

基于改进ResNet34的皮肤病分类研究

汪鑫阳, 李延玲*

青海民族大学数学与统计学院, 青海 西宁

收稿日期: 2026年7月20日; 录用日期: 2026年8月13日; 发布日期: 2026年8月25日

摘要

针对皮肤镜图像分类中深层卷积神经网络特征提取不充分及梯度消失的问题, 本文提出了一种基于改进ResNet34的皮肤病分类模型(Enhanced-ResNet34)。该模型在ResNet34的架构基础上, 引入了自适应归一化层以提升小批量训练的稳定性, 并构建了融合SE注意力机制与随机深度正则化的改进残差块, 以增强关键病理特征的表达能力并降低过拟合风险。同时, 在网络末端接入密集连接块以加强特征复用。实验采用HAM10000和ISIC2019两个公开数据集进行验证, 通过迁移学习、高级数据增强及网络结构优化的逐步消融实验表明, 本文模型在HAM10000验证集上的准确率达到92.6%, 显著优于基线模型。此外, 跨数据集泛化实验证明, 该模型在ISIC2019数据集上仍保持87.0%的分类准确率, 具备良好的鲁棒性与临床应用潜力。

关键词

皮肤病分类, 深度学习, 注意力机制, 数据增强, 迁移学习

Research on Skin Disease Classification Based on Improved ResNet34

Xinyang Wang, Yanling Li*

School of Mathematics and Statistics, Qinghai Minzu University, Xining Qinghai

Received: July 20, 2026; accepted: August 13, 2026; published: August 25, 2026

Abstract

To address the issues of insufficient feature extraction and vanishing gradients in deep convolutional neural networks for dermoscopic image classification, this paper proposes an enhanced skin disease classification model based on ResNet34 (Enhanced-ResNet34). Building upon the standard ResNet34 architecture, the model introduces an adaptive normalization layer to improve training stability with small batches. Furthermore, an improved residual block integrating the Squeeze-and-Excitation (SE)

*通讯作者。

文章引用: 汪鑫阳, 李延玲. 基于改进 ResNet34 的皮肤病分类研究[J]. 应用数学进展, 2026, 15(8): 293-307.
DOI: 10.12677/aam.2026.158354

attention mechanism and Stochastic Depth regularization is constructed to enhance the representation of critical pathological features and mitigate overfitting risks. Concurrently, a dense connection block is incorporated at the end of the network to strengthen feature reuse. The proposed model was evaluated on two public datasets: HAM10000 and ISIC2019. Through stepwise ablation experiments involving transfer learning, advanced data augmentation, and network structure optimization, the model achieved an accuracy of 92.6% on the HAM10000 validation set, significantly outperforming the baseline model. Additionally, cross-dataset generalization experiments demonstrated that the model maintained a classification accuracy of 87.0% on the ISIC2019 dataset, indicating strong robustness and potential for clinical applications.

Keywords

Skin Disease Classification, Deep Learning, Attention Mechanism, Data Augmentation, Transfer Learning

Copyright © 2026 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

皮肤癌是全球发病率增长最快的恶性肿瘤之一，若能在早期(厚度小于 1 mm)诊断并切除，患者 5 年生存率可高达 98%；反之则急剧下降。因此，早期筛查与精准分类是降低死亡率的关键。

皮肤镜(Dermoscopy)作为非侵入性成像技术，能清晰展示皮损微观结构，但传统人工诊断依赖经验且主观性强。近年来，深度学习在医学图像分析领域取得突破[1]。

本文提出了一种融合注意力机制与密集连接的增强型深度神经网络 ResNet34 模型(Enhanced-ResNet34)，通过对 ResNet34 微观结构的重构，融合多维度优化策略以提升诊断性能。

2. 数据集与预处理

2.1. 数据集

本研究采用的两个公开基准数据集为 HAM10000 和 ISIC2019，这两个数据集在皮肤癌分类研究中被广泛使用，具有极高的权威性和代表性。

2.1.1. HAM1000 数据集

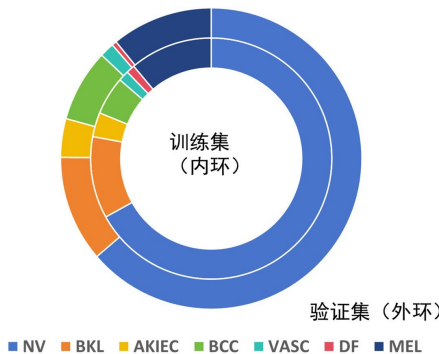


Figure 1. Sample distribution of the HAM10000 dataset
图 1. HAM10000 数据集样本分布

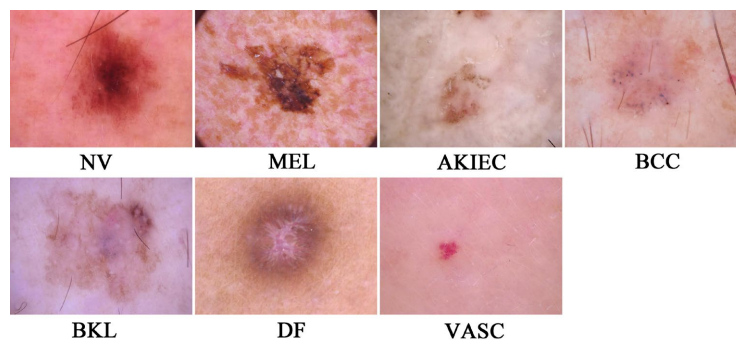


Figure 2. Sample images of the HAM10000 dataset

图 2. HAM10000 数据集样本图像

HAM10000 (Human Against Machine)数据集, 分布见图 1, 共包含 10,015 张训练集及 193 张验证集的皮肤镜图像。图像涵盖了七种常见的皮肤病变类型, 图像见图 2: 黑色素瘤(MEL)、黑素细胞痣(NV)、基底细胞癌(BCC)、光化性角化病(AKIEC)、良性角化病(BKL)、皮肤纤维瘤(DF)、血管性病变(VASC)。

2.1.2. ISIC2019 数据集

ISIC2019 数据集包含 25,331 张训练集和 6191 张验证集图像, 各种类分布见图 3, 并在病变类别上进行了改变, 有黑色素瘤(MEL)、黑素细胞痣(NV)、基底细胞癌(BCC)、光化性角化病(AKIEC)、良性角化病(BKL)、皮肤纤维瘤(DF)、血管病变(VASC), 并新增了鳞状细胞癌(SCC), 新增图像见图 4。

