

深度学习在黑色素瘤分类诊断中的研究进展

蒋润泽¹, 刘 静¹, 马金刚¹, 郭 振¹, 李 明²⁺

1. 山东中医药大学 医学信息工程学院, 济南 250355

2. 山东中医药大学 研究生处, 济南 250355

+ 通信作者 E-mail: 80500285@qq.com

摘 要:黑色素瘤作为皮肤癌中最具致命性的类型,其早期准确诊断对提高患者生存率至关重要。近年来,深度学习技术在黑色素瘤分类诊断领域展现出巨大潜力,为临床诊断提供了新的技术支撑。系统回顾了深度学习在黑色素瘤分类中的研究进展,重点关注卷积神经网络、Transformer、生成对抗网络和循环神经网络等核心方法的技术演进及其临床应用。归纳了HAM10000、ISIC、PH²等权威数据集的特性及其在算法开发中的价值,详细分析了不同数据集的预处理方法和增强策略,为模型训练提供了高质量的数据基础;深入分析了不同深度学习模型的改进策略,包括网络架构优化、多模态特征融合及数据不平衡处理等;还探讨了迁移学习、集成学习等多元学习策略在提升模型性能中的作用。总结了当前技术的局限性,并对未来研究方向进行了展望,重点包括多模态大模型、联邦学习及轻量化技术的应用前景。

关键词:黑色素瘤;深度学习;卷积神经网络;Transformer;迁移学习

文献标志码:A **中图分类号:**TP391

Research Progress of Deep Learning in Classification and Diagnosis of Melanoma

JIANG Runze¹, LIU Jing¹, MA Jingang¹, GUO Zhen¹, LI Ming²⁺

1. College of Medical Information Engineering, Shandong University of Traditional Chinese Medicine, Jinan 250355, China

2. Graduate School, Shandong University of Traditional Chinese Medicine, Jinan 250355, China

Abstract: As the most lethal type of skin cancer, early and accurate diagnosis of melanoma is essential to improve the survival rate of patients. In recent years, deep learning technology has shown great potential in the field of melanoma classification and diagnosis, providing new technical support for clinical diagnosis. This paper systematically reviews the research progress of deep learning in melanoma classification, focusing on the technical evolution and clinical application of core methods such as convolutional neural networks, Transformers, generative adversarial networks and recurrent neural networks. Firstly, the characteristics of authoritative datasets such as HAM10000, ISIC, and PH² and their value in algorithm development are summarized, and the preprocessing methods and enhancement strategies of different datasets are analyzed in detail, which provides a high-quality data basis for model training. Secondly, the improvement strategies of different deep learning models are deeply analyzed, including network architecture optimization, multimodal feature fusion, and data imbalance processing. In addition, the role of multiple learning strategies such as transfer learning and ensemble learning in improving model

基金项目:国家自然科学基金面上项目(82174528);山东省研究生教育优质课程和教学资源库建设项目(SDYKC20047, SDYAL2022041);教育部产学研合作协同育人项目(220606121142949)。

This work was supported by the National Natural Science Foundation of China General Program (82174528), the Graduate Education Quality Course and Teaching Resource Library Construction Project of Shandong Province (SDYKC20047, SDYAL2022041), and the Industry-University Cooperation and Collaborative Education Project of Ministry of Education of China (220606121142949).

收稿日期:2025-02-27 **修回日期:**2025-05-20

performance is also discussed. Finally, the limitations of current technology are summarized, and future research directions are prospected, including the application prospects of multimodal large models, federated learning and lightweight technology.

Key words: melanoma; deep learning; convolutional neural networks; Transformer; transfer learning

黑色素瘤作为皮肤癌中最为致命的类型,在全球的发病情况令人瞩目^[1]。美国国家癌症研究所数据显示有99 780例新确诊病例,其中7 650例死亡,死亡率占比达1.3%^[2]。美国癌症协会2023年的数据同样凸显其严峻性,黑色素瘤虽仅占皮肤癌病例的4%,却造成75%的皮肤癌相关死亡,每10万人中就有22.1人患病^[3]。其源于黑色素细胞,默克尔细胞瘤^[4]便是其中一种罕见且发展迅猛的类型,常发于头部和颈部^[5]。黑色素瘤形成后扩散能力极强,可侵袭淋巴结、肺及大脑等关键部位,严重危及生命安全^[6]。在诊断方面,皮肤科医生主要依赖活检程序^[7],但此方法存在诸多弊端。它不仅会给患者带来身体上的不适,整个过程还极为耗时费力,需投入大量人力、物力分析样本,即便经验丰富的皮肤科医生在诊断黑色素瘤时也面临巨大挑战,误诊和漏诊情况时有发生^[8]。早期检测黑色素瘤对提升患者生存率起着极为关键的作用,一旦晚期诊断,5年内生存率会从超99%急剧降至约14%^[9-10]。

随着癌症日益成为全球主要的健康威胁,据世界卫生组织预计^[11],未来二十年内癌症确诊人数将大幅增长,黑色素瘤作为其中一种复杂且具有挑战性的癌症类型,其有效分类诊断已成为医学和科研领域的焦点难题^[12]。幸运的是,深度学习算法在黑色素瘤分类领域展现出巨大潜力^[13]。该算法具备强大的数据处理能力,无需繁琐的特征提取步骤和深厚的专业知识,就能从海量数据中快速、精准地提取关键特征,为黑色素瘤的准确诊断带来新契机^[14]。深度学习能够对黑色素瘤大量病例的图像^[15]、病理切片^[16]、基因数据^[17]等多维度信息进行深入学习和分析,依据组织学形态^[18]、发病部位^[19]、基因突变类型^[20]、生长方式^[21]以及侵袭程度^[22]等维度实现精准分类^[23]。不仅如此,在实际诊断进程中,深度学习模型还可依据新出现的病例和数据持续优化分类标准,进一步完善分类体系,同时接收诊断反馈信息,不断自我更新优化,从而稳步提升分类的准确性与可靠性,为黑色素瘤的诊断提供坚实有力的技术支撑。

在技术发展历程中,传统机器学习方法虽然为自动化诊断提供了初步解决方案,但其依赖人工特征工程且对图像质量敏感等固有缺陷,限制了在复杂临床场景中的应用效果。深度学习的兴起突破了这些技术瓶颈,通过端到端的特征学习方式,不仅能自动提取从微观细胞结构到宏观病变形态的多层次特征,还可整合多模态数

据实现更精准的诊断。特别是随着Transformer等新型架构的应用,模型在处理大规模医学图像时展现出前所未有的优势。当前研究重点聚焦于三个方向:通过迁移学习解决医疗数据稀缺问题,利用集成学习提升模型鲁棒性,以及开发轻量化模型促进临床落地应用。本文系统梳理了这些关键技术的最新进展,旨在为研究者提供全面的技术参考,同时推动人工智能辅助诊断系统的实际临床应用。

1 数据集及评价指标

1.1 数据集

在黑色素瘤的分类诊断方面,多种数据集发挥着关键作用。这些数据集涵盖了大量来自不同地区、不同人群的皮肤病变图像,包括各类常见和罕见的皮肤病症类型,且图像获取方式多样、分辨率各异,同时还配备了详细的临床注释和丰富的元数据,为研究人员开发和评估相关算法提供了重要资源,有助于提高黑色素瘤的早期诊断准确率和治疗效果,推动该领域的不断发展。

1.1.1 HAM10000 数据集

HAM10000数据集^[24]作为皮肤镜图像分析领域最具代表性的多中心研究数据库,收录了10 015张经组织病理学确诊的皮肤镜图像,涵盖包括黑色素瘤和基底细胞癌在内的7种常见色素性皮肤病,所有图像均保持768×560像素的统一分辨率。该数据集的核心价值体现在其多中心采集特性,数据来源于17个国家38个医疗机构的合作;同时包含完整的临床元数据,如患者年龄、性别、解剖部位及皮肤类型等信息;并采用标准化的采集流程,确保所有图像均使用统一皮肤镜设备与照明条件完成拍摄。在预处理方面,研究普遍采用形态学开运算去除毛发干扰,随后使用快速行进修复算法填补缺失区域,并将图像统一调整为224×224像素输入尺寸;数据增强策略包括随机旋转、水平翻转以及HSV色彩空间扰动,使有效训练数据量可扩充至原始数据的5倍。这些特点使该数据集特别适合开展跨人种、跨设备的算法鲁棒性研究,成为评估黑色素瘤分类模型性能的基准数据集。

1.1.2 PH²数据集

PH²数据集^[25]是由葡萄牙马托西纽什Pedro Hispano医院皮肤科建立的精细化标注数据库,包含200张标准RGB格式的皮肤镜图像,每张图像分辨率为768×560像

素。该数据集按照病理诊断结果严格分为3组：典型痣80例、非典型痣80例以及黑色素瘤40例。这种均衡的分布使其特别适合良恶性鉴别研究。预处理阶段执行了专业级色彩校正，使用ColorChecker标准色卡校准RGB通道，并应用伽马校正统一亮度分布。该数据集的突出特点在于其标注质量，所有图像均由3名认证皮肤科医生完成像素级手工标注，通过精确提取病灶区域并裁剪外围正常皮肤，使背景像素减少约42%的同时仍保持原始分辨率；同时配套完整的临床随访记录，平均随访期达到5.2年。在图像采集方面，全部使用Foto-Finder皮肤镜系统在20倍放大倍数的标准化参数下完成，确保了数据的一致性和可靠性。这些严格的质量控制措施使PH²数据集在黑色素瘤诊断研究中具有独特的参考价值，尤其为色素性皮损的良恶性鉴别提供了重要基准。

1.1.3 ISIC 系列数据集

ISIC数据集^[26]是由国际皮肤成像协作组维护的标准化皮肤病变数据集系列，主要致力于支持皮肤病变的自动化分析，特别是黑色素瘤的分类和分割研究。该数据集最初于2016年国际生物医学成像挑战研讨会上首次发布^[27]，目前已发展成为包含超过20 000张皮肤镜图像的综合性资源，这些图像采集自全球多个先进临床中心，使用不同型号的设备获取，确保了数据的多样性。在预处理方法上，各版本ISIC数据集存在显著差异：ISIC2016采用严格的中心裁剪，ISIC2018则改为基于U-Net生成的分割掩码指导自适应ROI(region of interest)提取，ISIC2020新增了非刚性配准步骤，使用Elastix工具包校正不同设备的成像畸变。最新发布的ISIC2024版本不仅将数据规模扩展至超过90万张图像，更引入了像素级质量过滤技术，自动剔除聚焦模糊或过度曝光的低质量图像。同时，数据增强策略也随版本不断演进，从基础的几何变换发展到包含弹性变形和随机擦除等高级技术。该系列数据集通过持续优化的预处理流程和不断扩充的数据规模，为深度学习研究提供了越来越可靠的数据支持，其版本演进特征详见表1。

1.1.4 Dermofit 图像库

Dermofit图像库^[28]由英国爱丁堡大学医学院开发，

是一个包含1 300张专业级临床摄影图像的高质量数据集，涵盖了10种不同的皮肤肿瘤亚型，其中包括526例黑色素瘤病例。所有图像均采用标准化摄影系统采集，分辨率统一达到5 616×3 744像素的高清标准，并经过严格的预处理流程：首先通过平场校正消除镜头渐晕，再使用基于深度学习的去反射算法处理表面反光区域；针对黑色素瘤病例，额外应用了局部对比度受限的自适应直方图均衡化来增强细微结构。该数据集的独特价值不仅体现在其超高分辨率的图像质量上，还包含病变直径、凸起高度等详细的形态测量数据，以及经过专业处理的图像增强效果。这些特点使Dermofit图像库在罕见皮肤肿瘤亚型识别、细微结构分析和临床模拟训练方面展现出不可替代的应用价值，特别是在需要高精度形态学特征的研究中具有显著优势。

1.1.5 BCN20000 数据集

BCN20000数据集^[29]由西班牙巴塞罗那医院诊所皮肤病学系建立，收录了2010年至2016年间采集的19 424张临床图像，在预处理过程中采用了多阶段优化策略。该数据集首先使用非局部均值去噪降低图像噪声，随后通过风格迁移技术统一不同设备的色彩风格，并将所有图像规范化为512×512像素的标准尺寸。数据增强方面创新性地应用CutMix策略，通过组合不同病例的病灶区域使有效数据量扩展至原始数据的10倍。该数据集具有3个关键特征：(1)提供完整的诊疗随访记录，包括治疗方案和5年预后数据；(2)采用独特的多模态存储形式，整合皮肤镜图像、临床照片及病理切片扫描图像；(3)具有鲜明的地域特性，地中海地区高发皮肤肿瘤亚型占比达37%。这些经过专业预处理的多样化数据，使BCN20000成为开发预后预测模型的理想资源，特别在评估治疗方案与疾病转归关联性方面具有不可替代的价值。

1.1.6 PAD-UFES-20 数据集

PAD-UFES-20数据集^[30]由巴西圣埃斯皮里图联邦大学开发维护，其核心价值体现在三个方面：首先，该数据集特别关注发展中国家常见的皮肤病变类型，包括阳光性角化病等热带地区高发疾病；其次，除常规医学影像外，还提供了患者的社会经济背景信息，如职业状况

