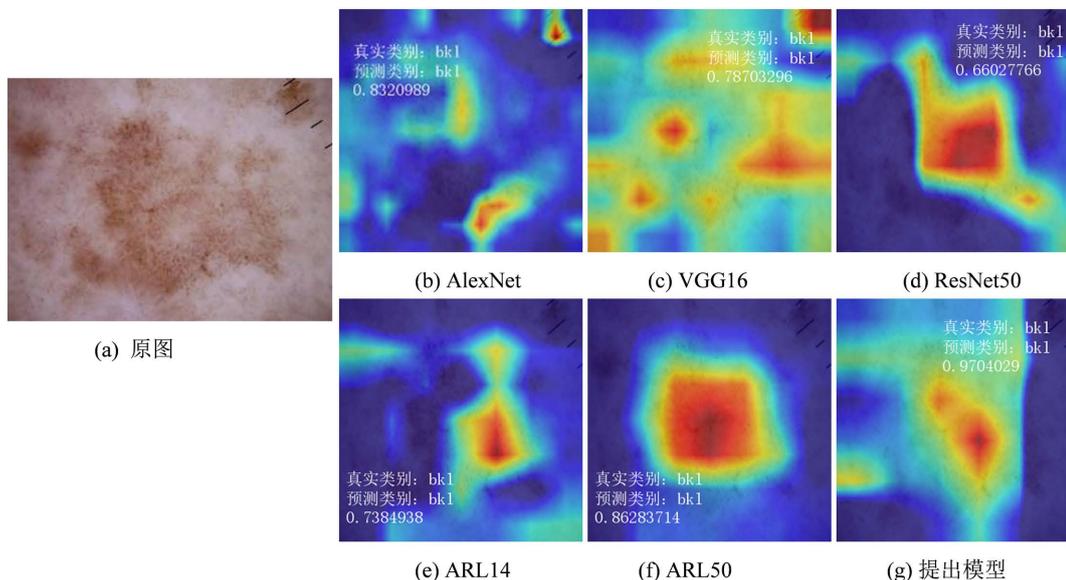


Figure 6. Experiment 2: Grad-CAM heat map analysis

图 6. 实验 2: Grad-CAM 热力图分析

图 7(a)展示了黑色素瘤(MEL)的原始病变图像,其中病变区域与周围正常皮肤的区分相对明显。图 7(b)~图 7(g)中,各算法均成功地将病变类型识别为 MEL,但对病变区域的关注度各有差异。AlexNet 主要关注病变区域的右上方,预测概率为 58.15%。VGG16 尽管识别类别正确,但由于高亮区域出现在非病变区,导致预测概率仅为 49.72%,标记为误识别。ResNet50 侧重于病变区域的右下方进行识别,预测概率达到 58.98%。而 ARL14 通过识别病变区域的下方实现了 76.94%的预测概率。ARL50 关注了不相关区域,预测概率为 71.38%。本文提出的模型则通过聚焦病变区域的左下方,并展现出较高的热力图活跃度,以 85.73%的准确率优于其他模型,说明其在病变识别精度与可靠性方面的优势。