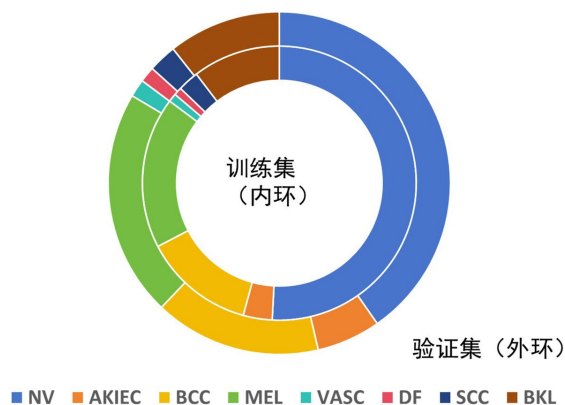
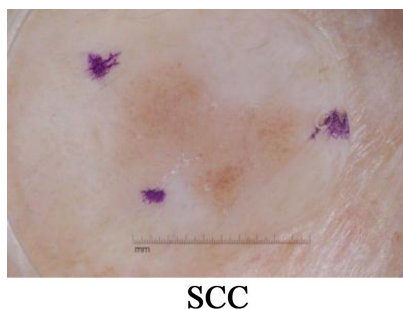


Figure 3. Sample distribution of the ISIC2019 dataset

图 3. ISIC2019 数据集样本分布



SCC

Figure 4. Augmented sample images of the ISIC2019 dataset

图 4. ISIC2019 数据集增加样本图像

2.2. 数据预处理

为了适配深度卷积神经网络的输入要求并提升模型的训练稳定性, 本文对原始皮肤镜图像进行了一系列基础的预处理操作。

2.2.1. 图像尺寸归一化与重采样

由于皮肤镜设备采集的原始图像分辨率各异, 直接输入网络会导致计算量巨大且难以批量处理。因此, 本文首先将所有图像统一调整至 224×224 像素的固定尺寸。在缩放过程中, 采用双线性插值法(Bilinear Interpolation)进行重采样。对于目标图像中的像素点 (x, y) , 其像素值 $f(x, y)$ 由源图像中距离最近的四个像素点加权计算得出:

$$f(x, y) = w_1 f(x_1, y_1) + w_2 f(x_2, y_1) + w_3 f(x_1, y_2) + w_4 f(x_2, y_2)$$

其中, $f(x_i, y_j)$ 为源图像中相邻四个像素点的值, w_i 为根据几何距离计算出的对应权重。该操作可以在保证图像视觉效果平滑的同时, 最大程度保留皮损的形态学特征, 以满足后续深度学习模型对输入张量维度的严格要求。

2.2.2. 色彩空间转换

皮肤镜图像通常包含丰富的颜色信息, 如黑色素分布、血管形态等, 这些信息对诊断至关重要。在预处理阶段, 本文确保所有图像均被转换为标准的 RGB 三通道色彩模式。对于极少数可能存在单通道或四通道的 RGBA 图像, 预处理流程会将其强制转换为三通道 RGB 格式, 从而保证输入数据的通道一致性。

2.2.3. 数据类型转换与张量标准化

在将图像送入神经网络计算前, 需将其从 NumPy 数组格式转换为 PyTorch 张量(Tensor)格式。随后, 为了加速模型收敛并防止梯度爆炸, 本文对像素值进行了归一化处理。

第一, 像素值缩放。将像素值从 $[0, 255]$ 线性映射至 $[0, 1]$, 第二, 标准化处理。由于本文模型需要使用预训练模型[2], 因此利用 ImageNet 数据集的统计均值和标准差, 对图像的 RGB 三个通道进行标准化计算。

2.2.4. 高级数据增强策略

为了解决医学图像数据集通常存在的样本量不足、类别分布不均以及模型过拟合等问题, 本文在训练阶段引入了基于 Albumentations 库的高级数据增强策略。采用 50% 概率的随机翻转和 $\pm 30^\circ$ 旋转模拟拍摄角度; 使用 CLAHE 增强局部细节并调整亮度对比度; 以 30% 概率执行 Coarse Dropout (随机遮挡), 迫使模型学习全局特征, 提升鲁棒性。

3. 本文模型介绍

3.1. 模型整体架构设计

为了有效解决皮肤镜图像分类任务中深层网络训练困难、特征表达能力不足以及过拟合等问题[3], 本文基于 ResNet34 架构, 提出了一种融合注意力机制与密集连接的增强型深度神经网络(Enhanced-ResNet34)。

本文模型在整体架构上保留了 ResNet 的残差学习思想[4], 但在细节模块上进行了深度优化, 见表 1。模型首先通过标准卷积层与最大池化层对输入图像进行初步特征提取[5]; 随后, 利用四个级联的 AdvancedBlock (改进型残差块) 构建深层特征提取网络; 在进入全连接层前, 引入了一个 DenseBlock (密

集连接块)以增强特征复用[6];最后通过全局平均池化与全连接分类器输出结果。该架构旨在通过多策略融合[7],提升模型对皮肤病变细微特征的识别能力。

Table 1. Network layer configuration table of the proposed model in this study

表 1. 本文模型网络层配置表

层级名称	输出尺寸	核心设置
Conv1	112×112	7×7 卷积核
MaxPool	56×56	3×3 最大池化
Layer1	56×56	3 个 AdvancedBlock
Layer2	28×28	4 个 AdvancedBlock
Layer3	14×14	6 个 AdvancedBlock
Layer4	7×7	3 个 AdvancedBlock
DenseBlock	7×7	3 层密集连接, 增长率 32, 输出通道 $512 + 3 \times 32$
Classifier	7 类输出	全局平均池化 + Dropout + Linear + Dropout

3.2. 模型效果评价指标

为了全面、客观地评估模型在训练过程中的学习状态及最终的预测性能,本文选取准确率(Accuracy)和损失(Loss)作为核心评价指标。

3.2.1. 准确率

$$\text{Accuracy} = \frac{N_{\text{correct}}}{N_{\text{total}}} \times 100\%$$

其中, N_{correct} 表示模型预测正确的样本数量, N_{total} 表示测试集中的总样本数量。准确率的取值范围为[0, 1], 数值越接近 1, 表明模型的分类性能越好。

3.2.2. 损失

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

其中, N 为样本总数, y_i 为第 i 个样本的真实标签, $\hat{y}_i$ 为模型预测该样本属于正类的概率。在训练过程中, 随着迭代次数(Epoch)的增加, 损失值应呈现下降趋势。