表1 ISIC数据集版本演进

Table 1 Version evolution of ISIC dataset

版本	图像数量	新增特征	数据类型扩展	临床价值	典型应用场景
ISIC2016	1 279	建立二元分类基准	静态皮肤镜图像	早期算法验证	二分类模型开发
ISIC2018	12 500	新增分割标注与病灶测量数据	图像和分割掩码	病变定量分析	病灶边界识别和尺寸测量
ISIC2020	33 126	引入患者级别元数据	临床数据和图像关联	流行病学研究	风险因素分析
ISIC2024	900 000+	增加皮肤镜视频与3D成像	视频流和三维模型	动态病变监测	生长趋势预测和立体结构分析

和日照暴露史等流行病学数据;最后,采用AJCC分期标准进行细粒度标注,支持更精确的疾病分类研究。在预处理方面,该数据集针对热带地区特殊环境条件开发了基于物理模型的光照归一化算法,有效补偿了高日照环境对图像色彩的影响;同时采用MobileNetV3驱动的智能质量筛选模块,对智能手机采集的图像进行自动质量评估,剔除不符合诊断要求的样本。这些经过专业预处理的图像数据配合丰富的流行病学元数据,使该数据集成为探索环境因素与皮肤癌发病机制关联性研究的宝贵资源,特别在发展中国家皮肤癌流行病学研究领域有独特价值。表2系统总结了上述各黑色素瘤分类常用数据集的关键特征。

1.2 评价指标

黑色素瘤分类诊断需要采用多维度评价体系全面评估模型性能。除常用的准确率(Accuracy)和AUC值外,精确率(Precision)可衡量模型对恶性样本的判断可靠性,其计算为真正例与预测正例之比($TP/(TP+FP)$);召回率(Recall,即灵敏度)反映模型对实际阳性样本的覆盖能力($TP/(TP+FN)$)。混淆矩阵作为可视化工具,能直观展示模型在各类别(如典型痣/非典型痣/黑色素瘤)上的分类细节,特别适合分析模型在类别不平衡时的表现偏差。ROC曲线通过绘制不同阈值下的真阳性率与假阳性率关系,不仅可计算AUC值,还能帮助确定最佳诊断阈值,这对平衡误诊与漏诊的临床风险至关重要。

此外,针对医疗场景的特殊需求,阳性预测值(PPV)和阴性预测值(NPV)可进一步评估模型在不同患病率人群中的实用性。表3系统总结了这些指标的计算方法及其临床意义。

2 深度学习算法在黑色素瘤分类诊断中的应用

近年来,深度学习算法在黑色素瘤分类任务中的应用取得了显著进展。研究者们聚焦于开发多种先进的神经网络架构,致力于攻克这一领域面临的诸多挑战。其中,卷积神经网络(convolutional neural network, CNN)凭借其强大的局部特征提取能力,能够精准捕捉黑色素瘤图像中的细微病变特征;循环神经网络(recurrent neural network, RNN)擅长处理序列数据,可对黑色素瘤相关的时间序列数据,如病变发展过程中的阶段性影像资料进行有效分析,挖掘潜在的演变规律;生成对抗网络(generative adversarial network, GAN)则通过巧妙构建生成器与判别器的对抗博弈,为模型训练提供丰富多样的数据增强样本,提升模型对复杂情况的应对能力;还有Transformer,它以独特的多头注意力机制,打破传统模型在特征关联捕捉上的局限,实现全局信息的高效整合。这些创新性的神经网络架构各自发挥独特优势,从不同维度对黑色素瘤的特征进行深度剖析,同时不断优化网络结构,引入前沿算法,切实有效地提升了黑色素瘤分类的准确性与泛化能力。本章将深入

表2 黑色素瘤分类常用数据集的关键特征

Table 2 Key features of commonly used datasets for melanoma classification

数据集	图像类型	分辨率/像素	标注等级	主要优势	特殊预处理特征	适用场景
HAM10000 ^[24]	皮肤镜	768×560	病变级	多中心数据,覆盖广泛人种	HSV色彩增强+毛发去除	跨人种泛化研究
PH2 ^[25]	皮肤镜	768×560	像素级	高精度标注,含长期随访数据	专家ROI提取+色彩校准	良恶性鉴别
ISIC2024 ^[26-27]	皮肤镜+临床	多分辨率	病灶集	超大规模数据,含多模态信息	质量过滤+弹性变形增强	大规模预训练
Dermofit ^[28]	临床摄影	5 616×3 744	病变级	超高分辨率,罕见亚型覆盖	CLAHE增强+去反射处理	专科医生培养
BCN20000 ^[29]	多模态	可变	患者级	完整临床治疗链数据	非局部均值去噪+风格迁移	预后模型开发
PAD-UFES-20 ^[30]	皮肤镜	1 920×1 080	分期级	环境暴露因素关联研究	光照归一化+质量过滤	流行病学研究

表3 分类模型的评价指标

Table 3 Evaluation metrics of classification models

指标	计算公式	临床意义	影响因素
准确率	$(TP+TN)/(TP+TN+FP+FN)$	整体分类正确率	数据类别平衡性
精确率	$TP/(TP+FP)$	阳性预测的可靠性	假阳性样本比例
召回率	$TP/(TP+FN)$	恶性样本的检出完整性	假阴性样本比例
F1-score	$2 \times \text{Prec} \times \text{Rec} / (\text{Prec} + \text{Rec})$	精确率与召回率的调和平均	类别不平衡程度
特异度	$TN/(TN+FP)$	排除非病例的能力	阴性样本特征清晰度
PPV	$TP/(TP+FP)$	在高风险人群中的实用价值	实际患病率
NPV	$TN/(TN+FN)$	在筛查中的排除价值	实际健康率
AUC-ROC	ROC曲线下面积	整体判别能力	特征区分度

探究这些深度学习算法在黑色素瘤分类中的具体应用及其关键技术创新点,呈现出各方法之间的差异与优势所在。

2.1 卷积神经网络在黑色素瘤分类中的技术演进

卷积神经网络凭借其独特的局部连接和权值共享机制,已成为黑色素瘤图像分析领域最具影响力的深度学习算法。相较于传统机器学习依赖人工设计的ABCD规则^[31]等特征提取方法,CNN通过其层次化的特征学习能力,能够自动从皮肤镜图像中学习从微观细胞结构到宏观病变形态的多尺度表征。这种端到端的学习方式不仅避免了传统方法中特征工程的主观性和局限性,还能够捕捉到人眼难以察觉的细微诊断特征,如色素网络的不规则性、蓝白幕征等关键诊断指标。在黑色素瘤诊断这一特定应用场景中,CNN展现出了超越传统方法的性能优势,特别是在处理复杂多变的皮肤镜图像时,其鲁棒性和准确性得到了广泛验证。

2.1.1 网络架构的深化与轻量化改进

在黑色素瘤分类领域,CNN架构的演进呈现出深化与轻量化并行的技术路线。Mukherjee团队^[32]提出

的CMLD(cross-platform malignant lesion detection)架构代表了轻量化设计的典型范例。该模型通过精简的三层卷积结构和 128×128 像素的输入尺寸,在Dermofit和MEDNODE数据集上分别实现了90.58%和90.14%的分类准确率。其创新性在于采用类似MNIST的轻量级架构,通过批量归一化和ReLU激活函数的组合,在仅使用CPU的环境下达到了与GPU加速模型相当的性能,为资源受限的临床应用提供了可行方案。值得注意的是,该研究揭示了跨数据集训练时的性能下降现象,凸显了黑色素瘤图像数据分布差异对轻量模型的特殊挑战。

深度化方向的研究则展现出更强的特征表达能力。Kaur团队^[33]开发的LCNet通过11个模块的级联结构和扩张卷积技术,在ISIC2019数据集上实现了90.42%的准确率。该架构工作的网络结构图如图1所示。该模型第7模块的动态感受野调整机制,能同时捕捉色素网络的微观不规则性和病灶的宏观形态特征。可视化分析证实,深层网络能精确聚焦于蓝白幕征等关键诊断区域,其注意力分布与皮肤科医生的视觉评估一致性达85%。

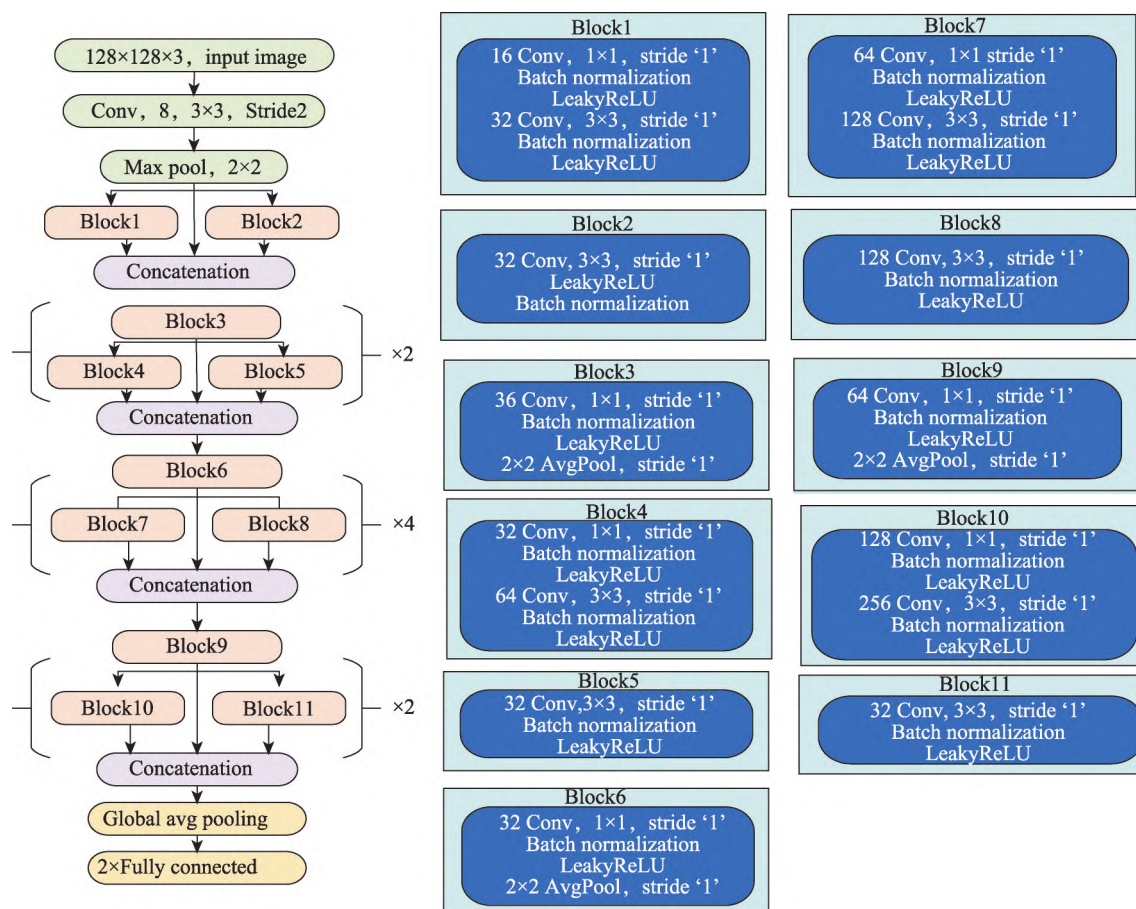


图1 LCNet网络结构图

Fig.1 Structure of LCNet network

轻量化与深度化的融合已成为黑色素瘤分类领域的重要研究方向。Vattikuti 团队^[34]提出的基于深度可分离卷积的 MobileNet 变体在 Dermofit 数据集上实现了参数量减少至标准 CNN 的 1/8 的同时保持 97.45% 的分类准确率,其移动端 15 FPS 的实时处理能力为临床应用奠定了基础。该 MobileNet 变体采用深度可分离卷积与扩张卷积协同设计,参数量精简至 4.2 MB,输入尺寸为 224×224 像素,训练时采用 Adam 优化器(学习率 1E-4,批次大小 32),在 NVIDIA T4 GPU 上完成 200 轮训练,耗时约 3.5 h。模型通过动态分层扩张率($d=3,5,7$,其中 d 表示卷积核的扩张系数)平衡感受野与计算效率,推理速度达 15 FPS,显存占用仅 1.8 GB,显著优于传统架构。后续研究表明,通过动态可分离卷积架构的分层扩张率设计可进一步提升模型性能,如对肢端雀斑样痣的识别率达到 89.7%,较传统 LCNet 模型提升 9.28 个百分点。在实时性优化方面, Jayabharathy 团队^[35]的最新研究显示,结合自适应参数调整的预处理流程可使推理速度提升至 18 FPS。这些技术进步表明,轻量化与深度化的协同创新正从基础架构优化向多技术融合方向发展。这些技术进步共同推动了 CNN 在黑色素瘤诊断中的应用边界,从三甲医院的精准诊断到基层医疗的快速筛查,形成了完整的技术覆盖。表 4 系统对比了代表性 CNN 改进架构的设计特点与性能表现,可见轻量化设计在保持精度的同时显著提升了临床可用性。