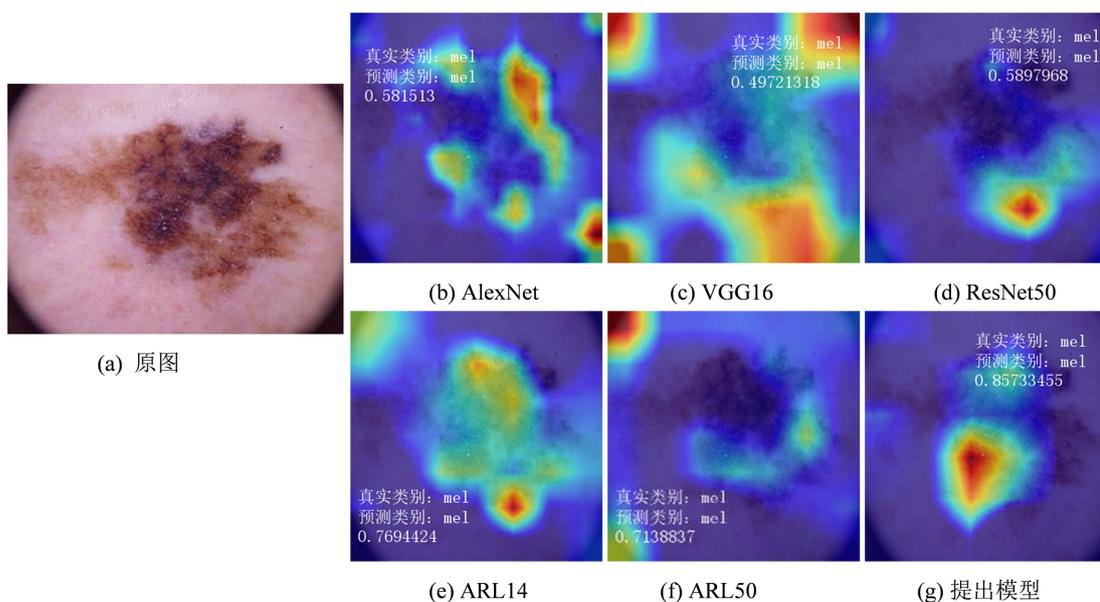


Figure 7. Experiment 3: Grad-CAM heat map analysis

图 7. 实验 3: Grad-CAM 热力图分析

图 8(a)展示了血管性皮肤病变(VASC)的原始病变图像。图 8(b)的预测概率达到了 99.76%，但热力图主要聚焦于病变周围非病变区域。图 8(c)虽成功预测出病变类型，但其热力图的分布广泛，未能实现对病变区域的精确定位。图 8(d)热力图集中于图像上方非病变区域，并将病变类型误诊为黑素细胞痣(NV)。图 8(e)的热力图虽正确标识出病变区域，然而预测概率只有 61.46%。图 8(f)热力图聚焦于病变周围区域，并将病变类型误诊为黑素细胞痣(NV)。图 8(g)的热力图准确标注了病变区域，而且预测概率高达 99.92%。

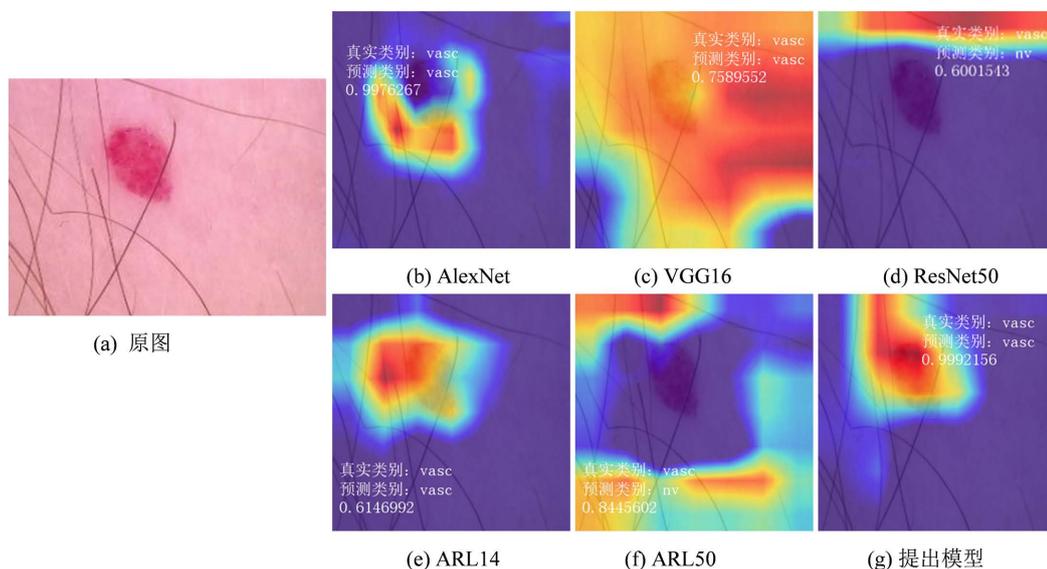


Figure 8. Experiment 4: Grad-CAM heat map analysis
图 8. 实验 4: Grad-CAM 热力图分析

为客观评估提出模型与对比模型的优劣，本文选用准确率(Accuracy)、精确度(Precision)以及 F1 得分(F1-score)作为评估标准。

Table 1. Evaluation index results (%) with HAM10000
表 1. 与 HAM10000 各算法评估指标结果(%)

模型	评估指标 (平均值)		
	Accuracy	Precision	F1-score
AlexNet	85.39	84.19	83.97
VGG16	84.90	82.35	83.37
ResNet50	85.63	83.44	83.81
ARL14	86.84	86.05	85.77
ARL50	84.06	83.77	83.64
提出模型	88.29	87.37	87.47