3.3. 核心组件设计

3.3.1. 自适应归一化层

针对医学图像数据量小、Batch Size 受限导致 Batch Normalization (BN)统计量不稳定的问题,本文设计了 Adaptive Norm 类, 优先采用组归一化策略。该组件将通道分为若干组, 在组内进行归一化计算。对于输入特征图 x , 其归一化公式为:

$$y = \frac{x - \mu_g}{\sqrt{\sigma_g^2 + \varepsilon}}$$

其中, μ_g 和 σ_g 分别为组 g 内的均值和标准差。这种设计摆脱了对 Batch 维度的依赖, 使得模型在小批次训练时依然能保持稳定的收敛性能。

3.3.2. 改进的残差块

传统的 ResNet 残差块在极深层网络中容易出现特征退化。本文构建的 AdvancedBlock 在标准残差路径上集成了 SE 注意力机制与随机深度正则化[8]。

引入 SE Layer 模块,通过全局平均池化和全连接层学习各通道的重要性权重。假设输入通道数为 C , 缩减比为 $r = 16$, 则权重向量 $s \in R^C$ 通过以下公式计算:

$$s = \sigma(W_2 \delta(W_1 z))$$

其中, z 为全局特征描述符, δ 为 ReLU 激活函数, σ 为 Sigmoid 激活函数。该权重与特征图相乘, 实现了对关键病理特征通道的增强。

随机深度, 在训练阶段, 以概率 p 随机跳过当前残差块, 直接输出恒等映射。这迫使网络在浅层也能学习到有效特征, 显著降低了过拟合风险。

3.3.3. 密集连接块

为了进一步加强特征传播并减少参数冗余, 本文在 ResNet34 的末端接入了一个 DenseBlock。包含 3 层 Bottleneck 结构卷积层。通过密集连接将前 L-1 层特征图拼接作为输入, 加强特征传播并减少参数冗余。

3.3.4. 分类器与损失函数

在特征提取完成后, 模型利用全局平均池化层将 7×7 的特征图转换为一维向量。随后, 通过一个包含两个 Dropout 层(丢弃率分别为 0.5 和 0.3)和两个全连接层的分类器进行最终判别。

Dropout 策略, 在全连接层之间引入高比例的 Dropout, 模拟集成学习的效果, 进一步防止过拟合。

损失函数, 采用交叉熵损失函数作为优化目标, 结合 AdamW 优化器进行参数更新。为了平衡不同类别的样本分布, 本文在训练中利用 ImageNet 的均值和标准差对输入图像进行了标准化预处理。

4. 模型逐步优化与泛化

4.1. 逐步优化策略与消融实验结果介绍

为了验证本文 Enhanced-ResNet34 模型的有效性, 实验采用分步迭代的优化思路, 从基线模型构建、迁移学习引入、骨干网络调整、数据增强优化、网络结构与参数优化五个维度逐步改进模型架构与训练策略。通过每一步单一变量的优化方式, 精准验证各策略对模型性能的提升效果。通过固定随机种子进行多次实验取平均值, 消除了随机性干扰。各阶段的平均准确率与损失值如下表 2 所示。

Table 2. Comparison of metrics during the stepwise optimization process of model

表 2. 模型逐步优化过程指标对比

阶段	策略变化	平均准确率	平均 LOSS
Step 1	随机初始化 EfficientNet-B3	70.2%	0.653
Step 2	加载 ImageNet 预训练权重	81.3%	0.452
Step 3	骨干网络更换为 ResNet34	83.6%	0.424
Step 4	引入 Albumentations 高级增强	88.5%	0.351
Step 5	引入 SE 注意力 + Bottleneck 结构优化 + 优化器调整	92.6%	0.185

4.2. 逐步优化策略

4.2.1. Step 1: 基线模型构建

本文第一步基于 EfficientNet-B3 网络构建基线模型, 采用随机初始化方式完成参数配置。EfficientNet 通过复合缩放策略统一调整网络宽度、深度和分辨率, 在参数量与计算量可控的前提下充分挖掘特征提取能力。选择随机初始化的核心目的是构建纯净的实验基线, 消除预训练权重等外部因素干扰, 为后续优化策略提供统一对比基准。

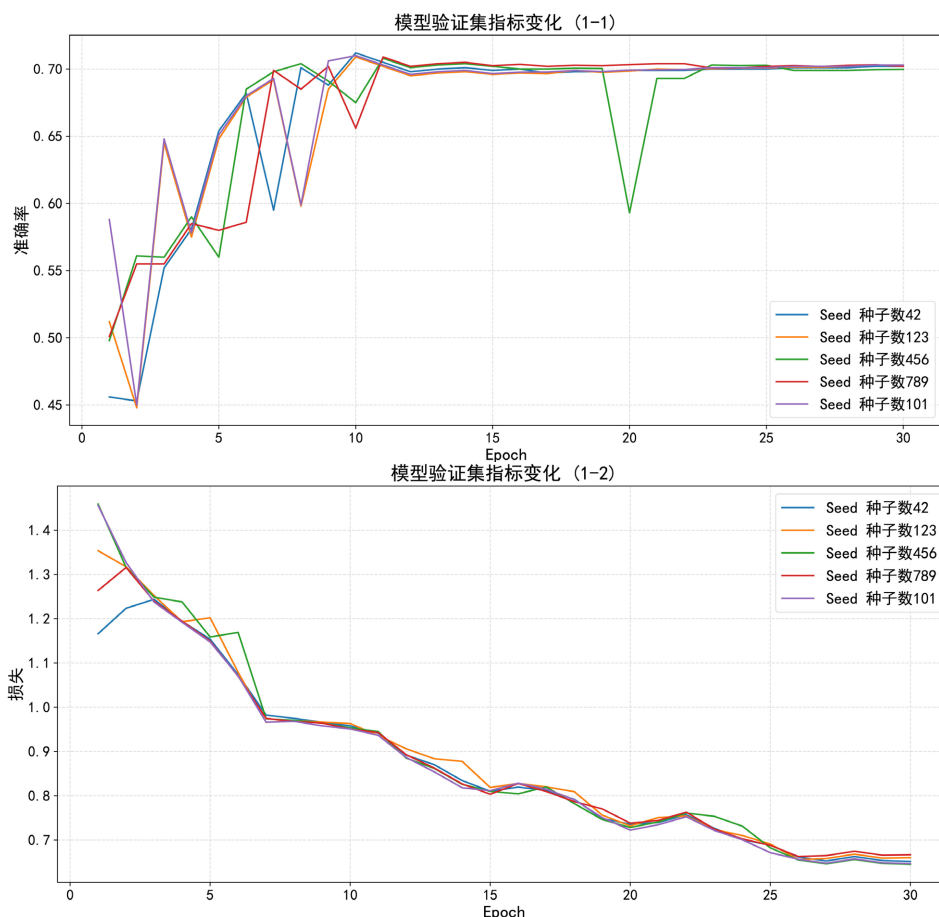


Figure 5. Validation accuracy (Top) and loss (Bottom) of the baseline model

图 5. 基线模型验证集准确率(上)与损失(下)