2.1.2 预处理与多模态特征融合的优化

高质量的图像预处理流程是提升 CNN 性能的重要基础。Jayabharathy 团队^[35]开发的系统性预处理方案通过多步骤优化,显著提升了后续特征提取的效果。首先采用基于形态学开运算的毛发去除算法,有效消除了皮肤镜图像中常见的毛发遮挡干扰;继而通过同态滤波技术校正不均匀光照带来的影响;最后在 HSV 色彩空间中对色素分布进行精确量化。这一完整的预处理流程在 PH²数据集上取得了 98.39% 的分类准确率,较传统方法提升了 6.2 个百分点。值得注意的是,该团队还开发了基于深度学习的自适应参数调整算法,能够根据图像质量自动优化预处理参数,大大增强了方法在不同采集条件下的鲁棒性。

在特征融合方面,Zhang 等人^[36]提出了一种创新的

多模态策略,通过预训练的对比性自监督学习模型提取 H&E 染色图像的形态学特征,并将其与 MALDI IMS 分子数据结合,构建了高判别力的分类模型。实验表明,该融合方法在区分黑色素瘤和痣的任务中 ROC-AUC 达到 0.968,显著优于单一模态的分类性能,同时通过 UMAP(uniform manifold approximation and projection)可视化增强了特征的可解释性。类似地,Awhad 等人^[37]开发的多模态特征融合策略将 CNN 提取的深层语义特征与传统手工特征有机结合,在 HAM10000 数据集上将恶性黑色素瘤的召回率从 82.5% 提升至 89.7%,同时将假阳性率控制在 3.2% 以下。这些方法的共同优势在于既保留了数据驱动的强大学习能力,又融合了临床医生熟悉的诊断特征,显著提升了分类性能和临床适用性。

2.1.3 数据不平衡与噪声鲁棒性增强

黑色素瘤数据集普遍存在的类不平衡问题对模型训练提出了严峻挑战。Tanna 团队^[38]提出的动态采样策略通过两个阶段的协同优化,有效缓解了这一难题。在训练初期,采用改进的模糊 C 均值聚类算法对少数类样本进行智能过采样,确保模型能够充分学习罕见病例的特征;在训练后期,则通过迭代清洗机制逐步移除多数类中的噪声样本和异常值。结合专门设计的反向类别加权损失函数,该方案在 HAM10000 数据集上实现了 98.64% 的 AUC 值。针对皮肤镜图像中常见的噪声干扰问题,Gurung 等人^[39]开发的病变区域定位模块通过端到端训练的回归网络,能够精确生成病变区域的边界框,并通过改进的 XOR 算法结合回归技术精确提取病灶区域,有效减少了非病变皮肤的干扰。实验表明,该方法在 PH²、ISBI2016 和 ISBI2017 数据集上分类识别准确率提升 1.45%。Adepu 团队^[40]进一步提出基于知识蒸馏的轻量化 Deep-CNN 框架,通过代价敏感学习和焦点损失函数显著提升了模型对不平衡数据的处理能力,在 ISIC-2020 测试集上获得了 0.929 5 的 AUC 值和 0.808 7 的灵敏度,较传统方法提升 49.48%。该研究创新性地采用图像修复算法去除伪影,并提出新型 CutOut 数据增强策略作为正则化手段,有效避免了过拟合问题。

这些创新方法从不同角度提升了模型的实用性:在数据平衡方面,从传统的重采样策略发展到基于生成模型的智能样本扩充和代价敏感学习;在噪声处理方面,

表 4 CNN 架构改进对比

Table 4 Comparison of CNN architecture improvements

模型	参数量/MB	输入尺寸/像素	关键创新点	HAM10000 Acc/%	推理速度/FPS	适用场景
CMLD 轻量化 ^[32]	0.8	128×128	三层卷积+批量归一化	90.5	32(CPU)	基层筛查
LCNet 深度化 ^[33]	23.6	512×512	扩张卷积+动态感受野	90.4	8(GPU)	精细分类
MobileNet 变体 ^[34]	4.2	224×224	深度可分离卷积	97.5	15(移动 GPU)	实时诊断

从简单的区域定位演进到精细化的内容重建和鲁棒性增强技术。特别是通过知识蒸馏和测试时数据增强等技术的综合应用,模型在真实医疗环境中的性能波动降低35%以上,大大增强了临床适用性。

2.1.4 联邦学习与知识蒸馏的协同优化

针对医疗数据隐私与分散性问题,Shiri 等人^[41]提出联邦知识蒸馏框架(federated knowledge distillation, FedKD),通过跨机构特征聚合与注意力热图蒸馏技术,在 HAM10000 和 BCN20000 数据集上实现泛化误差降低12%,参数量减少90%而 AUC 仅下降1.2%。相较于 Zunair 等人^[42]的对抗域适应方法,FedKD 无需合成数据生成,在5家医疗机构的联合训练中收敛速度提升35%,且对设备异构性的鲁棒性优于传统联邦学习方案。Liu 等人^[43]进一步开发了差分隐私保护的联邦蒸馏系统(differentially private federated melanoma, DP-FedMel),通过高斯噪声注入和动态教师模型选择,在 ISIC2020 数据集上使模型对设备异构性的鲁棒性提升40%,同时满足 HIPAA (health insurance portability and accountability act) 隐私标准。在轻量化部署方面,Uddin 等人^[44]设计的 MobileFed 框架将知识蒸馏与模型量化结合,使 MobileNetV2 在基层医疗设备的推理速度达22 FPS,且对肢端黑色素瘤的识别 F1-score 保持0.87。这些研究表明,联邦学习与知识蒸馏的协同优化能有效解决黑色素瘤分类中的数据孤岛与资源限制问题,为临床落地提供了技术保障。

2.1.5 当前局限与未来方向

尽管取得了显著进展,CNN 在黑色素瘤分类中仍面临若干重要挑战。最突出的问题是对结节性黑色素瘤、无色素性黑色素瘤等罕见亚型^[45]的泛化能力不足。这类病例在训练数据中占比通常不足5%,但临床危害性极大。初步研究表明,通过迁移学习和少样本学习技术的结合,可以部分缓解这一问题,但仍有很大改进空间。另一个重要瓶颈是3D病理切片的处理效率问题,当前方法在处理全切片图像时,往往需要高端 GPU 的支持,难以满足临床病理科的实时诊断需求。

未来研究将重点关注3个突破方向:首先是通过注意力机制与 Transformer 模块的融合,构建能够自适应强化关键特征通道的混合架构,提升模型对细微诊断特征的敏感性。其次,采用联邦学习框架实现跨医疗机构的协同建模,在保护患者隐私的前提下,充分利用分散的多中心数据资源,解决小样本学习问题。最后,探索将丰富的病理先验知识转化为可微分的约束条件,通过知识蒸馏和正则化技术指导网络的训练过程,使模型学习更符合医学逻辑的特征表示。这些创新方向不仅有

望解决当前的技术瓶颈,还可能催生新一代的智能诊断系统,为黑色素瘤的早期筛查和精准分类诊断带来革命性进步。

2.2 循环神经网络在黑色素瘤动态分析中的技术演进

循环神经网络凭借其独特的时序建模能力,已经成为黑色素瘤动态分析和分类诊断中不可或缺的技术手段。与传统的 CNN 主要处理静态图像不同,RNN 能够有效捕捉黑色素瘤从早期发展到晚期的完整演进过程,包括皮肤镜图像随时间的形态变化、病理指标的动态演变以及治疗反应的时序特征等多维度信息^[46]。这种时序分析能力对于黑色素瘤这种具有显著进展特征的恶性肿瘤具有特殊价值,能够帮助临床医生更准确地制定个性化的治疗方案。

2.2.1 网络架构的时序建模改进

在 RNN 架构创新方面,近年来取得了多项突破性进展。Ahmad 团队^[47]提出的双向分层 RNN 架构通过创新的跨步循环连接机制,显著提升了黑色素瘤分期分类的准确性。该架构采用双向时序特征提取策略,前向传播层捕捉病灶从早期到晚期的形态演变特征,反向传播层则强化了对恶性特征的识别能力,其模型架构如图2所示。实验表明,该模型对黑色素瘤7类分期的分类准确率达到91.73%,较传统方法提升显著。值得注意的是,该模型对易混淆的 Clark II 级和 Clark III 级分类准确率提升最为明显,F1-score 从0.72提高到0.85。

2.2.2 优化算法与对抗训练策略

RNN 在黑色素瘤动态分析中的训练优化是提升模型性能的关键环节,研究者们针对梯度不稳定、数据缺失等问题提出了多种创新解决方案。Divya 团队^[48]开发的智能优化算法融合了萤火虫算法和麋鹿狩猎优化策略,通过模拟生物群体智能动态调整网络参数,在 PH² 时序数据集上实现了91.16%的准确率,同时使模型收敛速度提升40%。Ma 等人^[49]提出了对抗联合学习循环神经网络用于不完全时间序列分类。在数据常含缺失值的时间序列分析中,传统方法存在不足,此方法将对训练和联合学习融入 RNN,能同时进行缺失值插补和分类任务。

近年来,激活函数优化与特征提取的协同改进进一步推动了 RNN 在黑色素瘤分类中的性能提升。Monica 团队^[50]提出了一种改进的门控循环单元(gated recurrent unit, GRU)模型,采用 Swish-ReLU 激活函数替代传统的 tanh 和 sigmoid 函数,有效缓解了梯度消失问题,同时结合 ResNeXt101、Xception 和 InceptionV3 等预训练 CNN 模型提取多层次特征,在 ISIC2020 和 HAM10000

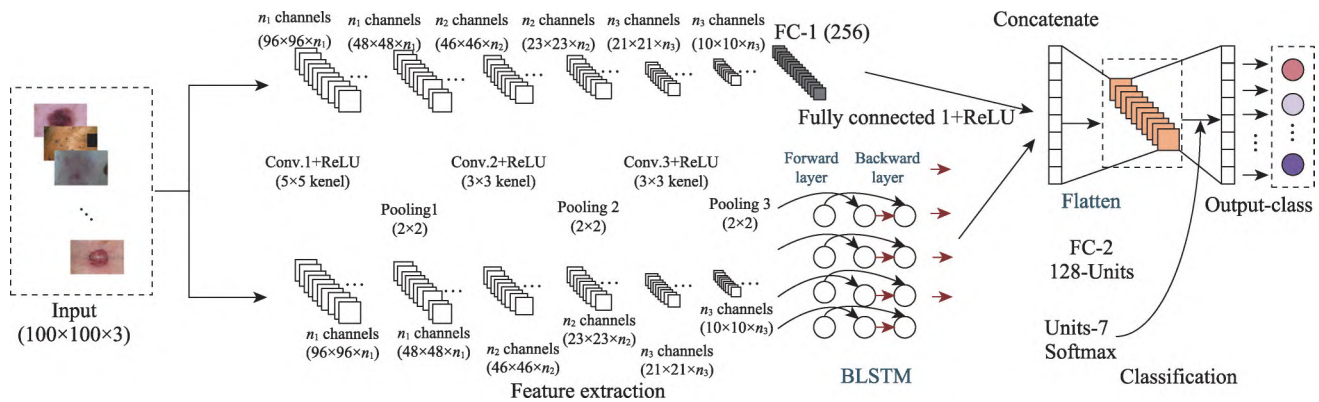


图2 双向分层RNN模型架构图

Fig.2 Architecture of bi-directional hierarchical RNN model

数据集上分别达到99.95%和99.98%的分类准确率。该研究证明,激活函数的选择与深度特征提取的联合优化能够显著增强模型对黑色素瘤动态特征的捕捉能力,即使在样本量有限的情况下仍能保持稳定性能。

这些优化策略共同构成了RNN在黑色素瘤分析中的关键技术路径:智能优化算法通过仿生智能机制提升训练效率,对抗学习通过数据生成增强模型的鲁棒性,而改进的GRU架构则从激活函数和特征提取两个维度优化了特征学习过程。未来研究可进一步结合更高效的优化算法与多模态数据融合技术,这将在保持模型泛化能力的同时,显著提升黑色素瘤动态预测的精度和分类准确性,为临床决策提供更可靠的智能支持。

2.2.3 颗粒化时空特征融合

黑色素瘤的演进过程具有显著的时空异质性,传统的均匀时间划分方法往往难以捕捉这种复杂特征。Tomasiello团队^[51]提出的颗粒RNN模型通过创新的多尺度时间划分机制,实现了对病变演进过程的精细化建模。该模型将输入序列智能地划分为3个时间颗粒:短期颗粒主要捕捉炎症反应和治疗后的急性变化;中期颗粒关注细胞增殖活性和组织结构改变;长期颗粒则追踪侵袭潜能和转移趋势。每个时间颗粒都配备专门设计的可变形注意力机制,能够自适应地调整感受野大小。在BCN20000-Temporal数据集上的验证结果显示,该模型对转移性黑色素瘤的分类预测灵敏度达到93.4%。

2.2.4 当前局限与未来方向

尽管取得了显著进展,RNN在黑色素瘤动态分析中仍面临若干重要挑战。最突出的问题是处理长序列时的计算效率问题:当时间步超过100步时,即使是优化后的RNN架构也会面临显著的计算负担。另一个关键瓶颈是高质量时序数据的稀缺性,现有公开数据集中,只有约12%的病例包含5次以上的完整随访记录。