表 1 的实验结果表明，所提出的模型在准确率(Accuracy)、精确率(Precision)、F1 得分(F1-score)均优于表中其他对比算法。该结果表明提出模型对数据集中存在的类别不平衡问题有较好的处理能力和较高的识别精度。

4. 结论

提出了一种基于可变形卷积 AlexNet 与软注意力机制的皮肤病变识别算法。提出改进的 AlexNet 可以提高模型辨析力的同时，降低模型的参数量，加快模型的训练和测试效率。提出的集成了软注意力机

制的模型可以使模型聚焦于皮肤病的关键特征区域，优化模型的特征提取和识别能力。提出的联合损失函数可以避免由数据的类别不平衡导致模型朝着错误方向收敛，增强模型的泛化能力和稳定性。在公开数据集进行了训练与测试，并与已有的 5 个皮肤病识别算法进行比较，实验结果表明，提出的算法可以有效解决数据集类别不平衡、模型复杂、准确率低的问题，能够提高识别精度，具有较强鲁棒性和泛化能力。

参考文献

- [1] SCF (2022) Skin Cancer Foundation. <https://www.skincancer.org/>
- [2] CDC (2022) Centers for Disease Control and Prevention. <https://www.cdc.gov/cancer/skin/>
- [3] Medhat, S., Abdel-Galil, H., Aboutabl, A.E., et al. (2023) Iterative Magnitude Pruning-Based Light-Version of Alex-Net for Skin Cancer Classification. *Neural Computing and Applications*, **36**, 1413-1428. <https://doi.org/10.1007/s00521-023-09111-w>
- [4] Hoang, L., Lee, S.H., Lee, E.J., et al. (2022) Multiclass Skin Lesion Classification Using a Novel Lightweight Deep Learning Framework for Smart Healthcare. *Applied Sciences*, **12**, Article 2677. <https://doi.org/10.3390/app12052677>
- [5] Raghavendra, P.V.S.P., Charitha, C., Begum, K.G., et al. (2023) Deep Learning-Based Skin Lesion Multi-Class Classification with Global Average Pooling Improvement. *Journal of Digital Imaging*, **36**, 2227-2248. <https://doi.org/10.1007/s10278-023-00862-5>
- [6] Yan, P., Wang, G., Chen, J., et al. (2023) Skin Lesion Classification Based on the VGG-16 Fusion Residual Structure. *International Journal of Imaging Systems and Technology*, **33**, 53-68. <https://doi.org/10.1002/ima.22798>
- [7] Wei, Z., Li, Q. and Song, H. (2022) Dual Attention Based Network for Skin Lesion Classification with Auxiliary Learning. *Biomedical Signal Processing and Control*, **74**, Article 103549. <https://doi.org/10.1016/j.bspc.2022.103549>
- [8] Zhang, J.P., et al. (2019) Attention Residual Learning for Skin Lesion Classification. *IEEE Transactions on Medical Imaging*, **38**, 2092-2103. <https://doi.org/10.1109/TMI.2019.2893944>
- [9] Carcagnì, P., Leo, M., Cuna, A., et al. (2019) Classification of Skin Lesions by Combining Multilevel Learnings in a DenseNet Architecture. *Image Analysis and Processing—ICIAP 2019: 20th International Conference*, Trento, Italy, 9-13 September 2019, 335-344. https://doi.org/10.1007/978-3-030-30642-7_30
- [10] Nguyen, V.D., Bui, N.D. and Do, H.K. (2022) Skin Lesion Classification on Imbalanced Data Using Deep Learning with Soft Attention. *Sensors*, **22**, Article 7530. <https://doi.org/10.3390/s22197530>
- [11] Dai, J.F., et al. (2017) Deformable Convolutional Networks. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017. <https://doi.org/10.1109/ICCV.2017.89>
- [12] Shaikh, M.A., et al. (2020) Attention Based Writer Independent Verification. 2020 *17th International Conference on Frontiers in Handwriting Recognition (ICFHR)*. Dortmund, 8-10 September 2020. <https://doi.org/10.1109/ICFHR2020.2020.00074>
- [13] Philipp, T., Rosendahl, C. and Kittler, H. (2018) The HAM10000 Dataset, a Large Collection of Multi-Source Dermatoscopic Images of Common Pigmented Skin Lesions. *Scientific Data*, **5**, Article No. 180161. <https://doi.org/10.1038/sdata.2018.161>
- [14] Selvaraju, R.R., et al. (2017) Grad-Cam: Visual Explanations from Deep Networks via Gradient-Based Localization. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017. <https://doi.org/10.1109/ICCV.2017.74>