实验结果见图 5, 随机初始化的 EfficientNet-B3 基线模型在验证集上的平均准确率为 70.2%, 平均损失值为 0.653。从准确率曲线可以看出, 模型在训练初期(前 10 个 Epoch)表现出剧烈震荡与不稳定性, 不同随机种子下性能差异显著, 表明模型对初始化参数高度敏感。尽管损失值随迭代呈现下降趋势并趋于平稳, 但整体收敛速度较慢, 最终准确率未能突破瓶颈。这种表现揭示了仅依靠从零开始训练的模型缺乏通用图像特征先验知识, 特征提取能力受限, 难以充分拟合数据分布, 需要通过后续策略进行优化。

4.2.2. Step 2: 迁移学习的引入

针对基线模型从零训练、特征学习效率低的问题, 本阶段引入迁移学习策略, 为 EfficientNet-B3 加载 ImageNet 数据集预训练权重。ImageNet 包含海量通用图像数据, 覆盖纹理、轮廓、色彩等基础视觉特

征。预训练权重使模型无需从零学习基础图像特征，可直接复用成熟的通用特征提取能力，仅针对特定任务数据集进行微调，大幅降低训练难度与收敛成本，同时有效缓解小样本数据集下的过拟合与收敛缓慢问题。

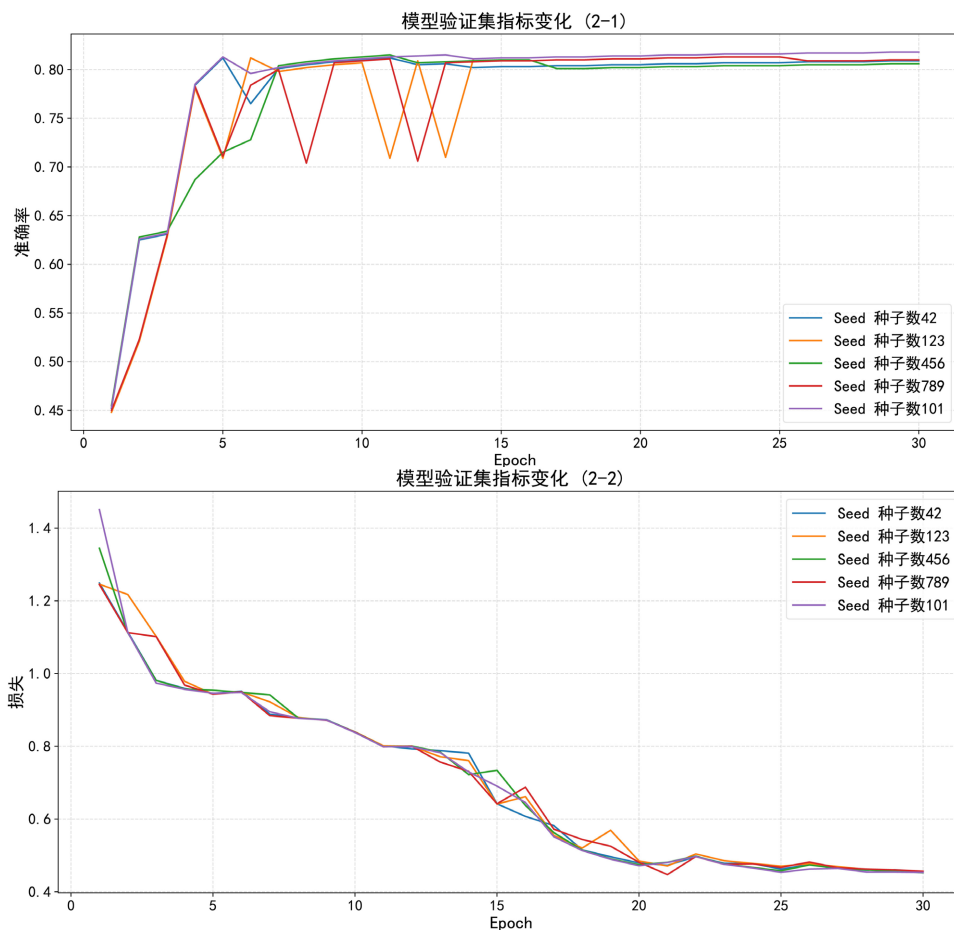


Figure 6. Validation accuracy (Top) and loss (Bottom) of transfer learning
图 6. 迁移学习验证集准确率(上)与损失(下)

实验结果见图 6，加载预训练权重的 EfficientNet-B3 模型性能实现了显著跃升。平均准确率从基线模型的 70.2% 提升至 81.3%，涨幅达 11.1%；平均损失值从 0.653 大幅下降至 0.452。观察训练曲线可见两个关键变化：第一，收敛速度显著加快，损失曲线在训练初期即呈现陡峭下降趋势，并在约第 10 个 Epoch 左右迅速趋于平稳，远优于基线模型；第二，训练稳定性增强，准确率曲线摆脱了初期剧烈震荡，不同随机种子的表现更加一致且快速攀升至高位平台期。这一结果有力证明，充分利用预训练权重中的通用图像特征先验知识，能够有效优化参数初始化状态，使其更快适应目标任务的数据分布，是突破性能瓶颈的高效策略。

4.2.3. Step 3: 骨干网络的适配性调整

为进一步提升模型在本文任务中的适配性，本阶段将骨干网络由 EfficientNet-B3 替换为 ResNet34。ResNet34 引入残差连接机制，有效解决了深层网络训练中的梯度消失、梯度爆炸及网络退化问题。相较于 EfficientNet-B3，其网络结构更简洁、梯度传播更稳定。同时，ResNet34 的层级特征提取结构能够分