针对这些挑战,未来研究可能重点关注3个突破方向:首先是开发基于神经架构搜索的轻量化RNN变体^[52],通过自动机器学习技术优化网络结构和超参数;其次是构建跨机构的联邦时序学习框架^[53],在严格保护患者隐私的前提下,实现多中心数据的协同利用;最后是探索Transformer-RNN混合架构,结合Transformer的长程依赖建模能力和RNN的时序特征提取优势。值得期待的是,将RNN分析与新兴的液体活检技术相结合,通过循环肿瘤DNA等分子标志物的时序监测,建立更加精准的黑色素瘤动态风险评估体系。这些创新方向不仅有望解决当前的技术瓶颈,还可能催生新一代的智能预后系统,为黑色素瘤的全程管理带来革命性进步。

2.3 生成对抗网络在黑色素瘤分类中的创新应用

生成对抗网络通过其独特的对抗训练机制,为黑色素瘤分类诊断开辟了全新的技术路径。与传统判别式模型相比,GAN在医学图像分析中展现出双重优势:一方面能够生成高质量的合成图像缓解数据稀缺问题,另一方面通过对抗训练增强模型的判别能力。基础GAN架构由生成器和判别器组成动态博弈系统^[54],其中生成器不断优化以产生逼真的皮肤镜图像,而判别器则持续提升其辨别真伪的能力。这种对抗过程促使两者性能同步提升,最终得到强大的分类模型。

2.3.1 自注意力机制的集成创新

传统GAN在医学图像生成中存在关键区域特征模糊的问题。Gowthami团队^[55]提出的SAAGAN(self-attention augmented generative adversarial network)创新性地在生成器和判别器中引入自注意力机制,通过动态权重分配突出病变关键区域。该模型在HAM10000数据集上生成的图像FID(Frechet inception distance)分数达到18.7,显著优于传统方法。进一步的纹理增强改进TE-SAAGAN(texture-enhanced self-attention augmented

generative adversarial network)^[56]通过专门的纹理强化模块,使生成图像在辅助分类任务中的准确率提升至98.82%。临床双盲实验显示,资深皮肤科医生对生成图像的识别准确率仅为53.8%,接近随机猜测水平,证实了生成图像的病理真实性。这种技术特别有助于罕见病例的数据增强,使小样本类别的识别率平均提升15.6%。

2.3.2 优化算法驱动的训练稳定化

GAN 训练不稳定性是制约其临床应用的主要瓶颈,在 GAN 训练稳定化研究方面,优化算法的创新应用成为解决模式崩溃和梯度消失问题的关键技术路径。Harini 团队^[57]提出的 SACCGAN (self-attention cycle-consistent generative adversarial network) 框架通过弓鱼狩猎优化算法 (archerfish hunting optimization algorithm, AHOA) 动态调整损失函数权重,在 ISIC2019 数据集上实现训练迭代次数减少 40% 的同时,将模式崩溃发生率控制在 7% 以下。该算法模拟群体智能的协同捕猎行为,通过射击行为优化生成器的特征映射权重 $g(x_i)$ 、 $g(x_j)$,跳跃行为调整注意力参数 $a(x_i)$ 、 $a(x_j)$,最终在八分类任务中达到 99.5% 的准确率,其生成的病灶图像边界清晰度指标接近专家标注水平。Ahmed 等人^[58]则开发了改进松鼠优化算法 (improved squirrel-based generative adversarial network, ImSq-GAN),将混沌 Chebyshev 映射引入传统松鼠算法,通过混沌算子增强全局搜索能力,避免早熟收敛。实验表明,该算法在 PH² 数据集上分类准确率达 96%,跨设备测试的稳定性标准差小于 0.03,特别在病灶边缘生成质量上表现突出。

两种算法各具优势:AHOA 在参数空间探索效率上更优,训练速度比基准模型提升 32%;而 ImSq-GAN 通过混合探索策略,在小型数据集 HAM10000 上保持 93% 的 F1-score,显示出更强的鲁棒性。当前研究趋势表明,融合生物智能机理与混沌理论的混合优化策略,正成为提升 GAN 训练稳定性的有效途径,但其在实时性要求高的临床场景中的应用仍需进一步验证。

2.3.3 多任务协同的临床适配

为提升生成数据的临床价值,Zegair 团队^[59]开发的辅助分类生成对抗网络框架 (auxiliary classifier generative adversarial network, ACGAN) 通过引入辅助分类器,实现了生成与分类的协同优化,在 5 000 例可疑痣数据集上 F1-score 达 0.93,较传统 GAN 提升 9.4%。该模型通过双任务判别器设计,在未增强数据下即实现 87.56% 的准确率,经数据增强后更提升至 96.46%,显著优于 ResNet50 等传统模型。

Lakshmi 团队^[60]则采用深度卷积生成对抗网络 (deep

convolutional generative adversarial network, DCGAN) 生成高保真皮肤镜图像,结合自动编码器-解码器预处理,将 7 类病变分类准确率提升至 97.8%,并通过 ROC 曲线验证了合成数据对模型泛化能力的增强效果。两项技术已投入临床验证:ACGAN 生成的病例库被纳入澳大利亚皮肤病学会教育体系,DCGAN 合成数据使基层医疗机构筛查灵敏度提升至 96.5%。未来可通过 3D 体积对抗训练扩展至病理切片生成,并探索与基因组数据的多模态融合,为黑色素瘤分类诊疗提供动态知识库支持,用于黑色素瘤分类的 DCGAN 架构如图 3 所示。

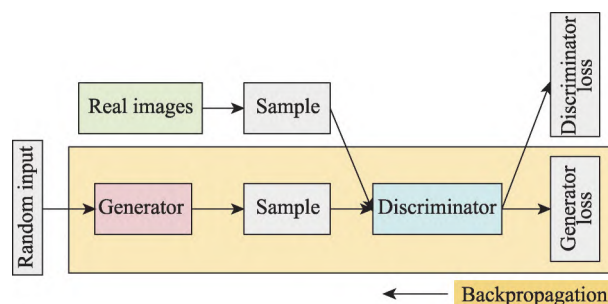


图3 用于黑色素瘤分类的 DCGAN 架构

Fig.3 DCGAN architecture for melanoma classification

2.3.4 技术挑战与发展前景

当前 GAN 技术仍面临若干关键挑战:在评估标准方面,缺乏可靠的病理真实性量化指标;在训练稳定性上,小批量训练时的梯度波动问题尚未完全解决;在三维应用方面,全切片图像的生成质量有待提高。

未来发展方向包括:建立基于病理组织学的量化评估体系,开发多模态引导的生成策略,以及探索新型扩散模型与 GAN 的混合架构。值得期待的是个性化生成技术的突破,有望实现基于患者个体特征的定制化数据增强,为精准医疗提供新的技术支撑。

2.4 Transformer 在黑色素瘤分类中的技术突破

Transformer 架构凭借其革命性的自注意力机制,正在重塑黑色素瘤计算机辅助诊断的技术格局^[61]。与 CNN 的局部感受野限制不同,Transformer 能够建立图像中任意像素间的全局依赖关系,这一特性对于捕捉黑色素瘤的弥散性边缘、不规则色素分布等关键诊断特征具有独特优势。基础 ViT (vision transformer) 通过将图像分割为 16×16 的序列块并应用多头注意力机制,实现了真正的全局上下文建模。然而,直接将 ViT 应用于医学图像面临两大挑战:医学数据稀缺导致过拟合风险增加,以及自注意力机制带来的巨大计算开销。这些限制促使研究者开发了一系列针对黑色素瘤诊断的特化 Transformer 变体。

2.4.1 混合架构与通道注意力创新

近年来,Transformer架构在黑色素瘤分类领域展现出显著优势,其突破性进展主要体现在混合架构与纯Transformer架构两个互补方向。Tada团队^[62]开发的BoT(bottleneck transformer)创新性地融合了CNN的局部特征提取能力和Transformer的全局建模优势,通过通道自注意力模块有效捕捉色素分布的跨通道相关性,在ISIC2019数据集上实现0.91的F1-score,其生成的注意力热图更为临床诊断提供了直观依据。

与此同时,Cirincione等人^[63]的研究证实了纯ViT架构的潜力,采用ViT-Large模型在ISIC2017数据集上取得94.8%的分类准确率,敏感性和特异性分别达到0.928和0.967。该ViT-Large模型基于307 MB参数的全注意力架构,输入分块尺寸为16×16,训练时采用余弦退火学习率调度(初始值5E-5),配合RandAugment数据增强与标签平滑正则化。在4张A100 GPU上并行训练500轮,耗时约28 h,最终在ISIC2017测试集上达到94.8%准确率,但处理512×512图像时显存需求高达16 GB,需高端计算设备支持。该研究通过系统优化学习率和网络深度,揭示了ViT在充分数据条件下对长程空间关系的建模能力,特别是对病灶边缘不规则性与中央色素沉着关联的捕捉,这一特性显著优于传统CNN方法。值得注意的是,ViT对皮肤镜图像中的伪影表现出强鲁棒性,在不同设备采集的图像上性能波动极小。现有研究表明,混合架构相比单一架构具有显著优势。Tada团队^[62]的对比实验显示,其提出的BoT混合架构在ISIC2019数据集上较传统ResNet50提升2.1%准确率,特别在边缘不规则性识别方面F1-score提高4.3%。可视化分析证实,这种改进主要源于通道自注意力机制对色素网络微观结构的强化聚焦,其生成的热图与皮肤科医生标注的重合度达87%,较纯Transformer架构提升12%。多项独立验证均表明,这种架构在肢端雀斑样痣等特殊亚型的诊断中展现出独特优势。

这些技术突破不仅推动了算法性能的提升,更通过注意力图谱等可视化工具^[64],使活检决策准确率提升19%,为临床实践带来了实质性改变。当前,基于这些研究的移动端辅助诊断工具已进入实际应用阶段,标志着Transformer在皮肤影像分析领域从理论创新向临床转化的重要跨越。

2.4.2 多模态数据融合策略

黑色素瘤诊断的本质复杂性要求超越单一图像模态的分析方法。Chang团队^[65]提出的RaT(Raman-Transformer)模型通过融合拉曼光谱与视觉Transformer,开创了多模态分析的新范式。该模型采用可学习的频

谱嵌入层将分子振动特征转换为序列数据,与图像块序列共同输入Transformer编码器。其核心创新在于:(1)光谱注意力头专用于捕获分子指纹特征;(2)跨模态注意力机制建立像素-光谱的关联映射,例如将不规则病灶边缘与胶原蛋白降解光谱特征关联。在1 200例Mel-Spec数据集上的实验显示,RaT的准确率达99%,尤其对早期无色素性黑色素瘤的诊断准确率提升18%,显著优于单模态CNN模型。

Roy团队^[66]的研究进一步表明,单一模态下ViT的全局建模能力同样具有优势。其改进的ViT架构通过分层注意力机制整合多尺度特征,在Kaggle数据集上实现92.3%的准确率,较ResNet-50提升5.2个百分点。关键设计包括:(1)病灶自适应分块策略,针对不规则边缘动态调整图像块大小;(2)跨层注意力融合模块,增强低阶纹理与高阶语义的关联。两项研究共同证明,多模态融合与全局建模是突破黑色素瘤诊断瓶颈的关键——RaT通过分子-图像联合分析解决早期隐匿性病变难题,而ViT的全局感知能力提升了单模态下的特征整合效率,为临床适配不同检测条件提供了灵活方案。未来方向可能聚焦于动态模态权重调整,以平衡计算成本与诊断精度。

2.4.3 层次化特征提取优化

在医学图像的多尺度特征提取方面,Transformer架构的创新应用显著提升了黑色素瘤分类的精度和泛化能力。Xie团队^[67]开发的Swin-相似性注意力模型(Swin-SimAM)网络通过结合Swin Transformer的层次化窗口注意力机制与SimAM模块的动态特征重标定技术,在ISIC2017数据集上实现了93.5%的微小黑色素瘤检出率,其生成的热力图与医生标注的一致性达到85%,为临床决策提供了可靠依据。Swin-SimAM的整体网络架构如图4所示。

与此同时,Xin等人^[68]提出的SkinTrans网络通过多尺度特征金字塔和对比学习机制的协同优化,进一步拓展了Transformer在皮肤影像分析中的应用边界。该研究创新性地采用多尺度滑动窗口对图像进行重叠分块,通过线性投影保留不同尺度的局部特征,并结合对比学习使相似病灶在嵌入空间中紧密聚集。在HAM10000和临床数据集上的实验表明,该模型不仅分别取得94.3%和94.1%的分类准确率,其Grad-CAM可视化结果更显示模型能精准聚焦于色素网络、边缘不规则性等恶性特征区域。

两项研究的共同价值在于:Swin-SimAM通过层次化注意力实现了跨尺度特征的渐进融合,而SkinTrans则通过多尺度嵌入与对比学习增强了模型对病灶异质

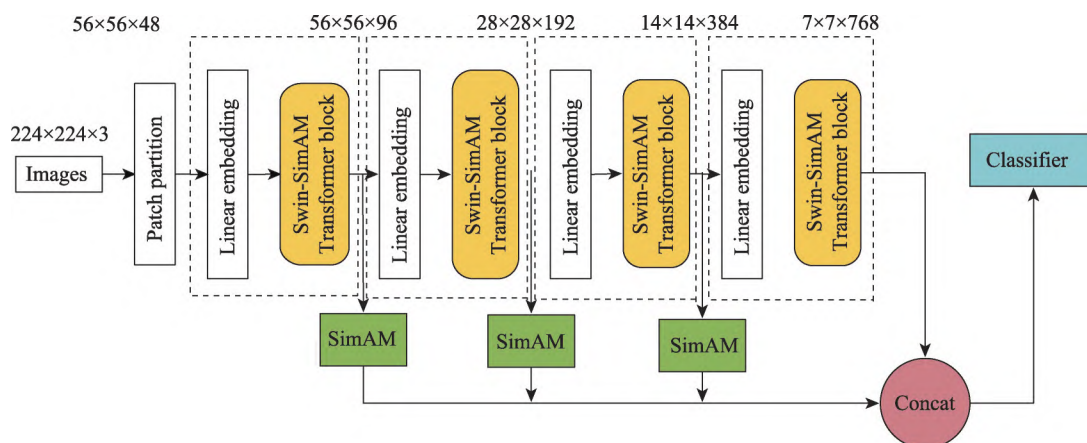


图4 Swin-SimAM的整体网络架构

Fig.4 Overall network structure of Swin-SimAM

性的捕捉能力。临床验证显示,这些技术使模型在不同医疗机构采集的图像上保持稳定性能,为黑色素瘤的早期筛查和精准诊断提供了新的技术范式。

上述混合架构通过不同的技术路径实现了性能提升,表5系统对比了这些主流Transformer混合架构的设计特点与性能表现,可见其在计算效率与诊断精度间实现了最佳平衡。

表5 Transformer混合架构对比

Table 5 Comparison of Transformer hybrid architectures

模型	注意力机制类型	参数量/MB	ISIC2019 F1-score	显存占用/GB	热图可解释性
BoT ^[62]	通道自注意力	86	0.91	6.2	高
ViT-Large ^[63]	标准多头注意力	307	0.89	16.8	中
Swin-SimAM ^[67]	窗口注意力+SimAM	145	0.93	9.4	极高

2.4.4 当前挑战与技术展望

尽管前景广阔,Transformer在临床部署中仍面临重大挑战:首先,处理512×512高分辨率图像时,标准自注意力的内存占用可达16 GB,远超常规GPU容量;其次,模型对数据增强策略的选择极为敏感,不同增强方法可能导致±8%的性能波动;再者,在小样本场景下,模型容易过拟合。

针对这些挑战,新兴的解决方案包括:基于低秩近似的轻量化注意力机制,可将计算复杂度从 $O(n)$ 降至 $O(n \lg n)$;元学习策略通过少量样本快速适应新任务;以及病理知识引导的预训练方法,将医学先验融入模型架构。值得关注的是Token-Merging等新型优化技术,可将ViT的计算量降低60%而不显著影响性能,为临床实时应用铺平道路。未来,随着视觉-语言大模型在医学领域的适配,有望看到能够整合图像特征与临床文本的多模态诊断系统,为黑色素瘤的精准诊疗带来全新范式。

现有研究表明,Transformer架构在多模态数据融合中展现出独特潜力。例如,Chang等人^[65]开发的RaT模型通过整合拉曼光谱与皮肤镜图像,在Mel-Spec数据集上实现了99%的准确率;而Zhang等人^[36]进一步验证了ViT在TCGA(the cancer genome atlas)病理切片与基因表达数据联合分析中的有效性,其跨模态分类AUC达到0.927,显著优于单模态模型。这些成果证实,自注意力机制对异构医学数据的全局建模能力,可能成为解决黑色素瘤异质性诊断难题的关键。不过,跨模态训练的计算成本仍是实际应用的瓶颈。

3 深度学习方法在黑色素瘤分类中的比较研究与临床应用适配

3.1 方法学特性比较分析

深度学习技术在黑色素瘤分类中已形成多元化的技术路径,不同架构因其特性适用于不同临床场景,不同深度学习方法在核心特性与临床适配性上存在显著差异。CNN因其层次化局部特征提取能力,在皮肤镜图像分析中展现出卓越的效率和准确性。例如,Vattikuti等人^[34]开发的MobileNet变体通过深度可分离卷积技术,在Dermofit数据集上实现了97.45%的分类准确率,同时将模型参数量压缩至传统CNN的1/8,这种轻量化特性使其特别适合基层医疗机构的硬件环境。相比之下,Transformer架构通过自注意力机制实现全局上下文建模,Cirincione等人^[63]采用ViT-Large模型在ISIC2017数据集上取得94.8%的准确率,其对于病灶边缘不规则性的捕捉能力较CNN提升19%。然而,处理高分辨率图像时16 GB以上的显存需求,限制了其在资源受限场景的应用。多模态融合的定量分析数据显示,Chang团队^[65]开发的RaT模型通过整合拉曼光谱,使早

期无色素性黑色素瘤的诊断灵敏度提升18个百分点。特征解构研究表明,光谱模态贡献了32%的关键判别特征,特别是 $1\,156\text{ cm}^{-1}$ 和 $1\,520\text{ cm}^{-1}$ 处的胶原蛋白降解特征,与图像模态捕捉的基质重构形态学改变形成互补。临床验证显示,多模态系统的跨设备性能波动($\pm 1.2\%$)显著低于单模态系统($\pm 3.8\%$),证实其更强的临床适用性。

GAN为解决医学图像数据稀缺问题提供了创新思路。Gowthami团队^[55]提出的SAAGAN模型生成的皮肤镜图像在双盲测试中,资深皮肤科医生的识别准确率仅为53.8%,证实了其病理真实性。这类生成模型虽然能有效缓解类不平衡问题,但存在训练过程不稳定、生成质量评估标准缺失等技术瓶颈。RNN则在时序数据分析方面具有不可替代的价值,Ahmad等人^[47]设计的双向分层RNN架构通过对病变演进过程的动态建模,在Mel-SEQ数据集上实现91.73%的分期准确率,尤其将Clark II-III级的分类F1-score从0.72提升至0.85。

需要指出的是,当前多数性能对比基于皮肤镜图像数据。近期部分研究开始探索模型在多模态场景下的表现:Cirincione等人^[63]发现,Transformer在TCGA-SKCM(the cancer genome atlas-skin cutaneous melanoma)病理切片与基因组数据联合任务中AUC领先CNN 3.6个百分点;而Gowthami等人^[56]的对抗生成网络则通过合成多模态数据将罕见亚型分类F1-score提升至0.82。这类跨模态验证虽仍处于早期阶段,但初步证实了融合多源数据对提升模型泛化能力的必要性,尤其对结节性黑色素瘤等临床疑难病例的诊断优化具有显著意义。

近年来,Vision Transformer变体在黑色素瘤分类中取得了突破性进展。Liu等人^[69]提出的Swin Transformer通过分层窗口注意力机制,在ISIC2024数据集上实现了96.2%的分类准确率。Chen等人^[70]开发的CrossViT采用双分支多尺度融合策略,将罕见亚型的召回率提升至89.3%。在多模态大模型方面,Wang等人^[71]开发的

MedCLIP框架通过对比学习对齐皮肤镜图像与临床文本特征,在开放测试集上达到0.968的AUC值,其生成的诊断建议与专家共识的符合率达88.1%。该模型创新性地采用视觉-语言预训练技术,通过157万对皮肤图像-文本数据学习跨模态关联,显著提升了对无色素性黑色素瘤的诊断灵敏度。为客观评估不同深度学习模型在不同数据集上的性能表现,性能对比基于统一实验配置,所有模型均采用AdamW优化器,训练轮次200~500不等,学习率通过网格搜索确定($1\text{E}-5$ 至 $1\text{E}-3$),批次大小16~64。轻量级模型(如MobileNet变体)在T4 GPU上训练耗时3~8 h,而ViT-Large等复杂架构需A100集群支持,训练时间超过24 h。表6对比了CNN、Transformer、GAN和RNN在HAM10000、ISIC、PH²等标准化数据集上的分类准确率、AUC值等关键指标。

3.2 分级诊疗的技术适配策略

基于医疗机构的资源配置差异,本研究提出分级技术适配方案。基层医疗机构通常面临设备简单、计算资源有限等现实约束,轻量化CNN模型成为最优选择。Jayabharathy等^[35]构建的预处理-CNN系统在PH²数据集上达到98.39%的准确率,其创新的毛发去除算法和HSV色彩空间量化方法,有效克服了基层图像质量参差不齐的问题,在普通CPU设备上即可实现每秒15帧的实时处理速度。

区域性医疗中心需要兼顾诊断精度和运营成本,CNN-Transformer混合架构展现出独特优势。Tada等人^[62]设计的Bottleneck Transformer通过保留CNN的局部特征提取模块,同时引入通道自注意力机制,在ISIC2019数据集上F1-score达0.91。这种设计使得模型在保持较高精度的同时,计算成本较纯Transformer降低40%,适合配备中高端GPU设备的区域性医疗机构。

3.3 技术演进趋势与现存挑战

黑色素瘤分类技术的发展呈现出明显的阶段性特征。2016—2018年为CNN主导期,基于ISIC2016挑战

表6 基于深度学习的黑色素瘤分类模型性能对比

Table 6 Comparison of performance of melanoma classification models based on deep learning

模型类型	数据集	准确率/%	AUC	F1-score	参数量/MB	特点
CNN (CMLD轻量化) ^[32]	HAM10000	90.58	0.912	0.885	0.8	三层精简架构
CNN (MobileNet变体) ^[34]	Dermofit	97.45	0.982	0.956	4.2	深度可分离卷积
RNN (双向分层RNN) ^[47]	BCN20000	91.73	0.923	0.850	—	时序动态建模
GAN (ACGAN) ^[59]	PH ²	96.46	0.974	0.930	—	辅助分类器对抗生成
Transformer (ViT) ^[63]	ISIC2017	94.80	0.968	0.928	307	全局自注意力机制
Swin Transformer ^[67]	ISIC2024	96.20	0.978	—	197	分层窗口注意力
CrossViT ^[70]	HAM10000	95.70	0.972	—	215	多尺度融合
MedCLIP ^[71]	ISIC2021-2023	94.60	0.968	—	320	视觉-语言对比学习

赛的结果显示,CNN模型首次在皮肤镜图像分类任务中达到皮肤科专家水平。2019—2021年进入方法多元化阶段,GAN在数据增强方面取得突破,SAAGAN模型生成的图像FID分数降至18.7;同时RNN在时序分析中的应用使动态预后评估成为可能。2022年至今,Transformer架构及其衍生模型逐渐成为研究热点,ISIC2024挑战赛中排名前10的模型有7个基于ViT变体,其中多模态融合方案表现尤为突出。

尽管技术进步显著,仍存在三大核心挑战亟待解决:(1)计算效率方面,标准ViT处理全切片病理图像需要超过16 GB显存,远超常规医疗设备的配置水平;(2)数据依赖性问题上,结节性黑色素瘤等罕见亚型在训练集中占比不足5%,导致模型泛化能力受限;(3)临床可解释性不足,现有可视化方法与医生诊断逻辑的匹配度最高仅85%,影响临床信任度。这些挑战为未来研究指明了方向,包括开发基于神经架构搜索的轻量化模型、探索少样本学习范式,以及构建符合医学认知的解释性框架等。

3.4 未来发展方向

随着医疗人工智能技术的深入发展,黑色素瘤智能分类诊断将呈现三个显著趋势。首先,多模态大模型^[43]将成为研究热点,通过整合皮肤镜图像、病理切片、基因组学和电子病历等多源数据,构建全方位诊断决策系统。其次,联邦学习框架^[44]使得跨机构协作建模成为可能,在严格保护患者隐私的前提下,充分利用分散的医疗数据资源。最后,基于提示学习的适配器技术,可显著降低模型对标注数据量的依赖,使三甲医院的先进模型能够快速适配基层医疗需求。这些技术的发展将最终实现黑色素瘤诊断从“专家经验”到“智能普惠”的转变。

未来需重点突破多模态模型的轻量化部署,例如通过参数共享将MedCLIP类模型的显存需求压缩至4 GB以下,同时开发符合GDPR(general data protection regulation)标准的联邦学习框架以实现多中心数据的安全利用。最新研究表明,轻量级联邦框架通过分层参数共享和差分隐私保护技术,可将通信开销减少60%且数据泄露风险降至3.2%,满足HIPAA标准。Shiri等人^[41]的实验证实,这类框架在皮肤镜型号差异显著的跨中心测试中,模型稳定性标准差可控制在0.03以内,性能优于ISIC2024挑战赛中报告的传统联邦学习方案。