层捕捉图像浅层纹理特征与深层语义特征, 对复杂场景、细微差异的图像特征识别精度更高, 更适配本文任务的数据分布特性。

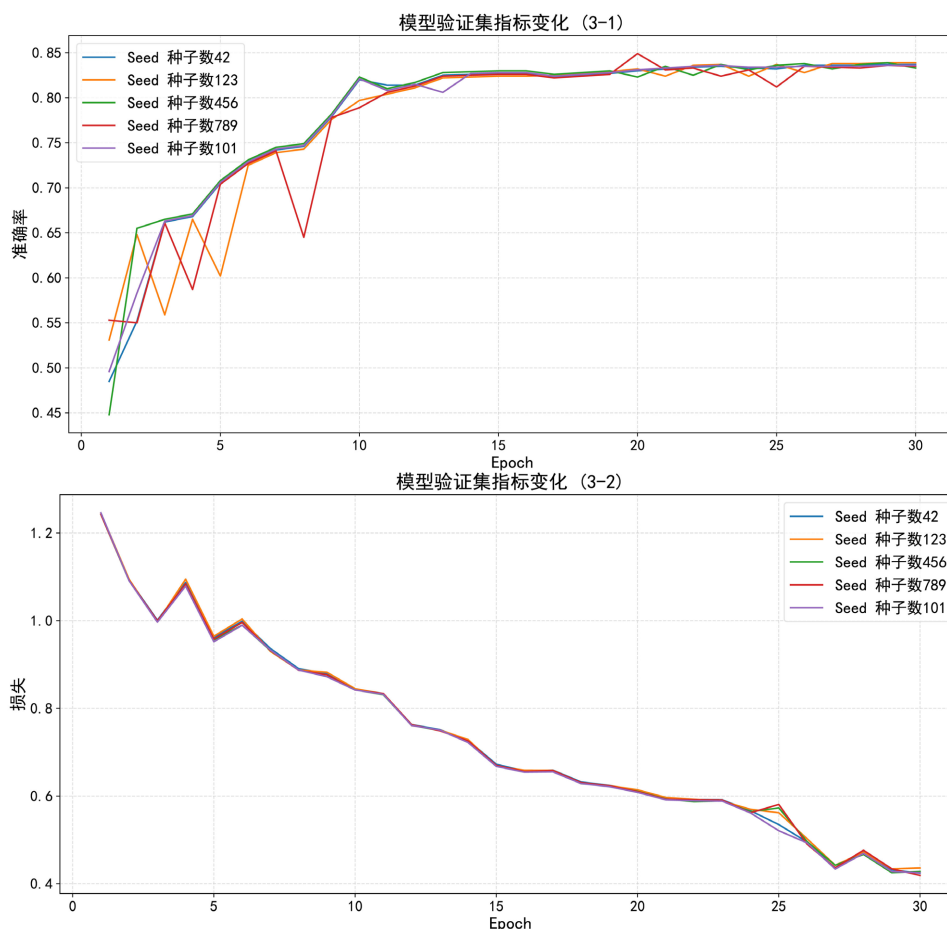


Figure 7. Validation accuracy (Top) and loss (Bottom) after network replacement
图 7. 更换网络验证集准确率(上)与损失(下)

实验结果见图 7, 经过本阶段优化, 模型平均准确率攀升至 83.6%, 较上一阶段提升了 2.3%; 平均损失值继续下探至 0.424。曲线呈现以下显著特征: 第一, 高位收敛与稳定性。尽管准确率起点较高, 模型依然保持了良好的上升势头, 且在训练后期(Epoch 20~30)表现出极高的稳定性, 未出现明显的过拟合反弹。第二, 深层优化的有效性。损失曲线呈现出平滑且持续的下降趋势, 最终稳定在较低水平。这验证了引入 ResNet34 作为骨干网络的有效性, 其特有的残差连接结构有效缓解了深层网络中的梯度传播问题, 增强了特征的复用与传递效率。通过优化骨干网络结构, 模型能够更精准地捕捉数据中的细微特征, 进一步缩小了预测值与真实标签之间的偏差。

4.2.4. Step 4: 数据增强策略

针对本文数据集样本场景单一、特征维度有限的问题, 本阶段引入 Albumentations 高级数据增强策略替代传统基础增强方式。Albumentations 支持随机色域变换、高斯模糊、噪声扰动、透视变换等多样化、高精度图像变换操作, 能够从空间、像素、色彩维度全方位扩充有效样本。同时, 该增强框架可实现动态实时增强, 在训练过程中持续生成差异化样本, 有效避免模型过拟合, 大幅提升模型的泛化能力与鲁棒性。