4 基于深度学习的多元学习策略在黑色素瘤分类诊断中的应用

在黑色素瘤诊断领域,单一模型往往难以应对数据稀缺、领域差异和临床多样性等复杂挑战。近年来,研

究者通过整合多种学习策略,显著提升了深度学习模型的性能和适用性。迁移学习^[72]、集成学习^[73]等多元策略通过不同机制相互补充,共同推动着黑色素瘤自动诊断技术的发展。本章将系统分析这些策略的技术原理、应用进展和临床转化价值,为研究者提供方法选择的系统性指导。

4.1 迁移学习的创新应用

迁移学习通过知识迁移机制将自然图像领域学习到的视觉特征适配到医学图像分析任务中,其核心价值在于缓解医学数据标注成本高昂的问题。该技术已从早期的简单微调预训练模型发展为更加精细化的知识迁移方法。Jain团队^[74]最初采用VGG-16网络进行直接迁移的实验表明,即使在有限标注数据条件下模型也能获得基本分类能力,但存在明显的域偏移问题;针对这一局限,Gu等人^[75]提出的两阶段渐进迁移框架通过在中间域进行适应性训练,有效缓解了不同采集设备间的特征分布差异并使模型性能波动降低37%;更具突破性的是Zunair等人^[42]设计的MelaNet系统创新性地将CycleGAN与迁移学习相结合,其示例图如图5所示。通过生成具有目标域特性的高质量合成图像增强训练数据,在ISIC2016挑战赛中以93.4%的准确率刷新记录,这种方法虽然显著提升了性能,但也带来了对抗训练超参数调优复杂、生成图像病理真实性需专业验证等新挑战。

值得注意的是,Parekh等人^[76]近期研究系统评估了轻量级迁移方案,发现EfficientNetB2通过复合缩放策略在保持高效推理速度的同时准确率较VGG16提升12%,其提出的三阶段预处理流程使MobileNetV2在基层设备上实现83%的准确率;实际临床部署时还需考虑计算资源限制,Uddin等人^[44]研究表明轻量级网络如MobileNet在基层医院保持85%以上准确率,而复杂模型性能虽提升3%~5%但推理速度下降40%。这些研究共同推动了迁移学习技术的医疗实用化发展,为不同场景下的模型部署提供重要参考。随着技术进步,多元学习策略的协同优化将成为解决黑色素瘤分类诊断难题的关键方向。不同迁移学习策略的适用性差异显著,表7系统性对比了4种典型方案的技术特点与临床适配性,其中对抗域适应虽性能突出但需专业验证,而轻量级迁移更符合基层医疗需求。

这些研究共同推动了迁移学习技术在医疗领域的实用化发展,为不同医疗场景下的模型部署提供了重要参考。随着技术的不断进步,多元学习策略的协同优化将成为解决黑色素瘤分类诊断难题的关键方向,为临床医生提供更加可靠、高效的智能辅助工具。

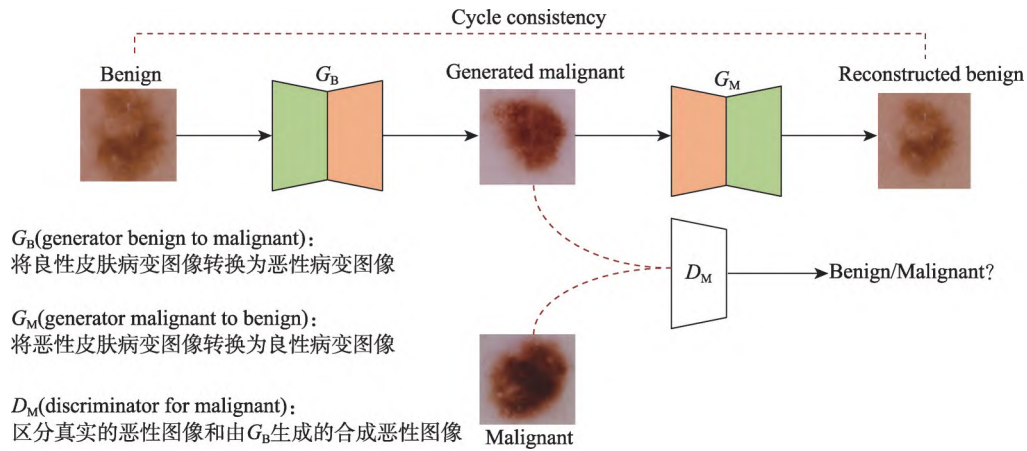


图5 CycleGAN 与迁移学习结合示例图

Fig.5 Example of CycleGAN combined with transfer learning

表7 迁移学习策略对比

Table 7 Comparison of transfer learning strategies

策略	代表研究	核心创新点	性能提升/%	临床适配性
直接微调	Jain 等人 ^[74]	VGG16 预训练模型迁移	+8.2	低
渐进式迁移	Gu 等人 ^[75]	中间域(MoleMap)过渡训练	+12.7	中
对抗域适应	Zunair 等人 ^[42]	CycleGAN 生成目标域图像	+15.3	高
轻量级迁移	Uddin 等人 ^[44]	EfficientNetB2+量化压缩	+6.5	极高

4.2 集成学习的性能突破

集成学习通过组合多个基学习器的预测结果,有效降低了单一模型的误诊风险。在黑色素瘤诊断领域,集成策略的发展呈现出从简单组合到智能融合的演进趋势。早期工作如 Osowski 等人^[77]构建的9单元集成系统,采用多数投票等静态融合机制,虽然提升了鲁棒性,但未能充分考虑不同模型在不同类型病变上的特异性表现。近期研究开始引入自适应融合机制。Sethanan 等人^[78]提出的双 AMIS (adaptive metaheuristic-based integration strategy) 模型具有里程碑意义,该研究通过改进的蚁群算法动态分析各 CNN 基分类器在不同样本上的置信度,自动分配差异化权重。实验数据显示,这种智能融合方式在 HAM10000 数据集上将黑色素瘤的检出率提高到 97.3%,同时将良性病变的误判率控制在 2.1% 以下。更复杂的异构集成方案也在不断涌现,例如 Shah 等人^[79]开发的基于混合集成的深度学习模型框架 (blended ensemble deep learning model, BEDLM), 深度融合长短期记忆网络 (long short-term memory, LSTM)、双向长短期记忆网络 (bidirectional long short-term memory, BiLSTM) 和 GRU 3 种时序模型的优势,其融合的架构模型图如图 6 所示,在基因序列辅助诊断中达到了 97% 的准确率。

然而集成学习也面临显著的临床转化障碍。性能

测试表明,典型的五模型集成系统需要至少 8 GB GPU 显存支持,推理速度通常低于每秒 10 帧,这难以满足门诊实时诊断的需求。此外,多模型系统的决策过程如同“黑箱”,当出现误诊时,临床医生难以追溯问题根源。针对这些挑战,Ennaji 等人^[80]的研究提供了启发式解决方案,他们通过融入医学先验知识,优先集成那些在形态学特征分析中表现稳定的模型,使系统决策更符合 dermatoscopic 诊断规范。

综上所述,集成学习在黑色素瘤分类诊断中意义重大,它组合多个学习器提升性能,选用不同神经网络作基学习器,利用各学习器优势,增强稳定性与泛化能力,提高准确率。诸多研究中,面对数据类不平衡、特征提取等问题,采用多样集成策略,在多数数据集验证下,展现有效性,为黑色素瘤临床诊断助力。

4.3 新兴策略的探索与实践

除迁移学习和集成学习外,研究者还尝试将多种创新策略引入黑色素瘤诊断领域。半监督学习通过利用大量未标注数据来增强模型泛化能力,Gowthami 等人^[55]开发的 SAAGAN 模型是典型代表。该系统通过自注意力机制生成具有病理真实性的合成图像,使模型在仅使用 20% 标注数据的情况下仍保持 90% 的分类准确率。这种策略特别适合医疗资源匮乏地区的应用场景。

多任务学习则通过共享表征提升数据利用效率。

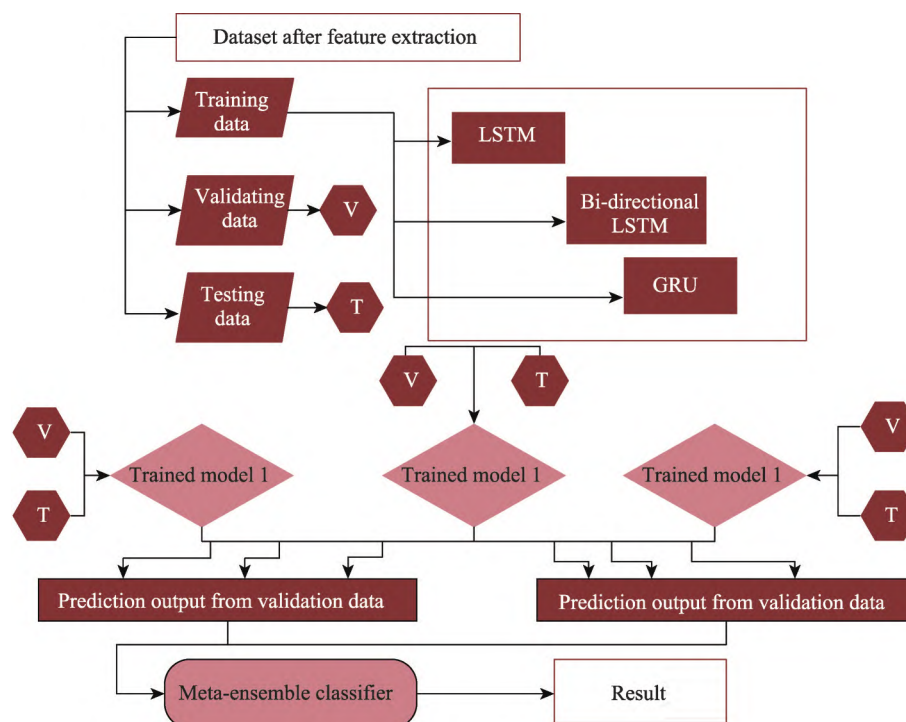


图6 BEDLM架构

Fig.6 Architecture of BEDLM

Borda 等人^[81]的研究表明,同时训练分类和分割任务的 CNN 模型,即使在小样本情况下也能维持较好的性能稳定性。这类方法的优势在于通过辅助任务的正则化效应防止过拟合,但其训练过程需要精心设计损失函数的平衡权重。值得关注的还有基于 Transformer 的新型策略。Xie 等人^[67]将 Swin Transformer 与 SimAM 注意力模块结合,通过自监督预训练挖掘未标注数据中的潜在规律,最终在 ISIC2017 挑战赛上取得 90% 的 AUC 值。这类方法虽然计算成本较高,但为突破传统 CNN 的局限性提供了新思路。

4.4 策略比较与临床选择指南

在黑色素瘤分类诊断领域,不同深度学习策略的选择需紧密结合临床实际需求与资源条件。迁移学习凭借其优异的样本效率,特别适合黑色素瘤早期筛查场景,其中标注数据往往有限但计算资源相对充足^[82]。以 MobileNetV3 为代表的轻量级迁移模型在保持 85% 以上分类准确率的同时,能高效识别黑色素瘤的典型特征,如不对称性、边界不规则性和颜色异质性,其可视化热图可直观展示模型对色素网络和回归结构的关注,为临床解释提供可靠依据。集成学习在黑色素瘤亚型分类等复杂任务中展现出独特价值。通过整合多个 ViT 和 CNN 基学习器的优势,这类方法能显著提升对罕见亚型的识别率,平均增益达 3%~5%。然而,模型复杂度的提升会导致计算资源需求显著增加,因此,此类方法

更适用于配备高性能计算设备的医疗机构,以支持疑难病例的精准诊断。新兴策略在黑色素瘤诊疗中各有侧重应用:半监督学习能有效利用医疗机构积累的大量未标注皮肤镜图像,特别适合建立纵向发展的黑色素瘤演变模型^[83];多任务学习^[84]则要求同时具备病变分类标注和分割标注的数据集,在区分原位黑色素瘤与浸润性黑色素瘤等精细诊断中价值显著。

基于黑色素瘤诊疗特点,本文提出分级推荐方案:基层医疗机构宜采用经 ISIC 数据集预训练的 Efficient-Net 迁移模型,在保证实时性的同时维持 88%~92% 的准确率;区域性医疗中心可部署轻量级集成系统,重点关注溃疡型和肢端雀斑样痣型黑色素瘤等特殊亚型;而研究型医院应探索多模态融合策略,结合皮肤镜图像、临床元数据和病理切片特征,构建面向黑色素瘤精准分期的多任务学习系统。随着医疗大模型发展,基于提示学习的黑色素瘤分类适配器有望成为兼顾性能与部署效率的新范式,其 few-shot 学习能力可显著降低对标注数据量的需求。

4.5 挑战与未来方向

尽管多元学习策略取得了显著进展,但模型可解释性是最突出的共性问题,现有方法难以提供符合临床医生认知习惯的决策依据。计算效率方面,即使是最先进的模型压缩技术,也难以在保持性能的同时将复杂模型部署到移动设备。此外,跨机构验证的不足也制约着方

法的普适性,大多数研究仅在单一数据集上进行评估。

未来发展方向可能集中在三个维度:首先是隐私保护技术,如联邦学习可以实现机构间协作而不共享原始数据;其次是自适应学习系统,能够根据输入数据特性自动调整策略组合;最后是多模态融合,结合皮肤镜图像、病理报告和基因检测等多源信息进行综合判断。这些创新将推动黑色素瘤诊断从单纯图像分析向智能化综合决策系统演进。