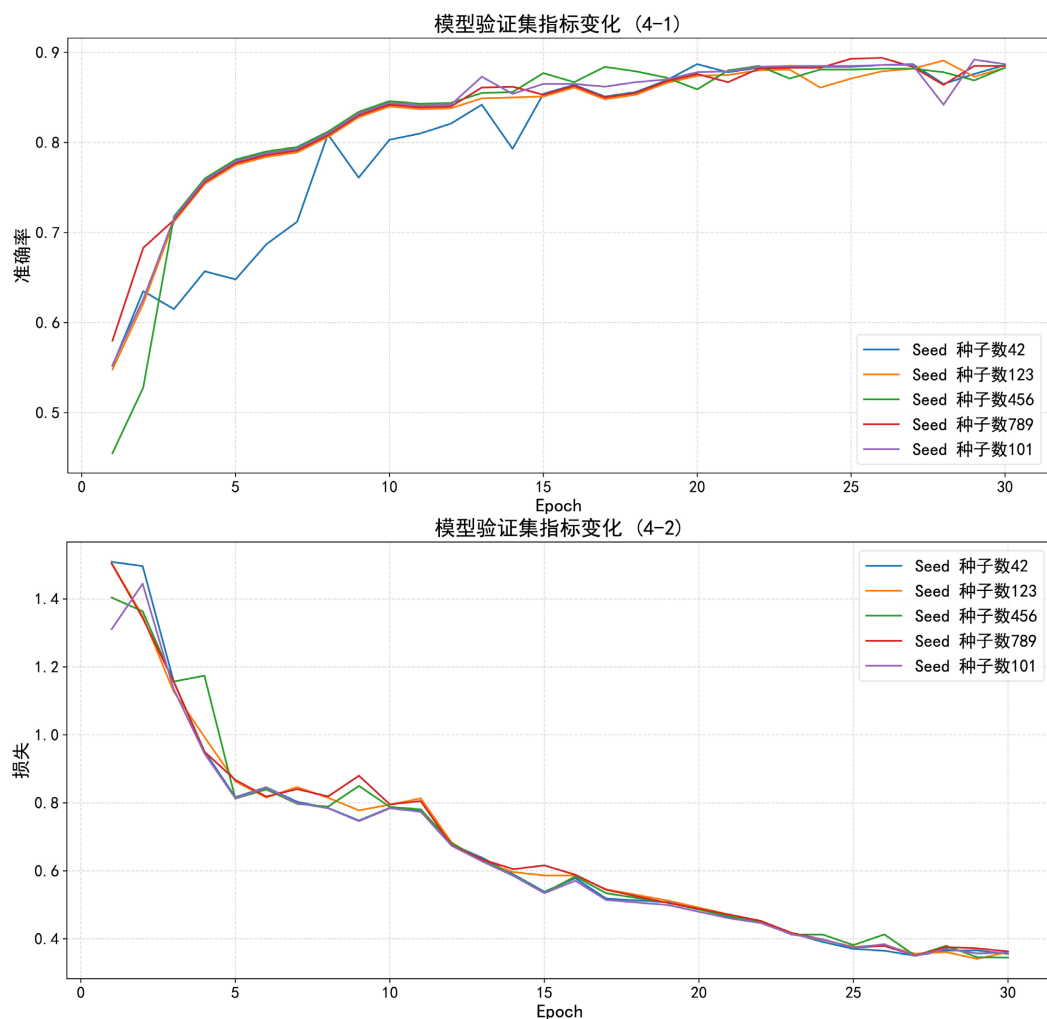


Figure 8. Validation accuracy (Top) and loss (Bottom) with data augmentation

图 8. 数据增强验证集准确率(上)与损失(下)

实验结果见图 8，通过丰富训练样本的多样性，模型平均准确率达到 88.5%，较上一阶段提升了 4.9%；平均损失值进一步降至 0.351。观察训练曲线可以发现：第一，泛化能力发生质变。准确率曲线在训练初期即表现出强劲的上升势头，且最终稳定在接近 0.9 的高位平台。不同随机种子下的曲线重合度较高，表明模型不再依赖特定的样本特征，而是学习到了更具鲁棒性的通用规律。第二，极低的收敛误差。损失曲线下降迅速且平稳，最终收敛至 0.4 以下的极低水平。这证实了高级数据增强有效弥补了原始数据集在数量和维度上的不足，迫使模型学习更本质的图像特征而非死记硬背，从而显著降低了预测误差。

4.2.5. Step 5: 网络结构深化

为突破模型上限，本阶段融合 SE 注意力机制、优化 DenseBlock 内部结构并调整优化器。其中，SE 注意力机制通过通道权重自适应分配，强化有效特征通道权重；针对 DenseBlock 部分，采用 Bottleneck 结构(1×1 卷积降维 + 3×3 卷积提取特征)替代普通卷积层，在保持特征提取能力的同时有效控制了参数增长量；搭配 AdamW 优化器策略，动态调整学习率。三者形成了特征精准化、结构高效化、训练稳定化的全方位优化体系。

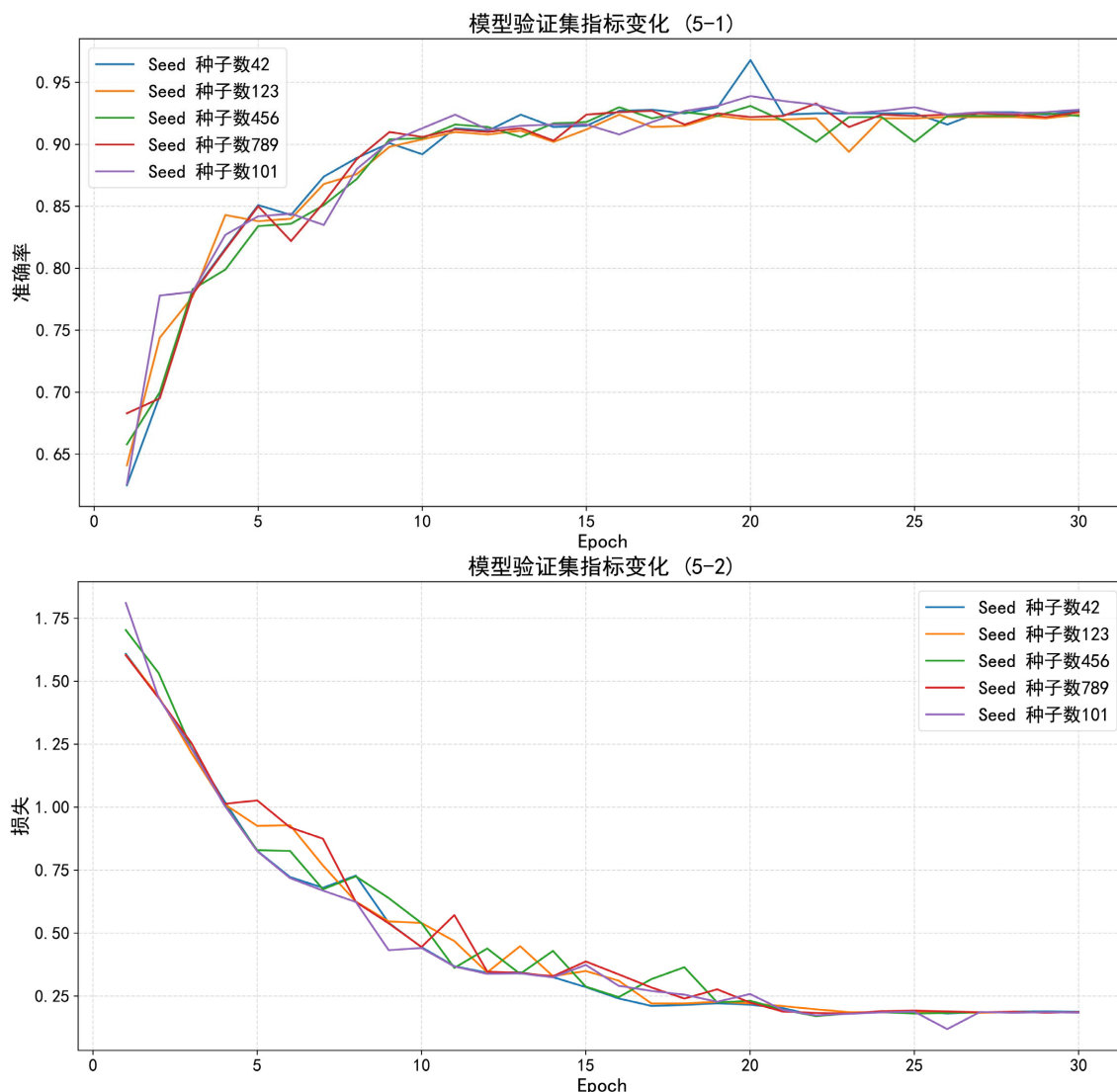


Figure 9. Validation accuracy (Top) and loss (Bottom) with deepened network structure

图 9. 结构深化验证集准确率(上)与损失(下)

实验结果见图 9, 多维度的网络优化策略产生了显著的协同效应。模型平均准确率攀升至 92.6% 的峰值, 平均损失值大幅降至 0.185。观察训练曲线可见以下决定性特征: 第一, 极致的预测精度。准确率曲线迅速突破 0.9 并稳定在高位, 不同随机种子下的表现高度一致且紧密收敛。这表明模型已经充分掌握了数据的核心分布规律, 具备了极强的特征筛选与识别能力。第二, 高效的收敛与泛化。损失曲线呈现出极其平滑的下降轨迹, 这证实了通过精简冗余计算和深化网络结构, 不仅加快了模型的迭代优化速度, 更有效抑制了过拟合风险, 实现了高精度与高稳定性的双重保障。经过这一系列综合改进, 模型成功克服了单一策略的局限性, 完成了从基础基线到高性能模型的完整蜕变。

为了进一步验证所提方法在细粒度分类任务中的有效性, 本文展示了最终模型在 HAM10000 数据集上的归一化混淆矩阵, 见图 10。从整体分布来看, 深蓝色高亮区域高度集中在主对角线上, 表明模型在 7 个类别上均取得了极高的分类准确率, 证明了经过多维度优化的网络结构能够有效提取皮肤病变的关键特征。特别是对于样本量较大的黑素细胞痣(NV)类, 显示出模型在处理主导类别时具有极强的鲁棒性。



Figure 10. Confusion matrix of the final model on the HAM10000 dataset
图 10. HAM10000 数据集最终模型混淆矩阵

4.3. 跨数据集泛化能力验证

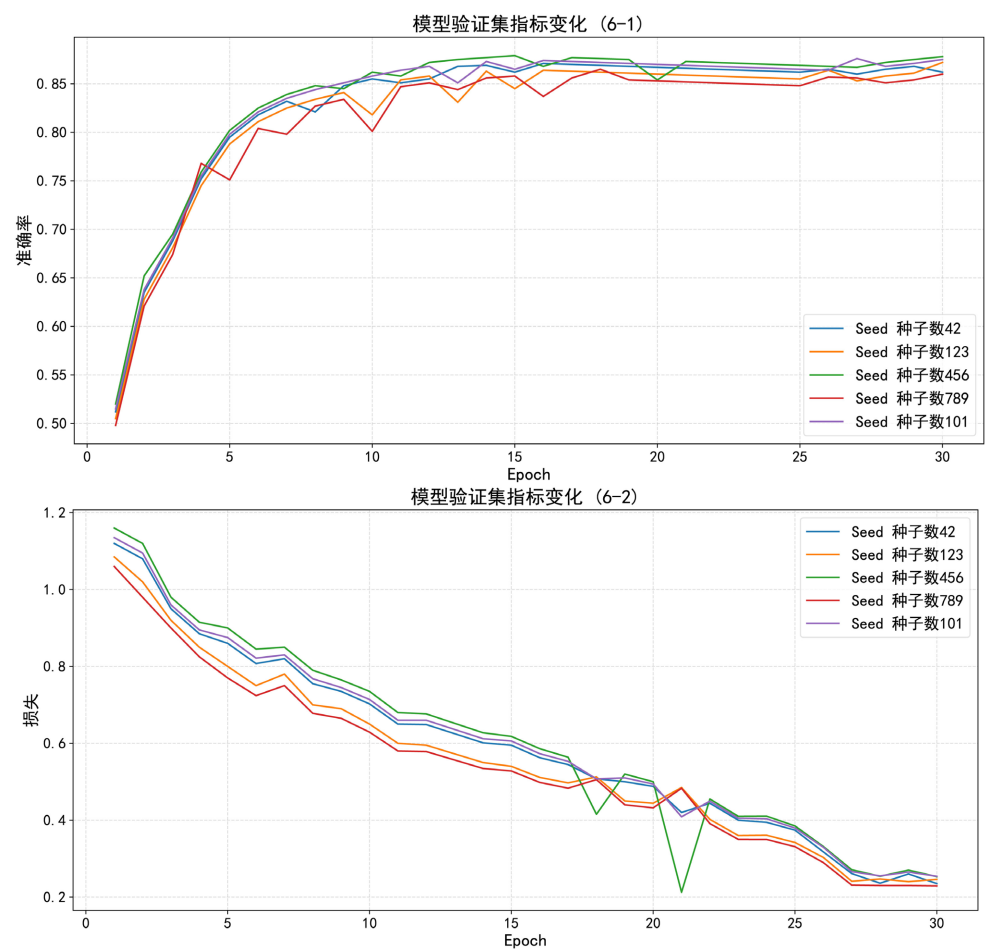


Figure 11. Validation accuracy (Top) and loss (Bottom) of the proposed model in this study on the ISIC2019 dataset
图 11. ISIC2019 本文模型验证集准确率(上)与损失(下)

为充分检验本文优化网络的通用适配能力, 消除单一数据集实验带来的结果偶然性, 本节开展跨数据集泛化验证实验。将最终模型迁移至 ISIC2019 数据集完成测试, 仅适配图像输入尺寸, 网络结构与训练超参完全保持不变, 以验证模型在不同样本规模、病变类别、图像分布下的迁移识别性能。

跨数据实验结果见图 11, 多维度网络优化体系具备良好的跨域协同效果。模型在 ISIC2019 验证集上平均分类准确率达到 87%, 平均损失稳定为 0.23, 显示出模型具有稳定可迁移的分类精度。准确率曲线稳步上升并快速收敛, 不同随机种子下曲线波动幅度小、性能表现高度统一。说明模型学习到的皮肤病变通用特征不受 HAM10000 数据分布约束, 能够自主适配 ISIC2019 新增病灶类型与海量样本特征, 完成跨数据集有效特征挖掘。在收敛特性方面, 损失曲线整体呈现下降趋势, 迭代后期无剧烈震荡, 最终稳定维持在较低水平。说明轻量化卷积结构、通道注意力与动态学习率优化器的协同作用, 能够缓解跨域数据偏移引发的收敛困难。

为进一步细化分析模型在多类别细粒度皮肤病变任务中的跨域分类效果, 本文绘制了 ISIC2019 数据集下模型归一化混淆矩阵, 见图 12。从混淆矩阵整体分布可以看出, 代表正确预测的深蓝色高亮区域主要集中在主对角线, 经综合优化的网络可在全新数据集下精准捕获各类皮肤病变独有判别特征, 完成稳定类别映射。