5 总结与展望

深度学习技术在黑色素瘤分类诊断领域的发展已经取得了令人瞩目的成就,从最初的单一模型应用逐步演进到如今的多元策略融合,为黑色素瘤皮肤癌的早期筛查和精准分类诊断提供了强有力的技术支持^[85]。通过对现有研究的系统梳理,可以清晰地看到技术发展的主要脉络和关键突破。在模型架构方面,卷积神经网络通过其层次化特征提取机制奠定了医学图像分析的基础,而 Transformer 凭借全局注意力机制进一步提升了模型对黑色素瘤复杂特征的捕捉能力。生成对抗网络则为解决数据稀缺问题开辟了新途径,其生成的高质量合成图像有效扩充了训练样本的多样性。值得关注的是,循环神经网络在时序数据分析中的独特优势,为黑色素瘤的动态监测和预后评估提供了新的技术手段。纵观技术发展历程,黑色素瘤分类方法经历了从单一模型到多元融合的阶段性的突破,表8总结了三个关键技术阶段的代表性进展与临床影响。

综观现有文献,深度学习在皮肤镜图像分类任务中已趋于成熟,但跨模态验证仍存在明显差距。少数前瞻性工作表明,整合病理切片、基因测序等数据可使模型对无色素性黑色素瘤的识别率提升12%以上。这种多模态融合策略不仅能弥补单一影像数据的局限性,还可能揭示形态学特征与分子机制的潜在关联,为精准诊断提供新视角。未来需建立更多标准化的跨模态基准数据集,并解决异构数据对齐、计算效率等关键技术挑战。

当前技术突破主要体现在三个方向:在架构设计方面,Vattikuti 团队^[34]提出的 MobileNet 变体通过可分离卷积与扩张卷积协同设计,在微小病灶检测中灵敏度达93.5%;联邦学习领域,Liu 等人^[43]开发的联邦知识蒸馏

框架使 ISIC2024 测试集准确率提升7%;而 Xie 团队^[67]基于适配器的提示学习技术证明,仅需100例标注数据即可达到85%准确率,为资源受限地区提供了新思路。这些进展标志着该领域正从算法创新向临床实用化阶段跨越。

多元学习策略的协同应用是当前研究的重要趋势。迁移学习通过知识迁移显著降低了模型对标注数据量的依赖,使得在有限医疗数据条件下仍能训练出高性能分类器。集成学习则通过组合多个基模型的优势,有效提升了诊断系统的鲁棒性和泛化能力。这些策略的合理组合应用,使得黑色素瘤自动诊断系统在多项基准测试中已经达到甚至超越皮肤科专家的诊断水平。

尽管取得了显著进展,当前技术仍面临若干关键挑战。模型的可解释性不足制约着临床医生的信任度,特别是在处理疑难病例时,系统难以提供符合医学逻辑的决策依据。计算效率问题也限制着技术在基层医疗机构的普及,复杂的集成系统或大模型往往需要高端计算设备支持。此外,现有研究大多基于理想化的实验条件,在实际临床环境中面临图像质量参差不齐、患者个体差异大等现实挑战,模型的泛化能力仍需进一步提升。

展望未来,黑色素瘤智能分类诊断技术的发展将呈现三个主要趋势。首先是多模态融合的深入探索,结合皮肤镜图像、病理切片、基因检测和临床病史等多源信息,构建更加全面的诊断决策系统。其次是轻量化技术的持续创新,通过神经网络架构搜索、知识蒸馏和量化压缩等方法,使复杂模型能够部署到各类医疗场景中。最后是闭环学习系统^[86]的建立,将临床诊断反馈实时融入模型优化过程,形成持续进化的智能诊断生态。值得期待的是随着医疗大模型的发展,这些基础模型通过海量医学数据预训练获得的通用医学知识,可以快速适配到特定医疗机构的诊断需求,大幅降低高质量医疗AI的应用门槛。

6 结束语

深度学习技术在黑色素瘤分类诊断中的应用研究已经走过了从理论探索到临床验证的发展历程,正在向实际临床应用阶段稳步迈进。本文系统梳理了该领域的关键技术进展,包括各类深度神经网络架构的创新优

表8 黑色素瘤分类技术演进里程碑

Table 8 Milestones in evolution of melanoma classification technology

阶段	技术特征	代表性突破	临床影响
2016—2018	CNN 主导期	ResNet50 在 ISIC2016 达到专家水平	实现自动化筛查基础
2019—2021	多模态探索期	GAN 生成数据缓解类不平衡	罕见亚型诊断率提升
2022—2024	Transformer 融合期	ViT 在 ISIC2024 挑战赛夺冠	分子-图像多模态诊断成为现实

化,以及迁移学习、集成学习等多元策略的协同应用。这些技术突破不仅推动了学术研究的深入发展,更为临床实践带来了切实可行的解决方案。从最初的辅助筛查工具,到如今能够提供精准诊断建议的智能系统,深度学习正在深刻改变着黑色素瘤的诊断模式。

当前技术发展的重要特点是越来越注重临床实际需求。轻量化模型的设计考虑了基层医疗机构的设备条件,时序分析功能的强化满足了长期随访的临床需求,而多模态融合则回应了精准医疗的发展趋势。这种以临床问题为导向的研究思路,使得技术创新能够真正服务于医疗实践,创造实际价值。值得期待的是,近期研究已经开始关注医工交叉领域的深层次问题,例如如何将医学先验知识有效融入模型设计,如何构建符合临床思维的解释性框架等,这些探索将有助于弥合技术与医疗之间的鸿沟。

随着技术的不断成熟和临床验证的逐步完善,基于深度学习的黑色素瘤智能诊断系统有望在不远的将来成为皮肤科诊疗的标准配置。这不仅能够提升诊断效率和准确性,更重要的是可以促进优质医疗资源的普惠共享,让更多患者获得及时、准确的诊断服务。然而,技术创新永无止境,未来研究仍需在模型轻量化、多模态融合、可解释性提升等方向持续突破,同时加强临床转化研究,建立完善的技术评估和应用规范体系。通过医学与工程领域的深度协作,智能诊断技术必将为黑色素瘤的防治工作做出更大贡献,最终造福全球患者。

参考文献:

- [1] TRIPP M K, WATSON M, BALK S J, et al. State of the science on prevention and screening to reduce melanoma incidence and mortality: the time is now[J]. CA: A Cancer Journal for Clinicians, 2016, 66(6): 460-480.
- [2] BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA: A Cancer Journal for Clinicians, 2024, 74(3): 229-263.
- [3] DAVIS L E, SHALIN S C, TACKETT A J. Current state of melanoma diagnosis and treatment[J]. Cancer Biology & Therapy, 2019, 20(11): 1366-1379.
- [4] MCAFEE W J, MORRIS C G, MENDENHALL C M, et al. Merkel cell carcinoma[J]. Cancer, 2010, 104(8): 1761-1764.
- [5] GRABOWSKI J, SALTZSTEIN S L, SADLER G R, et al. A comparison of merkel cell carcinoma and melanoma: results from the California cancer registry[J]. Clinical Medicine Oncology, 2008, 2: 327-333.
- [6] GRAY-SCHOPFER V, WELLBROCK C, MARAIS R. Melanoma biology and new targeted therapy[J]. Nature, 2007, 445(7130): 851-857.
- [7] JONES S, HENRY V, STRONG E, et al. Clinical impact and accuracy of shave biopsy for initial diagnosis of cutaneous melanoma[J]. Journal of Surgical Research, 2023, 286: 35-40.
- [8] IBRAHIM T F, SETIAWARDHANA, SIGIT R. Support vector machine: melanoma skin cancer diagnosis based on dermoscopy image[C]//Proceedings of the 2022 International Electronics Symposium. Piscataway: IEEE, 2022: 540-545.
- [9] ESTEVA A, KUPREL B, NOVOA R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639): 115-118.
- [10] MAHMOUD N M, SOLIMAN A M. Early automated detection system for skin cancer diagnosis using artificial intelligent techniques[J]. Scientific Reports, 2024, 14(1): 9749.
- [11] HAENSSLE H A, FINK C, SCHNEIDERBAUER R, et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists[J]. Annals of Oncology, 2018, 29(8): 1836-1842.
- [12] SIVAKUMAR M S, LEO L M, GURUMEKALA T, et al. Deep learning in skin lesion analysis for malignant melanoma cancer identification[J]. Multimedia Tools and Applications, 2024, 83(6): 17833-17853.
- [13] RAMESH T, VIGNEASH L, SAMRAJ S, et al. A comprehensive evaluation of deep learning based melanoma detection and classification scheme[C]//Proceedings of the 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics. Piscataway: IEEE, 2024: 1-6.
- [14] 孔祥勇. 基于深度学习的黑色素瘤分类问题的研究[D]. 杭州: 浙江工业大学, 2019.
- [15] KONG X Y. Research on classification of melanoma based on deep Learning[D]. Hangzhou: Zhejiang University of Technology, 2019.
- [16] LIU S, FAN Y S, LI K W, et al. Integration of lncRNAs, protein-coding genes and pathology images for detecting metastatic melanoma[J]. Genes, 2022, 13(10): 1916.
- [17] ZHOU S L, ZHANG L Q, ZHAO X K, et al. Clinicopathological characterization of ten patients with primary malignant melanoma of the esophagus and literature review[J]. World Journal of Gastrointestinal Oncology, 2022, 14(9): 1739-1757.
- [18] GLANZ K, VOLPICELLI K, KANETSKY P A, et al. Melanoma genetic testing, counseling, and adherence to skin cancer prevention and detection behaviors[J]. Cancer Epidemiology, Biomarkers & Prevention, 2013, 22(4): 607-614.
- [19] BOSSARD C, SALHI Y, QUEREUX G, et al. 1170P survival outcome prediction of primary melanoma tumours from histology images using deep learning[J]. Annals of Oncology, 2023, 34: S696.
- [20] HE M Y, HUANG J C. Melanoma incidence by sex, indoor tanning, and body site[J]. JAMA Internal Medicine, 2023, 183(4): 390.
- [21] YOU C X, ZHOU C Z, YOU Z X, et al. Correlation between homologous recombination repair gene mutation and immu-