Figure 12. Confusion matrix of the final model on the ISIC2019 dataset

图 12. ISIC2019 数据集最终模型混淆矩阵

综上所述, 实验充分证明 SE 注意力、深度可分离卷积与优化器组合优化体系具备良好通用性, 模型兼顾轻量化、高精度与强泛化能力, 在不同皮肤影像数据集场景下均具备可靠的临床识别潜力。

5. 与 SOTA 模型的性能对比

为全面评估本文提出的 Enhanced-ResNet34 模型的有效性, 选取了两份近年(2025 年)在皮肤病变分类领域具有代表性的前沿研究(SOTA)进行横向对比(见表 3)。所有对比实验均严格遵循公平原则, 对比指标统一采用准确率(Accuracy)。

DermViT 模型[9], 立足临床医学诊断逻辑对传统视觉 Transformer (ViT) [10]进行针对性改进, 构建了适配皮肤镜图像分类的专属架构。相较于该模型, 本文模型在 HAM10000 数据集准确率高出 7.26%, ISIC2019 数据集准确率高出 8.23%, 在兼顾特征精准提取的同时, 大幅提升了复杂场景下的分类稳定性。

基于两阶段知识蒸馏的皮肤病分类方案[2], 目标是平衡精度与部署成本。相较于该轻量模型, 本文模型在 HAM10000 数据集上准确率最高超出 3.97%, 在 ISIC2019 数据集优势最高达 5.61%, 充分证明本文多模块融合优化策略, 相比单一知识蒸馏方案, 能够更高效地挖掘皮肤病变细微病理特征。

综合横向对比结果可得, 本文 Enhanced-ResNet34 模型在单一数据集精准分类与跨数据集泛化适配两类任务中均优于同期 SOTA 模型, 同时保留了 ResNet 系列模型结构简洁、训练稳定的优势, 既规避了 ViT 模型参数量大、泛化能力弱的缺陷, 也弥补了轻量化蒸馏模型精度不足的短板, 更适配临床皮肤病智能辅助诊断的应用场景。

Table 3. Accuracy comparison with two State-of-the-Art (SOTA) models
表 3. 与两类 SOTA 模型分类准确率对比

模型	核心配置	HAM10000 (ACC)	ISIC2019 (ACC)
改进的 DermViT	DermViT (DCP + DHA + DFG)	85.34%	78.77%
两阶段知识蒸馏	两阶段知识蒸馏 + 轻量学生模型	88.63%~88.89%	81.39%~83.42%
Enhanced-ResNet34 (本文模型)	SE 注意力 + 随机深度 + DenseBlock	92.6%	87%

6. 结语

本文针对皮肤癌早期精准诊断的需求, 提出了一种融合注意力机制与密集连接的增强型 ResNet34 模型(Enhanced-ResNet34)。通过在 HAM10000 数据集上的系统性消融实验, 准确率达到 92.6%, 通过在 ISIC2019 数据集上的跨域验证, 本文模型在未针对该数据集进行特定超参数调整的情况下, 依然取得了 87.0%的分类准确率。通过与近年来的 SOTA 进行横向对比, 本文模型在两个数据集上准确率均高于 SOTA。这一结果有力地证明了 Enhanced-ResNet34 模型具有优秀的泛化能力, 能够适应不同来源、不同分布的皮肤镜图像, 缓解了单一数据集训练可能导致的过拟合与域偏移问题。

未来的研究工作将着眼于进一步优化模型的轻量化设计, 以适配移动端或边缘计算设备的实时诊断需求, 同时探索利用生成式对抗网络或扩散模型解决医学图像中少数类样本稀缺的问题, 从而推动人工智能辅助诊断技术在临床实践中的落地应用。

参考文献

[1] 宋双, 连剑. 基于 Swin Transformer 增强的 Conformer 在皮肤癌分类中的应用[J]. 信息技术与信息化, 2025(5): 81-84.

[2] Hussain, M., Khan, M.A., Damaševičius, R., Alasiry, A., Marzougui, M., Alhaisoni, M., *et al.* (2023) SkinNet-INIO: Multiclass Skin Lesion Localization and Classification Using Fusion-Assisted Deep Neural Networks and Improved Nature-Inspired Optimization Algorithm. *Diagnostics*, **13**, Article 2869. <https://doi.org/10.3390/diagnostics13182869>

[3] 鲁硕, 张魁星. 基于深度学习的皮肤癌图像分类研究进展[J]. 计算机系统应用, 2026, 35(2): 23-39.

[4] 谢克玮. 基于自蒸馏算法的皮肤病图像分类的研究[D]: [硕士学位论文]. 赣州: 江西理工大学, 2024.

[5] Almufareh, M.F., Tariq, N., Humayun, M. and Khan, F.A. (2024) Melanoma Identification and Classification Model Based on Fine-Tuned Convolutional Neural Network. *Digital Health*, **10**, Article 20552076241253757.

[6] Bibi, S., Khan, M.A., Shah, J.H., Damaševičius, R., Alasiry, A., Marzougui, M., *et al.* (2023) MSRNet: Multiclass Skin Lesion Recognition Using Additional Residual Block Based Fine-Tuned Deep Models Information Fusion and Best Feature Selection. *Diagnostics*, **13**, Article 3063. <https://doi.org/10.3390/diagnostics13193063>

[7] Kumar, H., Dwivedi, A., Mishra, A.K., Shukla, A.K., Sharma, B.K., Agarwal, R., *et al.* (2024) Transformer-Based Decoder of Melanoma Classification Using Hand-Crafted Texture Feature Fusion and Gray Wolf Optimization Algorithm. *MethodsX*, **13**, Article 102839. <https://doi.org/10.1016/j.mex.2024.102839>

-
- [8] 秦传东, 索志强. 融合改进的 ResNet50 与集成分类器的皮肤癌分类[J]. 计算机应用, 2026, 46(4): 1354-1362.
- [9] Zhang, X., Liu, Y., Ouyang, G., Chen, W., Xu, A., Hara, T., *et al.* (2025) DermViT: Diagnosis-Guided Vision Transformer for Robust and Efficient Skin Lesion Classification. *Bioengineering*, **12**, Article 421.
<https://doi.org/10.3390/bioengineering12040421>.
- [10] 黄金杰, 马媛雪. 优化 ViT 用于黑色素瘤分类: 特征筛选与 InfoNCE 损失的结合[J]. 光学精密工程, 2025, 33(16): 2649-2660.