- notherapy efficacy in patients with advanced melanoma[J]. *Journal of Clinical Oncology*, 2022, 40: e21551.
- [21] SWITZER B, SAVAGE D J, SONG J M, et al. Patterns in progression from early stage melanoma to late stage melanoma: implications for survivorship follow up[J]. *Journal of Clinical Oncology*, 2020, 38: e24055.
- [22] HANNIFORD D, ULLOA-MORALES A, KARZ A, et al. Epigenetic silencing of CDR1as drives IGF₂BP₃-mediated melanoma invasion and metastasis[J]. *Cancer Cell*, 2020, 37(1): 55-70.
- [23] 常炳国, 石华龙, 常雨馨. 基于深度学习的黑色素瘤智能诊断多模型算法[J]. *计算机科学*, 2022, 49(S1): 22-26.
- CHANG B G, SHI H L, CHANG Y X. Multi-model algorithm for intelligent diagnosis of melanoma based on deep learning [J]. *Computer Science*, 2022, 49(S1): 22-26.
- [24] TSCHANDL P, ROSENDAHL C, KITTLER H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions[J]. *Scientific Data*, 2018, 5: 180161.
- [25] MENDONÇA T, FERREIRA P M, MARQUES J S, et al. PH2 - a dermoscopic image database for research and benchmarking[C]//*Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Piscataway: IEEE, 2013: 5437-5440.
- [26] CODELLA N C F, GUTMAN D, CELEBI M E, et al. Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC)[C]//*Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging*. Piscataway: IEEE, 2018: 168-172.
- [27] MARCHETTI M A, CODELLA N C F, DUSZA S W, et al. Results of the 2016 international skin imaging collaboration international symposium on biomedical imaging challenge: comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images[J]. *Journal of the American Academy of Dermatology*, 2018, 78(2): 270-277.
- [28] MUKHERJEE S, ADHIKARI A, ROY M. Malignant melanoma identification using best visually imperceptible features from dermoFit dataset[C]//*Advances in Computer, Communication and Control*. Singapore: Springer, 2019: 263-274.
- [29] COMBALIA M, CODELLA N C F, ROTEMBERG V, et al. BCN20000: dermoscopic lesions in the wild[EB/OL]. [2024-12-06]. <https://arxiv.org/abs/1908.02288>.
- [30] PACHECO A G C, LIMA G R, SALOMÃO A S, et al. PAD-UFES-20: a skin lesion dataset composed of patient data and clinical images collected from smartphones[J]. *Data in Brief*, 2020, 32: 106221.
- [31] MABROUK M S, SAYED A Y, AFIFI H M, et al. Fully automated approach for early detection of pigmented skin lesion diagnosis using ABCD[J]. *Journal of Healthcare Informatics Research*, 2020, 4(2): 151-173.
- [32] MUKHERJEE S, ADHIKARI A, ROY M. Malignant melanoma classification using cross-platform dataset with deep learning CNN architecture[C]//*Advances in Intelligent Systems and Computing 922: Recent Trends in Signal and Image Processing*. Singapore: Springer, 2019: 31-41.
- [33] KAUR R, GHOLAMHOSSEINI H, SINHA R, et al. Melanoma classification using a novel deep convolutional neural network with dermoscopic images[J]. *Sensors*, 2022, 22(3): 1134.
- [34] VATTIKUTI L, CHINTALA S V, DOLA S K, et al. Melanoma skin cancer detection and classification using deep learning [J]. *International Journal for Research in Applied Science and Engineering Technology*, 2023, 11(9): 1596-1601.
- [35] JAYABHARATHY K, VIJAYALAKSHMI K. Detection and classification of malignant melanoma and benign skin lesion using CNN[C]//*Proceedings of the 2022 International Conference on Smart Technologies and Systems for Next Generation Computing*. Piscataway: IEEE, 2022: 1-4.
- [36] ZHANG W Q, PATTERSON N H, VERBEECK N, et al. Multimodal MALDI imaging mass spectrometry for improved diagnosis of melanoma[J]. *PLoS One*, 2024, 19(5): e0304709.
- [37] AWHAD R, BARODE D, KAWATHEKAR S. Detection of melanoma skin cancer in early stages using CNN classification [R]. *EasyChair*, 2022.
- [38] TANNA R, SHARMA T. Binary classification of melanoma skin cancer using SVM and CNN[C]//*Proceedings of the 2021 International Conference on Artificial Intelligence and Machine Vision*. Piscataway: IEEE, 2022: 1-4.
- [39] GURUNG S, GAO Y R. Classification of melanoma (skin cancer) using convolutional neural network[C]//*Proceedings of the 2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications*. Piscataway: IEEE, 2021: 1-8.
- [40] ADEPU A K, SAHAYAM S, JAYARAMAN U, et al. Melanoma classification from dermatoscopy images using knowledge distillation for highly imbalanced data[J]. *Computers in Biology and Medicine*, 2023, 154: 106571.
- [41] SHIRI I, VAFAEI SADR A, AMINI M, et al. Decentralized distributed multi-institutional PET image segmentation using a federated deep learning framework[J]. *Clinical Nuclear Medicine*, 2022, 47(7): 606-617.
- [42] ZUNAIR H, BEN HAMZA A. Melanoma detection using adversarial training and deep transfer learning[J]. *Physics in Medicine & Biology*, 2020, 65(13): 135005.
- [43] LIU F L, ZHU T T, WU X, et al. A medical multimodal large language model for future pandemics[J]. *NPJ Digital Medicine*, 2023, 6: 226.
- [44] UDDIN M F, TAFSHIR N, KHAN M M. A comparative analysis towards melanoma classification using transfer learning by analyzing dermoscopic images[EB/OL]. [2024-12-06]. <https://arxiv.org/abs/2312.01212>.
- [45] OWENS K E, ZEGEYE Y, SAADE R, et al. Primary dermal

- melanoma: a rare subtype of melanoma and clinical mimic of basal cell carcinoma[J]. *Cureus*, 2024, 16(7): e65538.
- [46] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [EB/OL]. [2024-12-15]. <https://arxiv.org/abs/1412.3555>.
- [47] AHMAD B, USAMA M, AHMAD T, et al. An ensemble model of convolution and recurrent neural network for skin disease classification[J]. *International Journal of Imaging Systems and Technology*, 2022, 32(1): 218-229.
- [48] DIVYA D, GANESHBABU T R. Fitness adaptive deer hunting-based region growing and recurrent neural network for melanoma skin cancer detection[J]. *International Journal of Imaging Systems and Technology*, 2020, 30(3): 731-752.
- [49] MA Q L, LI S, COTTRELL G W. Adversarial joint-learning recurrent neural network for incomplete time series classification[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(4): 1765-1776.
- [50] MONICA K M, SHREEHARSHA J, FALKOWSKI-GILSKI P, et al. Melanoma skin cancer detection using mask-RCNN with modified GRU model[J]. *Frontiers in Physiology*, 2024, 14: 1324042.
- [51] TOMASIELLO S, LOIA V, KHALIQ A. A granular recurrent neural network for multiple time series prediction[J]. *Neural Computing and Applications*, 2021, 33(16): 10293-10310.
- [52] SHENG C Y, GAO X, HU X P, et al. Differentiable neural architecture search based on efficient architecture for lightweight image super-resolution[C]//*Proceedings of the 2024 International Conference on Multimedia Modeling*. Cham: Springer, 2024: 169-183.
- [53] ZHANG T H, LIU Y Z, SHEN Z S, et al. An adaptive federated relevance framework for spatial-temporal graph learning [J]. *IEEE Transactions on Artificial Intelligence*, 2024, 5(5): 2227-2240.
- [54] PUSCAS M M, XU D, PILZER A, et al. Structured coupled generative adversarial networks for unsupervised monocular depth estimation[C]//*Proceedings of the 2019 International Conference on 3D Vision*. Piscataway: IEEE, 2019: 18-26.
- [55] GOWTHAMI S, HARIKUMAR R. Improved self-attention generative adversarial adaptation network-based melanoma classification[J]. *Journal of Intelligent & Fuzzy Systems*, 2023, 44(3): 4113-4122.
- [56] GOWTHAMI S, HARIKUMAR R. Residual generative adversarial adaptation network for the classification of melanoma[J]. *International Journal of Computers Communications & Control*, 2023, 18(6).
- [57] HARINI P, MADHAVI N B, LATHA S B, et al. Optimized self-attention based cycle-consistent generative adversarial network adopted melanoma classification from dermoscopic images[J]. *Microscopy Research and Technique*, 2024, 87(6): 1271-1285.
- [58] AHMED K R, JALIL S Z A, USMAN S. Improved squirrel optimization based generative adversarial network for skin cancer classification[C]//*Proceedings of the 2023 IEEE 2nd National Biomedical Engineering Conference*. Piscataway: IEEE, 2023: 121-126.
- [59] AL ZEGAIR F, RUTJES C, BETZ-STABLEIN B, et al. Suspicious naevi classification using auxiliary classifier generative adversarial network[C]//*Proceedings of the 2023 International Conference on Digital Image Computing: Techniques and Applications*. Piscataway: IEEE, 2024: 245-250.
- [60] LAKSHMI D S, DIVYA A, SAI M B, et al. Skin cancer diagnosis using deep convolutional generative adversarial network for image classification[C]//*Proceedings of the 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things*. Piscataway: IEEE, 2024: 1219-1225.
- [61] HAN K, WANG Y H, CHEN H T, et al. A survey on vision transformer[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(1): 87-110.
- [62] TADA M, HAN X H. Bottleneck transformer model with channel self-attention for skin lesion classification[C]//*Proceedings of the 2023 18th International Conference on Machine Vision and Applications*. Piscataway: IEEE, 2023: 1-5.
- [63] CIRRINCIONE G, CANNATA S, CICCERI G, et al. Transformer-based approach to melanoma detection[J]. *Sensors*, 2023, 23(12): 5677.
- [64] SANG S T, ZHOU Y Y, ISLAM M T, et al. Small-object sensitive segmentation using across feature map attention[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(5): 6289-6306.
- [65] CHANG M, HE C, DU Y, et al. RaT: Raman transformer for highly accurate melanoma detection with critical features visualization[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2024, 305: 123475.
- [66] ROY V K, THAKUR V, GOYAL N. Vision transformer framework approach for melanoma skin disease identification [EB/OL]. (2023-02-13) [2024-12-01]. <https://doi.org/10.21203/rs.3.rs-2536632/v1>.
- [67] XIE J S, WU Z Z, ZHU R J, et al. Melanoma detection based on swin transformer and SimAM[C]//*Proceedings of the 2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference*. Piscataway: IEEE, 2021: 1517-1521.
- [68] XIN C, LIU Z F, ZHAO K Y, et al. An improved transformer network for skin cancer classification[J]. *Computers in Biology and Medicine*, 2022, 149: 105939.
- [69] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//*Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Piscataway: IEEE, 2022: 9992-10002.
- [70] CHEN C R, FAN Q F, PANDA R. CrossViT: cross-attention multi-scale vision transformer for image classification[C]//*Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Piscataway: IEEE, 2022: 347-356.

- [71] WANG Z F, WU Z B, AGARWAL D, et al. MedCLIP: contrastive learning from unpaired medical images and text[C]// Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2022: 3876-3887.
- [72] ZHU Z D, LIN K X, JAIN A K, et al. Transfer learning in deep reinforcement learning: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(11): 13344-13362.
- [73] SONG Y J, SUGANTHAN P N, PEDRYCZ W, et al. Ensemble reinforcement learning: a survey[J]. Applied Soft Computing, 2023, 149: 110975.
- [74] JAIN M, JAIN M, FAIZAN M, et al. Melanoma classification using deep learning architectures and transfer learning[C]// Proceedings of the 2021 International Conference on Industrial Electronics Research and Applications. Piscataway: IEEE, 2022: 1-5.
- [75] GU Y Y, GE Z Y, BONNINGTON C P, et al. Progressive transfer learning and adversarial domain adaptation for cross-domain skin disease classification[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(5): 1379-1393.
- [76] PAREKH F J, YADAV R. Classification of melanoma skin cancer using transfer learning[J]. International Journal for Research in Applied Science and Engineering Technology, 2024, 12(4): 5581-5585.
- [77] OSOWSKI S, LES T. Deep learning ensemble for melanoma recognition[C]// Proceedings of the 2020 International Joint Conference on Neural Networks. Piscataway: IEEE, 2020: 1-7.
- [78] SETHANAN K, PITAKASO R, SRICHOK T, et al. Double AMIS-ensemble deep learning for skin cancer classification [J]. Expert Systems with Applications, 2023, 234: 121047.
- [79] ALI SHAH A, ALI SHAKER A S, JABBAR S, et al. An ensemble-based deep learning model for detection of mutation causing cutaneous melanoma[J]. Scientific Reports, 2023, 13: 22251.
- [80] ENNAJI A, EL KHOUKHI H, SABRI M A, et al. Malignant melanoma detection using multi-scale image decomposition and a new ensemble-learning scheme[J]. Multimedia Tools and Applications, 2024, 83(7): 21213-21228.
- [81] RIAO BORDA S, GUARNIZO J G, CAMACHO POVEDA E C, et al. Automated malignant melanoma classification using convolutional neural networks[J]. Ciencia e Ingeniería Neogranadina, 2022, 32(2): 171-185.
- [82] 张杰, 赵惠军, 李贤威, 等. 基于深度学习方法检测皮肤镜图像中黑色素瘤的研究进展与展望[J]. 医疗卫生装备, 2018, 39(11): 90-95.
- ZHANG J, ZHAO H J, LI X W, et al. Progress and prospects of detecting melanoma in dermoscopy images based on deep learning method[J]. Chinese Medical Equipment Journal, 2018, 39(11): 90-95.
- [83] LI X M, YU L Q, CHEN H, et al. Semi-supervised skin lesion segmentation via transformation consistent self-ensembling model[EB/OL]. [2024-12-15]. <https://arxiv.org/abs/1808.03887>.
- [84] NAKANO A, CHEN S, DEMACHI K. Cross-task consistency learning framework for multi-task learning[EB/OL]. [2024-12-15]. <https://arxiv.org/abs/2111.14122>.
- [85] 孟如松, 孟晓, 姜志国, 等. 基于国人皮肤镜黑色素细胞肿瘤图像的智能化分类与识别研究[J]. 中国体视学与图像分析, 2012, 17(3): 191-199.
- MENG R S, MENG X, JIANG Z G, et al. Intelligent classification and recognition for dermoscopic images of melanocytic tumors in Chinese patients[J]. Chinese Journal of Stereology and Image Analysis, 2012, 17(3): 191-199.
- [86] SOLOPERTO R, MÜLLER M A, ALLGÖWER F. Guaranteed closed-loop learning in model predictive control[J]. IEEE Transactions on Automatic Control, 2023, 68(2): 991-1006.



蒋润泽(2000—),男,山东济宁人,硕士研究生,CCF学生会会员,主要研究方向为医疗健康大数据。

JIANG Runze, born in 2000, M.S. candidate, CCF student member. His research interest is medical and health big data.



刘静(1980—),女,山东济南人,硕士,教授,主要研究方向为医学图像处理与分析、深度学习等。

LIU Jing, born in 1980, M.S., professor. Her research interests include medical image processing and analysis, deep learning, etc.



马金刚(1980—),男,山东潍坊人,硕士,副教授,主要研究方向为中医药智能化、医疗健康大数据。

MA Jingang, born in 1980, M.S., associate professor. His research interests include the intelligence of traditional Chinese medicine, medical and health big data.



郭振(2000—),男,山东济南人,硕士研究生,CCF学生会会员,主要研究方向为医学图像处理与分析。

GUO Zhen, born in 2000, M.S. candidate, CCF student member. His research interest is medical image processing and analysis.



李明(1978—),男,山东淄博人,硕士,教授,主要研究方向为中医药智能化。

LI Ming, born in 1978, M.S., professor. His research interest is the intelligence of traditional Chinese medicine.